# Perspective Mappings

David Eberly, Geometric Tools, Redmond WA 98052
https://www.geometrictools.com/



Created: August 16, 2011
Last Modified: April 1, 2019

## Contents

# 1 Introduction

This document describes various algorithms for perspective projections.

Section 1 shows how to create the perspective mapping between two convex quadrilaterals in the plane. The idea is to transform the convex quadrilaterals to canonical quadrilaterals using affine transformations and then to use a fractional linear transformation to map one canonical quadrilateral to the other canonical quadrilateral.

Section 2 provides a 3D geometric motivation for the perspective projection of one convex quadrilateral onto another. The idea is to place one quadrilateral in the $xy$-plane (the plane $z = 0$). One of the vertices is translated to the origin. The other quadrilateral has a vertex translated to the origin so that the translated quadrilateral lives on a plane containing the origin. This plane and an eyepoint for the perspective projection of the rotated quadrilateral onto the other is constructed, which leads to a 2D-to-2D mapping fractional linear transformation between the quadrilaterals. The construction can be used to map the first quadrilateral to a square. The square can then be mapped to the second quadrilateral. The composition of the two transformations is a perspective mapping from the first convex quadrilateral to the second.

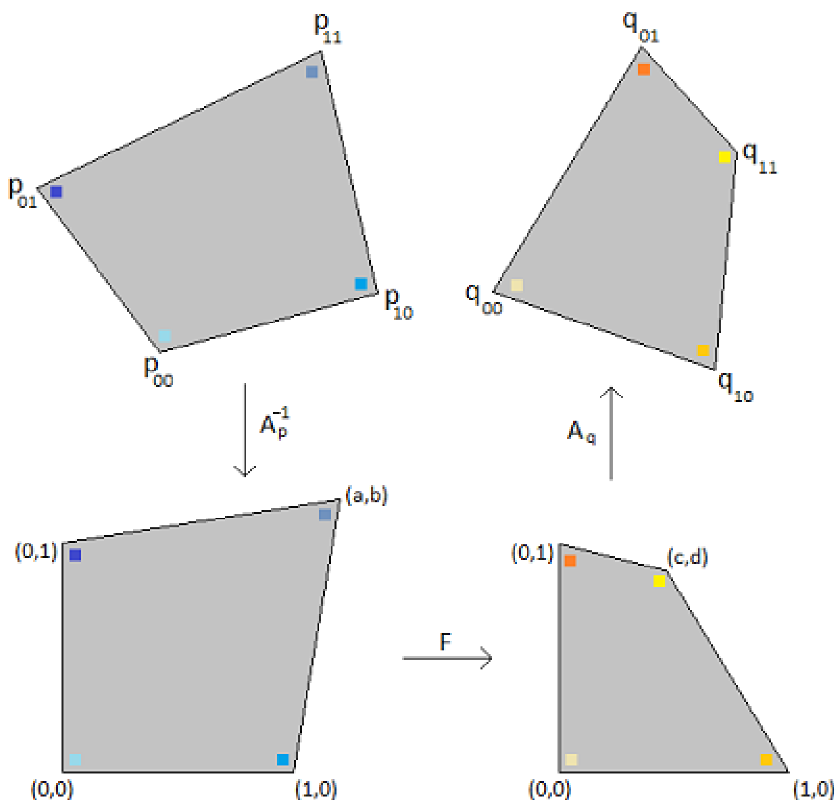
In 3 dimensions, a cuboid is a convex polyhedron with 6 quadrilateral faces. Section 3 describes a 4D geometric algorithm for the perspective projection between two cuboids. The standard application to computer graphics is the mapping of a view frustum to a canonical view volume with points $(x, y, z)$ subject to the constraints $|x| \leq 1$ and $|y| \leq 1$. The $z$-constraint is $z \in [0, 1]$ for DirectX or $z \in [-1, 1]$ for OpenGL. The perspective mapping is actually more general than that for mapping view frustums to canonical view volumes. It also supports perspective projection associated with parallax, where the viewport on the view plane is a convex quadrilateral that is not a rectangle. The standard projection matrices used in DirectX and OpenGL only allow for rectangular viewports.

Section 4 describes the extension to higher dimensions of the perspective mapping between cuboids to perspective mapping between hypercuboids. A hypercuboid is a bounded convex solid in $n$ dimensions whose faces are hypercuboids in $n - 1$ dimensions. The definition is recursive in dimension. For example, a 3D cuboid has quadrilateral faces where each quadrilateral is a 2D cuboid. The 2D cuboid consists of 1D cuboid "faces", which are line segments. As in Section 2, it is sufficient to map the first hypercuboid to a hypercube. The hypercube can then be mapped to the second hypercuboid. The composition of the two transformations is a perspective mapping from the first hypercuboid to the second.

# 2 Perspective Mapping Between Two Convex Quadrilaterals

The first convex quadrilateral has vertices $\mathbf{p}_{00}$, $\mathbf{p}_{10}$, $\mathbf{p}_{11}$ and $\mathbf{p}_{01}$, listed in counterclockwise order. The second quadrilateral has vertices $\mathbf{q}_{00}$, $\mathbf{q}_{10}$, $\mathbf{q}_{11}$ and $\mathbf{q}_{01}$, listed in counterclockwise order. The construction of the perspective mapping uses $3 \times 3$ homogeneous matrices and $3 \times 1$ homogeneous coordinates. Figure 1 shows the two quadrilaterals.

**Figure 1.** Two convex quadrilaterals. The left quadrilateral is to be mapped perspectively to the right quadrilateral. The transforms $A_p$ and $A_q$ are affine and $F$ is a fractional linear transformation.




