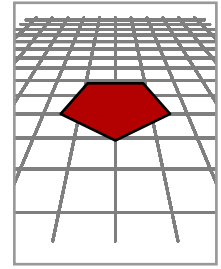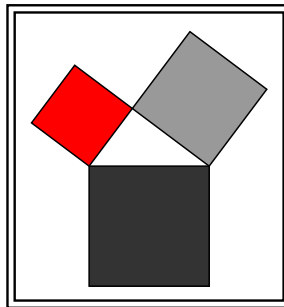# Perspective and homogeneous coordinates

Many image manipulation programs have a tool for modifying the perspective of images. Thus, I start out with this familiar picture

GIMP:1 and then I import it into my image manipulation program (which happens to be the **G**nu **I**mage **M**anipulation **P**rogram or **GIMP**). I next open the transform tool and choose the perspective option. When I click now in the image window, what I see is this:
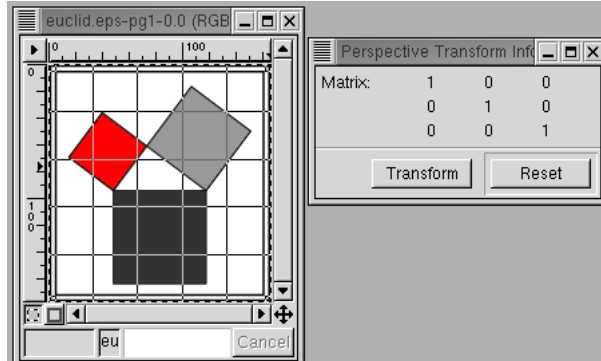
pective transform:1 The tool lays down a grid on top of my image, and simultaneously brings up what it calls a **Perspective Transform Info** window, in which is displayed a 3×3 matrix. It also shows some boxes at the corners of the grid that I can grab and move around. When I do so, the original image remains the same, and the grid adjusts itself to my choices. In effect, a geometrical transform of some kind is applied to the grid. The matrix changes, alas, in some apparently incomprehensible manner.
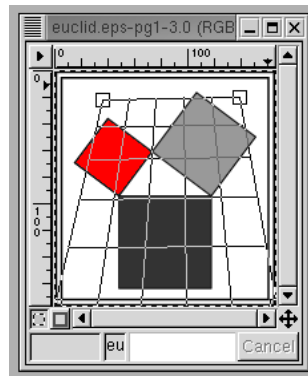
Finally, if I click on the **transform** button in the matrix window, the image itself is transformed:
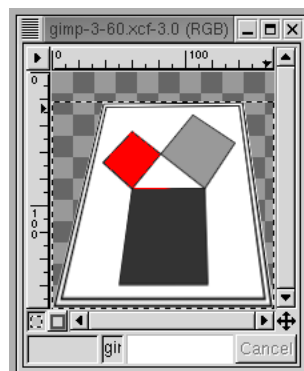
The original image has been distorted in such a way as to make it look as though it were being viewed from below—as though it were being viewed in a different **perspective**.

What kind of a transform are we seeing here? How is it different from others we have already seen, such as translation, rotation, shearing, and scale change? The image manipulation program has in fact other transform tools corresponding to those, such as the **rotate** and **scale** tools. All those other transforms have one property that the perspective transform doesn't—they all transform a pair of parallel lines into some other pair of parallel lines. They are all affine transforms, and all affine transforms have this property. So the perspective transform is really something new. As we'll see, in gaining flexibility we must sacrifice something else.

This chapter will analyze perspective transforms mathematically. The matrix window in the figures above turns out to be a major clue. The analysis will involve at first what seems like an enormous detour. Before we begin on that journey, I should motivate it by mentioning that perspective in 3D is a crucial ingredient in any graphics toolkit, and that the techniques to be explained here are, for that reason, ubiquitous. The advantage of working with this 2D image manipulation tool is that things are simpler to visualize. It should be thought of as a model for the more interesting 3D version.

