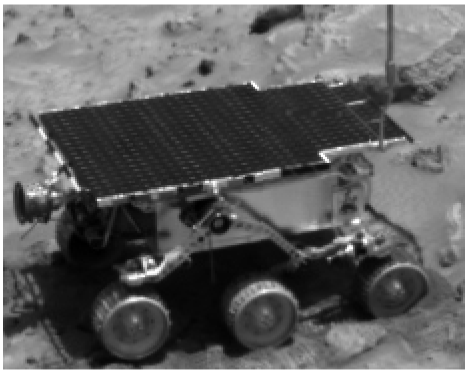
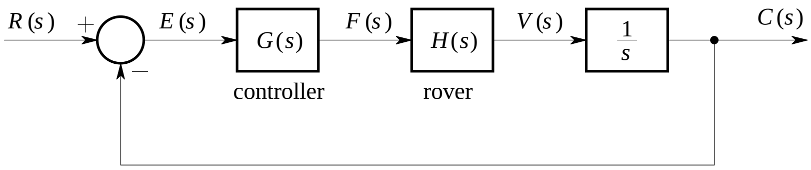
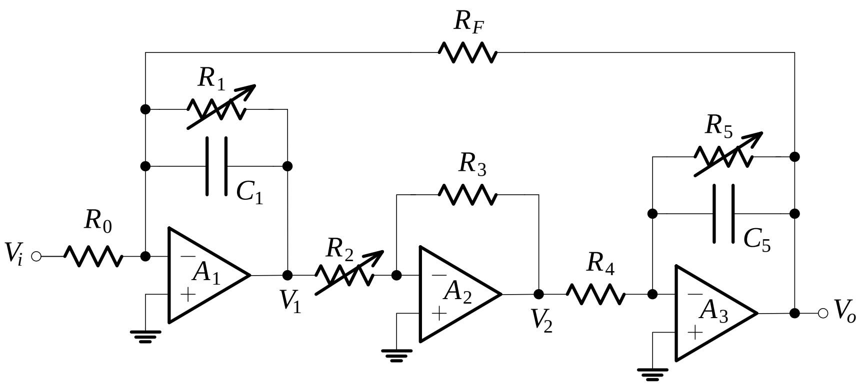
48540 Signals and Systems

Lecture Notes

2011







PMcL

Contents

Lecture 1A – Signals

Overview. Signal operations. Periodic & aperiodic signals. Deterministic & random signals. Energy & power signals. Common signals. Sinusoids.

Overview

Electrical engineers should never forget "the big picture".

Every day we take for granted the power that comes out of the wall, or the light that instantly illuminates the darkness at the flick of a switch. We take for granted the fact that electrical machines are at the heart of every manufacturing industry. There has been no bigger benefit to humankind than the supply of electricity to residential, commercial and industrial sites. Behind this "magic" is a large infrastructure of generators, transmission lines, transformers, protection relays, motors and motor drives.

We also take for granted the automation of once hazardous or laborious tasks, we take for granted the ability of electronics to control something as complicated as a jet aircraft, and we seem not to marvel at the idea of your car's engine having just the right amount of fuel injected into the cylinder with just the right amount of air, with a spark igniting the mixture at just the right time to produce the maximum power and the least amount of noxious gases as you tell it to accelerate up a hill when the engine is cold!

Why we study systems

We forget that we are now living in an age where we can communicate with anyone (and almost *anything*), anywhere, at anytime. We have a point-to-point telephone system, mobile phones, the Internet, radio and TV. We have never lived in an age so "information rich".

Electrical engineers are engaged in the business of designing, improving, extending, maintaining and operating this amazing array of systems.

You are in the business of becoming an engineer.

One thing that engineers need to do well is to break down a seemingly complex system into smaller, easier parts. We therefore need a way of describing these systems mathematically, and a way of describing the inputs and outputs of these systems - signals.

Signal Operations

There really aren't many things that you can do to signals. Take a simple FM modulator:

Example of signal operations

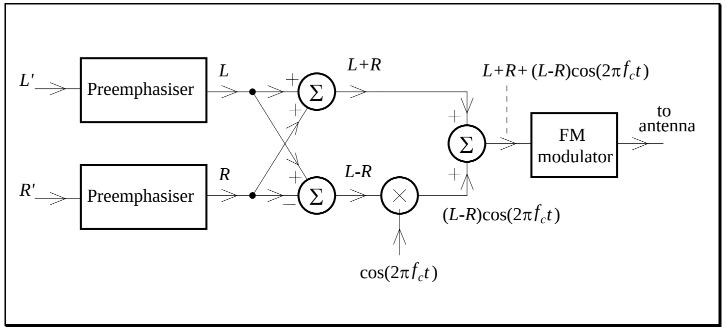


Figure 1A.1

Linear system operations

Some of the signals in this system come from "natural" sources, such as music or speech, some come from "artificial" sources, such as a sinusoidal oscillator. We can multiply two signals together, we can add two signals together, we can amplify, attenuate and filter. We normally treat all the operations as linear, although in practice some nonlinearity always arises.

We seek a way to analyse, synthesise and process signals that is mathematically rigorous, yet simple to picture. It turns out that Fourier analysis of signals and systems is one suitable method to achieve this goal, and the Laplace Transform is even more powerful. But first, let's characterise mathematically and pictorially some of the more common signal types.

Continuous and Discrete Signals

A continuous signal can be broadly defined as a quantity or measurement that varies continuously in relation to another variable, usually time. We say that the signal is a *function* of the independent variable, and it is usually described mathematically as a function with the argument in parentheses, e.g. $g(t)$. The parentheses indicate that *t* is a *real number*.

Common examples of continuous-time signals include temperature, voltage, audio output from a speaker, and video signals displayed on a TV. An example of a graph of a continuous-time signal is shown below:

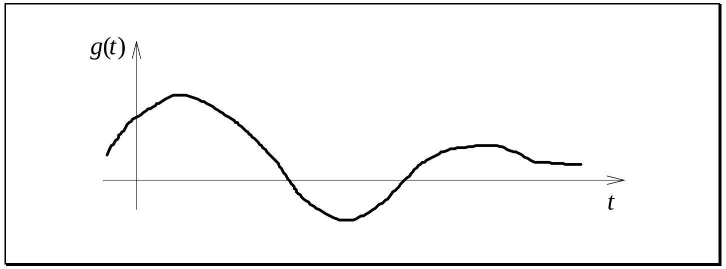


Figure 1A.2

A discrete signal is one which exists only at specific instants of the independent variable, usually time. A discrete signal is usually described mathematically as a function with the argument in brackets, e.g. $g[n]$. The brackets indicate that *n* is an *integer*.

Common examples of discrete-time signals include your bank balance, monthly sales of a corporation, and the data read from a CD. An example of a graph of a discrete-time signal is shown below:

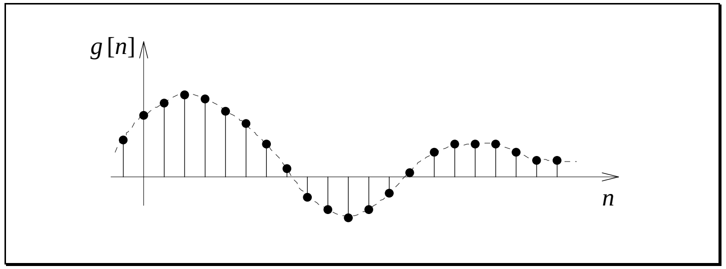


Figure 1A.3

Periodic and Aperiodic Signals

A periodic signal $g(t)$ is defined by the property:

Periodic function defined

$$\boxed{g(t) = g(t + T_0), \quad T_0 \neq 0} \tag{1A.1}$$

The smallest value of T_0 that satisfies Eq. (1A.1) is called the *period*. The *fundamental* frequency of the signal is defined as:

Fundamental frequency defined

$$\boxed{f_0 = \frac{1}{T_0}} \tag{1A.2}$$

A periodic signal remains unchanged by a positive or negative shift of any integral multiple of T_0. This means a periodic signal must begin at $t = -\infty$ and go on forever until $t = \infty$. An example of a periodic signal is shown below:

A periodic signal

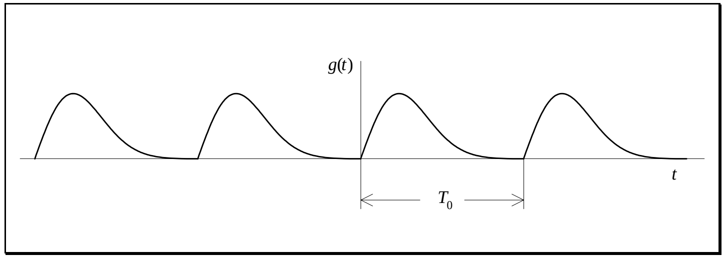


Figure 1A.

We can also have periodic discrete-time signals, in which case:

$$g[n] = g[n + T_0], \quad T_0 \neq 0 \tag{1A.3}$$

Aperiodic signals defined

An *aperiodic* signal is one for which Eq. (1A.1) or Eq. (1A.3) does not hold. Any finite duration signal is aperiodic.

Example

Find the period and fundamental frequency (if they exist) of the following signal:

$$g(t) = 3\cos 2\pi t + 4\sin 4\pi t \tag{1A.4}$$

We can graph this function to easily determine its period:

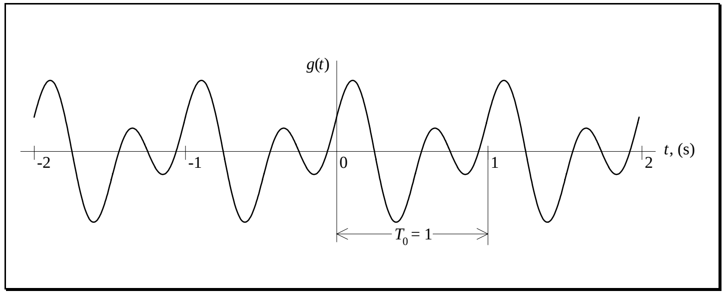


Figure 1A.5

From the graph we find that $T_0 = 1\ \text{s}$ and $f_0 = 1\ \text{Hz}$. It is difficult to determine the period mathematically (in this case) until we look at the signal's *spectrum* (Lecture 2A).

If we add two periodic signals, then the result *may* or *may not* be periodic. The result will only be periodic if an integral number of periods of the first signal coincides with an integral number of periods of the second signal:

$$\boxed{T_0 = qT_1 = pT_2 \text{ where } \frac{p}{q} \text{ is rational}} \tag{1A.5}$$

In Eq. (1A.5), the integers *p* and *q* must have no common factors.

We know that a sinusoid is a periodic signal, and we shall soon see that any signal composed of a sum of sinusoids with frequencies that are integral multiples of a fundamental frequency is also periodic.

Deterministic and Random Signals

Deterministic signals defined

A *deterministic* signal is one in which the past, present and future values are completely specified. For example, $g(t)=\cos 2\pi t$ and $g(t)=e^{-t}u(t)$ are obviously signals specified completely for all time.

Random signals defined

Random or *stochastic* signals cannot be specified at precisely one instant in time. This *does not* necessarily mean that any *particular* random signal is unknown - on the contrary they can be deterministic. For example, consider some outputs of a binary signal generator over 8 bits:

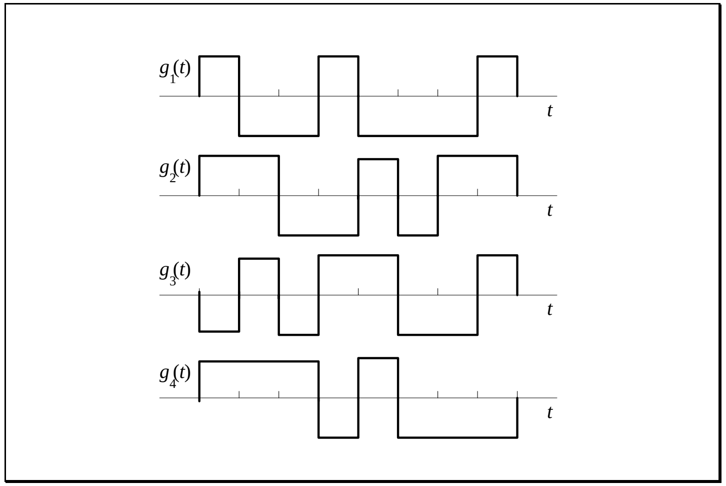


Figure 1A.6

Each of the possible $2^8=256$ waveforms is deterministic - the randomness of this situation is associated not with the waveform but with the uncertainty as to which waveform will occur. This is completely analogous to the situation of tossing a coin. We know the outcome will be a head or a tail - the uncertainty is the occurrence of a particular outcome in a given trial.

Random signals are information bearing signals

Random signals are most often the information bearing signals we are used to - voice signals, television signals, digital data (computer files), etc. Electrical "noise" is also a random signal.

Energy and Power Signals

For electrical engineers, the signals we wish to describe will be predominantly voltage or current. Accordingly, the *instantaneous* power developed by these signals will be either:

$$p(t)=\frac{v^2(t)}{R} \tag{1A.6}$$

or:

$$p(t)=Ri^2(t) \tag{1A.7}$$

In signal analysis it is customary to normalise these equations by setting $R=1$. With a signal $g(t)$ the formula for *instantaneous* power in both cases becomes:

$$\boxed{p(t)=g^2(t)} \tag{1A.8}$$

The dissipated energy, or the *total energy* of the signal is then given by:

The total energy of a signal

$$\boxed{E=\int_{-\infty}^{\infty} g^2(t)\,dt} \tag{1A.9}$$

A signal is classified as an *energy signal* if and only if the total energy of the signal is finite:

is finite for an energy signal

$$0<E<\infty \tag{1A.10}$$

The *average power* of a signal is correspondingly defined as:

The average power of a signal

$$P = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} g^2(t)\,dt \tag{1A.11}$$

If the signal is periodic with period T_0, we then have the special case:

and the average power of a periodic signal

$$P = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} g^2(t)\,dt \tag{1A.12}$$

Example

A sinusoid is a periodic signal and therefore has a finite power. If a sinusoid is given by $g(t) = A\cos(2\pi t/T_0 + \phi)$, then what is the average power?

The easiest way to find the average power is by performing the integration in Eq. (1A.12) graphically. For the arbitrary sinusoid given, we can graph the integrand $g^2(t) = A^2\cos^2(2\pi t/T_0 + \phi)$ as:

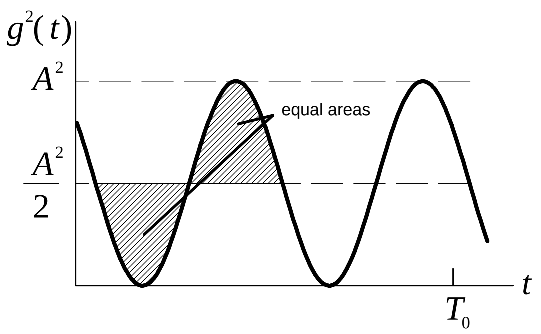


Figure 1A.7

Note that in drawing the graph we don't really need to know the identity $\cos^2(\theta) = (1+\cos 2\theta)/2$ – all we need to know is that if we start off with a

function uniformly oscillating between A and $-A$, then after squaring it oscillates uniformly between A^2 and 0. We can also see that the average value of the resulting waveform is $A^2/2$, because there are equal areas above and below this level. Therefore, if we integrate (i.e. find the area beneath the curve) over an interval spanning T_0, we must have an area equal to the average value times the span, i.e. $A^2T_0/2$. (This is the Mean Value Theorem for Integrals). So the average power is this area divided by the period:

The power of any arbitrary sinusoid of amplitude *A*

$$P = \frac{A^2}{2} \tag{1A.13}$$

This is a surprising result! The power of *any* sinusoid, no matter what its frequency or phase, is dependent only on its amplitude.

Confirm the above result by performing the integration algebraically.

A signal is classified as a *power signal* if and only if the average power of the signal is finite and non-zero:

are finite and non-zero for a power signal

$$0 < P < \infty \tag{1A.14}$$

We observe that if *E*, the energy of $g(t)$, is finite, then its power *P* is zero, and if *P* is finite, then *E* is infinite. It is obvious that a signal may be classified as one or the other but not both. On the other hand, there are some signals, for example:

$$g(t) = e^{-at} \tag{1A.15}$$

that cannot be classified as either energy or power signals, because both *E* and *P* are infinite.

It is interesting to note that periodic signals and most random signals are power signals, and signals that are both deterministic and aperiodic are energy signals.

Common Continuous-Time Signals

Superposition is the key to building complexity out of simple parts

Earlier we stated that the key to handling complexity is to reduce to many simple parts. The converse is also true - we can apply superposition to build complexity out of simple parts. It may come as a pleasant surprise that the study of only a few signals will enable us to handle almost any amount of complexity in a deterministic signal.

The Continuous-Time Step Function

The continuous-time step function defined

We define the continuous-time step function to be:

$$u(t)=\begin{cases}0, & t<0\\ 1/2, & t=0\\ 1, & t>0\end{cases} \tag{1A.16}$$

and graphed

Graphically:

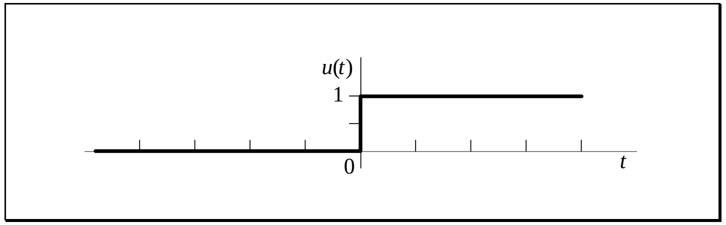


Figure 1A.8

The argument of a function determines its position

We will now make a very important observation: it is the *argument* of the function which determines the *position* of the function along the *t*-axis. Now consider the delayed step function:

$$u(t-t_0)=\begin{cases}0, & t<t_0\\ 1/2, & t=t_0\\ 1, & t>t_0\end{cases} \tag{1A.17}$$

We obtain the conditions on the values of the function by the simple substitution $t \to t-t_0$ in Eq. (1A.16).

Graphically, we have:

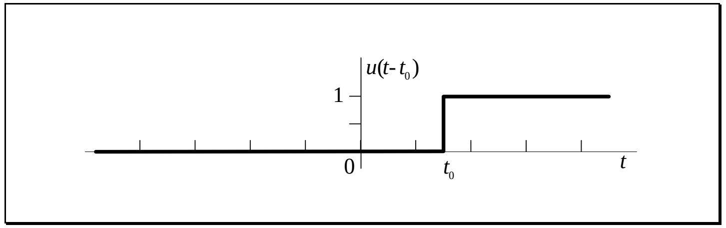


Figure 1A.9

We see that the argument $(t-t_0)$ simply shifts the origin of the original function to t_0. A positive value of t_0 shifts the function to the right - corresponding to a *delay* in time. A negative value shifts the function to the left - an *advance* in time.

as well as "width" and orientation

We will now introduce another concept associated with the argument of a function: if we divide it by a real constant - a *scaling* factor - then we regulate the orientation of the function about the point t_0, and usually change the "width" of the function. Consider the scaled and shifted step function:

$$u\left(\frac{t-t_0}{T}\right)=\begin{cases}0, & \dfrac{t}{T}<\dfrac{t_0}{T}\\ 1/2, & \dfrac{t}{T}=\dfrac{t_0}{T}\\ 1, & \dfrac{t}{T}>\dfrac{t_0}{T}\end{cases} \tag{1A.18}$$

In this case it is not meaningful to talk about the width of the step function, and the only purpose of the constant T is to allow the function to be reflected about the line $t=t_0$, as shown below:

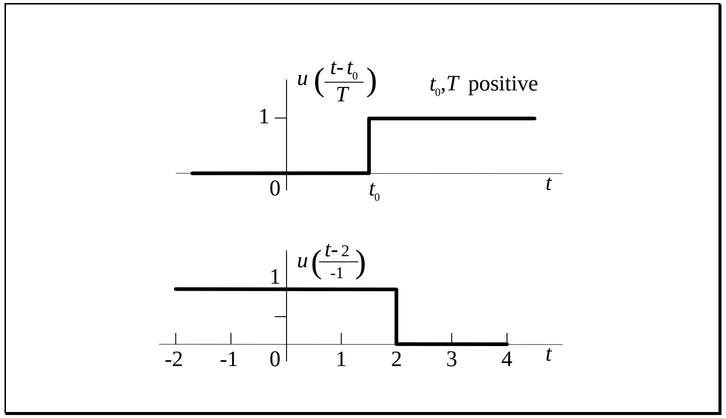


Figure 1A.10

Use Eq. (1A.18) to verify the bottom step function in Figure 1A.10.

The step function as a "switch"

The utility of the step function is that it can be used as a "switch" to turn another function on or off at some point. For example, the product given by $u(t-1)\cos 2\pi t$ is as shown below:

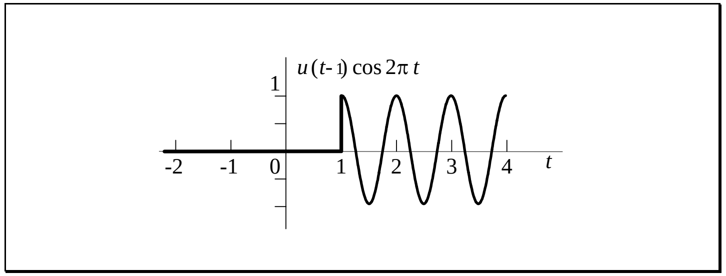


Figure 1A.11

The Rectangle Function

One of the most useful functions in signal analysis is the rectangle function, defined as:

The rectangle function defined

$$\operatorname{rect}(t)=\begin{cases}0, & |t|>\frac{1}{2}\\ 1/2, & |t|=\frac{1}{2}\\ 1, & |t|<\frac{1}{2}\end{cases} \tag{1A.19}$$

Graphically, this is a "rectangle" with a height, width and area of one:

and graphed

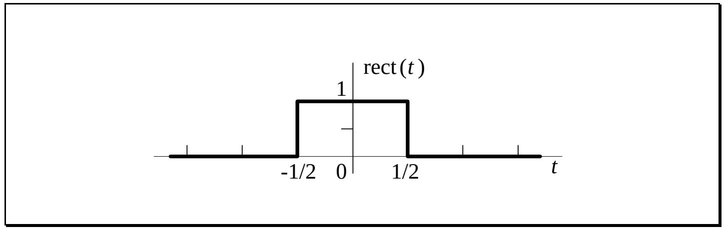


Figure 1A.12

If we generalise the argument, as we did for the step function, we get:

$$\operatorname{rect}\left(\frac{t-t_0}{T}\right)=\begin{cases}0, & \left|\frac{t-t_0}{T}\right|>\frac{1}{2}\\ 1/2, & \left|\frac{t-t_0}{T}\right|=\frac{1}{2}\\ 1, & \left|\frac{t-t_0}{T}\right|<\frac{1}{2}\end{cases} \tag{1A.20}$$

Graphically, it is a rectangle with a height of unity, it is centred on $t=t_0$, and both its width and area are equal to $|T|$:

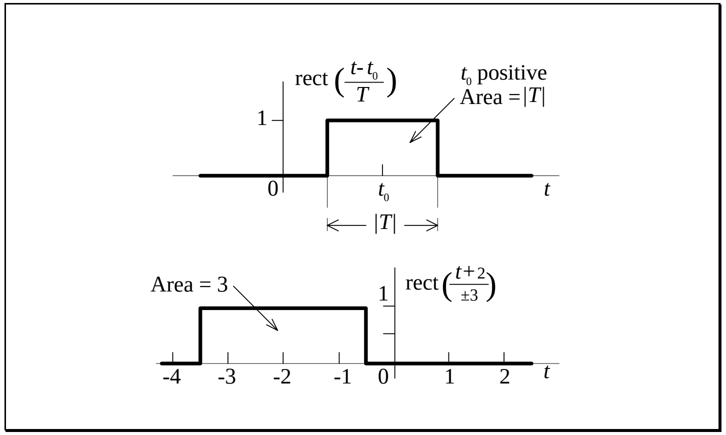


Figure 1A.13

The rectangle function can be used to turn another function on and off

In the time-domain the rectangle function can be used to represent a gating operation in an electrical circuit. Mathematically it provides an easy way to turn a function on and off.

Notice from the symmetry of the function that the sign of T has no effect on the function's orientation. However, the magnitude of T still acts as a scaling factor.

The Straight Line

It is surprising how something so simple can cause a lot of confusion. We start with one of the simplest straight lines possible. Let $g(t)=t$:

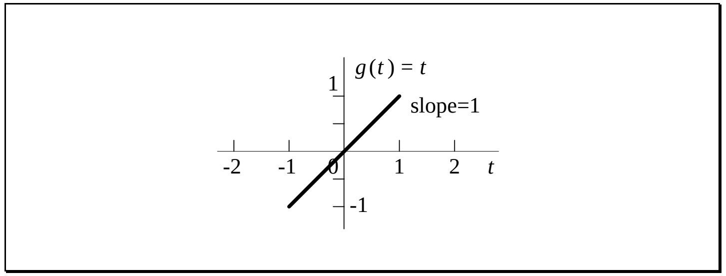


Figure 1A.14

Now shift the straight line along the t-axis in the standard fashion, to make $g(t)=t-t_0$:

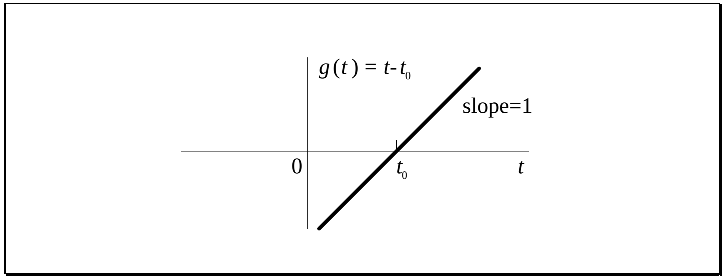


Figure 1A.15

To change the slope of the line, simply apply the usual scaling factor, to make:

The straight line defined

$$g(t)=\frac{t-t_0}{T} \tag{1A.21}$$

This is the equation of a straight line, with slope $1/T$ and t-axis intercept t_0:

and graphed

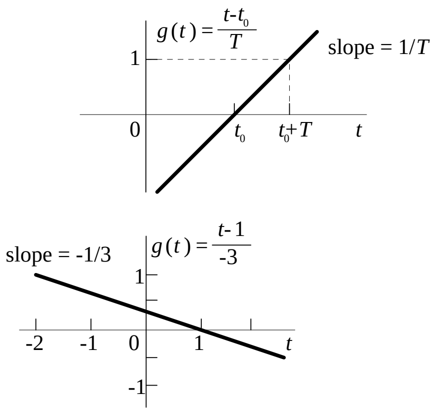


Figure 1A.16

We can now use our knowledge of the straight line and rectangle function to completely specify piece-wise linear signals.

Example

A function generator produces the following sawtooth waveform:

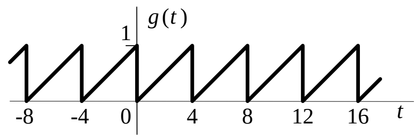


Figure 1A.17

We seek a mathematical description of this waveform. We start by recognising the fact that the waveform is periodic, with period $T_0 = 4$. First we describe only one period (say the one beginning at the origin). We recognise that the ramp part of the sawtooth is a straight line multiplied by a rectangle function:

$$g_0(t) = \frac{t}{4}\operatorname{rect}\left(\frac{t-2}{4}\right) \tag{1A.22}$$

Graphically, it is:

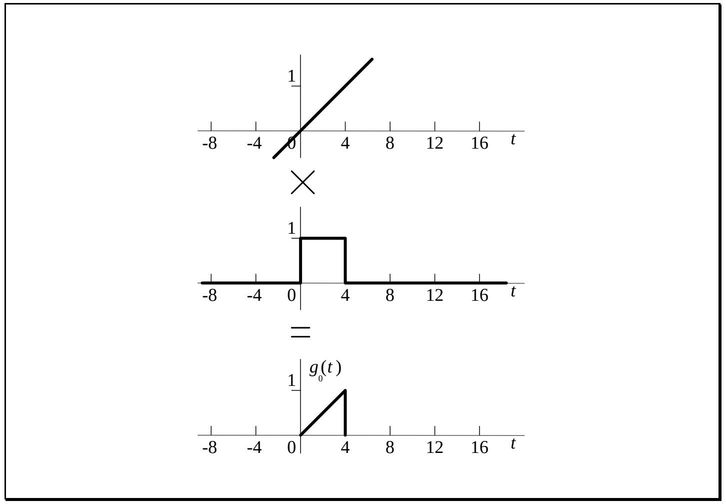


Figure 1A.18

The next period is just $g_0(t)$ shifted right (delayed in time) by 4. The next period is therefore described by:

$$g_1(t)=g_0(t-4)=\frac{t-4}{4}\mathrm{rect}\left(\frac{t-4-2}{4}\right) \tag{1A.23}$$

It is now easy to see the pattern. In general we have:

$$g_n(t)=g_0(t-4n)=\frac{t-4n}{4}\mathrm{rect}\left(\frac{t-4n-2}{4}\right) \tag{1A.24}$$

where $g_n(t)$ is the nth period and n is any integer. Now all we have to do is add up all the periods to get the complete mathematical expression for the sawtooth:

$$g(t)=\sum_{n=-\infty}^{\infty}\frac{t-4n}{4}\mathrm{rect}\left(\frac{t-4n-2}{4}\right) \tag{1A.25}$$

Example

Sketch the following waveform:

$$g(t)=(t-1)\mathrm{rect}(t-1.5)+\mathrm{rect}(t-2.5)+(-0.5t+2.5)\mathrm{rect}(0.5t-2) \tag{1A.26}$$

We can start by putting arguments into our "standard" form:

$$g(t)=(t-1)\mathrm{rect}(t-1.5)+\mathrm{rect}(t-2.5)+\left(\frac{t-5}{-2}\right)\mathrm{rect}\left(\frac{t-4}{2}\right) \tag{1A.27}$$

From this, we can compose the waveform out of the three specified parts:

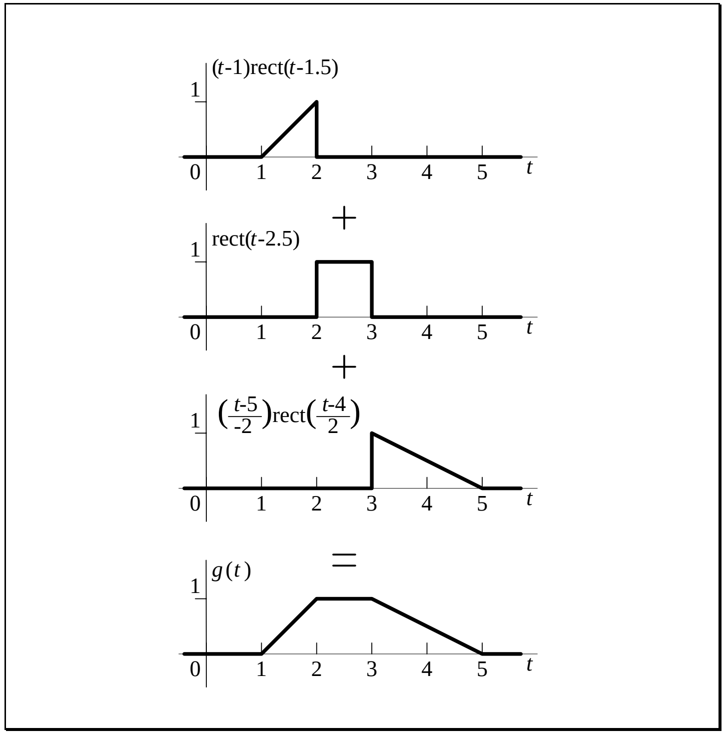


Figure 1A.19

The Sinc Function

The sinc function will show up quite often in studies of linear systems, particularly in signal spectra, and it is interesting to note that there is a close relationship between this function and the rectangle function. Its definition is:

The sinc function defined

$$\boxed{\operatorname{sinc}(t) = \frac{\sin \pi t}{\pi t}} \tag{1A.28}$$

Graphically, it looks like:

and graphed

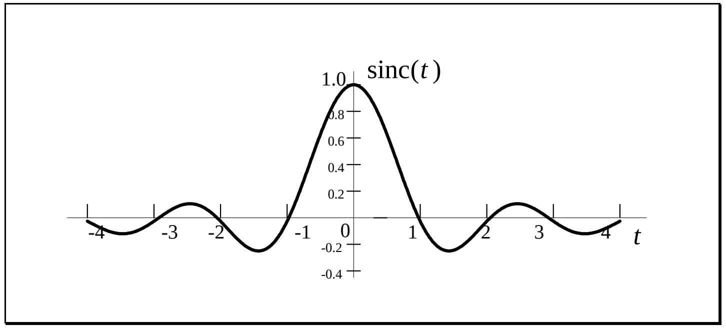


Figure 1A.20

Features of the sinc function

The inclusion of π in the formula for the sinc function gives it certain "nice" properties. For example, the zeros of $\operatorname{sinc}(t)$ occur at non-zero integral values, its width between the first two zeros is two, and its area (including positive and negative lobes) is just one. Notice also that it appears that the sinc function is undefined for $t = 0$. In one sense it is, but in another sense we can define a function's value at a singularity by approaching it from either side and averaging the limits.

This is not unusual – we did it explicitly in the definition of the step function, where there is obviously a singularity at the "step". We overcame this by calculating the limits of the function approaching zero from the positive and negative sides. The limit is 0 (approaching from the negative side) and 1

(approaching from the positive side), and the average of these two is $1/2$. We then made explicit the use of this value for a zero argument.

The limit of the sinc function as $t \to 0$ can be obtained using l'Hôpital's rule:

$$\lim_{t \to 0} \frac{\sin \pi t}{\pi t} = \lim_{t \to 0} \frac{\pi \cos \pi t}{\pi} = 1 \tag{1A.29}$$

Therefore, we say the sinc function has a value of 1 when its argument is zero.

With a generalised argument, the sinc function becomes:

$$\boxed{\operatorname{sinc}\left(\frac{t - t_0}{T}\right) = \frac{\sin \pi \left(\frac{t - t_0}{T}\right)}{\pi \left(\frac{t - t_0}{T}\right)}} \tag{1A.30}$$

Its zeros occur at $t_0 \pm nT$, its height is 1, its width between the first two zeros is $2|T|$ and its area is $|T|$:

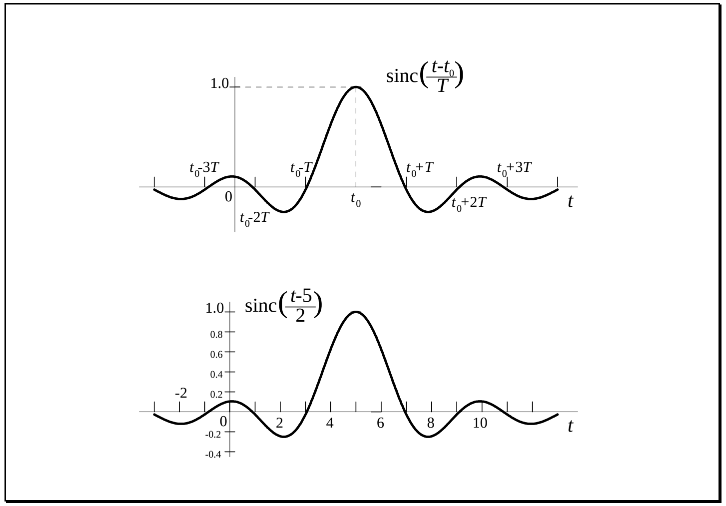


Figure 1A.21

The Impulse Function

The need for an impulse function

The impulse function, or Dirac delta function, is of great importance in the study of signal theory. It will enable us to represent "densities" at a single point. We shall employ the widely accepted symbol $\delta(t)$ to denote the impulse function.

An informal definition of the impulse function

The impulse function is often described as having an infinite height and zero width such that its area is equal to unity, but such a description is not particularly satisfying.

A more formal definition is obtained by first forming a sequence of functions, such as $1/|T|\operatorname{rect}(t/T)$, and then defining $\delta(t)$ to be:

A more formal definition of the impulse function

$$\delta(t)=\lim_{T\to 0}\frac{1}{|T|}\operatorname{rect}\left(\frac{t}{T}\right) \tag{1A.31}$$

As $|T|$ gets smaller and smaller the members of the sequence become taller and narrower, but their area remains constant as shown below:

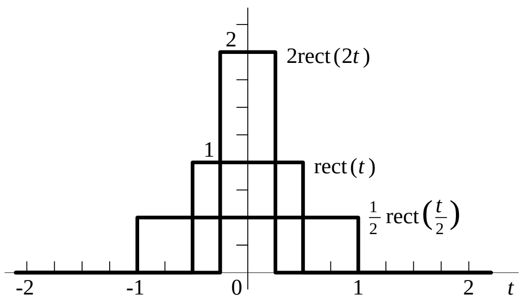


Figure 1A.22

This definition, however, is not very satisfying either.

From a physical point of view, we can consider the delta function to be so narrow that making it any narrower would in no way influence the results in which we are interested. As an example, consider the simple *RC* circuit shown below, in which a rectangle function is applied as $v_i(t)$:

An impulse function in the lab is just a very narrow pulse

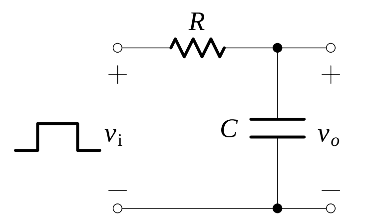


Figure 1A.23

We choose the width of the pulse to be *T*, and its height to be equal to $1/T$ such that its area is 1 as *T* is varied. The output voltage $v_o(t)$ will vary with time, and its exact form will depend on the relative values of *T* and the product *RC*.

If *T* is much larger than *RC*, as in the top diagram of Figure 1A.21, the capacitor will be almost completely charged to the voltage $1/T$ before the pulse ends, at which time it will begin to discharge back to zero.

As the duration of a rectangular input pulse get smaller

If we shorten the pulse so that $T \ll RC$, the capacitor will not have a chance to become fully charged before the pulse ends. Thus, the output voltage behaves as in the middle diagram of Figure 1A.21, and it can be seen that there is a considerable difference between this output and the proceeding one.

and smaller

If we now make *T* still shorter, as in the bottom diagram of Figure 1A.21, we note very little change in the shape of the output. In fact, as we continue to make *T* shorter and shorter, the only noticeable change is in the time it takes the output to reach a maximum, and this time is just equal to *T*.

and smaller, the output of a linear system approaches the "impulse response"

If this interval is too short to be resolved by our measuring device, the input is effectively behaving as a delta function and decreasing its duration further will have no observable effect on the output, which now closely resembles the impulse response of the circuit.

Graphical derivation of the impulse response of an *RC* circuit by decreasing a pulse's width while maintaining its area

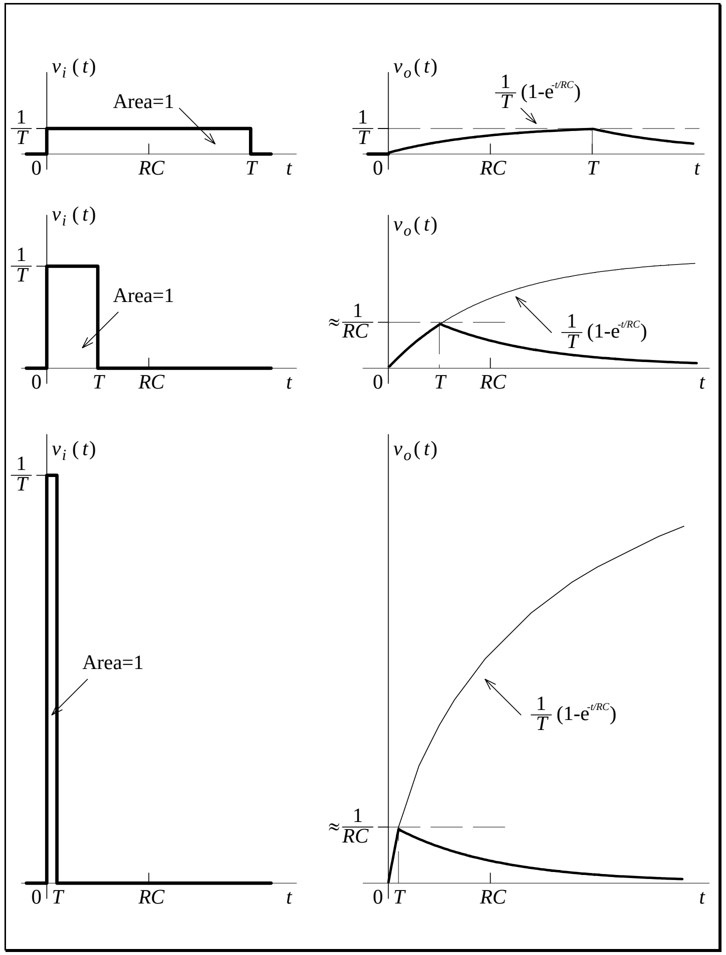


Figure 1A.24

With the previous discussion in mind, we shall use the following *properties* as the definition of the delta function: given the real constant t_0 and the arbitrary complex-valued function $f(t)$, which is continuous at the point $t=t_0$, then

The impulse function defined - as behaviour upon integration

$$\delta(t-t_0)=0, \qquad t \neq t_0 \tag{1A.32a}$$

$$\int_{t_1}^{t_2} f(t)\delta(t-t_0)dt = f(t_0) \quad t_1 < t_0 < t_2 \tag{1A.32b}$$

If $f(t)$ is discontinuous at the point $t=t_0$, Eq. (1A.32b) may still be used but the value of $f(t_0)$ is taken to be the average of the limiting values as t_0 is approached from above and from below. With this definition of the delta function, we do not specify the value of $\delta(t-t_0)$ at the point $t=t_0$. Note that we do not care about the exact form of the delta function itself but only about its behaviour under integration.

"Sifting" property of the impulse function defined

Eq. (1A.32b) is often called the *sifting* property of the delta function, because it "sifts" out a single value of the function $f(t)$.

Graphically we will represent $\delta(t-t_0)$ as a spike of unit height located at the point $t=t_0$, but observe that the *height of the spike* corresponds to the *area of the delta function*. Such a representation is shown below:

Graphical representation of an impulse function

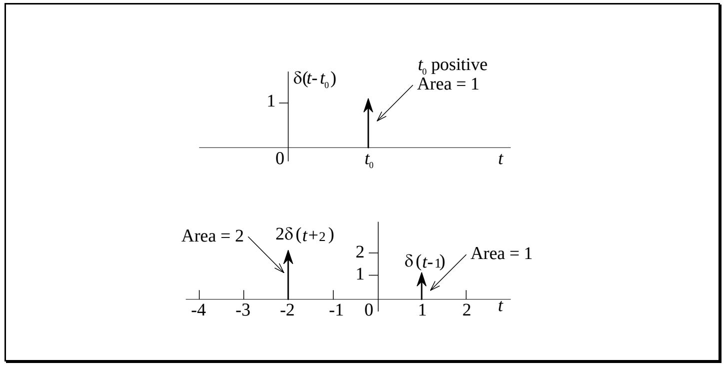


Figure 1A.25

Sinusoids

Sinusoids can be described in the frequency-domain

Here we start deviating from the previous descriptions of time-domain signals. All the signals described so far are aperiodic. With periodic signals (such as sinusoids), we can start to introduce the concept of "frequency content", i.e. a shift from describing signals in the time-domain to one that describes them in the frequency-domain. This "new" way of describing signals will be fully exploited later when we look at Fourier series and Fourier transforms.

Why Sinusoids?

The sinusoid with which we are so familiar today appeared first as the solution to differential equations that 17th century mathematicians had derived to describe the behaviour of various vibrating systems. They were no doubt surprised that the functions that had been used for centuries in trigonometry appeared in this new context. Mathematicians who made important contributions included Huygens (1596-1687), Bernoulli (1710-1790), Taylor (1685-1731), Ricatti (1676-1754), Euler (1707-1783), D'Alembert (1717-1783) and Lagrange (1736-1813).

Perhaps the greatest contribution was that of Euler (who invented the symbols π, e and $i=\sqrt{-1}$ which we call j). He first identified the fact that is highly significant for us:

The special relationship enjoyed by sinusoids and linear systems

For systems described by linear differential equations a sinusoidal input yields a sinusoidal output. (1A.33)

The output sinusoid has the same frequency as the input, it is however altered in *amplitude* and *phase*.

We are so familiar with this fact that we sometimes overlook its significance. *Only* sinusoids have this property with respect to linear systems. For example, applying a square wave input does not produce a square wave output.

It was Euler who recognised that an arbitrary sinusoid could be resolved into a pair of complex exponentials:

A sinusoid can be expressed as the sum of a forward and backward rotating phasor

$$\boxed{A\cos(\omega t+\phi)=Xe^{j\omega t}+X^{*}e^{-j\omega t}} \tag{1A.34}$$

where $X=A/2e^{j\phi}$. Euler found that the input/output form invariance exhibited by the sinusoid could be mathematically expressed most concisely using the exponential components. Thus if $A/2e^{j(\omega t+\phi)}$ was the input, then for a linear system the output would be $H.A/2e^{j(\omega t+\phi)}$ where H is a complex-valued function of ω.

We know this already from circuit analysis, and the whole topic of system design concerns the manipulation of *H*.

The second major reason for the importance of the sinusoid can be attributed to Fourier, who in 1807 recognised that:

Periodic signals are made up of sinusoids - Fourier Series

Any periodic function can be represented as the weighted sum of a family of sinusoids. (1A.35)

The way now lay open for the analysis of the behaviour of a linear system for any periodic input by determining the response to the individual sinusoidal (or exponential) components and adding them up (superposition).

This technique is called *frequency analysis*. Its first application was probably that of Newton passing white light through a prism and discovering that red light was unchanged after passage through a second prism.

Representation of Sinusoids

Sinusoids can be resolved into pairs of complex exponentials as in Eq. (1A.34). In this section, instead of following the historical, mathematical approach, we relate sinusoids to complex numbers using a utilitarian argument. Consider the sinusoid:

$$x(t) = A\cos\left(\frac{2\pi t}{T}\right) \tag{1A.36}$$

A is the *amplitude*, $2\pi/T$ is the *angular frequency*, and T is the *period*. Graphically, we have:

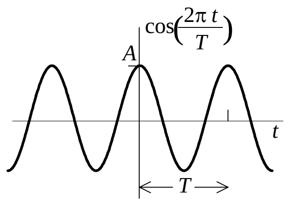


Figure 1A.26

Delaying the waveform by t_0 results in:

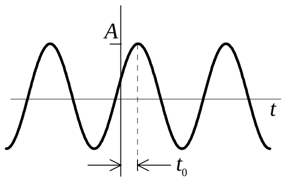


Figure 1A.27

Mathematically, we have:

$$\begin{aligned} x(t) &= A\cos\left(\frac{2\pi(t-t_0)}{T}\right) = A\cos\left(\frac{2\pi t}{T} - \frac{2\pi t_0}{T}\right) \\ &= A\cos\left(\frac{2\pi t}{T} + \phi\right) \end{aligned} \tag{1A.37}$$

A sinusoid expressed in the most general fashion

where $\phi = -2\pi t_0/T$. ϕ is the *phase* of the sinusoid. Note the negative sign in the definition of phase, which means a delayed sinusoid has negative phase.

We note that when $t_0 = T/4$ we get:

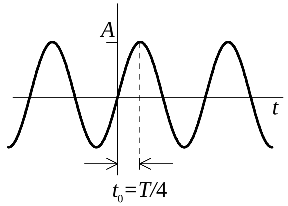


Figure 1A.28

and:

$$\begin{aligned} x(t) &= A\cos\left(\frac{2\pi t}{T} - \frac{\pi}{2}\right) \\ &= A\sin\left(\frac{2\pi t}{T}\right) \end{aligned} \tag{1A.38}$$

We can thus represent an arbitrary sinusoid as either:

$$x(t) = A\sin(\omega t + \phi) \text{ or } A\cos(\omega t - \pi/2 + \phi) \tag{1A.39}$$

Similarly:

$$x(t) = A\cos(\omega t + \phi) \text{ or } A\sin(\omega t + \pi/2 + \phi) \tag{1A.40}$$

We shall use the *cosine* form. If a sinusoid is expressed in the sine form, then we need to subtract $90°$ from the phase angle to get the cosine form.

Phase refers to the angle of a cosinusoid at *t*=0

When the *phase* of a sinusoid is referred to, it is the phase angle *in the cosine form* (in these lecture notes).

Resolution of an Arbitrary Sinusoid into Orthogonal Functions

An *arbitrary sinusoid* can be expressed as a weighted sum of a cos() and a -sin() term. Mathematically, we can derive the following:

A sinusoid can be broken down into two orthogonal components

$$\begin{aligned} A\cos(\omega t + \phi) &= A\cos\phi\cos\omega t - A\sin\phi\sin\omega t \\ &= [A\cos\phi]\cos\omega t + [A\sin\phi](-\sin\omega t) \end{aligned} \tag{1A.41}$$

Example

Lets resolve a simple sinusoid into the above components:

$$\begin{aligned} 2\cos(3t - 15°) &= 2\cos(-15°)\cos 3t + 2\sin(-15°)(-\sin 3t) \\ &= [1.93]\cos 3t + [-0.52](-\sin 3t) \end{aligned} \tag{1A.42}$$

Graphically, this is:

Any sinusoid can be decomposed into two orthogonal sinusoids of the same frequency, but different amplitude

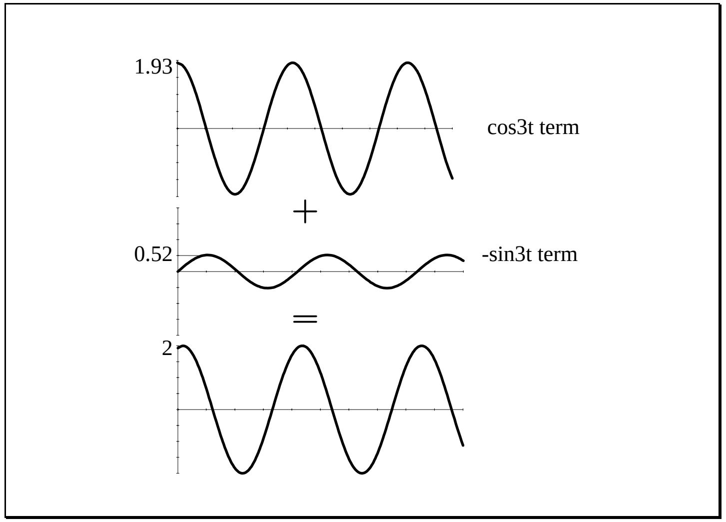


Figure 1A.29

The cos() and -sin() terms are said to be *orthogonal*. For our purposes orthogonal may be understood as the property that two vectors or functions mutually have when one cannot be expressed as a sum containing the other. We will look at basis functions and orthogonality more formally later on.

A quick definition of orthogonal

Representation of a Sinusoid by a Complex Number

There are three distinct pieces of information about a sinusoid - frequency, amplitude and phase. Provided the frequency is unambiguous, two real numbers are required to describe the sinusoid completely. Why not use a complex number to store these two real numbers?

Three real numbers completely specify a sinusoid

Suppose the convention is adopted that: the real part of a complex number is the $\cos\omega t$ amplitude, and the imaginary part is the $-\sin\omega t$ amplitude of the resolution described by Eq. (1A.41).

Two real numbers completely specify a sinusoid of a given frequency

Suppose we call the complex number created in this way the *phasor* associated with the sinusoid.

which we store as a complex number called a phasor

Our notation will be:

Time-domain and frequency-domain symbols

$x(t)$ – time-domain expression

$\overline{X}$ – phasor associated with $x(t)$

The reason for the bar over *X* will become apparent shortly.

Example

With the previous example, we had $x(t)=2\cos(3t-15°)$. Therefore the phasor associated with it, using our new convention, is $\overline{X}=1.93+j(-0.52)=2\angle-15°$.

A phasor can be "read off" a time-domain expression

We can see that the *magnitude* of the complex number is the *amplitude* of the sinusoid and the *angle* of the complex number is the *phase* of the sinusoid. In general, the correspondence between a sinusoid and its phasor is:

The phasor that corresponds to an arbitrary sinusoid

$$\boxed{x(t)=A\cos(\omega t+\phi) \quad \Leftrightarrow \quad \overline{X}=Ae^{j\phi}} \tag{1A.43}$$

Example

If $x(t)=3\sin(\omega t-30°)$ then we have to convert to our cos notation: $x(t)=3\cos(\omega t-120°)$. Therefore $\overline{X}=3\angle-120°$.

Note carefully that $\overline{X}\neq 3\cos(\omega t-120°)$. All we can say is that $x(t)=3\cos(\omega t-120°)$ is *represented* by $\overline{X}=3\angle-120°$.

Phasors make manipulating sinusoids of the same frequency easy

The convenience of complex numbers extends beyond their compact representation of the amplitude and phase. The sum of two phasors corresponds to the sinusoid which is the sum of the two component sinusoids represented by the phasors. That is, if $x_3(t)=x_1(t)+x_2(t)$ where $x_1(t)$, $x_2(t)$ and $x_3(t)$ are sinusoids with the same frequency, then $\overline{X}_3=\overline{X}_1+\overline{X}_2$.

Example

If $x_3(t)=\cos\omega t-2\sin\omega t$ then $\overline{X}_3=1\angle 0°-2\angle-90°=1+j2=2.24\angle 63°$ which corresponds to $x_3(t)=2.24\cos(\omega t+63°)$.

