

Exercise 35.12. GAUGE TRANSFORMATIONS IN A CURVED BACKGROUND**EXERCISES**

(a) Show that the infinitesimal coordinate transformation (35.65a) induces the change (35.65b) in the functional form of the metric perturbation.

(b) Discuss the relationship between this gauge transformation and the concept of a Killing vector (§25.2).

Exercise 35.13. TRANSVERSE-TRACELESS GAUGE FOR GRAVITATIONAL WAVES PROPAGATING IN A CURVED BACKGROUND

(a) Show that, in vacuum in a curved background spacetime, the gauge condition $\bar{h}_{\mu}^{\alpha|_{\alpha}} = 0$ is preserved by transformations whose generator satisfies the wave equation $\xi_{\mu}^{|\alpha}_{\alpha} = 0$.

(b) Locally (over distances much smaller than \mathcal{R}) linearized theory is applicable; so there exists such a transformation which makes [see equations (35.7b) and (35.8a)]

$$\bar{h} = 0 + \text{error}, \quad \bar{h}_{\mu\alpha} u^{\alpha} = 0 + \text{error}. \quad (35.69)$$

Here u^{α} is a vector field that is as nearly covariantly constant as possible ($u^{\alpha|_{\beta}} = 0$); i.e., it is a constant vector in the inertial coordinates of linearized theory; and the errors are small over distances much less than \mathcal{R} . Show that $\bar{h} = 0$ can be imposed globally along with $\bar{h}_{\mu\alpha}^{|\alpha} = 0$; i.e., show that, if it is imposed on an initial hypersurface, the propagation equation (35.68) preserves it.

(c) Show that in general, the background curvature prevents any vector field from being covariantly constant ($u^{\hat{\alpha}}|_{\beta} \sim u^{\hat{\alpha}}/\mathcal{R}$ at best); and from this show that $\bar{h}_{\mu\alpha} u^{\alpha} = 0$ cannot be imposed globally along with $\bar{h}_{\mu}^{\alpha|_{\alpha}} = 0$.

§35.15. STRESS-ENERGY TENSOR FOR GRAVITATIONAL WAVES

Turn now to an evaluation of the effective stress-energy tensor $T_{\mu\nu}^{(\text{GW})}$ of equation (35.61). The evaluation requires averaging various quantities over several wavelengths. Useful rules for manipulating quantities inside the averaging brackets $\langle \rangle$ are the following (see exercise 35.14 for justification).

The averaging process involved in "coarse-grain" viewpoint

(1) Covariant derivatives commute; e.g., $\langle h h_{\mu\nu|_{\alpha\beta}} \rangle = \langle h h_{\mu\nu|_{\beta\alpha}} \rangle$. The fractional errors made by freely commuting are $\sim (\lambda/\mathcal{R})^2$, well below the inaccuracy of the computation.

(2) Gradients average out to zero; e.g., $\langle (h_{|\alpha} h_{\mu\nu})_{|\beta} \rangle = 0$. Fractional errors made here are $\lesssim \lambda/\mathcal{R}$.

(3) As a corollary, one can freely integrate by parts, flipping derivatives from one h to the other; e.g., $\langle h h_{\mu\nu|_{\alpha\beta}} \rangle = \langle -h_{|\beta} h_{\mu\nu|_{\alpha}} \rangle$.

A straightforward but long calculation using these rules, using equation (35.58b) for $R_{\mu\nu}^{(2)}(h)$, using definition (35.63) of $\bar{h}_{\mu\nu}$, using the propagation equation (35.64), and using the definition (35.61) of $T_{\mu\nu}^{(\text{GW})}$, yields $\langle R^{(2)}(h) \rangle = 0$, and

$$T_{\mu\nu}^{(\text{GW})} = \frac{1}{32\pi} \langle \bar{h}_{\alpha\beta|_{\mu}} \bar{h}^{\alpha\beta|_{\nu}} - \frac{1}{2} \bar{h}_{|\mu} \bar{h}_{|\nu} - 2\bar{h}^{\alpha\beta|_{\beta}} \bar{h}_{\alpha|\mu|_{\nu}} \rangle. \quad (35.70)$$

Evaluation of effective stress-energy tensor for gravitational waves, $T_{\mu\nu}^{(\text{GW})}$

This is the result quoted in equation (35.23'), except that there one used an inertial

coordinate system of linearized theory, where covariant derivatives and ordinary derivatives are the same. In a gauge where $\bar{h}_{\mu}^{\alpha}{}_{|\alpha} = 0$, the last term vanishes. When, in addition, $\bar{h}_{\mu\nu}$ is traceless (see exercise 35.13), the second term vanishes; and there remains only

$$T_{\mu\nu}^{(\text{GW})} = \frac{1}{32\pi} \langle \bar{h}_{\alpha\beta}{}_{|\mu} \bar{h}^{\alpha\beta}{}_{|\nu} \rangle \quad \text{if } \bar{h}_{\mu}^{\alpha}{}_{|\alpha} = \bar{h} = 0. \quad (35.70')$$

Accuracy of expression
for $T_{\mu\nu}^{(\text{GW})}$

Properties of $T_{\mu\nu}^{(\text{GW})}$

These expressions for the effective stress-energy of a gravitational wave have fractional errors of order \mathcal{A} , due to the neglect of second-order corrections to $h_{\mu\nu}$; they also have fractional errors of order λ/\mathcal{R} , due to the averaging process, which makes no sense when λ approaches \mathcal{R} in magnitude. Since $\mathcal{A} \lesssim \lambda/\mathcal{R}$ (35.28), the dominant errors in $T_{\mu\nu}^{(\text{GW})}$ are $\sim \lambda/\mathcal{R}$.

To this accuracy, the stress-energy tensor for gravitational waves is on an equal footing with any other stress-energy tensor. It plays the same role in producing background curvature; and it enters into conservation laws in the same way. For example, one can show, either by direct calculation or from the identity $G^{(\text{B})\mu\nu}{}_{|\nu} = 0$, that

$$T^{(\text{GW})\mu\nu}{}_{|\nu} = 0 + \text{error}, \quad (35.71)$$

where the error $\sim (\lambda/\mathcal{R})(T^{(\text{GW})\mu\nu}/\mathcal{R})$ is negligible in the shortwave approximation.

Some of the properties of $T_{\mu\nu}^{(\text{GW})}$ have already been explored in §35.7. Further properties are explored in exercises 35.18 and 35.19.

EXERCISES

Exercise 35.14. BRILL-HARTLE AVERAGE

Isaacson (1968b) introduces the following averaging scheme, which he names “Brill-Hartle averaging.”

(a) In the small region, of size several times λ , where the averaging occurs, there will be a unique geodesic of $g_{\mu\nu}^{(\text{B})}$ connecting any two points \mathcal{P}' and \mathcal{P} ; so given a tensor $\mathbf{E}(\mathcal{P}')$ at \mathcal{P}' , one can parallel transport it along this geodesic to \mathcal{P} , getting there a tensor $\mathbf{E}(\mathcal{P}')_{-\mathcal{P}}$.

(b) Let $f(\mathcal{P}', \mathcal{P})$ be a weighting function that falls smoothly to zero when \mathcal{P}' and \mathcal{P} are separated by many wavelengths, and such that

$$\int f(\mathcal{P}', \mathcal{P}) \sqrt{-g^{(\text{B})}(\mathcal{P}')} d^4x' = 1. \quad (35.72)$$

(c) Then the average of the tensor field $\mathbf{E}(\mathcal{P}')$ over several wavelengths about the point \mathcal{P} is

$$\langle \mathbf{E} \rangle_{\mathcal{P}} \equiv \int \mathbf{E}(\mathcal{P}')_{-\mathcal{P}} f(\mathcal{P}', \mathcal{P}) \sqrt{-g^{(\text{B})}(\mathcal{P}')} d^4x'. \quad (35.73)$$

(i) Show that there exists an entity $g_{\mu}^{(\text{B})\alpha'}(\mathcal{P}, \mathcal{P}')$, whose primed index transforms as a tensor at \mathcal{P}' and whose unprimed index transforms as a tensor at \mathcal{P} , such that (for \mathbf{E} second rank)

$$E_{\alpha\beta}(\mathcal{P}')_{-\mathcal{P}} = g_{\alpha}^{(\text{B})\mu'} g_{\beta}^{(\text{B})\nu'} E_{\mu'\nu'}(\mathcal{P}'). \quad (35.74)$$

This entity is called the “bivector of geodesic parallel displacement”; see DeWitt and Brehme (1960) or Synge (1960a).

(ii) Rewriting expression (35.73) in coordinate language as

$$\langle E_{\alpha\beta}(x) \rangle = \int g_{\alpha}^{(B)\mu'}(x, x') g_{\beta}^{(B)\nu'}(x, x') E_{\mu'\nu'}(x') f(x, x') \sqrt{-g^{(B)}(x')} d^4x', \quad (35.73')$$

derive the three averaging rules cited at the beginning of §35.15. [For solution, see Appendix of Isaacson (1968b).]

Exercise 35.15. GEOMETRIC OPTICS

Develop geometric optics for gravitational waves of small amplitude propagating in a curved background. Pattern the analysis after geometric optics for electromagnetic waves (§22.5). In particular, let $\bar{h}_{\mu\nu}$ have an amplitude that varies slowly (on a scale $\ell \lesssim \mathcal{R}$) and a phase θ that varies rapidly ($\theta_{,\hat{\alpha}} \sim 1/\lambda$). Expand the amplitude in powers of λ/ℓ , so that

$$\bar{h}_{\mu\nu} = \Re \{ A_{\mu\nu} + \epsilon B_{\mu\nu} + \epsilon^2 C_{\mu\nu} + \dots \} e^{i\theta/\epsilon}. \quad (35.75)$$

Here ϵ is a formal expansion parameter, actually equal to unity, which reminds one that the terms attached to ϵ^n are proportional to $(\lambda/\mathcal{R})^n$. Define the following quantities (with $A_{\mu\nu}^*$ denoting the complex conjugate of $A_{\mu\nu}$):

$$\text{"wave vector": } k_{\alpha} \equiv \theta_{,\alpha} \quad (35.76a)$$

$$\text{"scalar amplitude": } \mathcal{A} \equiv \left(\frac{1}{2} A_{\mu\nu}^* A^{\mu\nu} \right)^{1/2} \quad (35.76b)$$

$$\text{"polarization": } e_{\mu\nu} \equiv A_{\mu\nu}/\mathcal{A}. \quad (35.76c)$$

By inserting expression (35.75) into the gauge condition (35.66) and the propagation equation (35.68), derive the fundamental equations of geometrical optics as follows.

(a) The rays (curves perpendicular to surfaces of constant phase) are null geodesics; i.e.

$$k_{\alpha} k^{\alpha} = 0, \quad (35.77a)$$

$$k_{\alpha|\beta} k^{\beta} = 0. \quad (35.77b)$$

(b) The polarization is orthogonal to the rays and is parallel transported along them;

$$e_{\mu\alpha} k^{\alpha} = 0, \quad (35.77c)$$

$$e_{\mu\nu|\alpha} k^{\alpha} = 0. \quad (35.77d)$$

(c) The scalar amplitude decreases as the rays diverge away from each other in accordance with

$$\mathcal{A}_{,\alpha} k^{\alpha} = -\frac{1}{2} k^{\alpha}_{|\alpha} \mathcal{A}. \quad (35.77e)$$

i.e.,

$$(\mathcal{A}^2 k^{\alpha})_{|\alpha} = 0 \text{ ("conservation of gravitons").} \quad (35.77f)$$

(d) The correction $B_{\mu\nu}$ to the amplitude obeys

$$B_{\mu\alpha} k^{\alpha} = i A_{\mu\alpha}^{|\alpha}, \quad (35.77g)$$

$$B_{\mu\nu|\alpha} k^{\alpha} = -\frac{1}{2} k^{\alpha}_{|\alpha} B_{\mu\nu} + \frac{1}{2} i A_{\mu\nu|\alpha}^{\alpha} + i R_{\alpha\mu\beta\nu}^{(B)} A^{\alpha\beta}. \quad (35.77h)$$

In accordance with exercise 35.13, specialize the gauge so that $\bar{h} = 0$, i.e.,

$$e_{\alpha}^{\alpha} = 0. \quad (35.77i)$$

Then show that the stress-energy tensor (35.70') for the waves is

$$T_{\mu\nu}^{(\text{GW})} = \frac{1}{32\pi} \mathcal{A}^2 k_\mu k_\nu. \quad (35.77j)$$

This has the same form as the stress-energy tensor for a beam of particles with zero rest mass (see §5.4). Show explicitly that $T^{(\text{GW})\mu\nu}_{\nu} = 0$.

Exercise 35.16. GRAVITONS

Show that geometric optics, as developed in the preceding exercise, is equivalent to the following: “A graviton is postulated to be a particle of zero rest mass and 4-momentum \mathbf{p} , which moves along a null geodesic ($\nabla_{\mathbf{p}}\mathbf{p} = 0$). It parallel transports with itself ($\nabla_{\mathbf{p}}\mathbf{e} = 0$) a transverse ($\mathbf{e} \cdot \mathbf{p} = 0$) traceless ($e_\alpha^\alpha = 0$) polarization tensor \mathbf{e} . Geometric optics is the theory of a stream of such gravitons moving through spacetime.” Exhibit the relationship between the quantities in this version of geometric optics and the quantities in the preceding version (e.g., $\mathbf{p} = \hbar\mathbf{k}$, where \hbar is Planck’s reduced constant $h/2\pi$).

Exercise 35.17. GRAVITATIONAL DEFLECTION OF GRAVITATIONAL WAVES

Show that gravitational waves of short wavelength passing through the solar system experience the same redshift and gravitational deflection as does light. (One should be able to infer this directly from exercise 35.15.)

Exercise 35.18. GAUGE INVARIANCE OF $T_{\mu\nu}^{(\text{GW})}$

Show that the stress-energy tensor $T_{\mu\nu}^{(\text{GW})}$ of equation (35.70) is invariant under gauge transformations of the form (35.65).

Exercise 35.19. $T_{\mu\nu}^{(\text{GW})}$ EXPRESSED AS THE AVERAGE OF A STRESS-ENERGY PSEUDOTENSOR

Calculate the average over several wavelengths of the Landau-Lifshitz stress-energy pseudotensor [equation (20.22)] for gravitational waves with $\lambda/\mathcal{R} \ll 1$. The result should be equal to $T_{\mu\nu}^{(\text{GW})}$. [Hint: Work in a gauge where $\bar{h}_\mu^\alpha{}_{|\alpha} = \bar{h} = 0$, to simplify the calculation.]

Exercise 35.20. SHORTWAVE APPROXIMATION FROM A VARIATIONAL VIEWPOINT

Readers who have studied the variational approach to gravitation theory in Chapter 21 may find attractive the following derivation of the basic equations of the shortwave approximation. It was devised, independently, by Sándor Kovács and Bernard Schutz, and by Bryce DeWitt (unpublished, 1971). MacCallum and Taub (1973) give a “non-Palatini” version.

(a) Define

$$g_{\mu\nu} \equiv g_{\mu\nu}^{(\text{B})} + h_{\mu\nu}, \quad \bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2} g_{\mu\nu}^{(\text{B})} h, \quad (35.78a)$$

$$W^\mu{}_{\beta\gamma} \equiv \frac{1}{2} g_{\beta\gamma}^{(\text{B})} (h_{\alpha\beta}{}_{|\gamma} + h_{\alpha\gamma}{}_{|\beta} - h_{\beta\gamma}{}_{|\alpha}). \quad (35.78b)$$

Raise and lower indices on $h_{\mu\nu}$ and $W^\mu{}_{\beta\gamma}$ with the background metric. Using the results of exercise 35.11, derive the following expression for the Lagrangian of the gravitational field:

$$\mathcal{L} \equiv \frac{1}{16\pi} (-g)^{1/2} R = \mathcal{L}' + \left(\begin{array}{l} \text{perfect divergence} \\ \text{of form } \partial\mathcal{L}^\alpha/\partial x^\alpha \end{array} \right) + \left(\begin{array}{l} \text{corrections of order} \\ \mathcal{A}^3/\lambda^2, R_{\mu\nu}^{(\text{B})}\mathcal{A}, \text{and smaller} \end{array} \right), \quad (35.78c)$$

where

$$\begin{aligned}\mathcal{L}' \equiv \frac{1}{16\pi} (-g^{(B)})^{1/2} [R^{(B)} - \bar{h}^{\mu\nu} (W^\alpha_{\mu\nu|\alpha} - W^\alpha_{\mu\alpha|\nu}) \\ + g^{(B)\mu\nu} (W^\alpha_{\beta\alpha} W^\beta_{\mu\nu} - W^\alpha_{\beta\nu} W^\beta_{\mu\alpha})].\end{aligned}\quad (35.78d)$$

[Hint: recall that

$$(-g^{(B)})^{1/2} B^\alpha_{|\alpha} = \partial [(-g^{(B)})^{1/2} B^\alpha] / \partial x^\alpha$$

for any B^α .] Drop the corrections of order \mathcal{A}^3/λ^2 from \mathcal{L} ; and, knowing in advance that the field equations will demand $R_{\mu\nu}^{(B)} \sim \mathcal{A}^2/\lambda^2$, drop also the corrections of order $R_{\mu\nu}^{(B)} \mathcal{A}$. Knowing that a perfect divergence contributes nothing in an extremization calculation, drop the divergence term from \mathcal{L} . Then \mathcal{L}' is the only remaining part of \mathcal{L} .

(b) Extremize $I \equiv \int \mathcal{L}' d^4x$ by the Palatini method (§21.2); i.e., abandon (temporarily) definition (35.78b) of $W^\mu_{\beta\gamma}$, and extremize I with respect to independent variations of $W^\mu_{\beta\gamma} = W^\mu_{\gamma\beta}$, $\bar{h}^{\mu\nu} = \bar{h}^{\nu\mu}$, and $g_{(B)}^{\mu\nu} = g_{(B)}^{\nu\mu}$. Show that extremization with respect to $W^\mu_{\beta\gamma}$ leads back to equation (35.78b) for $W^\mu_{\beta\gamma}$ in terms of $h_{\mu\nu}$. Show that extremization with respect to $\bar{h}^{\mu\nu}$, when combined with equations (35.78a,b), leads to the propagation equation for gravitational waves (35.64). Show that extremization with respect to $g^{(B)\mu\nu}$, when combined with equations (35.78a,b) and with the propagation equation (35.64), and when averaged over several wavelengths, leads to

$$G_{\mu\nu}^{(B)} = 8\pi T_{\mu\nu}^{(\text{GW})},$$

where $T_{\mu\nu}^{(\text{GW})}$ is given by equation (35.70). [Warning: The amount of algebra in this exercise is enormous, unless one chooses to impose the gauge conditions $\bar{h} = \bar{h}_\alpha^\beta|_\beta = 0$ from the outset.]

CHAPTER 36

GENERATION OF GRAVITATIONAL WAVES

Matter is represented by curvature, but not every curvature does represent matter; there may be curvature "in vacuo."

G. LEMAITRE in Schilpp (1949), p. 440

Generation of gravitational waves analyzed by electromagnetic analog

§36.1. THE QUADRUPOLE NATURE OF GRAVITATIONAL WAVES

Masses in an isolated, nearly Newtonian system move about each other. How much gravitational radiation do they emit?

For an order-of-magnitude estimate, one can apply the familiar radiation formulas of electromagnetic theory, with the replacement $e^2 \rightarrow -m^2$, which converts the static Coulomb law into Newton's law of attraction. This procedure treats gravity as though it were a spin-one (vector) field, rather than a spin-two (tensor) field; consequently, it introduces moderate errors in numerical factors and changes angular distributions. But it gives an adequate estimate of the total power radiated.

In electromagnetic theory, electric-dipole radiation dominates, with a power output or "luminosity," L , given (see §4.4 and Figure 4.6) by

$$L_{\text{electric dipole}} = (2/3)e^2\mathbf{a}^2$$

for a single particle with acceleration \mathbf{a} and dipole moment changing as $\ddot{\mathbf{d}} = e\ddot{\mathbf{x}} = e\mathbf{a}$;

$$L_{\text{electric dipole}} = (2/3)\ddot{\mathbf{d}}^2$$

for a general system with dipole moment \mathbf{d} . [Geometric units: luminosity in cm of mass-energy per cm of light travel time; charge in cm, $e = (G^{1/2}/c^2)e_{\text{conv}} = (2.87 \times 10^{-25} \text{ cm/esu}) \times (4.8 \times 10^{-10} \text{ esu}) = 1.38 \times 10^{-34} \text{ cm}$, acceleration in cm of distance per cm of time per cm of time. For conventional units, with e in esu or $(\text{g cm}^3/\text{sec}^2)^{1/2}$,

insert a factor c^{-3} on the right and get L in erg/sec]. The gravitational analog of the electric dipole moment is the mass dipole moment

Why gravitational waves cannot be dipolar

$$\mathbf{d} = \sum_{\text{particles } A} m_A \mathbf{x}_A.$$

Its first time-rate of change is the total momentum of the system,

$$\dot{\mathbf{d}} = \sum_{\text{particles } A} m_A \dot{\mathbf{x}}_A = \mathbf{p}.$$

The second time-rate of change of the mass dipole moment has to vanish because of the law of conservation of momentum, $\ddot{\mathbf{d}} = \dot{\mathbf{p}} = 0$. Therefore *there can be no mass dipole radiation in gravitation physics*.

The next strongest types of electromagnetic radiation are magnetic-dipole and electric-quadrupole. Magnetic-dipole radiation is generated by the second time-derivative of the magnetic moment, $\ddot{\mu}$. Here again the gravitational analog is a constant of the motion, the angular momentum,

$$\boldsymbol{\mu} = \sum_A (\text{position of } A) \times (\text{current due to } A) = \sum_A \mathbf{r}_A \times (m \mathbf{v}_A) = \mathbf{J};$$

so it cannot radiate. *Thus, there can be no gravitational dipole radiation of any sort.*

When one turns to quadrupole radiation, one finally gets a nonzero result (see Figure 36.1). The power output predicted by electromagnetic theory,

$$L_{\text{electric quadrupole}} = \frac{1}{20} \ddot{\mathbf{Q}}^2 \equiv \frac{1}{20} \ddot{Q}_{jk} \ddot{Q}_{jk},$$

$$Q_{jk} \equiv \sum_A e_A \left(x_{Aj} x_{Ak} - \frac{1}{3} \delta_{jk} r_A^2 \right)$$

(Q_{jk} here = Q_{jk} in much other literature), has as its gravitational counterpart

$$L_{\text{mass quadrupole}} = \frac{1}{5} \langle \ddot{\mathbf{I}}^2 \rangle \equiv \frac{1}{5} \langle \ddot{I}_{jk} \ddot{I}_{jk} \rangle, \quad (36.1)$$

$$I_{jk} \equiv \sum_A m_A \left(x_{Aj} x_{Ak} - \frac{1}{3} \delta_{jk} r_A^2 \right) = \int \rho \left(x_j x_k - \frac{1}{3} \delta_{jk} r^2 \right) d^3x. \quad (36.2)$$

Gravitational-wave power output expressed in terms of "reduced quadrupole moment" of source

Formula (36.1) contains the correct factor of $1/5$, which comes from tensor calculations (see §36.10), instead of the incorrect factor $1/20$ suggested by the electromagnetic analog; and the righthand side of (36.1) has been averaged ("⟨ ⟩") over several characteristic periods of the source to accord with one's inability to localize the energy of gravitational radiation inside a wavelength.

Source

Receptor

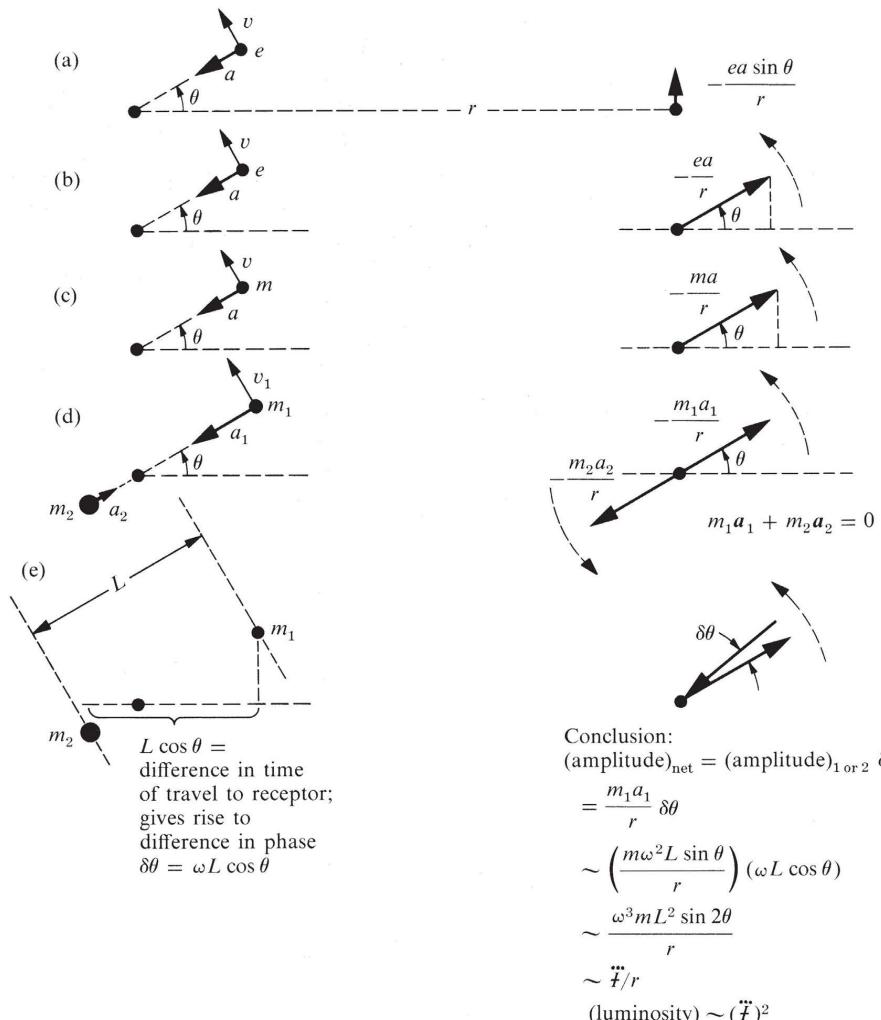


Figure 36.1

Why gravitational radiation is ordinarily weak. In brief, contributions to the amplitude of the outgoing wave from the mass dipole moments of the separate masses cancel, $(m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2)/r = 0$ (principle that action equals reaction).

(a) Radiation from an accelerated charge (see §4.4 and Figure 4.6).

(b) Representation of the field at the great distance r in terms of the typical rotating-vector diagram of electrical engineering; however, here, for ease of visualization, the *vertical* projection of the rotating vector gives the observed field (usual dipole-radiation field produced by a charge in circular orbit).

(c) Corresponding rotating-vector diagram for gravitational radiation, based on the simplified model of the gravitation field as a spin-one or vector field (to be contrasted with its true tensor character; hence details of angular distribution and total radiation as given by this simple diagram are not correct; but order of magnitude of luminosity is correct).

(d) The two masses m_1 and m_2 that hold each other in orbit give equal and opposite contributions to the amplitude of the outgoing wave because of the principle that action equals reaction. (In electromagnetic radiation from a hydrogen atom, the corresponding radiation amplitudes do not cancel: $e_{\text{elec}}\alpha_{\text{elec}} + e_{\text{prot}}\alpha_{\text{prot}} \sim e_{\text{elec}}\alpha_{\text{elec}} \neq 0$).

(e) In a better approximation, one has to allow for the difference in time of arrival at the receptor of the effects from the two masses. The two vectors that formerly opposed each other exactly are now drawn inclined, at the phase angle $\delta\theta$. The amplitude of the resulting field goes as \vec{r} , where \vec{r} is the reduced quadrupole moment; and the luminosity is proportional to \vec{r}^2 .

Notation: There is no ambiguity about the definition of the “second moment of the mass distribution” as it appears throughout the physics and mathematics literature

$$I_{jk} \equiv \int \rho x_j x_k d^3x.$$

Nor is there any ambiguity about how one constructs the moment of inertia tensor \mathcal{I}_{jk} from this second moment of the mass distribution:

$$\mathcal{I}_{jk} = \delta_{jk} \text{ trace } (I_{ab}) - I_{jk} = \int \rho(r^2 \delta_{jk} - x_j x_k) d^3x.$$

The moments that characterize a source radiating quadrupole gravitational radiation are here taken, equally unambiguously, to be the “trace-free part of the second moment of the mass distribution”:

$$\mathcal{I}_{jk} = I_{jk} - \frac{1}{3} \delta_{jk} \text{ trace } (I_{ab}) = I_{jk} - \frac{1}{3} \delta_{jk} I = \int \rho(x_j x_k - \frac{1}{3} \delta_{jk} r^2) d^3x. \quad (36.3)$$

This notation is adopted because it simplifies formulas, it simplifies calculations, it meshes well with much of the literature of gravitational-wave theory [e.g. Peters (1964), Peres and Rosen (1964)], and it is easy to remember. Another name for the quantities \mathcal{I}_{jk} is *reduced quadrupole moment*. This terminology makes clear the distinction between the quantities used here and the three-times-larger quantities that are called quadrupole moments in the standard text of Landau and Lifshitz (1962) and in the literature on nuclear quadrupole moments, and the 3/2-times-larger quantities used in the theory of spherical harmonics:

$$Q_{zz} \left(\begin{array}{l} \text{Landau and Lifshitz; also} \\ \text{nuclear quadrupole moments} \end{array} \right) = \int \rho(3z^2 - r^2) d^3x,$$

$$Q_{zz} \left(\text{theory of spherical harmonics} \right) = \int \rho\left(\frac{3}{2}z^2 - \frac{1}{2}r^2\right) d^3x,$$

$$\mathcal{I}_{zz} \left(\begin{array}{l} \text{reduced quadrupole moment;} \\ \text{unambiguous measure of} \\ \text{source strength adopted here} \end{array} \right) = \int \rho\left(z^2 - \frac{1}{3}r^2\right) d^3x.$$

Thus the \mathcal{I}_{jk} notation has the merit of circumventing the existing ambiguity in the literature.

That electromagnetic radiation is predominantly dipolar (spherical-harmonic index $l = 1$), and gravitational radiation is quadrupolar ($l = 2$) are consequences of a general theorem. Consider a classical radiation field, whose associated quantum mechanical particles have integer spin S , and zero rest mass. Resolve that radiation field into spherical harmonics—i.e., into multipole moments. All components with $l < S$ will vanish; in general those with $l \geq S$ will not; and this is independent of the nature of the source! [See, e.g., Couch and Newman (1972).] Since the lowest nonvanishing multipoles generally dominate for a slowly moving source (speeds $\ll c$), electromagnetic radiation ($S = 1$) is ordinarily dipolar ($l = S = 1$), while gravita-

Why gravitational waves are
ordinarily quadrupolar

tional radiation ($S = 2$) is ordinarily quadrupolar ($l = S = 2$). Closely connected with this theorem is the “topological fixed-point theorem” [e.g., Lifshitz (1949)], which distinguishes between scalar, vector, and tensor fields. For a scalar disturbance, such as a pressure wave, there is no difficulty in having a spherically symmetric source. Thus, over a sphere of a great radius r , there is no difficulty in having a pressure field that everywhere, at any one time, takes on the same value p . In contrast, there is no way to lay down on the surface of a 2-sphere a continuous vector field, the magnitude of which is non-zero and everywhere the same (“no way to comb smooth the hair on the surface of a billiard ball”). Likewise, there is no way to lay down on the surface of a 2-sphere a continuous non-zero transverse-traceless 2×2 matrix field that differs from one point to another at most by a rotation. Topology thus excludes the possibility of any spherically symmetric source of gravitational radiation whatsoever.

§36.2. POWER RADIATED IN TERMS OF INTERNAL POWER FLOW

Expression (36.1) for the power output can be rewritten in a form that is easier to use in order-of-magnitude estimates. Notice that the reduced quadrupole moment is

$$\begin{aligned} \ddot{t}_{jk} &\sim \frac{\left(\text{mass of that part of system which moves}\right) \times \left(\text{size of system}\right)^2}{\left(\text{time for masses to move from one side of system to other}\right)^3} = \frac{MR^2}{T^3} \\ &\sim \frac{M(R/T)^2}{T} \sim \frac{\left(\text{nonspherical part of kinetic energy}\right)}{T}; \\ \ddot{t}_{jk} &\sim L_{\text{internal}} \equiv \left(\frac{\text{power flowing from one side of system to other}}{T} \right); \end{aligned} \quad (36.4)$$

Gravitational-wave power output in terms of internal power flow of source

Consequently, equation (36.1) says that *the power output in gravitational waves (“luminosity”) is roughly the square of the internal power flow*

$$L_{\text{GW}} \sim (L_{\text{internal}})^2. \quad (36.5)$$

If this equation seems crazy (who but a fool would equate a power to the square of a power?), recall that in geometrized units power is dimensionless. The conversion factor to conventional units is

$$L_o \equiv c^5/G = 3.63 \times 10^{59} \text{ erg/sec} = 2.03 \times 10^5 M_{\odot}c^2/\text{sec}. \quad (36.6)$$

One may freely insert this factor of $L_o = 1$ wherever one wishes in order to feel more comfortable with the appearance of the equations. For example, one can rewrite equation (36.5) in the form

$$L_{\text{GW}}/L_{\text{internal}} \sim L_{\text{internal}}/L_o. \quad (36.7)$$

In applying the equation $L_{\text{GW}} \sim (L_{\text{internal}})^2$, one must be careful to ignore those internal power flows that cannot radiate at all, i.e., those that do not accompany a time-changing quadrupole moment. For example, in a star the internal power flows associated with spherical pulsation and axially symmetric rotation must be ignored.

Conservation of energy guarantees that radiation reaction forces will pull down the internal energy of the system at the same rate as gravitational waves carry energy away (see Box 19.1). The characteristic time-scale for radiation reaction to change the system markedly is

$$\begin{aligned}\tau_{\text{react}} &\sim [1/(\text{rate at which energy is lost})] \times [\text{energy in motions that radiate}] \\ &\sim [1/L_{\text{GW}}] \times [(L_{\text{internal}}) \times (\text{characteristic period } T \text{ of internal motions})] \\ &\sim (L_{\text{internal}}/L_{\text{GW}})T \sim (L_0/L_{\text{internal}})T.\end{aligned}\quad (36.8)$$

Characteristic time-scale for radiation-reaction effects

Consequently, *radiation reaction is important in one characteristic period only if the system achieves the enormous internal power flow*

$$L_{\text{internal}} \gtrsim L_o = 3.63 \times 10^{59} \text{ ergs/sec} = 1!$$

§36.3. LABORATORY GENERATORS OF GRAVITATIONAL WAVES

As a laboratory generator of gravitational waves, consider a massive steel beam of radius $r = 1$ meter, length $l = 20$ meters, density $\rho = 7.8 \text{ g/cm}^3$, mass $M = 4.9 \times 10^8 \text{ g}$ (490 tons), and tensile strength $t = 40,000$ pounds per square inch or $3 \times 10^9 \text{ dyne/cm}^2$. Let the beam rotate about its middle (so it rotates end over end), with an angular velocity ω limited by the balance between centrifugal force and tensile strength

$$\omega = (8t/\rho l^2)^{1/2} = 28 \text{ radians/sec.}$$

Power output from a rotating steel beam

The internal power flow is

$$\begin{aligned}L_{\text{internal}} &= \left(\frac{1}{2} I\omega^2\right)\omega = \frac{1}{24} Ml^2\omega^3 \\ &\approx 2 \times 10^{18} \text{ erg/sec} \approx 10^{-41} L_o.\end{aligned}$$

The order of magnitude of the power radiated is

$$L_{\text{GW}} \sim (10^{-41})^2 L_o \sim 10^{-23} \text{ erg/sec.} \quad (36.9)$$

(An exact calculation using equation (36.1) gives $2.2 \times 10^{-22} \text{ erg/sec}$; see Exercise 36.1.) Evidently the construction of a laboratory generator of gravitational radiation is an unattractive enterprise in the absence of new engineering or a new idea or both.

To rely on an astrophysical source and to build a laboratory or solar-system detector is a more natural policy to consider. Detection will be discussed in the next chapter. Here attention focuses on astrophysical sources.

EXERCISE**Exercise 36.1. GRAVITATIONAL WAVES FROM ROTATING BEAM**

A long steel beam of length l and mass M rotates end over end with angular velocity ω . Show that the power it radiates as gravitational waves is

$$L_{\text{GW}} = \frac{2}{45} M^2 l^4 \omega^6. \quad (36.10)$$

Use this formula to verify that the rod described in the text radiates 2.2×10^{-22} ergs/sec.

§36.4. ASTROPHYSICAL SOURCES OF GRAVITATIONAL WAVES: GENERAL DISCUSSION

Consider a highly dynamic astrophysical system (a star pulsating and rotating wildly, or a collapsing star, or an exploding star, or a chaotic system of many stars). If its mass is M and its size is R , then according to the virial theorem (exercise 39.6) its kinetic energy is $\sim M^2/R$. The characteristic time-scale for mass to move from one side of the system to the other, T , is

$$T \sim \frac{R}{(\text{mean velocity})} \sim \frac{R}{(M/R)^{1/2}} = \left(\frac{R^3}{M}\right)^{1/2} \quad (36.11a)$$

(\sim time of free fall; \sim time to turn one radian in Kepler orbit; see Chapter 25). Consequently, the internal power flow is

$$L_{\text{internal}} \sim \frac{(\text{kinetic energy})}{T} \sim \left(\frac{M^2}{R}\right) \left(\frac{M}{R^3}\right)^{1/2} \sim \left(\frac{M}{R}\right)^{5/2}. \quad (36.11b)$$

Power output from violent astrophysical sources, in terms of mass and radius

The gravitational-wave output or “luminosity” is the square of this quantity, or

$$L_{\text{GW}} \sim (M/R)^5 L_o. \quad (36.11c)$$

(If the system is rather symmetric, or if only a small portion of its mass is in motion, then its quadrupole moment does not change much, and the estimate of L_{GW} must be reduced accordingly. The wave amplitude goes down in proportion to the fraction of the mass in motion, and the power is reduced in proportion to the square of that fraction.)

Clearly, *the maximum power output occurs when the system is near its gravitational radius*; and because nothing, not even gravitational waves, can escape from inside the gravitational radius, *the maximum value of the output is $\sim L_o = 3.63 \times 10^{59}$ ergs/second, regardless of the nature of the system!*

Upper limit on power output

Actually, the above derivation of this limit and of equation (36.11c) uses approximations to general relativity that break down near the gravitational radius. [Velocities small compared to light are required in deriving the standard formula (36.1) for

L_{GW} (see §36.7); nearly Newtonian fields are required for the virial theorem arguments of (36.11a), as well as for the L_{GW} formula.] Nevertheless, in rough order of magnitude, equation (36.11c) is valid to quite near the Schwarzschild radius, say, $R \sim 3M$; and inside that point gravity is so strong that no system can resist collapse for an effective length of time much longer than $T \sim M$.

The time required for radiation-reaction forces to affect a system substantially [equation (36.8)] is of the order

$$\tau_{\text{react}} \sim (L_o/L_{\text{internal}})T \sim (R/M)^{5/2}T, \quad (36.11d)$$

where T is the characteristic time (36.11a) of rotation or free fall. (Note how one inserts and removes the factor $L_o = 1$ at will!) Consequently, *the effect of radiation reaction, as integrated over one period, is unimportant except when the system is near its gravitational radius.*

When a system such as a pulsating star is settling down into an equilibrium state, the radiation reaction will damp its internal motions. On the other hand, when the system, like a binary star system, is far from any state of equilibrium, then loss of energy (and angular momentum) to radiation under certain circumstances may speed up the angular velocity or speed up the internal motions and augment the radiation.

§36.5. GRAVITATIONAL COLLAPSE, BLACK HOLES, SUPERNOVAE, AND PULSARS AS SOURCES

Since $L_{\text{GW}} \sim (M/R)^5 L_o$, the most intense gravitational waves reaching Earth must come from a dynamic, deformed system near its gravitational radius (L_{GW} drops by a factor 100,000 with each increase by 10 of R !). The scenario of Figure 24.3 gives an impression of some of the dynamic processes that not only may happen but probably must happen. The sequence of events sketched out there includes pulses of gravitational radiation interspersed with intervals of continuous radiation of gradually increasing frequency. Pulse number one comes at the time of the original collapse of a star with white-dwarf core to a pancake-shaped neutron star. The details of what then goes on will differ enormously depending on the original mass and angular momentum of this “pancake.” In the illustration, this pancake fragments into a constellation of corevolving neutron stars, which then one by one undergo “pursuit and plunge.”

Whether in this kind of scenario or otherwise, perhaps the most favorable source of gravitational radiation is a star (the original very temporary “pancake” or one of the fragments therefrom) collapsing through its gravitational radius in a highly nonspherical manner. Such a star should terminate life with a last blast of gravitational waves, which carry off a sizeable fraction of its rest mass. Thus an order-of-magnitude estimate gives

$$\begin{aligned} (\text{energy radiated}) &= \int L_{\text{GW}} dt \sim L_o \cdot \left(\begin{array}{l} \text{time during which} \\ \text{peak luminosity occurs} \end{array} \right) \quad (36.12) \\ &\sim L_o M = M. \end{aligned}$$

Radiation reaction in astrophysical sources

Gravitational waves from:

- (1) stellar collapse and formation of a black hole

- (2) the fall of debris into a black hole

(Whether the energy radiated is $0.9M$, or $0.1M$, or $0.01M$ is not known for certain today; but it must lie in this range of orders of magnitude.) The radiation should be weak at low frequencies; it should rise to a peak at a frequency a little smaller than $1/M$; and it should cut off sharply for circular frequencies above $\omega \sim 1/M$.

Matter (“debris”; see Figure 24.3) falling into a black hole can also be a significant source of gravitational waves. The infalling matter will radiate only weakly when it is far from the gravitational radius; but as it falls through the gravitational radius (between $r \sim 4M$ and $r = 2M$), it should emit a strong burst. If m is the mass of an infalling lump of matter and M is the total mass of the black hole, then the total energy in the final burst is

$$E_{\text{radiated}} \sim m^2/M, \quad (36.13)$$

and it comes off in a time $\sim M$ with a power output of $L_{\text{GW}} \sim (m/M)^2 L_o$. (See exercise 36.2.) Actually, this is an extremely rough estimate of the energy output. In the limit where the infalling lump is small in both size and mass [(size of lump) \ll (gravitational radius of black hole); $m \ll M$; “delta-function lump”], one can perform an exact calculation of the spectrum and energy radiated by treating the lump and the waves as small perturbations on the Schwarzschild geometry of the black hole. The foundations for such a treatment were given by Zerilli (1970b). Zerilli’s formula was corrected and applied to the case of head-on impact by Davis, Ruffini, Press, and Price (1971). They predict the spectrum of Figure 36.2 and the total energy output

$$E_{\text{radiated}} = 0.0104m^2/M \quad (36.14)$$

for $m \ll M$ and (size of lump) $\ll M$.

A collision between black holes should also produce a strong burst of gravitational waves—through such collisions are probably very rare!

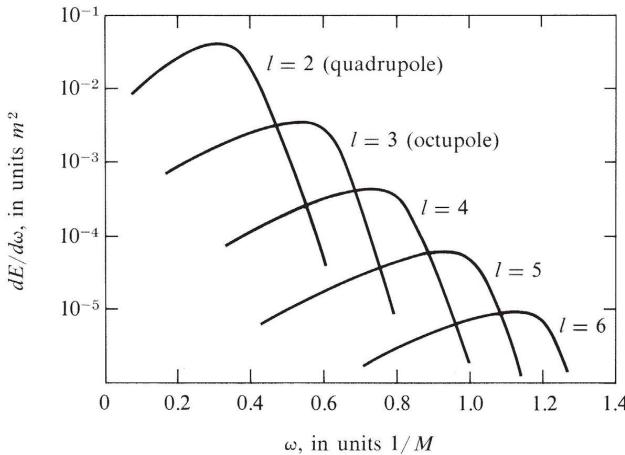
Not quite so rare, but still not common, are supernova explosions (about one per galaxy per 100 years). According to current theory as verified by pulsar observations, a supernova is triggered by the collapse of the core of a highly evolved star (see §24.3). The collapse itself and the subsequent wild gyrations of the collapsed core (neutron star) should produce a short, powerful burst of gravitational waves. The characteristics of the burst, as estimated with formulas (36.11), and assuming large departures from sphericity, are

$$\begin{aligned} (\text{energy radiated}) &\sim (\text{neutron-star binding energy}) \\ &\sim M^2/R \sim 0.1M \sim 10^{53} \text{ ergs}, \end{aligned}$$

$$(\text{mean frequency}) \sim 1/T \sim (M/R^3)^{1/2} \sim 0.03M^{-1} \sim 3000 \text{ Hz}, \quad (36.15)$$

$$(\text{power output}) \sim (M/R)^5 L_o \sim 10^{-5} L_o \sim 3 \times 10^{54} \text{ ergs/sec},$$

$$\left(\begin{array}{l} \text{time for gravitational} \\ \text{radiation to damp the} \\ \text{motion if turbulence,} \\ \text{heat conduction, and other} \\ \text{effects do not damp it} \\ \text{sooner} \end{array} \right) = \tau \sim M(M/R)^{-4} \sim 0.1 \text{ sec} \sim 300 \text{ periods.}$$

**Figure 36.2.**

Spectrum of the gravitational waves emitted by a “delta-function” lump of matter of mass m , falling head-on into a nonrotating (Schwarzschild) black hole of mass $M \gg m$. The total energy radiated is distributed among multipoles according to the empirical law

$$(\text{energy in } l\text{-pole waves}) \approx (0.44 m^2/M)e^{-2l},$$

and the total spectrum peaks at angular frequency

$$\omega_{\max} = 0.32/M.$$

These results were calculated by treating the infalling lump and the gravitational waves as small perturbations on the Schwarzschild geometry of the black hole. The relevant perturbation-theory equations were derived by Zerilli (1970), and were solved numerically to give these results by Davis, Ruffini, Press, and Price (1971).

In the last stages of the stellar pulsations, when the amplitude $\xi = \delta r$ has dropped to $\delta r/r \ll 1$, one can calculate the pulsation frequencies and damping times exactly by treating the fluid motions and gravitational waves as small perturbations of an equilibrium stellar model. The results of such a calculation, which are in good agreement with the above rough estimates, are shown in Box 36.1.

Long after the pulsations of the neutron star have been damped out by gravitational radiation reaction and by other forces, the star will continue to rotate; and as it rotates, carrying along with its rotation an off-axis-pointing magnetic moment, it will beam out the radio waves, light, and x-rays that astronomers identify as “pulsar radiation.” In this pulsar phase, gravitational radiation is important only if the star is somewhat deformed from axial symmetry (axial symmetry \Rightarrow constant quadrupole moment \Rightarrow no gravitational waves). According to estimates in exercise 36.3, a deformation that contains only 0.001 of the star’s mass could radiate 10^{38} ergs per second for the youngest known pulsar (Crab nebula); and the accompanying radiation reaction could be a significant source of the pulsar’s slowdown. However, it is not at all clear today (1973)—indeed, it seems unlikely—that the neutron star could support even so small a deformation.

(5) young pulsars

(continued on page 986)

Box 36.1 GRAVITATIONAL WAVES FROM PULSATING NEUTRON STARS

The table given here, taken from Thorne (1969a), shows various characteristics of the quadrupole oscillations of several typical neutron-star models. Note that the gravitational waves emitted by the most massive models (1) have frequencies $\nu = 1/T_n \sim 3,000$ Hz, (2) last for a time of $\sim \frac{1}{3}$ second, (3) damp out the stellar vibrations after only $\sim 1,000$ oscillations, and (4) carry off a total energy of $\sim (10^{54} \text{ ergs}) \times (\delta R/R)^2$, where $\delta R/R$ is the initial fractional amplitude of vibration of the star's surface.

These results are *not* based on the nearly Newtonian slow-

motion formalism of this chapter [equation (36.1), §§36.7 and 36.8], because that formalism is invalid here: the reduced wavelength of the radiation, $\lambda \sim 15$ km for waves from the most massive star, is not large compared to the star's gravitational radius, $2M \sim 6$ km; and the star's internal gravitational field is not weak (M/R as large as 0.29). Consequently, these results were derived using an alternative technique, which is valid for rapid motions and strong internal fields, but which assumes small perturbations away from the equilibrium stellar model. See Thorne (1969a) and papers cited therein for details.

QUADRUPOLE PULSATIONS OF NEUTRON STARS

Equation of state	ρ_c (g cm^{-3})	M/M_\odot	$2M/R$	n	T_n (msec)	τ_n (sec)	τ_n/T_n	E_M		Power			
								$\langle(\delta R/R)^2\rangle$	(ergs)	$\langle(\delta R/R)^2\rangle$	(ergs sec $^{-1}$)	$(\delta R/R)$	$(\delta r/r)_c$
H-W	3×10^{14}	0.405	0.0574	0	1.197	13.	11000	7.8×10^{50}	1.2×10^{50}	+	7.4	+	3.1
H-W	6×10^{15}	0.682	0.240	0	0.3109	0.19	610	2.8×10^{52}	2.9×10^{53}	+	5.2	+	3.7
				1	0.1713	0.28	1600	3.6×10^{51}	2.6×10^{52}	-	14.	-	3.3
				2	0.1179	1.3	11000	2.6×10^{50}	3.9×10^{50}	+	55.	+	5.9
				3	0.0938	24.	250000	8.9×10^{48}	$7. \times 10^{47}$	-	350.	-	24.
V_γ	5.15×10^{14}	0.677	0.159	0	0.6991	1.7	2400	5.7×10^{52}	$7. \times 10^{52}$	+	1.4	+	1.3
				1	0.2358	11.	47000	6.0×10^{50}	1.1×10^{50}	-	38.	-	4.7
V_γ	3×10^{15}	1.954	0.580	0	0.3777	0.22	600	1.7×10^{54}	1.6×10^{55}	+	1.9	+	3.1
				1	0.1556	1.6	10000	1.5×10^{54}	1.9×10^{54}	-	2.1	-	0.66
				2	0.1026	2.6	25000	5.2×10^{53}	4.0×10^{53}	+	2.9	+	0.40

The columns in the table have the following meanings.

Equation of state: the equation of state $p(\rho)$ used in constructing the equilibrium stellar model and in calculating the adiabatic index from $\gamma = [(\rho + p)/p] dp/d\rho$; H-W is the Harrison-Wheeler equation of state in the tabular form given by Hartle and Thorne (1968), Table 1; V_γ is the Levinger-Simmons-Tsuruta-Cameron V_γ equation of state in the tabular form given by Hartle and Thorne (1968), Table 2.

ρ_c : central density of total mass-energy for the equilibrium stellar model.

M/M_\odot : total mass-energy of the equilibrium model (i.e., the mass that governs distant Keplerian orbits), in units of the sun's mass.

$2M/R = 2GM/Rc^2$: ratio of the gravitational radius of the equilibrium model to its actual radius (radii are defined by $4\pi R^2 = \text{surface area}$).

n : the "order" of the pulsational normal mode under study (for all models given here, n is also the number of nodes in the radial relative eigenfunction, $\delta r/r$). Note: $n = 0$ is the fundamental (quadrupole) mode.

$T_n = 2\pi/\omega_n$: the pulsation period of the quasinormal mode measured in milliseconds.

τ_n : the damping time for the amplitude of the normal mode measured in seconds.

$\tau_n/T_n = \omega_n \tau_n / 2\pi$: the number of pulsation periods required for the amplitude to drop by a factor of $1/e$.

$E_{\mathcal{R}}/\langle(\delta R/R)^2\rangle$: energy of pulsation of the star, divided by the square of the relative amplitude of radial motion of the star's surface averaged over its surface.

Power/ $\langle(\delta R/R)^2\rangle$: the power radiated as gravitational waves, divided by the averaged square of the relative amplitude at the star's surface.

$(\delta R/R)(\delta r/r)_c^{-1}$: relative amplitude of radial motion at the star's surface divided by relative amplitude at the star's center.

$\delta\theta_s/\delta\theta_c = \delta\phi_s/\delta\phi_c$: amplitude of the angular displacement of the star's fluid at its surface divided by the same amplitude at its center.

Of the sources discussed in this section, most are “impulsive” rather than continuous (star collapsing through gravitational radius; debris falling into a black hole; collision between black holes; supernova explosion). They give rise to bursts of gravitational waves. An order-of-magnitude method of analyzing such bursts is spelled out in Box 36.2.

It is difficult and risky to pass from the above description of processes that should generate gravitational waves to an estimate of the characteristics of the waves that actually bathe the earth. For such an estimate, made in 1972 and subject to extensive revision as one’s understanding of the universe improves, see Press and Thorne (1972).

EXERCISES

Exercise 36.2. GRAVITATIONAL WAVES FROM MATTER FALLING INTO A BLACK HOLE

A lump of matter with mass m falls into a black hole of mass M . Show that a burst of gravitational waves is emitted with duration $\sim M$ and power $L_{\text{GW}} \sim (m/M)^2 L_o$, so that the total energy radiated is given in crude order of magnitude by equation (36.13).

Exercise 36.3. GRAVITATIONAL WAVES FROM A VIBRATING NEUTRON STAR

Idealize a neutron star as a sphere of incompressible fluid of mass M and radius R , with structure governed by Newton’s laws of gravity. Let the star pulsate in its fundamental quadrupole mode. Using Newtonian theory, calculate: the angular frequency of pulsation, ω ; the energy of pulsation E_{puls} ; the quantity $\frac{1}{5}\langle \ddot{\mathbf{r}}^2 \rangle$, which, according to equation (36.1), is the power radiated in gravitational waves, L_{GW} ; and the e -folding time, $\tau = E_{\text{puls}}/L_{\text{GW}}$, for radiating away the energy of the pulsations. Compare the answers with equations (36.15)—which are based on a much cruder approximation—and with the results in Box 36.1, which are based on much better approximations. [For solution, see Table 13 of Wheeler (1966).]

Exercise 36.4. PULSAR SLOWDOWN

The pulsar NPO532 in the Crab Nebula has a period of 0.033 seconds and is slowing down at the rate $dP/dt = 1.35 \times 10^{-5}$ sec/yr. Assuming the pulsar is a typical neutron star, calculate the rate at which it is losing rotational energy. If this energy loss is due primarily to gravitational radiation reaction, what is the magnitude of the star’s nonaxial deformation? [For solution, see Ferrari and Ruffini (1969); for a rigorous strong-field analysis, see Ipser (1970).]

§36.6. BINARY STARS AS SOURCES

Binary stars as sources of gravitational waves:

The most numerous sources of weak gravitational waves are binary star systems. Moreover, roughly half of all stars are in binary or multiple systems [see, for example, the compilation of Allen (1962)]. According to Kepler’s laws, two stars of masses m_1 and m_2 that circle each other have angular frequency ω and separation a coupled to each other by the formula

$$\omega^2 a^3 = m_1 + m_2 \equiv M.$$

Box 36.2 ANALYSIS OF BURSTS OF RADIATION FROM IMPULSE EVENTS*

	Electromagnetism	Gravitation
Typical moment relevant for radiation	$d_x(t)$	$\mathbf{f}_{xx}(t)$
Its Fourier transform	$(2\pi)^{-1/2} \int d_x \exp[i\omega t] dt$	$(2\pi)^{-1/2} \int \mathbf{f}_{xx} \exp[i\omega t] dt$
Name for this quantity	$d_x(\omega)$	$\mathbf{f}_{xx}(\omega)$
Time decomposition of total radiative energy loss ΔE	$c^{-3} \int \ddot{d}^2(t) dt$	$Gc^{-5} \int \ddot{\mathbf{f}}^2(t) dt$
Decomposition of ΔE according to circular frequency	$c^{-3} \int \ddot{d}^2(\omega) d\omega$	$Gc^{-5} \int \ddot{\mathbf{f}}^2(\omega) d\omega$
Integrand nearly constant with respect to ω from $\omega = 0$ up to a critical value of ω_{crit} , beyond which radiation falls off very fast	$\omega_{\text{crit}} \sim 1/\Delta t$	$\omega_{\text{crit}} \sim 1/\Delta t$
$-d \Delta E/d\omega$ for $\omega < \omega_{\text{crit}}$	$\sim c^{-3} \ddot{d}^2(0)$	$\sim Gc^{-5} \ddot{\mathbf{f}}^2(0)$
Zero frequency moment that enters this formula	$\sim (e_1 \Delta v_{x1} + e_2 \Delta v_{x2})$	$\Delta (\langle \text{Kinetic Energy} \rangle)_{xx}$
Rewrite of $-d \Delta E/d\omega$	$\sim (e \Delta v)^2/c^3$	$\sim G[\Delta (\langle \text{K.E.} \rangle)_{xx}]^2/c^5$
Total energy of pulse	$\sim \text{This}/\Delta t$	$\sim \text{This}/\Delta t$

* Box adapted from pp. 113 and 114 of Wheeler (1962).

As sample applications of this analysis, Wheeler (1962) cites the following:

Parameter	One atomic-nucleus fission of 180 MeV	Fission bomb yield 17 kilotons at 10% efficiency	Meteorite striking earth at escape velocity	Explosion of star when 10^{-4} of mass is released
Mass	$4 \times 10^{-22} \text{ g}$	10^4 g	10^9 g	$2 \times 10^{33} \text{ g}$
Velocity	$1.2 \times 10^9 \text{ cm/s}$	$4 \times 10^8 \text{ cm/s}$	$11 \times 10^5 \text{ cm/s}$	$4 \times 10^8 \text{ cm/s}$
Energy	$2.9 \times 10^{-4} \text{ erg}$	$7 \times 10^{20} \text{ erg}$	$6 \times 10^{20} \text{ erg}$	$1.8 \times 10^{50} \text{ erg}$
Fraction assumed relevant to radiative moment	1	0.1	1	0.1
Time integral of this moment $= \langle \text{K.E.} \rangle_{xx}$	$2.9 \times 10^{-4} \text{ erg}$	$7 \times 10^{19} \text{ erg}$	$6 \times 10^{20} \text{ erg}$	$1.8 \times 10^{49} \text{ erg}$
$\langle \text{K.E.} \rangle_{xx}/c^2$	$3.2 \times 10^{-25} \text{ g}$	0.08 g	0.67 g	$2 \times 10^{28} \text{ g}$
$\frac{dE}{d\omega} \sim \frac{G}{c} \left(\frac{\langle \text{K.E.} \rangle_{xx}}{c^2} \right)^2$	$2.3 \times 10^{-67} \frac{\text{erg}}{\text{rad/s}}$	$1.4 \times 10^{-20} \frac{\text{erg}}{\text{rad/s}}$	$1.0 \times 10^{-18} \frac{\text{erg}}{\text{rad/s}}$	$9 \times 10^{38} \frac{\text{erg}}{\text{rad/s}}$
Δt	10^{-21} s	10^{-8} s	10^{-3} s	10^4 s
$\Delta\omega \sim 1/\Delta t$	10^{21} rad/s	10^8 rad/s	10^3 rad/s	10^{-4} rad/s
$\Delta E_{\text{radiated}}$	10^{-46} erg	10^{-12} erg	10^{-15} erg	10^{35} erg
Assumed distance to detector	10^3 cm	10^3 cm	10^9 cm	10^{23} cm
$\Delta E/4\pi r^2$	$10^{-53} \text{ erg/cm}^2$	$10^{-19} \text{ erg/cm}^2$	$10^{-34} \text{ erg/cm}^2$	$10^{-12} \text{ erg/cm}^2$

The reader might find it informative to extend this table to the bursts of waves emitted by (1) debris falling into a black hole, (2) collisions between two black holes, and (3) a supernova explosion in which a star of two solar masses collapses to nuclear densities, ejecting half its mass in the process.

In this motion the kinetic energy is

$$(\text{kinetic energy}) = -\frac{1}{2} (\text{potential energy}) = \frac{1}{2} \frac{m_1 m_2}{a}.$$

The power that they radiate as gravitational waves can be estimated roughly as the square of the circulating power, $L \sim \omega \times (\text{kinetic energy})$; thus,

$$L_{\text{GW}} \sim \frac{\mu^2 M^3}{4a^5} L_o,$$

where $\mu = m_1 m_2 / M$ is the familiar reduced mass, and $M = m_1 + m_2$ is the total mass of this binary system.

An *exact* calculation based on equation (36.1) gives a result larger than this by a factor ~ 30 : for a binary system of semimajor axis a and eccentricity ϵ , the power output averaged over an orbital period is

$$(1) \text{ power output} \quad L_{\text{GW}} = \frac{32}{5} \frac{\mu^2 M^3}{a^5} f(\epsilon) L_o, \quad (36.16a)$$

where $f(\epsilon)$ is the dimensionless “correction function,”

$$f(\epsilon) = \left[1 + \frac{73}{24} \epsilon^2 + \frac{37}{96} \epsilon^4 \right] [1 - \epsilon^2]^{-7/2}. \quad (36.16b)$$

[See exercise 36.6 at end of §36.8; also Peters and Mathews (1963).]

(2) effects of radiation reaction

As the binary system loses energy by gravitational radiation, the stars spiral in toward each other (decrease of energy; tightening of gravitational binding). For circular orbits the energy, $E = -\frac{1}{2}m_1 m_2 / a = -\frac{1}{2}\mu M / a$, decreases as

$$\begin{aligned} dE/dt &= 1/2(\mu M/a^2)(da/dt) \\ &= -L_{\text{GW}} = -\frac{32}{5} \frac{\mu^2 M^3}{a^5}. \end{aligned}$$

Consequently, the evolution of the orbital radius is given by the formula

$$a = a_o(1 - t/\tau_o)^{1/4}, \quad (36.17a)$$

where $a_o = a_{\text{today}}$ and

$$\tau_o = \frac{1}{4} \left(\frac{-E}{L_{\text{GW}}} \right)_{\text{today}} = \frac{5}{256} \frac{a_o^4}{\mu M^2}. \quad (36.17b)$$

Thus, unless nongravitational forces intervene, the two stars will spiral together in a time τ_o (*spiral time*). For an elliptical orbit, the eccentricity also evolves. Radiation is emitted primarily at periastron. Therefore the braking forces of radiation reaction act there with greatest force. This effect deprives the stars of some of the kinetic energy of the excursions in their separation (“radial kinetic energy”). In consequence, the orbit becomes more nearly circular. [See Peters and Mathews (1963) for detailed calculations.]

The calculated power output, flux at Earth, and damping times are shown in Box 36.3 for several known binary stars and several interesting hypothetical cases. Notice that in the most favorable known cases the period is a few hours; the damping time is the age of the universe (could the absence of better cases be due to radiation reaction's having destroyed them?); the output of power in the form of gravitational waves is $\sim 10^{30}$ to 10^{32} ergs/sec (approaching the light output of the sun, 3.9×10^{33} ergs/sec); and the calculated flux at the Earth is $\sim 10^{-10}$ to 10^{-12} ergs/sec (too small to detect in 1973, but perhaps not too small several decades hence; see Chapter 37).

The hypothetical cases in Box 36.3 illustrate the general relations for astrophysical systems that were derived in §36.4—namely, that only as the system approaches its gravitational radius can L_{GW} approach L_o , and only then can damping remove nearly the whole energy in a single period.

§36.7. FORMULAS FOR RADIATION FROM NEARLY NEWTONIAN SLOW-MOTION SOURCES

Turn now from illustrative astrophysical sources to rigorous formulas valid for a wide variety of sources. One such formula has already been written down,

$$L_{\text{GW}} = \frac{1}{5} \langle \ddot{\vec{x}}_{jk} \ddot{\vec{x}}_{jk} \rangle, \quad (36.1)$$

but it has not yet been derived, nor has its realm of validity been discussed.

This formula for the power output is actually valid for any “nearly Newtonian, slow-motion source”—more particularly, for any source in which

$$(\text{size of source})/(\text{reduced wavelength of waves}) \ll 1, \quad (36.18a)$$

$$|\text{Newtonian potential}| \ll (\text{size of source})/(\text{reduced wavelength}), \quad (36.18b)$$

$$\frac{|\text{typical stresses}|}{(\text{mass density})} \ll \frac{(\text{size of source})}{(\text{reduced wavelength})}. \quad (36.18c)$$

It is not valid, except perhaps approximately, for fast-motion or strong-field sources. Moreover, there is no formalism available today which can handle effectively and *in general* the fast-motion case or the strong-field case.

The rest of this chapter is devoted to a detailed analysis of gravitational waves from nearly Newtonian, slow-motion sources. But the analysis (Track 2; §§36.9–36.11) will be preceded by a Track-1 summary in this section and the next.

For any source of size R and mean internal velocity v , the characteristic reduced wavelength ($\lambda = \lambda/2\pi$) of the radiation emitted is $\lambda \sim (\text{amplitude of motions})/v \lesssim R/v$. Consequently the demand (36.18a) that R/λ be $\ll 1$ [i.e., that the source be confined to a small region deep inside the near (nonradiation) zone] enforces the slow-motion constraint

$$v \ll 1.$$

(3) particular binaries observed by astronomers

The “nearly Newtonian, slow-motion approximation” for analyzing sources of gravitational waves

Box 36.3 GRAVITATIONAL RADIATION FROM SEVERAL BINARY STAR SYSTEMS^a

Type of system	Name	Period	$\frac{m_1}{M_\odot}$	$\frac{m_2}{M_\odot}$	Distance from earth (pc)	Spiral time ^b (ergs/sec)	L_{grav} (ergs/sec)	Flux at earth (erg/sec cm ²)
Solar System (Sun + Jupiter)	Solar System	11.86 yr.	1.0	9.56×10^{-4}	Earth is in near zone	2.5×10^{23} yr	5.2×10^{-10}	—
Typical resolved binaries from compilation of Van de Kamp (1958)	η Cas ξ Boo Sirius Fu 46	480 yr. 149.95 yr. 49.94 yr. 13.12 yr.	0.94 0.85 2.28 0.31	0.58 0.75 0.98 0.25	5.9 6.7 2.6 6.5	9.5×10^{24} 3.8×10^{23} 7.2×10^{21} 3.2×10^{21}	5.6×10^{10} 3.6×10^{12} 1.1×10^{15} 3.6×10^{14}	1.4×10^{-29} 6.7×10^{-28} 1.3×10^{-24} 7.1×10^{-26}
Typical eclipsing binaries from compilation of Gaposkin (1958)	β Lyr UWC Ma β Per WUMa	12.925 day 4.395 day 2.867 day 0.33 day	19.48 40.0 4.70 0.76	9.74 31.0 0.94 0.57	330 1470 30 110	7.0×10^{11} 8.2×10^8 3.2×10^{11} 6.2×10^9	0.057×10^{30} $49. \times 10^{30}$ 0.014×10^{30} 0.47×10^{30}	0.0004×10^{-11} 0.019×10^{-11} 0.013×10^{-11} 0.032×10^{-11}
Favorable cases from compilation of Braginsky (1965)	UV Leo V Pup i Boo YY Eri SW Lac WZ Sge	0.6 day 1.45 day 0.268 day 0.321 day 0.321 day 81 min	1.36 16.6 1.35 0.76 0.97 0.6	1.25 9.8 0.68 0.50 0.83 0.03	68 390 12 42 75 100	1.0×10^{10} 2.3×10^9 2.0×10^9 6.6×10^9 3.5×10^9 1.1×10^9 yr	0.63×10^{30} $65. \times 10^{30}$ 3.2×10^{30} 0.42×10^{30} 1.5×10^{30} 0.5×10^{30}	0.012×10^{-11} 0.36×10^{-11} $18. \times 10^{-11}$ 0.20×10^{-11} 0.21×10^{-11} 0.04×10^{-11}
Hypothetical binaries (neutron stars or black holes)	10^4 km 10^3 km 10^2 km 10 km	12.2 sec 0.39 sec 12.2 msec 0.39 msec	1.0 1.0 1.0 1.0	1.0 1.0 1.0 1.0	1000 1000 1000 1000	3.2 yr 2.8 hr 1.0 sec 0.10 msec	3.25×10^{41} 3.25×10^{46} 3.25×10^{51} 3.25×10^{56}	2.7×10^{-3} 2.7×10^2 2.7×10^7 2.7×10^{12}

^aBased on tables by Braginsky (1965) and by Ruffini and Wheeler (1971b).^bThe spiral time, τ_0 , as given by equation (36.17b) is the time for the two stars to spiral into each other if no nongravitational forces intervene.