1. The projective plane

In order to understand what the perspective tool does we must enlarge the usual plane by adding points 'at infinity'. Familiarity with the use of perspective in art, where lines meet at infinity and the horizon becomes a line on a canvas, should make this seem like a natural step. We have seen already a hint that this is necessary—two parallel lines meet only at infinity, but after a perspective transform is applied to them they may meet in a definite point of the image plane.
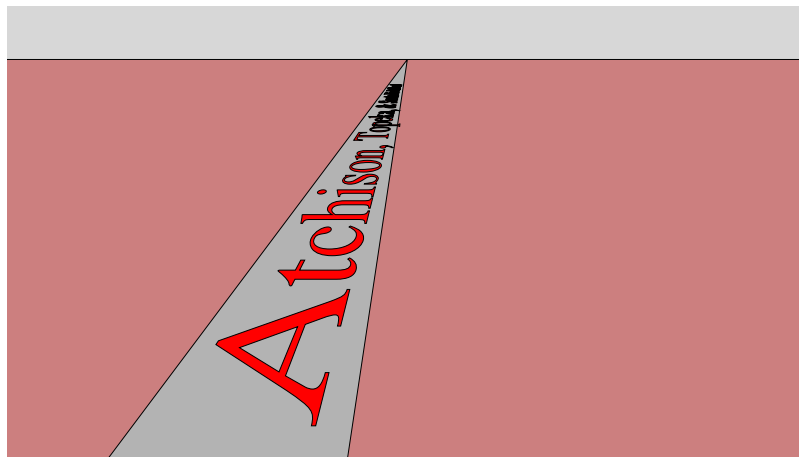
This means that the transform has converted some points at infinity into ordinary ones. We shall see also that it may move some ordinary points off to infinity as well.

points at infinity:3 How to enlarge the plane by adding points at infinity? There are several ways to do this. The one I'll follow is not quite the obvious one—I'll add one point at infinity for every **non-oriented direction**. In other words, on every line there exists one of these new infinite points, and we can reach it, so to speak, going infinitely far along the line—*in either direction*. The more obvious thing to do would be to add a point at infinity for every direction, so that every line contains two points at infinity, but this turns out not to be so useful.

projective plane:3 The extended plane we get by adding points at infinity to the usual plane in this way is called the **projective plane**. We can put coordinates on its points—there are two cases to be dealt with. (1) The ordinary finite points can be specified by the usual coordinates (x, y). (2) A direction can be specified by a vector $[x, y]$, but x and y are only determined up to a non-zero scalar multiple—$[cx, cy]$ has the same direction as $[x, y]$. Because of this **eous coordinates:3** ambiguity, the second are called **homogeneous coordinates**.

Any two parallel lines have the same direction, so they will contain the same unique point at infinity. In other words, with these new points added every pair of parallel lines will intersect at a single point. In fact, every two distinct lines, whether parallel or not, intersect in exactly one point. This is a very simple property of the new plane.

Adding the points at infinity in the way I have done it is intuitive if somewhat awkward. In particular it doesn't seem to give any clues about what GIMP does. And it is bothersome that the points at infinity are treated differently from the rest. There is a surprisingly elegant but equivalent way to carry out the construction.

We have already seen how to embed 2D points (x, y) into 3D by adding a third coordinate $z = 1$ to get $(x, y, 1)$. The ordinary 2D plane $z = 0$ becomes the plane $z = 1$. But it turns out that we can also embed the points at infinity into 3D? A direction $[x, y]$ in 2D is to be identified with the same direction $[x, y, 0]$ in the plane $z = 0$.

Thus far, the finite points and the infinite ones are still treated quite differently. But in fact *we can identify even the points in the plane $z = 1$ with directions in 3D*. A direction in 3D is essentially a vector $[x, y, z]$ (with one coordinate not 0), and this determines a unique line through the origin with that direction, that of all points (tx, ty, tz) as t ranges from $-\infty$ to ∞. That line will intersect the plane $z = 1$ in a unique point $(x/z, y/z, 1)$ (with $t = 1/z$) as long as $z \neq 0$. If z vanishes, then the line with direction $[x, y, 0]$ amounts to a point at infinity on the projective plane. In summary:

- *The points of the extended plane we get by adding unoriented directions to the usual plane are in exact correspondence with lines in 3D through the origin.*

To repeat: the 3D line with direction $[x, y, 0]$ corresponds to the direction $[x, y]$ in 2D, a point at infinity; and the line with direction $[x, y, z]$ where $z \neq 0$ corresponds to the 'finite' point $(x/z, y/z)$.

A line through the origin in 3D may be assigned the vector $[x, y, z]$ indicating its direction. These are **homogeneous coordinates** in the sense that multiplying them all by a non-zero constant doesn't change the direction. If $c \neq 0$ then the coordinates $[x, y, z]$ and $[cx, cy, cz]$ are associated to the same point of the projective plane. We have seen homogeneous coordinates before, when we assigned the coordinates $[A, B, C]$ to the line $Ax + By + C = 0$.

If (x, y, z) is any point in 3D with $z \neq 0$ then the point $(x/z, y/z, 1)$ on our copy of the usual 2D plane that is equivalent to it will be called its **projective normalization**. This corresponds in turn to the 2D point $(x/z, y/z)$, which I call its **2D projection**.

