F. Frequently Asked Questions

Contents

F1 Continuous-Time and Discrete-Time Signals

Q1.1 What is signum function? [Answer]

Q1.2 What is Dirac delta function? [Answer]

Q1.3 If a continuous-time signal is

$$x_a(t) = A\cos(\Omega_0 t + \phi)$$

and the corresponding discrete-time signal obtained by sampling $x_a(t)$ at rate F_T is

$$x[n] = A\cos(\omega_0 n + \phi),$$

what is the relationship between the continuous-time angular frequency Ω_0 and normalized discrete-time angular frequency ω_0 after sampling process? [Answer]

Q1.4 What is the bandwidth of a continuous-time signal? [Answer]

Q1.5 What is the relation between the bandwidth of a continuous-time signal and the bandwidth of the corresponding discrete-time signal after sampling process? [Answer]

Q1.6 For discrete-time sinusoidal signal $x[n] = \cos(\omega_0 n+\phi)$, why are frequencies around $\omega_0 = 2k\pi$ usually called *low* frequencies,and frequencies around $\omega_0 = (2k+1)\pi$ called *high* frequencies? [Answer]

Q1.7 What is zero-padding? [Answer]

Q1.8 What is the difference between total energy and average power of a continuous-time signal? [Answer]

Q1.9 What is the difference between total energy and average power of a discrete-time signal? [Answer]

Q1.10 What is the folding frequency? [Answer]

Q1.11 What are Nyquist frequency, Nyquist rate and Nyquist band? [Answer]

Q1.12 If we sample the DTFT (discrete-time Fourier transform) of a sequence $\{x[n]\}$ at N equally spaced points in the ω-axis in the range $0 \leq \omega \leq 2\pi$ starting at $\omega = 0$ and denote the N-point inverse DFT (discrete Fourier transform) of these N frequency samples as $\{y[n]\}$, what is the relation between $\{x[n]\}$ and $\{y[n]\}$?
[Answer]

Q1.13 By computing the DTFT of a discrete-time signal $g[n]$, we can analyze the frequency spectrum:

$$G(e^{j\omega}) = \sum_{n=-\infty}^{\infty} g[n]e^{-j\omega n}.$$

If $g[n]$ has infinite-length, we usually truncate $g[n]$ by multiplying a length-N window $w[n]$ to make it into a finite-length sequence $\gamma[n] = g[n] \cdot w[n]$. Is the frequency spectrum of $\gamma[n]$ different from the frequency spectrum of $g[n]$? [Answer]

Q1.14 Does there exist any real function having the same waveform (regardless of vertical and horizontal scaling) as its continuous-time Fourier-transform spectrum? [Answer]

Q1.15 Is convolution operation always associative for continuous-time signal? [Answer]

Q1.16 Is convolution operation always associative for discrete-time signal? [Answer]

Q1.1 What is signum function?

Answer: The signum function $f(x)$ outputs the sign of x:

$$\text{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}.$$

[Back to FAQs list.]

Q1.2 What is Dirac delta function?

Answer: *Dirac delta* function $\delta(\tau)$ is a function of τ with infinite height, i.e., as $\triangle$ goes to 0, i.e.,

$$\int_{-\infty}^{\infty} \delta(\tau)\, d\tau = 1.$$

One useful property of the Dirac delta function is the sifting property:

$$\int_{-\infty}^{\infty} x(\tau)\delta(t-\tau)\, d\tau = x(\tau)|_{\tau=t} = x(t).$$

Please see Useful Functions of the review material in the CD.

[Back to FAQs list.]

Q1.3 If a continuous-time signal is

$$x_a(t) = A\cos(\Omega_0 t + \phi)$$

and the corresponding discrete-time signal obtained by sampling $x_a(t)$ at rate F_T is

$$x[n] = A\cos(\omega_0 n + \phi),$$

what is the relationship between the continuous-time angular frequency Ω_0 and normalized discrete-time angular frequency ω_0 after sampling process?

Answer: The $\Omega_T = 2\pi F_T$ denotes the sampling angular frequency, then

$$\omega_0 = \frac{2\pi\Omega_0}{\Omega_T}.$$

is the normalized digital angular frequency of $x[n]$. For example, if the continuous-time signal $x_a(t) = \cos(26\pi t)$ is sampled at a sampling rate of 10Hz, then the normalized discrete-time angular frequency is

$$\omega_0 = \frac{2\pi \cdot 26\pi}{2\pi \cdot 10} = 2.6\pi.$$

Please see Section 2.3 in the text.

[Back to FAQs list.]

Q1.4 What is the bandwidth of a continuous-time signal?

Answer: Assume the continuous-time signal $x_a(t)$ is real and band-limited to Ω_m, i.e., the Continuous-Time Fourier Transform $X_a(j\Omega) = 0$, $|\Omega| > \Omega_m$. The bandwidth of $x_a(t)$ is $\Omega_m - (-\Omega_m) = 2\Omega_m$.

[Back to FAQs list.]

Q1.5 What is the relation between the bandwidth of a continuous-time signal and the bandwidth of the corresponding discrete-time signal after sampling process?

Answer: Assume the continuous-time signal $x_a(t)$ is real and band-limited to Ω_m, i.e., the Continuous-Time Fourier Transform $X(j\Omega) = 0$, $|\Omega| > \Omega_m$. Also assume the sampling period is T and no aliasing occurs, then the corresponding discrete-time signal $x[n]$ is band-limited to $\omega_m = \Omega_m T$, and the bandwidth of $x[n]$ is $\omega_m - (-\omega_m) = 2\omega_m = 2\Omega_m T$.

[Back to FAQs list.]

Q1.6 For discrete-time sinusoidal signal $x[n] = \cos(\omega_0 n + \phi)$, why are frequencies around $\omega_0 = 2k\pi$ usually called *low* frequencies,and frequencies around $\omega_0 = (2k+1)\pi$ called *high* frequencies?

Answer: Observe Figure 2.16 in the book, we see that the frequency of oscillation of the discrete-time sinusoidal sequence $x[n] = \cos(\omega_0 n)$ increases as ω_0 increases from 0 to π, and then the frequency of oscillation decreases as ω_0 increases from π to 2π. Because a frequency $\omega_0 + 2k\pi$ is indistinguishable from a frequency ω_0, and a frequency $\omega_0 + (2k+1)\pi$ is indistinguishable from a frequency $\omega_0 + \pi$. Therefore, frequencies in neighborhood of $\omega_0 = 2k\pi$ are usually called low frequencies, and frequencies around $\omega_0 = (2k+1)\pi$ are called high frequencies.

[Back to FAQs list.]

Q1.7 What is zero-padding?

Answer: Zero-padding is to append zeros to the end of the discrete-time signal $x[n]$. For example, if three zeros are padded to $x[n]$, then the new signal is $\{x[n], 0, 0, 0\}$.

[Back to FAQs list.]

Q1.8 What is the difference between total energy and average power of a continuous-time signal?

Answer: The total energy of a continuous-time signal $x_a(t)$ is the square integral over infinite time:

$$\lim_{T\to\infty} \int_{-T}^{T} |x(t)|^2 \, dt.$$

The average power of a continuous-time signal $x_a(t)$ is the total energy divided by time:

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 \, dt.$$

The definition of total energy can explained as the area under the squared signal $|x(t)|^2$, and it is a measurement of the strength of the signal $x(t)$ over infinite time. However, there are signals with infinite energy so we need to evaluate the average power of the signal $x(t)$ as a measurement of the strength over one unit time.

[Back to FAQs list.]

Q1.9 What is the difference between total energy and average power of a discrete-time signal?

