LESSON 23

Linearization and Discretization of Nonlinear Systems

SUMMARY

Many real-world systems are continuous-time in nature and quite a few are also nonlinear. The purpose of this lesson is to explain how to linearize and discretize a nonlinear differential equation model. We do this so that we will be able to apply our digital estimators to the resulting discrete-time system.

A nonlinear dynamical system is linearized about *nominal values* of its state vector and control input. For example, $\mathbf{x}^*(t)$ denotes the nominal value of $\mathbf{x}(t)$; it must provide a good approximation to the actual behavior of the system. The approximation is considered good if the difference between the nominal and actual solutions can be described by a set of linear differential equations, called *linear perturbation equations*. A linear perturbation state-variable model, which consists of a perturbation state equation and a perturbation measurement equation, is obtained in this lesson.

When a continuous-time dynamical system is excited by white noise, its discretization must be done with care, because white noise is not piecewise continuous. This means that we cannot assume a constant value for a continuous white noise process even in a very small time interval. We discretize such a system by first expressing its solution in closed form, then sampling the solution, and finally introducing a discrete-time disturbance that is statistically equivalent (through its first two moments) to an integrated white noise term that appears in the sampled solution. Tremendous simplifications of the discretized model occur if the plant matrices are approximately constant during a sampling interval.

When you complete this lesson, you will be able to linearize and discretize continuous-time nonlinear dynamical systems.

Many real-world systems are continuous-time in nature and quite a few are also nonlinear. For example, the state equations associated with the motion of a satellite of mass m about a spherical planet of mass M, in a planet-centered coordinate system, are nonlinear, because the planet's force field obeys an inverse square law. Figure 23-1 depicts a situation where the measurement equation is nonlinear. The measurement is angle ϕ_i and is expressed in a rectangular coordinate system; i.e., $\phi_i = \tan^{-1}[y/(x - l_i)]$. Sometimes the state equation may be nonlinear and the measurement equation linear, or vice versa, or they may both be nonlinear. Occasionally, the coordinate system in which we choose to work causes the two former situations. For example, equations of motion in a polar coordinate system are nonlinear, whereas the measurement equations are linear. In a polar coordinate system, where ϕ is a state variable, the measurement equation for the situation depicted in Figure 23-1 is $z_i = \phi_i$, which is linear. In a rectangular coordinate system, on the other hand, equations of motion are linear, but the measurement equations are nonlinear.

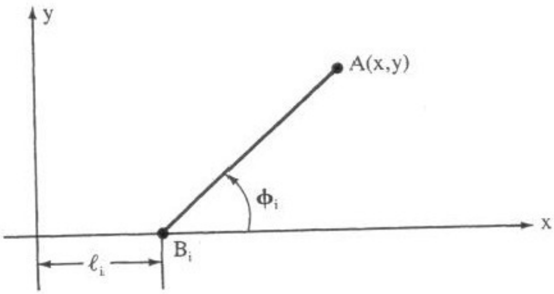

Figure 23-1 Coordinate system for an angular measurement between two objects A and B.

Finally, we may begin with a linear system that contains some unknown parameters. When these parameters are modeled as first-order Markov sequences, and these models are augmented to the original system, the augmented model is nonlinear, because the parameters that appeared in the original "linear" model are treated as states. We shall describe this situation in much more detail in Lesson 24.

The purpose of this lesson is to explain how to linearize and discretize a nonlinear differential equation model. We do this so that we will be able to apply our digital estimators to the resulting discrete-time system.

A DYNAMICAL MODEL

The starting point for this lesson is the nonlinear state-variable model

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] + \mathbf{G}(t)\mathbf{w}(t) \tag{23-1}$$

$$\mathbf{z}(t) = \mathbf{h}[\mathbf{x}(t), \mathbf{u}(t), t] + \mathbf{v}(t) \tag{23-2}$$

We shall assume that measurements are only available at specific values of time, at $t = t_i, i = 1, 2, \ldots$; thus, our measurement equation will be treated as a discrete-time equation, whereas our state equation will be treated as a continuous-time equation. State vector $\mathbf{x}(t)$ is $n \times 1$; $\mathbf{u}(t)$ is an $l \times 1$ vector of known inputs; measurement vector $\mathbf{z}(t)$ is $m \times 1$; $\dot{\mathbf{x}}(t)$ is short for $d\mathbf{x}(t)/dt$; nonlinear functions $\mathbf{f}$ and $\mathbf{h}$ may depend both implicitly and explicitly on t, and we assume that both $\mathbf{f}$ and $\mathbf{h}$ are continuous and continuously differentiable with respect to all the elements of $\mathbf{x}$ and $\mathbf{u}$; $\mathbf{w}(t)$ is a continuous-time white noise process, i.e., $\mathbf{E}\{\mathbf{w}(t)\} = \mathbf{0}$, and

$$\mathbf{E}\{\mathbf{w}(t)\mathbf{w}'(\tau)\} = \mathbf{Q}(t)\delta(t-\tau) \tag{23-3}$$

