The Kalman Filter
(part 2)

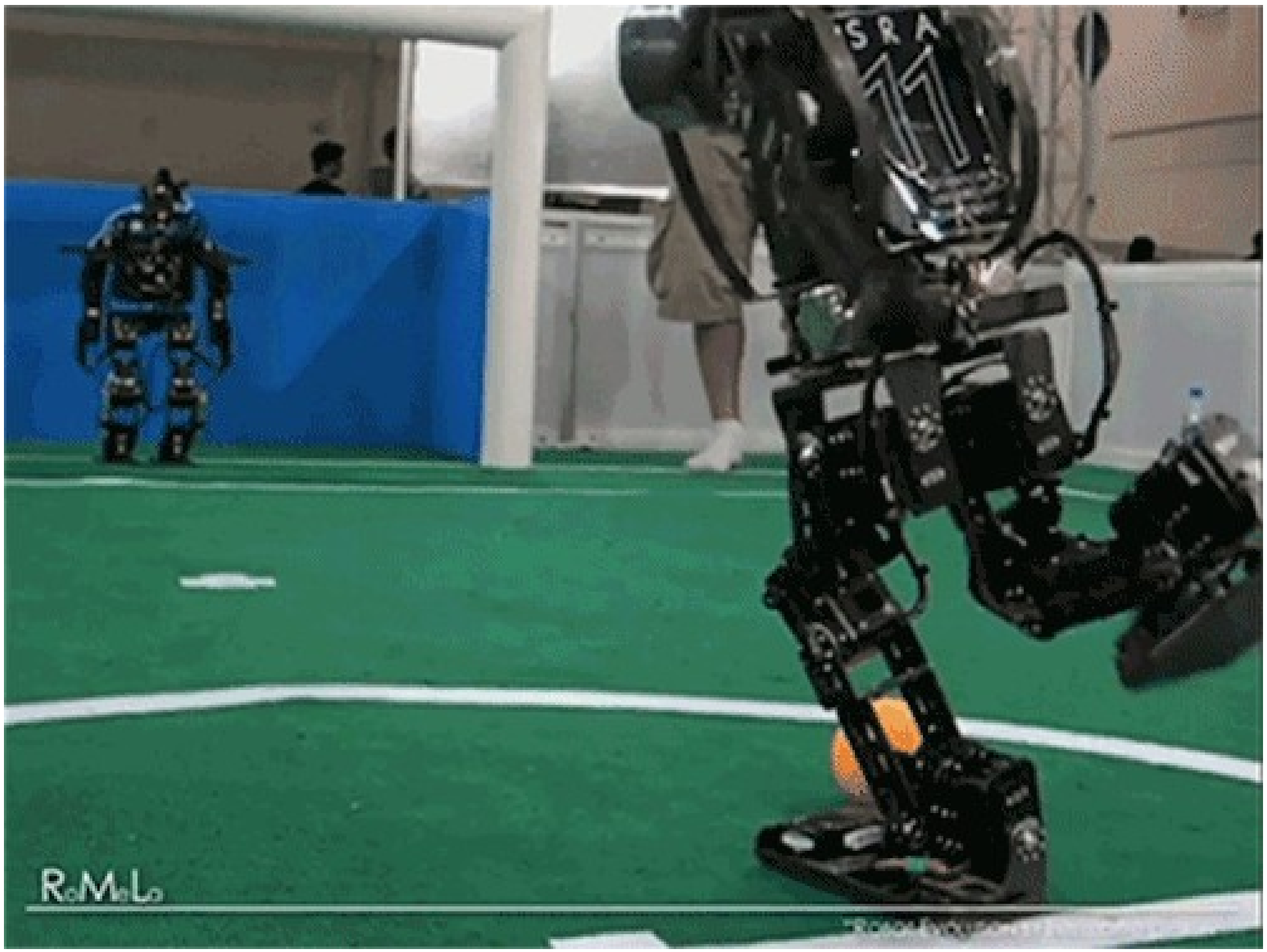
Reading Assignment

- Chapter 4 of PR
  - Focus on histogram and particle filters

Homework 1

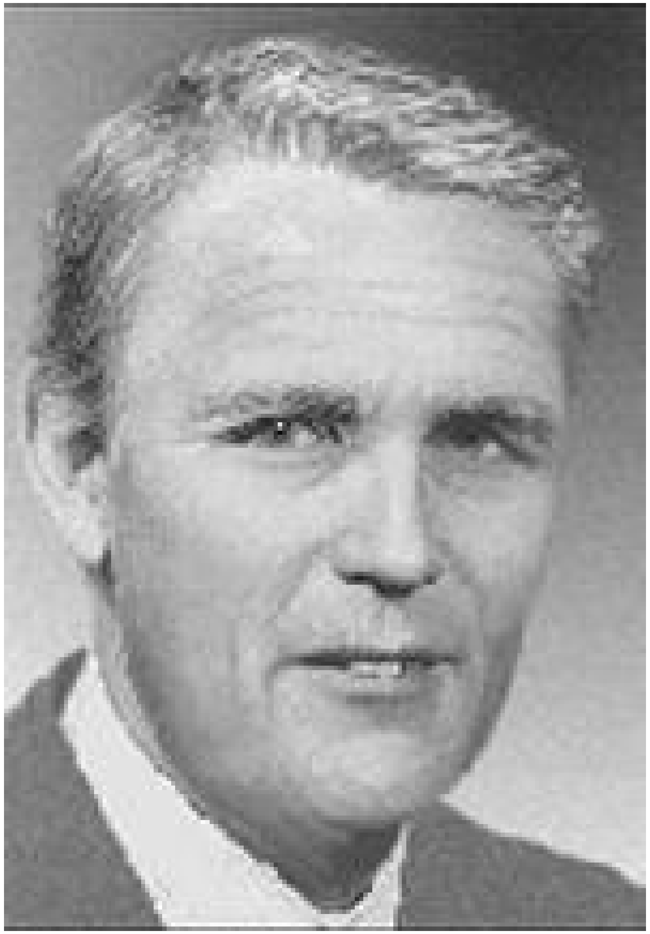
- See canvas – will preview at end of class

Something fun



Administrative Stuff

Rudolf Emil Kalman



[http://www.cs.unc.edu/~welch/kalman/kalmanBiblio.html]

Definition

- A Kalman filter is simply an optimal recursive data processing algorithm

- Under some assumptions the Kalman filter is optimal with respect to virtually any criterion that makes sense.

Definition

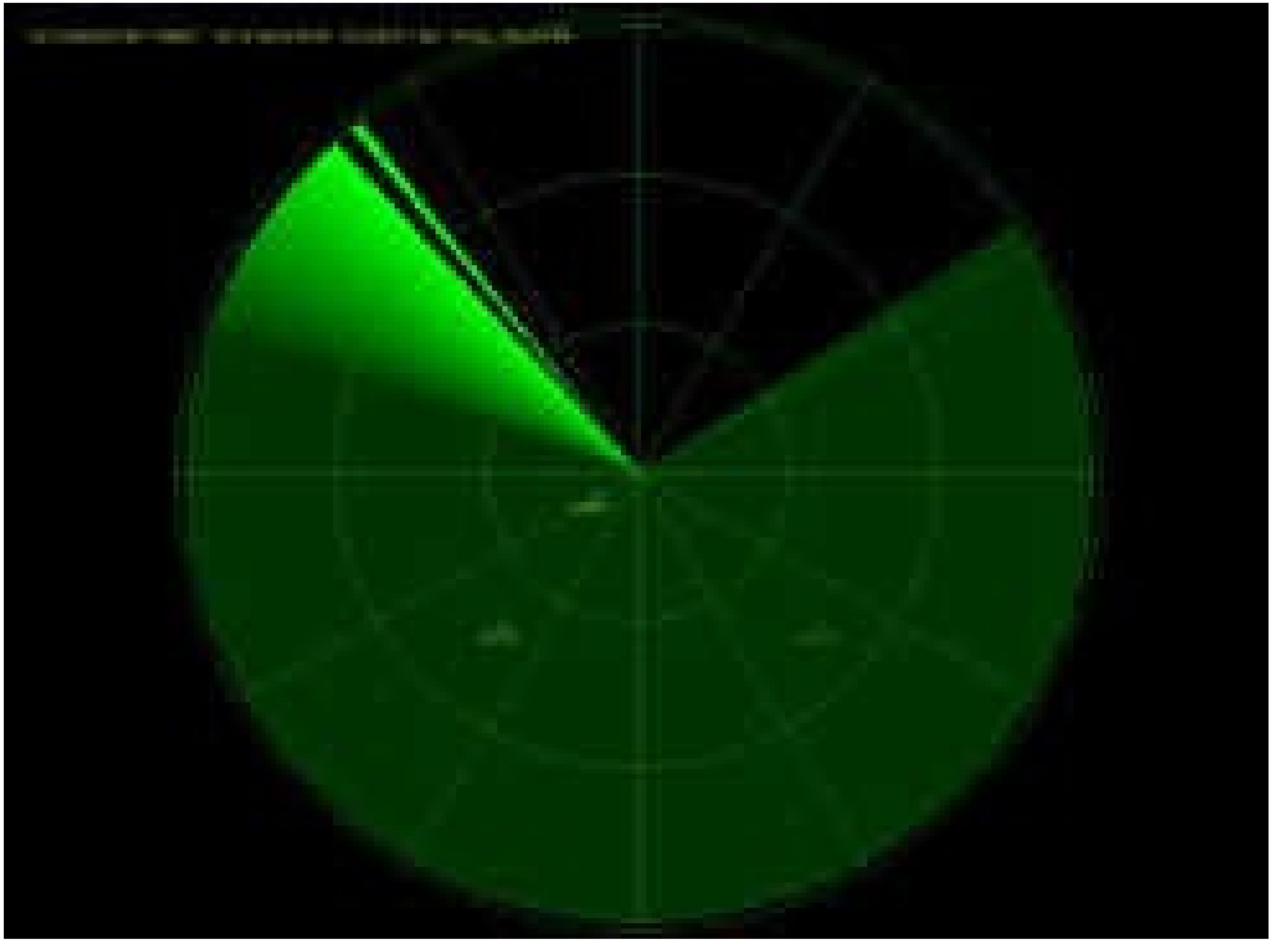
“The Kalman filter incorporates all information that can be provided to it. It processes all available measurements, regardless of their precision, to estimate the current value of the variables of interest.”

[Maybeck (1979)]

Why do we need a filter?

- No mathematical model of a real system is perfect
- Real world disturbances
- Imperfect Sensors

Application: Radar Tracking

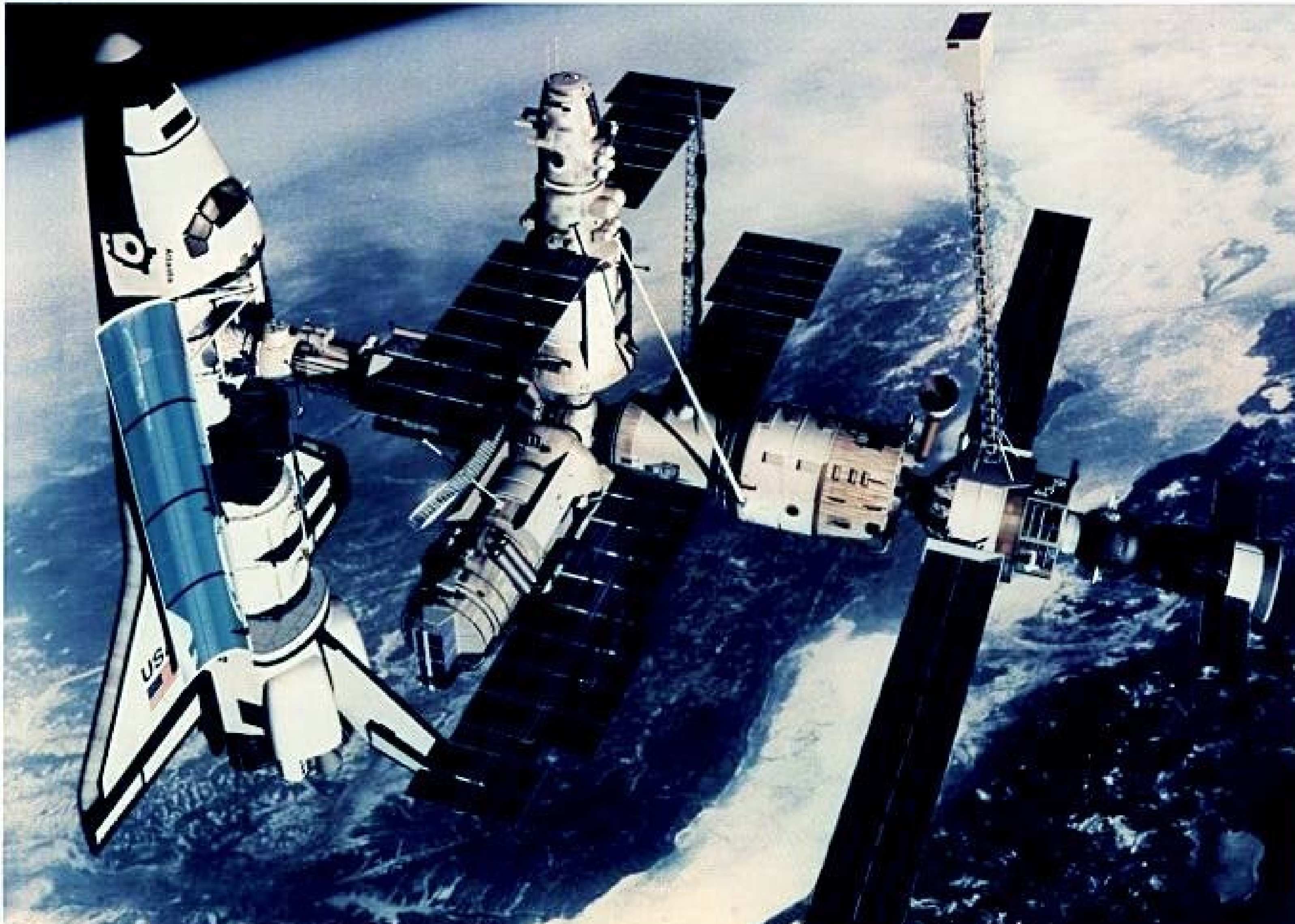


Application: Lunar Landing

https://github.com/chrislgarry/Apollo-11

National Aeronautics and
Space Administration

Shuttle Docking with Russian *Mir* Space Station



Application: Missile Tracking

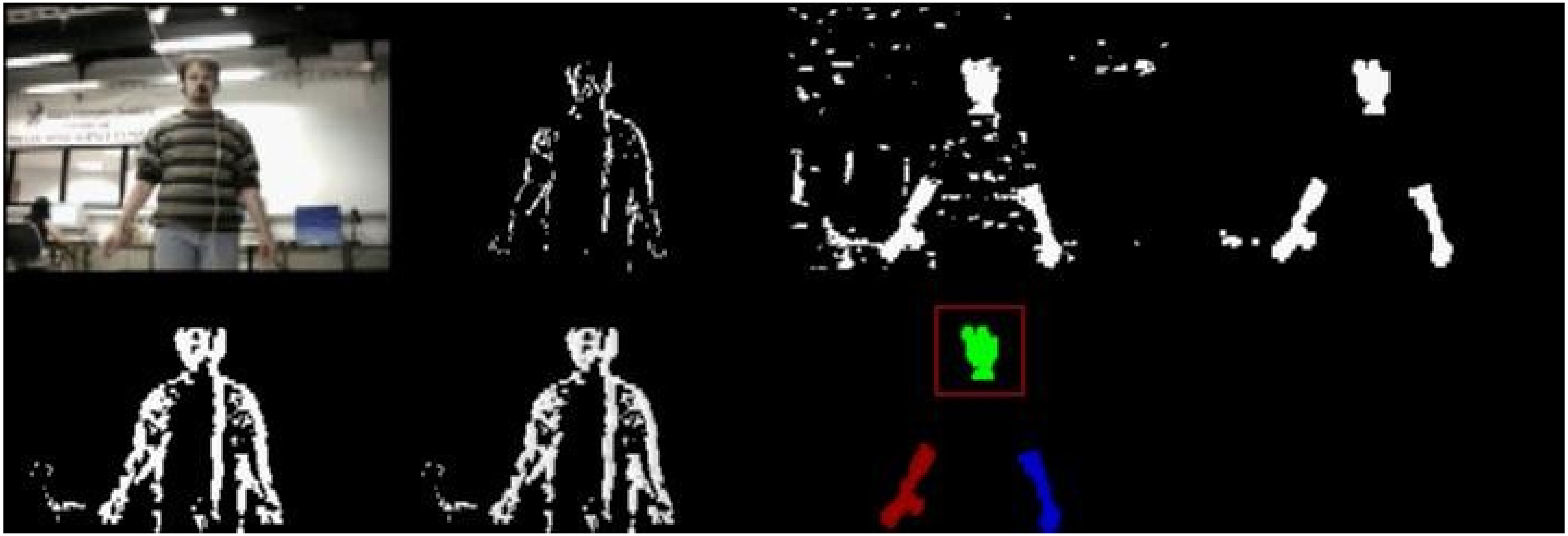
Application: Sailing

Application: Robot Navigation

Application: Other Tracking

Application: Head Tracking

Face & Hand Tracking



A Simple Recursive Example

- Problem Statement:

Given the measurement sequence:

$z_1, z_2, \ldots, z_n$ find the mean

[Brown and Hwang (1992)]

First Approach

1. Make the first measurement z_1

   Store z_1 and estimate the mean as $\mu_1 = z_1$

2. Make the second measurement z_2

   Store z_1 along with z_2 and estimate the mean as

   $\mu_2 = (z_1 + z_2)/2$

[Brown and Hwang (1992)]

First Approach (cont'd)

3. Make the third measurement z_3
   Store z_3 along with z_1 and z_2 and estimate the mean as

$$\mu_3 = (z_1+z_2+z_3)/3$$

[Brown and Hwang (1992)]

First Approach (cont'd)

n. Make the n-th measurement z_n

   Store z_n along with z_1 , z_2 ,…, z_{n-1} and estimate the mean as

$$\mu_n = (z_1 + z_2 + \ldots + z_n)/n$$

[Brown and Hwang (1992)]

Second Approach

1. Make the first measurement z_1

   Compute the mean estimate as

$$\mu_1 = z_1$$

   Store μ_1 and discard z_1

[Brown and Hwang (1992)]

Second Approach (cont'd)

2. Make the second measurement z_2

Compute the estimate of the mean as a weighted sum of the previous estimate μ_1 and the current measurement z_2:

$$\mu_2 = 1/2\ \mu_1 + 1/2\ z_2$$

Store μ_2 and discard z_2 and μ_1

[Brown and Hwang (1992)]

Second Approach (cont'd)

3. Make the third measurement z_3

   Compute the estimate of the mean as a weighted sum of the previous estimate μ_2 and the current measurement z_3:

$$\mu_3 = 2/3\ \mu_2 + 1/3\ z_3$$

Store μ_3 and discard z_3 and μ_2

[Brown and Hwang (1992)]

Second Approach (cont'd)

n. Make the n-th measurement z_n

Compute the estimate of the mean as a weighted sum of the previous estimate μ_{n-1} and the current measurement z_n:

$$\mu_n = (n-1)/n \ \mu_{n-1} + 1/n \ z_n$$

Store μ_n and discard z_n and μ_{n-1}

[Brown and Hwang (1992)]

Comparison

$$\hat{x}_1 = z_1$$

$$\hat{x}_2 = \frac{z_1 + z_2}{2}$$

$$\hat{x}_3 = \frac{z_1 + z_2 + z_3}{3}$$

$$\hat{x}_n = \frac{z_1 + z_2 + \ldots + z_n}{n}$$

Batch Method

$$\hat{x}_1 = z_1$$

$$\hat{x}_2 = \frac{1}{2}\hat{x}_1 + \frac{1}{2}z_2$$

$$\hat{x}_3 = \frac{2}{3}\hat{x}_2 + \frac{1}{3}z_3$$

$$\hat{x}_n = \frac{n-1}{n}\hat{x}_{n-1} + \frac{1}{n}z_n$$

Recursive Method

Analysis

- The second procedure gives the same result as the first procedure.

- It uses the result for the previous step to help obtain an estimate at the current step.

- The difference is that it does not need to keep the sequence in memory.

[Brown and Hwang (1992)]

Second Approach (rewrite the general formula)

$$\mu_n = (n-1)/n \ \mu_{n-1} + 1/n \ z_n$$

$$\mu_n = (n/n) \ \mu_{n-1} - (1/n) \ \mu_{n-1} + 1/n \ z_n$$

Second Approach
(rewrite the general formula)

$$\mu_n = (n-1)/n \ \mu_{n-1} + 1/n \ z_n$$

$$\mu_n = (n/n) \ \mu_{n-1} - (1/n) \ \mu_{n-1} + 1/n \ z_n$$

$$\mu_n = \underbrace{\mu_{n-1}}_{\text{Old Estimate}} + \underbrace{1/n}_{\text{Gain Factor}} \ \underbrace{(z_n - \mu_{n-1})}_{\text{Difference Between New Reading and Old Estimate}}$$

Old Estimate

Gain Factor

Difference Between New Reading and Old Estimate

Second Approach
(rewrite the general formula)

$$\begin{aligned}\hat{x}_n &= \left(\frac{n-1}{n}\right)\hat{x}_{n-1} + \left(\frac{1}{n}\right)z_n \\ &= \left(\frac{n}{n}\right)\hat{x}_{n-1} - \left(\frac{1}{n}\right)\hat{x}_{n-1} + \left(\frac{1}{n}\right)z_n \\ &= \underbrace{\hat{x}_{n-1}}_{\text{Old Estimate}} + \underbrace{\frac{1}{n}}_{\text{Gain Factor}}\underbrace{\left(z_n - \hat{x}_{n-1}\right)}_{\text{Difference Between New Reading and Old Estimate}}\end{aligned}$$

Old Estimate

Gain Factor

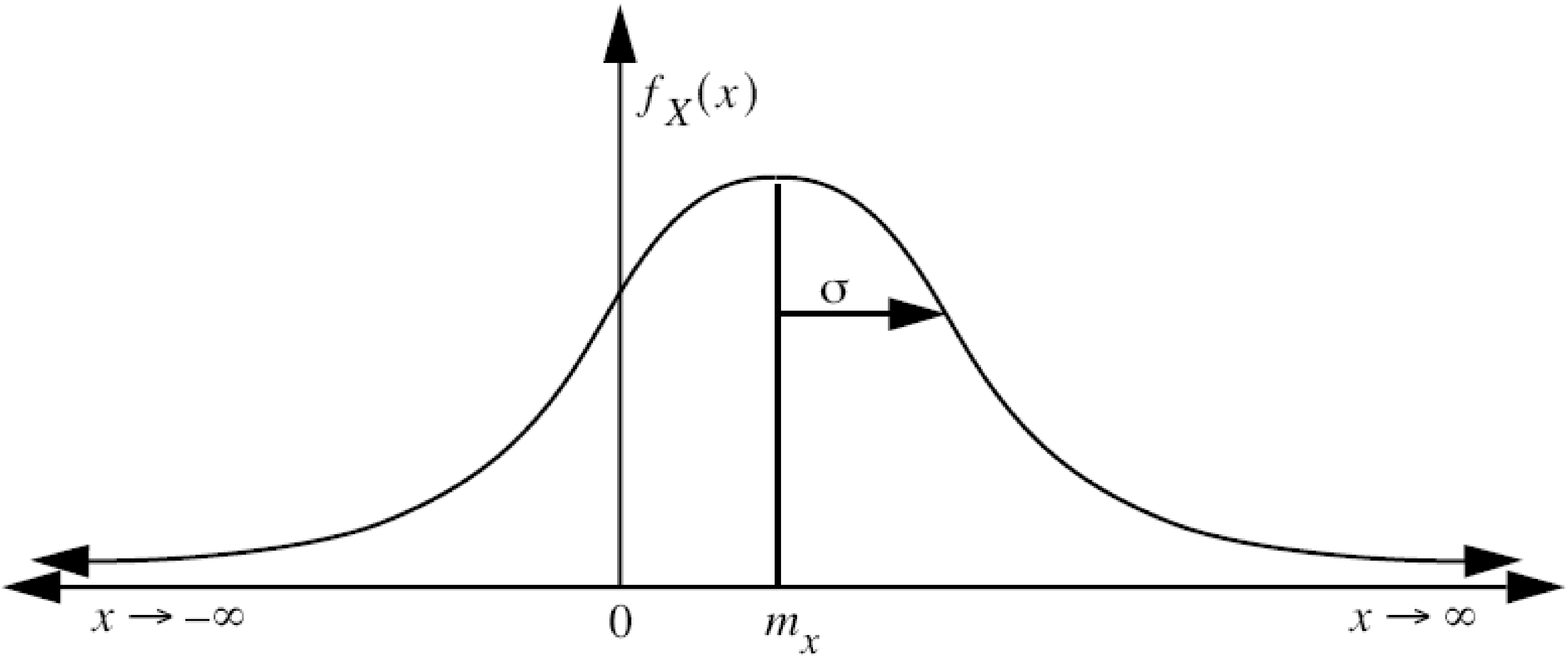
Difference Between New Reading and Old Estimate

Gaussian Properties

The Gaussian Function

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

Gaussian pdf



Properties

- If $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$
- Then $Y \sim N(a\mu + b, a^2\sigma^2)$

pdf for $Y = aX + b$

$$f_Y(y) = \frac{1}{\sqrt{2\pi a^2\sigma^2}} e^{-\frac{1}{2}\frac{(y-(a\mu+b))^2}{a^2\sigma^2}}$$

Properties

Finally, if X_1 and X_2 are independent (see Section 2.5 below), $X_1 \sim N(\mu_1, \sigma_1^2)$, and $X_2 \sim N(\mu_2, \sigma_2^2)$, then

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2), \tag{2.14}$$

and the density function becomes

$$f_X(x_1 + x_2) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-\frac{1}{2}\frac{(x - (\mu_1 + \mu_2))^2}{(\sigma_1^2 + \sigma_2^2)}}. \tag{2.15}$$