Affinely transform the source quadrilateral so that $\mathbf{p}_{00}$ is mapped to the origin $(0,0)$, $\mathbf{p}_{10} - \mathbf{p}_{00}$ is mapped to $(1,0)$ and $\mathbf{p}_{01} - \mathbf{p}_{00}$ is mapped to $(0,1)$. This is equivalent to computing $\mathbf{x} = (x_0, x_1)$ for which a point $\mathbf{p}$ in the quadrilateral is represented by

$$\mathbf{p} = \mathbf{p}_{00} + x_0(\mathbf{p}_{10} - \mathbf{p}_{00}) + x_1(\mathbf{p}_{01} - \mathbf{p}_{00}) \tag{1}$$

The representation of $\mathbf{p}_{11}$ leads to $(x_0, x_1) = (a, b)$, which is shown in the lower left quadrilateral of Figure 1. In homogeneous coordinates we have

$$\left[\begin{array}{c}\mathbf{p}\\ \hline 1\end{array}\right] = \left[\begin{array}{cc|c}\mathbf{p}_{10} - \mathbf{p}_{00} & \mathbf{p}_{01} - \mathbf{p}_{00} & \mathbf{p}_{00}\\ \hline 0 & 0 & 1\end{array}\right] \left[\begin{array}{c}\mathbf{x}\\ \hline 1\end{array}\right] = \left[\begin{array}{c|c}M_p & \mathbf{p}_{00}\\ \hline \mathbf{0}^{\mathsf{T}} & 1\end{array}\right] \left[\begin{array}{c}\mathbf{x}\\ \hline 1\end{array}\right] = A_p \left[\begin{array}{c}\mathbf{x}\\ \hline 1\end{array}\right] \tag{2}$$

where $M_p$ is a $2 \times 2$ matrix whose columns are $\mathbf{p}_{10} - \mathbf{p}_{00}$ and $\mathbf{p}_{01} - \mathbf{p}_{00}$ and where $\mathbf{0}^{\mathsf{T}}$ is the $1 \times 2$ vector of zeros. The equation defines the $3 \times 3$ matrix $A_p$. The inverse of the affine transformation maps $\mathbf{p}$ to $\mathbf{x}$,

$$\left[\begin{array}{c}\mathbf{x}\\ \hline 1\end{array}\right] = A_p^{-1} \left[\begin{array}{c}\mathbf{p}\\ \hline 1\end{array}\right] = \left[\begin{array}{c|c}M_p^{-1} & -M_p^{-1}\mathbf{p}_{00}\\ \hline \mathbf{0}^{\mathsf{T}} & 1\end{array}\right] \left[\begin{array}{c}\mathbf{p}\\ \hline 1\end{array}\right] = \left[\begin{array}{c}M_p^{-1}(\mathbf{p} - \mathbf{p}_{00})\\ \hline 1\end{array}\right] \tag{3}$$

Similarly, we can write points $\mathbf{q}$ in the target quadrilateral as

$$\mathbf{q} = \mathbf{q}_{00} + y_0(\mathbf{q}_{10} - \mathbf{q}_{00}) + y_1(\mathbf{q}_{01} - \mathbf{q}_{00}) \tag{4}$$

Define $\mathbf{y} = (y_0, y_1)$. The representation of $\mathbf{q}_{11}$ leads to $(y_0, y_1) = (c, d)$, which is shown in the lower right quadrilateral of Figure 1. In homogeneous coordinates,

$$\left[\begin{array}{c} \mathbf{q} \\ \hline 1 \end{array}\right] = \left[\begin{array}{cc|c} \mathbf{q}_{10} - \mathbf{q}_{00} & \mathbf{q}_{01} - \mathbf{q}_{00} & \mathbf{q}_{00} \\ \hline 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} \mathbf{y} \\ \hline 1 \end{array}\right] = \left[\begin{array}{c|c} M_q & \mathbf{q}_{00} \\ \hline \mathbf{0}^\mathsf{T} & 1 \end{array}\right] \left[\begin{array}{c} \mathbf{y} \\ \hline 1 \end{array}\right] = A_q \left[\begin{array}{c} \mathbf{y} \\ \hline 1 \end{array}\right] \tag{5}$$

where $M_q$ is a $2 \times 2$ matrix whose columns are $\mathbf{q}_{10} - \mathbf{q}_{00}$ and $\mathbf{q}_{01} - \mathbf{q}_{00}$. The equation defines the $3 \times 3$ matrix $A_q$.

In Figure 1, the two quadrilaterals at the bottom of the figure are called *canonical quadrilaterals*. The first is mapped perspectively to the second using a fractional linear transformation,

$$(y_0, y_1) = f(x_0, x_1) = \frac{(f_{00}x_0 + f_{01}x_1 + f_{02}, f_{10}x_0 + f_{11}x_1 + f_{12})}{f_{20}x_0 + f_{21}x_1 + 1} \tag{6}$$

We must determine the 8 unknown coefficients $f_{ij}$ from the 4 pairs of related 2-tuple vertices, which leads to a system of 8 linear equations in 8 unknowns.

The vertex $\mathbf{x} = (0, 0)$ maps to $\mathbf{y} = (0, 0)$, which leads to $f_{02} = 0$ and $f_{12} = 0$. The vertex $(1, 0)$ maps to $(1, 0)$, which leads to $f_{00}/(f_{20} + 1) = 1$ and $f_{10}/(f_{20} + 1) = 0$; therefore, $f_{00} = f_{20} + 1$ and $f_{10} = 0$. The vertex $(0, 1)$ maps to $(0, 1)$, which leads to $f_{01}/(f_{21} + 1) = 0$ and $f_{11}/(f_{21} + 1) = 1$; therefore, $f_{01} = 0$ and $f_{11} = f_{21}+1$. Finally, vertex $(a, b)$ maps to $(c, d)$, which leads to the linear equations $f_{00}a = c(f_{20}a+f_{21}b+1)$ and $f_{11}b = d(f_{20}a+f_{21}b+1)$. We have seen that $f_{02}$, $f_{12}$, $f_{10}$ and $f_{01}$ are zero. The remaining four equations have solution

$$f_{00} = \frac{c(a + b - 1)}{a(c + d - 1)}, \quad f_{11} = \frac{d(a + b - 1)}{b(c + d - 1)}, \quad f_{20} = \frac{a - c + bc - ad}{a(c + d - 1)}, \quad f_{21} = \frac{b - d - bc + ad}{b(c + d - 1)} \tag{7}$$

We can substitute these in equation (6), multiply numerator and denominator by $ab(c + d - 1)$, and factor $(a - c + bc - ad) = c(a + b - 1) - a(c + d - 1)$ and $(b - d - bc + ad) = d(a + b - 1) - b(c + d - 1)$ to obtain

$$(y_0, y_1) = \frac{(bc(a + b - 1)x_0, ad(a + b - 1)x_1)}{b(c(a + b - 1) - a(c + d - 1))x_0 + a(d(a + b - 1) - b(c + d - 1))x_1 + ab(c + d - 1)} \tag{8}$$

The factoring identifies subexpressions that can be computed once in a computer program, namely, $a$, $b$, $c$, $d$, $a + b - 1$ and $c + d - 1$.

