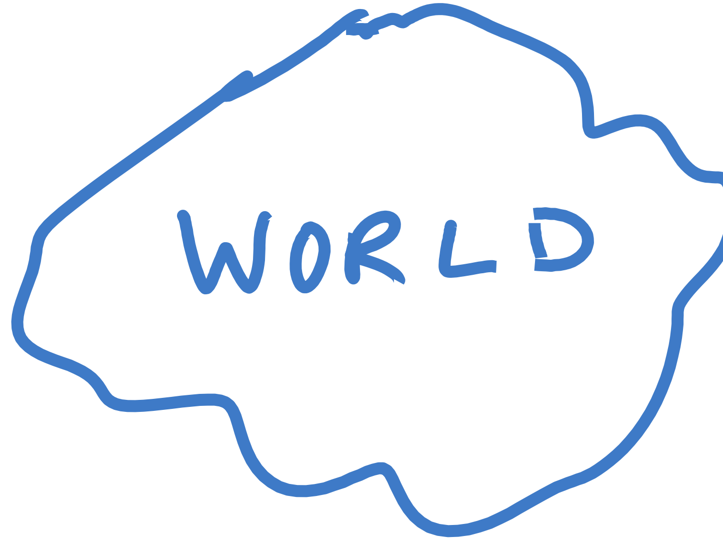
Fundamentals of Image Formation

Lecture 2

Jitendra Malik

A camera creates an image …

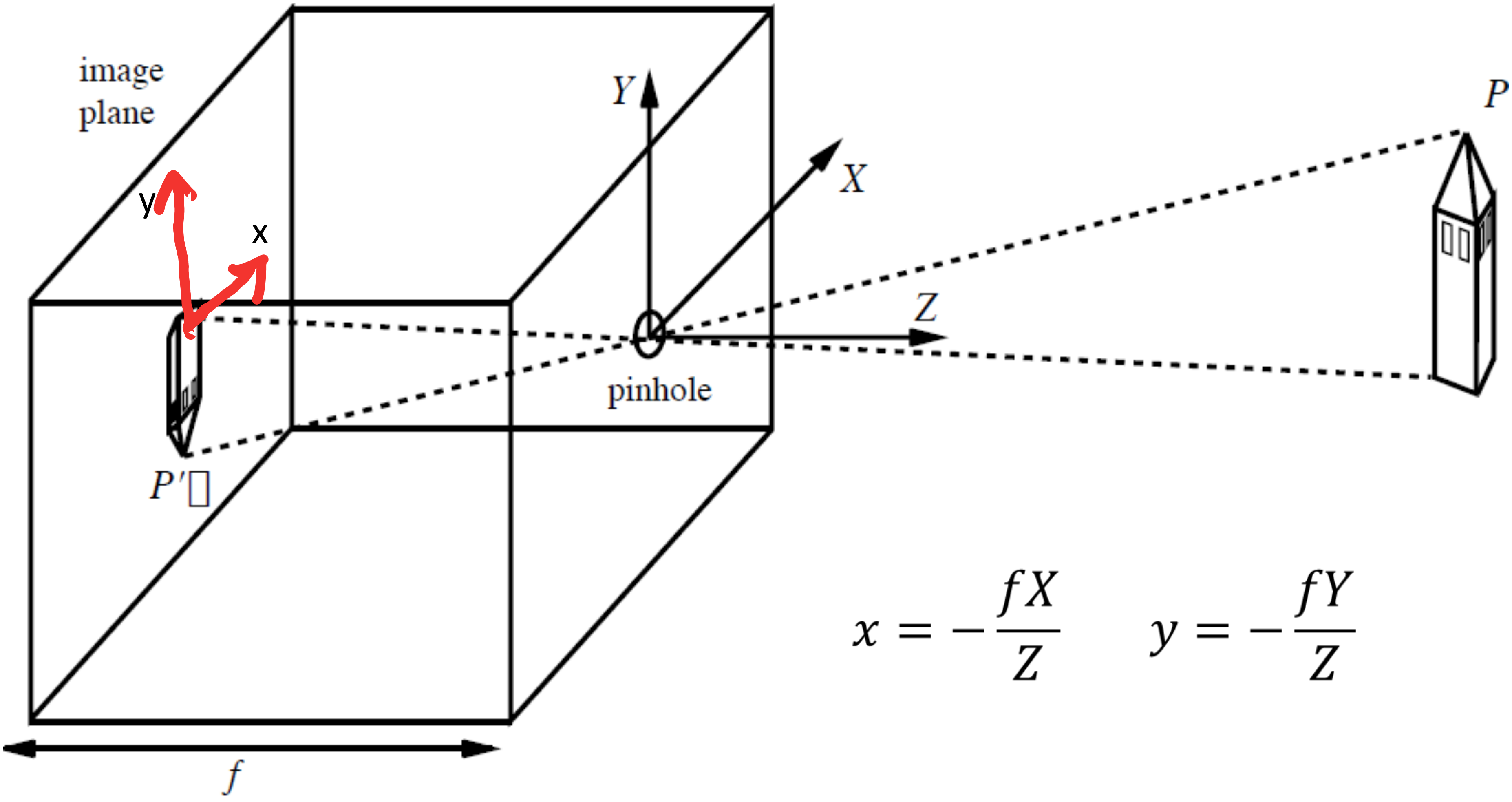


The image I(x,y) measures how much light is captured at pixel (x,y)

We want to know

- Where does a point (X,Y,Z) in the world get imaged?
- What is the brightness at the resulting point (x,y)?