There is one important relationship between the ordinary points and points at infinity. If we start at a point $P = (a, b)$ and head off to infinity in the direction $[x, y]$ then we pass through points $P + t[x, y] = (a + tx, b + ty)$. As t gets larger and larger, we should be approaching some point at infinity, shouldn't we? Sure enough, in terms of homogeneous coordinates we are passing through the points

$$[a + tx, b + ty, 1] \sim [a/t + x, b/t + y, 1/t] \longrightarrow [x, y, 0]$$

as $t \to \infty$.

2. Boy's surface

The projective plane has one very surprising property. If we move along a line to infinity we will pass through infinity and come back towards where we started—but from the opposite direction. We can see this in a simple **topological** model of the projective plane—i.e. one that preserves its qualitative features. In this model a point on the boundary of the square is identified with its opposite.
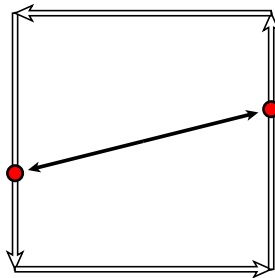
If we follow this procedure with an oriented square—i.e. pass it off through infinity and back again from the other direction—it comes back oriented the other way! This is shown in the following figures in the topological model:
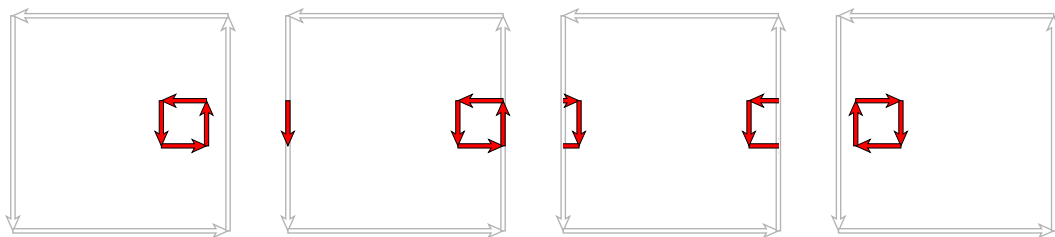
After a little thought, this makes sense. The opposite sides are identified, but with a twist much like that of the familiar Möbius strip. The projective plane is **non-oriented**, one-sided just like the Möbius strip.

There is still something unsatisfactory about my definition of the projective plane. If it really is a 2D surface, why can't we picture it as we do a sphere or a doughnut? Because it is nonorientable. Like the Klein bottle, which is also one-sided, it can only be realized in 3D as a surface that intersects itself. It is called **Boy's surface** after its

Boy:W.:4 discoverer, Werner Boy. Here are the original sketches (from his 1903 *Mathematische Annalen* article) showing how to construct a qualitatively correct version of it.

Fig. 19 a, b.

Fig. 19 c, d.

die positive Y-Achse und die r
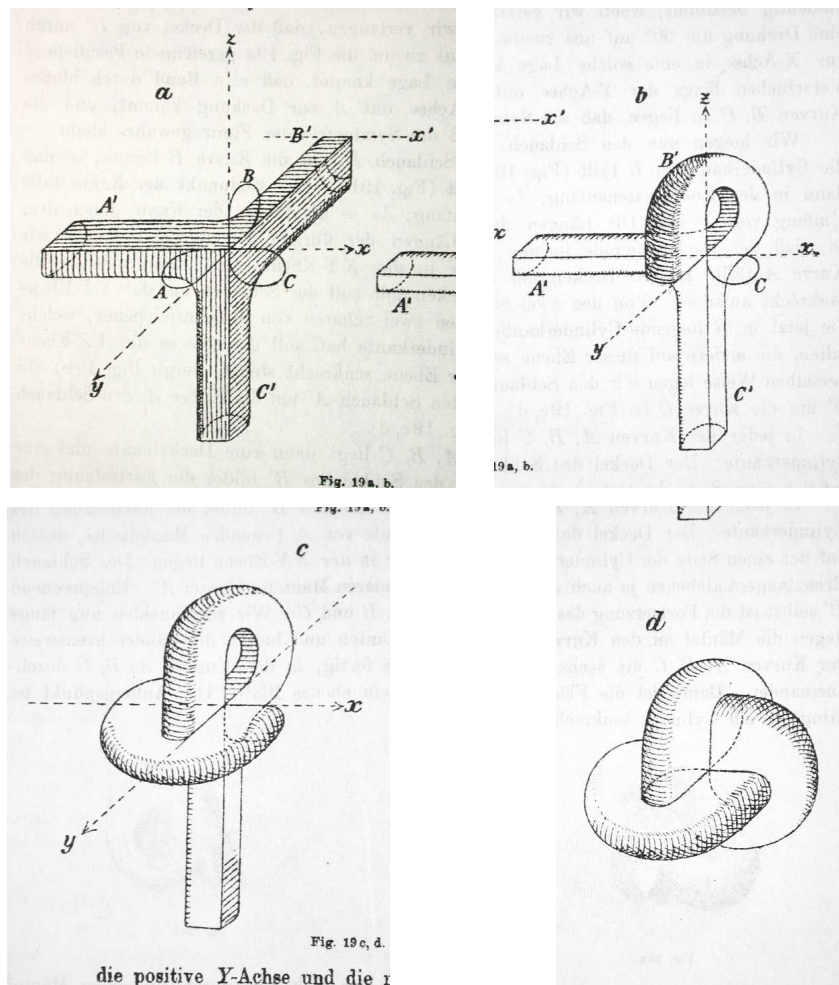
It is an interesting exercise to walk around on the surface and verify that it is one-sided. To see why it is a topological model for the projective plane is more difficult, and I won't really attempt it. One way to begin to understand what's at stake—keep in mind that each line through the origin in 3D intersects the unit sphere in exactly two opposite or **antipodal** points, say (x, y, z) and $(-x, -y, -z)$. Thus in modeling the projective plane as a surface, we are looking for a way to map the sphere onto a surface in which these two antipodal points collapse to one.