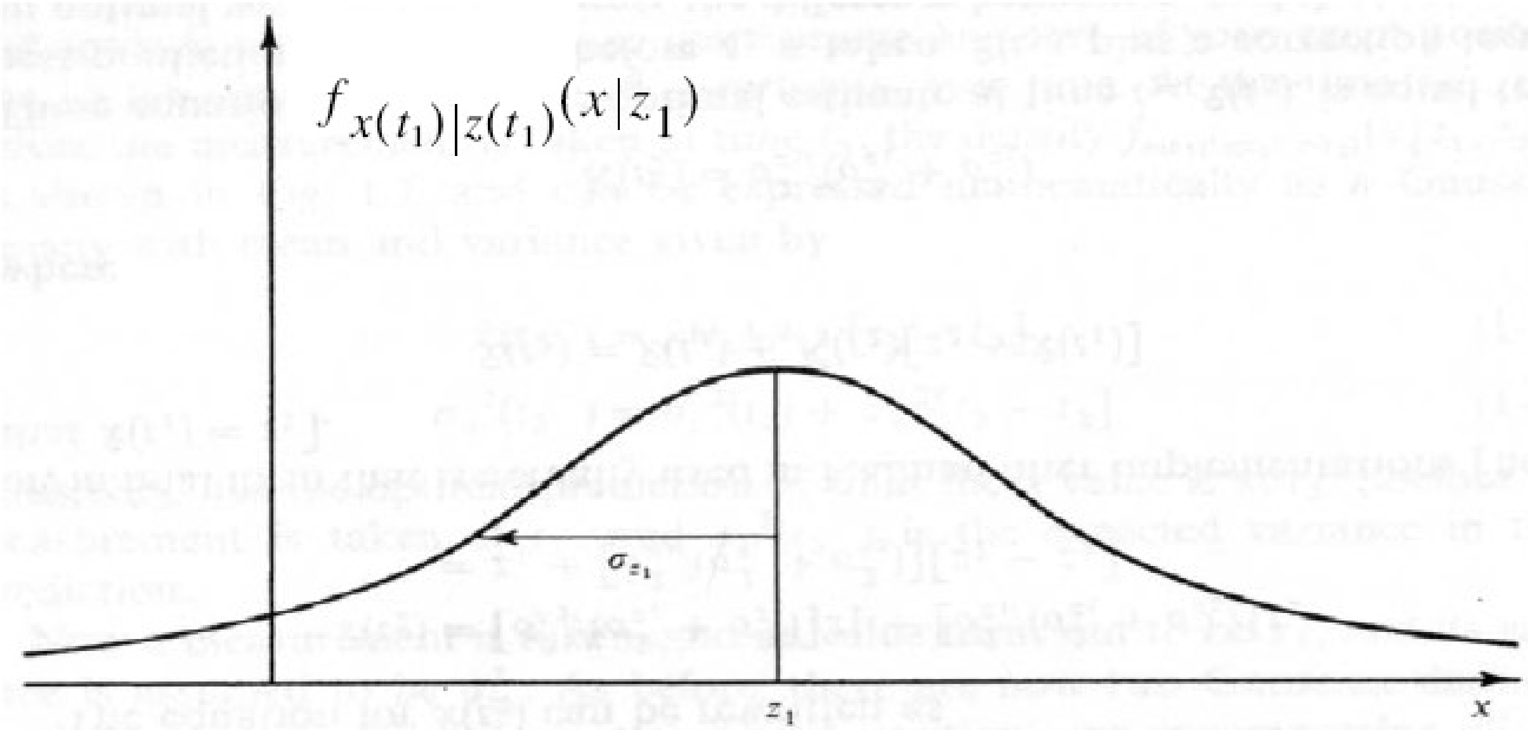
Summation and Subtraction

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$$

$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y$$

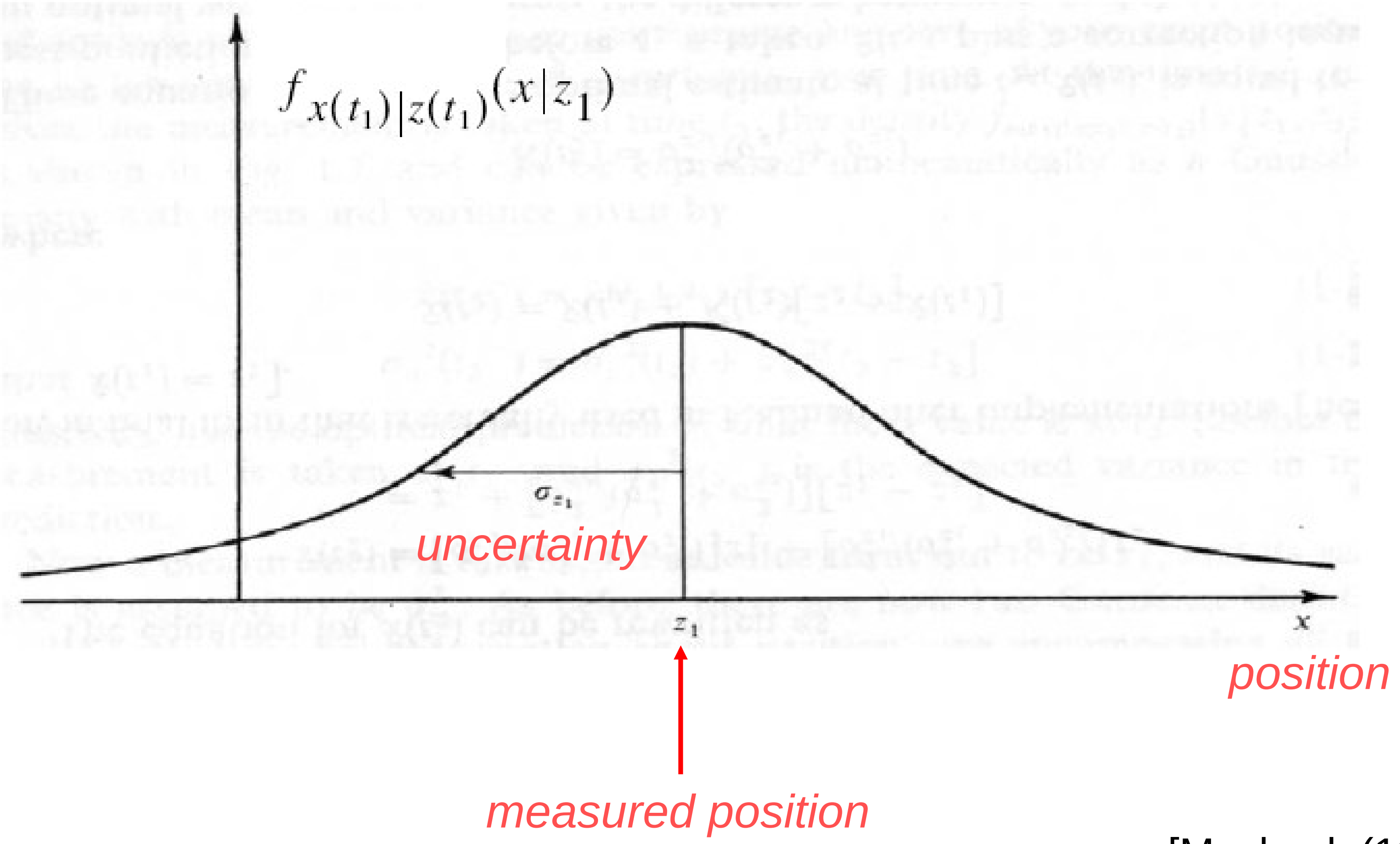
A simple example using diagrams

Conditional density of position based on measured value of z_1



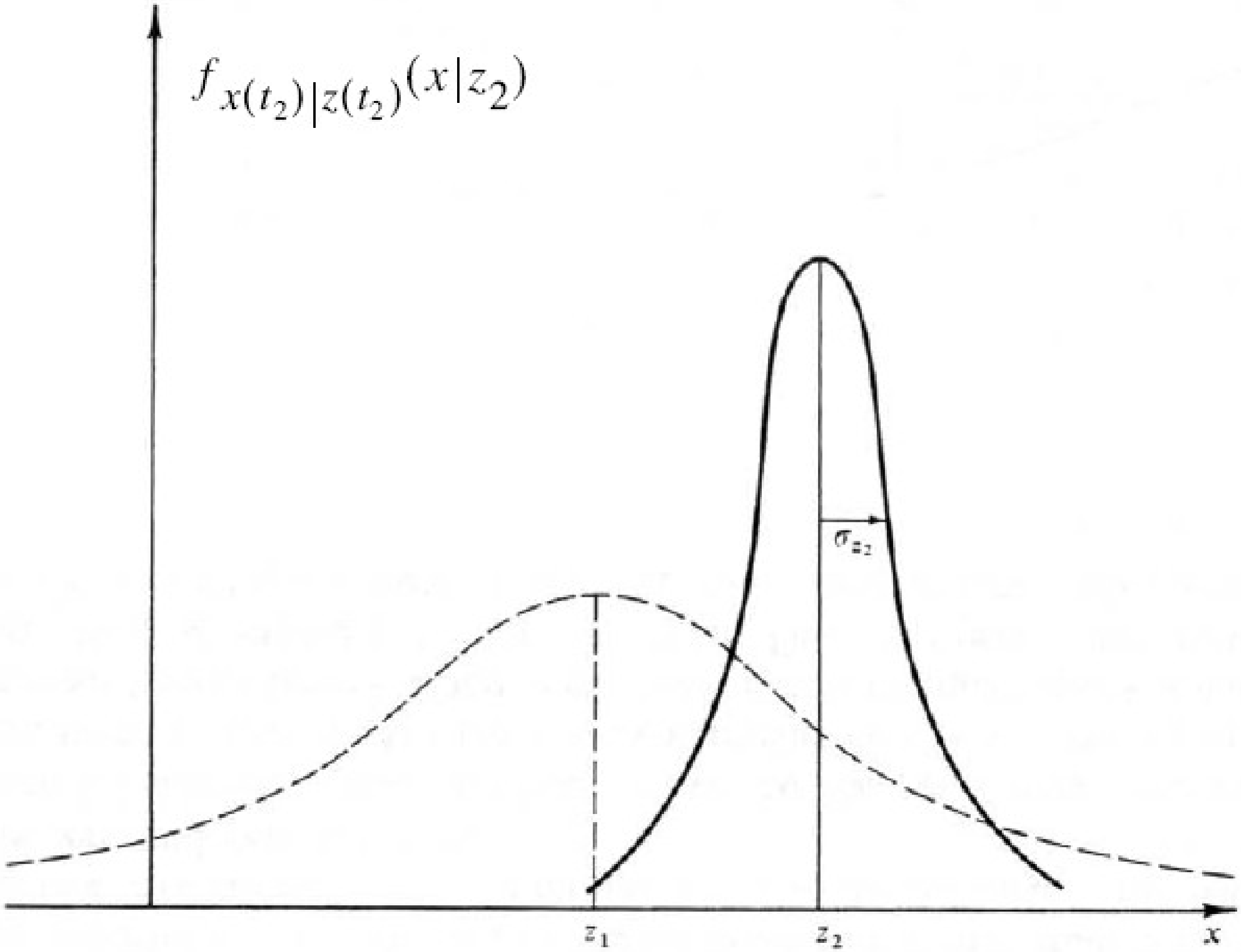
[Maybeck (1979)]

Conditional density of position based on measured value of z_1



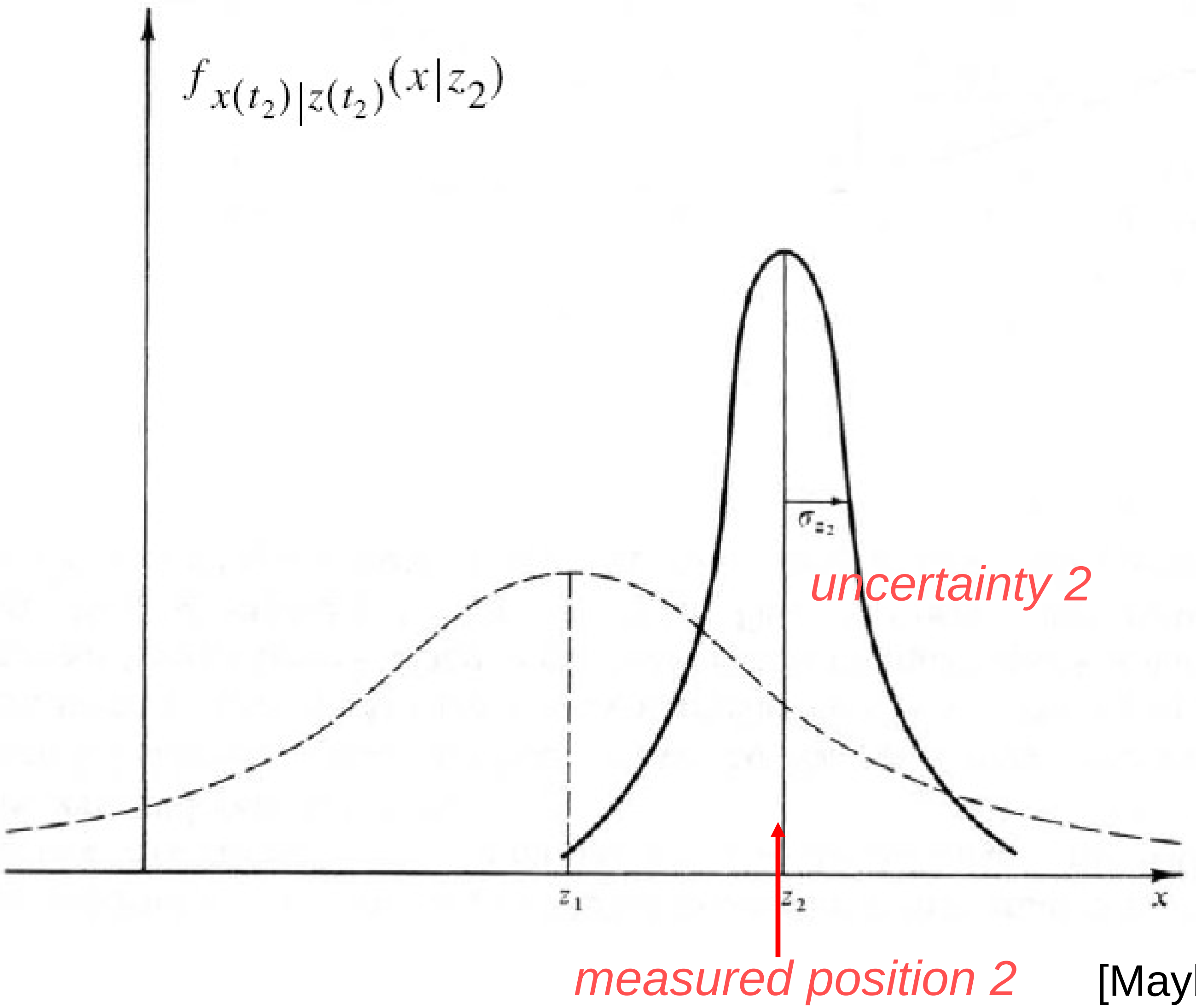
[Maybeck (1979)]

Conditional density of position based on measurement of z_2 alone



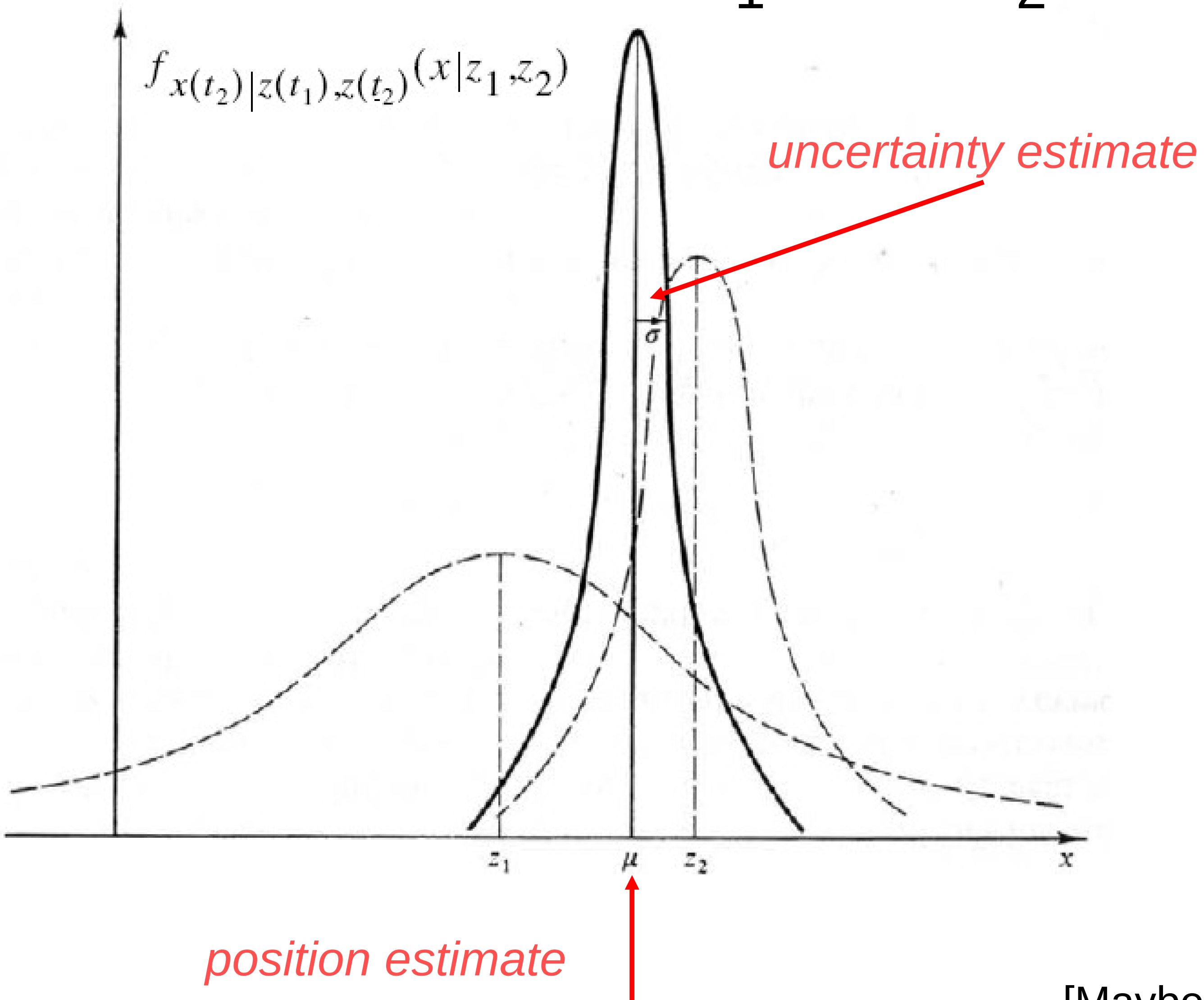
[Maybeck (1979)]

Conditional density of position based on measurement of z_2 alone



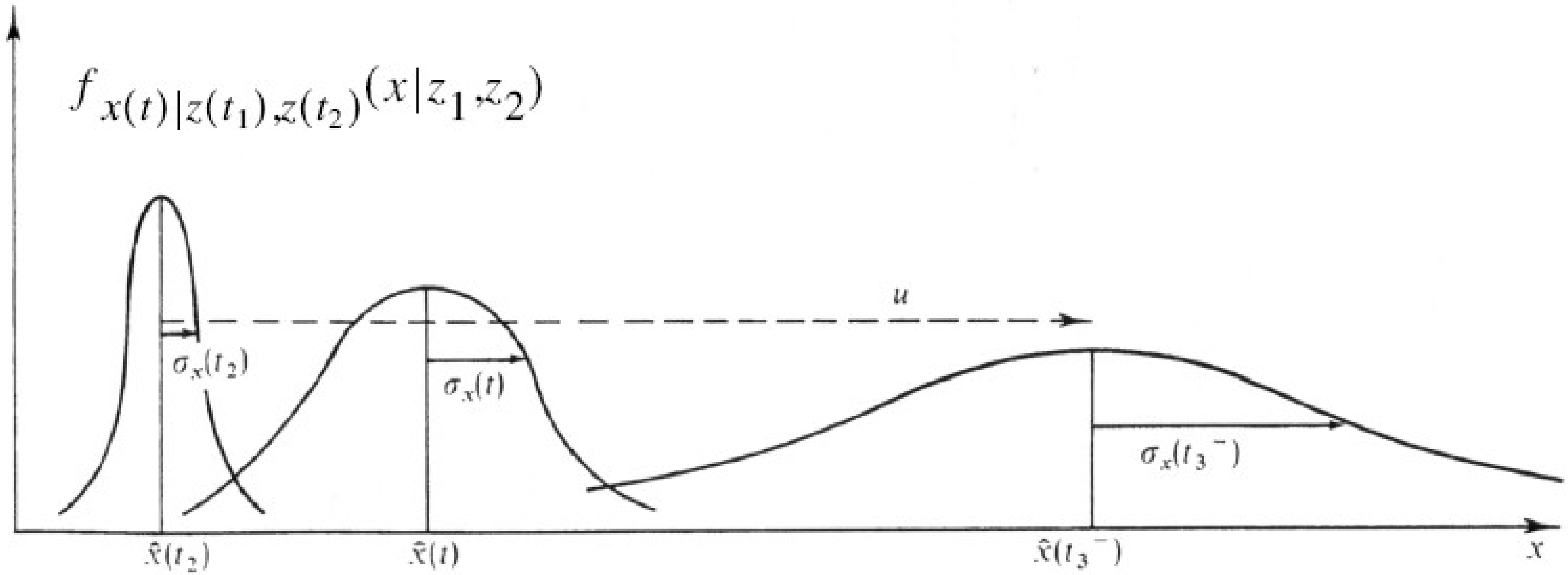
[Maybeck (1979)]

