Excursions in Computing Science:
Book 11c. Geometry and Gravity
Part I. Fields, Complementary Coordinates, Curved Space

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

July 22, 2019

I. Prefatory Notes

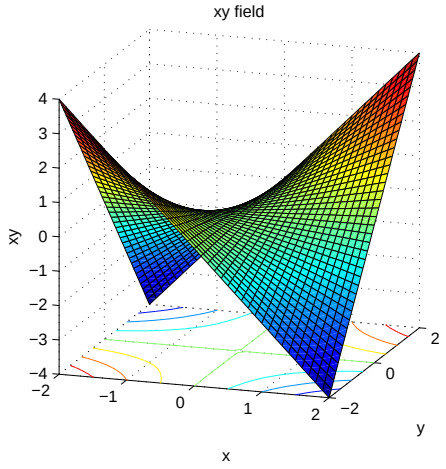
1. Fields and slopes. A function of two variables is called a "field" by the physicists, maybe because it might be describing a topography in which the height of the surface above, say sea level, is a function of two coordinates, say the longitude and latitude.

To be precise, physics would call this a *scalar* field (because the height is a single number at each point) in two dimensions (because of the two coordinates). We will soon encounter more ambitious types of field.

An example is

$$f(x, y) = xy$$

Here is what the surface xy looks like, plus a few of its values.



Field *xy* .

			y			
−9	−6	−3		3	6	9
−6	−4	−2		2	4	6
−3	−2	−1		1	2	3
			x			
3	2	1		−1	−2	−3
6	4	2		−2	−4	−6
9	6	3		−3	−6	−9

xy

What do we mean by the *slope* of this field?

The rate at which the surface rises depends on what *direction* we are going in. Given the two

coordinates, x and y, we evidently have two preferred directions to go in.

Looking at the numbers, we see that for $y = 1$, xy increases by 1 for each increase of x by 1: its slope is 1 when $y = 1$.

When $y = 2$ the slope with respect to x is 2.

When $y = 0$ the value of xy is constant (0) so its slope = 0.

When $y = -1$, xy goes down as x goes up, with a slope of -1.

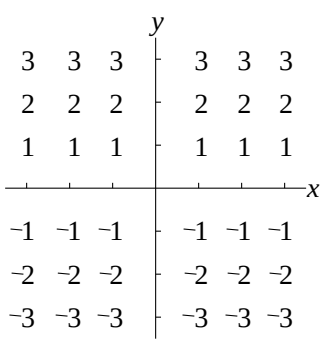
So it seems that the slope with respect to x of xy is y. This makes sense if we consider that we're fixing y at a constant value each time: the slope with respect to x of a constant times x is just that constant. So

$$\text{slope}_x xy = y$$

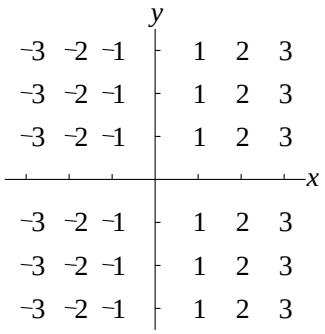
Similarly, we can fix x at different values and look at the slope as we go in the y-direction along that x-value. We get

$$\text{slope}_y xy = x$$

Slopes with respect to x and y of xy.



$\text{slope}_x = y$



$\text{slope}_y = x$

We can take this further and find *second slopes*. We can repeat the processes we did for the values of xy in the two cases of $\text{slope}_x xy = y$ and $\text{slope}_y xy = x$; or we can just write it formally.

$$\begin{aligned}
\text{slope}_x\text{slope}_x xy &= \text{slope}_x y = 0 \\
\text{slope}_y\text{slope}_x xy &= \text{slope}_y y = 1 \\
\text{slope}_x\text{slope}_y xy &= \text{slope}_x x = 1 \\
\text{slope}_y\text{slope}_y xy &= \text{slope}_y x = 0
\end{aligned}$$

(See how y is a constant with respect to x and vice-versa.)

The only nonzero result for xy is

$$\text{slope}_y\text{slope}_x xy = 1 = \text{slope}_x\text{slope}_y xy$$

Not surprisingly these two results are the same: it should not matter in which order we take the slopes.

2. "Reality" and coordinates. The physics notion of "field" is in a sense more solid than the mathematics notion of a function because it purports to describe something "real".

We'll take a limited view of "reality": "real" means existence independent of the coordinate system.

A coordinate system is a construct which we build in order to describe a field as a function, often in a way convenient to *us*, but quite unrelated to the "reality" of the field we're describing.

For example, if we rotated the coordinate system we've used to be able to describe our field as xy,

this description would no longer be correct, although the field did not go away and we must still be able to find a description of it.

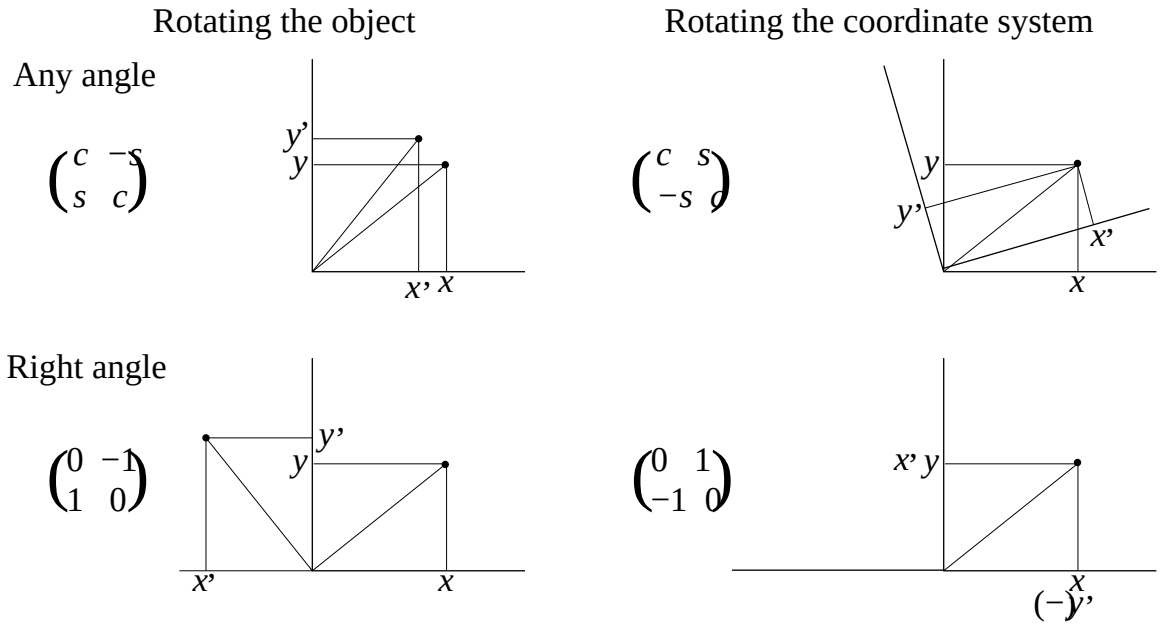
Here's the rotation from x, y to x', y'—and its inverse, which we'll find more useful.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \qquad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

where the rotation angle is θ, say, so c is $\cos\theta$ and s is $\sin\theta$.

You'll notice that the rotation matrix is opposite to the usual rotation, with $-s$ in the first column not the second. This is because we are rotating the *coordinate system* not the object.

Here's a diagram to help you work through the differences, first for an arbitrary angle (top) then for a right angle (bottom).



It's best to think of the arbitrary angle as small, so $c \approx 1$ and $s \approx 0$—but both positive. Then you can see why $x' < x$ and $y' > y$ when we rotate the object, but the opposite when we rotate the coordinate system.

We can redescribe the field in the new coordinate system by using the inverse transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} cx' - sy' \\ sx' + cy' \end{pmatrix}$$

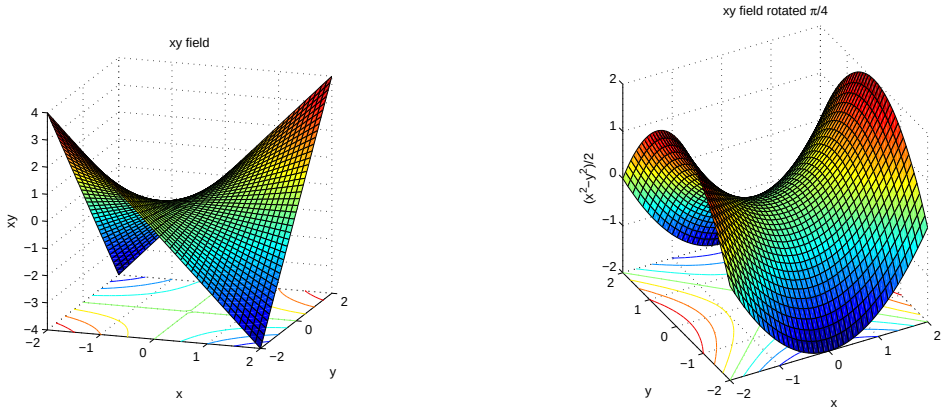
so

$$\begin{aligned} xy &= (cx' - sy')(sx' + cy') \\ &= cs(x'^2 - y'^2) + (c^2 - s^2)x'y' \\ &= \frac{1}{2}s_2(x'^2 - y'^2) + c_2 x'y' \end{aligned}$$

where the last line reexpresses the c, s expressions in terms of $c_2 = \cos(2\theta)$ and $s_2 = \sin(2\theta)$: this will be handy a little later in this Note.

Here are plots of the field in the two coordinate systems, side by side. This is meant to show that the field has not changed, although of course the sets of values for which it is plotted have conformed

to the corresponding coordinate systems. For concreteness, I've shown the rotation angle to be 45^o.



What in the math can we point to to say that the field has an existence independent of the coordinate suystem describing it? Nothing so far.

But if the field is "real" then so are its slopes, and we can find a connection between the transformation we've just seen describing the coordinate rotation, and the corresponding transformation of the slopes.

From our discussion in Note 1 about two-dimensional slopes, you should be able to do for x' and y' what we did there for x and y.

$$\begin{aligned} \text{slope}_{x'}(cs(x'^2 - y'^2) + (c^2 - s^2)x'y') &= 2csx' + (c^2 - s^2)y' \\ \text{slope}_{y'}(cs(x'^2 - y'^2) + (c^2 - s^2)x'y') &= -2csy' + (c^2 - s^2)x' \end{aligned}$$

This looks as if it could be written as a matrix equation so we'll put x' and y' into a vector and the slopes into another vector. We'll also here take a notational step and write, more briefly, and more conventionally.

$$\begin{aligned} \text{slope}_x &= \partial_x & \qquad \text{slope}_{x'} &= \partial_{x'} \\ \text{slope}_y &= \partial_y & \qquad \text{slope}_{y'} &= \partial_{y'} \end{aligned}$$

Thus

$$\begin{aligned} \begin{pmatrix} \partial_{x'} \\ \partial_{y'} \end{pmatrix} cs(x'^2 - y'^2) + (c^2 - s^2)x'y' &= \begin{pmatrix} 2cs & c^2 - s^2 \\ c^2 - s^2 & -2cs \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \\ &= \begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \end{aligned}$$

This matrix looks almost like a rotation matrix. We could make it so by swapping columns. And if we swap rows in the vector at the same time, the equation is still correct.

$$\begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} c_2 & s_2 \\ -s_2 & c_2 \end{pmatrix} \begin{pmatrix} y' \\ x' \end{pmatrix}$$

But our objective is to express $\partial_{x'}$ and $\partial_{y'}$ in terms of ∂_x and ∂_y. Now we know from Note 1 (and using the new notation)

$$\begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} xy = \begin{pmatrix} y \\ x \end{pmatrix}$$

and the transformation at the beginning of this Note is

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

so we won't do the swap just yet, but plug these two in.

$$\begin{aligned}
\begin{pmatrix} \partial_{x'} \\ \partial_{y'} \end{pmatrix} xy \quad (\text{in terms of } x',\ y') &= \begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \\
&= \begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\
&= \begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} \partial_y \\ \partial_x \end{pmatrix} xy \\
&= \begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix} \begin{pmatrix} s & c \\ c & -s \end{pmatrix} \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} xy \\
&= \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} xy
\end{aligned}$$

and this is the same matrix that gives x', y' from x, y.

(Warning: it's going to turn out that it is *not* the same matrix but the transpose of its inverse. We'll need new notation to find this out. In this example it's the same thing.)

Note that in the last line of the chain of equations above I took a quick step to multiply the s_2, c_2 matrix by the s, c matrix to get the c, s matrix: you can check this out by algebra, using $c^2 + s^2 = 1 = {c_2}^2 + {s_2}^2$, or you can do the Excursion that combines the double rotation with the single rotation back.

What about second slopes? We can write them as follows.

$$\begin{aligned}
\begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} (\partial_x \ \partial_y) xy &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\begin{pmatrix} \partial_{x'} \\ \partial_{y'} \end{pmatrix} (\partial_{x'} \ \partial_{y'}) xy &= \begin{pmatrix} \partial_{x'} \\ \partial_{y'} \end{pmatrix} (x' \ y') \begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix} \\
&= \begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix}
\end{aligned}$$

or we can apply our coordinate transformation twice to the final result

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c & -s \\ s & c \end{pmatrix} = \begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix}$$

So the same transformation works for second slopes as for first slopes, but just applied twice.

Every index transforms the same way. This is the math that says that an indexed variable has an existence independent of the coordinate system describing it.

The reason is that if we can write a relationship, say,

$$R_{\mathrm{jk}} - \frac{1}{2} R g_{\mathrm{jk}} = \kappa T_{\mathrm{jk}}$$

then the same relationship holds in a new coordinate system

$$R_{j'k'} - \frac{1}{2}Rg_{j'k'} = \kappa T_{j'k'}$$

after the coordinate transformation $\{j,k\} \rightarrow \{j',k'\}$.

The math has not changed even though the coordinate system has.

But note that matrix notation begins to get tricky. We must transpose matrices at the right places to get things right.

We need new notation which does not get tangled up in transposition.

Third slopes will make life impossible for matrices. The slopes themselves have three indices, not the two that matrices can express or the one that vectors can express. And to go from one three-index set of slopes to the transformed three-index set of slopes, we'll need to apply the transformation three times. Matrix/vector notation fails us.

3. Index notation and tensors. The new notation we need keeps track of indices explicitly. Thus, a transformation matrix, X, has elements

$$X_{jk}$$

Its transpose is

$$X_{kj}$$

and since each of these stand for an ordinary number, we can write them in any order in a product. All this is much easier than dealing with matrices. And we are not limited to one or two indices but can use any number.

We will, however, need indices for coordinates, too, so we must give up x and y and call them

$$x_1 \quad \text{and} \quad x_2$$

and similarly

$$x_{1'} \quad \text{and} \quad x_{2'}$$

Thus

$$x_{j'} = X_{j'k}x_k$$

expresses the transformation. We could have written

$$x_{j'} = \Sigma_k X_{j'k}x_k$$

with an explicit summation over the index k, but since k appears only on the right-hand side of the version that does not have the Σ (it is a "dummy index"), it is reasonable to interpret k as an index to be summed over. (This is the "Einstein summation convention".)

Note that the index notation in

$$x_{j'} = X_{j'k}x_k$$

generalizes the transformation to any number of dimensions. First, leaving x and y behind and introducing the parametrization x_k allows more (or less) than two dimensions. Second, replacing the explicit sum by Σ_k and then by the summation convention completes this parametrization.

To prove that the transformation from unprimed to primed coordinates applies also to slopes of all orders, we will need a further generalization for any linear transformation.

This involves perceiving that the (matrix) coefficients in a linear transformation are themselves slopes.

$$\begin{aligned} x' &= cx + sy = (\partial_x x')x + (\partial_y x')y \\ y' &= -sx + cy = (\partial_x y')x + (\partial_y y')y \end{aligned}$$

or

$$x_{j'} = (\partial_{x_k} x_{j'}) x_k$$

and, to avoid sub-subscripts (actually, I'm using midscripts, so "sub-midscripts" might best describe the above) we can rewrite this without ambiguity as

$$x_{j'} = (\partial_k x_{j'}) x_k$$

Now it is a rule of slopes that

$$\partial_{x'} f(x(x', y'), y(x', y')) = (\partial_x f)(\partial_{x'} x) + (\partial_y f)(\partial_{x'} y)$$

and similarly for $\partial_{y'} f$:

$$\partial_{y'} f(x(x', y'), y(x', y')) = (\partial_x f)(\partial_{y'} x) + (\partial_y f)(\partial_{y'} y)$$

So, together in matrix form, without the common f

$$\begin{pmatrix} \partial_{x'} \\ \partial_{y'} \end{pmatrix} = \begin{pmatrix} \partial_{x'} x & \partial_{x'} y \\ \partial_{y'} x & \partial_{y'} y \end{pmatrix} \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix}$$

which we compare with the inverse transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \partial_{x'} x & \partial_{y'} x \\ \partial_{x'} y & \partial_{y'} y \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

and we see immediately that the transformation from ∂_k to $\partial_{k'}$ is the transpose of the matrix of the transformation from x_k to $x_{k'}$.

$$\partial_{j'} = (\partial_{j'} x_k) \partial_k \quad \text{vs.} \quad x_k = (\partial_{j'} x_k) x_{j'}$$

We see that this notation is very compact: we must be very finicky to use it fluently.

The first-slope transformation

$$\partial_{k'} = (\partial_{k'} x_k) \partial_k$$

is the generalization of the first specific result in Note 2.

We can also show the second-slope transformation, provided that we make the *assumption* that the *transformation does not depend on* $x_{j'}$: $\partial_{j'}(\partial_{k'} x_k) = 0$.

Since

$$\partial_{k'} = (\partial_{k'} x_k) \partial_k$$

then

$$\begin{aligned} \partial_{j'} \partial_{k'} &= \partial_{j'}((\partial_{k'} x_k) \partial_k) \\ &= (\partial_{k'} x_k) \partial_{j'} \partial_k \\ &= (\partial_{k'} x_k)(\partial_{j'} x_j) \partial_j \partial_k \end{aligned}$$

or, using a single symbol with two indices for the second slope

$$\partial_{j'k'} = (\partial_{j'} x_j)(\partial_{k'} x_k) \partial_{jk}$$

This generalizes the matrix multiplication we had for two-dimensional second slopes in Note 2: we see that the same transformation (matrix) applies to each index; we also see where the matrix must be transposed if we look carefully at the order of indices in the last product $(\partial_{k'}x_k)\partial_{jk}$

The generalization also brings out explicitly the assumption we must make for it to work, This is an assumption we'll be prevented from making later, introducing new considerations to the mathematics.

We see the pattern and can go on to write the third-slope transformation without worrying about the limitations of matrices

$$\partial_{j'k'\ell'} = (\partial_{j'}x_j)(\partial_{k'}x_k)(\partial_{\ell'}x_\ell)\partial_{jk\ell}$$

(Of course, for the field xy the third slopes are zero, but you can now try another field, say x^2y^2.)

A mathematical construct such as this, which, when the coordinate system is changed, transforms the same way on each index, is called a *tensor*. Tensors imply a certain "reality", in the sense we've been discussing, for the entities they describe: tensor relationships do not change even if the coordinate system has been transformed.

For a coordinate transformation which is independent of the coordinates themselves (the assumption we made above), the slopes of a field are tensors. In a more general coordinate transformation, the slopes are not tensors, as we'll see in Note 11.

Since our notation can describe both tensors and non-tensors, we can discuss, generally, proto-tensors or *protors*.

4. Protors. Index notation is powerful, as we have just seen, but it is finicky, as we said, and it has two drawbacks in comparison with matrix notation. The first drawback is that it cannot suppress the indices and work with the whole aggregate of numbers that constitute a matrix or tensor or protor. The second drawback is related: that the only way to calculate specific examples is to write out all the (finicky) index values.

We cannot go back to matrices, but matrices are generalized by another structure (which, remarkably, was devised to manipulate the very large amounts of data held in secondary-storage *databases*).

This structure is the *relation*. The relational form of an array (vector, matrix, ..) is most commonly shown as a table giving indices and values all on the same footing.

Let's go back to the x, y to x', y' transformation in Note 2 for our first example.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

or

$$x_{j'} = X_{j'k}x_k$$

$x'($	j'	x'	$)$
	1	x'	
	2	y'	

or

$x'($	j'	x'	$)$
	1	x_1'	
	2	x_2'	

$X($	j'	k	X	$)$
	1	1	c	
	1	2	s	
	2	1	$-s$	
	2	2	c	

or

$x($	k	x	$)$
	1	x	
	2	y	

$x($	k	x	$)$
	1	x_1	
	2	x_2	

We are free to put as *data* (anything below the line under the *descriptors*) any symbols we like. The descriptors on the other hand contain all the indices, but no indexed variables, The symbol in front

of the parenthesis is the *relation name*. The symbols between the parentheses are the *attributes*. Note that it is possible to reuse the relation name as the attribute below which we can list the values in the vector or matrix.

So, how do we express matrix multiplication? We can do so independently of the data and using only the relation name and the attributes.

This is done with two independent but interacting "algebras", the *relational algebra* and the *attribute algebra*. Here are the expressions to combine X with x.

let `x'` **be equiv** + **of** `X * x` **by** `j'`
`x'` $\leftarrow$ [`j',x'`] **in** (`X` **ijoin** `x`)

The first statement is of the attribute algebra. Every symbol in it, X, x, j' and x', is an attribute of some relation. The new attribute, x', is defined by this statement to be a sum of the products of attributes X and x, grouped by each value of attribute j'.

The second statement is of the relational algebra. The symbols X, x and x' in it are relations and in addition the symbols j' and x' in the square brackets are attributes. This creates the new relation x' by "joining" relations X and x in a way which we'll describe next, and then "projects" only the attribute j and the newly defined atttribute x'.

We get away with polymorphic (ambiguous) reuse of names x, X and x' as both relations and attributes because the syntax makes clear which is which. But I'm not going to dwell on the syntax here, because it was devised for much more general operations than what we will need for tensors and protors. All we'll be needing is matrix and similar multiplications. For this, the index notation of Note 3 is sufficient, so we will not use relational or attribute algebra syntax after this, even though we will use relational representations to write out the detail and do the calculations.

But we must show what calculations are accomplished by the relational/attribute syntax so that we can adapt these calculations to tensors and protors in general.

We start out with the *join* (called **ijoin**, for "intersection join", in the syntax, and, more conventionally, "natural join" by the database community).

Here is how we join relations `X(j',k,X)` and `x(k,x)`: match rows of X and x that share values of k, and keep all attributes.

`X` **ijoin** `x(`	`j'`	`k`	`X`	`x`	`)`
	1	1	c	x	
	1	2	s	y	
	2	1	$-s$	x	
	2	2	c	y	

Here is how we apply the attribute algebra first to take the product $X \times x$, then to sum these products within groups of j'.

`X` **ijoin** `x(`	`j'`	`k`	`X`	`x`	`)`	`X*x`	`x'`
	1	1	c	x		cx	$cx + sy$
	1	2	s	y		sy	$cx + sy$
	2	1	$-s$	x		$-sx$	$-sx + cy$
	2	2	c	y		cy	$-sx + cy$

Finally, we complete the relational algebra statement by *project*ing only attributes j' and the new x'. (Note that duplicate rows are not included.)

`x'(`	`j'`	`x'`	`)`
	1	$cx + sy$	
	2	$-sx + cy$	

For a second example, let's try transforming the second slopes of xy. (We go back to using xy for the sake of concrete results.)

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}\begin{pmatrix} c & -s \\ s & c \end{pmatrix} = \begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix}$$

$$(\partial_{j'}x_j)(\partial_{k'}x_k)\partial_{j,k} = \partial_{j',k'}$$

Since we're dropping the relational/attribute algebra syntax we can avoid the redundant relation names.

Note that rows with value 0 can be omitted.

relation 1

j'	j	$\partial_{j'}x_j$
1	1	c
1	2	s
2	1	$-s$
2	2	c

relation 2

k'	k	$\partial_{k'}x_k$
1	1	c
1	2	s
2	1	$-s$
2	2	c

relation 3

j	k	$\partial_{j,k}$
1	2	1
2	1	1

Looking for common attributes to join on, we could join relations 1 and 3 (left to right, above) on j or relations 2 and 3 on k. We choose the latter. After the join we multiply values, sum, and project (the summing is trivial in this example: each sum has only one term).

k'	k	j	$\partial_{j'}x_j * \partial_{j,k}$
1	1	2	$c * 1$
1	2	1	$s * 1$
2	1	2	$-s * 1$
2	2	1	$c * 1$

k'	j	$\partial_{j'}x_j * \partial_{j,k}$
1	1	s
1	2	c
2	1	c
2	2	$-s$

Repeating this process with this result and relation 1 above is less trivial, both in multiplication and in summing. Note that rows may be rearranged at will, and so may columns as long as the attribute still labels its column.

j'	j	$\partial_{j'}x_j$
1	1	c
2	1	$-s$
1	2	s
2	2	c

j	k'	$\partial_{j'}x_j * \partial_{j,k}$
1	1	s
1	2	c
2	1	c
2	2	$-s$

(j'	j	k'	$\partial_{j'}x_j$	$\partial_{j'}x_j * \partial_{j,k}$)	$\Sigma\partial_{j'}x_j * \partial_{k'}x_k * \partial_{j,k}$
1	1	1	c	s	$2cs$
1	1	2	c	c	$c^2 - s^2$
2	1	1	$-s$	s	$c^2 - s^2$
2	1	2	$-s$	c	$-2cs$
1	2	1	s	c	(see 112)
1	2	2	s	$-s$	(see 111)
2	2	1	c	c	(see 211)
2	2	2	c	$-s$	(see 212)

So finally

j'	k'	$\partial_{j'k'}$
1	1	s_2
1	2	c_2
2	1	c_2
2	2	$-s_2$

In summary, a *relation* is a labelled set of columns with each row being unique and listed in no particular order. In the special relations we're using here, the last, or value, column has a value in each row such as a matrix element. All the preceeding columns are indices with integer (or sometimes symbolic) values. If the value column is 0 for any row, the entire row is omitted. Vectors are represented as *binary* relations (two columns) and matrices as *ternary* relations.

There is no reason to stop at ternary relations, with two index columns. Third slopes will be quaternary, with three index columns; fourth slopes will have pentuples for rows and so on. (So we could speak of "n-tuples" in general, being the rows of relations of "arity" n.)

To illustrate these higher orders we must pick another example field, such as $f(x, y) = x^2y^2$. Here are the untransformed slopes.

1st slopes

α	$\partial_\alpha x^2y^2$
1	$2xy^2$
2	$2x^2y$

2nd slopes

α	β	$\partial_{\alpha\beta} x^2y^2$
1	1	$2y^2$
1	2	$4xy$
2	1	$4xy$
2	2	$2x^2$

3rd slopes

α	β	γ	$\partial_{\alpha\beta\gamma} x^2y^2$
1	1	2	$4y$
1	2	1	$4y$
1	2	2	$4x$
2	1	1	$4y$
2	1	2	$4x$
2	2	1	$4x$

4th slopes

α	β	γ	δ	$\partial_{\alpha\beta\gamma\delta} x^2y^2$
1	1	2	2	4
1	2	1	2	4
2	1	1	2	4
1	2	2	1	4
2	1	2	1	4
2	2	1	1	4

The join-product-sum process works equally well at these higher arities.

We define a *protor* in terms of this relational representation of indexed variables, whether they are tensors or not.

An n-protor is an $(n + 1)$-ary relation with n index columns and one column of values. If these values are symbolic and would depend on positions as given by the coordinates, the protor is a *protor field*. Otherwise, that is, if the symbolic values depend only on a parameter such as the angle of rotation in c and s, or if the values are simply numbers, the protor is a *constant protor*.

Note that all vectors and matrices are protors, although they are not necessarily tensors.

Note again that any row whose value column is zero is simply omitted.

5. The protor calculator 1. The relational and attribute algebras are implemented in the *alg*ebraic *dat*a programming language *Aldat* which interprets syntax such as that used in Note 4. For two reasons we shall rewrite the parts relevant to protors in MATLAB. The first reason is that, because of the power of Aldat, it requires us to learn new ways of thinking about data aggregates, which we won't need if we stick to the familiar MATLAB. The second reason is that we'll need symbolic math, which is not implemented in any current version of Aldat, although it fits seamlessly into the attribute algebra.

For present purposes, the protor calculator consists of two MATLAB functions. First, `slopes()`,

```
slopes(in,vars)
```

e.g.,

```
>> syms x y
>> slopes(x*y,[x,y])
ans =
[ x, y]
[ y, x]
>> slp1x2y2 = slopes(x^2*y^2,[x,y])
slp1x2y2 =
[ x, 2*x*y^2]
[ y, 2*x^2*y]
>> slp2x2y2 = slopes(slp1x2y2,[x,y])
slp2x2y2 =
[ x, x, 2*y^2]
[ x, y, 4*x*y]
[ y, x, 4*x*y]
[ y, y, 2*x^2]
```

```
>> slp3x2y2 = slopes(slp2x2y2,[x,y])
slp3x2y2 =
[ x, x, y, 4*y]
[ x, y, x, 4*y]
[ x, y, y, 4*x]
[ y, x, x, 4*y]
[ y, x, y, 4*x]
[ y, y, x, 4*x]
>> slp4x2y2 = slopes(slp3x2y2,[x,y])
slp4x2y2 =
[ x, x, y, y, 4]
[ x, y, x, y, 4]
[ x, y, y, x, 4]
[ y, x, x, y, 4]
[ y, x, y, x, 4]
[ y, y, x, x, 4]
>> slp5x2y2 = slopes(slp4x2y2,[x,y])
slp5x2y2 = allzeros
```

From these examples we see how protors are represented as MATLAB arrays, one row per tuple. The first column of each result gives the index the slope has been taken with respect to. This is why it is appropriate to use the symbolic names rather than integers for the indices. (The MATLAB code rearranges the rows of each result, so you must look a little carefully to find the correspondences between input rows and output rows.)

The second MATLAB function is `joinred()`.

```
joinred(in1,col1,in2,col2)
```

e.g.,

```
>> syms c s
>> X = [1,1,c;1,2,s;2,1,-s;2,2,c]
X =
[ 1, 1,  c]
[ 1, 2,  s]
[ 2, 1, -s]
[ 2, 2,  c]
```

```
>> v = [1,x;2,y]
v =
[ 1, x]
[ 2, y]
>> joinred(X,2,v,1)
ans =
[ 1, c*x + s*y]
[ 2, c*y - s*x]
```

Here we joined X with v by matching the 2nd column of X with the 1st column of v.

The protor calculations to find the slopes of x^2y^2 transformed by

$(\alpha'$	α	X	$)$
1	1	c	
1	2	s	
2	1	$-s$	
2	2	c	

are

$$\begin{aligned}
\partial_{\alpha'} &= X_{\alpha'\alpha}\partial_{\alpha}x^2y^2 && \texttt{joinred(X,2,slp1x2y2,1)}\\
\partial_{\alpha'\beta'} &= X_{\alpha'\alpha}X_{\beta'\beta}\partial_{\alpha\beta}x^2y^2 && \texttt{joinred(X,2,joinred(X,2,slp2x2y2,2),2)}\\
\partial_{\alpha'\beta'\gamma'} &= X_{\alpha'\alpha}X_{\beta'\beta}X_{\gamma'\gamma}\partial_{\alpha\beta\gamma}x^2y^2 && \texttt{joinred(X,2,joinred(X,2,joinred(X,2,slp3x2y2,3),3),3)}\\
\partial_{\alpha'\beta'\gamma'\delta'} &= X_{\alpha'\alpha}X_{\beta'\beta}X_{\gamma'\gamma}X_{\delta'\delta}\partial_{\alpha\beta\gamma\delta}x^2y^2 && \texttt{joinred(X,2,joinred(X,2,joinred(X,2,joinred(X,2,slp4x2y2,4),4),4),4)}
\end{aligned}$$

with the following results

1st slopes

α'	$\partial_{\alpha'}x^2y^2$
x	$2xy(cy+sx)$
y	$2xy(cx-sy)$

2nd slopes

α'	β'	$\partial_{\alpha'\beta'}x^2y^2$
x	x	$2(cy)^2+8csxy+2(sx)^2$
x	y	$2cs(x^2-y^2)+4xy(c^2-s^2)$
y	x	$2cs(x^2-y^2)+4xy(c^2-s^2)$
y	y	$2(sy)^2-8csxy+2(cx)^2$

3rd slopes

α'	β'	γ'	$\partial_{\alpha'\beta'\gamma'}x^2y^2$
x	x	x	$12cs(sx+cy)$
x	x	y	$-4s^3x+8cs(cx-sy)+4c^3y$
x	y	x	$-4s^3x+8cs(cx-sy)+4c^3y$
x	y	y	$4c^3x-8cs(sx+cy)+4s^3y$
y	x	x	$-4s^3x+8cs(cx-sy)+4c^3y$
y	x	y	$4c^3x-8cs(sx+cy)+4s^3y$
y	y	x	$4c^3x-8cs(sx+cy)+4s^3y$
y	y	y	$-12cs(cx-sy)$

4th slopes

α'	β'	γ'	δ'	$\partial_{\alpha'\beta'\gamma'\delta'}x^2y^2$
x	x	x	x	$24c^2s^2$
x	x	x	y	$12cs(c^2-s^2)$
x	x	y	x	$12cs(c^2-s^2)$
x	x	y	y	$4c^2-16c^2s^2+4s^2$
x	y	x	x	$12cs(c^2-s^2)$
x	y	x	y	$4c^2-16c^2s^2+4s^2$
x	y	y	x	$4c^2-16c^2s^2+4s^2$
x	y	y	y	$-12cs(c^2-s^2)$
y	x	x	x	$12cs(c^2-s^2)$
y	x	x	y	$4c^2-16c^2s^2+4s^2$
y	x	y	x	$4c^2-16c^2s^2+4s^2$
y	x	y	y	$-12cs(c^2-s^2)$
y	y	x	x	$4c^2-16c^2s^2+4s^2$
y	y	x	y	$-12cs(c^2-s^2)$
y	y	y	x	$-12cs(c^2-s^2)$
y	y	y	y	$24c^2s^2$

It is important to note the column chosen for the joining attribute of the original slope: it is always the last index column.

This is because `joinred()` does not work with attribute names the way Aldat would, but with their positions. It outputs index columns in the order in which they appear in the input protors, omitting the join index from both protors.

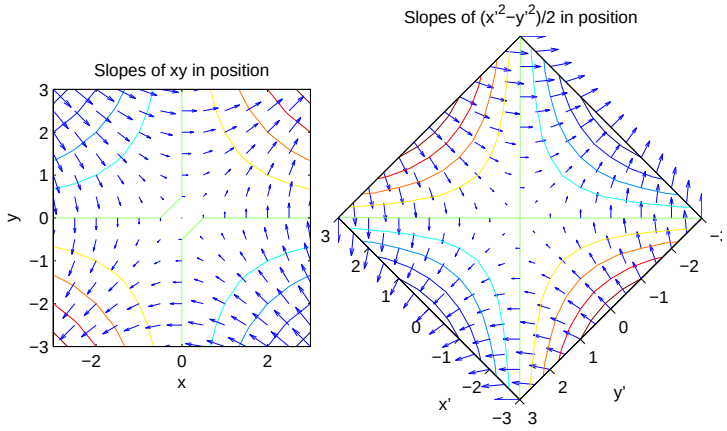
For fourth slopes, for example, with inputs X (four times) and $\partial_{\alpha\beta\gamma\delta}$, we have the following intermediate steps.

$$\begin{aligned} X_{\delta'\delta}, \partial_{\alpha\beta\gamma\delta} &\rightarrow \partial_{\delta'\alpha\beta\gamma} \\ X_{\gamma'\gamma}, \partial_{\delta'\alpha\beta\gamma} &\rightarrow \partial_{\gamma'\delta'\alpha\beta} \\ X_{\beta'\beta}, \partial_{\gamma'\delta'\alpha\beta} &\rightarrow \partial_{\beta'\gamma'\delta'\alpha} \\ X_{\alpha'\alpha}, \partial_{\beta'\gamma'\delta'\alpha} &\rightarrow \partial_{\alpha'\beta'\gamma'\delta'} \end{aligned}$$

(So we must be careful if we decide to check the intermediate steps of the calculation: the index columns are not labelled with their attributes so we must keep track of their positions.)

6. Divergence and Curl. The first slope of the *scalar field* xy, $\partial_\alpha xy$, is its *gradient*. This is a *vector field* giving the direction and magnitude of the steepest slope at each point.

Here are pictures of the gradients of xy and of xy rotated by 45 degrees to $cs(x'^2-y'^2)+(c^2-s^2)x'y'$.



The *divergence* and *curl* are two important properties of *vector* fields, producing respectively a scalar and a tensor field.

Instead of discussing the divergence and curl of the gradient we've just seen of xy (since both are in fact zero) we'll look at some new vector fields as examples.

They are

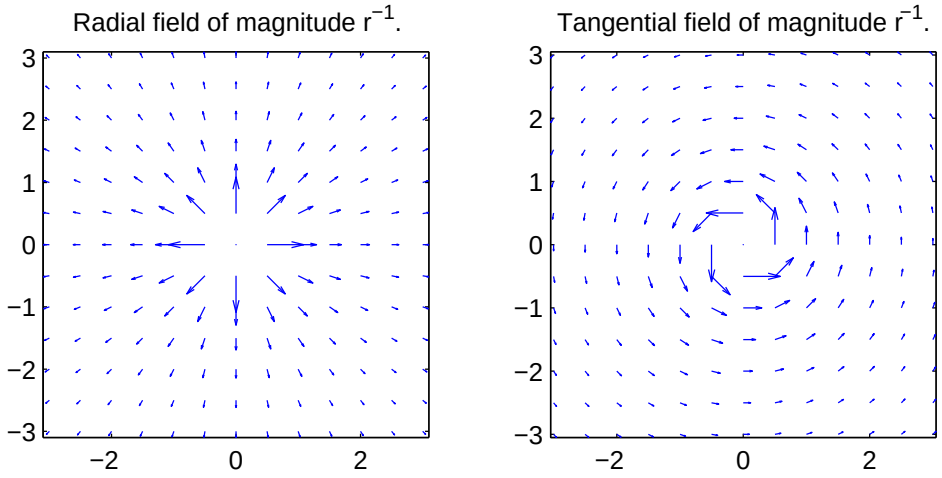
$$\begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} r^{p-1} \qquad \text{and} \qquad \begin{pmatrix} h_x \\ h_y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} r^{p-1}$$

where $r = \sqrt{x^2 + y^2}$ (and so

$$\begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} r = \begin{pmatrix} x \\ y \end{pmatrix} \tfrac{1}{r}).$$

The first can illustrate divergence but has zero curl. The second can illustrate curl but has zero divergence.

Here is what these two fields look like for $p = -1$. We can see intuitively that the first diverges and that the second curls.



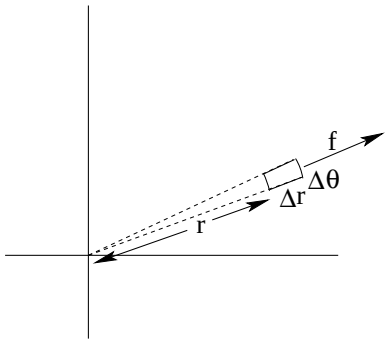
Both divergence and curl can be found from the first slopes.

$$\begin{aligned} \textbf{div} &= \partial_x f_x + \partial_y f_y \\ \textbf{curl} &= \partial_x h_y - \partial_y h_x \end{aligned}$$

We can derive these same results from a physical, and microscopic, viewpoint. We'll use three-dimensional vocabulary despite this being a two dimensional example. "Area" means the length of a segment of the boundary and "volume" means the area of the region enclosed by the boundary. We start with a "field" which can be thought of as a flow per unit area, convert that into a total flow at one boundary, find the net flow over the region (which is a difference, which over a small

enough region can be found by a slope), and finally find that net flow per unit volume in the region.

Divergence



This example is a field in the radial direction, with "flow" per unit "area" $f = r^p$ and "volume" of the region $= r\Delta\theta\Delta r$.

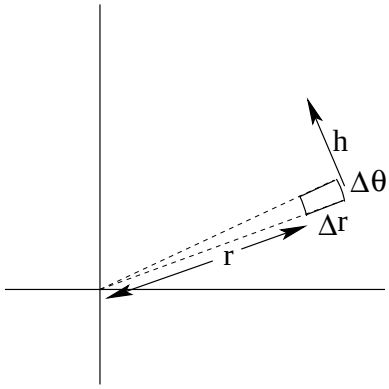
- total flow at boundary: $fr\Delta\theta = r^{p+1}\Delta\theta$;
- net flow from region = slope of this = $(p+1)r^p\Delta r\Delta\theta$;
- net flow per unit volume = $(p+1)r^{p-1}$ (which is zero if $p = -1$, the classical gravitational field in 2D).

In x-y coordinates, the field is $(x, y)r^{p-1}$ and

$$\begin{aligned}
\text{div} &= \partial_x(xr^{p-1}) + \partial_y(yr^{p-1}) \\
&= r^{p-1} + x(p-1)r^{p-2}\frac{x}{r} \\
&\quad + r^{p-1} + y(p-1)r^{p-2}\frac{y}{r} \\
&= 2r^{p-1} + (x^2+y^2)(p-1)r^{p-3} \\
&= (p+1)r^{p-1}
\end{aligned}$$

(In 3D this is $(p+2)r^{p-1}$ which is zero for $f = 1/r^2$; in d dimensions the generalization is $(p-1+d)r^{p-1}$.)

Curl



This example is a field in the tangential direction, also with "flow" per unit "area" $h = r^p$ and "volume" of the region $= r\Delta\theta\Delta r$.

- total flow at boundary: $hr\Delta\theta = r^{p+1}\Delta\theta$;
- net flow perpendicular to region = slope of this = $(p+1)r^p\Delta r\Delta\theta$;
- net flow per unit volume = $(p+1)r^{p-1}$ (which is again zero if $p = -1$).

