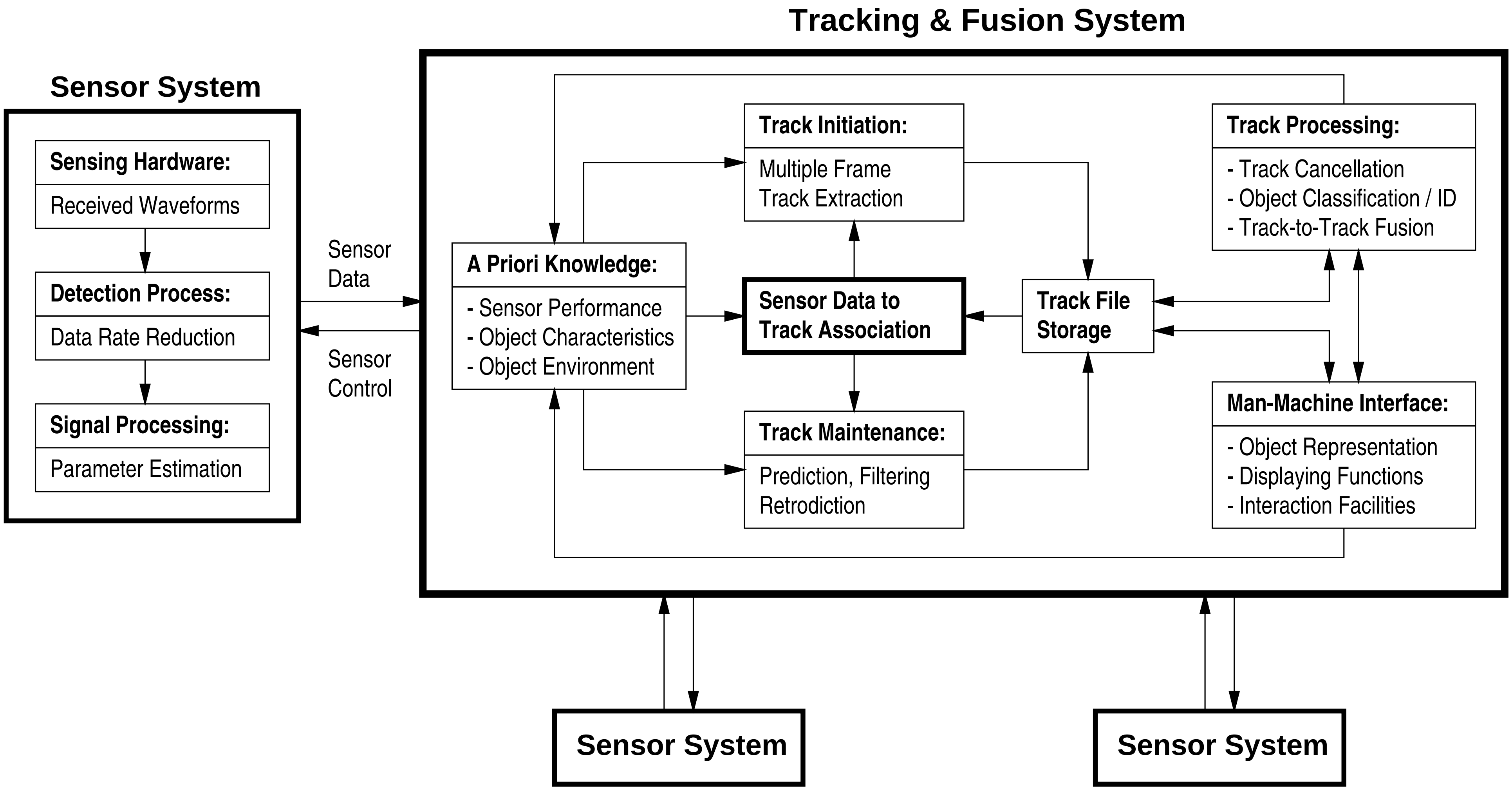
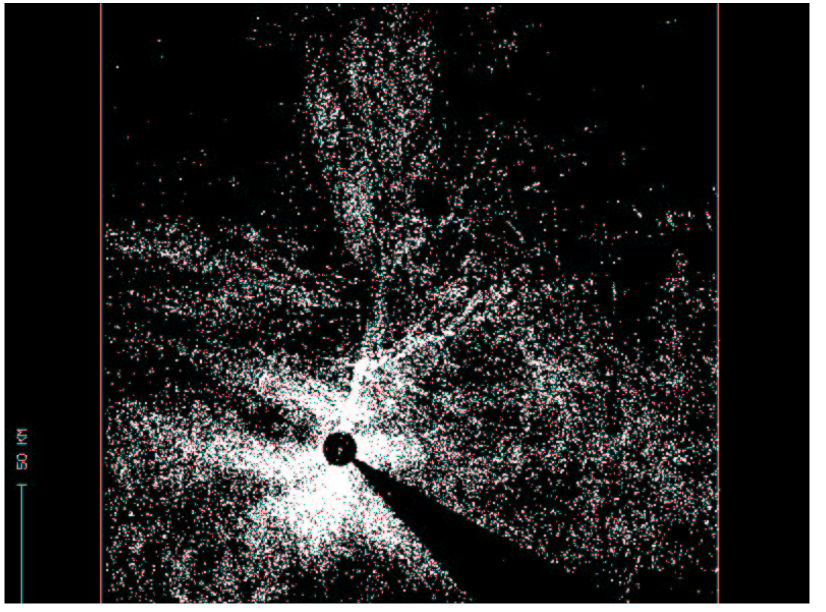
Introduction to Sensor Data Fusion Methods and Applications

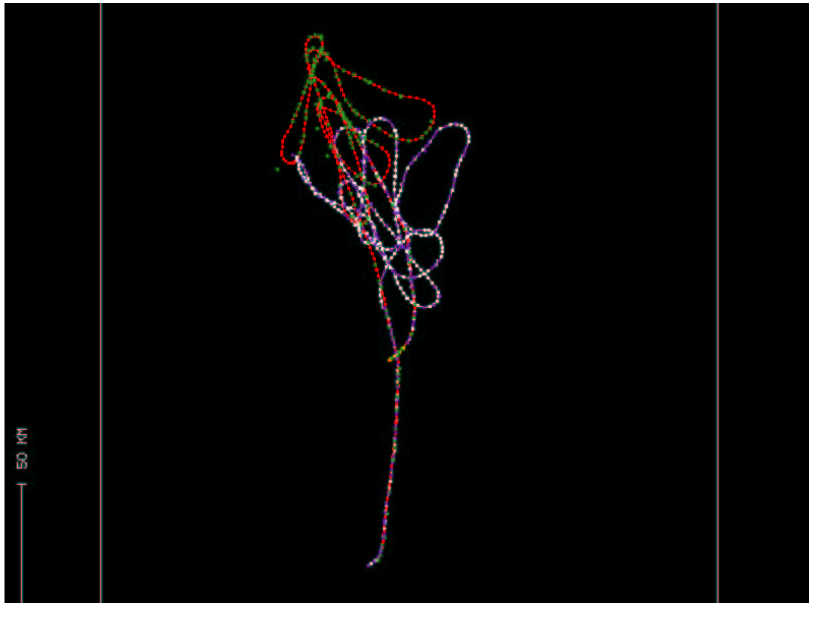
- **Last lecture: Why Sensor Data Fusion?**
  - Motivation, general context
  - Discussion of examples
- **Today: Steep climb to a first algorithm.**

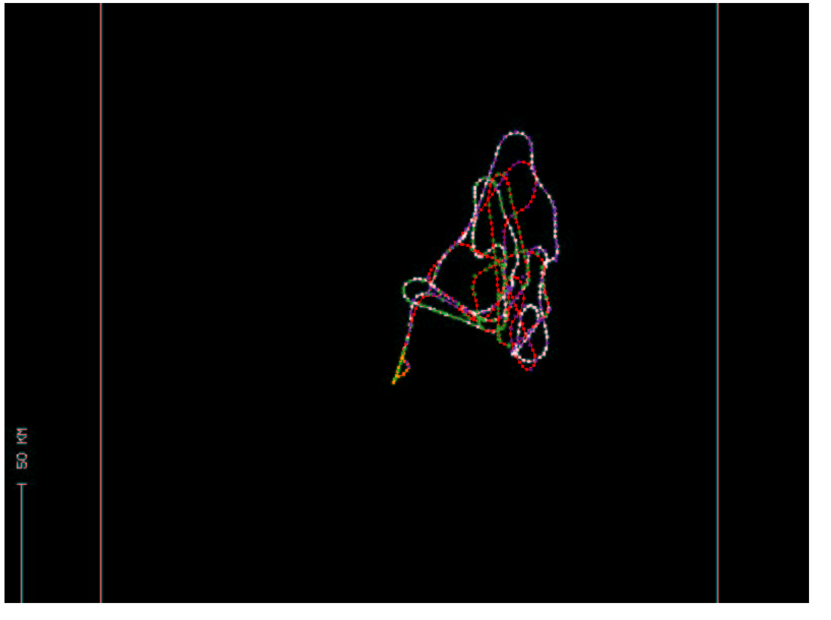
- **oral examination: 6 credit points after the end of the semester**
- **prerequisite: participate in the excercises, explain a good program**
- **job opportunities as research assistant in ongoing projects, practicum**
- **subsequently: bachelor at Fraunhofer FKIE, master / PhD possible**
- **slides/script: email to wolfgang.koch@fkie.fraunhofer.de, download**

A Generic Tracking and Sensor Data Fusion System









Fraunhofer
FKIE

Tracking Application: Ground Picture Production

GMTI Radar: Ground Moving Target Indicator

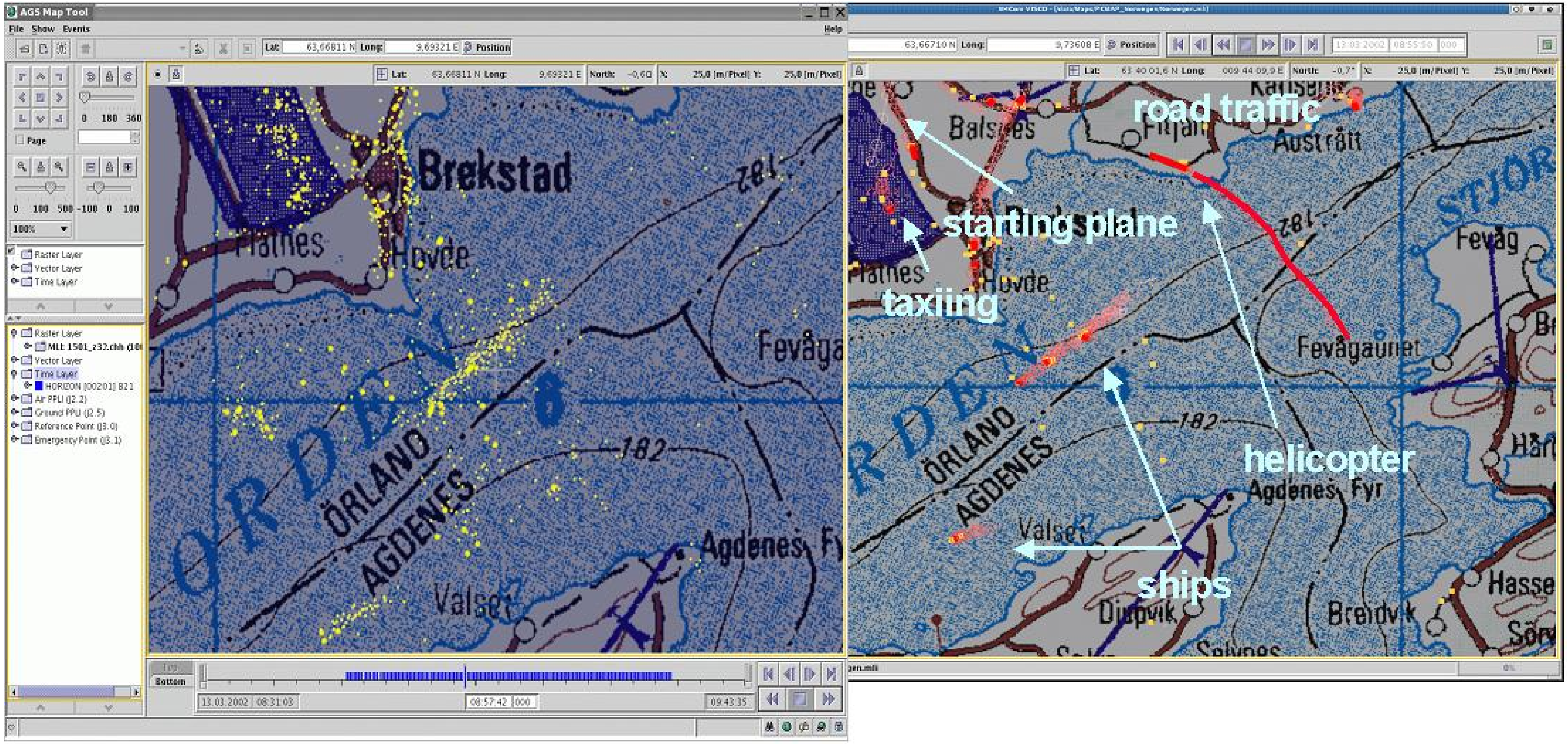
wide area, all-weather, day/night, real-time surveillance of a dynamically evolving ground or near-to-ground situation