Answer: The total energy of a discrete-time signal $x[n]$ is the square summation over infinite time:

$$\sum_{n=-\infty}^{\infty} |x[n]|^2.$$

The average power of a discrete-time signal $x[n]$ is the total energy divided by time:

$$\lim_{K\to\infty} \frac{1}{2K+1} \sum_{n=-K}^{K} |x[n]|^2.$$

The definition of total energy can explained as the summation of the squared signal $|x[n]|^2$, and it is a measurement of the strength of the signal $x[n]$ over infinite time. However, there are signals with infinite energy so we need to evaluate the average power of the signal $x[n]$ as a measurement of the strength over one unit time.

[Back to FAQs list.]

Q1.10 What is the folding frequency?

Answer: The definition of folding frequency is the same as the Nyquist frequency, which is the maximum frequency Ω_m of the band-limited signal $x_a(t)$ with continuous-time Fourier transform $X_a(j\Omega) = 0$, $|\Omega| > \Omega_m$.

[Back to FAQs list.]

Q1.11 What are Nyquist frequency, Nyquist rate and Nyquist band?

Answer: If the continuous-time signal $x_a(t)$ is band-limited in the frequency range $-\Omega_m \leq \Omega \leq \Omega_m$, then the Nyquist rate is the minimum sampling frequency $\Omega_T = 2\Omega_m$ such that no aliasing occurs during the sampling process. The Nyquist frequency, $\Omega_T/2 = \Omega_m$, equals one half of the Nyquist rate. The Nyquist band is the frequency range of $-\Omega_T/2 \leq \Omega \leq \Omega_T/2$.

[Back to FAQs list.]

Q1.12 If we sample the DTFT (discrete-time Fourier transform) of a sequence $\{x[n]\}$ at N equally spaced points in the ω-axis in the range $0 \leq \omega \leq 2\pi$ starting at $\omega = 0$ and denote the N-point inverse DFT (discrete Fourier transform) of these N frequency samples as $\{y[n]\}$, what is the relation between $\{x[n]\}$ and $\{y[n]\}$?

Answer: Denote the discrete-time Fourier transform of a sequence $\{x[n]\}$ as $X(e^{j\omega})$. If we sample $X(e^{j\omega})$ at N equally spaced points $\omega_k = 2\pi k/N$, $0 \leq k \leq N-1$, these N frequency samples can be regarded as an N-point DFT $Y[k]$ with the N-point inverse DFT a length-N sequence $\{y[n]\}$, $0 \leq n \leq N-1$.

It can be shown that the relation between $x[n]$ and $y[n]$ is

$$y[n] = \sum_{m=-\infty}^{\infty} x[n+mN], \quad 0 \leq n \leq N-1.$$

The above equation states that $\{y[n]\}$ is the summation of infinite shifted replicas of $x[n]$. Therefore, if $\{x[n]\}$, $0 \leq n \leq M-1$ is defined to be a finite-length sequence of length M and $M \leq N$, $\{x[n]\}$ can be fully recovered from $\{y[n]\}$. If $M > N$, the time-domain aliasing arises in generating $\{y[n]\}$ and we cannot recover $\{x[n]\}$ from $\{y[n]\}$. Please see Section 5.3 of the book.

[Back to FAQs list.]

Q1.13 By computing the DTFT of a discrete-time signal $g[n]$, we can analyze the frequency spectrum:

$$G(e^{j\omega}) = \sum_{n=-\infty}^{\infty} g[n]e^{-j\omega n}.$$

If $g[n]$ has infinite-length, we usually truncate $g[n]$ by multiplying a length-N window $w[n]$ to make it into a finite-length sequence $\gamma[n] = g[n] \cdot w[n]$. Is the frequency spectrum of $\gamma[n]$ different from the frequency spectrum of $g[n]$?

Answer: Let $\Psi_R(e^{j\omega})$ denote the DTFT of $w[n]$. The DTFT $\Gamma(e^{j\omega})$ of $\gamma[n]$ is given by the frequency-domain convolution:

$$\Gamma(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} G(e^{j\varphi})\Psi_R(e^{j(\omega-\varphi)})\, d\varphi.$$

Because the frequency spectrum $\Psi_R(e^{j\omega})$ consists of the main lobe and many sidelobes, the DTFT $\Gamma(e^{j\omega})$ of $\gamma[n]$ is different from the DTFT $G(e^{j\omega})$ of $g[n]$ in two aspects: (1) the main lobe width of the window frequency spectrum $\Psi_R(e^{j\omega})$ determines the frequency resolution of $\Gamma(e^{j\omega})$, (2) the relative sidelobe level of $\Psi_R(e^{j\omega})$ controls the *leakage* phenomenon, which is the spread of energy in frequency-domain from one single frequency to its neighborhood. The frequency resolution can be improved by increasing the window length, and the leakage phenomenon can be reduced by a proper choice of window (e.g., Rectangular, Hann, Hamming, Blackman). Please see Section 10.2 and Section 15.2 of the book.

[Back to FAQs list.]

Q1.14 Does there exist any real function having the same waveform (regardless of vertical and horizontal scaling) as its continuous-time Fourier-transform spectrum?

Answer: Gaussian function

$$h(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(t-\mu)^2/2\sigma^2}.$$

Note that convolving the square pulse function

$$x(t) = \begin{cases} +1 & , |t| \leq 1/2 \\ -1 & , 1/2 < |t| \end{cases}$$

with itself for infinite times will converge to the Gaussian function.

[Back to FAQs list.]

Q1.15 Is convolution operation always associative for continuous-time signal?

Answer: No. It is associative only for stable and single-sided continuous-time signals. For example, if

$$x_1(t) = A,\ x_2(t) = u(t),\ x_3(t) = \delta(t) - \delta(t-1),$$

then $x_1(t) \circledast (x_3(t) \circledast x_2(t)) \neq (x_1(t) \circledast x_3(t)) \circledast x_2(t)$.

[Back to FAQs list.]

Q1.16 Is convolution operation always associative for discrete-time signal?

Answer: No. It is associative only for stable and single-sided discrete-time signals. For example, if

$$x_1[n] = A,\ x_2[n] = u[n],\ x_3[n] = \delta[n] - \delta[n-1],$$

then $x_1[n] \circledast (x_3[n] \circledast x_2[n]) \neq (x_1[n] \circledast x_3[n]) \circledast x_2[n]$.

[Back to FAQs list.]

F2 Discrete-Time Systems

Q2.1 What is a linear time-invariant (LTI) system? How can the LTI property help us to analyze the system? [Answer]

Q2.2 What is the property of a zero-phase transfer function? Can a causal transfer function have zero-phase? [Answer]

Q2.3 How can we convert a zero-phase FIR transfer function to a causal FIR transfer function with identical magnitude responses? [Answer]

Q2.4 What are the properties of a minimum-phase transfer function? [Answer]

Q2.5 What is the difference between the phase delay and the group delay of an LTI discrete-time system? [Answer]

Q2.6 How does the group-delay of a discrete-time system affect a discrete-time signal composed of multiple sinusoids of different frequencies? [Answer]

Q2.7 What's the sufficient condition of the system

$$\sum_{k=0}^{N} d_k y[n-k] = \sum_{k=0}^{M} p_k x[n-k]$$

to be *linear* ? [Answer]

Q2.8 What is the impulse invariance method of IIR digital filter design? [Answer]

Q2.9 What is the frequency sampling method of FIR digital filter design? [Answer]

Q2.10 What is the least-mean-square error method of FIR digital filter design? [Answer]

Q2.11 What is the constrained least-square method of FIR digital filter design? [Answer]

Q2.12 What is the generalized multilevel FIR digital filter design? [Answer]

Q2.1 What is a linear time-invariant (LTI) system? How can the LTI property help us to analyze the system?

