Chapter 2
From Minkowski Spacetime to General Relativity

The aim here is to see how one can infer something about a statically supported charge in a gravitational field from the description, entirely within the special theory of relativity, of a uniformly accelerating charge in a spacetime with no gravitational fields. This will be an opportunity to introduce some of the mathematical machinery required to carry out calculations in this area. Section 2.1 describes coordinate systems that might be adopted by accelerating observers in flat spacetimes. Section 2.2 then proves that only when the observer has an eternally uniformly accelerating motion will the metric components in these coordinate systems have a static form. But the key part of the chapter is Sect. 2.3, which discusses different versions of the weak equivalence principle.

Section 2.4 is a detailed review of the weak field approximation usually used to make a connection between general relativity and Newtonian theory as an approximation to it. The aim here is to examine how one might interpret the GR view of a static and homogeneous gravitational field. Finally, Sect. 2.5 gives a critical review of the geodesic principle, which says that a particle left to its own devices in a curved spacetime, i.e., not subject to any external fields, will follow a geodesic. We discuss the demonstration of the latter 'principle' as a theorem of general relativity, and its limitations when the particle in question is not a point particle.

2.1 Semi-Euclidean Coordinate Systems

We begin therefore in Minkowski spacetime. We introduce a little extra generality by considering an accelerating observer AO with arbitrary acceleration along a straight line, specialising to the case of uniform acceleration later. This will allow us to bring out some of the particularities of the special case.

It is well known [11, p. 133] that an observer like AO can find well-adapted coordinates y^μ with the following properties (where the Latin index runs over $\{1,2,3\}$):

- First of all, any curve with all three y^i constant is timelike and any curve with y^0 constant is spacelike.

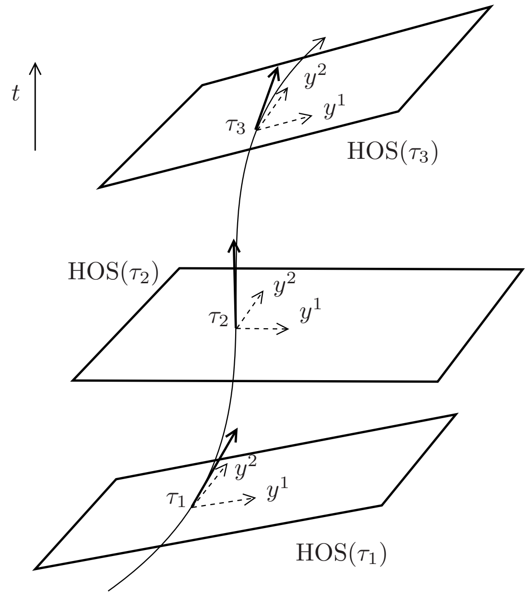

Fig. 2.1 Constructing a semi-Euclidean (SE) frame for an accelerating observer. View from an inertial frame with time coordinate t. The curve is the observer worldline given by (2.1). Three hyperplanes of simultaneity (HOS) are shown at three successive proper times τ_1, τ_2, and τ_3 of the observer. These hyperplanes of simultaneity are borrowed from the instantaneously comoving inertial observer, as are the coordinates y^1, y^2, and y^3 used to coordinatise them. Only two of the latter coordinates can be shown in the spacetime diagram

- At any point along the worldline of AO, the zero coordinate y^0 equals the proper time along that worldline.
- At each point of the worldline of AO, curves with constant y^0 which intersect it are orthogonal to it where they intersect it.
- The metric has the Minkowski form along the worldline of AO.
- The coordinates y^i are Cartesian on every hypersurface of constant y^0.
- The equation for the worldline of AO has the form $y^i = 0$ for $i = 1,2,3$.

Such coordinates could be called semi-Euclidean.

Let us suppose that the worldline of the accelerating observer is given in inertial coordinates in the inertial frame $\mathscr{I}$ by

$$t = \sigma \, , \qquad x = x(\sigma) \, , \qquad \frac{\mathrm{d}x}{\mathrm{d}\sigma} = v(\sigma) \, , \tag{2.1}$$

$$\frac{\mathrm{d}^2 x}{\mathrm{d}\sigma^2} = a(\sigma) \, , \qquad y(\sigma) = 0 = z(\sigma) \, , \tag{2.2}$$

using the time t in $\mathscr{I}$ to parametrise. The proper time $\tau(\sigma)$ of AO is given by

$$\frac{\mathrm{d}\tau}{\mathrm{d}\sigma} = (1 - v^2/c^2)^{1/2} \; . \tag{2.3}$$

The coordinates y^μ are constructed on an open neighbourhood of the AO worldline as follows (see Fig. 2.1). For an event (t,x,y,z) not too far from the worldline, there is a unique value of τ and hence also the parameter σ such that the point lies in the hyperplane of simultaneity (HOS) of AO when its proper time is τ. This hyperplane of simultaneity is given by

$$t - \sigma(\tau) = \frac{v\big(\sigma(\tau)\big)}{c^2}\Big[x - x\big(\sigma(\tau)\big)\Big] \; , \tag{2.4}$$

which solves, for any x and t, to give $\sigma(\tau)(x,t)$.

The semi-Euclidean coordinates attributed to the event (t,x,y,z) are, for the time coordinate y^0, (c times) the proper time τ found from (2.4) and, for the spatial coordinates, the spatial coordinates of this event in an instantaneously comoving inertial frame at proper time τ of AO. In fact, every other event in this instantaneously comoving inertial frame is attributed the same time coordinate $y^0 = c\tau$ and the appropriate spatial coordinates borrowed from this frame. Of course, the HOS of AO at time τ is also the one borrowed from the instantaneously comoving inertial frame.

There is just one detail to get out of the way: there are many different instantaneously comoving inertial frames for a given τ, and there are even many different ways to choose these frames as a smooth function of τ as one moves along the AO worldline, rotating back and forth around various axes in the original inertial frame $\mathscr{I}$ as τ progresses. We choose a sequence with no rotation about any space axis in the local rest frame. It can always be done by solving the Fermi–Walker transport equations. The semi-Euclidean coordinates are then given by

$$\begin{cases} y^0 = c\tau \; , \\[2mm] y^1 = \dfrac{[x - x(\sigma)] - v(\sigma)(t-\sigma)}{\sqrt{1 - v^2/c^2}} \; , \\[2mm] y^2 = y \; , \\[2mm] y^3 = z \; , \end{cases} \tag{2.5}$$

where $\sigma = \sigma(t,x)$ is found from (2.4). The inverse transformation, from semi-Euclidean coordinates to inertial coordinates, is given by

$$
\begin{cases}
t = \sigma(y^0) + \dfrac{v(y^0)}{c^2} y^1 \left[1 - \dfrac{v(y^0)^2}{c^2}\right]^{-1/2} , \\[2ex]
x = x(y^0) + y^1 \left[1 - \dfrac{v(y^0)^2}{c^2}\right]^{-1/2} , \\[2ex]
y = y^2 , \\[1ex]
z = y^3 ,
\end{cases}
\tag{2.6}
$$

where the function $\sigma(y^0)$ is just the expression relating inertial time to proper time for the accelerating observer, and the functions $x(y^0)$ and $v(y^0)$ should really be written $x\big(\sigma(y^0)\big)$ and $v\big(\sigma(y^0)\big)$, respectively.

The above relations are not very enlightening. They are only displayed to show that the idea of such coordinates can be made perfectly concrete. One calculates the metric components in this frame, viz.,

$$
g_{00} = 1/g^{00} = \left[1 + \frac{a(\sigma)[x - x(\sigma)]/c^2}{1 - v(\sigma)^2/c^2}\right]^2 , \tag{2.7}
$$

where $\sigma = \sigma(t,x)$ as found from (2.4), and

$$
g_{i0} = 0 = g_{0i} , \qquad g_{ij} = -\delta_{ij} , \qquad i,j \in \{1,2,3\} , \tag{2.8}
$$

and checks the list of requirements for the coordinates to be suitably adapted to the accelerating observer.

Although perfectly concrete, the coordinates are not perfectly explicit: the component g_{00} of the semi-Euclidean metric has been expressed in terms of the original inertial coordinates! To obtain explicit formulas, one needs to consider a specific motion $x(\sigma)$ of AO, the classic example being uniform acceleration:

$$
x(\sigma) = \frac{c^2}{g}\left[\left(1 + \frac{g^2\sigma^2}{c^2}\right)^{1/2} - 1\right] , \qquad t = \sigma , \tag{2.9}
$$

where g is some constant with units of acceleration. This does not look like a constant acceleration in the inertial frame:

$$
\frac{\mathrm{d}x}{\mathrm{d}\sigma} = \frac{g\sigma}{(1 + g^2\sigma^2/c^2)^{1/2}} , \qquad \frac{\mathrm{d}^2x}{\mathrm{d}\sigma^2} = \frac{g}{(1 + g^2\sigma^2/c^2)^{3/2}} . \tag{2.10}
$$

However, the 4-acceleration defined in the inertial frame $\mathscr{I}$ by

$$
a^\mu = \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} , \tag{2.11}
$$

where τ is the proper time, has constant magnitude. It turns out that

$$a^2 := a_\mu a^\mu = -g^2 ,$$

with a suitable convention for the signature of the metric.

In this case, the transformation from inertial to semi-Euclidean coordinates is

$$y^0 = \frac{c^2}{g} \tanh^{-1} \frac{ct}{x + c^2/g} , \tag{2.12}$$

$$y^1 = \left[\left(x + \frac{c^2}{g} \right)^2 - c^2 t^2 \right]^{1/2} - \frac{c^2}{g} , \qquad y^2 = y , \qquad y^3 = z , \tag{2.13}$$

and the inverse transformation is

$$t = \frac{c}{g} \sinh \frac{gy^0}{c^2} + \frac{y^1}{c} \sinh \frac{gy^0}{c^2} , \tag{2.14}$$

$$x = \frac{c^2}{g} \left(\cosh \frac{gy^0}{c^2} - 1 \right) + y^1 \cosh \frac{gy^0}{c^2} , \qquad y = y^2 , \qquad z = y^3 . \tag{2.15}$$

One finds the metric components to be

$$g_{00} = \left(1 + \frac{gy^1}{c^2} \right)^2 , \qquad g_{0i} = 0 = g_{i0} , \qquad g_{ij} = -\delta_{ij} , \tag{2.16}$$

for $i, j \in \{1, 2, 3\}$, in the semi-Euclidean frame. Interestingly, this metric is static, i.e., g_{00} is independent of y^0. It is the only semi-Euclidean metric that is, and this is one of the theoretical particularities of uniform acceleration.

Before proving this, it is worth pausing to wonder why AO should adopt semi-Euclidean coordinates. It must be comforting to attribute one's own proper time to events that appear simultaneous. But what events are simultaneous with AO? In the above construction, AO borrows the hyperplane of simultaneity of an inertially moving observer, who does not have the same motion at all. AO also borrows the lengths of this inertially moving observer. It is not obvious that an accelerating observer would set up such coordinates, or indeed that they have any significance whatever! A well-informed observer, who knew the theory of relativity, might well prefer to use the globally Minkowskian frame that happens to be available here, even though it is not adapted to her motion.

2.2 The SE Metric for Uniform Acceleration Is the Only Static SE Metric

We shall now show that a semi-Euclidean frame is static iff the acceleration of the observer is uniform in the sense of having constant pseudolength. We already have this in one direction: we have shown that the semi-Euclidean frame is static if the 4-acceleration is uniform. We now examine the metric for a general acceleration in one dimension and show that the condition for it to be static implies that the acceleration is uniform.

For this exercise, we shall need the partial derivatives

$$c\frac{\partial x}{\partial y^0} = \frac{v(\sigma)}{(1-v^2/c^2)^{1/2}} + \frac{y^1 a(\sigma) v(\sigma)/c^2}{(1-v^2/c^2)^2} , \tag{2.17}$$

$$\frac{\partial x}{\partial y^1} = \left[1 - \frac{v(\sigma)^2}{c^2}\right]^{-1/2} , \tag{2.18}$$

$$c\frac{\partial t}{\partial y^0} = \frac{1}{(1-v^2/c^2)^{1/2}} + \frac{a(\sigma) y^1}{c^2(1-v^2/c^2)^2} , \tag{2.19}$$

and

$$\frac{\partial t}{\partial y^1} = \frac{v(\sigma)}{c^2}\left[1 - \frac{v(\sigma)^2}{c^2}\right]^{-1/2} , \tag{2.20}$$

where $\sigma = \sigma(y^0)$, giving expressions entirely in terms of semi-Euclidean coordinates. These are obtained by differentiating (2.6) and using

$$\frac{\mathrm{d}\sigma(y^0)}{\mathrm{d}y^0} = \frac{1}{c(1-v^2/c^2)^{1/2}} . \tag{2.21}$$

Recall from (2.7) that the metric components are

$$g_{00} = 1/g^{00} = \left[1 + \frac{a(\sigma)[x - x(\sigma)]/c^2}{1-v^2/c^2}\right]^2 , \tag{2.22}$$

$$g_{01} = g_{10} = 0 , \qquad g_{11} = -1 , \tag{2.23}$$

with the rest also diagonal, and in the Minkowski way. So the semi-Euclidean metric is static iff

$$\frac{\partial}{\partial y^0}\left\{\frac{a(\sigma)[x - x(\sigma)]/c^2}{1-v^2/c^2}\right\} = 0 . \tag{2.24}$$

Note that x is one of the inertial coordinates, while σ refers to $\sigma(\tau)(x,t)$, the solution of

$$t-\sigma(\tau)=\frac{v\big(\sigma(\tau)\big)}{c^2}\big[x-x\big(\sigma(\tau)\big)\big]\,, \tag{2.25}$$

giving the HOS through the field point (x,t). Note, for example, that $x\big(\sigma(\tau)\big)$ is not just x.