$\mathbf{v}(t_i)$ is a discrete-time white noise sequence, i.e., $\mathbf{E}\{\mathbf{v}(t_i)\} = \mathbf{0}$ for $t = t_i, i = 1, 2, \ldots,$ and

$$\mathbf{E}\{\mathbf{v}(t_i)\mathbf{v}'(t_j)\} = \mathbf{R}(t_i)\delta_{ij} \tag{23-4}$$

And $\mathbf{w}(t)$ and $\mathbf{v}(t_j)$ are mutually uncorrelated at all $t = t_i$; i.e.,

$$\mathbf{E}\{\mathbf{w}(t)\mathbf{v}'(t_i)\} = \mathbf{0}, \qquad \text{for } t = t_i \quad i = 1, 2, \ldots \tag{23-5}$$

Note that, whereas it is indeed correct to refer to $\mathbf{R}(t_i)$ as a covariance matrix, it is incorrect to refer to $\mathbf{Q}(t)$ as a covariance matrix. It is $\mathbf{Q}(t)\delta(t-\tau)$ that is the covariance matrix. In the special case when $\mathbf{w}(t)$ is stationary so that $\mathbf{Q}(t) = \mathbf{Q}$, then $\mathbf{Q}$ is a *spectral intensity matrix*, because the Fourier transform of $\mathbf{E}\{\mathbf{w}(t)\mathbf{w}'(\tau)\}$, which is the power spectrum of $\mathbf{w}(t)$, equals $\mathbf{Q}$. In general, matrix $\mathbf{Q}(t)$ can be referred to as an *intensity matrix*.

EXAMPLE 23-1

Here we expand upon the previously mentioned satellite – planet example. Our example is taken from Meditch (1969, pp. 60 – 61), who states, "Assuming that the planet's force field obeys an inverse square law, and that the only other forces present are the satellite's two thrust forces $u_r(t)$ and $u_\theta(t)$ (see Figure 23-2), and that the satellite's initial position and velocity vectors lie in the plane, we know from elementary particle mechanics that the satellite's motion is confined to the plane and is governed by the two equations

$$\ddot{r} = r\dot{\theta}^2 - \frac{\gamma}{r^2} + \frac{1}{m}u_r(t) \tag{23-6}$$

and

$$\ddot{\theta} = -\frac{2\dot{r}\dot{\theta}}{r} + \frac{1}{m}u_\theta(t) \tag{23-7}$$

where $\gamma = GM$ and G is the universal gravitational constant.

"Defining $x_1 = r, x_2 = \dot{r}, x_3 = \theta, x_4 = \dot{\theta}, u_1 = u_r$, and $u_2 = u_\theta$, we have

$$\left.\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_1 x_4^2 - \frac{\gamma}{x_1^2} + \frac{1}{m}u_1(t) \\ \dot{x}_3 &= x_4 \\ \dot{x}_4 &= -\frac{2x_2x_4}{x_1} + \frac{1}{m}u_2(t) \end{aligned}\right\} \tag{23-8}$$

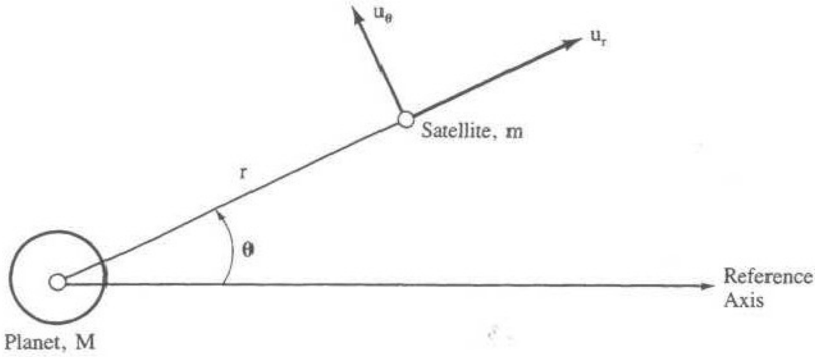

Figure 23-2 Schematic for satellite–planet system (Copyright 1969, McGraw-Hill).

which is of the form in (23-1). . . . Assuming . . . that the measurement made on the satellite during its motion is simply its distance from the surface of the planet, we have the scalar measurement equation

$$z(t) = r(t) - r_0 + \nu(t) = x_1(t) - r_0 + \nu(t) \tag{23-9}$$

where r_0 is the planet's radius."

Comparing (23-8) and (23-1), and (23-9) and (23-2), we conclude that

$$\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] = \text{col}\left[x_2, x_1 x_4^2 - \frac{\gamma}{x_1^2} + \frac{1}{m}u_1, x_4, -\frac{2x_2x_4}{x_1} + \frac{1}{m}u_2\right] \tag{23-10}$$

and

$$\mathbf{h}[\mathbf{x}(t), \mathbf{u}(t), t] = x_1 - r_0 \tag{23-11}$$

Observe that, in this example, only the state equation is nonlinear. □

PERTURBATION EQUATIONS

In this section we shall linearize our nonlinear dynamical model in (23-1) and (23-2) about nominal values of $\mathbf{x}(t)$ and $\mathbf{u}(t)$, $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$, respectively. If we are given a nominal input, $\mathbf{u}^*(t)$, then $\mathbf{x}^*(t)$ satisfies the nonlinear differential equation

$$\dot{\mathbf{x}}^*(t) = \mathbf{f}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] \tag{23-12}$$

and associated with $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$ is the following nominal measurement, $\mathbf{z}^*(t)$, where

$$\mathbf{z}^*(t) = \mathbf{h}[\mathbf{x}^*(t), \mathbf{u}^*(t)], \qquad t = t_i, \quad i = 1, 2, \ldots \tag{23-13}$$

