Excursions in Computing Science: Week iv. Space Math

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

May 24, 2022

I. Prefatory Notes

1. Matrix multiplication. Teacher, help your grade scholar master the multiplication of 2×2 matrices outlined below and then encourage hem to invent a few 2×2 matrices to exercise on. Try 3×3, 2×3, and other $n \times m$ matrices as well. A grade scholar who enjoys calculating will like this work for a while and will appreciate all the more the revelations later in these Notes of what matrices mean and how they can be applied.

Polynomials in Week iii add and subtract in fairly straightforward ways. They become more intriguing when multiplied, divided and factored. In these Notes we look at a quite different assemblage of numbers, the *matrix*.

A matrix is a rectangular array of numbers. We will focus on 2×2, square rectangles.

Here are two 2×2 matrices multiplied together.

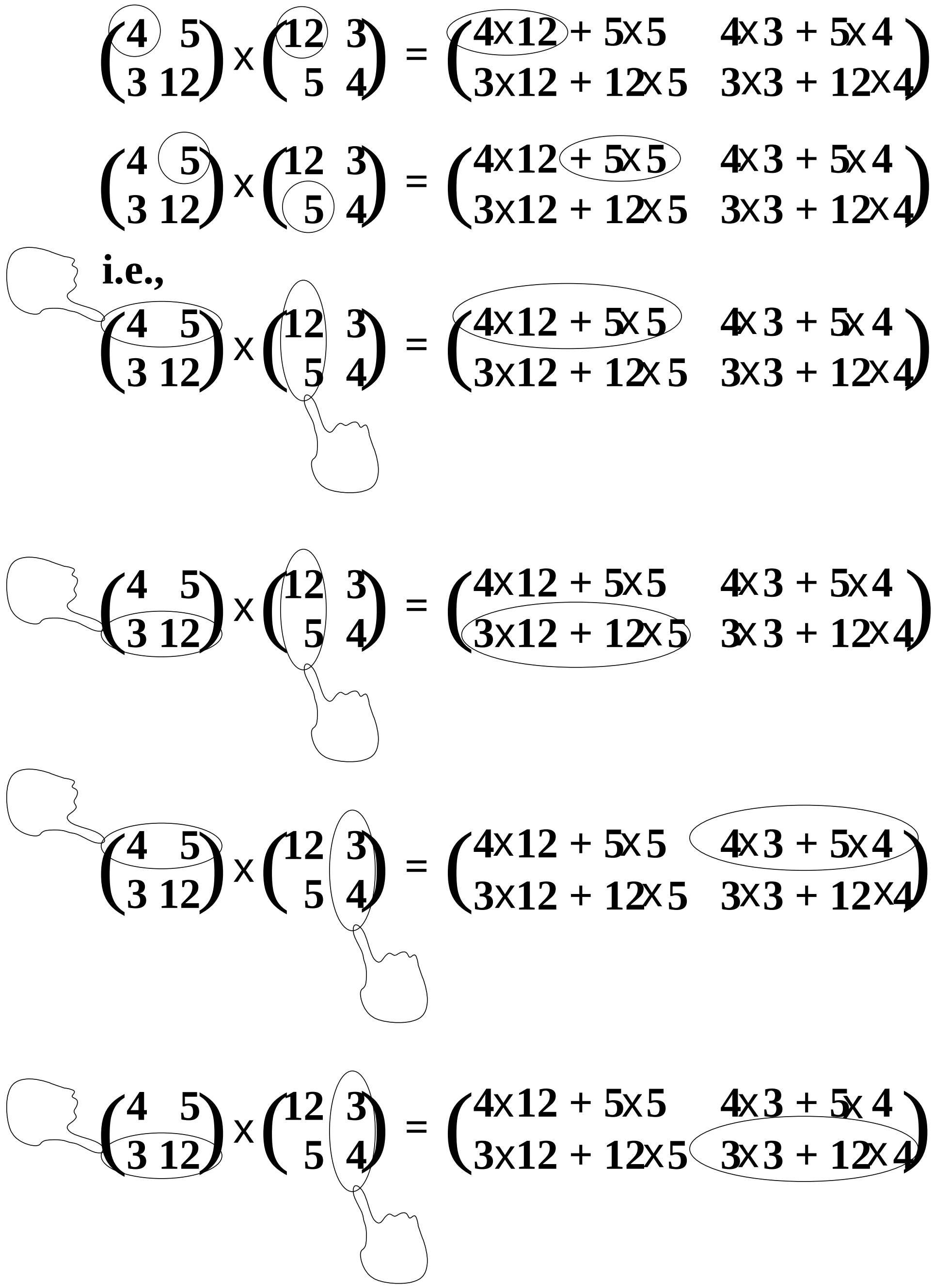
$$\begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} \times \begin{pmatrix} 12 & 3 \\ 5 & 4 \end{pmatrix} = \begin{pmatrix} 73 & 32 \\ 96 & 57 \end{pmatrix}$$

Here is how we get this answer.

$$\begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} \times \begin{pmatrix} 12 & 3 \\ 5 & 4 \end{pmatrix} = \begin{pmatrix} 4 \times 12 + 5 \times 5 & 4 \times 3 + 5 \times 4 \\ 3 \times 12 + 12 \times 5 & 3 \times 3 + 12 \times 4 \end{pmatrix}$$

A picture will help even more.

 The author gratefully acknowledges support from the taxpayers of Québec and of Canada who have paid his salary and research grants while this work was developed at McGill University, and from his students and their funding agencies. The present Notes are based on the teaching and ideas of Eugene Lehman.



Matrix multiplication is *not* necessarily commutative.

$$\begin{pmatrix} 12 & 3 \\ 5 & 4 \end{pmatrix} \times \begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} = \begin{pmatrix} 57 & 96 \\ 32 & 73 \end{pmatrix}$$

2. Vectors. Matrices do not have to be square. Here are two rather special 2×1 matrices.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Using the matrix multiplication rule

$$\begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

and

$$\begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 12 \end{pmatrix}$$

2×1 matrices are called *vectors*. (So are 1×2 matrices.)

The two vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

are special because any other vector can be made up from them.

$$\begin{pmatrix} x \\ y \end{pmatrix} = x \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \times \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

For example,

$$\begin{pmatrix} 4 \\ 3 \end{pmatrix} = 4 \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \times \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

This introduces two new operations on matrices: scalar multiplication and addition, both easy.

Scalar multiplication

$$2 \times \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

Addition

$$\begin{pmatrix} 4 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$$

3. Identity matrix. Notice how the first multiplication in Note 2 "selects" the first column of the matrix, and the second multiplication "selects" the second column.

We can actually lump together these two multiplications.

$$\begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix}$$

And, swapped

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix}$$

So we have a special square matrix, called the *identity*.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The identity matrix plays the same role in matrix multiplication that 1 does in number multiplication.

4. Matrix inverse. Given a matrix, what matrix multiplied by it gives the identity? This will be the *inverse* of the given matrix.

A fairly simple rule gives the inverse for a 2×2 matrix. The rule starts: Swap the diagonal elements and change the signs of the off-diagonal elements.

Try

$$\begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} \times \begin{pmatrix} 12 & -5 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} 33 & 0 \\ 0 & 33 \end{pmatrix}$$

This is almost the identity: we must just divide by 33.

Before we say what this 33 is, notice carefully just why the swap and the sign change give the off-diagonal zeroes in the result.

Try multiplying the two diagonal elements of the original matrix, then subtracting the product of the off-diagonal elements. This is called the *determinant* of the 2×2 matrix and in this case it is 33.

$$\begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} \times \frac{\begin{pmatrix} 12 & -5 \\ -3 & 4 \end{pmatrix}}{4 \times 12 - 5 \times 3} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

So the rest of the inversion rule is: Divide the new matrix by the determinant of the original matrix. Now you have the inverse of the original.

The convention is to use an exponent -1 to signify the inverse.

$$\begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} \times \begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

We do not usually talk about matrix division because the important operation is inversion, and inversion is enough to give us division.

$$\begin{pmatrix} 12 & 3 \\ 5 & 4 \end{pmatrix} \times \begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix}^{-1} = \begin{pmatrix} 135/33 & -48/33 \\ 48/33 & -9/33 \end{pmatrix}$$

is what we would mean if we could say

$$\begin{pmatrix} 12 & 3 \\ 5 & 4 \end{pmatrix} \div \begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix} = \begin{pmatrix} 135/33 & -48/33 \\ 48/33 & -9/33 \end{pmatrix}$$

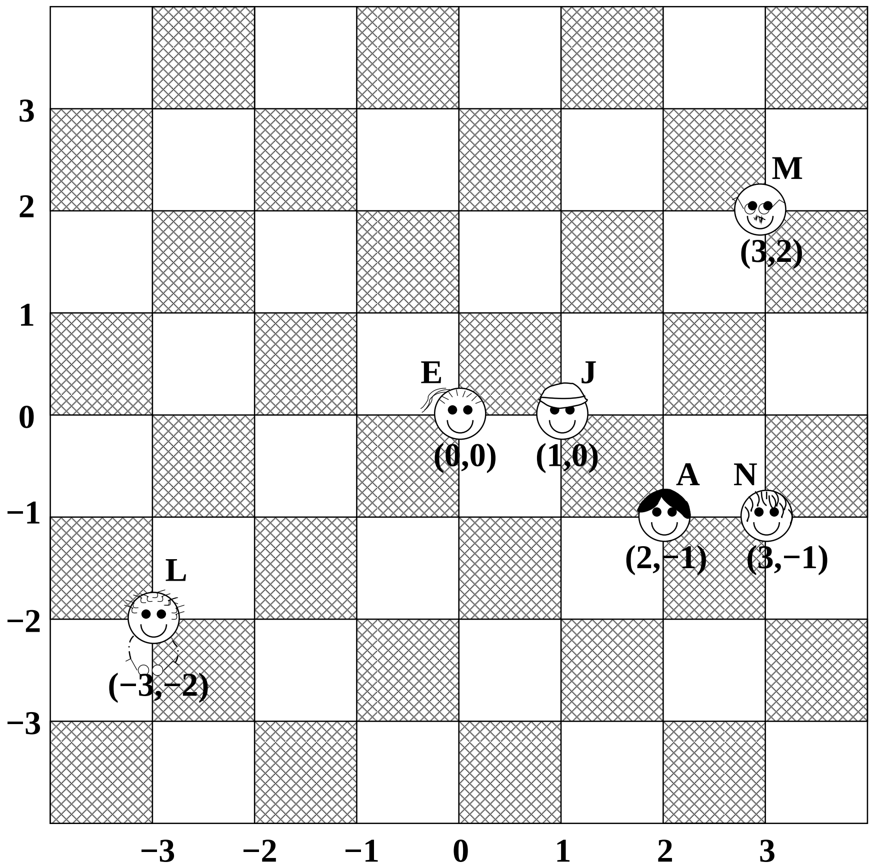
(By the way, the matrix we got here is in a special class: the off-diagonal element(s) above the diagonal differ only in their sign from their counterpart(s) below the diagonal. Can you see why this result had to turn out this way?)