These related conditions, $v \ll 1$ and $R \ll \lambda$, are satisfied by all presently conceived laboratory generators of gravitational waves. No one has seen how to bring a macroscopic mass up to a speed $v \sim 1$. These conditions are also satisfied by every gravitationally bound, nearly Newtonian system. Thus, for such a system of mass M , the condition for gravitational binding, $\frac{1}{2}Mv^2 \leq M^2/R$ guarantees that $v \lesssim (M/R)^{1/2} \ll 1$.

The conditions $M/R \ll R/\lambda$ and $|T^{jk}|/T^{00} \ll R/\lambda$ are satisfied by all nearly Newtonian sources of conceivable interest. Typical sources (e.g. binary stars) have

$$\frac{M}{R} \sim \frac{|T^{jk}|}{T^{00}} \sim \left(\frac{R}{\lambda}\right)^2 \ll \frac{R}{\lambda}$$

(virial theorem). In those rare cases where $(M/R \text{ or } |T^{jk}|/T^{00}) \gtrsim R/\lambda$ (e.g., a marginally stable, slowly vibrating star), the motion is so very slow that the radiation will be too weak to be interesting.

For any nearly Newtonian slow-motion system, there is a spacetime region deep inside the near zone ($r \ll \lambda$), but outside the boundary of the source ($r > R$), in which vacuum Newtonian gravitation theory is nearly valid. An observer in this Newtonian region can measure the Newtonian potential Φ and can expand it in powers of $1/r$:

$$\Phi = - \left(\frac{M}{r} + \frac{d_j n^j}{r^2} + \frac{3f_{jk} n^j n^k}{2r^3} + \dots \right), \quad \text{where } n^j = x^j/r. \quad (36.19a)$$

We can then give names to the coefficients in this expansion:

- $M \equiv$ “total mass-energy” = “active gravitational mass”;
- $d_j \equiv$ “dipole moment” [if he chooses the origin of coordinates carefully, he can make $d_j = 0$];
- $f_{jk} \equiv$ “reduced quadrupole moment” [because the system is nearly Newtonian, f_{jk} is given by expression (36.3)].

As this Newtonian potential reaches out into the radiation zone, the static portions of it ($-M/r - d_j n^j/r^2$) maintain their Newtonian form, unchanged. But the dynamic part ($-\frac{3}{2}f_{jk} n^j n^k/r^3$) ceases to be describable in Newtonian terms. As retardation effects become noticeable (at increasing r values), it gradually changes over into outgoing gravitational waves, which must be described in the full general theory of relativity, or in linearized theory, or in the “shortwave” approximation of §35.13.

If one chooses to use linearized theory in the radiation zone, and if one imposes the transverse-traceless gauge there ($h_{0\mu}^{TT} = 0$, $h_{jj}^{TT} = 0$, $h_{jk,k}^{TT} = 0$), then the gravitational waves take the form [derived later as equation (36.47)]

$$h_{jk}^{TT} = \frac{2}{r} \ddot{f}_{jk}^{TT}(t-r) + \text{corrections of order} \left[\frac{1}{r^2} \dot{f}_{jk}^{TT}(t-r) \right]. \quad (36.20) \quad (1) \text{ the wave field } h_{jk}^{TT}$$

Definitions of mass, dipole moment, and reduced quadrupole moment for a slow-motion source

Properties of gravitational waves in terms of reduced quadrupole moment:

Here $\ddot{\mathcal{F}}_{jk}^{TT}$ is the second time-derivative of the transverse-traceless part of the quadrupole moment (transverse to the radial direction; see §35.4); thus,

$$\begin{aligned}\ddot{\mathcal{F}}_{jk}^{TT} &= P_{ja} \ddot{\mathcal{F}}_{ab} P_{bk} - \frac{1}{2} P_{jk} P_{ab} \ddot{\mathcal{F}}_{ab}, \\ P_{ab} &= (\delta_{ab} - n_a n_b) \quad (\text{projection operator}), \\ n_a &= x^a/r \quad (\text{unit radial vector}).\end{aligned}\tag{36.21}$$

The effective stress-energy tensor for these outgoing waves (§35.7) has the same form as for a swarm of zero-mass particles traveling radially outward with the speed of light; at large distances its components of lowest nonvanishing order are

$$\begin{aligned}(2) \text{ effective stress-energy tensor} \quad T_{00}^{(\text{GW})} &= -T_{0r}^{(\text{GW})} = T_{rr}^{(\text{GW})} = \frac{1}{32\pi} \langle h_{jk,0}^{TT} h_{jk,0}^{TT} \rangle = \frac{1}{8\pi r^2} \langle \ddot{\mathcal{F}}_{jk}^{TT} \ddot{\mathcal{F}}_{jk}^{TT} \rangle \\ &= \frac{1}{8\pi r^2} \left(\ddot{\mathcal{F}}_{jk} \ddot{\mathcal{F}}_{jk} - 2n_i \ddot{\mathcal{F}}_{ij} \ddot{\mathcal{F}}_{jk} n_k + \frac{1}{2} (n_j \ddot{\mathcal{F}}_{jk} n_k)^2 \right),\end{aligned}\tag{36.22}$$

where $\langle \rangle$ denotes an average over several wavelengths. (Recall that one cannot localize the energy more closely than a wavelength!) The total power crossing a sphere of radius r at time t is

$$(3) \text{ total power radiated} \quad L_{\text{GW}}(t, r) = \int T^{(\text{GW})0r} r^2 d\Omega = \frac{1}{5} \langle \ddot{\mathcal{F}}_{jk}(t-r) \ddot{\mathcal{F}}_{jk}(t-r) \rangle.\tag{36.23}$$

(See exercise 36.9.) This is the formula with which this chapter began: equation (36.1).

The wave fronts are not precisely spherical. For example, for a binary star system the wave fronts in the equatorial plane must be spirals. This means that there is a tiny nonradial component of the momentum flux, which decreases in strength as $1/r^3$. Associated with this nonradial momentum is an angular momentum density (angular momentum relative to the system's center, $r = 0$), which drops off as $1/r^2$ [Peters (1964), as corrected by DeWitt (1971), p. 286]:

$$(4) \text{ density of angular momentum} \quad \mathcal{J}^i = \frac{1}{8\pi r^2} \epsilon^{ijk} \langle -6n_j \ddot{\mathcal{F}}_{km} \ddot{\mathcal{F}}_{mp} n_p + 9n_j \ddot{\mathcal{F}}_{km} n_m n_p \ddot{\mathcal{F}}_{pq} n_q \rangle.\tag{36.24}$$

The integral of this quantity over a sphere is the total angular momentum being transported outward per unit time,

$$(5) \text{ total angular momentum radiated} \quad -dJ_j/dt = \int \mathcal{J}_j r^2 d\Omega = \frac{2}{5} \epsilon^{jk\ell} \langle \ddot{\mathcal{F}}_{ka} \ddot{\mathcal{F}}_{a\ell} \rangle.\tag{36.25}$$

(See exercise 36.9.)

§36.8. RADIATION REACTION IN SLOW-MOTION SOURCES*

The conservation laws discussed in Box 19.1 and derived in §20.5 guarantee that the source must lose energy and angular momentum at the same rate as the gravitational waves carry them off. The agent that produces these losses is a tiny component of the spacetime curvature inside the source, which reverses sign if one changes from a (realistic) outgoing-wave boundary condition at infinity to the opposite (unrealistic) ingoing-wave condition. These “radiation-reaction” pieces of the curvature can be described in Newtonian language when the source obeys the nearly Newtonian, slow-motion conditions (36.18).

Outgoing-wave boundary condition gives rise to a Newtonian-type radiation-reaction potential

The dynamical part of the Newtonian potential, in its “standard form”

$$\Phi = -\frac{3}{2} I_{jk}(t)n_j n_k / r^3 + O(1/r^4); \text{ equation (36.18),}$$

has no retardation in it. (Newtonian theory demands action at a distance!) Consequently, there is no way whatsoever for the standard potential to decide, at large radii, whether to join onto outgoing waves or onto ingoing waves. Being undecided, it takes the middle track of joining onto standing waves (half outgoing, plus half ingoing). But this is not what one wants. It turns out (see §36.11) that the join can be made to purely outgoing waves if and only if Φ is augmented by a tiny “radiation-reaction” potential

$$\Phi = \Phi_{\text{standard Newtonian theory}} + \Phi^{(\text{react})}, \quad (36.26a)$$

$$\Phi^{(\text{react})} = \frac{1}{5} \frac{d^5 I_{jk}}{dt^5} x^j x^k. \quad (36.26b)$$

Form and magnitude of the radiation-reaction potential

If, instead, one sets $\Phi = \Phi_{\text{standard}} - \Phi^{(\text{react})}$, the potential will join onto purely ingoing waves.

In order of magnitude, the radiation-reaction potential is

$$\Phi^{(\text{react})} \sim \frac{1}{\lambda^5} (MR^2)r^2 \sim \frac{MR^2}{r^3} \left(\frac{r}{\lambda} \right)^5. \quad (36.27)$$

Consequently, near the source it is tiny compared to the standard Newtonian potential [a factor $(R/\lambda)^5 \sim v^5$ smaller!]. However, at the inner boundary of the radiation zone ($r \sim \lambda$), it is of the same order of magnitude as the dynamic, quadrupole part of the standard potential.

The radiation-reaction part of the Newtonian potential plays the same role as a producer of accelerations that any other part of the Newtonian potential does. Any particle in the Newtonian region experiences a gravitational acceleration given by

$$a_j = -\Phi_{,j} = -\Phi_{\text{standard},j} - \Phi_{,j}^{(\text{react})}. \quad (36.28)$$

Effects of the potential:

- (1) radiation-reaction accelerations

*The ideas and formalism described in this section were devised by Burke (1970), Thorne (1969b), and Chandrasekhar and Esposito (1970). Among the forerunners of these ideas were the papers of Peters (1964), and Peres and Rosen (1964).

Inside the source, this acceleration leads to energy and angular momentum losses given by

$$dE/dt = \int \rho a_j v_j d^3x \quad (36.29a)$$

and

$$dJ_j/dt = \int \epsilon_{jkl} x_k \rho a_l d^3x. \quad (36.29b)$$

- (2) loss of energy and angular momentum

(Here ρ is the density, v_j is the velocity, and a_j as above is the acceleration of the matter in the source.) Standard Newtonian theory conserves the energy and angular momentum. Therefore only the reaction part of the potential can produce losses:

$$\begin{aligned} dE/dt &= - \int \rho \Phi_{,j}^{(\text{react})} v_j d^3x, \\ dJ_j/dt &= - \int \epsilon_{jkl} x_k \rho \Phi_{,l}^{(\text{react})} d^3x. \end{aligned} \quad (36.30)$$

A straightforward calculation (exercise 36.5) using expression (36.26b) for the reaction potential yields, for the time-averaged losses,

$$\begin{aligned} dE/dt &= - \frac{1}{5} \langle \ddot{\vec{r}}_{jk} \ddot{\vec{r}}_{jk} \rangle, \\ dJ_j/dt &= - \frac{2}{5} \epsilon_{jkl} \langle \ddot{\vec{r}}_{ka} \ddot{\vec{r}}_{al} \rangle. \end{aligned} \quad (36.31)$$

Notice that these results agree with the energy and angular momentum carried off by the radiation as given by equations (36.1) and (36.25). The agreement is an absolute imperative. The laws of conservation of total energy and angular momentum demand it.

A slow-motion electromagnetic system emitting electric dipole radiation has a radiation-reaction potential

Radiation-reaction potential
for electromagnetic waves

$$A_j^{(\text{react})} = 0, \quad A_0^{(\text{react})} = -\Phi^{(\text{react})} = \frac{2}{3} \ddot{d}_j x^j, \quad (36.32)$$

which is completely analogous to $\Phi^{(\text{react})}$ of gravitation theory [see, e.g., Burke (1971)]. However, attention does not usually focus on this potential and the reaction forces it produces. Instead, it focuses on the reaction force in a special case: that of an isolated charge being accelerated by nonelectromagnetic forces. For such a charge, the reaction force is

$$\mathbf{F}^{(\text{React})} = \frac{2}{3} e^2 \ddot{\vec{x}}. \quad (36.33)$$

No such formula is relevant to gravitation theory, because there is no such thing as a gravitationally isolated, radiating particle (i.e., one accelerated by forces that have no coupling to gravity).

Exercise 36.5. ENERGY AND ANGULAR MOMENTUM LOSSES DUE TO RADIATION REACTION

Derive equations (36.31) for the rate at which gravitational radiation damping saps energy and angular momentum from a slow-motion source. Base the derivation on equations (36.26b) and (36.30).

Exercise 36.6. GRAVITATIONAL WAVES FROM BINARY STAR SYSTEMS

Apply the full formalism of §§36.7 and 36.8 to a binary star system with circular orbits. Calculate the angular distribution of the gravitational waves; the total power radiated; the total angular momentum radiated; the radiation-reaction forces; and the loss of energy and angular momentum due to radiation reaction. Compare the answers with the results quoted in §36.6. [For further details of the solution, see Peters and Mathews (1963).]

§36.9. FOUNDATIONS FOR DERIVATION OF RADIATION FORMULAS

Turn now from the formulas for radiation from a nearly Newtonian system in slow motion to a derivation of these formulas. Initially (this section) work in the full general theory of relativity without any approximations—not even that of slow motion. Impose only the constraint that the source be isolated, and that spacetime become asymptotically flat far away from it.

Use a coordinate system that becomes asymptotically Lorentz as rapidly as spacetime curvature permits, when one moves radially outward from the source toward infinity. Everywhere in this coordinate system, even inside the source, which may be relativistic, define

$$h_{\mu\nu} \equiv g_{\mu\nu} - \eta_{\mu\nu}. \quad (36.34)$$

The $h_{\mu\nu}$ are clearly not the components of a tensor. Neither is $\eta_{\mu\nu}$ the true metric tensor. Nevertheless, one is free to raise and lower indices on $h_{\mu\nu}$ with $\eta_{\mu\nu}$ and to define

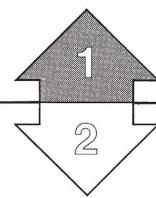
$$\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h; \quad h = h_\alpha^\alpha = h_{\alpha\beta} \eta^{\alpha\beta}. \quad (36.35) \quad (1) \text{ definition of } \bar{h}_{\mu\nu}$$

Moreover, one can always specialize the coordinates so that the four conditions

$$\bar{h}_{\mu,\alpha}^\alpha = 0 \quad (36.36)$$

are exactly satisfied everywhere, including the interior of the source.

With these definitions and conventions, $\bar{h}_{\mu\nu}$ becomes the gravitational field of linearized theory far from the source, and also inside the source if gravity is weak there. But if the interior gravity is strong ($|\bar{h}_{\mu\nu}|$ not $\ll 1$), $\bar{h}_{\mu\nu}$ in the interior has no connection whatsoever to linearized theory.

EXERCISES


The rest of this chapter is Track 2. Chapter 20 (conservation laws) is needed as preparation for it. It will be helpful in Chapter 39 (post-Newtonian formalism), but is not needed as preparation for any other chapters.

Derivation of formula for the gravitational-wave field produced by a slow-motion source:

(2) field equations in terms of $\bar{h}_{\mu\nu}$

(3) philosophy of controlled ignorance

(4) integral formulation of field equations

The exact Einstein field equations can be written in terms of $\bar{h}^{\mu\nu}$ as [cf. §20.3; in particular, combine equations (20.14), (20.18), and (20.3); and impose the coordinate condition (36.36)]

$$\bar{h}^{\mu\nu}_{,\alpha\beta}\eta^{\alpha\beta} = -16\pi(T^{\mu\nu} + t^{\mu\nu}), \quad (36.37)$$

where $T^{\mu\nu}$ are the components of stress-energy tensor, and $t^{\mu\nu}$ are quantities (components of the “stress-energy pseudotensor for the gravitational field”) that are of quadratic order and higher in $\bar{h}^{\mu\nu}$. Recall the “philosophy of controlled ignorance” expounded in §19.3. One is so ignorant that nowhere does one ever write down an explicit expression for $t^{\mu\nu}$ in terms of $\bar{h}^{\alpha\beta}$; and this ignorance is so controlled that one will never need such an expression in the calculations to follow! More specifically, the strength of the outgoing wave is proportional to the integral of a complicated expression over the interior of a system where “gravitational stresses” may be comparable to material stresses, $|t^{jk}| \sim |T^{jk}|$. No matter. All that will count for the radiation is the quadrupole part of the field. Moreover, that quadrupole moment is empirically definable by purely Newtonian measurements in the Newtonian region (1) well inside the wave zone, but (2) well outside the surface of the source. One does not have to know the inner workings of a star to define its mass (influence on Kepler orbits outside) nor does one have to know those inner workings to define its quadrupole moment *as sensed externally*.

Einstein’s equations (36.37), augmented by an outgoing-wave boundary condition, are equivalent to the integral equations

$$\bar{h}^{\mu\nu}(t, x^j) = 4 \int_{\text{all space}} \frac{[T^{\mu\nu} + t^{\mu\nu}]_{\text{ret}}}{|\mathbf{x} - \mathbf{x}'|} d^3x', \quad (36.38)$$

where

$$|\mathbf{x} - \mathbf{x}'| \equiv \left[\sum_j (x^j - x'^j)^2 \right]^{1/2}, \quad d^3x' \equiv dx'^1 dx'^2 dx'^3,$$

and the subscript “ret” means the quantity is to be evaluated at the retarded spacetime point

$$(t' = t - |\mathbf{x} - \mathbf{x}'|, x'^j).$$

These are integral equations because the unknowns, $\bar{h}^{\mu\nu}$, appear both outside and inside the integral (inside they are contained in $t^{\mu\nu}$). Notice that in passing from the wave equations (36.37) to the integral equations (36.38), one has cavalierly behaved as though $\bar{h}^{\mu\nu}$ were fields in flat spacetime. This is certainly not true; but the mathematical manipulations are valid nevertheless!—and the integral equations (36.38) are valid for any field point (t, x^j) , even inside the source.

§36.10. EVALUATION OF THE RADIATION FIELD IN THE SLOW-MOTION APPROXIMATION

(5) specialization to slow motion

Thus far the analysis has been exact. Now it is necessary to introduce the slow-motion assumption of §36.7: $R \ll \lambda$.

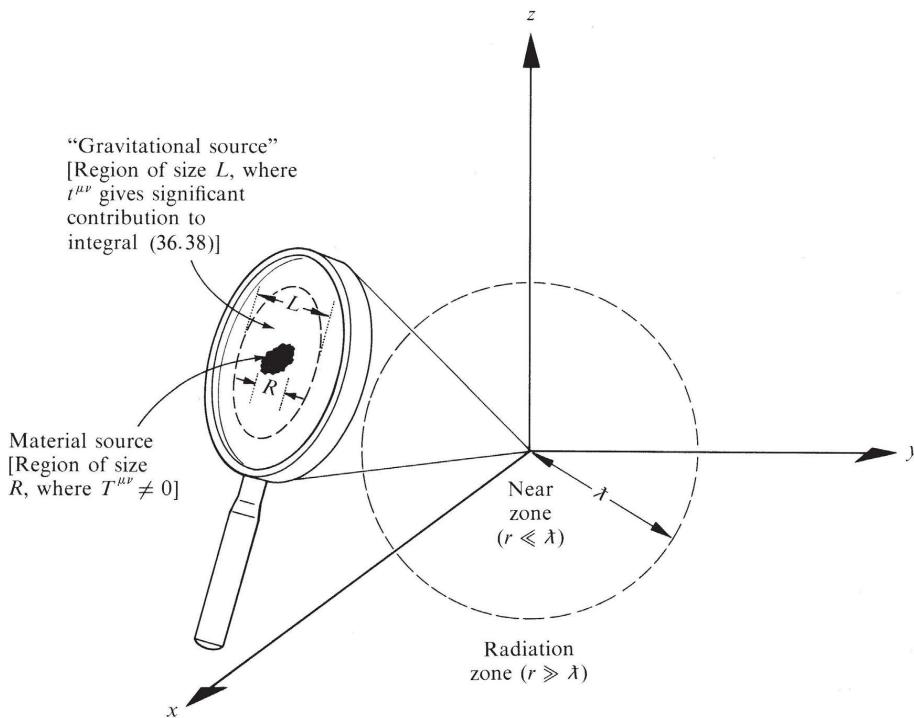


Figure 36.3.

A slow-motion source radiating gravitational waves. The origin of spatial coordinates is located inside the source. The size of the source, R , is very small compared to a reduced wavelength, $R \ll \lambda$. Significant contributions to the retarded integral (36.38) for $\bar{h}^{\mu\nu}$ come only from a region of size $L \sim R \ll \lambda$ surrounding the source, because outside the source—but in the near zone ($R \ll r \ll \lambda$)—the “stress-energy pseudotensor” $t^{\mu\nu}$ dies out as $1/r^4$ (see exercise 36.7).

In the radiation zone, $t^{\mu\nu}$ ceases to die out as $1/r^4$, and begins to die out as $1/r^2$; it is trying to describe (but cannot, really, without appropriate averaging) the stress-energy carried by the gravitational waves. If the source has been emitting waves long enough, contributions from the radiation zone to the retarded integral (36.38) may be nonnegligible:

$$[t^{\mu\nu}]_{\text{ret}} \sim \frac{1}{r'^2} \implies \int [t^{\mu\nu}]_{\text{ret}} d^3x' \sim \underbrace{\int \frac{1}{r'^2} r'^2 d\Omega' dr'}_{[\text{for } r > \lambda] \quad [\text{may have significant contributions from large } r']}$$

Such contributions are ignored in the text, in calculations of the radiated waves, because they have nothing whatsoever to do with the emission process itself. Rather, they are part of the propagation process treated in the last chapter. They include the background curvature produced by the stress-energy of the waves, scattering of waves off the background curvature, wave-wave scattering, etc.; and they are totally negligible in the neighborhood of the source itself ($r \lesssim 1,000 \lambda$, for example) because a slow-motion source radiates so very weakly.

Place the origin of spatial coordinates inside the source, as shown in Figure 36.3. For slow-motion systems, the only significant contributions to the retarded integrals (36.38) come from deep inside the near zone (from a region of size $L \sim R \ll \lambda$; see Figure 36.3). Confine attention to “field points” (points of observation) x^j far outside this “source region,”

$$|\mathbf{x}| \equiv r \gg L \gtrsim |\mathbf{x}'|, \quad (36.39a)$$

and expand the retarded integral (36.38) in powers of \mathbf{x}'/r —in just the same manner as was done in §19.1. (Such an expansion is justified by and requires the slow-motion assumption, $\lambda/R \sim \lambda/L \ll 1$.) The result is

$$\begin{aligned} \bar{h}^{\mu\nu}(t, \mathbf{x}) &= \frac{4}{r} \int [T^{\mu\nu}(\mathbf{x}', t-r) + t^{\mu\nu}(\mathbf{x}', t-r)] d^3x' \\ &\quad + O\left\{\frac{x^j}{r^2\lambda} \int x^j [T^{\mu\nu}(\mathbf{x}', t-r) + t^{\mu\nu}(\mathbf{x}', t-r)] d^3x'\right\}. \end{aligned} \quad (36.40)$$

(6) calculation of \bar{h}^{jk} in radiation zone

Of the ten components of $\bar{h}^{\mu\nu}$, only the six spatial ones, \bar{h}^{jk} , are of interest, since only they are needed in projecting out the transverse-traceless radiation field \bar{h}_{jk}^{TT} . The spatial components are expressed by equations (36.40) in terms of integrals over the “stress distribution” $T^{jk} + t^{jk}$. It will be convenient, in making comparisons with Newtonian theory, to reexpress \bar{h}^{jk} in terms of integrals over the “energy distribution” $T^{00} + t^{00}$. One can make the conversion with the help of the exact equations of motion $T^{\mu\nu}_{;\nu} = 0$, which have the special form

$$(T^{\mu\nu} + t^{\mu\nu})_{,\nu} = 0 \quad (36.41)$$

in the coordinate system being used [see equations (36.36) and (36.37); also the discussion in §20.3]. Applying these relations twice in succession, one obtains the identity

$$\begin{aligned} (T^{00} + t^{00})_{,00} &= -(T^{0\ell} + t^{0\ell})_{,\ell 0} = -(T^{\ell 0} + t^{\ell 0})_{,0\ell} \\ &= +(T^{\ell m} + t^{\ell m})_{,m\ell}. \end{aligned}$$

From this and the elementary chain rule for differentiation, it follows that

$$\begin{aligned} [(T^{00} + t^{00})x^j x^k]_{,00} &= (T^{\ell m} + t^{\ell m})_{,m\ell} x^j x^k \\ &= [(T^{\ell m} + t^{\ell m})x^j x^k]_{,m\ell} - 2[(T^{\ell j} + t^{\ell j})x^k + (T^{\ell k} + t^{\ell k})x^j]_{,\ell} \\ &\quad + 2(T^{jk} + t^{jk}), \end{aligned}$$

whence

$$\int (T^{jk} + t^{jk}) d^3x = \frac{1}{2} (d^2 I_{jk}/dt^2), \quad (36.42a)$$

where

$$I_{jk}(t) \equiv \int [T^{00}(t, \mathbf{x}) + t^{00}(t, \mathbf{x})] x^j x^k d^3x. \quad (36.42b)$$

(7) specialization to nearly Newtonian sources

Now introduce the nearly Newtonian assumption. It guarantees that gravitation contributes only a small fraction of the total energy:

$$t^{00} \sim (\Phi,)^2 \sim M^2/R^4 \sim (M/R)T^{00} \ll T^{00},$$

hence

$$I_{jk}(t) = \int T^{00}(t, \mathbf{x}) x^j x^k d^3x. \quad (36.42b')$$

The quantity I_{jk} thus represents the second moment of the mass distribution.

By combining equations (36.42) and (36.40), and by noting that inside the source $|t^{jk}| \sim |\Phi_{,j}\Phi_{,k}| \sim T^{00}|\Phi|$, one obtains

$$\begin{aligned}\bar{h}^{jk}(t, x) &= \frac{2}{r} \frac{d^2 I_{jk}(t-r)}{dt^2} + O\left[\frac{1}{r} \left(\frac{|T^{jk}|}{T^{00}} + |\Phi|\right) \frac{R}{\lambda} M\right] \\ &= \frac{2}{r} \frac{d^2 I_{jk}(t-r)}{dt^2} \underbrace{\left\{1 + O\left[\frac{|T^{jk}|}{T^{00}} + \frac{M}{R}\right]\right\}}_{R}.\end{aligned}\quad (36.43)$$

[negligible by assumptions (36.18)] \uparrow

Actually, what one wants are h_{jk}^{TT} , not \bar{h}^{jk} . They can be obtained by first lowering indices, using $\eta_{lm} = \delta_{lm}$, and then projecting out the TT part using the projection operator for radially traveling waves:

$$P_{lm} = \delta_{lm} - n_l n_m; \quad n_l = x^l/r \quad (36.44)$$

(see Box 35.1). (Because \bar{h}_{jk} and h_{jk} differ only in the trace, they have the same TT parts). The result is

$$h_{jk}^{TT}(t, x) = \frac{2}{r} \frac{d^2 I_{jk}^{TT}(t-r)}{dt^2}, \quad (36.45a)$$

where

$$I_{jk}^{TT} = P_{jl} I_{lm} P_{mk} - \frac{1}{2} P_{jk} (P_{lm} I_{mi}). \quad (36.45b)$$

This is not the best form in which to write the answer, because an external observer cannot measure directly the second moment of the mass distribution, I_{jk} . Fortunately, one can replace I_{jk} by the reduced quadrupole moment,

$$f_{jk} \equiv I_{jk} - \frac{1}{3} \delta_{jk} I = \int (T^{00} + t^{00}) \left(x^j x^k - \frac{1}{3} \delta_{jk} r^2 \right) d^3 x, \quad (36.46)$$

(8) conversion, by projection, to h_{jk}^{TT}

(9) reexpression of h_{jk}^{TT} in terms of reduced quadrupole moment

and write

$$h_{jk}^{TT}(t, x) = \frac{2}{r} \frac{d^2 f_{jk}^{TT}(t-r)}{dt^2}. \quad (36.47)$$

This is allowed because the TT parts of I_{jk} and f_{jk} are identical (exercise 36.8).

The reduced quadrupole moment f_{jk} has a well-defined, elementary physical significance for an observer confined to the exterior of the source. In the near zone ($r \ll \lambda$), but outside the source so that vacuum Newtonian theory is very nearly valid, the Newtonian potential is

$$\begin{aligned}\Phi &= -\frac{1}{2} h_{00} = -\frac{1}{2} h^{00} = -\frac{1}{2} \left(\bar{h}^{00} + \frac{1}{2} \bar{h} \right) = -\frac{1}{4} (\bar{h}^{00} + \bar{h}^{jj}) \\ &= - \int_{\text{all space}} \frac{[T^{00} + t^{00} + T^{jj} + t^{jj}]_{\text{ret}}}{|x - x'|} d^3 x'\end{aligned}$$

[see equation (36.38)]. Any nearly Newtonian, slow-motion source satisfies

$$|t^{00} + T^{jj} + t^{jj}| \ll T^{00}$$

[recall: $t^{\alpha\beta} \sim (\Phi_{,j})^2 \sim T^{00}|\Phi|$. Hence, one can write

$$\Phi(\mathbf{x}, t) = - \int \frac{[T^{00}(\mathbf{x}', t)]}{|\mathbf{x} - \mathbf{x}'|} d^3x'. \quad (36.48)$$

Expanding $|\mathbf{x} - \mathbf{x}'|^{-1}$ in powers of $1/r$, one obtains

$$\Phi = - \left(\frac{M}{r} + \frac{d_j x^j}{r^3} + \frac{3I_{jk} x^j x^k}{2r^5} + \dots \right) \text{ for } \begin{cases} r \ll \lambda, \text{ but } r \text{ nevertheless} \\ \text{less large enough that} \\ \text{vacuum Newtonian} \\ \text{theory is valid} \end{cases}, \quad (36.49a)$$

where

$$M = (\text{total mass-energy of source}) = \int T^{00} d^3x,$$

$$d_j \equiv (\text{dipole moment of source}) = \int T^{00} x^j dx^3, \quad (36.49b)$$

$$I_{jk} \equiv (\text{reduced quadrupole moment of source}) = \text{expression (36.46).}$$

Thus, the quantities I_{jk} , whose second time-derivatives determine the radiation field by equation (36.47), are precisely the components of the star's reduced quadrupole moment, as measured by an observer who explores its Newtonian potential Φ deep inside the near zone ($r \ll \lambda$) ("empirical quadrupole moment.").

The final answer (36.47) for the radiation field in terms of I_{jk}^{TT} was quoted in the summary of results given in §36.7. Also quoted there were expressions for the effective stress-energy tensor of the radiation and for the energy and angular momentum radiated [equations (36.22) to (36.25)]. Those expressions can be derived using the formalism of the shortwave approximation. (See exercise 36.9.)

EXERCISES

Exercise 36.7. MAGNITUDE OF $t^{\mu\nu}$

Consider a slow-motion source of gravitational waves. Show that far from the source, but in the near zone ($R \ll r \ll \lambda$) the components of the "stress-energy pseudotensor" $t^{\mu\nu}$ die out as $1/r^4$, but in the radiation zone ($r \gg \lambda$) they die out only as $1/r^2$.

Exercise 36.8. PROOF THAT THE TRANVERSE TRACELESS PARTS OF I_{jk} AND I_{jk} ARE IDENTICAL

Prove by direct computation that the TT parts of I_{jk} (36.42b) and I_{jk} (36.46) are identical, no matter where the observer is who does the TT projection (i.e., no matter what the unit vector n in the projection operator may be).

Exercise 36.9. ENERGY AND ANGULAR MOMENTUM RADIATED

(a) For the gravitational waves in asymptotically flat spacetime described by equation (36.47), calculate the smeared-out stress-energy tensor $T_{\mu\nu}^{(GW)}$ of equation (35.23). [Answer: equation (36.22).]

(b) Perform the integrals of equations (36.23) and (36.25) to obtain the total power and angular momentum radiated. [Hint: Derive and use the following averages over a sphere

$$\frac{1}{4\pi} \int n_i d\Omega = 0, \quad \frac{1}{4\pi} \int n_i n_j d\Omega = \frac{1}{3} \delta_{ij}, \quad \frac{1}{4\pi} \int n_i n_j n_k d\Omega = 0,$$

$$\frac{1}{4\pi} \int n_i n_j n_k n_l d\Omega = \frac{1}{15} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}).$$

Here $\mathbf{n} \equiv \mathbf{x}/|\mathbf{x}|$ is the unit radial vector.]

§36.11. DERIVATION OF THE RADIATION-REACTION POTENTIAL

Turn, finally, to a derivation of the radiation-reaction results quoted in §36.8. The analysis starts with the solution (36.43) for the spatial part of the radiation field in the original (i.e., not TT) gauge:

Derivation of formula for the radiation-reaction potential:

$$\bar{h}^{jk}(t, \mathbf{x}) = \frac{2}{r} \ddot{I}_{jk}(t - r). \quad (36.50)$$

Although this solution was originally derived by discarding all terms that die out faster than $1/r$, it is in fact an exact solution to the vacuum field equations $\bar{h}^{jk}_{,\alpha} = 0$ of linearized theory. This means that it is valid in the intermediate and near zones ($r \lesssim \lambda$, but $r > R$) as well as in the radiation zone.

Were one to replace the outgoing-wave condition by an ingoing-wave condition at infinity, the exact solution (36.50) for \bar{h}^{jk} would get replaced by

$$\bar{h}^{jk}(t, \mathbf{x}) = \frac{2}{r} \ddot{I}_{jk}(t + r).$$

Thus, in order to delineate the effects of the outgoing-wave boundary condition, one can write the exact solution in the form

$$\bar{h}_{jk}(t, \mathbf{x}) = \bar{h}^{jk}(t, \mathbf{x}) = \frac{2}{r} \ddot{I}_{jk}(t - \epsilon r), \quad \epsilon = \pm 1, \quad (36.51)$$

(1) formula for \bar{h}_{jk} anywhere outside source, with either outgoing or ingoing waves

and then focus attention on the effects of the sign of ϵ .

In the near zone ($r \ll \lambda$), but outside the nearly Newtonian source, this solution for \bar{h}_{jk} , as expanded in powers of r , becomes

$$\bar{h}_{jk} = 2 \left[\frac{I_{jk}^{(2)}}{r} - \epsilon I_{jk}^{(3)} + \frac{I_{jk}^{(4)}r}{2!} - \epsilon \frac{I_{jk}^{(5)}r^2}{3!} + \dots \right], \quad (36.52a)$$

(2) \bar{h}_{jk} specialized to near zone

where

$$I_{jk}^{(n)} \equiv d^n I_{jk}(t)/dt^n.$$

The corresponding forms of \bar{h}_{0j} and \bar{h}_{00} can be generated from this by the gauge conditions $\bar{h}_{\alpha\beta,\beta} = 0$; i.e., by $\bar{h}_{j0,0} = \bar{h}_{jk,k}$ and $\bar{h}_{00,0} = \bar{h}_{0j,j}$. The results are:

- (3) \bar{h}_{00} and \bar{h}_{0j} in near zone calculated by gauge conditions

$$\bar{h}_{0j} = 2 \left[-\frac{I_{jk}^{(1)}x^k}{r^3} + \frac{I_{jk}^{(3)}x^k}{2!r} - \epsilon \frac{2I_{jk}^{(4)}x^k}{3!} + \frac{3I_{jk}^{(5)}x^kr}{4!} - \epsilon \frac{4I_{jk}^{(6)}x^kr^2}{5!} \right] + (\text{static terms not associated with radiation}); \quad (36.52b)$$

$$\begin{aligned} \bar{h}_{00} = 2 & \left[\frac{(3x^jx^k - r^2\delta^{jk})}{r^5} I_{jk} - \frac{(x^jx^k - r^2\delta^{jk})}{2!r^3} I_{jk}^{(2)} - \epsilon \frac{2}{3!} I_{jj}^{(3)} \right. \\ & \left. + \frac{3(x^jx^k + r^2\delta^{jk})}{4!r} I_{jk}^{(4)} - \epsilon \frac{4(2x^jx^k + r^2\delta^{jk})}{5!} I_{jk}^{(5)} + \dots \right] \\ & + (\text{static and time-linear terms not associated with radiation}). \end{aligned} \quad (36.52c)$$

The leading term in these expressions rises as $1/r^3$ when one approaches the source:

$$\bar{h}_{00} \approx \frac{2(3x^jx^k - r^2\delta^{jk})}{r^5} I_{jk} = \frac{6I_{jk}n^jn^k}{r^3}.$$

It is precisely the leading term in the dynamic, quadrupole part of the Newtonian potential, $\Phi = -\frac{1}{2}h_{00} = -\frac{1}{4}\bar{h}_{00}$. All other terms without ϵ 's in front of them are corrections to the Newtonian potential. They produce effects like the perihelion shift of Mercury that in no way deplete the energy and angular momentum of the system.

- (4) plucking out the radiation-reaction potentials from $\bar{h}_{\alpha\beta}$

The terms with ϵ 's are associated with radiation reaction. Pluck the leading ones out and call them “reaction potentials”:

$$\begin{aligned} \bar{h}_{jk}^{(\text{react})} &= -2I_{jk}^{(3)} - \frac{1}{3}I_{jk}^{(5)}r^2, \\ \bar{h}_{0j}^{(\text{react})} &= -\frac{2}{3}I_{jk}^{(4)}x^k - \frac{1}{15}I_{jk}^{(6)}x^kr^2, \\ \bar{h}_{00}^{(\text{react})} &= -\frac{2}{3}I_{jj}^{(3)} - \frac{1}{15}(2x^jx^k + r^2\delta^{jk})I_{jk}^{(5)}. \end{aligned} \quad (36.53)$$

The corresponding metric perturbations $h_{\alpha\beta} = \bar{h}_{\alpha\beta} - \frac{1}{2}\bar{h}\eta_{\alpha\beta}$ are

$$\begin{aligned} h_{jk}^{(\text{react})} &= -2I_{jk}^{(3)} + \frac{2}{3}I_{ll}^{(3)}\delta_{jk} + 0(I_{jk}^{(5)}r^2), \\ h_{0j}^{(\text{react})} &= -\frac{2}{3}I_{jk}^{(4)}x^k + 0(I_{jk}^{(6)}r^3), \\ h_{00}^{(\text{react})} &= -\frac{4}{3}I_{ll}^{(3)} - \frac{1}{15}(x^jx^k + 3r^2\delta_{jk})I_{jk}^{(5)}. \end{aligned} \quad (36.54)$$

These reaction potentials in the near zone are understood most clearly by a change of gauge that brings them into Newtonian form. Set

$$x^\mu_{\text{new}} = x^\mu_{\text{old}} + \xi^\mu(x), \quad h_{\mu\nu_{\text{new}}} = h_{\mu\nu_{\text{old}}} - \xi_{\mu,\nu} - \xi_{\nu,\mu}$$

(5) conversion of
radiation-reaction
potentials to Newtonian
gauge

with

$$\begin{aligned} \xi_j &= -I_{jk}^{(3)}x^k + \frac{1}{3}I_{ll}^{(3)}x^j, \\ \xi_0 &= -\frac{2}{3}I_{ll}^{(2)} + \frac{1}{6}I_{jk}^{(4)}x^jx^k - \frac{1}{6}I_{ll}^{(4)}r^2. \end{aligned} \quad (36.55)$$

Then in the new gauge

$$\begin{aligned} h_{jk}^{(\text{react})} &= O(I_{jk}^{(5)}r^2), \quad h_{0j}^{(\text{react})} = O(I_{jk}^{(6)}r^3), \\ h_{00}^{(\text{react})} &= -\frac{2}{5}I_{jk}^{(5)}x^jx^k. \end{aligned} \quad (36.56)$$

This gauge is ideally suited to a Newtonian interpretation, since in it the geodesic equation for slowly moving particles has the form

$$d^2x^j/dt^2 = -\Phi_{,j}^{(\text{react})} + \left(\begin{array}{l} \text{terms not sensitive to} \\ \text{outgoing-wave condition} \end{array} \right), \quad (36.57)$$

with

$$\Phi^{(\text{react})} = -\frac{1}{2}h_{00}^{(\text{react})} = \frac{1}{5}I_{jk}^{(5)}x^jx^k. \quad (36.58)$$

Thus, the leading radiation-reaction effects (with fractional errors $\sim [\lambda/r]^2$) can be described in the near zone of a nearly Newtonian source by appending the term $\frac{1}{5}I_{jk}^{(5)}x^jx^k$ to the Newtonian potential. The resulting formalism and a qualitative version of the above derivation were presented in §36.8.

CHAPTER 37

DETECTION OF GRAVITATIONAL WAVES

I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be.

WILLIAM THOMSON, LORD KELVIN [(1889), p. 73]

§37.1. COORDINATE SYSTEMS AND IMPINGING WAVES

The detector is even easier to analyze than the generator or the transmission when one deals with gravitational waves within the framework of general relativity. Man's potential detectors all lie in the solar system, where gravity is so weak and spacetime so nearly flat that a plane gravitational wave coming in remains for all practical purposes a plane gravitational wave. (Angle of deflection of wave front passing limb of sun is only 1." 75.) Moreover, the nearest source of significant waves is so far away that, for all practical purposes, one can consider the waves as plane-fronted when they reach the Earth. Consequently, as they propagate in the z -direction past a detector, they can be described to high accuracy by the following transverse-traceless linearized expressions

Linearized description of gravitational waves propagating past Earth

Metric perturbation: $h_{xx}^{TT} = -h_{yy}^{TT} = A_+(t - z)$, $h_{xy}^{TT} = h_{yx}^{TT} = A_\times(t - z)$, (37.1a)

$$\text{Riemann tensor: } R_{x0x0} = -R_{y0y0} = -\frac{1}{2}\ddot{A}_+(t - z), \\ R_{x0y0} = R_{y0x0} = -\frac{1}{2}\ddot{A}_\times(t - z). \quad (37.1b)$$

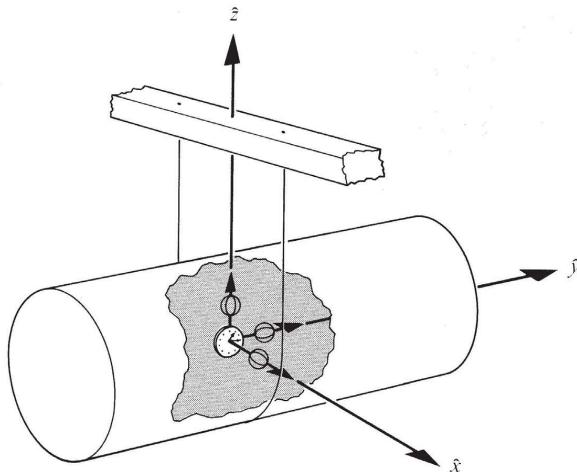


Figure 37.1.

The proper reference frame of a vibrating-bar detector. The bar hangs by a wire from a cross beam, which is supported by vertical posts (not shown) that are embedded in the Earth. Consequently, the bar experiences a 4-acceleration given, at the moment when this diagram is drawn, by $\mathbf{a} = g(\partial/\partial\hat{z})$, where g is the “local acceleration of gravity” ($g \sim 980 \text{ cm/sec}^2$). Later, the spatial axes will have rotated relative to the bar (“Foucoul pendulum effect” produced by Earth’s rotation), so the components of \mathbf{a} but not its magnitude will have changed.

The proper reference frame relies on an imaginary clock and three imaginary gyroscopes located at the bar’s center of mass (and shown above in a cut-away view). Coordinate time is equal to proper time as measured by the clock, and the directions of the spatial axes $\partial/\partial x^j$ are attached to the gyroscopes. The forces that prevent the gyroscopes from falling in the Earth’s field must be applied at the centers of mass of the individual gyroscopes (no torque!).

$$\text{Stress-energy: } T_{00}^{(\text{GW})} = T_{zz}^{(\text{GW})} = -T_{0z}^{(\text{GW})} = \frac{1}{16\pi} \langle \dot{A}_+^2 + \dot{A}_x^2 \rangle_{\text{time avg.}} \quad (37.1c)$$

(See exercise 37.1.)

To analyze most easily the response of the detector to these impinging waves, use not the TT coordinate system $\{x^\alpha\}$ (which is specially “tuned” to the waves), but rather use coordinates $\{\hat{x}^\alpha\}$ specially “tuned” to the experimenter and his detector. The detector might be a vibrating bar, or the vibrating Earth, or a loop of tubing filled with fluid (see Figures 37.1 and 37.2). But whatever it is, it will have a center of mass. Attach the spatial origin, $x^j = 0$, to this center of mass; and attach *orthonormal* spatial axes, $\partial/\partial x^j$, to (possibly imaginary) gyroscopes located at this spatial origin (Figure 37.1). If the detector is accelerating (i.e., not falling freely), make the gyroscopes accelerate with it by applying the necessary forces at their centers of mass (no torque!). Use, as time coordinate, the proper time $x^0 = \tau$ measured by a clock at the spatial origin. Finally, extend these locally defined coordinates \hat{x}^α throughout all spacetime in the “straightest” manner possible. (See

Proper reference frame of a detector

Track 2's §13.6 for full details.) The metric in this “proper reference frame of the detector” will have the following form

$$ds^2 = -(1 + 2a_j x^j)(dx^0)^2 + \delta_{jk} dx^j dx^k + O(|x^j|^2) dx^{\hat{\alpha}} dx^{\hat{\beta}}. \quad (37.2)$$

[equation (13.71) with $\omega^l = 0$.] Here a_j are the spatial components of the detector’s 4-acceleration. (Since \mathbf{a} must be orthogonal to the detector’s 4-velocity, a_0 vanishes.) Notice that, except for the acceleration term in g_{00} (“gravitational redshift term”; see §38.5 and exercise 6.6), this reference frame is locally Lorentz.

EXERCISES

Exercise 37.1. GENERAL PLANE WAVE IN TT GAUGE

Show that the most general linearized plane wave can be described in the transverse-traceless gauge of linearized theory by expressions (37.1). [Hint: Express the plane wave as a superposition (Fourier integral) of monochromatic plane waves, and describe each monochromatic plane wave by expressions (35.16). Use equations (35.10) and (35.23) to calculate $R_{\alpha\beta\gamma\delta}$ and $T_{\mu\nu}^{(\text{GW})}$.]

Exercise 37.2. TEST-PARTICLE MOTION IN PROPER REFERENCE FRAME

Show that a slowly moving test particle, falling freely through the proper reference frame of equation (37.2), obeys the equation of motion (geodesic equation)

$$d^2 x^j / d\hat{t}^2 = -a_j + O(|x^k|).$$

Thus, one can interpret $-a_j$ as the “local acceleration of gravity” (see caption of Figure 37.1).

§37.2. ACCELERATIONS IN MECHANICAL DETECTORS

Equations of motion for a mechanical detector

The proper reference frame of equation (37.2) is the closest thing that exists to the reference frame a Newtonian physicist would use in analyzing the detector. In fact, it is so nearly Newtonian that (according to the analysis of Box 37.1) *the equations of motion for a mechanical detector, when written in this proper reference frame, take their standard Newtonian form and can be viewed and dealt with in a fully Newtonian manner, with one exception: the gravitational waves produce a driving force of non-Newtonian origin, given by the familiar expression for geodesic deviation*

$$\begin{aligned} \left(\begin{array}{l} \text{force per unit mass (i.e., acceleration)} \\ \text{of a particle at } x^j \text{ relative to detector's} \\ \text{center of mass at } x^j = 0 \end{array} \right) &= \left(\frac{d^2 x^j}{d\hat{t}^2} \right)_{\text{due to waves}} \\ &= -(R_{j\hat{o}\hat{k}\hat{o}})_{\text{due to waves}} x^k. \end{aligned} \quad (37.3)$$

To use this equation, and to calculate detector cross sections later, one must know the components of the curvature tensor $R_{\hat{\alpha}\hat{\beta}\hat{\gamma}\hat{\delta}}$, and of the waves’ stress-energy tensor, $T_{\hat{\mu}\hat{\nu}}^{(\text{GW})}$, in the detector’s proper reference frame. One cannot calculate $R_{\hat{\alpha}\hat{\beta}\hat{\gamma}\hat{\delta}}$ directly

Box 37.1 DERIVATION OF EQUATIONS OF MOTION FOR A MECHANICAL DETECTOR

Consider a “mass element” in a mechanical detector (e.g., a cube of aluminum one millimeter on each edge if the detector is the bar of Figure 37.1; or an element of fluid with volume 1 mm^3 if the detector is the tube filled with fluid shown in part h of Figure 37.2). This mass element gets pushed and pulled by adjacent matter and electromagnetic fields, as the medium of the detector vibrates or flows or does whatever it is supposed to do. Let

$$\mathbf{f} \equiv \left(\begin{array}{l} \text{4-force per unit mass exerted on mass-element} \\ \text{by adjacent matter and by electromagnetic fields} \end{array} \right). \quad (1)$$

This 4-force per unit mass gives the mass element a 4-acceleration $\nabla_u \mathbf{u} = \mathbf{f}$; or, in terms of components in the detector’s proper reference frame, $f^j = Du^j/d\tau$. Assume that the mass element has a very small velocity ($v \ll 1$) in the detector’s proper reference frame (i.e., relative to the detector’s center of mass). Then, ignoring terms of $O(v^2)$, $O(|x^j|^2)$, and $O(|x^j|v)$, one has [see equation (37.2)]

$$d\hat{t}/d\tau = u^0 = 1 - a_j x^j \equiv 1 - \mathbf{a} \cdot \mathbf{x}, \quad (2)$$

and

$$f^j = d^2 x^j / d\tau^2 + \Gamma_{\hat{\alpha}\hat{\beta}}^j u^{\hat{\alpha}} u^{\hat{\beta}} = (d^2 x^j / d\hat{t}^2 + \Gamma_{\hat{0}\hat{0}}^j)(1 - 2\mathbf{a} \cdot \mathbf{x}). \quad (3)$$

Exercise 37.3 calculates $\Gamma_{\hat{0}\hat{0}}^j$ to precision of $O(|x^j|)$. Inserting its result and rearranging terms, one finds that

$$d^2 x^j / d\hat{t}^2 = (1 + 2\mathbf{a} \cdot \mathbf{x}) f^j - a^j (1 + \mathbf{a} \cdot \mathbf{x}) - R_{\hat{0}\hat{k}\hat{0}}^j x^k \quad (4)$$

(“equation of motion for mass element”).

Examine this equation, first from the viewpoint of an Einsteinian physicist, and then from the viewpoint of a Newtonian physicist.

The Einsteinian physicist recognizes $d^2 x^j / d\hat{t}^2$ as the “coordinate acceleration” of the mass element—but he keeps in mind that, to precision of $O(|x^j|^2)$, coordinate lengths and proper lengths are the same [see equation (37.2)]. The coordinate acceleration $d^2 x^j / d\hat{t}^2$ has three causes: (1) *the externally applied force*,

$$\begin{aligned} (1 + 2\mathbf{a} \cdot \mathbf{x}) f^j &= (d^2 x^j / d\hat{t}^2)_{\text{external force}} \\ &= (1 + 2\mathbf{a} \cdot \mathbf{x}) (d^2 x^j / d\tau^2)_{\text{external force}} \end{aligned} \quad (5a)$$

(the origin of the $\mathbf{a} \cdot \mathbf{x}$ correction is simply the conversion between coordinate time

Box 37.1 (continued)

and proper time); (2) the “inertial force” due to the acceleration of the reference frame,

$$-a^j(1 + \mathbf{a} \cdot \mathbf{x}) = (d^2x^j/d\hat{t}^2)_{\text{inertial force}} \quad (5b)$$

(see exercise 37.4 for explanation of the $\mathbf{a} \cdot \mathbf{x}$ correction); and (3) a “Riemann curvature force,” which will include Riemann curvature due to local, Newtonian gravitational fields (fields of Earth, sun, moon, etc.), plus Riemann curvature due to the impinging gravitational waves,

$$-(R^j_{0k0})_{\text{waves}}x^k - (R^j_{0k0})_{\text{Newton fields}}x^k = (d^2x^j/d\hat{t}^2)_{\text{curvature}} \quad (5c)$$

(linear superposition because all gravitational fields in the solar system are so weak). This “Riemann curvature force” is not, of course, “felt” by the mass element; it does not produce any 4-acceleration. Rather, like the inertial force, it originates in the choice of reference frame: The spatial coordinates x^j measure proper distance and direction away from the detector’s center of mass; and Riemann curvature tries to change this proper distance and direction (“relative acceleration;” “geodesic deviation”).

A Newtonian physicist views the equation of motion (4) in a rather different manner. Having been told that the spatial coordinates x^j measure proper distance and direction away from the detector’s center of mass, he thinks of them as the standard Euclidean spatial coordinates of Newtonian theory. He then rewrites equation (4) in the form

$$d^2x^j/d\hat{t}^2 = F^j - (R^j_{0k0})_{\text{waves}}x^k, \quad (6)$$

where

$$F^j \equiv \left(\begin{array}{l} \text{Newtonian force per unit mass} \\ \text{acting on mass element} \end{array} \right) \quad (7)$$

$$= (1 + 2\mathbf{a} \cdot \mathbf{x})f^j - a^j(1 + \mathbf{a} \cdot \mathbf{x}) - (R^j_{0k0})_{\text{Newton fields}}x^k.$$

The Newtonian physicist is free to express F^j in a form more familiar than this. He can ignore the subtleties of the $\mathbf{a} \cdot \mathbf{x}$ “redshift effects” because (1) they are small

$$|a^j(\mathbf{a} \cdot \mathbf{x})| \sim |f^j(\mathbf{a} \cdot \mathbf{x})| \lesssim |(R^j_{0k0})_{\text{waves}}x^k|; \quad (8)$$

and (2) they are steady in time, and therefore—by contrast with the equally small wave-induced forces—they cannot excite resonant motions of the detector. Also, he

can separate the “inertial acceleration,” $-a^{\hat{j}}$, into a contribution from the local acceleration of gravity at the detector’s center of mass, $-(\partial\Phi/\partial x^j)_{x^j=0}$, plus a contribution $-a^{\hat{j}}_{\text{absolute}}$ due to acceleration of the detector relative to the “absolute space” of Newtonian theory. Finally, he can rewrite the Riemann curvature due to Newtonian gravity in the familiar form $R^{\hat{j}}_{\hat{0}\hat{k}\hat{0}} = \partial^2\Phi/\partial x^{\hat{j}}\partial x^{\hat{k}}$. The net result is

$$\begin{aligned} F^{\hat{j}} = & \left[\begin{array}{l} \text{total Newtonian force per unit} \\ \text{mass acting on mass element} \end{array} \right] \\ & + f^{\hat{j}} \left[\begin{array}{l} \text{Newtonian force per unit mass exerted by} \\ \text{adjacent matter and by electromagnetic fields} \end{array} \right] \\ & - a^{\hat{j}}_{\text{absolute}} \left[\begin{array}{l} \text{inertial force per unit mass due to acceleration} \\ \text{of detector relative to Newtonian absolute space} \end{array} \right] \\ & - \left(\frac{\partial\Phi}{\partial x^{\hat{j}}} \right)_{\text{at mass element}} \left[\begin{array}{l} = -(\partial\Phi/\partial x^j)_{x^j=0} - (\partial^2\Phi/\partial x^{\hat{j}}\partial x^k)x^k \\ = \text{Newtonian gravitational acceleration} \end{array} \right]. \end{aligned} \quad (9)$$

Conclusion: The equation of motion for a mass element of a mechanical detector, when written in the detector’s proper reference frame, has the standard Newtonian form (6), with standard Newtonian driving forces (9), plus a driving force due to the gravitational waves given by

$$(d^2x^{\hat{j}}/dt^2)_{\text{due to waves}} = -(R^{\hat{j}}_{\hat{0}\hat{k}\hat{0}})_{\text{waves}}x^{\hat{k}}. \quad (10)$$

from the metric coefficients $g_{\hat{\alpha}\hat{\beta}}$ of expression (37.2); to do so one would need the unknown corrections of $O(|x^{\hat{j}}|^2)$. However, one can easily obtain $R^{\hat{\alpha}}_{\hat{\beta}\hat{\gamma}\hat{\delta}}$ and $T^{(\text{GW})}_{\hat{\mu}\hat{\nu}}$ from the corresponding components in the TT coordinate frame [equations (37.1)] by applying the transformation matrix $\|\partial x^\alpha/\partial x^{\hat{\mu}}\|$. To make the transformation trivial, orient the TT coordinate frame so that, to a precision of $O(h_{\mu\nu}) \ll 1$, it coincides with the detector’s proper reference frame near the detector’s center of mass at the moment of interest, $t = \hat{t} = 0$. Then the transformation matrix will be

$$\partial x^\alpha/\partial x^{\hat{\mu}} = \delta_\mu^\alpha + O(h_{\mu\nu}) + O(a_j x^j) + O(|\mathbf{a}| \hat{t}). \quad (37.4)$$

$$\left[\begin{array}{l} \text{corrections due to} \\ \text{ripples in spacetime} \\ \text{caused by waves} \end{array} \right] \left[\begin{array}{l} \text{redshift} \\ \text{corrections} \end{array} \right] \left[\begin{array}{l} \text{corrections due to relative} \\ \text{velocity of frames resulting} \\ \text{from detector’s acceleration} \end{array} \right]$$

The acceleration the detector experiences is typically

$$|\mathbf{a}| = \text{one “Earth gravity”} = 980 \text{ cm/sec}^2 \sim 1/(\text{light-year}).$$

Description of waves in frame of detector

Therefore to enormous precision $\|\partial x^\alpha / \partial x^{\hat{\mu}}\| = \|\delta_\mu^\alpha\|$, and components of tensors are the same in the two reference frames:

$$\begin{aligned} R_{\hat{x}\hat{0}\hat{x}\hat{0}} &= -R_{\hat{y}\hat{0}\hat{y}\hat{0}} = -\frac{1}{2}\ddot{A}_+, & R_{\hat{x}\hat{0}\hat{y}\hat{0}} &= R_{\hat{y}\hat{0}\hat{x}\hat{0}} = -\frac{1}{2}\ddot{A}_x, \\ T_{\hat{0}\hat{0}}^{(\text{GW})} &= T_{\hat{z}\hat{z}}^{(\text{GW})} = -T_{\hat{0}\hat{z}}^{(\text{GW})} = \frac{1}{16\pi}\langle \dot{A}_+^2 + \dot{A}_x^2 \rangle_{\text{time avg.}} \end{aligned} \quad (37.5)$$

[see equation (37.1)].

Combining equations (37.3) and (37.5), one obtains for the wave-induced accelerations relative to the center of mass of the detector

Explicit form of accelerations due to waves

$$\begin{aligned} \left(\frac{d^2\hat{x}}{dt^2} \right)_{\text{due to waves}} &= -R_{\hat{x}\hat{0}\hat{x}\hat{0}}\hat{x} - R_{\hat{x}\hat{0}\hat{y}\hat{0}}\hat{y} = \frac{1}{2}(\ddot{A}_+\hat{x} + \ddot{A}_x\hat{y}), \\ \left(\frac{d^2\hat{y}}{dt^2} \right)_{\text{due to waves}} &= -R_{\hat{y}\hat{0}\hat{y}\hat{0}}\hat{y} - R_{\hat{y}\hat{0}\hat{x}\hat{0}}\hat{x} = \frac{1}{2}(-\ddot{A}_+\hat{y} + \ddot{A}_x\hat{x}), \\ \left(\frac{d^2\hat{z}}{dt^2} \right)_{\text{due to waves}} &= 0. \end{aligned} \quad (37.6)$$

This analysis is valid only for “small” detectors ($L \ll \lambda$)

These expressions, like the equation of geodesic deviation, are valid only over regions small compared to one wavelength. Second derivatives of the metric (i.e., the components of the Riemann tensor) give a poor measure of geodesic deviation and of wave-induced forces over regions of size $L \gtrsim \lambda$. Thus, to analyze large detectors ($L \gtrsim \lambda$), one must abandon the “local mathematics” of the curvature tensor and replace it by “global mathematics”—e.g., an analysis in the TT coordinate frame using the metric components $h_{\mu\nu}$. For an example, see exercise 37.6.

All detectors of high sensitivity that have been designed up until now (1973) are small compared to a wavelength, and therefore can be analyzed using the techniques of Newtonian physics and the driving forces of equations (37.6).

It is useful to develop physical intuition for the driving forces, $-R_{\hat{0}\hat{k}\hat{0}}^j x^k$, produced by waves of various polarizations. Figure 35.2 is one aid to such intuition; Box 37.2 is another. [The reader may find it interesting to examine, compare, and reconcile them!]

EXERCISES

Exercise 37.3. CONNECTION COEFFICIENTS IN PROPER REFERENCE FRAME

- (a) Calculate $\Gamma_{\beta\hat{\gamma}}^{\hat{\alpha}}$ for the metric (37.2), ignoring corrections of $O(|x^j|)$. [Answer: Equations (13.69) with $\omega^i = 0$.]
- (b) Calculate $R_{\hat{0}\hat{k}\hat{0}}^j$ using the standard formula (8.44), and leaving spatial derivatives of the connection coefficients unevaluated because of the unknown corrections of $O(|x^j|)$ in $\Gamma_{\beta\hat{\gamma}}^{\hat{\alpha}}$. [Answer: $R_{\hat{0}\hat{k}\hat{0}}^j = \Gamma_{\hat{0}\hat{0},k}^{\hat{j}} - a^{\hat{j}}a^k$.]
- (c) Use the answer to part (b) to evaluate the $O(|x^j|)$ corrections to $\Gamma_{\hat{0}\hat{0}}^j$. [Answer:

$$\Gamma_{\hat{0}\hat{0}}^j = a^{\hat{j}}(1 + a_k x^k) + R_{\hat{0}\hat{k}\hat{0}}^j x^k + O(|x^k|^2). \quad (37.7)$$

Box 37.2 LINES OF FORCE FOR GRAVITATIONAL-WAVE ACCELERATIONS**A. Basic Idea**

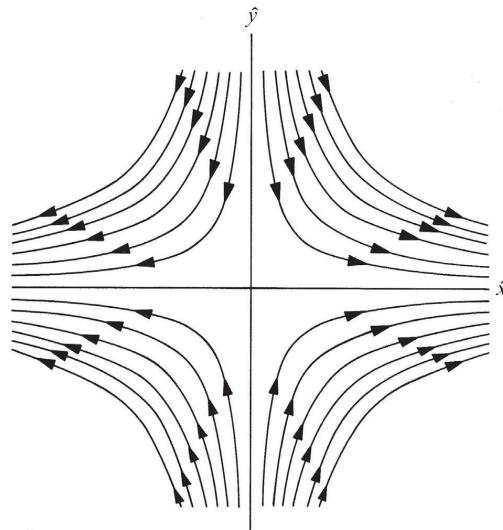
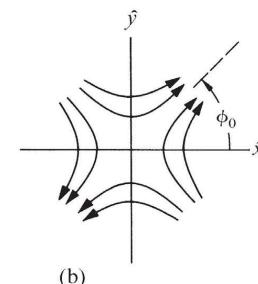
Consider a plane wave propagating in the \hat{z} direction. Discuss it entirely in the proper reference frame of a detector. The relative accelerations due to the wave are entirely transverse. Relative to the center of mass of the detector (origin of spatial coordinates) they are

$$\begin{aligned} d^2\hat{x}/d\hat{t}^2 &= \frac{1}{2}(\ddot{A}_+\hat{x} + \ddot{A}_x\hat{y}), \\ d^2\hat{y}/d\hat{t}^2 &= \frac{1}{2}(-\ddot{A}_+\hat{y} + \ddot{A}_x\hat{x}), \\ d^2\hat{z}/d\hat{t}^2 &= 0. \end{aligned} \quad (1)$$

Notice that these accelerations are divergence-free. Consequently they can be represented by “lines of force,” analogous to those of a vacuum electric field. At a value of $\hat{t} - \hat{z}$ where $\ddot{A}_x = 0$ (so polarization is entirely e_+), the lines of force are the hyperbolas shown here [sketch (a)]. The direction of the acceleration at any point is the direction of the arrow there; the magnitude of the acceleration is the density of force lines. Since acceleration is proportional to distance from center of mass, the force lines get twice as close together when one moves twice as far away from the origin in a given direction. When \ddot{A}_+ is positive, the arrows on the force lines are as shown in (a); when it is negative, they are reversed. As $|\ddot{A}_+|$ increases, the force lines move in toward the origin so their density goes up; as $|\ddot{A}_+|$ decreases, they move out toward infinity so their density goes down.

For polarization e_x the force lines are rotated by 45° from the above diagram. For intermediate polarization (values of $\hat{t} - \hat{z}$ where \ddot{A}_+ and \ddot{A}_x are both nonzero), the diagram is rotated by an intermediate angle [sketch (b)]

$$\phi_0 = \frac{1}{2} \text{arc tan } (\ddot{A}_x/\ddot{A}_+). \quad (2)$$

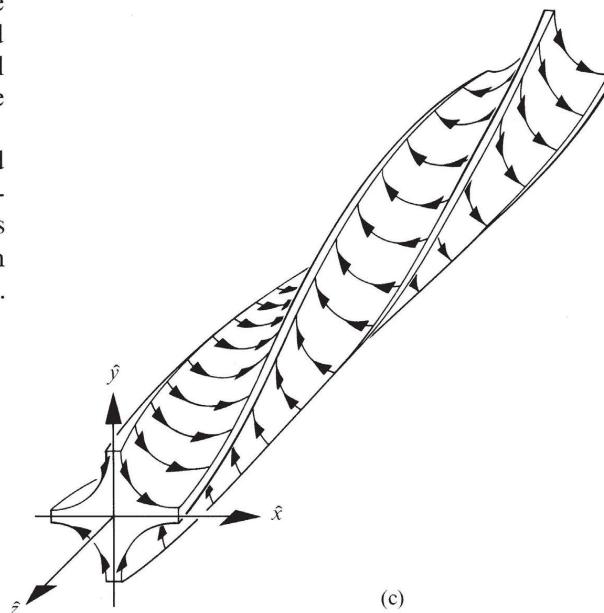
(a) Force lines for $\ddot{A}_x = 0, \ddot{A}_+ > 0$ 

(b)

Box 37.2 (continued)**B. Three-Dimensional Diagram**

At each value of $\hat{t} - \hat{z}$, the wave-produced accelerations have a specific polarization [orientation angle ϕ_0 of sketch (b)] and a specific amplitude (density of lines of force). Draw the lines of force in a three-dimensional $(\hat{x}, \hat{y}, \hat{z})$ diagram for fixed \hat{t} . Then as time passes the over-all diagram will remain unchanged in form, but will propagate with the speed of light in the \hat{z} direction.

