7 DISCRETE AND STOCHASTIC SYSTEMS

7.1 Discrete Systems

Recall our basic continuous system in state space form,

$$\begin{aligned} \dot{x} &= Ax + Bu \ , \\ y &= Cx \ . \end{aligned}$$

A control system that is to be implemented using a digital computer, as is usually the case, is in a discrete state space form,

$$\begin{aligned} x_{n+1} &= A_d x_n + B_d u_n \ , \\ y_n &= C_d x_n \ . \end{aligned}$$

The first thing we have to do is to be able to go from the continuous to the discrete model. We start with the solution to the state equations in the form

$$x(t) = e^{A(t-t_0)} x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)} B u(\tau) \, d\tau \ .$$

We can use this solution over one *sample* period T to obtain a difference equation. Let

$$\begin{aligned} t &= nT + T \ , \\ t_0 &= nT \ , \end{aligned}$$

and we get

$$x(nT+T) = e^{AT} x(nT) + \int_{nT}^{nT+T} e^{A(nT+T-\tau)} B u(\tau) \, d\tau \ .$$

Now assume that the input does not change within one sample period,

$$u(\tau) = u(nT) \quad \text{for} \quad nT \leq \tau < nT + T \ .$$

We refer to this operation as the zero–order hold with no delay. Then, by defining the auxiliary variable

$$\eta = nT + T - \tau \ ,$$

we get

$$x(nT+T) = e^{AT} x(nT) + \int_{0}^{T} e^{A\eta} B u(nT) \, d\eta \ .$$

Therefore, the system

$$\begin{aligned} \dot{x} &= Ax + Bu \ , \\ y &= Cx \ , \end{aligned}$$

becomes

$$\begin{aligned} x_{n+1} &= A_d x_n + B_d u_n \ , \\ y_n &= C_d x_n \ , \end{aligned}$$

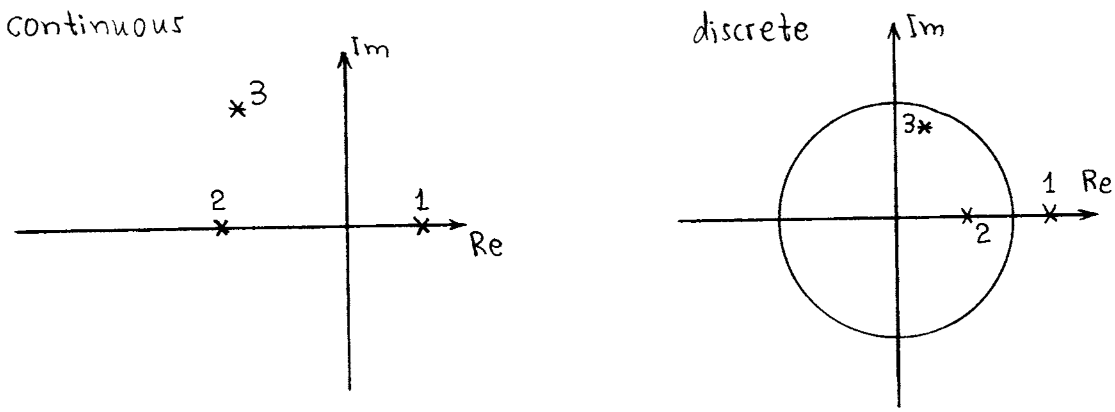

Figure 1: Stable poles for continuous and discrete systems

where

$$
\begin{aligned}
A_d &= e^{AT} , \\
B_d &= \int_0^T e^{A\eta} B \, d\eta , \\
C_d &= C ,
\end{aligned}
$$

and T is the sample period. The MATLAB command `c2d` automates the above conversion from continuous to discrete form.

A low sample period T; i.e., high sample rate, is in general desirable for good performance so that we can approximate the continuous model as closely as possible. This, however, will demand a fast computer and A/D and D/A converters. It should be emphasized here that low T is always with respect to the response time of the physical system. Low T for one system may be high for a different system. Low sample rate, high T, may lead to instabilities when the design is based on the continuous system. In such a case we should switch to a direct discrete design. This means that the continuous system is discretized first, and any compensator design is based on the discrete version. Fortunately this parallels the continuous design we have already developed.

We can place the poles of a discrete system to desirable locations by linear state variable feedback,

$$u_n = -K x_n ,$$

and if not all states are measurable we can use a discrete full–order estimator,

$$\widehat{x}_{n+1} = A_d \widehat{x}_n + B_d u_n + L(y_n - C_d \widehat{x}_n) .$$

