L4: Signals and transforms

Analog and digital signals

Fourier transforms

Z transform

Properties of transforms

This lecture is based on chapter 10 of [Taylor, TTS synthesis, 2009]

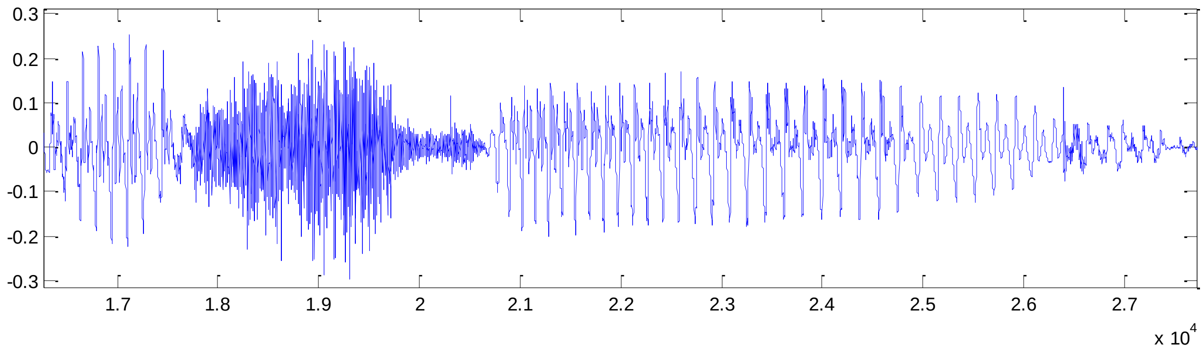
Analog signals

Signals

- A signal is a pattern of variation that encodes information
- A signal that varies over time is generally represented by a waveform
- A signal that varies continuously (i.e. speech) is called an analog signal and can be denoted as $x(t)$

Types of signals

- Periodic signals are those that repeat themselves over time, whereas aperiodic are those that do not
- Voiced speech signals are quasi-periodic since they do not repeat themselves exactly (i.e., due to jitter, shimmer and other causes)



Sinusoids

– Sinusoids are the basis for many DSP techniques as well as many processes in the physical world that oscillate
– A sinusoid signal can be represented by a sine or a cosine wave

$$x(t) = \sin(t)$$
$$x'(t) = \cos(t)$$

– The only difference between both signals is a phase shift φ of 90 degrees or $\pi/2$ radians

$$sin(t) = \cos(t - \pi/2)$$

Period and frequency

– Period (T): time elapsed between two repetitions of the signal
  - Measured in units of time (seconds)
– Frequency (F): # of times that a signal repeats per unit of time
  - Measured in hertz (Hz) (cycles per second)
  - Frequency is the reciprocal of period: $F = 1/T$

- To change the frequency of a sinusoidal, we multiply time by $2\pi F$, where F is measured in Hz

$$x(t) = \cos(2\pi F t + \varphi)$$

- To scale the signal, we then multiply by parameter A, its amplitude

$$x(t) = \mathrm{A}\cos(2\pi F t + \varphi)$$

- And to avoid having to write 2π every time, we generally use angular frequency ω, which has units of radians per second (1 cycle=2π rad)

$$x(t) = \mathrm{A}\cos(\omega t + \varphi)$$

ex4p1.m

Generate various sine waves with different phases, amplitudes and frequencies

General periodic signals

- Periodic signals do not have to be sinusoidal, they just have to meet

$$x(t) = x(t+T) = x(t+2T) = \cdots = x(t+nT)$$

  for some value $T = T_0$, which is called its fundamental period

- The reciprocal $F_0 = 1/T_0$ is called the fundamental frequency
- A harmonic frequency is any integer multiple of the fundamental frequency, $2F_0$, $3F_0$, …

Fourier synthesis

- It can be shown that ANY periodic signal can be represented as a sum of sinusoidals whose frequencies are harmonics of F_0

$$x(t) = a_0 cos(\varphi_0) + a_1 cos(\omega_0 t + \varphi_1) + a_2 cos(2\omega_0 t + \varphi_2) + \cdots$$

  - i.e., for appropriate values of the amplitudes a_k and phases φ_k

- which can be written in compact form as

$$x(t) = A_0 + \sum_{k=1}^{\infty} a_k cos(k\omega_0 t + \varphi_k)$$

- The above is known as the Fourier series

Exercise

– Synthesize a periodic square wave $x(t)$ as a sum of sinusoidals

$$x(t) = \begin{cases} 1 & 0 \le t \le T_0/2 \\ -1 & T_0/2 \le t \le T_0 \end{cases}$$

– It can be shown that the square wave can be generated by adding the odd harmonics, each having the same phase $\varphi = -\pi/2$, and the following amplitudes

$$a_k = \begin{cases} 4/(k\pi) & k = 1,3,5 \dots \\ 0 & k = 0,2,4 \dots \end{cases}$$

ex4p2.m
Generate code to reconstruct this signal

– This is all very interesting, but eventually we would like to do the reverse: estimate the parameters a_k from the signal $x(t)$

- This reverse problem is known as Fourier analysis, and will be described in a few slides

Sinusoids as complex exponentials

- A different representation of the sinusoidal $x(t) = \mathrm{A}\cos(\omega t + \varphi)$ greatly simplifies the mathematics
- This representation is based on Euler's formula

$$e^{j\theta} = cos\theta + jsin\theta$$

  - where $j = \sqrt{-1}$, and $e^{j\theta}$ is a complex number with real part $cos\theta$ and imaginary part $sin\theta$
- The inverse Euler formulas are

$$cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2}; sin\theta = \frac{e^{j\theta} - e^{-j\theta}}{2j}$$

- If we add amplitude A and set $\theta = \omega t + \varphi$

$$Ae^{j(\omega t+\varphi)} = Acos(\omega t + \varphi) + jAsin(\omega t + \varphi)$$

  - This representation seems quite crazy, but it does simplify the math

- As an example, consider the following decomposition of the complex sine wave

$$x(t) = Ae^{j(\omega t+\varphi)} = Ae^{j\varphi}e^{j\omega t} = Xe^{j\omega t}$$

  - Since $Ae^{j\varphi}$ is a constant, it can be combined with the amplitude

$$x(t) = Xe^{j\omega t}$$

    - such that the pure sine part $e^{j\omega t}$ is now free of phase information
  - In practice, real signals (i.e., speech) do not have imaginary part, so one can simply ignore it
- Combining this with the Fourier synthesis equation yields a more general expression

$$x(t) = \sum_{-\infty}^{\infty} a_k e^{jk\omega_0 t}$$

where $a_k = A_k e^{j\varphi_k}$

  - It can be shown that for real-valued signals, the complex amplitudes are conjugate symmetric ($a_k = -a_k$), so the negative harmonics do not add information and the signal can be reconstructed by summing from 0 to ∞

