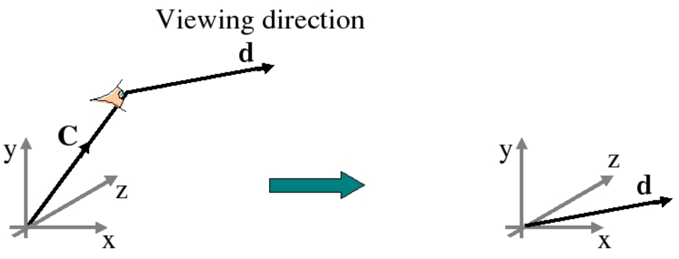
Lecture 2: Scene Transformation and Animation

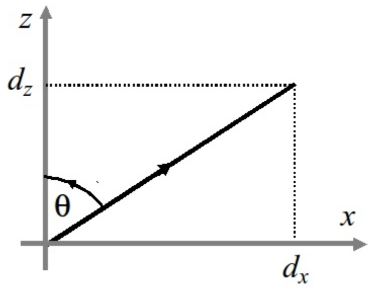
Flying Sequences

We will now consider a most important subject, namely, scene transformation. In any viewer centered application, such as a flight simulator or a computer game, we need to view the scene from a moving position. As the viewpoint changes we transform all the coordinates of the scene such that the viewpoint is the origin and the view direction is the z axis, before projecting and drawing the scene. Let us suppose that, in the coordinate system in which the scene is defined we wish to view it from the point $\mathbf{C} = [C_x, C_y, C_z]$, looking along the direction $\boldsymbol{d} = [d_x, d_y, d_z]$. The first step is to move the origin to $\mathbf{C}$ for which we use the transformation matrix $\mathcal{A}$.



$$\mathcal{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -C_x & -C_y & -C_z & 1 \end{bmatrix}$$

Following this, we wish to rotate about the y-axis so that $\mathbf{d}$ lies in the plane $x = 0$. Using the fact that $\mathbf{d}$ is defined by the co-ordinates $[d_x d_y d_z]$ and using the notation $v^2 = d_x^2 + d_z^2$ this is done by matrix $\mathcal{B}$.



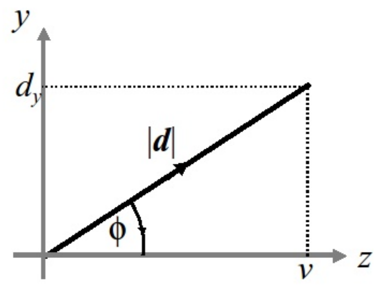
$$v = \sqrt{d_x^2 + d_z^2}$$

$$Cos\theta = d_z/v$$

$$Sin\theta = d_x/v$$

$$\mathcal{B} = \begin{bmatrix} d_z/v & 0 & d_x/v & 0 \\ 0 & 1 & 0 & 0 \\ -d_x/v & 0 & d_z/v & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Notice that we have avoided computing the Cos and Sin functions for this rotation by use of the direction cosine. To get the direction vector lying along the z axis a further rotation is needed. This time it is about the x axis using matrix $\mathcal{C}$.



$$v = \sqrt{d_x^2 + d_z^2}$$

$$Cos\phi = v/|d|$$

$$Sin\phi = d_y/|d|$$

$$\mathcal{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & v/|d| & d_y/|d| & 0 \\ 0 & -d_y/|d| & v/|d| & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Finally the transformation matrices are combined into one, and each point of the scene is transformed:

$$\mathcal{T} = \mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C}$$

and so for all the points

$$\mathbf{P_t} = \mathbf{P} \cdot \mathcal{T}$$

Problems with verticals

The concept of "vertical" is missing from the above analysis, and needs attention since it is easy to invert the vertical. This will be easily observed in the trivial example of figure 1, where an arrow whose base is at [0,0,-l] is being observed from the origin. A transformation based on rotating about the y axis first yields the correct solution. However, a transformation involving rotation about the x axis first inverts the image.

