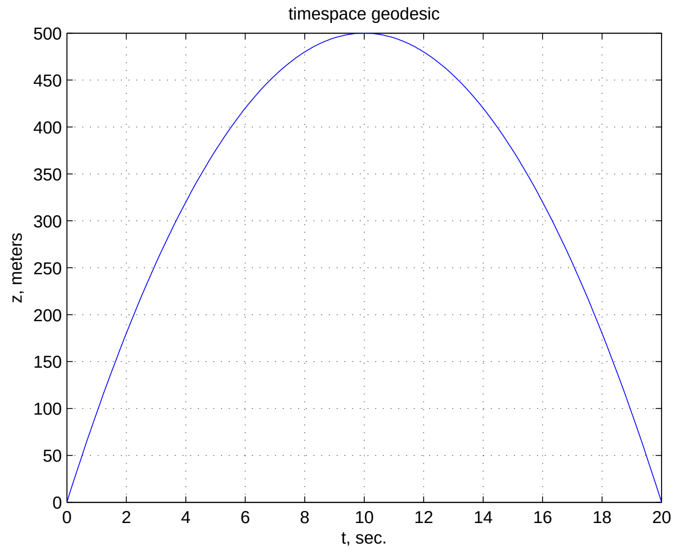
Excursions in Computing Science: Book 11c. Geometry and Gravity Part II. Gravity

T. H. Merrett*
McGill University, Montreal, Canada

May 24, 2022

I. Prefatory Notes

22. Curved timespace. Some military contemporaries of Isaac Newton decided to test a cannon without consulting him. To see how high it could shoot, they pointed it straight up and fired.



Twenty seconds later they had no cannon left, so they consulted.

Newton calculated the height $z(t)$ given the acceleration g due to gravity (constant near the Earth's surface, and downwards so we'll need a minus sign)

$$z(t) = \text{antislope}_t v(t)$$

where upwards velocity

$$\begin{aligned} v(t) &= \text{antislope}_t(-g) \\ &= v_0 - gt \end{aligned}$$

where v_0 is the (muzzle) velocity at time $t = 0$, so

$$z(t) = z_0 + v_0 t - \frac{1}{2} g t^2$$

and we can suppose the cannon is located at height 0 so that $z_0 = 0$.

For the ball to fly for 20 seconds, i.e., return to height 0 after 20 seconds,

$$0 = 20 v_0 - 400 g/2$$

making

$$v_0 = 10g = 100\text{m/s}$$

since $g = 10$ m/s (or close enough for mental calculation).

The maximum height happens when the slope of the curve is zero

$$\begin{aligned} 0 &= \text{slope}_t \left(v_0 t - \frac{1}{2} g t^2 \right) \\ &= v_0 - gt \end{aligned}$$

so that

$$t = v_0/g = 10 \text{ s}$$

and

$$\begin{aligned} z &= 10v_0 - 100g/2 \\ &= 500 \text{ m} \end{aligned}$$

Now if the parabola's path, as shown in the figure above, were a geodesic in the t-z timespace, that timespace would have to be curved. Or so it seems: we'll have to be careful.

Fortunately the factory had an even better cannon ready with muzzle velocity $\sqrt{2}$ higher. The soldiers tried to see how *far* the new cannon would shoot. Newton told them to raise it to an elevation of 45^o because with components $v_0 c = v_0 \cos\theta$ in the horizontal direction x and $v_0 s = v_0 \sin\theta$ in the vertical direction z he got two equations

$$\begin{aligned} x &= v_0 c t \\ z &= v_0 s t - \frac{1}{2} g t^2 \\ &= \frac{s}{c} x - \frac{g}{2} \frac{x^2}{v_0^2 c^2} \end{aligned}$$

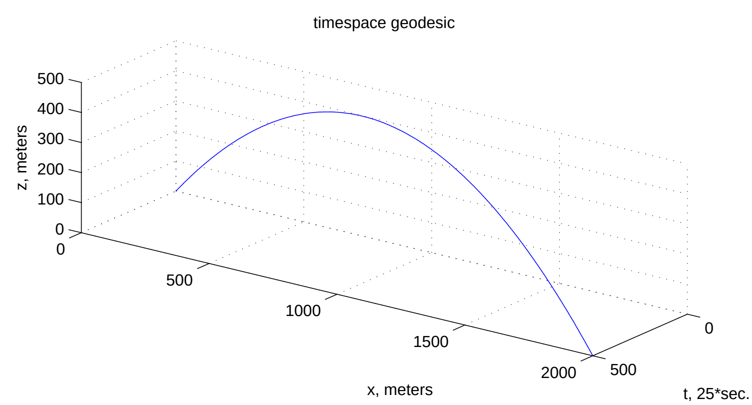
and $z = 0$ when $x = 0$ or $x = (2v_0 cs)/g = (v_0/g)\sin 2\theta$.

These two points at ground level are furthest apart when $\sin 2\theta$ is max, i.e., has zero slope

$$0 = \text{slope}_\theta \sin 2\theta = 2\cos 2\theta$$

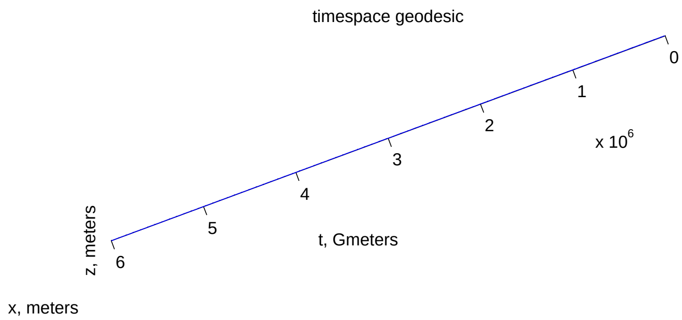
which happens when $2\theta = 90^o$.

So $v_0 = 100\sqrt{2}$ and x- and z-initial velocities are each 100 m/s, $v_x = 100 = v_z$. In 20 s the ball travelled 2000 m before landing.



Unfortunately, to make MATLAB show the time axis on a scale comparable to the x and z axes I had to multiply t by 25: what shows as 500 is actually, still, 20 sec.

Of course, really we should scale t by c, lightspeed, for all axes to be compared



We see that the curvature really is very small.

Just then the factory produced a new cannon more than twice as powerful as the original one. It could shoot with a muzzle velocity of 200 m/s horizontally plus up to 50 m/s vertically, $v_x \leq 200, v_z \leq 50$.

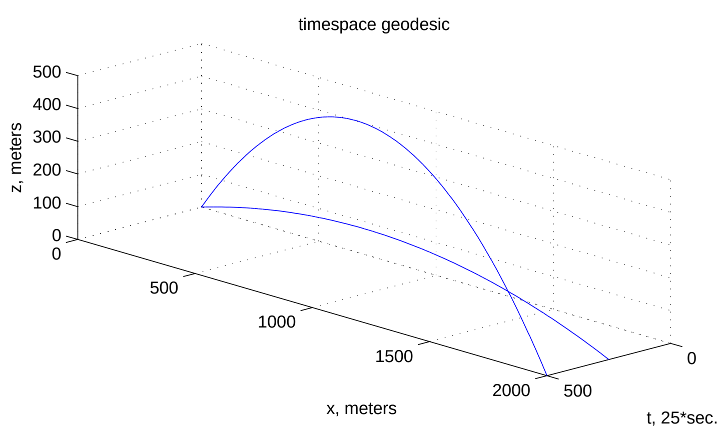
$$\begin{aligned} x &= 200t \\ z &= v_z t - \frac{1}{2}gt^2 \end{aligned}$$

so at $t = 10$s, the time needed to travel the 2000m shot by the $\sqrt{2}$ cannon, we need the full 50 m/s:

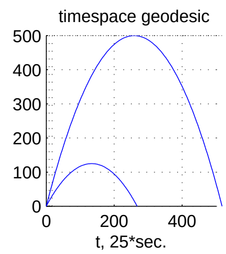
$$\begin{aligned} 0 &= 10v_z - \frac{100}{2}g \\ v_z &= 50 \text{ m/s} \end{aligned}$$

That is,

$$z = 50t - \frac{1}{2}gt^2$$



These two parabolas look very different from one another. But see what happens when we rotate out the x-axis and look at a time sequence as the cannon balls fly directly at us.



The tops of two parabolas now appear to match up.

We can show they are exactly the same by starting a new clock t' for the faster ball when $t = 5$ sec. for the slower ball.

$$t' = t - 5 \qquad\qquad t = t' + 5$$

$$\begin{aligned} z &= 100t - \frac{1}{2}gt^2 \\ &= 100t' + 500 - \frac{1}{2}10(t' - 5)^2 \\ &= 375 + 50t' - \frac{1}{2}10t'^2 \end{aligned}$$

Apart from the 375m, which is the height attained by the slower ball after 5 sec., this equation is identical to the parabola for the faster ball

$$z = 50t - \frac{1}{2}gt^2$$

So the curvature for both balls is somehow the same, independently of their speeds.

This suggests that it is valid to associate the gravitational field g with a curvature of timespace—the same curvature for all motions caused by the gravity.

And we've seen that the gravitational field g is very weak: when we measure time in meters (well, gigameters) the curvature is insignificant. But it is enough to return the cannonballs to Earth.

23. Gravitational redshift. To understand better this curvature of timespace, let's focus on time. We can show that a particle climbing up against gravity must experience a slowing of frequency (see Week 5 "Particles with periods").

From the point of view of light, this would be a reddening, but we'll work with a particle of mass m as well as of angular frequency ω.

Such a particle has energy

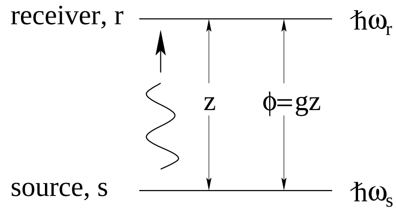
$$mc^2 = \hbar\omega$$

combining special relativity (Week 7a) and quantum ideas (Week 5).

The difference in energy due to the gravitational field g depends on the difference in height, giving a difference gz of gravitational potential $\phi = gz$, and an energy difference mgz, or

$$m\phi$$

Now consider a source s of particles and a receiver r placed above s.



For energy to be conserved, the particle must have two different frequencies at source ω_{s} and receiver ω_{r}

$$mc^2 = \hbar\omega_{\mathrm{s}} = \hbar\omega_{\mathrm{r}} + m\phi$$

So

$$\begin{aligned}
\omega_{\mathrm{r}} &= \omega_{\mathrm{s}} - \frac{m\phi}{\hbar} \\
\Delta\omega &= \omega_{\mathrm{r}} - \omega_{\mathrm{s}} \\
\frac{\Delta\omega}{\omega_{\mathrm{s}}} &= -\frac{m\phi}{\hbar mc^2/\hbar} \\
&= -\frac{\phi}{c^2}
\end{aligned}$$

So there will be fewer cycles per second when we're higher in the gravitational field. This is the same as saying that clocks higher in the gravitational field run *faster*

$$\frac{\Delta\tau}{\tau} = \frac{\phi}{c^2}$$

or

$$\tau_{\mathrm{r}} = \left(1 + \frac{\phi}{c^2}\right)\tau_{\mathrm{s}}$$

This is the *gravitational* time dilation, and a term for this must appear in the $(\Delta t)^2$ part of the metric.

For the time dimension alone, the small interval of *proper* time

$$(\Delta\tau)^2 = (\Delta s)^2 = -g_{00}(\Delta t)^2$$

where the minus sign reflects the fact that a boost (Lorentz transformation in special relativity—see Week III) is a shear and has $(\Delta r)^2 - (\Delta t)^2$ as the invariant.

Let's try for this element of the metric tensor (the 2 comes from $(1+a)^2 \approx 1 + 2a$)

$$g_{00} = -\left(1 + \frac{2\phi}{c^2}\right)$$

Since

$$\begin{aligned}
\omega &\propto \frac{1}{\Delta\tau} \\
\frac{\omega_{\mathrm{r}}}{\omega_{\mathrm{s}}} = \frac{\Delta\tau_{\mathrm{s}}}{\Delta\tau_{\mathrm{r}}} &= \frac{\left(1 + 2\phi_{\mathrm{s}}/c^2\right)^{1/2}\Delta t_{\mathrm{s}}}{\left(1 + 2\phi_{\mathrm{r}}/c^2\right)^{1/2}\Delta t_{\mathrm{r}}}
\end{aligned}$$

We've been writing the potential as the general term ϕ—it could have been gz throughout—but now we'll make an assumption: that the field is *static*, i.e., does not change in time.

Then you should be able to convince yourself that two light beams travelling from one place to another, e.g., from source to receiver, will arrive at the beginning and end of an interval with the same duration as the interval separating their departures: from the point of view of an observer at rest with respect to both source and receiver

$$\Delta t_{\mathrm{r}} = \Delta t_{\mathrm{s}}$$

So

$$\begin{aligned}
\frac{\omega_{\mathrm{r}}}{\omega_{\mathrm{s}}} &= \frac{\left(1+\frac{2\phi_{\mathrm{s}}}{c^2}\right)^{1/2}}{\left(1+\frac{2\phi_{\mathrm{r}}}{c^2}\right)^{1/2}} \\
&\approx \left(1+2\frac{\phi_{\mathrm{s}}-\phi_{\mathrm{r}}}{c^2}\right)^{1/2} \\
&= \left(1-2\frac{\phi}{c^2}\right)^{1/2} \\
&\approx 1-\frac{\phi}{c^2}
\end{aligned}$$

which is what we had.

So we have one of the terms in our metric for curved timespace:

$$g_{00} = -\left(1+\frac{2\phi}{c^2}\right)$$

Let's now specialize to a constant gravitational field, $\phi = gz$, and let's return to the useful convention that $c = 1$ in our units of measuring time and space

$$g_{00} = -(1+2gz)$$

And I'm going to take a guess that the other term in our 2-D t-z metric is

$$g_{11} = 1/(1+2gz)$$

This guess is based on Wheeler's result in the Excursion *Paraboloids* in Part I

$$g_{11} = 1/(1-a/r)$$

and the observation that, for a spherically symmetrical space around a source of mass M, the gravitational potential is

$$\phi = -\frac{M}{r}$$

so it is plausible to replace $-a/r$ by $2gz$ for uniform gravity g and height z.

We have, in t-z space,

$$(\Delta s)^2 = -(1+2gz)(\Delta t)^2 + (\Delta z)^2/(1+2gz)$$

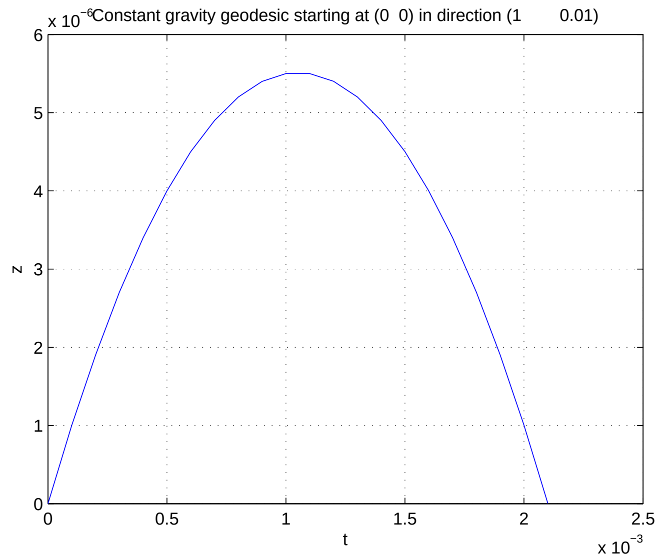
or

$$g_{\mathrm{dd}} = \begin{pmatrix} -(1+2gz) & \\ & 1/(1+2gz) \end{pmatrix}$$

From this, running `metr2conn()` from the protor calculator (Note 20, Part I), we get

$$
\begin{array}{llll}
\Gamma_{\mathrm{udd}}(\alpha & \beta & \gamma & \qquad\Gamma \qquad) \\
t & t & z & g/(1+2gz) \\
t & z & t & g/(1+2gz) \\
z & t & t & g(1+2gz) \\
z & z & z & -g/(1+2gz)
\end{array}
$$

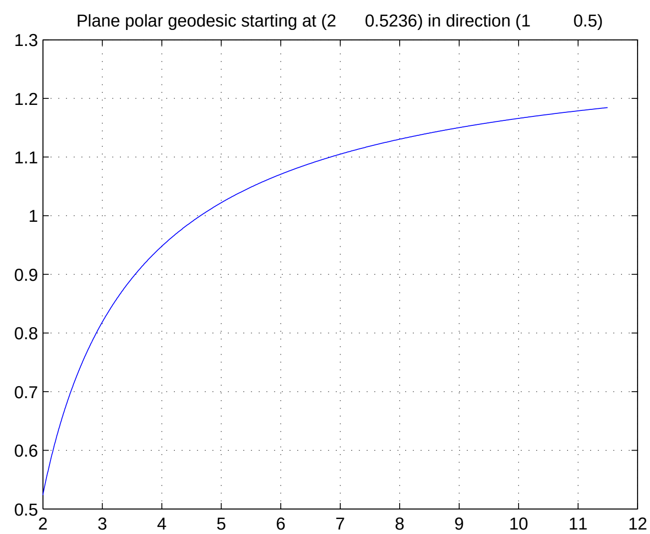
and, by adapting the MATLAB function `geodesicPlanePolar()` of Note 13, Part I to, say, `geodesicConstGrav()`, we get a nice parabolic trajectory



Despite this geodesic, the t-z space is *not* curved. Running `metr2curv()` on g_{dd} and its inverse g_{uu} gives zero curvature.

This is the way it should be. If the timespace were curved we would get tides, but these do not occur under constant gravitation.

It may be helpful to revisit the plane polar geodesic of Note 13, which is also in flat space but using a Cartesian plot instead of a polar plot.



We see that space need not be curved to support a curved geodesic.

24. Spherically symmetric gravity. We can combine the time component of the gravitational metric from Note 23

$$g_{00} = -\left(1 + \frac{2\phi}{c^2}\right)$$

with the space component of the metric for the paraboloid of rotation (Excursion *Paraboloid* of Part I)

$$g_{\mathrm{dd}} = \begin{pmatrix} \frac{1}{1-a/r} & & \\ & r^2 & \\ & & r^2 s^2 \end{pmatrix}$$

for a 3D spherically symmetric gravitational potential

$$\phi = -\frac{M}{r}$$

where M is the mass of the source of gravity at $r = 0$. (And $s = \sin\theta$.)

If we work with coordinates such that $c = 1$ it is plausible to set $a = 2M$ so that the full timespace metric is

$$g_{\mathrm{dd}} = \begin{pmatrix} \left(1 - \frac{2M}{r}\right) & & & \\ & \frac{1}{1-2M/r} & & \\ & & r^2 & \\ & & & r^2 s^2 \end{pmatrix}$$

Running this through `metr2curv()` in the protor calculator gives the affine connection

$\Gamma_{\mathrm{udd}}(\alpha$	β	γ	Γ $)$
t	t	r	$(M/r^2)/(1-2M/r)$
t	r	t	$(M/r^2)/(1-2M/r)$
r	t	t	$(M/r^2)(1-2M/r)$
r	r	r	$-(M/r^2)/(1-2M/r)$
r	θ	θ	$-r(1-2M/r)$
r	ϕ	ϕ	$-s^2(1-2M/r)$
θ	r	θ	$1/r$
θ	θ	r	$1/r$
θ	ϕ	ϕ	$-cs$
ϕ	ϕ	r	$1/r$
ϕ	r	ϕ	$1/r$
ϕ	ϕ	θ	c/s
ϕ	θ	ϕ	c/s

where $c = \cos\theta$ and $s = \sin\theta$ as usual, and the Riemann curvature (I've omitted the rows that can be obtained by symmetry)

$R_{\mathrm{dddd}}(\alpha$	β	γ	δ	R $)$
t	θ	t	θ	$(M/r)(1-2M/r)$
t	ϕ	t	ϕ	$(M/r)s^2(1-2M/r)$
r	t	r	t	$-2M/r^3$
r	θ	r	θ	$-(M/r)/(1-2M/r)$
r	ϕ	r	ϕ	$-(M/r)s^2(1-2M/r)$
θ	ϕ	θ	ϕ	$2Mrs^2$

The uddd form of the curvature is also given but is much longer and not worth writing out here. But its contraction, `contract(curv_uddd,1,3)` is zero for every term.

$$\begin{array}{cccc} R_{\rm dd}(\alpha & \beta & R &) \\ & t & t & 0 \\ & r & r & 0 \\ & \theta & \theta & 0 \\ & \phi & \phi & 0 \end{array}$$

and this is significant. It is the timespace form of the zero average curvature we found in Excursion *Paraboloid*.

$$R_{\rm dd} = 0$$

can be taken as the simplest statement of the law of gravity in Einstein's general relativity:

In the absence of mass, the average curvature is zero.

Compare this with the "free space" results of Note 7, Part I, on classical gravity. There the divergence of the gradient (the second slope) of the potential is zero. Here the contraction of the curvature (the second slope of the metric) is zero.

To simulate this gravitation we can plot, say, the orbit of Mercury around the Sun. We must revise the geodesic function yet again to one named after Karl Schwarzschild who produced it as a solution to Einstein's equations six months after Einstein published them:

```
geodesicSchwarzschild(direc,starte,n,step,M)
```

with the mass M of the Sun as a parameter.

