A General Relativity Coursebook - Course Manual

Edward Daw

June 2023

Contents

1 Guide for Teachers

1.1 Lecture plan

The coursebook grew out of notes for a course for 4^{th} year undergraduates at a middle-tier UK University. It is taught in a single semester, in which there are a total of 11 teaching weeks. There are two classes per week, each 50 minutes long. In principle, then, there are 22 lectures, but in reality one of these are usually not available, because I set a diagnostic quiz. Therefore I budget for 20 lectures in all. Given 20 lectures, here is how I usually organise my class.

Number	Subject matter	Sections
1.	Intro, Realm of validity, Tangent curves and planes	1.1-1.4
2.	Diagnostic quiz	1.8
3.	Spacetime, principle of equivalence, Einstein sum notation	1.5-1.7
4.	Polar coordinates and simple tensors	2.1-2.3
5.	Generalisation of tensor expressions, first index algebra	2.4-2.5
6.	Tensors in spacetime, contraction, index raising and lowering	2.7-2.9.2
7.	Index swaps, Symmetrisation/antisymmetrisation, dust, energy–momentum–stress tensor	2.9.3-3.3
8.	Divergence of energy–momentum–stress tensor, Euler-Lagrange equations	3.4-4.2
9.	Lagrangians in space, space-time, cyclic coordinates, constants of motion	4.3-4.6
10.	Tensor differentiation and Christoffel symbols	4.7-4.9
11.	Methods for determining the Christoffel symbols, geodesic equations	4.10-4.11
12.	Parallel transport and the Riemann curvature	5.1-5.3
13.	Symmetries of the Riemann tensor, Zero and related components	5.4
14.	Ricci tensor and scalar, Bianchi identities, Einstein's equations	5.5-5.7
15.	Schwarzschild solution derivation	6.1-6.9
16.	Kruskal coordinates, wormholes	6.10-6.12
17.	Cosmology, FRW metric, up to Ricci scalar	7.1-7.8
18.	From Ricci scalar to Friedmann equations, Newtonian limit	7.9-7.10
19.	Cosmological constant, Gravitational wave solutions for metric	7.14-8.5
20.	Gravitational waves in the lab frame and interferometry	8.6-8.8

Table 1: Lectures for a 4^{th} year undergraduate course in general relativity.

1.2 Problem chains

Though most of the problems in the book are independent of each other, certain problems are grouped together in sequence, so that the student will need ideally to have done all the previous problems in the sequence to attempt the next one. These sequences are identified below, along with an indication of the author's view of the overall difficulty of the sequence.

Sequence 1: Spherical polar coordinates. Problems 2.2, 2.3, 4.8, 4.9, 4.10, 4.11, 5.1, 5.3 Mid-grade.
Sequence 2: Antisymmetric tensors, curl. Problems 2.2, 2.8, 2.9, 2.10, 4.12, 4.13. Challenging.
Sequence 3: Precession of the perihelion of Mercury. 6.6, 6.7. Challenging.

2 Chapter 1

2.1 Diagnostic Quiz

Parts (a) to (o) are intended to test ability to apply the chain and product rules, and from parts (k) onwards to mix ordinary and partial derivatives.

a) $\frac{d}{dt}\left(e^{\theta}\right) = \frac{d\theta}{dt}e^{\theta}$.

b) $\frac{d}{dt}\left(e^{-r^2}\right) = -2r\left(\frac{dr}{dt}\right)e^{-r^2}$.

c) $\frac{d}{d\tau}\left(\cos(r\theta)\right) = -\sin(r\theta)\left(r\frac{d\theta}{d\tau} + \theta\frac{dr}{d\tau}\right)$.

d) $\frac{d}{dt}\left(\cosh\left(\frac{r}{\theta}\right)\right) = \sinh\left(\frac{r}{\theta}\right)\left(r\frac{d}{dt}\left(\theta^{-1}\right) + \theta^{-1}\frac{dr}{dt}\right) = \sinh\left(\frac{r}{\theta}\right)\left(\frac{1}{\theta}\frac{dr}{dt} - \frac{r}{\theta^2}\frac{d\theta}{dt}\right)$.

e) $\frac{d}{dt}\left(r\sinh\left(e^{\theta^2}\right)\right) = \frac{dr}{dt}\sinh\left(e^{\theta^2}\right) + r\frac{d}{dt}\left(\sinh\left(e^{\theta^2}\right)\right) = \frac{dr}{dt}\sinh\left(e^{\theta^2}\right) + 2r\theta e^{\theta^2}\frac{d\theta}{dt}\cosh\left(e^{\theta^2}\right)$.

f) $\frac{d}{dt}\left(r\frac{d\theta}{dt}\right) = \frac{dr}{dt}\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2}$.

g) $\frac{d}{dt}\left(r^2\frac{dr}{dt}\frac{d\theta}{dt}\right) = 2r\left(\frac{dr}{dt}\right)^2\frac{d\theta}{dt} + r^2\frac{d^2r}{dt^2}\frac{d\theta}{dt} + r^2\frac{dr}{dt}\frac{d^2\theta}{dt^2}$.

i) $\frac{d}{dt}\left(\frac{d\theta}{dt}/\frac{dr}{dt}\right) = \frac{d^2\theta}{dt^2}/\frac{dr}{dt} + \frac{d\theta}{dt}\frac{d}{dt}\left(1/\frac{dr}{dt}\right) = \frac{d^2\theta}{dt^2}/\frac{dr}{dt} - \frac{d\theta}{dt}\frac{d^2r}{dt^2}/\left(\frac{dr}{dt}\right)^2$.

j) $\frac{d}{dt}\left(\exp\left(\frac{d\theta}{dt}/\frac{dr}{dt}\right)\right) = \exp\left(\frac{d\theta}{dt}/\frac{dr}{dt}\right) \times$ answer to (i).

k) $\frac{\partial}{\partial r}\left(r\theta\right) = \theta$.