3. Projective transformations

The payoff in this business is that we can enlarge the collection of transformations we can apply to the points of the plane. We have seen that PostScript implicitly conceives of the affine transformation

$$x' = ax + cy + e$$
$$y' = bx + dy + f$$

in terms of the 3D linear transformation

$$[\,x'\quad y'\quad 1\,] = [\,x\quad y\quad 1\,]\begin{bmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{bmatrix}$$

This works because the 3D linear transformation corresponding to this particular matrix takes the plane $z = 1$ into itself. We now enlarge the transformations of the plane $z = 1$ by allowing right multiplication by an arbitrary 3×3 matrix which happens to be invertible—i.e. with non-zero determinant. Such a linear transformation takes lines through the origin into other lines through the origin, and therefore acts as a transformation of points of the projective plane—a **projective transformation**. It will not often take points with $z = 1$ into other points with $z = 1$, nor even necessarily into other points with $z \neq 0$, but it will usually do so. What does this mean in algebraic and geometrical terms?

Any 3×3 matrix acts as a transformation on the usual Euclidean space of 3 dimensions. Suppose T to be a 3×3 matrix, (x, y) a 2D point. How can T be used to transform this 2D point into another 2D point? (1) Extend it to a 3D point by adding on 1 as a last coordinate, so it becomes

$$(x, y, 1)$$

(2) The matrix T will transform this to another 3D point

$$(x, y, 1)\, T = (x', y', z')$$

(3) To get a 2D point back again, divide through by z' and throw away the final coordinate to get

$$(x'/z', y'/z') \ .$$

There will be problems when $z' = 0$, and we shall discuss these later.

4. The fundamental theorem

There is one fundamental theorem in this subject that makes the GIMP perspective tool practical. We have seen that the tool allows the user to move any one of four points around with the mouse to obtain the transformation desired. The mathematical assertion to go with this is:

- *Given four points P, Q, R, S in the projective plane with the property that no three of them lie on any one line. If P′, Q′, R′, S′ are another four points with the same property, then there exists a unique projective transformation taking P to P′, etc.*

I'll come back to this in a while, but place it in context first. Here is an analogous but simpler result:

- *Given three vectors u, v, w in 3D that are linearly independent. If u′, v′, w′ are any another three vectors, then there exists a unique 3×3 matrix transformation taking u to u′, etc.*

In the special case when u, v, and w are the three basis vectors

$$e_1 = [1, 0, 0]$$
$$e_2 = [0, 1, 0]$$
$$e_3 = [0, 0, 1]$$

the matrix can be written down explicitly. It is

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix},$$

whose rows are u', v', w'. In the general case, let T take e_1 to u etc. and T' take e_1 to u' etc.

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \xleftarrow{\;T\;} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} \xrightarrow{\;T'\;} \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix}$$

Then the transformation $T^{-1}T'$ with matrix

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix}^{-1} \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix}$$

takes u to u' etc.

Here is a consequence:

- *Given three points P, Q, R in the 2D plane, no two of which lie in a single line. If P', Q', R' are any another three 2D points, then there exists a unique affine transformation transformation taking P to P', etc.*

Embed these 2D points in 3D by making $z = 1$. According to the previous result, the 3D linear transformation whose matrix is

$$
\begin{bmatrix} x_P & y_P & 1 \\ x_Q & y_Q & 1 \\ x_R & y_R & 1 \end{bmatrix}^{-1} \begin{bmatrix} x_{P'} & y_{P'} & 1 \\ x_{Q'} & y_{Q'} & 1 \\ x_{R'} & y_{R'} & 1 \end{bmatrix}
$$

takes P etc. to P' etc. It is easy enough to check that it must then take the whole plane $z = 1$ into itself. From this it follows But then it can be checked in turn that this matrix has the form

$$
\begin{bmatrix} * & * & 0 \\ * & * & 0 \\ * & * & 1 \end{bmatrix}
$$

which means that it is the matrix of an affine transformation.

