Chapter 1

Introduction to Filter Design

1.1 Introduction

The study of *filter design* is one specific area within the discipline of *signal processing*. The word "filter" has Medieval Latin origins, from the word *filtrum*, or "felt," which was the material typically used to filter impurities from liquids. Today's definition is not conceptually different. We recognize that we use paper filter elements to hold back coffee grounds, we use filters to cleanse lubricating oil in our cars, and we use filters to purify air through air conditioning or heating systems in our homes. These common filters are examples of *discriminators*, in that they treat particles of different sizes very differently. They allow particles of small dimension (perhaps air or a liquid) to pass through, while holding back larger particles (perhaps dirt).

In the discipline of electrical engineering, when we refer to a "filter," we usually imply that the filter *discriminates* as a function of *frequency*. To illustrate such discrimination, consider the ideal *low pass filter* is discriminatory in that it allows lower frequency sinusoidal components to "pass through," and it stops higher frequency sinusoidal components from passing through. Similarly, a *high pass filter* is one that "prevents" lower frequency components and allows higher frequency signal components to "pass through." An ideal band pass filter is one which "passes" signal components within a band of frequencies, and stops signal components outside that band. Likewise, an ideal "notch" or band-reject filter is one which stops signal components within a band of frequencies more than those signal components outside that band. Finally, there is a class of filters called "all pass" filters, which essentially allows all frequency components to pass through. (In that case, the filter typically provides some desired phase response characteristic, while generally having a relatively constant magnitude response.)

In describing filter characteristics, it is sometimes useful to characterize a filter by its *frequency response*, which is a function that expresses exactly how the filter responds to sinusoidal inputs at different frequencies. The magnitude of the filter's frequency response, or *magnitude response*, is a graph of the filter's gain (or amplification) as a function of frequency. The phase of the filter's frequency response, or *phase response*, is a function which expresses the shift in phase from input to output caused by the filter as a function of frequency. An equally valid time domain characterization of a filter may be given by the filter's *impulse response*. The impulse response of a filter (or system) is, quite simply, the time domain response of that filter (or system) to an impulse input. In the case of a linear time-invariant (LTI) *analog* filter, a third way to characterize it is to describe the output (and its derivatives) as a function of the input (and its derivatives) using linear differential equation(s) with constant coefficients. In the case of a linear time-invariant (LTI) discrete-time filter, we can characterize this filter by difference equation(s) with constant coefficients.

1.2 The Ideal Low Pass Filter

With no loss in generality, we will consider a continuous-time ideal low-pass filter, and we will also consider the case where we have a linear phase characteristic (or "ideal delay" of T_D). An ideal low-pass filter is defined as a filter that exhibits a constant gain (K) up to some cutoff frequency, and a gain of zero beyond that frequency. The magnitude response of an ideal low pass continuous-time filter is shown in Figure 1.1. A logical question to ask would be, "… is it possible to design and build such a filter?" To answer this question, we will use the inverse Fourier transform to examine the nature of such a filter's impulse response. First, recognize that the frequency response of an ideal low-pass filter with linear phase characteristic is given by

$$H(f) = Ke^{-j2\pi fT_D}, |f| < f_c$$

and

$$H(f) = 0, otherwise \ . \tag{1}$$

The continuous time filter's impulse response is given by

$$h(t) = \int_{-\infty}^{\infty} H(f) e^{j2\pi ft} df \tag{2}$$

$$= \int_{-f_c}^{f_c} K e^{-j2\pi fT_D} e^{j2\pi ft} df$$

$$= \int_{-f_c}^{f_c} K e^{j2\pi f(t-T_D)} df$$

$$= \frac{K}{j2\pi(t-T_D)} \left(e^{j2\pi f_c(t-T_D)} - e^{-j2\pi f_c(t-T_D)} \right)$$