In x-y coordinates, the field is $(-y, x)r^{p-1}$ and

$$\begin{aligned}
\text{curl} &= \partial_x(xr^{p-1}) - \partial_y(-yr^{p-1}) \\
&= r^{p-1} + x(p-1)r^{p-2}\frac{x}{r} \\
&\quad -(-r^{p-1} - y(p-1)r^{p-2}\frac{y}{r}) \\
&= 2r^{p-1} + (x^2+y^2)(p-1)r^{p-3} \\
&= (p+1)r^{p-1}
\end{aligned}$$

This is a simple variant on the divergence. A similar discussion to that for divergence holds for higher dimensions.

We notice a remarkable thing in each case. The divergence and curl depend on the magnitude as well as the direction of the vector field, and for $p = -1$ both are zero. This is just the value of p we used to generate the intuitive pictures we showed above. Intuition says the one field diverges and the other curls, but calculation says that each field does neither.

The intuitive pictures look much the same with $p = -2$, and in this case divergence and curl are neither zero. (However, in three dimensions, $p = -2$ does make divergence and curl vanish in the analogous fields.)

In general, for vector fields f and h

$$\mathbf{div}.\mathbf{f} = \partial_j f_j$$

$$(\mathbf{curl} \times \mathbf{h})_{jk} \quad = \quad \partial_j h_k - \partial_k h_j$$

where **div.f** is just the notation for the dot product of two vectors giving a scalar, but **curl** $\times$ **h** is notation which can be considered to define a two-index tensor.

Thus in three dimensions

$$\begin{aligned}
\mathbf{div.f} &= \partial_x f_x + \partial_y f_y + \partial_z f_z \\
(\mathbf{curl} \times \mathbf{h})_{xy} &= \partial_x h_y - \partial_y h_x \\
(\mathbf{curl} \times \mathbf{h})_{yz} &= \partial_y h_z - \partial_z h_y \\
(\mathbf{curl} \times \mathbf{h})_{zx} &= \partial_z h_x - \partial_x h_z
\end{aligned}$$

See the Excursions for the examples

$$\begin{aligned}
\mathbf{f} &= \begin{pmatrix} x \\ y \\ z \end{pmatrix} r^{p-1} \\
\mathbf{h} &= \begin{pmatrix} z-y \\ x-z \\ y-x \end{pmatrix} r^{p-1}
\end{aligned}$$

We can show that divergence is invariant—it has the same form no matter what the coordinate system is—using a general argument in index notation.

We have found (Note 3)

$$\partial_{j'} = (\partial_{j'} x^j)\partial_j$$

so of course

$$\begin{aligned}
\partial_j &= (\partial_j x^{j'})\partial_{j'} \\
&= (\partial_j x^{j'})(\partial_{j'} x^k)\partial_k
\end{aligned}$$

and so

$$(\partial_j x^{j'})(\partial_{j'} x^k) \quad = \quad \delta_{jk} \stackrel{\text{def}}{=} \begin{cases} 1 \text{ if } j = k \\ 0 \text{ otherwise} \end{cases}$$

i.e., these two are inverses of each other.

We also know (Note 3)

$$x_{j'} = (\partial_j x_{j'}) x_j$$

because, if for coefficients $a, b, ..$

$$x_{j'} = a x_1 + b x_2 + ..$$

then

$$\begin{aligned}
a &= \partial_1 x_{j'} \\
b &= \partial_2 x_{j'}
\end{aligned}$$

and so on.

To go from the divergence $\partial_j f_j$ to $\partial_{j'} f_{j'}$ in the new coordinate system $x_{j'} = (\partial_j x_{j'}) x_j$ we need two

steps: 1) transform the slopes $\partial_{j'} = (\partial_{j'} x^j)\partial_j$; 2) transform the axes $f_j \to f_{j'}$ which is the same linear transformation as $x_j \to x_{j'}$.

We'll write it in reverse

$$\begin{aligned} \partial_{j'} f_{j'} &= (\partial_{j'} x_j)\partial_j(\partial_k x_{j'}) f_k \\ &= (\partial_{j'} x_j)(\partial_k x_{j'})\partial_j f_k \\ &= \partial_j f_j \end{aligned}$$

In the second line, we've repeated our earlier assumption that the transformation $(\partial_k x_{j'})$ does not depend on the coefficients x_j (Note 3). (If it does, the invariant form of the divergence is somewhat different, as we shall see in Note 15.)

In the third line we use the fact that the two transformations (for slopes and for coordinates) are inverses of each other (Note 3).

7. Classical gravity. Newton proposed that the gravity of a body such as the Sun or the Earth is a force, acting instantaneously at a distance, directed toward the body with a magnitude proportional to $1/r^2$ with r the distance from the body.

From the point of view of Euclidean geometry, this is an almost inevitable perspective on the influence of the body, which must be diluted with r in proportion to the area that the influence must affect at that distance: $4\pi r^2$.

In two dimensions, the "area" the influence gets spread over at distance r from the body is $2\pi r$. So 2D gravity would be a $1/r$ force.

Newtonian gravity came to be regarded as a field, x_j/r^3, or, in 2D x_j/r^2. These are both examples of the $x_j r^{p-1}$ field of Note 6, with $p = -2$ in 3D and $p = -1$ in 2D.

The *source* of this gravitational field would be the body that causes it, and was related by Newton to the mass of that body.

Mathematically, it makes sense to relate the *source* to the divergence of the field. (Since the field is incoming—directed towards the origin at the body—that divergence will be negative. Mathematically the slope of an inverse power is also negative.)

Here we run into a problem. We found in Note 6 that the divergence of $x_j r^{p-1}$ is $(p + d - 1)r^{p-1}$ in d dimensions: zero in both cases.

Well, of course, that's the divergence anywhere in the space with the r^{1-d} field in it. And there is no source anywhere in that space, so the divergence *should* be zero. (Indeed, from Euclid's perspective, it is quite natural that the divergence should be zero.)

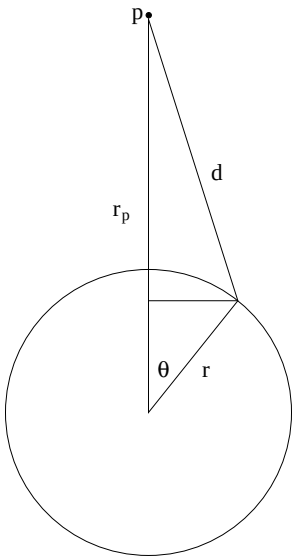
The source is only at the origin, $r = 0$, and we haven't put it in. Mathematically, the fields r^{1-d} have a singularity here, so physically we can't go all the way to $r = 0$.

We'll need a new field in the space around the origin that represents the body and the source.

Suppose a body of mass m_b occupies a sphere of radius r_b and has uniform *density* $\rho = m_b/(4\pi r_b^3/3)$. Then we can show that, anywhere *inside* that sphere (imagine a very thin hole bored all the way along a diameter), the field is radial and proportional to ρr, i.e.,

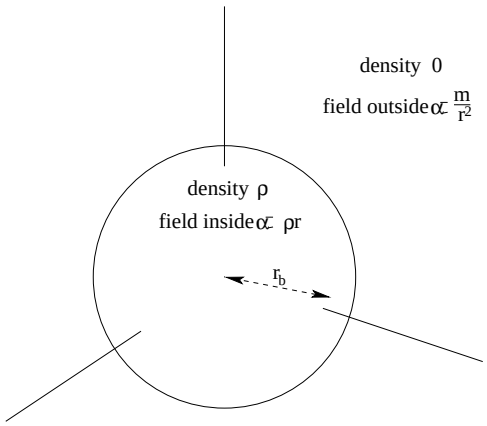
$$\rho \begin{pmatrix} x \\ y \end{pmatrix}$$

Then the divergence is nonzero and proportional to ρ. This mass density will be the source of the gravitational field.



In the figure, we imagine a spherical shell of matter which is creating a field at point p a distance r_p from the centre of the shell. We can put point p on the z-axis without losing generality, because the coordinate system can always be set up so that the z-axis contains p.

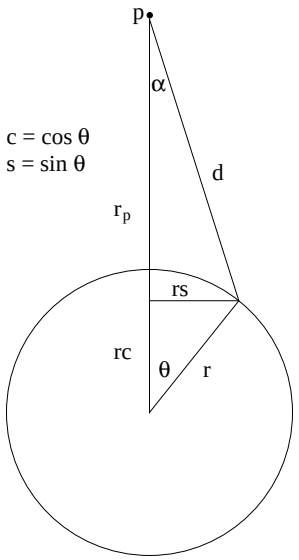
The angle θ is the usual coordinate of a location on the shell, and the angle ϕ is not shown because of symmetry about the z-axis.



We need two arguments. First that the spherical shell, of radius r, thickness Δr and mass $m = 4\pi r^2 \Delta r \rho$ attracts anything *outside* that shell (at radius $r_p > r$) proportionally to m/r_p^2 as if the mass were all at a point source at the origin.

The second argument is that the attraction of this same shell for anything *inside* it (at radius $r_p < r$) is zero.

These arguments require some setting up and rearranging of variables, followed by a two-dimensional antislope over (effectively) ϕ from 0 to 2π and θ from 0 to π.



First some geometry. Equating $(rs)^2 = (rs)^2$ in the two Pythagorean triangles sharing that side,

$$d^2 - (r_p - rc)^2 = r^2 - (rc)^2$$

so

$$2r_p rc = r_p^2 + r^2 - d^2$$

where $c = \cos\theta$. We're going to need $s = \sin\theta$ as well, because, with a small change in θ

$$\Delta(2r_p rc) = \Delta(r_p^2 + r^2 - d^2)$$

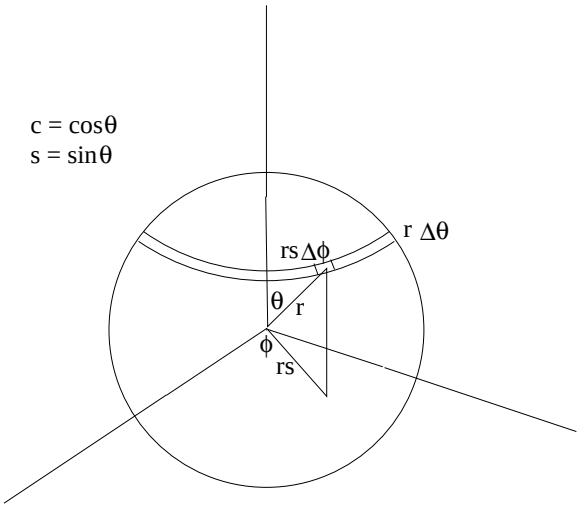
gives

$$-2r_p rs\Delta\theta = -2d\Delta d$$

by taking slopes with respect to θ on the left and with respect to d on the right: note that d and θ co-vary—as θ goes from 0 to π, d goes from $r_p - r$ to $r_p + r$.

Because of symmetry—each part of the shell is part of a ring all the same distance d from point p—we will need only the z-component of the field, not the component along line d. So we need a factor

$$\begin{aligned}\cos\alpha &= \frac{r_p - rc}{d} \\ &= \frac{2r_p^2 - 2r_p rc}{2r_p d} \\ &= \frac{d^2 + r_p^2 - r^2}{2r_p d}\end{aligned}$$



The contribution of a little part of the mass of the shell

$$\Delta m = \rho(\Delta r)(\Delta\theta)rs(\Delta\phi)$$

at distance d from p to the field at p is thus

$$\begin{aligned}
\Delta f &\approx \frac{\Delta m}{d^2}\cos\alpha \\
&= \frac{\rho r(\Delta r)rs(\Delta\theta)(\Delta\phi)}{d^2}\frac{d^2+r_p^2-r^2}{2r_p d} \\
&= \frac{\rho r(\Delta r)d(\Delta d)(\Delta\phi)}{r_p d^2}\frac{d^2+r_p^2-r^2}{2r_p d}
\end{aligned}$$

Including the whole ring, $\phi = 0$ to 2π, replaces $\Delta\phi$ by 2π. And including the whole shell, $\theta = 0$ to π, which is to say, $d = r_p - r$ to $r_p + r$, requires the antislope of this.

$$\begin{aligned}
f &\propto \text{antislope}_{d=r_p-r:r_p+r}\frac{2\pi r\rho}{2r_p^2}\frac{d^2+r_p^2-r^2}{d^2} \\
&= \frac{\pi r\rho\Delta r}{r_p^2}\text{antislope}_{d=r_p-r:r_p+r}\left(1+\frac{r_p^2-r^2}{d^2}\right) \\
&= \frac{\pi r\rho\Delta r}{r_p^2}\left(2r - \frac{r_p^2-r^2}{d}\Big|_{r_p-r}^{r_p+r}\right) \\
&= \frac{\pi r\rho\Delta r}{r_p^2}4r \\
&= \frac{4\pi r^2\Delta r\rho}{r_p^2} \\
&= \frac{m}{r_p^2}
\end{aligned}$$

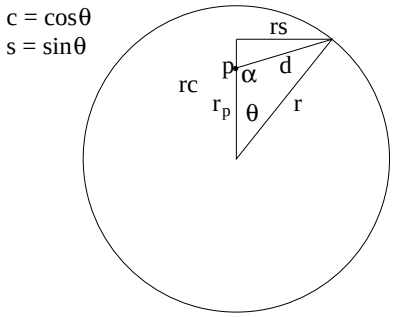
where $m = 4\pi r^2\Delta r\rho$ is the mass of the whole shell.

Thus the shell acts gravitationally as if the entire mass were concentrated at the point that is the

centre of the sphere.

A solid sphere is made up of concentric shells, and the same argument applies to each of them.

So a spherical body acts gravitationally on any point exterior to it as though its whole mass was concentrated at its centre.



What about a point *inside* a spherical shell? The argument above applies except that now, as θ ranges from 0 to π, d ranges from $r - r_p$ to $r + r_p$

$$\begin{aligned} f &\propto \frac{\pi r \rho \Delta r}{r_p^2}\left(2r_p - \frac{r_p^2 - r^2}{d}\Big|_{r-r_p}^{r+r_p}\right) \\ &= \frac{\pi r \rho \Delta r}{r_p^2} \times 0 \end{aligned}$$

So a spherical shell has *no* gravitational effect on a point inside it.

From these two results it follows that a point p in a narrow tunnel through a solid sphere of uniform density ρ, at a distance r from the centre, will experience a field, due to all shells of radius $< r$, proportional to

$$\frac{m}{r^2} = \frac{4\pi r^3 \rho/3}{r^2} = \frac{4}{3}\pi r \rho$$

Thus we have results for the gravitational field at points a) in free space with a spherical mass m centred a distance r away and b) within a sphere of mass density ρ.

To turn the proportionalities into equalities we need Newton's gravitational constant G_N so that the field represents the *acceleration* of the point p towards the centre of mass,

a) free space

$$f = -\frac{G_N m}{r^2}$$

b) within sphere

$$f = -\frac{4}{3}\pi r G_N \rho$$

These fields are vector fields, because the acceleration is directed, so strictly

a) free space

$$f_j = -\frac{G_N m}{r^3} x_j$$

b) within sphere

$$f_j = -\frac{4}{3}\pi x_j G_N \rho$$

Each field can, if we wish, be expressed as the gradient of a scalar field, called the *gravitational*

potential

a) free space

$$\boldsymbol{f} = G_N m \, \mathbf{grad} \frac{1}{r}$$

b) within sphere

$$\boldsymbol{f} = -\frac{2}{3}\pi G_N \rho \, \mathbf{grad} \, r^2$$

Going the other way, we notice an important property of the divergence of the gravitational field.

a) free space

$$\mathbf{div} \cdot \mathbf{f} = 0$$

b) within sphere

$$\mathbf{div} \cdot \mathbf{f} = -4\pi G_N \rho$$

These are the same result if we consider that, in free space, the density of matter $\rho = 0$. And this is a general result, which can be taken as an abstract formulation of classical gravity.

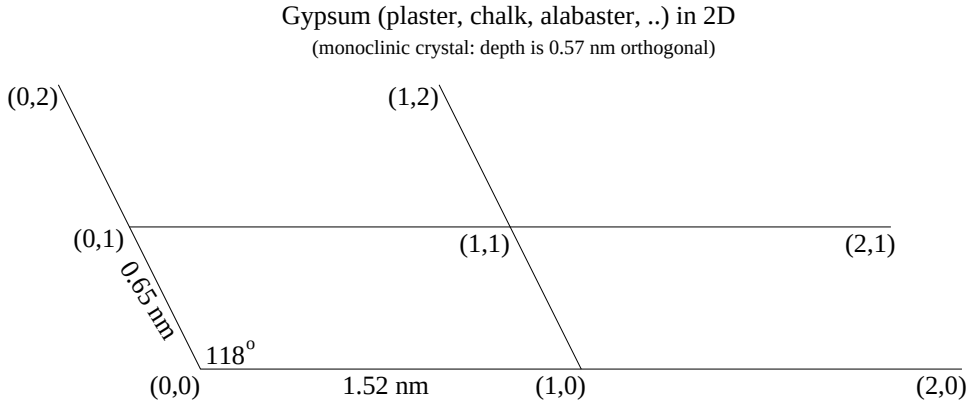
$$\mathbf{div} \cdot \mathbf{f} = -4\pi G_N \rho$$

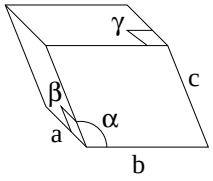
Or, in terms of the potential, ϕ ($\mathbf{f} = \mathbf{grad}\ \phi$),

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

Note for future reference that this is formed from *second slopes* of the potential ϕ.

8. Gypsum coordinates. It is sometimes appropriate to use a coordinate system in which the basis vectors are neither orthogonal nor of unit length. An example is a coordinate system which indexes sites in a crystal with integers. Here is a two-dimensional representation of gypsum.

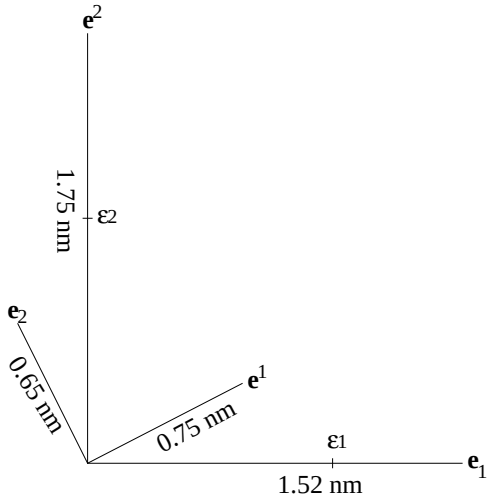




$\alpha \neq 90^o$
$\beta = 90^o$
$\gamma = 90^o$

Gypsum sand at White Sands, New Mexico MAM

Gypsum is a monoclinic crystal of $CaSO_4.2H_2O$, occurring naturally in great quantities in, for example, White Sands New Mexico and used for fertilizer, drywall ($\gamma\upsilon\psi o\varsigma$ means "plaster") and, as alabaster, for sculpture. It provides a simple example of a non-orthogonal coordinate system.



Here is a diagram showing three different sets of basis vectors for (2-dimensional) gypsum. The conventional basis $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$ is orthonormal. We've been using "midscripts" so far in these Notes, and will continue to do so for orthonormal basis vectors and for coordinate systems based on them.

The basis vectors that parallel the natural axes of gypsum, $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, are written with subscripts, and coordinates based on them will be written with superscripts.

A third pair of basis vectors, $\boldsymbol{e}^1$ and $\boldsymbol{e}^2$, is complementary, in the sense that $\boldsymbol{e}^j.\boldsymbol{e}_k = \delta^j_k$:

$$\begin{aligned} \boldsymbol{e}^1.\boldsymbol{e}_1 &= 1 & \qquad \boldsymbol{e}^1.\boldsymbol{e}_2 &= 0 \\ \boldsymbol{e}^2.\boldsymbol{e}_1 &= 0 & \qquad \boldsymbol{e}^2.\boldsymbol{e}_2 &= 1 \end{aligned}$$

i.e., the complementary basis is orthogonal to the gypsum basis, and normal to it in the above sense. The complementary basis vectors are written with superscripts and coordinates based on them will be written with subscripts.

Here are the cosines and sines for the gypsum basis and the complementary basis. (Numerical values are for (two-dimensional) gypsum. Angles are in degrees.)

$$c_1 = \cos(\angle \mathbf{e}_1) = \cos 0 = 1 \qquad c_2 = \cos(\angle \mathbf{e}_2) = \cos 118.43$$

$$s_1 = \sin(\angle \mathbf{e}_1) = \sin 0 = 0 \qquad s_2 = \sin(\angle \mathbf{e}_2) = \sin 118.43$$

$$c^1 = \cos(\angle \mathbf{e}^1) = \cos 28.43 = s_2 \qquad c^2 = \cos(\angle \mathbf{e}^2) = \cos 90 = -s_1$$

$$s^1 = \sin(\angle \mathbf{e}^1) = \sin 28.43 = -c_2 \qquad s^2 = \sin(\angle \mathbf{e}^2) = \sin 90 = c_1$$

The lengths of the basis vectors are in nanometers. We define $\sin(\angle \mathbf{e}_2 - \angle \mathbf{e}_1) = s_{21} = \sin 118.43$

$$\ell_1 = \mid \mathbf{e}_1 \mid = 1.52 \qquad \ell_2 = \mid \mathbf{e}_2 \mid = 0.65$$

$$\ell^1 = \mid \mathbf{e}^1 \mid = 1/(\ell_1 s_{21}) = 0.75 \qquad \ell^2 = \mid \mathbf{e}^2 \mid = 1/(\ell_2 s_{21}) = 1.75$$

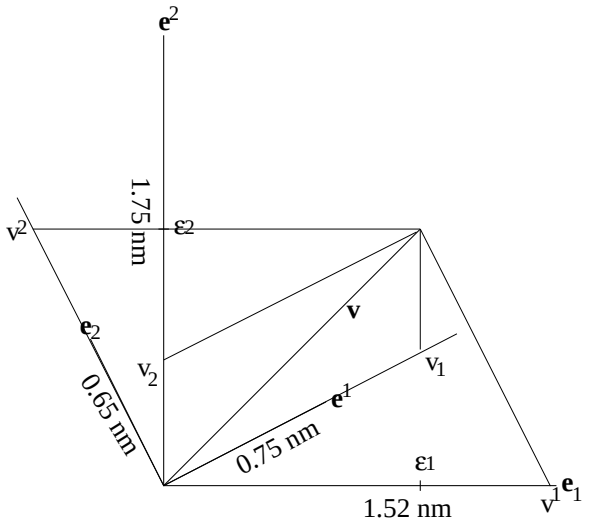
Here are the (gypsum) basis vectors in terms of the orthonormal basis.

$$\boldsymbol{e}_1 = \ell_1(c_1\boldsymbol{\epsilon}_1 + c_2\boldsymbol{\epsilon}_2) \qquad\qquad \boldsymbol{e}_2 = \ell_2(s_1\boldsymbol{\epsilon}_1 + s_2\boldsymbol{\epsilon}_2)$$

$$\begin{aligned}(\boldsymbol{e}_1 \;\; \boldsymbol{e}_2) &= (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2)\begin{pmatrix} \ell_1 c_1 & \ell_2 c_2 \\ \ell_1 s_1 & \ell_2 s_2 \end{pmatrix} \\ &= (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2)\begin{pmatrix} c_1 & c_2 \\ s_1 & s_2 \end{pmatrix}\begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}\end{aligned}$$

And similarly for the complementary basis,

$$(\boldsymbol{e}^1 \;\; \boldsymbol{e}^2) = (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2)\begin{pmatrix} c^1 & c^2 \\ s^1 & s^2 \end{pmatrix}\begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix}$$



Now we show an arbitrary vector $\boldsymbol{v}$. (To make the example concrete, the figure above shows $(v_1, v_2) = (1, 1)$ with respect to the orthonormal basis.) The *contravariant* coefficients v^1, v^2 and the *covariant* coefficients v_1, v_2 are *defined* below. Note the orthogonality of $\boldsymbol{e}^2$ and $\boldsymbol{e}_1$, and of $\boldsymbol{e}^1$ and $\boldsymbol{e}_2$, and how the coefficients are found by dropping lines from the end of $\boldsymbol{v}$ parallel to the *other* axis.

(We'll use this terminology from now on, so here is a useful mnemonic. Note the lengths of (numbers of letters in) the prefixes.

co:sub	covariant components are subscripted
contra:super	contravariant components are superscripted

Remember that components for an orthonormal basis are "midscripted".)

Contravariant

$$\begin{aligned}
\boldsymbol{v} &= v_1\boldsymbol{\epsilon}_1 + v_2\boldsymbol{\epsilon}_2 = (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= v^1\boldsymbol{e}_1 + v^2\boldsymbol{e}_2 \\
&= v^1\ell_1(c_1\boldsymbol{\epsilon}_1 + s_1\boldsymbol{\epsilon}_2) + v^2\ell_2(c_2\boldsymbol{\epsilon}_1 + s_2\boldsymbol{\epsilon}_2) \\
&= (v^1\ell_1c_1 + v^2\ell_2c_2)\boldsymbol{\epsilon}_1 + (v^1\ell_1s_1 + v^2\ell_2s_2)\boldsymbol{\epsilon}_2 \\
&= (\ell_1c_1 \;\; \ell_2c_2)\begin{pmatrix} v^1 \\ v^2 \end{pmatrix}\boldsymbol{\epsilon}_1 + (\ell_1s_1 \;\; \ell_2s_2)\begin{pmatrix} v^1 \\ v^2 \end{pmatrix}\boldsymbol{\epsilon}_2 \\
&= (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2)\begin{pmatrix} \ell_1c_1 & \ell_2c_2 \\ \ell_1s_1 & \ell_2s_2 \end{pmatrix}\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \\
&= (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2)\begin{pmatrix} c_1 & c_2 \\ s_1 & s_2 \end{pmatrix}\begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}\begin{pmatrix} v^1 \\ v^2 \end{pmatrix}
\end{aligned}$$

Covariant

$$\begin{aligned}
\boldsymbol{v} &= v_1\boldsymbol{\epsilon}_1 + v_2\boldsymbol{\epsilon}_2 = (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= v_1\boldsymbol{e}^1 + v_2\boldsymbol{e}^2 \\
&= v_1\ell^1(s_2\boldsymbol{\epsilon}_1 - c_2\boldsymbol{\epsilon}_2) + v_2\ell^2(-s_1\boldsymbol{\epsilon}_1 + c_1\boldsymbol{\epsilon}_2) \\
&= (v_1\ell^1s_2 - v_2\ell^2s_1)\boldsymbol{\epsilon}_1 + (v_1(-\ell^1c_2) + v_2\ell^2c_1)\boldsymbol{\epsilon}_2 \\
&= (\ell^1s_2 \;\; -\ell^2s_1)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\boldsymbol{\epsilon}_1 + (-\ell^1c_2 \;\; \ell^2c_1)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\boldsymbol{\epsilon}_2 \\
&= (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2)\begin{pmatrix} \ell^1s_2 & -\ell^2s_1 \\ -\ell^1c_2 & \ell^2c_1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2)\begin{pmatrix} s_2 & -s_1 \\ -c_2 & c_1 \end{pmatrix}\begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2)\begin{pmatrix} s_2 & -s_1 \\ -c_2 & c_1 \end{pmatrix}\begin{pmatrix} \frac{1}{\ell_1s_{21}} & \\ & \frac{1}{\ell_2s_{21}} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}
\end{aligned}$$

So we have the two relationships

$$\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = \left(\begin{pmatrix} c_1 & c_2 \\ s_1 & s_2 \end{pmatrix}\begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}\right)^{-1}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \qquad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \left(\begin{pmatrix} c^1 & c^2 \\ s^1 & s^2 \end{pmatrix}\begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix}\right)^{-1}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

The definition of the dual basis, from which we get the relationships linking c^j, s^j and ℓ^j to c_k, s_k and ℓ_k as

$$\begin{aligned}
\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} &= \begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{pmatrix}(\boldsymbol{e}^1 \;\; \boldsymbol{e}^2) \\
&= \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}\begin{pmatrix} c_1 & s_1 \\ c_2 & s_2 \end{pmatrix}\begin{pmatrix} \boldsymbol{\epsilon}_1 \\ \boldsymbol{\epsilon}_2 \end{pmatrix}(\boldsymbol{e}_1 \;\; \boldsymbol{e}_2)\begin{pmatrix} c^1 & c^2 \\ s^1 & s^2 \end{pmatrix}\begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix} \\
&= \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}\begin{pmatrix} c_1 & s_1 \\ c_2 & s_2 \end{pmatrix}\begin{pmatrix} c^1 & c^2 \\ s^1 & s^2 \end{pmatrix}\begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix} \\
&= \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}\begin{pmatrix} c_1 & s_1 \\ c_2 & s_2 \end{pmatrix}\begin{pmatrix} s_2 & -s_1 \\ -c_2 & c_1 \end{pmatrix}\begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix} \\
&= \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}\begin{pmatrix} s_{21} & \\ & s_{21} \end{pmatrix}\begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix} \\
&= \begin{pmatrix} \ell_1s_{21}\ell^1 & \\ & \ell_2s_{21}\ell^2 \end{pmatrix}
\end{aligned}$$

where the fourth line is due to the perpendicularities $\boldsymbol{e}^1 \perp \boldsymbol{e}_2$ and $\boldsymbol{e}^2 \perp \boldsymbol{e}_1$:

$$\begin{pmatrix} c^1 \\ s^1 \end{pmatrix} = \begin{pmatrix} s_2 \\ -c_2 \end{pmatrix} \qquad \begin{pmatrix} c^2 \\ s^2 \end{pmatrix} = \begin{pmatrix} -s_1 \\ c_1 \end{pmatrix}$$

and the values we gave at the beginning of this Note

$$\ell^1 = \frac{1}{\ell_1s_{21}} \qquad \ell^2 = \frac{1}{\ell_2s_{21}}$$

follow from the last line.

Similarly we can show that the dot product of $(v_1 \;\; v_2)$ with its superscripted counterpart gives the

length of an arbitrary vector $\boldsymbol{v}$.

$$\begin{aligned}
(v_1 \;\; v_2)\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} &= (v_1 \;\; v_2)\left(\begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix}\begin{pmatrix} c^1 & s^1 \\ c^2 & s^2 \end{pmatrix}\right)^{-1}\left(\begin{pmatrix} c_1 & c_2 \\ c_2 & s_2 \end{pmatrix}\begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}\right)^{-1}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= (v_1 \;\; v_2)\begin{pmatrix} s_2 & -c_2 \\ -s_1 & c_1 \end{pmatrix}^{-1} s_{21}\begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}\begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}^{-1}\begin{pmatrix} c_1 & c_2 \\ s_1 & s_2 \end{pmatrix}^{-1}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= (v_1 \;\; v_2)\begin{pmatrix} c_1 & c_2 \\ s_1 & s_2 \end{pmatrix}\frac{1}{s_{21}}s_{21}\begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}\begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix}^{-1}\begin{pmatrix} s_2 & -c_2 \\ -s_1 & c_1 \end{pmatrix}\frac{1}{s_{21}}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= (v_1 \;\; v_2)\begin{pmatrix} s_{21} & \\ & s_{21} \end{pmatrix}\frac{1}{s_{21}}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= (v_1 \;\; v_2)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}
\end{aligned}$$

So covariant and contravariant coordinates are complementary in the important sense that they must be used together to calculate lengths or distances and even angles between two different vectors.

$$(u_1 \;\; u_2)\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = (u^1 \;\; u^2)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \mid u \mid\mid v \mid \cos\theta$$

where θ is the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$, and

$$\mid u \mid^2 = (u_1 \;\; u_2)\begin{pmatrix} u^1 \\ u^2 \end{pmatrix} = (u^1 \;\; u^2)\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

and similarly for $\mid v \mid^2$ the square of the length of $\boldsymbol{v}$.

When we must use non-orthonormal basis vectors we'll find we need both contra- and co-variant coordinates.

Index notation. We can do all this more generally using index notation.

$$\boldsymbol{e}_j = \boldsymbol{\epsilon}_m X_{mj} = \boldsymbol{\epsilon}_m \ell_j t_{mj}$$

where the trigonometric terms $t_{1j} = c_j$ and $t_{2j} = s_j$;

$$\boldsymbol{e}^k = \boldsymbol{\epsilon}_n X_n{}^k = \boldsymbol{\epsilon}_n \ell^k t_n{}^k$$

with the trigonometric terms $t_1{}^k = c^k$ and $t_2{}^k = s^k$.

Find the superscripted X in terms of the subscripted X.

$$\begin{aligned}
\delta^k_j &= \boldsymbol{e}_j \boldsymbol{e}^k = \boldsymbol{\epsilon}_m X_{mj} \boldsymbol{\epsilon}_n X_n{}^k \\
&= X_{mj} X_n{}^k \boldsymbol{\epsilon}_m \boldsymbol{\epsilon}_n \\
&= X_{mj} X_n{}^k \delta_{mn} \\
&= X_{mj} X_m{}^k
\end{aligned}$$

So the superscripted X is the inverse of the subscripted X. To work this out in detail we must write the elements out in array form, making the trigonometric terms explicit again.

X_{mj}

m	1	2
$j=1$	$\ell_1 c_1$	$\ell_1 s_1$
2	$\ell_2 c_2$	$\ell_2 s_2$

$X_n{}^k$

k	1	2
$n=1$	$\ell^1 c^1$	$\ell^2 c^2$
2	$\ell^1 s^1$	$\ell^2 s^2$

$=$ $(X_{mj})^{-1}{}_n{}^k$

k	1	2
$n=1$	$\ell_2 s_2/\ell_1\ell_2 s_{21}$	$-\ell_1 s_1/\ell_1\ell_2 s_{21}$
2	$-\ell_2 c_2/\ell_1\ell_2 s_{21}$	$\ell_1 c_1/\ell_1\ell_2 s_{21}$

where $s_{21} = c_1 s_2 - c_2 s_1$. (These tables can be thought of as relations formatted so as to make it easy to see the inverse in the 2-by-2 case. Each could also be written as a three-column, four-row representation of the relation.)

There are multiple solutions to this, even with the restriction that $(c^k)^2 + (s^k)^2 = 1$. We choose the solution that gives positive lengths in the gypsum example.

$$\begin{aligned} \ell^1 &= \frac{1}{\ell_1 s_{21}} \\ \ell^2 &= \frac{1}{\ell_2 s_{21}} \end{aligned} \qquad \begin{aligned} c^1 &= s_2 \\ s^1 &= -c_2 \\ c^2 &= -s_1 \\ s^2 &= c_1 \end{aligned}$$

We can now show that the components of an arbitrary vector give its length as the dot product of subscripted with superscripted components.

$$\begin{aligned} v_m \boldsymbol{\epsilon}_m = \boldsymbol{v} &= v^j \boldsymbol{e}_j = v^j \boldsymbol{\epsilon}_m X_{mj} \\ &= v_j \boldsymbol{e}^j = v_j \boldsymbol{\epsilon}_m X_m{}^j \end{aligned}$$

$$\begin{aligned} v_m \boldsymbol{\epsilon}_m &= v^j X_{mj} \boldsymbol{\epsilon}_m \\ v_m &= v^j X_{mj} \\ v_m X_m{}^k &= v^j X_{mj} X_m{}^k = v^k \end{aligned} \qquad \begin{aligned} v_m \boldsymbol{\epsilon}_m &= v_j X_m{}^j \boldsymbol{\epsilon}_m \\ v_m &= v_j X_m{}^j \\ v_m X_{mk} &= v_j X_m{}^j X_{mk} = v_k \end{aligned}$$

The second line on each side above follows from the first because all the components of $\boldsymbol{\epsilon}_m$ are independent so for the sums to be equal, each of the corresponding coefficients must be equal.

The third lines follow because $X_m{}^j$ and X_{mk} are inverses, as we found from the orthonormality of the sub- and superscripted basis vectors to each other.

Thus

$$v^k v_k = v_m X_m{}^k X_{nk} v_n = v_m v_m$$

again because the two variants of X are inverses of each other.

The final right-hand side is the length-squared of $\boldsymbol{v}$ expressed in the orthonormal coordinate system. So the left-hand side, $v^k v_k$, equals this length-squared.

Generally, the dot product of two vectors is written as the sum of the products of corresponding sub- and superscripted coordinates.

(For triclinic crystals, such as turquoise, angles β and γ are also not right angles, and we would need to do all the above in three dimensions.)

We have started in this Note to use the same symbol to mean different, indeed opposite, things, and so the relational version of these protors will need an improved notation.

We had

$$X_{mj} X_m{}^k = \delta_j^k$$

so that X is its own inverse, depending on the position of the sub/superscript.

Representations which do not show sub/superscripts, such as matrix or relational representations, must be augmented to be clear.

So we'll say explicitly whether each index is mid, up (super) or down (sub) as follows

$$X_{\text{mu}} = (X_{\text{md}})^{-1T}$$

for matrices, or

$$X_{\text{md}}(m, j, X) \qquad \text{and} \qquad X_{\text{mu}}(m, k, X)$$

for relations.

We will call this "mud" notation for the three possible index positions.

Note that the indicators m, u and d are themselves typeset as midscripts, because midscript indices will not be used much from now on: the notation is effectively "ud" notation. (Note that we can "transpose", say X_{mu} as X_{um}, and so swap the index columns, but this makes no effective difference to any relational calculations.)

9. The metric. We can find a direct relationship between contravariant and covariant components by eliminating the orthonormal (midscripted) basis from the two versions of $\boldsymbol{v}$ in the previous Note.

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ s_1 & s_2 \end{pmatrix} \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \qquad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} c^1 & c^2 \\ s^1 & s^2 \end{pmatrix} \begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

So if we know only, say, the contravariant (superscripted) components, we can write the length in terms of them only.

$$\begin{aligned}
\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= \begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix}^{-1} \begin{pmatrix} c^1 & c^2 \\ s^1 & s^2 \end{pmatrix}^{-1} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= \begin{pmatrix} \ell^1 & \\ & \ell^2 \end{pmatrix}^{-1} \begin{pmatrix} c^1 & c^2 \\ s^1 & s^2 \end{pmatrix}^{-1} \begin{pmatrix} c_1 & c_2 \\ s_1 & s_2 \end{pmatrix} \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \\
&= \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix} s_{21} \begin{pmatrix} s_2 & -s_1 \\ -c_2 & c_1 \end{pmatrix}^{-1} \begin{pmatrix} c_1 & c_2 \\ s_1 & s_2 \end{pmatrix} \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \\
&= \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix} \frac{s_{21}}{s_{21}} \begin{pmatrix} c_1 & s_1 \\ c_2 & s_1 \end{pmatrix} \begin{pmatrix} c_1 & c_2 \\ s_1 & s_2 \end{pmatrix} \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \\
&= \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix} \begin{pmatrix} 1 & c_{21} \\ c_{21} & 1 \end{pmatrix} \begin{pmatrix} \ell_1 & \\ & \ell_2 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \\
&= \begin{pmatrix} \ell_1^2 & \ell_1\ell_2 c_{21} \\ \ell_1\ell_2 c_{21} & \ell_2^2 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix}
\end{aligned}$$

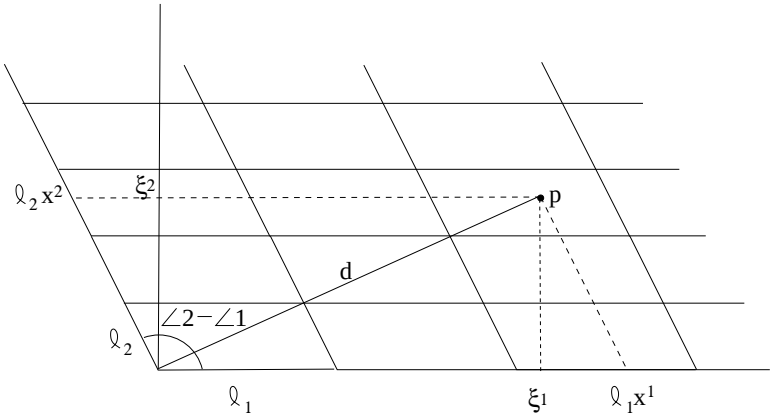
where $c_{21} = c_2c_1 + s_2s_1$ is $\cos(\angle\mathbf{e}_2 - \angle\mathbf{e}_1) = c_{21}$.

This matrix

$$\begin{pmatrix} \ell_1^2 & \ell_1\ell_2 c_{21} \\ \ell_1\ell_2 c_{21} & \ell_2^2 \end{pmatrix}$$

is called the *metric* g because it enables us to calculate lengths from a given set of coordinate components. Note that g is a symmetrical matrix. Check that it is a diagonal matrix if the coordinate axes are orthogonal to each other, and the identity matrix if they are orthonormal.

It is called the metric because it turns coordinates into lengths or distances. (This reinforces the notion, which is a fact, that unnormalized coordinate systems do not directly tell us about distances between points.) It can be found directly from distances. Here is a point p a distance d from the origin.



(All expressions in this diagram, except p, represent distances.)

In terms of the orthonormal coordinates

$$d^2 = {\xi_1}^2 + {\xi_2}^2$$

In terms of the gypsum coordinates x^1 and x^2

$$\begin{aligned} \xi_1 &= \ell_1 x^1 + \ell_2 c_{21} x^2 \\ \xi_2 &= \ell_2 s_{21} x^2 \\ d^2 &= \ell_1^2 (x^1)^2 + \ell_2^2 (x^2)^2 + 2\ell_1 \ell_2 c_{21} x^1 x^2 \end{aligned}$$

where $c_{21} = \cos(\angle 2 - \angle 1)$ and $s_{21} = \sin(\angle 2 - \angle 1)$, or, in matrix form,

$$d^2 = (x^1 \;\; x^2) \begin{pmatrix} \ell_1^2 & \ell_1 \ell_2 c_{21} \\ \ell_1 \ell_2 c_{21} & \ell_2^2 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}$$

So it is easy to read the metric off from the coordinate axes.