We can find the gain matrices K and L by poleplacement of

$$A_d - B_d K ,$$

and

$$A_d - L C_d .$$

We already know how to do the poleplacement design, the only thing we need to know is: When is a discrete system $x_{n+1} = Ax_n$ stable? We can see this by considering a scalar system. Consider the continuous system

$$\dot{x} = ax \ .$$

The solution is $x(t) = e^{at}x(0)$ so if $\Re\{a\} < 0$ the system will be stable. The discrete system

$$x_{n+1} = ax_n \ ,$$

has

$$\begin{aligned} x_1 &= ax_0 \ , \\ x_2 &= ax_1 = a^2x_0 \ , \\ x_3 &= ax_2 = a^3x_0 \ , \end{aligned}$$

and, finally,

$$x_n = a^n x_0 \ .$$

For stability, we want $x_n \to 0$ as $n \to \infty$, or $a^n \to 0$, which means that we want

$$|a| < 1 \ .$$

Therefore, the discrete time system $x_{n+1} = Ax_n$ is stable if and only if all eigenvalues of A have absolute value less that one; i.e., they are located inside the unit circle in the s–plane, see Figure 32. Since the continuous matrix A becomes e^{AT} when discretized, we can argue that an eigenvalue which is equal to λ for a continuous system, corresponds to an eigenvalue equal to $e^{\lambda T}$ for a discrete system with sample period T. By keeping this analogy in mind we can do in discrete time everything we did in continuous time. The corresponding MATLAB commands have the same names with simply the prefix `d` in front, for example `dlqr` will do the discrete LQR design.

As an example, consider the system

$$\dot{x} = x + u \ ,$$

which is open–loop unstable. A control law of the form $u = -2x$ places the closed loop pole of the continuous system at -1, this means that the continuous system has a time constant 1 second. Now let's discretize the system using a sample period T, we set the closed loop pole of the discrete system at e^{-T}. How different will be the discrete gain from the continuous gain 2? This should depend strictly on T. If T is very small compared to 1, the time constant of the system, then the two gains must be relatively close. Ten times smaller should be small enough. On the other hand, if T is of the same order of magnitude as 1, we have to compute the gain from the discrete design. This is illustrated by the results of Figure 33 where we present the discrete time gain for a discrete closed loop pole at e^{-T}, versus T for T from 0.01 sec to 1 sec. This corresponds to sample rates from 100 Hz to 1 Hz, respectively.

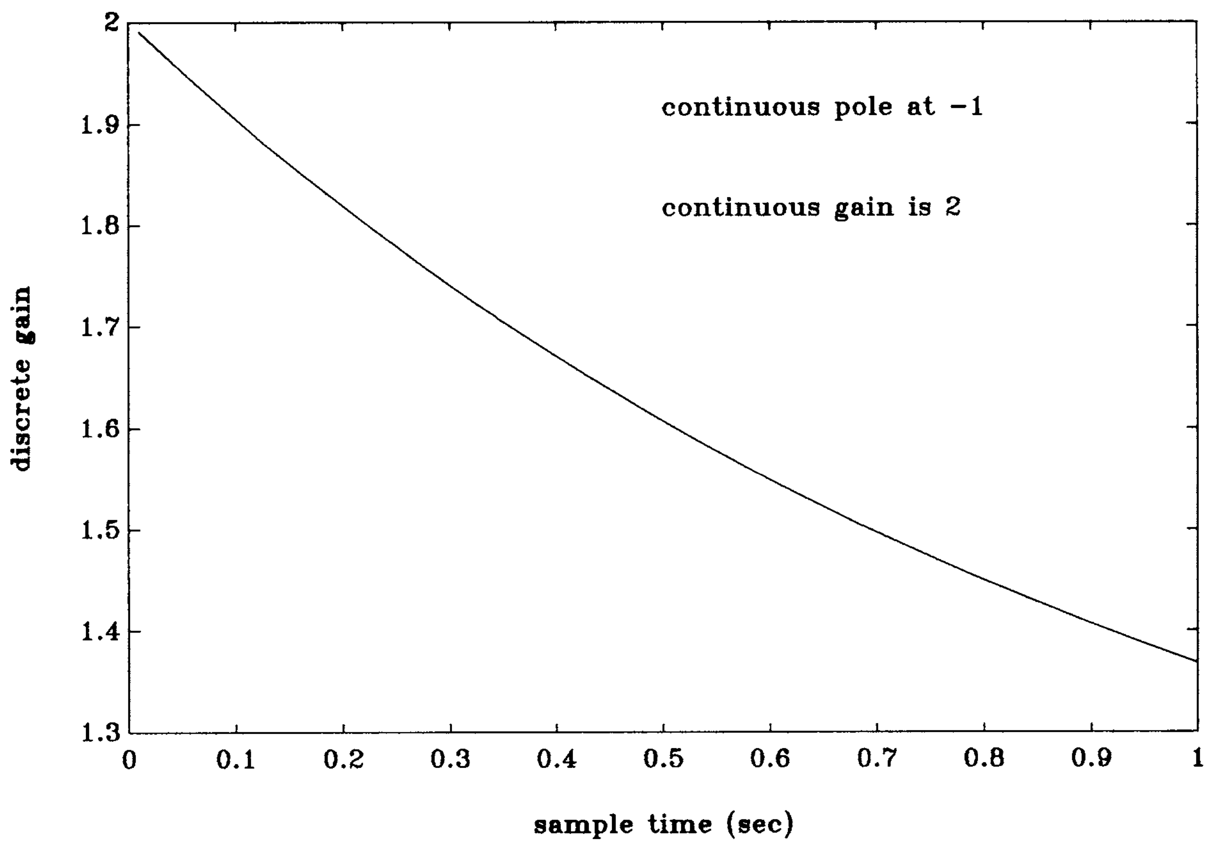

Figure 2: Discrete system example

7.2 Stochastic Processes

To this point we have treated the entire control/estimation problem as deterministic; everything had a known value at each time. In real world problems, however, there are quantities which we can only describe probabilistically, for example sensor characteristics or sea waves. There are unpredictable disturbances and measurement noise which occur during operation of real systems. These disturbances and noise can be modeled as stochastic processes. A very useful special class of stochastic processes is the Gauss–Markov process which can be completely described by the following:

1. Its mean value vector $\overline{x}$,
$$\underbrace{\overline{x}}_{(n\times 1)} \equiv E\left[x(t)\right] \ ,$$
which gives the expected value or ensemble average of all possible observations at time t; this is the most likely value.