l) $\frac{d}{dt}\frac{\partial}{\partial r}\left(\frac{r}{\theta}\right) = -\frac{1}{\theta^2}\frac{d\theta}{dt}$.

m) $\frac{\partial}{\partial r}\frac{d}{dt}\left(\frac{r}{\theta}\right) = \frac{\partial}{\partial r}\left(\frac{1}{\theta}\frac{dr}{dt} - \frac{r}{\theta^2}\frac{d\theta}{dt}\right) = -\frac{1}{\theta^2}\frac{d\theta}{dt}$.

n) $\frac{d}{dt}\frac{\partial}{\partial r}\left(\frac{r}{\theta^3}\right) = \frac{-3}{\theta^4}\frac{d\theta}{dt}$.

o) $\frac{d}{dt}\frac{\partial}{\partial r}\frac{\partial}{\partial\theta}\left(\theta^3 r^{-3}\right) = \frac{d}{dt}\frac{\partial}{\partial r}\left(3\theta^2 r^{-3}\right) = \frac{d}{dt}\left(-9\theta^2 r^{-4}\right) = -18\frac{\theta}{r^4}\frac{d\theta}{dt} + 36\frac{\theta^2}{r^5}\frac{dr}{dt}$.

Parts (p) to (s) are intended to test the ability to invert matrices.

(p) Either by the standard technique for inversion of a 2×2 matrix, or by recognising this as a rotation matrix, $R(\theta)$, and using $R^{-1}(\theta) = R(-\theta)$,

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

(q) Diagonal matrices are inverted by inverting each diagonal element, so

$$\begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix}^{-1} = \begin{pmatrix} 1/A & 0 & 0 \\ 0 & 1/B & 0 \\ 0 & 0 & 1/C \end{pmatrix}.$$

(r) This matrix is block diagonal, so take the reciprocal of A and use the standard inversion technique for the 2×2 sub-matrix excluding the elements in the row and column of the A element.

$$\left(\begin{array}{ccc} A & 0 & 0 \\ 0 & P & Q \\ 0 & R & S \end{array}\right)^{-1} = \left(\begin{array}{ccc} 1/A & 0 & 0 \\ 0 & \frac{S}{PS-RQ} & \frac{-Q}{PS-RQ} \\ 0 & \frac{-R}{PS-RQ} & \frac{P}{PS-RQ} \end{array}\right).$$

(s) This matrix is block diagonal. Either use the standard inversion technique for the 2×2 sub-matrix excluding the second and third row and column, or recognise this as a Lorentz boost matrix parallel to the z-axis, with relative velocity $v = \beta c$ directed parallel to the common z-axis between the two frames, and $\gamma = 1/\sqrt{1-\beta^2}$, so that the inverse is a matrix with the same form, the same γ and β replaced by $-\beta$. Consequently

$$\left(\begin{array}{cccc} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{array}\right)^{-1} = \left(\begin{array}{cccc} \gamma & 0 & 0 & +\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ +\beta\gamma & 0 & 0 & \gamma \end{array}\right).$$

Parts (t) to (ac) test the ability to solve ordinary differential equations. In the following, C and D are constants of integration.

(t) $y = -A^2 x + C$.

(u) $y = -\frac{A^2 x^2}{2} + C$.

(v) Separate and integrate, $dy/y = -A^2\,dx$, so $\ln|y| = -A^2 x + C$, so $y = De^{-A^2 x}$.

(w) Students should probably recognise the classical simple harmonic oscillator and just write the answer by inspection. They could also use a standard trial solution $y = e^{st}$, and solve for s; this leads naturally to complex solutions, and real linear combinations of these solutions lead you back to the same real result, $y = C\cos(Ax) + D\sin(Ax)$.

(x) Similar to (w), but this time you end up with real exponentials. $y = Ce^{+Ax} + De^{-Ax}$.

(y) Substitute $x = B\tan\theta$, leading to $y = (1/B)\arctan(x/B) + C$.

(z) $y = B\ln|B + x| + C$.

(aa) Write $x/(B + x)$ as $(B + x)/(B + x) - B/(B + x) = 1 - B/(B + x)$ and proceed as for (z), so $y = x - B\ln|B + x| + C$.

(ab) Substitute $x = B\sin\theta$, leading to $y = \arcsin(x/B) + C$.

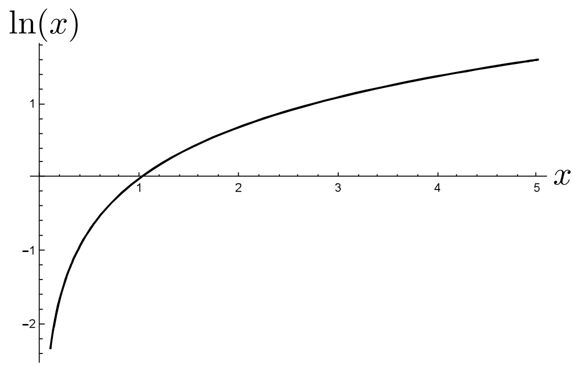
(ac) Substitute $z = y + 1$, then $d^2z/dx^2 = d^2y/dx^2$, substituting these results in, equation becomes $d^2z/dx^2 = z$. This is a special case of part (x), leading to $z = Ce^x + De^{-x}$ or $y = Ce^x + De^{-x} - 1$.

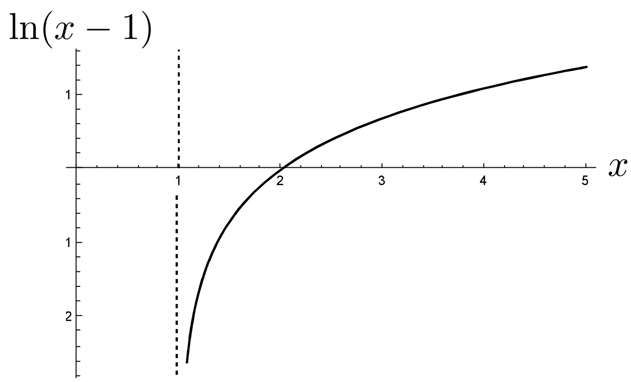
Parts (ad) to (ag) test the ability to sketch curves.