The Pinhole Camera

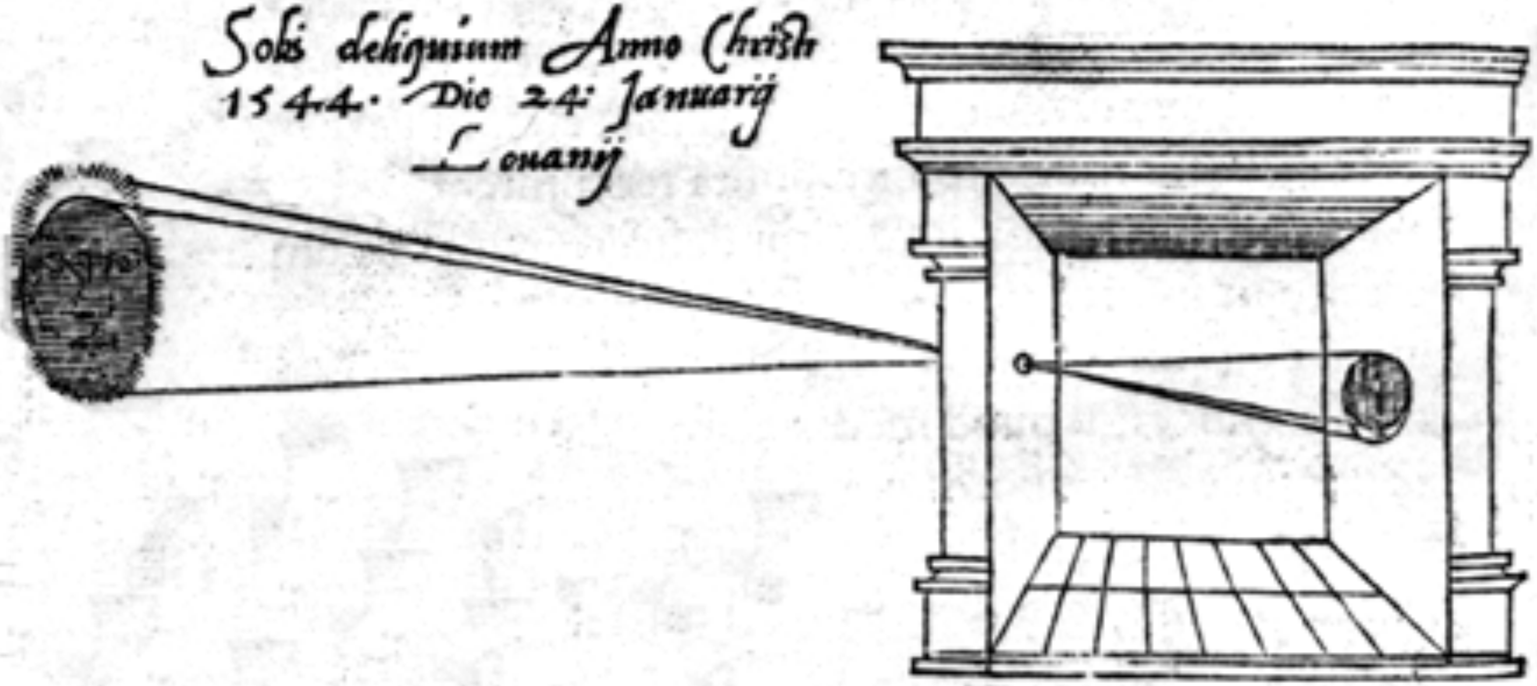


$$x = -\frac{fX}{Z} \qquad y = -\frac{fY}{Z}$$

Camera Obscura

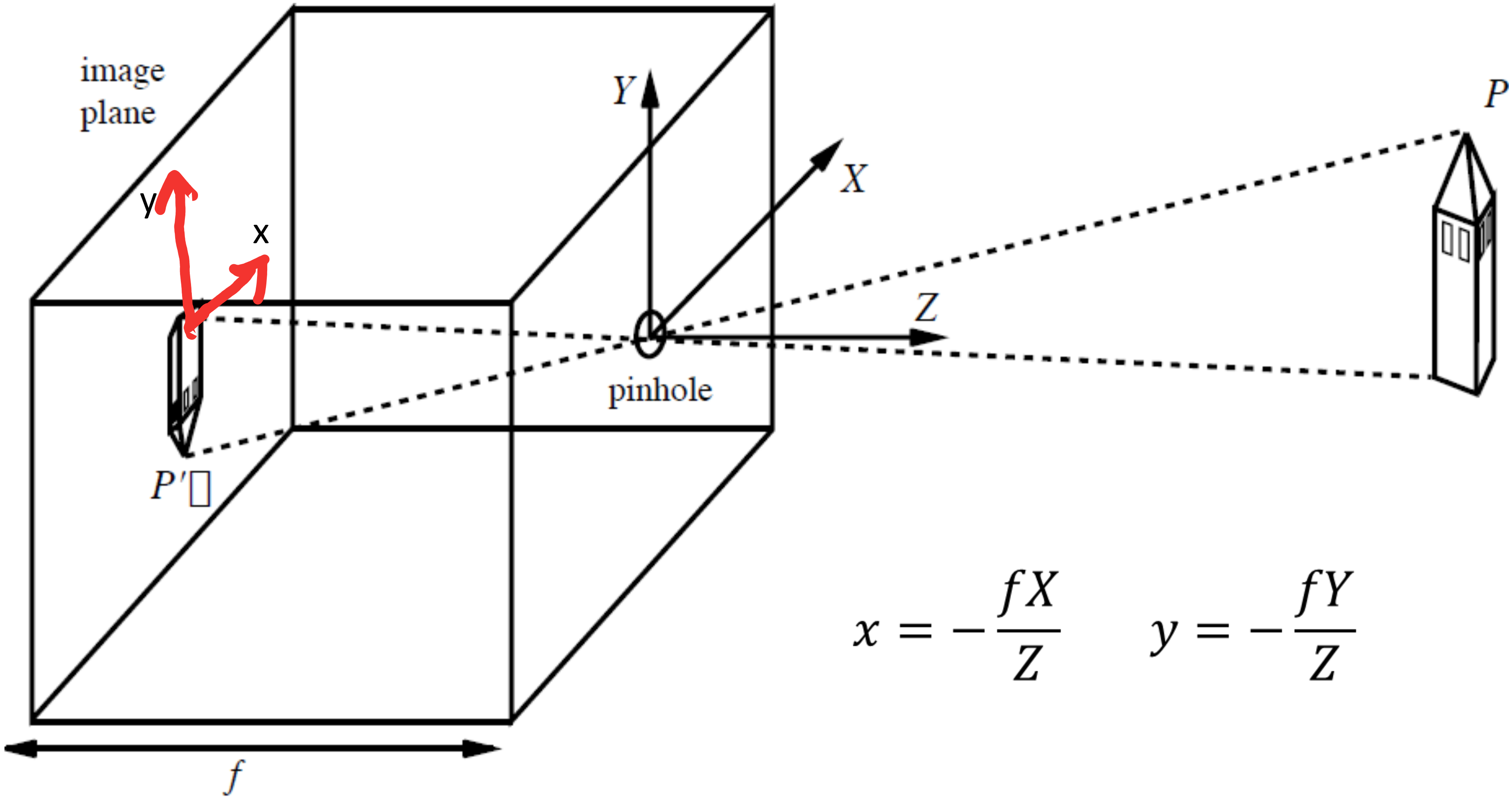
(Reinerus Gemma-Frisius, 1544)

illum in tabula per radios Solis, quàm in cœlo contingit: hoc eſt, ſi in cœlo ſuperior pars deliquiũ patiatur, in radiis apparebit inferior deficere, vt ratio exigit optica.



Sic nos exactè Anno .1544. Louanii eclipſim Solis obſeruauimus, inuenimusq; deficere paulò plus q̃ dex-

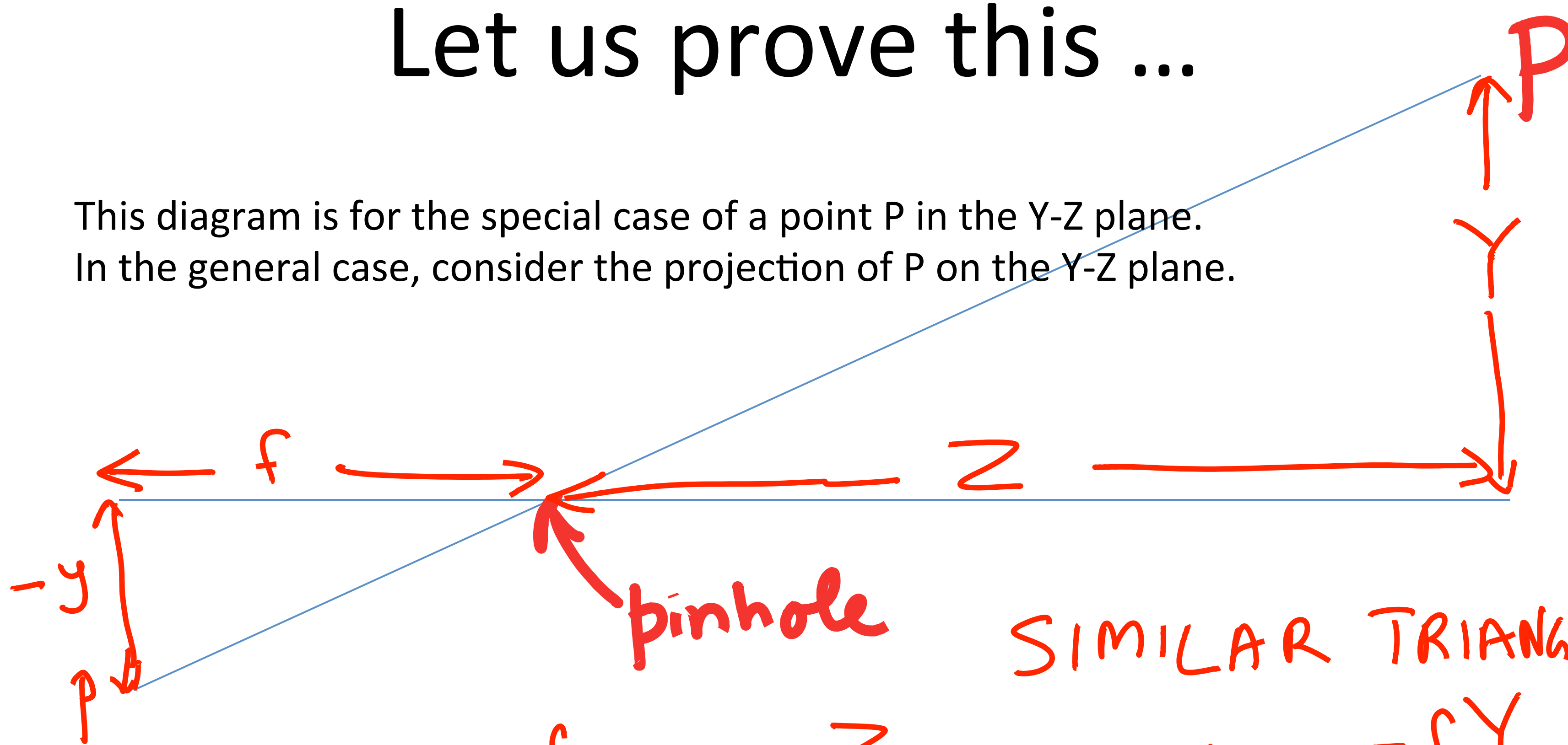
The Pinhole Camera



$$x = -\frac{fX}{Z} \qquad y = -\frac{fY}{Z}$$

Let us prove this …

This diagram is for the special case of a point P in the Y-Z plane.
In the general case, consider the projection of P on the Y-Z plane.