The aim now is to reexpress (2.24) as a condition on the trajectory $x(\sigma)$ of the observer, and show that the corresponding worldline is uniformly accelerated. To begin with,

$$\frac{\partial}{\partial y^0}\left\{\frac{a(\sigma)[x-x(\sigma)]}{c^2-v^2}\right\}=\frac{\mathrm{d}}{\mathrm{d}\sigma}\left\{\frac{a(\sigma)[x-x(\sigma)]}{c^2-v^2}\right\}\frac{\partial\sigma}{\partial y^0}+\frac{a(\sigma)}{c^2-v^2}\frac{\partial x}{\partial y^0}\,. \tag{2.26}$$

Now we have both $\partial x/\partial y^0$ and $\partial t/\partial y^0$ from (2.17) and (2.19). Note that, by (2.6),

$$y^1=[x-x(\sigma)](1-v^2/c^2)^{1/2}\,, \tag{2.27}$$

whence

$$\begin{aligned}\frac{\partial x}{\partial y^0}&=\frac{v(\sigma)}{(c^2-v^2)^{1/2}}+\frac{a(\sigma)v(\sigma)}{c^3(1-v^2/c^2)^2}[x-x(\sigma)](1-v^2/c^2)^{1/2}\\&=\frac{v(\sigma)}{(c^2-v^2)^{1/2}}\left\{1+\frac{a(\sigma)[x-x(\sigma)]}{c^2-v^2}\right\}\,.\end{aligned} \tag{2.28}$$

Furthermore, from the fact that $\sigma(x,t)=\sigma(y^0)$ and (2.21), we know that, when we put $x(y^0,y^1)$ and $t(y^0,y^1)$ into $\sigma(x,t)$, we get $\sigma(y^0)$, the function giving the Minkowski time corresponding to the proper time y^0. The latter has no dependence on y^1, and (2.21) gives

$$\frac{\mathrm{d}\sigma(y^0)}{\mathrm{d}y^0}=\frac{1}{(c^2-v^2)^{1/2}}\,. \tag{2.29}$$

We can now proceed from (2.26), noting that

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\left\{\frac{a(\sigma)[x-x(\sigma)]}{c^2-v^2}\right\}=\frac{a'(\sigma)[x-x(\sigma)]}{c^2-v^2}-\frac{a(\sigma)v(\sigma)}{c^2-v^2}+\frac{2a^2v[x-x(\sigma)]}{(c^2-v^2)^2}\,, \tag{2.30}$$

whence (2.24) is equivalent to

$$\left\{\frac{a'(\sigma)[x-x(\sigma)]}{c^2-v^2}-\frac{a(\sigma)v(\sigma)}{c^2-v^2}+\frac{2a^2v[x-x(\sigma)]}{(c^2-v^2)^2}\right\}\frac{1}{(c^2-v^2)^{1/2}}$$
$$+\frac{a(\sigma)}{c^2-v^2}\left\{1+\frac{a(\sigma)[x-x(\sigma)]}{c^2-v^2}\right\}\frac{v(\sigma)}{(c^2-v^2)^{1/2}}=0\,.$$

There is now some cancellation. We find that

$$\left[a'(\sigma)+\frac{3a(\sigma)^2 v(\sigma)}{c^2-v^2}\right][x-x(\sigma)]=0\,, \tag{2.31}$$

and since we are not allowing $x=x(\sigma)$ as an interesting solution, we conclude that

$$\text{metric static} \quad \Longleftrightarrow \quad a'(\sigma)+\frac{3a(\sigma)^2 v(\sigma)}{c^2-v^2}=0\,. \tag{2.32}$$

The next step is to look at the solutions to this equation.

As an aside, it is interesting to rederive (2.29) by this approach. We begin with the relation (2.25) that defines $\sigma(\tau)(x,t)$, thinking of it when x and t are written in their turn as functions of y^0 and y^1. Differentiating both sides with respect to y^0, we obtain

$$\frac{\partial t}{\partial y^0}-\frac{\partial \sigma}{\partial y^0}=\frac{a}{c^2}\frac{\partial \sigma}{\partial y^0}\big[x-x\big(\sigma(\tau)\big)\big]+\frac{v\big(\sigma(\tau)\big)}{c^2}\left[\frac{\partial x}{\partial y^0}-v\big(\sigma(\tau)\big)\frac{\partial \sigma}{\partial y^0}\right]\,. \tag{2.33}$$

We solve this for $\partial\sigma/\partial y^0$, obtaining

$$\frac{\partial \sigma}{\partial y^0}=\frac{\dfrac{\partial t}{\partial y^0}-\dfrac{v(\sigma)}{c^2}\dfrac{\partial x}{\partial y^0}}{1-\dfrac{v^2}{c^2}+\dfrac{a[x-x(\sigma)]}{c^2}}\,. \tag{2.34}$$

Likewise, by the same algebra,

$$\frac{\partial \sigma}{\partial y^1}=\frac{\dfrac{\partial t}{\partial y^1}-\dfrac{v(\sigma)}{c^2}\dfrac{\partial x}{\partial y^1}}{1-\dfrac{v^2}{c^2}+\dfrac{a[x-x(\sigma)]}{c^2}}\,. \tag{2.35}$$

Substituting (2.17) and (2.19) into (2.34) and doing a little algebra, we do indeed find the result (2.29) above, written here with the partial derivative, viz.,

$$\frac{\partial \sigma(y^0)}{\partial y^0}=\frac{1}{(c^2-v^2)^{1/2}}\,. \tag{2.36}$$

But even more convincingly, if we substitute (2.18) and (2.20) of p. 10, viz.,

$$\frac{\partial x}{\partial y^1}=\left[1-\frac{v(\sigma)^2}{c^2}\right]^{-1/2}\,, \tag{2.37}$$

and

$$\frac{\partial t}{\partial y^1} = \frac{v(\sigma)}{c^2}\left[1 - \frac{v(\sigma)^2}{c^2}\right]^{-1/2} , \tag{2.38}$$

into (2.35), we obtain zero! This is just because

$$\frac{\partial t}{\partial y^1} = \frac{v(\sigma)}{c^2}\frac{\partial x}{\partial y^1} .$$

Solving the Equation for a Static Metric

We now return to (2.32) and consider solutions. It is instructive to begin by checking that the uniform acceleration does solve the equation, since the calculation is not quite as simple as we might expect. We have the formulas for such an acceleration on p. 8, viz.,

$$v(\sigma) = \frac{g\sigma}{(1+g^2\sigma^2/c^2)^{1/2}} , \qquad a(\sigma) = \frac{g}{(1+g^2\sigma^2/c^2)^{3/2}} , \tag{2.39}$$

whence also

$$a'(\sigma) = -\frac{3g^3\sigma/c^2}{(1+g^2\sigma^2/c^2)^{5/2}} . \tag{2.40}$$

Then

$$\begin{aligned}\frac{3a(\sigma)^2 v(\sigma)}{c^2 - v^2} &= \frac{3g^2}{(1+g^2\sigma^2/c^2)^3}\frac{g\sigma}{(1+g^2\sigma^2/c^2)^{1/2}}\frac{1+g^2\sigma^2/c^2}{c^2}\\ &= \frac{3g^3\sigma/c^2}{(1+g^2\sigma^2/c^2)^{5/2}} .\end{aligned}$$

Equation (2.32) is therefore satisfied, which we knew it should be, and that is encouraging.

But are there other solutions? The last calculation can guide us intuitively. We observe that

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\left\{\frac{a(\sigma)}{[1-v(\sigma)^2/c^2]^{3/2}}\right\} = \frac{1}{[1-v(\sigma)^2/c^2]^{3/2}}\left[a'(\sigma) + \frac{3a(\sigma)^2 v(\sigma)}{c^2 - v^2}\right] ,$$

so that

$$\text{metric static} \quad \Longleftrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}\sigma}\left\{\frac{a(\sigma)}{[1-v(\sigma)^2/c^2]^{3/2}}\right\} = 0 . \tag{2.41}$$

Put another way,

$$\text{metric static} \quad \Longleftrightarrow \quad \frac{a(\sigma)}{[1-v(\sigma)^2/c^2]^{3/2}} = \kappa \,, \tag{2.42}$$

for some constant κ. Once again, we can check that this is satisfied by the uniform acceleration. We note that

$$1 - \frac{v(\sigma)^2}{c^2} = 1 - \frac{g^2\sigma^2/c^2}{1+g^2\sigma^2/c^2} = \frac{1}{1+g^2\sigma^2/c^2} \,,$$

so that the above uniform acceleration satisfies the condition iff $\kappa = g$.

We can completely solve the above differential equation, which is rather surprising in itself. We have

$$\frac{\mathrm{d}v}{[1-v(\sigma)^2/c^2]^{3/2}} = \kappa \mathrm{d}\sigma \,.$$

The left-hand side is integrated by first substituting $v = c\sin\theta$, so that

$$\mathrm{d}v = c\cos\theta\,\mathrm{d}\theta \,,$$

and

$$\begin{aligned}
\int \frac{\mathrm{d}v}{[1-v(\sigma)^2/c^2]^{3/2}} &= c\int \frac{\mathrm{d}\theta}{\cos^2\theta} \\
&= c\int (1+\tan^2\theta)\,\mathrm{d}\theta \\
&= c\tan\theta \,.
\end{aligned}$$

Of course,

$$\tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{v/c}{(1-v^2/c^2)^{1/2}} \,,$$

and the final, most general solution is

$$\kappa\sigma + \kappa' = \frac{v}{(1-v^2/c^2)^{1/2}} \,, \tag{2.43}$$

where κ and κ' are constants. We can rearrange this to obtain $v(\sigma)$. After a little manipulation,

$$v(\sigma) = \frac{\kappa\sigma + \kappa'}{\left[1+(\kappa\sigma+\kappa')^2/c^2\right]^{1/2}} \,. \tag{2.44}$$

This is to be compared with the uniform acceleration

$$v(\sigma) = \frac{g\sigma}{(1+g^2\sigma^2/c^2)^{1/2}} \,,$$

which is the solution when $\kappa = g$ and $\kappa' = 0$. Since κ' in (2.44) is just a readjustment of the zero of the variable σ, we see that all solutions of the static metric equation are in fact uniform accelerations, which is what had to be proved.

2.3 The Step to General Relativity

So what does all this tell us about gravitational fields? Suppose we are doing general relativity and we are given a spacetime with coordinates $\{y^\mu\}$ and interval

$$\mathrm{d}s^2 = \left(1+\frac{gy^1}{c^2}\right)^2 (\mathrm{d}y^0)^2 - (\mathrm{d}y^1)^2 - (\mathrm{d}y^2)^2 - (\mathrm{d}y^3)^2 \,, \tag{2.45}$$

where g is a constant. We might begin by calculating the connection coefficients. They turn out to be fairly simple functions of the coordinates. In fact using the usual formula

$$\Gamma^\mu_{\sigma\tau} = \frac{1}{2} g^{\mu\nu}\left(g_{\nu\sigma,\tau} + g_{\nu\tau,\sigma} - g_{\sigma\tau,\nu}\right) \tag{2.46}$$

for the Levi-Civita connection coefficients, where commas denote ordinary coordinate derivatives, we find that all the coefficients are zero except

$$\Gamma^1_{00} = \frac{g}{c^2}\left(1+\frac{gy^1}{c^2}\right) \,, \qquad \Gamma^0_{01} = \Gamma^0_{10} = \frac{g}{c^2}\left(1+\frac{gy^1}{c^2}\right)^{-1} \,. \tag{2.47}$$

We would then calculate the Riemann curvature tensor, finding it to be identically zero. This would immediately tell us that we were dealing with some non-inertial coordinatisation of a flat spacetime, in which there were no tidal effects.

One ought to point out a somewhat delicate matter regarding the above interval, which comes from the metric (2.16): when $y^1 = -c^2/g$, the metric is degenerate because $g_{00} = 0$, and g^{00} is not defined. However, one would soon observe that this was due to the choice of coordinates, rather like the singularity at the event horizon in the Schwarzschild metric. Such singularities can be removed by a better choice of coordinates. In the present case, one would transform by (2.14) and (2.15) and the metric would reduce to the Minkowski form η_{ij} of (1.1).

According to the usual rules of general relativity, one would then argue as follows: the coordinates (t,x,y,z) are those that would be set up in a freely falling frame, in which the connection coefficients are zero and freely falling bodies not subjected to any forces follow straight lines. There are two things to note:

- In a general, curved spacetime, one can always find such coordinates locally in the neighbourhood of any given event. This is an application of what is sometimes called the weak equivalence principle WEP and it is inherent in the manifold formulation of general relativity when one adopts the Levi-Civita connection (zero torsion). In that case, for any event (point) P in spacetime, there is always a choice of coordinates in some neighbourhood of that event for which the connection coefficients are zero at P and the metric takes the Minkowski form at P. By continuity, the connection remains close to zero and the metric close to η over some small enough neighbourhood of P.

- In general relativity, one adopts the idea that free fall is not due to a force, i.e., gravity is no longer a force in this view. This is a linguistic adjustment made to accord with the idea that forces and accelerations are intimately linked. In general relativity, freely falling bodies follow geodesics (with some provisos, as explained in Sect. 2.5), and the geodesic equation says precisely that their four-acceleration is zero.