A perhaps even easier way of finding the metric from the coordinate axes is

$$g_{\mathrm{dd}} = \begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{pmatrix} \cdot (\boldsymbol{e}_1 \;\; \boldsymbol{e}_2) = \begin{pmatrix} \boldsymbol{e}_1 . \boldsymbol{e}_1 & \boldsymbol{e}_1 . \boldsymbol{e}_2 \\ \boldsymbol{e}_2 . \boldsymbol{e}_1 & \boldsymbol{e}_2 . \boldsymbol{e}_2 \end{pmatrix}$$

For gypsum this is also

$$g_{\mathrm{dd}} = \begin{pmatrix} \ell_1^2 & \ell_1 \ell_2 c_{21} \\ \ell_1 \ell_2 c_{21} & \ell_2^2 \end{pmatrix}$$

From this, g can be decomposed into the product of the matrix that transforms from orthonormal coordinates with the transpose of that matrix.

$$\begin{aligned} g_{\mathrm{dd}} &= \begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{pmatrix} \cdot (\boldsymbol{e}_1 \;\; \boldsymbol{e}_2) \\ &= \begin{pmatrix} \ell_c c_1 & \ell_1 s_1 \\ \ell_2 c_2 & \ell_2 s_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_1 \\ \boldsymbol{\epsilon}_2 \end{pmatrix} (\boldsymbol{\epsilon}_1 \;\; \boldsymbol{\epsilon}_2) \begin{pmatrix} \ell_c c_1 & \ell_1 c_2 \\ \ell_2 s_1 & \ell_2 s_2 \end{pmatrix} \\ &= \begin{pmatrix} \ell_c c_1 & \ell_1 s_1 \\ \ell_2 c_2 & \ell_2 s_2 \end{pmatrix} \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \begin{pmatrix} \ell_c c_1 & \ell_1 c_2 \\ \ell_2 s_1 & \ell_2 s_2 \end{pmatrix} \\ &= \begin{pmatrix} \ell_c c_1 & \ell_1 s_1 \\ \ell_2 c_2 & \ell_2 s_2 \end{pmatrix} \begin{pmatrix} \ell_c c_1 & \ell_1 c_2 \\ \ell_2 s_1 & \ell_2 s_2 \end{pmatrix} \end{aligned}$$

or, generally (and we see the transpose only on the apparently swapped positions of m and k)

$$\begin{aligned} g_{jk} &= \boldsymbol{e}_j.\boldsymbol{e}_k \\ &= X_{mj}\boldsymbol{\epsilon}_j.\boldsymbol{\epsilon}_n X_{nk} \\ &= X_{mj}\delta_{mn}X_{nk} \\ &= X_{mj}X_{mk} \end{aligned}$$

Similarly g^{-1} gives lengths if we must use the covariant components only.

$$\begin{aligned} (v_1 \;\; v_2)\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} &= (v^1 \;\; v^2)\begin{pmatrix} \ell_1^2 & \ell_1\ell_2 c_{21} \\ \ell_1\ell_2 c_{21} & \ell_2^2 \end{pmatrix}\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \\ &= (v_1 \;\; v_2)\frac{1}{\ell_1^2\ell_2^2(1-c_{21})}\begin{pmatrix} \ell_2^2 & -\ell_1\ell_2 c_{21} \\ -\ell_1\ell_2 c_{21} & \ell_1^2 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \end{aligned}$$

Thus, in mud notation, $g_{\text{uu}} = {g_{\text{dd}}}^{-1}$.

10. Fields in gypsum. We can also transform fields such as the radial field (in orthonormal coordinates)

$$(f_1 \;\; f_2) = r^{p-1}(x_1 \;\; x_2)$$

and the tangential field

$$(h_1 \;\; h_2) = r^{p-1}(-x_2 \;\; x_1)$$

to gypsum coordinates.

First we review the results for fields in orthogonal coordinates.

$$\begin{aligned} r &= \sqrt{{x_1}^2 + {x_2}^2} \\ \begin{pmatrix} \partial_1 \\ \partial_2 \end{pmatrix} r &= \frac{1}{r}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{aligned}$$

So, for f

$$\begin{aligned} \begin{pmatrix} \partial_1 \\ \partial_2 \end{pmatrix}(f_1 \;\; f_2) &= \begin{pmatrix} \partial_1 \\ \partial_2 \end{pmatrix} r^{p-1}(x_1 \;\; x_2) \\ &= r^{p-1}\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + (p-1)r^{p-3}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}(x_1 \;\; x_2) \end{aligned}$$

and

$$\begin{aligned} \mathbf{div}.\boldsymbol{f} &= (p+1)r^{p-1} \\ \mathbf{curl} \times \boldsymbol{f} &= 0 \end{aligned}$$

And for h

$$\begin{aligned} \begin{pmatrix} \partial_1 \\ \partial_2 \end{pmatrix}(h_1 \;\; h_2) &= \begin{pmatrix} \partial_1 \\ \partial_2 \end{pmatrix} r^{p-1}(-x_2 \;\; x_1) \\ &= r^{p-1}\begin{pmatrix} & 1 \\ -1 & \end{pmatrix} + (p-1)r^{p-3}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}(x_1 \;\; x_2) \end{aligned}$$

and

$$\begin{aligned}\mathbf{div}.\boldsymbol{h} &= 0\\ \mathbf{curl}\times\boldsymbol{h} &= (p+1)r^{p-1}\end{aligned}$$

Now we transform to gypsum coordinates. We'll be interested in the contravariant $\boldsymbol{x}$.

$$(x_1\ \ x_2)\begin{pmatrix}\boldsymbol{\epsilon}_1\\ \boldsymbol{\epsilon}_2\end{pmatrix} = (x^1\ \ x^2)\begin{pmatrix}\boldsymbol{e}_1\\ \boldsymbol{e}_2\end{pmatrix} = (x^1\ \ x^2)\begin{pmatrix}\ell_1c_1 & \ell_1s_1\\ \ell_2c_2 & \ell_2s_2\end{pmatrix}\begin{pmatrix}\boldsymbol{\epsilon}_1\\ \boldsymbol{\epsilon}_2\end{pmatrix}$$

and in both contravariant and covariant $\boldsymbol{f}$ and $\boldsymbol{h}$

$$(f^1\ \ f^2)\begin{pmatrix}\ell_1c_1 & \ell_1s_1\\ \ell_2c_2 & \ell_2s_2\end{pmatrix} = (f_1\ \ f_2) = (f_1\ \ f_2)\begin{pmatrix}\ell_1c_1 & \ell_1c_2\\ \ell_2s_1 & \ell_2s_2\end{pmatrix}^{-1}$$

(note the transposed inverse), and similarly for $\boldsymbol{h}$.

The divergence is a scalar and so must combine covariant and contravariant, taking the covariant slope of the contravariant coordinates.

$$\mathbf{div}.\boldsymbol{f} = \partial_j f^j \qquad\qquad \mathbf{div}.\boldsymbol{h} = \partial_j h^j$$

where we note that the covariant ∂_j means slope_{x^j}, the slope with respect to the contravariant coordinates.

The curl is generally a two-index tensor. (Since it is antisymmetric, it has only one independent component in two dimensions, so we've been getting away with writing it as a scalar.) We don't want a mixed tensor (one index up, the other down) so of the four possibilities, $\mathbf{curl}_{jk}$, $\mathbf{curl}^k_j$, $\mathbf{curl}^j_k$ and $\mathbf{curl}^{jk}$ we'll choose the first—both indices covariant—because it will be the easiest to use in future. Thus

$$\mathbf{curl}\times\boldsymbol{f} = \partial_1 f_2 - \partial_2 f_1 \qquad\qquad \mathbf{curl}\times\boldsymbol{h} = \partial_1 h_2 - \partial_2 h_1$$

Here are the preliminaries for $\boldsymbol{x}$

$$(x_1\ \ x_2) = (x^1\ \ x^2)\begin{pmatrix}\ell_1c_1 & \ell_1s_1\\ \ell_2c_2 & \ell_2s_2\end{pmatrix}$$

so

$$\begin{aligned}r &= \sqrt{{x_1}^2 + {x_2}^2}\\ &= \sqrt{(\ell_1c_1x^1 + \ell_2c_2x^2)^2 + (\ell_1s_1x^1 + \ell_2s_2x^2)^2}\\ &= \sqrt{(\ell_1x^1)^2 + 2\ell_1\ell_2c_{21}x^1x^2 + (\ell_2x^2)^2}\end{aligned}$$

where, again, $c_{21} = c_1c_2 + s_1s_2$ is the cosine of the angle between the axes $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. (And we'll soon need the counterpart, $s_{21} = c_1s_2 - c_2s_1$, the sine of that angle.)

We will need

$$\begin{pmatrix}\partial_1\\ \partial_2\end{pmatrix} r = \frac{1}{r}\begin{pmatrix}\ell_1^2 & \ell_1\ell_2c_{21}\\ \ell_1\ell_2c_{21} & \ell_2^2\end{pmatrix}\begin{pmatrix}x^1\\ x^2\end{pmatrix}$$

Here are the divergences of $\boldsymbol{f}$ and $\boldsymbol{h}$ in gypsum coordinates.

$$(f^1\ \ f^2) = (f_1\ \ f_2)\begin{pmatrix}\ell_1c_1 & \ell_1s_1\\ \ell_2c_2 & \ell_2s_2\end{pmatrix}^{-1}$$

$$
\begin{aligned}
&= r^{p-1}(x^1 \;\; x^2)\begin{pmatrix} \ell_1 c_1 & \ell_1 s_1 \\ \ell_2 c_2 & \ell_2 s_2 \end{pmatrix}\begin{pmatrix} \ell_1 c_1 & \ell_1 s_1 \\ \ell_2 c_2 & \ell_2 s_2 \end{pmatrix}^{-1} \\
&= r^{p-1}(x^1 \;\; x^2)\begin{pmatrix} \partial_1 \\ \partial_2 \end{pmatrix}(f^1 \;\; f^2) \\
&= r^{p-1}\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + (p-1)r^{p-3}\begin{pmatrix} \ell_1^2 & \ell_1\ell_2 c_{21} \\ \ell_1\ell_2 c_{21} & \ell_2^2 \end{pmatrix}\begin{pmatrix} x^1 \\ x^2 \end{pmatrix}(x^1 \;\; x^2)
\end{aligned}
$$

from which, summing the diagonals and using the above expression for r,

$$\mathbf{div}.\boldsymbol{f} = \partial_j f^j = (p+1)r^{p-1}$$

as before: divergence is an invariant scalar.

Next, h

$$
\begin{aligned}
(h^1 \;\; h^2) &= (h_1 \;\; h_2)\begin{pmatrix} \ell_1 c_1 & \ell_1 s_1 \\ \ell_2 c_2 & \ell_2 s_2 \end{pmatrix}^{-1} \\
&= r^{p-1}(-x_2 \;\; x_1)\begin{pmatrix} \ell_1 c_1 & \ell_1 s_1 \\ \ell_2 c_2 & \ell_2 s_2 \end{pmatrix}^{-1} \\
&= r^{p-1}(x^1 \;\; x^2)\begin{pmatrix} -\ell_1 s_1 & \ell_1 c_1 \\ -\ell_2 s_2 & \ell_2 c_2 \end{pmatrix}\begin{pmatrix} \ell_1 c_1 & \ell_1 s_1 \\ \ell_2 c_2 & \ell_2 s_2 \end{pmatrix}^{-1} \\
&= r^{p-1}(x^1 \;\; x^2)\begin{pmatrix} -\ell_1 s_1 & \ell_1 c_1 \\ -\ell_2 s_2 & \ell_2 c_2 \end{pmatrix}\begin{pmatrix} \ell_2 s_2 & -\ell_1 s_1 \\ -\ell_2 s_2 & \ell_2 c_2 \end{pmatrix}\frac{1}{\ell_1\ell_2 s_{21}} \\
&= r^{p-1}(x^1 \;\; x^2)\begin{pmatrix} -\ell_1\ell_2 c_{21} & \ell_1^2 1 \\ -\ell_2^2 & \ell_1\ell_2 c_{21} \end{pmatrix}\frac{1}{\ell_1\ell_2 s_{21}} \\
\begin{pmatrix} \partial_1 \\ \partial_2 \end{pmatrix}(h^1 \;\; h^2) &= r^{p-1}\begin{pmatrix} -\ell_1\ell_2 c_{21} & \ell_1^2 1 \\ -\ell_2^2 & \ell_1\ell_2 c_{21} \end{pmatrix}\frac{1}{\ell_1\ell_2 s_{21}} \\
&\quad +(p-1)r^{p-3}\begin{pmatrix} \ell_1^2 & \ell_1\ell_2 c_{21} \\ \ell_1\ell_2 c_{21} & \ell_2^2 \end{pmatrix}\begin{pmatrix} x^1 \\ x^2 \end{pmatrix}(x^1 \;\; x^2)\begin{pmatrix} \ell_1^2 & \ell_1\ell_2 c_{21} \\ \ell_1\ell_2 c_{21} & \ell_2^2 \end{pmatrix}\frac{1}{\ell_1\ell_2 s_{21}} \\
&= r^{p-1}\begin{pmatrix} -\ell_1\ell_2 c_{21} & \ell_1^2 1 \\ -\ell_2^2 & \ell_1\ell_2 c_{21} \end{pmatrix}\frac{1}{\ell_1\ell_2 s_{21}} \\
&\quad +(p-1)r^{p-3}\begin{pmatrix} \ell_1^2 x^1 + \ell_1\ell_2 c_{21} x^2 \\ \ell_1\ell_2 c_{21} x^1 + \ell_2^2 x^2 \end{pmatrix}(-\ell_1\ell_2 c_{21} x^1 - \ell_2^2 x^2 \quad \ell_1^2 x^1 + \ell_1\ell_2 c_{21} x^2)
\end{aligned}
$$

and, summing the diagonals

$$\mathbf{div}.\boldsymbol{h} = 0$$

These results on divergence can be captured generally, using

$$
\begin{aligned}
f_m &= X_{mj} f^j \\
&= X_m{}^k f_k \\
X_{mj} X_m{}^k &= \delta_j^k \\
X_{mj} X_n{}^j &= \delta_{mn} \\
X_{mj} X_{mk} &= g_{jk} \\
X_m{}^j X_m{}^k &= g^{jk}
\end{aligned}
\qquad\qquad
\begin{aligned}
X_{mj} &= \partial_j x_m \\
X_m{}^k &= \partial^k x_m \\
X_m{}^k &= \partial_m x^k
\end{aligned}
$$

$$
\begin{aligned}
\partial_j f^j &= X_{mj}\partial_m(X_n{}^j f_n) \\
&= X_{mj} X_n{}^j \partial_m f_n \\
&= \delta_{mn}\partial_m f_n \\
&= \partial_m f_m
\end{aligned}
$$

where we've used the inverse of $X_{^nj}$ in the first line and in the second line, and the assumption that the transformations X do not depend on the coordinates.

Here are the curls of $\boldsymbol{f}$ and $\boldsymbol{h}$ in gypsum coordinates.

$$\begin{aligned}
(f_1 \;\; f_2) &= (f_{\scriptscriptstyle 1} \;\; f_{\scriptscriptstyle 2}) \begin{pmatrix} \ell_1 c_1 & \ell_2 c_2 \\ \ell_1 s_1 & \ell_2 s_2 \end{pmatrix} \\
&= r^{p-1}(x^1 \;\; x^2) \begin{pmatrix} \ell_1 c_1 & \ell_1 s_1 \\ \ell_2 c_2 & \ell_2 s_2 \end{pmatrix} \begin{pmatrix} \ell_1 c_1 & \ell_2 c_2 \\ \ell_1 s_1 & \ell_2 s_2 \end{pmatrix} \\
&= r^{p-1}(x^1 \;\; x^2) \begin{pmatrix} \ell_1^2 & \ell_1 \ell_2 c_{21} \\ \ell_1 \ell_2 c_{21} & \ell_2^2 \end{pmatrix} \\
\begin{pmatrix} \partial_1 & \partial_2 \end{pmatrix} (f_1 \;\; f_2) &= r^{p-1} \begin{pmatrix} \ell_1^2 & \ell_1 \ell_2 c_{21} \\ \ell_1 \ell_2 c_{21} & \ell_2^2 \end{pmatrix} \\
&\quad + (p-1) r^{p-3} \begin{pmatrix} \ell_1^2 & \ell_1 \ell_2 c_{21} \\ \ell_1 \ell_2 c_{21} & \ell_2^2 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} (x^1 \;\; x^2) \begin{pmatrix} \ell_1^2 & \ell_1 \ell_2 c_{21} \\ \ell_1 \ell_2 c_{21} & \ell_2^2 \end{pmatrix}
\end{aligned}$$

and, since all the matrices are symmetric,

$$\mathbf{curl} \times \boldsymbol{f} = 0$$

Next, $\boldsymbol{h}$

$$\begin{aligned}
(h_1 \;\; h_2) &= (h_{\scriptscriptstyle 1} \;\; h_{\scriptscriptstyle 2}) \begin{pmatrix} \ell_1 c_1 & \ell_2 c_2 \\ \ell_1 s_1 & \ell_2 s_2 \end{pmatrix} \\
&= r^{p-1}(x^1 \;\; x^2) \begin{pmatrix} -\ell_1 s_1 & \ell_1 c_1 \\ -\ell_2 s_2 & \ell_2 c_2 \end{pmatrix} \begin{pmatrix} \ell_1 c_1 & \ell_2 c_2 \\ \ell_1 s_1 & \ell_2 s_2 \end{pmatrix} \\
&= r^{p-1}(x^1 \;\; x^2) \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} \ell_1 \ell_2 s_{21} \\
\begin{pmatrix} \partial_1 & \partial_2 \end{pmatrix} (f_1 \;\; f_2) &= r^{p-1} \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} \ell_1 \ell_2 s_{21} \\
&\quad + (p-1) r^{p-3} \begin{pmatrix} \ell_1^2 & \ell_1 \ell_2 c_{21} \\ \ell_1 \ell_2 c_{21} & \ell_2^2 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} (x^1 \;\; x^2) \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} \ell_1 \ell_2 s_{21} \\
&= r^{p-1} \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} \ell_1 \ell_2 s_{21} \\
&\quad + (p-1) r^{p-3} \begin{pmatrix} \ell_1^2 x^1 + \ell_1 \ell_2 c_{21} x^2 \\ \ell_1 \ell_2 c_{21} x^1 + \ell_2^2 x^2 \end{pmatrix} (-x^2 \;\; x^1) \ell_1 \ell_2 s_{21}
\end{aligned}$$

so that, subtracting the off-diagonals and using the contravariant r

$$\begin{aligned}
\mathbf{curl} \times \boldsymbol{h} &= (2r^{p-1} + (p-1) r^{p-3} r^2) \ell_1 \ell_2 s_{21} \\
&= (p+1) r^{p-1} \det X
\end{aligned}$$

where

$$\ell_1 \ell_2 s_{21} = \det \begin{pmatrix} \ell_1 c_1 & \ell_2 c_2 \\ \ell_1 s_1 & \ell_2 s_2 \end{pmatrix}$$

the determinant of the transformation or, alternatively the square root of

$$g = \det \begin{pmatrix} \ell_1^2 & \ell_1 \ell_2 c_{21} \\ \ell_1 \ell_2 c_{21} & \ell_2^2 \end{pmatrix}$$

11. Polar coordinates. A commonly-used alternative coordinate system is polar, with the distance from the origin r and the angle θ as the two coordinates.

$$\begin{aligned} r &= \sqrt{x^2+y^2} & \qquad x &= r\cos\theta \\ \theta &= \operatorname{atan}(y,x) & \qquad y &= r\sin\theta \end{aligned}$$

We introduce the idea of *local* coordinate systems in which an observer at point (r,θ) describes a vector $(\Delta r, \Delta\theta)$ in terms of how much r and θ must change from the observer's standpoint (r,θ) to the top of the vector at $(r+\Delta r, \theta+\Delta\theta)$. (That's in terms of coordinates. In terms of lengths we need $(r+\Delta r, r\theta + r\Delta\theta)$.)

In order to measure distances correctly, the local basis vectors must have length 1 in the r direction and have length r in the (orthogonal) θ direction, because the length s satisfies

$$s^2 = (\Delta r)^2 + (r\Delta\theta)^2$$

So we have

$$\begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} = \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix}$$

(Please remember that, since these are *vectors*, i.e., two-dimensional *intervals*, they are described independently of their locations, i.e., their starting points.)

This is an orthogonal but not normalized coordinate system with metric

$$g_{\mathrm{dd}} = \begin{pmatrix} \boldsymbol{e}_r.\boldsymbol{e}_r & \boldsymbol{e}_\theta.\boldsymbol{e}_r \\ \boldsymbol{e}_r.\boldsymbol{e}_\theta & \boldsymbol{e}_\theta.\boldsymbol{e}_\theta \end{pmatrix} = \begin{pmatrix} 1 & \\ & r^2 \end{pmatrix}$$

So the contravariant coordinates for a vector $\boldsymbol{v}$ in this basis are given by

$$\begin{aligned} (v_x \;\; v_y) &= \boldsymbol{v} = (v^r \;\; v^\theta) \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} \\ &= (v^r \;\; v^\theta) \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix} \end{aligned}$$

Because the polar coordinate basis is not orthonormal it must have a different dual basis with metric

$$g_{\mathrm{uu}} = {g_{\mathrm{dd}}}^{-1} = \begin{pmatrix} \boldsymbol{e}^r.\boldsymbol{e}^r & \boldsymbol{e}^\theta.\boldsymbol{e}^r \\ \boldsymbol{e}^r.\boldsymbol{e}^\theta & \boldsymbol{e}^\theta.\boldsymbol{e}^\theta \end{pmatrix} = \begin{pmatrix} 1 & \\ & 1/r^2 \end{pmatrix}$$

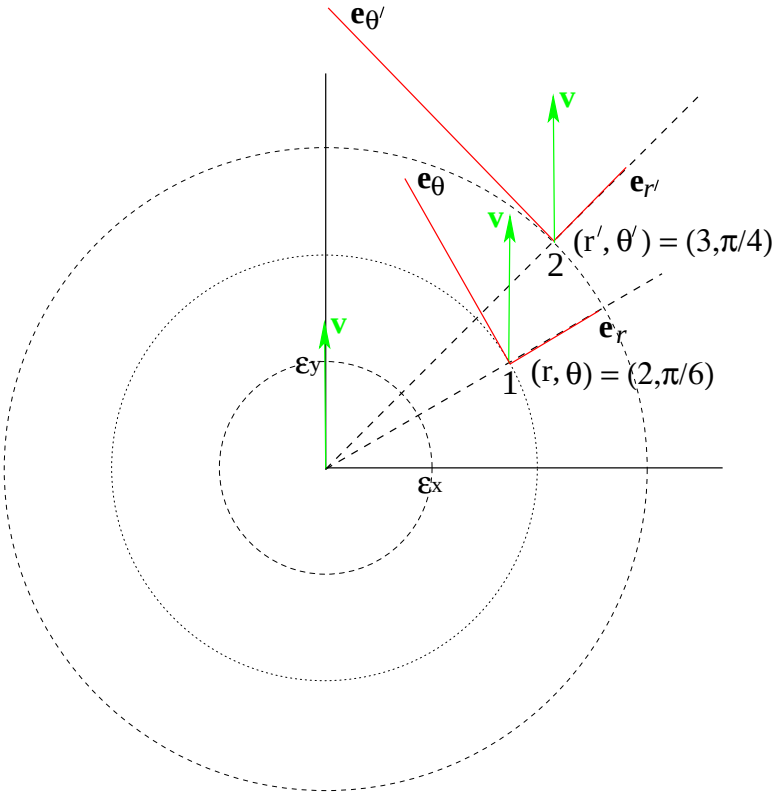
and with transformation matrix the transpose of the inverse of the $\boldsymbol{\epsilon}_j \to e_k$ transformation

$$\begin{pmatrix} \boldsymbol{e}^r \\ \boldsymbol{e}^\theta \end{pmatrix} = \begin{pmatrix} c & s \\ -s/r & c/r \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix}$$

And we can check

$$\begin{aligned} g_{\mathrm{uu}} = \begin{pmatrix} \boldsymbol{e}^r \\ \boldsymbol{e}^\theta \end{pmatrix} (\boldsymbol{e}^r \;\; \boldsymbol{e}^\theta) &= \begin{pmatrix} c & s \\ -s/r & c/r \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix} (\boldsymbol{\epsilon}_x \;\; \boldsymbol{\epsilon}_y) \begin{pmatrix} c & -s/r \\ s & c/r \end{pmatrix} \\ &= \begin{pmatrix} c & s \\ -s/r & c/r \end{pmatrix} \begin{pmatrix} c & -s/r \\ s & c/r \end{pmatrix} = \begin{pmatrix} 1 & \\ & 1/r^2 \end{pmatrix} \\ (\boldsymbol{e}_r \;\; \boldsymbol{e}_\theta) \begin{pmatrix} \boldsymbol{e}^r \\ \boldsymbol{e}^\theta \end{pmatrix} &= (\boldsymbol{\epsilon}_x \;\; \boldsymbol{\epsilon}_y) \begin{pmatrix} c & -rs \\ s & rc \end{pmatrix} \begin{pmatrix} c & s \\ -s/r & c/r \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix} \\ &= (\boldsymbol{\epsilon}_x \;\; \boldsymbol{\epsilon}_y) \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix} = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \end{aligned}$$

Here is a vector $\boldsymbol{v}$, shown in three different locations, two of them at two different local polar coordinate systems.



The components of the vector $\boldsymbol{v}$ in the three systems are

$$\begin{aligned}(v_x \;\; v_y) &= (0 \;\; \sqrt{2}) \\ (v^r \;\; v^\theta) &= (v_x \;\; v_y)\begin{pmatrix} c & -s/r \\ s & c/r \end{pmatrix} \\ (v_r \;\; v_\theta) &= (v_x \;\; v_y)\begin{pmatrix} c & -rs \\ s & rc \end{pmatrix}\end{aligned}$$

$$\begin{aligned}(r \;\; \theta) &= (2 \;\; \pi/6) \\ (c \;\; s) &= (\sqrt{3}/2 \;\; 1/2) \\ (v^r \;\; v^\theta) &= (1/\sqrt{2} \;\; \sqrt{3}/(2\sqrt{2})) \\ (v_r \;\; v_\theta) &= (1/\sqrt{2} \;\; \sqrt{6})\end{aligned} \qquad \begin{aligned}(r \;\; \theta) &= (3 \;\; \pi/4) \\ (c \;\; s) &= (1/\sqrt{2} \;\; 1/\sqrt{2}) \\ (v^r \;\; v^\theta) &= (1 \;\; 1/3) \\ (v_r \;\; v_\theta) &= (1 \;\; 3)\end{aligned}$$

and you can easily check the contravariant components $(1 \;\; 1/3)$ at $(r \;\; \theta) = (3 \;\; \pi/4)$ in the diagram. And note that the length of $\boldsymbol{v}$ is $\sqrt{2}$ in all systems.

$$2 = (0 \;\; \sqrt{2})\begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{\sqrt{3}}{2\sqrt{2}} \end{pmatrix}\begin{pmatrix} 1/\sqrt{2} \\ \sqrt{6} \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{3} \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

Now let's work with contravariant coordinates at the two locations, calling the first location 1 (at, say, $(2 \;\; \pi/6)$) and the second location 2 (at, say, $(3 \;\; \pi/4)$), and labelling the c and s accordingly.

$$\begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} c_1 & -r_1 s_1 \\ s_1 & r_1 c_1 \end{pmatrix}\begin{pmatrix} {v_1}^r \\ {v_1}^\theta \end{pmatrix} = \begin{pmatrix} c_1 & -s_1 \\ s_1 & c_1 \end{pmatrix}\begin{pmatrix} 1 & \\ & r_1 \end{pmatrix}\begin{pmatrix} {v_1}^r \\ {v_1}^\theta \end{pmatrix}$$

$$\begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} c_2 & -r_2 s_2 \\ s_2 & r_2 c_2 \end{pmatrix} \begin{pmatrix} {v_2}^r \\ {v_2}^\theta \end{pmatrix} = \begin{pmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{pmatrix} \begin{pmatrix} 1 & \\ & r_2 \end{pmatrix} \begin{pmatrix} {v_2}^r \\ {v_2}^\theta \end{pmatrix}$$

To go directly from location 1 to location 2 we must eliminate the orthonormal coordinates $(v_1 \;\; v_2)$.

$$\begin{aligned} \begin{pmatrix} {v_1}^r \\ {v_1}^\theta \end{pmatrix} &= \begin{pmatrix} 1 & \\ & 1/r_1 \end{pmatrix} \begin{pmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix} \\ &= \begin{pmatrix} 1 & \\ & 1/r_1 \end{pmatrix} \begin{pmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{pmatrix} \begin{pmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{pmatrix} \begin{pmatrix} 1 & \\ & r_2 \end{pmatrix} \begin{pmatrix} {v_2}^r \\ {v_2}^\theta \end{pmatrix} \\ &= \begin{pmatrix} 1 & \\ & 1/r_1 \end{pmatrix} \begin{pmatrix} c_{21} & -s_{21} \\ s_{21} & c_{21} \end{pmatrix} \begin{pmatrix} 1 & \\ & r_2 \end{pmatrix} \begin{pmatrix} {v_2}^r \\ {v_2}^\theta \end{pmatrix} \end{aligned}$$

where c_{21} and s_{21} are again, respectively, the cos and sin of the difference in angles $\theta_2 - \theta_1$.

This transformation between the coordinate system at point 1 and the coordinate system at point 2 introduces something new.

The transformation now depends on the coordinates (r_1, θ_1) and (r_2, θ_2). It will no longer be possible to commute the operation of taking slopes with the transformation.

$$\begin{aligned} \partial_r X(r,\theta) &\neq X(r,\theta)\partial_r \\ \partial_\theta X(r,\theta) &\neq X(r,\theta)\partial_\theta \end{aligned}$$

So we will have to adjust our idea of tensors (Note 3) to this new situation: a tensor will still transform identically on all indices, but the transformation will need to be more subtle, particularly where slopes are concerned.

We'll soon need the above transformation from $\boldsymbol{v}_2$ to $\boldsymbol{v}_1$ in index notation.

$$\begin{aligned} {v_1}^j &= {X_{12}}^j_{j'} {v_2}^{j'} \\ {X_{12}}^j_{j'} &= {X_{1m}}^j {X_{2m}}_{j'} \\ {v_2}^{j'} &= {X_{21}}^{j'}_j {v_1}^j \\ {X_{21}}^{j'}_j &= {X_{2m}}^{j'} {X_{1m}}_j \\ {X_{12}}^j_\ell {X_{21}}^\ell_k &= \delta^j_k \\ {v_1}_j &= {X_{21}}^{j'}_j {v_2}_{j'} \\ {v_2}_{j'} &= {X_{12}}^j_{j'} {v_1}_j \end{aligned}$$

12. The affine connection. To adjust the operation of slope to the new, coordinate-dependent, transformation, we must now consider transformations between two coordinate systems located arbitrarily closely together, say

$$\begin{aligned} r_1 &= r & \qquad r_2 &= r + \Delta r \\ \theta_1 &= \theta & \qquad \theta_2 &= \theta + \Delta\theta \end{aligned}$$

for sufficiently small Δr, $\Delta\theta$.

Then

$$\begin{pmatrix} {v_1}^r \\ {v_1}^\theta \end{pmatrix} = \begin{pmatrix} 1 & \\ & 1/r_1 \end{pmatrix} \begin{pmatrix} c_{21} & -s_{21} \\ s_{21} & c_{21} \end{pmatrix} \begin{pmatrix} 1 & \\ & r_2 \end{pmatrix} \begin{pmatrix} {v_2}^r \\ {v_2}^\theta \end{pmatrix}$$

becomes

$$\begin{pmatrix} {v_1}^r \\ {v_1}^\theta \end{pmatrix} = \begin{pmatrix} 1 & \\ & 1/r \end{pmatrix} \begin{pmatrix} c_{21} & -s_{21} \\ s_{21} & c_{21} \end{pmatrix} \begin{pmatrix} 1 & \\ & r+\Delta r \end{pmatrix} \begin{pmatrix} {v_2}^r \\ {v_2}^\theta \end{pmatrix}$$

and since $\Delta\theta$ is small, $c_{21} = \cos(\theta_2 - \theta_1) = \cos(\Delta\theta) \approx 1$ and $s_{21} = \sin(\theta_2 - \theta_1) = \sin(\Delta\theta) \approx \Delta\theta$, we now have

$$\begin{aligned} \begin{pmatrix} {v_1}^r \\ {v_1}^\theta \end{pmatrix} &\approx \begin{pmatrix} 1 & \\ & 1/r \end{pmatrix} \begin{pmatrix} 1 & -\Delta\theta \\ \Delta\theta & 1 \end{pmatrix} \begin{pmatrix} 1 & \\ & r+\Delta r \end{pmatrix} \begin{pmatrix} {v_2}^r \\ {v_2}^\theta \end{pmatrix} \\ &\approx \left(\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 0 & -r\Delta\theta \\ \Delta\theta/r & \Delta r/r \end{pmatrix} \right) \begin{pmatrix} {v_2}^r \\ {v_2}^\theta \end{pmatrix} \end{aligned}$$

eliminating powers of Δr and $\Delta\theta$ higher than the first.

We can write this transformation as

$$X = I + \Delta X$$

and we see that ΔX is linear in the small change $(\Delta r,\ \Delta\theta)$.

So we'd like to express ΔX as the product of something with the vector $(\Delta r\ \ \Delta\theta)$.

Unfortunately there is no way that "something" can be a matrix—try it! We're going to need to go to three indices,

$$\Delta X^j_\ell = \Gamma^j_{k\ell}\Delta^k$$

so we'll switch to relations (Note 4).

ΔXud(j	ℓ	ΔX)
r	θ	$-r\Delta\theta$
θ	r	$\Delta\theta/r$
θ	θ	$\Delta r/r$

Δ(k	Δ)
r	Δr
θ	$\Delta\theta$

And here's the Γ that will do it—try it!

$\Gamma_{\text{udd}}(j$	k	ℓ	Γ)
r	θ	θ	$-r$
θ	r	θ	$1/r$
θ	θ	r	$1/r$

Γ is called the *affine connection* because it is a linear connection between coordinate systems which are translated and linearly deformed from one another. It is also commonly named after Elwin Bruno Christoffel.

Notice that Γ is symmetric on k and ℓ.

Notice also that the index positions are dictated by the contravariance of the coordinates being transformed.

$$\begin{aligned} {v_1}^j &= (\delta^j_k + \Delta x^j_k){v_2}^k \\ \Delta X^j_\ell &= \Gamma^j_{k\ell}\Delta^k \end{aligned}$$

The dummy indices, over which summations are performed, are alternatively up and down, to keep the result a tensor, and the other indices match partners across the = signs.

Because it arises from very small changes in the coordinates, we can obtain Γ directly from slopes. We'll investigate the slopes of the basis vectors $\boldsymbol{e}_k$

$$\partial_j \boldsymbol{e}_k$$

Since the basis vectors located at point x^j change as x^j changes, $\partial_j \boldsymbol{e}_k$ reflects this change. We'll use transformation X from Note 10 and earlier

$$\begin{aligned} f^j &= X_m{}^j f_m \\ f^j \boldsymbol{e}_j = \boldsymbol{f} &= f_m \boldsymbol{\epsilon}_m \\ &= f_m \delta_{mn} \boldsymbol{\epsilon}_n \\ &= f_m X_m{}^j X_{nj} \boldsymbol{\epsilon}_n \end{aligned}$$

so

$$\boldsymbol{e}_j = X_{nj} \boldsymbol{\epsilon}_n$$

or

$$\boldsymbol{\epsilon}_n = X_n{}^j \boldsymbol{e}_j$$

Thus

$$\begin{aligned} \partial_j \boldsymbol{e}_k &= \partial_j (X_{nk} \boldsymbol{\epsilon}_n) \\ &= \partial_j (X_{nk}) \boldsymbol{\epsilon}_n \\ &= \partial_j (X_{nk}) X_n{}^\ell \boldsymbol{e}_\ell \\ &= \Gamma^\ell_{jk} \boldsymbol{e}_\ell \end{aligned}$$

where we've defined

$$\Gamma^\ell_{jk} \stackrel{\text{def}}{=} \partial_j (X_{nk}) X_n{}^\ell$$

Let's see if they are the same as before in polar coordinates.

$$\begin{pmatrix} \mathbf{e}^r \\ \mathbf{e}^\theta \end{pmatrix} = \begin{pmatrix} c & s \\ -s/r & c/r \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix}$$

∂_{d}(j ∂)

j	∂
r	slope_r
θ	slope_θ

X_{du}(n k X)

n	k	X
1	r	c
1	θ	$-rs$
2	r	s
2	θ	rc

$(\partial X)_{\text{ddu}}$(j k n ∂)

j	k	n	∂
θ	r	1	$-s$
θ	r	2	c
r	θ	1	$-s$
r	θ	2	c
θ	θ	1	$-rc$
θ	θ	2	$-rs$

X_{mu}(n ℓ X)

n	ℓ	X
1	r	c
1	θ	$-s/r$
2	r	s
2	θ	c/r

$(\partial X \bowtie X)_{\mathrm{ddmu}}$(j	k	n	ℓ	∂XX)
θ	r	1	r	$-cs$
θ	r	2	r	cs
θ	r	1	θ	s^2/r
θ	r	2	θ	c^2/r
r	θ	1	r	$-cs$
r	θ	2	r	cs
r	θ	1	θ	s^2/r
r	θ	2	θ	c^2/r
θ	θ	1	r	$-rc^2$
θ	θ	2	r	$-rs^2$
θ	θ	1	θ	cs
θ	θ	2	θ	$-cs$

$(X\partial X)_{\mathrm{ddu}}$(j	k	ℓ	$X\partial X$)
θ	r	θ	$1/r$
r	θ	θ	$1/r$
θ	θ	r	$-r$

Γ_{udd}(ℓ	j	k	Γ)
r	θ	θ	$-r$
θ	r	θ	$1/r$
θ	θ	r	$1/r$

Γ_{udd} is what we got before.

We can also show the symmetry: since

$$X_{mj} = \partial_j x_m$$

is itself a slope

$$\begin{aligned}\Gamma^\ell_{jk} &= (\partial_j(\partial_k x_m))X_n{}^\ell \\ &= (\partial_k(\partial_j x_m))X_n{}^\ell \\ &= \Gamma^\ell_{kj}\end{aligned}$$

And we can show that

$$\partial_j \boldsymbol{e}^k = -\Gamma^k_{j\ell}\boldsymbol{e}^\ell$$

To do this, note that

$$\begin{aligned}g_{kk'} &= X_{mk}X_{mk'} \\ \partial_j g_{kk'} &= X_{mk}\partial_j X_{mk'} + X_{mk'}\partial_j X_{mk} \\ &= g_{\ell k}X_m{}^\ell \partial_j X_{mk'} + g_{\ell k'}X_m{}^\ell \partial_j X_{mk} \\ &= g_{\ell k}\Gamma^\ell_{jk'} + g_{\ell k'}\Gamma^\ell_{jk}\end{aligned}$$

where we've used g_{dd} to lower indices in X_{mu}. We use this new relationship in the following.

We assume an analogue to Γ which we'll call H and we'll show that $H = -\Gamma$. We know, from the definition of Γ,

$$\Gamma^\ell_{jk}\boldsymbol{e}_\ell = \partial_j \boldsymbol{e}_k$$

Here is the analogous definition of H and what follows from it.

$$\begin{aligned}H^k_{j\ell}\boldsymbol{e}^\ell &= \partial_j \boldsymbol{e}^k = \partial_j(g^{kk'}\boldsymbol{e}_{k'}) \\ &= \partial_j(g^{kk'})\boldsymbol{e}_{k'} + g^{kk'}\partial_j \boldsymbol{e}_{k'} \\ &= g^{\ell k}H^{k'}_{j\ell}\boldsymbol{e}_{k'} + g^{\ell k'}H^k_{j\ell}\boldsymbol{e}_{k'} + g^{kk'}\partial_j \boldsymbol{e}_{k'} \\ &= H^{k'}_{j\ell}g^{\ell k}g_{k'\ell'}\boldsymbol{e}^{\ell'} + H^k_{j\ell}\boldsymbol{e}^\ell + g^{kk'}\Gamma^\ell_{jk'}g_{\ell\ell'}\boldsymbol{e}^{\ell'} \\ 0 &= H^{k'}_{j\ell}g^{\ell k}g_{k'\ell'} + g^{kk'}\Gamma^\ell_{jk'}g_{\ell\ell'} \\ &= H^\ell_{jk'}g^{\ell k}g_{\ell\ell'} + g^{kk'}\Gamma^\ell_{jk'}g_{\ell\ell'}\end{aligned}$$

and it suffices for the latter to hold that

$$H^{\ell}_{jk'} = -\Gamma^{\ell}_{jk'}$$

In the above derivation the following explanations may help. In line 3 we introduced H as the analogue of Γ in taking the slope of the metric. In line 5 we cancelled two terms and then equated the coefficients of $\boldsymbol{e}^{\ell'}$ because that vector itself is not zero. In line 6 we exchanged names of two dummy indices: $k' \leftrightarrow \ell$.

When we introduced the metric g_{dd} in Note 9 we got away from the orthonormal coordinate system we had started with and became able to work purely in the general coordinate systems we're now transforming into each other.

We should be able to do that, too, with the affine connection Γ. We have derived Γ in terms of the slope of the orthonormal-to-general transformation, X, but we also know that g is the product of X with its transpose.

So can we find Γ_{udd} in terms of slopes of g_{dd}? A clue is the relationship we've just derived. Since $g_{kk'} = X_{mk}X_{mk'}$ we showed that $\partial_j g_{kk'} = g_{\ell k}\Gamma^{\ell}_{jk'} + g_{\ell k'}\Gamma^{\ell}_{jk}$

We can write this down twice more permuting the indices and then add the first two and and subtract the third to eliminate all but one of the Γ_{udd} s.

$$\begin{aligned}
\partial_j g_{kk'} &= g_{\ell k}\Gamma^{\ell}_{jk'} + g_{\ell k'}\Gamma^{\ell}_{jk} \\
\partial_{k'} g_{jk} &= g_{\ell j}\Gamma^{\ell}_{k'k} + g_{\ell k}\Gamma^{\ell}_{k'j} \\
\partial_k g_{k'j} &= g_{\ell k'}\Gamma^{\ell}_{kj} + g_{\ell j}\Gamma^{\ell}_{kk'} \\
2g_{\ell k}\Gamma^{\ell}_{jk'} &= \partial_j g_{kk'} + \partial_{k'} g_{jk} - \partial_k g_{k'j} \\
\Gamma^{\ell}_{jk'} &= \frac{1}{2} g^{\ell k}(\partial_j g_{kk'} + \partial_{k'} g_{jk} - \partial_k g_{k'j})
\end{aligned}$$

Note that the symmetry of Γ_{udd} follows from the symmetry of g_{dd}.

