

7.1 Digital Filter Response

A digital filter can be described in several ways. These include, but are not limited to, difference equations, block diagram, impulse response, and the system function. Each model is useful in the description of systems and their behavior, and they are all related. They provide different views of the same entity.

A good place to start is with a difference equation. The difference equation is an important model in modeling random processes as well as in describing digital systems. It is a completely general model of a time-invariant discrete linear system. We will develop difference equations as a digital system model and then relate that approach to other common models.

After we have developed the description tools, we will show how they can be used to model random processes by using to construct random processes with various properties. These will enable us to employ various analytical tools on processes with known properties so that we can gain insight into the usefulness and limitation of each tool. This will prepare us for encounters with random processes with unknown structures.

7.1.1 Difference Equation Model

Let the input sequence to a digital filter be denoted by $x(n)$ and the corresponding response by $y(n)$. The current output value can depend upon the current input value, past input values and past output values. This can be expressed as

$$\begin{aligned} y(n) + a_1y(n-1) + a_2y(n-2) + \cdots + a_py(n-p) \\ = b_0x(n) + b_1x(n-1) + \cdots + b_qx(n-q) \end{aligned} \quad (7.1)$$

The difference equation is completely described by the coefficients $\{1, a_1, \dots, a_p\}$ and $\{b_0, b_1, \dots, b_p\}$. Note that we will always assume (without loss of generality) that $a_0 = 1$. The above equations can be solved for the current output value.

$$\begin{aligned} y(n) = b_0x(n) + b_1x(n-1) + \cdots + b_qx(n-q) - a_1y(n-1) \\ - a_2y(n-2) - \cdots - a_py(n-p) \end{aligned} \quad (7.2)$$

A block diagram of this equation is shown in Figure 7.1 on page 128. The storage elements in the top row preserve a history of the past inputs and those in the bottom row preserve a history of the past outputs.

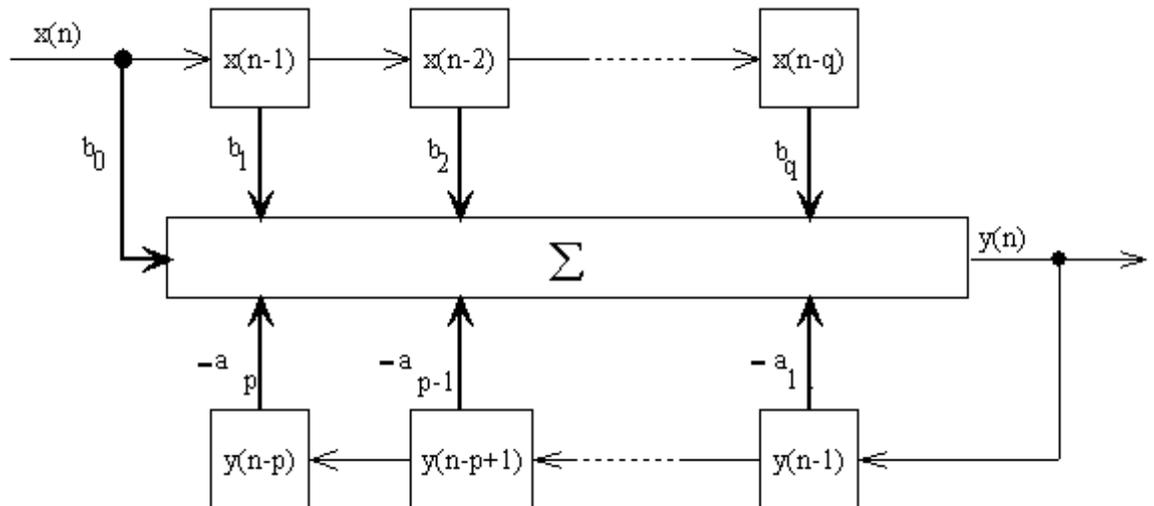


Figure 7.1: Block diagram of a digital filter realization of the difference equation.

Example 7.1.1 Single Pole Filter This filter is characterized by a simple difference equation

$$y(n) + ay(n - 1) = x(n) \quad (7.3)$$

To see its behavior, consider the simple input sequence $[\dots 0, 1, 0, 0, 0, 0, \dots]$ which has a single nonzero input value. Without loss of generality we can choose the index origin so that $x(0) = 1$ and $x(n) = 0$ for $n \neq 0$. The current output can be written in terms of the current input and the immediate past output.

$$y(n) = x(n) - ay(n - 1)$$

This equation is illustrated by the block diagram in Figure 7.2 on page 129. We will assume the initial state $y(-1) = 0$. Then it is a simple matter to write out a few of the values of the output sequence beginning with $n = 0$.

n	0	1	2	3	4
$y(n)$	1	$-a$	a^2	$-a^3$	a^4

By direct substitution into the difference equation you can show that the general solution for this input and initial condition is $y(n) = -(a)^n$ for $n \geq 0$. The output is clearly bounded if and only if $|a| \leq 1$. We will see later when