If $\mathbf{u}(t)$ is an input derived from a feedback control law, so that $\mathbf{u}(t) = \mathbf{u}[\mathbf{x}(t), t]$, then $\mathbf{u}(t)$ can differ from $\mathbf{u}^*(t)$, because $\mathbf{x}(t)$ will differ from $\mathbf{x}^*(t)$. On the other hand, if $\mathbf{u}(t)$ does not depend on $\mathbf{x}(t)$ then usually $\mathbf{u}(t)$ is the same as $\mathbf{u}^*(t)$, in which case $\delta\mathbf{u}(t) = \mathbf{0}$. Throughout this lesson, we shall assume that $\delta\mathbf{u}(t) \neq \mathbf{0}$ and that $\mathbf{x}^*(t)$ exists. We discuss two methods for choosing $\mathbf{x}^*(t)$ in Lesson 24. Obviously, one is just to solve (23-12) for $\mathbf{x}^*(t)$.

Note that $\mathbf{x}^*(t)$ must provide a good approximation to the actual behavior of the system. The approximation is considered good if the difference between the nominal and actual solutions can be described by a system of linear differential equations, called *linear perturbation equations*. We derive these equations next. Let

$$\delta\mathbf{x}(t) = \mathbf{x}(t) - \mathbf{x}^*(t) \tag{23-14}$$

and

$$\delta\mathbf{u}(t) = \mathbf{u}(t) - \mathbf{u}^*(t) \tag{23-15}$$

then

$$\frac{d}{dt}\delta\mathbf{x}(t) = \delta\dot{\mathbf{x}}(t) = \dot{\mathbf{x}}(t) - \dot{\mathbf{x}}^*(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] + \mathbf{G}(t)\mathbf{w}(t) - \mathbf{f}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] \tag{23-16}$$

Fact 1. *When* $\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$ *is expanded in a Taylor series about* $\mathbf{x}^*(t)$ *and* $\mathbf{u}^*(t)$*, we obtain*

$$\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] = \mathbf{f}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] + \mathbf{F}_\mathbf{x}[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{x}(t) + \mathbf{F}_\mathbf{u}[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{u}(t) + \text{higher-order terms} \tag{23-17}$$

where $\mathbf{F}_\mathbf{x}$ *and* $\mathbf{F}_\mathbf{u}$ *are* $n \times n$ *and* $n \times l$ *Jacobian matrices; i.e.,*

$$\mathbf{F}_\mathbf{x}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \begin{pmatrix} \partial f_1/\partial x_1^* & \cdots & \partial f_1/\partial x_n^* \\ \vdots & \ddots & \vdots \\ \partial f_n/\partial x_1^* & \cdots & \partial f_n/\partial x_n^* \end{pmatrix} \tag{23-18}$$

and

$$\mathbf{F}_\mathbf{x}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \begin{pmatrix} \partial f_1/\partial u_1^* & \cdots & \partial f_1/\partial u_l^* \\ \vdots & \ddots & \vdots \\ \partial f_n/\partial u_1^* & \cdots & \partial f_n/\partial u_l^* \end{pmatrix} \tag{23-19}$$

In these expressions $\partial f_i/\partial x_j^*$ *and* $\partial f_i/\partial u_j^*$ *are short for*

$$\frac{\partial f_i}{\partial x_j^*} = \left.\frac{\partial f_i[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial x_j(t)}\right|_{\mathbf{x}(t)=\mathbf{x}^*(t),\, \mathbf{u}(t)=\mathbf{u}^*(t)} \tag{23-20}$$

and

$$\frac{\partial f_i}{\partial u_j^*} = \left.\frac{\partial f_i[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial u_j(t)}\right|_{\mathbf{x}(t)=\mathbf{x}^*(t),\, \mathbf{u}(t)=\mathbf{u}^*(t)} \tag{23-21}$$

Proof. The Taylor series expansion of the ith component of $\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$ is

$$f_i[\mathbf{x}(t), \mathbf{u}(t), t] = f_i[\mathbf{x}^*(t), \mathbf{u}^*(t), t] + \frac{\partial f_1}{\partial x_1^*}[x_1(t) - x_1^*(t)] + \cdots$$
$$+ \frac{\partial f_i}{\partial x_n^*}[x_n(t) - x_n^*(t)] + \frac{\partial f_i}{\partial u_1^*}[u_1(t) - u_1^*(t)] + \cdots \tag{23-22}$$
$$+ \frac{\partial f_i}{\partial u_l^*}[u_l(t) - u_l^*(t)] + \text{higher-order terms}$$

where $i = 1, 2, \ldots, n$. Collecting these n equations together in vector-matrix format, we obtain (23-17), in which $\mathbf{F_x}$ and $\mathbf{F_u}$ are defined in (23-18) and (23-19), respectively. □

Substituting (23-17) into (23-16) and neglecting the "higher-order terms", we obtain the following *perturbation state equation*:

$$\delta\dot{\mathbf{x}}(t) = \mathbf{F_x}[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{x}(t)$$
$$+ \mathbf{F_u}[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{u}(t) + \mathbf{G}(t)\mathbf{w}(t) \tag{23-23}$$

Observe that, even if our original nonlinear differential equation is not an explicit function of time {i.e., $\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t)]$}, our perturbation state equation is always time-varying because Jacobian matrices $\mathbf{F_x}$ and $\mathbf{F_u}$ vary with time, because $\mathbf{x}^*$ and $\mathbf{u}^*$ vary with time.