GMTI Tracking: Some Characteristic Aspects

backbone of a ground picture: moving target tracks

- airborne, dislocated, mobile sensor platforms
- vehicles, ships, ‘low-flyers’, radars, convoys
- occlusions: Doppler-blindness, topography
- road maps, terrain information, tactical rules
- dense target / dense clutter situations: MHT

Examples of GMTI Tracks (live exercise)



Fraunhofer FKIE

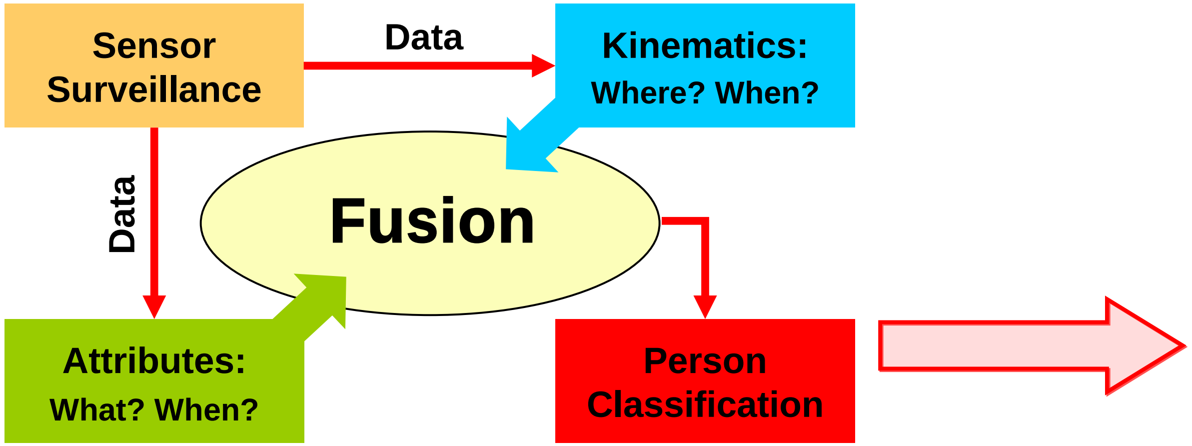
Multiple Sensor Security Assistance Systems

General Task

Covert & Automated Surveillance of a Person Stream: Identification of Anomalous Behavior

Towards a Solution

Exploit Heterogeneous Multiple Sensor Systems.



On Characterizing Tracking / Fusion Performance

a well-understood paradigm: *air surveillance with multiple radars*
Many results can be transfered to other sensors (IR, E/O, sonar, acoustics).

Sensor Data Fusion: 'tracks' represent the available information on the targets associated to them with appropriate quality measures, thus providing answers to:

When? Where? How many? To which direction? How fast, accelerating? What?

On Characterizing Tracking / Fusion Performance

a well-understood paradigm: *air surveillance with multiple radars*
Many results can be transfered to other sensors (IR, E/O, sonar, acoustics).

Sensor Data Fusion: 'tracks' represent the available information on the targets associated to them with appropriate quality measures, thus providing answers to:

When? Where? How many? To which direction? How fast, accelerating? What?

By sensor data fusion we wish to establish one-to-one associations between:

targets in the field of view $\leftrightarrow$ identified tracks in the tracking computer

Strictly speaking, this is only possible under ideal conditions regarding the sensor performance and underlying target situation. The tracking/fusion performance can thus be measured by its **deficiencies when compared with this ideal goal**.

1. Let a target be detected at first by a sensor at time t_a. Usually, a time delay is involved until a confirmed track has finally been established at time t_e (track extraction). A 'measure of deficiency' is thus:
   - extraction delay $t_e - t_a$.

1. Let a target be detected at first by a sensor at time t_a. Usually, a time delay is involved until a confirmed track has finally been established at time t_e (track extraction). A 'measure of deficiency' is thus:
    - extraction delay $t_e - t_a$.

2. Unavoidably, false tracks will be extracted in case of a high false return density (e.g. clutter, jamming/detection), i.e. tracks related to unreal or unwanted targets. Corresponding 'deficiencies' are:
    - mean number of falsely extracted targets per time,
    - mean life time of a false track before its deletion.

1. Let a target be detected at first by a sensor at time t_a. Usually, a time delay is involved until a confirmed track has finally been established at time t_e (track extraction). A 'measure of deficiency' is thus:
   - extraction delay $t_e - t_a$.

2. Unavoidably, false tracks will be extracted in case of a high false return density (e.g. clutter, jamming/detection), i.e. tracks related to unreal or unwanted targets. Corresponding 'deficiencies' are:
   - mean number of falsely extracted targets per time,
   - mean life time of a false track before its deletion.

3. A target should be represented by one and the same track until leaving the field of view. Related performance measures/deficiencies:
   - mean life time of tracks related to true targets,
   - probability of an 'identity switch' between targets,
   - probability of a target not being represented by a track.

4. The track inaccuracy (error covariance of a state estimate) should be as small as possible. The deviations between estimated and actual target states should at least correspond with the error covariances produced (consistency). If this is not the case, we speak of a ‘track loss’.

4. The track inaccuracy (error covariance of a state estimate) should be as small as possible. The deviations between estimated and actual target states should at least correspond with the error covariances produced (consistency). If this is not the case, we speak of a ‘track loss’.

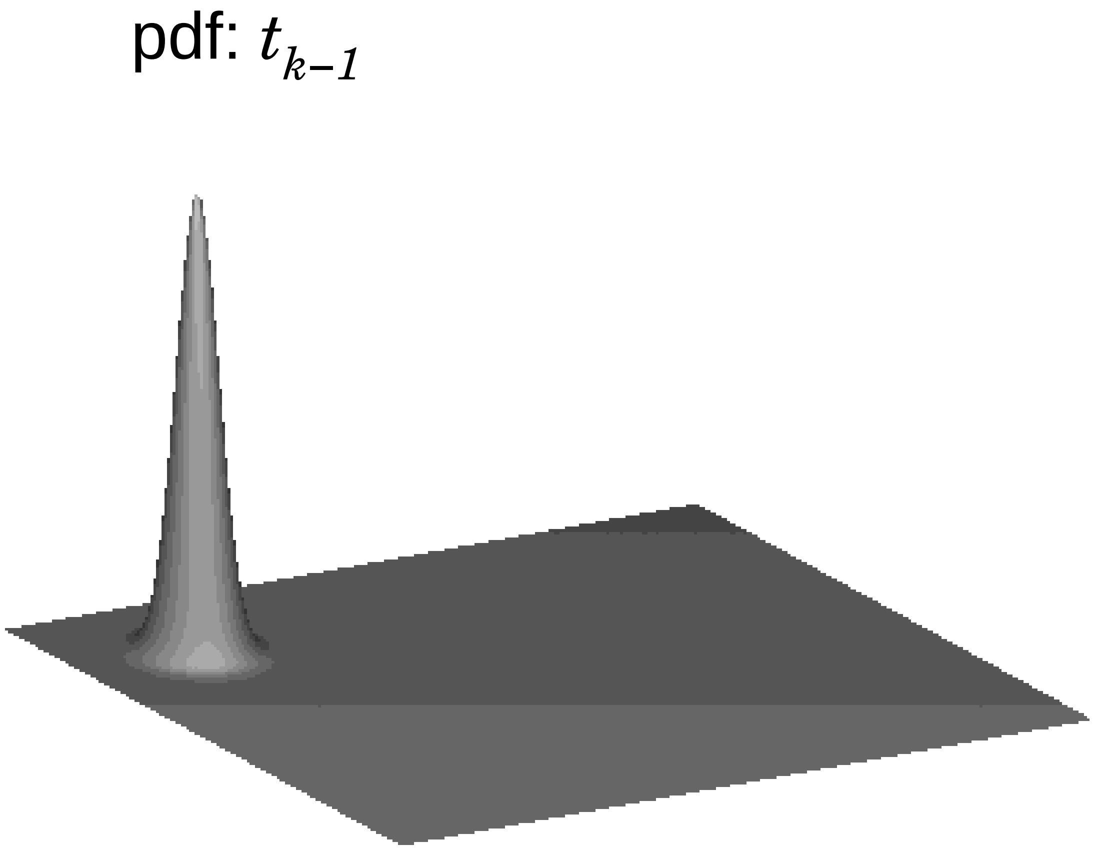
   - A track must really represent a target!

Challenges:

- low detection probability • high clutter density • low update rate
- agile targets • dense target situations • formations, convoys
- target-split events (formation, weapons) • jamming, deception

Basic Tasks:

- models: sensor, target, environment $\rightarrow$ *physics*
- data association problems $\rightarrow$ *combinatorics*
- estimation problems $\rightarrow$ *probability, statistics*
- process control, realization $\rightarrow$ *computer science*



'Probability densities functions (pdf)' $p(\mathrm{x}_{k-1}|\mathcal{Z}^{k-1})$ **represent imprecise knowledge on the 'state'** x_{k-1} **based on imprecise measurements** $\mathcal{Z}^{k-1}$**.**

Characterize an object by *quantitatively describable* properties: object state

Examples:

- object position x on a strait line: $x \in \mathbb{R}$
- kinematic state $\mathbf{x} = (\mathbf{r}^\top, \dot{\mathbf{r}}^\top, \ddot{\mathbf{r}}^\top)^\top$, $\mathbf{x} \in \mathbb{R}^9$

  position $\mathbf{r} = (x, y, z)^\top$, velocity $\dot{\mathbf{r}}$, acceleration $\ddot{\mathbf{r}}$
- joint state of two objects: $\mathbf{x} = (\mathbf{x}_1^\top, \mathbf{x}_2^\top)^\top$
- kinematic state $\mathbf{x}$, object extension $\mathbf{X}$

  z.B. ellipsoid: symmetric, positively definite matrix
- kinematic state $\mathbf{x}$, object class class