Fourier analysis

– Given a periodic signal $x(t)$, the coefficients a_k can be derived from the Fourier analysis equation:

$$a_k = \frac{1}{T_0}\int_0^{T_0} x(t)e^{-jk\omega_0 t}dt$$

– Example: compute the Fourier analysis for $x(t) = b_n e^{jn\omega_0 t}$

$$a_k = \frac{1}{T_0}\int_0^{T_0} b_n e^{jn\omega_0 t}e^{-jk\omega_0 t}dt = \frac{b_n}{T_0}\int_0^{T_0} e^{j(n-k)\omega_0 t}dt$$

- For $n = k$, the integrand is 1, the integral is T_0, and $a_k = b_n$ (as expected)
- For $n \neq k$, we have the integral of a (complex) sine wave over a multiple of its period, which integrates to zero:

$$\int_0^{T_0} e^{jk\omega_0 t}dt = 0$$

– The representation of periodic signal $x(t)$ in terms of its harmonic coefficients a_k is known as the spectrum

- Hence, we can represent a signal in the time domain (a waveform) or in the frequency domain (a spectrum)

Magnitude and phase spectrum

- Rather than plotting the spectrum of a signal in terms of its real and imaginary parts, one generally looks at the magnitude and phase
- The human ear is largely insensitive to phase information
  - As an example, if you play a piano note and then again a while later, both sound identical
- This result holds when you have a complex signal
  - If we synthesize a signal with the same magnitude spectrum of a square wave and arbitrary phase it will sound the same as the square wave
  - Nonetheless, the two waveforms may look very different (see below)!
- For these reasons, one generally studies just the magnitude spectrum

Exercise

<u>ex4p3.m</u>

- Synthesize a square wave of the previous exercise, now with a different phase
- Plot both and show they look very different
- Play both and show they sounds similar

The Fourier transform

– In general we will need to analyze non-periodic signals, so the previous Fourier synthesis/analysis equations will not suffice
– Instead, we use the Fourier transform, defined as

$$X(j\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$$

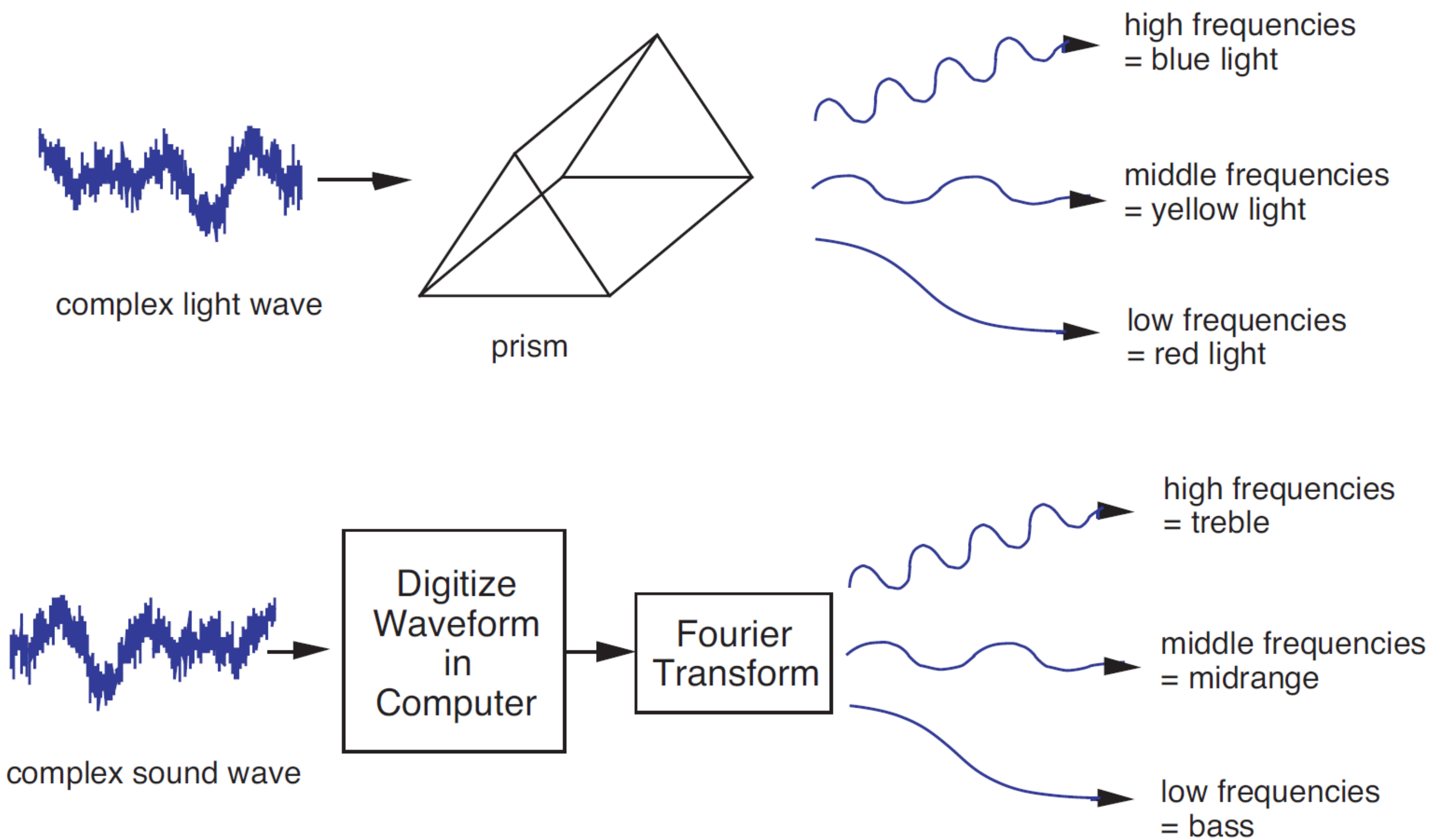
  - Compare with the Fourier analysis equation $a_k = \frac{1}{T_0}\int_0^{T_0} x(t)e^{-jk\omega_0 t}dt$
    – The integral is over $(-\infty, \infty)$ since the signal is aperiodic
    – The result is a continuous function over frequency, rather than over a discrete set of harmonics

– And the inverse Fourier transform is defined as

$$x(t) = \int_{-\infty}^{\infty} X(j\omega)e^{-j\omega t}dt$$

  - For which the same discussion holds when compared to the Fourier synthesis equation $x(t) = \sum_{-\infty}^{\infty} a_k e^{jk\omega_0 t}$

The Fourier transform as a sound “prism”



From [Sethares (2007). Rhythms and transforms]