The convexity of the $\mathbf{p}$-quadrilateral is characterized by $a + b - 1 > 0$ and implies that both numerator coefficients $bc(a+b-1)$ and $ad(a+b-1)$ are positive. The denominator is a function of the form $k_0x_0+k_1x_1+k_2$ where $(x_0, x_1$ is in the canonical quadrilateral shown in Figure 1. The containment constraints are $x_0 \geq 0$, $x_1 \geq 0$, $(1 - b)x_0 + a(x_1 - 1) \leq 0$ and $b(x_0 - 1) + (1 - a)x_1 \leq 0$. The minimum of the denominator subject to the linear inequality constraints is solved as a linear programming problem that is easily solved. The minimum must occur at a vertex of the domain. The denominator at $(0, 0)$ is $ab(c+d-1)$, which is positive because the convexity of the $\mathbf{q}$-quadrilateral is characterized by $c + d - 1 > 0$. The denominator at $(1, 0)$ is $bc(a + b - 1)$, which is positive. The denominator at $(0, 1)$ is $ad(a + b - 1)$, which is positive. Finally, the denominator at $(a, b)$ is $ab(a + b - 1)$, which is positive. The minimum of the denominator is positive, which means equation (8) has no singularities in its domain.

The fractional linear transformation from the **q**-quadrilateral to the **p**-quadrilateral is the inverse of the function in equation (8),

$$(x_0, x_1) = \frac{(da(c+d-1)y_0, cb(c+d-1)y_1)}{d(a(c+d-1)-c(a+b-1))y_0 + c(b(c+d-1)-d(a+b-1))y_1 + cd(a+b-1)} \tag{9}$$

This may be constructed by simply swapping $(x_0, x_1)$ and $(y_0, y_1)$ and by swapping $(a, b)$ and $(c, d)$ in equation (8).

The perspective mapping from the **p**-quadrilateral to the **q**-quadrilateral is provided by a $3\times3$ homogeneous matrix and a perspective divide. Define $s = a + b - 1$ and $t = c + d - 1$. The homogeneous matrix is

$$F = \left[\begin{array}{cc|c} bcs & 0 & 0 \\ 0 & ads & 0 \\ \hline b(cs-at) & a(ds-bt) & abt \end{array}\right] = \left[\begin{array}{c|c} D & \mathbf{0} \\ \hline \boldsymbol{\ell}^\mathsf{T} & \lambda \end{array}\right] \tag{10}$$

where $D$ is a $2\times2$ diagonal matrix, $\mathbf{0}$ is the $2\times1$ vector of zeros, $\boldsymbol{\ell}^\mathsf{T}$ is a $1\times2$ vector and $\lambda$ is a scalar. The perspective mapping is

$$\left[\begin{array}{c} \mathbf{q} \\ \hline 1 \end{array}\right] \sim \left[\begin{array}{c} \mathbf{q}' \\ \hline w \end{array}\right] = A_q F A_p^{-1} \left[\begin{array}{c} \mathbf{q} \\ \hline 1 \end{array}\right] \tag{11}$$

where the similarity symbol indicates that the left-hand side is obtained via the perspective divide: $\mathbf{q} = \mathbf{q}'/w$. The $3\times3$ homography matrix $H = A_q F A_p^{-1}$ is

$$\begin{aligned} H &= \left[\begin{array}{c|c} M_q & \mathbf{q}_{00} \\ \hline \mathbf{0}^\mathsf{T} & 1 \end{array}\right] \left[\begin{array}{c|c} D & \mathbf{0} \\ \hline \boldsymbol{\ell}^\mathsf{T} & \lambda \end{array}\right] \left[\begin{array}{c|c} M_p^{-1} & -M_p^{-1}\mathbf{p}_{00} \\ \hline \mathbf{0}^\mathsf{T} & 1 \end{array}\right] \\ &= \left[\begin{array}{c|c} (M_q D + \mathbf{q}_{00}\boldsymbol{\ell}^\mathsf{T})M_p^{-1} & -(M_q D + \mathbf{q}_{00}\boldsymbol{\ell}^\mathsf{T})M_p^{-1}\mathbf{p}_{00} + \lambda\mathbf{q}_{00} \\ \hline \boldsymbol{\ell}^\mathsf{T} M_p^{-1} & -\boldsymbol{\ell}^\mathsf{T} M_p^{-1}\mathbf{p}_{00} + \lambda \end{array}\right] \end{aligned} \tag{12}$$

Some algebra will show that

$$H \left[\begin{array}{c} \mathbf{p}_{00} \\ \hline 1 \end{array}\right] = \left[\begin{array}{c} abt\,\mathbf{q}_{00} \\ \hline abt \end{array}\right],\ H \left[\begin{array}{c} \mathbf{p}_{10} \\ \hline 1 \end{array}\right] = \left[\begin{array}{c} bcs\,\mathbf{q}_{10} \\ \hline bcs \end{array}\right],\ H \left[\begin{array}{c} \mathbf{p}_{01} \\ \hline 1 \end{array}\right] = \left[\begin{array}{c} ads\,\mathbf{q}_{01} \\ \hline ads \end{array}\right],\ H \left[\begin{array}{c} \mathbf{p}_{11} \\ \hline 1 \end{array}\right] = \left[\begin{array}{c} abs\,\mathbf{q}_{00} \\ \hline abs \end{array}\right] \tag{13}$$

The perspective division leads to the correct mapping of vertices of the **p**-quadrilateral to the **q**-quadrilateral. Generally, the mapping of **p** to **q** is

$$\mathbf{q} = \frac{(M_q D + \mathbf{q}_{00}\boldsymbol{\ell}^\mathsf{T})M_p^{-1}(\mathbf{p} - \mathbf{p}_{00}) + \lambda\mathbf{q}_{00}}{\boldsymbol{\ell}^\mathsf{T} M_p^{-1}(\mathbf{p} - \mathbf{p}_{00}) + \lambda} \tag{14}$$