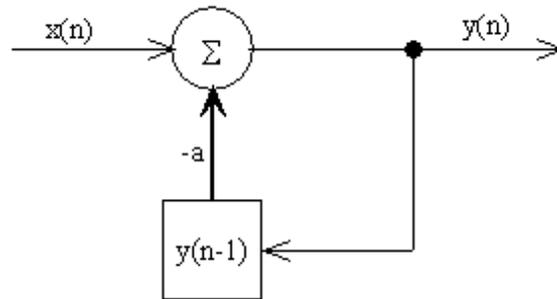


Figure 7.2: Block diagram of a single pole filter. The system stores one value, the immediately preceding output. This value establishes the system state.

we discuss the system function that this system has a “single pole” whose location is determined by a , and which is stable if the “pole” is inside the unit circle.

This example illustrates the use of an impulse input to gain insight into the behavior of a system. The impulse response can be used as a general tool for this purpose.

Impulse Response

The impulse response of a digital filter is the sequence that is generated when the input sequence is $x(n) = \delta(n)$, where

$$\delta(n) = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases} \quad (7.4)$$

The response to this input sequence is traditionally referred to as $h(n)$. The impulse response is sufficient to provide a complete description of the behavior of any linear time-invariant system, and is therefore equivalent to the difference equation description or the block diagram. The response of a system to a general input sequence $x(n)$ is

$$y(n) = \sum_{m=-\infty}^{\infty} h(m)x(n-m) \quad (7.5)$$

Clearly, $y(n) = h(n)$ when $x(n) = \delta(n)$.

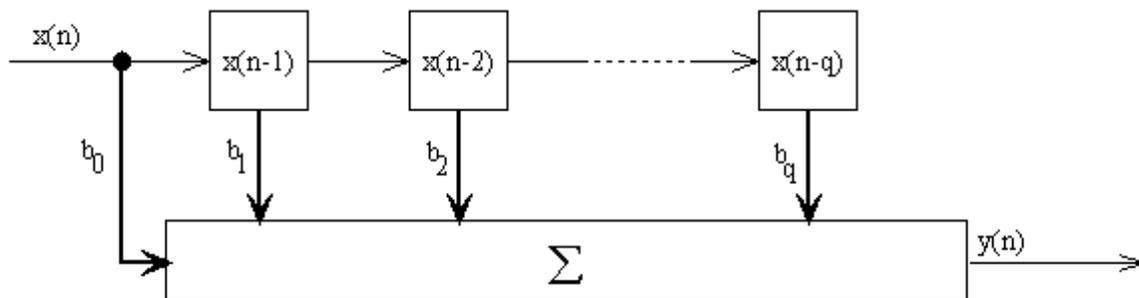


Figure 7.3: Block diagram of a finite impulse response digital filter. The response of a single input sample persists for q additional steps.

A special class of systems has an impulse response of finite duration. These systems are characterized by the absence of a feedback path. That is, $a_1 = a_2 = \dots = a_p = 0$. The block diagram of a finite impulse response (FIR) system is shown in Figure.7.3.

The impulse response of the FIR filter is

$$h(n) = \begin{cases} 0, & n < 0 \\ b(n), & 0 \leq n \leq q \\ 0, & n > q \end{cases} \quad (7.6)$$

This relationship provides one means to realize any impulse response that one may need. The algorithm must store as many of the past inputs as necessary to provide a good approximation through use of a finite number of coefficients. The practical issue is the size of the memory, determined by q , that is practical.

FIR Applications

The FIR filter is often used to smooth a random process to suppress noise and bring out a slower-varying signal. This is shown in the first example below. Another application is in the detection of a signal in a noisy background with a matched filter. We will look briefly at these applications.

Example 7.1.2 Weighted Average Suppose that we would like to smooth a waveform by averaging over several values of the input. We may want to

reduce the weight given to input values that are farther in the past. Let r be a number in the range $0 \leq r \leq 1$. Let the impulse response be defined by

$$h(n) = \begin{cases} 0, & n < 0 \\ r^n, & 0 \leq n \leq q \\ 0, & n > q \end{cases} \quad (7.7)$$

Then the output, using (7.5), is

$$y(n) = \sum_{m=0}^q r^m x(n-m) \quad (7.8)$$

If $r = 1$ we give equal weight to the q past outputs and if $r < 1$ we give less weight to those farther in the past. This can be effective in reducing the amount of noise on a slowly varying signal, as shown in Figure ?? on page ?. In the example shown in the figure the signal is a sine wave that has a period of 100 samples. The noise can be reduced by averaging over an interval of about 10 samples without seriously distorting the signal.

