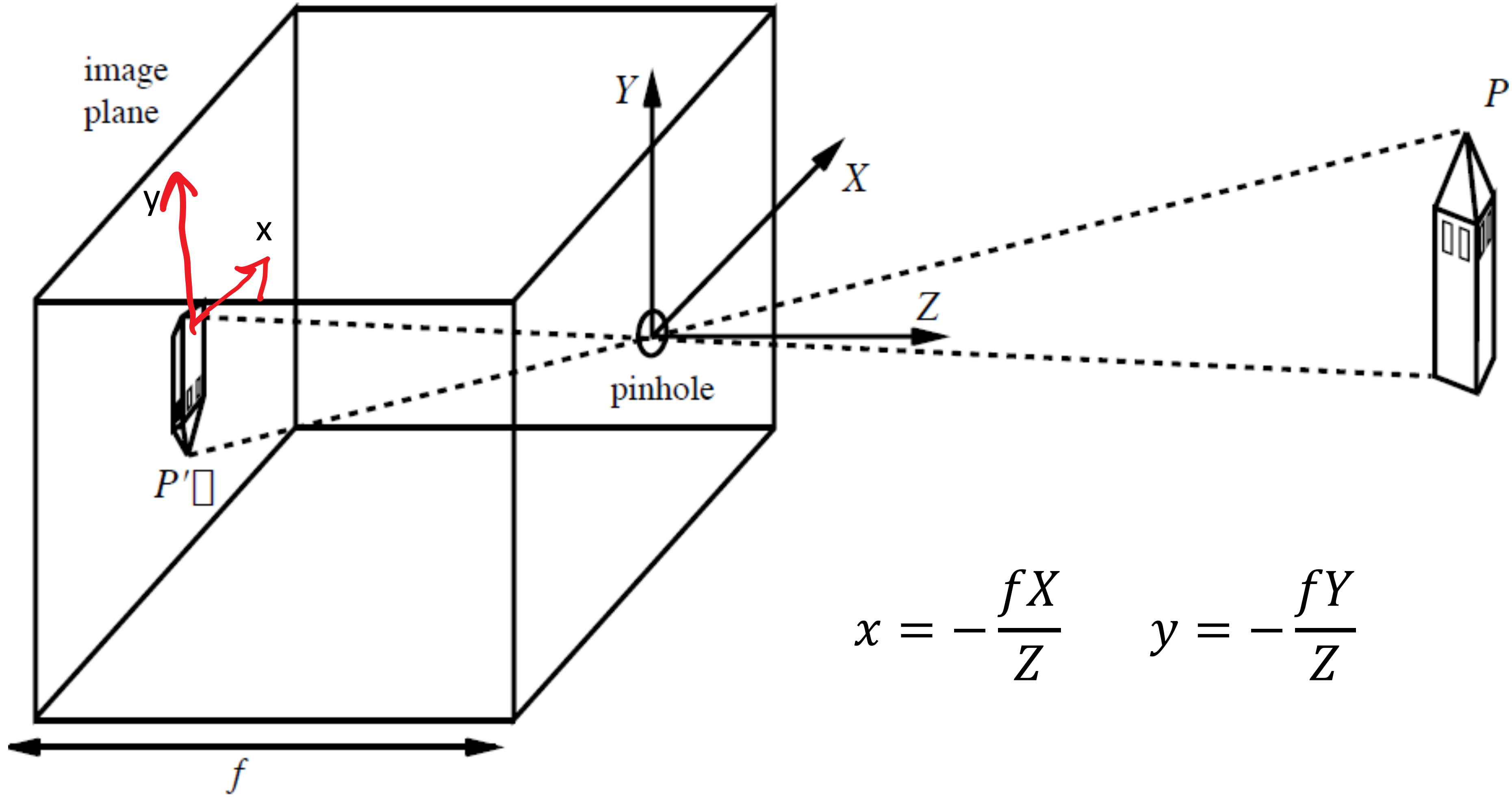
Dynamic Perspective

How a moving camera reveals scene depth and egomotion parameters

Jitendra Malik

UC Berkeley

The Pinhole Camera



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

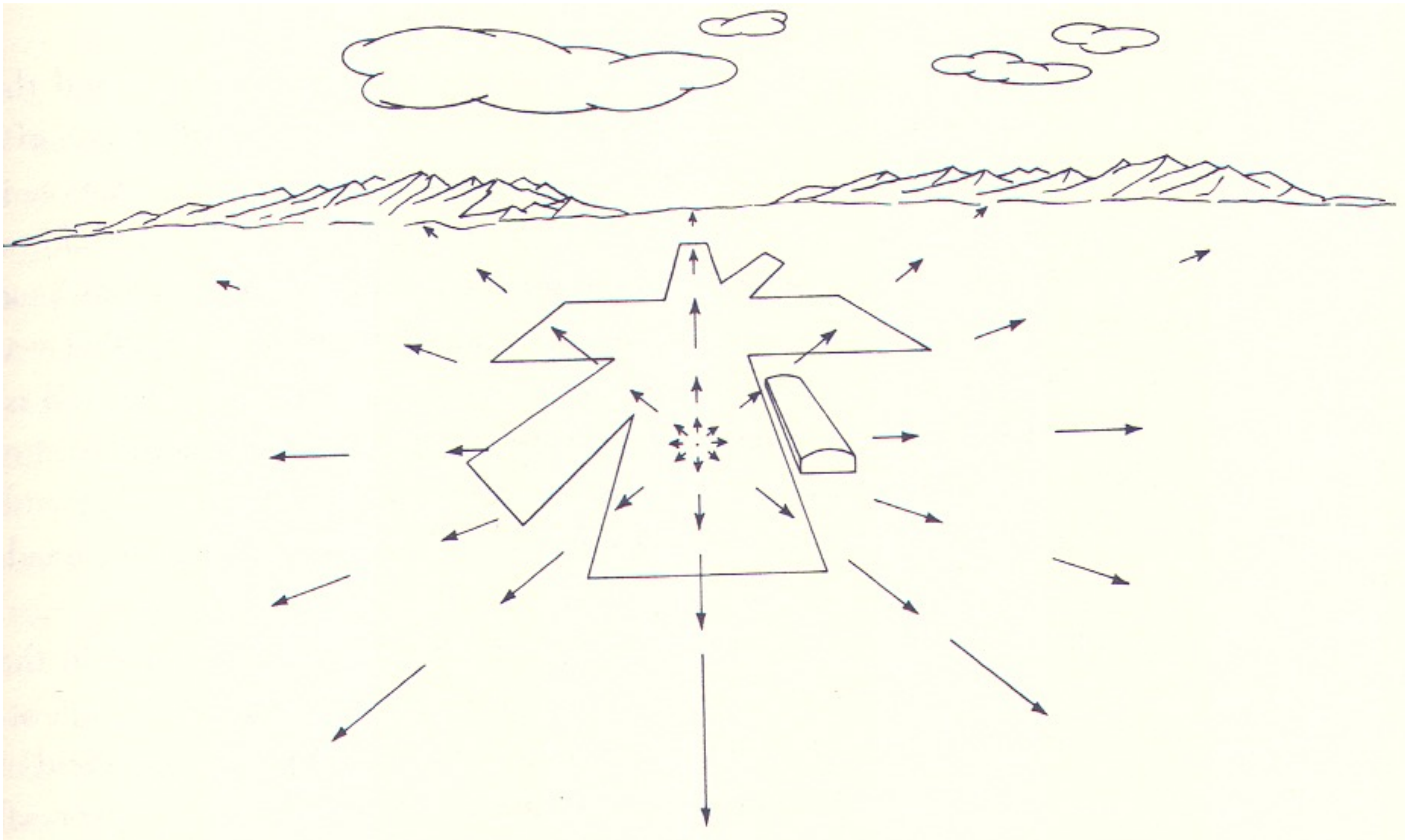
Suppose the camera moves with respect to the world…

