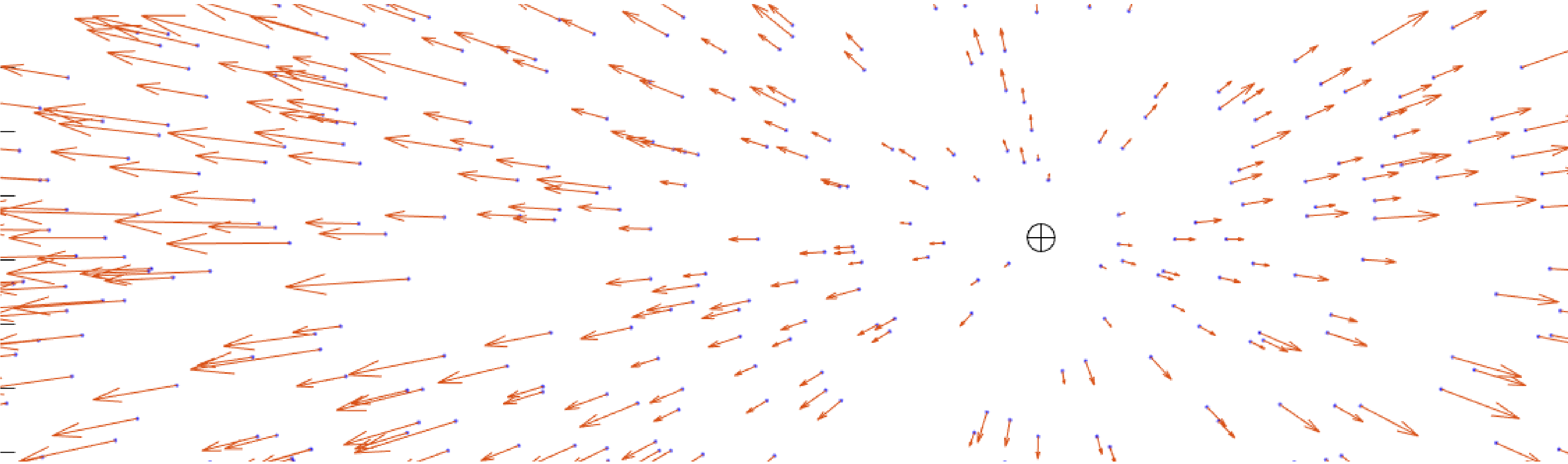
Perception:
3D Velocities from Optical Flow

Kostas Daniilidis

Which direction is this quadrotor moving?

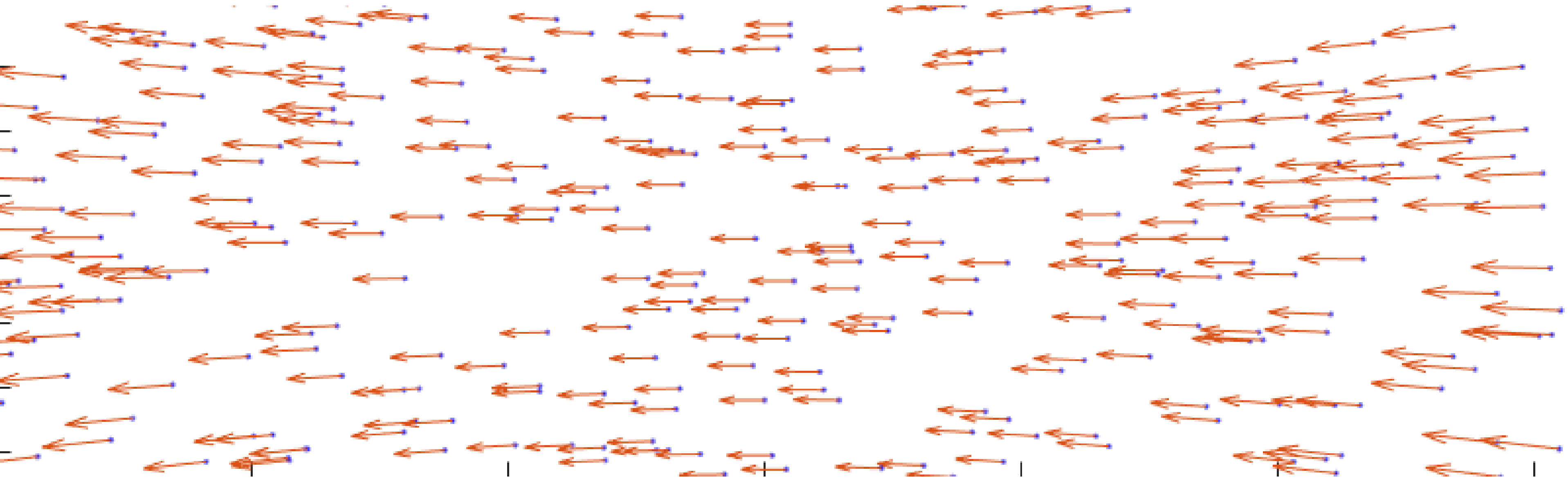
Which direction is this car moving?