Answer: The LTI system is a system satisfying both the linear and time-invariant(or shift-invariant in the discrete case) properties. This enables us to analyze the system output by the convolution operation:

$$y[n] = x[n] \circledast h[n],$$

where $h[n]$ is the impulse response of the system and $x[n]$ is the input. Please see Section 2.5 in the book.

[Back to FAQs list.]

Q2.2 What is the property of a zero-phase transfer function? Can a causal transfer function have zero-phase?

Answer: A zero-phase transfer function $F(e^{j\omega})$ has a real and nonnegative frequency response. For a length-N causal digital filter with impulse response

$$f[n], \quad 0 \le n \le N-1,$$

it is impossible for the DTFT

$$F(e^{j\omega}) = \sum_{n=0}^{N-1} f[n] e^{-j\omega n}$$

to be real and nonnegative unless $f[n]$ is a delta function with nonnegative magnitude. Please see Section 7.2 in the book.

[Back to FAQs list.]

Q2.3 How can we convert a zero-phase FIR transfer function to a causal FIR transfer function with identical magnitude responses?

Answer: A zero-phase FIR transfer function is of the form

$$F(z) = a_0 + \frac{a_1}{2}(z + z^{-1}) + \frac{a_2}{2}(z^2 + z^{-2}) + \cdots + \frac{a_M}{2}(z^M + z^{-M}),$$

with a_0, a_1,..., a_M real numbers. Its frequency response is of the form

$$\begin{aligned} F(e^{j\omega}) &= a_0 + \frac{a_1}{2}(e^{j\omega} + e^{-j\omega}) + \frac{a_2}{2}(e^{j2\omega} + e^{-j2\omega}) + \cdots + \frac{a_M}{2}(e^{jM\omega} + e^{-jM\omega}) \\ &= a_0 + a_1 \cos(\omega) + a_2 \cos(2\omega) + \cdots + a_M \cos(M\omega). \end{aligned}$$

A causal transfer function $G(z)$ with the same magnitude response as that of $H(z)$ is given by

$$\begin{aligned} G(z) =& z^{-M} F(z) = \frac{a_M}{2} + \frac{a_{M-1}}{2} z^{-1} + \cdots + \frac{a_1}{2} z^{-(M-1)} + a_0 z^{-M} \\ &+ \frac{a_1}{2} z^{-(M+1)} + \cdots + \frac{a_{M-1}}{2} z^{-(2M-1)} + \frac{a_M}{2} z^{-2M}. \end{aligned}$$

[Back to FAQs list.]

Q2.4 What are the properties of a minimum-phase transfer function?

Answer: A causal stable transfer function with all zeros inside the unit circle is called a minimum-phase transfer function. The properties of a minimum-phase transfer function are (1) minimum group delay, and (2) maximum energy compactness. Please see Section 7.2 in the book.

[Back to FAQs list.]

Q2.5 What is the difference between the phase delay and the group delay of an LTI discrete-time system?

Answer: If the input to an LTI system $H(e^{j\omega})$ is a sinusoidal signal with frequency ω_0

$$x[n] = A\cos(\omega_0 n + \phi),$$

then the output is the input multiplied by $|H(e^{j\omega})|$ and delayed by $\theta(\omega_0)$ in phase

$$\begin{aligned} y[n] &= A|H(e^{j\omega})|\cos\Big(\omega_0\Big(n + \frac{\theta(\omega_0)}{\omega_0}\Big) + \phi\Big) \\ &= A|H(e^{j\omega})|\cos(\omega_0(n - \tau_p(\omega_0)) + \phi), \end{aligned}$$

where $\tau_p(\omega_0) = -\frac{\theta(\omega_0)}{\omega_0}$ is called the phase delay. However, if the input signal has many sinusoidal components with different frequencies and all of them are not harmonically related, then each sinusoidal frequency component will have different phase delay. Therefore, the group delay given by

$$\tau_g(\omega) = -\frac{d\theta(\omega)}{d\omega}$$

is used to indicate the time delay between the continuous-time underlying waveforms of input signal and output signal. Please see Section 3.9 of the book.

[Back to FAQs list.]

Q2.6 What is the effect of non-constant group delay?

Answer: The group delay is a measure of the linearity of the phase function as a function of the frequency and is the time delay between waveforms of underlying continuous-time signals whose sampled versions, sampled at $t = nT$, are precisely the input and the output discrete-time signals. The waveform of the underlying continuous-time output signal shows distortion when the group delay of the LTI system is not constant over the bandwidth of the input signal. We show two examples in the following.

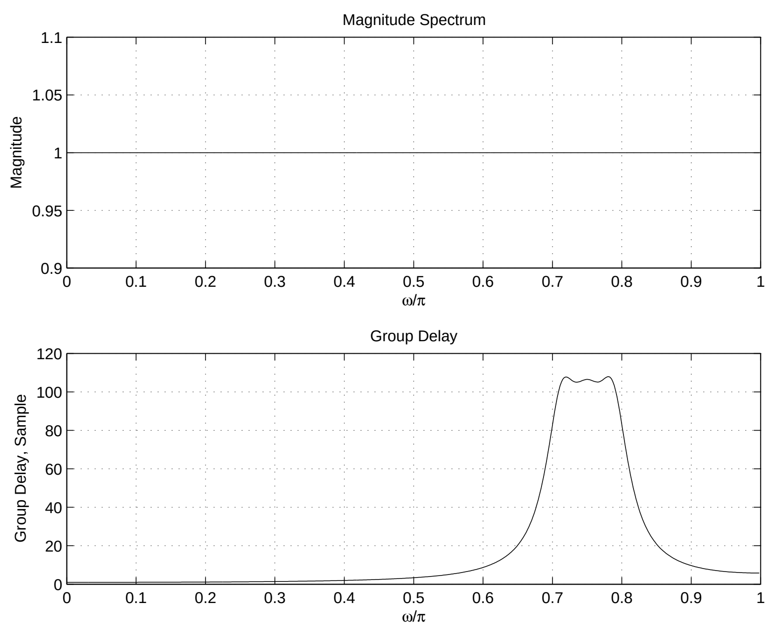

Figure F2.1: Allpass transfer function.

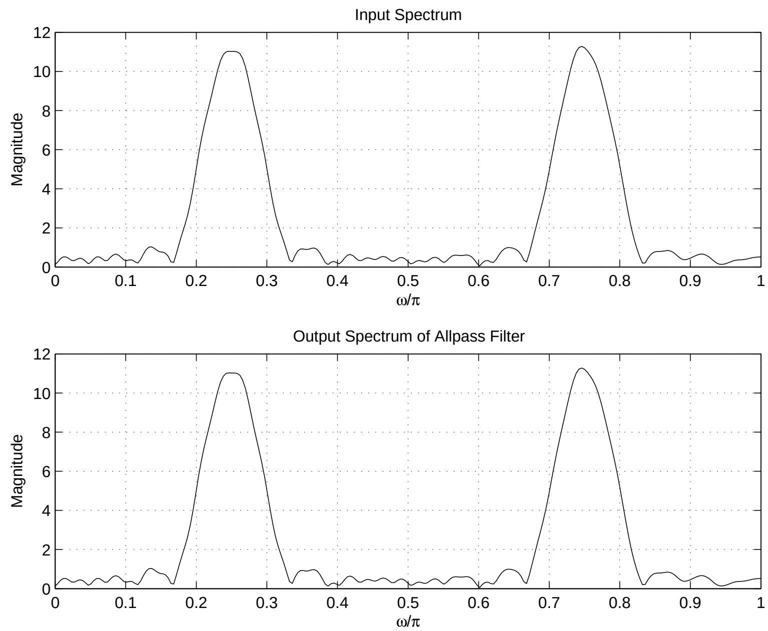

Figure F2.2: Input/output spectrum.

