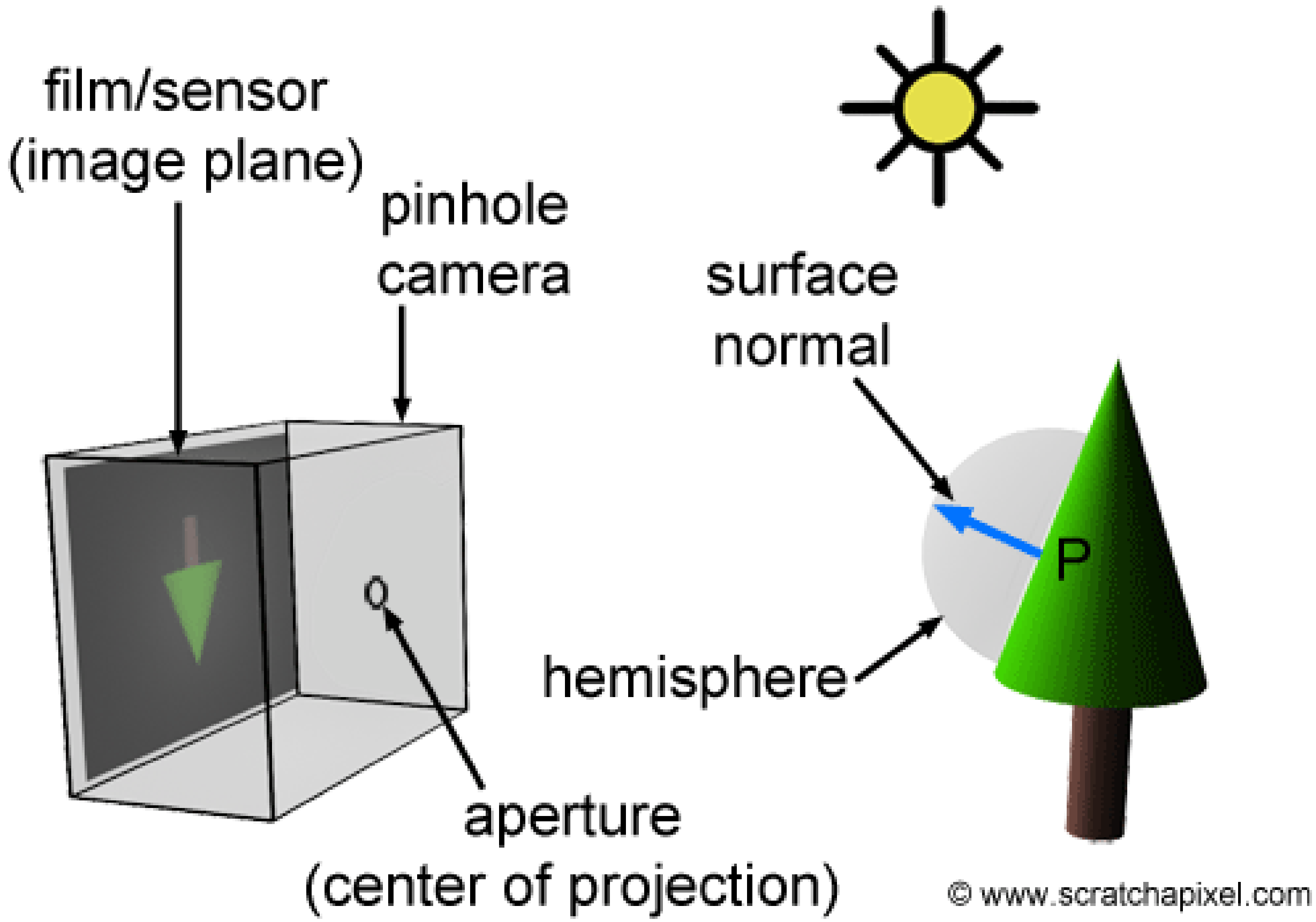
Camera model and calibration

Karel Zimmermann, zimmerk@cmp.felk.cvut.cz
(some images taken from **Tomáš Svoboda's**)
Czech Technical University in Prague, Center for Machine Perception
http://cmp.felk.cvut.cz

Last update: March 13, 2017

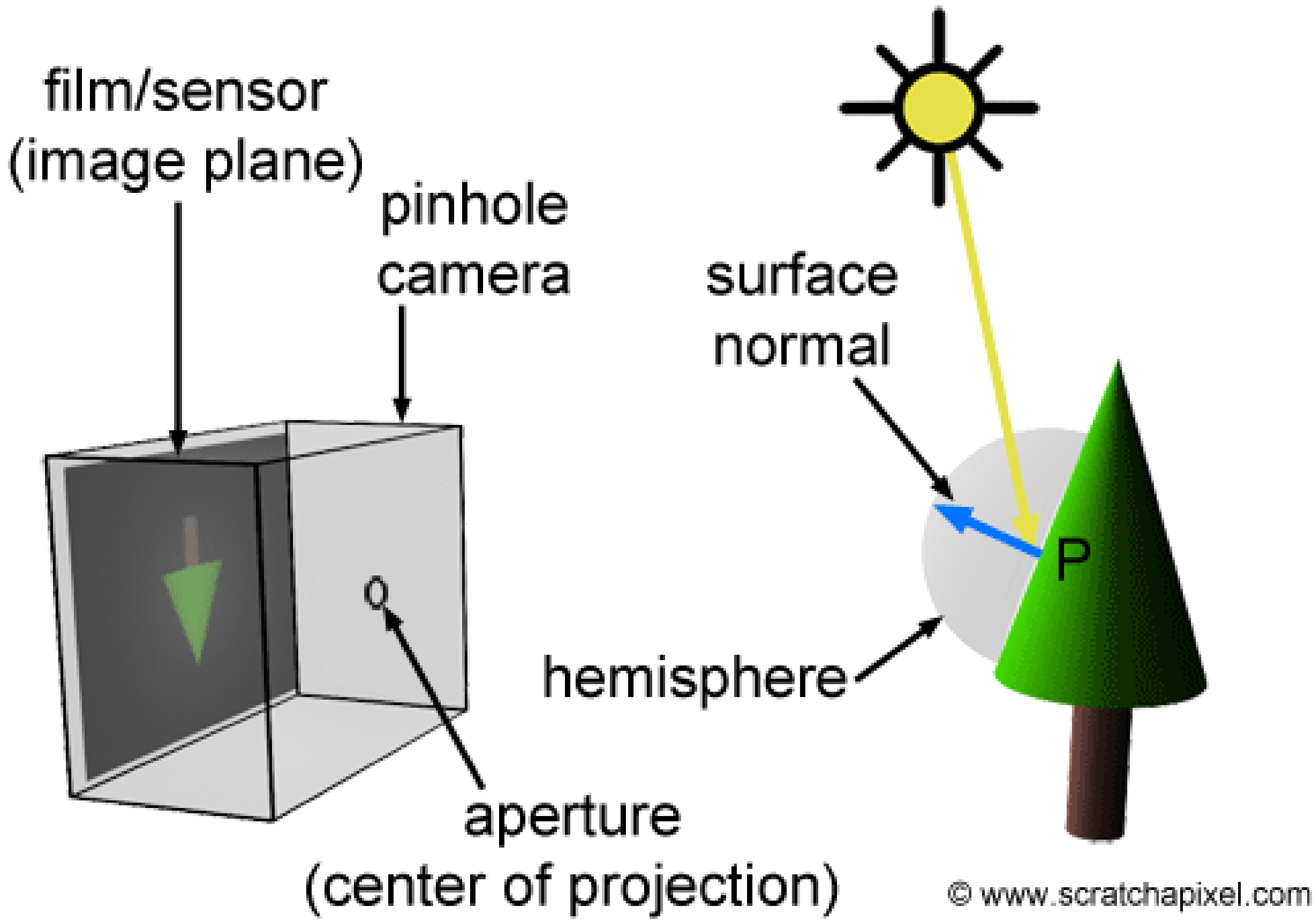
Pinhole camera principle



[1]

[1] http://en.wikipedia.org/wiki/Pinhole_camera

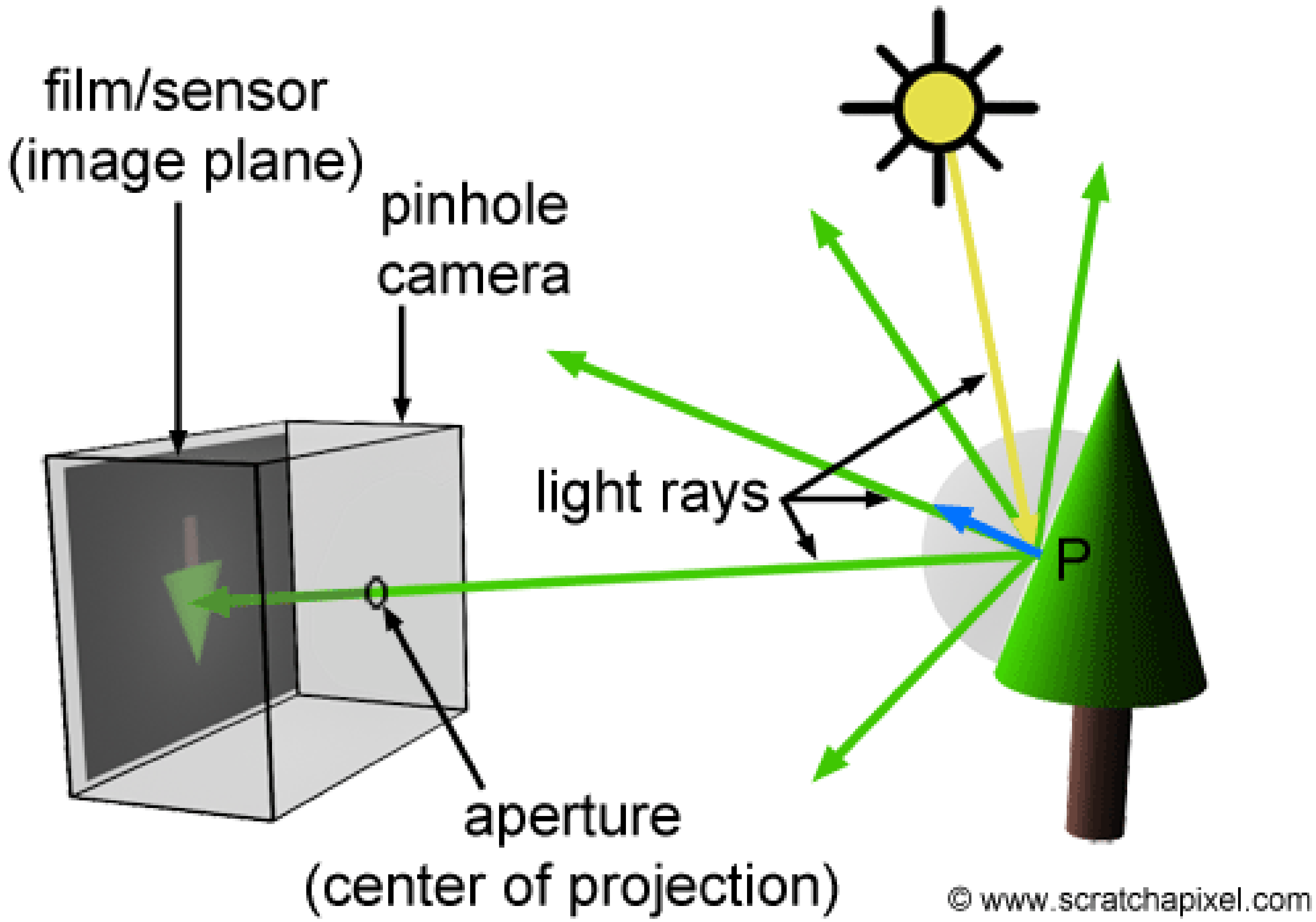
Pinhole camera principle



[2]

[2] http://en.wikipedia.org/wiki/Pinhole_camera

Pinhole camera principle



[3]

[3] http://en.wikipedia.org/wiki/Pinhole_camera

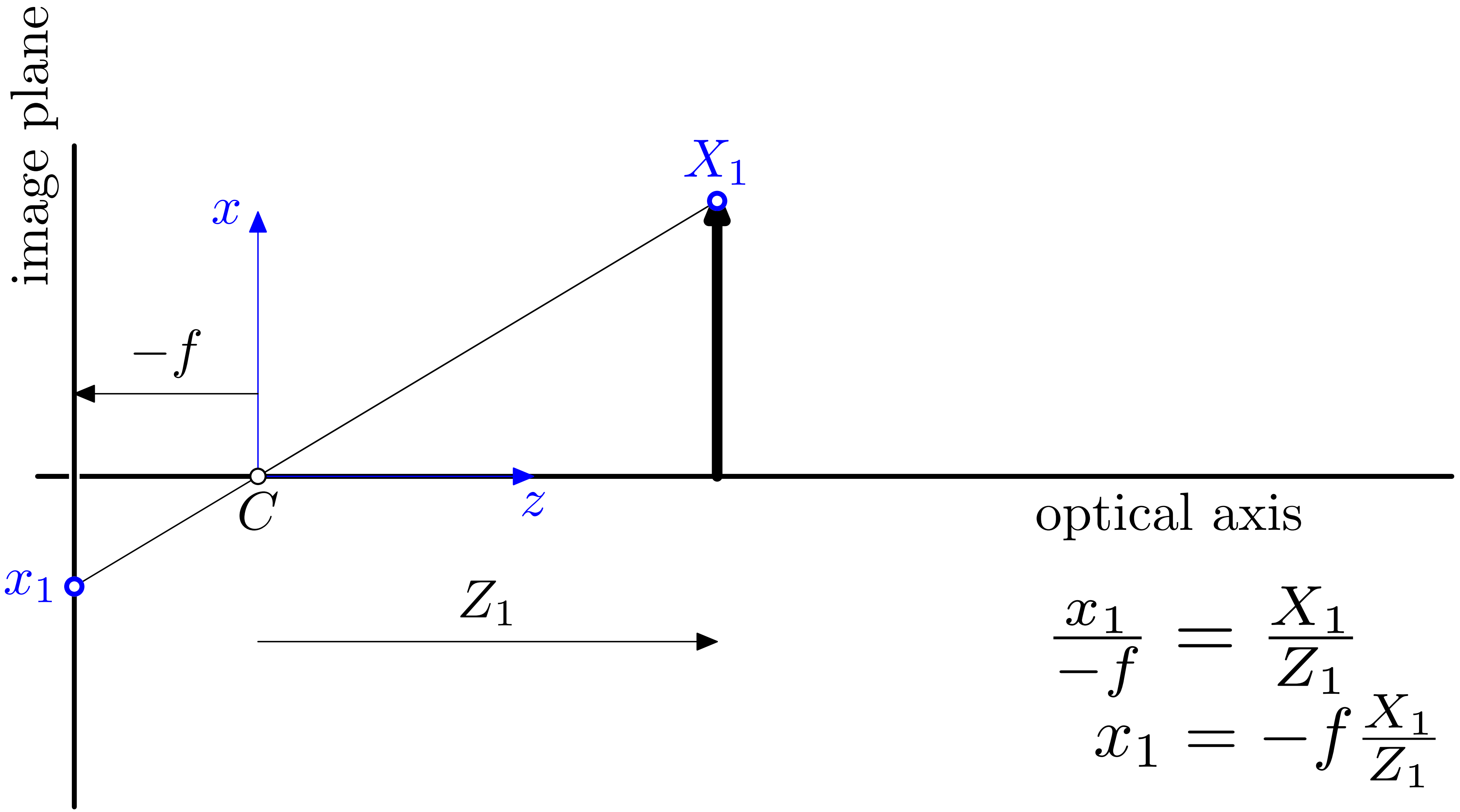
Camera Obscura — room-sized

[4]