If the camera translates only



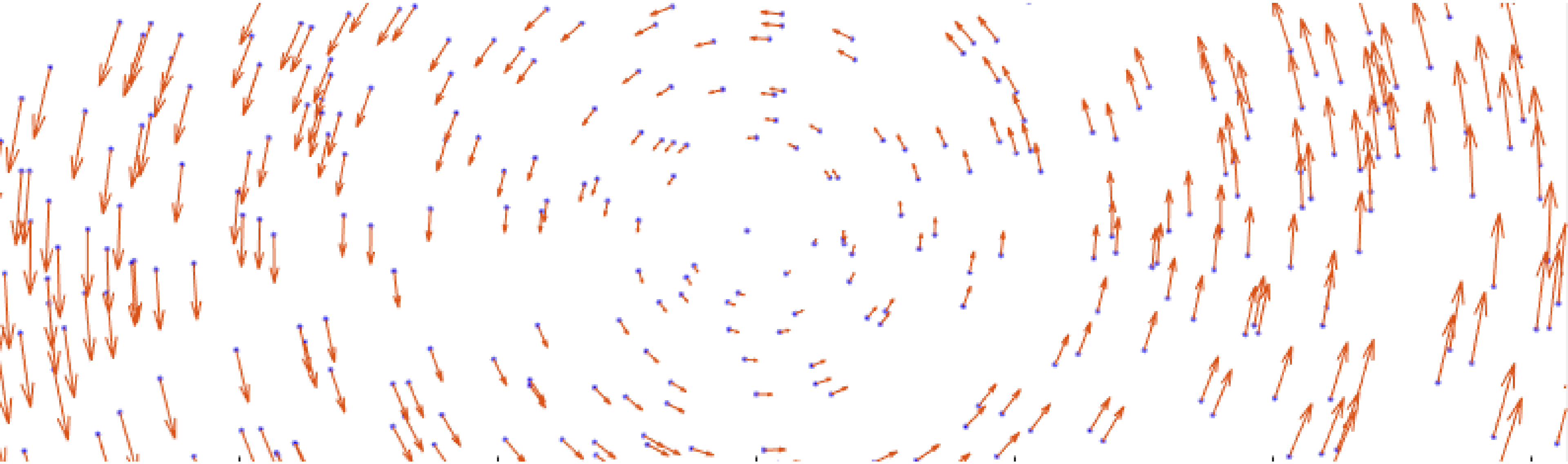
… the vector field is radially expanding
from a point called the Focus of Expansion and we
can quite easily infer our direction of motion.

If the camera rotates only ..



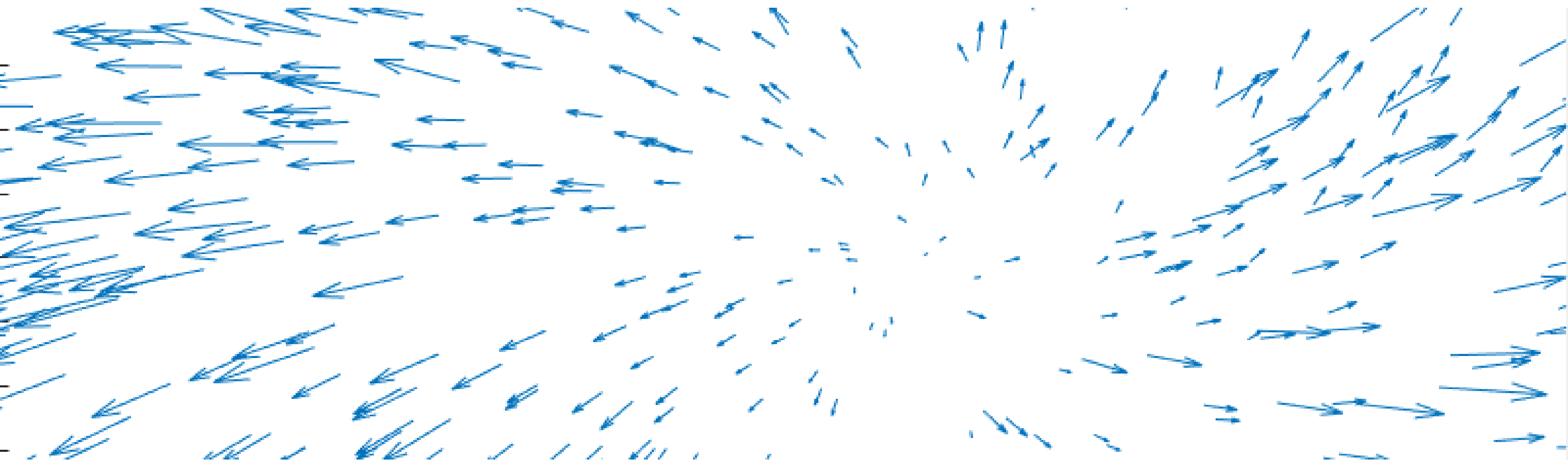
…. for example around the vertical axis the pattern is almost horizontal and vectors are longer left and right.