If its determinant is zero, a matrix is not invertible. (Why?). Such a matrix is called *singular*. Singular matrices play the role in matrix "division" that 0 plays in number division. But note that there will be more than one singular matrix.

5. Vectors in space. Now let's see what all these matrices and their strange operations might mean and might be useful for.

We start with vectors, specifically the "column vectors" (2×1 matrices) we have been using. These are just pairs of numbers, and so are useful for working with two-dimensional space.

Here is a view from the ceiling of a classroom with a floor tiled with large dark and light linoleum tiles, and of the six people currently in the classroom. (It's not that they are all looking at the ceiling and not paying attention, but that I couldn't draw both the floor and the faces of the people at their desks in any other way.)



Everybody's name (one letter each) is also shown, and so are their positions (two numbers—a vector—each).

Positions must be measured from some starting point, and by convention they are all measured from the *origin*, the point (0,0).

So we had to show the origin and, for symmetry, it appears in the centre of the picture. It could be anywhere else, such as the bottom left-hand corner (a frequent scientific convention) or the top left-hand corner (the usual computer graphics convention) or somewhere completely outside the picture.

Wherever it has been put, the origin is the point of reference for all positions, hence its name.

Putting the origin at the centre of the picture allows us to show negative numbers on the same footing as positive.

Note how "M" is positioned 3 tile widths right of "E" (who happens to be sitting at the origin) and 2 tile heights above. So these two numbers form the two components of the position of "M".

"L" on the other hand is directly opposite "M" relative to the origin. By convention (again) rightwards and upwards are indicated by positive numbers and leftwards and downwards by negative. So "L"'s position consists of two numbers which are the respective negatives of "M"'s numbers, -3 and -2.

Vectors can be written either horizontally as 1×2 matrices or vertically as 2×1 matrices. It was convenient to write them horizontally in the picture but in the text we will stick to column vectors. These are more common than row vectors, and I myself have some trouble with left and right which I do not have with top and bottom. It is important to distinguish the first from the second element since the first element of a vector conventionally describes the left-right direction in space while the second describes the up-down direction. In row vectors the first element is the left one. In column vectors it is the top one.

Here are the vectors corresponding to the positions of the six people in the classroom. One other vector is added because we do not have anybody seated at the position given by the second special vector.

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \overset{\text{J}}{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} \quad \overset{\text{E}}{\begin{pmatrix} 0 \\ 0 \end{pmatrix}} \quad \overset{\text{N}}{\begin{pmatrix} 3 \\ -1 \end{pmatrix}} \quad \overset{\text{A}}{\begin{pmatrix} 2 \\ -1 \end{pmatrix}} \quad \overset{\text{L}}{\begin{pmatrix} -3 \\ -2 \end{pmatrix}} \quad \overset{\text{M}}{\begin{pmatrix} 3 \\ 2 \end{pmatrix}}$$

These seven vectors can also be lumped into a single 2×7 matrix.

$$\begin{pmatrix} 0 & 1 & 0 & 3 & 2 & -3 & 3 \\ 1 & 0 & 0 & -1 & -1 & -2 & 2 \end{pmatrix}$$

6. Positions and intervals. So far the vectors just stand for positions in space. They can also stand for intervals.

For example, the interval from “A” to “N” is N − A:

$$\begin{pmatrix} 3 \\ -1 \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Note that this is the same vector as the position of “J”.

So vectors representing intervals also represent them relative to the origin: they don’t start at the first position.

We would really need four numbers to give both the interval and its starting point. But we already have these four numbers in the two vectors N and A. So it is economical just to take the two numbers in N − A as the interval. But this can be a confusing convention and takes getting used to.

A similar convention also holds when we interpret ordinary numbers as positions along a line (such as the Celcius temperature scale) or as intervals on the line (such as how much the temperature went up today (positive interval) or down last night (negative interval)).

Thus we can interpret addition and subtraction of vectors. Two vectors representing positions can be subtracted to give the vector representing the interval between. Two vectors representing position and interval respectively can be added to give the new position (again a vector) that is the given interval away from the first position.

7. Transforming space. How can we interpret multiplication? By the rule for matrix multiplication we cannot multiply two vectors (except only if the first is a row vector and the second a column vector, but we are sticking to column vectors): why?

So we must return to multiplying 2×2 matrices and column vectors.

Recall from Note 2 that the two special column vectors “select” the two columns of the matrix when multiplied by the matrix. But these special vectors just describe the intervals of one step (tile) rightward and one step upward in the classroom space. So we can easily read the effect of multiplying these special vectors by the matrix.

For an example, I’m going to modify the matrix a little from

$$\begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix}$$

by dividing the first column by 5 and the second by 13. (What is special about the triplets 3,4,5 and 5,12,13? The answer may give a hint about why I am making this change, but it will not become clear until Week 2.)

$$\begin{pmatrix} 4/5 & 5/13 \\ 3/5 & 12/13 \end{pmatrix} = \begin{pmatrix} 0.8 & 0.38 \\ 0.6 & 0.92 \end{pmatrix}$$

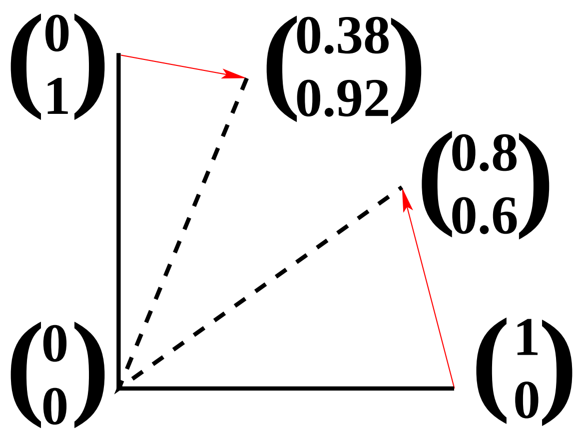
Here is the effect of this matrix on the left-right unit vector (a “unit” vector has length 1)

$$\begin{pmatrix} 0.8 & 0.38 \\ 0.6 & 0.92 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix}$$

and here is the effect on the up-down unit vector

$$\begin{pmatrix} 0.8 & 0.38 \\ 0.6 & 0.92 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.38 \\ 0.92 \end{pmatrix}$$

and we see that multiplying by the matrix has had the effect of bending the left-right unit vector upwards to a new vector, and bending the up-down unit vector rightwards to another new vector.



The red arrows show the changes to the special vectors.

Recall also from Note 2 that any vector is a combination of the two special vectors. So any vector is part left-right unit vector and part up-down unit vector. The left-right part will be bent by the matrix multiplication in the way we have just seen. The up-down part will be bent leftwards as we also saw.

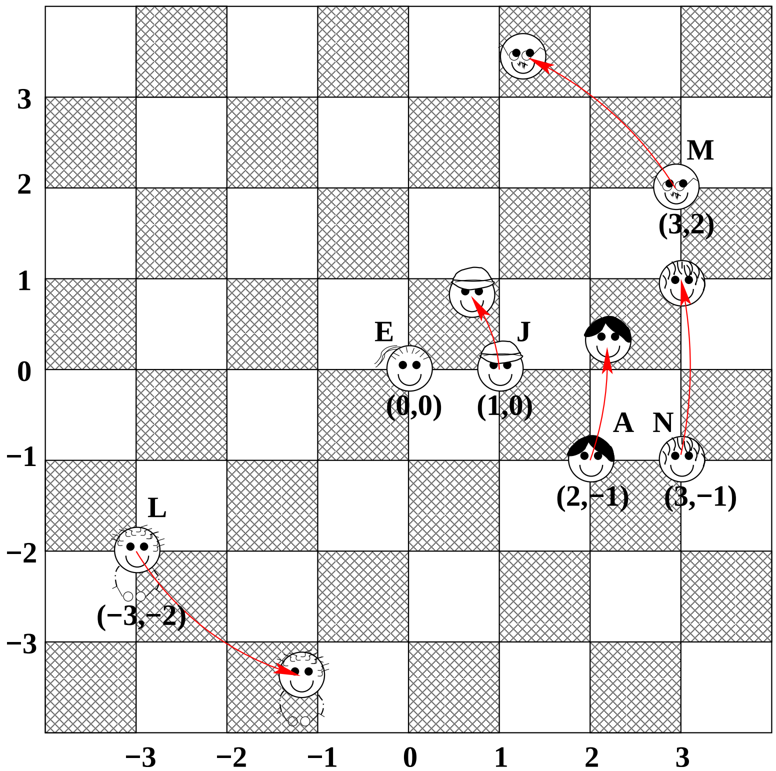
8. Rotations.
To turn, turn will be our delight,
Till by turning, turning we come round right.
Joseph Brackett (1797-1882)

Let’s look at this in a special case, the matrix

$$\begin{pmatrix} 4/5 & -3/5 \\ 3/5 & 4/5 \end{pmatrix} = \begin{pmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{pmatrix}$$

Multiply every position in the classroom by this matrix and see where everybody moves to. (I’ll use the lumped vectors, the 2×7 matrix, to write this more compactly.)

$$\begin{pmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{pmatrix} \times \begin{pmatrix} 0 & 1 & 0 & 3 & 2 & -3 & 3 \\ 1 & 0 & 0 & -1 & -1 & -2 & 2 \end{pmatrix} = \begin{pmatrix} -.6 & .8 & 0 & 3 & 2.2 & -1.2 & 1.2 \\ .8 & .6 & 0 & 1 & 0.4 & -3.4 & 3.4 \end{pmatrix}$$



9. Shear. All rotation matrices have determinant 1. (Check this for Note 8.) We can find other matrices which also have determinant 1.

