Appendix D Discrete-Time (ARMAX-type) Models

This Appendix describes discrete-time models of the following types:

1. Autoregressive (AR)
2. Moving average (MA)
3. Autoregressive moving average (ARMA)
4. Autoregressive moving average with exogenous inputs (ARMAX)
5. Autoregressive integrated moving average (ARIMA)

D1 Discrete-Time Models as Integrated Continuous Processes

Discrete-time constant parameter models are appropriate when measurements are sampled at discrete fixed intervals, and system inputs are held constant over sampling intervals. For all of the above types except ARIMA, it is assumed that the system is a wide-sense stationary random process.

A discrete model can be derived from a continuous linear model of the form

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{u}(t) + \mathbf{G}\mathbf{q}(t)\\ \mathbf{y}(t) &= \mathbf{H}\,\mathbf{x}(t) + \mathbf{v}(t)\end{aligned} \tag{D1-1}$$

where $\mathbf{x}(t)$ is the *n*-element state vector, $\mathbf{u}(t)$ is a known *l*-element input, $\mathbf{q}(t)$ is a *p*-element ($p \le n$) vector of random process noise, $\mathbf{y}(t)$ is an *m*-element ($m \le n$) measurement vector, $\mathbf{v}(t)$ is an *m*-element measurement noise vector, $\mathbf{F}$ is an $n \times n$ matrix, $\mathbf{L}$ is $n \times l$, $\mathbf{G}$ is $n \times p$ and $\mathbf{H}$ is $m \times n$. Some elements of the $\mathbf{v}$ vector may be zero. It is usually assumed that

$$\begin{aligned}E[\mathbf{q}(t)] = \mathbf{0},\quad & E[\mathbf{q}^T(t)\,\mathbf{q}(\tau)] = \mathbf{Q}_s\delta(t-\tau)\\ E[\mathbf{v}(t)] = \mathbf{0},\quad & E[\mathbf{v}^T(t)\mathbf{v}(\tau)] = \mathbf{R}_s\delta(t-\tau)\\ & E[\mathbf{q}^T(t)\mathbf{v}(\tau)] = \mathbf{0}\end{aligned}$$

where $\delta(t-\tau)$ is the Dirac delta function

$$\delta(t) = \begin{cases} 0 & t \ne 0 \\ \infty & t = 0 \end{cases} \quad \text{with} \quad \int_{-\infty}^{\infty} \delta(t)\,dt = 1 \;.$$

$\mathbf{Q}_s$ and $\mathbf{R}_s$ are *power spectral densities* (PSD) of the white noise processes. By integrating the $\dot{\mathbf{x}}(t)$ equations over the sampling interval T, treating $\mathbf{q}(t)$ and $\mathbf{v}(t)$ as constant over T, one obtains equations of the form

$$\begin{aligned}\mathbf{x}_i &= \mathbf{A}\,\mathbf{x}_{i-1} + \mathbf{B}\,\mathbf{u}_{i-1} + \mathbf{C}\mathbf{q}_{i-1}\\ \mathbf{y}_i &= \mathbf{D}\,\mathbf{x}_i + \mathbf{v}_i\end{aligned} \tag{D1-2}$$

where we use the notation $\mathbf{x}_i = \mathbf{x}(t_i), \mathbf{x}_{i-1} = \mathbf{x}(t_{i-1}), \mathbf{v}_i = \mathbf{v}(t_i)$, etc. Integration of equations (D1-1) to obtain equation (D1-2) is discussed in Section 2.3. Notice that $\mathbf{v}_i$ and $\mathbf{q}_{i-1}$ are also zero-mean, uncorrelated, white random sequences, but $\mathbf{Cq}_{i-1}$ is equal to the integrated effect of $\mathbf{Gq}(t)$ over the interval T, taking into account the dynamics of $\mathbf{x}(t)$ represented by the first line of equation (D1-1).

Equation (D1-2) is not quite in the form commonly used to define ARMAX models. The measurement noise $\mathbf{v}_i$ is sometimes dropped from the discrete model since it is assumed that the only noise in the system affects both current and past values of $\mathbf{y}_i$. Separate measurement noise can be added later if desired. Also, use of $\mathbf{x}_i$ in $\mathbf{y}_i = \mathbf{D}\,\mathbf{x}_i$ limits the number of zeroes in the resulting ARMA model to one less than the number of poles. To allow the number of zeroes to be equal to the number of poles, the measurement equation is modified so that either a portion of $\mathbf{q}(t)$ directly affects $\mathbf{y}_i$, or $\mathbf{y}_i$ uses variables that will be equal to the value of $\mathbf{x}_i$ at the next time step. In other words, equation (D1-2) is replaced with

$$\begin{aligned} \mathbf{x}_i &= \mathbf{A}\,\mathbf{x}_{i-1} + \mathbf{B}\,\mathbf{u}_{i-1} + \mathbf{C}\mathbf{q}_{i-1} \\ \mathbf{y}_i &= \mathbf{D}(\mathbf{A}\mathbf{x}_i + \mathbf{B}\mathbf{u}_i + \mathbf{C}\mathbf{q}_i) \end{aligned} . \tag{D1-3}$$

Notice that the current control and process noise now directly affect the measurement.

For purposes of this derivation, we now assume that the system is single-input, single-output so that $\mathbf{y}, \mathbf{v}, \mathbf{u}$ and $\mathbf{q}$ in equation (D1-3) are scalars, **B** and **C** are column vectors, and **D** is a row vector. The derivation can be extended to multi-input, multi-output (MIMO) systems, but the notation and derivation become cumbersome. Also, to keep the derivation simple, we temporarily assume that **x** is of dimension two. Again this is not a general restriction. Using these assumptions we write