Used by the art department at the UNC at Chapel Hill

[4] http://en.wikipedia.org/wiki/Camera_obscura

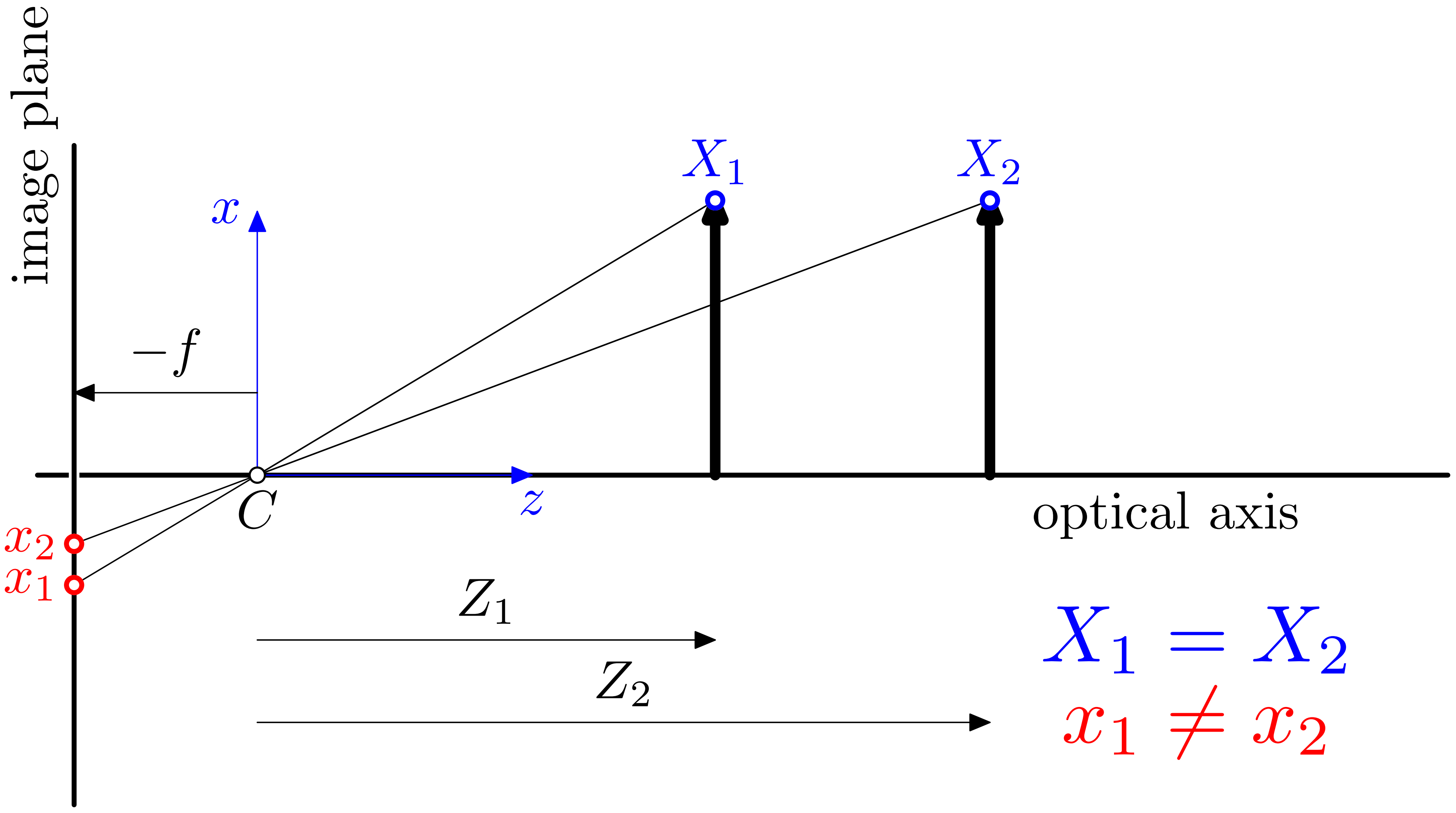
1D Pinhole camera projects 2D to 1D



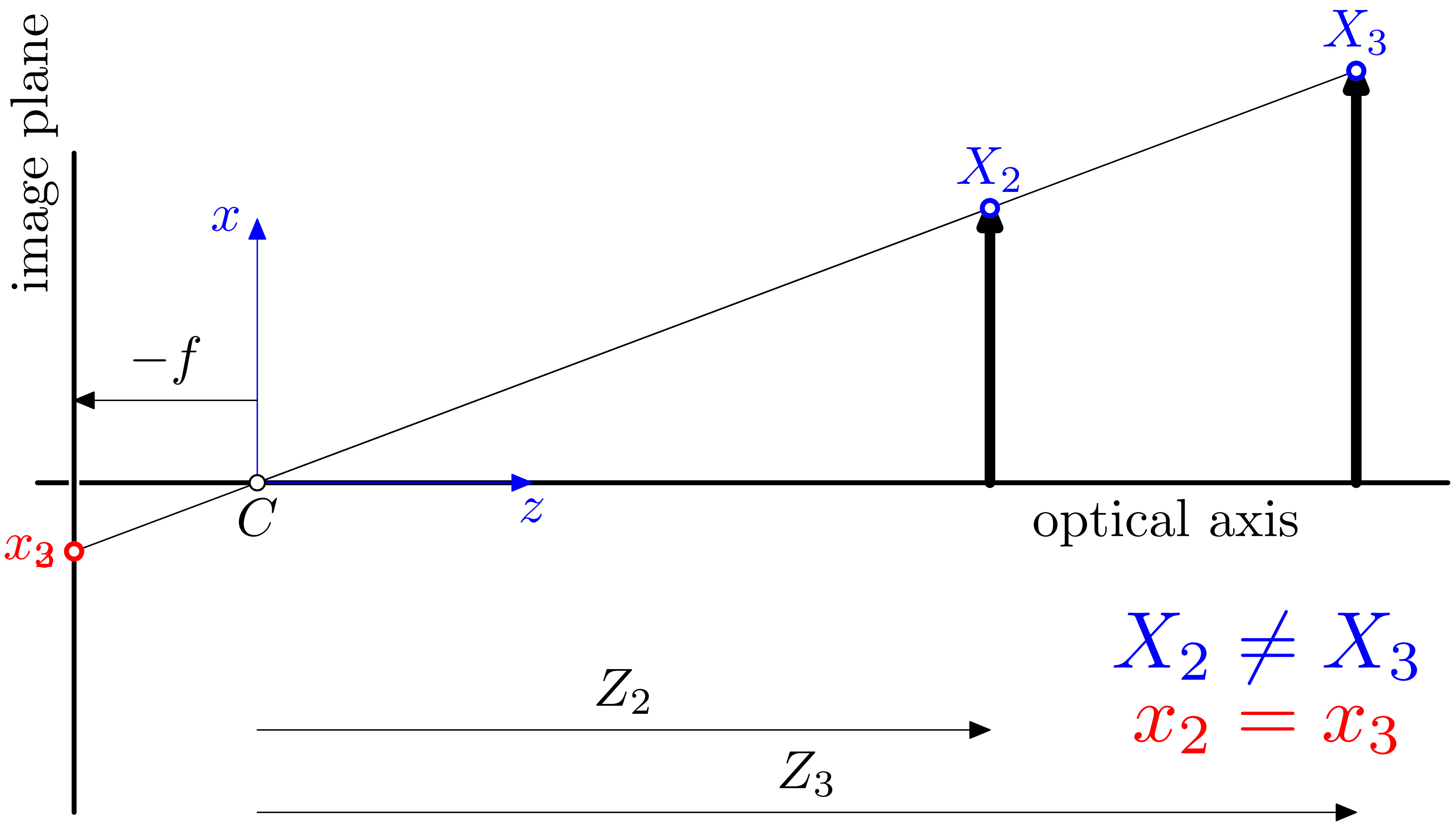
$$\frac{x_1}{-f} = \frac{X_1}{Z_1}$$

$$x_1 = -f\,\frac{X_1}{Z_1}$$

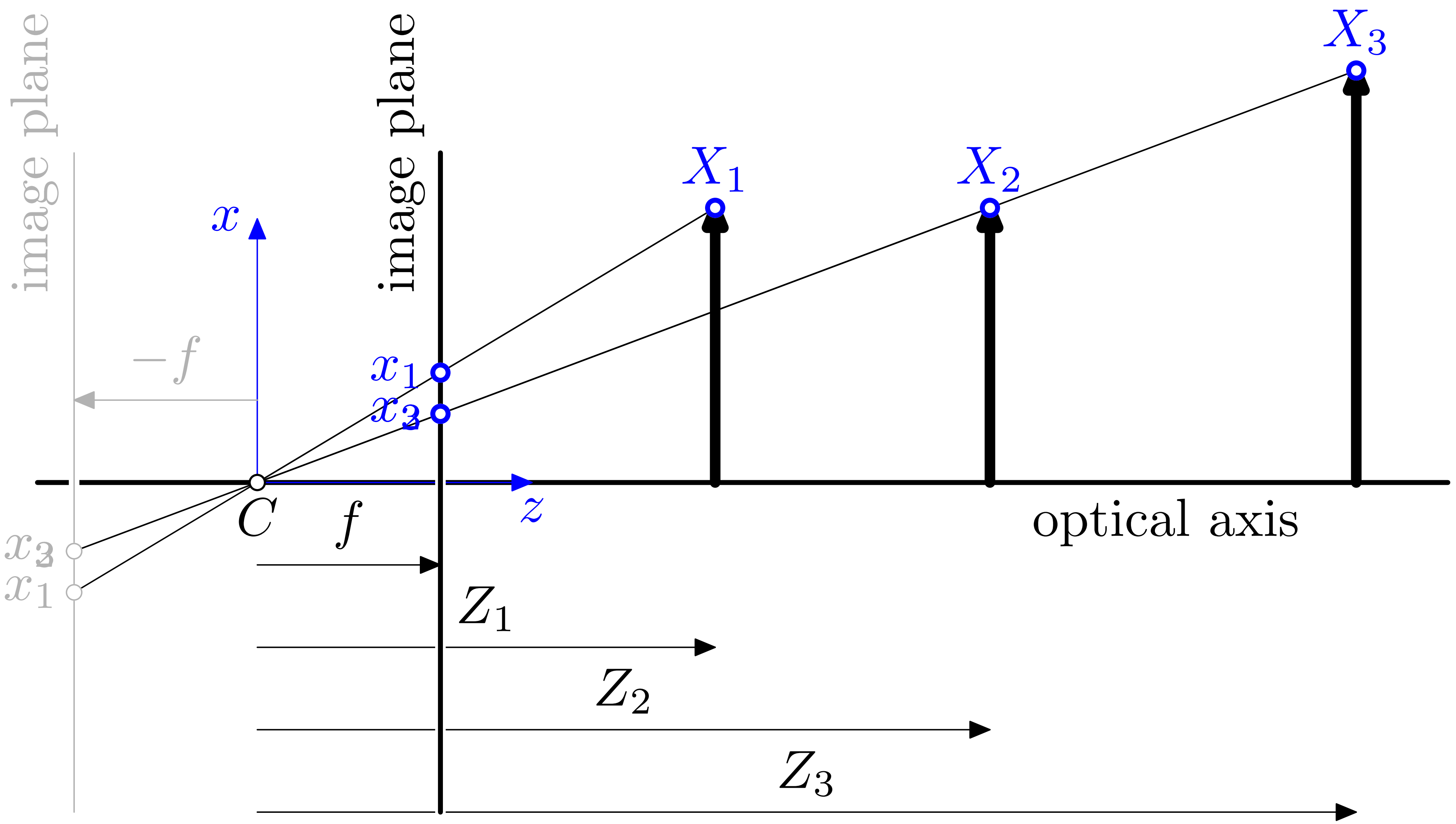
Distant objects are smaller



1D Pinhole camera projects 2D to 1D



1D Pinhole camera projects 2D to 1D

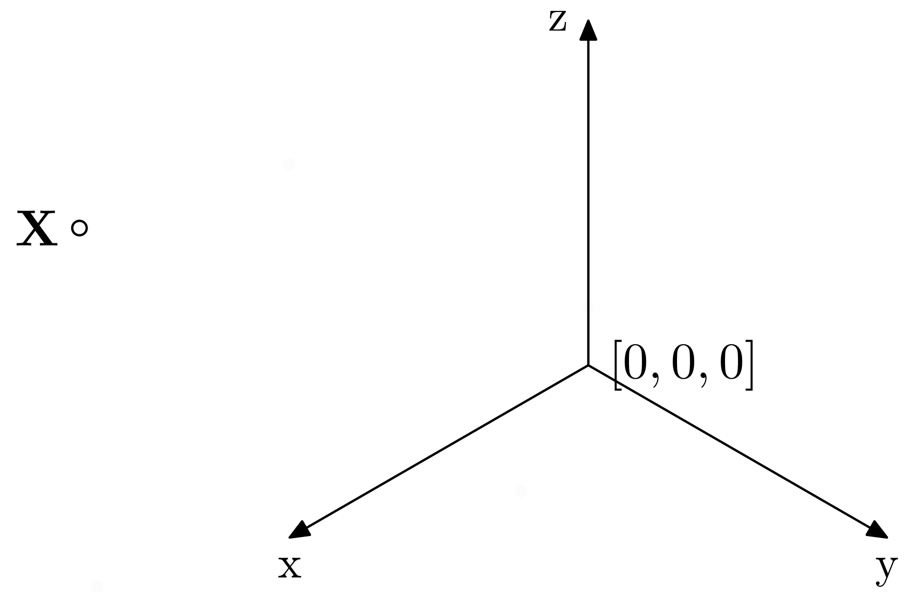


- We move image plane in front of C to get rid of $(-)$ sign.

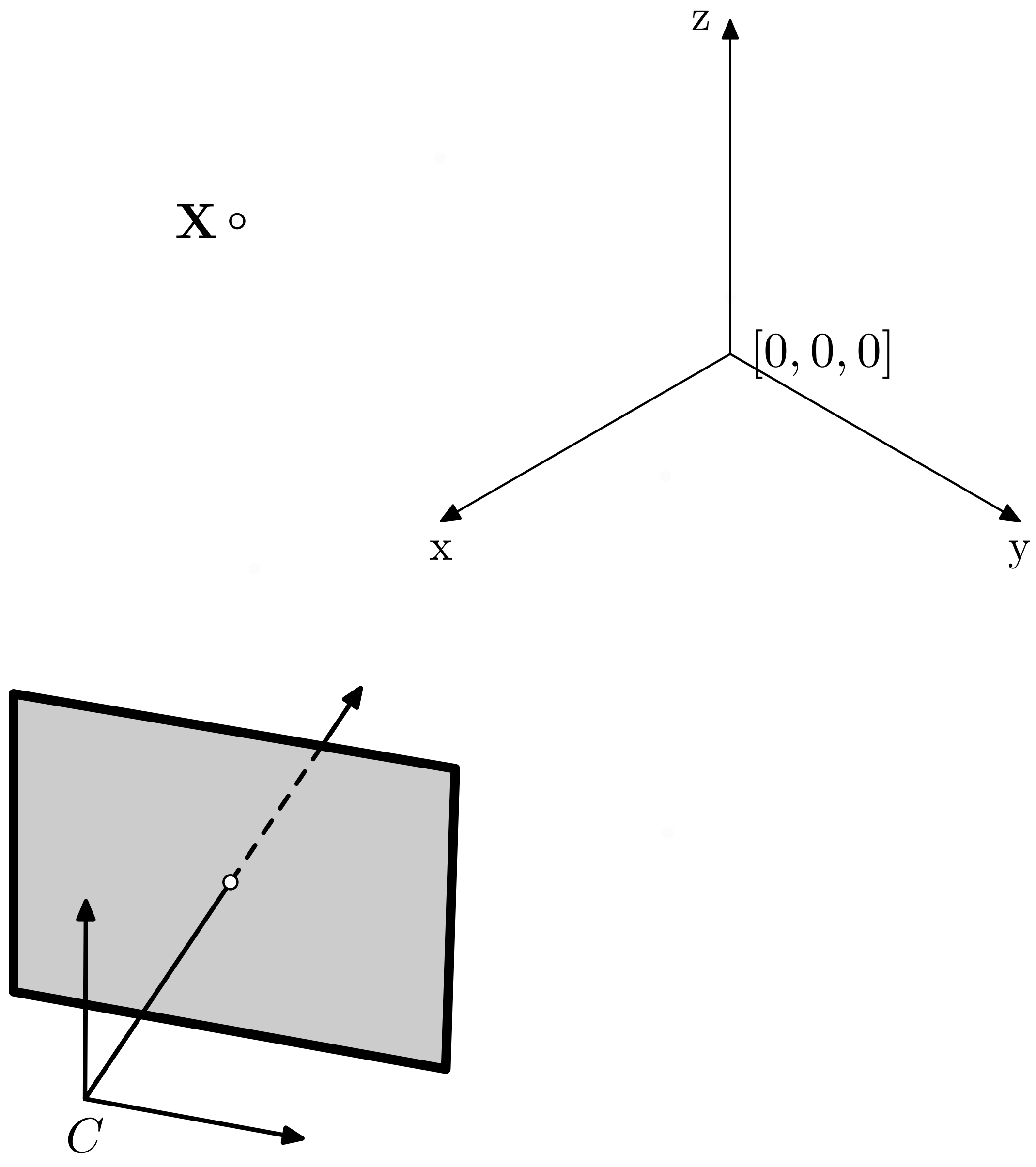
$$x = f\frac{X}{Z}$$

How does the 3D world
project to the 2D image plane?