Next, let

$$\delta\mathbf{z}(t) = \mathbf{z}(t) - \mathbf{z}^*(t) \tag{23-24}$$

Fact 2. *When* $\mathbf{h}[\mathbf{x}(t), \mathbf{u}(t), t]$ *is expanded in a Taylor series about* $\mathbf{x}^*(t)$ *and* $\mathbf{u}^*(t)$, *we obtain*

$$\mathbf{h}[\mathbf{x}(t), \mathbf{u}(t), t] = \mathbf{h}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] + \mathbf{H_x}[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{x}(t)$$
$$+ \mathbf{H_u}[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{u}(t) + \text{higher-order terms} \tag{23-25}$$

where $\mathbf{H_x}$ *and* $\mathbf{H_u}$ *are* $m \times n$ *and* $m \times l$ *Jacobian matrices; i.e.,*

$$\mathbf{H_x}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \begin{pmatrix} \partial h_1/\partial x_1^* & \cdots & \partial h_1/\partial x_n^* \\ \vdots & \ddots & \vdots \\ \partial h_m/\partial x_1^* & \cdots & \partial h_m/\partial x_n^* \end{pmatrix} \tag{23-26}$$

and

$$\mathbf{H_u}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \begin{pmatrix} \partial h_1/\partial u_1^* & \cdots & \partial h_1/\partial u_l^* \\ \vdots & \ddots & \vdots \\ \partial h_m/\partial u_1^* & \cdots & \partial h_m/\partial u_l^* \end{pmatrix} \tag{23-27}$$

In these expressions $\partial h_i/\partial x_j^*$ *and* $\partial h_i/\partial u_j^*$ *are short for*

$$\frac{\partial h_i}{\partial x_j^*} = \left.\frac{\partial h_i[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial x_j(t)}\right|_{\mathbf{x}(t)=\mathbf{x}^*(t), \mathbf{u}(t)=\mathbf{u}^*(t)} \tag{23-28}$$

and

$$\frac{\partial h_i}{\partial u_j^*} = \left.\frac{\partial h_i[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial u_j(t)}\right|_{\mathbf{x}(t)=\mathbf{x}^*(t), \mathbf{u}(t)=\mathbf{u}^*(t)} \quad \square \tag{23-29}$$

We leave the derivation of this fact to the reader, because it is analogous to the derivation of the Taylor series expansion of $\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$.

Substituting (23-25) into (23-24) and neglecting the "higher-order terms," we obtain the following *perturbation measurement equation:*

$$\begin{aligned}\delta\mathbf{z}(t) &= \mathbf{H}_x[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{x}(t) \\ &\quad + \mathbf{H}_\mathbf{u}[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{u}(t) + \mathbf{v}(t), \qquad t = t_i, \quad i = 1, 2, \ldots\end{aligned} \tag{23-30}$$

Equations (23-23) and (23-30) constitute our linear perturbation equations, or our linear *perturbation state-variable model.*

An interesting article on the linearization of an equation very similar to (23-1), in which the derivatives needed in (23-18), (23-19), (23-26), and (23-27) are computed by numerical differentiation, is Taylor and Antoniotti (1993).

EXAMPLE 23-2

Returning to our satellite–planet Example 23-1, we find that

$$\mathbf{F}_\mathbf{x}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \mathbf{F}_\mathbf{x}[\mathbf{x}^*(t)] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ x_4^2(t) + \dfrac{2\gamma}{x_1^3(t)} & 0 & 0 & 2x_1(t)x_4(t) \\ 0 & 0 & 0 & 1 \\ \dfrac{2x_2(t)x_4(t)}{x_1^2(t)} & \dfrac{-2x_4(t)}{x_1(t)} & 0 & \dfrac{-2x_2(t)}{x_1(t)} \end{pmatrix}_*$$

$$\mathbf{F}_\mathbf{u}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \mathbf{F}_\mathbf{u} = \begin{pmatrix} 0 & 0 \\ \dfrac{1}{m} & 0 \\ 0 & 0 \\ 0 & \dfrac{1}{m} \end{pmatrix}$$

$$\mathbf{H}_\mathbf{x}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \mathbf{H}_\mathbf{x} = (1 \quad 0 \quad 0 \quad 0)$$

and

$$\mathbf{H}_\mathbf{u}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \mathbf{0}$$

In the equation for $\mathbf{F}_\mathbf{x}[\mathbf{x}^*(t)]$, the notation $(\quad)_*$ means that all $x_i(t)$ within the matrix are nominal values, i.e., $x_i(t) = x_i^*(t)$.

