Lecture 12

Perspective:

image projection
translation
vanishing points

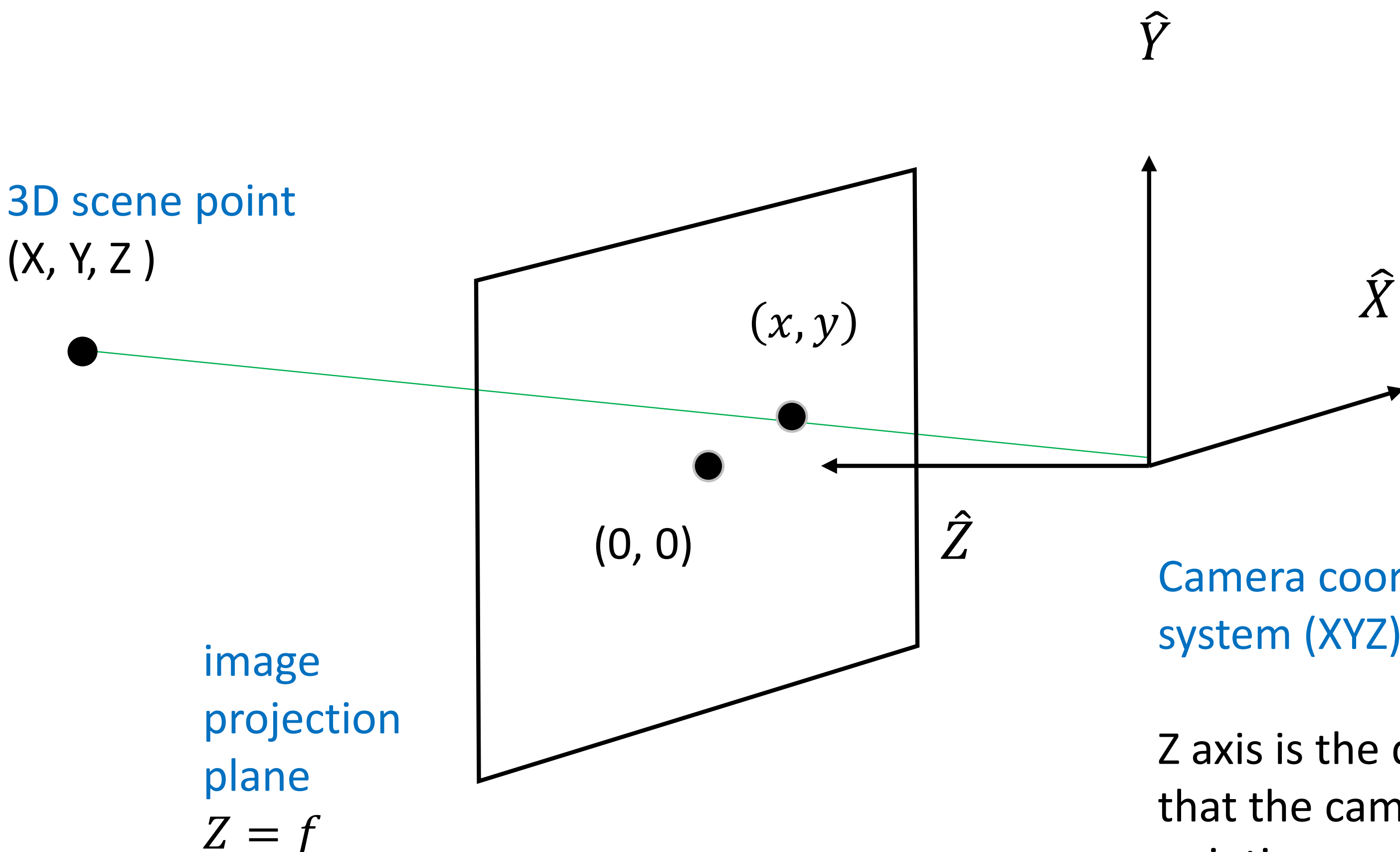
Mon. Oct. 19, 2020

Perspective techniques in painting were discovered in the 15 century.

If you have taken an art class, then you'll be familiar with how to draw using perspective.

Today, we'll look at the basic models of perspective used in 3D vision.

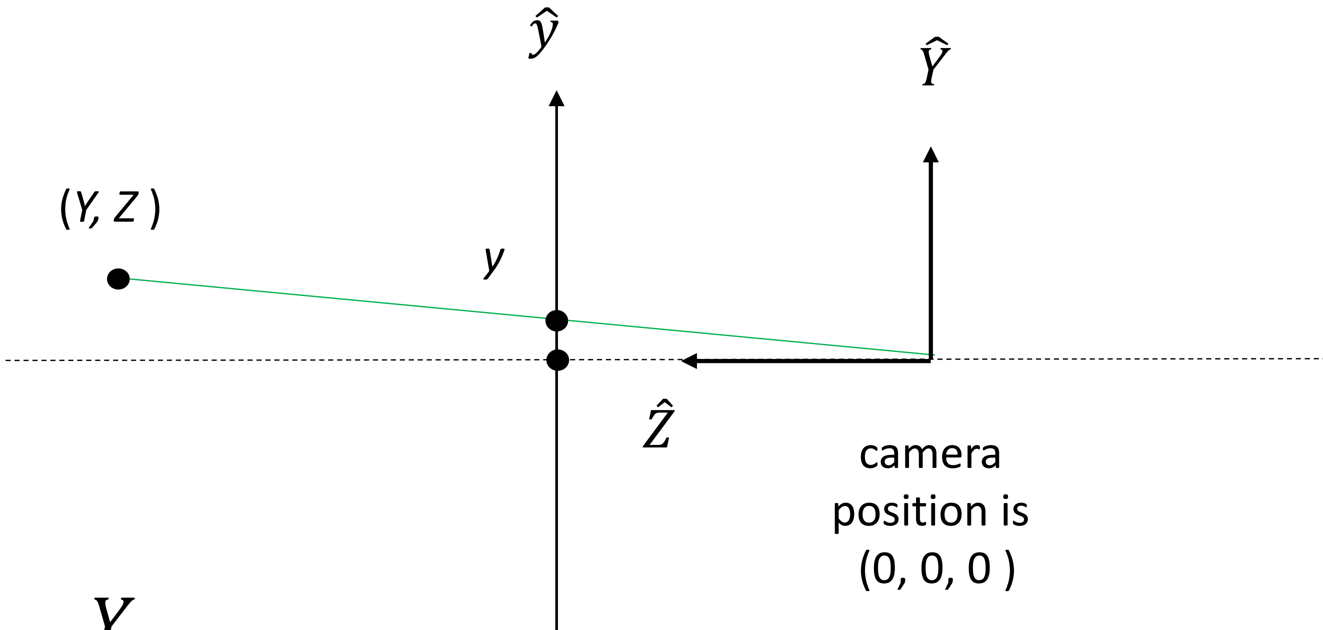
Image projection



Camera coordinate system (XYZ)

Z axis is the direction that the camera is pointing.

View from side (YZ)



$$\frac{y}{f} = \frac{Y}{Z}$$

Image
projection
plane
$Z = f$

View from above (XZ)

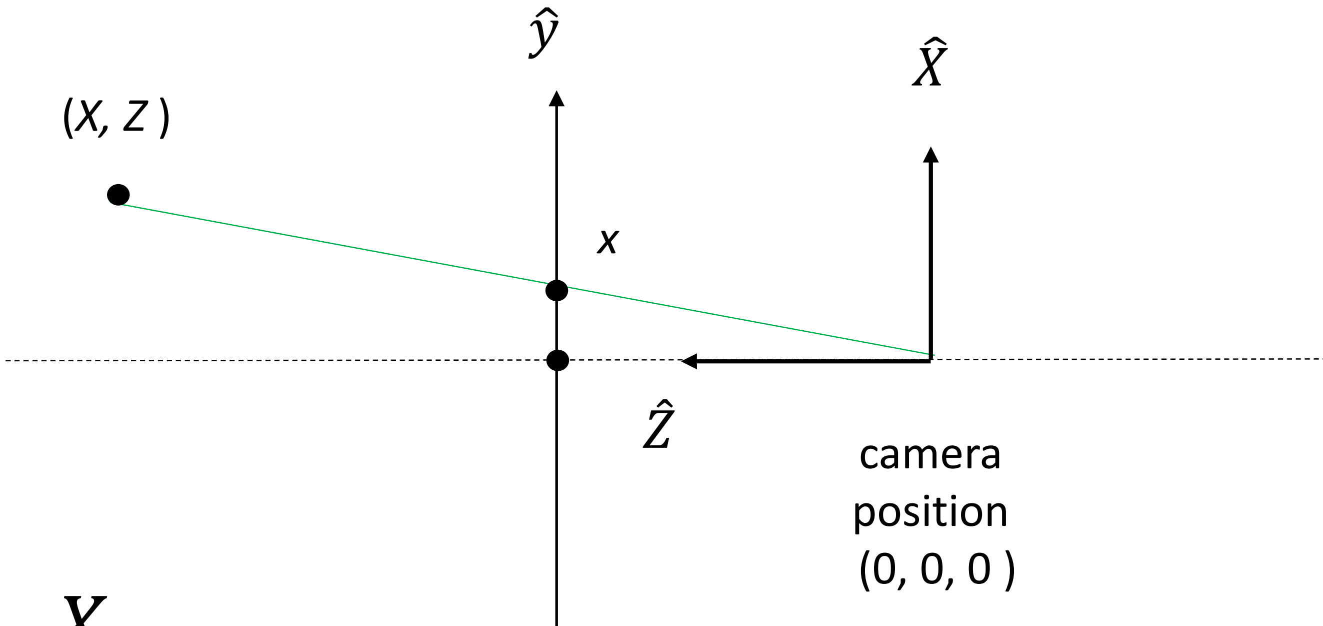
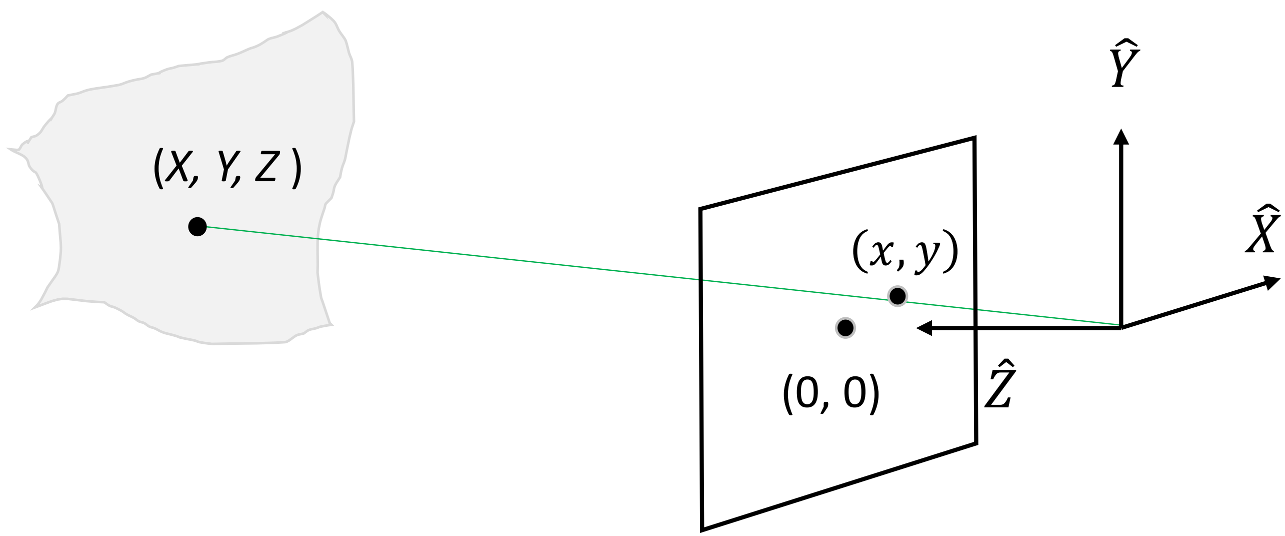


Image
projection
plane
$Z = f$

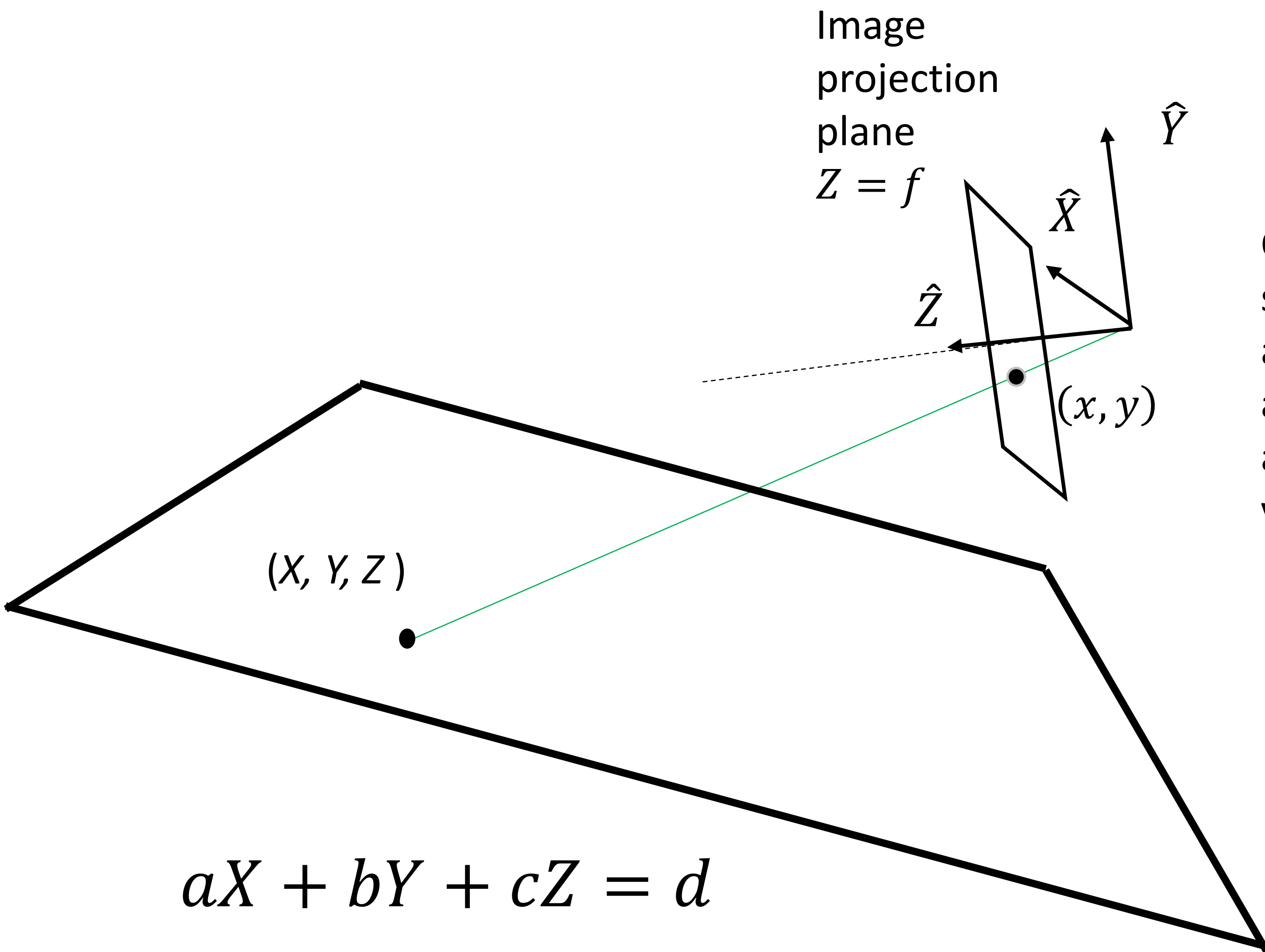
$$\frac{x}{f} = \frac{X}{Z}$$

Depth Map $Z(x, y)$



The mapping from image positions (x, y) to depth Z values on a 3D surface is called a “depth map”. We write $Z(x, y)$.