## 3 Perspective Mapping from a Convex Quadrilateral to a Square

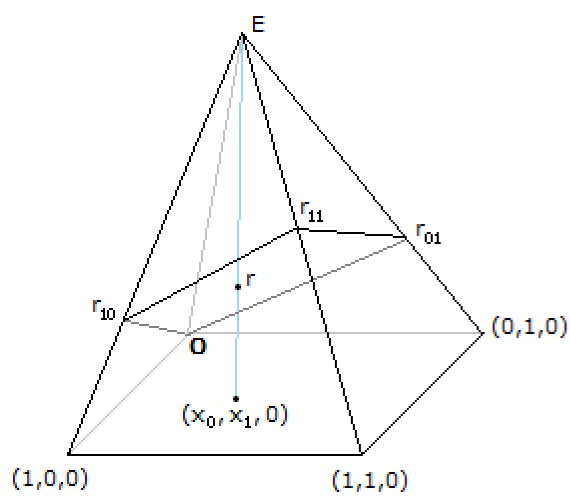
Let the square vertices be $(0, 0)$, $(1, 0)$, $(1, 1)$, and $(0, 1)$. The corresponding quadrilateral vertices are $\mathbf{q}_{00}$, $\mathbf{q}_{10}$, $\mathbf{q}_{11}$, and $\mathbf{q}_{01}$. To map the quadrilateral to the square, consider the problem as a perspective projection in three dimensions. The view plane is $y_2 = 0$ and the eyepoint is $\mathbf{E} = (e_0, e_1, e_2)$ where $e_2 \neq 0$ (the eyepoint is not on the view plane). The perspective projection of a point $\mathbf{r} = (r_0, r_1, r_2)$ to the view plane is the point $\mathbf{p} = (x_0, x_1, 0)$, which lies on the ray originating at $\mathbf{E}$ and contains $\mathbf{r}$. The intersection of the ray with the plane is the projection point: $\mathbf{p} = \mathbf{E} + t(\mathbf{r} - \mathbf{E})$ for some $t$. Some algebra leads to $t = e_2/(e_2 - r_2)$, $x_0 = (e_2 r_0 - e_0 r_2)/(e_2 - r_2)$, and $x_1 = (e_2 r_1 - e_1 r_2)/(e_2 - r_2)$.

Embed the square vertices in the view plane as $(i_0, i_1, 0)$ for $i_0$ and $i_1$ in $\{0, 1\}$. Embed the quadrilateral vertices as $(\mathbf{q}_{i_0 i_1} - \mathbf{q}_{00}, 0)$, which translates one of the vertices to the origin. Rotate the embedded quadrilateral vertices so that they lie in a plane containing the origin $\mathbf{0}$ and having normal vector $\mathbf{N} = (n_0, n_1, n_2)$. Let the rotated points be denoted $\mathbf{r}_{i_0 i_1} = R(\mathbf{q}_{i_0 i_1} - \mathbf{q}_{00}, 0)$, where $R$ is a rotation matrix corresponding to the normal; thus, $\mathbf{N} \cdot \mathbf{r}_{i_0 i_1} = 0$. The problem is to construct $\mathbf{E}$ and $\mathbf{N}$ so that

$$\mathbf{E} + t_{i_0 i_1} (\mathbf{r}_{i_0 i_1} - \mathbf{E}) = (i_0, i_1, 0), \quad i_j \in \{0, 1\} \tag{15}$$

for some scalars $t_{i_0 i_1}$. Equation (15) says that the embedded quadrilateral vertices are projected perspectively to the square vertices. Figure 2 shows the geometric configuration.

**Figure 2.** Projection of a quadrilateral onto a square. The rotated quadrilateral points $\mathbf{r}_{i_0 i_1}$ are projected to the square points $(i, j, 0)$. The figure shows a rotated quadrilateral interior point projected to a square interior point.




The vector $\mathbf{q}_{11} - \mathbf{q}_{00}$ can be represented as a linear combination of two quadrilateral edges, namely,

$$\mathbf{q}_{11} - \mathbf{q}_{00} = a_0 (\mathbf{q}_{10} - \mathbf{q}_{00}) + a_1 (\mathbf{q}_{01} - \mathbf{q}_{00}) \tag{16}$$

The coefficients $a_0$ and $a_1$ are constructed as the solution of two linear equations in two unknowns. These coefficients show up in the equations for the perspective projection. The convexity of the quadrilateral guarantees that $a_0 \geq 0$, $a_1 \geq 0$, and $a_0 + a_1 > 1$. Rotation preserves the relative positions of the vertices, so we know that

$$\mathbf{r}_{11} = a_0 \mathbf{r}_{10} + a_1 \mathbf{r}_{01} \tag{17}$$

Equation (15) contains four relationships,

$$\begin{aligned}
\mathbf{E} + t_{00}\,(\mathbf{r}_{00} - \mathbf{E}) &= (0,0,0) \\
\mathbf{E} + t_{10}\,(\mathbf{r}_{10} - \mathbf{E}) &= (1,0,0) \\
\mathbf{E} + t_{01}\,(\mathbf{r}_{01} - \mathbf{E}) &= (0,1,0) \\
\mathbf{E} + t_{11}\,(\mathbf{r}_{11} - \mathbf{E}) &= (1,1,0)
\end{aligned} \tag{18}$$

It must be that $t_{00} = 1$, because $\mathbf{r}_{00} = \mathbf{0}$. Dotting the last three equations with the normal vector leads to

$$(1 - t_{10})\mathbf{N} \cdot \mathbf{E} = n_0, \;\; (1 - t_{01})\mathbf{N} \cdot \mathbf{E} = n_1, \;\; (1 - t_{11})\mathbf{N} \cdot \mathbf{E} = n_0 + n_1 \tag{19}$$

Geometrically, it is clear that $\mathbf{N} \cdot \mathbf{E} \neq 0$, so these three equations imply

$$t_{11} = t_{10} + t_{01} - 1 \tag{20}$$

Substituting equations (17) and (20) into the last line of equation (18) produces

$$(2 - t_{10} - t_{01})\mathbf{E} + (t_{10} + t_{01} - 1)(a_0\mathbf{r}_{10} + a_1\mathbf{r}_{01}) = (1,1,0) \tag{21}$$

Adding the second and third lines of equation (18),

$$(2 - t_{10} - t_{01})\mathbf{E} + t_{10}\mathbf{r}_{10} + t_{01}\mathbf{r}_{01} = (1,1,0) \tag{22}$$

Subtracting equation (22) from (21), we have

$$[a_0(t_{10} + t_{01} - 1) - t_{10}]\mathbf{r}_{10} + [a_1(t_{10} + t_{01} - 1) - t_{01}]\mathbf{r}_{01} = \mathbf{0} \tag{23}$$

The linear independence of $\mathbf{r}_{10}$ and $\mathbf{r}_{01}$ guarantees that the coefficients in equation (23) must both be zero. We now have two linear equations in the two unknowns $t_{10}$ and $t_{01}$. These are easily solved, and together with equation (20) produce

$$t_{00} = 1, \;\; t_{10} = \frac{a_0}{a_0 + a_1 - 1}, \;\; t_{01} = \frac{a_1}{a_0 + a_1 - 1}, \;\; t_{11} = \frac{1}{a_0 + a_1 - 1} \tag{24}$$

It turns out that computing $t_{11}$ explicitly in terms of $a_0$ and $a_1$ is not necessary to construct the fractional linear transformations.

