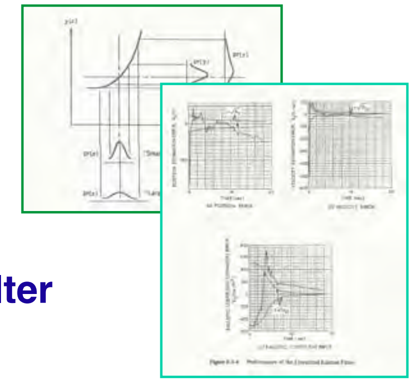
Nonlinear State Estimation

Extended Kalman Filters

Robert Stengel
Optimal Control and Estimation, MAE 546
Princeton University, 2015

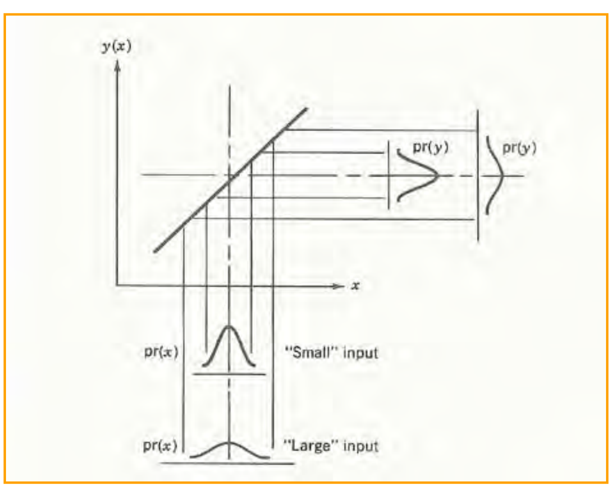
- Deformation of the probability distribution
- Neighboring-optimal estimator
- Extended Kalman-Bucy filter
- Hybrid extended Kalman filter
- Quasilinear extended Kalman filter



Linear Transformation of a Probability Distribution

- Linear transformation does not change the shape of a probability distribution

$$\boxed{y = kx}$$



$$E(y) = \int_{-\infty}^{\infty} y \,\mathrm{pr}(x)\,dx = \bar{y}$$
$$= k\int_{-\infty}^{\infty} x \,\mathrm{pr}(x)\,dx = k\,\bar{x}$$

$$E\left[(y-\bar{y})^2\right] = \sigma_y^2 = \int_{-\infty}^{\infty} (y-\bar{y})^2 \,\mathrm{pr}(x)\,dx$$
$$= k^2\int_{-\infty}^{\infty} (x-\bar{x})^2 \,\mathrm{pr}(x)\,dx = k^2\sigma_x^2$$
$$\sigma_y = k\sigma_x$$

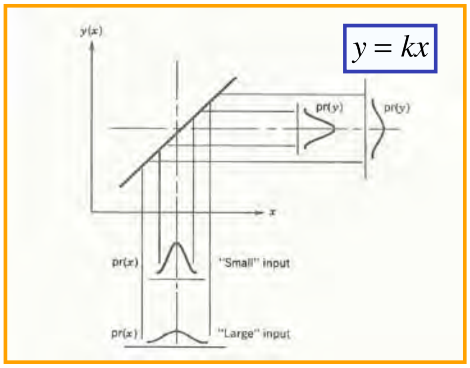
... and higher central moments scale as well

Linear Transformation of a Gaussian Probability Distribution

- **Probability distribution of *y* is Gaussian as well**

$$\mathrm{pr}(x)=\frac{1}{\sqrt{2\pi}\,\sigma_x}e^{-\frac{(x-\bar{x})^2}{2\sigma_x^2}}$$

$$\mathrm{pr}(y)=\frac{1}{\sqrt{2\pi}\,\sigma_y}e^{-\frac{(y-\bar{y})^2}{2\sigma_y^2}}=\frac{1}{\sqrt{2\pi}\,k\sigma_x}e^{-\frac{(y-k\bar{x})^2}{2k^2\sigma_x^2}}$$



Skew is zero:

$$E\left[(y-\bar{y})^3\right]=\int_{-\infty}^{\infty}(y-\bar{y})^3\,\mathrm{pr}(x)\,dx$$

$$=k^3\int_{-\infty}^{\infty}(x-\bar{x})^3\,\mathrm{pr}(x)\,dx=0$$

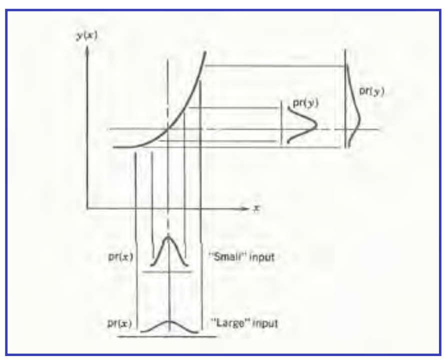
Nonlinear Transformation of a Probability Distribution

$$y=f(x)$$

$$y(x)=y(x_o)+\Delta y(\Delta x)$$

$$=f(x_o)+\left.\frac{\partial f}{\partial x}\right|_{x=x_o}\Delta x+\cdots$$

$$\therefore\ \Delta y(\Delta x)\approx\left.\frac{\partial f}{\partial x}\right|_{x=x_o}\Delta x$$

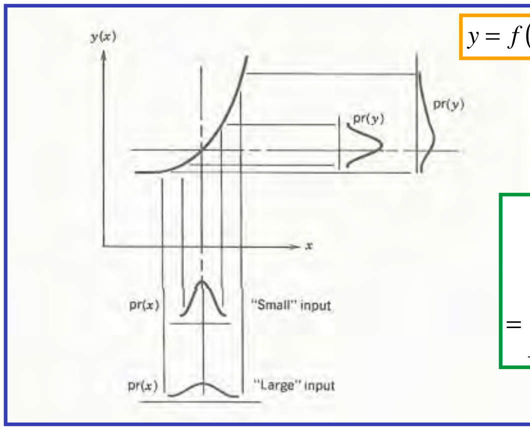


$$\mathrm{Pr}\left[y(x_o),y(x_o+\Delta x)\right]y(x)=\mathrm{Pr}\left[x_o,x_o+\Delta x\right]$$

$$\mathrm{pr}\left[y(x_o)\right]\Delta y=\mathrm{pr}\left[y(x_o)\right]\left(\left.\frac{\partial f}{\partial x}\right|_{x=x_o}\Delta x\right)=\mathrm{pr}\left[x_o\right]\Delta x$$

$$\therefore\ \text{in the limit}_{\Delta x\to 0}\ \Rightarrow \mathrm{pr}\left[y(x_o)\right]=\frac{\mathrm{pr}\left[x_o\right]}{\left|(\partial f/\partial x)_{x=x_o}\right|}$$