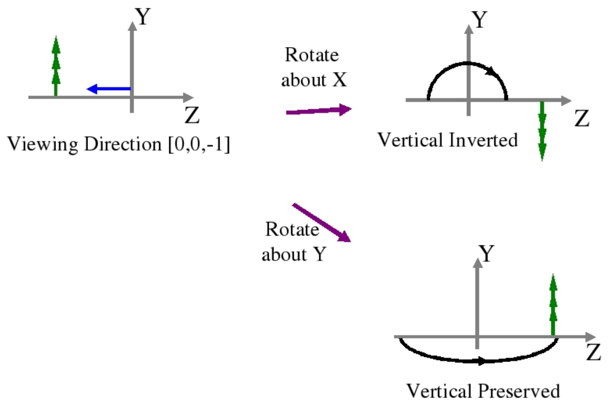


Figure 1: Inversion of the Vertical

Rotation about a general line

A very similar problem to the viewer centered transformation concerns animation of objects in a fixed scene. Let us suppose that we want to rotate one sub-object about some line in the Cartesian space where the scene is defined. Let the line be: $\mathbf{L} + \mu\mathbf{d}$ where $\mathbf{L} = [L_x, L_y, L_z]$ is the position vector of any point on the line and $\mathbf{d}$ is a unit direction vector along the line. We use a translation to move the origin so that it is on the line (matrix $\mathcal{A}$ but transforming the origin to $\mathbf{L}$ rather than $\mathbf{C}$) followed by two rotations to make the z axis coincident with the direction vector $\mathbf{d}$ (matrices $\mathcal{B}$ and $\mathcal{C}$). Now we can perform a rotation of the object about the z-axis using the standard rotation matrix $\mathcal{R}_z$ defined previously. Finally we need to restore the coordinate system as it was, such that the viewpoint is the same as before. To do this we simply invert the transformation matrices $\mathcal{A}$ $\mathcal{B}$ and $\mathcal{C}$. As before we multiply all the individual transformation matrices together to make one matrix which is then applied to all the points.

$\mathcal{T} = \mathcal{A} \cdot \mathcal{B} \cdot \mathcal{C} \cdot \mathcal{R}_z \cdot \mathcal{C}^{-1} \cdot \mathcal{B}^{-1} \cdot \mathcal{A}^{-1}$

and for all the points $\mathbf{P_t} = \mathbf{P} \cdot \mathcal{T}$

Other object transformations for graphical animation are performed similarly. For example to make an object shrink we move the origin to its centre, perform a scaling and then restore the origin to its original position.

Projection by Matrix Multiplication

If we use homogeneous co-ordinates then it is also possible to compute projection by multiplication by a projection matrix. Placing the centre of projection at the origin and using $z = f$ as the projection plane gives us matrix $\mathcal{M}_p$ for perspective projection. Matrix $\mathcal{M}_o$ is for orthographic projection:

$$\mathcal{M}_p = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/f \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \mathcal{M}_o = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

It is not immediately obvious that matrix $\mathcal{M}_p$ produces the correct perspective projection. Let us transform an arbitrary point $\mathbf{V}$ with homogeneous co-ordinates $[x, y, z, 1]$ by using matrix multiplication. For the projected point $\mathbf{P}$ we get:

$$\mathbf{P} = \mathbf{V} \cdot \mathcal{M}_p = [x, y, z, z/f]$$

This point must be normalised into a Cartesian coordinate. To do this we divide the first three co-ordinates by the value of the fourth and we get:

$$\mathbf{P_c} = [xf/z, yf/z, f, 1]$$

which is the projected point. It is interesting to note that the projection matrix is obviously singular (it has a row of zeros) and, therefore, it has no inverse. This must be so because it is impossible to reconstruct a 3D object from its 2D projection without other information. Projection matrices can of course be combined with the other matrices. Indeed the popularity of the orthographic projection is that it simplifies the amount of calculations since the z row and column fall to zero. Although a simplification applies when the perspective projection is used, there is also the need to normalise the resulting homogenous coordinates, and this adds to the computation time.

Homogenous coordinates

We now take a second look at homogeneous coordinates, and their relation to vectors. Previously, we described the fourth ordinate as a scale factor, and ensured that, with the exception of the projection transformation, it was always normalised to 1. As an alternative, we can consider the fourth ordinate as indicating a type as follows. Informally we acknowledge that a normal Cartesian coordinate is a special form of vector, which we call a position vector, and usually denote using bold face capital letters. Hence, we can say that a normalised homogenous coordinate is the same as a position vector.

By contrast, if we consider a homogenous coordinate where the last ordinate is zero - $[x, y, z, 0]$ - we find that we cannot normalise it in the usual way because of the divide by zero. So clearly a homogenous coordinate of this form cannot be directly associated with a point in Cartesian space. However, it still has a magnitude and direction, and hence we can consider it to be a direction vector. Thus homogenous coordinates fall into two classes, those with the final ordinate non-zero, which can be normalised into position vectors or points, and those with zero in the final ordinate which are direction vectors, or vectors in the pure sense of the word.