We have already encountered our first equivalence principle here, and it is so deeply embedded in the theory that it would be disastrous if the theory itself contradicted it. Fortunately, this does not happen, at least, not on the basis of the discussion in [2]. For one thing, the equivalence principle cited in the quotation from Bondi and Gold on p. 3 refers to electromagnetic radiation fields, whereas the weak equivalence principle WEP mentioned here does not yet have anything to say about electromagnetism.

Note also that, since the proof concerning the existence of this convenient form of the metric and connection coefficients only requires one to assume that the connection is the usual Levi-Civita connection entirely determined by the metric, and this follows if one assumes zero torsion, i.e., that the connection coefficients are symmetric in their two lower indices, one might say that this equivalence principle hardly looks like a principle at all: the whole structure of the theory as it is usually formulated implies the weak principle of equivalence. On the other hand, one might also say that it is precisely the prerogative of a principle to be inescapable.

So what would the general relativist deduce from the discovery of a global inertial frame? There would appear to be several possibilities, all of which are equivalent as far as general relativity is concerned. One might say that one was in a perfectly flat region of spacetime far from any source of curvature (gravity), and that the coordinates $\{y^\mu\}$ were simply those adopted by a uniformly accelerating observer clever enough to set up a system so well adapted to her motion. This is what we were considering before. But one might also say that the coordinates (t,x,y,z) are those of an observer freely falling in a static and homogeneous gravitational field (SHGF). The fact that there are no tidal effects, i.e., the curvature is zero, is what inspires us to say that the field is static and homogeneous.

Presumably, in the second case, if the observer looked around, she would find some source, i.e., some distribution of energy (e.g., mass energy) to which one could attribute the presence of this SHGF. It is hard to imagine what it might be and we shall discuss this below (see Sect. 2.4). However, a practically-minded observer with knowledge of physics would probably just say that, in reality, the curvature is not quite zero but that this is a good approximation over some region of spacetime, of the kind usually made in Earth-based laboratories, given the accuracy with which measurements can be made.

But the Earth-based observer would presumably have a considerable advantage over one who was simply presented with the interval (2.45) and could not see the source. In the laboratory, one can measure the acceleration relative to the floor, for example, assuming that it is at rest in some sense (or can be treated as being at rest) relative to the gravitational source. At some risk of confusion, let us call this the acceleration due to gravity, but bearing in mind that there is no acceleration due to gravity in general relativity (there is only acceleration when one is not allowed to free-fall). Without sight of the gravitational source or other reference point, our relativist in possession of the above interval could not even determine this acceleration due to gravity. Looking at the expression in (2.45), one might want to say that it was equal to g, because the interval does indeed single out this value. But in a certain sense, the value g is just an artefact of the choice of coordinates.

One should have no doubt about this. Starting with the Minkowski frame, one could have chosen any uniformly accelerating observer, with value g' say, and obtained new coordinates $\{y'^{\mu}\}$ such that the interval takes the form

$$\mathrm{d}s^2 = \left(1 + \frac{g'y'^1}{c^2}\right)^2 (\mathrm{d}y'^0)^2 - (\mathrm{d}y'^1)^2 - (\mathrm{d}y'^2)^2 - (\mathrm{d}y'^3)^2 \ . \tag{2.48}$$

The fact is, of course, that in this specific context, one cannot say how much of the acceleration is due to gravity and how much is due to some other effect, e.g., the kind of accelerating effect we were presumably imagining at the outset when we considered a uniformly accelerating body in a flat spacetime (with no gravity).

One might appeal to the well known weak field approximation in which the theory begins to look like Newton's theory. Could this give us some reason to think that g rather than g' was the acceleration due to gravity in a given context? We shall discuss this in some detail in Sect. 2.4, to bring out another rather subtle point, but let us just see globally what is involved. One writes the metric in the form

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

where $h_{\mu\nu}$ is assumed to be small. A range of other assumptions is usually made about the metric components and allowed coordinate transformations which maintain this range of assumptions, but there is no need to go into too much detail here. In the case of (2.45), one reads off

$$h_{00} \approx 2gy^1/c^2 \ , \tag{2.49}$$

in a region where $gy^1/c^2 \ll 1$. One then considers the geodesic equation giving the worldline of a freely falling object relative to these coordinates and concludes that, if the motion is indeed entirely due to a gravitational potential Φ, then one must have

$$h_{00} = 2\Phi/c^2 \ . \tag{2.50}$$

This would give

$$\Phi \approx gy^1 \ . \tag{2.51}$$

In this Newtonian view, the potential causes a force per unit mass of

$$(-\Phi_{,1}, -\Phi_{,2}, -\Phi_{,3}) = (-g, 0, 0) , \tag{2.52}$$

so that we are indeed explaining the quantity g appearing in the metric by the effects of gravity, on this interpretation.

There is no surprise here, but of course the metric components (2.48) would give us a gravitational potential Φ' such that

$$(-\Phi'_{,1}, -\Phi'_{,2}, -\Phi'_{,3}) = (-g', 0, 0) , \tag{2.53}$$

if we could assume that the motion were indeed entirely due to a gravitational potential Φ'. And this is precisely what we do not know: how much of g or g' comes from a gravitational effect?

This highlights a fundamental point about general relativity which might be stated as a form of the equivalence principle WEP2: from local observations alone, one cannot tell whether the fields one observes via the spacetime metric are due to one's own motion or a gravitational potential, or both, and in the latter case, in what proportions. The word 'local' is essential here, for two reasons:

- If one could look around, one would presumably see any sources of gravitational potential operating in the neighbourhood. One could then apply Einstein's equations to try to understand the metric and one could perhaps try to measure the 'acceleration due to gravity' by measuring a rate of change of proper distance to the gravitational source for objects in free fall.
- An acceleration of the observer which changes the metric components via an appropriate choice of the observer's coordinates cannot introduce curvature, so if there is curvature, one can attribute it to a gravitational potential. However, one must be able to survey a large enough region of spacetime to be able to determine the curvature, if there is any.

Considering the second point, it might seem that data from any neighbourhood of a given event, no matter how small, would suffice to establish the curvature if there were any. The reason why a certain size of neighbourhood is required is just that one is always restricted by measurement accuracy. For any given level of measurement accuracy, supposing that one has arranged by coordinate choice for the metric to be Minkowskian and the connection components zero at some given point, there will be some neighbourhood of the event in which the metric appears to be Minkowskian and the connection components zero. More about this in a moment.

Referring to the first of the above points, if one has the distribution of energy and momentum relative to some coordinate system in the spacetime, one then knows the metric, connection, and curvature relative to those coordinates by solving Einstein's equations (at least insofar as they and the relevant boundary conditions are determinate). On this basis one would then like to say which aspects of a motion were due to gravity and which were an artefact of the coordinates. For example, within the context of the weak field approximation, one can consider the effects of a massive point particle, with energy–momentum distribution

$$T^{00} = M\delta(\mathbf{x}) \,, \tag{2.54}$$

where M is its mass and δ is a point-support distribution on the spacelike hypersurface with coordinates $\mathbf{x}$, whence Einstein's equations reduce to something like

$$h_{00} = -\frac{2MG}{|\mathbf{x}|} \,,$$

and one deduces the form of Φ, viz.,

$$\Phi = -\frac{GM}{|\mathbf{x}|} \,.$$

This is discussed more carefully in Sect. 2.4. The present scenario would appear to escape somewhat from the kind of constraint imposed by Einstein's equations. For one thing, it is not obvious that any energy–momentum distribution would lead via Einstein's equations to the static, homogeneous gravitational field proposed to explain the intervals (2.45) or (2.48).

Of course, as mentioned above, one can easily imagine an approximate context, and this context is precisely the one in which the neighbourhood over which one can carry out measurements of the gravitational potential as embodied by the metric is too small to allow one to detect the curvature. This is what allows the above version WEP2 of the weak equivalence principle: if one is restricted to a small enough neighbourhood, where 'small enough' is dictated by one's measurement accuracy, then one is forced into the situation of a spacetime that could be attributed the intervals (2.45) or (2.48).

This follows from the weak equivalence principle WEP first mentioned on p. 15. At a given event in spacetime, one arranges for the metric to have Minkowski form and the Levi-Civita connection components to be zero. But the latter are simply related to the coordinate derivatives of the metric components and the fact that they are zero at the given event means that the coordinate derivatives of the metric components are also zero there. This in turn means that the metric components relative to these coordinates will change very slowly from the Minkowski form in any small neighbourhood of the event. However, it should be noted that the coordinate derivatives of the connection components, which depend on the curvature components, are not generally zero, so the connection components may be changing rapidly from zero over any neighbourhood of the event.

Since the above statement of WEP2 refers to an acceleration of the observer which changes the metric components via an appropriate choice of the observer's coordinates, one ought to say more about what might be an appropriate choice of the observer's coordinates. We already encountered this question at the beginning of this chapter for an observer accelerating in a flat spacetime. With coordinates satisfying the conditions listed on p. 5, the equation for a geodesic that happens to intersect the observer worldline at some event, i.e., describing a free-falling object encountered by the observer, takes the form [11, p. 135].

$$\frac{\mathrm{d}^2y^i}{\mathrm{d}y^{0\,2}} + \underbrace{a^i}_{\substack{\text{inertial}\\ \text{force}}} + \underbrace{2\Omega^i{}_j \frac{\mathrm{d}y^j}{\mathrm{d}y^0}}_{\substack{\text{Coriolis}\\ \text{force}}} - \underbrace{a^j \frac{\mathrm{d}y^j}{\mathrm{d}y^0} \frac{\mathrm{d}y^i}{\mathrm{d}y^0}}_{\substack{\text{relativistic}\\ \text{correction}}} = 0\,, \tag{2.55}$$

with suitable interpretation of the terms, where a^i and $\Omega^i{}_j$ are determined by the connection coefficients according to

$$a^i := \Gamma^i_{00}\,, \qquad \Gamma^0_{0i} = \Gamma^0_{i0} = a^i \quad (i = 1,2,3)\,, \tag{2.56}$$

$$\Omega^i{}_j := \Gamma^i_{j0} = \Gamma^i_{0j} = -\Gamma^j_{0i} \quad (i,j = 1,2,3)\,. \tag{2.57}$$

Furthermore, one can arrange for the antisymmetric rotation matrix $\Omega^i{}_j$ to be zero, viz., $\Omega^i{}_j = 0$, for $i,j = 1,2,3$, by suitable choice of coordinates (basically, using Fermi–Walker transport). It is important to note that (2.55) holds exactly only on the observer worldline, although it will hold approximately nearby.

One can do something very similar even in a curved spacetime [11, p. 181]. An observer with quite arbitrary motion can find well-adapted coordinates y^μ with the following properties (where the Latin index runs over $\{1,2,3\}$):

- First of all, any curve with all three y^i constant is timelike and any curve with y^0 constant is spacelike.
- At any point along the observer worldline, the zero coordinate y^0 equals the proper time along that worldline.
- At each point of the observer worldline, curves with constant y^0 which intersect it are orthogonal to it where they intersect it.
- The metric has the Minkowski form along the observer worldline.
- Any purely spatial geodesic through the observer worldline on a hypersurface $y^0 = \text{constant}$ satisfies

$$\frac{\mathrm{d}^2y^i}{\mathrm{d}s^2} = 0\,, \tag{2.58}$$

  where s is an affine parameter for the geodesic.
- The equation for the observer worldline has the form $y^i = 0$ for $i = 1,2,3$.

Such coordinates are said to be normal. Note the difference with the flat spacetime case: the coordinates y^i are not necessarily Cartesian on every hypersurface of constant y^0 as they were on p. 5, but we have $\mathrm{d}^2y^i/\mathrm{d}s^2 = 0$ for any purely spatial geodesic through the observer worldline on a hypersurface $y^0 = \text{constant}$. Put another way, in a semi-Euclidean frame (for flat spacetime), the fifth condition (2.58) is true for any purely spatial geodesic, not just for spatial geodesics through the observer worldline.

Carrying out the above construction, one obtains precisely the same relations (2.55)–(2.57) for the worldline of any freely falling object where it encounters the observer worldline.

In the last two examples, the observer chose coordinates in such a way that she always remained at the origin of the spatial hypersurfaces given by keeping her time coordinate constant. In this situation, even though the connection components associated with the rotation matrix $\Omega^i{}_j$ can be made equal to zero, those related to a^i cannot. So with this appropriate choice of coordinates, said to be adapted to the observer's worldline, one cannot make the connection components equal to zero on the observer worldline, unless the latter happens to be a geodesic, i.e., unless the observer happens to be in free fall.

Of course, as mentioned before, the observer can chose any coordinates she likes in general relativity. In a flat spacetime, the best choice may well be some set of globally inertial coordinates, and in a curved spacetime, some set of locally inertial coordinates, even though the observer then has motion relative to those coordinates. But there is a reason for looking at the worldline of a freely falling object relative to the kind of coordinates an observer might consider to be adapted to her own worldline, and this is that it brings out the idea that an object can appear to be accelerating even when it is not, in particular, when it is the observer that is accelerating rather than the observed object. Note how the components a^i of the observer 4-acceleration become interpreted as the components of the inertial acceleration of the test particle relative to such coordinates.

This leads to the idea that one cannot then say how much of the motion of an object relative to one's coordinates can be attributed to gravitational effects and how much to one's choice of coordinates. Let us refer to this as WEP3, another contender for the post of equivalence principle. One is familiar with this idea from the Newtonian theory of gravity in the pre-relativistic context: because inertial mass and passive gravitational mass are equal (an experimental result, always open to further testing as techniques improve), an observer cannot tell whether the acceleration of a test particle relative to her Euclidean coordinate system is due to some gravitational field or due to her own acceleration and the consequent acceleration of her comoving Euclidean coordinate system. (The acceleration of that Euclidean system must not be rotational here.) In pre-relativistic theory, one has no difficulty setting up Euclidean coordinate systems in spatial hypersurfaces and time is the same for everyone. What becomes of WEP3 in GR?