$$\begin{aligned} y_i &= \mathbf{D}(\mathbf{A}\,\mathbf{x}_i + \mathbf{B}\,u_i + \mathbf{C}q_i) \\ &= \mathbf{D}\,\mathbf{A}(\mathbf{A}\,\mathbf{x}_{i-1} + \mathbf{B}\,u_{i-1} + \mathbf{C}q_{i-1}) + \mathbf{D}(\mathbf{B}\,u_i + \mathbf{C}q_i) \\ &= \mathbf{D}\mathbf{A}^2(\mathbf{A}\,\mathbf{x}_{i-2} + \mathbf{B}\,u_{i-2} + \mathbf{C}q_{i-2}) + \mathbf{DA}(\mathbf{B}\,u_{i-1} + \mathbf{C}q_{i-1}) + \mathbf{D}(\mathbf{B}\,u_i + \mathbf{C}q_i) \\ &= \mathbf{D}\mathbf{A}^3\mathbf{x}_{i-2} + \mathbf{D}\mathbf{A}^2(\mathbf{B}u_{i-2} + \mathbf{C}q_{i-2}) + \mathbf{DA}(\mathbf{B}u_{i-1} + \mathbf{C}q_{i-1}) + \mathbf{D}(\mathbf{B}\,u_i + \mathbf{C}q_i) \end{aligned} \tag{D1-4}$$

and

$$\begin{bmatrix} y_{i-1} \\ y_{i-2} \end{bmatrix} = \begin{bmatrix} \mathbf{DA}^2 \\ \mathbf{DA} \end{bmatrix} \mathbf{x}_{i-2} + \begin{bmatrix} \mathbf{DA} \\ 0 \end{bmatrix} (\mathbf{B}u_{i-2} + \mathbf{C}q_{i-2}) + \mathbf{DB} \begin{bmatrix} u_{i-1} \\ u_{i-2} \end{bmatrix} + \mathbf{DC} \begin{bmatrix} q_{i-1} \\ q_{i-2} \end{bmatrix} . \tag{D1-5}$$

Notice that **DB, DC**, **DAB,** and **DAC** are scalars. We assume that the matrix

$$\begin{bmatrix} \mathbf{DA}^2 \\ \mathbf{DA} \end{bmatrix}$$

in equation (D1-5) is full rank so that it can be inverted to obtain

$$\begin{aligned}\mathbf{x}_{i-2} &= \begin{bmatrix}\mathbf{DA}^2\\ \mathbf{DA}\end{bmatrix}^{-1}\left(\begin{bmatrix}y_{i-1}\\ y_{i-2}\end{bmatrix} - \begin{bmatrix}\mathbf{DA}\\ 0\end{bmatrix}(\mathbf{B}u_{i-2}+\mathbf{C}q_{i-2}) - \mathbf{DB}\begin{bmatrix}u_{i-1}\\ u_{i-2}\end{bmatrix} - \mathbf{DC}\begin{bmatrix}q_{i-1}\\ q_{i-2}\end{bmatrix}\right)\\ &= \begin{bmatrix}k_{11}\\ k_{21}\end{bmatrix} y_{i-1} + \begin{bmatrix}k_{12}\\ k_{22}\end{bmatrix} y_{i-2}\\ &\quad - \begin{bmatrix}k_{11}\mathbf{DA}+k_{12}\mathbf{D}\\ k_{21}\mathbf{DA}+k_{22}\mathbf{D}\end{bmatrix}(\mathbf{B}u_{i-2}+\mathbf{C}q_{i-2}) - \begin{bmatrix}k_{11}\mathbf{D}\\ k_{21}\mathbf{D}\end{bmatrix}(\mathbf{B}u_{i-1}+\mathbf{C}q_{i-1})\end{aligned} \tag{D1-6}$$

where we have defined

$$\begin{bmatrix}k_{11} & k_{12}\\ k_{21} & k_{22}\end{bmatrix} \triangleq \begin{bmatrix}\mathbf{DA}^2\\ \mathbf{DA}\end{bmatrix}^{-1},$$

a 2×2 matrix. Substituting equation (D1-6) in equation (D1-4), we obtain

$$\begin{aligned}y_i &= \mathbf{DA}^3\begin{bmatrix}k_{11}\\ k_{21}\end{bmatrix} y_{i-1} + \mathbf{DA}^3\begin{bmatrix}k_{12}\\ k_{22}\end{bmatrix} y_{i-2}\\ &\quad + \mathbf{DA}^2\left(\mathbf{I} - \mathbf{A}\begin{bmatrix}k_{11}\mathbf{DA}+k_{12}\mathbf{D}\\ k_{21}\mathbf{DA}+k_{22}\mathbf{D}\end{bmatrix}\right)(\mathbf{B}u_{i-2}+\mathbf{C}q_{i-2}) \quad .\\ &\quad + \mathbf{D}\left(\mathbf{A} - \mathbf{A}^3\begin{bmatrix}k_{11}\mathbf{D}\\ k_{21}\mathbf{D}\end{bmatrix}\right)(\mathbf{B}u_{i-1}+\mathbf{C}q_{i-1}) + \mathbf{D}(\mathbf{B}u_i + \mathbf{C}q_i)\end{aligned} \tag{D1-7}$$

Since **D** is a row vector (for scalar y_i), all multiplications of column vectors by **D** produce a scalar result. Hence equation (D1-7) can be written as

$$y_i = -\sum_{j=1}^{2}\alpha_j y_{i-j} + \sum_{j=0}^{2}\beta_j u_{i-j} + \sum_{j=0}^{2}\gamma_j q_{i-j} \tag{D1-8}$$

where α_j, β_j and γ_j are implicitly defined from the matrix multiplications in equation (D1-7):

$$\alpha_1 = \mathbf{DA}^3 \begin{bmatrix} k_{11} \\ k_{21} \end{bmatrix}, \quad \alpha_2 = \mathbf{DA}^3 \begin{bmatrix} k_{12} \\ k_{22} \end{bmatrix}$$

$$\beta_0 = \mathbf{DB}, \quad \beta_1 = \mathbf{D}\left(\mathbf{A} - \mathbf{A}^3 \begin{bmatrix} k_{11}\mathbf{D} \\ k_{21}\mathbf{D} \end{bmatrix} \right)\mathbf{B}, \quad \beta_2 = \mathbf{DA}^2 \left(\mathbf{I} - \mathbf{A} \begin{bmatrix} k_{11}\mathbf{DA} + k_{12}\mathbf{D} \\ k_{21}\mathbf{DA} + k_{22}\mathbf{D} \end{bmatrix} \right)\mathbf{B} \quad . \quad \text{(D1-9)}$$

$$\gamma_0 = \mathbf{DC}, \quad \gamma_1 = \mathbf{D}\left(\mathbf{A} - \mathbf{A}^3 \begin{bmatrix} k_{11}\mathbf{D} \\ k_{21}\mathbf{D} \end{bmatrix} \right)\mathbf{C}, \quad \gamma_2 = \mathbf{DA}^2 \left(\mathbf{I} - \mathbf{A} \begin{bmatrix} k_{11}\mathbf{DA} + k_{12}\mathbf{D} \\ k_{21}\mathbf{DA} + k_{22}\mathbf{D} \end{bmatrix} \right)\mathbf{C}$$

