ECE 310 Fall 2023

Lecture 28

Practical filter design

Corey Snyder

Learning Objectives

After this lecture, you should be able to:

- Describe the key features of a practical LTI filter's magnitude response including passband, stopband, and transition band.
- Understand the key design tradeoffs and desirable features in practical filter design.
- Draw the direct form I and direct form II block diagrams of a filter described by an LCCDE.

Recap from previous lecture

We concluded our discussion of the DFT and FFT last lecture by demonstrating how the FFT may provide an efficient implementation of linear convolution. We now turn our attention to a new practical application: digital filter design. This lecture will establish much of our terminology and design considerations for creating practically realizable digital filters.

1 Ideal digital filters

Thus far, we have indirectly discussed the concept of *filtering* or choosing a *digital filter* for a given application. There are many ways to design and apply filters to discrete-time signals. In this course, we will primarily focus on LTI *frequency-selective filters*. We use the term "frequency-selective" to describe how these filters identify different bands of frequency to pass through or stop. As such, we refer to the *passband* as the range(s) of frequencies which a filter allows to pass with no distortion. Conversely, the *stopband* gives the range(s) of frequencies which the filter suppresses or *attenuates* as strongly as possible. We begin by describing ideal filters that can perfectly pass or stop any given frequency.

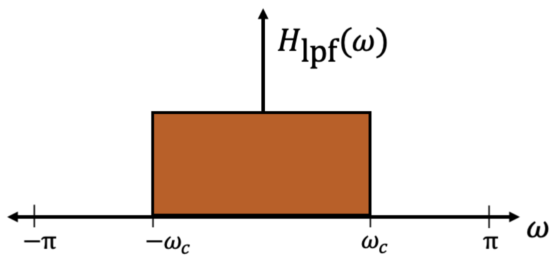
The classic application of discrete-time filters is signal de-noising. In this scenario, an input signal is corrupted with some noise and we design an appropriate filter that will stop the noise frequencies and let the true signal frequencies pass through. Designing and applying filters using frequency-domain analysis is a powerful technique that is a key element of this course.

1.1 Common filter structures

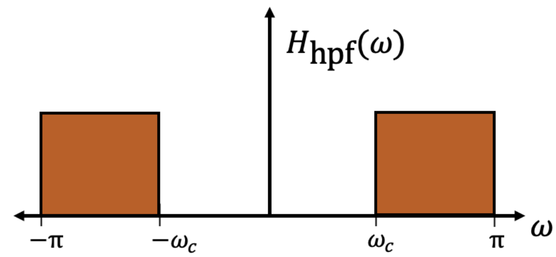
There are a few canonical filter structures that we will define in this section. For each structure, we will give the ideal frequency response for any choice of passband and stopband. Figure 1 provides a summary illustration of these frequency responses.

Low-pass filter. A low-pass filter allows low frequencies to pass while stopping high frequencies:

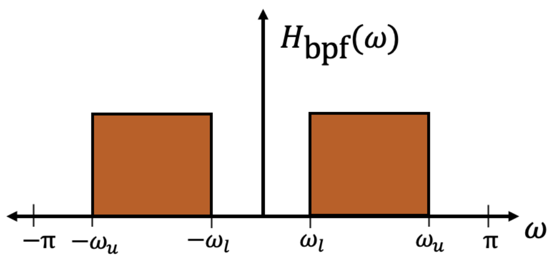
$$H_{\text{lpf}}(\omega) = \begin{cases} 1, & |\omega| \leq \omega_c \\ 0, & \omega_c < |\omega| \leq \pi \end{cases}. \tag{1}$$



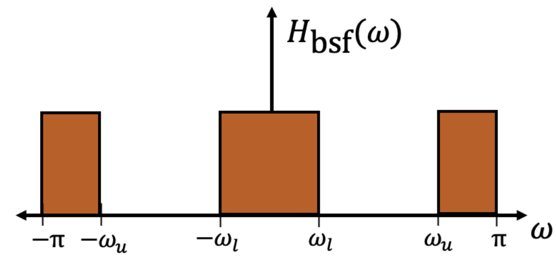
(a) Low-pass filter



(b) High-pass filter



(c) Band-pass filter



(d) Band-stop filter

Figure 1: Illustration of the common digital filter structures

We refer to ω_c as the *cutoff frequency* of a low-pass filter.