Formalisation of the Relationship between Phasor and Sinusoid

Using Euler's expansion:

$$e^{j\theta}=\cos\theta+j\sin\theta \tag{1A.44}$$

we have:

$$Ae^{j\phi}e^{j\omega t}=Ae^{j(\omega t+\phi)}=A\cos(\omega t+\phi)+jA\sin(\omega t+\phi) \tag{1A.45}$$

We can see that the sinusoid $A\cos(\omega t+\phi)$ represented by the phasor $\overline{X}=Ae^{j\phi}$ is equal to $\mathrm{Re}\{\overline{X}e^{j\omega t}\}$. Therefore:

A phasor / time-domain relationship

$$\boxed{x(t)=\mathrm{Re}\{\overline{X}e^{j\omega t}\}} \tag{1A.46}$$

This can be visualised as:

Graphical interpretation of rotating phasor / time-domain relationship

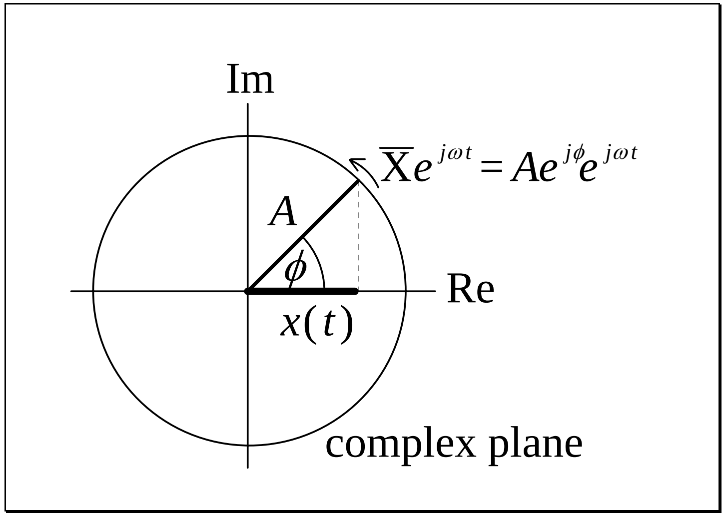


Figure 1A.30

Euler’s Complex Exponential Relationships for Cosine and Sine

Euler’s expansion:

$$e^{j\theta} = \cos\theta + j\sin\theta \tag{1A.47}$$

can be visualized as:

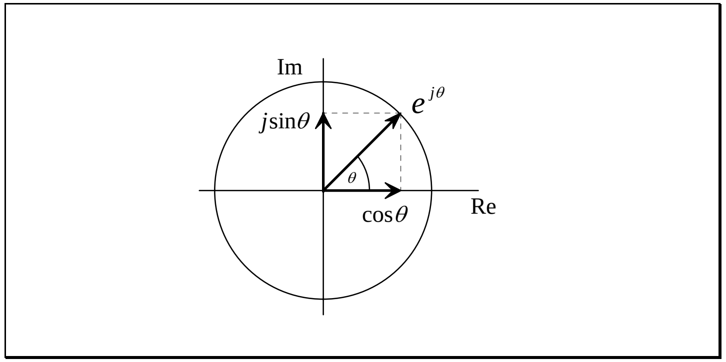


Figure 1A.31

By mirroring the vectors about the real axis, it is obvious that:

$$e^{-j\theta} = \cos\theta - j\sin\theta \tag{1A.48}$$

and:

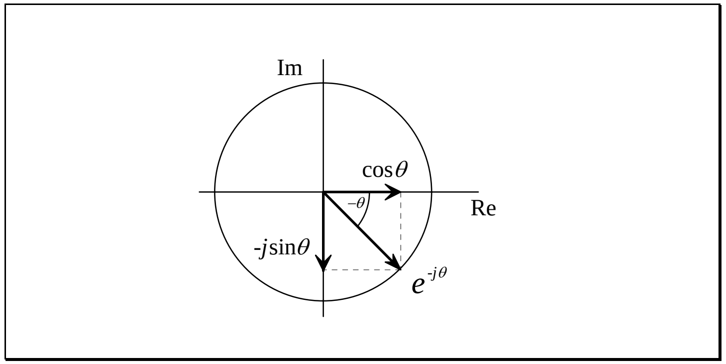


Figure 1A.32

By adding Eqs. (1A.45) and (1A.46), we can write:

Cos represented as a sum of complex exponentials

$$\boxed{\cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2}} \tag{1A.49}$$

and:

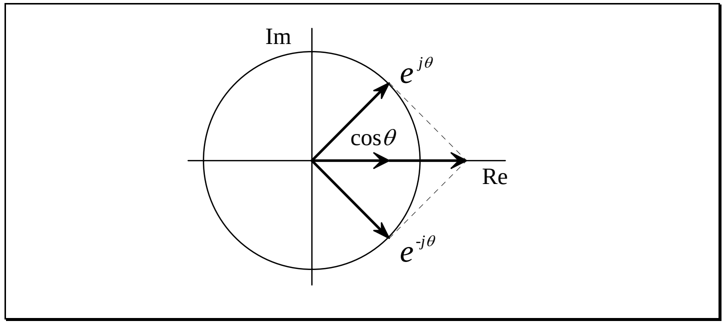


Figure 1A.33

By subtracting Eqs. (1A.45) and (1A.46), we can write:

Sin represented as a sum of complex exponentials

$$\boxed{\sin\theta = \frac{e^{j\theta} - e^{-j\theta}}{j2}} \tag{1A.50}$$

and:

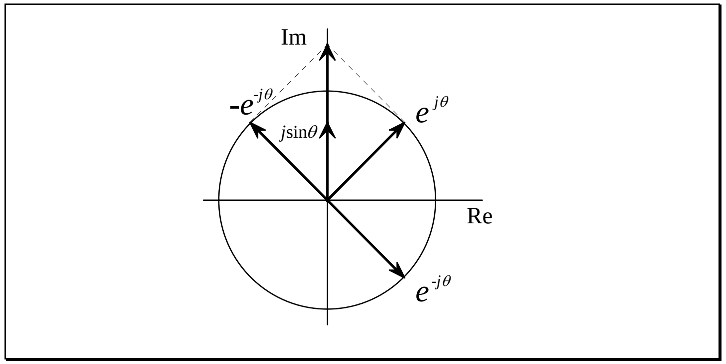


Figure 1A.34

A New Definition of the Phasor

To avoid the mathematical "clumsiness" of needing to take the real part, another definition of the phasor is often adopted. In circuit analysis we use the definition of the phasor as given before. In communication engineering we will find it more convenient to use a new definition:

A new phasor definition

$$\boxed{X = \frac{A}{2}e^{j\phi}, \quad \text{or } X = \frac{\overline{X}}{2}} \tag{1A.51}$$

We realise that this phasor definition, although a unique and *sufficient* representation for every sinusoid, is just half of the sinusoid's full representation. Using Euler's complex exponential expansion of cos, we get:

A better phasor / time-domain relationship

$$\begin{aligned} A\cos(\omega t + \phi) &= \frac{A}{2}e^{j\phi}e^{j\omega t} + \frac{A}{2}e^{-j\phi}e^{-j\omega t} \\ &= Xe^{j\omega t} + X^* e^{-j\omega t} \end{aligned} \tag{1A.52}$$

The two terms in the summation represent two *counter-rotating* phasors with angular velocities ω and $-\omega$ in the complex plane, as shown below:

Graphical interpretation of counter-rotating phasors / time-domain relationship

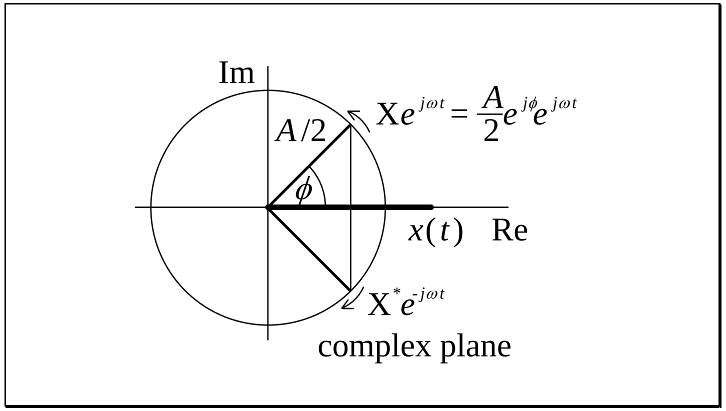


Figure 1A.35

Graphical Illustration of the Relationship between the Two Types of Phasor and their Corresponding Sinusoid

Consider the first representation of a sinusoid: $x(t) = \mathrm{Re}\{\overline{X}e^{j\omega t}\}$. Graphically, $x(t)$ can be "generated" by taking the projection of the rotating phasor formed by multiplying $\overline{X}$ by $e^{j\omega t}$, onto the real axis:

A sinusoid can be generated by taking the real part of a rotating complex number

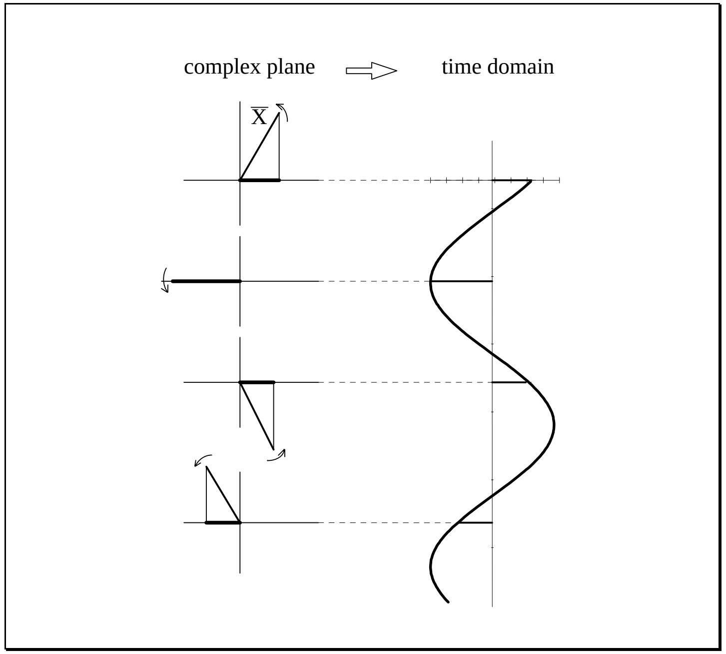


Figure 1A.36

Now consider the second representation of a sinusoid: $x(t) = Xe^{j\omega t} + X^* e^{-j\omega t}$. Graphically, $x(t)$ can be "generated" by simply adding the two complex conjugate counter-rotating phasors $Xe^{j\omega t}$ and $X^* e^{-j\omega t}$. The result will always be a real number:

A sinusoid can be generated by adding up a forward rotating complex number and its backward rotating complex conjugate

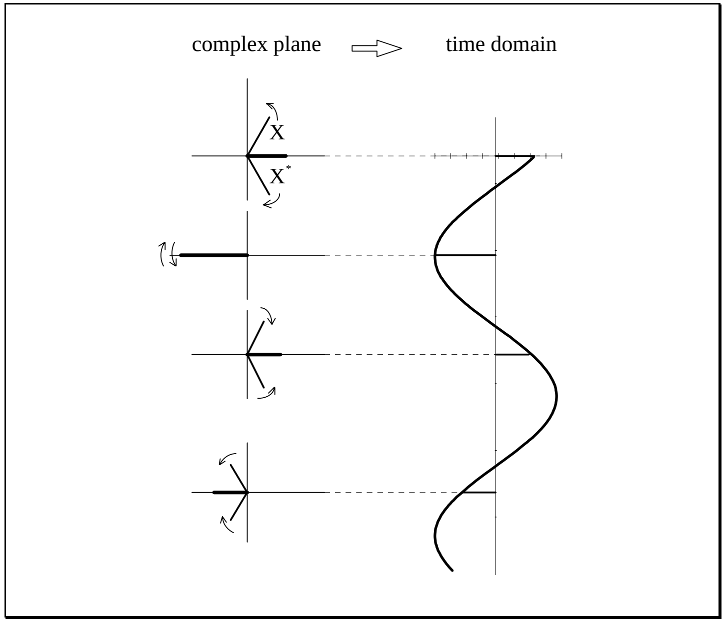


Figure 1A.37

Negative Frequency

Negative frequency just means the phasor is going clockwise

The phasor $X^* e^{-j\omega t}$ rotates with a speed ω but in the clockwise direction. Therefore, we consider this counter-rotating phasor to have a *negative frequency*. The concept of negative frequency will be very useful when we start to manipulate signals in the frequency-domain.

The link between maths and graphs should be well understood

You should become very familiar with all of these signal types, and you should feel comfortable representing a sinusoid as a complex exponential. Being able to manipulate signals mathematically, while at the same time imagining what is happening to them graphically, is the key to readily understanding signals and systems.

Common Discrete-Time Signals

A lot of discrete-time signals are obtained by "sampling" continuous-time signals at regular intervals. In these cases, we can simply form a discrete-time signal by substituting $t = nT_s$ into the mathematical expression for the continuous-time signal, and then rewriting it using discrete notation.

Superposition is the key to building complexity out of simple parts

Example

Consider the sinusoidal function $g(t) = \cos(2\pi f_0 t)$. If we "sample" this signal at discrete times, we get:

$$g(t)\big|_{t=nT_s} = g(nT_s) = \cos(2\pi f_0 nT_s) \tag{1A.53}$$

Since this is valid only at times which are multiples of T_s, it is a discrete-time signal and can be written as such:

$$g[n] = \cos(2\pi f_0 nT_s) = \cos(\Omega n) \tag{1A.54}$$

where $\Omega = 2\pi f_0 T_s$ is obviously a constant. A graph of this discrete-time signal is given below:

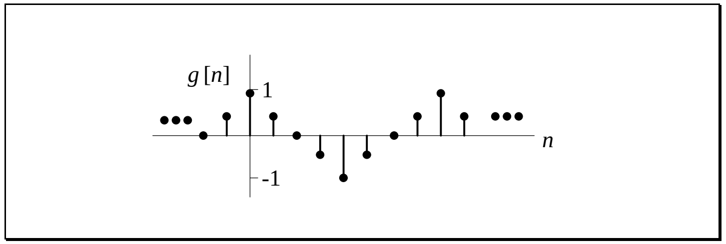


Figure 1A.38

This "sampling" process only works when the continuous-time signal is smooth and of finite value. Therefore, the discrete-time versions of the rectangle and impulse are defined from first principles.

The Discrete-Time Step Function

We define the discrete-time step function to be:

The discrete-time step function defined

$$u[n] = \begin{cases} 0, & n < 0 \\ 1, & n \geq 0 \end{cases} \tag{1A.55}$$

Graphically:

and graphed

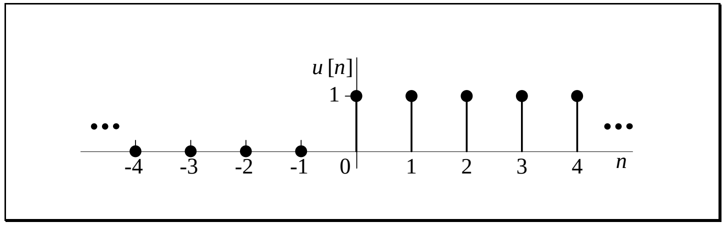
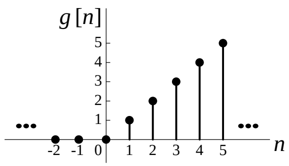


Figure 1A.39

Note that this is *not* a sampled version of the continuous-time step function, which has a discontinuity at $t = 0$. We *define* the discrete-time step to have a value of 1 at $n = 0$ (instead of having a value of $1/2$ if it were obtained by sampling the continuous-time signal).

Example

A discrete-time signal has the following graph:



We recognise that the signal is increasing linearly after it "turns on" at $n = 0$. Therefore, an expression for this signal is $g[n] = nu[n]$.

The Unit-Pulse Function

There is no way to sample an impulse, since its value is undefined. However, we shall see that the discrete-time *unit-pulse*, or Kronecker delta function, plays the same role in discrete-time systems as the impulse does in continuous-time systems. It is defined as:

The unit-pulse defined

$$\delta[n] = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases} \tag{1A.56}$$

Graphically:

and graphed

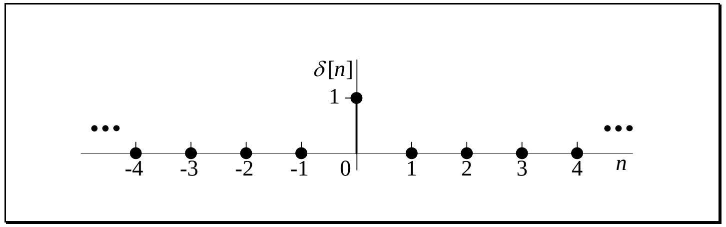
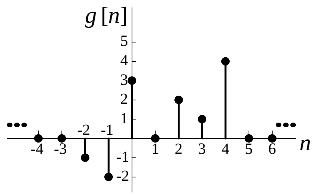


Figure 1A.40

Example

An arbitrary-looking discrete-time signal has the following graph:



With no obvious formula, we can express the signal as the sum of a series of delayed and weighted unit-pulses. Working from left to right, we get: $g[n] = -\delta[n+2] - 2\delta[n+1] + 3\delta[n] + 2\delta[n-2] + \delta[n-3] + 4\delta[n-4]$.

Summary

- Signals can be characterized with many attributes – continuous or discrete, periodic or aperiodic, energy or power, deterministic or random. Each characterization tells us something useful about the signal.

- Signals written mathematically in a standard form will benefit our analysis in the future. We use arguments of the form $\left(\frac{t-t_0}{T}\right)$.

- Sinusoids are special signals. They are the *only* real signals that retain their shape when passing through a linear time-invariant system. They are often represented as the sum of two complex conjugate, counter-rotating phasors.

- The phasor corresponding to $x(t) = A\cos(\omega t + \phi)$ is $X = \frac{A}{2}e^{j\phi}$.

- Most discrete-time signals can be obtained by "sampling" continuous-time signals. The discrete-time step function and unit-pulse are defined from first principles.

References

Kamen, E.W. & Heck, B.S.: *Fundamentals of Signals and Systems Using MATLAB®*, Prentice-Hall, Inc., 1997.

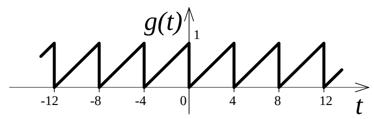
Quiz

Encircle the correct answer, cross out the wrong answers. [one or none correct]

1.
The signal $\cos(10\pi t + \pi/3) - \sin(11\pi t - \pi/2)$ has:

(a) period = 1 s (b) period = 2 s (c) no period

2.



The periodic sawtooth waveform can be represented by:

(a) $g(t) = \sum_{n=-\infty}^{\infty} \frac{(t-4n)}{4} \text{rect}\left(\frac{t-2}{4}\right)$

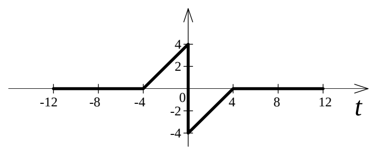
(b) $g(t) = \sum_{n=-\infty}^{\infty} \frac{(t-4n)}{4} \text{rect}\left(\frac{t-2n}{4}\right)$

(c) $g(t) = \sum_{n=-\infty}^{\infty} \frac{(t-4)}{4} \text{rect}\left(\frac{t-2-4n}{4}\right)$

3.
The sinusoid $-5\sin(314t - 80°)$ can be represented by the phasor $\overline{X}$:

(a) $-5\angle -80°$ (b) $5\angle 100°$ (c) $5\angle 10°$

4.



The energy of the signal shown is:

(a) 16 (b) 64 (c) 0

5.
Using forward and backward rotating phasor notation, $X = 3\angle 30°$. If the angular frequency is 1 rads^{-1}, $x(2)$ is:

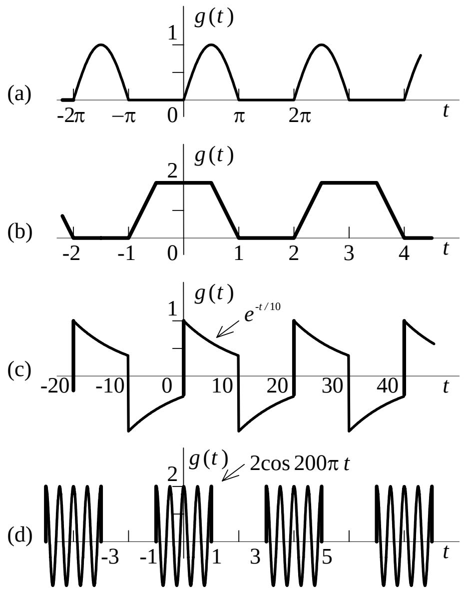
(a) 5.994 (b) 5.088 (c) -4.890

Answers: 1. b 2. x 3. c 4. x 5. c

Exercises

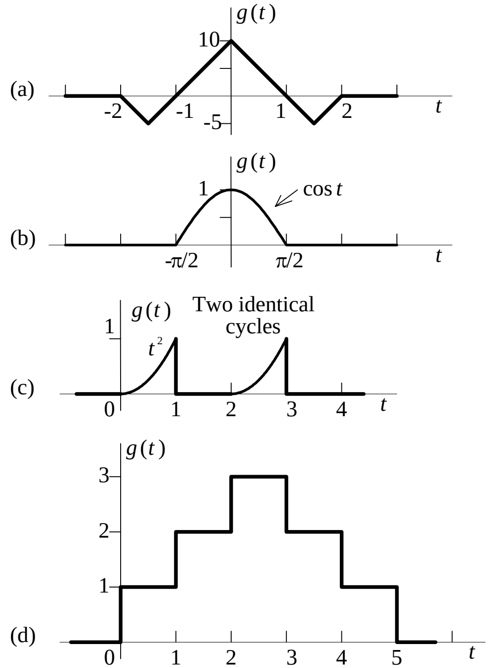
1.

For each of the periodic signals shown, determine their time-domain expressions, their periods and average power:



2.

For the signals shown, determine their time-domain expressions and calculate their energy:



3.

Using the sifting property of the impulse function, evaluate the following integrals:

(i) $\int_{-\infty}^{\infty}\delta(t-2)\sin\pi t\,dt$

(iv) $\int_{-\infty}^{\infty} f(t_1-t)\,\delta(t-t_2)\,dt$

(ii) $\int_{-\infty}^{\infty}\delta(t+3)e^{-t}\,dt$

(v) $\int_{-\infty}^{\infty} e^{j\pi t}\,\delta(t+2)\,dt$

(iii) $\int_{-\infty}^{\infty}\delta(1-t)(t^3+4)\,dt$

(vi) $f(t)*\delta(t-t_0)$

4.

Show that: $\delta\left(\frac{t-t_0}{T}\right)=|T|\delta(t-t_0)$

Hint: Show that: $\int_{-\infty}^{\infty} f(t)\delta\left(\frac{t-t_0}{T}\right)dt=|T|f(t_0)$

This is a case where common sense does not prevail. At first glance it might appear that $\delta[(t-t_0)/T]=\delta(t-t_0)$ because "you obviously can't scale something whose width is zero to begin with." However, it is the impulse function's behaviour upon integration that is important, and not its shape! Also notice how it conforms nicely to our notion of the impulse function being a limit of a narrow rectangle function. If a rectangle function is scaled by *T*, then its area is also *T*.

5.

Complete the following table:

(a) $x(t)=27\cos(100\pi t+15°)$	$\overline{X}=$
(b) $x(t)=-5\sin(2t-80°)$	$\overline{X}=$
(c) $x(t)=100\sin(2\pi t+45°)$	phasor diagram is:
(d) $x(t)=16\cos(314t+30°)$	amplitude of $\cos(314t)$ term = amplitude of $-\sin(314t)$ term =
(e) $\overline{X}=27-j11$	$x(t)=$
(f) $\overline{X}=100\angle-60°$	$x(t)=$
(g) $x(t)=4\cos(2\pi t+45°)-2\sin(2\pi t)$	$\overline{X}=$
(h) $x(t)=2\cos(t+30°)$	$X=$ $X^*=$
(i) $\overline{X}=3\angle 30°$, $\omega=1$	$x(2)=$
(j) $\overline{X}=3\angle 30°$	$X=$ $X^*=$

Leonhard Euler (1707-1783) *(Len´ ard Oy´ ler)*

The work of Euler built upon that of Newton and made mathematics the tool of analysis. Astronomy, the geometry of surfaces, optics, electricity and magnetism, artillery and ballistics, and hydrostatics are only some of Euler's fields. He put Newton's laws, calculus, trigonometry, and algebra into a recognizably modern form.

Euler was born in Switzerland, and before he was an adolescent it was recognized that he had a prodigious memory and an obvious mathematical gift. He received both his bachelor's and his master's degrees at the age of 15, and at the age of 23 he was appointed professor of physics, and at age 26 professor of mathematics, at the Academy of Sciences in Russia.

Among the symbols that Euler initiated are the sigma (Σ) for summation (1755), *e* to represent the constant 2.71828…(1727), *i* for the imaginary $\sqrt{-1}$ (1777), and even *a*, *b*, and *c* for the sides of a triangle and *A*, *B*, and *C* for the opposite angles. He transformed the trigonometric ratios into functions and abbreviated them $\sin$, $\cos$ and $\tan$, and treated logarithms and exponents as functions instead of merely aids to calculation. He also standardised the use of π for 3.14159…

His 1736 treatise, *Mechanica*, brought a universality to Newton mechanics with the guidance of a rigorous mathematical notation. An introduction to pure mathematics, *Introductio in analysin infinitorum*, appeared in 1748 which treated algebra, the theory of equations, trigonometry and analytical geometry. In this work Euler gave the formula $e^{ix}=\cos x+i\sin x$. It did for calculus what Euclid had done for geometry. Euler also published the first two complete works on calculus: *Institutiones calculi differentialis*, from 1755, and *Institutiones calculi integralis*, from 1768.

Euler's work in mathematics is vast. He was the most prolific writer of mathematics of all time. After his death in 1783 the St Petersburg Academy continued to publish Euler's unpublished work for nearly 50 more years!

Some of his phenomenal output includes: books on the calculus of variations; on the calculation of planetary orbits; on artillery and ballistics; on analysis; on shipbuilding and navigation; on the motion of the moon; lectures on the differential calculus. He made decisive and formative contributions to geometry, calculus and number theory. He integrated Leibniz's differential calculus and Newton's method of fluxions into mathematical analysis. He introduced beta and gamma functions, and integrating factors for differential equations. He studied continuum mechanics, lunar theory, the three body problem, elasticity, acoustics, the wave theory of light, hydraulics, and music. He laid the foundation of analytical mechanics. He proved many of Fermat's assertions including Fermat's Last Theorem for the case $n = 3$. He published a full theory of logarithms of complex numbers. Analytic functions of a complex variable were investigated by Euler in a number of different contexts, including the study of orthogonal trajectories and cartography. He discovered the Cauchy-Riemann equations used in complex variable theory.

Euler made a thorough investigation of integrals which can be expressed in terms of elementary functions. He also studied beta and gamma functions. As well as investigating double integrals, Euler considered ordinary and partial differential equations. The calculus of variations is another area in which Euler made fundamental discoveries.

He considered linear equations with constant coefficients, second order differential equations with variable coefficients, power series solutions of differential equations, a method of variation of constants, integrating factors, a method of approximating solutions, and many others. When considering vibrating membranes, Euler was led to the Bessel equation which he solved by introducing Bessel functions.

Euler made substantial contributions to differential geometry, investigating the theory of surfaces and curvature of surfaces. Many unpublished results by Euler in this area were rediscovered by Gauss.

Euler considered the motion of a point mass both in a vacuum and in a resisting medium. He analysed the motion of a point mass under a central force and also

considered the motion of a point mass on a surface. In this latter topic he had to solve various problems of differential geometry and geodesics.

He wrote a two volume work on naval science. He decomposed the motion of a solid into a rectilinear motion and a rotational motion. He studied rotational problems which were motivated by the problem of the precession of the equinoxes.

He set up the main formulas for the topic of fluid mechanics, the continuity equation, the Laplace velocity potential equation, and the Euler equations for the motion of an inviscid incompressible fluid.

He did important work in astronomy including: the determination of the orbits of comets and planets by a few observations; methods of calculation of the parallax of the sun; the theory of refraction; consideration of the physical nature of comets.

Euler also published on the theory of music...

Euler did not stop working in old age, despite his eyesight failing. He eventually went blind and employed his sons to help him write down long equations. Euler died of a stroke after a day spent: giving a mathematics lesson to one of his grandchildren; doing some calculations on the motion of balloons; and discussing the calculation of the orbit of the planet Uranus, recently discovered by William Herschel.

Lecture 1B – Systems

Differential equations. System modelling. Discrete-time signals and systems. Difference equations. Discrete-time block diagrams. Discretization in time of differential equations. Convolution in LTI discrete-time systems. Convolution in LTI continuous-time systems. Graphical description of convolution. Properties of convolution. Numerical convolution.

Linear Differential Equations with Constant Coefficients

Modelling of real systems involves approximating the real system to such a degree that it is tractable to our mathematics. Obviously the more assumptions we make about a system, the simpler the model, and the more easily solved. The more accurate we make the model, the harder it is to analyse. We need to make a trade-off based on some specification or our previous experience.

A lot of the time our modelling ends up describing a continuous-time system that is linear, time-invariant (LTI) and finite dimensional. In these cases, the system is described by the following equation:

Linear differential equation

$$\boxed{y^{(N)}(t)+\sum_{i=0}^{N-1} a_i y^{(i)}(t)=\sum_{i=0}^{M} b_i x^{(i)}(t)} \tag{1B.1}$$

where:

$$y^{(N)}(t)=\frac{d^N y(t)}{dt^N} \tag{1B.2}$$

Initial Conditions

The above equation needs the N initial conditions:

$$\boxed{y(0^-),\, y^{(1)}(0^-),\, \ldots,\, y^{(N-1)}(0^-)} \tag{1B.3}$$

We take 0^- as the time for initial conditions to take into account the possibility of an impulse being applied at $t=0$, which will change the output instantaneously.

First-Order Case

For the first order case we can express the solution to Eq. (1B.1) in a useful (and familiar) form. A first order system is given by:

First-order linear differential equation

$$\boxed{\frac{dy(t)}{dt}+ay(t)=bx(t)} \tag{1B.4}$$

To solve, first multiply both sides by an *integrating factor* equal to e^{at}. This gives:

$$e^{at}\left[\frac{dy(t)}{dt}+ay(t)\right]=e^{at}bx(t) \tag{1B.5}$$

Thus:

$$\frac{d}{dt}\left[e^{at}y(t)\right]=e^{at}bx(t) \tag{1B.6}$$

Integrating both sides gives:

$$e^{at}y(t)-y(0^-)=\int_{0^-}^{t}e^{a\tau}bx(\tau)d\tau, \quad t\geq 0 \tag{1B.7}$$

Finally, dividing both sides by the integrating factor gives:

First glimpse at a convolution integral – as the solution of a first-order linear differential equation

$$y(t)=e^{-at}y(0^-)+\int_{0^-}^{t}e^{-a(t-\tau)}bx(\tau)d\tau, \quad t\geq 0 \tag{1B.8}$$

Use this to solve the simple revision problem for the case of the unit step.

The two parts of the response given in Eq. (1B.8) have the obvious names zero-input response (ZIR) and zero-state response (ZSR). It will be shown later that the ZSR is given by a convolution between the system's impulse response and the input signal.

System Modelling

In modelling a system, we are nearly always after the input/output relationship, which is a differential equation in the case of continuous-time systems. If we're clever, we can break a system down into a connection of simple components, each having a relationship between cause and effect.

Electrical Circuits

The three basic linear, time-invariant relationships for the resistor, capacitor and inductor are respectively:

$$v(t)=Ri(t) \tag{1B.9a}$$

$$i(t)=C\frac{dv(t)}{dt} \tag{1B.9b}$$

$$v(t)=L\frac{di(t)}{dt} \tag{1B.9c}$$

Cause / effect relationships for electrical systems

Mechanical Systems

In linear translational systems, the three basic linear, time-invariant relationships for the inertia force, damping force and spring force are respectively:

$$F(t)=M\frac{d^2x(t)}{dt} \tag{1B.10a}$$

$$F(t)=k_d\frac{dx(t)}{dt} \tag{1B.10b}$$

$$F(t)=k_s x(t) \tag{1B.10c}$$

Cause / effect relationships for mechanical translational systems

Where $x(t)$ is the position of the object under study.

For rotational motion, the relationships for the inertia torque, damping torque and spring torque are:

Cause / effect relationships for mechanical rotational systems

$$F(t) = I\,\frac{d^2\theta(t)}{dt} \tag{1B.11a}$$

$$F(t) = k_d\,\frac{d\theta(t)}{dt} \tag{1B.11b}$$

$$F(t) = k_s\theta(t) \tag{1B.11c}$$

Finding an input-output relationship for signals in systems is just a matter of applying the above relationships to a conservation law: for electrical circuits it is one of Kirchhoff's laws, in mechanical systems it is D'Alembert's principle.

Discrete-time Systems

Discrete-time systems are important…

A discrete-time signal is one that takes on values only at discrete instants of time. Discrete-time signals arise naturally in studies of economic systems – amortization (paying off a loan), models of the national income (monthly, quarterly or yearly), models of the inventory cycle in a factory, etc. They arise in science, eg. in studies of population, chemical reactions, the deflection of a weighted beam. They arise all the time in electrical engineering, because of digital control eg. radar tracking system, processing of electrocardiograms, digital communication (CD, mobile phone, Internet). Their importance is probably now reaching that of continuous-time systems in terms of analysis and design – specifically because today signals are processed digitally, and hence they are a special case of discrete-time signals.

…especially as microprocessors play a central role in today's signal processing

It is now cheaper and easier to perform most signal operations inside a microprocessor or microcontroller than it is with an equivalent analog continuous-time system. But since there is a great depth to the analysis and design techniques of continuous-time systems, and since most physical systems are continuous-time in nature, it is still beneficial to study systems in the continuous-time domain.

Linear Difference Equations with Constant Coefficients

Linear, time-invariant, discrete-time systems can be modelled with the difference equation:

$$y[n] + \sum_{i=1}^{N} a_i y[n-i] = \sum_{i=0}^{M} b_i x[n-i] \tag{1B.12}$$

Linear time-invariant (LTI) difference equation

Solution by Recursion

We can solve difference equations by a direct numerical procedure.

There is a MATLAB® function available for download from the Signals and Systems web site called `recur` *that solves the above equation.*

Complete Solution

By solving Eq. (1B.12) recursively it is possible to generate an expression for the complete solution $y[n]$ in terms of the initial conditions and the input $x[n]$.

First-Order Case

Consider the first-order linear difference equation:

$$y[n] + ay[n-1] = bx[n] \tag{1B.13}$$

First-order linear difference equation

with initial condition $y[-1]$. *By successive substitution, show that:*

$$\begin{aligned} y[0] &= -ay[-1] + bx[0] \\ y[1] &= a^2 y[-1] - abx[0] + bx[1] \\ y[2] &= -a^3 y[-1] + a^2 bx[0] - abx[1] + bx[2] \end{aligned} \tag{1B.14}$$

From the pattern, it can be seen that for $n \geq 0$,

First look at a convolution summation – as the solution of a first-order linear difference equation

$$\boxed{y[n] = (-a)^{n+1} y[-1] + \sum_{i=0}^{n} (-a)^{n-i} bx[i]} \tag{1B.15}$$

This solution is the discrete-time counterpart to Eq. (1B.8).

Discrete-Time Block Diagrams

An LTI discrete-time system can be represented as a block diagram consisting of adders, gains and delays. The gain element is shown below:

A discrete-time gain element

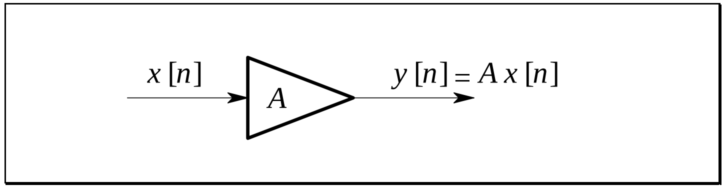


Figure 1B.1

The *unit-delay* element is shown below:

A discrete-time unit-delay element

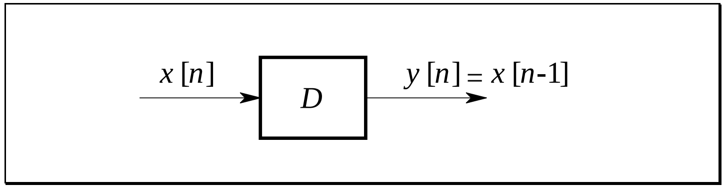
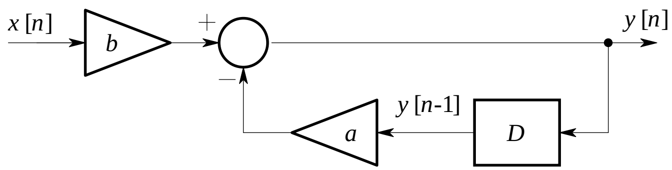


Figure 1B.2

Such an element is normally implemented by the memory of a computer, or a digital delay line.

Example

Using these two elements and an adder, we can construct a representation of the discrete-time system given by $y[n] + ay[n-1] = bx[n]$. The system is shown below:



Discretization in Time of Differential Equations

Often we wish to use a computer for the solution of continuous-time differential equations. We can: if we are careful about interpreting the results.

First-Order Case

Let's see if we can approximate the first-order linear differential equation given by Eq. (1B.4) with a discrete-time equation. We can approximate the continuous-time derivative using *Euler's approximation*, or forward difference:

Approximating a derivative with a difference

$$\boxed{\left.\frac{dy(t)}{dt}\right|_{t=nT} \approx \frac{y(nT+T) - y(nT)}{T}} \tag{1B.16}$$

If T is suitably small and $y(t)$ is continuous, the approximation will be accurate. Substituting this approximation into Eq. (1B.4) results in a discrete-time approximation given by the difference equation:

The first-order difference equation approximation of a first-order differential equation

$$\boxed{y[n] \approx (1 - aT) y[n-1] + bTx[n-1]} \tag{1B.17}$$

The discrete values $y[n]$ are approximations to the solution $y(nT)$.

Show that $y[n]$ *gives approximate values of the solution* $y(t)$ *at the times* $t = nT$ *with arbitrary initial condition* $y[-1]$ *for the special case of zero input.*

Second-order Case

We can generalize the discretization process to higher-order differential equations. In the second-order case the following approximation can be used:

The second-order difference equation approximation of a second-order derivative

$$\boxed{\left.\frac{d^2 y(t)}{dt^2}\right|_{t=nT} \approx \frac{y(nT+2T)-2y(nT+T)+y(nT)}{T^2}} \tag{1B.18}$$

Now consider the second-order differential equation:

$$\boxed{\frac{d^2 y(t)}{dt^2}+a_1\frac{dy(t)}{dt}+a_0 y(t)=b_1\frac{dx(t)}{dt}+b_0 x(t)} \tag{1B.19}$$

Show that the discrete-time approximation to the solution $y(t)$ *is given by:*

An n^{th}-order differential equation can be approximated with an n^{th}-order difference equation

$$\boxed{\begin{aligned} y[n] &= (2-a_1T)y[n-1]-(1-a_1T+a_0T^2)y[n-2] \\ &\quad +b_1Tx[n-1]+(b_0T^2-b_1T)x[n-2] \end{aligned}} \tag{1B.20}$$

Convolution in Linear Time-invariant Discrete-time Systems

Although the linear difference equation is the most basic description of a linear discrete-time system, we can develop an *equivalent* representation called the convolution representation. This representation will help us to determine important system properties that are not readily apparent from observation of the difference equation.

One advantage of this representation is that the output is written as a linear combination of past and present input signal elements. *It is only valid when the system's initial conditions are all zero.*

First-Order System

We have previously considered the difference equation:

Linear difference equation

$$y[n]+ay[n-1]=bx[n] \tag{1B.21}$$

and showed by successive substitution that:

The complete response

$$y[n]=(-a)^{n+1}y[-1]+\sum_{i=0}^{n}(-a)^{n-i}bx[i] \tag{1B.22}$$

By the definition of the convolution representation, we are after an expression for the output *with all initial conditions zero*. We then have:

The zero-state response (ZSR) – a convolution summation

$$y[n]=\sum_{i=0}^{n}(-a)^{n-i}bx[i] \tag{1B.23}$$

In contrast to Eq. (1B.21), we can see that Eq. (1B.23) depends exclusively on present and past values of the input signal. One advantage of this is that we may directly observe how each past input affects the present output signal. For example, an input $x[i]$ contributes an amount $(-a)^{n-i}bx[i]$ to the totality of the output at the n^{th} period.

Unit-Pulse Response of a First-Order System

A discrete-time system's unit-pulse response defined

The output of a system subjected to a unit-pulse response $\delta[n]$ is denoted $h[n]$ and is called the *unit-pulse response*, or *weighting sequence* of the discrete-time system. It is very important because it completely characterises a system's behaviour. It may also provide an experimental or mathematical means to determine system behaviour.

For the first-order system of Eq. (1B.21), if we let $x[n]=\delta[n]$, then the output of the system to a unit-pulse input can be expressed using Eq. (1B.23) as:

$$y[n]=\sum_{i=0}^{n}(-a)^{n-i}b\delta[i] \tag{1B.24}$$

which reduces to:

$$y[n]=(-a)^{n}b \tag{1B.25}$$

The unit-pulse response for this system is therefore given by:

A first-order discrete-time system's unit-pulse response

$$h[n]=(-a)^{n}bu[n] \tag{1B.26}$$

General System

For a general linear time-invariant (LTI) system, the response to a delayed unit-pulse $\delta[n-i]$ must be $h[n-i]$.

Since $x[n]$ can be written as:

$$x[n]=\sum_{i=0}^{\infty}x[i]\delta[n-i] \tag{1B.27}$$

and since the system is LTI, the response $y_i[n]$ to $x[i]\delta[n-i]$ is given by:

$$y_i[n]=x[i]h[n-i] \tag{1B.28}$$

The response to the sum Eq. (1B.27) must be equal to the sum of the $y_i[n]$ defined by Eq. (1B.28). Thus the response to $x[n]$ is:

$$\boxed{y[n]=\sum_{i=0}^{\infty}x[i]h[n-i]} \tag{1B.29}$$

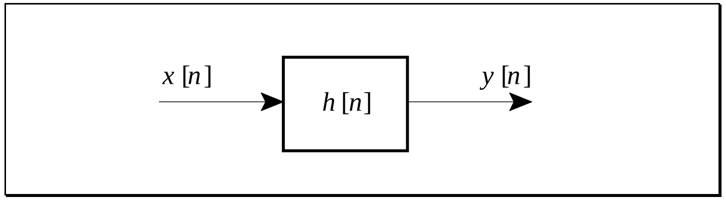
Convolution summation defined for a discrete-time system

This is the convolution representation of a discrete-time system, also written as:

$$\boxed{y[n]=h[n]*x[n]} \tag{1B.30}$$

Convolution notation for a discrete-time system

Graphically, we can now represent the system as:



Graphical notation for a discrete-time system using the unit-pulse response

Figure 1B.3

It should be pointed out that the convolution representation is not very efficient in terms of a digital implementation of the output of a system (needs lots more memory and calculating time) compared with the difference equation.

Convolution is *commutative* which means that it is also true to write:

$$\boxed{y[n]=\sum_{i=0}^{\infty}h[i]x[n-i]} \tag{1B.31}$$

Discrete-time convolution can be illustrated as follows. Suppose the unit-pulse response is that of a filter of finite length *k*. Then the output of such a filter is:

$$\begin{aligned} y[n] &= h[n] * x[n] \\ &= \sum_{i=0}^{\infty} h[i]x[n-i] \\ &= h[0]x[n] + h[1]x[n-1] + \ldots + h[k]x[n-k] \end{aligned} \tag{1B.32}$$

Graphically, this summation can be viewed as two buffers, or arrays, sliding past one another. The array locations that overlap are multiplied and summed to form the output at that instant.

Graphical view of the convolution operation in discrete-time

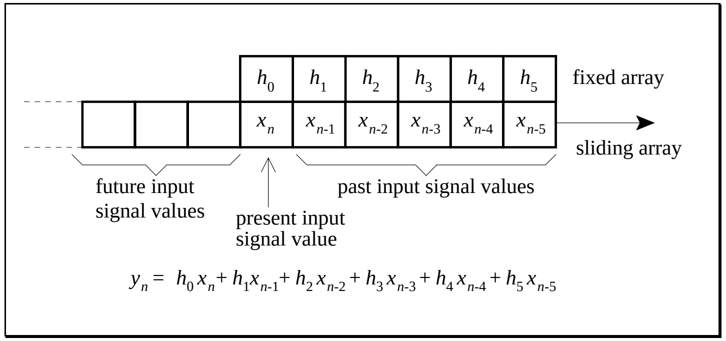


Figure 1B.4

In other words, the output at time *n* is equal to a linear combination of past and present values of the input signal, *x*. The system can be considered to have a memory because at any particular time, the output is still responding to an input at a previous time.

Discrete-time convolution can be implemented by a *transversal digital filter*:

Transversal digital filter performs discrete-time convolution

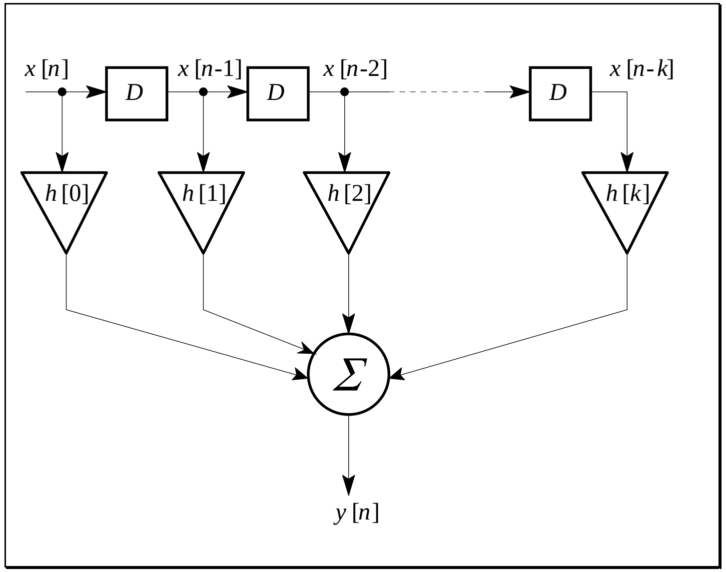


Figure 1B.5

MATLAB® can do convolution for us. Use the `conv` *function.*

System Memory

A system's memory can be roughly interpreted as a measure of how significant past inputs are on the current output. Consider the two unit-pulse responses below:

System memory depends on the unit-pulse response…

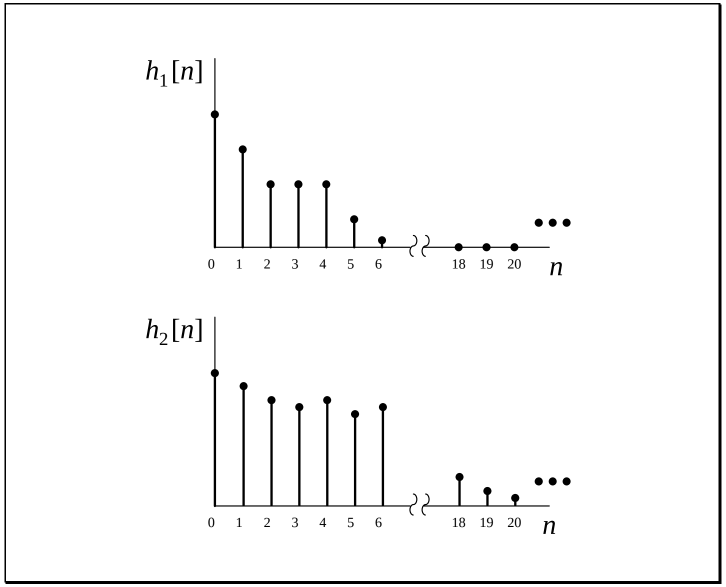


Figure 1B.6

System 1 depends strongly on inputs applied five or six iterations ago and less so on inputs applied more than six iterations ago. The output of system 2 depends strongly on inputs 20 or more iterations ago. System 1 is said to have a shorter memory than system 2.

…specifically - on how long it takes to decay to zero.

It is apparent that a measure of system memory is obtained by noting how quickly the system unit-pulse response decays to zero: *the more quickly a system's weighting sequence goes to zero, the shorter the memory*. Some applications require a short memory, where the output is more readily influenced by the most recent behaviour of the input signal. Such systems are *fast responding*. A system with long memory does not respond as readily to changes in the recent behaviour of the input signal and is said to be *sluggish*.

System Stability

BIBO stability defined

A system is stable if its output signal remains bounded in response to any bounded signal. (1B.33)

If a bounded input (BI) produces a bounded output (BO), then the system is termed BIBO stable. This implies that:

$$\lim_{i\to\infty} h[i] = 0 \tag{1B.34}$$

This is something not readily apparent from the difference equation. A more thorough treatment of system stability will be given later.

What can you say about the stability of the system described by Eq. (1B.21)?

Convolution in Linear Time-invariant Continuous-time Systems

Deriving convolution for the continuous-time case

The input / output relationship of a continuous time system can be specified in terms of a convolution operation between the input and the impulse response of the system.

Recall that we can consider the impulse as the limit of a rectangle function:

Start with a rectangle input

$$x_r(t) = \frac{1}{T}\operatorname{rect}\left(\frac{t}{T}\right) \tag{1B.35}$$

as $T \to 0$. The system response to this input is:

and the output response.

$$y(t) = y_r(t) \tag{1B.36}$$

and since:

As the input approaches an impulse function

$$\lim_{T\to 0} x_r(t) = \delta(t) \tag{1B.37}$$

then:

then the output approaches the impulse response

$$\lim_{T \to 0} y_r(t) = h(t) \tag{1B.38}$$

Now expressing the general input signal as the limit of a staircase approximation as shown in Figure 1B.7:

Treat an arbitrary input waveform as a sum of rectangles

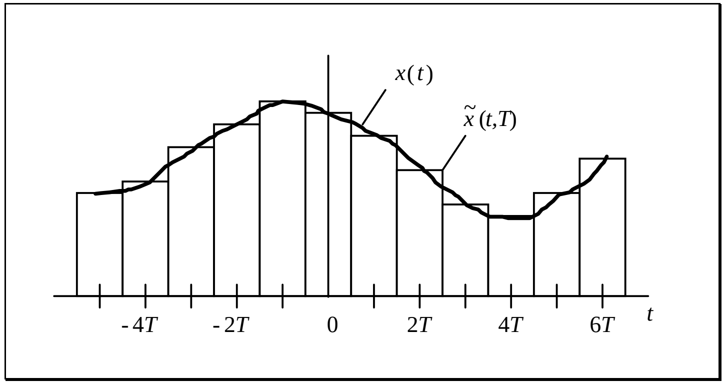


Figure 1B.7

we have:

which get smaller and smaller and eventually approach the original waveform

$$x(t) = \lim_{T \to 0} \tilde{x}(t,T) \tag{1B.39}$$

where:

$$\tilde{x}(t,T) = \sum_{i=-\infty}^{\infty} x(iT)\,\mathrm{rect}\left(\frac{t-iT}{T}\right) \tag{1B.40}$$

We can rewrite Eq. (1B.40) using Eq. (1B.35) as:

The staircase is just a sum of weighted rectangle inputs…

$$\tilde{x}(t,T) = \sum_{i=-\infty}^{\infty} x(iT)T x_r(t-iT) \tag{1B.41}$$

Since the system is time-invariant, the response to $x_r(t-iT)$ is $y_r(t-iT)$. Therefore the system response to $\tilde{x}(t,T)$ is:

$$\tilde{y}(t,T) = \sum_{i=-\infty}^{\infty} x(iT)T y_r(t-iT) \tag{1B.42}$$

…and we already know the output…

because superposition holds for linear systems. The system response to $x(t)$ is just the response:

$$y(t) = \lim_{T \to 0} \tilde{y}(t,T) = \lim_{T \to 0} \sum_{i=-\infty}^{\infty} x(iT) y_r(t-iT)T \tag{1B.43}$$

…even in the limit as the staircase approximation approaches the original input

When we perform the limit, $x(iT) \to x(\tau)$, $y_r(t-iT) \to h(t-\tau)$ and $T \to d\tau$. Hence the output response can be expressed in the form:

$$\boxed{y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)\,d\tau} \tag{1B.44}$$

Convolution integral for continuous-time systems defined

If the input $x(t) = 0$ for all $t < 0$ then:

$$y(t) = \int_{0}^{\infty} x(\tau)h(t-\tau)\,d\tau \tag{1B.45}$$

Convolution integral if the input starts at time t=0

If the input is causal, then $h(t-\tau) = 0$ for negative arguments, i.e. when $\tau > t$. The upper limit in the integration can then be changed so that:

$$y(t) = \int_{0}^{t} x(\tau)h(t-\tau)\,d\tau \tag{1B.46}$$

Convolution integral if the input starts at time t=0, and the system is causal

Once again, it can be shown that convolution is *commutative* which means that it is also true to write (*compare with Eq. (1B.31)*):

$$\boxed{y(t) = \int_{0}^{t} h(\tau)x(t-\tau)\,d\tau} \tag{1B.47}$$

Alternative way of writing the convolution integral

With the convolution operation denoted by an asterisk, "*", the input / output relationship becomes:

Convolution notation for a continuous-time system

$$\boxed{y(t) = h(t) * x(t)} \tag{1B.48}$$

Graphically, we can represent the system as:

Graphical notation for a continuous-time system using the impulse response

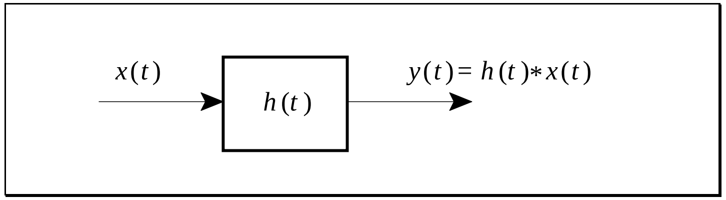


Figure 1B.8

It should be pointed out, once again, that the convolution relationship is *only valid when there is no initial energy stored in the system. ie. initial conditions are zero. The output response using convolution is just the ZSR.*

Graphical Description of Convolution

Consider the following continuous-time example which has a *causal* impulse response function. A causal impulse response implies that there is no response from the system until an impulse is applied at $t = 0$. In other words, $h(t) = 0$ for $t < 0$. Let the impulse response of the system be a decaying exponential, and let the input signal be the unit-step:

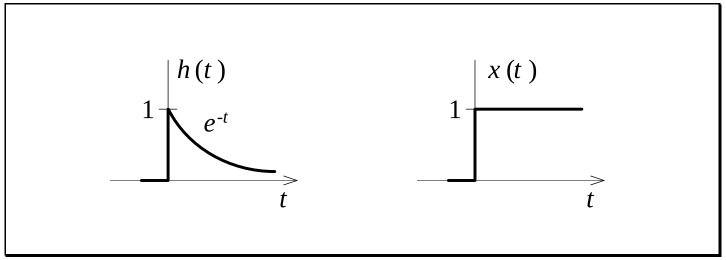


Figure 1B.9

Using graphical convolution, the output $y(t)$ can be obtained. First, the input signal is flipped in time about the origin. Then, as the time "parameter" t advances, the input signal "slides" past the impulse response – in much the same way as the input values slide past the unit-pulse values for discrete-time convolution. You can think of this graphical technique as the continuous-time version of a digital transversal filter (you might like to think of it as a discrete-time system and input signal, with the time delay between successive values so tiny that the finite summation of Eq. (1B.30) turns into a continuous-time integration).