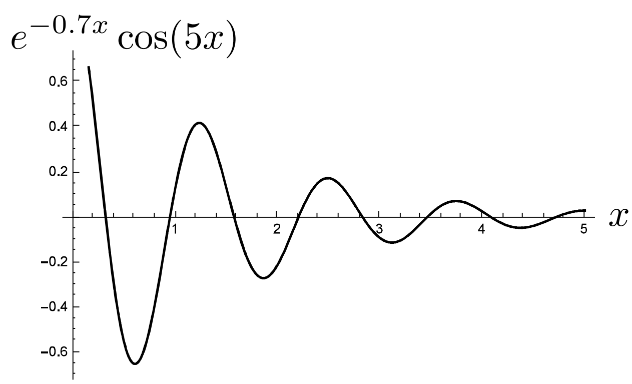
(ad) Plot of $\ln(x)$ versus x.
(ae) Plot of $\ln(x - 1)$ versus x.

(af) Plot of $e^{-0.7x}\cos(5x)$ versus x.

(ag) Plot of $(x + 1)/(x - 2)$ versus x.

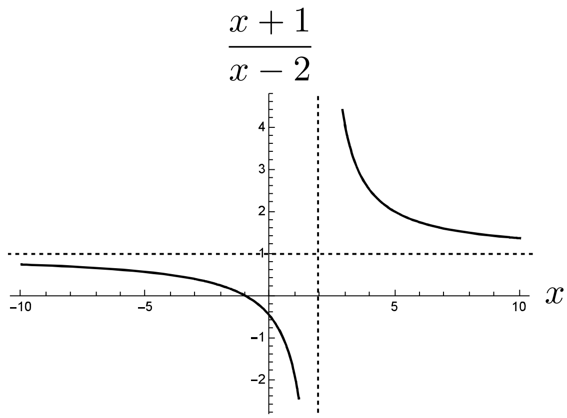






2.2 Problems from Chapter 1

1.1 Starting from the definition of the derivative in Equation 1.1, show that the first derivative of x^n, where n is a positive integer, is nx^{n-1}.



Substitute x^n into the Equation.

$$\begin{aligned}\frac{dy}{dx} &= \lim_{\varepsilon\to 0}\frac{(x+\varepsilon)^n - x^n}{\varepsilon}\\ &= \lim_{\varepsilon\to 0}\frac{x^n\left(1+\frac{\varepsilon}{x}\right)^n - x^n}{\varepsilon}.\end{aligned}$$

Expand $(1+\varepsilon/x)$ binomially, to first order. The leading x^n terms cancel, and in the remaining term the factors of ε cancel, so you can take the limit, leading to the desired result.

$$\begin{aligned}\frac{dy}{dx} &\simeq \lim_{\varepsilon\to 0}\frac{x^n\left(1+\frac{n\varepsilon}{x}\right) - x^n}{\varepsilon}\\ &= nx^{n-1}.\end{aligned}$$

1.2 A triangle drawn on the surface of a sphere of radius R consists of three segments of great circles that join at three vertices. It is a right angled triangle if one of the three interior angles of the figure is 90°. It is a small spherical triangle if the lengths of the three segments are all much less than the radius R. If the three sides are of length a, b, and c, and if there is a right angle between the sides of length b and c, then the cosine rule of spherical trigonometry simplifies to $\cos a = (\cos b)\times(\cos c)$.

(a) Show that to second order in a, b and c or products of these, we have $a^2 \simeq b^2 + c^2$.

(b) Show that to fourth order the explicit form of the modified Pythagoras' theorem is

$$a^2 \simeq b^2 + c^2 - \frac{1}{3}b^2c^2.$$

Hint: when you encounter a term in a^4, you will need to re-express this in terms of terms in b and c. You can substitute in the second order result from part (a) so long as you can argue that any other terms beyond this approximation result in corrections at higher order than b^4, c^4 or b^2c^2, and can therefore be neglected at fourth order.

In my experience students find part (b) of this problem challenging.
(1.2a) Substitute in the expansions of the cosines to second order in their arguments, multiply out

and collect terms, neglecting the fourth order term that appears on the right.

$$
\begin{aligned}
\cos(a) &= \cos(b)\cos(c) \\
\left(1-\frac{a^2}{2}\right) &\simeq \left(1-\frac{b^2}{2}\right)\left(1-\frac{c^2}{2}\right) \\
&\simeq 1-\frac{b^2}{2}-\frac{c^2}{2} \\
a^2 &\simeq b^2+c^2
\end{aligned}
$$

(1.2b) This time keep terms up to 4^{th} order in a, b, and c in the cosine expansions. Again, multiply out the brackets on the right and neglect any sixth order terms.

$$
\begin{aligned}
1-\frac{a^2}{2}+\frac{a^4}{24} &\simeq \left(1-\frac{b^2}{2}+\frac{b^4}{24}\right)\left(1-\frac{c^2}{2}+\frac{c^4}{24}\right) \\
&\simeq 1-\frac{b^2}{2}+\frac{b^4}{24}-\frac{c^2}{2}+\frac{c^4}{24}+\frac{b^2c^2}{4} \\
a^2 &\simeq b^2+c^2-\frac{b^4}{12}-\frac{c^4}{12}-\frac{b^2c^2}{2}+\frac{a^4}{12}
\end{aligned}
$$

The last step is to realise that the rightmost term can be expanded using the result from part (a), because any 4^{th} order corrections to that result will appear as 6^{th} order or higher and are therefore neglected to 4th order. This yields an equation with all the 4th order terms that appear, and no higher order terms remaining. To fourth order

$$
\begin{aligned}
a^2 &\simeq b^2+c^2-\frac{b^4}{12}-\frac{c^4}{12}-\frac{b^2c^2}{2}+\frac{(b^2+c^2)^2}{12} \\
&\simeq b^2+c^2-\frac{b^4}{12}-\frac{c^4}{12}-\frac{b^2c^2}{2}+\frac{b^4+c^4+2b^2c^2}{12} \\
a^2 &\simeq b^2+c^2-\frac{b^2c^2}{2}+\frac{b^2c^2}{6} \\
&\simeq b^2+c^2-\frac{b^2c^2}{3}.
\end{aligned}
$$

Students may be interested in how this result is consistent with Equation 1.26. The answer is that we can write

$$
\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \simeq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -bc/6 \\ -bc/6 & 0 \end{pmatrix}.
$$

This is consistent with with Equation 1.26, and reproduces the relationship we have proved.