Digital signals

A digital signal $x[n]$ is a sequence of numbers

$$x[n] = \cdots x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots$$

- Each point in the sequence is called a sample, and the distance (in time) between two samples is called the sampling period T_S
  - Likewise, the sample rate or sample frequency is $F_S = 1/T_S$
- For a given sample frequency F_S, the highest frequencies that $x[n]$ can represent is $F_S/2$; this is known as the <u>Nyquist frequency</u>
- The <u>bit range</u> describes the dynamic range of the digital signal, and is given by the number of bits used to store the signal
  - With *16* bits, you can represent 2^{16} values, from -32768 to 32767

Normalized frequency

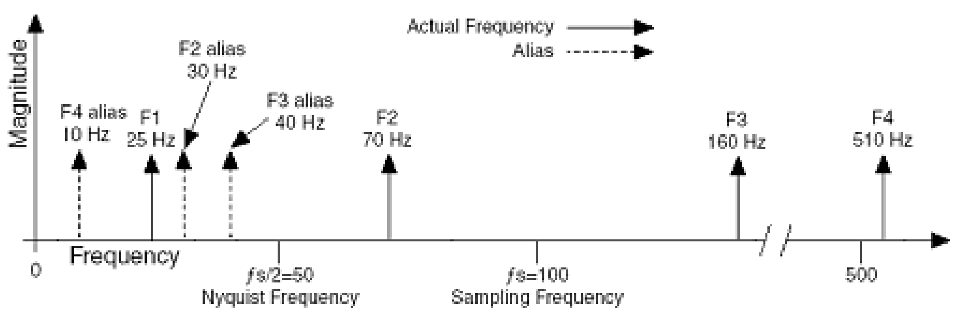
- With digital signals, one generally uses the normalized frequency

$$\hat{\omega} = \omega/F_S = 2\pi F/F_S$$

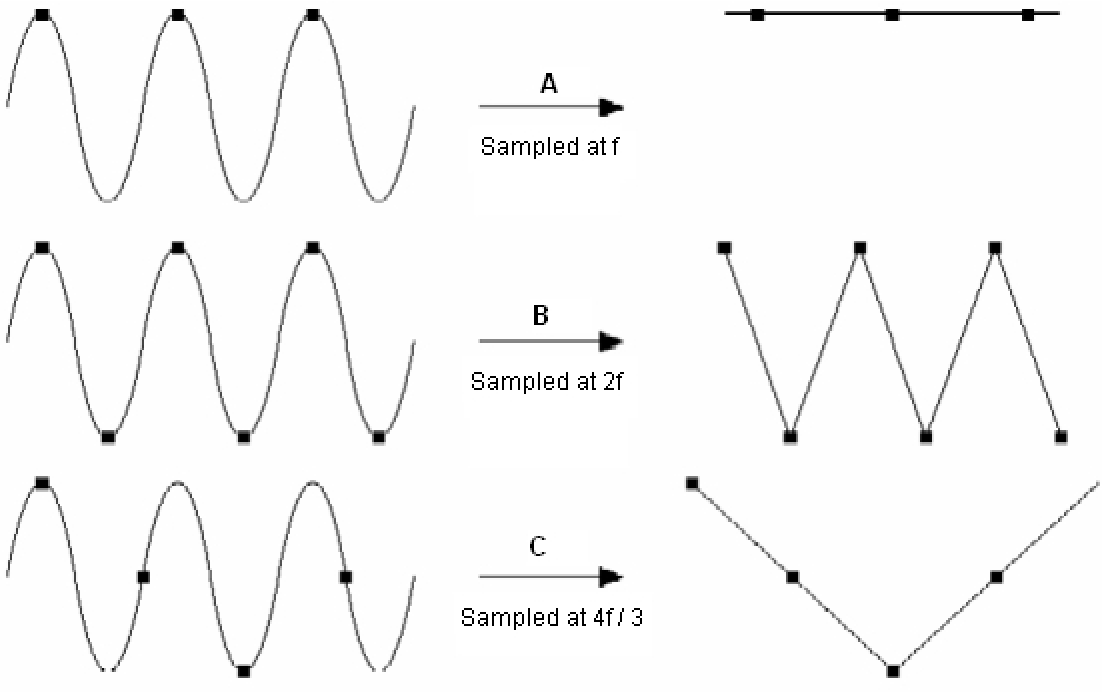
  - This will come in handy when you try to convert indices in the FFT into real frequencies (Hz)

Aliasing

– Occurs when the signal contains frequencies above $F_S/2$ (video)
– These frequencies appear as mirrored within the Nyquist range
– Assume a signal with frequencies at 25Hz, 70Hz, 160Hz, and 510Hz, and a sampling frequency $F_S = 100Hz$
  - When sampled, the 25Hz component appears correctly
  - However, the remaining components appear mirrored
    – Alias F2: |100 – 70| = 30 Hz
    – Alias F3: |2×100 – 160| = 40 Hz
    – Alias F4: |5×100 – 510| = 10 Hz



http://zone.ni.com/devzone/cda/tut/p/id/3016



aliasing

aliasing

http://zone.ni.com/cms/images/devzone/tut/a/0f6e74b4493.jpg

The discrete-time Fourier transform (DTFT)

- Taking the expression of the Fourier transform $X(j\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$ and noticing that $x[n] = x(nT_S)$, the DTFT can be derived by numerical (rectangular) integration

$$X(j\omega) = \sum_{-\infty}^{\infty} \left(x(nT_S)e^{-j\omega nT_S}\right) \times T_S$$

- which, using the normalized frequency $\hat{\omega}$, becomes

$$X\left(e^{j\hat{\omega}}\right) = \sum_{-\infty}^{\infty} x[n]e^{-j\hat{\omega}n}$$

  - where the multiplicative term T_S has been neglected

- Note that the DTFT is discrete in time but still continuous in frequency
- In addition, it requires an infinite sum, which is not useful for computational reasons

Discrete Fourier transform (DFT)

- The DFT is obtained by "sampling" the spectrum at N discrete frequencies $\omega_k = 2\pi F_s/N$, which yields the transform

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn}$$

- Interpretation
  - For each required frequency value $X[k]$, we compute the inner value of our signal $x[n]$ with sine wave $\exp(-j2\pi kn/N)$
  - The result is a complex number that describes the magnitude and phase of $x[n]$ at that frequency
- Frequency resolution
  - Note that the number of time samples in $x[n]$ is the same as the number of discrete frequencies in $X[k]$
  - Therefore, the longer the waveform, the better frequency resolution we can achieve
    - As we saw in the previous lectures, though, with speech there is a limit to how long of a sequence we want to use since the signal is not stationary