$$= \frac{K}{\pi(t-T_D)} \sin\left[2\pi f_c(t-T_D)\right] \tag{3}$$

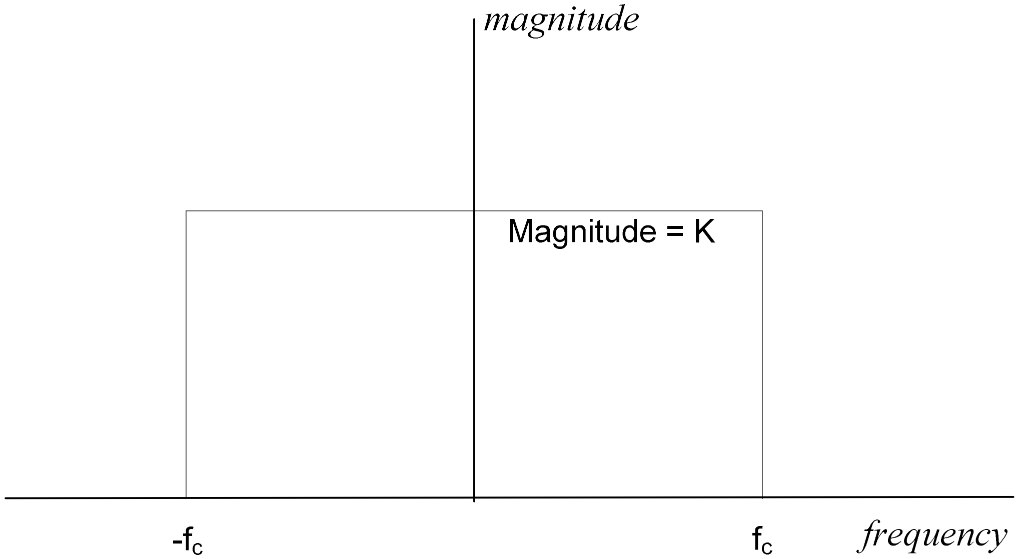

Figure 1.1. Magnitude response for ideal low pass filter.

where T_D represents some amount of pure delay (equating to a linear phase shift term of $e^{-j2\pi fT_D}$ in the frequency domain.) For purposes of illustration, a plot of h(t) from Equation (3) is shown in Figure 1.2, for T_D =2.5seconds, K=1, and f_c =1. Note that we constrained the plot to a time period between -5 and 5 seconds, however the impulse response extends over infinite time, for both positive and negative t, thereby violating causality. Said another way, *one cannot design a causal, ideal, low pass filter*, because a band-limited frequency response corresponds to an infinitely wide impulse response!

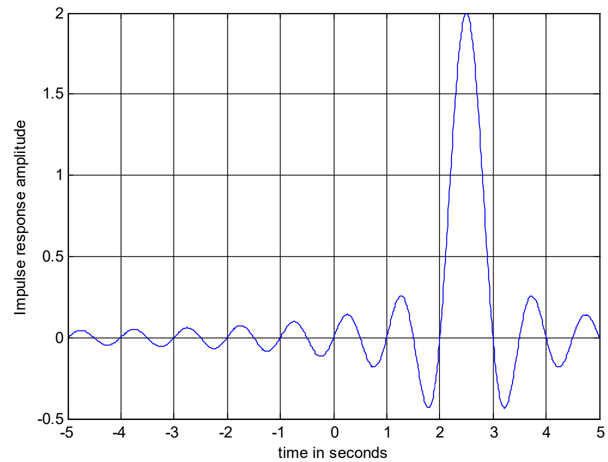

Figure 1.2. Impulse response for ideal low pass filter. Observe this impulse response is non-causal.

1.3 The *Real-World* Filter

"Real-world" filters are filters whose impulse responses or frequency responses *approximate* the response of an ideal filter. Naturally there are a number of possible *best approximations* to an ideal response, depending on how one defines *best*, and it is these various approximations that will be studied in subsequent chapters.

Before we embark on designing filters, we must first understand filter *specifications*. Filter specifications are generally derived for a specific filtering application, and are usually expressed in terms of pass band limit(s) and stop band limit(s). Figure 1.3 shows the magnitude response for a "real-world" low pass filter, where ω_p represents the upper extent of the pass band, and ω_s represents the lower extent of the stop band. The frequency range $0 < \omega < \omega_p$ is called the filter ***pass band***, because Fourier signal components presented to the filter input within the range of $0 < \omega < \omega_p$ are generally "passed through" to the filter output. The frequency range $\omega_p < \omega < \omega_s$ is called the ***transition band***, because the filter magnitude response "transitions" from the pass band to the stop band in this region. The frequency range $\omega > \omega_s$ is called the ***stop band*** of the filter, because Fourier signal components presented to the filter input within that range are generally "stopped," or more precisely, they are *attenuated*. Attenuation is defined as the inverse of gain, or the negative of the gain if gain is measured on a logarithmic scale (as we will soon see!).