1.3 Physics problems on spheres are commonly done using spherical polar coordinates at fixed radius, (θ, ϕ), where θ is the colatitude, the angle between the line joining the sphere centre and the north pole and the line joining the sphere centre and the point in question, and ϕ is the azimuthal angle between an arbitrary great circle running through the north and south poles (in the case of the Earth this arbitrary great circle also passes through Greenwich) on the sphere, and the plane running through the north and south poles and the point in question. A three sided figure on the surface of a sphere is constructed by starting at the North pole, then descending at constant ϕ to some value of θ, next moving at fixed θ through an interval of ϕ, so remaining at constant latitude and moving east or west, then finally travelling back at fixed ϕ to the north pole. Is this three sided figure a spherical triangle? If not, why not?

No, the three sided figure described isn't a spherical triangle. This is because a line of constant latitude is not a geodesic. To see this, imagine a line of constant latitude very close to the North

Pole. At high latitudes, it's obviously a circle, and a circle is not a straight line. More generally, the only line of constant latitude which is a geodesic is the equator. If you look at the flight paths of planes, which tend to follow geodesics if they can because they minimise the use of fuel, they hardly ever follow lines of latitude. Alternatively, try stretching elastic over the surface of a globe, and you'll find that the geodesic paths followed by stretched elastic veer north (south) of the latitude lines in the northern (southern) hemisphere.

1.4 Generalising from Equations 1.15 in two dimensions and 1.21 in four dimensions, how many independent second order derivatives are there of the metric coefficients in D dimensions?

In N dimensions, the number of independent metric coefficients is $N(N+1)/2$. This is easiest determined by trying 2, 3 and 4 dimensions. The general formula can be proved by induction, as follows: Assume it's right in N dimensions. Adding another dimension contributes $N+1$ independent new numbers, a whole single column in the matrix which is now $N+1 \times N+1$. The total number of independent coefficients in $N+1$ dimensions is $N(N+1)/2 + (N+1)$, which simplifies to $(N(N+1) + 2(N+1))/2$, or $(N+1)((N+1)+1)/2$. This is of the same form as $N(N+1)/2$, with N replaced by $N+1$. Therefore, assuming the result for N dimensions, and adding another dimension, we obtain a result of the same form for $N+1$ dimensions. Finally, the result holds for 1 dimension, yielding only 1 independent number. This concludes the proof.

1.5 A lift is falling freely under gravity. An observer in the lift has set up a laser pointer attached to the wall of the lift pointing in her coordinate systems at right angles so that the emerging beam is emitted normal to the wall. Bearing in mind the principle of equivalence, what is the pathway of the photons across the lift. What would an observer outside the lift standing on the Earth's surface observe about the light beam as the lift falls? It's a glass lift so both observers can see the beam.

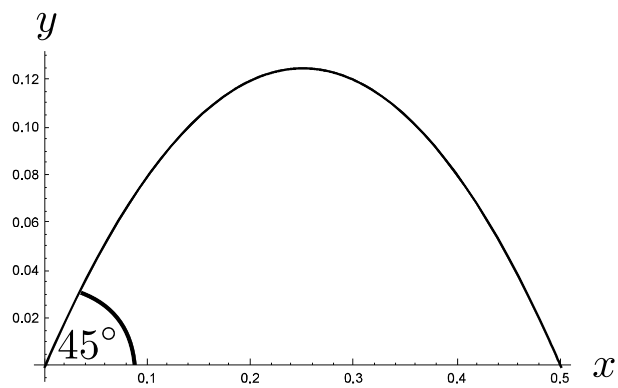
By the principle of equivalence, within the lift, the laser light will follow a straight line trajectory between one side of the lift and the other, in a Cartesian coordinate system stationary with respect to the lift. To the observer outside the lift, the lift is accelerating downwards. Because the speed of light is finite, the pathway taken by the light in a 'flat Earth' approximation is a parabola. This is not the right result, because the Earth isn't really flat, but it does correctly predict that the light beam bends towards the Earth.

1.6 Specialist aircraft such as the KC-105 provided early trainee astronauts with a zero gravity environment. Incidentally, the KC-105 was nicknamed 'the vomit comet'! Apparently the first experience of a zero gravity environment can be uncomfortable. A KC-105 starts its zero gravity flight segment at a height h flying at an initial velocity $\vec{v}$ directed $45°$ above the horizontal. The passengers remain in zero gravity until the plane again reaches the same altitude h. Take the x and y axes to be parallel to and perpendicular to the Earth's surface, respectively.

(a) **Sketch the trajectory of the aircraft during its zero gravity segment.**

(b) **Assuming the maximum vertical displacement of the plane is much less than the radius of the Earth, give formulae for the horizontal and vertical components of the aircraft displacement during the zero gravity flight segment. Take the start of the zero gravity segment to be at time $t = 0$ and position $x = 0$, $y = h$, the subsequent motion to be in the XY plane, and the horizontal component of the initial velocity to be in the direction of increasing x. Express your answers in terms of the quantities defined above and the local gravitational acceleration magnitude g directed downwards.**

(1.6a) Just as for the elevator, the aircraft must follow the trajectory of a freely falling object in order that objects inside to have the same acceleration vector as the aircraft walls, and hence float inside the aircraft without accelerating towards the walls. If the initial velocity of the aircraft is directed as described, then the subsequent trajectory is a parabola, as discussed in Section 1.1.