High-pass filter. High-pass filters allow high frequencies to pass while stopping low frequencies:

$$H_{\mathrm{hpf}}(\omega) = \begin{cases} 0, & |\omega| \leq \omega_c \\ 1, & \omega_c < |\omega| \leq \pi \end{cases}. \tag{2}$$

High-pass filters and low-pass filters have complementary structures.

Band-pass filter. Different from low-pass and high-pass filters, band-pass filters have two stopbands and one passband:

$$H_{\mathrm{bpf}}(\omega) = \begin{cases} 1, & \omega_l \leq |\omega| < \omega_u \\ 0, & \text{otherwise} \end{cases}. \tag{3}$$

Here, ω_l and ω_u define the lower and upper cutoff frequencies of the passband

Band-stop filter. The band-stop filter is the complementary filter structure to the band-pass filter. Instead, we have two passbands and a single stopband:

$$H_{\mathrm{bsf}}(\omega) = \begin{cases} 0, & \omega_l \leq |\omega| < \omega_u \\ 1, & \text{otherwise} \end{cases}. \tag{4}$$

1.2 Issues with ideal filters

We must consider whether we can use these ideal filters in practice. For example, we may look at the low-pass filter $H_{\mathrm{lpf}}(\omega)$ and obtain the corresponding impulse response via the inverse DTFT to find that

$$h_{\mathrm{lpf}}[n] = \frac{\sin(\omega_c n)}{\pi n}, \quad -\infty < n < \infty. \tag{5}$$

There are a few issues then with using this ideal filter:

1. **The impulse response is infinitely long**. We cannot use the convolution sum for infinitely long impulse responses.

2. **The system is non-causal**. This ideal impulse response extends to $-\infty < n < \infty$. For real-time applications, we will be unable to use this filter since it requires future samples. We may shift our filter and truncate its length as a practical solution and in fact we will discuss this in later lectures about practical LTI filter design. This will introduce linear phase (and uniform group delay) to our frequency response and fix the infinite length issues. Still, such a workaround represents only an approximation of our ideal low-pass filter so this solution does not come for free!

3. **The system is not BIBO stable**. The impulse response $h_{\text{lpf}}[n]$ is not absolutely summable; therefore, the ideal low-pass filter is not BIBO stable. This point is of more theoretical importance than practical importance, however, since an infinitely long signal would be necessary to yield an unbounded output.

Note that the above concerns apply equally to the other three filter structures because we can obtain each of the other filters by modifying a low-pass filter:

- High-pass filters: We can create a high-pass filter using $H_{\text{hpf}}(\omega) = 1 - H_{\text{lpf}}$.

- Band-pass filters: We can modulate a low-pass filter to create a band-pass filter using $h_{\text{bpf}}[n] = \cos(\omega_0 n) h_{\text{lpf}}[n]$.

- Band-stop filters: We can modulate a low-pass filter to create a band-pass filter then obtain $H_{\text{bsf}}(\omega) = 1 - H_{\text{bpf}}$.

In our filter design lectures, we will commonly focus on low-pass filters for this reason, but all the practices we develop apply well to other canonical filter structures.

2 Digital filter design considerations

We cannot achieve these ideal filter structures in practice; thus, we must consider how practical filters may look and define key terminology that will guide our discussion of digital filter design. Figure 2 demonstrates many of these key terms we must define by illustrating the magnitude response of a low-pass filter on a decibel (dB) scale.

Passband. A passband is any range of frequencies we let pass through from the input signal. Note that we may have multiple passbands like in a band-stop filter.