Conditional density of position based on data z_1 and z_2



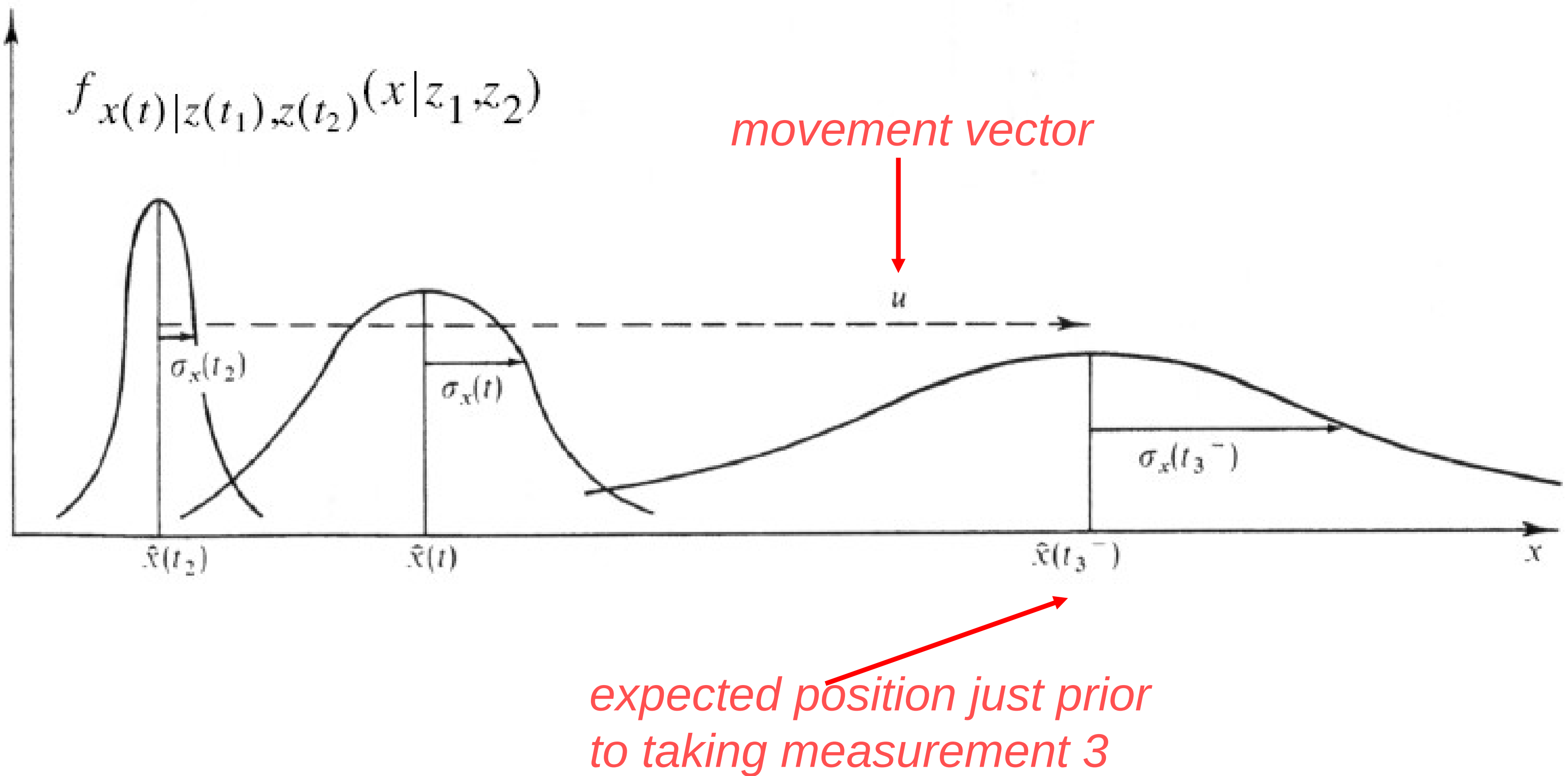
[Maybeck (1979)]

Propagation of the conditional density



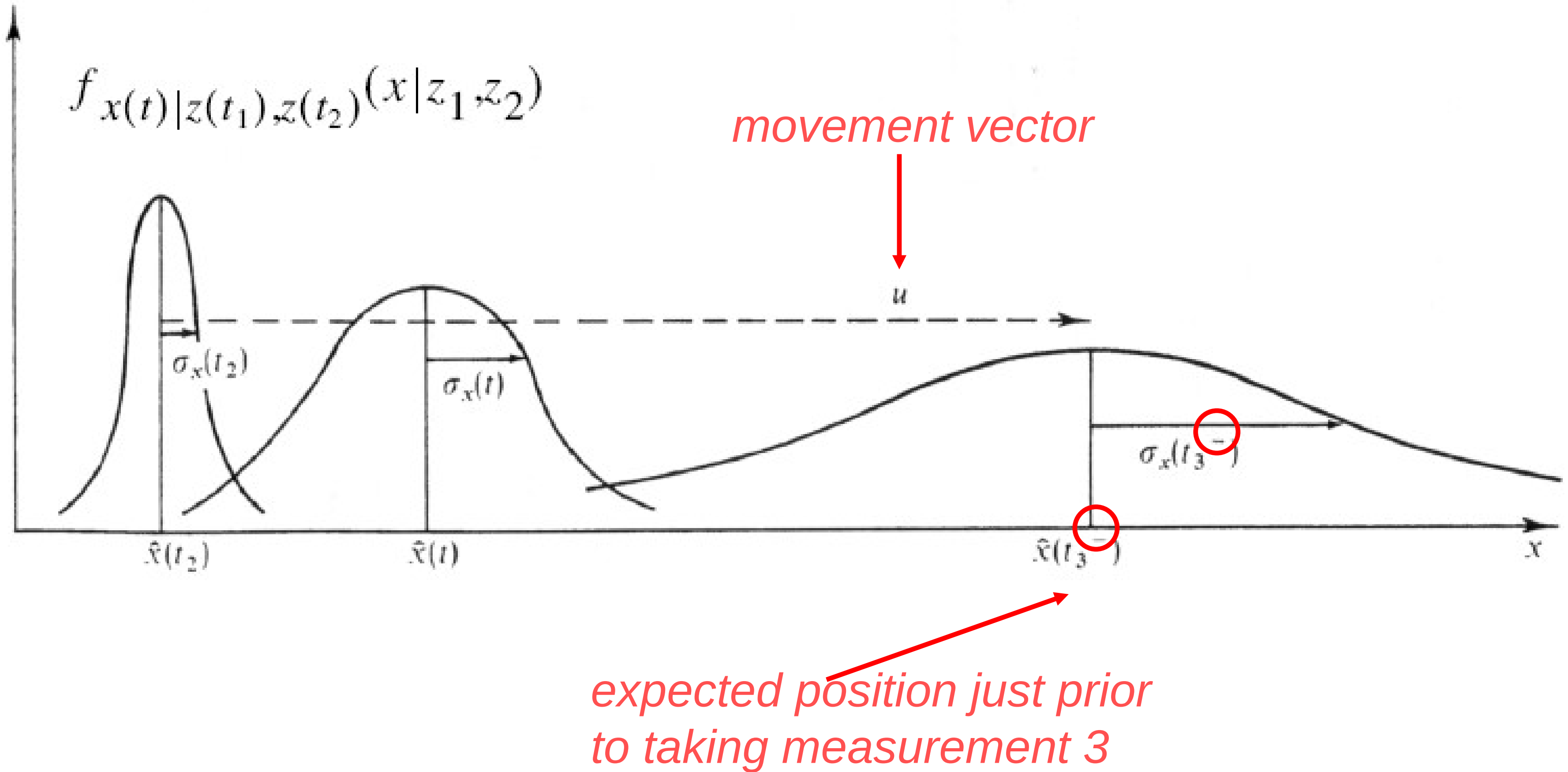
[Maybeck (1979)]

Propagation of the conditional density



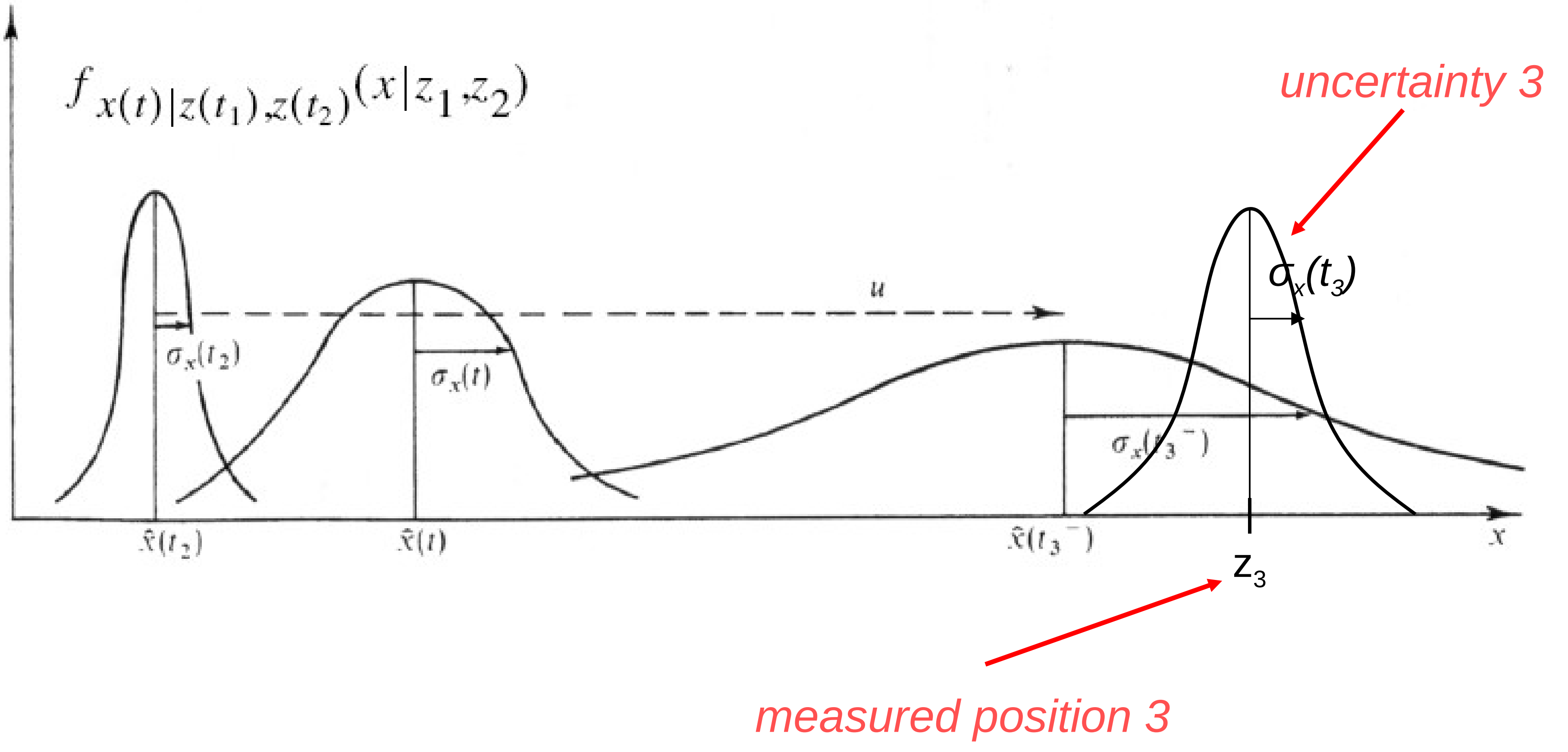
[Maybeck (1979)]

Propagation of the conditional density

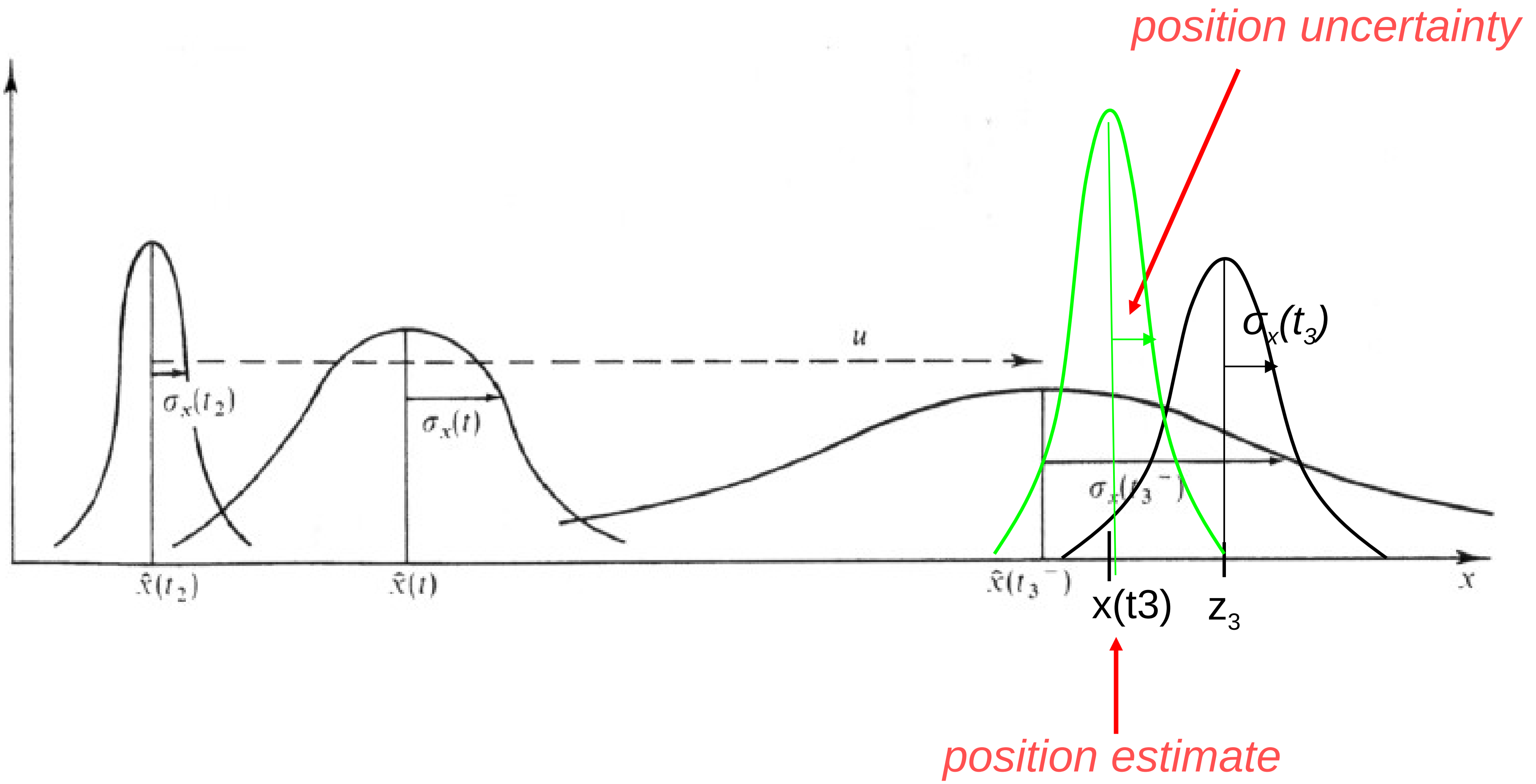


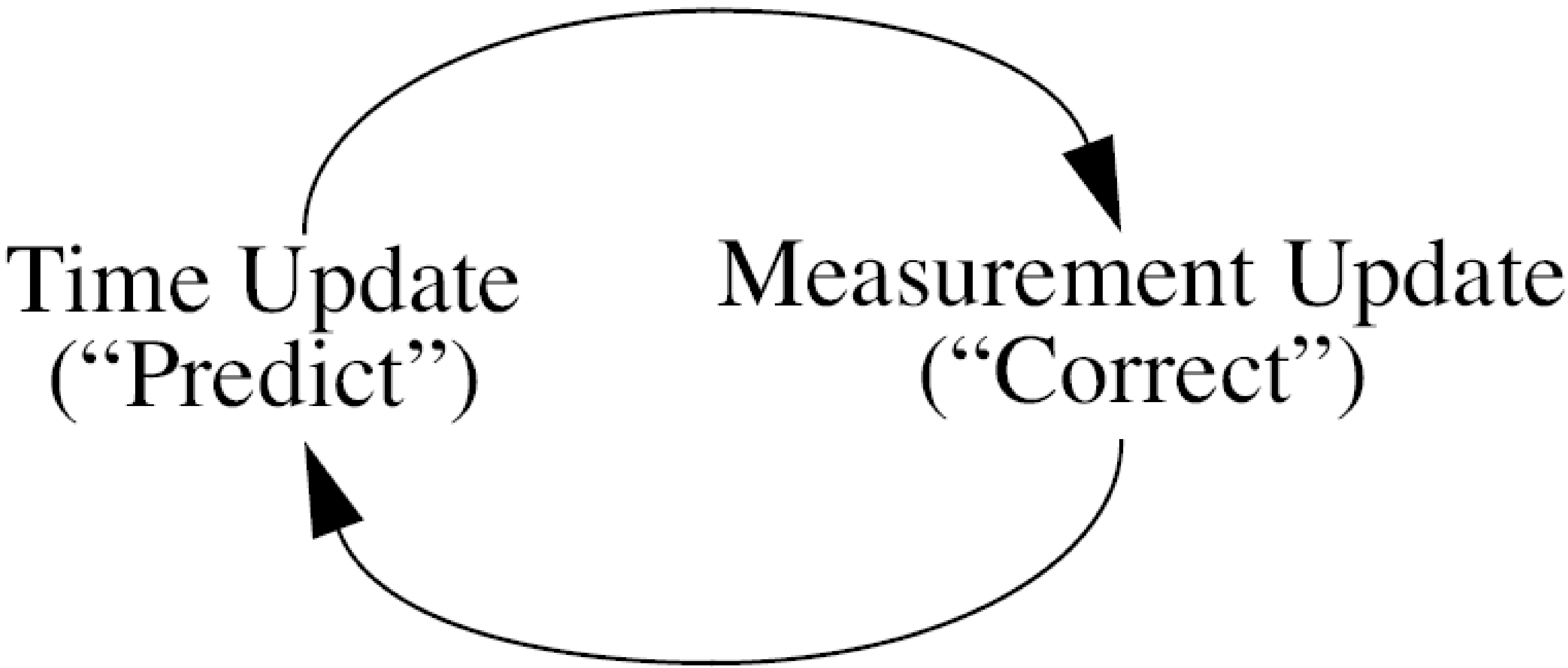
[Maybeck (1979)]