– Both the DTFT and DFT have inverse transforms, defined by

$$x[n] = \frac{1}{2\pi}\int_{-\infty}^{\infty} X\left(e^{j\widehat{\omega}}\right)e^{j\widehat{\omega}n}d\widehat{\omega}$$

$$x[n] = \frac{1}{N}\sum_{n=0}^{N-1} X[k]e^{j\frac{2\pi}{N}kn}$$

The DFT as a matrix multiplication

- Denoting $W_n = e^{-j2\pi/N}$, the DFT can be expressed as

$$X[k] = \sum_{n=0}^{N-1} x[n](W_n)^{kn}$$

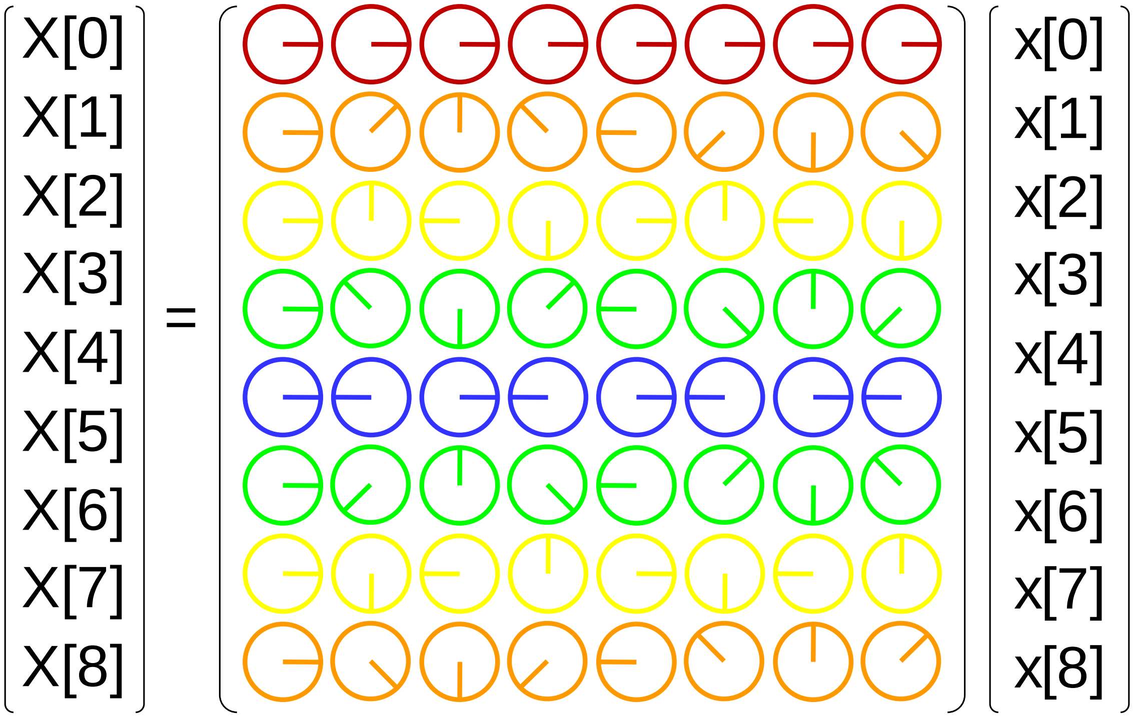
- Or using matrix notation:

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ \\ X(N-1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & W_N & W_N^2 & & W_N^{N-1} \\ 1 & W_N^2 & W_N^4 & & W_N^{2(N-1)} \\ 1 & & & & \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & & W_N^{(N-1)(N-1)} \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ \\ x(N-1) \end{bmatrix}$$

  - So the DFT can also be thought of as a projection of the time series data by means of a complex-valued matrix

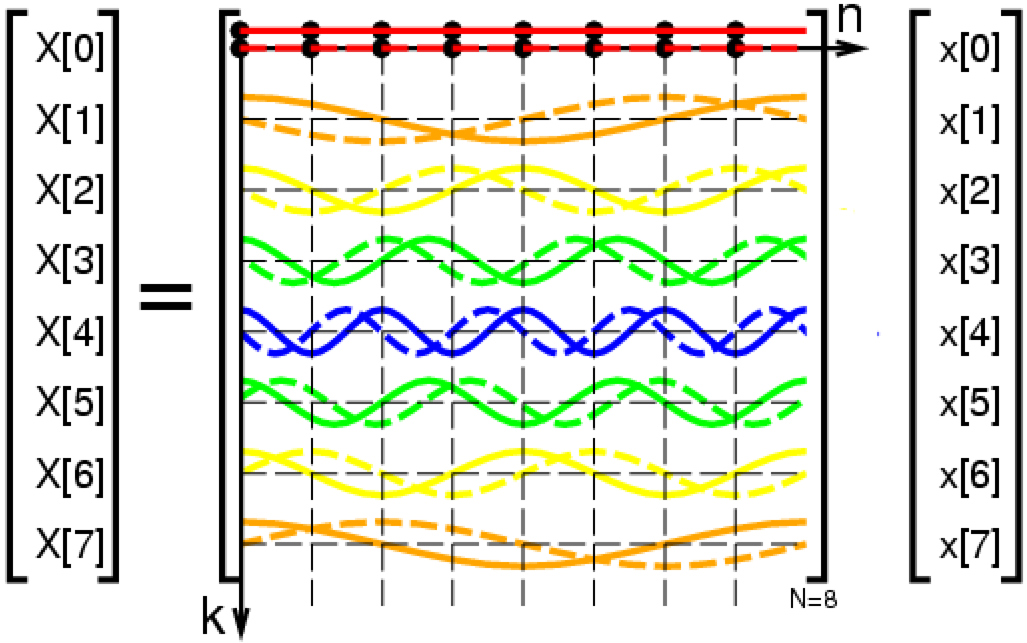
http://en.wikipedia.org/wiki/DFT_matrix

– Note that the k^{th} row of the DFT matrix consist of a unitary vector rotating clockwise with a constant increment of $2\pi k/N$



- The second and last row are complex conjugates
- The third and second-to-last rom are complex conjugates…

– So, expressing these rotating unitary vectors in terms of the underlying sine waves, we obtain