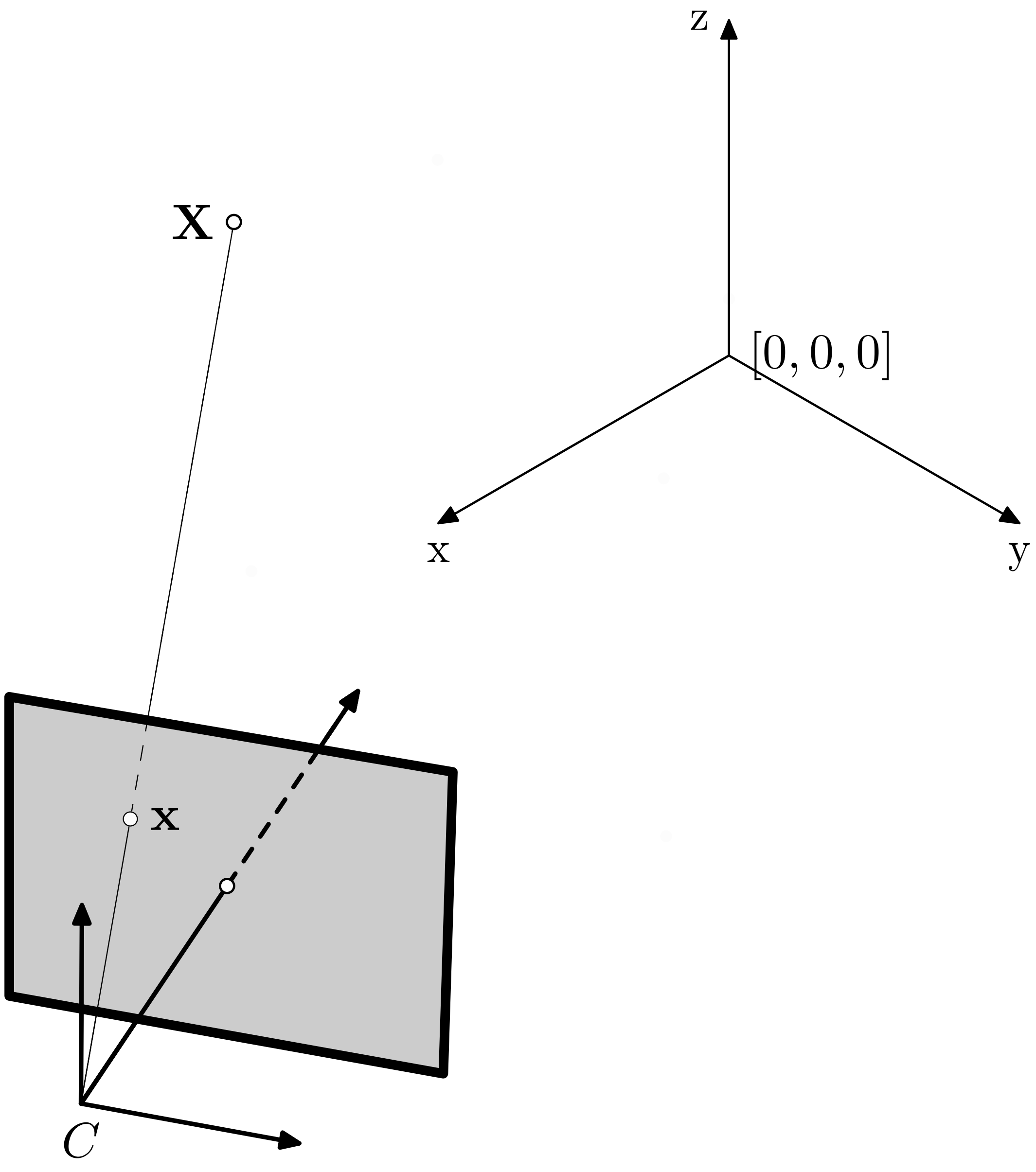
A 3D point $\mathbf{X}$ in a world coordinate system



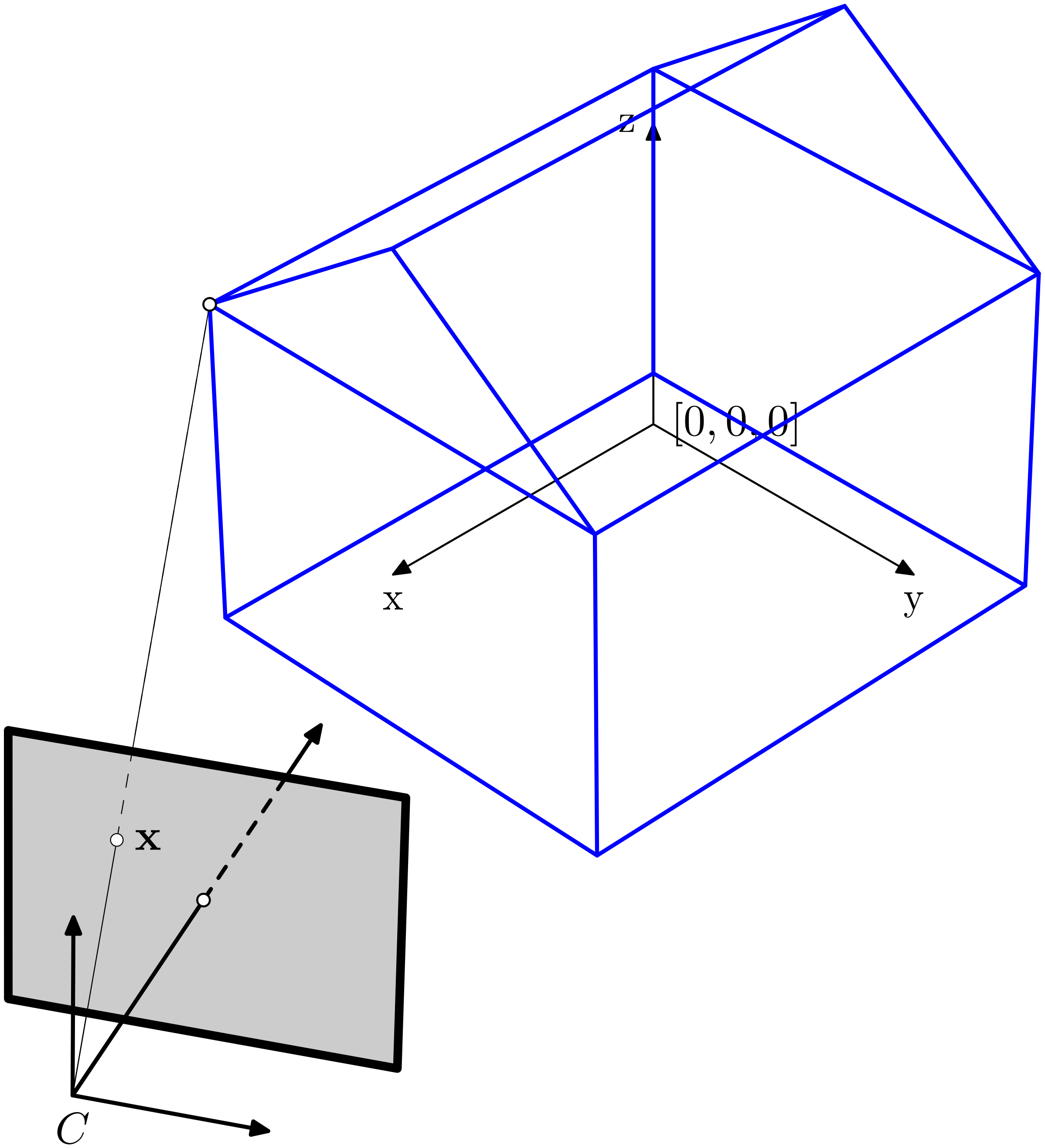
A pinhole camera observes a scene



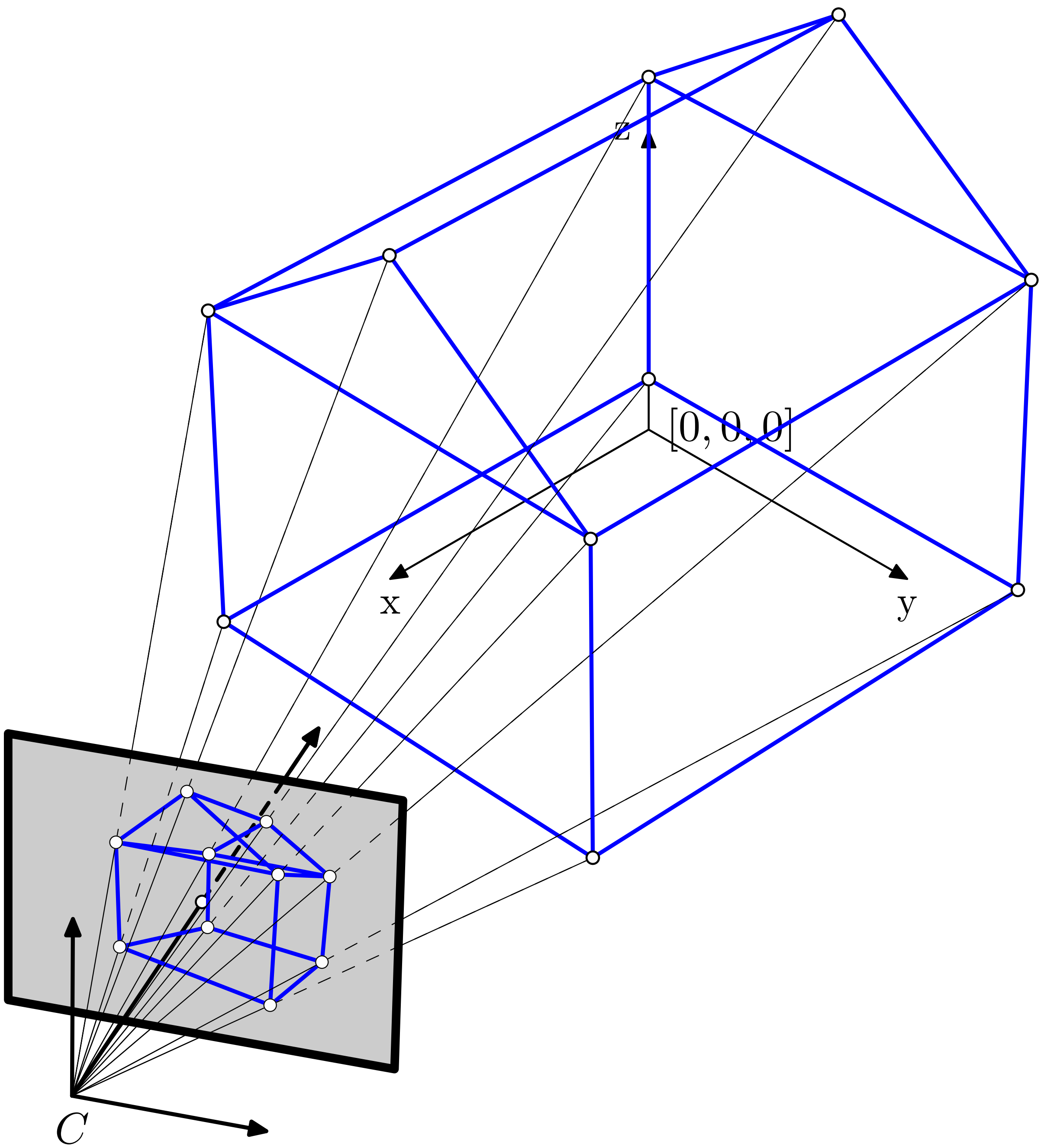
Point $\mathbf{X}$ projects to the image plane, point $\mathbf{x}$



Scene projection



Scene projection



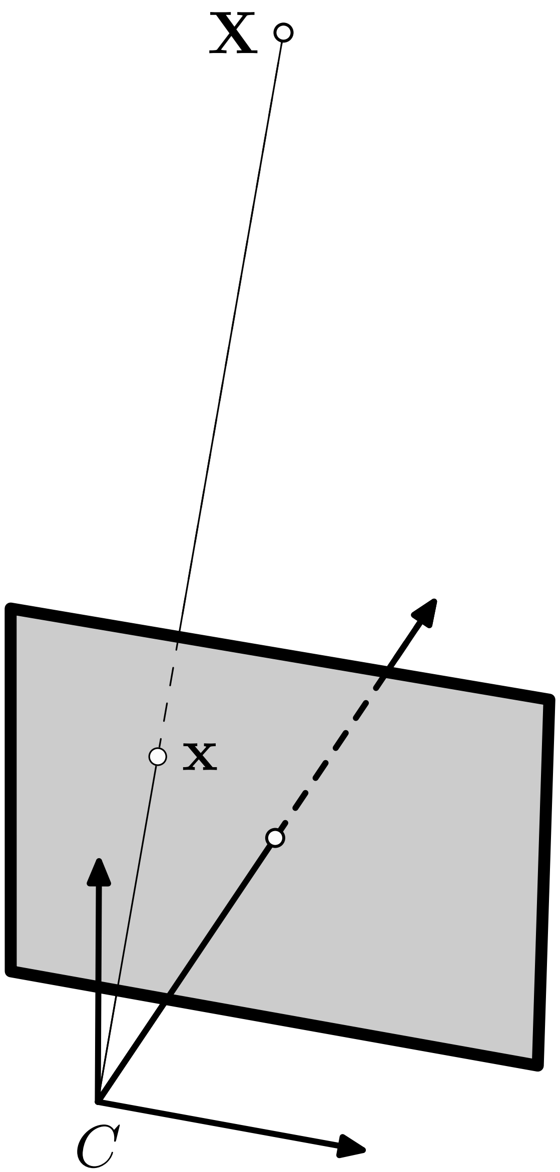
3D → 2D Projection

We remember that:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f\frac{X}{Z} \\ \\ f\frac{Y}{Z} \end{bmatrix}$$

$$\begin{bmatrix} Zx \\ Zy \end{bmatrix} = \begin{bmatrix} fX \\ fY \end{bmatrix}$$

$$\begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix} = \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}$$



3D $\longrightarrow$ 2D Projection

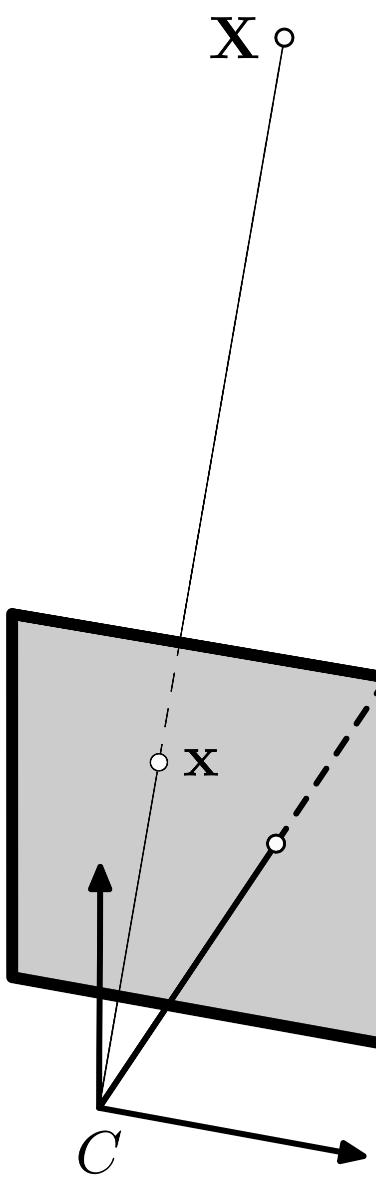
Homogeneous coordinates:
$\mathbf{x} = [x\ y\ 1]^\top$ and $\mathbf{X} = [X\ Y\ Z\ 1]^\top$ yield

$$\begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix} = \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}$$

$$\lambda \mathbf{x} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \mathbf{X}$$

$$\lambda_{[1\times1]} \mathbf{x}_{[3\times3]} = \mathtt{K}_{[3\times3]} \left[\ \mathtt{I} | \mathbf{0}\ \right] \mathbf{X}_{[4\times1]}$$

but . . .