Equation (D1-8) is the general form for a second-order ARMAX model. This form can be extended to an *n*-element model as

$$\boxed{y_i = -\sum_{j=1}^{n} \alpha_j y_{i-j} + \sum_{j=0}^{n} \beta_j u_{i-j} + \sum_{j=1}^{n} \gamma_j q_{i-j} + q_i} \, . \qquad \text{(D1-10)}$$

The u_i terms are the exogenous inputs. If all $\beta_j = 0$, the model is ARMA. If all $\alpha_j = 0$, and $\beta_j = 0$, the model is MA. If all $\gamma_j = 0$ and $\beta_j = 0$, the model is AR. Notice that the summations for y_{i-j}, u_{i-j} and q_{i-j} in equation (D1-10) may be separately truncated at lower order than *n*.

The purpose of this derivation was to show that discrete ARMAX-type models could adequately characterize the behavior of physical systems represented by the stochastic continuous model shown in equation (D1-1). To obtain the ARMAX model, it was necessary to make a number of restrictive assumptions. These include:

1. The sampling interval is constant with no missing measurements,
2. the system is a stationary random process,
3. system inputs are constant over the sampling interval (not required but usually assumed),
4. the continuous model is linear,
5. measurement noise $\mathbf{v}_i$ is dropped from the discrete model and is replaced with direct feed of $\mathbf{u}_i$ and $\mathbf{q}_i$ into the measurement $\mathbf{y}_i$,
6. the system is single-input, single-output.

Although separate measurement noise was removed for derivation purposes, it can be added back into the model at the measurement output. This may be desirable when implementing the model in a Kalman filter—both to make the model more realistic and to avoid numerical problems. The ARMAX structure can be extended to MIMO systems, but it is difficult to derive MIMO ARMAX models from continuous system models.

It is usually not practical to compute the model parameters α_i, β_i and λ_i directly from a continuous model—even if it is available—so they must be determined empirically from measured data. There are many different ways in which this may be done. This topic is explored in Chapter 12.

We now examine the characteristics of each model type. Further information on discrete models and the relationship to z-transforms, correlation functions, and power spectra may be found in Kay (1988), Marple (1987), Kay and Marple (1981), Oppenheim and Schafer (1975), Rabiner and Gold (1975), Priestley (1981), Box, et al (2008), Shumway and Stoffer (2006), Åstrom (1980), Brockwell and Davis (2006), and DeRusso et al. (1965).

D2 Autoregressive (AR) Process

An AR process (without control inputs) is defined as

$$y_i + \alpha_1 y_{i-1} + \alpha_2 y_{i-2} + ... + \alpha_n y_{i-n} = q_i \quad . \tag{D2-1}$$

AR processes contain only poles (no zeroes) since the z-transform is

$$\frac{Y(z)}{Q(z)} = \frac{1}{1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + ... + \alpha_n z^{-n}} \tag{D2-2}$$

where z^{-1} is the unit delay operator,

$$\begin{aligned} z^{-1} &= \mathrm{e}^{-j2\pi f/T} \\ &= \cos(j2\pi f/T) - j\sin(j2\pi f/T) \end{aligned} ,$$

$j = \sqrt{-1}$ and f is frequency in Hz. The denominator of equation (D2-2) can be factored to obtain

$$\frac{Y(z)}{Q(z)} = \frac{\kappa}{(1 - p_1 z^{-1})(1 - p_2 z^{-1}) \cdots (1 - p_n z^{-1})} \tag{D2-3}$$

where $p_1, p_2, \ldots$ are the z-plane poles and

$$\kappa = \frac{\prod_{i=1}^{n} (1 - p_i)}{1 + \sum_{i=1}^{n} \alpha_i} .$$

If the roots of the denominator all lie within the unit circle in the complex plane, the model is *stable* (output remains bounded) and *causal* (output only depends on past inputs, not future). Stability can be understood by considering the first-order AR model $y_i = -\alpha_1 y_{i-1} + q_i$ with the pole located on or outside the unit circle. When $\alpha_1 = 1$, y_i will alternate in sign at each time step with an added perturbation q_i that may either increase or decrease the magnitude of y_i, but will tend to make the variance of y increase. When $\alpha_1 < -1$, y_i will increase at each step even without the effect of q_i.

When $\alpha_1 > 1$, y_i will alternate in sign (provided that q_i is smaller than y_i) at each step, but y_i will still increase in magnitude.

The AR model is ideal for modeling narrowband signals since the output of the transfer function equation (D2-3) will peak near the frequencies represented by the poles. However, it cannot accurately model a transfer function with nulls unless the expansion is carried to infinite order.

Equation (D2-1) can be represented in state-space form as:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}_i = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0. & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -\alpha_1 & -\alpha_2 & -\alpha_3 & \dots & -\alpha_{n-1} & -\alpha_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}_{i-1} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ q_{i-1} \end{bmatrix} \quad . \qquad \text{(D2-4)}$$

$$y_i = q_i - \sum_{j=1}^{n} \alpha_j x_{j,i}$$

This form of the state-space model is called the *controllable companion form* since a single control input on the last state (i.e., treat q_{i-1} as a known control input) can be found that will drive the output y to any desired value if the system is stable. The companion form has certain benefits in control system analysis when studying the effects of feedback. Other equivalent state-space representations are also possible. For example, equation (D2-5) shows the *observable companion form*. Other structures include direct, cascade, parallel, Lagrange, and frequency sampling (Rabiner and Gold 1975). It should be noted that high-order companion matrices tend to be nearly singular and high-order controllable canonical forms tend to be uncontrollable (Kenney and Laub 1988).

