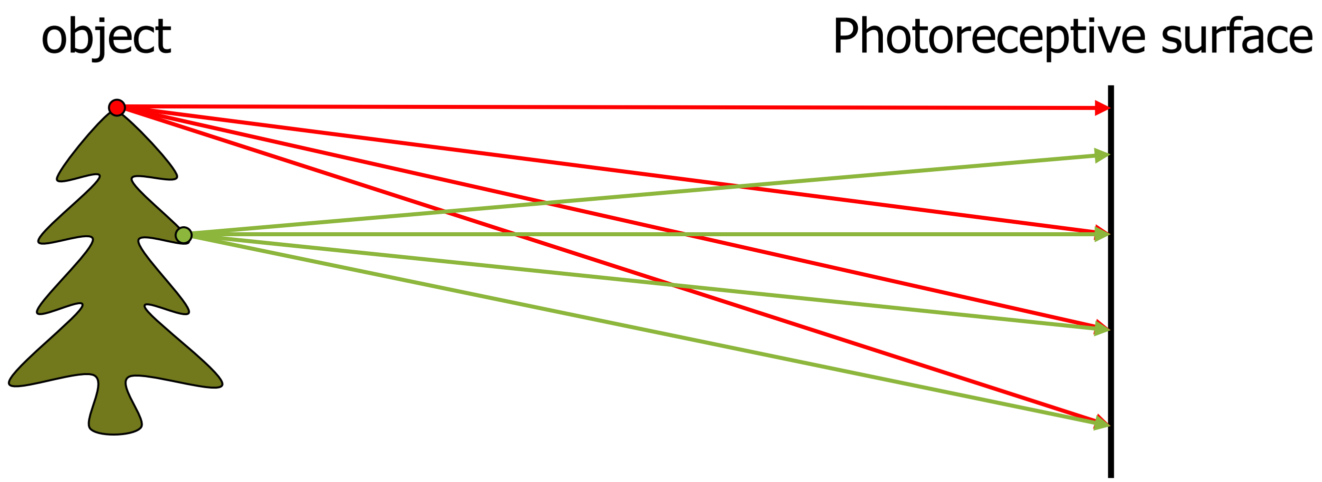
The camera

Sony Cybershot WX1

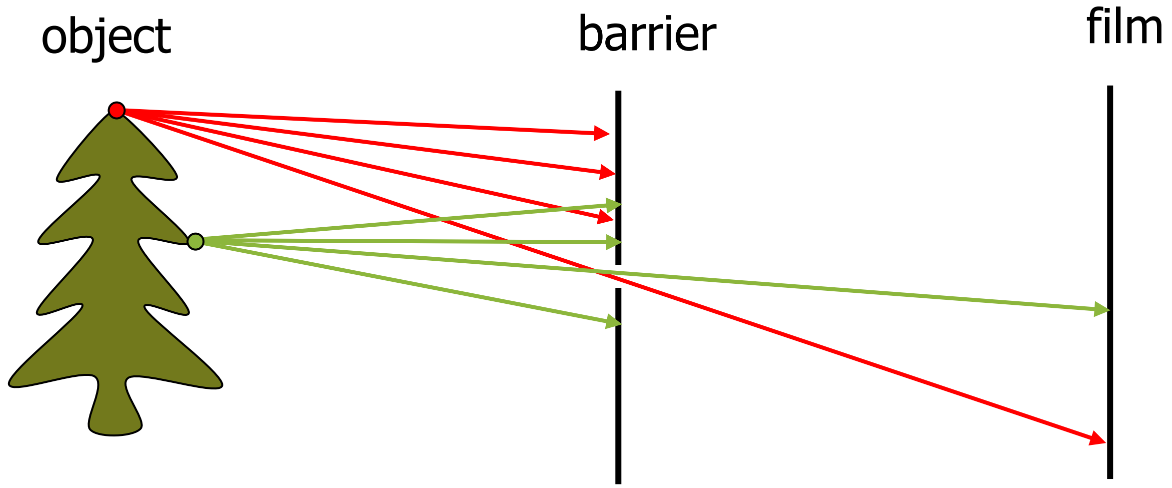
The camera | image formation

- If we place a piece of film in front of an object, do we get a reasonable image?



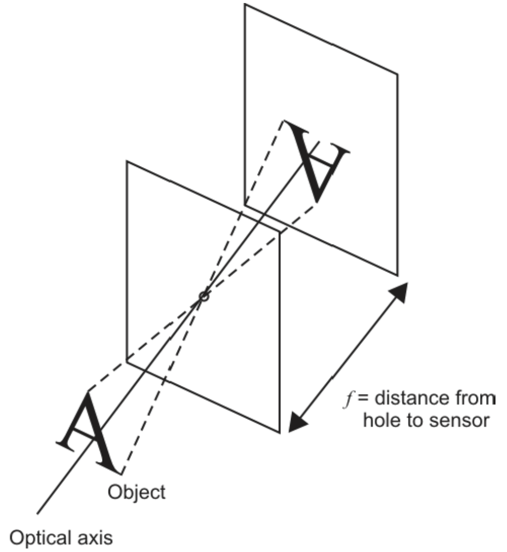
The camera | image formation

- If we place a piece of film in front of an object, do we get a reasonable image?
- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the **aperture**



The camera | camera obscura (pinhole camera)