Exercise 1. *Prove the last two claims: (1) If a linear transformation takes three non-linear points on the plane $z = 1$ into other points on the same plane, then it takes all points on the plane $z = 1$ into points on the plane $z = 1$; (2) that if it takes the plane $z = 1$ to itself then it derives from an affine 2D transformation.*

Now let's return to the original claim about projective transformations. Translating it through the definition of projective points, it follows from this:

- *Given four vectors P, Q, R, S in 3D such that no three of them lie in one plane. If P', Q', R', S' are another four vectors with the same property, then there exists an invertible 3×3 matrix taking P to a scalar multiple of P', etc. It is unique up to a single scalar multiplication.*

In effect, the transformation takes the line through P to that through P' etc. The proof will construct the matrix explicitly. The basic idea is similar to that of the first result in this section. I first look at the special case

$$
\begin{aligned}
P &= e_1 = [1, 0, 0] \\
Q &= e_2 = [0, 1, 0] \\
R &= e_3 = [0, 0, 1 \\
S &= e_{123} = [1, 1, 1] \ .
\end{aligned}
$$

This special case is simple. Let M be the matrix with rows equal to P', Q', R':

$$
M = \begin{bmatrix} P' \\ Q' \\ R' \end{bmatrix} \ .
$$

By assumption on the vectors P' etc. it will be invertible. And it takes $[1, 0, 0]$ to P', $[0, 1, 0]$ to Q', $[0, 0, 1]$ to R', as you can calculate. By assumption also we can write

$$
S' = c_{P'} P' + c_{Q'} Q' + c_{R'} R' \ .
$$

We can find these coefficients by solving the 3×3 system of linear equations with coefficient matrix

$$
\begin{bmatrix} P' \\ Q' \\ R' \end{bmatrix} \ .
$$

The matrix

$$
\begin{bmatrix} c_{P'} P' \\ c_{Q'} Q' \\ c_{R'} R' \end{bmatrix}
$$

will have the same effect on the projective points $(1, 0, 0)$ etc. as M, and will take $[1, 1, 1]$ to S'.

In the general case, we can find transformations T taking e_1 to a multiple of P etc, and T' taking e_1 to a multiple of P' etc. But then the composite $T^{-1}T'$ takes P etc. to a multiple of P' etc.

5. Projective lines

A projective 2D point is now understood to be a row vector $[x, y, z]$, where x etc. are determined only up to a common non-zero-scalar multiple, and are not all equal to 0. If $[A, B, C]$ is another similar triple, then the set of points where

$$Ax + By + Cz = 0$$

is a plane—in fact a union of lines through the origin in 3D—and is called a **projective line**. Its intersection with the plane $z = 1$ is the familiar 2D line $Ax + By + C = 0$. We shall consider projective lines as column vectors, so $Ax + By + Cz$ is a matrix product.

Normally, the equation $Ax + By + C = 0$ with $A = B = 0$ makes no sense. But now in the projective plane, it becomes the plane $Cz = 0$ or just $z = 0$. It is the set of projective points at infinity, which is often called the **line at infinity**, or in effect the horizon.

If

$$T = \begin{bmatrix} t_{1,1} & t_{1,2} & t_{1,3} \\ t_{2,1} & t_{2,2} & t_{2,3} \\ t_{3,1} & t_{3,2} & t_{3,3} \end{bmatrix}$$

is a projective transformation, it acts on 2D points by taking

$$(x, y) \mapsto \left(\frac{t_{1,1}x + t_{2,1}y + t_{3,1}}{t_{1,3}x + t_{2,3}y + t_{3,3}}, \frac{t_{1,2}x + t_{2,2}y + t_{3,2}}{t_{1,3}x + t_{2,3}y + t_{3,3}}, \right) .$$

It will take a finite point off to infinity if the denominator vanishes. This will happen on the line whose coordinates are the last column of T.

6. A remark about solving linear systems

A system of n linear equations in n unknowns can be written as a matrix equation

$$xA = b$$

where A is an $n \times n$ matrix and has solution

$$x = bA^{-1} .$$

When A is 2×2 or 3×3 there is a formula for the inverse which is easy to program and relatively efficient. The same formula is valid in any number of dimensions, but for 4×4 matrices or larger it becomes impractical compared to other numerical methods, for example those involving Gauss elimination.

This formula involves the determinant of several matrices. The determinant $\det(A)$ of a square matrix A is a number that can be calculated in a variety of ways. One way is directly from its definition, which is, for a general matrix, rather complicated.