Propagation of the conditional density



Updating the conditional density after the third measurement

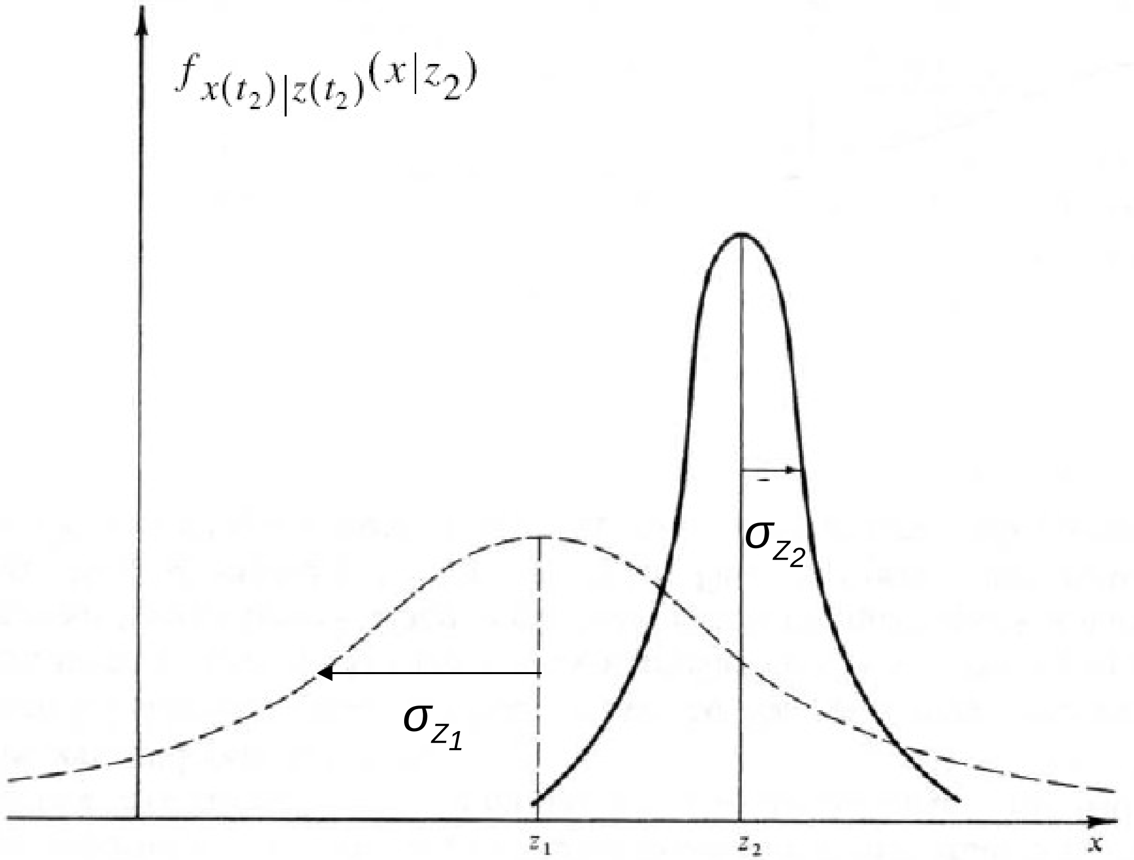



Time Update
("Predict")
Measurement Update
("Correct")

Questions?

Now let’s do the same thing
…but this time we’ll use math

How should we combine the two measurements?



[Maybeck (1979)]

Calculating the new mean

$$\mu = \boxed{\text{Scaling Factor 1}}\, z_1 + \boxed{\text{Scaling Factor 2}}\, z_2$$

Calculating the new mean

$$\mu = \boxed{\textit{Scaling Factor 1}}\, z_1 + \boxed{\textit{Scaling Factor 2}}\, z_2$$

$$\mu = [\sigma_{z_2}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)] z_1 + [\sigma_{z_1}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)] z_2$$

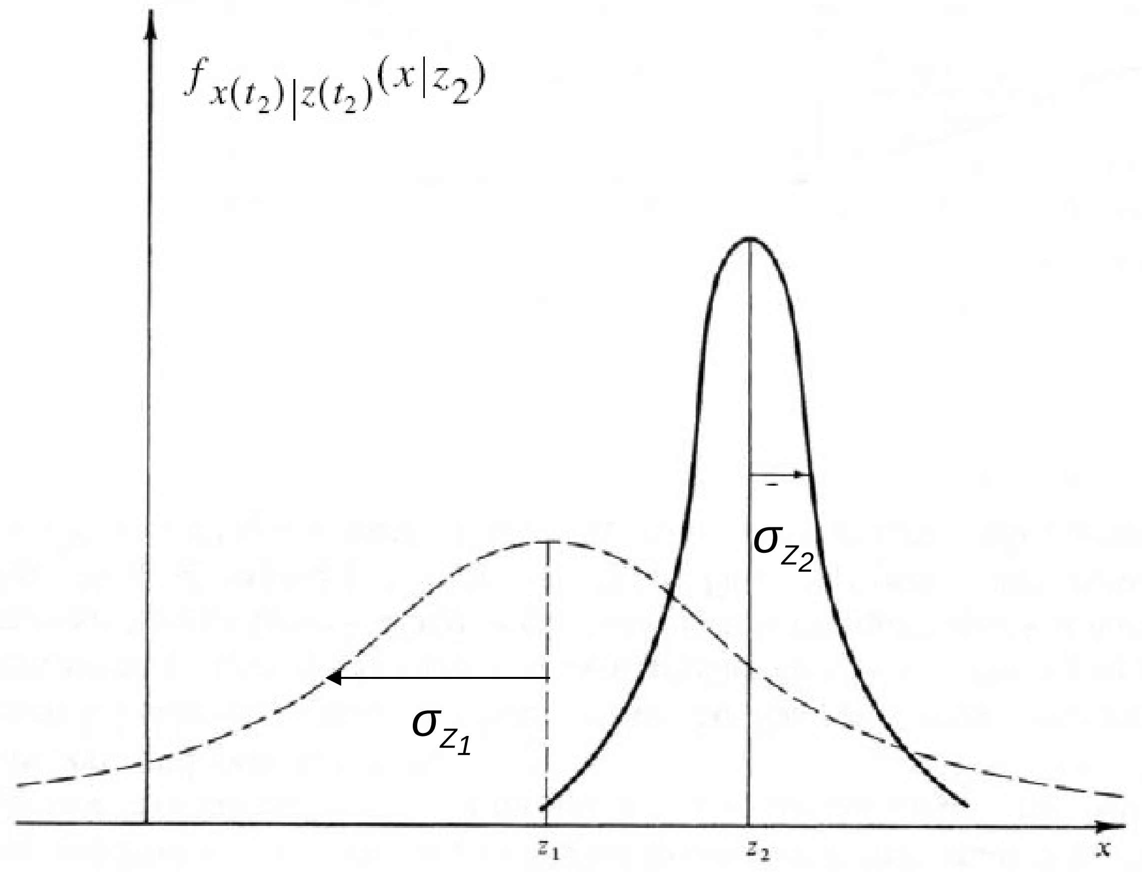
Calculating the new mean

$$\mu = \boxed{\textit{Scaling Factor 1}}\, z_1 + \boxed{\textit{Scaling Factor 2}}\, z_2$$

$$\mu = [\sigma_{z_2}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)] z_1 + [\sigma_{z_1}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)] z_2$$

Why is this not z_1?

Calculating the new variance



[Maybeck (1979)]

Calculating the new variance

$$\sigma^2 = \boxed{\textit{Scaling Factor 1}}\ \sigma_{z_1}^2 + \boxed{\textit{Scaling Factor 2}}\ \sigma_{z_2}^2$$

Remember the Gaussian Properties?

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2$$

Remember the Gaussian Properties?

- If $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$
- Then $Y \sim N(a\mu + b, a^2\sigma^2)$

This is a^2 not a

The scaling factors must be squared!

$$\sigma^2 = \boxed{\textit{Scaling Factor 1}}\ \sigma_{z_1}^2 \quad \boxed{\textit{Scaling Factor 2}}\ \sigma_{z_2}^2$$

$$\left[\sigma_{z_2}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)\right]^2 \qquad \left[\sigma_{z_1}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)\right]^2$$

The scaling factors must be squared!

$$\sigma^2 = \boxed{\textit{Scaling Factor 1}}\ \sigma_{z_1}^2 + \boxed{\textit{Scaling Factor 2}}\ \sigma_{z_2}^2$$

$$[\sigma_{z_2}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]^2 \qquad [\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]^2$$

$$\sigma^2 = [\sigma_{z_2}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]^2\sigma_{z_1}^2 + [\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]^2\sigma_{z_2}^2$$

Therefore the new variance is

$$1/\sigma^2 = (1/\sigma_{z_1}^2) + (1/\sigma_{z_2}^2)$$

Try to derive this on your own.

Another Way to Express The New Position

$$\hat{x}(t_2) = [\sigma_{z_2}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_2$$

$$= z_1 - z_1 + [\sigma_{z_2}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_2$$

$$= z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)][z_2 - z_1]$$

[Maybeck (1979)]

Another Way to Express The New Position

$$\hat{x}(t_2) = [\sigma_{z_2}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)] z_1 + [\sigma_{z_1}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)] z_2$$

$$= z_1 + [\sigma_{z_1}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)][z_2 - z_1]$$

$$\hat{x}(t_2) = \hat{x}(t_1) \quad + \quad K(t_2) \quad [z_2 - \hat{x}(t_1)]$$

[Maybeck (1979)]

Another Way to Express The New Position

$$\hat{x}(t_2) = \hat{x}(t_1) + K(t_2)[z_2 - \hat{x}(t_1)]$$

$$K(t_2) = \sigma_{z_1}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)$$

[Maybeck (1979)]

The equation for the variance can also be rewritten as

$$\sigma_x^2(t_2) = \sigma_x^2(t_1) - K(t_2)\sigma_x^2(t_1)$$

[Maybeck (1979)]

Adding Movement

$$dx/dt = u + w$$

[Maybeck (1979)]

Adding Movement

$$\hat{x}(t_3^-) = \hat{x}(t_2) + u[t_3 - t_2]$$

$$\sigma_x^2(t_3^-) = \sigma_x^2(t_2) + \sigma_w^2[t_3 - t_2]$$

[Maybeck (1979)]

Adding Movement

$$\hat{x}(t_3) = \hat{x}(t_3^-) + K(t_3)[z_3 - \hat{x}(t_3^-)]$$

$$\sigma_x^2(t_3) = \sigma_x^2(t_3^-) - K(t_3)\sigma_x^2(t_3^-)$$

$$K(t_3) = \sigma_x^2(t_3^-)/[\sigma_x^2(t_3^-) + \sigma_{z_3}^2]$$

[Maybeck (1979)]

Properties of K

- If the measurement noise is large K is small

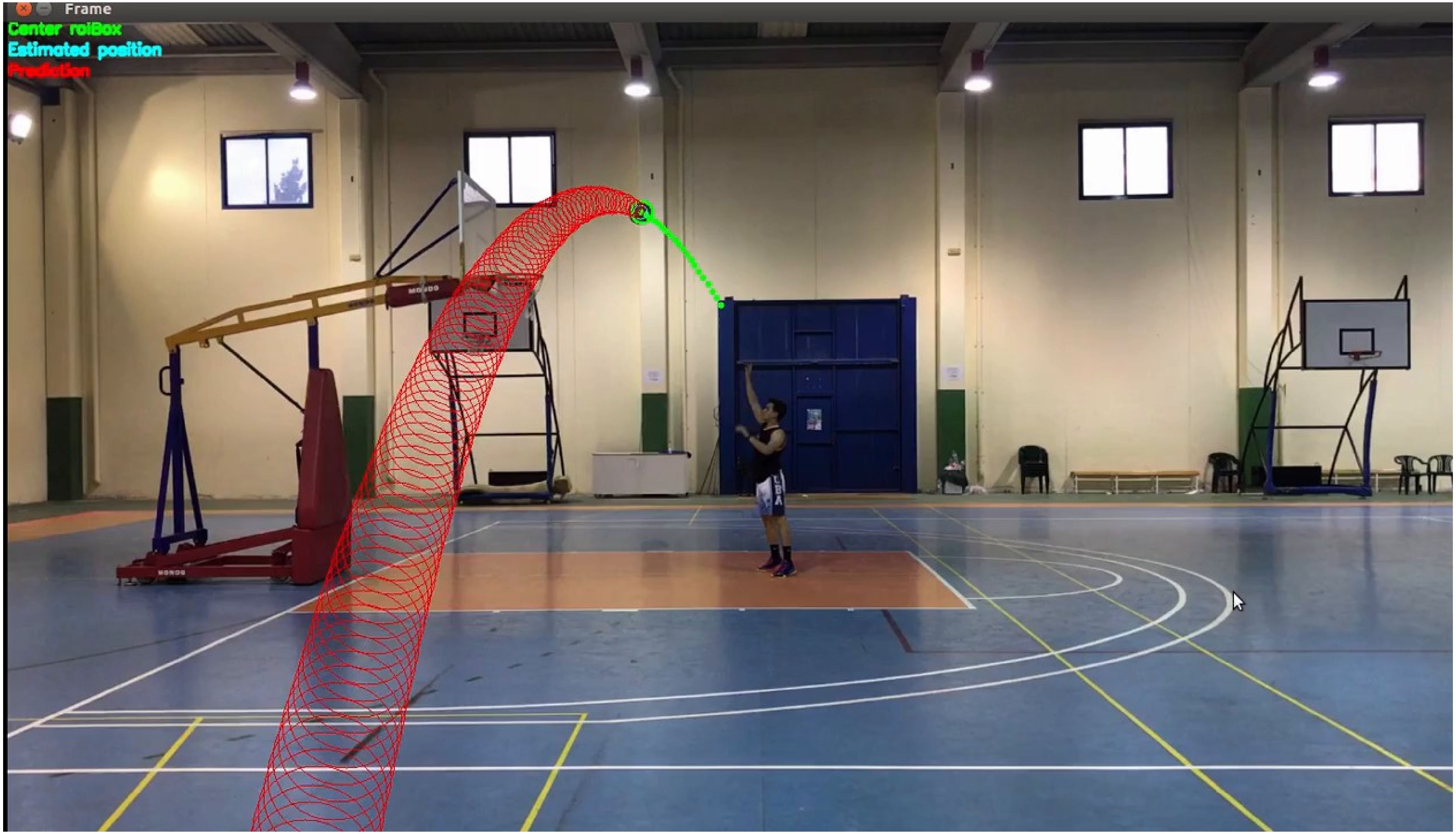
$$K(t_3) = \sigma_x^2(t_3^-)/[\sigma_x^2(t_3^-) + \sigma_{z_3}^2]$$

$$\sigma_{z_3}^2 \rightarrow \infty \ , \ K(t_3) \rightarrow 0$$

[Maybeck (1979)]