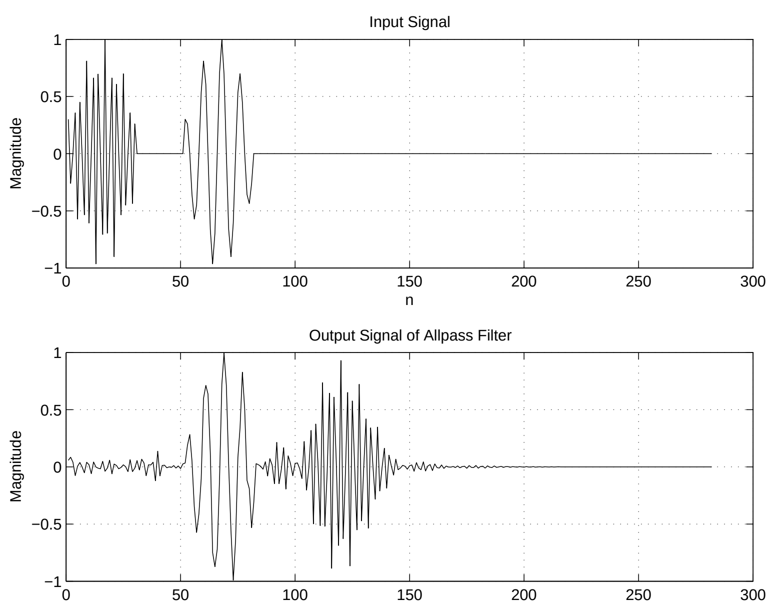

Figure F2.3: Input/output signal.

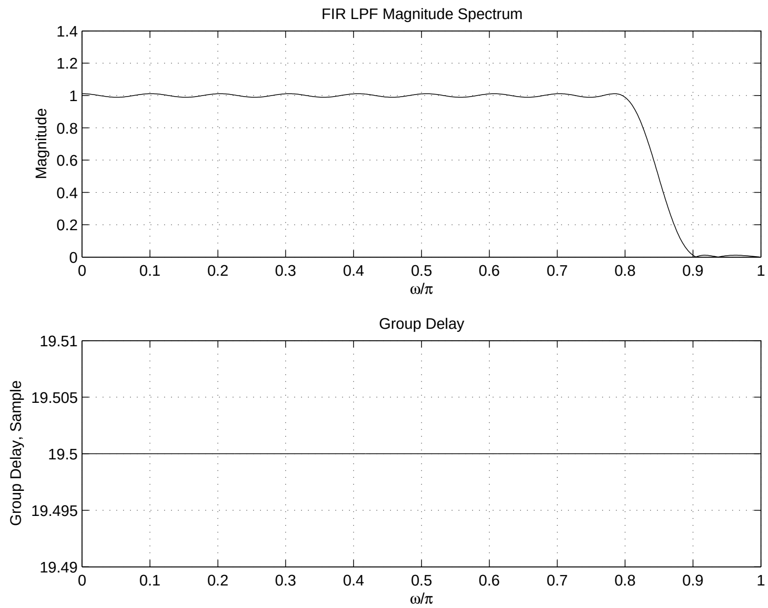

Figure F2.4: FIR LPF transfer function.

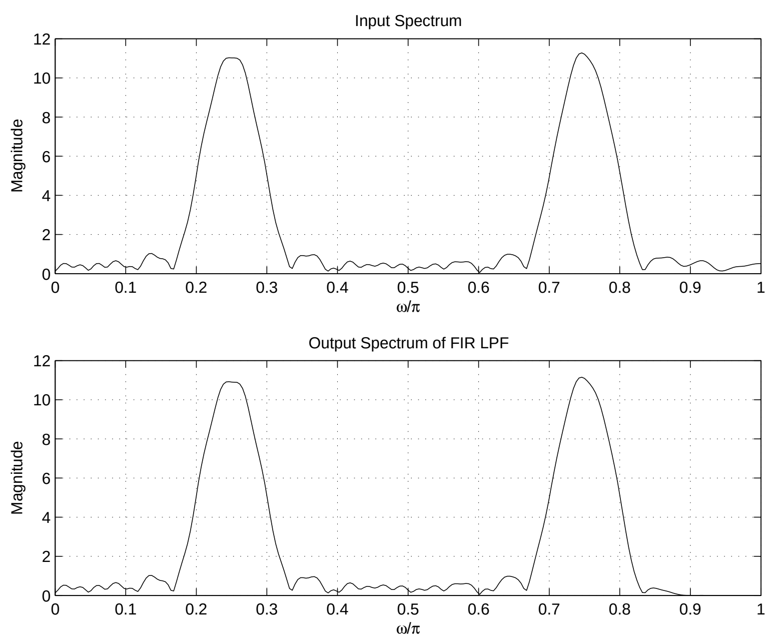

Figure F2.5: Input/output spectrum.

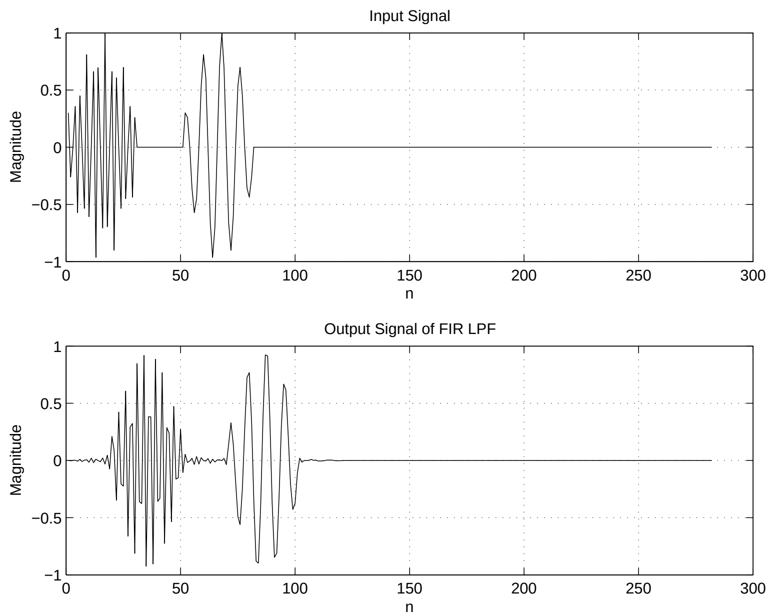

Figure F2.6: Input/output signal.

In Figure F2.1-F2.3, we show an example of non-constant group delay. Figure F2.1 shows the spectrum of an allpass filter with group delay approximately equal 100 samples at the frequency $\omega = 0.75\pi$. Figure F2.2 shows an input signal consisting of two sinusoidal pulses of frequencies $\omega = 0.25\pi$ and $\omega = 0.75\pi$. Figure F2.3 shows that the high-frequency pulse starts at sample $= 0$ and the low-frequency pulse starts at sample $= 50$. After the allpass filter the $\omega = 0.25\pi$ pulse remains at the same sample location as input, while the $\omega = 0.75\pi$ pulse is delayed by 100 samples at output.