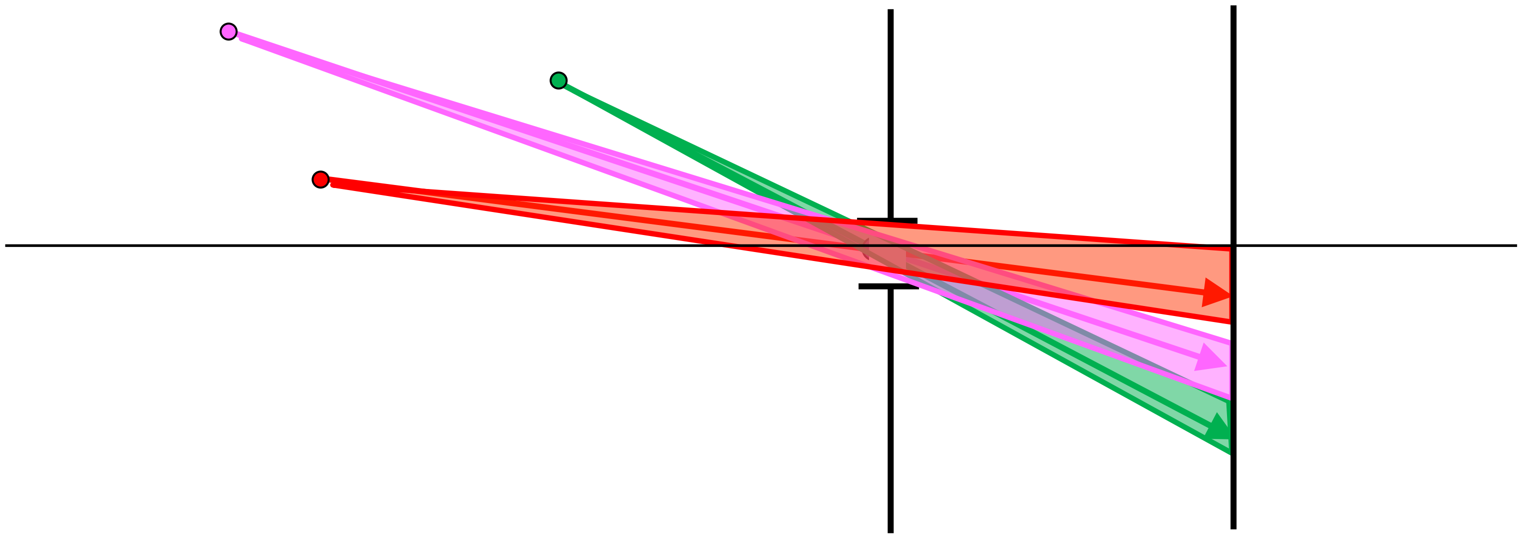
- Pinhole model:
  - Captures **beam of rays** – all rays through a single point
  - The point is called **Center of Projection** or **Optical Center**
  - An “inverted” image is formed on the **Image Plane**
- We will use the pinhole camera model to describe how the image is formed



Gemma-Frisius (1508–1555)

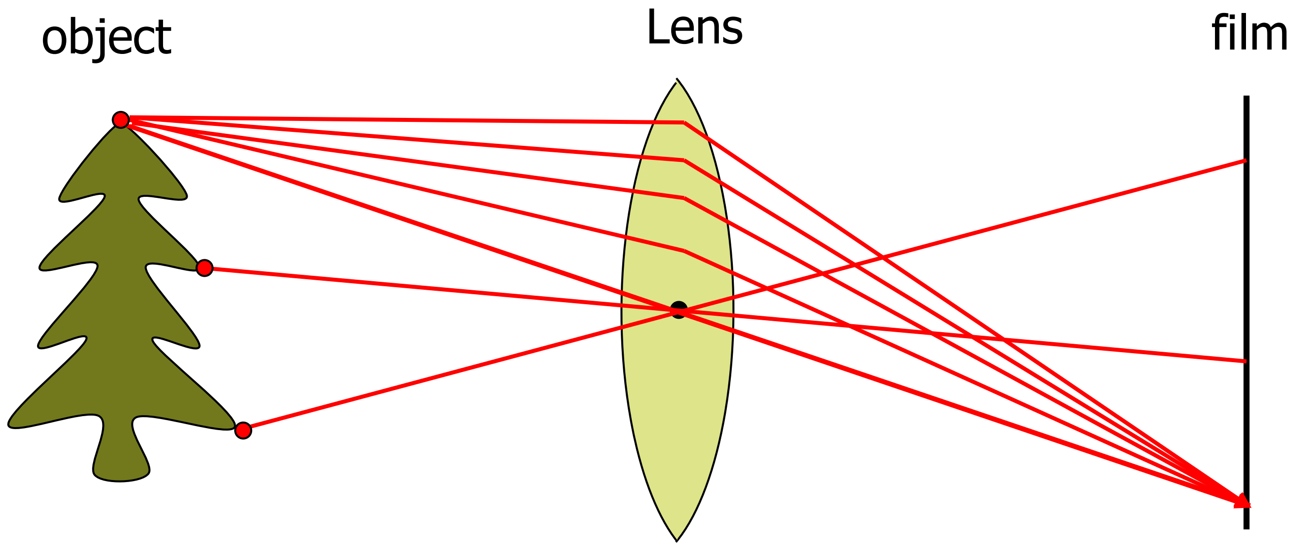
The camera | why use a lens?