Sketch (c) shows such a diagram for righthand circularly polarized waves of unchanging amplitude. Note: The authors are not aware of diagrams such as these [(a), (b), (c) above] and their use in analyzing detector response prior to William H. Press (1970).

**Exercise 37.4. WHY THE $a \cdot x$?**

Explain the origin of the $a \cdot x$ correction in equation (5b) of Box 37.1. [Hint: Take the viewpoint of an observer at rest at the spatial origin who watches two freely falling particles respond to the inertial force. At time $\hat{t} = 0$, put one particle at the origin and the other at x^j . As time passes, the separation of the particles in their common Lorentz frame remains fixed; so there develops a Lorentz contraction from the viewpoint of the observer at $x^j = 0$.]

Exercise 37.5. ORIENTATION OF POLARIZATION DIAGRAM

Derive equation (2) of Box 37.2.

§37.3. TYPES OF MECHANICAL DETECTORS

Eight types of mechanical detectors:

Figure 37.2 shows eight different types of mechanical detectors for gravitational waves. (By “mechanical detector” is meant a detector that relies on the relative

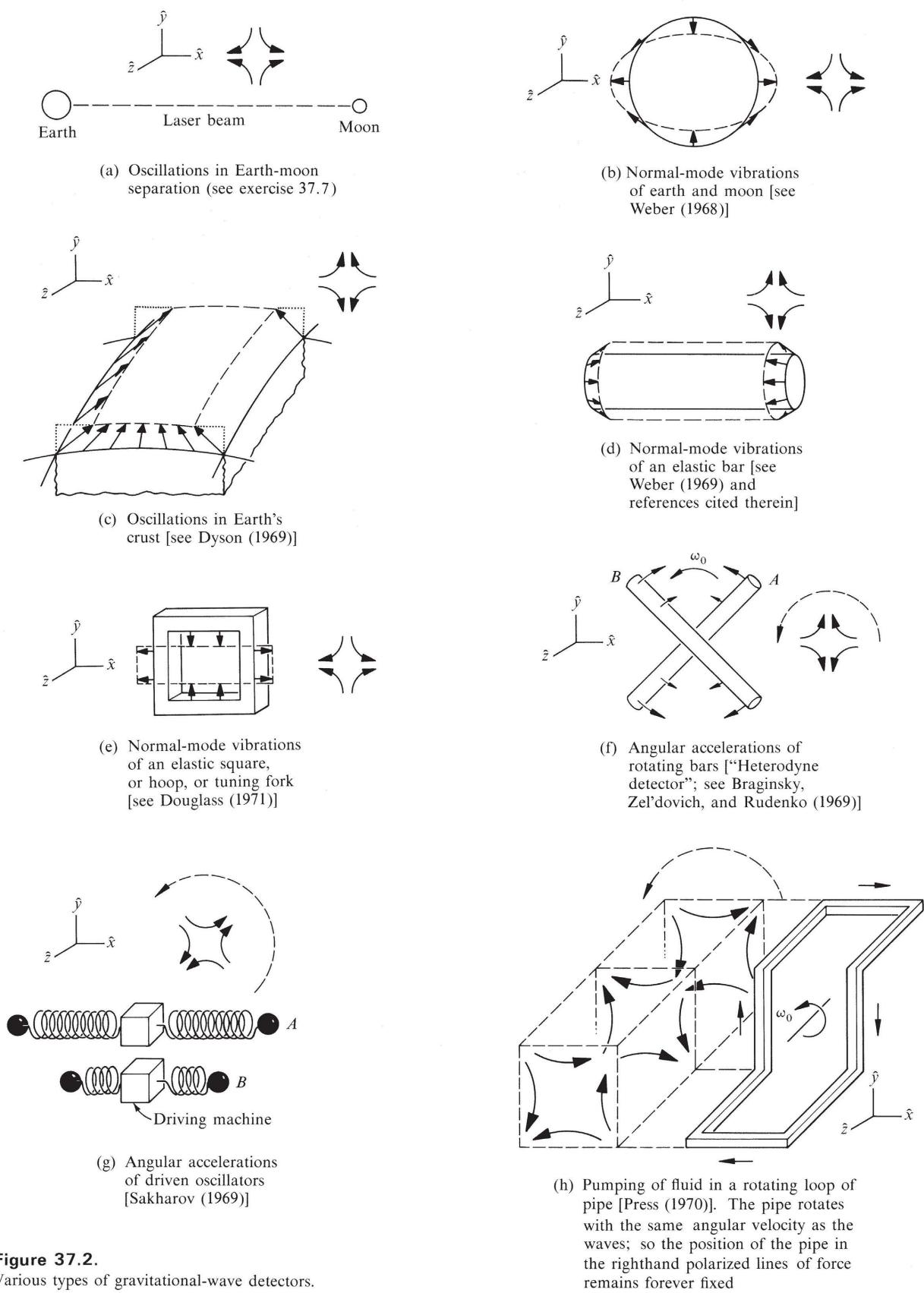


Figure 37.2.
Various types of gravitational-wave detectors.

motions of matter. Nonmechanical detectors are described in §37.9, near end of this chapter.) These eight detectors, and others, can be analyzed easily using the force-line diagrams of Box 37.2. A *qualitative* discussion of each of the eight detectors is given below. (A full quantitative analysis for each one would entail experimental technicalities for which general relativity is irrelevant, and which are beyond the scope of this book. However, some quantitative details are spelled out in §§37.5–37.8.)

1. The Relative Motions of Two Freely Falling Bodies

(1) freely falling bodies

As a gravitational wave passes two freely falling bodies, their proper separation oscillates (Figure 37.3). This produces corresponding oscillations in the redshift and round-trip travel times for electromagnetic signals propagating back and forth between the two bodies. Either effect, oscillating redshift or oscillating travel time, could be used in principle to detect the passage of the waves. Examples of such detectors are the Earth-Moon separation, as monitored by laser ranging [Fig. 37.2(a)]; Earth-spacecraft separations as monitored by radio ranging; and the separation between two test masses in an Earth-orbiting laboratory, as monitored by redshift measurements or by laser interferometry. Several features of such detectors are explored in exercises 37.6 and 37.7. As shown in exercise 37.7, such detectors have so low a sensitivity that they are of little experimental interest.

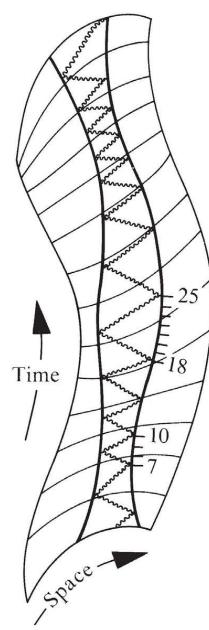


Figure 37.3.

Time of round-trip travel between two geodesics responds to oscillations in the curvature of spacetime (diagram is schematic only; symbolic of a laser pulse sent from the Earth to a corner reflector on the Moon and back at a time when a very powerful, long-wavelength gravitational wave passes by; the wave would have to be powerful because a direct measure of distance to better than 10 cm is difficult, and such precision produces a much less sensitive indicator of waves than the vibrations in length [10^{-14} cm or less] of a Weber bar; see exercise 37.7). The geodesics are curved toward each other in regions where the relevant component of the Riemann curvature tensor, call it $R_{\hat{x}\hat{t}\hat{x}\hat{t}}$, has one sign, and curved away from each other in regions where it has the opposite sign. The diagram allows one to see at a glance the answer to an often expressed puzzlement: Is not any change in round-trip travel time mere trumpery flummery? The metric perturbation, $\delta h_{\mu\nu}$, of the wave changes the scale of distances slightly but also correspondingly changes the scale of time. Therefore does not every possibility of any really meaningful and measurable effect cancel out? Answer: (1) The widened separation between the geodesics is not a local effect but a cumulative one. It does not arise from the local value of $\delta h_{\mu\nu}$ directly or even from the local value of the curvature. It arises from an accumulation of the bending process over an entire half-period of the gravitational wave. (2) The change in separation of the geodesics is a true change in proper distance, and shows up in a true change in proper time (see “ticks” on the world line of one of the particles). See exercise 37.6. Note: When one investigates the separation between the geodesics, not over a single period, as here, but over a large number of periods, he finds a cumulative, systematic, net slow bending of the rapidly wiggling geodesics toward each other. This small, attractive acceleration is evidence in gravitation physics of the effective mass-energy carried by the gravitational waves (see Chapter 35).

2. Normal-Mode Vibrations of the Earth and Moon

A gravitational wave sweeping over the Earth will excite its quadrupole modes of vibration, since the driving forces in the wave have quadrupole spatial distributions [see Fig. 37.2(b)]. The fundamental quadrupole mode of the Earth has a period of 54 minutes, while that of the moon has a period of 15 minutes. Thus, the Earth and Moon should selectively pick out the 54-minute and 15-minute components of any passing wave train. Section 37.7 will analyze quantitatively the interaction between the wave and solid-body vibrations. By comparing that analysis with seismometer studies of the Earth's vibrations, Weber (1967) put the first observational limit ever on the cosmic flux of gravitational waves:

$$I_\nu \equiv \frac{d \text{ flux}}{d \text{ frequency}} < 3 \times 10^7 \text{ erg cm}^{-2} \text{ sec}^{-1} \text{ Hz}^{-1} \text{ at } \nu = 3.1 \times 10^{-4} \text{ Hz.} \quad (37.8)$$

3. Oscillations in the Earth's Crust

If the neutron star in a pulsar is slightly deformed from axial symmetry, its rotation will produce gravitational waves. The period of the waves is half the period of the pulsar (rotation of star through 180° produces one period of waves)—i.e., it should range from 0.017 sec for NP0532 (Crab Pulsar) to 1.87 sec. for NP0527. Such a wave train cannot excite the 54-minute quadrupole vibration or any of the other normal, low-frequency modes of vibration of the Earth. The kind of vibrations it *can* excite allow themselves in principle to be described in the language of normal modes. However, they are clearly and more conveniently envisaged as vibrations of localized regions of the Earth; or, more particularly, vibrations of the Earth's crust.

Dyson (1969) has analyzed the response of an elastic solid, such as the Earth, to an incident, off-resonance gravitational wave. He shows that the response depends on irregularities in the elastic modulus for shear waves, and that it is strongest at a free surface [Figure 37.2(c)]. For the fraction of gravitational-wave energy crossing a flat surface that is converted into energy of elastic motion of the solid, he finds the expression

$$(\text{fraction}) = (8\pi G\rho/\omega^2)(s/c)^3 \times \sin^2\theta |\cos\theta|^{-1} [1 + \cos^2\theta + (s/v) \sin^2\theta]. \quad (37.9)$$

Here s and v are the velocities of shear waves and compressive waves, respectively, and θ is the angle between the direction of propagation of the waves and the normal to the surface. Considering a flux of 2×10^{-5} erg/cm² sec (an optimistic but conceivable value for waves from a pulsar) incident horizontally ($\theta = \pi/2$; “divergent” factor $|\cos\theta|^{-1}$ cancels out in calculation!), and taking s to be 4.5×10^5 cm/sec and ω to be 6 rad/sec, he calculates that the 1-Hz horizontal displacement produced in the surface has an amplitude of $\xi_0 \sim 2 \times 10^{-17}$ cm, too small by a factor of the order of 10^5 to be detected against background seismic noise. He points to the possibilities of improvements, especially via resonance (elastic waves reflected back and forth between two surfaces; Antarctic ice sheet).

(2) Earth and Moon

(3) Earth's crust

4. Normal-Mode Vibrations of an Elastic Bar

(4) elastic bar

As of 1972, the most often-discussed type of detector is the aluminum bar invented by Joseph Weber (1960, 1961) [see Figures 37.1 and 37.2(d)]. Weber's bars are cylindrical in shape, with length 153 cm, diameter 66 cm, and weight 1.4×10^6 g. Each bar is suspended by a wire in vacuum and is mechanically decoupled from its surroundings. Around its middle are attached piezoelectric strain transducers, which couple into electronic circuits that are sensitive to the bar's fundamental end-to-end mode of oscillation (frequency $\nu = 1,660$ Hz). When a gravitational wave hits the bar broadside, as shown in Figure 37.2(d), the relative accelerations carried by the wave will excite the fundamental mode of the bar. As of 1972, Weber observes sudden, simultaneous excitations in two such bars, one at the University of Maryland, near Washington, D.C.; the other at Argonne National Laboratory, near Chicago [see Weber (1969, 1970a,b)]. No one has yet come forward with a workable explanation for Weber's coincidences other than gravitational waves from outer space. However, the history of physics is rich with instances where supposedly new effects had to be attributed in the end to long familiar phenomena. Therefore it would seem difficult to rate the observed events as "battle-tested." To achieve that confidence rating would seem to require confirmation with different equipment, or under different circumstances, or both; experiments to provide that confirmation are now (1972) underway. If one makes this tentative assessment, one can be excused for expressing at the same time the greatest admiration for the experimental ingenuity, energy, and magnificent persistence that Joseph Weber has shown in his more than decade-long search for the most elusive radiation on the books of physics.

Mechanical detectors of the above four types represent systems on which measurements have been made; so practical difficulties and realizable noise levels can be estimated properly. In the continuing search for improved methods, more elaborate detectors are being studied, and in 1972 one can list a number of interesting proposals, as below. For these it is hard to know how much development would be required in order to achieve the desired performance.

5. Normal-Mode Vibrations of Elastic Bodies of Other Shapes

(5) elastic bodies of other shapes

The "bar" of a Weber detector need not be cylindrical in shape. For a discussion of a detector with the shape of a hollow square, a hoop, or a tuning fork, see Douglass (1971); such a detector might allow its fundamental frequency to be adjusted for the most favorable response, with given mass, or given maximum dimension, or both. Sections 37.4 and 37.7 and exercises 37.9 to 37.12 analyze in detail the operation of a "vibrating-bar" detector of arbitrary shape. See also Douglass and Tyson (1971).

6. Angular Accelerations of Rotating Bars

(6) rotating bars
("heterodyne detector")

All the potential detectors described thus far respond in the most obvious of manners to the tidal accelerations of a gravitational wave: relative distances oscillate in and

out. But the tidal accelerations contain, in addition to a length-changing component, also a tangential, rotation-producing component. In picture (a) of Box 37.2, the length-changing component dominates near the \hat{x} and \hat{y} axes, whereas the rotation-producing component dominates half-way between the axes. Vladimir B. Braginsky was the first to propose a detector that responds to the rotation-producing accelerations [see Braginsky, Zel'dovich, and Rudenko (1969); Braginsky and Nazarenko (1971)]. It consists of two metal rods, oriented perpendicular to each other, and rotating freely with angular velocity ω_0 in their common plane [see Fig. 37.2(f)]. (The rotation is relative to the gyroscopes of the proper reference frame of the detector; equivalently, it is relative to the Lorentz frame local to the detector.) Let monochromatic gravitational waves of angular frequency $\omega = 2\omega_0$ (change of phase per unit of time equals twice the angular velocity at which the pattern of lines of force turns) impinge broadside on the rotating rods. The righthand circularly polarized component of the waves will then rotate with the rods; so their orientation in its lines-of-force diagram will remain forever fixed. With the orientation of Fig. 37.2(f), rod A will undergo angular acceleration, while rod B will decelerate. The experimenter can search for the constant relative angular acceleration of the two rods (constant so long as the angle between them does not depart significantly from 90°). Better yet, the experimenter can (all too easily) adjust the rotation rate ω_0 so it does not quite match the waves' frequency ω . Then for $\frac{1}{2}\omega_0/|\omega - 2\omega_0|$ rotations, rod A will accelerate and B will decelerate; then will follow $\frac{1}{2}\omega_0/|\omega - 2\omega_0|$ rotations in which A decelerates and B accelerates, and so on (frequency beating). The experimenter can search for oscillations in the relative orientation of the rods. One need not worry about the lefthand polarized waves marring the experiment. Since they do not rotate with the rods, their angular accelerations average out over one cycle.

Such a device is called a "heterodyne detector" by Braginsky. He envisages that such detectors might be placed in free-fall orbits about the Earth late in the 1970's. Heterodyne detectors would work most efficiently for long monochromatic wave trains such as those from pulsars; but even for short bursts of waves they may be more sensitive than vibrating bars [see Braginsky and Nazarenko (1971)].

7. Angular Accelerations of Driven Oscillators

Andrei D. Sakharov (1969) has proposed a different type of detector for the angular accelerations of a gravitational wave. Instead of two rotating bars, it consists of two identical, driven oscillators, initially parallel and nonrotating, but oscillating out of phase with each other. Each oscillator experiences angular accelerations in one direction at one phase of a passing wave, and in the opposite direction at the next phase, but the torques do not cancel out. When the oscillator is maximally distended, it experiences a greater torque (acceleration \propto length; torque \propto length²) than when it is maximally contracted. Consequently, if the driven oscillations have the same angular frequency as a passing, monochromatic wave, and if the phases are as shown in Figure 37.2(g), then oscillator A will receive an angular acceleration in the righthand direction, while B receives an angular acceleration in the lefthand direction.

(7) rotation of driven oscillators

8. Pumping of Fluid in a Rotating Loop of Pipe

(8) fluid in pipe

A third type of detector that responds to angular accelerations has been described by William Press (1970). This detector would presumably be far less sensitive than others, and therefore not worth constructing; but it is intriguing in its novel design; and it illustrates features of gravitational waves ignored by other detectors. Press's detector consists of a loop of rotating pipe, containing a superfluid. The shape of the pipe and its constant rotation rate are chosen so that the gravitational waves will pump the fluid around inside the pipe. One conceivable pipe design (a bad one to build in practice, but an easy one to analyze) is shown in Fig. 37.2(h). Note that use is made of the variation in tidal acceleration along the direction of propagation of the wave as well as perpendicular to that direction. To analyze the response of the fluid to a righthand circularly polarized wave, one can mentally place the rotating pipe in the three-dimensional line-of-force diagram of Box 37.2(c).

EXERCISES

Exercise 37.6. RELATIVE MOTION OF FREELY FALLING BODIES AS A DETECTOR OF GRAVITATIONAL WAVES [see Figures 37.2(a) and 37.3.]

Consider two test bodies initially at rest with respect to each other in flat, empty spacetime. (The case where other, gravitating bodies are nearby can be treated without too much more difficulty; but this exercise concerns only the simplest example!) A plane, nearly monochromatic gravitational wave, with angular frequency ω and polarization e_+ , impinges on the bodies, coming from the $-z$ direction. As shown in exercise 35.5, the bodies remain forever at rest in those TT coordinates that constituted the bodies' global inertial frame before the wave arrived. Calculate, for arbitrary separations $(\Delta x, \Delta y, \Delta z)$ of the test bodies, the redshift and the round-trip travel time of photons going back and forth between them. Compare the answer, for large $\Delta x, \Delta y, \Delta z$, with the answer one would have obtained by using (without justification!) the equation of geodesic deviation. Physically, why does the correct answer *oscillate* with increasing separation? Discuss the feasibility and the potential sensitivity of such a detector using current technology.

Exercise 37.7. EARTH-MOON SEPARATION AS A GRAVITATIONAL-WAVE DETECTOR

In the early 1970's one can monitor the Earth-moon separation using laser ranging to a precision of 10 cm, with successive observations separated by at least one round-trip travel time. Suppose that no oscillations in round-trip travel time are observed except those (of rather long periods) to be expected from the Earth-moon-sun-planets gravitational interaction. What limits can one then place on the energy flux of gravitational waves that pass the Earth? The mathematical formula for the answer should yield numerically

$$\text{Flux} \lesssim 10^{18} \text{ erg/cm}^2 \text{ sec for } 0.3 \text{ cycle/sec} \lesssim \nu \lesssim 1 \text{ cycle/day}, \quad (37.10a)$$

corresponding to a limit on the mass density in gravitational waves of

$$\text{Density} \lesssim 10^{-13} \text{ g/cm}^3. \quad (37.10b)$$

Why is this an uninteresting limit?

§37.4. VIBRATING, MECHANICAL DETECTORS: INTRODUCTORY REMARKS

The remainder of this chapter (except for §37.9) gives a detailed analysis of vibrating, mechanical detectors (Earth; Weber bar; “bars” with complex shapes; and so on).

The details of the analysis and its applications depend in a crucial way on the values of two dimensionless numbers: (1) the ratio τ_{GW}/τ_0 , where

$$\tau_{\text{GW}} \equiv \left(\begin{array}{l} \text{characteristic time scale for changes in} \\ \text{gravitational-wave amplitude and spectrum} \end{array} \right), \quad (37.11\text{a})$$

$$\tau_0 \equiv \left(\begin{array}{l} e\text{-folding time for detector vibrations (in)} \\ \text{normal mode of interest) to die out as} \\ \text{a result of internal damping} \end{array} \right); \quad (37.11\text{b})$$

and (2) the ratio $\bar{E}_{\text{vibration}}/kT$, where

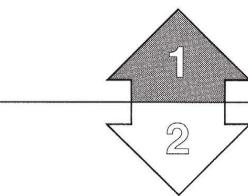
$$\bar{E}_{\text{vibration}} \equiv \left(\begin{array}{l} \text{mean value of detector's vibration energy (in)} \\ \text{normal mode of interest) while waves are} \\ \text{passing and driving detector} \end{array} \right), \quad (37.12\text{a})$$

$$\begin{aligned} kT &\equiv \left(\begin{array}{l} \text{Boltzman's} \\ \text{constant} \end{array} \right) \times \left(\begin{array}{l} \text{detector's} \\ \text{temperature} \end{array} \right) \\ &= \left(\begin{array}{l} \text{Mean energy in normal mode} \\ \text{of interest when gravitational} \\ \text{waves are not exciting it} \end{array} \right). \end{aligned} \quad (37.12\text{b})$$

When $\tau_{\text{GW}} \gg \tau_0$, the detector views the radiation as having a “*steady flux*,” and it responds with steady-state vibrations; when $\tau_{\text{GW}} \ll \tau_0$ (short burst of waves), the waves deal a “*hammer blow*” to the detector. When $\bar{E}_{\text{vibration}} \gg kT$, the driving force of the waves dominates over the detector’s random, internal, Brownian-noise forces (“*wave-dominated detector*”); when $\bar{E}_{\text{vibration}} \lesssim kT$, the driving force of the waves must compete with the detector’s random, internal, Brownian-noise forces (“*noisy detector*”).

Sections 37.5 to 37.7 deal with wave-dominated detectors ($\bar{E}_{\text{vibration}} \gg kT$). The key results of those sections are summarized in Box 37.3, which appears here as a quick preview (though it may not be fully understandable in advance). Section 37.8 treats noisy detectors.

Warning: Throughout the rest of this chapter prime attention focuses on the concept of cross section. This is fine for a first introduction to the theory of detectors. But cross section is not the entire story, especially when one wishes to study the detailed wave-form of the radiation. And sometimes (e.g., for the detector of Figure 37.2a), it is *none* of the story. A first-rate experimenter designing a new detector will not deal primarily in cross sections any more than a radio engineer will in designing a new radio telescope. Attention will also focus heavily on the bandwidth



The rest of this chapter is Track 2. No earlier track-2 material is needed as preparation for it, nor is it needed for any later chapter.

Definitions: “*steady flux*,” “*hammer-blow waves*,” “*wave-dominated detector*,” “*noisy detector*”

Design of detectors requires much more than the concept of cross section

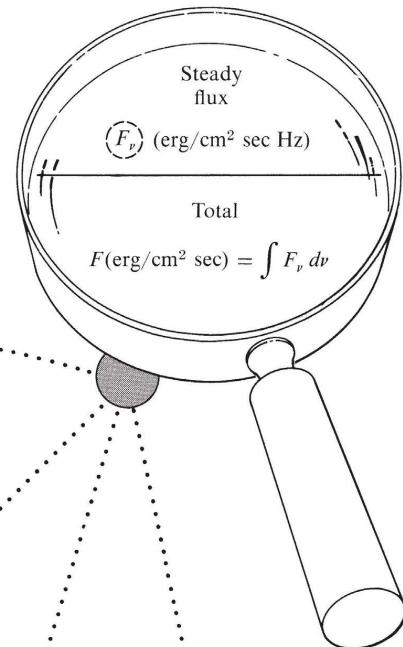
(continued on page 1022)

Box 37.3 WAYS TO USE CROSS SECTION FOR WAVE-DOMINATED DETECTORS
A. To Calculate Rate at which Detector Extracts Energy from a Steady Flux of Radiation

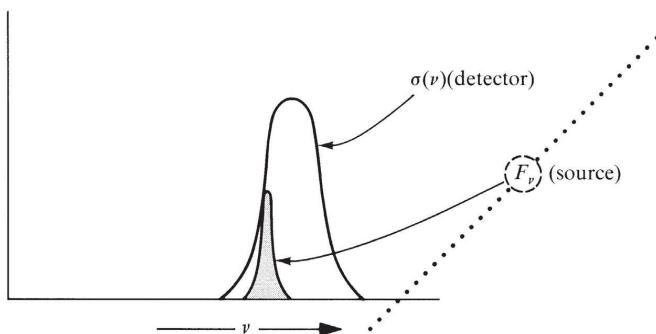
$$(\tau_{\text{GW}} \gg \tau_0)$$

1. Frequency distribution of radiation arbitrary:

$$\begin{aligned} & \left(\text{steady rate at which detector extracts energy from gravitational waves} \right) \\ & = \int \underbrace{(F_p(v)}_{\text{erg/cm}^2 \text{ sec Hz}} \cdot \underbrace{\sigma(v)}_{\text{cm}^2} \cdot \underbrace{dv}_{\text{Hz}} \end{aligned}$$



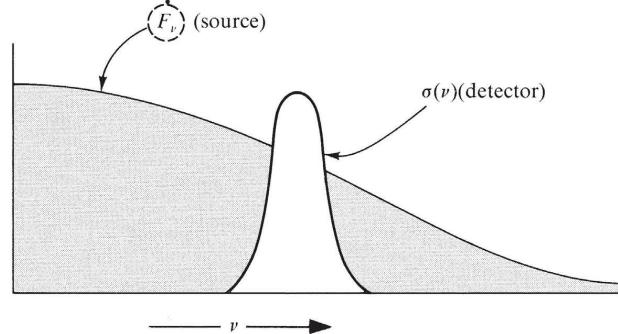
2. Frequency spread of radiation small compared to line width of detector:



$$\begin{aligned} & \left(\text{Steady rate at which detector extracts energy from gravitational waves} \right) \\ & = \sigma(v_{\text{source}}) \int (F_p) dv = \sigma F \end{aligned}$$

3. Frequency spread of radiation large compared to line width of detector:

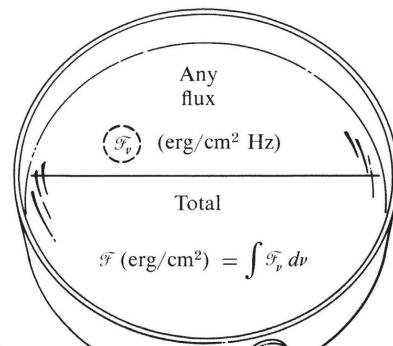
$$\begin{aligned} & \left(\text{steady rate at which detector extracts energy from gravitational waves} \right) \\ & = \underbrace{(F_p v_{\text{detector}})}_{\text{erg/cm}^2 \text{ sec Hz}} \cdot \underbrace{\int \sigma(v) dv}_{\text{"resonance integral", cm}^2 \text{ Hz}} \end{aligned}$$



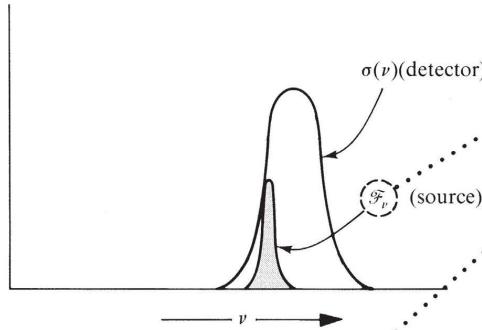
B. To Calculate *Total* Energy Deposited in Detector by *any* Passing Wave train

1. If frequency distribution of radiation is arbitrary:

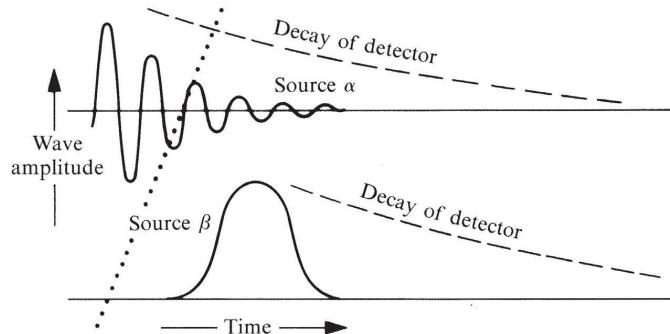
$$\left(\text{total energy deposited} \right) = \int \underbrace{\mathcal{F}_p(v)}_{\text{erg/cm}^2 \text{ Hz}} \cdot \underbrace{\sigma(v)}_{\text{cm}^2} \cdot \underbrace{dv}_{\text{Hz}}$$



2. If frequency spread of radiation is small compared to line width of detector ("monochromatic waves"):



$$\left(\text{total energy deposited} \right) = \underbrace{\sigma(v_{\text{source}})}_{\text{cm}^2} \int \underbrace{\mathcal{F}_p}_{\text{erg/cm}^2} dv = \sigma \mathcal{F}$$

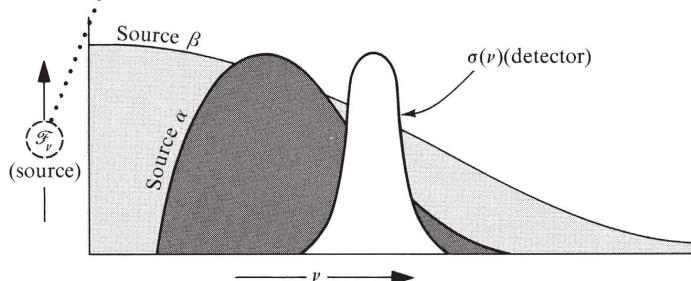


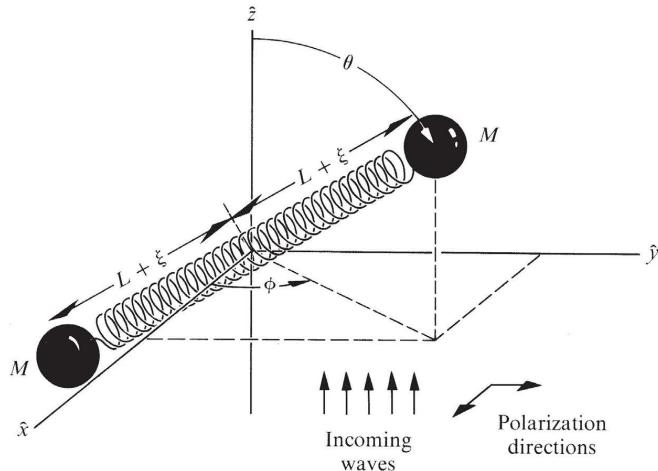
3. If frequency spread of radiation is large compared to line width of detector (as it *must* be for hammer-blown radiation, where):

$$\Delta v_{\text{source}} \gtrsim 1/4\pi\tau_{\text{GW}} \gg 1/4\pi\tau_0 = \Delta v_{\text{detector}}$$

$$(\text{total energy deposited}) =$$

$$\underbrace{\mathcal{F}_p(v_{\text{detector}})}_{\text{erg/cm}^2 \text{ Hz}} \underbrace{\int \sigma(v) dv}_{\text{cm}^2 \text{ Hz, 'resonance integral'}}$$



**Figure 37.4.**

An idealized detector (vibrator) responding to linearly polarized gravitational waves.

of the antenna, and on other, more detailed characteristics of its response, on coupling of the antenna to the displacement sensor, on response characteristics of the sensor, on antenna noise, on sensor noise, and so on. For an overview of these issues, and for discussions of detectors for which the concept of cross section is useless, see, e.g., Press and Thorne (1972).

§37.5. IDEALIZED WAVE-DOMINATED DETECTOR, EXCITED BY STEADY FLUX OF MONOCHROMATIC WAVES

Idealized detector: oscillator driven by a steady flux of monochromatic waves:

(1) derivation of equation of motion

Begin with the case of a *wave-dominated detector* ($\bar{E}_{\text{vibration}} \gg kT$) being driven by a steady flux of radiation ($\tau_{\text{GW}} \gg \tau_0$). Deal at first, not with a solid bar of arbitrary shape, but rather with the idealized detector of Figure 37.4: an oscillator made of two masses M on the ends of a spring of equilibrium length $2L$. Let the detector have a natural frequency of vibration ω_0 and a damping time $\tau_0 \gg 1/\omega_0$, so that its equation of motion (in the detector's proper reference frame) is

$$\ddot{\xi} + \dot{\xi}/\tau_0 + \omega_0^2 \xi = \text{driving acceleration.} \quad (37.13)$$

Let gravitational waves of polarization \mathbf{e}_+ and angular frequency ω impinge on the detector from the $-\hat{z}$ direction; and let the polar angles of the detector relative to the wave-determined $\hat{x}, \hat{y}, \hat{z}$ -axes be θ and ϕ .

The incoming waves are described by equations (37.1) with the amplitude

$$A_x = 0, \quad A_+ = \mathcal{A}_+ e^{-i\omega(t-z)}. \quad (37.14)$$

(Here and throughout one must take the real part of all complex expressions.)

Assume that the detector is much smaller than a wavelength, so that one can set $z \approx \hat{z} = 0$ throughout it. Then the tidal accelerations produced by the wave

$$\left(\frac{d^2\hat{x}}{dt^2} \right)_{\text{due to wave}} = -R_{\hat{x}\hat{y}\hat{y}}x^j = -\frac{1}{2}\omega^2\mathcal{Q}_+e^{-i\omega t}\hat{x},$$

$$\left(\frac{d^2\hat{y}}{dt^2} \right)_{\text{due to wave}} = -R_{\hat{y}\hat{y}\hat{y}}x^j = +\frac{1}{2}\omega^2\mathcal{Q}_+e^{-i\omega t}\hat{y},$$

have as their component along the oscillator

$$\begin{aligned} \frac{d^2\xi}{dt^2} &= \frac{\hat{x}}{L} \frac{d^2\hat{x}}{dt^2} + \frac{\hat{y}}{L} \frac{d^2\hat{y}}{dt^2} + \frac{\hat{z}}{L} \frac{d^2\hat{z}}{dt^2} = -\frac{1}{2}\omega^2\mathcal{Q}_+Le^{-i\omega t}\frac{\hat{x}^2 - \hat{y}^2}{L^2} \\ &= -\frac{1}{2}\omega^2\mathcal{Q}_+Le^{-i\omega t}\sin^2\theta\cos 2\phi. \end{aligned}$$

Consequently, the equation of motion for the oscillator is

$$\ddot{\xi} + \dot{\xi}/\tau_0 + \omega_0^2\xi = -\frac{1}{2}\omega^2\mathcal{Q}_+Le^{-i\omega t}\sin^2\theta\cos 2\phi. \quad (37.15)$$

The driving force varies as $\cos 2\phi$ because of the “spin-2” nature of gravitational waves: a rotation through 180° in the transverse plane leaves the waves unchanged; a rotation through 90° reverses the phase. The $\sin^2\theta$ term results from the transverse nature of the waves [one factor of $\sin \theta$ to account for projection onto the detector’s direction], plus their tidal-force nature [another factor of $\sin \theta$ to account for (relative force) \propto (distance in transverse plane)].

The straightforward steady-state solution of the equation of motion (37.15) is

$$\xi = \frac{\frac{1}{2}\omega^2\mathcal{Q}_+L\sin^2\theta\cos 2\phi}{\omega^2 - \omega_0^2 + i\omega/\tau_0} e^{-i\omega t}. \quad (37.16)$$

(2) oscillator amplitude as function of frequency and orientation

When the incoming waves are near resonance with the detector, $|\omega \pm \omega_0| \lesssim 1/\tau_0$, the oscillator is excited to large amplitude. Otherwise the excitation is small. Focus attention henceforth on near-resonance excitations; then equation (37.16) can be simplified (*note*: ω_0 is positive, but ω may be negative or positive):

$$\xi = \frac{\frac{1}{4}\omega_0\mathcal{Q}_+L\sin^2\theta\cos 2\phi}{|\omega| - \omega_0 + \frac{1}{2}\text{sgn}(\omega)i/\tau_0} e^{-i\omega t}. \quad (37.16')$$

One measure of the detector’s usefulness is its cross section for absorbing gravitational-wave energy. The steady-state vibrational energy in a detector with the above amplitude and with 2 masses M is

$$E_{\text{vibration}} = 2 \cdot \frac{1}{2} \cdot M \cdot (\dot{\xi}^2)_{\text{max}} = \frac{\frac{1}{16}ML^2\omega_0^4\mathcal{Q}_+^2\sin^4\theta\cos^2 2\phi}{(|\omega| - \omega_0)^2 + (1/2\tau_0)^2}. \quad (37.17)$$

This energy is being dissipated internally at a rate $E_{\text{vibration}}/\tau_0$. If one ignores reradiation of energy as gravitational waves (a negligible process!), one can equate the dissipation rate to the rate at which the detector absorbs energy from the incoming waves—which in turn equals the “cross section” σ times the incoming flux:

$$E_{\text{vibration}}/\tau_0 = -dE_{\text{waves}}/dt = \sigma T^{0z(\text{GW})} = \frac{1}{32\pi} \sigma \omega^2 \mathcal{Q}_+^2.$$

- (3) cross sections for polarized radiation

Consequently, *near resonance* ($|\omega \pm \omega_0| \ll \omega_0$), the cross section for interception of gravitational-wave energy is

$$\sigma = \frac{2\pi ML^2(\omega_0^2/\tau_0) \sin^4\theta \cos^2 2\phi}{(|\omega| - \omega_0)^2 + (1/2\tau_0)^2}, \quad \text{for polarized radiation.} \quad (37.18)$$

This expression applies to monochromatic radiation. However, experience with many other kinds of waves has taught that one often has to deal with a broad continuum of frequencies, with the “bandwidth” of the incident radiation far greater than the width of the detector resonance (see Box 37.3). Under these conditions, the relevant quantity is not the cross section itself, but the “resonance integral” of the cross section,

$$\int_{\text{resonance}} \sigma d\nu = \int \sigma(d\omega/2\pi) = 2\pi ML^2 \omega_0^2 \sin^4\theta \cos^2 2\phi, \quad (37.19)$$

for polarized radiation.

- (4) cross sections for unpolarized radiation

Before examining the magnitude of this cross section, scrutinize its directionality (the “antenna-beam pattern”). The factor of $\sin^4\theta \cos^2 2\phi$ refers to linearly polarized, \mathbf{e}_+ radiation (see Figure 37.4). For the orthogonal mode of polarization, \mathbf{e}_x , $\cos^2 2\phi$ is to be replaced by $\sin^2 2\phi$; and for unpolarized (incoherent mixture) radiation or circularly polarized radiation, the cross section is the average of these two expressions; thus

$$\sigma = \frac{\pi ML^2(\omega_0^2/\tau_0) \sin^4\theta}{(|\omega| - \omega_0)^2 + (1/2\tau_0)^2} \quad \text{for unpolarized radiation.} \quad (37.20)$$

Notice that this unpolarized cross section is peaked, with half-width 33° , about the equatorial plane of the detector. Averaged over all possible directions of incoming waves, the cross section is

$$\begin{aligned} \langle \sigma \rangle_{\text{all directions}} &= \frac{1}{2} \int_0^\pi \sigma \sin \theta d\theta = \frac{8}{15} \sigma_{\text{max}} \\ &= \frac{(8\pi/15)ML^2(\omega_0^2/\tau_0)}{(|\omega| - \omega_0)^2 + (1/2\tau_0)^2} \quad \text{for unpolarized radiation.} \end{aligned} \quad (37.21)$$

One can rewrite the above cross sections in several suggestive forms. For example, on resonance, the cross section (37.21) reads

$$\langle \sigma \rangle_{\text{all directions}} = \frac{4\pi^2}{15} \frac{4M}{2\pi/\omega_0} (\omega_0 \tau_0)(2L)^2.$$

Recall that $\omega_0\tau_0$ defines the “ Q ” of a detector, $1/Q \equiv$ (fraction of *energy* dissipated per radian of oscillation). Note that $2\pi/\omega_0$ is the wavelength λ_0 of resonant radiation. Finally, denote by $r_g = 4M$ the gravitational radius of the detector. In terms of these three familiar quantities, find for the cross section the formula

$$\begin{aligned}\langle\sigma\rangle_{\text{all directions}} &= \frac{\text{(cross section for absorbing waves on resonance)}}{(2L)^2} \\ &\quad \text{("geometric" cross section of detector)} \\ &= (4\pi^2/15)(r_g/\lambda_0)Q \quad \text{for unpolarized radiation} \\ &\quad \text{on resonance.}\end{aligned}\quad (37.22)$$

Magnitude of cross sections
for any resonant detector

This relation holds in order of magnitude for any resonant detector. It shows starkly that gravitational-wave astronomy must be a difficult enterprise. How large could you make the factor r_g/λ_0 , given a reasonable budget? Weber’s 1970 detectors have $2L_{\text{effective}} \approx 1$ meter, $r_g \approx (0.74 \times 10^{-28} \text{ cm/g}) \times (10^6 \text{ g}) \approx 10^{-22} \text{ cm}$, $\nu_0 = \omega_0/2\pi = 1,660 \text{ Hz}$, $\lambda_0 \approx 200 \text{ km}$, $r_g/\lambda_0 \approx \frac{1}{2} \times 10^{-29}$, $\tau_0 \approx 20 \text{ sec}$, $Q \approx 2 \times 10^5$; so that

$$\sigma_{\text{Weber}} \approx 3 \times 10^{-20} \text{ cm}^2 \text{ on resonance.} \quad (37.23)$$

Flux required to excite a
Weber-type detector

What flux of gravitational-wave energy would have to be incident to excite a cold detector ($\sim 0^\circ \text{ K}$) into roughly steady-state vibrations with a vibration energy of (Boltzmann’s constant) \times (room temperature) $\sim 4 \times 10^{-14} \text{ erg}$? The vibrator, if cooled enough to be wave-dominated, dissipates its energy at the rate $E_{\text{vibration}}/\tau_0 \sim 2 \times 10^{-15} \text{ erg/sec}$. The incident flux has to make up this loss, at the rate

$$T_{00}^{(\text{GW})}\sigma \sim 2 \times 10^{-15} \text{ erg/sec}, \quad (37.24a)$$

implying an incident flux of the order of $2 \times 10^{-15}/3 \times 10^{-20} \sim 10^5 \text{ erg/cm}^2 \text{ sec}$. Moreover, this flux has to be concentrated in the narrow range of resonance

$$\nu \approx \nu_0 \pm 1/4\pi\tau_0 = (1660 \pm 0.004) \text{ Hz.} \quad (37.24b)$$

By anybody’s standards, this is a very high flux of gravitational radiation for such a small bandwidth ($\sim 10^7 \text{ erg/cm}^2 \text{ sec Hz}$, as compared to the flux of blackbody gravitational radiation, $8\pi\nu^2kT/c^2 = 3 \times 10^{-27} \text{ erg/cm}^2 \text{ sec Hz}$, that would correspond to Planck equilibrium at the same temperature; the large factor of difference is a direct reflection of the difference in rate of damping of the oscillator by friction and by gravitational radiation).

Equation (37.22) makes it seem that an optimal detector must have a large Q . This is not necessarily so. Recall that the bandwidth, $\Delta\omega \approx \omega_0/Q$, over which the cross section is large, decreases with increasing Q . When an incoming steady flux of waves of bandwidth $\Delta\omega \gg \omega_0/Q \equiv 1/\tau_0$ and of specific flux

A large Q is not necessarily
optimal

$$F_\nu(\text{erg/cm}^2 \text{ sec Hz})$$

drives the detector, it deposits energy at the rate

$$\left(\frac{\text{rate of deposit}}{\text{of energy}} \right) = \frac{dE}{dt} = \int_{\text{resonance}} F_\nu \sigma d\nu = F_\nu(\nu_0) \int_{\text{resonance}} \sigma d\nu.$$

↑
for radiation with
bandwidth $\Delta\nu \gg 1/\tau_0$

Consequently, the relevant measure of detector effectiveness will be the integral of the cross section over the resonance, $\int \sigma d\nu$ (37.19). (See next section.) This frequency-integrated cross section is independent of the detector's Q , so one must use more sophisticated reasoning (e.g., signal-to-noise theory) in deciding whether a large Q is desirable. (See §37.8).

§37.6. IDEALIZED, WAVE-DOMINATED DETECTOR, EXCITED BY ARBITRARY FLUX OF RADIATION

Response of idealized detector to an arbitrary, non-monochromatic flux:

(1) derivation

Let plane-polarized waves of polarization e_+ but *arbitrary* spectrum [equation (37.1) with $A_x = 0$] impinge on the idealized detector of Figure 37.4. Then the equation of motion for the detector is the same as for monochromatic waves [equation (37.15)], but with $-\omega^2 \mathcal{A}_+ e^{-i\omega t}$ replaced by \ddot{A}_+ :

$$\ddot{\xi} + \dot{\xi}/\tau_0 + \omega_0^2 \xi = \frac{1}{2} \ddot{A}_+ L \sin^2\theta \cos 2\phi. \quad (37.26)$$

[By now one is fully accustomed to the fact that all analyses of detectors (when the detector is much smaller than the wavelength of the waves) are performed in the proper reference frame, with coordinates $\hat{t}, \hat{x}, \hat{y}, \hat{z}$. Henceforth, for ease of eyesight, abandon the “hats” on these “proper coordinates,” and denote them as merely t, x, y, z .]

Fourier-analyze the waves and the detector displacement,

$$A_+(t) = (2\pi)^{-1/2} \int_{-\infty}^{+\infty} \widetilde{A}_+(\omega) e^{-i\omega t}, \quad (37.27a)$$

$$\xi(t) = (2\pi)^{-1/2} \int_{-\infty}^{+\infty} \widetilde{\xi}(\omega) e^{-i\omega t}; \quad (37.27b)$$

and conclude from equation (37.26) that

$$\widetilde{\xi} = \frac{\frac{1}{2} \omega^2 \widetilde{A}_+ L \sin^2\theta \cos 2\phi}{\omega^2 - \omega_0^2 + i\omega/\tau_0}.$$

This Fourier amplitude is negligible unless $|\omega \pm \omega_0| \ll \omega_0$; consequently, without loss of accuracy, one can rewrite it as

$$\ddot{\xi} = \frac{\frac{1}{4} \omega_0 \tilde{A}_+ L \sin^2 \theta \cos 2\phi}{|\omega| - \omega_0 + \frac{1}{2} \operatorname{sgn}(\omega) i / \tau_0}. \quad (37.28)$$

[Compare with the steady-state amplitude (37.16').]

Ask how much total energy is deposited in the detector by the gravitational waves. Do *not* seek an answer by examining the amplitude of the vibrations, $\xi(t)$, directly; since that amplitude is governed by *both* internal damping and the driving force of the waves, it does not reflect directly the energy deposited. To get the total energy deposited, integrate over time the force acting on each mass multiplied by its velocity:

$$\left(\begin{array}{c} \text{total energy} \\ \text{deposited} \end{array} \right) = \int_{-\infty}^{+\infty} 2 \underbrace{\left(\frac{1}{2} M \ddot{A}_+ L \sin^2 \theta \cos 2\phi \right)}_{[2 \text{ masses}]} \dot{\xi} dt. \underbrace{\dot{\xi}}_{[\text{force on each mass}]} \underbrace{dt}_{[\text{velocity of each mass}]}$$

Use Parseval's theorem (one of the most powerful tools of mathematical physics!) to replace the time integral by a frequency integral

$$\left(\begin{array}{c} \text{total energy} \\ \text{deposited} \end{array} \right) = \Re \int_{-\infty}^{+\infty} (ML \sin^2 \theta \cos 2\phi) (-\omega^2 \tilde{A}_+^*) (-i\omega \ddot{\xi}) d\omega.$$

Then use equation (37.28) to rewrite this entirely in terms of the wave amplitude

$$\left(\begin{array}{c} \text{total energy} \\ \text{deposited} \end{array} \right) = \int_{-\infty}^{+\infty} \left[\frac{2\pi(\omega_0^2/\tau_0)ML^2 \sin^4 \theta \cos^2 2\phi}{(|\omega| - \omega_0)^2 + (1/2\tau_0)^2} \right] \left[\frac{\omega^2 |\tilde{A}_+|^2}{16\pi} \right] d\omega. \quad (37.29)$$

The first term in this expression is precisely the cross section for monochromatic waves, derived in the last section (37.18). The second term has an equally simple interpretation: the total energy that the gravitational waves carry past a unit surface area of detector is

$$\begin{aligned} \mathcal{F}(\text{ergs/cm}^2) &= \int T_{00}^{(\text{GW})} dt = \int \frac{1}{16\pi} \dot{A}_+^2 dt \\ &= \int \frac{\omega^2 |\tilde{A}_+|^2}{16\pi} d\omega = \int \frac{\omega^2 |\tilde{A}_+|^2}{8} d\nu \end{aligned} \quad (37.30)$$

(Parseval's theorem again!). Consequently, the energy per unit frequency interval, per unit area carried by the waves is

$$\mathcal{F}_\nu(\text{ergs/cm}^2 \text{ Hz}) = \frac{1}{8} \omega^2 |\tilde{A}_+|^2 \quad (37.31)$$

[for $-\infty < \nu < +\infty$; double this for $0 < \nu < +\infty$, a convention we use for the rest of this chapter]. This is 2π times the second term in (37.29).

Combining equations (37.18), (37.29), and (37.31), then, one finds

$$\left(\begin{array}{c} \text{total energy} \\ \text{deposited} \end{array} \right) = \int \sigma(\nu) \mathcal{F}_\nu(\nu) d\nu. \quad (37.32) \quad \begin{matrix} (2) \text{ answer} \\ \left(\begin{array}{c} \text{energy} \\ \text{deposited} \end{array} \right) = \int \sigma \mathcal{F}_\nu d\nu \end{matrix}$$

How one can measure energy deposited

This is the total energy deposited, regardless of the spectrum of the waves, and regardless of whether they come in a steady flux for a long time, or in a short burst, or in any other form. It is perfectly general—so long as the detector is wave-dominated ($E_{\text{vibration}} \gg kT$) while the waves are driving it.

How can an experimenter measure the total energy deposited? He cannot measure it directly, in general, but he can measure a quantity equal to it: the total energy that goes into internal damping, i.e., into “friction.” Energy is removed by “friction” at a rate $E_{\text{vibration}}/\tau_0$, when the vibration energy is much greater than kT (during period of wave-dominance). Therefore, the experimenter can measure

$$\left(\begin{array}{l} \text{total energy} \\ \text{deposited} \end{array} \right) = \frac{1}{\tau_0} \int E_{\text{vibration}} dt, \quad \text{in general.} \quad (37.33)$$

↑
[integrate over the period that $E_{\text{vibration}} \gg kT$]

In the special case of “hammer-blow waves” (τ_{GW} = duration of waves $\ll \tau_0$), the vibration energy is driven “instantaneously” from $\sim kT$ to a peak value, $E_{\text{vibration}}^{\text{peak}} \gg kT$, and then decays exponentially back to $\sim kT$; thus

$$\left(\begin{array}{l} \text{total energy} \\ \text{deposited} \end{array} \right) = \frac{1}{\tau_0} \int_0^{\infty} E_{\text{vibration}}^{\text{peak}} e^{-t/\tau_0} dt = E_{\text{vibration}}^{\text{peak}} \quad (37.34)$$

for hammer-blow waves.

When the waves are steady for a long period of time ($\tau_{\text{GW}} \gg \tau_0$), with specific flux

$$F_{\nu} = \mathcal{F}_{\nu}/\tau_{\text{GW}} \quad (\text{ergs/cm}^2 \text{ sec Hz}),$$

then the energy will be deposited at a constant rate

$$(dE/dt) = (\text{total energy deposited})/\tau_{\text{GW}};$$

and equation (37.32) can be rewritten

$$\left(\begin{array}{l} \text{rate of deposit} \\ \text{of energy} \end{array} \right) = \int \sigma(\nu) F_{\nu} d\nu, \quad \text{for steady waves } (\tau_{\text{GW}} \gg \tau_0). \quad (37.35)$$

Equations (37.32) and (37.35) are the key equations for application of the concept of cross section to realistic situations. They are applicable not only to polarized radiation, but also to unpolarized radiation and to radiation coming in from all directions, if one merely makes sure to use the appropriate cross section [equation (37.20) or (37.21) instead of (37.18)]. For examples of their application, see Box 37.3.

§37.7. GENERAL WAVE-DOMINATED DETECTOR, EXCITED BY ARBITRARY FLUX OF RADIATION

The cross sections of the idealized spring-plus-mass detector can be put into a form more elegant than equations (37.18) to (37.21)—a form that makes contact with many

branches of physics, and is valid for *any* vibrating resonant detector whatsoever.

Introduce the “Einstein A -coefficients,” which describe the rate at which a unit amount of detector energy is lost to internal damping and to reradiation of gravitational waves:

$$A_{\text{diss}} \equiv \left(\frac{\text{rate at which energy is dissipated internally}}{\text{energy in oscillations of detector}} \right) = \frac{1}{\tau_0}, \quad (37.36a)$$

$$A_{\text{GW}} \equiv \left(\frac{\text{rate at which energy is reradiated}}{\text{energy in oscillations}} \right). \quad (37.36b)$$

Cross sections reexpressed in terms of “Einstein A -coefficients”

For the idealized detector of Figure 37.4, the standard formula (36.1) for the emission of gravitational waves yields

$$(\text{power reradiated}) = \frac{32}{15} \omega^6 M^2 L^2 \langle \xi^2 \rangle_{\text{time avg.}} \quad (37.37)$$

(see exercise 37.8). Consequently

$$A_{\text{GW}} = \frac{16}{15} M L^2 \omega^4. \quad (37.38)$$

One can use these relations to rewrite the detector cross sections in terms of A_{diss} , A_{GW} , and the reduced wavelength

$$\lambda \equiv 1/\omega \quad (37.39)$$

of the radiation. For example, the cross section (37.21)—now with $\omega \geq 0$ —is

$$\langle \sigma \rangle_{\text{all directions}} = \frac{1}{2} \pi \lambda^2 \frac{A_{\text{GW}} A_{\text{diss}}}{(\omega - \omega_0)^2 + (A_{\text{diss}}/2)^2} \quad \begin{array}{l} \text{for unpolarized} \\ \text{radiation} \end{array} \quad (37.40)$$

(recall the assumption $|\omega - \omega_0| \ll \omega_0$ in all cross-section formulas) and the corresponding integral over the resonance is

$$\int \langle \sigma \rangle_{\text{all directions}} d\nu = \frac{1}{2} \pi \lambda_0^2 A_{\text{GW}} \text{ for polarized radiation.} \quad (37.41)$$

These expressions for the cross section are comprehensive in their application. They apply to any vibrating, resonant, gravitational-wave detector whatsoever, as one sees from the “detailed balance” calculation of exercise 37.9, and from the dynamic calculations of exercise 37.10. They also apply, with obvious changes in statistical factors and notation, to compound-nucleus reactions in nuclear physics (“Breit-Wigner formula”; see Blatt and Weisskopf, pp. 392–94, 408–10, 555–59), to the absorption of photons by atoms and molecules, to reception of electromagnetic waves by a television antenna, etc. Equation (37.41) says in effect, “Calculate the rate at which the oscillator is damped by emission of gravitational radiation; multiply that rate by the geometric factor familiar in all work with antennas, $\frac{1}{2} \pi \lambda_0^2$, and immediately obtain the resonance integral of the cross section. The result is expressed in geometric

Generality of the A -coefficient formalism

Scattering of radiation by detector

units (cm). To get the resonance integral in conventional units, multiply by the conversion factor $c = 3 \times 10^{10}$ cm Hz.

The ‘dynamic analysis’ of the idealized masses-on-spring detector, as developed in the last section, is readily extended to a vibrating detector of arbitrary shape (Earth; Weber’s bar; an automobile fender; and so on). The extension is carried out in exercise 37.10 and its main results are summarized in Box 37.4.

Part of the energy that goes into a detector is reradiated as scattered gravitational radiation. For any detector of laboratory dimensions with laboratory damping coefficients, this fraction is fantastically small. However, in principle one can envisage a larger system and conditions where the reradiation is not at all negligible. In such an instance one is dealing with scattering. No attempt is made here to analyze such scattering processes. For a simple order-of-magnitude treatment, one can use the same type of Breit-Wigner scattering formula that one employs to calculate the scattering of neutrons at a nuclear resonance or photons at an optical resonance. A still more detailed account will analyze the correlation between the polarization of the scattered radiation and the polarization of the incident radiation. The kind of formalism useful here for gravitational radiation with its tensor character will be very much like that now used to treat polarization of radiation with a spin-1 character. Here notice especially the “Madison Convention” [Barschall and Haeberli (1971)] developed by the collaborative efforts of many workers after experience during many years with a variety of conflicting notations. Considering the way in which the best notation that is available today for spin-1 radiation was evolved, one can only feel that it is too early to canonize any one notation for describing the scattering parameters for an object that is scattering gravitational radiation.

EXERCISES

Exercise 37.8. POWER RERADIATED

The idealized gravitational wave detector of Figure 37.4 vibrates with angular frequency ω . Show that the power it radiates as gravitational waves is given by equation (37.37).

Exercise 37.9. CROSS SECTIONS CALCULATED BY DETAILED BALANCE

Use the principle of detailed balance to derive the cross sections (37.41) for a vibrating, resonant detector of any size, shape, or mass (e.g., for the vibrating Earth, or Weber’s vibrating cylinder, or the idealized detector of Figure 37.4). [*Hints:* Let the detector be in thermal equilibrium with a bath of blackbody gravitational waves. Then it must be losing energy by reradiation as rapidly as it is absorbing it from the waves. (Internal damping can be ignored because, in true thermal equilibrium, energy loss by internal damping will match energy gain from random internal Brownian forces.) In detail, the balance of energy in and out reads [with I_ν = “specific intensity,” equation (22.48)]

$$\begin{aligned} & [4\pi I_\nu(\nu = \nu_0)]_{\text{blackbody}} \times \int \langle \sigma \rangle_{\text{all directions}} d\nu \\ & = A_{\text{GW}} \times (\text{Energy in normal mode of detector}). \end{aligned}$$

Solve for $\int \langle \sigma \rangle d\nu$, using the familiar form of the Planck spectrum and the fact that gravitational waves have two independent states of polarization.] *Note:* Because detailed balance

Box 37.4 VIBRATING, RESONANT DETECTOR OF ARBITRARY SHAPE**A. Physical Characteristics of Detector**

1. Detector is a solid object (Earth, Weber bar, automobile fender, . . .) with density distribution $\rho(x)$ and total mass $M = \int \rho d^3x$.
2. Detector has normal modes of vibration. The n th normal mode is characterized by:

ω_n = angular frequency;

$$\tau_n = \left(\begin{array}{l} \text{e-folding time for vibration energy} \\ \text{to decay as result of internal damping} \end{array} \right) \gg 1/\omega_n; \quad (1)$$

$\mathbf{u}_n(x)$ = eigenfunction (defined here to be dimensionless and real).

The eigenfunctions \mathbf{u}_n are orthonormalized, so that

$$\int \rho \mathbf{u}_n \cdot \mathbf{u}_m d^3x = M \delta_{nm}. \quad (2)$$

3. During a normal-mode vibration with $E_{\text{vibration}} \gg kT$, a mass element originally at \hat{x} receives the displacement

$$\delta x = \xi = \mathbf{u}_n(x) \mathcal{B}_n e^{-i\omega_n t - t/\tau_n}, \quad (3a)$$

↑
[constant amplitude]

the density at fixed x changes by

$$\delta \rho = -\nabla \cdot (\rho \mathbf{u}_n) \mathcal{B}_n e^{-i\omega_n t - t/\tau_n}, \quad (3b)$$

and the moment of inertia tensor oscillates

$$\delta I_{jk} = I_{(n)jk} \mathcal{B}_n e^{-i\omega_n t - t/\tau_n}. \quad (3c)$$

Here $I_{(n)jk}$ is the “moment of inertia factor for the n th normal mode”:

$$\begin{aligned} I_{(n)jk} &\equiv \int -(\rho u_n^l)_{,l} x^j x^k d^3x \\ &= \int \rho (u_n^j x^k + u_n^k x^j) d^3x \end{aligned} \quad (4)$$

[dimensions: mass \times length, multiply by \mathcal{B}_n (length) to get I_{jk}].

The corresponding “reduced quadrupole factor for the n th normal mode” is

$$f_{(n)jk} \equiv I_{(n)jk} - \frac{1}{3} I_{(n)ll} \delta_{jk}. \quad (5)$$

Box 37.4 (continued)**B. Cross Sections for Detector (exercise 37.10)**

1. For *polarized radiation* with propagation direction \mathbf{n} and polarization tensor \mathbf{e} :

$$h_{jk} = A(t - \mathbf{n} \cdot \mathbf{x})e_{jk},$$

$$e_{jk}n_k = 0, \quad e_{jj} = 0, \quad e_{jk}e_{jk} = 2; \quad (6)$$

$$\sigma_n(\nu) = \sigma_n(\omega/2\pi) = \frac{\pi}{4} \frac{|\mathcal{H}_{(n)jk}e_{jk}|^2}{M} \frac{\omega_n^2/\tau_n}{(|\omega| - \omega_n)^2 + (1/2\tau_n)^2}, \quad (7a)$$

$$\int_{\text{resonance}} \sigma_n d\nu = \frac{\pi}{4} \frac{|\mathcal{H}_{(n)jk}e_{jk}|^2}{M} \omega_n^2. \quad (7b)$$

2. For *unpolarized radiation* (random mixture of polarizations) with propagation direction \mathbf{n} , cross sections are

$$\sigma_n(\nu) = \sigma_n(\omega/2\pi) = \frac{\pi}{4} \frac{(\mathcal{H}_{(n)jk}^{\text{TT}})^2}{M} \frac{\omega_n^2/\tau_n}{(|\omega| - \omega_n)^2 + (1/2\tau_n)^2}, \quad (8a)$$

$$\int_{\text{resonance}} \sigma_n d\nu = \frac{\pi}{4} \frac{(\mathcal{H}_{(n)jk}^{\text{TT}})^2}{M} \omega_n^2. \quad (8b)$$

Here $\mathcal{H}_{(n)jk}^{\text{TT}}$ is the transverse-traceless part of $\mathcal{H}_{(n)jk}$ (transverse and traceless relative to the propagation direction \mathbf{n}):

$$\mathcal{H}_{(n)}^{\text{TT}} = P\mathcal{H}_{(n)}P - \frac{1}{2}P \text{ trace } (P\mathcal{H}_{(n)}), \quad P_{jk} \equiv \delta_{jk} - n_j n_k. \quad (9)$$

(See Box 35.1)

3. Cross sections for *unpolarized radiation, averaged over all directions*, are

$$\langle \sigma_n(\nu) \rangle_{\text{all directions}} = \frac{1}{2}\pi\lambda^2 \frac{A_{\text{GW}}A_{\text{diss}}}{(|\omega| - \omega_n)^2 + (A_{\text{diss}}/2)^2}, \quad (10a)$$

$$\int_{\text{resonance}} \langle \sigma_n \rangle_{\text{all directions}} d\nu = \frac{1}{2}\pi\lambda^2 A_{\text{GW}}, \quad (10b)$$

where the Einstein A coefficients are

$$A_{\text{diss}} = 1/\tau_n, \quad (11)$$

$$A_{\text{GW}} = \frac{1}{5} \frac{(\mathcal{H}_{(n)jk})^2}{M} \omega_n^4. \quad (12)$$

C. Spectrum Radiated by an Aperiodic Source (exercise 37.11)

It is instructive to compare these formulas with expressions for the radiation emitted by an aperiodic source.

- Fourier-analyze the reduced quadrupole factor of the source

$$\tilde{I}_{jk}(t) = (2\pi)^{-1/2} \int_{-\infty}^{+\infty} \tilde{I}_{jk}(\omega) e^{-i\omega t} d\omega.$$

- Then the total energy per unit frequency ($\nu \geq 0$) radiated over all time, into a unit solid angle about the direction \mathbf{n} , and with polarization tensor \mathbf{e} , is

$$\frac{dE}{d\nu d\Omega} = \frac{1}{8} \sum_{\omega = \pm 2\pi\nu} |\tilde{I}_{jk} e_{jk}|^2 \omega^6 \quad (13a)$$

[compare with equations (7)]. Summed over polarizations, this is

$$\frac{dE}{d\nu d\Omega} = \frac{1}{2} \sum_{j,k} |\tilde{I}_{jk}^{\text{TT}}|^2 \omega^6 \quad (13b)$$

[compare with equations (8)]. Here $\nu \geq 0$.

- The total energy radiated per unit frequency, integrated over all directions, still with $\nu \geq 0$, is

$$dE/d\nu = \frac{4\pi}{5} \sum_{j,k} |\tilde{I}_{jk}|^2 \omega^6 \quad (14)$$

[compare with equations (10)–(12)].

can be applied to any kind of resonant system in interaction with any kind of thermal bath of radiation or particles, equations (37.40) and (37.41), with appropriate changes of statistical factors, have wide generality.

Exercise 37.10. NORMAL-MODE ANALYSIS OF VIBRATING, RESONANT DETECTORS

Derive all the results for vibrating, resonant detectors quoted in Box 37.4. Pattern the derivation after the treatment of the idealized detector in §37.6. [Guidelines: (a) Let the detector be driven by the polarized waves of equation (6), Box 37.4; and let it be wave-dominated ($E_{\text{vibration}} \gg kT$). Show that the displacements $\delta\mathbf{x} = \xi(\mathbf{x}, t)$ of its mass elements are described by

$$\xi = \sum_n B_n(t) \mathbf{u}_n(x), \quad (37.42a)$$

where the time-dependent amplitude for the n th mode satisfies the driven-oscillator equation

$$\ddot{B}_n + \dot{B}_n/\tau_n + \omega_n^2 B_n = R_n(t), \quad (37.42b)$$

and where the curvature-induced driving term is

$$\begin{aligned} R_n(t) &= -R_{j\hat{0}k\hat{0}} \int (\rho/M) u_n^j x^k d^3x \\ &= \frac{1}{4} \tilde{A}(I_{(n)jk} e_{jk})/M. \end{aligned} \quad (37.42c)$$

(See Box 37.4 for notation.)