Figure F2.4-F2.6 show an example of constant group delay. Figure F2.4 is the designed linear-phase FIR lowpass filter with group delay of 19.5 samples. Figure F2.5 gives input and out spectrums. Figure F2.6 shows that both $\omega = 0.25\pi$ pulse and $\omega = 0.75\pi$ pulse are delayed by the same amount of 19.5 samples.

[Back to FAQs list.]

Q2.7 What's the sufficient condition of the system

$$\sum_{k=0}^{N} d_k y[n-k] = \sum_{k=0}^{M} p_k x[n-k]$$

to be *linear* ?

Answer: Zero initial conditions, i.e., $y[-1] = y[-2] = \cdots = y[-N] = 0$. Please refer to Problem 2.67 in the book.

[Back to FAQs list.]

Q2.8 What is the impulse invariance method of IIR digital filter design?

Answer: The basic idea behind the impulse invariance method is to develop a causal stable IIR transfer function $G(z)$ whose impulse response $g[n]$ is exactly identical to the uniformly sampled version of the impulse response $h_a(t)$ of a prototype causal stable analog transfer function $H_a(s)$; that is,

$$g[n] = h_a(nT), \; n = 0, \; 1, \; 2, ..., \; \infty \tag{F2.1}$$

where T is the sampling period. It can be shown that the relation between the transfer function $G(z)$ and $H_a(s)$ is given by

$$\begin{aligned} G(z) &= \mathcal{Z}\{g[n]\} = \mathcal{Z}\{h_a(nT)\} \\ &= \frac{1}{T}\sum_{k=-\infty}^{\infty} H_a(s + j\frac{2\pi k}{T})\big|_{s=\frac{1}{T}\ln z}. \end{aligned} \tag{F2.2}$$

The corresponding relation between the frequency response $G(e^{j\omega})$ of the digital transfer function and the frequency response $H_a(j\Omega)$ is given by

$$G(e^{j\omega}) = \frac{1}{T}\sum_{k=-\infty}^{\infty} H_a(j\frac{\omega}{T} + j\frac{2\pi k}{T}). \tag{F2.3}$$

It follows from the above expression that the frequency response of the desired digital transfer function is given by the sum of the frequency responses of the original analog transfer function and its frequency shifted versions, shifted by $\pm 2\pi k/T$, with the overall sum scaled by the factor $1/T$. As a result, if the analog frequency response $H_a(j\Omega)$ is bandlimited with

$$H_a(j\Omega) \text{ for } |\Omega| \geq \frac{\pi}{T}, \tag{F2.4}$$

then

$$G(e^{j\omega}) = \frac{1}{T}H_a(j\frac{\omega}{T}) \text{ for } 0 \leq |\omega| < \pi, \tag{F2.5}$$

and there is no aliasing. If the above condition (F2.5) does not hold, then there will be aliasing.

[Back to FAQs list.]

Q2.9 What is the frequency sampling method of FIR digital filter design?

Answer: In this approach, the specified frequency response $H_d(e^{j\omega})$ is first uniformly sampled at N equally spaced points $\omega_k = 2\pi k/N$, $k = 0,\ 1, ...,\ N-1$, providing N *frequency samples*. These N frequency samples constitute an N-point DFT $H[k]$ whose N-point inverse-DFT thus yields the impulse response coefficients $h[n]$ of the FIR filter of length N. The basic assumption here is that the specified frequency response is uniquely characterized by the N frequency samples and, hence, can be fully recovered from these samples.

Now, $H_d(e^{j\omega})$ is a periodic function of ω with a Fourier series representation given by

$$H_d(e^{j\omega}) = \sum_{k=-\infty}^{\infty} h_d[n]e^{-j\omega n}. \tag{F2.6}$$

Its Fourier coefficients $h_d[n]$ are thus given by

$$h_d[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} H_d(e^{j\omega})e^{j\omega n}\,d\omega. \tag{F2.7}$$

It is instructive to develop the relation between $h_d[n]$ and $h[n]$.

From Eq.(F2.6),

$$H[k] = H_d(e^{j\omega_k}) = H_d(e^{j(2\pi k/N)}) = \sum_{l=-\infty}^{\infty} h_d[\ell]W_N^{k\ell}, \tag{F2.8}$$

where $W_N = e^{-j(2\pi/N)}$. An inverse-DFT of $H[k]$ yields

$$h[n] = \frac{1}{N}\sum_{k=0}^{N-1} H[k]W_N^{-kn}. \tag{F2.9}$$

Substituting Eq.(F2.8) in Eq.(F2.9), we get

$$\begin{aligned} h[n] &= \frac{1}{N}\sum_{k=0}^{N-1}\sum_{\ell=-\infty}^{\infty} h_d[\ell]W_N^{k\ell}W_N^{-kn} \\ &= \sum_{\ell=-\infty}^{\infty} h_d[\ell]\big[\frac{1}{N}\sum_{k=0}^{N-1} W_N^{-k(n-\ell)}\big]. \end{aligned} \tag{F2.10}$$

Making use of the identity

$$\frac{1}{N}\sum_{k=0}^{N-1} W_N^{-k(n-r)} = \begin{cases} 1, & \text{for } r = n + mN, \\ 0, & \text{otherwise.} \end{cases} \tag{F2.11}$$

in Eq.(F2.10), we finally arrive at the desired relation

$$h[n] = \sum_{m=-\infty}^{\infty} h_d(n + mN),\ 0 \le n \le N-1. \tag{F2.12}$$

The above relation indicates that $h[n]$ is obtained from $h_d[n]$ by adding an infinite number of shifted replicas of $h_d[n]$ to $h_d[n]$, with each replica shifted by an integer multiple of N sampling instants, and observing the sum only for the interval $0 \le n \le N-1$. Thus, if $h_d[n]$ is a finite-length sequence of length of less than or equal to N, then $h[n] = h_d[n]$ for $0 \le n \le N-1$, otherwise there is a time-domain aliasing of samples with $h[n]$ bearing no resemblance to $h_d[n]$.

It can be shown that the transfer function of the desired FIR filter is given by

$$H(z) = \frac{1 - z^{-N}}{N}\sum_{k=0}^{N-1}\frac{H[k]}{1 - W_N^{-k} z^{-1}}. \tag{F2.13}$$

On the unit circle, Eq.(F2.13) reduces to

$$H(e^{j\omega}) = \frac{e^{-j\omega[(N-1)/2]}}{N}\sum_{k=0}^{N-1} H[k]\frac{\sin(\frac{\omega N - 2\pi k}{2})}{\sin(\frac{\omega N - 2\pi k}{2N})} e^{j\pi k[(N-1)/N]}. \tag{F2.14}$$

It can be shown that

$$H(e^{j\omega})\big|_{\omega=2\pi k/N} = H[k],\ k = 0,\ 1, ...,\ N-1. \tag{F2.15}$$

That is, the FIR filter designed via the frequency sampling approach has exactly the specified frequency samples $H[k] = H_d(e^{e^{j\omega}})$, where $\omega_k = 2\pi k/N$, $k =\ 0,\ 1, ...\ N-1$, whether or not the length of $h_d[n]$ is less than or equal to N.

In general, the mininum stopband attenuation of the desired FIR filter is much smaller than the desired value. The reason behind the overall unsatisfactory magnitude response is that the impulse response $h_d[n]$ corresponding to the ideal filter is of infinite length, and as a result of the relation of Eq.(F2.12), there is a severe time-domain aliasing in determining the impulse response coefficients of the FIR filter $h[n]$.

[Back to FAQs list.]

Q2.10 What is the least-mean-square error method of FIR digital filter design?