An example is a shear matrix. The matrix we started with in Note 1 shears space as we saw in Note 7: it squeezes it in one direction and lets it squirt out in another direction, like a toothpaste tube. That matrix also does other things to the space so let's see if we can purify the notion of shear.

First, we can make the distortion symmetrical. This takes a "symmetric" matrix, such as

$$\begin{pmatrix} 4/5 & 3/5 \\ 3/5 & 4/5 \end{pmatrix} = \begin{pmatrix} 0.8 & 0.6 \\ 0.6 & 0.8 \end{pmatrix}$$

However, the determinant is no longer 1. (What is it?)

To get determinant 1 for the symmetric matrix

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix}$$

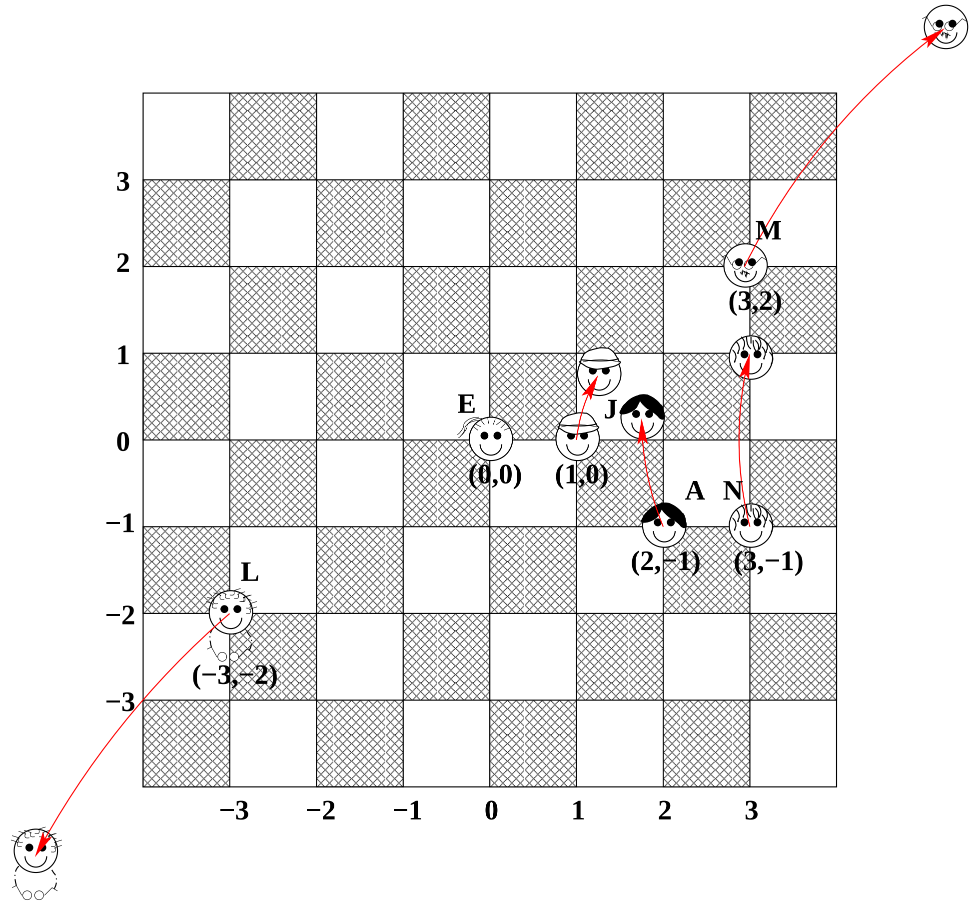
we need $a^2 - b^2 = 1$. This can also be done with a Pythagorean triple.

For example, $(5/4)^2 - (3/4)^2 = 1$, so

$$\begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix} = \begin{pmatrix} 1.25 & 0.75 \\ 0.75 & 1.25 \end{pmatrix}$$

is a symmetric, det=1 matrix. We call such matrices (pure) "shear" matrices.

Here is the effect on the classroom space.



10. Diagonalizing matrices. The shear matrix of the previous Note appears to stretch the space in the direction half-way between the horizontal x and the vertical y axes. And it shrinks the space in the direction half-way between the vertical y axis and the horizontal $-x$ axis.

In fact, calling these directions

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

respectively, we can see that the "stretch" is a doubling and the "shrink" is a halving.

$$\begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$\begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1/2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

This leads us to wonder if *rotating* the shear matrix could be thought of as producing a pure stretch and a pure shrink in these new directions respectively.

Let's look at the shear matrix S and the rotation matrix R formally. The effect of S on a vector v is Sv, the product. Similarly the effect of R on v is Rv. Now let's rotate the sheared vector.

$$R(Sv) = RSv = RSIv = RS(R^{-1}R)v = (RSR^{-1})(Rv)$$

where I've supposed that R has an *inverse* R^{-1} (Note 4). I've stuck in that inverse multiplied by R because the product is the identity I and changes nothing. The reason I've done this is to find out the effect of rotation on the matrix S.

This effect is that we can rotate the matrix S by multiplying: RSR^{-1}.

We can see this by noting that $R(Sv)$ is the the rotated result of the shear, and the "rotated shear" is the effect of the shear on the rotated v, i.e., on Rv. And the math shows that this rotated shear must be RSR^{-1}.

Let's try it. To rotate the x and y axes in the new directions

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

we need the matrix

$$R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

(I've divided everything by $\sqrt{2}$ because this matrix must be a rotation so its determinant must be 1. Or, $c^2 + s^2 = 1$: what are c and s in this case?)

Let's do it.

$$\begin{aligned} RSR^{-1} &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2/4 & 8/4 \\ -2/4 & 8/4 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 4/4 & 0 \\ 0 & 16/4 \end{pmatrix} \\ &= \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix} \end{aligned}$$

This rotation *diagonalizes* the matrix. It has the effect of transforming the axes to new axes along which the shear becomes pure stretch and pure shrink respectively.

(There is a subtle reason why we *should* have done the transformation $R^{-1}SR$. Try this. Is the new "x' axis stretched and the new "y" axie shrunk?)

11. Summary

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

1. Matrix multiplication.
2. Vectors.
3. Identity matrix.
4. Matrix inverse.
5. Vectors in space.
6. Positions and intervals.
7. Transforming space.
8. Rotations.
9. Shear.
10. Diagonalization.

II. The Excursions

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

1. "Transpose" the operations in Note 2 by rewriting each 2×1 matrix as a 1×2 matrix. When you transpose each matrix in a multiplication, note that the multiplication rule can no longer work. So you must also exchange the two matrices. Try

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \times \begin{pmatrix} 4 & 3 \\ 5 & 12 \end{pmatrix} = \begin{pmatrix} 4 & 3 \end{pmatrix}$$

Using these two ideas, rework all the matrix calculations in these Notes into their transposes.

2. What is

$$\frac{\begin{pmatrix} 12 & -5 \\ -3 & 4 \end{pmatrix}}{4 \times 12 - 5 \times 3} \times \begin{pmatrix} 4 & 5 \\ 3 & 12 \end{pmatrix}$$

3. Are the following matrices singular?

$$\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & 6 \\ 1 & 3 \end{pmatrix} \quad \begin{pmatrix} 2 & 8 \\ 1 & 4 \end{pmatrix}$$

What is the pattern? Do all singular matrices obey this pattern? Can any non-singular matrix obey it? How does this pattern transform space? (Draw the effect on the two special unit vectors.)

4. What is the condition that the determinant of the antisymmetric matrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

be 1? How can this be achieved by Pythagorean triples?

5. a) Write down a rotation matrix based on the Pythagorean triple 5, 12, 13.
b) Multiply this both ways with the rotation matrix from Note 8: does matrix multiplication "commute" for rotation matrices? (An operation, *, is *commutative* if $a * b = b * a$ for any a and b. In Note 1 we saw that matrix multiplication does *not* commute for arbitrary matrices.) Why should rotation matrices commute?
c) What are the inverses of these rotation matrices?
d) What is the vector that is twice the angle from horizontal as that made by

$$\begin{pmatrix} 4/5 \\ 3/5 \end{pmatrix}$$

What is the corresponding Pythagorean triple?

6. a) Show that $m^2 - n^2, 2mn$ and $m^2 + n^2$ is a Pythagorean triple for any two integers m and n.
b) Write a MATLAB program to find all the Pythagorean triples up to m and n equal to some given upper value.
c) Run your program, say for m and n up to 9. How many of the triples you get are not multiples of some other triples in your list?