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}_i = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & -\alpha_n \\ 1 & 0 & 0 & \dots & 0 & -\alpha_{n-1} \\ 0 & 1 & 0 & \dots & 0 & -\alpha_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -\alpha_2 \\ 0 & 0 & 0 & \dots & 1 & -\alpha_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}_{i-1} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ q_{i-1} \end{bmatrix} \quad . \qquad \text{(D2-5)}$$

$$y_i = q_i + x_{n-1,i} - \alpha_1 x_{n,i}$$

For a third-order system, the state-space model of equation (D2-4) can be represented in block-diagram form as shown Figure D2-1.

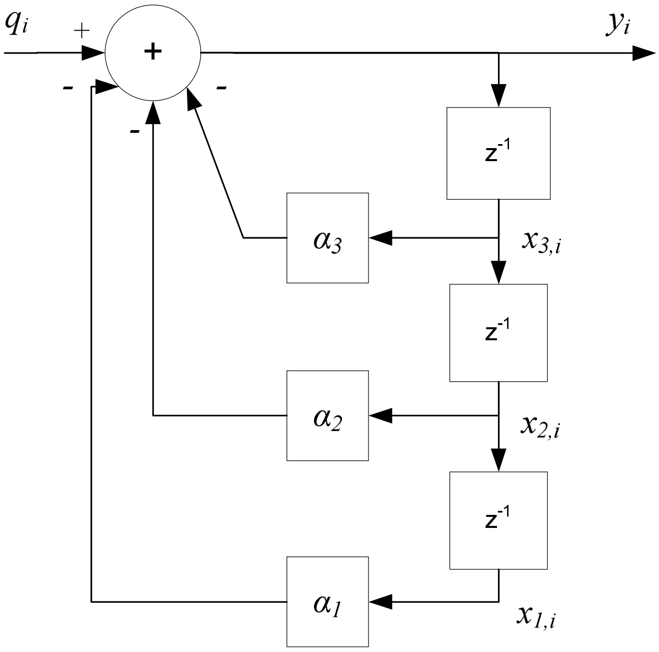

Figure D2-1: Third-order AR model.

Example D2-1: Third-order AR Model
Consider the third-order model that may be used to represent colored noise in a dynamic system:

$$\begin{aligned}\frac{Y(z)}{Q(z)} &= \frac{1}{(1-0.8z^{-1})\left[1-(0.65+j0.65)z^{-1}\right]\left[1-(0.65-j0.65)z^{-1}\right]} \\ &= \frac{1}{(1-0.8z^{-1})(1-1.3z^{-1}+0.845z^{-2})} \\ &= \frac{1}{1-2.1z^{-1}+1.885z^{-2}-0.676z^{-3}}\end{aligned} \quad . \qquad \text{(D2-6)}$$

with roots $z = 0.8,\ (0.65+j0.65),\ (0.65-j0.65)$ as shown in Figure D2-2. Since q_i is assumed to be white (uncorrelated) noise, the PSD of q is the same at all frequencies:

$$S_q(f) = S_{q0}, \quad |f| < \frac{1}{2T} \quad .$$

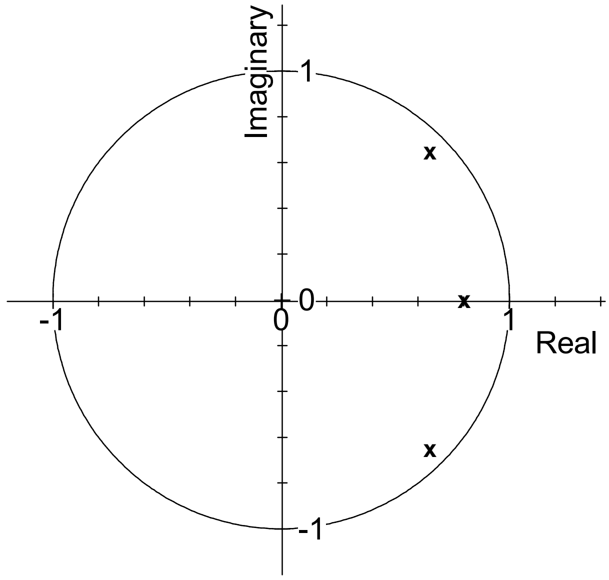

Figure D2-2: Pole locations in complex z-plane for third-order AR model.

The PSD of the real output y_i is

$$
\begin{aligned}
S_y(f) &= \left|Y(z)\right|^2_{z=\exp(j2\pi f/T)} = Y(z)Y(z^{-1})\Big|_{z=\exp(j2\pi f/T)} \\
&= S_{q0}\left[\frac{1}{(1-0.8z^{-1})(1-1.3z^{-1}+0.845z^{-2})}\right]\left[\frac{1}{(1-0.8z)(1-1.3z+0.845z^2)}\right]\Bigg|_{z=\exp(j2\pi f/T)} \\
&= \frac{S_{q0}}{\left(1.64-1.6\,\mathrm{Re}(z)\right)\left(1.714025-4.797\,\mathrm{Re}(z)+3.380\,\mathrm{Re}(z)^2\right)}\Bigg|_{z=\exp(j2\pi f/T)}
\end{aligned}
\tag{D2-7}
$$

evaluated along the unit circle $z=\cos(j2\pi f/T)+j\sin(j2\pi f/T)$ so that Re(z) = cos($2\pi f/T$). $S_y(f)$ is plotted in Figure D2-3 as a function of frequency for $S_{q0}=1$. Notice that the PSD peaks at a frequency of about 0.125 Hz, which corresponds to the location of the complex pole pair z = (0.65 +j 0.65), (0.65 -j 0.65). Since the poles are only 0.081 from the unit circle, the damping is low and the PSD peak is pronounced. Also notice that the PSD magnitude drops starting from frequency zero, and this is due to the pole at z = (0.80+j0).