The affine connection Γ_{udd} depends only on the metric g_{dd} (and it involves slopes of g_{dd}). This means that we can work with spaces—such as the 2-D surface of a sphere—for which there is no obvious orthonormal coordinate system to build on.

The properties of a coordinate system are determined by the metric alone.

13. Parallel transport and geodesics. The components of a vector $\boldsymbol{v}$ also vary, as we have seen, as the location of the coordinate system giving those components is changed. We can describe this variance using Γ_{udd}.

Since

$$\boldsymbol{v} = v^j \boldsymbol{e}_j$$

we can find the component v^k

$$\boldsymbol{v} \cdot \boldsymbol{e}^k = v^j \boldsymbol{e}_j \boldsymbol{e}^k = v^j \delta^k_j = v^k$$

So the slope of a component

$$\begin{aligned}
\partial_j v^k &= \partial_j(\boldsymbol{v} \cdot \boldsymbol{e}^k) \\
&= \partial_j(\boldsymbol{v}) \cdot \boldsymbol{e}^k + \boldsymbol{v} \cdot \partial_j \boldsymbol{e}^k \\
&= \partial_j(\boldsymbol{v}) \cdot \boldsymbol{e}^k - \boldsymbol{v} \cdot \Gamma^k_{jj'} \boldsymbol{e}^{j'} \\
&= \partial_j(\boldsymbol{v}) \cdot \boldsymbol{e}^k - \Gamma^k_{jj'} v^{j'}
\end{aligned}$$

where $\partial_j(\boldsymbol{v}) \cdot \boldsymbol{e}^k$ is the kth contravariant component of the slope of the vector $\boldsymbol{v}$.

What if $\partial_j(\boldsymbol{v}) = 0$? This means that the vector does not change as we move its location from one local coordinate system to another. That is, since the origin of $\boldsymbol{v}$ can be anywhere in the space (a vector is an *interval*, not an absolute position), the vector $\boldsymbol{v}$ can be thought of as moving parallel to itself.

This is *parallel transport*.

$$\partial_j v^k = -\boldsymbol{v} \cdot \Gamma^k_{jj'} \boldsymbol{e}^{j'}$$

We can use it to create *geodesics*. A geodesic may be described as the shortest path between two points. Or we can explore a more useful definition.

In a flat space, a geodesic is a straight line, the original concept given by the shortest-path definition.

In a curved space (which we haven't come to yet) such as the 2-D surface of a sphere, a geodesic is a segment of a "great circle", the line resulting from cutting the sphere by a plane through its centre.

It is fairly apparent, and an alternate definition which is better for us is, that a geodesic can be generated by a small vector, moved repeatedly parallel to itself by placing its origin at its previous tip.

$$v'^k = v^k + \Delta v^k$$

where

$$\Delta v^k = (\partial_j v^k)\Delta^j = -\Gamma^k_{jj'} v^{j'} \Delta^j$$

so

$$v'^k = (\delta^k_{j'} - \Gamma^k_{jj'}\Delta^j)v^{j'}$$

(Note the sign difference from the beginning of Note 12: why?)

We can illustrate geodesic generation for polar coordinates in a plane because we already have, from Note 12 (in the mud notation a $_{\mathrm{d}}$ and a $_{\mathrm{u}}$ are omitted because they are dummy variables and summed over)

$$\Gamma_{\mathrm{u-d}}\Delta_{-} = \begin{pmatrix} 0 & -r\Delta^\theta \\ \Delta^\theta/r & \Delta^r/r \end{pmatrix}$$

A MATLAB function `[v, origin] = geodPlanePolarStep(v,origin)` can calculate

$$v = v - \begin{pmatrix} 0 & -r\Delta^\theta \\ \Delta^\theta/r & \Delta r/r \end{pmatrix} v$$

and the new origin

$$\begin{pmatrix} r + \Delta^r \\ \theta + \Delta^\theta \end{pmatrix}$$

given vector

$$v = \begin{pmatrix} \Delta^r \\ \Delta^\theta \end{pmatrix}$$

and its origin (r, θ).

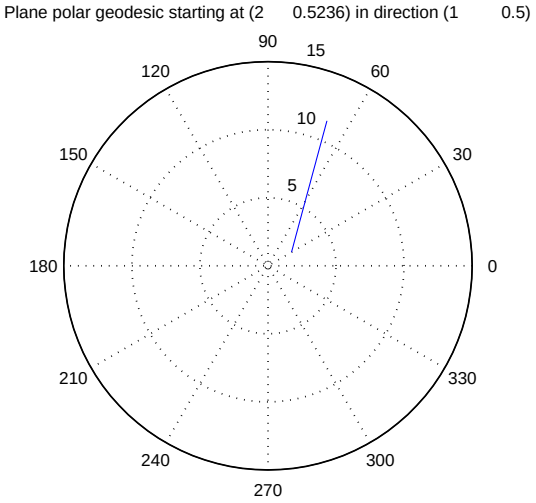
This step can be iterated `n` times by

```
ends = geodesicPlanePolar(v,start,n,stepsize)
```

which reduces `v` to a vector in the same direction but divided in length down to `stepsize`, and starting the iteration at origin `start`

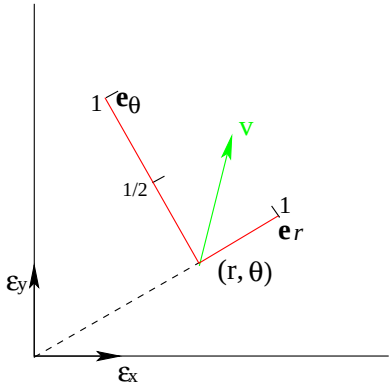
Here is a result with, initially

direction of v	start	n	stepsize
1	2	1000	0.01
1/2	$\pi/6$		



We see the 1000 steps as a straight line—a geodesic in this flat space, never mind that it is all calculated in terms of varying polar coordinates.

Note that the vector $\boldsymbol{v} = (1 \;\; 1/2)^{\mathrm{T}}$ appears in the initial coordinate system at $(r, \theta) = (2, \pi/6)$ as a vector from the origin at $(r, \theta) = (2, \pi/6)$ to the point $(1\Delta r, (1/2)r\Delta\theta)$



14. Absolute slopes. From Note 13 we have

$$(\partial_j \boldsymbol{v}) \cdot \boldsymbol{e}^k = \partial_j v^k + \Gamma^k_{j\ell} v^\ell$$

and we can see that neither $\partial_j v^k$ nor $\Gamma^k_{j\ell} v^\ell$ is a tensor, even though v^k is. Hence neither ∂_j nor $\Gamma^k_{j\ell}$ is a tensor.

For v^k transforms as

$$v_2{}^{k'} = X_{21}{}^{k'}_{k} \, v_1{}^k$$

but

$$\begin{aligned} \partial_{2j'} v_2{}^{k'} &= X_{21}{}^{j}_{j'} \partial_{1j} (X_{21}{}^{k'}_{k} \, v_1{}^k) \\ &\neq X_{21}{}^{j}_{j'} X_{21}{}^{k'}_{k} (\partial_{1j} v_1{}^k) \end{aligned}$$

because X_{21} now generally depends on the coordinate ${x_1}^j$.

We want to claim that $(\partial_j \boldsymbol{v}) \cdot \boldsymbol{e}^k$ *is* a tensor, because $\boldsymbol{v}$ and $\boldsymbol{e}^k$ both certainly are tensors, being vectors independent of any coordinate system.

Let's check an example, using the radial field $\boldsymbol{f} = \boldsymbol{r} r^{p-1}$ in polar coordinates.

$$\begin{aligned}
\begin{pmatrix} f^r \\ f^\theta \end{pmatrix} &= \begin{pmatrix} c & s \\ -s/r & c/r \end{pmatrix} \begin{pmatrix} f_x \\ f_y \end{pmatrix} \\
&= \begin{pmatrix} c & s \\ -s/r & c/r \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} r^{p-1} \\
&= \begin{pmatrix} c & s \\ -s/r & c/r \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} r^{p} \\
&= \begin{pmatrix} r^p \\ 0 \end{pmatrix}
\end{aligned}$$

and

$$\begin{pmatrix} \partial_r \\ \partial_\theta \end{pmatrix} = \begin{pmatrix} c & s \\ -sr & cr \end{pmatrix} \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix}$$

In Cartesian coordinates (x, y) the metric g is independent of x, y: it is the identity matrix $g_{jk} = \delta_{jk}$. So the affine connection, being a combination of slope g with respect to the coordinates, is 0.

So we have in orthogonal coordinates (we'll call them the 0-system, so the affine connection ${\Gamma'}_1 = \Gamma_0$)

$$\begin{aligned}
\left(\begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} \boldsymbol{f} \right) \cdot (\boldsymbol{\epsilon}_x \;\; \boldsymbol{\epsilon}_y) &= \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} (f_x \;\; f_y) + {\Gamma_0}_{kj\ell} f_\ell \\
&= \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} (f_x \;\; f_y) \\
&= \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} (x \;\; y) r^{p-1} \\
&= \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} r^{p-1} + (p-1) r^{p-3} \begin{pmatrix} x^2 & xy \\ xy & y^2 \end{pmatrix}
\end{aligned}$$

where $\Gamma_0 = 0$ gives the second line and the fourth line was worked out in Note 10.

In polar coordinates (system 1)

$$\begin{aligned}
\left(\begin{pmatrix} \partial_r \\ \partial_\theta \end{pmatrix} \boldsymbol{f} \right) \cdot (\boldsymbol{e}^r \;\; \boldsymbol{e}^\theta) &= \begin{pmatrix} \partial_r \\ \partial_\theta \end{pmatrix} (f^r \;\; f^\theta) + {\Gamma_1}^k{}_{j\ell} f^\ell \\
&= \begin{pmatrix} \partial_r \\ \partial_\theta \end{pmatrix} (r^p \;\; 0) + {\Gamma_1}_{\mathrm{udd}} \bowtie f_{\mathrm{u}} \\
&= p \begin{pmatrix} 1 & \end{pmatrix} r^{p-1} + \begin{pmatrix} & \\ & 1 \end{pmatrix} r^{p-1} \\
&= \begin{pmatrix} p & \\ & 1 \end{pmatrix} r^{p-1}
\end{aligned}$$

where $\Gamma_{\text{1udd}} \bowtie f_{\text{u}}$ in line 2 expresses the join of

$\Gamma_{\text{1udd}}(k$	j	ℓ	$\Gamma\ )$
r	θ	θ	$-r$
θ	r	θ	$1/r$
θ	θ	r	$1/r$

with

$f_{\text{u}}(\ell$	$f\ )$
r	r^p

But we can get this result by tensor transform of $(\partial_j \boldsymbol{f}) \cdot \boldsymbol{e}^k$ (using $x = cr, y = sr$):
We need

$$\begin{aligned}
\begin{pmatrix} \boldsymbol{e}^r \\ \boldsymbol{e}^\theta \end{pmatrix} &= \begin{pmatrix} 1 & \\ & 1/r^2 \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} \\
&= \begin{pmatrix} 1 & \\ & 1/r^2 \end{pmatrix} \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix} \\
&= \begin{pmatrix} c & s \\ -s/r & c/r \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix}
\end{aligned}$$

and

$$\begin{aligned}
(\boldsymbol{e}^r \ \ \boldsymbol{e}^\theta) \begin{pmatrix} v_r \\ v_\theta \end{pmatrix} &= (\boldsymbol{\epsilon}_x \ \ \boldsymbol{\epsilon}_y) \begin{pmatrix} c & -s/r \\ s & c/r \end{pmatrix} \begin{pmatrix} v_r \\ v_\theta \end{pmatrix} \\
&= (\boldsymbol{\epsilon}_x \ \ \boldsymbol{\epsilon}_y) \begin{pmatrix} v_x \\ v_y \end{pmatrix}
\end{aligned}$$

so

$$\begin{pmatrix} v_r \\ v_\theta \end{pmatrix} = \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

and, since ∂_j transforms as does v_j,

$$\begin{pmatrix} \partial_r \\ \partial_\theta \end{pmatrix} = \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix}$$

Thus

$$\begin{aligned}
\left(\begin{pmatrix} \partial_r \\ \partial_\theta \end{pmatrix} \boldsymbol{f} \right) \cdot (\boldsymbol{e}^r \ \ \boldsymbol{e}^\theta) &= \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \boldsymbol{f} \cdot (\boldsymbol{\epsilon}_x \ \ \boldsymbol{\epsilon}_y) \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} \begin{pmatrix} c & -s/r \\ s & c/r \end{pmatrix} \\
&= \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \left[\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} r^{p-1} + (p-1) \begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix} r^{p-1} \right] \begin{pmatrix} c & -s/r \\ s & c/r \end{pmatrix} \\
&= \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} r^{p-1} + (p-1) \begin{pmatrix} 1 \end{pmatrix} r^{p-1} \\
&= \begin{pmatrix} p & \\ & 1 \end{pmatrix} r^{p-1}
\end{aligned}$$

Here $(\partial_j \boldsymbol{f}) \cdot \boldsymbol{e}^k = \partial_j f^k + \Gamma^k_{j\ell} f^\ell$ transforms as a tensor—although the individual components each do not.
We conclude that $(\partial_j \boldsymbol{f}) \cdot \boldsymbol{e}^k$ is a tensor and should be used instead of $\partial_j f^k$ which is not.
We need a notation for the "absolute slope"

$$\mathcal{D}_j f^k \stackrel{\text{def}}{=} (\partial_j \boldsymbol{f}) \cdot \boldsymbol{e}^k = \partial_j f^k + \Gamma^k_{j\ell} f^\ell$$

This tells us how much the vector field $\boldsymbol{v}$ changes itself as we move along coordinate j, without being misled by "changes" in its components due to the changing coordinate system.

15. Gradient, divergence and curl with absolute slope. Now that we have absolute slopes, we use them everywhere we formerly used slopes. In Cartesian coordinates, or coordinate systems whose transformation from Cartesian is independent of the coordinates, the results are the same as before. In other coordinate systems the regular slopes are wrong because they do not transform as tensors. The absolute slopes do.

A source of variety, and confusion, is that there are both covariant and contravariant absolute slopes. We've introduced the covariant one $\mathcal{D}_j$, which is always used, for reasons this Note tries to make clear. But there is also the contravariant absolute slope, $\mathcal{D}^j$, which is equally valid in principle if not in practice.

a) Here are the four possible combinations of absolute slope

Co-co	Co-contra
$\mathcal{D}_j v_k = \partial_j v_k - \Gamma^\ell_{jk} v_\ell$	$\mathcal{D}_j v^k = \partial_j v^k + \Gamma^k_{j\ell} v^\ell$
Contra-co	Contra-contra
$\mathcal{D}^j v_k = \partial^j v_k - g^{jj'} \Gamma^\ell_{j'k} v_\ell$	$\mathcal{D}^j v^k = \partial^j v^k + g^{jj'} \Gamma^k_{j'\ell} v^\ell$

The bottom two are simple transformations of the two above, using the metric g_{uu} to raise the index j from down to up.

The relationship between the top two needs work. Here we derive co-co from co-contra.

We must show that all that changes between the two is the sign on the affine connection Γ_{udd}—and of course a necessary repositioning of the indices to keep up and down consistent.

$$\begin{aligned}
\mathcal{D}_j v_k &= \mathcal{D}_j (g_{kk'} v^{k'}) \\
&= g_{kk'} \mathcal{D}_j v^{k'} \\
&= g_{kk'} \partial_j v^{k'} + g_{kk'} \Gamma^{k'}_{j\ell} v^\ell \\
&= \partial_j (g_{kk'} v^{k'}) - (\partial_j g_{kk'}) v^{k'} + g_{kk'} \Gamma^{k'}_{j\ell} v^\ell \\
&= \partial_j (g_{kk'} v^{k'}) - (g_{k'\ell'} \Gamma^{\ell'}_{jk} + g_{k\ell} \Gamma^\ell_{jk'}) v^{k'} + g_{kk'} \Gamma^{k'}_{j\ell} v^\ell \\
&= \partial_j v_k - (g_{k'\ell'} \Gamma^{\ell'}_{jk} + g_{k\ell} \Gamma^\ell_{jk'} - g_{k\ell} \Gamma^\ell_{jk'}) v^{k'} \\
&= \partial_j v_k - g_{k'\ell'} \Gamma^{\ell'}_{jk} v^{k'} \\
&= \partial_j v_k - \Gamma^\ell_{jk} v_\ell
\end{aligned}$$

In the above, the second line follows from the important fact that the absolute slope of the metric is zero: the metric is constant. We show this below, but first continue commenting on the above derivation.

The third line is the co-contra absolute slope which we showed in Note 14.

The fourth line uses the rule for the slope of a product.

The fifth line uses the expansion of ∂g from Note 12.

Then it is just index rearrangement, cancellation and lowering the index of v.

Now we show that the metric is always constant. We use

$$\Gamma^\ell_{jk} = \frac{1}{2} g^{\ell\rho} (\partial_j g_{k\rho} + \partial_k g_{j\rho} - \partial_\rho g_{jk})$$

from Note 12 and an extension, which we do not prove, of the absolute slope (co-co) applied to variables of more than one index.

$$\begin{aligned}
\mathcal{D}_j g_{kk'} &= \partial_j g_{kk'} - \Gamma^\ell_{jk} g_{kk'} - \Gamma^\ell_{jk'} g_{k\ell} \\
&= \partial_j g_{kk'} - \frac{1}{2}\delta^\rho_{k'}(\partial_j g_{k\rho} + \partial_k g_{j\rho} - \partial_\rho g_{jk}) - \frac{1}{2}\delta^\rho_k(\partial_j g_{k'\rho} + \partial_{k'} g_{j\rho} - \partial_\rho g_{jk'}) \\
&= \partial_j g_{kk'} - \frac{1}{2}(\partial_j g_{kk'} + \partial_k g_{jk'} - \partial_{k'} g_{jk} + \partial_j g_{kk'} + \partial_{k'} g_{jk} - \partial_k g_{jk'}) \\
&= 0
\end{aligned}$$

b) There are two possible gradients, now using absolute slopes in place of the ordinary slope.

Co	Contra
$\mathcal{D}_j S = \partial_j S$	$\mathcal{D}^j S = \partial^j S$

These need no affine connection because the field whose slope is being found has no indices.

c) One of the four varieties of curl, co-co, has a particularly simple form, because of the symmetry of Γ.

Co-co

$$\begin{aligned}
\mathcal{D}_j v_k - \mathcal{D}_k v_j &= \partial_j v_k - \Gamma^\ell_{jk} v_\ell \\
&\quad -\partial_k v_j + \Gamma^\ell_{kj} v_\ell \\
&= \partial_j v_k - \partial_k v_j
\end{aligned}$$

Co-contra

$$\begin{aligned}
\mathcal{D}_j v^k - \mathcal{D}_k v^j &= \partial_j v^k + \Gamma^k_{j\ell} v^\ell \\
&\quad -\partial_k v^j - \Gamma^j_{k\ell} v^\ell
\end{aligned}$$

Contra-co

$$\begin{aligned}
\mathcal{D}^j v_k - \mathcal{D}^k v_j &= \partial^j v_k - g^{jj'}\Gamma^\ell_{jj'} v^\ell \\
&\quad -\partial^k v_j + g^{kk'}\Gamma^\ell_{kk'} v^\ell
\end{aligned}$$

Contra-contra

$$\begin{aligned}
\mathcal{D}^j v^k - \mathcal{D}^k v^j &= \partial^j v^k - \partial^k v^j \\
&\quad +g^{jj'} g^{kk'}(\partial_{j'} g_{\ell k'} - \partial_{k'} g_{\ell j'}) v^\ell
\end{aligned}$$

d) We saved divergence for last because the co-contra form can be simplified, with some work, to a form involving $\sqrt{g}$, the square root of the determinant of g_{dd}.

We'll give an unexplained derivation immediately below, and explain it after presenting all four

versions.

Co-co

$$\begin{aligned}\mathcal{D}_j v_j &= \partial_j v_j - \Gamma^\ell_{jj} v_\ell \\ &= \partial_j v_j - \frac{1}{2} g^{\ell\rho}(\partial_j g_{j\rho} + \partial_j g_{j\rho} - \partial_\rho g_{jj}) v_\ell \\ &= \partial_j v_j + \frac{1}{2} g^{\ell\rho}(\partial_\rho g_{jj}) v_\ell - g^{\ell\rho}(\partial_j g_{j\rho}) v_\ell\end{aligned}$$

Co-contra

$$\begin{aligned}\mathcal{D}_j v^j &= \partial_j v^j + \Gamma^j_{j\ell} v^\ell \\ &= \partial_j v^j + \frac{1}{2} g^{j\ell'}(\partial_j g_{\ell\ell'} + \partial_\ell g_{j\ell'} - \partial_{\ell'} g_{j\ell}) v^\ell \\ &= \partial_j v^j + \frac{1}{2} g^{j\ell'} \partial_\ell g_{j\ell'} v^\ell \\ &\quad + (\frac{1}{2} g^{j\ell'} \partial_j g_{\ell\ell'} - \frac{1}{2} g^{j\ell'} \partial_{\ell'} g_{j\ell}) v^\ell \\ &= \partial_j v^j + \frac{1}{2} g^{j\ell'} \partial_\ell g_{j\ell'} v^\ell \\ &= \partial_j v^j + \frac{1}{2} \mathrm{tr}(G^{-1} \partial_\ell G) v^\ell \\ &= \partial_j v^j + \frac{1}{2} (\partial_\ell \ln \det G) v^\ell \\ &= \partial_j v^j + (\partial_\ell \ln \sqrt{g}) v^\ell \\ &= \partial_j v^j + \frac{1}{\sqrt{g}} (\partial_\ell \sqrt{g}) v^\ell \\ &= \frac{1}{\sqrt{g}} \partial_\ell (\sqrt{g} v^\ell)\end{aligned}$$

Contra-co

$$\mathcal{D}^j v_j = \partial^j v_j - g^{jj'} \Gamma^\ell_{j'j} v_\ell$$

Contra-contra

$$\mathcal{D}^j v^j = \partial^j v^j + g^{jj'} \Gamma^j_{j\ell} v^\ell$$

The co-contra version is the one that is used. The derivation needs some amplification. In line 5, G is defined as the matrix g_{dd}, the covariant metric tensor. Line 6 starts a derivation, for any matrix G, which takes us back to line 5: see below. Line 7 uses the conventional definition $g = \det G$.

Here is the derivation that takes us from line 5 to line 6 in the co-contra argument above.

$$\begin{aligned}\partial_\ell \ln \det G &\approx \frac{1}{\Delta x^\ell} \Delta \ln \det G \\ &= \frac{1}{\Delta x^\ell} (\ln \det(G + \Delta G) - \ln \det G) \\ &= \frac{1}{\Delta x^\ell} \ln \left(\frac{\det(G + \Delta G)}{\det G} \right) \\ &= \frac{1}{\Delta x^\ell} \ln \det(G^{-1}(G + \Delta G)) \\ &= \frac{1}{\Delta x^\ell} \ln \det(1 + G^{-1} \Delta G) \\ &\approx \frac{1}{\Delta x^\ell} \ln(1 + \mathrm{tr}(G^{-1} \Delta G)) \\ &\approx \frac{1}{\Delta x^\ell} \mathrm{tr}(G^{-1} \Delta G) \\ &\approx \mathrm{tr}(G^{-1} \partial_\ell G)\end{aligned}$$

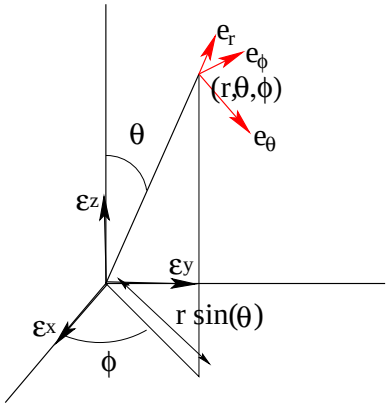
And even this derivation needs comment. The jump from det in line 5 to trace in line 6 follows because the ΔG produces dth powers of shrinkingly small quantities in the determinant everywhere

except on the diagonal. For example, in 2D, $d = 2$:

$$\begin{vmatrix} 1+g^{11}\Delta g_{11} & g^{12}\Delta g_{12} \\ g^{21}\Delta g_{21} & 1+g^{22}\Delta g_{22} \end{vmatrix} = 1 + g^{11}\Delta g_{11} + g^{22}\Delta g_{22} + \mathcal{O}(\Delta^2)$$

The step in line 7 uses the expansion $\ln(1+x) = x + x^2/2 + ..$, stopping at the linear term.

16. Spherical polar coordinates. The three-dimensional extension of plane polar coordinates involves two angles.



from which we can read off the metric

$$g_{\rm dd} = \begin{pmatrix} 1 & & \\ & r^2 & \\ & & r^2 s^2 \end{pmatrix}$$

where we can continue to abbreviate $\sin\theta$ as s because there is no dependence on angle ϕ in the metric. (And $c = \cos\theta$.)

From the metric we can get the affine connection

$\Gamma_{\rm udd}(j$	k	ℓ	Γ $)$
r	θ	θ	$-r$
r	ϕ	ϕ	$-rs^2$
θ	r	θ	$1/r$
θ	θ	r	$1/r$
θ	ϕ	ϕ	$-cs$
ϕ	r	ϕ	$1/r$
ϕ	ϕ	r	$1/r$
ϕ	θ	ϕ	c/s
ϕ	ϕ	θ	c/s

We can move on to parallel transport and geodesics on the *surface* of a sphere. This means retaining in the affine connection only those terms independent of r,

$\Gamma_{\rm udd}(j$	k	ℓ	Γ $)$
θ	ϕ	ϕ	$-cs$
ϕ	θ	ϕ	c/s
ϕ	ϕ	θ	c/s

To transport a vector $\boldsymbol{v}$ parallel to itself, winding up as vector $\boldsymbol{v'}$ we had in Note 13

$$v'^j = (\delta^j_k - \Gamma^j_{k\ell}\Delta^\ell)v^k$$

Here the small change Δ^ℓ in coordinates is

$\Delta_{\mathrm{u}}(\ell$	Δ	$)$
θ	Δ^θ	
ϕ	Δ^ϕ	

so the reduced join $\Gamma\Delta = \Gamma \bowtie \Delta$ is

$\Gamma\Delta_{\mathrm{ud}}(j$	k	$\Gamma\Delta$ $)$
θ	ϕ	$-cs\Delta\phi$
ϕ	θ	$-cs\Delta\phi$
ϕ	ϕ	$c/s\Delta\theta$

$$\begin{pmatrix} 0 & -cs\Delta\phi \\ (c/s)\Delta\phi & (c/s)\Delta\theta \end{pmatrix}$$

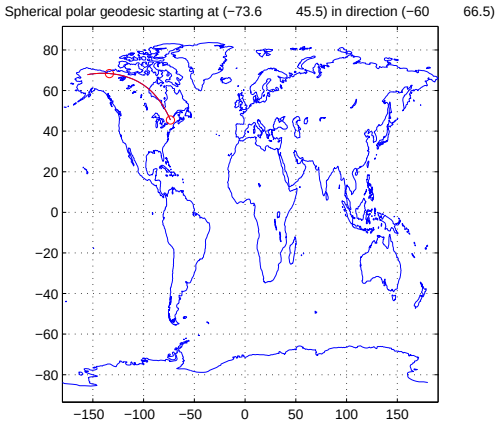
given as a relation on the left and as a matrix on the right.

A MATLAB routine like that for plane polar geodesics in Note 13 can use parallel transport to track geodesics on the surface of the Earth.

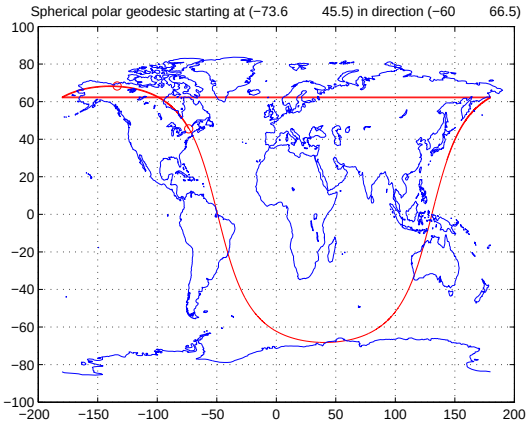
Here is the plot from the invocation

```
ends = geodesicLongLat([-60;66.5],Montreal',Inuvik',80,0.01)
```

where Montreal is at $(-73.6, 45.5)$ degrees and the initial direction is not directly towards Inuvik at $(-133^o43', 68^o21')$ but in direction `(-60,66.50)`, greatly shortened by a factor 100. (This is an *equirectangular* projection from the sphere onto the plane: ϕ and θ are each mapped directly (linearly) to x and y respectively.)

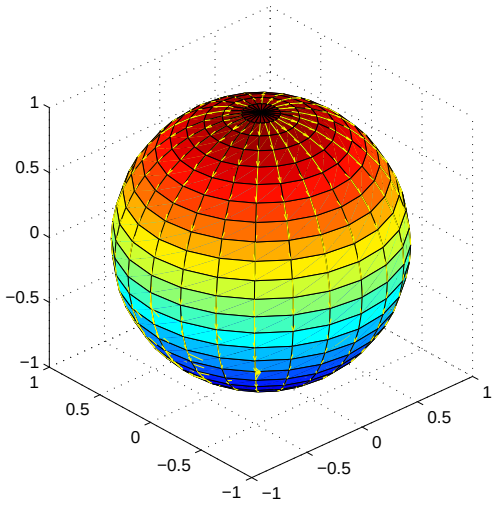


Here is the great circle continuing entirely around the globe: I did not lift the pen as it switched from -180 to 180 degrees of longitude.



We can use our general results from Note 15 on divergence and curl to examine a "magnetic" field and a "rotation" field, respectively, on the surface of a sphere.

Divergence. Here is the "magnetic" field, $v^\theta = 1 = v_\theta$ and $v^\phi = 0 - v_\phi$.



The divergence must be a scalar and only co-contra and contra-co forms can give scalars, so we try both.

Co-contra

$$\begin{aligned}
\partial_j V^j + \Gamma^j_{j\ell} V^\ell &= \cos\theta/\sin\theta \\
\frac{1}{\sqrt{g}}\partial_\ell(\sqrt{g}V^\ell) &= \frac{1}{\sin\theta}\partial_\theta(\sin\theta V^\theta) \\
&= \cos\theta/\sin\theta
\end{aligned}$$

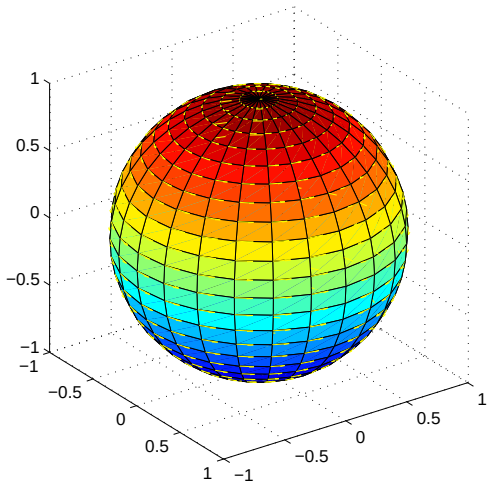
Contra-co

$$\begin{aligned}
\partial_\mu V^\mu + \Gamma^\mu_{\mu\lambda} V^\lambda &= \cos\theta/\sin\theta \\
\frac{1}{\sqrt{g}}\partial_\lambda(\sqrt{g}V^\lambda) &= \frac{1}{\sin\theta}\partial_\theta(\sin\theta V^\theta) \\
&= \cos\theta/\sin\theta
\end{aligned}$$

These results make sense as the divergence of the magnetic field from north pole to south pole:

θ	0	$\pi/4$	$\pi/2$	$3\pi/4$	π
$\cos\theta$	1	$1/\sqrt{2}$	0	$-1/\sqrt{2}$	-1
$\sin\theta$	0	$1/\sqrt{2}$	1	$1/\sqrt{2}$	0
$\cos\theta/\sin\theta$	∞	1	0	-1	$-\infty$

Curl. Here is the "rotation" field, $v^\theta = 0 = v_\theta$, $v_\phi = \sin\theta$ and $v^\phi = 1/\sin\theta$. (Note that we must now accommodate the non-normalized coordinate system by multiplying or dividing by $\sin\theta$ for the down and up second coordinates respectively.

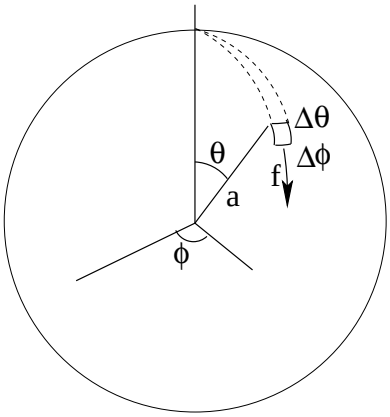


Here is the result from the simplest approach, co-co.

$$\begin{aligned}
\mathcal{D}_j V_k - \mathcal{D}_k V_j &= \partial_j V_k - \partial_k V_j \\
&= \partial_\theta V_\phi - \partial_\phi V_\theta \\
&= \partial_\theta \sin\theta - \partial_\phi 0 \\
&= \cos\theta
\end{aligned}$$

We can check these results with microscopic calculations following on from the microscopic derivations of **div** and **curl** we made in Note 6 for flat space.

Divergence



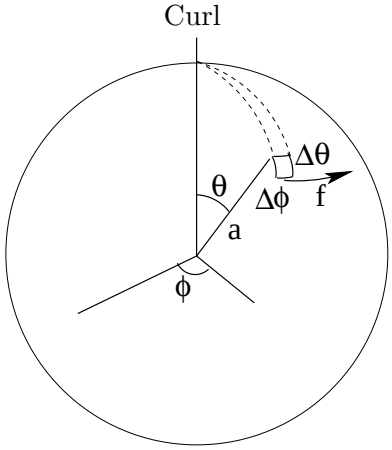
This example is a field of constant magnitude f in the θ direction, the "area" it applies to is $a\sin\theta\Delta\phi$ and the "volume of the region is $(a\Delta\theta)(a\sin\theta\Delta\phi) = a^2 sin\theta\Delta\theta\Delta\phi$.

- total flow at boundary = $fa\sin\theta\Delta\phi$;
- net flow from region = slope of this = $f\cos\theta\Delta\theta\Delta\phi$;
- net flow per unit area =

  $$(f/a)(\cos\theta/\sin\theta)$$

  which agrees with the divergence above and is also plausibly inversely proportional to the radius a of the sphere: at infinite radius the surface would be flat and have no divergence (except maybe at the poles).

Curl

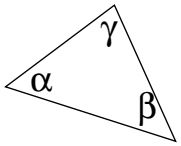


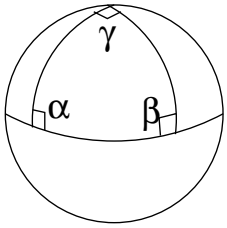
This example is a field of constant magnitude $f\sin\theta$ in the ϕ direction (for a covariant system), the "area" it applies to is again $a\sin\theta\Delta\phi$ and the "volume of the region is again $(a\Delta\theta)(a\sin\theta\Delta\phi) = a^2 sin\theta\Delta\theta\Delta\phi$.

- total flow at boundary = $fa\sin^2\theta\Delta\phi$;
- net flow perpendicular to region = slope of this = $2f\cos\theta\sin\theta\Delta\theta\Delta\phi$;
- net flow per unit area = $(f/a)(2\cos\theta)$, which agrees with the curl above (apart from the 2) and is again plausibly inversely proportional to the radius a of the sphere.

17. Curvature. How can we tell if we live in curved space? Suppose we were two-dimensional creatures who can measure angles in a triangle. Then in flat, Euclidean, space, the angles would total π in any triangle. But in curved space, the total would differ from π. Here is the example of a sphere (which has "positive curvature") with a particular triangle whose *excess angle* is $\pi/2$. (Other triangles on a sphere will have different excess angles, but which will all be positive.)

Excess Angle for Triangles



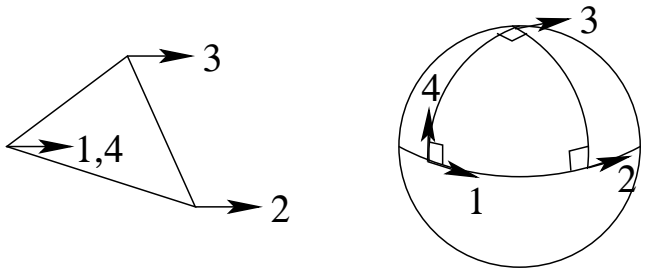


$\alpha+\beta+\gamma=\pi$ $\qquad$ $\alpha=\beta=\gamma=\frac{\pi}{2}$

$$\text{excess angle} = \frac{3\pi}{2} - \pi = \frac{\pi}{2}$$

The excess angle may be measured by parallel transport. In a flat space a vector, transported parallel to itself around a closed curve (for example a triangle), will return to its original orientation. In a curved space it will not. Here is the same triangle on the sphere: we see that the final position, 4, of the vector differs by $\pi/2$ from its original position, 1.

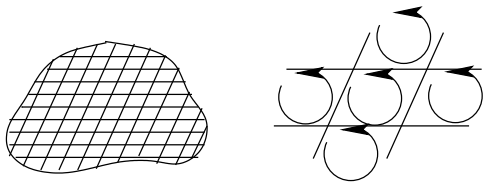
Excess Angle by Parallel Transport



$$\Delta A = \frac{\pi}{2} \quad \text{excess angle}$$

We no longer need triangles but can parallel-transport a vector around an arbitrary closed curve. The diagram shows that this is the same as parallel-transporting the vector around each of a whole bunch of parallelograms decomposing the area contained within the curve: each internal edge will be traversed in both directions, cancelling out the effects of all internal edges and leaving only the effect of the boundary, i.e., the curve itself.

For Arbitrary Curve

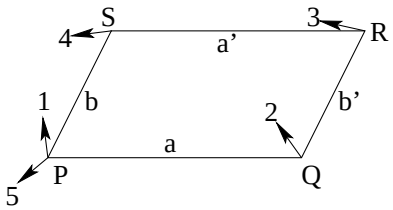


.. all loops cancel,
except on outer boundary

.. so just add up contributions
from each parallelogram

So we can focus on an infinitesimal parallelogram.

For Infinitesimal Parallelogram



The analysis of this involves parallel-transporting the vector, which we'll call **v**, around the parallelogram, and parallel-transporting the edges a and b to give the other edges, respectively a' and b', of the parallelogram itself.

There are three preliminaries.

First, a and b must be parallel-transported to a' and b'.

$$b'^{\nu} \;=\; b^{\nu} - \Gamma_{P}{}^{\nu}_{\gamma\delta} a^{\gamma} b^{\delta}$$

$$a'^{\nu} = a^{\nu} - \Gamma_{P}{}^{\nu}_{\gamma\delta} b^{\gamma} a^{\delta}$$

because the change in coordinate for moving a, $\Delta x = b$ and the change in coordinate for moving b, $\Delta x = a$.

Note that in the four dimensions of timespace it is conventional to use Greek indices. More Greek letters are available and free of alternative meanings than Roman. We will find that curvature becomes really interesting in four or more dimensions. So we're going to switch to Greek letters for the remaining Notes.

Second, note that Γ has a midscript, P, to indicate that the Γ has been calculated as at point P. Since we will also need Γ at points Q and S, this brings us to the second preliminary. We will need the slope of Γ to approximate the infinitesimal transportation from P to these points.

$$\begin{aligned} \Gamma_{Q}{}^{\mu}_{\nu\lambda} &= \Gamma_{P}{}^{\mu}_{\nu\lambda} + a^{\alpha} \mathrm{slope}_{\alpha} \Gamma_{P}{}^{\mu}_{\nu\lambda} \\ \Gamma_{S}{}^{\mu}_{\nu\lambda} &= \Gamma_{P}{}^{\mu}_{\nu\lambda} + b^{\alpha} \mathrm{slope}_{\alpha} \Gamma_{P}{}^{\mu}_{\nu\lambda} \end{aligned}$$

The third preliminary is a similar transportation of the vector $\mathbf{v}$ which is to be transported around the entire parallelogram.

$$\begin{aligned} v_{Q}{}^{\lambda} &= v_{P}{}^{\lambda} - \Gamma_{P}{}^{\lambda}_{\alpha\beta} a^{\alpha} v_{P}{}^{\beta} \\ v_{S}{}^{\lambda} &= v_{P}{}^{\lambda} - \Gamma_{P}{}^{\lambda}_{\alpha\beta} b^{\alpha} v_{P}{}^{\beta} \end{aligned}$$

We'll work out the expression for the transport of $\mathbf{v}$ from P to Q to R, from which we can easily get the transport of $\mathbf{v}$ from P to S to R by swapping a and b. Then we put the first together with the negative of the second to get the transport of $\mathbf{v}$ from P to Q to R to S and back to P.

So we start work with only the first of each of the above three pairs of preliminary equations. We break down $\Delta_{PQR} = \Delta_{PQ} + \Delta_{QR}$, where $\Delta_{PQ} v = v_Q - v_P$ etc.

$$\begin{aligned} \Delta_{PQ} v^{\mu} &= -\Gamma_{P}{}^{\mu}_{\nu\lambda} a^{\nu} v_{P}{}^{\lambda} \\ \Delta_{QR} v^{\mu} &= -\Gamma_{Q}{}^{\mu}_{\nu\lambda} b'^{\nu} v_{Q}{}^{\lambda} \\ &= -(\Gamma_{P}{}^{\mu}_{\nu\lambda} + a^{\alpha} \mathrm{slope}_{\alpha} \Gamma_{P}{}^{\mu}_{\nu\lambda})(b^{\nu} - \Gamma_{P}{}^{\nu}_{\gamma\delta} a^{\gamma} b^{\delta})(v_{P}{}^{\lambda} - \Gamma_{P}{}^{\lambda}_{\alpha\beta} a^{\alpha} v_{P}{}^{\beta}) \\ &= -\Gamma_{P}{}^{\mu}_{\nu\lambda} b^{\nu} v_{P}{}^{\lambda} + \Gamma_{P}{}^{\mu}_{\nu\lambda} \Gamma_{P}{}^{\nu}_{\gamma\delta} a^{\gamma} b^{\delta} v_{P}{}^{\lambda} + \Gamma_{P}{}^{\mu}_{\nu\lambda} b^{\nu} \Gamma_{P}{}^{\lambda}_{\alpha\beta} a^{\alpha} v_{P}{}^{\beta} - a^{\alpha} (\mathrm{slope}_{\alpha} \Gamma_{P}{}^{\mu}_{\nu\lambda}) b^{\nu} v_{P}{}^{\lambda} + .. \end{aligned}$$

where the terms left out in the last equality are those of orders higher than ab: the approximation is linear.