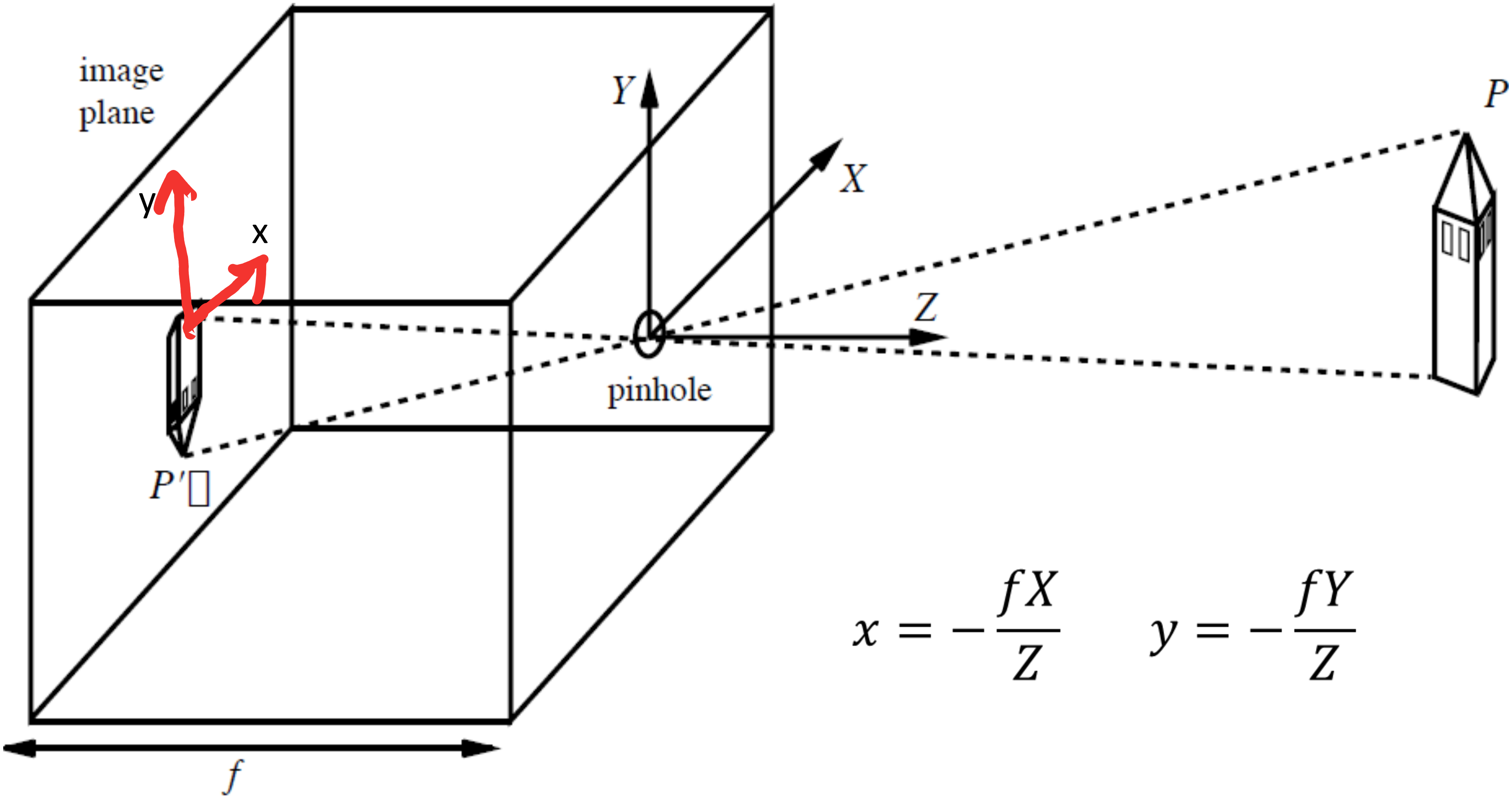
SIMILAR TRIANGLES

$$\frac{f}{-y} = \frac{Z}{Y} \Rightarrow y = \frac{-fY}{Z}$$

This is true even if the point P is not in the YZ plane.

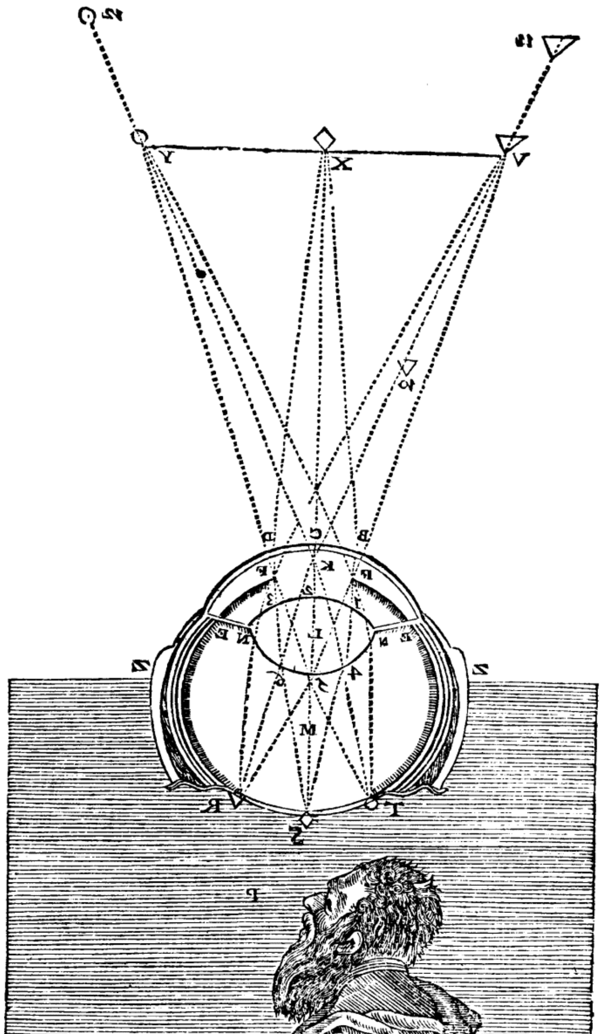
By similar reasoning $x = \dfrac{-fX}{Z}$