When $t = 0$, there is obviously no overlap between the impulse response and input signal. The output must be zero since we have assumed the system to be in the zero-state (all initial conditions zero). Therefore $y(0) = 0$. This is illustrated below:

Graphical illustration of continuous-time - "snapshot" at t=0

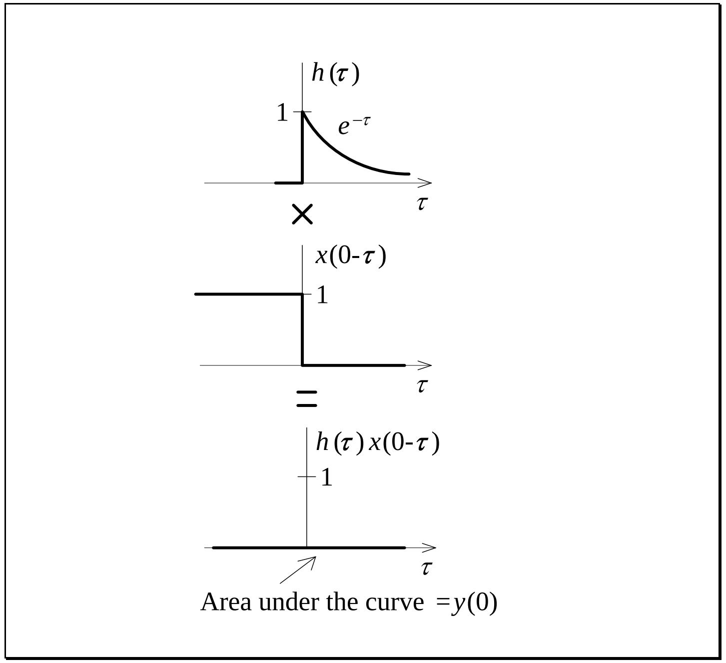


Figure 1B.10

Letting time "roll-on" a bit further, we take a snapshot of the situation when $t=1$. This is shown below:

Graphical illustration of continuous-time - "snapshot" at t=1

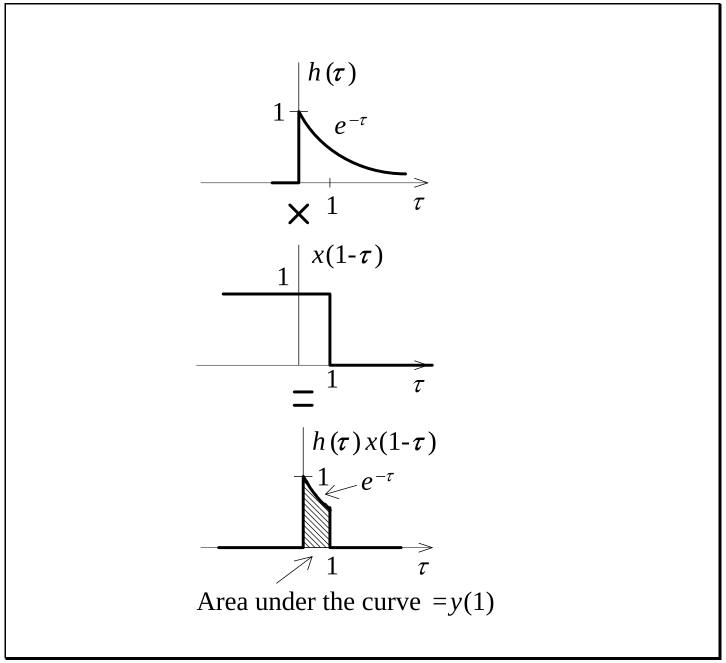
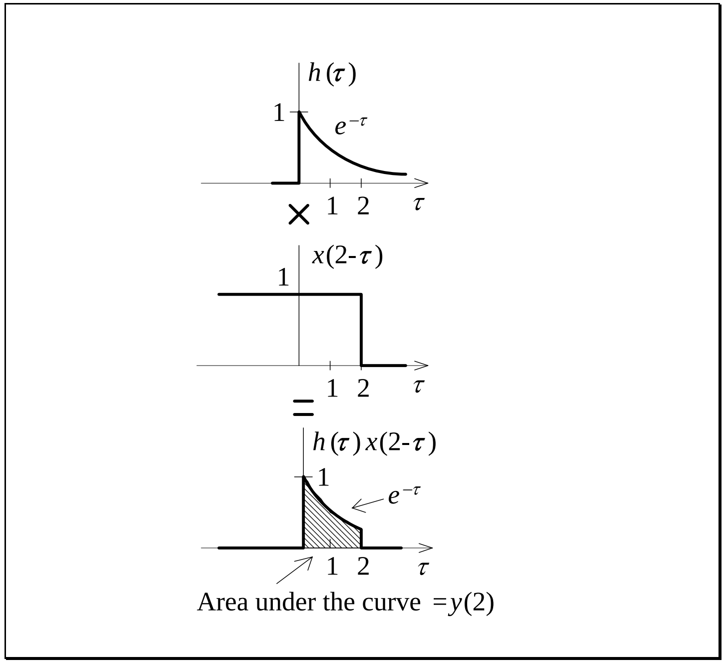


Figure 1B.11

The output *value* at $t=1$ is now given by:

$$\begin{aligned} y(1) &= \int_0^1 h(\tau)x(1-\tau)d\tau \\ &= \int_0^1 e^{-\tau}d\tau = \left[e^{-\tau}\right]_0^1 = 1-e^{-1} \approx 0.63 \end{aligned} \tag{1B.49}$$

Taking a snapshot at $t=2$ gives:



Graphical illustration of continuous-time - "snapshot" at t=2

Figure 1B.12

The output *value* at $t=2$ is now given by:

$$\begin{aligned} y(2) &= \int_0^2 h(\tau)x(2-\tau)d\tau \\ &= \int_0^2 e^{-\tau}d\tau = \left[e^{-\tau}\right]_0^2 = 1-e^{-2} \approx 0.86 \end{aligned} \tag{1B.50}$$

If we keep evaluating the output for various values of t, we can build up a graphical picture of the output for all time:

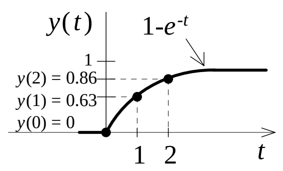


Figure 1B.13

In this simple case, it is easy to verify the graphical solution using Eq. (1B.47). The output value at any time t is given by:

$$\begin{aligned} y(t) &= \int_0^t h(\tau)x(t-\tau)d\tau \\ &= \int_0^t e^{-\tau}d\tau = \left[e^{-\tau}\right]_0^t = 1-e^{-t} \end{aligned} \tag{1B.51}$$

In more complicated situations, it is often the graphical approach that provides a quick insight into the form of the output signal, and it can be used to give a rough sketch of the output without too much work.

Properties of Convolution

In the following list of continuous-time properties, the notation $x(t) \rightarrow y(t)$ should be read as "the input $x(t)$ produces the output $y(t)$". Similar properties also hold for discrete-time convolution.

Convolution properties

$$ax(t) \rightarrow ay(t) \tag{1B.52a}$$

$$x_1(t)+x_2(t) \rightarrow y_1(t)+y_2(t) \tag{1B.52b}$$

$$a_1x_1(t)+a_2x_2(t) \rightarrow a_1y_1(t)+a_2y_2(t) \tag{1B.52c}$$

Linearity

$$x(t-t_0) \rightarrow y(t-t_0) \tag{1B.52d}$$

Time-invariance

Convolution is also associative, commutative and distributive with addition, all due to the linearity property.

Numerical Convolution

We have already looked at how to discretize a continuous-time system by discretizing a system's input / output differential equation. The following procedure provides another method for discretizing a continuous-time system. It should be noted that the two different methods produce two *different* discrete-time representations.

Computers work with discrete data

We start by thinking about how to simulate a continuous-time convolution with a computer, which operates on discrete-time data. The integral in Eq. (1B.47) can be discretized by setting $t = nT$:

$$y(nT) = \int_0^{nT} h(\tau)x(nT-\tau)d\tau \tag{1B.53}$$

By effectively reversing the procedure in arriving at Eq. (1B.47), we can break this integral into regions of width T:

$$\begin{aligned} y(nT) &= \int_0^T h(\tau)x(nT-\tau)d\tau \\ &\quad + \int_T^{2T} h(\tau)x(nT-\tau)d\tau + \ldots \\ &\quad + \int_{iT}^{(i+1)T} h(\tau)x(nT-\tau)d\tau + \ldots \end{aligned} \tag{1B.54}$$

which can be rewritten using the summation symbol:

$$y(nT) \approx \sum_{i=0}^{n} \int_{iT}^{iT+T} h(\tau)x(nT-\tau)d\tau \tag{1B.55}$$

If T is small enough, $h(\tau)$ and $x(nT-\tau)$ can be taken to be constant over each interval:

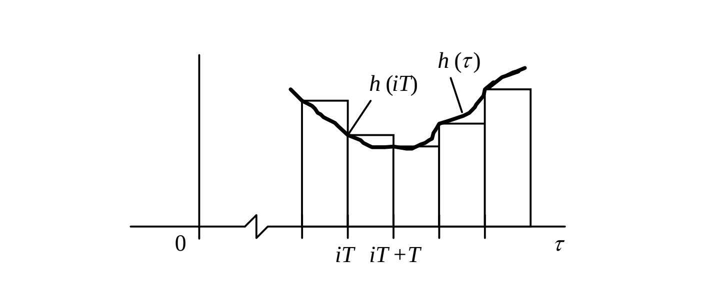


Figure 1B.14

That is, apply Euler's approximation:

$$\begin{aligned} h(\tau) &\approx h(iT) \\ x(nT-\tau) &\approx x(nT-iT) \end{aligned} \tag{1B.56}$$

so that Eq. (1B.55) becomes:

$$y(nT) \approx \sum_{i=0}^{n} \int_{iT}^{iT+T} h(iT)x(nT-iT)d\tau \tag{1B.57}$$

Since the integrand is constant with respect to τ, it can be moved outside the integral which is easily evaluated:

$$y(nT) \approx \sum_{i=0}^{n} h(iT)x(nT-iT)T \tag{1B.58}$$

We approximate the integral with a summation

Writing in the notation for discrete-time signals, we have the following input / output relationship:

$$\boxed{y[n] \approx \sum_{i=0}^{n} h[i]x[n-i]T, \quad n = 0, 1, 2, \ldots} \tag{1B.59}$$

Convolution approximation for causal systems with inputs applied at t=0

This equation can be viewed as the convolution-summation representation of a linear time-invariant system with unit-pulse response $Th[n]$, where $h[n]$ is the sampled version of the impulse response $h(t)$ of the original continuous-time system.

Convolution with an Impulse

One very important particular case of convolution that we will use all the time is that of convolving a function with a delayed impulse. We can tackle the problem three ways: graphically, algebraically, or by using the concept that a system performs convolution. Using this last approach, we can surmise what the solution is by recognising that the convolution of a function $h(t)$ with an impulse is equivalent to applying an impulse to a system that has an impulse response given by $h(t)$:

Applying an impulse to a system creates the impulse response

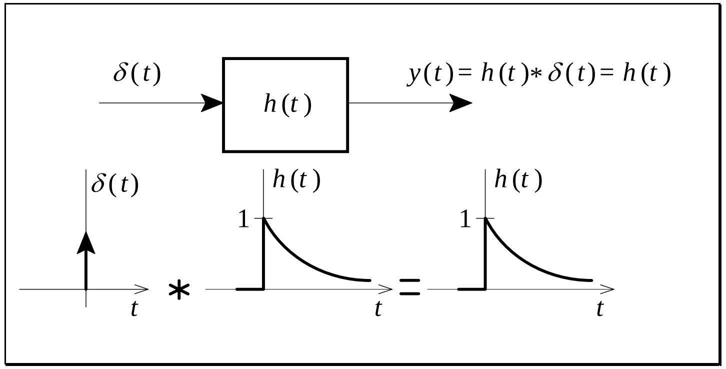


Figure 1B.15

The output, by definition, is the impulse response, $h(t)$. We can also arrive at this result algebraically by performing the convolution integral, and noting that it is really a sifting integral:

$$\delta(t) * h(t) = \int_{-\infty}^{\infty} \delta(\tau) h(t-\tau) d\tau = h(t) \tag{1B.60}$$

If we now apply a delayed impulse to the system, and since the system is time-invariant, we should get out a delayed impulse response:

Applying a delayed impulse to a system creates a delayed impulse response

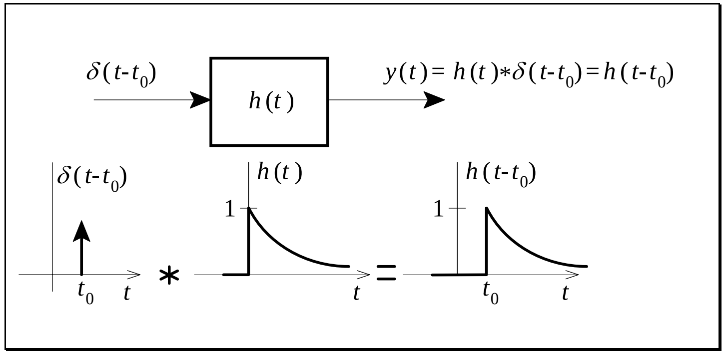


Figure 1B.16

Again, using the definition of the convolution integral and the sifting property of the impulse, we can arrive at the result algebraically:

$$\begin{aligned} \delta(t-t_0) * h(t) &= \int_{-\infty}^{\infty} \delta(\tau - t_0) h(t-\tau) d\tau \\ &= h(t-t_0) \end{aligned} \tag{1B.61}$$

Therefore, in general, we have:

$$\boxed{f(x) * \delta(x-x_0) = f(x-x_0)} \tag{1B.62}$$

This can be represented graphically as:

Convolving a function with an impulse shifts the original function to the impulse's location

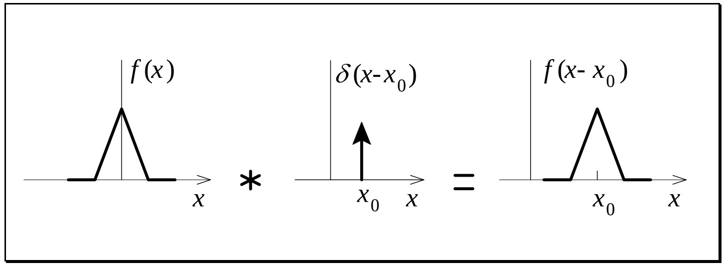


Figure 1B.17

Summary

- Systems are predominantly described by differential or difference equations – they are the equations of *dynamics*, and tell us how outputs and various *states* of the system change with time for a given input.
- Most systems can be derived from simple cause / effect relationships, together with a few conservation laws.
- Discrete-time signals occur naturally and frequently – they are signals that exist only at discrete points in time. Discrete-time systems are commonly implemented using microprocessors.
- We can approximate continuous-time systems with discrete-time systems by a process known as discretization – we replace differentials with differences.
- Convolution is another (equivalent) way of representing an input / output relationship of a system. It shows us features of the system that were otherwise "hidden" when written in terms of a differential or difference equation.
- Convolution introduces us to the concept of an impulse response for a continuous-time system, and a unit-pulse response for a discrete-time system. Knowing this response, we can determine the output for *any* input, *if the initial conditions are zero*.
- A system is BIBO stable if its impulse response decays to zero in the continuous-time case, or if its unit-pulse response decays to zero in the discrete-time case.
- Convolving a function with an impulse shifts the original function to the impulse's location.

References

Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB®*, Prentice-Hall International, Inc., 1997.

Exercises

1.

The following continuous-time functions are to be uniformly sampled. Plot the discrete signals which result if the sampling period T is (i) $T = 0.1\,\text{s}$, (ii) $T = 0.3\,\text{s}$, (iii) $T = 0.5\,\text{s}$, (iv) $T = 1\,\text{s}$. How does the sampling time affect the accuracy of the resulting signal?

(a) $x(t) = 1$ (b) $x(t) = \cos 4\pi t$ (c) $x(t) = \cos 10\pi t$

2.

Plot the sequences given by:

(a) $y_1[n] = 3\delta[n+1] - \delta[n] + 2\delta[n-1] + 1/2\,\delta[n-2]$

(b) $y_2[n] = -4\delta[n] - \delta[n-2] + 3\delta[n-3]$

3.

From your solution in Question 2, find $a[n] = y_1[n] - y_2[n]$. Show graphically that the resulting sequence is equivalent to the sum of the following delayed unit-step sequences:

$$a[n] = 3u[n+1] - u[n-1] - 1/2\,u[n-2] - 9/2\,u[n-3] + 3u[n-4]$$

4.

Find $y[n] = y_1[n] + y_2[n]$ when:

$$y_1[n] = \begin{cases} 0, & n = -1, -2, -3, \ldots \\ (-1)^{n^2-1}, & n = 0, 1, 2, \ldots \end{cases}$$

$$y_2[n] = \begin{cases} 0, & n = -1, -2, -3, \ldots \\ 1/2\left(1 + (-1)^n\right), & n = 0, 1, 2, \ldots \end{cases}$$

5.

The following series of numbers is known as the Fibonacci sequence:

$$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34\ldots$$

(a) Find a difference equation which describes this number sequence $y[n]$, when $y[0]=0$.

(b) By evaluating the first few terms show that the following formula also describes the numbers in the Fibonacci sequence:

$$y[n]=\frac{1}{\sqrt{5}}\left[\left(0.5+\sqrt{1.25}\right)^n-\left(0.5-\sqrt{1.25}\right)^n\right]$$

(c) Using your answer in (a) find $y[20]$ and $y[25]$. Check your results using the equation in (b). Which approach is easier?

6.

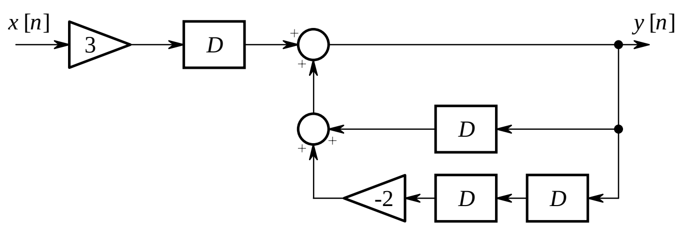
Construct block diagrams for the following difference equations:

(i) $y[n]=y[n-2]+x[n]+x[n-1]$

(ii) $y[n]=2y[n-1]-y[n-2]+3x[n-4]$

7.

(i) Construct a difference equation from the following block diagram:



(ii) From your solution calculate $y[n]$ for n = 0, 1, 2 and 3 given $y[-2]=-2$, $y[-1]=-1$, $x[n]=0$ for $n<0$ and $x[n]=(-1)^n$ for n = 0, 1, 2 …

8.

(a) Find the unit-pulse response of the linear systems given by the following equations:

(i) $y[n]=\frac{T}{2}\left(x[n]+x[n-1]\right)+y[n-1]$

(ii) $y[n]=x[n]-0.75x[n-1]+0.5y[n-1]$

(b) Determine the first five terms of the response of the equation in (ii) to the input:

$$x[n]=\begin{cases}0, & n=-2,-3,-4,\ldots\\ 1, & n=-1\\ (-1)^n, & n=0,1,2,\ldots\end{cases}$$

using (i) the basic difference equation, (ii) graphical convolution and (iii) the convolution summation. (Note $y[n]=0$ for $n\le -2$).

9.

For the single input-single output continuous- and discrete-time systems characterized by the following equations, determine which coefficients must be zero for the systems to be

(a) linear

(b) time invariant

(i) $a_1\left(\frac{d^3y}{dt^3}\right)^2+a_2\frac{d^2y}{dt^2}+\left(a_3+a_4y+a_5\sin t\right)\frac{dy}{dt}+a_6y=a_7x$

(ii) $a_1y^2[n+3]+a_2y[n+2]+\left(a_3+a_4y[n]+a_5\sin(n)\right)y[n+1]+a_6y[n]=a_7x[n]$

10.

To demonstrate that nonlinear systems do not obey the principle of superposition, determine the first five terms of the response of the system:

$$y[n] = 2y[n-1] + x^2[n]$$

to the input:

$$x_1[n] = \begin{cases} 0, & n = -1, -2, -3, \ldots \\ 1, & n = 0, 1, 2, \ldots \end{cases}$$

If $y_1[n]$ denotes this response, show that the response of the system to the input $x[n] = 2x_1[n]$ is not $2y_1[n]$.

Can convolution methods be applied to nonlinear systems? Why?

11.

A system has the unit-pulse response:

$$h[n] = 2u[n] - u[n-2] - u[n-4]$$

Find the response of this system when the input is the sequence:

$$\delta[n] - \delta[n-1] + \delta[n-2] - \delta[n-3]$$

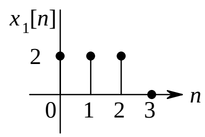
using (i) graphical convolution and (ii) convolution summation.

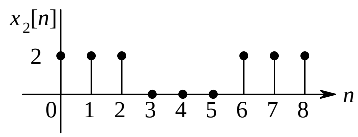
12.

For $x_1[n]$ and $x_2[n]$ as shown below find

(i) $x_1[n] * x_1[n]$ (ii) $x_1[n] * x_2[n]$ (iii) $x_2[n] * x_2[n]$

using (a) graphical convolution and (b) convolution summation.



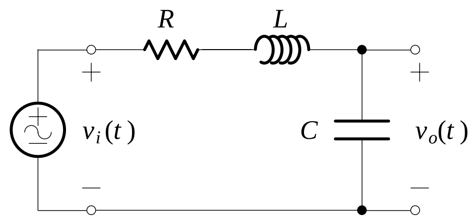


13.

Use MATLAB® and discretization to produce approximate solutions to the revision problem.

14.

Use MATLAB® to graph the output voltage of the following *RLC* circuit:

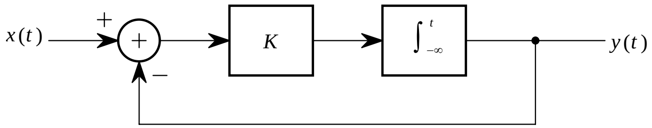


when $R = 2$, $L = C = 1$, $v_o(0) = 1$, $\dot{v}_o(0) = -1$ and $v_i(t) = \sin(t)u(t)$.

Compare with the exact solution: $v_o(t) = 0.5[(3+t)e^{-t} - \cos(t)]$, $t \geq 0$. How do you decide what value of T to use?

15.

A feedback control system is used to control a room's temperature with respect to a preset value. A simple model for this system is represented by the block diagram shown below:



In the model, the signal $x(t)$ represents the commanded temperature change from the preset value, $y(t)$ represents the produced temperature change, and t is measured in minutes. Find:

a) the differential equation relating $x(t)$ and $y(t)$,

b) the impulse response of the system, and

c) the temperature change produced by the system when the gain K is 0.5 and a step change of $0.75°$ is commanded at $t = 4\text{ min}$.

d) Plot the temperature change produced.

e) Use MATLAB® and numerical convolution to produce approximate solutions to this problem and compare with the theoretical answer.

16.

Use MATLAB® and the numerical convolution method to solve Q14.

17.

Quickly changing inputs to an aircraft rudder control are smoothed using a digital processor. That is, the control signal is converted to a discrete-time signal by an A/D converter, the discrete-time signal is smoothed with a discrete-time filter, and the smoothed discrete-time signal is converted to a continuous-time, smoothed, control signal by a D/A converter. The smoothing filter has the unit-pulse response:

$$h[nT] = \left(0.5^n - 0.25^n\right)u[nT], \quad T = 0.25\text{ s}$$

Find the zero-state response of the discrete-time filter when the input signal samples are:

$$x[nT] = \{1, 1, 1\}, \quad T = 0.25\text{ s}$$

Plot the input, unit-pulse response, and output for $-0.75 \le t \le 1.5\text{ s}$.

18.

A wave staff measures ocean wave height in meters as a function of time. The height signal is sampled at a rate of 5 samples per second. These samples form the discrete-time signal:

$$s[nT] = \cos(2\pi(0.2)nT + 1.1) + 0.5\cos(2\pi(0.3)nT + 1.5)$$

The signal is transmitted to a central wave-monitoring station. The transmission system corrupts the signal with additive noise given by the MATLAB® function:

```
function n0=drn(n)
  N=size(n,2);
  rand('seed', 0);
  no(1)=rand-0.5;
  for I=2:N;
    no(i)=0.2*no(i-1)+(rand-0.5);
end
```

The received signal plus noise, $x[nT]$, is processed with a low-pass filter to reduce the noise.

The filter unit-pulse response is:

$$h[nT] = \{0.182(0.76)^n - 0.144(0.87)^n\cos(0.41n) + 0.194(0.87)^n\sin(0.41n)\}u[nT]$$

Plot the sampled height signal, $s[nT]$, the filter input signal, $x[nT]$, the unit-pulse response of the filter, $h[nT]$, and the filter output signal $y[nT]$, for $0 \le t \le 6\text{ s}$.

Gustav Robert Kirchhoff (1824-1887)

Kirchhoff was born in Russia, and showed an early interest in mathematics. He studied at the University of Königsberg, and in 1845, while still a student, he pronounced Kirchhoff's Laws, which allow the calculation of current and voltage for *any* circuit. They are the Laws electrical engineers apply on a routine basis – they even apply to non-linear circuits such as those containing semiconductors, or distributed parameter circuits such as microwave striplines.

He graduated from university in 1847 and received a scholarship to study in Paris, but the revolutions of 1848 intervened. Instead, he moved to Berlin where he met and formed a close friendship with Robert Bunsen, the inorganic chemist and physicist who popularized use of the "Bunsen burner".

In 1857 Kirchhoff extended the work done by the German physicist Georg Simon Ohm, by describing charge flow in three dimensions. He also analysed circuits using topology. In further studies, he offered a general theory of how electricity is conducted. He based his calculations on experimental results which determine a constant for the speed of the propagation of electric charge. Kirchhoff noted that this constant is approximately the speed of light – but the greater implications of this fact escaped him. It remained for James Clerk Maxwell to propose that light belongs to the electromagnetic spectrum.

Kirchhoff's most significant work, from 1859 to 1862, involved his close collaboration with Bunsen. Bunsen was in his laboratory, analysing various salts that impart specific colours to a flame when burned. Bunsen was using coloured glasses to view the flame. When Kirchhoff visited the laboratory, he suggested that a better analysis might be achieved by passing the light from the flame through a prism. The value of spectroscopy became immediately clear. Each element and compound showed a spectrum as unique as any fingerprint, which could be viewed, measured, recorded and compared.

Spectral analysis, Kirchhoff and Bunsen wrote not long afterward, promises "the chemical exploration of a domain which up till now has been completely closed.” They not only analysed the known elements, they discovered new ones. Analyzing salts from evaporated mineral water, Kirchhoff and Bunsen detected a blue spectral line – it belonged to an element they christened *caesium* (from the Latin *caesius*, sky blue). Studying lepidolite (a lithium-based mica) in 1861, Bunsen found an alkali metal he called rubidium (from the Latin *rubidius*, deepest red). Both of these elements are used today in atomic clocks. Using spectroscopy, ten more new elements were discovered before the end of the century, and the field had expanded enormously – between 1900 and 1912 a “handbook” of spectroscopy was published by Kayser in six volumes comprising five thousand pages!

“[Kirchhoff is] a perfect example of the true German investigator. To search after truth in its purest shape and to give utterance with almost an abstract self-forgetfulness, was the religion and purpose of his life.” – Robert von Helmholtz, 1890.

Kirchhoff’s work on spectrum analysis led on to a study of the composition of light from the Sun. He was the first to explain the dark lines (Fraunhofer lines) in the Sun's spectrum as caused by absorption of particular wavelengths as the light passes through a gas. Kirchhoff wrote “It is plausible that spectroscopy is also applicable to the solar atmosphere and the brighter fixed stars.” We can now analyse the collective light of a hundred billion stars in a remote galaxy billions of light-years away – we can tell its composition, its age, and even how fast the galaxy is receding from us – simply by looking at its spectrum!

As a consequence of his work with Fraunhofer’s lines, Kirchhoff developed a general theory of emission and radiation in terms of thermodynamics. It stated that a substances capacity to emit light is equivalent to its ability to absorb it at the same temperature. One of the problems that this new theory created was the “blackbody” problem, which was to plague physics for forty years. This fundamental quandary arose because heating a black body – such as a metal bar – causes it to give off heat and light. The spectral radiation, which depends only on the temperature and not on the material, could not be predicted by classical physics. In 1900 Max Planck solved the problem by discovering quanta, which had enormous implications for twentieth-century science.

In 1875 he was appointed to the chair of mathematical physics at Berlin and he ceased his experimental work. An accident-related disability meant he had to spend much of his life on crutches or in a wheelchair.

Lecture 2A – Fourier Series, Spectra

Orthogonality. Inner product. Orthogonality in power signals. Orthogonality in energy signals. Trigonometric Fourier series. Spectrum. Compact trigonometric Fourier series. Complex exponential Fourier series. Power. Filters.

Orthogonality

The idea of breaking a complex phenomenon down into easily comprehended components is quite fundamental to human understanding. Instead of trying to commit the totality of something to memory, and then, in turn having to think about it in its totality, we identify *characteristics*, perhaps associating a scale with each characteristic. Our memory of a person might be confined to the characteristics gender, age, height, skin colour, hair colour, weight and how they rate on a small number of personality attribute scales such as optimist-pessimist, extrovert-introvert, aggressive-submissive etc.

The concept of breaking the complex down into the simple

We only need to know how to travel east and how to travel north and we can go from any point on earth to any other point. An artist (or a television tube) needs only three primary colours to make any colour.

In choosing the components it is most efficient if they are *independent* as in gender and height for a person. It would waste memory capacity, for example, to adopt as characteristics both current age and birthdate, as one could be predicted from the other, or all three of total height and height above and below the waist.

Independent characteristics are efficient descriptors

Orthogonality in Mathematics

Vectors and functions are similarly often best represented, memorised and manipulated in terms of a set of magnitudes of independent components. Recall that any vector $\mathbf{A}$ in 3 dimensional space can be expressed in terms of any three vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ which do not lie in the same plane as:

Specifying a 3D vector in terms of 3 components

$$\mathbf{A} = A_1\mathbf{a} + A_2\mathbf{b} + A_3\mathbf{c} \tag{2A.1}$$

where A_1, A_2 and A_3 are appropriately chosen constants.

Vector in 3D space showing components

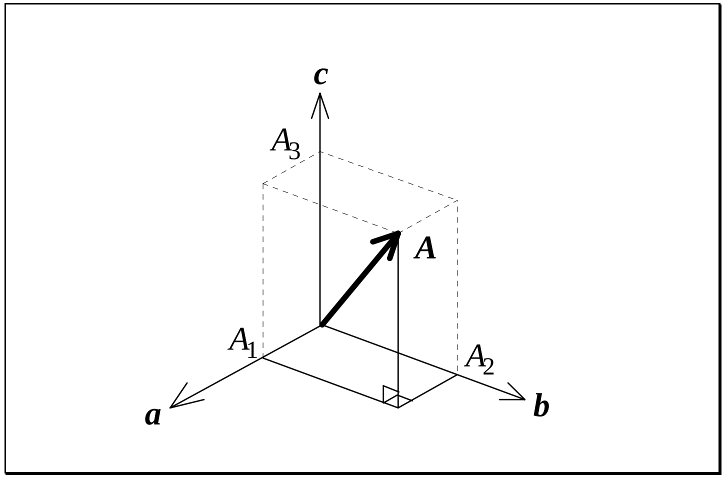


Figure 2A.1

Basis set described as set of linearly independent vectors

The vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are said to be *linearly independent* for no one of them can be expressed as a linear combination of the other two. For example, it is impossible to write $\mathbf{a}=\alpha\mathbf{b}+\beta\mathbf{c}$ no matter what choice is made for α and β (because a linear combination of $\mathbf{b}$ and $\mathbf{c}$ must stay in the plane of $\mathbf{b}$ and $\mathbf{c}$). Such a set of linearly independent vectors is said to form a *basis set* for three dimensional vector space. They *span* three dimensional space in the sense that any vector $\mathbf{A}$ can be expressed as a linear combination of them.

Orthogonality defined for vectors

If $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are mutually perpendicular they form an *orthogonal* basis set. Orthogonal means that the *projection* of one component onto another is zero. In vector analysis the projection of one vector onto another is given by the dot product. Hence if $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are orthogonal:

Finding the components of a vector

$$\mathbf{A}\cdot\mathbf{a}=A_1\mathbf{a}\cdot\mathbf{a}+A_2\mathbf{a}\cdot\mathbf{b}+A_3\mathbf{a}\cdot\mathbf{c}=A_1a^2 \tag{2A.2}$$

Hence we have an easy way of finding the components when the basis set is orthogonal - just project the vector onto each of the components in the basis set and normalise:

A vector described in terms of orthogonal components

$$\mathbf{A}=\left(\frac{\mathbf{A}\cdot\mathbf{a}}{a^2}\right)\mathbf{a}+\left(\frac{\mathbf{A}\cdot\mathbf{b}}{b^2}\right)\mathbf{b}+\left(\frac{\mathbf{A}\cdot\mathbf{c}}{c^2}\right)\mathbf{c} \tag{2A.3}$$

Orthonormal defined

If a^2, b^2 and c^2 are 1, the set of vectors are not just *orthogonal* they are *orthonormal*.

Orthogonal components is a general concept with wide application

The above description of a vector in three dimensional space is exactly analogous to resolving a colour into three primary (orthogonal) components. Suppose we had an orange coloured light. We could *project* it through red, green and blue filters and find the intensity of each of the three components. The orange original could be synthesised once again by red, green and blue lights of appropriate intensity.

Inner Product

The definition of the dot product for vectors in space can be extended to any general vector space. Consider two *n*-dimensional vectors:

Two *n*-dimensional vectors

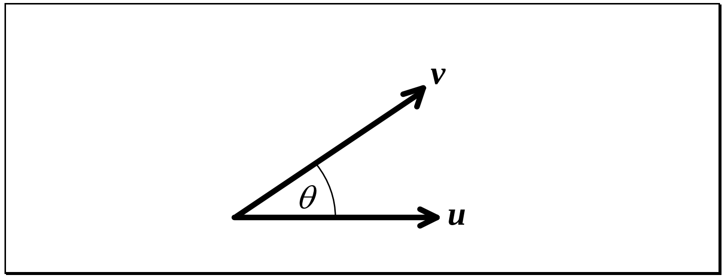


Figure 2A.2

The inner product can be written mathematically as:

Inner product for vectors defined

$$\langle\mathbf{u},\mathbf{v}\rangle=\mathbf{u}^T\mathbf{v}=\mathbf{v}^T\mathbf{u} \tag{2A.4}$$

For example, in 3 dimensions:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{2A.5}$$

$$= u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$= |\mathbf{u}||\mathbf{v}| \cos\theta$$

If $\theta = 90°$ the two vectors are orthogonal and $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. They are linearly independent, that is, one vector cannot be written in terms of the other.

Real functions such as $x(t)$ and $y(t)$ on a given interval $0 \le t \le T$ can be considered a "vector space", since they obey the same laws of addition and scalar multiplication as spatial vectors. For functions, we can define the inner product by the integral:

Inner product for functions defined

$$\boxed{\langle x, y \rangle = \int_0^T x(t) y(t) dt} \tag{2A.6}$$

This definition ensures that the inner product of functions behaves in exactly the same way as the inner product of vectors. Just like vectors, we say that "two functions are orthogonal" if their inner product is zero:

Orthogonality for functions defined

$$\boxed{\langle x, y \rangle = 0} \tag{2A.7}$$

Orthogonality in Power Signals

Consider a finite power signal, $x(t)$. The average power of $x(t)$ is:

$$P_x = \frac{1}{T} \int_0^T x^2(t) dt = \frac{1}{T} \langle x, x \rangle \tag{2A.8}$$

Now consider two finite power signals, $x(t)$ and $y(t)$. The average value of the product of the two signals observed over a particular interval, *T*, is given by the following expression:

$$\frac{1}{T} \int_0^T x(t) y(t) dt = \frac{1}{T} \langle x, y \rangle \tag{2A.9}$$

This average can also be interpreted as a measure of the correlation between $x(t)$ and $y(t)$. If the two signals are orthogonal, then over a long enough period, *T*, the average of the product tends to zero, since $\langle x, y \rangle = 0$. In this case, when the signals are added together the total power (i.e. the mean square value), is:

The power of a signal made up of orthogonal components is the sum of the component signal powers

$$\begin{aligned} P_{x+y} &= \frac{1}{T} \langle x+y, x+y \rangle \\ &= \frac{1}{T} \int_0^T [x(t) + y(t)]^2 dt \\ &= \frac{1}{T} \int_0^T [x^2(t) + 2x(t)y(t) + y^2(t)] dt \\ &= \frac{1}{T} \langle x, x \rangle + \frac{2}{T} \langle x, y \rangle + \frac{1}{T} \langle y, y \rangle \\ &= P_x + P_y \end{aligned} \tag{2A.10}$$

This means that the total power in the combined signal can be obtained by adding the power of the individual orthogonal signals.

Orthogonality in Energy Signals

Consider two finite energy signals in the form of pulses in a digital system. Two pulses, $p_1(t)$ and $p_2(t)$, are orthogonal over a time interval, T, if:

$$\int_0^T x(t)y(t)dt = \langle x, y \rangle = 0 \tag{2A.11}$$

Similar to the orthogonal finite power signals discussed above, the total energy of a pulse produced by adding together two orthogonal pulses can be obtained by summing the individual energies of the separate pulses.

The energy of a signal made up of orthogonal components is the sum of the component signal energies

$$\begin{aligned} E_{x+y} &= \langle x+y, x+y \rangle \\ &= \int_0^T [x(t)+y(t)]^2 dt \\ &= \int_0^T [x^2(t)+2x(t)y(t)+y^2(t)]dt \\ &= \langle x,x \rangle + 2\langle x,y \rangle + \langle y,y \rangle \\ &= E_x + E_y \end{aligned} \tag{2A.12}$$

For example, Figure 2A.3, illustrates two orthogonal pulses because they occupy two completely separate portions of the time interval 0 to T. Therefore, their product is zero over the time period of interest which means that they are orthogonal.

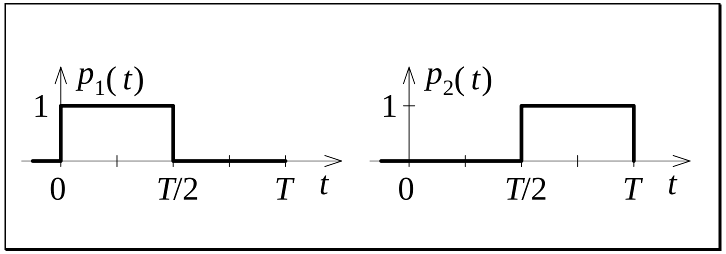


Figure 2A.3

The Fourier Series

The Fourier series is one way of representing periodic functions as a weighted sum of orthogonal *functions*. The equivalent of the dot product (or the light filter) for obtaining a projection in this case is the inner product given by:

Definition of "inner product" for a function

$$\langle g(t), \phi_n(t) \rangle = \int_{-T_0/2}^{T_0/2} g(t)\phi_n(t)dt \tag{2A.13}$$

This is the "projection" of $g(t)$ onto $\phi_n(t)$, the nth member of the orthogonal basis set. The equivalent relationships hold between orthogonal functions as do between orthogonal vectors:

The "projection" of a function onto itself gives a number

$$\int_{-T_0/2}^{T_0/2} \phi_n(t)\phi_m(t)dt = c_n^2 \qquad (\mathbf{a}\cdot\mathbf{a} = a^2) \qquad n=m \tag{2A.14}$$

and:

The "projection" of a function onto an orthogonal function gives zero

$$\int_{-T_0/2}^{T_0/2} \phi_n(t)\phi_m(t)dt = 0 \qquad (\mathbf{a}\cdot\mathbf{b} = 0) \qquad n\neq m \tag{2A.15}$$

When $c_n^2 = 1$ the basis set of functions $\phi_n(t)$ (all n) are said to be orthonormal.

There are many possible orthogonal basis sets for representing a function over an interval of time T_0, for example the infinite set of Walsh functions shown below:

Example of an orthogonal basis set – the Walsh functions

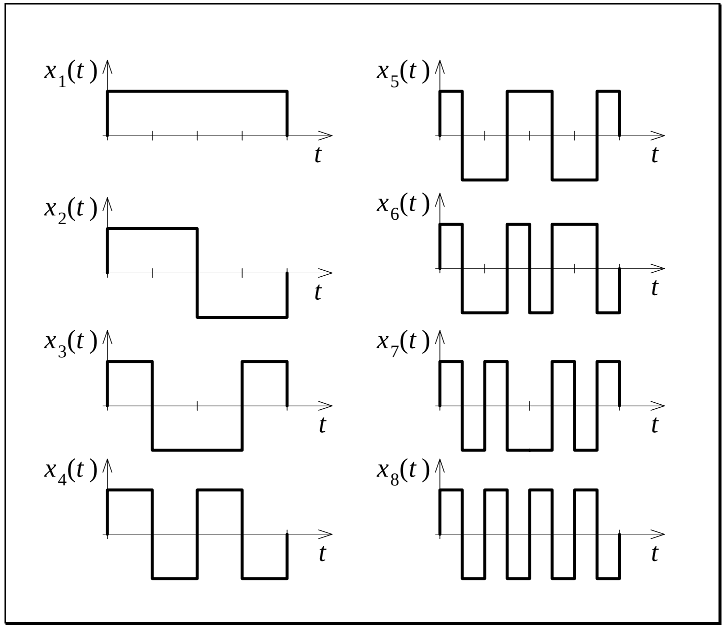


Figure 2A.4

The basis set chosen for the Fourier series is the set of pairs of sines and cosines with frequencies nf_0. They were chosen for the two reasons outlined in Lecture 1A - for linear systems a sinusoidal input yields a sinusoidal output, and they have a compact notation using complex numbers. The constant c_n^2 in this case is either $T_0/2$ or T_0, as can be seen from the following relations:

The orthogonal functions for the Fourier series are sinusoids

$$\int_{-T_0/2}^{T_0/2} \cos(2\pi m f_0 t)\cos(2\pi n f_0 t)dt = \begin{cases} T_0 & m=n=0 \\ T_0/2 & m=n\neq 0 \\ 0 & m\neq n \end{cases} \tag{2A.16a}$$

$$\int_{-T_0/2}^{T_0/2} \sin(2\pi m f_0 t)\sin(2\pi n f_0 t)dt = \begin{cases} 0 & m=n=0 \\ T_0/2 & m=n\neq 0 \\ 0 & m\neq n \end{cases} \tag{2A.16b}$$

$$\int_{-T_0/2}^{T_0/2} \cos(2\pi m f_0 t)\sin(2\pi n f_0 t)dt = 0 \qquad \text{all } m,\ n \tag{2A.16c}$$

In an analogous way to the representation of a vector given in Eq. (2A.1), we let any periodic function $g(t)$ with period T_0 be expressed as a sum of orthogonal components:

The trigonometric Fourier series defined

$$g(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi n f_0 t) + b_n \sin(2\pi n f_0 t) \tag{2A.17}$$

The frequency $f_0 = 1/T_0$ is the fundamental frequency and the frequency nf_0 is the nth harmonic frequency. The right-hand side of Eq. (2A.17) is known as a Fourier series, with a_n and b_n known as Fourier series coefficients.

Now look back at Eqs. (2A.3) and (2A.14). If we want to determine the coefficients in the Fourier series, all we have to do is "project" the function onto each of the components of the basis set and normalise by dividing by c_n^2:

How to find the trigonometric Fourier series coefficients

$$a_0 = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} g(t)dt \tag{2A.18a}$$

$$a_n = \frac{2}{T_0}\int_{-T_0/2}^{T_0/2} g(t)\cos(2\pi n f_0 t)dt \tag{2A.18b}$$

$$b_n = \frac{2}{T_0}\int_{-T_0/2}^{T_0/2} g(t)\sin(2\pi n f_0 t)dt \tag{2A.18c}$$

Compare these equations with Eq. (2A.2). These equations tell us how to "filter out" one particular component of the Fourier series. Note that frequency 0 is DC, and the coefficient a_0 represents the average, or DC part of the periodic signal $g(t)$.

Example

Find the Fourier series for the rectangular pulse train $g(t)$ shown below:

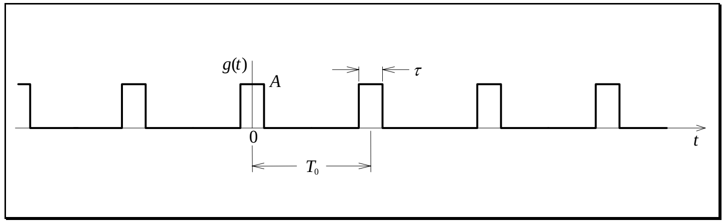


Figure 2A.5

Here the period is T_0 and $f_0 = 1/T_0$. Using Eqs. (2A.18), we have for the Fourier series coefficients:

$$a_0 = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} g(t)dt = \frac{1}{T_0}\int_{-\tau/2}^{\tau/2} A dt = \frac{A\tau}{T_0} = A\tau f_0 \tag{2A.19a}$$

$$a_n = \frac{2}{T_0}\int_{-\tau/2}^{\tau/2} A\cos(2\pi n f_0 t)dt = \frac{2A}{\pi n}\sin(n\pi f_0 \tau) = 2A\tau f_0 \operatorname{sinc}(n f_0 \tau) \tag{2A.19b}$$

$$b_n = \frac{2}{T_0}\int_{-\tau/2}^{\tau/2} A\sin(2\pi n f_0 t)dt = 0 \tag{2A.19c}$$

We can therefore say:

$$\sum_{n=-\infty}^{\infty} A\operatorname{rect}\left(\frac{t-nT_0}{\tau}\right) = A\tau f_0 + \sum_{n=1}^{\infty} 2A\tau f_0 \operatorname{sinc}(n f_0 \tau)\cos(2\pi n f_0 t) \tag{2A.20}$$

This expression is quite unwieldy, and what is its use?

Consider the case where $\tau = T_0/5$. We could draw up a table of the Fourier series coefficients as a function of n:

The trigonometric Fourier series coefficients can be tabled

n	a_n	b_n
0	0.2*A*	0
1	0.3742*A*	0
2	0.3027*A*	0
3	0.2018*A*	0
4	0.0935*A*	0
5	0	0
6	-0.0624*A*	0
etc.	-0.0865*A*	0

An even better way of representing the Fourier series coefficients is to graph them:

but a graph of the trigonometric Fourier series coefficients is better

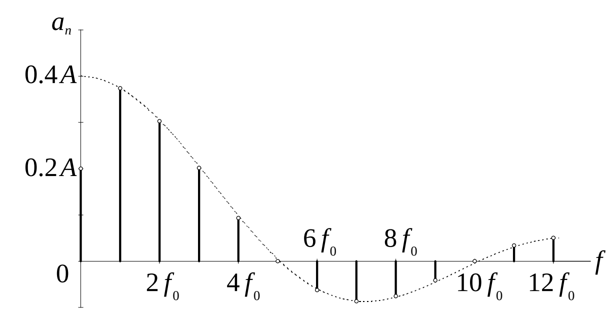


Figure 2A.6

The Spectrum

The spectrum defined

The information about the amplitudes of the orthogonal components in the Fourier series is called the spectrum. There are many ways of graphically representing the spectrum.

One simple way is to graph the coefficients in the Fourier series as a function of frequency, as in Figure 2A.6, but there are “better” ways.

The Compact Trigonometric Fourier Series

The trigonometric Fourier series in Eq. (2A.17) can be written in a more compact and meaningful way as follows:

The compact trigonometric Fourier series defined

$$\boxed{g(t) = \sum_{n=0}^{\infty} A_n \cos(2\pi n f_0 t + \phi_n)} \qquad (2A.21)$$

By expanding $A_n \cos(2\pi n f_0 t + \phi_n)$ it is easy to show that:

$$\boxed{\begin{aligned} a_n &= A_n \cos\phi_n \\ b_n &= -A_n \sin\phi_n \end{aligned}} \qquad (2A.22)$$

and therefore:

and associated constants

$$A_n = \sqrt{a_n^2 + b_n^2} \qquad (2A.23a)$$

$$\phi_n = -\tan^{-1}\left(\frac{b_n}{a_n}\right) \qquad (2A.23b)$$

From the compact Fourier series it follows that $g(t)$ consists of sinusoidal signals of frequencies 0, f_0, $2f_0$, …, nf_0, …. The nth harmonic, $A_n \cos(2\pi n f_0 t + \phi_n)$, has amplitude A_n and phase ϕ_n. We can plot the magnitude spectrum $(A_n \text{ vs. } f)$ and the phase spectrum $(\phi_n \text{ vs. } f)$ for a given periodic signal.

Notice how we can store the amplitude and phase information *at a particular frequency* using our phasor notation of Lecture 1A. We can *represent* the *n*th harmonic by the phasor:

Harmonic phasors defined

$$\overline{G}_n = A_n e^{j\phi_n} = A_n \cos\phi_n + jA_n \sin\phi_n \tag{2A.24}$$

$\overline{G}_n$ can be derived directly from the Fourier series coefficients using Eq. (2A.22):

and related to the trigonometric Fourier series coefficients

$$\overline{G}_n = a_n - jb_n \tag{2A.25}$$

Substituting for a_n and b_n from Eqs. (2A.18a)-(2A.18c) results in:

$$\begin{aligned}\overline{G}_0 &= a_0 = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} g(t)dt \\ \overline{G}_n &= a_n - jb_n \\ &= \frac{2}{T_0}\int_{-T_0/2}^{T_0/2} g(t)\cos(2\pi nf_0t)dt - j\frac{2}{T_0}\int_{-T_0/2}^{T_0/2} g(t)\sin(2\pi nf_0t)dt \\ &= \frac{2}{T_0}\int_{-T_0/2}^{T_0/2} g(t)[\cos(2\pi nf_0t) - j\sin(2\pi nf_0t)]dt\end{aligned} \tag{2A.26}$$

which can be simplified using Euler's identity to give:

Obtaining the harmonic phasors directly

$$\overline{G}_n = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} g(t)e^{-j2\pi nf_0t}dt \quad n = 0 \tag{2A.27a}$$

$$\overline{G}_n = \frac{2}{T_0}\int_{-T_0/2}^{T_0/2} g(t)e^{-j2\pi nf_0t}dt \quad n \neq 0 \tag{2A.27b}$$

Note that the negative sign in $\overline{G}_n = a_n - jb_n$ comes from the fact that in the phasor representation of a sinusoid the real part of $\overline{G}_n$ is the amplitude of the *cos* component and the imaginary part of $\overline{G}_n$ is the amplitude of the *-sin* component.

The expression for the compact Fourier series, as in Eq. (2A.21), can now be written as:

Fourier series expressed as a sum of harmonic phasors projected onto the real axis

$$g(t) = \mathrm{Re}\left(\sum_{n=0}^{\infty} \overline{G}_n e^{j2\pi nf_0t}\right) \tag{2A.28}$$

Each term in the sum is a phasor rotating at a different speed. $g(t)$ is the projection of the instantaneous "vector" sum of these phasors.

We can now think of a spectrum as a graph of the *phasor value* as a function of frequency. For this representation, we still need a magnitude spectrum and a phase spectrum. We call this representation "a single sided" spectrum.

A spectrum is a graph of phasor values vs. frequency

Example

A single frequency sinusoid $A\cos(2\pi f_0t + \phi)$ has single sided magnitude and phase spectra:

The single-sided spectrum of a sinusoid

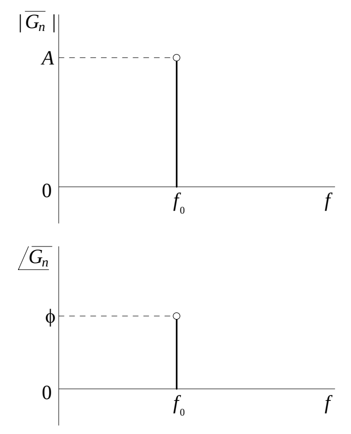


Figure 2A.7

Example

Find the compact Fourier series coefficients for the rectangular pulse train $g(t)$ shown below:

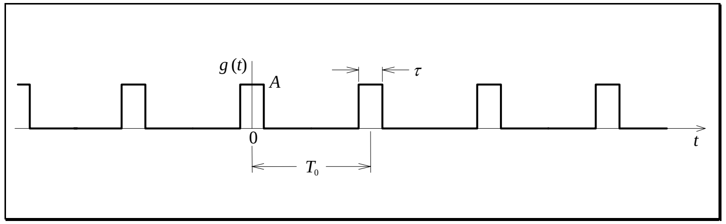


Figure 2A.8

Again, the period is T_0 and $f_0 = 1/T_0$. Using Eqs. (2A.27), we have for the compact Fourier series coefficients (harmonic phasors):

$$\overline{G}_0 = \frac{1}{T_0}\int_{-\tau/2}^{\tau/2} A\,dt = Af_0\tau \tag{2A.29a}$$

$$\begin{aligned}
\overline{G}_n &= \frac{2}{T_0}\int_{-\tau/2}^{\tau/2} Ae^{-j2\pi nf_0 t}dt \\
&= 2Af_0\left[\frac{-1}{j2\pi nf_0}e^{-j2\pi nf_0 t}\right]_{-\tau/2}^{\tau/2} \\
&= \frac{-2A}{j2\pi n}\left[e^{-j\pi nf_0\tau} - e^{j\pi nf_0\tau}\right] \\
&= \frac{2A}{\pi n}\sin(\pi nf_0\tau) = 2Af_0\tau\left[\frac{\sin(\pi nf_0\tau)}{\pi nf_0\tau}\right] \\
&= 2Af_0\tau\,\mathrm{sinc}(nf_0\tau)
\end{aligned} \tag{2A.29b}$$

A real number for a harmonic phasor just means the phase is zero

This is the same result as Eqs. (2A.19) because $\overline{G}_n$ is real, which just means the phase is zero. In general we will have a complex number as in Eq. (2A.24).

Symmetry in the Time and Frequency Domains

Symmetry in the Fourier series

We obtained a real number in the previous example because the function $g(t)$ was an even function. An even function is one which possesses symmetry about the $g(t)$ axis. Mathematically, it means $g(t) = g(-t)$. It can be expressed as a sum of cosine waves only (cosines are even functions). If we had an odd function, we would get purely imaginary $\overline{G}_n$, which means it can be expressed as a sum of sine waves (sine waves are odd functions). Mathematically, an odd function is such that $g(t) = -g(-t)$. Note also that *our* choice of the arbitrary origin determines whether a periodic function (extending from $-\infty$ to ∞ by definition) is even or odd – we should choose the origin to make the mathematics easy.

For the case of $\tau = T_0/5$, the single sided magnitude and phase spectra associated with Eqs. (2A.29) are:

Single-sided spectrum for a rectangular pulse train

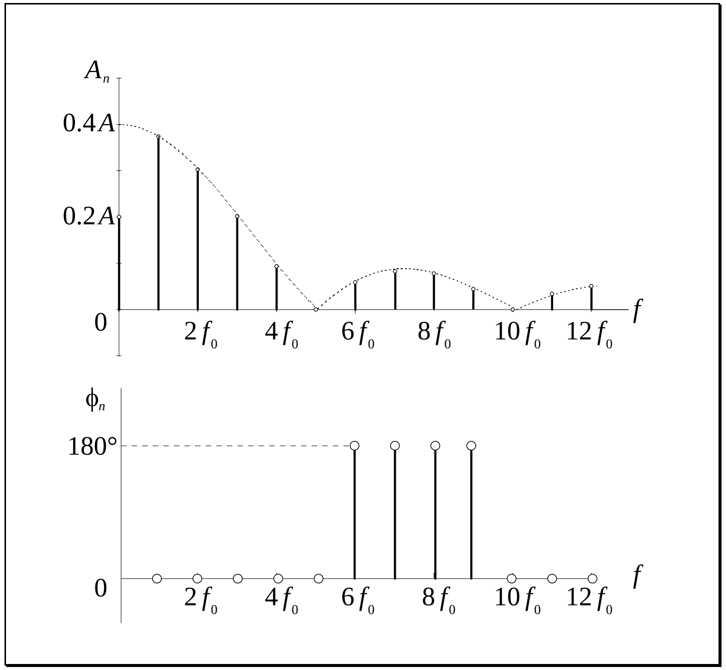


Figure 2A.9

Since the case of real $\overline{G}_n$ occurs so frequently, we often incorporate the phase information in the "magnitude" spectrum for convenience:

Single-sided magnitude spectrum with phase information for an even rectangular pulse train

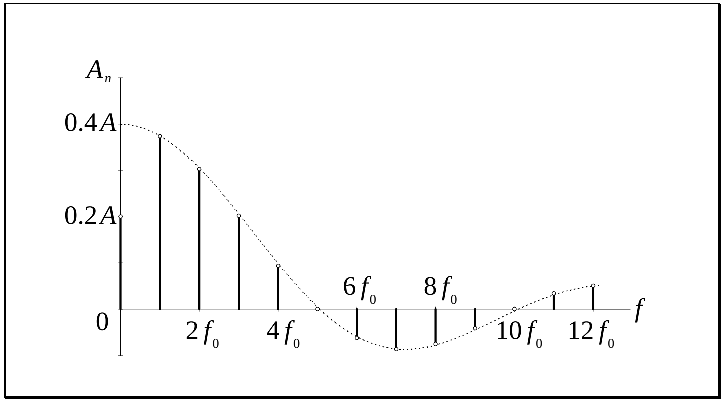


Figure 2A.10

A square pulse train (50% duty cycle square wave) is a special case of the rectangular pulse train, when $\tau = T_0/2$. For this case $g(t)$ is:

A 50% duty cycle square wave

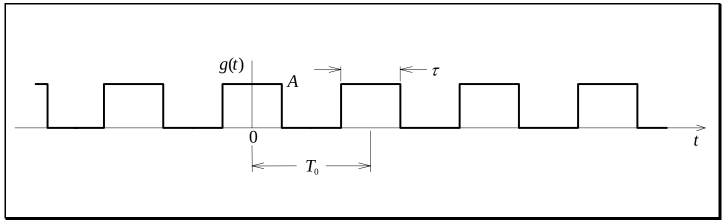


Figure 2A.11

The magnitude spectrum is:

and its magnitude spectrum

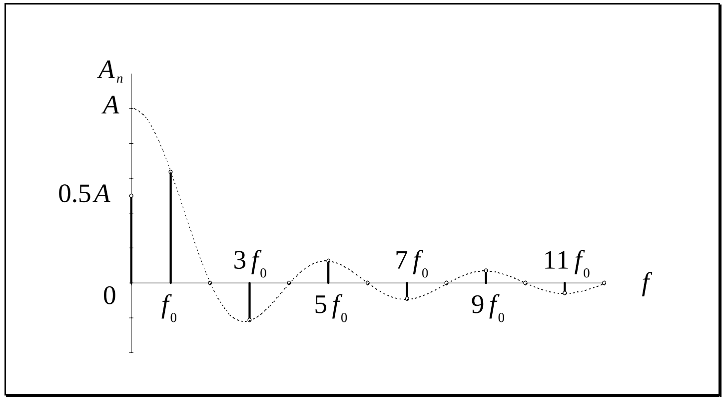


Figure 2A.12

Notice how the width of the rectangular pulse determines the sinc function envelope. The spectrum is still only non-zero at discrete frequencies spaced f_0 Hz apart. In this case the envelope happens to make all the amplitudes of the even harmonics zero. This is due to the square wave possessing *half-wave* symmetry. Mathematically, this is expressed as $g(t) = -g\left(t - T_0/2\right)$ if we ignore any DC component. Half-wave symmetry is not dependent on our choice of origin. An example of a half-wave symmetric function is:

Half-wave symmetry implies no even harmonics

A half-wave symmetric waveform

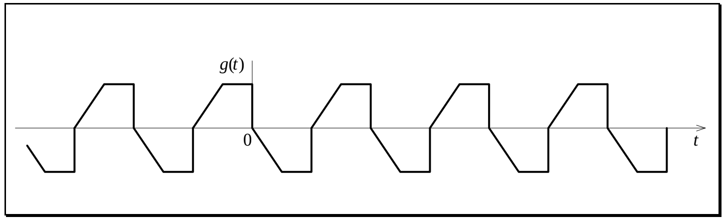


Figure 2A.13

Any half-wave symmetric function will have *only* odd harmonics.