Nonlinear Transformation of a Gaussian Probability Distribution



$$E(y) = \bar{y} = \int_{-\infty}^{\infty} y \, \mathrm{pr}(x) dx$$

$$= \int_{-\infty}^{\infty} f(x) \, \mathrm{pr}(x) dx \neq k\bar{x} \quad \textbf{(in general)}$$

$$E\left[(y-\bar{y})^2\right] = \int_{-\infty}^{\infty} (y-\bar{y})^2 \, \mathrm{pr}(x) dx$$

$$= \int_{-\infty}^{\infty} \left[f(x)-\bar{y}\right]^2 \mathrm{pr}(x) dx \neq k^2 (x-\bar{x})^2 \quad \textbf{(in general)}$$

$$\textbf{Skew}: E\left[(y-\bar{y})^3\right] = \int_{-\infty}^{\infty} (y-\bar{y})^3 \, \mathrm{pr}(x) dx$$

$$\neq 0 \quad \text{(in general)}$$

Probability distribution of *y* is not Gaussian

Nonlinear Dynamic Systems with Random Inputs and Measurement Error

Continuous-time system with random inputs and measurement error

$$\dot{\mathbf{x}}(t) = \mathbf{f}\left[\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t), t\right], \quad \mathbf{x}(0) \text{ given}$$
$$\mathbf{z}(t) = \mathbf{h}\left[\mathbf{x}(t), \mathbf{u}(t), \mathbf{n}(t)\right]$$

$$E\left[\mathbf{w}(t)\mathbf{w}^T(\tau)\right] = \mathbf{Q}'_C \delta(t-\tau)$$
$$E\left[\mathbf{n}(t)\mathbf{n}^T(\tau)\right] = \mathbf{R}_C \delta(t-\tau)$$
$$E\left[\mathbf{w}(t)\mathbf{n}^T(\tau)\right] = 0$$

Discrete-time system with random inputs and measurement error

$$\mathbf{x}(t_{k+1}) = \mathbf{f}\left[\mathbf{x}(t_k), \mathbf{u}(t_k), \mathbf{w}(t_k), t\right], \quad \mathbf{x}(0) \text{ given}$$
$$\mathbf{z}(t) = \mathbf{h}\left[\mathbf{x}(t_k), \mathbf{u}(t_k), \mathbf{n}(t_k)\right]$$

$$E\left(\mathbf{w}_j \mathbf{w}_k^T\right) = \mathbf{Q}'_k \delta_{jk}$$
$$E\left(\mathbf{n}_j \mathbf{n}_k^T\right) = \mathbf{R}_k \delta_{jk}$$
$$E\left(\mathbf{w}_j \mathbf{n}_k^T\right) = 0$$

State Propagation for Nonlinear Dynamic Systems

Continuous-time system

$$\dot{\mathbf{x}}(t) = \mathbf{f}\left[\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t), t\right], \quad \mathbf{x}(0) \text{ given}$$

$$\mathbf{x}(t_{k+1}) = \mathbf{x}(t_k) + \int_{t_k}^{t_{k+1}} \mathbf{f}\left[\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t), t\right] dt$$

$$= \mathbf{x}(t_k) + \Delta\mathbf{x}(t_k)$$

Explicit numerical integration

Discrete-time system

$$\mathbf{x}(t_{k+1}) = \mathbf{f}\left[\mathbf{x}(t_k), \mathbf{u}(t_k), \mathbf{w}(t_k), t\right], \quad \mathbf{x}(0) \text{ given}$$

$$\mathbf{x}(t_{k+1}) = \mathbf{x}(t_k) + \Delta\mathbf{f}\left[\mathbf{x}(t_k), \mathbf{u}(t_k), \mathbf{w}(t_k), t_k\right]$$

$$= \mathbf{x}(t_k) + \Delta\mathbf{x}(t_k)$$

Explicit numerical summation

In both cases, the state propagation can be expressed as

$$\boxed{\mathbf{x}(t_{k+1}) = \mathbf{x}(t_k) + \Delta\mathbf{x}(t_k)}$$

Nonlinear Propagation of the Mean

The underlying deterministic model

$$\mathbf{x}(t_{k+1}) = \mathbf{x}(t_k) + \Delta\mathbf{f}\left[\mathbf{x}(t_k), \mathbf{u}(t_k), \mathbf{w}(t_k), t_k\right]$$

$$= \mathbf{x}(t_k) + \Delta\mathbf{x}(t_k)$$

Propagation of the <u>mean</u> (continuous- or discrete-time) for a random variable

$$E\left[\mathbf{x}(t_{k+1})\right] \triangleq \overline{\mathbf{x}}(t_{k+1}) = E\left[\mathbf{x}(t_k) + \Delta\mathbf{x}(t_k)\right]$$

$$= E\left[\mathbf{x}(t_k)\right] + E\left[\Delta\mathbf{x}(t_k)\right] \triangleq \overline{\mathbf{x}}(t_k) + \Delta\overline{\mathbf{x}}(t_k)$$

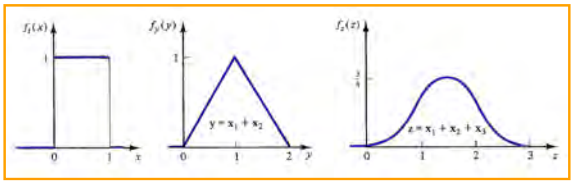
"State Estimate" usually defined as an <u>estimate of the mean</u> of the associated (non-Gaussian) random process

Probability Distribution Propagation for Nonlinear Dynamic Systems

Consequently

$$\boxed{\overline{\mathbf{x}}(t_{k+n}) = \overline{\mathbf{x}}(t_k) + \Delta\overline{\mathbf{x}}(t_k) + \Delta\overline{\mathbf{x}}(t_{k+1}) + \cdots + \Delta\overline{\mathbf{x}}(t_{k+n-1})}$$

Central limit theorem: probability distribution of x(t_{k+n}) approaches a Gaussian distribution for large *k+n*



$\cdots$even if $\text{pr}\left[\Delta x(t_{k+1}, t_k)\right]$ is not Gaussian

Mean and variance are dominant measures of the probability distribution of a system's state, x(t)

Nonlinear Propagation of the Covariance

Second central moment

$$E\left\{\left[\mathbf{x}(t_{k+1}) - \overline{\mathbf{x}}(t_{k+1})\right]\left[\mathbf{x}(\tau_{k+1}) - \overline{\mathbf{x}}(\tau_{k+1})\right]^T\right\} \triangleq \mathbf{P}(t_{k+1})$$

$$= E\left\langle\left\{\left[\mathbf{x}(t_k) + \Delta\mathbf{x}(t_k)\right] - \left[\overline{\mathbf{x}}(t_k) + \Delta\overline{\mathbf{x}}(t_k)\right]\right\}\left\{\left[\mathbf{x}(\tau_k) + \Delta\mathbf{x}(\tau_k)\right] - \left[\overline{\mathbf{x}}(\tau_k) + \Delta\overline{\mathbf{x}}(\tau_k)\right]\right\}^T\right\rangle$$

$$\delta\mathbf{x}(t_k) \triangleq \mathbf{x}(t_k) - \overline{\mathbf{x}}(t_k)$$

$$\mathbf{P}(t_{k+1}) = E\left\langle\left\{\delta\mathbf{x}(t_k) + \delta\left[\Delta\mathbf{x}(t_k)\right]\right\}\left\{\delta\mathbf{x}(\tau_k) + \delta\left[\Delta\mathbf{x}(\tau_k)\right]\right\}^T\right\rangle$$