At this time we may attempt to construct $\mathbf{E}$ and $\mathbf{N}$, but this is not necessary to actually construct the perspective projection. Instead, consider a rotated quadrilateral point $\mathbf{r}$ that is projected to a square point $(x_0, x_1, 0)$, as illustrated in Figure 2. We may write $\mathbf{r}$ as a linear combination of two rotated quadrilateral edges,

$$\mathbf{r} = y_0\mathbf{r}_{10} + y_1\mathbf{r}_{01} \tag{25}$$

The ray equation for the pair of points is

$$(x_0, x_1, 0) = \mathbf{E} + t(\mathbf{r} - \mathbf{E}) = \mathbf{E} + t(y_0\mathbf{r}_{10} + y_1\mathbf{r}_{01} - \mathbf{E}) \tag{26}$$

for some scalar $t$. Substituting our now known values for $t_{i_0 i_1}$ into the second and third lines of equation (18), solving for $\mathbf{r}_{10}$ and $\mathbf{r}_{01}$, replacing in equation (26), and grouping like terms produces

$$\left(1 - \frac{ty_0(1 - t_{10})}{t_{10}} - \frac{ty_1(1 - t_{01})}{t_{01}} - t\right)\mathbf{E} + \left(\frac{ty_0}{t_{10}} - x_0\right)(1,0,0) + \left(\frac{ty_1}{t_{01}} - x_1\right)(0,1,0) = (0,0,0) \tag{27}$$

By linear independence of $\mathbf{E}$, $(1,0,0)$, and $(0,0,1)$, the coefficients in equation (27) must all be zero. This leads to the fractional linear transformation that maps the quadrilateral to the square

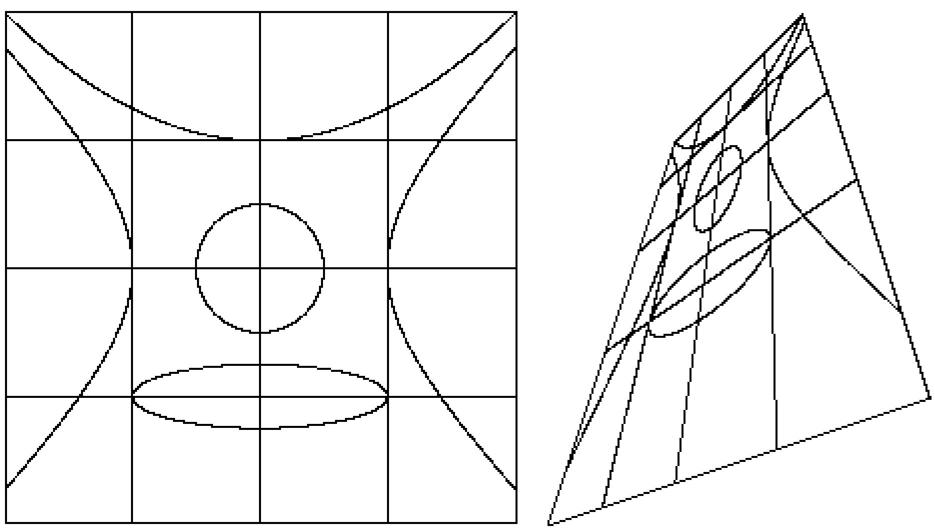
$$(x_0, x_1) = \frac{(a_1(a_0+a_1-1)y_0, a_0(a_0+a_1-1)y_1)}{a_0a_1 + a_1(a_1-1)y_0 + a_0(a_0-1)y_1} \tag{28}$$

The fractional linear transformation that maps the square to the quadrilateral is

$$(y_0, y_1) = \frac{(a_0x_0, a_1x_1)}{(a_0+a_1-1) + (1-a_1)x_0 + (1-a_0)x_1} \tag{29}$$

**Example**. Intuitively, perspective transformations map lines to lines. They also map conic sections to conic sections. The proof is straightforward. Let the perspective transformation be $y_0 = (a_0x_0+a_1x_1+a_2)/(c_0x_0+c_1x_1+c_2)$ and $y_1 = (b_0x_0+b_1x_1+b_2)/(c_0x_0+c_1x_1+c_2)$. The inverse transformation is of the same form, so it suffices to show a conic section in $(y_0, y_1)$ coordinates satisfying $Ay_0^2 + By_0y_1 + Cy_1^2 + Dy_0 + Ey_1 + F = 0$ is mapped to a conic section in $(x_0, x_1)$ coordinates satisfying $\bar{A}x_0^2 + \bar{B}x_0x_1 + \bar{C}x_1^2 + \bar{D}x_0 + \bar{E}x_1 + \bar{F} = 0$. Substituting the formulas for $y_0$ and $y_1$ into the quadratic equation, multiplying by $(c_0x_0 + c_1x_1 + c_2)^2$, expanding the products, and grouping the appropriate terms yields a quadratic in $x_0$ and $x_1$. The left image in Figure 3 shows a square containing several conic sections. The top curve is a parabola, the left and right curves are hyperbolas, the center curve is a circle, and the bottom curve is an ellipse. The grid consists of straight lines. The right image in Figure 3 shows a perspective mapping of these conic sections.

**Figure 3.** Projection of conic sections in a square to conic sections in a quadrilateral.



# 4 Perspective Mapping from a Cuboid to a Cube

The ideas of the previous section extend to 3D. A cuboid is a convex polyhedron that has 6 planar faces. For example, a typical view frustum is a cuboid. A cube is a special case of a cuboid, all six faces are squares, and opposite faces are parallel. Let the vertices of the cuboid be denoted $\mathbf{q}_{i_0i_1i_2}$ where $i_j \in \{0,1\}$. The canonical cube vertices are $(i_0, i_1, i_1)$.