2. Its correlation matrix C,
$$\underbrace{C(t,\tau)}_{(n\times n)} \equiv E\left\{\left[x(t)-\overline{x}(t)\right]\left[x(\tau)-\overline{x}(\tau)\right]^T\right\} \ ,$$
which is a symmetric matrix and gives the relationship between the deviation from the mean at time t to the deviation from the mean at a different time τ.

When $t=\tau$, this correlation matrix becomes the covariance matrix which measures the mean square deviation of the state vector from the mean; i.e.,

$$\underbrace{X(t)}_{(n\times n)} \equiv C(t,t) = E\left\{\left[x(t)-\overline{x}(t)\right]\left[x(t)-\overline{x}(t)\right]^T\right\} \ .$$

At any time t, the state $x(t)$ is normally distributed (Gaussian distribution) about the mean and the diagonal elements of $X(t)$ give the variance (standard deviation squared) for the associated elements of x.

A special Gauss–Markov process is the purely random process. This is an idelized, very jittery process which is completely uncorrelated from one time to the next. This is a useful model for disturbances or noise which change very rapidly compared with the time response of a system. The correlation matrix for a purely random process is

$$C(t,\tau) = \underbrace{Q(t)}_{(n\times n)} \ \delta(t-\tau) \ ,$$

where $Q(t)$ is the *power spectral density*, and $\delta(t-\tau)$ is the Dirac delta function; this is zero everywhere except at $t=\tau$ where it assumes a “value” such that $\int_{-\infty}^{+\infty}\delta(t-\tau)d\tau = 1$. This can be viewed as the limit of a sequence of impulses of random magnitude (equal plus and minus so the mean is zero; average square magnitude is $\sigma^2(t)$) and random time of occurence. For such a sequence,

$$Q(t) \approx 2\left[\sigma(t)\right]^2 \beta(t) \ ,$$

where $\beta(t)$ is the average number of occurences per unit time.

The key behind using Gauss–Markov processes is that a Gauss–Markov process can always be represented by a state vector of a linear dynamical system forced by a Gaussian purely random process where the initial state vector is Gaussian. Thus,

$$\dot{x} = Ax + \Gamma w \ ,$$

where

$$\begin{aligned}
&E[w(t)] = \overline{w} = 0 \ , \\
&E\left[w(t)w^T\right] = Q(t)\delta(t-\tau) \ , \\
&E\left[x(t_0)\right] = \overline{x}_0 \ , \\
&E\left\{\left[x(t_0) - \overline{x}_0\right]\left[x(t_0) - \overline{x}_0\right]^T\right\} = X_0 \ , \\
&E\left\{\left[w(t) - \overline{w}\right]\left[x(t_0) - \overline{x}_0\right]^T\right\} = 0 \ .
\end{aligned}$$

The forcing disturbance w and the initial state $x(t_0)$ are completely independent or uncorrelated. Recall the state property for deterministic systems: knowing the current state and the state equation completely defines the future for zero control. The *Markov property* is completely parallel to this: knowing the current state mean $\overline{x}_0$ and covariance matrix X_0 completely defines the future mean and covariance for zero control when subjected to the disturbance described by $\overline{w} = 0$ and Q. The *Gaussian property* states that the state will always be normally distributed about the mean value in accordance with the variance (standard deviation squared) given by the diagonal elements of the covariance matrix. Thus for one state x, it will be within one standard deviation σ of $\overline{x}$ 68.3% of the time; within 2σ of $\overline{x}$ 95.5% of the time; within 3σ of $\overline{x}$ 99.7% of the time. For multiple states these percentages decrease as shown in the following table:

n	σ	2σ	3σ
1	68.3	95.5	99.7
2	39.4	86.5	98.9
3	20.0	73.9	97.1

The mean value vector of a Gauss–Markov process obeys the state differential equation

$$\dot{\overline{x}} = A\overline{x} + \Gamma\overline{w} \ , \quad \overline{x}(t_0) = \overline{x}_0 \ .$$

The covariance matrix obeys equation

$$\dot{X} = AX + XA^T + \Gamma Q \Gamma^T \ , \quad X(t_0) = X_0 \ ,$$

which is completely independent and which could be calculated in advance. Note that the term $AX + XA^T$ represents the effect of the system dynamics and it may decrease X for