- The ideal pinhole: only one ray of light reaches each point on the film
  - ➪ image can be very dim; gives rise to diffraction effects
- Making the pinhole bigger (i.e. aperture) makes the image blurry



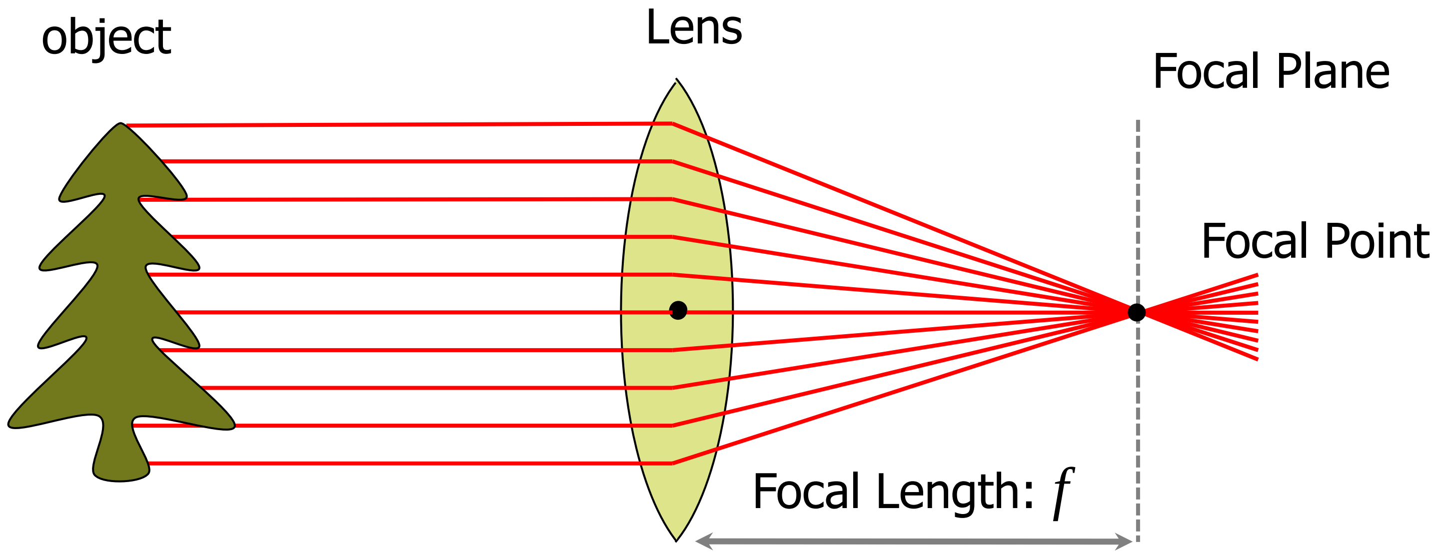
The camera | why use a lens?

- A lens focuses light onto the film
- Rays passing through the **optical center** are not deviated



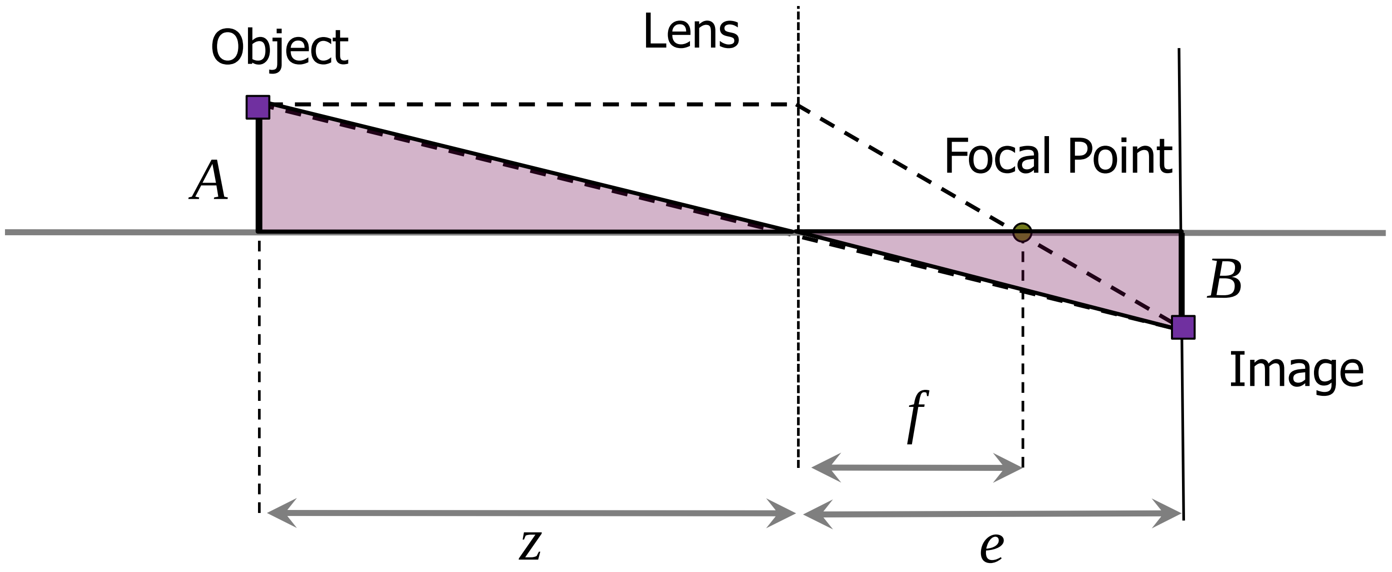
The camera | why use a lens?