The Newtonian idea of a test particle moving only due to gravitational forces (a motion called free fall) is replaced in GR by the principle that such a test particle will follow a timelike geodesic with the proper time of the particle as affine parameter. In fact this 'principle' can be deduced from Einstein's equations, a point which deserves its own short section at the end of this chapter. This so-called geodesic principle immediately builds in the idea that the purely gravitational motions of any two objects are going to be the same if they start at the same spacetime event and with the same 4-velocity, regardless of their inertial masses, or indeed any detail of their internal physical makeup. Their motions depend only on the geometrical structure of the spacetime as given by the metric (which itself determines the connection and curvature), and of course initial conditions. We shall see in Sect. 2.5 that this is only an approximation, a point to be confirmed in Chap. 12. Anyway, such motion is described as inertial, or non-accelerational, and herein lies a difficulty for the pre-relativistic equivalence principle mentioned in the last paragraph.

The problem is this. The pre-relativistic equivalence principle can compare a test particle that is freely falling under the effect of a gravitational force (and gravity is a force in this theory) with one that has the special property of moving inertially (and appears to accelerate because the observer is accelerating relative to the special inertial frames that exist in this theory). In this theory one does of course have a difficulty in saying why some frames should be inertial while others are not, but this is another matter. The point is that, apart from the freely falling locally inertial frames of GR, there are no other, distinct inertial frames. Put another way, if one were to say in GR that one could not distinguish the motion of a test particle subject only to gravitational effects from one with 'inertial acceleration', one would be committing tautology: the test particle with 'inertial acceleration' would presumably be one subject only to gravitational effects, i.e., with no acceleration, but appearing to accelerate as in equations like (2.55). One would be talking about the very same test particle on each side of the comparison.

So is there nothing left of WEP3 in general relativity? Could an accelerating observer (on a non-geodesic worldline) not at least mistake the motion of a freely falling test object as an acceleration by setting up the coordinates described above and obtaining an equation of motion like (2.55) for the worldline of the test object relative to her coordinates? If such were the case, one could then say that acceleration of the observer and 'gravitational acceleration' of a test particle were indistinguishable. This certainly looks to be the case. On the other hand, we have to ask whether the observer would not realise in some direct way that she was accelerating. For example, if she knew the metric, connection and curvature components relative to her adapted coordinates, and it suffices to know only the metric components over some neighbourhood in order to achieve this, would she not realise that she was accelerating?

Of course, she would. Although the metric components are zero actually on the observer worldline, they are not zero in any neighbourhood of the worldline, even if it happens to be a geodesic. Worse, the connection components, and hence the coordinate derivatives of the metric components, are not even zero on the worldline, unless it is a geodesic. So in the case of an accelerating observer, the metric will not generally have the Minkowski form in these coordinates as soon as one leaves the observer worldline. If the observer can measure the metric accurately enough, she will know that she is accelerating and the apparent acceleration of the freely falling object can be attributed to coordinate choice.

The point about the above tautology is that, in the pre-relativistic case one was comparing two different test particles, one undergoing gravitational acceleration and the other in inertial motion (subject to no forces), whereas in the GR version of WEP3 one considers just one test particle in free fall, i.e., unaccelerating, and notes that it may appear to be accelerating relative to suitable coordinates. So this idea has changed somewhat, but it survives. It becomes something like this: locally, an observer cannot tell whether a test particle is freely falling or accelerating. The word 'locally' is there to remind us that the observer must not be able to survey a sufficiently wide neighbourhood of her worldline to detect the fact that it is really herself who accelerates, if that be the case.

We need to consider one more claim, which is in fact made by Bondi and Gold [2]: that it is impossible to distinguish between the action on a particle of matter of a constant acceleration or of static support in a gravitational field. This looks very like WEP2, except that it mentions a test particle. WEP2 says that, in a flat spacetime, an observer cannot decide whether there is a non-tidal gravitational field or not. So if the observer is moving with the test particle, she will not be able to say whether the particle is supported, with her, in a gravitational field, or whether the particle is undergoing a uniform acceleration, with her, depending on whether there is a gravitational field or not, respectively. Of course, one could say more simply that, in the GR understanding of Minkowski spacetime, static support in an SHGF interpretation of the spacetime is in fact just a uniform acceleration. If this can be viewed as a version of the weak equivalence principle, let us label it WEP4.

The measurements here refer to determination of the metric. If the test particles one is observing are charged, one expects to find EM fields and we have not yet addressed the question as to whether these fields could tell us more. This is in part the subject of the present book. It should be noted that the above versions WEP2 and WEP3 of the weak equivalence principle are not as fundamental as WEP on p. 15. When we come to examine EM fields, there is no reason a priori to suppose that the statements WEP2 or WEP3 should extend to cover such effects.

In fact one should ask what really is the utility of these ideas. Once GR has been constructed with WEP as an immediate consequence of the theory, and deployed physically with the help of the strong equivalence principle SEP (discussed further in Chap. 4), what need has one for any other consequences they may have? This will become a more pressing point in Chap. 12 which discusses the worldlines that charged particles are expected to have when falling 'freely' in curved spacetimes.

But to return to the main issue of the present chapter, as a flat spacetime, the spacetime with interval (2.45) or (2.48) escapes totally from the constraints imposed by Einstein's equations: the gravitational effects one might attribute to this scenario are non-tidal, i.e., the curvature is in fact zero. Put another way, although there might be some energy–momentum distribution that would lead via Einstein's equations to the SHGF proposed to explain the intervals, the completely empty energy–momentum distribution will also lead to this solution.

So the main point to bear in mind as we proceed is this: in the spacetime described by (2.45) or (2.48), without seeing any sources for gravitational effects, an observer cannot say whether she is uniformly accelerating in a flat spacetime (and using semi-Euclidean coordinates) or sitting still relative to coordinates that describe a spacetime containing a static, homogeneous gravitational field, or even a mixture of both. Note that sitting still in coordinates that describe such a spacetime is indeed an accelerating motion according to general relativity because one does not have geodesic motion. One might say that an observer cannot distinguish whether she has a uniform acceleration in a flat spacetime or a stationary state in a static, homogeneous gravitational field. The proviso is of course that one cannot see other things in the neighbourhood, such as sources of gravitation, or possibly other things, to be discussed in this book, viz., electromagnetic fields.

Furthermore, in the context of the spacetime described by (2.45) or (2.48), the weak principle of equivalence WEP as stated on p. 15 plays a key role. This principle is a fundamental building block in the general theory of relativity (without torsion). To recapitulate, it states that, given any spacetime event, one can always find coordinates in some neighbourhood of that event for which the metric has approximately the Minkowski form and the connection coefficients are approximately zero. But in the spacetime discussed here, everything works out exactly, i.e., there are coordinates for which the metric is everywhere exactly the Minkowski metric and the connection coefficients are everywhere exactly zero. This means that there are not even any higher order effects to help one make the kind of distinction described above.

And it will be worth remembering for future discussion that the gravitational field described by (2.45) or (2.48), if there is one, is necessarily static and homogeneous, for there are no tidal effects. One would like to say that this precludes any spatiotemporal variation in the field. However, it turns out that this conclusion is not so obvious as one would think, and this will lead to more subtleties in Sect. 15.7.

2.4 Weak Field Approximation

In the last section, we mentioned the weak field approximation, because it provides a point of contact between theory and observation that might help one to analyse the motions of test particles. The reason this gets a further section here is that there is a point to be noted about the way coordinates, and objects expressed relative to coordinates, are often somewhat naively interpreted. There are two parts to this:

- The first is rather standard, obtaining the approximate relation (2.50) between the metric and a Newtonian gravitational potential.
- The second concerns the relation between gravitational source and metric via Einstein's equation in this approximate context.

The following is adapted from DeWitt's Stanford lectures [24].

We consider a situation in which coordinates, called quasi-canonical coordinates, can be found, defined by the property that

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} , \qquad \text{where } |h_{\mu\nu}| \lesssim \varepsilon \ll 1 , \quad \forall \mu, \nu . \tag{2.59}$$

Such a coordinate system is not unique. It remains quasi-canonical under rotations and low-velocity Lorentz boosts. (Note that the $\eta_{\mu\nu}$ term in the metric would be fixed even by high-velocity boosts, but the $h_{\mu\nu}$ might be made large by such a transformation.) It also remains quasi-canonical under general coordinate transformations $\bar{x}^{\mu} = x^{\mu} + \xi^{\mu}$ restricted only by the requirement

$$|\xi^\mu{}_{,\nu}| \lesssim \varepsilon \qquad \text{and} \quad |h_{\mu\nu,\sigma}\xi^\sigma| \lesssim \varepsilon^2 \ , \quad \forall \mu, \nu \ . \tag{2.60}$$

One then finds a new $\bar{h}_{\mu\nu}$ defined by $\bar{h}_{\mu\nu} := \bar{g}_{\mu\nu} - \eta_{\mu\nu}$, whence

$$\begin{aligned}
\eta_{\mu\nu} + \bar{h}_{\mu\nu} = \bar{g}_{\mu\nu} &= \frac{\mathrm{d}x^\sigma}{\mathrm{d}\bar{x}^\mu}\frac{\mathrm{d}x^\tau}{\mathrm{d}\bar{x}^\nu} g_{\sigma\tau} \\
&= (\delta^\sigma{}_\mu - \xi^\sigma{}_{,\mu})(\delta^\tau{}_\nu - \xi^\tau{}_{,\nu})(\eta_{\sigma\tau} + h_{\sigma\tau}) \\
&= \eta_{\mu\nu} + h_{\mu\nu} - \xi_{\nu,\mu} - \xi_{\mu,\nu} + O(\varepsilon^2) \ ,
\end{aligned} \tag{2.61}$$

using the conditions in (2.60). So $h_{\mu\nu}$ undergoes the transformation

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \xi_{\mu,\nu} - \xi_{\nu,\mu} \ , \qquad \xi_\mu := \eta_{\mu\nu}\xi^\nu \ , \tag{2.62}$$

for such a change of coordinates. Note that in these approximations, one can always use the Minkowski metric to raise and lower indices.

The choice of quasi-canonical coordinate system can be restricted by imposing a supplementary condition

$$l^{\mu\nu}{}_{,\nu} = 0 \ , \tag{2.63}$$

where

$$l_{\mu\nu} := h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h \ , \qquad h := h_\sigma{}^\sigma \ , \tag{2.64}$$

$$h_{\mu\nu} = l_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}l \ , \qquad l := l_\sigma{}^\sigma = -h \ . \tag{2.65}$$

Now $l_{\mu\nu}$ transforms by

$$\begin{aligned}
\bar{l}_{\mu\nu} &= \bar{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\bar{h} \\
&= l_{\mu\nu} - \xi_{\mu,\nu} - \xi_{\nu,\mu} + \eta_{\mu\nu}\xi^\sigma{}_{,\sigma} \ .
\end{aligned} \tag{2.66}$$

It follows that

$$\begin{aligned}
\bar{l}^{\mu\nu}{}_{,\nu} &= l^{\mu\nu}{}_{,\nu} - \xi^\mu{}_{,}{}^\nu{}_\nu - \xi^\nu{}_{,}{}^\mu{}_\nu + \xi^\sigma{}_{,\sigma}{}^\mu \\
&= l^{\mu\nu}{}_{,\nu} - \xi^\mu{}_{,}{}^\nu{}_\nu \ .
\end{aligned}$$

So $\bar{l}_{\mu\nu}$ can be made to satisfy the supplementary condition (2.63) by choosing ξ^μ as a solution of the wave equation with a source:

$$\Box^2\xi^\mu := \xi^\mu{}_{,}{}^\nu{}_\nu = l^{\mu\nu}{}_{,\nu} \ . \tag{2.67}$$

The supplementary condition (2.63) does not fix the coordinate system completely, because it is still possible to carry out rotations, low-velocity Lorentz boosts, and coordinate transformations satisfying the homogeneous wave equation

$$\Box^2 \xi^\mu = 0 \,. \tag{2.68}$$

For gravitational sources with low velocities and when there is no gravitational radiation around, one can impose another condition on the above quasi-canonical coordinate systems. Indeed, one expects there to exist quasi-stationary coordinates. These are quasi-canonical coordinate systems, fixed in some intuitive sense with respect to the gravitational sources, in which all time derivatives $h_{\mu\nu,0}$ may be neglected compared with the spatial derivatives $h_{\mu\nu,i}$ $(i = 1,2,3)$. Note the interpretation of the zero coordinate as time and the other three coordinates as space, on the basis of the almost Minkowskian form of the metric. If such coordinates exist, the gravitational field itself is said to be quasi-stationary.

Any two quasi-stationary coordinate systems can be related by a rotation or a general transformation $\bar{x}^\mu = x^\mu + \xi^\mu$ in which the ξ^μ satisfy the previous conditions and in addition have negligible time derivatives. Under the latter, the components of $h_{\mu\nu}$ and $l_{\mu\nu}$ transform according to

$$\bar{h}_{00} = h_{00} \,, \qquad \bar{h}_{0i} = h_{0i} - \xi_{0,i} \,, \qquad \bar{h}_{ij} = h_{ij} - \xi_{i,j} - \xi_{j,i} \,, \tag{2.69}$$

$$\bar{l}_{00} = l_{00} - \xi_{i,i} \,, \qquad \bar{l}_{0i} = l_{0i} - \xi_{0,i} \,, \qquad \bar{l}_{ij} = l_{ij} - \xi_{i,j} - \xi_{j,i} + \delta_{ij}\xi_{k,k} \,. \tag{2.70}$$

The supplementary condition (2.63) is then imposed on $\bar{l}_{\mu\nu}$ by choosing ξ^μ to satisfy

$$\nabla^2 \xi^\mu := \xi^\mu{}_{,ii} = l^{\mu i}{}_{,i} \,. \tag{2.71}$$

Now h_{00} is the same for all these quasi-stationary coordinate systems, since it is not affected by the general coordinate transformations $\bar{x}^\mu = x^\mu + \xi^\mu$ and it is not affected by rotations. It is this uniqueness that allows one to make contact between the formalism of general relativity and observation, or at least, between Einstein's theory of gravity and Newton's.