(1.6b) This is an application of the standard constant acceleration equations. With a 45° degree initial angle of elevation, the horizontal and vertical components of the initial velocity are the same. Let $v = |\vec{v}|$. Both horizontal and vertical velocity components are $v/\sqrt{2}$.

$$\begin{aligned} x &= \frac{vt}{\sqrt{2}} \\ y &= h + \frac{vt}{\sqrt{2}} - \frac{gt^2}{2} \end{aligned}$$

Substitute t from the equation for x into the equation for y.

$$\begin{aligned} y &= h + x - \frac{gx^2}{v^2} \\ &= h + x\left(1 - \frac{gx}{v^2}\right) \end{aligned}$$

The height y is h when $x = 0$ and when $x = v^2/g$.

3 Chapter 2

2.1 A portion of the two dimensional plane can be described in plane hyperbolic coordinates (ε, η), where

$$x = \varepsilon \cosh \eta$$
$$y = \varepsilon \sinh \eta,$$

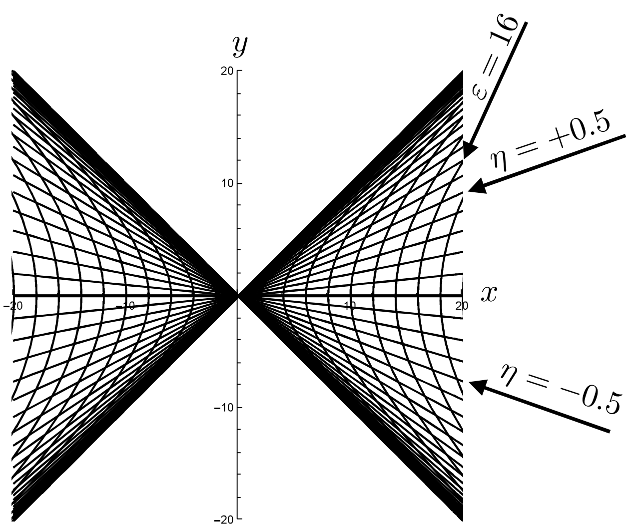
and η and ε are real numbers.

(a) Sketch and describe the lines of constant ε.

Squaring the expressions for x and y, and taking the difference, we obtain $x^2 - y^2 = \varepsilon^2$, so that $y = \pm\sqrt{x^2 - \varepsilon^2}$. Lines of constant ε are therefore hyperbolas centered on the x axis. The limiting curve is where $\varepsilon = 0$, in which case you have $y = \pm x$, which corresponds to straight lines of gradient ± 1 passing through the origin. For real ε, all points on all hyperbolas obey $|y| \leq |x|$.

(b) Sketch and describe the lines of constant η.

Taking the ratio $y/x = \tanh \eta$, so that $y = x \tanh \eta$. Lines of constant η are therefore straight lines passing through the origin having gradient η. Since $\tanh \eta$ has range $(-1, 1)$, the lines again populate the region defined by $|y| \leq |x|$.



(c) What portion of the XY plane is not covered in plane hyperbolic coordinates?

See above discussion and figure. The region where $|y| > |x|$ is not covered.

(d) Work out expressions for the basis vectors $\vec{e_\eta}$ and $\vec{e_\varepsilon}$ in terms of the coordinates ε, η, $\vec{e_x}$ and $\vec{e_y}$.

Start by writing $\vec{r} = x\vec{e_x} + y\vec{e_y} = \varepsilon\cosh(\eta)\vec{e_x} + \varepsilon\sinh(\eta)\vec{e_y}$. Therefore we obtain

$$\begin{aligned}
\vec{e_\varepsilon} &= \frac{d\vec{r}}{d\varepsilon} = \cosh(\eta)\vec{e_x} + \sinh(\eta)\vec{e_y} \\
\vec{e_\eta} &= \frac{d\vec{r}}{d\eta} = \varepsilon\sinh(\eta)\vec{e_x} + \varepsilon\cosh(\eta)\vec{e_y}.
\end{aligned}$$

(e) Work out the components of the metric in plane hyperbolic coordinates in terms of η and ε.

The metric coefficients are the dot products between all possible pairs of basis vectors $\vec{e_\varepsilon}$ and $\vec{e_\eta}$, so

$$\begin{aligned}
(g_{\mu\nu}) = \begin{pmatrix} \vec{e_\varepsilon}\cdot\vec{e_\varepsilon} & \vec{e_\varepsilon}\cdot\vec{e_\eta} \\ \vec{e_\eta}\cdot\vec{e_\varepsilon} & \vec{e_\varepsilon}\cdot\vec{e_\varepsilon} \end{pmatrix} &= \begin{pmatrix} \cosh^2(\eta)+\sinh^2(\eta) & 2\varepsilon\sinh(\eta)\cosh(\eta) \\ 2\varepsilon\sinh(\eta)\cosh(\eta) & \varepsilon^2(\cosh^2(\eta)+\sinh^2(\eta)) \end{pmatrix} \\
&= \begin{pmatrix} \cosh(2\eta) & \varepsilon\sinh(2\eta) \\ \varepsilon\sinh(2\eta) & \varepsilon^2\cosh(2\eta) \end{pmatrix}
\end{aligned}$$

(f) Work out an expression for the transformation between the plane hyperbolic components $(d\varepsilon, d\eta)$ of a small displacement $d\vec{r}$ in the portion of the plane covered by the plane hyperbolic coordinate system to the same small displacement in terms of the Cartesian components of the displacement (dx, dy).