- A lens focuses light onto the film
- Rays passing through the **optical center** are not deviated
- All rays parallel to the **optical axis** converge at the **focal point**



The camera | this lens equation

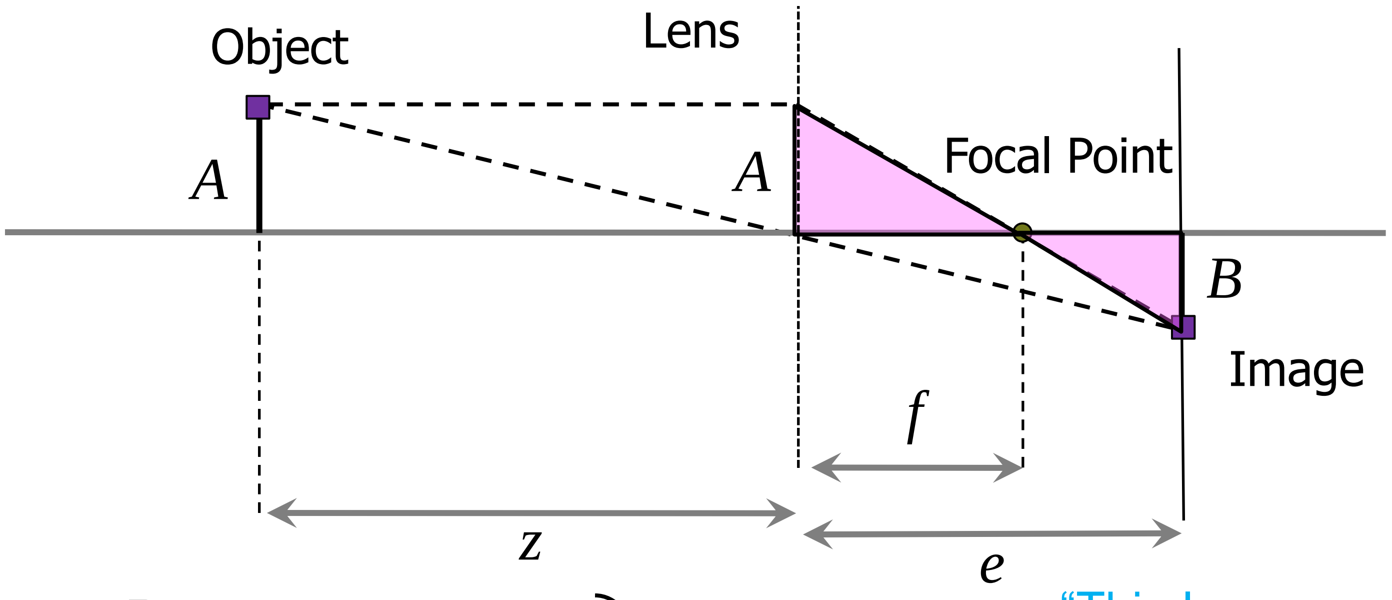
- What is the relationship between f, z, and e ?



- Similar Triangles: $\frac{B}{A} = \frac{e}{z}$

The camera | this lens equation

- What is the relationship between f, z, and e ?



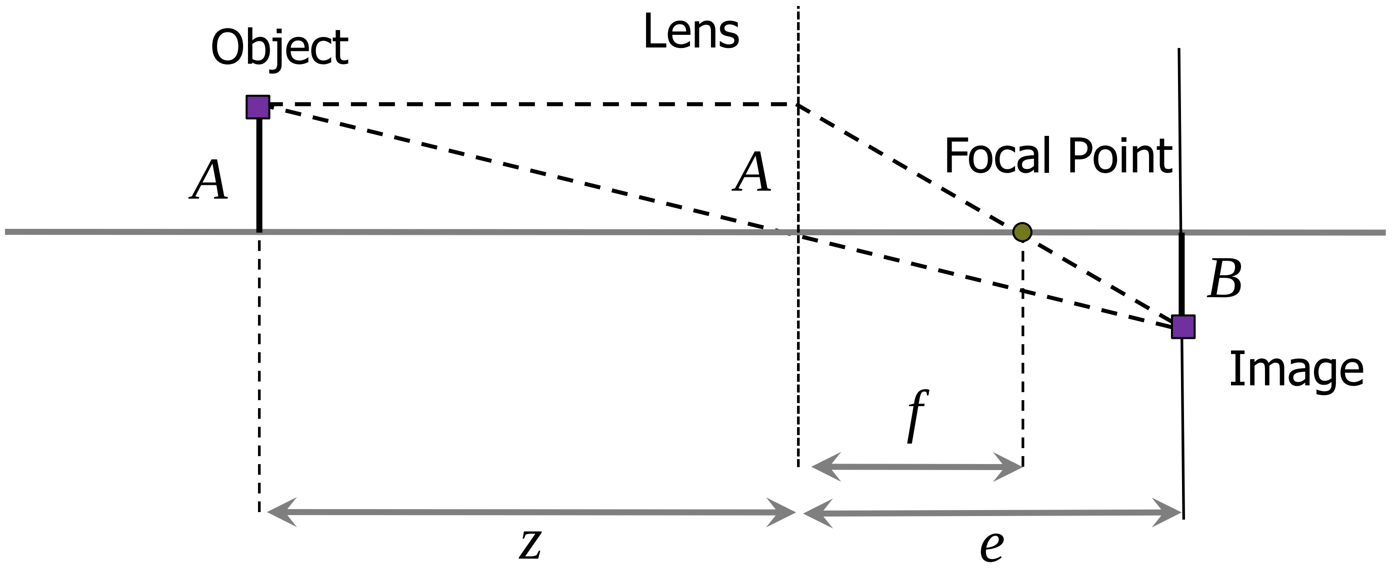
- Similar Triangles: $\frac{B}{A} = \frac{e}{z}$, $\frac{B}{A} = \frac{e-f}{f} = \frac{e}{f} - 1$

$$\frac{e}{f} - 1 = \frac{e}{z} \Rightarrow \boxed{\frac{1}{f} = \frac{1}{z} + \frac{1}{e}}$$

"Thin lens equation"

The camera | pinhole approximation

- What happens if $z \gg f$?

