CO317 - Graphics (60005)

Lecture 1 - Projections and Transformations

Two Dimensional Graphics

At the lowest level, in every operating system, graphics processing operates on the pixels in a window with primitives, such as;

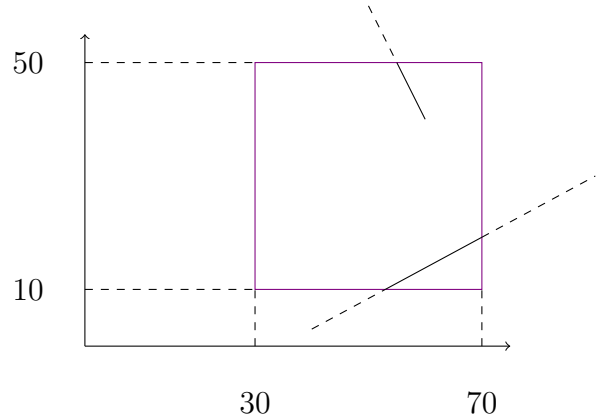
- `SetPixel(int x, int y, int colour);`
- `DrawLine(int xs, int ys, int xf, int yf);`

However, we'd like to be able to draw scenes from a three-dimensional world and have it appear in two-dimensional graphics primitives.

World Coordinate System

In order to achieve independence when drawing objects, we define a world coordinate system. For example, let our world be defined in meters, we can then allow a pixel to represent a millimetre. A viewing area is a window, and is defined as part of our 3D world. **Clipping** occurs when we attempt to draw outside (dashed) of the window;

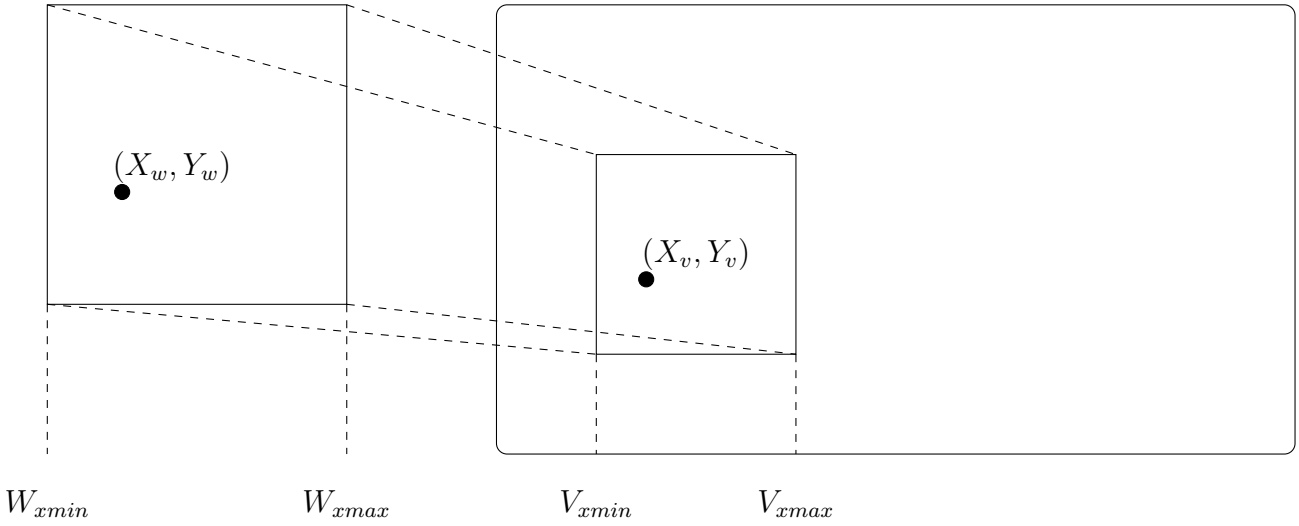
```
SetWindow(30, 10, 70, 50)
DrawLine(40, 3, 90, 30)
DrawLine(50, 60, 60, 40)
```



However, this isn't as trivial to do in 3D, as it cannot simply be left to the operating system. While we can represent 3D objects as a series of 2D commands, it's inefficient and expensive for the OS to perform the clipping (therefore we should do this manually).