Answer: For the design of a linear-phase FIR filter with a minimum mean-square error criterion, the error measure is

$$\varepsilon = \sum_{i=1}^{K} \left\{ W(\omega_i) \left[\breve{H}(\omega_i) - D(\omega_i) \right] \right\}^2, \tag{F2.16}$$

where $\breve{H}(\omega)$ is the amplitude response of the designed filter, $D(\omega)$ is the desired amplitude response, and $W(\omega)$ is the weighting function. Since the amplitude response for all four types of linear-phase FIR filters can be expressed in the form

$$\breve{H}(\omega) = Q(\omega) \sum_{k=0}^{L} \tilde{a}[k] \cos(\omega k), \tag{F2.17}$$

where $Q(\omega)$, $\tilde{a}[k]$, and L are given in Section 10.3 of the book. Hence, the mean-square error of Eq. (F2.16) is a function of the filter parameters $\tilde{a}[k]$. To arrive at the minimum value of ε, we set

$$\frac{\partial \varepsilon}{\partial \tilde{a}[k]} = 0, \qquad 0 \le k \le L,$$

which results in a set of $(L+1)$ linear equations that can be solved for $\tilde{a}[k]$.

Without any loss of generality, we consider here the design of a Type 1 linear-phase FIR filter. In this case, $Q(\omega) = 1$, $\tilde{a}[k] = a[k]$, and $L = M$. The expression for the mean-square error then takes the form

$$\begin{aligned} \varepsilon &= \sum_{i=1}^{K} \left\{ W(\omega_i) \left[\sum_{k=0}^{M} a[k] \cos(\omega_i k) - D(\omega_i) \right] \right\}^2 \\ &= \sum_{i=1}^{K} \left\{ \sum_{k=0}^{M} W(\omega_i) a[k] \cos(\omega_i k) - W(\omega_i) D(\omega_i) \right\}^2 . \end{aligned} \tag{F2.18}$$

Using the notation,

$$\mathbf{H} = \begin{bmatrix} W(\omega_1) & W(\omega_1)\cos(\omega_1) & \cdots & W(\omega_1)\cos(M\omega_1) \\ W(\omega_2) & W(\omega_2)\cos(\omega_2) & \cdots & W(\omega_2)\cos(M\omega_2) \\ \vdots & \vdots & \ddots & \vdots \\ W(\omega_K) & W(\omega_K)\cos(\omega_K) & \cdots & W(\omega_K)\cos(M\omega_K) \end{bmatrix},$$

$$\mathbf{a} = \left[a[0]\ a[1]\ \cdots\ a[M]\right]^T,$$

and

$$\mathbf{d} = \left[W(\omega_1)D(\omega_1)\ W(\omega_2)D(\omega_2)\ \cdots\ W(\omega_K)D(\omega_K)\right]^T,$$

we can express Eq. (F2.18) in the form

$$\varepsilon = \mathbf{e}^T\mathbf{e},$$

where

$$\mathbf{e} = \mathbf{H}\mathbf{a} - \mathbf{d}.$$

The minimum mean-square solution is then obtained by solving the *normal equations* [Par87]:

$$\mathbf{H}^T\mathbf{H}\mathbf{a} = \mathbf{H}^T\mathbf{d}.$$

If $K \geq M$, which is typically the case, the above equation should be solved using an iterative method such as the *Levinson-Durbin algorithm* [Lev47], [Dur59], as the direct solution is often ill-conditioned. A similar formulation can be carried out for the other three types of linear-phase FIR filters. Note that the design approach outlined here can be used to design a linear-phase FIR filter meeting any arbitrarily shaped desired response.

[Back to FAQs list.]

Q2.11 What is the constrained least-square method of FIR digital filter design?

Answer: FIR filters with constraints on their frequency response can be designed using the least-mean-squares approach by incorporating the constraints into the design algorithm. To illustrate this approach, assume, without any loss of generality, that the filter to be designed is a Type 1 linear-phase FIR filter of order $N = 2M$ with an amplitude response given by Eq. (10.62) in the book and constrained to have a null at ω_o. This can be written as a single equality constraint $\mathbf{Ga} = \mathbf{d}$, where

$$\begin{aligned} \mathbf{G} &= [1,\ \cos(\omega_o),\ \cos(2\omega_o), \cdots, \cos(M\omega_o)], \\ \mathbf{a} &= [a[0],\ a[1],\ \cdots,\ a[M]]^T \\ \mathbf{d} &= [0]. \end{aligned} \tag{F2.19}$$

In the general case, if there are r constraints, then $\mathbf{G}$ is an $r \times (M+1)$ matrix and $\mathbf{d}$ is an $r \times 1$ column vector. Such a filter can be designed by using the constrained least-square method. This method minimizes the square error

$$\varepsilon = \left(\frac{1}{\pi} \int_0^{\pi} W(\omega)[\breve{H}(\omega) - D(\omega)]^2 \, d\omega \right)^{1/2}, \tag{F2.20}$$

subject to the side constraints

$$\mathbf{Ga} = \mathbf{d}. \tag{F2.21}$$

As before, $D(\omega)$ is the desired amplitude response and $W(\omega)$ is the weighting function. The side constraints of Eq. (F2.21) need not be linear, but the solution is more easily obtained if they are.

To minimize ε^2 subject to the constraints, we first form the Lagrangian:

$$\Phi = \varepsilon^2 + \boldsymbol{\mu}^T \cdot [\mathbf{Ga} - \mathbf{d}], \tag{F2.22}$$

where

$$\boldsymbol{\mu} = [\mu_1,\ \mu_2,\ \ldots, \mu_r]^T$$

is the vector of the so-called Lagrange multipliers. We can derive the necessary conditions for the minimization of ε^2 by setting the derivatives of Φ with respect to the filter parameters $a[k]$ and the Lagrange multipliers μ_i to zero resulting in the following equations:

$$\begin{aligned} \mathbf{R}\mathbf{a} + \mathbf{G}^T \boldsymbol{\mu} &= \mathbf{c}, \\ \mathbf{G}\mathbf{a} &= \mathbf{d}, \end{aligned} \tag{F2.23}$$

where the coefficients of the vector $\mathbf{c} = [c[0],\ c[1],\ \ldots,\ c[M]]$ are given by

$$\begin{aligned} c[0] &= \frac{1}{\pi}\int_0^\pi W(\omega)D(\omega)\,d\omega, \\ c[k] &= \frac{1}{\pi}\int_0^\pi W(\omega)D(\omega)\cos(k\omega)\,d\omega, \qquad 1 \le k \le M, \end{aligned}$$

and the (i,k)th element $R_{i,k}$ of the matrix $\mathbf{R}$ is given by

$$R_{i,k} = \int_0^\pi W(\omega)\cos(i\omega)\cos(k\omega)\,d\omega.$$

$$\begin{aligned} \mathbf{R}\mathbf{a} + \mathbf{G}^T \boldsymbol{\mu} &= \mathbf{c}, \\ \mathbf{G}\mathbf{a} &= \mathbf{d}, \end{aligned} \tag{F2.24}$$

where the coefficients of the vector $\mathbf{c} = [c[0],\ c[1],\ \ldots,\ c[M]]$ are given by

$$\begin{aligned} c[0] &= \frac{1}{\pi}\int_0^\pi W(\omega)D(\omega)\,d\omega, \\ c[k] &= \frac{1}{\pi}\int_0^\pi W(\omega)D(\omega)\cos(k\omega)\,d\omega, \qquad 1 \le k \le M, \end{aligned}$$

and the (i,k)th element $R_{i,k}$ of the matrix $\mathbf{R}$ is given by

$$R_{i,k} = \int_0^\pi W(\omega)\cos(i\omega)\cos(k\omega)\,d\omega.$$

The two matrix equations of Eq. (F2.24) can be written as

$$\begin{bmatrix} \mathbf{R} & \mathbf{G}^T \\ \mathbf{G} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \boldsymbol{\mu} \end{bmatrix} = \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix}. \tag{F2.25}$$