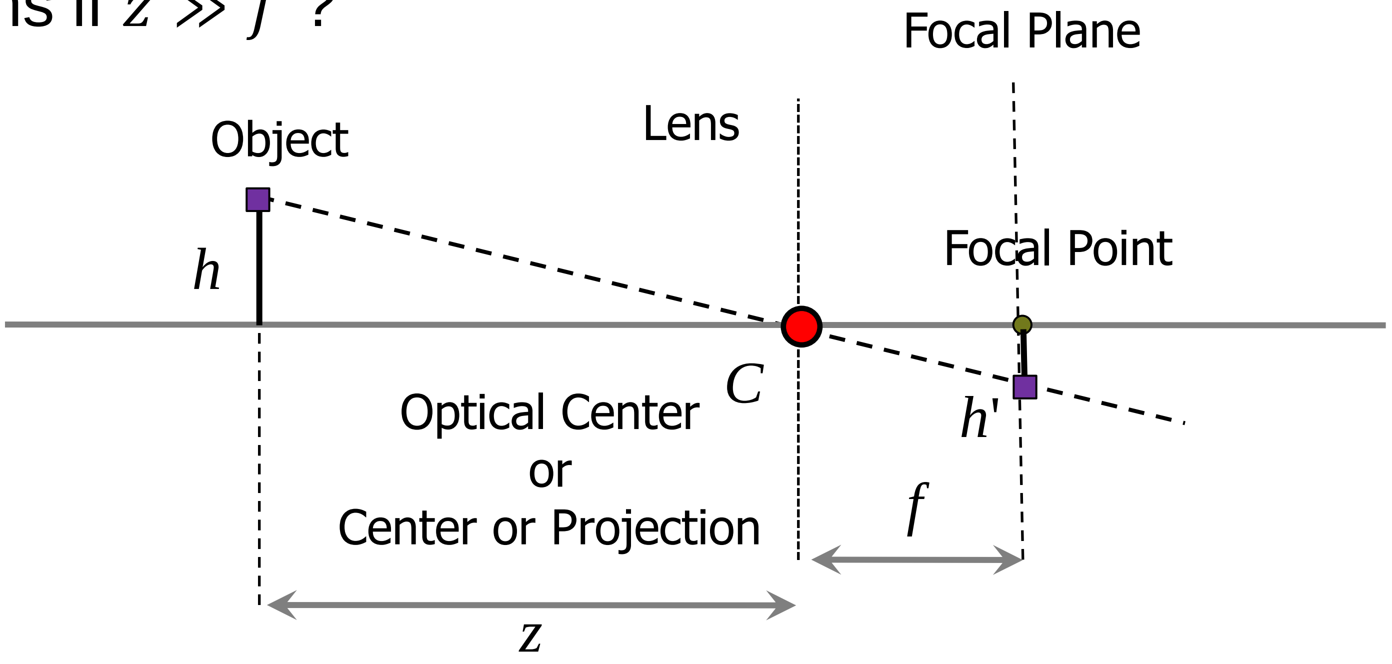


- We need to adjust the image plane such that objects at infinity are in focus

$$\frac{1}{f} = \underbrace{\boxed{\frac{1}{z}}}_{\cong 0} + \frac{1}{e} \Rightarrow \frac{1}{f} \approx \frac{1}{e} \Rightarrow f \approx e$$

The camera | pinhole approximation

- What happens if $z \gg f$?