Thus (and we drop the P midscript since it now appears everywhere),

$$\Delta_{PQR} v^{\mu} = -\Gamma^{\mu}_{\nu\lambda} a^{\nu} v^{\lambda} - \Gamma^{\mu}_{\nu\lambda} b^{\nu} v^{\lambda} + \Gamma^{\mu}_{\nu\lambda} \Gamma^{\nu}_{\gamma\delta} a^{\gamma} b^{\delta} v^{\lambda} + \Gamma^{\mu}_{\nu\lambda} b^{\nu} \Gamma^{\lambda}_{\alpha\beta} a^{\alpha} v^{\beta} - a^{\alpha} (\mathrm{slope}_{\alpha} \Gamma^{\mu}_{\nu\lambda}) b^{\nu} v^{\lambda} + ..$$

Now for the inverse of the path back: swap a and b.

$$\Delta_{PSR} v^{\mu} = -\Gamma^{\mu}_{\nu\lambda} b^{\nu} v^{\lambda} - \Gamma^{\mu}_{\nu\lambda} a^{\nu} v^{\lambda} + \Gamma^{\mu}_{\nu\lambda} \Gamma^{\nu}_{\gamma\delta} b^{\gamma} a^{\delta} v^{\lambda} + \Gamma^{\mu}_{\nu\lambda} a^{\nu} \Gamma^{\lambda}_{\alpha\beta} b^{\alpha} v^{\beta} - b^{\alpha} (\mathrm{slope}_{\alpha} \Gamma^{\mu}_{\nu\lambda}) a^{\nu} v^{\lambda} + ..$$

Subtracting the latter from the former we get a lot of cancellation and we can factor the abv part out of every term.

$$\Delta_{PQRSP} v^{\mu} = -R^{\mu}_{\lambda\alpha\beta} a^{\alpha} b^{\beta} v^{\lambda}$$

where

$$R^{\mu}_{\lambda\alpha\beta} \stackrel{\mathrm{def}}{=} \mathrm{slope}_{\alpha} \Gamma^{\mu}_{\beta\lambda} - \mathrm{slope}_{\beta} \Gamma^{\mu}_{\alpha\lambda} + \Gamma^{\mu}_{\alpha\nu} \Gamma^{\nu}_{\beta\lambda} - \Gamma^{\mu}_{\beta\nu} \Gamma^{\nu}_{\alpha\lambda}$$

is the *Riemann curvature tensor*. (Note the grouping of λ and μ: this is the convention.)

Understanding this tensor is not immediate, so let's try examples.

First, **plane polar coordinates**, with metric $g_{\alpha\beta}$ and affine connection $\Gamma^{\alpha}_{\beta\gamma}$

g_dd(α	β	g)
r	r	1
θ	θ	r^2

Γ_udd(α	β	γ	Γ)
θ	θ	r	$1/r$
θ	r	θ	$1/r$
r	θ	θ	$-r$

The slopes of the affine connection are with respect to r. The first three lines below are calculated and then rearranged into the second three lines for the second term in the curvature tensor.

slope$_\alpha\Gamma^{\mu}_{\lambda\beta}$($\alpha$	μ	λ	β	slopeΓ)
	r	θ	θ	r	$-1/r^2$
	r	θ	r	θ	$-1/r^2$
	r	r	θ	θ	-1
$-$slope$_\beta\Gamma^{\mu}_{\lambda\alpha}$	r	θ	θ	r	$1/r^2$
	θ	θ	r	r	$1/r^2$
	θ	r	θ	r	-1

giving four terms, since the first of each triple cancel each other.

Here are the two products of affine connections with each other. The first is calculated (five lines) and the second is a rearrangement, swapping α with β and not retaining the join index, ν.

$\Gamma^{\mu}_{\nu\alpha}\Gamma^{\nu}_{\lambda\beta}$($\mu$	ν	α	λ	β	ΓΓ)
	θ	θ	r	θ	r	$1/r^2$
	θ	θ	r	r	θ	$1/r^2$
	r	θ	θ	θ	r	-1
	r	θ	θ	r	θ	-1
	θ	r	θ	θ	θ	r^2
$-\Gamma^{\mu}_{\nu\beta}\Gamma^{\nu}_{\lambda\alpha}$	θ		r	θ	r	$-1/r^2$
	θ		θ	r	r	$-1/r^2$
	r		r	θ	θ	1
	r		θ	r	θ	1
	θ		θ	θ	θ	$-r^2$

Here there are three pairs of cancellations—the first, fourth and fifth in each of the five-row sets—leaving again four terms.

Finally, combining the two sets of four terms—one set from the slope calculations and one set from the product—we find they all cancel. So the excess angle will be zero in plane polar coordinates. That is exactly what should happen, since the space is flat. It has zero curvature.

The second example is curved: the **surface of a sphere** with metric $g_{\alpha\beta}$ and affine connection $\Gamma^{\alpha}_{\beta\gamma}$.

$g_{\rm dd}$(α	β	g)
θ	θ	r^2
ϕ	ϕ	r^2s^2

$\Gamma_{\rm udd}$(α	β	γ	Γ)
θ	ϕ	ϕ	$-sc$
ϕ	θ	ϕ	c/s
ϕ	ϕ	θ	c/s

where $s = \sin\theta$ and $c = \cos\theta$. So the slopes are with respect to θ. We follow the same procedure as for plane polar.

$\texttt{slope}_\alpha\Gamma^\mu_{\lambda\beta}(\alpha$	μ	λ	β	$\texttt{slope}\Gamma$ $)$
	θ	θ	ϕ	ϕ	s^2-c^2
	θ	ϕ	θ	ϕ	$-1/s^2$
	θ	ϕ	ϕ	θ	$-1/s^2$
$-\texttt{slope}_\beta\Gamma^\mu_{\lambda\alpha}$	ϕ	θ	ϕ	θ	c^2-s^2
	ϕ	ϕ	θ	θ	$1/s^2$
	θ	ϕ	ϕ	θ	$1/s^2$

Cancellations again reduce these six rows to four.

$\Gamma^\mu_{\nu\alpha}\Gamma^\nu_{\lambda\beta}(\mu$	ν	α	λ	β	$\Gamma\Gamma$ $)$
	θ	ϕ	ϕ	θ	ϕ	$-c^2$
	θ	ϕ	ϕ	ϕ	θ	$-c^2$
	ϕ	ϕ	θ	θ	ϕ	c^2/s^2
	ϕ	ϕ	θ	ϕ	θ	c^2/s^2
	ϕ	θ	ϕ	ϕ	ϕ	$-c^2$
$-\Gamma^\mu_{\nu\beta}\Gamma^\nu_{\lambda\alpha}$	θ		ϕ	θ	ϕ	c^2
	θ		θ	ϕ	ϕ	c^2
	ϕ		ϕ	θ	θ	$-c^2/s^2$
	ϕ		θ	ϕ	θ	$-c^2/s^2$
	ϕ		ϕ	ϕ	ϕ	c^2

Three pairs of cancellations reduce these also to four rows.

But the four rows from the slopes combine with the four rows from the products to produce a non-zero curvature tensor.

$\texttt{R}^\mu_{\lambda\alpha\beta}(\mu$	λ	α	β	$\texttt{R}$ $)$
θ	ϕ	θ	ϕ	s^2
ϕ	θ	θ	ϕ	-1
θ	ϕ	ϕ	θ	$-s^2$
ϕ	θ	ϕ	θ	1

Let's look at

$$\Delta v^\mu = -R^\mu_{\lambda\alpha\beta}a^\alpha b^\beta v^\lambda$$

directly on the sphere for a small parallelogram of sides

$a\_\mathrm{u}(\alpha$	a $)$
ϕ	$\Delta\phi$

$b\_\mathrm{u}(\beta$	b $)$
θ	$\Delta\theta$

The result is, with θ as the first and ϕ as the second index,

$$\Delta v = -\begin{pmatrix} & -s^2 \\ 1 & \end{pmatrix} v\Delta\theta\Delta\phi$$

Δv is thus orthogonal to v, and it turns v counterclockwise, as you can see in the two base cases

$$\begin{matrix} v \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{matrix} \quad \begin{matrix} \Delta v \\ \begin{pmatrix} 0 \\ -1 \end{pmatrix} \end{matrix} \Delta\theta\Delta\phi \qquad\qquad \begin{matrix} v \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{matrix} \quad \begin{matrix} \Delta v \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{matrix} s^2\Delta\theta\Delta\phi$$

The factor of s^2 reflects a) a factor $s = \sin\theta$ reducing the area $\Delta\theta\Delta\phi$ as θ moves from equator to (either) pole, and b) a factor s multiplying the ϕ component of the test vector v to give its length.

Something neat happens if we lower the one up-index

$g_{\mu\delta}R^{\delta}_{\lambda\alpha\beta}(\mu$	δ	λ	α	β	gR	$)$
θ	θ	ϕ	θ	ϕ	r^2s^2	
ϕ	ϕ	θ	θ	ϕ	$-r^2s^2$	
θ	θ	ϕ	ϕ	θ	$-r^2s^2$	
ϕ	ϕ	θ	ϕ	θ	r^2s^2	

If we remove the column for δ, the join index, we see that $R_{\mu\lambda\alpha\beta}$ has some *symmetries*

$R_{\mu\lambda\alpha\beta}(\mu$	λ	α	β	gR	$)$
θ	ϕ	θ	ϕ	r^2s^2	
ϕ	θ	θ	ϕ	$-r^2s^2$	
θ	ϕ	ϕ	θ	$-r^2s^2$	
ϕ	θ	ϕ	θ	r^2s^2	

namely

$$\begin{aligned} R_{\mu\lambda\alpha\beta} &= -R_{\lambda\mu\alpha\beta} \\ R_{\mu\lambda\alpha\beta} &= -R_{\mu\lambda\beta\alpha} \\ R_{\mu\lambda\alpha\beta} &= R_{\alpha\beta\mu\lambda} \end{aligned}$$

Let's see if we can confirm these symmetries.

We must anticipate two advanced results. The first is that, by the meaning of tensors, a tensor relationship which holds in one coordinate system also holds in any coordinate system. Excursion *Absolute slope and curvature* shows that the curvature tensor is indeed a tensor in this definition.

The second advanced result is that, in a small enough region, a curved space is locally flat, and that in this case, we can find coordinates in which the slope of the metric vanishes—although the slope of this slope may not. See Excursion *Local flatness*.

So we look at the curvature tensor in such a coordinate system. The affine connection, being a combination of slopes of the metric, vanishes, but its slope does not.

$$\begin{aligned} R^{\mu}_{\lambda\alpha\beta} &= \partial_\alpha\Gamma^{\mu}_{\lambda\beta} - \partial_\beta\Gamma^{\mu}_{\lambda\alpha} \\ \Gamma^{\mu}_{\lambda\beta} &= \frac{1}{2}g^{\mu\delta}(\partial_\beta g_{\delta\lambda} + \partial_\lambda g_{\delta\beta} - \partial_\delta g_{\lambda\beta}) \\ \partial_\alpha\Gamma^{\mu}_{\lambda\beta} &= \frac{1}{2}g^{\mu\delta}(\partial_\alpha\partial_\beta g_{\delta\lambda} + \partial_\alpha\partial_\lambda g_{\delta\beta} - \partial_\alpha\partial_\delta g_{\lambda\beta}) \\ \partial_\beta\Gamma^{\mu}_{\lambda\alpha} &= \frac{1}{2}g^{\mu\delta}(\partial_\beta\partial_\alpha g_{\delta\lambda} + \partial_\beta\partial_\lambda g_{\delta\alpha} - \partial_\beta\partial_\delta g_{\lambda\alpha}) \\ R^{\mu}_{\lambda\alpha\beta} &= \frac{1}{2}g^{\mu\delta}(\partial_\alpha\partial_\lambda g_{\delta\beta} - \partial_\beta\partial_\lambda g_{\delta\alpha} + \partial_\beta\partial_\delta g_{\lambda\alpha} - \partial_\alpha\partial_\delta g_{\lambda\beta}) \\ R_{\mu\lambda\alpha\beta} &= \frac{1}{2}(\partial_\alpha\partial_\lambda g_{\mu\beta} - \partial_\beta\partial_\lambda g_{\mu\alpha} + \partial_\beta\partial_\mu g_{\lambda\alpha} - \partial_\alpha\partial_\mu g_{\lambda\beta}) \\ &= A - B + C - D \end{aligned}$$

In the above, the first line is the short, locally flat, version of the curvature tensor; the second line is the definition of the affine connection; the third line is the slope of the affine connection, given that $g^{\mu\delta}$ has zero slope; the fourth line swaps α and β in the third line to give the term to be subtracted; the fifth line is the difference of the third and fourth lines; the sixth line gets rid of the leading g term by lowering the first index; and the seventh line abbreviates each of the respective terms so that we can use them in the following rearrangements of the indices.

$$R_{\mu\lambda\alpha\beta} = A - B + C - D$$

$$\begin{aligned} R_{\lambda\mu\alpha\beta} &= D - C + B - A \\ R_{\mu\lambda\beta\alpha} &= B - A + D - C \\ R_{\alpha\beta\mu\lambda} &= C - B + A - D \end{aligned}$$

which proves the symmetries.

There is another symmetry which is not obvious from the 2-dimensional example of the surface of a sphere. This is the cyclic symmetry

$$\begin{aligned} R_{\mu\lambda\alpha\beta} + R_{\mu\beta\lambda\alpha} + R_{\mu\alpha\beta\lambda} &= A - B + C - D + E - A + F - C + B - E + D - F \\ &= 0 \end{aligned}$$

where A, B, C and D are as before, taking advantage of the symmetry of $\partial_\alpha\partial_\beta$ and of $g_{\alpha\beta}$, and where E and F are additional terms which, as we see, cancel.

The cyclic symmetry reduces to the swap symmetries we saw first, unless all the indices are different. Thus the cyclic symmetry says nothing for a space under four dimensions.

Back in two dimensions, the single independent number in the Riemann tensor is simply related to the **Gaussian curvature**,

$$K = -\frac{R_{1212}}{g}$$

Karl Friedrich Gauss' "Theorema Egregium" (1827?) established the curvature (that is, the "inner" curvature) of a 2D surface in terms of the metric tensor and its derivatives—that is by knowing only measurements which creatures living *in* that surface can make. K relates the excess angle to the area bounded by the curve traversed, in the limit that the area becomes arbitrarily small.

$$\Delta\theta = K\Delta\text{area}$$

Bernhard Riemann subsequently generalized this result to any number of dimensions.

On a sphere we can talk about large areas because the curvature is constant, i.e., independent of location on the sphere. (The symmetry of the sphere should make this apparent: what is there to differentiate any particular location from any other?) So the excess angle for the 1/8 of the sphere at the beginning of this Note is $\pi/2$ and its area is

$$\frac{4\pi r^2}{8} = \frac{\pi}{2}r^2$$

so $K = 1/r^2$.

The derivation of K from $R_{\mu\lambda\alpha\beta}$ usually goes from the full expression for R above to the even more cumbersome expression Gauss derived for K. Here is a simpler derivation.

For a small angle formed by changing a vector $\mathbf{v}$ by $\Delta\mathbf{v}$ the definition of angle in radians allows us to say

$$\Delta\theta = \frac{|\ \mathbf{\Delta v}\ |}{|\ \mathbf{v}\ |}$$

We know that

$$\Delta v^\mu = -R^\mu_{\lambda\alpha\beta}a^\alpha b^\beta v^\lambda$$

and can work with this to get the infinitesimal area and the R_{1212}.

In two dimensions, indices take on only two values which we'll call 1 and 2. And the area of a parallelogram of sides $\mathbf{a}$ and $\mathbf{b}$ is, using Cartesian coordinates

$$\begin{aligned} a_x b_y - a_y b_x &= |\ a\ ||\ b\ |\ (\cos\phi\sin(\phi+\theta) - \cos(\phi+\theta)\sin\phi) \\ &= |\ a\ ||\ b\ |\ \sin(\phi+\theta-\phi) \\ &= |\ a\ ||\ b\ |\ \sin\theta \end{aligned}$$

where **a** is inclined angle ϕ from the x-axis and **b** is inclined angle θ from **a**. The final result is the area between two vectors of the magnitudes given and separated by angle θ.

We'll actually need $a^1b^2 - a^2b^1$ but the transformation from (x, y) to non-orthonormal coordinates is familiar (e.g., Note 8, Note 11) and we can invent a transformation matrix to write

$$\begin{pmatrix} a^1 \\ a^2 \end{pmatrix} = \begin{pmatrix} p & r \\ q & s \end{pmatrix}^{-1} \begin{pmatrix} a_x \\ a_y \end{pmatrix} = \frac{1}{\sqrt{g}} \begin{pmatrix} s & -r \\ -q & p \end{pmatrix} \begin{pmatrix} a_x \\ a_y \end{pmatrix}$$

where $\sqrt{g} = ps - qr$, and similarly for $\boldsymbol{b}$.

So

$$\begin{aligned} a^1b^2 - a^2b^1 &= \frac{1}{g}[(sa_x - ra_y)(-qb_x + pb_y) - (-qa_x + pa_y)(sb_x - rb_y)] \\ &= \frac{1}{g}[(ps - qr)a_xb_y - (ps - qr)a_yb_x] \\ &= \frac{1}{\sqrt{g}}(a_xb_y - a_yb_x) \end{aligned}$$

We can always find a coordinate system in which the metric tensor is diagonal, so we can assume this without losing generality. (Make the basis orthogonal: choose a coordinate transformation which is an orthogonal matrix, so its inverse is its transpose. This is a diagonalizing transformation.)

Putting this all together

$$\begin{aligned} \Delta v^\mu &= -R^\mu_{\lambda\alpha\beta} a^\alpha b^\beta v^\lambda \\ &= -R^\mu_{\lambda\alpha\beta}(a^\alpha b^\beta - a^\beta b^\alpha)v^\lambda \qquad \text{restricting } \alpha < \beta \\ &= -g^{\mu\nu} R_{\nu\lambda\alpha\beta}(a^\alpha b^\beta - a^\beta b^\alpha)v^\lambda \\ &= -g^{\mu\mu} R_{\mu\lambda\alpha\beta}(a^\alpha b^\beta - a^\beta b^\alpha)v^\lambda \\ &= -\begin{pmatrix} g^{11}R_{1212}(a^1b^2 - a^2b^1)v^2 \\ g^{22}R_{2112}(a^1b^2 - a^2b^1)v^1 \end{pmatrix} \\ &= \begin{pmatrix} g^{11}v^2 \\ -g^{22}v^1 \end{pmatrix}(-R_{1212})(a^1b^2 - a^2b^1) \\ &= \begin{pmatrix} g^{11}v^2 \\ -g^{22}v^1 \end{pmatrix}(-R_{1212})(a_xb_y - a_yb_x)/\sqrt{g} \\ &= \mathcal{V}(-R_{1212})(a_xb_y - a_yb_x)/\sqrt{g} \end{aligned}$$

In the second line we restrict $\alpha < \beta$ and use the antisymmetry of $R^\mu_{\lambda\alpha\beta}$ on the last two indices. In the third line we change $R^\mu_{\lambda\alpha\beta}$ to $R_{\mu\lambda\alpha\beta}$. In the fourth line we use our assumption that g^{jk} is diagonal. In the fifth line we restrict to the two indices 1 and 2 and to the four nonzero values of $R_{\mu\lambda\alpha\beta}$. The seventh line rewrites the sixth in terms of the area in orthonormal coordinates. The last line gives the vector part a name $\mathcal{V}$ to use in the following.

The norm of this is just the norm of the vector part $\mathcal{V}$ times the factors outside (which we see are coming pretty close to the product of R_{1212} and the area over g). We start with the square of the norm.

$$\begin{aligned} |\,\mathcal{V}\,|^2 &= g_{jk}\mathcal{V}^j\mathcal{V}^k \\ &= g_{11}(g^{11}v^2)^2 + g_{22}(g^{22}v^1)^2 \\ &= g^{11}(v^2)^2 + g^{22}(v^1)^2 \\ &= (g_{22}(v^2)^2 + g_{11}(v^1)^2)/g \\ &= |\,\mathbf{v}\,|^2/g \end{aligned}$$

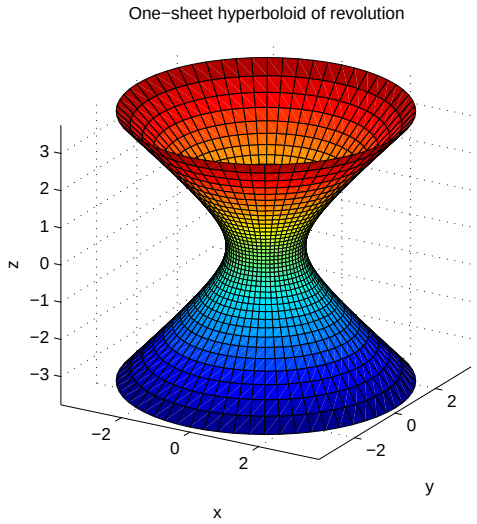
So the norm of $\mathcal{V}$ is the norm of the original vector $\mathbf{v}$ divided by $\sqrt{g}$. Note incidentally that $\mathcal{V}$ is orthogonal to $\mathbf{v}$: take the inner product $\mathbf{v}^j g_{jk} \mathcal{V}^k$. This means that the norm of the new vector $\mathbf{v} + \mathbf{\Delta v}$ is the same as that of $\mathbf{v}$.

From all this we see that the excess angle, which is the ratio of norms, is

$$\begin{aligned} \Delta\theta &= \frac{|\,\mathbf{\Delta v}\,|}{|\,\mathbf{v}\,|} \\ &= -R_{1212}(a_x b_y - a_y b_x)/g \\ &= -R_{1212}\Delta\text{area}/g \end{aligned}$$

as claimed.

18. Negative curvature. A two-dimensional surface has negative curvature if it curves oppositely in opposite dimensions. A hyperboloid of revolution is a good example.



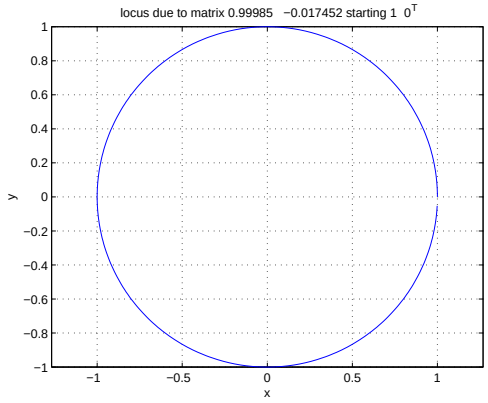
This is an x-z hyperbola rotated about the z-axis, so we should study hyperbolas first.

In Week iv we found that hyperbolas emerge from shear matrices in the same way that circles emerge from rotation matrices. Here are the results of applying a very small rotation or a very small shear to a point in space, then applying it repeatedly to the previous result.

Rotation

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$

$$c^2 + s^2 = 1$$

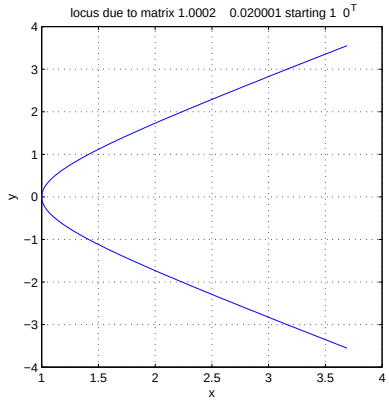


Here, $c = 0.99985$ and $s = 0.017452$:
a very small rotation.

Shear

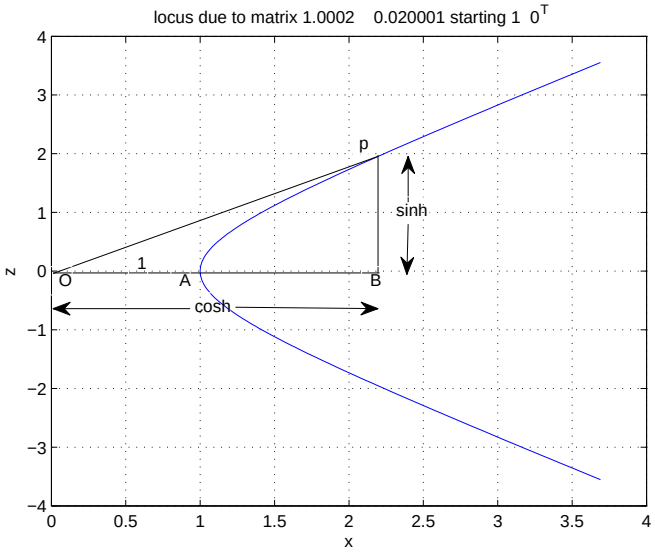
$$\begin{pmatrix} ch & sh \\ sh & ch \end{pmatrix}$$

$$ch^2 - sh^2 = 1$$



Here, $ch = 1.00020$ and $sh = 0.02000$:
a very small shear.

I've used symbols *ch* and *sh* respectively for "hyperbolic cosine" and "hyperbolic sine"[1] because they give the x and y coordinates, respectively, just as do cos and sin for circles.



The discussion in Excursion *What functions for ch, sh?* argues for the following parametrization.

$$\begin{aligned} ch &= \cosh p = \frac{1}{2}(e^p + e^{-p}) \\ sh &= \sinh p = \frac{1}{2}(e^p - e^{-p}) \end{aligned}$$

[1] I preferred to use the Hebrew *khaf* and *shin* for these but LaTeX does not support Hebrew other than ℵ.

Now three dimensions: the hyperboloid. The extension is easy: the only change is to rotate in the x-y plane.

$$\begin{aligned} x &= \cosh(p)\cos(\theta) \\ y &= \cosh(p)\sin(\theta) \\ z &= \sinh(p) \end{aligned}$$

We are interested in the metric and hence the affine connection and hence the curvature tensor.

The incremental distance as we change θ obeys

$$(\Delta s)^2 = \cosh(p)^2(\Delta\theta)^2$$

because $\cosh(p) = ch$ is the radius of that section of the rotated hyperbola.

Similarly, but less obviously,

$$(\Delta s)^2 = \cosh(2p)(\Delta p)^2$$

because the vertical distance squared (i.e., dependent on p but not on θ) from the origin to any point $(r, z) = (\sqrt{x^2+y^2}, z)$ on the hyperboloid—that is, the radius of the arc from p to $p+\Delta p$—is

$$\cosh(p)^2 + \sinh(p)^2 = \cosh(2p)$$

So the metric is

$$g_{\mathrm{dd}} = \begin{pmatrix} \cosh(2p) & \\ & \cosh(p)^2 \end{pmatrix}$$

We can follow the now well-practiced calculation to find the affine connection.

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\beta}(\mathrm{slope}_\nu\ g_{\mu\beta} + \mathrm{slope}_\mu\ g_{\nu\beta} - \mathrm{slope}_\beta\ g_{\mu\nu})$$

$\Gamma_{\mathrm{udd}}(\lambda$	μ	ν	Γ $)$
p	p	p	th_2
p	θ	θ	$-\frac{1}{2}th_2$
θ	p	θ	th
θ	θ	p	th

where $th = \tanh(p)$ and $th_2 = \tanh(2p)$.

The next step is the Riemann curvature tensor. I'll find $-R_{\mathrm{dddd}}$ because the 2-dimensional curvature is $-R_{1212}/g$.

$$-R_{\mu\nu\alpha\beta} = g_{\mu\lambda}(\mathrm{slope}_\alpha\ \Gamma^\lambda_{\nu\beta} - \mathrm{slope}_\beta\ \Gamma^\lambda_{\nu\alpha} + \Gamma^\lambda_{\gamma\alpha}\Gamma^\gamma_{\nu\beta} - \Gamma^\lambda_{\gamma\beta}\Gamma^\gamma_{\nu\alpha})$$

$-R_{\mathrm{dddd}}(\mu$	ν	α	β	R $)$
p	θ	p	θ	$-\frac{1}{2}th\ th_2 - \frac{1}{ch_2}$
p	θ	θ	p	$\frac{1}{2}th\ th_2 + \frac{1}{ch_2}$
θ	p	p	θ	$-\frac{1}{2}th\ th_2 + 1$
θ	p	θ	p	$\frac{1}{2}th\ th_2 - 1$

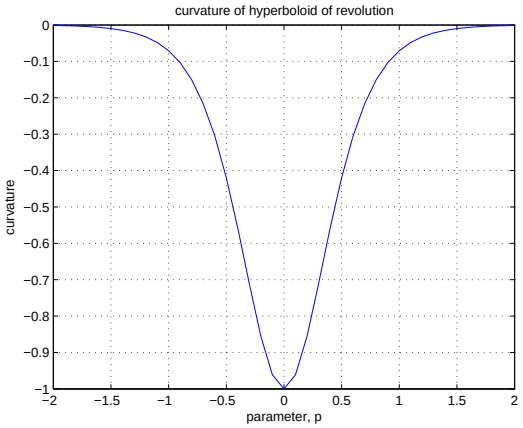
making it seem for a moment that a symmetry is broken. But

$$-\frac{1}{2}th\ th_2 - \frac{1}{ch_2} = -\frac{ch^2}{ch_2} = \frac{1}{2}th\ th_2 - 1$$

Furthermore it is easy to see that the expression for $-R_{1212}$ is always negative, so the Gaussian curvature $K = -R_{1212}/g$ is also always negative.

$$\begin{aligned} K &= -\frac{R_{1212}}{g} \\ &= -\frac{\frac{ch^2}{ch_2}}{ch_2 ch^2} \\ &= -\frac{1}{{ch_2}^2} \end{aligned}$$

We also see that the curvature is not constant.



The (negative) curvature is maximum at $p = 0$, the most bent part of the hyperbola and hyperboloid. It approaches zero asymptotically as the hyperboloid approaches a 45-degree cone, which can be made from a flat sheet of paper and so has zero curvature.

There is no surface embedded in our usual Euclidean 3-dimensional space which has *constant* negative curvature. But we can find an esoteric such surface, a "pseudo-sphere", if we look at yet another way to represent *ch* and *sh* for a hyperbola.

parameter	ch	sh	ch^{-1}	sh^{-1}
$-i\frac{\pi}{4} < p < i\frac{\pi}{4}$	$\cos(p)$	$i\sin(p)$	$\mathrm{acos}(ch)$	$\mathrm{asin}(-i\ sh)$

Note, for example, that

$$ch^2 + sh^2 = \cos^2 p - \sin^2 p = \cos(2p) = ch_2$$

This is based on a central result on two-dimensional numbers (and i is just the "number"—imagine that!—representing the 90-degree rotation)

$$\begin{aligned} e^{i\theta} &= \cos\theta + i\sin\theta \\ e^{p} &= \cosh p + \sinh p \end{aligned}$$

where I've given also the corresponding relationship on hyperbolic cosine and sine.

So we've just set parameter $p = i\theta$. But this has the effect of making one axis "imaginary": instead of $(x, z) = (\cos p, i\sin p)$ we can use $(x, \zeta) = (\cos p, \sin p)$ where $\zeta = -iz$.

Let's explore the surface of a sphere in x-y-ζ coordinates, i.e.,

$$\begin{aligned} x &= \cos\theta\cos\phi \\ y &= \cos\theta\sin\phi \\ \zeta &= \sin\theta \end{aligned}$$

In the metric we can retain $r\sin\theta$ because it gives a horizontal direction, i.e., parallel to the x-y plane. But we must change the other direction so that, where $r^2 = x^2 + y^2$,

$$r^2 + (i\zeta)^2 = r^2 - \zeta^2$$

is a distance. So the metric has a sign change in the (θ,θ) entry (which we've been writing first, before (ϕ,ϕ), for spheres).

$$g\_\mathrm{dd} = \begin{pmatrix} -r^2 & \\ & r^2 s^2 \end{pmatrix}$$

reverting to the abbreviations s for $\sin\theta$ (and c for $\cos\theta$).

Following our usual calculations, we get affine connection and curvature tensor for this pseudo-sphere with some sign changes from the results for the sphere.

$\Gamma\_\mathrm{udd}(\lambda$	μ	ν	Γ	)
θ	ϕ	ϕ	cs	
ϕ	θ	ϕ	c/s	
ϕ	ϕ	θ	c/s	

$-R\_\mathrm{uddd}(\mu$	ν	α	β	R	)
θ	ϕ	θ	ϕ	$-s^2$	
θ	ϕ	ϕ	θ	s^2	
ϕ	θ	θ	ϕ	-1	
ϕ	θ	ϕ	θ	1	

So

$-R\_\mathrm{dddd}(\mu$	ν	α	β	R	)
θ	ϕ	θ	ϕ	$r^2 s^2$	
θ	ϕ	ϕ	θ	$-r^2 s^2$	
ϕ	θ	θ	ϕ	$-r^2 s^2$	
ϕ	θ	ϕ	θ	$r^2 s^2$	

And

$$K = -\frac{R_{1212}}{g} = \frac{r^2 s^2}{-r^4 s^2} = -\frac{1}{r^2}$$

a negative constant over the whole surface.

19. The Ricci tensor. In Excursion *Number of components of the curvature tensor* we show that the Riemann curvature tensor has $d^2(d^2-1)/12$ independent components in d dimensions. It is of interest to compare this with the Δ_d independent components of a symmetric d-by-d matrix, e.g., a symmetric d-dimensional tensor of two indices.

d	1	2	3	4
$d^2(d^2-1)/12$	0	1	6	20
Δ_d	1	3	6	10

So, for up to three dimensions, a 2-index tensor can in principle contain all the information of the full Riemann curvature tensor.

Such a 2-index tensor is the *Ricci tensor*

$$R_{\lambda\beta} = R^{\mu}_{\lambda\mu\beta} = g^{\mu\delta} R_{\delta\lambda\mu\beta}$$

and we can see that it is symmetric by the symmetries of $g^{\mu\delta} = g^{\delta\mu}$ and $R_{\delta\lambda\alpha\beta} = R_{\alpha\beta\delta\lambda}$

The Ricci tensor is arrived at by *contracting* two indices of the Riemann tensor, one up and one down so that Ricci is a tensor too.

If we contract again we get the *curvature scalar*

$$R = g^{\lambda\beta}R_{\lambda\beta} = g^{\lambda\beta}g^{\mu\delta}R_{\delta\lambda\mu\beta}$$

We can go the other way and write the Riemann tensor explicitly in terms of the curvature scalar R in $d = 2$ dimensions and in terms of R and the Ricci tensor in $d \leq 3$ dimensions.

2D. To start, you should check that the symmetries of $R_{\mu\lambda\alpha\beta}$ require, in two dimensions when $\mu, \lambda, \alpha, \beta \epsilon \{1, 2\}$, that the only nonzero components out of the 16 possible will be

$$R_{1212} = -R_{1221} = -R_{2112} = R_{2121}$$

Then you can show

$$\begin{aligned} R &= g^{\lambda\beta}g^{\mu\delta}R_{\delta\lambda\mu\beta} \\ &= (g^{11}g^{22} - g^{12}g^{21})R_{1212} + (g^{22}g^{11} - g^{21}g^{12})R_{2121} \\ &= \frac{2}{g}R_{1212} \end{aligned}$$

where $g = g_{11}g_{22} - (g_{12})^2$ is the determinant of the inverse of g_{uu}, $g = 1/(g^{11}g^{22} - (g^{12})^2)$.

Still in two dimensions, you can also show that

$$R_{\mu\lambda\alpha\beta} = (g_{\mu\alpha}g_{\lambda\beta} - g_{\mu\beta}g_{\lambda\alpha})\frac{R_{1212}}{g}$$

Here, for this and the previous result for R, are the nonzero components of the relations $R_{\mu\lambda\alpha\beta}$ and $g_{\mu\alpha}g_{\lambda\beta} - g_{\mu\beta}g_{\lambda\alpha}$

μ	λ	α	β	R	$gg-$
1	2	1	2	r	g
1	2	2	1	$-r$	$-g$
2	1	1	2	$-r$	$-g$
2	1	2	1	r	g

where $r = R_{1212}$ and $g = g_{11}g_{22} - g_{12}g_{21}$.

So we have the Riemann curvature in two dimensions in terms of the curvature scalar.

$$R_{\mu\lambda\alpha\beta} = \frac{1}{2}(g_{\mu\alpha}g_{\lambda\beta} - g_{\mu\beta}g_{\lambda\alpha})R$$

Incidentally

$$\begin{aligned} R_{\lambda\beta} &= g^{\mu\delta}(g_{\delta\mu}g_{\lambda\beta} - g_{\delta\beta}g_{\lambda\mu})\frac{R_{1212}}{g} \\ &= (\delta^{\mu}_{\mu}g_{\lambda\beta} - \delta^{\mu}_{\beta}g_{\lambda\mu})\frac{R_{1212}}{g} \\ &= (2g_{\lambda\beta} - g_{\lambda\beta})\frac{R_{1212}}{g} \\ &= g_{\lambda\beta}\frac{R_{1212}}{g} = g_{\lambda\beta}\frac{R}{2} \end{aligned}$$

3D+. In three dimensions we need the Ricci tensor as well as R to give $R_{\mu\lambda\alpha\beta}$ completely.

$$R_{\mu\lambda\alpha\beta} = g_{\mu\alpha}R_{\lambda\beta} - g_{\mu\beta}R_{\lambda\alpha} - g_{\lambda\alpha}R_{\mu\beta} + g_{\lambda\beta}R_{\mu\alpha} - \frac{1}{2}(g_{\mu\alpha}g_{\lambda\beta} - g_{\mu\beta}g_{\lambda\alpha})R$$

In three or more dimensions d this generalizes to

$$R_{\mu\lambda\alpha\beta} = \frac{1}{d-2}(g_{\mu\alpha}R_{\lambda\beta} - g_{\mu\beta}R_{\lambda\alpha} - g_{\lambda\alpha}R_{\mu\beta} + g_{\lambda\beta}R_{\mu\alpha}) - \frac{1}{(d-1)(d-2)}(g_{\mu\alpha}g_{\lambda\beta} - g_{\mu\beta}g_{\lambda\alpha})R + C_{\mu\lambda\alpha\beta}$$

where $C_{\mu\lambda\alpha\beta}$ is the *Weyl tensor*, which in four or more dimensions, picks up all the leftovers not accounted for by the Ricci tensor and the scalar R.

Excursion *Paraboloid* part (e) illustrates a use of the Ricci tensor.

20. More protor calculator. Building on Note 5 and the Excursions implementing slope() and `joinred()`, we have one major function to add to the protor calculator,

```
function(conn,curv_uddd,curv_dddd) = metr2curv(metr,invmetr,indices)
```

Fairly self-evidently, `metr2curv()` computes the Riemann curvature tensor (`curv_uddd` and `curv_dddd` forms) via the affine connection (`conn`) given the metric tensor (`metr` and its inverse `invmetr`). It requires `metr` and `invmetr` in protor form, and a list of symbolic indices for the `slope()` operation.

Here is a sample invocation for the three-dimensional paraboloid of Excursion *Paraboloid* (part (e)).

```
>> syms a r theta phi
>> gparab3 = [r,r,1/(1-a/r);theta,theta,r^2;phi,phi,r^2*sin(theta)^2]
gparab3 =
[     r,     r,     -1/(a/r - 1)]
[ theta, theta,              r^2]
[   phi,   phi, r^2*sin(theta)^2]
>> invgparab3 = [r,r,1-a/r;theta,theta,1/r^2;phi,phi,1/r^2/sin(theta)^2]
invgparab3 =
[     r,     r,              1 - a/r]
[ theta, theta,                1/r^2]
[   phi,   phi, 1/(r^2*sin(theta)^2)]
>> [connParab3,curvParab3_uddd,curvParab3_dddd] = metr2curv(gparab3,invgparab3,[r,theta,phi])
connParab3 =
[   phi,   phi,     r,                       1/r]
[   phi,   phi, theta,    cos(theta)/sin(theta)]
[   phi,     r,   phi,                       1/r]
[   phi, theta,   phi,    cos(theta)/sin(theta)]
[     r,   phi,   phi, r*sin(theta)^2*(a/r - 1)]
[     r,     r,     r,      a/(2*r^2*(a/r - 1))]
[     r, theta, theta,              r*(a/r - 1)]
[ theta,   phi,   phi,   -cos(theta)*sin(theta)]
[ theta,     r, theta,                       1/r]
[ theta, theta,     r,                       1/r]
curvParab3_uddd =
[   phi,     r,     r,   phi,                    a/(2*r^3*(a/r - 1))]
[     r,   phi,     r,   phi,                 (a*sin(theta)^2)/(2*r)]
[     r, theta,     r, theta,                              a/(2*r)]
[ theta,     r,     r, theta,                    a/(2*r^3*(a/r - 1))]
[   phi, theta, theta,   phi,                                   a/r]
[ theta,   phi, theta,   phi, - sin(theta)^2 - sin(theta)^2*(a/r - 1)]
[   phi,     r,   phi,     r,                   -a/(2*r^3*(a/r - 1))]
```

```
32 [     r,   phi,   phi,     r,                 -(a*sin(theta)^2)/(2*r)]
33 [     r, theta, theta,     r,                                -a/(2*r)]
34 [ theta,     r, theta,     r,                   -a/(2*r^3*(a/r - 1))]
35 [   phi, theta,   phi, theta,                                    -a/r]
36 [ theta,   phi,   phi, theta,   sin(theta)^2 + sin(theta)^2*(a/r - 1)]
37 curvParab3_dddd =
38 [   phi,     r,   phi,     r,            -(a*sin(theta)^2)/(2*r*(a/r - 1))]
39 [   phi,     r,     r,   phi,             (a*sin(theta)^2)/(2*r*(a/r - 1))]
40 [   phi, theta,   phi, theta,                          -a*r*sin(theta)^2]
41 [   phi, theta, theta,   phi,                           a*r*sin(theta)^2]
42 [     r,   phi,   phi,     r,             (a*sin(theta)^2)/(2*r*(a/r - 1))]
43 [     r,   phi,     r,   phi,            -(a*sin(theta)^2)/(2*r*(a/r - 1))]
44 [     r, theta,     r, theta,                        -a/(2*r*(a/r - 1))]
45 [     r, theta, theta,     r,                         a/(2*r*(a/r - 1))]
46 [ theta,   phi,   phi, theta,  r^2*(sin(theta)^2 + sin(theta)^2*(a/r - 1))]
47 [ theta,   phi, theta,   phi, -r^2*(sin(theta)^2 + sin(theta)^2*(a/r - 1))]
48 [ theta,     r,     r, theta,                          a/(2*r*(a/r - 1))]
49 [ theta,     r, theta,     r,                        -a/(2*r*(a/r - 1))]
50 >> riccParab3 = contract(curvParab3_uddd,1,3)
52 riccParab3 =
52 [   phi,   phi, -sin(theta)^2 + sin(theta)^2*(a/r - 1) - (a*sin(theta)^2)/(2*r)]
53 [     r,     r,                                              -a/(r^3*(a/r - 1))]
54 [ theta, theta,                                                        -a/(2*r)]
55 >> Xparab3_ud = [r,r,sqrt(1-a/r);theta,theta,1/r;phi,phi,1/r/sin(theta)]
56 Xparab3_ud =
57 [     r,     r,  (1 - a/r)^(1/2)]
58 [ theta, theta,              1/r]
59 [   phi,   phi, 1/(r*sin(theta))]
60 >> riccParb3norm = transform(riccParab3,Xparab3_ud)
61 riccParb3norm =
62 [   phi,   phi, -(sin(theta)^2 + sin(theta)^2*(a/r - 1)
                                  - (a*sin(theta)^2)/(2*r))/(r^2*sin(theta)^2)]
63 [     r,     r,
                                                                        a/r^3]
64 [ theta, theta,
                                                                   -a/(2*r^3)]
```

In the above code, I've added line numbers for this discussion. You can compare the results with the calculations in Excursion *Paraboloid* (part(e)) as we go along.