Image Projection of a Scene Plane



Camera coordinate system can be arbitrary. E.g. the Y axis is not necessarily aligned with gravity vector.

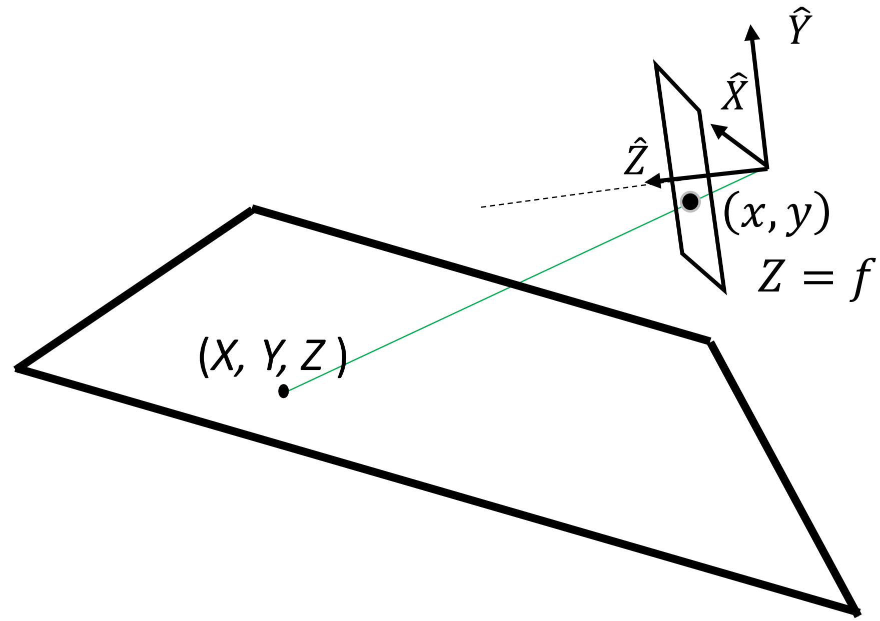
$$aX + bY + cZ = d$$

What is the depth map $Z(x, y)$?

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

$$\frac{y}{f} = \frac{Y}{Z}$$

$$aX + bY + cZ = d$$



Multiplying the last one by $\frac{f}{Z}$ and substituting gives …

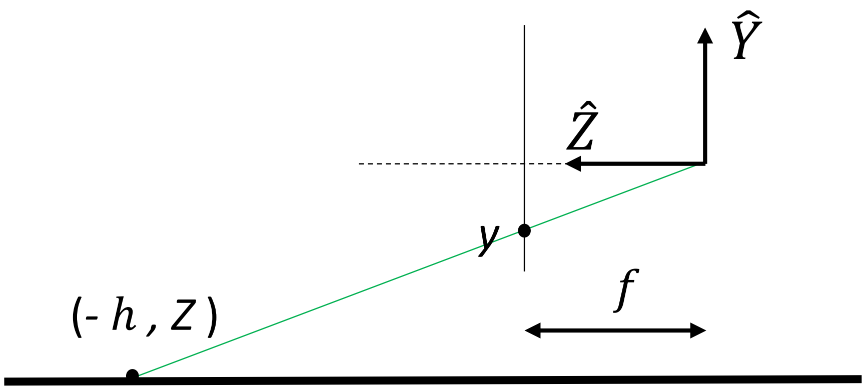
$$ax + by + cf = \frac{fd}{Z}$$

$$Z(x, y) = \frac{fd}{ax+by+cf} .$$

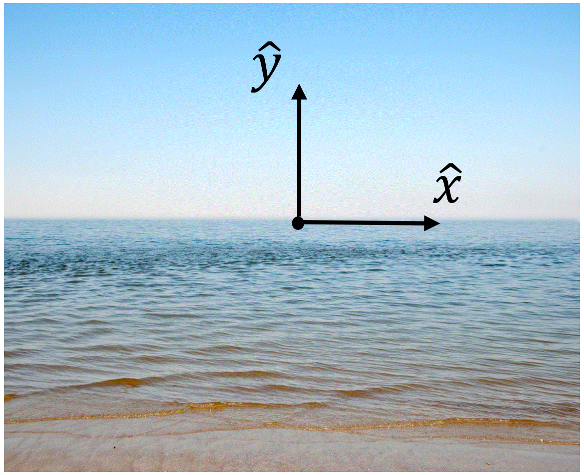
Example (ground and horizon)

What is the depth map $Z(x, y)$?

Example



Ground plane $Y = -h$

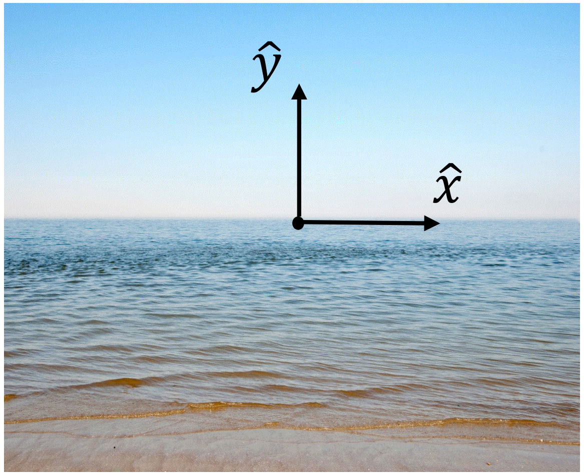


$$aX + bY + cZ = d$$

$$a = 0, b = 1, c = 0, d = -h$$

$$Z(x, y) = \frac{-fh}{y}$$

$$Z(x, y) = \frac{fd}{ax+by+cf}\ .$$



$$Z(x, y) = -\frac{fh}{y}$$

What happens to depth map as $y \rightarrow 0$?

Also note the depth map above the *horizon* $(y > 0)$ is undefined.

Lecture 12

Perspective:

image projection

camera translation

vanishing points

Mon. Oct. 19, 2020

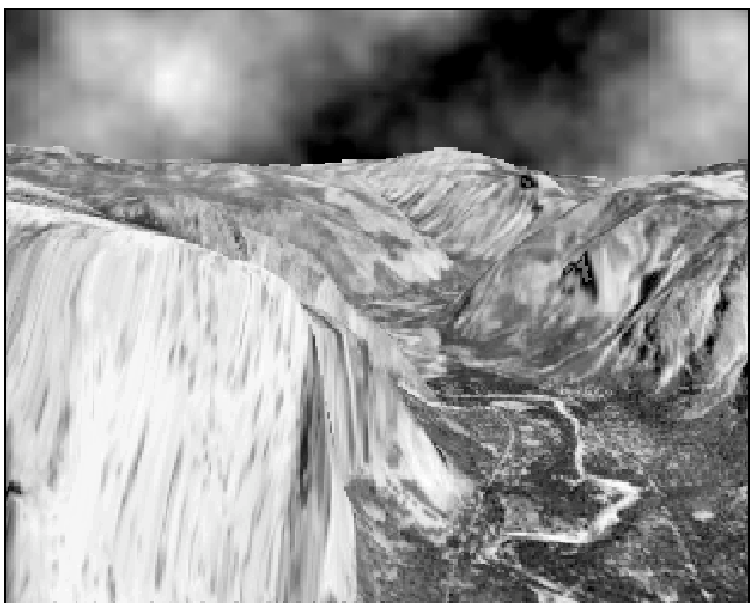
What is the motion field *produced by* a moving observer (camera) ?

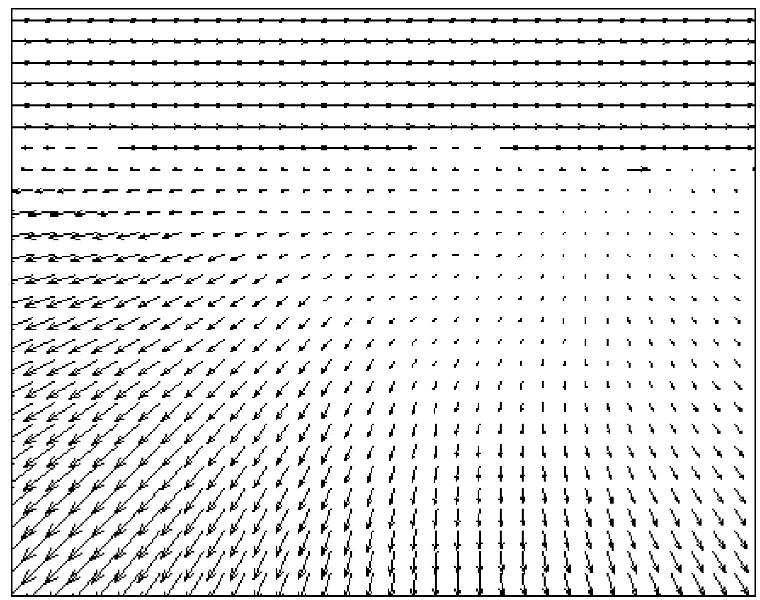
What is the motion field *produced by* a moving observer (camera) ?

Assume the scene is static, and camera is moving.

Assume camera translation only.

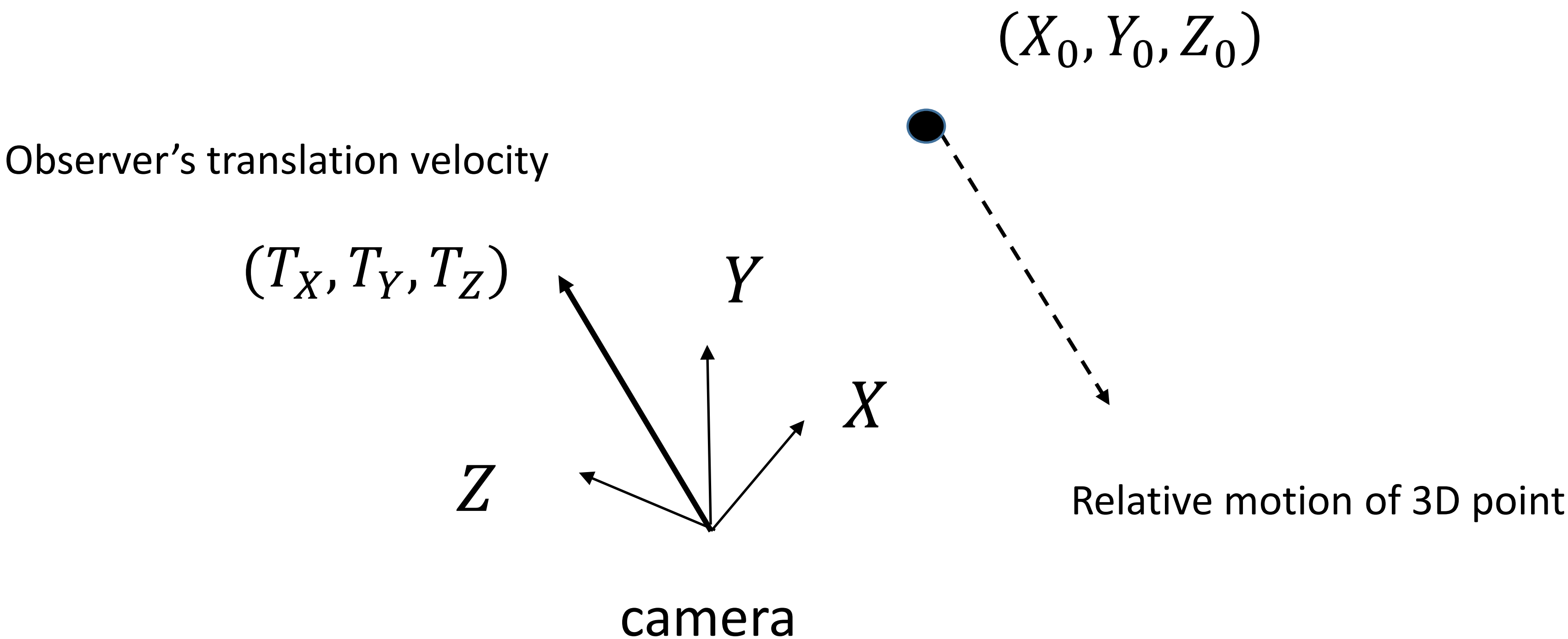
We will discuss camera rotation next lecture.