The Pinhole Camera



$$x = -\frac{fX}{Z} \qquad y = -\frac{fY}{Z}$$

The image is inverted



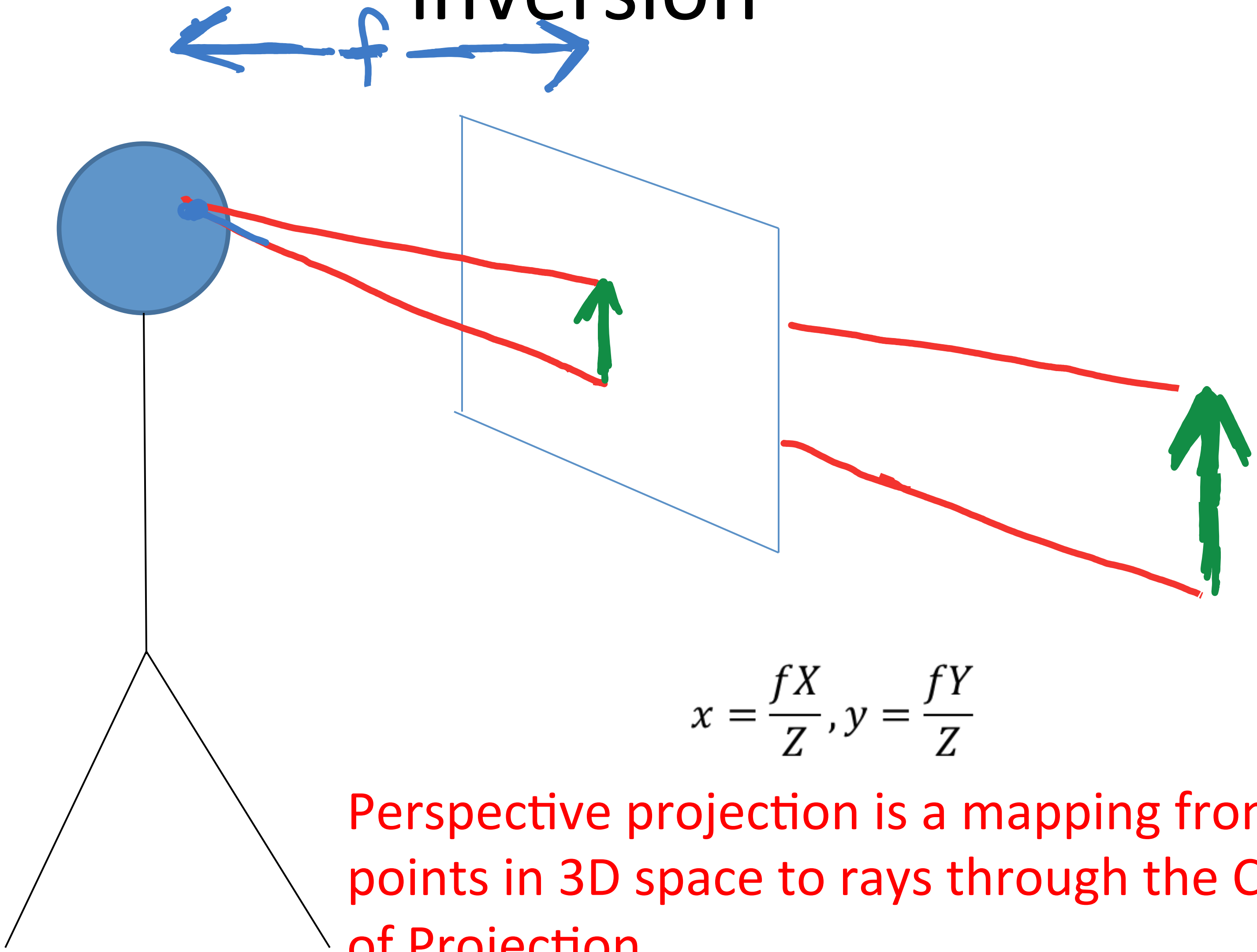
From Descartes(1637), La Dioptrique

This was pointed out by Kepler in 1604

But this is no big deal. The brain can interpret it the right way. And for a camera, software can simply flip the image top-down and right-left. After this trick, we get

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

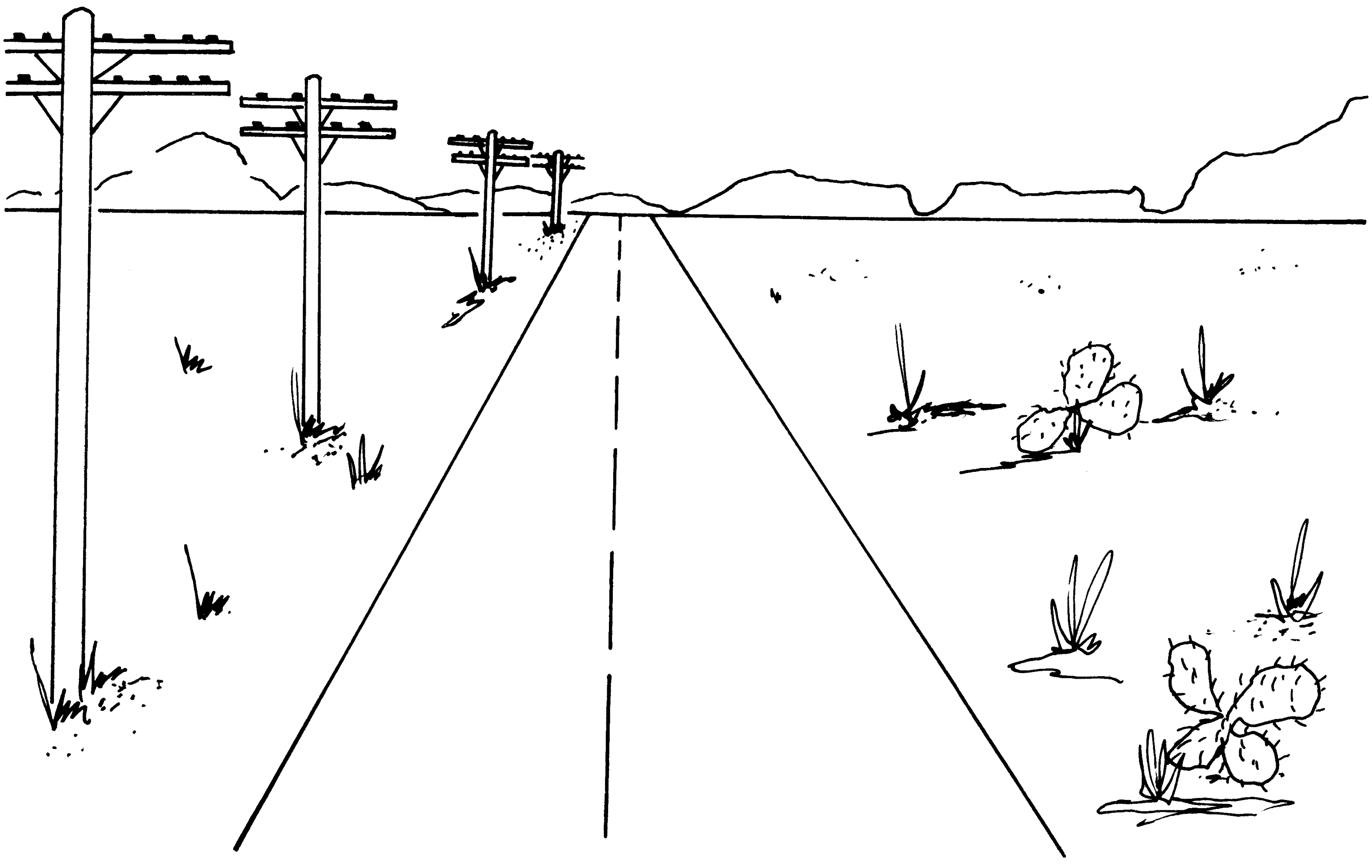
A projection model that avoids inversion



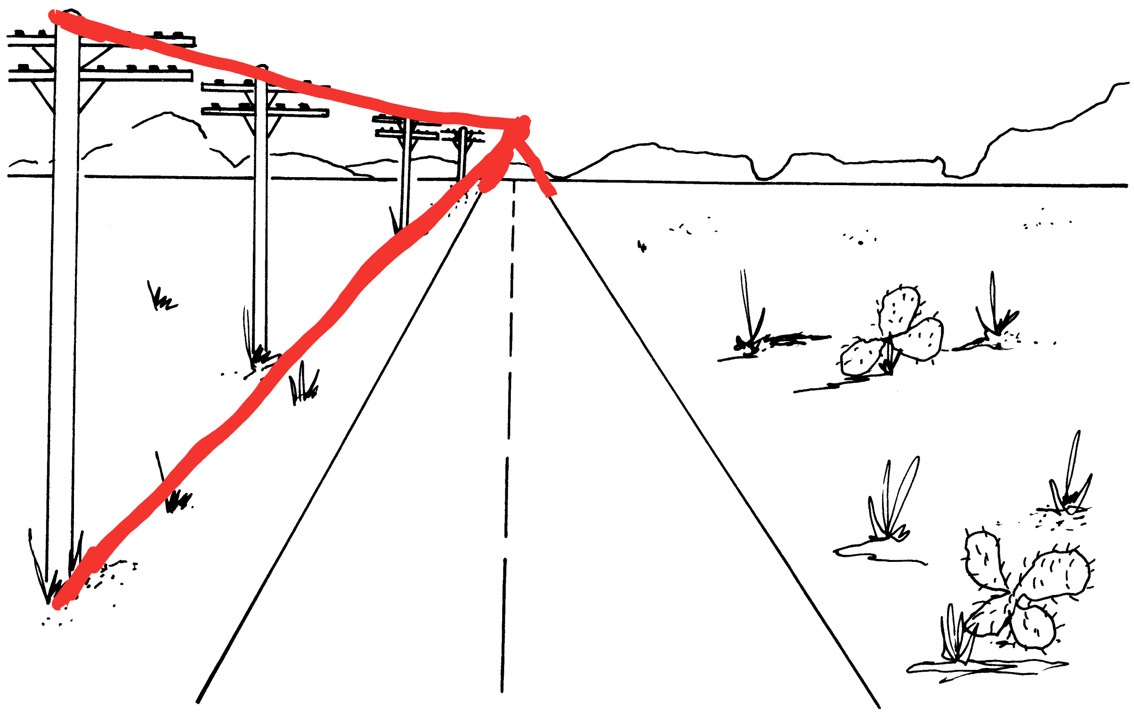
$$x = \frac{fX}{Z}, y = \frac{fY}{Z}$$

Perspective projection is a mapping from points in 3D space to rays through the Center of Projection