- When a point (X,Y,Z) in the world moves relative to the camera, its projection in the image (x,y) moves as well.
- This movement in the image plane is called optical flow. Suppose the point (x,y) moves to ($x+\Delta x$, $y+\Delta y$) in time Δt, then

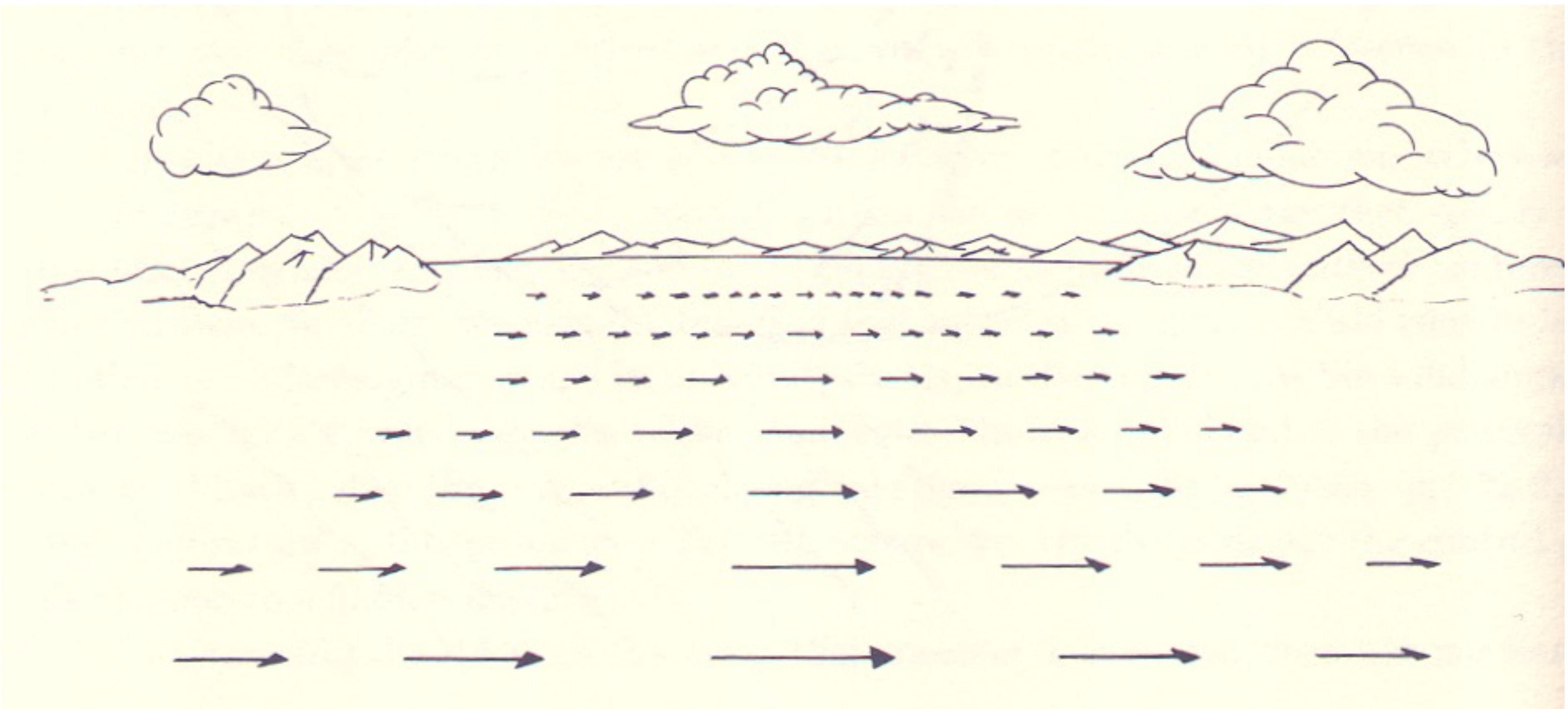
$$u = \frac{\Delta x}{\Delta t}, v = \frac{\Delta y}{\Delta t}$$

are the two components of the optical flow at (x,y)