To make M/r dimensionless (instead of having physical dimensions M/L) we must figure out how to express mass as a length. The missing constants are Newton's G_N and lightspeed c. The physical dimensions of c are of course L/T. Those of G_N are $L^3/(MT^2)$ since (see Note 7 of Part I) acceleration

$$a = -G_N M/r^2$$

or, dimensionally

$$\frac{L}{T^2} = \frac{G_N M}{L^2}$$

Thus the physical dimensions of G_N/c^2 are L/M so this is the conversion constant. It works out to be (see excursion *Newton's constant* in Part I)

$$\frac{G_N}{c^2} = 734 \times 10^{-30} \frac{\rm m}{\rm kg}$$

and since the solar mass $M = 2 \times 10^{30}$ kg we have

$$M = 1.47\mathrm{Km}$$

We'll use gigameters as units of the simulation, since Mercury's orbit ranges from 46 Gm (perihelion) to 70 Gm (aphelion), so $M = 1.47 \times 10^{-6}$

The `start` vector has four components

$$\begin{array}{rcl} t &=& 0 \\ r &=& 70 \\ \theta &=& \pi/2 \\ \phi &=& 0 \end{array}$$

where we choose θ to be in the equatorial plane—and we'll find that it stays there.

The initial direction `direc` is also a four-vector

$$\begin{aligned}\Delta t &= 1\\ \Delta r &= 0\\ \Delta\theta &= 0\\ \Delta\phi &= 0.00000185\end{aligned}$$

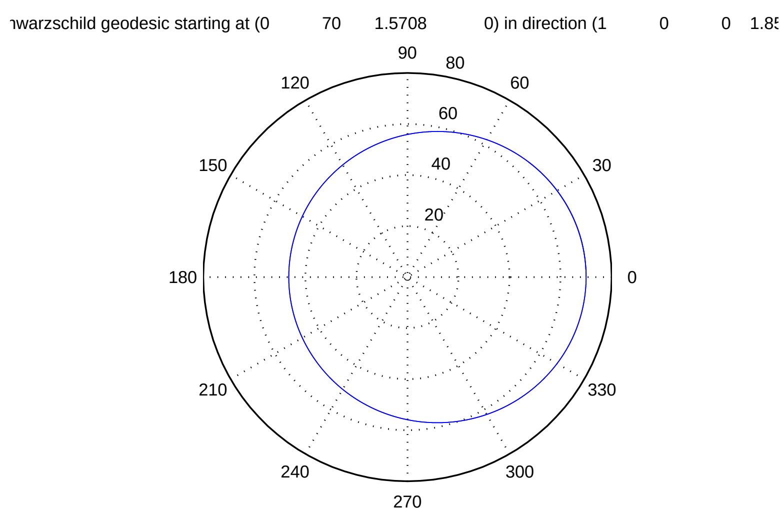
so that the initial planetary motion is purely tangential in the $\theta = \pi/2$ plane, and the initial velocity must be both slow enough and fast enough to keep it in orbit. I've tuned the angular velocity $\Delta\phi\Delta t$ so that the orbit, starting at aphelion, has the right perihelion of 46Gm.

The supporting function `geodSchwarzschildStep(x,origin,M)` uses the affine connection in a matrix which parallel-transports the 4-vector $\mathbf{x}$. With $m = M/r^2$ and $a = 1 - 2M/r$, and of course $s = \sin\theta$ and $c = \cos\theta$, this matrix is

$$\begin{pmatrix} m\Delta r/a & m\Delta t/a & & \\ ma\Delta t & -m\Delta r/a & -ra\Delta\theta & -rs^2a\Delta\phi \\ & \Delta\theta/r & \Delta r/r & -cs\Delta\phi \\ & \Delta\phi/r & c\Delta\phi/s & \Delta r/r + c\Delta\theta/s \end{pmatrix}$$

Here's the result of

```
ends = geodesicSchwarzschild(direc,start,3*10^4,100,1.47*10^-6)
```



The plot for this simulation is not fine enough to show the tiny precession of Mercury's orbit that astronomers had observed but had not been able to explain by Newton's theory. Newton's elliptical orbit now, under general relativity, appears to rotate around the Sun, with the angle of its major axis changing by only 43 seconds of arc (there are 3600 seconds in a degree) per *century*.

This metric exactly calculates the anomaly in Mercury's orbit. It also predicts that light from distant stars will bend in the gravitational field of the Sun—or of a galaxy for that matter ("gravitational lensing").

The bending of light is a prediction which distinguishes general relativity from an elegant theory proposed in 1914, the year before Einstein published, by Gunnar Nordström, which added a fifth dimension to timespace in order to accommodate gravity in addition to Maxwell's electromagnetism.

In 1919, as soon as World War I was over, the British astrophysicist Arthur Stanley Eddington led an expedition to observe a solar eclipse in the southern hemisphere, and to confirm the bending of the light from a star directly behind the Sun.

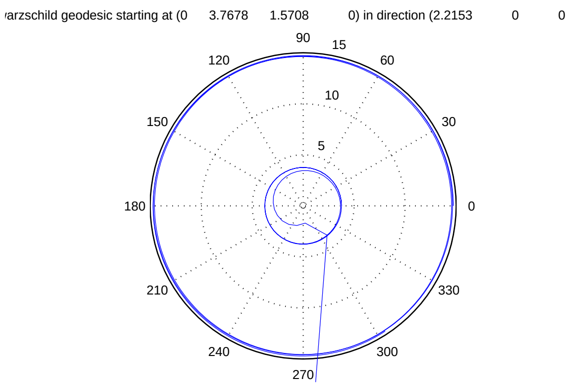
25. Schwarzschild orbits and black holes. We can discover four basic types of orbit in the Schwarzschild metric if we pick our parameters carefully.

The parameters in the following are starting 4-position $x^\alpha = t, r, \theta, \phi$ and starting 4-velocity $u^\alpha = \partial_\tau x^\alpha$ where τ is the relativistic proper time $(\Delta\tau)^2 = (\Delta t)^2 - (\Delta r)^2 - (r\Delta\theta)^2 - (rs\Delta\phi)^2$, and M.

At the outset we'll use $M = 1$ and we'll stick to the equatorial plane $\theta = \pi/2$ and $u^\theta = 0$. We'll always start at $t = 0$ and $\phi = 0$.

For the two types of bound orbit, we'll start with $u^r = 0$.

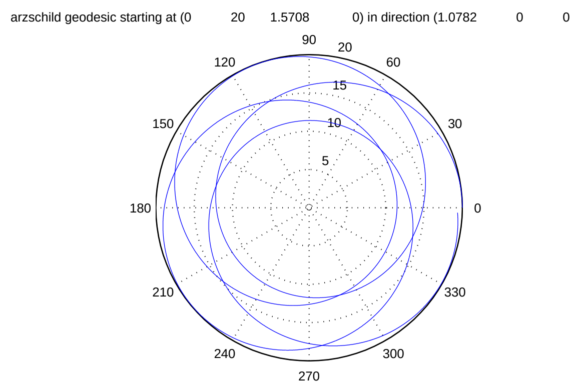
Circular orbits



	r	u^t	u^ϕ
outer	14.7222	1.1207	0.0198
inner	3.7678	2.2153	0.3029

```
geodesicSchwarzschild(Uu,Xu,1000,1,1)
```

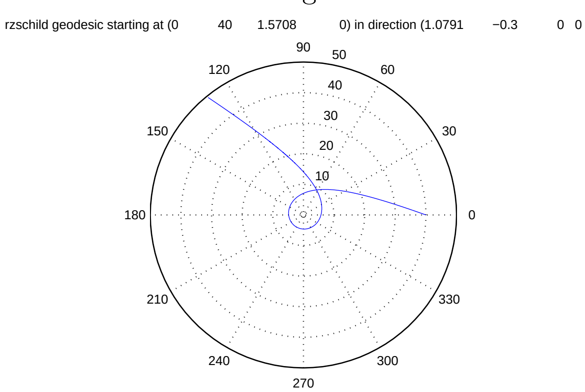
Precessing orbits



r	u^t	u^ϕ
20	1.0982	0.0107

```
geodesicSchwarzschild(Uu,Xu,15000,0.1,1)
```

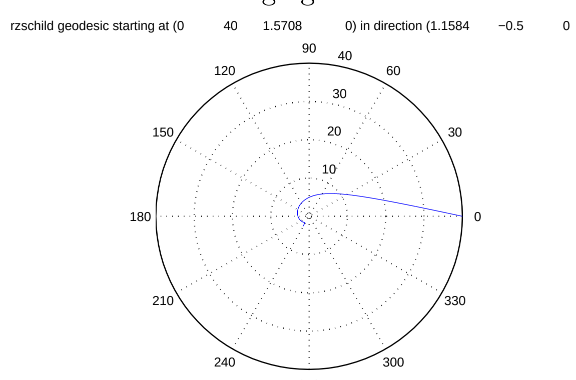
Scattering orbits



r	u^r	u^t	u^ϕ
40	-0.3	1.0791	0.0027

```
geodesicSchwarzschild(Uu,Xu,3000,0.1,1)
```

Plunging orbits



r	u^r	u^t	u^ϕ
40	-0.5	1.1584	0.0027

```
geodesicSchwarzschild(Uu,Xu,8787,0.01,1)
```

These are all quite different from Newtonian orbits, which are elliptical (including circular) if bound and hyperbolic if unbound—with a parabolic orbit in between (as in the case of a test body starting at rest at $r = \infty$).

The circular orbits are Newtonian, but there are only two of them and the inner one is unstable as is apparent from the picture of the simulation (the sudden zig off to infinity is an artefact of the calculation: the body has hit the "singularity" that we'll discuss at the end of this Note).

The precessing orbits alternate between two radii and could be approximated by ellipses with the Sun at a focus, except for the precession. The picture shows a greatly exaggerated version of Mercury's precession.

The scattering orbit does not look at all like the Newtonian parabola or hyperbola and is highly relativistic.

The plunging orbit is impossible under Newton. The Newtonian potential $-M/r$ if augmented by the effect of angular momentum ℓ becomes $-M/r + \ell^2/(2r^2)$ and is positive-infinite at $r = 0$. So it repels any falling body with nonzero ℓ. But relativistic effects overcome this and result in a potential barrier of only finite height.

By examining the Schwarzschild metric in terms of potential and kinetic energy we can understand all these orbits qualitatively and also arrive at the quantitative parameters I used for the simulation by `geodesicSchwarzschild()`.

We must work with 4-velocities and we must know two preliminary things about them.

First,

$$\mathbf{u} \cdot \mathbf{u} = g_{\alpha\beta} u^\alpha u^\beta = -1$$

as we can see by examining a coordinate system in which the body is at rest. Then $u^j = 0, j = r, \theta, \phi$ (or $j = x, y, z$) and only $u^t = \partial_\tau t$ is nonzero. But in this frame $\Delta\tau = \Delta t$ so $u^t = 1$. Since $g_{tt} = -1$ we have $\mathbf{u} \cdot \mathbf{u} = -1$. Because this is a tensor equation, it is true in any coordinate system.

Second, symmetries of the system give rise to conservation laws, just as we saw in Note 38 of Book 8c. The way to do this in relativity is by means of *symmetry vectors* (also named after Wilhelm Killing).

The symmetry vectors for the Schwarzschild metric are

$$\xi_{\mathrm{u}} = (1, 0, 0, 0) \qquad \text{and} \qquad \eta_{\mathrm{u}} = (0, 0, 0, 1)$$

respectively because the metric is independent of t and of ϕ.

The conserved quantities are given by the dot product of the symmetry vector and the 4-velocity as it travels along a geodesic, i.e., is in "free fall". For the Schwarzschild metric these are worth giving special names to because they are related to energy (conserved because of the time-independence) and angular momentum (conserved because of the ϕ-independence).

So we have

$$\begin{aligned} e &= -\xi \cdot \mathbf{u} = \left(1 - \frac{2M}{r}\right) u^t \\ \ell &= \eta \cdot \mathbf{u} = r^2 s^2 u^\phi \end{aligned}$$

We can now use these in

$$\begin{aligned} -1 &= \mathbf{u} \cdot \mathbf{u} \\ &= -\left(1 - \frac{2M}{r}\right)(u^t)^2 + \left(1 - \frac{2M}{r}\right)^{-1} (u^r)^2 + (ru^\phi)^2 \\ &= -\left(1 - \frac{2M}{r}\right)^{-1} e^2 + \left(1 - \frac{2M}{r}\right)^{-1} (u^r)^2 + \frac{\ell^2}{r^2} \end{aligned}$$

where the second line incorporates our assumptions $\theta = \pi/2$ and $u^\theta = 0$—which we can see as a consequence of the conservation of angular momentum.

To compare with Newton's effective energy (per unit mass of the test body)

$$\frac{1}{2}(u^r)^2 + V_N$$

where

$$V_N = -\frac{M}{r} + \frac{\ell^2}{2r^2}$$

we need to isolate e^2

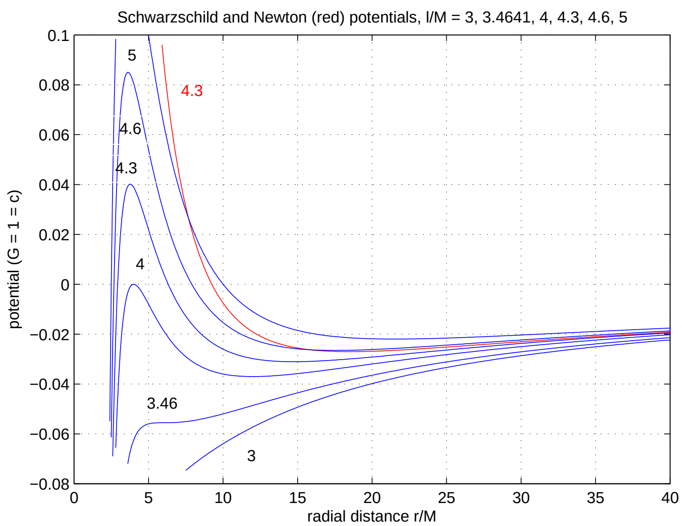
$$\begin{aligned} e^2 &= (u^r)^2 + \left(1 + \frac{\ell^2}{r^2}\right)\left(1 - \frac{2M}{r}\right) \\ &= (u^r)^2 + 1 - \frac{2M}{r} + \frac{\ell^2}{r^2} - \frac{2M\ell^2}{r^3} \end{aligned}$$

So the effective potential is the last three terms of

$$\frac{1}{2}(e^2 - 1) = \frac{1}{2}(u^r)^2 - \frac{M}{r} + \frac{\ell^2}{2r^2} - \frac{M\ell^2}{r^3}$$

and note that, compared with Newton, above, there is an extra term, in r^{-3}.

We can plot this for various values of ℓ/M and compare it with Newton (in red).



We see the finite central barrier and the ℓ/M dependence of the maxima and minima.

The effective potential goes to 0 when r goes to infinity and when

$$0 = Mr^2 - \ell^2 r/2 + M\ell^2$$

so

$$r_0 = \frac{\ell^2}{4M}\left(1 \pm \sqrt{1 - 16\left(\frac{M}{\ell}\right)^2}\right)$$

which become a single value at $\ell/M = 4$.

The minimum and maximum are at

$$\begin{aligned} 0 &= \partial_r\left(-\frac{M}{r} + \frac{\ell^2}{2r^2} - \frac{M\ell^2}{r^3}\right) \\ &= \frac{M}{r^2} - \frac{\ell^2}{r^3} + \frac{3M\ell^2}{r^4} \end{aligned}$$

i.e.,

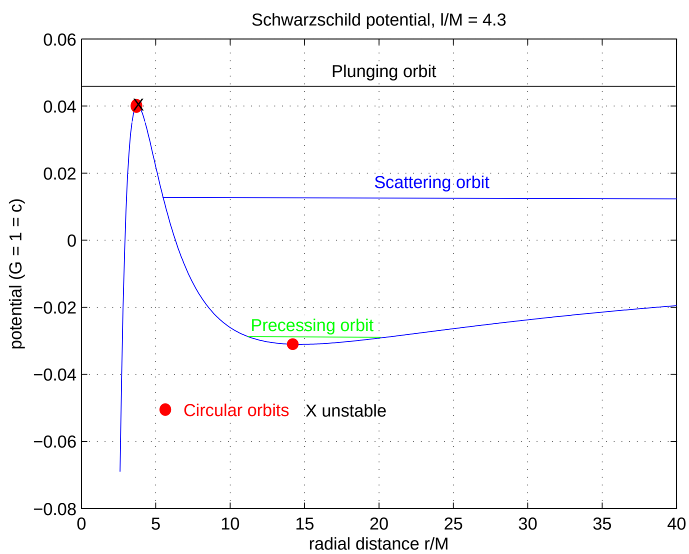
$$0 = Mr^2 - \ell^2 r + 3M\ell^2$$

so

$$r_\pm = \frac{\ell^2}{2M}\left(1 \pm \sqrt{1 - 12\left(\frac{M}{\ell}\right)^2}\right)$$

which becomes a single value at $\ell/M = \sqrt{12} = 3.46$.

The plot suggests that $\ell/M = 4.3$ is a good choice to explore by simulation. Here are the energy levels corresponding to the four types of orbit.



The values I used in the simulation arise from $\ell = \eta \cdot \mathbf{u}$ and $\mathbf{u} \cdot \mathbf{u} = -1$ and are as follows

$$\begin{aligned} u^\phi &= \frac{\ell}{r^2} \\ u^t &= \left(1 + \left(\frac{\ell}{r}\right)^2 + \frac{(u^r)^2}{1 - 2/r}\right)^{1/2} \Big/ \left(1 - \frac{2}{r}\right) \end{aligned}$$

where $u^r = 0$ for the circular and precessing orbits and u^r was fixed by trial and error for the scattering and plunging orbots (so the energy levels shown for these orbits in the above figure are only approximate).

Note that I stopped the simulation for the plunging orbit after 8787 steps of size 0.01. One more step would zing the orbiting body straight out away from the Sun and it goes so far that a few more steps blow up the calculation and MATLAB won't plot at all.

This, as I said for the unstable circular orbit, is an artefact of the calculation. What happens is that the body crosses the $r = 2M$ barrier at which the denominator of

$$\frac{1}{1 - \frac{2M}{r}}$$

goes to zero and changes sign. For this simulation, the region

$$r \leq 2M$$

is unphysical and should be avoided.

The singularity at $r = 2M$ is actually not physical. It can be removed by changing coordinates (see excursion *Orthonormal polars* of Part I). But $r = 2M$ is a physical boundary, called the *Schwarzschild radius*. It is buried deeply inside normal stars and planets ($r_{\text{Sol}} = 2.95$ km, $r_{\text{Earth}} = 8.84$ mm). But a sufficiently massive star can overcome the repulsions of its electrons and protons and even the Pauli exclusion principle (weeks 6 and 7a) of its electrons and nucleons to collapse below this "Schwarzschild radius": for a nonrotating star (which is what Schwarzschild discusses) the critical mass is several solar masses.

The result of such a collapse is called a *black hole* because even light cannot escape from it. The first plot in this Note shows that black holes are impossible in Newtonian gravity (because of the shape of the potential) but not in relativistic gravity.

The actual radius of the Sun, 0.696 Gm, is 473 thousand times its mass M_{Sol}, so we can see why, for the orbit of Mercury at 46–70 Gm, the plots of the previous Note won't show any difference from Newton. This Note is really discussing only black holes.

How much energy can we extract from a Schwarzschild black hole? We drop a body into the black hole from infinity along the $\ell/M = 3.46$ (the $\sqrt{12}$) line, which gives the body a temporary stopping point where we can calculate the energy before dropping it the rest of the way into the hole.

The radius of this innermost stable circular orbit (ISCO) is found from the condition that the square root vanishes in

$$r = \frac{\ell^2}{2M}\left(1 \pm \sqrt{1 - 12\left(\frac{M}{\ell}\right)^2}\right)$$

namely

$$\frac{\ell^2}{M^2} = 12$$

so

$$\frac{r}{M} = 6$$

Since we assume it is orbiting, so $u^r = 0$, the energy per unit mass is just the effective potential (per unit mass)

$$\mathcal{E} = -\frac{M}{6M} + \frac{12M^2}{2(6M)^2} - \frac{12M^3}{(6M)^3} = -\frac{1}{18}$$

Now $\mathcal{E} = (e^2 - 1)/2$ by definition and to make the next step clearer I'll put lightspeed c back in, so

$$\begin{aligned} \mathcal{E} &= \frac{e^2 - c^2}{2} = \frac{c^2}{2}\left(\left(\frac{e}{c}\right)^2 - 1\right) \\ \frac{e}{c} &= \sqrt{1 + \frac{2\mathcal{E}}{c^2}} \approx 1 + \frac{\mathcal{E}}{c^2} \end{aligned}$$

where we take the nonrelativistic limit $\mathcal{E} \ll c^2$ (i.e., putting the mass m of the body back in, total $m\mathcal{E} \ll mc^2$ the rest energy).

So $ec = \mathcal{E} + c^2$ where $\mathcal{E}$ is potential (and 0 kinetic) energy. This just makes ec the total non-GR (general-relativistic) energy. This is the energy seen by an observer in the flat space an infinite distance away from the black hole.

Back to $c = 1$ we have

$$e = \sqrt{1 + 2\mathcal{E}} = \sqrt{1 - \frac{2}{18}} = \sqrt{\frac{8}{9}}$$

as the energy seen by that observer—in units of mc^2—per unit mass. This is less than 1, about 94% of 1, so the body has given up 6% of its energy.

Compare this 6% to the 0.6% we got for deuterium fusion in Note 11 of Week 7a: a black hole can convert an order of magnitude more mass to energy than thermonuclear fusion.

26. Rotationally symmetric gravity. The next free-space configuration (where the average curvature, and hence the Ricci tensor, vanishes) to be solved came in 1963, almost half a century after Schwarzschild's 1916 solution of the spherically symmetric case.

It's probably safe to say that all gravitating bodies rotate, even if they can otherwise be considered spherically symmetric for simplicity.

A body rotating about its $\theta = 0$ axis is not excluded by symmetry from having ϕ-t cross terms in its metric, because reversing ϕ is an acceptable symmetry operation (rotate in the other direction) as long as t is reversed at the same time.

In the following I'm going to cheat because Roy Kerr in 1963 did the hard work of finding the metric that sets the Ricci tensor to zero under rotation. I'll be pulling parts of the solution apparently out of the air. But I hope at least to motivate the form of Kerr's result. The proof of the handwaving will be that Ricci $= 0$ in the end.