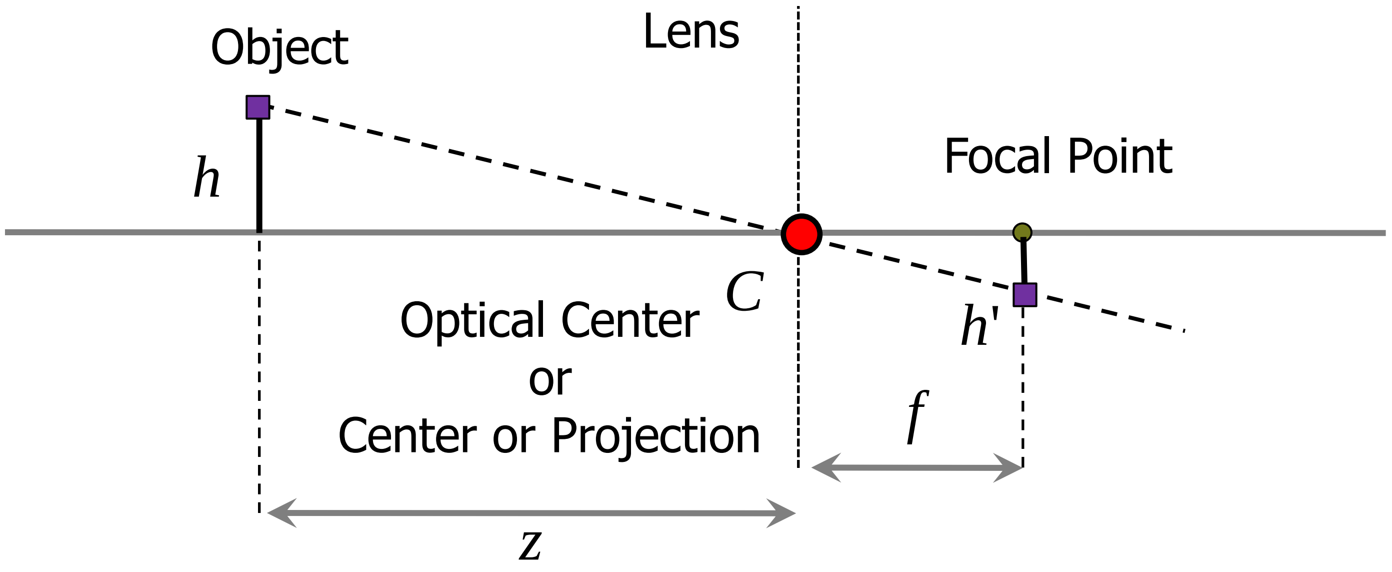


- We need to adjust the image plane such that objects at infinity are in focus

$$\frac{1}{f} = \underbrace{\boxed{\frac{1}{z}}}_{\cong 0} + \frac{1}{e} \Rightarrow \frac{1}{f} \approx \frac{1}{e} \Rightarrow f \approx e$$

The camera | pinhole approximation

- What happens if $z \gg f$?

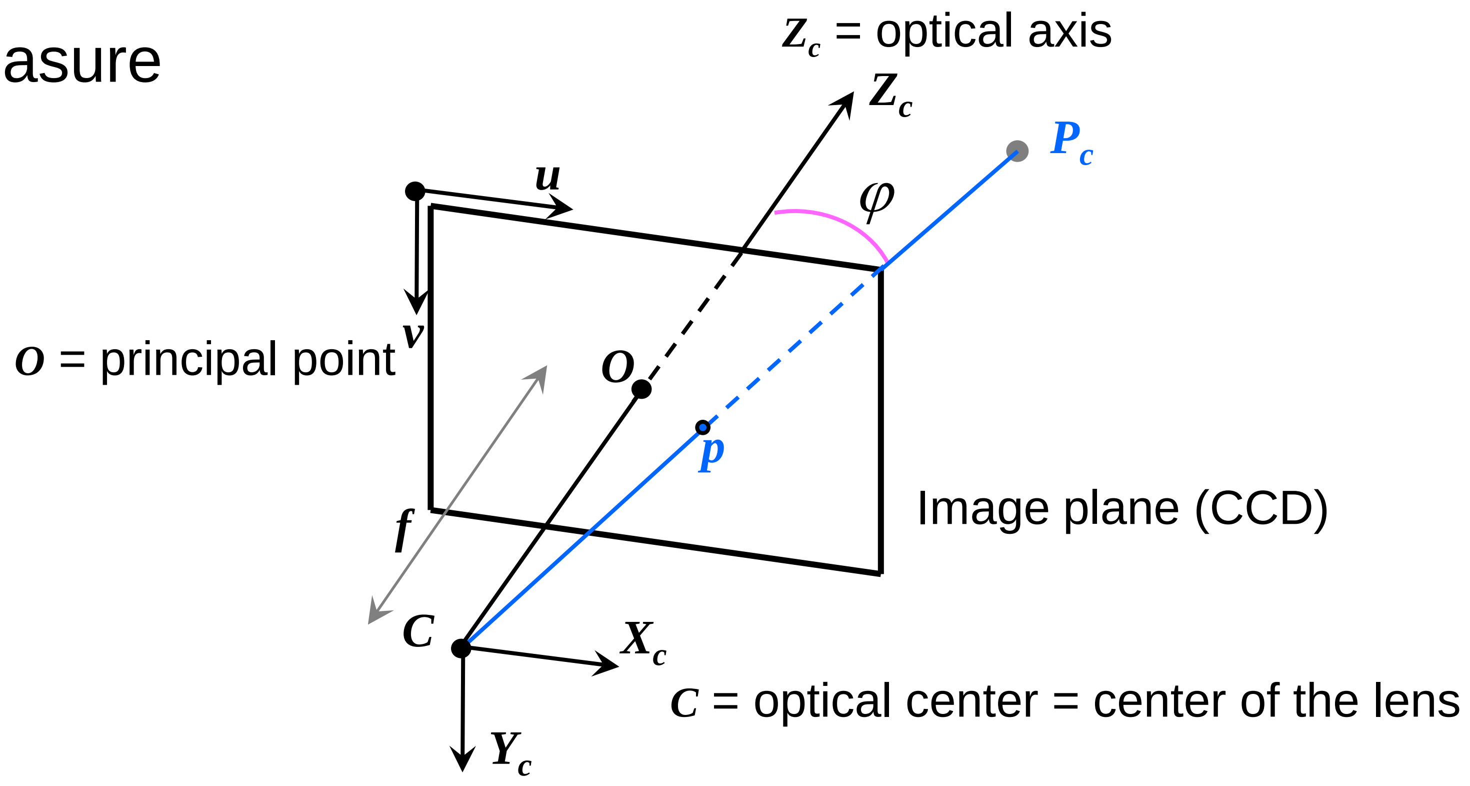


$$\frac{h'}{h} = \frac{f}{z} \Rightarrow h' = \frac{f}{z} h$$

The camera | perspective camera

- For convenience, the image plane is usually represented in front of C such that the image preserves the same orientation (i.e. not flipped)

- A camera does not measure distances but angles!