Gibson's example I: Optical flow for a pilot landing a plane



Gibson's example II:
Optical flow from the side window of a car



Outline

$\begin{bmatrix} u \\ v \end{bmatrix} (x,y)$

- Derive equation relating optical flow field to scene depth Z(x,y) and the motion of the camera t, ω
- The translational component of the flow field is the more important one – it is what tells Z(x,y) and the translation t
- The rotational component of the flow field reveals information about ω

How does a point X in the scene move?

Assume that the camera moves with translational velocity $\mathbf{t} = (t_x, t_y, t_z)$ and angular velocity $\omega = (\omega_x, \omega_y, \omega_z)$. Eq.(3.1) is used to characterize the movement of $\mathbf{X}$,

$$\dot{\mathbf{X}} = -\mathbf{t} - \omega \wedge \mathbf{X}, \tag{3.1}$$

which can be written out in coordinates as Eq.(3.2):

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = - \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} - \begin{bmatrix} \omega_y Z - \omega_z Y \\ \omega_z X - \omega_x Z \\ \omega_x Y - \omega_y X \end{bmatrix}. \tag{3.2}$$

Now consider the effect of projection…

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = -\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} - \begin{bmatrix} \omega_y Z - \omega_z Y \\ \omega_z X - \omega_x Z \\ \omega_x Y - \omega_y X \end{bmatrix}. \tag{3.2}$$

Assume the image plane lies at $f = 1$, then $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$. Taking the derivative, we have

$$\dot{x} = \frac{\dot{X}Z - \dot{Z}X}{Z^2}, \dot{y} = \frac{\dot{Y}Z - \dot{Z}Y}{Z^2}. \tag{3.3}$$

$$x = \frac{X}{Z} \quad \dot{x} = \dot{\left(\frac{X}{Z}\right)} = \frac{\dot{X}Z - \dot{Z}X}{Z^2}$$

$$\dot{x} = \frac{\{-t_x - (w_y Z - w_z Y)\}Z - \{-t_z - (w_x Y - w_y X)\}X}{Z^2}$$

$$= -\frac{t_x}{Z} - w_y + w_z \frac{Y}{Z} \ldots$$

Now consider the effect of projection…

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = -\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} - \begin{bmatrix} \omega_y Z - \omega_z Y \\ \omega_z X - \omega_x Z \\ \omega_x Y - \omega_y X \end{bmatrix}. \tag{3.2}$$

Assume the image plane lies at $f = 1$, then $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$. Taking the derivative, we have

$$\dot{x} = \frac{\dot{X}Z - \dot{Z}X}{Z^2}, \dot{y} = \frac{\dot{Y}Z - \dot{Z}Y}{Z^2}. \tag{3.3}$$

Substitute $\dot{X}, \dot{Y}, \dot{Z}$ in Eq.(3.3) using Eq.(3.2), plug in $x = \frac{X}{Z}, y = \frac{Y}{Z}$, and simplify it, we get

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \frac{1}{Z}\begin{bmatrix} -1 & 0 & x \\ 0 & -1 & y \end{bmatrix}\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} + \begin{bmatrix} xy & -(1+x^2) & y \\ 1+y^2 & -xy & -x \end{bmatrix}\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \tag{3.4}$$

The Optical Flow Equations

$$\underbrace{\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \frac{1}{Z}\begin{bmatrix} -1 & 0 & x \\ 0 & -1 & y \end{bmatrix}\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}}_{\text{TRANSLATION COMPONENT}} + \underbrace{\begin{bmatrix} xy & -(1+x^2) & y \\ 1+y^2 & -xy & -x \end{bmatrix}\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}}_{\text{ROTATION COMPONENT}}$$

$$\dot{x} = \frac{1}{Z}\left(-t_x + x t_z\right) + xy\,w_x - (1+x^2)\,w_y + y\,w_z$$

$$\dot{y} = \frac{1}{Z}\left(-t_y + y t_z\right) + (1+y^2)\,w_x - xy\,w_y - x\,w_z$$

Optical flow for pure translation

If the motion of the camera is purely translational, the terms due to rotation in Eq. (3.4) can be dropped and the flow field becomes

$$u(x,y) = \frac{-t_x + xt_z}{Z(x,y)}, v(x,y) = \frac{-t_y + yt_z}{Z(x,y)}. \tag{3.5}$$

Optical flow for pure translation along Z axis

If the motion of the camera is purely translational, the terms due to rotation in Eq. (3.4) can be dropped and the flow field becomes

$$u(x,y) = \frac{-t_x + xt_z}{Z(x,y)}, v(x,y) = \frac{-t_y + yt_z}{Z(x,y)}. \tag{3.5}$$