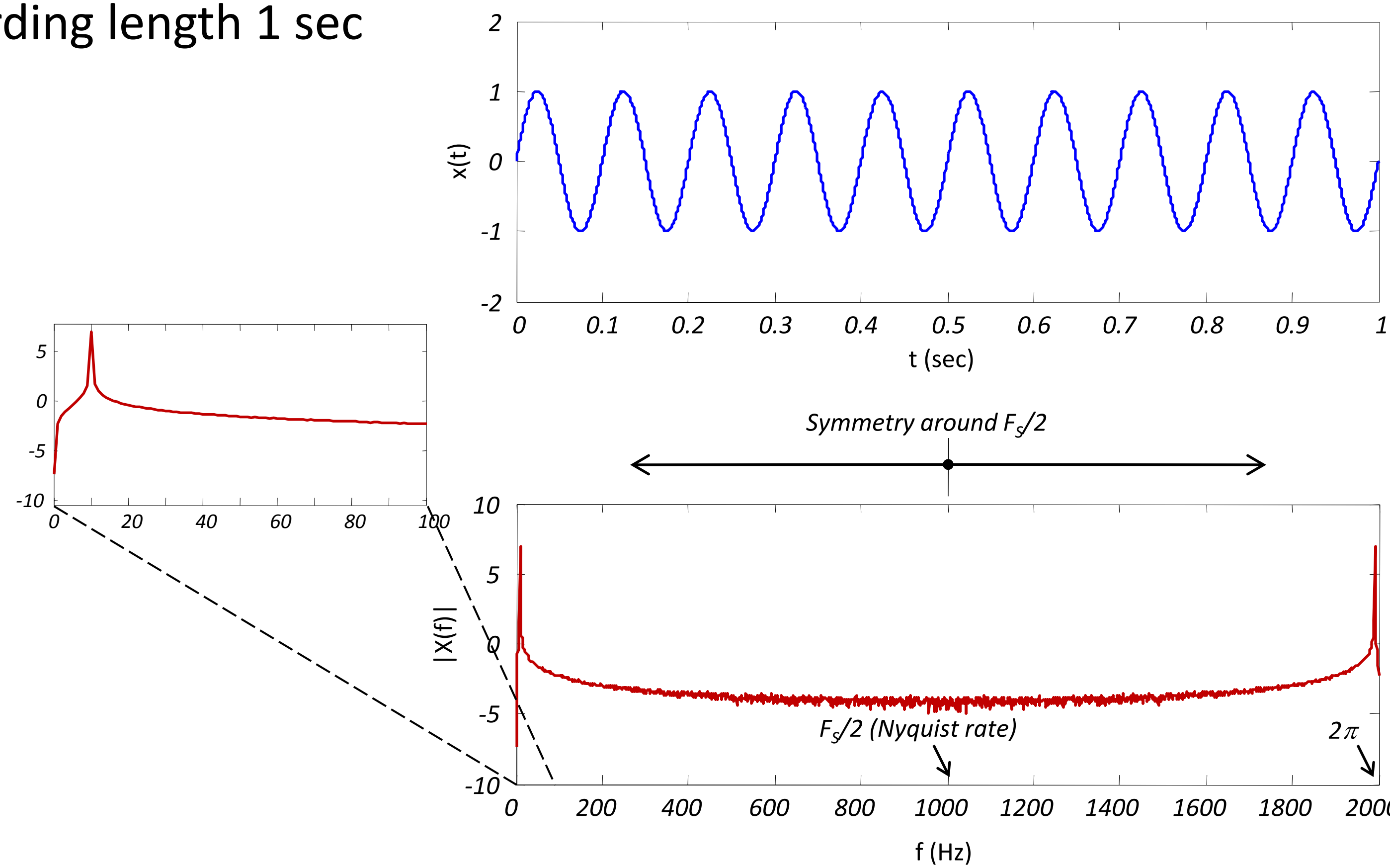
- where the solid line represents the real part and the dashed line represent the imaginary part of the corresponding sine wave

– Note how this illustration brings us back to the definition of the DFT as an inner product between our signal x[k] and a complex sine wave

http://en.wikipedia.org/wiki/DFT_matrix

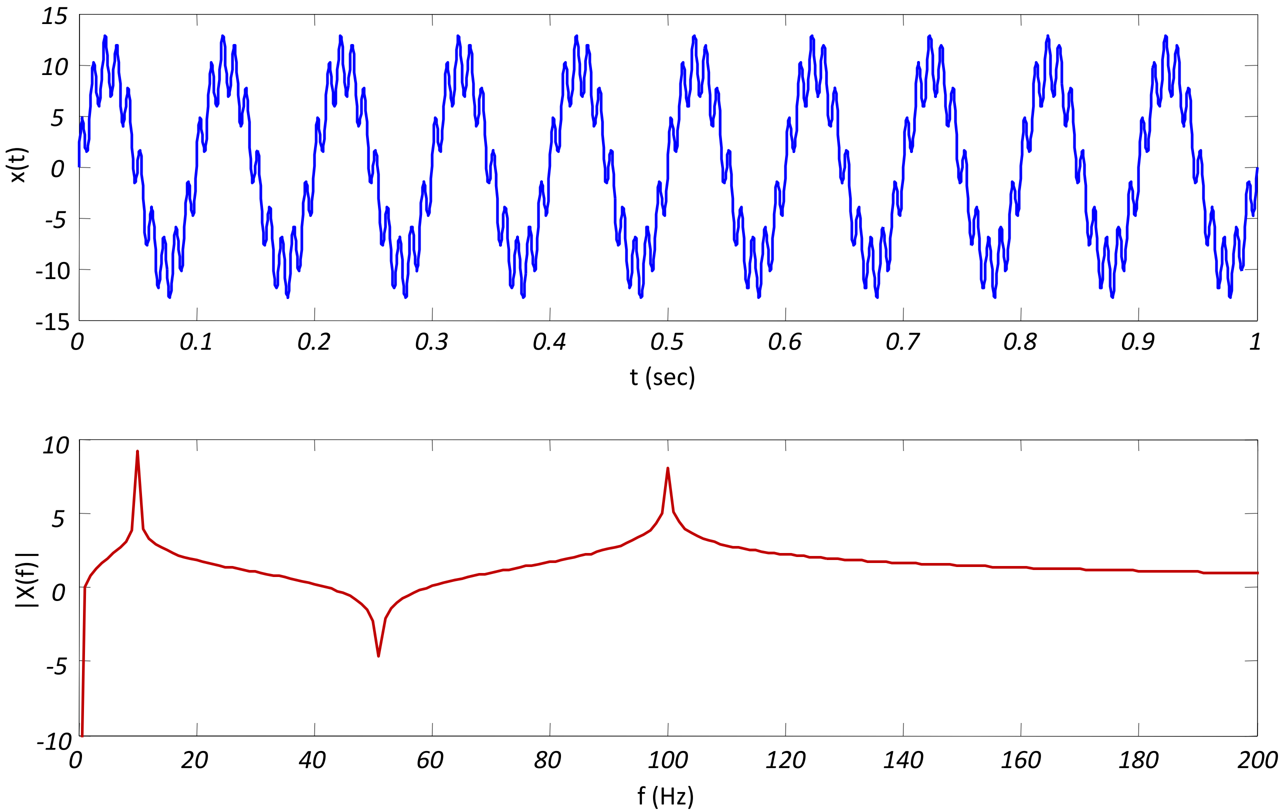
Example (1)

- Sampling rate $F_S = 2kHz$
- Signal $x(t) = \sin(2\pi 10t)$
- Recording length 1 sec