Perspective projection| from scene points to pixels

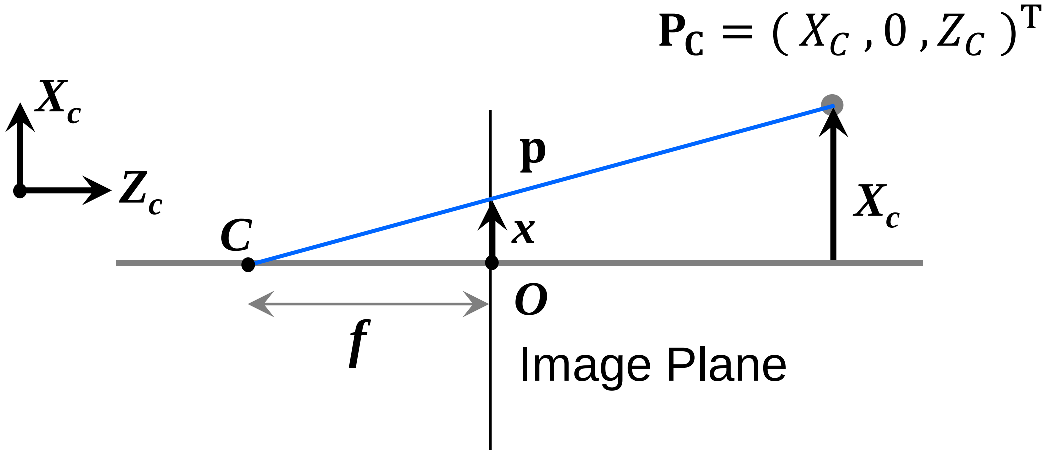
- The Camera point $\mathbf{P_C} = (X_C, 0, Z_C)^\mathrm{T}$ projects to $\mathbf{p} = (x, y)$ onto the image plane

- From similar triangles:

$$\frac{x}{f} = \frac{X_c}{Z_c} \Rightarrow x = \frac{fX_c}{Z_c}$$

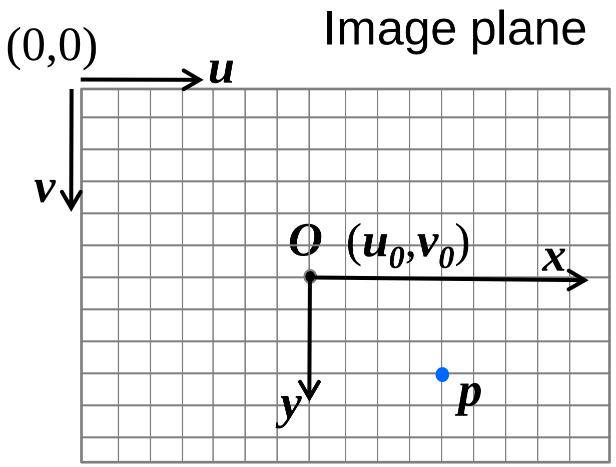
- Similarly, in the general case:

$$\frac{y}{f} = \frac{Y_c}{Z_c} \Rightarrow y = \frac{fY_c}{Z_c}$$



Perspective projection| from scene points to pixels

- To convert $\mathbf{p}$, from the local image plane coordinates (x, y) to the pixel coordinates (u, v), we need to account for:
  - The pixel coordinates of the camera optical center $O = (u_0, v_0)$
  - Scale factor k for the pixel-size



$$u = u_0 + kx \Rightarrow u_0 + k\frac{fX_C}{Z_C}$$

$$v = v_0 + ky \Rightarrow v_0 + k\frac{fY_C}{Z_C}$$

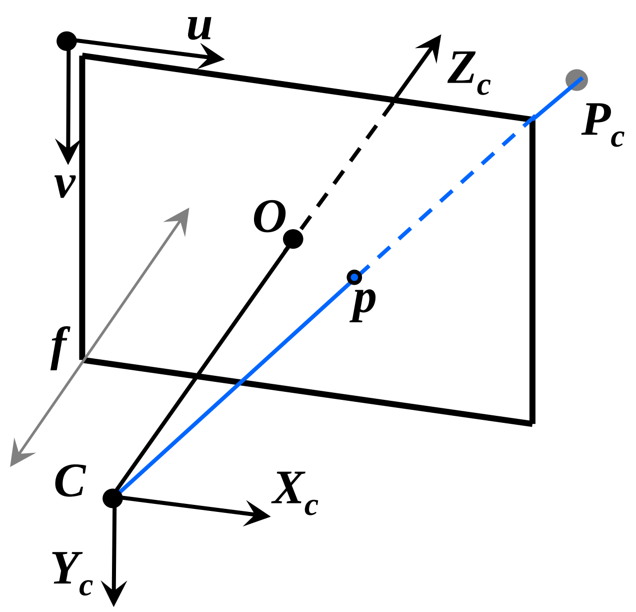
- Use Homogeneous Coordinates for linear mapping from 3D to 2D, by introducing an extra element (scale):

$$p = \begin{pmatrix} u \\ v \end{pmatrix} \Rightarrow \tilde{p} = \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