Solving the above equation we get

$$
\begin{aligned}
\boldsymbol{\mu} &= (\mathbf{G}\,\mathbf{R}^{-1}\mathbf{G}^T)^{-1}(\mathbf{G}\,\mathbf{R}^{-1}\mathbf{c}-\mathbf{d}),\\
\mathbf{a} &= \mathbf{R}^{-1}(\mathbf{c}-\mathbf{G}^T\boldsymbol{\mu}).
\end{aligned} \tag{F2.26}
$$

When the integrals needed to form $\mathbf{R}$ and $\mathbf{c}$ cannot be calculated simply, then $\mathbf{R}$ and $\mathbf{c}$ can be approximated using the discrete forms

$$\mathbf{R} \cong \mathbf{H}^T\mathbf{H}, \qquad \mathbf{c} \cong \mathbf{H}^T\mathbf{d}.$$

In the special case when the error function is not weighted, i.e., $W(\omega) = 1$, $\mathbf{R}$ becomes an identity matrix, and the c_i are simply the coefficients of the Fourier series expansion of $D(\omega)$. As a result, Eq. (F2.26) reduces to

$$
\begin{aligned}
\boldsymbol{\mu} &= (\mathbf{G}\,\mathbf{G}^T)^{-1}(\mathbf{G}\,\mathbf{c}-\mathbf{d}),\\
\mathbf{a} &= \mathbf{c}-\mathbf{G}^T\boldsymbol{\mu}.
\end{aligned}
$$

One useful application of the constrained least-square approach is the design of filters by a criterion that takes into account both the square error and the peak-ripple error (or Chebyshev error). The constrained least-square approach to filter design allows a compromise between the square error and the Chebyshev criteria, and produces the filter with least-square error and the best Chebyshev error filter as special cases.

The constrained least-square filter design method can be used to design both linear-phase and minimum-phase FIR filters without specifying explicitly the transition bands [Sel96]. It minimizes the weighted integral-square error of Eq. (F2.20) over the whole frequency range such that the local minima and maxima of $\breve{H}(\omega)$ remain within the specified lower and upper bound functions $L(\omega)$ and $U(\omega)$. As ε defined above is simply the $\mathcal{L}_2$ norm of the error function $[\breve{H}(\omega) - D(\omega)]$, it has also been referred to as the $\mathcal{L}_2$ error. For lowpass filter design with a cutoff frequency ω_o, the functions $L(\omega)$ and $U(\omega)$ are defined by

$$
\begin{aligned}
L(\omega) &= 1-\delta_p,\ U(\omega) = 1+\delta_p, \quad \text{for } 0 \le \omega \le \omega_o,\\
L(\omega) &= -\delta_s,\ U(\omega) = \delta_s, \quad \text{for } \omega_o \le \omega \le \pi.
\end{aligned} \tag{F2.27}
$$

Because this design problem has inequality constraints, an iterative algorithm is employed to minimize the error ε of Eq. (F2.20) subject to the constraints on the values of $\breve{H}(\omega_i)$, where the frequency points ω_i are contained in a constraint set $S = \{\omega_1,\ \omega_2, \ldots, \omega_m\}$ with $\omega_i \in [0, \pi]$. Let the set S be partitioned into two sets, the set S_ℓ containing the frequency points $\omega_i, 1 \le i \le q$, where the equality constraint

$$\breve{H}(\omega) = L(\omega)$$

is imposed, and the set S_u containing the frequency points $\omega_i, q+1 \le i \le m$, where the equality constraint

$$\breve{H}(\omega) = U(\omega)$$

is imposed. Then the equality constrained problem is solved on each iteration.

When the Lagrange multipliers are all nonnegative, Kuhn-Tucker conditions [Fle87] state that the solution of the equality constrained problem minimizes ε of Eq. (F2.20) while satisfying the inequality constraints:

$$\begin{array}{rcll} \breve{H}(\omega_i) & \ge & L(\omega_i), & 1 \le i \le q, \\ \breve{H}(\omega_i) & \le & U(\omega_i), & q+1 \le i \le m. \end{array}$$

The constrained least-square design algorithm therefore consists of the following steps:

Step 1: Initialization. Choose the constraint set to be an empty set, i.e., $S = \emptyset$.

Step 2: Minimization with Equality Constraints. Solve Eq. (F2.26) for the Lagrange multipliers by minimizing the mean-square error ε of Eq. (F2.20) satisfying the equality constraints $\breve{H}(\omega_i) = L(\omega_i)$ for $\omega_i \in S_\ell$ and $\breve{H}(\omega_i) = U(\omega_i)$ for $\omega_i \in S_u$.

Step 3: Kuhn-Tucker Conditions. If a Lagrange multiplier μ_j is negative, then remove the corresponding frequency ω_j from the constrained set S, and return to Step 2. Otherwise, calculate the coefficients $a[k]$ using Eq. (F2.26) and proceed to Step 4.

Step 4: Multiple Exchange of Constraint Set. Set the constraint set S equal to $S_\ell \cup S_u$. Note that at the frequency points ω_i in S_ℓ, $\partial\breve{H}(\omega)/\partial\omega|_{\omega=\omega_i} = 0$ and $\breve{H}(\omega_i) \le L(\omega_i)$. Likewise, at the frequency points ω_i in S_u, $\partial\breve{H}(\omega)/\partial\omega|_{\omega=\omega_i} = 0$ and $\breve{H}(\omega_i) \ge U(\omega_i)$.

Step 5: Convergence Check. The algorithm converges if $\breve{H}(\omega) \geq L(\omega) - \Delta$ for all ω_i in S_ℓ, and if $\breve{H}(\omega) \leq U(\omega) + \Delta$ for all ω_i in S_u. Otherwise, go back to Step 2.

In Step 5, Δ is a very small number, typically 10^{-6}, chosen a priori based on the desired numerical accuracy. For an additional discussion on the algorithm and its properties, see [Sel96], [Sel98].

[Back to FAQs list.]

Q2.12 What is the generalized multilevel FIR digital filter design?

Answer: A generalization of the multilevel L-band FIR digital filter of Figure 10.1 in the book is obtained by making the constant magnitude levels in each a linear function of ω resulting a frequency response given by [Cha95]

$$H_{ML}(e^{j\omega}) = m_\ell \omega + d_\ell, \quad \omega_{\ell-1} < \omega < \omega_\ell, \ \ 1 \leq \ell \leq L,$$

where

$$m_\ell = \frac{A_\ell^- - A_{\ell-1}^+}{\omega_\ell - \omega_{\ell-1}}$$

is the slope of the ℓth segment, and

$$d_\ell = \frac{\omega_\ell A_{\ell-1}^+ - \omega_{\ell-1}^-}{\omega_\ell - \omega_{\ell-1}}$$

is the intercept of the ℓth segment. It can be shown that the impulse response of the generalized multilevel FIR digital filter is given by

$$\begin{aligned} h_{ML}[n] = \frac{2}{n^2\pi} \sum_{\ell=1}^{L} m_\ell \sin \frac{\omega_{\ell-1} + \omega_\ell}{2} \cdot \sin \frac{(\omega_{\ell-1} - \omega_\ell)n}{2} \\ \frac{1}{n\pi} \sum_{\ell=1}^{L-1} (A_\ell^- - A_{\ell-1}^+) \sin \omega_\ell n, \quad n \leq 0, \qquad \text{(F2.28)} \end{aligned}$$

$$h_{ML}[n] = \frac{1}{2\pi} \sum_{\ell=1}^{L} (A_\ell^- - A_{ell+1}^+)(\omega_\ell - \omega_{\ell-1}), \qquad \text{(F2.29)}$$

assuming $A_\ell^- = 0$.