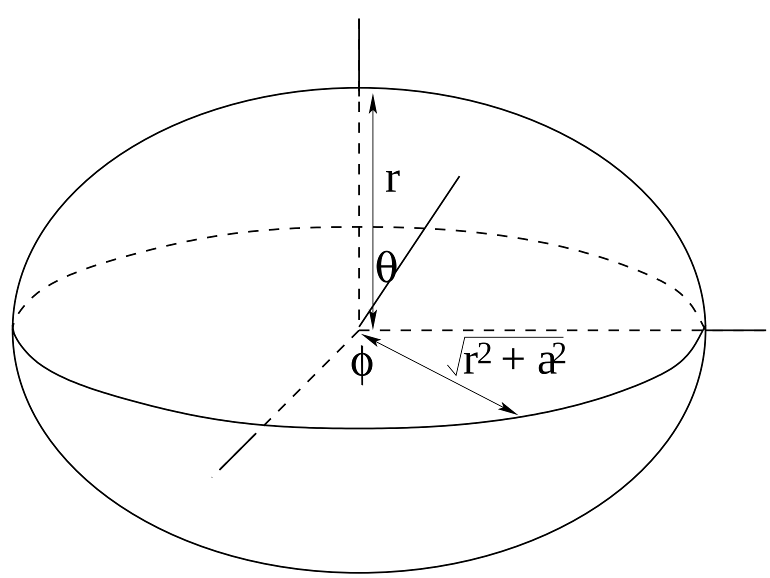
I'll elaborate on the ϕ-t cross term mentioned above, but first I'll consider two special cases: the ellipsoid in flat Cartesian space that is a sphere with a rotational bulge; and the no-rotation Schwarzschild limiting case. As well as the four coordinates t, r, θ and ϕ, we have two parameters, the mass M and the rotational angular momentum J, related to M for a rotating sphere of radius R and angular velocity ω by $J \propto MR^2\omega$.

First, if $M = 0$, we have flat space.

$$(\Delta s)^2 = -(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$$

because there is no gravity—but described with coordinates for an ellipsoid of rotation

$$\begin{aligned} x &= \sqrt{r^2 + a^2}\sin\theta\cos\phi \\ y &= \sqrt{r^2 + a^2}\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned}$$



By taking slopes with respect to r, θ and ϕ and abbreviating $s = \sin\theta$ and $c = \cos\theta$ we find

$$\begin{aligned}
\Delta x &= \frac{\Delta r}{\sqrt{r^2+a^2}} rs\cos\phi + \sqrt{r^2+a^2}(c\cos\phi\Delta\theta - s\sin\phi\Delta\phi) \\
\Delta y &= \frac{\Delta r}{\sqrt{r^2+a^2}} rs\sin\phi + \sqrt{r^2+a^2}(c\sin\phi\Delta\theta + s\cos\phi\Delta\phi) \\
\Delta z &= c\Delta r - rs\Delta\theta
\end{aligned}$$

The metric

$$(\Delta s)^2 = -(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$$

becomes

$$(\Delta s)^2 = -(\Delta t)^2 + \frac{r^2+a^2c^2}{r^2+a^2}(\Delta r)^2 + (r^2+a^2c^2)(\Delta\theta)^2 + (r^2+a^2)s^2(\Delta\phi)^2$$

This is the $M = 0$ limit, so to help us generalize later we define new symbols T, ρ, Δ and Φ in the limit $M = 0$

$$(\Delta s)^2 = -T_{M=0}(\Delta t)^2 + \frac{\rho^2_{M=0}}{\Delta_{M=0}}(\Delta r)^2 + \rho^2_{M=0}(\Delta\theta)^2 + \Phi_{M=0}(\Delta\phi)^2$$

Second, when $J = 0$, we get the spherically symmetric Schwarzschild metric

$$\begin{aligned}
(\Delta s)^2 &= -\left(1 - \frac{2M}{r}\right)(\Delta t)^2 + \frac{(\Delta r)^2}{1 - 2M/r} + r^2(\Delta\theta)^2 + r^2s^2(\Delta\phi)^2 \\
&= -T_{J=0}(\Delta t)^2 + \frac{\rho^2_{J=0}}{\Delta_{J=0}}(\Delta r)^2 + \rho^2_{J=0}(\Delta\theta)^2 + \Phi_{J=0}(\Delta\phi)^2
\end{aligned}$$

where I've again introduced T, ρ, Δ and Φ but this time in the $J = 0$ limit.

We can get some insights by comparing T, ρ, Δ and Φ in the two limiting cases but of course to generalize both involves either repeating Kerr's labour or drawing on his results.

$M = 0$	$J = 0$	try
$\rho^2_{M=0} = r^2 + a^2c^2$	$\rho^2_{J=0} = r^2$	$a = J/M = MR^2\omega/M = R^2\omega$ $\rho^2 = r^2 + a^2c^2$
$T_{M=0} = 1$	$T_{J=0} = 1 - 2M/r$ $= 1 = 2Mr/r^2$	$T = 1 - 2Mr/\rho^2$
$\Delta_{M=0} = r^2 + a^2$	$\Delta_{J=0} = \rho^2\,(1 - 2M/r)$	$\Delta = r^2 + a^2 - 2Mr$
$\Phi_{M=0} = r^2 + a^2$	$\Phi_{J=0} = r^2$	

Third, to find the last generalization, Φ, we introduce the ϕ-t cross term in the form $A(B\Delta\phi - \Delta t)^2$.

This is plausible if we think of an angular velocity $B = \Delta\phi/\Delta t$ and we can anticipate Kerr's result $B = (Rs)^2\omega = as^2$.

Since A must be part of the generalization of T we can try $A = -T + 1 = 2Mr/\rho^2$.

Then the cross term is

$$\frac{2Mr}{\rho^2}(a^2s^4(\Delta\phi)^2 - 2as^2(\Delta\phi)(\Delta t) + (\Delta t)^2)$$

and this tells us what to add to $\Phi_{M=0}$

$$\Phi = r^2 + a^2 + \frac{2Ma^2rs^2}{\rho^2}$$

Putting it all together we have the Kerr metric

$$(\Delta s)^2 = -\left(1 - \frac{2Mr}{\rho^2}\right)(\Delta t)^2 - \frac{4Mars^2}{\rho^2}\Delta\phi\Delta t + \frac{\rho^2}{\Delta}(\Delta r)^2 + \rho^2(\Delta\theta)^2 + \left(r^2 + a^2 + \frac{2Ma^2rs^2}{\rho^2}\right)s^2(\Delta\phi)^2$$

As I say, the proof of all this is that the Ricci tensor for this metric vanishes, which was Kerr's starting point. Using the protor calculator and MATLAB's `simplify()` we can confirm

```
gKerr = [t,t,2*M*r/rho2-1;t,phi,-2*M*r*a*s2/rho2;phi,t,-2*M*r*a*s2/rho2;...
      r,r,rho2/Delta;theta,theta,rho2;phi,phi,(r^2+a^2+2*M*r*a^2*s2/rho2)*s2]
detTPhigKerr = (2*M*r/rho2-1)*(r^2+a^2+2*M*r*a^2*s2/rho2)*s2 - (2*M*r*a*s2/rho2)^2;
invgKerr = [t,t,(r^2+a^2+2*M*r*a^2*s2/rho2)*s2/detTPhigKerr;...
      t,phi,2*M*r*a*s2/rho2/detTPhigKerr;phi,t,2*M*r*a*s2/rho2/detTPhigKerr;...
      r,r,Delta/rho2;theta,theta,1/rho2;phi,phi,(2*M*r/rho2-1)/detTPhigKerr]
connKerr = simplify(metr2conn(gKerr,invgKerr,[t,r,theta,phi]))
curv_udddKerr = conn2curv(connKerr,[t,r,theta,phi])
ricciKerr = simplify(contract(curv_udddKerr,1,3))
```

27. Kerr orbits and black holes. Because it does not have spherical symmetry a Kerr object has much more complicated orbits than a Schwarzschild object. We cannot, for a start, guarantee that an orbit will stay in one plane. Only in the equatorial plane will this be true, because of the axial symmetry. We will limit ourselves to this plane, i.e., $\theta = \pi/2, c = \cos\theta = 0$ and $s = \sin\theta = 1$.

To calculate and plot orbits we must adapt the geodesic calculator for Schwarzschild geometry (or the earlier calculator for spherical or plane polar coordinates, for instance) to the MATLAB function `geodesicKerrEq()` supported by `geodKerrEqStep()`.

The latter contains the Kerr-Equatorial form of the matrix containing $\Delta t, \Delta r$ and $\Delta\phi$ (but not $\Delta\theta$ which is zero in the equatorial plane, so that we can also leave out the θ coordinate and need use only a 3-by-3 matrix).

This matrix is based on the affine connection for the Kerr-Equatorial geometry. This has 13 components of which the independent ones are

$\Gamma_{\text{udd}}(\alpha$	β	γ	Γ	)
t	t	r	$M(r^2+a^2)/r^2/\Delta$	
t	t	ϕ	$-M(3r^2+a^2)/r^2/\Delta$	
r	t	t	$M\Delta/r^4$	
r	t	ϕ	$-Ma\Delta/r^4$	
r	r	r	$-(Mr-a^2)/r/\Delta$	
r	ϕ	ϕ	$(Ma^2-r^3)\Delta/r^4$	
ϕ	t	r	$Ma/r^2/\Delta$	
ϕ	r	ϕ	$-(M(2r^2+a^2)-r^3)/r^2/\Delta$	

where $\Delta = r^2 - 2Mr + a^2$ and, since this is mostly in the denominator, it is important to note that $\Delta = 0$ if

$$r = r_{\pm} = M \pm \sqrt{M^2 - a^2}$$

This also imposes the restriction $\mid a \mid \leq M$ (M being positive). The case of $\mid a \mid = M$ is called an *extremal* black hole because this gives the maximum angular momentum $\mid J \mid = \mid a \mid M \leq M^2$

The matrix for the geodesic calculations is

$$(\Gamma\Delta)^\alpha_\beta = \Gamma^\alpha_{\beta\gamma}\Delta x^\gamma$$

with $(\Delta x^t, \Delta x^r, \Delta x^\phi) = (\Delta t, \Delta r, \Delta \phi)$.

For the precessing, scattering and plunging orbits, I ran `geodesicKerrEq()` with exactly the same parameters as for the Schwarzschild orbits in Note 25, plus the new, Kerr, parmeter a which I set to its maximum value $a = M$ (and I'm using $M = 1$).

However, it is important to run the calculations for ϕ-velocities, u^ϕ, in both directions, because there will be a difference depending on whether the orbiting body is co-rotating with the black hole (counterclockwise since we'll take a to be positive), or counter-rotating.

Here are the two precessing orbits from the invocation

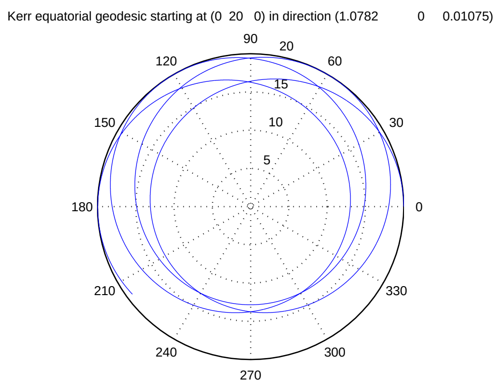
```
geodesicKerrEq(Uu,Xu,15000,0.1,1,1)
```

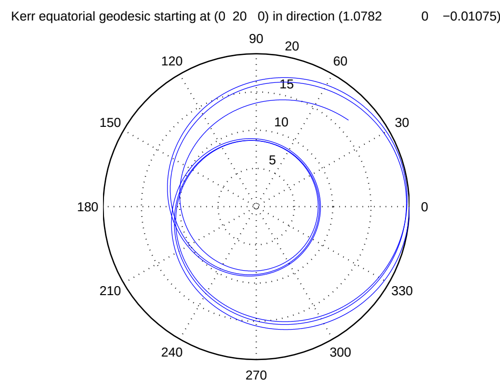
(the last two parameters being M and a respectively, preceded by the number of iterations and the step size) where

$$(u^t, u^r, u^\phi) = (1.0982, 0, \pm 0.0107)$$

and

$$(x^t, x^r, x^\phi) = (0, 20, 0)$$





The first precesses about 60 degrees each orbit and alternates between radii 20 and 13. Compare this with the Schwarzschild case of ∼100-degree precessing and radii 20 to about 12.

The counter-rotating case precesses hardly at all and the inner radius is about 8.

The differences illustrate "frame dragging" by the rotating Kerr object. It is effectively dragging timespace around itself as it rotates.

The scattering orbit for co-rotation becomes a plunging orbit for counter-rotation.

```
geodesicKerrEq(Uu,Xu,3000,-1,1,1)
```

with

$$(u^t, u^r, u^\phi) = (1.0791, 0.3, 0.0027)$$

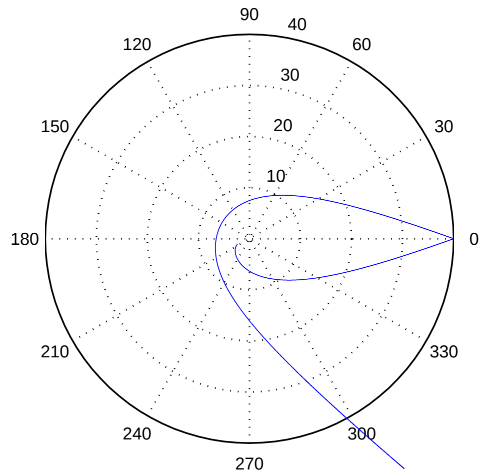
for co-rotation, and `geodesicKerrEq(Uu,Xu,1200,0.1,1,1)`
with

$$(u^t, u^r, u^\phi) = (1.0791, 0.3, -0.0027)$$

for counter-rotation, and

$$(x^t, x^r, x^\phi) = (0, 40, 0)$$

for both.



We can work out an effective potential for Kerr-Equatorial orbits but now it depends not only on ℓ but also in e—which are again names for conserved quantities calculated from the same two symmetry (Killing) vectors we used in Note 25.

$$\begin{aligned} -e &= \xi \cdot \mathbf{u} = -\left(1 - \frac{2M}{r}\right) u^t - \frac{2aM}{r} u^\phi \\ \ell &= \eta \cdot \mathbf{u} = -\frac{2aM}{r} u^t + \left(r^2 + a^2 + \frac{2Ma^2}{r}\right) u^\phi \end{aligned}$$

Solving this 2-by-2 system for u^t, u^ϕ

$$\begin{pmatrix} u^t \\ u^\phi \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} r^2 + a^2 + 2Ma/r & -2Ma/r \\ 2Ma/r & 1 - 2M/r \end{pmatrix} \begin{pmatrix} e \\ \ell \end{pmatrix}$$

Then we use, as in Note 25

$$\begin{aligned} -1 &= \mathbf{u} \cdot \mathbf{u} \\ &= (u^t \;\; u^\phi \;\; u^r) \begin{pmatrix} -(1 - 2M/r) & -2aM/r & \\ -2aM/r & r^2 + a^2 + 2Ma^2/r & \\ & & r^2/\Delta \end{pmatrix} \begin{pmatrix} u^t \\ u^\phi \\ u^r \end{pmatrix} \\ &= \frac{1}{\Delta^2} (e \;\; \ell \;\; u^r) \begin{pmatrix} A & B & \\ -B & C & \\ & & \Delta \end{pmatrix} \begin{pmatrix} -C & -B & \\ -B & A & \\ & & \frac{r^2}{\Delta} \end{pmatrix} \begin{pmatrix} A & -B & \\ B & C & \\ & & \Delta \end{pmatrix} \begin{pmatrix} e \\ \ell \\ u^r \end{pmatrix} \\ &= \frac{1}{\Delta} (e \;\; \ell \;\; u^r) \begin{pmatrix} -(r^2 + a^2 + 2Ma^2/r) & 2aM/r & \\ 2aM/r & 1 - 2M/r & \\ & & r^2 \end{pmatrix} \begin{pmatrix} e \\ \ell \\ u^r \end{pmatrix} \end{aligned}$$

where I introduced abbreviations A, B, C to make the matrix multiplication easier (and $\Delta = AC + B^2$). So

$$-(ru^2)^2 = \Delta - e^e \left(r^2 + a^2 + \frac{2Ma^2}{r}\right) - e\ell \frac{4Ma}{r} + \ell^2 \left(1 - \frac{2M}{r}\right)$$

Since $\Delta = r^2 + a^2 - 2Mr$, when we divide both sides by r^2 we can see another effective potential energy as a sum of the same powers of $1/r$, namely $1/r, 1/r^2$ and $1/r^3$, as in the Schwarzschild case in Note 25.

After rearranging we have

$$\frac{e^2 - 1}{2} = \frac{u^r u^r}{2} - V_{\text{eff}}$$

with

$$V_{\text{eff}} = -\frac{M}{r} + \frac{\ell^2 - a^2(e^2 - 1)}{2r^2} - \frac{M(ae - \ell)^2}{r^3}$$

This effective potential depends on e as well as on ℓ. And its dependence on ℓ includes a dependence on the sign of ℓ, unlike the Schwarzschild case which depends only on ℓ^2.

Thus there will be a difference between co-rotating orbits ($\ell > 0$) and counter-rotating orbits ($\ell < 0$) as we saw above.

Because of the three inverse powers of r, the effective potential has the same qualitative shape as that for Schwarzschild, except when $\ell = ae$, when there is no $-1/r^3$ term to dominate it at small r and the $+1/2r^2$ term provides an infinite potential barrier against plunging orbits.

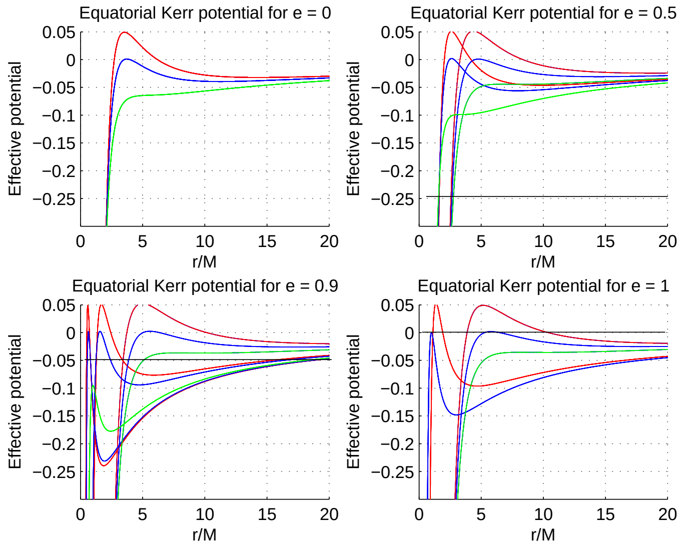
Here are some plots of the effective potential for various ℓ at $e = 0$, 0.5, 0.9 and 1. The values of ℓ were chosen to give maximum height at about 0.05 (red plots) and about 0 (blue plots). The green plots show the effective potential giving the innermost stable circular orbit (ISCO) for the respective energies. Here are those values of ℓ.

	$e = 0$		$e = 0.5$	
	co-rot	con-rot	co-rot	con-rot
red	4.1	−4.1	3.42	−4.75
blue	3.74	−3.74	3.13	−4.3
green	3.1	−3.1	2.45	−3.75

	$e = 0.9$		$e = 1$	
	co-rot	con-rot	co-rot	con-rot
red	3.72	−5.75	2.46	−5.34
	1.505			
blue	2.47	−4.75	2	−4.85
	1.54			
green	(1.8)	−4.2	—	−4.3

(Recall that all these plots are for the extremal Kerr black hole.)

Here are the plots.



Where the plots don't overlap, the triplet on the right is the counter-rotating set. Note that for $e = 0.9$ two sets of co-rotating plots appear. And also I couldn't find a green potential of the same slope as the others but plotted the closest I could get.

The horizontal black line in three of the plots is the level of $(e^2 - 1)/2$ fixed for each plot.

We can actually find the ISCO over all e and ℓ by arguing that the radial acceleration must vanish to keep the orbit circular and that the shape of the potential at the orbit must be concave with respect to r to keep it stable—or, actually, flat if it is *just* stable.

Since acceleration in the r direction is caused by a force which is given by the r-shape of the potential, circularity requires

$$0 = \partial_r V_{\text{eff}} = \frac{M}{r^2} - \frac{2P}{r^3} + \frac{3Q}{r^4}$$

where I've abbreviated

$$V_{\text{eff}} = -\frac{M}{r} + \frac{P}{r^2} - \frac{Q}{r^3}$$

The stability condition is

$$0 \leq \partial_{rr} V_{\text{eff}} = -\frac{2M}{r^3} + \frac{6P}{r^4} - \frac{12Q}{r^5}$$

with equality holding for the *innermost* case.