If the camera rotates only around optical axis



.... for example around the optical axis
we obtain a curling vector field.

If we combine translation and rotation..



…. It is hard to tell whereto we are moving!

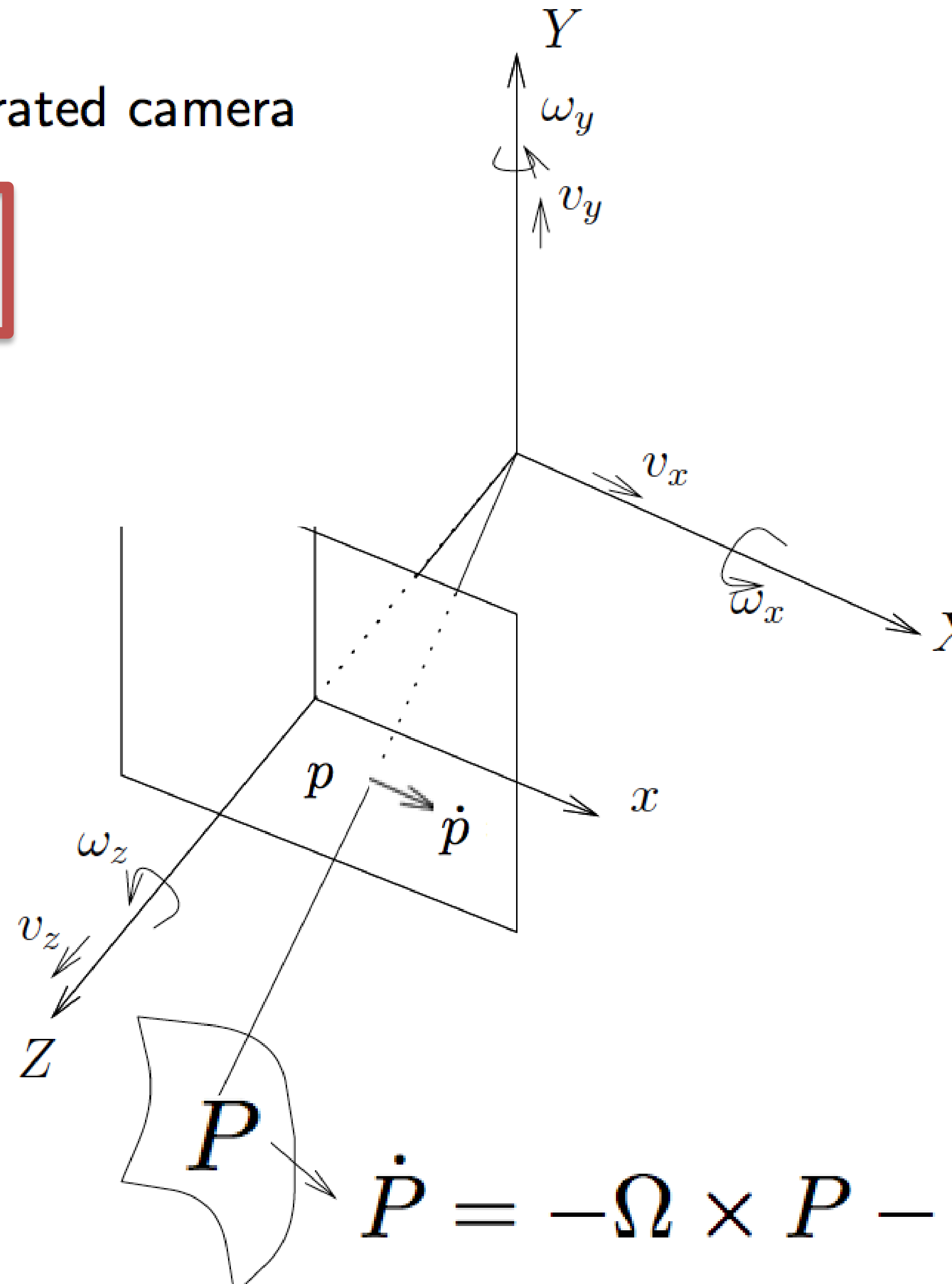
Projection equations for calibrated camera