Typically, filter gain or attenuation specifications are given in terms of "dB," which can be considered a logarithmic measure of gain or attenuation. More specifically, if we define "gain" at frequency ω_o as the magnitude of the ratio of voltage output to voltage input at frequency ω_o,

$$|G(\omega_o)| = \left|\frac{V_{out}(\omega_o)}{V_{in}(\omega_o)}\right| ,$$

we can specify this gain in dB as

$$G_{dB}(\omega_o) = 20\log_{10}|G(\omega_o)| = 20\log_{10}\left|\frac{V_{out}(\omega_o)}{V_{in}(\omega_o)}\right| .$$

Since attenuation is the reciprocal of gain (on a linear scale), or the negative of gain (on a dB scale), we can say that

$$\left|A(\omega_o)\right| = \left|\frac{V_{in}(\omega_o)}{V_{out}(\omega_o)}\right| \; ,$$

and

$$A_{dB}(\omega_o) = 20\log_{10}\left|A(\omega_o)\right| = 20\log_{10}\left|\frac{V_{in}(\omega_o)}{V_{out}(\omega_o)}\right| = -G_{dB}(\omega_o) \, .$$

Hence, for a filter with maximum pass band gain of G_{max}=0dB (shown below in Figure 1.3), we might speak of a filter having 20dB of *attenuation* at a specific frequency (perhaps in the stop band), which would mean the filter would have a *gain* of -20dB at that frequency. (Remember, "linear" gain and attenuation are reciprocals of one another, meaning that gain and attenuation expressed in dB are negatives of one another.)

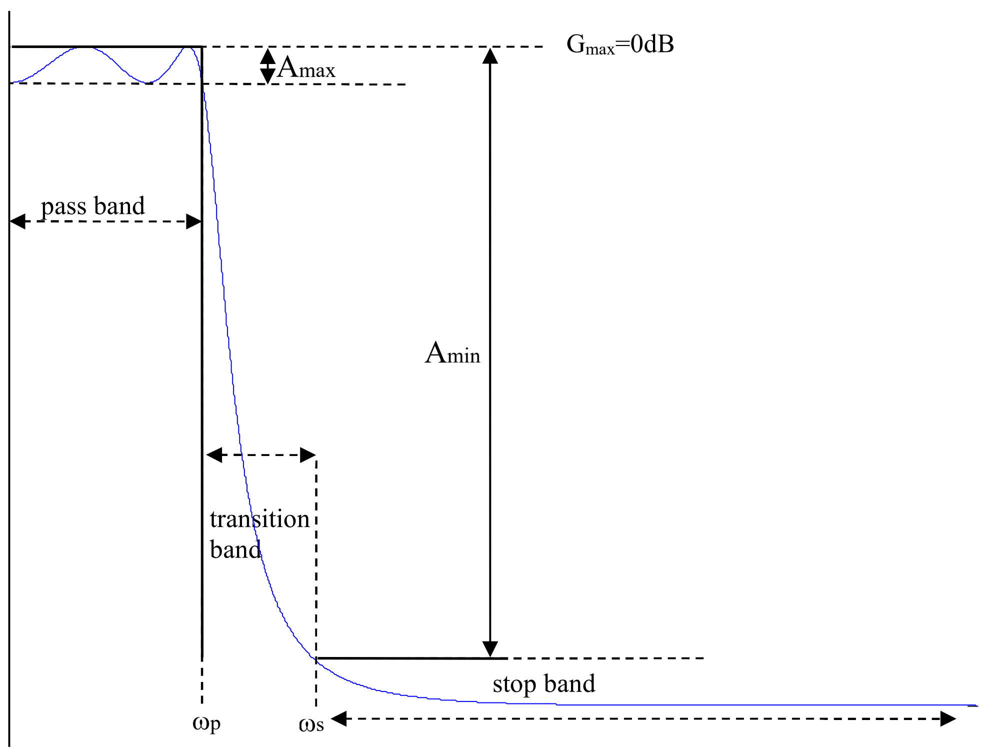

Figure 1.3. Frequency response magnitude (gain) for low pass filter (drawn on a linear frequency axis).

In Figure 1.3 we show a typical magnitude response characteristic for a low pass filter, assuming the maximum gain G_{max} is 0dB. In addition to ω_p and ω_s, A_{max} and A_{min} are part of your filter specifications. A_{max} is typically defined as the *maximum allowed pass band attenuation, in dB.* A_{min} is typically defined as the *minimum acceptable stop band attenuation, in dB.* (Of course, in the case of an ideal low pass filter, we would like a perfectly flat pass band with A_{max}=0, a perfectly flat stop band with infinite attenuation, and a transition band of zero width, but we also know that we cannot design a perfect filter!) If we allow a relaxed constraint on A_{max}, we are *increasing* the value of A_{max}, and we allow our filter approximation to be a poorer approximation to the ideal filter in the pass band. Similarly, if we allow a relaxed constraint on A_{min}, we are *decreasing* the value of A_{min}, and allowing our filter approximation to be a poorer approximation to the ideal filter in the stop band. Recall that an ideal filter specification would be $\omega_p = \omega_s$, A_{max} = 0dB, and $A_{min} = \infty$ dB.