"Yosemite sequence" (rendered)

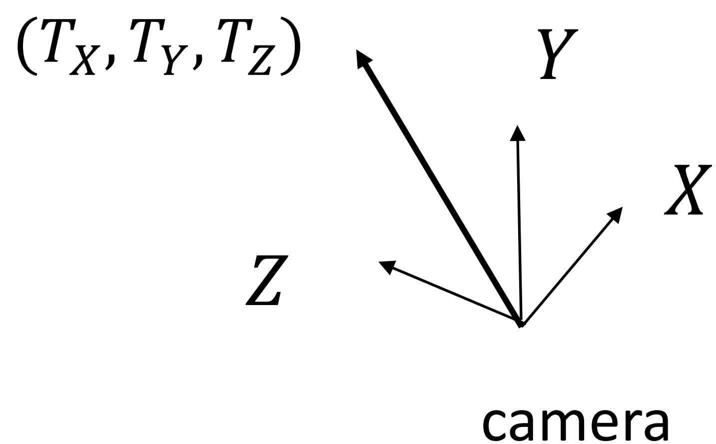
Motion field produced by a *translating* observer

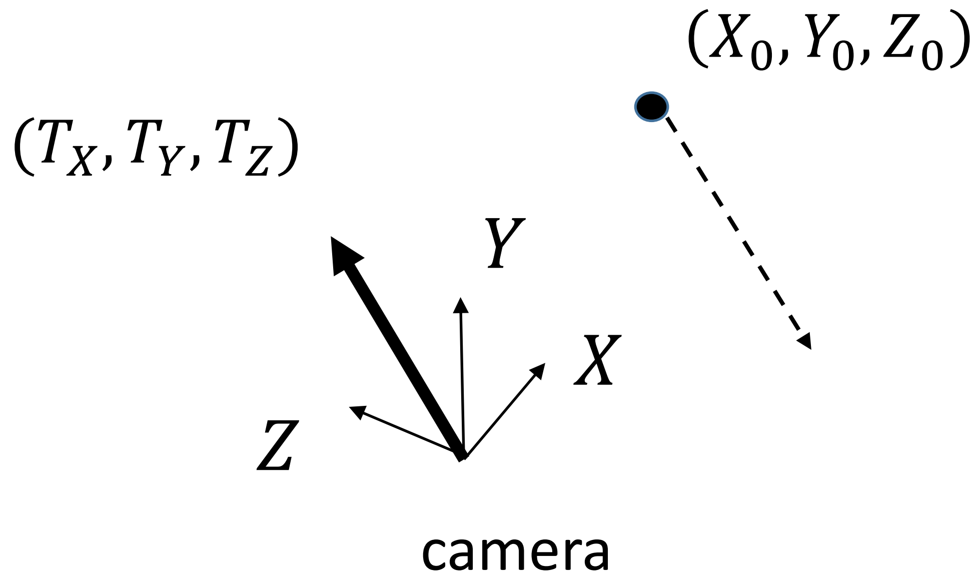


Motion field produced by a translating observer

We will decompose the motion field into two components:

- due to forward motion T_Z
- due to lateral motion (T_X, T_Y)



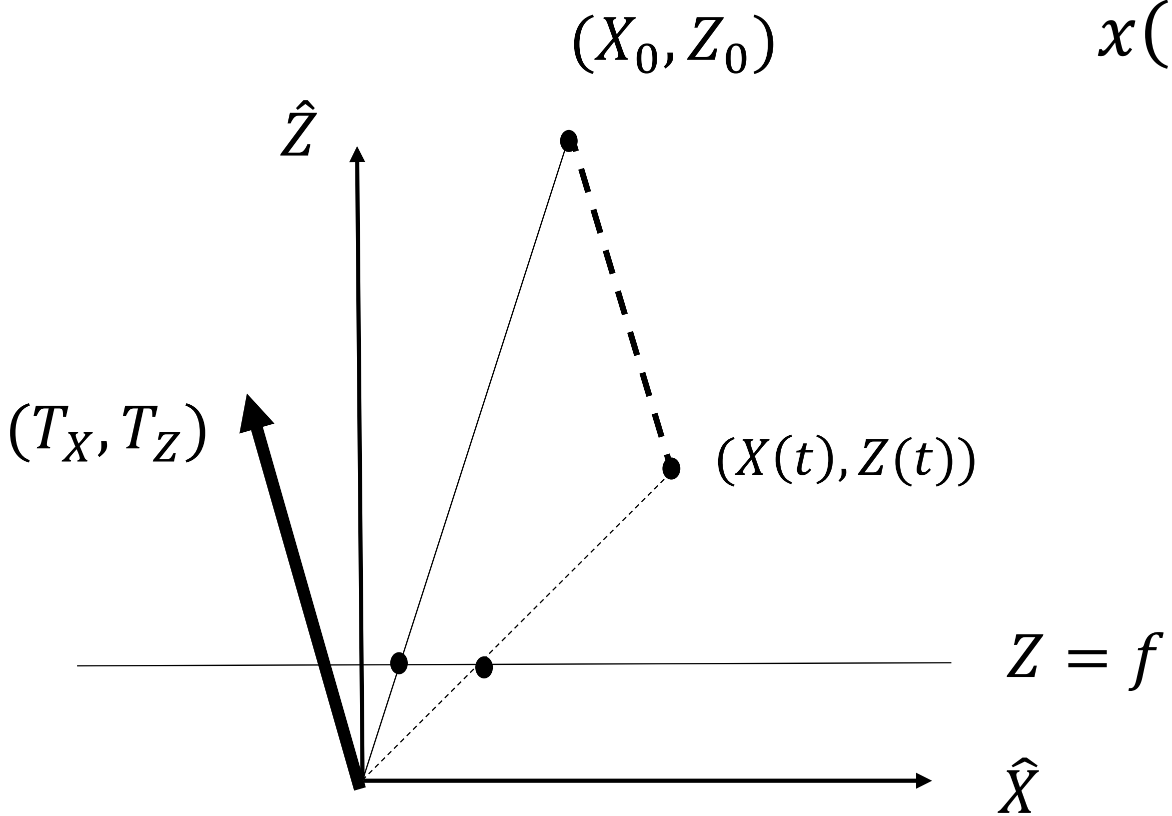


The 3D path of the scene point in the camera's coordinate system is:

$$(X(t), Y(t), Z(t)) = \ (X_0, Y_0, Z_0) + t\left(-T_x, -T_y, -T_z\right)$$

The *relative* 3D velocity of the scene point $(-T_X, -T_Y, -T_Z)$

What is the **image path** $(x(t), y(t))$ of the scene point on the projection plane?

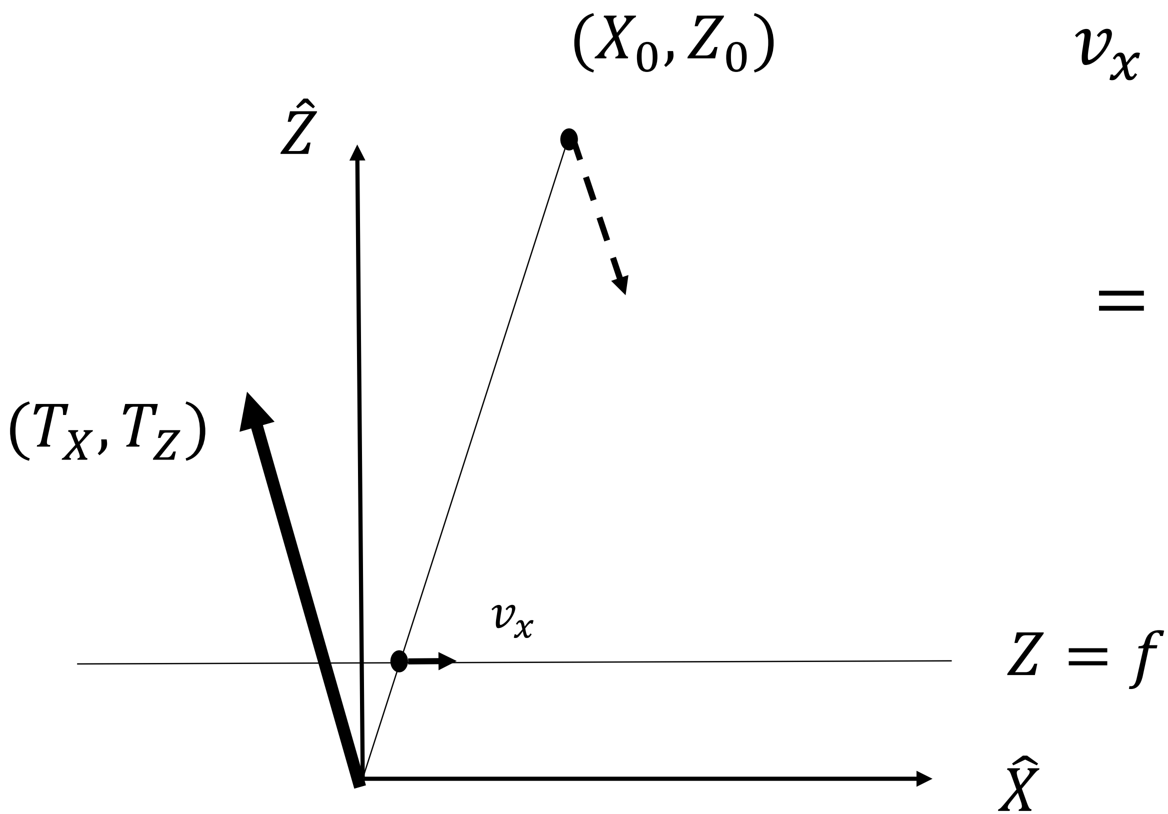


$$x(t) = f\,\frac{X(t)}{Z(t)}$$

$$= f\;\frac{X_o - T_X\,t}{Z_o - T_Z\,t}$$

Notation: $(x(t), y(t))$ is a position in the plane $Z = f$.

What is the **image velocity** (v_x, v_y) of the scene point?



$$v_x = \frac{dx(t)}{dt} \Big|_{t=0}$$

$$= f \, \frac{-T_X Z_0 + T_Z X_0}{{Z_o}^2}$$

What is the **image velocity** (v_x, v_y) of the scene point?

$$(v_x, v_y) = \left(f \ \frac{-T_X Z_0 + T_Z X_0}{{Z_o}^2} \ , f \ \frac{-T_Y Z_0 + T_Z Y_0}{{Z_o}^2} \right)$$

Previous slide

Same derivation for Y.

What is the **image velocity** (v_x, v_y) of the scene point?

$$(v_x, v_y) = \left(f\,\frac{-T_X Z_0 + T_Z X_0}{{Z_o}^2}, f\,\frac{-T_Y Z_0 + T_Z Y_0}{{Z_o}^2} \right)$$

$$= \frac{f}{Z_o}\left(-T_X, -T_Y\right) + \frac{T_Z}{Z_o}\underbrace{\left(\frac{f X_o}{Z_o}, \frac{f Y_0}{Z_o}\right)}_{=(x,y)}$$

Lateral translation component

Forward translation component

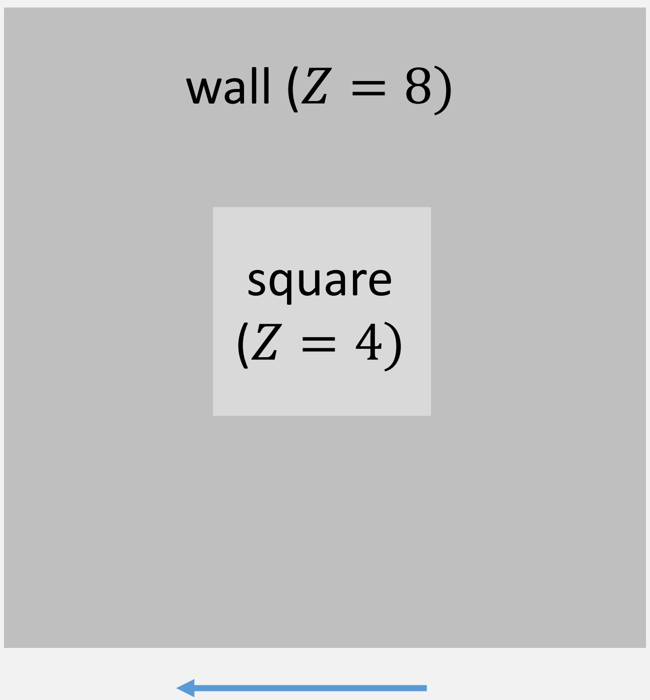
Both components depend on inverse depth.

Let’s look at some examples.

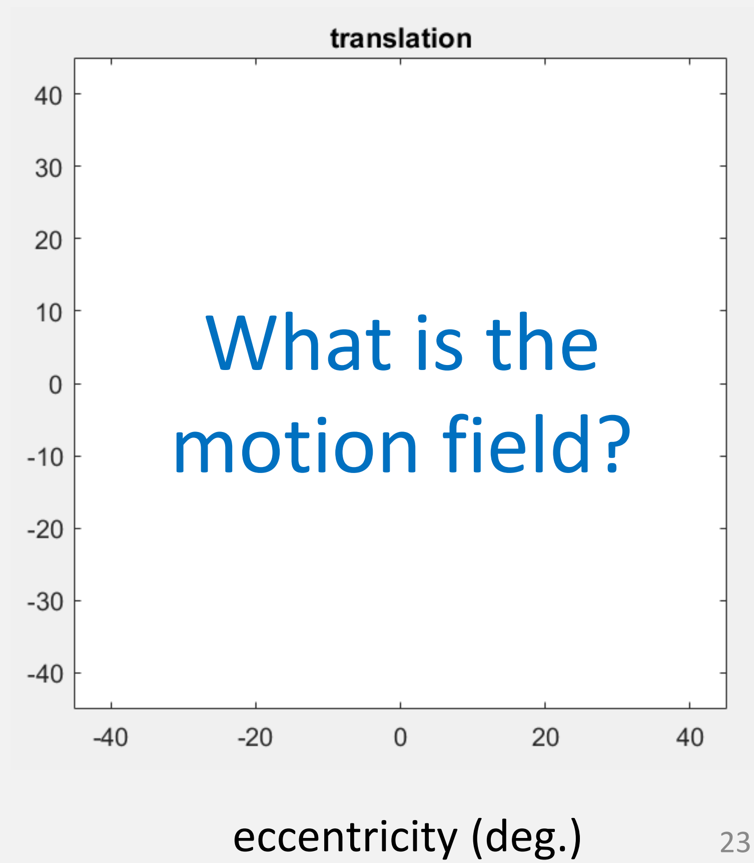
Lateral translation ($T_Z = 0$)

$$(v_x, v_y) = \frac{f}{Z_o}(-T_X, -T_Y)$$

Example:



