This document summarizes the non-isometry 4×4 matrices that we use in computer graphics. There is one matrix for the viewport transformation, and there are two matrices for the projections. The inverses to these three matrices are also important. This tutorial also describes the operation of homogeneous division, just so that we can relate it to the matrices.

1 Homogeneous division

Suppose that we have a 4×1 homogeneous vector

$$\vec{v} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}.$$

Thus far in the course, w has always been 1, but that soon changes. If we have a non-zero w, then we can divide the entire vector by w to get a new vector with a 1 in its last component:

$$\frac{1}{w}\vec{v} = \begin{bmatrix} x/w \\ y/w \\ z/w \\ 1 \end{bmatrix}.$$

This operation is called *homogeneous division*.

Concretely, here's what you need to know in CS 311. You need to know what I mean when I say "homogeneous division". You need to know that homogeneous division fails when $w = 0$. Those are the main things. They're not difficult.

2 Optional: What It Means

The meaning of homogeneous division, and of homogeneous coordinates in general, is a topic in mathematics called *projective geometry*. I give you a short explanation in this section. If you want to know more, then talk to me in person.

Projective space is like our ordinary three-dimensional space, but with extra points "at infinity". They arise in a really simple way. Homogeneous vectors

$$\vec{v} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$$

represent points in projective space. So you might think, "There are four numbers in there, so aren't we talking about four-dimensional space?" No. There are two peculiarities — one major and one minor — that you must keep in mind.

- The major peculiarity is that each point in this space is represented by (infinitely) many four-dimensional vectors. Namely, if two four-dimensional vectors are scalar multiples of each other, then they represent the same point. In particular, if the vector $\vec{v}$ above has a non-zero w-component, then $\vec{v}$ and $\frac{1}{w}\vec{v}$ represent the same point.

- The minor peculiarity is that the zero vector, which consists of four zeros, does not represent any point in this space. When doing projective geometry, just pretend that there is no zero vector.

It follows that the points in projective space fall into two camps, as we now explore.

The first camp consists of those points $\vec{v}$ where the last component w is non-zero. For such a point, we can perform homogeneous division to make the last component 1, without changing the point that we're talking about. So the first camp is essentially the points of the form

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}.$$

These points naturally correspond to points (x, y, z) in ordinary three-dimensional space.

The second camp consists of those points $\vec{v}$ where the last component w is zero. These points are not part of ordinary three-dimensional space. They lie outside it. They are called *points at infinity*. To see why, consider for example the sequence of 3D points

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 10 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 10^2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 10^3 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 10^4 \\ 0 \\ 0 \end{bmatrix}, \ldots$$

which zoom along the x-axis. Intuitively, the sequence approaches

$$\begin{bmatrix} \infty \\ 0 \\ 0 \end{bmatrix}.$$

But ∞ is not a number. Don't try calculating with it; it will bring you only misery. Remarkably, the situation improves if we shift to projective geometry. The homogeneous version of the sequence is

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10^2 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10^3 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10^4 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \ldots.$$

But, according to the rules of projective geometry, those points are identical to

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 10^{-1} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 10^{-2} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 10^{-3} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 10^{-4} \end{bmatrix}, \ldots,$$

and they approach

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

which is a legitimate point in projective space. More generally, if you have sequence of three-dimensional points

$$c \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} cx \\ cy \\ cz \end{bmatrix}$$

with c going to infinity, then the projective version is

$$\begin{bmatrix} cx \\ cy \\ cz \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ c^{-1} \end{bmatrix},$$

which approaches

$$\begin{bmatrix} x \\ y \\ z \\ 0 \end{bmatrix}$$

as c goes to infinity. This is the sense in which projective points, where the last coordinate is 0, are points "at infinity".

If you take a course in drawing, then you might learn about vanishing points, where parallel lines in the world intersect in the drawing. Vanishing points are points at infinity. For example, we saw above that the sequence

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10^2 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10^3 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10^4 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \ldots,$$

which lies along the line $y = 0$ in the x-y-plane, approaches

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Consider also the sequence

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10^2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10^3 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 10^4 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \ldots$$

which lies along the line $y = 1$ in the x-y-plane. It approaches the same point at infinity. So the two lines, which are parallel, intersect at that point!