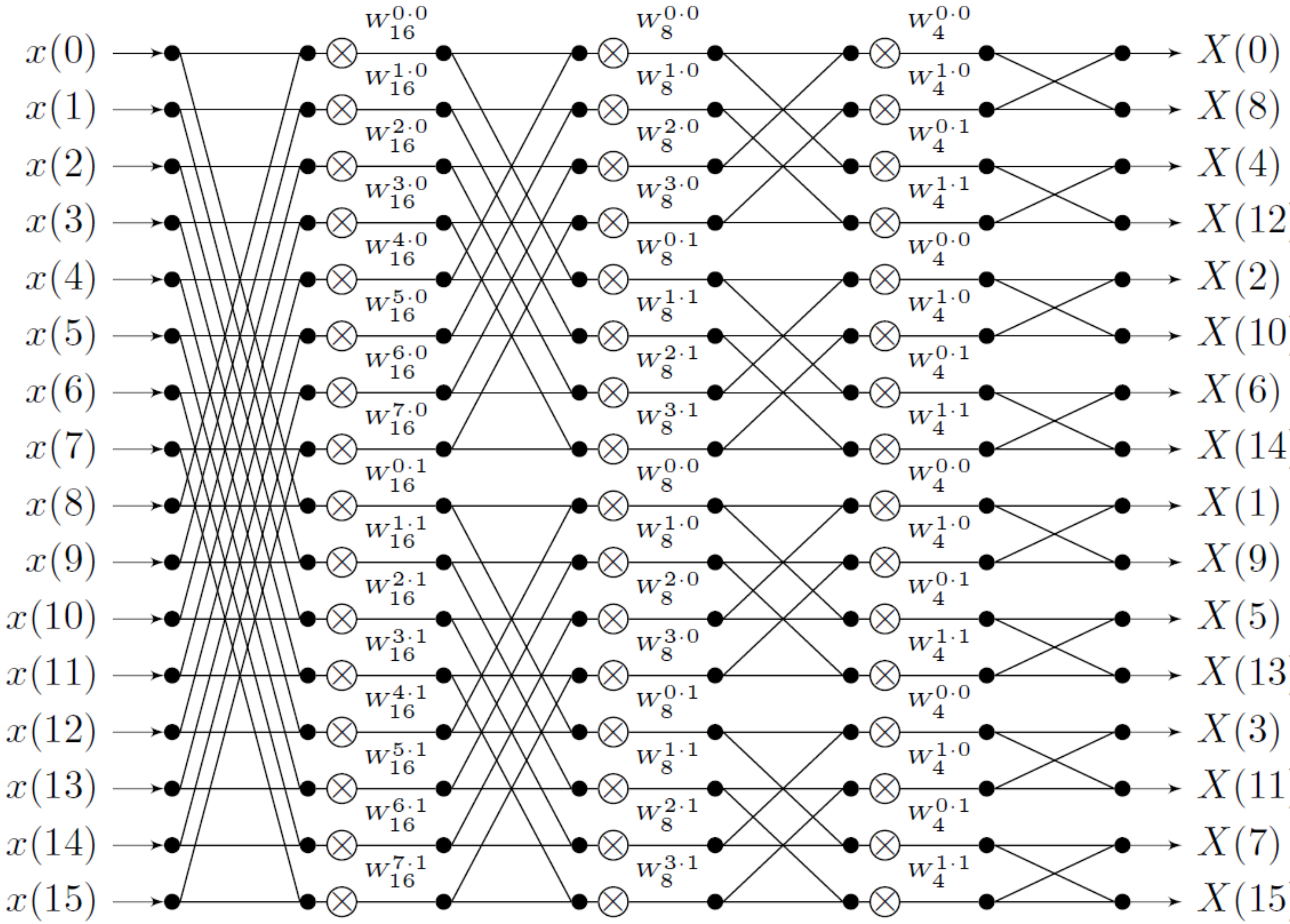
Example (2)

- Sampling rate $F_S = 2kHz$
- Signal $x(t) = 10\sin(2\pi 10t) + 3\sin(2\pi 100t)$
- Recording length 1 sec



Fast Fourier Transform (FFT)

- The FFT is an efficient implementation of the DFT
  - The DFT runs in $O(N^2)$, whereas FFT algorithms run in $O(N\log_2 N)$
- Several FFT algorithms exists, but the most widely used are radix-2 algorithms, which require $N = 2^k$ samples
  - If the time signal does not have the desired number of samples, one simply "pads" the signal with extra zeros



$x(0)$
$x(1)$
$x(2)$
$x(3)$
$x(4)$
$x(5)$
$x(6)$
$x(7)$
$x(8)$
$x(9)$
$x(10)$
$x(11)$
$x(12)$
$x(13)$
$x(14)$
$x(15)$
$w_{16}^{0\cdot 0}$
$w_{16}^{1\cdot 0}$
$w_{16}^{2\cdot 0}$
$w_{16}^{3\cdot 0}$
$w_{16}^{4\cdot 0}$
$w_{16}^{5\cdot 0}$
$w_{16}^{6\cdot 0}$
$w_{16}^{7\cdot 0}$
$w_{16}^{0\cdot 1}$
$w_{16}^{1\cdot 1}$
$w_{16}^{2\cdot 1}$
$w_{16}^{3\cdot 1}$
$w_{16}^{4\cdot 1}$
$w_{16}^{5\cdot 1}$
$w_{16}^{6\cdot 1}$
$w_{16}^{7\cdot 1}$
$w_{8}^{0\cdot 0}$
$w_{8}^{1\cdot 0}$
$w_{8}^{2\cdot 0}$
$w_{8}^{3\cdot 0}$
$w_{8}^{0\cdot 1}$
$w_{8}^{1\cdot 1}$
$w_{8}^{2\cdot 1}$
$w_{8}^{3\cdot 1}$
$w_{8}^{0\cdot 0}$
$w_{8}^{1\cdot 0}$
$w_{8}^{2\cdot 0}$
$w_{8}^{3\cdot 0}$
$w_{8}^{0\cdot 1}$
$w_{8}^{1\cdot 1}$
$w_{8}^{2\cdot 1}$
$w_{8}^{3\cdot 1}$
$w_{4}^{0\cdot 0}$
$w_{4}^{1\cdot 0}$
$w_{4}^{0\cdot 1}$
$w_{4}^{1\cdot 1}$
$w_{4}^{0\cdot 0}$
$w_{4}^{1\cdot 0}$
$w_{4}^{0\cdot 1}$
$w_{4}^{1\cdot 1}$
$w_{4}^{0\cdot 0}$
$w_{4}^{1\cdot 0}$
$w_{4}^{0\cdot 1}$
$w_{4}^{1\cdot 1}$
$w_{4}^{0\cdot 0}$
$w_{4}^{1\cdot 0}$
$w_{4}^{0\cdot 1}$
$w_{4}^{1\cdot 1}$
$X(0)$
$X(8)$
$X(4)$
$X(12)$
$X(2)$
$X(10)$
$X(6)$
$X(14)$
$X(1)$
$X(9)$
$X(5)$
$X(13)$
$X(3)$
$X(11)$
$X(7)$
$X(15)$