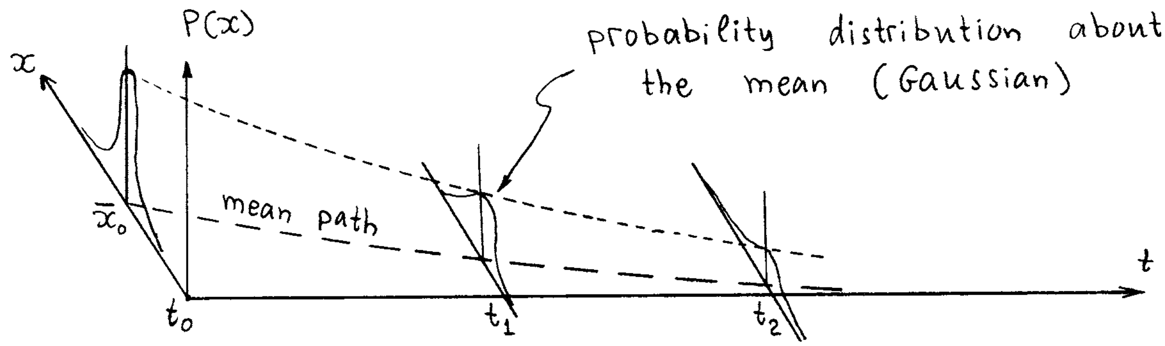

Figure 3: Response of first order system to noise

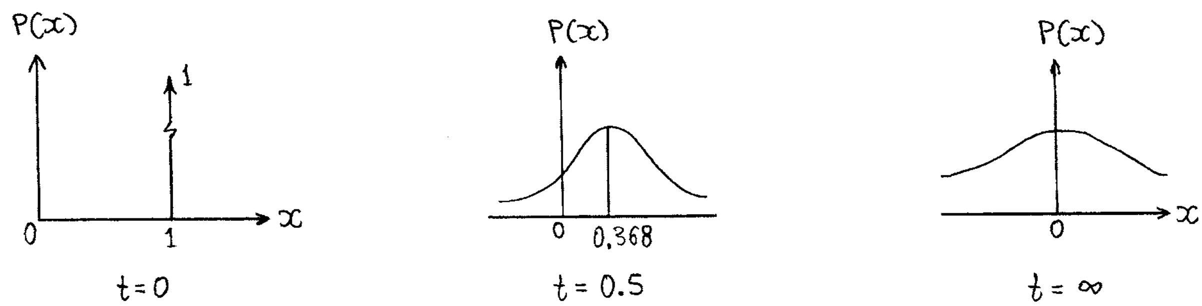

Figure 4: "Snapshots" of Figure 34 at different times

a stable system, while the other term $\Gamma Q \Gamma^T$ represents the effect of the disturbance and it always increases X since we have a positive definite Q.

We can visualize this by considering a simple first order system so that all the above matrices are scalars. A stable first order system with an initial mean $\overline{x}_0$ and small standard deviation σ_0 could be released while subjected to noise. It could respond as shown in Figure 34 for large noise. As an example suppose we have the system

$$\dot{x} + 2x = w \ ,$$

where w is zero mean, purely random (white noise), x is exactly 1 at $t = 0$. At this time, x is released and the disturbance w with power spectral density $Q = q = 3$ begins to act on the system. We want to determine the mean and the covariance of the response. The mean will follow the state equation

$$\dot{\overline{x}} = A\overline{x} + \Gamma\overline{w} = -2\overline{x} \ , \quad \overline{x}(0) = 1 \ , \quad A = -2 \ ,$$

and $\overline{w} = 0$ since w is white noise. The solution for the mean is

$$\overline{x}(t) = e^{-2t} \ .$$

The covariance will follow equation

$$\dot{X} = AX + XA^T + \Gamma Q \Gamma^T \ , \quad Q = q = 3 \ , \quad \Gamma = 1 \ ,$$

or

$$\dot{X} = -2X - 2X + q \ ,$$

and, with exact knowledge at $t = 0$, the initial condition is $X(0) = 0$. The solution is

$$X(t) = 0.75 \left(1 - e^{-4t}\right) = \sigma^2 \ ,$$

the variance of $x(t)$ about its mean $\overline{x}(t)$, refer to Figure 35.

normally distributed			
t	$\overline{x}$	X	σ
0	1	0	0
0.5	0.368	0.648	0.805
∞	0	0.750	0.866

Most physical disturbances can be modeled by one of the following special cases:

1. *White noise*: A stationary, purely random Gauss–Markov process with zero correlation time (see below) and constant power spectral density,

$$C = Q\delta(t - \tau) \ .$$

2. *Random bias*: A random, unpredictable constant with infinite correlation time and constant correlation. In this case we introduce