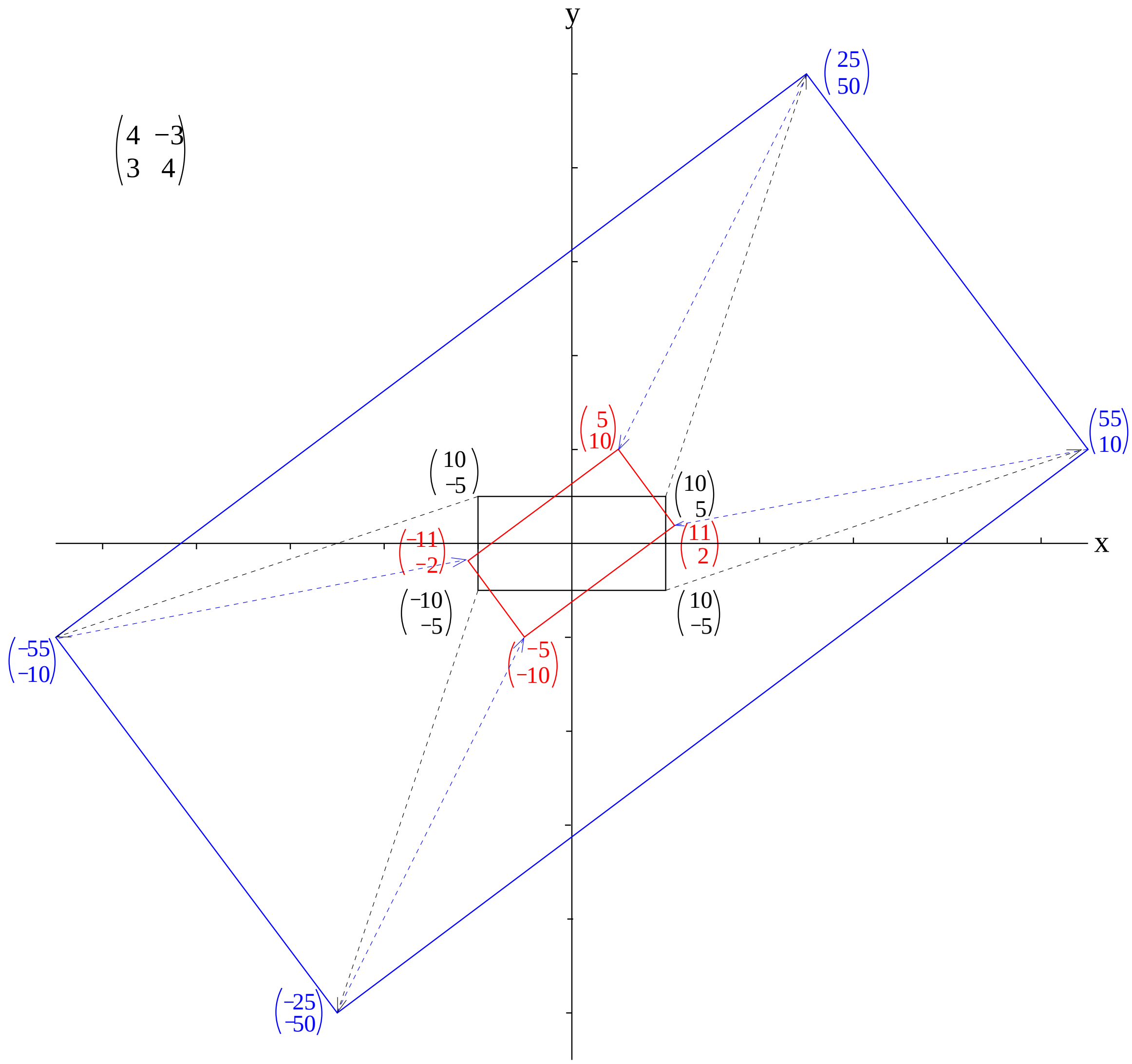
7. Suppose the class in Note 5 is all sitting around a table with corners

$$\begin{pmatrix} 10 \\ 5 \end{pmatrix}, \begin{pmatrix} 10 \\ -5 \end{pmatrix}, \begin{pmatrix} -10 \\ -5 \end{pmatrix} \, and \begin{pmatrix} -10 \\ 5 \end{pmatrix}$$

a) Show that the effect of

$$\begin{pmatrix} 4 & -3 \\ 3 & 4 \end{pmatrix}$$

on the (black) table is the new blue table in

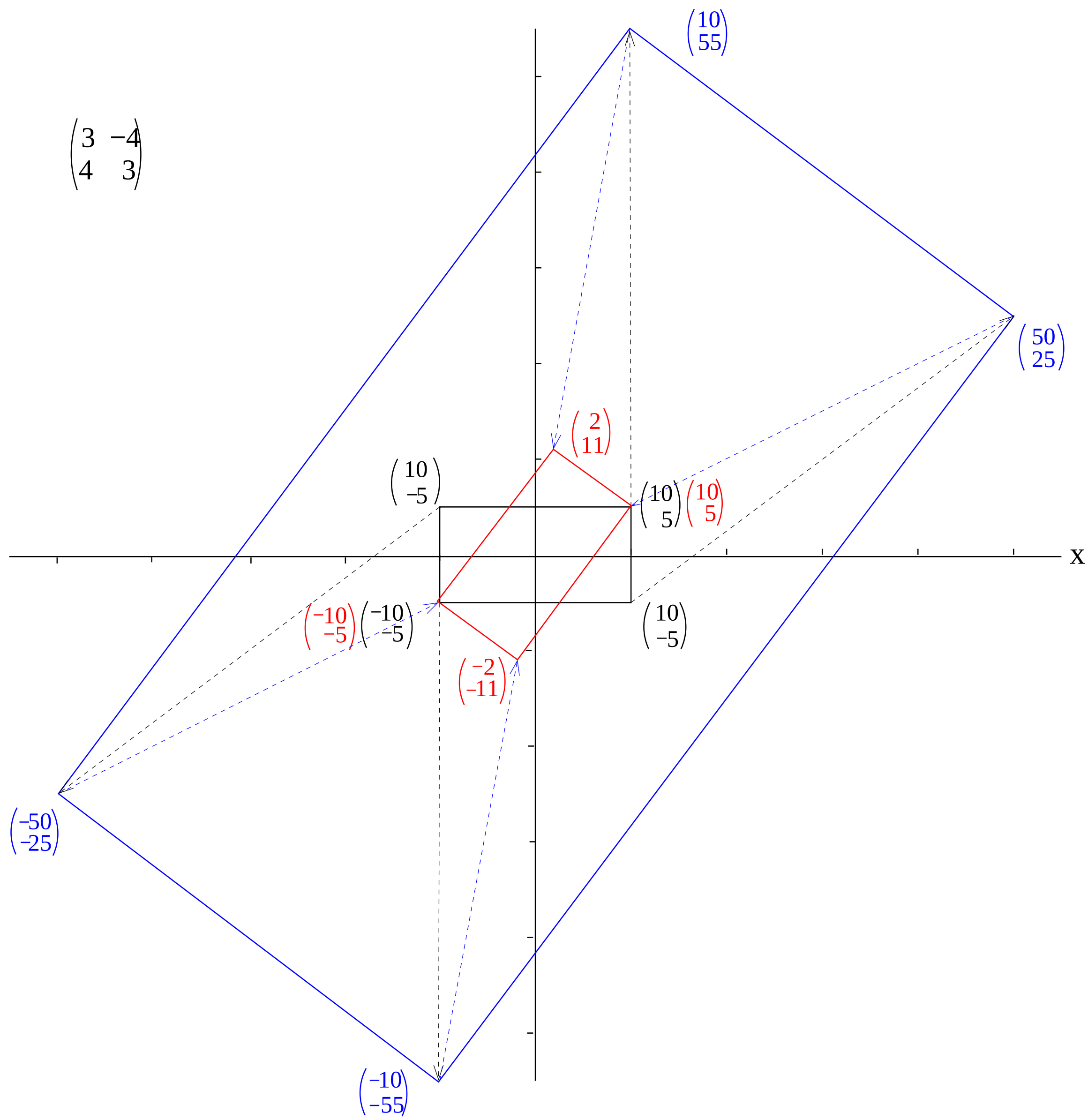


b) Confirm that points *along* the black lines, such as

$$\begin{pmatrix} 10 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -5 \end{pmatrix}, \begin{pmatrix} -10 \\ 4 \end{pmatrix} \text{ or } \begin{pmatrix} 9 \\ 5 \end{pmatrix}$$

and not just the corners, map to corresponding points on the blue lines.
c) What matrix gives the new red table shown above?
d) Do the same for

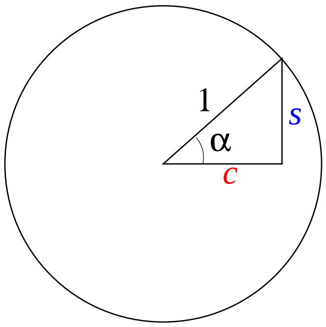
$$\begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}$$



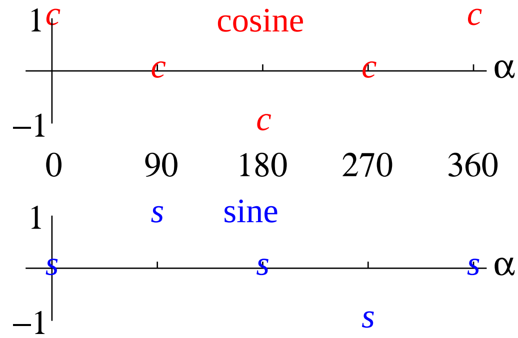
e) This rotates through a larger angle. What matrix would rotate through the still larger angle of 90 degrees?
f) Note that the *rotation* in (d) maps one corner of the table to another corner (and the same for their negatives or "opposites"). Find a rotation matrix which maps the whole table into itself, by mapping each corner into its opposite corner. What matrix just leaves the table alone by mapping every corner back into itself?
g) The two different matrices you have found in (f) are the *symmetry rotations* of the table. If the table were square, it would have two more symmetry rotations: can you find them?
h) Can you find two more matrices which *reflect* (i) the top of the rectangle into the bottom and (ii) the left side of the rectangle into the right side? These are also symmetry operations: they also map the rectangle onto itself.

8. What are the determinants of the two matrices in the previous excursion which map the black tables into the blue tables? By what factor(s) do the *areas* of the tables increase by these mappings? By what factor(s) do the *linear* dimensions of the table increase? Why does dividing the matrix by 5 make the determinant 1? What would happen in three dimensions?

9. **Cosine and sine.** a) For angles α of 0, 90, 180 and 270 degrees calculate the values for c and s in the table.