Observe that the linearized satellite–planet system is time varying, because its linearized planet matrix, $\mathbf{F}_\mathbf{x}[\mathbf{x}^*(t)]$, depends on the nominal trajectory $\mathbf{x}^*(t)$. $\square$

DISCRETIZATION OF A LINEAR TIME-VARYING STATE-VARIABLE MODEL

In this section we describe how we discretize the general linear, time-varying state-variable model

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{G}(t)\mathbf{w}(t) \tag{23-31}$$

$$\mathbf{z}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{v}(t), \qquad t = t_i \quad i = 1, 2, \ldots \tag{23-32}$$

The application of this section's results to the perturbation state-variable model is given in the following section.

In (23-31) and (23-32), $\mathbf{x}(t)$ is $n \times 1$, control input $\mathbf{u}(t)$ is $l \times 1$, process noise $\mathbf{w}(t)$ is $p \times 1$, and $\mathbf{z}(t)$ and $\mathbf{v}(t)$ are each $m \times 1$. Additionally, $\mathbf{w}(t)$ is a continuous-time white noise process, $\mathbf{v}(t_i)$ is a discrete-time white noise sequence, and $\mathbf{w}(t)$ and $\mathbf{v}(t_i)$ are mutually uncorrelated at all $t = t_i, i = 1, 2, \ldots$; i.e., $\mathbf{E}\{\mathbf{w}(t)\} = \mathbf{0}$ for all t, $\mathbf{E}\{\mathbf{v}(t_i)\} = \mathbf{0}$ for all t_i, and

$$\mathbf{E}\{\mathbf{w}(t)\mathbf{w}'(\tau)\} = \mathbf{Q}(t)\delta(t-\tau) \tag{23-33}$$

$$\mathbf{E}\{\mathbf{v}(t_i)\mathbf{v}'(t_j)\} = \mathbf{R}(t_i)\delta_{ij} \tag{23-34}$$

and

$$\mathbf{E}\{\mathbf{w}(t)\mathbf{v}'(t_i)\} = \mathbf{0}, \qquad \text{for } t = t_i, \quad i = 1, 2, \ldots \tag{23-35}$$

It is tempting to discretize (23-31) by setting $d\mathbf{x}(t) = [\mathbf{x}(t_{k+1}) - \mathbf{x}(t_k)]/\Delta t$, and setting $t = t_k$ on the right-hand side of (23-31). Whereas $\mathbf{u}(t_k)$ is a well-defined function at all time points, including t_k, $\mathbf{w}(t_k)$ is not, because $\mathbf{w}(t_k)$ is a continuous-time white noise process, which, in general is a mathematical fiction. See Lesson 26 for additional discussions on continuous-time white noise processes.

Consequently, our approach to discretizing state equation (23-31) begins with the solution of that equation.

Theorem 23-1. *The solution to state equation (23-31) can be expressed as*

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t, \tau)[\mathbf{B}(\tau)\mathbf{u}(\tau) + \mathbf{G}(\tau)\mathbf{w}(\tau)]d\tau \tag{23-36}$$

where state transition matrix $\mathbf{\Phi}(t, \tau)$ is the solution to the following matrix homogeneous differential equation:

$$\left.\begin{aligned} \dot{\mathbf{\Phi}}(t, \tau) &= \mathbf{F}(t)\mathbf{\Phi}(t, \tau) \\ \mathbf{\Phi}(t, t) &= \mathbf{I} \end{aligned}\right\} \quad \square \tag{23-37}$$

This result should be a familiar one to many readers of this book. See the Supplementary Material at the end of this lesson for its proof.

Next, we assume that $\mathbf{u}(t)$ is a piecewise constant function of time for $t \in [t_k, t_{k+1}]$ and set $t_0 = t_k$ and $t = t_{k+1}$ in (23-36), to obtain

$$\mathbf{x}(t_{k+1}) = \mathbf{\Phi}(t_{k+1}, t_k)\mathbf{x}(t_k) + \left[\int_{t_k} \mathbf{\Phi}(t_{k+1}, \tau)\mathbf{B}(\tau)d\tau\right]\mathbf{u}(t_k)$$

$$+ \int_{t_k}^{t_{k+1}} \mathbf{\Phi}(t_{k+1}, \tau)\mathbf{G}(\tau)\mathbf{w}(\tau)d\tau \qquad (23\text{-}38)$$

which can also be written as

$$\mathbf{x}(k+1) = \mathbf{\Phi}(k+1, k)\mathbf{x}(k) + \mathbf{\Psi}(k+1, k)\mathbf{u}(k) + \mathbf{w}_d(k) \qquad (23\text{-}39)$$

where

$$\mathbf{\Phi}(k+1, k) = \mathbf{\Phi}(t_{k+1}, t_k) \qquad (23\text{-}40)$$

$$\mathbf{\Psi}(k+1, k) = \int_{t_k}^{t_{k+1}} \mathbf{\Phi}(t_{k+1}, \tau)\mathbf{B}(\tau)d\tau \qquad (23\text{-}41)$$

and $\mathbf{w}_d(k)$ is a discrete-time white Gaussian sequence that is *statistically equivalent through its first two moments* to

$$\int_{t_k}^{t_{k+1}} \mathbf{\Phi}(t_{k+1}, \tau)\mathbf{G}(\tau)\mathbf{w}(\tau)d\tau$$