Matched Filter A FIR filter can also be used for an operation called *matched filtering*. Suppose that one wants to detect the presence of a waveform $s(n)$ that is of some finite duration, but that it is observed in the presence of additive noise. Then the observed samples are $x(n) = s(n) + \xi(n)$, where $\xi(n)$ represents the noise process. In this example we will assume that the noise is “white” so that $E[\xi(m)\xi(n)] = \sigma^2\delta(m-n)$. Let the duration of $s(n)$ be $q+1$ so that the signal can be represented by the samples $\{s(0), s(1), \dots, s(q)\}$. Then the maximum-likelihood test for the presence of s in an observed sequence x is to form the weighted sum

$$u = \sum_{n=0}^q x(n)s(n) \quad (7.9)$$

and compare the result u to a decision threshold η . If $u > \eta$ then the signal is judged to be present in the sample sequence $x(n)$. We can substitute $x(n) = s(n) + \xi(n)$ (case of signal present) and $x(n) = \xi(n)$ (case of signal absent) into (7.9) to determine the expected value of u in each case. When the signal is *absent* the expected value of u will be (assuming that the average value of the noise is zero)

$$E[u] = \sum_{n=0}^q E[\xi(n)s(n)] = \sum_{n=0}^q s(n)E[\xi(n)] = 0 \quad (7.10)$$

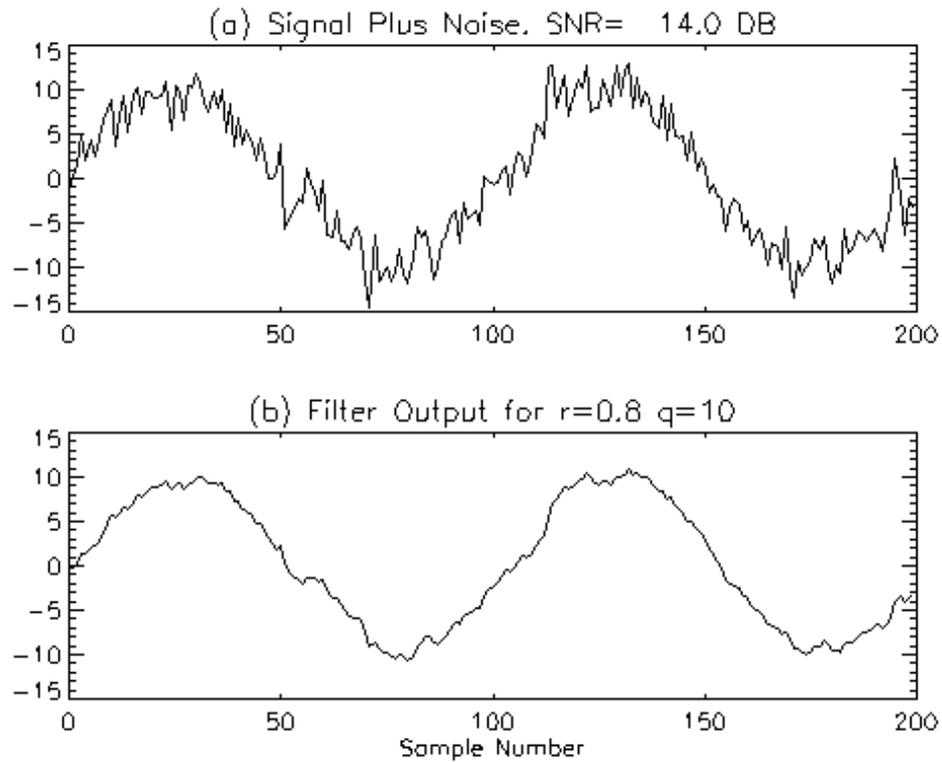


Figure 7.4: The averaging filter is effective in reducing the amount of noise on a low-frequency signal.

When the signal is *present*, the expected value of u will be

$$\begin{aligned}
 E[u] &= \sum_{n=0}^q E[(s(n) + \xi(n))s(n)] & (7.11) \\
 &= \sum_{n=0}^q s^2(n) + s(n)E[\xi(n)] \\
 &= \mathcal{E}_s
 \end{aligned}$$

We should be able to determine which is the case by evaluating (7.9) with an observed $x(n)$ and comparing the result to a decision threshold. The

parameter \mathcal{E}_s is called the “energy” in s , and is equal to

$$\mathcal{E}_s = \sum_{n=0}^q s^2(n) \quad (7.12)$$

The decision threshold is often set midway between the output that would be expected with signal absent and that which would be expected with signal present. That would produce a threshold of $\eta = \mathcal{E}_s/2$. More extensive analysis based on risk analysis can be used to refine this setting, but it is sufficient for the purpose of illustrating the principle.