Some perspective phenomena…

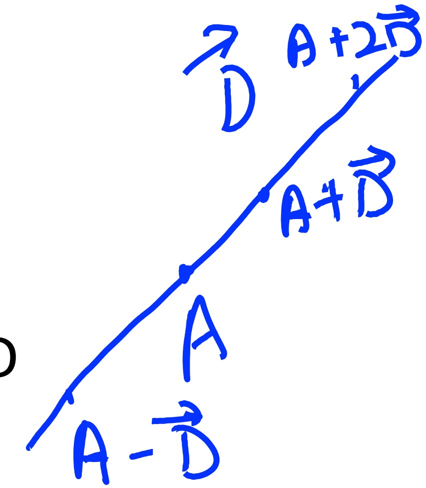


Parallel lines converge to a vanishing point



Each family of parallel lines has its own vanishing point

Proof



Let there be a point A and a direction vector D in three dimensional space.

$$\begin{bmatrix}X\\Y\\Z\end{bmatrix} = \begin{bmatrix}A_x\\A_y\\A_z\end{bmatrix} + \lambda \begin{bmatrix}D_x\\D_y\\D_z\end{bmatrix} \qquad -\infty \leftarrow \lambda \rightarrow \infty$$

$$x = \frac{fX}{Z} = \frac{f(A_x + \lambda D_x)}{A_z + \lambda D_z}$$

Let us consider $\lambda \to \infty$

$$x = \frac{f \lambda D_x}{\lambda D_z} = \frac{f D_x}{D_z}$$

This expression does not depend on A

Coordinates of the projected point are

$$\frac{f D_x}{D_z}$$ for the x-coordinate

(and by doing the same process for y-coordinate

$$\frac{f D_y}{D_z}$$ for the y-coordinate.

Thus $\left(\frac{f D_x}{D_z} , \frac{f D_y}{D_z} \right)$ are

the coordinates of the vanishing point

..

Each family of parallel lines has its own vanishing point

But this isn't true of the vertical lines. They stay parallel. Why?

Vanishing point in vector notation

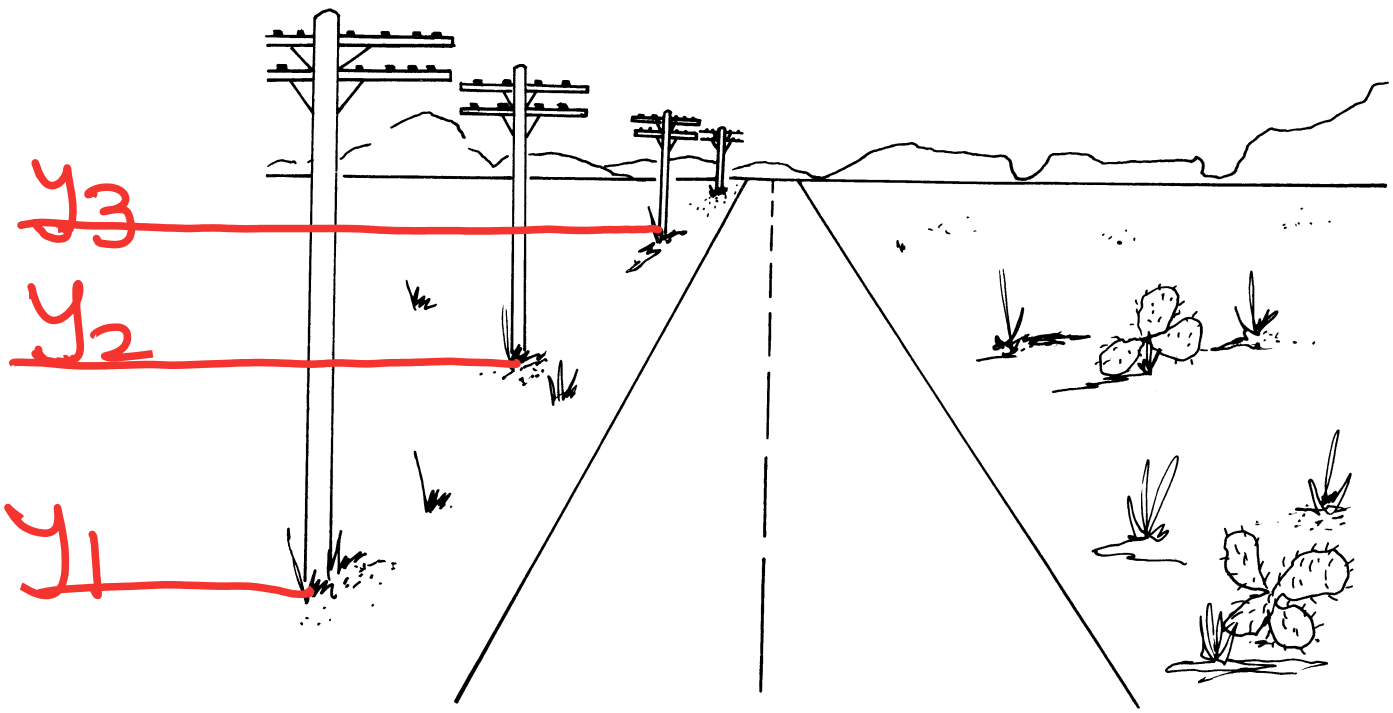
$$\mathbf{p} = f\frac{\mathbf{X}}{Z}$$

A line of points in 3D can be represented as $\mathbf{X} = \mathbf{A} + \lambda\mathbf{D}$, where $\mathbf{A}$ is a fixed point, $\mathbf{D}$ a unit vector parallel to the line, and λ a measure of distance along the line. As λ increases points are increasingly further away and in the limit:

$$\lim_{\lambda\to\infty} \mathbf{p} = f\frac{\mathbf{A} + \lambda\mathbf{D}}{A_Z + \lambda D_Z} = f\frac{\mathbf{D}}{D_Z}$$

i.e. the image of the line terminates in a *vanishing point* with coordinates $(fD_X/D_Z, fD_Y/D_Z)$, unless the line is parallel to the image plane ($D_Z = 0$). Note, the vanishing point is unaffected (invariant to) line position, $\mathbf{A}$, it only depends on line orientation, $\mathbf{D}$. Consequently, the family of lines parallel to $\mathbf{D}$ have the same vanishing point.