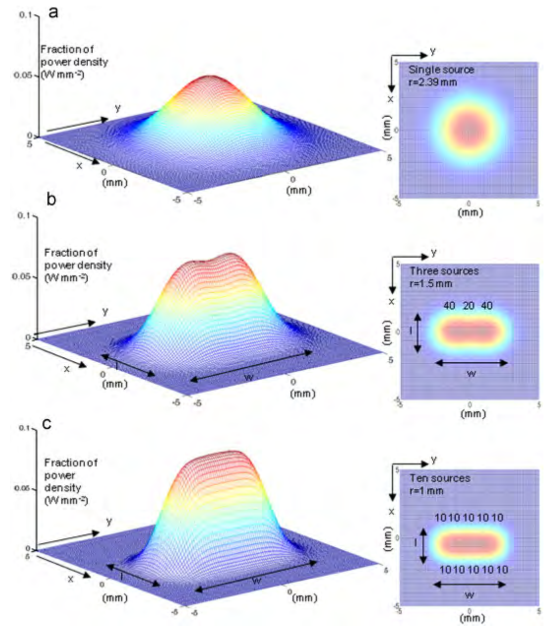
Covariance matrix

$$= E\left[\delta\mathbf{x}(t_k)\delta\mathbf{x}^T(\tau_k)\right] + E\left\{\delta\mathbf{x}(t_k)\delta\left[\Delta\mathbf{x}(\tau_k)\right]\right\} + E\left\{\delta\left[\Delta\mathbf{x}(t_k)\right]\delta\mathbf{x}^T(\tau_k)\right\} + E\left\{\delta\left[\Delta\mathbf{x}(t_k)\right]\delta\left[\Delta\mathbf{x}(\tau_k)\right]\right\}$$

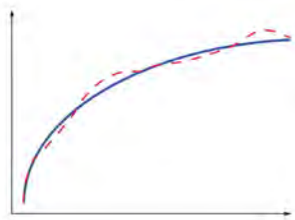
$$\triangleq \mathbf{P}(t_k) + \mathbf{M}(t_k) + \mathbf{M}^T(t_k) + \Delta\mathbf{P}(t_k)$$

Gaussian and Non-Gaussian Probability Distributions

- (Almost) all random variables have means and standard deviations
- Minimizing estimate error covariance tends to minimize the "spread" of the error in many (but not all) non-Gaussian cases
- <u>Central Limit Theorem</u> implies that estimate errors tend toward normal distribution



Neighboring-Optimal Estimator



Neighboring-Optimal Estimator

$$\dot{\mathbf{x}}(t) = \mathbf{f}\left[\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t), t\right]$$
$$\mathbf{z}(t) = \mathbf{h}\left[\mathbf{x}(t)\right] + \mathbf{n}(t)$$

- **Assume**
  - **Nominal solution exists**
  - **Disturbance and measurement errors are small**
  - **State stays close to the nominal solution**
  - **Mean and variance are good approximators of probability distribution**

$$\mathbf{x}_o(t), \mathbf{u}_o(t), \mathbf{w}_o(t) \text{ known in } \left[0, t_f\right]$$

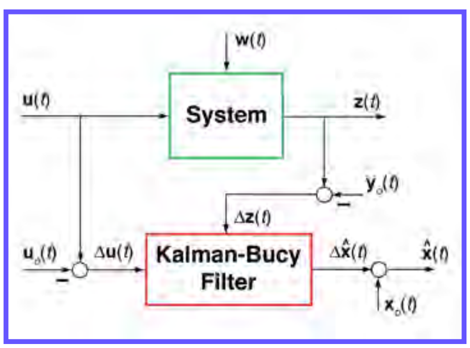
$$\dot{\mathbf{x}}_o(t) + \Delta\dot{\mathbf{x}}(t) = \mathbf{f}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), \mathbf{w}_o(t), t\right] + \left[\mathbf{F}(t)\Delta\mathbf{x}(t) + \mathbf{G}(t)\Delta\mathbf{u}(t) + \mathbf{L}(t)\Delta\mathbf{w}(t)\right]$$
$$\mathbf{z}_o(t) + \Delta\mathbf{z}(t) = \mathbf{h}\left[\mathbf{x}_o(t)\right] + \mathbf{n}_o(t) + \left[\mathbf{H}(t)\Delta\mathbf{x}(t) + \Delta\mathbf{n}(t)\right]$$

Jacobian matrices evaluated along the nominal path

$$\mathbf{F}(t) \triangleq \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), \mathbf{w}_o(t)\right]; \quad \mathbf{G}(t) \triangleq \frac{\partial \mathbf{f}}{\partial \mathbf{u}}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), \mathbf{w}_o(t)\right]; \quad \mathbf{L}(t) \triangleq \frac{\partial \mathbf{f}}{\partial \mathbf{w}}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), \mathbf{w}_o(t)\right]$$

Neighboring-Optimal Estimator

Estimate the perturbation from the nominal path



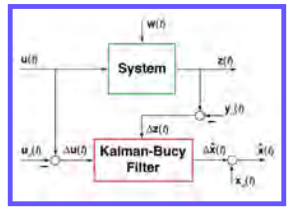
$$\Delta\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}(t)\Delta\hat{\mathbf{x}}(t) + \mathbf{G}(t)\Delta\mathbf{u}(t) + \mathbf{K}_C(t)\left[\Delta\mathbf{z}(t) - \mathbf{H}(t)\Delta\hat{\mathbf{x}}(t)\right]$$

$\mathbf{K}_C(t)$ is the LTV Kalman-Bucy gain matrix
LTV Kalman filter could be used to estimate the state at discrete instants of time

$$\hat{\mathbf{x}}(t) \simeq \mathbf{x}_{Nom}(t) + \Delta\hat{\mathbf{x}}(t)$$

Neighboring-Optimal Estimator Example

(from Gelb, 1974)



Radar tracking a falling body (one dimension)

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ d-g \\ 0 \end{bmatrix}; \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} altitude \\ velocity \\ ballistic\ coefficient, \beta \end{bmatrix}$$

$$z(t) = x_1(t) + n(t)$$

where

$$Drag = d = \frac{\rho V^2(t)}{2\beta(t)} = \frac{\rho x_2^2(t)}{2x_3(t)}$$

$$Density = \rho = \rho_o e^{-altitude/k} = \rho_o e^{-x_1(t)/k}$$

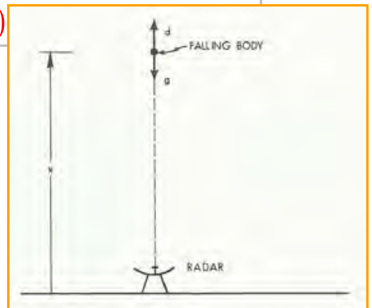


Figure 6.1-2 Geometry of the One-Dimensional Tracking Problem

Estimate Error for Example

(from Gelb, 1974)

$$\mathbf{w}(t) = \mathbf{0}$$

$$\mathbf{P}(0) = \begin{bmatrix} 500 & 0 & 0 \\ 0 & 20{,}000 & 0 \\ 0 & 0 & 2.5\times10^5 \end{bmatrix}$$

- **Pre-computed nominal trajectory**
- **Filter gains also may be pre-computed**
- **Position and velocity errors diverge**
- **Ballistic coefficient estimate error exceeds filter estimate**
- **Example of parameter estimation**