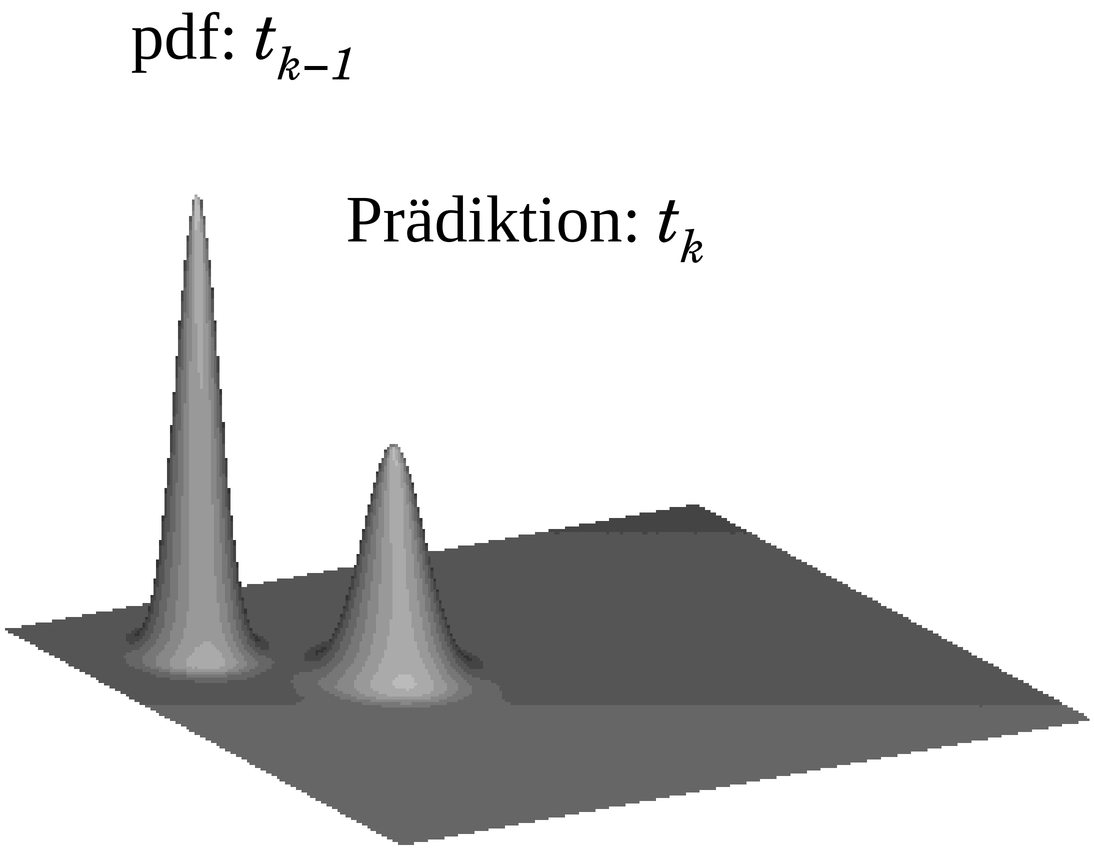
  z.B. bird, sailing plane, helicopter, passenger jet, ...

Learn unknown object states from imperfect measurements and describe by functions $p(\mathbf{x})$ imprecise knowledge mathematically precisely!

How to deal with probability density functions?

- pdf $p(x)$: Extract *probability statements* about the RV x by integration!

- naïvely: *positive* and *normalized* functions ($p(x) \geq 0$, $\int dx\, p(x) = 1$)



Exploit imprecise knowledge on the dynamical behavior of the object.

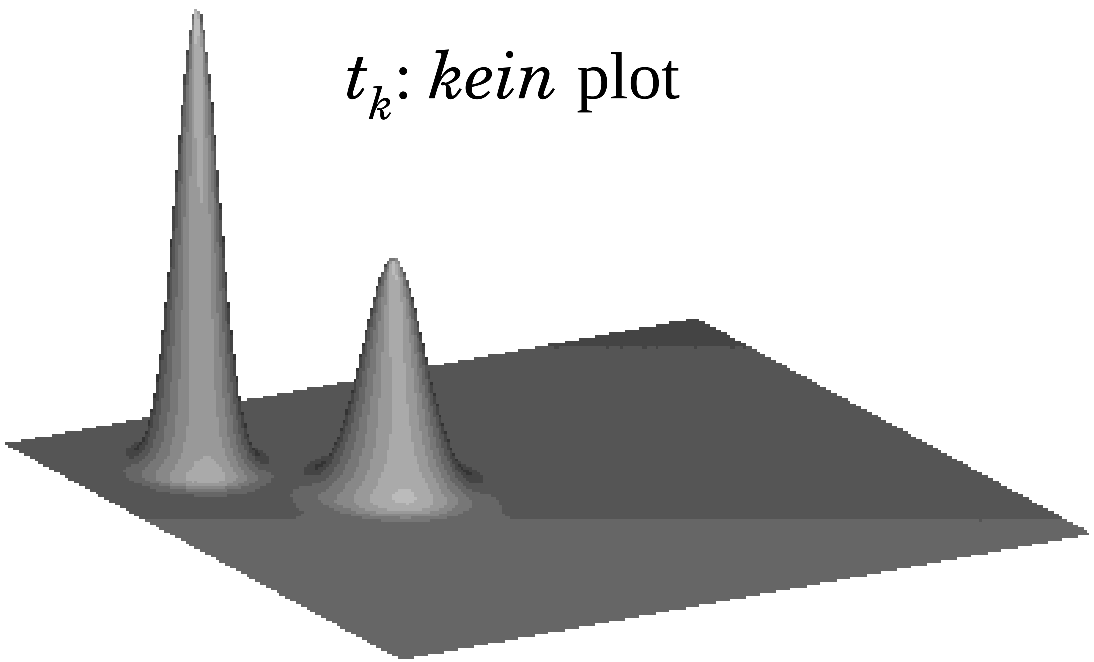
$$\underbrace{p(\mathbf{x}_k|\mathcal{Z}^{k-1})}_{\text{prediction}} = \int d\mathbf{x}_{k-1}\ \underbrace{p(\mathbf{x}_k|\mathbf{x}_{k-1})}_{\text{dynamics}}\ \underbrace{p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1})}_{\text{old knowledge}}.$$

Fraunhofer FKIE

How to deal with probability density functions?

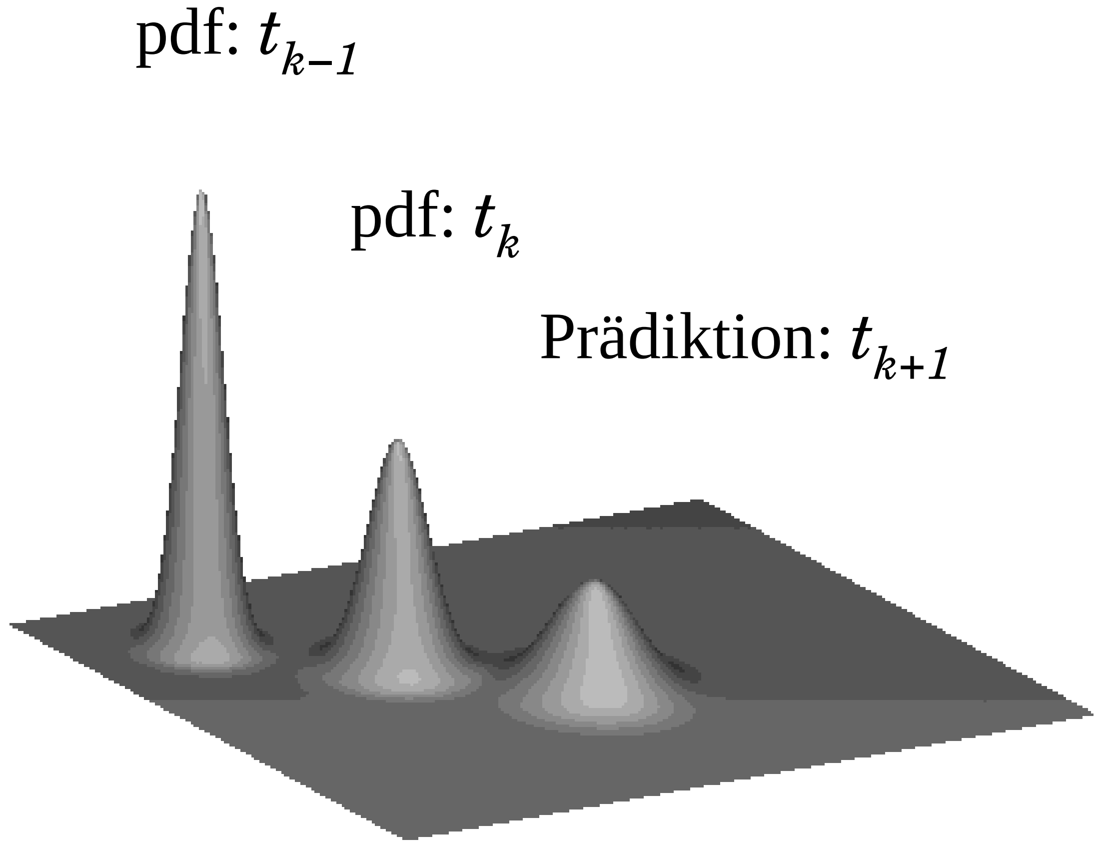
- pdf $p(x)$: Extract *probability statements* about the RV x by integration!
- naïvely: *positive* and *normalized* functions ($p(x) \geq 0$, $\int dx\, p(x) = 1$)
- *conditional pdf* $p(x|y) = \frac{p(x,y)}{p(y)}$: Impact of information on y on RV x?
- *marginal density* $p(x) = \int dy\, p(x,y) = \int dy\, p(x|y)\, p(y)$: Enter y!

$$
\begin{aligned}
p(\mathbf{x}_k|\mathcal{Z}^{k-1}) &= \int d\mathbf{x}_{k-1}\, p(\mathbf{x}_k, \mathbf{x}_{k-1}|\mathcal{Z}^{k-1}) \\
&= \int d\mathbf{x}_{k-1}\, p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathcal{Z}^{k-1})\, p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1}) \\
&= \int d\mathbf{x}_{k-1}\, p(\mathbf{x}_k|\mathbf{x}_{k-1})\, p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1})
\end{aligned}
$$

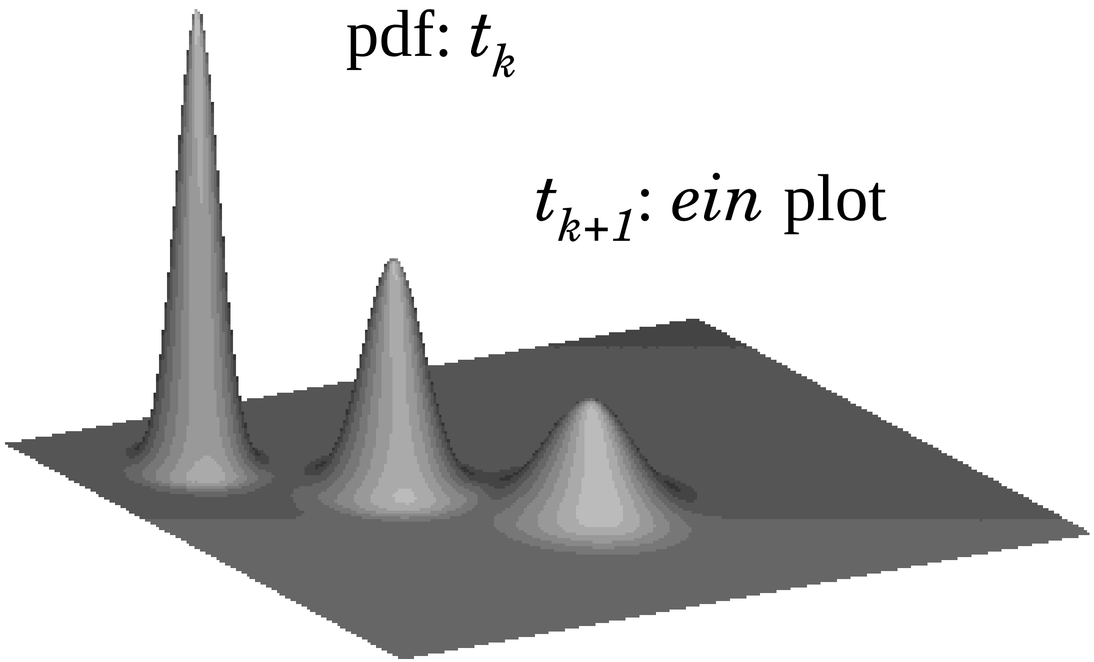


missing sensor detection: 'data processing' = prediction

(not always: exploitation of 'negative' sensor evidence)