Perspective projection| from scene points to pixels

$$u = u_0 + kx \Rightarrow u_0 + k\frac{fX_C}{Z_C}$$

$$v = v_0 + ky \Rightarrow v_0 + k\frac{fY_C}{Z_C}$$



- Expressed in matrix form and homogeneous coordinates:

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} kf & 0 & u_0 \\ 0 & kf & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$

- Or alternatively

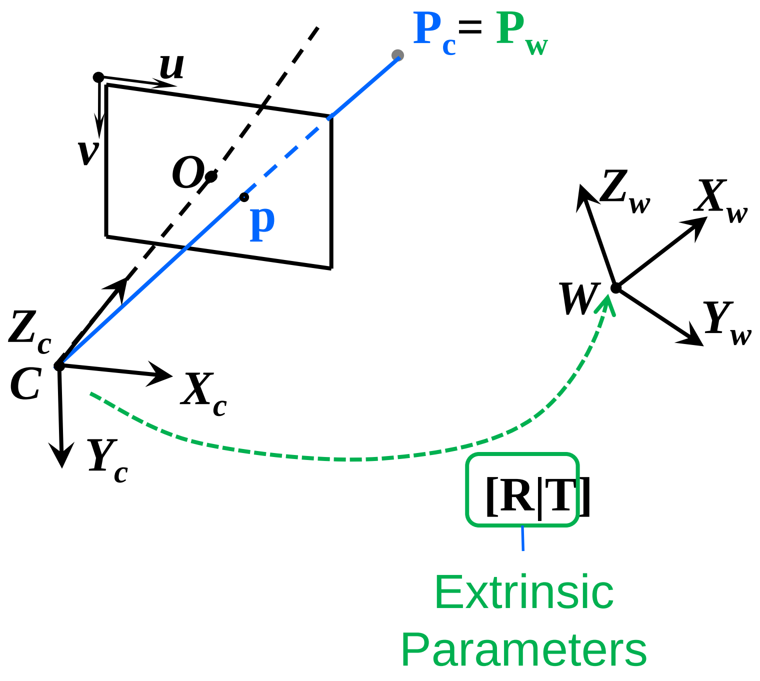
$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} \alpha & 0 & u_0 \\ 0 & \alpha & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = K \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$

Focal length in pixels

Intrinsic parameters matrix

Perspective projection| from scene points to pixels

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} & R & | & T & \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

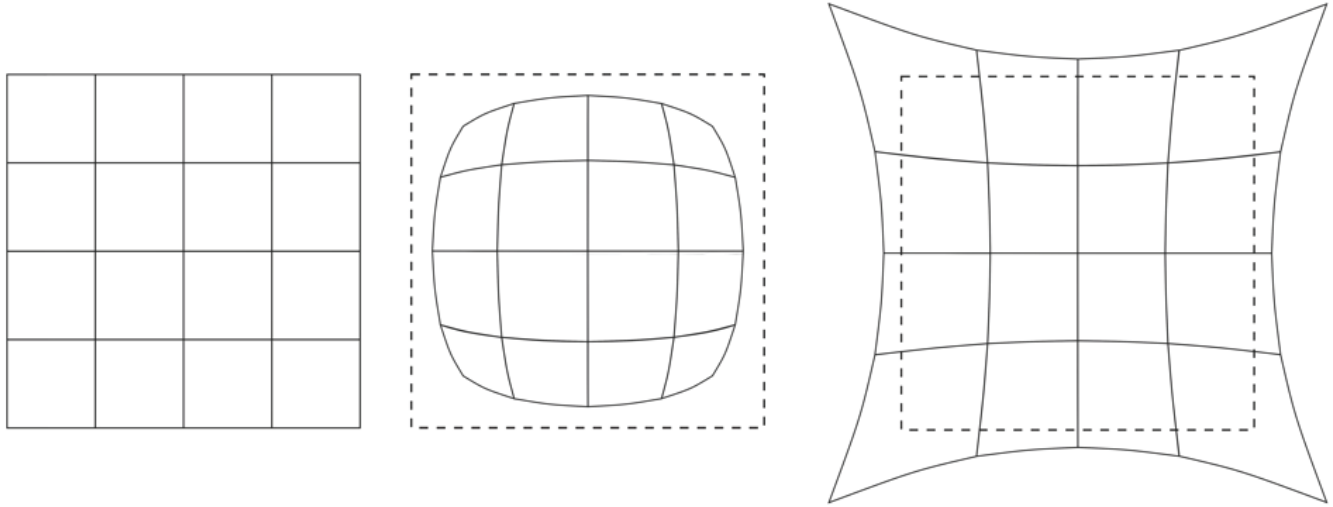


$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$

Perspective Projection Matrix

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boxed{K[R|T]} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Perspective projection| radial distortion



No distortion

Barrel distortion

Pincushion

Perspective projection| radial distortion

- The standard model of radial distortion is a transformation from the ideal coordinates (u, v) (i.e., undistorted) to the real observable coordinates (distorted) (u_d, v_d)

- The amount of distortion of the coordinates of the observed image is a nonlinear function of their radial distance. For most lenses, a simple quadratic model of distortion produces good results

$$\begin{bmatrix} u_d \\ v_d \end{bmatrix} = (1 + k_1 r^2) \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}$$

where

$$r^2 = (u - u_0)^2 + (v - v_0)^2$$