Normalisation

A normalisation process is required to convert from device independent commands (where screen resolution isn't taken into account) to drawing commands using pixels. Consider a point in the world coordinate window (X_w, Y_w), and its corresponding result on the viewport (pixel coordinates; (X_v, Y_v));



The expressions are similar for Y;

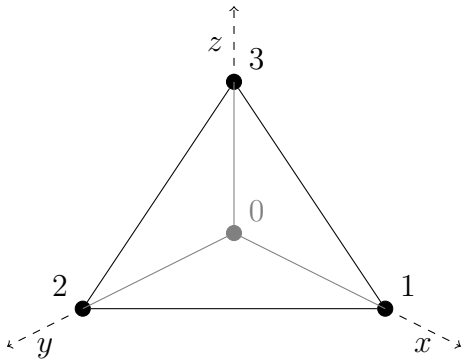
$$\frac{(X_w - W_{xmin})}{(W_{xmax} - W_{xmin})} = \frac{(X_v - V_{xmin})}{(V_{xmax} - V_{xmin})} \Rightarrow X_v = \frac{(X_w - W_{xmin})(V_{xmax} - V_{xmin})}{(W_{xmax} - W_{xmin})} + V_{xmin}$$

This gives us the resulting pair of linear equations (intuitively), where the constants found from the known values W_{xmin}, V_{xmax}, etc. are used to define the normalisation;

$$X_v = AX_w + B$$
$$Y_v = CY_w + D$$

Polygon Rendering

Most graphics applications deal with very simple objects - flat / planar polyhedra, referred to as **faces** or **facets**. These are graphic primitives, and can be used to approximate any shape. Consider the following tetrahedron, consisting of four vertices;



For this, we need a mixture of different data, including numerical data about the actual 3D coordinates of the vertices, as well as topological data regarding what vertices are connected to what. This can be represented in the following tables;

vertex data

index	location
0	$(0, 0, 0)$
1	$(1, 0, 0)$
2	$(0, 1, 0)$
3	$(0, 0, 1)$

face data

index	vertices
0	0 1 3
1	0 2 1
2	0 3 2
3	1 2 3

This separation allows for the vertices to move without affecting the faces.

Projections

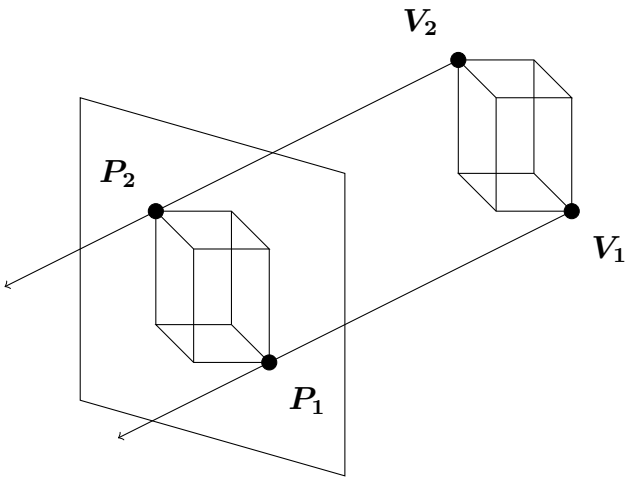
In order to draw a 3D wire frame, the points must first be converted into a 2D representation, via a **projection**, which can then be drawn with simple drawing primitives. Intuitively, we have an observer (a focal point) where all viewing rays converge. The observer is located between a projection surface P and an object V. While it's possible to project onto any surface, we only consider linear projections onto a flat surface.