Lines 1–11 set up the symbolic variables and the metric and its inverse.

Line 12 is the invocation of `metr2curv()` with these parameters.

Lines 13–23 show the affine connection calculated by `metr2curv()`.

Lines 24–36 show the uddd Riemann tensor and lines 37–49 show R_{dddd}, both as calculated by `metr2curv()`: to get the results of Excursion *Paraboloid*, some of these must be simplified manually, because MATLAB is sloppy about symbolic calculation.

Line 50 invokes another function `contract()` to contract Riemann R_{uddd} to R_{dd} (Ricci), and lines 51–54 show the result, in agreement with Excursion *Paraboloid*.

Lines 55–59 define the normalizing transformation matrix and line 60 uses the function `transform()`, discussed in Note 5, to take the final step in Excursion *Paraboloid*: compare with the result in lines

61–64.

The implementation of `metr2curv()` is based on other functions and so the full code is short.

```
% function [conn,curv_uddd,curv_dddd] = metr2curv(metr,invmetr,indices)  THM  140130
% Use metr2conn(), conn2curv(), matmult() to calculate Riemann_dddd from metr
% metric_dd), invmetr (inverse metric_uu) and  indices
% In metr: the metric tensor dd: 2 cols indices, 1 col values
% invmetr: the inverse metric tensor uu: 2 cols indices, 1 col values
% indices syms: (row) vector of all index values (Greek, Latin)
% Out curv: the Riemann_dddd curvature tensor
% Uses metr2conn(), conn2curv(), joinred()
function [conn,curv_uddd,curv_dddd] = metr2curv(metr,invmetr,indices)
conn = metr2conn(metr,invmetr,indices);         % display test only
% NB may need to improve MATLAB's simplification of  conn  before continuing:
curv_uddd = conn2curv(conn,indices);
syms allzeros                                   %140429 deal with flat space
if curv_uddd==allzeros
  curv_dddd = allzeros;
else
  curv_dddd = joinred(metr,2,curv_uddd,1);
end
```

The two steps are encapsulated in two new functions.

First, find the affine connection.

```
conn = metr2conn(metr,invmetr,indices);
```

This in turn requires a couple of auxiliary functions not discussed in Note 5, and we can give the full code.

```
% function conn = metr2conn(metr,invmetr,indices)           THM           140129
% Combine slopes(), nozeros(), swapcols(), mergesum(), joinmult(), scalarmult()
% and reduce() to calculate the affine connection given the metric, its inverse
% and the set of indices to differentiate with respect to.
% In metr: the metric tensor - 2 columns of indices dd, 1 column of values
% invmetr: the inverse of the metric tensor - ditto but indices uu
% indices: (row) vector of all indices involved, usde by slope()
% NB all entries in these arrays are  syms : invocation must be preceded by,
% e.g., syms theta phi r; g = [theta,theta,r^2;phi,phi,r^2*sin(theta)^2]; etc.
% Out conn: the Christoffel symbols - 3 columns of indices udd, 1 col. values
function conn = metr2conn(metr,invmetr,indices)
di_g = nozeros(slopes(metr,indices));           % display test only
preG = nozeros(mergesum(1,nozeros(...
mergesum(1,di_g,-1,swapcols(di_g,1,2))),1,swapcols(di_g,1,3)));
conn = reduce(scalarmult(1/2,joinmult(invmetr,2,preG,2)),2);
```

The two auxiliary functions for `metr2conn()` are a merge

```
% function out = mergesum(sign1,in1,sign2,in2)             THM            140127
% Merge two relations on their (commensurate) index columns, incorporating the
% given sign for the value column and summing (with signs) if match.
% In sign1, sign2, integer: the respective signs for the two value columns
% hell, it can be any numeric weight or even another sym
% in1, in2: the size#In2-ary relations of size#In1 tuples each to merge
```

```
% Out out: the merged result.
% Used by metr2conn(), conn2curv(), abslopes(), mergesum()
```

and a routine to swap indices

```
% function out = swapcols(in,col1,col2)                    THM           140127
% Swap two index columns (i.e., avoid last column) in symbolic relation  in .
% In in: sizeIn2-ary relation stored as array 1 row per tuple
% col1, integer: column to be swapped with col2
% col2, integer: column to be swapped with col1
% Out out: the relation with index columns swapped
% Used by metr2conn(), conn2curv(), abslopes(), mergesum()
% Related but not reduced to movecols(), permcols()
```

These are straightforward to implement except that, for `mergesum()`, MATLAB precludes inequality comparisons on symbolic variables so I could not use merge logic but had to find matches first and then append unmatched tuples from each operand.

Here is a sample run of `mergesum()`, simple enough to show how it needs to be built.

```
1 >> mergesum(1,[1,1,phi;2,2,theta],2,[1,1,theta;2,1,phi])
2 ans =
3 [ 1, 1, phi + 2*theta]
4 [ 2, 2,         theta]
5 [ 2, 1,         2*phi]
```

Here is one for `swapcols()` showing how the specifed columns are simply interchanged.

```
 1 digk1k2 =
 2 [  xi,  xi,  xi,  2*k1^2*xi]
 3 [  xi,  xi, eta,  eta*k1*k2]
 4 [  xi, eta,  xi,  eta*k1*k2]
 5 [ eta,  xi, eta,   k1*k2*xi]
 6 [ eta, eta,  xi,   k1*k2*xi]
 7 [ eta, eta, eta, 2*eta*k2^2]
 8 >> digk1k2_12 = swapcols(digk1k2,1,2)
 9 digk1k2_12 =
10 [  xi,  xi,  xi,  2*k1^2*xi]
11 [  xi,  xi, eta,  eta*k1*k2]
12 [ eta,  xi,  xi,  eta*k1*k2]
13 [  xi, eta, eta,   k1*k2*xi]
14 [ eta, eta,  xi,   k1*k2*xi]
15 [ eta, eta, eta, 2*eta*k2^2]
```

The second step of `metr2curv()` uses the affine connection to get the Riemann curvature.

```
function curv = conn2curv(conn,indices)
```

uses `slopes()`, `nozeros()`, `swapcols()`, `mergesum()`, `joinmult()`, `reduce()`, and `permcols()` in straightforward ways to generate R_{uddd} from Γ_{udd}. In addition it needs

```
% function out = permcols(in,perm)                          THM           140128
% Permute index columns of relation  in  according to perm
% e.g., alpha,mu,nu,beta->mu,nu,alpha,beta uses [2,3,1,4]
```

```
% In in: relation to be permuted
% perm integer (row) vector: the permutation
% Out out: the permuted relation
% Used by conn2curv(), bianchi()
% Related but not reducible to movecols(), swapcols()
```

to permute columns, as in

```
 1 gamma =
 2 [ eta, eta, eta, eta*k2^2*(eta^2*k2^2 + 1) + eta*k1^2*k2^2*xi^2]
 3 [ eta,  xi,  xi,  eta*k2*k1^3*xi^2 + eta*k2*(eta^2*k2^2 + 1)*k1]
 4 [  xi, eta, eta,    k1*xi*eta^2*k2^3 + k1*xi*(k1^2*xi^2 + 1)*k2]
 5 [  xi,  xi,  xi,   k1^2*xi*(k1^2*xi^2 + 1) + eta^2*k1^2*k2^2*xi]
 6 >> permcols(gamma,[2,3,1])   % NB for this e.g., same as [3,2,1]
 7 ans =
 8 [ eta, eta, eta, eta*k2^2*(eta^2*k2^2 + 1) + eta*k1^2*k2^2*xi^2]
 9 [  xi,  xi, eta,  eta*k2*k1^3*xi^2 + eta*k2*(eta^2*k2^2 + 1)*k1]
10 [ eta, eta,  xi,    k1*xi*eta^2*k2^3 + k1*xi*(k1^2*xi^2 + 1)*k2]
11 [  xi,  xi,  xi,   k1^2*xi*(k1^2*xi^2 + 1) + eta^2*k1^2*k2^2*xi]
```

With these hints you should be able to implement the two new functions, `metr2curv()` and `contract()` and their auxiliaries.

(A further addition to the protor calculator is given in Note 30 (Part II): `curv2einstein()` calculates the "Einstein tensor" given the curvature and metric tensors.)

21. Summary.

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

This long Part evolves the mathematics underlying general relativity, Einstein's theory of gravity. It works from the notion of fields (in physics) to how to account for the (mathematical) coordinate system used to describe a field, to the notions of divergence and curl with the former illustrated by a discussion of classical gravity. It stretches the discussion of coordinate systems to include non-orthonormal coordinates (using the gypsum crystal as example) and introduces the *metric*. It stretches coordinates further to consider coordinate systems which themselves depend on the coordinates (e.g., polar coordinates) and introduces the *affine connection*, which is needed to calculate slopes, for instance, because the coordinates themselves change as we move from place to place. This leads to parallel transport and geodesics and to "absolute slopes" which tell us how the field itself changes from point to point independently of the coordinate changes. Divergence and curl must be rewritten using absolute slopes.

The final stretch is to curved space, Gauss' result that we can detect intrinsic curvature of the space we live in without having to get "outside", and Riemann's capturing of this in general in the curvature tensor. Curvature can be positive, zero (flat) or negative. Finally we look at a simplification of the four-index Riemann curvature tensor, with loss of information in four or more dimensions, to the two-index Ricci tensor.

This mathematics needs new notation, the *index notation* and, for calculations, a generalization of matrices to a special kind of *relation* we call a *protor*. MATLAB implementations of some of the suite of programs giving the *protor calculator* are outlined in the course of the Part.

Here is the table of contents.

Here are some useful results.

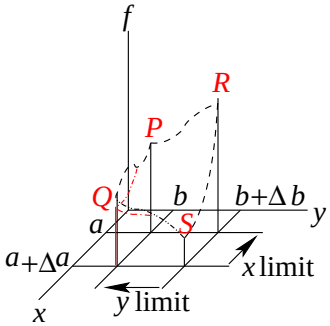
	Metric	Affine connection	Riemann curvature
Cartesian	$g_{\mathrm{dd}}(j\ k\ g)$: $(x\ x\ 1)$, $(y\ y\ 1)$		
Gypsum	$g_{\mathrm{dd}}(j\ k\ g)$: $(x\ x\ \ell_1^2)$, $(x\ y\ \ell_1\ell_2 c_{21})$, $(y\ x\ \ell_1\ell_2 c_{21})$, $(y\ y\ \ell_2^2)$		
Plane polar	$g_{\mathrm{dd}}(j\ k\ g)$: $(r\ r\ 1)$, $(\theta\ \theta\ r^2)$	$\Gamma_{\mathrm{udd}}(j\ \mathtt{k}\ \ell\ \Gamma)$: $(r\ \theta\ \theta\ -r)$, $(\theta\ r\ \theta\ 1/r)$, $(\theta\ \theta\ r\ 1/r)$	
Spherical polar	$g_{\mathrm{dd}}(j\ k\ g)$: $(r\ r\ 1)$, $(\theta\ \theta\ r^2)$, $(\phi\ \phi\ r^2 s^2)$, $s=\sin\theta\quad c=\cos\theta$	$\Gamma_{\mathrm{udd}}(j\ k\ \ell\ \Gamma)$: $(r\ \theta\ \theta\ -r)$, $(r\ \phi\ \phi\ -rs^2)$, $(\theta\ r\ \theta\ 1/r)$, $(\theta\ \theta\ r\ 1/r)$, $(\theta\ \phi\ \phi\ -cs)$, $(\phi\ r\ \phi\ 1/r)$, $(\phi\ \phi\ r\ 1/r)$, $(\phi\ \theta\ \phi\ c/s)$, $(\phi\ \phi\ \theta\ c/s)$	
Spherical 2D	$g_{\mathrm{dd}}(j\ k\ g)$: $(\theta\ \theta\ r^2)$, $(\phi\ \phi\ r^2 s^2)$	$\Gamma_{\mathrm{udd}}(j\ k\ \ell\ \Gamma)$: $(\theta\ \phi\ \phi\ -cs)$, $(\phi\ \theta\ \phi\ c/s)$, $(\phi\ \phi\ \theta\ c/s)$	$R_{\mathrm{dddd}}(\mu\ \lambda\ \alpha\ \beta\ R)$: $(\theta\ \phi\ \theta\ \phi\ r^2 s^2)$, $(\phi\ \theta\ \theta\ \phi\ -r^2 s^2)$, $(\theta\ \phi\ \phi\ \theta\ -r^2 s^2)$, $(\phi\ \theta\ \phi\ \theta\ r^2 s^2)$
Hotplate	$g_{\mathrm{dd}}(j\ k\ g)$: $(\rho\ \rho\ 1/(1+k^2\rho^2))$, $(\phi\ \phi\ \rho^2/(1+k^2\rho^2))$	(See Spherical 2D)	(See Spherical 2D)
Hyperboloid	$g_{\mathrm{dd}}(j\ k\ g)$: $(p\ p\ \cosh(2p))$, $(\theta\ \theta\ \cosh^2(p))$	$\Gamma_{\mathrm{udd}}(\lambda\ \mu\ \nu\ \Gamma)$: $(p\ p\ p\ \tanh(2p))$, $(p\ \theta\ \theta\ -\frac{1}{2}\tanh(2p))$, $(\theta\ p\ \theta\ \tanh(p))$, $(\theta\ \theta\ p\ \tanh(p))$	$R_{\mathrm{dddd}}(\mu\ \nu\ \alpha\ \beta\ R)$: $(p\ \theta\ p\ \theta\ \cosh^2(p)/\cosh(2p))$, $(p\ \theta\ \theta\ p\ -\cosh^2(p)/\cosh(2p))$, $(\theta\ p\ p\ \theta\ -\cosh^2(p)/\cosh(2p))$, $(\theta\ p\ \theta\ p\ \cosh^2(p)/\cosh(2p))$
Paraboloid 2D	$g_{\mathrm{dd}}(j\ k\ g)$: $(r\ r\ 1/(1-a/r))$, $(\theta\ \theta\ r^2)$	$\Gamma_{\mathrm{udd}}(\alpha\ \beta\ \gamma\ \Gamma)$: $(r\ r\ r\ \frac{a}{2r(a-r)})$, $(r\ \theta\ \theta\ a-r)$, $(\theta\ r\ \theta\ 1/r)$, $(\theta\ \theta\ r\ 1/r)$	$R_{\mathrm{dddd}}(\alpha\ \beta\ \gamma\ \delta\ R)$: $(r\ \theta\ r\ \theta\ -a/2/(a-r))$, $(r\ \theta\ \theta\ r\ a/2/(a-r))$, $(\theta\ r\ r\ \theta\ a/2/(a-r))$, $(\theta\ r\ \theta\ r\ -a/2/(a-r))$
Paraboloid 3D	$g_{\mathrm{dd}}(j\ k\ g)$: $(r\ r\ 1/(1-a/r))$, $(\theta\ \theta\ r^2)$, $(\phi\ \phi\ r^2\sin^2\theta)$	$\Gamma_{\mathrm{udd}}(\alpha\ \beta\ \gamma\ \Gamma)$: $(r\ r\ r\ a/2r(a-r))$, $(r\ \theta\ \theta\ a-r)$, $(r\ \phi\ \phi\ (a-r)\sin^2\theta)$, $(\theta\ r\ \theta\ 1/r)$, $(\theta\ \theta\ r\ 1/r)$, $(\theta\ \phi\ \phi\ -\cos\theta\sin\theta)$, $(\phi\ r\ \phi\ 1/r)$, $(\phi\ \phi\ r\ 1/r)$, $(\phi\ \theta\ \phi\ \cos\theta/\sin\theta)$, $(\phi\ \phi\ \theta\ \cos\theta/\sin\theta)$	$R_{\mathrm{dddd}}(\alpha\ \beta\ \gamma\ \delta\ R)$: $(r\ \theta\ r\ \theta\ -a/2/(a-r))$, $(r\ \theta\ \theta\ r\ a/2/(a-r))$, $(\theta\ r\ r\ \theta\ a/2/(a-r))$, $(\theta\ r\ \theta\ r\ -a/2/(a-r))$, $(r\ \phi\ r\ \phi\ -a\sin^2\theta/2/(a-r))$, $(r\ \phi\ \phi\ r\ a\sin^2\theta/2/(a-r))$, $(\phi\ r\ r\ \phi\ a\sin^2\theta/2/(a-r))$, $(\phi\ r\ \phi\ r\ -a\sin^2\theta/2/(a-r))$, $(\theta\ \phi\ \theta\ \phi\ -ar\sin^2(\theta))$, $(\theta\ \phi\ \phi\ \theta\ ar\sin^2(\theta))$, $(\phi\ \theta\ \theta\ \phi\ ar\sin^2(\theta))$, $(\phi\ \theta\ \phi\ \theta\ -ar\sin^2(\theta))$

22. Curved timespace.
23. Gravitational redshift.
24. Spherically symmetric gravity.
25. Schwarzschild orbits and black holes.
26. Rotationally symmetric gravity.
27. Kerr orbits and black holes.
28. Tides.
29. Light orbits.
30. The source of gravity.
31. Cosmology.
32. Negative pressure and dark energy.
33. Gravitational irreversibility.
34. Alternatives to geometry.

II. **The Excursions**

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

1. **Slopes commute?** (Note 1.) Does $\text{slope}_y \text{slope}_x = \text{slope}_x \text{slope}_y$? The formulae in the diagram below makes it appear simple to show that it is generally true, but it is subtle because of the different orders of taking the limits. The first limit in $\partial_x \partial_y$ (slope_x slope_y) is the y-limit shown—bringing R and S to P and Q respectively—while the second limit brings Q and S to P and R. This is no problem unless there is a discontinuity in the $PQRS$ region. Suppose there were a Queen-Anne-style "eye" discontinuity just along the PQ edge, including Q, as shown in red. Then taking the x-limit first will have no problems—the Q to P shrinkage will take place along either the upper or the lower path (it doesn't matter) and avoid the discontinuity; then the RS to PQ shrinkage will be straightforward. But taking the y-limit first will hit the discontinuity at Q and give an undefined slope for the ratio shown with the red division line.



$$\text{slope}_x \ \text{slope}_y f(x,y) \simeq \frac{\dfrac{f(a+\Delta a, b+\Delta b) - f(a+\Delta a, b)}{\Delta b} - \dfrac{f(a, b+\Delta b) - f(a,b)}{\Delta b}}{\Delta a}$$

$$\text{slope}_y \ \text{slope}_x f(x,y) \simeq \frac{\dfrac{f(a+\Delta a, b+\Delta b) - f(a, b+\Delta b)}{\Delta a} - \dfrac{f(a+\Delta a, b) - f(a,b)}{\Delta a}}{\Delta b}$$

2. (Note 2.) Use a rotation matrix twice to show that $c_2 = \cos(2\theta)$ and $s_2 = \sin(2\theta)$.

3. (Note 2.) Show that

$$\begin{pmatrix} s_2 & c_2 \\ c_2 & -s_2 \end{pmatrix} \begin{pmatrix} s & c \\ c & -s \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$$

4. a) Given $f(g(x))$ show that $\partial_x f = (\partial_g f)(\partial_x g)$:
start with

$$\begin{aligned}\partial_x f &\approx \frac{f(g(x+\Delta x)) - f(g(x))}{\Delta x} \\ &= \frac{f(g(x+\Delta x)) - f(g(x))}{g(x+\Delta x) - g(x)} \frac{g(x+\Delta x) - g(x)}{\Delta x}\end{aligned}$$

b) Given $f(g(x), h(x))$ show that $\partial_x f = (\partial_g f)(\partial_x g) + (\partial_h f)(\partial_x h)$:
start with

$$\begin{aligned}\partial_x f &\approx \frac{f(g(x+\Delta x), h(x+\Delta x)) - f(g(x), h(x))}{\Delta x} \\ &= \frac{f(g(x+\Delta x), h(x+\Delta x)) - f(g(x+\Delta x), h(x)) + f(g(x+\Delta x), h(x)) - f(g(x), h(x))}{\Delta x}\end{aligned}$$

c) Take the next step and confirm the "rule of slopes" cited in Note 3.

5. Look up the programming language Aldat, e.g., http://www.cs.mcgill.ca/∼tim/aldat/welcome.html and, for tutorial detail, http://www.cs.mcgill.ca/∼tim/cs612/

6. Implement `slopes(in,vars)` from Note 5 in MATLAB. Use MATLAB's `diff(expr,var)` repeatedly to do the symbolic slope calculations, building up the extra indices in the new first column. It should also call a function `nozeros(in)` at the end to eliminate any rows with zero in the value column. (I cheated in showing the examples because my implementation of `slopes()` does not call `nozeros()`, so I should have shown the invocations as `nozeros(slopes(..))`.)

7. a) Implement `joinred(in1,col1,in2,col2)` from Note 5 in MATLAB. You can write separate functions `reduce(in,col)` and `joinmult(in1,col1,in2,col2)` so that `joinred()` is implemented as

```
nozeros(reduce(joinmult(in1,col1,in2,col2),col1))
```

The task of `joinmult()` is to implement the natural join (**ijoin**) on two relations `in1` and `in2` using `col1` and `col2` as join columns, and to multiply the two last (value) columns of each relation. `Reduce()` then groups the index columns into sets differing only on `col1`, and sums the last (value) column.
b) Your implementation should manage columns as specified at the end of Note 5. You can show, also, that

```
joinred(joinred(..(joinred(protor,1,X,2),..),1,X,2),1,X,2)
```

will apply coordinate transformation `X` to every index in `protor`, and so you can implement `transform(protor,X)` for a protor of any number of indices.

8. For the three-dimensional examples from Note 6,

$$\begin{aligned}\mathbf{f} &= \begin{pmatrix} x \\ y \\ z \end{pmatrix} r^{p-1} \\ \mathbf{h} &= \begin{pmatrix} z-y \\ x-z \\ y-x \end{pmatrix} r^{p-1}\end{aligned}$$

show that

$$\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} (f_x \;\; f_y \;\; f_z) = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} r^{p-1} + \begin{pmatrix} x^2 & xy & xz \\ yx & y^2 & yz \\ zx & zy & z^2 \end{pmatrix} (p-1) r^{p-3}$$

and

$$\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} (h_x \;\; h_y \;\; h_z) = \begin{pmatrix} & 1 & -1 \\ -1 & & 1 \\ 1 & -1 & \end{pmatrix} r^{p-1} + \begin{pmatrix} (z-y)x & (x-z)x & (y-x)x \\ (z-y)y & (x-z)y & (y-x)y \\ (z-y)z & (x-z)z & (y-x)z \end{pmatrix} (p-1) r^{p-3}$$

.

9. In Note 6 show that

$$\begin{pmatrix} x \\ y \end{pmatrix} r^{p-1} = \mathbf{grad} \frac{r^{p+1}}{p+1}$$

10. What is the significance of the divergence vanishing for $p = -1$ in two dimensions and for $p = -2$ in three?

11. Apply to curl the argument in Note 6 for the invariance of the divergence, assuming that the coordinate transformation does not depend on the coordinates, and show that its form is invariant.

12. Show that the vector field that is the slopes of $f(x, y) = xy$ (Note 1) has zero divergence and curl (Note 6).
It does not, however, have zero *shear*, defined as

$$\mathbf{shear}\ q = \partial_x q_y + \partial_y q_x$$

Show that this form is also invariant under the assumption that the transformation coefficients are independent of the coordinates.

13. **Newton's constant.** Given that the acceleration due to gravity at the surface of the Earth is about 10m/s^2, the radius of the Earth is about 6.4Mm (6.4×10^6m), and the mass of the Earth (found from the orbit of the Moon or of artificial satellites) is about 6Ykg (Y = yotta $= 10^{24}$), what is Newton's gravitational constant G_N (see Note 7) in m-s-kg units (and what are the units)?

14. **Gauss' law** says

$$\int_V \mathbf{div}.\mathbf{f} = \int_S \mathbf{f}.\mathbf{n}$$

using conventional integral signs to mean, respectively, the antislope taken over the whole volume V and over the whole surface S. Show that the classical gravitational field

$$\mathbf{f}.\mathbf{n} = \frac{4}{3}\pi r G_N \rho$$

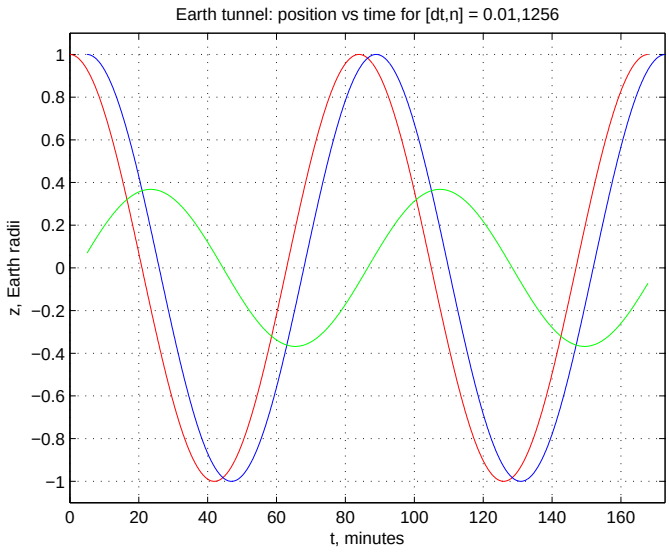
where $\mathbf{f}.\mathbf{n}$ is the part of $\mathbf{f}$ normal to the surface of a sphere of radius r, satisfies Gauss' law when the volume and surface are those of that sphere.

15. **Message through the Earth.** If you could drill a tunnel through the Earth, going straight from a point on the surface, through the centre, to the opposite point on the surface, you could drop a bottle with a message in it and it would take 42 minutes for the trip to the other side.
This Excursion is not to calculate the 42 minutes (although the data in Excursion *Newton's constant* on the Earth's radius and surface gravitational acceleration will enable you to calculate it afterwards), but to plot the motions of the bottle.
Your MATLAB program can imagine the tunnel running down the z-axis and the bottle

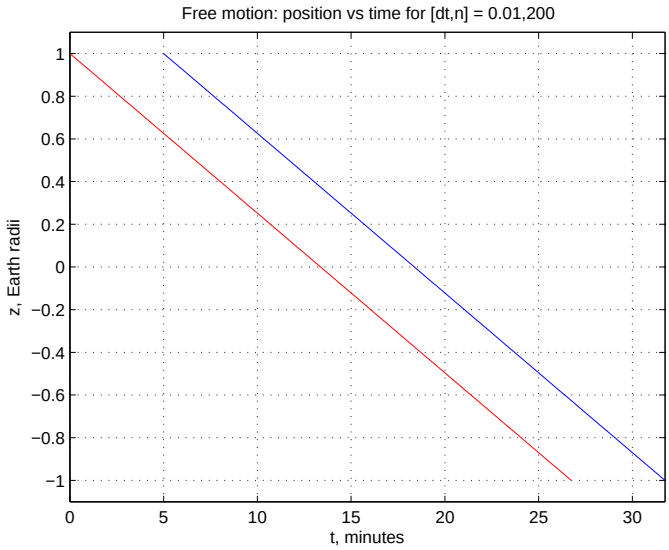
starting with zero velocity at $z = 1$ Earth radius. Then, since acceleration is proportional to z (Note 7)

$$\begin{aligned} v &= v - z \times \Delta t \\ z &= z + v \times \Delta t \end{aligned}$$

is the heart of your program. You should get a nice cosine curve, if you use small enough Δt, with period 2π (and you can adjust the plot so it shows 84 minutes instead of 2π).
Of more interest is the difference between this bottle and a second bottle you dropped five minutes later with a P.S.
Here are the trajectories of the two bottles (red for the message, blue for the P.S.) and of the difference in their z positions (green).

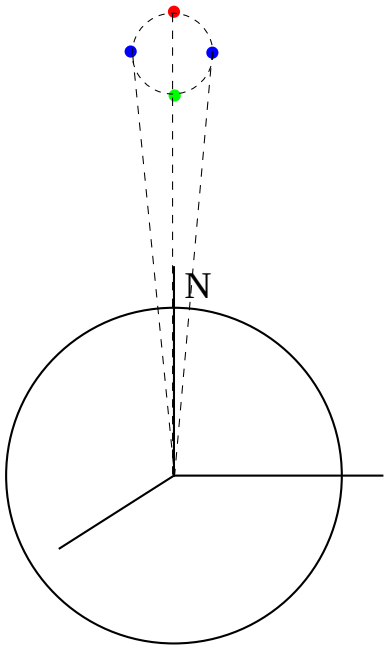


Note that the difference is also sinusoidal and goes periodically to zero: the two bottles each fall back down the tunnel if they are not caught at the other side, and so will meet each other a few minutes after the message reaches the surface. (So, anticipating the P.S. and wanting to avoid a lot of broken glass going up and down, you actually drilled a parallel tunnel for the P.S., both tunnels still effectively through the centre of the Earth, so there will be no crashes.)
This is not the usual trajectory for bodies moving without (non-gravitational) influence. For example, two bodies in uniform motion maintain their separation and do not meet—or meet at most once.



We will be asking: what kind of geodesic (Note 13) must bodies follow in order to meet periodically. Preparing for that is the purpose of this Excursion.
(Wheeler [Whe90, Ch.4] has an elaborate qualitative discussion of this project.)
You can also do all these calculations analytically and even figure out when the two bodies pass each other (if it's not obvious).

16. **Falling traffic lights** with two blue UFOs. The second kind of gravitational influence happens outside the Earth. One body falling is again not very informative, so we shall look at two bodies. In fact we'll look at four, quartering a vertical circle about 200 earth radii above the North pole.
Here is a traffic light, disassembled into its red and green lights, but dropped with red still on top and green still on the bottom, plus two blue UFOs falling with it on either side.



I've exaggerated the horizontal distance but reduced the vertical in this picture. I've also shown the circle defined by the four test masses in dashed lines, and the fall lines for each mass. We'll call the angles from the vertical of the two non-vertical dashed lines (the "UFO"s' fall lines) θ and $-\theta$, a little in defiance of the convention for spherical polar coordinates.
It is clear that the two outlying bodies will approach each other as they follow their fall lines. What about the upper and lower bodies?
Your MATLAB program can start with

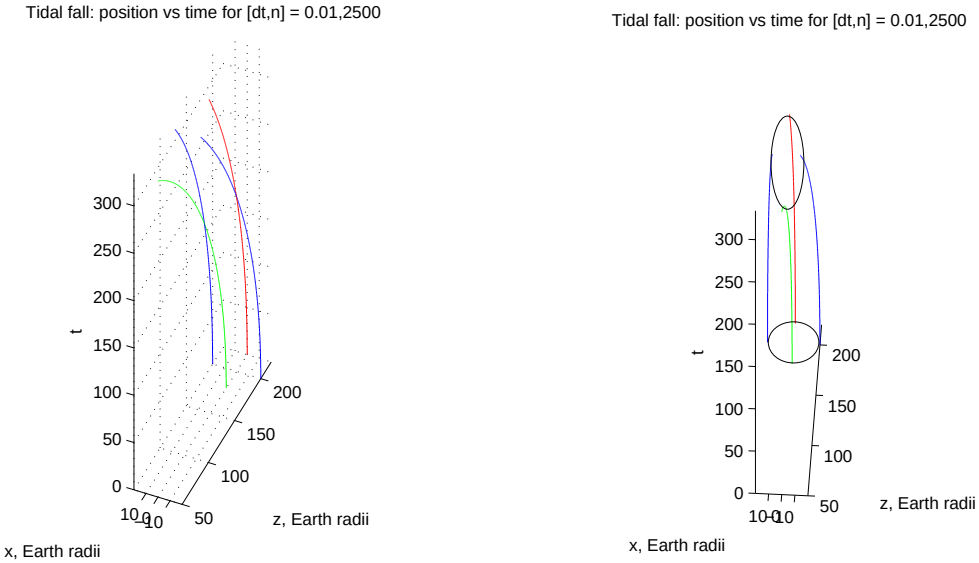
$$\begin{aligned} z &= [180, 200, 220, 200] \\ s &= \sin[0.0, 0.1, 0.0, -0.1] \\ x &= z. * s \end{aligned}$$

and zero velocities v_z and v_x for all four bodies.

Then the loop must be based on (Note 7)

$$\begin{aligned} a_z &= -100./z. \wedge 2 \\ a_x &= a_z. * s \\ v_z &= v_z + a_z * \Delta t \\ v_x &= v_x + a_x * \Delta t \end{aligned}$$

Here is the result from two perspectives. In the second I've drawn the circle at the start and the ellipse at the end showing how the circle was distorted. (Note that MATLAB plots the time *upwards*.
The gravitational effect this shows is what the gravity of the Moon does to the oceans of the Earth, and how the gravity of the Earth locked the Moon's rotation so that one side now always faces us.



While the bodies at θ and $-\theta$ approach each other along their fall lines, the other two, upper and lower, bodies draw apart from each other because the $1/r^2$ acceleration is greater for the lower body.
This decreasing and increasing separations of pairs of bodies, depending on their orientation relative to each other, is characteristic of gravitation outside but near a source of gravity. We also will want to ask, as in the previous Excursion, what kind of geodesics (Note 13) are being followed.
(Wheeler [Whe90, Ch.5] also discusses this qualitatively. You can imagine the two skydivers

on the cover of that book as the two test bodies at θ and $-\theta$. There would have to be a third skydiver below them, and a fourth above, taking the picture, to complete the configuration. And they'd be jumping from a lot further than the atmosphere.)
Again, you can do these calculations analyticlly.

17. Calculate and confirm the lengths I've given in Note 8 for $\boldsymbol{e}^1$ and $\boldsymbol{e}^2$, their angles, and the three ways of showing that the length of $\boldsymbol{v}$ is $\sqrt{2}$.

18. Show directly that the metric for a coordinate system in which the axes $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$ respectively rotated towards each other each by an angle (in one case positive and in the other case negative) given by c and s as

$$\begin{pmatrix} 1 & 2cs \\ 2cs & 1 \end{pmatrix}$$

Check this against the more general result of Note 9.

19. Show that $g_{\mathrm{ud}} = I = g_{\mathrm{du}}$, i.e., $g^j_k = \delta^j_k$ the identity matrix. Hint: how must you use $g^{j\ell}$ to change $g_{\ell k}$ to g^j_k?

20. From Note 10 show that

$$r^2 = (x^1 \;\; x^2) \begin{pmatrix} \ell_1^2 & \ell_1\ell_2 c_{21} \\ \ell_1\ell_2 c_{21} & \ell_2^2 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}$$

so

$$\begin{pmatrix} \partial_1 \\ \partial_2 \end{pmatrix} r = \frac{1}{r} \begin{pmatrix} \ell_1^2 & \ell_1\ell_2 c_{21} \\ \ell_1\ell_2 c_{21} & \ell_2^2 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}$$

21. Using $g_{jk} = X_{m}{}_j X_{m}{}_k$, $g^{jk} = X_{m}{}^j X_{m}{}^k$ and $X_{m}{}^j X_{n}{}^j$ show that $g_{j\ell} g^{\ell k} = \delta^k_j$ and $\partial_j = X_{m}{}_j \partial_m$.

22. Using the notation at the end of Note 11, with g and X meaning either of g_1 or g_2 and X_1 or X_2 respectively, show that

$$\begin{aligned} g_{jj'} &= X_{m}{}_j X_{m}{}_{j'} \\ g^{jj'} &= X_{m}{}^j X_{m}{}^{j'} \end{aligned}$$

23. **Levi-Civita symbol, alternating tensors.** Show that the Levi-Civita symbol is not a tensor.
a) Levi-Civita defined a family of symbols, with one index for each dimension, in terms of even and odd permutations. Thus

ϵ_{mm}(j	k	ϵ)
x	y	1
y	x	-1

ϵ_{mmm}(j	k	ℓ	ϵ)
x	y	z	1
y	z	x	1
z	x	y	1
y	x	z	-1
x	z	y	-1
z	y	x	-1

Let's transform ϵ_{jk} from Cartesian x, y, where $g_C = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ to $\epsilon^{j'k'}$ in polar r, θ, where $g_P = \begin{pmatrix} 1 & \\ & r^2 \end{pmatrix}$ (Note 11).

$$\begin{aligned}
\begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} &= \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix} \\
(\boldsymbol{e}_r \;\; \boldsymbol{e}_\theta) \begin{pmatrix} v^r \\ v^\theta \end{pmatrix} &= (\boldsymbol{\epsilon}_x \;\; \boldsymbol{\epsilon}_y) \begin{pmatrix} c & -rs \\ s & rc \end{pmatrix} \begin{pmatrix} v^r \\ v^\theta \end{pmatrix} \\
&= (\boldsymbol{\epsilon}_x \;\; \boldsymbol{\epsilon}_y) \begin{pmatrix} v_x \\ v_y \end{pmatrix} \\
\begin{pmatrix} v^r \\ v^\theta \end{pmatrix} &= \begin{pmatrix} c & s \\ -s/r & c/r \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix}
\end{aligned}$$

and the transformation is $v^{j'} = X_j{}^{j'} v_j$ with

$X_j{}^{j'}$(j	j'	X)
x	r	c
y	r	s
x	θ	$-s/r$
y	θ	c/r

So $X_j{}^{j'} X_k{}^{k'} \epsilon_{jk}$ is

j'	k'	XXϵ
r	θ	$1/r$
θ	r	$-1/r$

which is $\epsilon^{j'k'}/\sqrt{g}$ where $g = \det g_P$ and $\epsilon^{j'k'}$ is again the Levi-Civita symbol.

$\epsilon^{j'k'}$(j'	k'	ϵ)
r	θ	1
θ	r	-1

So the Levi-Civita symbol does not transform into itself and is not a tensor.
b) But the corresponding *alternating tensor* is

$$\frac{1}{\sqrt{g}} \epsilon^{jk}$$

(Note that g disappears in Cartesian coordinates because

$$\frac{1}{\sqrt{g_C}} \epsilon_{jk} = \epsilon_{jk}.)$$

c) The Levi-Civita symbol has an important connection with determinants. For example

$$g = \begin{vmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{vmatrix} = g_{11} g_{22} - g_{12} g_{21} = g_{1j'} g_{2l'} \epsilon^{j'k'}$$

Now notice

$$g_{2j'}g_{1k'}\epsilon^{j'k'} = g_{21}g_{12} - g_{11}g_{22} = -g$$

while

$$\begin{aligned} g_{1j'}g_{1k'}\epsilon^{j'k'} &= g_{11}g_{12} - g_{12}g_{11} = 0 \\ g_{2j'}g_{2k'}\epsilon^{j'k'} &= g_{21}g_{22} - g_{22}g_{21} = 0 \end{aligned}$$

So we can immediately get the covariant form of the Levi-Civita symbol

$$\epsilon_{jk} = g_{jj'}g_{kk'}\epsilon^{j'k'} = g\epsilon^{jk}$$

And, for the alternating tensor,

$$\sqrt{g}\epsilon_{jk} = g\frac{1}{\sqrt{g}}\epsilon^{jk}$$

so the co- and contravariant forms are $\sqrt{g}\epsilon_{jk}$ and $\epsilon^{jk}/\sqrt{g}$ respectively.
This generalizes to the d-index alternating tensor in d dimensions.
d) The alternating tensor can also be used to show that the curl is a tensor, given that the absolute slope (Note 14) is a tensor. In three dimensions

$$\begin{aligned} (\mathbf{curl} \times \boldsymbol{v})_{jk} &= \sqrt{g}\epsilon_{jk\ell}\frac{1}{\sqrt{g}}\epsilon^{\ell mn}D_m v_n \\ &= D_j v_k - D_k v_j \\ &= \partial_j v_k - \partial_k v_j \end{aligned}$$

(and Note 15 derives the third line).
The product of the three tensors involved must give a tensor.
(In 3D it is possible to define the curl as a "pseudo-vector", using only one alternating tensor

$$(\mathbf{curl} \times \boldsymbol{v})^{\ell} = \frac{1}{\sqrt{g}}\epsilon^{\ell mn}D_m v_n$$

but pseudo-vectors are odd beasts because they do not transform properly under reflection. They have three components only in 3D, and should generally be thought of as antisymmetric 2-index tensors as in the first definition above.)
e) Write a MATLAB function `out = leviCivita(n)` which uses recursion to calculate the Levi-Civita symbol on n indices. Don't try to use it for $n >\approx 5$: why?