Passband ripple. Passband ripple is the amount of variation in the gain of a passband. In Fig. 2, we see that the intended gain is 0 dB (gain of 1 on a linear scale). The passband ripple then tells us how much the gain of the passband deviates from this desired 0 dB gain. We want passband ripple to be as small as possible and zero ideally like with any ideal filter structure. You may sometimes see the symbol δ_p to denote passband ripple on a dB scale.

Stopband. A stopband is any range of frequencies we attenuate or remove from the input signal. Note that we may have multiple stopbands like in a band-pass filter.

Stopband attenuation. The stopband attenuation is the highest gain achieved by a given stopband. We would like our stopband attenuation to be as strong as possible, i.e. the lowest gain possible or a very negative number in dB. Thus, the stopband attenuation sets the worst case for how much we reduce the magnitude of undesirable frequencies in any stopband. Figure 2 shows a filter with a stopband attenuation of approximately -23dB. Since we calculate dB as $20\log_{10}(|H(\omega)|)$, this corresponds to roughly a 1/10 reduction in magnitude.

Transition band. A transition band is any range of frequencies that transition the magnitude response between a passband and a stopband. We define any transition band as lying between a pair of passband

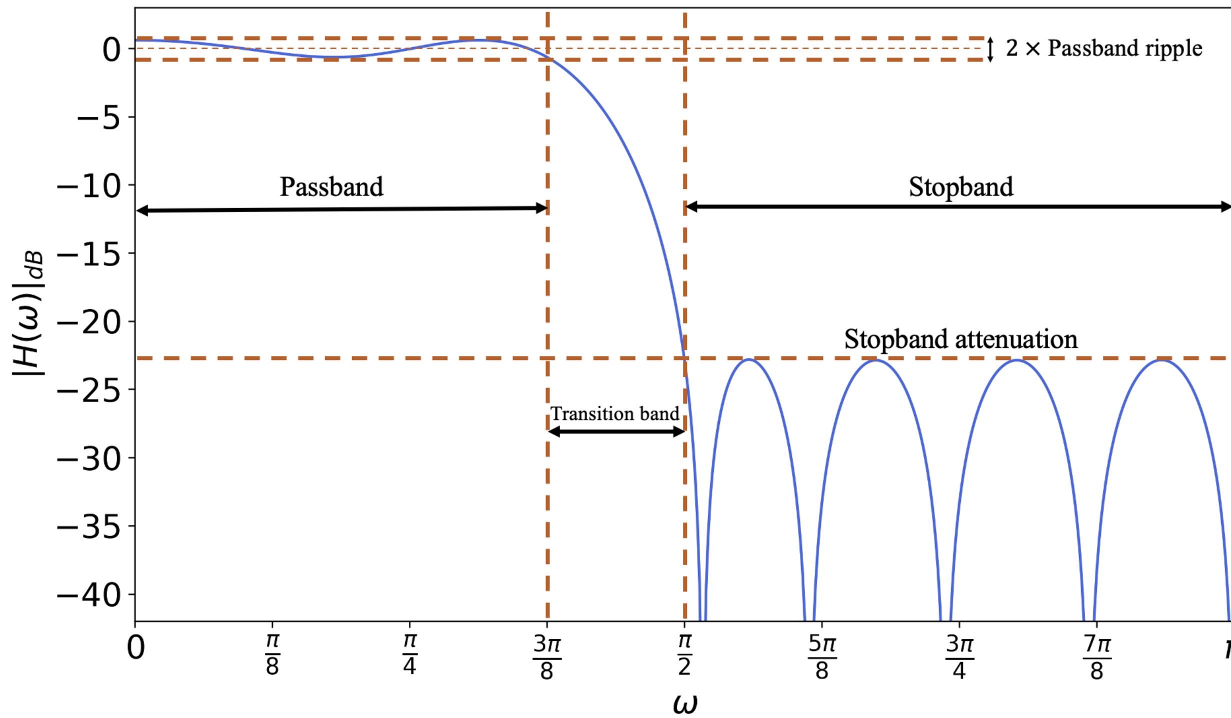


Figure 2: Illustration of key specifications for practical digital filters.

and stopbands edges, e.g. end of a passband and start of a stopband. With ideal digital filters, we had no transition bands since there were perfectly sharp cutoffs between passband and stopband. We may also have multiple transition bands in a given frequency response.