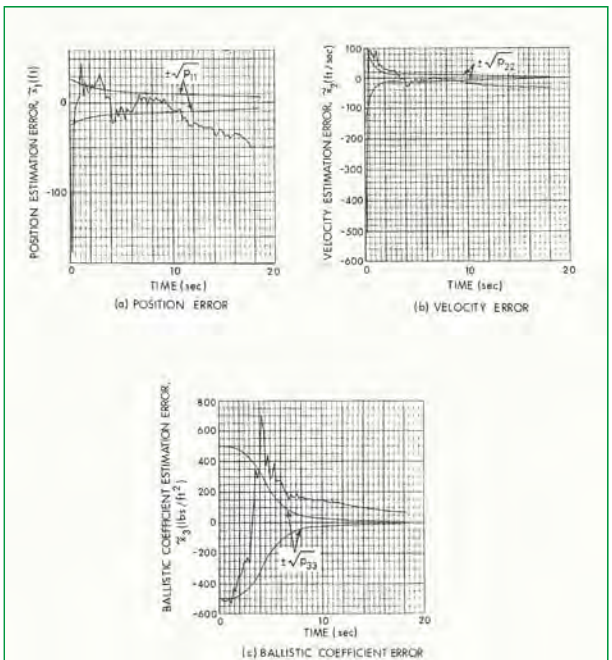


Figure 6.1-4 Performance of the Linearized Kalman Filter

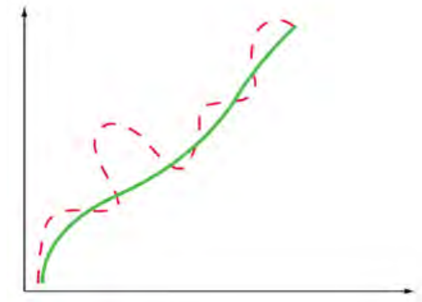
Extended Kalman Filters

Extended Kalman-Bucy Filter

$$\dot{\mathbf{x}}(t) = \mathbf{f}\left[\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t), t\right]$$
$$\mathbf{z}(t) = \mathbf{h}\left[\mathbf{x}(t)\right] + \mathbf{n}(t)$$

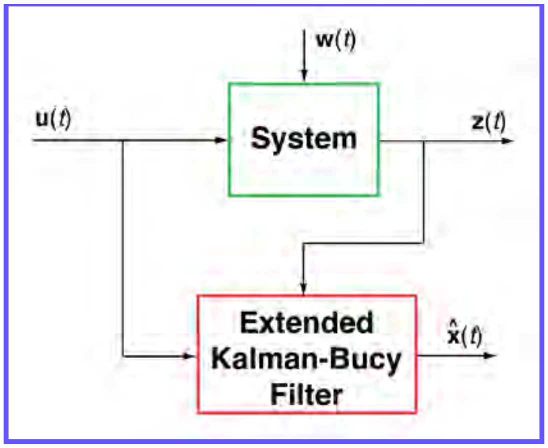
$$\mathbf{F}(t) \triangleq \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), \mathbf{w}_o(t)\right]; \quad \mathbf{G}(t) \triangleq \frac{\partial \mathbf{f}}{\partial \mathbf{u}}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), \mathbf{w}_o(t)\right];$$
$$\mathbf{L}(t) \triangleq \frac{\partial \mathbf{f}}{\partial \mathbf{w}}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), \mathbf{w}_o(t)\right]; \quad \mathbf{H}(t) \triangleq \frac{\partial \mathbf{h}}{\partial \mathbf{x}}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), \mathbf{w}_o(t)\right]$$

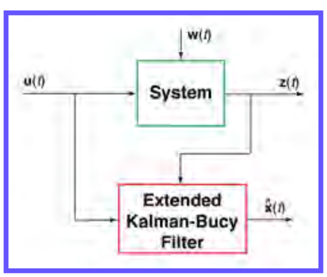
- **Assume**
  - No nominal solution is reliable or available
  - Disturbances and measurement errors may not be small



Extended Kalman-Bucy Filter

- In the estimator
  - Replace the linear dynamic model by the nonlinear model
  - Compute the filter gain matrix using the linearized model
  - Make linear update to the state estimate propagated by the nonlinear model





Extended Kalman-Bucy Filter

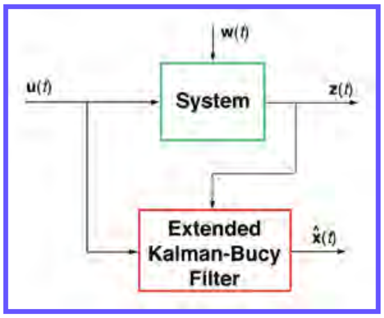
State Estimate: Nonlinear propagation plus linear correction

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}\left[\hat{\mathbf{x}}(t),\mathbf{u}(t),t\right] + \mathbf{K}_C(t)\left\{\mathbf{z}(t) - \mathbf{h}\left[\hat{\mathbf{x}}(t)\right]\right\}$$

$$\mathbf{x}(t_{k+1}) = \hat{\mathbf{x}}(t_k) + \int_{t_k}^{t_{k+1}} \left\langle \mathbf{f}\left[\hat{\mathbf{x}}(t),\mathbf{u}(t),t\right] + \mathbf{K}_C(t)\left\{\mathbf{z}(t) - \mathbf{h}\left[\hat{\mathbf{x}}(t)\right]\right\}\right\rangle dt$$

Filter Gain

$$\mathbf{K}_C(t) = \mathbf{P}(t)\mathbf{H}^T\left[\hat{\mathbf{x}}(t),\mathbf{u}(t)\right]\mathbf{R}_C^{-1}(t)$$



Extended Kalman-Bucy Filter

Covariance Estimate

$$\dot{\mathbf{P}}(t) = \mathbf{F}\left[\hat{\mathbf{x}}(t),\mathbf{u}(t)\right]\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T\left[\hat{\mathbf{x}}(t),\mathbf{u}(t)\right]$$
$$+\mathbf{L}\left[\hat{\mathbf{x}}(t),\mathbf{u}(t)\right]\mathbf{Q}'_C(t)\mathbf{L}^T\left[\hat{\mathbf{x}}(t),\mathbf{u}(t)\right] - \mathbf{K}_C(t)\mathbf{H}\left[\hat{\mathbf{x}}(t),\mathbf{u}(t)\right]\mathbf{P}(t)$$

- **Linear Kalman-Bucy filter**
  - State estimate is affected by the covariance estimate
  - Covariance estimate is not affected by the state estimate
  - Consequently, the covariance estimate is unaffected by the output, $\mathbf{z}(t)$
- **Extended Kalman-Bucy filter**
  - State estimate is affected by the covariance estimate
  - Covariance estimate is affected by the state estimate
  - Therefore, the covariance estimate is affected by the output, $\mathbf{z}(t)$

Extended Kalman-Bucy Filter Example

(from Gelb, 1974)

- Early tracking error is large
- Position, velocity, and ballistic coefficient errors converge to estimated bounds
- Filter gains must be computed on-line