In later chapters we will discover that these four specifications (ω_p, ω_s, A_{max} and A_{min}), coupled with the desired filter approximation, will determine completely the required filter order, or complexity, of a low pass or high pass filter that meets a set of desired specifications. Finally, we will see in later chapters that it is sometimes convenient to plot a filter's magnitude response in dB on a logarithmic frequency axis. On these plots, it is often helpful to observe the response over a "decade" (or multiple decades), or over an "octave" (or multiple octaves) of frequency. A *decade* change in frequency is defined as a change in frequency by a factor of 10 (either up or down). An *octave* change in frequency is defined as a change in frequency by a factor of 2.

Often the magnitude response vertical scale is logarithmic, and gain is expressed in decibels, or dB. A dB is typically defined as a logarithmic measure of a ratio of "power out" to "power in" for a system, so that gain in dB (G_{dB}) is typically defined as

$$\boxed{G_{dB} = 10\log_{10}\left(\frac{P_{out}}{P_{in}}\right)} \tag{4}$$

Recognize that one can also calculate the ratio of powers from (4) in terms of input and output voltages, so that

$$G_{dB} = 10\log_{10}\left(\frac{{V_{out}}^2 / R_{out}}{{V_{in}}^2 / R_{in}}\right) \tag{5}$$

Notice that power delivered is proportional to the square of voltage, and if resistance is held constant (i.e. $R_{out} = R_{in}$), then

$$G_{dB} = 20\log_{10}\left|\frac{V_{out}}{V_{in}}\right| \tag{6}$$

which is consistent with the expression we had earlier in this section.

****

Example 1.1. You are given a system such that the measured output power is 1W when the input power supplied is 2W. Express the system's gain in dB.

Solution: By definition, the system's gain in dB is given by $G_{dB} = 10\log_{10}\left(\frac{P_{out}}{P_{in}}\right) = 10\log_{10}\left(\frac{1}{2}\right)$ = -3.01dB.

This means that the output power is 3dB below the input power, or the signal has been attenuated by 3dB.

****

Example 1.2. You have designed an audio amplifier such that the RMS voltage level on the output is 10 times larger than the input voltage level, over a specific range of frequencies. Express the system's gain in terms of dB.

Solution: The system's gain is

$G_{dB} = 20\log_{10}\left(\frac{V_{out}}{V_{in}}\right) = 20\log_{10}\left(\frac{10x}{x}\right) = 20dB$.

****

Example 1.3. A sinusoidal voltage waveform of 1.0V peak is applied to the input to a filter. The filter output is a phase shifted sinusoid of the same frequency, with amplitude of 0.5V peak. Express the filter's gain in terms of dB.

Solution: The system's gain is $G_{dB} = 20\log_{10}\left(\frac{V_{out}}{V_{in}}\right) = 20\log_{10}\left(\frac{0.5}{1}\right) = -6.02dB$.

We can say that the filter gain is -6dB at this frequency, or equivalently, the filter provides 6dB attenuation at this frequency.

****

Example 1.4. (a) How many decades is a factor of 100 in frequency? (b) How many decades corresponds to one octave? (c) How many octaves corresponds to one decade?

Solution: a) $10^x = 100$, so x=2 decades. b) $10^x = 2$, so $x=\log_{10}(2)=.3$ decades. c) $2^x = 10$, so $x= \log_2(10)=3.3$ octaves.

Problems

Problems 1.1-1.4: For the problems listed below, determine the following information:

a) filter type (low pass, high pass, band pass, notch)
b) range of frequencies for pass band(s)
c) range of frequencies for stop band(s)
d) range of frequencies for transition band(s)
e) sketch of magnitude response showing specific frequencies and gains.

(Note below that HF means "high frequency," DC means "zero frequency," CF means "center frequency, and Max Gain means maximum pass band gain).

1.1 DC Gain= Max Gain =0 dB, Amax=1.0dB, Amin=30.0dB, ω_p=1000 rad/sec, and ω_s=3000 rad/sec.

1.2 HF Gain = Max Gain = 20 dB, Amax=0.5dB, Amin=30 dB, ω_s=2000 rad/sec, and ω_p=4000 rad/sec.

1.3 CF Gain = Max Gain = 10dB, Amax=2.0dB, Amin=20 dB, ω_{p_1}=2000 rad/sec, ω_{p_2}=4000 rad/sec, ω_{s_1}=1000 rad/sec. and ω_{s_2}=8000 rad/sec.

1.4 DC Gain = HF Gain = Max Gain = 0 dB, Amax=1.0dB, Amin=35dB, ω_{p_1}=1000 rad/sec, ω_{p_2}=8000 rad/sec, ω_{s_1}=2000 rad/sec. and ω_{s_2}=4000 rad/sec.

1.5 Suppose you have a low pass filter that has a high frequency magnitude response characteristic with a slope of -18dB/octave. Express the slope in units of dB/decade.