Transition bandwidth. The transition bandwidth of a transition band is the frequency separation between the passband and stopband edges that outline the transition band. We would like for transition bandwidth to be as small as possible since this reduces the range of frequencies that we do not clearly pass or stop. The transition band in Fig. 2 has a transition bandwidth of $\pi/8$.

For practical filter design, we are trying to achieve a frequency response that is close to an ideal filter that performs the desired frequency selective operation. We mentioned above how we would like small passband ripple, strong stopband attenuation, and narrow transition bandwidth. All these factors push a practical filter closer to its ideal counterpart; however, we know that in practice they are not all achievable. Most often, we are concerned about how to balance stopband attenuation and transition bandwidth against the length or complexity of the digital filter. We will explore these tradeoffs in upcoming lectures. For now, we conclude with two common representations we use to describe and visualize practical digital filters.

3 Block diagrams

We can express any practical digital filter, finite impulse response (FIR) or infinite impulse response (IIR), as a linear constant coefficient difference equation (LCCDE). Recall that we express a causal LCCDE with input signal $x[n]$ and output signal $y[n]$ as

$$y[n] = \sum_{i=1}^{K} b_i y[n-i] + \sum_{j=0}^{M-1} c_j x[n-j],\ 0 \leq K < \infty,\ 1 \leq M < \infty. \tag{6}$$

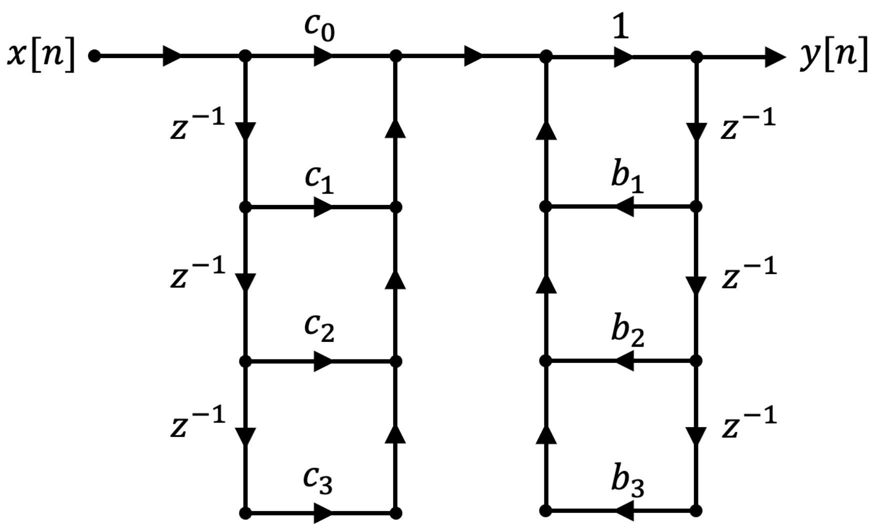


Figure 3: Direct form I diagram for the LCCDE of Eqn. 7.

Furthermore, we introduced block diagrams as a way of depicting any LCCDE. We will return to those block diagrams now, but instead use a slightly simplified notation known as *flow diagrams*. We will demonstrate two forms known as *direct form I* and *direct form II*, though others do exist. To do so, we will illustrate the following LCCDE with both forms when $K = 3$ and $M = 4$:

$$y[n] = b_1 y[n-1] + b_2 y[n-2] + b_3 y[n-3] + c_0 x[n] + c_1 x[n-1] + c_2 x[n-2] + c_3 x[n-3]. \tag{7}$$

3.1 Direct form I

The direct form I structure is perhaps the most intuitive digital filter flow graph. We partition the diagram into an input half followed by an output/feedback half. Figure 3 depicts the direct form I structure of Eqn. 7. We denote any multiplication or delay using a triangle labeled with the given coefficient or delay symbol, respectively. Any triangle without a label may be assumed to be identity or multiplication by one. Any junction is either a point for a signal to split in two directions or for signals to be added where they meet. Altogether, we can count the number of shifts, multiplications, and additions necessary for each filter structure. For Fig. 3, we have seven multiplies (ignoring multiplication by one), six additions, and six delays.