[Back to FAQs list.]

F3 Numbers and Polynomials

Q3.1 What is a prime number? [Answer]

Q3.2 When are two integers relatively prime? [Answer]

Q3.3 What is the greatest common divisor (gcd) of two positive integers? [Answer]

Q3.4 How can we find the greatest common divisor of two positive integers? [Answer]

Q3.5 What is the greatest common divisor of two polynomials? [Answer]

Q3.6 How can we find the greatest common divisor of two polynomials? [Answer]

Q3.7 What is the modulo operation of two integers? [Answer]

Q3.8 What is the multiplicative inverse of a integer? [Answer]

Q3.9 What is the modulo operation of two polynomials? [Answer]

Q3.1 What is a prime number?

Answer: Prime numbers are numbers that have only improper divisors. For example, 17 is only divisible by 1 and 17, 3 only by 1 and 3. The number 1 is not counted among the prime numbers so that the sequence is $\{2, 3, 5, 7, 11, 13, 17, 19, ...\}$.

[Back to FAQs list.]

Q3.2 When are two integers relatively prime?

Answer: If two integers a and b have no common divisor (except 1), then a and b are called *coprime* or *relatively prime*.

[Back to FAQs list.]

Q3.3 What is the greatest common divisor (GCD) of two positive integers?

Answer: If a and b are any two integers, then the greatest common divisor is the largest divisor common to both a and b. For example, $360 = 2^3 \cdot 3^2 \cdot 5$ and $96 = 2^5 \cdot 3$ so the greatest common divisor of 96 and 360 is $2^3 \cdot 3 = 24$. We usually use the notation $\gcd(A, B)$ to denote the GCD of the numbers A and B.

[Back to FAQs list.]

Q3.4 How can we find the greatest common divisor of two positive integers?

Answer: The GCD of two positive numbers can be determined by using the Euclid's algorithm. We illustrate this method by determining the GCD of 3145 and 2992. First, we divide the larger number(3145) by the smaller number(2992) and get the remainder(153):

$$3145 = 2992 \cdot 1 + 153.$$

Next, we divide the smaller number(2992) by the remainder above(153) and get the remainder(85):

$$2992 = 153 \cdot 19 + 85.$$

Then divide the divisor above(153) by the remainder above(85) and get the remainder(68):

$$153 = 85 \cdot 1 + 68.$$

Repeat this recurrence process till the final remainder is 0:

$$\begin{aligned} 85 &= 68 \cdot 1 + 17, \\ 68 &= 17 \cdot 4 + 0. \end{aligned}$$

The GCD of 3145 and 2992 is the last nonzero remainder 17.

[Back to FAQs list.]

Q3.5 What is the greatest common divisor of two polynomials?

Answer: The greatest common divisor of two polynomials is the unique common divisor of the highest degree. For example,

$$\begin{aligned} a(z) &= 4z^{-3} + 8z^{-2} + 5z^{-1} + 3 = (1 + z^{-1} + 2z^{-2})(3 + 2z^{-1}), \\ b(z) &= 2z^{-2} + 11z^{-1} + 12 = (4 + z^{-1})(3 + 2z^{-1}), \end{aligned}$$

so the greatest common divisor is $3 + 2z^{-1}$.

[Back to FAQs list.]

Q3.6 How can we find the greatest common divisor of two polynomials?

Answer: The greatest common divisor of two polynomials can be found by continued fraction expansion. For example,

$$\begin{aligned} \frac{a(z)}{b(z)} &= \frac{4z^{-3} + 8z^{-2} + 5z^{-1} + 3}{2z^{-2} + 11z^{-1} + 12}, \\ &= 2z^{-1} - 7 + 29\frac{2z^{-1} + 3}{2z^{-2} + 11z^{-1} + 12}, \\ &= 2z^{-1} - 7 + 29\frac{1}{(2z^{-2} + 11z^{-1} + 12)/(2z^{-1} + 3)}, \\ &= 2z^{-1} - 7 + 29\frac{1}{z^{-1} + 4}, \end{aligned}$$

so the greatest common divisor is $2z^{-1} + 3$.

[Back to FAQs list.]

Q3.7 What is the modulo operation of two integers?

Answer: The modulo operation of integer X over integer N is the residue of X divided by N and is denoted as $\langle X \rangle_N$. When negative numbers are used, $\langle X \rangle_N$ has the same sign as N. For example, $\langle 67 \rangle_{13} = 67 - 5 \cdot 13 = 2$, $\langle 67 \rangle_{-13} = 67 - (-13) \cdot (-6) = -11$ and $\langle -67 \rangle_{13} = -67 - 13 \cdot (-6) = 11$.

[Back to FAQs list.]

Q3.8 What is the multiplicative inverse of a integer?

Answer: The notation

$$\langle X^{-1} \rangle_N$$

denotes the multiplicative inverse of X evaluated modulo N.
If $\langle X^{-1} \rangle_N = \alpha$, then $\langle X\alpha \rangle_N = 1$. For example, $\langle 3^{-1} \rangle_4 = 3$ because $\langle 3 \cdot 3 \rangle_4 = 1$, and $\langle 8^{-1} \rangle_5 = 2$ because $\langle 8 \cdot 2 \rangle_5 = 1$.

[Back to FAQs list.]

Q3.9 What is the modulo operation of two polynomials?

Answer: In the case of polynomial, the operation $a(z)$ mod $b(z)$ is the residue $r(z)$ after the polynomial division $a(z)/b(z)$. For example, if $a(z) = 4z^{-3} + 2z^{-2} + 5z^{-1} + 1$ and $b(z) = z^{-2} + 3z^{-1} + 4$ then the residue after the division

$$\frac{a(z)}{b(z)} = 4z^{-1} - 10 + \frac{19z^{-1} + 41}{z^{-2} + 3z^{-1} + 4}$$

is $19z^{-1} + 41$. Therefore, $a(z) \bmod b(z) = r(z) = 19z^{-1} + 41$.

[Back to FAQs list.]

References

[Dur59] J. Durbin. Efficient estimation of parameters in moving average model. *Biometrika*, 46:306–316, 1959.

[Fle87] R. Fletcher. *Practical Methods of Optimization*. Wiley, New York, NY, 1987.

[Lev47] N. Levinson. The wiener RMS criterion in filter design and prediction. *J. Math. Phys.*, 25:261–278, 1947.

[Cha95] M.V. Chan, J.A. Heinen, and R.J. Niederjohn. Formulas for the impulse response of a digital filter with an arbitrary piecewise-linear frequency response. *IEEE Trans. on Signal Processing*, 43:308–310, 1995.

[Par87] T.W. Parks and C.S. Burrus. *Digital Filter Design*. Wiley, New York, NY, 1987.

[Sel96] I.W. Selesnick, M. Lang, and C. S. Burrus. Constrained least square design of FIR filters without specified transition bands. *IEEE Trans. on Signal Processing*, 44:1879–1892, 1996.

[Sel98] I.W. Selesnick, M. Lang, and C. S. Burrus. A Modified algorithm for constrained least square design of multiband FIR filters without specified transition bands. *IEEE Trans. on Signal Processing*, 46:497–501, 1998.