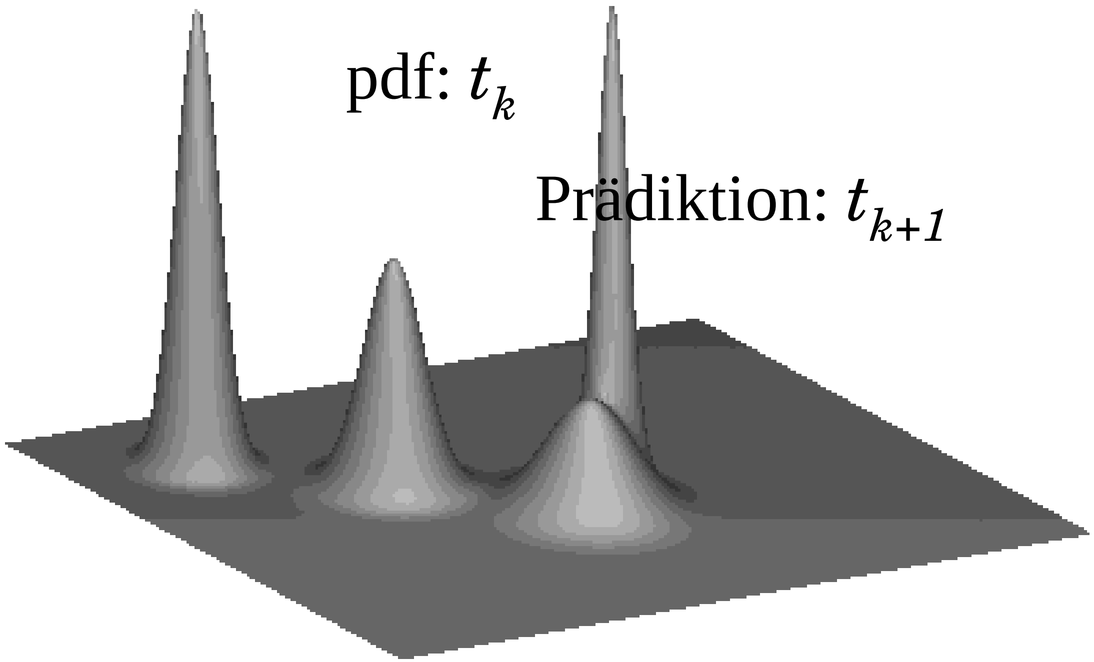


missing sensor information: increasing knowledge dissipation



sensor information on the kinematical object state

Fraunhofer FKIE



BAYES' formula: $$\underbrace{p(\mathbf{x}_{k+1}|\mathcal{Z}^{k+1})}_{\textbf{new knowledge}} = \frac{p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1})\, p(\mathbf{x}_{k+1}|\mathcal{Z}^k)}{\int d\mathbf{x}_{k+1}\, p(\underbrace{\mathbf{z}_{k+1}}_{\textbf{plot}}|\mathbf{x}_{k+1})\, \underbrace{p(\mathbf{x}_{k+1}|\mathcal{Z}^k)}_{\textbf{prediction}}}$$

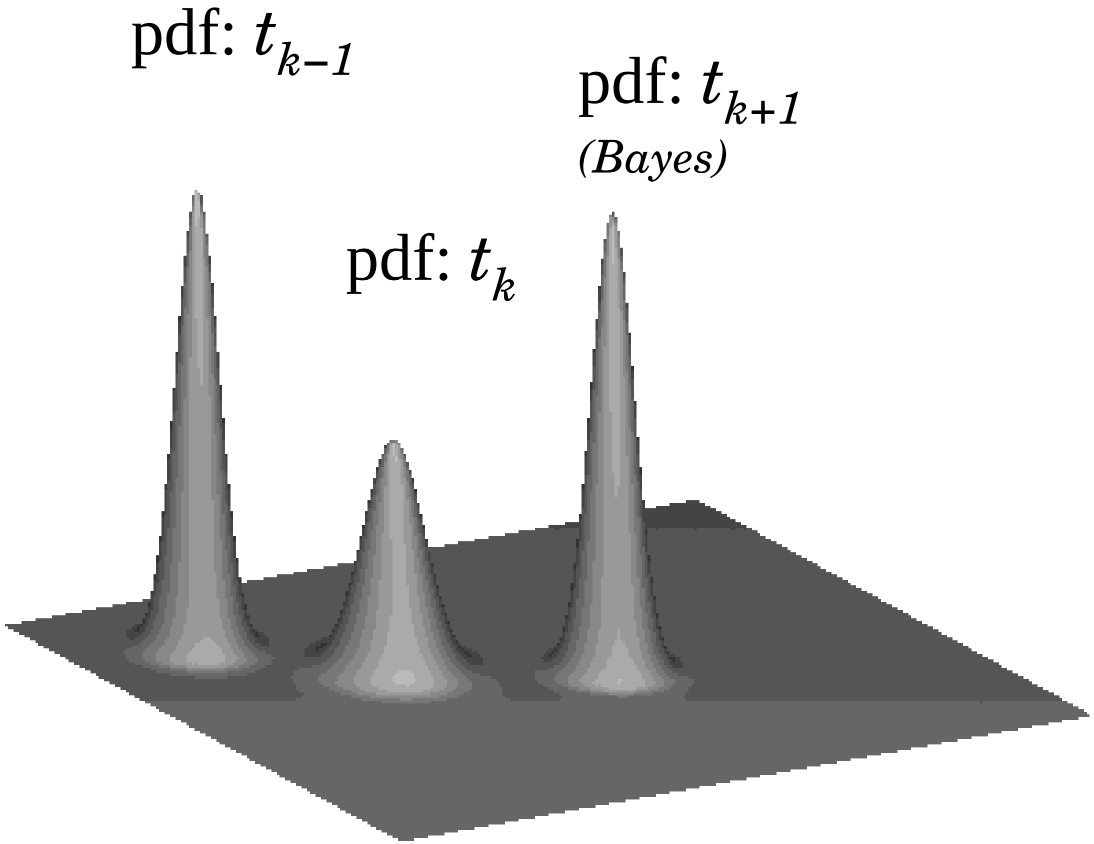
Fraunhofer FKIE

How to deal with probability density functions?

- pdf $p(x)$: Extract *probability statements* about the RV x by integration!
- naïvely: *positive* and *normalized* functions ($p(x) \geq 0$, $\int dx\, p(x) = 1$)
- *conditional pdf* $p(x|y) = \frac{p(x,y)}{p(y)}$: Impact of information on y on RV x?
- *marginal density* $p(x) = \int dy\, p(x,y) = \int dy\, p(x|y)\, p(y)$: Enter y!
- Bayes: $p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\int dx\, p(y|x)p(x)}$: $p(x|y) \leftarrow p(y|x), p(x)$!

$$p(x|y)\, p(y) = p(x,y) = p(y,z) = p(y|x)\, p(x)$$

Fraunhofer FKIE



filtering = **sensor data processing**

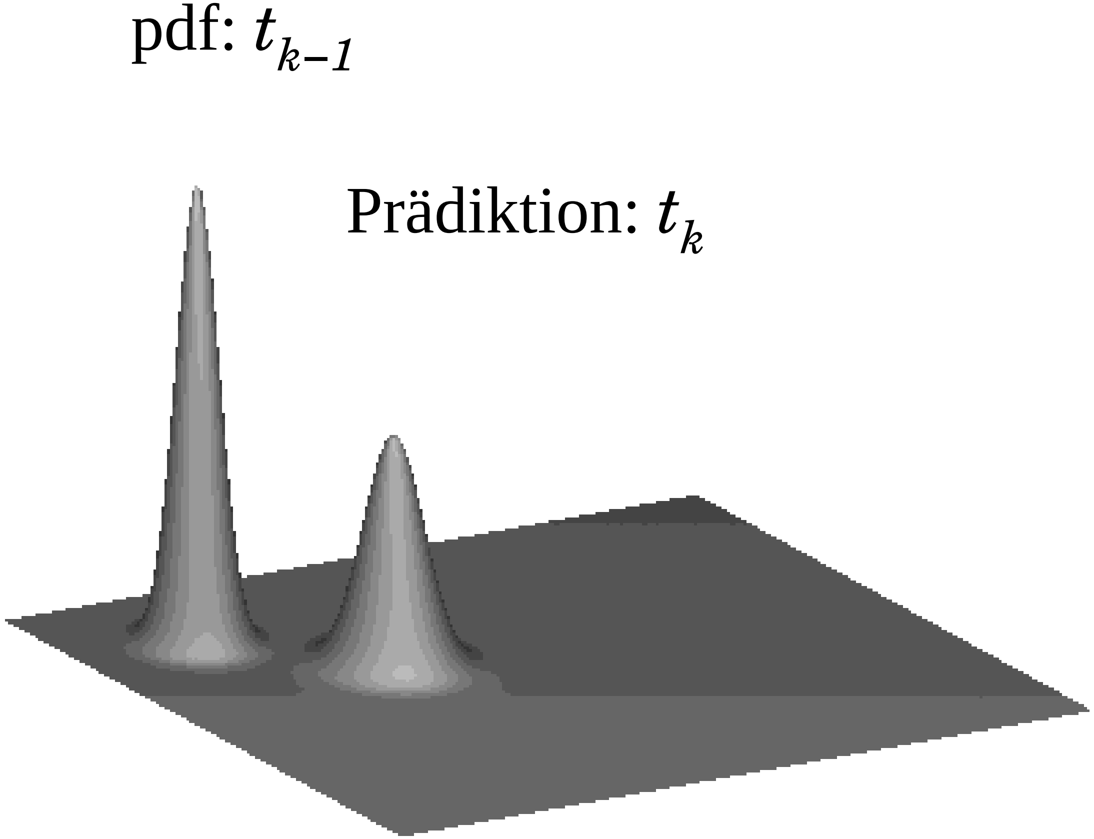
Target or Object Tracking: Basic Idea

Iterative updating of conditional probability densities!

kinematic target state x_k at time t_k, **accumulated sensor data** $\mathcal{Z}^k$

a priori knowledge: target dynamics models, sensor model

- **prediction:** $p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1}) \xrightarrow{\text{dynamics model}} p(\mathbf{x}_k|\mathcal{Z}^{k-1})$
- **filtering:** $p(\mathbf{x}_k|\mathcal{Z}^{k-1}) \xrightarrow[\text{sensor model}]{\text{sensor data } Z_k} p(\mathbf{x}_k|\mathcal{Z}^k)$
- **retrodiction:** $p(\mathbf{x}_{l-1}|\mathcal{Z}^k) \xleftarrow[\text{dynamics model}]{\text{filtering output}} p(\mathbf{x}_l|\mathcal{Z}^k)$



Exploit imprecise knowledge on the dynamical behavior of the object.

$$\underbrace{p(\mathbf{x}_k|\mathcal{Z}^{k-1})}_{\text{prediction}} = \int d\mathbf{x}_{k-1}\ \underbrace{p(\mathbf{x}_k|\mathbf{x}_{k-1})}_{\text{dynamics}}\ \underbrace{p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1})}_{\text{old knowledge}}.$$