We can multiply these two equations on both sides by r and r^2 respectively, and combine them with

$$\mathcal{E} = V_{\text{eff}} = -\frac{M}{r} + \frac{P}{r^2} - \frac{Q}{r^3}$$

(where I've abbreviated $(e^2 - 1)/2 = \mathcal{E}$) because, for the orbit to be circular, we also have $u^r = 0$. This gives

$$\begin{pmatrix} -M & P & -Q \\ M & -2P & 3Q \\ -2M & 6P & -12Q \end{pmatrix} \begin{pmatrix} 1/r \\ 1/r^2 \\ 1/r^3 \end{pmatrix} = \begin{pmatrix} \mathcal{E} \\ 0 \\ 0 \end{pmatrix}$$

Gaussian elimination (Week 8 Note 5) brings this to

$$\begin{pmatrix} -M & P & -Q \\ & -P & 2Q \\ & & -2Q \end{pmatrix} \begin{pmatrix} 1/r \\ 1/r^2 \\ 1/r^3 \end{pmatrix} = \begin{pmatrix} \mathcal{E} \\ \mathcal{E} \\ 2\mathcal{E} \end{pmatrix}$$

and working from the bottom to the top of this triangular matrix (either by more elementary row operations or by algebra) gives

$$\begin{aligned} Q &= -\mathcal{E}r^3 \\ P &= -3\mathcal{E}r^2 \\ M &= -3\mathcal{E}r \end{aligned}$$

These make three equations in the three unkowns, r, e and ℓ. We can make M disappear by introducing $\rho = r/M, \alpha = a/M$ and $\lambda = \ell/M$. Then we can get $\mathcal{E}$ (and so e) in terms of ρ from $M = -3\mathcal{E}r$

$$\begin{aligned} \rho &= \frac{r}{M} = -\frac{1}{3\mathcal{E}} \\ \mathcal{E} &= -\frac{1}{3\rho} \end{aligned}$$

and since

$$\begin{aligned} \mathcal{E} &= (e^2-1)/2 \\ e^2 &= 1+2\mathcal{E} \\ &= 1-\frac{2}{3\rho} \\ e &= \pm\sqrt{1-\frac{2}{3\rho}} \end{aligned}$$

and we'll consider only non-negative e

Next, $Q = -\mathcal{E}r^3$ gives λ in terms of ρ and e

$$\begin{aligned} -\mathcal{E}r^3 &= Q = M(\ell - ae)^2 \\ (\lambda - \alpha e)^2 &= -\mathcal{E}\rho^3 = \frac{\rho^2}{3} \\ \lambda &= \alpha e \pm \frac{\rho}{\sqrt{3}} \end{aligned}$$

Finally $P = -3\mathcal{E}r^2$ leads us to a complicated equation involving only ρ.

$$\begin{aligned} -3\mathcal{E}r^2 &= P = \frac{\ell^2}{2} - a^2\mathcal{E} \\ 0 &= (a^2 - 3r^2)\mathcal{E} - \frac{\ell^2}{2} \\ 0 &= (\alpha^2 - 3\rho^2)\mathcal{E} - \frac{\lambda^2}{2} \\ &= \frac{3\rho^2 - \alpha^2}{3\rho} - \frac{\lambda^2}{2} \\ 0 &= 3\rho^2 - \frac{3}{2}\rho\lambda^2 - \alpha^2 \\ &= 3\rho^2 - \frac{3}{2}\rho\left(\alpha e \pm \frac{\rho}{\sqrt{3}}\right)^2 - \alpha^2 \\ &= 3\rho^2 - \frac{3}{2}\rho\left(\alpha^2 e^2 \pm \frac{2\alpha e\rho}{\sqrt{3}} + \frac{\rho^2}{3}\right) - \alpha^2 \end{aligned}$$

$$
\begin{aligned}
&= 3\rho^2 - \frac{3}{2}\rho\left(\alpha^2\left(1-\frac{2}{3\rho}\right) \pm \frac{2\alpha\rho}{\sqrt{3}}\sqrt{1-\frac{2}{3\rho}} + \frac{\rho^3}{3}\right) - \alpha^2 \\
&= -\frac{\rho^3}{3} \mp \alpha\rho^2\sqrt{3}\sqrt{1-\frac{2}{3\rho}} + 3\rho^2 - \frac{3}{2}\alpha^2\rho
\end{aligned}
$$

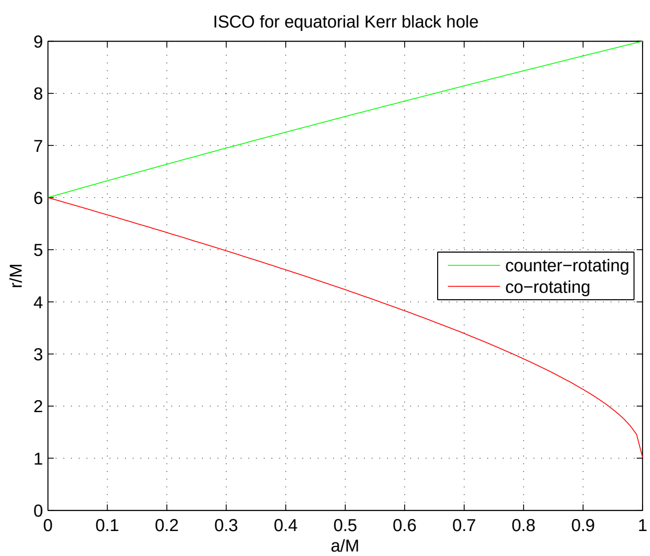
It seems less work to use Newton's method (Week v Note 5) than to solve this analytically to find how ρ depends on α. But first we can note the special case $\alpha = 1$ ($a = M$). This can be solved with $\rho = 1$ and $e = 1/\sqrt{3}$, from which $\lambda = 2/\sqrt{3}$ (using the upper signs in $\lambda = \alpha e \pm \rho/\sqrt{3}$):

$$
\begin{aligned}
0 &= 3\rho^2 - \frac{3}{2}\rho\lambda^2 - \alpha^2 \\
&= 3 - \frac{3}{2}\frac{4}{3} - 1
\end{aligned}
$$

Newton's method requires the slope with respect to ρ of the final complicated equation, which is

$$
\begin{aligned}
&\partial_\rho\left(-\frac{\rho^3}{3} \mp \alpha\rho^2\sqrt{3}\sqrt{1-\frac{2}{3\rho}} + 3\rho^2 - \frac{3}{2}\alpha^2\rho\right) \\
= &-\frac{3}{2}\rho^2 \mp 2\sqrt{3}\alpha\rho\sqrt{1-\frac{2}{3\rho}} \mp \frac{\alpha}{\sqrt{3}\sqrt{1-2/3\rho}} + 6\rho - \frac{3}{2}\alpha^2
\end{aligned}
$$

Here are the two solutions of this as plots of ρ versus α.

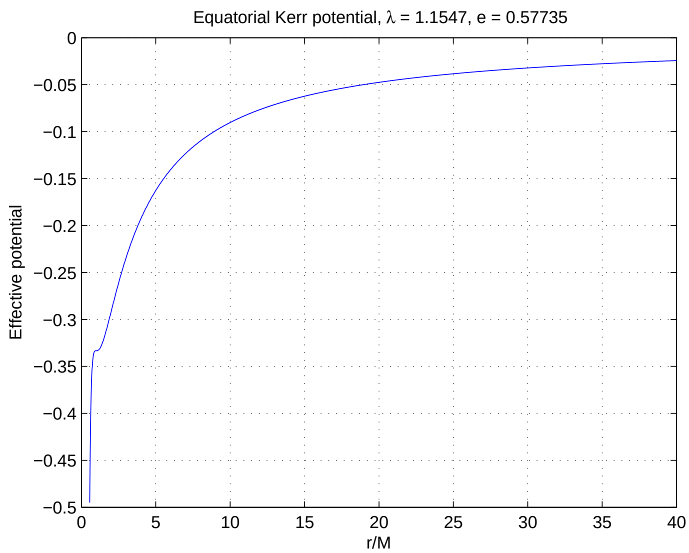


We can see the red curve reaching $\rho = 1$ at $\alpha = 1$. We can also see the non-rotating limit at $\alpha = 0$ where $\rho = 6$ in agreement with the innermost stable circular orbit for Schwarzschild in Note 25.

The two solutions correspond to the sign of the angular momentum λ. If λ is positive the body orbits in the direction the Kerr black hole is rotating (co-rotating), otherwise it is counter-rotating.

(By the way, the Newton's method calculation is very sensitive to the first guess chosen each time. I used $\rho = 6 - 4\alpha$ to start the co-rotating solution and $\rho = 6 + 3\alpha$ to start the counter-rotating solution. The latter evidently remains essentially linear.)

Here is the effective potential V_{eff} for the ISCO values $e = 1/\sqrt{3}, \lambda = 2/\sqrt{3}$ at $\alpha = 1$ (invoking `potenKerrEq(alpha,lambda,e,'b')`).



We can see the little ledge that the co-rotating orbiting body can get caught on. If it has fallen in to the ledge from infinity, it releases

$$1 - e = 1 - \frac{1}{\sqrt{3}} = 42\%$$

of its rest energy. Compare this with the 6% found in Note 25 for the Schwarzschild black hole: the Kerr black hole can in principle release almost half as much energy as total annihilation.

28. Tides. Schwarzschild tidal forces are identical to Newton's. But they are caused by the curvature of timespace as given by the full Riemann tensor. So far we have used only the average curvature, expressed in the Ricci tensor. We need at least two bodies to tell us about the full curvature. Excursion *Falling traffic lights* of Part I shows an example of four bodies.

Each body falls along its own geodesic and the equation of geodesic deviation shows the effect of the curvature on the separation χ between the two bodies. In orthonormal coordinates with $\mathbf{e}_{\hat{\tau}}$, the timelike axis (using proper time τ because the coordinate system is falling with the body), this equation is

$$\partial_{\hat{\tau}\hat{\tau}}\chi^{\hat{\alpha}} = -R^{\hat{\alpha}}_{\hat{\tau}\hat{\beta}\hat{\tau}}\chi^{\hat{\beta}}$$

The bodies are falling radially (even if they are in a satellite or on a planet in orbit) so the four-velocity $\mathbf{u} = \mathbf{e}_{\hat{\tau}}$. The four basis vectors $\mathbf{u} = \mathbf{e}_{\hat{\alpha}}$, with, say, $\hat{\alpha} = (\hat{\tau}, \hat{r}, \hat{\theta}, \hat{\phi})$, are parallel-propagated along the geodesic so their absolute slopes in that direction are zero

$$\mathbf{D_u e}_{\hat{\alpha}} = 0$$

It is not hard to see that the chain rule for slopes says we should calculate

$$u^{\beta} D_{\beta}(\mathbf{e}_{\hat{\alpha}})^{\gamma} = 0$$

So now we must find the orthonormal basis that has $\mathbf{e}_{\hat{\tau}} = \mathbf{u}$. First, we can use the work of Note 25 to find $\mathbf{u}$.

If we consider bodies falling freely and radially from infinity, the conserved quantities from Note 25 are

$$e = 1 \qquad \text{and} \qquad \ell = 0$$

So the equation we get in Note 25 from $\mathbf{u} \cdot \mathbf{u} = -1$

$$\frac{e^2 - 1}{2} = \frac{1}{2}(u^r)^2 - \frac{M}{r} + \frac{\ell^2}{2r^2} - \frac{M\ell^2}{r^3}$$

becomes

$$0 = \frac{1}{2}(u^r)^2 - \frac{M}{r}$$

so

$$u^r = \sqrt{\frac{2M}{r}}$$

For u^t we had in Note 25

$$e = \left(1 - \frac{2M}{r}\right) u^t$$

so

$$u^t = \left(1 - \frac{2M}{r}\right)^{-1}$$

Thus we have

$$\mathbf{u} = \left(\left(1 - \frac{2M}{r}\right)^{-1}, \sqrt{\frac{2M}{r}}, 0, 0\right)$$

This is $\mathbf{e}_{\hat{\tau}}$. From spherical symmetry, $\mathbf{e}_{\hat{\theta}}$ and $\mathbf{e}_{\hat{\phi}}$ cannot mix with $\mathbf{e}_{\hat{\tau}}$ or $\mathbf{e}_{\hat{r}}$ so for orthonormality

$$\begin{aligned} \mathbf{e}_{\hat{\theta}} &= (0, 0, 1/r, 0) \\ \mathbf{e}_{\hat{\phi}} &= (0, 0, 0, 1/rs) \end{aligned}$$

because

$$1 = \mathbf{e}_{\hat{\theta}} \cdot \mathbf{e}_{\hat{\theta}} = e^{\theta}_{\hat{\theta}} g_{\theta\theta} e^{\theta}_{\hat{\theta}} = \frac{1}{r} r^2 \frac{1}{r}$$

and similarly for $\mathbf{e}_{\hat{\phi}}$ since $g_{\phi\phi} = r^2 s^2$.

Finally we must find $\mathbf{e}_{\hat{r}}$ orthonormal to all the others

$$\mathbf{e}_{\hat{r}} = \left(-\sqrt{\frac{2M}{r}}\left(1 - \frac{2M}{r}\right)^{-1}, 1, 0, 0\right)$$

(The minus sign on the square root is because the bodies are infalling.

To get the Riemann curvature we can use `transform()` (Note 20 in Part I) on $R_{\rm dddd}$ (Note 24)

```
transform(curv_ddddSchwarz,xSchwFreeFall,[],'dddd')
```

where `curv_ddddSchwarz` represents the protor $R_{\rm dddd}$ and `xSchwFreeFall` is the protor version of the transformation matrix coming from the orthonormal basis.

`xSchwFreeFall`(α	β	val	)
t	t	$(1-2M/r)^{-1}$	
t	r	$-\sqrt{2M/r}$	
r	t	$-\sqrt{2M/r}(1-2M/r)^{-1}$	
r	r	1	
θ	θ	$1/r$	
ϕ	ϕ	$1/rs$	

The six independent components of the result simplify to

$R_{\rm dddd}$($\hat{\gamma}$	$\hat{\delta}$	$\hat{\beta}$	$\hat{\epsilon}$	R	)
t	r	t	r	$-2M/r^3$	
t	θ	t	θ	M/r^3	
t	ϕ	t	ϕ	M/r^3	
r	θ	r	θ	$-M/r^3$	
r	ϕ	r	ϕ	$-M/r^3$	
θ	ϕ	θ	ϕ	$2M/r^3$	

Of these, we're interested in those with t under $\hat{\delta}$ and $\hat{\epsilon}$, so we use symmetry to swap the values under $\hat{\gamma}$ with $\hat{\delta}$ and $\hat{\beta}$ with $\hat{\epsilon}$ in the first three rows

$$
\begin{array}{cccccc}
R_{\mathrm{dddd}}(\hat{\gamma} & \hat{\delta} & \hat{\beta} & \hat{\epsilon} & R &) \\
r & t & r & t & -2M/r^3 & \\
\theta & t & \theta & t & M/r^3 & \\
\phi & t & \phi & t & M/r^3 &
\end{array}
$$

Then we must raise the $\hat{\gamma}$ index to $\hat{\alpha}$ by multiplying by $g^{\hat{\gamma}\hat{\alpha}} = \delta^{\hat{\gamma}\hat{\alpha}} = 1$ for $\hat{\gamma}, \hat{\alpha} \neq t$ in the Minkowski metric.

So, finally, the tidal accelerations are

$$
\begin{aligned}
\partial_{\hat{\tau}\hat{\tau}} \chi^{\hat{r}} &= \frac{2M}{r^3} \chi^{\hat{r}} \\
\partial_{\hat{\tau}\hat{\tau}} \chi^{\hat{\theta}} &= -\frac{M}{r^3} \chi^{\hat{\theta}} \\
\partial_{\hat{\tau}\hat{\tau}} \chi^{\hat{\phi}} &= -\frac{M}{r^3} \chi^{\hat{\phi}}
\end{aligned}
$$

where χ is the separation between the bodies, e.g., the diameter of a planet experiencing the tidal forces.

It is not straightforward to turn these accelerations into tidal height, even for a featureless planet with uniform ocean depth (if we're talking water tides, which is the best known kind of tide on Earth). A featureless earth, covered entirely by ocean of fixed depth, would experience 1/2-meter tides from the Moon and 1/4-meter tides from the Sun. These sometimes add together ("spring tides") or subtract ("neap tides") depending on the relative positions of Sun and Moon. Note that the inverse r^3 makes the solar tide weak relative to the lunar tide, despite the immensely greater mass M of the Sun.

By contrast, the highest tidal differences on Earth are some 10 meters in the Bay of Fundy, caused by geographical features and in particular the sort of resonance you get if you slosh back and forth in a full bathtub with the right frequency.

29. Light orbits. The "four-velocity" $\mathbf{u}$ we've been using, in Notes 25 and 27, for particles with mass satisfies

$$\mathbf{u} \cdot \mathbf{u} = -1$$

and is defined as a slope with respect to the proper time τ

$$u^\alpha = \partial_\tau x^\alpha$$

for position $\mathbf{x}$ (say $\mathbf{x} = (t, r, \theta, \phi)$, or we'll also be using $\mathbf{x} = (t, x, y, z)$).

Both of these considerations must change when we deal with massless particles such as the photons of light, because massless particles travel at lightspeed and their proper time τ is always zero, as we see below. Now

$$\mathbf{u} \cdot \mathbf{u} = 0$$

and

$$u^\alpha = \partial_\lambda x^\alpha$$

for some parameter λ. Let's explore.

For a massy particle (i.e., a particle with nonzero rest mass: the usual term, "massive", can be misleading), velocity $\vec{v} = (v^x, y^y, v^z)$ gives rise to the contraction and dilation in the Lorentz transformation of special relativity (Week 3 Note 7) with the factor

$$\gamma = \frac{1}{\sqrt{1 - v^2}}$$

and the four-velocity

$$\mathbf{u} = (\gamma, \gamma\vec{v}) = (\gamma, \gamma v^x, \gamma v^y, \gamma v^z)$$

so, of course

$$\begin{aligned}\mathbf{u}\cdot\mathbf{u} &= (\gamma, \gamma\vec{v})\begin{pmatrix} -1 & \\ & I \end{pmatrix}\begin{pmatrix} \gamma \\ \gamma\vec{v} \end{pmatrix} \\ &= -\gamma^2(1-v^2) = -1\end{aligned}$$

But if $v^2 = 1$ (lightspeed) γ is undefined and the relationship cannot hold.

Furthermore, proper time, the invariant

$$\begin{aligned}\tau^2 = \mathbf{x}\cdot\mathbf{x} &= (t, \vec{x})\begin{pmatrix} -1 & \\ & I \end{pmatrix}\begin{pmatrix} t \\ \vec{x} \end{pmatrix} \\ &= x^2 - t^2 = v^2t^2 - t^2 = 0\end{aligned}$$

cannot be used as a parameter to track progress through timespace (as the "world line" of the particle) because it is zero everywhere. So we need a new way of parametrizing in order, for instance, to find slopes and hence velocities.

A simple way to use a parameter λ analogously to τ with

$$u^\alpha = \partial_\lambda x^\alpha$$

and the force-free acceleration of empty space

$$a^\alpha = \partial_\lambda u^\alpha = 0$$

would be to write (illustrated for lightspeed along the x-axis)

$$x^\alpha = u^\alpha \lambda$$

with

$$\mathbf{u} = (1, 1, 0, 0)$$

so that $x = ct$ in $c = 1$ units.

Note though that we could equally have

$$x^\alpha = u^\alpha (2\lambda)$$

with

$$\mathbf{u} = (1/2, 1/2, 0, 0)$$

so that

$$u^\alpha = \partial_{2\lambda} x^\alpha$$

That is, λ is unique as a parameter only up to multiplication by a constant. This will be important later in this Note.

For Schwarzschild orbits of light everything is the same as Note 25 except $\mathbf{u}\cdot\mathbf{u} = 0$

$$\begin{aligned} e &= -\xi\cdot\mathbf{u} = \left(1 - \frac{2M}{r}\right) u^t \\ \ell &= \eta\cdot\mathbf{u} = r^2 s^2 u^\phi \end{aligned}$$

but now, for the equatorial orbit $\theta = \pi/2, s = 1$

$$0 = -\left(1 - \frac{2M}{r}\right)^{-1} e^2 + \left(1 - \frac{2M}{r}\right)^{-1} (u^r)^2 + \frac{\ell^2}{r^2}$$

And from this

$$0 = -e^2 + \frac{\ell^2}{r^2}\left(1 - \frac{2M}{r}\right) + (u^r)^2$$

so

$$\left(\frac{e}{\ell}\right)^2 = \left(\frac{u^r}{\ell}\right)^2 + W_{\text{eff}}$$

with

$$W_{\text{eff}} = \frac{1}{r^2}\left(1 - \frac{2M}{r}\right)$$

This form is dictated by the fact that both e and ℓ involve slopes with respect to λ (respectively $u^t = \partial_\lambda t$ and $u^\phi = \partial_\lambda \phi$) and that λ is arbitrary up to multiplication by a constant. So only the *ratio* e/ℓ can be meaningful physically.

The effective potential for light has only one extremum, a maximum at

$$0 = \partial_r W_{\text{eff}} = \frac{6M}{r^4} - \frac{2}{r^3}$$

or $r = 3M$ (NB this is outside the Schwarzschild horizon at $r = 2M$) with value $1/27M^2$.

Thus there are three types of orbit for light: a deflected orbit for $(e/\ell)^2 < 1/27M^2$, an unstable circular orbit at the maximum $(e/\ell)^2 = 1/27M^2$, and a plunging orbit at $(e/\ell)^2 > 1/27M^2$.

So a star can bend light, as Einstein predicted in 1915 and two English expeditions observed in 1919 right after ceasing hostilities with Einstein's country.

Light can also plunge into a star when $(e/\ell)^2 > 1/27M^2$, and not escape.

Finally, light can orbit a star sufficiently dense that its radius is $< 3M$ (so that the orbiting light remains outside the star).

For a Kerr orbit of light we also change only $\mathbf{u} \cdot \mathbf{u} = 0$ from the discussion of Note 27.

$$\begin{aligned}
0 &= \mathbf{u} \cdot \mathbf{u} = \frac{1}{\Delta}(e \;\; \ell \;\; u^r) \begin{pmatrix} -\left(r^2 + a^2 + 2Ma/r\right) & 2Ma/r & \\ 2Ma/r & 1 - 2M/r & \\ & & r^2 \end{pmatrix} \begin{pmatrix} e \\ \ell \\ u^r \end{pmatrix} \\
0 &= -\left(r^2 + a^2 + \frac{2Ma}{r}\right) e^2 + \frac{4Ma}{r^2}\frac{e}{\ell} + \left(1 - \frac{2M}{r}\right) \ell^2 + r^2 (u^r)^2 \\
0 &= -\left(1 + \left(\frac{a}{r}\right)^2 + \left(\frac{a}{r}\right)^2 \frac{2M}{r}\right)\left(\frac{e}{\ell}\right)^2 + \frac{4Ma}{r^2}\frac{e}{\ell} + \frac{1}{r^2}\left(1 - \frac{2M}{r}\right) + \left(\frac{u^r}{\ell}\right)^2 \\
\left(\frac{u^r}{\ell}\right)^2 &= \left(\frac{e}{\ell}\right)^2 - W_{\text{eff}}
\end{aligned}$$

with

$$W_{\text{eff}} = \frac{1}{r^2}\left(1 - \left(\frac{ea}{\ell}\right)^2 + \frac{2M}{r}\left(1 - \frac{ea}{\ell}\right)^2\right)$$

This has the same r-dependence as Schwarzschild, and equals it when $a = 0$. So the orbits are qualitatively the same, except for the e/ℓ dependence and in particular the dependence on the sign of ℓ.