α	c	s
0	1	0
90	0	1
180	−1	0
270	0	−1
360	1	0



Confirm that c and s plot as shown as a function of the angle α.
b) What we've been calling c and s are short for "cosine" and "sine" respectively. These depend on the angle α and are usually written $\cos(\alpha)$ and $\sin(\alpha)$ respectively. Add to the table and to the plots the values for cos(45) and sin(45): note that they must be equal and that the sum of their squares must be 1.

$$\cos(45) = \sin(45) \qquad\qquad \cos^2(45) + \sin^2(45) = 1$$

(In the table write the exact values. Calculate these to one decimal place for the plots.) What about α = 135, 225 and 315 degrees? (Watch the signs!)
c) Now add the values of c and s for α = 30, 60, 120, 150, 210, 240, 300 and 330 degrees. Note that a 30-60-90-degree triangle is half of an equilateral triangle: show that with hypotenuse 1 the two other sides are $1/2$ and $\sqrt{3}/2$.
d) What are cos(45−30) and sin(45−30)? This will give you the values for c and s for α = 15, 75, 105, 165, 195, 255, 285 and 345 degrees. Hint: if you rotate by 45 degrees then counter-rotate by 30 degrees, you must get a net rotation of 45−30 = 15 degrees.
Show that

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

i.e., following the rotation

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

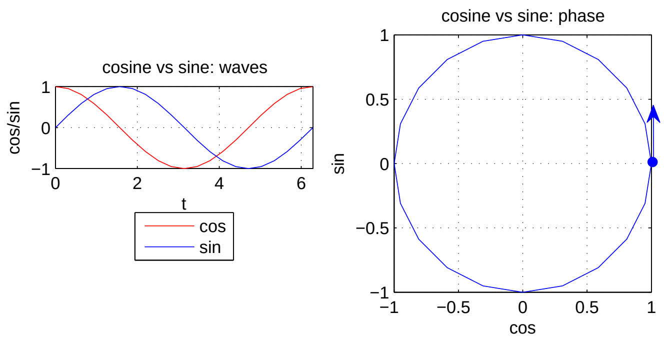
by the rotation

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

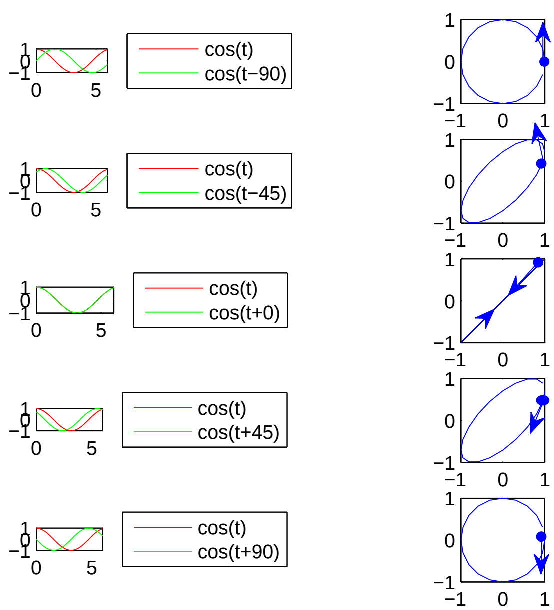
gives the rotation through zero degrees, so that the second is the counter-rotation to the first. How does this relate to the inverse in Note 4?
Then show that the rotation which is the combination of two rotations is

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix} \times \begin{pmatrix} C & -S \\ S & C \end{pmatrix} = \begin{pmatrix} cC - sS & -(sC + cS) \\ sC + cS & cC - sS \end{pmatrix}$$

e) What are $c_2 = \cos(2\alpha)$ and $s_2 = \sin(2\alpha)$ in terms of $c = \cos(\alpha)$ and $s = \sin(\alpha)$?
f) **Comparing waves.** Cos(t) and sin(t) both make waves when plotted against t. You can compare them in two ways. First. plot them both against t. Second, plot them against each other. This makes a circle, generated counterclockwise as t increases.



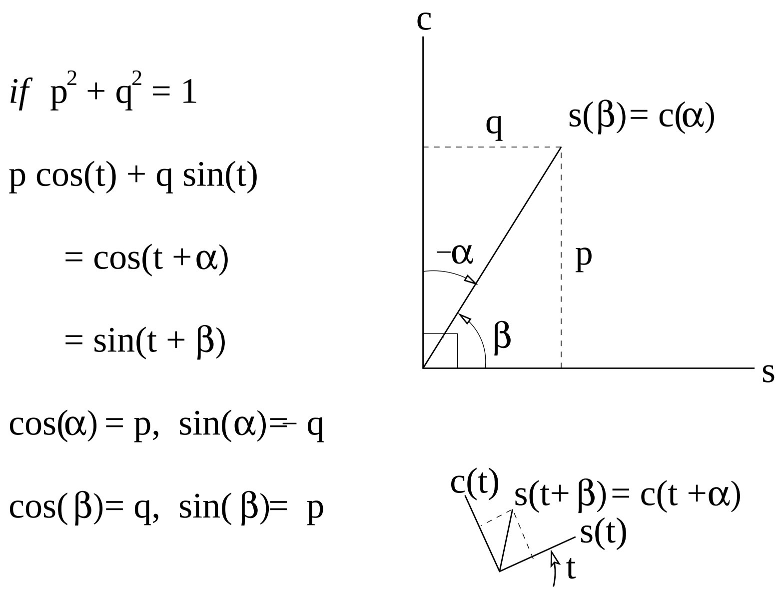
Use the table above and your additional calculations to confirm the circle and its direction.
g) Show that $\sin(t) = \cos(t - 90)$, measuring angles in degrees. Confirm the following wave comparisons. Note the change of direction of rotation as the added constant angle changes sign.



What happens if you continue the added angle from +90 all the way around to −90 degrees?
What happens if you compare $\sin(t)$ with $\sin(t + \alpha)$ for various constant angles α?
The angle α is called the "phase difference" between the two.
h) Use (d) above to show that, if $p^2 + q^2 = 1$,

$$\begin{aligned} p\cos(t) + q\sin(t) &= \cos(t + \alpha) \\ &= \sin(t + \beta) \end{aligned}$$

with $\cos(\alpha) = p$, $\sin(\alpha) = -q$, $\cos(\beta) = q$ and $\sin(\beta) = p$. Hence verify the following "vector" representation of a combination of sin() and cos().

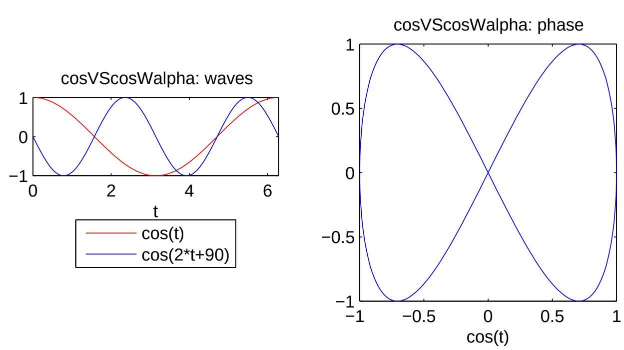


The small diagram underneath says that the whole system of $\sin(t)$, $\cos(t)$ and $\cos(t + \alpha)$ $= \sin(t + \beta)$ rotates as t increases while preserving the relationship among the three lines. Note that sin() is the horizontal axis and cos() the vertical: $\cos(t)$ "leads" $\sin(t)$ during this rotation, or $\sin(t)$ "lags" $\cos(t)$. Explain how $\cos(t+\alpha) = \sin(t+\beta)$ leads and lags $\sin(t)$ and $\cos(t)$, respectively.

g) Express $a\cos(t) + b\sin(t)$ in terms of phase-changed cos() or sin() when $r = \sqrt{a^2 + b^2} \neq 1$.

h) Write a MATLAB program to compare $\cos(t)$ and $\cos(w \times t + \alpha)$ for multiplier w and phase difference α. Try it with w simple integers such as 2 or 3 and then simple fractions such as 1/2 or 2/3. Start with $\alpha = 90$ degrees, then get more adventurous. (Be careful to run it far enough to show the whole picture: `rat(w,0.1)` is a useful MATLAB function.

Here is a simple call, `cosVScosWalpha(2,90)`



10. **2-dimensional numbers: a digression from matrices.** a) The rotation matrices in Note 8 and in the previous excursion have the form

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

where c and s are numbers such that $c^2 + s^2 = 1$. Check that this true for the two matrices that map black rectangles into red rectangles in the previous excursion.
b) Using matrix addition (Note 2), show that the above can be written

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix} = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} + \begin{pmatrix} 0 & -s \\ s & 0 \end{pmatrix}$$

Then, using scalar multiplication (Note 2), show that it can further be written

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

c) These are special matrices. What happens when

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is multiplied by *any* 2×2 matrix (or *any* 2×1 vector)? What happens when

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

is multiplied by *any* 2×1 vector? (Try it first on the two special "unit" vectors of Notes 2 and 7.)
Why is the first matrix usualy called the *identity* matrix (Note 3.)? It is usually given the special name I. And let's give the second matrix the special name i.
So we can write the rotation as

$$cI + si$$

d) Rotation matrices behave just like ordinary numbers. They commute when multiplied by each other, which matrices in general do not do. So in principle we could treat I and i as numbers (which c and s already are). I, as a number, acts just like 1, so we can now write the rotation

$$c + si$$

or, more conventionally, as

$$c + is$$

e) But the i thing is very weird. As we saw in (c) i rotates through 90 degrees, a right angle. But numbers are 1-dimensional, along a line from $-\infty$ to ∞. How can they be rotated through a right angle?
Well, they can if we think that i introduces a *second dimension.*
We can say i stands for "imagine that!".
And so we can imagine *2-dimensional* numbers that obey all the arithmetic laws that regular "1-dimensional" numbers do.
f) Practice this by multiplying the matrices

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix} \times \begin{pmatrix} C & -S \\ S & C \end{pmatrix} = \begin{pmatrix} cC - sS & -(sC + cS) \\ sC + cS & cC - sS \end{pmatrix}$$

and comparing it with the corresponding multiplication of 2-dimensional numbers

$$\begin{aligned}(c+is)\times(C+iS) &= cC+isiS+i(sC+cS)\\ &= cC-sS+i(sC+cS)\end{aligned}$$

Does the result say the same thing in both cases?
g) "But, wait!" you say. "How did $isiS$ become $-sS$?"
Well, $isiS = iisS = i^2sS$ by commutativity of number multiplication. So we just need to know what i^2 means.
Well, multiplying by i means rotating through a right angle. So multiplying by i again means rotating through a *second* right angle. Two right angles make 180 degrees.
Now think of the 1-dimensional number line. If I rotated it by 180 degrees, I would map 1 into -1, -1 into 1, and every other number into its negative. That is, rotating through two right angles just changes the sign of any 1-dimensional number. So $i^2 = -1$.
h) Practice comparing matrix rotations and 2-dimensional number rotations for specific Pythagorean triples, e.g., $c = 4/5$ and $s = 3/5$ as in Note 7. Every time you come up with an i^2 just replace it by -1.