The Complex Exponential Fourier Series

The complex exponential Fourier series is the most mathematically convenient and useful representation of a periodic signal. Recall that Euler's formulas relating the complex exponential to cosines and sines are:

$$e^{j\theta} = \cos\theta + j\sin\theta \tag{2A.30a}$$

$$\cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2} \tag{2A.30b}$$

$$\sin\theta = \frac{e^{j\theta} - e^{-j\theta}}{j2} \tag{2A.30c}$$

Substitution of Eqs. (2A.30b) and (2A.30c) into the trigonometric Fourier series, Eq. (2A.17), gives:

$$\begin{aligned} g(t) &= a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi n f_0 t) + b_n \sin(2\pi n f_0 t) \\ &= a_0 + \sum_{n=1}^{\infty} \left(\frac{a_n}{2} e^{j2\pi n f_0 t} + \frac{a_n}{2} e^{-j2\pi n f_0 t} \right. \\ &\qquad \left. - j\frac{b_n}{2} e^{j2\pi n f_0 t} + j\frac{b_n}{2} e^{-j2\pi n f_0 t} \right) \\ &= a_0 + \sum_{n=1}^{\infty} \left(\left[\frac{a_n}{2} - j\frac{b_n}{2} \right] e^{j2\pi n f_0 t} \right. \\ &\qquad \left. + \left[\frac{a_n}{2} + j\frac{b_n}{2} \right] e^{-j2\pi n f_0 t} \right) \end{aligned} \tag{2A.31}$$

This can be rewritten in the form:

The complex exponential Fourier series

$$\boxed{g(t) = \sum_{n=-\infty}^{\infty} G_n e^{j2\pi n f_0 t}} \tag{2A.32}$$

where:

The relationship between complex exponential and trigonometric Fourier series coefficients

$$G_0 = a_0 = A_0 \tag{2A.33a}$$

$$\left.\begin{aligned} G_n &= \frac{a_n - jb_n}{2} = \frac{A_n}{2} e^{j\phi_n} \qquad (2A.33b) \\ G_{-n} &= \frac{a_n + jb_n}{2} = \frac{A_n}{2} e^{-j\phi_n} \qquad (2A.33c) \end{aligned}\right\} n \ge 1$$

From Eqs. (2A.27) and (2A.33), we can see that an alternative way of writing the Fourier series coefficients, in *one* neat formula instead of *three*, is:

The complex exponential Fourier series coefficients defined

$$\boxed{G_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g(t) e^{-j2\pi n f_0 t} dt} \tag{2A.34}$$

Thus, the trigonometric and complex exponential Fourier series are not two different series but represent two different *ways of writing* the same series. The coefficients of one series can be obtained from those of the other.

Harmonic phasors can also have negative frequency

The complex exponential Fourier series can also be viewed as being based on the compact Fourier series but uses the fact that we can use the alternative phasor definition G_n instead of $\overline{G}_n$. In this case we remember that $G_n = \overline{G}_n / 2$ and that for every forward rotating phasor G_n there is a corresponding backward rotating phasor G_n^*. Eqs. (2A.28) and (2A.27) then become:

The complex exponential Fourier series and coefficients defined

$$g(t) = \sum_{n=-\infty}^{\infty} G_n e^{j2\pi n f_0 t} \tag{2A.35a}$$

$$G_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g(t) e^{-j2\pi n f_0 t} dt \tag{2A.35b}$$

where:

The symmetry of the complex exponential Fourier series coefficients

$$G_{-n} = G_n^* \tag{2A.36}$$

Thus, if:

$$G_n = |G_n| e^{j\phi_n} \tag{2A.37}$$

then:

$$G_{-n} = |G_n| e^{-j\phi_n} \tag{2A.38}$$

A double-sided spectrum shows negative frequency phasors

$|G_n|$ is the magnitude and ϕ_n is the phase of G_n. For a real $g(t)$, $|G_{-n}| = |G_n|$, and the *double-sided* magnitude spectrum $|G_n|$ vs. f is an even function of f. Similarly, the phase spectrum ϕ_n vs. f is an odd function of f because $\phi_{-n} = -\phi_n$.

Example

Find the complex exponential Fourier series for the rectangular pulse train $g(t)$ shown below:

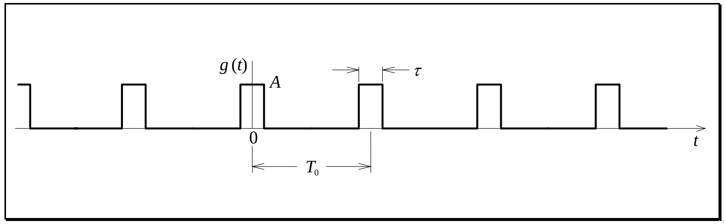


Figure 2A.14

The period is T_0 and $f_0 = 1/T_0$. Using Eq. (2A.35b), we have for the complex exponential Fourier series coefficients:

$$\begin{aligned} G_n &= \frac{1}{T_0} \int_{-\tau/2}^{\tau/2} A e^{-j2\pi n f_0 t} dt \\ &= \frac{-A}{j2\pi n} \left[e^{-j\pi n f_0 \tau} - e^{j\pi n f_0 \tau} \right] \\ &= \frac{A}{\pi n} \sin(\pi n f_0 \tau) \\ &= A f_0 \tau \operatorname{sinc}(n f_0 \tau) \end{aligned} \tag{2A.39}$$

For the case of $\tau = T_0/5$, the double-sided magnitude spectrum is then:

The double-sided magnitude spectrum of a rectangular pulse train

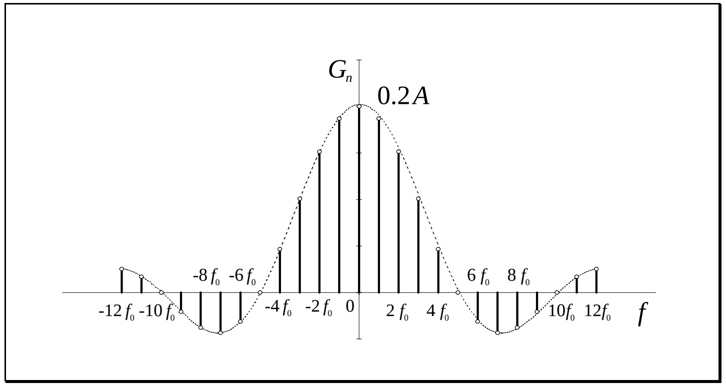


Figure 2A.15

For the case of $\tau = T_0/2$, the spectrum is:

and for a 50% duty cycle square wave

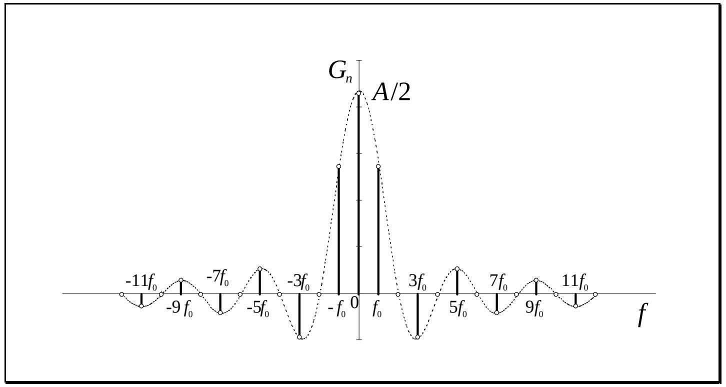


Figure 2A.16

Thus, the Fourier series can be represented by spectral lines at all harmonics of $f_0 = 1/T_0$, where each line varies according to the complex quantity G_n. In particular, for a rectangular pulse train, G_n follows the envelope of a sinc function, with amplitude $A\tau/T_0$ and zero crossings at integer multiples of $1/\tau$.

Another case of interest, which is fundamental to the analysis of digital systems, is when we allow each pulse in a rectangular pulse train to turn into an impulse, i.e. $\tau \to 0$ and $A \to \infty$ such that $A\tau = 1$. In this case, each pulse in Figure 2A.14 becomes an impulse of unit strength, and $g(t)$ is simply a uniform train of unit impulses, as shown below:

A uniform train of impulses,

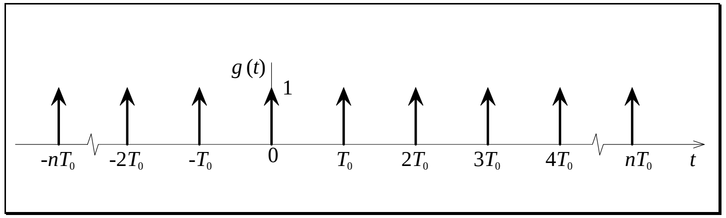


Figure 2A.17

The result of Eq. (2A.39) is still valid if we take the appropriate limit:

$$G_n = \lim_{\tau \to 0} A f_0 \tau \operatorname{sinc}(n f_0 \tau) = f_0 \tag{2A.40}$$

Thus, with $A\tau = 1$, the amplitude of the sinc function envelope is $Af_0\tau = f_0$, and when $\tau \to 0$, $\operatorname{sinc}(nf_0\tau) = \operatorname{sinc}(0) = 1$. Therefore:

its Fourier series,

$$\boxed{\begin{aligned} g(t) &= \sum_{n=-\infty}^{\infty} G_n e^{j2\pi n f_0 t} \\ \sum_{n=-\infty}^{\infty} \delta(t - nT_0) &= f_0 \sum_{n=-\infty}^{\infty} e^{j2\pi n f_0 t} \end{aligned}} \tag{2A.41}$$

The spectrum has components of frequencies nf_0, n varying from $-\infty$ to ∞, including 0, all with an equal strength of f_0, as shown below:

and its spectrum

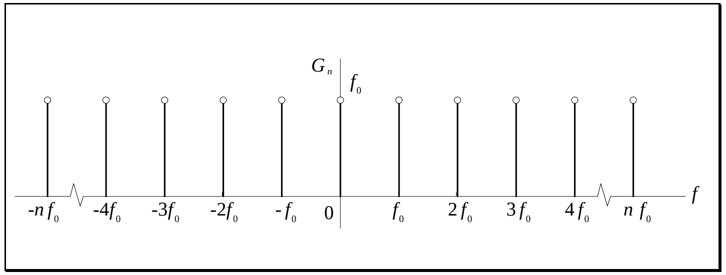


Figure 2A.18

The complex exponential Fourier series is so convenient we will use it almost exclusively. Therefore, when we refer to the spectrum of a signal, we are referring to its double-sided spectrum. It turns out that the double-sided spectrum is the easiest to use when describing signal operations in systems. It also enables us to calculate the average power of a signal in an easy manner.

Power

Calculating power using orthogonal components

One of the advantages of representing a signal in terms of a set of orthogonal components is that it is very easy to calculate its average power. Because the components are orthogonal, the total average power is just the sum of the average powers of the orthogonal components.

For example, if the double-sided spectrum is being used, since the magnitude of the sinusoid represented by the phasor G_n is $2|G_n|$, and the average power of a sinusoid of amplitude A is $P = A^2/2$, the total power in the signal $g(t)$ is:

Power for a double-sided spectrum

$$\boxed{P = \sum_{n=-\infty}^{\infty} |G_n|^2 = \sum_{n=-\infty}^{\infty} G_n G_n^*} \tag{2A.42}$$

Note that the DC component only appears once in the sum ($n=0$). Its power contribution is G_0^2 which is correct.

Example

How much of the power of a 50% duty cycle rectangular pulse train is contained in the first three harmonics?

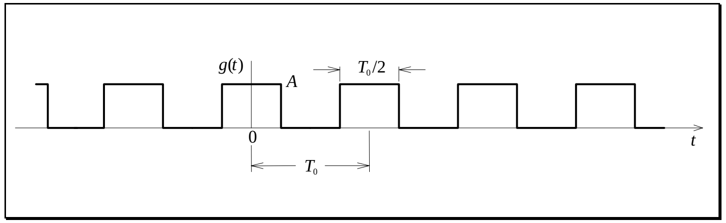


Figure 2A.19

We first find the total power in the time-domain:

$$P = \frac{1}{T_0}\int_{-T_0/4}^{T_0/4} A^2 dt = \frac{A^2}{2} \tag{2A.43}$$

To find the power in each harmonic, we work in the frequency domain.

Draw the double-sided spectrum:

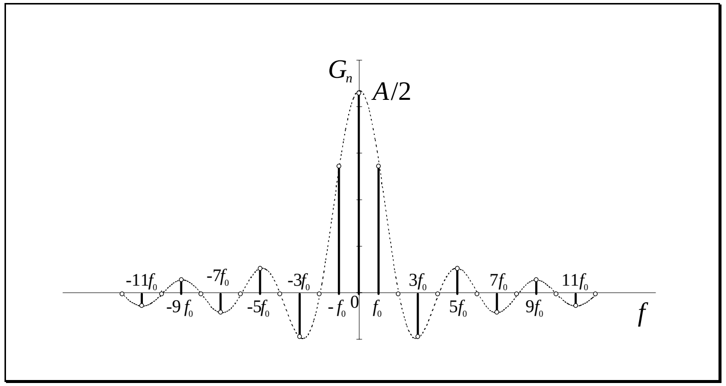


Figure 2A.20

We have for the DC power contribution:

$$P_0 = \frac{A^2}{4} \tag{2A.44}$$

which is 50% of the total power. The DC plus fundamental power is:

$$P_0 + P_1 = \frac{A^2}{4} + \frac{2A^2}{\pi^2} = 0.4526A^2 \tag{2A.45}$$

which is 90.5% of the total power. The DC, fundamental and 3$^{\text{rd}}$ harmonic power is:

$$P_0 + P_1 + P_3 = \frac{A^2}{4} + \frac{2A^2}{\pi^2} + \frac{2A^2}{9\pi^2} = 0.4752A^2 \tag{2A.46}$$

which is 95% of the total power. Thus, the spectrum makes obvious a characteristic of a periodic signal that is not obvious in the time domain. In this case, it was surprising to learn that 95% of the power in a square wave is contained in the frequency components up to the 3$^{\text{rd}}$ harmonic. This is important – we may wish to lowpass filter this signal for some reason, but retain most of its power. We are now in a position to give the cutoff frequency of a lowpass filter to retain any amount of power that we desire.

Filters

A filter acts to change each component sinusoid of a periodic function

Filters are devices that shape the input signal's spectrum to produce a new output spectrum. They shape the input spectrum by changing the amplitude and phase of each component sinusoid. This frequency-domain view of filters has been with us implicitly – we specify the filter in terms of a transfer function $H(s)$. When evaluated at $s = j\omega$ the transfer function is a complex number. The magnitude of this complex number, $|H(j\omega)|$, *multiplies* the corresponding *magnitude* of the component phasor of the input signal. The phase of the complex number, $\angle H(j\omega)$, *adds* to the *phase* of the component phasor.

Visual view of a filter's operation

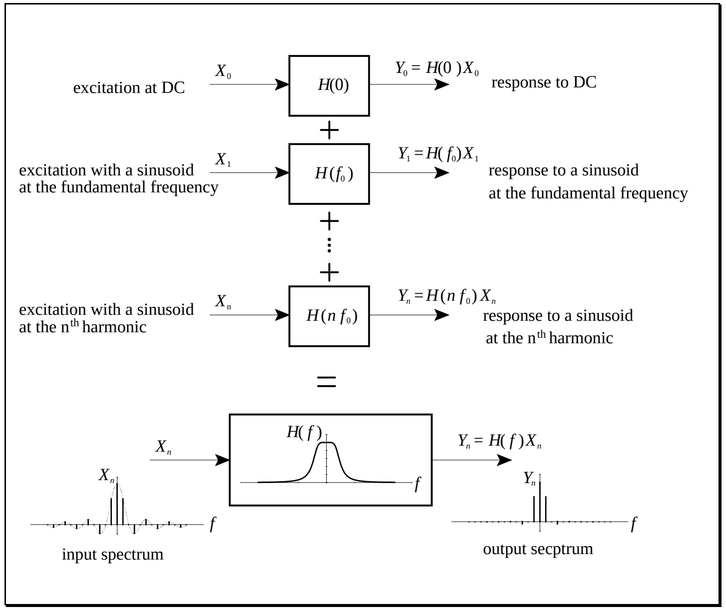


Figure 2A.21

Recall from Lecture 1A that it is the sinusoid that possesses the special property with a linear system of "a *sinusoid* in gives a *sinusoid* out". We now have a view that "a *sum of sinusoids* in gives a *sum of sinusoids* out", or more simply: "a *spectrum* in gives a *spectrum* out". The input spectrum is changed by the *frequency response* of the system to give the output spectrum.

This view of filters operating on individual components of an input signal has been implicit in the characterisation of systems via the frequency response. Experimentally, we determine the frequency response of a system by performing the operations in the top half of Figure 2A.21. That is, we apply different sinusoids (including DC which can be thought of as a sinusoid of zero frequency) to a system and measure the resulting amplitude change and phase shift of the output sinusoid. We then build up a picture of $H(f)$ by plotting the experimentally derived points on a graph (if log scales are chosen then we have a Bode plot). After obtaining the frequency response, we should be able to tell what happens when we apply any periodic signal, as shown in the bottom half of Figure 2A.21. The next example illustrates this process.

Example

Let's see what happens to a square wave when it is "passed through" a 3rd order Butterworth filter. For the filter, we have:

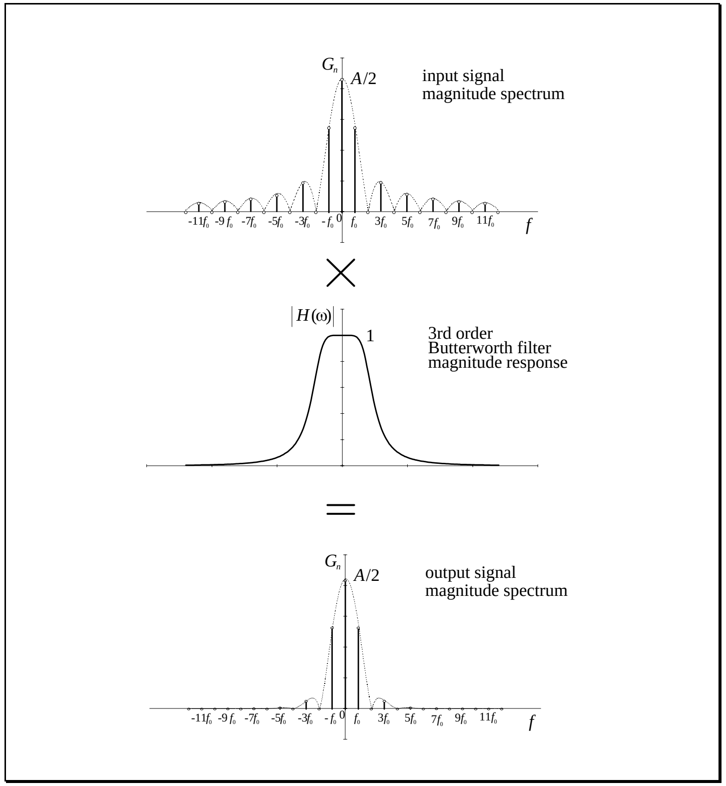
A filter is defined in terms of its magnitude and phase response

$$\left|H(j\omega)\right| = \frac{1}{\sqrt{1+(\omega/\omega_0)^6}} \tag{2A.47a}$$

$$\angle H(j\omega) = \tan^{-1}\left(\frac{2\omega_0^2\omega - \omega^3}{2\omega_0\omega^2 - \omega_0^3}\right) \tag{2A.47b}$$

Here, ω_0 is the "cutoff" frequency of the filter. It does not represent the angular frequency of the fundamental component of the input signal – filters know nothing about the signals to be applied to them. Since the filter is linear, superposition applies. For each component phasor of the input signal, we multiply by $\left|H(j\omega)\right|e^{j\angle H(j\omega)}$. We then reconstruct the signal by adding up the "filtered" component phasors.

This is an operation best performed and thought about graphically. For the case of the filter cutoff frequency set at twice the input signal's fundamental frequency we have for the output magnitude spectrum:



Filter output magnitude spectrum obtained graphically using the input signal's magnitude spectrum and the filter's magnitude response

Figure 2A.22

We could perform a similar operation for the phase spectrum.

If we now take the output spectrum and reconstruct the time-domain waveform it represents, we get:

The output signal in the time-domain obtained from the output spectrum

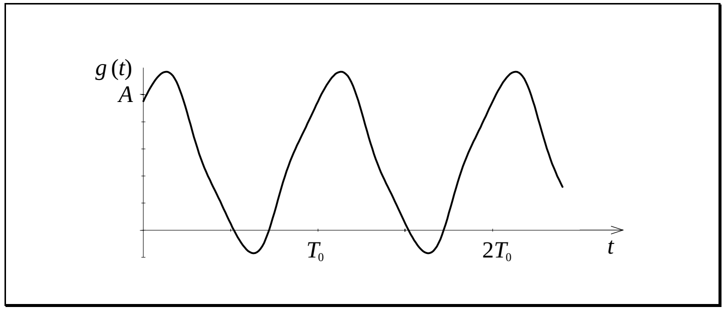


Figure 2A.23

Some features to look for in a spectrum

This looks like a shifted sinusoid (DC + sine wave) with "a touch of 3rd harmonic distortion". With practice, you are able to recognise the components of waveforms and hence relate them to their magnitude spectrum as in Figure 2A.22. How do we know it is 3rd harmonic distortion? Without the DC component the waveform exhibits half-wave symmetry, so we know the 2nd harmonic (an even harmonic) is zero.

If we extend the cutoff frequency to ten times the fundamental frequency of the input square wave, the filter has less effect:

Sharp transitions in the time-domain are caused by high frequencies

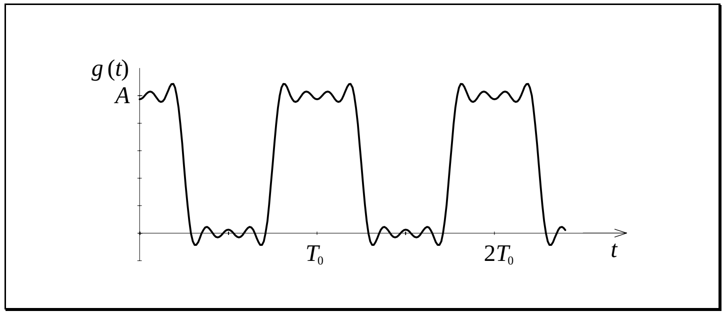
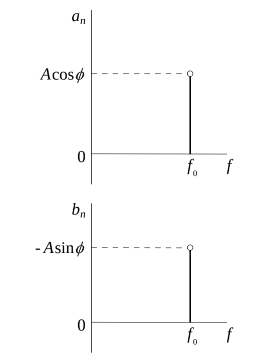
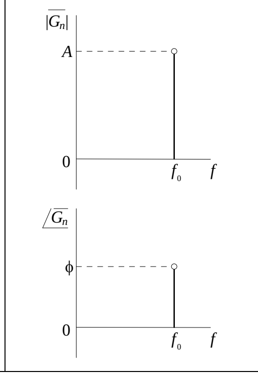
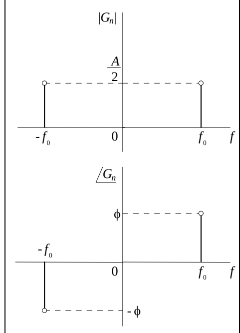


Figure 2A.24

From this example it should be apparent that high frequencies are needed to make sharp transitions in the time-domain.

Relationships Between the Three Fourier Series Representations

The table below shows the relationships between the three different representations of the Fourier series:

	Trigonometric	Compact Trigonometric	Complex Exponential
Fourier Series	$g(t)=a_0+\sum_{n=1}^{\infty}a_n\cos(2\pi n f_0 t) + b_n\sin(2\pi n f_0 t)$	$g(t)=\sum_{n=0}^{\infty}A_n\cos(2\pi n f_0 t+\phi_n) = \mathrm{Re}\left(\sum_{n=0}^{\infty}\overline{G}_n e^{j2\pi n f_0 t}\right)$	$g(t)=\sum_{n=-\infty}^{\infty}G_n e^{j2\pi n f_0 t}$
Fourier Series Coefficients	$a_0=\frac{1}{T_0}\int_{-T_0/2}^{T_0/2}g(t)dt$ $a_n=\frac{2}{T_0}\int_{-T_0/2}^{T_0/2}g(t)\cos(2\pi n f_0 t)dt$ $b_n=\frac{2}{T_0}\int_{-T_0/2}^{T_0/2}g(t)\sin(2\pi n f_0 t)dt$	$\overline{G}_n=\frac{1}{T_0}\int_{-T_0/2}^{T_0/2}g(t)e^{-j2\pi n f_0 t}dt$ $(n=0)$ $\overline{G}_n=\frac{2}{T_0}\int_{-T_0/2}^{T_0/2}g(t)e^{-j2\pi n f_0 t}dt$ $(n\neq 0)$	$G_n=\frac{1}{T_0}\int_{-T_0/2}^{T_0/2}g(t)e^{-j2\pi n f_0 t}dt$
Spectrum of a single sinusoid $A\cos(2\pi f_0 t+\phi)$			

Summary

- All periodic waveforms are made up of a sum of sinusoids – a Fourier series. There are three equivalent notations for the Fourier series.
- The trigonometric Fourier series expresses a periodic signal as a DC (or average) term and a sum of harmonically related cosinusoids and sinusoids.
- The compact trigonometric Fourier series expresses a periodic signal as a sum of cosinusoids of varying amplitude and phase.
- The complex exponential Fourier series expresses a periodic signal as a sum of counter-rotating harmonic phasors ("sum of harmonic phasors").
- The coefficients of the basis functions in the Fourier series are called "Fourier series coefficients". For the complex exponential Fourier series, they are complex numbers, and are just the phasor representation of the sinusoid at that particular harmonic frequency.
- Fourier series coefficients can be found for any periodic waveform by taking "projections" of the periodic waveform onto each of the orthogonal basis functions making up the Fourier series.
- A spectrum of a periodic waveform is a graph of the amplitudes and phases of constituent basis functions.
- The most convenient spectrum is the double-sided spectrum. It is graph of the complex exponential Fourier series coefficients. We usually have to graph a "magnitude" spectrum and "phase" spectrum.

References

Haykin, S.: *Communication Systems*, John-Wiley & Sons, Inc., New York, 1994.

Lathi, B. P.: *Modern Digital and Analog Communication Systems*, Holt-Saunders, Tokyo, 1983.

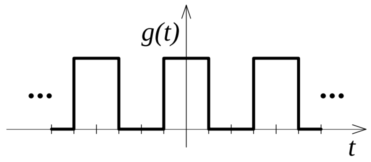
Quiz

Encircle the correct answer, cross out the wrong answers. [one or none correct]

1.
The signal $3\cos(2\pi t)+4\sin(4\pi t)$ has:

(a) $A_3=3$, $\phi_3=0$ (b) $G_{\pm1}=3/2\angle\mp\pi/2$ (c) $a_2=0$, $b_2=4$

2.



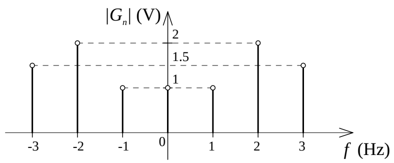
The Fourier series of the periodic signal will have no:

(a) DC term (b) odd harmonics (c) even harmonics

3.
The double-sided amplitude spectrum of a real signal always possesses:

(a) even symmetry (b) odd symmetry (c) no symmetry

4.



Amplitude spectrum of a signal. The power (across 1Ω) is:

(a) 14.5 W (b) 10 W (c) 15.5 W

5.
The phase of the 3rd harmonic component of the periodic signal $g(t)=\sum_{n=-\infty}^{\infty} A\,\mathrm{rect}\left(\frac{t-1-4n}{0.2}\right)$ is:

(a) $90°$ (b) $-90°$ (c) $0°$

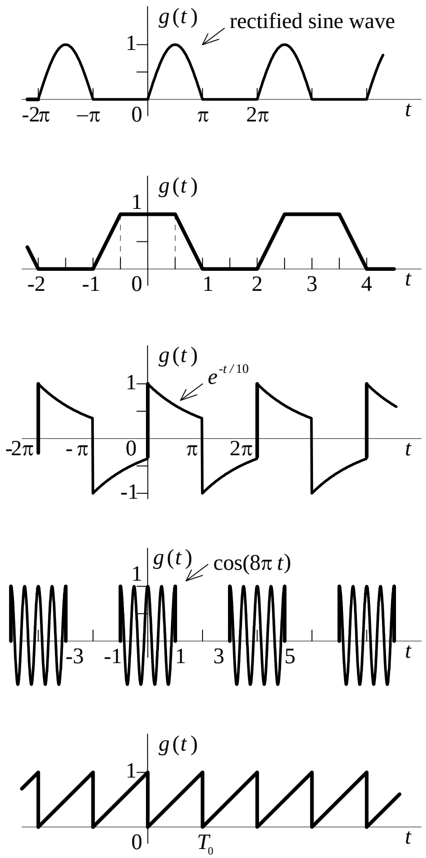
Answers: 1. c 2. c 3. a 4. c 5. a

Exercises

1.

For each of the periodic signals shown, find the complex exponential Fourier series coefficients. Plot the magnitude and phase spectra in MATLAB®, for $-25 \le n \le 25$.

(a) (b) (c) (d) (e)



2.

A signal $g(t) = 2\cos(100\pi t) - \sin(100\pi t)$.

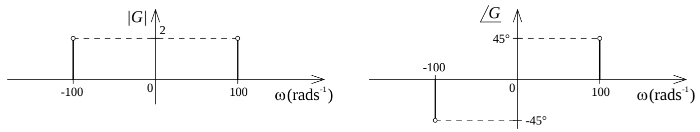
(a) Plot the single-sided magnitude spectrum and phase spectrum.

(b) Plot the double-sided magnitude spectrum and phase spectrum.

(c) Plot the real and imaginary components of the double-sided spectrum.

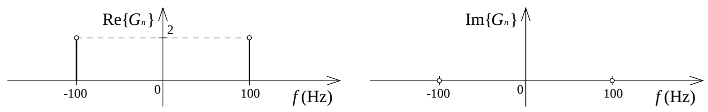
3.

What is $\overline{G}$ for the sinusoid whose spectrum is shown below.



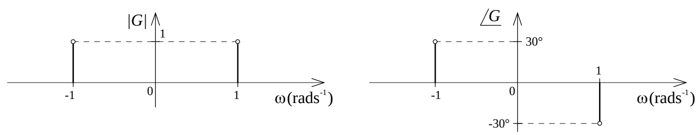
4.

What is $g(t)$ for the signal whose spectrum is sketched below.



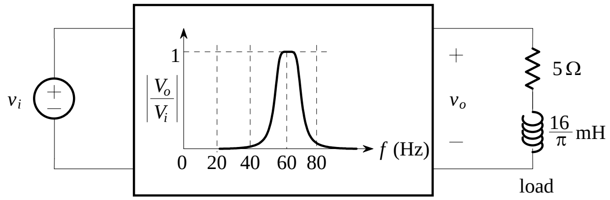
5.

Sketch photographs of the counter rotating phasors associated with the spectrum below at $t = 0$, $t = 1$ and $t = 2$.



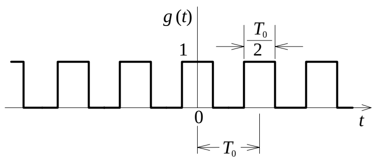
6.

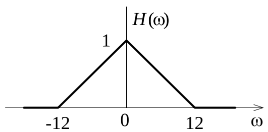
A voltage waveform $v_i(t)$ with a period of 0.08s is defined by: $v_i = 60\text{ V},\ 0 < t < 0.01\text{ s}$; $v_i = 0\text{ V},\ 0.01 < t < 0.08\text{ s}$. The voltage v_i is applied as the source in the circuit shown below. What average power is delivered to the load?



7.

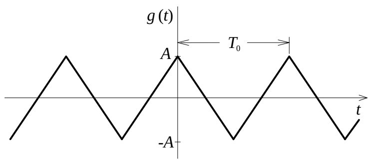
A periodic signal $g(t)$ is transmitted through a system with transfer function $H(\omega)$. For three different values of T_0 ($T_0 = 2\pi/3,\ \pi/3$, and $\pi/6$) find the double-sided magnitude spectrum of the output signal. Calculate the power of the output signal as a percentage of the power of the input signal $g(t)$.





8.

Estimate the bandwidth B of the periodic signal $g(t)$ shown below if the power of all components of $g(t)$ within the band B is to be at least 99.9 percent of the total power of $g(t)$.



Joseph Fourier (1768-1830) *(Jo´ sef Foor´ yay)*

Fourier is famous for his study of the flow of heat in metallic plates and rods. The theory that he developed now has applications in industry and in the study of the temperature of the Earth’s interior. He is also famous for the discovery that many functions could be expressed as infinite sums of sine and cosine terms, now called a trigonometric series, or Fourier series.

Fourier first showed talent in literature, but by the age of thirteen, mathematics became his real interest. By fourteen, he had completed a study of six volumes of a course on mathematics. Fourier studied for the priesthood but did not end up taking his vows. Instead he became a teacher of mathematics. In 1793 he became involved in politics and joined the local Revolutionary Committee. As he wrote:-

> *As the natural ideas of equality developed it was possible to conceive the sublime hope of establishing among us a free government exempt from kings and priests, and to free from this double yoke the long-usurped soil of Europe. I readily became enamoured of this cause, in my opinion the greatest and most beautiful which any nation has ever undertaken.*

Fourier became entangled in the French Revolution, and in 1794 he was arrested and imprisoned. He feared he would go to the guillotine but political changes allowed him to be freed. In 1795, he attended the Ecole Normal and was taught by, among others, Lagrange and Laplace. He started teaching again, and began further mathematical research. In 1797, after another brief period in prison, he succeeded Lagrange in being appointed to the chair of analysis and mechanics. He was renowned as an outstanding lecturer but did not undertake original research at this time.

In 1798 Fourier joined Napoleon on his invasion of Egypt as scientific adviser. The expedition was a great success (from the French point of view) until August 1798 when Nelson’s fleet completely destroyed the French fleet in the Battle of the Nile, so that Napoleon found himself confined to the land he was occupying. Fourier acted as an administrator as French type political institutions and administrations were set up. In particular he helped establish educational facilities in Egypt and carried out archaeological explorations.

The Institute d'Égypte was responsible for the completely serendipitous discovery of the Rosetta Stone in 1799. The three inscriptions on this stone in two languages and three scripts (hieroglyphic, demotic and Greek) enabled Thomas Young and Jean-François Champollion, a protégé of Fourier, to invent a method of translating hieroglyphic writings of ancient Egypt in 1822.

While in Cairo, Fourier helped found the Institute d'Égypte and was put in charge of collating the scientific and literary discoveries made during the time in Egypt. Napoleon abandoned his army and returned to Paris in 1799 and soon held absolute power in France. Fourier returned to France in 1801 with the remains of the expeditionary force and resumed his post as Professor of Analysis at the Ecole Polytechnique.

Napoleon appointed Fourier to be Prefect at Grenoble where his duties were many and varied – they included draining swamps and building highways. It was during his time in Grenoble that Fourier did his important mathematical work on the theory of heat. His work on the topic began around 1804 and by 1807 he had completed his important memoir *On the Propagation of Heat in Solid Bodies*. It caused controversy – both Lagrange and Laplace objected to Fourier's expansion of functions as trigonometric series.

This extract is from a letter found among Fourier's papers, and unfortunately lacks the name of the addressee, but was probably intended for Lagrange.

…it was in attempting to verify a third theorem that I employed the procedure which consists of multiplying by $\cos x\,dx$ *the two sides of the equation*

$$\phi(x) = a_0 + a_1 \cos x + a_2 \cos 2x + \ldots$$

and integrating between $x = 0$ *and* $x = \pi$. *I am sorry not to have known the name of the mathematician who first made use of this method because I would have cited him. Regarding the researches of d'Alembert and Euler could one not add that if they knew this expansion they made but a very imperfect use of it. They were both persuaded that an arbitrary…function could never be resolved in a series of this kind, and it does not seem that any one had developed a constant in cosines of multiple arcs* [i.e. found a_1, $a_2, \ldots$, with $1 = a_1 \cos x + a_2 \cos 2x + \ldots$ for $-\pi/2 < x < \pi/2$] *the first problem which I had to solve in the theory of heat.*

Other people before Fourier had used expansions of the form $f(x) \sim \sum_{r=-\infty}^{\infty} a_r \exp(irt)$ but Fourier's work extended this idea in two totally new ways. One was the "Fourier integral" (the formula for the Fourier series coefficients) and the other marked the birth of Sturm-Liouville theory (Sturm and Liouville were nineteenth century mathematicians who found solutions to

many classes of partial differential equations arising in physics that were analogous to Fourier series).

Napoleon was defeated in 1815 and Fourier returned to Paris. Fourier was elected to the Académie des Sciences in 1817 and became Secretary in 1822. Shortly after, the Academy published his prize winning essay *Théorie analytique de la chaleur* (*Analytical Theory of Heat*). In this he obtains for the first time the equation of heat conduction, which is a partial differential equation in three dimensions. As an application he considered the temperature of the ground at a certain depth due to the sun's heating. The solution consists of a yearly component and a daily component. Both effects die off exponentially with depth but the high frequency daily effect dies off much more rapidly than the low frequency yearly effect. There is also a phase lag for the daily and yearly effects so that at certain depths the temperature will be completely out of step with the surface temperature.

All these predictions are confirmed by measurements which show that annual variations in temperature are imperceptible at quite small depths (this accounts for the permafrost, i.e. permanently frozen subsoil, at high latitudes) and that daily variations are imperceptible at depths measured in tenths of metres. A reasonable value of soil thermal conductivity leads to a prediction that annual temperature changes will lag by six months at about 2–3 metres depth. Again this is confirmed by observation and, as Fourier remarked, gives a good depth for the construction of cellars.

As Fourier grew older, he developed at least one peculiar notion. Whether influenced by his stay in the heat of Egypt or by his own studies of the flow of heat in metals, he became obsessed with the idea that extreme heat was the natural condition for the human body. He was always bundled in woollen clothing, and kept his rooms at high temperatures. He died in his sixty-third year, "thoroughly cooked".

References

Körner, T.W.: *Fourier Analysis*, Cambridge University Press, 1988.

Lecture 2B – The Fourier Transform

Fourier transform. Continuous spectra. Existence of the Fourier transform. Finding Fourier transforms. Symmetry between the time-domain and frequency-domain. Time shifting. Frequency shifting. Fourier transform of sinusoids. Relationship between the Fourier series and Fourier transform. Fourier transform of a uniform train of impulses. Standard Fourier transforms. Fourier transform properties.

The Fourier Transform

The Fourier series is used to represent a periodic function as a weighted sum of sinusoidal (or complex exponential) functions. We would like to extend this result to functions that are not periodic. Such an extension is possible by what is known as the Fourier transform representation of a function.

Developing the Fourier transform

To derive the Fourier transform, we start with an aperiodic signal $g(t)$:

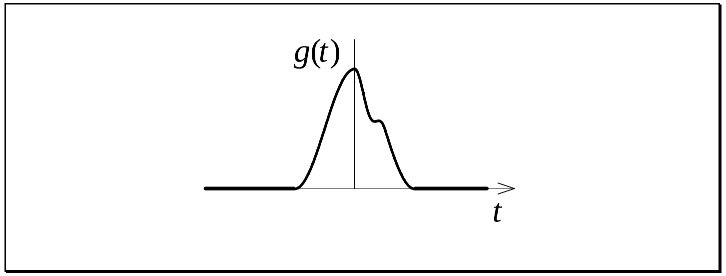


Figure 2B.1

Now we construct a new periodic signal $g_p(t)$ consisting of the signal $g(t)$ repeating itself every T_0 seconds:

Make an artificial periodic waveform from the original aperiodic waveform

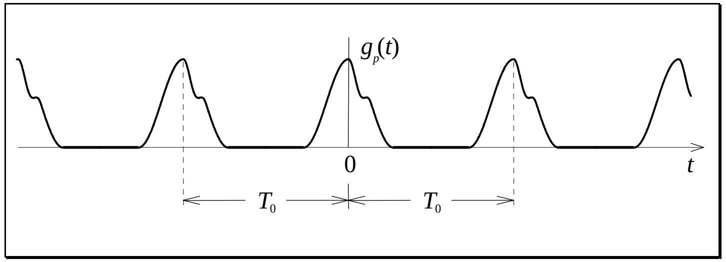


Figure 2B.2

The period T_0 is made long enough so that there is no overlap between the repeating pulses. This new signal $g_p(t)$ is a periodic signal and so it can be represented by an exponential Fourier series.

In the limit, if we let T_0 become infinite, the pulses in the periodic signal repeat after an infinite interval, and:

$$\lim_{T_0 \to \infty} g_p(t) = g(t) \tag{2B.1}$$

Thus, the Fourier series representing $g_p(t)$ will also represent $g(t)$, in the limit $T_0 \to \infty$.

The exponential Fourier series for $g_p(t)$ is:

The Fourier series for our artificial periodic waveform

$$g_p(t) = \sum_{n=-\infty}^{\infty} G_n e^{j2\pi n f_0 t} \tag{2B.2a}$$

$$G_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g_p(t) e^{-j2\pi n f_0 t} dt \tag{2B.2b}$$

In the time-domain, as $T_0 \to \infty$, the pulse train becomes a single non-periodic pulse located at the origin, as in Figure 2B.1.

In the frequency-domain, as T_0 becomes larger, f_0 becomes smaller and the spectrum becomes "denser" (the discrete frequencies are closer together). In the limit as $T_0 \to \infty$, the fundamental frequency $f_0 \to df$ and the harmonics become infinitesimally close together. The individual harmonics no longer exist as they form a continuous function of frequency. In other words, the spectrum exists for every value of *f* and is no longer a discrete function of *f* but a continuous function of *f*.

As seen from Eq. (2B.2b), the amplitudes of the individual components become smaller, too. The *shape* of the frequency spectrum, however, is unaltered (all G_n's are scaled by the same amount, *T*). In the limit as $T_0 \to \infty$, the magnitude of each component becomes infinitesimally small, until they eventually vanish!

Clearly, it no longer makes sense to define the concept of a spectrum as the amplitudes and phases of certain harmonic frequencies. Instead, a new concept is introduced called *spectral density*.

To illustrate, consider a rectangular pulse train, for which we know:

$$G_n = A f_0 \tau \operatorname{sinc}(n f_0 \tau) \tag{2B.3}$$

The spectrum for $\tau = 1/5$ s and $T_0 = 1$ s looks like:

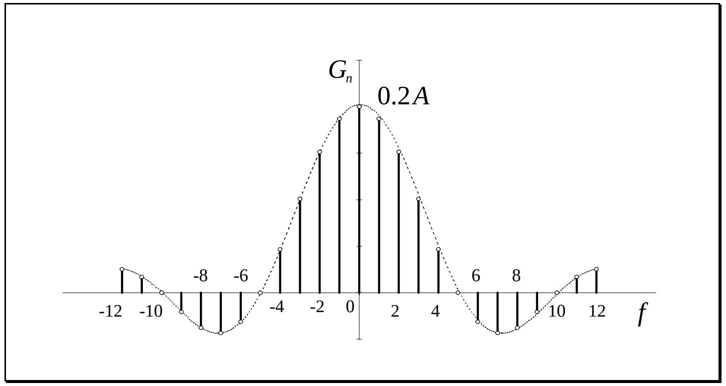


Figure 2B.3

For $T_0 = 5$ s we have:

As the period increases, the spectral lines get closer and closer, but smaller and smaller

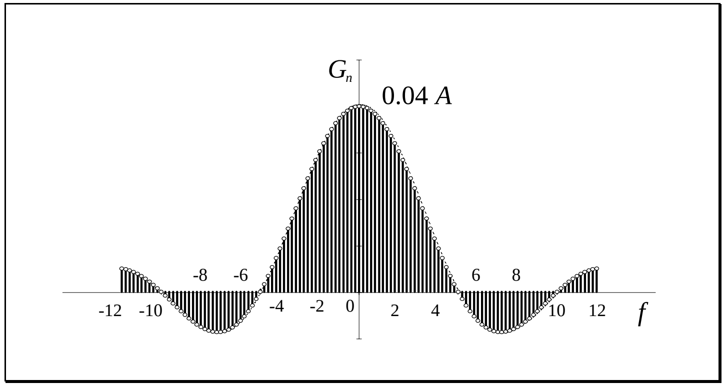


Figure 2B.4

The envelope retains the same shape (it was never a function of T_0 anyway), but the amplitude gets smaller and smaller with increasing T_0. The spectral lines get closer and closer with increasing T_0. In the limit, it is impossible to draw the magnitude spectrum as a graph of $|G_n|$ because the amplitudes of the harmonic phasors have reduced to zero.

It is possible, however, to graph a new quantity:

$$G_n T_0 = \int_{-T_0/2}^{T_0/2} g_p(t) e^{-j2\pi n f_0 t} dt \tag{2B.4}$$

which is just a rearrangement of Eq. (2B.2b). We suspect that the product $G_n T_0$ will be finite as $T_0 \to \infty$, in the same way that the area remained constant as $T \to 0$ in the family of rect functions we used to "explain" the impulse function.

As $T_0 \to \infty$, the frequency of any "harmonic" nf_0 must now correspond to the general frequency variable which describes the continuous spectrum.

In other words, n must tend to infinity as f_0 approaches zero, such that the product is finite:

$$nf_0 \to f, \qquad T_0 \to \infty \tag{2B.5}$$

With the limiting process as defined in Eqs. (2B.1) and (2B.5), Eq. (2B.4) becomes:

$$G_n T_0 \to \int_{-\infty}^{\infty} g(t) e^{-j2\pi f t} dt \tag{2B.6}$$

The right-hand side of this expression is a function of *f* (and *not* of *t*), and we represent it by:

The Fourier transform defined

$$\boxed{G(f) = \int_{-\infty}^{\infty} g(t) e^{-j2\pi f t} dt} \tag{2B.7}$$

Therefore, $G(f) = \lim_{T_0 \to \infty} G_n T_0$ has the dimensions amplitude per unit frequency, and $G(f)$ can be called the *spectral density*.

To express $g(t)$ in the time-domain using frequency components, we apply the limiting process to Eq. (2B.2a). Here, we now notice that:

$$f_0 \to df, \qquad T_0 \to \infty \tag{2B.8}$$

That is, the fundamental frequency becomes infinitesimally small as the period is made larger and larger. This agrees with our reasoning in obtaining Eq. (2B.5), since there we had an infinite number, *n*, of infinitesimally small discrete frequencies, $f_0 = df$, to give the finite continuous frequency *f*. In order to apply the limiting process to Eq. (2B.2a), we multiply the summation by $T_0 f_0 = 1$:

$$g_p(t) = \sum_{n=-\infty}^{\infty} G_n T_0 e^{j2\pi n f_0 t} f_0 \tag{2B.9}$$

and use Eq. (2B.8) and the new quantity $G(f) = G_n T_0$. In the limit the summation becomes an integral, $nf_0 \to ndf \to f$, $g_p(t) \to g(t)$, and:

The inverse Fourier transform defined

$$\boxed{g(t) = \int_{-\infty}^{\infty} G(f) e^{j2\pi ft} df} \tag{2B.10}$$

Eqs. (2B.7) and (2B.10) are collectively known as a Fourier transform pair. The function $G(f)$ is the *Fourier transform* of $g(t)$, and $g(t)$ is the *inverse Fourier transform* of $G(f)$.

To recapitulate, we have shown that:

The Fourier transform and inverse transform side-by-side

$$g(t) = \int_{-\infty}^{\infty} G(f) e^{j2\pi ft} df \tag{2B.11a}$$

$$G(f) = \int_{-\infty}^{\infty} g(t) e^{-j2\pi ft} dt \tag{2B.11b}$$

These relationships can also be expressed symbolically by a Fourier transform pair:

Notation for a Fourier transform pair

$$\boxed{g(t) \Leftrightarrow G(f)} \tag{2B.12}$$

Continuous Spectra

Making sense of a continuous spectrum

The concept of a continuous spectrum is sometimes bewildering because we generally picture the spectrum as existing at discrete frequencies and with finite amplitudes. The continuous spectrum concept can be appreciated by considering a simple analogy.

Consider a beam loaded with weights of $G_1, G_2, G_3, \ldots, G_n$ kilograms at uniformly spaced points $x_1, x_2, x_3, \ldots, x_n$ as shown in (a) below:

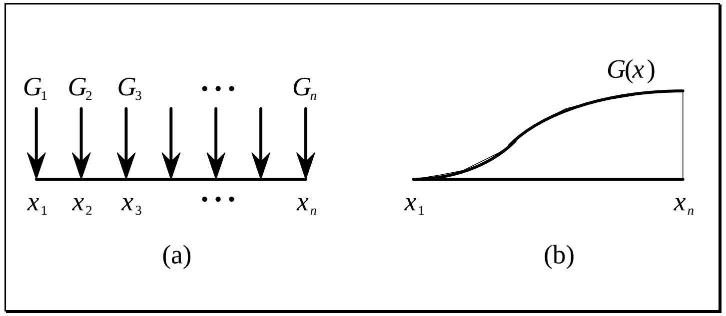


Figure 2B.5

The beam is loaded at *n* discrete points, and the total weight *W* on the beam is given by:

We can have either individual finite weights

$$W = \sum_{i=1}^{n} G_i \tag{2B.13}$$

Now consider the case of a continuously loaded beam, as shown in Figure 2B.5(b) above. The loading density $G(x)$, in kilograms per meter, is a function of *x*. The total weight on the beam is now given by a continuous sum of the infinitesimal weights at each point- that is, the integral of $G(x)$ over the entire length:

or continuous infinitesimal weights

$$W = \int_{x_1}^{x_n} G(x) dx \tag{2B.14}$$

In the discrete loading case, the weight existed only at discrete points. At other points there was no load. On the other hand, in the continuously distributed case, the loading exists at every point, but at any one point the load is zero. The load in a small distance *dx*, however, is given by $G(x)dx$. Therefore $G(x)$ represents the relative loading at a point *x*.

An exactly analogous situation exists in the case of a signal and its frequency spectrum. A periodic signal can be represented by a sum of discrete exponentials with finite amplitudes (harmonic phasors):

$$g(t)=\sum_{n=-\infty}^{\infty} G_n e^{j2\pi n f_0 t} \tag{2B.15}$$

For a nonperiodic signal, the distribution of exponentials becomes continuous; that is, the spectrum exists at every value of *f*. At any one frequency *f*, the amplitude of that frequency component is zero. The total contribution in an infinitesimal interval *df* is given by $G(f)e^{j2\pi ft}df$, and the function $g(t)$ can be expressed in terms of the continuous sum of such infinitesimal components:

We can get a signal from infinitesimal sinusoids – if we have an infinite number of them!

$$g(t)=\int_{-\infty}^{\infty} G(f)e^{j2\pi ft}df \tag{2B.16}$$

An electrical analogy could also be useful: just replace the discrete loaded beam with an array of filamentary conductors, and the continuously loaded beam with a current sheet. The analysis is the same.

Existence of the Fourier Transform

Dirichlet (1805-1859) investigated the sufficient conditions that could be imposed on a function $g(t)$ for its Fourier transform to exist. These so-called *Dirichlet conditions* are:

The Dirichlet conditions are sufficient but not necessary conditions for the FT to exist

1. On any finite interval: (2B.17a)

   a) $g(t)$ is bounded

   b) $g(t)$ has a finite number of maxima and minima

   c) $g(t)$ has a finite number of discontinuities.

2. $g(t)$ is “absolutely integrable”, i.e. $\int_{-\infty}^{\infty}|g(t)|dt<\infty$ (2B.17b)

Note that these are *sufficient* conditions and not *necessary* conditions. Use of the Fourier transform for the analysis of many useful signals would be impossible if these were necessary conditions.

Any signal with finite energy:

$$E=\int_{-\infty}^{\infty}|g(t)|^2 dt<\infty \tag{2B.18}$$

is absolutely integrable, and so all energy signals are Fourier transformable.

Many signals of interest to electrical engineers are not energy signals and are therefore not absolutely integrable. These include the step function and all periodic functions. It can be shown that signals which have infinite energy but which contain a finite amount of power and meet Dirichlet condition 1 do have valid Fourier transforms.

For practical purposes, for the signals or functions we may wish to analyse as engineers, we can use the rule of thumb that if we can draw a picture of the waveform $g(t)$, then it has a Fourier transform $G(f)$.

Finding Fourier Transforms

Example

Find the spectrum of the following rectangular pulse:

A common and useful signal…

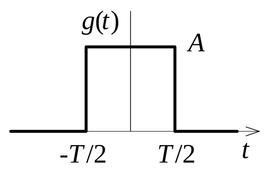


Figure 2B.6

We note that $g(t) = A\operatorname{rect}(t/T)$. We then have:

$$
\begin{aligned}
G(f) &= \int_{-\infty}^{\infty} A \operatorname{rect}\left(\frac{t}{T}\right) e^{-j2\pi f t}\, dt \\
&= A \int_{-T/2}^{T/2} e^{-j2\pi f t}\, dt \\
&= \frac{A}{-j2\pi f}\left(e^{-j\pi f T} - e^{j\pi f T}\right) \\
&= \frac{A}{\pi f}\sin(\pi f T) \\
&= AT \operatorname{sinc}(fT)
\end{aligned}
\tag{2B.19}
$$

Therefore, we have the Fourier transform pair:

and its transform

$$\boxed{A\operatorname{rect}\left(\frac{t}{T}\right) \Leftrightarrow AT\operatorname{sinc}(fT)} \tag{2B.20}$$

We can also state this graphically:

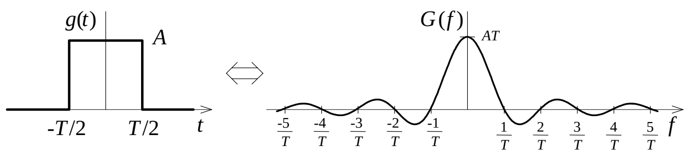


Figure 2B.7

There are a few interesting observations we can make in general by considering this result. One is that a time limited function has frequency components approaching infinity. Another is that compression in time (by making T smaller) will result in an expansion of the frequency spectrum ("wider" sinc function in this case). Another is that the Fourier transform is a linear operation – multiplication by A in the time domain results in the spectrum being multiplied by A.