3.2 Direct form II

The direct form II structure is less straightforward, however, we may find some savings in filter implementation. Figure 4 shows the direct form II implementation of Eqn. 7. We can show how this structure is equivalent to the desired LCCDE by considering the intermediate labeled signal $v[n]$. First, we note that the desired transfer function $H(z)$ for this LCCDE is given by

$$H(z) = \frac{c_0 + c_1 z^{-1} + c_2 z^{-2} + c_3 z^{-3}}{1 - b_1 z^{-1} - b_2 z^{-2} - b_3 z^{-3}}. \tag{8}$$

Now, we will express $x[n]$ and $y[n]$ in terms of $v[n]$, respectively. Looking at Fig. 4, we observe that

$$v[n] = x[n] + b_1 v[n-1] + b_2 v[n-2] + b_3 v[n-3] \tag{9}$$
$$y[n] = c_0 v[n] + c_1 v[n-1] + c_2 v[n-2] + c_3 v[n-3]. \tag{10}$$

Taking the z-transformof both sides and rearranging, we have

$$X(z) = V(z)(1 - b_1 z^{-1} - b_2 z^{-2} - b_3 z^{-3}) \tag{11}$$

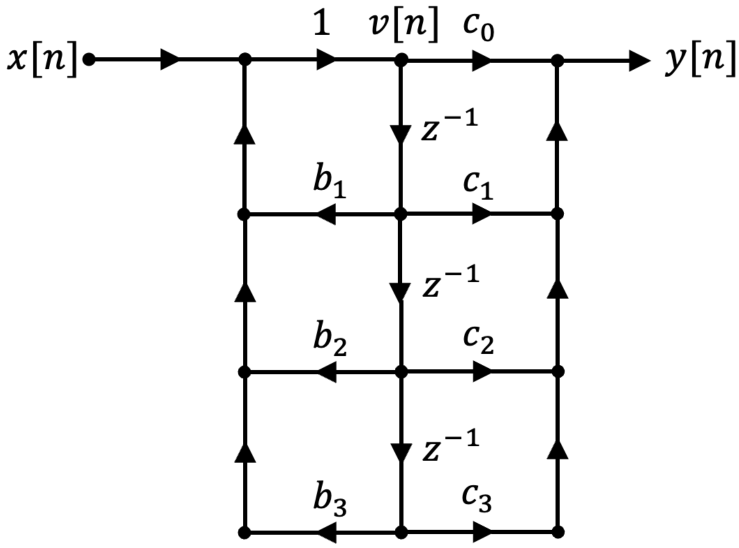


Figure 4: Direct form II diagram for the LCCDE of Eqn. 7.

$$Y(z) = V(z)(c_0 + c_1 z^{-1} + c_2 z^{-2} + c_3 z^{-3}). \tag{12}$$

We then can compute the transfer function via

$$H(z) = \frac{Y(z)}{X(z)} = \frac{c_0 + c_1 z^{-1} + c_2 z^{-2} + c_3 z^{-3}}{1 - b_1 z^{-1} - b_2 z^{-2} - b_3 z^{-3}}, \tag{13}$$

and we find the same transfer function! This direct form II structure has the benefit of requiring seven multiplies, six additions multiplies, but now only three delays. We may derive a direct form II diagram from a direct form I diagram by overlapping the first and second halves of the direct form I diagram along their shared delay paths. Doing so will move the c_j coefficients to the right side of the diagram and the b_i coefficients to the left side.