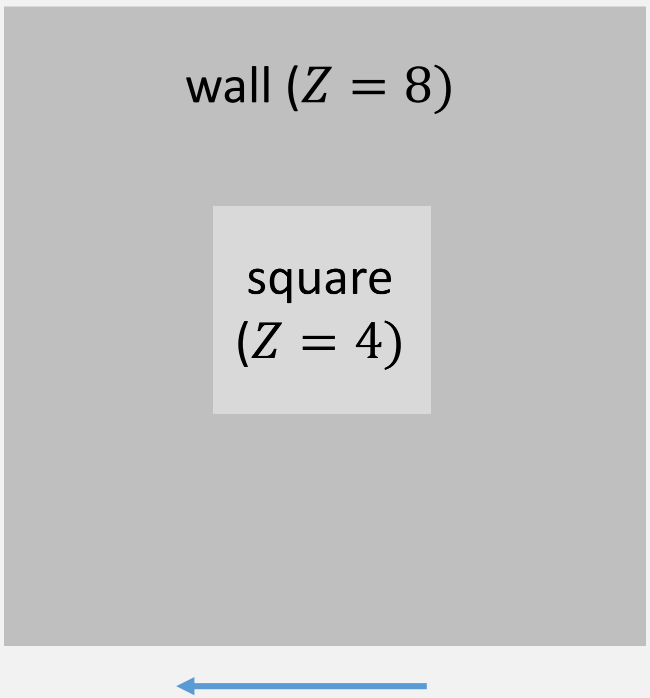
$(T_X, T_Y = 0)$



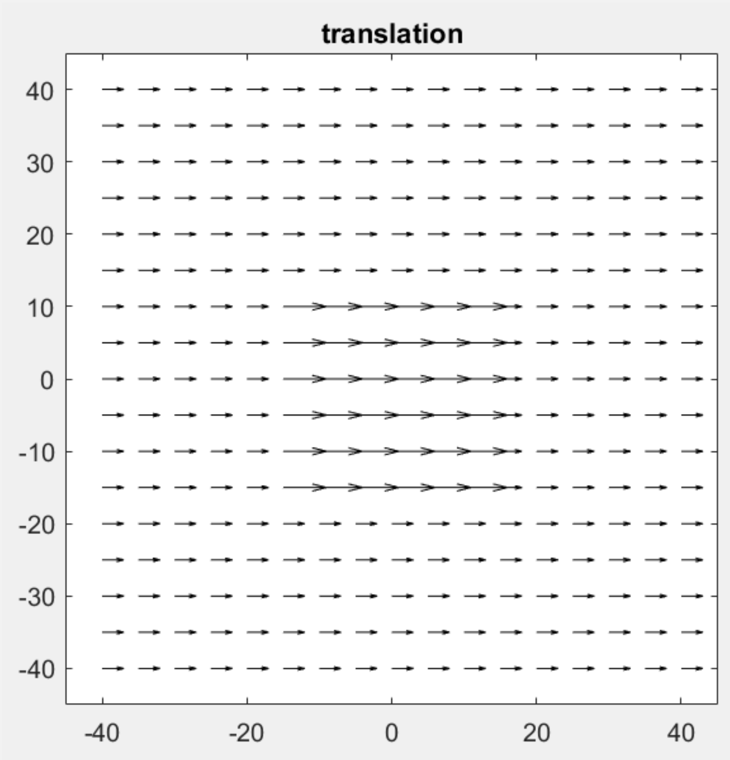
Lateral translation $(T_Z = 0)$

$$(v_x, v_y) = \frac{f}{Z_o}(-T_X, -T_Y)$$

Example:



$(T_X \neq 0, T_Y = 0)$

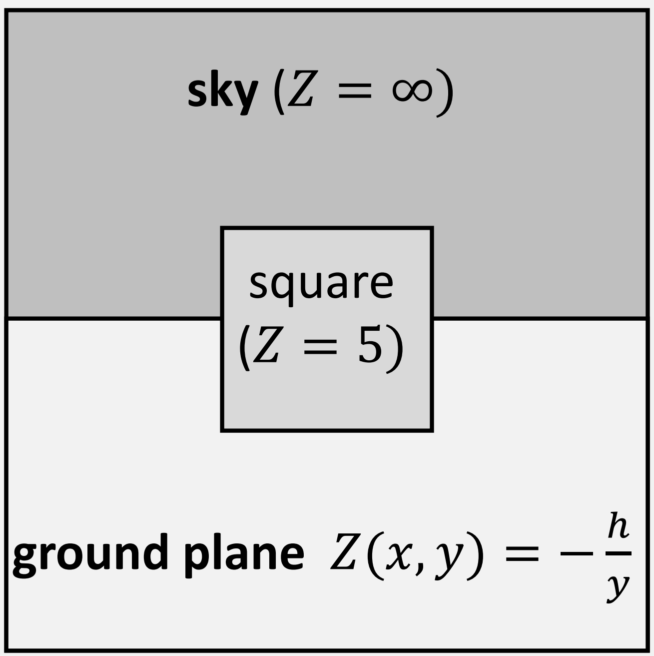


eccentricity (deg.)

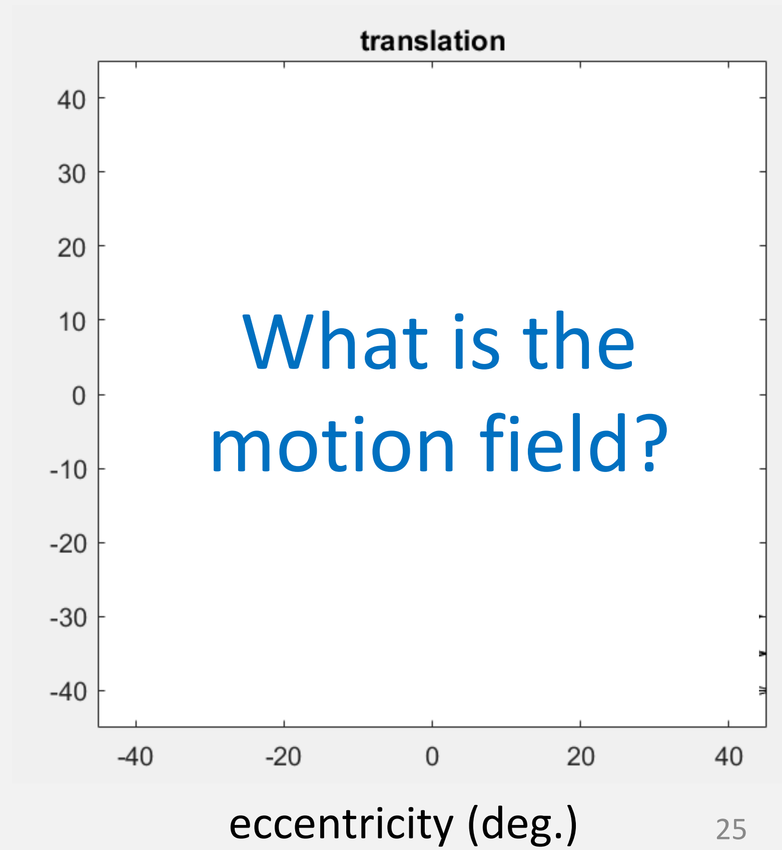
Lateral translation $(T_Z = 0)$

$$\left(v_x, v_y\right) = \frac{f}{Z_o}\left(-T_X, -T_Y\right)$$

Example:



$(T_X \neq 0, T_Y = 0)$

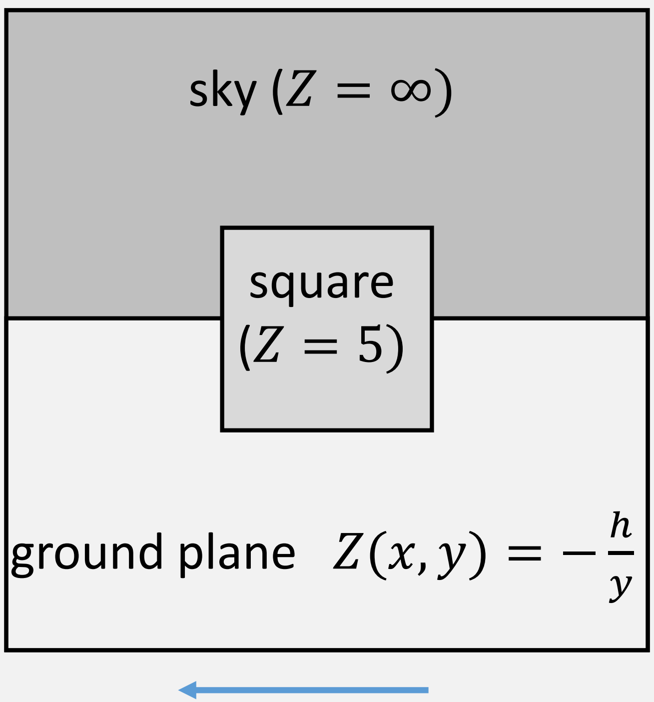


Lateral translation ($T_Z = 0$)

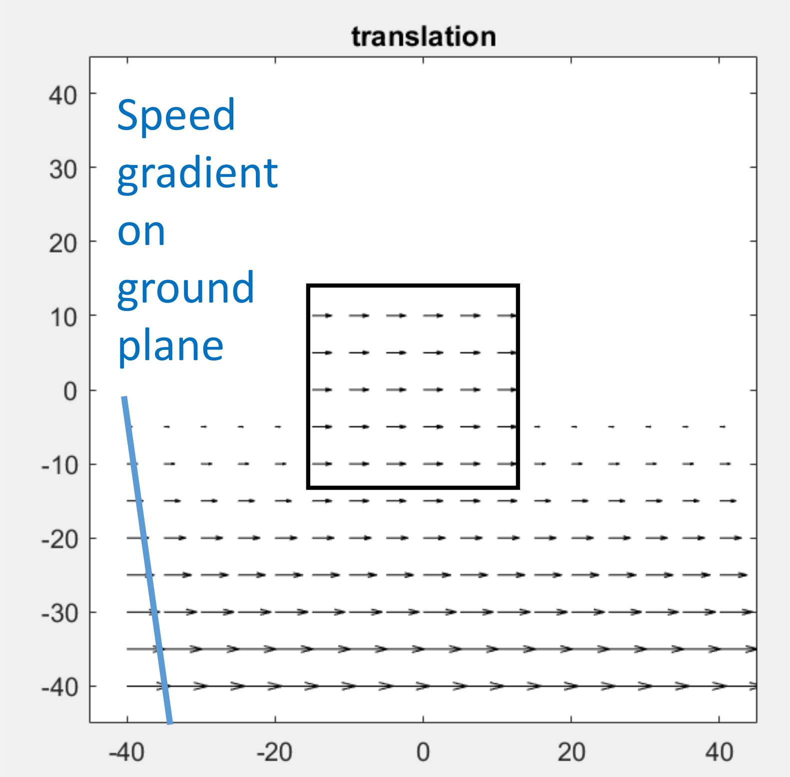
$$(v_x, v_y) = \frac{f}{Z_o}\,(-T_X, -T_Y)$$

$$= -\frac{fy}{h}\,(-T_X, 0)$$

Example:



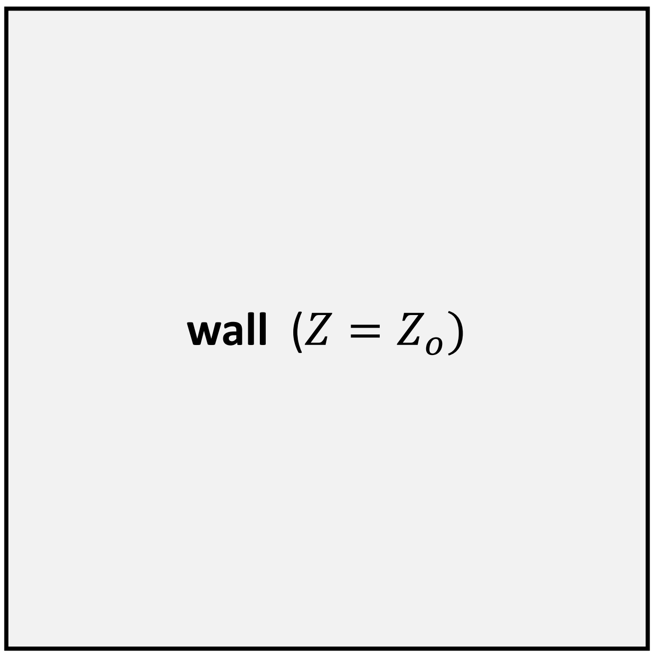
($T_X \neq 0$, $T_Y = 0$)

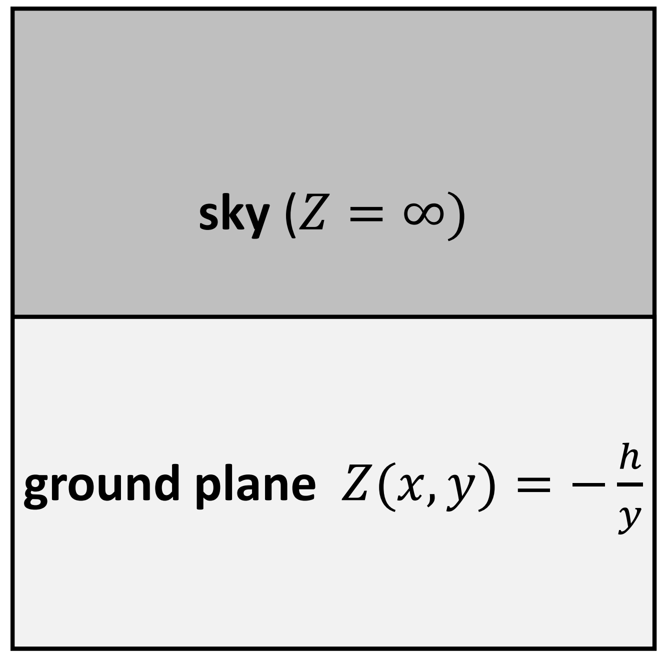


eccentricity (deg.)