30. The source of gravity. For Newton's gravity we found at the end of Note 7 in Part I that the source is *mass density*, ρ. We had there that the second slope of the Newtonian potential ϕ

$$\mathbf{div} \cdot \mathbf{grad}\phi = 4\pi G_N \rho$$

For Einstein's gravity we have found in Part II, particularly Notes 24, 26 and 29, that the metric tensor plays the role of the potential and its second slope is the curvature tensor.

In Part II so far we have considered only gravity in free space, containing no mass, albeit *influenced* by mass with either spherical or rotational symmetry. In free space, the *average* curvature is zero, as expressed by the Ricci tensor $R_{\alpha\beta} = R^{\mu}_{\alpha\mu\beta}$

$$R_{\alpha\beta} = 0$$

In the presence of matter, Einstein had to extend the right-hand side of this equation to a non-zero tensor, giving a four-dimensional generalization of the Newtonian density scalar ρ.

Here is a simple motivation for the *stress-energy-momentum* tensor that he came up with, using the energy-momentum 4-vector (see the 2-D time-and-space version of this in Note 4 of Week 7a) together with the timespace 4-vector giving an increment of time and space.

$$T_{\rm dd} = \begin{pmatrix} \Delta E/\Delta\mathcal{V}_4 \\ \Delta p_x/\Delta\mathcal{V}_4 \\ \Delta p_y/\Delta\mathcal{V}_4 \\ \Delta p_z/\Delta\mathcal{V}_4 \end{pmatrix} (\Delta t \ \ \Delta x \ \ \Delta y \ \ \Delta z)$$

where I've written energentum as a density

$$\Delta(\text{energentum})/\Delta\mathcal{V}_4$$

where $\Delta\mathcal{V}_4$ is the "4-volume", effectively $\Delta t\Delta x\Delta y\Delta z$, and so the density expresses rates of change in time as well as rates of change across space.

Multiplied out and cancelling $\Delta t, \Delta x, \Delta y$ and Δz, respectively, this tensor has the interpretation

$$T_{\rm dd} = \left(\begin{array}{c|c} \text{energy density} & \text{energy flux} \\ \hline \text{mom. density} & \text{mom. flux} \end{array}\right)$$

where "density" refers as usual to amount per spatial volume $1/\Delta\mathcal{V}_3 = \Delta t/\Delta\mathcal{V}_4 = 1/\Delta x\Delta y\Delta z$, and "flux" refers to rate of flow across an area, e.g., $\Delta x/\Delta\mathcal{V}_4 = 1/\Delta t\Delta y\Delta z$.

The stress tensor, discussed in Excursion *Elasticity and viscosity* of Book 9c Part III, is force per unit area, which is the rate of change of momentum per unit area, e.g., $(\Delta p/\Delta\mathcal{V}_4)\Delta x = \Delta p/\Delta t\Delta y\Delta z$. It is symmetrical, as we argued there, so that there is no spontaneous rotation.

Because of the equivalence of mass and energy we also have equality between momentum density (the lower part of the left column of $T_{\rm dd}$) and energy flux (the right part of the top row of $T_{\rm dd}$).

$$\begin{aligned} \text{momentum density} &= (\text{mass density}) \times (\text{mean velocity of mass flow}) \\ &= (\text{energy density}) \times (\text{mean velocity of energy flow}) \\ &= \text{energy flux} \end{aligned}$$

Thus the entire "stress-energy" tensor (the usual appelation drops "momentum") is symmetrical. That's fortunate because it must match the symmetrical Ricci tensor.

Unfortunately, we cannot just say

$$R_{\alpha\beta} = T_{\alpha\beta} \qquad \text{(wrong!)}$$

because energy and momentum are *conserved*

$$D_\alpha T_{\alpha\beta} = 0$$

and the Ricci tensor does not go along with this

$$D_\alpha R_{\alpha\beta} \neq 0$$

(Note that because we conventionally work only with the covariant absolute slope I've had to raise the indices: check that this does not disturb anything.)

So we must find a variant of the Ricci tensor which does go along. It will turn out to be

$$R^{\alpha\beta} - \frac{1}{2}Rg^{\alpha\beta}$$

where we introduced the *curvature scalar* R in Note 9 of Part I.

The route to this is a little long, if straightforward, so you should first check that it works in free space, i.e., is zero when $R_{\alpha\beta} = 0$.

First we establish the *Jacobi identity* for any non-commuting operators.

$$\begin{aligned}
& [D_\lambda, [D_\mu, D_\nu]] + [D_\nu, [D_\lambda, D_\mu]] + [D_\mu, [D_\nu, D_\lambda]] \\
= \; & [D_\lambda, D_\mu D_\nu - D_\nu D_\mu] + [D_\nu, D_\lambda D_\mu - D_\mu D_\lambda] + [D_\mu, D_\nu D_\lambda - D_\lambda D_\nu] \\
= \; & D_\lambda(D_\mu D_\nu - D_\nu D_\mu) - (D_\mu D_\nu - D_\nu D_\mu)D_\lambda \\
& + D_\nu(D_\lambda D_\mu - D_\mu D_\lambda) - (D_\lambda D_\mu - D_\mu D_\lambda)D_\nu \\
& + D_\mu(D_\nu D_\lambda - D_\lambda D_\nu) - (D_\nu D_\lambda - D_\lambda D_\nu)D_\mu \\
= \; & 0
\end{aligned}$$

where you can find the cancellations.

Second we show the *Bianchi identity* involving specifically the absolute slope D_μ and the Riemann curvature which we found, in the Excursion *Absolute slope and curvature* of Part I, to satisfy

$$[D_\mu, D_\nu]A_\alpha = -R^{\gamma}_{\alpha\mu\nu}A_\gamma$$

and

$$[D_\mu, D_\nu]D_\lambda A_\alpha = -R^{\gamma}_{\lambda\mu\nu}D_\gamma A_\alpha - R^{\gamma}_{\alpha\mu\nu}D_\lambda A_\gamma$$

and, in Note 17 of Part I, to satisfy symmetries listed there under pairwise swaps of indices, plus the cyclic symmetry

$$R_{\mu\nu\alpha\beta} + R_{\mu\beta\nu\alpha} + R_{\mu\alpha\beta\nu} = 0$$

(and similarly if the first index is raised by multiplying by, say, $g^{\lambda\mu}$).

Here we go. We write each of the three terms in the Jacobi identity in terms of R. Here's the first

$$\begin{aligned}
[D_\lambda, [D_\mu, D_\nu]]A_\alpha &= D_\lambda[D_\mu, D_\nu]A_\alpha - [D_\mu, D_\nu]D_\lambda A_\alpha \\
&= -D_\lambda(R^{\gamma}_{\alpha\mu\nu}A_\gamma) + R^{\gamma}_{\lambda\mu\nu}D_\gamma A_\alpha + R^{\gamma}_{\alpha\mu\nu}D_\lambda A_\gamma \\
&= -(D_\lambda R^{\gamma}_{\alpha\mu\nu})A_\gamma - R^{\gamma}_{\alpha\mu\nu}D_\lambda A_\gamma + R^{\gamma}_{\alpha\mu\nu}D_\lambda A_\gamma + R^{\gamma}_{\lambda\mu\nu}D_\gamma A_\alpha \\
&= -(D_\lambda R^{\gamma}_{\alpha\mu\nu})A_\gamma + R^{\gamma}_{\lambda\mu\nu}D_\gamma A_\alpha
\end{aligned}$$

Permuting indices gives us the other two.

$$[D_\nu, [D_\lambda, D_\mu]]A_\alpha = -(D_\nu R^{\gamma}_{\alpha\lambda\mu})A_\gamma + R^{\gamma}_{\nu\lambda\mu}D_\gamma A_\alpha$$

$$[D_\mu, [D_\nu, D_\lambda]]A_\alpha = -(D_\mu R^{\gamma}_{\alpha\nu\lambda})A_\gamma + R^{\gamma}_{\lambda\mu\nu}D_\gamma A_\alpha$$

Adding the three

$$\begin{aligned}
0 &= -(D_\lambda R^{\gamma}_{\alpha\mu\nu} + D_\nu R^{\gamma}_{\alpha\lambda\mu} + D_\mu R^{\gamma}_{\alpha\nu\lambda})A_\gamma + (R^{\gamma}_{\lambda\mu\nu} + R^{\gamma}_{\nu\lambda\mu} + R^{\gamma}_{\lambda\mu\nu})D_\gamma A_\alpha \\
&= -(D_\lambda R^{\gamma}_{\alpha\mu\nu} + D_\nu R^{\gamma}_{\alpha\lambda\mu} + D_\mu R^{\gamma}_{\alpha\nu\lambda})A_\gamma
\end{aligned}$$

Since A_γ is arbitrary, the expression in parentheses is zero. This is the R_{uddd} form of the Bianchi identity.

We can get the R_{dddd} form directly since $Dg = 0$ for the metric tensor g.

$$\begin{aligned}
0 &= D_\lambda R^\gamma_{\alpha\mu\nu} + D_\nu R^\gamma_{\alpha\lambda\mu} + D_\mu R^\gamma_{\alpha\nu\lambda} \\
&= D_\lambda g^{\gamma\beta} R_{\beta\alpha\mu\nu} + D_\nu g^{\gamma\beta} R_{\beta\alpha\lambda\mu} + D_\mu g^{\gamma\beta} R_{\beta\alpha\nu\lambda} \\
&= g^{\gamma\beta}(D_\lambda R_{\beta\alpha\mu\nu} + D_\nu R_{\beta\alpha\lambda\mu} + D_\mu R_{\beta\alpha\nu\lambda})
\end{aligned}$$

This is the R_{dddd} form of the Bianchi identity.

Now we move on. Contracting twice then raising an index tells us that the divergence of $R^{\alpha\beta} - \frac{1}{2}Rg^{\alpha\beta}$ vanishes (note some deliberate index swapping in order to change signs)

$$\begin{aligned}
0 &= g^{\beta\mu}(D_\lambda R_{\beta\alpha\mu\nu} - D_\nu R_{\beta\alpha\mu\lambda} + D_\mu R_{\beta\alpha\nu\lambda}) \\
&= D_\lambda R_{\alpha\nu} - D_\nu R_{\alpha\lambda} + D_\mu g^{\beta\mu} R_{\beta\alpha\nu\lambda}) \\
0 &= g^{\alpha\nu}(D_\lambda R_{\alpha\nu} - D_\nu R_{\alpha\lambda} - D_\mu g^{\beta\mu} R_{\alpha\beta\nu\lambda}) \\
&= D_\lambda R - D_\nu g^{\alpha\nu} R_{\alpha\lambda} - D_\mu g^{\beta\mu} R_{\beta\lambda}) \\
&= D_\lambda R - D_\nu R^\nu_\lambda - D_\mu R^\mu_\lambda) \\
&= D_\lambda R - 2D_\nu R^\nu_\lambda \\
0 &= g^{\mu\lambda}(D_\lambda R - 2D_\nu R^\nu_\lambda) \\
&= D_\lambda g^{\mu\lambda} R - 2D_\nu R^{\nu\mu} \\
&= D_\nu(g^{\mu\nu} R - 2R^{\nu\mu})
\end{aligned}$$

The result of all this is that

$$D_\alpha\left(R^{\alpha\beta} - \frac{1}{2}Rg^{\alpha\beta}\right) = 0$$

so that the *Einstein tensor*

$$G^{\alpha\beta} = R^{\alpha\beta} - \frac{1}{2}Rg^{\alpha\beta}$$

is the suitable left-hand side of the *Einstein equation*

$$G^{\alpha\beta} = 8\pi G_N T^{\alpha\beta}$$

relating the average curvature of timespace to the stress-energy tensor giving the density of matter in that space.

The constant $8\pi G_N$ is very similar to the $4\pi G_N$ of Newtonian gravity, and comes from it so that Einsteinian gravity matches up in the limit of weak gravity and small velocities—i.e., small timespace curvature.

In geometrical units, of course, $G_N = 1$ as well as $c = 1$.

31. Cosmology. Remarkably, the simplest example of the full Einstein equation is the Universe, as modelled by Einstein in 1917, Friedmann in 1922, Lemaître in 1927 and Robertson and Walker independently in 1935.

On the scale of many clusters of galaxies, the universe is modelled as a perfect fluid, having energy density and pressure (the diagonal part of the stress tensor), and with all inhomogeneities smoothed out. As well as homogeneity (same everywhere in space) the universe is assumed to be isotropic (same in all spatial directions). Thus the spatial part of the metric must be flat or have constant curvature, either positive or negative.

Friedmann allowed the universe to be dynamic, i.e., changing in time, according to some function $a(t)$ to be determined. His metric is deceptively simple

$$(\Delta s)^2 = -(\Delta t)^2 + a^2(t)((\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2)$$

or

$$(\Delta s)^2 = -(\Delta t)^2 + a^2(t)((\Delta r)^2 + r^2((\Delta\theta)^2 + s^2(\Delta\phi)^2)$$

for flat space.

For curved space we can pick radii of curvature 1 and -1 because the function $a(t)$ will do the overall scaling. So the coordinates for a 3-sphere and for a 3-hyperboloid are given respectively by $r = \sin\chi$ and $r = \sinh\chi$ where χ is the extra angle (beyond θ and ϕ) needed for the extra dimension.

$$\begin{aligned}
x &= \left\{ \begin{array}{c} \sin\chi \\ \sinh\chi \end{array} \right\} \sin\theta\cos\phi \\
y &= \left\{ \begin{array}{c} \sin\chi \\ \sinh\chi \end{array} \right\} \sin\theta\sin\phi \\
z &= \left\{ \begin{array}{c} \sin\chi \\ \sinh\chi \end{array} \right\} \cos\theta \\
w &= \left\{ \begin{array}{c} \cos\chi \\ \cosh\chi \end{array} \right\}
\end{aligned}$$

So the three metrics are

$$\begin{aligned}
(\Delta s)^2 &= -(\Delta t)^2 + a^2(t)\left((\Delta\chi)^2 + \left\{ \begin{array}{c} \sin^2\chi \\ \chi^2 \\ \sinh^2\chi \end{array} \right\} ((\Delta\theta)^2 + \sin^2\theta(\Delta\phi)^2\right) \\
(\Delta s)^2 &= -(\Delta t)^2 + a^2(t)\left(\frac{(\Delta r)^2}{1-kr^2} + r^2((\Delta\theta)^2 + \sin^2\theta(\Delta\phi)^2\right)
\end{aligned}$$

since

$$\begin{aligned}
\Delta r &= \text{slope}_\chi r\Delta\chi \\
&= \left\{ \begin{array}{c} \cos\chi\Delta\chi \\ \Delta\chi \\ \cosh\chi\Delta\chi \end{array} \right.
\end{aligned}$$

so

$$(\Delta\chi)^2 = \left\{ \begin{array}{c} (\Delta r)^2/(1-\sin^2\chi) \\ (\Delta r)^2 \\ (\Delta r)^2/(1+\sinh^2\chi) \end{array} \right.$$

and we've invented a "signature" for the curvature

$$k = \left\{ \begin{array}{rl} 1 & \text{positive curvature, closed universe} \\ 0 & \text{zero curvature, flat universe} \\ -1 & \text{negative curvature, open universe} \end{array} \right.$$

The function $a(t)$ must be pinned down by the Einstein equation, so we must calculate the curvature tensor and hence the Einstein tensor. The protor calculator (Notes 5 and 20 of Part I) can do this but MATLAB cannot handle the unknown function $a(t)$. So I arbitrarily set $a(t) = bt^2 + dt$ then picked the slopes $\dot{a}(t) = 2bt + d$ and $\ddot{a}(t) = 2b$ out of the results of the curvature calculation, and then translated them back.

The curvature tensor below has been transformed to orthonormal coordinates.

$$\mathbf{e}_{\hat{t}} = (1, 0, 0, 0)$$

$$
\begin{aligned}
\mathbf{e}_{\hat{r}} &= (0, \frac{\sqrt{1-kr^2}}{a}, 0, 0) \\
\mathbf{e}_{\hat{\theta}} &= (0, 0, \frac{1}{ar}, 0) \\
\mathbf{e}_{\hat{\phi}} &= (0, 0, 0, \frac{1}{ar\sin\theta})
\end{aligned}
$$

The `dddd` curvature tensor, leaving out the half of the terms that have different first and third indices (they differ in signs from their counterparts but do not contribute to the Ricci tensor) is

$$
\begin{array}{rccccr}
R_{\mathrm{dddd}}(& \hat{\alpha} & \hat{\beta} & \hat{\gamma} & \hat{\delta} & R \quad) \\
& t & r & t & r & -x \\
& r & t & r & t & -x \\
& t & \theta & t & \theta & -x \\
& \theta & t & \theta & t & -x \\
& t & \phi & t & \phi & -x \\
& \phi & t & \phi & t & -x \\
& r & \theta & r & \theta & y \\
& \theta & r & \theta & r & y \\
& r & \phi & r & \phi & y \\
& \phi & r & \phi & r & y \\
& \theta & \phi & \theta & \phi & y \\
& \phi & \theta & \phi & \theta & y
\end{array}
$$

where $x = \ddot{a}/a$ and $y = (k + \dot{a}^2)/a^2$.

The Einstein tensor can be calculated by hand from this, but here is the new MATLAB function in the protor calculator.

```
% function eins = curv2einstein(metr,curv)              THM              150423
% From curv_dddd and metric find Einstein tensor (adapted to dddd 150427)
% in    metr protor: the uu form of the metric
%       curv protor: the dddd form of curvature (e.g., from metr2curv)
% out   eins protor: the Einstein tensor
function eins = curv2einstein(metr,curv)
% raise first index of curv and contract for ricci
ricci = simplify(contract(simplify(joinred(metr,2,curv,1)),1,3));
% raise first index of ricci and contract for scalar
r = simplify(contract(simplify(joinred(metr,2,ricci,1)),1,2));
eins = simplify(mergesum(1,ricci,-r/2,metr));
```

The result is

$$
G_{\mathrm{dd}} = \begin{pmatrix} 3y & & & \\ & -2x-y & & \\ & & -2x-y & \\ & & & -2x-y \end{pmatrix}
$$

or

$$
\begin{array}{rccc}
G_{\mathrm{dd}}(& \hat{\alpha} & \hat{\beta} & G \qquad) \\
& t & t & 3(k+\dot{a}^2)/a^2 \\
& r & r & -(k+\dot{a}^2+2a\ddot{a})/a^2 \\
& \theta & \theta & -(k+\dot{a}^2+2a\ddot{a})/a^2 \\
& \phi & \phi & -(k+\dot{a}^2+2a\ddot{a})/a^2
\end{array}
$$

(Note that in these calculations, the orthonormalized metric is just $\eta_{\mathrm{dd}} = \mathrm{diag}[-1,1,1,1]$ so raising and lowering indices is just a matter of changing sign if the index is t.)

That's the left-hand side of the Einstein equation.

On the right, we need the stress-energy tensor, which, in orthonormalized coordinates for the ideal fluid with energy density ρ and pressure p, is very simple.

$$T_{\mathrm{dd}} = \begin{pmatrix} \rho & & & \\ & p & & \\ & & p & \\ & & & p \end{pmatrix}$$

So two *Friedmann equations* result

F1: $$3(k+\dot{a}^2)/a^2 = 8\pi\rho$$

F2: $$-(k+\dot{a}^2+2a\ddot{a})/a^2 = 8\pi p$$

We can combine these to get the first law of thermodynamics, so, beyond the first law, we need only one of the two.

For a perfect fluid the first law says

$$\Delta E = -p\Delta\mathcal{V}_3$$

for energy E and 3-volume $\mathcal{V}_3$, where

$$\begin{aligned} E &= \rho\mathcal{V}_3 \\ \mathcal{V}_3 &= a^3\Delta x\Delta y\Delta z \end{aligned}$$

using coordinates x, y, z which "co-move" with the fluid and are independent of time t.

So

$$\Delta(\rho a^3\Delta x\Delta y\Delta z) = -p\Delta(a^3\Delta x\Delta y\Delta z)$$

and we can factor out the common constant

$$\Delta(\rho a^3) = -p\Delta(a^3)$$

i.e., since $\Delta f = \mathrm{slope}_t f\Delta t$

$$\mathrm{slope}_t(\rho a^3) = -p\ \mathrm{slope}_t(a^3)$$

Thus the first law becomes the form we want to derive from the Friedmann equations

$$\dot{\rho}a^3 + 3\rho a^2\dot{a} = -3pa^2\dot{a}$$

or

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho+p)$$

Now let's derive this from the Friedmann equations

F1: $$3(k+\dot{a}^2)/a^2 = 8\pi\rho$$

$\dot{\mathrm{F1}}$:

$$\begin{aligned} 6\dot{a}\ddot{a} &= 8\pi\dot{\rho}a^2 + 16\pi\rho a\dot{a} \\ 4\pi a^2\dot{\rho} &= 3\dot{a}\ddot{a} - 8\pi\rho a\dot{a} \\ \dot{\rho} &= \frac{3}{4\pi}\frac{\dot{a}\ddot{a}}{a^2} - \frac{2}{a}\rho\dot{a} \end{aligned}$$

F2,1:

$$
\begin{aligned}
-\left(2\frac{\ddot{a}}{a}+\frac{k+\dot{a}^2}{a^2}\right) &= 8\pi\rho \\
-\left(2\frac{\ddot{a}}{a}+\frac{8\pi\rho}{3}\right) &= 8\pi\rho \\
\frac{\ddot{a}}{a} &= -\frac{8\pi}{2}\left(\frac{\rho}{3}+p\right)
\end{aligned}
$$

$\dot{\text{F}}1$:

$$
\begin{aligned}
\dot{\rho} &= -\frac{3}{4\pi}\frac{8\pi}{2}\frac{\dot{a}}{a}\left(\frac{\rho}{3}+p\right)-2\frac{\dot{a}}{a}\rho \\
&= -3\frac{\dot{a}}{a}(\rho+p) \\
0 &= \dot{\rho}+3\frac{\dot{a}}{a}(\rho+p)
\end{aligned}
$$

which is the first law.