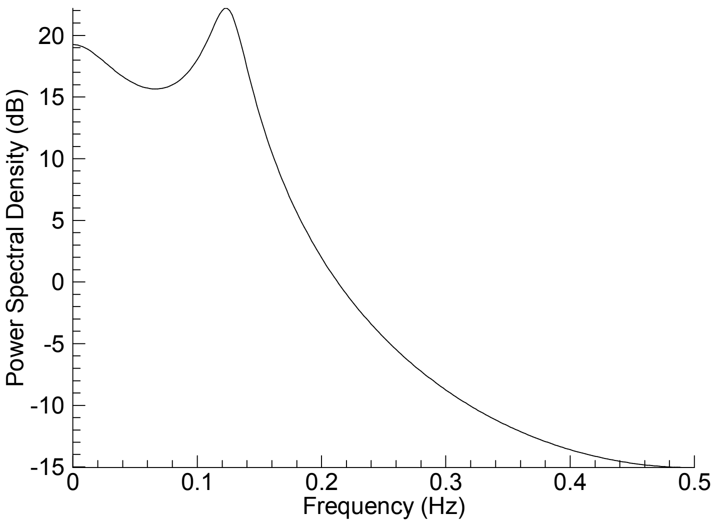

Figure D2-3: PSD of example D2-1 (third-order AR model).

D3 Moving Average (MA) Process

An MA process is simply a weighed average of past inputs, with no feedback:

$$y_i = q_i + \gamma_1 q_{i-1} + \gamma_2 q_{i-2} ... + \gamma_p q_{i-p} \ . \tag{D3-1}$$

Again we have ignored control inputs. The transfer function for the MA model is

$$\begin{aligned} \frac{Y(z)}{Q(z)} &= 1 + \gamma_1 z^{-1} + \gamma_2 z^{-2} + ... + \gamma_n z^{-n} \\ &= \kappa(1 - r_1 z^{-1})(1 - r_2 z^{-1}) \cdots (1 - r_n z^{-1}) \end{aligned} \tag{D3-2}$$

where $r_1, r_2, \ldots$ are the z-plane zeroes and

$$\kappa = \frac{1 + \sum_{i=1}^{n} \gamma_i}{\prod_{i=1}^{n} (1 - r_i)} \ .$$

MA models are ideal for modeling nulls in the PSD, but cannot accurately model peaks.

A block diagram for a third-order MA model is shown in Figure D3-1. Unlike the AR model, it is not possible to predict future y_i from just past values since y_i depends on past values of the random inputs u_i (which are not observed). To predict one time step in the future it is necessary to implement the model in a Kalman filter and estimate the states from measurements y.

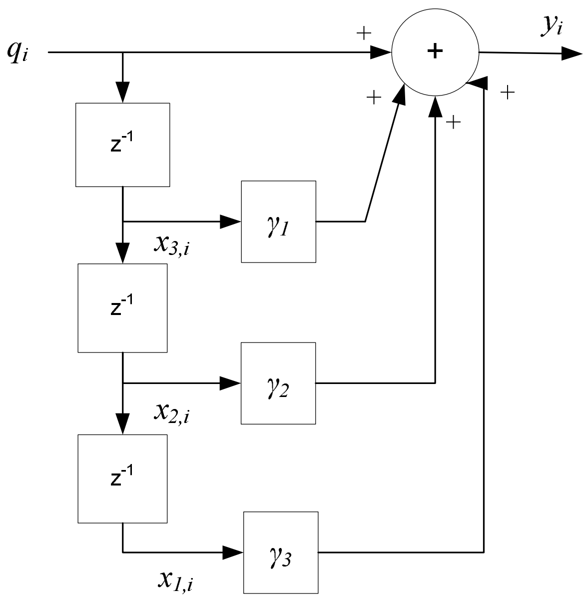

Figure D3-1: Third-order MA model.

Example D3-1: Third-order MA Process

This example places the zeros of the MA transfer function close to the imaginary axis:

$$\begin{aligned}\frac{Y(z)}{Q(z)} &= (1-0.3z^{-1})\left[1-(0.16+j0.89)z^{-1}\right]\left[1-(0.16-j0.89)z^{-1}\right] \\ &= (1-0.3z^{-1})(1-0.32z^{-1}+0.8177z^{-2}) \\ &= (1-0.62z^{-1}+0.9137z^{-2}-0.24531z^{-3})\end{aligned} \qquad \text{(D3-3)}$$

Figure D3-2 shows the location of the zeroes in the z-plane and Figure D3-3 shows the PSD. Again the zeroes are close to the unit circle (lightly damped) and the PSD null at 0.25 Hz is sharp. The zero at 0.3+j0 does not have a strong affect on the PSD because it is close to the origin.

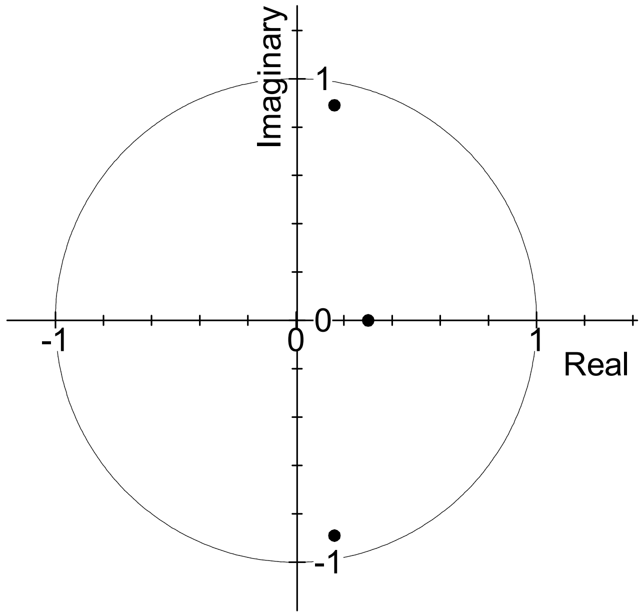

Figure D3-2: Zero locations for third-order MA model.