(b) Fourier-analyze the amplitudes of the detector and waves,

$$B_n(t) = (2\pi)^{-1/2} \int_{-\infty}^{+\infty} \tilde{B}_n(\omega) e^{-i\omega t} d\omega, \quad (37.42d)$$

$$A(t) = (2\pi)^{-1/2} \int_{-\infty}^{+\infty} \tilde{A}(\omega) e^{-i\omega t} d\omega, \quad (37.42e)$$

and solve the equation of motion (37.42b,c) to obtain, in the neighborhood of resonance,

$$\tilde{B}_n = \frac{\frac{1}{8} \omega_n \tilde{A}(I_{(n)jk} e_{jk})/M}{|\omega| - \omega_n + \frac{1}{2} i/\tau_n} \quad \text{for } |\omega \pm \omega_n| \ll \omega_n. \quad (37.42f)$$

(c) Calculate the total energy deposited in the detector by integrating

$$\left(\begin{array}{l} \text{energy} \\ \text{deposited} \end{array} \right) = \int (\text{Force per unit volume}) \cdot (\text{Velocity}) d^3x dt.$$

Thereby obtain

$$\left(\begin{array}{l} \text{energy deposited in} \\ \text{nth normal mode} \end{array} \right) = \frac{1}{4} (I_{(n)jk} e_{jk}) \int \tilde{A} \dot{B}_n dt.$$

(d) Apply Parseval's theorem and combine with expression (37.42f) to obtain

$$\left(\begin{array}{l} \text{energy deposited in} \\ \text{nth normal mode} \end{array} \right) = \int \sigma_n(\nu) \tilde{\mathcal{F}}_\nu(\nu) d\nu, \quad (37.43)$$

where σ_n is given by equation (7a) of Box 37.4, and (for $-\infty < \omega < +\infty$)

$$\tilde{\mathcal{F}}_\nu(\nu) = \tilde{\mathcal{F}}_\nu(\omega/2\pi) = \frac{1}{8} \omega^2 |\tilde{A}|^2. \quad (37.44)$$

(e) Show that $\tilde{\mathcal{F}}_\nu(\nu)$ is the total energy per unit area per unit frequency carried by the waves past the detector.

(f) Obtain all the remaining cross sections quoted in Box 37.4 by appropriate manipulations of this cross section. Use the mathematical tools for projecting out and integrating "transverse-traceless parts," which were developed in Box 35.1 and exercise 36.9.

Exercise 37.11. SPECTRUM OF ENERGY RADIATED BY A SOURCE

Derive the results quoted in the last section of Box 37.4.

Exercise 37.12. PATTERNS OF EMISSION AND ABSORPTION

The elementary dumbbell oscillator of Figure 37.4, initially unexcited, has a cross section for absorption of unpolarized gravitational radiation proportional to $\sin^4\theta$, and when excited radiates with an intensity also proportional to $\sin^4\theta$ (Chapter 36). The patterns of emission and absorption are identical. Any other dumbbell oscillator gives the same pattern, apart from a possible difference of orientation. Consider a nonrotating oscillator of general shape undergoing free vibrations in a single nondegenerate (and therefore nonrotatory) mode, or excited from outside by *unpolarized* radiation.

- (a) Show that its pattern of emission is identical with its pattern of absorption. [Hint: Make the comparisons suggested in the last few parts of Box 37.4.]
- (b) Show that this emission pattern (\equiv absorption pattern), apart from three Euler angles that describe the orientation of this pattern in space, and apart from a fourth parameter that determines total intensity, is uniquely fixed by a single ("fifth") parameter.
- (c) Construct diagrams for the pattern of intensity for the two extreme values of this parameter and for a natural choice of parameter intermediate between these two extremes.
- (d) Define the parameter in question in terms of a certain dimensionless combination of the principal moments of the reduced quadrupole tensor.

Exercise 37.13. MULTIMODE DETECTOR

Consider a cylindrical bar of length very long compared to its diameter. Designate the fundamental mode of end-to-end vibration of the bar as " $n = 1$," and call the mode with $n - 1$ nodes in its eigenfunction the " n th mode." Show that the cross section for the interception of unpolarized gravitational waves at the n th resonance, integrated over that resonance, and averaged over direction, is given by the formula [Ruffini and Wheeler (1971b)]

$$\int_{\substack{n\text{th} \\ \text{resonance;}} \atop \text{random}} \sigma(v)dv = \frac{32}{15\pi} \frac{v^2}{c^2} \frac{M}{n^2} \text{ for } n \text{ odd (zero for even } n\text{)}, \quad (37.45)$$

where v is the speed of sound in the bar expressed in the same units as the speed of light, c ; and M is the mass of the bar (geometric units; multiply the righthand side by the factor $G/c = 2.22 \times 10^{-18} \text{ cm}^2 \text{ Hz/g}$ when employing conventional units). Show that this expression gives $\int \sigma dv = 1.0 \times 10^{-21} \text{ cm}^2 \text{ Hz}$ for the lowest mode of Weber's bar. Multimode detectors are (1973) under construction by William Fairbank and William Hamilton, and by David Douglass and John A. Tyson.

Exercise 37.14. CROSS SECTION OF IDEALIZED MODEL OF EARTH FOR ABSORPTION OF GRAVITATIONAL RADIATION

The observed period of quadrupole vibration of the earth is 54 minutes [see, e.g., Bolt (1964) or Press (1965) for survey and bibliography]. To analyze that mode of vibration, with all due allowance for elasticity and the variation of density in the earth, is a major enterprise. Therefore, for a first estimate of the cross section of the earth for the absorption of quadrupole radiation, treat it as a globe of fluid of uniform density held in the shape of a sphere by gravitational forces alone (zero rigidity). Let the surface be displaced from $r = a$ to

$$r = a + a\alpha P_2(\cos \theta), \quad (37.46a)$$

where θ is polar angle measured from the North Pole and α is the fractional elongation of the principal axis. The motion of lowest energy compatible with this change of shape is described by the velocity field

$$\xi^x = -\frac{1}{2}\alpha x, \quad \xi^y = -\frac{1}{2}\alpha y, \quad \xi^z = \alpha z \quad (37.46b)$$

(zero divergence, zero curl).

(a) Show that the sum of the kinetic energy and the gravitational potential energy is

$$E = -(3/5)(M^2/a)(1 - \alpha^2/5) + (3/20)Ma^2\dot{\alpha}^2. \quad (37.46c)$$

(b) Show that the angular frequency of the free quadrupole vibration is

$$\omega = (16\pi/15)^{1/2}\rho^{1/2}. \quad (37.46d)$$

(c) Show that the reduced quadrupole moments are

$$t_{xx} = t_{yy} = -Ma^2\alpha/5, \quad t_{zz} = 2Ma^2\alpha/5. \quad (37.46e)$$

(d) Show that the rate of emission of vibrational energy, averaged over a period, is

$$-\langle dE/dt \rangle = (3/125)M^2a^4\omega^6\alpha_{\text{peak}}^2. \quad (37.46f)$$

(e) Show that the exponential rate of decay of energy by reason of gravitational wave damping, or “gravitational radiation line broadening,” is

$$A_{\text{GW}} = (4/25)Ma^2\omega^4. \quad (37.46g)$$

(f) Show that the resonance integral of the absorption cross section for radiation incident from random directions with random polarization is

$$\int_{\text{resonance}} \langle \sigma(\nu) \rangle d\nu = (\pi/2)\lambda^2 A_{\text{GW}} = (2\pi/25)Ma^2/\lambda^2. \quad (37.46h)$$

(g) Evaluate this resonance integral. Note: This model of a globe of fluid of uniform density would imply for the earth, with average density 5.517 g/cm³, a quadrupole vibration period of 94 min, as compared to the observed 54 min; and a moment of inertia (2/5)Ma² as compared to the observed 0.33Ma². Ruffini and Wheeler (1971b) have estimated correction factors for both effects and give for the final resonance integral $\sim 5 \text{ cm}^2 \text{ Hz}$.

§37.8. NOISY DETECTORS

When the bandwidth of the incoming waves is large compared to the resonance width of the detector, the waves deposit a total energy in the detector given by

$$\begin{aligned} \text{(total energy deposited)} &= \int \sigma \mathcal{F}_\nu d\nu = \mathcal{F}_\nu(\nu = \nu_0) \int \sigma d\nu. \\ [\text{ergs}]^\uparrow &\qquad\qquad\qquad [\text{erg cm}^{-2}\text{Hz}^{-1}]^\uparrow \quad [\text{cm}^2\text{Hz}]^\uparrow \end{aligned}$$

At least, this is so if the detector is wave-dominated (i.e., if $E_{\text{vibration}} \gg kT$ while waves act; i.e., if initial amplitude of oscillation, produced by Brownian forces, is too small to interfere constructively or destructively with the amplitude due to waves).

Unfortunately, all experiments today (1973) are faced with noisy detectors. Nobody has yet found waves so strong, or constructed a detector so sensitive, that the detector is wave-dominated. Consequently, a key experimental task today is to pick a small signal out of large noise. Many techniques for doing this have been developed and used in a variety of fields of physics, as well as in astronomy, psychology and engineering [see, e.g., Davenport and Root (1958), Blackman and Tukey (1959), and

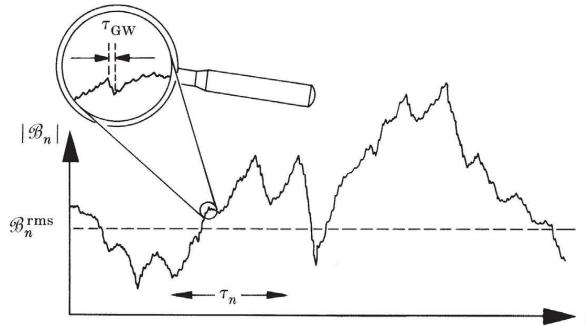


Figure 37.5.

Detection of hammer-blow gravitational waves with a noisy detector. Detection of even a weak pulse is possible if the time of the pulse is short enough. The amplitude \mathcal{B}_n of the detector's vibrations changes by an amount $\sim \mathcal{B}_n^{\text{rms}} (\Delta t / \tau_n)^{1/2}$ during a time interval Δt , due to thermal fluctuations (random-walk, Brownian-noise forces). Depicted in the inset is a change in amplitude produced by a burst of waves of duration τ_{GW} arriving out of phase with the detector's thermal motions (energy extracted by waves!). The waves are detectable because

$$\Delta |\mathcal{B}_n|_{\text{due to waves}} \gg \mathcal{B}_n^{\text{rms}} (\tau_{\text{GW}} / \tau_n)^{1/2},$$

even though $\Delta |\mathcal{B}_n| \ll \mathcal{B}_n^{\text{rms}}$.

references given there]. The key point is always to find some feature of the signal that is statistically more prominent than the same feature of the noise, plus a correlation to show that it arises from the expected signal source and not from elsewhere (“protection from systematic error”). Thus to detect steady gravitational waves from a pulsar, one might seek to define very precisely two numbers $\langle N^2 \rangle$ and $\langle (N + S)^2 \rangle = \langle N^2 \rangle + \langle S^2 \rangle$, where N and S are the noise and signal amplitudes respectively. A long series of observations (with the pulsar out of the antenna beam) gives one value of $\langle N^2 \rangle$. Another equally long series of observations, interspersed with the first series, will be expected in zeroth approximation to give the same value of $\langle N^2 \rangle$. In the next approximation one recognizes and calculates the influence of normal statistical fluctuations. In an illustrative example, theory, confirmed by statistical tests of other parameters drawn from the same data, guarantees that the fluctuations are less than $10^{-5} \langle N^2 \rangle$ with 95 per cent confidence (only 5 per cent chance of exceeding $10^{-5} \langle N^2 \rangle$; this limit is set by time and money, not by absolute limitations of physics). Let the second series of observations be carried out only at times when the pulsar is in the antenna beam. Let it give

$$\{\langle N^2 \rangle + \langle S^2 \rangle\}_{\text{2d series}} = (1 + 7.3 \times 10^{-5}) \{\langle N^2 \rangle\}_{\text{1st series}}.$$

Then in first approximation one can say that $\langle S^2 \rangle$ lies with 95 per cent confidence in the limits $(7.3 \pm 1.0) \times 10^{-5} \langle N^2 \rangle$.

Many conceivable sources of gravitational radiation produce bursts rather than a steady signal strength (Figure 37.5). Thus one is led to ask in what features “hammer-blow radiation” ($\tau_{\text{GW}} \ll \tau_0$) differs from noise. The “Brownian motion” noise in the detector may be thought of as arising from large numbers of small

Rate-of-change of detector amplitude as a tool for extracting burst signals from thermal noise

(molecular) energy exchanges with a heat bath. The calculations below estimate the typical rate of change of amplitude that a series of such molecular “knocks” can produce in a detector, and compare it with the rapid amplitude change produced by a “hammer-blow” pulse of radiation. The calculations show that sudden thermally induced changes, even of very small amplitude, are rare. Thus sudden changes are a suitable feature for the observations to focus on. The actual detection of pulses requires a more extended analysis, however, which goes beyond the estimates made below. Such an analysis would calculate the probabilities that rare events (sudden changes in amplitude) occur by chance (i.e., due to thermal fluctuations) in specified periods of time, the still smaller probabilities that they occur in coincidence between two or more detectors, and the correlations with postulated sources.

Consider a realistic detector of the type described in Box 37.4. But examine it at a time when it is *not* radiation-dominated. Then its motions are being driven by internal Brownian forces (thermal fluctuations), and perhaps also by an occasional burst of gravitational waves. Focus attention on a particular normal mode (mode “ n ”), and describe that mode’s contribution to the vibration of the detector by the vector field

$$\delta\mathbf{x} = \xi = \mathcal{B}_n(t)e^{-i\omega_n t}\mathbf{u}_n(x). \quad (37.47)$$

Description of thermal noise
in resonant detector

Since \mathbf{u}_n is dimensionless with mean value unity ($\int \rho \mathbf{u}_n^2 d^3x = M$), the complex number $\mathcal{B}_n(t)$ is the mass-weighted average of the amplitudes of motion of the detector’s mass elements. This amplitude changes slowly with time (rate $\ll \omega_n$) as a result of driving by Brownian forces; but averaged over time it has a magnitude corresponding to a vibration energy of kT :

$$\langle E_{\text{vibration}} \rangle = 2 \left\langle \frac{1}{2} \int \rho \dot{\xi}^2 d^3x \right\rangle = \frac{1}{2} M \omega_n^2 \langle |\mathcal{B}_n|^2 \rangle = kT; \quad (37.48)$$

i.e.,

$$\mathcal{B}_n^{\text{rms}} \equiv \langle |\mathcal{B}_n|^2 \rangle^{1/2} = (2kT/M\omega_n^2)^{1/2}. \quad (37.49)$$

Example: for Weber’s detector ($M \sim 10^3$ kg, $\omega_0 \sim 10^4$ /sec), the fundamental mode at room temperature has

$$\mathcal{B}_0^{\text{rms}} = \left(\frac{2 \times 1.38 \times 10^{-16} \times 300 \text{ erg}}{10^6 \text{ g} \times 10^8 \text{ sec}^{-2}} \right)^{1/2} = 3 \times 10^{-14} \text{ cm}. \quad (37.50)$$

One’s hope for detecting weak hammer-blow radiation lies not in an examination of the detector’s vibration amplitude (or energy), but in an examination of its rate of change (Figure 37.5). The time-scale for large Brownian fluctuations in amplitude ($|\Delta\mathcal{B}_n| \sim \mathcal{B}_n^{\text{rms}}$), when the detector is noisy, is the same as the time scale τ_n for internal forces to damp the detector, when it is driven to $E_{\text{vibration}} \gg kT$. Thus, *the amplitude \mathcal{B}_n does a “random walk” under the influence of Brownian forces, with the mean time for “large walks” ($|\Delta\mathcal{B}_n| \sim \mathcal{B}_n^{\text{rms}}$) being $\Delta t \approx \tau_n$.* The change in \mathcal{B}_n over shorter times Δt is smaller by the “ $1/\sqrt{N}$ factor,” which always enters into random-walk processes:

$$\sqrt{N} = \left(\frac{\text{number of vibration cycles in time } \tau_n}{\text{number of vibration cycles in time } \Delta t} \right)^{1/2} = \left(\frac{\tau_n}{\Delta t} \right)^{1/2}; \quad (37.51)$$

$$\langle |\Delta \mathcal{B}_n^{(\text{thermal})}| \rangle \approx \mathcal{B}_0^{\text{rms}} \left(\frac{\Delta t}{\tau_n} \right)^{1/2} = \left(\frac{2kT}{M\omega_n^2} \right)^{1/2} \left(\frac{\Delta t}{\tau_n} \right)^{1/2} \text{ during time } \Delta t. \quad (37.52)^*$$

Now suppose that “hammer-blow” radiation (burst of duration $\Delta t = \tau_{\text{GW}} \ll \tau_n$) strikes the detector, producing a change $\Delta \mathcal{B}_n^{(\text{GW})}$ in the detector’s amplitude. This change in amplitude, because it comes so quickly, (1) superposes linearly on any change in amplitude produced in the same time interval by the action of Brownian-motion forces; and (2) is therefore independent in value of the presence or absence of Brownian-motion forces, i.e., independent of all thermal agitation. Therefore $\Delta \mathcal{B}_n^{(\text{GW})}$ (a quantity with *both* magnitude and phase!) is identical to what it would have been if the detector were at zero temperature:

Effect of a burst of waves on
a noisy, resonant detector

$$\underbrace{\frac{1}{2} M\omega_n^2 |\Delta B_n^{(\text{GW})}|^2}_{\text{energy that would be deposited if detector were at zero temperature}} = \int \sigma_n(\nu) \mathcal{F}_\nu(\nu) d\nu = \mathcal{F}_\nu(\omega_n/2\pi) \int \sigma_n(\nu) d\nu;$$

↑
For hammer-blow radiation, bandwidth of radiation is always \gg bandwidth of detector;
see Box 37.4

i.e.,

$$|\Delta \mathcal{B}_n^{(\text{GW})}| = \left(\frac{2\mathcal{F}_\nu(\omega_n/2\pi) \int \sigma_n d\nu}{M\omega_n^2} \right)^{1/2}. \quad (37.53)$$

This wave-induced change in amplitude will be distinguishable from thermal changes only if it is significantly bigger than the thermal changes (37.52) expected during the same length of time τ_{GW} :

$$\begin{aligned} |\Delta \mathcal{B}_n^{(\text{GW})}| &\gg \langle |\Delta \mathcal{B}_n^{(\text{thermal})}| \rangle \text{ during time } \tau_{\text{GW}} \\ \text{equivalently: } F_\nu(\omega_n/2\pi) &\gg \left(\frac{kT}{\int \sigma_n d\nu} \right) \left(\frac{\tau_{\text{GW}}}{\tau_n} \right) \end{aligned} \quad \left. \begin{array}{l} \text{criteria for} \\ \text{detectability} \end{array} \right\} \quad (37.54)$$

Criteria for detectability of burst

Of course, if one is equipped only to measure the magnitude of the detector’s amplitude or energy, and not its phase, these criteria for detectability are not quite sufficient. The wave-induced change in squared amplitude (proportional to change in energy) will depend on the relative phases of the initial amplitude and amplitude change

*For a fuller derivation and discussion of this formula, see, e.g., Braginsky (1970). Two key points covered there are: (1) a statistical version of the formula, which describes the probability that in time Δt the amplitude will change by a given amount, from a given initial value; and (2) quantum-mechanical corrections, which come into play in the limit as $\tau_n \rightarrow \infty$, but which are unimportant for detectors of the early 1970’s.

$$\begin{aligned}
 \Delta|\mathcal{B}_n|^2 &= |\mathcal{B}_n^{(\text{initial})} + \Delta\mathcal{B}_n^{(\text{GW})}|^2 - |\mathcal{B}_n^{(\text{initial})}|^2 \\
 &\approx 2|\mathcal{B}_n^{(\text{initial})}||\Delta\mathcal{B}_n^{(\text{GW})}| && \text{if in phase} \\
 &\approx 0 && \text{if phase difference is } \pm\pi/2 \\
 &\approx -2|\mathcal{B}_n^{(\text{initial})}||\Delta\mathcal{B}_n^{(\text{GW})}| && \text{if phase difference is } \pi.
 \end{aligned} \tag{37.55}$$

Thus, only a burst that arrives in phase with the initial motion of the detector or with reversed phase will be measurable. But for such a burst, the criteria (37.54) are sufficient.

Ways to improve sensitivity of detector

Equations (37.54) make it clear that *there are three ways to improve the sensitivity of vibratory detectors to hammer-blow radiation: (1) increase the detector's integrated cross-section [which can be done only by increasing the rate A_{GW} at which it reradiates gravitational waves; see equations (10b) and (11b) of Box 37.4]; (2) cool the detector; (3) increase the detector's damping time.*

Box 37.5 applies the above detectability criteria to some detectors that seem feasible in the 1970's, and to some bursts of waves predicted by theory. The conclusions of that comparison give one hope!

To be complete, the above discussion should have analyzed not only noise in the detector, but also the noise in the sensor which one uses to measure the amplitude of the detector's displacements. However, the theory of displacement sensors is beyond the scope of this book. For a brief discussion and for references, see Press and Thorne (1972).

§37.9. NON-MECHANICAL DETECTORS

Non-mechanical detectors

When gravitational waves flow through matter, they excite it into motion. Such excitations are the basis for all detectors described thus far. But gravitational waves interact not only with matter; they also interact with electromagnetic fields; and those interactions can also be exploited in detectors. One of the most promising detectors that may be built in the future, one designed by Braginsky and Menskii (1971), relies on a resonant interaction between gravitational waves and electromagnetic waves. It is described in Box 37.6.

§37.10. LOOKING TOWARD THE FUTURE

The future of gravitational-wave astronomy

As this book is being written, it is not at all clear whether the experimental results of Joseph Weber constitute a genuine detection of gravitational waves. (See §37.4, part 4.) But whether they do or not, gravitational-wave astronomy has begun, and seems to have a bright future. The technology of 1973 appears sufficient for the construction of detectors that will register waves from a star that collapses to form a black hole anywhere in our galaxy (Box 37.5); and detectors of the late 1970's and early 1980's may well register waves from pulsars and from supernovae in other galaxies. The technical difficulties to be surmounted in constructing such detectors are enormous. But physicists are ingenious; and with the impetus provided by Joseph Weber's pioneering work, and with the support of a broad lay public sincerely interested in pioneering in science, all obstacles will surely be overcome.

**Box 37.5 DETECTABILITY OF HAMMER-BLOW WAVES
FROM ASTROPHYSICAL SOURCES: TWO EXAMPLES**
(The following calculations are accurate only to
within an order of magnitude or so)

**A. Waves from a Star of Ten Solar Masses Collapsing to Form
a Black Hole; 1972 Detector with 1975 (?) Sensor**

- Predicted characteristics of radiation:

$$\begin{aligned} \text{(intensity at Earth)} &= \mathcal{F}_\nu \sim \frac{M_\odot}{4\pi(\text{distance})^2\nu} \\ &\sim (2 \times 10^5 \text{ ergs/cm}^2 \text{ Hz}) [(\text{distance to center of galaxy})/(\text{distance})]^2, \\ (\text{frequency of waves}) &= \nu \sim 10^3 \text{ Hz}, \\ (\text{bandwidth of waves}) &= \Delta\nu \sim 10^3 \text{ Hz}, \\ (\text{duration of burst}) &= \tau_{\text{GW}} \sim 10^{-3} \text{ sec to } 10^{-1} \text{ sec.} \end{aligned}$$

- Detector properties: A Weber bar, vibrating in its fundamental mode, with

$$\begin{aligned} M &= 10^6 \text{ g}, \quad \int \sigma d\nu = 10^{-21} \text{ cm}^2 \text{ Hz} \text{ (exercise 37.13),} \\ \nu_0 &= \omega_0/2\pi = 1,660 \text{ Hz}, \quad T = 3 \text{ K (liquid Helium temperature),} \\ \tau_0 &= 20 \text{ seconds,} \\ \mathcal{B}_0^{\text{rms}} &= \left(\frac{2 \times 1.37 \times 10^{-16} \times 3 \text{ erg}}{10^6 \text{ g} \times 10^8 \text{ sec}^{-2}} \right)^{1/2} = 3 \times 10^{-15} \text{ cm,} \\ |\Delta\mathcal{B}_0^{(\text{thermal})}| &= (3 \times 10^{-15} \text{ cm})(10^{-3}/20)^{1/2} = 2 \times 10^{-17} \text{ cm,} \\ &\quad \text{during } \Delta t = 10^{-3} \text{ sec,} \\ |\Delta\mathcal{B}_0^{(\text{thermal})}| &= 2 \times 10^{-16} \text{ cm, during } \Delta t = 0.1 \text{ sec.} \end{aligned}$$

- Effect of waves [equation (37.53)]:

$$\begin{aligned} |\Delta\mathcal{B}_0^{(\text{GW})}| &= \left(\frac{2 \times 2 \times 10^5 \times 10^{-21} \text{ ergs}}{10^6 \times 10^8 \text{ sec}^{-2}} \right)^{1/2} \left(\frac{\text{center of Galaxy}}{\text{distance}} \right) \\ &= 2 \times 10^{-15} \text{ cm} \left(\frac{\text{distance to}}{\text{center of Galaxy}} \right). \end{aligned}$$

- Conclusion: Gravitational waves from a massive star collapsing to form a black hole anywhere in our galaxy are readily detectable, if one can construct a “sensor” to measure changes in vibration amplitudes of magnitude $\lesssim 10^{-15}$ cm on time scales < 0.1 seconds. This does appear to be feasible with 1972 technology; see Press and Thorne (1972).

Box 37.5 (continued)

B. Waves from a Supernova Explosion in the Virgo Cluster of Galaxies; a Detector that might be constructable by late 1970's or early 1980's

- Predicted characteristics of radiation:

$$\begin{aligned} \text{(intensity at Earth)} &= F_\nu \sim \frac{0.03 M_\odot}{4\pi(11 \text{ megaparsecs})^2 \nu} \\ &\sim 4 \times 10^{-3} \text{ ergs/cm}^2 \text{ Hz}, \end{aligned}$$

$$\text{(frequency of waves)} = \nu \sim 10^3 \text{ Hz},$$

$$\text{(bandwidth of waves)} \sim \Delta\nu \sim 10^3 \text{ Hz},$$

$$\text{(duration of burst)} = \tau_{\text{GW}} \sim 0.3 \text{ sec, or } \tau_{\text{GW}} \sim 2 \times 10^{-3} \text{ sec.*}$$

- Detector: A Weber-type bar made not of metal, but of a 1,000-kg monocrystal of quartz, cooled to a temperature of 3×10^{-3} K. (For such a monocrystal, it is thought that the damping time would increase in inverse proportion to temperature, $\tau_0 \propto 1/T$) Estimated properties of such a detector:

$$M \sim 10^6 \text{ g}, \quad \int \sigma d\nu = 10^{-21} \text{ cm}^2 \text{ Hz} \text{ (same as for Weber bar),}$$

$$\nu_0 = \omega_0/2\pi \sim 1,500 \text{ Hz}, \quad T = 3 \times 10^{-3} \text{ K},$$

$$\tau_0 \sim 10^6 \text{ sec,}$$

$$\mathcal{B}_0^{\text{rms}} = \left(\frac{2 \times 1.37 \times 10^{-16} \times 3 \times 10^{-3} \text{ erg}}{10^6 \text{ g} \times 10^8 \text{ sec}^{-2}} \right)^{1/2} = 1 \times 10^{-16} \text{ cm,}$$

$$|\Delta\mathcal{B}_0^{\text{(thermal)}}| = (1 \times 10^{-16} \text{ cm}) \left(\frac{0.3 \text{ or } 2 \times 10^{-3}}{10^6} \right)^{1/2} = \begin{cases} 6 \times 10^{-20} \text{ cm,} \\ \text{or} \\ 5 \times 10^{-21} \text{ cm.} \end{cases}$$

- Effect of waves [equation (37.53)]:

$$|\Delta\mathcal{B}_0^{\text{(GW)}}| = \left(\frac{2 \times 4 \times 10^{-3} \times 10^{-21} \text{ ergs}}{10^6 \times 10^8 \text{ sec}^{-2}} \right)^{1/2} = 3 \times 10^{-19} \text{ cm.}$$

- Conclusion: Gravitational waves are detectable from a supernova in the Virgo cluster, if one can construct a sensor to measure changes in vibration amplitudes of magnitude $\lesssim 10^{-19}$ cm on time scales of $\lesssim 0.1$ seconds; and if one can construct a detector with the above characteristics.

*For the duration of waves from a supernova explosion, two time scales appear to be relevant: (1) the time required for the final stages of the collapse of the white-dwarf core to a neutron star or a neutron-star pancake, $\tau \sim (\text{dimensions of neutron star})/(\text{speed of sound in nuclear matter}) \sim 2 \times 10^{-3}$ sec ("pulse of gravitational radiation"); and (2) the time required for a vibrating neutron star to lose its energy of vibration by gravitational radiation ("damped train of waves"), $\tau \sim 0.3$ sec.

Box 37.6 A NONMECHANICAL DETECTOR OF GRAVITATIONAL WAVES
[Braginsky and Menskii (1971)]

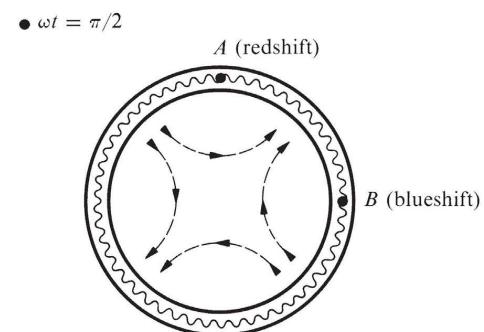
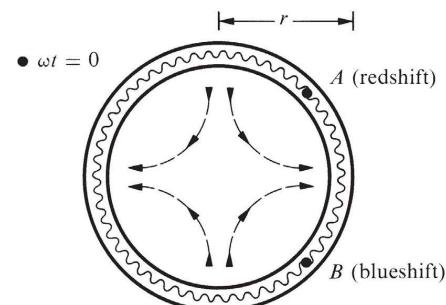
The Idea in Brief

(see diagram at right)

A toroidal waveguide contains a monochromatic train of electromagnetic waves, traveling around and around it. Gravitational waves propagate perpendicular to the plane of the torus. If the circuit time for the EM waves is twice the period of the gravitational waves, then one circularly polarized component of the gravitational waves will stay always in phase with the traveling EM waves. Result: a resonance develops. In one region of the EM wave train, gravitational tidal forces always "push" the waves forward (*blue shift!*) in another region the tidal forces "push" backward (*red shift!*). An EM frequency difference builds up linearly with time; a phase difference builds up quadratically.

ω = (angular frequency of gravitational waves) = (rate of change of phase of waves with time) = (two times angular velocity with which pattern of "lines of force" rotates)

r = (radius of torus), is adjusted so the speed of propagation of EM waves in waveguide is $v = \frac{1}{2}\omega r$.



Outline of Quantitative Analysis

- Let waveguide fall freely in an Earth orbit. Orient axes of waveguide's proper reference frame (\equiv local Lorentz frame) so (1) waveguide lies in \hat{x}, \hat{y} -plane, and (2) gravitational waves propagate in \hat{z} direction.

- Let gravitational waves have amplitudes

$$A_+ - iA_\times = \mathcal{A}e^{-i\omega(t-z)} \quad (1)$$

[Recall: $\hat{t} \approx t$, $\hat{z} \approx z$; i.e., proper frame and TT coordinates almost agree.] Then in plane of waveguide ($z = 0$),

$$\begin{aligned} R_{\hat{x}0}\hat{x}0 &= -R_{\hat{y}0}\hat{y}0 = \frac{1}{2}\omega^2\mathcal{A} \cos(\omega t) \\ R_{\hat{x}0}\hat{y}0 &= R_{\hat{y}0}\hat{x}0 = \frac{1}{2}\omega^2\mathcal{A} \sin(\omega t) \end{aligned} \quad (2)$$

- Consider two neighboring parts of the EM wave, one at $\phi = \alpha + \frac{1}{2}\omega t$; the other at $\phi = \alpha + \delta\alpha + \frac{1}{2}\omega t$. Treat them as photons. Each moves along a null geodesic, except for

[EM waves propagate counterclockwise; gravitational line-of-force diagram rotates counterclockwise; they stay in phase.]

Box 37.6 (continued)

the deflective guidance of the wave guide. Thus, their wave vectors \mathbf{k} satisfy

$$\nabla_{\mathbf{k}} \mathbf{k} = \left(\begin{array}{l} \text{deflective "acceleration"} \\ \text{of waveguide} \end{array} \right); \quad (3)$$

and the difference $\delta\mathbf{k} = \nabla_{\mathbf{n}} \mathbf{k}$ between the wave vectors of the two parts of the wave (difference measured via parallel transport) satisfies the equation

$$\begin{aligned} \nabla_{\mathbf{k}} \delta\mathbf{k} &= \nabla_{\mathbf{k}} \nabla_{\mathbf{n}} \mathbf{k} = [\nabla_{\mathbf{k}}, \nabla_{\mathbf{n}}] \mathbf{k} + \nabla_{\mathbf{n}} \nabla_{\mathbf{k}} \mathbf{k} \quad (4) \\ &= \mathbf{Riemann} (\dots, \mathbf{k}, \mathbf{k}, \mathbf{n}) + \nabla_{\mathbf{n}} \nabla_{\mathbf{k}} \mathbf{k} \end{aligned}$$

[deflective acceleration of wave guide]

The waveguide influences the direction of propagation of the waves, but not their frequency. Thus only **Riemann** enters into the 0 component of the above equation:

$$k^{\hat{\alpha}} \delta k^{\hat{\beta}}_{,\hat{\alpha}} = R^{\hat{\alpha}}_{\hat{\beta}\hat{\gamma}} k^{\hat{\alpha}} k^{\hat{\beta}} n^{\hat{\gamma}}. \quad (5)$$

4. Let $k^{\hat{0}} = \omega_e$ be the angular frequency of the electromagnetic wave. The direction of the space component \mathbf{k} of the propagation 4-vector is along the purely spatial vector \mathbf{n} ; so

$$k^{\hat{0}} = \omega_e, \quad \mathbf{k} = (v\omega_e/r\delta\alpha)\mathbf{n}, \quad n^{\hat{0}} = 0. \quad (6)$$

Use these relations to rewrite equation (5) as

$(d\delta\omega_e/d\hat{t})_{\text{moving with photons}}$

$$= (v\omega_e/r\delta\alpha) R^{\hat{\alpha}}_{\hat{i}\hat{j}} n^{\hat{i}} n^{\hat{j}}. \quad (7)$$

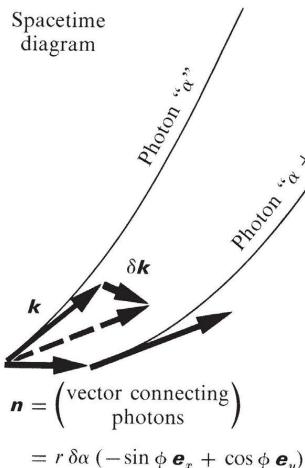
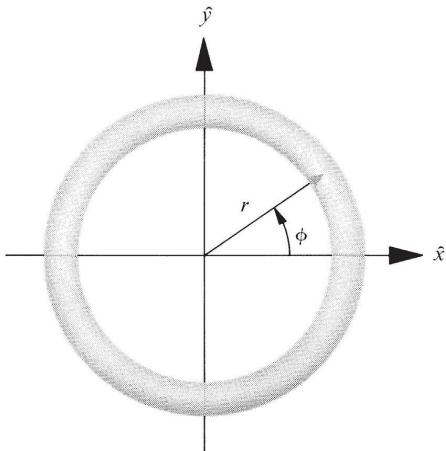
5. Combine the expression for \mathbf{n} in the spacetime diagram with equations (2) and (7), and with the world line $\phi = \alpha + \frac{1}{2}\omega t$ for the photons, to obtain

$(d\delta\omega_e/d\hat{t})_{\text{moving with photons}}$

$$= -\frac{1}{2} v\omega_e \omega^2 dr (\cos 2\alpha) \delta\alpha. \quad (8)$$

6. Integrate over time and over α to obtain

$$\omega_e = \omega_{eo} \left[1 - \frac{1}{4} \mathcal{Q}v (\sin 2\alpha)(\omega r)(\omega \hat{t}) \right]. \quad (9)$$



PART **IX**

EXPERIMENTAL TESTS OF GENERAL RELATIVITY

Wherein the reader is tempted by a harem of charming gravitation theories (and some not so charming), is saved from his foolish passions by an army of experiments, cleaves unto his faithful spouse, Geometrodynamics, vows to lead an honest life hereafter, and becomes a True Believer.

CHAPTER 38

TESTING THE FOUNDATIONS OF RELATIVITY

Provando e riprovando
(Verify the one and disprove the other)

GALILEO

§38.1. TESTING IS EASIER IN THE SOLAR SYSTEM THAN IN REMOTE SPACE

For the first half-century of its life, general relativity was a theorist's paradise, but an experimentalist's hell. No theory was thought more beautiful, and none was more difficult to test.

The situation has changed. In the last few years general relativity has become one of the most exciting and fruitful branches of experimental physics. A half-century late, the march of technology has finally caught up with Einstein's genius—not only on the astronomical front, but also in laboratory experiments.

On the astronomical front, observers search for phenomena in which relativity is important, and study them: cosmology, pulsars, quasars, gravitational waves, black holes. Unfortunately, in pulsars and quasars, and in the sources of cosmological radiation and gravity waves, gravitational effects are tightly interwoven with the local hydrodynamics and local plasma physics. There is little hope of separating the several effects sufficiently sharply to get *clean* tests of the nature of gravity. Instead, astrophysicists must put the laws of gravity into their calculations along with all the other laws of physics and the observational data; and they must then seek, as output, information about the doings of matter and fields “way out there.”

Thus, for clean tests of general relativity one turns to the laboratory—but to a laboratory that is much larger today than formerly: a laboratory that includes the entire solar system.

Clean tests of general relativity are currently confined to solar system

Capabilities of technology in 1970's

In the solar system all relativistic effects are tiny. Nonetheless, some of them are measurable with a precision, in the 1970's, of one part in 1,000 of their whole magnitude or better (see Box 38.1).

§38.2. THEORETICAL FRAMEWORKS FOR ANALYZING TESTS OF GENERAL RELATIVITY

There are now possible many experiments for testing general relativity. But most of them are expensive; very expensive. They involve atomic clocks flown on space-

Box 38.1 TECHNOLOGY OF THE 1970's CONFRONTED WITH RELATIVISTIC PHENOMENA

<i>Quantity to be measured</i>	<i>Magnitude of relativistic effects</i>	<i>Precision of a one-day measurement in the early 1970's</i>
Angular separation of two sources on the sky	Solar deflection of starlight (1) if light ray grazes edge of Sun, $1''.75$ (2) if light ray comes in perpendicular to Earth-sun line, $0''.004$	(a) With optical telescope, $\sim 1''$ (b) Angular separation of two quasars with radio telescope (differential measurement from day to day, not absolute measurement) in 1970, $\sim 0''.1$ in mid 1970's, $\sim 0''.001$
Distance between two bodies in solar system	(a) Perihelion shift per Earth year (1) for Mercury, 120 km (2) for Mars, 15 km (b) Relativistic time delay for radio waves from Earth, past limb of sun, to Venus (one way), $1 \times 10^{-4} \text{ sec} = 30 \text{ km}$ (c) Periodic relativistic effects in Earth-moon separation (1) in general relativity, 100 cm (2) in Jordan-Brans-Dicke theory, 100 cm; $(840 \text{ cm})/(2 + \omega)$	(a) Separation of another planet (Mercury, Venus, Mars) from Earth, by bouncing radar signals off it, $\sim 0.3 \text{ km}$ (b) Separation of a radio transponder (on another planet or in a space craft) from Earth, by measuring round-trip radio travel time, $\sim 3 \times 10^{-8} \text{ sec} = 10 \text{ m} = 0.01 \text{ km}$ (c) Earth-moon separation by laser ranging, $\sim 10 \text{ cm}$
Difference in lapse of proper time between two world lines in solar system	(a) Clock on Earth vs. clock in synchronous Earth orbit, $\Delta t/t \sim 6 \times 10^{-10}$ (b) Clock on Earth vs. clock in orbit about sun, $\Delta t/t \sim 10^{-8}$	Stability of a hydrogen maser clock, $\Delta t/t \sim 10^{-13}$ for t up to one year

craft; radar signals bounced off planets; radio beacons and transponders landed on planets or orbited about them; etc. Because of the expense, it is crucial to have as good a theoretical framework as possible for comparing the relative values of various experiments—and for proposing new ones, which might have been overlooked.

Such a framework must lie outside general relativity. It must scrutinize the foundations of Einstein's theory. It must compare Einstein's theory with other viable theories of gravity to see which experiments can distinguish between them. It must be a "theory of theories."

At present, in 1973, there are two different frameworks in broad use. One, devised largely by Dicke (1964b),* assumes almost nothing about the nature of gravity. It is used to design and discuss experiments for testing, at a very fundamental level, the nature of spacetime and gravity. Within it, one asks such questions as: Do all bodies respond to gravity with the same acceleration? Is space locally isotropic in its intrinsic properties? What are the theoretical implications of local isotropy? What types of fields, if any, are associated with gravity: scalar fields, vector fields, tensor fields, affine fields? Although some of the experiments that tackle these questions will be discussed below, this book will not attempt a detailed exposition of the Dicke framework.

The second framework in broad use is the "parametrized post-Newtonian (PPN) formalism." It has been developed to higher and higher levels of sophistication by Eddington (1922), Robertson (1962), Schiff (1962, 1967), Nordtvedt (1968b, 1969), Will (1971c), and Will and Nordtvedt (1972).

The PPN formalism is an approximation to general relativity, and also to a variety of other contemporary theories of gravity, called "metric theories." It is a good approximation whenever, as in the solar system, the sources of the field gravitate weakly ($|\Phi|/c^2 \ll 1$) and move slowly ($v^2/c^2 \ll 1$). The PPN formalism contains a set of ten parameters whose values differ from one theory to another. Solar-system experiments (measurements of perihelion shift, light deflection, etc.) can be regarded as attempts to measure some of these PPN parameters, and thereby to determine which metric theory of gravity is correct—general relativity, Brans-Dicke (1961)-Jordan (1959) theory, one of Bergmann's (1968) scalar-tensor theories, one of Nordström's theories, Whitehead's (1922) theory, or something else. [For reviews of Nordström and Whitehead, see Whitrow and Morduch (1965), Will (1971b), and Ni (1972). For a significant nonmetric theory, see Cartan (1920) and Trautman (1972).]

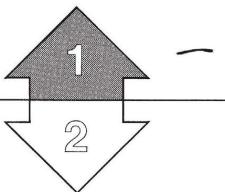
Chapter 39 will discuss the concept of a metric theory of gravity and will construct the PPN formalism; and then Chapter 40 will use the PPN formalism to analyze the systematics of the solar system, and to discuss a variety of past and future experiments that distinguish between various metric theories of gravity. But first, as a prelude to those topics, this chapter will examine experiments that test the foundations of general relativity—foundations on which most other metric theories also rest. For a more detailed discussion of most of these experiments, see Dicke (1964b).

Theoretical frameworks for analyzing gravitation experiments:

(1) Dicke framework

(2) PPN framework

*See Thorne and Will (1971), or Will (1972), for expositions of both frameworks and a comparison of them.



The rest of this chapter is Track 2.
No earlier Track-2 material is needed as preparation for it, but Chapter 7 (incompatibility of gravity and special relativity) will be helpful.
This chapter is not needed as preparation for any later chapter, but it will be helpful in Chapters 39 and 40 (other theories; PPN formalism; experimental tests).

Eötvös-Dicke experiment to test uniqueness of free fall

§38.3. TESTS OF THE PRINCIPLE OF THE UNIQUENESS OF FREE FALL: EÖTVÖS-DICKE EXPERIMENT

One fundamental building block common to Einstein's theory of gravity and to almost all other modern theories is the principle of "uniqueness of free fall":* "The world line of a freely falling test body is independent of its composition or structure." By "test body" is meant an electrically neutral body, small enough that (1) its self-gravitational energy, as calculated using standard Newtonian theory, can be neglected compared to its rest mass ($M/R \ll 1$), and (2) the coupling of its multipole moments to inhomogeneities of the gravitational field can be neglected.†

The uniqueness of free fall permits one to regard spacetime as filled with a set of curves, the test-body trajectories, which are unique aside from parametrization. Through each event, along each timelike or null direction in spacetime, there passes one and only one test-body trajectory. Describe these trajectories mathematically: that is a central imperative of any theory of gravity.

When translated into Newtonian language, the uniqueness of free fall states that any two test bodies must fall with the same acceleration in a given external gravitational field. Experimental tests of this principle search for differences in acceleration from one body to another. The most precise experiments to date are of a type devised by Baron Lorand von Eötvös (Box 38.2), redesigned and pushed to much higher precision by the Princeton group of Robert H. Dicke (Box 38.3), and extended with modifications by the Moscow group of Vladimir B. Braginsky. (See Figure 1.6 and Box 1.2 for experimental details.)

These Eötvös-Dicke experiments are "null experiments." They balance the acceleration of one body against the acceleration of another, and look for tiny departures from equilibrium. The reason is simple. Null experiments typically have much higher precision than experiments measuring the value of a nonzero quantity.

Eötvös, Pekar, and Fekete (1922) checked to an accuracy of 5 parts in 10^9 that the Earth imparts the same acceleration to wood, platinum, copper, asbestos, water, magnalium (90% Al, 10% Mg), copper sulphate, and tallow. Renner (1935) checked, to 7 parts in 10^{10} , the Earth's acceleration of platinum, copper, bizmuth, brass, glass, ammonium fluoride, and an alloy of 30% Mg, 70% Cu. Dicke, and later Braginsky, chose to use the sun's gravitational acceleration rather than the Earth's, since the alternation in the direction of the sun's pull every 12 hours lends itself to amplification by resonance. (See Figure 1.6.) Roll, Krotkov, and Dicke (1964) reported an

*R. H. Dicke calls this principle "The weak equivalence principle." We prefer to avoid confusion with the equivalence principle (Chapter 16).

†In general relativity, one often uses an alternative definition of test body, which places no constraint on the self-gravitational energy [abandon condition (1) while retaining (2)]. Such a definition is preferable, in principle, because the theory of matter has not been developed sufficiently to decide whether (and no objective test has ever been proposed to decide whether), gravitational energy at the subnuclear scale is a small fraction, a large fraction, or the entirety of the rest mass. But for present purposes a definition constraining test bodies to have $M/R \ll 1$ is preferable for two reasons. First, most theories of gravity that currently "compete" with Einstein's (a) agree with the principle of uniqueness of free fall when the macroscopic, Newtonian, self-gravitational energy is neglected ($M/R \ll 1$), but (b) violate that principle when macroscopic, Newtonian self-gravitational energy is taken into account. See §40.9 for details. Second, the test bodies used in the Eötvös-Dicke experiment have M/R so small that their macroscopic, Newtonian, self-gravitational energies are, in fact, negligible ($M/R \sim E_{\text{grav}}/M \sim 10^{-27}$).

agreement of 1 part in 10^{11} between the sun's acceleration of aluminum and gold, while Braginsky and Panov (1971) reported agreement to 1 part in 10^{12} for aluminum and platinum.

From this agreement, one can infer the response of neutrons, protons, electrons, virtual electron-positron pairs, nuclear binding energy, and electrostatic energy to the sun's gravity. Gold is 60% neutrons, while aluminum is only 50% neutrons. Therefore even from the 1964 results one could conclude that neutrons and protons must have the same acceleration to within $[0.6 - 0.5 = 0.1]^{-1}$ parts in $10^{11} = 1$ part in 10^{10} . Similarly, electrons must accelerate the same as nucleons to 2 parts in 10^7 ; virtual pairs (being more abundant in gold than in aluminum) must accelerate the same to 1 part in 10^4 ; nuclear binding energy, to 1 part in 10^7 ; and electrostatic energy to 3 parts in 10^9 .

This accuracy of testing gives one confidence in the principle of the *uniqueness of free fall*.

(continued on page 1054)

Theoretical implications of
Eötvös-Dicke experiment

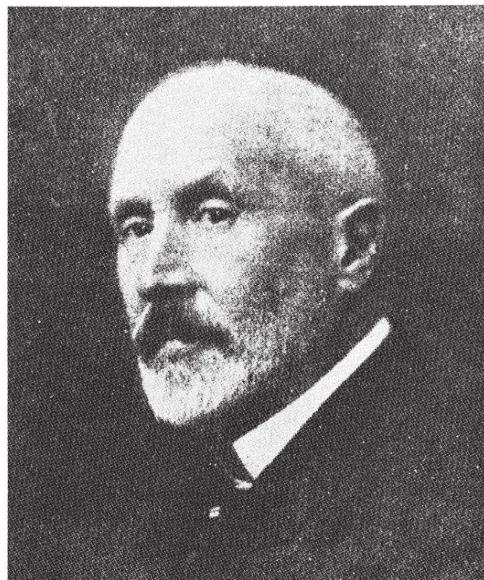
Box 38.2 BARON LORAND VON EÖTVÖS
Budapest, July 27, 1848—Budapest, April 8, 1919

Eötvös (pronounced ut'vûsh) studied at Heidelberg with Kirchhoff, Helmholtz, and Bunsen and at Königsberg with Neumann and Richelot. His 1870 Heidelberg Ph.D. thesis dealt with an issue of relativity: can the motion of a light source relative to an "ether" be detected by comparing the light intensities in the direction of the motion and in the opposite direction?

Studies of his at the same time resulted in the Eötvös law of capillarity, (surface tension) $\approx 2.12 (T_{\text{crit}} - T)/(\text{specific volume})^{2/3}$. Eötvös, made professor of physics at Budapest in 1872, concentrated on gravity from 1886 onward. He developed and extended the original Michell-Cavendish torsion balance, which measured not only $\Phi_{,xx}$ and $\Phi_{,xy}$ (where Φ is the gravitational potential) but also $\Phi_{,xz}$ and $\Phi_{,yz}$, all to a precision destined to be unexcelled for decades. He showed that the so-called "ratio between gravitational mass and inertial mass" cannot vary from material to material by more than 5 parts in 10^9 . He investigated the paleomagnetism of bricks and other ceramic objects, and studied the shape of the earth. He served (June 1894—January 1895) as minister of public instruction and religious affairs (a cabinet position held in earlier years by his father). He founded a school which trained high-school teachers, to whose leavening influence one can give some of the credit for such outstanding scientists as von Karman, von Neuman, Teller, and Wigner. He served one year as rector of the University of Budapest.

"I can never forget the moment when my train rushed into the railroad station of Heidelberg along the banks of the Neckar. . . I cannot forget my happiness that now I could breathe the same air as those men of science whose fame attracted me there."

[EÖTVÖS IN 1887, AS QUOTED IN FEJÉR AND MIKOLA (1918), P. 259.]

Box 38.2 (continued)*Photograph by A. Szekely 1913*

"Insofar as it is permitted on the basis of a few experiments, we can therefore declare that μ , that is, the weakening of the Earth's attraction through the intervening compensator quadrants, is less than one part in 5×10^{10} the absorption (of gravity) by the entire earth along a diameter is less than about one part in 800.

"We have carried out a series of observations which surpassed all previous ones in precision, but in no case could we discover any detectable deviation from the law of proportionality of gravitation and inertia."

[EÖTVÖS, PEKÁR, AND FEKETE (1922).]

"Science shall never find that formula by which its necessary character could be proved. Actually science itself might cease if we were to find the clue to the secret."

[EÖTVÖS, PRESIDENTIAL ADDRESS TO THE HUNGARIAN ACADEMY OF SCIENCES, 1890,
AS QUOTED IN FEJÉR AND MIKOLA (1918), P. 280.]

"We should consider it as one of the most astonishing errors of the present age that so many people listen to the words of pseudoprophets who, in place of the dogmas of religion, offer scientific dogmas with medieval impatience but without historical justification."

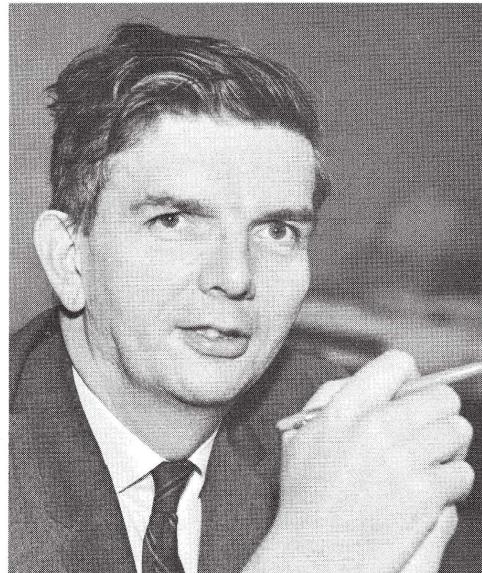
[EÖTVÖS, 1877, AS QUOTED IN FEJÉR AND MIKOLA (1918), P. 280.]

**Box 38.3 ROBERT HENRY DICKE May 6, 1916, St. Louis, Missouri
Cyrus Fogg Brackett Professor of Physics at Princeton University**

During 1941–1946, Dicke was a leader in replacing the outmoded concept of lumped circuit elements by a new microwave analysis based on symmetry considerations, conservation laws, reciprocity relations, and the scattering matrix—concepts that led, among others, to the lock-in amplifier and the microwave radiometer. Searching for a means to reduce the Doppler width of spectral lines for precision measurements, Dicke discovered recoilless radiation in atomic systems held in a box or in a buffer gas. This development led to (1) the discovery of the basic idea of the gas-cell atomic clock and (2) a much more precise measurement of the gyromagnetic ratio of electrons in the 1s and 2s levels of hydrogen and of the hyperfine structure of atomic hydrogen.

A fundamental paper by Dicke in 1954 set forth the theory of coherent radiation processes and of the superradiant state, and laid the foundation for the future development of the laser and the maser, to which he also contributed. His patent no. 2,851,652 (filed May 21, 1956) was the first disclosure of a device for the generation of infrared radiation by a coherent process, and supplied the first suggestion for combining the use of an etalon resonator with an amplifying gas.

Beginning in the 1960's, Dicke brought his talent for precision measurement to the service of experimental cosmology, and with his collaborators: (1)



checked the equivalence principle with the up-to-then unprecedented accuracy of 1 part in 10^{11} ; (2) determined the solar oblateness; and (3) suggested that the primordial cosmic-fireball radiation, a tool for seeing deeper into the past history of the universe than has ever before been possible, should be observable, and therefore should be hunted down and found.

"For want of a better term, a gas which is radiating strongly because of coherence will be called 'superradiant.' . . . As the system radiates it passes to states of lower m with r unchanged—to the 'superradiant' region m ~ 0"

(1954)

"Possibilities are examined for the excitation of optical 'superradiant' states of gas"

(1957)

Box 38.3 (continued)

"A 'gravitational oblateness' of [the sun of] 5×10^{-5} would require the abandonment of Einstein's purely geometrical theory of gravitation. . . Such a flattening [of the sun] could be understood as the effect of a rather rapidly rotating interior. . . The answer appears to be that in the past, and to this day, the solar corona with its magnetic field has acted as a brake on the surface of the sun"

(1964a)

"New measurements of the solar oblateness have given a value for the fractional difference of equatorial and polar radii of $(5.0 \pm 0.7) \times 10^{-5}$ "

[DICKE AND GOLDENBERG (1967)]

"[The universe must] have aged sufficiently for there to exist elements other than hydrogen. It is well-known that carbon is required to make physicists"

(1961)

"The question of the constancy of such dimensionless numbers is to be settled not by definition but by measurements"

[BRANS AND DICKE (1961)]

"The geophysical data lead to an upper limit of 3 parts in 10^{13} per year on the rate of change of the fine-structure constant"

[DICKE AND PEEBLES (1962)]

§38.4. TESTS FOR THE EXISTENCE OF A METRIC GOVERNING LENGTH AND TIME MEASUREMENTS, AND PARTICLE KINEMATICS

Special relativity, general relativity, and all other metric theories of gravity assume the existence of a metric field and predict that this field determines the rates of ticking of atomic clocks and the lengths of laboratory rods by the familiar relation $-dr^2 = ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta$.

Experimental evidence for existence of a metric

The experimental evidence for a metric comes largely from elementary particle physics. It is of two types: *first*, experiments that measure time intervals directly, e.g., measurements of the time dilation of the decay times of unstable particles;* *second*, experiments that reveal the fundamental role played by the Lorentz group in particle kinematics and elsewhere in particle physics.† To cast out the metric tensor entirely would leave one with no theoretical framework adequate for interpreting such experiments.

*For a 2 per cent test of time dilation with muons of $(1 - v^2)^{-1/2} \sim 12$ in a storage ring, see Farley, Bailey, Brown, Giesch, Jöstlein, van der Meer, Picasso, and Tannenbaum (1966). For earlier time-dilation experiments see Frisch and Smith (1963); Durbin, Loar, and Havens (1952); and Rossi and Hall (1941).

†See p. 18 of Lichtenberg (1965) for a discussion of Lorentz invariance, spin and statistics, the TCP theorem, and relevant experiments.

Notice what particle-physics experiments *do* and *do not* tell one about the metric tensor, \mathbf{g} . *First*, they *do not* guarantee that there exist global Lorentz frames, i.e., coordinate systems extending throughout all of spacetime, in which $g_{\alpha\beta} = \eta_{\alpha\beta}$. However, they *do* suggest that at each event \mathcal{P} there exist orthonormal frames with $\mathbf{e}_\alpha(\mathcal{P}) \cdot \mathbf{e}_\beta(\mathcal{P}) = \eta_{\alpha\beta}$, which are related to each other by Lorentz transformations. These orthonormal frames provide one with a definition of the inner product between any two vectors at a given event—and, thereby, they define the metric field.

Second, particle experiments *do not* guarantee that freely falling particles move along geodesics of the metric field, i.e., along straight lines in local Lorentz frames. (Here, in §§38.4 and 38.5, the phrase “local Lorentz frame” means a “normal” coordinate system at an event \mathcal{P} , in which $g_{\alpha\beta}(\mathcal{P}) = \eta_{\alpha\beta}$ and $g_{\alpha\beta,\gamma}(\mathcal{P}) = 0$. The term “inertial frame” is avoided because no assertions are made, yet, about test-body motion.) In particular, one does not know from elementary-particle experiments whether the local Lorentz frames in the laboratory are freely falling (so they fly up from the center of the earth and then fall back with Newtonian acceleration $g = 980 \text{ cm/sec}^2$), whether they are forever at rest relative to the laboratory walls, or whether they undergo some other type of motion. All one is led to believe is that a metric determines the nature of the spacetime intervals ($d\tau^2 = -g_{\mu\nu} dx^\mu dx^\nu$) measured by atomic clocks, that the various local Lorentz frames in the laboratory therefore move with uniform velocity relative to each other (they are connected by Lorentz transformations), and that electric and magnetic fields and the energies and momenta of particles undergo Lorentz transformations in the passage from one local Lorentz frame to another.

Third, elementary particle experiments *do* suggest that the times measured by atomic clocks depend only on velocity, not on acceleration. The measured squared interval is $ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta$, independently of acceleration (until the acceleration becomes so great it disturbs the structure of the clock; see §16.4 and Box 16.3). Equivalently, but more physically, the time interval measured by a clock moving with velocity v^j relative to a local Lorentz frame is

$$d\tau = (-\eta_{\alpha\beta} dx^\alpha dx^\beta)^{1/2} = [1 - (v^x)^2 - (v^y)^2 - (v^z)^2]^{1/2} dt, \quad (38.1)$$

independently of the clock’s acceleration d^2x^j/dt^2 . If this were not so, then particles moving in circular orbits in strong magnetic fields would exhibit different decay rates than freely moving particles—which they do not [Farley *et al.* (1966)].*

§38.5. TESTS OF GEODESIC MOTION: GRAVITATIONAL REDSHIFT EXPERIMENTS

The uniqueness of free fall, as tested by the Dicke-Eötvös experiments, implies that spacetime is filled with a family of preferred curves, the test-body trajectories. There

Particle experiments do *not* guarantee existence of global Lorentz frames, or geodesic motion for test particles

Particle experiments *do* suggest proper time is independent of acceleration

*The experiment of Farley *et al.* is a 2 percent check of acceleration-independence of the muon decay rate for energies $E/m = (1 - v^2)^{-1/2} \sim 12$ and for accelerations, as measured in the muon rest frame, of $a = 5 \times 10^{20} \text{ cm/sec}^2 = 0.6 \text{ cm}^{-1}$.

Physical meaning of a comparison between test-body trajectories and geodesics of metric

Pound-Rebka-Snider redshift experiment as a test of geodesic motion

is also another family of preferred curves, the *geodesics* of the metric \mathbf{g} . It is tempting to identify these geodesics with the test-body trajectories. Einstein's geometric theory of gravity makes this identification ("equivalence principle"). One might conceive of theories that reject this identification. What is the experimental evidence on this point?

In order to see what kinds of experiments are relevant, it is helpful to elucidate the physical significance of the geodesics.

A geodesic of \mathbf{g} is most readily identified locally by the fact that it is a straight line in the local Lorentz frames. Put differently, a body's motion is unaccelerated as measured in a local Lorentz frame if and only if the body moves along a geodesic of \mathbf{g} . Hence, to determine whether test-body trajectories are geodesics, one must compare experimentally the motion of the spatial origin of a local Lorentz frame (as defined by atomic-clock readings) with the motion of a test body (material particle).

It is easy to study experimentally the motions of test bodies; relative to an earth-bound laboratory, they accelerate downward with $g = 980 \text{ cm/sec}^2$; and this acceleration can be measured at a given location on the Earth to a precision of 1 part in 10^6 .

Unfortunately, it is much more difficult to measure the motion of a local Lorentz frame, once again as defined by atomic-clock readings. The only direct experimental handle one has on this today, with sufficient precision to be interesting, is gravitational redshift experiments. (See §§7.2–7.5 and §25.4 for theoretical discussions of the gravitational redshift in the framework of general relativity.)

The redshift experiment of highest precision is that of Pound and Rebka (1960), as improved by Pound and Snider (1965); see Figure 38.1. It used the Mossbauer effect to measure the redshift of 14.4 keV gamma rays from Fe^{57} . The emitter and absorber of the gamma rays were placed at rest at the bottom and top of a tower at Harvard University, separated by a height $h = 74 \text{ feet} = 22.5 \text{ meters}$. The measured redshift agreed, to 1 percent precision, with the general relativistic prediction of

$$\Delta\lambda/\lambda = gh = 2.5 \times 10^{-15}. \quad (38.2)$$

This result tells one that the local Lorentz frames are not at rest relative to the Earth's surface; rather, they are accelerating downward with the same acceleration, g , as acts on a free particle (to within 1 percent precision). To arrive at this conclusion, one analyzes the experiment in the laboratory reference frame, where everything (the experimental apparatus, the Earth, the Earth's gravitational field) is static. Relative to the laboratory a local Lorentz frame, momentarily at rest, accelerates downward (horizontal accelerations being ruled out by symmetry) with some unknown acceleration a . Equivalently, the laboratory accelerates upward (in $+z$ direction) with acceleration a relative to the local Lorentz frame. Consequently, the spacetime metric in the laboratory frame has the standard form

$$ds^2 = -(1 + 2az) dt^2 + dx^2 + dy^2 + dz^2 + O(|x^j|^2) dx^\alpha dx^\beta, \quad (38.3)$$

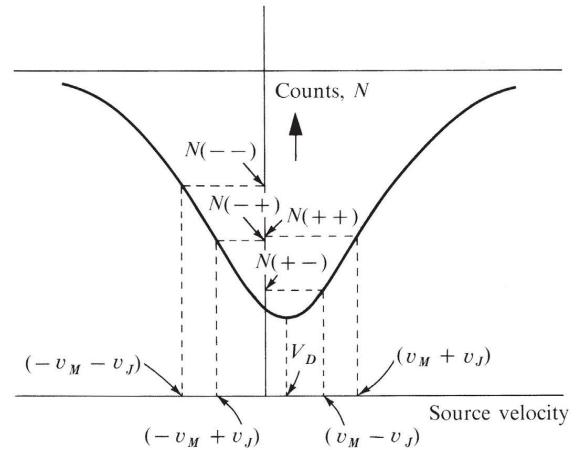
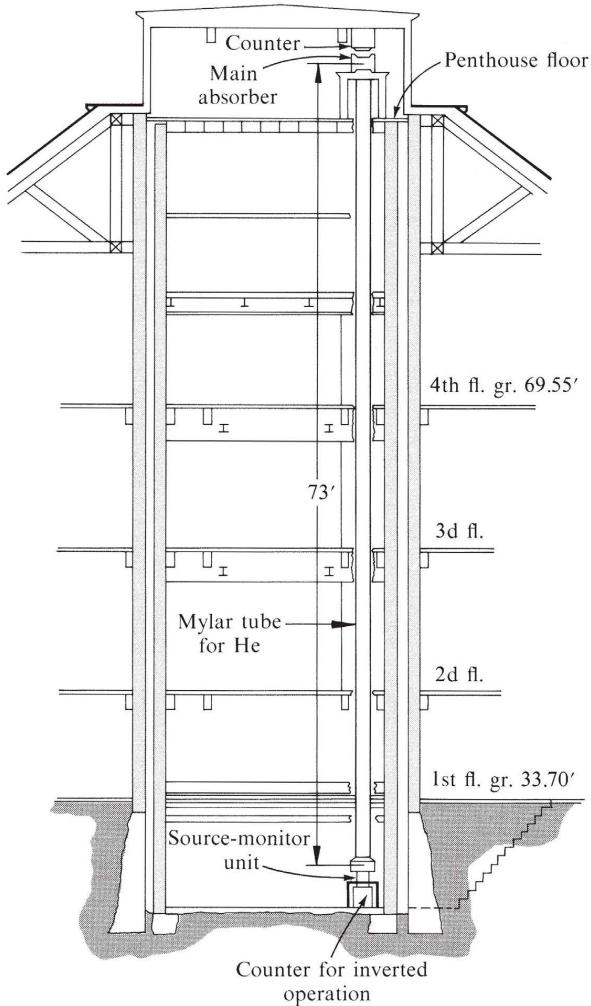


Figure 38.1.