To project perspectively the cuboid to the cube, we can embed the problem in 4D. The view hyperplane is $y_3 = 0$ and the eyepoint is $\mathbf{E} = (e_0, e_1, e_2, e_3)$ where $e_3 \neq 0$. The cuboid is translated to the origin and then rotated into a hyperplane with normal $\mathbf{N} = (n_0, n_1, n_2, n_3)$. The rotation matrix is $R$ and the rotated and translated vertices are $\mathbf{r}_{i_0 i_1 i_2} = R(\mathbf{q}_{i_0 i_1 i_2} - \mathbf{q}_{000}, 0)$. Our goal is that

$$\mathbf{E} + t_{i_0 i_1 i_2} (\mathbf{r}_{i_0 i_1 i_2} - \mathbf{E}) = (i_0, i_1, i_2 0), \quad i_j \in \{0, 1\} \tag{30}$$

for some scalars $t_{i_0 i_1 i_2}$.

The vector $\mathbf{q}_{111} - \mathbf{q}_{000}$ can be represented as a linear combination of three cuboid edges, namely,

$$\mathbf{q}_{111} - \mathbf{q}_{000} = a_0 (\mathbf{q}_{100} - \mathbf{q}_{000}) + a_1 (\mathbf{q}_{010} - \mathbf{q}_{000}) + a_2 (\mathbf{q}_{001} - \mathbf{q}_{000}) \tag{31}$$

The coefficients $a_0$, $a_1$, and $a_2$ are constructed as the solution of three linear equations in three unknowns. These coefficients show up in the equations for the perspective projection. The convexity of the cuboid guarantees that $a_0 \geq 0$, $a_1 \geq 0$, $a_2 \geq 0$, and $a_0 + a_1 + a_2 > 1$. Rotation preserves the relative positions of the vertices, so we know that

$$\mathbf{r}_{111} = a_0 \mathbf{r}_{100} + a_1 \mathbf{r}_{010} + a_2 \mathbf{r}_{001} \tag{32}$$

Dotting the relationships in equation (30) with the normal vector produces

$$\begin{aligned} (1 - t_{100})\mathbf{N} \cdot \mathbf{E} &= n_0, & (1 - t_{110})\mathbf{N} \cdot \mathbf{E} &= n_0 + n_1 \\ (1 - t_{010})\mathbf{N} \cdot \mathbf{E} &= n_1, & (1 - t_{101})\mathbf{N} \cdot \mathbf{E} &= n_0 + n_2 \\ (1 - t_{001})\mathbf{N} \cdot \mathbf{E} &= n_2, & (1 - t_{011})\mathbf{N} \cdot \mathbf{E} &= n_1 + n_2 \\ & & (1 - t_{111})\mathbf{N} \cdot \mathbf{E} &= n_0 + n_1 + n_2 \end{aligned} \tag{33}$$

Because $\mathbf{N} \cdot \mathbf{E} \neq 0$, we obtain

$$t_{110} = t_{100} + t_{010} - 1, \;\; t_{101} = t_{100} + t_{001} - 1, \;\; t_{011} = t_{010} + t_{001} - 1, \;\; t_{111} = t_{100} + t_{010} + t_{001} - 2 \tag{34}$$

Substituting the relationship for $t_{111}$ from equation (34) into the equation of (30) involving $t_{111}$ leads to

$$(3 - t_{100} - t_{010} - t_{001})\mathbf{E} + (t_{100} + t_{010} + t_{001} - 2)(a_0 \mathbf{r}_{100} + a_1 \mathbf{r}_{010} + a_2 \mathbf{r}_{001}) = (1, 1, 1) \tag{35}$$

Adding the relationships of equation (30) involving $t_{100}$, $t_{010}$, and $t_{001}$ produces

$$(3 - t_{100} - t_{010} - t_{001})\mathbf{E} + t_{100}\mathbf{r}_{100} + t_{010}\mathbf{r}_{010} + t_{001}\mathbf{r}_{001} = (1, 1, 1) \tag{36}$$

Subtracting equation (36) from (35), we have

$$\begin{aligned} &[a_0(t_{100} + t_{010} + t_{001} - 2) - t_{100}]\mathbf{r}_{100} + [a_1(t_{100} + t_{010} + t_{001} - 2) - t_{010}]\mathbf{r}_{010} \\ &\quad + [a_2(t_{100} + t_{010} + t_{001} - 2) - t_{001}]\mathbf{r}_{001} = \mathbf{0} \end{aligned} \tag{37}$$

The linear independence of $\mathbf{r}_{100}$, $\mathbf{r}_{010}$, and $\mathbf{r}_{001}$ guarantees that the coefficients in equation (37) must all be zero. We now have three linear equations in three unknowns $t_{100}$, $t_{010}$, and $t_{001}$. These are easily solved, and together with previous equations produce

$$\begin{aligned} t_{000} &= 1, & t_{110} &= \tfrac{+a_0 + a_1 - a_2 + 1}{a_0 + a_1 + a_2 - 1} \\ t_{100} &= \tfrac{2a_0}{a_0 + a_1 + a_2 - 1}, & t_{101} &= \tfrac{+a_0 - a_1 + a_2 + 1}{a_0 + a_1 + a_2 - 1} \\ t_{010} &= \tfrac{2a_1}{a_0 + a_1 + a_2 - 1}, & t_{011} &= \tfrac{-a_0 + a_1 + a_2 + 1}{a_0 + a_1 + a_2 - 1} \\ t_{001} &= \tfrac{2a_2}{a_0 + a_1 + a_2 - 1}, & t_{111} &= \tfrac{+a_0 + a_1 + a_2 + 1}{a_0 + a_1 + a_2 - 1} \end{aligned} \tag{38}$$

It turns out that computing $t_{110}$, $t_{101}$, $t_{011}$, and $t_{111}$ explicitly in terms of $a_0$, $a_1$, and $a_2$ is not necessary to construct the fractional linear transformations.