We can gain intuition by considering the even more special case of translation along the optical axis, i.e. $t_z \neq 0, t_x = 0, t_y = 0$, the flow field in Eq.(3.5) becomes

$$u(x,y) = \frac{xt_z}{Z(x,y)}, v(x,y) = \frac{yt_z}{Z(x,y)}; \tag{3.6}$$

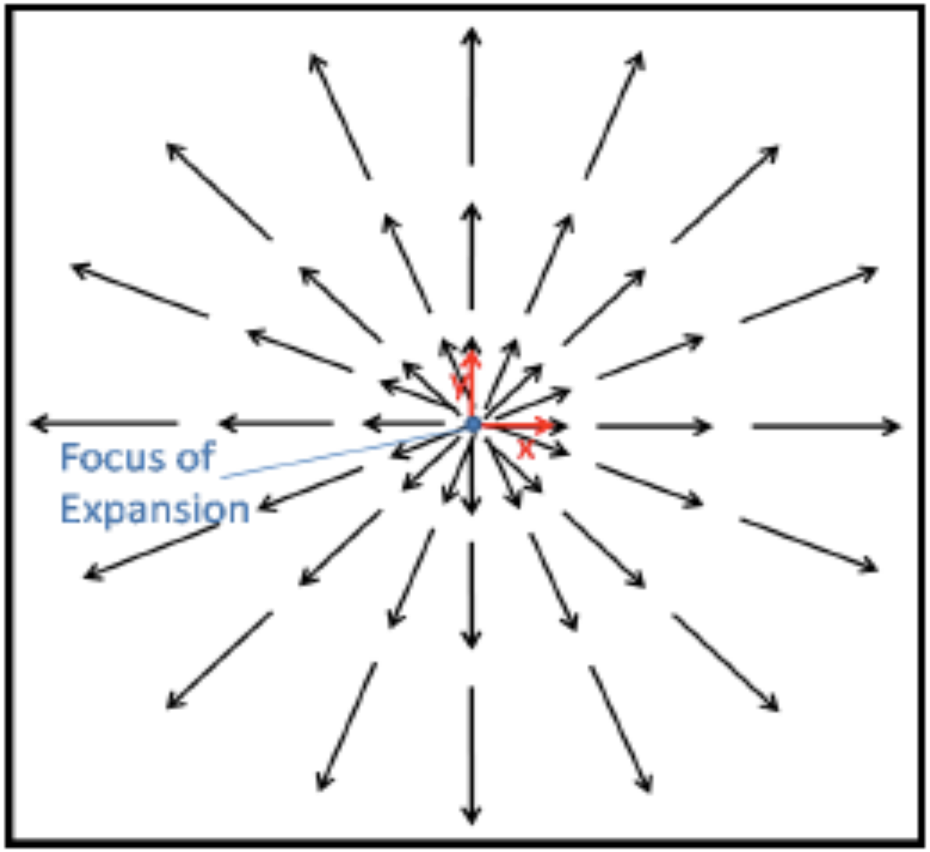
or equivalently

$$[u,v]^T(x,y) = \frac{t_z}{Z}[x,y]^T \tag{3.7}$$

Optical flow for pure translation along Z-axis

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{t_Z}{Z} \begin{bmatrix} x \\ y \end{bmatrix}$$

· $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ at the origin



The optical flow vector is a scalar multiple of the position vector

Scale Factor Ambiguity

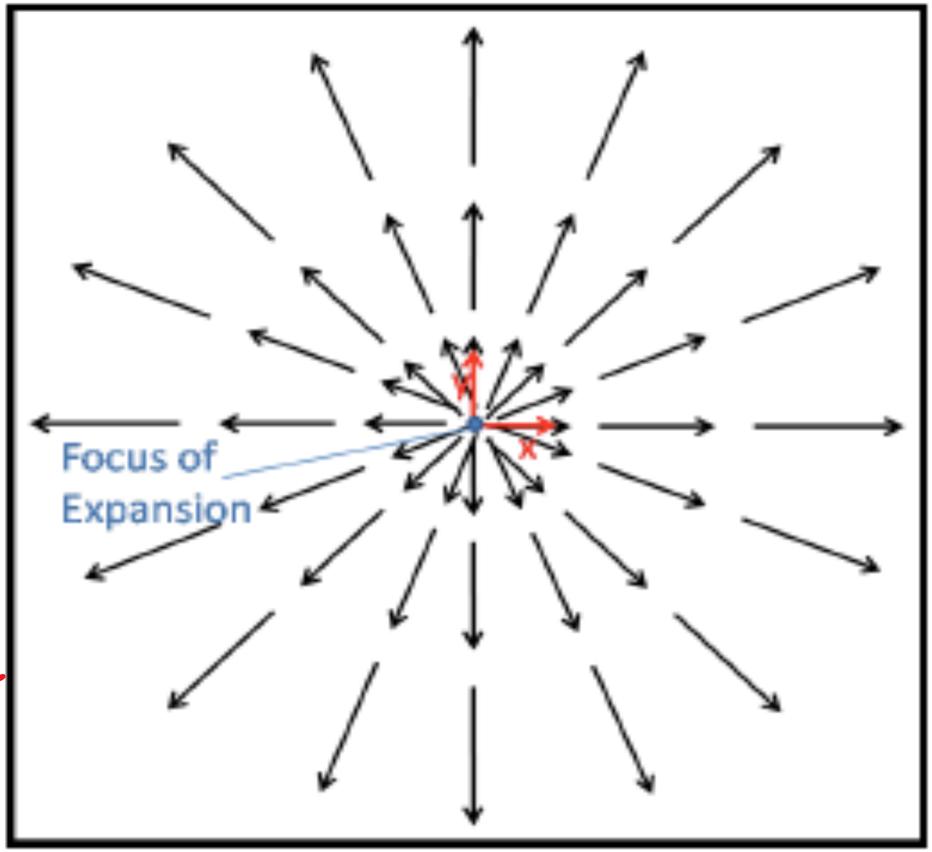
$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{t_z}{Z}\begin{bmatrix} x \\ y \end{bmatrix}$$

If

$t_z \mapsto k t_z$

$Z \mapsto k Z$

$\begin{bmatrix} u \\ v \end{bmatrix}$ unchanged



However we can compute the time to contact

Optical flow for general translation

$$u(x,y) = \frac{-t_x + xt_z}{Z(x,y)}, v(x,y) = \frac{-t_y + yt_z}{Z(x,y)}.$$

When is this (0,0)?