The value of u can be produced by passing the values of x into a FIR filter whose weights are determined by s . We can see this by constructing a digital filter with the coefficients $b(m) = s(q - m)$ for $0 \leq m \leq q$. The response of such a filter to the sequence $x(n)$ would be

$$y(n) = \sum_{m=0}^q b(m)x(n - m) = \sum_{m=0}^q s(q - m)x(n - m) \quad (7.13)$$

At the point $n = q$ the output value is $y(q) = u$, which can be seen by $n = q$ in the above equation and then making the index change $k = q - m$.

$$y(q) = \sum_{m=0}^q s(q - m)x(q - m) = \sum_{k=0}^q s(k)x(k) = u \quad (7.14)$$

Example 7.1.3 Matched Filter A signal $s(n) = 10e^{-n/6} \sin(2\pi(.005)n)$ $0 \leq n \leq 30$ will be detected when an object is present. No signal will be detected when the object is absent. The detected waveform is $x(n) = s(n) + \xi(n)$ when the object is present and $x(n) = \xi(n)$ when the object is not present. The decision can be made by comparing the output of a matched filter to a decision threshold. In this case the matched filter coefficients are $b(n) = s(10 - n) = 10e^{(30-n)/6} \sin(2\pi(.15 - 0.005n))$

$$b(n) = s(10 - n) = 10e^{(30-n)/6} \sin(2\pi(.15 - 0.005n)), \quad 0 \leq n \leq 10$$

A plot of the signal is shown in Figure 7.5 on page 135, where we see that it has the characteristic of a damped sinusoidal pulse. When the same signal is observed with additive noise with standard deviation $\sigma = 1.5$, the result is a noisy waveform such as the middle plot of the figure. The response of a

matched filter to this signal is shown in the lower panel of Figure 7.5. We observe that the filter output builds up slowly as each sample of the noisy waveform is entered. The final value of the filter output is used in making a decision about the presence of the signal.

The probability that the matched filter output may be on the wrong side of the decision threshold is explored in Figure 7.6 on page 136. A total of 300 tests were run in which noise was generated and added to the signal waveform and then filtered. The upper figure shows a record of the final filter output value under the assumption of signal present (top line) and only noise (bottom line). The line dividing the records is at a level $\mathcal{E}_s/2$, and corresponds to the decision threshold. It is noted that a few points would cross the line from either direction, corresponding to either missing the target or a false alarm. A pair of histograms of the results are shown in the lower plot, where we can see the distribution of output values from the matched filter under the two hypotheses.

The examples of a smoothing filter and a matched filter give us an idea about how FIR digital filters can be specified and used in the processing of noisy signals. At this point we know that a digital filter can be described by either its difference equation or by its impulse response, and we know how to relate the impulse response of a FIR filter to its difference equation. In the case of the matched filter of Example 7.1.3 we used the impulse response to determine the FIR filter parameters.

7.1.2 System Function Model

It will be useful to have other ways to describe and specify the behavior of digital filters. In particular, we would like to have a natural method to look at them in the frequency domain.

Continuous-time system

In continuous systems we use the system function $H(s)$ or $H(j\omega)$ to describe the frequency response. The system function describes the response of a system to the special input $x(s, t) = e^{st}$. The response is $y(s, t) = H(s)e^{st}$. Here the parameter s is introduced into the arguments of the input and output functions to make it clear that it is a value that can be specified in

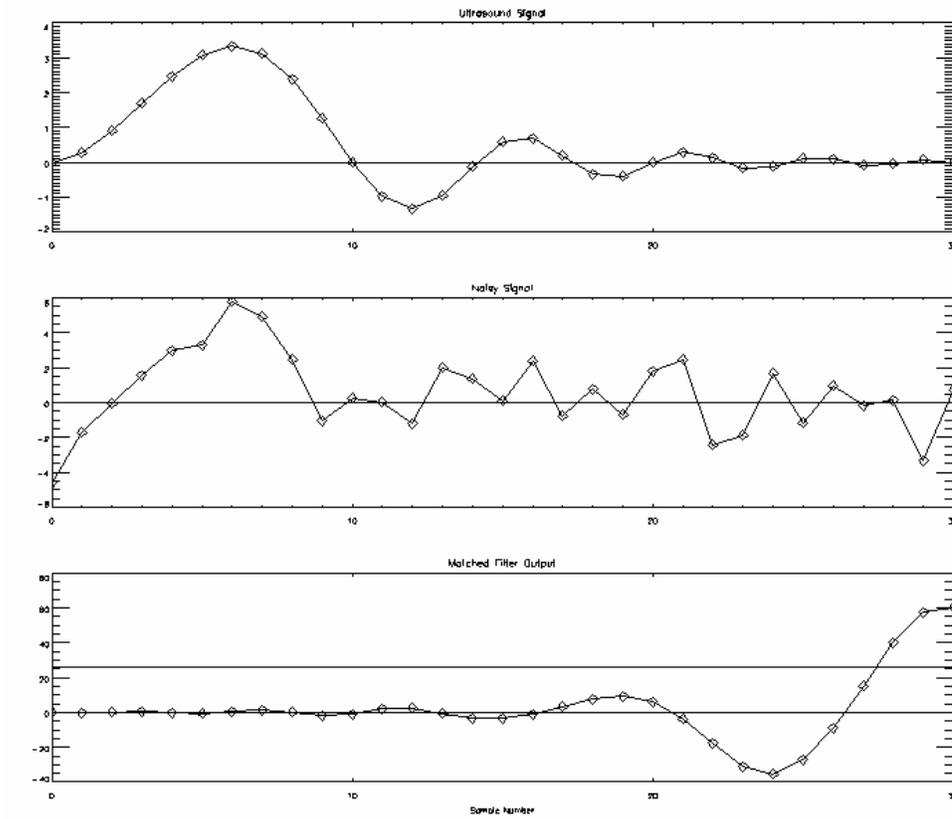


Figure 7.5: The signal is shown in the top panel. A typical received signal plus noise is shown in the middle panel. The response of a matched filter to the waveform signal plus noise waveform is shown in the lower panel. The final value of the filter output is used in the decision rule.