Using the total derivative formula (Equation 2.18 for example), we obtain

$$\begin{aligned}
d\varepsilon &= \left(\frac{\partial\varepsilon}{\partial x}\right)_y dx + \left(\frac{\partial\varepsilon}{\partial y}\right)_x dy \\
d\eta &= \left(\frac{\partial\eta}{\partial x}\right)_y dx + \left(\frac{\partial\eta}{\partial y}\right)_x dy
\end{aligned}$$

Start with $x^2 - y^2 = \varepsilon^2$ from part (a). Then differentiate with respect to x and y, yielding

$$\begin{aligned}
x &= \varepsilon\left(\frac{\partial\varepsilon}{\partial x}\right)_y \\
\left(\frac{\partial\varepsilon}{\partial x}\right)_y &= \frac{x}{\varepsilon} = \cosh\eta \\
-y &= \varepsilon\left(\frac{\partial\varepsilon}{\partial y}\right)_x \\
\left(\frac{\partial\varepsilon}{\partial y}\right)_x &= \frac{-y}{\varepsilon} = -\sinh(\eta).
\end{aligned}$$

Now start with $\tanh(\eta) = y/x$, again from part (a). Then differentiate with respect to x and y, yielding

$$
\begin{aligned}
\operatorname{sech}^2(\eta)\left(\frac{\partial\eta}{\partial x}\right)_y &= \frac{-y}{x^2} \\
\frac{1}{\cosh^2(\eta)}\left(\frac{\partial\eta}{\partial x}\right)_y &= \frac{-y}{\varepsilon^2\cosh^2(\eta)} \\
\left(\frac{\partial\eta}{\partial x}\right)_y &= \frac{-\sinh(\eta)}{\varepsilon}. \\
\operatorname{sech}^2(\eta)\left(\frac{\partial\eta}{\partial y}\right)_x &= \frac{1}{x} \\
\frac{1}{\cosh^2(\eta)}\left(\frac{\partial\eta}{\partial y}\right)_x &= \frac{1}{\varepsilon\cosh(\eta)} \\
\left(\frac{\partial\eta}{\partial y}\right)_x &= \frac{\cosh(\eta)}{\varepsilon}.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
d\varepsilon &= \cosh(\eta)dx - \sinh(\eta)\vec{e_y} \\
d\eta &= \frac{-\sinh(\eta)}{\varepsilon}dx + \frac{\cosh(\eta)}{\varepsilon}\vec{e_y}.
\end{aligned}
$$

(g) Show that $d\vec{r} = dx\,\vec{e_x} + dy\,\vec{e_y} = d\eta\,\vec{e_\eta} + d\varepsilon\,\vec{e_\varepsilon}$. Use explicit substitution as was carried out in Section 2.3 for the case of plane polar coordinates, rather than by asserting the general results.

Write $d\vec{r} = d\varepsilon\vec{e_\varepsilon} + d\eta\vec{e_\eta}$, and substitute in from the above and from part (d).

$$
\begin{aligned}
d\vec{r} = &\left(\left(\frac{\partial\varepsilon}{\partial x}\right)_y dx + \left(\frac{\partial\varepsilon}{\partial y}\right)_x dy\right)(\cosh(\eta)\vec{e_x} + \sinh(\eta)\vec{e_y}) + \\
&\left(\left(\frac{\partial\eta}{\partial x}\right)_y dx + \left(\frac{\partial\eta}{\partial y}\right)_x dy\right)(\varepsilon\sinh(\eta)\vec{e_x} + \varepsilon\cosh(\eta)\vec{e_y}) \\
&= (\cosh(\eta)dx - \sinh(\eta)dy)(\cosh(\eta)\vec{e_x} + \sinh(\eta)\vec{e_y}) + \\
&\left(\frac{-\sinh(\eta)}{\varepsilon}dx + \frac{\cosh(\eta)}{\varepsilon}dy\right)(\varepsilon\sinh(\eta)\vec{e_x} + \varepsilon\cosh(\eta)\vec{e_y}) \\
&= \cosh^2(\eta)dx\vec{e_x} - \sinh(\eta)\cosh(\eta)dx\vec{e_y} - \sinh(\eta)\cosh(\eta)dy\vec{e_x} - \sinh^2(\eta)dy\vec{e_y} \\
&- \sinh^2(\eta)dx\vec{e_y} - \sinh(\eta)\cosh(\eta)dx\vec{e_y} + \cosh(\eta)\sinh(\eta)dy\vec{e_x} + \cosh^2(\eta)dy\vec{e_y} \\
&= dx\vec{e_x} + dy\vec{e_y}.
\end{aligned}
$$

Alternatively some approach this from the other end. Start with

$$
\begin{aligned}
dx &= \left(\frac{\partial x}{\partial\varepsilon}\right)_\eta d\varepsilon + \left(\frac{\partial x}{\partial\eta}\right)_\varepsilon d\eta = \cosh(\eta)d\varepsilon + \varepsilon\sinh(\eta)d\eta \\
dy &= \left(\frac{\partial y}{\partial\varepsilon}\right)_\eta d\varepsilon + \left(\frac{\partial y}{\partial\eta}\right)_\varepsilon d\eta = \sinh(\eta)d\varepsilon + \varepsilon\cosh(\eta)d\eta.
\end{aligned}
$$

We need expresions for $\vec{e_x}$ and $\vec{e_y}$ in terms of $\vec{e_\varepsilon}$ and $\vec{e_\eta}$. From part (d) we have

$$
\begin{aligned}
\vec{e_\varepsilon} &= \cosh(\eta)\vec{e_x} + \sinh(\eta)\vec{e_y} \\
\vec{e_\eta} &= \varepsilon\sinh(\eta)\vec{e_x} + \varepsilon\cosh(\eta)\vec{e_y}.
\end{aligned}
$$

Multiply the upper equation by $\varepsilon\cosh(\eta)$ and the lower equation by $\sinh(\eta)$.