$u = \dfrac{-t_x + x t_z}{Z(x,y)}$ $\quad u = 0 \Rightarrow \dfrac{-t_x + x t_z}{Z} = 0 \Rightarrow x = \dfrac{t_x}{t_z}$

Similarly $v = 0$ if $y = \dfrac{t_y}{t_z}$

$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ at point $\left(\dfrac{t_x}{t_z}, \dfrac{t_y}{t_z}\right)$

With respect to the FOE, the flow vectors are radially outward

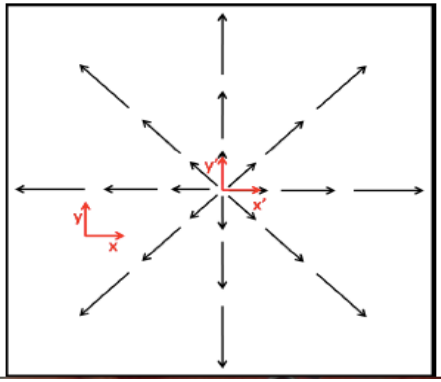
Suppose we change the origin to the FOE by applying the following co-ordinate change to Eq.(3.5),

$$x' = x - \frac{t_x}{t_z}, y' = y - \frac{t_y}{t_z}, \tag{3.9}$$

then the optical flow field becomes

$$[u, v]^T(x', y') = \frac{t_z}{Z}[x', y']^T. \tag{3.10}$$

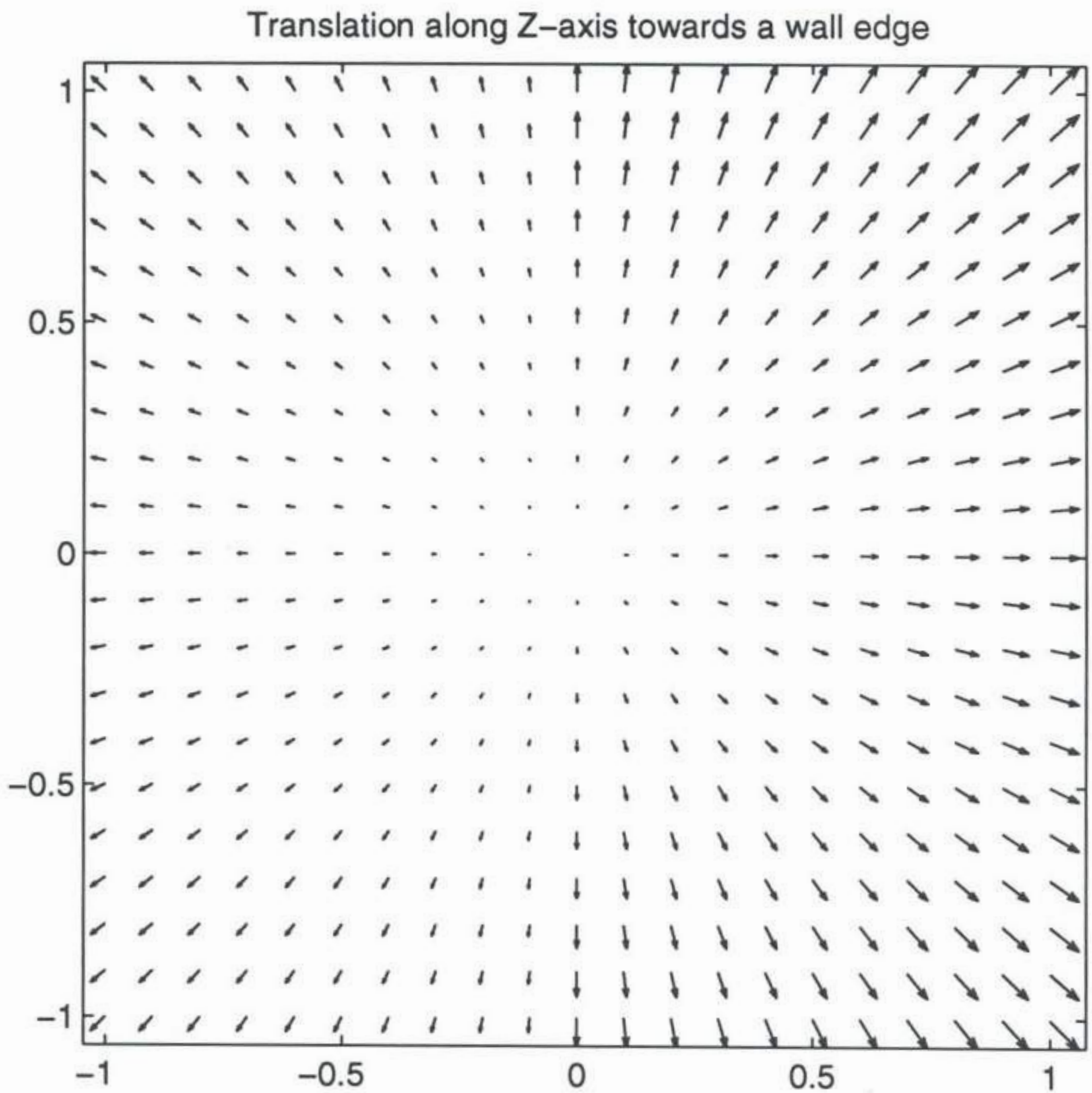
which should look very familiar. Thus the general case too corresponds to optical flow vectors pointing outwards from the FOE, justifying the choice of the term. Figure 3.3 shows such an optical vector field.