[4]for the notation conventions, see the talk notes

Transform $\mathbf{X}$ into the camera coordinate system

Rotate the vector:

$$\mathbf{X}^e = \mathtt{R}_w^\top(\mathbf{X}_w^e - \mathbf{C}_w)$$

Use homogeneous coordinates to get a matrix equation

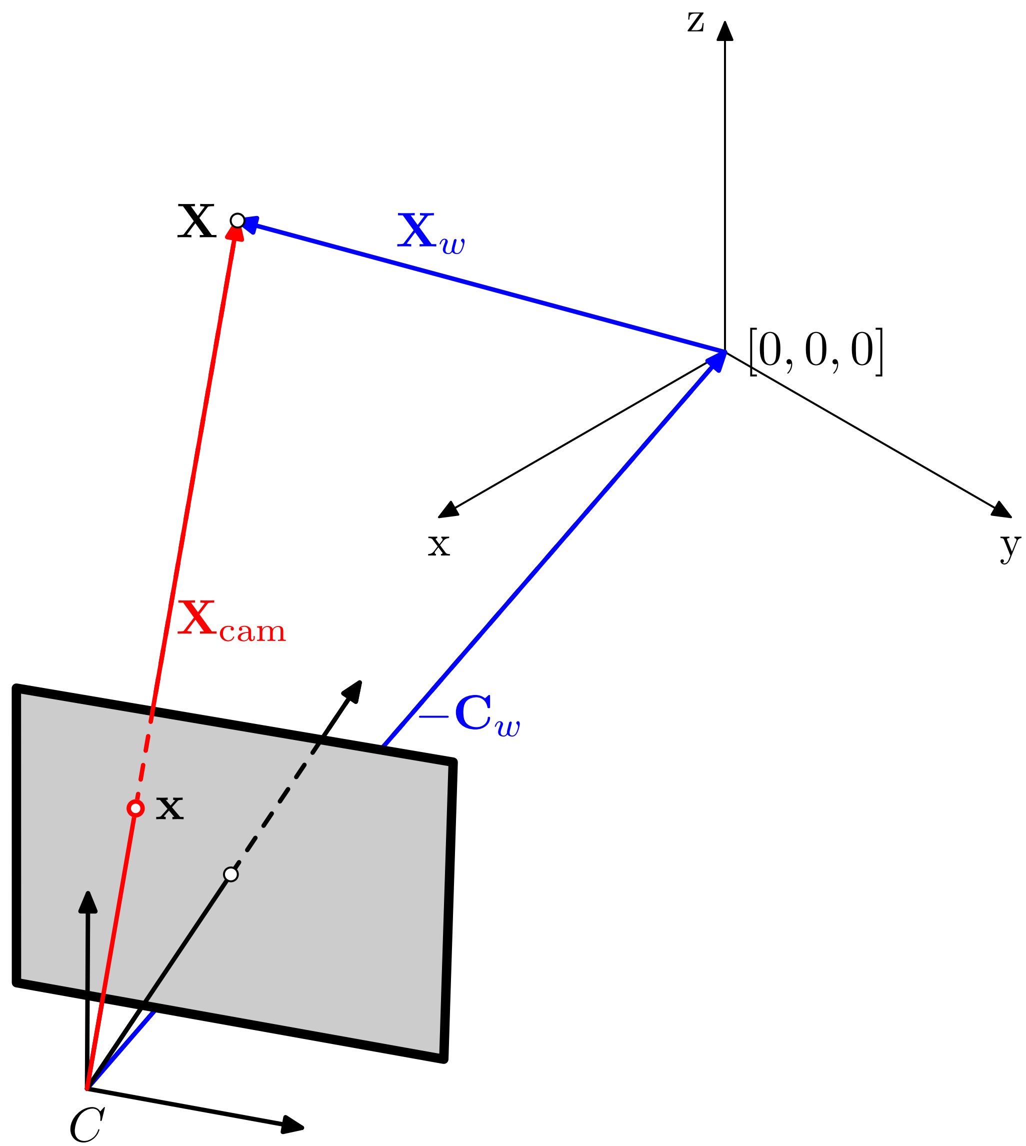
$$\mathbf{X} = \begin{bmatrix} \mathtt{R}_w^\top & -\mathtt{R}_w^\top \mathbf{C}_w \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{X}_w$$

Translation of the world in the camera

$$\mathbf{t} = -\mathtt{R}_w^\top \mathbf{C}_w$$

Rotation of the world in the camera

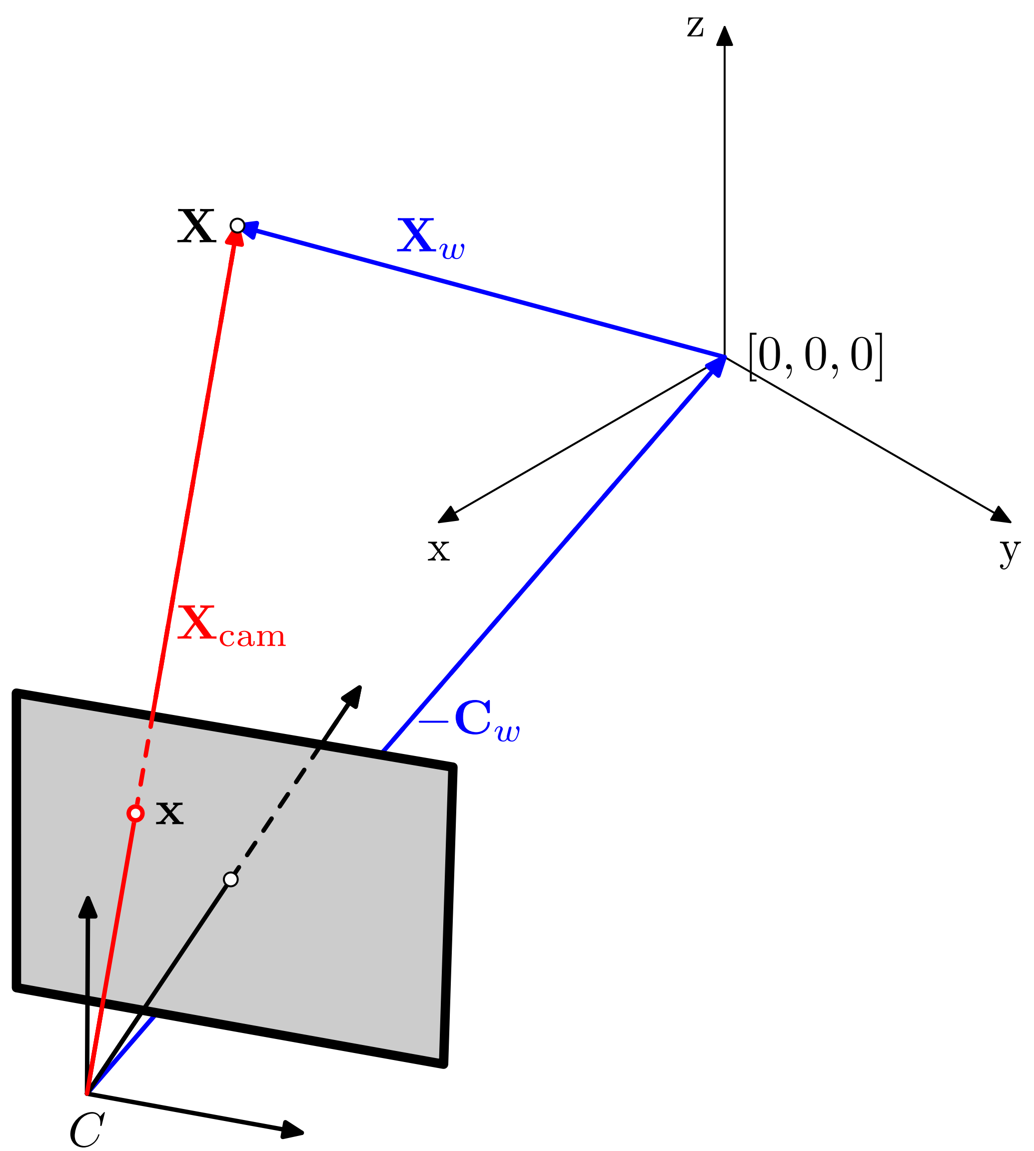
$$\mathtt{R} = \mathtt{R}_w^\top$$



Camera matrix

- $\mathtt{t}$ and $\mathtt{R}$ are called External parameters of the camera.
- The matrix $\mathtt{K}$ is called Internal parameters of the camera.

$$\begin{aligned} \lambda \mathbf{x} &= \mathtt{K}\mathbf{X} = \\ &= \mathtt{K} \begin{bmatrix} \mathtt{R} & \mathbf{t} \end{bmatrix} \mathbf{X}_w = \\ &= \mathtt{P}\mathbf{X}_w \end{aligned}$$

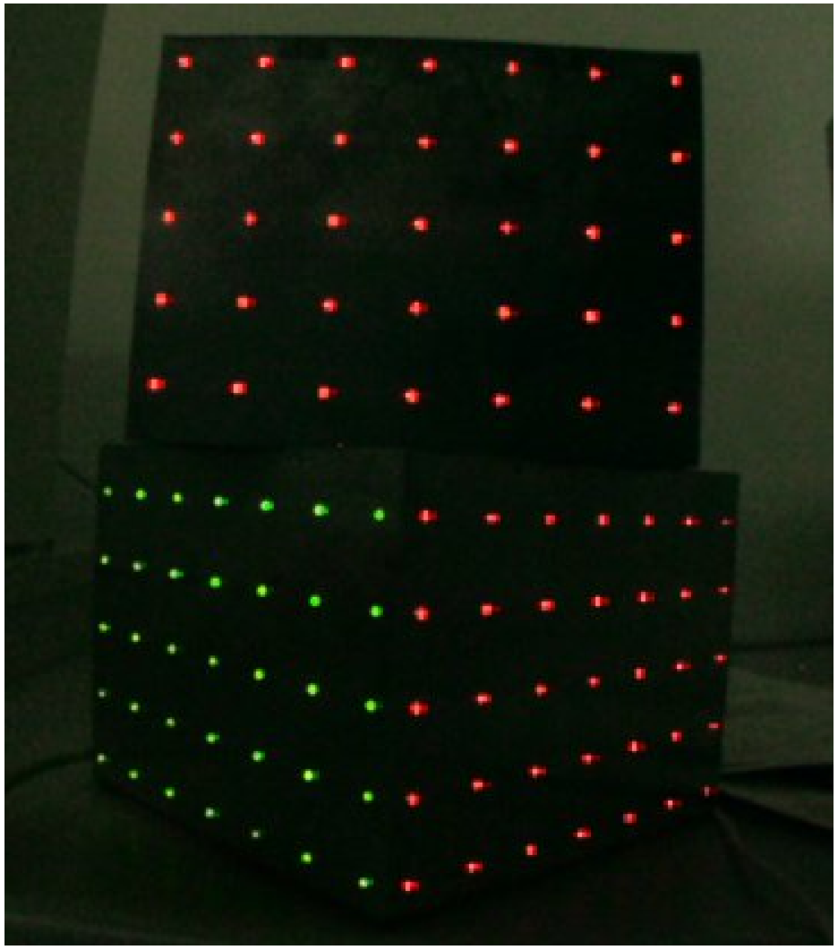


- We omit the world index and write simply $\lambda \mathbf{x} = \mathtt{P}\mathbf{X}$

Estimation of camera parameters—camera calibration

The goal: estimate the 3×4 camera projection matrix $\mathtt{P}$ from calibration object.

Assume: known correspondence between 2D camera coordinates $[u, v]^\top$ and 3D point $[X\ Y\ Z]^\top$ with known coordinates



Estimation of camera parameters—camera calibration

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1^\top \\ \mathbf{p}_2^\top \\ \mathbf{p}_3^\top \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\frac{\lambda u}{\lambda} = \frac{\mathbf{p}_1^\top \mathbf{X}}{\mathbf{p}_3^\top \mathbf{X}} \quad \text{and} \quad \frac{\lambda v}{\lambda} = \frac{\mathbf{p}_2^\top \mathbf{X}}{\mathbf{p}_3^\top \mathbf{X}}$$

Re-arrange and assume $\lambda \neq 0$ to get set of homegeneous equations

$$\begin{aligned} u\mathbf{X}^\top \mathbf{p}_3 - \mathbf{X}^\top \mathbf{p}_1 &= 0 \\ v\mathbf{X}^\top \mathbf{p}_3 - \mathbf{X}^\top \mathbf{p}_2 &= 0 \end{aligned}$$

Estimation of camera parameters—camera calibration

$$
\begin{aligned}
u\mathbf{X}^\top \mathbf{p}_3 - \mathbf{X}^\top \mathbf{p}_1 &= 0 \\
v\mathbf{X}^\top \mathbf{p}_3 - \mathbf{X}^\top \mathbf{p}_2 &= 0
\end{aligned}
$$

Re-shuffle into a matrix form:

$$
\underbrace{\begin{bmatrix} -\mathbf{X}^\top & \mathbf{0}^\top & u\mathbf{X}^\top \\ \mathbf{0}^\top & -\mathbf{X}^\top & v\mathbf{X}^\top \end{bmatrix}}_{\mathtt{A}_{[2\times 12]}} \underbrace{\begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix}}_{\mathbf{p}_{[12\times 1]}} = \mathbf{0}_{[2\times 1]}
$$

A correspondece $\mathbf{u}_i \leftrightarrow \mathbf{X}_i$ forms two homogeneous equations. $\mathtt{P}$ has 12 parameters but scale does not matter. We need at least 6 2D $\leftrightarrow$ 3D pairs to get a solution. We constitute $\mathtt{A}_{[\geq 12\times 12]}$ data matrix and solve

$$
\mathbf{p}^* = \operatorname{argmin} \|\mathtt{A}\mathbf{p}\| \ \text{ subject to } \ \|\mathbf{p}\| = 1
$$

which is a constrained LSQ problem. $\mathbf{p}^*$ minimizes algebraic error

Solution of constrained LSQ problem

We solve $\mathbf{p}^* = \mathrm{argmin} \|\mathtt{A}\mathbf{p}\| \;\; \text{subject to} \;\; \|\mathbf{p}\| = 1$ by Lagrange function

$$\begin{aligned} L(\mathbf{p}, \lambda) &= \|\mathtt{A}\mathbf{p}\| + \lambda(1 - \|\mathbf{p}\|) = \\ &= \mathbf{p}^\top \mathtt{A}^\top \mathtt{A}\mathbf{p} + \lambda(1 - \mathbf{p}^\top \mathbf{p}) \end{aligned}$$

Solution of constrained LSQ problem

We solve $\mathbf{p}^* = \mathrm{argmin}\|\mathtt{A}\mathbf{p}\| \;\; \text{subject to} \;\; \|\mathbf{p}\| = 1$ by Lagrange function

$$\begin{aligned} L(\mathbf{p},\lambda) &= \|\mathtt{A}\mathbf{p}\| + \lambda(1 - \|\mathbf{p}\|) = \\ &= \mathbf{p}^\top \mathtt{A}^\top \mathtt{A}\mathbf{p} + \lambda(1 - \mathbf{p}^\top\mathbf{p}) \end{aligned}$$

Critical points:

$$\begin{aligned} \frac{\partial L(\mathbf{p},\lambda)}{\partial \mathbf{p}} &= 2\mathtt{A}^\top\mathtt{A}\mathbf{p} - 2\lambda\mathbf{p} = \mathbf{0} \\ \frac{\partial L(\mathbf{p},\lambda)}{\partial \lambda} &= 1 - \mathbf{p}^\top\mathbf{p} = \mathbf{0} \end{aligned}$$

Solution of constrained LSQ problem

We solve $\mathbf{p}^* = \mathrm{argmin} \|\mathtt{A}\mathbf{p}\| \;\; \text{subject to} \;\; \|\mathbf{p}\| = 1$ by Lagrange function

$$\begin{aligned} L(\mathbf{p}, \lambda) &= \|\mathtt{A}\mathbf{p}\| + \lambda(1 - \|\mathbf{p}\|) = \\ &= \mathbf{p}^\top \mathtt{A}^\top \mathtt{A}\mathbf{p} + \lambda(1 - \mathbf{p}^\top \mathbf{p}) \end{aligned}$$

Critical points:

$$\begin{aligned} \frac{\partial L(\mathbf{p}, \lambda)}{\partial \mathbf{p}} &= 2\mathtt{A}^\top \mathtt{A}\mathbf{p} - 2\lambda \mathbf{p} = \mathbf{0} \\ \frac{\partial L(\mathbf{p}, \lambda)}{\partial \lambda} &= 1 - \mathbf{p}^\top \mathbf{p} = \mathbf{0} \end{aligned}$$

First equation is characteristic equation $(\mathtt{A}^\top \mathtt{A} - \lambda \mathtt{I})\mathbf{p} = \mathbf{0}$, every eigen-vector $\mathbf{p}$ of $\mathtt{A}^\top \mathtt{A}$ with unit length is critical point.