The Z transform

- The Z transform is defined as

$$X(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

  - which is the familiar DTFT for $z = e^{j\widehat{\omega}}$

- The Z transform is the most practical of all the transforms in digital signal processing because it allows us to manipulate signals and filters as polynomials (in z^{-1})

$$X(z) = \sum_{n=-\infty}^{\infty} x[n] (z^{-1})^{n}$$

The Laplace transform

- A generalization of the Z transform for continuous-time signals
- The Laplace transform is to the Fourier transform what the Z transform is to the DTFT
- The Laplace transform is not required here since we will always work with discrete-time signals (i.e., after they are sampled)

Frequency domain for digital signals

- For analog signals, the frequency domain extends from $-\infty$ to ∞
  - For digital signals, however, we know that the Nyquist frequency $(F_S/2)$ is the highest that can be represented by the signal
  - Thus, the spectrum for $|f| > F_S/2$ contains no new information
- What happens beyond the Nyquist range?
  - It can be shown that the spectrum repeats itself at multiples of the Nyquist frequency, or at multiples of 2π for the normalized frequency $\hat{\omega}$
  - In other words, the spectrum of a digital signal is <u>periodic</u>
  - For this reason, the spectrum is described as $X\left(e^{j\hat{\omega}}\right)$ rather than as $X(\hat{\omega})$

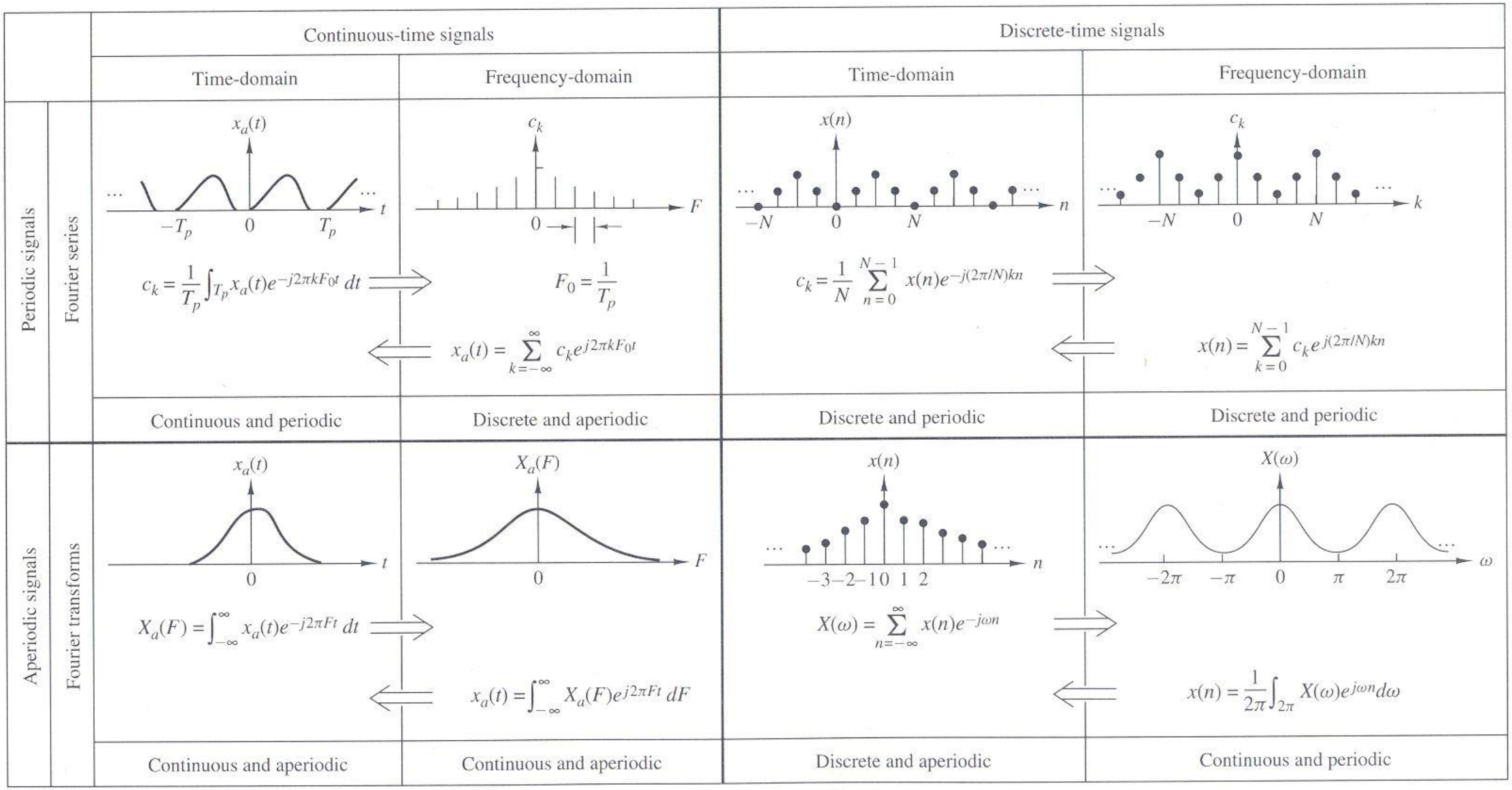


Figure 4.27 Summary of analysis and synthesis formulas.

[Proakis & Malonakis, 1996]

Properties of the transforms

A number of properties hold for all these transforms

	Discrete Time	Continuous time
Discrete Frequency	Discrete Fourier Transform (DFT)	Fourier analysis, periodic waveform
Continuous Frequency	Discrete Fourier Transform (DTFT)	Fourier Transform
Continuous variable	z-transform	Laplace Transform

Linearity

- Taking the Fourier transform as an example, this means that
  - if $x(t) = \alpha s_1(t) + \beta s_2(t)$,
  - then $X(j\omega) = \alpha X_1(j\omega) + \beta X_2(j\omega)$

Time and frequency duality

- It can be shown that

$$g(t) \overset{F}{\rightarrow} G(\omega)$$
$$g(\omega) \overset{IF}{\rightarrow} G(t)$$

  - To convince yourself, note that the Fourier transform and its inverse have very similar forms

Time delay

- Delaying a signal by t_d is equivalent to multiplying its Fourier transform by $e^{-j\omega t_d}$

$$x(t) \leftarrow X(j\omega)$$
$$x(t-t_d) \leftarrow X_d(j\omega) = X(j\omega)\, e^{-j\omega t_d}$$

  - Note that $e^{-j\omega t_d}$ does not affect the magnitude of $X_d(j\omega)$, only its phase by a linear delay of t_d, as we should expect

Frequency shift

- From the duality principle, we can then infer that multiplying a signal by $e^{j\omega_0 t}$ causes a shift of ω_0 in its Fourier transform

$$x(t)e^{j\omega_0 t} \leftarrow X\big(j(\omega-\omega_0)\big)$$

- Thus, a shift in the frequency domain corresponds to modulation in the time domain
  - To see this, note that the Fourier transform of signal $x(t) = e^{j\omega_0 t}$ (a sine wave) is $2\pi\delta(\omega-\omega_0)$, that is, a single impulse a frequency ω_0
- This property will become handy when we introduce the STFT

Scaling

- Compression of a signal in time will stretch its Fourier transform, and vice versa

$$x(at) \xrightarrow{F} \frac{1}{|a|} X(j\omega/a)$$

Impulse properties

- If we compress the time signal more and more, we reach a unit impulse $\delta[n]$, which has zero width
- As expected from the scaling property, the Fourier transform of an impulse will then be infinitely stretched (it is 1 at all frequencies)

$$\delta(t) \xrightarrow{F} 1$$

- and by virtue of the duality property

$$1 \xrightarrow{IF} \delta(\omega)$$

  - which is also intuitive, since a constant signal (a DC offset) has no energy at frequencies other than zero

Convolution

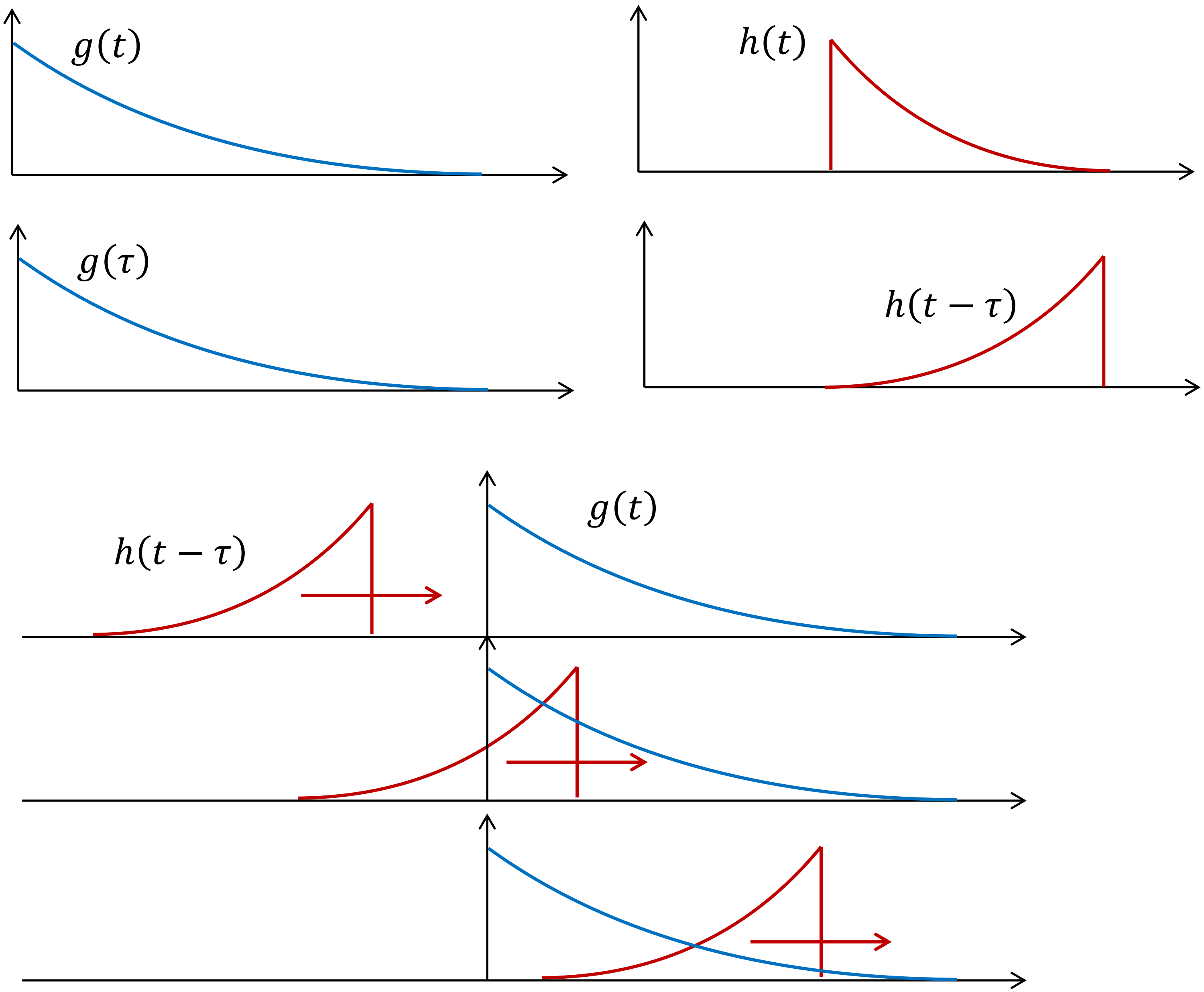
- Convolution is defined as the overlap between two functions when one is passed over the other

$$f(t) = g(t) \otimes h(t) = \int g(\tau) h(t-\tau) d\tau$$

  - Convolution is similar to correlation, with the exception that in convolution one of the signals is "flipped"
- Taking Fourier transforms on both sides, it can be shown that

$$F(\omega) = G(\omega) H(\omega)$$

  - Recall a similar expression when we discussed the vocal tract filter?
- In other words, convolution in the time domain corresponds to multiplication in the frequency domain


$g(t)$
$h(t)$
$g(\tau)$
$h(t-\tau)$
$h(t-\tau)$
$g(t)$

Stochastic signals

- All the transforms wee seen so far integrate/sum over an infinite sequence, which is meaningful only if the result is <u>finite</u>
  - This is the case for all periodic and many non-periodic signals, but is not always true; in the latter case, the Fourier transform does not exist
- As an example, for stochastic signals generated from a random process, e.g., “noisy” fricative sounds
  - It is hard to describe them in the time domain due to their random nature
  - The Fourier/Z transforms cannot be used as defined
- To avoid these issues, we analyze averages from these signals through the autocorrelation function (a measure of self-similarity)

$$R(j) = \sum_{n=-\infty}^{\infty} x[n]x[n-j]$$

  - which is the expected value of the product of signal $x[n]$ with a time-shifted version of itself
- The autocorrelation function does have a Fourier transform, which is known as the <u>power spectral density</u> of the signal

Example

ex4p4.m

Compute the autocorrelation of a noisy signal, and then compute its power spectral density