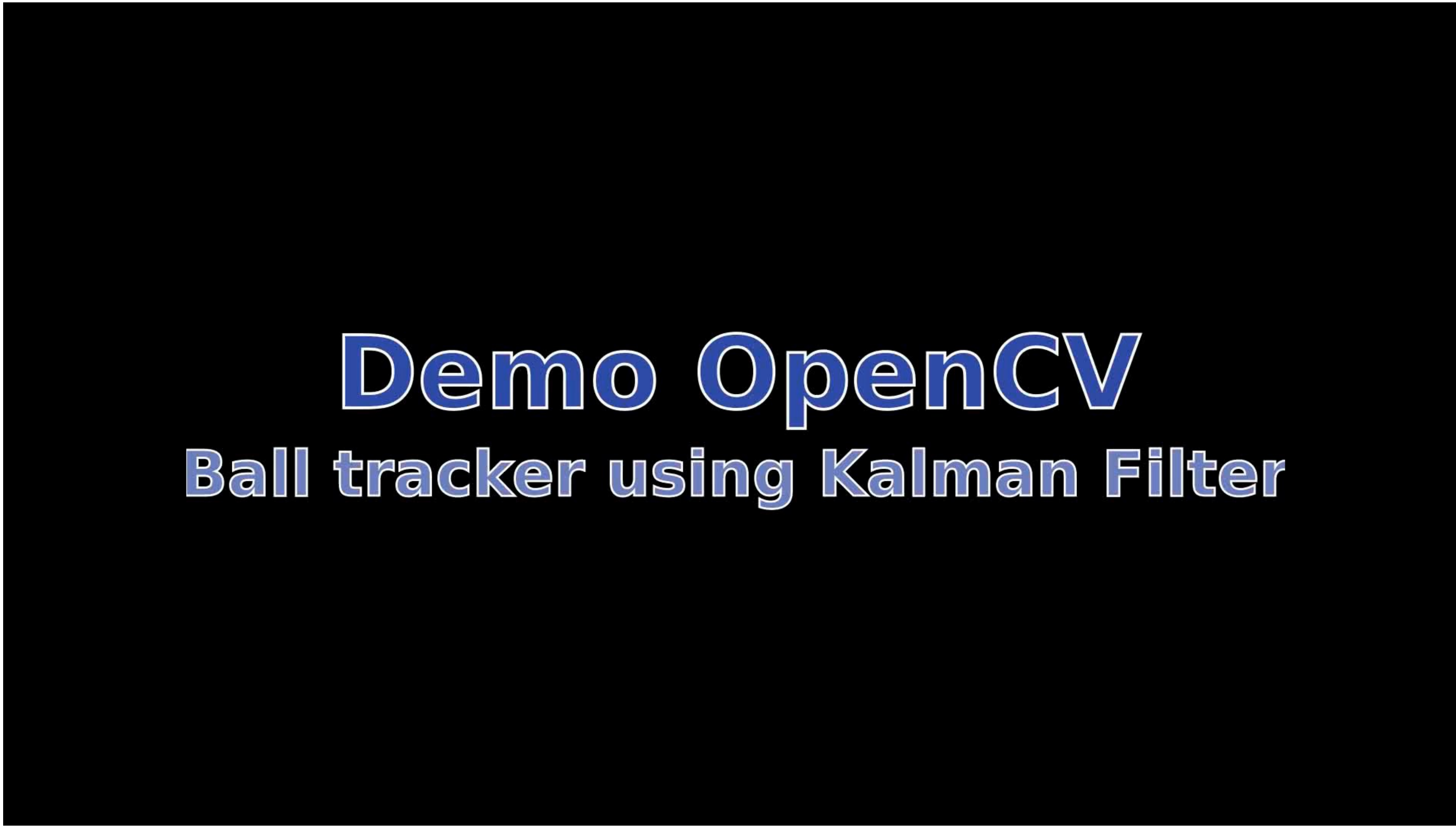
The Kalman Filter (part 2)

Example Applications



https://www.youtube.com/watch?v=MxwVwCuBEDA

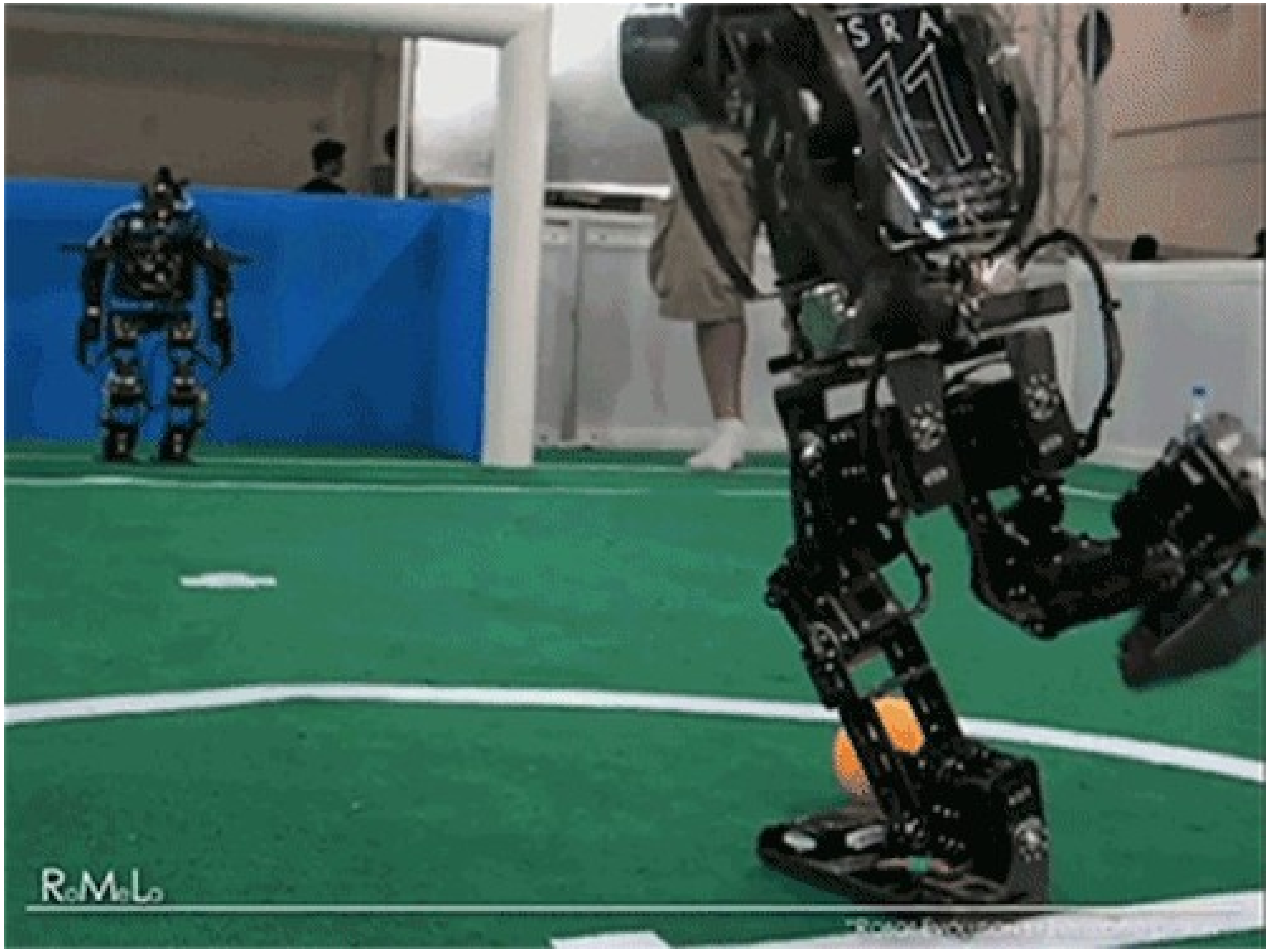
https://github.com/pabsaura/Prediction-of-Trajectory-with-kalman-filter-and-open-cv



https://www.youtube.com/watch?v=sG-h5ONsj9s

https://www.myzhar.com/blog/tutorials/tutorial-opencv-ball-tracker-using-kalman-filter/

Something fun

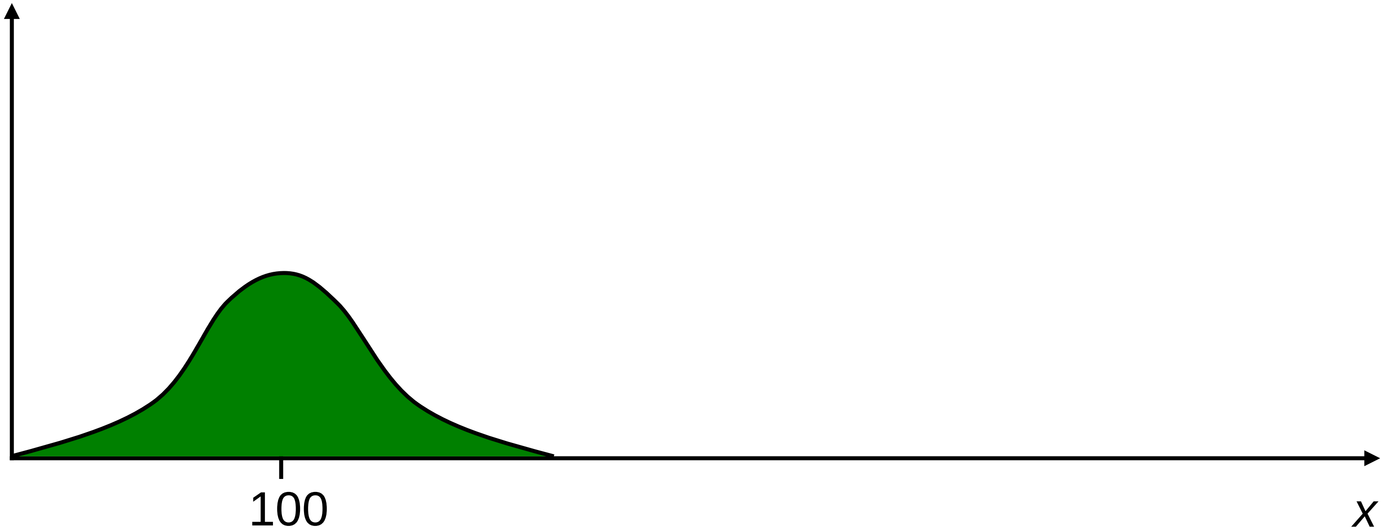


Another Example

[www.cse.lehigh.edu/~spletzer/cse398_Spring05/lec011_Localization2.ppt]

A Simple Example

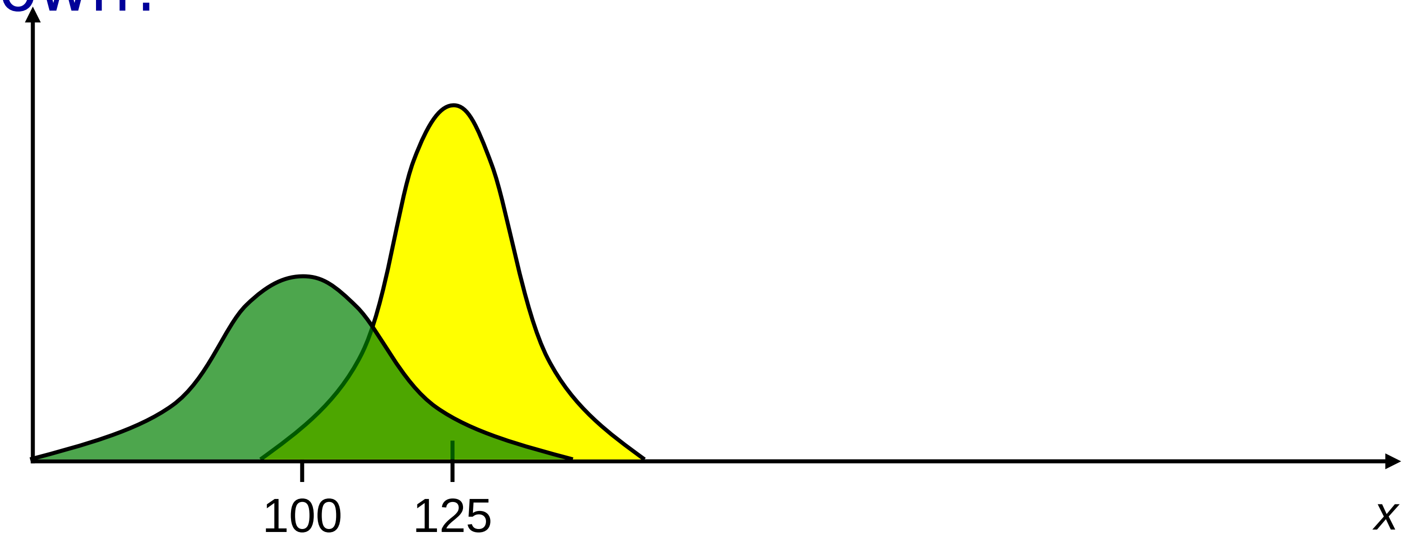
- Consider a ship sailing east with a perfect compass trying to estimate its position.
- You estimate the position x from the stars as $z_1=100$ with a precision of $\sigma_x=4$ miles



[www.cse.lehigh.edu/~spletzer/cse398_Spring05/lec011_Localization2.ppt]

A Simple Example (cont'd)

- Along comes a more experienced navigator, and she takes her own sighting z_2
- She estimates the position $x = z_2 = 125$ with a precision of σ_x=3 miles
- How do you merge her estimate with your own?