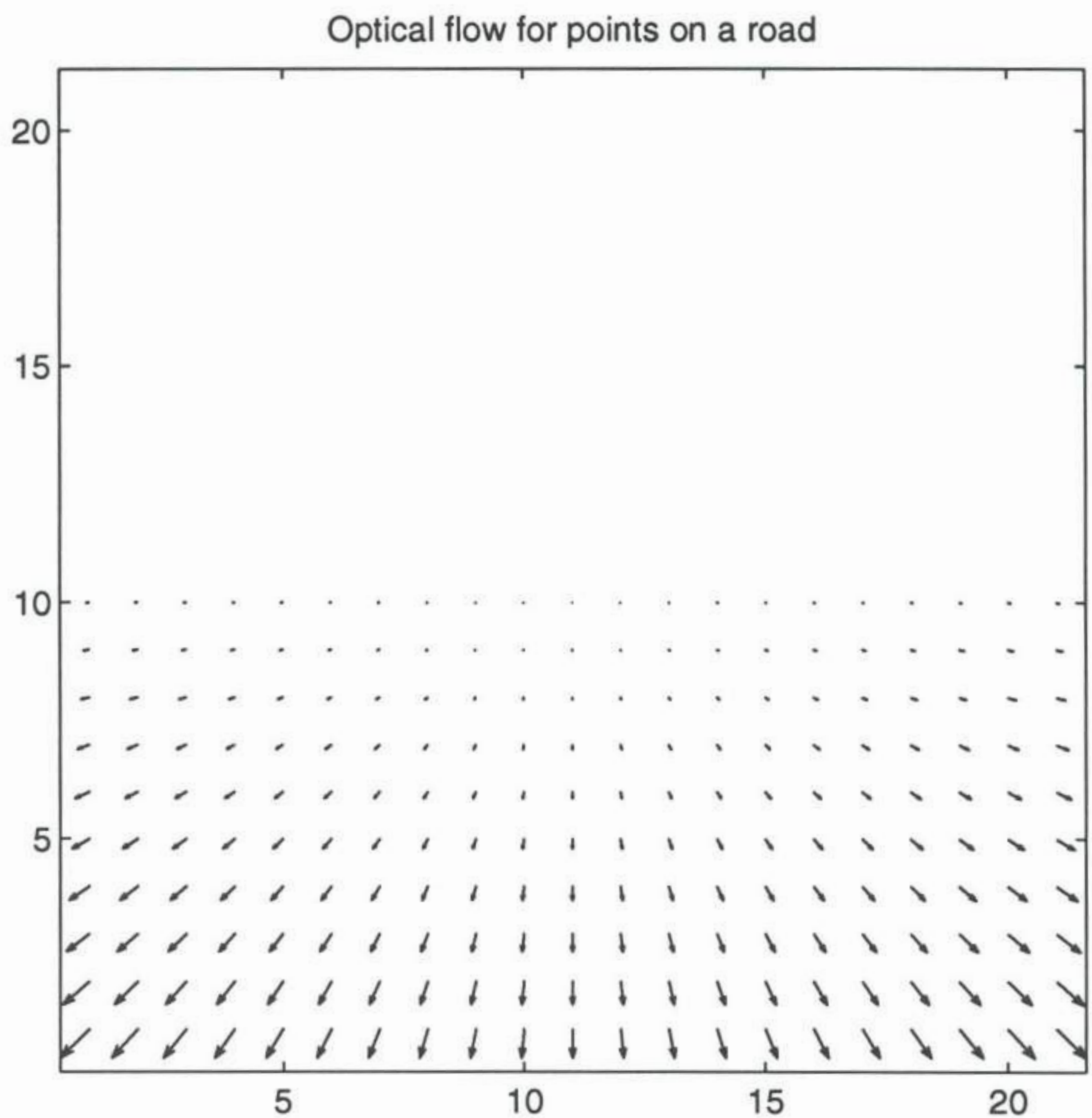


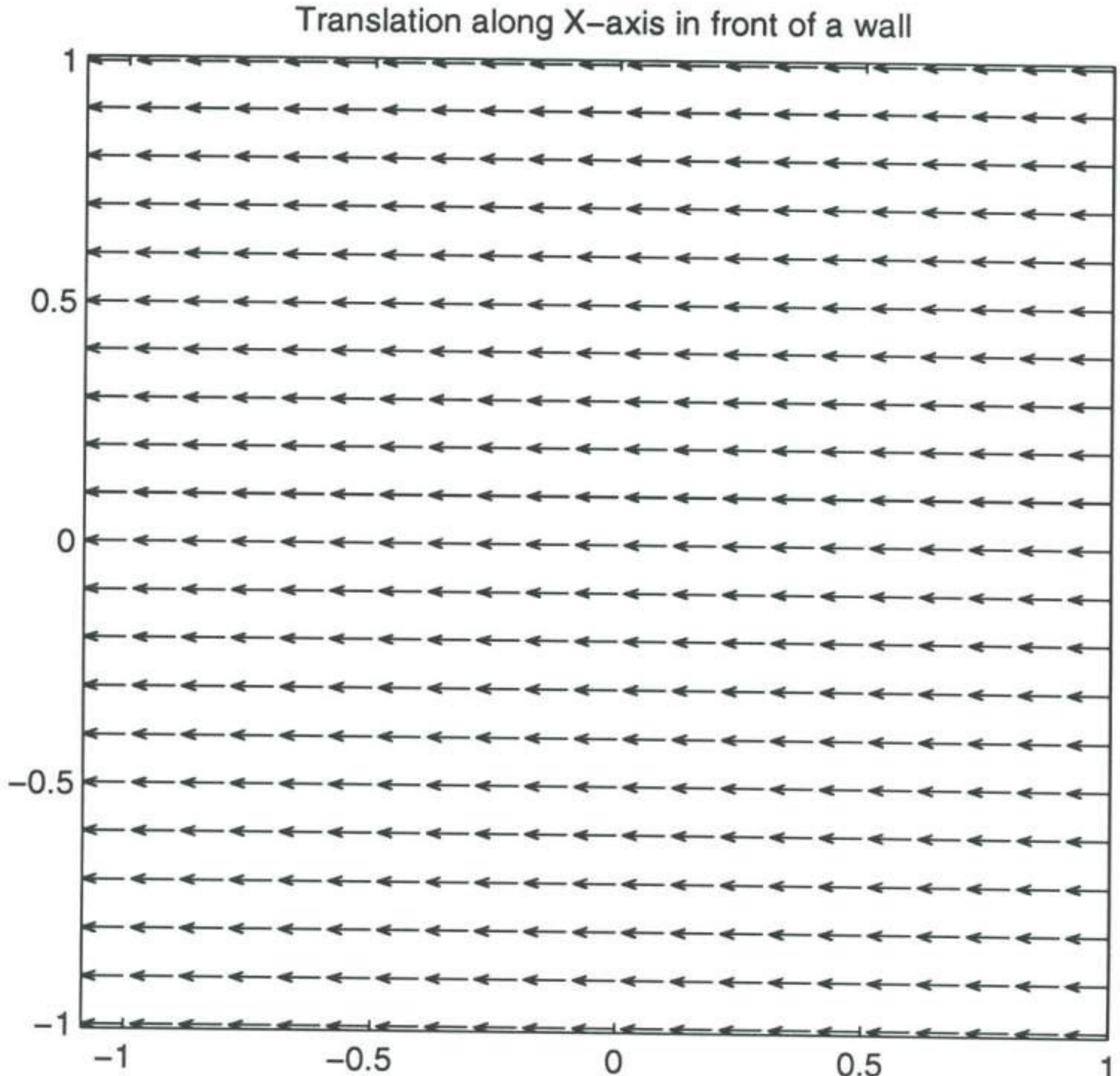
The Optical Flow Equations

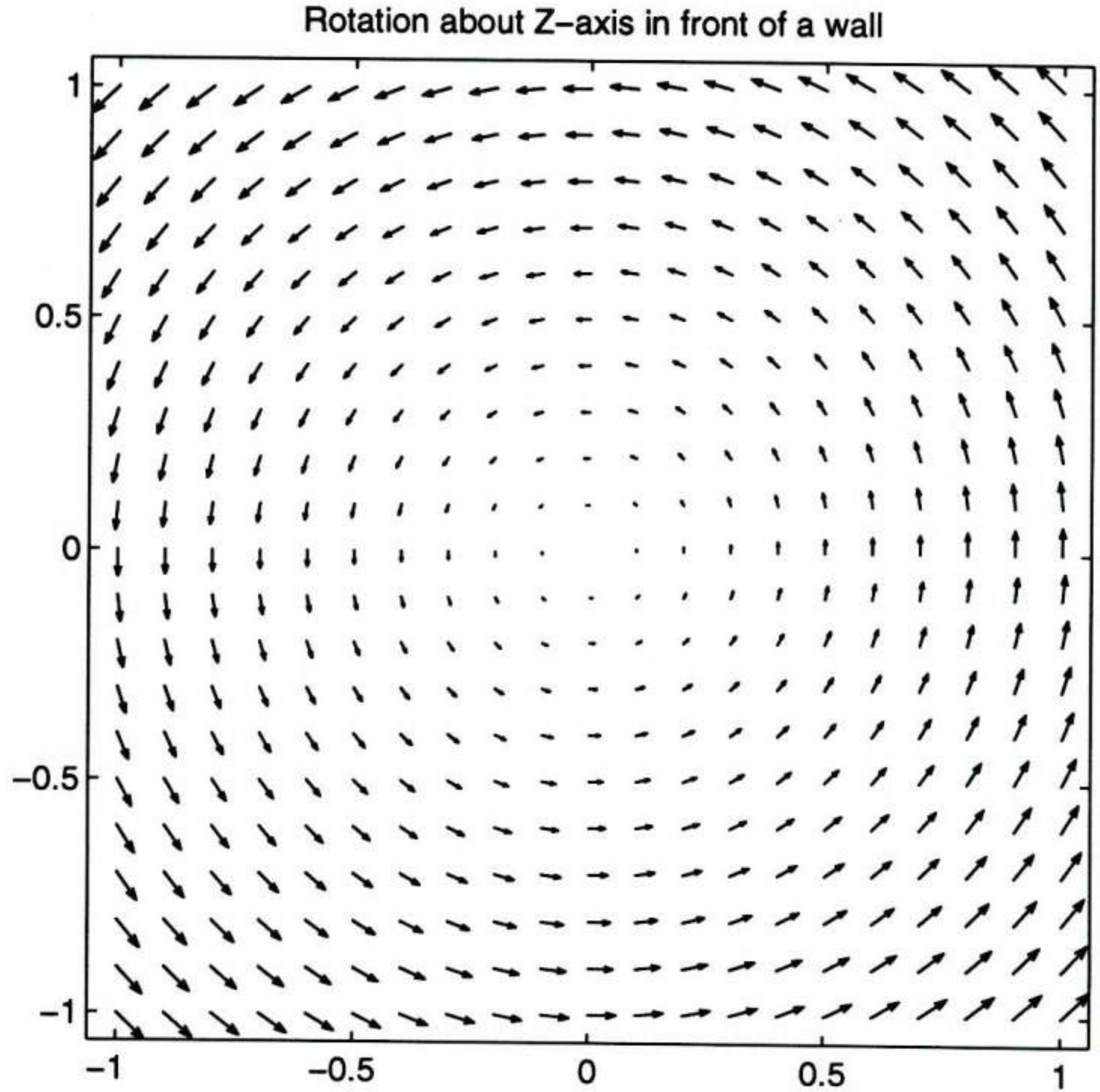
$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \underbrace{\frac{1}{Z}\begin{bmatrix} -1 & 0 & x \\ 0 & -1 & y \end{bmatrix}\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}}_{\text{TRANSLATION COMPONENT}} + \underbrace{\begin{bmatrix} xy & -(1+x^2) & y \\ 1+y^2 & -xy & -x \end{bmatrix}\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}}_{\text{ROTATION COMPONENT}}$$

If the motion is pure rotation, we can determine $\underset{\sim}{\omega}$ from the flow field.

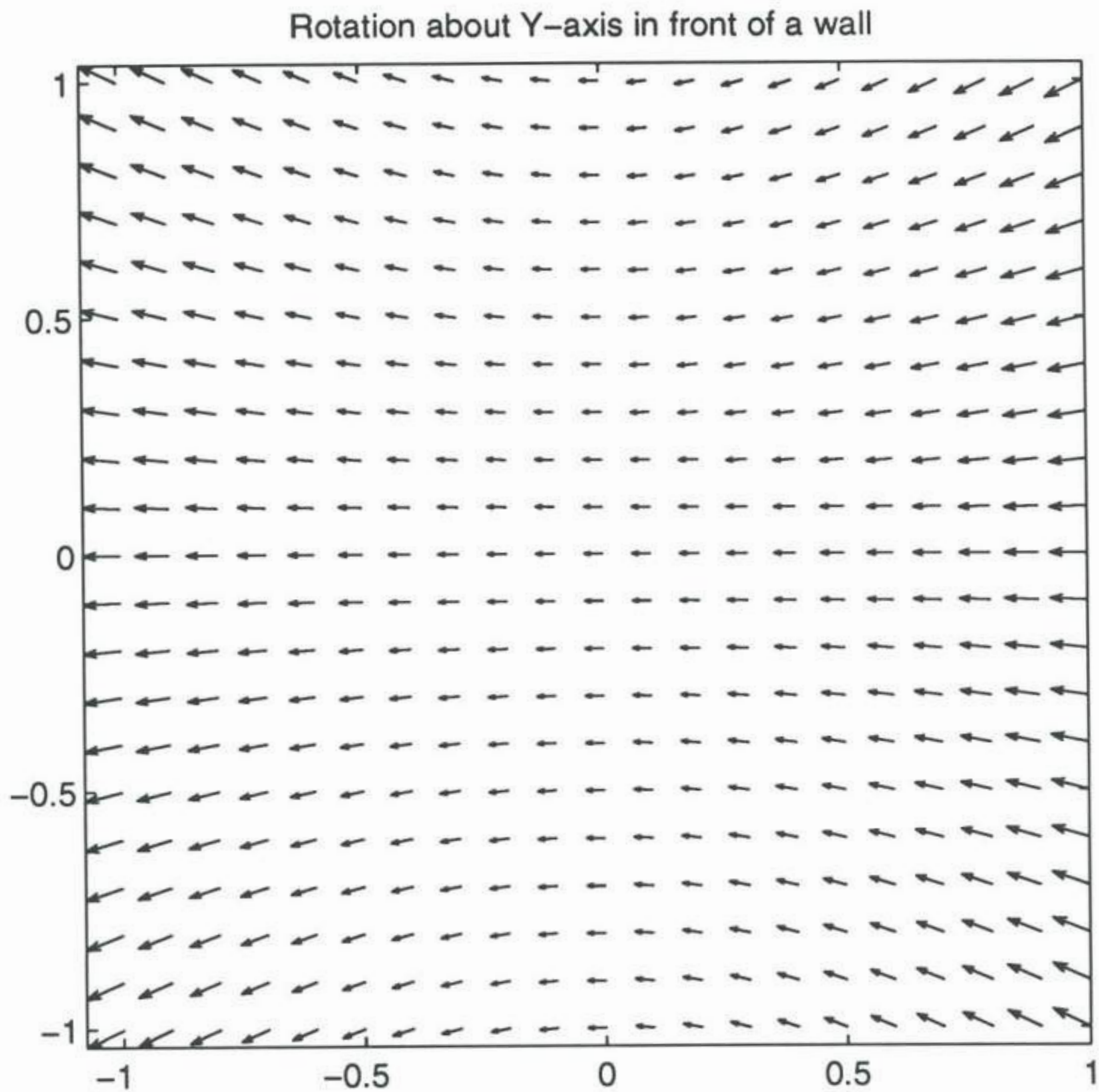

Translation along Z–axis towards a wall edge
1
0.5
0
–0.5
–1
–1
–0.5
0
0.5
1


Optical flow for points on a road
20
15
10
5
5
10
15
20


Translation along X–axis in front of a wall
1
0.5
0
–0.5
–1
–1
–0.5
0
0.5
1



Rotation about Z–axis in front of a wall
1
0.5
0
−0.5
−1
−1
−0.5
0
0.5
1



Rotation about Y−axis in front of a wall
1
0.5
0
−0.5
−1
−1
−0.5
0
0.5
1