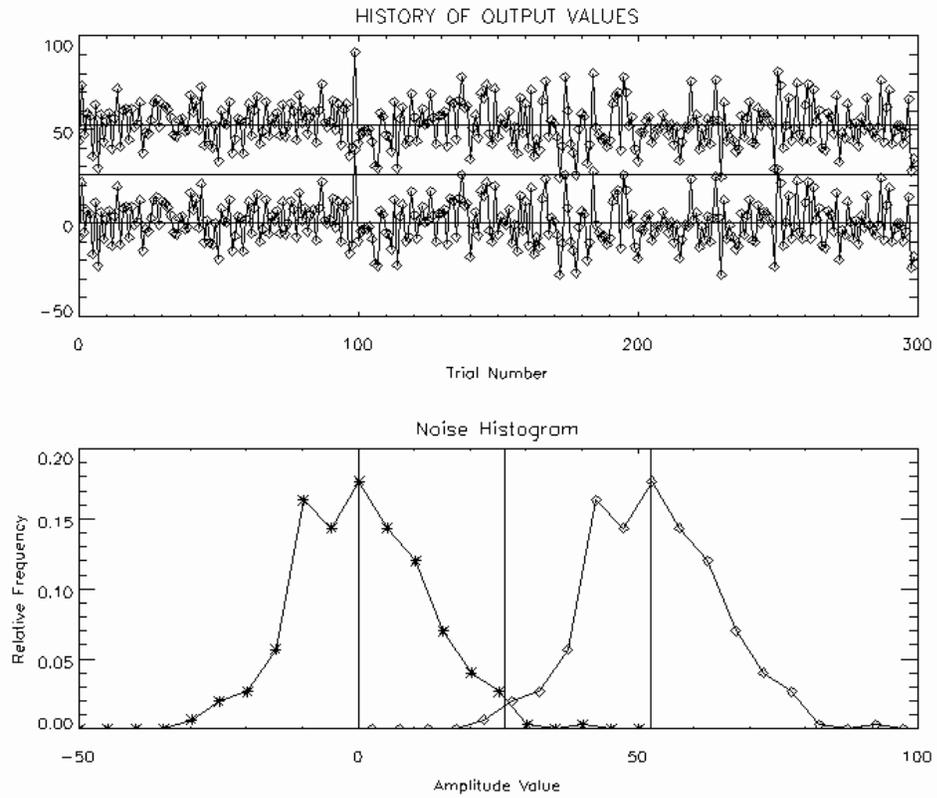


Figure 7.6: The upper panel shows two plots of the final matched filter value for a sequence of 300 trials of detecting the ultrasound signal in white noise. The upper trace is for signal plus noise and the lower trace with only noise. The lower panel shows a histogram of the filter output values for the two cases. The vertical markers are at 0, $\mathcal{E}_s/2$, and \mathcal{E}_s .

our analysis. This approach to finding the system function is valid for *any* linear time-invariant system.

Example 7.1.4 Consider the continuous-time system that is described by a differential equation

$$\begin{aligned} y(t) + a_1y^{(1)}(t) + a_2y^{(2)}(t) + \cdots + a_py^{(p)}(t) \\ = b_0x(t) + b_1x^{(1)}(t) + \cdots + b_qx^{(q)}(t) \end{aligned} \quad (7.15)$$

Let us substitute $x(t) = e^{st}$ and $y(t) = H(s)e^{st}$. Then $x^{(k)}(t) = s^k e^{st}$ and $y(t) = s^k H(s)e^{st}$, and we find

$$\begin{aligned} (1 + a_1s + a_2s^2 + \cdots + a_ps^p)H(s)e^{st} \\ = (b_0 + b_1s + \cdots + b_qs^q)e^{st} \end{aligned} \quad (7.16)$$

This equation will be valid for all values of t provided the coefficients of e^{st} on each side of the equation are equal. This allows us to solve for the system function.

$$H(s) = \frac{b_0 + b_1s + \cdots + b_qs^q}{1 + a_1s + a_2s^2 + \cdots + a_ps^p} \quad (7.17)$$

This function is valid for all values of s for which the denominator is nonzero, including complex and imaginary values. This tells us that any system that is described by a linear differential equation with constant coefficients has a system function that is a ratio of polynomials in s . The coefficients are the coefficients from the differential equation.