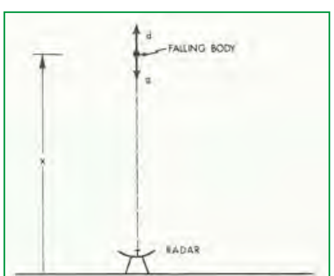


Figure 6.1-2 Geometry of the One-Dimensional Tracking Problem

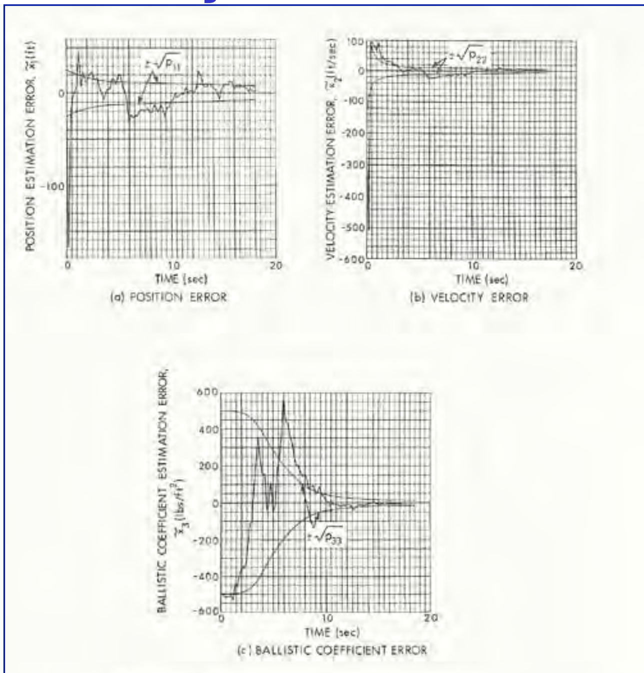


Figure 6.1-3 Performance of the Extended Kalman Filter

Hybrid Extended Kalman Filter

Numerical integration of propagation equations

State Estimate (–)

$$\hat{\mathbf{x}}\left[t_k(-)\right] = \mathbf{x}\left[t_{k-1}(+)\right] + \int_{t_{k-1}}^{t_k} \mathbf{f}\left[\hat{\mathbf{x}}\left[\tau(+)\right], \mathbf{u}(\tau)\right] d\tau$$

Covariance Estimate (–)

$$\mathbf{P}\left[t_k(-)\right] = \mathbf{P}\left[t_{k-1}(+)\right] + \int_{t_{k-1}}^{t_k} \left[\mathbf{F}(\tau)\mathbf{P}(\tau) + \mathbf{P}(\tau)\mathbf{F}^T(\tau) + \mathbf{L}(\tau)\mathbf{Q}'_C(\tau)\mathbf{L}^T(\tau)\right] d\tau$$

Jacobian matrices must be calculated

Hybrid Extended Kalman Filter

Recursive estimate updates

Filter Gain

$$\mathbf{K}(t_k) = \mathbf{P}\left[t_k(-)\right]\mathbf{H}^T(t_k)\left[\mathbf{H}(t_k)\mathbf{P}\left[t_k(-)\right]\mathbf{H}^T(t_k) + \mathbf{R}(t_k)\right]^{-1}$$

State Estimate (+)

$$\hat{\mathbf{x}}\left[t_k(+)\right] = \hat{\mathbf{x}}\left[t_k(-)\right] + \mathbf{K}(t_k)\left\langle \mathbf{z}(t_k) - \mathbf{h}\left\{\hat{\mathbf{x}}\left[t_k(-)\right]\right\}\right\rangle$$

Covariance Estimate (+)

$$\mathbf{P}\left[t_k(+)\right] = \left[\mathbf{I}_n - \mathbf{K}(t_k)\mathbf{H}(t_k)\right]\mathbf{P}\left[t_k(-)\right]$$

Iterated Extended Kalman Filter

(Gelb, 1974)

Re-apply the update equations to the updated solution to improve the estimate before proceeding

Re-linearize output matrix before each new update

State Estimate (+)

$$\hat{\mathbf{x}}_{k,i+1}(+)=\hat{\mathbf{x}}_k(-)+\mathbf{K}_{k,i}\left\{\mathbf{z}_k-\mathbf{h}\left[\hat{\mathbf{x}}_{k,i}(+)\right]-\mathbf{H}_k\left[\hat{\mathbf{x}}_{k,i}(+)\right]\left[\hat{\mathbf{x}}_k(-)-\hat{\mathbf{x}}_{k,i}(+)\right]\right\},\quad \hat{\mathbf{x}}_{k,0}(+)=\hat{\mathbf{x}}_k(-)$$

Arbitrary # of iterations: $i=0,1,\cdots$

Filter Gain

$$\mathbf{K}_{k,i}=\mathbf{P}_k(-)\mathbf{H}_k^T\left[\hat{\mathbf{x}}_{k,i}(+)\right]\left\{\mathbf{H}_k\left[\hat{\mathbf{x}}_{k,i}(+)\right]\mathbf{P}_k(-)\mathbf{H}_k^T\left[\hat{\mathbf{x}}_{k,i}(+)\right]+\mathbf{R}_k\right\}^{-1}$$

Covariance Estimate (+)

$$\mathbf{P}_{k,i+1}(+)=\left\{\mathbf{I}_n-\mathbf{K}_{k,i}\mathbf{H}_k\left[\hat{\mathbf{x}}_{k,i}(+)\right]\right\}\mathbf{P}_k(-)$$

Two-Dimensional Example

(from Gelb, 1974)

Same falling sphere dynamics, with offset radar

$$z(t_k)=\left\{r_1^2(t_k)+\left[x(t_k)-r_2\right]^2\right\}^{1/2}+n(t)$$

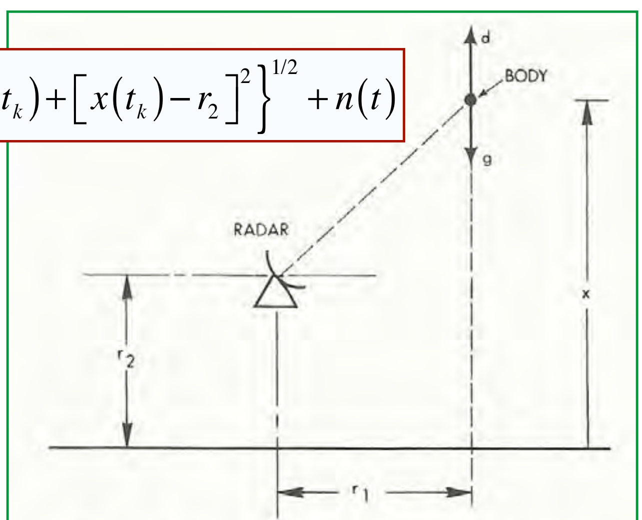


Figure 6.1-5 Geometry of the Two-Dimensional Tracking Problem

Two-Dimensional Example

(from Gelb, 1974)

Comparison of alternative nonlinear estimators
100-trial Monte Carlo evaluation

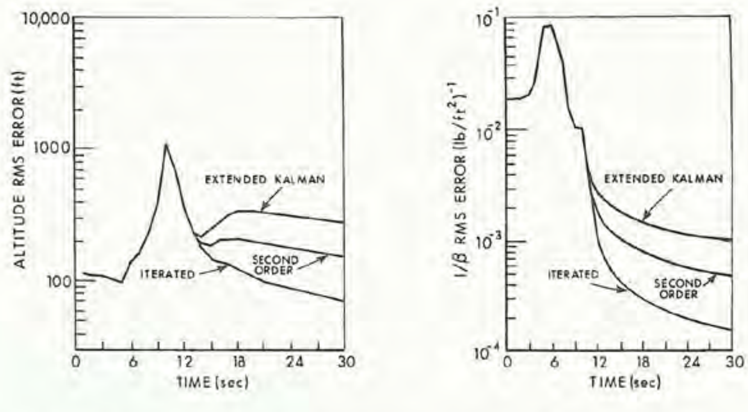


Figure 6.1-6 Comparative Performance of Several Tracking Algorithms (Ref. 5)