For 2×2 matrices, the formula is quite simple. If

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$$

then

$$\det(A) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1} .$$

But if A has size $n \times n$, then its determinant is a sum of signed products

$$\pm a_{1,c_1} a_{2,c_2} \ldots a_{n,c_n}$$

where the n-tuple $(c_1, \ldots, c_n)$ ranges over all permutations of $(1, 2, \ldots, n)$, so there are $n!$ in all—a potentially very large number. How is the sign determined? It is equal to the parity $\mathrm{sgn}(c)$ of the number of **inversions** of the permutation c, the pairs $c_i > c_j$ with $i < j$. So

$$\det A = \sum_c \mathrm{sgn}(c)\, a_{1,c_1} a_{2,c_2} \ldots a_{n,c_n}$$

where c ranges over all permutations of $[1, n]$.

For $n = 3$ we have the following parity assignments:

| Permutation	Inversions	Parity
$(1, 2, 3)$	$\emptyset$	1
$(2, 3, 1)$	$(3, 1), (2, 1)$	1
$(3, 1, 2)$	$(3, 1), (3, 2)$	1
$(2, 1, 3)$	$(1, 2)$	-1
$(1, 3, 2)$	$(3, 2)$	-1
$(3, 2, 1)$	$(3, 1), (3, 2), (2, 1)$	-1

If

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$
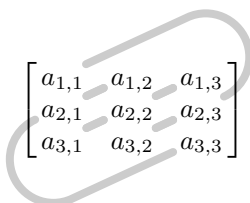
then this assignment of signs tells us that its determinant is

$$D = a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} - a_{1,1}a_{2,3}a_{3,2} - a_{1,3}a_{2,2}a_{3,1}\,.$$

The formula for this 3×3 determinant can be remembered easily if you note that the $+$ terms are products going in a sense from upper left to lower right
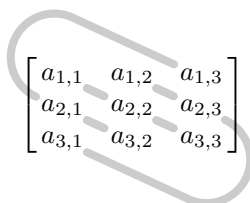
and the $-$ terms from lower left to upper right:

If A is any $n \times n$ matrix, its its **cofactor matrix** A^*, sometimes called its **adjoint**, is a new $n \times n$ matrix derived from it. For any $1 \le i \le n$, $1 \le j \le n$ the associated (i, j)-th **minor** is the determinant of the $(n-1) \times (n-1)$ matrix you get by cutting out from A the i-th row and j-th column. The cofactor matrix is that obtained from A whose (i, j)-entry $A^*_{i,j}$ is that minor, multiplied by $(-1)^{i+j}$. For example, the following figure shows that the cofactor entry $A^*_{1,2}$ is $-(a_{2,1}a_{3,3} - a_{2,3}a_{3,1})$.

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

The cofactor matrix is involved in another formula for the determinant:

$$\det(A) = \sum_{j=1}^{n} a_{i,j} A_{i,j}^{*}$$

for any choice of i. I'll leave it as an exercise to see that the two general formulas agree.

matrix:transpose:10 I recall also that the **transpose** ^{t}A of a matrix A is the matrix obtained by reflecting it around its principal diagonal. The formula for the determinant has as immediate consequence:

- *The inverse of a matrix A is the transpose of its cofactor matrix, divided by its determinant.*

For example, if

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

then its cofactor matrix is

$$A^{*} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix},$$

its transpose cofactor matrix is

$$^{t}A^{*} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

and its inverse is

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

You can check this if you multiply it by A.

The cofactor matrix of the 3×3 matrix $A = (a_{i,j})$ is

$$A^{*} = \begin{bmatrix} \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} & -\begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} & \begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} \\ -\begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} & \begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} & -\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} \\ \begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{vmatrix} & -\begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{2,1} & a_{2,3} \end{vmatrix} & \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} \end{bmatrix}.$$

Note that the signs form a checkerboard pattern like this:

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}.$$

You can check explicitly that $A\,^{t}A^{*} = I$.

Exercise 2. *Construct a PostScript procedure with one argument, a 3×3 matrix, that returns its inverse matrix. In this exercise, a 3×3 matrix should be an array of 3 arrays of 3 numbers, which are its rows:*

```
[[1 0 0] [0 1 0] [0 0 1]]
```

is the identity matrix.

7. The GIMP perspective tool, revisited

GIMP:10 How does the Theorem relate to the perspective tool? We start off with a rectangular image with corners at C_i for $i = 1, 2, 3, 4$. The mouse picks four points P_i somewhere in the image and calculates the perspective transformation taking each C_i to P_i, or equivalently the 3×3 matrix taking each

$$(C_{i,x}, C_{i,y}, 1) \longmapsto (P_{i,x}, P_{i,y}, 1)$$

in the sense of homogeneous coordinates. We can calculate this matrix explicitly. Let

$$\varepsilon_1 = (1, 0, 0)$$
$$\varepsilon_2 = (0, 1, 0)$$
$$\varepsilon_3 = (0, 0, 1)$$
$$\varepsilon_4 = (1, 1, 1)$$

Let σ be the 3×3 matrix taking the ε_i to the C_i, τ the one taking them to the P_i. Then $\sigma^{-1}\tau$ takes the C_i to the P_i.