The experiment of Pound and Rebka (1959) and Pound and Snider (1965) on the gravitational redshift of photons rising 22.5 meters against gravity through a helium-filled tube in a shaft in the Jefferson Physical Laboratory of Harvard University. The source of Co^{57} had an initial strength greater than a curie. The 14.4 keV gamma rays had to pass in through an absorber enriched in Fe^{57} to reach the large-window proportional counters. Both source and absorber were placed in temperature-regulated ovens. The velocity of the source consisted of two parts: one steady (v_M), to put the center of the emission line on the part of the transmission curve that is nearly straight; and the other alternating between $+v_J$ and $-v_J$, to sweep the transmission curve in this straight region; similarly when the steady velocity was $-v_M$. The departure from symmetry between the two cases $+v_M$ and $-v_M$ allows one to determine the offset v_D (effect of gravitational redshift) from the zero-gravity case of stationary emitter and stationary absorber. The final result for the redshift was (0.9990 ± 0.0076) times the value 4.905×10^{-15} of $2gh/c^2$ predicted from the principle of equivalence (difference between “up” experiment and “down” experiment). Diagrams adapted from Pound and Snider (1965).

which Track-2 readers have met in §§6.6 and 13.6; and Track-1 readers have met and used in Box 16.2. Moreover, in the laboratory frame the metric is static, gravity is static, and the experimental apparatus is static. Therefore the crest of each electromagnetic wave that climbs upward must follow a world line $t(z)$ identical in form to the world lines of the crests before and after it; thus,

$$\begin{aligned} \text{wave crest } \#0: t &= t_0(z), \\ \text{wave crest } \#1: t &= t_0(z) + \Delta t, \\ &\vdots \\ \text{wave crest } \#n: t &= t_0(z) + n \Delta t. \end{aligned}$$

[Here, as in Schild's argument (§7.3) that redshift implies spacetime curvature, no assumption is made about the form of the wave-crest world lines $t_0(z)$; see Figure 7.1.] Hence, expressed in *coordinate time*, the interval between reception of successive wave crests is the same as the interval between emission. Both are Δt . But the atomic clocks of the experiment (Fe^{57} nuclei) are assumed to measure proper time $\Delta\tau \equiv (-g_{\alpha\beta} \Delta x^\alpha \Delta x^\beta)^{1/2}$, not coordinate time. Thus

$$\begin{aligned} \frac{\lambda_{\text{received}}}{\lambda_{\text{emitted}}} &= \frac{\Delta\tau_{\text{received}}}{\Delta\tau_{\text{emitted}}} = \frac{(1 + az_{\text{received}}) \Delta t}{(1 + az_{\text{emitted}}) \Delta t} \\ &= 1 + a(z_{\text{received}} - z_{\text{emitted}}); \end{aligned}$$

i.e.,

$$\frac{\Delta\lambda}{\lambda} = ah \quad \left[\begin{array}{l} \text{theoretical prediction based on assumptions} \\ (i) \text{ that atomic clocks measure } \Delta\tau = (-g_{\alpha\beta} \Delta x^\alpha \Delta x^\beta)^{1/2}; \\ (ii) \text{ that electromagnetic radiation has the form of a} \\ \text{ wave train;} \\ (iii) \text{ that local Lorentz frames accelerate downward} \\ \text{ with acceleration } a \text{ relative to the laboratory.} \end{array} \right] \quad (38.4)$$

Direct comparison with the experimental result (38.2) reveals that *local Lorentz frames in an Earthbound laboratory accelerate downward with the same acceleration g as acts on a test particle (to within 1 per cent precision)*.

[The above discussion is basically a reworked version of Schild's proof (§7.2) that the redshift experiment implies spacetime is curved. After all, how could spacetime possibly be flat if Lorentz frames in Washington, Moscow, and Peking all accelerate toward the Earth's center with $g = 980 \text{ cm/sec}^2$?]

Other redshift experiments

Of all redshift experiments, the Pound-Rebka-Snider experiment is the easiest to interpret theoretically, because it was performed in a uniform gravitational field. Complementary to it is the experiment by Brault (1962), which measured the redshift of the sodium D₁ line emitted on the surface of the sun and received at Earth (Figure 38.2). To a precision of 5 per cent, he found a redshift of $GM_\odot/R_\odot c^2$, where M_\odot and R_\odot are the mass and radius of the sun. This is just the redshift to be expected if

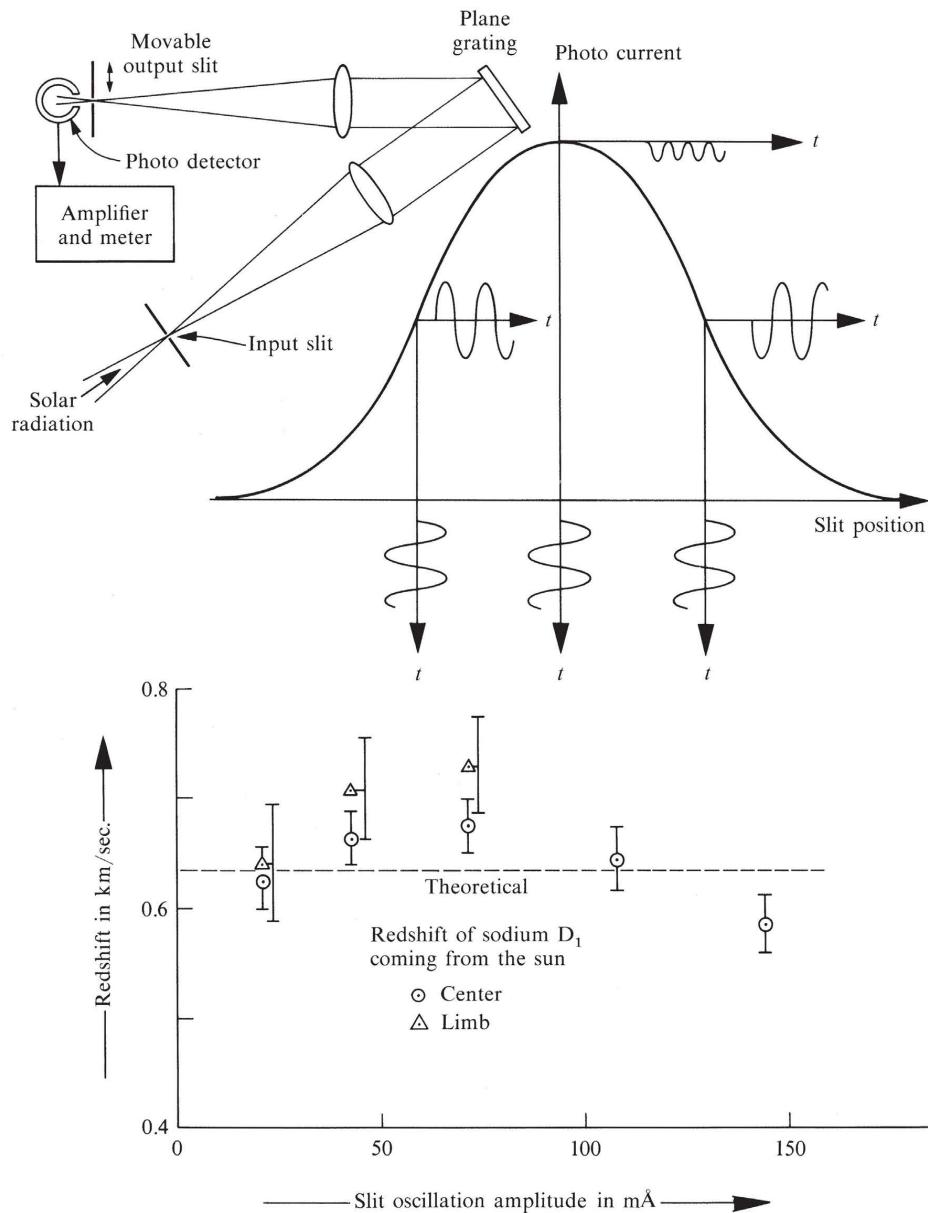


Figure 38.2.

The measurement by Brault (1962) of the redshift of the D_1 line of sodium gives 1.05 ± 0.05 of the gravitational redshift predicted by general relativity. This strong line, in contrast to the weak lines used by earlier investigators (1) is emitted high in the sun's atmosphere, above the regions strongly disturbed by the pressure and convective shifts, and yet lower than the chromosphere, and (2) comes closer to standing up cleanly above the background than any other line in the visible spectrum. Brault built a new photoelectric spectrometer (upper diagram), with its slit vibrated mechanically back and forth across a narrow region of the spectrum, to define the position of the line peak (1) electronically, (2) independently of subjective judgment, and (3) with a precision greater by a factor of the order of ten than that afforded by conventional visual methods. The slit is considered set on a line when its mean position is such that the photomultiplier current contains no signal at the frequency of the modulation. The redshift measured in this way is corrected for orbital motion and for rotation of the sun and the Earth to give the points in circles and triangles in the lower diagram. Extrapolation to zero vibration of the slit gives the cited number for the redshift. Figure adapted from thesis of Brault (1962).

the local Lorentz frames, at each point along the photon trajectory, fall in step with freely falling test bodies.*

In summary, redshift experiments reveal that, to a precision of several percent, the local Lorentz frames at the Earth's surface and near the sun are unaccelerated relative to freely falling test bodies. Equivalently, test bodies move along straight lines in the local Lorentz frames. Equivalently, *the test-body trajectories are geodesics of the metric \mathbf{g} .*

§38.6. TESTS OF THE EQUIVALENCE PRINCIPLE

Tests of the equivalence principle:

(1) geodesic motion

Of all the principles at work in gravitation, none is more central than the equivalence principle. As enunciated in §16.2, it states: “*In any and every local Lorentz frame, anywhere and anytime in the universe, all the (nongravitational) laws of physics must take on their familiar special-relativistic forms.*”

That test bodies move along straight lines in local Lorentz frames (geodesic motion) is one aspect of the equivalence principle. Other aspects are the universality of Maxwell’s equations

$$F^{\alpha\beta}_{,\beta} = 4\pi J^\alpha \text{ and } F_{\alpha\beta,\gamma} + F_{\beta\gamma,\alpha} + F_{\gamma\alpha,\beta} = 0 \quad (38.5)$$

in all local Lorentz frames; the universality of the law of local energy-momentum conservation

$$T^{\alpha\beta}_{,\beta} = 0; \quad (38.6)$$

and the universality of the values of the dimensionless constants that enter into the local laws of physics:

$$\alpha_e \equiv \frac{e^2}{\hbar c} = \frac{1}{137.0360 \dots} = \begin{cases} \text{electromagnetic fine-} \\ \text{structure constant} \end{cases}; \quad (38.7)$$

$$\frac{m_{\text{neutron}}}{m_{\text{proton}}} = 1.00138 \dots, \quad \frac{m_{\text{electron}}}{m_{\text{proton}}} = \frac{1}{1836.12 \dots}, \quad \text{etc.}$$

(Attention here is confined to dimensionless constants, since only they are independent of one’s arbitrary choice of units of measure.)

If one focuses attention on a given event and asks about invariance of the form of the physical laws [equations (38.5), (38.6), etc.] from one Lorentz frame to another, one is then in the province of special relativity. Here a multitude of experiments verify the equivalence principle (see §38.4).

If one asks about variations in the form of the laws from one event to another, one opens up a Pandora’s box of possibilities that one hardly dares to contemplate. However, no experimental evidence has ever given the slightest warrant to consider any such “departure from democracy” in the action of the laws of physics. Moreover, astronomical observations provide strong evidence that the laws of physics are the

(2) physical laws are locally Lorentz-invariant

(3) laws do not vary from event to event

*For a review of other, less-precise redshift experiments, see Bertotti, Brill, and Krotkov (1962).

same in distant stellar systems as in the solar system, and the same in distant galaxies as in our own Galaxy. (See, in Box 29.5, Edwin Hubble's expressions of joy upon discovering this.)

Constancy of the dimensionless "constants" from event to event can be tested to high precision, if one assumes constancy of the physical laws. Dirac (1937, 1938), Teller (1948), Jordan (1955, 1959), Gamow (1967), and others have proposed that the fine-structure "constant" α_e might be a slowly varying scalar field, perhaps governed by a cosmological equation. However, rather stringent limits on such variations follow from data on the fine-structure splitting of the spectral lines of quasars and radio galaxies. For the quasar 3C 191 with redshift $z = 1.95$, Bahcall, Sargent, and Schmidt (1967) find $\alpha_e(3C 191)/\alpha_e(\text{Earth}) = 0.97 \pm 0.5$. With a cosmological interpretation of the quasar redshift, this corresponds to a limit $(1/\alpha_e)(d\alpha_e/dt) \lesssim 1/10^{11}$ years. An even tighter limit has been obtained from radio-galaxy data, where there is no question about the interpretation of the redshift. Bahcall and Schmidt (1967) measured fine-structure splitting in five radio galaxies with $z \approx 0.20$, corresponding to an emission of light 2×10^9 years ago. They obtained $\alpha_e(z = 0.20)/\alpha_e(\text{Earth}) = 1.001 \pm 0.002$, which yields the limit $|(1/\alpha_e)(d\alpha_e/dt)| \lesssim 1/10^{12}$ years.

Dyson (1972) points out that comparison of the rate of beta decay of Re^{187} in times past (via osmium-rhenium abundance ratios in old ores) with the rate of beta-decay today provides a means to check on any possible variation of α_e with time more sensitive than redshift data and more sensitive than any changes in rates of alpha decay and fission between early times and now. He summarizes the available data on Re^{187} and arrives at the limit

$$|(1/\alpha_e)(d\alpha_e/dt)| \lesssim 5/10^{15} \text{ years.}$$

For further evidence of the constancy of the fundamental constants see Minkowski and Wilson (1956), Dicke (1959a,b), Dicke and Peebles (1962b).

Spatial variations of α_e , $m_{\text{neutron}}/m_{\text{proton}}$, and other "constants" in the solar system can be sought by means of Eötvös-type experiments. The reasoning [by Dicke (1969)] leading from such experiments to limits on any spatial variation of the constants is indirect. It recalls the reasoning used in standard treatises on polar molecules to deduce the acceleration of a polarizable molecule pulled on by an inhomogeneous electric field. It proceeds as follows.

Suppose one of the dimensionless "constants," " α ," depends on position. This will lead to a position-dependence of the total mass-energy of a laboratory test body. For example, if α_e depends on position, then the coulomb energy of an atomic nucleus will also ($E_{\text{coul}} \propto e^4 \propto \alpha_e^2$; $\delta M/E_{\text{coul}} = 2 \delta \alpha_e/\alpha_e$). One can calculate the change in a test body's mass-energy when it is moved from x^μ to $x^\mu + \delta x^\mu$ by assuming no change at all in the body's structure during its displacement:

$$\delta M = (\partial M/\partial \alpha)_{\text{fixed structure}} (\partial \alpha/\partial x^\mu) \delta x^\mu. \quad (38.8)$$

After the displacement, a weakening of internal forces (due, e.g., to a decrease of α)

- (4) fundamental constants do not vary from event to event

Eötvös-type experiments as tests for spatial variation of fundamental constants

may cause a change in structure, but that change will be accompanied by a conversion of internal potential energy into internal kinetic energy, which conserves M .

Now consider the following thought experiment [an elaboration of the argument by which Einstein first derived the gravitational redshift (§7.2)]: Take n particles, each with mass-energy μ . Make the particles with a structure such that a negligible fraction of μ is associated with the “constant” of interest, α :

$$(1/\mu)(\partial\mu/\partial\alpha) = 0. \quad (38.9)$$

Place these particles at a height h in a (locally) uniform Newtonian field. Combine them together there, releasing binding energy $E_B(h)$, to form a composite body of mass

$$M = n\mu - E_B(h) \quad (38.10)$$

which depends in a significant manner on the “constant” α ,

$$(1/M)(\partial M/\partial\alpha) \neq 0. \quad (38.11)$$

Lower this body, *and* the released binding energy tied up in a little bag, a distance δh . The total force acting is (in Newtonian language)

$$F = Ma + E_B(h)g. \quad (38.12)$$

Here g is acceleration experienced by the type of mass-energy that is independent of α when it is in free fall. In contrast, “free” fall of the assembled body M is not really free fall, because of the supplementary “polarization force” pulling on this object. Hence the assembled body in “free” fall experiences an acceleration, a , a little different from g . However, the mass that is accelerated is precisely M , and therefore the force required to produce this acceleration is given by the product Ma . The energy gained in lowering the body and the bag is

$$E(\text{down}) = F\delta h = Ma\delta h + E_B(h)g\delta h.$$

Put this energy in the bag.

At $h - \delta h$ use some of the energy from the bag to pull the body apart into its component particles. The energy required is

$$\begin{aligned} E_B(h - \delta h) &= n\mu - M(h - \delta h) = n\mu - M(h) + \frac{\partial M}{\partial\alpha} \frac{d\alpha}{dh} \delta h \\ &= E_B(h) + \frac{\partial M}{\partial\alpha} \frac{d\alpha}{dh} \delta h; \end{aligned}$$

so an energy

$$\begin{aligned} E_{\text{bag}} &= E_B(h) + E(\text{down}) - E_B(h - \delta h) \\ &= \left[Ma + E_B(h)g - \frac{\partial M}{\partial\alpha} \frac{d\alpha}{dh} \right] \delta h \end{aligned} \quad (38.13)$$

is left in the bag. Use this energy to raise the n particles and the bag back up to

height h . Assume total energy conservation, so that there will be no extra energy and no deficit when the n particles and bag have returned to the original state back at height h . This means that E_{bag} must be precisely the right amount of energy to do the raising:

$$E_{\text{bag}} = n\mu g \delta h = [M + E_B(h)]g \delta h. \quad (38.14)$$

Combining expressions (38.13) and (38.14) for E_{bag} , discover that

$$a - g = \frac{1}{M} \frac{\partial M}{\partial \alpha} \frac{d\alpha}{dh}. \quad (38.15)$$

Thus, under the assumption of total energy conservation (no perpetual-motion machines!), a spatial dependence of a physical “constant” α will lead to the anomaly (38.15) in the acceleration of a body whose mass depends on α .

Coulomb energy, which is proportional to α_e^2 , amounts in a gold nucleus to 0.4 per cent of the mass, and to 0.1 per cent in an aluminum nucleus. Hence, a spatial variation in α_e should lead to a fractional difference in the gravitational accelerations of these two nuclei equal to

$$\left| \frac{a_{Au} - a_{Al}}{g} \right| \approx \frac{1}{g} 2 \frac{0.003}{\alpha_e} \frac{d\alpha_e}{dh} \lesssim 1 \times 10^{-11};$$

i.e.,

$$\frac{1}{\alpha_e} \left| \frac{d\alpha_e}{dh} \right| \lesssim 1 \times 10^{-9} g \approx 1 \times 10^{-9} \text{ cm/sec}^2 = 1 \times 10^{-30}/\text{cm}$$

at the Earth due to the sun.

Here use is made of the limit (1×10^{-11}) from Dicke's experiment (§38.3), and the acceleration $g = 0.6 \text{ cm/sec}^2$ due to the sun at Earth.

Notice that this says the gradient of $\ln \alpha_e$ is less than 1×10^{-9} the gradient of the Newtonian potential!

§38.7. TESTS FOR THE EXISTENCE OF UNKNOWN LONG-RANGE FIELDS

Whether or not one accepts the assumption that test bodies move on geodesics of the metric, it remains conceivable that previously unknown long-range fields (fields with “ $1/r$ ” fall-off at large distances) are somehow associated with gravity.

If “new” long-range fields (not metric, not electromagnetic) do exist, waiting to be discovered, then there are two ways by which they could influence matter. First, they could *couple directly to matter*, producing, for example, slight deviations from geodesic motion (deviations smaller than the limits of §38.5), or slight dependences of masses of particles on position (dependences smaller than the limits of §38.6). Second (and harder to detect), they could *couple indirectly to matter* by being mere

Possible existence of new long-range fields associated with gravity

Direct vs. indirect coupling

participants in field equations that determine the geometry of spacetime. This section will describe tests for direct-coupling effects. Theories with fields that couple indirectly will be described in Box 39.1, and tests for such fields will be discussed in Chapter 40.

Experimental limits on
direct-coupling fields:

(1) Hughes-Drever
experiment

(2) ether-drift experiments

Dicke (1964b), using his framework for analyzing tests of gravitation theories (§38.2), has shown that several null experiments place stringent limits on unknown, direct-coupling, long-range fields.

One of these experiments is the “Hughes-Drever Experiment” [Hughes, Robinson, and Beltran-Lopez (1960); Drever (1961)]. It can be thought of as a search for a symmetric second-rank tensor field $h_{\alpha\beta}$ that produces slight deviations of test-body trajectories from geodesics of the metric $g_{\alpha\beta}$. Unless one’s experiments happen to be made in a region of spacetime where $h_{\alpha\beta}$ is a constant multiple of $g_{\alpha\beta}$ (“mere rescaling of all lengths and times by a constant factor”), this tensor field must produce anisotropies in the properties of spacetime—which, in turn, will cause anisotropies in the inertial mass of a nucleon, and in turn will cause in an atomic nucleus relative shifts of degenerate energy levels with different magnetic quantum numbers. The Hughes-Drever experiment places stringent limits on such shifts, and thereby on a possible tensor field $h_{\alpha\beta}$. To quote Dicke (1964, p. 186), “If two [tensor] fields are present with the one strongly anisotropic in a coordinate system chosen to make the other isotropic, the strength of [direct] coupling to one must be only of the order of 10^{-22} that of the other. . . . [Moreover], on the moving Earth with ever-changing velocity, anisotropy would be expected at some season.” From the experiments of Hughes and Drever, then, one concludes that there is not the slightest evidence for the presence of a second tensor field. For further details see Dicke and Peebles (1962a).

Another series of experiments, called “*ether-drift experiments*,” places stringent limits on any unknown, long-range vector field that couples directly to mass-energy. One can imagine such a field of cosmological origin. Being cosmological, the 4-vector would most naturally be expected to point in the same direction as the 4-vector \mathbf{u} of the “cosmological fluid” (identical with the time direction \mathbf{e}_0 of a frame in which the cosmic microwave radiation is isotropic). The 4-vector of the new field would then have spatial components in any other frame. In principle an observer could use them to discern his direction of motion and speed relative to the mean rest frame of the universe. The ether-drift experiments search for effects of such a field.

For example, the experiment of Turner and Hill (1964) searches for a dependence of clock rates on such a vector field, by examining the transverse Doppler shift as a function of direction for an emitter on the rim of a centrifuge and a receiver at its center (Figure 38.3). If there is any effect, it would most naturally be expected to have the form

$$\frac{\left(\begin{array}{l} \text{rate of clock moving relative} \\ \text{to universe with speed } \beta \end{array} \right)}{\left(\begin{array}{l} \text{rate of clock at rest} \\ \text{relative to universe} \end{array} \right)} = 1 + \gamma \beta^2, \quad \gamma \text{ a small constant.} \quad (38.17)$$

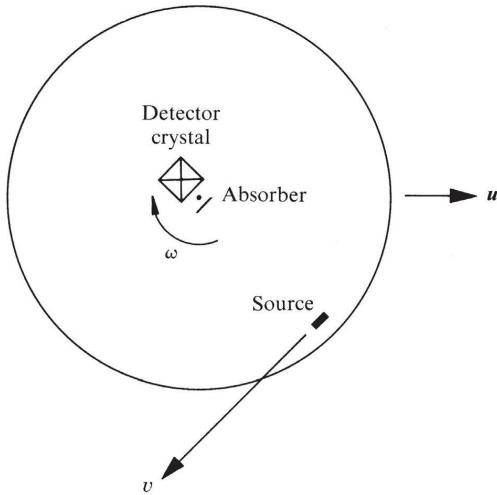


Figure 38.3.

The experiment of Turner and Hill (1964) looks for a dependence of proper clock rate (the clock being a Co^{57} source placed near the rim of the centrifuge) on velocity relative to the distant matter of the universe; or, in operational terms, relative to a “new local field” described by a 4-vector. The 14.4 keV gamma rays from the Co^{57} already experience a second-order Doppler shift of 1.3 parts in 10^{13} . One searches for an additional shift $\gamma\beta^2$ where $\beta = u + v(e_x \cos \omega t + e_y \sin \omega t)$ is the velocity relative to the frame in which the scalar field is purely timelike. The transmission of the gamma rays through the Fe^{57} absorber will drop linearly with any such additional shift, and will be noted as a drop in the counting rate of the NaI crystal. The source was 10 cm from the axis of rotation and the centrifuge turned at 15,000 rpm. The value of γ deduced from the experiment was $(1 \pm 4) \times 10^{-5}$.

A clock at the center of the centrifuge has $\beta = u = ue_x$, whereas one on the rim has $\beta = u + v(e_x \cos \omega t + e_y \sin \omega t)$. Thus, the shift between rim and disk should vary with position

$$\Delta\lambda/\lambda = -\Delta\nu/\nu = -2\gamma uv \cos \omega t + \text{usual transverse shift.}$$

The data of Turner and Hill, using the Mössbauer effect, show that

$$|\gamma| < 4 \times 10^{-5}. \quad (38.18)$$

Hence, a cosmological vector field, if present, has only a weak direct coupling to matter.

For further discussion of these experiments and references on others like them, see Dicke (1964b).

CHAPTER 39

OTHER THEORIES OF GRAVITY AND THE POST-NEWTONIAN APPROXIMATION

§39.1. OTHER THEORIES

Among all bodies of physical law none has ever been found that is simpler or more beautiful than Einstein's geometric theory of gravity (Chapters 16 and 17); nor has any theory of gravity ever been discovered that is more compelling.

As experiment after experiment has been performed, and one theory of gravity after another has fallen by the wayside a victim of the observations, Einstein's theory has stood firm. No purported inconsistency between experiment and Einstein's laws of gravity has ever surmounted the test of time.

Query: Why then bother to examine alternative theories of gravity? *Reply:* To have "foils" against which to test Einstein's theory.

To say that Einstein's geometrodynamics is "battle-tested" is to say it has won every time it has been tried against a theory that makes a different prediction. How then does one select new antagonists for decisive new trials by combat?

Not all theories of gravity are created equal. Very few, among the multitude in the literature, are sufficiently viable to be worth comparison with general relativity or with future experiments. The "worthy" theories are those which satisfy *three criteria for viability: self-consistency, completeness, and agreement with past experiment.*

Role of alternative gravitation theories as foils for experimental tests

Criteria for viability of a theory:

(1) self-consistency

Self-consistency is best illustrated by describing several theories that fail this test. The classic example of an internally inconsistent theory is the spin-two field theory of gravity [Fierz and Pauli (1939); Box 7.1 here], which is equivalent to linearized general relativity (Chapter 18). The field equations of the spin-two theory imply that all gravitating bodies move along straight lines in global Lorentz reference frames, whereas the equations of motion of the theory insist that gravity deflects

bodies away from straight-line motion. (When one tries to remedy this inconsistency, one finds oneself being “bootstrapped” up to general relativity; see route 5 of Box 17.2.) Another self-inconsistent theory is that of Kustaanheimo (1966). It predicts zero gravitational redshift when the wave version of light (Maxwell theory) is used, and nonzero redshift when the particle version (photon) is used.

Completeness: To be complete a theory of gravity must be capable of analyzing from “first principles” the outcome of every experiment of interest. It must therefore mesh with and incorporate a consistent set of laws for electromagnetism, quantum mechanics, and all other physics. No theory is complete if it *postulates* that atomic clocks measure the “interval” $d\tau = (-g_{\alpha\beta} dx^\alpha dx^\beta)^{1/2}$ constructed from a particular metric. Atomic clocks are complex systems whose behavior must be calculated from the fundamental laws of quantum theory and electromagnetism. No theory is complete if it *postulates* that planets move on geodesics. Planets are complex systems whose motion must be calculated from fundamental laws for the response of stressed matter to gravity. For further discussion see §§16.4, 20.6, and 40.9.

(2) completeness

Agreement with past experiment: The necessity that a theory agree, to within several standard deviations, with the “four standard tests” (gravitational redshift, perihelion shift, electromagnetic-wave deflection, and radar time-delay) is obvious. Equally obvious but often forgotten is the need to agree with the expansion of the universe (historically the ace among all aces of general relativity) and with observations at the more everyday, Newtonian level. Example: Birkhoff’s (1943) theory predicts the same redshift, perihelion shift, deflection, and time-delay as general relativity. But it requires that the pressure inside gravitating bodies equal the total density of mass-energy, $p = \rho$; and, as a consequence, it demands that sound waves travel with the speed of light. Of course, this prediction disagrees violently with experiment. Therefore, Birkhoff’s theory is not viable. Another example: Whitehead’s (1922) theory of gravity was long considered a viable alternative to Einstein’s theory, because it makes exactly the same prediction as Einstein for the “four standard tests.” Not until the work of Will (1971b) was it realized that Whitehead’s theory predicts a time-dependence for the ebb and flow of ocean tides that is completely contradicted by everyday experience (see §40.8).

(3) agreement with past experiment

§39.2. METRIC THEORIES OF GRAVITY

Two lines of argument narrow attention to a restricted class of gravitation theories, called *metric theories*.

Why attention focuses on metric theories of gravity

The first line of argument constitutes the theme of the preceding chapter. It examined experiment after experiment, and reached two conclusions: (1) *spacetime possesses a metric*; and (2) *that metric satisfies the equivalence principle* (the standard special relativistic laws of physics are valid in each local Lorentz frame). *Theories of gravity that incorporate these two principles are called metric theories.** In brief, Chapter 38 says, “For any adequate description of gravity, look to a metric theory.”

*For a slightly narrower definition of metric theories, see Thorne and Will (1971).

Exception: Cartan's (1922b, 1923) theory ["general relativity plus torsion"; see Trautman (1972)] is nonmetric, but agrees with experiment and is experimentally indistinguishable from general relativity with the technology of the 1970's.

The second line of argument pointing to metric theories begins with the issue of completeness (preceding section). To be complete, a theory must incorporate a self-consistent version of all the nongravitational laws of physics. No one has found a way to incorporate the rest of physics with ease except to introduce a metric, and then invoke the principle of equivalence. Other approaches lead to dismaying complexity, and usually to failure of the theory on one of the three counts of self-consistency, completeness, and agreement with past experiment. *All the theories known to be viable in 1973 are metric*, except Cartan's. [See Ni (1972b); Will (1972).]

How metric theories differ

In only one significant way do metric theories of gravity differ from each other: their laws for the generation of the metric. In general relativity theory, the metric is generated directly by the stress-energy of matter and of nongravitational fields. In Dicke-Brans-Jordan theory (Box 39.1, p. 1070), matter and nongravitational fields generate a scalar field ϕ ; then ϕ acts together with the matter and other fields to generate the metric. Expressed in the language of §38.7, ϕ is a "new long-range field" that couples indirectly to matter. As another example, a theory devised by Ni (1970, 1972) (Box 39.1) possesses a flat-space metric η and a universal time coordinate t ("prior geometry"; see §17.6); η acts together with matter and nongravitational fields to generate a scalar field ϕ ; and then η , t , and ϕ combine to create the physical metric g that enters into the equivalence principle.

All three of the above theories—Einstein, Dicke-Brans-Jordan, Ni—were viable in the summer of 1971, when this section was written. But in autumn 1971 Ni's theory, and many other theories that had been regarded as viable, were proved by Nordtvedt and Will (1972) to disagree with experiment. This is an example of the rapidity of current progress in experimental tests of gravitation theory!

Henceforth, in this chapter and the next, attention will be confined to metric theories of gravity and their comparison with experiment.

§39.3. POST-NEWTONIAN LIMIT AND PPN FORMALISM

Weak-field, slow-motion expansion of a metric theory

The solar system, where experiments to distinguish between metric theories are performed, has weak gravity,

$$|\Phi| = |\text{Newtonian potential}| \lesssim 10^{-6}; \quad (39.1a)$$

moreover, the matter that generates solar-system gravity moves slowly

$$v^2 = (\text{velocity relative to solar-system center of mass})^2 \lesssim 10^{-7} \quad (39.1b)$$

and has small stress and internal energies

$$|T_{jk}|/\rho_o = (\text{stress divided by baryon "mass" density}) \lesssim 10^{-6}, \quad (39.1c)$$

$$\Pi = (\rho - \rho_o)/\rho_o = \left(\frac{\text{internal energy density per unit baryon "mass" density}}{\text{unit baryon "mass" density}} \right) \lesssim 10^{-6}. \quad (39.1d)$$

[Here the *baryon “mass” density* ρ_o , despite its name, and despite the fact it is sometimes even more misleadingly called “density of rest mass-energy,” is actually a measure of the number density of baryons n , and nothing more. It is defined as the product of n with some standard figure for the mass per baryon, μ_0 , in some well-defined standard state; thus,

$$\rho_o \equiv n\mu_0.] \quad (39.1e)$$

Consequently, the analysis of solar-system experiments using any metric theory of gravity can be simplified, without significant loss of accuracy, by a simultaneous expansion in the small parameters $|\Phi|$, v^2 , $|T_{jk}|/\rho_o$, and Π . Such a “weak-field, slow-motion expansion” gives: (1) flat, empty spacetime in “zero order”; (2) the Newtonian treatment of the solar system in “first order”; and (3) post-Newtonian corrections to the Newtonian treatment in “second order”.

The formalism of Newtonian theory plus post-Newtonian corrections is called the “*post-Newtonian approximation*.” Each metric theory has its own post-Newtonian approximation. Despite the great differences between metric theories themselves, their post-Newtonian approximations are very similar. They are so similar, in fact, that one can construct a single post-Newtonian theory of gravity, devoid of any reference to indirectly coupling fields (ϕ in Dicke-Brans-Jordan; η , t , and ϕ in Ni; see Box 39.1), that contains the post-Newtonian approximation of every conceivable metric theory as a special case. This all-inclusive post-Newtonian theory is called the “*Parametrized Post-Newtonian (PPN) Formalism*.” It contains a set of parameters (called “*PPN parameters*”) that can be specified arbitrarily. One set of values for these parameters makes the PPN formalism identical to the post-Newtonian limit of general relativity; another set of values makes it the post-Newtonian limit of Dicke-Brans-Jordan theory, etc.

Subsequent sections of this chapter present a version of the PPN formalism devised by Clifford M. Will and Kenneth Nordtvedt, Jr. (1972). [See also Will (1972).] This version, containing ten PPN parameters, encompasses as special cases nearly every metric theory of gravity known to the authors. The few exceptions [Whitehead (1922) and theories reviewed by Will (1973)] all disagree with experiment. One can include them in the PPN formalism by adding additional terms and parameters.

The ten parameters are described heuristically in Box 39.2, for the convenience of readers who would skip the full details of the formalism (§§39.4–39.12).

How accurate is the PPN formalism? Or, stated more precisely, how accurately does the post-Newtonian approximation agree with the metric theory from which it comes? In the solar system, where $|\Phi|$, v^2 , $|T_{jk}|/\rho_o$, and Π are all $\lesssim 10^{-6}$, the post-Newtonian approximation makes fractional errors of $\lesssim 10^{-6}$ in quantities of post-Newtonian order, and fractional errors of $\lesssim 10^{-12}$ in quantities of Newtonian order. For example, it misrepresents the deflection of light by $\lesssim 10^{-6} \times$ (post-Newtonian deflection) $\sim 10^{-6}$ seconds of arc. And it ignores relativistic deformations of the Earth’s orbit of magnitude $< 10^{-12} \times$ (one astronomical unit) ~ 10 centimeters. Clearly, there is no need in the 1970’s to use higher-order corrections to the post-Newtonian approximation; and hence no need to construct a “parametrized post-post-Newtonian framework.” However, in the words of Shapiro (1971b): “If one projects from the achievements in the last decade, it is not unreasonable to predict

Post-Newtonian approximation

PPN formalism

Accuracy of PPN formalism in solar system

Box 39.1 THE THEORIES OF DICKE-BRANS-JORDAN AND OF NI
A. Dicke-Brans-Jordan

References: Brans and Dicke (1961); Jordan (1959). [Notes: This is the special case $\eta = -1$ of Jordan's theory. An alternative mathematical representation of the theory is given by Dicke (1962).]

Fields associated with gravity:

ϕ , a long-range scalar field;

\mathbf{g} , the metric of spacetime (from which are constructed the covariant derivative ∇ and the curvature tensors, in the usual manner).

Equations by which these fields are determined:

The trace of the stress-energy tensor generates ϕ via the curved-spacetime wave equation

$$\square\phi = \phi_{,\alpha}^{\alpha} = \frac{8\pi}{3 + 2\omega} T,$$

where ω is the dimensionless "Dicke coupling constant."

The stress-energy tensor and ϕ together generate the metric (i.e., the spacetime curvature) via the field equations

$$G_{\alpha\beta} = \frac{8\pi}{\phi} T_{\alpha\beta} + \frac{\omega}{\phi^2} \left(\phi_{,\alpha}\phi_{,\beta} - \frac{1}{2} g_{\alpha\beta}\phi_{,\mu}\phi^{\mu} \right) + \frac{1}{\phi} (\phi_{;\alpha\beta} - g_{\alpha\beta}\square\phi),$$

where $G_{\alpha\beta}$ is the Einstein tensor.

Variational principle for these equations:

$$\delta \int [\phi R - \omega(\phi_{,\alpha}\phi^{\alpha}/\phi) + 16\pi L](-g)^{1/2} d^4x = 0,$$

where R is the scalar curvature and L is the matter Lagrangian.

Equivalence principle is satisfied:

The special-relativistic laws of physics are valid, without change, in the local Lorentz frames of the metric \mathbf{g} .

Consequence: the scalar field does not exert any direct influence on matter; its only role is that of participant in the field equations that determine the geometry of spacetime. It is an "indirectly coupling field" in the sense of §38.7.

This theory is self-consistent, complete, and for $\omega > 5$ in "reasonable" accord (two standard deviations or better) with all pre-1973 experiments.

B. Ni

References: Ni (1970, 1972)

Fields associated with gravity:

η , a flat "background metric" ("prior geometry" in sense of §17.6). There exist,

by assumption, coordinate systems (“background Lorentz frames”) in which everywhere at once $\eta_{00} = -1$, $\eta_{0j} = 0$, and $\eta_{jk} = \delta_{jk}$.
 t , a scalar field called the “universal time coordinate” (“prior geometry” in sense of §17.6), which is so “tuned” to the background metric that

$$t_{|\alpha\beta} = 0, \quad t_{,\alpha} t_{,\beta} \eta^{\alpha\beta} = -1,$$

where “ $|$ ” denotes covariant derivative with respect to η .

This means there exists a background Lorentz frame (the “rest frame of the universe”) in which $x^0 = t$.

ϕ , a scalar field called the “scalar gravitational field”.
 \mathbf{g} , the metric of spacetime (from which are constructed the covariant derivative ∇ and the curvature tensors, in the usual manner).

Equations by which these fields are determined:

The stress-energy of spacetime generates the scalar gravitational field ϕ via the wave equation

$$\begin{aligned} \square\phi \equiv \phi^{\alpha}_{;\alpha} &= -2\pi T^{\alpha\beta} \partial g_{\alpha\beta} / \partial \phi \\ &= 4\pi T^{\alpha\beta} [\eta_{\alpha\beta} e^{-2\phi} + (e^{2\phi} + e^{-2\phi}) t_{,\alpha} t_{,\beta}]. \end{aligned}$$

ϕ , η , and t together determine the metric of spacetime through the algebraic relation

$$\mathbf{g} = e^{-2\phi} \eta + (e^{-2\phi} - e^{2\phi}) \mathbf{dt} \otimes \mathbf{dt}.$$

Note: In the “rest frame of the universe” that is presupposed in this theory, this metric reduces to

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta = -e^{2\phi} dt^2 + e^{-2\phi}(dx^2 + dy^2 + dz^2).$$

Variational principle for the field equation for ϕ :

$$\delta \int (-2\phi^{\alpha}_{;\alpha} + 16\pi L)(-\mathbf{g})^{1/2} d^4x = 0,$$

where L is the matter Lagrangian.

Equivalence principle is satisfied:

The special-relativistic laws of physics are valid, without change, in the local Lorentz frames of the spacetime metric \mathbf{g} .

Consequence: ϕ , η , and t do not exert any direct influence on matter; they are “indirectly coupling fields” in the sense of §38.7.

This theory is self-consistent and complete. If the solar system were at rest in the “rest frame of the universe”, the theory would agree with all experiments to date—except, possibly, the expansion of the universe. But the motion of the solar system through the universe leads to serious disagreement with experiment (Will and Nordtvedt 1972; §40.8).

Box 39.2 HEURISTIC DESCRIPTION OF THE TEN PPN PARAMETERS

Para-meter	<i>What it measures, relative to general relativity^a</i>	<i>Value in General Relativity</i>	<i>Value in Dicke-Brans-Jordan Theory^b</i>	<i>Value in Ni's Theory^b</i>
γ	How much space curvature (g_{jk}) is produced by unit rest mass?	1	$\frac{1+\omega}{2+\omega}$	1
β	How much nonlinearity is there in the superposition law for gravity (g_{00})?	1	1	1
β_1	How much gravity (g_{00}) is produced by unit kinetic energy ($\frac{1}{2}\rho_o v^2$)?	1	$\frac{3+2\omega}{4+2\omega}$	1
β_2	How much gravity (g_{00}) is produced by unit gravitational potential energy ($\rho_o U$)?	1	$\frac{1+2\omega}{4+2\omega}$	1
β_3	How much gravity (g_{00}) is produced by unit internal energy ($\rho_o \Pi$)?	1	1	1
β_4	How much gravity (g_{00}) is produced by unit pressure (p)?	1	$\frac{1+\omega}{2+\omega}$	1
ξ	How much <i>more</i> gravity (g_{00}) is produced by radial kinetic energy [$\frac{1}{2}\rho_o(v \cdot \hat{r})^2$]—i.e., kinetic energy of motion toward observer—than by transverse kinetic energy?	0	0	0
η	How much <i>more</i> gravity (g_{00}) is produced by radial stress [$\hat{r} \cdot t \cdot \hat{r}$] than by transverse stress?	0	0	0
Δ_1	How much dragging of inertial frames (g_{0j}) is produced by unit momentum ($\rho_o v$)?	1	$\frac{10+7\omega}{14+7\omega}$	$-\frac{1}{7}$
Δ_2	How much easier is it for momentum ($\rho_o v$) to drag inertial frames radially (toward the observer) than in a transverse direction?	1	1	1

^aThese heuristic descriptions are based on equations (39.23).^bFor expositions of these theories see Box 39.1. For derivation of their PPN values and of PPN values for other theories, see Ni (1972).

that in the 1980's techniques will be available to detect second-order effects of general relativity. At that point the ratio of theoretical to experimental relativists may take a sharp turn downwards."

Actually, there are a few exceptions to the claim that the post-Newtonian approximation suffices for the 1970's. These exceptions occur where the external universe impinges on and influences the solar system. For example, gravitational waves propagating into the solar system from distant sources (Chapters 35–37) are ignored by every post-Newtonian approximation and by the PPN framework. They must be treated using a full metric theory or a weak-field, "fast-motion" approximation

to such a theory. Similarly, time-dependence of the “gravitational constant” (§40.8), induced in some theories by expansion of the universe, is beyond the scope of the PPN formalism, as is the expansion itself.

The PPN formalism is used not only in interpreting experimental tests of gravitation theories, but also as a powerful tool in theoretical astrophysics. By specializing all the PPN parameters to unity, except $\xi = \eta = 0$, one obtains the post-Newtonian approximation to Einstein’s theory of gravity. This post-Newtonian approximation can then be used (and has been used extensively) to calculate general relativistic corrections to such phenomena as the structure and stability of stars.*

Applications of PPN formalism to astrophysics

Historical and Notational Notes

The earliest parametrizations of the post-Newtonian approximation were performed, and used in interpreting solar system experiments, by Eddington (1922), Robertson (1962), and Schiff (1962, 1967). However, they dealt solely with the vacuum gravitational field outside an isolated, spherical body (the sun). Nordtvedt (1968b, 1969) devised the first full PPN formalism, capable of treating all aspects of the solar system; he treated the sun, planets, and moon as made from “gases” of point-particles (atoms) that interact gravitationally and electromagnetically. Will (1971c) later used techniques devised by Chandrasekhar (1965a) to modify Nordtvedt’s formalism, so that it employs a stressed, continuous-matter description of celestial bodies. The version of the formalism presented here, devised by Will and Nordtvedt (1972), generalizes all previous versions to acquire “post-Galilean invariance” [see Chandrasekhar and Contopoulos (1967)]. The most detailed and up-to-date review article on the PPN formalism is Will (1972).

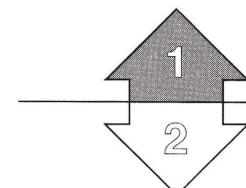
History and notation of PPN formalism

In the literature of post-Newtonian physics and the PPN formalism, the Newtonian potential is described traditionally not by Φ , but by

$$U \equiv -\Phi \equiv + \int \frac{\rho_o(\mathbf{x}') d^3x'}{|\mathbf{x} - \mathbf{x}'|}. \quad (39.2)$$

To avoid confusion, this chapter and the next will use U , although the rest of the book uses Φ .

Turn now to a detailed, Track-2 exposition of the PPN formalism.



EXPOSITION OF PPN FORMALISM:
Coordinate system

§39.4. PPN COORDINATE SYSTEM

The PPN formalism covers the solar system (or whatever system is being analyzed) with coordinates $(t, x_j) \equiv (t, x^j)$ that are as nearly globally Lorentz as possible:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \quad |h_{\alpha\beta}| \lesssim M_\odot/R_\odot \sim 10^{-6}. \quad (39.3)$$

*See, e.g., a long series of papers by Chandrasekhar and his associates in the *Astrophysical Journal*, beginning with Chandrasekhar (1965a,b,c).

The rest of this chapter is Track 2. No earlier Track-2 material is needed as preparation for it, but the following will be helpful:

- (1) Chapter 7 (incompatibility of gravity and special relativity)
- (2) §17.6 (no prior geometry);
- (3) §§36.9–36.11 (generation of gravitational waves); and
- (4) Chapter 38 (tests of foundations).

This chapter is not needed as preparation for any later chapter, but it will be helpful in Chapter 40 (solar-system tests)

(In this sense the PPN formalism is like linearized theory; see Chapter 18.) The velocity of the coordinate system (i.e., 4-velocity of its spatial origin) is so chosen that the solar system is approximately at rest in these coordinates. (Whether the center of mass of the solar system is precisely at rest, or is moving with some low velocity $v \lesssim (M_\odot/R_\odot)^{1/2} \sim 10^{-3} \sim 300$ km/sec, is a matter for the user of the formalism to decide. For more on the options, see §§39.9 and 39.12.)

The PPN coordinates provide one with a natural “3 + 1” split of spacetime into space plus time. That split is conveniently treated using the notation of three-dimensional, flat-space vector analysis—even though spacetime and the three-dimensional hypersurfaces $x^0 = \text{constant}$ are both curved. The resultant three-dimensional formalism will look more like Newtonian theory than like general relativity—as, indeed, one wishes it to; after all, one’s goal is to study small relativistic corrections to Newtonian theory!

§39.5. DESCRIPTION OF THE MATTER IN THE SOLAR SYSTEM

Description of matter

Relative to the PPN coordinates, the matter of the solar system (idealized as a stressed medium) has a coordinate-velocity field

$$v_j \equiv dx_j/dx^0. \quad (39.4)$$

Choose an event \mathcal{P} , and in its neighborhood transform to an orthonormal frame that moves with the matter there. Orient the spatial axes \mathbf{e}_j of this comoving frame so that they coincide as accurately as possible with the PPN coordinate axes. (This requirement will be made more precise in §39.10.) In the orthonormal comoving frame, define the following quantities, which describe the state of the matter:

$$(\text{density of total mass-energy}) \equiv \rho; \quad (39.5a)$$

$$\begin{aligned} &(\text{baryon “mass” density}) \equiv \rho_o \\ &\equiv (\text{number density}) \times (\text{standard rest mass per baryon, } \mu_0,) \\ &\quad (\text{of baryons, } n) \times (\text{for matter in some standard state}); \end{aligned} \quad (39.5b)$$

$$(\text{specific internal energy density}) \equiv \Pi \equiv (\rho - \rho_o)/\rho_o; \quad (39.5c)$$

$$(\text{components of stress tensor}) \equiv t_{ij} \equiv \mathbf{e}_i \cdot \mathbf{T} \cdot \mathbf{e}_j; \quad (39.5d)$$

$$(\text{pressure}) \equiv p \equiv \frac{1}{3} (t_{xx} + t_{yy} + t_{zz}) \quad (39.5e)$$

$$\equiv (\text{average of stress over all directions}).$$

Anisotropies (i.e., shears) in the stress are important only in planets such as the Earth; and even there they are dominated by the isotropic pressure:

$$t_{ij} = p \delta_{ij} + p \times [\text{corrections } \ll 1]. \quad (39.6)$$

For many purposes, especially inside the sun, one can ignore the anisotropies, thereby approximating the solar-system matter as a perfect fluid.*

The isotropic part of the radiation field gives a significant contribution to the pressure, p , and the density of internal energy, $\rho_o \Pi$, inside the sun. However, the anisotropic radiation flux is ignored in the stress-energy tensor. This approximation is allowable because in the sun the outward energy flux carried by radiation is less than 10^{-15} of the internal energy density $\rho_o \Pi$; in planets it is even less.

§39.6. NATURE OF THE POST-NEWTONIAN EXPANSION

For any gravitationally bound configuration such as the solar system, the Newtonian approximation imposes limits on the sizes of various dimensionless physical quantities (see exercise 39.1):

$$\begin{aligned} \epsilon^2 &\equiv \text{maximum value of Newtonian potential } U & (39.7) \\ &\gtrsim \text{values anywhere of } U, v^2, p/\rho_o, |t_{ij}|/\rho_o, \Pi. \end{aligned}$$

Relative magnitudes of expansion parameters

(The Newtonian potential at the center of the sun is $\epsilon^2 \sim 10^{-5}$. The values of p/ρ_o , t_{ij}/ρ_o , and Π there are also $\sim 10^{-5}$, and they are much smaller elsewhere. The orbital velocities of the planets are all less than $100 \text{ km/sec} = 3 \times 10^{-4}$, so $v^2 < 10^{-7}$.) Moreover, changes with time of all quantities at fixed x_j are due primarily to the motion of the matter. As a result, time derivatives are small by $O(\epsilon)$ compared to space derivatives,

$$\left| \frac{\partial A / \partial t}{\partial A / \partial x_j} \right| \sim |v_j| \lesssim \epsilon \text{ for any quantity } A, \quad (39.8)$$

although *not* in the radiation zone, where outgoing gravitational waves flow (distance \gtrsim one light year from Sun). Consequently, the radiation zone must be excluded from the analysis when one makes a post-Newtonian expansion. To treat it requires different techniques, e.g., those of Chapter 36.

Conditions 39.7 and 39.8 suggest that one expand the metric coefficients in powers of the small parameter ϵ , treating U , v^2 , p/ρ_o , t_{ij}/ρ_o , and Π as though they were all of $O(\epsilon^2)$ (often they are smaller!), and treating time derivatives as $O(\epsilon)$ smaller than space derivatives.

In this “post-Newtonian” expansion, terms odd in ϵ (i.e., terms such as

$$\int \frac{\rho_o(\mathbf{x}', t) v_j(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} d^3x' \sim \frac{M}{R} v \sim \epsilon^3 \quad (39.9)$$

whose total number of v 's and $(\partial/\partial t)$'s is odd) change sign under time reversal,

Rules of the expansion

*In the solar system, post-Newtonian corrections due to anisotropic stresses are so much smaller than other post-Newtonian corrections that there is no hope of measuring them in the 1970's. For this reason, elsewhere in the literature (but not in this book) the PPN formalism treats all stresses at the post-Newtonian level as isotropic pressures, thereby setting to zero the PPN parameter η of §§39.8–39.11.

whereas terms even in ϵ do not. Time reversal ($x^{\bar{0}} = -x^0$) also changes the sign of g_{0j} ($g_{\bar{0}j} = -g_{0j}$), but leaves g_{00} and g_{jk} unchanged. Therefore, g_{0j} must contain only terms odd in ϵ ; whereas g_{00} and g_{jk} must contain only even terms. (Actually, this ceases to be the case when radiation damping enters the picture. In the real world one always insists on outgoing-wave boundary conditions. But time reversal converts outgoing waves to ingoing waves; so an extra sign change is required to convert back to out. Therefore, radiation damping terms in the near-zone metric are even in ϵ for g_{0j} , but odd for g_{00} and g_{jk} . However, radiation damping does not come into play until order ϵ^5 beyond Newtonian theory—see Chapter 36—so it will be ignored here.)

Expanded form of metric

The form of the expansion is already known through Newtonian order (see §17.4, with Φ replaced by $-U$): Newtonian gravity is only obtained when one demands that

$$\begin{aligned} g_{00} &= -1 + 2U + [\text{terms } \lesssim \epsilon^4], \\ g_{0j} &= [\text{terms } \lesssim \epsilon^3], \\ g_{ij} &= \delta_{ij} + [\text{terms } \lesssim \epsilon^2]. \end{aligned} \tag{39.10}$$

The stated limits on the higher-order corrections are dictated by demanding that the space components of the geodesic equation agree with the Newtonian equation of motion:

$$\begin{aligned} \frac{d^2x_j}{dt^2} &\approx \frac{d^2x_j}{d\tau^2} = -\Gamma^j_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} \approx -\Gamma^j_{\alpha\beta} \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt} \\ &= -\Gamma^j_{00} - 2\Gamma^j_{0k}v_k - \Gamma^j_{kl}v_kv_l \\ &= U_{,j} + \text{terms of order } \{\epsilon g_{0k,j}; \epsilon^2 g_{kl,j}\}. \end{aligned} \tag{39.11}$$

One would get the wrong Newtonian limit if g_{0k} were $O(\epsilon)$ or greater, and if $g_{kl} - \delta_{kl}$ were $O(1)$ or greater.

The above pattern continues on to all orders in the expansion. Thus in the geodesic equation, and also in the law of local conservation of energy-momentum $T^{\alpha\beta}_{;\beta} = 0$, g_{00} always goes hand-in-hand with ϵg_{0k} and $\epsilon^2 g_{jk}$ (see exercise 39.2). Therefore, the post-Newtonian expansion has the form summarized in Box 39.3.

EXERCISES

Exercise 39.1. ORDERS OF MAGNITUDE IN GRAVITATIONALLY BOUND SYSTEMS

Use Newtonian theory to derive conditions (39.7) for any gravitationally bound system. [Hint: Such concepts as orbital velocities, the speeds of sound and shear waves, the virial theorem, and hydrostatic equilibrium are relevant.]

Exercise 39.2. PATTERN OF TERMS IN POST-NEWTONIAN EXPANSION

Verify the statements in the paragraph following equation (39.11). In particular, suppose that one wishes to evaluate the coordinate acceleration, d^2x_j/dt^2 , to accuracy $\epsilon^{2N}U_{,j}$ for some

integer N . Show that this undertaking requires a knowledge of g_{00} to accuracy ϵ^{2N+2} , of g_{0k} to ϵ^{2N+1} , and of g_{jk} to ϵ^{2N} . Also suppose that one knows T^{00} to accuracy $\rho_o \epsilon^{2N}$, T^{0j} to $\rho_o \epsilon^{2N+1}$, and T^{jk} to $\rho_o \epsilon^{2N+2}$ [see, e.g., equations (39.13) for $N = 0$ and (39.42) for $N = 2$]. Show that to calculate $T^{0\alpha}_{;\alpha}$ with accuracy $\epsilon^{2N+1} \rho_{o,j}$ and $T^{j\alpha}_{;\alpha}$ with accuracy $\epsilon^{2N+2} \rho_{o,j}$, one must know g_{00} to ϵ^{2N+2} , g_{0k} to ϵ^{2N+1} , and g_{jk} to ϵ^{2N} . This dictates the pattern of Box 39.3.

§39.7. NEWTONIAN APPROXIMATION

At Newtonian order the metric has the form (39.10); and the 4-velocity and stress-energy tensor have components, relative to the PPN coordinate system,

$$u^0 = +1 + O(\epsilon^2), \quad u^j = v_j + O(\epsilon^3); \quad (39.12)$$

$$\begin{aligned} T^{00} &= \rho_o + O(\rho_o \epsilon^2), & T^{0j} &= \rho_o v_j + O(\rho_o \epsilon^3), \\ T^{jk} &= t_{jk} + \rho_o v_j v_k + O(\rho_o \epsilon^4) \end{aligned} \quad (39.13)$$

(see exercise 39.3). Two sets of equations govern the structure and evolution of the solar system. (1) The Einstein field equations. As was shown in §18.4, and also in §17.4, in the Newtonian limit Einstein's equations reduce to Laplace's equation

$$U_{,jj} = -4\pi\rho_o, \quad (39.14a)$$

which has the “action-at-a-distance” solution

$$U(\mathbf{x}, t) = \int \frac{\rho_o(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} d^3x'. \quad (39.14b)$$

Box 39.3 POST-NEWTONIAN EXPANSION OF THE METRIC COEFFICIENTS

<i>Level of approximation (and papers expanding general relativity to this level)</i>	<i>Order or value of terms</i>		
	g_{00}	g_{0j}	g_{jk}
flat, empty spacetime	-1	0	δ_{jk}
Newtonian approximation	$2U$	0	0
post-Newtonian approximation [(Fock (1959); Chandrasekhar (1965a))]	+ terms $\sim \epsilon^4$	+ terms $\sim \epsilon^3$	+ terms $\sim \epsilon^2$
post-post-Newtonian approximation [Chandrasekhar and Nutku (1969)]	+ terms $\sim \epsilon^6$	+ terms $\sim \epsilon^5$	+ terms $\sim \epsilon^4$
radiation damping [Chandrasekhar and Esposito (1970)]	+ terms $\sim \epsilon^7$	+ terms $\sim \epsilon^6$	+ terms $\sim \epsilon^5$

(2) The law of local energy-momentum conservation, $T^{\alpha\beta}_{;\beta} = 0$. The time component of this law reduces to the conservation of rest mass

$$\partial\rho_o/\partial t + \partial(\rho_o v_j)/\partial x_j = 0 + \text{fractional errors of } O(\epsilon^2); \quad (39.15a)$$

and the space components reduce to Newton's second law of motion, " $F = ma$ ":

$$\rho_o dv_j/dt = \rho_o(\partial U/\partial x_j) - \partial t_{jk}/\partial x_k + \text{fractional errors of } O(\epsilon^2), \quad (39.15b)$$

$$d/dt \equiv (\text{time derivative following the matter}) \equiv \partial/\partial t + v_k \partial/\partial x_k \quad (39.16)$$

(see exercise 39.3).

Equations (39.14)–(39.16), together with equations of state describing the planetary and solar matter, are the foundations for all Newtonian calculations of the structure and motion of the sun and planets. Notice that the internal energy density $\rho_o \Pi$ nowhere enters into these equations. It is of no importance to Newtonian hydrodynamics. It matters for the sun's thermal-energy balance; but that is irrelevant here.

EXERCISES

Exercise 39.3. NEWTONIAN APPROXIMATION

(a) Derive equations (39.13) for the components of the stress-energy tensor in the PPN coordinate frame. [Hint: In the rest frame of the matter ("comoving orthonormal frame") $T_{\hat{0}\hat{0}} = \rho = \rho_o + O(\epsilon^2)$, $T_{\hat{0}\hat{j}} = 0$, $T_{\hat{j}\hat{k}} = t_{jk}$; see equations (39.5). Lorentz-transform these components by a pure boost with ordinary velocity $-v_j$ to obtain $T_{\alpha\beta}$.]

(b) Show that, in the PPN coordinate frame, $T^{0\alpha}_{;\alpha} = 0$ reduces to equation (39.15a), and $T^{j\alpha}_{;\alpha} = 0$, when combined with (39.15a), reduces to equation (39.15b).]

Exercise 39.4. A USEFUL FORMULA

Derive from equations (39.15) the following useful formula, valid for any function $f(x, t)$:

$$\frac{d}{dt} \int \rho_o(x, t) f(x, t) d^3x = \int \rho_o(x, t) \frac{df(x, t)}{dt} d^3x + \text{fractional errors of } O(\epsilon^2). \quad (39.17)$$

Here both integrals are extended over all of space; and df/dt is the derivative following the matter (39.16).

Exercise 39.5. STRESS TENSOR FOR NEWTONIAN GRAVITATIONAL FIELD

Define a "stress tensor for the Newtonian gravitational field U " as follows:

$$t_{jk} \equiv \frac{1}{4\pi} \left(U_{,j} U_{,k} - \frac{1}{2} \delta_{jk} U_{,\ell} U_{,\ell} \right). \quad (39.18)$$

Show that the equations of motion for the matter (39.15b) can be rewritten in the forms

$$\rho_o \frac{dv_j}{dt} = - \frac{\partial}{\partial x^k} (t_{jk} + t_{jk}) + \text{fractional errors of } O(\epsilon^2), \quad (39.19)$$

$$(\rho_o v_j)_{,t} + (t_{jk} + t_{jk} + \rho_o v_j v_k)_{,k} = 0 + \text{fractional errors of } O(\epsilon^2). \quad (39.19')$$

Exercise 39.6. NEWTONIAN VIRIAL THEOREMS

(a) From equation (39.19') show that

$$\frac{d^2 I_{jk}}{dt^2} = 2 \int (t_{jk} + t_{jk} + \rho_0 v_j v_k) d^3x + \text{fractional errors of } O(\epsilon^2), \quad (39.20a)$$

where I_{jk} is the second moment of the system's mass distribution,

$$I_{jk} = \int \rho_0 x_j x_k d^3x.$$

This is called the “time-dependent tensor virial theorem.”

(b) From this infer that, if $\langle \dots \rangle_{\text{long time}}$ denotes an average over a long period of time, then

$$\left\langle \int (t_{jk} + t_{jk} + \rho_0 v_j v_k) d^3x \right\rangle_{\text{long time}} = O \left(\int \rho_0 \epsilon^4 d^3x \right). \quad (39.20b)$$

This is called the “tensor virial theorem.”

(c) By contraction of indices and use of equations (39.18), (39.14a), and (39.5e), derive the (ordinary) *virial theorems*:

$$\frac{1}{2} \frac{d^2 I}{dt^2} = \int \rho_0 v^2 d^3x - \int \frac{1}{2} \rho_0 U d^3x + 3 \int p d^3x + \text{fractional errors of } O(\epsilon^2), \quad (39.21a)$$

where I is the trace of the second moment of the mass distribution

$$I = I_{jj} = \int \rho_0 r^2 d^3x;$$

and

$$\underbrace{\left(\int \rho_0 v^2 d^3x \right)}_{2 \times \begin{pmatrix} \text{kinetic} \\ \text{energy} \end{pmatrix}} - \underbrace{\frac{1}{2} \int \rho_0 U d^3x}_{\begin{pmatrix} \text{gravitational} \\ \text{energy} \end{pmatrix}} + \underbrace{3 \int p d^3x}_{3 \times \begin{pmatrix} \text{pressure} \\ \text{integral} \end{pmatrix}} = O \left(\int \rho_0 \epsilon^4 d^3x \right). \quad (39.21b)$$

Exercise 39.7. PULSATION FREQUENCY FOR NEWTONIAN STAR

Use the ordinary, time-dependent virial theorem (39.21a) to derive the following equation for the fundamental angular frequency of pulsation of a nonrotating, Newtonian star:

$$\omega^2 = (3\bar{\Gamma}_1 - 4) \frac{|\text{star's self-gravitational energy}|}{(\text{trace of second moment of star's mass distribution})}; \quad (39.22a)$$

$$\bar{\Gamma}_1 = \left(\begin{array}{l} \text{pressure-weighted average} \\ \text{of adiabatic index} \end{array} \right) \equiv \frac{\int \Gamma_1 p d^3x}{\int p d^3x}. \quad (39.22b)$$

In the derivation assume that the pulsations are “homologous”—i.e., that a fluid element with equilibrium position x^j (relative to center of mass $x^j = 0$) gets displaced to $x^j + \xi^j(x, t)$, where

$$\xi^j = (\text{small constant}) x^j e^{-i\omega t}.$$

Assume nothing else. Notes: (1) The result (39.22) was derived differently in Box 26.2 and used in §24.4. (2) The assumption of homologous pulsation is fully justified if $|\Gamma_1 - 4/3| = \text{constant} \ll 1$; see Box 26.2. (3) The result (39.22) is readily generalized to slowly

rotating Newtonian stars; see, e.g., Chandrasekhar and Lebovitz (1968). It can also be generalized to nonrotating post-Newtonian stars using general relativity (Box 26.2), or using the PPN formalism for any metric theory [Ni (1973)]. And it can be generalized to slowly rotating, post-Newtonian stars [see, e.g., Chandrasekhar and Lebovitz (1968)].

§39.8. PPN METRIC COEFFICIENTS

Post-Newtonian corrections to metric:

(1) rules governing forms

(2) construction of corrections

The post-Newtonian corrections $k_{\alpha\beta}$ to the metric coefficients $g_{\alpha\beta}$ are calculated, in any metric theory of gravity, by lengthy manipulations of the field equations. (See, e.g., exercise 39.14 near the end of this chapter for general relativity.) But without ever picking some one theory, and without ever writing down any set of field equations, one can infer the *forms* of the post-Newtonian corrections $k_{\alpha\beta}$. Their forms are fixed by the following constraints: (1) They must be of post-Newtonian order ($k_{00} \sim \epsilon^4$, $k_{0j} \sim \epsilon^3$, $k_{ij} \sim \epsilon^2$). (2) They must be dimensionless. (3) k_{00} must be a scalar under rotations, k_{0j} must be components of a 3-vector, and k_{jk} must be components of a 3-tensor. (4) The corrections must die out at least as fast as $1/r$ far from the solar system, so that the coordinates become globally Lorentz and spacetime becomes flat at $r = \infty$. (5) For simplicity, one can assume that the metric components are generated only by ρ_o , $\rho_o \Pi$, t_{ij} , p , products of these with the velocity v_j , and time-derivatives of such quantities;* but not by their spatial gradients. [This assumption of simplicity is satisfied by all metric theories examined up to 1973, except Whitehead (1922) and theories reviewed by Will (1973)—which disagree with experiment.] Note the further justification for this assumption in exercise 39.8.

Begin with the corrections to the spatial components, $k_{ij} \sim \epsilon^2$. There are only two functionals of ρ_o , p , Π , t_{jk} , v_j , that die out at least as fast as $1/r$, are dimensionless, are $O(\epsilon^2)$, and are second-rank, symmetric 3-tensors; they are

$$\delta_{ij}U(\mathbf{x}, t); \quad U_{ij}(\mathbf{x}, t) = \int \frac{\rho_o(\mathbf{x}', t)(x_i - x'_i)(x_j - x'_j)}{|\mathbf{x} - \mathbf{x}'|^3} d^3x'. \quad (39.23a)$$

Thus, k_{ij} must be $k_{ij} = 2\gamma \delta_{ij}U + 2\Gamma U_{ij}$, for some constant “PPN parameters” γ and Γ . By an infinitesimal coordinate transformation [$x_{i\text{NEW}} = x_{i\text{OLD}} + \Gamma \partial\chi/\partial x_i$, with $\chi(\mathbf{x}, t) = -\int \rho_o(\mathbf{x}', t)|\mathbf{x} - \mathbf{x}'| d^3x'$] one can set $\Gamma = 0$, thereby obtaining

$$g_{ij} = \delta_{ij} + k_{ij} = \delta_{ij}(1 + 2\gamma U) + O(\epsilon^4). \quad (39.23b)$$

*One allows for time derivatives because retarded integrals contain such terms when expanded to post-Newtonian order; thus,

$$\int \frac{\rho_o(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|)}{|\mathbf{x} - \mathbf{x}'|} d^3x' = \int \left[\frac{\rho_o(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} - \frac{\partial \rho_o(\mathbf{x}', t)}{\partial t} + \dots \right] d^3x'.$$

However, it turns out that, with a suitable choice of coordinates (“gauge”), all time-odd retarded terms [e.g., $\int (\partial \rho_o / \partial t) d^3x$] vanish, except at “the post $^{5/2}$ -Newtonian order” and at higher orders of approximation; there they lead to radiation damping (see Box 39.3). For example, $\int (\partial \rho_o / \partial t) d^3x = (d/dt) \int \rho_o d^3x$ vanishes by virtue of the conservation of baryon number.