$$\dot{x}_{n+1} = 0 \ , \quad x_{n+1}(t_0) = \text{random} \ .$$

Figure 5: White noise and random bias

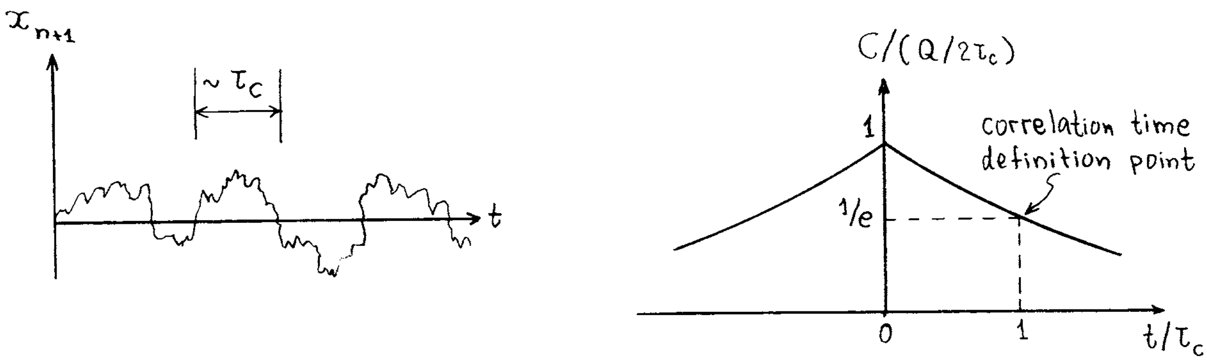

Figure 6: Exponentially correlated noise

We add a new, constant state to the system of equations; i.e., we augment the state equations and estimate x_{n+1} along with the rest of the states x_i, $i = 1, \ldots, n$, as we have already seen before. As examples, a disturbance which changes rapidly compared to the dominant dynamics of the system can be modeled as white noise; e.g. wave effects on the steering of a large tanker. A disturbance which changes very slowly compared to the dominant dynamics of the system can be modeled as a random bias; e.g. tidal current on ship steering.

3. *Exponentially correlated noise*: Between the two extremes where white noise and random bias models are appropriate, are disturbances which change on the same time scale as the dominant dynamics of the system. These disturbances have finite, non–zero correlation times τ_c. The simplest can be modeled as a first order system driven by white noise; i.e.,

$$\tau_c \dot{x}_{n+1} + x_{n+1} = w \ .$$

In these cases the state vector can be augmented with x_{n+1}. Disturbances which change with about the same dynamics as the system must be modeled with a finite τ_c; e.g. the force and moment produced by a passing ship during underway replenishment. The above equation is called a *shaping filter* because it "shapes" white noise w to produce another disturbance x_{n+1} which is called "colored" noise. The correlation time is the same as the time constant of the disturbance variation, this can be obtained by considering the physics of the problem. For example, if the disturbance is the force produced by a passing ship we can take τ_c to be approximately the time it takes to travel a ship length.

To complete the model for the exponentially correlated disturbance it is necessary to specify the power spectral density of the white noise w. This is given by,

$$q = 2\sigma^2\tau_c \ ,$$

where

$$\begin{aligned} \sigma &= \text{root mean sqaue (RMS) noise level,} \\ \tau_c &= \text{correlation time.} \end{aligned}$$

The same formula is also used in design to establish the power spectral density of disturbances modeled as white noise. In that case the correlation times (modeled as zero) are actually small nonzero quantities compared to the time constants of the system. In practice this can be the integration time step in simulations, or the sample time in experiments.

More complex models for modeling disturbances are also possible, this is a trade–off between accuracy and simplicity. Of great interest to naval engineering is the modeling of the disturbance due to waves. The simplest approach would be to model this as white noise, this is very accurate for large ships. For smaller vessels it might be worth modeling the *periodic* nature of the disturbance. There are a couple of ways to do this. If we assume a sinusoidal wave as the dominant model for waves in the area, we can use a second order model driven by white noise w,

$$\ddot{y} + \omega^2 y = w \ ,$$

where ω is the assumed frequency of the dominant wave (usual periods of sea waves are in the 6 to 15 sec range), and y is the amplitude of the disturbance. In state space form then we need to augment our system with two additional equations

$$\begin{aligned} \dot{x}_{n+1} &= x_{n+2} \ , \\ \dot{x}_{n+2} &= -\omega^2 x_{n+1} + w \ , \end{aligned}$$

where $y = x_{n+1}$ and $\dot{y} = x_{n+2}$. More accurate descriptions of the seaway are also used. A typical description follows the so–called Pierson–Moskowitz wave spectrum given by

$$S(\omega) = \frac{a}{\omega^5} e^{-b/\omega^4} \ ,$$

where a, b are constants describing the particular seaway. Such a spectrum can be simulated by feeding a white noise signal into a suitable shaping filter. As an example, for a significant wave height (the average of the highest one third of all wave heights) of 7 m and a mean wave period of 9.4 seconds we have $a = 0.78$ and $b = 0.063$. Then the rational spectrum

$$S_R(\omega) = \frac{b_2^2\omega^2}{\omega^6 + (a_1^2 - 2a_2)\omega^4 + (a_2^2 - 2a_1a_3)\omega^2 + a_3^2} \ ,$$

with $a_1 = 0.5$, $a_2 = 0.33$, $a_3 = 0.07$, and $b_2 = 0.415$ can be used as an approximation of $S(\omega)$ for the chosen sea state. When both S and S_R are plotted versus ω the agreement is good. For details see the article “Control of yaw and roll by a rudder/fin stabilization system” by Kallstrom in the Proceedings of the Sixth Ship Control Systems Symposium, 1981. A stochastic process with spectral density given by S_R can be obtained as output from the filter

$$G(s) = \frac{b_2 s}{s^3 + a_1 s^2 + a_2 s + a_3} \ ,$$

with white noise as input. A similar model can be built for approximating the wave slope spectrum,

$$S_s(\omega) = \frac{\omega^4}{g^2} S(\omega) \ ,$$

and either one or both wave height and wave slope models can then be used for realistic design and simulations.

7.3 Kalman Filter

We present now the maximum likelihood, stochastic observer or filter for a nonstationary Gauss–Markov process. This will be seen to be completely parallel to the deterministic observer discussed in Section 3. We will sketch the derivation of the continuous time *Kalman* filter using calculus of variations in a manner which parallels our derivation of the optimal control law in 6.6.