Fraunhofer FKIE

The Multivariate GAUSSian Pdf

- *wanted:* probabilities 'concentrated' around a center $\mathbf{x}$

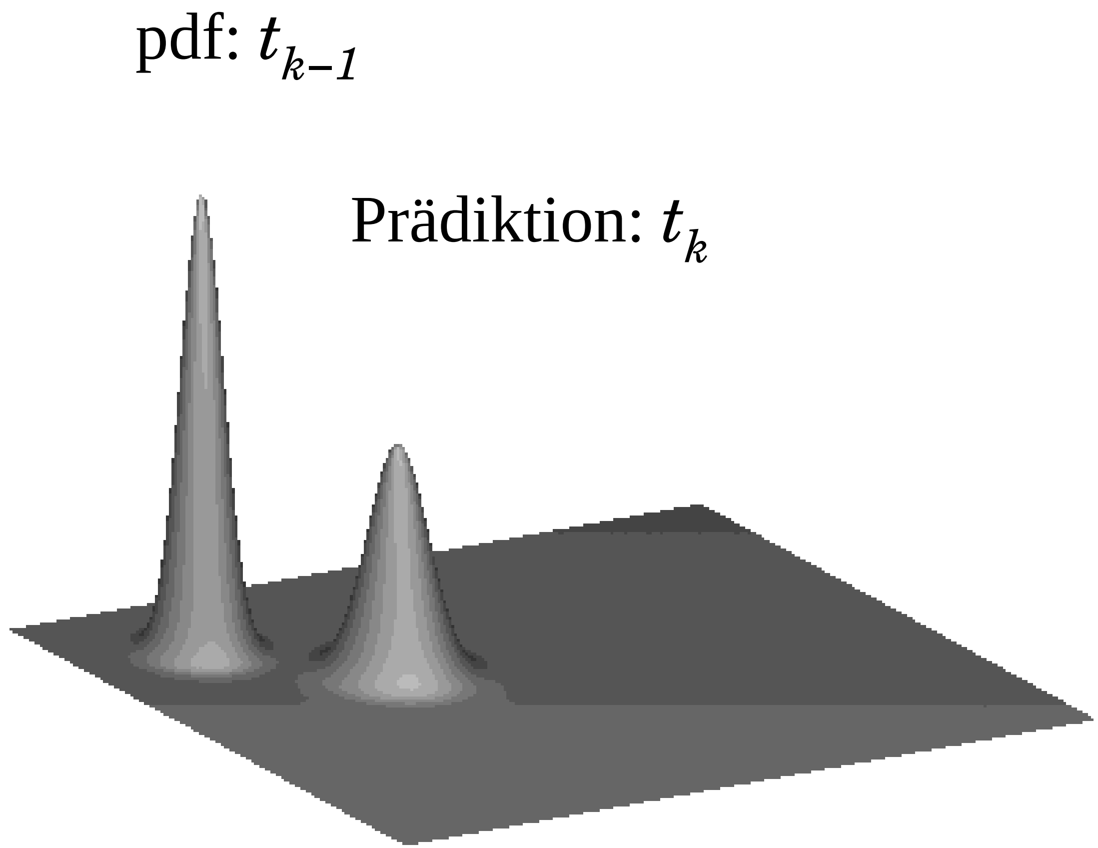
- *quadratic distance:* $q(x) = \frac{1}{2}(x - \mathbf{x})\mathbf{P}^{-1}(x - \mathbf{x})^\top$

  $q(x)$ defines an ellipsoid around $\mathbf{x}$, its volume and orientation being determined by a matrix $\mathbf{P}$ (symmetric: $\mathbf{P}^\top = \mathbf{P}$, positively definite: all eigenvalues > 0).

- *first attempt:* $p(x) = \mathrm{e}^{-q(x)} / \int dx\, \mathrm{e}^{-q(x)}$ (normalized!)

$$p(x) = \mathcal{N}(x;\, \mathbf{x},\, \mathbf{P}) = \frac{1}{\sqrt{|2\pi\mathbf{P}|}}\, \mathrm{e}^{-\frac{1}{2}(x-\mathbf{x})^\top \mathbf{P}^{-1}(x-\mathbf{x})}$$

- GAUSS*ian Mixtures:* $p(x) = \sum_i p_i\, \mathcal{N}(x;\, \mathbf{x}_i,\, \mathbf{P}_i)$ (weighted sums)



Exploit imprecise knowledge on the dynamical behavior of the object.

$$\underbrace{p(\mathbf{x}_k|\mathcal{Z}^{k-1})}_{\text{prediction}} = \int d\mathbf{x}_{k-1}\ \underbrace{\mathcal{N}\big(\mathbf{x}_k;\ \mathbf{F}\mathbf{x}_{k-1},\ \mathbf{D}\big)}_{\text{dynamics}}\ \underbrace{\mathcal{N}\big(\mathbf{x}_{k-1};\ \mathbf{x}_{k-1|k-1},\ \mathbf{P}_{k-1|k-1}\big)}_{\text{old knowledge}}.$$

Fraunhofer FKIE

A Useful Product Formula for GAUSSians

$$\mathcal{N}(\mathbf{z};\, \mathbf{F}\mathbf{x},\, \mathbf{D})\, \mathcal{N}(\mathbf{x};\, \mathbf{y},\, \mathbf{P}) = \underbrace{\mathcal{N}(\mathbf{z};\, \mathbf{F}\mathbf{y},\, \mathbf{S})}_{\text{independent of } \mathbf{x}}\, \mathcal{N}(\mathbf{x};\, \mathbf{y} + \mathbf{W}\boldsymbol{\nu},\, \mathbf{P} - \mathbf{W}\mathbf{S}\mathbf{W}^\top)$$

$$\boldsymbol{\nu} = \mathbf{z} - \mathbf{F}\mathbf{y}, \quad \mathbf{S} = \mathbf{F}\mathbf{P}\mathbf{F}^\top + \mathbf{D}, \quad \mathbf{W} = \mathbf{P}\mathbf{F}^\top\mathbf{S}^{-1}.$$

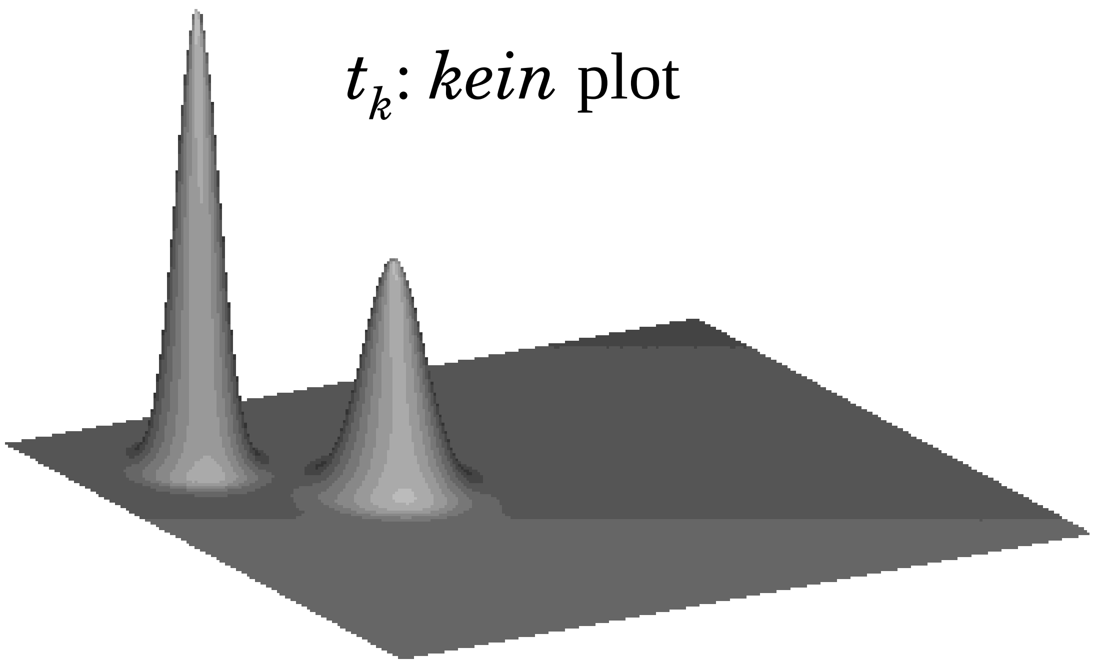
Kalman filter: $\mathbf{x}_k = (\mathbf{r}_k^\top, \dot{\mathbf{r}}_k^\top)^\top$, $\mathcal{Z}^k = \{\mathbf{z}_k, \mathcal{Z}^{k-1}\}$

initiation: $p(\mathbf{x}_0) = \mathcal{N}\big(\mathbf{x}_0;\ \mathbf{x}_{0|0},\ \mathbf{P}_{0|0}\big)$, initial ignorance: $\mathbf{P}_{0|0}$ 'large'

prediction: $\mathcal{N}\big(\mathbf{x}_{k-1};\ \mathbf{x}_{k-1|k-1},\ \mathbf{P}_{k-1|k-1}\big) \xrightarrow[\mathbf{F}_{k|k-1},\ \mathbf{D}_{k|k-1}]{\text{dynamics model}} \mathcal{N}\big(\mathbf{x}_k;\ \mathbf{x}_{k|k-1},\ \mathbf{P}_{k|k-1}\big)$

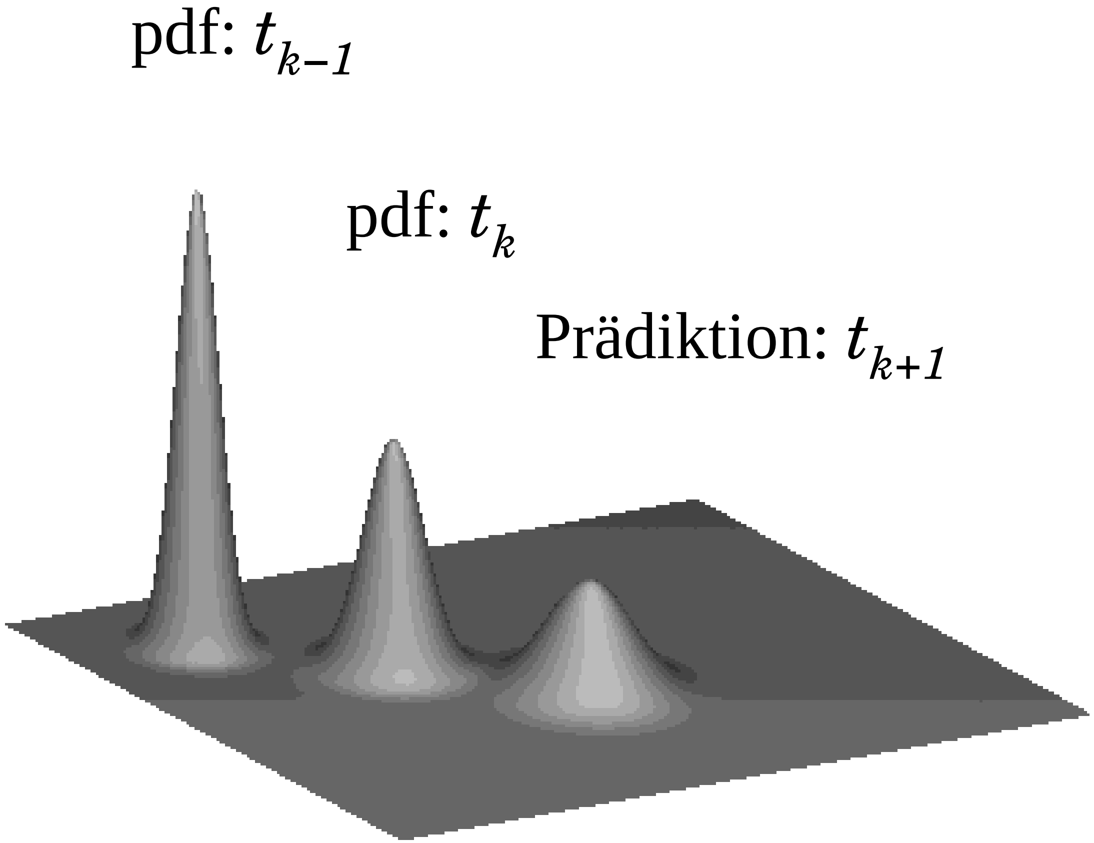
$$\mathbf{x}_{k|k-1} = \mathbf{F}_{k|k-1}\mathbf{x}_{k-1|k-1}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k|k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k|k-1}{}^\top + \mathbf{D}_{k|k-1}$$