Next consider $k_{0j} \sim \epsilon^3$. Trial and error yield only two vector functionals that die out as $1/r$ or faster, are dimensionless, and are $O(\epsilon^3)$. They are

$$V_j(x, t) = \int \frac{\rho_o(x', t)v_j(x', t)}{|x - x'|} d^3x', \quad (39.23c)$$

$$W_j(x, t) = \int \frac{\rho_o(x', t)[(x - x') \cdot v(x', t)](x_j - x'_j) d^3x'}{|x - x'|^3}. \quad (39.23d)$$

Thus, k_{0j} must be a linear combination of these, involving unknown constants (PPN parameters) Δ_1 and Δ_2 :

$$g_{0j} = k_{0j} = -\frac{7}{2}\Delta_1 V_j - \frac{1}{2}\Delta_2 W_j + O(\epsilon^5). \quad (39.23e)$$

Finally consider $k_{00} \sim \epsilon^4$. Trial and error yields a variety of terms, which can all be combined together with the Newtonian part of g_{00} to give

$$g_{00} = -1 + 2U + k_{00} = -1 + 2U - 2\beta U^2 + 4\Psi - \xi \mathcal{A} - \eta \mathcal{D}, \quad (39.23f)$$

where

$$\Psi(x, t) = \int \frac{\rho_o(x', t)\psi(x', t)}{|x - x'|} d^3x', \quad (39.23g)^*$$

$$\psi = \beta_1 v^2 + \beta_2 U + \frac{1}{2}\beta_3 \Pi + \frac{3}{2}\beta_4 p/\rho_o,$$

$$\mathcal{A}(x, t) = \int \frac{\rho_o(x', t)[(x - x') \cdot v(x', t)]^2}{|x - x'|^3} d^3x', \quad (39.23h)$$

$$\mathcal{D}(x, t) = \int \frac{\left[t_{jk}(x', t) - \frac{1}{3}\delta_{jk}t_{11}(x', t) \right] (x_j - x'_j)(x_k - x'_k)}{|x - x'|^3} d^3x'. \quad (39.23i)$$

Also, $\beta, \beta_1, \beta_2, \beta_3, \beta_4, \xi, \eta$ are unknown constants (PPN parameters). Elsewhere in the literature the term $-\eta \mathcal{D}$ in g_{00} is ignored (see footnote on p. 1075).

Yet another term is possible: one could have set

$$g_{00} = [\text{value in equation (39.23f)}] - \Sigma \int \int \frac{\rho_o(x', t)\rho_o(x'', t)[(x - x') \cdot (x' - x'')] d^3x' d^3x''}{|x - x'| |x' - x''|^3}, \quad (39.24)$$

where Σ is another PPN parameter. [It can be shown, using the Newtonian equations (39.14)–(39.16), that this expression dies out as $1/r$ far from the solar system.] If

*WARNING: *Throughout* the literature the notation Φ is used where we use Ψ for the functional (39.23g), and ϕ is used for our ψ . We are forced to violate the standard notation to avoid confusion with the Newtonian potential $\Phi = -U$. However, we urge that nobody else violate the standard notation!

such a Σ term had been included, then one could have removed it by making the infinitesimal coordinate transformation

$$x_{\text{new}}^0 = x_{\text{old}}^0 - \frac{1}{2} \Sigma \int \frac{\rho_o(\mathbf{x}', t)[(\mathbf{x} - \mathbf{x}') \cdot \mathbf{v}(\mathbf{x}', t)]}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (39.25)$$

Rigidity of coordinate system

(see exercise 39.9). Thus, there is no necessity to include the Σ term.

The absence of the Σ term from g_{00} means that the time coordinate has been fixed rigidly up through post-Newtonian order:

$$x^0 \text{ has uncertainties only of } O(R_\odot \epsilon^5) \sim 10^{-14} \text{ seconds.} \quad (39.26a)$$

The space coordinates are also fixed rigidly through post-Newtonian order:

$$x^j \text{ has uncertainties only of } O(R_\odot \epsilon^4) \sim 0.1 \text{ cm,} \quad (39.26b)$$

because any transformation of the form

$$x_{\text{new}}^j = x_{\text{old}}^j + \text{position-dependent terms of } O(\epsilon^2 R_\odot)$$

Summary of PPN metric and parameters

would destroy the form (39.23b) of the space part of the metric.

In summary, for almost every metric theory of gravity yet invented, accurate through post-Newtonian order the metric coefficients have the form (39.23). One theory is distinguished from another by the values of its ten “post-Newtonian parameters” $\beta, \beta_1, \beta_2, \beta_3, \beta_4, \gamma, \xi, \eta, \Delta_1$ and Δ_2 . These are determined by comparing the field equations of the given theory with the form (39.23) of the post-Newtonian metric. The parameter values for general relativity and for several other theories are given in Box 39.2, along with a heuristic description of each parameter.

EXERCISES

Exercise 39.8. ABSENCE OF “METRIC-GENERATES-METRIC” TERMS IN POST-NEWTONIAN LIMIT

In writing down the post-Newtonian metric corrections, one might be tempted to include terms that are generated by the Newtonian potential acting alone, without any *direct* aid from the matter. After all, general relativity and other metric theories are nonlinear; so the two-step process (matter) $\rightarrow U \rightarrow$ (post-Newtonian metric corrections) seems quite natural. *Show* that such terms are not needed, because the equations (39.14)–(39.16) of the Newtonian approximation enable one to reexpress them in terms of direct integrals over the matter distribution. In particular, show that

$$\int \frac{\partial^2 U(\mathbf{x}', t)/\partial x_j' \partial t}{|\mathbf{x} - \mathbf{x}'|} d^3x' = 2\pi[V_j(\mathbf{x}, t) - W_j(\mathbf{x}, t)] \quad (39.27)$$

where V_j and W_j are defined by equations (39.23c,d); also show that

$$\begin{aligned} \int \frac{[\partial U(\mathbf{x}', t)/\partial x_j][\partial U(\mathbf{x}', t)/\partial x_j]}{|\mathbf{x} - \mathbf{x}'|} d^3x' \\ = -2\pi[U(\mathbf{x}, t)]^2 + 4\pi \int \frac{\rho_o(\mathbf{x}', t)U(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} d^3x'. \end{aligned} \quad (39.28)$$

Note that the terms on the righthand sides of (39.27) and (39.28) are already included in the expressions (39.23e,f) for g_{0j} and g_{00} .

Exercise 39.9. REMOVAL OF Σ TERM FROM g_{00}

Show that the coordinate transformation (39.25) removes the Σ term from the metric coefficient g_{00} of equation (39.24), as claimed in the text.

Exercise 39.10. VERIFICATION OF FORMS OF POST-NEWTONIAN CORRECTIONS

Verify the claims in the text immediately preceding equations (39.23a,b,c,f).

§39.9. VELOCITY OF PPN COORDINATES RELATIVE TO “UNIVERSAL REST FRAME”

Thus far it has been assumed tacitly that the center of mass of the solar system is at rest in the PPN coordinate system. Is this really a permissible assumption? Put differently, can one always so adjust the PPN coordinate system that its origin moves with any desired velocity (e.g., that of the solar system); or is the PPN coordinate system rigidly and irrevocably attached to some “universal rest frame”?

In general relativity, the geometry of curved spacetime picks out no preferred coordinate frames (except in cases with special symmetry). Therefore, one expects the velocity of the PPN coordinate frame to be freely specifiable. Put differently, one expects the entire PPN formalism, for general relativity, to be invariant under Lorentz transformations of the PPN coordinates [combined, perhaps, with “infinitesimal coordinate transformations” to maintain the gauge conditions that the “ Σ ” and “ U_{jk} ” terms of (39.24) and (39.23a) be absent]. By contrast, in Ni’s theory of gravity (Box 39.1) the geometry of spacetime *always* picks out a preferred coordinate frame: the “rest frame of the universe.” One would not be surprised, in this case, to find the PPN coordinate frame rigidly attached to the universal rest frame.

The above intuition is correct, according to calculations by Will (1971d) and by Will and Nordtvedt (1972). When dealing with general relativity and other theories with little or no “prior geometry,” one can freely specify the velocity of the PPN coordinate system (at some initial instant of time). But for theories like Ni’s, with a preferred “universal rest frame” (“preferred-frame theories”), only in the preferred frame can the post-Newtonian metric assume the form derived in the last section [equations (39.23)]. This restriction on the PPN metric does not mean that one is confined, in preferred-frame theories, to perform all calculations in the universal rest frame. Rather, it means that for such theories the PPN metric requires generalization to take account of coordinate-frame motion relative to the universal rest frame.

The required generalization can be achieved by subjecting the PPN metric (39.23) to (1) a Lorentz boost from the preferred frame $\{x_{\text{OLD}}^\alpha\}$ to a new PPN frame $\{x_{\text{NEW}}^\alpha\}$, which moves with velocity w , plus (2) a change of gauge designed to keep the metric coefficients as simple as possible. The boost-plus-gauge-change is [Will and Nordtvedt (1972)]

Preferred-frame theories of gravity

Generalization of PPN metric to moving frames

$$\mathbf{x}_{\text{OLD}} = \mathbf{x}_{\text{NEW}} + \frac{1}{2}(\mathbf{x}_{\text{NEW}} \cdot \mathbf{w})\mathbf{w} + \left(1 + \frac{1}{2}\mathbf{w}^2\right)\mathbf{w}t_{\text{NEW}} \quad (39.29\text{a})$$

$$+ O(\epsilon^5 t_{\text{NEW}} + \epsilon^4 \mathbf{x}_{\text{NEW}}),$$

$$\begin{aligned} t_{\text{OLD}} &= t_{\text{NEW}} \left(1 + \frac{1}{2}\mathbf{w}^2 + \frac{3}{8}\mathbf{w}^4\right) + \left(1 + \frac{1}{2}\mathbf{w}^2\right)\mathbf{x}_{\text{NEW}} \cdot \mathbf{w} \\ &\quad + \underbrace{\left(\frac{1}{2}\Delta_2 + \xi - 1\right)w_j \frac{\partial \chi}{\partial x_{\text{NEW}}^j}}_{\uparrow \text{[gauge change]}} + O(\epsilon^6 t_{\text{NEW}} + \epsilon^5 \mathbf{x}_{\text{NEW}}), \end{aligned} \quad (39.29\text{b})$$

$$\chi(t_{\text{NEW}}, \mathbf{x}_{\text{NEW}}) \equiv - \int \rho_o(t_{\text{NEW}}, \mathbf{x}'_{\text{NEW}}) |\mathbf{x}_{\text{NEW}} - \mathbf{x}'_{\text{NEW}}| d^3x'_{\text{NEW}}. \quad (39.29\text{c})$$

[Note: One insists, in the spirit of the post-Newtonian approximation, that the velocity \mathbf{w} of the new PPN frame relative to the universal rest frame be no larger than the characteristic internal velocities of the system:

$$|\mathbf{w}| \lesssim \epsilon. \quad (39.30)$$

This change of coordinates produces corresponding changes in the velocity of the matter

$$\begin{aligned} v_{\text{OLD}} &= \frac{dx_{\text{OLD}}}{dt_{\text{OLD}}} = v_{\text{NEW}} \left(1 - \mathbf{w} \cdot \mathbf{v}_{\text{NEW}} - \frac{1}{2}\mathbf{w}^2\right) \\ &\quad + \mathbf{w} \left(1 - \frac{1}{2}\mathbf{w} \cdot \mathbf{v}_{\text{NEW}}\right) + O(\epsilon^5). \end{aligned} \quad (39.31)$$

A long but straightforward calculation (exercise 39.11) yields the following components for the metric in the new PPN coordinates. [Note: The subscripts NEW are here and hereafter dropped from the notation.]

Final form of metric

$$g_{jk} = \delta_{jk}(1 + 2\gamma U) + O(\epsilon^4), \quad (39.32\text{a})$$

$$g_{0j} = -\frac{7}{2}\Delta_1 V_j - \frac{1}{2}\Delta_2 W_j + \left(\alpha_2 - \frac{1}{2}\alpha_1\right)w_j U - \alpha_2 w_k U_{kj} + O(\epsilon^5), \quad (39.32\text{b})$$

$$\begin{aligned} g_{00} &= -1 + 2U - 2\beta U^2 + 4\Psi - \xi \mathcal{A} - \eta \mathcal{D} \\ &\quad + (\alpha_2 + \alpha_3 - \alpha_1)\mathbf{w}^2 U + (2\alpha_3 - \alpha_1)w_j V_j - \alpha_2 w_j w_k U_{kj} + O(\epsilon^6). \end{aligned} \quad (39.32\text{c})$$

Here α_1 , α_2 , and α_3 are certain combinations of PPN parameters

$$\alpha_1 = 7\Delta_1 + \Delta_2 - 4\gamma - 4, \quad (39.33\text{a})$$

$$\alpha_2 = \Delta_2 + \xi - 1, \quad (39.33\text{b})$$

$$\alpha_3 = 4\beta_1 - 2\gamma - 2 - \xi. \quad (39.33\text{c})$$

The “gravitational potentials” U , V_j , W_j , Ψ , \mathcal{A} , and \mathcal{D} appearing here are to be calculated in the new, “moving” PPN coordinate system by the same prescriptions

as one used in the universal rest frame. Thus, their functional forms are the same as previously, but their values at any given event are different (see exercise 39.11):

$$U(\mathbf{x}, t) = \int \frac{\rho_o(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} d^3x'; \quad (39.34a)$$

$$V_j(\mathbf{x}, t) = \int \frac{\rho_o(\mathbf{x}', t)v_j(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} d^3x'; \quad (39.34b)$$

$$W_j(\mathbf{x}, t) = \int \frac{\rho_o(\mathbf{x}', t)[(\mathbf{x} - \mathbf{x}') \cdot \mathbf{v}(\mathbf{x}', t)](x_j - x'_j)}{|\mathbf{x} - \mathbf{x}'|^3} d^3x'; \quad (39.34c)$$

$$\Psi(\mathbf{x}, t) = \int \frac{\rho_o(\mathbf{x}', t)\psi(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} d^3x',$$

$$\psi = \beta_1 v^2 + \beta_2 U + \frac{1}{2}\beta_3 \Pi + \frac{3}{2}\beta_4 p/\rho_o; \quad (39.34d)$$

$$\mathcal{A}(\mathbf{x}, t) = \int \frac{\rho_o(\mathbf{x}', t)[(\mathbf{x} - \mathbf{x}') \cdot \mathbf{v}(\mathbf{x}', t)]^2}{|\mathbf{x} - \mathbf{x}'|^3} d^3x'; \quad (39.34e)$$

$$\mathcal{D}(\mathbf{x}, t) = \int \frac{\left[t_{jk}(\mathbf{x}', t) - \frac{1}{3}\delta_{jk}t_{ll}(\mathbf{x}', t) \right] (x_j - x'_j)(x_k - x'_k)}{|\mathbf{x} - \mathbf{x}'|^3} d^3x'. \quad (39.34f)$$

The quantity U_{jk} is the gravitational potential defined in equation (39.23a):

$$U_{jk}(\mathbf{x}, t) = \int \frac{\rho_o(\mathbf{x}', t)(x_j - x'_j)(x_k - x'_k)}{|\mathbf{x} - \mathbf{x}'|^3} d^3x'. \quad (39.34g)$$

Notice that the velocity \mathbf{v} of the PPN coordinate system relative to the universal rest frame appears explicitly in the PPN metric only if one or more of the coefficients $\alpha_1, \alpha_2, \alpha_3$ is nonzero. Thus, theories with $\alpha_1 = \alpha_2 = \alpha_3 = 0$ (e.g., general relativity) possess *no* preferred universal rest frame in the post-Newtonian limit; all their PPN frames are “created equal.” By contrast, theories with at least one of $\alpha_1, \alpha_2, \alpha_3$, nonzero (e.g., Ni’s theory) *do* possess a preferred frame.

The generalized form (39.32) of the PPN metric, by virtue of the process used to construct it, is invariant under a Lorentz boost plus a gauge adjustment [“Post-Galilean transformation”; see Chandrasekhar and Contopoulos (1967)]:

$$\begin{aligned} \mathbf{x}_{\text{OLD}} &= \mathbf{x}_{\text{NEW}} + \frac{1}{2}(\mathbf{x}_{\text{NEW}} \cdot \boldsymbol{\beta})\boldsymbol{\beta} + \left(1 + \frac{1}{2}\beta^2\right)\boldsymbol{\beta}t_{\text{NEW}} \\ &\quad + O(\epsilon^5 t_{\text{NEW}} + \epsilon^4 \mathbf{x}_{\text{NEW}}), \\ t_{\text{OLD}} &= \left(1 + \frac{1}{2}\beta^2 + \frac{3}{8}\beta^4\right)t_{\text{NEW}} + \left(1 + \frac{1}{2}\beta^2\right)\mathbf{x}_{\text{NEW}} \cdot \boldsymbol{\beta} \\ &\quad + \left(\frac{1}{2}\Delta_2 + \xi_1 - 1\right)\boldsymbol{\beta} \cdot \nabla_{\text{NEW}}\chi + O(\epsilon^6 t_{\text{NEW}} + \epsilon^5 \mathbf{x}_{\text{NEW}}). \end{aligned} \quad (39.35)$$

Post-Galilean invariance

Of course, it is also invariant under spatial rotations.

EXERCISE**Exercise 39.11. TRANSFORMATION TO MOVING FRAME**

Show that the change of coordinates (39.29) changes the PPN metric coefficients from the form (39.23) to the form (39.32). [Hints: (1) Keep firmly in mind the fact that the potentials U , V_j , W_j , \mathcal{A} , and \mathcal{D} are not scalar fields. Each coordinate system possesses its own potentials. For example, by using equations (39.29) in the integral for U_{OLD} , one finds

$$\begin{aligned} U_{\text{OLD}}(x_{\text{OLD}}, t_{\text{OLD}}) &= \int \frac{\rho_0(x'_{\text{OLD}}, t_{\text{OLD}})}{|x_{\text{OLD}} - x'_{\text{OLD}}|} d^3x'_{\text{OLD}} \\ &= \left[U_{\text{NEW}} - w_j(V_{j\text{NEW}} - W_{j\text{NEW}}) + \frac{1}{2} w_j w_k \chi_{jk} \right]_{x_{\text{NEW}}, t_{\text{NEW}}} + O(\epsilon^6). \end{aligned} \quad (39.36)$$

(2) The law of baryon conservation (39.44) may be useful.]

§39.10. PPN STRESS-ENERGY TENSOR

The motion of the solar system is governed by the equations $T^{\alpha\beta}_{;\beta} = 0$. Before studying them, one must calculate the post-Newtonian corrections to the stress-energy tensor in the PPN coordinate frame. This requires a transformation from the comoving, orthonormal frame $\omega^{\hat{\alpha}}$, where

$$T^{\hat{0}\hat{0}} = \rho_0(1 + \Pi), \quad T^{\hat{0}\hat{j}} = 0, \quad T^{\hat{j}\hat{k}} = t_{jk}, \quad (39.37)$$

to the coordinate frame. One can effect this transformation in two stages: stage 2 is a transformation

$$\omega^{\tilde{\alpha}} \equiv [1 - U + O(\epsilon^4)] dt \quad (39.38a)$$

$$+ \left[\frac{7}{2} A_1 V_j + \frac{1}{2} A_2 W_j + \left(\frac{1}{2} \alpha_1 - \alpha_2 \right) w_j U + \alpha_2 w_k U_{kj} + O(\epsilon^5) \right] dx^j,$$

$$\omega^{\tilde{j}} \equiv [(1 + \gamma U) \delta_{jk} + O(\epsilon^4)] dx^k + O(\epsilon^5) dt, \quad (39.38b)$$

between the coordinate frame and an orthonormal frame attached to it; stage 1 is a pure Lorentz transformation (boost) between the two orthonormal frames $\omega^{\tilde{\alpha}}$ and $\omega^{\hat{\alpha}}$. The 4-velocity of this boost is minus the 4-velocity of the matter, which has components

$$u^j = v_j u^0, \quad u^0 = 1 + \frac{1}{2} v^2 + U + O(\epsilon^4) \quad \text{in coord. frame}; \quad (39.39)$$

$$\left. \begin{aligned} u^{\tilde{j}} &= v_{\tilde{j}} u^{\tilde{0}}, \quad u^{\tilde{0}} = 1 + \frac{1}{2} v^2 + O(\epsilon^4), \\ v_{\tilde{j}} &= v_j [1 + (1 + \gamma) U] \end{aligned} \right\} \text{in } \omega^{\tilde{\alpha}} \text{ frame.} \quad (39.40)$$

Combining the boost, which has ordinary velocity $\beta_{\tilde{j}} = -v_{\tilde{j}}$, with the transformation (39.38), and then inverting, one obtains the result (exercise 39.12)

$$dx^\alpha = A^\alpha_\beta \omega^\beta, \quad \begin{cases} \omega^\beta = \text{orthonormal comoving basis,} \\ dx^\alpha = \text{PPN coordinate basis;} \end{cases}$$

Transformation from rest frame of matter to PPN coordinate frame

$$A^0_{\hat{0}} = 1 + \frac{1}{2} v^2 + U + O(\epsilon^4),$$

$$\begin{aligned} A^0_{\hat{j}} &= v_j \left[1 + \frac{1}{2} v^2 + (2 + \gamma)U \right] - \frac{7}{2} A_1 V_j - \frac{1}{2} A_2 W_j \\ &\quad + \left(\alpha_2 - \frac{1}{2} \alpha_1 \right) w_j U - \alpha_2 w_k U_{kj} + O(\epsilon^5), \end{aligned} \quad (39.41)$$

$$A^j_{\hat{0}} = v_j \left[1 + \frac{1}{2} v^2 + U \right] + O(\epsilon^5),$$

$$A^j_{\hat{k}} = (1 - \gamma U) \delta_{jk} + \frac{1}{2} v_j v_k + O(\epsilon^4).$$

This transformation, when applied to the stress-energy tensor (39.37) yields, in the PPN coordinate frame,

$$T^{00} = \rho_o (1 + \Pi + v^2 + 2U) + O(\rho_o \epsilon^4), \quad (39.42a)$$

$$T^{0j} = \rho_o (1 + \Pi + v^2 + 2U) v_j + t_{j\hat{m}} v_m + O(\rho_o \epsilon^5), \quad (39.42b)$$

$$\begin{aligned} T^{jk} &= t_{jk} (1 - 2\gamma U) + \rho_o (1 + \Pi + v^2 + 2U) v_j v_k \\ &\quad + \frac{1}{2} (v_j t_{k\hat{m}} v_m + v_k t_{j\hat{m}} v_m) + O(\rho_o \epsilon^6). \end{aligned} \quad (39.42c)$$

Stress-energy tensor in coordinate frame

Exercise 39.12. THE TRANSFORMATION BETWEEN COMOVING FRAME AND PPN FRAME

EXERCISE

Carry out the details of the derivation of the transformation matrix (39.41); and in the process calculate the correction of $O(\epsilon^4)$ to $A^0_{\hat{0}}$.

§39.11. PPN EQUATIONS OF MOTION

The post-Newtonian corrections to the Newtonian equations of motion (39.15) and (39.16) are derived from the law of conservation of baryon number $(\rho_o u^\alpha)_{;\alpha} = 0$, and from the law of conservation of local energy-momentum, $T^{\alpha\beta}_{;\beta} = 0$. The simplest of the equations of motion is the conservation of baryon number. Its exact expression is $(\rho_o u^\alpha)_{;\alpha} = (1/\sqrt{-g})(\sqrt{-g}\rho_o u^\alpha)_{,\alpha} = 0$. Define a new quantity

$$\begin{aligned} \rho^* &\equiv \rho_o \left(1 + \frac{1}{2} v^2 + 3\gamma U \right) \\ &= \rho_o u^0 \sqrt{-g} + O(\rho_o \epsilon^4) \end{aligned} \quad (39.43)$$

[see (39.39) for u^0 , and (39.32) for the metric]. Then rest-mass conservation takes on the same form as at the Newtonian order (39.15a), except now it is more accurate:

Law of baryon conservation

$$\rho^*,_t + (\rho^* v_j),_j = 0 + \text{errors of } O(\rho_o, \epsilon^5). \quad (39.44)$$

The next simplest equation of motion is $T^{0\alpha},_\alpha = 0$. Straightforward evaluation, using the metric of equations (39.32) and the stress-energy tensor of equations (39.42), yields

$$[\rho_o(1 + \Pi + v^2 + 2U)],_t + [\rho_o(1 + \Pi + v^2 + 2U)v_j + t_{\hat{j}\hat{m}}v_m],_j + (3\gamma - 2)\rho_o U,_t + (3\gamma - 3)\rho_o v_k U,_k = O(\rho_o, \epsilon^5). \quad (39.45)$$

By subtracting equation (39.44) from this, and using the Newtonian equations of motion (39.15) and (39.16) to simplify several terms where the Newtonian approximation is adequate, one obtains

Law of energy conservation

$$\rho_o d\Pi/dt + t_{\hat{j}\hat{k}}v_{j,k} = 0 + \text{errors of } O(\rho_o, \epsilon^5). \quad (39.46)$$

Notice that this is nothing but the first law of thermodynamics (local energy conservation) with energy flow through the matter being neglected. (Neglecting energy flow was justified in §39.5.) This first law of thermodynamics is actually a post-Newtonian equation in the context of hydrodynamics, rather than a Newtonian equation, because Π does not affect the hydrodynamic motion at Newtonian order (see §39.7).

The last of the equations of motion, $T^{ja},_\alpha = 0$, reduces to the post-Newtonian Euler equation

Post-Newtonian Euler equation

$$\begin{aligned} & \rho^* \frac{dv_j}{dt} - \rho^* U,_j + [t_{\hat{j}\hat{k}}(1 + 3\gamma U)],_k - t_{\hat{j}\hat{k}},_k \left(\frac{1}{2} \mathbf{v}^2 + \Pi \right) - \frac{t_{\hat{j}\hat{k}}t_{\hat{k}\hat{l}},_l}{\rho^*} \\ & + \rho^* \frac{d}{dt} \left[(2\gamma + 2)Uv_j - \frac{1}{2}(7\Delta_1 + \Delta_2)V_j - \frac{1}{2}\alpha_1 Uw_j \right] - v_j \rho^* U,_t + v_k t_{\hat{k}\hat{j},t} \\ & + \frac{1}{2} \Delta_2 \rho^*(V_j - W_j),_t + \frac{1}{2} \rho^* [(7\Delta_1 + \Delta_2)v_k + (\alpha_1 - 2\alpha_3)w_k] V_{k,j} \\ & - \rho^* \left[2\Psi - \frac{1}{2}\xi\mathcal{A} - \frac{1}{2}\eta\mathcal{D} - \frac{1}{2}\alpha_2 w_i w_k U_{ik} + \alpha_2 w_i (V_i - W_i) \right]_j \\ & - \rho^* U,_j \left[\gamma \mathbf{v}^2 - \frac{1}{2}\alpha_1 \mathbf{w} \cdot \mathbf{v} + \frac{1}{2}(\alpha_2 + \alpha_3 - \alpha_1)\mathbf{w}^2 - (2\beta - 2)U + 3\gamma p/\rho^* \right] \\ & + \frac{1}{2}(v_{j,k} t_{\hat{k}\hat{m}} v_m - t_{\hat{j}\hat{m}} v_{m,k} v_k) + \frac{1}{2}[v_m(t_{\hat{m}\hat{j}} v_k),_k - v_j(t_{\hat{k}\hat{l}} v_k),_l] = 0. \end{aligned} \quad (39.47)$$

Partial derivatives are denoted by commas; d/dt is the time-derivative following the matter [equation (39.16)].

Equations (39.44), (39.46), and (39.47) are a complete set of equations of motion at the post-Newtonian order.

Exercise 39.13. EQUATIONS OF MOTION**EXERCISES**

Carry out the details of the derivation of the equations of motion (39.44), (39.46), and (39.47). As part of the derivation, calculate the following values of the Christoffel symbols in the PPN coordinate frame:

$$\begin{aligned}\Gamma^0_{00} &= -U_{,t} + O(U_{,j}\epsilon^3), \quad \Gamma^0_{0j} = -U_{,j} + O(U_{,j}\epsilon^2), \\ \Gamma^0_{jk} &= \gamma U_{,t} \delta_{jk} + \frac{7}{2} \Delta_1 V_{(j,k)} + \frac{1}{2} \Delta_2 W_{(j,k)} + \left(\frac{1}{2} \alpha_1 - \alpha_2\right) w_{(j} U_{,k)} \\ &\quad + \alpha_2 w_i U_{i(j,k)} + O(U_{,j}\epsilon^3). \\ \Gamma^i_{00} &= -U_{,j} + \left[(\beta + \gamma) U^2 - 2\Psi + \frac{1}{2} \xi \mathcal{A} + \frac{1}{2} \eta \mathcal{D} + \frac{1}{2} (\alpha_1 - \alpha_2 - \alpha_3) w^2 U \right. \\ &\quad \left. + \frac{1}{2} (\alpha_1 - 2\alpha_3) w_i V_i + \frac{1}{2} \alpha_2 w_i w_k U_{ik} \right]_{,j} - \frac{7}{2} \Delta_1 V_{j,t} - \frac{1}{2} \Delta_2 W_{j,t} \\ &\quad + \left(\alpha_2 - \frac{1}{2} \alpha_1\right) w_j U_{,t} - \alpha_2 w_i U_{ij,t} + O(U_{,j}\epsilon^4), \\ \Gamma^j_{0k} &= \gamma U_{,t} \delta_{jk} - \left(\frac{7}{2} \Delta_1 + \frac{1}{2} \Delta_2\right) V_{[j,k]} - \frac{1}{2} \alpha_1 w_{[j} U_{,k]} + O(U_{,j}\epsilon^3), \\ \Gamma^j_{kl} &= -\gamma (U_{,j} \delta_{kl} - 2U_{,(k} \delta_{l)j}) + O(U_{,j}\epsilon^2).\end{aligned}\tag{39.48}$$

Here square brackets on tensor indices denote antisymmetrization, and round brackets denote symmetrization. As part of the derivation, it may be useful to prove and use the relations

$$\chi(t, \mathbf{x}) = - \int \rho_o(t, \mathbf{x}') |\mathbf{x} - \mathbf{x}'| d^3x', \tag{39.49a}$$

$$\chi_{,jk} = -\delta_{jk} U + U_{jk}, \tag{39.49b}$$

$$\chi_{,it} = V_i - W_i + O(\epsilon^5), \tag{39.49c}$$

$$W_{[k,j]} = V_{[k,j]}. \tag{39.50}$$

Here χ is the function originally defined in equation (39.29c).

Exercise 39.14. POST-NEWTONIAN APPROXIMATION TO GENERAL RELATIVITY

Perform a post-Newtonian expansion of Einstein's field equations, thereby obtaining the values cited in Box 39.2 for the PPN parameters of general relativity. The calculations might best follow the approach of Chandrasekhar (1965a): Set $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$, and assume

$$h_{00} = O(\epsilon^2) + O(\epsilon^4), \quad h_{0j} = O(\epsilon^3), \quad h_{jk} = O(\epsilon^2). \tag{39.51}$$

Choose the space and time coordinates so that the four "gauge conditions"

$$\left. \begin{aligned} h_{jk,k} - \frac{1}{2} h_{,j} &= O(\epsilon^4/R_\odot) \\ h_{0k,k} - \frac{1}{2} h_{kk,0} &= O(\epsilon^5/R_\odot) \end{aligned} \right\} \text{with } h = h_{\alpha\beta} \eta^{\alpha\beta} = -h_{00} + h_{tt} \tag{39.52}$$

are satisfied.

- (a) Show that the spatial gauge conditions are the post-Newtonian approximations to those (35.1a) used in the study of weak gravitational waves, but that the temporal gauge condition is not.

(b) Use these gauge conditions and the post-Newtonian limit in equations (8.24) and (8.47) to obtain for the Ricci tensor, accurate to linearized order,

$$R_{00} = -\frac{1}{2} h_{00,mm} + O(\epsilon^4/R_\odot^2), \quad R_{jk} = -\frac{1}{2} h_{jk,mm} + O(\epsilon^4/R_\odot^2), \quad (39.53a)$$

$$R_{0j} = -\frac{1}{2} h_{0j,mm} - \frac{1}{4} h_{00,0j} + O(\epsilon^5/R_\odot^2). \quad (39.53b)$$

(c) Combine these with the Newtonian form (39.13) of the stress-energy tensor, and with equation (39.27), to obtain the following metric coefficients, accurate to linearized order:

$$\begin{aligned} h_{00} &= 2U + k_{00} + O(\epsilon^6), & h_{0j} &= -\frac{7}{2} V_j - \frac{1}{2} W_j + O(\epsilon^5), \\ &\quad \uparrow & &\quad \text{[unknown post-Newtonian correction]} \\ h_{jk} &= 2U \delta_{jk} + O(\epsilon^4). \end{aligned} \quad (39.54)$$

Here U , V_j , and W_j are to be regarded as defined by equations (39.34a,b,c). By comparing these metric coefficients with equations (39.32), discover that

$$\gamma = 1, \quad \Delta_1 = 1, \quad \Delta_2 = 1 \quad (39.55)$$

for general relativity.

(d) With this knowledge of the metric in linearized order, one can carry out the analysis of §39.10 (using $\gamma = \Delta_1 = \Delta_2 = 1$ throughout), to obtain the post-Newtonian corrections to the stress-energy tensor [equation (39.42) with $\gamma = 1$].

(e) Calculate, similarly, the post-Newtonian corrections to the Ricci tensor component R_{00} , using $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$, using $h_{\alpha\beta}$ as given in equations (39.54), and using the gauge conditions (39.52). The answer should be

$$R_{00} = \left(-U - \frac{1}{2} k_{00} - U^2 \right)_{,mm} + 4UU_{,mm} + O(\epsilon^6/R_\odot^2). \quad (39.56)$$

(f) Evaluate the Einstein equation $R_{00} = 8\pi(T_{00} - \frac{1}{2}g_{00}T)$, accurate to post-Newtonian order, and solve it to obtain the post-Newtonian metric correction

$$k_{00} = -2U^2 + 4\Psi, \quad (39.57)$$

where Ψ is given by equations (39.34d) with $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 1$. By comparing with equations (39.32c) and (39.34d), discover that

$$\beta = \beta_1 = \beta_2 = \beta_3 = \beta_4 = 1, \quad \xi = \eta = 0 \quad (39.58)$$

for general relativity.

(g) Knowing the full post-Newtonian metric, and the full post-Newtonian stress-energy tensor, one can carry out the calculations of §39.11 (using $\gamma = \beta = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \Delta_1 = \Delta_2 = 1$, $\xi = \eta = 0$) to obtain the post-Newtonian equations of motion for the matter [equations (39.44), (39.46), and (39.47)].

§39.12. RELATION OF PPN COORDINATES TO SURROUNDING UNIVERSE

One crucial issue remains to be clarified: What is the orientation of the PPN coordinate system relative to the surrounding universe? More particularly: Does the PPN coordinate system rotate relative to the “fixed stars on the sky,” or is it “rigidly attached” to them, in some sense? In order to answer this question, imagine using the PPN formalism to analyze the solar system. Make no assumptions about the solar system’s velocity through the PPN coordinate frame. Then, as one moves outward from the Sun, past the Earth’s orbit, past Pluto’s orbit, and on out toward interstellar space, one sees the PPN coordinate frame become more and more Lorentz in its global properties [$g_{\alpha\beta} = \eta_{\alpha\beta} + O(M_\odot/r)$]. Thus, *far from the solar system the PPN coordinates become a “Lorentz frame moving through the galaxy.”* This means, of course, that the spatial axes of the PPN coordinate frame behave as though they were attached to gyroscopes far outside the solar system. Equivalently: The PPN coordinate system Fermi-Walker-transports its spatial axes through the spacetime geometry of the galaxy and universe.

Solar system’s PPN coordinate frame is attached to a local Lorentz frame of Galaxy

§39.13. SUMMARY OF PPN FORMALISM

The PPN formalism, as constructed in this chapter, is summarized in Box 39.4. Much of the recent literature uses a different set of PPN parameters than are used in this book; for a translation from one parameter set to the other, see Box 39.5.

Exercise 39.15. MANY-BODY SYSTEM IN POST-NEWTONIAN LIMIT OF GENERAL RELATIVITY

EXERCISE

Consider, in the post-Newtonian limit of general relativity, a system made up of many gravitationally interacting bodies with separations large compared to their sizes (*example:* the solar system). Idealize each body to be spherically symmetric, to be free of internal motions, and to have isotropic internal stresses, $t_{jk} = \delta_{jk} p$. Let the world line of the center of body A , in some chosen PPN coordinate frame, be $x_A(t)$; and let the (coordinate) velocity of the center of body A be

$$v_A(t) = dx_A/dt. \quad (39.59a)$$

The total mass-energy of body A as measured in its neighborhood (rest mass-energy plus internal energy plus self-gravitational energy) is given by

$$M_A = \int_{\mathcal{V}_A} \left(1 + \Pi - \frac{1}{2} U_{\text{self}} \right) d(\text{rest mass}) + O(M_A \epsilon^4), \quad (39.59b)$$

where U_{self} is the body’s own Newtonian potential (no contributions from other bodies), and \mathcal{V}_A is the interior of the body.

(continued on page 1094)

Box 39.4 SUMMARY OF THE PPN FORMALISM**I. Variables**

- $\rho_o(\mathbf{x}, t)$: baryon “mass” density (§39.3), as measured in rest frame
 $\Pi(\mathbf{x}, t)$: specific internal energy (dimensionless; §39.3), as measured in rest frame
 $t_{jk}(\mathbf{x}, t)$: components of stress referred to orthonormal axes of rest frame
 $v_j(\mathbf{x}, t)$: coordinate velocity of matter (i.e., rest frame) relative to PPN coordinates
 $U(\mathbf{x}, t), \Psi(\mathbf{x}, t), \mathcal{A}(\mathbf{x}, t), \mathcal{D}(\mathbf{x}, t), V_j(\mathbf{x}, t), W_j(\mathbf{x}, t), U_{jk}(\mathbf{x}, t)$: gravitational potentials
 $\gamma, \beta, \beta_1, \beta_2, \beta_3, \beta_4, \Delta_1, \Delta_2, \xi, \eta$: parameters whose values distinguish one theory from another (see Box 39.2)
 w : velocity of PPN coordinate frame relative to “universal rest frame” [relevant only for theories with nonzero α_1, α_2 , or α_3 ; see eq. (39.33)].

II. Equations governing evolution of these variables

- ρ_o : conservation of rest mass, equation (39.44)
 Π : first law of thermodynamics, equation (39.46)
 t_{jk} : determined in terms of ρ_o , Π , and other material variables (chemical composition, strains, etc.) by equations of state and the usual theory of a stressed medium—which is not discussed here
 v_j : equations of motion (“ $\mathbf{F} = m\mathbf{a}$ ”), equations (39.47)
 $U, \Psi, \mathcal{A}, \mathcal{D}, V_j, W_j, U_{jk}$: source equations (39.34)

III. Quantities to be calculated from these variables

- $g_{00}(\mathbf{x}, t), g_{0j}(\mathbf{x}, t), g_{jk}(\mathbf{x}, t)$: these components of metric in PPN coordinate frame are expressed in terms of gravitational potentials by equations (39.32)
 $u^0(\mathbf{x}, t), u^j(\mathbf{x}, t)$: these components of matter 4-velocity in PPN coordinate frame are given by equations (39.39)
 $T^{00}(\mathbf{x}, t), T^{0j}(\mathbf{x}, t), T^{jk}(\mathbf{x}, t)$: these components of stress-energy tensor in PPN coordinate frame are given by equations (39.42)

IV. Relation between rest frame, PPN coordinates, and the universe

1. Orthonormal basis ω^α of rest frame, where t_{jk} are defined, is related to PPN coordinate basis $d\mathbf{x}^\alpha$ by equations (39.41)
2. Far from the sun, the PPN coordinates become asymptotically Lorentz; i.e., they form an inertial frame moving through the spacetime geometry of the galaxy and the universe.
3. Gives no account of expansion of universe or of cosmic gravitational waves impinging on solar system.

Box 39.5 PPN PARAMETERS USED IN LITERATURE: A TRANSLATOR'S GUIDE

The original “point-particle version” of the PPN formalism [Nordtvedt (1968b)], and the original “perfect-fluid version” [Will (1971c)] used different sets of PPN parameters. This book has adopted Will’s set, and has added the parameter η characterizing effects of anisotropic stresses. More recently, Will and Nordtvedt have jointly adopted a revised set of parameters, described below.

A. Translation Table

<i>Will-Nordtvedt revised parameters^a</i>	<i>Revised parameters in notation of this book^b</i>	<i>Revised parameters in notation of Nordtvedt (1968b)^c</i>
γ	γ	γ
β	β	β
α_1	$7\Delta_1 + \Delta_2 - 4\gamma - 4$	$8\Delta - 4\gamma - 4$
α_2	$\Delta_2 + \xi - 1$	$\alpha''' - 1$
α_3	$4\beta_1 - 2\gamma - 2 - \xi$	$4\alpha'' - \alpha''' - 2\gamma - 1$
ξ_1	ξ	$\alpha''' - \chi$
ξ_2	$2\beta + 2\beta_2 - 3\gamma - 1$	$2\beta - \alpha' - 1$
ξ_3	$\beta_3 - 1$	absent
ξ_4	$\beta_4 - \gamma$	absent

^aRevised parameters are used by Will and Nordtvedt (1972), Nordtvedt and Will (1972), Will (1972), and Ni (1973).

^bNotation of this book is used by Will (1971a,b,c,d), Ni (1972), and Thorne, Ni, and Will (1971).

^cNordtvedt’s original “point-particle” parameters were used by Nordtvedt (1968b, 1970, 1971a,b).

B. Significance of Revised Parameters

$\alpha_1, \alpha_2, \alpha_3$ measure the extent of and nature of “preferred-frame effects”; see §39.9. Any theory of gravity with at least one α nonzero is called a *preferred-frame theory*.

$\xi_1, \xi_2, \xi_3, \xi_4, \alpha_3$ measure the extent of and nature of breakdowns in global conservation laws. A theory of gravity possesses, at the post-Newtonian level, all 10 global conservation laws (4 for energy-momentum, 6 for angular momentum; see Chapters 19 and 20) if and only if $\xi_1 = \xi_2 = \xi_3 = \xi_4 = \alpha_3 = 0$. See Will (1971d), Will and Nordtvedt (1972), Will (1972), for proofs and discussion. Any theory with $\xi_1 = \xi_2 = \xi_3 = \xi_4 = \alpha_3 = 0$ is called a *conservative theory*.

In general relativity and the Dicke-Brans-Jordan theory, all α ’s and ξ ’s vanish. Thus, general relativity and Dicke-Brans-Jordan are conservative theories with no preferred-frame effects.

(a) Show that, when written in the chosen PPN coordinate frame, this expression for M_A becomes

$$M_A = \int_{\mathcal{V}_A} \rho_o \left(1 + \Pi + \frac{1}{2} v_A^2 + 3U - \frac{1}{2} U_{\text{self}} \right) d^3x + O(M_A \epsilon^4). \quad (39.59c)$$

Use equations (39.43), (39.44), and (39.46) to show that M_A is conserved as the bodies move about, $dM_A/dt = 0$.

(b) Pick an event (t, \mathbf{x}) outside all the bodies, and at time t denote

$$\mathbf{r}_A \equiv \mathbf{x}_A - \mathbf{x}, \quad \mathbf{r}_{AB} \equiv \mathbf{x}_A - \mathbf{x}_B, \quad r_A \equiv |\mathbf{r}_A|, \quad r_{AB} \equiv |\mathbf{r}_{AB}|. \quad (39.59d)$$

Show that the general-relativistic, post-Newtonian metric (39.32) at the chosen event has the form

$$g_{ik} = \delta_{jk} \left(1 + 2 \sum_A \frac{M_A}{r_A} \right) + O(\epsilon^4), \quad (39.60a)$$

$$g_{0j} = - \sum_A \frac{M_A}{r_A} \left[\frac{7}{2} v_{Aj} + \frac{1}{2} \frac{(\mathbf{v}_A \cdot \mathbf{r}_A) r_{Aj}}{r_A^2} \right] + O(\epsilon^5), \quad (39.60b)$$

$$\begin{aligned} g_{00} = & -1 + 2 \sum_A \frac{M_A}{r_A} - 2 \left(\sum_A \frac{M_A}{r_A} \right)^2 + 3 \sum_A \frac{M_A v_A^2}{r_A} \\ & - 2 \sum_A \sum_{B \neq A} \frac{M_A M_B}{r_A r_{AB}} + O(\epsilon^6). \end{aligned} \quad (39.60c)$$

[Hint: From the Newtonian virial theorem (39.21a), applied to body A by itself in its own rest frame, conclude that

$$\int_{\mathcal{V}_A} \left(3p - \frac{1}{2} \rho_o U_{\text{self}} \right) d^3x = O(M_A \epsilon^4), \quad (39.61)$$

where the integral is performed in the PPN frame.]

(c) Perform an infinitesimal coordinate transformation,

$$t_{\text{OLD}} = t_{\text{NEW}} - \frac{1}{2} \sum_A \frac{M_A (\mathbf{r}_A \cdot \mathbf{v}_A)}{r_A}, \quad \mathbf{x}_{\text{OLD}} = \mathbf{x}_{\text{NEW}}, \quad (39.62)$$

to bring the metric (39.60) into the standard form originally devised by Einstein, Infeld, and Hoffmann (1938), and by Eddington and Clark (1938):

$$g_{jk} = \delta_{jk} \left(1 + 2 \sum_A \frac{M_A}{r_A} \right) + O(\epsilon^4), \quad (39.63a)$$

$$g_{0j} = -4 \sum_A \frac{M_A}{r_A} v_{Aj} + O(\epsilon^5), \quad (39.63b)$$

$$\begin{aligned} g_{00} = & -1 + 2 \sum_A \frac{M_A}{r_A} - 2 \left(\sum_A \frac{M_A}{r_A} \right)^2 + 3 \sum_A \frac{M_A v_A^2}{r_A} \\ & - 2 \sum_A \sum_{B \neq A} \frac{M_A M_B}{r_A r_{AB}} - \frac{\partial^2 \chi}{\partial t^2} + O(\epsilon^6), \end{aligned} \quad (39.63c)$$

where χ [equation (39.49a)] is given by

$$\chi = - \sum_A M_A r_A.$$

(d) The equations of motion for the bodies can be obtained in either of two ways: by performing a volume integral of the Euler equation (39.48) over the interior of each body; or by invoking the general arguments of §20.6. The latter way is the easier. Use it to conclude that any chosen body K moves along a geodesic of the metric obtained by omitting the terms $A = K$ from the sums in (39.63). Show that the geodesic equation for body K reduces to

$$\begin{aligned} \frac{d^2 \mathbf{x}_K}{dt^2} \equiv \frac{d\mathbf{v}_K}{dt} &= \sum_{A \neq K} \mathbf{r}_{AK} \frac{M_A}{r_{AK}^3} \left[1 - 4 \sum_{B \neq K} \frac{M_B}{r_{BK}} - \sum_{C \neq A} \frac{M_C}{r_{CA}} \left(1 - \frac{\mathbf{r}_{AK} \cdot \mathbf{r}_{CA}}{2r_{CA}^2} \right) \right. \\ &\quad \left. + \mathbf{v}_K^2 + 2\mathbf{v}_A^2 - 4\mathbf{v}_A \cdot \mathbf{v}_K - \frac{3}{2} \left(\frac{\mathbf{v}_A \cdot \mathbf{r}_{AK}}{r_{AK}} \right)^2 \right] \\ &- \sum_{A \neq K} (\mathbf{v}_A - \mathbf{v}_K) \frac{M_A \mathbf{r}_{AK} \cdot (3\mathbf{v}_A - 4\mathbf{v}_K)}{r_{AK}^3} \\ &+ \frac{7}{2} \sum_{A \neq K} \sum_{C \neq A} \mathbf{r}_{CA} \frac{M_A M_C}{r_{AK} r_{CA}^3}. \end{aligned} \quad (39.64)$$

Equations (39.63) and (39.64) are called the Einstein-Infeld-Hoffman (“EIH”) equations for the geometry and evolution of a many-body system. They are used widely in analyses of planetary orbits in the solar system. For example, the Caltech Jet Propulsion Laboratory uses them, in modified form, to calculate ephemerides for high-precision tracking of planets and spacecraft. The above method of deriving the EIH equations and metric was devised by Fock (1959). For a similar calculation in the Dicke-Brans-Jordan theory, see Estabrook (1969); and for a derivation of the analogous many-body equations in the full PPN formalism, see Will (1972).

CHAPTER 40

SOLAR-SYSTEM EXPERIMENTS

§40.1. MANY EXPERIMENTS OPEN TO DISTINGUISH GENERAL RELATIVITY FROM PROPOSED METRIC THEORIES OF GRAVITY

No audience will show up for a fight if in everyone's eyes the challenger has zero chance to win. No battle-hungry promoter desperately trying to finance the fight can afford to put into the ring against the champion any but the best contender that he can find. Against Einstein's metric theory of gravity, the judgment of the day (as §39.2 showed) leaves one no option except to put up another theory of gravity that is also metric (or metric plus torsion).

To put on a contest, then, is to design and perform an experiment that distinguishes general relativity from some not completely implausible metric theory of gravity. This chapter describes such experiments—some already performed; some to be performed in the future—and analyses their significance using the PPN formalism of Chapter 39.

In most of the experiments to be described, one investigates the motion of the moon, planets, spacecraft, light rays, or gyroscopes through the spacetime geometry of the solar system. That spacetime geometry is very complicated. It includes the spherical fields of the sun and all the planets, nonspherical fields due to their quadrupolar and higher-order deformations, and fields due to their momentum and angular momentum. Moreover, the spacetime geometry results—or at least in the post-Newtonian formalism it is viewed as resulting—from a *nonlinear* superposition of all these fields.*

This chapter analyzes experiments using PPN formalism

Complexity of solar system's spacetime geometry

*Of course, from the point of view of Einstein's full general relativity theory, all that realistically counts is the one and only curved-spacetime geometry of the real physical world. All these "individual fields" are mere bookkeepers' discourse, and they are best abandoned (they cease to be useful) when one passes from the post-Newtonian limit to the full Einstein theory.

Fortunately for this discussion, several of the most important experiments are free of almost all these complications. The effects they measure are associated entirely with the spherical part of the sun's gravitational field. A description of these experiments will come first (§§40.2–40.5), and then attention will turn to experiments that are more complex in principle.

To discuss central-field experiments, one needs an expression for the external gravitational field of an idealized, isolated, static, spherical sun. In general relativity, such a gravitational field is described by the Schwarzschild line element,

$$ds^2 = - \left(1 - \frac{2M_\odot}{r}\right) dt^2 + \frac{dr^2}{1 - 2M_\odot/r} + r^2(d\theta^2 + \sin^2\phi d\phi^2).$$

Idealization of geometry to that of isolated, static, spherical sun:

(1) in Schwarzschild coordinates

But this line element is not what one wants, for two reasons: (1) it is “too accurate”; (2) it is written in the “wrong” coordinate system.

Why too accurate? Because it is simple only when unperturbed and unmodified; whereas some modified theories show up new effects that are so complex they are tractable only in the post-Newtonian approximation. Why wrong coordinate system? Because physicists, astronomers, and other celestial mechanics have adopted the fairly standard convention of using “isotropic coordinates” rather than “Schwarzschild coordinates” when analyzing the solar system. Example: post-Newtonian expansions, including the PPN formalism of Chapter 39, almost always use isotropic coordinates. Another example: the relativistic ephemeris for the solar system, prepared by the Caltech Jet Propulsion Laboratory [Ohandley *et al.* (1969); Anderson (1973)] and used extensively throughout the world, employs isotropic coordinates.

Modify the Schwarzschild line element, then. First transform to isotropic coordinates (Exercise 31.7); then expand the metric coefficients in powers of M_\odot/r , to post-Newtonian accuracy. Thereby obtain

$$\begin{aligned} ds^2 &= - \left[1 - 2\frac{M_\odot}{r} + 2\left(\frac{M_\odot}{r}\right)^2\right] dt^2 + \left[1 + 2\frac{M_\odot}{r}\right] [dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)] \\ &= - \left[1 - 2\frac{M_\odot}{r} + 2\left(\frac{M_\odot}{r}\right)^2\right] dt^2 + \left[1 + 2\frac{M_\odot}{r}\right] [dx^2 + dy^2 + dz^2]. \end{aligned} \tag{40.1}$$

(2) in isotropic coordinates

Here r, θ, ϕ are related to x, y, z in the usual manner:

$$r = (x^2 + y^2 + z^2)^{1/2}, \quad \theta = \tan^{-1}[z/(x^2 + y^2)^{1/2}], \quad \phi = \tan^{-1}(y/x); \tag{40.2}$$

and r is the new, “isotropic” radial coordinate, not to be confused with the Schwarzschild r . (The reader who has not studied §39.6 will discover in the next section why one keeps terms of order M_\odot^2/r^2 in g_{00} but not in g_{jk} .) Note: this post-Newtonian expression for the metric is a special case of the result derived in exercise 19.3.

If one calculates the gravitational field of the same source (the sun) in the same post-Newtonian approximation in other metric theories of gravity, one obtains a very similar result:

(3) in PPN formalism

$$\begin{aligned} ds^2 &= - \left[1 - 2 \frac{M_\odot}{r} + 2\beta \left(\frac{M_\odot}{r} \right)^2 \right] dt^2 + \left[1 + 2\gamma \frac{M_\odot}{r} \right] [dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)] \\ &= - \left[1 - 2 \frac{M_\odot}{r} + 2\beta \left(\frac{M_\odot}{r} \right)^2 \right] dt^2 + \left[1 + 2\gamma \frac{M_\odot}{r} \right] [dx^2 + dy^2 + dz^2] \end{aligned} \quad (40.3)$$

(see exercise 40.1). Here γ and β are two of the ten PPN parameters described in Box 39.2. Recall from that box that γ measures “the amount of space curvature produced by unit rest mass,” while β measures “the amount of nonlinearity in the superposition law for g_{00} .” These heuristic descriptions find their mathematical counterparts in the above form for the idealized metric surrounding a spherically symmetric center of attraction.

By measuring the parameter γ to high precision, one can distinguish between general relativity ($\gamma = 1$) and the Dicke-Brans-Jordan theory [$\gamma = (1 + \omega)/(2 + \omega)$, where ω is the “Dicke coupling constant”]; see Box 39.2. But general relativity and Dicke-Brans-Jordan predict the same value for β ($\beta = 1$). This identity does not mean that β is unworthy of measurement. A value $\beta \neq 1$ is predicted by other theories [see Ni (1972)]; so measurements of β are useful in distinguishing such theories from general relativity.

Actually, the above form (40.3) for the sun’s metric is not fully general. In any theory with a preferred “universal rest frame” (e.g., Ni’s theory; Box 39.1), there are additional terms in the metric due to motion of the sun relative to that preferred frame (exercise 40.1):

(4) including preferred-frame effects

$$\begin{aligned} ds^2 &= (\text{expression 40.3}) + (\alpha_2 + \alpha_3 - \alpha_1) \frac{M_\odot}{r} w^2 dt^2 + 2 \left(\alpha_2 - \frac{1}{2} \alpha_1 \right) \frac{M_\odot}{r} w_j dx^j dt \\ &\quad - \alpha_2 \left[\frac{M_\odot}{r^3} x^j x^k - \frac{I_\odot}{r^5} \left(x^j x^k - \frac{1}{3} r^2 \delta_{jk} \right) \right] w_j dt (2 dx^k + w_k dt). \end{aligned} \quad (40.3')$$

In these “preferred-frame terms” $I_\odot \equiv I_{jj} = \int \rho r^2 d^3x$ is the trace of the second moment of the sun’s mass distribution;

$$\begin{aligned} \alpha_1 &= 7\Delta_1 + \Delta_2 - 4\gamma - 4, \\ \alpha_2 &= \Delta_2 + \zeta - 1, \\ \alpha_3 &= 4\beta_1 - 2\gamma - 2 - \zeta \end{aligned}$$

are combinations of PPN parameters; and w is the sun’s velocity (\equiv velocity of coordinate system) relative to the preferred frame. (Theories such as general relativity and Dicke-Brans-Jordan, which possess no preferred frame, have $\alpha_1 = \alpha_2 = \alpha_3 = 0$, and therefore have no preferred-frame terms in the metric.) For ease of exposition, all equations and calculations in this chapter will ignore the preferred-frame terms; but the consequences of those terms will be discussed and references analyzing them will be cited.

Exercise 40.1. PPN METRIC FOR IDEALIZED SUN [Track 2]

Show that for an isolated, static, spherical sun at rest at the origin of the PPN coordinate system, the PPN metric (39.32) reduces to expressions (40.3), (40.3'). As part of the reduction, show that the sun's total mass-energy is given by

$$M_{\odot} = \int_0^{R_{\odot}} \rho_0 (1 + 2\beta_2 U + \beta_3 \Pi + 3\beta_4 p/\rho_0) 4\pi r^2 dr. \quad (40.4)$$

[Warning: One must not look at this formula and immediately think: "The contribution of rest mass is $\int \rho_0 4\pi r^2 dr$, the contribution of gravitational energy is $\int 2\beta_2 \rho_0 U 4\pi r^2 dr$, etc." Rather, in making any such interpretation one must remember that (1) spacetime is curved, so $4\pi r^2 dr$ is not *proper volume* as measured by physical meter sticks; also (2) virial theorems (exercise 39.6) and other integral theorems can be used to change the form of the integrand. For further discussion see exercises 40.9 and 40.10 below.]

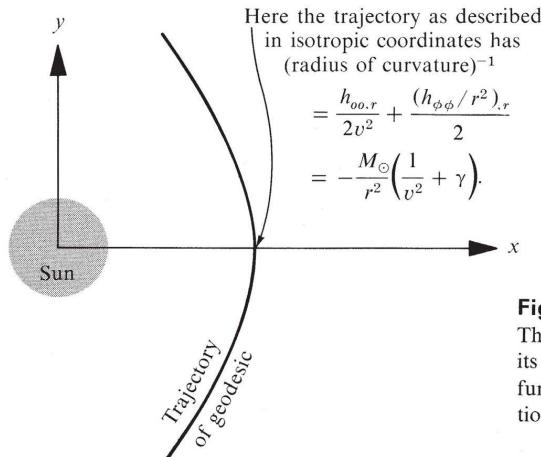
EXERCISES**§40.2. THE USE OF LIGHT RAYS AND RADIO WAVES
TO TEST GRAVITY**

In the Newtonian limit, planetary and spacecraft orbits are strongly influenced by gravity; but light propagation and radio-wave propagation (at "infinite" velocity) are not influenced at all. For this reason, experimental studies of orbits are beset by the problem of separating the relativistic effects from much larger standard Newtonian effects. By contrast, experimental studies of light and radio-wave propagation do not contend with any such overpowering Newtonian background. Not surprisingly, they are to date (1973) the clearest and most definitive of the solar-system experiments.

Light rays and radio waves give "clean" tests of relativity

Mathematically, the parameter that distinguishes a light ray from a planet is its high speed. In the geodesic equation, the magnitude of the velocity determines which metric coefficients can influence the motion. Consider, for example, a weak, static field $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$, and a particle at $(x, y, z) = (r, 0, 0)$ moving with velocity $(v_x, v_y, v_z) = (0, v, 0)$; see Figure 40.1. Here the effect of gravity on the trajectory of the particle can be characterized by the quantity

$$\begin{aligned} & \left(\text{curvature of trajectory in 3-dimensional,} \right) = \left(\text{radius of curvature} \right)^{-1} \\ & \text{nearly Euclidean, space} \quad \text{of trajectory} \\ & = \frac{d^2 x}{dy^2} = \frac{d\tau}{dy} \frac{d}{d\tau} \left(\frac{d\tau}{dy} \frac{dx}{d\tau} \right) = \frac{1}{u^y} \frac{d}{d\tau} \left(\frac{u^x}{u^y} \right) = \frac{1}{(u^y)^2} \frac{du^x}{d\tau} \\ & = - \frac{(1 - v^2)}{v^2} \Gamma^x_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} = - \frac{1}{v^2} \Gamma^x_{\alpha\beta} \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt} \\ & = - \Gamma^x_{00} v^{-2} - 2\Gamma^x_{0y} v^{-1} - \Gamma^x_{yy} \\ & = \frac{1}{2} h_{00,x} v^{-2} + (h_{0y,x} - h_{0x,y}) v^{-1} + \left(\frac{1}{2} h_{yy,x} - h_{xy,y} \right). \end{aligned}$$

**Figure 40.1.**

The bending of the trajectory of a test body at its point of closest approach to the sun, as a function of its 3-velocity. (See text for computation and discussion.)

Reexpressed in spherical coordinates, in the terminology of the idealized solar line element (40.3), this formula says

$$\begin{aligned} \left(\text{curvature of trajectory} \right)_{\text{in 3-space}} &= \frac{1}{2} h_{00,r} v^{-2} + \frac{1}{2} (h_{\phi\phi}/r^2)_{,r} \\ &\approx -(M_\odot/r^2)(v^{-2} + \gamma) \end{aligned} \quad (40.5)$$

for a particle at its point of closest approach to the sun. (Compare with exercise 25.21.) Note that here γ is a PPN parameter; it is *not* $(1 - v^2)^{-1/2}$.

Notice what happens as one boosts the velocity of the particle. For slow velocities [$v^2 \sim$ (post-Newtonian expansion parameter ϵ^2) $\approx M_\odot/R_\odot$], the Newtonian part of h_{00} dominates completely; and the tiny post-Newtonian corrections come equally from the ϵ^4 part of h_{00} , the ϵ^3 part of h_{0j} , and the ϵ^2 part of h_{jk} . [This was the justification for expanding h_{00} to $O(\epsilon^4)$, h_{0j} to $O(\epsilon^3)$, and h_{jk} to $O(\epsilon^2)$ in the post-Newtonian limit; see §39.6.] But as v increases, the ordering of the terms changes. In the high- v regime ($v \sim 1 \gg \epsilon^2$), the bending of the trajectory has become almost imperceptible because of the high forward momentum of the particle and the short time it receives transverse momentum from the sun. What bending is left is due to the ϵ^2 (Newtonian) part of h_{00} , and the ϵ^2 (post-Newtonian) part of h_{jk} . Nothing else can have a significant influence. Notice, moreover, that—even when one allows for “preferred-frame” effects—these dominant terms,

$$h_{00} = 2U = 2M_\odot/r \text{ and } h_{jk} = 2\gamma U \delta_{jk} = 2\gamma(M_\odot/r) \delta_{jk},$$

depend only on the Newtonian potential $U \equiv -\Phi$ and the PPN parameter γ .

This is a special case of a more general result: *Aside from fractional corrections of $\epsilon^2 \lesssim 10^{-6}$, relativistic effects on light and radio-wave propagation are governed entirely by the Newtonian potential U and the PPN parameter γ .* These relativistic effects include the gravitational redshift (discussed in the last chapter; independent

Light rays are governed solely by Newtonian potential and PPN parameter γ

of γ), the gravitational deflection of light and radio waves (discussed below; dependent on γ), and the “relativistic time-delay” (discussed below; dependent on γ).

§40.3. “LIGHT” DEFLECTION

Consider a light or radio ray coming into a telescope on Earth from a distant star or quasar. Do not assume, as in the usual discussion (exercises 18.6 and 25.24), that the ray passes near the sun. The deflection by the sun’s gravitational field will probably be measurable, in the middle or late 1970’s, even when the ray passes far from the sun! [The calculation that follows is due to Ward (1970), but Shapiro (1967) first derived the answer.]

Orient the PPN spherical coordinates of equation (40.3) so that the ray lies in the “plane” $\theta = \pi/2$. By symmetry, if it starts out in this plane far from the Earth, it must lie in this plane always. Let the incoming ray enter the solar system along the line $\phi = 0$; and let the Earth be located at $r = r_E$, $\phi = \phi_E$ when the ray reaches it. (See Figure 40.2.) One wishes to calculate the angle α between the incoming light ray and the center of the sun, as measured in the orthonormal frame $(\mathbf{e}_r, \mathbf{e}_\phi)$ of an observer on Earth. If the sun had zero mass (flat, Euclidean space), α would be $\pi - \phi_E$ (see Figure 40.2). However, the sun produces a deflection: $\alpha = \pi - \phi_E + \delta\alpha$. The deflection angle $\delta\alpha$ is the true objective of the calculation.

In the calculation, ignore the Earth’s orbital and rotational motions. They lead to aberration, for which correction can be made by the usual formula of special relativity (Lorentz transformation in the neighborhood of the telescope.) Also ignore deflection of the light ray due to the Earth’s gravitational field (deflection angle $\sim 2M_E/R_E \sim 0''.0003$), which might be detectable in the late 1970’s.

Light deflection:

(1) derivation

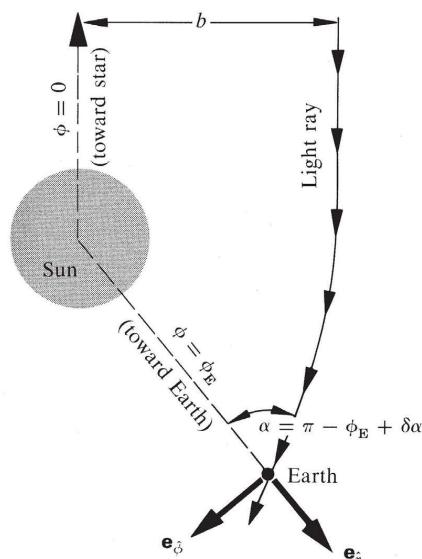


Figure 40.2.

Coordinates used in the text for calculating the deflection of light. Notice that in this diagram ϕ increases in the clockwise direction.

As the first step in calculating the deflection angle, determine the trajectory of the ray in the r, ϕ -plane. This can be calculated either using the geodesic equation, or using the eikonal method of geometric optics (Hamilton-Jacobi method; §22.5 and Box 25.4). The result of such a calculation (exercise 40.2) is an equation connecting r with ϕ ; thus,

$$\frac{b}{r} = \sin \phi + \frac{(1 + \gamma)M_{\odot}}{b}(1 - \cos \phi). \quad (40.6)$$

Notice that b has a simple geometric interpretation: far from the sun, the ray trajectory is $\phi = b/r + O(M_{\odot}b/r^2)$. Consequently, b is the impact parameter in the usual sense of classical scattering theory (see Figure 40.2). The ray makes its closest approach to the sun (assuming it is not intercepted by the Earth first) at the PPN coordinate radius

$$r_{\min} = b \left[1 - \frac{(1 + \gamma)M_{\odot}}{b} \right] \approx b. \quad (40.7)$$

Thus, b can also be thought of as the radius of the ray's "perihelion."