The idea is to consider a freely falling particle, assumed to be moving slowly compared with light. It has worldline $z^\mu(t)$, where $t = x^0$, and this is assumed to satisfy the geodesic equation (but see Sect. 2.5 for more on that). The slow motion implies that

$$\left|\frac{\mathrm{d}z^i}{\mathrm{d}t}\right| \ll \left|\frac{\mathrm{d}z^0}{\mathrm{d}t}\right| = 1 \,. \tag{2.72}$$

The geodesic equation [see (2.95) on p. 36] becomes

$$\frac{\mathrm{d}^2 z^\mu}{\mathrm{d}t^2} + \Gamma^\mu_{\nu\sigma}\frac{\mathrm{d}z^\nu}{\mathrm{d}t}\frac{\mathrm{d}z^\sigma}{\mathrm{d}t} = h(t)\frac{\mathrm{d}z^\mu}{\mathrm{d}t} \,, \tag{2.73}$$

where

$$h(t) = -\frac{\mathrm{d}^2 t/\mathrm{d}\tau^2}{(\mathrm{d}t/\mathrm{d}\tau)^2} = \frac{\mathrm{d}^2\tau/\mathrm{d}t^2}{\mathrm{d}\tau/\mathrm{d}t} ,$$

using t as parameter rather than the proper time τ of the test particle. Note that t is not an affine parameter and this explains the term in h. The spatial part of this equation is

$$\frac{\mathrm{d}^2 z^i}{\mathrm{d}t^2} + \Gamma^i_{00} + 2\Gamma^i_{0j}\frac{\mathrm{d}z^i}{\mathrm{d}t} + \Gamma^i_{jk}\frac{\mathrm{d}z^j}{\mathrm{d}t}\frac{\mathrm{d}z^k}{\mathrm{d}t} = h(t)\frac{\mathrm{d}z^i}{\mathrm{d}t} ,$$

in which we neglect the last term on the left. We have

$$h^{\mu\nu} = \eta^{\mu\sigma}\eta^{\nu\rho}h_{\sigma\rho} ,$$

where $\eta^{\mu\sigma} = \eta_{\mu\sigma}$. Then note that, from

$$\Gamma^\mu_{\nu\sigma} = \frac{1}{2}g^{\mu\rho}\left(\partial_\nu g_{\sigma\rho} + \partial_\sigma g_{\nu\rho} - \partial_\rho g_{\nu\sigma}\right) ,$$

our approximation gives

$$\Gamma^\mu_{\nu\sigma} \approx \frac{1}{2}(\eta^{\mu\rho} - h^{\mu\rho})\left(\partial_\nu h_{\sigma\rho} + \partial_\sigma h_{\nu\rho} - \partial_\rho h_{\nu\sigma}\right) .$$

When we insert this into the above formula, we will be able to use

$$\Gamma^\mu_{\nu\sigma} \approx \frac{1}{2}\eta^{\mu\rho}(\partial_\nu h_{\sigma\rho} + \partial_\sigma h_{\nu\rho} - \partial_\rho h_{\nu\sigma}) .$$

Note that we have used the approximation

$$g^{\mu\rho} \approx \eta^{\mu\rho} - h^{\mu\rho} ,$$

which follows from

$$(\eta^{\mu\rho} - h^{\mu\rho})(\eta_{\rho\nu} + h_{\rho\nu}) \approx \delta^\mu_\nu .$$

Now

$$\Gamma^i_{00} \approx \frac{1}{2}\eta^{i\rho}\left(\partial_0 h_{0\rho} + \partial_0 h_{0\rho} - \partial_\rho h_{00}\right) \approx -\frac{1}{2}\eta^{ij}\,\partial_j h_{00} ,$$

so we can conclude that

$$\Gamma^i_{00} \approx \frac{1}{2}\delta^{ij}\,\partial_j h_{00} .$$

We have neglected $\partial_0 h_{\mu\nu}$ in comparison to $\partial_i h_{\mu\nu}$, $i = 1, 2, 3$. Also,

$$\Gamma^i_{0j} \approx \frac{1}{2}\eta^{i\rho}(\partial_0 h_{j\rho} + \partial_j h_{0\rho} - \partial_\rho h_{0j}) \approx -\frac{1}{2}\delta^{ik}(\partial_j h_{0k} - \partial_k h_{0j}) ,$$

once again neglecting $\partial_0 h_{\mu\nu}$.

We must now approximate the right-hand side of the geodesic equation (2.73). Since

$$\left(\frac{\mathrm{d}\tau}{\mathrm{d}t}\right)^2 = g_{\mu\nu}\frac{\mathrm{d}z^\mu}{\mathrm{d}t}\frac{\mathrm{d}z^\nu}{\mathrm{d}t} ,$$

we have

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} \approx (1+h_{00})^{1/2} \approx 1+\frac{1}{2}h_{00} .$$

Consequently,

$$\frac{\mathrm{d}^2\tau}{\mathrm{d}t^2} \approx \frac{1}{2}h_{00,0} .$$

We now have a formula for the function $h(t)$:

$$h(t) \approx \frac{1}{2}h_{00,0}\left(1-\frac{1}{2}h_{00}\right) \approx \frac{1}{2}h_{00,0} .$$

We can therefore neglect the term on the right-hand side of the geodesic equation (2.73).

Finally, the geodesic equation becomes approximately

$$m\frac{\mathrm{d}^2 z^i}{\mathrm{d}t^2} \approx -m\delta^{ij}\partial_j\left(\frac{1}{2}h_{00}\right) + m\delta^{ik}(\partial_j h_{0k} - \partial_k h_{0j})\frac{\mathrm{d}z^j}{\mathrm{d}t} ,$$

where m is the mass of the test particle. The second term on the right-hand side is velocity dependent and can be neglected on the grounds of the slow motion hypothesis (2.72) or attributed to a rotational effect and removed by declaring a non-rotating, quasi-canonical coordinate system to be one in which

$$\partial_j h_{0k} - \partial_k h_{0j} = 0 .$$

In this case,

$$m\frac{\mathrm{d}^2 z^i}{\mathrm{d}t^2} \approx -m\delta^{ij}\partial_j\left(\frac{1}{2}h_{00}\right) .$$

We can now make the identification [see (2.50) on p. 17]

$$h_{00} = 2\Phi , \tag{2.74}$$

so that

$$\frac{\mathrm{d}^2 z^i}{\mathrm{d}t^2} = -\frac{1}{2}h_{00,i} = -\Phi_{,i} , \tag{2.75}$$

where we make the interpretation that Φ is what would be called the Newtonian gravitational potential.

But who would call this a Newtonian potential? If one could find an observer who would set up these coordinates in her neighbourhood, and if this observer were to pretend that they were Minkowskian coordinates in a flat spacetime under the jurisdiction of special relativity, in which gravity really is a force field modelled by a potential, then observation of freely falling particles would lead this observer (under these illusions!) to attribute a Newtonian gravitational potential to this region of spacetime. This is much more sophisticated than the Newtonian idea that accelerating observers can pretend that they are not accelerating and attribute free particle motions to some gravitational potential. What makes the Newtonian version so simple is just the fact that time and distance are the same in the accelerating frame.

Note that, in the Newtonian spacetime, the accelerating observer who pretends she is not accelerating and attributes free particle motions to a gravitational field will be at pains to understand what sourced that field. We shall come in a moment to the question of the source. But let us just return to the observation on p. 18 that the metric components (2.48) on p. 17 would give us a quite different Newtonian gravitational potential. Of course, to get the metric components into the form (2.48), we have changed coordinates, and what has happened is that the quantity h_{00} in the above theory has not remained unchanged. This in turn is telling us that we have made a disallowed coordinate transformation for the above theory. This is not one of the coordinate changes that kept the coordinates in the quasi-stationary and quasi-canonical form. And yet the semi-Euclidean coordinates giving the intervals (2.45) or (2.48) are stationary, and they are quasi-canonical in DeWitt's sense, provided we restrict to coordinate values satisfying

$$\left|\frac{gy^1}{c^2}\right| \ll 1 \qquad \text{or} \qquad \left|\frac{g'y'^1}{c^2}\right| \ll 1\,, \quad \text{respectively}\,.$$

So DeWitt's allowed sets of coordinate transformations would appear to divide coordinate systems into disjoint classes, each of which can be considered as quasi-stationary and quasi-canonical. Note in passing that the above restriction on the coordinate values keeps us well away from the singularity in these metrics, which occurs when $y^1 = -c^2/g$.

But the key point to note here is that, in order to see a gravitational potential from an initial context, viz., general relativity, in which gravity is not modelled that way, we have to pretend that slightly non-inertial coordinates are in fact inertial, in the sense that they are things we can understand in that way, as though we were doing Newtonian gravitation. This point will be picked up in Sect. 15.7 when we consider Boulware's claim that there is an extremely strong gravitational field near the horizon of any uniformly accelerating semi-Euclidean observer.

Now if one knew the source of gravity, e.g., a star, one would expect this to help in establishing what part of a particle motion could be attributed to gravity and what part could be explained as an artefact of the coordinates. It would at least establish that there is some gravity. So let us return to the case of a massive point source described by (2.54) on p. 19, viz.,

$$T^{00} = M\delta(\mathbf{x}) , \tag{2.76}$$

and see rather briefly how DeWitt's theory in [24] makes this second link between GR and observation. DeWitt's aim was to establish the value of the constant κ in Einstein's equations

$$G^{\mu\nu} := R^{\mu\nu} - \frac{1}{2}Rg^{\mu\nu} = \frac{1}{2\kappa}g^{-1/2}T^{\mu\nu} , \tag{2.77}$$

where $G^{\mu\nu}$ is the Einstein tensor, $R^{\mu\nu}$ the Ricci tensor, g the modulus of the determinant of the covariant metric components $g_{\mu\nu}$, and $T^{\mu\nu}$ the energy–momentum tensor for the matter and energy distribution in the spacetime. Our aim will of course be to see whether this analysis throws any light on the SHGF.

In the quasi-canonical coordinate systems described above, it is first shown that the curvature tensor takes the approximate form

$$R_{\mu\nu\sigma\tau} = -\frac{1}{2}(h_{\mu\sigma,\nu\tau} + h_{\nu\tau,\mu\sigma} - h_{\mu\tau,\nu\sigma} - h_{\nu\sigma,\mu\tau}) , \tag{2.78}$$

also called the linearised Riemann tensor, which turns out to be invariant under the approximate coordinate transformation law (2.62) on p. 25. This invariance is basically due to the fact that, unlike the metric tensor, the Riemann tensor is already first order in small quantities. DeWitt remarks that the physical significance of this invariance is that the presence or absence of a real gravitational field is characterised by the presence or absence of a nonzero Riemann tensor. Since this tensor represents the gravitational field, the weak field approximation is going to provide an invariant characterisation of the field, that is, an expression for the field that is independent of which quasi-canonical coordinate system we are using.

The problem for the SHGF is of course that the Riemann tensor is zero, so it would not count as a real gravitational field in DeWitt's remark. And as we just noted when examining the worldlines of free particles, different SE observers can attribute any Newtonian potential they like to the spacetime neighbourhood, precisely because changing from one uniformly accelerating SE coordinate system to another does not keep one inside the same class of quasi-canonical coordinate systems.

The linearised Ricci tensor and curvature scalar are

$$R_{\mu\nu} = \eta^{\sigma\tau}R_{\mu\sigma\nu\tau} = -\frac{1}{2}\eta^{\sigma\tau}\left(h_{\mu\nu,\sigma\tau} + h_{\sigma\tau,\mu\nu} - h_{\sigma\nu,\mu\tau} - h_{\mu\tau,\sigma\nu}\right) , \tag{2.79}$$

$$R = \eta^{\mu\nu}R_{\mu\nu} = -h_{,\mu}{}^{\mu} + h^{\mu\nu}{}_{,\mu\nu} . \tag{2.80}$$

The linearised Einstein equations are then

$$\frac{1}{2\kappa}T^{\mu\nu} = -\frac{1}{2}\left(l^{\mu\nu}{}_{,\sigma}{}^{\sigma} - l^{\mu\sigma}{}_{,}{}^{\nu}{}_{\sigma} - l^{\nu\sigma}{}_{,}{}^{\mu}{}_{\sigma} + \eta^{\mu\nu}l^{\sigma\tau}{}_{,\sigma\tau}\right) . \tag{2.81}$$

This is where the supplementary condition (2.63), viz., $l^{\mu\nu}{}_{,\nu} = 0$, comes into its own. When coordinates are chosen so that this condition is satisfied, which is described as choosing the de Donder gauge, the linearised Einstein equations take the simple form

$$\Box^2 l^{\mu\nu} := l^{\mu\nu}{}_{,\sigma}{}^{\sigma} = -\frac{1}{\kappa} T^{\mu\nu} . \tag{2.82}$$

We can now tackle the point mass source at the origin of some coordinate system of the kind singled out above, for which we propose $T^{00} = M\delta(\mathbf{x})$, with all other components of the energy–momentum tensor being zero.