Forward translation $(T_X = T_Y = 0)$

$$(v_x, v_y) = \frac{T_Z}{Z_o}(x, y)$$



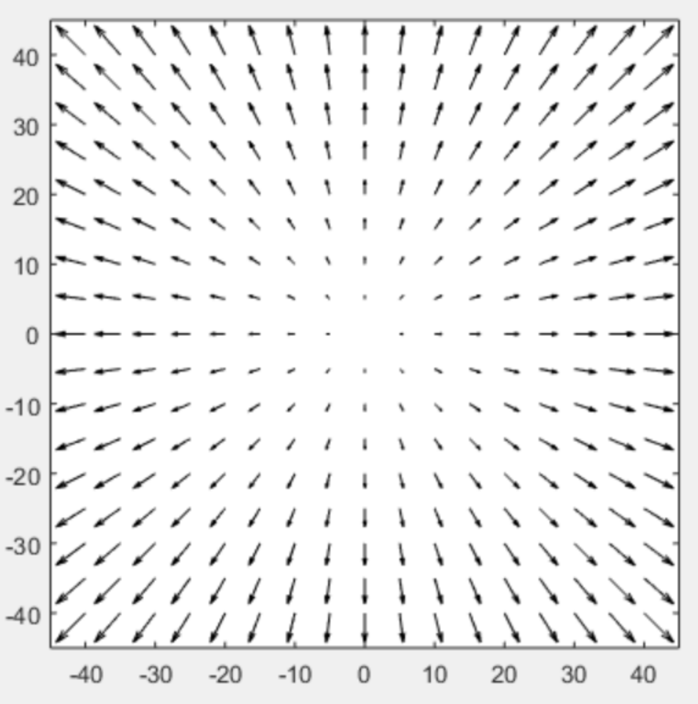


What are the motion fields?

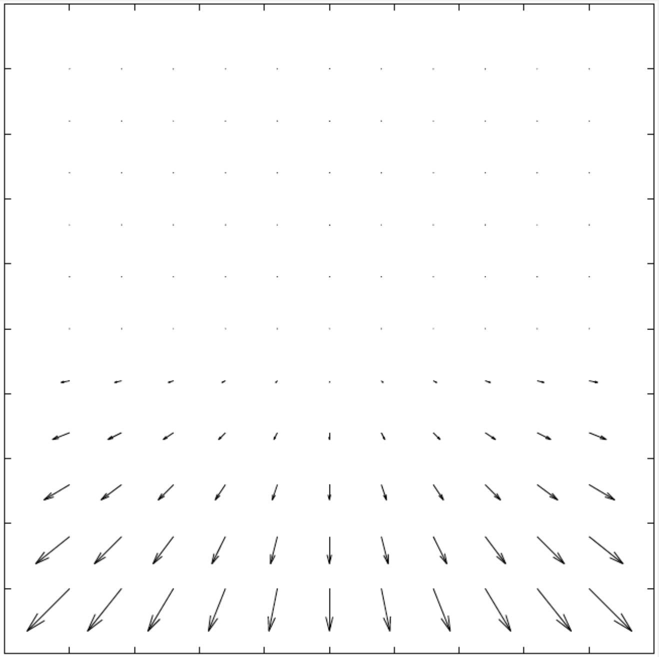
Forward translation $(T_X = T_Y = 0)$

$$(v_x, v_y) = \frac{T_Z}{Z(x,y)} \; (x, y)$$

$Z(x, y)$ is a depth map.
(x, y) is a position in projection plane.



wall
$Z = Z_o$



sky and ground plane
(see Exercises)

Lecture 12

Perspective:

image projection

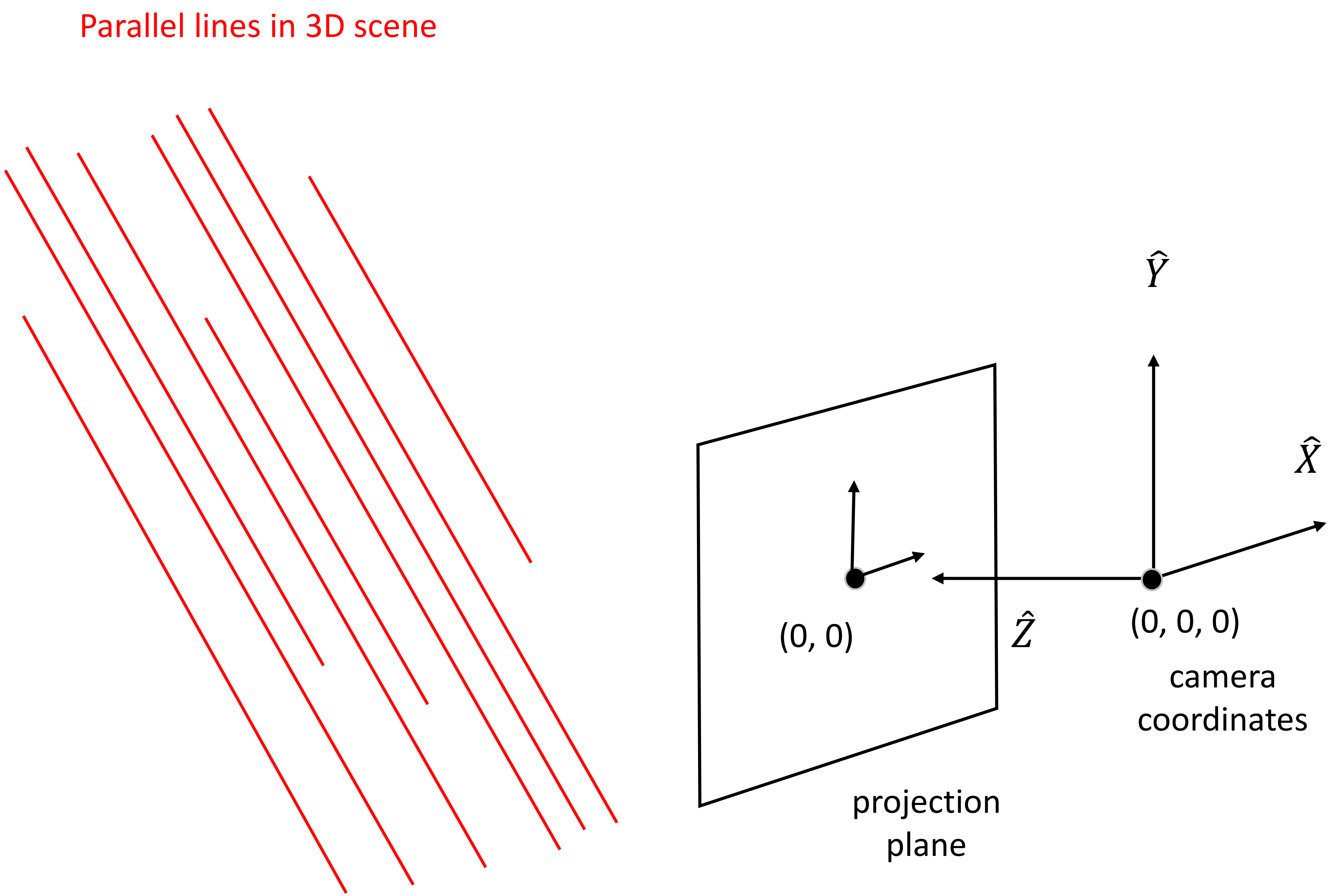
translation

vanishing points

Mon. Oct. 19, 2020

Vanishing Points

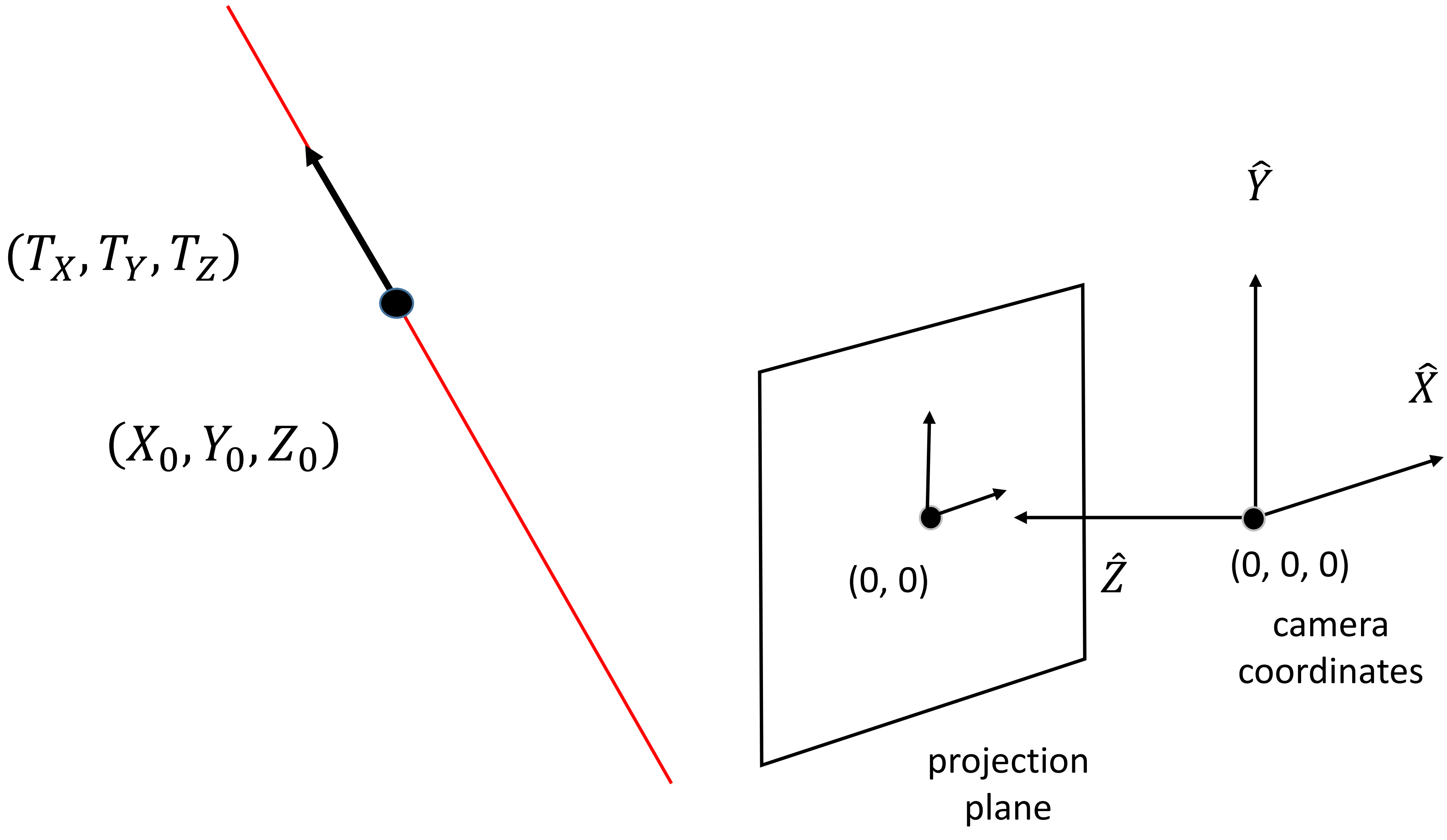
Parallel lines in the scene meet “at infinity”. The location of this meeting point *in the image projection plane* is called a vanishing point.



Parallel lines in 3D scene
$\hat{Y}$
$\hat{X}$
$\hat{Z}$
(0, 0)
(0, 0, 0)
camera coordinates
projection plane