Time and frequency characteristics reflect an inverse relationship

---

Letting A=1 and T=1 in Eq. (2B.20) results in what we shall call a "standard" transform:

$$\boxed{\operatorname{rect}(t) \Leftrightarrow \operatorname{sinc}(f)} \tag{2B.21}$$

One of the most common transform pairs – commit it to memory!

It is standard in the sense that we cannot make any further simplifications. It expresses a fundamental relationship between the rect and sinc functions without any complicating factors.

We may now state our observations more formally. We see that the linearity property of the Fourier transform pair can be defined as:

$$\boxed{ag(t) \Leftrightarrow aG(f)} \tag{2B.22}$$

Linearity is obeyed by transform pairs

Time scaling property of transform pairs

We can also generalise the time scaling property to:

$$\boxed{g\left(\frac{t}{T}\right) \Leftrightarrow |T| G(fT)} \tag{2B.23}$$

Thus, expansion in time corresponds to compression in frequency and vice versa.

We could (and should in future) derive our Fourier transforms by starting with a standard Fourier transform and then applying appropriate Fourier transform pair properties. For example, we could arrive at Eq. (2B.20) without integration by the following:

Applying standard properties to derive a Fourier transform

$$\operatorname{rect}(t) \Leftrightarrow \operatorname{sinc}(f) \text{ (standard transform)}$$

$$\operatorname{rect}\left(\frac{t}{T}\right) \Leftrightarrow T\operatorname{sinc}(fT) \text{ (scaling property)}$$

$$A\operatorname{rect}\left(\frac{t}{T}\right) \Leftrightarrow AT\operatorname{sinc}(fT) \text{ (linearity)} \tag{2B.24}$$

We now only have to derive enough standard transforms and discover a few Fourier transform properties to be able to handle almost any signal and system of interest.

Symmetry between the Time-Domain and Frequency-Domain

From the definition of the inverse Fourier transform:

$$g(t) = \int_{-\infty}^{\infty} G(f) e^{j2\pi ft} df \tag{2B.25}$$

we can make the transformations $t = -f$ and $f = x$ to give:

$$g(-f) = \int_{-\infty}^{\infty} G(x) e^{-j2\pi xf} dx \tag{2B.26}$$

Now since the value of a definite integral is independent of the variable of integration, we can make the change of variable $x = t$ to give:

$$g(-f) = \int_{-\infty}^{\infty} G(t) e^{-j2\pi tf} dt \tag{2B.27}$$

Notice that the right-hand side is precisely the definition of the Fourier transform of the function $G(t)$. Thus, there is an almost symmetrical relationship between the transform and its inverse. This is summed up in the duality property:

Duality property of transform pairs

$$\boxed{G(t) \Leftrightarrow g(-f)} \tag{2B.28}$$

Example

Consider a rectangular pulse in the frequency-domain. This represents the frequency response of an ideal low-pass filter, with a cut-off frequency of *B*.

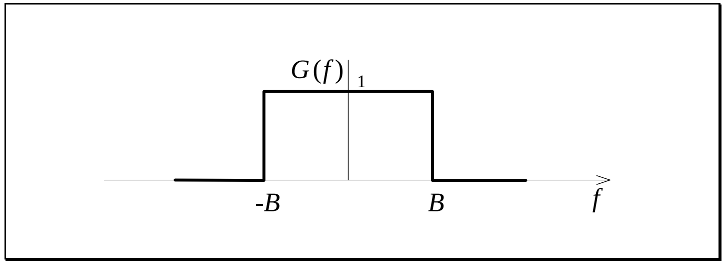


Figure 2B.8

We are interested in finding the inverse Fourier transform of this function. Using the duality property, we know that if $\operatorname{rect}(t) \Leftrightarrow \operatorname{sinc}(f)$ then $\operatorname{sinc}(t) \Leftrightarrow \operatorname{rect}(-f)$. Since the rect function is symmetric, this can also be written as $\operatorname{sinc}(t) \Leftrightarrow \operatorname{rect}(f)$. But our frequency-domain transform is $G(f) = \operatorname{rect}(f/2B)$, so we need to apply the time-scale property in reverse to arrive at:

$$2B\operatorname{sinc}(2Bt) \Leftrightarrow \operatorname{rect}(f/2B) \tag{2B.29}$$

Fourier transform of a sinc function

Graphically:

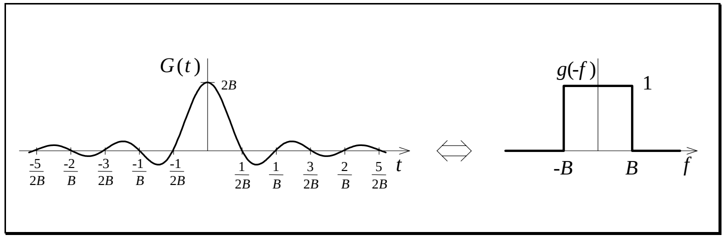


Figure 2B.9

Example

We wish to find the Fourier transform of the impulse function $\delta(t)$. We can find by direct integration (using the sifting property of the impulse function):

$$G(f) = \int_{-\infty}^{\infty} \delta(t) e^{-j2\pi ft} dt = 1 \tag{2B.30}$$

That is, we have another standard Fourier transform pair:

$$\boxed{\delta(t) \Leftrightarrow 1} \tag{2B.31}$$

Fourier transform of an impulse function

It is interesting to note that we could have obtained the same result using Eq. (2B.20). Let the height be such that the area *AT* is always equal to 1, then we have from Eq. (2B.20):

$$\frac{1}{T}\operatorname{rect}\left(\frac{t}{T}\right) \Leftrightarrow \operatorname{sinc}(fT) \tag{2B.32}$$

Now let $T \to 0$ so that the rectangle function turns into the impulse function $\delta(t)$. Noting that $\operatorname{sinc}(0)$ is unity, we arrive at Eq. (2B.31). Graphically, we have:

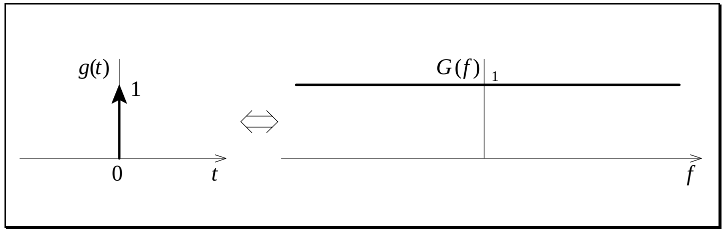


Figure 2B.10

This result says that an impulse function is composed of all frequencies from DC to infinity, and all contribute equally.

This example highlights one very important feature of finding Fourier transforms - there is usually more than one way to find them, and it is up to us (with experience and ability) to find the easiest way. However, once the Fourier transform has been obtained for a function, it is unique - this serves as a check for different methods.

Example

We wish to find the Fourier transform of the constant 1 (if it exists), which is a DC signal. Choosing the path of direct evaluation of the integral, we get:

$$G(f)=\int_{-\infty}^{\infty} 1 e^{-j2\pi ft} dt \tag{2B.33}$$

which appears intractable. However, we can approach the problem another way. If we let our rectangular function of Figure 2B.6 take on infinite width, it becomes a DC signal. As the width of the rectangular pulse becomes larger and larger, the sinc function becomes narrower and narrower, and its amplitude increases. Eventually, with infinite pulse width, the sinc function has no width and infinite height – n impulse function. Using Eq. (2B.20) with $T \to \infty$, and noting that the area under the sinc function is $1/T$ (see Lecture 1A) we therefore have:

Fourier transform of a constant

$$\boxed{1 \Leftrightarrow \delta(f)} \tag{2B.34}$$

Yet another way to arrive at this result is through recognising that a certain amount of symmetry exists in equations for the Fourier transform and inverse Fourier transform.

Applying the duality property to Eq. (2B.31) gives:

$$1 \Leftrightarrow \delta(-f) \tag{2B.35}$$

Since the impulse function is an even function, we get:

$$\boxed{1 \Leftrightarrow \delta(f)} \tag{2B.36}$$

which is the same as Eq. (2B.34). Again, two different methods converged on the same result, and one method (direct integration) seemed impossible! It is therefore advantageous to become familiar with the properties of the Fourier transform.

Time Shifting

A time shift to the right of T seconds (i.e. a time delay) can be represented by $g(t-t_0)$. A time shift to the left of T seconds can be represented by $g(t+t_0)$. The Fourier transform of a time shifted function to the right is:

$$g(t-t_0) \Leftrightarrow \int_{-\infty}^{\infty} g(t-t_0) e^{-j2\pi ft} dt \tag{2B.37}$$

Letting $x = t - t_0$ then $dx = dt$ and $t = x + t_0$ and we have:

Time shift property defined

$$\boxed{\begin{aligned} g(t-t_0) &\Leftrightarrow \int_{-\infty}^{\infty} g(x) e^{-j2\pi f(x+t_0)} dx \\ g(t-t_0) &\Leftrightarrow G(f) e^{-j2\pi f t_0} \end{aligned}} \tag{2B.38}$$

Therefore, a time shift to the right of t_0 seconds in the time-domain is equivalent to multiplication by $e^{-j2\pi f t_0}$ in the frequency-domain. Thus, a time delay simply causes a *linear phase change* in the spectrum – the magnitude spectrum is left unaltered. For example, the Fourier transform of $g(t) = A\,\mathrm{rect}\left(\frac{t-T/2}{T}\right)$ is $G(f) = AT\,\mathrm{sinc}(fT)e^{-j\pi T}$ and is shown below:

Time shifting a function just changes the phase

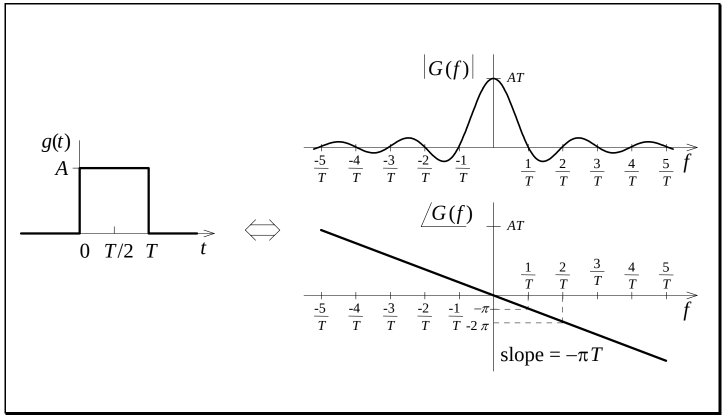


Figure 2B.11

Frequency Shifting

Similarly to the last section, a spectrum $G(f)$ can be shifted to the right, $G(f-f_0)$, or to the left, $G(f+f_0)$. This property will be used to derive several standard transforms and is particularly important in communications where it forms the basis for modulation and demodulation.

The inverse Fourier transform of a frequency shifted function to the right is:

$$\int_{-\infty}^{\infty} G(f-f_0) e^{j2\pi ft} df \Leftrightarrow G(f-f_0) \tag{2B.39}$$

Letting $x = f - f_0$ then $dx = df$ and $f = x + f_0$ and we have:

Frequency shift property defined

$$\boxed{\begin{aligned} \int_{-\infty}^{\infty} G(x) e^{j2\pi(x+f_0)t} dx &\Leftrightarrow G(f-f_0) \\ g(t) e^{j2\pi f_0 t} &\Leftrightarrow G(f-f_0) \end{aligned}} \tag{2B.40}$$

Therefore, a frequency shift to the right of f_0 Hertz in the frequency-domain is equivalent to multiplication by $e^{j2\pi f t_0}$ in the time-domain. Similarly, a shift to the left by f_0 Hertz in the frequency-domain is equivalent to multiplication by $e^{-j2\pi f t_0}$ in the time-domain. The sign of the exponent in the shifting factor is opposite to that with time shifting – this can also be explained in terms of the duality property.

For example, consider "amplitude modulation", making use of the of the Euler expansion for a cosine:

Multiplication by a sinusoid in the time-domain shifts the original spectrum up and down by the carrier frequency

$$\begin{aligned} g(t)\cos(2\pi f_c t) &= g(t)\left[\tfrac{1}{2} e^{j2\pi f_c t} + \tfrac{1}{2} e^{-j2\pi f_c t}\right] \\ &\qquad\qquad \Downarrow \qquad\qquad \Downarrow \\ &\Leftrightarrow \tfrac{1}{2} G(f-f_c) + \tfrac{1}{2} G(f+f_c) \end{aligned} \tag{2B.41}$$

The Fourier Transform of Sinusoids

Using Eq. (2B.34) and the frequency shifting property, we know that:

$$e^{j2\pi f_0 t} \Leftrightarrow \delta(f - f_0) \qquad \text{(2B.42a)}$$

$$e^{-j2\pi f_0 t} \Leftrightarrow \delta(f + f_0) \qquad \text{(2B.42b)}$$

Therefore, by substituting into Euler's relationship for the cosine and sine:

$$\cos(2\pi f_0 t) = \tfrac{1}{2} e^{j2\pi f_0 t} + \tfrac{1}{2} e^{-j2\pi f_0 t} \qquad \text{(2B.43a)}$$

$$\begin{aligned}\sin(2\pi f_0 t) &= \tfrac{1}{2j} e^{j2\pi f_0 t} - \tfrac{1}{2j} e^{-j2\pi f_0 t} \\ &= -\tfrac{j}{2} e^{j2\pi f_0 t} + \tfrac{j}{2} e^{-j2\pi f_0 t}\end{aligned} \qquad \text{(2B.43b)}$$

we get the following transform pairs:

Standard transforms for cos and sin

$$\boxed{\cos(2\pi f_0 t) \Leftrightarrow \frac{1}{2}\delta(f - f_0) + \frac{1}{2}\delta(f + f_0)} \qquad \text{(2B.44a)}$$

$$\boxed{\sin(2\pi f_0 t) \Leftrightarrow -\frac{j}{2}\delta(f - f_0) + \frac{j}{2}\delta(f + f_0)} \qquad \text{(2B.44b)}$$

Graphically:

Spectra for cos and sin

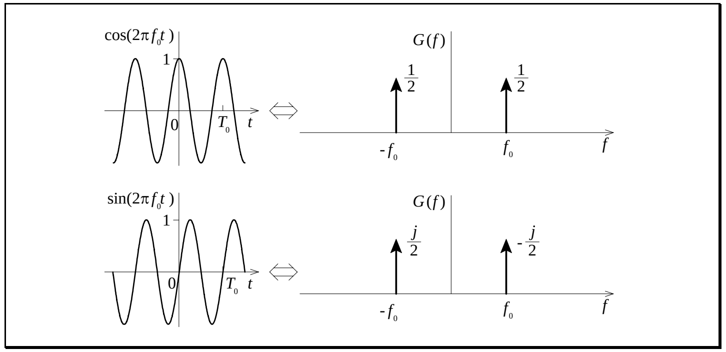


Figure 2B.12

Relationship between the Fourier Series and Fourier Transform

From the definition of the Fourier series, we know we can express any periodic waveform as a sum of harmonic phasors:

$$g(t) = \sum_{n=-\infty}^{\infty} G_n e^{j2\pi n f_0 t} \qquad \text{(2B.45)}$$

This sum of harmonic phasors is just a linear combination of complex exponentials. From Eq. (2B.42a), we already know the transform pair:

$$e^{j2\pi n f_0 t} \Leftrightarrow \delta(f - nf_0) \qquad \text{(2B.46)}$$

We therefore expect, since the Fourier transform is a linear operation, that we can easily find the Fourier transform of any periodic signal. Scaling Eq. (2B.46) by a complex constant, G_n, we have:

$$G_n e^{j2\pi n f_0 t} \Leftrightarrow G_n \delta(f - nf_0) \qquad \text{(2B.47)}$$

Summing over all harmonically related exponentials, we get:

$$\boxed{\sum_{n=-\infty}^{\infty} G_n e^{j2\pi n f_0 t} \Leftrightarrow \sum_{n=-\infty}^{\infty} G_n \delta(f - nf_0)} \qquad \text{(2B.48)}$$

Therefore, in words:

The spectrum of a periodic signal is a weighted train of impulses – each weight is equal to the Fourier series coefficient at that frequency, G_n (2B.49)

Example

Find the Fourier transform of the rectangular pulse train $g(t)$ shown below:

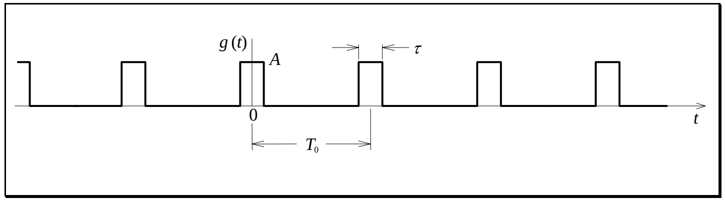


Figure 2B.13

We have already found the Fourier series coefficients for this waveform:

$$G_n = Af_0\tau\operatorname{sinc}\left(nf_0\tau\right) \tag{2B.50}$$

For the case of $\tau = T_0/5$, the spectrum is then a weighted train of impulses, with spacing equal to the fundamental of the waveform, f_0:

The double-sided magnitude spectrum of a rectangular pulse train is a weighted train of impulses

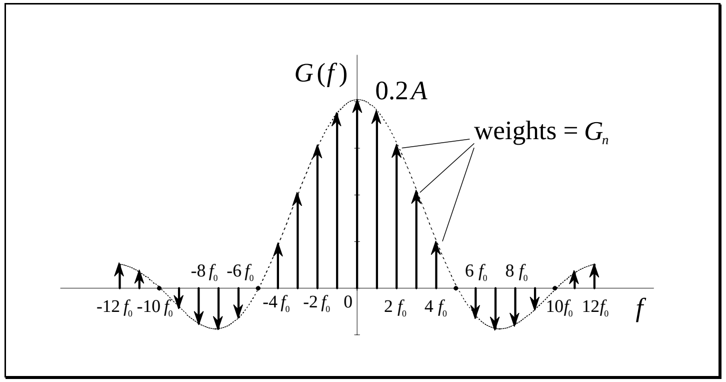


Figure 2B.14

Pairs of impulses in the spectrum correspond to a sinusoid in the time-domain

This should make intuitive sense – the spectrum is now defined as "spectral density", or a graph of *infinitesimally small phasors* spaced *infinitesimally close together*. If we have an impulse in the spectrum, then that must mean a *finite phasor* at a *specific frequency*, i.e. a sinusoid (we recognise that a *single* sinusoid has a *pair* of impulses for its spectrum – see Figure 2B.12).

The Fourier Transform of a Uniform Train of Impulses

We will encounter a uniform train of impulses frequently in our analysis of communication systems:

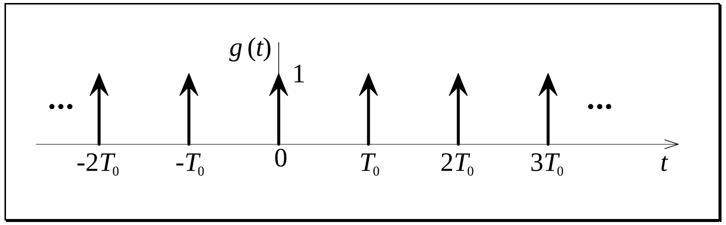


Figure 2B.15

To find the Fourier transform of this waveform, we simply note that it is periodic, and so Eq. (2B.48) applies. We have already found the Fourier series coefficients of the uniform train of impulses as the limit of a rectangular pulse train:

$$G_n = f_0 \tag{2B.51}$$

Using Eq. (2B.48), we therefore have the following Fourier transform pair:

The Fourier transform of a uniform train of impulses is a uniform train of impulses

$$\boxed{\sum_{k=-\infty}^{\infty}\delta\left(t-kT_0\right) \Leftrightarrow f_0\sum_{n=-\infty}^{\infty}\delta\left(f-nf_0\right)} \tag{2B.52}$$

Graphically:

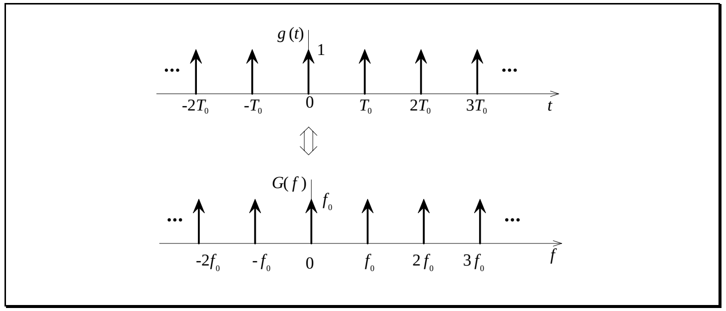


Figure 2B.16

Standard Fourier Transforms

$$\mathrm{rect}(t) \Leftrightarrow \mathrm{sinc}(f) \tag{F.1}$$

$$\delta(t) \Leftrightarrow 1 \tag{F.2}$$

$$e^{-t}u(t) \Leftrightarrow \frac{1}{1+j2\pi f} \tag{F.3}$$

$$A\cos(2\pi f_0 t+\phi) \Leftrightarrow \frac{A}{2}e^{j\phi}\delta(f-f_0)+\frac{A}{2}e^{-j\phi}\delta(f+f_0) \tag{F.4}$$

$$\sum_{k=-\infty}^{\infty}\delta(t-kT_0) \Leftrightarrow f_0\sum_{n=-\infty}^{\infty}\delta(f-nf_0) \tag{F.5}$$

Fourier Transform Properties

Assuming $g(t) \Leftrightarrow G(f)$.

$$ag(t) \Leftrightarrow aG(f) \tag{F.6}$$ Linearity

$$g\left(\frac{t}{T}\right) \Leftrightarrow |T|G(fT) \tag{F.7}$$ Scaling

$$g(t-t_0) \Leftrightarrow G(f)e^{-j2\pi f t_0} \tag{F.8}$$ Time-shifting

$$g(t)e^{j2\pi f_0 t} \Leftrightarrow G(f-f_0) \tag{F.9}$$ Frequency-shifting

$$G(t) \Leftrightarrow g(-f) \tag{F.10}$$ Duality

$$\frac{d}{dt}g(t) \Leftrightarrow j2\pi f G(f) \tag{F.11}$$ Time-differentiation

$$\int_{-\infty}^{t} g(\tau)d\tau \Leftrightarrow \frac{1}{j2\pi f}G(f)+\frac{G(0)}{2}\delta(f) \tag{F.12}$$ Time-integration

$$g_1(t)g_2(t) \Leftrightarrow G_1(f)*G_2(f) \tag{F.13}$$ Multiplication

$$g_1(t)*g_2(t) \Leftrightarrow G_1(f)G_2(f) \tag{F.14}$$ Convolution

$$\int_{-\infty}^{\infty} g(t)dt = G(0) \tag{F.15}$$

Area

$$\int_{-\infty}^{\infty} G(f)df = g(0) \tag{F.16}$$

Summary

- Aperiodic waveforms do not have a Fourier series – they have a Fourier transform. Periodic waveforms also have a Fourier transform if we allow for the existence of impulses in the transform.

- A spectrum, or spectral density, of *any* waveform is a graph of the amplitude and phase of the Fourier transform of the waveform.

- To find the Fourier transform of a signal, we start with a known Fourier transform pair, and apply any number of Fourier transform pair properties to arrive at the solution. We can check if necessary by using the definition of the Fourier transform and its inverse.

- The Fourier transform of a periodic signal is a weighted train of impulses – each impulse occurs at a harmonic of the fundamental and is weighted by the corresponding Fourier series coefficient at that frequency.

References

Haykin, S.: *Communication Systems*, John-Wiley & Sons, Inc., New York, 1994.

Lathi, B. P.: *Modern Digital and Analog Communication Systems*, Holt-Saunders, Tokyo, 1983.

Exercises

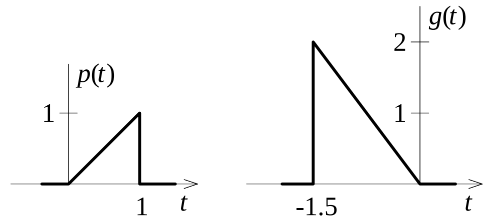
1.

Show that:

$$e^{-a|t|} \Leftrightarrow \frac{2a}{a^2 + (2\pi f)^2}$$

2.

The Fourier transform of a pulse $p(t)$ is $P(f)$. Find the Fourier transform of $g(t)$ in terms of $P(f)$.



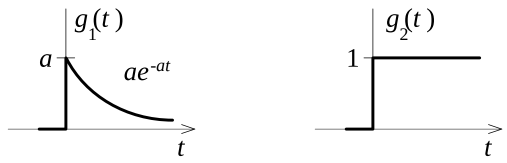
3.

Show that:

$$g(t-t_0) + g(t+t_0) \Leftrightarrow 2G(f)\cos(2\pi f t_0)$$

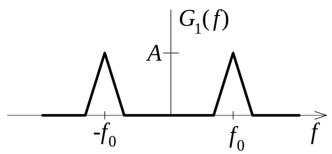
4.

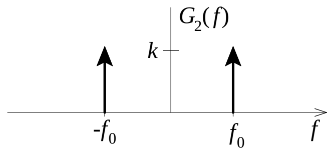
Use $G_1(f)$ and $G_2(f)$ to evaluate $g_1(t) * g_2(t)$ for the signals shown below:



5.

Find $G_1(f) * G_2(f)$ for the signals shown below:

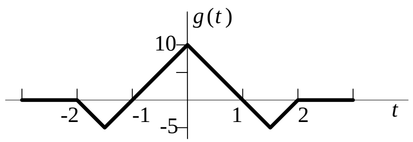




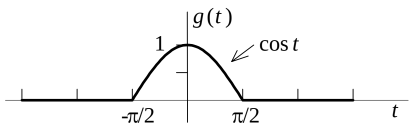
6.

Find the Fourier transforms of the following functions:

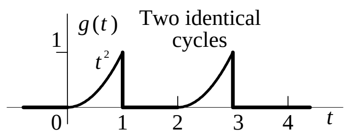
(a)



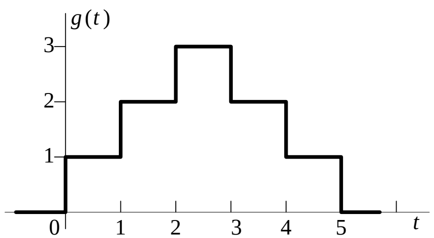
(b)



(c)



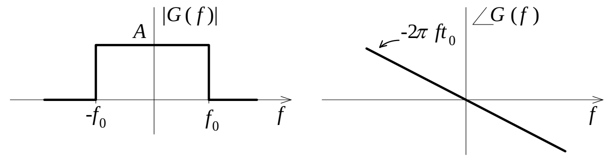
(d)



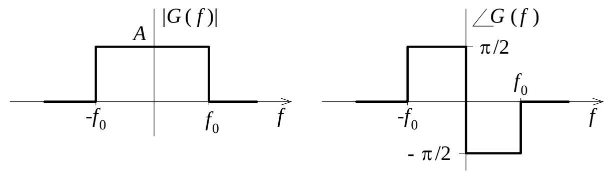
7.

Determine signals $g(t)$ whose Fourier transforms are shown below:

(a)

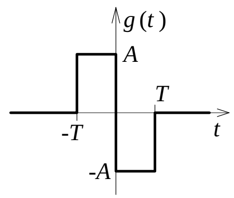


(b)



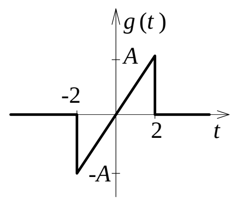
8.

Using the time-shifting property, determine the Fourier transform of the signal shown below:



9.

Using the time-differentiation property, determine the Fourier transform of the signal below:



William Thomson (Lord Kelvin) (1824-1907)

William Thomson was probably the first true electrical engineer. His engineering was firmly founded on a solid bedrock of mathematics. He invented, experimented, advanced the state-of-the-art, was entrepreneurial, was a businessman, had a multi-disciplinary approach to problems, held office in the professional body of his day (the Royal Society), published papers, gave lectures to lay people, strived for an understanding of basic physical principles and exploited that knowledge for the benefit of mankind.

William Thomson was born in Belfast, Ireland. His father was a professor of engineering. When Thomson was 8 years old his father was appointed to the chair of mathematics at the University of Glasgow. By age 10, William Thomson was attending Glasgow University. He studied astronomy, chemistry and natural philosophy (physics, heat, electricity and magnetism). Prizes in Greek, logic (philosophy), mathematics, astronomy and physics marked his progress. In 1840 he read Fourier's *The Analytical Theory of Heat* and wrote:

> *…I had become filled with the utmost admiration for the splendour and poetry of Fourier… I took Fourier out of the University Library; and in a fortnight I had mastered it - gone right through it.*

At the time, lecturers at Glasgow University took a strong interest in the approach of the French mathematicians towards physical science, such as Lagrange, Laplace, Legendre, Fresnel and Fourier. In 1840 Thomson also read Laplace's *Mécanique Céleste* and visited Paris.

In 1841 Thomson entered Cambridge and in the same year he published a paper on *Fourier's expansions of functions in trigonometrical series*. A more important paper *On the uniform motion of heat and its connection with the mathematical theory of electricity* was published in 1842.

The examinations in Cambridge were fiercely competitive exercises in problem solving against the clock. The best candidates were trained as for an athletics contest. Thomson (like Maxwell later) came second. A day before he left Cambridge, his coach gave him two copies of Green's *Essay on the*

Application of Mathematical Analysis to the Theories of Electricity and Magnetism.

After graduating, he moved to Paris on the advice of his father and because of his interest in the French approach to mathematical physics. Thomson began trying to bring together the ideas of Faraday, Coulomb and Poisson on electrical theory. He began to try and unify the ideas of "action-at-a-distance", the properties of the "ether" and ideas behind an "electrical fluid". He also became aware of Carnot's view of heat.

In 1846, at the age of twenty two, he returned to Glasgow on a wave of testimonials from, among others, De Morgan, Cayley, Hamilton, Boole, Sylvester, Stokes and Liouville, to take up the post of professor of natural philosophy. In 1847-49 he collaborated with Stokes on hydrodynamic studies, which Thomson applied to electrical and atomic theory. In electricity Thomson provided the link between Faraday and Maxwell. He was able to mathematise Faraday's laws and to show the formal analogy between problems in heat and electricity. Thus the work of Fourier on heat immediately gave rise to theorems on electricity and the work of Green on potential theory immediately gave rise to theorems on heat flow. Similarly methods used to deal with linear and rotational displacements in elastic solids could be applied to give results on electricity and magnetism. The ideas developed by Thomson were later taken up by Maxwell in his new theory of electromagnetism.

Thomson's other major contribution to fundamental physics was his combination of the almost forgotten work of Carnot with the work of Joule on the conservation of energy to lay the foundations of thermodynamics. The thermodynamical studies of Thomson led him to propose an absolute temperature scale in 1848 (The Kelvin absolute temperature scale, as it is now known, was defined much later after conservation of energy was better understood).

The Age of the Earth

In the first decades of the nineteenth century geological evidence for great changes in the past began to build up. Large areas of land had once been under water, mountain ranges had been thrown up from lowlands and the evidence of fossils showed the past existence of species with no living counterparts. Lyell, in his *Principles of Geology* sought to explain these changes "by causes now in operation". According to his theory, processes – such as slow erosion by wind and water; gradual deposition of sediment by rivers; and the cumulative effect of earthquakes and volcanic action – combined over very long periods of time to produce the vast changes recorded in the Earth's surface. Lyell's so-called 'uniformitarian' theory demanded that the age of the Earth be measured in terms of hundreds of millions and probably in terms of billions of years. Lyell was able to account for the disappearance of species in the geological record but not for the appearance of new species. A solution to this problem was provided by Charles Darwin (and Wallace) with his theory of evolution by natural selection. Darwin's theory also required vast periods of time for operation. For natural selection to operate, the age of the Earth had to be measured in many hundreds of millions of years.

Such demands for vast amounts of time run counter to the laws of thermodynamics. Every day the sun radiates immense amounts of energy. By the law of conservation of energy there must be some source of this energy. Thomson, as one of the founders of thermodynamics, was fascinated by this problem. Chemical processes (such as the burning of coal) are totally insufficient as a source of energy and Thomson was forced to conclude that gravitational potential energy was turned into heat as the sun contracted. On this assumption his calculations showed that the Sun (and therefore the Earth) was around 100 million years old.

However, Thomson's most compelling argument concerned the Earth rather than the Sun. It is well known that the temperature of the Earth increases with depth and

> *this implies a continual loss of heat from the interior, by conduction outwards through or into the upper crust. Hence, since the upper crust does not become hotter from year to year there must be a…loss of heat from the whole earth. It is possible that no cooling may result from this loss of heat but only an exhaustion of potential energy which in this case could scarcely be other than chemical.*

Since there is no reasonable mechanism to keep a chemical reaction going at a steady pace for millions of years, Thomson concluded "…that the earth is merely a warm chemically inert body cooling". Thomson was led to believe that the Earth was a solid body and that it had solidified at a more or less uniform temperature. Taking the best available measurements of the conductivity of the Earth and the rate of temperature change near the surface, he arrived at an estimate of 100 million years as the age of the Earth (confirming his calculations of the Sun's age).

The problems posed to Darwin's theory of evolution became serious as Thomson's arguments sank in. In the fifth edition of *The Origin of Species*, Darwin attempted to adjust to the new time scale by allowing greater scope for evolution by processes other than natural selection. Darwin was forced to ask for a suspension of judgment of his theory and in the final chapter he added

> *With respect to the lapse of time not having been sufficient since our planet was consolidated for the assumed amount of organic change, and this objection, as argued by [Thomson], is probably one of the gravest yet advanced, I can only say, firstly that we do not know at what rate species change as measured by years, and secondly, that many philosophers are not yet willing to admit that we know enough of the constitution of the universe and of the interior of our globe to speculate with safety on its past duration.*

(Darwin, *The Origin of Species*, Sixth Edition, p.409)

The chief weakness of Thomson's arguments was exposed by Huxley

A variant of the adage: "Garbage in equals garbage out".

> *…this seems to be one of the many cases in which the admitted accuracy of mathematical processes is allowed to throw a wholly inadmissible appearance of authority over the results obtained by them. Mathematics may be compared to a mill of exquisite workmanship, which grinds you stuff of any degree of fineness; but nevertheless, what you get out depends on what you put in; and as the grandest mill in the world will not extract wheat-flour from peascods, so pages of formulae will not get a definite result out of loose data.*
>
> (*Quarterly Journal of the Geological Society of London*, Vol. 25, 1869)

However, Thomson's estimates were the best available and for the next thirty years geology took its time from physics, and biology took its time from geology. But Thomson and his followers began to adjust his first estimate down until at the end of the nineteenth century the best physical estimates of the age of the Earth and Sun were about 20 million years whilst the minimum the geologists could allow was closer to Thomson's original 100 million years.

Then in 1904 Rutherford announced that the radioactive decay of radium was accompanied by the release of immense amounts of energy and speculated that this could replace the heat lost from the surface of the Earth.

> *The discovery of the radioactive elements…thus increases the possible limit of the duration of life on this planet, and allows the time claimed by the geologist and biologist for the process of evolution.*
>
> (Rutherford quoted in Burchfield, p.164)

A problem for the geologists was now replaced by a problem for the physicists. The answer was provided by a theory which was just beginning to be gossiped about. Einstein's theory of relativity extended the principle of conservation of energy by taking matter as a form of energy. It is the conversion of matter to heat which maintains the Earth's internal temperature and supplies the energy radiated by the sun. The ratios of lead isotopes in the Earth compared to meteorites now leads geologists to give the Earth an age of about 4.55 billion years.

The Transatlantic Cable

The invention of the electric telegraph in the 1830s led to a network of telegraph wires covering England, western Europe and the more settled parts of the USA. The railroads, spawned by the dual inventions of steam and steel, were also beginning to crisscross those same regions. It was vital for the smooth and safe running of the railroads, as well as the running of empires, to have speedy communication.

Attempts were made to provide underwater links between the various separate systems. The first cable between Britain and France was laid in 1850. The operators found the greatest difficulty in transmitting even a few words. After 12 hours a trawler accidentally caught and cut the cable. A second, more heavily armoured cable was laid and it was a complete success. The short lines worked, but the operators found that signals could not be transmitted along submarine cables as fast as along land lines without becoming confused.

In spite of the record of the longer lines, the American Cyrus J. Fields proposed a telegraph line linking Europe and America. Oceanographic surveys showed that the bottom of the Atlantic was suitable for cable laying. The connection of existing land telegraph lines had produced a telegraph line of the length of the proposed cable through which signals had been passed extremely rapidly. The British government offered a subsidy and money was rapidly raised.

Faraday had predicted signal retardation but he and others like Morse had in mind a model of a submarine cable as a hosepipe which took longer to fill with water (signal) as it got longer. The remedy was thus to use a thin wire (so that less electricity was needed to charge it) and high voltages to push the signal through. Faraday's opinion was shared by the electrical adviser to the project, Dr Whitehouse (a medical doctor).

Thomson's researches had given him a clearer mathematical picture of the problem. The current in a telegraph wire in air is approximately governed by the wave equation. A pulse on such a wire travels at a well defined speed with no change of shape or magnitude with time. Signals can be sent as close together as the transmitter can make them and the receiver distinguish them.

In undersea cables of the type proposed, capacitive effects dominate and the current is approximately governed by the diffusion (i.e. heat) equation. This equation predicts that electric pulses will last for a time that is proportional to the length of the cable squared. If two or more signals are transmitted within this time, the signals will be jumbled at the receiver. In going from submarine cables of 50 km length to cables of length 2400 km, retardation effects are 2500 times worse. Also, increasing the voltage makes this jumbling (called intersymbol interference) worse. Finally, the diffusion equation shows that the wire should have as large a diameter as possible (small resistance).

Whitehouse, whose professional reputation was now involved, denied these conclusions. Even though Thomson was on the board of directors of Field’s company, he had no authority over the technical advisers. Moreover the production of the cable was already underway on principles contrary to Thomson’s. Testing the cable, Thomson was astonished to find that some sections conducted only half as well as others, even though the manufacturers were supplying copper to the then highest standards of purity.

Realising that the success of the enterprise would depend on a fast, sensitive detector, Thomson set about to invent one. The problem with an ordinary galvanometer is the high inertia of the needle. Thomson came up with the mirror galvanometer in which the pointer is replaced by a beam of light.

In a first attempt in 1857 the cable snapped after 540 km had been laid. In 1858, Europe and America were finally linked by cable. On 16 August it carried a 99-word message of greeting from Queen Victoria to President Buchanan. But that 99-word message took 16½ hours to get through. In vain, Whitehouse tried to get his receiver to work. Only Thomson’s galvanometer was sensitive enough to interpret the minute and blurred messages coming through. Whitehouse ordered that a series of huge two thousand volt induction coils be used to try to push the message through faster – after four weeks of

this treatment the insulation finally failed; 2500 tons of cable and £350 000 of capital lay useless on the ocean floor.

In 1859 eighteen thousand kilometres of undersea cable had been laid in other parts of the world, and only five thousand kilometres were operating. In 1861 civil war broke out in the United States. By 1864 Field had raised enough capital for a second attempt. The cable was designed in accordance with Thomson’s theories. Strict quality control was exercised: the copper was so pure that for the next 50 years ‘telegraphist’s copper’ was the purest available.

Once again the British Government supported the project – the importance of quick communication in controlling an empire was evident to everybody. The new cable was mechanically much stronger but also heavier. Only one ship was large enough to handle it and that was Brunel’s *Great Eastern*. She was fives time larger than any other existing ship.

This time there was a competitor. The Western Union Company had decided to build a cable along the overland route across America, Alaska, the Bering Straits, Siberia and Russia to reach Europe the long way round. The commercial success of the cable would therefore depend on the rate at which messages could be transmitted. Thomson had promised the company a rate of 8 or even 12 words a minute. Half a million pounds was being staked on the correctness of the solution of a partial differential equation.

In 1865 the *Great Eastern* laid cable for nine days, but after 2000 km the cable parted. After two weeks of unsuccessfully trying to recover the cable, the expedition left a buoy to mark the spot and sailed for home. Since communication had been perfect up until the final break, Thomson was confident that the cable would do all that was required. The company decided to build and lay a new cable and then go back and complete the old one.

Cable laying for the third attempt started on 12 July 1866 and the cable was landed on the morning of the 27th. On the 28th the cable was open for business and earned £1000. Western Union ordered all work on their project to be stopped at a loss of $3 000 000.

On 1 September after three weeks of effort the old cable was recovered and on 8 September a second perfect cable linked America and Europe. A wave of knighthoods swept over the engineers and directors. The patents which Thomson held made him a wealthy man.

For his work on the transatlantic cable Thomson was created Baron Kelvin of Largs in 1892. The Kelvin is the river which runs through the grounds of Glasgow University and Largs is the town on the Scottish coast where Thomson built his house.

Other Achievements

There are many other factors influencing local tides – such as channel width – which produce phenomena akin to resonance in the tides. One example of this is the narrow Bay of Fundy, between Nova Scotia and New Brunswick, where the tide can be as high as 21m. In contrast, the Mediterranean Sea is almost tideless because it is a broad body of water with a narrow entrance.

Thomson worked on several problems associated with navigation – sounding machines, lighthouse lights, compasses and the prediction of tides. Tides are primarily due to the gravitational effects of the Moon, Sun and Earth on the oceans but their theoretical investigation, even in the simplest case of a single ocean covering a rigid Earth to a uniform depth, is very hard. Even today, the study of only slightly more realistic models is only possible by numerical computer modelling. Thomson recognised that the forces affecting the tides change periodically. He then approximated the height of the tide by a trigonometric polynomial – a Fourier series with a finite number of terms. The coefficients of the polynomial required calculation of the Fourier series coefficients by numerical integration – a task that "…required not less than twenty hours of calculation by skilled arithmeticians." To reduce this labour Thomson designed and built a machine which would trace out the predicted height of the tides for a year in a few minutes, given the Fourier series coefficients.

Michelson (of Michelson-Morley fame) was to build a better machine that used up to 80 Fourier series coefficients. The production of 'blips' at discontinuities by this machine was explained by Gibbs in two letters to *Nature*. These 'blips' are now referred to as the "Gibbs phenomenon".

Thomson also built another machine, called the *harmonic analyser*, to perform the task "which seemed to the Astronomer Royal so complicated and difficult that no machine could master it' of computing the Fourier series coefficients from the record of past heights. This was the first major victory in the struggle "to substitute brass for brain" in calculation.

Thomson introduced many teaching innovations to Glasgow University. He introduced laboratory work into the degree courses, keeping this part of the work distinct from the mathematical side. He encouraged the best students by offering prizes. There were also prizes which Thomson gave to the student that he considered most deserving.

Thomson worked in collaboration with Tait to produce the now famous text *Treatise on Natural Philosophy* which they began working on in the early 1860s. Many volumes were intended but only two were ever written which cover kinematics and dynamics. These became standard texts for many generations of scientists.

In later life he developed a complete range of measurement instruments for physics and electricity. He also established standards for all the quantities in use in physics. In all he published over 300 major technical papers during the 53 years that he held the chair of Natural Philosophy at the University of Glasgow.

During the first half of Thomson's career he seemed incapable of being wrong while during the second half of his career he seemed incapable of being right. This seems too extreme a view, but Thomson's refusal to accept atoms, his opposition to Darwin's theories, his incorrect speculations as to the age of the Earth and the Sun, and his opposition to Rutherford's ideas of radioactivity, certainly put him on the losing side of many arguments later in his career.

William Thomson, Lord Kelvin, died in 1907 at the age of 83. He was buried in Westminster Abbey in London where he lies today, adjacent to Isaac Newton.

References

Burchfield, J.D.: *Lord Kelvin and The Age of the Earth*, Macmillan, 1975.
Körner, T.W.: *Fourier Analysis*, Cambridge University Press, 1988.
Encyclopedia Britannica, 2004.
Morrison, N.: *Introduction to Fourier Analysis*, John Wiley & Sons, Inc., 1994.
Thomson, S.P.: *The Life of Lord Kelvin*, London, 1976.
Kelvin & Tait: *Treatise on Natural Philosophy*, Appendix B.

Lecture 3A – Frequency-Domain Analysis

Sinusoidal response. Arbitrary response. Ideal filters. Sampling. Reconstruction. Aliasing. Amplitude modulation (DSB-SC). Demodulation. Fourier series. Windowing. Multiplication and convolution. Filtering.

Frequency-Domain Analysis

Since we can now represent signals in terms of a Fourier series (for periodic signals) or a Fourier transform (for aperiodic signals), we seek a way to describe a *system* in terms of frequency. That is, we seek a model of a linear, time-invariant system governed by continuous-time differential equations that expresses its behaviour with respect to frequency, rather than time. The concept of a *signal's spectrum* and a *system's frequency response* will be seen to be of fundamental importance in the frequency-domain characterisation of a system.

We have a description of signals in the frequency domain - we need one for systems

The power of the frequency-domain approach will be seen as we are able to determine a system's output given *almost any* input. Fundamental signal operations can also be explained easily – such as modulation / demodulation and sampling / reconstruction – in the frequency domain that would otherwise appear bewildering in the time domain.

Response to a Sinusoidal Input

We have already seen that the output of a LTI system is given by:

$$y(t) = h(t) * x(t) \tag{3A.1}$$

Starting with a convolution description of a system

if initial conditions are zero.

Suppose the input to the system is:

$$x(t) = A\cos(\omega_0 t + \phi) \tag{3A.2}$$

We apply a sinusoid

We have already seen that this can be expressed (thanks to Euler) as:

A sinusoid is just a sum of two complex conjugate counter-rotating phasors

$$\begin{aligned} x(t) &= \frac{A}{2}e^{j\phi}e^{j\omega_0 t} + \frac{A}{2}e^{-j\phi}e^{-j\omega_0 t} \\ &= Xe^{j\omega_0 t} + X^* e^{-j\omega_0 t} \end{aligned} \tag{3A.3}$$

Where X is the *phasor* representing $x(t)$. Inserting this into Eq. (3A.1) gives:

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\tau)\left(Xe^{j\omega_0(t-\tau)} + X^* e^{-j\omega_0(t-\tau)}\right)d\tau \\ &= \int_{-\infty}^{\infty} h(\tau)e^{-j\omega_0\tau}Xe^{j\omega_0 t}d\tau + \int_{-\infty}^{\infty} h(\tau)e^{j\omega_0\tau}X^* e^{-j\omega_0 t}d\tau \\ &= \left[\int_{-\infty}^{\infty} h(\tau)e^{-j\omega_0\tau}d\tau\right]Xe^{j\omega_0 t} + \left[\int_{-\infty}^{\infty} h(\tau)e^{j\omega_0\tau}d\tau\right]X^* e^{-j\omega_0 t} \end{aligned} \tag{3A.4}$$

This rather unwieldy expression can be simplified. First of all, if we take the Fourier transform of the impulse response, we get:

The Fourier transform of the impulse response appears in our analysis…

$$\boxed{H(\omega) = \int_{-\infty}^{\infty} h(t)e^{-j\omega t}dt} \tag{3A.5}$$

where obviously $\omega = 2\pi f$. Now Eq. (3A.4) can be written as:

$$y(t) = H(\omega_0)Xe^{j\omega_0 t} + H(-\omega_0)X^* e^{-j\omega_0 t} \tag{3A.6}$$

If $h(t)$ is real, then:

$$H(-\omega) = H^*(\omega) \tag{3A.7}$$

which should be obvious by looking at the definition of the Fourier transform. Now let:

…and relates the output phasor with the input phasor!

$$\boxed{Y = H(\omega_0)X} \tag{3A.8}$$

This equation is of fundamental importance! It says that the output phasor to a system is equal to the input phasor to the system, scaled in magnitude and changed in angle by an amount equal to $H(\omega_0)$ (a complex number). Also:

$$Y^* = H^*(\omega_0)X^* = H(-\omega_0)X^* \tag{3A.9}$$

Eq. (3A.6) can now be written as:

$$y(t) = Ye^{j\omega_0 t} + Y^* e^{-j\omega_0 t} \tag{3A.10}$$

which is just another way of writing the sinusoid:

The magnitude and phase of the input sinusoid change – according to the Fourier transform of the impulse response

$$y(t) = A|H(\omega_0)|\cos(\omega_0 t + \phi + \angle H(\omega_0)) \tag{3A.11}$$

Hence the response resulting from the sinusoidal input $x(t) = A\cos(\omega_0 t + \phi)$ is also a sinusoid with the same frequency ω_0, but with the amplitude scaled by the factor $|H(\omega_0)|$ and the phase shifted by an amount $\angle H(\omega_0)$.

Frequency, magnitude and phase response defined

The function $H(\omega)$ is termed the frequency response. $|H(\omega)|$ is called the magnitude response and $\angle H(\omega)$ is called the phase response. Note that the system impulse response and the frequency response form a Fourier transform pair:

The impulse response and frequency response form a Fourier transform pair

$$\boxed{h(t) \Leftrightarrow H(f)} \tag{3A.12}$$

We now have an easy way of analysing systems with sinusoidal inputs: simply determine $H(f)$ and apply $Y = H(f_0)X$.

Two ways to find the frequency response

There are two ways to get $H(f)$. We can find the system impulse response $h(t)$ and take the Fourier transform, or we can find it directly from the differential equation describing the system.

Example

For the simple *RC* circuit below, find the response to an arbitrary sinusoid, assuming no stored energy in the system (zero initial conditions). This is termed the steady-state response, since the input is assumed to be valid for all time.

Finding the frequency response of a simple system

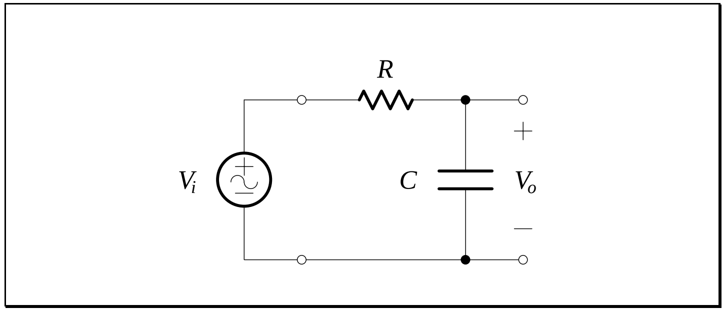


Figure 3A.1

The input/output differential equation for the circuit is:

$$\frac{dv_o(t)}{dt} + \frac{1}{RC} v_o(t) = \frac{1}{RC} v_i(t) \quad (3A.13)$$

which is obtained by KVL. Since the input is a sinusoid, which is really just a sum of conjugate complex exponentials, we know from Eq. (3A.6) that the output is $V_o = H(\omega_0) A e^{j(\omega_0 t + \phi)}$ if the input is $V_i = A e^{j(\omega_0 t + \phi)}$. Note that V_i and V_o are complex numbers, and if the factor $e^{j\omega_0 t}$ were suppressed they would be *phasors*. The differential equation Eq. (3A.13) becomes:

$$\frac{d}{dt}\left[H(\omega_0) A e^{j(\omega_0 t+\phi)}\right] + \frac{1}{RC}\left[H(\omega_0) A e^{j(\omega_0 t+\phi)}\right] = \frac{1}{RC}\left[A e^{j(\omega_0 t+\phi)}\right] \quad (3A.14)$$

and thus:

$$j\omega_0 H(\omega_0) A e^{j(\omega_0 t+\phi)} + \frac{1}{RC} H(\omega_0) A e^{j(\omega_0 t+\phi)} = \frac{1}{RC} A e^{j(\omega_0 t+\phi)} \quad (3A.15)$$

Dividing both sides by $V_i = A e^{j(\omega_0 t+\phi)}$ gives:

$$j\omega_0 H(\omega_0) + \frac{1}{RC} H(\omega_0) = \frac{1}{RC} \quad (3A.16)$$

and therefore:

$$H(\omega_0) = \frac{1/RC}{j\omega_0 + 1/RC} \quad (3A.17)$$

which yields for an arbitrary frequency:

Frequency response of a lowpass *RC* circuit

$$H(\omega) = \frac{1/RC}{j\omega + 1/RC} \quad (3A.18)$$

This is the frequency response for the simple *RC* circuit. As a check, we know that the impulse response is:

Impulse response of a lowpass *RC* circuit

$$h(t) = 1/RC\, e^{-t/RC} \quad (3A.19)$$

Using your standard transforms, show that the frequency response is the Fourier transform of the impulse response.

The magnitude function is:

Magnitude response of a lowpass *RC* circuit

$$|H(\omega)| = \frac{1/RC}{\sqrt{\omega^2 + (1/RC)^2}} \quad (3A.20)$$

and the phase function is:

Phase response of a lowpass *RC* circuit

$$\angle H(\omega) = -\tan^{-1}(\omega RC) \tag{3A.21}$$

Plots of the magnitude and phase function are shown below:

A graph of the frequency response – in this case as magnitude and phase

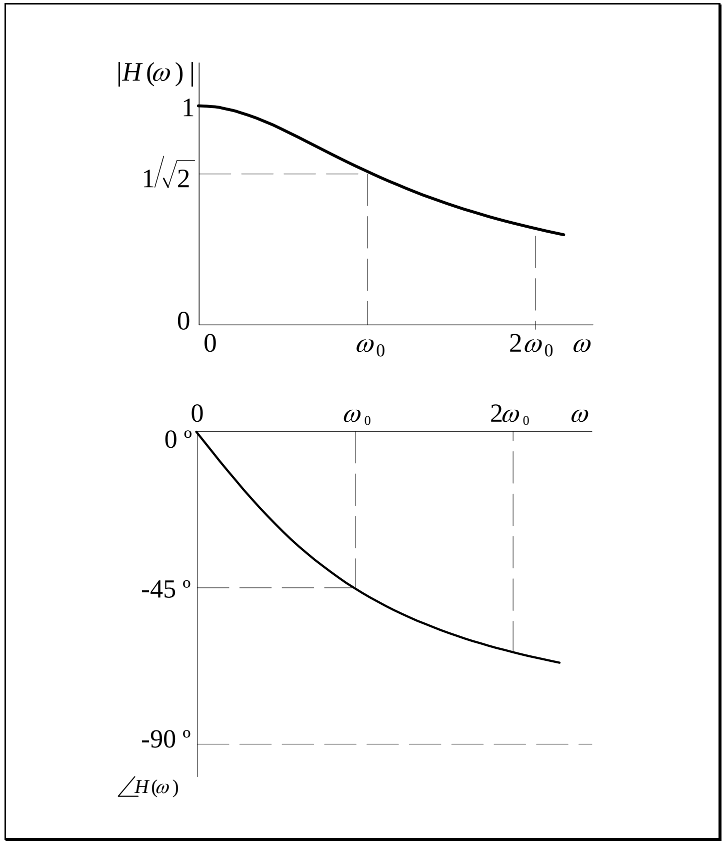


Figure 3A.2

The system's behaviour described in terms of frequency

The behaviour of the *RC* circuit is summarized by noting that it passes low frequency signals without any significant attenuation and without producing any significant phase shift. As the frequency increases, the attenuation and the phase shift become larger. Finally as the frequency increases to ∞, the *RC*

Filter terminology defined

circuit completely "blocks" the sinusoidal input. As a result of this behaviour, the circuit is an example of a *lowpass* filter. The frequency $\omega_0 = 1/RC$ is termed the *cutoff* frequency. The *bandwidth* of the filter is also equal to ω_0.

Response to an Arbitrary Input

If we can do one sinusoid, we can do an infinite number…

It should now be obvious how we handle arbitrary inputs.

Periodic Inputs

For periodic inputs, we can express the input signal by a complex exponential Fourier series:

which is just a Fourier series for a periodic signal

$$x(t) = \sum_{n=-\infty}^{\infty} X_n e^{jn\omega_o t} \tag{3A.22}$$

It follows from the previous section that the output response resulting from the complex exponential input $X_n e^{jn\omega_0 t}$ is equal to $H(n\omega_0)X_n e^{jn\omega_0 t}$. By linearity, the response to the periodic input $x(t)$ is:

$$y(t) = \sum_{n=-\infty}^{\infty} H(n\omega_0) X_n e^{jn\omega_o t} \tag{3A.23}$$

Since the right-hand side is a complex exponential Fourier series, the output $y(t)$ must be periodic, with fundamental frequency equal to that of the input, ie. the output has the same period as the input.