$x = \frac{X}{Z}$, $y = \frac{Y}{Z}$

or in vector notation $\boxed{p = \frac{1}{Z}P}$

Differentiating w.r.t. time yields:

$$\dot{p} = \frac{\dot{P}}{Z} - \frac{\dot{Z}}{Z}p$$



$$\dot{P} = -\Omega \times P - V$$

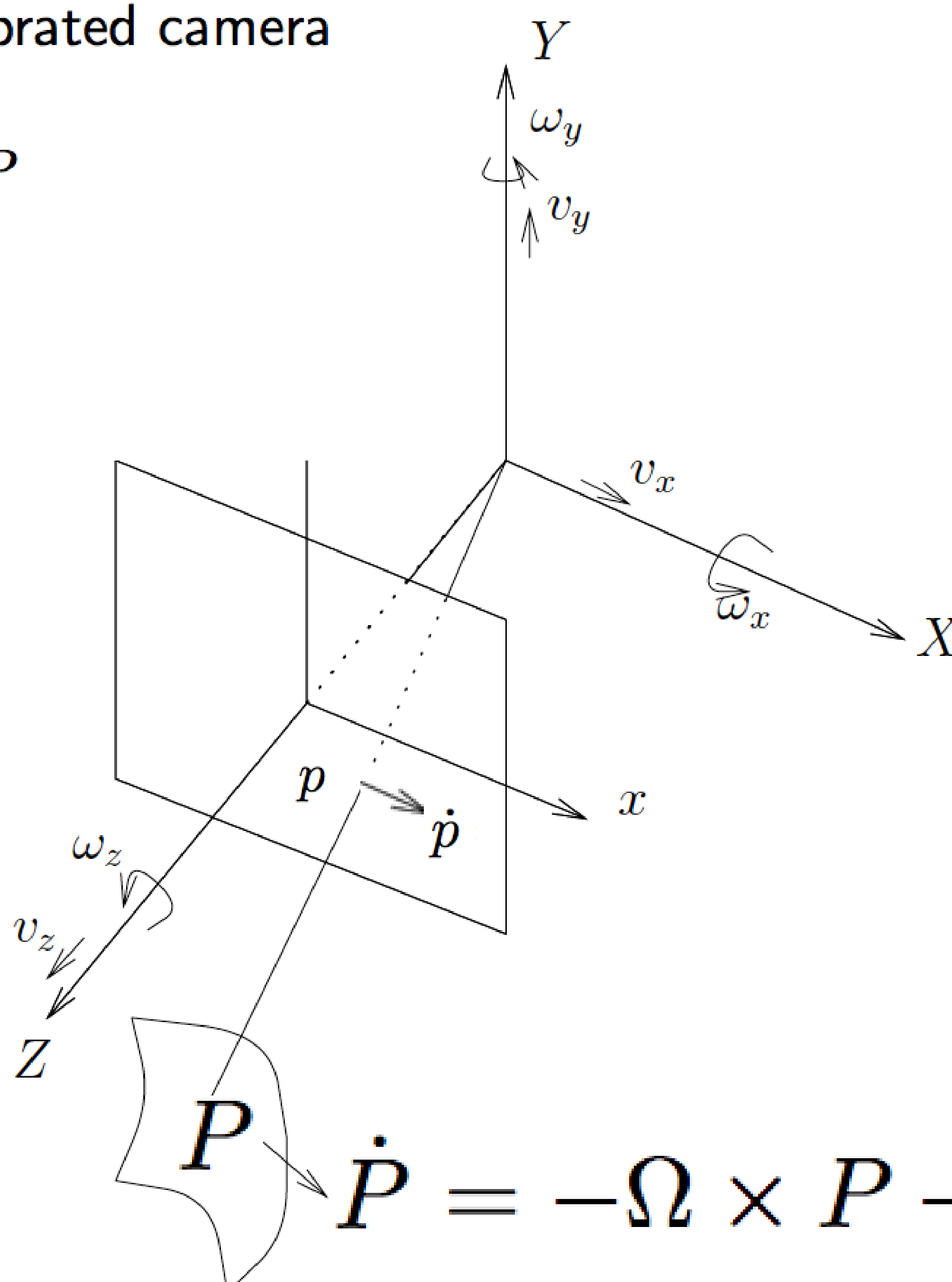
Projection equations for calibrated camera

$x = \frac{X}{Z}$, $y = \frac{Y}{Z}$

or in vector notation $p = \frac{1}{Z}P$

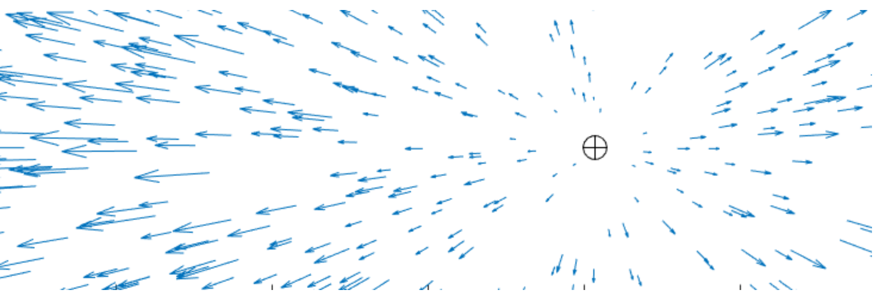
Differentiating w.r.t. time yields:

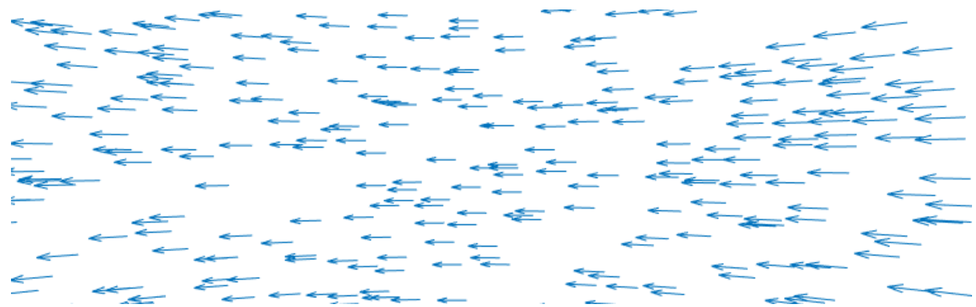
$$\boxed{\dot{p} = \frac{\dot{P}}{Z} - \frac{\dot{Z}}{Z}p}$$

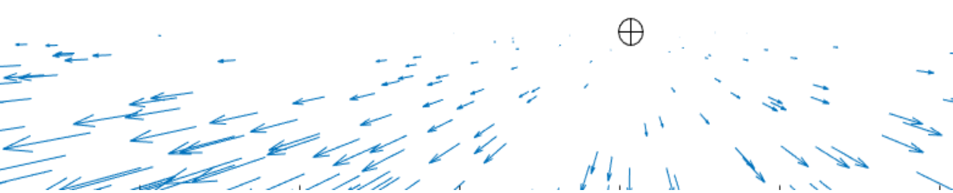


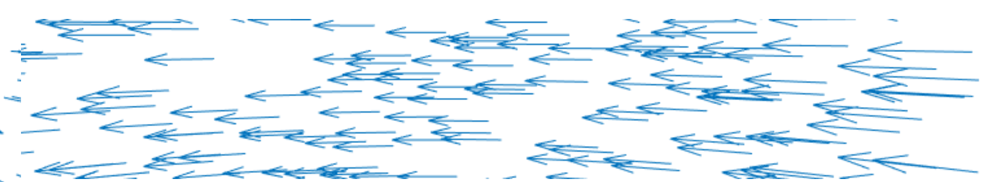
$$\dot{P} = -\Omega \times P -$$

$$\dot{p} = \underbrace{\frac{1}{Z}\begin{bmatrix} xV_z - V_x \\ yV_z - V_y \end{bmatrix}}_{\text{translational flow}} + \underbrace{\begin{bmatrix} xy & -(1+x^2) & y \\ (1+y)^2 & -xy & -x \end{bmatrix}}_{\text{rotational flow independent of depth}} \Omega$$









$$\dot{p} = \underbrace{\frac{1}{Z}\begin{bmatrix} xV_z - V_x \\ yV_z - V_y \end{bmatrix}}_{\text{translational flow}} + \underbrace{\begin{bmatrix} xy & -(1+x^2) & y \\ (1+y)^2 & -xy & -x \end{bmatrix}}_{\text{rotational flow independent of depth}} \Omega$$

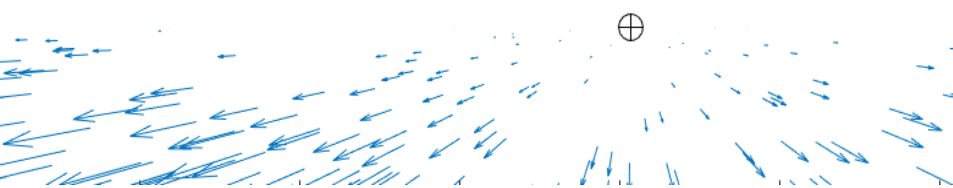
If Z is known, $\dot{p}$ is linear in V and Ω.

Having at least 3 optical flow vectors not on collinear points and corresponding depths we can solve for the 3D velocities from 6 equations.

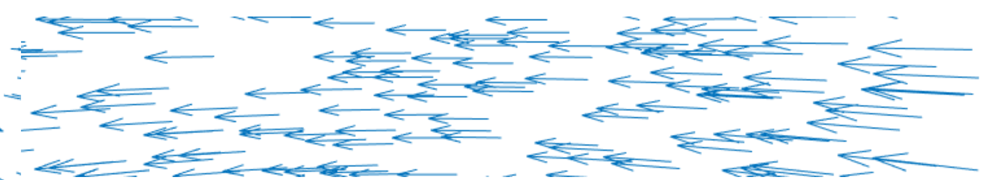
If the field is purely rotational then we have no information about depth.

$$\dot{p} = \underbrace{\begin{bmatrix} xy & -(1+x^2) & y \\ (1+y)^2 & -xy & -x \end{bmatrix}}_{\text{rotational flow independent of depth}} \Omega$$

Translation does not move points at infinity.