Observe that the dependence of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ on the two time points t_k and t_{k+1} rationalizes our use of double arguments for them, which is why we have used the double-argument notation for $\mathbf{\Phi}$ and $\mathbf{\Psi}$ in our basic state-variable model in Lesson 15. The mean and covariance matrices of $\mathbf{w}_d(k)$ are

$$\mathbf{E}\{\mathbf{w}_d(k)\} = \mathbf{E}\left\{\int_{t_k}^{t_{k+1}} \mathbf{\Phi}(t_{k+1}, \tau)\mathbf{G}(\tau)\mathbf{w}(\tau)d\tau\right\} = \mathbf{0} \qquad (23\text{-}42)$$

and

$$\begin{aligned}
\mathbf{E}\{\mathbf{w}_d(k)\mathbf{w}_d'(k)\} &\triangleq \mathbf{Q}_d(k+1, k) \\
&= \mathbf{E}\left\{\int_{t_k}^{t_{k+1}} \mathbf{\Phi}(t_{k+1}, \tau)\mathbf{G}(\tau)\mathbf{w}(\tau)d\tau \right. \\
&\qquad \left. \int_{t_k}^{t_{k+1}} \mathbf{w}'(\xi)\mathbf{G}'(\xi)\mathbf{\Phi}'(t_{k+1}, \xi)d\xi\right\} \\
&= \int_{t_k}^{t_{k+1}} \mathbf{\Phi}(t_{k+1}, \tau)\mathbf{G}(\tau)\mathbf{Q}(\tau)\mathbf{G}'(\tau)\mathbf{\Phi}'(t_{k+1}, \tau)d\tau \qquad (23\text{-}43)
\end{aligned}$$

respectively.

Observe, from the right-hand side of Equations (23-40), (23-41), and (23-43), that these quantities can be computed from knowledge about $\mathbf{F}(t)$, $\mathbf{B}(t)$, $\mathbf{G}(t)$, and $\mathbf{Q}(t)$. In general, we must compute $\mathbf{\Phi}(k+1, k)$, $\mathbf{\Psi}(k+1, k)$, and $\mathbf{Q}_d(k+1, k)$ using numerical integration, and these matrices change from one time interval to the next because $\mathbf{F}(t)$, $\mathbf{B}(t)$, $\mathbf{G}(t)$, and $\mathbf{Q}(t)$ usually change from one time interval to the next.

Because our measurements have been assumed to be available only at sampled values of t at $t = t_i$, $i = 1, 2, \ldots$, we can express (23-32) as

$$\mathbf{z}(k+1) = \mathbf{H}(k+1)\mathbf{x}(k+1) + \mathbf{v}(k+1) \tag{23-44}$$

Equations (23-39) and (23-44) constitute our discretized state-variable model.

EXAMPLE 23-3

Great simplifications of the calculations in (23-40), (23-41) and (23-43) occur if $\mathbf{F}(t)$, $\mathbf{B}(t)$, $\mathbf{G}(t)$, and $\mathbf{Q}(t)$ are approximately constant during the time interval $[t_k, t_{k+1}]$, i.e., if

$$\left.\begin{array}{llll} \mathbf{F}(t) \simeq \mathbf{F}_k, & \mathbf{B}(t) \simeq \mathbf{B}_k & \mathbf{G}(t) \simeq \mathbf{G}_k, & \text{and} \\ \mathbf{Q}(t) \simeq \mathbf{Q}_k, & \text{for } t \in [t_k, t_{k+1}] & & \end{array}\right\} \tag{23-45}$$

To begin, (23-37) is easily integrated to yield

$$\mathbf{\Phi}(t, \tau) = e^{\mathbf{F}_k(t-\tau)} \tag{23-46}$$

Hence,

$$\mathbf{\Phi}(k+1, k) = e^{\mathbf{F}_k T} = \mathbf{\Phi}(k) \tag{23-47}$$

where we have assumed that $t_{k+1} - t_k = T$. The matrix exponential is given by the infinite series

$$e^{\mathbf{F}_k T} = \mathbf{I} + \mathbf{F}_k T + \mathbf{F}_k^2 \frac{T^2}{2} + \mathbf{F}_k^3 \frac{T^3}{3!} + \cdots \tag{23-48}$$

and, for sufficiently small values of T,

$$e^{\mathbf{F}_k T} \simeq \mathbf{I} + \mathbf{F}_k T \tag{23-49}$$

We use this approximation for $e^{\mathbf{F}_k T}$ in deriving simpler expressions for $\mathbf{\Psi}(k+1, k)$ and $\mathbf{Q}_d(k+1, k)$. Comparable results can be obtained for higher-order truncations of $e^{\mathbf{F}_k T}$.