The frequency response simply multiplies the input Fourier series coefficients to produce the output Fourier series coefficients

It can be seen that the only thing we need to determine is new Fourier series coefficients, given by:

$$\boxed{Y_n = H(n\omega_0) X_n} \tag{3A.24}$$

The output *magnitude* spectrum is just:

$$\boxed{|Y_n| = |H(n\omega_0)||X_n|} \quad (3A.25)$$

Don't forget – the frequency response is just a frequency dependent complex number

and the output *phase* spectrum is:

$$\boxed{\angle Y_n = \angle H(n\omega_0) + \angle X_n} \quad (3A.26)$$

These relationships describe how the system "processes" the various complex exponential components comprising the periodic input signal. In particular, Eq. (3A.25) determines if the system will pass or attenuate a given component of the input. Eq. (3A.26) determines the phase shift the system will give to a particular component of the input.

Aperiodic Inputs

If we can do finite sinusoids, we can do infinitesimal sinusoids too!

Start with the convolution integral again

Taking the Fourier transform of both sides of the time domain input/output relationship of an LTI system:

$$y(t) = h(t) * x(t) \quad (3A.27)$$

we get:

and transform to the frequency domain

$$Y(f) = \int_{-\infty}^{\infty} [h(t) * x(t)] e^{-j\omega t} dt \quad (3A.28)$$

Substituting the definition of convolution, we get:

$$Y(f) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} h(\tau) x(t-\tau) d\tau\right] e^{-j\omega t} dt \quad (3A.29)$$

This can be rewritten in the form:

$$Y(f) = \int_{-\infty}^{\infty} h(\tau) \left[\int_{-\infty}^{\infty} x(t-\tau) e^{-j\omega t} dt\right] d\tau \quad (3A.30)$$

Using the change of variable $\lambda = t - \tau$ in the second integral gives:

$$Y(f) = \int_{-\infty}^{\infty} h(\tau) \left[\int_{-\infty}^{\infty} x(\lambda) e^{-j\omega(\lambda+\tau)} d\lambda\right] d\tau \quad (3A.31)$$

Factoring out $e^{-j\omega\tau}$ from the second integral, we can write:

$$Y(f) = \left[\int_{-\infty}^{\infty} h(\tau) e^{-j\omega\tau} d\tau\right] \left[\int_{-\infty}^{\infty} x(\lambda) e^{-j\omega\lambda} d\lambda\right] \quad (3A.32)$$

which is:

$$\boxed{Y(f) = H(f) X(f)} \quad (3A.33)$$

Convolution in the time-domain is multiplication in the frequency-domain

This is also a proof of the "convolution in time property" of Fourier transforms.

Eq. (3A.33) is the *frequency-domain* version of the equation given by Eq. (3A.27). It says that the spectrum of the output signal is equal to the product of the frequency response and the spectrum of the input signal.

The output spectrum is obtained by multiplying the input spectrum by the frequency response

The output *magnitude* spectrum is:

$$\boxed{|Y(f)| = |H(f)||X(f)|} \quad (3A.34)$$

The magnitude spectrum is scaled

and the output *phase* spectrum is:

$$\boxed{\angle Y(f) = \angle H(f) + \angle X(f)} \quad (3A.35)$$

The phase spectrum is added to

Note that the frequency domain description applies to all inputs that can be Fourier transformed, including sinusoids if we allow impulses in the spectrum. Periodic inputs are then a special case of Eq. (3A.33).

By similar arguments together with the duality property of the Fourier transform, it can be shown that convolution in the frequency-domain is equivalent to multiplication in the time-domain.

Convolution in the frequency-domain is multiplication in the time-domain

Ideal Filters

A first look at frequency-domain descriptions - filters

Now that we have a feel for the frequency-domain description and behaviour of a system, we will briefly examine a very important application of electronic circuits – that of frequency selection, or filtering. Here we will examine ideal filters – the topic of real filter design is rather involved.

Ideal filters pass sinusoids within a given frequency range, and reject (completely attenuate) all other sinusoids. An example of an ideal lowpass filter is shown below:

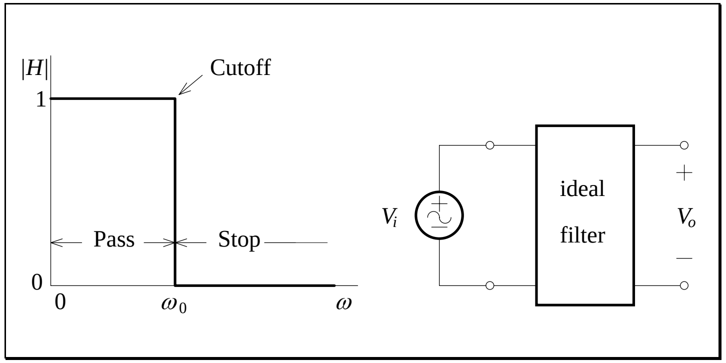


Figure 3A.3

Filter types

Other basic types of filters are highpass, bandpass and bandstop. All have similar definitions as given in Figure 3A.3. Frequencies that are passed are said to be in the *passband*, while those that are rejected lie in the *stopband*. The point where passband and stopband meet is called ω_0, the *cutoff* frequency. The term *bandwidth* as applied to a filter corresponds to the width of the passband.

An ideal lowpass filter with a bandwidth of B Hz has a magnitude response:

$$\left|H(f)\right| = K\text{rect}\left(\frac{f}{2B}\right) \tag{3A.36}$$

Phase Response of an Ideal Filter

Most filter specifications deal with the magnitude response. In systems where the filter is designed to pass a particular "waveshape", phase response is extremely important. For example, in a digital system we may be sending 1's and 0's using a specially shaped pulse that has "nice" properties, e.g. its value is zero at the centre of all other pulses. At the receiver it is passed through a lowpass filter to remove high frequency noise. The filter introduces a delay of D seconds, but the output of the filter is as close as possible to the desired pulse shape.

A particular phase response is crucial for retaining a signal's "shape"

This is illustrated below:

A filter introducing delay, but retaining signal "shape"

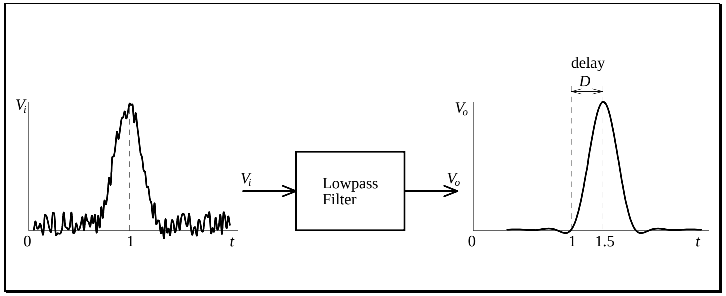


Figure 3A.4

To the pulse, the filter just looks like a delay. We can see that distortionless transmission through a filter is characterised by a constant delay of the input signal:

Distortionless transmission defined

$$\boxed{v_o(t) = Kv_i(t - D)} \tag{3A.37}$$

In Eq. (3A.37) we have also included the fact that all frequencies in the passband of the filter can have their amplitudes multiplied by a constant without affecting the waveshape. Also note that Eq. (3A.37) applies only to frequencies in the passband of a filter - we do not care about any distortion in a stopband.

We will now relate these distortionless transmission requirements to the phase response of the filter. From Fourier analysis we know that any periodic signal can be decomposed into an infinite summation of sinusoidal signals. Let one of these be:

$$v_i = A\cos(\omega t + \phi) \tag{3A.38}$$

From Eq. (3A.37), the output of the filter will be:

$$\begin{aligned} v_o &= KA\cos[\omega(t-D)+\phi] \\ &= KA\cos(\omega t - \omega D + \phi) \end{aligned} \tag{3A.39}$$

The input and output signals differ only by the gain K and the phase angle which is:

Distortionless transmission requires a linear phase

$$\boxed{\theta = -D\omega} \tag{3A.40}$$

That is, the phase response must be a straight line with negative slope that passes through the origin.

In general, the requirement for the phase response in the passband to achieve distortionless transmission through the filter is:

Group delay defined

$$\boxed{\frac{d\theta}{d\omega} = -D} \tag{3A.41}$$

The delay D in this case is referred to as the *group delay*. (This means the group of sinusoids that make up the waveshape have a delay of D).

The ideal lowpass filter can be expressed completely as:

$$\boxed{H(f) = K\operatorname{rect}\left(\frac{f}{2B}\right)e^{-j2\pi fD}} \tag{3A.42}$$

Example

We would like to determine the maximum frequency for which transmission is practically distortionless in the following simple filter:

Model of a short piece of co-axial cable, or twisted pair

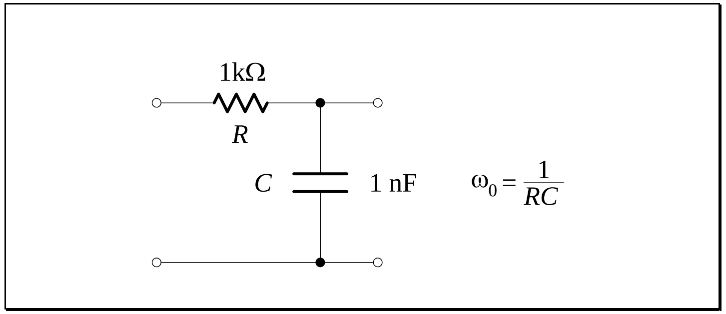


Figure 3A.5

We would also like to know the group delay caused by this filter.

We know the magnitude and phase response already:

$$|H(\omega)| = \frac{1}{\sqrt{1+(\omega/\omega_0)^2}} \tag{3A.43a}$$

$$\angle H(\omega) = -\tan^{-1}(\omega/\omega_0) \tag{3A.43b}$$

These responses are shown below:

The deviation from linear phase and constant magnitude for a simple first-order filter

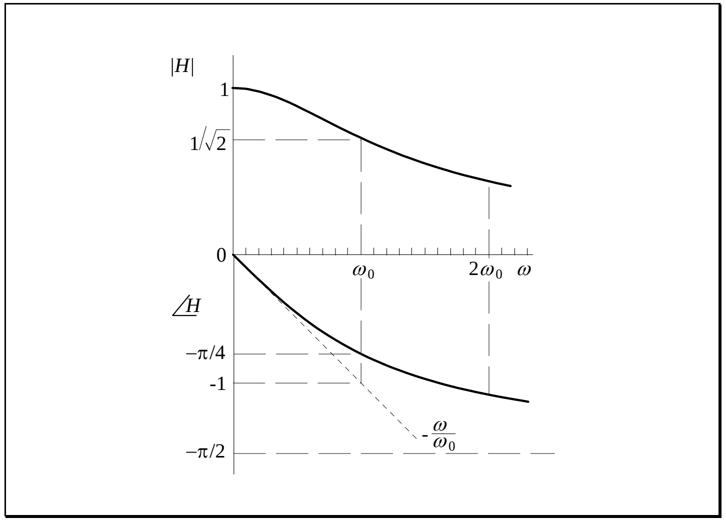


Figure 3A.6

Suppose we can tolerate a deviation in the magnitude response of 1% in the "passband". We then have:

$$\frac{1}{\sqrt{1+\left(\omega/\omega_0\right)^2}} \geq 0.99 \qquad \text{(3A.44)}$$

$$\omega \leq 0.1425\omega_0$$

Also, suppose we can tolerate a deviation in the delay of 1% in the "passband". We first find an expression for the delay:

$$D = -\frac{d\theta}{d\omega} = \frac{1}{1+\left(\omega/\omega_0\right)^2}\frac{1}{\omega_0} \qquad \text{(3A.45)}$$

and then impose the condition that the delay be within 1% of the delay at DC:

$$\frac{1}{1+\left(\omega/\omega_0\right)^2} \geq 0.99 \qquad \text{(3A.46)}$$

$$\omega \leq 0.1005\omega_0$$

We can see from Eqs. (3A.44) and (3A.46) that we must have $\omega \leq 0.1\omega_0$ for practically distortionless transmission. The delay for $\omega \leq 0.1\omega_0$, according to Eq. (3A.45), is approximately given by:

Approximate group delay for a first-order lowpass circuit

$$\boxed{D \approx \frac{1}{\omega_0}} \qquad \text{(3A.47)}$$

For the values shown in Figure 3A.6, the group delay is approximately 1 μs. In practice, variations in the magnitude transfer function up to the half-power frequency are considered tolerable (this is the bandwidth BW of the filter). Over this range of frequencies, the phase deviates from the ideal linear characteristic by at most $-\pi/4-(-1)=0.2146$ radians (see Figure 3A.6). Frequencies well below ω_0 are transmitted practically without distortion, but frequencies in the vicinity of ω_0 will suffer some distortion.

The ideal filter is unrealizable. To show this, take the inverse Fourier transform of the ideal filter's frequency response, Eq. (3A.42):

$$h(t) = 2BK\mathrm{sinc}\left[2B(t-D)\right] \qquad \text{(3A.48)}$$

An ideal filter is unrealizable because it is non-causal

It should be clear that the impulse response is not zero for $t<0$, and the filter is therefore not causal (how can there be a response to an impulse at $t=0$ before it is applied?). One way to design a real filter is simply to multiply Eq. (3A.48) by $u(t)$, the unit step.

Sampling

Sampling is one of the most important things we can do to a continuous-time signal – because we can then process it digitally

Sampling is one of the most important operations we can perform on a signal. Samples can be quantized and then operated upon digitally (digital signal processing). Once processed, the samples are turned back into a continuous-time waveform. (eg. CD, mobile phone!) Here we demonstrate how, if certain parameters are right, a sampled signal can be reconstructed from its samples almost perfectly.

Ideal sampling involves multiplying a waveform by a train of impulses. The weights of the impulses are the sample values to be used by a digital signal processor (computer). An ideal sampler is shown below:

An ideal sampler multiplies a signal by a uniform train of impulses

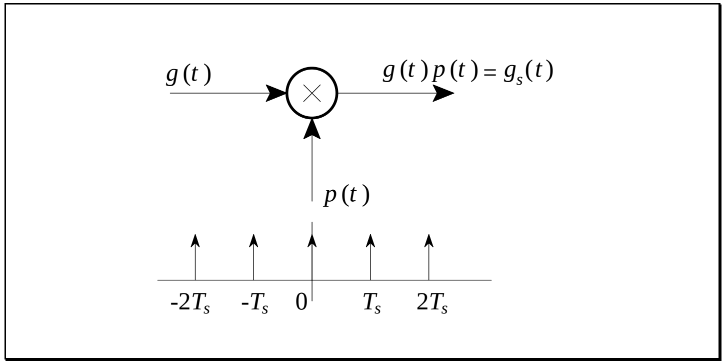


Figure 3A.7

Let $g(t)$ be a time domain signal. If we multiply it by $\sum_{k=-\infty}^{\infty}\delta(t-kT_s)$ we get an ideally sampled version:

$$g_s(t)=g(t)\cdot\sum_{k=-\infty}^{\infty}\delta(t-kT_s) \tag{3A.49}$$

In the time domain, we have:

An ideal sampler produces a train of impulses - each impulse is weighted by the original signal

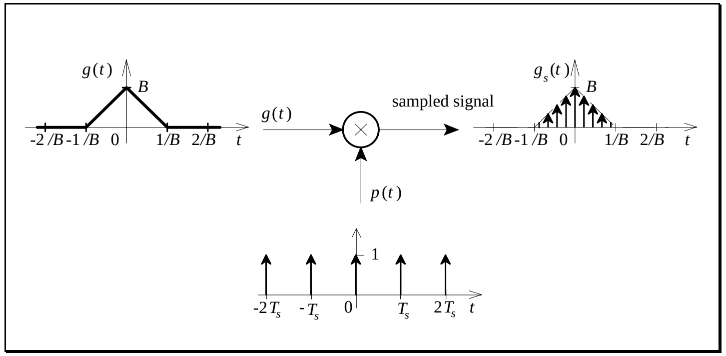


Figure 3A.8

Taking the Fourier transform of both sides of Eq. (3A.49):

$$\begin{aligned} G_s(f) &= G(f) * f_s\sum_{n=-\infty}^{\infty}\delta(f-nf_s) \\ &= f_s\sum_{n=-\infty}^{\infty}G(f-nf_s) \end{aligned} \tag{3A.50}$$

Graphically, in the frequency domain:

A sampled signal's spectrum is a scaled replica of the original, periodically repeated

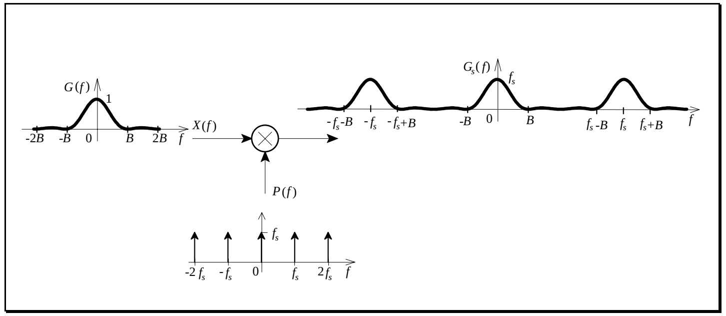


Figure 3A.9

Thus the Fourier transform of the sampled waveform is a scaled replica of the original, periodically repeated along the frequency axis. Spacing between repeats is equal to the sampling frequency (the inverse of the sampling interval).

Similarly, if we sample a continuous Fourier transform $G(f)$ by multiplying it with $\sum_{n=-\infty}^{\infty}\delta(f-nf_s)$ and take the inverse Fourier transform we get:

$$\mathrm{F}^{-1}\left[G_s(f)\right]=g(t)*T_s\sum_{k=-\infty}^{\infty}\delta(t-kT_s) \tag{3A.51}$$

which is a periodic repeat of the inverse transform of the original signal.

Sampling in one domain implies periodicity in the other

Thus, sampling in the frequency domain results in periodicity in the time domain. We already know this! We know a periodic time domain signal can be synthesised from sinusoids with frequencies nf_0, ie. has a transform consisting of impulses at frequencies nf_0.

We now see the general pattern: *Sampling in one domain implies periodicity in the other*.

Reconstruction

If a sampled signal $g_s(t)$ is applied to an ideal lowpass filter of bandwidth *B*, the only component of the spectrum $G_s(f)$ that is passed is just the original spectrum $G(f)$.

We recover the original spectrum by lowpass filtering

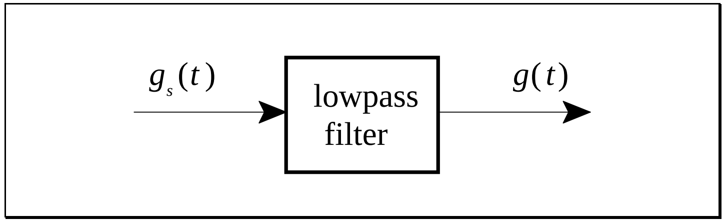


Figure 3A.10

Hence the output of the filter is equal to $g(t)$, which shows that the original signal can be completely and exactly reconstructed from the sampled waveform $g_s(t)$.

In the time domain:

A weighted train of impulses turns back into the original signal after lowpass filtering…

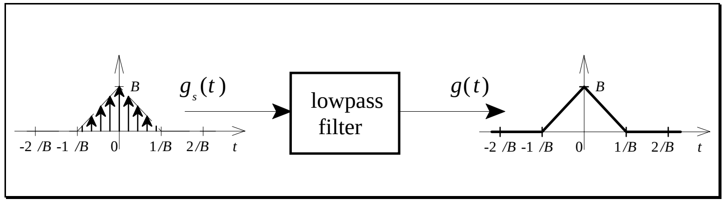


Figure 3A.11

and in the frequency domain:

but it's a much clearer operation in the frequency domain!

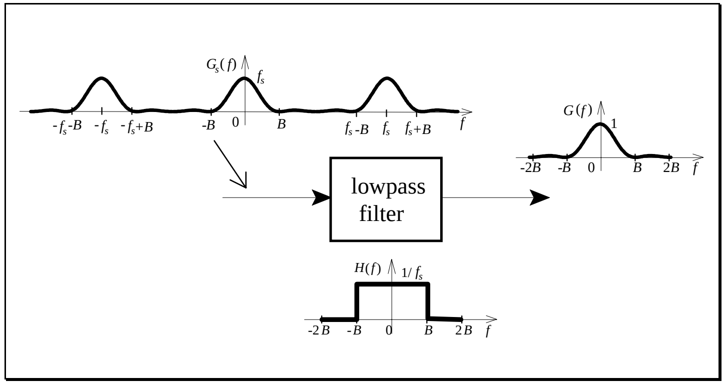


Figure 3A.12

We can't sample and reconstruct perfectly, but we can get close!

There are some limitations to perfect reconstruction though. One is that time-limited signals are not bandlimited (e.g. rect function). Any time-limited signal therefore cannot be perfectly reconstructed, since there is no sample rate high enough to ensure repeats of the original spectrum do not overlap. However, many signals are *essentially bandlimited*, which means spectral components higher than, say *B*, do not make a significant contribution to either the shape or energy of the signal.

Aliasing

We have to ensure no spectral overlap when sampling

We saw that sampling in one domain implies periodicy in the other. If the function being made periodic has an extent that is smaller than the period, there will be no resulting overlap and hence it will be possible to recover the continuous (unsampled) function by windowing out just one period from the domain displaying periodicy.

Nyquist's criterion is the formal expression of the above fact. It states:

Nyquists's criterion

Perfect reconstruction of a sampled signal is possible if the sampling rate is greater than twice the bandwidth of the signal being sampled

$$f_s > 2B \tag{3A.52}$$

Foldover frequency defined

Nyquist frequency defined

To avoid aliasing, we have to sample at a rate $f_s > 2B$. The frequency $f_s/2$ is called the spectral foldover frequency, and it is determined only by the *selected sample rate*, and it may be selected independently of the characteristics of the signal being sampled. The frequency *B* is termed the Nyquist frequency, and it is a function only of the signal and is independent of the selected sampling rate. *Do not confuse these two independent entities!* The Nyquist frequency is a *lower bound* for the foldover frequency in the sense that failure to *select* a foldover frequency at or above the Nyquist frequency will result in spectral aliasing and loss of the capability to reconstruct a continuous-time signal from its samples without error. The Nyquist frequency for a signal which is *not bandlimited* is infinity; that is, there is no finite sample rate that would permit errorless reconstruction of the continuous-time signal from its samples.

Nyquist rate defined

The Nyquist *rate* is defined as $f_N = 2B$, and is not to be confused with the similar term Nyquist *frequency*. The Nyquist rate is $2B$, whereas the Nyquist frequency is *B*. To prevent aliasing, we need to sample at a rate greater than the Nyquist rate, i.e. $f_s > f_N$.

To illustrate aliasing, consider the case where we have not selected the sample rate higher than twice the bandwidth of a lowpass signal:

An illustration of aliasing in the frequency-domain

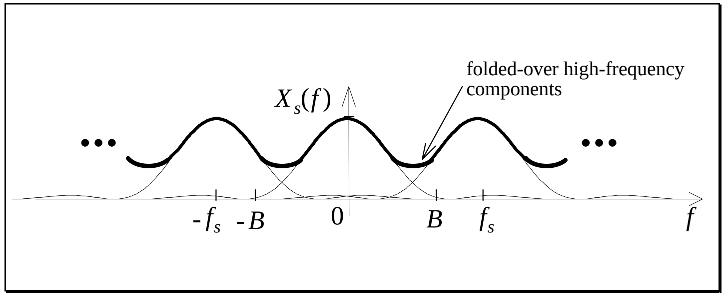


Figure 3A.13

If the sampled signal $x_s(t)$ is lowpass filtered with cutoff frequency *B*, the output spectrum of the filter will contain high-frequency components of $x(t)$ folded-over to low-frequency components.

How to avoid aliasing

To summarise – we can avoid aliasing by either:

1. Selecting a sample rate higher than twice the bandwidth of the signal (equivalent to saying that the foldover frequency is greater than the bandwidth of the signal); or

2. By bandlimiting (using a filter) the signal so that its bandwidth is less than half the sample rate.

Summary of the Sampling and Reconstruction Process

The sampling and reconstruction process in both the time-domain and frequency-domain

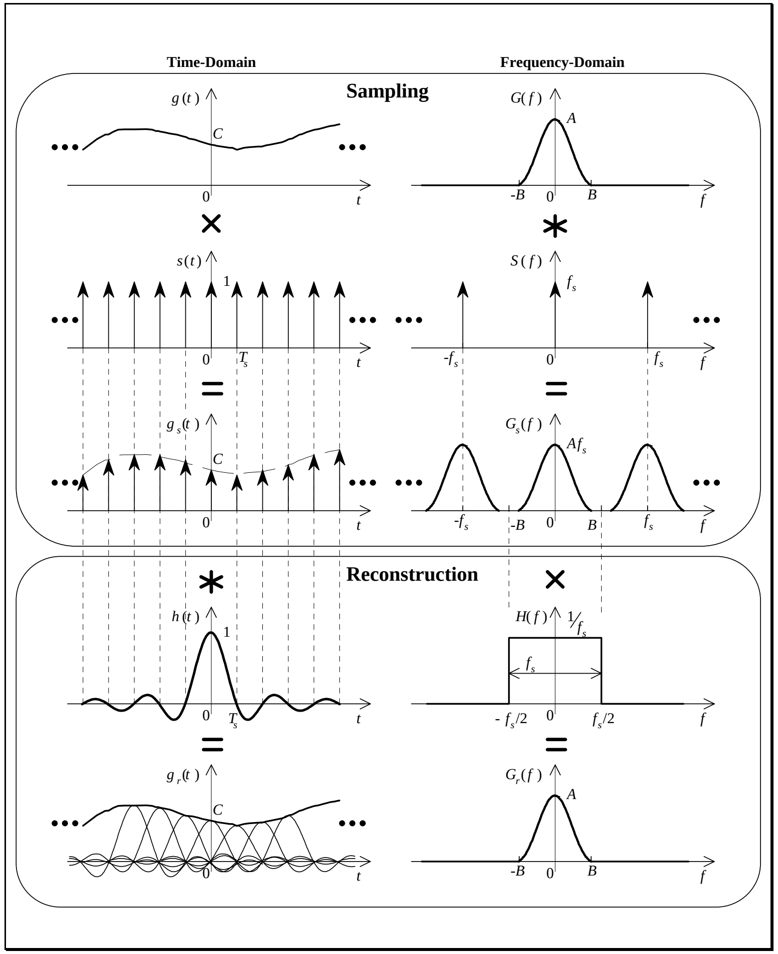


Figure 3A.14

Modulation

Another practical and very important application of Fourier analysis is when we consider an operation called *modulation*.

Let $x(t)$ be a signal such as an audio signal that is to be transmitted through a cable or the atmosphere. In amplitude modulation (AM), the signal modifies (or modulates) the *amplitude* of a *carrier* sinusoid $\cos(\omega_c t)$. In one form of AM transmission, the signal $x(t)$ and the carrier $\cos(\omega_c t)$ are simply multiplied together. The process is illustrated below:

Double side-band – suppressed carrier (DSB-SC) modulation

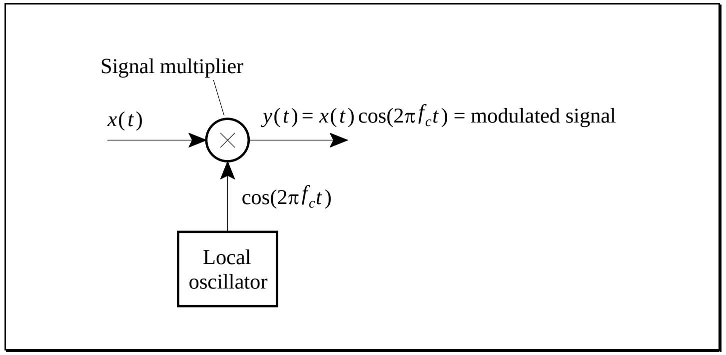


Figure 3A.15

The *local oscillator* in Figure 3A.15 is a device that produces the sinusoidal signal $\cos(\omega_c t)$. The multiplier is implemented with a non-linear device, and is usually an integrated circuit at low frequencies.

By the multiplication property of Fourier transforms, the output spectrum is obtained by *convolving* the spectrum of $x(t)$ with the spectrum of $\cos(2\pi f_c t)$. We now restate a very important property of convolution involving an impulse:

$$\boxed{X(f) * \delta(f - f_0) = X(f - f_0)} \tag{3A.53}$$

The output spectrum of the modulator is therefore:

DSB-SC up-translates the baseband spectrum

$$\boxed{\begin{aligned} Y(f) &= X(f) * \frac{1}{2}\left[\delta(f - f_c) + \delta(f + f_c)\right] \\ &= \frac{1}{2}\left[X(f - f_c) + X(f + f_c)\right] \end{aligned}} \qquad (3A.54)$$

The spectrum of the modulated signal is a replica of the signal spectrum but "shifted up" in frequency. If the signal has a bandwidth equal to B then the modulated signal spectrum has an upper sideband from f_c to $f_c + B$ and a lower sideband from $f_c - B$ to f_c, and the process is therefore called *double-sideband transmission*, or DSB transmission for short. An example of modulation is given below in the time-domain:

DSB-SC in the time-domain

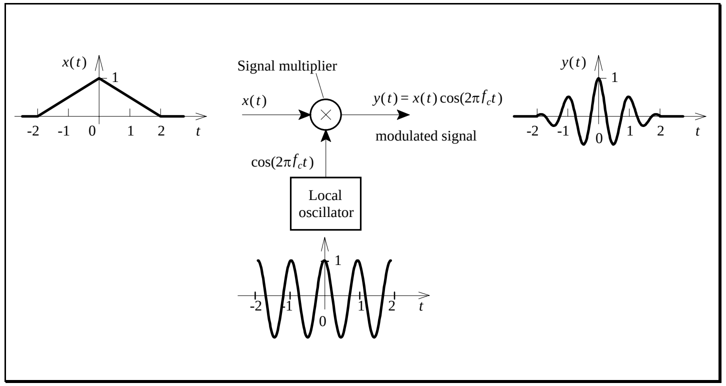


Figure 3A.16

And in the frequency domain:

DSB-SC in the frequency-domain

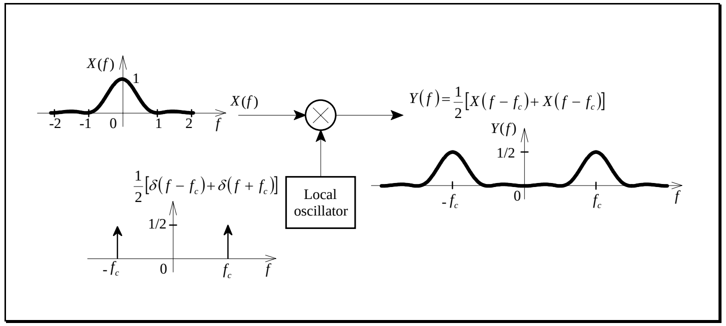


Figure 3A.17

Modulation lets us share the spectrum, and achieves practical propagation

The higher frequency range of the modulated signal makes it possible to achieve good propagation in transmission through a cable or the atmosphere. It also allows the "spectrum" to be shared by independent users – e.g. radio, TV, mobile phone etc.

Demodulation

The reconstruction of $x(t)$ from $x(t)\cos(\omega_c t)$ is called *demodulation*. There are many ways to demodulate a signal, here we will consider one common method called *synchronous* or *coherent demodulation*.

Coherent demodulation of a DSB-SC signal

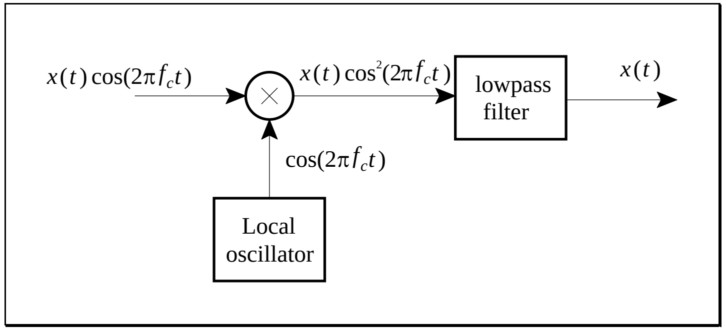


Figure 3A.18

Coherent demodulation requires a carrier at the receiver that is synchronized with the transmitter

The first stage of the demodulation process involves applying the modulated waveform $x(t)\cos(\omega_c t)$ to a multiplier. The other signal applied to the multiplier is a local oscillator which is assumed to be synchronized with the carrier signal $\cos(\omega_c t)$, i.e. there is no phase shift between the carrier and the signal generated by the local oscillator.

The output of the multiplier is:

$$\begin{aligned}&\frac{1}{2}\left[X\left(f-f_c\right)+X\left(f+f_c\right)\right]*\frac{1}{2}\left[\delta\left(f-f_c\right)+\delta\left(f+f_c\right)\right]\\&=\frac{1}{2}X\left(f\right)+\frac{1}{4}\left[X\left(f-2f_c\right)+X\left(f+2f_c\right)\right]\end{aligned} \tag{3A.55}$$

$x(t)$ can be "extracted" from the output of the multiplier by lowpass filtering with a cutoff frequency of B Hz and a gain of 2.

Another way to think of this is in the time-domain:

$$x(t)\cos^2\left(2\pi f_c t\right)=x(t)\frac{1}{2}\left(1+\cos\left(4\pi f_c t\right)\right) \tag{3A.56}$$

Therefore, it is easy to see that lowpass filtering, with a gain of 2, will produce $x(t)$. An example of demodulation in the time-domain is given below:

Demodulation in the time-domain

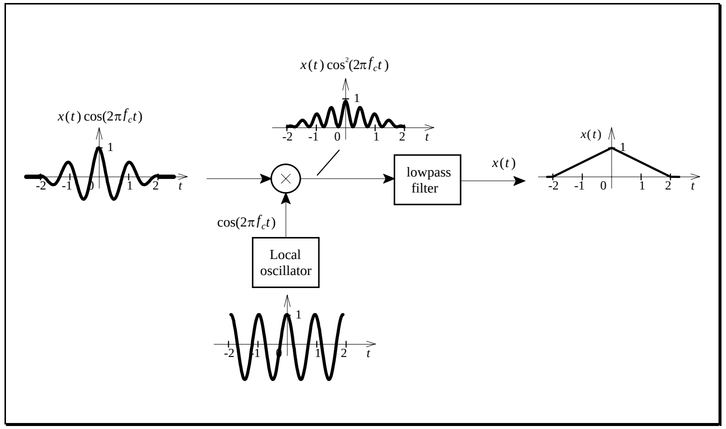


Figure 3A.19

The operation of demodulation is best understood in the frequency domain:

Demodulation in the frequency-domain

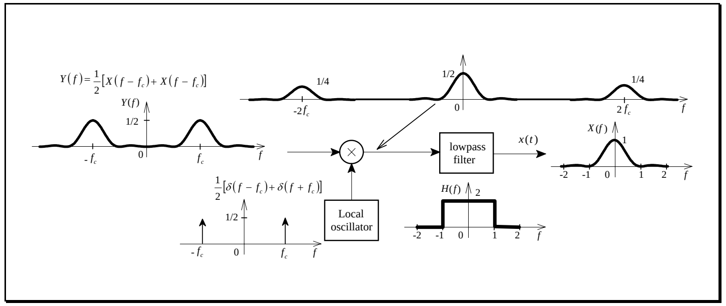


Figure 3A.20

Summary of DBS-SC Modulation and Demodulation

The DSB-SC modulation and demodulation process in both the time-domain and frequency-domain

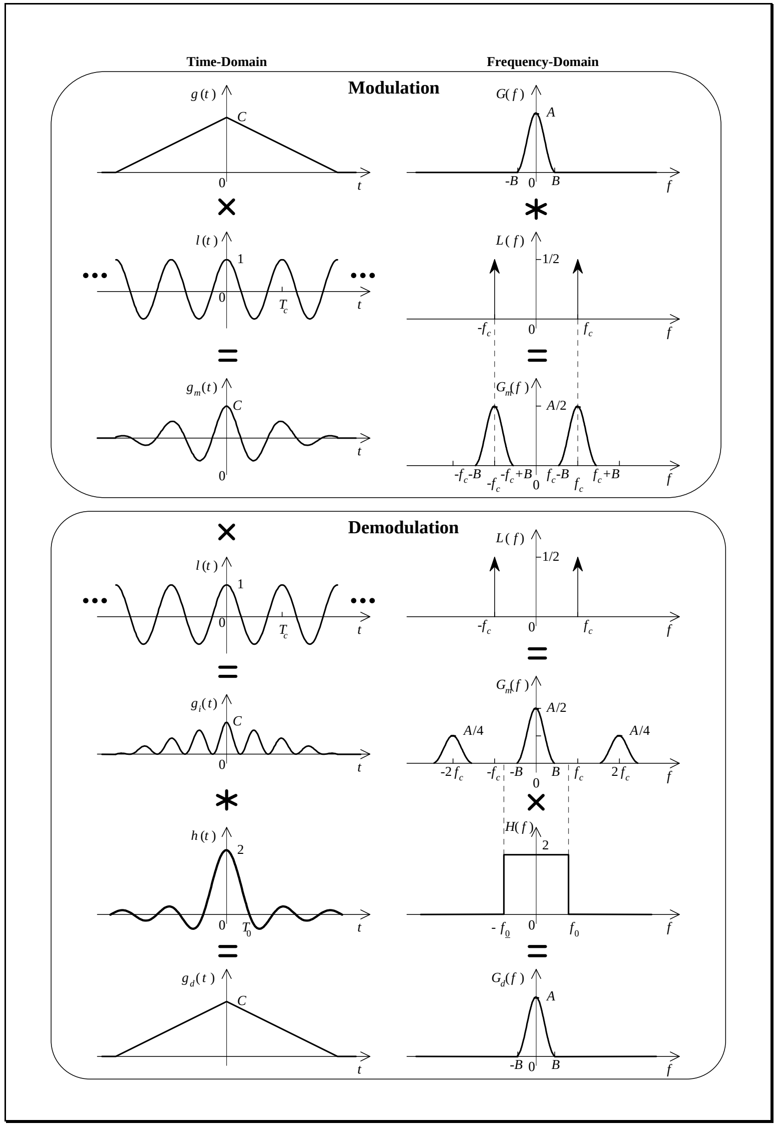


Figure 3A.21

Finding the Fourier Series of a Periodic Function From the Fourier Transform of a Single Period

The quick way to determine Fourier series coefficients

It is usually easier to find the Fourier transform of a single period than performing the integration needed to find Fourier series coefficients (because all the standard Fourier properties can be used). This method allows the Fourier series coefficients to be determined directly from the Fourier transform, provided the period is known. Don’t forget, only periodic functions have Fourier series representation.

Suppose we draw one period of a periodic waveform:

A single period

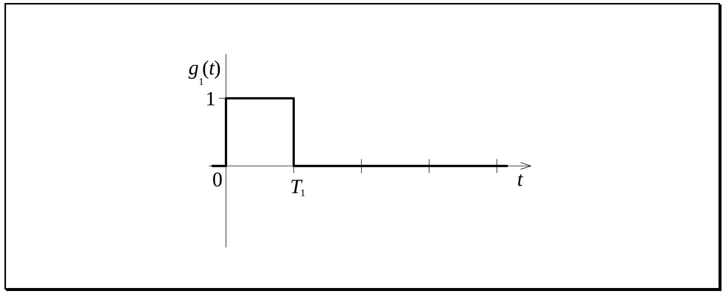


Figure 3A.22

We can create the periodic version by convolving $g_1(t)$ with a train of unit impulse functions with spacing equal to the period, T_0:

Convolved with a uniform impulse train

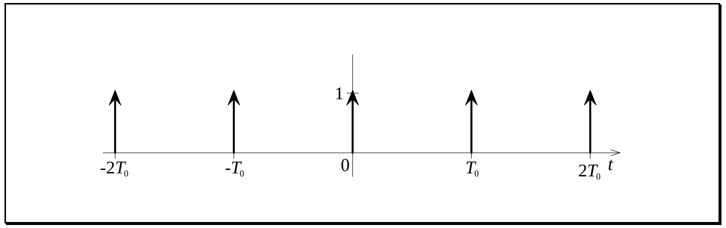


Figure 3A.23

That is, we need to convolve $g_1(t)$ with $\sum_{k=-\infty}^{\infty}\delta(t-kT_0)$. Thus, $g_p(t)$, the periodic version is:

$$g_p(t)=g_1(t)*\sum_{k=-\infty}^{\infty}\delta(t-kT_0) \tag{3A.57}$$

gives the periodic waveform

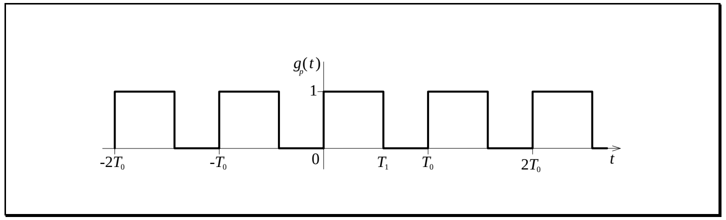


Figure 3A.24

Using the convolution multiplication rule:

$$F\left[g_p(t)\right]=F\left[g_1(t)\right]\cdot F\left[\sum_{k=-\infty}^{\infty}\delta(t-kT_0)\right]$$

$$=G_1(f)\cdot f_0\sum_{n=-\infty}^{\infty}\delta(f-nf_0) \tag{3A.58}$$

In words, the Fourier transform of the periodic signal consists of impulses located at harmonics of $f_0=1/T_0$, whose weights are:

Fourier series coefficients from the Fourier transform of one period

$$\boxed{G_n=f_0G_1(nf_0)} \tag{3A.59}$$

These are the Fourier series coefficients.

In Figure 3A.24 we have:

$$G_n=f_0T_1\operatorname{sinc}(nf_0T_1)e^{-j\pi nf_0T_1} \tag{3A.60}$$

Graphically, the operation indicated by Eq. (3A.58) takes the original spectrum and multiplies it by a train of impulses – effectively creating a weighted train of impulses:

Sampling in the frequency domain produces a periodic waveform in the time-domain

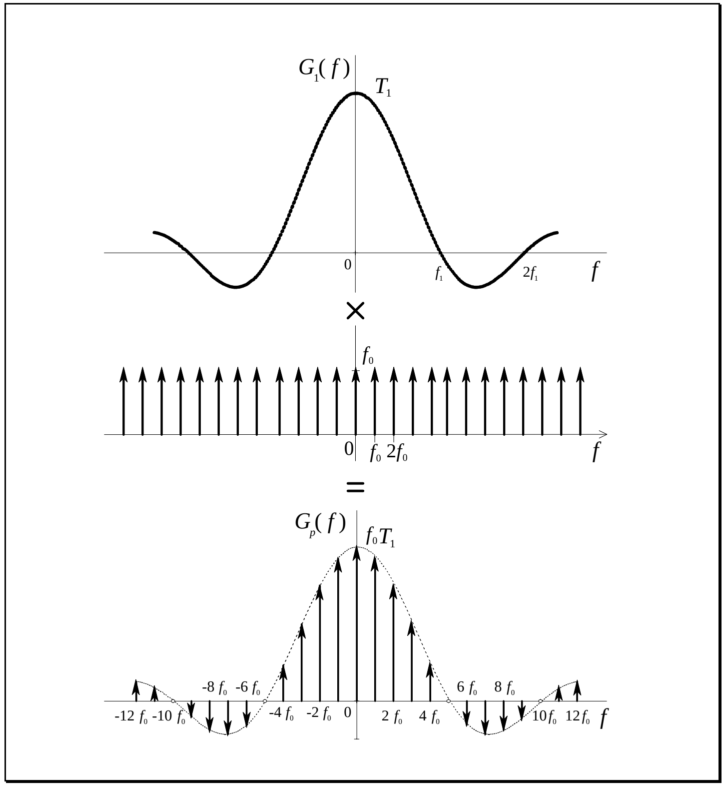


Figure 3A.25

According to Eq. (3A.59), the Fourier series coefficients are just the weights of the impulses in the spectrum of the periodic function. To get the n^{th} Fourier series coefficient, use the weight of the impulse located at nf_0.

This is in perfect agreement with the concept of a continuous spectrum. Each frequency has an infinitesimal amplitude sinusoid associated with it. If an impulse exists at a certain frequency, then there is a finite amplitude sinusoid at that frequency.

Remember that pairs of impulses in a spectrum represent a sinusoid in the time-domain

Windowing in the Time Domain

Some practical effects of looking at signals over a finite time

Often we wish to deal with only a segment of a signal, say from $t=0$ to $t=T$. Sometimes we have no choice, as this is the only part of the signal we have access to - our measuring instrument has restricted us to a "window" of duration T beginning at $t=0$. Outside this window the signal is forced to be zero. How is the signal's Fourier transform affected when the signal is viewed through a window?

Windowing defined

Windowing in the time domain is equivalent to multiplying the original signal $g(t)$ by a function which is non-zero over the window interval and zero elsewhere. So the Fourier transform of the windowed signal is the original signal convolved with the Fourier transform of the window.

Example

Find the Fourier transform of $\sin(2\pi t)$ when it is viewed through a rectangular window from 0 to 1 second:

A rectangular window applied to a sinusoid

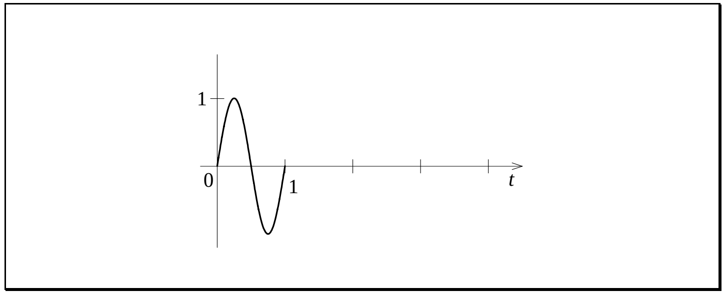


Figure 3A.26

The viewed signal is:

$$g_w(t)=\sin(2\pi t)\operatorname{rect}(t-0.5) \tag{3A.61}$$

The Fourier transform will be:

$$\begin{aligned}
&\mathcal{F}[\sin(2\pi t)] * \mathcal{F}[\operatorname{rect}(t-0.5)] \\
&= \left[\frac{-j}{2}\delta(f-1)+\frac{j}{2}\delta(f+1)\right] * \operatorname{sinc}(f)e^{-j\pi f} \\
&= -\frac{j}{2}\operatorname{sinc}(f-1)e^{-j\pi(f-1)}+\frac{j}{2}\operatorname{sinc}(f+1)e^{-j\pi(f+1)}
\end{aligned} \tag{3A.62}$$

Graphically, the magnitude spectrum of the windowed signal is:

Spectrum of a rectangular windowed sinusoid

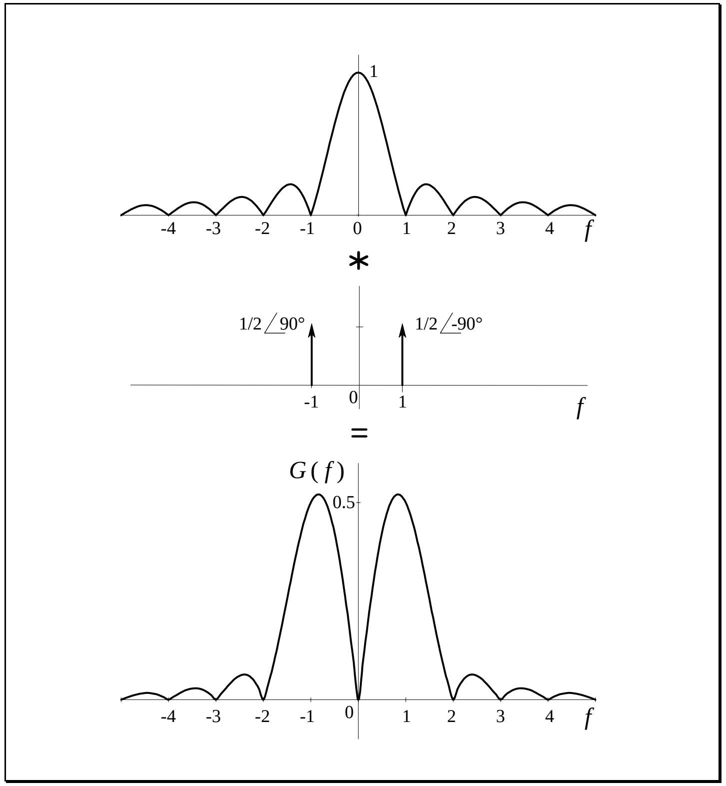


Figure 3A.27

If the window were changed to 4 seconds, we would then have:

The longer we look, the better the spectrum

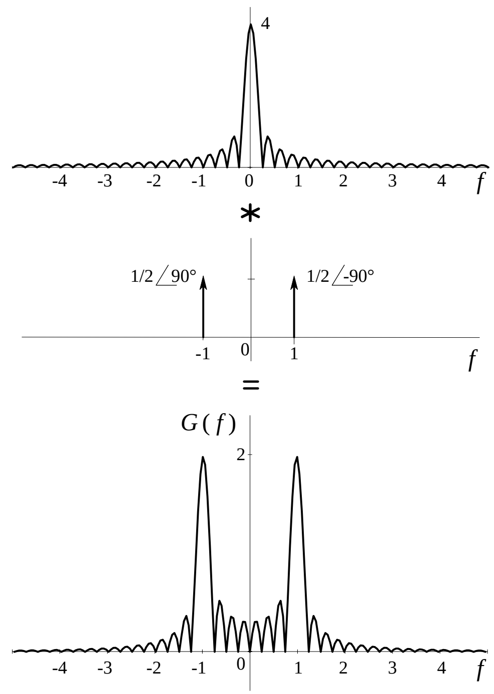


Figure 3A.28

Obviously, the longer the window, the more accurate the spectrum becomes.

Example

Find the Fourier transform of $\operatorname{sinc}(t)$ when it is viewed through a window from -2 to 2 seconds:

Viewing a sinc function through a rectangular window

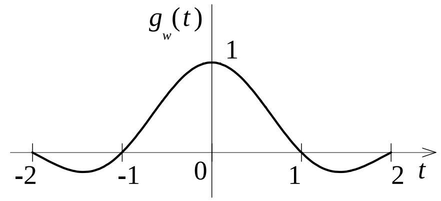


Figure 3A.29

We have:

$$g_w(t) = \operatorname{sinc}(t)\operatorname{rect}(t/4) \tag{3A.63}$$

and:

$$\mathcal{F}\left[g_w(t)\right] = \operatorname{rect}(f) * 4\operatorname{sinc}(4f) \tag{3A.64}$$

Graphically:

Ripples in the spectrum caused by a rectangular window

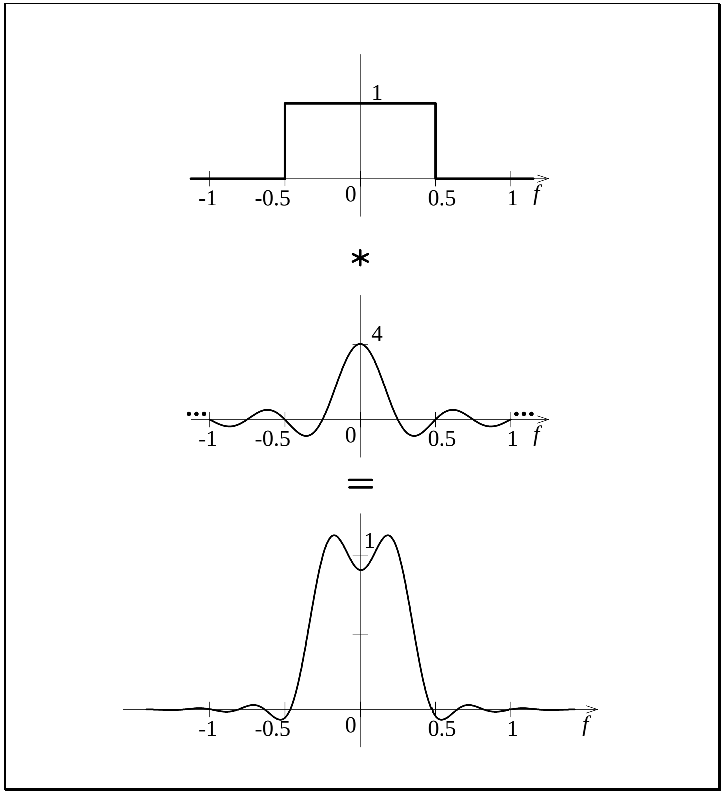


Figure 3A.30

We see that windowing in the time domain by T produces ripples in the frequency domain with an approximate spacing of $2/T$ between peaks. (In a magnitude spectrum, there would be a peak every $1/T$).

Practical Multiplication and Convolution

We have learned about multiplication and convolution as mathematical operations that we can apply in either the time or frequency domain. Remember however that the time and frequency domain are just two ways of describing the same signal - a time varying voltage that we measure across two terminals. What do we need to physically do to our signal to perform an operation equivalent to say multiplying its frequency domain by some function?

Two physical operations we can do on signals are *multiplication* (with another signal) and *filtering* (with a filter with a defined transfer function).

Multiplication in the time-domain is convolution in the frequency-domain

Multiplication in time domain ≡ convolution in frequency domain

You can buy a "4 quadrant multiplier" as an IC from any electronics supplier. Depending on the bandwidth of the signal they can handle, the price varies from several dollars to over a hundred dollars. They have a pair of terminals for the two signals to be multiplied together. At higher frequencies any non-linear device can be used for multiplication.

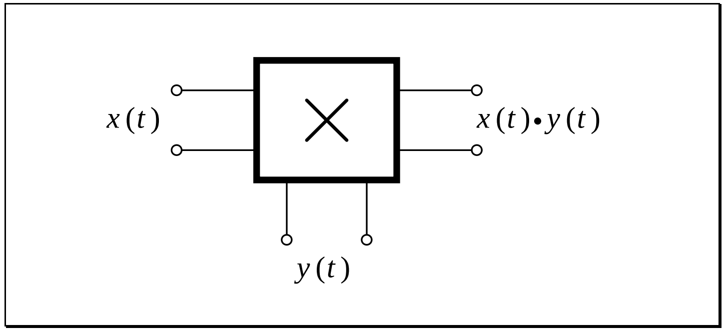


Figure 3A.31

Time domain output is *x*(*t*) multiplied with *y*(*t*)	(3A.65a)
Frequency domain output is *X*(*f*) convolved with *Y*(*f*)	(3A.65b)

Convolution in the time-domain is multiplication in the frequency-domain

Convolution in time domain ≡ multiplication in frequency domain

You can design a network, either passive or active that performs as a filter. Its characteristics can be specified by its transfer function (a plot of magnitude vs frequency and a second plot of phase vs frequency or a complex expression of frequency) or equivalently by its impulse response which is the inverse Fourier transform of the transfer function.

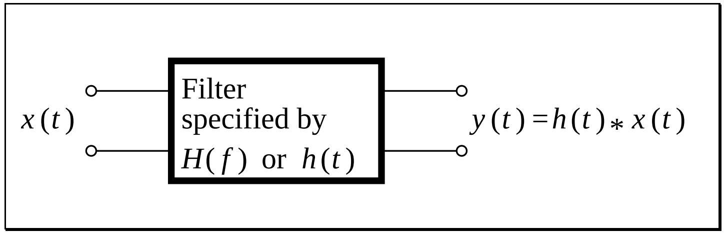


Figure 3A.32

Time domain output is $x(t)$ convolved with $h(t)$	(3A.66)
Frequency domain output is $X(f)$ multiplied with $H(f)$	(3A.66b)

What does a Filter do to a Signal?

Filtering a periodic signal

Passing a periodic signal through a filter will distort the signal (so called linear distortion) because the filter will change the relative amplitudes and phases of the sinusoids that make up its Fourier series. Once we can calculate the amplitude change and phase shift that the filter imposes on an arbitrary sinusoid we are in a position to find out how each sinusoid is affected, and hence synthesise the filtered waveform. In general, the output Fourier transform is just the input Fourier transform multiplied by the filter frequency response $H(f)$.