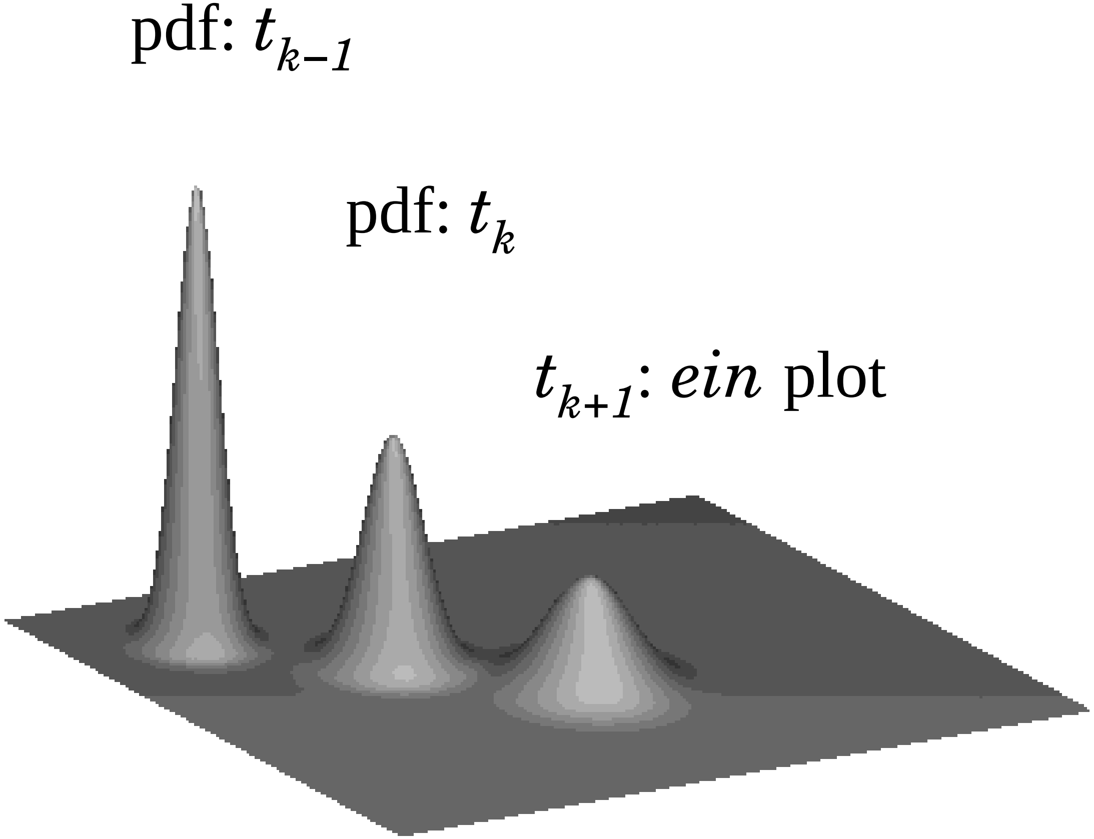


missing sensor detection: ‘data processing’ = prediction

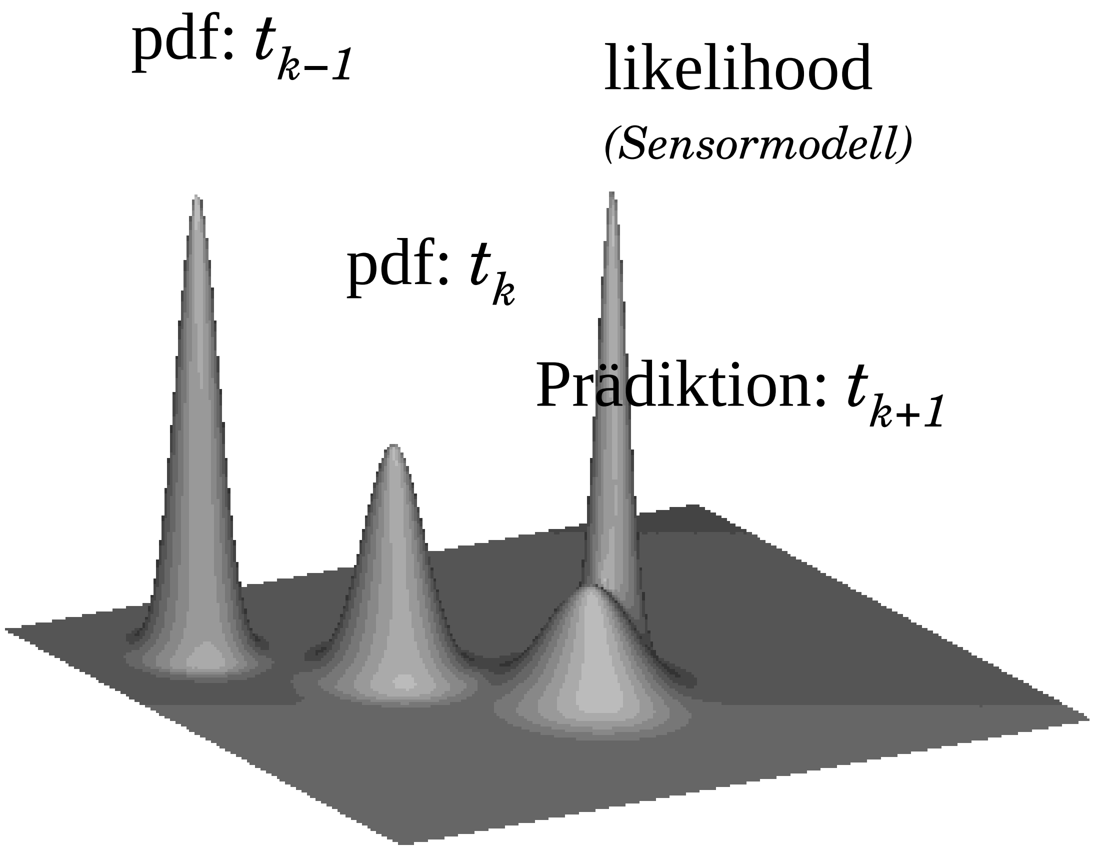
(not always: exploitation of ‘negative’ sensor evidence)



missing sensor information: increasing knowledge dissipation

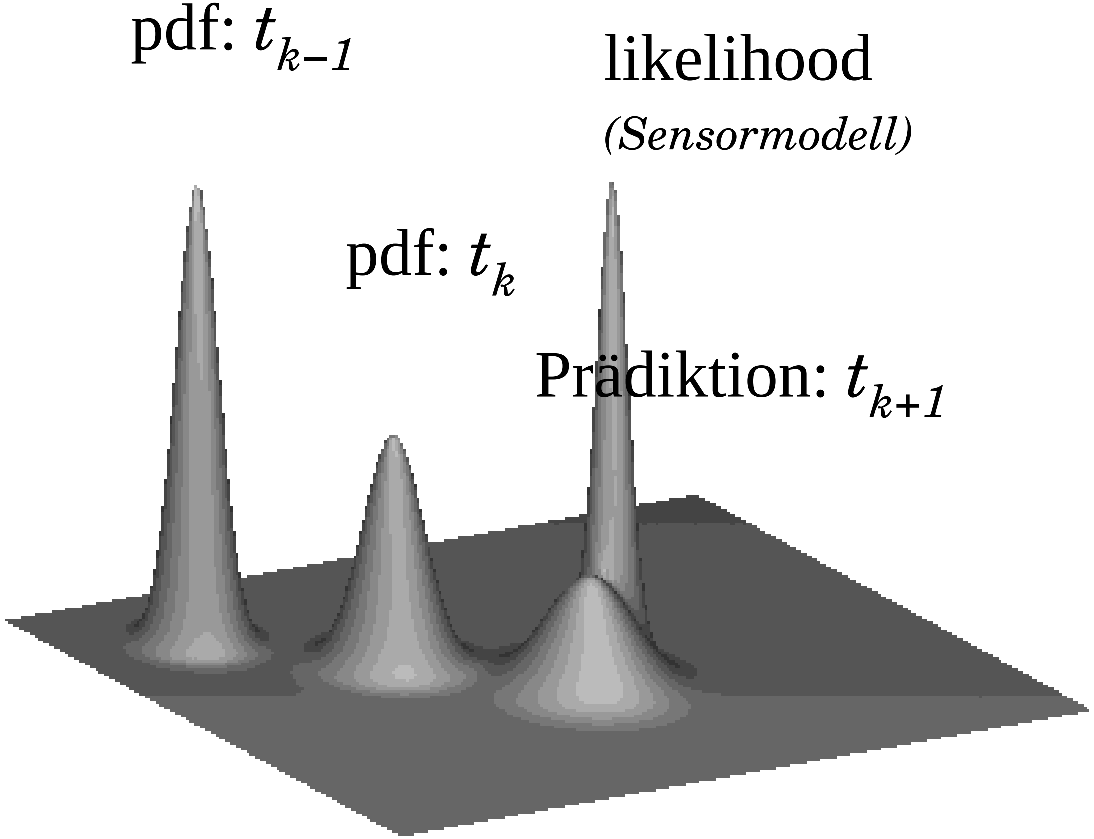


sensor information on the kinematical object state



Bayes' formula:

$$\underbrace{p(\mathbf{x}_{k+1}|\mathcal{Z}^{k+1})}_{\text{new knowledge}} = \frac{p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1})\, p(\mathbf{x}_{k+1}|\mathcal{Z}^k)}{\int d\mathbf{x}_{k+1}\, p(\underbrace{\mathbf{z}_{k+1}}_{\text{plot}}|\mathbf{x}_{k+1})\, \underbrace{p(\mathbf{x}_{k+1}|\mathcal{Z}^k)}_{\text{prediction}}}$$



BAYES' formula:

$$\underbrace{p(\mathbf{x}_{k+1}|\mathcal{Z}^{k+1})}_{\text{new knowledge}} = \frac{\mathcal{N}\big(\mathbf{z}_k;\ \mathbf{H}\mathbf{x}_k,\ \mathbf{R}\big)\ \mathcal{N}\big(\mathbf{x}_k;\ \mathbf{x}_{k|k-1},\ \mathbf{P}_{k|k-1}\big)}{\int d\mathbf{x}_{k+1}\ \mathcal{N}\big(\mathbf{z}_k;\ \mathbf{H}\mathbf{x}_k,\ \mathbf{R}\big)\ \underbrace{\mathcal{N}\big(\mathbf{x}_k;\ \mathbf{x}_{k|k-1},\ \mathbf{P}_{k|k-1}\big)}_{\text{prediction}}}$$

A Useful Product Formula for Gaussians

$$\mathcal{N}\big(\mathbf{z};\, \mathbf{H}\mathbf{x},\, \mathbf{R}\big)\ \mathcal{N}\big(\mathbf{x};\, \mathbf{y},\, \mathbf{P}\big) = \underbrace{\mathcal{N}\big(\mathbf{z};\, \mathbf{H}\mathbf{y},\, \mathbf{S}\big)}_{\text{independent of } \mathbf{x}}\ \mathcal{N}\big(\mathbf{x};\, \mathbf{y} + \mathbf{W}\boldsymbol{\nu},\, \mathbf{P} - \mathbf{W}\mathbf{S}\mathbf{W}^\top\big)$$

$$\boldsymbol{\nu} = \mathbf{z} - \mathbf{H}\mathbf{y}, \quad \mathbf{S} = \mathbf{H}\mathbf{P}\mathbf{H}^\top + \mathbf{D}, \quad \mathbf{W} = \mathbf{P}\mathbf{H}^\top\mathbf{S}^{-1}.$$

Kalman filter: $\mathbf{x}_k = (\mathbf{r}_k^\top, \dot{\mathbf{r}}_k^\top)^\top$, $\mathcal{Z}^k = \{\mathbf{z}_k, \mathcal{Z}^{k-1}\}$