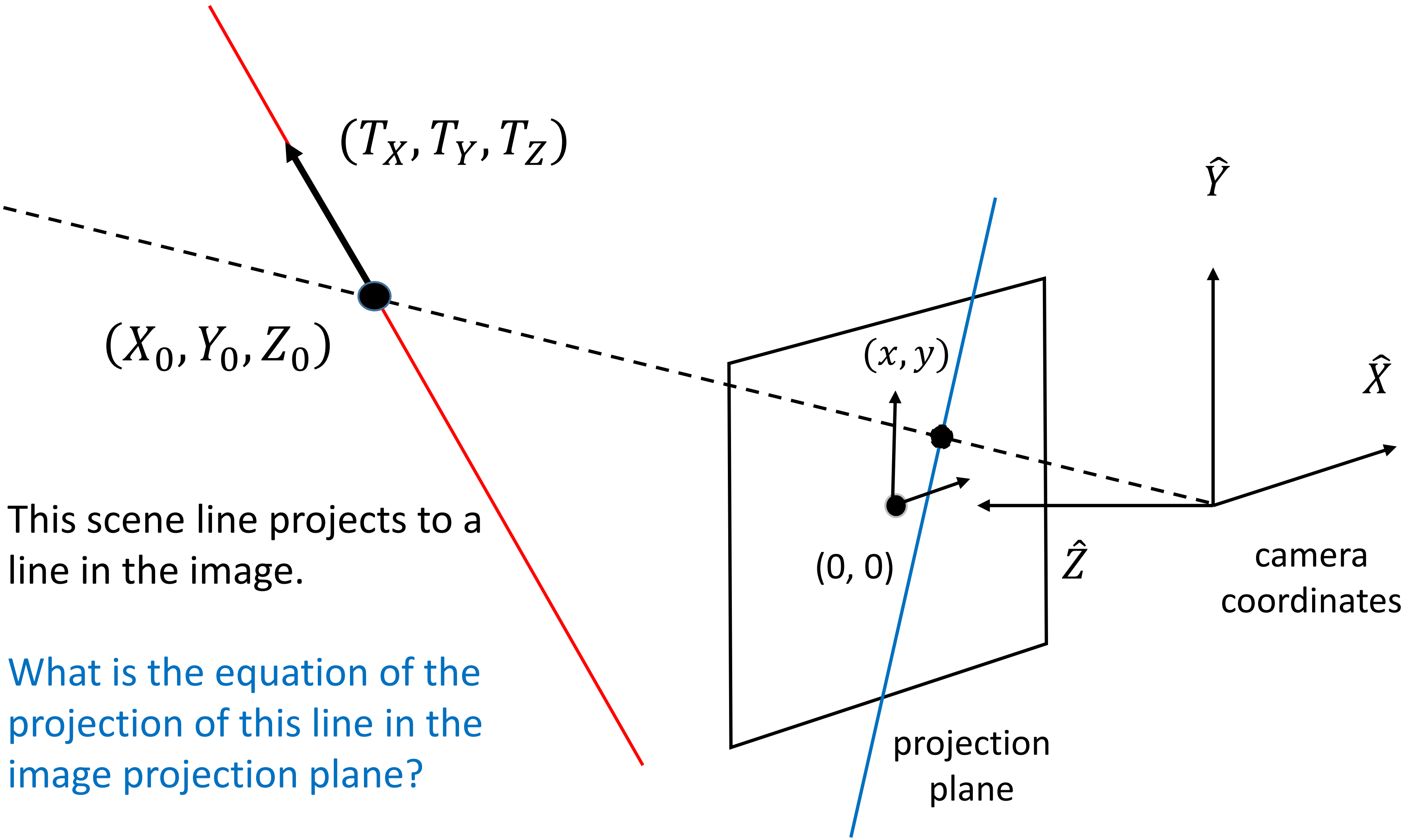
Parametric equation of a scene line

$$\big(X(t), Y(t), Z(t)\big) = (X_0, Y_0, Z_0) + (T_X, T_Y, T_Z)\, t$$



Parametric equation of a scene line

$$\big(X(t), Y(t), Z(t)\big) = \ (X_0, Y_0, Z_0) + (T_X, T_Y, T_Z)\ t$$



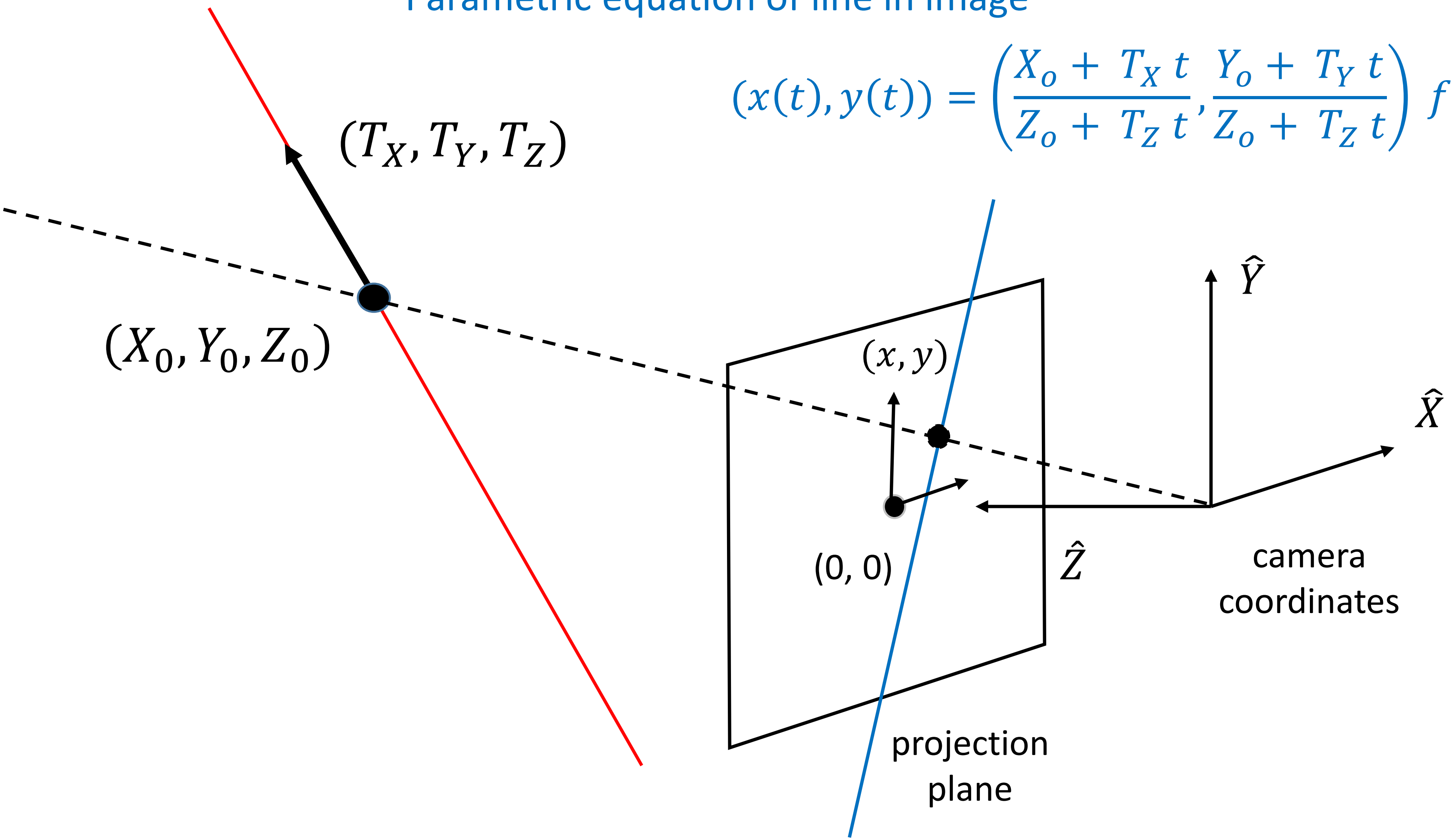
This scene line projects to a line in the image.

What is the equation of the projection of this line in the image projection plane?

$$\frac{x}{f} = \frac{X}{Z} \qquad \frac{y}{f} = \frac{Y}{Z}$$

Parametric equation of line in image

$$(x(t), y(t)) = \left(\frac{X_o + T_X\, t}{Z_o + T_Z\, t}, \frac{Y_o + T_Y\, t}{Z_o + T_Z\, t} \right) f$$

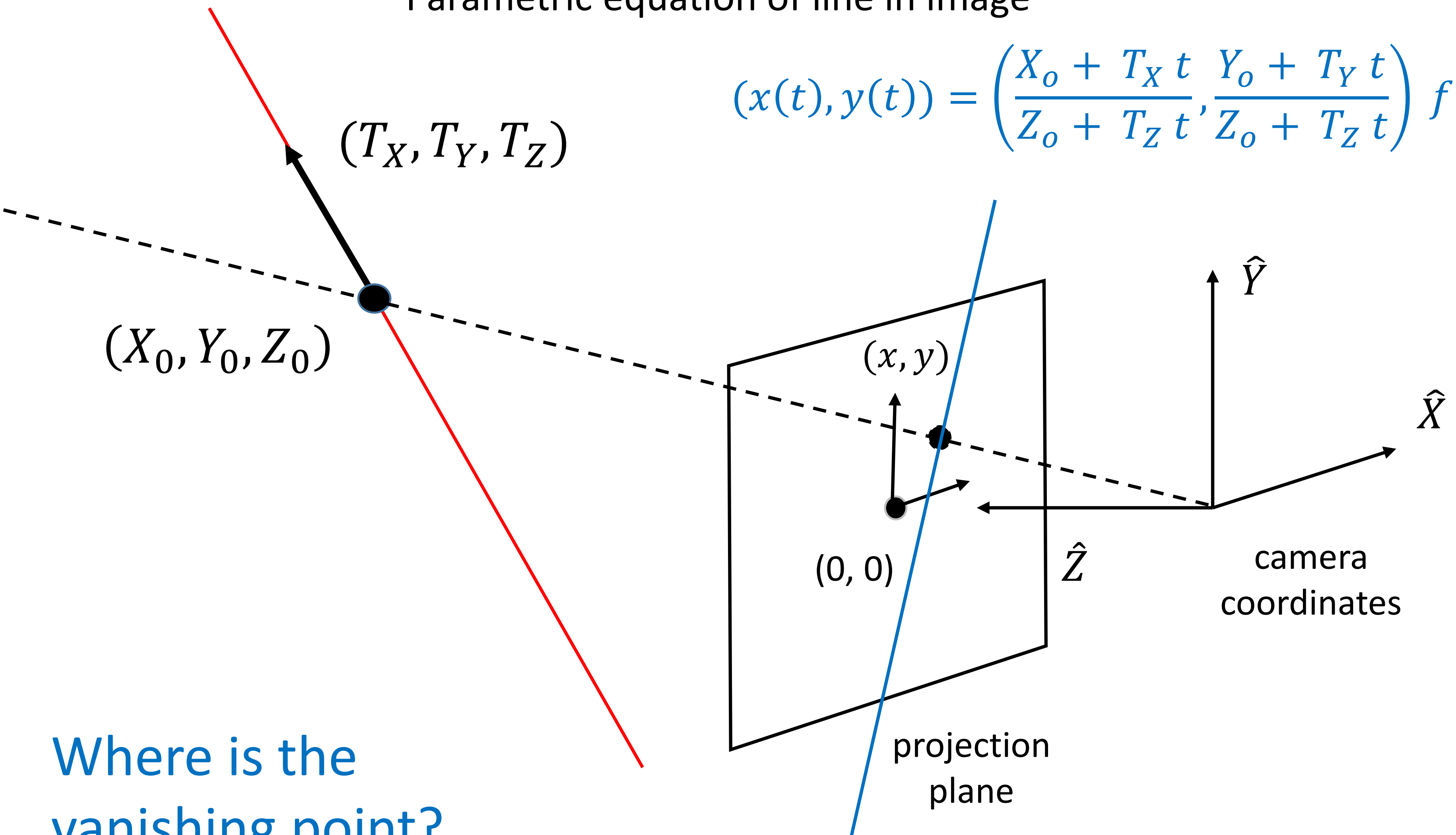


Parametric equation of scene line

$$\big(X(t), Y(t), Z(t)\big) = \ (X_0, Y_0, Z_0) + (T_X, T_Y, T_Z)\ t$$

Parametric equation of line in image

$$(x(t), y(t)) = \left(\frac{X_o + T_X\ t}{Z_o + T_Z\ t}, \frac{Y_o + T_Y\ t}{Z_o + T_Z\ t}\right) f$$

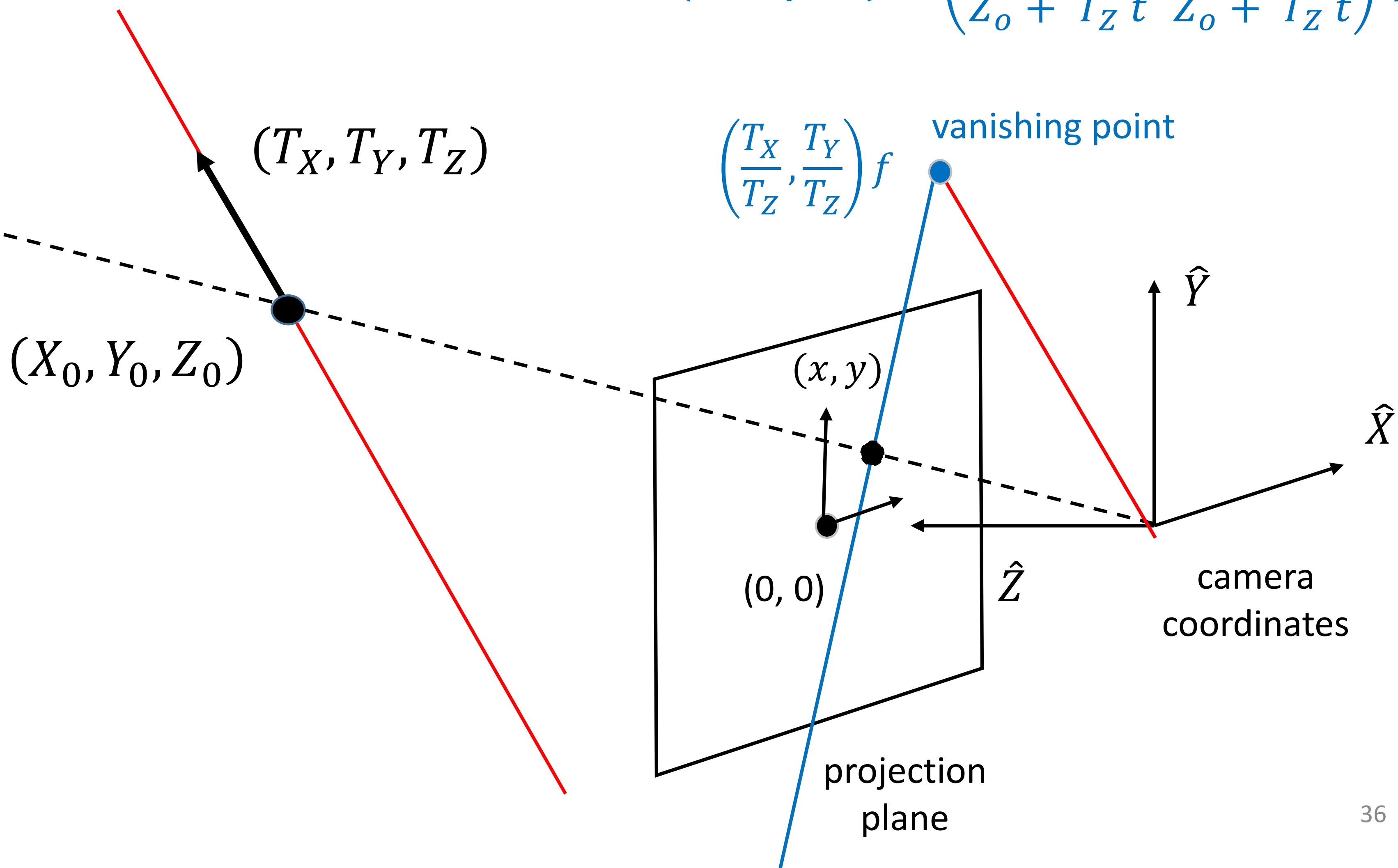


Where is the vanishing point?