Of course there is problem of circularity here, because we do not know a priori whether the spacetime around such an object will allow us to find a stationary, quasi-canonical system. Since the point mass is at least stationary relative to the spatial coordinates $\mathbf{x}$, one might well expect a stationary system, but it is only with hindsight that we could justify this process. So we simply push ahead by requiring the supplementary condition $l^{\mu\nu}{}_{,\nu} = 0$ and observing that $l^{\mu\nu}$ can then have only one nonzero component, viz., l^{00}, because (2.82) implies $\Box^2 l^{\mu\nu} = 0$ whenever $(\mu,\nu) \neq (0,0)$. [Note that the supplementary condition can still be imposed in a quasi-stationary coordinate system, as evidenced by (2.71) on p. 26.] Equation (2.82) also gives

$$\nabla^2 l^{00} = \frac{1}{\kappa} T^{00} = \frac{1}{\kappa} M\delta(\mathbf{x}) . \tag{2.83}$$

The solution satisfying

$$\lim_{|\mathbf{x}|\to\infty} l^{00} = 0 , \tag{2.84}$$

and only taken seriously within the context of the weak field approximation when $|\mathbf{x}| \gg GM$, is

$$l^{00}(x) = -\frac{M}{4\pi\kappa}\frac{1}{|\mathbf{x}|} = \frac{1}{4\pi\kappa G}\Phi , \tag{2.85}$$

where we have introduced the usual definition for the Newtonian potential of a point particle:

$$\Phi(x) = -\frac{GM}{|\mathbf{x}|} . \tag{2.86}$$

The last manoeuvre is a sleight of hand! Although Φ may be the same function of x as the Newtonian potential, x itself is not a Cartesian coordinate for spacetime, except in an approximate sense. We now retrieve $h_{\mu\nu}$ from

$$l = l^{00} = l_{00} , \qquad h_{00} = l_{00} - \frac{1}{2} l = \frac{1}{2} l_{00} = \frac{1}{8\pi\kappa G}\Phi ,$$

$$h_{ij} = l_{ij} + \frac{1}{2}\delta_{ij} l = \frac{1}{2}\delta_{ij} l_{00} = \delta_{ij}\frac{1}{8\pi\kappa G}\Phi \,, \qquad h_{0i} = l_{0i} = 0 \,.$$

As noted in (2.69) on p. 26, if we transform to another quasi-stationary coordinate system in the same class, h_{ij} and h_{0i} will take other values, but h_{00} will remain unchanged to this level of approximation. Note, however, that it is the boundary condition (2.84) which rules out the addition of some affine function of the space coordinates to the proposed solution for l^{00}, and such an affine function, e.g., gz, is precisely the kind of linear potential that leads to a static, homogeneous component to the gravitational field. If one switched to some other class of quasi-stationary and quasi-canonical coordinate system, one would then have to change the boundary condition (2.84) in order to obtain an approximation to Newtonian gravity.

In his presentation [24], DeWitt uses this theory to deduce the value of κ in Einstein's equations (2.77). For general relativity to agree with Newtonian theory about the motion of bodies under the action of gravity in this weak field approximation, we noted that $h_{00} = 2\Phi$ [see (2.74) on p. 28]. It follows that

$$\kappa = \frac{1}{16\pi G} \,, \tag{2.87}$$

whence the full Einstein equations take the form

$$G^{\mu\nu} := R^{\mu\nu} - \frac{1}{2} R g^{\mu\nu} = 8\pi G g^{-1/2} T^{\mu\nu} \,. \tag{2.88}$$

Now we were looking at all this in the hope that it might throw light on the SHGF. What could source a static, homogeneous gravitational field in the Newtonian theory? The answer is of course an infinite plane of matter of infinitesimal thickness and unchanging, uniform surface density σ. Let us take this as the plane $z = 0$. The Newtonian gravitational potential at a height z above the plane is the sum of all the potentials due to surface elements on the plane. Since sets of points on the plane that are equidistant from the field point constitute rings, one carries out a simple integration over the whole plane by dividing it into thin rings centered on the point directly below the field point. The result is

$$\Phi(z) = 2\pi G\sigma z \qquad (z > 0) \,. \tag{2.89}$$

The force field per unit mass in this Newtonian gravitational (NG) theory is then

$$(-\Phi_{,1}, -\Phi_{,2}, -\Phi_{,3}) = (0, 0, -2\pi G\sigma) \,,$$

which is constant in space and time, and pointing downward for a field point above the plane. When $z < 0$, the sign of Φ is reversed, so that we have a spatiotemporally constant gravitational force field pulling up toward the plane $z = 0$. But note that for a given point $z > 0$, this same matter plane could be anywhere below it and the gravitational effect would be the same. Presumably it could even be moving back and forth up the z axis and the observer at some z that always remained above it would not notice, according to NG.

Let us thus consider the possibility

$$T^{00} = \sigma\delta(z) , \tag{2.90}$$

in the weak field approximation of general relativity, where σ is the constant mass per unit area of a plane through the spatial origin of some coordinate system, taking $T^{\mu\nu} = 0$ whenever $(\mu,\nu) \neq (0,0)$. If we can arrange with hindsight for the coordinates to be quasi-stationary and quasi-canonical, and satisfy the supplementary condition as in the analysis for a point particle source, then Einstein's equations take the simple form

$$\nabla^2 l^{00} = \frac{1}{\kappa}T^{00} = \frac{1}{\kappa}\sigma\delta(z) . \tag{2.91}$$

Of course, assuming that l^{00} is a function of z alone, this becomes

$$\frac{\mathrm{d}^2 l^{00}}{\mathrm{d}z^2} = \frac{\sigma}{\kappa}\delta(z) , \tag{2.92}$$

which is easy to solve because

$$\delta(z) = \frac{\mathrm{d}^2}{\mathrm{d}z^2}\left(\frac{|z|}{2}\right) .$$

We find

$$l^{00} = c_1 + c_2 z + \frac{\sigma}{2\kappa}|z| ,$$

where c_1 and c_2 are constants. Since we expect reflection symmetry in the spatial plane $z = 0$, this implies $c_1 = 0 = c_2$, and the proposed solution is

$$l^{00} = \frac{\sigma}{2\kappa}|z| .$$

This time there is no solution with

$$\lim_{|\mathbf{x}|\to\infty} l^{00} = 0 ,$$

but in fact we are expecting our approximations to be valid close to $z = 0$. We observe with hindsight that this solution does satisfy the supplementary condition $l^{\mu\nu}{}_{,\nu} = 0$. Once again, $l = l^{00} = l_{00}$, and inserting the value of κ,

$$h_{00} = l_{00} - \frac{1}{2}l = \frac{1}{2}l_{00} = 4\pi G\sigma|z| .$$

Since $h_{00} = 2\Phi$ is supposed to give the corresponding Newtonian potential, we then find

$$\Phi = 2\pi G\sigma|z| , \tag{2.93}$$

the same as in the purely Newtonian theory (NG) [see (2.89)].

This looks very nice, but there are two reasons why we have not really achieved our uniformly accelerating SE metric:

- The value of $g_{00} := \eta_{00} + h_{00}$ differs from the SE metric when $z < 0$, and this is obviously an essential part of the solution!
- The components h_{ij} of the above solution are not zero, as they would be for the SE metric.

Still, in general relativity, if we had an infinite plane of constant mass density σ, we might expect something like the uniformly accelerating SE metric, which is a flat space solution, on either side of the plane, just because in NG we find that this gives a solution with no tidal effects (implying zero curvature in GR). But if this SE metric does satisfy the Einstein equations for such a plane, where is it, and what is it doing? In the uniformly accelerating SE coordinates, it is amusing to imagine that the plane $y^1 = 0$ would do the trick, proposing

$$T^{00} = \sigma\delta(z) , \qquad T^{\mu\nu} = 0 \quad \forall\, (\mu,\nu) \neq (0,0) .$$

In fact, rather than trying to solve the full Einstein equations for this idealistic energy–momentum distribution, one can calculate the Einstein tensor for the metric

$$g_{00} = \left(1 + \frac{g|z|}{c^2}\right)^2 , \qquad g_{0i} = 0 = g_{i0} , \qquad g_{ij} = -\delta_{ij} , \quad i,j \in \{1,2,3\} . \tag{2.94}$$

It is interesting to note that this gives the following results. The connection coefficients are all zero except for

$$\Gamma^3_{00} = \frac{g}{c^2}\left(1 + \frac{g|z|}{c^2}\right)\theta(z) , \qquad \Gamma^0_{30} = \frac{g}{c^2}\left(1 + \frac{g|z|}{c^2}\right)^{-1}\theta(z) = \Gamma^0_{03} ,$$

where

$$\theta(z) := \begin{cases} -1 \text{ for } z < 0 , \\ +1 \text{ for } z > 0 , \end{cases}$$

and the components of the Ricci tensor are all zero except for

$$R_{00} = \frac{g}{c^2}\delta(z) , \qquad R_{33} = -\frac{g}{c^2}\delta(z) .$$

The curvature scalar is then

$$R = \frac{2g}{c^2}\delta(z) ,$$

and the Einstein tensor is

$$G_{00} = 0 = G_{33} \;, \qquad G_{11} = \frac{g}{c^2}\delta(z) = G_{22} \;, \qquad G_{\mu\nu} = 0 \quad \forall\, \mu \neq \nu \;.$$

Raising the indices, one finds that the energy–momentum distribution required to give this metric is specified by

$$T^{00} = 0 = T^{33} \;, \qquad T^{11} = \frac{g}{8\pi G c^2}\delta(z) = T^{22} \;, \qquad T^{\mu\nu} = 0 \quad \forall\, \mu \neq \nu \;.$$

The interpretation of this result is left to the reader. What is clear is that Einstein's equations were never intended to fulfill our Newtonian dreams.

So let us make a brief conclusion. What concerned us in the last section was how an observer at fixed SE spatial coordinate might try to understand the uniformly accelerating SE metric, or at least the way freely falling particles move around relative to her coordinate systems. We have seen that, on a local basis, any Newtonian gravitational potential whatever could be viewed as responsible for such motions. Furthermore, in a Newtonian world, where such effects could be generated by an infinite matter plane, one could measure accelerations of test particles relative to this source and distinguish them from accelerations of the observer, whereas in the world of general relativity one cannot make this kind of distinction so easily, because one cannot identify even an idealistic source so easily. On the other hand, one expects the SE metric to be a good approximation in certain spacetime regions, e.g., in a laboratory situated a long way from a star. In the latter case, which comes under the jurisdiction of the Schwarzschild solution to Einstein's equations, one can to a certain extent and with a certain proviso use the proper distance from the source to distinguish coordinate accelerations (of a test particle) due to the source and those due to choice of coordinates. The inevitable level of approximation is what determines the extent to which this will succeed, while the fact that proper distance depends on a subjective choice of spacelike hypersurface stands as a proviso.

In the highly idealistic, uniformly accelerating semi-Euclidean spacetime, one has to accept that one does not know where the gravitational source is located so that one cannot use it to make deductions about the motion of free particles. In that case, even though the curvature is zero, whence there are no tidal effects in this spacetime, one has to accept that different SE observers can choose to explain free particle motions by quite different Newtonian potentials in the weak field approximation. And we shall see in Chap. 15 that some authors interpret this gravitational field as having different strengths at different locations in spacetime. But it should be remembered that such an observer is not using locally inertial coordinates. In order to see a gravitational potential from an initial context, viz., general relativity, in which gravity is not modelled that way, we have to pretend that slightly non-inertial coordinates are in fact inertial, in the sense that they are things we can understand in that way, as though we were doing Newtonian gravitation.

To put it another way, in the weak field approximation to free fall, one can only get a Newtonian potential if the connection coefficients are not all zero at the event in question, whence the coordinate system is not locally inertial. Furthermore, this potential does not make a distinction between what one might call artefacts of the choice of coordinates, which might possibly be ascribed to the motion of an observer sitting at the spatial origin of that coordinate system, and gravitational effects that are due to the energy–momentum distribution. Even in the context of the weak field analysis for a point particle source described above, Einstein's equations allow the addition of any affine function of a space coordinate to the proposed Newtonian potential, and this affine function is only ruled out by boundary conditions imposed by the requirement that things look roughly as they would according to NG. And finally, that Newtonian potential may be the usual function of the space coordinates for such a source, but those coordinates have to be assimilated with the Cartesian coordinates of Newtonian physics to make this interpretation.

2.5 Geodesic Principle

This section can be considered as a slightly off-beat review of some basic relativity theory. We return to the interesting question of the geodesic principle brought up in Sect. 2.3. This states that, when point particles are not acted upon by forces (apart from gravitational effects), their trajectories take the form

$$\frac{\mathrm{d}^2x^i}{\mathrm{d}s^2}+\Gamma^i_{jk}\frac{\mathrm{d}x^j}{\mathrm{d}s}\frac{\mathrm{d}x^k}{\mathrm{d}s}=0\,, \tag{2.95}$$

where $x^i(s)$ gives the worldline as a function of the proper time s of the particle and Γ^i_{jk} are the connection coefficients in the given coordinate system. In the literature, this is often derived from an action principle. One writes the worldline as a function $x^i(\lambda)$ of some arbitrary parameter λ, whence the appropriate action for the worldline between two points $P_1=x(\lambda_1)$ and $P_2=x(\lambda_2)$ of spacetime is

$$s(P_1,P_2):=\int_{\lambda_1}^{\lambda_2}\left(g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right)^{1/2}\mathrm{d}\lambda=\int_{\lambda_1}^{\lambda_2}L\mathrm{d}\lambda=\int_{\lambda_1}^{\lambda_2}\mathrm{d}s\,, \tag{2.96}$$

with Lagrangian

$$L:=\left(g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right)^{1/2}. \tag{2.97}$$

The Euler–Lagrange equations extremising the action under variation of the worldline are

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial L}{\partial\dot{x}^i}\right)-\frac{\partial L}{\partial x^i}=0\,, \tag{2.98}$$

with $\dot{x}^i:=\mathrm{d}x^i/\mathrm{d}\lambda$, and these lead to the above geodesic equation (2.95).