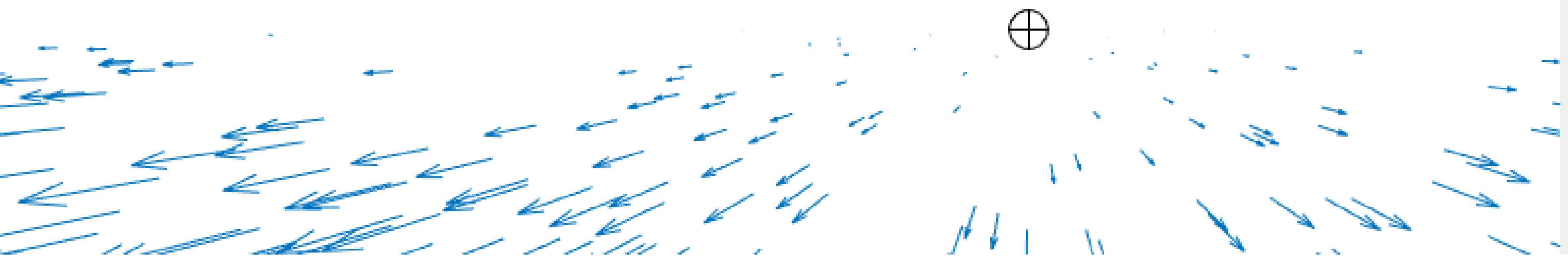


If we look at the horizontal plane points at infinity still rotate.



Translational Flow:

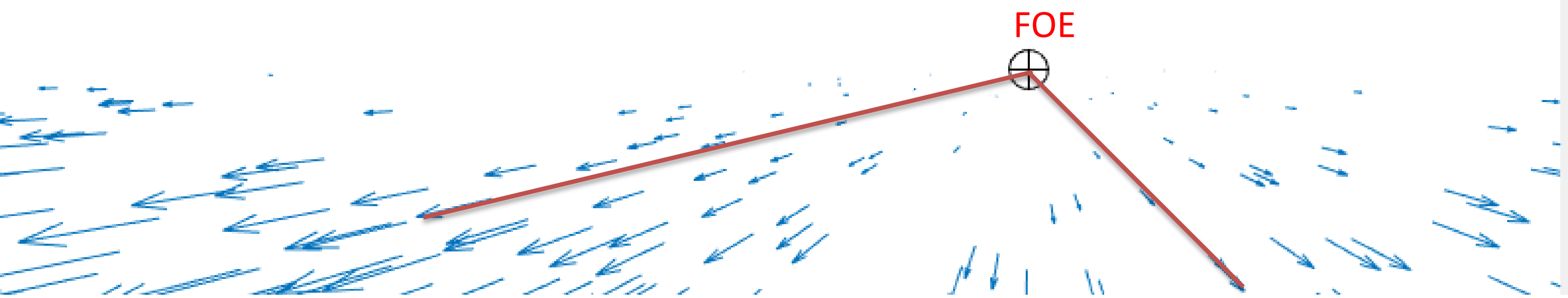
$$\dot{p}_{\text{trans}} = \frac{V_z}{Z} \begin{bmatrix} x - \frac{V_x}{V_z} \\ y - \frac{V_y}{V_z} \end{bmatrix}$$



By intersecting the lines spanned by $\dot{p}_{\text{trans}}$, we can obtain the Focus of Expansion (FOE) also called Epipole

$$FOE = (V_x/V_z, V_y/V_z)$$

FOE can also be at infinity if $V_z = 0$.



Translational Flow:

$$\dot{p}_{\text{trans}} = \frac{V_z}{Z} \begin{bmatrix} x - \frac{V_x}{V_z} \\ y - \frac{V_y}{V_z} \end{bmatrix}$$

The time to collision (which birds and insects estimate) is

$$\frac{Z}{V_z}$$

$$\frac{V_z}{Z} = \frac{\|\dot{p}_{\text{trans}}\|}{\|p - \vec{FOE}\|}$$

Points at the same radial distance from FOE have flow vector lengths proportional to inverse depth (or inverse time to collision).

From

$$\dot{p}_{trans}^T (p \times V) = 0$$

we obtain the following coplanarity condition

$$V^T (p \times \dot{p}_{trans}) = 0$$

which says that image point, flow, and linear velocity lie on the same plane.

We can obtain V from two points

$$V \sim (p_1 \times \dot{p_1}) \times (p_2 \times \dot{p_2})$$

From

$$\dot{p}_{trans}^T(p \times V) = 0$$

we obtain the following coplanarity condition

$$V^T(p \times \dot{p}_{trans}) = 0$$

which says that image point, flow, and linear velocity lie on the same plane.