[www.cse.lehigh.edu/~spletzer/cse398_Spring05/lec011_Localization2.ppt]

A Simple Example (cont'd)

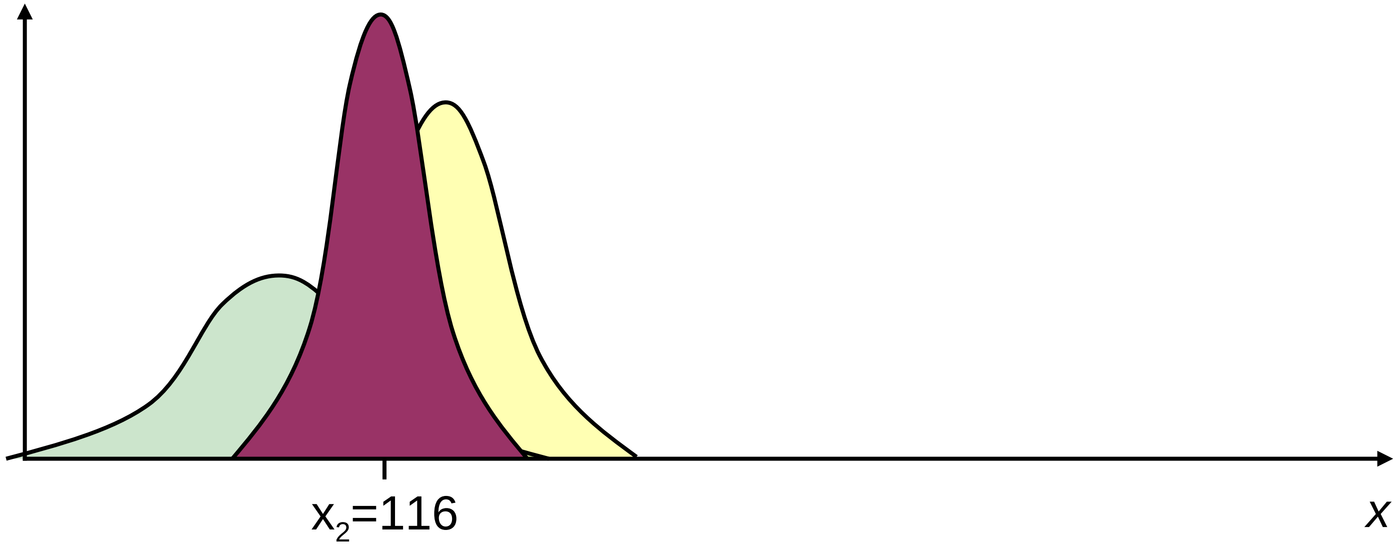
$$\mu = \left[\frac{\sigma_{z_2}^2}{\sigma_{z1}^2 + \sigma_{z_2}^2}\right] z_1 + \left[\frac{\sigma_{z_1}^2}{\sigma_{z1}^2 + \sigma_{z_2}^2}\right] z_2$$

$$= \left[\frac{9}{16+9}\right] 100 + \left[\frac{16}{16+9}\right] 125 = 116$$

$$\frac{1}{\sigma^2} = \frac{1}{\sigma_{z_1}^2} + \frac{1}{\sigma_{z_2}^2}$$

$$\frac{1}{\sigma^2} = \frac{1}{9} + \frac{1}{16} = \frac{25}{144}$$

$$\Rightarrow \sigma = 2.4$$



[www.cse.lehigh.edu/~spletzer/cse398_Spring05/lec011_Localization2.ppt]

A Simple Example (cont'd)

- With the distributions being Gaussian, the best estimate for the state is the mean of the distribution, so…

$$x_2 = \left[\frac{\sigma_{z_2}^2}{\sigma_{z_1}^2+\sigma_{z_2}^2}\right] z_1 + \left[\frac{\sigma_{z_1}^2}{\sigma_{z_1}^2+\sigma_{z_2}^2}\right] z_2$$

or alternately

$$= z_1 + \left[\frac{\sigma_{z_1}^2}{\sigma_{z_1}^2+\sigma_{z_2}^2}\right]\left(z_2 - z_1\right)$$

$$= z_1 + K_2\left(z_2 - z_1\right)$$

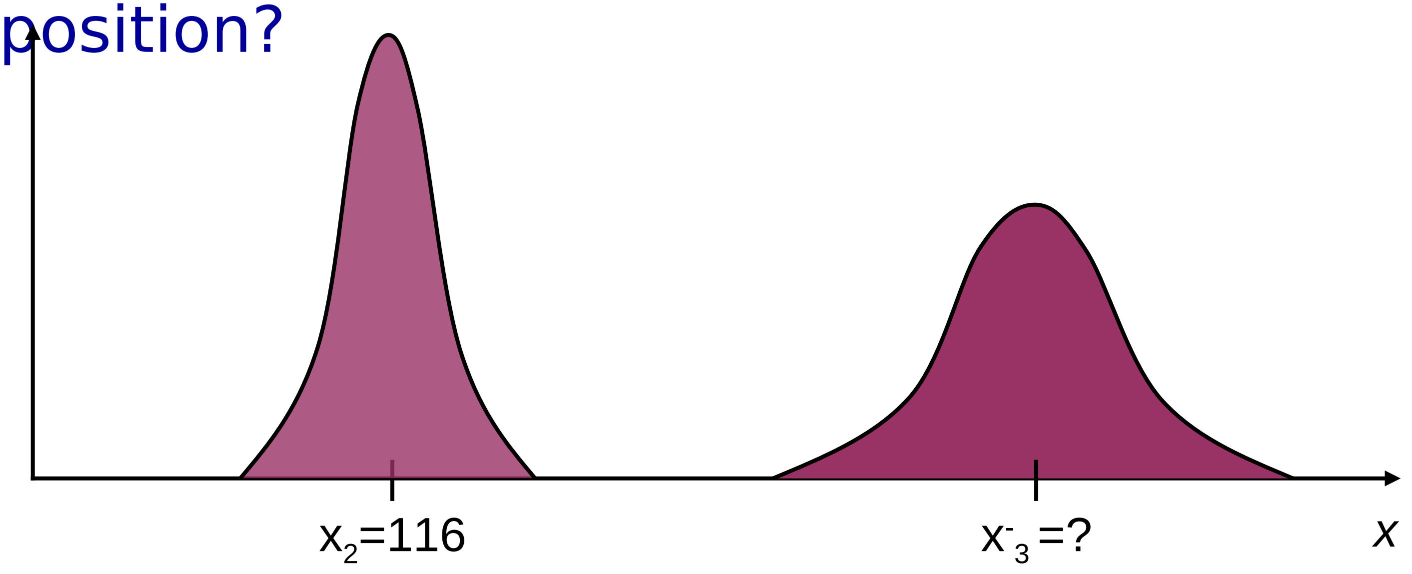
Correction Term

where K_t is referred to as the *Kalman gain*, and must be computed at each time step

[www.cse.lehigh.edu/~spletzer/cse398_Spring05/lec011_Localization2.ppt]

A Simple Example (cont'd)

- OK, now you fall asleep on your watch. You wake up after 2 hours, and you now have to re-estimate your position
- Let the velocity of the boat be nominally 20 miles/hour, but with a variance of $\sigma^2_w=4$ miles²/hour
- What is the best estimate of your current position?



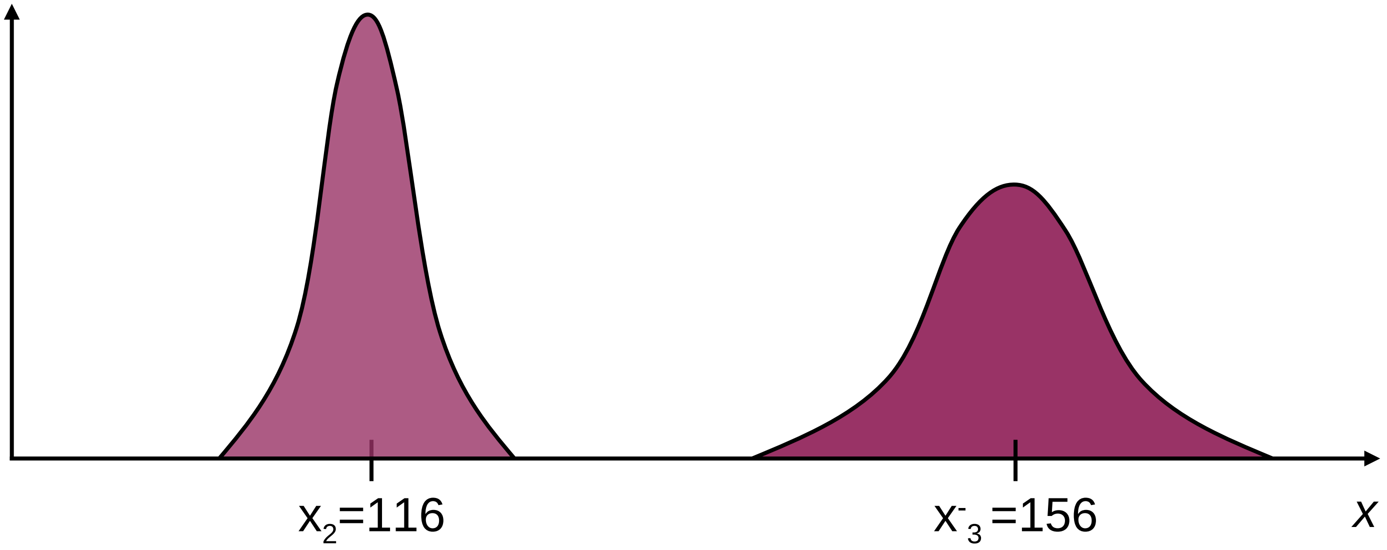
[www.cse.lehigh.edu/~spletzer/cse398_Spring05/lec011_Localization2.ppt]

A Simple Example (cont'd)

- The next effect is that the gaussian is translated by a distance and the variance of the distribution is increased to account for the uncertainty in dynamics

$$x_3^- = x_2 + v\Delta t \qquad \Rightarrow x_3^- = 116 + 40 = 156$$

$$\sigma_3^{2-} = \sigma_2^2 + \sigma_w^{2-}\,\Delta t \qquad \Rightarrow \sigma_3^{2-} = 5.76 + 8 = 13.76$$



[www.cse.lehigh.edu/~spletzer/cse398_Spring05/lec011_Localization2.ppt]

A Simple Example (cont'd)

- OK, this is not a very accurate estimate. So, since you've had your nap you decide to take another measurement and you get $z_3=165$ miles
- Using the same update procedure as the first update, we obtain

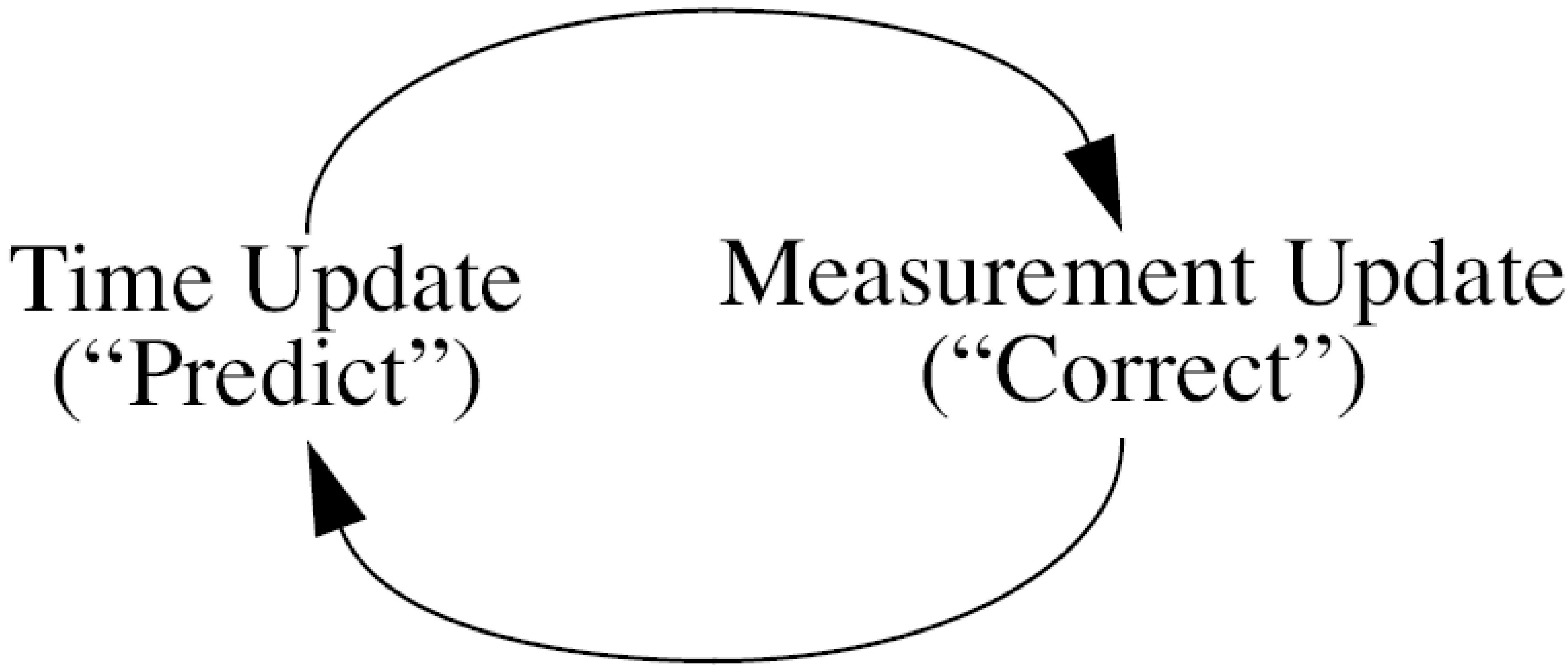
$$x_3 = x_3^- + K_3\left(z_3 - x_3^-\right)$$

$$\sigma_3^2 = \sigma_3^{2-} - K_3\sigma_3^{2-}$$
$$= 13.76 - \left[\frac{13.76}{13.76+16}\right]13.76 = 7.40$$

and so on…

The Predictor-Corrector Approach

- In this example, prediction came from using knowledge of the vehicle dynamics to estimate its change in position
- An analogy with a robot would be integrating information from the robot kinematics (*i.e.* you give it a desired [x, y, α] velocities for a time Δt) to estimate changed in position
- The correction is accomplished through making exteroceptive observations and then fusing this with your current estimate
- This is akin to updating position estimates using landmark information, etc.
- In practice, the prediction rate is typically much higher than the correction


Time Update
("Predict")
Measurement Update
("Correct")

Calculating the new mean

$$\mu = [\textit{Scaling Factor 1}]z_1 + [\textit{Scaling Factor 2}]z_2$$

$$\mu = [\sigma_{z_2}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_1 + [\sigma_{z_1}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_2$$

Calculating the new variance

$$\sigma^2 = \boxed{\textit{Scaling Factor 1}}\ \sigma^2_{z_1} + \boxed{\textit{Scaling Factor 2}}\ \sigma^2_{z_2}$$

$$\left[\sigma^2_{z_2} / (\sigma^2_{z_1} + \sigma^2_{z_2})\right]^2 \qquad \left[\sigma^2_{z_1} / (\sigma^2_{z_1} + \sigma^2_{z_2})\right]^2$$

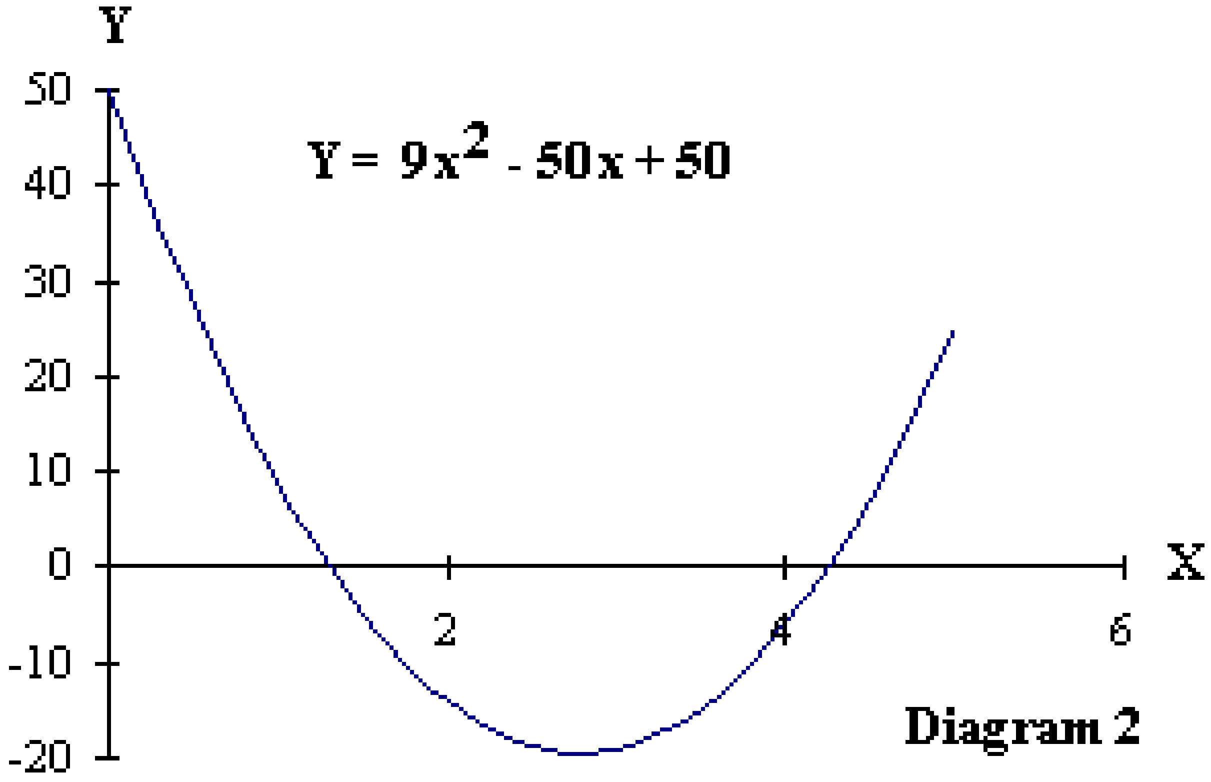
$$\sigma^2 = \left[\sigma^2_{z_2} / (\sigma^2_{z_1} + \sigma^2_{z_2})\right]^2 \sigma^2_{z_1} + \left[\sigma^2_{z_1} / (\sigma^2_{z_1} + \sigma^2_{z_2})\right]^2 \sigma^2_{z_2}$$

What makes these scaling factors special? Are there other ways to combine the two measurements?