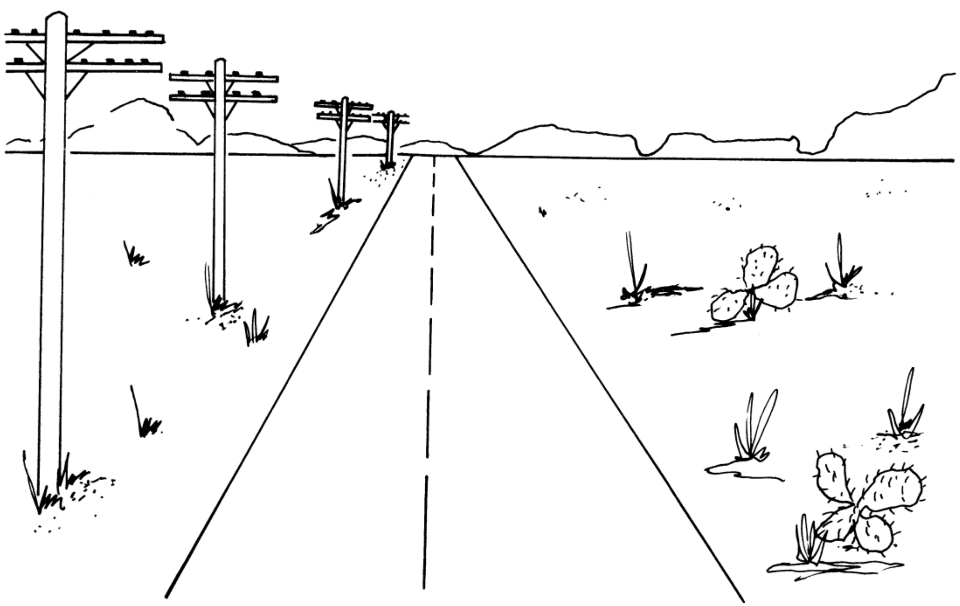
Nearer objects are lower in the image



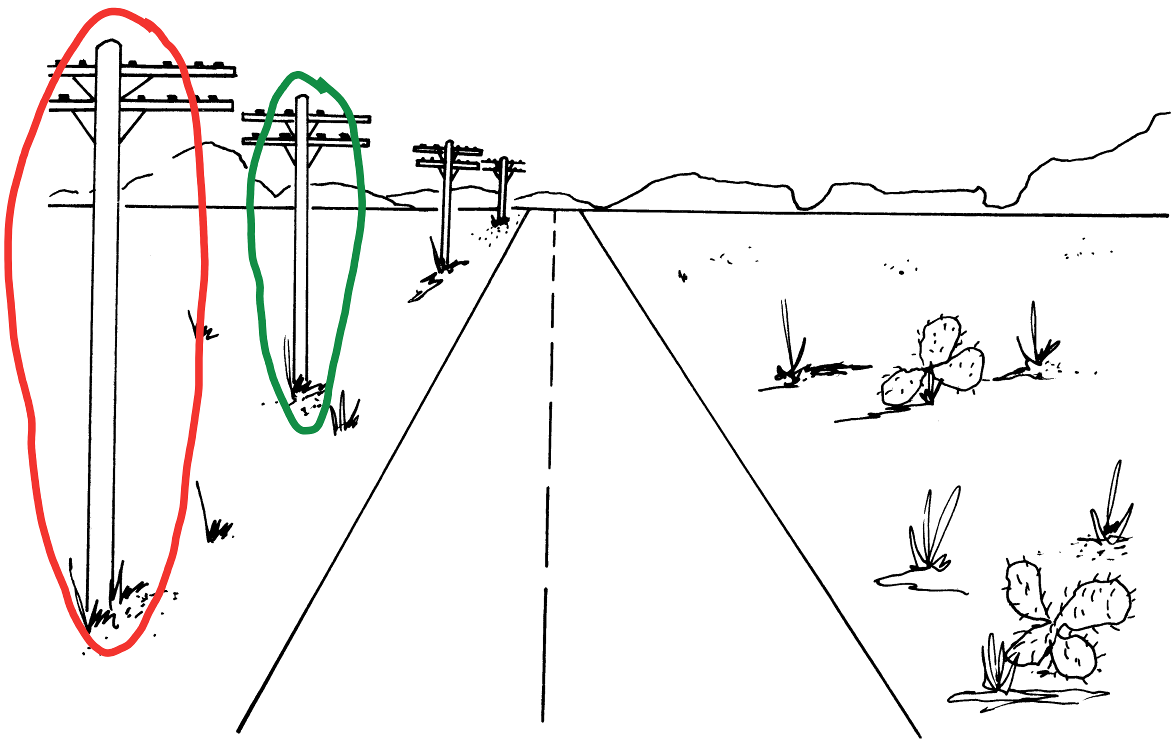
Proof

The equation of the ground plane is $Y = -h$

A point on the ground plane will have y-coordinate $y = -fh/Z$

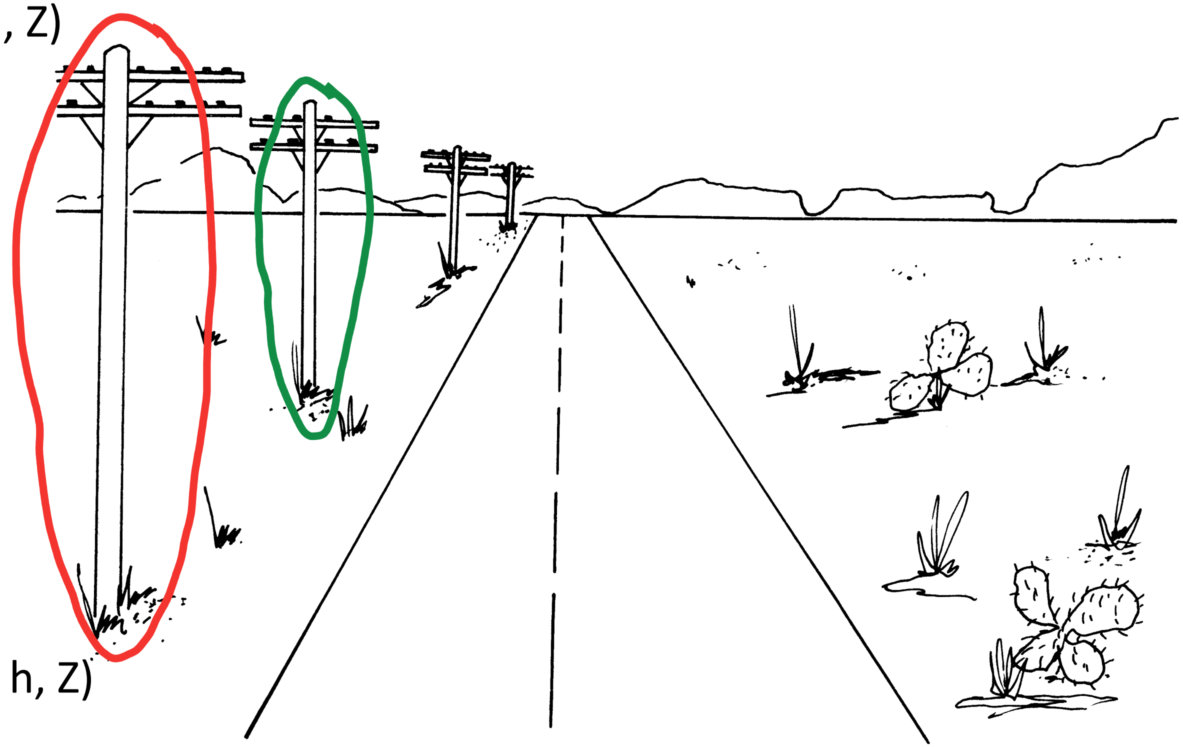


Nearer objects look bigger



Nearer objects look bigger

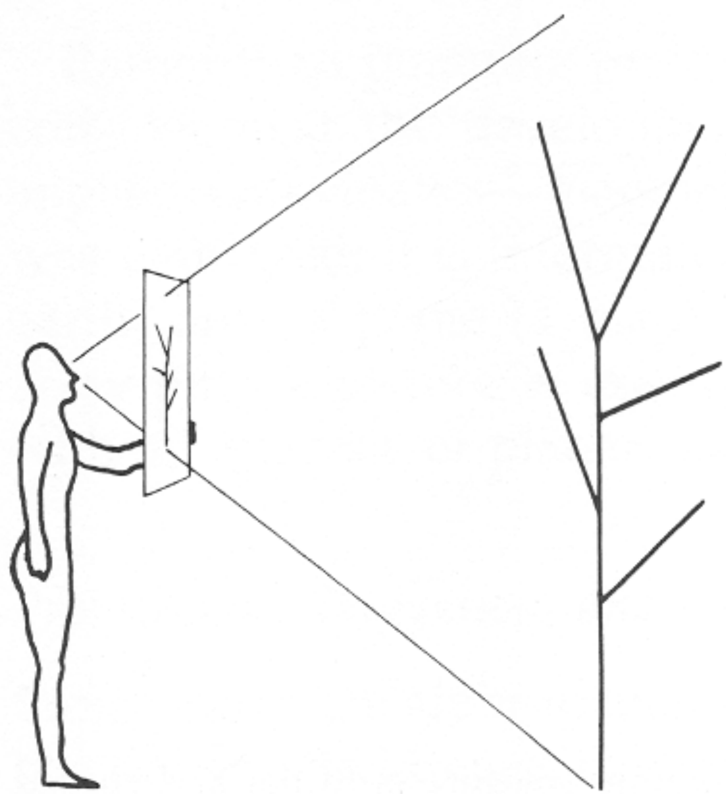
Top at (X, L – h, Z)

