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

This document only briefly explains the geometric intuition and meaning. See class for more intuition and meaning.

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.

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 below. 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". Homogeneous vectors

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

represent points in projective space. There are two odd features of this relationship. First, the homogeneous $\vec{0}$ vector does not represent any point in projective space. Second, multiple vectors can represent the same point. Precisely, two vectors represent the same point if they differ by homogeneous division (or homogeneous multiplication). Therefore the points in homogeneous space fall into two camps.

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 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. In a sense, they lie outside it. We call them *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.

2 Viewport

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

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

The inverse matrix is

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

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. Do these two computations give the same answer?

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

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

Interpreted as points in everyday three-dimensional space, they 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). For example, the first and last points map to

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

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

At the risk of not being understood, let me 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.

3 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 transformation matrix is

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

The inverse matrix is

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

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.

4 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 transformation matrix is

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

The inverse matrix is

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

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.