Recall our classical full order observer design

$$\dot{} = A + Bu + L(y - C) \ .$$

In general, we would like to place the observer poles as negative as possible, this will create large elements of the observer gain matrix L. The larger the L, the faster the error in the observer dynamics will decay to zero. A very large L, however, will amplify undesirable noise which is always present in real systems. Therefore, there seems to be a limit on L which should depend on the level of noise in the system; this in turn should be directly related to the quality of our sensors and the disturbances. The Kalman filter *is* this best value for L and it provides an optimal stochastic observer, just like the linear quadratic regulator provided an optimal controller.

Consider the system

$$\dot{x} = Ax + Bu + \Gamma w \ ,$$

where w is a purely random process, and

$$E[x(t_0)] = \overline{x}_0 \ ,$$
$$E\left\{[x(t_0) - \overline{x}_0][x(t_0) - \overline{x}_0]^T\right\} = P_0 \ ,$$

which is the covariance of the error in the estimate of the state (t_0) at t_0. Initially we assume that $(t_0) = \overline{x}_0$: the most likely estimate at t_0 is the mean value at that time. In general,

$$P(t) = E\left\{[(t) - x(t)][(t) - x(t)]^T\right\} = E\left[(t)^T(t)\right] \ ,$$

where $\equiv -x$ is the error in the estimate of the state. The disturbance w in the state equations is a purely random process with

$$
\begin{aligned}
&E[w(t)] = 0\ , \quad \text{zero mean,}\\
&E[w(t)w^T(t)] = Q(t)\delta(t-\tau)\ ,
\end{aligned}
$$

where Q is the power spectral density matrix. We want to estimate the state vector (t) using a set of noisy measurements,

$$
y = Cx + v\ ,
$$

where the measurement noise v is another purely random process with

$$
\begin{aligned}
&E[v(t)] = 0\ , \quad \text{zero mean,}\\
&E[v(t)v^T(t)] = R(t)\delta(t-\tau)\ .
\end{aligned}
$$

What we want to do is to generate an estimate of both x and w which enter the state equations. This can be done in a least square sense if we minimize the cost function

$$
J = \tfrac{1}{2}\left[(x_0 - \overline{x}_0)^T P_0^{-1}(x_0 - \overline{x}_0)\right] + \tfrac{1}{2}\int_{t_0}^{t_f}\left[w^T Q^{-1} w + (y - Cx)^T R^{-1}(y - Cx)\right] dt\ .
$$

Observe that the first term minimizes the error in the initial estimate; the second term minimizes the error in the estimate of w; and the third term minimizes the error in the estimate of x. The minimization is subject to the constraints

$$
\begin{aligned}
\dot{x} &= Ax + Bu + \Gamma w\ ,\\
y &= Cx + v\ .
\end{aligned}
$$

Following a process similar to the LQR design, we can define the Hamiltonian

$$
H = \tfrac{1}{2}\left[w^T Q^{-1} w + (y - Cx)^T R^{-1}(y - Cx)\right] + \lambda^T(Ax + Bu + \Gamma w)\ ,
$$

and formulate the Euler–Lagrange equations, as before. We can find then that the optimal observer has the familiar form,

$$
\dot{} = A + Bu + L(y - C)\ , \quad (t_0) = \overline{x}_0\ ,
$$

where L is the Kalman filter gain matrix

$$
L = PC^T R^{-1}\ ,
$$

and P is the solution of the forward matrix Riccati differential equation

$$
\begin{aligned}
&\dot{P} = AP + PA^T + \Gamma Q \Gamma^T - PC^T R^{-1} CP\ ,\\
&P(t_0) = P_0\ .
\end{aligned}
$$

In the steady state case, these results become

$$
L = PC^T R^{-1}\ ,
$$

where now P is the solution to the algebraic Riccati equation

$$AP + PA^T + \Gamma Q \Gamma^T - PC^T R^{-1} CP = 0 \; .$$

The positive definite solution defines P, the covariance of the error in the estimate of the state .

As an example, consider the system

$$\begin{aligned} \dot{x} &= -2x + w \; , \quad \text{so } A = -2, \; \Gamma = 1, \\ y &= x + v \; , \quad \text{so } C = 1. \end{aligned}$$

The disturbance w is exponentially correlated with a correlation time

$$\tau_w = 0.01 \; ,$$

and root mean square value

$$\sigma_w = 1.2 \; .$$

The measurement noise v is also exponentially correlated with correlation time

$$\tau_v = 0.01 \; ,$$

but with an RMS value

$$\sigma_v = 0.2 \; .$$

We want to design a Kalman filter to produce a best estimate of x from y. The system has the time constant

$$T = 0.5 \gg 0.01 \; ,$$

so we can model both the disturbance and noise as white noise compared with the dynamics of the system. The power spectral densities are estimated as

$$\begin{aligned} \text{for } w \; &: \quad Q \approx 2\sigma_w^2 \tau_w = 2(1.2)^2 0.01 = 0.0288 \; , \\ \text{for } v \; &: \quad R \approx 2\sigma_v^2 \tau_v = 2(0.2)620.01 = 0.0008 \; . \end{aligned}$$