Example 7.1.5 Consider a continuous-time linear system that is described by an impulse response function $h(t)$. The response of the system to any input $x(t)$ is

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t - \tau)d\tau \quad (7.18)$$

Now let $x(t) = e^{st}$ and $y(t) = H(s)e^{st}$. Then we must have

$$H(s)e^{st} = \int_{-\infty}^{\infty} h(\tau)e^{s(t-\tau)}d\tau = e^{st} \int_{-\infty}^{\infty} h(\tau)e^{-s\tau}d\tau$$

Therefore,

$$H(s) = \int_{-\infty}^{\infty} h(\tau)e^{-s\tau}d\tau \quad (7.19)$$

This result shows that the system function and the impulse response are a transform pair for time-invariant continuous-time linear systems. Either function is just an alternative way of looking at the behavior of the system.

A major use for the system function is in describing the behavior of the system in the frequency domain. Every sinusoidal waveform can be written in terms of exponentials. For example,

$$\cos(\omega t) = \frac{1}{2}(e^{j\omega t} + e^{-j\omega t})$$

Hence, the response to $x(t) = \cos(\omega t)$ is $y(t) = \frac{1}{2}H(j\omega)e^{j\omega t} + \frac{1}{2}H(-j\omega)e^{-j\omega t}$. For a linear system with real parameters, such as real coefficients in the differential equation or a real impulse response, the response to a real input, such as a cosine wave, must also be real. This requires that the system function have conjugate symmetry. $H(-j\omega) = H^*(j\omega)$. Every complex number can be written in a polar coordinate form. If Z is a complex number then it can be written in the form $Z = |Z|e^{j\theta}$. We can apply this to the system function since it is just a complex number at any fixed value of ω . Hence the output is $y(t) = \frac{1}{2}|H(j\omega)|e^{j(\omega t + \theta)} + \frac{1}{2}|H(j\omega)|e^{-j(\omega t + \theta)} = |H(j\omega)|\cos(\omega t + \theta)$. Therefore, $|H(j\omega)|$ is the “gain” of the system at frequency ω and θ is the phase shift at frequency ω .

Example 7.1.6 *Suppose that we have a system for which we know that the impulse response is the simple exponential function*

$$h(t) = e^{-at}u(t)$$

where $u(t)$ is the unit step function and a is a positive constant. The system function is

$$H(s) = \int_0^{\infty} e^{-at}e^{-st}dt = \frac{1}{s+a}$$

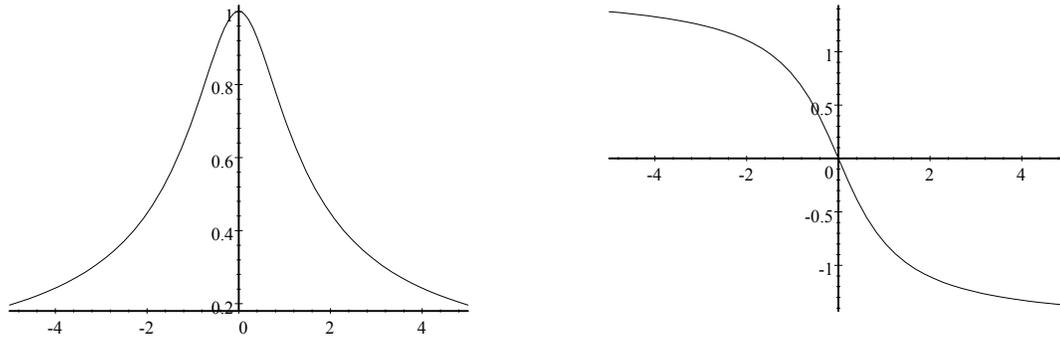
The analysis of Example 7.1.4 shows that the difference equation must be

$$x(t) = 1 + a\frac{dy}{dt}$$

The value of the system function for $s = j\omega$ is

$$H(j\omega) = \frac{1}{a + j\omega} = \frac{1}{\sqrt{a^2 + \omega^2}}e^{-j\text{Arctan}(\omega/a)}$$

The response to $x(t) = \cos(\omega t + \alpha)$ will be $y(t) = \frac{1}{\sqrt{a^2 + \omega^2}} \cos(\omega t + \alpha - \text{Arctan}(\omega/a))$.



Magnitude of system function vs ω/a

A plot of the magnitude and phase of the system function shows that it passes frequencies below about $\omega = a$ and rejects higher frequency. .

Discrete-time system

The system function for a discrete-time system is based on the difference equation rather than the differential equation. Just as in a continuous system, an exponential input to a discrete system will produce an exponential response. The difference equation is reproduced below from (7.1)

$$y(n) + a_1 y(n-1) + a_2 y(n-2) + \dots + a_p y(n-p) = b_0 x(n) + b_1 x(n-1) + \dots + b_q x(n-q)$$

Let us now substitute the exponential sequence $x(n) = z^n$. Here z is a complex number, which may be expressed in the usual complex number forms wherever that becomes useful in the analysis. Let us assume that the response is of the form $y(n) = H(z)z^n$. Substitution into the difference equation then produces

$$(1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_p z^{-p}) z^n H(z) = (b_0 + b_1 z^{-1} + \dots + b_q z^{-q}) z^n$$

This can be solved for the system function

$$H(z) = \frac{b_0 + b_1 z^{-1} + \dots + b_q z^{-q}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_p z^{-p}} \quad (7.20)$$

The close relationship between the system function, the difference equation and the block diagram shown in Figure 7.1 on page 128 is evident. The system function is defined wherever the denominator is not zero. These locations are called the “poles” of the system function.

z -Transform The z -transform plays the same role in discrete system representation that the Laplace transform plays in continuous system representation. For any sequence $x(n)$.