Consider now how vector addition works with these definitions. If we add two direction vectors, we add the ordinates as before and we obtain a direction vector. ie:

$$[x_i, y_i, z_i, 0] + [x_j, y_j, z_j, 0] = [x_i + x_j, y_i + y_j, z_i + z_j, 0]$$

This is the normal vector addition rule which operates independently of Cartesian space. However, if we add a direction vector to a position vector we obtain a position vector or point:

$$[X_i, Y_i, Z_i, 1] + [x_j, y_j, z_j, 0] = [X_i + x_j, Y_i + y_j, Z_i + z_j, 1]$$

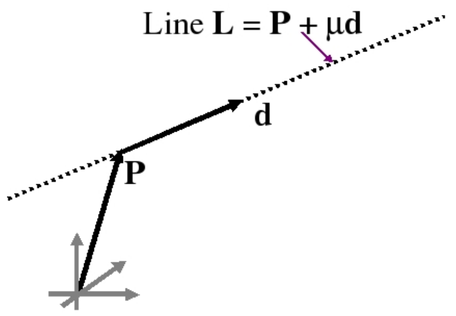


Figure 2: Definition of a straight line

This is a nice result, because it ties in with our definition of a straight line in Cartesian space, which is the sum of a point and a scaled direction as shown in figure 2.

Now, consider a general affine transformation matrix. We ignore the possibility of doing perspective projection or shear, so that the last column will always be $[0, 0, 0, 1]^T$, and the matrix will be of the form shown in figure 3, with the rows viewed as three direction vectors and a position vector.

$$\begin{array}{cl} \begin{bmatrix} q_x & q_y & q_z & 0 \\ r_x & r_y & r_z & 0 \\ s_x & s_y & s_z & 0 \\ C_x & C_y & C_z & 1 \end{bmatrix} & \begin{array}{l} \text{Direction vector} \\ \text{Direction vector} \\ \text{Direction vector} \\ \text{Position vector} \end{array} \end{array}$$

Figure 3: The structure of a space transformation matrix

To see what the individual rows mean we consider the effect of the transformation in simple cases. For example take the unit vectors along the Cartesian axes eg $\boldsymbol{i} = [1, 0, 0, 0]$:

$$[1, 0, 0, 0] \begin{bmatrix} q_x & q_y & q_z & 0 \\ r_x & r_y & r_z & 0 \\ s_x & s_y & s_z & 0 \\ C_x & C_y & C_z & 1 \end{bmatrix} = [q_x, q_y, q_z, 0]$$

In other words the direction vector of the top row represents the direction in which the x axis points after transformation, and similarly we find that $\boldsymbol{j} = [0, 1, 0, 0]$ will be transformed to direction $[r_x, r_y, r_z, 0]$ and $\boldsymbol{k} = [0, 0, 1, 0]$ will be transformed to $[s_x, s_y, s_z, 0]$. Similarly, we can see the effect of the bottom row by considering the transformation of the origin which has homogeneous coordinate $[0, 0, 0, 1]$. This will be transformed to $[Cx, Cy, Cz, 1]$.

$$[0, 0, 0, 1] \begin{bmatrix} q_x & q_y & q_z & 0 \\ r_x & r_y & r_z & 0 \\ s_x & s_y & s_z & 0 \\ C_x & C_y & C_z & 1 \end{bmatrix} = [C_x, C_y, C_z, 1]$$

Notice also that the zero in the last ordinate ensures that direction vectors will not be affected by the translation, whereas all position vectors will be moved by the same factor. Notice also that if we do not shear the object the three vectors $\mathbf{q}$ $\mathbf{r}$ and $\mathbf{s}$ will remain orthogonal so that $\mathbf{q} \cdot \mathbf{r} = \mathbf{r} \cdot \mathbf{s} = \mathbf{q} \cdot \mathbf{s} = 0$. Unfortunately however, this analysis does not help us to determine the transformation matrix, though we will need it later in the course when we look at annimation. In general it would be more natural to assume that we know the vectors $\boldsymbol{u}$,$\boldsymbol{v}$, and $\boldsymbol{w}$ which form the view centered coordinate system. To see how to write down a transformation matrix from thses vectors we need to introduce the notion of the dot product as a projection onto a line. This is most readily seen in two dimensions as shown in Figure 4. By dropping perpendiculars from the point $\mathbf{P}$ to the line defined by vector $\boldsymbol{u}$ and vector $\boldsymbol{v}$ we see that the distances from the origin are respectively $\mathbf{P} \cdot \boldsymbol{u}$ and $\mathbf{P} \cdot \boldsymbol{v}$.

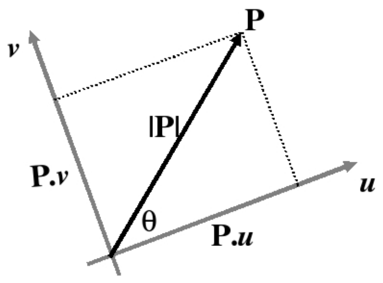


Figure 4: The dot product as projection

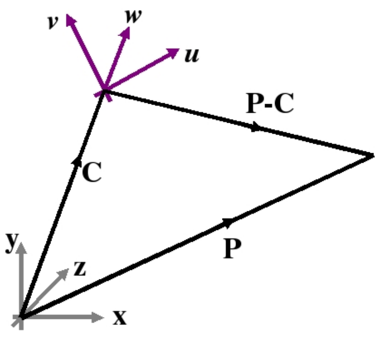


Figure 5: Three dimensional point projection

$\mathbf{P} \cdot \boldsymbol{u} = |\mathbf{P}||\boldsymbol{u}|Cos\theta = |\mathbf{P}|Cos\theta$ since $\boldsymbol{u}$ is a unit vector. Now, suppose that we wish to rotate the scene so that the new x and y axes were the $\boldsymbol{u}$ and $\boldsymbol{v}$ vectors, then the x ordinate would be defined by $\mathbf{P} \cdot \boldsymbol{u}$ and the y by $\mathbf{P} \cdot \boldsymbol{v}$.