24. **Physical tensors.** Look up the tensors of *polarizability, conductivity* and *inertia* in [FLS64, Ch.31]. They are all 2-index tensors defined as relating one vector to another. So they all transform properly and so *are* tensors.
Feynman also discusses the tensor of *stress*, defined as linking each of the three components of force applied to each of three directions of surface area of a very small cube internal to a body—he suggests a block of jello. This is a 2-index protor and to show that it is a tensor we must show that it, too, relates two vectors to each other. Study Feynman's argument.
He goes on to discuss a corresponding 2-index *strain* tensor, which gives the responses of the jello to the stress. The relationship between stress and strain is, in general, given by a 4-index protor, which must also be a tensor since both stress and strain are tensors. And the piezoelectric field, a vector, relates to the stress tensor via the 3-index *piezoelectric coefficients*, also, necessarily, a tensor.

25. **Orthonormal polars.** Working with an orthonormal basis often gives simpler results and is an easy transformation if the basis is already orthogonal, as in polar coordinates. We have

$$\begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} = \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix}$$
$$g_{\mathrm{dd}} = \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \begin{pmatrix} c & -rs \\ s & rc \end{pmatrix} = \begin{pmatrix} 1 & \\ & r^2 \end{pmatrix}$$

i.e., $(\Delta s)^2 = (\Delta r)^2 + (r\Delta\theta)^2$. We would like

$$\begin{pmatrix} \boldsymbol{e}_{\hat{r}} \\ \boldsymbol{e}_{\hat{\theta}} \end{pmatrix} = \begin{pmatrix} 1 & \\ & 1/r \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} = \begin{pmatrix} 1 & \\ & 1/r \end{pmatrix} \begin{pmatrix} c & s \\ -rs & rc \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix}$$
$$g_{\mathrm{dd}} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} c & -s \\ s & c \end{pmatrix} = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$$

so $(\Delta s)^2 = (\Delta\hat{r})^2 + (\Delta\hat{\theta})^2$.
We can describe the new basis as

$$\begin{aligned} \boldsymbol{e}_{\hat{r}} &= [1, 0] \\ \boldsymbol{e}_{\hat{\theta}} &= [0, 1/r] \end{aligned}$$

What is the orthonormal basis for the metric

$$(\Delta s)^2 = -(1 - 2m/r)(\Delta t)^2 + (\Delta r)^2/(1 - 2m/r) + r^2((\Delta\theta)^2 + (\sin\theta\Delta\phi)^2)?$$

Is there a singularity in the new metric?

26. What is the affine connection for gypsum coordinates?

27. Use the derivation at the end of Note 12 that

$$\Gamma^\ell_{jk'} = \frac{1}{2} g^{\ell k}(\partial_j g_{kk'} + \partial_{k'} g_{jk} - \partial_k g_{k'j})$$

to calculate the affine connection for the plane polar metric.

∂_{d}(j ∂)

r	slope_r
θ	slope_θ

g_{dd}(k k' g)

r	r	1
θ	θ	r^2

$\partial g_{\mathrm{ddd}}$(j k k' ∂g)

r	θ	θ	2r

$\partial g_{\mathrm{ddd}}$(j k k' ∂g)

r	θ	θ	2r
θ	θ	r	2r
θ	r	θ	−2r

g_{uu}(ℓ k g)

r	r	1
θ	θ	$1/r^2$

$(\frac{1}{2}\partial g \bowtie g)_{\mathrm{d-du}}$(j k k' ℓ ∂gg)

r	θ	θ	θ	1/r
θ	r	θ	θ	1/r
θ	θ	r	r	−r

Γ_{udd}(ℓ j k' Γ)

r	θ	θ	−r
θ	r	θ	1/r
θ	θ	r	1/r

28. Show that the result

$$\begin{pmatrix} {v_1}^r \\ {v_1}^\theta \end{pmatrix} = \left(\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 0 & -r\Delta\theta \\ \Delta\theta/r & \Delta r/r \end{pmatrix} \right) \begin{pmatrix} {v_2}^r \\ {v_2}^\theta \end{pmatrix}$$

at the beginning of Note 12 gives

$$\Delta v^j = -\Gamma^j_{k\ell} \Delta^k v^\ell$$

consistent with result

$$\partial_j \boldsymbol{e}^k = -\Gamma^k_{j\ell} \boldsymbol{e}^\ell$$

at the end of Note 12.

29. Note 12 derives the affine connection for plane polar coordinates more directly than by applying the general formula as in the previous Excursion, by taking the slopes of the components of a vector $\boldsymbol{v}$. Give another direct derivation by taking the slopes of the basis vectors.
Find

$$\Delta \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_{r'} \\ \boldsymbol{e}_{\theta'} \end{pmatrix} - \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix}$$

by finding $\begin{pmatrix} \boldsymbol{e}_{r'} \\ \boldsymbol{e}_{\theta'} \end{pmatrix}$ from $\begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix}$ and $\begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix}$ from $\begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix}$ then combining to get $\begin{pmatrix} \boldsymbol{e}_{r'} \\ \boldsymbol{e}_{\theta'} \end{pmatrix}$ directly from $\begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix}$ in terms of $r' = r + \Delta r$, $c_{21} = \cos(\theta' - \theta)$ and $s_{21} = \sin(\theta' - \theta)$ as Δr and $\Delta\theta$ shrink.

$$\begin{aligned}
\begin{pmatrix} \boldsymbol{e}_{r'} \\ \boldsymbol{e}_{\theta'} \end{pmatrix} &= \begin{pmatrix} c' & s' \\ -r's' & r'c' \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \end{pmatrix} \\
&= \begin{pmatrix} c' & s' \\ -r's' & r'c' \end{pmatrix} \begin{pmatrix} c & -s/r \\ s & c/r \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} \\
&= \begin{pmatrix} c_{21} & s_{21}/r \\ -r's_{31} & r'c_{21}/r \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} \\
&\to \begin{pmatrix} 1 & \Delta\theta/r \\ -(r+\Delta r)\Delta\theta & r(r+\Delta r)/r \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} \\
&= \left(\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} & \Delta\theta/r \\ -r\Delta\theta & \Delta r/r \end{pmatrix} \right) \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix}
\end{aligned}$$

So

$$\begin{aligned}
\Delta \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix} &= \begin{pmatrix} \partial_r \boldsymbol{e}_r & \partial_\theta \boldsymbol{e}_r \\ \partial_r \boldsymbol{e}_\theta & \partial_\theta \boldsymbol{e}_\theta \end{pmatrix} \\
&= \begin{pmatrix} & \Delta\theta/r \\ -r\Delta\theta & \Delta r/r \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \end{pmatrix}
\end{aligned}$$

and because of the independence of Δr and $\Delta\theta$ from each other, we can get the following from $\partial_j \boldsymbol{e}_k = \Gamma^\ell_{jk} \boldsymbol{e}_\ell$

$\Gamma_{\mathrm{udd}}(\ell$	j	k	Γ	)
r	θ	θ	$-r$	
θ	r	θ	$1/r$	
θ	θ	r	$1/r$	

30. Following the arguments of Note 14, show that the field

$$(f_x \;\; f_y) = (-y \;\; x) r^{p-1}$$

becomes $(f^r \;\; f^\theta) = (0 \;\; r^{p-1})$.
Show that

$$\left(\begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} \boldsymbol{f} \right) \cdot (\boldsymbol{\epsilon}_x \;\; \boldsymbol{\epsilon}_y) = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} r^{p-1} + (p-1) \begin{pmatrix} -xy & x^2 \\ -y^2 & xy \end{pmatrix} r^{p-3}$$

Show that

$$\left(\begin{pmatrix} \partial_r \\ \partial_\theta \end{pmatrix} \boldsymbol{f} \right) \cdot (\boldsymbol{e}^r \;\; \boldsymbol{e}^\theta) = \begin{pmatrix} & p/r^2 \\ 1 & \end{pmatrix} r^p$$

a) by taking the absolute slope, and
b) by transforming from

$$\left(\begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} \boldsymbol{f} \right) \cdot (\boldsymbol{\epsilon}_x \;\; \boldsymbol{\epsilon}_y)$$

Show that absolute slopes

$$\begin{pmatrix} & p/{r_1}^2 \\ 1 & \end{pmatrix} {r_1}^p$$

and

$$\begin{pmatrix} & p/{r_2}^2 \\ 1 & \end{pmatrix} {r_2}^p$$

transform into each other in the limit as $(r_1 \;\; \theta_1)$ and $(r_2 \;\; \theta_2)$ become arbitrarily close.

31. Check that the transformation to spherical polar coordinates from Cartesian is $\boldsymbol{e}_j = X_{mj} \boldsymbol{\epsilon}_m$ so $\boldsymbol{\epsilon}_m = X_m{}^j \boldsymbol{e}_j$:

$$\begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \\ \boldsymbol{e}_\phi \end{pmatrix} = \begin{pmatrix} s_\theta c_\phi & s_\theta s_\phi & c_\theta \\ -r c_\theta c_\phi & -r c_\theta s_\phi & r s_\theta \\ -r s_\theta s_\phi & r s_\theta c_\phi & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \\ \boldsymbol{\epsilon}_z \end{pmatrix} \qquad \begin{pmatrix} \boldsymbol{\epsilon}_x \\ \boldsymbol{\epsilon}_y \\ \boldsymbol{\epsilon}_z \end{pmatrix} = \begin{pmatrix} s_\theta c_\phi & -c_\theta c_\phi / r & -s_\phi/(r s_\theta) \\ s_\theta s_\phi & -c_\theta s_\phi / r & c_\phi / r s_\theta \\ c_\theta & s\theta / r & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_\theta \\ \boldsymbol{e}_\phi \end{pmatrix}$$

where $c_\theta = \cos\theta, s_\theta = \sin\theta, c_\phi = \cos\phi$ and $s_\phi = \sin\phi$.
Check that $X_{mj} X_m{}^k = \delta^k_j$ and that $X_{mj} X_{mk} = g_{jk}$ with g_{jk} the spherical metric of Note 16.

32. For spherical polar coordinates (Note 16) derive the affine connection from the metric.
First, using $\partial_{\mathrm{d}}(k, \partial)$,

$g_{\mathrm{dd}}(\ell$	ρ	g	$)$
r	r	1	
θ	θ	r^2	
ϕ	ϕ	$r^2 s^2$	

$\partial g_{rmddd}(k$	ℓ	ρ	∂g $)$
r	θ	θ	$2r$
r	ϕ	ϕ	$2rs^2$
θ	ϕ	ϕ	$2r^2 cs$
θ	r	θ	$2r$
ϕ	r	ϕ	$2rs^2$
ϕ	θ	ϕ	$2r^2 cs$
θ	θ	r	$-2r$
ϕ	ϕ	r	$-2rs^2$
ϕ	ϕ	θ	$-2r^2 cs$

Next we join on ρ, which we can later remove to give $\Gamma^j_{k\ell}$.

$g_{\rm uu}(\ell$	ρ	g	)
r	r	1	
θ	θ	$1/r^2$	
ϕ	ϕ	$1/(rs)^2$	

$g_{\rm uu} \bowtie \partial g_{\rm ddd}(j$	ρ	k	ℓ	$d\partial g$	)
r	r	θ	θ	$-r$	
r	r	ϕ	ϕ	$-rs^2$	
θ	θ	r	θ	$1/r$	
θ	θ	θ	r	$1/r$	
θ	θ	ϕ	ϕ	$-cs$	
ϕ	ϕ	r	ϕ	$1/r$	
ϕ	ϕ	ϕ	r	$1/r$	
ϕ	ϕ	θ	ϕ	c/s	
ϕ	ϕ	ϕ	θ	c/s	

Finally we combine the affine connection with small changes in the spherical polar coordinates, using $\Delta_{\rm u}(\mu, \Delta)$, to get the matrix for $\Gamma\Delta$. (The join index, μ, can be removed to get the matrix on β and λ. Note that the (ϕ, ϕ) entry is the sum of two terms.)

$\Gamma\Delta_{\rm ud-}(\beta$	λ	μ	$\Gamma\Delta$	)
r	θ	θ	$-r\Delta\theta$	
r	ϕ	ϕ	$-rs^2\Delta\phi$	
θ	r	θ	$\Delta\theta/r$	
θ	θ	r	$\Delta r/r$	
θ	ϕ	ϕ	$-cs\Delta\phi$	
ϕ	r	ϕ	$\Delta\phi/r$	
ϕ	ϕ	r	$\Delta r/r$	
ϕ	θ	ϕ	$\Delta\phi c/s$	
ϕ	ϕ	θ	$\Delta\theta c/s$	

$$\begin{array}{cc} & \lambda \\ & r \\ \beta & \theta \\ & \phi \end{array} \begin{array}{c} \begin{array}{ccc} r & \theta & \phi \end{array} \\ \begin{pmatrix} 0 & -r\Delta\theta & -rs^2\Delta\phi \\ \Delta\theta/r & \Delta r/r & cs\Delta\phi \\ \Delta\phi/r & \Delta\phi c/s & \Delta r/r + \Delta\theta c/s \end{pmatrix} \end{array}$$

This is for spherical polar coordinates in all three dimensions. Now we restrict ourselves to the two-dimensional surface of a sphere of fixed radius. All terms containing r can be dropped, and $\Delta r = 0$. Here is the $\Gamma\Delta$ matrix.

$$\begin{array}{cc} & \lambda \\ \beta & \theta \\ & \phi \end{array} \begin{array}{c} \begin{array}{cc} \theta & \phi \end{array} \\ \begin{pmatrix} 0 & -cs\Delta\phi \\ \Delta\phi c/s & \Delta\theta c/s \end{pmatrix} \end{array}$$

33. Develop the MATLAB function `geodesicLongLat(..)` used to plot geodesics on the surface of the Earth in Note 16.

34. Show that two objects each following a separate geodesic on the sphere (Note 16: great circle) are separated from each other by a distance $\cos\phi\sqrt{2(1-\cos\alpha)}$ where the great circles are in planes separated by angle α, which is small enough that the angle ϕ, increasing with time, can be considered to give the location of each body as it progresses around the great circle. The two objects could be the two bottles in the Excursion *Message through the Earth*: compare that result with this.

35. Show that co-contra gives the same result, $\cos\theta$, for the rotation (curl) field in Note 16 as co-co. Show that the two contra- approaches get $\cos\theta/\sin\theta$ and that this is plausible.

36. **Hotplate geometry.** a) Show that the circumference of a circle drawn on a sphere of radius R, with the circle's centre at $\theta = 0$ and radius $\rho = r\theta$ as measured along the surface of the sphere is

$$\pi R \sin\left(\frac{\rho}{R}\right) = 2\pi\rho\left(1 - \frac{\rho^2}{3!R^2} + ..\right)$$

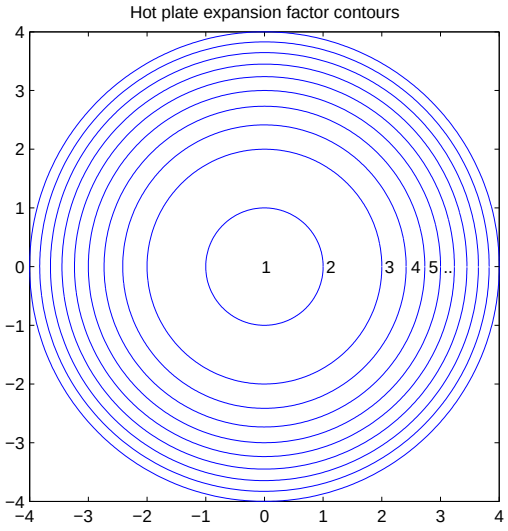
Compared with the circumference of $2\pi\rho$ that a Euclidean geometer would expect, show that this corresponds to an *expansion of the rulers* used to measure the circumference by a factor

$$\frac{1}{1-\frac{\rho^2}{3!R^2}+..} \approx 1+\frac{\rho^2}{3!R^2}$$

b) Now imagine a flat 2D "hot plate" which heats up rulers and everything else on it so that they expand by a factor

$$f(\rho,\phi) = 1+k^2\rho^2$$

That is, in this case, the plate is cool in the middle and gets hotter in concentric rings as we move from the centre to edge.



(In general, f is an *expansion field*, depending on both coordinates, ρ and ϕ, but for now it depends only on ρ and we will just write it f.)
What is the geometry of the hot plate? Careful! Rulers on the hot plate expand no matter which direction they're lined up with, so not only will the circumference appear to shrink inversely as the rulers expand, but the radius of a circle will not be measured as ρ but as about

$$1\ (\rho = 0 \text{ to } \Delta\rho) + \frac{1}{f}\ (\rho = \Delta\rho \text{ to } 2\Delta\rho) + \frac{1}{f}\ (\rho = 2\Delta\rho \text{ to } 3\Delta\rho) + ..$$

or, exactly

$$r = \text{antislope}\ |_0^\rho\ \frac{1}{f} = \text{antislope}\ |_0^\rho\ \frac{1}{1+k^2\rho^2}$$

(To calculate this, recall that slopes of inverse functions are reciprocals, e.g.,

$$y = x^2 \quad : \quad \text{slope}_x y = 2x \quad ; \quad \text{slope}_y x = \frac{1}{2x} = \frac{1}{2\sqrt{y}}$$

So

$$t = \tan u \quad : \quad \text{slope}_u t = 1+t^2 \quad ; \quad \text{slope}_t u = \frac{1}{1+t^2}$$

Hence antislope $\frac{1}{1+t^2} = u = \text{atan } t$.)
So

$$r = \text{antislope } |_0^p \frac{1}{1+k^2\rho^2} = \frac{1}{k}\text{atan}k\rho$$

or

$$\rho = \frac{1}{k}\tan kr$$

and

$$\begin{aligned}\frac{2\pi\rho}{f} &= \frac{2\pi\rho}{1+k^2\rho^2} \\ &= \frac{2\pi}{k}\frac{\tan kr}{1+\tan^2 kr} \\ &= \frac{2\pi}{k}\sin kr\cos kr \\ &= \frac{2\pi}{2k}\sin 2kr\end{aligned}$$

Compare this to a sphere of radius R

$$\frac{2\pi\rho}{f} = 2\pi R\sin\frac{r}{R}$$

and we have that the geometry of the $1+k^2\rho^2$ hot plate is that of a sphere of radius $R = 1/2k$.
c) To show that the mapping from sphere to $1 + k^2\rho^2$ hot plate is complete we must show that the metrics are the same.
For a sphere

$$(\Delta s)^2 = (\Delta\rho)^2 + R^2\sin^2\frac{\rho}{R}(\Delta\phi)^2$$

where ρ is the radial distance as measured on the surface of the sphere.
For the hot plate, the distance is that of plane polar coordinates, all divided by the expansion factor

$$\begin{aligned}(\Delta s)^2 &= \frac{1}{f^2}((\Delta\rho)^2 + \rho^2(\Delta\phi)^2) \\ &= \frac{1}{(1+k^2\rho^2)^2}((\Delta\rho)^2 + \rho^2(\Delta\phi)^2)\end{aligned}$$

Now

$$\Delta\rho = (\text{slope}_r\rho)\Delta r = (1+\tan^2 kr)\Delta r$$

so

$$\frac{\Delta\rho}{f} = \Delta r$$

Also, as above,

$$\frac{\rho}{f} = \frac{1}{k}\frac{\tan kr}{1+\tan^2 kr} = \frac{1}{2k}\sin 2kr$$

So

$$(\Delta s)^2 = (\Delta r)^2 + \frac{1}{(2k)^2}\sin^2(2kr)(\Delta\phi)^2$$

where r is the radial distance as measured by the ruler on the hot plate.
The correspondence is exact (except, as Feynman points out [Fey97, p.120], at the one point at infinity (the south pole)).
d) The idea of a "hot plate" expansion field $f(\rho,\phi)$ is not limited to imitating the surface of a sphere. Create other interesting geometries. (See [**?**] for a proposal by Poincaré for a hotplate geometry with temperature $T = a^2 - r^2$ which is coolest (0) on the boundary at $r = a$.)

37. **Absolute slope and curvature.** The curvature tensor, derived using parallel transport at the beginning of Note 17, is also the commutator of the absolute slope.

$$[D_k, D_{k'}]v_j = D_k D_{k'} v_j - D_{k'} D_k v_j = -R^{\ell}_{jkk'} v_\ell$$

We use

$$\begin{aligned} D_{k'} v_j &= \partial_{k'} v_j - \Gamma^{\ell}_{k'j} v_\ell \\ D_k t_{k'j} &= \partial_k t_{k'j} - \Gamma^{\ell}_{kj} t_{k'\ell} - \Gamma^{\ell}_{kk'} t_{\ell j} \end{aligned}$$

So that

$$\begin{aligned} D_k D_{k'} v_j &= \partial_k D_{k'} v_j - \Gamma^{\ell}_{kj} D_{k'} v_\ell - \Gamma^{\ell}_{kk'} D_\ell v_j \\ &= \partial_k(\partial_{k'} v_j - \Gamma^{\ell}_{k'j} v_\ell) - \Gamma^{\ell}_{kj}(\partial_{k'} v_\ell - \Gamma^{\ell'}_{k'\ell} v_{\ell'}) - \Gamma^{\ell}_{kk'}(\partial_\ell v_j - \Gamma^{\ell'}_{\ell j} v_{\ell'}) \\ D_{k'} D_k v_j &= \partial_{k'} D_k v_j - \Gamma^{\ell}_{k'j} D_k v_\ell - \Gamma^{\ell}_{k'k} D_\ell v_j \\ &= \partial_{k'}(\partial_k v_j - \Gamma^{\ell}_{kj} v_\ell) - \Gamma^{\ell}_{k'j}(\partial_k v_\ell - \Gamma^{\ell'}_{k\ell} v_{\ell'}) - \Gamma^{\ell}_{k'k}(\partial_\ell v_j - \Gamma^{\ell'}_{\ell j} v_{\ell'}) \end{aligned}$$

So, subtracting and doing a lot of cancellation, and one swap of dummy variables ℓ and ℓ',

$$\begin{aligned} [D_k, D_{k'}]v_j &= (\partial_{k'}\Gamma^{\ell}_{kj} - \partial_k \Gamma^{\ell}_{k'j} + \Gamma^{\ell'}_{kj}\Gamma^{\ell}_{k'\ell'} - \Gamma^{\ell'}_{k'j}\Gamma^{\ell}_{k\ell'}) v_\ell \\ &= -R^{\ell}_{jkk'} v_\ell \end{aligned}$$

It is also true that

$$[D_k, D_{k'}]v^j = R^{j}_{\ell kk'} v^\ell$$

and that higher-rank tensors just generate additional terms, e.g.,

$$[D_k, D_{k'}]t^j_{j'} = R^{j}_{\ell kk'} t^{\ell}_{j'} - R^{\ell}_{j'kk'} t^{j}_{\ell}$$

Since the absolute slope is a tensor, so are products and commutators of absolute slopes. Thus, the curvature protor is indeed a tensor. It is also the only second slope of the metric tensor, g_{dd}, that is a tensor.

38. **Bianchi identities.** Using the result of Excursion *Absolute slope and curvature*, the cyclic symmetry of the curvature tensor from Note 17

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

and, of course, the other symmetries and antisymmetries of R_{dddd}, prove the "Bianchi identities"

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

a) For any (non-commutative) operator D show the "Jacobi identity"

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

b) Expand the first of the three terms of the Jacobi identity, applied to an arbitrary vector, with D now the absolute slope and replace the commutators by the curvature tensor using the above Excursion.

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

Similarly for the second and third

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

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

So

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

and the second term vanishes because it is just the triple of the cyclic symmetry of R_{dddd} with the first index raised, say by multiplying with $g^{\alpha\gamma}$. Thus

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

giving the Bianchi identities sought (after using the other symmetries to swap indices around). We use the arbitrariness of A_γ to get the first line - which itself gives the Bianchi identities for R_{uddd}. We use the constancy of the metric under absolute slope to step from the second line to the third.

39. **Divergence-free curvature.** Use the Bianchi identities to show that the "Einstein tensor"

$$R^{\mu\nu} - \frac{1}{2}Rg^{\mu\nu}$$

has zero divergence.
Contract the R_{dddd} Bianchi identities first with $g^{\beta\mu}$ then with $g^{\alpha\nu}$ and finally raise an index with $g^{\mu\lambda}$.

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

Note that the constancy of the metric means its absolute slope is zero so that we can add $\Lambda g_{\mu\nu}$ to the Einstein tensor and still have zero divergence

$$D_\mu(R^{\mu\nu} - \frac{1}{2}Rg^{\mu\nu} + \Lambda g^{\mu\nu}) = 0$$

Λ has significance in cosmology and so is called the cosmological constant—even though Einstein, who introduced it in 1917 (to stop the universe from contracting under gravity), later considered it "the biggest blunder of my life": with it he could have anticpated Hubble's 1929 discovery of the expanding universe.

40. **Number of components of the curvature tensor.** The symmetries of R in Note 17 severely reduce the number of independent terms in the curvature tensor. Instead of d^4 terms we can show that there are

$$\frac{1}{12}d^2(d^2-1)$$

independent numbers in d dimensions.

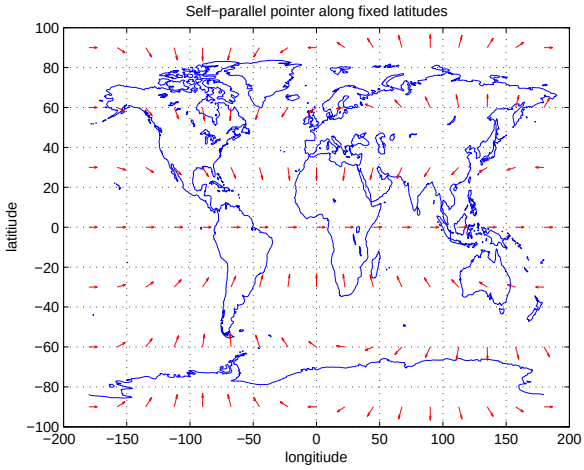
d	$d^2(d^2-1)/12$
1	0
2	1
3	6
4	20

(The single value in 2 dimensions is, for instance, the r^2s^2 we found above for the surface of the sphere. This value appears in four different entries with varying sign, but it is one independent value.)
Here is the demonstration, first in three dimensions then in four.
$R_{\mu\lambda\alpha\beta}$ can be examined in terms of the $\mu\lambda$ pair of indices and the $\alpha\beta$ pair. Each of these pairs, in three dimensions, can take on $\Delta_2 = 3$ values: 12, 13, 23: call them x, y and z, respectively. The reverses, 21, 31, 32, just change sign because of the antisymmetry, and the values 11, 22, 33 represent 0. (Δ_n is the nth triangular number.)
$R_{\mu\lambda\alpha\beta}$ is all possible combinations of x, y and z, with the symmetry of swapping any two of these index-pairs also limiting these values to $\Delta_{\Delta_2} = \Delta_3 = 6$, as we calculated above with the formula.

	x	y	z
x	a	b	c
y	b	d	e
z	c	e	f

(We can also see, with a little thought, that there are 4 different occurrences, with varying sign, of the values for a, d and f, and 8 for b, c and e. This leaves 45 zeros among the $3^4 = 81$ elements of the 3-D tensor.)
In four dimensions, the same argument gives $\Delta_{\Delta_3} = \Delta_6 = 21$ different terms. This must be corrected, because the cyclic symmetry takes one of these different values away. The number of ways we can select four different values from d numbers is $d(d-1)(d-2)(d-3)/4!$, i.e., $d!4$ ("d choose 4") or S^4_{d-3} (the $d-3$rd 4-simplex number). That is 1 when $d = 4$ in the 4-D case. So the number is reduced to 20 as in the above table.
Given that $\Delta_n = n(n+1)/2$ you can combine $\Delta_{\Delta_{d-1}}$ with S^4_{d-3} to get the above formula for the number of independent elements in the curvature tensor.

41. **Angular excess at fixed latitude.** Here is how angular excess develops as we traverse a curve on a sphere.

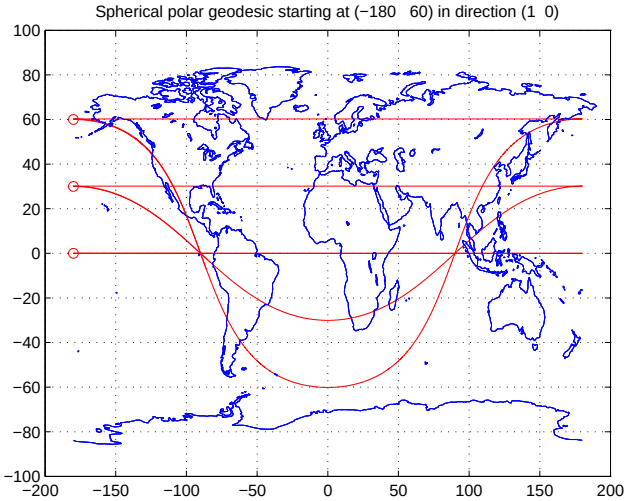


As we proceed along the Equator (lat $= 0$, $\theta = \pi/2$), for instance, holding a pointer in front of us which we maintain parallel to itself for the whole circuit, the pointer stays in the direction we are going.

If we do the same thing arbitrarily close to the North Pole, we must keep turning left while the pointer stays in its fixed direction. This is because the geodesic is a "polar equator", proceeding straight ahead of us at our start, down to the Equator, then the South Pole, then back up behind us. Each time we turn to the left, our (right) arm holding the stick must bend a little further to the right, giving the increasingly clockwise direction shown at the top of the map. The pointer makes a full rotation as we walk around the Pole.

The effect is less extreme at lower latitudes. For instance, on the Tropic of Cancer (lat $=$ 30, $\theta = \pi/2 - \pi/6 = \pi/3$), the pointer will wind up pointing backwards, having made half a rotation.

Effectively the pointer keeps trying to follow a "straight line" (a great circle or geodesic) while we keep turning away. So as we turn away the pointer diverges increasingly from the direction we're headed. A picture of great circles touching each of three latitudes at hypothetical starting points at longitude -180^o ($\phi = 0$) will help visualize this process. We must imagine a new great circle redefining the direction the pointer wants to keep at each step we take along the latitude.



Another way of looking at this is to think of walking with one foot on each side of the latitude circle we're following. At the Equator, we can take equal-length strides with each foot. But at the Tropic of Cancer our northern foot must take shorter strides than our southern foot, because the latitude circle the northern foot follows is shorter than the latitude circle to the south. So we are effectively turning all the time towards the north.
We can calculate all this. We'll use the affine connection for a sphere directly instead of the curvature tensor.

$\Gamma_{udd}(\alpha$	β	γ	Γ $)$
θ	ϕ	ϕ	$-cs$
ϕ	θ	ϕ	c/s
ϕ	ϕ	θ	c/s

remembering that $c = \cos\theta$ and $s = \sin\theta$.
Since, for parallel transport, the change in coordinates of a vector $\mathbf{v}$ is due entirely to the change in position

$$\Delta v^\beta = -\Gamma^\beta_{\mu\lambda}\Delta^\mu v^\lambda$$

we can write the slope of $\mathbf{v}$ with respect to position Δ

$$\text{slope}_\mu v^\beta \approx \frac{\Delta x^\beta}{\Delta^\mu} = -\Gamma^\beta_{\mu\lambda} v^\lambda$$

Specializing to the spherical surface we have two slope equations for slopes with respect to longitude ϕ as we stick to fixed latitude θ.

$$\begin{aligned} \text{slope}_\phi v^\theta &= cs\ v^\phi \\ \text{slope}_\phi v^\phi &= -(c/s)\ v^\theta \end{aligned}$$

These are "coupled" slope equations which we must separate into pure equations in one variable each. We can do that by taking the slope again and then using the other equation to eliminate.

$$\begin{aligned} \text{slope}_\phi\text{slope}_\phi v^\theta &= cs\ \text{slope}_\phi v^\phi \\ &= -c^2 v^\theta \\ \text{slope}_\phi\text{slope}_\phi v^\phi &= -(c/s)\ \text{slope}_\phi v^\theta \\ &= -c^2 v^\phi \end{aligned}$$

Both equations are the same so the two variables, v^θ and v^ϕ, have the same behaviour. From the first picture above we can see that this behaviour is oscillatory—a full cycle in the traversal at the North Pole, half a cycle in the traversal at the tropic of cancer, and zero cycles when we traverse the Equator. We could try $e^{ic\phi}$ (or $e^{-ic\phi}$) as a solution for each of the variables $v = v^\theta(\phi)$ and $v = v^\phi(\phi)$ but we would like to keep the numbers one-dimensional so we try instead

$$\begin{aligned} v^\theta &= A\cos c\phi + B\sin c\phi \\ v^\phi &= C\cos c\phi + D\sin c\phi \\ \text{slope}_\phi v^\theta &= -cA\sin c\phi + cB\cos c\phi \\ \text{slope}_\phi v^\phi &= -cC\sin c\phi + cD\cos c\phi \end{aligned}$$

where I've shown the slopes, too, enough to make clear that the second slopes will just give back $-c^2$ times the original variables.

Each of these solutions has two undetermined constants. If you take the second slopes you'll see that the constants don't get pulled out of the original proposed solution, so the slopes alone cannot determine the constants. (Typically, a slope equation has one undetermined constant for each slope.)
We must use starting conditions to fix the constants. We decided to start at $\phi = 0$ the International Date Line, with an initial pointer

$$\begin{pmatrix} v_0^\theta \\ v_0^\phi \end{pmatrix}$$

So setting $\phi = 0$ in the solutions we see that

$$\begin{aligned} A &= v_0^\theta \\ C &= v_0^\phi \end{aligned}$$

and using the slopes of the solutions and the original equations we also find (hint: both sides of the equations below were multiplied by c)

$$\begin{aligned} B &= sv_0^\phi \\ D &= -\frac{1}{s}v_0^\theta \end{aligned}$$

So now we know the pointer at any position, assuming that we have walked along a fixed latitude starting at the International Date Line, with initial pointer. Let's say it is initiallly pointing straight ahead of us along the line of latitude:

$$\begin{pmatrix} v_0^\theta \\ v_0^\phi \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Embodying these solutions into the function `fixedLatCircle(noTheta,noPhi,v0)` and invoking `fixedLatCircle(7,17,[1;0])` gives the arrows shown above, a 7 by 17 grid. (Hint: multiply the equation for $v^\phi(\phi)$ by $s = \sin\theta$ to show arrows of the same length at all latitudes on the equiangular projection.)
We can find the angular difference between the initial and final position $\phi = 2\pi$ at each latitude.

$$\begin{aligned} \cos\alpha &= \frac{g_{jk}v_0^j v^k(2\pi)}{g_{jk}v_0^j v_0^k} \\ |\,\mathbf{v}_0\,|^2\cos\alpha &= v_0^\theta v^\theta(2\pi) + s^2 v_0^\phi v^\phi(2\pi) \\ &= (v_0^\theta)^2\cos 2\pi c + v_0^\theta v_0^\phi s\,\sin 2\pi c + s^2(-v^\phi v^\theta(\sin 2\pi c)/s + (v_0^\phi)^2\cos 2\pi c) \\ &= ((v_0^\theta)^2 + s^2(v_0^\phi)^2)\cos 2\pi c \\ &= |\,\mathbf{v}_0\,|^2\cos 2\pi c \end{aligned}$$

so the angle the pointer finally winds up pointing at is

$$\alpha = 2\pi c$$

The excess angle, because we've gone around a full circle each time, with "interior angles" summing to 2π, is 2π minus this

$$\Delta = 2\pi(1-c) = 2\pi(1-\cos\theta_0)$$

In the above calculations I have set the radius of the sphere to 1 because it does not figure in either the plot or the angle calculation. For the final step, next, I set it to r, so the metric now

$$g_{\mathrm{dd}} = \begin{pmatrix} r^2 & \\ & r^2 s^2 \end{pmatrix}$$

To check the curvature (which is constant on a sphere) we need the area contained by this constant-latitude circle with the North Pole at the centre.

$$\begin{aligned} \text{area} &= \text{antislope}_{\phi=0:2\pi}\text{antislope}_{\theta=0:\theta_0}\sqrt{g} \\ &= r^2\text{antislope}_{\phi=0:2\pi}\text{antislope}_{\theta=0:\theta_0}\sin\theta \\ &= 2\pi r^2(-\cos\theta)_0^{\theta_0} \\ &= 2\pi r^2(1-\cos\theta_0) \end{aligned}$$

and the ratio of these two is the curvature, $1/r^2$.

42. **Local flatness.** In Note 17 I claimed that curved space is "locally flat" in the sense that the slope of the metric vanishes, and so, then, do the affine connections or Christoffel symbols. (It is in fact the definition of a *manifold* that we can always find a transformation to a coordinate system in which the slope of the metric vanishes. We'll deal always with manifolds.)
The way to do this is to measure the coordinates along geodesics.
In Note 13 we gave the parallel transport that creates a geodesic as

$$x'^{\alpha} = (\delta^{\alpha}_{\lambda} - \Gamma^{\alpha}_{\nu\lambda}\Delta^{\nu})x^{\lambda}$$

where Δ^{ν} is a small displacement in the ν-direction from the start point of the geodesic.
(In fact, if we choose orthogonal geodesics we can make an orthonormal basis so that the metric is just the identity matrix up to first order in displacements.)
We'll take this as the transformation from new local coordinates $\overline{x}^{\lambda} = \Delta^{\lambda}$ in the basis of the geodesics, to original local coordinates x^{α} which are small changes in the global coordinate system

$$x^{\alpha} = (\delta^{\alpha}_{\lambda} - \Gamma^{\alpha}_{\nu\lambda}\overline{x}^{\nu})\overline{x}^{\lambda}$$

We can call this the transformation (matrix)

$$X^{\alpha}_{\lambda} = \delta^{\alpha}_{\lambda} - \Gamma^{\alpha}_{\nu\lambda}\overline{x}^{\nu}$$

If the geodesic in the original coordinate frame is $g_{\alpha\beta}$ then its transformation is

$$\overline{g}_{\lambda\mu} = X^{\alpha}_{\lambda}X^{\beta}_{\mu}g_{\alpha\beta}$$

which we can see easily in matrix terms

$$\begin{aligned} (\overline{x}^1, \overline{x}^2)\overline{g_{\mathrm{dd}}}\begin{pmatrix} \overline{x}^1 \\ \overline{x}^2 \end{pmatrix} &= \Delta S \\ &= (x^1, x^2)g_{\mathrm{dd}}\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \\ &= (\overline{x}^1, \overline{x}^2)X^{\mathrm{T}}g_{\mathrm{dd}}X\begin{pmatrix} \overline{x}^1 \\ \overline{x}^2 \end{pmatrix} \end{aligned}$$

So $\overline{g_{\mathrm{dd}}} = X^{\mathrm{T}}g_{\mathrm{dd}}X$ and, of course, in index notation transposition is conveyed simply by the indices.

We expand $g_{\rm dd}$, using a Taylor's series, in terms of small displacements $\overline{x}^\gamma$ about the point P we've chosen as origin

$$g(P+\overline{x})_{\alpha\beta} = g(P)_{\alpha\beta} + \partial_\nu g(P)_{\alpha\beta}\overline{x}^\nu + \mathcal{O}(\overline{x}^2)$$

(If we write any function $f(P+x)$ as a series in powers of x

$$f(P+x) = f(P) + ax + bx^2 + cx^3 + ..$$

we can find successive slopes with respect to x and then let $x \to 0$

$$\begin{aligned} \text{slope} f(P+x) &= a + 2bx + 3cx^2 + .. \to a \\ \text{slope slope} f(P+x) &= 2b + 6cx + .. \to 2b \\ \text{slope slope slope} f(P+x) &= 6c + .. \to 3!c \\ &\vdots \end{aligned}$$

and so we get the coefficient a for the first order in x, b for second order, and so on.)
Here is the whole transformation, taken to first order in $\overline{x}^\nu$

$$\begin{aligned} \overline{g}_{\lambda\mu} &= X^\alpha_\lambda X^\beta_\mu g_{\alpha\beta} \\ &= (\delta^\alpha_\lambda - \Gamma^\alpha_{\nu\lambda}\overline{x}^\nu)(\delta^\beta_\mu - \Gamma^\beta_{\nu\mu}\overline{x}^\nu)(g(P)_{\alpha\beta} + \partial_\nu g(P)_{\alpha\beta}\overline{x}^\nu + ..) \\ &= g(P)_{\lambda\mu} + [\partial_\nu g(P)_{\lambda\mu} - \Gamma^\alpha_{\nu\lambda} g(P)_{\alpha\mu} - \Gamma^\beta_{\nu\mu} g(P)_{\beta\lambda}]\overline{x}^\nu + .. \end{aligned}$$

and the term in square brackets, multiplying $\overline{x}^\nu$, vanishes because of the geodesic.
We can illustrate this for the surface of a sphere.