Second-order filter includes additional terms in **f[.]** *and* **h[.]**

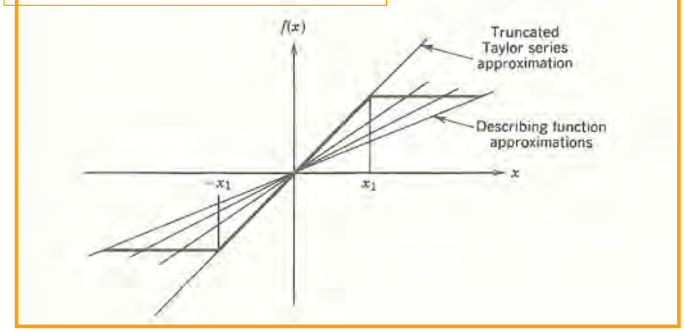
Quasilinearization (Describing Functions)

Quasilinearization

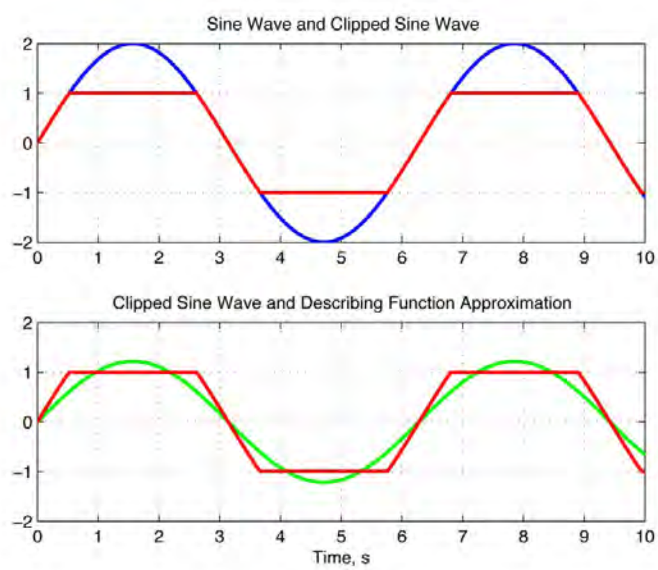
- **True linearization**: slope of a nonlinear curve at the evaluation point
- **Quasilinearization**: amplitude-dependent slope of a nonlinear curve at the evaluation point
- **Describing function**: quasilinear function of affine form:

$$Describing\ Function = Bias + Scale\ Factor\left(x - x_o\right)$$

Saturation (Limiting) Nonlinearity



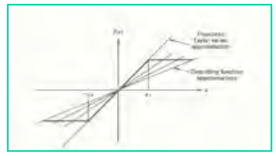
Comparison of Clipped Sine Wave with Describing Function Approximation



Deterministic Describing Function of the Saturation Function

- Describing function depends on wave form and amplitude of the input
- Saturation function

$$y = f(x) = \begin{cases} a, & x \ge a \\ x, & -a < x < a \\ -a, & x \le -a \end{cases}$$



- Describing function input = Nonlinear function input
- Sinusoidal input

$$\boxed{x(t) = A\sin\omega t}$$

- Clipped sine wave

$$y(t) = f[A\sin\omega t] = \begin{cases} a, & x \ge a \\ A\sin\omega t, & -a < x < a \\ -a, & x \le -a \end{cases}$$

Sinusoidal-Input Describing Function of the Saturation Function

(from Graham and McRuer, 1961)

- Approximate nonlinear function by linear function

$$\boxed{f(x) \approx d_0 + d_1(x - x_o)}$$

- Most readily calculated as the first term of a Fourier series for *y(t)*
- For symmetric input ($x_0 = 0$) to symmetric nonlinearity, $d_0 = 0$, and

$$d_1 = \frac{2A}{\pi}\left[\sin^{-1}\left(\frac{a}{A}\right) + \left(\frac{a}{A}\right)\sqrt{1 - \left(\frac{a}{A}\right)^2}\right]$$

a : Saturation limit
A : Input amplitude

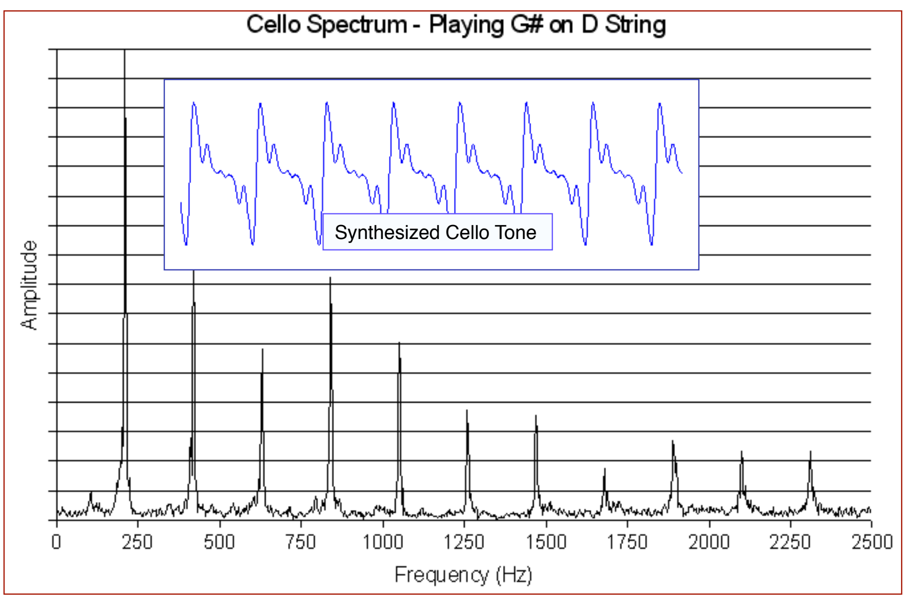
Sinusoidal-Input Describing Function of the Saturation Function

Describing function output

$$y_D(t) = d_1 \sin \omega t = \left\{ \frac{2A}{\pi} \left[\sin^{-1}\left(\frac{a}{A}\right) + \left(\frac{a}{A}\right)\sqrt{1-\left(\frac{a}{A}\right)^2} \right] \right\} \sin \omega t$$

See "Describing Function Analysis of Nonlinear Simulink Models" **in *Simulink Control Design 3.1***

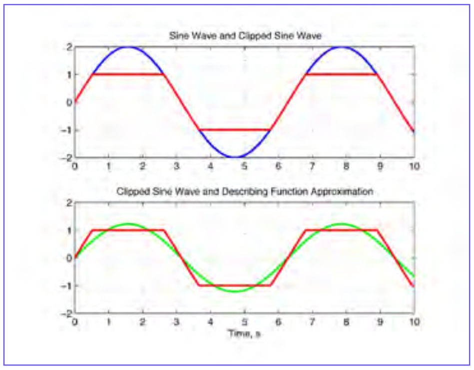
Nonlinearity Introduces Harmonics in Output