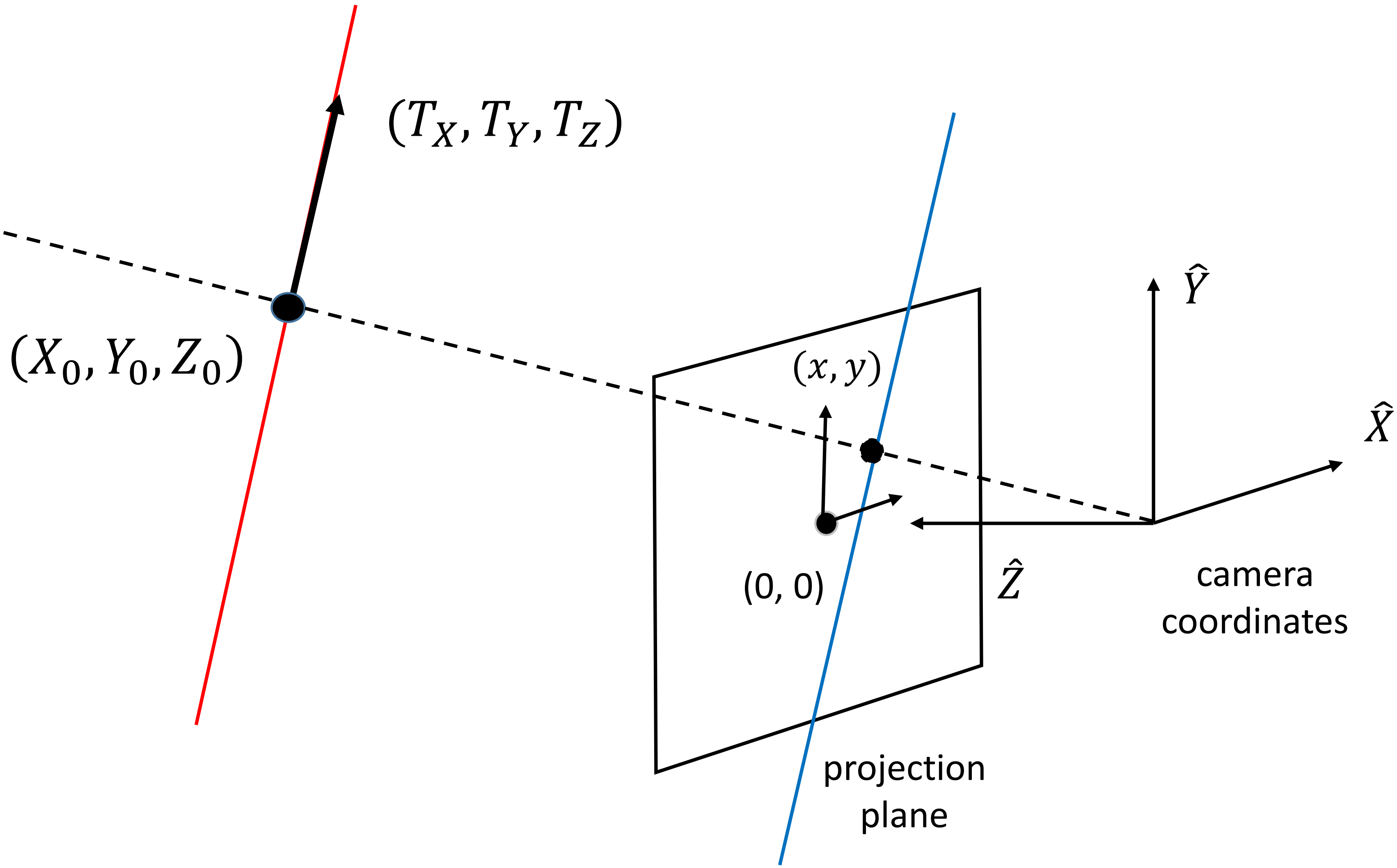
$$\big(X(t), Y(t), Z(t)\big) = \ (X_0, Y_0, Z_0) + (T_X, T_Y, T_Z)\ t$$

The vanishing point is defined by letting $t \rightarrow \infty$.

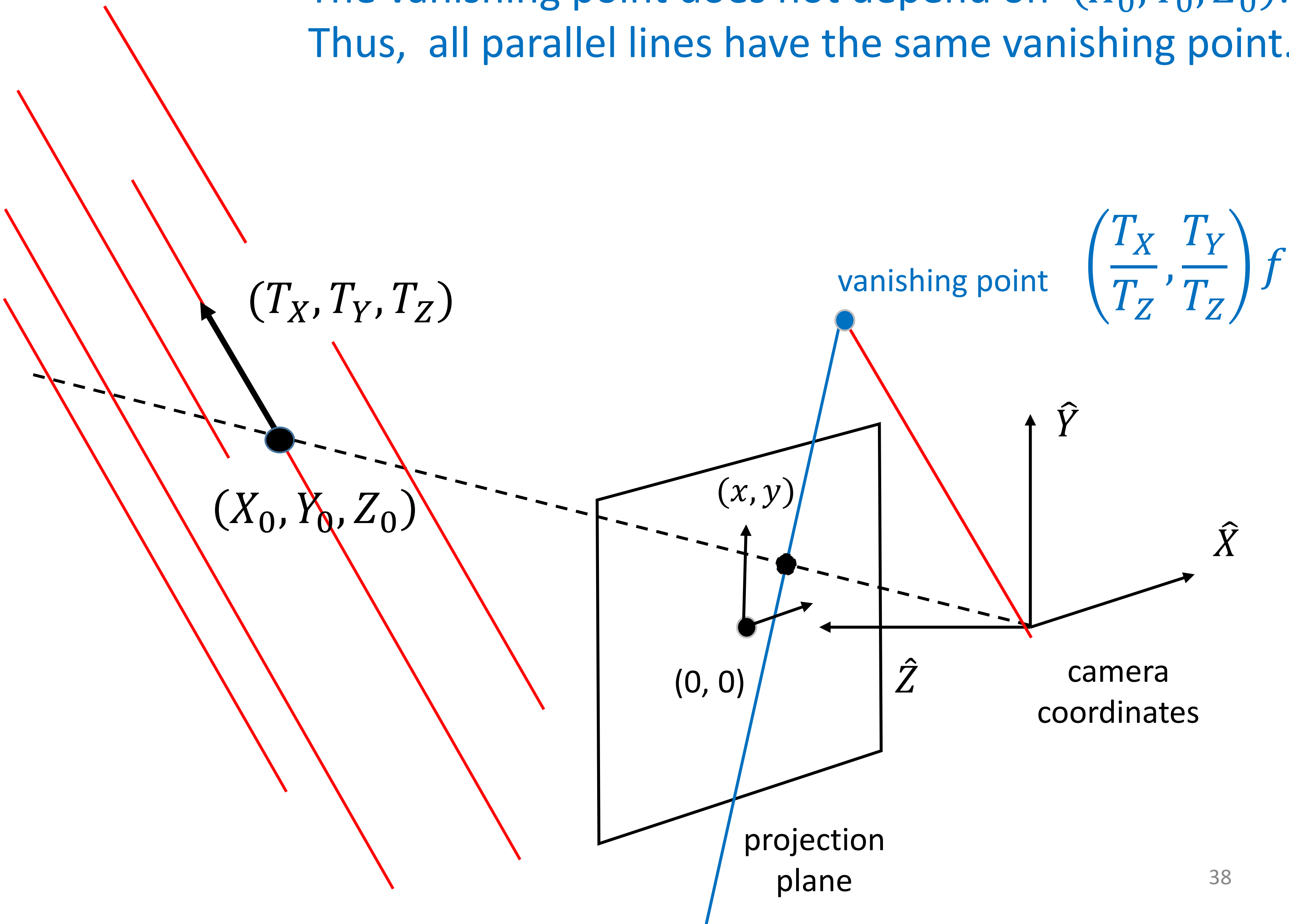
$$(x(t), y(t)) = \left(\frac{X_o + T_X\, t}{Z_o + T_Z\, t}, \frac{Y_o + T_Y\, t}{Z_o + T_Z\, t}\right) f$$



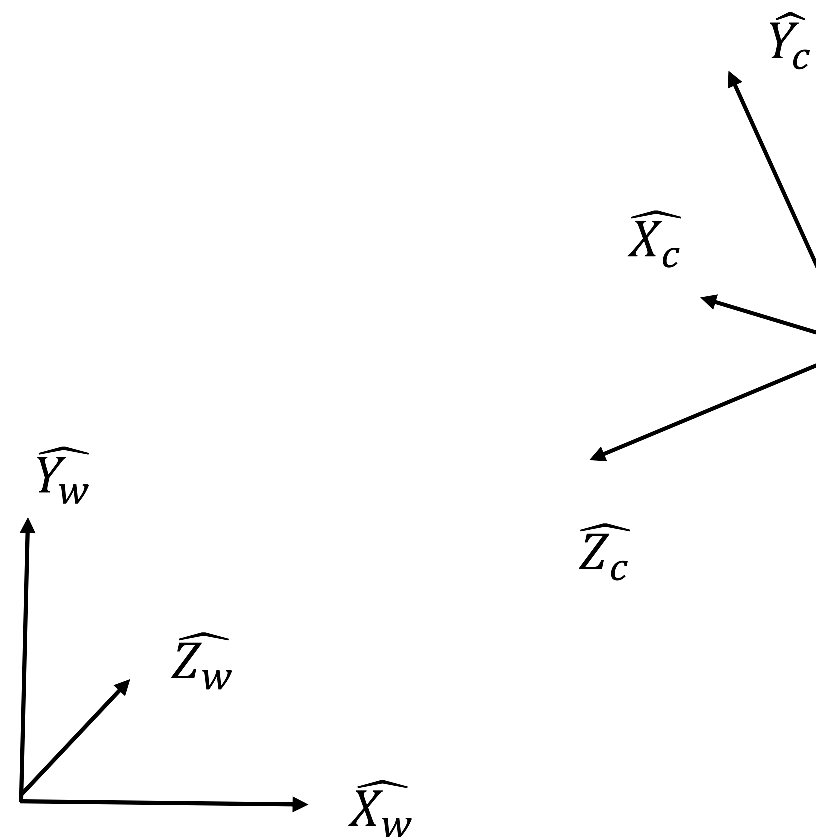
If $T_Z = 0$, then argument on previous slide doesn't work. This is the case that the scene lines are parallel to the image plane. In this case, the vanishing point is "at infinity".



The vanishing point does not depend on (X_0, Y_0, Z_0).
Thus, all parallel lines have the same vanishing point.



Camera versus World Coordinates



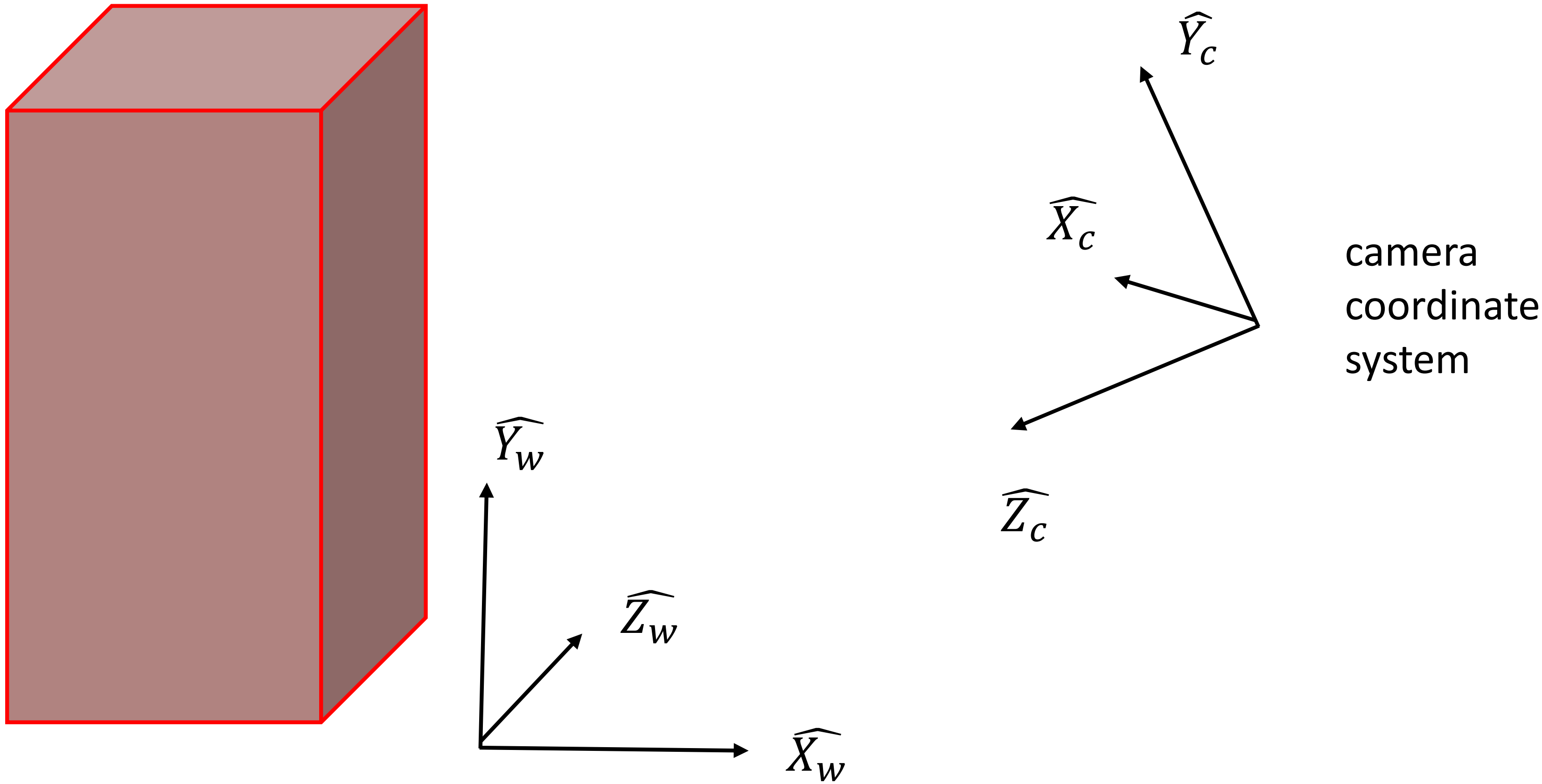
camera coordinate system

world coordinate system

Suppose a scene has three orthogonal sets of parallel lines. This is typical of manmade environments.