Notice that the ray returns to $r = \infty$, not at an angle $\phi = \pi$, but rather at

$$\phi(r = \infty) = \pi + 2(1 + \gamma)M_{\odot}/b. \quad (40.8a)$$

Thus, the total deflection angle is

$$\begin{aligned} (\text{angle of total deflection}) &= 2(1 + \gamma)M_{\odot}/b \\ &= \frac{1}{2}(1 + \gamma)1''.75 \text{ for a ray that} \\ &\quad \text{just grazes the sun.} \end{aligned} \quad (40.8b)$$

But this is not the quantity of primary interest. Rather, one seeks the position of the star as seen by an astronomer on Earth. The angle $\alpha = \pi - \phi_E + \delta\alpha$ between the sun and the star as measured by the astronomer is given by (see Figure 40.2)

$$\begin{aligned} \tan(\pi - \phi_E + \delta\alpha) &= -\tan \phi_E + \delta\alpha/\cos^2 \phi_E \\ &= \frac{u^{\hat{\phi}}}{u^{\hat{r}}} = \left[\frac{(1 + \gamma M_{\odot}/r)r d\phi/d\lambda}{(1 + \gamma M_{\odot}/r) dr/d\lambda} \right]_E = \left[\frac{r d\phi}{dr} \right]_E \\ &= -\left[\frac{(b/r) d\phi}{d(b/r)} \right]_E, \end{aligned} \quad (40.9)$$

where $u^{\beta} = dx^{\beta}/d\lambda$ are the components of a tangent to the ray at the Earth. By inserting into this equation expression (40.6) for the trajectory of the ray, one obtains

$$\begin{aligned} \tan \phi_E - \frac{\delta\alpha}{\cos^2 \phi_E} &= \frac{\sin \phi_E + [(1 + \gamma)M_{\odot}/b](1 - \cos \phi_E)}{\cos \phi_E + [(1 + \gamma)M_{\odot}/b] \sin \phi_E} \\ &= \tan \phi_E - [(1 + \gamma)M_{\odot}/b](1 - \cos \phi_E)/\cos^2 \phi_E. \end{aligned} \quad (40.10)$$

Thus, the deflection angle measured at the Earth is

$$\delta\alpha = \frac{(1 + \gamma)M_{\odot}}{b}(1 + \cos\alpha) = \frac{(1 + \gamma)M_{\odot}}{r_E} \left(\frac{1 + \cos\alpha}{1 - \cos\alpha} \right)^{1/2}. \quad (40.11)$$
(2) formula for deflection angle

It ranges from zero when the ray comes in opposite to the sun's direction ($\alpha = \pi$), through the value

$$(1 + \gamma)M_{\odot}/r_E = \frac{1}{2}(1 + \gamma)0''.0041 \quad (40.12)$$

when the ray comes in perpendicular to the Earth-Sun line ($\alpha = \pi/2$), to the "classical value" of $\frac{1}{2}(1 + \gamma) \times 1''.75$ when the ray comes in grazing the sun's limb.

All experiments to date (1972) have examined the case of grazing passage. The experimental results are stated and discussed in Box 40.1. They show that the PPN parameter γ has its general relativistic value of 1 to within an uncertainty of about 20 percent.

By the middle or late 1970's, measurements of the deflection of radio waves from quasars should determine γ to much better than 1 percent. Also, by that time radio astronomers may be making progress toward setting up high-precision coordinates on the sky using very long baseline interferometry. If so, they will have to use equation (40.11) to compensate for the "warping" of the coordinates caused by the sun's deflection of radio waves in all regions of the sky, not just near the solar limb.

Experimental measurements
of deflection

Exercise 40.2. TRAJECTORY OF LIGHT RAY IN SUN'S GRAVITATIONAL FIELD

EXERCISE

Derive equation (40.6) for the path of a light ray in isotropic coordinates (40.3) in the sun's "equatorial plane." Use one or more of three alternative approaches: (1) direct integration of the geodesic equation (the hardest approach!); (2) computation based on the three integrals of the motion

$$\mathbf{k} \cdot \mathbf{k} = 0, \quad \mathbf{k} \cdot (\partial/\partial t) = k_0, \quad \mathbf{k} \cdot (\partial/\partial\phi) = k_\phi = -bk_0;$$

$\mathbf{k} \equiv d/d\lambda$ = tangent vector to geodesic

(see §§25.2 and 25.3); (3) computation based on the Hamilton-Jacobi method (Box 25.4), which for photons (zero rest mass) reduces to the "eikonal method" of geometric optics (see §22.5).

§40.4. TIME-DELAY IN RADAR PROPAGATION

Another effect of spacetime curvature on electromagnetic waves is a relativistic delay in the round-trip travel time for radar signals. It was first pointed out by Shapiro (1964); see also Muhleman and Reichley (1964, 1965).

Radar time delay:

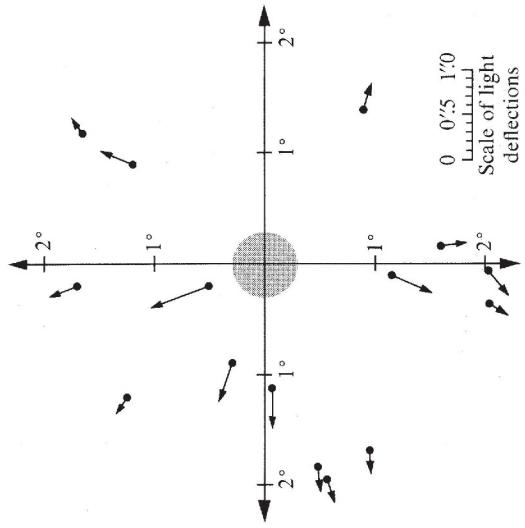
(continued on page 1106)

Box 40.1 DEFLECTION OF LIGHT AND RADIO WAVES BY SUN: EXPERIMENTAL RESULTS

Eclipse Measurements

Until 1968 every experiment measured the deflection of starlight during total eclipse of the sun. The measurements were beset by difficulties such as poor weather, optical distortions due to temperature changes, and the strange propensity of eclipses to attain maximum time of totality in jungles, in the middles of oceans, in deserts, and in arctic tundras. Lists of all the results and references are given by Bertotti, Brill and Krotkov (1962), and by Klüber (1960). Dicke (1964b) summarizes the results as follows:

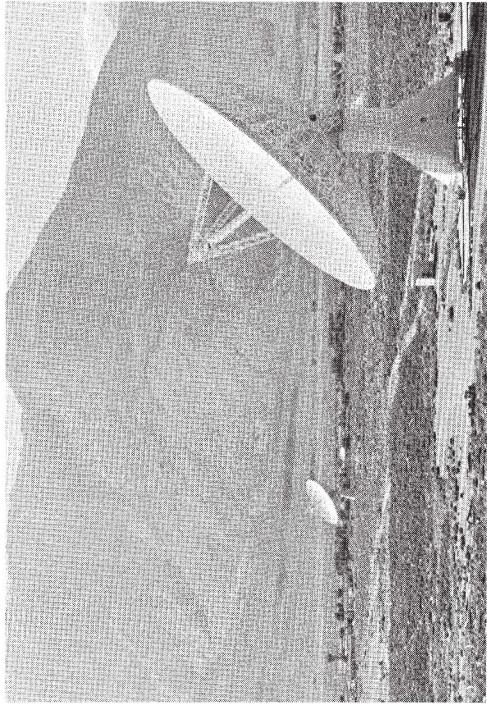
"The analyses [of the experimental data] scatter from a deflection at the limb of the sun of 1.43 seconds of arc to 2.7 seconds [compared to a general relativistic value of 1.75 seconds]. The scatter would not be too bad if one could believe that the technique was free of systematic errors. It appears that one must consider this observation uncertain to at least 10 percent, and perhaps as much as 20 percent." This result corresponds to an uncertainty in γ of 20 to 40 percent.



Measurements on the Deflection of Radio Waves

Each October 8 the sun, as seen from the Earth, passes in front of the quasar 3C279. By monitoring the angular separation between 3C279 and a nearby quasar 3C273, radio astronomers can measure the deflection by the sun of the 3C279 radio waves. The monitoring uses radio interferometers. [See references cited in table for discussion of the technique.] Technology of the early 1970's should permit measurements to a precision 0.001 seconds of arc or better, if the two ends of the interferometer are separated by several thousand kilometers ("transcontinental" or "transworld" baseline). But as of 1971 the only successful experiments were less ambitious; they used baselines of less than 10 kilometers. A summary of these pre-1971, short-baseline results is shown in the table.

Observed light deflections (mean of two instruments) of the 15 best measured stars within $2^{\circ}5'$ of the sun's center in the total solar eclipse of September 21, 1922 at Wallal, Western Australia, as determined by Campbell and Trumpler (1928). The arrows represent in size and direction the observed light deflections relative to the reference stars (5° to 10° from the sun's center). (See Box 1.6 for Einstein's description of the deflection in terms of the curvature of geometry near the sun).

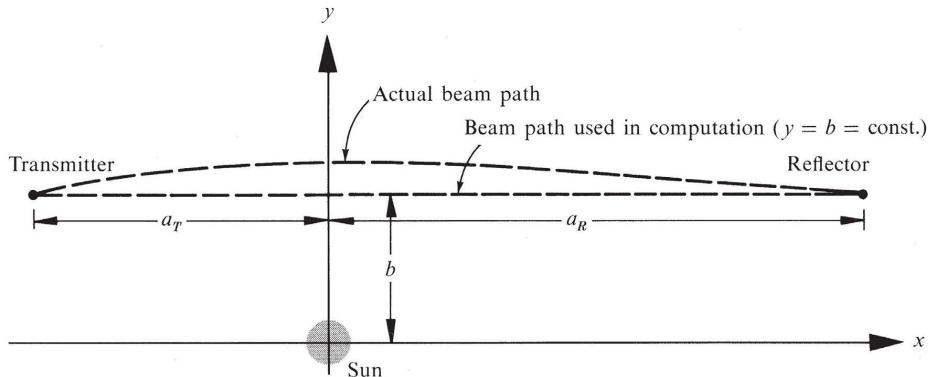


The 90-foot (background) and 130-foot (foreground) radio interferometer system at Caltech's Owens Valley Radio Observatory. These were used by Seielstad, Sramek, and Weiler (1970) in their pioneering measurement of the deflection of quasar radio waves by the sun. During the experiment the two antennas were separated by 1.07 kilometers. (Photo by Alan Moffet.)

Experimental results^a

Dates of observation	Observatory	Experimenters and reference	Number of telescopes and separations	Wave lengths	$\frac{1}{2}(1 + \gamma) = \frac{\text{(Observed deflection)}}{\text{(Einstein prediction)}}$	Formal standard error	One-sigma error
Sept. 30-Oct. 15 1969	Owens Valley (Caltech)	Seielstad, Sramek, Weiler (1970)	2, 1.07 km	3.1 cm	1.01	±0.12	±0.12
Oct. 2-Oct. 10 1969	Goldstone (Caltech-JPL)	Muhleman, Ekers, Fomalont (1970)	2, 21.56 km	12.5 cm	1.04	±0.05	+0.15 -0.10
Oct. 2-Oct. 12 1970	National Radio Astronomy Observatory (USA)	Sramek (1971)	3, 0.80 km, 1.90 km, 2.70 km	11.1 cm, 3.7 cm	0.90	±0.05	±0.05
Sept. 30-Oct. 15 1970	Mullard Radio Astronomy Observatory (Cambridge Univ.)	Hill (1971)	3, 0.66 km, 1.41 km	11.1 cm, 6.0 cm	1.07	±0.17	±0.17

^aHere (observed deflection)/(Einstein prediction) is the number $\frac{1}{2}(1 + \gamma)$ obtained by fitting the observational data to the PPN prediction (40.11). [For these experiments the ray passes near the solar limb; so (40.11) reduces to $\delta\alpha = \frac{1}{2}(1 + \gamma)(M_{\odot}/b)$.] The "formal standard error" is obtained from the data by standard statistical techniques. However, it is not usually a good measure of the certainty of the result, because it fails to take account of systematic errors. The quoted "one-sigma error" is the experimenters' best estimate of the combined statistical and systematic uncertainties. The experimenters estimate a probability of 68 percent that the true result is within "1σ" of their measured value; a probability of 95 percent that it is within "2σ"; etc.

**Figure 40.3.**

Diagram, in the PPN coordinate system, for the calculation of the relativistic time delay.

- (1) foundations for calculation; Fermat's principle

Let a radar transmitter on Earth send a radio wave out to a reflector elsewhere in the solar system, and let the reflector return the wave to Earth. Calculate the round-trip travel time, as measured by a clock on Earth. For simplicity of calculation, idealize both Earth and reflector as nonrotating and as at rest in the static, spherical gravitational field of the Sun. At the end of the calculation, the effects of rotation and motion will be discussed separately. Also ignore time dilation of the transmitter's clock due to the Earth's gravitational field; it is easily corrected for, and it is so small that it will not come into play in these radar experiments before the middle or late 1970's. The gravitational effects of the other planets on the radio waves are too small to be discernible in the foreseeable future, unless the beam grazes the limb of one of them. However, the effects of dispersion in the solar wind and corona are discernible and must be corrected for. These corrections will not be discussed here, since they are free of any general-relativistic influence.

The calculation of the round-trip travel time can be simplified by using a general-relativistic version of Fermat's principle: *In any static field ($g_{0j} = 0$, $g_{\alpha\beta,0} = 0$) consider all null curves between two points in space, $x^j = a^j$ and $x^j = b^j$. Each such null curve, $x^j(t)$, requires a particular coordinate time interval Δt to get from a^j to b^j . The curves of extremal Δt are the null geodesics of spacetime.* The proof of this theorem is outlined in exercise 40.3.

Because of Fermat's principle, the lapse of coordinate time between transmission of the radar beam and reflection at the reflector, t_{TR} , is the same for a straight path in the PPN coordinates as for the slightly curved path which the beam actually follows. (The two differ by a fractional amount $\Delta t_{TR}/t_{TR} \sim (\text{angle of deflection})^2 \lesssim 10^{-12}$, which is far from discernable.) Hence, in the computation one can ignore the gravitational bending of the beam.

- (2) details of calculation

Adopt Cartesian PPN coordinates with the sun at the origin; the transmitter, sun, and reflector in the $z = 0$ "plane"; and the transmitter-reflector line along the x direction (see Figure 40.3). The transmitter is at $(x, y) = (-a_T, b)$ in the PPN coordinates, and the reflector is at $(x, y) = (a_R, b)$. Recall that for a null ray $ds^2 = 0 =$

$g_{00} dt^2 - g_{xx} dx^2$. It follows that the lapse of coordinate time between transmission and reflection is

$$t_{\text{TR}} = \int_{-a_T}^{a_R} \left(\frac{g_{xx}}{-g_{00}} \right)^{1/2} dx = \int_{-a_T}^{a_R} \left[1 + \frac{(1+\gamma)M_\odot}{\sqrt{x^2 + b^2}} \right] dx.$$

Integration yields

$$t_{\text{TR}} = a_R + a_T + (1+\gamma)M_\odot \ln \left[\frac{(a_R + \sqrt{a_R^2 + b^2})(a_T + \sqrt{a_T^2 + b^2})}{b^2} \right]. \quad (40.13)$$

The lapse of coordinate time in round-trip travel has twice this magnitude. The lapse of proper time measured by an Earth-based clock is

$$\begin{aligned} \Delta\tau &= |g_{00}|_{\text{Earth}}^{1/2} 2t_{\text{TR}}, \\ \Delta\tau &= 2(a_R + a_T) \left(1 - \frac{M_\odot}{\sqrt{a_R^2 + b^2}} \right) \\ &\quad + 2(1+\gamma)M_\odot \ln \left[\frac{(a_R + \sqrt{a_R^2 + b^2})(a_T + \sqrt{a_T^2 + b^2})}{b^2} \right] \end{aligned} \quad (40.14)$$

This is the lapse of time on an Earth-based clock, aside from corrections for the orbital and rotational motion of the Earth, for the orbital motion of the reflector, for dispersion of radiation traversing the solar wind and corona, and for time dilation in the Earth's gravitational field.

Any reader is reasonable who objects to the form (40.14) in which the time-delay has been written. The quantities a_R , a_T , and b are coordinate positions in the PPN coordinate system, rather than numbers the astronomer can measure directly. They differ from coordinate positions in other, equally good coordinate systems by amounts of the order of $M_\odot \sim 1.5$ km. The objection is not mathematical in its origin. The quantities a_R , a_T , and b are perfectly well-defined [with post-post-Newtonian uncertainties of order $b(M_\odot/b)^2 \lesssim 10^{-6}$ km], because the PPN coordinate system is perfectly well-defined. But they are not quantities which the experimenter can measure directly, with precision anywhere near that required to test the relativistic terms in the time-delay formula (40.14).

In practice, fortunately, the experimenter does not need to measure a_R , a_T , or b with high precision. Instead, he checks the time-delay formula by measuring the changes in $\Delta\tau$ as the Earth and reflector move in their orbits about the Sun; i.e., he measures $\Delta\tau$ as a function of Earth-based time τ . Notice that when the beam is passing near the sun ($b \ll a_R$, $b \ll a_T$; but $db/d\tau \gg da_R/d\tau$ and $db/d\tau \gg da_T/d\tau$ because the Earth's and reflector's orbits are nearly circular), the change of b in the \ln term of (40.14) dominates all other relativistic corrections to the Newtonian delay; consequently (using $db/d\tau \sim 10$ km/sec for typical experiments)

$$\begin{aligned} \frac{d\Delta\tau}{d\tau} - (\text{Constant Newtonian part}) &\approx -4(1+\gamma) \frac{M_\odot}{b} \frac{db}{d\tau} \\ &\sim 4(1+\gamma) \left(\frac{1.5 \text{ km}}{10^6 \text{ km}} \right) \left(\frac{10 \text{ km}}{\text{sec}} \right) \sim \frac{30 \mu\text{sec}}{\text{day}}. \end{aligned} \quad (40.15)$$

(3) formula for delay

(4) comparison with experiment

Such differential shifts in round-trip travel time—which rise as the Earth-reflector line moves toward the Sun and falls as it moves away—are readily observable.

In practice, in order to obtain precisions better than about 20 percent in the determination of the parameter γ by time-delay measurements, one must carefully collect and analyze data for a large fraction of a year—from a time when the beam is far from the sun ($b \sim a_T \sim 10^8$ km), to the time of superior conjunction ($b \sim R_\odot \sim 10^6$ km), and on around to a time of distant beam again. Such a long “arc” of data is needed to determine the reflector’s orbit with high precision, and to take full advantage of the slow, logarithmic falloff of $\Delta\tau$ with b (40.14). When the beam is far from the sun ($b \gg R_\odot$), the simplifying assumptions behind equation (40.15) are not valid; and the relativistic time-delay gets intertwined with the orbital motions of the Earth and the reflector. The analysis then remains straightforward, but its details are so complex that one resorts to numerical integrations on a computer to carry it out. Because the orbital motions enter, the time-delay data then contain information about other metric parameters (β is the dominant one) in addition to γ .

The experimental results as of 1971 are described in Box 40.2. They yield a value for the PPN parameter γ that is more accurate than the value from light and radio-wave deflection experiments:

(5) experimental result for γ

$$\gamma = 1.02 \pm 0.08. \quad (40.16)$$

Future experiments using spacecraft may improve the precision of γ to ± 0.001 or better.

EXERCISE

Exercise 40.3. FERMAT’S PRINCIPLE

Prove Fermat’s principle for a static gravitational field. [Hint: The proof might proceed as follows. Write down the geodesic equation in four-dimensional spacetime using an affine parameter λ . Convert from the parameter λ to coordinate time t , and use $ds^2 = 0$ to obtain

$$g_{jk} \frac{d^2x^k}{dt^2} + \Gamma_{jkl} \frac{dx^k}{dt} \frac{dx^l}{dt} - \Gamma_{j00} \frac{g_{kl}}{g_{00}} \frac{dx^k}{dt} \frac{dx^l}{dt} + \frac{d^2t/d\lambda^2}{(dt/d\lambda)^2} g_{jk} \frac{dx^k}{dt} = 0.$$

Combine with the time part of the geodesic equation

$$\frac{d^2t/d\lambda^2}{(dt/d\lambda)^2} = -2\Gamma_{0k0} \frac{dx^k/dt}{g_{00}}$$

and use the expression for the Christoffel symbols in terms of the metric to obtain

$$\gamma_{jk} \frac{d^2x^k}{dt^2} + \frac{1}{2} (\gamma_{jk,l} + \gamma_{jl,k} - \gamma_{kl,j}) \frac{dx^k}{dt} \frac{dx^l}{dt} = 0, \quad \gamma_{jk} \equiv -\frac{g_{jk}}{g_{00}}.$$

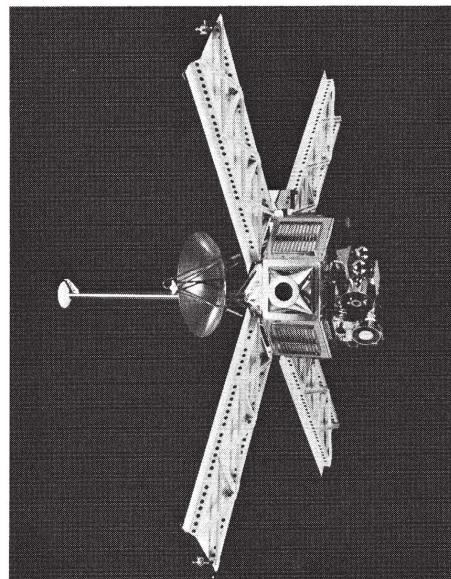
Then notice that this is a geodesic equation with affine parameter t in a three-dimensional manifold with metric γ_{jk} . The familiar extremum principle for this geodesic is

$$\delta \int_{a^j}^{b^j} (\gamma_{jk} dx^j dx^k)^{1/2} dt = \delta \int_{a^j}^{b^j} dt = 0,$$

which is precisely Fermat’s principle!]

**Box 40.2 RADAR TIME DELAY IN THE SOLAR SYSTEM:
EXPERIMENTAL RESULTS**

Two types of experiments have been performed to measure the relativistic effects [proportional to $\frac{1}{2}(1 + \gamma)$; equation (40.14)] in the round-trip radar travel time in the solar system. In one type ("passive" experiment) the reflector is the surface of the planet Venus or the planet Mercury. In the other type ("active" experiment) the "reflector" is electronic equipment on board a spacecraft that receives the signal and transmits it back to Earth ("transponder"). Passive experiments suffer from noise due to topography of the reflecting planet (earlier radar return from mountain tops than from valley floors), and they suffer from weakness of the returned signal. Active experiments suffer from buffeting of the spacecraft by solar wind, buffeting by fluctuations in solar radiation pressure, and buffeting by leakage from gas jets ("outgassing"). Experiments of the future will solve these problems by placing a transponder on the surface of a planet or on a "drag-free" (buffeting-free) spacecraft. But experiments of the present and future must both contend with fluctuating time delays due to dispersion in the fluctuating solar wind and corona. Fortunately, these are smaller than the relativistic effects, except when



The Mariner VI spacecraft (mock-up), which was the reflector in a 1970 measurement of $\frac{1}{2}(1 + \gamma)$ by radar time delay [photo courtesy the Caltech Jet Propulsion Laboratory].

the beam passes within 2 or 3 solar radii of the sun.

The results of experiments performed before 1972 are listed in the table.

<i>Dates of observation</i>	<i>Radar telescopes</i>	<i>Reflector</i>	<i>Experimenters and reference</i>	<i>Experimental result^a</i>			
				<i>Wave length</i>	$\frac{(\text{Observed delay})}{(\text{Einstein prediction})}$	<i>Formal standard error</i>	<i>One-sigma error</i>
November 1966 to August 1967	Haystack (MIT)	Venus and Mercury	Shapiro (1968)	3.8 cm	0.9		± 0.2
1967 through 1970	Haystack (MIT), and Arecibo (Cornell)	Venus and Mercury	Shapiro, Ash, <i>et al.</i> (1971)	3.8 cm, and 70 cm.	1.015	± 0.02	± 0.05
October 1969 to January 1971	Deep Space Network (NASA)	Mariner VI and VII spacecraft	Anderson, <i>et al.</i> (1971)	14 cm.	1.00	± 0.014	± 0.04

^aHere (observed delay)/(Einstein prediction) is the value of $\frac{1}{2}(1 + \gamma)$ obtained by fitting the observational data, $\Delta\tau(\tau)$, to a more sophisticated version of the PPN prediction (40.14). This more sophisticated version includes the gravitational influences of all the planets on the orbits of reflector and Earth; also the effect of the moon on the Earth's orbit and the effect of the Earth's rotation on the travel time; also, to as good an extent as possible, the delay due to dispersion in the solar corona and wind. "Formal standard error" and "one-sigma error" are defined in the table in Box 40.1.

§40.5. PERIHELION SHIFT AND PERIODIC PERTURBATIONS IN GEODESIC ORBITS

Perihelion shift for geodesic orbits around spherical sun, ignoring preferred-frame effects

The light-deflection and time-delay experiments both measured γ . To measure other PPN parameters, one must examine the effects of gravity on slowly moving bodies; this was the message of §40.2.

Begin with the simplest of cases: the geodesic orbit of a test body in the sun's spherical gravitational field, ignoring all gravitational effects of the planets, of solar oblateness, and of motion relative to any preferred frame. The PPN metric then has the form (40.3):

$$\begin{aligned} ds^2 = & - \left[1 - 2 \frac{M_{\odot}}{r} + 2\beta \frac{M_{\odot}^2}{r^2} \right] dt^2 \\ & + \left[1 + 2\gamma \frac{M_{\odot}}{r} \right] [dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)]. \end{aligned} \quad (40.3)$$

Orient the coordinates so the test body moves in the equatorial "plane" $\theta = \pi/2$; and calculate the shape $r(\phi)$ of its nearly Keplerian, nearly elliptical geodesic orbit. The result, accurate to order M_{\odot}/r beyond Newtonian theory, is

$$r = \frac{(1 - e^2)a}{1 + e \cos[(1 - \delta\phi_0/2\pi)\phi]}, \quad (40.17)$$

where a and e are constants of integration, and $\delta\phi_0$ is defined by

$$\begin{aligned} \delta\phi_0 &= \frac{(2 - \beta + 2\gamma)}{3} \frac{6\pi M_{\odot}}{a(1 - e^2)} \\ &= \frac{6\pi M_{\odot}}{a(1 - e^2)} \text{ in general relativity.} \end{aligned} \quad (40.18)$$

(For derivation, see exercise 40.4.)

Notice that, if $\delta\phi_0$ were zero—as it is in the Newtonian limit—then the orbit (40.18) would be an ellipse with semimajor axis a and eccentricity e (see Box 25.4). The constant $\delta\phi_0$ merely makes the ellipse precess: for r to go through a complete circuit, from perihelion to aphelion to perihelion again, $(1 - \delta\phi_0/2\pi)\phi$ must change by 2π ; so ϕ must change by $2\pi + \delta\phi_0$. Thus, *the perihelion shifts forward by an angle $\delta\phi_0$ with each circuit around the ellipse*.

Relative to what does the perihelion shift? (1) Relative to the PPN coordinate system; (2) relative to inertial frames at the outskirts of the solar system (since the PPN coordinates are tied to those frames; see §39.12); (3) relative to a frame determined by the "fixed stars in the sky" (since the inertial frames at the outskirts of the solar system, inertial frames elsewhere in our galaxy, and inertial frames in our cluster of galaxies should not rotate significantly relative to each other); (4) relative to the perihelia of (other) planets, which themselves are shifting at calculable rates that decrease as one moves outward in the solar system from Mercury to Venus to Earth to

The perihelion shift is not the only relativistic effect contained in the orbital motion for a test body. There are other effects, but they are all periodic rather than cumulative with time; so, with the limited technology of the pre-space-age era, it was impossible to detect them. But the technology of the 1970's is bringing them within reach. Moreover, many space-age experiments are necessarily of short duration (\lesssim one orbit)—particularly those involving spacecraft and transponders landed on planets. For these, the periodic perturbations in an orbit are of almost as much experimental value as the cumulative perihelion shift. The periodic effects are not obvious in the PPN orbital equation (40.17); it looks like the simplest of precessing ellipses. But the quantities the observer measures directly are not a , e , and $\delta\phi_0$. Instead, he measures the time evolution of round-trip radar travel times, $\Delta\tau(\tau)$, and of angular positions on the sky $[\theta_0(\tau), \phi_0(\tau)]$. To compute these quantities is perfectly straightforward in principle, but in practice is a very complex task. The calculations predict relativistic effects periodic with the frequency of the orbit and all its harmonics. The amplitudes of these effects, for the lower harmonics, must obviously be of the order of $M_\odot \sim 1 \text{ km} \sim 10 \mu\text{sec} \sim 0''.01$ arc on the sky. (The distance $M_\odot = 1.48 \text{ km}$ is the characteristic length for all relativistic effects in the sun's spherical field!)

The most favorable orbits for experimental tests of the perihelion shift and of periodic effects are those that go nearest the sun and have the highest eccentricity [see equation (40.18)]—the orbits of Mercury, Venus, Earth, Mars, and the asteroid Icarus. But how does one know that these orbits are geodesics? After all, planets are not “test bodies”; they themselves produce nonnegligible curvature in spacetime. It turns out (see §40.9 for full discussion) that there should exist tiny deviations from geodesic motion, but they are too small to compete with the perihelion shift or with the periodic effects discussed above, at least for these five bodies.

Extensive astronomical observations of planetary orbits, dating back to the mid-nineteenth century and aided by radar since 1966, have yielded accurate values for planetary perihelion shifts (accurate to ± 0.4 seconds of arc per century for Mercury). From the data, which are summarized and discussed in Box 40.3, one obtains the value

$$\frac{1}{3}(2 - \beta + 2\gamma) = 1.00 \begin{cases} +0.01 \\ -0.10 \end{cases} \quad (40.19a)$$

for the ratio of observed relativistic shift to general relativistic prediction. Combining this result with the radar-delay value for γ (40.16), one obtains a value

$$\beta = 1.0 \begin{cases} +0.4 \\ -0.2 \end{cases} \quad (40.19b) \quad \text{Experimental result for } \beta$$

for the PPN parameter β . (Recall: β measures the “amount of nonlinearity in the superposition law for g_{00} .”)

The periodic effects in the planetary orbits have not yet (1973) been studied experimentally.

The above discussion and Box 40.3 have ignored the motion of the solar system relative to the preferred frame (if one exists); i.e., they have ignored the terms (40.3')

Periodic perturbations in geodesic orbits

Comparison of theory with planetary orbits

Box 40.3 PERIHELION SHIFTS; EXPERIMENTAL RESULTS

Relativistic corrections to Newtonian theory are not the only cause of shift in the perihelion of a planetary orbit. Any departure of the Newtonian gravitational field from its idealized, spherical, inverse-square-law form also produces a shift. Such nonsphericities and resulting shifts are brought about by (1) the gravitational pulls of other planets, and (2) deformation of the sun ("solar oblateness"; "quadrupole moment"). In addition, when the primary data are optical positions of planets on the sky (right ascension and declination as functions of time), there is an apparent perihelion shift caused by the precession of the Earth's axis ("general precession"; observer not on a "stable platform"; see exercise 16.4).

The perihelion shifts due to a general precession and to the gravitational pulls of other planets can be calculated with high precision. But in 1973 there is no fully reliable way to determine the solar quadrupole moment. It is conventional to quantify the sun's quadrupole moment by a dimensionless parameter J_2 , which appears in the following expression for the Newtonian potential,

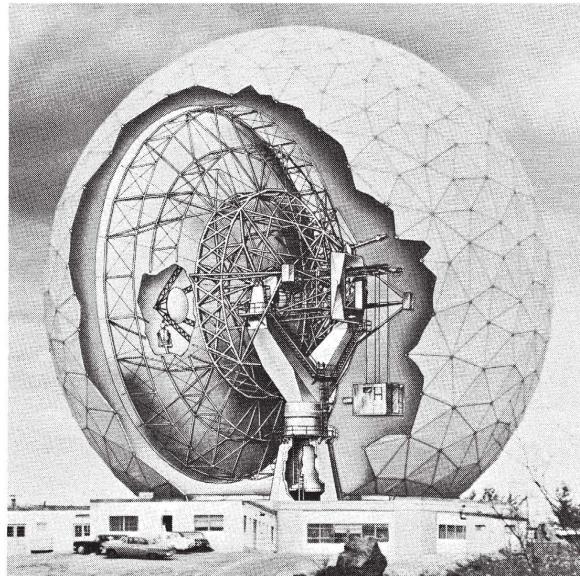
$$U = \frac{M_\odot}{r} \left[1 - J_2 \frac{R_\odot^2}{r^2} \left(\frac{3 \cos^2 \theta - 1}{2} \right) \right].$$

If the sun were rotating near breakup velocity, J_2 would be near 1. Very careful measurements of the optical shape of the sun [Dicke and Goldenberg (1967)] show a flattening, which suggests J_2 may be near 3×10^{-5} .

The total perihelion shift produced by relativity plus solar quadrupole moment is (see exercise 40.5)

$$\delta\phi = \frac{6\pi M_\odot}{a(1-e^2)} \lambda_p,$$

$$\lambda_p \equiv \frac{2-\beta+2\gamma}{3} + J_2 \frac{R_\odot^2/M_\odot}{2a(1-e^2)}.$$



The Haystack radar antenna, which Irwin Shapiro and his group have used to collect extensive data on the systematics of the inner part of the solar system. Those data are rapidly becoming the most important source of information about perihelion shifts. (Picture courtesy of Lincoln Laboratories, MIT.)

Note that relativistic and quadrupole shifts have different dependences on the semimajor axis a and eccentricity e of the orbit. This difference in dependence allows one to obtain values for both the quadrupole moment parameter J_2 , and the PPN parameter $\frac{1}{3}(2-\beta+2\gamma)$ by combining measurements of $\delta\phi$ for more than one planet.

The experimental results, as of 1972, are as follows.

I. Data for Mercury from optical studies [Clemence (1943, 1947)]*
 (general relativity with no solar oblateness predicts 43.03"/century)

Quantity	Value
(a) Total observed shift per century	5599".74 ± 0".41
(b) Contribution to shift caused by observer not being in an inertial frame far from the sun ("general precession" as evaluated in 1947)	5025".645 ± 0".50
(c) Shift per century produced by Newtonian gravitation of other planets	531".54 ± 0".68
(d) Residual shift per century to be attributed to general relativity plus solar oblateness	42".56 ± 0".94
(e) Residual shift if one uses the 1973 value for the "general precession"	41".4 ± 0".90
(f) Corresponding value of λ_p (see above)	$\lambda_p = 0.96 \pm 0.02$

II. 1970 Results of Shapiro (1970, 1971a,b), Shapiro *et al.* (1972)

(a) Values of λ_p obtained by reanalyzing all the world's collection of optical data, and combining it with radar data	$\{ (\lambda_p)_{\text{Mercury}} = 1.00 \pm 0.01$
(b) Value of J_2 obtained by comparing the observed shifts for Mercury and Mars	$\{ (\lambda_p)_{\text{Mars}} = 1.07 \pm 0.10$
	$J_2 \lesssim 3 \times 10^{-5}$

III. Theoretical implications of Shapiro's results

(a) Value of $(2 - \beta + 2\gamma)/3$	$1.00 \{ +0.01$
(b) Value of β obtained by combining with γ from time delay experiments [equation (40.16)]	$1.0 \{ +0.4$

* Clemence (1947) notes, "The observations cannot be made in a Newtonian frame of reference. They are referred to the moving equinox, that is, they are affected by the precession of the equinoxes, and the determination of the precessional motion is one of the most difficult problems of observational astronomy, if not the most difficult. In the light of all these hazards, it is not surprising that a difference of opinion could exist regarding the closeness of agreement between the observed and theoretical motions."

in the sun's metric. When one takes account of these terms, one finds an additional contribution to the perihelion shift, given for small eccentricities $e \ll 1$ by

$$\delta\phi_0 = -\alpha_1 \frac{\pi}{2e} \left(\frac{M_\odot}{a} \right)^{1/2} \mathbf{w} \cdot \mathbf{Q} - \alpha_2 \frac{\pi}{4} [(\mathbf{w} \cdot \mathbf{P})^2 - (\mathbf{w} \cdot \mathbf{Q})^2] + \alpha_3 \frac{\pi}{e} \left(\frac{|\mathcal{Q}_\odot|}{M_\odot} \right) \left(\frac{\omega_\odot a^2}{M_\odot} \right) \mathbf{w} \cdot \mathbf{Q} \quad (40.20)$$

Perihelion shift due to preferred-frame forces

[see Nordtvedt and Will (1972)]. Here M_\odot , \mathcal{Q}_\odot , and ω_\odot are the sun's mass, self-gravitational energy, and rotational angular velocity; \mathbf{w} is the sun's velocity relative to the preferred frame; a and e are the semimajor axis and eccentricity of the orbit; \mathbf{P} is the unit vector pointing from the sun to the perihelion; and \mathbf{Q} is a unit vector orthogonal to \mathbf{P} and lying in the orbital plane. Comparison with observations for

Mercury—and combination with limits on α_1 and α_2 discussed below [equations (40.46b) and (40.48)]—yields the limit

Experimental result for α_3

$$\left| \alpha_3 \frac{\mathbf{w} \cdot \mathbf{Q}}{200 \text{ km/sec}} \right| \lesssim 2 \times 10^{-5}. \quad (40.21a)$$

Since the velocity of the sun around the Galaxy is ~ 200 km/sec, and the peculiar motion of the Galaxy relative to other nearby galaxies is ~ 200 km/sec, a value $w \sim 200$ km/sec is reasonable. Moreover, there is no reason to believe that w and Q are orthogonal, so one is fairly safe in concluding

$$|\alpha_3| = |4\beta_1 - 2\gamma - 2 - \xi| \lesssim 2 \times 10^{-5} \quad (40.21b)$$

This is a stringent limit on theories that possess universal rest frames. For example, with great certainty it rules out a theory devised by Coleman (1971), which has $\beta = \gamma = 1$, but $\alpha_3 = -4$; see Ni (1972).

The future of orbital experiments

Looking toward the future, one cannot expect data on orbits of spacecraft to give decisive tests of general relativity, despite the high precision (~ 10 meters in 1972) with which spacecraft can be tracked. Spacecraft are buffeted by the solar wind. They respond to fluctuations in this wind and in the pressure of solar radiation, and respond also to “outgassing” from leaky jets. Unless one can develop excellent “drag-free” or “conscience-guided” spacecraft, one must therefore continue to rely on planets as the source of data on geodesics. However, planetary data themselves can be greatly improved in the future by placing radar transponders on the surfaces of planets or in orbit about them, by improvements in radar technology, and by the continued accumulation of more and more observations.

EXERCISES

Exercise 40.4. DERIVATION OF PERIHELION SHIFT IN PPN FORMALISM

[See exercise 25.16 for a derivation in general relativity, accurate when gravity is strong ($2M/r$ as large as $\frac{1}{3}$) but the orbital eccentricity is small. The present exercise applies to any “metric theory” and to any eccentricity, but it assumes gravity is weak ($2M/r \ll 1$) and ignores motion relative to any universal rest frame.] Derive equation (40.17) for the shape of any bound orbit of a test particle moving in the equatorial plane of the PPN gravitational field (40.3). Keep only “first-order” corrections beyond Newtonian theory (first order in powers of M_\odot/r). [Sketch of solution using Hamilton-Jacobi theory (Box 25.4): (1) Hamilton-Jacobi equation, referred to a test body of unit mass, is

$$\begin{aligned} -1 &= g^{\alpha\beta} \tilde{S}_{,\alpha} \tilde{S}_{,\beta} \\ &= - \left[1 + 2 \frac{M_\odot}{r} + (4 - 2\beta) \left(\frac{M_\odot}{r} \right)^2 \right] \left(\frac{\partial \tilde{S}}{\partial t} \right)^2 + \left[1 - 2\gamma \frac{M_\odot}{r} \right] \left[\left(\frac{\partial \tilde{S}}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial \tilde{S}}{\partial \phi} \right)^2 \right]. \end{aligned}$$

(2) Solution to Hamilton-Jacobi equation is

$$\begin{aligned} \tilde{S} &= -\tilde{E}t + \tilde{L}\phi \pm \int^r \left\{ -(1 - \tilde{E}^2) + \frac{2M_\odot}{r} [1 - (1 + \gamma)(1 - \tilde{E}^2)] \right. \\ &\quad \left. - \frac{\tilde{L}^2}{r^2} \left[1 - \frac{2M_\odot^2}{\tilde{L}^2} (2 - \beta + 2\gamma) \right] \right\}^{1/2} dr, \end{aligned} \quad (40.22)$$

where post-post-Newtonian corrections have been discarded. In discarding post-post-Newtonian corrections, recall that \tilde{E} is the conserved energy per unit rest mass and \tilde{L} is the angular momentum per unit rest mass (see Box 25.4). Consequently one has the order-of-magnitude relations

$$1 - \tilde{E}^2 \sim (\text{velocity of test body})^2 \sim M_\odot/r$$

and

$$(M_\odot/\tilde{L})^2 \sim (M_\odot/rv)^2 \sim M_\odot/r.$$

(3) The shape of the orbit is determined by the “condition of constructive interference,” $\partial\tilde{S}/\partial\tilde{L} = 0$:

$$\begin{aligned}\phi = \pm \int & \left\{ -\frac{1 - \tilde{E}^2}{\tilde{L}^2} + \frac{2M_\odot}{\tilde{L}^2 r} [1 - (1 + \gamma)(1 - \tilde{E}^2)] \right. \\ & \left. - \frac{1}{r^2} \left[1 - \frac{2M_\odot^2}{\tilde{L}^2} (2 - \beta + 2\gamma) \right] \right\}^{-1/2} d(1/r).\end{aligned}$$

(4) This integral is readily evaluated in terms of trigonometric functions. For a bound orbit ($\tilde{E} < 1$), the integral is

$$\phi = \left(1 + \frac{\delta\phi_0}{2\pi}\right) \cos^{-1} \left[\frac{(1 - e^2)a}{er} - \frac{1}{e} \right]$$

where

$$\begin{aligned}a &\equiv \frac{M_\odot}{1 - \tilde{E}^2} [1 - (1 + \gamma)(1 - \tilde{E}^2)], \\ 1 - e^2 &\equiv \left(\frac{\tilde{L}}{M_\odot}\right)^2 (1 - \tilde{E}^2) \left[1 + 2(1 + \gamma)(1 - \tilde{E}^2) - 2\left(\frac{M_\odot}{\tilde{L}}\right)^2 (2 - \beta + 2\gamma) \right], \\ \delta\phi_0 &\equiv \frac{1}{3} (2 - \beta + 2\gamma) 6\pi (M_\odot/\tilde{L})^2.\end{aligned}\quad (40.23)$$

(5) Straightforward manipulations bring this result into the form of equations (40.17) and (40.18).]

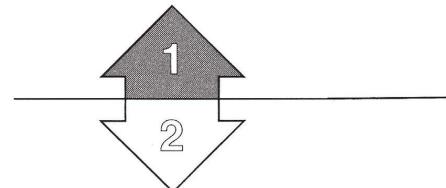
Exercise 40.5. PERIHELION SHIFT FOR OBLATE SUN

(a) The Newtonian potential for an oblate sun has the form

$$U = \frac{M_\odot}{r} \left(1 - J_2 \frac{R_\odot^2}{r^2} \frac{3 \cos^2\theta - 1}{2} \right), \quad (40.24)$$

where J_2 is the “quadrupole-moment parameter.” One knows that $J_2 \lesssim 3 \times 10^{-5}$. Show that if an oblate sun is at rest at the origin of the PPN coordinate system, the metric of the surrounding spacetime [equations (39.32)] can be put into the form

$$\begin{aligned}ds^2 = & - \left[1 - 2 \frac{M_\odot}{r} - 2J_2 \left(\frac{M_\odot R_\odot^2}{r^3} \right) \left(\frac{3 \cos^2\theta - 1}{2} \right) + 2\beta \left(\frac{M_\odot}{r} \right)^2 \right] dt^2 \\ & + \left[1 + 2\gamma \frac{M_\odot}{r} \right] [dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)] \\ & + \text{corrections of post-post-Newtonian magnitude.}\end{aligned}\quad (40.25)$$



(b) Let a test particle move in a bound orbit in the equatorial plane. Use Hamilton-Jacobi theory to show that its orbit is a precessing ellipse [equation (40.17)] with a precession per orbit given by

$$\delta\phi_0 = \frac{2 - \beta + 2\gamma}{3} \frac{6\pi M_\odot}{a(1 - e^2)} + J_2 \frac{3\pi R_\odot^2}{a^2(1 - e^2)^2}. \quad (40.26)$$

For the significance of this result, see Box 40.3.

The rest of this chapter is Track 2. No earlier Track-2 material is needed as preparation for it, but the following will be helpful:
 (1) Chapter 6 (accelerated observers);
 (2) §17.6 (no prior geometry); and
 (3) Chapters 38 and 39 (tests of foundations; other theories; PPN formalism).
 It is not needed as preparation for any later chapter.

3-body effects in lunar orbit:

(1) theory

(2) prospects for measurement

§40.6. THREE-BODY EFFECTS IN THE LUNAR ORBIT

The relativistic effects discussed thus far all involve the spherical part of the sun's external gravitational field, and thus they can probe only the PPN parameters β and γ plus the "preferred-frame" parameters α_1 , α_2 , and α_3 . Attempts to measure other PPN parameters can focus on three-body interactions (discussed here), the dragging of inertial frames by a rotating body (§40.7), anomalies in the locally measured gravitational constant (§40.8), and deviations of planetary and lunar orbits from geodesics (§40.9).

There is no better place to study three-body interactions than the Earth-moon orbit. The pulls of the Earth, the moon, and the sun all contribute. Perturbations in the motion of Earth and moon about their common center of gravity can be measured with high precision using laser ranging (earth-moon separation measured to ~ 10 cm in early 1970's) and using a radio beacon on the moon's surface (angular position on sky potentially measurable to better than $0''.001$ of arc).

Over and above any Newtonian three-body interactions, the Earth and the sun, acting together in a nonlinear manner, should produce relativistic perturbations in the lunar orbit that are barely within the range of this technology. These effects depend on the familiar parameters γ (measuring space curvature) and β [measuring amount of nonlinear superposition, $(U_{\text{Earth}} + U_{\text{sun}})^2$, in g_{00}]. In addition, they depend on β_2 , which regulates the extent to which the sun's potential, U_{sun} , acting inside the Earth, affects the strength of the Earth's gravitational pull, causing it to vary as the Earth moves nearer and farther from the sun. These effects are expected to depend also on ζ , Δ_1 , and Δ_2 , which regulate the extent to which the Earth's orbital momentum and anisotropies in kinetic energy (caused by the sun) gravitate.

Bromberg (1958), Baierlein (1967), and Krogh and Baierlein (1968) have calculated the three dominant three-body effects in the Earth-moon orbit using general relativity and the Dicke-Brans-Jordan theory. These effects are noncumulative and have amplitudes of ~ 100 cm, ~ 20 cm, and ~ 10 cm. The 100-cm effect [which was originally discovered by de Sitter (1916)] is known to depend only on γ . The precise dependence of the other effects on the PPN parameters is not known.

The prospects for measuring these effects in the 1970's are dim; they are masked by peculiarities in the orbit of the moon that have nothing to do with relativity.

§40.7. THE DRAGGING OF INERTIAL FRAMES

The experiments discussed thus far study the motion of electromagnetic waves, spacecraft, planets, and asteroids through the solar system. An entirely different type of experiment measures changes in the orientation of a gyroscope moving in the gravitational field of the Earth. This experiment is particularly important because it can measure directly the “dragging of inertial frames” by the angular momentum of the Earth.

It is useful, before specializing to a rotating Earth, to derive a general expression for the precession of a gyroscope in the post-Newtonian limit. (Track-1 readers, and others who have not studied Chapters 6 and 39, may have difficulty following the derivation. No matter. It is the answer that counts!)

Let S^α be the spin of the gyroscope (i.e., its angular momentum vector), and let u^α be its 4-velocity. The spin is always orthogonal to the 4-velocity, $S^\alpha u_\alpha = 0$ (see Box 5.6). Assume that any nongravitational forces acting on the gyroscope are applied at its center of mass, so that there is no torque in its proper reference frame. Then the gyroscope will “Fermi-Walker transport” its spin along its world line (see §6.5):

$$\nabla_u \mathbf{S} = \mathbf{u}(\mathbf{a} \cdot \mathbf{S}), \quad \mathbf{a} \equiv \nabla_u \mathbf{u} = \text{4-acceleration.} \quad (40.27)$$

The objective of the calculation is to write down and analyze this transport equation in the post-Newtonian limit.

The gyroscope moves relative to the PPN coordinate grid with a velocity $v_j \equiv dx^j/dt \equiv dx_j/dt$. Assume that $v_j \lesssim \epsilon$, where ϵ is the post-Newtonian expansion parameter ($\epsilon^2 \approx M_\odot/R_\odot$). As the gyroscope moves, it carries with itself an orthonormal frame $\mathbf{e}_{\hat{\alpha}}$, which is related to the PPN coordinate frame by a pure Lorentz boost, plus a renormalization of the lengths of the basis vectors [transformation (39.41)]. The spin is a purely spatial vector ($S^0 = 0$) in this comoving frame; its length $(S_{\hat{j}} S_{\hat{j}})^{1/2}$ remains fixed (conservation of angular momentum); and its direction is regulated by the Fermi-Walker transport law.

The basis vectors $\mathbf{e}_{\hat{\alpha}}$ of the comoving frame are *not* Fermi-Walker transported, by contrast with the spin. Rather, they are tied by a pure boost (no rotation!) to the PPN coordinate grid, which in turn is tied to an inertial frame far from the solar system, which in turn one expects to be fixed relative to the “distant stars.” Thus, by calculating the precession of the spin relative to the comoving frame,

$$dS_{\hat{j}}/d\tau \equiv \epsilon_{\hat{j}\hat{k}\hat{l}} \Omega_{\hat{k}} S_{\hat{l}}, \quad (40.28)$$

one is in effect evaluating the spin’s angular velocity of precession, $\Omega_{\hat{j}}$, relative to a frame fixed on the sky by the distant stars.

Calculate $dS_{\hat{j}}/d\tau$:

$$\frac{dS_{\hat{j}}}{d\tau} = \nabla_u (\mathbf{S} \cdot \mathbf{e}_{\hat{j}}) = (\nabla_u \mathbf{S}) \cdot \mathbf{e}_{\hat{j}} + \mathbf{S} \cdot (\nabla_u \mathbf{e}_{\hat{j}}) = \mathbf{S} \cdot \nabla_u \mathbf{e}_{\hat{j}}. \quad (40.29)$$

Here use is made of the fact that $\nabla_u \mathbf{S}$ is in the \mathbf{u} direction [equation (40.27)] and

Gyroscope precession:

(1) general analysis

is thus orthogonal to \mathbf{e}_j . The quantity $\mathbf{S} \cdot \nabla_u \mathbf{e}_j$ is readily evaluated in the PPN coordinate frame. In the evaluation, one uses as metric coefficients [equations (39.32)] the expressions

$$\begin{aligned} g_{00} &= -1 + 2U + O(\epsilon^4), & g_{jk} &= \delta_{jk}(1 + 2\gamma U) + O(\epsilon^4), \\ g_{0j} &= -\frac{7}{2}\Delta_1 V_j - \frac{1}{2}\Delta_2 W_j + \left(\begin{array}{l} \text{"preferred-} \\ \text{frame terms"} \end{array} \right) + O(\epsilon^5); \end{aligned} \quad (40.30)$$

one takes as the components of \mathbf{e}_j and \mathbf{S} [obtained via the transformation (39.41)] the expressions

$$\begin{aligned} e_j^0 &= v_j + O(\epsilon^3), & e_j^k &= (1 - \gamma U)\delta_{jk} + \frac{1}{2}v_k v_j + O(\epsilon^4), \\ S^0 &= v_j S_j + O(\epsilon^3 S_j), \\ S^k &= (1 - \gamma U)S_k + \frac{1}{2}v_k v_j S_j + O(\epsilon^4 S_j); \end{aligned} \quad (40.31)$$

and one uses the relation

$$dv_j/d\tau = a_j + U_{,j} + O(\epsilon^2 U_{,j}) \quad (40.32)$$

where a_j (assumed $\lesssim U_{,j}$) are the components of the 4-acceleration. One finds (see exercise 40.6) for the precession of the spin the result

$$dS_j/d\tau = \mathbf{S} \cdot \nabla_u \mathbf{e}_j = S_k[v_{[j}a_{k]} + g_{0[k,j]} - (2\gamma + 1)v_{[j}U_{,k]}].$$

Rewritten in *three-dimensional vector form* this result becomes

- (2) general PPN formula for precession

$$d\mathbf{S}/d\tau = \boldsymbol{\Omega} \times \mathbf{S}, \quad (40.33a)$$

$$\boldsymbol{\Omega} \equiv -\frac{1}{2}\mathbf{v} \times \mathbf{a} - \frac{1}{2}\nabla \times \mathbf{g} + \left(\gamma + \frac{1}{2}\right)\mathbf{v} \times \nabla U, \quad (40.33b)$$

$$\mathbf{g} \equiv g_{0j}\mathbf{e}_j. \quad (40.33c)$$

In this final answer it does not matter whether the 3-vectors entering into $\boldsymbol{\Omega}$ are evaluated in the coordinate frame or in the comoving orthonormal frame, since \mathbf{e}_j and $\partial/\partial x_j$ differ only by corrections of order ϵ^2 .

Equations (40.33) describe in complete generality at the post-Newtonian level of approximation the precession of the gyroscope spin \mathbf{S} relative to the comoving orthonormal frame that is rotationally tied to the distant stars.

- (3) specialization: Thomas precession

For an electron with spin \mathbf{S} in orbit around a proton, only the first term, $-\frac{1}{2}\mathbf{v} \times \mathbf{a}$, is present (no gravity). This term leads to the Thomas precession, which plays an important role in the fine structure of atomic spectra [see, e.g., Ruark and Urey (1930)]. For other ways of deriving the Thomas precession, see exercise 6.9 and §41.4.

The Thomas precession comes into play for a gyroscope on the surface of the Earth (\mathbf{a} = Newtonian acceleration of gravity), but not for a gyroscope in a freely moving satellite.

If one ignores the rotation of the Earth and preferred-frame effects, and puts the PPN coordinate frame at rest relative to the center of the Earth, then g_{0j} vanishes and Ω is given by

$$\begin{aligned}\Omega &= \mathbf{v} \times \left[-\frac{1}{2}\mathbf{a} + \left(\gamma + \frac{1}{2}\right)\nabla U \right] \\ &= \gamma\mathbf{v} \times \nabla U \text{ for gyroscope on Earth's surface} \\ &= \left(\gamma + \frac{1}{2}\right)\mathbf{v} \times \nabla U \text{ for gyroscope in orbit.}\end{aligned}\quad (40.34)$$

The general-relativistic term $(\gamma + \frac{1}{2})\mathbf{v} \times \nabla U$ is caused by the motion of the gyroscope through the Earth's curved, static spacetime geometry. Notice that it depends solely on the same parameter γ as is tested by electromagnetic-wave experiments. In order of magnitude, for a gyroscope in a near-Earth, polar orbit,

$$\Omega \approx \frac{3}{2} \left(\frac{M_E}{R_E} \right)^{1/2} \left(\frac{M_E}{R_E^2} \right) \approx 8 \text{ seconds of arc per year.} \quad (40.35)$$

The general-relativistic precession $\frac{3}{2}\mathbf{v} \times \nabla U$ was derived by W. de Sitter (1916) for the "Earth-moon gyroscope" orbiting the sun. Eleven years later L. H. Thomas (1927) derived the special relativistic precession $-\frac{1}{2}\mathbf{v} \times \mathbf{a}$ for application to atomic physics.

The Earth's rotation produces off-diagonal terms, g_{0j} , in the PPN metric (exercise 40.7):

$$\mathbf{g} = g_{0j}\mathbf{e}_j = -\left(\frac{7}{4}\mathcal{A}_1 + \frac{1}{4}\mathcal{A}_2\right)\frac{\mathbf{J} \times \mathbf{r}}{r^3}. \quad (40.36)$$

Here \mathbf{J} is the Earth's angular momentum. These off-diagonal terms contribute an amount

$$\Omega = -\frac{1}{2}\nabla \times \mathbf{g} = \left(\frac{7}{8}\mathcal{A}_1 + \frac{1}{8}\mathcal{A}_2\right)\frac{1}{r^3} \left[-\mathbf{J} + \frac{3(\mathbf{J} \cdot \mathbf{r})\mathbf{r}}{r^2} \right] \quad (40.37)$$

to the precession of the gyroscope. Notice that this contribution, unlike the others, is independent of the linear velocity of the gyroscope. One can think of it in the following way.

The gyroscope is rotationally at rest relative to the inertial frames in its neighborhood. It and the local inertial frames rotate relative to the distant galaxies with the angular velocity Ω because the Earth's rotation "drags" the local inertial frames along with it. Notice that near the north and south poles the local inertial frames rotate in the same direction as the Earth does (Ω parallel to \mathbf{J}), but near the equator they rotate in the opposite direction (Ω antiparallel to \mathbf{J} ; compare Ω with the magnetic field of the Earth!). Although this might seem paradoxical at first, an analogy devised by Schiff makes it seem more reasonable.* Consider a rotating, solid sphere immersed in a viscous fluid. As it rotates, the sphere will drag the fluid along with it. At various points in the fluid, set down little rods, and watch how the fluid

- (4) specialization: precessions due to acceleration and Earth's Newtonian potential

- (5) specialization: precession due to Earth's rotation

*This analogy can be made mathematically rigorous; see footnote on p. 255 of Thorne (1971); see also, §21.12 on Mach's principle.

rotates them as it flows past. Near the poles the fluid will clearly rotate the rods in the same direction as the star rotates. But near the equator, because the fluid is dragged more rapidly at small radii than at large, the end of a rod closest to the sphere is dragged by the fluid more rapidly than the far end of the rod. Consequently, the rod rotates in the direction opposite to the rotation of the sphere.

In order of magnitude, the precessional angular velocity caused by the Earth's rotation is

$$\Omega \sim J_E/R_E^3 \sim 0.1 \text{ seconds of arc per year.} \quad (40.37')$$

- (6) prospects for measuring precession

Both this precession, and the larger one [equation (40.35)] due to motion through the Earth's static field, may be detectable in the 1970's. Equipment aimed at detecting them via a satellite experiment is now (1973) under construction at Stanford University; see Everitt, Fairbank, and Hamilton (1970); also O'Connell (1972).*

The gyroscope precession produced by motion of the Earth relative to the preferred frame (if any) is too small to be of much interest.

*The dragging of inertial frames by a rotating body plays important roles elsewhere in gravitation physics, e.g., in the definition of angular momentum for a gravitating body (§19.2), and in black-hole physics (Chapter 33). The effect was first discussed and calculated by Thirring and Lense (1918). More recent calculations by Brill and Cohen (1966) of idealized situations where the effect may be large give insight into the mechanism of the effect. See also the discussion of Mach's principle in §21.12.

EXERCISES

Exercise 40.6. PRECESSIONAL ANGULAR VELOCITY

Derive equations (40.33) for the precession of a gyroscope in the post-Newtonian limit. Base the derivation on equations (40.29)–(40.32).

Exercise 40.7. OFF-DIAGONAL TERMS IN METRIC ABOUT THE EARTH

Idealize the Earth as an isolated, rigidly rotating sphere with angular momentum \mathbf{J} . Use equations (39.34b,c) and (39.27) to show that (in three-dimensional vector notation)

$$\mathbf{V} \equiv V_j \mathbf{e}_j = \mathbf{W} \equiv W_j \mathbf{e}_j = \frac{1}{2} \mathbf{J} \times \mathbf{r}/r^3 \quad (40.38)$$

outside the Earth, in the Earth's PPN rest frame. From this, infer equation (40.36).

Exercise 40.8. SPIN-CURVATURE COUPLING

Consider a spinning body (e.g., the Earth or a gyroscope or an electron) moving through curved spacetime. Tidal gravitational forces produced by the curvature of spacetime act on the elementary pieces of the spinning body. These forces should depend not only on the positions of the pieces relative to the center of the object, but also on their relative velocities. Moreover, the spin of the body,

$$\mathbf{S} \equiv \int (\rho \mathbf{r} \times \mathbf{v}) d(\text{volume}) \quad \text{in comoving orthonormal frame,}$$

is a measure of the relative positions and velocities of its pieces. Therefore one expects the spin to couple to the tidal gravitational forces—i.e., to the curvature of spacetime—producing

deviations from geodesic motion. Careful solution of the PPN equations of Chapter 39 for general relativity reveals [Papapetrou (1951), Pirani (1956)] that such coupling occurs and causes a deviation of the worldline from the course that it would otherwise take; thus,

$$m \frac{Du^\alpha}{d\tau} = -S_\mu u_\nu \frac{D^2 u_\beta}{d\tau^2} \epsilon^{\alpha\mu\nu\beta} + \frac{1}{2} (\epsilon^{\lambda\mu\rho\tau} R^{\alpha\nu}_{\lambda\mu}) u_\nu S_\rho u_\tau. \quad (40.39)$$

Evaluate, in order of magnitude, the effects of the supplementary term on planetary orbits in the solar system.

[Answer: They are much too small to be detected. However, there are two other material places to look for the effect: (1) when a rapidly spinning neutron star, or a black hole endowed with substantial angular momentum enters the powerful tidal field of another neutron star or black hole; and (2) when an individual electron, or the totality of electrons in the “Dirac sea of negative energy states,” enter a still more powerful tidal field (late phase of gravitational collapse). Such a tidal field, or curvature, pulls oppositely on electrons with the two opposite directions of spin [Pirani (1956); DeWitt (1962), p. 338; Schwinger (1963a,b)] just as an electric field pulls oppositely on electrons with the two opposite signs of charge [“vacuum polarization”; see especially Heisenberg and Euler (1936)]. In principle, the tidal field pulling on the spin of an electron need not be due to “background” spacetime curvature; it might be due to a nearby massive spinning object, such as a “live” black hole (chapter 33) [“gravitational spin-spin coupling”; O’Connell (1972)].

§40.8. IS THE GRAVITATIONAL CONSTANT CONSTANT?

The title and subject of this section are likely to arouse confusion. Throughout this book one has used geometrized units, in which $G = c = 1$. Therefore, one has locked oneself into a viewpoint that forbids asking whether the gravitational constant changes from event to event.

False! One can perfectly well ask the question in the context of $G = c = 1$, if one makes clear what is meant by the question.

In §§1.5 and 1.6, c was *defined* to be a certain conversion factor between centimeters and seconds; and G/c^2 was defined to be a certain conversion factor between grams and centimeters. These definitions by fiat do not guarantee, however, that a Cavendish experiment* to measure the attraction between two bodies will yield

$$\text{Force} = -Gm_1 m_2 / r^2 = -m_1 m_2 / r^2.$$

If general relativity correctly describes classical gravity, and if the values of the conversion factors G and c have been chosen precisely right, then any Cavendish experiment, anywhere in the universe, *will* yield “Force = $-m_1 m_2 / r^2$ ”. But if the

*See any standard textbook for a description of Cavendish experiments. By his original version of the experiment, with two separated spheres suspended by fine wires, Henry Cavendish (1798) inferred the mass and hence the density of the Earth. He reported: “By a mean of the experiments made with the wire first used, the density of the Earth comes out 5.48 times greater than that of water; and by a mean of those made with the latter wire it comes out the same; and . . . the extreme results do not differ from the mean more than 0.38, or 1/14 of the whole.” The most precise method of measuring G today [Rose *et al.* (1969)] gives $G_c = (6.674 \pm .004) \times 10^{-8} \text{ cm}^3/\text{g sec}^2$ (one standard deviation).

"Cavendish gravitational constant," G_C , defined

Changes of G_C with time

Spatial variations in G_C

Dicke-Brans-Jordan theory, or almost any other metric theory gives the correct description of gravity, the force in the Cavendish experiment will depend on where and when the experiment is performed, as well as on m_1 , m_2 , and r . To discuss Cavendish experiments as tests of gravitation theory, then, one must introduce a new proportionality factor

$$G_C \equiv G_{\text{Cavendish}} \equiv (\text{"Cavendish gravitational constant"}), \quad (40.40)$$

which enters into the Newtonian force law

$$\text{Force} = -G_C m_1 m_2 / r^2. \quad (40.41)$$

This Cavendish constant will be unity in general relativity, but in most other metric theories it will vary from event to event in spacetime.

In some theories, such as Dicke-Brans-Jordan, the Cavendish constant is determined by the distribution of matter in the universe. As a result, the expansion of the universe changes its value:

$$\frac{1}{G_C} \frac{dG_C}{dt} \sim - \left(\frac{0.1 \text{ to } 1}{\text{age of universe}} \right) \sim \frac{-1}{10^{10} \text{ or } 10^{11} \text{ years}}$$

[see, e.g. Brans and Dicke (1961)]. A variety of observations place limits on such time variations. Big time changes in G_C during the last 4.6 billion years would have produced marked effects on the Earth, the sun, and the entire solar system. The expected geophysical effects have been summarized and compared with observations by Dicke and Peebles (1965). It is hard to draw firm conclusions because of the complexity of the geophysics involved, but a fairly certain limit is

$$(1/G_C)(dG_C/dt) \lesssim 1/10^{10} \text{ years} \quad (\text{geophysical}). \quad (40.42a)$$

Eventually, high-precision measurements of the orbital motions of planets will yield a better limit. For the present, planetary observations show

$$(1/G_C)(dG_C/dt) \lesssim 4/10^{10} \text{ years} \quad (\text{planetary orbits}) \quad (40.42b)$$

[Shapiro, Smith, *et al.* (1971)]. These limits are tight enough to begin to be interesting, but not yet tight enough to disprove any otherwise viable theories of gravity.

If G_C is determined by the distribution of matter in the universe, then it should depend on where in the universe one is, as well as when. In particular, as one moves from point to point in the solar system, closer to the Sun and then farther away, one should see G_C change. Indeed this is the case in most metric theories of gravity, though not in general relativity. Analyses of Cavendish experiments using the PPN formalism reveal spatial variation in G_C given by

$$\Delta G_C = -2G_C(\beta + \gamma - \beta_2 - 1)U \quad (40.43)$$

[Nordtvedt (1970, 1971a); Will (1971b)].

The amplitude of these variations along the Earth's elliptical orbit is $\Delta G_C/G_C \sim 10^{-10}$, if $\beta + \gamma - \beta_2 - 1 \sim 1$. This is far too small to measure directly in the

1970's. Despite great ingenuity and effort, the most accurate experiments measuring the value of G_C have precisions in 1972 no better than 1 part in 10^4 [see Beams (1971)]. Experiments to search for yearly variations in G_C on Earth without measuring the actual value ("null-type experiments") can surely be performed with better precision than 1 in 10^4 —but not with precisions approaching 1 in 10^{10} . On the other hand, *indirect* consequences of a spatial variation of G_C in the solar system are almost certainly measurable (see §40.9 below).