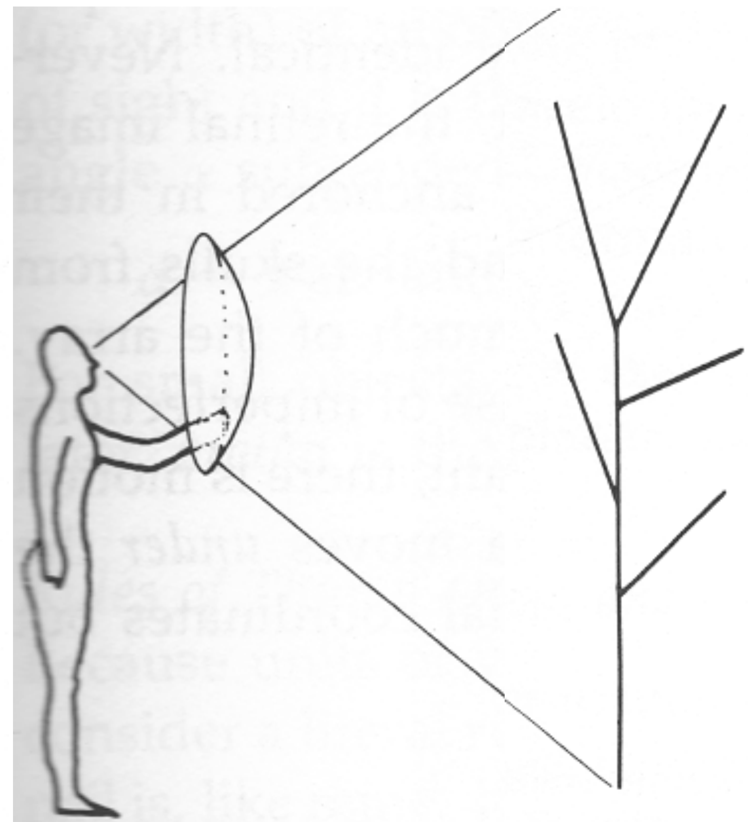


Bottom at (X, - h, Z)

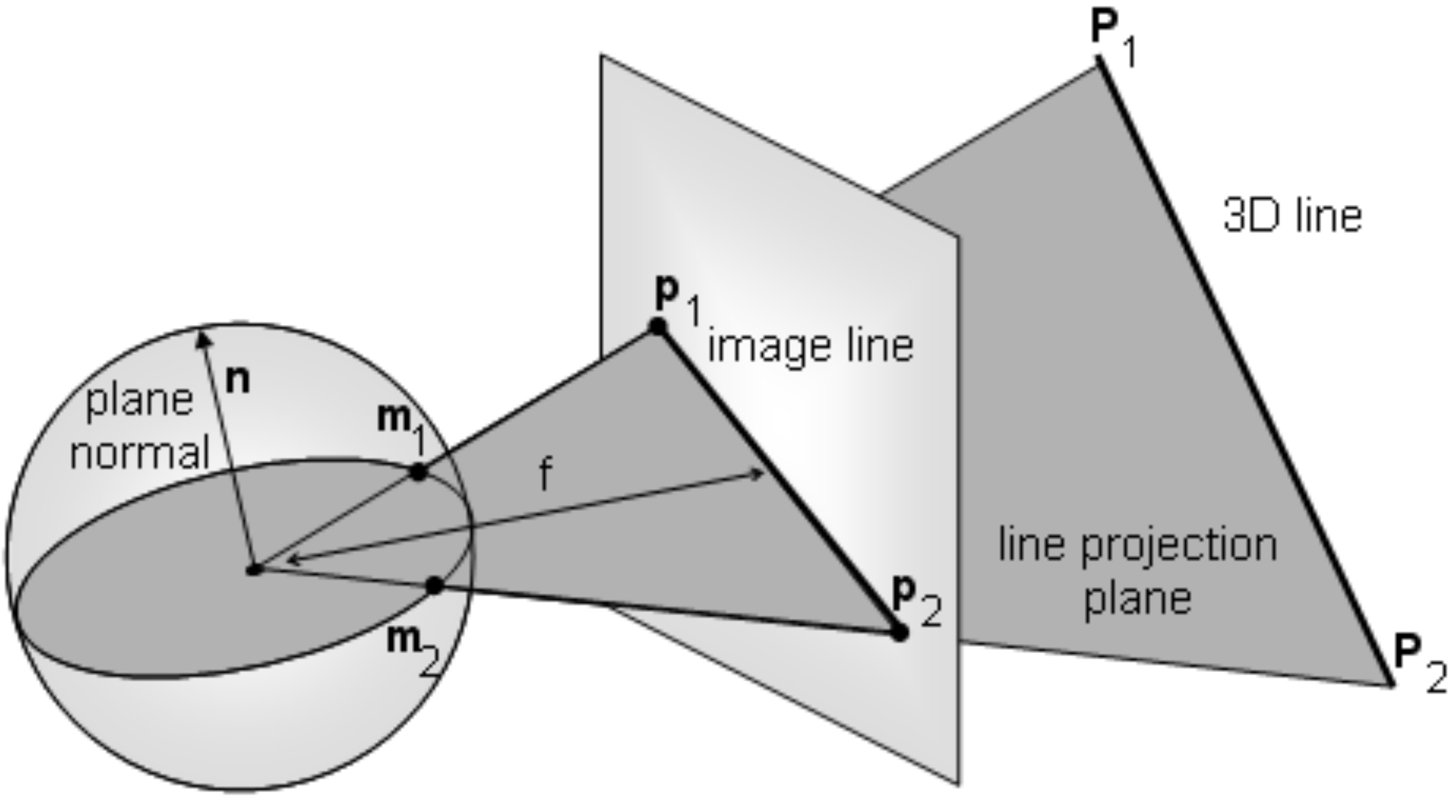
It is straightforward to calculate the projection of the top & bottom of the pole. The difference is the “apparent height”

Perspective projection is a mapping from 3D points to rays through the center of projection; the imaging surface could be planar or spherical