So, from the Einstein equation for the Friedmann-Lemaître-Robertson-Walker universe we get the first law of thermodynamics and the one equation usually called the Friedmann equation

$$3\frac{k+\dot{a}^2}{a^2}=8\pi\rho$$

Solving this equation is the business of cosmology. The simplest assumption is that the universe has three components: a pressureless "dust" of matter, i.e., galaxies; a uniform radiation pressure (photons and other relativistic particles such as neutrinos); and the volume-independent energy of the "cosmological constant" (Excursion *Divergence-free curvature* in Part I).

With these and the choices of closed, flat and open universe given by the signature k this simple cosmology becomes very rich.

32. Negative pressure and dark energy. Since gravity has no repulsive component but is purely attractive it is evident that $a(t)$ in the previous Note must shrink, and the universe along with it. But the cosmological constant (Excursion *Divergence-free curvature* in Part I) offers a way out, as Einstein intended.

This changes the Einstein equation of Note 30 from

$$G^{\alpha\beta}=8\pi G_N T^{\alpha\beta}$$

to

$$G^{\alpha\beta}+\Lambda g^{\alpha\beta}=8\pi G_N T^{\alpha\beta}$$

where the extra term on the geometry side (left side) does not change the divergence from zero.

The significance of zero divergence is that the source side (right side), the stress-energy tensor, must have zero divergence in order that energy-momentum (energentum) be conserved [Whe90, Ch.7].

So it actually makes some sense to swap the Λg term over to the right-hand side and call it the *vacuum energentum.*

$$G^{\alpha\beta}=8\pi G_N(T^{\alpha\beta}_{\neg\Lambda}+T_\Lambda g^{\alpha\beta})$$

where

$$T_\Lambda = -\frac{\Lambda}{8\pi G_N}$$

and $T^{\alpha\beta}_{\neg\Lambda}$ is the former $T^{\alpha\beta}$, the non-vacuum part of the stress-energy tensor.

We'll now write

$$T^{\alpha\beta} = T^{\alpha\beta}_{\neg\Lambda} + T_\Lambda g^{\alpha\beta}$$

to incorporate both components into the stress-energy tensor.

Now suppose that we are in vacuum only—no other matter or stress-energy—and using orthonormal coordinates so

$$g_{\mathrm{uu}} = \eta_{\mathrm{uu}} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

and

$$T_{\mathrm{uu}} = \begin{pmatrix} \rho_\Lambda & & & \\ & p_\Lambda & & \\ & & p_\Lambda & \\ & & & p_\Lambda \end{pmatrix} = T_\Lambda \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} = T_\Lambda \eta_{\mathrm{uu}}$$

as in Note 31 with ρ_Λ the vacuum energy density and p_Λ the vacuum pressure. (Why does $T_{\mathrm{uu}} = T_{\mathrm{dd}}$?)

Thus

$$\rho_\Lambda = -T_\Lambda = \frac{\Lambda}{8\pi G_N} = -p_\Lambda$$

That the vacuum pressure is the negative of the vacuum energy density (which is non-negative) is very significant, and we can see this in two other ways.

First, we use the first law of thermodynamics, adapted from Note 31 with $\rho =$ constant: $3\rho a^2 \dot{a} = -3pa^2\dot{a}$ so $\rho = -p$.

Second, we consider what must happen to a system with constant energy density. If we expand the volume, the energy will increase, so that takes work against some force, which must be "pressure" resisting expansion—a negative pressure contrary to the usual sense in which pressure *causes* expansion.

We must now see how negative pressure, counterintuitively, gives rise to a *repulsive* force. This is a relativistic effect, due to the pressure contribution to gravity via the stress-energy tensor.

In the Newtonian limit, this relativistic effect changes the gravity equation from

$$\nabla^2\phi = 4\pi G_N \rho$$

to

$$\begin{aligned} \nabla^2\phi &= 4\pi G_N(\rho + 3p) \\ &= 4\pi G_N(\rho - 3\rho) \\ &= -8\pi G_N \rho \\ &= -\Lambda \end{aligned}$$

for pressure p. (Excursion *Newtonian limit* finds the $\rho + 3p$ term I've used here.) $\bigtriangledown^2$ is for us inessential notation for the second slope in three spatial dimensions.

Since the second slope $\bigtriangledown^2$ of the potential ϕ is a negative constant, ϕ must itself be a downward-pointing parabola as a function of the radius r.

This potential sloped down everywhere and so gives rise to a repulsive force, and one that increases in magnitude with r.

Einstein used the cosmological constant in this way to overcome gravitational attraction in order to force the universe to be static, as everybody then believed it was. When Hubble in 1929 showed that the universe is not static but expanding[1] and Eddington in 1930 proved that Einstein's "static" universe is unstable. Einstein told George Gamow that he had blundered, but the blunder, if any, was to miss predicting a dynamic universe (which Friedmann did in 1922) rather than introducing the cosmological constant. The repulsive effect of a constant vacuum energy density has come to be known as "dark energy" and observations show that about 18 times more dark energy is needed than all the energy residing in matter that we can see, Hence the adjective "dark"—but bear in mind that the phrase is just another name for something we don't know.

33. Gravitational irreversibility. In Notes 25 and 27 we found that Schwarzschild and Kerr black holes, respectively, can in principle release 6% and 42% of the rest energy of an infalling body. In 1969 Roger Penrose discovered a mechanism for extracting the rotational energy of a Kerr black hole.[2]

While the Penrose process may not account for much of the astronomically observed black-hole energy output (see Excursion *Blandford-Znajek mechanism*) it is theoretically significant because it leads to four laws of black hole dynamics analogous to the four laws of thermodynamics.

Penrose imagined a body falling into the ergosphere (see Excursion *Ergosphere*) of a rotating black hole and splitting into two. One of the two pieces is ejected but the other falls in. (Wheeler [Whe90, p.215] imagines a body carrying a "sacrificial victim". Misner, Thorne and Wheeler [MTW73, p.908] imagine a super-tech garbage disposal system: it's best to dump our trash where we are ignorant of the effects of doing so, and we are in principle—for now—ignorant of the innards of a black hole.)

Since energentum is conserved

$$\mathbf{p}_{\text{in}} = \mathbf{p}_{\text{out}} + \mathbf{p}_{\text{sacr}}$$

the symmetry vector $\xi = (1, 0, 0, 0)$ (Note 25) allows us to say

$$\mathbf{p}_{\text{out}} \cdot \xi = \mathbf{p}_{\text{in}} \cdot \xi - \mathbf{p}_{\text{sacr}} \cdot \xi$$

and these are energies, at least they are *outside* the ergosphere.

In Excursion *Ergosphere* we saw that *inside* the ergosphere

$$\xi \cdot \xi = g_{tt} > 0$$

so ξ is spacelike, not timelike (as is the four-velocity component $\mathbf{u^t}$ in the Excursion).

This means that $E_{\text{sacr}} = -\mathbf{p}_{\text{sacr}} \cdot \xi$ is not an energy but momentum and so can be either positive or negative.

If it is negative then

$$E_{\text{out}} = E_{\text{in}} - E_{\text{sacr}} > E_{\text{in}}$$

[1] Hubble used Cepheid variable stars, whose brightness is correlated to the rate at which they pulse, to discover a linear relationship between their distance and the amount by which their light is redshifted due to their motion away from us. In 1998 and 1999 Reiss, Schmidt et al., and Perlmutter reported that the very distant galaxies are accelerating away from us. They observed redshifts from type 1a supernovae, which have fixed brightness some million times that of the Cepheid variables.

[2] In this Note and the next I am exploring somewhat beyond the domain of the present Book and so will break my habit of restricting citations to the Excursions.

What the Penrose process has done has been to obtain energy by reducing the rotational angular momentum of the black hole: that negative linear momentum of the sacritifical body opposes the black-hole rotation.

So Penrose gives us a means to extract energy from a rotating black hole, which we can do until finally the rotation stops, the ergosphere collapses, and the process stops. But we could drop more bodies in, in the normal way without extracting energy, to replace the black-hole mass (= energy lost) and bring its rotation back up.

Could we continue to extract energy in this way without net change to the black hole? Demetrios Christodoulou [Chr70] in 1970 found that, like thermodynamic processes, the Penrose process is irreversible, except for a limiting reversible case, like the equilibrium case of thermodynamic (Book 9c, Parts II,III).

What we need is an inequality relating mass change ΔM with angular momentum change ΔL, say

$$\Delta M \geq \Omega_H \Delta L$$

At the horizon we can find

$$M = E = \Omega_H L$$

where Ω_H describes the angular velocity of the horizon—which means the angular velocity of light rays trapped on the horizon.

Thus on the horizon (well, I glossed over a subtlety: see [Har03, p.325])

$$\Delta M = \Omega_H \Delta L$$

and since Ω_H is the maximum possible angular velocity, anywhere outside the horizon

$$\Delta M > \Omega_H \Delta L$$

This angular velocity Ω_H appears in the tangent vector to the horizon

$$\ell_{\mathrm{u}} = (1, 0, 0, \Omega_H)$$

and ℓ must be a null vector

$$0 = \ell_{\mathrm{u}} \cdot \ell_{\mathrm{u}} = \ell^\alpha g_{\alpha\beta} \ell^\beta$$

since it describes light rays.

Focussing on the equatorial plane of the Kerr black hole, $\theta = \pi/2$ so $s = \sin\theta = 1$ and $c = \cos\theta = 0$ (so $\rho^2 = r^2$ in Note 26), and on the outer horizon $r_+ = M + \sqrt{M^2 - a^2}$ the metric is

$$(\Delta s)^2 = -\left(1 - \frac{2M}{r_+}\right)(\Delta t)^2 - \frac{4Ma}{r_+}(\Delta)(\Delta\phi) + \left(r_+^2 + a^2 - \frac{2Ma^2}{r_+}\right)(\Delta\phi)^2$$

It is useful to know that

$$\begin{aligned} r_+^2 + a^2 &= M^2 + 2M\sqrt{M^2 - a^2} + M^2 \\ &= 2M(M + \sqrt{M^2 - a^2}) \\ &= 2Mr_+ \end{aligned}$$

So we have

$$\begin{aligned} 0 &= \ell^\alpha g_{\alpha\beta} \ell^\beta \\ &= \left(\frac{2M}{r_+} - 1\right) - \frac{4Ma}{r_+}\Omega_H + \left(2Mr_+ + \frac{2Ma^2}{r_+}\right)\Omega_H^2 \end{aligned}$$

and if we try $\Omega_H = a/2Mr_+$

$$\begin{aligned}
0 &= 2Mr_+ - r_+^2 - 4Mar_+\frac{a}{2Mr_+} + 2M(r_+^2+a^2)r_+\left(\frac{a}{2Mr_+}\right)^2 \\
&= 2Mr_+ - r_+^2 - 2a^2 + (2Mr_+)^2\left(\frac{a}{2Mr_+}\right)^2 \\
&= 2Mr_+ - 2Mr_+
\end{aligned}$$

On the horizon, then

$$\Delta M = \Omega_H \Delta L = \frac{a}{2Mr_+}\Delta L$$

Christodoulou introduced a quantity called the *irreducible mass*

$$M_{\rm ir}^2 = \frac{1}{2}Mr_+ = \frac{1}{2}M(M+\sqrt{M^2-a^2})$$

which is exactly the mass^2 of a Schwarzschild black hole where $a=0$. The change in this as M and $L = aM$ change is

$$\Delta M_{\rm ir} = \partial_M M_{\rm ir}\Delta M + \partial_L M_{\rm ir}\Delta L \geq 0$$

We can see this inequality by taking the slope

$$\begin{aligned}
\partial_M r_+ &= \partial_M\left(M+\sqrt{M^2-\frac{L^2}{M^2}}\right) \\
&= 1 + \frac{M + L^2/M^2}{\sqrt{M^2-a^2}} \\
\partial_L r_+ &= \partial_L\left(M+\sqrt{M^2-\frac{L^2}{M^2}}\right) \\
&= -\frac{L/M^2}{\sqrt{M^2-a^2}}
\end{aligned}$$

so

$$\begin{aligned}
2M_{\rm ir}\partial_M M_{\rm ir} &= \frac{1}{2}r_+ + \frac{1}{2}M\partial_M r_+ \\
&= \frac{1}{2}\left(M+\sqrt{M^2-a^2}+M+\frac{M^2+a^2}{\sqrt{M^2-a^2}}\right) \\
&= \frac{1}{2}\left(\frac{2M\sqrt{M^2-a^2}+M^2-a^2+M^2+a^2}{\sqrt{M^2-a^2}}\right) \\
&= \frac{Mr_+}{\sqrt{M^2-a^2}}
\end{aligned}$$

and

$$\begin{aligned}
2M_{\rm ir}\partial_L M_{\rm ir} &= \frac{1}{2}M\partial_L r_+ \\
&= -\frac{1}{2}\frac{a}{\sqrt{M^2-a^2}}
\end{aligned}$$

Finally

$$\begin{aligned}
\Delta M_{\text{ir}} &= \frac{1}{2M_{\text{ir}}\sqrt{M^2-a^2}}\left(Mr_+\Delta M - \frac{a}{2}\Delta L\right) \\
&= \frac{MR_+}{2M_{\text{ir}}\sqrt{M^2-a^2}}\left(\Delta M - \frac{a}{2Mr_+}\Delta L\right) \\
&\geq 0
\end{aligned}$$

from the inequality

$$\Delta M \geq \frac{a}{2Mr_+}\Delta L$$

Christodoulou's result is that *the irredicuble mass of a black hole never decreases.* This is exactly analogous to the second law of thermodynamice, which says that entropy never decreases. Given the presence of irreversibility we should expect an entropy-like quantity to underly it.

We can anticipate the next bit of history by relating the irreducible mass to the surface area of the horizon.

Given the line element for a sphere we can directly find its surface area

$$\begin{aligned}
(\Delta\Sigma)^2 &= r^2(\Delta\theta)^2 + r^2 s^2(\delta\phi)^2 \\
A &= \text{antislope}_{\theta=0:\pi}(r\ \text{antislope}_{\phi=0:2\pi}(rs)) \\
&= 2\pi r^2 \text{antislope}_{\theta=0:\pi}\sin\theta \\
&= -2\pi r^2\cos\theta\ |_0^\pi \\
&= 4\pi r^2
\end{aligned}$$

For the $r = r_+$ horizon of a Kerr black hole $\Delta r = 0$ and at any fixed time $\Delta t = 0$ so

$$\begin{aligned}
(\Delta\Sigma)^2 &= \rho_+^2(\Delta\theta)^2 + \left(r_+^2 + a^2 + \frac{2Mr_+a^2s^2}{\rho_+^2}\right)s^2(\Delta\phi)^2 \\
&= \rho_+^2(\Delta\theta)^2 + \left(\frac{2Mr_+\rho_+^2 + 2Mr_+a^2s^2}{\rho_+^2}\right)s^2(\Delta\phi)^2 \\
&= \rho_+^2(\Delta\theta)^2 + \frac{2Mr_+}{\rho_+^2}(r_+^2 + a^2)s^2(\Delta\phi)^2 \\
&= \rho_+^2(\Delta\theta)^2 + \left(\frac{2Mr_+}{\rho_+^2}\right)^2 s^2(\Delta\phi)^2
\end{aligned}$$

where $\rho_+^2 = r_+^2 + a^2c^2$ and, of course, $c^2 + s^2 = \cos^2\theta + \sin^2\theta = 1$.

So

$$\begin{aligned}
A &= \text{antislope}_{\theta=0:\pi}\left(\rho_+\text{antislope}_{\phi=0:2\pi}\left(\frac{2Mr_+}{\rho_+}s\right)\right) \\
&= 4\pi Mr_+\text{antislope}_{\theta=0:\pi}\sin\theta \\
&= 8\pi Mr_+ \\
&= 16\pi M_{\text{ir}}^2
\end{aligned}$$

(In the Schwarzschild case $a = 0$ and $A = 4\pi(2M)^2$, in agreement with the horizon radius $2M$: the geometry is Euclidean at least in the θ-Φ direction.)

So Christodoulou's result is that *the area of the horizon of a black hole never decreases.*

Christodoulou did not put it this way but Jacob Bekenstein, another Princeton graduate student, did, after a discussion in which Wheeler wondered about the disappearance of the entropy of combining cold tea with hot tea by dropping both cups into a black hole [Whe90, p.221].

Bekenstein returned to Wheeler several months after this 1970 conversation and said that the entropy does not disappear: the surface area of the horizon *is* the entropy of the black hole [containing, somehow, the information about everything that fell in to build the black hole].

Since entropy implies temperature and temperature implies radiation, but black holes supposedly don't radiate, Brandon Carter and Stephen Hawking were dubious. But to his surprise Hawking found a mechanism for radiating, rather like the Penrose process but for any black hole.

Hawking radiation requires quantum physics—it is a step in the attempt to link general relativity with quantum theory—but we can sketch the idea.

Heisenberg's uncertainty principle, $\Delta E \Delta t \geq \hbar$, allows spontaneous creation of particles of energy ΔE as long as they are gone again in time $\Delta t = \hbar/\Delta E$. This can happen with a matter-antimatter pair, such as an electron and a positron, which appear then immediately annihilate each other. But if they appear at a black hole horizon one particle may head across the horizon and so never return to be annihilated, leaving the other particle free to escape: Hawking radiation.

Hawking's quantum analysis was able to find the constant (4) for the relationship between entropy and horizon area

$$S_H = \frac{k_B}{4\hbar} A$$

where Bekenstein's dimensional analysis found only the Boltzmann constant k_B and Planck's constant $\hbar$.

This entropy is the Bekenstein-Hawking entropy, The associated Hawking temperature (for a spherical black hole) is

$$k_B T = \frac{c^3 \hbar}{8\pi G_N M}$$

It is inversely proportional to the mass and so very very small for ordinary black holes of several to several billion solar masses. (For this equation, $c \neq 1 \neq G_N$.)

But very tiny black holes, which *may* have been created by the Big Bang, can be hot, and so radiate away mass with increasing rapidity, becoming ever hotter until they explode.

Hartle [Har03, §13.3] discusses Hawking radiation and black hole evaporation.

The connection between "black hole dynamics" and thermodynamics is evidently very strong. In 1973, Bardeen, Carter and Hawking [MTW73, p.888] formulated the "four laws of black hole dynamics" here taken from Starinets [Sta14]

	Black holes		Thermodynamics
0	The horizon of a stationary black hole has constant surface gravity	0	The system in thermal equilibrium has a constant temperature T
1	In perturbations of stationary black holes, the change of mass M is related to the change of charge Q, angular momentum J and horizon area A by $dM = (k/8\pi)dA + \Omega dJ + \Phi dQ$	1	In thermodynamic processes the change of energy E is related to the change of entropy S (plus relevant work terms) by $dE = TdS - pdV - \mu dN$
2	The horizon area A is a non-decreasing function of time	2	The entropy S is a non-decreasing function of time
3	It is impossible to achieve zero surface gravity by a physical process	3	It is impossible to achieve zero temperature by a physical process

Misner, Thorne and Wheeler discuss the proof of the second law in detail [MTW73, §34.5]. They note two major assumptions in the proof.

First, the proof assumes no negative energy density. It seems that we can have warp drives (or wormholes) or the second law but not both (see Excursion *Negative energy: warp drives*).

Second, it assumes cosmic censorship [Har03, pp.310,311,319] which conjectures that there are no "naked singularities": the "singularities" at the centres of black holes, or at the beginning of the Big Bang, are all hidden decently away behind horizons. (Is that a problem for the Big Bang?)

In 1993 Gerard t'Hooft pointed out that the two-dimensional nature of black hole entropy might have broader significance, leading to the "holographic principle" of t'Hooft and Leonard Susskind: states in a region of timespace can be reproduced by information residing on its surface boundary [Che10, p.175].

34. Alternatives to geometry. Geometry was central to Einstein's conception of gravity and to our discussion here so far, But we know that Einstein's theory is incomplete—it does not account for 96% of the energy of the cosmos ("dark matter" (20%) is not included at all, and the "dark energy" (76%), related to the cosmological constant, is not explained); it does not describe the centres of black holes except as "singularities" or the beginning of the Big Bang; it does not mesh with quantum physics or unify with the three forces of the "Standard Model" of particle theory—and alternatives have been proposed. So it is worth while to look at five of these.

a) *Whitehead.* The very distinguished mathematician and philosopher Alfred North Whitehead (1861–1947) reacted almost immediately to general relativity in three books published in 1919–22 [Whi19, Whi20, Whi22] (but in no journal of physics that I know of). He says [Whi20, p.181]

> Where I differ from Einstein is that he conceives of this quantity which I call the impetus as merely expressing the characteristics of the space and time to be adopted and thus ends by talking of the gravitational field expressing a curvature in the space-time manifold. I cannot attach any clear conception to his interpretation of space and time. My formulae differ slightly from his though they agree in those instances where his results have been verified.