Finally, the perspective transform tool first shows the effect of this matrix on a grid, and then when it has been accepted, applies it to every pixel in the image. Incidentally, the matrix that the tool displays is that of the projective transformation that we have constructed, except that GIMP follows mathematical conventions rather than those of computer graphics, so it is in fact the transpose of the one seen here.

Exercise 3. *Construct a PostScript program that reproduces with some flexibility the GIMP window at the beginning of this chapter. That is to say, it displays two 5×5 grids, one the transform of the other by a projective transformation corresponding to any given 3×3 matrix. Then set that matrix equal to the transpose of the one in the tool window:*

$$\begin{bmatrix} 0.535 & -0.219 & 39 \\ 0 & 0.5 & 24 \\ 0 & -0.00205 & 1 \end{bmatrix}$$

and verify that what you see is what you should see. There is one thing to be careful about—the origin of GIMP coordinates is at the upper left, with y increasing downwards. So start your drawing by translating your origin to the upper left corner, and then writing `1 -1 scale`.

8. Projections in 2D

ective projection:11 One special case of a transformation nicely described by 3×3 homogeneous matrices is projection onto a line from a point. It is not a projective transformation, since it collapses 2D onto 1D, and is hence singular, with no inverse. Suppose $P = (x_P, y_P)$ is the point from which things are projected, $f(x, y) = Ax + By + C = 0$ the line onto which things are projected. A point Q is transformed to the intersection of the line PQ with the line $f(x, y) = 0$, assuming it is not parallel to that line. This will be the point $R = (1 - t)P + tQ$ where $f(R) = 0$. We solve

$$f(R) = (1 - t)f(P) + tf(Q)$$
$$= 0$$
$$t(f(Q) - f(P)) = -f(P)$$
$$t = \frac{-f(Q)}{f(P) - f(Q)}$$
$$R = \frac{f(P)Q - f(Q)P}{f(P) - f(Q)} \ .$$

If we embed 2D in 3D this, then P becomes $(x_P, y_P, 1)$, Q becomes $(x, y, 1)$, and

$$R = (f(P)Q - f(Q)P, f(P) - f(Q))$$
$$= f(P)(x, y, 1) - f(Q)(x_P, y_P, 1)$$
$$= f(P)Q - f(Q)P \ .$$

This already looks promising, since we no longer have to contemplate dividing by 0—instead, when PQ is parallel to $f = 0$ the intersection will be a point at infinity. It becomes even more promising if we note that the expression $f(P)Q - f(Q)P$ is linear as a function of Q, so the projection arises from a linear transformation in 3D, or in other words a 3×3 matrix.

Exercise 4. *What is the matrix?*

9. Perspective in 3D

One way in which projective transformations originate is in viewing a plane in perspective in 3D. I'll look at just one example. Let's place a an observer in 3D at the point $E = (0, 0, e)$ with $e > 0$, looking down the z-axis in the negative direction. We are going to have him looking at images by projecting them onto the plane $z = 0$. That is to say, we are going to map points $P = (x, y, z)$ onto the point of intersection of the line from P to E with the viewing plane. This map is explicitly

$$(x, y, z) \longmapsto \left(\frac{ex}{e - z}, \frac{ey}{e - z} \right) \ .$$

I'll now restrict this map to a plane of the form $y = y_0$, so we have a 2D map

$$(x, z) \longmapsto \left(\frac{ex}{e - z}, \frac{ey_0}{e - z} \right) \ .$$

I claim that this can be seen as a projective transformation—that is to say, calculated in terms of a 3×3 homogeneous matrix acting on homogeneous coordinates. This is not quite obvious. Later on when we look at 3D drawing in more detail, we'll see it as a special case of a very general fact, but it's worthwhile doing a simple case first.
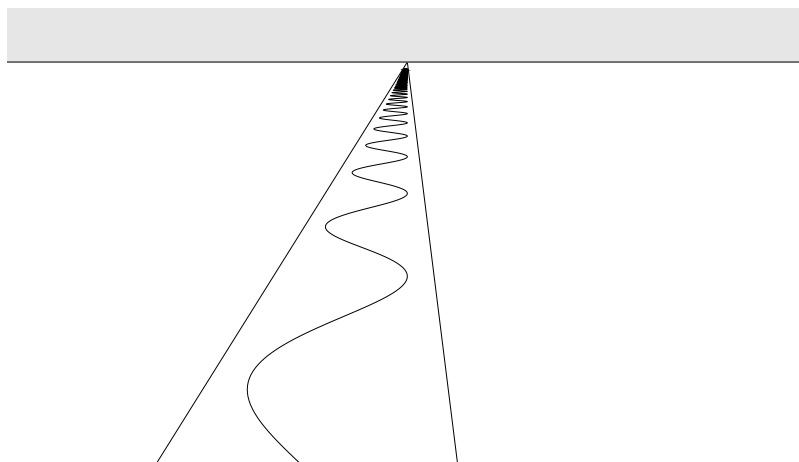
First we change (x, z) into $(x, z, 1)$, and then into (x, z, w). The map above then becomes the restriction to $(x, z, 1)$ of the homogeneous linear map