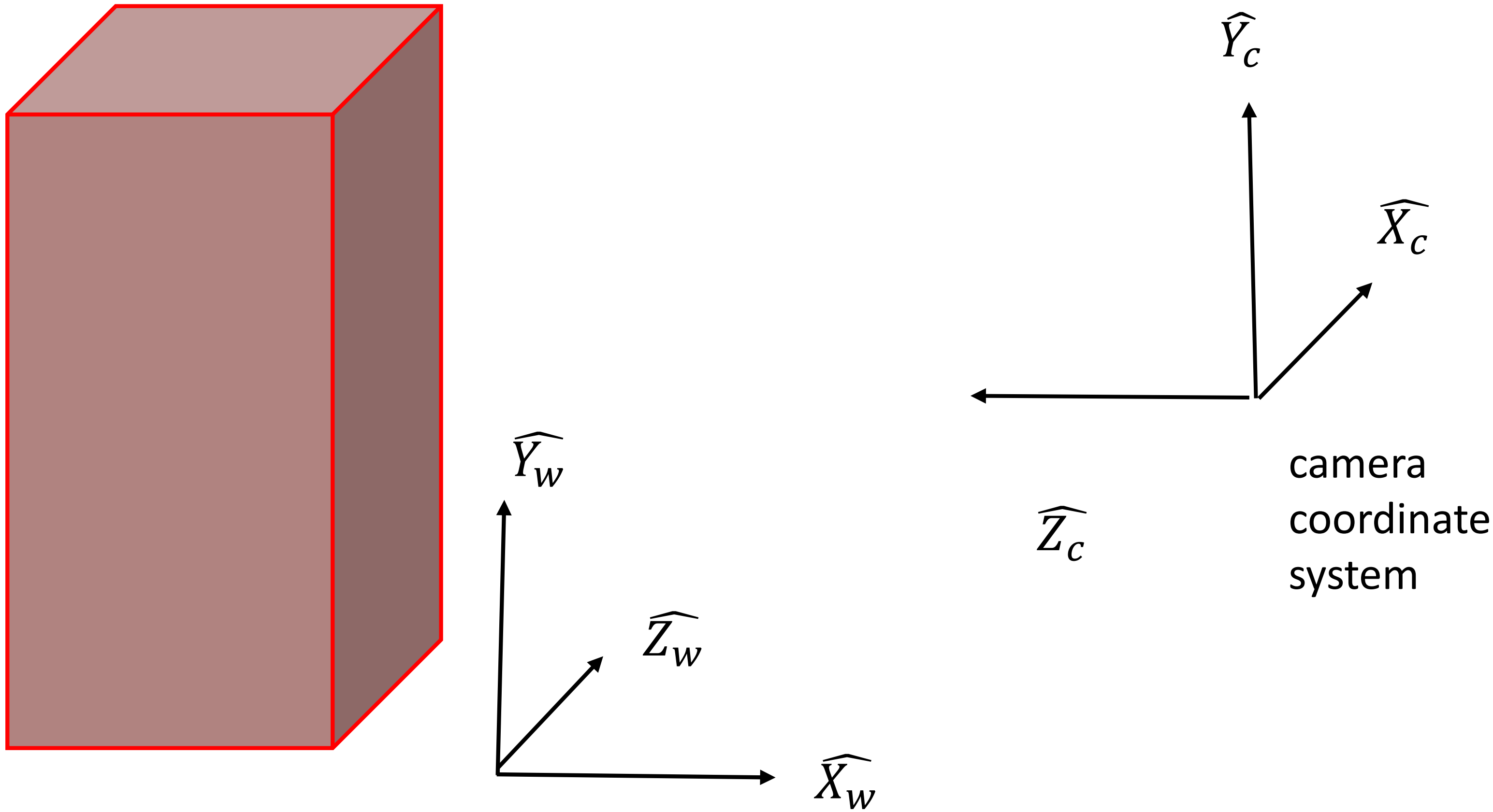
Often we define the world coordinates according to these lines.

This yields three vanishing points. How many are *finite* ?



world coordinate system

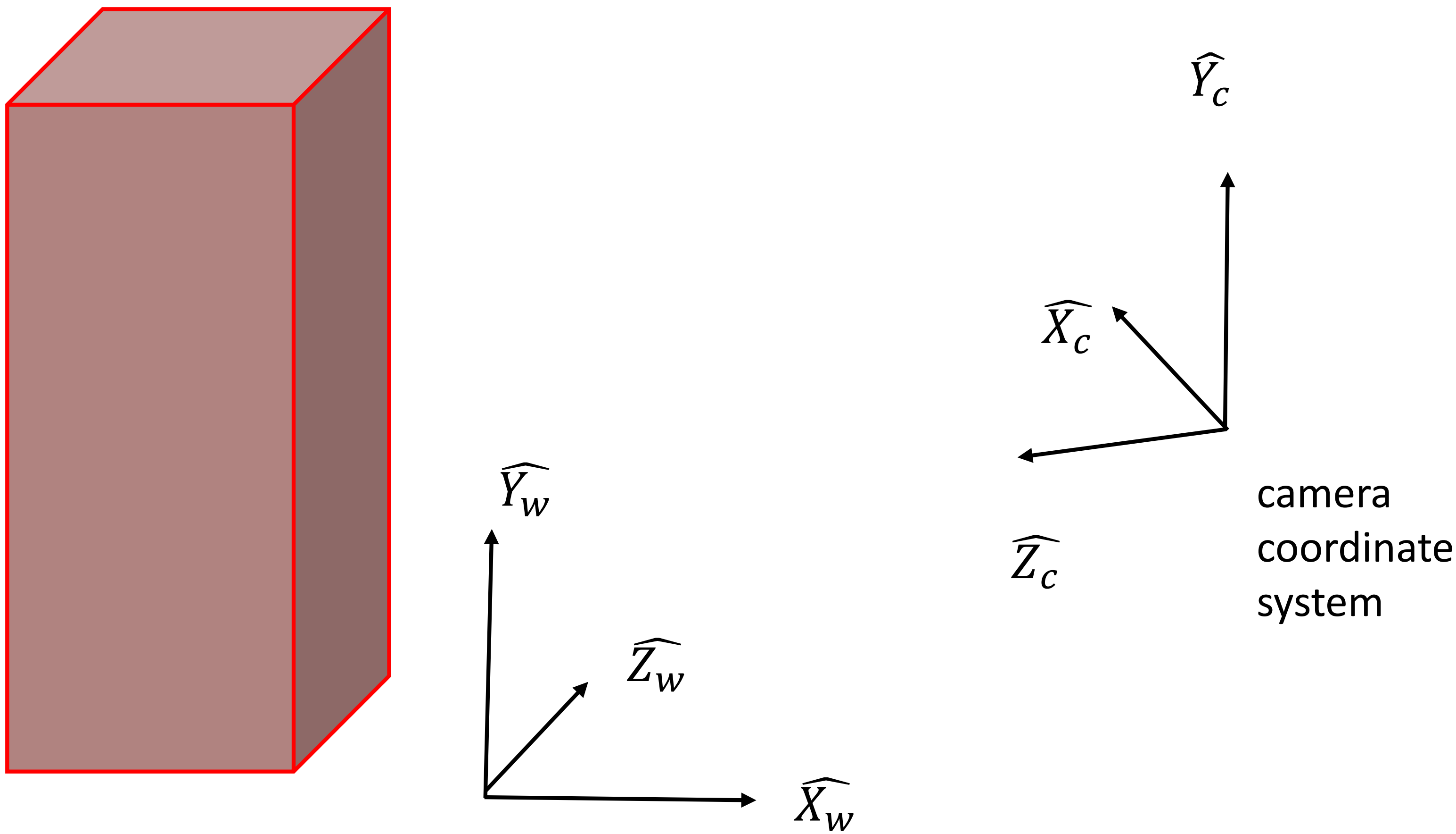
Case 1 (1 point perspective):
The camera Z axis is parallel to one of the scene axes.
In this case, the other two scene axes must be parallel to the image plane.
This yields one finite vanishing point, namely Z axis (and two infinite ones).



world coordinate system

1 Point Perspective Examples

Case 2 (2 point perspective):
Exactly one of the scene axes is parallel to the image plane (often Y).
This yields two finite vanishing points (and one infinite one).

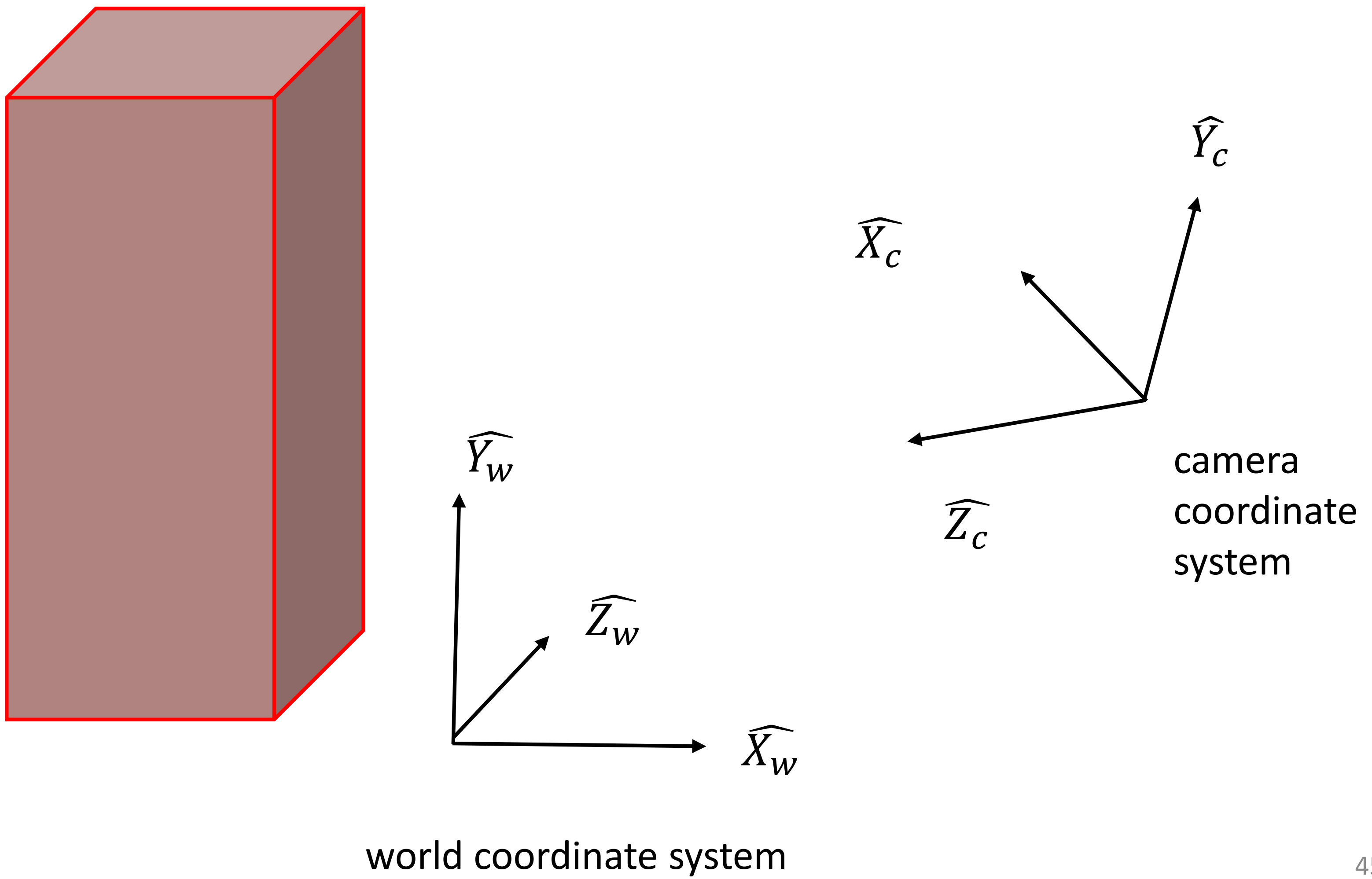


world coordinate system

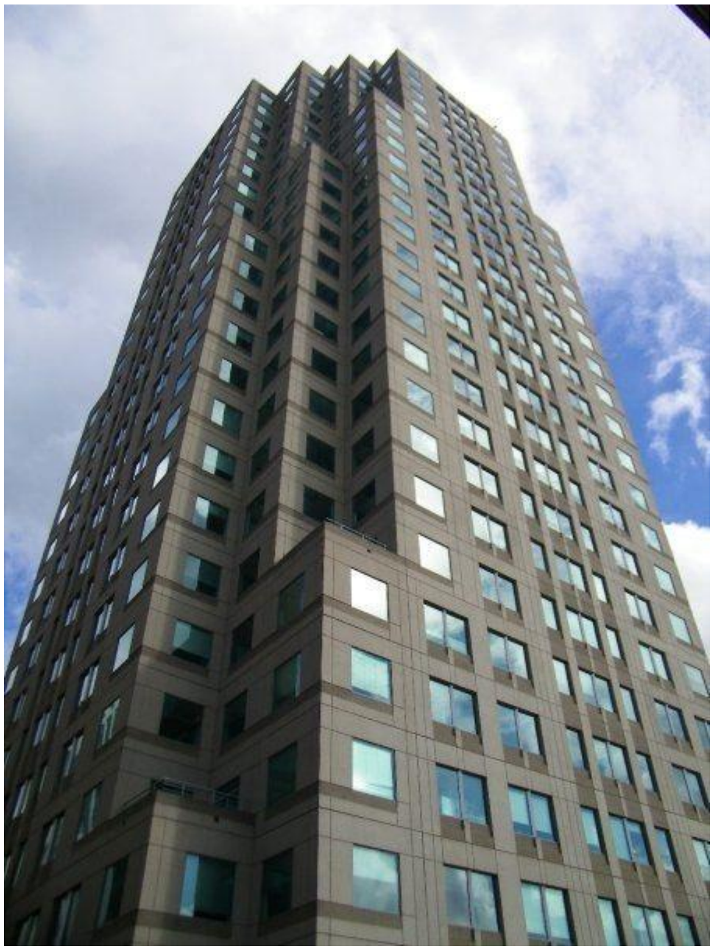
Example: 2 point perspective

The vanishing points don’t have to fall within the image window.
(The size of the window in the projection plane is a separate issue which we will discuss later.)

Case 3 (3 point perspective):
None of the scene axes is parallel to the image plane.
This yields three finite vanishing points (and no infinite ones).



Examples: 3 point perspective



COMP 558 Overview

Part 1 : 2D Vision

RGB

Image filtering

Edge detection

Least Squares Estimation

Robust Estimation: Hough transform & RANSAC

Features 1: corners

Image Registration: the Lucas-Kanade method

Scale space

Histogram-based Tracking

Features 2: SIFT

Part 2 : 3D Vision

Perspective: projection, translation, vanishing points

Rotation, Homogeneous coordinates

Camera intrinsics and Extrinsics

Least Squares methods (SVD)

Camera Calibration

Homographies & rectification

Stereo and Epipolar Geometry 1 & 2

Stereo correspondence

Features 3: CNN's

Object classification and detection

Segmentation (time permitting)

Cameras and Photography

RGBD Cameras