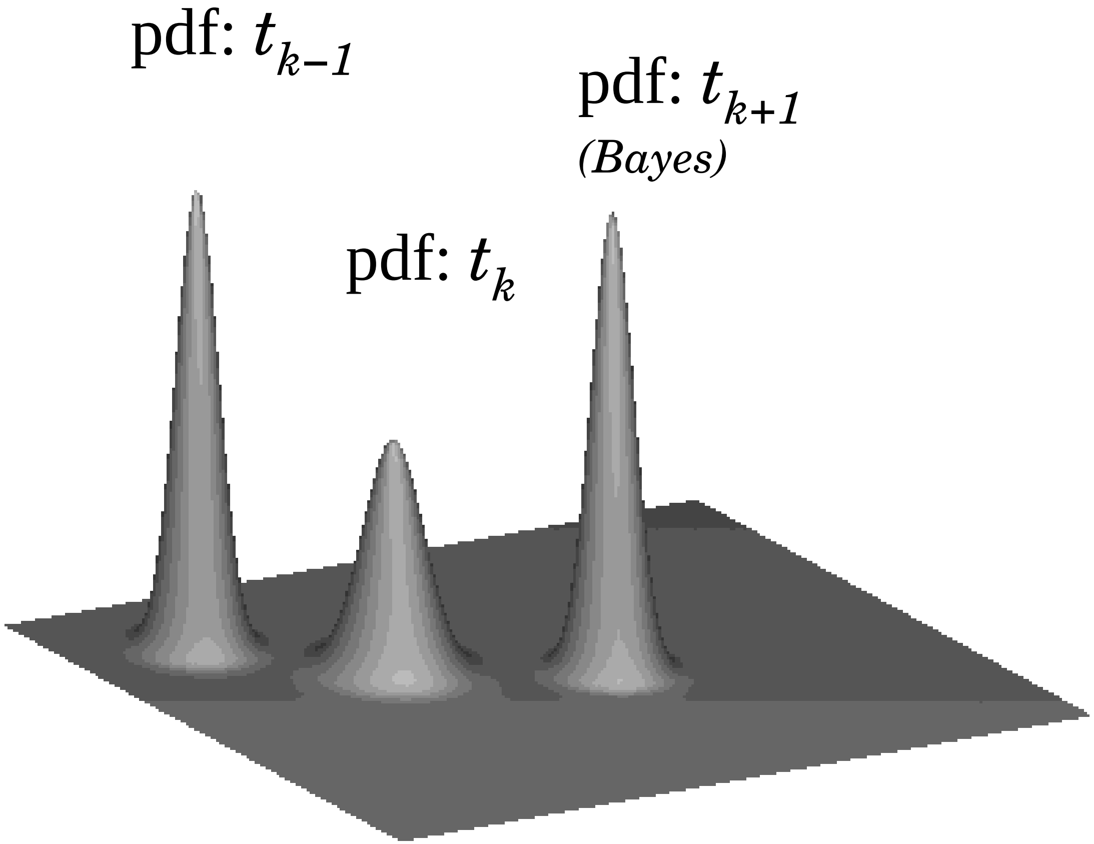
initiation: $p(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0;\, \mathbf{x}_{0|0},\, \mathbf{P}_{0|0})$, initial ignorance: $\mathbf{P}_{0|0}$ 'large'

prediction: $\mathcal{N}(\mathbf{x}_{k-1};\, \mathbf{x}_{k-1|k-1},\, \mathbf{P}_{k-1|k-1}) \xrightarrow[\mathbf{F}_{k|k-1},\, \mathbf{D}_{k|k-1}]{\text{dynamics model}} \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k-1},\, \mathbf{P}_{k|k-1})$

$$\mathbf{x}_{k|k-1} = \mathbf{F}_{k|k-1}\mathbf{x}_{k-1|k-1}$$
$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k|k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k|k-1}{}^\top + \mathbf{D}_{k|k-1}$$

filtering: $\mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k-1},\, \mathbf{P}_{k|k-1}) \xrightarrow[\text{sensor model: } \mathbf{H}_k,\, \mathbf{R}_k]{\text{current measurement } \mathbf{z}_k} \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k},\, \mathbf{P}_{k|k})$

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{W}_{k|k-1}\boldsymbol{\nu}_{k|k-1}, \qquad \boldsymbol{\nu}_{k|k-1} = \mathbf{z}_k - \mathbf{H}_k\mathbf{x}_{k|k-1}$$
$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{W}_{k|k-1}\mathbf{S}_{k|k-1}\mathbf{W}_{k|k-1}{}^\top, \qquad \mathbf{S}_{k|k-1} = \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k{}^\top + \mathbf{R}_k$$
$$\mathbf{W}_{k|k-1} = \mathbf{P}_{k|k-1}\mathbf{H}_k{}^\top\mathbf{S}_{k|k-1}{}^{-1} \qquad \text{'KALMAN gain matrix'}$$



filtering = sensor data processing

Summary: BAYESian (Multi-) Sensor Tracking

- **Basis:** In the course of time one or several sensors produce measurements of targets of interest. Each target is characterized by its current state vector, being expected to change with time.

- **Objective:** Learn as much as possible about the individual target states at each time by analyzing the ‘time series’ which is constituted by the sensor data.

Summary: BAYESian (Multi-) Sensor Tracking

- **Basis:** In the course of time one or several sensors produce measurements of targets of interest. Each target is characterized by its current state vector, being expected to change with time.

- **Objective:** Learn as much as possible about the individual target states at each time by analyzing the 'time series' which is constituted by the sensor data.

- **Problem:** imperfect sensor information: inaccurate, incomplete, and eventually ambiguous. Moreover, the targets' temporal evolution is usually not well-known.

Summary: BAYESian (Multi-) Sensor Tracking

- **Basis:** In the course of time one or several sensors produce measurements of targets of interest. Each target is characterized by its current state vector, being expected to change with time.
- **Objective:** Learn as much as possible about the individual target states at each time by analyzing the 'time series' which is constituted by the sensor data.
- **Problem:** imperfect sensor information: inaccurate, incomplete, and eventually ambiguous. Moreover, the targets' temporal evolution is usually not well-known.
- **Approach:** Interpret measurements and target vectors as random variables (RVs). Describe by probability density functions (pdf) what is known about them.

Summary: BAYESian (Multi-) Sensor Tracking

- **Basis:** In the course of time one or several sensors produce measurements of targets of interest. Each target is characterized by its current state vector, being expected to change with time.

- **Objective:** Learn as much as possible about the individual target states at each time by analyzing the 'time series' which is constituted by the sensor data.

- **Problem:** imperfect sensor information: inaccurate, incomplete, and eventually ambiguous. Moreover, the targets' temporal evolution is usually not well-known.

- **Approach:** Interpret measurements and target vectors as random variables (RVs). Describe by probability density functions (pdf) what is known about them.

- **Solution:** Derive iteration formulae for calculating the pdfs! Develop a mechanism for initiation! By doing so, exploit all background information available! Derive state estimates from the pdfs along with appropriate quality measures!

How to deal with probability density functions?

- pdf $p(x)$: Extract *probability statements* about the RV x by integration!
- naïvely: *positive* and *normalized* functions $(p(x) \geq 0, \int dx\, p(x) = 1)$
- *conditional pdf* $p(x|y) = \frac{p(x,y)}{p(y)}$: Impact of information on y on RV x?
- *marginal density* $p(x) = \int dy\, p(x,y) = \int dy\, p(x|y)\, p(y)$: Enter y!
- Bayes: $p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\int dx\, p(y|x)p(x)}$: $p(x|y) \leftarrow p(y|x), p(x)$!

How to deal with probability density functions?

- pdf $p(x)$: Extract *probability statements* about the RV x by integration!
- naïvely: *positive* and *normalized* functions ($p(x) \geq 0$, $\int dx\, p(x) = 1$)
- *conditional pdf* $p(x|y) = \frac{p(x,y)}{p(y)}$: Impact of information on y on RV x?
- *marginal density* $p(x) = \int dy\, p(x,y) = \int dy\, p(x|y)\, p(y)$: Enter y!
- Bayes: $p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\int dx\, p(y|x)p(x)}$: $p(x|y) \leftarrow p(y|x), p(x)$!
- *certain knowledge* on x: $p(x) = \delta(x-y)$ '$=$' $\lim_{\sigma\to 0} \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\frac{(x-y)^2}{\sigma^2}}$

Fraunhofer FKIE

How to deal with probability density functions?

- pdf $p(x)$: Extract *probability statements* about the RV x by integration!
- naïvely: *positive* and *normalized* functions ($p(x) \geq 0$, $\int dx\, p(x) = 1$)
- *conditional pdf* $p(x|y) = \frac{p(x,y)}{p(y)}$: Impact of information on y on RV x?
- *marginal density* $p(x) = \int dy\, p(x,y) = \int dy\, p(x|y)\, p(y)$: Enter y!
- Bayes: $p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\int dx\, p(y|x)p(x)}$: $p(x|y) \leftarrow p(y|x), p(x)$!
- *certain knowledge* on x: $p(x) = \delta(x-y)$ '=' $\lim_{\sigma\to 0} \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\frac{(x-y)^2}{\sigma^2}}$
- *transformed RV* $y = t[x]$: $p(y) = \int dx\, p(y,x) = \int dx\, p(y|x)\, p_x(x)$

How to deal with probability density functions?

- pdf $p(x)$: Extract *probability statements* about the RV x by integration!
- naïvely: *positive* and *normalized* functions ($p(x) \geq 0$, $\int dx\, p(x) = 1$)
- *conditional pdf* $p(x|y) = \frac{p(x,y)}{p(y)}$: Impact of information on y on RV x?
- *marginal density* $p(x) = \int dy\, p(x,y) = \int dy\, p(x|y)\, p(y)$: Enter y!
- Bayes: $p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\int dx\, p(y|x)p(x)}$: $p(x|y) \leftarrow p(y|x), p(x)$!
- *certain knowledge* on x: $p(x) = \delta(x-y)$ '=' $\lim_{\sigma \to 0} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-y)^2}{\sigma^2}}$
- *transformed* $RV\, y = t[x]$: $p(y) = \int dx p(y,x) = \int dx p(y|x) p_x(x) =$
  $\int dx\, \delta(y - t[x])\, p_x(x) =: [\mathcal{T} p_x](y)$ ($\mathcal{T} : p_x \mapsto p$, "transfer operator")

Fraunhofer FKIE

Create your own ground truth generator!

Exercise 2.1

Consider an object that moves in two dimensions on the trajectory:

$$\mathbf{r}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = A \begin{pmatrix} \sin(\omega t) \\ \sin(2\omega t) \end{pmatrix} \quad \text{with} \quad A = \frac{v^2}{q}, \quad \omega = \frac{q}{2v}$$

and speed and acceleration parameters: $v = 300\,\frac{\mathrm{m}}{\mathrm{s}}$, $q = 9\,\frac{\mathrm{m}}{\mathrm{s}^2}$!

1. Plot the trajectory. Why is it periodical? What is its period $T = T(v, q)$?
2. Show for the velocity and acceleration vector:

$$\dot{\mathbf{r}}(t) = v \begin{pmatrix} \cos(\omega t)/2 \\ \cos(2\omega t) \end{pmatrix}, \quad \ddot{\mathbf{r}}(t) = -q \begin{pmatrix} \sin(\omega t)/4 \\ \sin(2\omega t) \end{pmatrix}!$$

3. Calculate for each instance of time t the tangential and normal vectors in $\mathbf{r}(t)$:

$$\mathbf{t}(t) = \frac{1}{|\dot{\mathbf{r}}(t)|} \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix}, \quad \mathbf{n}(t) = \frac{1}{|\dot{\mathbf{r}}(t)|} \begin{pmatrix} -\dot{y}(t) \\ \dot{x}(t) \end{pmatrix}!$$

4. Plot $|\dot{\mathbf{r}}(t)|$, $|\ddot{\mathbf{r}}(t)|$, $\ddot{\mathbf{r}}(t)\mathbf{t}(t)$ and $\ddot{\mathbf{r}}(t)\mathbf{n}(t)$ over a period T!
5. Discuss the temporal behaviour based on the trajectory $\mathbf{r}(t)$!
6. What are the maximum speeds and accelerations, v_{max}, q_{max}?