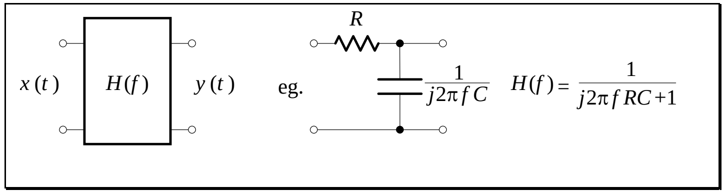


Figure 3A.33

$$\boxed{Y(f)=H(f)\cdot X(f)} \tag{3A.67}$$

For a periodic signal:

$$X(f)=\sum_{n=-\infty}^{\infty} X_n\delta(f-nf_0) \tag{3A.68}$$

and therefore:

Spectrum of a filtered periodic signal

$$\boxed{Y(f)=\sum_{n=-\infty}^{\infty} H(nf_0)X_n\delta(f-nf_0)} \tag{3A.69}$$

Example

A periodic signal has a Fourier series as given in the following table:

Input Fourier series table

N	Amplitude	Phase
0	1	
1	2	-30°
2	3	-90°

What is the Fourier series of the output if the signal is passed through a filter with transfer function $H(f) = j4\pi f$? The period is 2 seconds.

There are 3 components in the input signal with frequencies 0, 0.5 Hz and 1 Hz. The complex gain of the filter at each frequency is:

Filter gain table

n	$H(f)$	Gain	Phase shift
0	0	0	
1	$j4\pi \times 0.5$	2π	90°
2	$j4\pi \times 1$	4π	90°

Hence, the output Fourier series table is:

Output Fourier series table

n	Amplitude	Phase
0	0	
1	4π	60°
2	12π	0°

Example

Suppose the same signal was sent through a filter with transfer function as sketched:

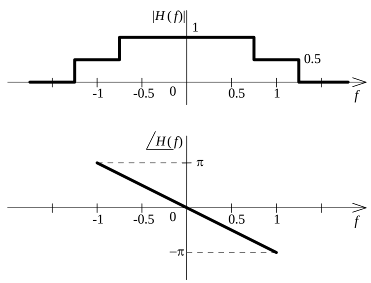


Figure 3A.34

The output Fourier series is:

Output Fourier series table

n	Amplitude	Phase
0	1	
1	2	-120°
2	1.5	-270°

How to Use MATLAB® to Check Fourier Series Coefficients

MATLAB® is a software package that is particularly suited to signal processing. It has instructions that will work on vectors and matrices. A vector can be set up which gives the samples of a signal. Provided the sample spacing meets the Nyquist criterion the instruction `G=fft(g)` returns a vector containing N times the Fourier series coefficients, where $\mathtt{G}(1)=N\cdot G_0$ is the DC term, $\mathtt{G}(2)=N\cdot G_1$, $\mathtt{G}(3)=N\cdot G_2$ etc. and $\mathtt{G}(\mathtt{N})=N\cdot G_{-1}$, $\mathtt{G}(\mathtt{N-1})=N\cdot G_{-2}$ etc. where N is the size of the vector. `G=ifft(g)`does the inverse Fourier transform.

Example

Suppose we want to find the Fourier series coefficients of $g(t)=\cos(2\pi t)$. Note period = 1 s.

Step 1

Choose sample frequency - since highest frequency present is 1 Hz, choose 4 Hz (minimum is > 2 Hz).

Step 2

Take samples over one period starting at $t=0$. Note $N=4$.

$$\mathtt{g}=\begin{bmatrix}1 & 0 & -1 & 0\end{bmatrix}$$

Step 3

Find $\mathtt{G}=\mathtt{fft}(\mathtt{g})=\begin{bmatrix}0 & 2 & 0 & 2\end{bmatrix}$.

Hence $G_0=0$, $G_1=2/4=1/2$, $G_{-1}=2/4=1/2$. $G_{\pm 2}$ should be zero if the Nyquist criterion is met. These are in fact the Fourier series coefficients of $g(t)$.

Example

Find the Fourier series coefficients of a 50% duty cycle square wave.

Step 1

In this case the spectrum is ∞ so we can never choose f_s high enough. There is always some error. Suppose we choose 8 points of one cycle.

Step 2

$$\mathtt{g}=\begin{bmatrix}1 & 1 & 0.5 & 0 & 0 & 0 & 0.5 & 1\end{bmatrix}$$

Note: if samples occur on transitions, input half way point.

Step 3

$$\mathtt{G}=\mathtt{fft}(\mathtt{g})=\begin{bmatrix}4 & 2.4142 & 0 & -0.4142 & 0 & 0.4142 & 0 & 2.4142\end{bmatrix}$$

Therefore $G_1=2.4142/8=0.3015$. The true value is 0.3183. Using 16 points, $G_1=0.3142$.

You should read Appendix A – The Fast Fourier Transform, and look at the example MATLAB® code in the "FFT - Quick Reference Guide" for more complicated and useful examples of setting up and using the FFT.

Summary

- Sinusoidal inputs to linear time-invariant systems yield sinusoidal outputs. The output sinusoid is related to the input sinusoid by a complex-valued function known as the *frequency response*, $H(f)$.
- The frequency response of a system is just the Fourier transform of the impulse response of the system. That is, the impulse response and the frequency response form a Fourier transform pair: $h(t) \Leftrightarrow H(f)$.
- The frequency response of a system can be obtained directly by performing analysis in the *frequency domain*.
- The output of an LTI system due to *any* input signal is obtained most easily by considering the spectrum: $Y(f) = H(f)X(f)$. This expresses the important property: convolution in the time-domain is equivalent to multiplication in the frequency-domain.
- Filters are devices best thought about in the frequency-domain. They are *frequency* selective devices, changing both the *magnitude* and *phase* of frequency components of an input signal to produce an output signal.
- *Linear* phase is desirable in filters because it produces a constant *delay*, thereby preserving waveshape.
- *Sampling* is the process of converting a continuous-time signal into a discrete-time signal. It is achieved, in the ideal case, by multiplying the signal by a train of impulses.
- *Reconstruction* is the process of converting signal samples back into a continuous-time signal. It is achieved by passing the samples through a lowpass filter.
- *Aliasing* is an effect of sampling where spectral overlap occurs, thus destroying the ability to later reconstruct the signal. It is caused by not meeting the *Nyquist criterion*: $f_s > 2B$.

- *Modulation* shifts a baseband spectrum to a higher frequency range. It is achieved in many ways – the simplest being the multiplication of the signal by a carrier sinusoid.
- *Demodulation* is the process of returning a modulated signal to the baseband. Modulation and demodulation form the basis of modern communication systems.
- *Fourier series coefficients* can be obtained from the Fourier transform of one period of the signal by the formula: $G_n = f_0 G_1(nf_0)$.
- Using finite-duration signals in the time-domain is called *windowing*. Windowing affects the spectrum of the original signal.

References

Haykin, S.: *Communication Systems*, John-Wiley & Sons, Inc., New York, 1994.

Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB®*, Prentice-Hall, 1997.

Lathi, B. P.: *Modern Digital and Analog Communication Systems*, Holt-Saunders, Tokyo, 1983.

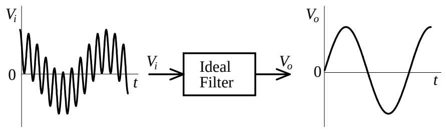
Quiz

Encircle the correct answer, cross out the wrong answers. [one or none correct]

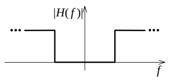
1.
The convolution of $x(t)$ and $y(t)$ is given by:

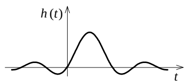
(a) $\int_{-\infty}^{\infty} x(\tau)y(t-\tau)d\tau$ (b) $\int_{-\infty}^{\infty} x(\tau)y(\tau-t)d\tau$ (c) $\int_{-\infty}^{\infty} X(\lambda)Y(f-\lambda)d\lambda$

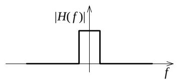
2.



The Fourier transform of the impulse response of the filter resembles:

(a) 

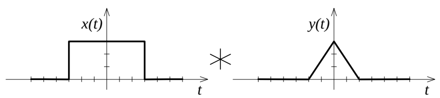
(b) 

(c) 

3.
The Fourier transform of one period of a periodic waveform is $G(f)$. The Fourier series coefficients, G_n, are given by:

(a) $nf_0G(f_0)$ (b) $G(nf_0)$ (c) $f_0G(nf_0)$

4.



The peak value of the convolution is:

(a) 9 (b) 4.5 (c) 6

5.
The scaling property of the Fourier transform is:

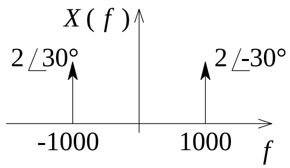
(a) $g(at) \Leftrightarrow G\left(\frac{f}{a}\right)$ (b) $g(at) \Leftrightarrow \frac{1}{|a|}G(f)$ (c) $ag(t) \Leftrightarrow aG(f)$

Answers: 1. a 2. c 3. c 4. x 5. x

Exercises

1.

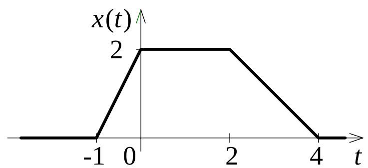
Write an expression for the time domain representation of the voltage signal with double sided spectrum given below:



What is the power of the signal?

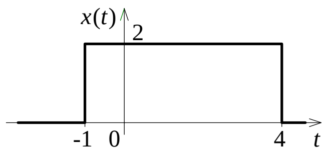
2.

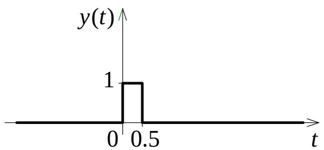
Use the integration and time shift rules to express the Fourier transform of the pulse below as a sum of exponentials.



3.

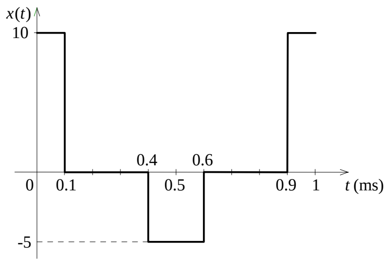
Sketch the convolution of the two functions shown below.





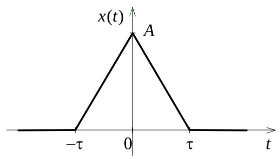
4.

Calculate the magnitude and phase of the 4 kHz component in the spectrum of the periodic pulse train shown below. The pulse repetition rate is 1 kHz.



5.

By relating the triangular pulse shown below to the convolution of a pair of identical rectangular pulses, deduce the Fourier transform of the triangular pulse:



6.

The pulse $x(t) = 2B\operatorname{sinc}(2Bt)\operatorname{rect}(Bt/8)$ has ripples in the amplitude spectrum. What is the spacing in frequency between positive peaks of the ripples?

7.

A signal is to be sampled with an ideal sampler operating at 8000 samples per second. Assuming an ideal low pass antialiasing filter, how can the sampled signal be reconstituted in its original form, and under what conditions?

8.

A train of impulses in one domain implies what in the other?

9.

The following table gives information about the Fourier series of a periodic waveform, $g(t)$, which has a period of 50 ms.

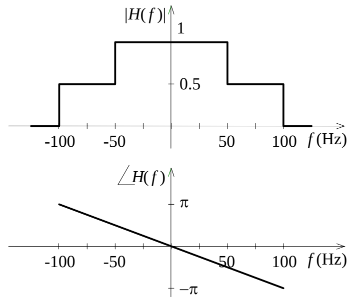
Table 1

Harmonic #	Amplitude	Phase (º)
0	1	
1	3	-30
2	1	-30
3	1	-60
4	0.5	-90

(a) Give the frequencies of the fundamental and the 2nd harmonic. What is the signal power, assuming that $g(t)$ is measured across $50\ \Omega$?

(b) Express the third harmonic as a pair of counter rotating phasors. What is the value of the third harmonic at $t = 20$ ms?

(c) The periodic waveform is passed through a filter with transfer function $H(f)$ as shown below.



Draw up a table in the same form as Table 1 of the Fourier series of the output waveform. Is there a DC component in the output of the amplifier?

10.

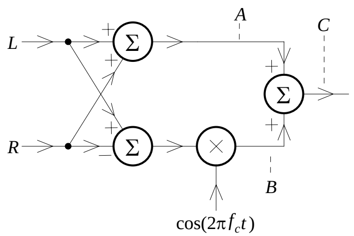
A signal, bandlimited to 1 kHz, is sampled by multiplying it with a rectangular pulse train with repetition rate 4 kHz and pulse width 50 μs. Can the original signal be recovered without distortion, and if so, how?

11.

Sketch the Fourier transform of the waveform $g(t) = (1 + \cos(2\pi t))\sin(20\pi t)$.

12.

A scheme used in stereophonic FM broadcasting is shown below:



The input to the left channel (L) is a 1 kHz sinusoid, the input to the right channel (R) is a 2 kHz sinusoid. Draw the spectrum of the signal at points A, B and C if $f_c = 38 \text{ kHz}$.

13.

Draw a block diagram of a scheme that could be used to recover the left (L) and right (R) signals of the system shown in Question 12 if it uses the signal at C as the input.

14.

A signal is to be analysed to identify the relative amplitudes of components which are known to exist at 9 kHz and 9.25 kHz. To do the analysis a digital storage oscilloscope takes a record of length 2 ms and then computes the Fourier series. The 18$^{\text{th}}$ harmonic thus computed can be non-zero even when no 9 kHz component is present in the input signal. Explain.

15.

Use MATLAB® to determine the output of a simple RC lowpass filter subjected to a square wave input given by:

$$x(t) = \sum_{n=-\infty}^{\infty} \operatorname{rect}(t - 2n)$$

for the cases: $1/RC = 1$, $1/RC = 10$, $1/RC = 100$. Plot the time domain from $t = -3$ to 3 and Fourier series coefficients up to the 50$^{\text{th}}$ harmonic for each case.

Lecture 3B – The Laplace Transform

Laplace transform. Region of convergence. Properties of the Laplace transform. Evaluation of inverse Laplace transforms. Solutions of integro-differential equations. System transfer function. Block diagrams. Block-diagram reduction.

The Laplace Transform

The Laplace transform is a generalisation of the Fourier transform. We need it for two reasons:

The need for the Laplace transform

1. there are some functions of interest, such as the ramp function $r(t) = tu(t)$ which do not have a Fourier transform.

2. we wish to determine a system's response from a specific time, $t = 0$, and also include any initial conditions in the system's response.

The Laplace transform fulfils both of these objectives. By adding an exponential *convergence factor* to the definition of the Fourier transform, we get:

$$X(\sigma + j\omega) = \int_{-\infty}^{\infty} x(t) e^{-(\sigma + j\omega)t} dt \tag{3B.1}$$

Letting $s = \sigma + j\omega$, this becomes:

The two-sided Laplace transform

$$X(s) = \int_{-\infty}^{\infty} x(t) e^{-st} dt \tag{3B.2}$$

which is called the *two-sided Laplace transform*. By making the integral only valid for time $t \geq 0$, we can incorporate initial conditions into the *s*-domain description of signals and systems. This is termed the *one-sided Laplace transform*, which will be referred to simply as the Laplace transform:

The one-sided Laplace transform

$$\boxed{X(s) = \int_{0^-}^{\infty} x(t) e^{-st} dt} \tag{3B.3}$$

The Laplace transform variable, s, is termed *complex frequency* – it is a complex number in the complex plane:

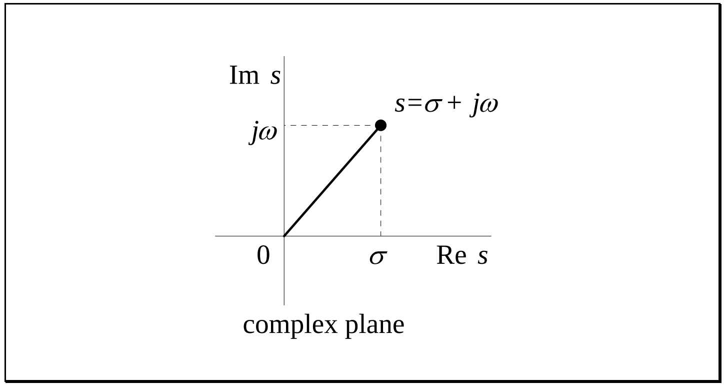


Figure 3B.1

The inverse Laplace transform is defined by:

The inverse Laplace transform

$$x(t)=\frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty}X(s)e^{st}\,ds \tag{3B.4}$$

It is common practice to use a bidirectional arrow to indicate a Laplace transform pair, as follows:

The Laplace transform pair defined

$$x(t)\leftrightarrow X(s) \tag{3B.5}$$

Region of Convergence (ROC)

The *region of convergence* (ROC) for the Laplace transform $X(s)$, is the set of values of s (the region in the complex plane) for which the integral in Eq. (3B.3) converges.

Example

To find the Laplace transform of a signal $x(t)=e^{-at}u(t)$ and its ROC, we substitute into the definition of the Laplace transform:

$$X(s)=\int_{0^-}^{\infty}e^{-at}u(t)e^{-st}\,dt \tag{3B.6}$$

Because $u(t)=0$ for $t<0$ and $u(t)=1$ for $t\geq 0$,

$$\begin{aligned}X(s)&=\int_{0^-}^{\infty}e^{-at}e^{-st}\,dt=\int_{0^-}^{\infty}e^{-(s+a)t}\,dt\\&=-\frac{1}{s+a}e^{-(s+a)t}\bigg|_{0}^{\infty}\end{aligned} \tag{3B.7}$$

Note that s is complex and as $t\to\infty$, the term $e^{-(s+a)t}$ does not necessarily vanish. Here we recall that for a complex number $z=\alpha+j\beta$,

$$e^{-zt}=e^{-(\alpha+j\beta)t}=e^{-\alpha t}e^{-j\beta t} \tag{3B.8}$$

Now $\left|e^{-j\beta t}\right|=1$ regardless of the value of βt. Therefore, as $t\to\infty$, $e^{-zt}\to 0$ only if $\alpha>0$, and $e^{-zt}\to\infty$ if $\alpha<0$. Thus:

$$\lim_{t\to\infty}e^{-zt}=\begin{cases}0 & \operatorname{Re} z>0\\ \infty & \operatorname{Re} z<0\end{cases} \tag{3B.9}$$

Clearly:

$$\lim_{t\to\infty} e^{-(s+a)t} = \begin{cases} 0 & \operatorname{Re}(s+a) > 0 \\ \infty & \operatorname{Re}(s+a) < 0 \end{cases} \tag{3B.10}$$

Use of this result in Eq. (3B.7) yields:

$$X(s) = \frac{1}{s+a} \quad \operatorname{Re}(s+a) > 0 \tag{3B.11}$$

or:

$$\boxed{e^{-at}u(t) \leftrightarrow \frac{1}{s+a} \quad \operatorname{Re} s > -a} \tag{3B.12}$$

The ROC of $X(s)$ is $\operatorname{Re} s > -a$, as shown in the shaded area below:

A signal and the region of convergence of its Laplace transform in the *s*-plane

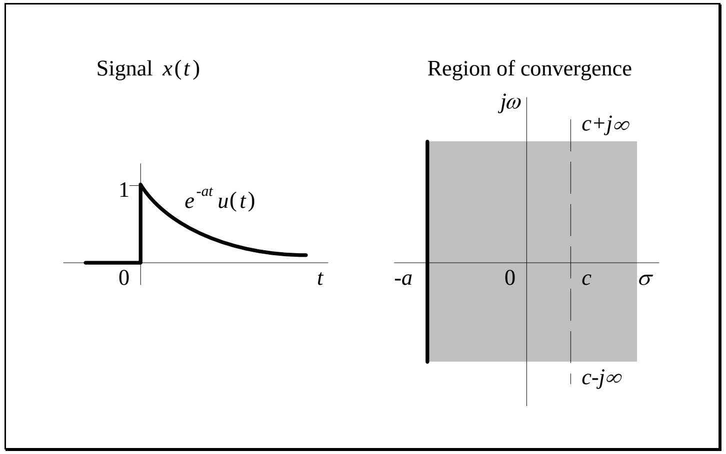


Figure 3B.2

This fact means that the integral defining $X(s)$ in Eq. (3B.7) exists only for the values of s in the shaded region in Figure 3B.2. For other values of s, the integral does not converge.

The ROC is required for evaluating the inverse Laplace transform, as defined by Eq. (3B.4). The operation of finding the inverse transform requires integration in the complex plane. The path of integration is along $c + j\omega$, with ω varying from $-\infty$ to ∞. This path of integration must lie in the ROC for $X(s)$. For the signal $e^{-at}u(t)$, this is possible if $c > -a$. One possible path of integration is shown (dotted) in Figure 3B.2. We can avoid this integration by compiling a table of Laplace transforms.

The ROC is needed to establish the convergence of the Laplace transform

If all the signals we deal with are restricted to $t \geq 0$, then the inverse transform is unique, and we do not need to specify the ROC.

The ROC is not needed if we deal with causal signals only

Finding a Fourier Transform using a Laplace Transform

If the function $x(t)$ is zero for all $t < 0$, and the Laplace transform exists at the point $s = 0 + j\omega$, then:

$$X(\omega) = X(s)\big|_{s=j\omega} \tag{3B.13}$$

The Fourier transform from the Laplace transform

That is, the Fourier transform $X(\omega)$ is just the Laplace transform $X(s)$ with $s = j\omega$.

Example

To find the Fourier transform of a signal $x(t) = e^{-3t}u(t)$, we substitute $s = j\omega$ into its Laplace transform:

$$X(\omega) = \frac{1}{s+3}\bigg|_{s=j\omega} = \frac{1}{j\omega+3} = \frac{1}{3+j2\pi f} \tag{3B.14}$$

A quick check from our knowledge of the Fourier transform shows this to be correct (because the Laplace transform includes the $j\omega$ axis in its region of convergence).

We can also view the Laplace transform geometrically, if we are willing to split the transform into its magnitude and phase (remember $X(s)$ is a complex number). The magnitude of $X(s) = 1/(s+3)$ can be visualised by graphing $|X(s)|$ as the height of a surface spread out over the *s*-plane:

The graph of the magnitude of the Laplace transform over the *s*-plane forms a surface with poles and zeros

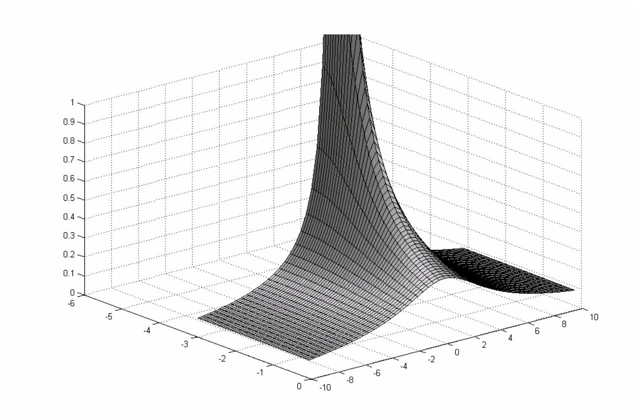


Figure 3B.3

There are several things we can notice about the plot above. First, note that the surface has been defined only in the ROC, i.e. for $\operatorname{Re} s > -3$. Secondly, the surface approaches an infinite value at the point $s = -3$. Such a point is termed a *pole*, in obvious reference to the surface being analogous to a tent (a *zero* is a point where the surface has a value of zero).

We can completely specify $X(s)$, apart from a constant gain factor, by drawing a so-called *pole-zero plot*:

A pole-zero plot is a shorthand way of representing a Laplace transform

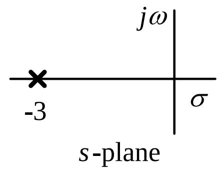


Figure 3B.4

A pole-zero plot locates all the critical points in the *s*-plane that completely specify the function $X(s)$ (to within an arbitrary constant), and it is a useful analytic and design tool.

Thirdly, one cut of the surface has been fortuitously placed along the imaginary axis. If we graph the height of the surface along this cut against ω, we get a picture of the magnitude of the Fourier transform vs. ω:

The Fourier transform is obtained from the $j\omega$-axis of a plot of the Laplace transform over the entire *s*-plane

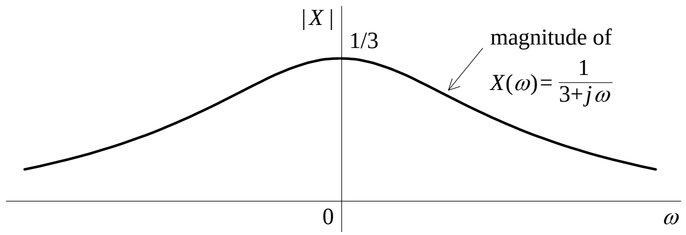


Figure 3B.5

With these ideas in mind, it should be apparent that a function that has a Laplace transform with a ROC in the right-half plane does not have a Fourier transform (because the Laplace transform surface will never intersect the $j\omega$-axis).

Finding Laplace Transforms

Like the Fourier transform, it is only necessary from a practical viewpoint to find Laplace transforms for a few standard signals, and then formulate several properties of the Laplace transform. Then, finding a Laplace transform of a function will consist of starting with a known transform and successively applying known transform properties.

Example

To find the Fourier transform of a signal $x(t)=\delta(t)$, we substitute into the Laplace transform definition:

$$X(s)=\int_{0^-}^{\infty}\delta(t)e^{-st}dt \tag{3B.15}$$

Recognising this as a sifting integral, we arrive at a standard transform pair:

The Laplace transform of an impulse

$$\boxed{\delta(t)\leftrightarrow 1} \tag{3B.16}$$

Thus, the Laplace transform of an impulse is 1, just like the Fourier transform.

Example

To find the Laplace transform of the unit-step, just substitute $a=0$ into Eq. (3B.12). The result is:

The Laplace transform of a unit-step

$$\boxed{u(t)\leftrightarrow \frac{1}{s}} \tag{3B.17}$$

This is a frequently used transform in the study of electrical circuits and control systems.

Example

To find the Laplace transform of $\cos(\omega_0 t)u(t)$, we recognise that:

$$\cos(\omega_0 t)u(t)=\frac{1}{2}\left[e^{j\omega_0 t}+e^{-j\omega_0 t}\right]u(t) \tag{3B.18}$$

From Eq. (3B.12), it follows that:

$$\begin{aligned}\frac{1}{2}\left[e^{j\omega_0 t}+e^{-j\omega_0 t}\right]u(t) &\leftrightarrow \frac{1}{2}\left[\frac{1}{s-j\omega_0}+\frac{1}{s+j\omega_0}\right]\\ &=\frac{s}{s^2+\omega_0^2}\end{aligned} \tag{3B.19}$$

and so we have another standard transform:

$$\boxed{\cos(\omega_0 t)u(t)\leftrightarrow \frac{s}{s^2+\omega_0^2}} \tag{3B.20}$$

A similar derivation can be used to find the Laplace transform of $\sin(\omega_0 t)u(t)$.

Most of the Laplace transform properties are inherited generalisations of Fourier transform properties. There are a few important exceptions, based on the fact that the Laplace transform is one-sided (from 0 to ∞), whereas the Fourier transform is double-sided (from $-\infty$ to ∞) and the Laplace transform is more general in the sense that it covers the entire *s*-plane, not just the $j\omega$-axis.

Differentiation Property

One of the most important properties of the Laplace transform is the differentiation property. It enables us to directly transform a *differential* equation into an *algebraic* equation in the complex variable *s*. It is much easier to solve algebraic equations than differential equations!

The Laplace transform of the derivative of a function is given by:

$$L\left[\frac{dx}{dt}\right]=\int_{0^-}^{\infty}\frac{dx}{dt}e^{-st}dt \tag{3B.21}$$

Integrating by parts, we obtain:

$$L\left[\frac{dx}{dt}\right]=x(t)e^{-st}\Big|_{0^-}^{\infty}+s\int_{0^-}^{\infty}x(t)e^{-st}dt \tag{3B.22}$$

For the Laplace integral to converge [i.e. for $X(s)$ to exist], it is necessary that $x(t)e^{-st}\to 0$ as $t\to\infty$ for the values of s in the ROC for $X(s)$. Thus:

The Laplace transform differentiation property

$$\boxed{\frac{d}{dt}x(t)\leftrightarrow sX(s)-x(0^-)} \tag{3B.23}$$

Standard Laplace Transforms

$$u(t)\leftrightarrow\frac{1}{s} \tag{L.1}$$

$$\delta(t)\leftrightarrow 1 \tag{L.2}$$

$$e^{-t}u(t)\leftrightarrow\frac{1}{s+1} \tag{L.3}$$

$$(\cos\omega t)u(t)\leftrightarrow\frac{s}{s^2+\omega^2} \tag{L.4}$$

$$(\sin\omega t)u(t)\leftrightarrow\frac{\omega}{s^2+\omega^2} \tag{L.5}$$

Laplace Transform Properties

Assuming $x(t)\leftrightarrow X(s)$.

$$ax(t)\leftrightarrow aX(s) \tag{L.6}$$ Linearity

$$x\left(\frac{t}{T}\right)\leftrightarrow |T|X(sT) \tag{L.7}$$ Scaling

$$x(t-c)u(t-c)\leftrightarrow e^{-cs}X(s) \tag{L.8}$$ Time shifting

$$e^{at}x(t)\leftrightarrow X(s-a) \tag{L.9}$$ Multiplication by exponential

$$t^N x(t)\leftrightarrow(-1)^N\frac{d^N}{ds^N}X(s) \tag{L.10}$$ Multiplication by *t*

$$\frac{d}{dt}x(t)\leftrightarrow sX(s)-x(0^-) \tag{L.11}$$ Differentiation

$$\int_0^t x(\tau)d\tau\leftrightarrow\frac{1}{s}X(s) \tag{L.12}$$ Integration

$$x_1(t)*x_2(t)\leftrightarrow X_1(s)X_2(s) \tag{L.13}$$ Convolution

$$x(0)=\lim_{s\to\infty}sX(s) \tag{L.14}$$ Initial-value theorem

$$\lim_{t\to\infty}x(t)=\lim_{s\to 0}sX(s) \tag{L.15}$$ Final-value theorem

Evaluation of the Inverse Laplace Transform

Finding the inverse Laplace transform requires integration in the complex plane, which is normally difficult and time consuming to compute. Instead, we can find inverse transforms from a table of Laplace transforms. All we need to do is express $X(s)$ as a sum of simpler functions of the forms listed in the table. Most of the transforms of practical interest are *rational functions*, that is, ratios of polynomials in *s*. Such functions can be expressed as a sum of simpler functions by using partial fraction expansion.

Rational Laplace Transforms

A rational Laplace transform of degree *N* can be expressed as:

$$X(s)=\frac{b_M s^M+b_{M-1}s^{M-1}+\ldots+b_1 s+b_0}{a_N s^N+a_{N-1}s^{N-1}+\ldots+a_1 s+a_0} \tag{3B.24}$$

This can also be written:

$$X(s)=\frac{b_M s^M+b_{M-1}s^{M-1}+\ldots+b_1 s+b_0}{a_N(s-p_1)(s-p_2)\ldots(s-p_N)} \tag{3B.25}$$

where the p_i are called the *poles* of $X(s)$. If the poles are all distinct, then the partial fraction expansion of Eq. (3B.25) is:

Rational Laplace transforms written in terms of *poles*

$$\boxed{X(s)=\frac{c_1}{s-p_1}+\frac{c_2}{s-p_2}+\ldots+\frac{c_N}{s-p_N}} \tag{3B.26}$$

We call the coefficients c_i "residues", which is a term derived from complex variable theory. They are given by:

Residues defined

$$\boxed{c_i=\left[(s-p_i)X(s)\right]_{s=p_i}} \tag{3B.27}$$

Taking the inverse Laplace transform of Eq. (3B.26) using standard transform (L.3) and property (L.7), gives us:

The time-domain *form* depends *only* on the *poles*

$$x(t)=c_1e^{p_1t}+c_2e^{p_2t}+\ldots+c_Ne^{p_Nt},\quad t\geq 0 \tag{3B.28}$$

Note that the form of the time-domain expression is determined *only by the poles* of $X(s)$!

Example

Find the inverse Laplace transform of:

$$Y(s)=\frac{2(s+1)}{(s+2)(s+4)} \tag{3B.29}$$

By expansion into partial fractions we have:

$$\frac{2(s+1)}{(s+2)(s+4)}=\frac{c_1}{s+2}+\frac{c_2}{s+4} \tag{3B.30}$$

To find c_1, multiply both sides of Eq. (3B.30) by $(s+2)$:

$$\frac{2(s+1)}{(s+4)}=c_1+\frac{c_2(s+2)}{s+4} \tag{3B.31}$$

As this equation must be true for all values of *s*, set $s=-2$ to remove c_2:

$$\frac{2(-2+1)}{(-2+4)}=c_1=-1 \tag{3B.32}$$

Heaviside's cover-up rule

An equivalent way to find c_1, without performing algebraic manipulation by hand, is to mentally cover up the factor $(s+2)$ on the left-hand side, and then evaluate the left-hand side at a value of *s* that makes the factor $(s+2)=0$, i.e. at $s=-2$. This mental procedure is known as *Heaviside's cover-up rule*.

Applying Heaviside's cover-up rule for c_2 results in the *mental* equation:

$$\frac{2(-4+1)}{(-4+2)} = c_2 = 3 \tag{3B.33}$$

Therefore, the partial fraction expansion is:

$$Y(s) = \frac{-1}{s+2} + \frac{3}{s+4} \tag{3B.34}$$

The inverse Laplace transform can now be easily evaluated using standard transform (L.3) and property (L.7):

$$y(t) = -e^{-2t} + 3e^{-4t}, \quad t \geq 0 \tag{3B.35}$$

Note that the continual writing of $u(t)$ after each function has been replaced by the more notationally convenient condition of $t \geq 0$ on the total solution.

If there is a pole $p_1 = \sigma + j\omega$, then the complex conjugate $p_1^* = \sigma - j\omega$ is also a pole (that's how we get real coefficients in the polynomial). In this case the residues of the two poles are complex conjugate and:

$$X(s) = \frac{c_1}{s-p_1} + \frac{c_1^*}{s_1 - p_1^*} + \frac{c_3}{s-p_3} + \ldots + \frac{c_N}{s-p_N} \tag{3B.36}$$

The inverse transform is:

$$x(t) = c_1 e^{p_1 t} + c_1^* e^{p_1^* t} + \ldots + c_N e^{p_N t}, \quad t \geq 0 \tag{3B.37}$$

which can be expressed as:

Complex-conjugate poles lead to a sinusoidal response

$$x(t) = 2|c_1| e^{\sigma t} \cos(\omega t + \angle c_1) + c_3 e^{p_3 t} + \ldots + c_N e^{p_N t}, \quad t \geq 0 \tag{3B.38}$$

Now suppose the pole p_1 of $X(s)$ is repeated r times. Then the partial fraction expansion of Eq. (3B.25) is:

Partial fraction expansion with repeated poles

$$\boxed{X(s) = \frac{c_1}{s-p_1} + \frac{c_2}{(s-p_1)^2} + \ldots + \frac{c_r}{(s-p_1)^r} + \frac{c_{r+1}}{(s-p_{r+1})} + \ldots + \frac{c_N}{(s-p_N)}} \tag{3B.39}$$

The residues are given by Eq. (3B.27) for the distinct poles $r+1 \leq i \leq N$ and:

Residue defined for repeated poles

$$\boxed{c_{r-i} = \frac{1}{i!}\left[\frac{d^i}{ds^i}(s-p_1)^r X(s)\right]_{s=p_1}} \tag{3B.40}$$

for the repeated poles $0 \leq i \leq r-1$.

Example

Find the inverse Laplace transform of:

$$Y(s) = \frac{s+4}{(s+1)(s+2)^2} \tag{3B.41}$$

Expanding into partial fractions we have:

$$\frac{s+4}{(s+1)(s+2)^2} = \frac{c_1}{s+1} + \frac{c_2}{s+2} + \frac{c_3}{(s+2)^2} \tag{3B.42}$$

Find c_1 by multiplying both sides by $(s+1)$ and setting $s=-1$ (Heaviside's cover-up rule):

$$\frac{-1+4}{(-1+2)^2} = c_1 = 3 \tag{3B.43}$$

To find c_3, multiply throughout by $(s+2)^2$ and set $s=-2$:

$$\frac{-2+4}{(-2+1)} = c_3 = -2 \tag{3B.44}$$

Note that Heaviside's cover-up rule only applies to the repeated partial fraction with the highest power. To find c_2, we have to use Eq. (3B.40). Multiplying throughout by $(s+2)^2$ gives:

$$\frac{s+4}{s+1}=\frac{c_1}{s+1}(s+2)^2+c_2(s+2)+c_3 \tag{3B.45}$$

Now to get rid of c_3, differentiate with respect to *s*:

$$\frac{d}{ds}\left[\frac{s+4}{s+1}\right]=\frac{d}{ds}\left[\frac{c_1(s+2)^2}{s+1}\right]+\frac{d}{ds}\left[c_2(s+2)\right] \tag{3B.46}$$

The differentiation of the quotients can be obtained using:

$$\frac{d}{dx}\left[\frac{u}{v}\right]=\frac{v\frac{du}{dx}-u\frac{dv}{dx}}{v^2} \tag{3B.47}$$

Therefore, the LHS of our problem becomes:

$$\frac{d}{ds}\left[\frac{s+4}{s+1}\right]=\frac{(s+1)-(s+4)}{(s+1)^2}=\frac{-3}{(s+1)^2} \tag{3B.48}$$

The second term on the RHS becomes:

$$\frac{d}{ds}\left[c_2(s+2)\right]=c_2 \tag{3B.49}$$

Differentiation of the c_1 term will result in an $(s+2)$ multiplier. Therefore, if *s* is set to -2 in the equation after differentiation, we can resolve c_2:

$$c_2=\frac{-3}{(-2+1)^2}=-3 \tag{3B.50}$$

Therefore, the partial fraction expansion is:

$$Y(s)=\frac{3}{s+1}-\frac{3}{s+2}-\frac{2}{(s+2)^2} \tag{3B.51}$$

The inverse Laplace transform can now be easily evaluated using (L.3), (L.7) and (L.10):

$$y(t)=3e^{-t}-3e^{-2t}-2te^{-2t},\quad t\geq 0 \tag{3B.52}$$

Multiple poles produce coefficients that are polynomials of *t*

A multiple pole will, in general, produce a coefficient of the exponential term which is a polynomial in *t*.

MATLAB® can be used to obtain the poles and residues for a given rational function $X(s)$.

Example

Given:

$$X(s)=\frac{s^2-2s+1}{s^3+3s^2+4s+2}$$

calculate $x(t)$.

Using MATLAB® we just do:

```
num = [1 -2 1];
den = [1 3 4 2];
[r,p] = residue(num,den);
```

which returns vectors of the residues and poles in `r` and `p` respectively.

In summary, there are three types of response from an LTI system:

- Real poles: the response is a sum of exponentials.
- Complex poles: the response is a sum of exponentially damped sinusoids.
- Multiple poles: the response is a sum of exponentials but the coefficients of the exponentials are polynomials in *t*.

Thus, we see that any high order rational function $X(s)$ can be decomposed into a combination of first-order factors.

Transforms of Differential Equations

The time-differentiation property of the Laplace transform sets the stage for solving linear differential equations with constant coefficients. Because $d^k y/dt^k \leftrightarrow s^k Y(s)$, the Laplace transform of a differential equation is an algebraic equation that can be readily solved for $Y(s)$. Next we take the inverse Laplace transform of $Y(s)$ to find the desired solution $y(t)$.

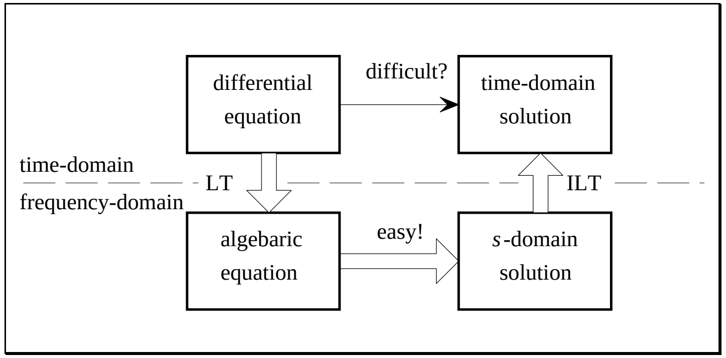


Figure 3B.6

Example

Solve the second-order linear differential equation:

$$\left(D^2+5D+6\right)y(t)=(D+1)x(t) \tag{3B.53}$$

for the initial conditions $y(0^-)=2$ and $\dot{y}(0^-)=1$ and the input $x(t)=e^{-4t}u(t)$.

The equation is:

$$\frac{d^2y}{dt^2}+5\frac{dy}{dt}+6y=\frac{dx}{dt}+x \tag{3B.54}$$

Let $y(t)\leftrightarrow Y(s)$. Then from property (L.11):

$$\begin{aligned}\frac{dy}{dt}&\leftrightarrow sY(s)-y(0^-)=sY(s)-2\\ \frac{d^2y}{dt^2}&\leftrightarrow s^2Y(s)-sy(0^-)-\dot{y}(0^-)=s^2Y(s)-2s-1\end{aligned} \tag{3B.55}$$

Also, for $x(t)=e^{-4t}u(t)$, we have:

$$X(s)=\frac{1}{s+4} \quad \text{and} \quad \frac{dx}{dt}\leftrightarrow sX(s)-x(0^-)=\frac{s}{s+4}-0=\frac{s}{s+4} \tag{3B.56}$$

Taking the Laplace transform of Eq. (3B.54), we obtain:

$$\left[s^2Y(s)-2s-1\right]+5\left[sY(s)-2\right]+6Y(s)=\frac{s}{s+4}+\frac{1}{s+4} \tag{3B.57}$$

Collecting all the terms of $Y(s)$ and the remaining terms separately on the left-hand side, we get:

$$\left(s^2+5s+6\right)Y(s)-(2s+11)=\frac{s+1}{s+4} \tag{3B.58}$$

so that:

$$\left(s^2+5s+6\right)Y(s)=\underbrace{(2s+11)}_{\text{initial condition terms}}+\underbrace{\frac{s+1}{s+4}}_{\text{input terms}} \tag{3B.59}$$

Therefore:

$$\begin{aligned}Y(s)&=\underbrace{\frac{2s+11}{s^2+5s+6}}_{\text{zero-input component}}+\underbrace{\frac{s+1}{(s+4)\left(s^2+5s+6\right)}}_{\text{zero-state component}}\\ &=\left[\frac{7}{s+2}-\frac{5}{s+3}\right]+\left[\frac{-1/2}{s+2}+\frac{2}{s+3}-\frac{3/2}{s+4}\right]\end{aligned} \tag{3B.60}$$

Taking the inverse transform yields:

$$y(t) = \underbrace{7e^{-2t} - 5e^{-3t}}_{\text{zero-input response}} + \underbrace{-\tfrac{1}{2}e^{-2t} + 2e^{-3t} - \tfrac{3}{2}e^{-4t}}_{\text{zero-state response}}$$
$$= \tfrac{13}{2}e^{-2t} - 3e^{-3t} - \tfrac{3}{2}e^{-4t}, \quad t \geq 0 \tag{3B.61}$$

The Laplace transform method gives the total response, which includes zero-input and zero-state components. The initial condition terms give rise to the zero-input response. The zero-state response terms are exclusively due to the input.

Consider the N^{th} order input/output differential equation:

Systems described by differential equations

$$\boxed{\frac{d^N y(t)}{dt^N} + \sum_{i=0}^{N-1} a_i \frac{d^i y(t)}{dt^i} = \sum_{i=0}^{M} b_i \frac{d^i x(t)}{dt^i}} \tag{3B.62}$$

If we take the Laplace transform of this equation using (L.11), assuming zero initial conditions, we get:

transform into rational functions of s

$$Y(s) = \frac{b_M s^M + b_{M-1}s^{M-1} + \ldots + b_1 s + b_0}{s^N + a_{N-1}s^{N-1} + \ldots + a_1 s + a_0} X(s) \tag{3B.63}$$

Now define:

Transfer function defined

$$\boxed{Y(s) = H(s)X(s)} \tag{3B.64}$$

The function $H(s)$ is called the *transfer function* of the system since it specifies the transfer from the input to the output in the s-domain (assuming no initial energy). This is true for any system. For the case described by Eq. (3B.62), the transfer function is a rational polynomial given by:

$$H(s) = \frac{b_M s^M + b_{M-1}s^{M-1} + \ldots + b_1 s + b_0}{s^N + a_{N-1}s^{N-1} + \ldots + a_1 s + a_0} \tag{3B.65}$$

Do the revision problem assuming zero initial conditions.

System Transfer Function

For a linear time-invariant system described by a convolution integral, we can take the Laplace transform and get:

$$Y(s) = H(s)X(s) \tag{3B.66}$$

which shows us that:

$$\boxed{h(t) \leftrightarrow H(s)} \tag{3B.67a}$$

The relationship between time-domain and frequency-domain descriptions of a system

The transfer function is the Laplace transform of the impulse response!	(3B.67b)

Instead of writing $H(s)$ as in Eq. (3B.65), it can be expressed in the factored form:

Transfer function factored to get zeros and poles

$$\boxed{H(s) = \frac{b_M (s - z_1)(s - z_2)\cdots(s - z_M)}{(s - p_1)(s - p_2)\cdots(s - p_N)}} \tag{3B.68}$$

where the z's are the zeros of the system and the p's are the poles of the system. This shows us that apart from a constant factor b_M, the poles and zeros determine the transfer function completely. They are often displayed on a *pole-zero diagram*, which is a plot in the s-domain showing the location of all the poles (marked by x) and all the zeros (marked by o).

You should be familiar with direct construction of the transfer function for electric circuits from previous subjects.

Block Diagrams

Block diagrams are transfer functions

Systems are often represented as interconnections of *s*-domain "blocks", with each block containing a transfer function.

Example

Given the following electrical system:

A simple first-order system

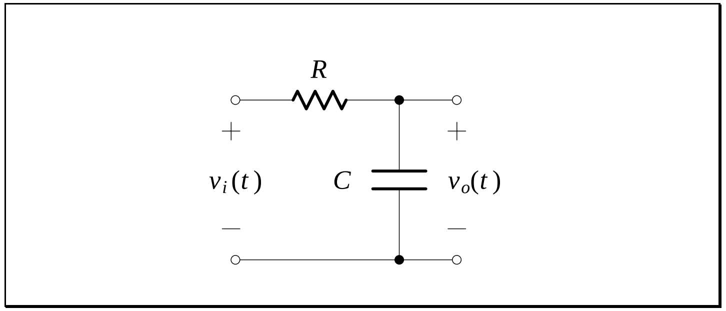


Figure 3B.7

we can perform KVL around the loop to get the differential equation:

$$v_o(t) + RC\frac{dv_o(t)}{dt} = v_i(t) \tag{3B.69}$$

Assuming zero initial conditions, the Laplace transform of this is:

$$V_o(s) + sRCV_o(s) = V_i(s) \tag{3B.70}$$

and therefore the system transfer function is:

$$\frac{V_o(s)}{V_i(s)} = \frac{1}{1+sRC} = \frac{1}{1+sT} \tag{3B.71}$$

where $T = RC$, the time constant. Therefore the block diagram is:

Block diagram of the simple lowpass *RC* circuit

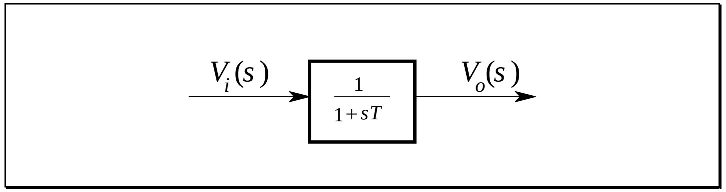


Figure 3B.8

Note that there's no hint of what makes up the inside of the block, except for the input and output signals. It could be a simple *RC* circuit, a complex passive circuit, or even an active circuit (with op-amps). The important thing the block does is hide all this detail.

Notation

We use the following notation, where $G(s)$ is the transfer function:

A block represents multiplication with a transfer function

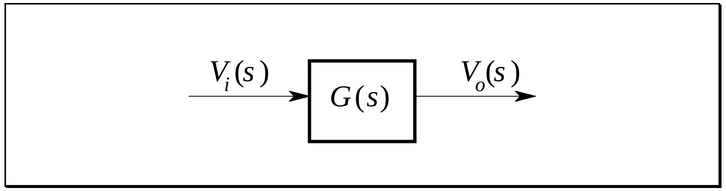


Figure 3B.9

Most systems have several blocks interconnected by various forwards and backwards paths. Signals in block diagrams can not only be transformed by a transfer function, they can also be added and subtracted.

Addition and subtraction of signals in a block diagram

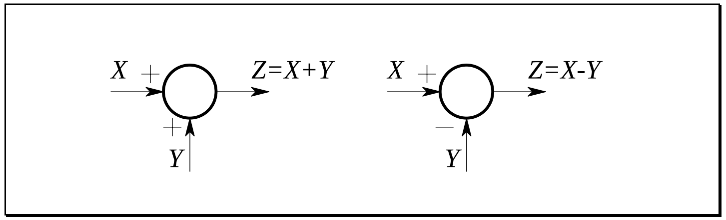


Figure 3B.10

Cascading Blocks

Cascading blocks implies multiplying the transfer functions

Blocks can be connected in cascade.

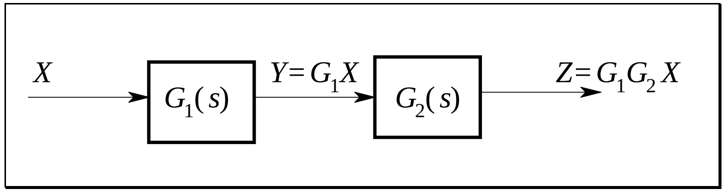


Figure 3B.11

Care must be taken when cascading blocks. Consider what happens when we try to create a second-order circuit by cascading two first-order circuits:

A circuit which IS NOT the cascade of two first-order circuits

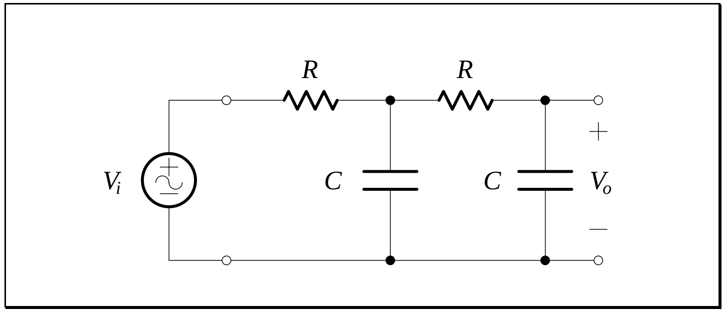


Figure 3B.12

Show that the transfer function for the above circuit is:

$$\frac{V_o}{V_i} = \frac{\left(1/RC\right)^2}{s^2 + \left(3/RC\right)s + \left(1/RC\right)^2} \tag{3B.72}$$

Compare with the following circuit:

A circuit which IS the cascade of two first-order circuits

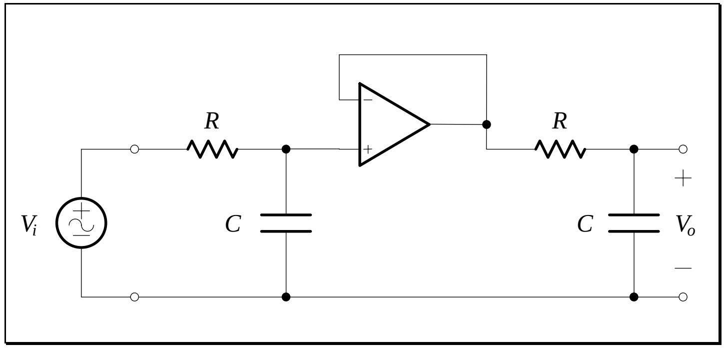


Figure 3B.13

which has the transfer function:

$$\begin{aligned}\frac{V_o}{V_i} &= \frac{1/RC}{s + 1/RC}\,\frac{1/RC}{s + 1/RC} \\ &= \frac{\left(1/RC\right)^2}{s^2 + \left(2/RC\right)s + \left(1/RC\right)^2}\end{aligned} \tag{3B.73}$$

We can only cascade circuits if they are buffered

They are different! In the first case, the second network *loads* the first (i.e. they interact). We can only cascade circuits if the "outputs" of the circuits present a low impedance to the next stage, so that each successive circuit does not "load" the previous circuit. Op-amp circuits of both the inverting and non-inverting type are ideal for cascading.

Standard Form of a Feedback Control System

Perhaps the most important block diagram is that of a *feedback connection*, shown below:

Standard form for the feedback connection

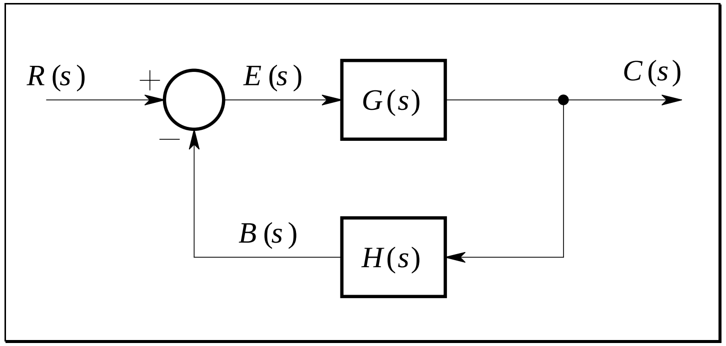


Figure 3B.14

We have the following definitions:

$G(s)$ = forward path transfer function

$H(s)$ = feedback path transfer function

$R(s)$ = **r**eference, input, or desired output

$C(s)$ = **c**ontrolled variable, or output

$B(s)$ = output multiplied by $H(s)$

$E(s)$ = actuating **e**rror signal

$R(s)-C(s)$ = system error

$\frac{C(s)}{R(s)}$ = closed-loop transfer function

$G(s)H(s)$ = loop gain

To find the transfer function, we solve the following two equations which are self-evident from the block diagram:

$$\begin{aligned} C(s) &= G(s)E(s) \\ E(s) &= R(s)-H(s)C(s) \end{aligned} \tag{3B.74}$$

Then the output $C(s)$ is given by:

$$\begin{aligned} C(s) &= G(s)R(s)-G(s)H(s)C(s) \\ C(s)[1+G(s)H(s)] &= G(s)R(s) \end{aligned} \tag{3B.75}$$

and therefore:

Transfer function for the standard feedback connection

$$\boxed{\frac{C(s)}{R(s)} = \frac{G(s)}{1+G(s)H(s)}} \tag{3B.76}$$

Similarly, we can show that:

Finding the error signal's transfer function for the standard feedback connection

$$\boxed{\frac{E(s)}{R(s)} = \frac{1}{1+G(s)H(s)}} \tag{3B.77}$$

and:

$$\frac{B(s)}{R(s)} = \frac{G(s)H(s)}{1+G(s)H(s)} \tag{3B.78}$$

Notice how all the above expressions have the *same* denominator.

We define $1+G(s)H(s)=0$ as the *characteristic equation* of the *differential equation* describing the system. Note that for negative feedback we get $1+G(s)H(s)$ and for positive feedback we get $1-G(s)H(s)$.