As in the 2D problem, we can construct the fractional linear transformations for the projection without explicitly constructing $\mathbf{E}$ and $\mathbf{N}$. We may write a rotated interior cuboid point $\mathbf{r}$ as

$$\mathbf{r} = y_0\mathbf{r}_{100} + y_1\mathbf{r}_{010} + y_2\mathbf{r}_{001} \tag{39}$$

The ray equation for the pair of points is

$$(x_0, x_1, x_2, 0) = \mathbf{E} + t(\mathbf{r} - \mathbf{E}) = \mathbf{E} + t(y_0\mathbf{r}_{100} + y_1\mathbf{r}_{010} + y_2\mathbf{r}_{001} - \mathbf{E}) \tag{40}$$

for some scalar $t$. Substituting the values for $t_{i_0i_1i_2}$ into equation (30); solving for $\mathbf{r}_{100}$, $\mathbf{r}_{010}$, and $\mathbf{r}_{001}$; replacing in equation (40), and grouping like terms produces

$$\begin{aligned}&\left(1 - \tfrac{ty_0(1-t_{100})}{t_{100}} - \tfrac{ty_1(1-t_{010})}{t_{010}} - \tfrac{ty_2(1-t_{001})}{t_{001}}\right)\mathbf{E} \\ &+ \left(\tfrac{ty_0}{t_{100}} - x_0\right)(1,0,0,0) + \left(\tfrac{ty_1}{t_{010}} - x_1\right)(0,1,0,0) + \left(\tfrac{ty_2}{t_{001}} - x_2\right)(0,0,1,0) = (0,0,0,0)\end{aligned} \tag{41}$$

By linear independence of $\mathbf{E}$, $(1,0,0,0)$, $(0,1,0,0)$, and $(0,0,1,0)$, the coefficients in equation (41) must all be zero. This leads to the fractional linear transformation that maps the cube to the cuboid,

$$(y_0, y_1, y_2) = \frac{2(a_0x_0, a_1x_1, a_2x_2)}{d_0x_0 + d_1x_1 + d_2x_2 + d_3} \tag{42}$$

where $d_0 = a_0 - a_1 - a_2 + 1$, $d_1 = -a_0 + a_1 - a_2 + 1$, $d_2 = -a_0 - a_1 + a_2 + 1$, and $d_3 = a_0 + a_1 + a_2 + 1$. The fractional linear transformation that maps the cube to the cuboid is

$$(x_0, x_1, x_2) = \frac{d_3(a_1a_2y_0, a_0a_2y_1, a_0a_1y_2)}{2a_0a_1a_2 - a_1a_2d_0y_0 - a_0a_2d_1y_1 - a_0a_1d_2y_2} \tag{43}$$

**Example**. When drawing a view volume in 3D graphics, one typically creates a view frustum with near plane, far plane, left plane, right plane, bottom plane, and top plane. The vertices of geometric primitives are transformed from model space to world space, then from world space to view space (camera space). The view-space vertices are finally transformed to clip space (homogeneous coordinates) by a $4 \times 4$ perspective projection matrix. The rasterizer does the clipping in homogeneous coordinates and then performs the perspective divide to obtain the normalized window coordinates $(x, y) \in [-1, 1]^2$ and normalized depth $z \in [0, 1]$ (for DirectX; OpenGL maps depth to $[-1, 1]$). The viewport is the full rectangular window, but it may be chosen to be an subwindow specified as an axis-aligned rectangle.

It is possible to specify a viewport that is a convex quadrilateral (nonrectangular). Let this viewport have vertices in camera coordinates, $\mathbf{q}_{i_0i_10}$ for $i_j \in \{0, 1\}$, and let them be counterclockwise oriented. This polygon acts as the near face of a cuboidal view volume. The quadrilateral may be extruded toward the far plane of the view frustum, thus constructing the far face of the cuboidal view volume. Let the camera forward direction be $\mathbf{D}$. Let $n$ be the near-plane distance from the eyepoint and let $f$ be the far-plane distance from the eyepoint. The viewport vertex $\mathbf{q}_{i_0i_10}$ is on the near plane, so $\mathbf{D} \cdot \mathbf{q}_{i_0i_1} = n$. The corresponding far-plane vertex is $\mathbf{q}_{i_0i_11} = (f/n)\mathbf{q}_{i_0i_1}$ so that $\mathbf{D} \cdot \mathbf{q}_{i_0i_11} = f$.

We may construct a projection matrix that maps the cuboidal view volume to the cube $[-1, 1]^2 \times [0, 1]$. Define $\mathbf{U}_0 = \mathbf{q}_{100} - \mathbf{q}_{000}$, $\mathbf{U}_1 = \mathbf{q}_{010} - \mathbf{q}_{000}$, and $\mathbf{U}_2 = \mathbf{q}_{001} - \mathbf{q}_{000}$. Define the matrix $M = [\mathbf{U}_0\ \mathbf{U}_1\ \mathbf{U}_2]$

whose columns are the specified vectors. Define $\mathbf{a}$ to be the column vector whose components are the $a_i$ in the fractional linear transformations. Then

$$\mathbf{a} = M^{-1}(\mathbf{q}_{111} - \mathbf{q}_{000}) \tag{44}$$

In homogeneous block form,

$$\begin{bmatrix} \mathbf{a} \\ 1 \end{bmatrix} = \begin{bmatrix} M & \mathbf{q}_{000} \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{q}_{111} \\ 1 \end{bmatrix} = \begin{bmatrix} M^{-1} & -M^{-1}\mathbf{q}_{000} \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{q}_{111} \\ 1 \end{bmatrix} = \begin{bmatrix} M^{-1}(\mathbf{q}_{111} - \mathbf{q}_{000}) \\ 1 \end{bmatrix} \tag{45}$$

Define $H$ to be the homogeneous matrix whose upper-left block is the matrix $M$ and whose translation component is the vector $\mathbf{q}_{000}$.

The projection of equation (43) maps the cuboidal volume to the canonical cube $[0,1]^3$. We need to perform one additional scaling and translation to map $(x_0, x_1)$ from $[0,1]^2$ to $[-1,1]^2$. With this additional transformation, the cuboid-to-cube mapping may be written in homogeneous form as

$$P = \begin{bmatrix} \frac{a_2}{a_0}(2d_3 + d_0) & \frac{a_2}{a_1} d_1 & d_2 & -2a_2 \\ \frac{a_2}{a_0} d_0 & \frac{a_2}{a_1}(2d_3 + d_1) & d_2 & -2a_2 \\ 0 & 0 & d_3 & 0 \\ -\frac{a_2}{a_0} d_0 & -\frac{a_2}{a_1} d_1 & -d_2 & 2a_2 \end{bmatrix} \tag{46}$$

where the output $(y_0, y_1, y_2, w)$ is processed by the rasterizer in the usual manner, performing the perspective divide by $w$ to obtain the cube points.

The projection matrix from camera space to clip space is therefore $PH^{-1}$. The matrix $H^{-1}$ maps camera coordinates to cuboid coordinates relative to three linearly independent edges of the cuboid. These points are then perspectively projected by $P$ into the cube.