Our filter is given by

$$\dot{} = -2 + L(y-) \; , \quad L = PR^{-1} \; .$$

To find P we use the algebraic Riccati equation

$$\begin{aligned} & -2P + P(-2) + 0.0288 - P\frac{1}{0.0008}P = 0 \Rightarrow \\ & P^2 + 0.0032P - 0.00002304 = 0 \Rightarrow \\ & P = 0.00346 \; , \end{aligned}$$

the positive root. Then

$$L = PR^{-1} = \frac{0.00346}{0.0008} = 4.3246 \; ,$$

giving

$$\dot{} = -2 + 4.3246(y-) = -6.3246 + 4.3246y \ .$$

The error in the estimate produced by the filter is

$$\begin{aligned}
\dot{} &= \dot{} - \dot{x} \\
&= A + Bu + L(Cx + v - C) - Ax - Bu - \Gamma w \\
&= (A - LC) + Lv - \Gamma w \\
&= (-2 - 4.3246) + 4.3246v - w \\
&= -6.3246 + 4.3246v - w \ .
\end{aligned}$$

The eigenvalue of the filter is at -6.3246 which is well to the left of the system eigenvalue, -2, so the estimate will converge fast compared to the system.

7.4 The LQG Compensator

Recall that the separation principle allowed us to design the controller and the estimator separately and then use instead of x in the control law. The same principle states here that the optimal way to control a system

$$\dot{x} = Ax + Bu + \Gamma w \ ,$$

is to use a Kalman stochastic observer to estimate the state from the noisy measurements

$$y = Cx + v \ ,$$

and then use this estimate with the optimal deterministic linear controller we have already developed. The optimal controller can be derived from the LQR design, or we can use any kind of state feedback and feedforward we desire. The key is that we have no control over the poles of the observer here, nor can we choose the Q and R matrices that enter the Kalman filter design. These are set by the quality of our sensors and the level of the disturbances. After computing L from the Riccati equation, we should find the observer poles from the eigenvalues of $(A - LC)$ and make sure that they are more negative (the dominant pole) than the dominant poles of the controller. This can be done directly if we use poleplacement or indirectly by changing the weighting matrices in the LQR design. In case that the controller poles are not satisfactory, it is time to get better sensors!

The above combination of the optimal controller (LQR) and the optimal stochastic observer (Kalman filter) is called the Linear Quadratic Gaussian (LQG) compensator. This theoretical result produces a control system which is completely parallel to the deterministic observer and controller derived previously, except that now the controller and observer gain matrices are theoretically derived to yield optimal performance in the presence of stochastic disturbances w and measurement noise v.

Summarizing the total design problem, we have:

- State x,

$$\dot{x} = Ax + Bu + \Gamma w \ , \quad x(t_0) = x_0 \ ,$$

with

$$\begin{aligned} &E[ww^T] = Q\delta(t - \tau) \ , \quad \text{white noise} \ , \\ &E[w] = 0 \ , \\ &\text{covariance} \;\; X = \left[(x - \overline{x})(x - \overline{x})^T\right] \ . \end{aligned}$$

- Estimate ,

$$\dot{} = A + Bu + L(y - C) \ , \quad (t_0) = \overline{x}_0 \ ,$$

with covariance

$$\widehat{X} = \left[(-\overline{x})(-\overline{x})^T\right] \ .$$

- Error in estimate ,

$$= -x \ ,$$

with covariance

$$P = \left[(-x)(-x)^T\right] = E\left[{}^T\right] \ .$$

- Measurements y,

$$y = Cx + v \ ,$$

with

$$\begin{aligned} &E\left[vv^T\right] = R\delta(t - \tau) \ , \quad \text{white noise} \ , \\ &E[v] = 0 \ . \end{aligned}$$

- Controller,

$$u = -K \ ,$$

- Controller gain K,

$$K = R^{-1}B^TS \ ,$$

where

$$A^TS + SA - SBR^{-1}B^TS + Q = 0 \ ,$$

- Estimator gain L,

$$L = PC^TR^{-1} \ ,$$

where

$$AP + PA^T + \Gamma Q\Gamma^T - PC^TR^{-1}CP = 0 \ .$$

It should of course be emphasized that the matrices Q and R that enter the controller design are completely different than those in the observer design. The block diagram of the LQG design is shown in Figure 38.