$g_{\rm dd}(\alpha$	β	g	$)$
θ	θ	r^2	
ϕ	ϕ	$r^2 s^2$	

$\Gamma_{\rm udd}(\alpha$	β	γ	Γ	$)$
θ	ϕ	ϕ	$-sc$	
ϕ	θ	ϕ	c/s	
ϕ	ϕ	θ	c/s	

By analogy with the discussion of geodesics in Note 16 we have the transformation as a matrix

$$\begin{array}{cc} & \begin{array}{cc} \lambda & \\ 1 & 2 \end{array} \\ \begin{array}{c} \alpha \\ \\ \end{array} & \begin{array}{c} 1 \\ 2 \end{array} \left(\begin{array}{cc} 1 & cs\overline{x}^2 \\ -\overline{x}^2 c/s & 1 - \overline{x}^1 c/s \end{array} \right) \end{array}$$

and the transformation of $g_{\rm dd}$ is

$$\left(\begin{array}{cc} 1 & -\overline{x}^2 c/s \\ cs\overline{x}^2 & 1 - \overline{x}^1 c/s \end{array} \right) \left(\begin{array}{cc} r^2 & \\ & r^2(s^2 + 2\overline{x}^1 cs) \end{array} \right) \left(\begin{array}{cc} 1 & cs\overline{x}^2 \\ -\overline{x}^2 c/s & 1 - \overline{x}^1 c/s \end{array} \right)$$

Note the first-order term added to $s^2 = \sin^2(\theta)$ to take it to $\sin^2(\theta + \overline{x}^1)$.
Worked out to first order this gives

$$\left(\begin{array}{cc} r^2 & \\ & r^2 s^2 \end{array} \right)$$

which is the metric we started with: the two infinitesimal directions do lie along geodesics, and they are orthogonal, diagonalizing the metric. The basis vectors are not normalized to unit length, however, so the metric is not the identity matrix. The result is independent of $\overline{x}$ to first order, so the first slope of the metric with respect to the parameter θ is zero, so the spherical surface is locally flat in these coordinates.
You can show that the metric is also unchanged and has zero first slope for plane polar

coordinates

$g_{\mathrm{dd}}(\alpha$	β	g	$)$
r	r	1	
θ	θ	r^2	

$\Gamma_{\mathrm{udd}}(\alpha$	β	γ	Γ	$)$
θ	θ	r	$1/r$	
θ	r	θ	$1/r$	
r	θ	θ	$-r$	

and for the hyperboloid of revolution (see Note 18)

$g_{dd}(\alpha$	β	g	$)$
p	p	$\cosh(2p)$	
θ	θ	$\cosh(p)^2$	

$\Gamma\_\mathrm{udd}(\alpha$	β	γ	Γ	$)$
p	p	p	$\tanh(2p)$	
p	θ	θ	$-\frac{1}{2}\tanh(2p)$	
θ	p	θ	$\tanh(p)$	
θ	θ	p	$\tanh(p)$	

It is interesting to take the above expansions to second order. The metrics we start and end with are (now omitting the overlines on the x-coordinates):
For the surface of a sphere ($c = \cos\theta, s = \sin\theta, c_2 = \cos 2\theta$):

$$\begin{pmatrix} r^2 & \\ & r^2(s^2 + 2csx^\theta + 2c_2 x^\theta x^\theta) \end{pmatrix} \qquad \begin{pmatrix} 1 + c^2 x^\phi x^\phi & -c^2 x^\theta x^\phi \\ -c^2 x^\theta x^\phi & s^2 + c^2 s^2 x^\phi x^\phi - (1+s^2) x^\theta x^\theta \end{pmatrix}$$

For plane polar coordinates:

$$\begin{pmatrix} r^2 & \\ & r^2 + 2rx^r + 2x^r x^r \end{pmatrix} \qquad \begin{pmatrix} 1 + x^\theta x^\theta & -x^r x^\theta \\ -x^r x^\theta & r^2 + r^2 x^\theta x^\theta + x^r x^r \end{pmatrix}$$

And for the hyperboloid of revolution ($c = \cosh p, s = \sinh p, c_2 = \cosh 2p$):

$$\begin{pmatrix} c_2 + 2s_2 x^p + 2c_2 x^p x^p & \\ & c^2 + 2csx^p + c_2 x^p x^p \end{pmatrix} \qquad \begin{pmatrix} c_2 + ((2-s_2^2)/c_2) x^p x^p + s^2 x^\theta x^\theta & ((s_2^2/(2c_2) - s^2) x^p x^\theta \\ ((s_2^2/(2c_2) - s^2) x^p x^\theta & c^2 + (s_2^2/(4c_2)) x^\theta x^\theta + (1-s^2) x^p x^p \end{pmatrix}$$

All the resulting metrics are symmetrical, as they should be. This means that they can each be diagonalized. The eigenvalues, λ, of a symmetrical 2-by-2 matrix satisfy

$$\begin{pmatrix} A & B \\ B & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

So

$$\begin{pmatrix} A - \lambda & B \\ B & D - \lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

which requires the determinant to vanish

$$\begin{aligned} 0 &= \begin{Vmatrix} A - \lambda & B \\ B & D - \lambda \end{Vmatrix} \\ &= \lambda^2 - (A+D)\lambda + AD - B^2 \end{aligned}$$

So

$$\lambda = \frac{1}{2}\left((A+D) \pm \sqrt{(A-D)^2 + 4B^2} \right)$$

These two possible values for λ are the two elements in the diagonalized matrix. Note that in each of the three matrices B is quadratic in x, in fact involves a product of two different xs. So the $4B^2$ term will be quartic. We'll limit the diagonal matrices to quadratic in x, so the B terms vanish and we're left with the matrix

$$\begin{pmatrix} A & \\ & D \end{pmatrix}$$

in each case: just the diagonal elements of the metrics shown above to second order.
What are the eigenvectors corresponding to these two eigenvalues, and hence what is the orthogonal matrix (its transpose is its inverse) that diagonalizes the metric?
In each case we can write

$$\overline{g}_{\alpha\beta} = g(<\text{origin}>)_{\alpha\beta} + Q_{\alpha\beta\mu\nu}x^\mu x^\nu$$

where $g(<\text{origin}>)$ means $g()$ evaluated at $x^\mu = 0$ and where the symbol $Q_{\alpha\beta\mu\nu}$ has the following meanings.

For sphere ($<\text{origin}>=\theta$)

Q_{dddd}				
$(\alpha$	β	μ	ν	Q $)$
θ	θ	ϕ	ϕ	r^2c^2
ϕ	ϕ	θ	θ	$r^2(c^2-2)$
ϕ	ϕ	ϕ	ϕ	r^2s^2

For polar ($<\text{origin}>=r$)

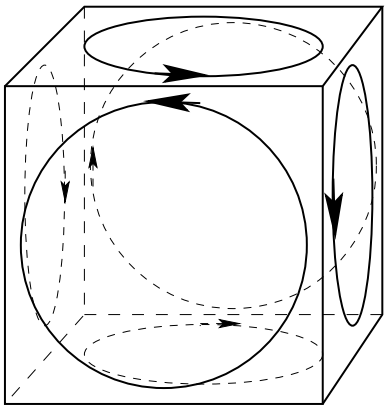
Q_{dddd}				
$(\alpha$	β	μ	ν	Q $)$
θ	θ	θ	θ	1
ϕ	ϕ	θ	θ	r^2
ϕ	ϕ	ϕ	ϕ	-1

For hyperboloid ($<\text{origin}>=p$)

Q_{dddd}				
$(\alpha$	β	μ	ν	Q $)$
p	p	p	p	$(2-s_2^2)/c^2$
θ	θ	p	p	$1-s^2$
θ	θ	θ	θ	$s_2^2/(4c_2)$

I've taken the diagonal form of the metric in each case: the nondiagonal forms will add two terms to Q each time. What are they?

43. **The boundary of boundary is zero.** In three dimensions, curvature can be calculated, as in Note 17, by the difference resulting from parallel transport of a vector around loops in the six faces of an infinitesimal cube.



We see that these loops counteract each other totally along the twelve edges of the cube because each edge borders two faces.
What does this say about total curvature, i.e., the sum of all six curvature results?
Note that the faces are the boundary of the cube and the edges are the boundaries of the boundary. See [Whe90, Ch.7] for a longer discussion.
In a four-dimensional infinitesimal hypercube, what is the boundary of the boundary, and what does it say about the sum of all curvatures?

44. a) Calculate the curvature tensor for spherical coordinates

$$(\Delta s)^2 = (\Delta r)^2 + r^2(\Delta\theta)^2 + r^2\sin\theta(\Delta\phi)^2$$

b) Calculate the curvature tensor for spherical coordinates limited to the surface of a sphere

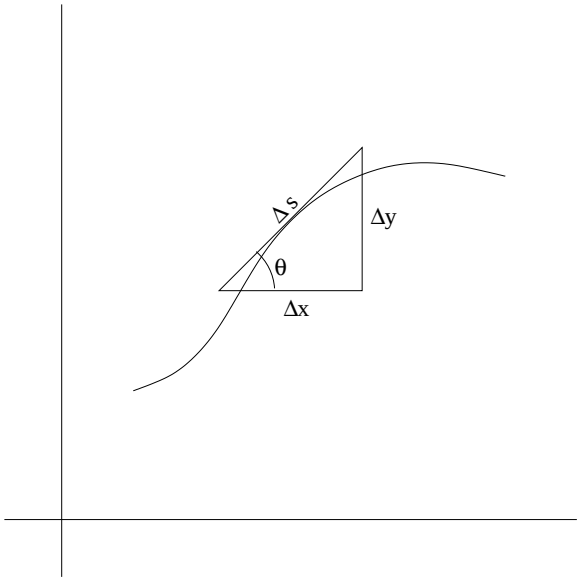
$$(\Delta s)^2 = r^2(\Delta\theta)^2 + r^2\sin\theta(\Delta\phi)^2$$

45. What is the rule for the number of vertices, edges, faces and other parts of a d-dimensional hypercube? (Columns to the right of the column headed d.)

Δ_d	$\mid R(d) \mid$	d	V	E	F		
0	0	0	1				
1	0	1	2	1			
3	1	2	4	4	1		
6	6	3	8	12	6	1	
10	20	4	16	32	24	8	1

The columns to the left are Δ_d, the triangular numbers, and $\mid R(d) \mid$, the number of different elements of the curvature tensor from Excursion *Number of components of the curvature tensor*.
Why do the number of different elements of the curvature tensor differ from the number of faces F?

46. **Curvature of lines.** The curvature of a line is defined as the rate of change, per unit length along the line, of the angle of the tangent to the line.



a) Show that the curvature of a circle is the reciprocal of its radius r.
The arc length subtended by angle θ is $s = r\theta$, so

$$\text{slope}_s\theta = 1/\text{slope}_\theta s = 1/r$$

b) Show that the length Δs along a small segment of any curve $y(x)$ is

$$\Delta s = \sqrt{1 + (\text{slope}_x y)^2}\Delta x$$

so

$$\text{slope}_x s = \sqrt{1 + (\text{slope}_x y)^2}$$

c) Similarly, show that for a parametric curve $x(p), y(p)$

$$\Delta s \approx \sqrt{(\text{slope}_p x)^2 + (\text{slope}_p y)^2}\Delta p$$

So

$$\text{slope}_p s = \sqrt{(\text{slope}_p x)^2 + (\text{slope}_p y)^2}$$

d) For curve $y(x)$ again show that

$$\begin{aligned}
\text{slope}_s\theta &= \text{slope}_s x \ \text{slope}_x\theta \\
&= \text{slope}_x\theta / \text{slope}_x s \\
&= \text{slope}_x(\text{atan}(\text{slope}_x y))/\text{slope}_x s \\
&= \frac{\text{slope}_x\text{slope}_x y}{1+(\text{slope}_x y)^2}\frac{1}{\sqrt{1+(\text{slope}_x y)^2}} \\
&= \frac{\text{slope}_x\text{slope}_x y}{(1+(\text{slope}_x y)^2)^{3/2}}
\end{aligned}$$

So the radius of the "osculating circle", which is the circle that matches the curvature of the curve at the point (x, y) under consideration, is

$$r = \frac{(1+(\text{slope}_x y)^2)^{3/2}}{\text{slope}_x\text{slope}_x y}$$

This is known as the radius of curvature, and is the reciprocal of the curvature.
e) For the parametric curve $x(p), y(p)$ show that

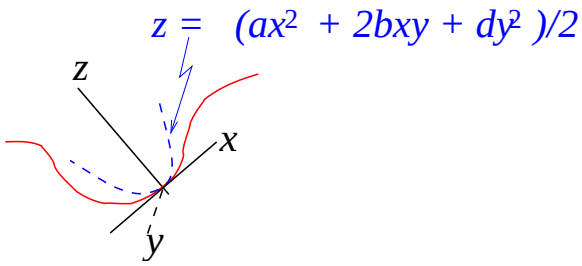
$$\begin{aligned}
\text{slope}_s\theta &= \text{slope}_s p \ \text{slope}_x\theta \\
&= \text{slope}_p\theta / \text{slope}_p s \\
&= \frac{\text{slope}_p \text{atan}(\text{slope}_p y/\text{slope}_p x)}{\text{slope}_p s} \\
&= \frac{\text{slope}_p(\text{slope}_p y/\text{slope}_p x)}{1+(\text{slope}_p y)^2}\frac{1}{\sqrt{(\text{slope}_p x)^2+(\text{slope}_p y)^2}} \\
&= \frac{\text{slope}_p^2 y \ \text{slope}_p x - \text{slope}_p^2 x \ \text{slope}_p y}{((\text{slope}_p x)^2+(\text{slope}_x y)^2)^{3/2}}
\end{aligned}$$

and the radius of curvature is the reciprocal of this.
f) Show that a line has no *intrinsic* curvature, i.e., a 1-dimensional creature living in the line can never tell if the line is curved or not.

47. For a curved surface in two dimensions, show that

$$z = \frac{\xi^2}{2\rho_1} + \frac{\eta^2}{2\rho_2}$$

for suitable internal (to the surface) coordinates (ξ, η), where ρ_1 and ρ_2 are the radii of the two osculating circles [MTW73, p.335].
Here is a one-dimensional cross-section of the surface (red) showing the tangent plane (coordinates (x, y)) at a given point, and a paraboloid of revolution used to approximate the surface at that point.



a) We want to transform (x, y) to (ξ, η) so that $z = (k_1\xi^2 + k_2\eta^2)/2$ instead of $z = (ax^2 + bxy + dy^2)/2$. We rotate within the tangential plane by an angle θ using $c = \cos\theta$ and $s = \sin\theta$.

$$\begin{aligned} z &= \frac{1}{2}(x\ \ y)\begin{pmatrix} a & b \\ b & d \end{pmatrix} y \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \frac{1}{2}(\xi\ \ \eta)\begin{pmatrix} c & -s \\ s & c \end{pmatrix}\begin{pmatrix} a & b \\ b & d \end{pmatrix}\begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} \xi \\ \eta \end{pmatrix} \\ &= \frac{1}{2}(\xi\ \ \eta)\begin{pmatrix} ac^2 - 2bcs + ds^2 & (a-d)cs + b(c^2 - s^2) \\ (a-d)cs + b(c^2 - s^2) & as^2 + 2bcs + dc^2 \end{pmatrix}\begin{pmatrix} \xi \\ \eta \end{pmatrix} \end{aligned}$$

To make this diagonal we want

$$z = \frac{1}{2}(\xi\ \ \eta)\begin{pmatrix} k_1 & \\ & k_2 \end{pmatrix}\begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

so we need

$$0 = (a-d)cs + b(c^2 - s^2) = (a-d)s_2/2 + bc_2$$

where $s_2 = \sin(2\theta)$ and $c_2 = \cos(2\theta)$, so

$$\frac{s_2}{c_2} = \frac{2b}{d-a}$$

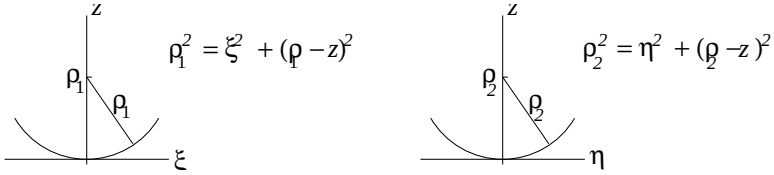
or

$$\theta = \operatorname{atan}\frac{2b}{d-a}$$

And, of course,

$$\begin{aligned} k_1 &= ac^2 - 2bcs + ds^2 \\ k_2 &= as^2 + 2bcs + dc^2 \end{aligned}$$

b) We want to show that $k_1 = 1/\rho_1$ and $k_2 = 1/\rho_2$, the radii of the two osculating circles, and hence that k_1 and k_2 are linear curvatures.
So, in directions ξ and η we approximate the surface by circles of radius ρ_1 and ρ_2 respectively.



The discussion is the same for each circle, so here it is for the ξ-circle with radius ρ_1.

$$\begin{aligned} 0 &= z^2 - 2\rho_1 z + \xi^2 \\ z &= \rho_1 \pm \rho_1\sqrt{1 - \xi^2/(2\rho_1^2)} \\ &\approx \rho_1(1 \pm (1 - \xi^2/(2\rho_1^2))) \\ &= \xi^2/(2\rho_1) \end{aligned}$$

where we choose the $-$ sign because that is the lower part of the circle seeking $z(\xi)$.
Combining ξ- and η- directions,

$$z = \frac{\xi^2}{2\rho_1} + \frac{\eta^2}{2\rho_2}$$

which, from (a), tells us $k_1 = 1/\rho_1$ and $k_2 = 1/\rho_2$, so k_1 and k_2 must be curvatures.
c) Show that the metric of the surface is

$$\begin{pmatrix} k_1^2\xi^2 + 1 & k_1 k_2 \xi\eta \\ k_1 k_2 \xi\eta & k_2^2\eta^2 + 1 \end{pmatrix}$$

using

$$(\Delta s)^2 = (\Delta z)^2 + (\Delta\xi)^2 + (\Delta\eta)^2 = (k_1\xi\Delta\xi + k_2\eta\Delta\eta)^2 + (\Delta\xi)^2 + (\Delta\eta)^2$$

48. In four dimensions, why are there six degrees of freedom for space transformations (three Lorentz, three rotation) and why must the Riemann curvature have twenty independent components [FMW03, pp.104,5]?
a) Show that a p-index transformation of d-dimensional variables $x^\alpha \to x'^\alpha$ has dS_d^p elements, where S_d^p is the number for the p-dimensional simplex of size d (see, e.g., Week ii, Note 6), that is

$$\begin{aligned} S_d^1 &= d &&: 1, 2, 3, 4, .. \\ S_d^2 &= \Delta_d &&: 1, 3, 6, 10, .. \\ S_d^3 &= \Diamond\!\!\!\!\!+_d &&: 1, 4, 10, 20, .. \end{aligned}$$

and so on (e.g., Week i Notes 1–6).
Since the transformation elements have the forms of partial slopes

$$\begin{aligned} p = 1 \text{ index} &\qquad \partial_{\alpha'} x^\mu \\ p = 2 \text{ indices} &\qquad \partial_{\alpha'}\partial_{\beta'} x^\mu \\ p = 3 \text{ indices} &\qquad \partial_{\alpha'}\partial_{\beta'}\partial_{\gamma'} x^\mu \end{aligned}$$

etc., we must count partial slopes while remembering that $\partial_{\alpha'}\partial_{\beta'} = \partial_{\beta'}\partial_{\alpha'}$.
Here it is in $d = 2$ dimensions.

$p = 1$ $\begin{matrix} \partial_{1'}x^1 & \partial_{2'}x^1 \\ \partial_{1'}x^2 & \partial_{2'}x^2 \end{matrix}$ $dS_d^1 = d \times d = 4$

$p = 2$

x^1	$\partial_{1'}$	$\partial_{2'}$
$\partial_{1'}$	$\surd$	$\surd$
$\partial_{2'}$		$\surd$

$dS_d^2 = d\Delta_d = 2 \times 3 = 6$
and similarly for x^2.

$p = 3$

x^1	$\partial_{1'}$	$\partial_{1'}$	$\partial_{2'}$
	$\partial_{1'}$	$\surd$	$\surd$
	$\partial_{2'}$		$\surd$

x^1	$\partial_{2'}$	$\partial_{1'}$	$\partial_{2'}$
	$\partial_{1'}$		
	$\partial_{2'}$		$\surd$

$dS_d^3 = d \Diamond\!\!\!\!\!+_d = 2 \times 4 = 8$
and similarly for x^2.

Once you've made the argument in general, here's a summary of dS_d^p for $d, p = 1..4$

$d\backslash p$	1	2	3	4
1	1	1	1	1
2	4	6	8	10
3	9	18	30	45
4	16	40	80	140

b) Show that the number of $(p-1)$st slopes of the d-dimensional metric tensor g_{dd} is $\Delta_d S_d^{p-1}$.
For $p=1$ we have the Δ_d components of the (symmetric) metric tensor: $S_d^0 = 1$.
For $p=2$, each of these components has $S_d^1 = d$ slopes.
For $p=3$, by the reasoning in (a), each component has $S_d^2 = \Delta_d$ slopes.
Check the generalization and then calculate

$d\backslash p$	1	2	3	4
1	1	1	1	1
2	3	6	9	12
3	6	18	36	60
4	10	40	100	200

c) When $d=4$ compare the bottom rows of the two tables

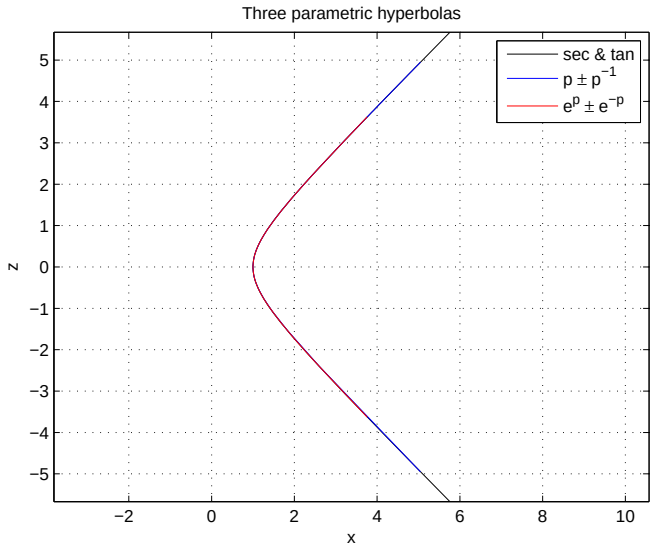
p	1	2	3	4
freedom to transform:	16	40	80	
metric components:	10	40	100	
net freedom:	6	0	-20	

Note that the Riemann curvature tensor, with 20 independent components, is formed by second slopes of the metric tensor.
Note that the 6 first slopes of the metric can in principle be transformed away: this gives us a locally flat space to work with.

49. **What functions for ch, sh?** The two symbols in Note 18, *ch* and *sh*, will be *functions* of some parameter, just as c and s (cos() and sin()) are functions of the *angle*. What functions? What parameter?
The choice of functions does not matter, in principle, just as long as they satisfy the constraint $ch^2 - sh^2 = 1$. In practice, one will turn out better than the rest. Let's explore. Each of the following satisfy the constraint on *ch* and *sh*.

parameter	ch	sh	ch^{-1}	sh^{-1}
$-\frac{\pi}{4} < \theta < \frac{\pi}{4}$	$\frac{1}{\cos(2\theta)} = \sec(2\theta)$	$\frac{\sin(2\theta)}{\cos(2\theta)} = \tan(2\theta)$	$\frac{1}{2}\text{asec}\, ch$	$\frac{1}{2}\text{atan}\, sh$
$0 < p < \infty$	$\frac{1}{2}(p + \frac{1}{p})$	$\frac{1}{2}(p - \frac{1}{p})$	$ch + sh = ch + \sqrt{ch^2 - 1}$	$ch + sh = sh + \sqrt{sh^2 + 1}$
$-\infty < p < \infty$	$\frac{1}{2}(e^p + e^{-p})$	$\frac{1}{2}(e^p - e^{-p})$	$\ln(ch + sh) = \ln(ch + \sqrt{ch^2 - 1})$	$\ln(ch + sh) = \ln(sh + \sqrt{sh^2 + 1})$

The overlapping plot shows that all give the same hyperbola: the exact dependence on the parameter does not matter. The only relationship that matters is $ch^2 - sh^2 = 1$.
(I've written the first parameter as a half-angle so it ranges between the asymptotes of the hyperbola. This does not guarantee that the point (ch, sh) is located at that angle from the origin.)



In all these cases, we can find functions corresponding to the other trigonometric functions:

$$\begin{aligned} th &= \frac{sh}{ch} && \tan \\ sch &= \frac{1}{ch} && \sec \\ csh &= \frac{1}{sh} && \csc \\ cth &= \frac{1}{th} && \cot \end{aligned}$$

and relationships such as

$$th^2 + sch^2 = 1$$

But we can only get an analogue of the trigonometric doubling relationships

$$\begin{aligned} c^2 - s^2 &= c_2 \\ 2cs &= s_2 \end{aligned}$$

and their generalizations (here c_2 and s_2 are $\cos(2\theta)$ and $\sin(2\theta)$ respectively) with the third option.

parameter	$ch^2 + sh^2$	$2ch \times sh$
$-\frac{\pi}{4} \leq \theta \leq \frac{\pi}{4}$	$\frac{1+\sin(2\theta)^2}{\cos(2\theta)^2}$	$2\frac{\sin(2\theta)}{\cos(2\theta)^2}$
$0 \leq p \leq \infty$	$ch(p^2)$	$sh(p^2)$
$-\infty \leq p \leq \infty$	$ch(2p)$	$sh(2p)$

This explains why the third choice of parameter and functions is used:

$$\begin{aligned} \cosh p &= \frac{1}{2}(e^p + e^{-p}) \\ \sinh p &= \frac{1}{2}(e^p - e^{-p}) \end{aligned}$$

Furthermore, the slopes of these hyperbolic functions remind us of trig.

$$\begin{aligned} \text{slope } \cosh(p) &= \sinh(p) \\ \text{slope } \sinh(p) &= \cosh(p) \end{aligned}$$

So this is the convenient parametrization.
Now that convenience—well, let's call it elegance—has dictated our choice of parametrization, what does the parameter mean? Not very obviously, and not very helpfully, it is the area OAP, shown in Note 18, between the line connecting the origin with the point and the hyperbola itself.
We can see this most easily by taking the antislope with respect to z (in the figure on p.3) of the difference (remember, $x^2 - z^2 = 1$)

$$x_{\text{on the hyperbola}} - z\frac{ch}{sh} = \sqrt{z^2+1} - z\frac{\sqrt{sh^2-1}}{sh}$$

This is taken from $z = 0$ to sh.
The antislope is (from a table of integrals then applying the limits then finding a term which cancels the second term above)

$$\ln(sh + \sqrt{sh^2+1}) = \ln(sh+ch) = p$$

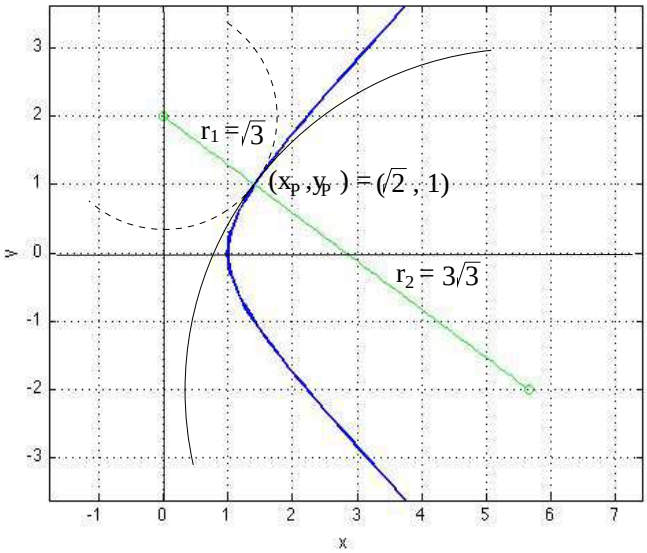
It is just the parameter we came up with and which gives the functions all those neat trig-like properties.

50. How far does the notion of negative curvature in Note 18 fit the tidal motion of Excursion *Tides* for the discussion of classical gravity in Note 7?

51. **Hyperboloid.** Show that the Gaussian curvature for the hyperboloid is the product of the reciprocals of its two radii of curvature

$$K = \frac{1}{r_1}\frac{1}{r_2}$$

a) The radius of curvature r_1 is the distance to the z-axis along a line normal to the hyperboloid.



Radii of curvature for hyperboloid of revolution

(Note in this figure that the dashed osculating circle is actually in a plane perpendicular to the page.)
Show that $r_1 = ch_2{}^{1/2} = \cosh(2p)^{1/2}$ in the notation of Note 18 ([Whe90, p.137] has a nice picture).

$$\begin{aligned} r &= \cosh p \\ z &= \sinh p \\ \Delta r &= \text{slope}_p r \Delta p = \sinh p \Delta p \\ \Delta z &= \text{slope}_p z \Delta p = \cosh p \Delta p \end{aligned}$$

So the equation of the straight line tangent to the hyperboloid is, for some a,

$$z = \frac{\Delta z}{\Delta r} r + a = r \coth p + a$$

And the equation of the normal—the line perpendicular to the hyperboloid—is

$$z = -\frac{\Delta r}{\Delta z} r + b = -r \tanh p + b$$

Since $(r, z) = (\cosh p, \sinh p)$ is on both of these lines, then in particular

$$z = 2 \sinh p$$

so the normal line is

$$z = -(r \cosh p + 2 \sinh p)$$

This crosses the z-axis at $(0, 2 \sinh p)$ giving distance to $(\cosh p, \sinh p)$

$$r_1^2 = \cosh^2 p + \sinh^2 p = \cosh(2p)$$

b) The other radius of curvature r_2 is clearly in the opposite direction. To find it we can use the general results in Excursion *Curvature of lines*.
We use the parametric form of the hyperbola, $x = \cosh p$ and $y = \sinh p$ to get

$$r_2 = \frac{(\sinh^2 p + \cosh^2 p)^{3/2}}{\sinh^2 p - \cosh^2 p} = -(\cosh(2p))^{3/2}$$

c) Thus

$$\frac{1}{r_1} \frac{1}{r_2} = \frac{1}{\cosh^2(2p)}$$

which is what we found for the Gaussian curvature K in Note 18.

52. **Paraboloid.** Show that the radii of curvature of the paraboloid of revolution generated by

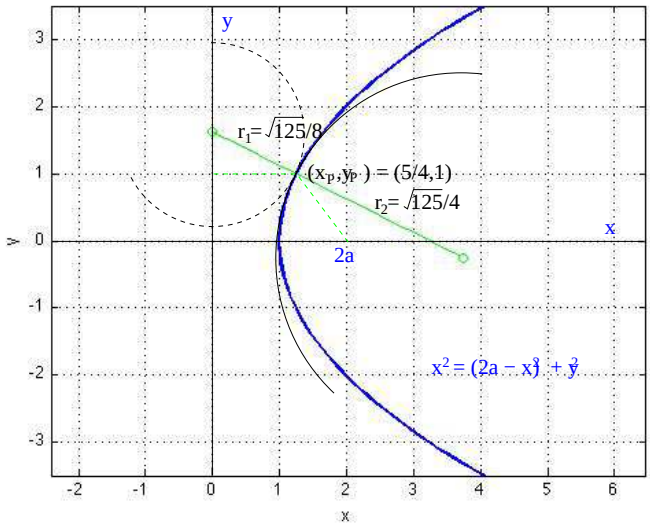
$$x = a \left(1 + \left(\frac{y}{2a} \right)^2 \right)$$

are

$$r_1 = a \left(1 + \left(\frac{y}{2a} \right)^2 \right)^{3/2}$$

and

$$r_2 = -2a \left(1 + \left(\frac{y}{2a} \right)^2 \right)^{3/2}$$



(Note in the figure that the dashed osculating circle is actually in a plane perpendicular to the page.)
a) The two radii are directed along the normal to the paraboloid. The centre of the osculating circle with radius r_1 is the axis of revolution $x = 0$.
The tangent to the parabola has slope $\text{slope}_x y$

$$\begin{aligned} 4ax &= y^2 + 4a^2 \\ 4a &= 2y\,\text{slope}_x y \\ \text{slope}_x y &= \frac{2a}{y} \end{aligned}$$

so the slope of the normal is the negative reciprocal of this and the equation of the normal line is

$$y = -\frac{y_P}{2a}x + b$$

where (x_P, y_P) is the point on the parabola where the normal crosses, and b can be found by setting $(x, y) = (x_P, y_P)$

$$b = y_P\left(1 + \frac{x_P}{2a}\right)$$

so

$$y = y_P\left(1 + \frac{x_P - x}{2a}\right)$$

When $x = 0$, on the axis of rotation,

$$y = b = y_P\left(1 + \frac{x_P}{2a}\right)$$

The distance from this point to (x_P, y_P) is

$$\begin{aligned} r_1^2 &= x_P^2 + \left(\frac{x_P y_P}{2a}\right)^2 \\ &= x_P^2\left(1 + \left(\frac{y_P}{2a}\right)^2\right) \\ &= a^2\left(1 + \left(\frac{y_P}{2a}\right)^2\right)^2\left(1 + \left(\frac{y_P}{2a}\right)^2\right) \end{aligned}$$

giving the first result above.
b) The second radius comes from the result for $y(x)$ in the Excursion *Curvature of lines*.

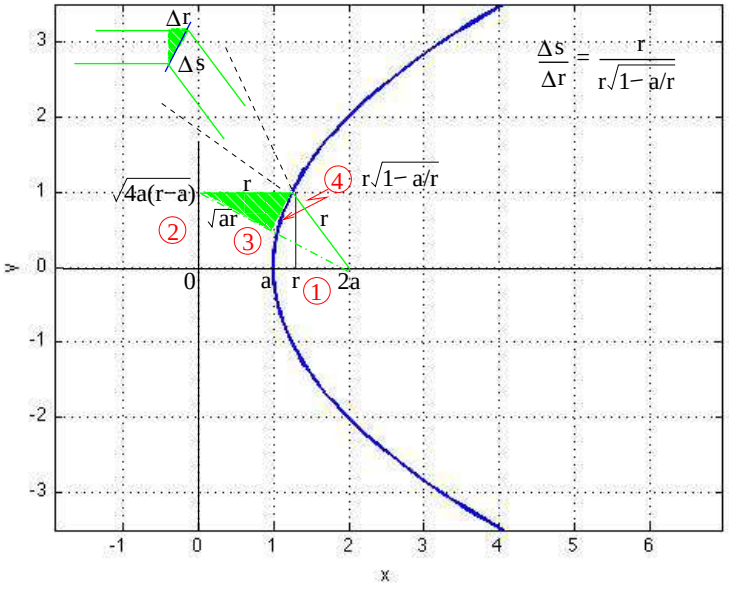
$$\begin{aligned} r_2 &= \frac{(1+(\text{slope}_x y)^2)^{3/2}}{\text{slope}_x \text{slope}_x y} \\ &= \frac{\left(1+\left(\frac{2a}{y}\right)^2\right)^{3/2}}{-\frac{2a}{y}\frac{2a}{y}} \\ &= -2a\frac{\left(1+\left(\frac{2a}{y}\right)^2\right)^{3/2}}{\left(\frac{2a}{y}\right)^3} \\ &= -2a\left(\left(\frac{y}{2a}\right)^2+1\right)^{3/2} \end{aligned}$$

giving the second result above.
Thus the Gaussian curvature

$$\begin{aligned} K &= \frac{1}{r_1 r_2} = \frac{1}{2r_1^2} \\ &= -\frac{1}{2a^2}\left(1+\left(\frac{y}{2a}\right)^2\right)^3 \\ &= -\frac{a}{2x^3} \end{aligned}$$

c) Calculate the metric to affine connection to curvature using the general results of Note 18 and show that this gives the same Gaussian curvature for the paraboloid of rotation.
First the metric. Here is Wheeler's argument [Whe90, pp.156–8], based on Apollonius, that

$$(\Delta s)^2 = \frac{(\Delta r)^2}{1-a/r}$$



The two shaded green triangles are similar, the smaller one being microscopic in size but formed between two sets of very close line pairs connecting the point P on the parabola a) to the axis of rotation (the lines in the pair are parallel) and b) to the focus (the lines in the pair are not parallel but are close enough together to be considered so).
The four steps in the figure are subtraction (1) followed by three Pythagorean triangles (2), (3), (4).
Including the dependence on the angle θ in the dimension perpendicular to the page, we have the metric

$$(\Delta s)^2 = \frac{(\Delta r)^2}{1 - a/r} + r^2(\Delta\theta)^2$$

From this, show that the affine connection is

$\Gamma_{\mathrm{udd}}(\alpha$	β	γ	Γ $)$
r	r	r	$\frac{a}{2r(a-r)}$
r	θ	θ	$a - r$
θ	r	θ	$1/r$
θ	θ	r	$1/r$

and that the Riemann curvature is

$R_{\mathrm{dddd}}(\alpha$	β	γ	δ	R $)$
r	θ	r	θ	$-a/2(a-r)$
r	θ	θ	r	$a/2(a-r)$
θ	r	r	θ	$a/2(a-r)$
θ	r	θ	r	$-a/2(a-r)$

so the Gaussian curvature is

$$\begin{aligned} K &= -\frac{R_{r\theta r\theta}}{g} \\ &= \frac{a/2/(a-r)}{r^2(1-a/r)} \\ &= -\frac{a}{2r^3} \end{aligned}$$

We relate this to part (b) above by noting that r and x have the same meaning in the two diagrams.
d) What are the curvatures in three dimensions?
Check first that what we've just solved is a two-dimensional problem, namely the surface of the paraboloid, even though we've shown an extraneous third dimension (the vertical dimension in the figures) so that we can visualize the paraboloid.
The two dimensions in the problem are

(a) the circular dimension in which we circle around the axis of rotation (symmetry) a distance r from it (in parts (a) and (b) this distance is called x)—the (θ, θ) component of the metric is r^2;

(b) the inward-outward dimension in which we approach or recede from the centre of symmetry along the paraboloid—the (r, r) component of the metric is $1/(1 - a/r)$.

The curvature of this two-dimensional space we have just seen to be

$$K = -\frac{a}{2r^3}$$

The three-dimensional extension we'd like to look at goes from circular symmetry to spherical symmetry. To keep in mind an extraneous dimension in which to embed the new, three-dimensional construct, we must think of an infinity of "axes of rotation" pointing in all directions so that we have circular symmetry in all directions.
Fortunately we don't have to visualize four dimensions (and the "fourth" dimension has nothing to do with time, by the way) because curvature always involves only two-dimensional elements.
So we must think, in the three "real" dimensions, of the curved space we're now considering, about a (cubic) sample volume of six equal (two-dimensional) faces.
The "circular" dimension of the two-dimensional paraboloid of the previous parts of this Excursion now becomes two circular dimensions, each maintaining constant distance r from the centre of symmetry. We'll imagine two of the faces of our sample volume as small parts of the spherical surfaces of radius r and $r + \Delta r$ respectively.
So the curvature of the closer of these faces is

$$K_{\text{facing}} = \frac{1}{r_1^2} = \frac{a}{r^3}$$

This leaves four more faces of our sample volume, each one sideways-on to the centre of symmetry. Since two edges of any of these faces is a constant distance from the centre, traversing them is along a circular dimension. The other two edges are in the inwards-outwards dimension, approaching and receding from the centre.
So each of the sideways-on faces has curvature

$$K_{\text{sideways}} = \frac{1}{r_1 r_2} = -\frac{a}{2r^3}$$

There are two of these sideways edges for each facing edge, so that the total curvature can be found from three faces.

$$\begin{aligned} K_{\text{total}} &= K_{\text{facing}} + 2K_{\text{sideways}} \\ &= \frac{a}{r^3} - 2\frac{a}{2r^3} \\ &= 0 \end{aligned}$$

This is an important property of the paraboloid of "two-dimensional revolution" (i.e., about a sphere) and plays the same role in gravity that the divergence-free $1/r^2$ field plays in classical 3-D gravity (Note 7).
e) Obtain these same three curvatures from the Ricci tensor.
Extend the 2-D metric of the paraboloid to 3-D by changing the circle (metric component r^2) to a sphere (metric components r^2, $r^2 \sin^2\theta$).
With metric

$$\begin{pmatrix} \frac{1}{1-a/r} & & \\ & r^2 & \\ & & r^2 \sin^2\theta \end{pmatrix}$$

find the affine connection

$\Gamma_{\mathrm{udd}}(\alpha$	β	γ	Γ $)$
r	r	r	$a/2r(a-r)$
r	θ	θ	$a-r$
r	ϕ	ϕ	$(a-r)\sin^2\theta$
θ	r	θ	$1/r$
θ	θ	r	$1/r$
θ	ϕ	ϕ	$-\cos\theta\sin\theta$
ϕ	r	ϕ	$1/r$
ϕ	ϕ	r	$1/r$
ϕ	θ	ϕ	$\cos\theta/\sin\theta$
ϕ	ϕ	θ	$\cos\theta/\sin\theta$

and the Ricci tensor

$R_{\mathrm{dd}}(\alpha$	β	R $)$
r	r	$\frac{a}{r^3(1-ra/r)}$
θ	θ	$-\frac{a}{2r}$
ϕ	ϕ	$-\frac{a}{2r}\sin^2\theta$

Transform to orthonormal coordinates (Excursion *Orthonormal polars*)

$$\begin{pmatrix} \sqrt{1-a/r} & & \\ & 1/r & \\ & & 1/(r\sin\theta) \end{pmatrix} \begin{pmatrix} \frac{a}{r^3(1-a/r)} & & \\ & -\frac{a}{2r} & \\ & & -\frac{a}{2r}\sin^2\theta \end{pmatrix} \begin{pmatrix} \sqrt{1-a/r} & & \\ & 1/r & \\ & & 1/(r\sin\theta) \end{pmatrix}$$
$$= \begin{pmatrix} \frac{a}{r^3} & & \\ & -\frac{a}{2r^3} & \\ & & -\frac{a}{2r^3} \end{pmatrix}$$

as required.
Note that the total curvature is the trace of this matrix.

53. Look up [Wei72, pp.144–6] for a derivation of the Riemann curvature tensor in terms of the curvature scalar, the Ricci tensor and the Weyl tensor in three or more dimensions (Note 19).

54. Implement the MATLAB functions `metr2curv()` and `contract()` of Note 20 and their auxiliary functions.

55. Look up Karl Friedrich Gauss (1777–1855). What is Gaussian curvature? What other ideas are named after Gauss—in numerical analysis, in statistics, in field theory, in magnetism?

56. Look up Georg Friedrich Bernhard Riemann (1826–1866), Gauss' student. What is Riemann curvature? What other ideas are named after Riemann—Riemann surface, Riemann's conjecture?

57. Apart from books cited in particular Excursions above, three other books I've learned from are [McC57, Che10] and [Har03].

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

References

[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.

[Fey97] Richard P Feynman. *Six Not-So-Easy Pieces*. Addison-Wesley Longmn Inc., Reading, MA, 1997.

[FLS64] R. P. Feynman, R. B. Leighton, and M. Sands. *The Feynman Lectures on Physics, Volume II*. Addison-Wesley, 1964.

[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.

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

[McC57] A. J. McConnell. *Applications of Tensor Analysis*. Dover Publictions, Inc., New York, 1957. Originally *Applications of the Absolute Differential Calculus*, 1931.

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

[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.