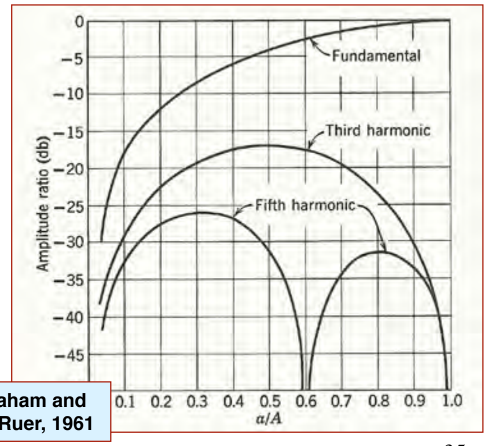


Harmonic Describing Functions of Saturation

Fourier series of symmetrically clipped sine wave includes symmetric harmonic terms

$$y_D(t) = d_1 \sin \omega t + d_3 \sin(3\omega t + \varphi_3) + d_5 \sin(5\omega t + \varphi_5) + \cdots$$





Graham and McRuer, 1961

Describing Function Derived from Expected Values of Input and Output

$$\textit{Describing Function} = \textit{Bias} + \textit{Scale Factor}\,(x - x_o)$$

Approximate nonlinear function by linear function

$$f(x) \approx d_0 + d_1(x - x_o)$$

Statistical representation of fit error

$$J = E\left\{\left[f(x) - d_0 - d_1(x - x_o)\right]^2\right\}$$

Minimize fit error to find d_0 and d_1

$$\frac{\partial J}{\partial d_0} = 0; \qquad \frac{\partial J}{\partial d_1} = 0$$

Random-Input Describing Function

$$\tilde{x} \triangleq x - x_o$$

Bias and scale factor

$$d_0 = E\left[f(x)\right] - d_1 E[\tilde{x}]$$

$$d_1 = \frac{E\left[\tilde{x}f(x)\right] - d_0 E[\tilde{x}]}{E\left[\tilde{x}^2\right]}$$

by elimination

$$d_0 = \frac{E\left[\tilde{x}^2\right]E\left[f(x)\right] - E[\tilde{x}]E\left[\tilde{x}f(x)\right]}{E\left[\tilde{x}^2\right] - E^2[\tilde{x}]}$$

$$d_1 = \frac{E\left[\tilde{x}f(x)\right] - E[\tilde{x}]E\left[f(x)\right]}{E\left[\tilde{x}^2\right] - E^2[\tilde{x}]}$$

Random-Input Describing Function

If $f(x)$ and $\tilde{x}$ both have zero mean

$$d_0 = 0$$

$$d_1 = \frac{E\left[\tilde{x}f(x)\right]}{E\left[\tilde{x}^2\right]} = \frac{E\left[xf(x)\right]}{E\left[x^2\right]}$$

Describing function for symmetric function

$$\textit{Describing Function} = d_0 + d_1\left(x - x_o\right) = d_1 x$$

$$= \frac{E\left[xf(x)\right]}{E\left[x^2\right]} x$$

Random-Input Describing Function for the Saturation Function

(from Graham and McRuer, 1961)

Describing function input: White noise with standard deviation, σ

$$x(t) \sim N(0,\sigma) \sim \text{Zero-mean white noise with standard deviation, } \sigma$$

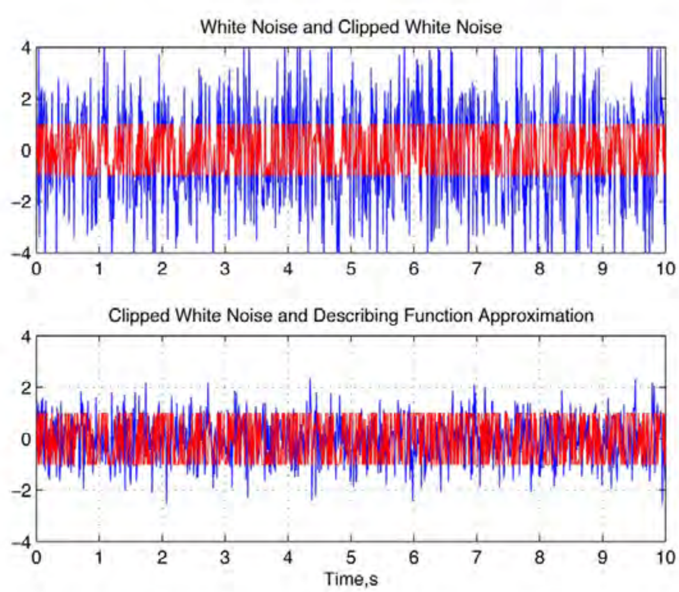
Describing function output

$$y_D(t) = \frac{E\left[\tilde{x}f(x)\right]}{E\left[\tilde{x}^2\right]} x(t) = d_1 x(t) = \operatorname{erf}\left(\frac{a}{\sqrt{2}\sigma}\right) x(t)$$

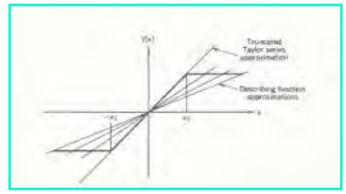
where the **error function**, erf(.), is

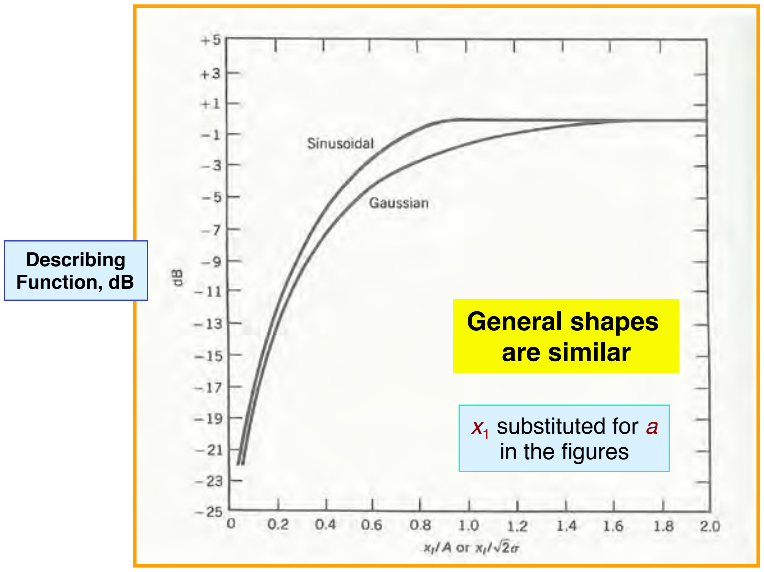
$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-\lambda^2} \, d\lambda$$

Comparison of Clipped White Noise with Describing Function Approximation



Random-Input and Sinusoidal Describing Functions of the Saturation Function





Multivariate Describing Functions

- Let **f(x)** be a nonlinear vector function of a vector
- The quasilinear describing function approximation is

$$\boxed{\begin{aligned}&\mathbf{f}(\mathbf{x}) \approx \mathbf{b} + \mathbf{D}(\mathbf{x} - \hat{\mathbf{x}})\\ &where \quad \hat{\mathbf{x}} = E(\mathbf{x})\end{aligned}}$$