Solution of constrained LSQ problem

We solve $\mathbf{p}^* = \mathrm{argmin}\|\mathtt{A}\mathbf{p}\| \ \text{ subject to } \ \|\mathbf{p}\| = 1$ by Lagrange function

$$\begin{aligned} L(\mathbf{p}, \lambda) &= \|\mathtt{A}\mathbf{p}\| + \lambda(1 - \|\mathbf{p}\|) = \\ &= \mathbf{p}^\top \mathtt{A}^\top \mathtt{A}\mathbf{p} + \lambda(1 - \mathbf{p}^\top \mathbf{p}) \end{aligned}$$

Critical points:

$$\begin{aligned} \frac{\partial L(\mathbf{p}, \lambda)}{\partial \mathbf{p}} &= 2\mathtt{A}^\top \mathtt{A}\mathbf{p} - 2\lambda\mathbf{p} = \mathbf{0} \\ \frac{\partial L(\mathbf{p}, \lambda)}{\partial \lambda} &= 1 - \mathbf{p}^\top \mathbf{p} = \mathbf{0} \end{aligned}$$

First equation is characteristic equation $(\mathtt{A}^\top \mathtt{A} - \lambda \mathtt{I})\mathbf{p} = \mathbf{0}$, every eigen-vector $\mathbf{p}$ of $\mathtt{A}^\top \mathtt{A}$ with unit length is critical point.

Since cost function $\|\mathtt{A}\mathbf{p}\|$ in these eigen-vectors is equal to their eigen-values $\|\mathtt{A}\mathbf{p}\| = \mathbf{p}^\top \mathtt{A}^\top \mathtt{A}\mathbf{p} = \mathbf{p}^\top \lambda \mathbf{p} = \lambda \mathbf{p}^\top \mathbf{p} = \lambda\|\mathbf{p}\| = \lambda.$ the solution is the eigen-vector of $\mathtt{A}^\top \mathtt{A}$ with the smallest eigen-value.

Decomposition of $\mathtt{P}$ into the calibration parameters

$$\mathtt{P} = \left[\begin{array}{cc} \mathtt{KR} & \mathtt{K}\mathbf{t} \end{array}\right] \quad \text{and} \quad \mathbf{C} = -\mathtt{R}^{-1}\mathbf{t}$$

We know that $\mathtt{R}$ should be 3×3 orthonormal, and $\mathtt{K}$ upper triangular.

```
P = P./norm(P(3,1:3));

[K,R] = rq(P(:,1:3));
t = inv(K)*P(:,4);
C = -R'*t;
```

References

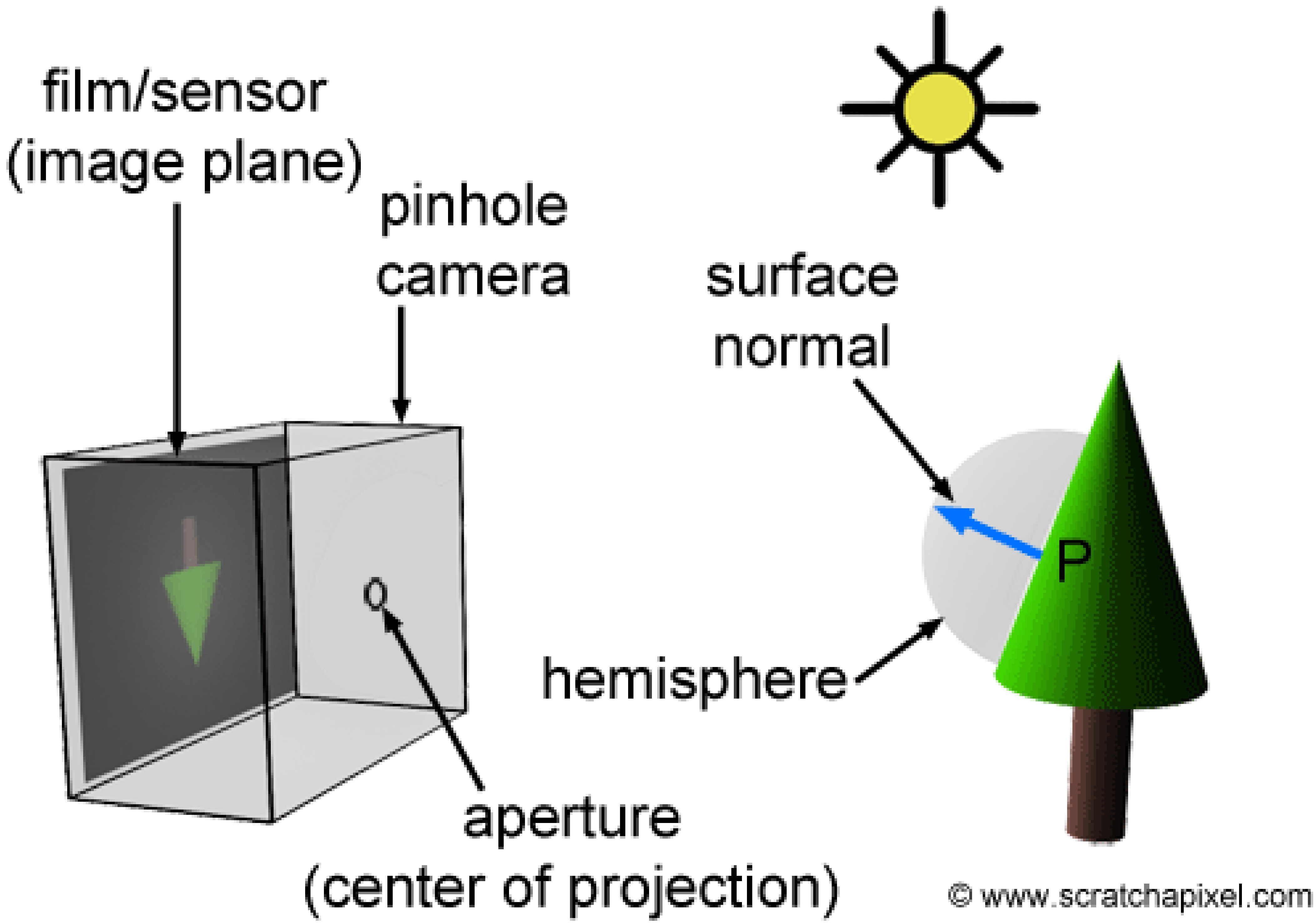
The book [2] is the ultimate reference. It is a must read for anyone wanting use cameras for 3D computing.

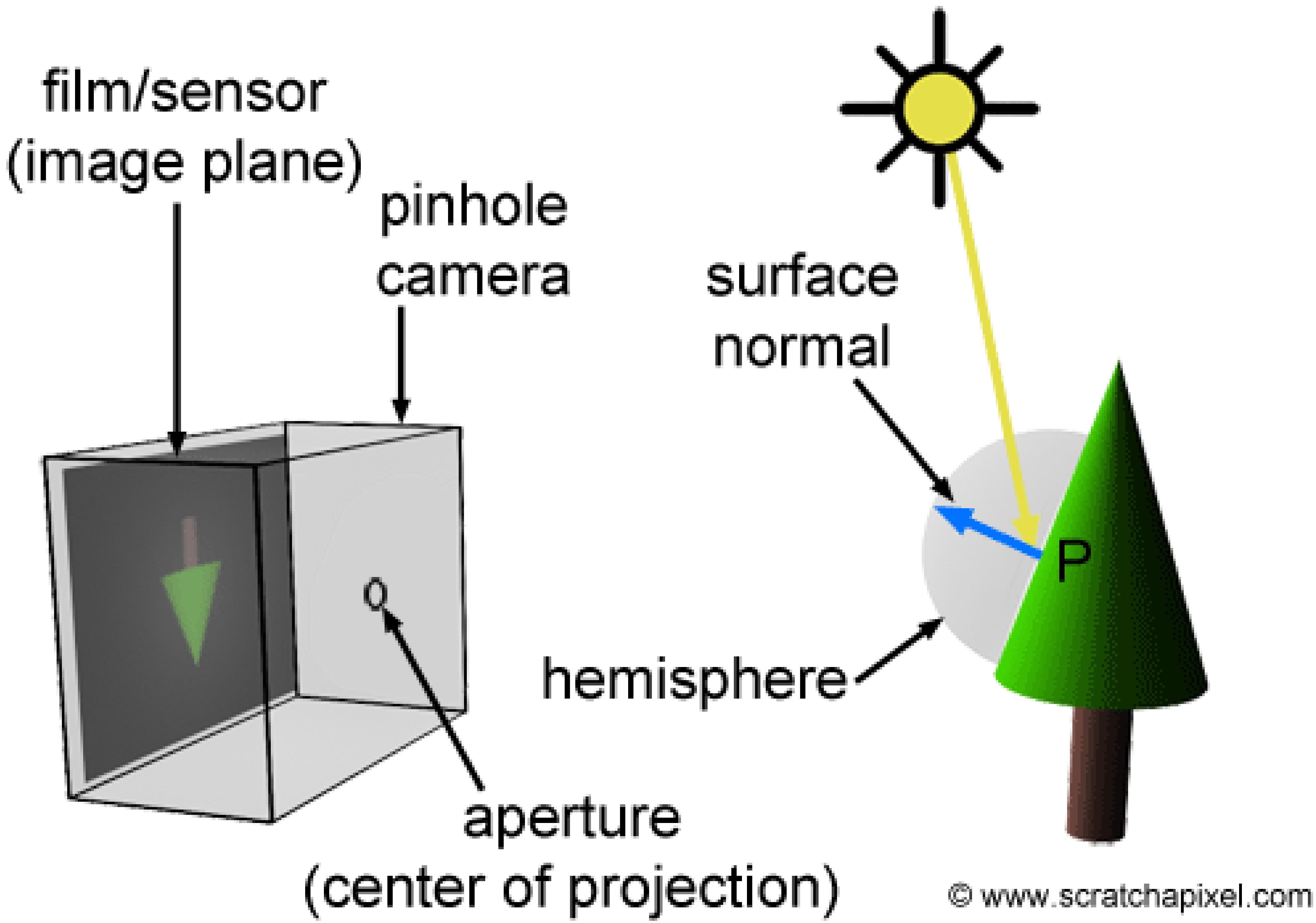
Details about matrix decompositions used throughout the lecture can be found at [1]

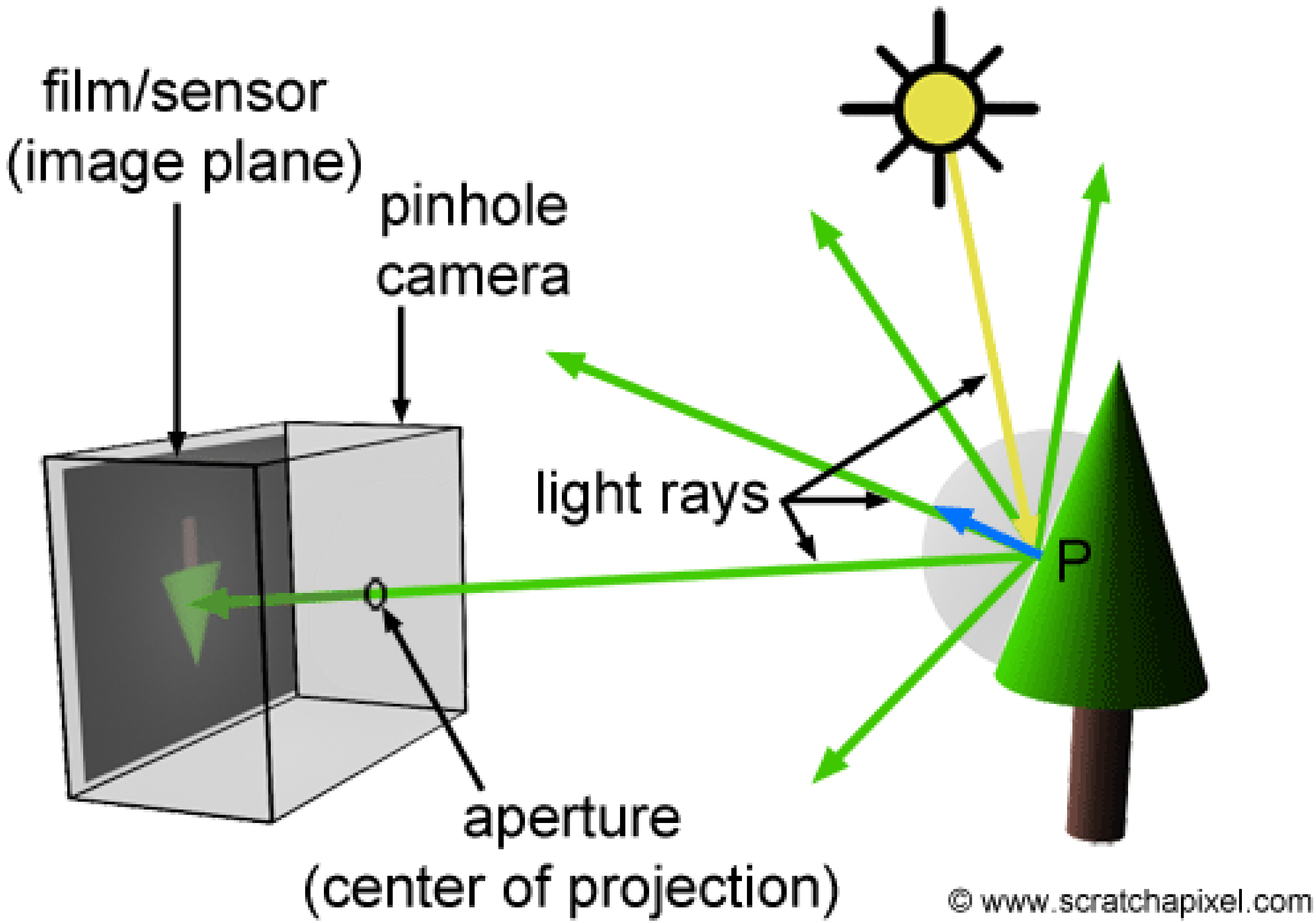
[1] Gene H. Golub and Charles F. Van Loan. Matrix Computation. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, USA, 3rd edition, 1996.

[2] Richard Hartley and Andrew Zisserman. Multiple view geometry in computer vision. Cambridge University, Cambridge, 2nd edition, 2003.