The action for some particles labelled by a is

$$\mathscr{A} = -\sum_a cm_a \int \mathrm{d}s_a \ , \tag{2.99}$$

where m_a is the mass of particle a and s_a is its proper time. This is the action *because* variation of the worldline of particle a gives its equation of motion as

$$\frac{\mathrm{d}^2 a^i}{\mathrm{d}s_a^2} + \Gamma^i_{jk} \frac{\mathrm{d}a^j}{\mathrm{d}s_a} \frac{\mathrm{d}a^k}{\mathrm{d}s_a} = 0 \ . \tag{2.100}$$

Likewise, if some of the particles are charged with charge e_a for particle a, and there are some EM fields F_{ik}, one declares the action to be

$$\mathscr{A} = -\sum_a cm_a \int \mathrm{d}s_a - \frac{1}{16\pi c} \int F_{ik} F^{ik} (-g)^{1/2} \mathrm{d}^4 x - \sum_a \frac{e_a}{c} \int A_i \mathrm{d}a^i \ , \tag{2.101}$$

where A_i is a 4-vector potential from which F_{ij} derives, simply *because* variation of the worldline of particle a gives its equation of motion as

$$\frac{\mathrm{d}^2 a^i}{\mathrm{d}s_a^2} + \Gamma^i_{jk} \frac{\mathrm{d}a^j}{\mathrm{d}s_a} \frac{\mathrm{d}a^k}{\mathrm{d}s_a} = \frac{e_a}{m_a} F^i{}_j \frac{\mathrm{d}a^j}{\mathrm{d}s_a} \ , \tag{2.102}$$

the minimal generalisation of the Lorentz force law to a curved manifold, while variation of A_i gives the EM field equations as the minimal extension of Maxwell's equations to the curved spacetime. Of course, these actions are designed to give appropriate field equations, and we are just decreeing here that the appropriate field equation for the particle labelled by a is a geodesic equation, or an equation like (2.102).

So what is the physical motivation? The appropriate physical argument supporting (2.95) is an application of the strong principle of equivalence, and it is here that we discover exactly how we are to link what happens in the curved manifold with measurements in our own world. We start from an action, but at some point we must say what the point of contact would be with physical reality. Let us suppose we impose a strong principle of equivalence, that is, we say roughly speaking that any physical interaction other than gravitation behaves in a locally inertial frame as though gravity were absent (but see Chap. 4 for more details). Relative to such a frame, any particle that is not subject to (non-gravitational) forces will then move in a straight line with uniform velocity, i.e., it will follow the trajectory described by

$$\frac{\mathrm{d}v^i}{\mathrm{d}s} = 0 \ , \tag{2.103}$$

where v^i is its 4-velocity. This is expressed covariantly through (2.95). If there are non-gravitational forces, we start with $F = ma$ in the locally inertial frame and we find (2.95) with a force term on the right-hand side. It is quite clear that we still have a version of Newton's second law $F = ma$, so we have certainly not explained inertia and inertial effects by this ploy, but merely extended this equation of motion to the new theory.

Let us look more closely at the claim that (2.103) is expressed covariantly through (2.95). A cheap way is to set the connection coefficients equal to zero in (2.95). This is basically the observation that the two equations are the same relative to Cartesian coordinates in a flat spacetime. This misses out some of the machinery of the connection construction that lies at the heart of non-Euclidean geometry, and this is not the place to expose all that. However, the following is an amusing way to look a little more deeply into this affair.

Let us set the scene by assigning a connection to the manifold in the following way: for each point and coordinate neighbourhood of that point, we have a function whose value at any point in the neighbourhood is a set of N^3 numbers Γ^a_{bc}, where N is the dimension of the manifold. Given another neighbourhood, the quantities transform by the rule

$$\Gamma^{a'}_{b'c'} = \Gamma^d_{ef} X^{a'}_d X^e_{b'} X^f_{c'} + X^{a'}_d X^d_{c'b'} \ ,$$

where we have used the notation

$$X^{a'}_d := \frac{\partial x^{a'}}{\partial x^d} \ , \qquad X^d_{c'b'} := \frac{\partial^2 x^d}{\partial x^{c'} x^{b'}} \ .$$

One can show with a little algebra that this transformation is transitive.

Now suppose U is a coordinate neighbourhood of M and $\lambda^a(u)$ is defined along a curve γ, given by $x^a(u)$. Then

$$\dot{\lambda}^a = \frac{\mathrm{d}\lambda^a}{\mathrm{d}u}$$

is not a vector, in general. We define the absolute derivative of λ^a to be

$$\frac{\mathrm{D}\lambda^a}{\mathrm{d}u} = \frac{\mathrm{d}\lambda^a}{\mathrm{d}u} + \Gamma^a_{bc}\lambda^b \frac{\mathrm{d}x^c}{\mathrm{d}u} \ ,$$

for any N^3 numbers Γ^a_{bc} associated with each point and coordinate neighbourhood. If this is a vector field for all vector fields λ^a along γ, then it can be shown that Γ^a_{bc} must transform as stated above, using the quotient theorem for tensors. Conversely, if

$$\frac{\mathrm{D}\lambda^a}{\mathrm{d}u} = \frac{\mathrm{d}\lambda^a}{\mathrm{d}u} + \Gamma^a_{bc}\lambda^b \frac{\mathrm{d}x^c}{\mathrm{d}u} \ ,$$

and Γ^a_{bc} transforms as above, then $\mathrm{D}\lambda^a/\mathrm{d}u$ is a vector.

We can now extend the Euclidean idea of parallelism. Suppose P has coordinates x^a and some neighbouring point Q has coordinates $x^a + \delta x^a$. We seek some kind of bijection from $T_P(M)$ to $T_Q(M)$, so that we can call corresponding vectors parallel. We demand that the correspondence be linear and reduce to the identity when Q approaches P. (The point about linearity is that, if v is parallel to w, then λv ought to be parallel to λw, for any $\lambda \in \mathbb{R}$, and so on.)

Let $\lambda^a + \delta^* \lambda^a$ be the vector at Q which is parallel to λ^a at P. Then there exists Y^a_b, dependent on P and Q, but not on λ^a, such that

$$\lambda^a + \delta^* \lambda^a = Y^a_b \lambda^b \ .$$

To first order in δx^a, we would like to obtain

$$Y^a_b = \delta^a_b - \Gamma^a_{bc} \delta x^c \ ,$$

noting that Γ^a_{bc} depends only on P, and δx^c depends on both P and Q. This will be possible provided that

$$\delta^* \lambda^a = -\Gamma^a_{bc} \lambda^b \delta x^c \ . \tag{2.104}$$

We see that

$$\frac{\mathrm{D}\lambda^a}{\mathrm{d}u} = \lim_{\delta u \to 0} \frac{\delta \lambda^a - \delta^* \lambda^a}{\delta u} \ ,$$

where $\delta \lambda^a = \lambda^a(u + \delta u) - \lambda^a(u)$. Note that $\delta \lambda^a - \delta^* \lambda^a$ is a vector at Q.

For a flat space, parallelism is well-understood, and it is from this context that we wish to generalise. In Cartesian coordinates x^a, the connection components are zero, but in general coordinates $x^{a'}$, they are given by

$$\Gamma^{a'}_{b'c'} = \frac{\partial x^{a'}}{\partial x^d} \frac{\partial^2 x^d}{\partial x^{b'} \partial x^{c'}} \ ,$$

according to the transformation equation stipulated above.

Assuming this to be the general definition of connection coefficients in a flat space, relative to some arbitrary curvilinear coordinate system $\{x^{a'}\}$ expressed as functions of a Cartesian coordinate system $\{x^a\}$ (kept fixed throughout), it is instructive to derive the general transformation law for the connection in this flat space context. That is, we consider some other curvilinear coordinate system $x^{a''}$, in which the connection has coefficients

$$\Gamma^{e''}_{f''g''} = \frac{\partial x^{e''}}{\partial x^d} \frac{\partial^2 x^d}{\partial x^{f''} \partial x^{g''}} \ ,$$

and then check that

$$\Gamma^{e''}_{f''g''} = \Gamma^{a'}_{b'c'} \frac{\partial x^{e''}}{\partial x^{a'}} \frac{\partial x^{b'}}{\partial x^{f''}} \frac{\partial x^{c'}}{\partial x^{g''}} + \frac{\partial x^{e''}}{\partial x^{h'}} \frac{\partial^2 x^{h'}}{\partial x^{f''} \partial x^{g''}} \ .$$

This motivates the definition and transformation of connection coefficients in a general curved space, where the connection cannot be made zero on any neighbourhood.

Still working within our flat space, we can now see how these non-zero connection coefficients implement parallelism in general (non-Cartesian) coordinates.

In the Cartesian coordinates x^a, a vector v^a at x_0 is parallel to the vector with the same components v^a at $x_0 + \delta x$. Let us see what this looks like in the non-Cartesian primed coordinate system. Let us write the primed components of each of these two vectors. Firstly, at x_0,

$$v^{b'} = \left.\frac{\partial x^{b'}}{\partial x^a}\right|_{x_0} v^a \ ,$$

and secondly, at $x_0 + \delta x$,

$$v^{b'} + \delta^* v^{b'} = \left.\frac{\partial x^{b'}}{\partial x^a}\right|_{x_0+\delta x} v^a \ .$$

We can evaluate the last term to first order in δx, to obtain

$$v^{b'} + \delta^* v^{b'} = \left\{ \left.\frac{\partial x^{b'}}{\partial x^a}\right|_{x_0} + \frac{\partial}{\partial x^c}\left.\frac{\partial x^{b'}}{\partial x^a}\right|_{x_0} \delta x^c \right\} v^a \ .$$

Comparing with (2.104), we are expecting to find

$$-\left.\Gamma^{b'}_{d'e'}\right|_{x_0} v^{d'} \delta x^{e'} = \left.\frac{\partial^2 x^{b'}}{\partial x^c \partial x^a}\right|_{x_0} \delta x^c \, v^a \ .$$

Inserting

$$v^a = \frac{\partial x^a}{\partial x^{d'}} v^{d'} \qquad \text{and} \qquad \delta x^{c'} = \frac{\partial x^c}{\partial x^{e'}} \delta x^{e'} \ ,$$

we are hoping to find that

$$-\left.\Gamma^{b'}_{d'e'}\right|_{x_0} = \left.\frac{\partial^2 x^{b'}}{\partial x^c \partial x^a}\right|_{x_0} \frac{\partial x^c}{\partial x^{e'}} \frac{\partial x^a}{\partial x^{d'}} \ .$$

The manipulation here is straightforward, first writing the right-hand side as

$$\frac{\partial}{\partial x^{e'}} \left\{ \frac{\partial x^{b'}}{\partial x^a} \right\} \frac{\partial x^a}{\partial x^{d'}} \ ,$$

and then as

$$\frac{\partial}{\partial x^{e'}} \left\{ \frac{\partial x^{b'}}{\partial x^a} \frac{\partial x^a}{\partial x^{d'}} \right\} - \frac{\partial x^{b'}}{\partial x^a} \frac{\partial^2 x^a}{\partial x^{e'} \partial x^{d'}} \ .$$

The first term here is zero, and we are left with exactly what we defined to be $-\Gamma^{b'}_{d'e'}$.

The moral of this story is that, for general coordinates in a flat space we require non-zero connection coefficients, and we carry over both the transformation equation and the way they come into vector parallelism when we generalise to manifolds which do not have Cartesian coordinate systems. All this is supposed to comfort the reader that the covariant generalisation of

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}s^2} = 0 \tag{2.105}$$

is the geodesic equation

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}s^2} + \Gamma^i_{jk} \frac{\mathrm{d}x^j}{\mathrm{d}s} \frac{\mathrm{d}x^k}{\mathrm{d}s} = 0 \, . \tag{2.106}$$

This is then taken as the equation of motion of a point particle upon which no forces are acting, unless one counts gravity as a force. We observe that there is no mention of any parameters characterising the point particle. In particular there is no mention of its inertial mass. Of course there is no mention of parameters describing its inner make-up. After all, it is supposed to be a point particle. One has to wonder what would happen to a slightly spatially extended particle, i.e., with a world tube that intersects spatial hypersurfaces in a small region rather than a single mathematical point. This object might be spinning in some sense, or contain a charge distribution, for example. We shall return to this point in a moment, but let us begin with the disappearance of the inertial mass since this is directly relevant to the notion of equivalence principle.

Indeed, one of the great equivalence principles of physics, in fact an experimental result that already features in Newtonian physics, is the equivalence of passive gravitational mass, the measure of how much a particle is supposed to feel the Newtonian force of gravity, and the inertial mass, the measure of how much a particle is supposed to resist being accelerated. This meant that, in Newtonian gravitational theory, an observer could not tell whether the acceleration of a test particle relative to her Euclidean coordinate system was due to some gravitational field or due to her own acceleration and the consequent acceleration of her comoving Euclidean coordinate system. (The acceleration of that Euclidean system must not be rotational here.) In pre-relativistic theory, one had no difficulty setting up Euclidean coordinate systems in spatial hypersurfaces and time was the same for everyone.