He says [Whi22, p.v]

> I should be very willing to believe that each permanent space is either uniformly elliptic or uniformly hyperbolic, if any observations are more simply explained by such a hypothesis.

as opposed to the Euclidean space he uses in his theory: his philosphical objection to general relativity is its central conception that "mass grips spacetime, telling it how to curve" (and "spacetime grips mass, telling it how to move") [Whe90, pp.12,11].

Whitehead's theory itself is presented in chapter IV of [Whi22] and elaborated in specific examples in the next thirteen (brief) chapters. Whitehead derives the three predictions that were used at the time to check Einstein's theory—the precession of Mercury, the bending of light, and the gravitational redshift—but his theory goes on to find gravitational effects on electromagnetic phenomena, and so makes further predictions, notably about spectral lines from the Sun, including the "limb effect" (the spectra from the centre of the solar disk are not the average of the spectra from the "limbs", the points on the solar equator at its two visible end points, where the surface is rotating respectively towards and away from us).

Because of Whitehead's eschewal of geometry he is able to give a very nice formal presentation of tensors [Whi22, Part 3] but I find his physical theory quite unintuitive. It was translated in 1951 by John Leighton Synge [Syn51] in three lectures at the University of Maryland. (Synge there points out some attention paid by Eddington in 1924[3] when he pointed out that the Schwarzschild metric could be derived from Whitehead.

Apart from Eddington's and Synge's brief attentions, physicists seem to have ignored Whitehead's "principle of relativity". The theory does not lead to Einstein's as a special case, so it disagrees with Einstein, whose general relativity has led us so far and has not yet been found wanting in any specific test.

But since we now know Einstein is incomplete, Whitehead may deserve further attention. I don't think he's going to get it.

The problem is cultural more than simple testability of two competing theories. As a mathematician, Whitehead would want a finished formalism, free of any scaffolding used to build it, and this obscures the motivating ideas. As a philosopher, he would want all ambiguities eliminated before beginning to theorize. But science doesn't work in these ways, needing intuitive bases for intellectual reproducibility and overcoming apparent ambiguities by dialogue. Such cultural differences, like ideological disagreements, are seldom resolved by rational debate, and must usually wait for opponents to die off.

b) *Feynman.* In 1995 a course on gravitation given by Richard Feynman at Caltech in 1962–3 was finally published [FMW03]. Eschewing Einstein's geometrical approach because "today, physics students know about quantum theory and mesons and the fundamental particles, which were unknown in Einstein's day", Feynman imagined that on "say a planet such as Venus, we have scientists ... who know just what to do about nucleons, mesons, etc., but who do not know about gravitation".

"Now what would the Venutians do", Feynman asks, with a sudden discovery of gravity? "They would probably try to interpret it in terms of the field theories which are familiar to them".

So in lecture 3 he starts with a particle, the graviton, a hypothetical carrier of the gravitational attraction, and argues that it must have spin 2. He builds a theory from this and argues that gravitons must couple through what looks like the trace of a stress-energy tensor. Lecture 5 works out the deflection of light and the gravitational red shift, but not quite the precession of Mercury. The latter is fixed in lecture 6, which concludes with the Einstein equation finally arising from all this discussion of (special) relativistic quantum field theory.

Although he starts without geometry, Feynman does come to Einstein's general relativity in lecture 6 and to the possibility of a geometric interpretation in lecture 8.

His motivation in all this is to find a quantum theory of gravity, something he did not succeed in doing, and he approved for publication only 16 of the 27 lectures he actually gave.

Feynman was not the first to try this route: in their Foreword John Preskill and Kip Thorne describe predecessors Fierz and Pauli (1939), Gupta (1954) and Kraichnan (1955,1956).

Subsequently, Steven Weinberg's text [Wei72] also avoids geometry in order to find a better fit with quantum physics: "I believe that the geometrical approach has driven a wedge between general relativity and the theory of elementary particles". But Weinberg also fully includes Einstein's theory in his results.

c) *Jacobson.* Also in 1995 Ted Jacobson published a letter [Jac95] which also derives Einstein's equations from a non-geometrical starting point, and is much easier to follow than Feynman, given our background in relativity (these Notes) and in thermodynamics (Book 9c).

Jacobson speaks for himself.

[3]By 1927 Eddington had reservations about Whitehead's "Principle of Relativity" [Edd30, p.146]. He also remarks, with cheerful lack of angst, on ambiguities in science, mentioning on p.194 William Bragg's comment that physicists are classical on Mondays, Wednesdays and Fridays but quantum on Tuesdays, Thursdays and Saturdays, adding that "Perhaps that ought to make us feel a little sympathetic towards the man whose philosophy of the universe takes one form on weekdays and another form on Sundays."

> How did classical General Relativity know that horizon area would turn out to be a form of entropy, and that surface gravity is a temperature? In this letter I will answer that question by turning the logic around and deriving the Einstein equation from the proportionality of entropy and horizon area together with the fundamental relation $dQ = TdS$ connecting heat Q, entropy S, and temperature T

Black hole horizons are "causal horizons" in that we cannot know what lies behind them. So are light cones, or any null surface—a surface along which light travels and so has zero proper time. Jacobson:

> In thermodynamics, heat is energy that flows between degrees of freedom that are not macroscopically observable. In spacetime dynamics, we shall define heat as energy that flows across a causal horizon. It can be felt via the gravitational field it generates, but its particular form or nature is unobservable from outside the horizon. .. This sort of horizon is a null hypersurface .. The outside world is separated from the system not by a diathermic wall, but by a causality barrier.
> That causal horizons should be associated with entropy is suggested by the observation that they hide information. In fact, the overwhelming majority of the information that is hidden resides in correlations between vacuum fluctuations just inside and outside of the horizon.

He goes on to challenge continuity in field theory.

> Because of the infinite number of short wavelength field degrees of freedom near the horizon, the associated "entanglement entropy" is divergent in continuum quantum field theory. If, on the other hand, there is a fundamental cutoff length ℓ_c, then the entanglement entropy is finite and proportional to the horizon area in units of ℓ_c^2, as long as the radius of curvature of spacetime is much longer than ℓ_c.

He is arguing here not for a completely discrete quantum field theory, but at least one with a low-wavelength cutoff.

A remarkable conclusion is that gravity is not a fundamental force. A sound

> wave is a travelling disturbance of local density, which propagates via a myriad of incoherent collisions. Since the sound field is only a statistically defined observable on the fundamental phase space of the multiparticle system, it should not be canonically quantized as if it were a fundamental field, even though there is no question that the individual molecules are quantum mechanical. By analogy, the viewpoint developed here suggests that it may not be correct to canonically quantize the Einstein equations, even if they describe a phenomenon that is ultimately quantum mechanical.

Erik Verlinde [Ver10] subsequently took this argument further, to any 2-dimensional (holographic) surface (" by thinking about accelerated observers, one can in principle locate holographic screens anywhere in space") and to the assertion that gravity (and hence timespace) is *emergent* ("On one side [of the surface] there is space, on the other side nothing yet.") Verlinde explicitly argues that gravity is no longer a fundamental force. He gives as a mechanism the "entropic force", illustrated by polymer molecules giving rise to Hooke's law of elasticity, and relates these to the thought experiment that led Bekenstein to link entropy with horizon surface area. Verlinde uses Jacobson's argument in his final steps of deriving Einstein from these entropic considerations.

I wonder if an entropic basis to gravity will establish Hawking's "chronology protection" conjecture that "the laws of physics prevent the formation of time machines for backward time travel" [ER12, p.191].

d) *Sen and Ashtekar.* In 1986 Abhay Ashtekar published [Ash86] a reformulation of general relativity in Hamiltonian form (Note 39 of Book 8c—Part IV), which makes it amenable to quantization.

The paper is highly mathematical, involving SU(2) symmetry, soldering forms, cotangent bundles and Yang-Mills theory. Smolin [Smo01, pp.125,40,224] discusses the effect of this and two companion papers, and its inspiration by Amitaba Sen.

e) *Barbour.* Smolin [Smo13, p.168] describes a theory resulting from work by Julian Barbour and Niall Ó Murchoda which is dual to general relativity (the way particles are dual to waves in quantum physics—see Week 5—i.e., equivalent yet complementary, but in a way also conducive to quantization) in which size is relative instead of simultaneity. In this dual, called *shape dynamics*, it is possible to have a single time for all observers.

35. Summary.

(These notes show the trees. Try to see the forest!)

Einstein decided that it could not be a coincidence that inertial mass (m in Newton's $F = ma$) is always identical to gravitational mass (m in Newton's $F = G_N Mm/r^2$ or $F = mg$) and so he declared this to be the *principle of equivalence.* It then follows that acceleration and gravitation are indistinguishable from each other, another manifestation of the principle of equivalence.

Einstein also knew that his special relativity ruled out Newton's instantaneous action-at-a-distance as the mechanism for gravitational attraction, because no influence can exceed the speed of light.

Special relativity also told him that geometry could serve as the intermediary for gravity—or acceleration—when he thought of a merry-go-round. Its geometry could not be Euclidean because the circumference, moving at a constant (circular) velocity would cause the circumference to contract (in the direction of motion) and so become less than π times the diameter.

So he imagined gravity—or that centrifugal acceleration of the merry-go-round—bending space itself. Since special relativity combines space with time, gravity necessarily bends timespace as a whole.

The math and the details occupied him for a decade. He took up Gauss' and Riemann's study of manifolds (curved spaces that are everywhere locally flat) and used free-fall as the locally flat frame. Gravity curves timespace and two or more test bodies will show this curvature as motion and as tides. Mass causes the curvature, which is zero on average where there is no matter.

The universe as a whole can be considered curved. In simple models that curvature is constant, so it can be positive, like the two-dimensional surface of a sphere, and finite even though unbounded; or negative, like a saddle surface. We cannot imagine a four-dimensional extension of the surface of a ball, or even a three-dimensional one, but the mathematics of Gauss and Riemann can cope.

We discussd the math in Part I. Part II starts by motivating curved timespace. It discusses three cases of matterless curvature (constant gravity which is flat, spherical gravity, and gravity of rotating spheres), looking at orbits of planets and of light, and looking at tides. We move on to the source of gravity and Einstein's completion of his work, and its simplest application, which is to the universe. We discuss some technological possibilities and we look beyond geometry as will apparently be necessary in the challenge to combine gravity with quantum physics.

II. **The Excursions**

You've seen lots of ideas. Now *do* something with them!

1. a) Is timespace curved by the Schwarzschild metric in 2D

$$g_{\mathrm{dd}} = \begin{pmatrix} -(1-2M/r) & \\ & 1/(1-2M/r) \end{pmatrix}?$$

   b) What about a variant of the constant-gravity metric discussed in Note 23

$$g_{\mathrm{dd}} = \begin{pmatrix} -(1+2gz)^2 & \\ & 1/(1+2gz)^2 \end{pmatrix}?$$

   c) How does the geodesic of this latter differ from that of Note 23 (compare plots)?

2. **Schwarzschild metric.** Lillian and Hugh Lieber [LL45, pp.215–55] provide an accessible derivation of the Schwarzschild metric—the timespace metric found in Note 24—from the vanishing of the Ricci tensor, i.e., the reverse of the argument in Note 24.
   They start by writing down a parametrized metric which must have spherical symmetry (p.233)
$$(\Delta s)^2 = -e^{\nu}(\Delta t)^2 + e^{\lambda}(\Delta r)^2 + e^{\mu} r^2((\Delta\theta)^2 + s^2(\Delta\phi)^2)$$
   where ν, λ and μ are all functions only of r.
   They invoke the *isotropy* of space (same in all directions) and an assumption that the physics is *static* to argue that there can be no cross terms ($\Delta r\Delta\theta$ or $\Delta\phi\Delta t$) in this metric. (Actually this metric could be generalized slightly so that ν, λ and μ are also functions of t.)
   They immediately collapse $e^{\mu}r^2$ into just r^2 by absorbing the effect of μ into λ (p.236).
   They then work directly with the expansion of the curvature via the affine connection to the metric. They use the notation of Eddington [Edd23] and I should here note the links between their notation and ours.

theirs	ours
$\{\mu\nu, \lambda\}$	$\Gamma^{\lambda}_{\mu\nu}$
$B^{\alpha}_{\sigma\tau\rho}$	$R^{\alpha}_{\sigma\tau\rho}$
$G_{\sigma\tau}$	$R_{\sigma\tau} = R^{\alpha}_{\sigma\alpha\tau}$

   Their $(\Delta s)^2$ is the negative of ours, and they use indices {1,2,3,4} where we use $\{r, \theta, \phi, t\}$ respectively. On p.247 they have

$$
\begin{array}{lllll}
\Gamma_{\mathrm{udd}}(\lambda & \mu & \nu & \Gamma &) \\
t & t & r & \nu'/2 & \\
r & r & r & \lambda'/2 & \\
r & t & t & e^{\nu-\lambda}\nu'/2 & \\
r & \theta & \theta & -e^{-\lambda}r & \\
r & \phi & \phi & -e^{-\lambda}rs^2 & \\
\theta & r & \theta & 1/r & \\
\theta & \phi & \phi & -cs & \\
\phi & r & \phi & 1/r & \\
\phi & \theta & \phi & c/s &
\end{array}
$$

where I've not repeated values that can be found by symmetry, where $\nu' = \partial_r\nu(r)$ and $\lambda' = \partial_r\lambda(r)$, and where $c = \cos\theta$ and $s = \sin\theta$.
You can compare these results with [Har03, p.546] which includes a time dependence in ν and λ as well.
Lieber and Lieber show that R_{dd} is symmetrical and that it boils down to five nonzero terms (the diagonal and $R_{r\theta}$), then just to the diagonal and then they note that $R_{\phi\phi} = s^2 R_{\theta\theta}$ (p.250).
They now set the final three terms of R_{dd} to zero, according to Einstein's law of gravity.
So first we have

$$0 = e^{\lambda-\nu}R_{tt} + R_{rr}$$

giving

$$\lambda' = -\nu'$$

so

$$\lambda = \text{const} - \nu$$

and the fact that timespace must be flat far from the origin sets that arbitrary constant to 0 (p.252).
So we have the result that $g_{rr} = -1/g_{tt}$ which we have already used in Notes 23 and 24.
Next, from $0 = R_{\theta\theta}$ they show

$$e^{\nu} = 1 - \text{const}/r$$

and they set that constant of integration to $2M$ (p.253).
This completes the argument that reverses the argument of Note 24.
(Note that MATLAB does not support slopes of unspecified functions, so to run the Lieber calculation on `metr2curv()` you'll have to "kludge", e.g., by pretending $\lambda(r) = ar^2$ and $\nu(r) = br^2$, say, and then reconstructing $\lambda', \lambda", \nu', \nu"$ from the results.)

3. a) Work through Lieber [LL45, pp.267–72] on the precession of Mercury. b) Work through Lieber [LL45, pp.289–6] on the deflection of light.

4. How do the orbits of planets and the bending of light by the gravity of the Sun show us that time must be involved in the curvature of timespace? Hint: consider the differences in velocities of planets and light.

5. What is the mass of the Earth (5.97×10^{24} kg) in units of length?

6. The claim in Note 25 that (symmetry vector)·(4-velocity) is a conserved quantity comes from, e.g., Hartle [Har03, p.177]. The demonstration depends on describing the geodesic ("the shortest distance between two points") using the Euler-Lagrange equation (Note 37 of Book 8c).

7. Hartle [Har03, Ch.24] discusses the astrophysics from Chandrasekhar white dwarf stars to the maximum size of a neutron star at ≈ 7 solar masses (Note 25).

8. Look up Roy Kerr (1934–). Fulvio Melia's account [Mel09] conveys his efforts at the University of Texas in 1963. (Note 26.)

9. **The ergosphere.** While it is in principle possible to hover, with constant r, θ and ϕ, arbitrarily close to the horizon of a Schwarzschild black hole, frame dragging (Note 27) of a Kerr black hole makes this impossible.
   The Kerr horizons at

$$\begin{aligned} 0 &= \Delta = r^2 - 2Mr + a^2 \\ r_\pm &= M \pm \sqrt{M^2 - a^2} \end{aligned}$$

   are double: inner r_- and outer r_+. We will try to hover at the outer one.
   Hovering requires that u^r, u^θ and u^ϕ all vanish, leaving only

$$-1 = u^t \cdot u^t = -\left(1 - \frac{2Mr}{\rho^2}\right) \mid u^t \mid^2$$

   which requires that $-g_{tt} = (1 - 2Mr/\rho^2)$ be non-negative, so

$$2Mr \leq \rho^2 = r^2 + a^2 c^2$$

   and for equality

$$r_{E\pm} = M \pm \sqrt{M^2 - a^2 c^2}$$

   So this surface is also double, the outer one being outside r_+ and the inner one inside r_- (Cheng [Che10, §8.4.2] has a helpful picture).
   For $r_{E-} < r < r_{E+}$, $-g_{tt}$ will be negative and u^t is not timelike and so impossible. Clearly

$$r_+ = M + \sqrt{M^2 - a^2} \geq M + \sqrt{M^2 - a^2 c^2} = r_{E+}$$

   since $c^2 = \cos^2\theta \leq 1$, so we cannot get down to r_+ except at the poles where $c^2 = 1$.
   r_{E+} is the outer surface of what is called the "*ergosphere*".
   Check Cheng's argument that the physical singularity from a Kerr black hole is a *ring* of radius a inside r_{E-}.
   Check Cheng's discussion of the *Penrose process* for extracting energy from a rotating black hole. Hartle [Har03, §15.5] shows that the surface area of a black hole does not decrease, even when it gives up rotational energy through the Penrose process.

10. Explain the $e = 0.9$ plots for the effective equatorial Kerr potential of Note 27.

11. In Note 28 to find the orthonormal basis for radial freefall we didn't need to use $\mathbf{D_u e}_{\hat{\alpha}} = 0$ but we can check that it does using `abslopes()`. If `xSchwFreeFall_t`, `xSchwFreeFall_r`, `xSchwFreeFall_th` and `xSchwFreeFall_ph` are the protor forms of $\mathbf{e}_{\hat{\tau}}, \mathbf{e}_{\hat{r}}, \mathbf{e}_{\hat{\theta}}$ and $\mathbf{e}_{\hat{\phi}}$ respectively we get zeros by running
    `joinred(xSchwFreeFall_t,1,abslopes(`x`,[r,theta],[1],connSchwarz),1)`
    where `connSchwarz` is the Schwarzschild affine connection (Note 24) and x is, in turn, each of `xSchwFreeFall_t`, `xSchwFreeFall_r`, `xSchwFreeFall_th` and `xSchwFreeFall_ph`.

12. Use `joinred()` in the protor calculator to check that the basis for radial freefall given in Note 28 for the Schwarzschild metrc is indeed orthonormal, giving the Minkowski metric of special relativity

$$\begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

13. The simple orthonormal basis for the Schwarzschild metric

$$\begin{aligned}
\mathbf{e}_{\hat{\tau}} &= ((1-2M/r)^{-1/2}, 0, 0, 0) \\
\mathbf{e}_{\hat{r}} &= (0, (1-2M/r)^{1/2}, 0, 0) \\
\mathbf{e}_{\hat{\theta}} &= (0, 0, 1/r, 0) \\
\mathbf{e}_{\hat{\phi}} &= (0, 0, 0, 1/rs)
\end{aligned}$$

gives the same transformed $R_{\hat{\alpha}\hat{\beta}\hat{\gamma}\hat{\delta}}$ that we found for freefall in Note 28.

14. The movie “Interstellar” features a planet where only an hour passes for every seven years aboard the mothership waiting in space, and where there are 4-hour tides apparently some thousand times greater than on Earth. Given that the well-known relativist Kip Thorne supposedly advised the film’s director, this must be at least plausible. What is likely true about the planet’s location and what is still implausible?
For such a gravitational time shift, the planet must be in close orbit around a tremendous mass such as a black hole.
The time distortion puts the planet’s orbit almost at the horizon, $r = r_H$, since

$$1 - \frac{r_H}{r} = \frac{1 \text{ hr}}{7\text{yr}} = \frac{1}{61320} \stackrel{\text{def}}{=} \epsilon$$

so

$$\frac{r_H}{r} = 1 - \epsilon \approx 1$$

A Schwarzschild black hole has an innermost stable circular orbit (ISCO: see Note 25) at $R = 6M$, too much bigger than $r_H = 2M$ to satisfy the above.
A Kerr black hole has an ISCO at $r = M$, if it is extremal (see Note 27), just skimming the horizon at $r_H = M$. So it is an extremal Kerr black hole.
Calculating Kerr tides is not straightforward so let’s pretend we can use Schwarzschild/Newton tides, especially $\partial_{\hat{\tau}\hat{\tau}}\chi^{\hat{r}} = 2M\chi^{\hat{r}}/r^3$ (Note 28). If the black hole is big enough, the tides at its horizon can be, paradoxically, as small as we like. (Yes, 1000*tides on Earth are actually quite small.) This is because if $r = M$ the tidal force $\propto M/r^3 = 1/M^2$.
We’ll compare the tides of “Interstellar” with the solar tides at Earth orbit so that we can get the black hole mass M in terms of solar mass, which we’ll call m. We’ll call Earth’s orbital distance r (it’s close enough to a circle), and for the “Interstellar” ISCO $R = M$. Thus, calling the 1000 factor simply f, we are comparing

$$\frac{M}{R^3} = f\frac{m}{r^3}$$

so

$$\begin{aligned}
M^3 &= \frac{Mr^3}{mf} \\
M &= \sqrt{\frac{r^3}{mf}}
\end{aligned}$$

For Earth $r = 150$Gm and for Sol $m = 1477$Tm, so if $f = 1000$

$$M = 48 \times 19^{12} = 1477 \times 32 \times 10^9$$

or 32 GSol.
Black holes of billions of solar masses apparently do exist. So the time shift and size of the tides seem plausible in “Interstellar”.