$$(x, z, w) \longmapsto (ex, ey_0, w, ew - z) \ .$$

or

$$\begin{bmatrix} x & z & w \end{bmatrix} \longmapsto \begin{bmatrix} x & z & w \end{bmatrix} \begin{bmatrix} e & & \\ & & -1 \\ & ey_0 & e \end{bmatrix} \ .$$

And of course these transforms may be combined with the techniques of Chapter 8.

Exercise 5. *Construct a PostScript procedure with four 2D points P_i as arguments, returning the 3×3 matrix taking*

$$(0, 0) \mapsto P_0$$
$$(1, 0) \mapsto P_1$$
$$(0, 1) \mapsto P_2$$
$$(1, 1) \mapsto P_3 \ .$$

References

Gray:J.:13   1. D. A. Brannan, M. Esplen, and J. Gray, **Geometry**, Cambridge University Press, 1999. The literature on projective geometry is large. With the advent of computers, the goals of the subject have become less abstract, so that older books tend to look a bit out of date. This is a recent comprehensive text that illustrates the trend towards less abstraction. Lots of pictures.

Kemp:M.:13   2. M. Kemp, **The Science of Art: optical themes in western art from Brunelleschi to Seurat**, Yale University Pres, 1990. The literature on perspective is even more vast. This is an extremely thorough history, with lots of implicit suggestions for projects. in mathematical graphics.

Blinn:J.:13   3. Three books by the well known expert in computer graphics, Jim Blinn—

> **Jim Blinn's Corner: a Trip Down the Graphics Pipeline**, Morgan Kaufmann, 1996
> **Jim Blinn's Corner: Dirty Pixels**, Morgan Kaufmann, 1998
> **Jim Blinn's Corner: Notation, Notation, Notation**, Morgan Kaufmann, 2003

—offer a well motivated outsider's look at geometry. The essays are taken from Blinn's column in the journal *IEEE Computer Graphics and Applications*. Most are interesting, and some are absolutely delightful. Especially valuable are his essays on homogeneous coordinates, but there's a lot of other mathematics and mathematical pictures to be found in these volumes.

Glassner:A.:13   4. Andrew Glassner, **Andrew Glassner's Notebook: Recreational Computer Graphics**, Morgan Kaufmann, 1999. I mention this book here because it has much in common with Blinn's, although not much with the topic of this chapter. Like Blinn's books, it is based on a column in *IEEE Computer Graphics and Applications*, and amounts to a collection of improved essays from that column. More mathematical, less practical than Blinn's books. For that reason, not quite so interesting to a mathematician. Much fancier pictures!

GIMP:13   5. The program **GIMP** is available on essentially all operating systems, and without cost for Linux and Windows. Invaluable as an image manipulation tool for most. Almost as capable as **Photoshop** and—for better or worse— about as easy to use. But for simple tasks it's fine. The **GIMP** home page is at

> `http://www.gimp.org/`

Boy:W.:13   6. Werner Boy, 'Über die Curvatura integra', *Mathematische Annalen* (**57**), 1903, 151–184.

This paper can be seen in the digitized edition of the Annalen at

> `http://134.76.163.65/agora_docs/25917TABLE_OF_CONTENTS.html`

which is part of the Göttingen digitalization project.

Cohn-Vossen:S.:13   7. Stephan Cohn-Vossen and David Hilbert, **Geometry and the Imagination**, Chelsea, 1952. Werner Boy was a student of Hilbert's, and his surface is said to have been a surprise to his advisor. Several variants of the Apéry:F.:13 surface are discussed on pp. 317–321 of Hilbert & Cohn-Vossen, but alas! without any history. Franços Apéry has written the beautiful book **Models of the real projective plane** (Vieweg, 1987) on the same subject. It includes many handsome computer-generated pictures, and explains how they were generated. Many other computer-generated pictures of the surface immersed in 3D can be found on the 'Net, but I find Boy's original figures easiest to understand.