$$\begin{aligned}\varepsilon\cosh(\eta)\vec{e_\varepsilon} &= \varepsilon\cosh^2(\eta)\vec{e_x} + \varepsilon\sinh(\eta)\cosh(\eta)\vec{e_y}\\ \sinh(\eta)\vec{e_\eta} &= \varepsilon\sinh^2(\eta)\vec{e_x} + \varepsilon\sinh(\eta)\cosh(\eta)\vec{e_y}.\end{aligned}$$

Subtract the lower equation from the upper one.

$$\begin{aligned}\varepsilon\cosh(\eta)\vec{e_\varepsilon} - \sinh(\eta)\vec{e_\eta} &= \varepsilon\vec{e_x}\\ \vec{e_x} &= \cosh(\eta)\vec{e_\varepsilon} - \frac{\sinh(\eta)}{\varepsilon}\vec{e_\eta}.\end{aligned}$$

This time multiply the upper equation by $\varepsilon\sinh(\eta)$ and the lower equation by $\cosh(\eta)$.

$$\begin{aligned}\varepsilon\sinh(\eta)\vec{e_\varepsilon} &= \varepsilon\sinh(\eta)\cosh(\eta)\vec{e_x} + \varepsilon\sinh^2(\eta)\vec{e_y}\\ \cosh(\eta)\vec{e_\eta} &= \varepsilon\sinh(\eta)\cosh(\eta)\vec{e_x} + \varepsilon\cosh^2(\eta)\vec{e_y}.\end{aligned}$$

Subtract the upper equation from the lower one.

$$\begin{aligned}\cosh(\eta)\vec{e_\eta} - \varepsilon\sinh(\eta)\vec{e_\varepsilon} &= \varepsilon\vec{e_y}\\ \vec{e_y} &= -\sinh(\eta)\vec{e_\varepsilon} + \frac{\cosh(\eta)}{\varepsilon}\vec{e_\eta}.\end{aligned}$$

Therefore

$$\begin{aligned}d\vec{r} &= dx\vec{e_x} + dy\vec{e_y}\\ &= (\cosh(\eta)d\varepsilon + \varepsilon\sinh(\eta)d\eta)\left(\cosh(\eta)\vec{e_\varepsilon} - \frac{\sinh(\eta)}{\varepsilon}\vec{e_\eta}\right)\\ &+ (\sinh(\eta)d\varepsilon + \varepsilon\cosh(\eta)d\eta)\left(-\sinh(\eta)\vec{e_\varepsilon} + \frac{\cosh(\eta)}{\varepsilon}\vec{e_\eta}\right)\\ &= \cosh^2(\eta)d\varepsilon\vec{e_\varepsilon} - \frac{\sinh(\eta)\cosh(\eta)}{\varepsilon}d\varepsilon\vec{e_\eta} + \sinh(\eta)\cosh(\eta)d\eta\vec{e_\varepsilon} - \sinh^2(\eta)d\eta\vec{e_\eta}\\ &- \sinh^2(\eta)d\varepsilon\vec{e_\varepsilon} + \frac{\sinh(\eta)\cosh(\eta)}{\varepsilon}d\varepsilon\vec{e_\eta} - \varepsilon\sinh(\eta)\cosh(\eta)\vec{e_\varepsilon} + \cosh^2(\eta)d\eta\vec{e_\eta}.\\ &= d\varepsilon\vec{e_\varepsilon} + d\eta\vec{e_\eta}.\end{aligned}$$

2.2 The conversion from Cartesian to spherical polar coordinates takes place by way of the following substitutions: $x = r\sin\theta\cos\phi$, $y = r\sin\theta\cos\phi$, and $z = r\cos\theta$, where θ is the so-called colatitude, (the latitude θ^{L} is by convention the angle of elevation above the equator, $\theta^{\mathrm{L}} = \pi - \theta$), and ϕ is the longitude (in navigation), or the azimuth (in astronomy and geometry).

(a) Find expressions for the basis vectors $\vec{e_r}$, $\vec{e_\theta}$ and $\vec{e_\phi}$ in terms of r, θ, ϕ, $\vec{e_x}$, $\vec{e_y}$ and $\vec{e_z}$.

$$\begin{aligned}\vec{r} = x\vec{e_x} + y\vec{e_y} + z\vec{e_z} &= r\sin\theta\cos\phi\vec{e_x} + r\sin\theta\sin\phi\vec{e_y} + r\cos\theta\vec{e_z}.\\ \vec{e_r} = \frac{\partial\vec{r}}{\partial r} &= \sin\theta\cos\phi\vec{e_x} + \sin\theta\sin\phi\vec{e_y} + \cos\theta\vec{e_z}.\\ \vec{e_\theta} = \frac{\partial\vec{r}}{\partial\theta} &= r\cos\theta\cos\phi\vec{e_x} + r\cos\theta\sin\phi\vec{e_y} - r\sin\theta\vec{e_z}\\ \vec{e_\phi} = \frac{\partial\vec{r}}{\partial\phi} &= -r\sin\theta\sin\phi\vec{e_x} + r\sin\theta\cos\phi\vec{e_y}.\end{aligned}$$

(b) Find expressions for the coefficients of the metric in spherical polar coordinates.

Take all the inner products of the three basis vectors with each other.

$$(g_{\mu\nu}) = \begin{pmatrix} \vec{e_r}\cdot\vec{e_r} & \vec{e_r}\cdot\vec{e_\theta} & \vec{e_r}\cdot\vec{e_\phi} \\ \vec{e_\theta}\cdot\vec{e_r} & \vec{e_\theta}\cdot\vec{e_\theta} & \vec{e_\theta}\cdot\vec{e_\phi} \\ \vec{e_\phi}\cdot\vec{e_r} & \vec{e_\phi}\cdot\vec{e_\theta} & \vec{e_\phi}\cdot\vec{e_\phi} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}$$

(c) Find expressions for the coefficients g^{ij} of the inverse metric.

$$(g^{\mu\nu}) = (g_{\mu\nu})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/r^2 & 0 \\ 0 & 0 & 1/(r^2\sin^2\theta) \end{pmatrix}$$

2.3 Consider a spherical surface, the set of all points at the same radius R from the origin. In terms of spherical coordinates a point on the surface can be specified by coordinates (θ, ϕ), which are the angle coordinates of the spherical coordinate system. You don't need R to be a coordinate because R is fixed. Work out the components of the metric in this coordinate system. Hint: because there are only two coordinates, the metric has three independent components. However, you'll need to start by describing points on the spherical surface in terms of their three Cartesian components in order to arrive at the result.