# 5 Perspective Mapping from a Hypercuboid to a Hypercube

The ideas in the previous sections generalize to higher dimensions. Let the $d$-dimensional hypercuboid have vertices $\mathbf{q}_I$, where $I = (i_1, i_2, \ldots i_d)$ is a multiindex with components $i_j \in \{0, 1\}$. The canonical hypercube vertices are $\mathbf{I}$, with bold face used to distinguish between vectors and multiindices.

The view hyperplane is $y_d = 0$ and the eyepoint is $\mathbf{E} = (e_1, e_2, \ldots, e_d)$ with $e_d \neq 0$. The hypercuboid is translated to the origin and then rotated into a hyperplane with normal $\mathbf{N} = (n_1, n_2, \ldots, n_d)$. Let $R$ be the corresponding rotation matrix and define the rotated hypercuboid vertices by $\mathbf{r}_I = R(\mathbf{q}_I - \mathbf{q}_O)$, where $O$ is the multiindex of all zeros. Our goal is that

$$\mathbf{E} + t_I(\mathbf{r}_I - \mathbf{E}) = (\mathbf{I}, 0) \tag{47}$$

for some scalars $t_I$. Let $U$ be the multiindex of all ones. The vector $\mathbf{q}_U - \mathbf{q}_O$ may be written as

$$\mathbf{q}_U - \mathbf{q}_O = \sum_{j=1}^{d} a_j(\mathbf{q}_{B_j} - \mathbf{q}_O) \tag{48}$$

where $B_j$ is the multiindex with 0 in all components except for a 1 in component $j$. The coefficients are nonnegative and $\sum_{j=1}^{d} a_j > 0$. By linearity of the rotation matrix,

$$\mathbf{r}_U = \sum_{j=1}^{d} a_j \mathbf{r}_{B_j} \tag{49}$$

Let $I = (i_1, i_2, \ldots, i_d)$ be a multiindex with at least one nonzero component. Suppose that $I$ has $m$ components that are 1; it is the case that $1 \le m \le d$. Let those components be $i_{j_1}$ through $i_{j_m}$. Dotting the ray equation with the normal vector, we have

$$(1 - t_I)\mathbf{N} \cdot \mathbf{E} = \sum_{k=1}^{m} n_{i_k} \tag{50}$$

Whenever $m \ge 2$, the $t_I$ variable is related to $t_{B_j}$ variables by summations. Specifically, $I = \sum_{k=1}^{m} i_k B_{i_k}$ and

$$(1 - t_I)\mathbf{N} \cdot \mathbf{E} = \sum_{k=1}^{m} n_{i_k} = \sum_{k=1}^{m} (1 - t_{B_{i_k}})\mathbf{N} \cdot \mathbf{E} \tag{51}$$

Therefore, once we determine the relationships between $t_{B_i}$ and $a_j$, the $t_I$ with 2 or more 1-valued indices are determined. It turns out that for the construction of the fractional linear transformations, we only need to know $t_U$ explicitly in terms of the $a_j$,

$$t_U = \sum_{i=1}^{d} t_{B_i} - (d-1) \tag{52}$$

Substituting the relationship for $t_U$ into equation (47) leads to

$$\left(d - \sum_{i=1}^{d} t_{B_i}\right)\mathbf{E} + \left(\sum_{j=1}^{d} t_{B_j} - (d-1)\right)\left(\sum_{i=1}^{d} a_i \mathbf{r}_{B_i}\right) = \mathbf{U} \tag{53}$$

Adding the relationships of equation (47) involving the $t_{B_i}$ produces

$$\left(d - \sum_{i=1}^{d} t_{B_i}\right)\mathbf{E} + \sum_{i=1}^{d} t_{B_i}\mathbf{r}_{B_i} = \mathbf{U} \tag{54}$$

Subtracting the two equations, we have

$$\sum_{i=1}^{d}\left[a_i\left(\sum_{j=1}^{d} t_{B_j} - (d-1)\right) - t_{B_i}\right]\mathbf{r}_{B_i} = \mathbf{0} \tag{55}$$

The linear independence of the $\mathbf{r}_{B_i}$ guarantees that all the coefficients in this equation are zero. This leads to an invertible linear system of $d$ equations in the $d$ unknowns $t_{B_1}$ through $t_{B_d}$.

As in dimensions 2 and 3, we need not construct $\mathbf{E}$ and $\mathbf{N}$ explicitly. We may write a rotated interior hypercuboid point $\mathbf{r}$ as

$$\mathbf{r} = \sum_{i=1}^{d} y_i \mathbf{r}_{B_i} \tag{56}$$

The ray equation for the pair of points is

$$(x_1, \ldots, x_d, 0) = \mathbf{E} + t(\mathbf{r} - \mathbf{E}) = \mathbf{E} + t\left(\sum_{i=1}^{d} y_i \mathbf{r}_{B_i} - \mathbf{E}\right) \tag{57}$$

for some scalar $t$. Substituting the values for $t_{B_i}$ into equation (47), solving for $\mathbf{r}_{B_i}$, replacing in equation (57), and grouping like terms produces

$$\left(1 - \sum_{i=1}^{d} \frac{t y_i (1 - t_{B_i})}{t_{B_i}}\right) \mathbf{E} + \sum_{i=1}^{d} \left(\frac{t y_i}{t_{B_i}} - x_i\right) (\mathbf{B}_i, 0) = \mathbf{0} \tag{58}$$

By linear independence of $\mathbf{E}$ and the $(\mathbf{B}_i, 0)$, the coefficients in equation (58) must all be zero. This leads to the fractional linear transformation that maps the hypercube to the hypercuboid,

$$y_i = \frac{(d-1) a_i x_i}{\sum_{k=1}^{d} c_k x_k + c_{d+1}}, \quad 1 \le i \le d \tag{59}$$

where

$$c_k = \sum_{j=1}^{d} b_{kj} a_j + 1, \ 1 \le k \le d; \quad c_{d+1} = \sum_{j=1}^{d} a_j - 1; \quad b_{kj} = \begin{cases} d-2, & k = j \\ -1, & k \ne j \end{cases} \tag{60}$$

The fractional linear transformation that maps the hypercuboid to the hypercube is

$$x_i = \frac{\left[\prod_{j=1}^{d} a_j\right] c_{d+1} (y_i / a_i)}{\left[\prod_{j=1}^{d} a_j\right] \left((d-1) - \sum_{j=1}^{d} c_j (y_j / a_j)\right)}, \quad 1 \le i \le d \tag{61}$$

The products of the $a_j$ appear as if they cancel, but if any of the $a_j$ are zero, you need to formally multiply through by the product and then evaluate the expression using the $a_j$.