End


film/sensor
(image plane)
pinhole
camera
surface
normal
P
hemisphere
aperture
(center of projection)
© www.scratchapixel.com


film/sensor
(image plane)
pinhole
camera
surface
normal
P
hemisphere
aperture
(center of projection)
© www.scratchapixel.com


film/sensor
(image plane)
pinhole
camera
light rays
P
aperture
(center of projection)
© www.scratchapixel.com

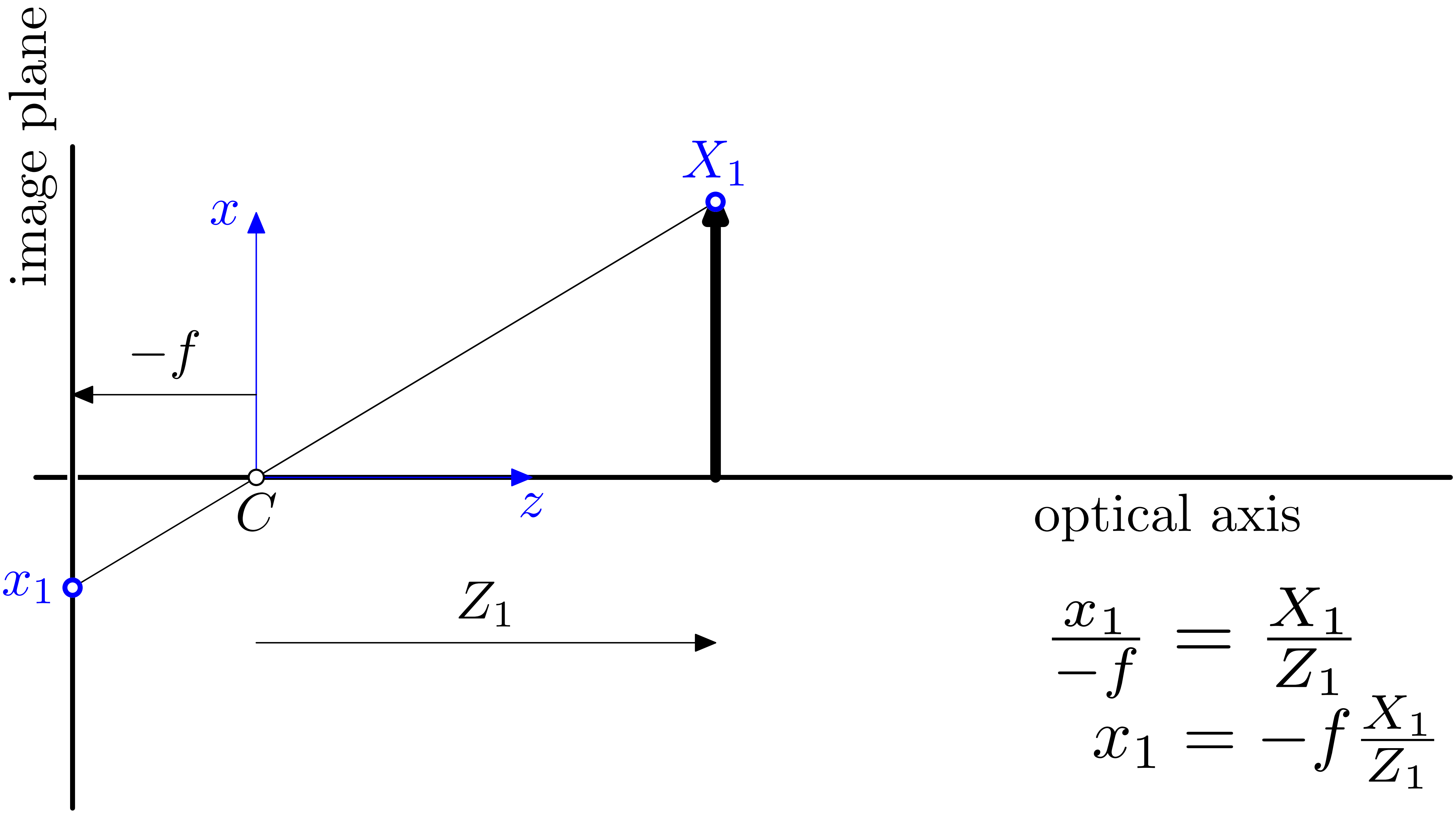


image plane
X_1
x
$-f$
C
z
optical axis
x_1
Z_1
$\frac{x_1}{-f} = \frac{X_1}{Z_1}$
$x_1 = -f\,\frac{X_1}{Z_1}$

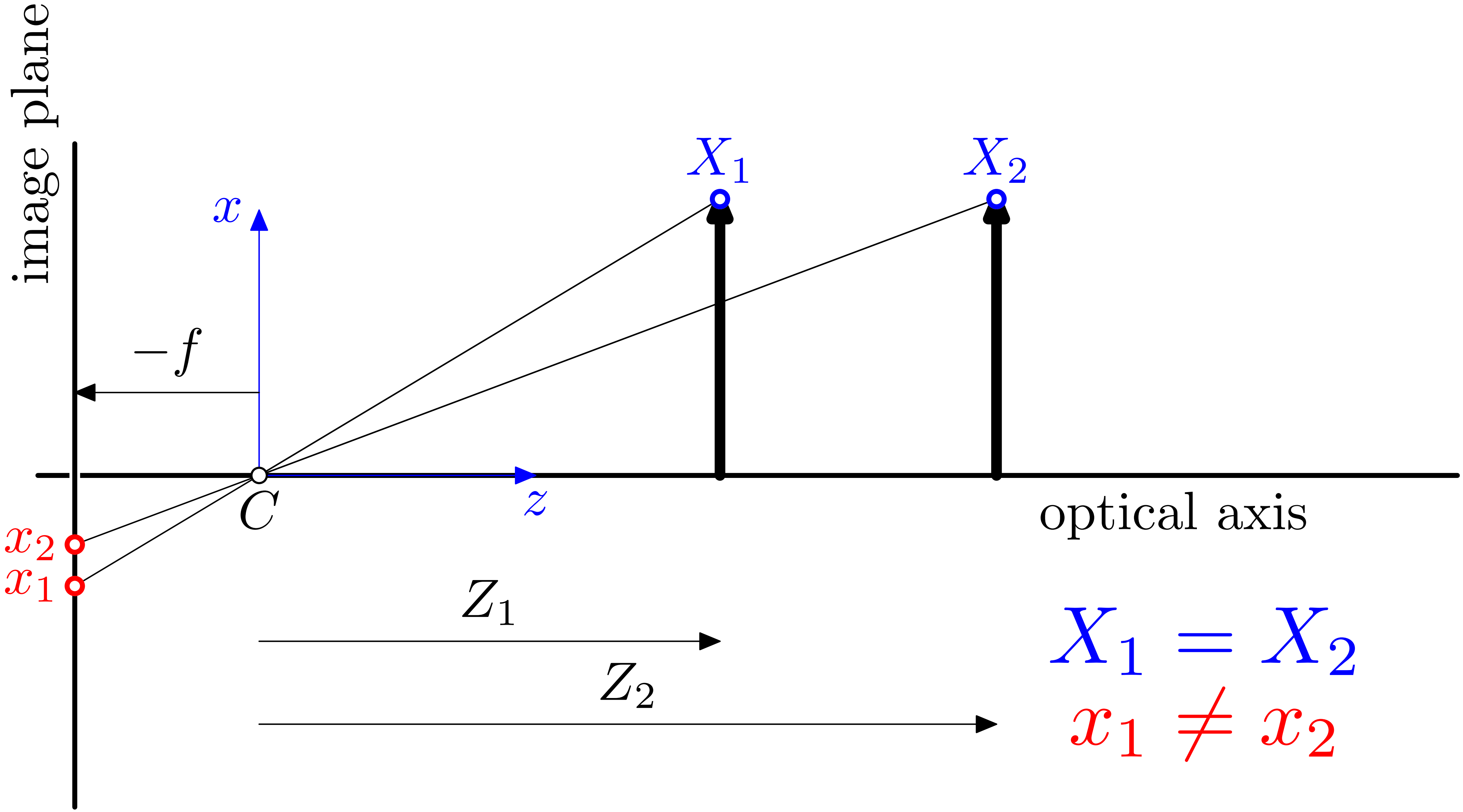


image plane
X_1
X_2
x
$-f$
C
z
optical axis
x_2
x_1
Z_1
Z_2
$X_1 = X_2$
$x_1 \neq x_2$

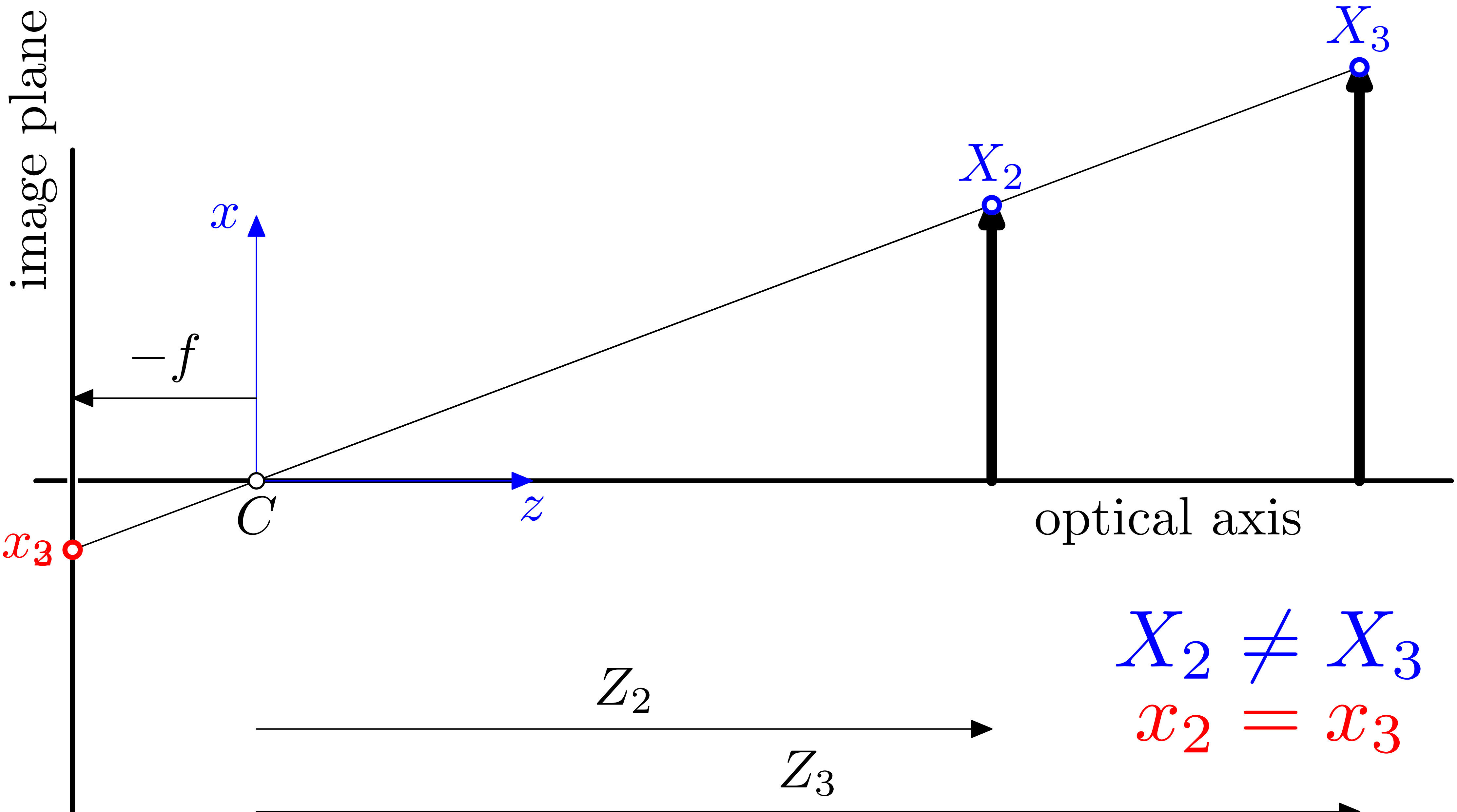


image plane
X_3
X_2
x
$-f$
C
z
optical axis
x_3
x_2
$X_2 \neq X_3$
$x_2 = x_3$
Z_2
Z_3