The basis vectors $\vec{e_\theta}$ and $\vec{e_\phi}$ are the same as in problem 2.2, except the radius r is a fixed scale R. The metric coefficients are subset of 4 of the 9 elements from problem 2.2 involving dot products between these two basis vectors. These can be with our without R, since it is a constant that can be factored out. If you leave it in, subsequent calculations of the Ricci scalar will reveal the dependence of the curvature on R.

$$(g_{\mu\nu}) = R^2 \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}$$

2.4 Consider a function $f(x, y)$ of two variables x and y. In turn, x and y are both functions of two further variables u and v, $x(u, v)$ and $y(u, v)$, By considering the expression for the total derivative of a function of two variables from the notes, show that

$$\left(\frac{\partial f}{\partial u}\right)_v = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial u}\right)_v + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial u}\right)_v .$$

In the case where $x(u, v)$ and $y(u, v)$ are invertible, so that $u(x, y)$ and $v(x, y)$ are also well defined functions, show in addition that

$$\left(\frac{\partial f}{\partial y}\right)_x = \left(\frac{\partial f}{\partial u}\right)_v \left(\frac{\partial u}{\partial y}\right)_x + \left(\frac{\partial f}{\partial v}\right)_u \left(\frac{\partial v}{\partial y}\right)_x .$$

Start with the total derivative df in terms of x and y,

$$df = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy$$

Substitute in the total derivatives dx and dy in terms of u and v,

$$dx = \left(\frac{\partial x}{\partial u}\right)_v du + \left(\frac{\partial x}{\partial v}\right)_u dv \qquad\qquad dy = \left(\frac{\partial y}{\partial u}\right)_v du + \left(\frac{\partial y}{\partial v}\right)_u dv$$

so that

$$\begin{aligned} df &= \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial u}\right)_v du + \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial v}\right)_u dv + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial u}\right)_v du + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial v}\right)_u dv \\ &= \left(\left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial u}\right)_v + \left(\frac{\partial f}{\partial y}\right)_u \left(\frac{\partial y}{\partial u}\right)_v\right) du + \left(\left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial v}\right)_u + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial v}\right)_u\right) dv \end{aligned}$$

Comparing the coefficients of du and dv between the above expressions for df, we recover the desired result.

If as stated $x(u, v)$ and $y(u, v)$ are invertible, we may start by considering the total derivative of f with respect to u and v, and then write the total derivatives du and dv in terms of x and y, therefore by the same manipulations as above we can prove the desired result for $\partial f/\partial y$.

2.5 By equating the Kronecker delta, δ^k_l with the partial derivative $\partial x^k/\partial x^l$, and using the chain rule for partial derivatives discussed in Problem 2.4, show that the components of δ^k_l transform as the components of a tensor of rank $\frac{1}{1}$.

Start by writing the Kronecker delta as the partial derivative of one component of x^j with respect to another coordinate x^k in the same coordinate system.

$$\delta^j_k = \frac{\partial x^j}{\partial x^k}.$$

Use the chain rule to express the relationships between partial derivatives in different coordinate systems. This is the proof in question 2.4 generalised to a function of n variables. In this case n is the number of dimensions in the space, and hence the number of coordinates x^k, where k runs over that number of dimensions.

$$\frac{\partial x^j}{\partial x^k} = \frac{\partial x^j}{\partial x^{m'}}\frac{\partial x^{m'}}{\partial x^k}$$

Apply this rule again on the right hand partial derivative.

$$\frac{\partial x^{m'}}{\partial x^k} = \frac{\partial x^{m'}}{\partial x^{q'}}\frac{\partial x^{q'}}{\partial x^k}.$$

Substitute this lower result into the upper one.

$$\begin{aligned}\delta^j_k &= \frac{\partial x^j}{\partial x^{m'}}\frac{\partial x^{m'}}{\partial x^{q'}}\frac{\partial x^{q'}}{\partial x^k}\\ &= \frac{\partial x^j}{\partial x^{m'}}\frac{\partial x^{q'}}{\partial x^k}\delta^{m'}_{q'}.\end{aligned}$$

This is the transformation law for the components of a tensor of rank $\frac{1}{1}$. The primed and unprimed indices may be interchanged,

$$\delta^{j'}_{k'} = \frac{\partial x^{j'}}{\partial x^m}\frac{\partial x^q}{\partial x^{k'}}\delta^m_q.$$

The general rules for the transformation laws for the components of tensors of arbitrary rank are given in Equation 2.32.

2.6 In Section 2.2 we saw that our basis vectors defined by $\vec{e_i} = \partial\vec{r}/\partial x^i$ are not generally of unit length. Show that a unit vector $\hat{e}_i$ and the basis vector $\vec{e_i}$ in the same direction as previously defined are related by (no Einstein sum over index j here) $\vec{e_j} = \sqrt{g_{jj}}\hat{e}_j$, or equivalently $\hat{e}_j = \vec{e_j}/\sqrt{g^{jj}}$.

Since $g_{ii} = \vec{e_i}\cdot\vec{e_i} = |\vec{e_i}|^2$, we arrive at $|\vec{e_i}| = \sqrt{g_{ii}}$. Therefore a unit vector $\hat{e}_i$ is obtained by dividing $\vec{e_i}$ by its magnitude, the required result.

2.7 In Section 2.9.4 we discussed creating symmetric and antisymmetric objects under interchange of indices. Consider a tensor of rank $\frac{0}{3}$ in three spatial dimensions with no