So where did the inertial mass of the particle go? If we look back to (2.102), viz.,

$$m_a \left(\frac{\mathrm{d}^2 a^i}{\mathrm{d}s_a^2} + \Gamma^i_{jk} \frac{\mathrm{d}a^j}{\mathrm{d}s_a} \frac{\mathrm{d}a^k}{\mathrm{d}s_a} \right) = e_a F^i{}_j \frac{\mathrm{d}a^j}{\mathrm{d}s_a} \, , \tag{2.107}$$

we find the inertial mass m_a multiplying the acceleration term in the equation to give a force on the right that is determined by an external field F_{ij} and a coupling constant e_a characterising the particle. This is a typical equation of motion when there is some non-gravitational force (in this case electromagnetic) acting on the particle. Now in Newtonian gravitational theory, if the external field happens to be a gravitational potential Φ, one gets an equation of motion like this:

$$m_a\left(\frac{\mathrm{d}^2a^i}{\mathrm{d}s_a^2}+\overline{\Gamma}^i_{jk}\frac{\mathrm{d}a^j}{\mathrm{d}s_a}\frac{\mathrm{d}a^k}{\mathrm{d}s_a}\right)=m_a h^{ij}\Phi_{,j}\ , \tag{2.108}$$

where $(h^{ij})=\mathrm{diag}\,(0,1,1,1)$ and $\overline{\Gamma}^i_{jk}$ is the connection appropriate to Newtonian spacetime and relative to whatever coordinates we have chosen to describe it. The coupling factor on the right-hand side is just m_a, the same factor as we have on the left-hand side. The only relevant characteristic of our point particle thus cancels out.

This explains how the inertial mass disappears from the equation, but how do we get rid of the gravitational potential we have just introduced? Of course, we can absorb it into the connection, following the much more detailed account of all this in [11]. We now have a new connection

$$\Gamma^i_{jk}:=\overline{\Gamma}^i_{jk}+h^{il}\Phi_{,l}t_jt_k\ , \tag{2.109}$$

where $(t_i):=(1,0,0,0)$, so that $t_i=\partial t/\partial x^i$. Equation (2.108) becomes

$$\frac{\mathrm{d}^2a^i}{\mathrm{d}s_a^2}+\Gamma^i_{jk}\frac{\mathrm{d}a^j}{\mathrm{d}s_a}\frac{\mathrm{d}a^k}{\mathrm{d}s_a}=0\ , \tag{2.110}$$

still in this Newtonian context. So the equality of inertial and passive gravitational mass allows us to treat the trajectories of particles subjected only to gravitational effects as geodesics of a non-flat connection, because we do expect this new connection in (2.109) to be non-flat in general.

As Friedman says in [11], the equality of inertial and passive gravitational mass implies the existence of a connection Γ such that freely falling objects follow geodesics of Γ. This does not work for other types of interaction, where the ratio of m_a to the coupling factor, e.g., e_a for a charged particle, is not the same for all bodies. The worldlines of charged particles in an EM field cannot be construed as the geodesics of any single connection, because m_a/e_a in (2.102) varies from one particle to another. To quote from Friedman [11, p. 197]:

> The principle of equivalence (strict proportionality) of inertial and gravitational mass must therefore be true if any theory of gravitation like general relativity, in which gravitational interaction is explained by the dependence of a nonflat connection on the distribution of matter, is to be possible. But, of course, general relativity is not the only theory of this type. Classical gravitational theory can also be formulated in this way by taking advantage of the very same equivalence [of inertial and passive gravitational mass].

Friedman's book is recommended for anyone who thinks that Newtonian gravitational theory cannot be given a fully covariant and totally geometric treatment. The essential difference with general relativity is that, in this treatment of Newtonian gravity, there is a flat connection $\overline{\Gamma}$ living alongside the non-flat connection Γ of (2.109). The deep fact here is that, in general relativity, the non-flat connection is the *only* connection of the spacetime. We return to this point when we introduce the strong equivalence principle in Chap. 4.

Anyway, we now know what a point particle is supposed to do. But what about a spatially extended particle? Well, it turns out that the geodesic principle is not a principle at all, and neither is it likely to be better than an approximation for a real particle that cannot be treated as a mathematical point. This will be directly relevant in Chap. 12 when we discuss DeWitt and Brehme's extension of the Lorentz–Dirac equation to curved spacetimes [4], so it is worth spelling things out right away.

The point is that the geodesic principle follows from Einstein's equations in general relativity. We cannot go into all the details here. They can be found in the standard textbooks. However, it is worth seeing a proof of sorts. Recall that Einstein's equations can be written

$$G_{ij} = -\kappa T_{ij} \,, \tag{2.111}$$

where

$$G_{ij} := R_{ij} - \frac{1}{2} g_{ij} R \tag{2.112}$$

is the Einstein tensor expressed in terms of the Ricci tensor R_{ij} and curvature scalar R, κ is a constant that turns out to be expressible as

$$\kappa = \frac{8\pi G}{c^4} \,, \tag{2.113}$$

and T_{ij} is the energy–momentum tensor expressing the distribution of mass and energy in the spacetime.

Now the covariant derivative of the Einstein tensor is zero in many circumstances, in particular when something called the torsion is zero. But the torsion is indeed often zero. In fact, it is sourced by the spin currents of matter in such a way that, in contrast to curvature, it does not propagate in spacetime, so it could only be nonzero in regions where there is matter or energy with some rotational property. A very clear, though somewhat sophisticated account of all this can be found in [20, Chap. 5]. Anyway, in a region where there is no spinning matter, Einstein's equation (2.111) implies that the covariant derivative of the energy–momentum tensor is zero. This is what we shall use to derive the geodesic 'principle'.

To do this, we shall consider an almost-pointlike particle. So when almost-point particles are not acted upon by forces (apart from gravitational effects), we would like to show that their trajectories take the form

$$\frac{\mathrm{d}^2 x^m}{\mathrm{d}s^2} + \Gamma^m_{kj} \frac{\mathrm{d}x^k}{\mathrm{d}s} \frac{\mathrm{d}x^j}{\mathrm{d}s} = 0 \,. \tag{2.114}$$

We consider a small blob of dustlike (i.e., zero pressure) matter with density ρ and velocity field

$$v^i = \frac{\mathrm{d}x^i}{\mathrm{d}s} \,. \tag{2.115}$$

This equation expresses the fact that we view each component dust particle as having its own path $x^i(s)$. The energy–momentum tensor for this matter is then

$$T^{ik} = \rho \frac{\mathrm{d}x^i}{\mathrm{d}s}\frac{\mathrm{d}x^k}{\mathrm{d}s} \tag{2.116}$$

and we are saying that Einstein's field equation (2.111) implies that

$$T^{ik}{}_{;k} = 0 \,. \tag{2.117}$$

We analyse (2.117) by inserting (2.116) and the result is the geodesic equation (2.95). For completeness, here is the argument. We have

$$\rho_{,k}\frac{\mathrm{d}x^i}{\mathrm{d}s}\frac{\mathrm{d}x^k}{\mathrm{d}s} + \rho\left(\frac{\partial}{\partial x^k}\frac{\mathrm{d}x^i}{\mathrm{d}s} + \Gamma^i_{km}\frac{\mathrm{d}x^m}{\mathrm{d}s}\right)\frac{\mathrm{d}x^k}{\mathrm{d}s} + \rho\frac{\mathrm{d}x^i}{\mathrm{d}s}\left(\frac{\partial}{\partial x^k}\frac{\mathrm{d}x^k}{\mathrm{d}s} + \Gamma^k_{km}\frac{\mathrm{d}x^m}{\mathrm{d}s}\right) = 0 \,. \tag{2.118}$$

If we did not have the idea of a velocity field v^i, it would be difficult to interpret partial derivatives of $\mathrm{d}x^i/\mathrm{d}s$ with respect to the coordinates. But as things are, we can say

$$\frac{\mathrm{d}x^k}{\mathrm{d}s}\frac{\partial}{\partial x^k}\frac{\mathrm{d}x^i}{\mathrm{d}s} = \frac{\mathrm{d}x^k}{\mathrm{d}s}\frac{\partial v^i}{\partial x^k} = \frac{\mathrm{d}v^i}{\mathrm{d}s} = \frac{\mathrm{d}^2x^i}{\mathrm{d}s^2} \,. \tag{2.119}$$

The terms in the second bracket of (2.118) are

$$\frac{\partial v^k}{\partial x^k} + \Gamma^k_{km}v^m = \operatorname{div} v \,, \tag{2.120}$$

and the whole thing can now be expressed by

$$\operatorname{div}(\rho v)\frac{\mathrm{d}x^i}{\mathrm{d}s} + \rho\left(\frac{\mathrm{d}^2x^i}{\mathrm{d}s^2} + \Gamma^i_{km}\frac{\mathrm{d}x^k}{\mathrm{d}s}\frac{\mathrm{d}x^m}{\mathrm{d}s}\right) = 0 \,. \tag{2.121}$$

By mass conservation,

$$\operatorname{div}(\rho v) = 0 \,, \tag{2.122}$$

and the result follows.

This proof purports to show that each constitutive particle of the blob follows a geodesic. But then we did not allow these particles to jostle one another. For example, we have zero pressure, as attested by the form of the energy–momentum tensor in (2.116). And we did not allow the particles to generate any torsion by revolving about the center of energy of the blob. And neither did we endow them with electric charge. It is in this sense that the geodesic 'principle' is in fact just an approximation, unless the test particle is not a blob, but a mathematical point.

Some argue that inertia is explained in general relativity, precisely because of the above proof (or better variants of it). This point of view is expressed in the philosophical study by Brown [21, p. 141]:

> GR is the first in the long line of dynamical theories, based on that profound Aristotelian distinction between natural and forced motions of bodies, that *explains* inertial motion.

This is not the view taken here, for reasons to be explained shortly. However, other issues discussed in Brown's book, in particular what he refers to as the dynamical approach to spacetime structure, should mark a turning point in our understanding of relativity theories that is exactly in line with the approach advocated in the present book. Another quote concerning the explanation of inertial motion is [21, Sect. 9.3]:

> Inertia, in GR, is just as much a consequence of the field equations as gravitational waves. For the first time since Aristotle introduced the fundamental distinction between natural and forced motions, inertial motion is part of the dynamics. It is no longer a miracle.

Here is an argument against that view. Recall the discussion just after (2.102) on p. 37. It was pointed out that actions like (2.99) and (2.101) are designed to give appropriate field equations, and that the appropriate field equation for the particle labelled by a is a geodesic equation, or an equation like (2.102). Now in GR, one adds a gravitational part to the action, viz.,

$$\mathscr{A}_{\mathrm{grav}} := \frac{c^3}{16\pi G}\int R(-g)^{1/2}\mathrm{d}^4x\,. \tag{2.123}$$

Some textbooks motivate this as follows. When the metric is varied in $\mathscr{A}_{\mathrm{grav}}$, a constant multiple of the Einstein tensor pops out. The point about this is the observation that, when the metric is varied in an action like (2.101), the energy–momentum tensor T_{ij} pops out. One gets a sum of contributions to this tensor from the matter as encapsulated in the action term

$$-\sum_a cm_a\int \mathrm{d}s_a\,,$$

and from the EM fields as encapsulated in the action term

$$-\frac{1}{16\pi c}\int F_{ik}F^{ik}(-g)^{1/2}\mathrm{d}^4x\,.$$

Setting the variation of the full action with respect to the metric equal to zero, one then obtains the Einstein equations, with the Einstein tensor on one side and the total energy–momentum on the other side.

Now the covariant derivative of the Einstein tensor is zero (assuming zero torsion) and this could in principle be worrying, because the Einstein equation then implies that the covariant derivative of the total energy–momentum is zero. However, there is a general result that the energy–momentum tensor derived from an action of the form

$$\int L(-g)^{1/2} \mathrm{d}^4 x \tag{2.124}$$

by varying the metric always has zero covariant derivative when L is a scalar, and more sophisticated versions, e.g., invariance of the matter action under the group of diffeomorphisms is sufficient to guarantee zero covariant derivative of the corresponding energy–momentum tensor on shell if the torsion is zero [20, Sect. 6.5]. (As mentioned above, if the torsion is not zero, the covariant derivative of the Einstein tensor is not zero either. This case is not considered here.) Of course, the action $\mathscr{A}$ in (2.99) does not have the form (2.124) but one expects some general theorem to ensure that the resulting energy–momentum tensor will have zero covariant derivative on shell, i.e., when the field equations, that is, the geodesic equations, are satisfied.

So it looks as though the geodesic equations, and their variants with a force on one side, are built in by construction of the action. It is no surprise therefore that they should pop out again when we set the covariant derivative of the energy–momentum tensor equal to zero. Perhaps one should be more suspicious of arguments from actions. They are neat, and bring a level of unity in the sense that one can derive several dynamical equations from the same action by varying different items. On the other hand, we are only getting out what we put in somewhere else.

One way to make progress with explaining inertia might be to pay more attention to the fact that test particles are not likely to be well modelled by mathematical points. We have already mentioned the spinning particle and the effect of curvature on its motion. A possibly different effect (although possibly similar?) occurs when the particle is a source of some classical force field, typically electromagnetic. The spatially extended particle then exerts forces on itself and in simple cases it can be shown that these forces oppose acceleration in flat spacetime and explain why a force is needed to keep the particle off a geodesic in curved spacetime. If all inertia were due to these self-forces, the geodesic equation, or relevant extension of $F = ma$ would then be replaced by an equation of the form $\sum F = 0$, where the F summed over include self-forces.

Springer
http://www.springer.com/978-3-540-68469-5

Uniformly Accelerating Charged Particles
A Threat to the Equivalence Principle
Lyle, S.
2008, XV, 361 p., Hardcover
ISBN: 978-3-540-68469-5