Characterize an object by *quantitatively describable* properties: object state

Examples:

- object position x on a strait line: $x \in \mathbb{R}$
- kinematic state $\mathbf{x} = (\mathbf{r}^\top, \dot{\mathbf{r}}^\top, \ddot{\mathbf{r}}^\top)^\top$, $\mathbf{x} \in \mathbb{R}^9$

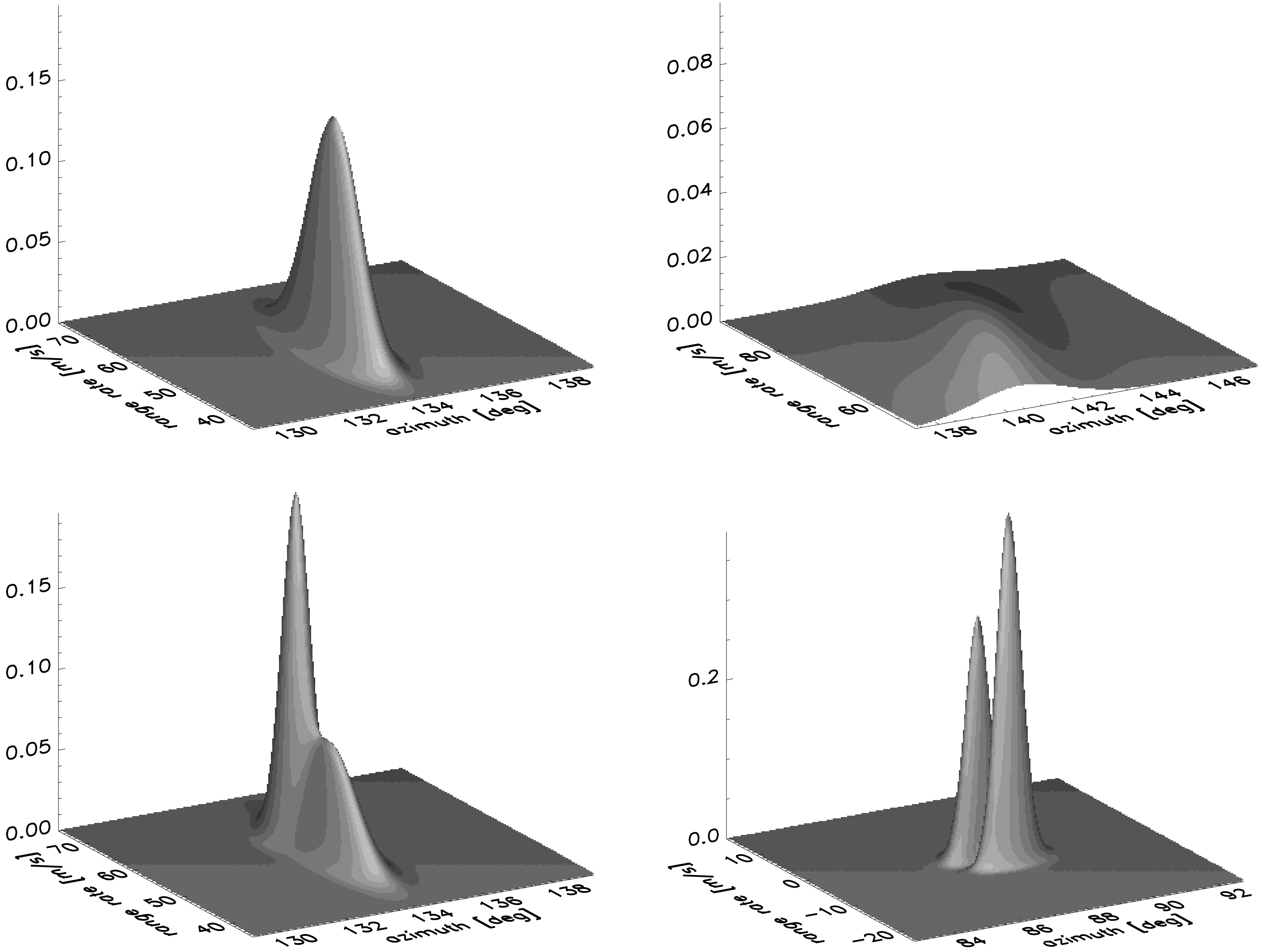
  position $\mathbf{r} = (x, y, z)^\top$, velocity $\dot{\mathbf{r}}$, acceleration $\ddot{\mathbf{r}}$
- joint state of two objects: $\mathbf{x} = (\mathbf{x}_1^\top, \mathbf{x}_2^\top)^\top$
- kinematic state $\mathbf{x}$, object extension $\mathbf{X}$

  z.B. ellipsoid: symmetric, positively definite matrix
- kinematic state $\mathbf{x}$, object class class

  z.B. bird, sailing plane, helicopter, passenger jet, ...

Learn unknown object states from imperfect measurements and
describe by functions $p(\mathbf{x})$ imprecise knowledge mathematically precisely!

Interpret unknown object states as *random variables*, x [1D] or $\mathbf{x}$, $\mathbf{X}$ [vector / matrix variate]), characterized by corresponding *probability density functions* (pdf).



The concrete shape of the pdf $p(\mathbf{x})$ contains the full knowledge on $\mathbf{x}$!

Information on a random variable (RV) can be extracted by integration from the corresponding pdf. !

at present: one dimensional case:

How probable is it that $x \in (a, b) \subseteq \mathbb{R}$ ***holds?***

Answer: $P\{x \in (a, b)\} = \int_a^b dx\, p(x) \quad \Rightarrow \quad p(x) \geq 0$

in particular: $P\{x \in \mathbb{R}\} = \int_{-\infty}^{\infty} dx\, p(x) = 1$ (normalzation)

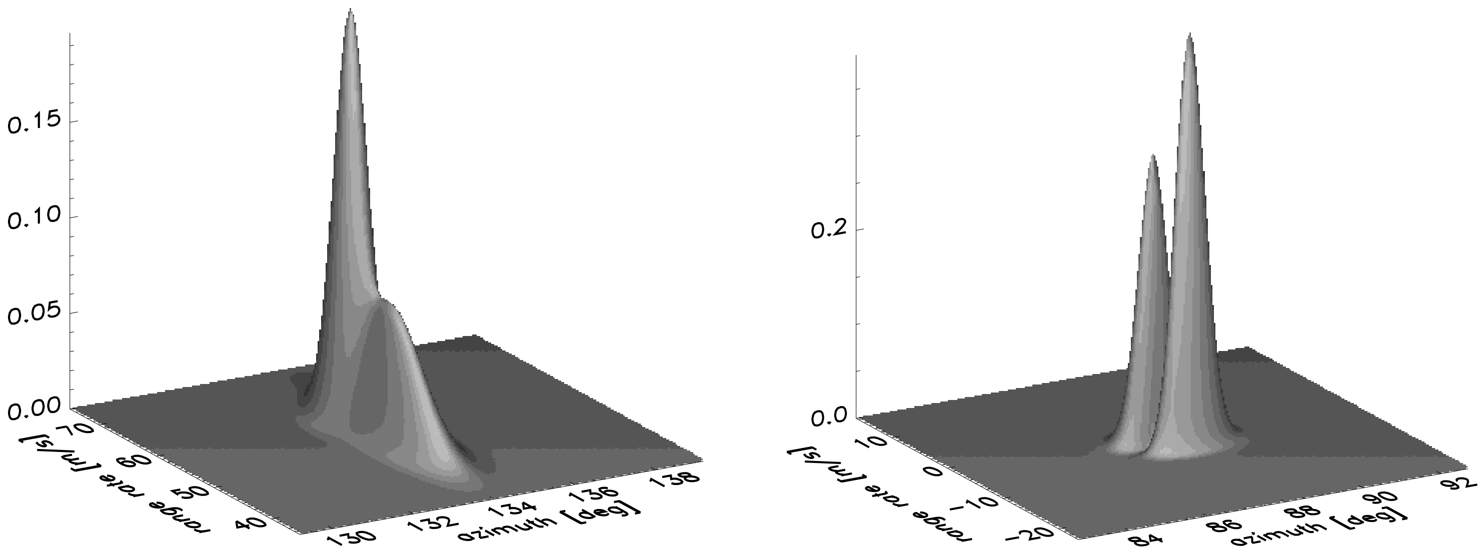
intuitive interpretation: "*the object is somewhere in* $\mathbb{R}$"

loosely: $p(x)\, dx$ **is probabity for** x **having a value between** x **and** $x + dx$

How to characterize the properties of a pdf?

specifically: How to associate a single “expected” value to a RV?

The maximum of the pdf is sometimes but not always useful!



Fraunhofer FKIE

How to characterize the properties of a pdf?

specifically: How to associate a single "expected" value to a RV?

The maximum of the pdf is sometimes but not always useful! ($\rightarrow$ examples)

instead: Calculate the centroid of the pdf!

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} dx \, x \, p(x) = \bar{x}$$ "expectation value"

more generally: Consider functions $g : x \mapsto g(x)$ of the RV x!

$$\mathbb{E}[g(x)] = \int_{-\infty}^{\infty} dx \, g(x) \, p(x),$$ "expectation value of the *observable* g"

Example: Consider the observable $\frac{1}{2}mx^2$ (kinetic energy, x = speed)

An important observable: the “error” of an estimate

- **Quality:** How useful is an expectation value $\bar{x} = \mathbb{E}[x]$?

Consider special obervables as distance measure:

$$g(x) = |x - \bar{x}| \quad \text{oder} \quad g(x) = (x - \bar{x})^2$$

quadratic measures: computationally more comfortable!

‘expected error’ of the expectation value $\bar{x}$:

$$\boxed{\mathbb{V}[x] = \mathbb{E}[(x - \bar{x})^2],\ \sigma_x = \sqrt{\mathbb{V}[x]}}$$

variance, standard deviation

Exercise 2.2

Show that $\boxed{\mathbb{V}[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2}$ holds.

Expectation value of the observable x^2 also called “2nd moment” of the pdf of x.

Calculate expectation and variance of the uniform density of a RV $x \in \mathbb{R}$ in the intervall $[a, b]$.

Exercise 2.3

$$p(x) = \mathcal{U}(\underbrace{x}_{\text{ZV}}; \underbrace{a, b}_{\text{Parameter}}) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{sonst} \end{cases}$$

!

Pdf correctly normalized? $$\int_{-\infty}^{\infty} dx \, \mathcal{U}(x; a, b) = \frac{1}{b-a} \int_a^b dx = 1$$

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} dx \, x \, \mathcal{U}(x; a, b) = \frac{b+a}{2}$$

$$\mathbb{V}[x] = \frac{1}{b-a} \int_a^b dx \, x^2 - \mathbb{E}[x]^2 = \frac{1}{12}(b-a)^2$$

Important example: x *normally distributed* over $\mathbb{R}$ (Gauss)

- *wanted:* probabilities concentrated around μ
- quadratic distance: $||x-\mu||^2 = \frac{1}{2}(x-\mu)^2/\sigma^2$ (mathematically convenient!)
- Parameter σ is a measure of the "width" of the pdf: $||\sigma||^2 = \frac{1}{2}$
- for 'large' distances, i.e. $||x - \mu||^2 \gg \frac{1}{2}$, the pdf shall decay quickly.
- simplest approach: $\tilde{p}(x) = \mathrm{e}^{-||x-\mu||^2}$ $\quad$ ($> 0\ \forall x \in \mathbb{R}$, normalization?)
- Normalized for $p(x) = \tilde{p}(x) / \int_{-\infty}^{\infty} dx\, \tilde{p}(x)$!

Formula collection delivers: $\int_{-\infty}^{\infty} dx\, \tilde{p}(x) = \sqrt{2\pi}\sigma$

An admissible pdf with the required properties is obviously given by:

$$\mathcal{N}(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Exercise 2.4

Show for the Gaussian density $p(x) = \mathcal{N}(x; \mu, \sigma)$:

$$\boxed{\mathbb{E}[x] = \mu,\ \mathbb{V}[x] = \sigma^2}$$

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} dx\, x\, \mathcal{N}(x; \mu, \sigma) = \mu$$

$$\mathbb{V}[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2 = \sigma^2$$

Use substitution and partial integration!

Use $\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\frac{1}{2}x^2} = \sqrt{2\pi}$!