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n} \quad (7.21)$$

The inverse z -transform can be defined. That is not necessary in this analysis, and we will simply illustrate the computation of the z -transform for some example functions.

Example 7.1.7 Let $x(n) = A^n u(n)$, where $u(n)$ is the unit step function and A is a constant. Then

$$X(z) = \sum_{n=0}^{\infty} A^n z^{-n} = \sum_{n=0}^{\infty} (A/z)^n$$

This is just a geometric series of the form $(1 + r + r^2 + \dots)$, which converges to $1/(1 - r)$ provided $|r| < 1$. Hence,

$$X(z) = \frac{1}{1 - A/z} = \frac{z}{z - A}$$

The denominator has a root at $z = A$, which corresponds to the ratio in the exponential sequence.

Delay Operator

Suppose that the z -transform of $x(n)$ is $X(z)$. Then, the z -transform of $x(n - k)$ is $X(z)z^{-k}$. This can be shown by substitution of $x(n - k)$ into (7.21). We can therefore refer to z^{-1} as the delay operator.

Convolution Property

Let $x_1(n)$ and $x_2(n)$ be two sequences with z -transform relationships

$$\begin{aligned} x_1(n) &\leftrightarrow X_1(z) \\ x_2(n) &\leftrightarrow X_2(z) \end{aligned}$$

Let $x(n) = x_1(n) * x_2(n)$ be the convolution operation. Then the z -transform is determined by the product $X(z) = X_1(z)X_2(z)$.

Impulse Function

The function $\delta(n)$ defined in (7.4) has the z -transform $\Delta(z) = 1$, as can be easily established from the definition (7.21).

Impulse Response

The impulse response of a discrete linear time-invariant system is called $h(n)$. It is the response of the system to an input $\delta(n)$. The output for any other input sequence $x(n)$ is given by the convolution

$$y(n) = \sum_{m=-\infty}^{\infty} x(m)h(n-m) = \sum_{m=-\infty}^{\infty} h(m)x(n-m) \quad (7.22)$$

Upon using the convolution property we find that

$$Y(z) = X(z)H(z) \quad (7.23)$$

In particular, if the input is an impulse then $X(z) = 1$ and the z -transform of the response is just $H(z)$. Therefore, $h(n)$ and $H(z)$ are a transform pair

$$h(n) \leftrightarrow H(z) \quad (7.24)$$

Frequency Response

Just as in the continuous case, we can use an exponential input to a discrete system to determine its frequency response. Let $x(n) = e^{j\omega n}$ be an input sequence. The output, from (7.22) is

$$y(n) = \sum_{m=-\infty}^{\infty} h(m)e^{j\omega(n-m)} = e^{j\omega n} \sum_{m=-\infty}^{\infty} h(m)e^{-j\omega m}$$

The summation is just $H(z)$ with $z = e^{j\omega}$. Hence, an exponential input sequence produces the exponential output sequence

$$y(n) = H(e^{j\omega})e^{j\omega n}$$

where

$$H(e^{j\omega}) = \sum_{m=-\infty}^{\infty} h(m)e^{-j\omega m}$$

The system response $H(e^{j\omega})$ provides the same kind of information about a discrete system that $H(j\omega)$ provides about a continuous system. The position of the frequency variable in the exponent means that the function is periodic with period 2π .

Example 7.1.8 *The system whose block diagram is shown in Figure 7.2 on page 129 has the impulse response*

$$h(n) = (-a)^n u(n)$$

The frequency response of the system is

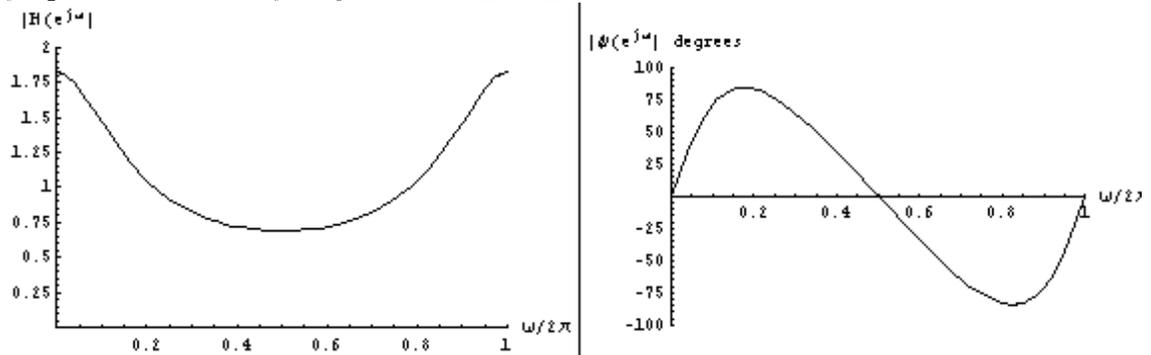
$$H(e^{j\omega}) = \sum_{m=-\infty}^{\infty} (-a)^m e^{-j\omega m} = \sum_{m=-\infty}^{\infty} (-ae^{-j\omega})^m = \frac{1}{1 + ae^{j\omega}}$$