11. **Three-dimensional rotations.** Since a rotation is always in a plane (2 dimensions), writing rotations in three dimensions is not in principle harder than in two. Only there will be different rotations in different planes. We can look at basic rotations in the xy, yz and zx planes. Each will leave one direction invariant, so corresponding to that direction there will be a diagonal 1 in the now 3-by-3 matrix.

$$R_{xy} = \begin{pmatrix} c & -s & \\ s & c & \\ & & 1 \end{pmatrix} \qquad R_{yz} = \begin{pmatrix} 1 & & \\ & c & -s \\ & s & c \end{pmatrix} \qquad R_{zx} = \begin{pmatrix} c & & s \\ & 1 & \\ -s & & c \end{pmatrix}$$

Show that $R_{yz}R_{xy} \neq R_{xy}R_{yz}$ for a 90-degree angle a) by matrix multiplication and b) by performing these two rotations in the different orders on some unsymmetrical 3-D object such as a chair.
c) What would R_{xy} look like in four dimensions? How many different basic 4-dimensional rotations can there be? (How many ways can two dimensions be chosen from among the four?)

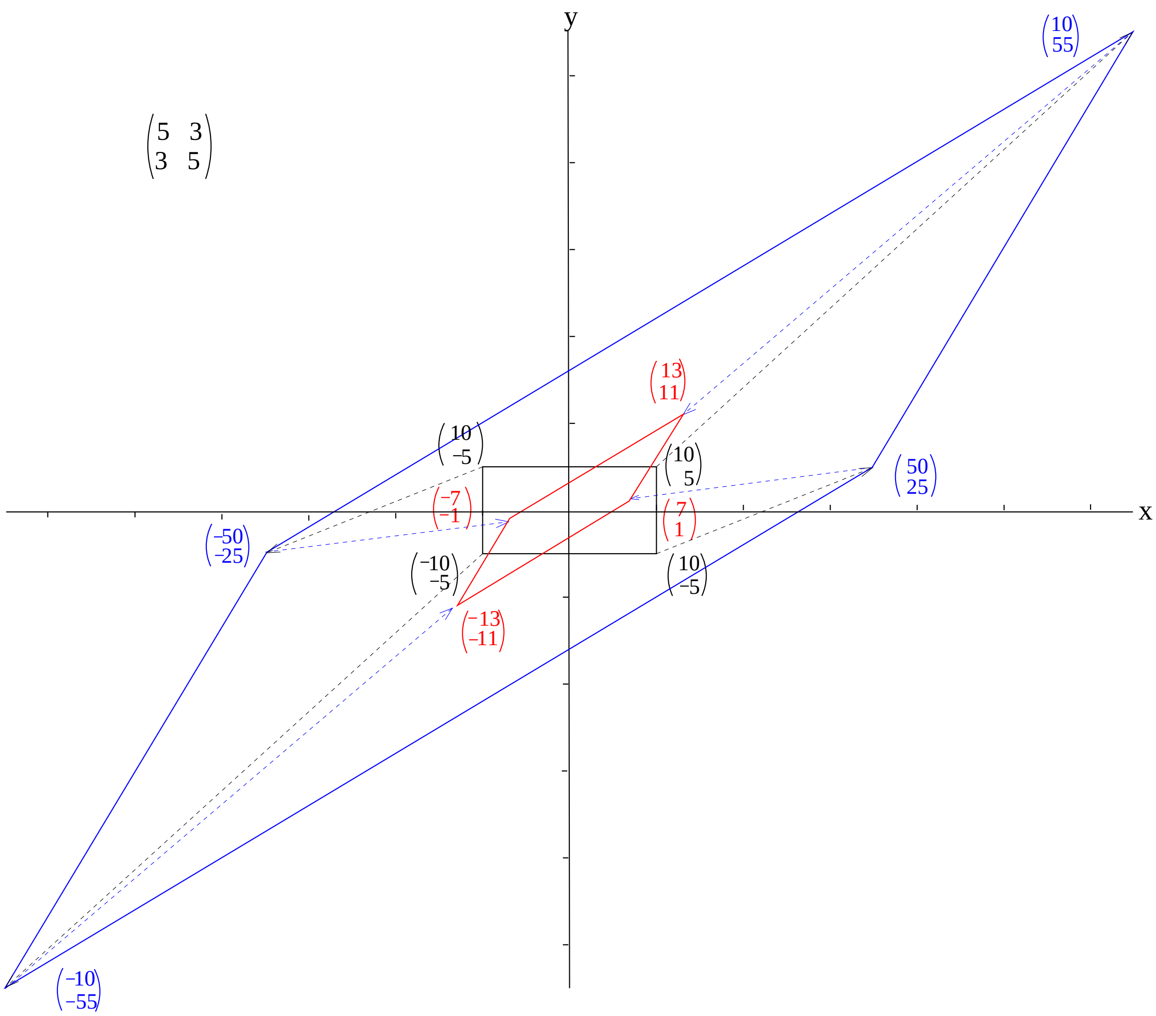
12. Calculate the effect of the first symmetric matrix of Note 9 on the seven vectors of the classroom in Note 8 and compare this with the shear matrix by drawing the transformed space.

13. a) Write down a shear matrix based on the Pythagorean triple 5, 12, 13.
b) Multiply this both ways with the shear matrix from Note 9: does matrix multiplication commute for shear matrices?
c) What are the inverses of these shear matrices?
d) What are the invariant vectors of your new shear matrix?

14. a) Show that the matrix

$$\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

applied to the table the class is all sitting around gives the distorted blue table in



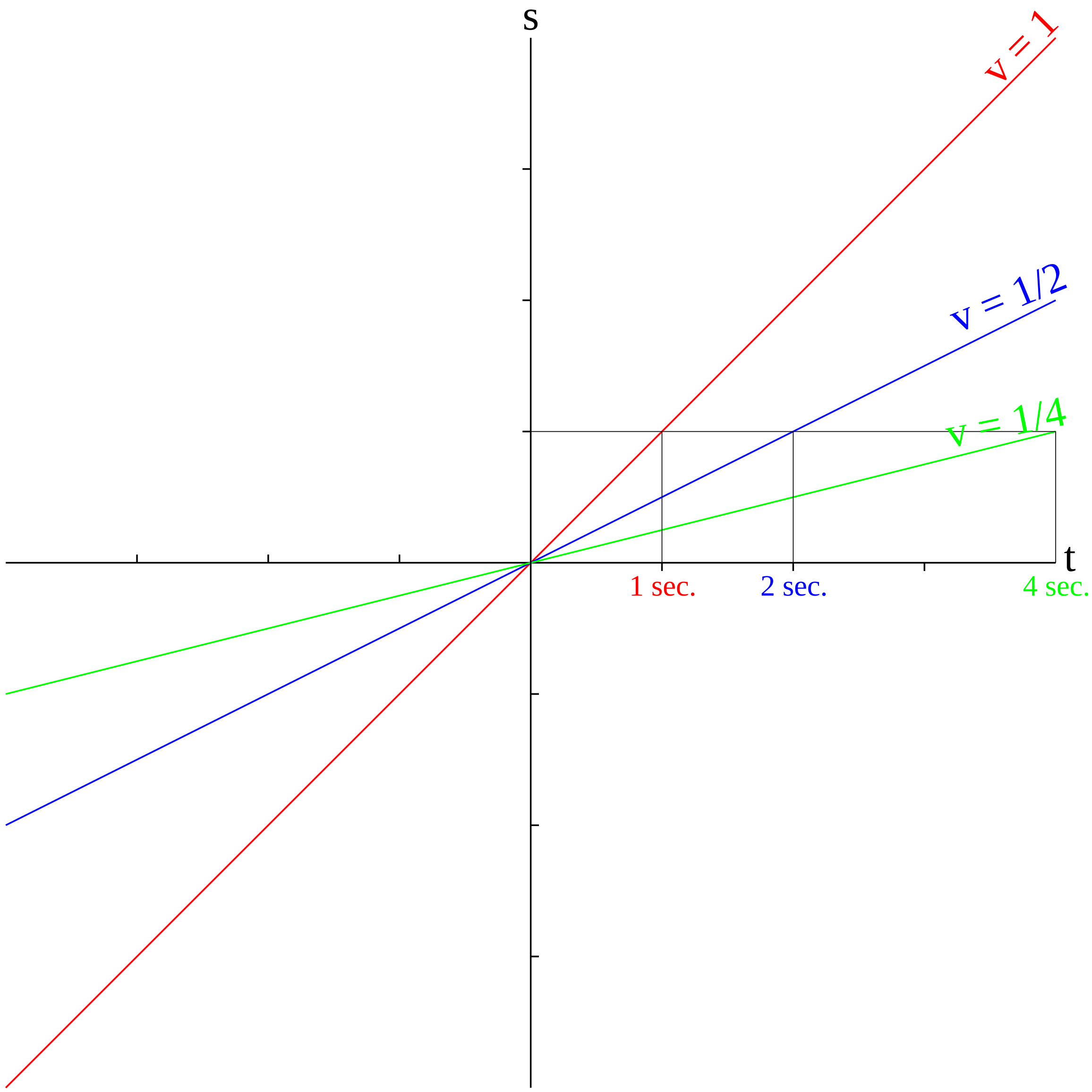
b) Find the shear matrix that gives the distorted red table.

15. a) How do

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

transform under the shear matrix of Note 9? Draw the new space.
b) Which of these vectors are appropriately called "invariant" vectors of the matrix?

16. Rotations and shears have so far been discussed for two-dimensional *space* and I've labelled the axes x and y.
We can also consider shear in *timespace* with the axes labelled t for time and s for space.
What connects time and space is *velocity*. If we travel, time becomes space: travel twice as fast and the same time allows us to cover twice the distance.
Here are three different velocities in timespace.