Substituting (23-46) into (23-41), we find that

$$\begin{aligned} \mathbf{\Psi}(k+1, k) &= \int_{t_k}^{t_{k+1}} \mathbf{\Phi}(t_{k+1}, \tau)\mathbf{B}_k d\tau = \int_{t_k}^{t_{k+1}} e^{\mathbf{F}_k(t_{k+1}-\tau)}\mathbf{B}_k d\tau \\ &\simeq \int_{t_k}^{t_{k+1}} [\mathbf{I} + \mathbf{F}_k(t_{k+1} - \tau)]\mathbf{B}_k d\tau \\ &\simeq \mathbf{B}_k T + \mathbf{F}_k \mathbf{B}_k t_{k+1} T - \mathbf{F}_k \mathbf{B}_k \int_{t_k}^{t_{k+1}} \tau d\tau \\ &\simeq \mathbf{B}_k T + \mathbf{F}_k \mathbf{B}_k \frac{T^2}{2} \simeq \mathbf{B}_k T = \mathbf{\Psi}(k) \end{aligned} \tag{23-50}$$

where we have truncated $\mathbf{\Psi}(k+1, k)$ to its first-order term in T. Proceeding in a similar manner for $\mathbf{Q}_d(k+1, k)$, it is straightforward to show that

$$\mathbf{Q}_d(k+1, k) \simeq \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k' T = \mathbf{Q}_d(k) \tag{23-51}$$

Note that (23-47), (23-49), (23-50), and (23-51), while much simpler than their original expressions, can still change in values from one time interval to another because of their dependence on k. Note, also, that under the piecewise constant condition the double arguments for $\mathbf{\Psi}$ and $\mathbf{Q}_d$ are not needed. □

Applying the results of the preceding section to the perturbation state-variable model in (23-23) and (23-30), we obtain the following *discretized perturbation state-variable model*:

$$\delta\mathbf{x}(k+1) = \mathbf{\Phi}(k+1,k;^*)\delta\mathbf{x}(k) + \mathbf{\Psi}(k+1,k;^*)\delta\mathbf{u}(k) + \mathbf{w}_d(k) \tag{23-52}$$

$$\delta\mathbf{z}(k+1) = \mathbf{H_x}(k+1;^*)\delta\mathbf{x}(k+1) + \mathbf{H}_u(k+1;^*)\delta\mathbf{u}(k+1) + \mathbf{v}(k+1) \tag{23-53}$$

The notation $\mathbf{\Phi}(k+1,k;^*)$, for example, denotes the fact that this matrix depends on $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$. More specifically,

$$\mathbf{\Phi}(k+1,k;^*) = \mathbf{\Phi}(t_{k+1},t_k;^*) \tag{23-54}$$

where

$$\begin{aligned}\dot{\mathbf{\Phi}}(t,\tau;^*) &= \mathbf{F_x}[\mathbf{x}^*(t),\mathbf{u}^*(t),t]\mathbf{\Phi}(t,\tau;^*)\\ \mathbf{\Phi}(t,t;^*) &= \mathbf{I}\end{aligned} \tag{23-55}$$

Additionally,

$$\mathbf{\Psi}(k+1,k;^*) = \int_{t_k}^{t_{k+1}} \mathbf{\Phi}(t_{k+1},\tau;^*)\mathbf{F_u}[\mathbf{x}^*(\tau),\mathbf{u}^*(\tau),\tau]d\tau \tag{23-56}$$

and

$$\mathbf{Q}_d(k+1,k;^*) = \int_{t_k}^{t_{k+1}} \mathbf{\Phi}(t_{k+1},\tau;^*)\mathbf{G}(\tau)\mathbf{Q}(\tau)\mathbf{G}'(\tau)\mathbf{\Phi}'(t_{k+1},\tau;^*)d\tau \tag{23-57}$$

COMPUTATION

To simulate the nonlinear dynamical system, given by (23-1) and (23-2), use Simulink®.

Linearization can be accomplished within SIMULINK.

There are no M-files available to **discretize** a time-varying differential equation model, such as the perturbation state equation in (23-31). When the matrices in the time-varying differential equation model are piecewise constant over the sampling interval, as described in Example 23-3, then discretization can be accomplished with the aid of the matrix exponential, an M-file for which can be found in MATLAB:

expm: Matrix exponential: $\mathbf{y} = \text{expm}(\mathbf{x})$

You will need to write an M-file to produce Equations (23-47), (23-41), and (23-43). You will then be able to implement the discrete-time state-equation in (23-39). Because we have assumed that measurements are given in sampled form, no discretization is needed in order to implement the discrete-time measurement equation in (23-44). If, however, your actual measurements are analog in nature, you must first digitize them (see Lesson 26 for a discussion on how to relate the statistics of continuous-time and discrete-time white noise).

Supplementary Material

PROOF OF THEOREM 23-1

Theorem 23-1 provides the solution to a time-varying state equation in terms of the state transition matrix $\mathbf{\Phi}(t, \tau)$. Our derivation of (23-36) for $\mathbf{x}(t)$ proceeds in three (classical) steps: (1) obtain the homogeneous solution to vector differential equation (23-31); (2) obtain the particular solution to (23-31); and (3) obtain the complete solution to (23-31).