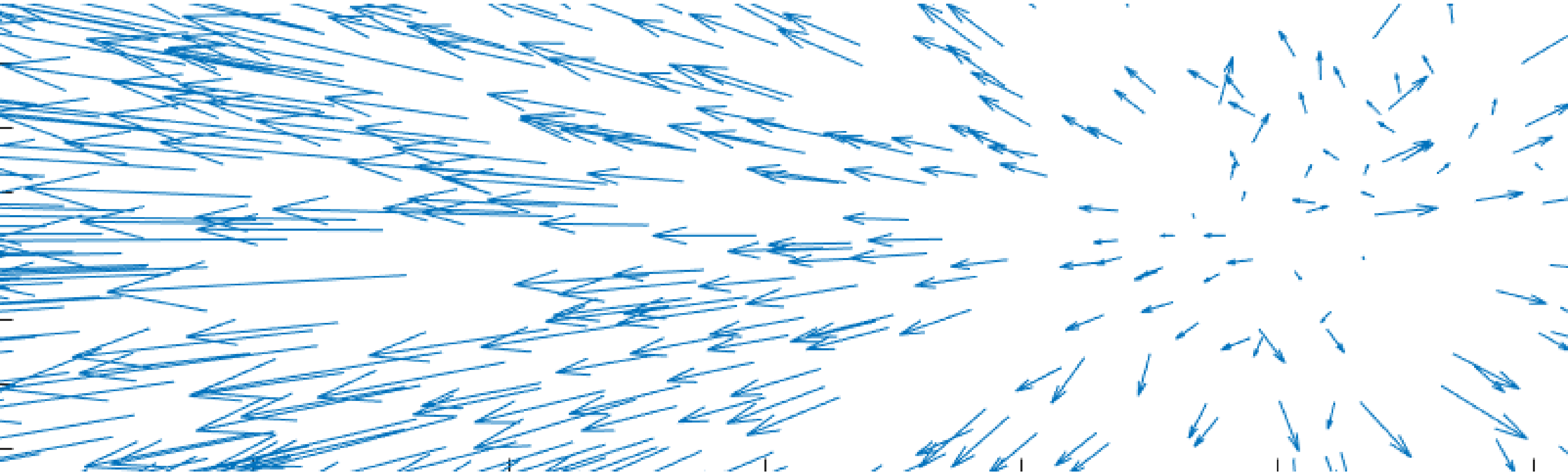
We can obtain V from two points

$$V \sim (p_1 \times \dot{p_1}) \times (p_2 \times \dot{p_2})$$

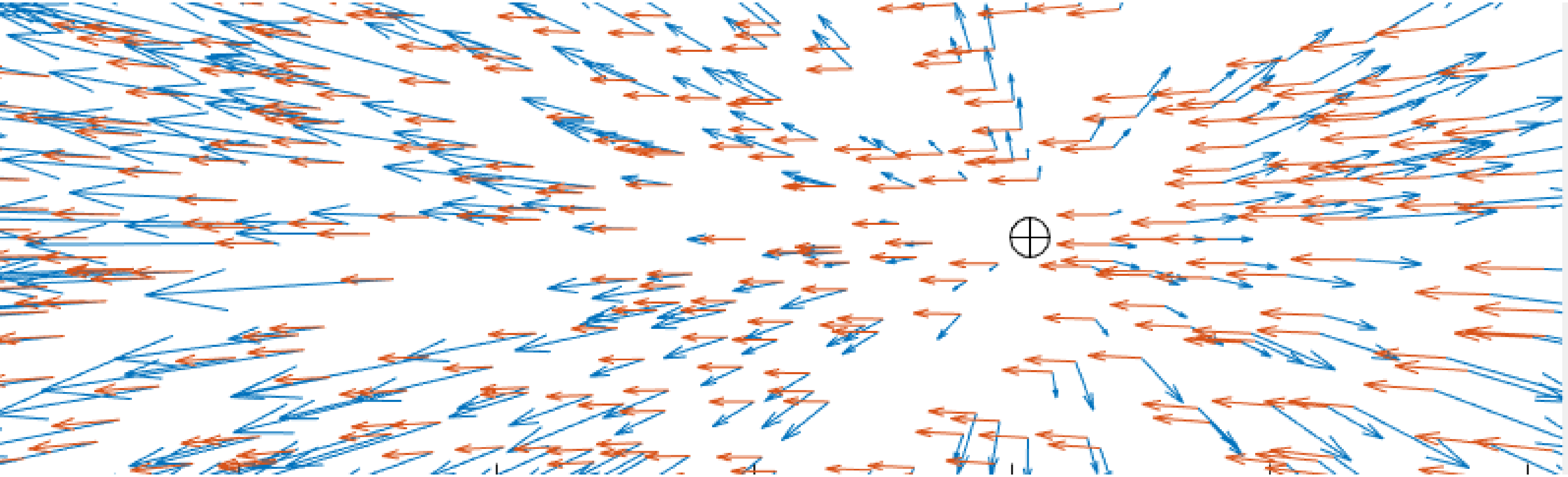
and from n points we obtain a homogeneous system

$$\underbrace{\begin{pmatrix} (p_1 \times \dot{p}_1)^T \\ p_2 \times \dot{p}_2)^T \\ \dots \\ p_n \times \dot{p}_n)^T \end{pmatrix}}_{A} V = 0 \tag{1}$$