If we travel at 1 m/sec it takes us 1 sec to go 1 m; if we travel at 1/2 m/sec it takes us 2 sec to go 1 m; if we travel at 1/4 m/sec then we need 4 sec to go that 1 m.
A shear transformation is called a *boost* in timespace. Let's look at

$$\begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix}$$

Show that this moves

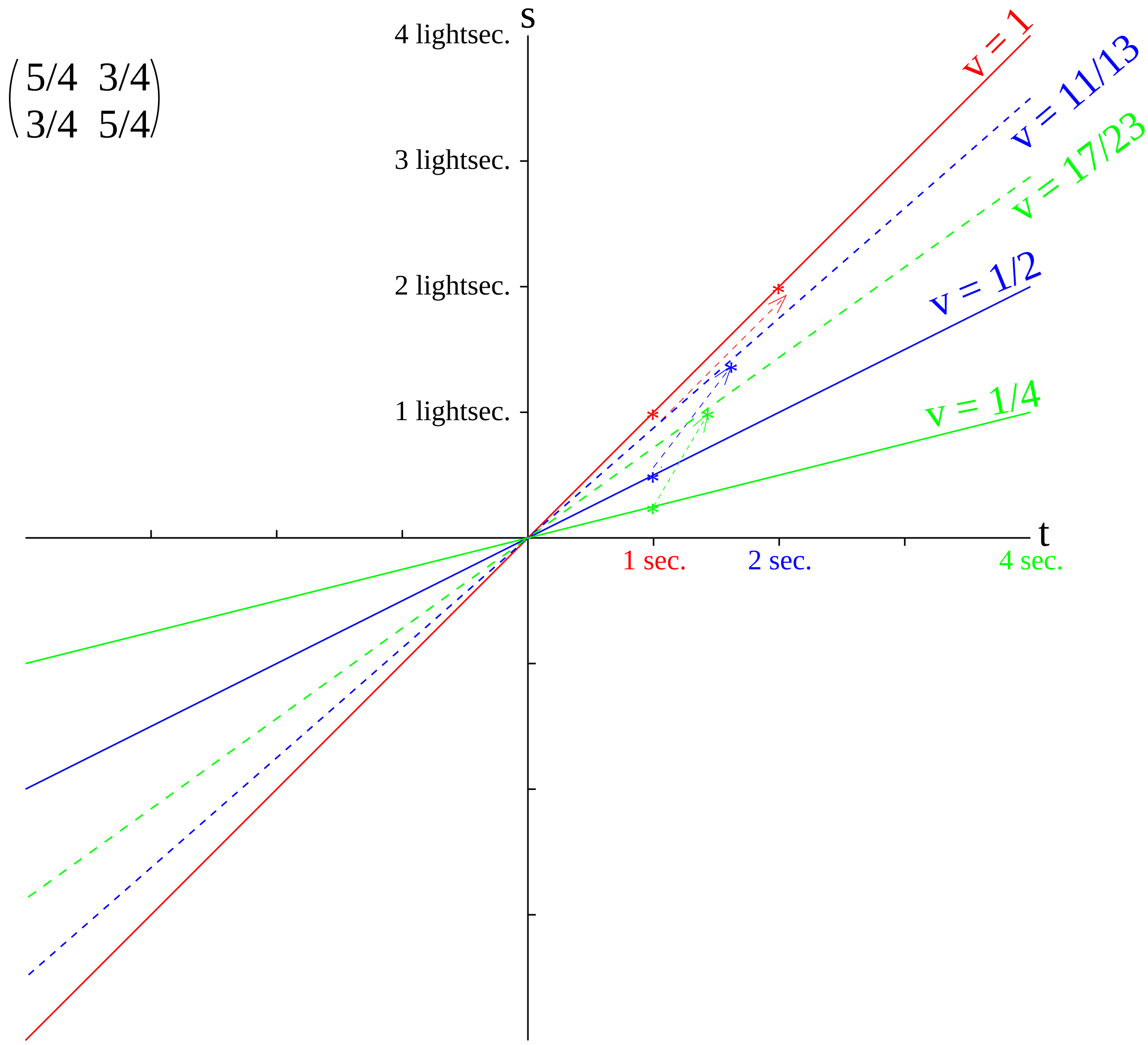
$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} to \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 1/2 \end{pmatrix} to \begin{pmatrix} 13/8 \\ 11/8 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 1/4 \end{pmatrix} to \begin{pmatrix} 23/16 \\ 17/16 \end{pmatrix}$$

and so moves the velocity lines passing through the first of each of these pairs of points to the dashed velocity lines passing through the second of each pair.

(When we are interested only in velocities, why does it not matter where on the lines the points lie?)



Interpret the changes in the lines as "boosts" in the velocities. The lowest velocity, 1/4, gets the biggest boost, to 17/23. The middle velocity, 1/2, gets not quite so big a boost, to 11/13. The highest velocity, 1, stays exactly the same.
If we measure the space units not in meters but in light-seconds, $v = 1$ corresponds to lightspeed. Einstein said nothing can go faster than light. We have just seen the mathematics behind that. Boosts correspond to accelerations and nothing can accelerate lightspeed because the $v = 1$ line is an invariant line of the boost (shear) transformation: see previous Excursion.

17. **Reflections, etc.** a) What do the matrices

$$F_y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
$$F_{xy} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

do to an x-y-space such as the classroom?
b) Show that any 2-by-2 shear matrix of the sort discussed in Note 9 is a combination of the

identity matrix I and F_{xy}

$$S = aI + bF_{xy}$$

What restriction must we assume on the relationship between a and b to get the kinds of shear in Note 9?

c) Show that *any* 2-by-2 matrix is a combination of I, the 90-degree rotation matrix R_{90}, F_y and F_{xy}

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \alpha I + \beta R_{90} + \gamma F_{xy} + \delta F_y$$

What are α, β, γ and δ in terms of a, b, c and d?

18. Draw the classroom of Note 8, or the class table of previous excursions, as transformed by any of the 2×2 matrices discussed in these Notes or that you have invented yourself.

19. The "MAT" in the MATLAB programming language stands for matrices. The TI81 calculator and its successors can also do matrix operations. Learn how to use these or equivalent software to check the calculations in these Notes and your own exercises.

20. **Symmetry.** a) Draw a rectangle with corners labelled `a, b, c` and `d` at points

$$\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} -2 \\ -1 \end{pmatrix}$$

respectively. What effect do the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

have on the rectangle? On `a, b, c` and `d`? Is it appropriate to call these matrices, respectively, R_0 and R_{180}?

What effect do the matrices

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

have on the rectangle? On `a, b, c` and `d`? Is it appropriate to call these matrices, respectively, F_v and F_h?

b) Multiply each of these four matrices by every other one, and confirm the matrix "multiplication table"

$\times$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

c) To see the pattern in this, rewrite it using the names I gave above instead of the matrices themselves. Here is how it should come out: I've replaced R_{180} by just R and R_0 by I for the "identity" matrix.

×	I	R	F_v	F_h
I	I	R	F_v	F_h
R	R	I	F_h	F_v
F_v	F_v	F_h	I	R
F_h	F_h	F_v	R	I

This set of four operators (which we initially wrote as matrices) forms a *group*, a mathematical construct which has a multiplication table such as above, with a) only the original operators from the top row and first column appear in the body of the table—i.e., multiplying group elements keeps within the group, and b) each column and each row has no repeating operators. (A precise characterization of groups is given in Book 8c Note 4.) The *group of the rectangle* is the set of four operators that transform any rectangle into itself. These are the *symmetry* operators of the rectangle.

d) Find the eight matrices that make up the group of the *square*, give them appropriate symbols or short names, and write down their multiplication table. Notice that the group of the rectangle is a *subgroup* of the group of the square. What does this mean? Does a square have more symmetries than a rectangle?

Is there a subgroup within the group of the rectangle?

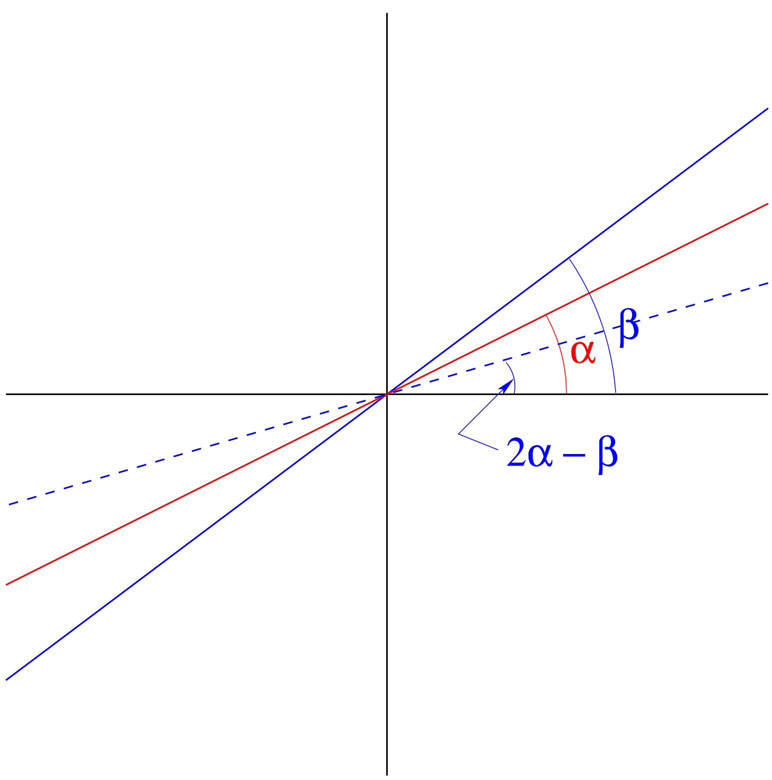
21. **Projection and reflection.** a) Why is the matrix

$$P = (c, s)P \begin{pmatrix} c \\ s \end{pmatrix} = \begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix}$$

called a *projection* matrix? Try transforming any vector with it

$$P \begin{pmatrix} x \\ y \end{pmatrix} = ?$$

b) Suppose the solid blue line in the figure (at angle β) is reflected in the red line (at angle α) to produce the dashed blue line. Show that the angle of this reflected line is $2\alpha - \beta$.



c) Call the *reflection* matrix that transforms the blue line into the dashed blue line F. Why does $P = (I + F)/2$?
d) From this, show that the matrix that reflects the 2-D space in a mirror with orientation