- They minimize the error between the prediction and the true value of X.
- They are optimal in the least-squares sense.

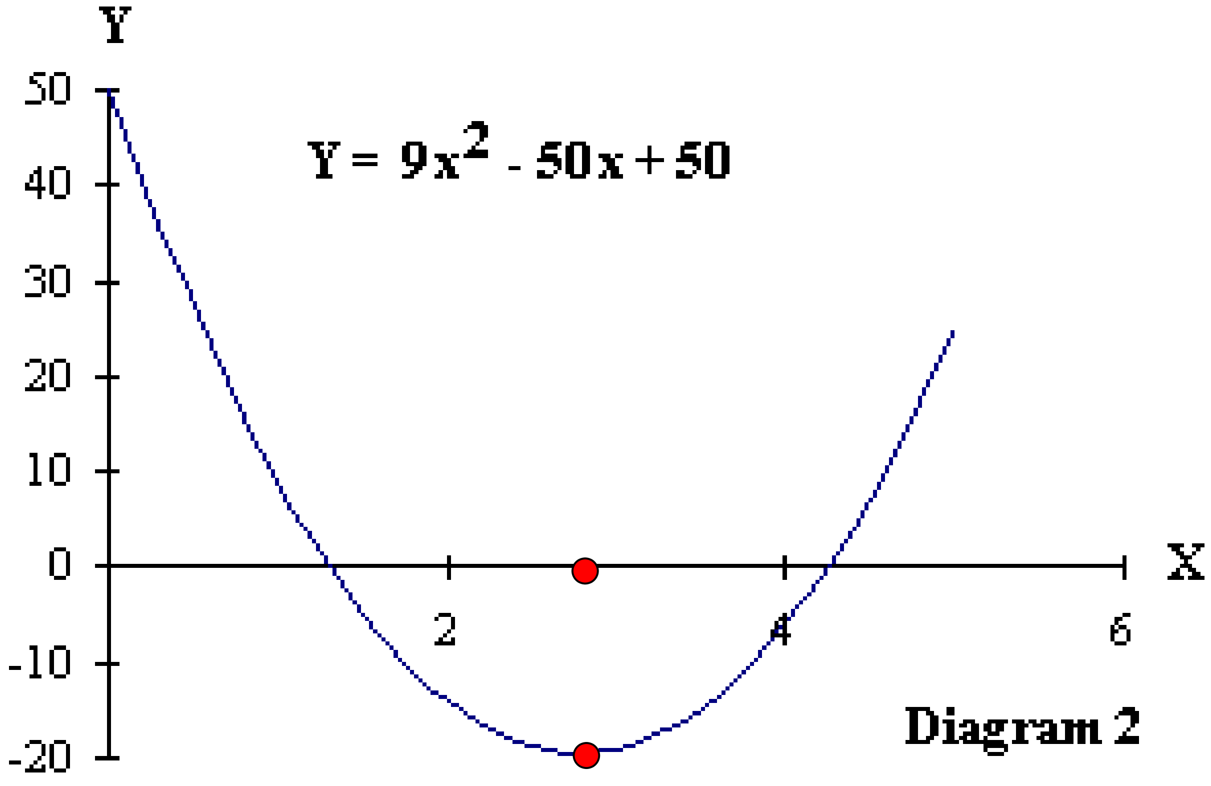
Minimize the error

What is the minimum value?



[http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/function.gif]

What is the minimum value?



[http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/function.gif]

Finding the Minimum Value

- $Y = 9x^2 - 50x + 50$
- $dY/dX = 18x – 50$
- $0 = 18x – 50$
- $X = 50/18 = 2.68….$
- Min Y =

Finding the Minimum Value

- $Y = 9x^2 - 50x + 50$

- $dY/dx = 18x - 50 = 0$

- The minimum is obtained when $x = 50/18 = 2.77777(7)$

- The minimum value is

$$Y(x_{min}) = 9*(50/18)^2 - 50*(50/18) + 50 = -19.44444(4)$$

Start with two measurements

$$z_1 = x + v_1 \text{ and } z_2 = x + v_2$$

- v_1 and v_2 represent zero mean noise

Formula for the estimation error

- The new estimate is

$$\hat{x} = s_1 z_1 + s_2 z_2$$

- The error is

$$e = \hat{x} - x$$

Expected value of the error

$$
\begin{aligned}
E[e] &= E[\hat{x} - x] \\
&= E[s_1 z_1 + s_2 z_2 - x]
\end{aligned}
$$

Expected value of the error

$$
\begin{aligned}
E[e] &= E[\hat{x} - x] \\
&= E[s_1 z_1 + s_2 z_2 - x] \\
&= E[s_1(x + v_1) + s_2(x + v_2) - x]
\end{aligned}
$$

Expected value of the error

$$\begin{aligned}
E[e] &= E[\hat{x} - x] \\
&= E[s_1 z_1 + s_2 z_2 - x] \\
&= E[s_1(x + v_1) + s_2(x + v_2) - x] \\
&= s_1 E[x] + s_1 E[v_1] + s_2 E[x] + s_2 E[v_2] - E[x]
\end{aligned}$$

Expected value of the error

$$
\begin{aligned}
E[e] &= E[\hat{x} - x] \\
&= E[s_1 z_1 + s_2 z_2 - x] \\
&= E[s_1(x + v_1) + s_2(x + v_2) - x] \\
&= s_1 E[x] + s_1 E[v_1] + s_2 E[x] + s_2 E[v_2] - E[x] \\
&= s_1 E[x] + 0 + s_2 E[x] + 0 - E[x]
\end{aligned}
$$

Expected value of the error

$$
\begin{aligned}
E[e] &= E[\hat{x} - x] \\
&= E[s_1 z_1 + s_2 z_2 - x] \\
&= E[s_1(x + v_1) + s_2(x + v_2) - x] \\
&= s_1 E[x] + s_1 E[v_1] + s_2 E[x] + s_2 E[v_2] - E[x] \\
&= s_1 E[x] + 0 + s_2 E[x] + 0 - E[x] \\
&= s_1 x + s_2 x - x = 0
\end{aligned}
$$

- If the estimate is unbiased this should hold

Therefore, $s_1 + s_2 - 1 = 0$

which can be rewritten as $s_2 = 1 - s_1$

Find the Mean Square Error

$$E[e^2] = E[(\hat{x} - x)^2]$$

$$= \quad ?$$

$$
\begin{aligned}
E[e^2] &= E[(\hat{x} - x)^2] \\
&= E[\hat{x}^2 - 2\hat{x}x + x^2] \\
&= E[(s_1 z_1 + s_2 z_2)^2 - 2(s_1 z_1 + s_2 z_2)x + x^2] \\
&= E[(s_1(x + v_1) + s_2(x + v_2))^2 - 2(s_1(x + v_1) + s_2(x + v_2))x + x^2] \\
&= E[s_1^2(x + v_1)^2 + 2s_1 s_2(x + v_1)(x + v_2) + s_2^2(x + v_2)^2 - 2s_1(x + v_1)x - 2s_2(x + v_2) \\
&= E[\underline{\underline{s_1^2 x^2}} + \underline{2s_1^2 v_1 x} + s_1^2 v_1^2 + \underline{\underline{2s_1 s_2 x^2}} + \underline{2s_1 s_2 v_1 x} + \underline{2s_1 s_2 v_2 x} + 2s_1 s_2 v_1 v_2 + \\
&\quad + \underline{\underline{s_2^2 x^2}} + \underline{2s_2^2 v_2 x} + s_2^2 v_2^2 - \underline{\underline{2s_1 x^2}} - \underline{2s_1 v_1 x} - \underline{\underline{2s_2 x^2}} - \underline{2s_2 v_2 x} + \underline{\underline{x^2}}] \\
&= E[(s_1^2 + 2s_1 s_2 + s_2^2 - 2s_1 - 2s_2 + 1)x^2 + \\
&\quad +2(s_1^2 v_1 + s_1 s_2 v_1 + s_1 s_2 v_2 + s_2^2 v_2 - s_1 v_1 - s_2 v_2)x + \\
&\quad +s_1^2 v_1^2 + 2s_1 s_2 v_1 v_2 + s_2^2 v_2^2] \\
&= \{(s_1 + s_2)^2 - 2(s_1 + s_2) + 1\} E[x^2] + \\
&\quad +2\{s_1^2 E[v_1] + s_1 s_2 E[v_1] + s_1 s_2 E[v_2] + s_2^2 E[v_2] - s_1 E[v_1] - s_2 E[v_2]\} E[x] + \\
&\quad +s_1^2 E[v_1^2] + 2s_1 s_2 E[v_1]E[v_2] + s_2^2 E[v_2^2] \\
&= (1 - 2 + 1)E[x^2] + 2(0 + 0 + 0 + 0 - 0 - 0)E[x] + s_1^2 E[v_1^2] + 0 + s_2^2 E[v_2^2] \\
&= s_1^2 E[v_1^2] + s_2^2 E[v_2^2] \\
&= s_1^2 \sigma_1^2 + s_2^2 \sigma_2^2 \\
&= s_1^2 \sigma_1^2 + (1 - s_1)^2 \sigma_2^2
\end{aligned}
$$

Mean Square Error

$$E[e^2] \quad = \quad s_1^2\sigma_1^2 + (1 - s_1)^2\sigma_2^2$$

Minimize the mean square error

$$E[e^2] = s_1^2\sigma_1^2 + (1 - s_1)^2\sigma_2^2$$

$$\begin{aligned}
\frac{\mathrm{dE[e^2]}}{\mathrm{ds_1}} &= 2s_1\sigma_1^2 - 2(1 - s_1)\sigma_2^2 \\
&= 2s_1\sigma_1^2 + 2s_1\sigma_2^2 - 2\sigma_2^2 \\
&= 2s_1(\sigma_1^2 + \sigma_2^2) - 2\sigma_2^2 = 0
\end{aligned}$$

Finding S_1

$$2s_1(\sigma_1^2 + \sigma_2^2) - 2\sigma_2^2 = 0$$

$$2s_1(\sigma_1^2 + \sigma_2^2) = 2\sigma_2^2$$

- Therefore

$$s_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

Finding S_2

$$
\begin{aligned}
s_2 &= 1 - s_1 \\
&= 1 - \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2} \\
&= \frac{\sigma_1^2+\sigma_2^2-\sigma_2^2}{\sigma_1^2+\sigma_2^2} \\
&= \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}
\end{aligned}
$$

Finally we get what we wanted

$$\hat{x} = s_1 z_1 + s_2 z_2$$
$$= \left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right) z_1 + \left(\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right) z_2$$

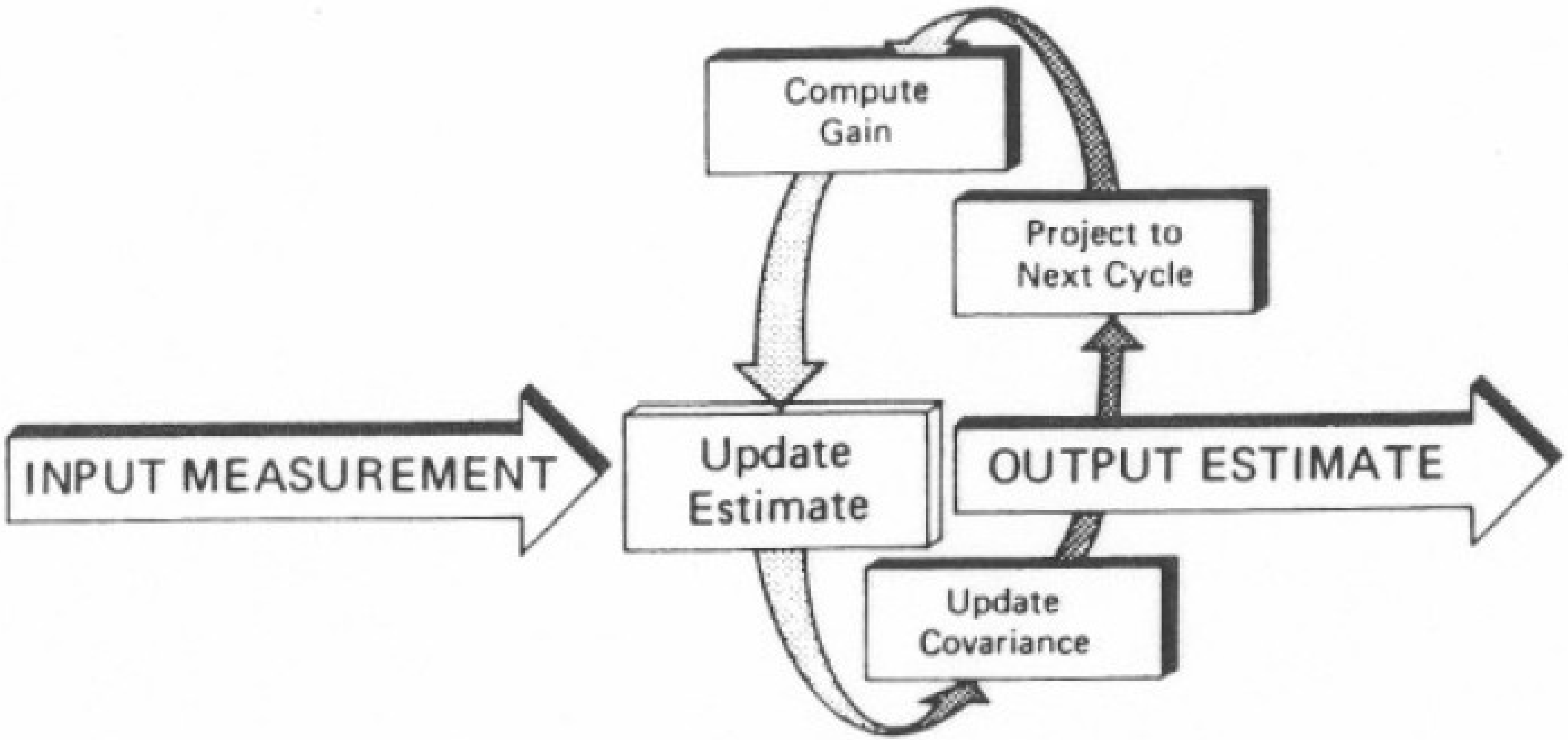
Finding the new variance

$$
\begin{aligned}
\sigma^2 &= s_1^2\sigma_1^2 + s_2^1\sigma_2^2 \\
&= \left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^2 \sigma_1^2 + \left(\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2 \sigma_2^2 \\
&= \frac{\sigma_2^4\sigma_1^2 + \sigma_1^4\sigma_2^2}{\left(\sigma_1^2+\sigma_2^2\right)^2} \\
&= \frac{\sigma_1^2\sigma_2^2\left(\sigma_1^2+\sigma_2^2\right)}{\left(\sigma_1^2+\sigma_2^2\right)^2} \\
&= \frac{\sigma_1^2\sigma_2^2}{\left(\sigma_1^2+\sigma_2^2\right)} \\
&= \frac{1}{\left(\frac{\sigma_1^2+\sigma_2^2}{\sigma_2^2\sigma_1^2}\right)} \\
&= \frac{1}{\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}}
\end{aligned}
$$

Formula for the new variance

$$\boxed{\frac{1}{\sigma^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$$

Kalman Filter Diagram



[Brown and Hwang (1992)]

Overview of Homework 1

THE END