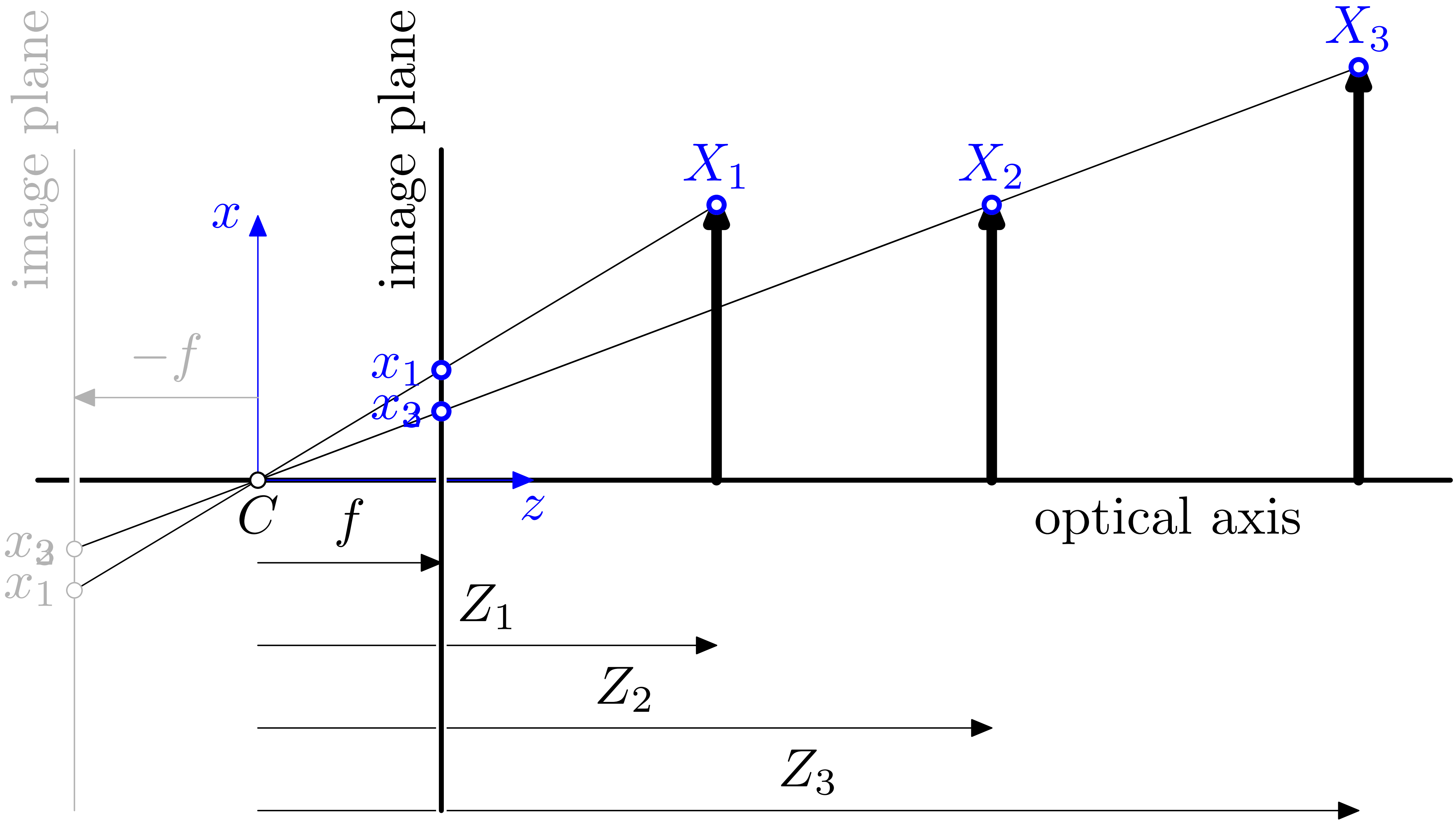
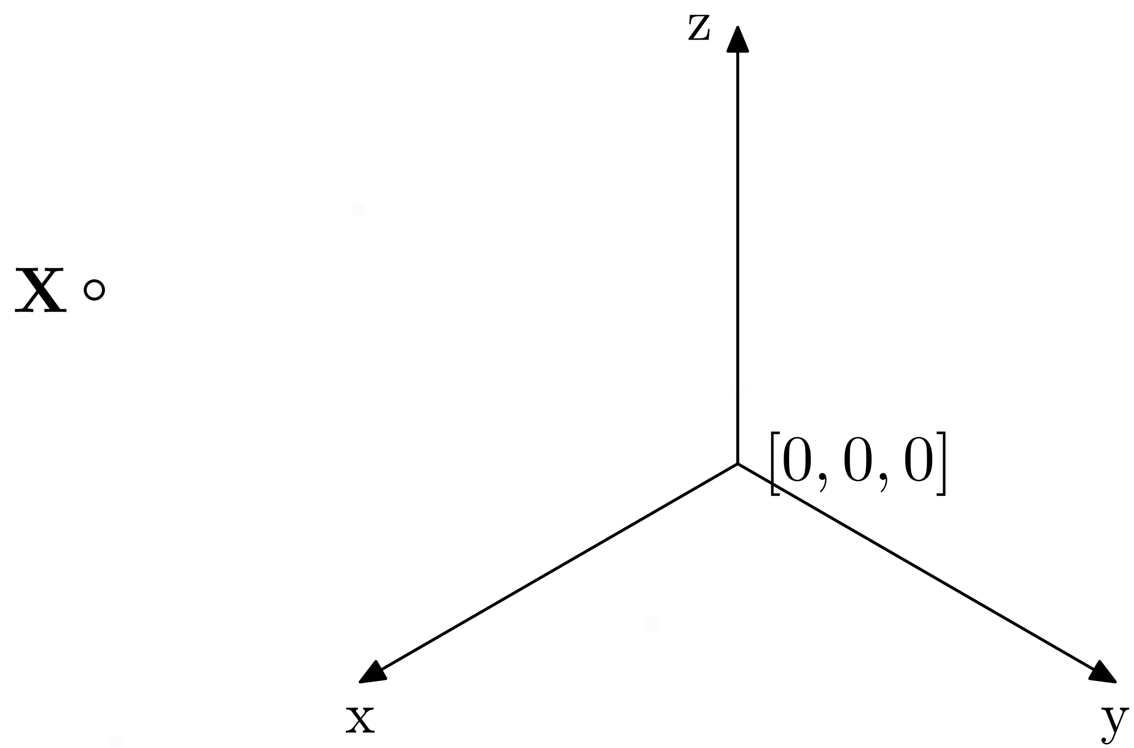
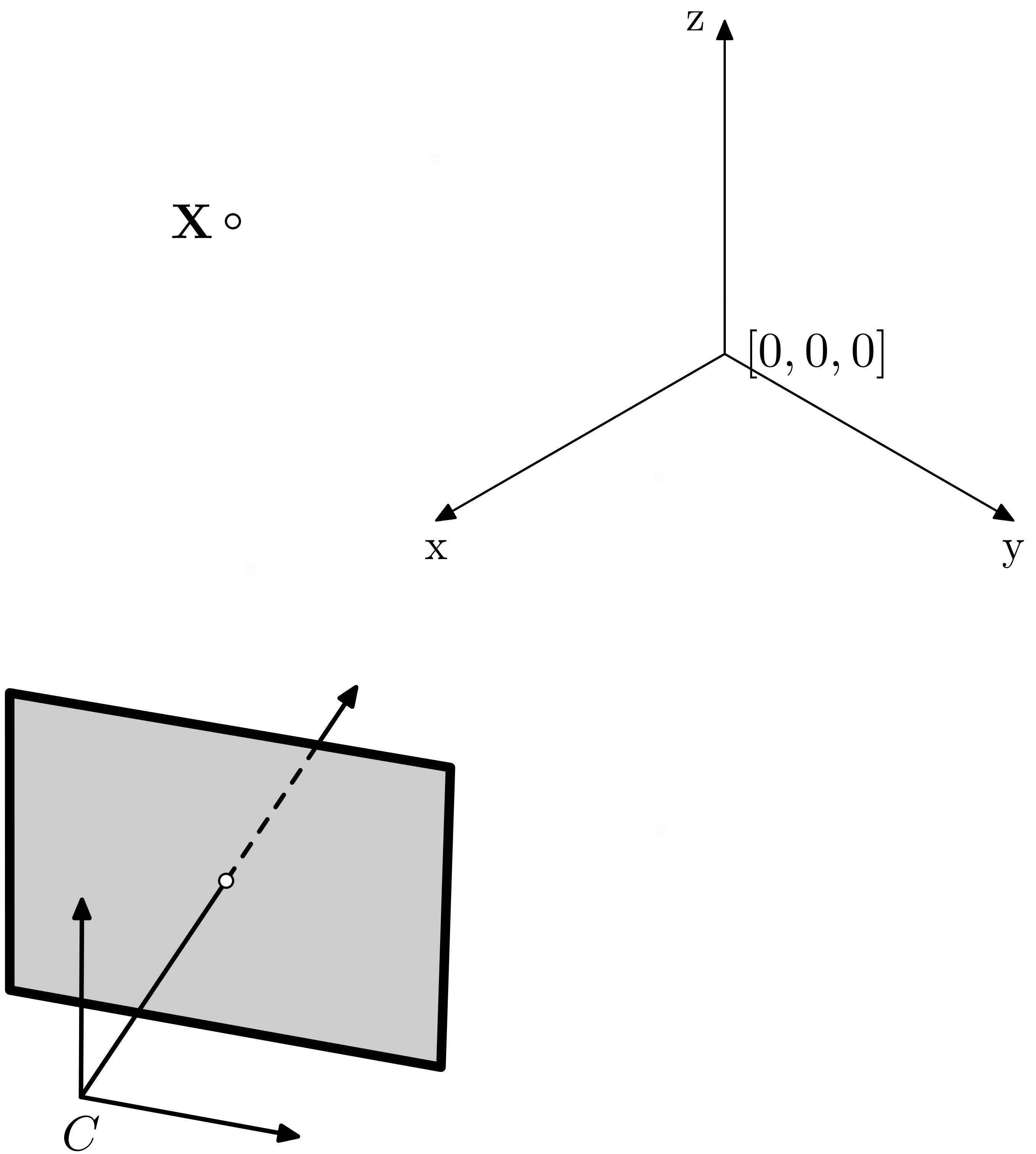
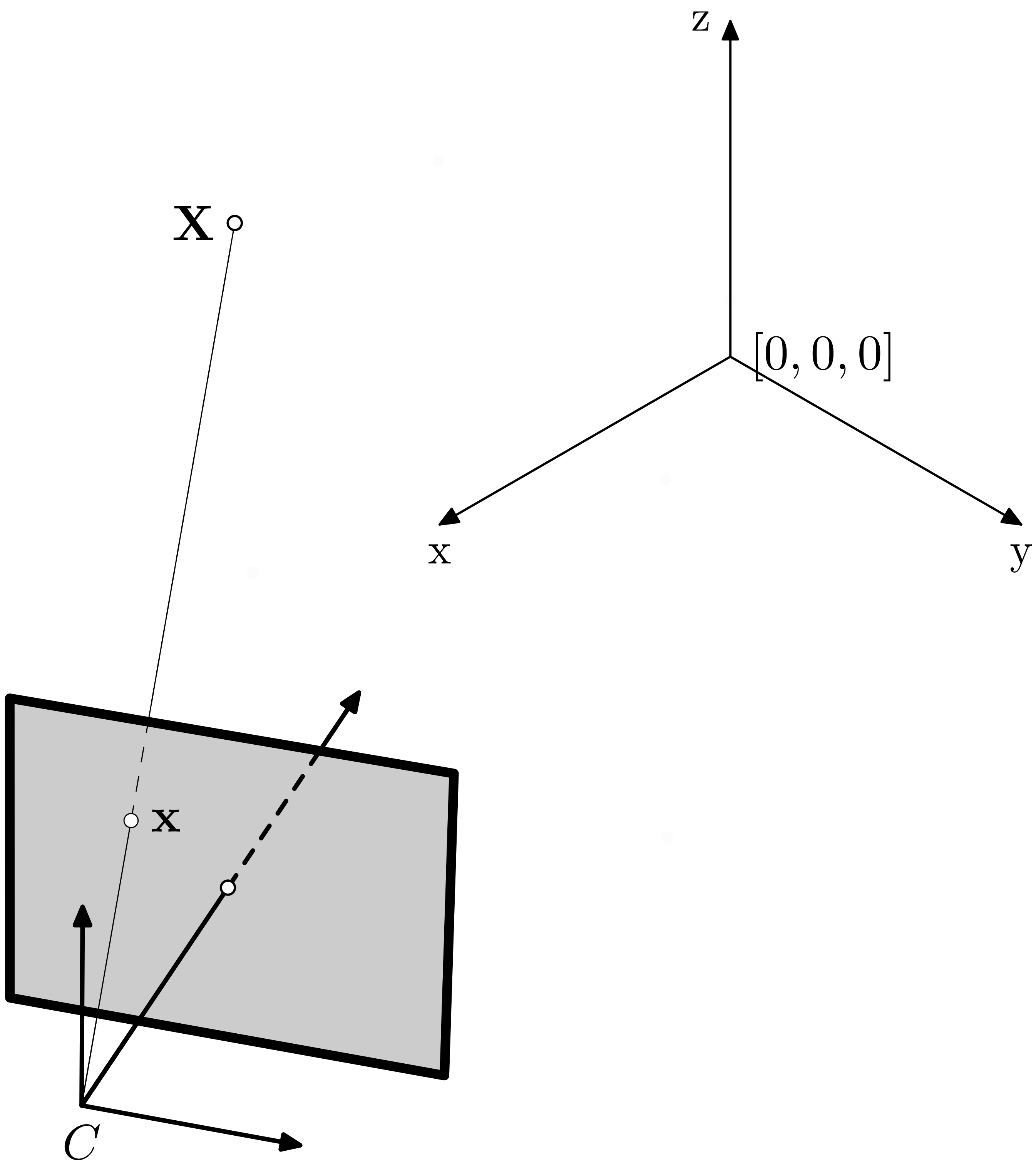


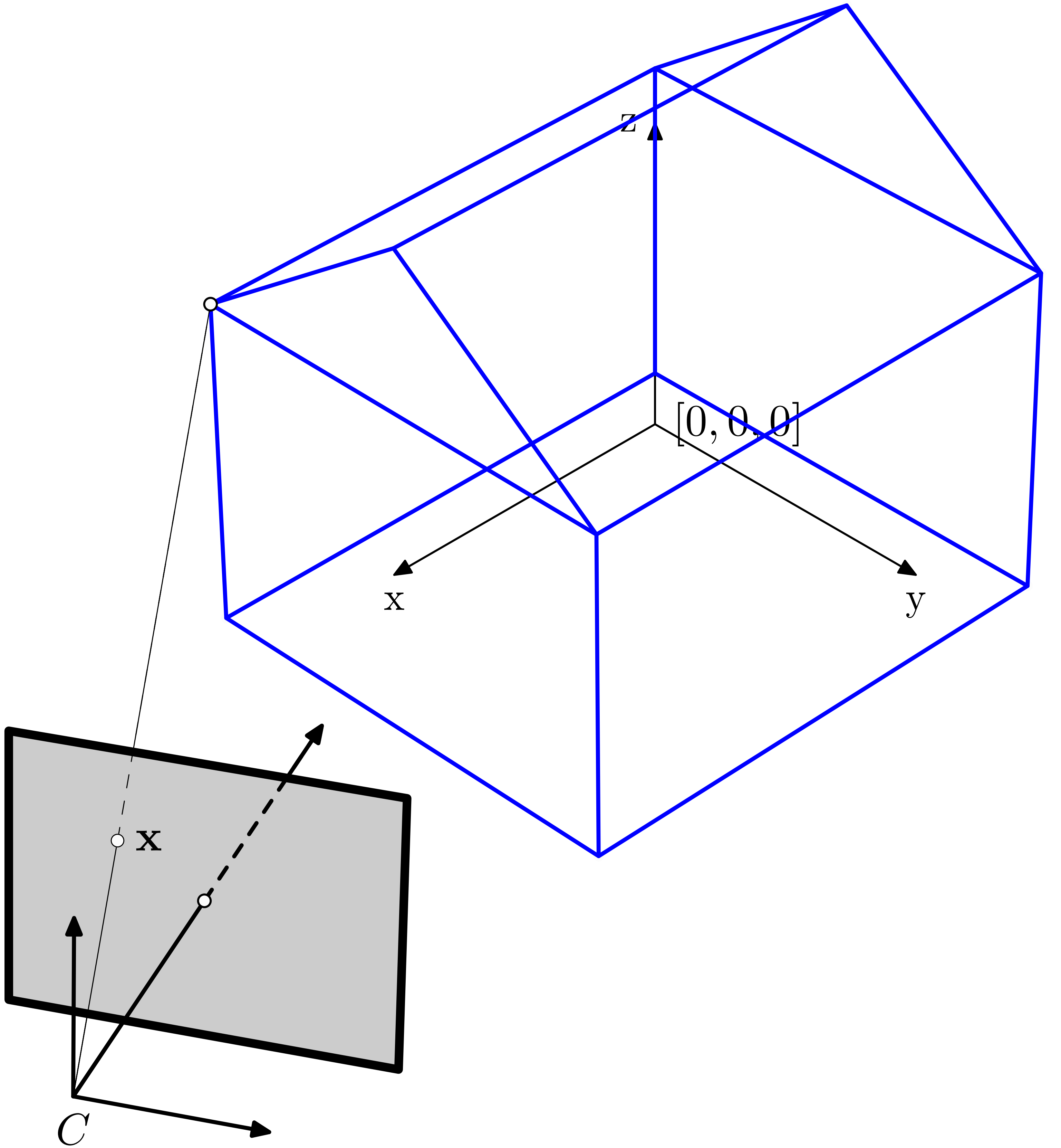
image plane
image plane
X_3
X_1
X_2
x
$-f$
x_1
x_3
x_2
z
C
f
optical axis
x_3
x_2
x_1
Z_1
Z_2
Z_3