3 Viewport

A viewport with bottom-left corner at $(0, 0)$ and top-right corner at (r, t) has viewport transformation matrix

$$V = \begin{bmatrix} \frac{r}{2} & 0 & 0 & \frac{r}{2} \\ 0 & \frac{t}{2} & 0 & \frac{t}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The inverse matrix is

$$V^{-1} = \begin{bmatrix} \frac{2}{r} & 0 & 0 & -1 \\ 0 & \frac{2}{t} & 0 & -1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Optional exercise: Start with an arbitrary homogeneous vector $\vec{v}$. Transform $\vec{v}$ by the viewport and then perform homogeneous division. Also perform homogeneous division on $\vec{v}$ and then transform by the viewport. Show that these two computations give the same answer.

To understand what the viewport transformation does, consider the eight vectors

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.$$

In each of the eight vectors, the last coordinate is 1, so we can interpret the vector as an ordinary 3D point. These eight points define the corners of a cube of side length 2. This cube is called "the viewing volume in normalized device coordinates". The viewport maps this cube to a certain parallelepiped (box), whose eight vertices are respectively

$$\begin{bmatrix} r \\ t \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} r \\ t \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} r \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} r \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ t \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ t \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

This parallelepiped is called "the viewing volume in screen coordinates".

I now restate the same facts in different language. The viewport transforms the cube $[-1, 1] \times [-1, 1] \times [-1, 1]$ to the parallelepiped $[0, r] \times [0, t] \times [0, 1]$. The transformation is of a simple kind, being accomplished by a stretch and a translation in each dimension. The transformation preserves the "sense" in each dimension. In particular, in the third dimension, $[-1, 1]$ maps to $[0, 1]$ such that -1 goes to 0 and 1 goes to 1.

4 Orthographic Projection

An orthographic projection has a viewing volume that is a parallelepiped. In the first direction it extends from ℓ (left) to r (right). In the second it extends from b (bottom) to t (top), and in the third from f (far) to n (near). In Cartesian product notation, it is the set $[\ell, r] \times [b, t] \times [f, n]$. The projection matrix is

$$P = \begin{bmatrix} \frac{2}{r-\ell} & 0 & 0 & \frac{-r-\ell}{r-\ell} \\ 0 & \frac{2}{t-b} & 0 & \frac{-t-b}{t-b} \\ 0 & 0 & \frac{-2}{n-f} & \frac{n+f}{n-f} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The inverse matrix is

$$P^{-1} = \begin{bmatrix} \frac{r-\ell}{2} & 0 & 0 & \frac{r+\ell}{2} \\ 0 & \frac{t-b}{2} & 0 & \frac{t+b}{2} \\ 0 & 0 & \frac{f-n}{2} & \frac{f+n}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Geometrically, the orthographic projection maps its viewing volume $[\ell, r] \times [b, t] \times [f, n]$ to the cube $[-1, 1] \times [-1, 1] \times [-1, 1]$. The transformation is of a simple kind. In the first two dimensions it consists of stretching and translation. In the third dimension, it consists of a negation, stretching, and translation. In particular, in the third dimension, $[f, n]$ maps to $[-1, 1]$ such that f goes to 1 and n goes to -1.

5 Perspective Projection

A perspective projection has a viewing volume that is a frustum (truncated pyramid). In the third dimension it extends from f to n. The plane where the third coordinate equals n is called the *near plane*. On the near plane, the frustum extends from ℓ to r in the first dimension and from b to t in the second. On the far plane, the frustum extends from $\ell \cdot \frac{f}{n}$ to $r \cdot \frac{f}{n}$ in the first dimension and from $b \cdot \frac{f}{n}$ to $t \cdot \frac{f}{n}$ in the second. The projection matrix is

$$P = \begin{bmatrix} \frac{-2n}{r-\ell} & 0 & \frac{r+\ell}{r-\ell} & 0 \\ 0 & \frac{-2n}{t-b} & \frac{t+b}{t-b} & 0 \\ 0 & 0 & \frac{n+f}{n-f} & \frac{-2nf}{n-f} \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

The inverse matrix is

$$P^{-1} = \begin{bmatrix} \frac{r-\ell}{-2n} & 0 & 0 & \frac{r+\ell}{-2n} \\ 0 & \frac{t-b}{-2n} & 0 & \frac{t+b}{-2n} \\ 0 & 0 & 0 & -1 \\ 0 & 0 & \frac{n-f}{-2nf} & \frac{n+f}{-2nf} \end{bmatrix}.$$

Like orthographic projection, perspective projection maps its viewing volume to the cube $[-1, 1] \times [-1, 1] \times [-1, 1]$. In the third dimension, f goes to 1 and n goes to -1.