Orthographic Projections

The simplest form of a projection is an **orthographic projection**. The assumptions made are that the viewpoint is located at $z = -\infty$, and the plane of projection is $z = 0$. With the viewing point being infinitely far away, the rays become parallel. This gives all projectors the same direction;

$$\boldsymbol{d} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

This gives the following, with each projection line having the equation $\boldsymbol{P} = \boldsymbol{V} + \mu\boldsymbol{d}$;



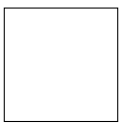
By substituting in the direction $\boldsymbol{d}$ we have determined, it gives the following Cartesian equations for each component;

$$P_x = V_x + 0$$
$$P_y = V_y + 0$$
$$P_z = V_z - \mu$$

However, since we have the projection plane $z = 0$, we also know that $P_z = 0$, therefore we don't need to solve for μ. From this, we can determine the projected location is the 3D x and y components of the vertex;

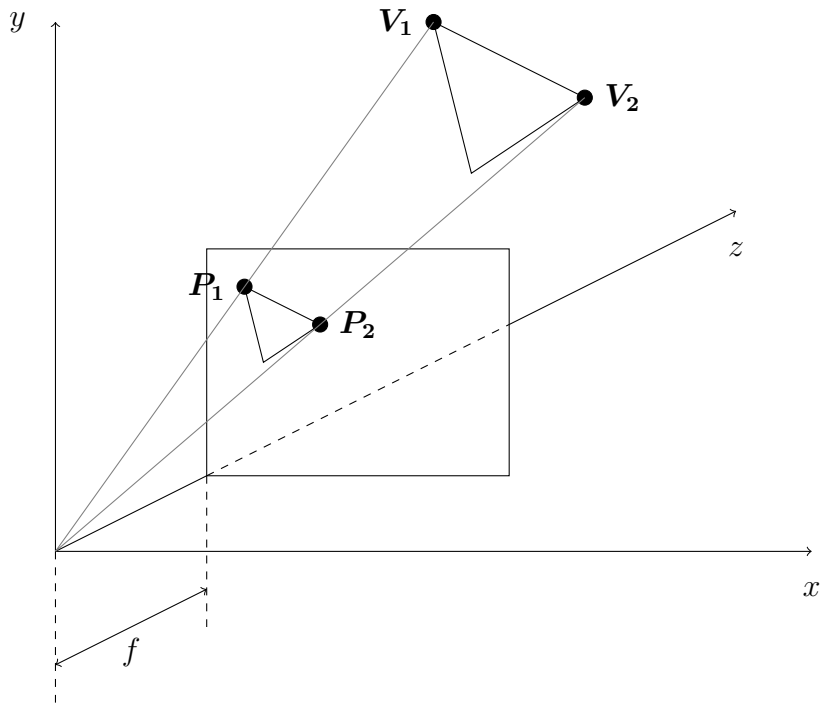
$$\boldsymbol{P} = \begin{bmatrix} V_x \\ V_y \\ 0 \end{bmatrix}$$

Viewing the wireframe for a cube directly from a face would look like the following;



Perspective Projection

While orthographic projections are fine when depth isn't a consideration (such as objects mostly being at the same distance from the viewer), it's insufficient for close work, where we want details to be realistic. The difference here is that we are no longer at an infinite distance (instead being at the origin), and the projection plane is $z = f$ (where f stands for focal length);



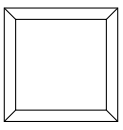
This gives us the following equation (since all projectors must go through the origin);

$$\boldsymbol{P} = \mu \boldsymbol{V}$$

We can work out the value of μ, let it be μ_p as follows;