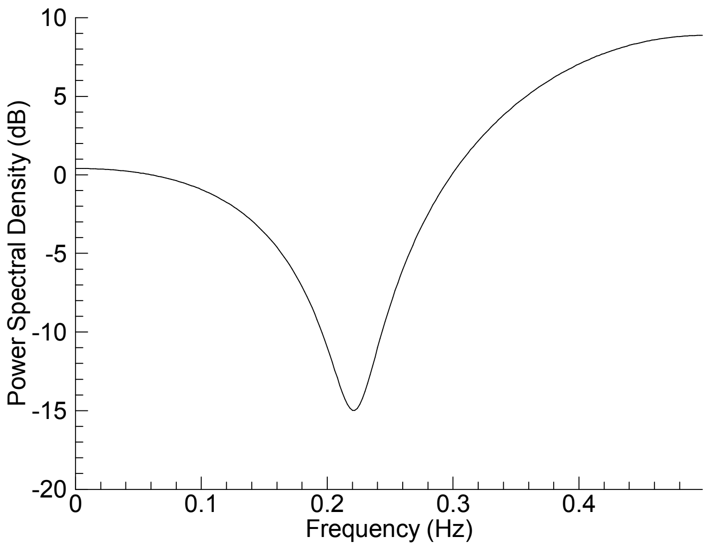

Figure D3-3: PSD of Example D3-1 (third-order MA model).

D4 Autoregressive Moving Average (ARMA) Process

As the name implies, an ARMA process has both poles and zeroes. The number of zeroes is generally less than or equal to the number of poles, but this is not an inherent restriction. Poles must be inside the unit circle for a *stable* and *causal* system. If all poles and zeroes are inside the unit circle it is called a *minimum phase* or *invertible* system; that is, the input noise q_i can be reconstructed from knowledge of the past outputs y_j for $j \le i$. The ARMA difference equation is:

$$y_i + \alpha_1 y_{i-1} + \alpha_2 y_{i-2} + ... + \alpha_n y_{i-n} = q_i + \gamma_1 q_{i-1} + \gamma_2 q_{i-2} ... + \gamma_l q_{i-l} \quad . \tag{D4-1}$$

The transfer function for the ARMA model is

$$\frac{Y(z)}{Q(z)} = \frac{1 + \gamma_1 z^{-1} + \gamma_2 z^{-2} + \ldots + \gamma_l z^{-l}}{1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \ldots + \alpha_n z^{-n}} \tag{D4-2}$$

which can be factored to obtain

$$\frac{Y(z)}{Q(z)} = \frac{\kappa (z - r_1)(z - r_2) \cdots (z - r_l)\, z^{n-l}}{(z - p_1)(z - p_2) \cdots (z - p_n)} \tag{D4-3}$$

where

$$\kappa = \left(\frac{\prod_{i=1}^{n} (1 - p_i)}{1 + \sum_{i=1}^{n} \alpha_i} \right) \left(\frac{1 + \sum_{i=1}^{l} \gamma_i}{\prod_{i=1}^{l} (1 - r_i)} \right) .$$

It is important that common factors appearing in both the numerator and denominator of equation (D4-3) be removed since they can cause observability problems when ARMA models are used for estimation purposes.

The zeroes, r_i, and poles, p_i, may be real or complex. Complex poles or zeroes must occur in complex conjugate pairs for a real system. Hence equation (D4-3) can be written in terms of first-order factors for real poles and zeroes, and second-order factors for complex conjugate poles or zeroes. For example, the transfer function for complex conjugate pole (p, p^*) and zero (r, r^*) pairs can be written as

$$\frac{(z - r)(z - r^*)}{(z - p)(z - p^*)} = \frac{z^2 - 2\,\mathrm{Re}(r)\, z + |r|^2}{z^2 - 2\,\mathrm{Re}(p)\, z + |p|^2} \tag{D4-4}$$

where $|\cdot|^2 = \mathrm{Re}(\cdot)^2 + \mathrm{Im}(\cdot)^2$ and $\mathrm{Im}(\cdot)$ is the imaginary part of the complex number.

A third-order ARMA model with added measurement noise v_i can be represented in state-space form as:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_i = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_1 & -a_2 & -a_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{i-1} + \begin{bmatrix} 0 \\ 0 \\ q_{i-1} \end{bmatrix}$$

$$y_i = \begin{bmatrix} \gamma_3 - \alpha_1 & \gamma_2 - \alpha_2 & \gamma_1 - \alpha_3 \end{bmatrix} \mathbf{x}_i + q_i + v_i \qquad \text{(D4-5)}$$

This is shown in the block diagram of Figure (D4-1). This structure uses the minimum number of delays and is sometimes called *canonical*, although the description is ambiguous as other three-delay structures are also possible.

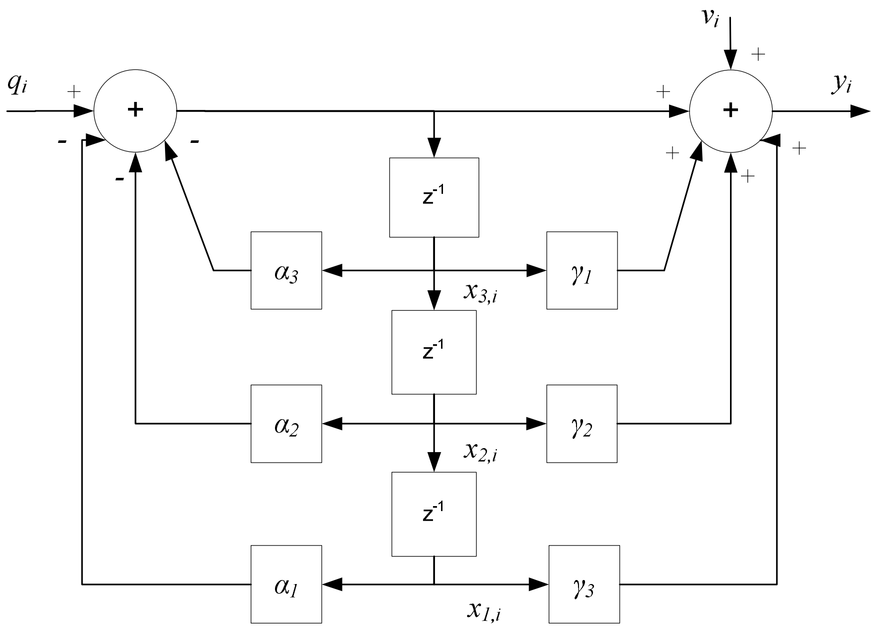

Figure D4-1: Third-order ARMA model (with added measurement noise).