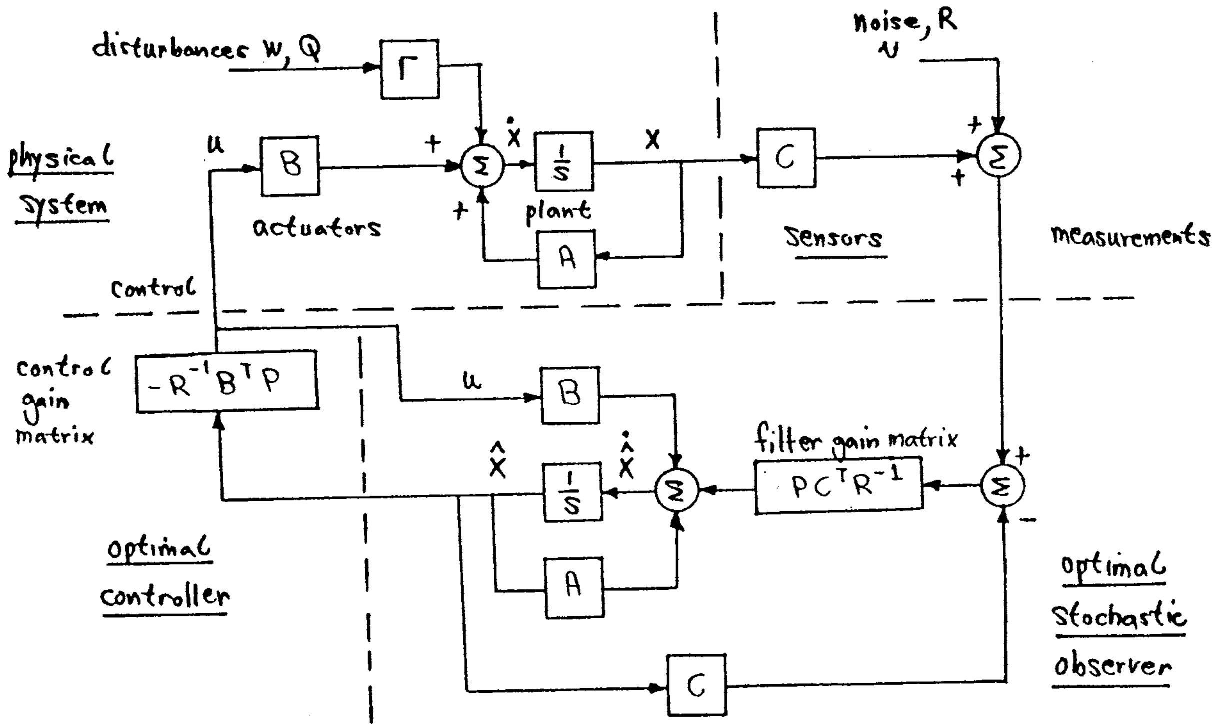

Figure 7: Compensator block diagram

7.5 Linear Quadratic Gaussian Compensator block diagram

With the optimal design developed above the performance can be evaluated either probabilistically or deterministically using computer simulation. Here we will develop the *root mean square* (RMS) response which can be easily computed for linear systems and can serve as a comparison index for different designs. If we wish to establish the response of the state about zero (the states are defined as deviations from nominal), we can begin with the filter response. Using the previous equations,

$$\begin{aligned} \dot{} &= (A - BK) + L[C(x -) + v] \ , \\ \dot{} &= (A - BK) - LC + Lv \ , \quad (t_0) = \overline{x}_0 = 0 \ . \end{aligned}$$

The dynamics in the error are governed by,

$$\begin{aligned} \dot{} &= \dot{} - \dot{x} \\ &= A + Bu + L(Cx + v - C) - Ax - Bu - \Gamma w \\ &= (A - LC) + Lv - \Gamma w \ , \end{aligned}$$

with

$$(t_0) = \overline{x}_0 - x(t_0) = -x(t_0) \ .$$

From the $\dot{}$ and $\dot{}$ equations we can see that is statistically independent of , so

$$E\left[(t_0)^T(t_0)\right] = 0 \ ,$$

and

$$E\left[(t)^T(t)\right] = 0 \ .$$

We can, therefore, establish the covariance of the state to be given by

$$\begin{aligned} X &= E\left[x(t)x^T(t)\right] \\ &= E\left[(-)(-)^T\right] \\ &= E\left[{}^T\right] - E\left[{}^T\right] - E\left[{}^T\right] - E\left[{}^T\right] \\ &= E\left[{}^T\right] - E\left[{}^T\right] \ . \end{aligned}$$

This gives

$$X(t) = \widehat{X}(t) + P(t) \ ,$$

or, at steady state,

$$X = \widehat{X} + P \ ,$$

which says that

$$\begin{aligned} (\text{covariance of state}) \;&=\; (\text{covariance of estimate of state}) \\ &\qquad + (\text{covariance of error in estimate of state}) \ . \end{aligned}$$

We already know how to obtain P and thus we need $\widehat{X}$ to obtain X, and the RMS response of the state x which is given by the square root of the diagonal terms in X. If we use the above equation in the definition of the covariance $\widehat{X}$ we can finally obtain the following differential equation for $\widehat{X}$,

$$\dot{\widehat{X}} = (A - BK)\widehat{X} + \widehat{X}(A - BK)^T + PC^T R^{-1} CP = 0 \; , \quad \widehat{X}(t_0) = 0 \; ,$$

which in the steady state yields the linear matrix equation,

$$(A - BK)\widehat{X} + \widehat{X}(A - BK)^T + PC^T R^{-1} CP = 0 \; ,$$

which can be solved for $\widehat{X}$ and then used in $X = \widehat{X} + P$ to obtain X.

The root mean square (RMS) use of the controls u can be derived directly from the definition of its covariance,

$$U \equiv E\left[uu^T\right] = E\left[(-K)(-K)^T\right] = KE\left[^T\right]K^T = K\widehat{X}K^T \; .$$

The square root of the associated diagonal elements of X and U give the RMS value of the states and controls, respectively, when the system is subjected to the disturbances w described by Q and the measurement noise described by R. The above equations are estimates of the RMS value of the response of a system and can be used for comparing different control and estimator designs. It should be borne in mind that they are not valid for nonlinear systems; they can not be used when the control effort saturates, for example. In these cases the associated RMS values of the variables of interest should be computed numerically by simulation.