In Ni's theory of gravity (Box 39.1), and other two-tensor or vector-tensor theories like it, where the prior geometry picks out a preferred "universal rest frame," the Cavendish constant G_C can depend on velocity relative to the preferred frame. For Cavendish experiments with two equal masses separated by distances large compared to their sizes, G_C varies as

$$\Delta G_C = G_C \left[\frac{1}{2} (\alpha_2 + \alpha_3 - \alpha_1) v^2 - \frac{1}{2} \alpha_2 (\mathbf{v} \cdot \mathbf{n})^2 \right] \quad (40.44)$$

[Will (1971b)]. Here \mathbf{v} is the velocity of the Cavendish apparatus relative to the preferred frame, and \mathbf{n} is the unit vector between the two masses. For experiments where one body is a massive sphere (e.g., the Earth), and the other is a small object on the sphere's surface, G_C varies as

$$\begin{aligned} \Delta G_C / G_C &= \frac{1}{2} [(\alpha_3 - \alpha_1) + \alpha_2(1 - I/MR^2)] v^2 \\ &\quad - \frac{1}{2} \alpha_2(1 - 3I/MR^2)(\mathbf{v} \cdot \mathbf{n})^2 \end{aligned} \quad (40.44')$$

[Nordtvedt and Will (1972)]. Here M and R are the mass and radius of the sphere, and

$$I = \int (\rho r^2) 4\pi r^2 dr$$

is the trace of the second moment of its mass distribution. Consequences of these effects for planetary orbits have not yet been spelled out, but consequences for Earthbound experiments have.

Think of a Cavendish experiment in which one mass is the Earth, and the other is a gravimeter on the Earth's surface. The gravimeter gives a reading for the "local acceleration of gravity,"

$$g = G_C m_{\text{Earth}} / r_{\text{Earth}}^2. \quad (40.45)$$

As the Earth turns, so the unit vector \mathbf{n} between its center and the gravimeter rotates, G_C and hence g will fluctuate with a period of 12 sidereal hours and an amplitude

$$(\Delta g/g)_{\text{amplitude}} = \frac{1}{4} \alpha_2 v^2 \cos^2 \theta_m.$$

Here θ_m is the minimum, as the Earth rotates, of the angle between \mathbf{v} (constant vector) and \mathbf{n} (rotating vector). (Note: we have used the value $I/MR^2 \simeq 0.5$ for the Earth.) These fluctuations will produce tides in the Earth of the same type as are

Dependence of G_C on velocity

Anomalies in Earth tides due to anisotropies in G_C :

produced by the moon and sun. As of 1972, gravimeter measurements near the Earth's equator show no sign of any anomalous 12-sidereal-hour effects down to an amplitude of $\sim 10^{-9}$ [Will (1971b)]. Consequently,

(1) experimental value of α_2

$$|\alpha_2|^{1/2}v \cos \theta_m = |\Delta_2 + \zeta - 1|^{1/2}v \cos \theta_m \lesssim 6 \times 10^{-5} \sim 20 \text{ km/sec.} \quad (40.46a)$$

Using a rough estimate of $v \sim 200$ km/sec for the Earth's velocity relative to the universal rest frame, and $\theta_m \lesssim 60^\circ$ for the angle between v and the Earth's equatorial plane, one obtains the rough limit

$$|\alpha_2| = |\Delta_2 + \zeta - 1| \lesssim 0.03. \quad (40.46b)$$

[This limit does not affect the three theories in Box 39.1; of them, only Ni's theory has prior geometry and a universal rest frame; and it predicts isotropic effects in $\Delta G_C/G_C$ [equation (40.44)], but no anisotropic effects. However, other theories with universal rest frames—e.g. Papapetrou's (1954a,b,c) theory—are ruled out by this limit; see Ni (1972), Nordtvedt and Will (1972).]

(2) experimental disproof of Whitehead theory

Whitehead's theory of gravity (which is a two-tensor theory with a rather different type of prior geometry from Ni's) predicts that the galaxy should produce velocity-independent anisotropies in G_C . These, in turn, would produce Earth tides with periods of 12 sidereal hours and amplitudes of

$$\Delta g/g \sim 2 \times 10^{-7} \sim 100 \times \left(\begin{array}{l} \text{experimental limit on} \\ \text{such amplitudes} \end{array} \right)$$

[Will (1971b)]. The absence of such tides proves Whitehead's theory to be incorrect—a feat of disproof beyond the power of all redshift, light-deflection, time-delay, and perihelion-shift measurements. (For all these “standard experiments,” the predictions of Whitehead and Einstein are identical!)

Equation (40.44') predicts a periodic annual variation of the Cavendish constant on Earth, as the Earth moves around the sun:

$$\begin{aligned} v &= \left(\begin{array}{l} \text{velocity of Earth} \\ \text{relative to sun} \end{array} \right) + \left(\begin{array}{l} \text{velocity of sun relative} \\ \text{to preferred frame} \end{array} \right) \equiv v_E + w; \\ (\Delta G_C/G_C)_{\substack{\text{averaged over} \\ \text{Surface of Earth}}} &= \frac{1}{2} \left(\frac{2}{3} \alpha_2 + \alpha_3 - \alpha_1 \right) \left(w^2 + v_E^2 + 2w \cdot v_E \right). \end{aligned} \quad (40.47)$$

↑
[varies sinusoidally with period of one year]

Anomalies in Earth rotation rate due to dependence of G_C on velocity

This annual variation, assuming all PPN parameters are of order unity, is 1,000 times larger than the one produced by the Earth's motion in and out through the sun's gravitational potential [equation (40.43)]. In response to this changing Cavendish constant, the Earth's self-gravitational pull should change, and the Earth should “breathe” inward (greater pull) and outward (relaxed pull). The resulting annual variations in the Earth's moment of inertia should produce annual changes in its rotation rate ω (changes in “length of day” as measured by atomic clocks):

$$\delta\omega/\omega \sim 0.1 \left(\frac{2}{3} \alpha_2 + \alpha_3 - \alpha_1 \right) w \cdot v_E$$

[Nordtvedt and Will (1972)]. Comparison with the measured annual variations of rotation rate (all of which geophysicists attribute to seasonal changes in the Earth's atmosphere) yields the following limit

$$\left| \frac{2}{3} \alpha_2 + \alpha_3 - \alpha_1 \right| \leq 0.2. \quad (40.48) \quad \begin{matrix} \text{Experimental value of} \\ \frac{2}{3}\alpha_2 + \alpha_3 - \alpha_1 \end{matrix}$$

[See Nordtvedt and Will (1972)]. This limit rules out several preferred-frame theories of gravity, including that of Ni (Boxes 39.1 and 39.2).

The experimental results (40.21), (40.46), and (40.48), when combined, place the following very rough limits on any theory that possesses a Universal rest frame:

$$\begin{aligned} |\alpha_1| &= |7\Delta_1 + \Delta_2 - 4\gamma - 4| \leq 0.2, \\ |\alpha_2| &= |\Delta_2 + \xi - 1| \leq 0.03, \\ |\alpha_3| &= |4\beta_1 - 2\gamma - 2 - \xi| \leq 2 \times 10^{-5}. \end{aligned} \quad (40.49)$$

These limits completely disprove all theories with preferred frames that have been examined to date except one devised by Will and Nordtvedt [see Ni (1972); Nordtvedt and Will (1972)].

In some theories of gravity, the result of a Cavendish experiment depends on the chemical composition and internal structure of the test bodies (exercises 40.9 and 40.10). Kruezer (1968) has performed the most accurate search for such effects to date. He finds that G_C is the same for fluorine and bromine to a precision of

$$\left| \frac{G_C(\text{bromine}) - G_C(\text{fluorine})}{G_C} \right| \leq 5 \times 10^{-5}. \quad (40.50)$$

Dependence of G_C on chemical composition

Exercise 40.9. CAVENDISH CONSTANT FOR IDEALIZED SUN

EXERCISES

Idealize the sun as a static sphere of perfect fluid at rest at the origin of the PPN coordinates. Then its external gravitational field has the form (40.3), with M_\odot given by (40.4). Consequently, a test body of mass m , located far away at radius r , is accelerated by a gravitational force

$$\text{Force} = -mM_\odot/r^2. \quad (40.51a)$$

(a) Calculate the mass of the sun, M , in the sense of the amount of energy required to construct it by adding one spherical shell of matter on top of another, working from the inside outward. [Answer:

$$\begin{aligned} M &= \underbrace{\int_0^{R_\odot} \rho_0(1 + \Pi + 3\gamma U)4\pi r^2 dr}_{\text{rest mass + internal energy}} - \underbrace{\frac{1}{2} \int_0^{R_\odot} \rho_0 U 4\pi r^2 dr}_{\text{gravitational potential energy}} \\ &= \int_0^{R_\odot} \rho_0 \left[1 + \Pi + \left(3\gamma - \frac{1}{2} \right) U \right] 4\pi r^2 dr. \end{aligned} \quad (40.51b)$$

(b) Use the virial theorem [equation (39.21b)] to rewrite equation (40.4) in the form

$$M_\odot = \int_0^{R_\odot} \rho_0 \left[1 + \beta_3 \Pi + \left(2\beta_2 + \frac{1}{2} \beta_4 \right) U \right] 4\pi r^2 dr. \quad (40.51c)$$

(c) Combine the above equations with the definition

$$\text{Force} = -G_C m M / r^2 \quad (40.51d)$$

of the Cavendish constant for r far outside the sun, to obtain

$$G_C = \frac{\left(\begin{array}{l} \text{mass of sun as defined by its effect in} \\ \text{bending world line of a faraway test particle} \end{array} \right)}{\left(\begin{array}{l} \text{mass-energy as defined by applying law of} \\ \text{conservation of energy to the steps in the} \\ \text{construction of the sun} \end{array} \right)} \quad (40.52)$$

$$= 1 + \int (\rho_0/M_0)[(\beta_3 - 1)\Pi + \frac{1}{2}(4\beta_2 + \beta_4 - 6\gamma + 1)U]4\pi r^2 dr.$$

Unless $\beta_3 = 1$, and $4\beta_2 + \beta_4 - 6\gamma + 1 = 0$ (as they are, of course, in Einstein's theory), G_C will depend on the sun's internal structure! Specialize equation (40.52) to "conservative theories of gravity (Box 39.5), and explain why the result is what one would expect from equation (40.43).

Exercise 40.10. CAVENDISH CONSTANT FOR ANY BODY

Extend the analysis of exercise 40.9 to a source that is arbitrarily stressed and has arbitrary shape and internal velocities (subject to the constraints $v^2 \ll 1$, $|I_{jk}|/\rho_0 \ll 1$, $U \ll 1$, $\Pi \ll 1$, of the post-Newtonian approximation). Assume that the body is at rest relative to the universal rest frame. Show that G_C depends on the internal structure of the source unless

$$2\beta_1 - \beta_4 = 1, \quad 4\beta_2 + \beta_4 - 6\gamma = -1, \quad \beta_3 = 1, \quad \xi = 0, \quad \eta = 0. \quad (40.53)$$

Of course, these PPN constraints are all satisfied by Einstein's theory.

§40.9. DO PLANETS AND THE SUN MOVE ON GEODESICS?

Crucial to solar-system experiments is the question of whether the sun and the planets move on geodesics of spacetime. This question is complicated by the contributions to the spacetime curvature made by the moving body itself.

The sense in which general relativity predicts geodesic motion for planets and sun

To elucidate the question—and to obtain an answer in the framework of general relativity—consider an "Einstein elevator" type of argument. The astronomical object under consideration has an outer boundary, and each point on this boundary describes a world line. These world lines define a world tube. Some distance outside of this world tube construct a "buffer zone" as in §20.6. Tailor its inner and outer dimensions, according to the mass and moments of the object and the curvature of the enveloping space ("strength of the tide-producing force of the external gravitational field"), in such a way that the departure ϵ (cf. §20.6) of the metric from flatness in this buffer zone takes on values equal at most to twice the extremal achievable value ϵ_{extrem} (a minimum with respect to variations in r , a maximum

with respect to variations in direction; in other words, a minimax). Then, apart from errors of order ϵ_{extrem} , the object can be regarded as moving in an asymptotically flat space. The law of conservation of total 4-momentum applies. It assures one that the object moves in a (locally) straight line with uniform velocity. Consider, next, a “background geometry” that agrees just outside the buffer zone with the actual geometry to accuracy ϵ_{extrem} or better, but that inside is a source-free solution of Einstein’s field equation. Then, to an accuracy governed by the magnitude of ϵ_{extrem} , the locally straight line along which the astronomical object moves will be a geodesic of this background geometry.

Insofar as one can give any well-defined meaning to the departure of the actual motion from this geodesic (a task complicated by the fact that the background geometry does not actually exist!), one can calculate this departure by making use of the PPN formalism or some other approximation scheme [see, e.g., Taub (1965)]. This deviation springs ordinarily in substantial measure, and sometimes almost wholly, from a coupling between the Riemann curvature tensor of the external field and the multipole moments of the astronomical object (angular momentum associated with rotation; quadrupole and higher moments associated with deformation; see, e.g., exercises 40.8 and 16.4). This coupling is important for the Earth-moon system, but one need not use relativity to calculate it; Newtonian theory does the job to far greater accuracy than needed—or would, if one understood the interiors of the Earth and the moon well enough! For the planets and sun, the effect is negligible. (Exercise: use Newtonian theory to prove so!).

Thus, in general relativity as applied to the solar system, one can approximate the orbit of the sun, the Earth-moon mass center, and each other planet, as a geodesic of that “background spacetime geometry” which would exist if its own curvature effects were absent. This is the approach used to analyze the perihelion shift for planets in §40.5 in the context of general relativity, and to derive in exercise 39.15 the post-Newtonian “many-body equations of motion.”

In most other metric theories of gravity, including the Dicke-Brans-Jordan theory, there are substantial departures from geodesic motion. The “Einstein elevator” argument fails in these theories because spacetime is endowed not only with a metric, but also with a long-range field that couples indirectly (cf. §§38.7 and 39.2) to massive, gravitating bodies.

This phenomenon is best understood in terms of Dicke’s argument about the influence of spatial variations of the fundamental constants on experiments of the Eötvös-Dicke type (see §38.6). In a theory where the Cavendish gravitational constant G_C depends on position (as it does not and cannot in general relativity), a body with significant self-gravitational energy E_{grav} must fall, in a perfectly uniform external Newtonian gravitational field, with an anomalous acceleration:

$$\left(\begin{array}{l} \text{acceleration of} \\ \text{massive body} \end{array} \right) - \left(\begin{array}{l} \text{acceleration of} \\ \text{test body} \end{array} \right) = \frac{1}{M} \left(\frac{\partial E_{\text{grav}}}{\partial G_C} \right) \nabla G_C \quad (40.54)$$

$$= \frac{E_{\text{grav}}}{MG_C} \nabla G_C$$

Deviations from geodesic motion:

(1) due to curvature coupling

(2) due to spatial dependence of gravitational constant (Nordtvedt effect)

[see equation (38.15)]. In Dicke-Brans-Jordan theory, G_C is essentially the reciprocal of the scalar field; and it contains a small part that is proportional to the Newtonian potential, U [equation (40.43) with the appropriate values of the parameters from Box 39.2]. As a result, the sun falls with an acceleration smaller by one part in 10^6 than the acceleration of a test body; Jupiter falls with an acceleration one part in 10^9 smaller; and the Earth, one part in 10^{10} smaller. Translated into relativistic language: the scalar field, by influencing the gravitational self-energy of a massive body, produces deviations from geodesic motion.

One can use the full PPN formalism of Chapter 29 to calculate the motion of massive bodies in any metric theory of gravity. Nordtvedt (1968b) and Will (1971a) have done this. They find that a massive body at rest in a uniform external field experiences a (Newtonian-type) PPN coordinate acceleration given by

$$\frac{d^2x_j}{dt^2} = E_{jk} \frac{\partial U}{\partial x_k},$$

where E_{jk} is a quantity depending on the body's structure:

$$E_{jk} = \delta_{jk} \left\{ 1 - (7\Delta_1 - 3\gamma - 4\beta) \frac{E_{\text{grav}}}{m} \right\} - (2\beta + 2\beta_2 - 3\gamma + \Delta_2 - 2) \frac{\Omega_{jk}}{m}, \quad (40.55)$$

$$\Omega_{jk} = -\frac{1}{2} \int \frac{\rho_o \rho'_o (x_j - x'_j)(x_k - x'_k)}{|x - x'|^3} d^3x d^3x', \quad E_{\text{grav}} = \sum \Omega_{jj}.$$

Here m is the body's total mass-energy, Ω_{jk} is the “Chandrasekhar potential-energy tensor,” and E_{grav} is the body's self-gravitational energy. [Note: Dicke's method of calculating the anomalous acceleration (40.54) breaks down in theories that are not “conservative” (Box 39.5).]

In general relativity, the combinations of PPN coefficients appearing in E_{jk} vanish; so $E_{jk} = \delta_{jk}$, and the body falls with the usual acceleration—i.e., it moves along a geodesic. But in most other theories of gravity $E_{jk} \neq \delta_{jk}$; the body does *not* move on a geodesic; and its acceleration may even be in a different direction than the gradient of the Newtonian potential!

This predicted departure from geodesic motion is called the “Nordtvedt effect.” The possibility of such an effect was first noticed in passing by Dicke (1964c), but was discovered independently and explored in great detail by Nordtvedt (1968a,b). The Nordtvedt effect in a theory other than general relativity produces a number of phenomena in the solar system that are potentially observable. [See Nordtvedt (1971b) for an enumeration and references.] The effect most suitable for a test is a “polarization” of the Earth-moon orbit due to the fact that the moon should fall toward the sun with a greater acceleration than does the Earth. This “polarization” results in an eccentricity in the orbit that points always toward the Sun and has the amplitude

$$\begin{aligned}\delta r &= 840 \left[3\gamma + 4\beta - 7\Delta_1 - \frac{1}{3}(2\beta + 2\beta_2 - 3\gamma + \Delta_2 - 2) \right] \text{cm} \quad (40.56) \\ &= 67 \text{ meters} \quad \text{in Ni's theory (Boxes 39.1 and 39.2)} \\ &= \frac{8.4}{2+\omega} \text{ meters} \quad \text{in Dicke-Brans-Jordan theory (Boxes 39.1 and 39.2)} \\ &= 0 \quad \text{in Einstein's theory.}\end{aligned}$$

Box 40.4 CATALOG OF EXPERIMENTS

Type of experiment	Description of experiment	Where discussed
I. Tests of foundations of general relativity	1. Tests of uniqueness of free fall (Eötvös-Dicke-Braginsky experiments) 2. Tests for existence of metric (time dilation of particle decays; role of Lorentz group in particle kinematics; etc.) 3. Searches for new, direct-coupling, long-range fields (Hughes Drever experiment; ether-drift experiments) 4. Gravitational redshift experiments 5. Constancy, in space and time, of the nongravitational physical constants	§38.3; Figure 1.6; Box 1.1 §38.4 §38.7; Figure 38.3 §38.5; Figures 38.1 and 38.2; §§7.2, 7.3, and 7.4 §38.6
II. Post-Newtonian (“solar-system”) experiments	1. Deflection of light and radio waves by Sun 2. Relativistic delay in round-trip travel time for radar beams passing near Sun 3. Perihelion shifts and periodic perturbations in planetary orbits 4. Three-body effects in the Lunar orbit 5. Precession of gyroscopes (“geodetic precession” and precession due to dragging of inertial frames by Earth’s rotation) 6. Spatial variation of the Cavendish gravitational constant in the solar system 7. Dependence of the Cavendish gravitational constant on the chemical composition of the gravitating body 8. Earth tides with sidereal periods 9. Annual variations in Earth rotation rate 10. Periodicities in Earth-Moon separation due to breakdown of geodesic motion	§40.3; Box 40.1 §40.4; Box 40.2 §40.5; Box 40.3 §40.6 §40.7 §§40.8 and 40.9 §40.8 §40.8 §40.8 §40.9
III. Cosmological observations	1. Change of Cavendish gravitational constant with time in solar system 2. Large-scale features of universe (expansion, isotropy, homogeneity; existence and properties of cosmic microwave radiation; . . .) 3. Agreement of various measures of age of universe (age from expansion; ages of oldest stars; age of solar system)	§40.8 Chapters 27–30; especially Chapter 29 §29.7
IV. Gravitational-Wave experiments	Existence of waves; propagation speed; polarization properties; . . .	Chapters 35–37; especially Chapter 37

2

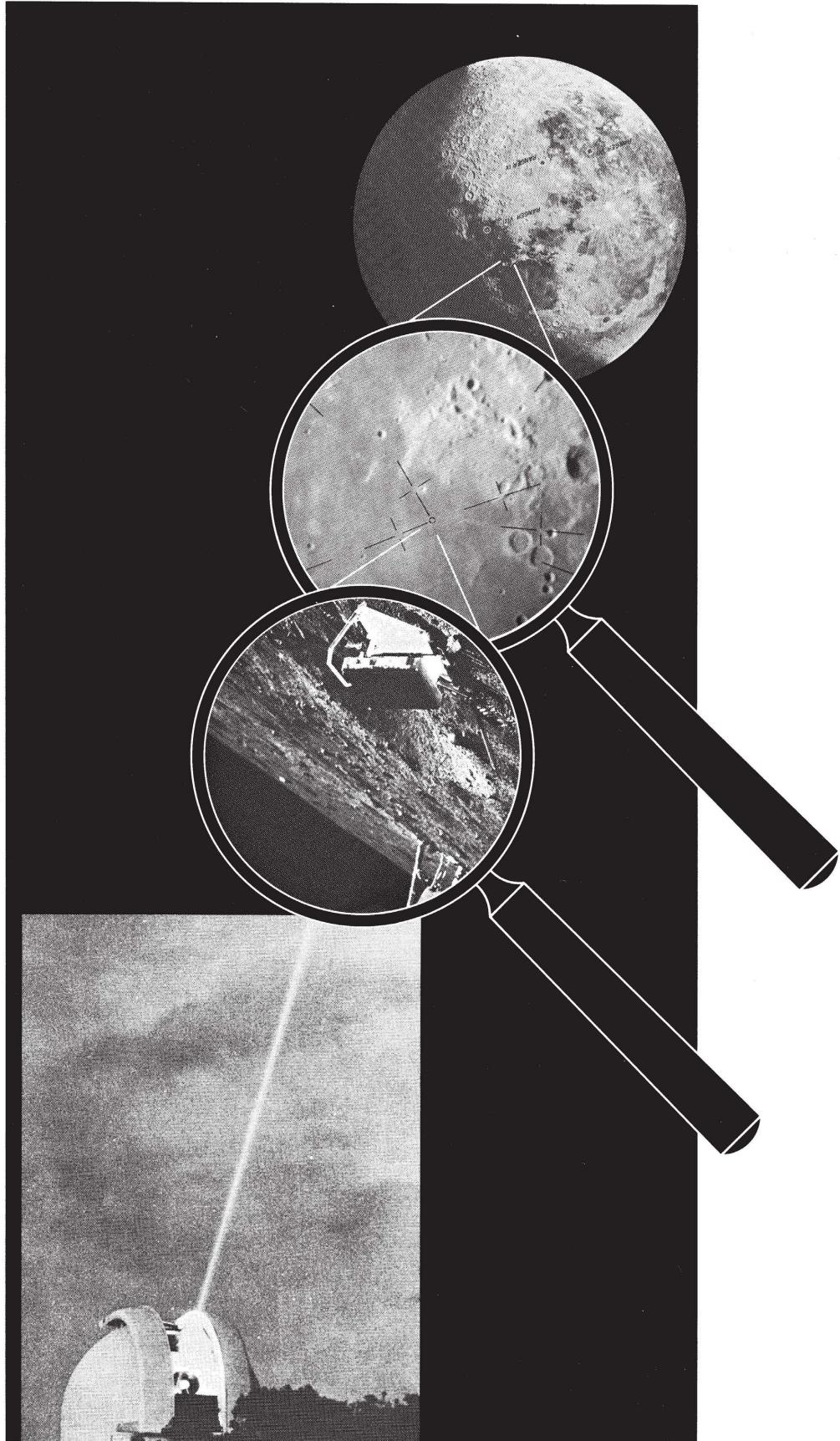


Figure 40.4. (facing page)

Measuring the separation between earth and moon by determining the time-delay (about 2.5 sec) between the emission of light from a laser on the earth and the return of this light to the earth. A key element in the program is a corner reflector, the first of which was landed on the moon July 20, 1969, by the Apollo 11 flight crew. In November 1971, there were three such reflectors on the moon: two American, and one French-built and Soviet-landed. A pulsed ruby laser projects a beam out of the 107-inch reflecting telescope of the McDonald Observatory of the University of Texas, on Mount Locke, 119 miles east of El Paso. This beam makes a spot of light on the moon's surface about 3.2 km in diameter. Laser light is bounced straight back to the earth by the "laser ranging retroreflectors" (LR³). Each consists of an aluminum panel of 46 cm by 46 cm with 100 fused silica corner cubes each 3.8 cm in diameter. The first reflector ever set up appears in the first inset, near the lunar landing module. It is tilted with respect to the landscape of the moon. The photograph was made shortly before astronauts Neil A. Armstrong and Edwin E. Aldrin, Jr., took off for the earth. The second inset is a photograph made by D. G. Currie of the field of view in the guiding eyepiece of the McDonald 107-inch telescope in an interval when the laser was not firing at the Apollo 11 site. One guides the telescope to Tranquility Base (small circle) by aligning fiducial marks on more visible moonscape features. In November 1971, the LR³ experiment and continuing time-of-flight measurements were the responsibility of the National Aeronautics and Space Administration and a Lunar Retroreflecting Ranging Team of representatives from several centers of research. One of the members of this team, Carroll Alley, of the University of Maryland, is hereby thanked for his kindness in providing the photographs used in this montage. Thanks to this NASA work, the distance between the laser source on the earth and the reflectors on the moon is known with an accuracy now better than half a meter. The astronauts left behind on the moon not only LR³ and a seismometer and other equipment, but also a plaque: "We came in peace for all mankind."

By the mid 1970's, lunar laser-ranging data will probably be able to determine the amplitudes of this polarization to a precision of one meter or better [see Bender *et al.* (1971); also Figure 40.4].

§40.10. SUMMARY OF EXPERIMENTAL TESTS OF GENERAL RELATIVITY

No longer is general relativity "a theorist's Paradise, but an experimentalist's Hell." It is now a Paradise for all—as one can see quickly by perusing the catalog of experiments given in Box 40.4 on page 1129. Moreover, general relativity has emerged from each of its tests unscathed—a remarkable 1973 tribute to the 1915 genius of Albert Einstein.

PART X

FRONTIERS

Wherein the reader—who, during a life of continued variety for forty chapters (besides the Preface), was eight chapters a mathematician, four times enticed (once by an old friend), four chapters a cosmologist, and four chapters a transported astrophysicist in the land of black holes, and who at last inherited a wealth of experiments, lived honest, and became a True Believer—now ventures forth in search of new frontiers to conquer.

CHAPTER 41

SPINORS

§41.1. REFLECTIONS, ROTATIONS, AND THE COMBINATION OF ROTATIONS

Spinors and their applications in relativity grew out of the analysis of “rotations,” first in space, then in spacetime. Take a cube (Figure 41.1). Rotate it about one axis through 90° . Then pick another axis at right angles to the first. About it rotate the cube again through 90° . In this way the cube is carried from the orientation marked “Initial” to that marked “Final.” How can one make this net transformation in a single step, with a single rotation? In other words, what is the law for the combination of rotations?

Were rotations described by vectors, then one could apply the law of combination of vectors. The resultant of two vectors of the same magnitude (90°) separated by a right angle, is a single vector that (1) lies in the same plane and (2) has the magnitude $2^{1/2} \times 90^\circ = 127.28^\circ$. Both predictions are wrong. To turn the cube from initial to final orientation in a single turn, (1) take an axis running from the center through the vertex *A* and (2) rotate through 120° .

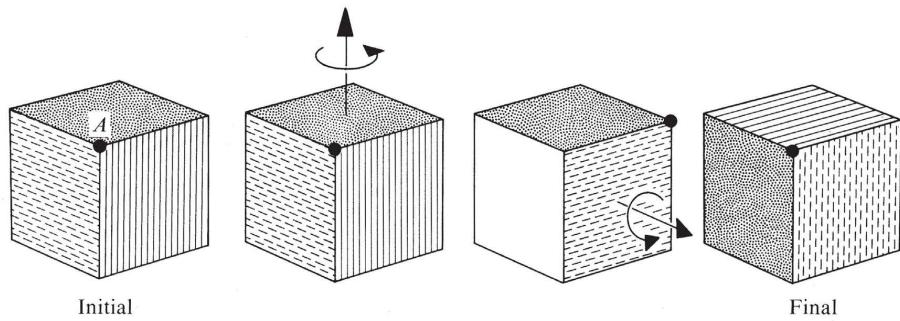
What computational algorithm can ever reproduce a law of combination of rotations apparently so strange? On the evening of October 16, 1843, William Rowan Hamilton was walking with his wife along the Royal Canal in Dublin when the answer leaped to his mind, the fruit of years of reflection. With his knife he then and there carved on a stone on Brougham Bridge the formulas*

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1,$$

This chapter is entirely Track 2. No earlier Track-2 material is needed as preparation for it, nor is it needed as preparation for any later chapter.

The problem of combining rotations

*In the same city on June 21, 1972 President Eamon de Valera told one of the authors that, while in jail one evening in 1916, scheduled to be shot the next morning, he wrote down the formula of which he was so fond, $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$.

**Figure 41.1.**

Rotation about the vertical axis through 90° , followed by rotation about the horizontal axis through 90° , gives a net change in orientation that can be achieved by a single rotation through 120° about an axis emergent from the center through the corner A .

which in today's notation,

$$\sigma_x = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = i\mathbf{i}, \quad \sigma_y = \begin{vmatrix} 0 & -i \\ i & 0 \end{vmatrix} = i\mathbf{j}, \quad \sigma_z = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} = i\mathbf{k}, \quad (41.1)$$

take the form

$$\begin{aligned} \sigma_x^2 &= \sigma_y^2 = \sigma_z^2 = 1, \\ \sigma_x \sigma_y &= -\sigma_y \sigma_x = i\sigma_z \text{ (and cyclic permutations).} \end{aligned} \quad (41.2)$$

Rotation operators:
(1) defined

To any rotation is associated a quantity (Hamilton's "quaternion;" today's "spin matrix" or "spinor transformation" or "rotation operator")

$$R = \cos(\theta/2) - i \sin(\theta/2)(\sigma_x \cos \alpha + \sigma_y \cos \beta + \sigma_z \cos \gamma), \quad (41.3)$$

(2) as tools in combining rotations

where θ is the angle of rotation and α, β, γ are the angles between the axis of rotation and the coordinate axes. A rotation described by R_1 followed by a rotation described by R_2 gives a net change in orientation described by the single rotation

$$R_3 = R_2 R_1. \quad (41.4)$$

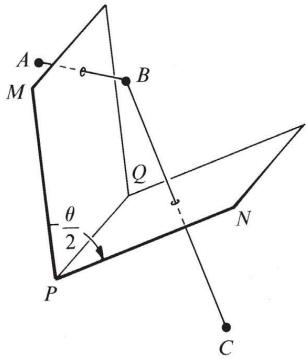
This is Hamilton's formula for the combination of two rotations (steps toward it by Euler in 1776; obtained by Gauss in 1819 but never published by him).

In the example in Figure 41.1,

$$\begin{aligned} R_1(\text{rotation by } \theta = 90^\circ \text{ about } z\text{-axis}) &= (1 - i\sigma_z)/2^{1/2}, \\ R_2(\text{rotation by } \theta = 90^\circ \text{ about } x\text{-axis}) &= (1 - i\sigma_x)/2^{1/2}, \end{aligned}$$

and the product of the two is

$$\begin{aligned} R_2 R_1 &= (1 - i\sigma_x + i\sigma_y - i\sigma_z)/2 \\ &= \cos 60^\circ - i \sin 60^\circ (\sigma_x/3^{1/2} - \sigma_y/3^{1/2} + \sigma_z/3^{1/2}). \end{aligned}$$

**Figure 41.2.**

Reflection in the plane MPQ carries A to B . Reflection in the plane NPQ carries B to C . The combination of the two reflections in the two planes separated by the angle $\theta/2$ produces the same end result (transformation from A to C) as rotation through the angle θ about the line PQ .

According to Hamilton's rule (41.3), this result implies a net rotation through 120° about a line that makes equal angles with the x -axis, the y -axis, and the z -axis, in conformity with what one already saw in Figure 41.1 (axis of rotation running from center of cube through the corner A).

What one has just done in the special example one can do in the general case: obtain the parameters $\theta_3, \alpha_3, \beta_3, \gamma_3$ of the net rotation (four unknowns!) by identifying the four coefficients of the four Hamilton units $1, -i\sigma_x, -i\sigma_y, -i\sigma_z$ on both sides of the equation $R_3 = R_2R_1$. In this way one arrives at the four prequaternion formulas of Olinde Rodrigues (1840) for the combination of the two rotations.

Why do half-angles put in an appearance? And what is behind the law of combination of rotations? The answer to both questions is the same: a rotation through the angle θ about a given axis may be visualized as the consequence of successive reflections in two planes that meet along that axis at the angle $\theta/2$ (Figure 41.2). Two rotations therefore imply four reflections. However, it can be arranged that reflections no. 2 and no. 3 take place in the same plane, the plane that includes the two axes of rotation. Then reflection no. 3 exactly undoes reflection no. 2. By now there remain only reflections no. 1 and no. 4, which together constitute one rotation: the net rotation that was desired (Figures 41.3 and 41.4).

The rotation

$$R = \cos(\theta/2) - i \sin(\theta/2)(\sigma_x \cos \alpha + \sigma_y \cos \beta + \sigma_z \cos \gamma) \quad (41.3)$$

Geometric reason that half angles appear in rotation operators

Algebraic properties of rotation operators

is undone by the inverse rotation

$$R^{-1} = \cos(\theta/2) + i \sin(\theta/2)(\sigma_x \cos \alpha + \sigma_y \cos \beta + \sigma_z \cos \gamma). \quad (41.3')$$

Thus the product of the two rotation operators

$$RR^{-1} = R^{-1}R = 1 \quad (41.5)$$

is an operator, the unit operator, that leaves unchanged everything that it acts on. The reciprocal R^{-1} of the combination $R = R_2R_1$ of two rotations is

$$R^{-1} = R_1^{-1}R_2^{-1} \quad (41.5')$$

(reverse order of factors!), as one verifies by substitution into (41.5).

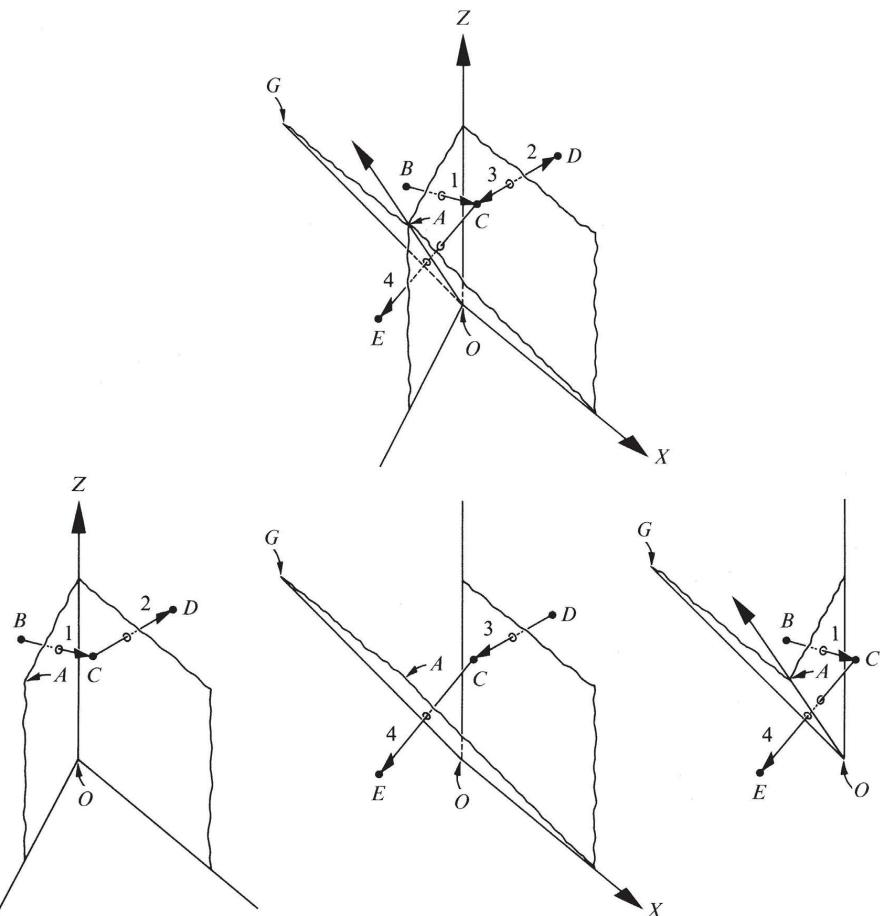


Figure 41.3.

Composition of two rotations seen in terms of reflections. The first rotation (for instance, 90° about OZ in the example of Figure 41.1.) is represented in terms of reflection 1 followed by reflection 2 (the planes of the two reflections being separated by $90^\circ/2 = 45^\circ$ in the example). The second reflection appears as the resultant of reflections 3 and 4. But the reflections 2 and 3 take place in the common plane ZOX . Therefore one reflection undoes the other. Thus the sequence of four operations 1234 collapses to the two reflections 1 and 4. Their place in turn is taken by a single rotation about the axis OA .

The conjugate transpose, M^* , of a matrix M is obtained by taking the conjugate complex of every element in the matrix and then interchanging rows and columns. By direct inspection of matrix expressions (41.1) one sees that $\sigma_x^* = \sigma_x$, $\sigma_y^* = \sigma_y$, $\sigma_z^* = \sigma_z$. Such matrices are said to be Hermitian. The conjugate transpose of the product $M = PQ$ of two matrices is the product $M^* = Q^*P^*$ of the individual conjugate transposed matrices taken in the reverse order. For the rotation matrix written down above, note that $R^* = R^{-1}$. Such a matrix is said to be unitary. The

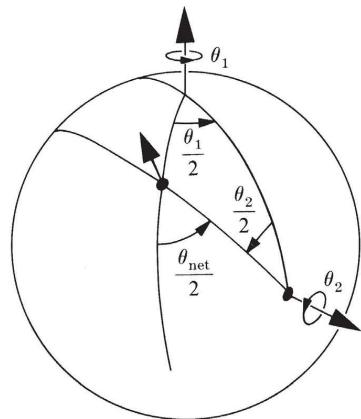


Figure 41.4.
Law of composition of rotations epitomized by a spherical triangle in which each of the three important angles represents *half* an angle of rotation.

determinant of a unitary matrix may be seen to have absolute value unity from the following line of argument:

$$\begin{aligned} 1 &= \det(\text{unit matrix}) = \det(RR^{-1}) \\ &= \det(RR^*) = \det R \det R^* \\ &= |\det R|^2. \end{aligned} \quad (41.6)$$

In actuality the determinant of the rotation spin matrix is necessarily unity ("uni-modular matrix") as shown in the following exercises

Exercise 41.1. ELEMENTARY FEATURES OF THE ROTATION MATRIX

Write equation (41.3) in the form

$$R(\theta) = \cos(\theta/2) - i \sin(\theta/2)(\boldsymbol{\sigma} \cdot \mathbf{n}),$$

and establish the following properties:

- (a) $(\boldsymbol{\sigma} \cdot \mathbf{n})^2 = 1 \equiv \text{unit matrix};$
- (b) $\text{tr}(\boldsymbol{\sigma} \cdot \mathbf{n}) = 0$ (tr means "trace," i.e., sum of diagonal elements);
- (c) $\underbrace{[R, (\boldsymbol{\sigma} \cdot \mathbf{n})]}_{\substack{\downarrow \\ \text{commutator}}} \equiv R(\boldsymbol{\sigma} \cdot \mathbf{n}) - (\boldsymbol{\sigma} \cdot \mathbf{n})R = 0;$
- (d) $\frac{dR}{d\theta} = -\frac{i}{2}(\boldsymbol{\sigma} \cdot \mathbf{n})R. \quad (41.7)$

[Note that if one thinks of θ as increasing with angular velocity ω , so $d\theta/dt = \omega = \text{constant}$, then this last equation reads

$$\frac{dR}{dt} = -\frac{i}{2}(\boldsymbol{\sigma} \cdot \boldsymbol{\omega})R \quad (41.7')$$

where $\boldsymbol{\omega} = \omega \mathbf{n}.$]

EXERCISES

Exercise 41.2. ROTATION MATRIX HAS UNIT DETERMINANT

Recall from exercise 5.5 that for any matrix M one has

$$d[\ln(\det M)] = \text{tr}(M^{-1} dM)$$

and use this to show that $\det R$ in (41.7) is constant, and therefore equal to $(\det R)_{\theta=0} = 1$.

§41.2. INFINITESIMAL ROTATIONS

Infinitesimal rotations

A given rotation can be obtained by performing in turn two rotations of half the magnitude, or four rotations of a fourth the magnitude, or eight of an eighth the magnitude, and so on. Thus one arrives in the limit at the concept of an infinitesimal rotation described by the spin matrix

$$R = 1 - (i/2)(\sigma_x d\theta_{yz} + \sigma_y d\theta_{zx} + \sigma_z d\theta_{xy})$$

or

$$R = 1 - (i d\theta/2)(\boldsymbol{\sigma} \cdot \mathbf{n}). \quad (41.8)$$

Here the quantities

$$\begin{aligned} d\theta_{yz} &= -d\theta_{zy} = n^x d\theta = \cos \alpha d\theta, \\ d\theta_{zx} &= -d\theta_{xz} = n^y d\theta = \cos \beta d\theta, \\ d\theta_{xy} &= -d\theta_{yx} = n^z d\theta = \cos \gamma d\theta, \end{aligned} \quad (41.9)$$

are the components of the infinitesimal rotation in the three indicated planes. An infinitesimal rotation in the (x, y) -plane through the angle $d\theta_{xy}$ transforms the vector $\mathbf{x} = (x, y, z)$ into a new vector with changed components x' and y' but with unchanged component $z' = z$. More generally, the infinitesimal rotation (41.8) considered in this same “active” sense* produces the transformation

$$\mathbf{x} \longrightarrow \mathbf{x}',$$

with

$$\begin{aligned} x' &= x - (d\theta_{xy})y - (d\theta_{xz})z, \\ y' &= -(d\theta_{yx})x + y - (d\theta_{yz})z, \\ z' &= -(d\theta_{zx})x - (d\theta_{zy})y + z. \end{aligned} \quad (41.10)$$

Representation of a 3-vector as a spin matrix

Spinor calculus provides an alternative (and shorthand!) means to calculate the foregoing effect of a rotation on a vector. Associate with the vector \mathbf{x} the spin matrix

$$X = x\sigma_x + y\sigma_y + z\sigma_z = (\mathbf{x} \cdot \boldsymbol{\sigma}), \quad (41.11)$$

* An “active” transformation changes one vector into another, while leaving unchanged the underlying reference frame (if there is one). By contrast, a “passive” transformation leaves all vectors unchanged, but alters the reference frame. All transformations in previous chapters of this book were passive.

and with the vector \mathbf{x}' a corresponding spin matrix or quaternion X' . Then the effect of the rotation is summarized in the formula

$$X \longrightarrow X' = RXR^*. \quad (41.12)$$

Rotation of a 3-vector described in spin-matrix language

Test this formula for the general infinitesimal rotation (41.10). It reads

$$(\mathbf{x}' \cdot \boldsymbol{\sigma}) = [1 - (i d\theta/2)(\boldsymbol{\sigma} \cdot \mathbf{n})](\mathbf{x} \cdot \boldsymbol{\sigma})[1 + (i d\theta/2)(\boldsymbol{\sigma} \cdot \mathbf{n})]$$

or, to the first order in the quantity $d\theta$,

$$(\mathbf{x}' \cdot \boldsymbol{\sigma}) = (\mathbf{x} \cdot \boldsymbol{\sigma}) + (i d\theta/2)[(\mathbf{x} \cdot \boldsymbol{\sigma})(\boldsymbol{\sigma} \cdot \mathbf{n}) - (\boldsymbol{\sigma} \cdot \mathbf{n})(\mathbf{x} \cdot \boldsymbol{\sigma})]. \quad (41.13)$$

The product of spin matrices $A = (\mathbf{a} \cdot \boldsymbol{\sigma})$ and $B = (\mathbf{b} \cdot \boldsymbol{\sigma})$ built from two distinct vectors \mathbf{a} and \mathbf{b} is

$$AB = (\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) = a^x b^x \sigma_x^2 + a^x b^y \sigma_x \sigma_y + \dots,$$

or, according to (41.2),

$$AB = (\mathbf{a} \cdot \mathbf{b}) + i(\mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma}. \quad (41.14)$$

Employ this formula to evaluate the right-hand side of (41.13). In the square brackets, the terms in $(\mathbf{x} \cdot \mathbf{n})$ have opposite signs and cancel. In contrast, the terms in $(\mathbf{n} \times \mathbf{x})$ have the same sign. They combine to cancel the factor 2 in $(d\theta/2)$. End up with

$$(\mathbf{x}' \cdot \boldsymbol{\sigma}) = (\mathbf{x} \cdot \boldsymbol{\sigma}) + d\theta(\mathbf{n} \times \mathbf{x}) \cdot \boldsymbol{\sigma}$$

or

$$\mathbf{x}' = [1 + (d\theta)\mathbf{n} \times] \mathbf{x} \quad (41.15)$$

in agreement with (41.10), as was to be shown.

A finite rotation about a given axis can be considered as the composition of infinitesimal rotations about that axis. To see this composition in simplest form, rewrite the spin matrix (41.8) associated with the general infinitesimal rotation as

$$R(d\theta) = e^{-(i d\theta/2)(\boldsymbol{\sigma} \cdot \mathbf{n})} \quad (41.16)$$

Composition of finite rotation from infinitesimal rotations

(exponential function defined by its power-series expansion). Note that $(\boldsymbol{\sigma} \cdot \mathbf{n})$ commutes (a) with unity and (b) with itself, and in addition (c) has a unit square. Therefore the calculation of the exponential function proceeds no differently here, for spin matrices, than for everyday algebra. The composition of the spin matrices for infinitesimal rotations about an unchanging axis proceeds by adding exponents, to give

$$R(\theta) = e^{-i(\theta/2)(\boldsymbol{\sigma} \cdot \mathbf{n})}, \quad (41.17)$$

which can also be obtained immediately from equation (41.7). This expression can be put in another form by developing the power series; thus,

$$\begin{aligned}
 R(\theta) &= \sum_{p=0}^{\infty} (1/p!)(-i\theta\boldsymbol{\sigma} \cdot \mathbf{n}/2)^p \\
 &= \sum_{\text{even } p} (1/p!)(-i\theta/2)^p + (\boldsymbol{\sigma} \cdot \mathbf{n}) \sum_{\text{odd } p} (1/p!)(-i\theta/2)^p \\
 &= \cos(\theta/2) - i \sin(\theta/2)(\boldsymbol{\sigma} \cdot \mathbf{n})
 \end{aligned} \tag{41.18}$$

in agreement with the expression (41.3) originally given for a spinor transformation. The effect of one infinitesimal rotation after another after another . . . on a vector is given by

$$X' = R(d\theta) \dots R(d\theta) X R^*(d\theta) \dots R^*(d\theta),$$

with the consequence that even for a finite rotation $R = R(\theta)$ one is correct in employing the formula

$$X' = RXR^*. \tag{41.19}$$

EXERCISE

Exercise 41.3. MORE PROPERTIES OF THE ROTATION MATRIX

Show that for $X = \mathbf{x} \cdot \boldsymbol{\sigma}$ one has the commutation relation

$$[(\boldsymbol{\sigma} \cdot \mathbf{n}), X] = 2i(\mathbf{n} \times \mathbf{x}) \cdot \boldsymbol{\sigma}.$$

Use this to obtain, from equation (41.19) in the form $X = RX_0R^*$ [where X_0 is constant, while $R(\theta)$ is given by equation (41.17)], the formula

$$\frac{d}{d\theta}(\mathbf{x} \cdot \boldsymbol{\sigma}) = (\mathbf{n} \times \mathbf{x}) \cdot \boldsymbol{\sigma}.$$

Why is this equivalent to the standard definition

$$\frac{dx}{dt} = \omega \times \mathbf{x}$$

for the angular velocity? Reverse the argument to show that equation (41.7') correctly defines the rotation $R(t)$ resulting from a time-dependent angular velocity $\omega(t)$, even though the simple solution $R = \exp[-\frac{1}{2}it(\boldsymbol{\sigma} \cdot \boldsymbol{\omega})]$ of this equation can no longer be written when $\boldsymbol{\omega}$ is not constant.

§41.3. LORENTZ TRANSFORMATION VIA SPINOR ALGEBRA

4-vectors and Lorentz transformations in spin-matrix language

Generate a rotation by two reflections in space? Then why not generate a Lorentz transformation by two reflections in spacetime? If for this purpose one has to turn from a real half-angle between the two planes of reflection to a complex half-angle, that development will come as no surprise; nor will it be a surprise that one can

still represent the effect of the Lorentz transformation by a matrix multiplication of the form

$$X \longrightarrow X' = LXL^*. \quad (41.20)$$

Here the “Lorentz spin transformation matrix” L is a generalization of the rotation matrix, R . Also the “coordinate-generating spin matrix” X is now generalized from (41.11) to

$$X = t + (\mathbf{x} \cdot \boldsymbol{\sigma}) \quad (41.21)$$

or

$$X = \begin{vmatrix} t+z & x-iy \\ x+iy & t-z \end{vmatrix}. \quad (41.22)$$

It is demanded that this matrix be Hermitian

$$X = X^*. \quad (41.23)$$

Then and only then are the coordinates (t, x, y, z) real. The conjugate transpose of the transformed spin matrix must also be Hermitian—and is:

$$\begin{aligned} (X')^* &= (LXL^*)^* \\ &= (L^*)^*(X)^*(L)^* = LXL^* = X'. \end{aligned} \quad (41.24)$$

Therefore the new coordinates (t', x', y', z') are guaranteed to be real, as desired. This reality requirement is a rationale for the form of the spin-matrix transformation (41.20), with L appearing on one side of X and L^* on the other.

A Lorentz transformation is defined by the circumstance that it leaves the interval invariant:

$$t'^2 - x'^2 - y'^2 - z'^2 = t^2 - x^2 - y^2 - z^2. \quad (41.25)$$

Note that the determinant of the matrix X as written out above has the value

$$\det X = t^2 - x^2 - y^2 - z^2. \quad (41.26)$$

Consequently the requirement for the preservation of the interval may be put in the form

$$\det X' = \det X \quad (41.27)$$

or

$$(\det L)(\det X)(\det L^*) = \det X. \quad (41.28)$$

This requirement is fulfilled by demanding

$$\det L = 1 \quad (41.29)$$

[it is not a useful generalization to multiply every element of L here by a common phase factor $e^{i\delta}$, and therefore multiply $\det L$ by $e^{2i\delta}$, because the net effect of this phase factor is nil in the formula $X' = LXL^*$].

Infinitesimal Lorentz transformations

The spin matrix associated with a rotation, whether finite or infinitesimal, already satisfied the condition $\det L = 1$ [proved in exercise (41.2)]. This condition, being algebraic, will continue to hold when the real angles $d\theta_{yz}$, $d\theta_{zx}$, $d\theta_{xy}$, are replaced by complex angles, $d\theta_{yz} + i d\alpha_x$, $d\theta_{zx} + i d\alpha_y$, $d\theta_{xy} + i d\alpha_z$. The spin-transformation matrix acquires in this way a total of six parameters, as needed to describe the general infinitesimal Lorentz transformation. Thus the spin matrix for the general infinitesimal Lorentz transformation can be put in the form

$$\begin{aligned} L &= 1 - (i/2)(\sigma_x d\theta_{yz} + \sigma_y d\theta_{zx} + \sigma_z d\theta_{xy}) \\ &\quad + (1/2)(\sigma_x d\alpha_x + \sigma_y d\alpha_y + \sigma_z d\alpha_z) \\ &= 1 - (i d\theta/2)(\boldsymbol{\sigma} \cdot \mathbf{n}) + (\boldsymbol{\sigma} \cdot d\boldsymbol{\alpha}/2). \end{aligned} \quad (41.30)$$

The effect of this transformation upon the coordinates is to be read out from the formula

$$X \longrightarrow X' = LXL^*$$

or

$$\begin{aligned} t' + (\boldsymbol{\sigma} \cdot \mathbf{x}') &= [1 - (i d\theta/2)(\boldsymbol{\sigma} \cdot \mathbf{n}) + (\boldsymbol{\sigma} \cdot d\boldsymbol{\alpha}/2)] \\ &\quad \times [t + (\boldsymbol{\sigma} \cdot \mathbf{x})][1 + (i d\theta/2)(\boldsymbol{\sigma} \cdot \mathbf{n}) + (\boldsymbol{\sigma} \cdot d\boldsymbol{\alpha}/2)] \end{aligned} \quad (41.31)$$

Employ equation (41.14) for $(\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \mathbf{B})$ to reduce the right side to the form

$$t + (\boldsymbol{\sigma} \cdot \mathbf{x}) + (\boldsymbol{\sigma} \cdot d\boldsymbol{\alpha})t + d\theta(\mathbf{n} \times \mathbf{x}) \cdot \boldsymbol{\sigma} + (\mathbf{x} \cdot d\boldsymbol{\alpha}).$$

Now compare coefficients of 1, σ_x , σ_y and σ_z , respectively, on both sides of the equation, and find

$$\begin{aligned} t' &= t + (\mathbf{x} \cdot d\boldsymbol{\alpha}) \\ \mathbf{x}' &= \mathbf{x} + t d\boldsymbol{\alpha} + d\theta(\mathbf{n} \times \mathbf{x}), \end{aligned} \quad (41.32)$$

in agreement with the conventional expression for an infinitesimal Lorentz transformation or “boost” of velocity $d\boldsymbol{\alpha}$, in active form, as was to be shown.

The composition of such infinitesimal Lorentz transformations gives a finite Lorentz transformation. The result, however, can be calculated easily only when all infinitesimal transformations commute. Thus assume that $d\theta$ and $d\boldsymbol{\alpha}$ are in a fixed ratio, so

$$\omega \equiv \mathbf{n} \frac{d\theta}{d\tau} \text{ and } \mathbf{a} \equiv \frac{d\boldsymbol{\alpha}}{d\tau}$$

are constants, with τ a parameter. Then integration with respect to τ (composition of infinitesimal transformations) gives a finite transformation $L = \exp[-\frac{1}{2}i\tau\boldsymbol{\sigma} \cdot (\omega + i\mathbf{a})]$. For $\tau = 1$, so $\theta\mathbf{n} = \omega\tau$, $\boldsymbol{\alpha} = \mathbf{a}\tau$, this reads

$$L = \exp[(\boldsymbol{\alpha} - i\theta\mathbf{n}) \cdot \boldsymbol{\sigma}/2]. \quad (41.33)$$

Composition of finite Lorentz transformations from infinitesimal transformations

In the special case of a pure boost (no rotation; $\theta = 0$), the exponential function is evaluated along the lines indicated in (41.18), with the result

$$L = \cosh(\alpha/2) + (\mathbf{n}_\alpha \cdot \boldsymbol{\sigma}) \sinh(\alpha/2). \quad (41.34)$$

Here $\mathbf{n}_\alpha = \boldsymbol{\alpha}/\alpha$ is a unit vector in the direction of the boost. The corresponding Lorentz transformation itself is evaluated from the formula

$$X' = LXL^*$$

or

$$\begin{aligned} t' + (\mathbf{x}' \cdot \boldsymbol{\sigma}) &= [\cosh \alpha/2 + (\mathbf{n}_\alpha \cdot \boldsymbol{\sigma}) \sinh \alpha/2][t + (\mathbf{x} \cdot \boldsymbol{\sigma})] \\ &\quad \times [\cosh \alpha/2 + (\mathbf{n}_\alpha \cdot \boldsymbol{\sigma}) \sinh \alpha/2]. \end{aligned} \quad (41.35)$$

Simplify with the help of the relations

$$\cosh^2(\alpha/2) + \sinh^2(\alpha/2) = \cosh \alpha,$$

$$2 \sinh(\alpha/2) \cosh(\alpha/2) = \sinh \alpha,$$

and

$$(\mathbf{n}_\alpha \cdot \boldsymbol{\sigma})(\mathbf{x} \cdot \boldsymbol{\sigma})(\mathbf{n}_\alpha \cdot \boldsymbol{\sigma}) = 2(\mathbf{n}_\alpha \cdot \mathbf{x})(\mathbf{n}_\alpha \cdot \boldsymbol{\sigma}) - (\mathbf{n}_\alpha \cdot \mathbf{n}_\alpha)(\mathbf{x} \cdot \boldsymbol{\sigma}),$$

and on both sides of the equation compare coefficients of 1 and $\boldsymbol{\sigma}$, to find

$$t' = (\cosh \alpha)t + (\sinh \alpha)(\mathbf{n}_\alpha \cdot \mathbf{x}),$$

$$\begin{aligned} \mathbf{x}' &= [(\sinh \alpha)\mathbf{n}_\alpha t + (\cosh \alpha)(\mathbf{n}_\alpha \cdot \mathbf{x})\mathbf{n}_\alpha] \text{ ("in-line part of transformation")} \\ &\quad + [\mathbf{x} - (\mathbf{x} \cdot \mathbf{n}_\alpha)\mathbf{n}_\alpha] \text{ ("perpendicular part of } \mathbf{x} \text{ unchanged").} \end{aligned} \quad (41.36)$$

In this way one verifies that the quantity α is the usual "velocity parameter," connected with the velocity itself by the relations

$$\begin{aligned} (1 - \beta^2)^{-1/2} &= \cosh \alpha, \\ \beta(1 - \beta^2)^{-1/2} &= \sinh \alpha, \\ \beta &= \tanh \alpha. \end{aligned} \quad (41.37)$$

That velocity parameters add for successive boosts in the same direction shows nowhere more clearly than in the representation (41.33) of the spin-transformation matrix:

$$\begin{aligned} L(\alpha_2)L(\alpha_1) &= \exp[\alpha_2(\mathbf{n}_\alpha \cdot \boldsymbol{\sigma})/2] \exp[\alpha_1(\mathbf{n}_\alpha \cdot \boldsymbol{\sigma})/2] = \exp[(\alpha_2 + \alpha_1)(\mathbf{n}_\alpha \cdot \boldsymbol{\sigma})/2] \\ &= L(\alpha_2 + \alpha_1). \end{aligned} \quad (41.38)$$

Turn from this special case, and ask how to get the resultant of two arbitrary Lorentz transformations, each of which is a mixture of a rotation and a boost. No simpler method offers itself to answer this question than to use formula (41.33) together with the equation

$$L(\text{resultant}) = L_2L_1. \quad (41.39)$$

§41.4. THOMAS PRECESSION VIA SPINOR ALGEBRA

A spinning object, free of all torque, but undergoing acceleration, changes its direction as this direction is recorded in an inertial frame of reference. This is the

Thomas precession [see exercise 6.9 and first term in equation (40.33b)]. This precession accounts for a factor two in the effective energy of coupling of spin and orbital angular momentum of an atomic electron. In a nucleus it contributes a little to the coupling of the spin and orbit of a nucleon. The evaluation of the Thomas precession affords an illustration of spin-matrix methods in action.

**Origin of Thomas precession:
composition of two boosts
is not a pure boost**

The precession in question can be discussed quite without reference either to angular momentum or to mass in motion. It is enough to consider a sequence of inertial frames of reference $S(t)$ with these two features. (1) To whatever point the motion has taken the mass at time t , at that point is located the origin of the frame $S(t)$. (2) The inertial frame $S(t + dt)$ at the next succeeding moment has undergone no rotation with respect to the inertial frame $S(t)$, as rotation is conceived by an observer in that inertial frame. However, it has undergone a rotation ("Thomas precession") as rotation is conceived and defined in the laboratory frame of reference.

How is it possible for "no rotation" to appear as "rotation"? The answer is this: one pure boost, followed by another pure boost in another direction, does not have as net result a third pure boost; instead, the net result is a boost plus a rotation. This idea is not new in kind. Figure 41.1 illustrated how a rotation about the z -axis followed by a rotation about the x -axis had as resultant a rotation about an axis with not only an x -component and a z -component but also a y -component. What is true of rotations is true of boosts: they defy the law for the addition of vectors.

**Derivation of Thomas
precession using spin
matrices**

Let the frame S_0 coincide with the laboratory frame, and let the origin of this laboratory frame be where the moving frame is at time t . Let $S(t)$ be a Lorentz frame moving with this point at time t . Let one pure boost raise its velocity relative to the laboratory from β to $\beta + d\beta$. The resulting final configuration cannot be reached from S_0 by a pure boost. Instead, first turn S_0 relative to the laboratory frame ("rotation R associated with the Thomas precession") and then send it by a simple boost to the final configuration. Only one choice of this rotation will be right to produce match-up. Thus, distinguishing the spin matrices for pure boosts and pure rotations by the letters B and R , one has the relation

$$B(\beta + d\beta)R(\omega dt) = "B(d\beta)"B(\beta) \quad (41.40)$$

out of which to find the angular velocity ω of the Thomas precession. The quotation marks in " $B(d\beta)$ " carry a double warning: (1) the velocity of transformation that boosts $S(t)$ to $S(t + dt)$ is not $(\beta + d\beta) - \beta = d\beta$ (law of vector addition—or subtraction—not applicable to velocity), and (2) " $B(d\beta)$ " does not appear as a pure boost in the laboratory frame. It appears as a pure boost only in the comoving frame.

Take care of the second difficulty first. It is only a difficulty because the formalism for combination of transformations, $R_3 = R_2R_1$, as developed in §41.1 presupposes all operations R_1, R_2, \dots , to be defined and carried out in the laboratory reference frame. In contrast, the quantity " $B(d\beta)$ " is understood to imply a pure boost as defined and carried out in the comoving frame. Such an operation can be fitted into the formalism as follows. (1) Undo any velocity that the object already has. In other words apply the operator $B^{-1}(\beta)$. Then the object is at rest in the laboratory frame. Then apply the necessary small pure boost, $B(a_{\text{comoving}} d\tau)$, where a_{comoving}

is the acceleration as it will be sensed by the object and $d\tau$ is the lapse of proper time as it will be sensed by the object. At the commencement of this brief acceleration the object is at rest relative to the laboratory. What is a pure boost to it is a pure boost relative to the laboratory. It is also a pure boost in the spin-matrix formalism. Then transform back from laboratory to moving frame. Thus have the relation

$$\text{“}B(d\beta)\text{”} = B(\beta)B(a_{\text{comoving}} d\tau)B^{-1}(\beta). \quad (41.41)$$

The equation for the determination of the Thomas precession now reads

$$B(\beta + d\beta)R(\omega dt) = B(\beta)B(a_{\text{comoving}} d\tau) \quad (41.42)$$

or, with all unknowns put on the left,

$$R(\omega dt)B^{-1}(a_{\text{comoving}} d\tau) = B^{-1}(\beta + d\beta)B(\beta). \quad (41.43)$$

The first task, to replace the erroneous value of the velocity change ($d\beta$) by a correct value ($a_{\text{comoving}} d\tau$), is now made part of the problem along with the evaluation of the Thomas precession itself.

Principles settled, the calculation proceeds by inserting the appropriate expressions for all four factors in (41.43), and evaluating both sides of the equation to the first order of small quantities, as follows:

$$1 - (i dt\omega + d\tau a) \cdot \sigma/2 = [\cosh(\alpha'/2) - (\mathbf{n}_\alpha \cdot \sigma) \sinh(\alpha'/2)] \\ \times [\cosh(\alpha/2) + (\mathbf{n}_\alpha \cdot \sigma) \sinh(\alpha/2)]. \quad (41.44)$$

Here α and \mathbf{n}_α are the velocity parameter and unit vector that go with the velocity β ; $\alpha' = \alpha + d\alpha$, and $\mathbf{n}_{\alpha'} = \mathbf{n}_\alpha + d\mathbf{n}_\alpha$, go with $\beta + d\beta$. Develop the righthand side of (41.44) by the methods of calculus, writing $\alpha' = \alpha + d\alpha$ and $\mathbf{n}_{\alpha'} = \mathbf{n}_\alpha + d\mathbf{n}_\alpha$, and applying the rule for the differentiation of a product. Equate coefficients of $-\sigma/2$ and $-i\sigma/2$ on both sides of the equation. Thus find

$$a_{\text{comoving}} d\tau = (d\alpha)\mathbf{n}_\alpha + (\sinh \alpha) d\mathbf{n}_\alpha \quad (41.45)$$

and

$$\omega dt = [2 \sinh^2(\alpha/2)] d\mathbf{n}_\alpha \times \mathbf{n}_\alpha. \quad (41.46)$$

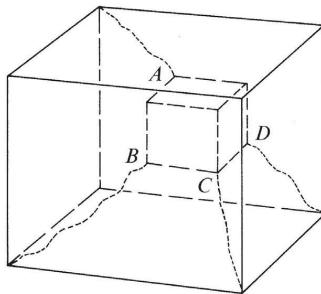
Angular velocity of Thomas precession

The one expression gives the change of velocity as seen in a comoving inertial frame. The other gives the precession as seen in the laboratory frame. For low velocities the expression for the Thomas precession reduces to

$$\omega = \mathbf{a} \times \beta/2. \quad (41.47)$$

Here \mathbf{a} is the acceleration. Only the component perpendicular to the velocity β is relevant for the precession.

For an elementary account of the importance of the Thomas precession in atomic physics, see, for example, Ruark and Urey (1930).

**Figure 41.5.**

“Orientation-entanglement relation” between a cube and the walls of a room. A 360° rotation of the cube entangles the threads. A 720° rotation might be thought to entangle them still more—but instead makes it possible completely to disentangle them.

§41.5. SPINORS

Orientation-entanglement
relation

Paint each face of a cube a different color. Then connect each corner of the cube to the corresponding corner of the room with an elastic thread (Figure 41.5). Now rotate the cube through $2\pi = 360^\circ$. The threads become tangled. Nothing one can do will untangle them. It is impossible for every thread to proceed on its way in a straight line. Now rotate the cube about the same axis by a further 2π . The threads become still more tangled. However, a little work now completely straightens out the tangle (Figure 41.6). Every thread runs as it did in the beginning in a straight line from its corner of the cube to the corresponding corner of the room. More generally, rotations by $0, \pm 4\pi, \pm 8\pi, \dots$, leave the cube in its standard “orientation-entanglement relation” with its surroundings, whereas rotations by $\pm 2\pi, \pm 6\pi, \pm 10\pi, \dots$, restore to the cube only its orientation, not its orientation-entanglement relation with its surroundings. Evidently there is something about the geometry of orientation that is not fully taken into account in the usual concept of orientation; hence the concept of “orientation-entanglement relation” or (briefer term!) “version” (Latin *versor*, turn). Whether there is also a detectable difference in the physics (contact potential between a metallic object and its metallic surroundings, for example) for two inequivalent versions of an object is not known [Aharonov and Susskind (1967)].

In keeping with the distinction between the two inequivalent versions of an object, the spin matrix associated with a rotation,

$$R = \cos(\theta/2) - i(\mathbf{n} \cdot \boldsymbol{\sigma}) \sin(\theta/2), \quad (41.48)$$

reverses sign on a rotation through an odd multiple of 2π . This sign change never shows up in the law of transformation of a vector, as summarized in the formula

$$X \rightarrow X' = RXR^* \quad (41.49)$$

Spinor defined

(two factors R ; sign change in each!). The sign change does show up when one turns from a vector to a 2-component quantity that transforms according to the law

$$\xi \rightarrow \xi' = R\xi. \quad (41.50)$$