- Cost function = <u>Trace</u> of the error covariance matrix

$$\begin{aligned}J &= E\Big[\mathbf{Tr}\Big\{\big[\mathbf{f}(\mathbf{x}) - \mathbf{b} - \mathbf{D}(\mathbf{x} - \hat{\mathbf{x}})\big]\big[\mathbf{f}(\mathbf{x}) - \mathbf{b} - \mathbf{D}(\mathbf{x} - \hat{\mathbf{x}})\big]^T\Big\}\Big]\\ &= E\Big[\mathbf{Tr}\Big\{\big[\mathbf{f}(\mathbf{x}) - \mathbf{b} - \mathbf{D}(\mathbf{x} - \hat{\mathbf{x}})\big]^T\big[\mathbf{f}(\mathbf{x}) - \mathbf{b} - \mathbf{D}(\mathbf{x} - \hat{\mathbf{x}})\big]\Big\}\Big]\end{aligned}$$

- **Quasilinear extended Kalman-Bucy filter**
  - **F, G, and L replaced by describing function matrices**

Describing Function Matrices for State Estimation

Minimize fit error to find b and D

$$\frac{\partial J}{\partial \mathbf{b}} = \mathbf{0}; \qquad \frac{\partial J}{\partial \mathbf{D}} = \mathbf{0}$$

Describing function bias *(see text)*

$$\mathbf{b} = E\left[\mathbf{f}(\mathbf{x})\right] \triangleq \hat{\mathbf{f}}(\mathbf{x})$$

Describing function scaling matrix is a function of the covariance inverse *(see text)*

$$\mathbf{D} = E\left\{\left[\mathbf{f}(\mathbf{x})\tilde{\mathbf{x}}^T\right] - E\left[\mathbf{f}(\mathbf{x})\right]\tilde{\mathbf{x}}^T\right\}\mathbf{P}^{-1}$$

where

$$\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}; \quad \mathbf{P} = E\left(\tilde{\mathbf{x}}\,\tilde{\mathbf{x}}^T\right)$$

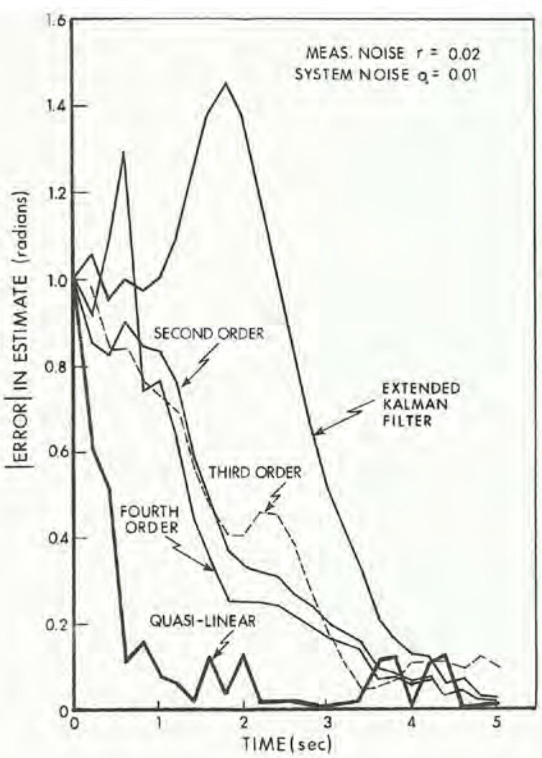


Figure 6.2-5 Estimation Errors for Five Filters (Ref. 6)

Monte Carlo Comparison of Quasilinear Filter with Extended Kalman Filter and Three Others

(from Gelb, 1974)

$$\dot{x}(t) = -\sin x(t) + w(t)$$
$$z(t) = 0.5\sin(2x_k) + n_k$$

Next Time:
Sigma Points (Unscented Kalman) Filters

plus Brief Introduction to Particle, Batch Least-Squares, Backward-Smoothing, Gaussian Mixture Filters

Supplemental Material

Quasilinear Filter Propagation Equations

True linearization Jacobians replaced by quasilinear (describing function) Jacobians

$$\hat{\mathbf{x}}\left[t_k(-)\right] = \mathbf{x}\left[t_{k-1}(+)\right] + \int_{t_{k-1}}^{t_k} \mathbf{f}\left[\hat{\mathbf{x}}\left[\tau(+)\right], \mathbf{u}(\tau)\right] d\tau$$

$$\mathbf{P}\left[t_k(-)\right] = \mathbf{P}\left[t_{k-1}(+)\right] + \int_{t_{k-1}}^{t_k} \left[\mathbf{D_F}(\tau)\mathbf{P}(\tau) + \mathbf{P}(\tau)\mathbf{D_F}^T(\tau) + \mathbf{D_L}(\tau)\mathbf{Q}'_C(\tau)\mathbf{D_L^T}(\tau)\right] d\tau$$

$\mathbf{D_F}$: $(n \times n)$ Stability matrix of $\mathbf{f}(\mathbf{x})$ containing describing function elements

Covariance propagation is state-dependent

Quasilinear System Stability Matrix

True linearization of nonlinear system equation

$$\begin{aligned}\dot{\mathbf{x}}_o(t) + \Delta\dot{\mathbf{x}}(t) &\approx \mathbf{f}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), t\right] + \mathbf{F}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), t\right]\Delta\mathbf{x}(t) + \cdots \\ &= \mathbf{f}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), t\right] + \mathbf{F}(t)\Delta\mathbf{x}(t) + \cdots\end{aligned}$$

Quasilinearization of nonlinear system equation

Some or all elements of stability matrix are state-dependent

$$\begin{aligned}\dot{\mathbf{x}}_o(t) + \Delta\dot{\mathbf{x}}(t) &\approx \mathbf{f}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), t\right] + \mathbf{F}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), t\right]\Delta\mathbf{x}(t) + \cdots \\ &= \mathbf{f}\left[\mathbf{x}_o(t), \mathbf{u}_o(t), t\right] + \mathbf{D_F}\left[E\left[\Delta\mathbf{x}(t)\right], \mathbf{x}_o(t), \mathbf{u}_o(t), t\right]\Delta\mathbf{x}(t) + \cdots\end{aligned}$$

Quasilinear Filter Gain and Updates

Filter Gain

$$\mathbf{K}(t_k) = \mathbf{P}[t_k(-)]\mathbf{D}_{\mathbf{H}}^T(t_k)\left[\mathbf{D}_{\mathbf{H}}(t_k)\mathbf{P}[t_k(-)]\mathbf{D}_{\mathbf{H}}^T(t_k) + \mathbf{R}(t_k)\right]^{-1}$$

State Estimate Update

$$\hat{\mathbf{x}}[t_k(+)] = \hat{\mathbf{x}}[t_k(-)] + \mathbf{K}(t_k)\left\langle \mathbf{z}(t_k) - \mathbf{h}\left\{\hat{\mathbf{x}}[t_k(-)]\right\}\right\rangle$$

Covariance Estimate Update

$$\mathbf{P}[t_k(+)] = \left[\mathbf{I}_n - \mathbf{K}(t_k)\mathbf{D}_{\mathbf{H}}(t_k)\right]\mathbf{P}[t_k(-)]$$

$\mathbf{D}_{\mathbf{H}}$: Output matrix of $\mathbf{h}(\mathbf{x})$ containing describing function elements