The gain and phase shift of the system at frequency ω are equal the magnitude and phase angle of $H(e^{j\omega})$. These are

$$|H(e^{j\omega})| = \frac{1}{\sqrt{(1+a \cos \omega)^2 + a^2 \sin^2 \omega}}$$

$$\phi(e^{j\omega}) = -\tan^{-1} \frac{a \sin \omega}{1+a \cos \omega}$$

The periodic character of both functions is evident, since all of the terms in each expression are periodic. The gain and phase shift are plotted in Figure ?? where one period of the response is shown for a value $a = -0.45$ for the filter parameter. The system emphasizes low frequencies more than high frequencies. This is a “low-pass filter.” Note that the operational range of frequencies in a useful system is $0 \leq \omega \leq 0.5$.



Gain and phase shift of a single-pole system function with parameter $a = -0.45$.

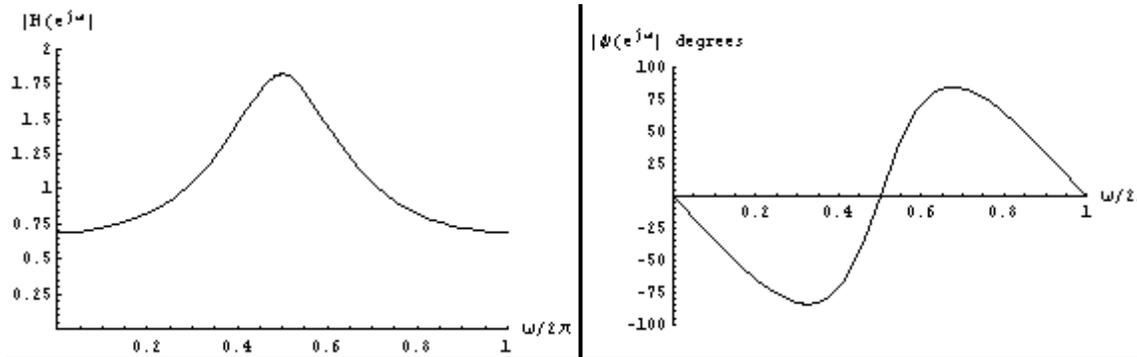


Figure 7.7: Gain and phase shift of a single-pole system function with a parameter $a = 0.45$. Note that this creates a high-pass filter.

The gain and phase-shift of a system with the system function

$$H(e^{j\omega}) = \frac{1}{1 - 2be^{-j\omega} \cos \omega_0 + b^2 e^{-j2\omega}}$$

is shown in Figure 7.8. The location of the peak is determined by the value of ω_0 and $b = 1.01$. One would expect that this filter would have a “narrow-

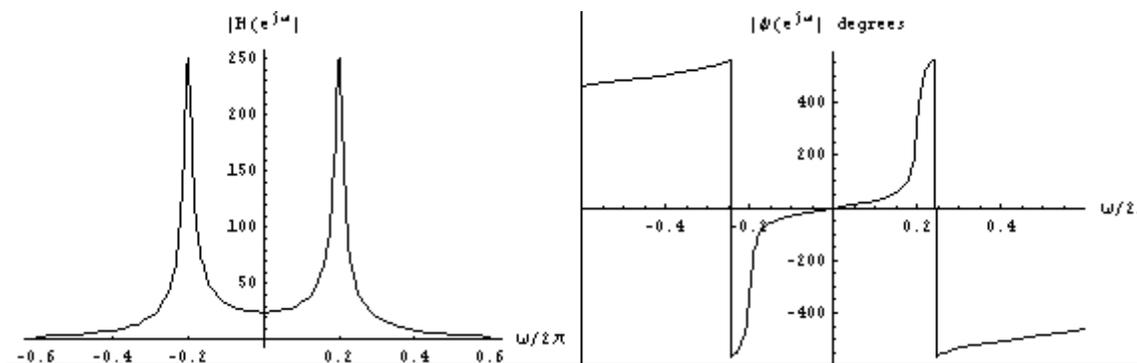


Figure 7.8: Gain and phase shift of a filter with a single harmonic resonance at $\omega_0 = 0.2$ and $b = 1.01$.

band” output, no matter what the input spectrum was.

7.1.3 System Output

The response of a system can be calculated in a number of ways. The difference equation can be implemented directly, transforms can be used, and the output can be calculated by a general implementation of a digital filter. The program `filter` is one useful computational implementation. An example of such a computation is provided by `sigdemo3` and `sigdemo5`. Finally, `sigdemo4` shows how the spectrum estimate can be improved by averaging many repetitions of power spectrum computations.