Homogeneous Solution to (23-31)

The homogeneous equation corresponding to (23-31) is

$$\frac{d\mathbf{x}_h(t)}{dt} = \mathbf{F}(t)\mathbf{x}_h(t), \qquad \text{for } t \geq t_0 \text{ and arbitrary } \mathbf{x}_0 \tag{23-58}$$

Let

$$\mathbf{x}_h(t) = \mathbf{X}(t)\mathbf{x}(t_0) \tag{23-59}$$

where $\mathbf{X}(t)$ is an arbitrary unknown $n \times n$ matrix. Then $d\mathbf{x}_h(t)/dt = [d\mathbf{X}(t)/dt]\mathbf{x}(t_0)$, so that $[d\mathbf{X}(t)/dt]\mathbf{x}(t_0) = \mathbf{F}(t)\mathbf{X}(t)\mathbf{x}(t_0)$, from which it follows that

$$\frac{d\mathbf{X}(t)}{dt} = \mathbf{F}(t)\mathbf{X}(t) \tag{23-60}$$

Letting $\mathbf{x}_h(t_0) = \mathbf{x}(t_0)$, we find, from (23-59), that $\mathbf{X}(t_0) = \mathbf{I}$, which is the initial condition for (23-60).

$\mathbf{X}(t)$ is called the *fundamental matrix* of the system in (23-31); it depends only on $\mathbf{F}(t)$.

Particular Solution to (23-31)

We use the *Lagrange variation of parameter technique* as follows. Let

$$\mathbf{x}_p(t) = \mathbf{X}(t)\mathbf{y}(t) \tag{23-61}$$

where $\mathbf{y}(t)$ is an unknown $n \times 1$ vector. Differentiate $\mathbf{x}_p(t)$ and substitute the result, as well as (23-61), into (23-31) to see that

$$\frac{d\mathbf{X}(t)}{dt}\mathbf{y}(t) + \mathbf{X}(t)\frac{d\mathbf{y}(t)}{dt} = \mathbf{F}(t)\mathbf{X}(t)\mathbf{y}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{G}(t)\mathbf{w}(t)$$

But $d\mathbf{X}(t)/dt = \mathbf{F}(t)\mathbf{X}(t)$, so this last equation reduces to

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{X}^{-1}(t)[\mathbf{B}(t)\mathbf{u}(t) + \mathbf{G}(t)\mathbf{w}(t)] \tag{23-62}$$

Meditch (1969, pp. 32-33) shows that $\mathbf{X}(t)$ is nonsingular so that $\mathbf{X}^{-1}(t)$ exists. Integrate (23-62) with respect to time to see that

$$\mathbf{y}(t) = \int_{t_0}^{t} \mathbf{X}^{-1}(\tau)[\mathbf{B}(\tau)\mathbf{u}(\tau) + \mathbf{G}(\tau)\mathbf{w}(\tau)]d\tau \tag{23-63}$$

so that

$$\mathbf{x}_p(t) = \mathbf{X}(t) \int_{t_0}^{t} \mathbf{X}^{-1}(\tau)[\mathbf{B}(\tau)\mathbf{u}(\tau) + \mathbf{G}(\tau)\mathbf{w}(\tau)]d\tau \tag{23-64}$$

Complete Solution to (23-31)

The complete solution to (23-31) equals the sum of the homogeneous and particular solutions; i.e.,

$$\mathbf{x}(t) = \mathbf{x}_h(t) + \mathbf{x}_p(t) = \mathbf{X}(t)\mathbf{x}(t_0) + \mathbf{X}(t) \int_{t_0}^{t} \mathbf{X}^{-1}(\tau)[\mathbf{B}(\tau)\mathbf{u}(\tau) + \mathbf{G}(\tau)\mathbf{w}(\tau)]d\tau \tag{23-65}$$

Let

$$\boldsymbol{\Phi}(t, \tau) \triangleq \mathbf{X}(t)\mathbf{X}^{-1}(\tau) \tag{23-66}$$

Obviously, $\boldsymbol{\Phi}(t, t) = \mathbf{I}$, and $\boldsymbol{\Phi}(t, t_0) = \mathbf{X}(t)\mathbf{X}^{-1}(t_0) = \mathbf{X}(t)$. Putting these last three equations into (23-65), we find that

$$\mathbf{x}(t) = \boldsymbol{\Phi}(t, t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \boldsymbol{\Phi}(t, \tau)[\mathbf{B}(\tau)\mathbf{u}(\tau) + \mathbf{G}(\tau)\mathbf{w}(\tau)]d\tau \tag{23-67}$$

for $t \geq t_0$, which is the solution to (23-31) that is stated in the main body of this lesson as (23-36).

All that is left is to obtain the differential equation for generating $\boldsymbol{\Phi}(t, \tau)$. From (23-66), we find that

$$\begin{aligned}\frac{d\boldsymbol{\Phi}(t, \tau)}{dt} &= \frac{d\mathbf{X}(t)}{dt}\mathbf{X}^{-1}(\tau) = \mathbf{F}(t)\mathbf{X}(t)\mathbf{X}^{-1}(\tau) \\ &= \mathbf{F}(t)\boldsymbol{\Phi}(t, \tau)\end{aligned} \tag{23-68}$$

subject, of course, to $\boldsymbol{\Phi}(t, t) = \mathbf{I}$.

SUMMARY QUESTIONS

1. Suppose our nonlinear state equation does not depend explicitly on t. Then the perturbation state equation will be: