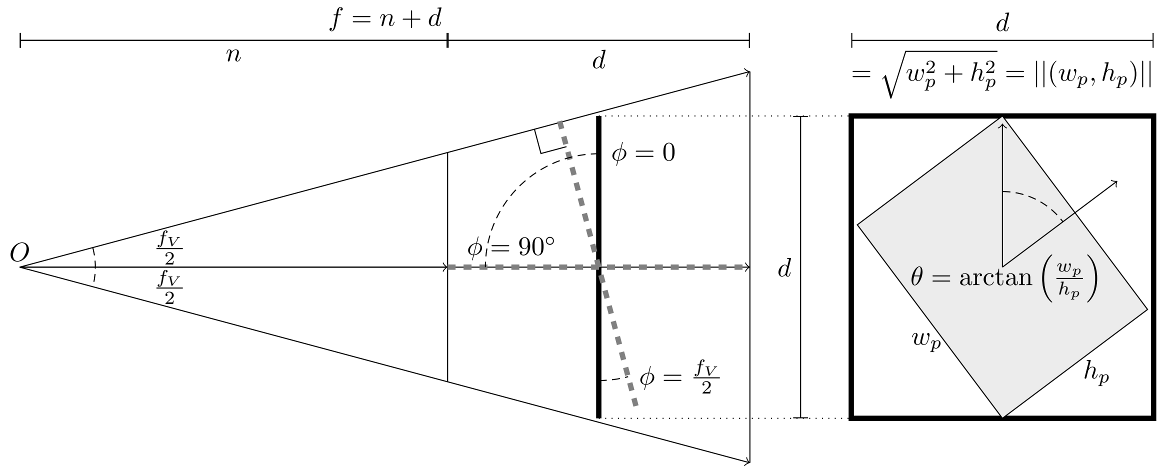
Derivations

We seek the general perspective transform $\mathcal{F}$ of a quadrilateral,

$$M = \mathcal{F}\left(w_p, h_p, \theta, \phi, f_V\right)$$

which produces a 4×4 warp matrix M by which to transform the vertices of an image of width w_p and height h_p pixels, given a camera axis rotation θ, out-of-plane (tilt) rotation ϕ and vertical field-of-view f_V.



The above figure illustrates the relationships between w_p, h_p, θ, ϕ and f_V for the particular configuration (Tilt equal to $\frac{f_V}{2}$, Rotation equal to $\arctan\left(\frac{w_p}{h_p}\right)$) which results in the biggest image to process (The extreme case).
Let us assign to each corner of the image a coordinate in the object coordinate system as follows, assuming the convention that the observer faces down the negative z-axis, with the positive y-axis pointing up and the positive x-axis pointing to the right:

Corner	**Object Coordinate System**
Top Left	$\left(-\frac{w_p}{2}, \frac{h_p}{2}, 0\right)$
Top Right	$\left(\frac{w_p}{2}, \frac{h_p}{2}, 0\right)$
Bottom Right	$\left(\frac{w_p}{2}, -\frac{h_p}{2}, 0\right)$
Bottom Left	$\left(-\frac{w_p}{2}, -\frac{h_p}{2}, 0\right)$

More specifically, we seek

$$\boxed{\mathcal{F} = P T R_\phi R_\theta}$$

where R_θ is the rotation matrix around the z-axis, R_ϕ is the rotation matrix around the x-axis, T is the translation matrix that displaces the coordinate system down the z-axis and P is the projection matrix for a square viewport with vertical field-of-view f_V.

R_θ and R_ϕ are trivial to find; they are

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$R_\phi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi & 0 \\ 0 & \sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

T is more difficult to find, but not excessively so. We must first define a variable d representing the side length of the square that would wholy contain any rotation of the image. This is plainly

$$d = \sqrt{w_p^2 + h_p^2}$$

Now, given f_V, we solve for the hypotenuse of a right triangle with an opposite side of length $\frac{d}{2}$ subtending an angle $\frac{f_V}{2}$. This measure we will denote h.

$$h = \frac{d}{2\sin\left(\frac{f_V}{2}\right)}$$

It is h that describes by how much to translate the object coordinate system down the z-axis. The matrix T is therefore defined as:

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -h \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{d}{2\sin\left(\frac{f_V}{2}\right)} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The last matrix, the perspective matrix P, relies on a few more assumptions. Let us define two more values, n and f, based on d and h computed above:

$$n = h - \frac{d}{2}$$

$$f = h + \frac{d}{2}$$

These represent the distances to the near plane and far plane, respectively. Given that the aspect ratio is 1 (we are rotating the square in which the image is embedded), we can now define a perspective matrix P as follows:

$$P = \begin{bmatrix} \frac{n}{n\tan\left(\frac{f_V}{2}\right)} & 0 & 0 & 0 \\ 0 & \frac{n}{n\tan\left(\frac{f_V}{2}\right)} & 0 & 0 \\ 0 & 0 & -\frac{(f+n)}{f-n} & -\frac{2fn}{f-n} \\ 0 & 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} \cot\left(\frac{f_V}{2}\right) & 0 & 0 & 0 \\ 0 & \cot\left(\frac{f_V}{2}\right) & 0 & 0 \\ 0 & 0 & -\frac{(f+n)}{f-n} & -\frac{2fn}{f-n} \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

One more thing needs to be elaborated upon: the size of the output image. A desirable property of the warp is that, for the special case $\theta = 0, \phi = 0$, it should leave the image unchanged (same size, no distortion), except for centering it. This can be accomplished if the viewport transform maps to a square image of side length l, where l is defined as:

$$l = d\sec\left(\frac{f_V}{2}\right)$$

This follows from the fact that the distance from the center point of the image to the point directly above it on the upper plane of the viewing frustrum is $h\tan\left(\frac{f_V}{2}\right)$, and the square side length is twice that distance since it also extends downwards to the lower plane. Hence,

$$l = 2h\tan\left(\frac{f_V}{2}\right) = 2\frac{d}{2\sin\left(\frac{f_V}{2}\right)}\tan\left(\frac{f_V}{2}\right) = d\frac{\tan\left(\frac{f_V}{2}\right)}{\sin\left(\frac{f_V}{2}\right)} = d\sec\left(\frac{f_V}{2}\right)$$