To demonstrate that an ARMA model can be implemented using state-space representations other than companion forms, we expand equation (D1-23) in partial fractions (provided that $l < n$), to obtain

$$\frac{Y(z)}{Q(z)} = \frac{K_1}{z - p_1} + \frac{K_2}{z - p_2} + \cdots + \frac{K_n}{z - p_n} \qquad \text{(D4-6)}$$

where the complex coefficients K_i are determined by the method of residues (DeRusso et al. 1965):

$$K_i = \left[\frac{Y(z)}{Q(z)}(z - p_i) \right]_{z=p_i} . \qquad \text{(D4-7)}$$

If repeated roots of the form $1/(z-p_i)^r$ appear in the transfer function, then equation (D4-6) will contain terms

$$\frac{K_i}{(z-p_i)} + \frac{K_{i+1}}{(z-p_i)^2} + ... + \frac{K_{i+r-1}}{(z-p_i)^r} \qquad \text{(D4-8)}$$

and the *K* coefficients are computed as

$$K_{i+r-1-j} = \frac{1}{j!} \left[\frac{d^j}{dz^j} \left(\frac{Y(z)}{Q(z)}(z-p_i)^r \right) \right]_{z=p_i} \quad \text{for } j = 0,...,r-1 \qquad \text{(D4-9)}$$

Repeated roots at $z = 0$ are sometimes used to account for transport delays if the delay is not explicitly modeled.

Example D4-1: Partial Fraction Expansion of Third-order ARMA Model

Consider the third-order ARMA model

$$\frac{Y(z)}{Q(z)} = \frac{(z-r_1)(z-r_1^*)(z-r_3)}{(z-p_1)(z-p_1^*)(z-p_3)} \qquad \text{(D4-10)}$$

with complex conjugate poles (p_1, p_1^*) and zeros (r_1, r_1^*). To use a partial fraction expansion, the order of the numerator polynomial must be lower than the order of the denominator. Hence we must either use a preliminary step of long division to write

$$\frac{Y(z)}{Q(z)} = 1 + \frac{R(z)}{Q(z)}$$

where *R*(*z*) is second order, or to multiply by (z/z) to obtain

$$\frac{Y(z)}{Q(z)} = z\left(\frac{Y(z)}{z\,Q(z)} \right)$$

and then expand $Y(z)/[z\,Q(z)]$ in partial fractions. We use the second method to obtain

$$\frac{Y(z)}{Q(z)} = z\left[\frac{K_1}{z-p_1} + \frac{K_1^*}{z-p_1^*} + \frac{K_3}{z-p_3} + \frac{K_4}{z} \right] \qquad \text{(D4-11)}$$

where

$$K_1 = \frac{(p_1 - r_1)(p_1 - r_1^*)(p_1 - r_3)}{(p_1 - p_1^*)(p_1 - p_3)p_1}, \quad K_3 = \frac{(p_3 - r_1)(p_3 - r_1^*)(p_3 - r_3)}{(p_3 - p_1)(p_3 - p_1^*)p_3}, \quad K_4 = \frac{r_1 r_1^* r_3}{p_1 p_1^* p_3}$$

The complex pole pairs can be combined to obtain

$$\frac{Y(z)}{Q(z)} = \frac{2\,\mathrm{Re}(K_1)\,z^2 - 2\left(\mathrm{Re}(K_1)\,\mathrm{Re}(p_1) + \mathrm{Im}(K_1)\,\mathrm{Im}(p_1)\right)z}{z^2 - 2\,\mathrm{Re}(p_1)z + |p_1|^2} + \frac{K_3 z}{z - p_3} + K_4 . \quad \text{(D4-12)}$$

A state-space representation of equation (D4-12) is

$$\begin{aligned}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_i &= \begin{bmatrix} 0 & 1 & 0 \\ -\alpha_1 & -\alpha_2 & 0 \\ 0 & 0 & -\alpha_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{i-1} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} q_{i-1} \\
y_i &= \begin{bmatrix} -\gamma_1\alpha_1 & \gamma_2 - \gamma_1\alpha_2 & -\gamma_3\alpha_3 \end{bmatrix} \mathbf{x}_i + (K_4 + \gamma_1 + \gamma_3) q_i + v_i
\end{aligned} \quad \text{(D4-13)}$$

where

$$\begin{aligned}
&\alpha_1 = |p_1|^2, \quad \alpha_2 = -2\,\mathrm{Re}(p_1), \quad \alpha_3 = -p_3 \\
&\gamma_1 = 2\,\mathrm{Re}(K_1), \quad \gamma_2 = -2\left(\mathrm{Re}(K_1)\,\mathrm{Re}(p_1) + \mathrm{Im}(K_1)\,\mathrm{Im}(p_1)\right), \quad \gamma_3 = K_3
\end{aligned} . \quad \text{(D4-14)}$$

Figure D4-2 shows the block diagram. This structure can be less sensitive to the numerical singularity problems of the companion forms.