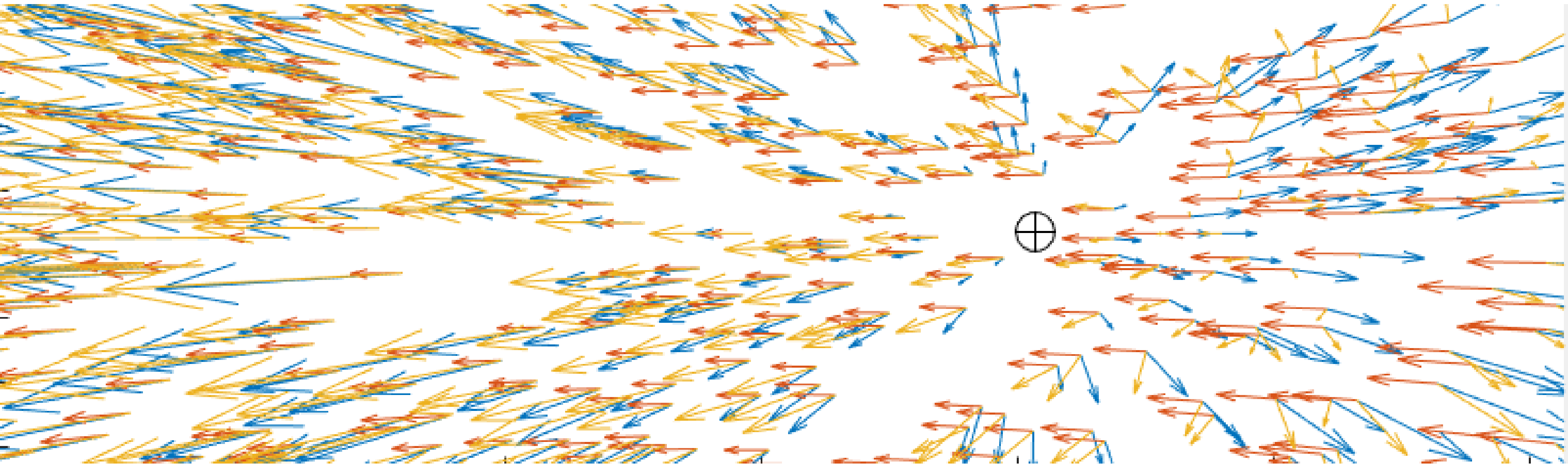
Then V is the nullspace of A which can be obtained from SVD.



.. But how can we split a mixed optical flow field?



.. As addition of two vector fields?



Both V and Ω unknown

Recall that

$$\dot{p} = \frac{1}{Z} F(x,y)V + G(x,y)\Omega$$

This is can be written linearly in inverse depths and Ω:

$$\dot{p} = \left[F(x,y)V \;\; G(x,y) \right] \begin{bmatrix} \frac{1}{Z} \\ \Omega \end{bmatrix}$$

For n points we can write out a system of equations:

$$\begin{pmatrix} \dot{p}_1 \\ \dot{p}_2 \\ \dots \\ \dot{p}_n \end{pmatrix} = \Phi(V) \begin{pmatrix} \frac{1}{Z_1} \\ \frac{1}{Z_1} \\ \dots \\ \frac{1}{Z_N} \\ \Omega \end{pmatrix}$$

The Φ matrix is a 2N by (N+3) matrix and is a function of V

$$\dot{d} = \Phi(V) \begin{pmatrix} \frac{1}{Z_1} \\ \frac{1}{Z_1} \\ \dots \\ \frac{1}{Z_N} \\ \Omega \end{pmatrix}$$

The Φ matrix is a 2N by (N+3) matrix and is a function of V

$$\dot{d} = \Phi(V) \begin{pmatrix} \frac{1}{Z_1} \\ \frac{1}{Z_1} \\ \dots \\ \frac{1}{Z_N} \\ \Omega \end{pmatrix}$$

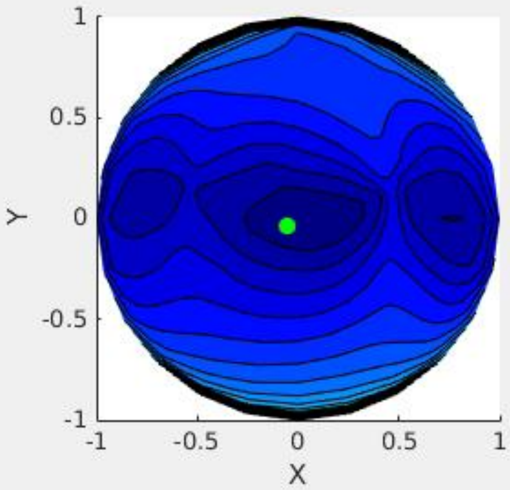
If we solve for the unknown vector of inverse depths and Ω we obtain

$$\Phi^+(V)\dot{d}$$

which we can insert back in the objective function.

A search on the sphere yields then V:

$$\arg\min_{V \in S^2} \|\dot{d} - \Phi(V)\Phi(V)^+\dot{d}\|^2$$



Error function on the sphere of all translation directions (foci of expansion)