At least three things are still implausible to me. First, the time differential is between two observers, and the man left on board the mothership would have to be far enough away to be almost uninfluenced by the black hole gravity: not in orbit around the planet as the movie led me to understand (let alone what an orbit around a planet in Kerr ISCO would look like). Second, the crew landing on the planet stayed four hours (their time) and experienced two high tides, implying a planetary "day" of eight hours. But a planet that close to a black hole would be tidally locked, like our Moon, with one face always to its primary, and the tides would be stationary,

Third, the planet has liquid water—indeed it seems to be a water world. Two ways a planet could be warmed, apart from orbiting within the "Goldilocks" (habitable) zone of some star, would be radioactivity—such as what keeps the Earth's core molten—or tidal flexing—such as Jupiter's effect on its moon Europa (which may be what allows it to retain liquid water, albeit under some kilometers of outer ice crust).

15. Check the calculations for your imaginary visit to four black holes in [Tho94, Prologue].

16. The discussion in Note 28 is based on Hartle [Har03, pp.453,438,441], leaving steps of the argument out. Complete the argument.

17. Follow Hartle's argument [Har03, p.206] that the ratio $\mid \ell/e \mid = d$, the distance the incoming light ray initially is from the line parallel to it, passing through the centre of the Schwarzschild star in Note 29.

18. Note 29 was also taken from Hartle [Har03, pp.84,91,209,318]. Fill in the gaps. Work through the calculation from p.210 on the deflection of light predicted by Einstein.

19. Write a MATLAB function to calculate light geodesics for the Schwarzschild metric and use Note 29 to provide the starting parameters.

20. Hartle [Har03, §§22.1,22.2] and Cheng [Che10, §14.2.2] discuss the stress-energy tensor (Note 30), unfortunately for us assuming that it is accessible to their readers via the electromagnetic field. Misner, Thorne and Wheeler [MTW73, §5.7] discuss its symmetry.

21. Misner, Thorne and Wheeler [MTW73, Box 15.1] and, more accessibly, Wheeler [Whe90, Ch.7] discusses the connection between "the boundary of a boundary is zero" (see the Excursion of that name in Part I) and the Bianchi identities of Note 30.

22. Hartle [Har03, Ch.18] and Cheng [Che10, §9.4, Ch.10] work through many details of cosmology. D'Inverno [d'I92, p.335] offers a classification of Friedmann universes (Note 31).

23. **Newtonian limit.** Hartle [Har03, §22.4] and Cheng [Che10, §14.2.2] discuss the Newtonian limit of Einstein gravity. This is needed in Note 30 to establish the constant $8\pi G_N$ in the Einstein equation, and in Note 32 to show that the cosmological constant gives rise to a repulsive force. Here is a synopsis of Cheng's §11.1 and Box 14.1 for the latter (with $\kappa = 8\pi G_N$).
    In
$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \kappa T_{\mu\nu}$$
    we must shift the R part of the geometrical side (left side) to the source side (right side). So we take the trace of both sides
$$g^{\nu\mu}R_{\mu\nu} - \frac{1}{2}Rg^{\nu\mu}g_{\mu\nu} = \kappa g^{\nu\mu}T_{\mu\nu}$$
    so
$$R - 2R = \kappa T$$

and we can use the trace T instead of $-R$

$$R_{\mu\nu} = \kappa\left(T_{\mu\nu} - \frac{1}{2}Tg_{\nu\mu}\right)$$

In the Newtonian limit, T_{00} dominates and we can set $g_{\mathrm{dd}} = \eta_{\mathrm{dd}} = \mathrm{diag}[-1,1,1,1]$

$$\begin{aligned}
R_{00} &= \kappa\left(T_{00} - \frac{1}{2}T\eta_{00}\right) \\
&= \kappa\left(T_{00} - \frac{1}{2}\eta^{\nu\mu}T_{\mu\nu}\eta_{00}\right) \\
&= \kappa\left(T_{00} + \frac{1}{2}(-T_{00} + T_{11} + T_{22} + T_{33})\right) \\
&= \frac{\kappa}{2}(T_{00} + T_{11} + T_{22} + T_{33}) \\
&= \frac{\kappa}{2}(\rho + 3p)
\end{aligned}$$

From this we were able to use in Note 32 $\bigtriangledown^2\phi = R_{00} = 4\pi G_N(\rho + 3p)$
How does this also get the constant $8\pi G_N$ for the Einstein equation from the Newtonian constant $4\pi G_N$?

24. **Inflation.** That the universe is expanding (Note 32) suggested that at some point in time (at and prior to which physics cannot go) it was concentrated in (almost) zero space at some incredible temperature. (The strongest evidence for this was the 1965 discovery of Penzias and Wilson of a constant background radio signal, uniform and isotropic, created only some 360,000 years after that original point in time (and so at a much lower temperature but still higher than its apparent temperature today of 2.7°K) explained by Dicke et al. in 1965 and predicted in 1948 by George Gamow and by Alpher and Herman.)
In 1981 Alan Guth proposed that the universe could have expanded by some 10^{30} times in about 10^{-35} seconds right after the "big bang" as a result of the temperature cooling below a critical T_c [Gut97] [Che10, §11.2.2]
The mechanism is a Higgs field (similar to the one discovered in 2012 at the Large Hadron Collider at CERN) with a potential

$$\phi(r) = \alpha_0(T - T_c)r^2 + \lambda r^4$$

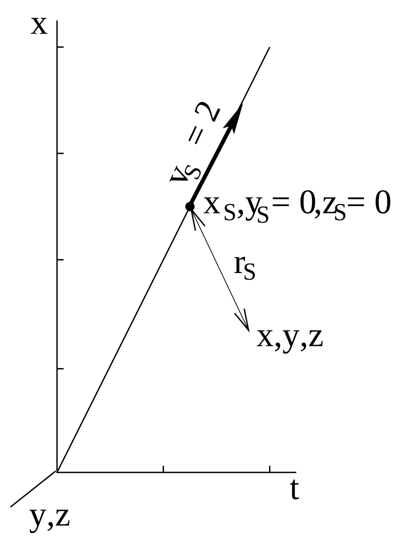
so that the quadratic part is negative once the temperature drops below T_c—and corresponds to the $\bigtriangledown^2\phi = -\Lambda$ of Note 32—but the quartic part is positive and so eventually stops the expansion.
Guth proposed that this enormous expansion would account for the apparent *flatness* of timespace observed by the astronomers and needed for the Friedmann-type cosmologies. He also pointed out that the rate of expansion, vastly greater than lightspeed (this is OK in *general* relativity, as long as it is *timespace* that is expanding; it is not OK for things moving *within* timespace, where special relativity always holds locally), explains why parts of the universe which we can observe are too far apart from each other ever to have influenced each other, and yet are remarkably uniform in properties, as if they had at some time been in equilibrium, (Or maybe light used to move much faster [Smo06, p.232].)

25. Guth [Gut97] gives a history of cosmology: early in chapter 3, inflationary notably in chapters 10, 12–14. In his Epilogue, Guth points out the inflation is a class of theories—in about fifty different forms as of publication in 1997. (And note that that was a year before the discovery of "dark energy".

26. In his Appendix A, Guth [Gut97] shows that a gravitational field has negative energy, using the example of a spherical shell collapsing under its own gravity to a smaller spherical shell and thus creating a gravitational field in the space between the two radii while releasing energy doing so. Smolin [Smo97, p.171,Ch.12] says that gravitating systems tend "to become more heterogeneous as time goes on" rather than succumbing to the increasing disorder of increasing energy; indeed, putting energy into an orbital system raises the orbit and slows the satellite, which is a *decrease* in temperature as contrasted with the opposite effect in a gas. Connect these ideas.

27. **Negative energy: warp drives.** The conventional way to use the Einstein equation was to construct the desired stress-energy tensor for the right-hand side and then solve the equation to find the metric.
That's what we did for cosmology in Note 31, although we also started with a shape for the metric (to capture the observed homogeneity and isotropy of the universe), leaving only one function to be determined by the Friedmann equations resulting from the Einstein equation. That's also what we did in Note 24 for Schwarzschild's spherically symmetric geometry and in Note 25 for Kerr's axially symmetric geometry. Even with $T_{\text{uu}} \neq 0$ in both of these cases, finding the metric was not easy.
In the 1970s Kip Thorne [Tho94] turned this process around, saying let's pretend we have advanced technology capable of doing whatever we want with energentum and stress. We'll choose the metric we want and find from that the stress-energy tensor we need.
Among other things, Thorne constructed wormholes [Har03, pp.151,273] [Che10, pp123,154] (and formulated the "grandfather paradox" of time travel tidily in terms of billiard balls).
In this excursion we look at Miguel Alcubierre's warp drive [Alc94] [Har03, pp.144,489].
Suppose we are in a spaceship S at position $(x_S, y_S, z_S) = (v_S t, 0, 0)$ and travelling at constant speed v_S along the x-axis, say $v_S = 2$. Then

$$(x_S, y_S, z_S) = (v_s t, 0, 0)$$



An arbitrary point (x, y, z) in space is a distance

$$r_S = \sqrt{(x - x_S)^2 + y^2 + z^2} = \sqrt{(x - x_S)^2 + \rho^2}$$

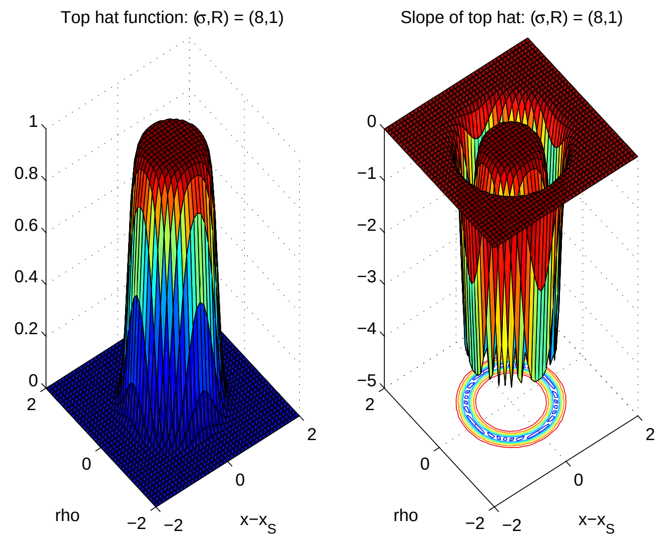
from the spaceship, where we define $\rho^2 = y^2 + z^2$ to reduce the two other spatial dimensions to one for visualization.
We want the metric to include a spherical region around the ship, the centre of which is

flat space but whose boundary shell is highly warped. This can be given by the "top hat" function,

$$f(r_S) = \frac{\tanh(\sigma(r_S + R)) - \tanh(\sigma(r_S - R))}{2\tanh(\sigma R)}$$

This has two parameters, σ which specifies (inversely) the width of the shell, and R which is its radius.
Here is the top hat function and its slope,

$$f'(r_S) = \frac{1}{2\tanh(\sigma R)}\left(\frac{1}{\cosh^2(\sigma(r_S + R))} - \frac{1}{\cosh^2(\sigma(r_S - R))}\right)$$



The values chosen for the parameters in these plots are $\sigma = 8$ and $R = 1$. Note that the slope, as the radius increases, drops to a large negative value before returning to zero.
In addition to the top hat function $f(r_S)$ the Alcubierre metric includes the desired velocity for the spaceship, $v_S(t) = \partial_t x_S(t)$. We'll assume v_S is constant and such that x_S is initially zero.
Here is the metric.

$$(\Delta s)^2 = -(\Delta t)^2 + (\Delta x - v(t)f(r_S)\Delta t)^2 + (\Delta\rho)^2$$

(this is in two spatial dimensions; in three $\Delta\rho$ would expand $(\Delta\rho)^2 = (\Delta x)^2 + (\Delta y)^2$), or

$$g_{\mathrm{dd}} = \begin{pmatrix} v^2f^2 - 1 & -vf & \\ -vf & 1 & \\ & & 1 \end{pmatrix} \qquad g_{\mathrm{uu}} = \begin{pmatrix} -1 & -vf & \\ -vf & 1 - v^2f^2 & \\ & & 1 \end{pmatrix}$$

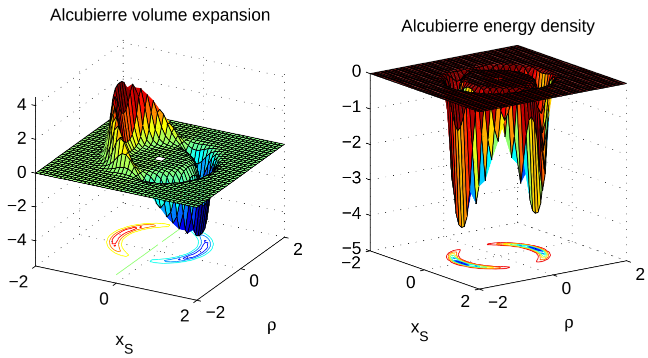
From this, two results follow. The volume expansion

$$\theta = v_S\partial_{x_S}f(r_S) = v_S\frac{x_S}{r_S}\partial_{r_S}f(r_S)$$

also known as York time, and the energy density

$$G^{00} = -\left(v_S\frac{\rho}{2r_S}\partial_{r_S}f(r_S)\right)^2$$

These both follow from the slope of the top hat function and we can plot them.



In the first figure note the negative expansion of the volume of space in front of the spacecraft as it travels in the x direction, and the positive expansion behind it.
In the second figure (and in the above expression for G^{00}: the energy density is, of course, $G^{00}/8\pi G_N$) note that the energy density is negative. This is something new. Dark energy, in its role in the expansion of the universe, involves negative *pressure* but positive energy density. To construct a warp drive (or a wormhole, for that matter) we need a further form of energy which we don't understand, so we call it "exotic matter".
I have not persuaded MATLAB to calculate G^{00} and suspect that more powerful calculation is needed. Hartle [Har03, p.489] refers to the Mathematica program for his book. Alcubierre cites advanced analytical techniques by James York.

28. **Negative mass.** Would a negative mass fall upward [ER12, Ch.11]?
Newton says
$$F = -G_N mM/r^2 \qquad\qquad F = ma$$
where m is the mass of the test body. In the first, m is the gravitational mass. In the second m is the inertial mass. They are different concepts but written with the same symbol m because they have always been found to be the same. Einstein raised this apparent coincidence to the *prinicple of equivalence*.
So which way does the force F go if m is negative?
What if the mass M of the attracting body were negative, with m still negative? With m positive?
If two test bodies of masses m and $-m$ ($m > 0$) were placed near each other and far from all other influences, what would they do? Are energy $(1/2)mv^2$ and momentum mv conserved?

29. **Blandford-Znajek mechanism.** Hartle [Har03, Box 15.1,p.326] describes in outline an electromagnetic mechanism by which a charged rotating black hole can supply the kind of energies observed, say, at the radio-astronomy source Cygnus A in the constellation Cygnus (Hartle p.289), or quasars (quasi-stellar objects) [Whe90, p.224]. Pursue this in connection with the discussion in Note 33.

30. What happens to the entropy of a black hole which evaporates by Hawking radiation (Note 33)? Hartle [Har03, §13.3] discusses Hawking radiation and black hole evaporation.

31. **Gravitational waves.** One prediction of general relativity which seems very hard to reconcile with space of unchanging curvature (such as Whitehead's space in Note 34a) is gravitational waves. These are too weak to have been detected yet on earth, but Russel Hulse and

Joseph Taylor discovered in 1974 a binary pulsar—a pair of neutron stars in orbit around each other, one of which is magnetized and sweeps us with a radio beam as it rotates [Har03, pp.250,534]—and observed orbital energy loss corresponding closely to the emission of gravitational waves [Che10, §15.4.3].
Study linearized general relativity and gravitational waves, e.g., in Hartle [Har03, Ch.16] or Cheng [Che10, Ch.15].

32. How can constant acceleration of an observer create a horizon in otherwise empty space, and associate a temperature with that horizon? (Note that you must be accelerating to hover close to a black hole horizon.) (Note 34.) See [Smo01, Ch.6].

33. For another take on geometry see [Tho94, Ch.11] which notes the equivalence of rubber rulers and flexible clocks to curved timespace, and also considers black-hole horizons as electrically conducting membranes.

34. > It is not true that to a spacetime there corresponds a geometry, defined by a Riemannian manifold. What is true is that to the physical spacetime there correspond [sic] an infinite class of manifolds and metrics, which are related one to the other by certain transformations, which are called diffeomorphisms. These transformations preserve only those relations among the fields that describe physically observable properties of space and time.
    >
    > The fact that in general relativity the physical world is represented not by geometries but only by those relationships inherent in geometries which are preserved by these transformations, is often missed in discussions of the theory. .. The argument that led to [Einstein's] understanding [of the difference] is called the "hole argument" and on a technical level it is the key to the interpretation of the theory, I refer the reader to the discussions of it by Sachel and Barbour, which have settled the matter for most physicists I know who have considered the matter.

    (in the Notes to Part IV [Smo97, p.333]): find out about diffeomorphisms.

35. Thorne [Tho94, Chs.1,2] reports thoroughly on the development of Einstein's thinking on special and general relativity. Seven of Einstein's original papers are translated in [LEMW52]. Thorne, in his next seven chapters, gives the turbulent history of the acceptance and development of the idea of black holes.

36. Compare Whitehead's treatment of tensors [Whi22, Part III] with section B of Einstein's 1916 paper [LEMW52, pp.120–42].

37. Any part of the Prefatory Notes that needs working through.

References

[Alc94] Miguel Alcubierre. The warp drive: hyper-fast travel within general relativity. *Classical and Quantum Gravity*, 11(5):L73–7, 1994.

[Ash86] Abhay Ashtekar. New variables for classical and quantum gravity. *Phys.Rev.Lett.*, 57(18):2244–7, Nov. 1986.

[Che10] Ta-Pei Cheng. *Relativity, Gravitation and Cosmology: A Basic Introduction.* Oxford Master Series in Particle Physics, Astrophysics, and Cosmology. Oxford University Press, Oxford, 2010.

[Chr70] Demetrios Christodoulou. Reversible and irreversible transformations in black-hole physics. *Phys. Rev. Lett.*, 25(22):1596–7, Nov. 1970.

[d'I92] Ray d'Inverno. *Introducing Einstein's Relativity*. Clarendon Press, Oxford, 1992.

[Edd23] A S Eddington. *The Mathematical Theory of Relativity*. Cambridge University Press, Cambridge, 1923.

[Edd30] A S Eddington. *The Nature of the Physical World*. Cambridge University Press, Cambridge, 1930. Gifford Lectures, University of Edinburgh, 1927.

[ER12] Allen Everett and Thomas Roman. *Time Travel and Warp Drives: A Scientific Guide to Shortcuts through Time and Space*. U. Chicago Press, Chicago, 2012.

[FMW03] Richard P. Feynman, Fernando B. Morinigo, and William G. Wagner. *Feynman Lectures on Gravitation*. Westview Press, Boulder and Oxford, 2003. Caltech Lectures, 1962–3.

[Gut97] Alan Guth. *The Inflationary Universe: The Quest for a New Theory of Cosmic Origins*. Addison-Wesley Publishing Co. Inc., Reading, Mass, 1997.

[Har03] James B Hartle. *Gravity: An Introduction to Einstein's General Relativity*. Addison Wesley, San Francisco, 2003.

[Jac95] Ted Jacobson. Thermodynamics of spacetime: The Einstein equation of state. *Physical Review Letters*, 75(7):1260–3, Aug. 1995. arXiv:gr-qc/9504004.

[LEMW52] H A Lorentz, A Einstein, H Minkowski, and H Weyl. *The Principle of Relativity*. Dover Publications, Inc., New York, 1952. Translated by W Perrett and G B Jeffery.

[LL45] Lillian R Lieber and Hugh Gray Lieber. *The Einstein Theory of Relativity*. Farrar & Rinehart, Inc., New York, 1936,1945.

[Mel09] Fulvio Melia. *Cracking the Einstein Code: Relativity and the Birth of Black Hole Physics*. U. Chicago Press, Chicago, 2009.

[MTW73] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman and Co., New York, 1973.

[Smo97] Lee Smolin. *The Life of the Cosmos*. Oxford University Press, New York, 1997.

[Smo01] Lee Smolin. *Three Roads to Quantum Gravity*. Basic Books, New York, 2001.

[Smo06] Lee Smolin. *The Trouble with Physics:The Rise of String Theory, the Fall of a Science, and What Comes Next*. TheHoughton MifflinCompany, New York, 2006.

[Smo13] Lee Smolin. *Time Reborn: From the Crisis in Physics to the Future of the Universe*. Alfred A Knopf, Canada, Toronto, 2013.

[Sta14] Andrei Starinets. Black holes in Einstein's gravity and beyond. Slides from talk given 2014/9/20. Last accessed 150509. https://www2.physics.ox.ac.uk/research/rudolf-peierls-centre-for-theoretical-physics/saturday-mornings-of-theoretical-physics, Sept. 2014.

[Syn51] J. L. Synge. Whitehead's principle of relativity. Mimeographed notes, compiled to LaTeXand posted to arXiv 05/5/19 by A.J.Coleman: arXiv:/0505027v/2, 1951.

[Tho94] Kip S. Thorne. *Black Holes and Time Warps: Einstein's Outrageous Legacy*. W.W.Norton & Company, New York, 1994.

[Ver10] Erik Verlinde. On the origin of gravity and the laws of Newton. arxiv.org/abs/1001.0785, Jan. 2010.

[Wei72] Steven Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity.* John Wiley and Sons, Inc., New York, Toronto, 1972.

[Whe90] J. A. Wheeler. *A Journey into Gravity and Spacetime.* W. H. Freeman and Company, New York, 1990.

[Whi19] A N Whitehead. *An Enquiry Concerning the Principles of Natural Knowledge.* Cambridge University Press, Cambridge, 1919.

[Whi20] A N Whitehead. *The Concept of Nature.* Cambridge University Press, Cambridge, 1920. Tanner Lectures, Trinity College, Nov. 1919.

[Whi22] A N Whitehead. *The Principle of Relativity with Applications to Physical Science.* Cambridge University Press, Cambridge, 1922.