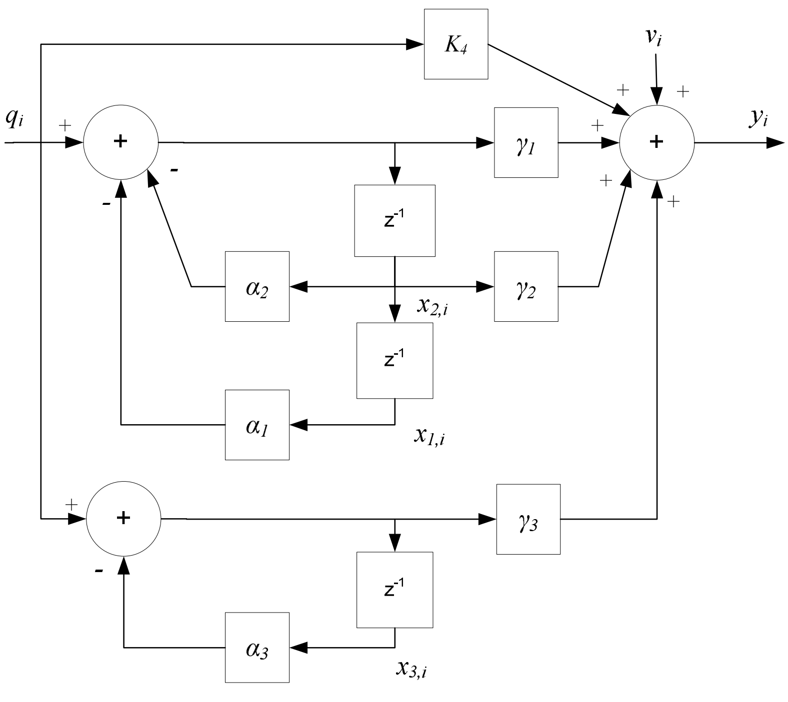

Figure D4-2: Alternate Third-Order ARMA Model (with added measurement noise)

Example D4-2: Third-order ARMA Process

Combining the poles from Example D2-1 with the zeros of Example D3-1 we get the transfer function equation (D4-15) and PSD of y_i in Figure D4-3. As expected, the peak and null of the individual AR and MA processes are evident.

$$\frac{Y(z)}{Q(z)} = \frac{(1-0.3z^{-1})(1-0.32z^{-1}+0.8177z^{-2})}{(1-0.8z^{-1})(1-1.3z^{-1}+0.845z^{-2})} \qquad \text{(D4-15)}$$

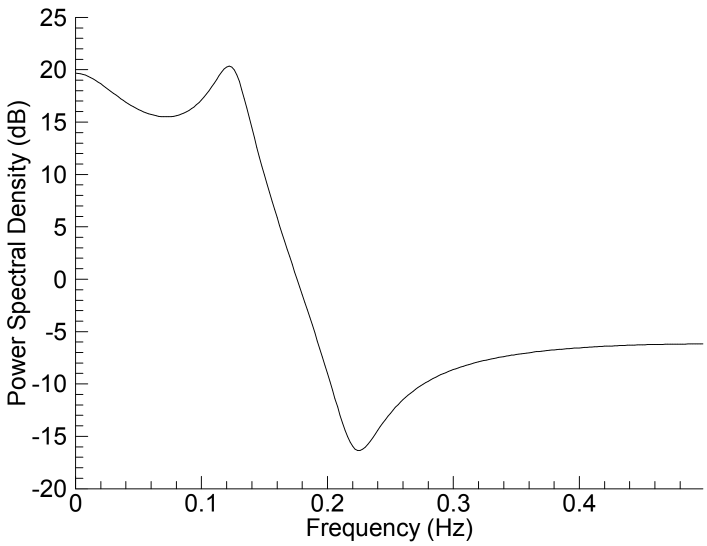

Figure D4-3: PSD of Example D4-2 (third-order ARMA model)

D5 Autoregressive Moving Average Process with Exogenous Inputs (ARMAX)

Exogenous inputs u_i are added to an ARMA process to obtain an ARMAX process:

$$\begin{aligned} y_i + \alpha_1 y_{i-1} + \alpha_2 y_{i-2} + \ldots + \alpha_n y_{i-n} &= \beta_0 u_i + \beta_1 u_{i-1} + \ldots \beta_m u_{i-m} \\ &\quad + q_i + \gamma_1 q_{i-1} + \gamma_2 q_{i-2} \ldots + \gamma_q q_{i-l} \end{aligned} . \qquad \text{(D5-1)}$$

Adding exogenous inputs does not change the characteristics of ARMA processes described in the previous section, except that the output power spectra will also include the effects of the u_i terms.

We demonstrate how AR, MA, ARMA, or ARMAX processes can be extended to MIMO cases using a two-input, two-output, third-order ARMAX model. One possible implementation in state-space form is:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_i = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\alpha_1 & -\alpha_2 & -\alpha_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{i-1} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_{i-1} + \begin{bmatrix} 0 \\ 0 \\ q_{i-1} \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_i = \begin{bmatrix} \gamma_{13}-\alpha_1 & \gamma_{12}-\alpha_2 & \gamma_{11}-\alpha_3 \\ \gamma_{23}-\alpha_1 & \gamma_{22}-\alpha_2 & \gamma_{21}-\alpha_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_i + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_i + \begin{bmatrix} q_i \\ q_i \end{bmatrix} \quad \text{(D5-2)}$$

We have abused the subscript notation somewhat, but the meaning should be clear. In this example the system has a different response to the exogenous inputs u_i than to the noise q_i.

D6 Autoregressive Integrated Moving Average (ARIMA) Process

Measured outputs of many systems have nonzero means, trends, or other types of possibly random long-term behavior. Hence the measurements are not statistically stationary, but successive differences of measurements may be stationary. In this case the system may be treated as an ARIMA process. Depending on the type of nonstationary behavior, there are several approaches for handling the nonstationary behavior. When the time series has slow random changes in level, it may be appropriate to work with short sections of the data and treat each section as a stationary ARMA or ARMAX process. In other cases, first, second or higher order differences of the data may be computed and then modeled as ARMA/ARMAX processes. Alternately the differencing can be included in an ARIMA model. For further information on linear nonstationary models see Box et al. (2008, chapter 4).

D7 Discrete Model Summary

We have shown that discrete models can be developed from discretely sampled continuous dynamic models, but they impose the restrictions that the sampling interval must be uniform with no data dropouts, and the system must be statistically stationary. Nonetheless, discrete ARMAX-type models are often used in applications when it is difficult to develop models from first principles. For example, they are sometimes used in process control (Åstrom 1980, Levine 1996) and biomedical (Bronzino 2006, Lu et al. 2001, Guler 2002) applications. When ARMAX coefficients cannot be computed from first principles, they must be determined empirically from input-output data. This topic is discussed in Chapter 12.