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

must be

$$F = \begin{pmatrix} c^2 - s^2 & 2cs \\ 2cs & s^2 - c^2 \end{pmatrix}$$

e) Show that therefore

$$F = \begin{pmatrix} c_2 & s_2 \\ s_2 & c_2 \end{pmatrix}$$

where the double rotation

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

(We say that c_2 is the cosine of twice the angle that c is cosine of and s is sine of, and s_2 is sine of twice this same angle. See part (e) of the excursion above on "Cosine and sine" for combined rotations.)
f) Incidentally, if the red line has orientation

$$\frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

and the blue line has orientation

$$\frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

show that the reflected (dashed blue) line has orientation

$$\frac{1}{25} \begin{pmatrix} 24 \\ 7 \end{pmatrix}$$

22. **Diagonalizing matrices.** Here is a formal way to find the directions which diagonalize any matrix, and also to find the values on the final diagonal. The directions are called "eigenvectors" and the diagonal values "eigenvalues", from the German word "eigen" meaning "own" or "proper".
For a matrix A suppose v is an eigenvector:

$$Av = \lambda v$$

That is, A operating on v preserves the direction of v, which is to say, gives a multiple λ of v: λ is the corresponding eigenvalue.
We can rewrite this

$$(A - \lambda I)v = 0$$

where I is the identity and "0" in this context means the "zero" vector of all zeros.
Unlike ordinary arithmetic, this equation does not mean that v (or $A - \lambda I$) must be zero. But it does mean that the *determinant* (Note 4) is:

$$\det(A - \lambda I) = 0$$

For 2×2 matrices A and I this gives the quadratic equation

$$\lambda^2 - \text{trace}A + \det A = 0$$

We know what determinants are. The *trace* of a matrix is just the sum of its diagonal elements.
a) Show that this is true for an arbitrary 2×2 matrix

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

b) Work out and solve the quadratic equation for the shear matrix of Notes 9 and 10

$$\begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix}$$

(If you don't yet know how to solve a quadratic, try substituting first one then the other of the two diagonal values we got in Note 10: $\lambda = 2$ and $\lambda = 1/2$.)
c) What happens if we start to diagonalize a rotation matrix in this way?
After we find the eigenvalues, the next step is to find the eigenvectors. Let's do this by example for the shear matrix. Call either of the eigenvalues we just found λ and try to find the components of v which make

$$(A - \lambda I)v = 0$$

For instance

$$\begin{aligned}
\begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \left(\begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix} - 2\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}\right)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= \left(\begin{pmatrix} 5/4-2 & 3/4 \\ 3/4 & 5/4-2 \end{pmatrix}\right)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= \left(\begin{pmatrix} -3/4 & 3/4 \\ 3/4 & -3/4 \end{pmatrix}\right)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
&= \frac{3}{4}\begin{pmatrix} -v_1 + v_2 \\ v1 - v_2 \end{pmatrix}
\end{aligned}$$

So $v_2 = v_1$ in both terms. We can set both = 1, or, to "normalize" the eigenvector so it has unit length, we can set both to $1/\sqrt{2}$.
d) Repeat this for the second eigenvalue $\lambda = 1/2$ for the shear matrix example, and confirm that the resulting eigenvector is what we had in Note 10.
Not every matrix can be diagonalized. Here is a family of 2×2 matrices, for any a, b or c, which cannot.

$$A = \begin{pmatrix} c-a & b \\ -a^2/b & c+a \end{pmatrix}$$

e) Using the eigenvalue procedure (leading up to (a) and (b) above) show that the two eigenvalues are the same, each c.
Although these matrices cannot be exactly diagonalized, they can be almost diagonalized, to

$$J = \begin{pmatrix} c & 0 \\ 1 & c \end{pmatrix}$$

which is called "Jordan canonical form".
Let's see this, by finding the "eigenvectors", the columns P_1 and P_2 of the matrix P. This

time we will write $P^{-1}AP$ with the inverse first.

$$\begin{aligned} P^{-1}AP &= J \\ AP &= PJ \\ \begin{pmatrix} c-a & b \\ -a^2/b & c+a \end{pmatrix}[P_1, P_2] &= [P_1, P_2]\begin{pmatrix} c & 0 \\ 1 & c \end{pmatrix} \\ &= [cP_1 + P_2, cP_2] \end{aligned}$$

So P_2 behaves normally and we can solve for it.

$$\begin{aligned} AP_2 &= cP_2 \\ (A - cI)P_2 &= 0 \\ \left(\begin{pmatrix} c-a & b \\ -a^2/b & c+a \end{pmatrix} - c\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}\right)\begin{pmatrix} p \\ q' \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} -a & b \\ -a^2/b & a \end{pmatrix}\begin{pmatrix} p \\ q' \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

and you can show $q' = pa/b$.
But P_1 is different and is called a "generalized eigenvector".

$$(A - cI)P_1 = P_2$$

and so we must take an extra step.

$$\begin{aligned} (A - cI)^2 P_1 &= 0 \\ \begin{pmatrix} -a & b \\ -a^2/b & a \end{pmatrix}^2 \begin{pmatrix} q \\ r \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} q \\ r \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

and you can show that there is no restriction on q or r.
So

$$P = \begin{pmatrix} q & p \\ r & pa/b \end{pmatrix}$$

f) Now for this example, find the product $P^{-1}AP$ and by comparing it to J find the final relation on the p, q, r making up P, namely $p = br - aq$. How did we happen to miss this equation in finding the generalized eigenvectors?
g) Show that the equality of the eigenvalues and the presence of the 0 in J are essential by tweaking J in two different ways. First try

$$\begin{pmatrix} c+\epsilon & 0 \\ 1 & c \end{pmatrix}$$

and show that this diagonalizes exactly, with eigenvalues $c + \epsilon$ and c. Then try

$$\begin{pmatrix} c & \epsilon \\ 1 & c \end{pmatrix}$$

and show that this diagonalizes exactly, with eigenvalues $c \pm \sqrt{\epsilon}$. Since ϵ can be arbitrarily small in each case, what does this say about calculating with J on computer?
h) J in the slightly different guise

$$G = \begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix}$$

(dividing by c and supposing $v = 1/c$) is called the Galilean transformation and connects time and space via velocity v. Figure out how this works. Hint: if you start at s and travel at velocity v for time t you arrive at $s + vt$. Einstein disagreed with Galileo and replaced G by a shear transformation

$$L = \frac{1}{\sqrt{1-v^2}} \begin{pmatrix} 1 & v \\ v & 1 \end{pmatrix}$$

Show that this also has determinant 1. Thus "space math" becomes "timespace math".

23. **Array addressing.** When a matrix is stored on a computer or calculator its elements must be stored in a memory which is organized with only one index giving each address, not the two indices that the matrix uses.

Thus the matrix

$$\begin{pmatrix} 3 & 0 \\ -4 & 1 \end{pmatrix}$$

might be stored as

3	−4	0	1

Here, the mapping from the indices

$$\begin{pmatrix} 1\,1 & 1\,2 \\ 2\,1 & 2\,2 \end{pmatrix}$$

is to the addresses

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

or, since you'll see that it is much handier to start everything at zero, it is the mapping from the indices

$$\begin{pmatrix} 0\,0 & 0\,1 \\ 1\,0 & 1\,1 \end{pmatrix}$$

to the addresses

$$\begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix}$$

If the row index is j and the column index is k, show that the address is

$$a = j + 2k$$

and, for an m-by-n matrix

$$a = j + m * k$$

Here are the addresses for a 3-by-4

j					
0	0	3	6	9	
1	1	4	7	10	
2	2	5	8	11	
	0	1	2	3	k

a) Use the address formula to check that it generates the 3-by-4 addresses shown above, and convince yourself that it will work for any m-by-n matrix.
b) Matrices are a special case of *arrays* in a programming language: an m-by-n matrix is called an m-by-n array. But we could also have 3-dimensional arrays, say, m-by-n-by-p, or 4-dimensional arrays, say m-by-n-by-p-by-q, and so on.
For $m = 4, n = 3$ and $p = 1$, calculate all the 3-D addresses given by the formula

$$a = j + m * k + m * n * \ell$$

where ℓ is the third index, running from 0 to $p - 1$.
c) The formula $a = j + m * k$ is one of two possibilities for 2-D arrays, both working in the same general way. This one is said to store the array in "row-major" order. What is the formula that will store a 2-D array in "column-major" order?
How many different addressing possibilities are there in 3-D? 4-D?

24. **Simplex arrays.** Extending the previous Excursion, what if we wanted to store a *triangular* matrix? For example, a *symmetric* matrix could be stored in half the space if we did not store the (j, k) element as well as the (k, j) element.

$$\begin{pmatrix} 5 & 4 & 3 & 9 \\ 4 & 5 & 6 & 3 \\ 3 & 6 & 4 & 5 \\ 9 & 3 & 5 & 4 \end{pmatrix}$$

could just as well be stored as

$$\begin{pmatrix} 5 & 4 & 3 & 9 \\ & 5 & 6 & 3 \\ & & 4 & 5 \\ & & & 4 \end{pmatrix}$$

Here is a way of addressing it

$$\begin{array}{c|ccccc} k & & & & & \\ 0 & 0 & 1 & 3 & 6 & \\ 1 & & 2 & 4 & 7 & \\ 2 & & & 5 & 8 & \\ 3 & & & & 9 & \\ \hline & 0 & 1 & 2 & 3 & j \end{array}$$

a) Show that

$$a = k + \Delta_j$$

generates all these addresses, still starting all our counts from 0. (Δ_k is the kth triangular number: see Notes 1 and 2 of Week i.)
Note that this formula does not need to know the overall size of the matrix.
What can you say about the addresses in the top row, above?
b) For a 4-by-4-by-4 array, calculate all the addresses generated by

$$a = \ell + \Delta_k + \diamondplus_j l$$

being careful in what order you write down the result: try j horizontal and k vertical for $\ell = 0$, then ditto for $\ell = 1$ and so on.
What can you say now about the addresses in the top row when $k = 0 = \ell$?
c) In the 4-by-4 example, check that $k \leq j$ always. What is the corresponding relationship among the indices j, k and ℓ in the 4-by-4-by-4 example? What must we do to look up an element (of the symmetric matrix so stored) whose indices do not satisfy these constraints, e.g., element $j, k = 2, 1$?
d) How does all this discussion generalize to a "simplex array" of any number of dimensions?

25. Any part of the Preliminary Notes that needs working through.