The generalisation to three dimensions is shown in figure 5, in which a point $\mathbf{P}$ is transformed into the (u, v, w) axis system, translated by vector $\mathbf{C}$, is given by:

$$P^t_x = (\mathbf{P} - \mathbf{C}) \cdot \boldsymbol{u}$$
$$P^t_y = (\mathbf{P} - \mathbf{C}) \cdot \boldsymbol{v}$$
$$P^t_z = (\mathbf{P} - \mathbf{C}) \cdot \boldsymbol{w}$$

Expressing this as a transformation matrix we get:

$$[P^t_x, P^t_y, P^t_z, 1] = [P_x, P_y, P_z, 1] \begin{bmatrix} u_x & v_x & w_x & 0 \\ u_y & v_y & w_y & 0 \\ u_z & v_z & w_z & 0 \\ -\mathbf{C}\cdot\boldsymbol{u} & -\mathbf{C}\cdot\boldsymbol{v} & -\mathbf{C}\cdot\boldsymbol{w} & 1 \end{bmatrix}$$

So now by maintaining values of $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ and $\boldsymbol{w}$ throughout an animation sequence, for example by adjusting their values in response to mouse or joystick commands, we can simply write down the correct scene transformation matrix for viewing the virtual world from the correct position.

Finally we will derive the matrix used for a view centered transformation with the viewer at point $\mathbf{C}$ and looking in direction $\mathbf{d}$ by finding the new axis system $\mathbf{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$, and then deducing the transformation matrix. The direction vector for viewing is $\mathbf{d} = [d_x, d_y, d_z]$ and this must lie along the $\boldsymbol{w}$ direction, so $\boldsymbol{w} = \mathbf{d}/|\mathbf{d}|$. We first find any two vectors, $\mathbf{p}$ and $\mathbf{q}$,which have the same directions as $\mathbf{u}$ and $\mathbf{v}$, but do not necessarily have unit length. We can constrain the system to preserve horizontals and verticals in two ways: (i) Vector $\mathbf{p}$ will be in the direction of the new x-axis, so to preserve the horizontal, we choose: $p_y = 0$, and (ii) Vector $\mathbf{q}$ will be the new y axis (vertical) and should have a positive y component (so that the picture is not upside down) so we choose (arbitrarily): $q_y = 1$. Substituting our constraints in the Cartesian components we get:

$\mathbf{p} = [p_x, 0, p_z]$ and $\mathbf{q} = [q_x, 1, q_z]$

Now we know that, in a left hand axis system:

$\boldsymbol{k} = \boldsymbol{i} \times \boldsymbol{j}$ so therefore we can write: $\mathbf{d} = \mathbf{p} \times \mathbf{q}$

since we have not as yet constrained the magnitude of $\mathbf{p}$. Evaluating the cross product we get three Cartesian equations:

$d_x = p_y q_z - p_z q_y \qquad d_y = p_z q_x - p_x q_z \qquad d_z = p_x q_y - p_y q_x$

Substituting the values $p_y = 0$ and $q_y = 1$ these equations simplify to

$d_x = -p_z \qquad d_y = p_z q_x - p_x q_z \qquad d_z = p_x$

Thus we have solved for vector $\mathbf{p} = [d_z, 0, -d_x]$. To solve for $\mathbf{q}$ we need to express the condition that $\mathbf{p}$ and $\mathbf{q}$ are orthogonal and so have zero dot product, whence:

$\mathbf{p} \cdot \mathbf{q} = d_z q_x + 0 - d_x q_z = 0$ thus $q_z = d_z q_x / d_x$

and from the cross product already evaluated:

$d_y = p_z q_x - p_x q_z = -d_x q_x - d_z q_z = -d_x q_x - d_z^2 q_x / d_x$

so $q_x = -d_y d_x / (d_x^2 + d_z^2)$

and $\mathbf{q} = (-d_y d_x/(d_x^2 + d_z^2), 1, d_y d_z/(d_x^2 + d_z^2))$

finally we have that: $\boldsymbol{u} = \mathbf{p}/|\mathbf{p}|$ and $\boldsymbol{v} = \mathbf{q}/|\mathbf{q}|$

Thus we can deduce the whole of the transformation matrix.