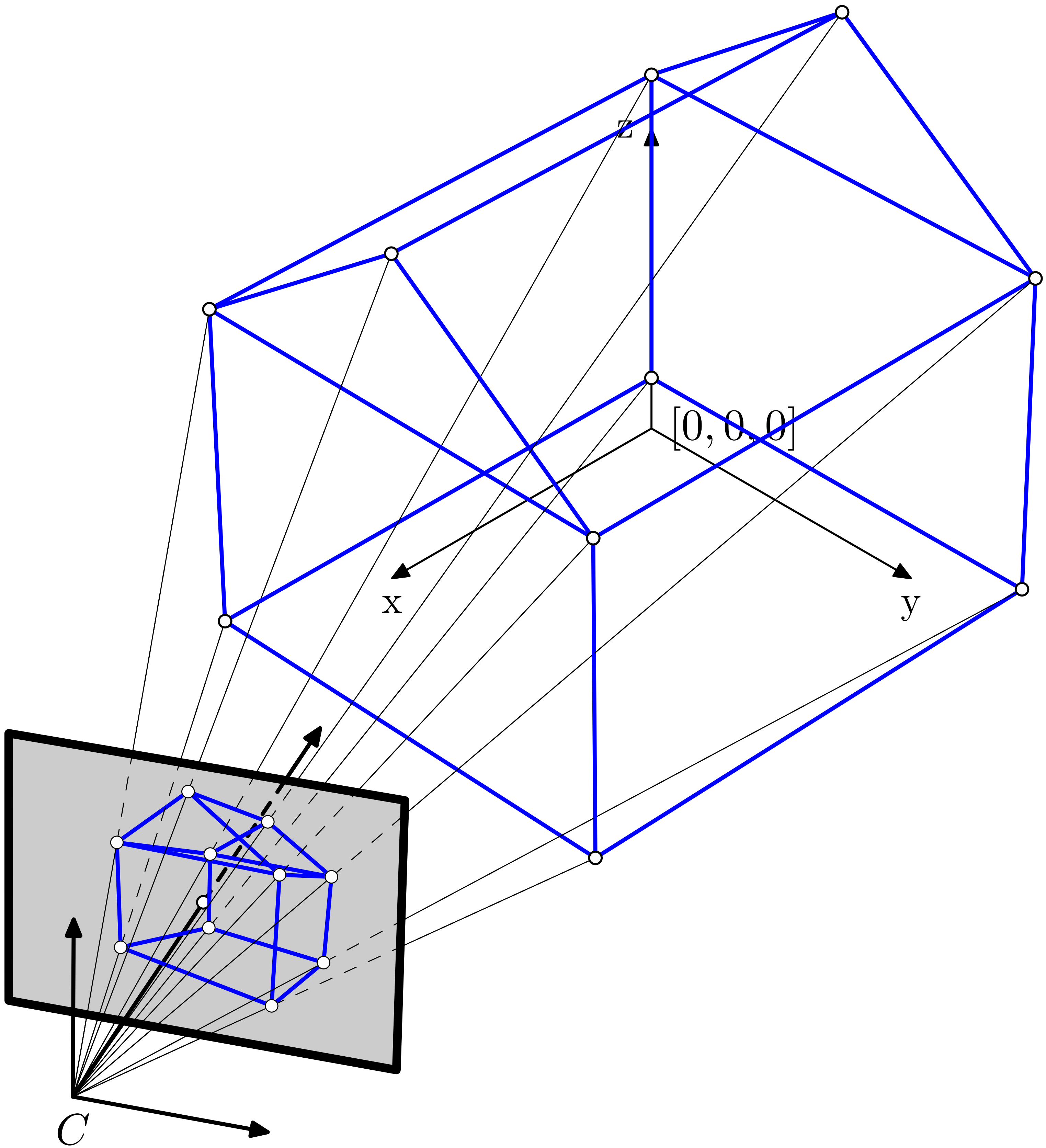

z
$\mathbf{X}$
$[0,0,0]$
x
y



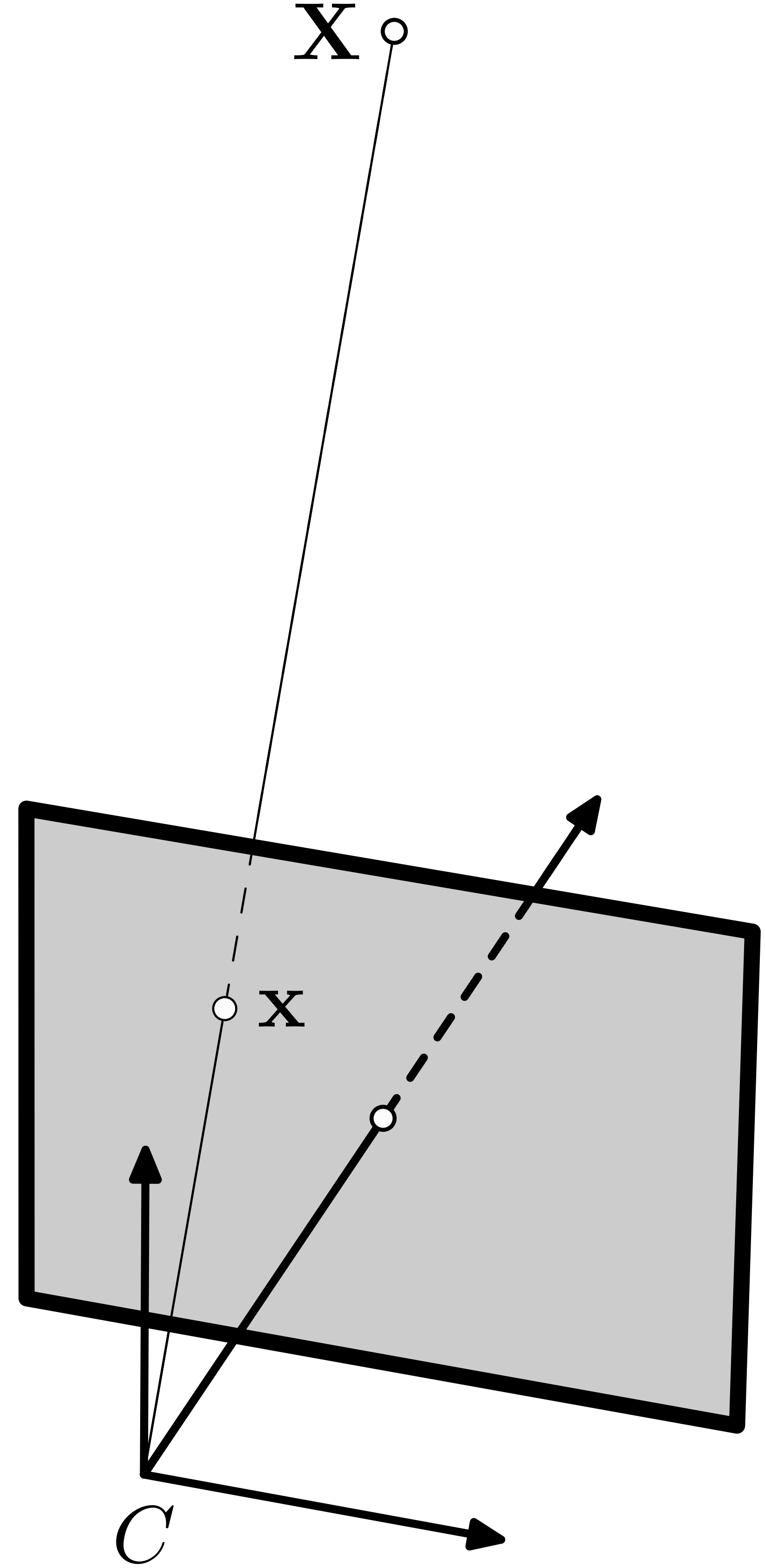
z
X
[0, 0, 0]
x
y
C

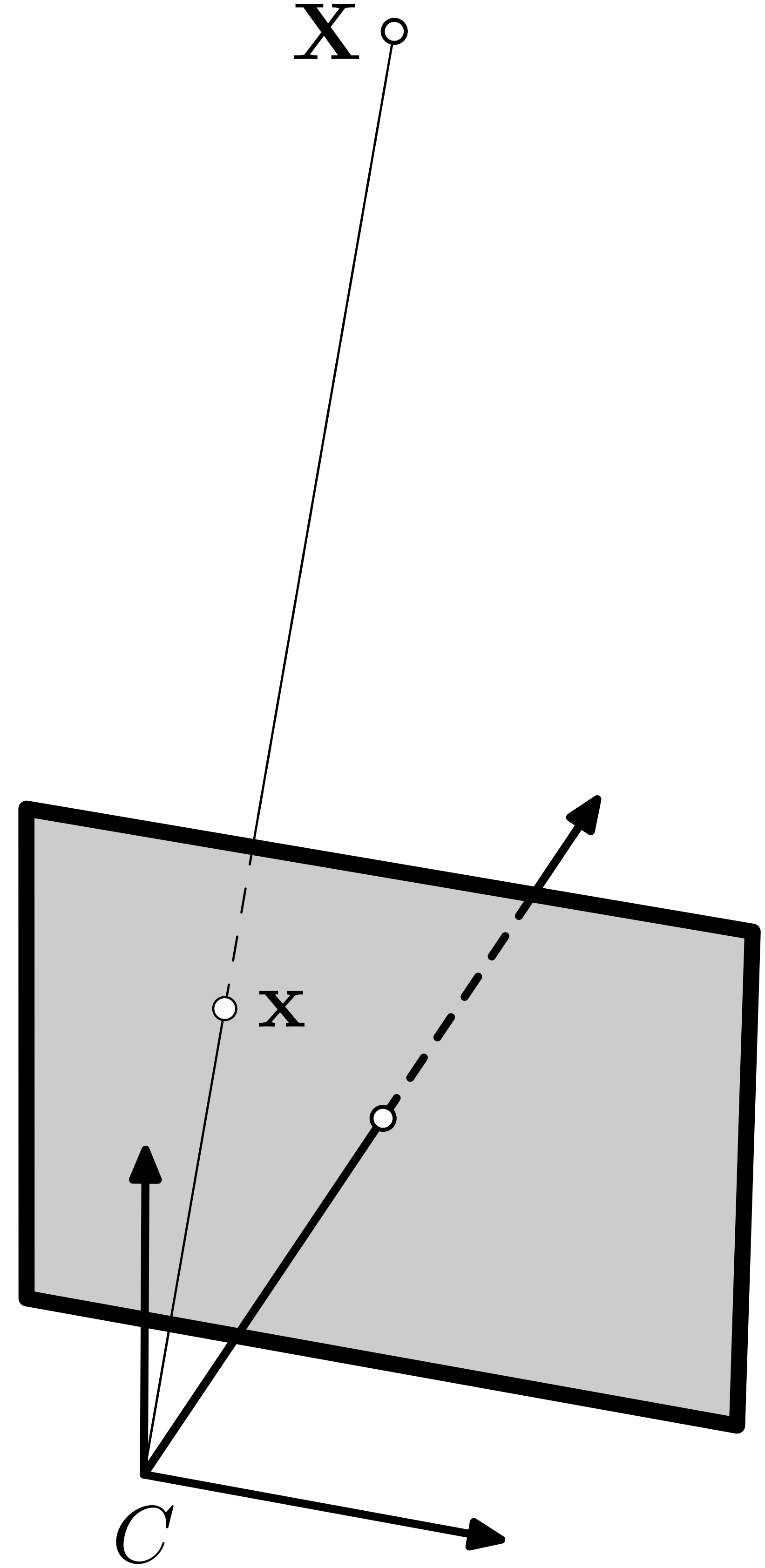

z
[0, 0, 0]
x
y
X
x
C

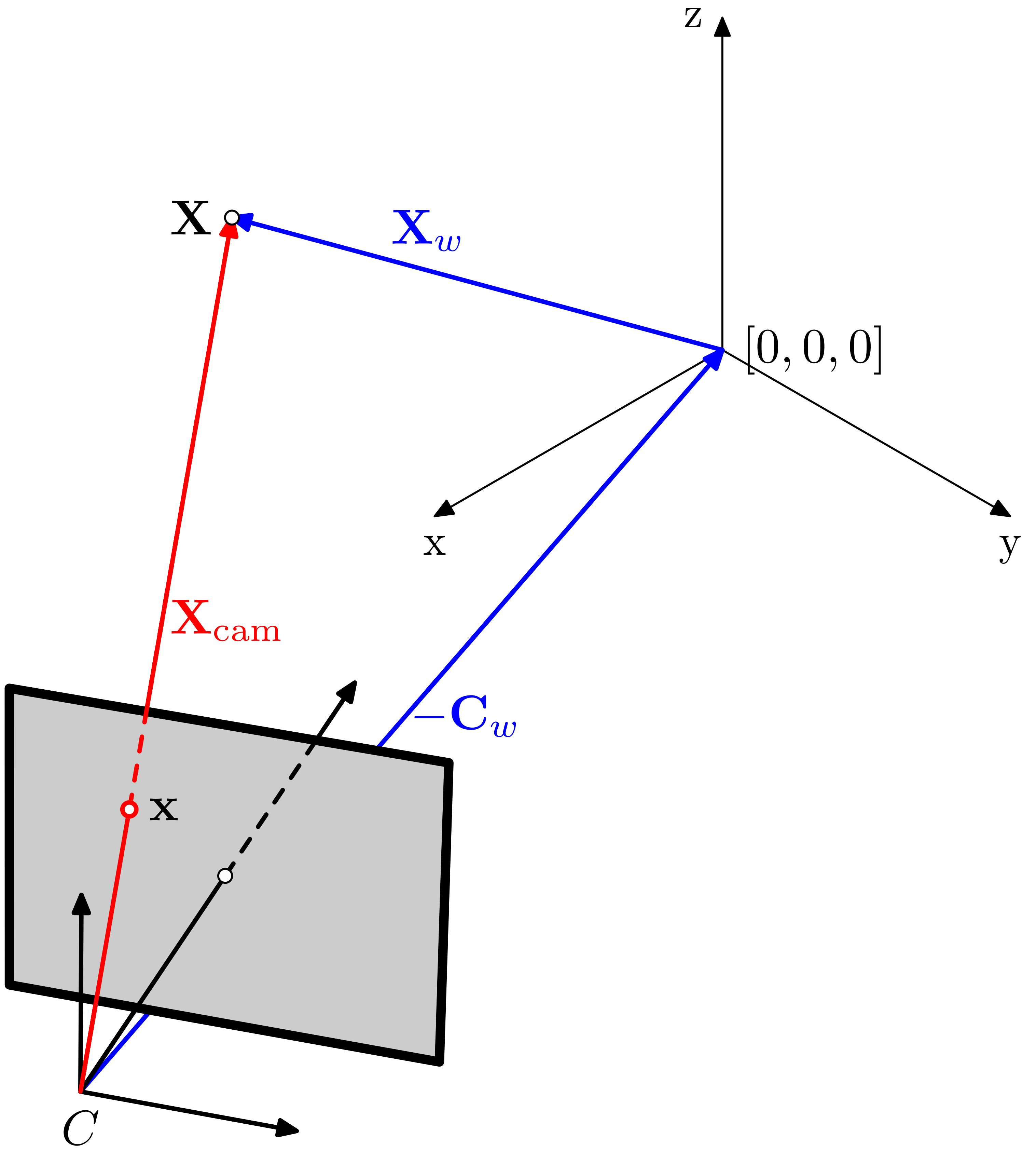

z
[0, 0, 0]
x
y
x
C



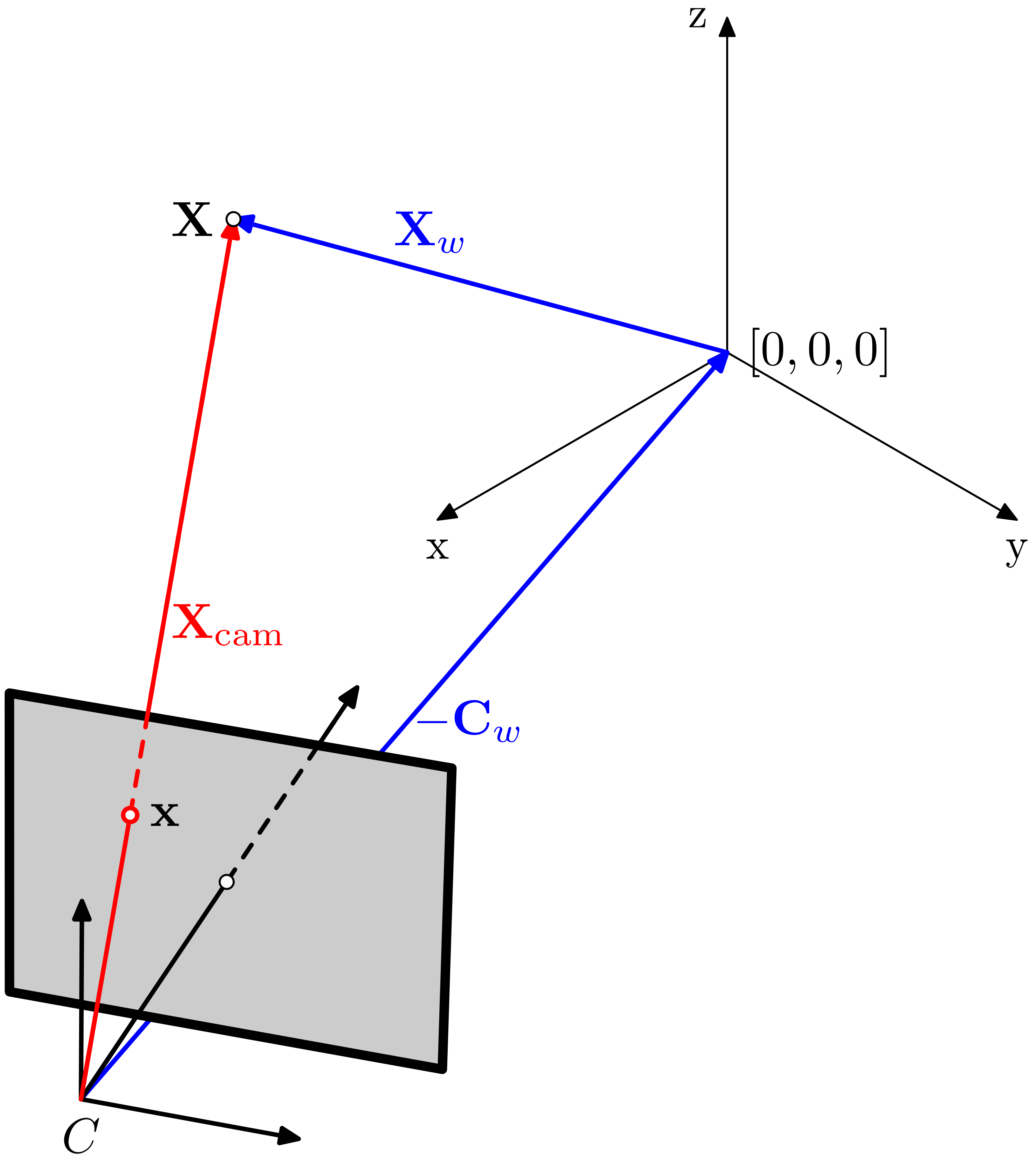
z
[0, 0, 0]
x
y
C


X
x
C


X
x
C



z
$\mathbf{X}$
$\mathbf{X}_w$
$[0, 0, 0]$
x
y
$\mathbf{X}_{\text{cam}}$
$-\mathbf{C}_w$
$\mathbf{x}$
C



z
X
$\mathbf{X}_w$
[0, 0, 0]
x
y
$\mathbf{X}_{\text{cam}}$
$-\mathbf{C}_w$
x
C