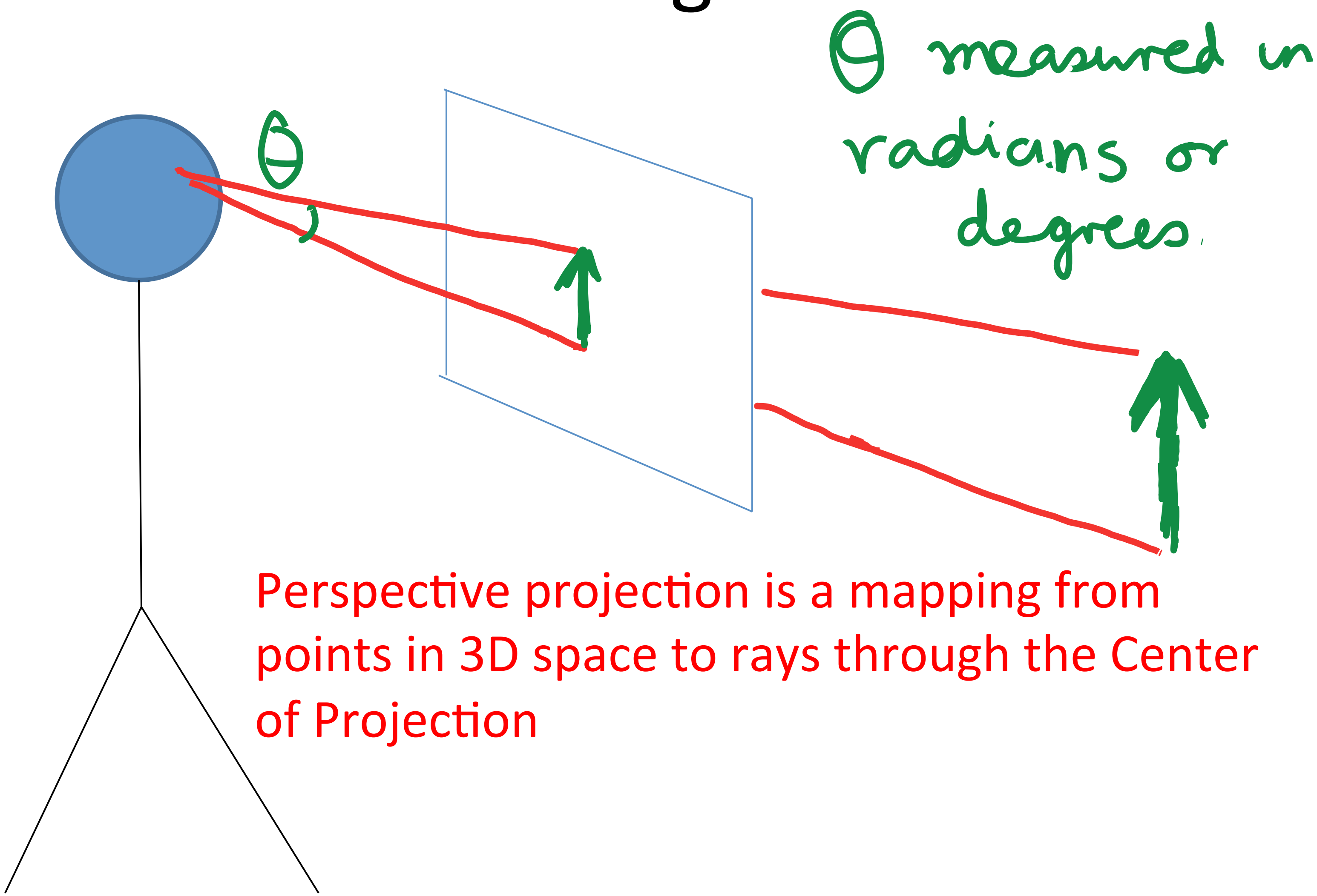


Projection of a line in planar perspective is a line, in spherical perspective is a great circle



(ρ, θ, ϕ) gets mapped to $(1, \theta, \phi)$.

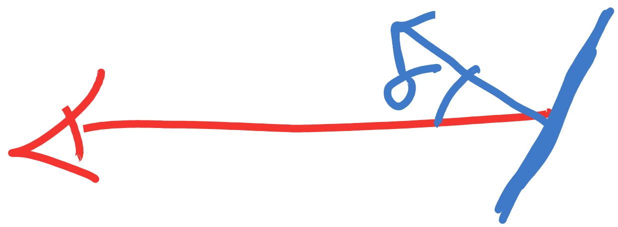
The natural measure of image size is visual angle



Perspective projection is a mapping from points in 3D space to rays through the Center of Projection

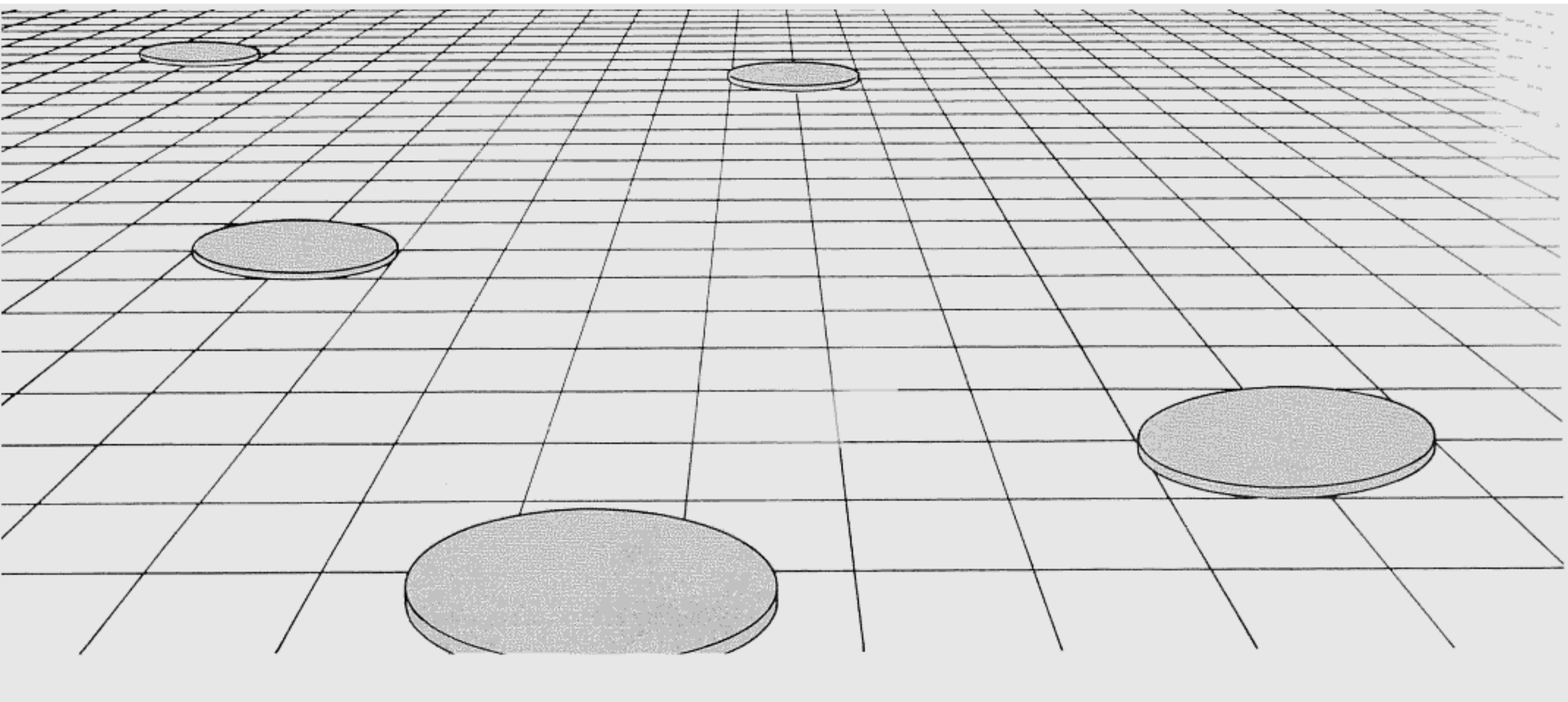
Two main effects of perspective projection

1. Distance – farther objects project to smaller sizes on the image plane. The scaling factor is 1/Z
2. Foreshortening – objects that are slanted with respect to the line of sight project to smaller sizes on the image plane. The scaling factor is cos σ



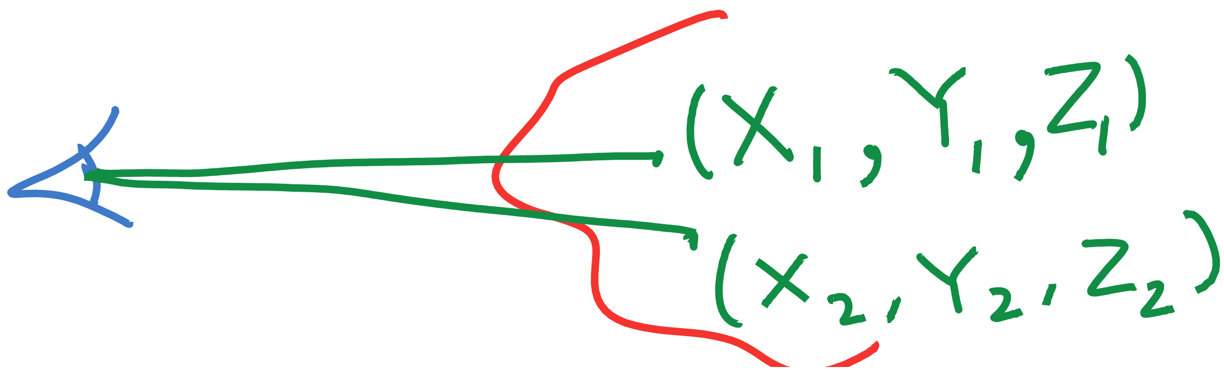
σ is the angle between the line of sight and the surface normal

The slabs that are far away not only look smaller, but also more foreshortened



Orthographic projection

Approximation to perspective when the object is relatively far away compared to the depth variation in it



The idea is as follows: If the depth Z of points on the object varies within some range $Z_0 \pm \Delta Z$, with $\Delta Z \ll Z_0$, then the perspective scaling factor f/Z can be approximated by a constant $s = f/Z_0$. The equations for projection from the scene coordinates (X, Y, Z) to the image plane become $x = sX$ and $y = sY$. Note that scaled orthographic projection is an approximation that is valid only for those parts of the scene with not much internal depth variation;

Cartoon. (Drawing by S. Harris; © 1975 The New Yorker Magazine, Inc.)