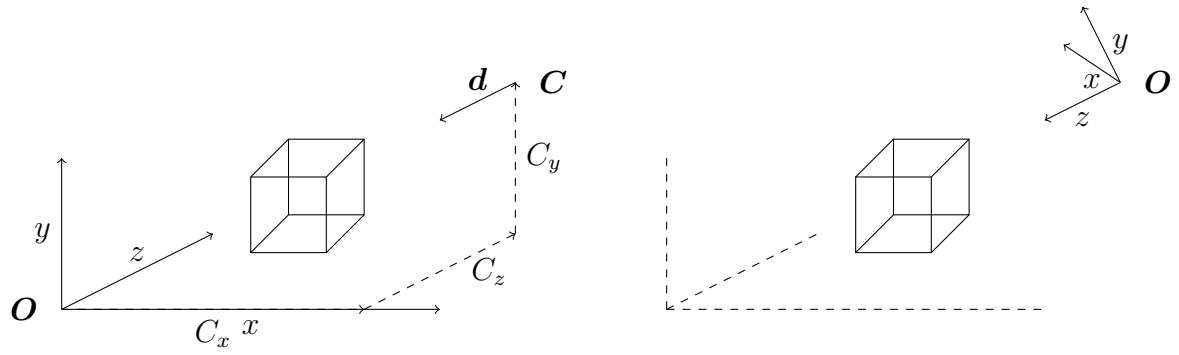
$$\begin{aligned}
P_z &= f && \text{by projection plane} \\
\mu_p &= \frac{P_z}{V_z} \\
&= \frac{f}{V_z} \\
P_x &= \mu_p V_x \\
&= \frac{f V_x}{V_z} \\
P_y &= \mu_p V_y \\
&= \frac{f V_y}{V_z}
\end{aligned}$$

Viewing the wireframe for a cube directly from a face would look like the following (note the difference to the orthographic projection);



Transformations

Scenes are defined in a particular coordinate system, but we want to be able to draw a scene from any angle. To do so, it's easier to have the viewpoint at the origin, and the z-axis as the direction of view. As such, we need to be able to **transform** the coordinates of a scene.



These are done by the application of transformation matrices. For example, a standard transformation to make an object twice as big from the origin;

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Translation

However, being restricted to matrix operations with $\mathbb{R}^{3\times 3}$ means that we cannot represent translations (for example, a shift of two units on the x-axis, such that $x' = x + 2$). The solution to this is to use 4D **homogenous coordinates**, where we assume the fourth dimension is fixed to 1.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Frequently the last ordinate is 1, however in general it is a scale factor;

$$\underbrace{(p_x, p_y, p_z, s)}_{\text{homogeneous}} \Leftrightarrow \underbrace{\left(\frac{p_x}{s}, \frac{p_y}{s}, \frac{p_z}{s}\right)}_{\text{Cartesian}}$$

Affine Transformations

Affine transformations preserve parallel lines. Most of the transformations we require are affine, with the most important being scaling, rotation, and translation;

- **scaling** by (s_x, s_y, s_z)

$$\begin{bmatrix} s_x & 0 & 0 & 1 \\ 0 & s_y & 0 & 1 \\ 0 & 0 & s_z & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \\ 1 \end{bmatrix} = \begin{bmatrix} s_x p_x \\ s_y p_y \\ s_z p_z \\ 1 \end{bmatrix}$$

- **rotation**

  In order to define a rotation, we need both an axis and an angle, with the simplest rotations being about the Cartesian axes. The following matrices are used for rotations of θ about each of the axes;

$$\boldsymbol{\mathcal{R}_x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{\mathcal{R}_y} = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{\mathcal{R}_z} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

It's important to note that **rotations have a direction**. In this course, we use a left-handed coordinate system, where the rotation is anti-clockwise when looking along the axis of rotation (think about the origin being closer to you, and the axis going off to ∞ away from you).

- **translation** by (t_x, t_y, t_z)

$$\begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \\ 1 \end{bmatrix} = \begin{bmatrix} p_x + t_x \\ p_y + t_y \\ p_z + t_z \\ 1 \end{bmatrix}$$

However, perspective projections are an example of a non-affine transformation, as it doesn't preserve parallels. Intuitively, it's not invertible (singular), as we cannot convert from a photograph to a 3D model.

Note that we should be careful when we combine transformation. As matrix multiplication isn't commutative, we should read a sequence of matrices multiplied together from right to left, with the right-most matrix, before the vector, being the **first** transformation to be applied.