Characteristic equations for positive and negative feedback

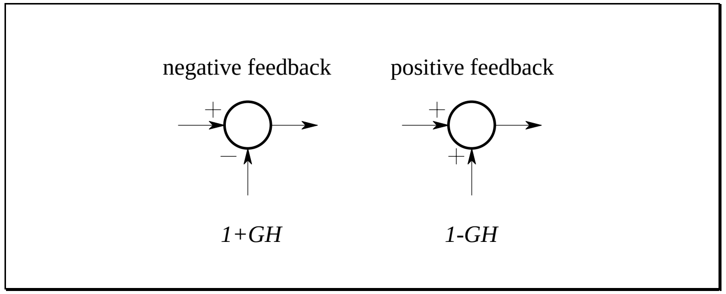
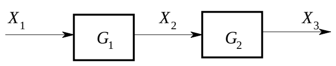
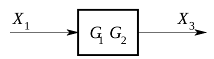
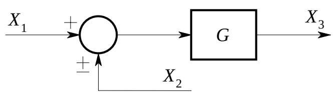
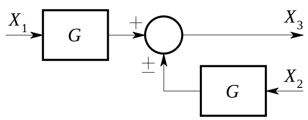
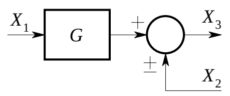
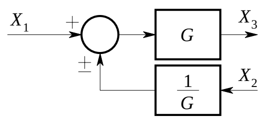
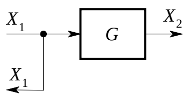
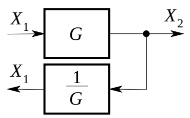
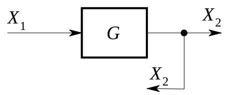
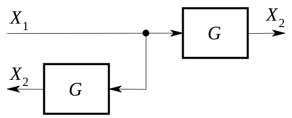
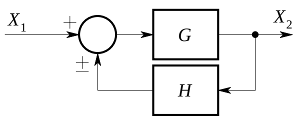
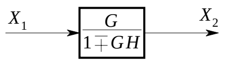


Figure 3B.15

Block Diagram Transformations

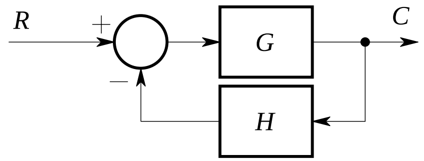
Simplifying block diagrams

We can manipulate the signals and blocks in a block diagram in order to simplify it. The overall system transfer function is obtained by combining the blocks in the diagram into just one block. This is termed block-diagram reduction.

	Original Diagram	Equivalent Diagram
Combining blocks in cascade		
Moving a summing point behind a block		
Moving a summing point ahead of a block		
Moving a pick-off point behind a block		
Moving a pick-off point ahead of a block		
Eliminating a feedback loop		

Example

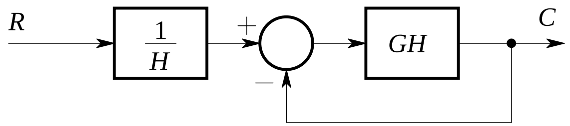
Given:



then alternatively we can get:

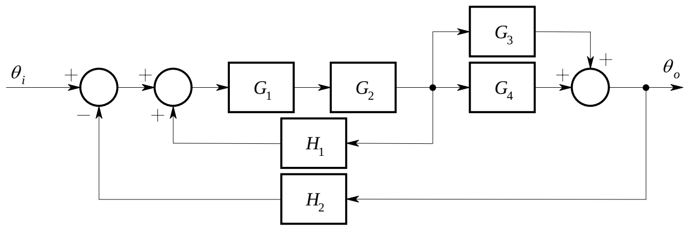
$$\begin{aligned} C &= G(R - HC) \\ &= GH\left(\frac{R}{H} - C\right) \end{aligned}$$

which is drawn as:

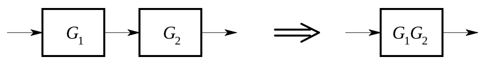


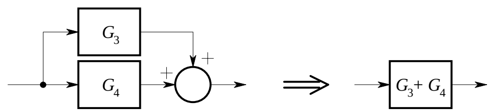
Example

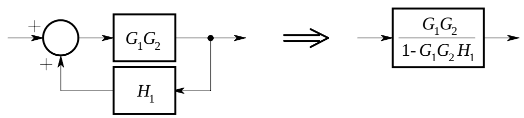
Given:



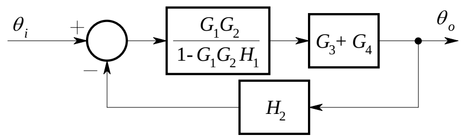
we put:



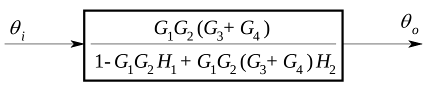




Therefore we get:

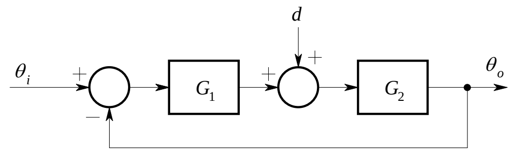


which simplifies to:

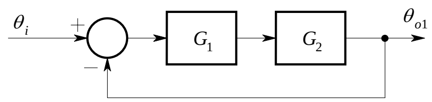


Example

Consider a system where *d* is a disturbance input which we want to suppress:



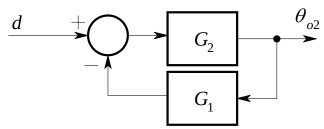
Using superposition, consider θ_i only:



Therefore:

$$\frac{\theta_{o1}}{\theta_i}=\frac{G_1G_2}{1+G_1G_2}$$

Considering *d* only:



we have:

$$\frac{\theta_{o2}}{\theta_i}=\frac{G_2}{1+G_1G_2}$$

Therefore, the total output is:

$$\theta_o=\theta_{o1}+\theta_{o2}=\frac{G_1G_2}{1+G_1G_2}\theta_i+\frac{G_2}{1+G_1G_2}d$$

Therefore, use a small G_2 and a large G_1!

Summary

- The Laplace transform is a generalization of the Fourier transform. To find the Laplace transform of a function, we start from a known Laplace transform pair, and apply any number of Laplace transform pair properties to arrive at the solution.

- Inverse Laplace transforms are normally obtained by the method of partial fractions using residues.

- Systems described by differential equations have rational Laplace transforms. The Laplace transforms of the input signal and output signal are related by the *transfer function* of the system: $Y(s) = H(s)X(s)$. There is a one-to-one correspondence between the coefficients in the differential equation and the coefficients in the transfer function.

- The impulse response and the transfer function form a Laplace transform pair: $h(t) \leftrightarrow H(s)$.

- The transfer function of a system can be obtained by performing analysis in the *s*-domain.

- A block diagram is composed of blocks containing transfer functions and adders. They are used to diagrammatically represent systems. All single-input single-output systems can be reduced to one equivalent transfer function.

References

Haykin, S.: *Communication Systems*, John-Wiley & Sons, Inc., New York, 1994.

Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB®*, Prentice-Hall, 1997.

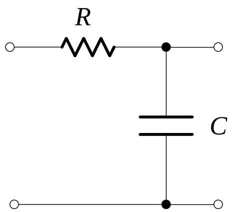
Lathi, B. P.: *Modern Digital and Analog Communication Systems*, Holt-Saunders, Tokyo, 1983.

Exercises

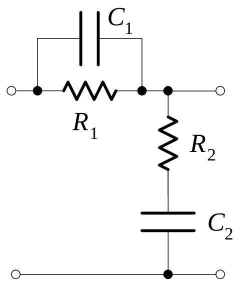
1.

Obtain transfer functions for the following networks:

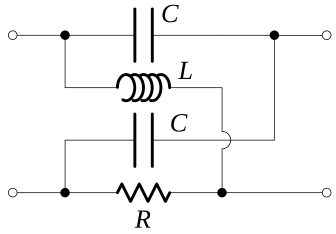
a)



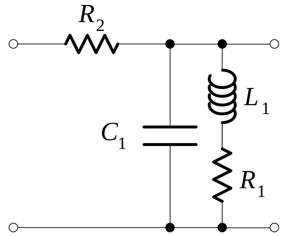
b)



c)



d)



2.

Obtain the Laplace transforms for the following integro-differential equations:

a) $L\dfrac{di(t)}{dt} + Ri(t) + \dfrac{1}{C}\int_0^t i(\tau)d\tau = e(t)$

b) $M\dfrac{d^2x(t)}{dt^2} + B\dfrac{dx(t)}{dt} + Kx(t) = 3t$

c) $J\dfrac{d^2\theta(t)}{dt^2} + B\dfrac{d\theta(t)}{dt} + K\theta(t) = 10\sin\omega t$

3.

Find the $f(t)$ which corresponds to each $F(s)$ below:

a) $$F(s)=\frac{s^2+5}{s^3+2s^2+4s}$$

b) $$F(s)=\frac{s}{s^4+5s^2+4}$$

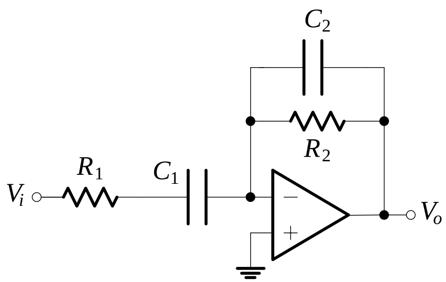
c) $$F(s)=\frac{3s+1}{5s^3(s-2)^2}$$

4.

Use the final value theorem to determine the final value for each $f(t)$ in 3 a), b) and c) above.

5.

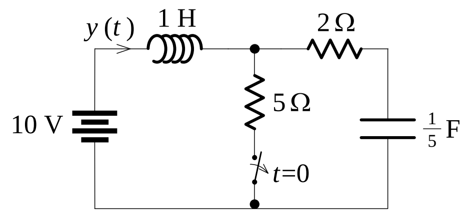
Find an expression for the transfer function of the following network (assume the op-amp is ideal):



6.

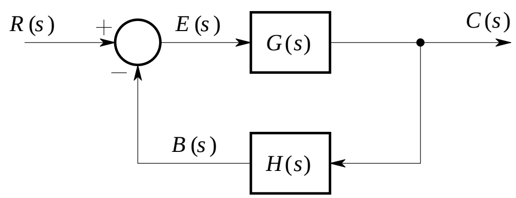
In the circuit below, the switch is in a closed position for a long time before $t=0$, when it is opened instantaneously.

Find the inductor current $y(t)$ for $t \geq 0$.



7.

Given:



a) Find expressions for:

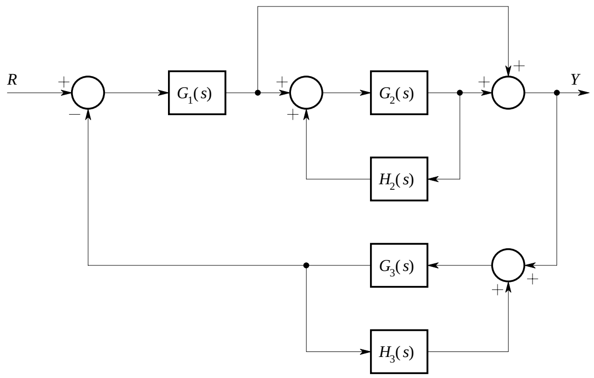
(i) $\dfrac{C(s)}{R(s)}$ (ii) $\dfrac{E(s)}{R(s)}$ (iii) $\dfrac{B(s)}{R(s)}$

b) What do you notice about the denominator in each of your solutions?

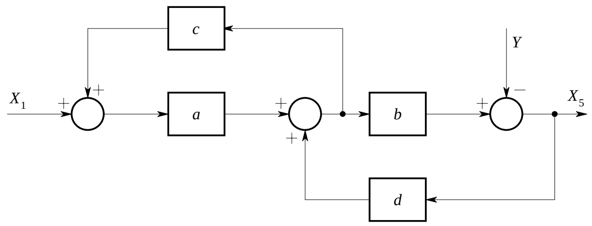
8.

Using block diagram reduction techniques, find the transfer functions of the following systems:

a)



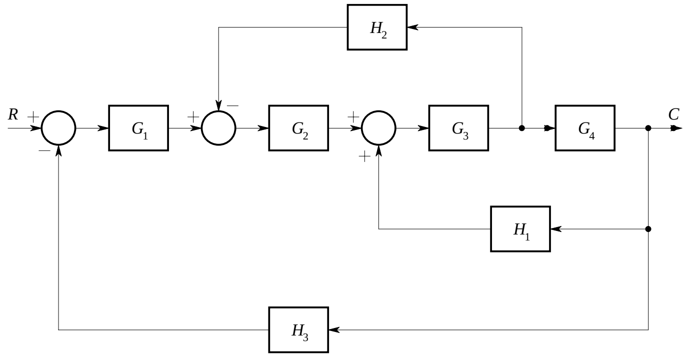
b)



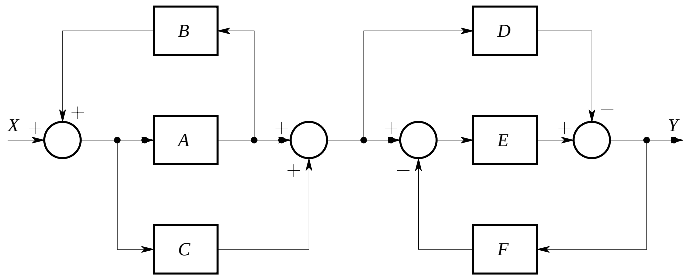
9.

Use block diagram reduction techniques to find the transfer functions of the following systems:

a)



b)



Pierre Simon de Laplace (1749-1827)

The application of mathematics to problems in physics became a primary task in the century after Newton. Foremost among a host of brilliant mathematical thinkers was the Frenchman Laplace. He was a powerful and influential figure, contributing to the areas of celestial mechanics, cosmology and probability.

Laplace was the son of a farmer of moderate means, and while at the local military school, his uncle (a priest) recognised his exceptional mathematical talent. At sixteen, he began to study at the University of Caen. Two years later he travelled to Paris, where he gained the attention of the great mathematician and philosopher Jean Le Rond d'Alembert by sending him a paper on the principles of mechanics. His genius was immediately recognised, and Laplace became a professor of mathematics.

He began producing a steady stream of remarkable mathematical papers. Not only did he make major contributions to difference equations and differential equations but he examined applications to mathematical astronomy and to the theory of probability, two major topics which he would work on throughout his life. His work on mathematical astronomy before his election to the Académie des Sciences included work on the inclination of planetary orbits, a study of how planets were perturbed by their moons, and in a paper read to the Academy on 27 November 1771 he made a study of the motions of the planets which would be the first step towards his later masterpiece on the stability of the solar system.

In 1773, before the Academy of Sciences, Laplace proposed a model of the solar system which showed how perturbations in a planet's orbit would not change it's distance from the sun. For the next decade, Laplace contributed a stream of papers on planetary motion, clearing up discrepancies in the orbit's of Jupiter and Saturn, he showed how the moon accelerates as a function of the Earth's orbit, he introduced a new calculus for discovering the motion of celestial bodies, and even a new means of computing planetary orbits which led to astronomical tables of improved accuracy.

The 1780s were the period in which Laplace produced the depth of results which have made him one of the most important and influential scientists that the world has seen. Laplace let it be known widely that he considered himself the best mathematician in France. The effect on his colleagues would have been only mildly eased by the fact that Laplace was right!

In 1784 Laplace was appointed as examiner at the Royal Artillery Corps, and in this role in 1785, he examined and passed the 16 year old Napoleon Bonaparte.

In 1785, he introduced a field equation in spherical harmonics, now known as Laplace's equation, which is found to be applicable to a great deal of phenomena, including gravitation, the propagation of sound, light, heat, water, electricity and magnetism.

Laplace presented his famous nebular hypothesis in 1796 in *Exposition du systeme du monde*, which viewed the solar system as originating from the contracting and cooling of a large, flattened, and slowly rotating cloud of incandescent gas. The *Exposition* consisted of five books: the first was on the apparent motions of the celestial bodies, the motion of the sea, and also atmospheric refraction; the second was on the actual motion of the celestial bodies; the third was on force and momentum; the fourth was on the theory of universal gravitation and included an account of the motion of the sea and the shape of the Earth; the final book gave an historical account of astronomy and included his famous nebular hypothesis which even predicted black holes. Laplace stated his philosophy of science in the *Exposition*:-

"Your Highness, I have no need of this hypothesis. "
- Laplace, to Napoleon on why his works on celestial mechanics make no mention of God.

> *If man were restricted to collecting facts the sciences were only a sterile nomenclature and he would never have known the great laws of nature. It is in comparing the phenomena with each other, in seeking to grasp their relationships, that he is led to discover these laws...*

Exposition du systeme du monde was written as a non-mathematical introduction to Laplace's most important work. Laplace had already discovered the invariability of planetary mean motions. In 1786 he had proved that the eccentricities and inclinations of planetary orbits to each other always remain small, constant, and self-correcting. These and many of his earlier results formed the basis for his great work the *Traité du Mécanique Céleste* published in 5 volumes, the first two in 1799.

The first volume of the *Mécanique Céleste* is divided into two books, the first on general laws of equilibrium and motion of solids and also fluids, while the second book is on the law of universal gravitation and the motions of the centres of gravity of the bodies in the solar system. The main mathematical approach was the setting up of differential equations and solving them to describe the resulting motions. The second volume deals with mechanics applied to a study of the planets. In it Laplace included a study of the shape of the Earth which included a discussion of data obtained from several different expeditions, and Laplace applied his theory of errors to the results.

In 1812 he published the influential study of probability, *Théorie analytique des probabilités*. The work consists of two books. The first book studies generating functions and also approximations to various expressions occurring in probability theory. The second book contains Laplace's definition of probability, Bayes's rule (named by Poincaré many years later), and remarks on mathematical expectation. The book continues with methods of finding probabilities of compound events when the probabilities of their simple components are known, then a discussion of the method of least squares, and inverse probability. Applications to mortality, life expectancy, length of marriages and probability in legal matters are given.

After the publication of the fourth volume of the *Mécanique Céleste*, Laplace continued to apply his ideas of physics to other problems such as capillary action (1806-07), double refraction (1809), the velocity of sound (1816), the theory of heat, in particular the shape and rotation of the cooling Earth (1817-1820), and elastic fluids (1821).

Lecture 4A – Transfer Functions

Stability. Step response. Sinusoidal response. Arbitrary response. Frequency response.

Overview

Transfer functions are obtained by Laplace transforming a system's input/output differential equation, or by analysing a system directly in the *s*-domain. From the transfer function, we can derive many important properties of a system.

The transfer function tells us lots of things

Stability

To look at stability, let's examine the rational system transfer function that ordinarily arises from linear differential equations:

$$H(s) = \frac{b_M s^M + b_{M-1} s^{M-1} + \ldots + b_1 s + b_0}{s^N + a_{N-1} s^{N-1} + \ldots + a_1 s + a_0} \tag{4A.1}$$

The transfer function $H(s)$ is the Laplace transform of the impulse response $h(t)$. Since it's the poles of a Laplace transform that determine the system's time-domain response, then the poles of $H(s)$ determine the form of $h(t)$. In particular, for real and complex non-repeated poles:

$$\frac{c}{s+p} \leftrightarrow ce^{-pt} \tag{4A.2a}$$

$$\frac{A\cos\phi(s+\sigma) - A\sin\phi\omega}{(s+\sigma)^2 + \omega^2} \leftrightarrow Ae^{-\sigma t}\cos(\omega t + \phi) \tag{4A.2b}$$

The impulse response is always some sort of exponential

If $H(s)$ has repeated poles, then $h(t)$ will contain terms of the form $ct^i e^{-pt}$ and/or $At^i e^{-\sigma t}\cos(\omega t + \phi)$.

From the time-domain expressions, it follows that $h(t)$ converges to zero as $t \to \infty$ if and only if:

Conditions on the poles for a stable system

$$\boxed{\operatorname{Re}(p_i) < 0} \tag{4A.3a}$$

where p_i are the poles of $H(s)$.

This is equivalent to saying:

Stability defined

A system is stable if all the poles of the transfer function lie in the open left-half *s*-plane (4A.3b)

Unit-Step Response

Consider a system with rational transfer function $H(s) = B(s)/A(s)$. If an input $x(t)$ is applied for $t \geq 0$ with no initial energy in the system, then the transform of the resulting output response is:

$$Y(s) = \frac{B(s)}{A(s)} X(s) \tag{4A.4}$$

Suppose $x(t)$ is the unit-step function $u(t)$, so that $X(s) = 1/s$. Then the transform of the step response is:

Transform of step-response for any system

$$Y(s) = \frac{B(s)}{A(s)s} \tag{4A.5}$$

Using partial fractions, this can be written as:

$$Y(s) = \frac{H(0)}{s} + \frac{E(s)}{A(s)} \tag{4A.6}$$

where $E(s)$ is a polynomial in s and the residue of the pole at the origin was given by:

$$c = \left[sY(s)\right]_{s=0} = H(0) \tag{4A.7}$$

Taking the inverse Laplace transform of $Y(s)$, we get the time-domain response to the unit-step function:

The complete response consists of a transient part and a steady-state part

$$\boxed{y(t) = H(0) + y_{\mathrm{tr}}(t), \quad t \geq 0} \tag{4A.8}$$

where $y_{\mathrm{tr}}(t)$ is the inverse Laplace transform of $E(s)/A(s)$. If the system is stable so that all the roots of $A(s) = 0$ lie in the open left-half plane, then the term $y_{\mathrm{tr}}(t)$ converges to zero as $t \to \infty$, in which case $y_{\mathrm{tr}}(t)$ is the *transient part of the response*.

So, if the system is stable, the step response contains a transient that decays to zero and it contains a constant with value $H(0)$. The constant $H(0)$ is the *steady-state* value of the step response.

Transients are important, especially for control system design

An analysis of the transient response is very important because we may wish to design a system with certain time-domain behaviour. For example, we may have a requirement to reach 99% of the steady-state value within a certain time, or we may wish to limit any oscillations about the steady-state value to a certain amplitude etc. This will be examined for the case of first-order and second-order systems.

First-Order Systems

For the first-order transfer function:

$$H(s) = \frac{p}{s+p} \tag{4A.9}$$

The unit-step response is:

Step-response of a first-order system

$$y(t) = 1 - e^{-pt}, \quad t \geq 0 \tag{4A.10}$$

which has been written in the form of Eq. (4A.8). If the system is stable, then *p* lies in the open left-half plane, and the second term decays to zero. The rate at which the transient decays to zero depends on how far over to the left the pole is. Since the total response is equal the constant "1" plus the transient response, the rate at which the step response converges to the steady-state value is equal to the rate at which the transient decays to zero. This may be an important design consideration.

Time constant defined

An important quantity that characterizes the rate of decay of the transient is the *time constant T*. It is defined as $1/p$, assuming $p > 0$. You are probably familiar with the concept of a time constant for electric circuits (eg. $T = RC$ for a simple *RC* circuit), but it is a concept applicable to all first-order systems. The smaller the time constant, the faster the rate of decay of the transient.

Second-Order Systems

Now consider the second-order system given by the transfer function:

Standard form of a second-order lowpass transfer function

$$H(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{4A.11}$$

The real parameters in the denominator are:

Damping ratio and natural frequency defined

$$\zeta = \text{damping ratio} \tag{4A.12a}$$

$$\omega_n = \text{natural frequency} \tag{4A.12b}$$

If we write:

$$H(s) = \frac{\omega_n^2}{(s+p_1)(s+p_2)} \tag{4A.13}$$

then the poles of $H(s)$ are given by the quadratic formula:

Pole locations for a second-order system

$$p_1 = -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1} \tag{4A.14a}$$

$$p_2 = -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1} \tag{4A.14b}$$

There are three cases to consider.

Distinct Real Poles ($\zeta > 1$) - Overdamped

In this case, the transfer function can be expressed as:

$$H(s) = \frac{\omega_n^2}{(s+p_1)(s+p_2)} \tag{4A.15}$$

and the transform of the unit-step response is given by:

$$Y(s) = \frac{\omega_n^2}{(s+p_1)(s+p_2)s} \tag{4A.16}$$

Rewriting as partial fractions and taking the inverse Laplace transform, we get the unit-step response:

Step response of a second-order overdamped system

$$y(t) = 1 + c_1 e^{-p_1 t} + c_2 e^{-p_2 t}, \quad t \geq 0 \tag{4A.17}$$

Therefore, the transient part of the response is given by the sum of two exponentials:

Transient part of the step response of a second-order overdamped system

$$y_{\mathrm{tr}}(t) = c_1 e^{-p_1 t} + c_2 e^{-p_2 t} \tag{4A.18}$$

and the steady-state value:

Steady-state part of the step response of a second-order overdamped system

$$y_{\mathrm{ss}}(t) = H(0) = \frac{\omega_n^2}{p_1 p_2} = 1 \tag{4A.19}$$

It often turns out that the transient response is *dominated* by one pole (the one closer to the origin – why?), so that the step response looks like that of a first-order system.

Repeated Real Poles ($\zeta = 1$) – Critically Damped

In this case, the transfer function has the factored form:

$$H(s) = \frac{\omega_n^2}{(s+\omega_n)^2} \tag{4A.20}$$

Expanding $H(s)/s$ via partial fractions and taking the inverse transform yields the step response:

Unit-step response of a second-order critically damped system

$$y(t) = 1 - (1 + \omega_n t)e^{-\omega_n t} \tag{4A.21}$$

Hence, the transient response is:

Transient part of the unit-step response of a second-order critically damped system

$$y_{\mathrm{tr}}(t) = -(1 + \omega_n t)e^{-\omega_n t} \tag{4A.22}$$

and the steady-state response is:

Steady-state part of the unit-step response of a second-order critically damped system

$$y_{\mathrm{ss}}(t) = H(0) = 1 \tag{4A.23}$$

as before.

Complex Conjugate Poles ($0 < \zeta < 1$) – Underdamped

For this case, we define the damped frequency:

Damped frequency defined

$$\omega_d = \omega_n \sqrt{1-\zeta^2} \tag{4A.24}$$

so that the poles are located at:

Underdamped pole locations are complex conjugates

$$p_{1,2} = -\zeta\omega_n \pm j\omega_d \tag{4A.25}$$

The transfer function is then:

$$H(s) = \frac{\omega_n^2}{(s+\zeta\omega_n)^2 + \omega_d^2} \tag{4A.26}$$

and the transform of the unit-step response $Y(s) = H(s)/s$ can be expanded as:

$$Y(s) = \frac{1}{s} - \frac{(s+\zeta\omega_n)+\zeta\omega_n}{(s+\zeta\omega_n)^2+\omega_d^2} \tag{4A.27}$$

Thus, from the transform Eq. (4A.2b), the unit-step response is:

Step response of a second-order underdamped system

$$y(t) = 1 - \frac{\omega_n}{\omega_d} e^{-\zeta\omega_n t} \sin\left(\omega_d t + \cos^{-1}\zeta\right) \tag{4A.28}$$

Verify the above result.

The transient response is an exponentially decaying sinusoid with frequency ω_d rads[-1]. Thus second-order systems with complex poles have an oscillatory step response.

Second-Order Pole Locations

It is convenient to think of the second-order response in terms of the pole-zero plot of the second-order transfer function. Only two parameters determine the pole locations: ζ and ω_n. It is instructive how varying either parameter moves the pole locations around the s-plane. A graph of some pole locations for various values of ζ and one value of ω_n are shown below:

Second-order pole locations

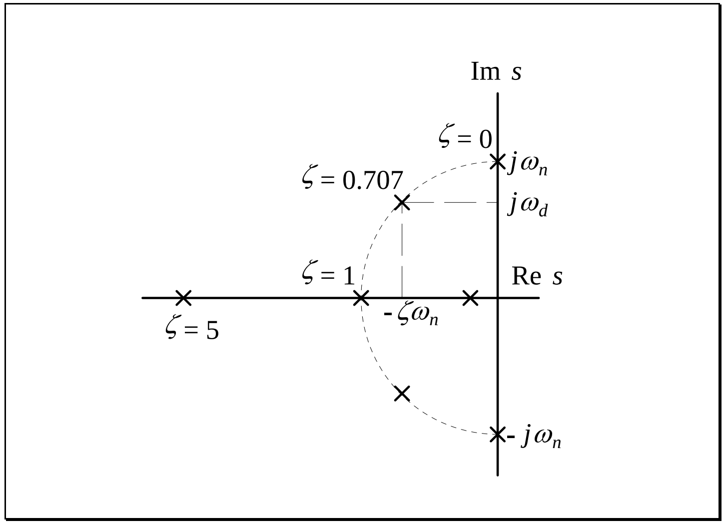


Figure 4A.1

We can see, for fixed ω_n, that varying ζ from 0 to 1 causes the poles to move from the imaginary axis along a circular arc with radius ω_n until they meet at the point $s = -\omega_n$. If $\zeta = 0$ then the poles lie on the imaginary axis and the transient never dies out – we have a *marginally stable* system. As ζ is increased, the response becomes less oscillatory and more and more damped, until $\zeta = 1$. Now the poles are real and repeated, and there is no sinusoid in the response. As ζ is increased, the poles move apart on the real axis, with one moving to the left, and one moving toward the origin. The response becomes more and more damped due to the right-hand pole getting closer and closer to the origin.

How the damping ratio, ζ, varies the pole location

Sinusoidal Response

Again consider a system with rational transfer function $H(s)=B(s)/A(s)$. To determine the system response to the sinusoidal input $x(t)=C\cos\omega_0 t$, we first find the Laplace transform of the input:

$$X(s)=\frac{Cs}{s^2+\omega_0^2}=\frac{Cs}{(s+j\omega_0)(s-j\omega_0)} \tag{4A.29}$$

The transform $Y(s)$ of the ZSR is then:

$$Y(s)=\frac{B(s)}{A(s)}\frac{Cs}{(s+j\omega_0)(s-j\omega_0)} \tag{4A.30}$$

The partial fraction expansion of Eq. (4A.30) is:

$$Y(s)=\frac{(C/2)H(j\omega_0)}{(s-j\omega_0)}+\frac{(C/2)H^*(j\omega_0)}{(s+j\omega_0)}+\frac{E(s)}{A(s)} \tag{4A.31}$$

You should confirm the partial fraction expansion.

The inverse Laplace transform of both sides of Eq. (4A.31) yields:

The sinusoidal response of a system

$$y(t)=\frac{C}{2}\left[H(j\omega_0)e^{j\omega_0 t}+H^*(j\omega_0)e^{-j\omega_0 t}\right]+y_{\mathrm{tr}}(t) \tag{4A.32}$$

and from Euler's identity, this can be written:

is sinusoidal (plus a transient)

$$y(t)=C\left|H(j\omega_0)\right|\cos(\omega_0 t+\angle H(j\omega_0))+y_{\mathrm{tr}}(t) \tag{4A.33}$$

When the system is stable, the $y_{\mathrm{tr}}(t)$ term decays to zero and we are left with the steady-state response:

Steady-state sinusoidal response of a system

$$y_{\mathrm{ss}}(t)=C\left|H(j\omega_0)\right|\cos(\omega_0 t+\angle H(j\omega_0)),\quad t\geq 0 \tag{4A.34}$$

This result is exactly the same expression as that found when performing Fourier analysis, except there the expression was for all time, and hence there was no transient! This means that the frequency response function $H(\omega)$ can be obtained directly from the transfer function:

Frequency response from transfer function

$$\boxed{H(\omega)=H(j\omega)=H(s)\big|_{s=j\omega}} \tag{4A.35}$$

Arbitrary Response

Suppose we apply an arbitrary input $x(t)$ that has rational Laplace transform $X(s)=C(s)/D(s)$ where the degree of $C(s)$ is less than $D(s)$. If this input is applied to a system with transfer function $H(s)=B(s)/A(s)$, the transform of the response is:

$$Y(s)=\frac{B(s)C(s)}{A(s)D(s)} \tag{4A.36}$$

This can be written as:

$$Y(s)=\frac{F(s)}{D(s)}+\frac{E(s)}{A(s)} \tag{4A.37}$$

Taking the inverse transform of both sides gives:

The Laplace transform of an output signal always contains a steady-state term and a transient term

$$y(t)=y_{\mathrm{ss}}(t)+y_{\mathrm{tr}}(t) \tag{4A.38}$$

where $y_{\mathrm{ss}}(t)$ is the inverse transform of $F(s)/D(s)$ and $y_{\mathrm{tr}}(t)$ is the inverse transform of $E(s)/A(s)$.

The *form* of the response is determined by the poles *only*

The important point to note about this simple analysis is that the *form* of the transient response is determined by the poles of the system transfer function $H(s)$ regardless of the particular form of the input signal $x(t)$, while the *form* of the steady-state response depends directly on the poles of the input $X(s)$, regardless of what the system transfer function $H(s)$ is!

Summary

- A system is stable if all the poles of the transfer function lie in the open left-half *s*-plane.
- The complete response of a system consists of a transient response and a steady-state response. The transient response consists of the ZIR and a part of the ZSR. The steady-state response is part of the ZSR. The transfer function gives us the ZSR *only*!
- The step response is an important response because it occurs so frequently in engineering applications – control systems in particular. Second-order systems exhibit different step responses depending on their pole locations – overdamped, critically damped and underdamped.
- The frequency response of a system can be obtained from the transfer function by setting $s = j\omega$.
- The poles of the *system* determine the *transient* response.
- The poles of the *signal* determine the *steady-state* response.

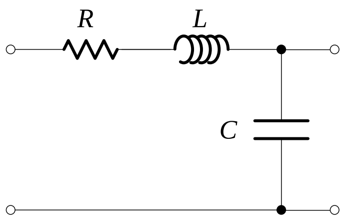
References

Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB®*, Prentice-Hall, 1997.

Exercises

1.

Determine whether the following circuit is *stable* for any element values, and for any bounded inputs:



2.

Suppose that a system has the following transfer function:

$$H(s) = \frac{8}{s+4}$$

a) Compute the system response to the following inputs. Identify the steady-state solution and the transient solution.

(i) $x(t) = u(t)$

(ii) $x(t) = tu(t)$

(iii) $x(t) = 2\sin(2t)u(t)$

(iv) $x(t) = 2\sin(10t)u(t)$

b) Use MATLAB® to compute the responses numerically. Plot the responses and compare them to the responses obtained analytically in part a).

3.

Consider three systems which have the following transfer functions:

(i) $H(s) = \dfrac{32}{s^2 + 4s + 16}$

(ii) $H(s) = \dfrac{32}{s^2 + 8s + 16}$

(iii) $H(s) = \dfrac{32}{s^2 + 10s + 16}$

For each system:

a) Determine if the system is critically damped, underdamped, or overdamped.

b) Calculate the step response of the system.

c) Use MATLAB® to compute the step response numerically. Plot the response and compare it to the plot of the response obtained analytically in part b).

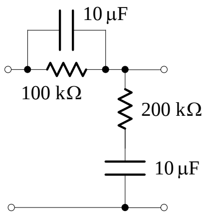
4.

For the circuit shown, compute the steady-state response $y_{ss}(t)$ resulting from the inputs given below assuming that there is no initial energy at time $t = 0$.

a) $x(t) = u(t)$

b) $x(t) = 10\cos(t)u(t)$

c) $x(t) = \cos(5t - \pi/6)u(t)$



Oliver Heaviside (1850-1925)

The mid-Victorian age was a time when the divide between the rich and the poor was immense (and almost insurmountable), a time of unimaginable disease and lack of sanitation, a time of steam engines belching forth a steady rain of coal dust, a time of horses clattering along cobblestoned streets, a time when social services were the fantasy of utopian dreamers. It was into this smelly, noisy, unhealthy and class-conscious world that Oliver Heaviside was born the son of a poor man on 18 May, 1850.

A lucky marriage made Charles Wheatstone (of Wheatstone Bridge fame) Heaviside’s uncle. This enabled Heaviside to be reasonably well educated, and at the age of sixteen he obtained his first (and last) job as a telegraph operator with the Danish-Norwegian-English Telegraph Company. It was during this job that he developed an interest in the physical operation of the telegraph cable. At that time, telegraph cable theory was in a state left by Professor William Thomson (later Lord Kelvin) – a diffusion theory modelling the passage of electricity through a cable with the same mathematics that describes heat flow.

I remember my first look at the great treatise of Maxwell’s…I saw that it was great, greater and greatest, with prodigious possibilities in its power. – Oliver Heaviside

By the early 1870’s, Heaviside was contributing technical papers to various publications – he had taught himself calculus, differential equations, solid geometry and partial differential equations. But the greatest impact on Heaviside was Maxwell’s treatise on electricity and magnetism – Heaviside was swept up by its power.

In 1874 Heaviside resigned from his job as telegraph operator and went back to live with his parents. He was to live off his parents, and other relatives, for the rest of his life. He dedicated his life to writing technical papers on telegraphy and electrical theory – much of his work forms the basis of modern circuit theory and field theory.

In 1876 he published a paper entitled *On the extra current* which made it clear that Heaviside (a 26-year-old unemployed nobody) was a brilliant talent. He had extended the mathematical understanding of telegraphy far beyond Thomson's submarine cable theory. It showed that inductance was needed to permit finite-velocity wave propagation, and would be the key to solving the problems of long distance telephony. Unfortunately, although Heaviside's paper was correct, it was also unreadable by all except a few – this was a trait of Heaviside that would last all his life, and led to his eventual isolation from the "academic world". In 1878, he wrote a paper *On electromagnets, etc.* which introduced the expressions for the AC impedances of resistors, capacitors and inductors. In 1879, his paper *On the theory of faults* showed that by "faulting" a long telegraph line with an inductance, it would actually improve the signalling rate of the line – thus was born the idea of "inductive loading", which allowed transcontinental telegraphy and long-distance telephony to be developed in the USA.

Now all has been blended into one theory, the main equations of which can be written on a page of a pocket notebook. That we have got so far is due in the first place to Maxwell, and next to him to Heaviside and Hertz. – H.A. Lorentz

When Maxwell died in 1879 he left his electromagnetic theory as twenty equations in twenty variables! It was Heaviside (and independently, Hertz) who recast the equations in modern form, using a symmetrical vector calculus notation (also championed by Josiah Willard Gibbs (1839-1903)). From these equations, he was able to solve an enormous amount of problems involving field theory, as well as contributing to the *ideas* behind field theory, such as energy being carried by fields, and not electric charges.

Rigorous mathematics is narrow, physical mathematics bold and broad. – Oliver Heaviside

A major portion of Heaviside's work was devoted to "operational calculus". This caused a controversy with the mathematicians of the day because although it seemed to solve physical problems, it's mathematical rigor was not at all clear. His knowledge of the physics of problems guided him correctly in many instances to the development of suitable mathematical processes. In 1887 Heaviside introduced the concept of a *resistance operator*, which in modern terms would be called *impedance*, and Heaviside introduced the symbol Z for it. He let p be equal to time-differentiation, and thus the resistance operator for an inductor would be written as pL. He would then treat p just like an algebraic quantity, and solve for voltage and current in terms of a power series in p. In other words, Heaviside's operators allowed the reduction of the *differential* equations of a physical system to equivalent *algebraic* equations.

Heaviside was fond of using the unit-step as an input to electrical circuits, especially since it was a very practical matter to send such pulses down a telegraph line. The unit-step was even called the Heaviside step, and given the symbol $H(t)$, but Heaviside simply used the notation **1**. He was tantalizingly close to discovering the impulse by stating "… $p \cdot \mathbf{1}$ means a function of t which is wholly concentrated at the moment $t = 0$, of total amount 1. It is an impulsive function, so to speak…[it] involves only ordinary ideas of differentiation and integration pushed to their limit."

Paul Dirac derived the modern notion of the impulse, when he used it in 1927, at age 25, in a paper on quantum mechanics. He did his undergraduate work in electrical engineering and was both familiar with all of Heaviside's work and a great admirer of his.

Heaviside also played a role in the debate raging at the end of the 19th century about the age of the Earth, with obvious implications for Darwin's theory of evolution. In 1862 Thomson wrote his famous paper *On the secular cooling of the Earth*, in which he imagined the Earth to be a uniformly heated ball of molten rock, modelled as a semi-infinite mass. Based on experimentally derived thermal conductivity of rock, sand and sandstone, he then mathematically allowed the globe to cool according to the physical law of thermodynamics embedded in Fourier's famous partial differential equation for heat flow. The resulting age of the Earth (100 million years) fell short of that needed by Darwin's theory, and also went against geologic and palaeontologic evidence. John Perry (a professor of mechanical engineering) redid Thomson's analysis using discontinuous diffusivity, and arrived at approximate results that could (based on the conductivity and specific heat of marble and quartz) put the age of the Earth into the billions of years. But Heaviside, using his operational calculus, was able to solve the diffusion equation for a finite spherical Earth. We now know that such a simple model is based on faulty premises – radioactive decay within the Earth maintains the thermal gradient without a continual cooling of the planet. But the power of Heaviside's methods to solve remarkably complex problems became readily apparent.

The practice of eliminating the physics by reducing a problem to a purely mathematical exercise should be avoided as much as possible. The physics should be carried on right through, to give life and reality to the problem, and to obtain the great assistance which the physics gives to the mathematics. – Oliver Heaviside, *Collected Works*, Vol II, p.4

Throughout his "career", Heaviside released 3 volumes of work entitled *Electromagnetic Theory*, which was really just a collection of his papers. Heaviside shunned all honours, brushing aside his honorary doctorate from the University of Göttingen and even refusing to accept the medal associated with his election as a Fellow of the Royal Society, in 1891.

4A.18

In 1902, Heaviside wrote an article for the *Encyclopedia Britannica* entitled *The theory of electric telegraphy*. Apart from developing the wave propagation theory of telegraphy, he extended his essay to include “wireless” telegraphy, and explained how the remarkable success of Marconi transmitting from Ireland to Newfoundland might be due to the presence of a permanently conducting upper layer in the atmosphere. This supposed layer was referred to as the “Heaviside layer”, which was directly detected by Edward Appleton and M.A.F. Barnett in the mid-1920s. Today we merely call it the “ionosphere”.

Heaviside spent much of his life being bitter at those who didn’t recognise his genius – he had disdain for those that could not accept his mathematics without formal proof, and he felt betrayed and cheated by the scientific community who often ignored his results or used them later without recognising his prior work. It was with much bitterness that he eventually retired and lived out the rest of his life in Torquay on a government pension. He withdrew from public and private life, and was taunted by “insolently rude imbeciles”. Objects were thrown at his windows and doors and numerous practical tricks were played on him.

Heaviside should be remembered for his vectors, his field theory analyses, his brilliant discovery of the distortionless circuit, his pioneering applied mathematics, and for his wit and humor. – P.J. Nahin

Today, the historical obscurity of Heaviside’s work is evident in the fact that his vector analysis and vector formulation of Maxwell’s theory have become “basic knowledge”. His operational calculus was made obsolete with the 1937 publication of a book by the German mathematician Gustav Doetsch – it showed how, with the Laplace transform, Heaviside’s operators could be replaced with a systematic, routine method.

The last five years of Heaviside’s life, with both hearing and sight failing, were years of great privation and mystery. He died on 3rd February, 1925.

References

Nahin, P.: *Oliver Heaviside: Sage in Solitude*, IEEE Press, 1988.

Lecture 4B – Frequency Response

Frequency response function. Bode plots. Approximate Bode plots. Transfer function synthesis. Digital filters.

Overview

An examination of a system's frequency response is useful in several respects. It can help us determine things such as the DC gain and bandwidth, how well a system meets the stability criterion, and whether the system is robust to disturbance inputs.

Despite all this, remember that the time- and frequency-domain are inextricably related – we can't alter the characteristics of one without affecting the other. This will be demonstrated for a second-order system later.

Frequency Response Function

Recall that for a LTI system characterized by $H(s)$, and for a sinusoidal input $x(t) = A\cos(\omega_0 t)$, the steady-state response is:

$$y_{ss}(t) = A\left|H(\omega_0)\right|\cos\left(\omega_0 t + \angle H(\omega_0)\right) \tag{4B.1}$$

where $H(\omega)$ is the *frequency response* function, obtained by setting $s = j\omega$ in $H(s)$. Thus, the system behaviour for sinusoidal inputs is completely specified by the *magnitude response* $\left|H(\omega)\right|$ and the *phase response* $\angle H(\omega)$.

The definition above is precisely how we determine the frequency response experimentally – we input a sinusoid and, in the steady-state, measure the magnitude and phase change at the output.

Determining the Frequency Response from a Transfer Function

We can get the frequency response of a system by manipulating its transfer function. Consider a simple first-order transfer function:

$$H(s)=\frac{K}{s+p} \tag{4B.2}$$

The sinusoidal steady state corresponds to $s=j\omega$. Therefore, Eq. (4B.2) is, for the sinusoidal steady state:

$$H(\omega)=\frac{K}{j\omega+p} \tag{4B.3}$$

The complex function $H(\omega)$ can also be written using a complex exponential in terms of magnitude and phase:

$$\boxed{H(\omega)=\left|H(\omega)\right|e^{j\angle H(\omega)}} \tag{4B.4a}$$

which is normally written in polar coordinates:

The transfer function in terms of magnitude and phase

$$\boxed{H(\omega)=\left|H(\omega)\right|\angle H(\omega)} \tag{4B.4b}$$

We plot the magnitude and phase of $H(\omega)$ as a function of ω or f. We use both linear and logarithmic scales.

If the logarithm (base 10) of the *magnitude* is multiplied by 20, then we have the *gain* of the transfer function in decibels (dB):

The magnitude of the transfer function in dB

$$\boxed{\left|H(\omega)\right|_{\mathrm{dB}}=20\log\left|H(\omega)\right|\ \mathrm{dB}} \tag{4B.5}$$

A *negative* gain in decibels is referred to as *attenuation*. For example, -3 dB gain is the same as 3 dB attenuation.

The phase function is usually plotted in degrees.

For example, in Eq. (4B.2), let $K=p=\omega_0$ so that:

$$H(\omega)=\frac{1}{1+j\omega/\omega_0} \tag{4B.6}$$

The magnitude function is found directly as:

$$\left|H(\omega)\right|=\frac{1}{\sqrt{1+(\omega/\omega_0)^2}} \tag{4B.7}$$

and the phase is:

$$\angle H(\omega)=-\tan^{-1}\left(\frac{\omega}{\omega_0}\right) \tag{4B.8}$$

Magnitude Responses

A magnitude response is the magnitude of the transfer function for a sinusoidal steady-state input, plotted against the frequency of the input. Magnitude responses can be classified according to their particular properties. To look at these properties, we will use linear magnitude versus linear frequency plots. For the simple first-order *RC* circuit that you are so familiar with, the magnitude function given by Eq. (4B.7) has three frequencies of special interest corresponding to these values of $\left|H(\omega)\right|$:

The magnitude response is the magnitude of the transfer function in the sinusoidal steady state

$$\begin{aligned}\left|H(0)\right|&=1\\ \left|H(\omega_0)\right|&=\frac{1}{\sqrt{2}}\approx 0.707\\ \left|H(\infty)\right|&\to 0\end{aligned} \tag{4B.9}$$

The frequency ω_0 is known as the *half-power frequency*. The plot below shows the complete magnitude response of $H(\omega)$ as a function of ω, and the circuit that produces it:

A simple lowpass filter

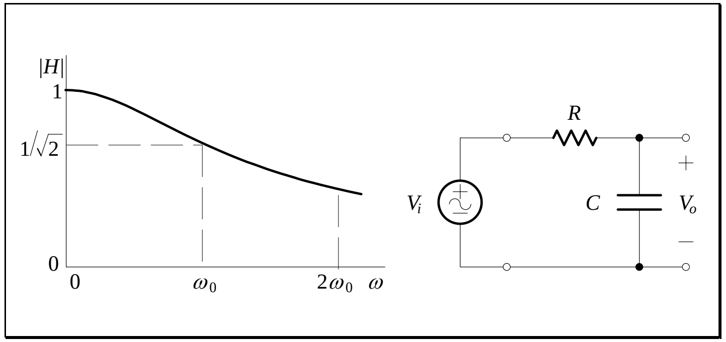


Figure 4B.1

An idealisation of the response in Figure 4B.1, known as a *brick wall*, and the circuit that produces it are shown below:

An ideal lowpass filter

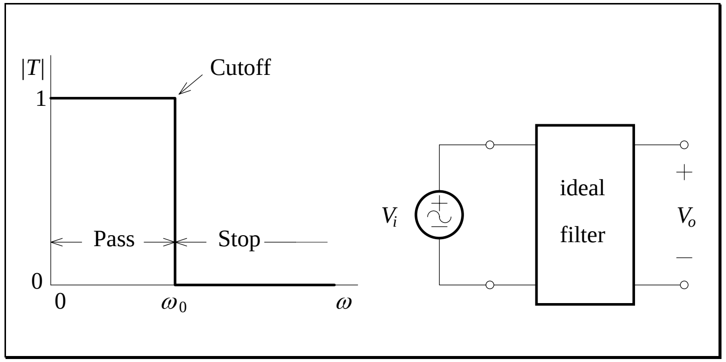


Figure 4B.2

Pass and stop bands defined

For the ideal filter, the output voltage remains fixed in amplitude until a critical frequency is reached, called the *cutoff frequency*, ω_0. At that frequency, and for all higher frequencies, the output is zero. The range of frequencies with output is called the *passband*; the range with no output is called the *stopband*. The obvious classification of the filter is a *lowpass* filter.

Even though the response shown in the plot of Figure 4B.1 differs from the ideal, it is still known as a lowpass filter, and, by convention, the half-power frequency is taken as the cutoff frequency.

If the positions of the resistor and capacitor in the circuit of Figure 4B.1 are interchanged, then the resulting circuit is:

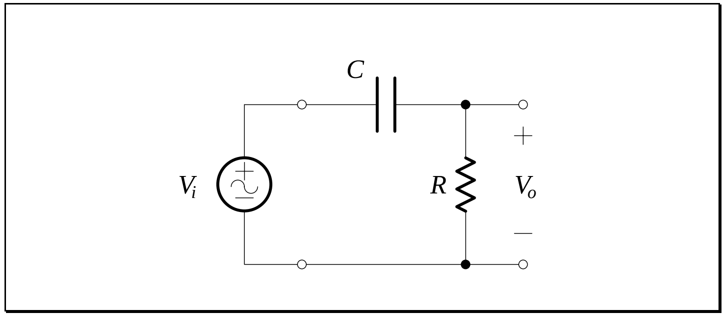


Figure 4B.3

Show that the transfer function is:

$$H(s)=\frac{s}{s+1/RC} \tag{4B.10}$$

Letting $1/RC=\omega_0$ again, and with $s=j\omega$, we obtain:

$$H(\omega)=\frac{j\omega/\omega_0}{1+j\omega/\omega_0} \tag{4B.11}$$

The magnitude function of this equation, at the three frequencies given in Eq. (4B.9), is:

$$\begin{aligned}|H(0)|&=0\\ |H(\omega_0)|&=\frac{1}{\sqrt{2}}\approx 0.707\\ |H(\infty)|&\to 1\end{aligned} \tag{4B.12}$$

The plot below shows the complete magnitude response of $H(\omega)$ as a function of ω, and the circuit that produces it:

A simple highpass filter

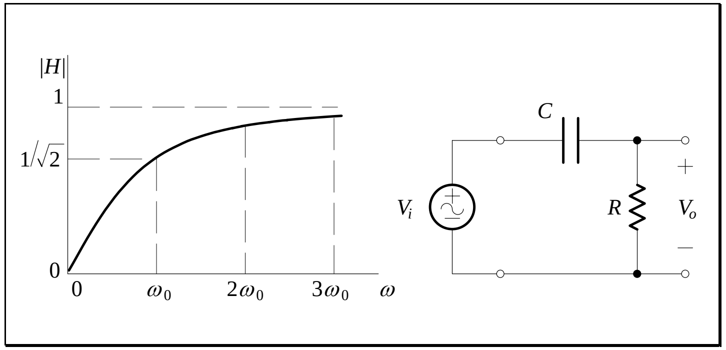


Figure 4B.4

This filter is classified as a *highpass filter*. The ideal brick wall highpass filter is shown below:

An ideal highpass filter

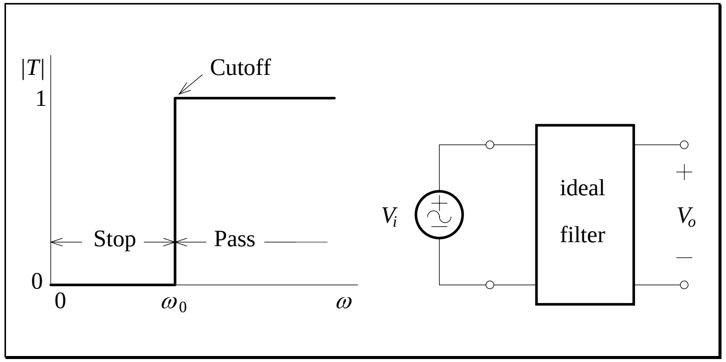


Figure 4B.5

The cutoff frequency is ω_0, as it was for the lowpass filter.

Phase Responses

Phase response is obtained in the sinusoidal steady state

Like magnitude responses, phase responses are only meaningful when we look at sinusoidal steady-state signals. A transfer function for a sinusoidal input is:

$$H(\omega)=\frac{Y}{X}=\frac{|Y|\angle\theta}{|X|\angle 0}=|H|\angle\theta \tag{4B.13}$$

where the input is taken as the phase reference (zero phase).

For the bilinear transfer function:

$$H(\omega)=K\frac{j\omega+z}{j\omega+p} \tag{4B.14}$$

the phase is:

The phase of the bilinear transfer function

$$\boxed{\theta=\angle K+\tan^{-1}\left(\frac{\omega}{z}\right)-\tan^{-1}\left(\frac{\omega}{p}\right)} \tag{4B.15}$$

Lead and lag circuits defined

We use the sign of this phase angle to classify systems. Those giving positive θ are known as *lead* systems, those giving negative θ as *lag* systems.

For the simple *RC* circuit of Figure 4B.5, for which $H(\omega)$ is given by Eq. (4B.6), we have:

$$\theta=-\tan^{-1}\left(\frac{\omega}{\omega_0}\right) \tag{4B.16}$$

Since θ is negative for all ω, the circuit is a lag circuit. When $\omega=\omega_0$, $\theta=-\tan^{-1}(1)=-45°$.

A complete plot of the phase response is shown below:

Lagging phase response for a simple lowpass filter

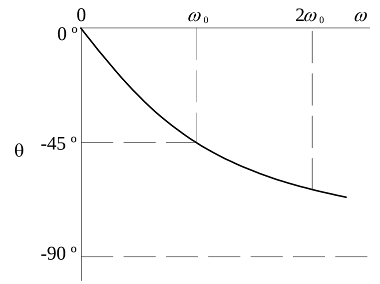


Figure 4B.6

For the circuit in Figure 4B.5, *show that the phase is given by:*

$$\theta = 90° - \tan^{-1}\left(\frac{\omega}{\omega_0}\right) \qquad (4B.17)$$

The phase response has the same shape as Figure 4B.6 but is shifted upward by $90°$:

Leading phase response for a simple highpass filter

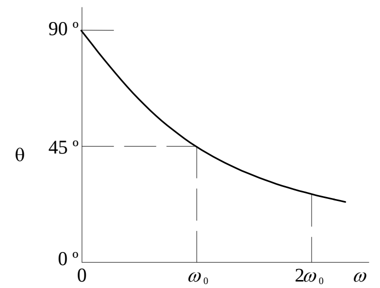


Figure 4B.7

The angle θ is positive for all ω, and so the circuit is a lead circuit.

Frequency Response of a Lowpass Second-Order System

Starting from the usual definition of a lowpass second-order system transfer function, we get the following frequency response function:

$$H(\omega) = \frac{\omega_n^2}{\omega_n^2 - \omega^2 + j2\zeta\omega_n\omega} \qquad (4B.18)$$

The magnitude is:

The magnitude response of a lowpass second order transfer function

$$\boxed{\left|H(j\omega)\right| = \frac{\omega_n^2}{\sqrt{\left(\omega_n^2 - \omega^2\right)^2 + \left(2\zeta\omega_n\omega\right)^2}}} \qquad (4B.19)$$

and the phase is:

The phase response of a lowpass second order transfer function

$$\boxed{\theta = -\tan^{-1}\left(\frac{2\zeta\omega_n\omega}{\omega_n^2 - \omega^2}\right)} \qquad (4B.20)$$

The magnitude and phase functions are plotted below for $\zeta = 0.4$:

Typical magnitude and phase responses of a lowpass second order transfer function

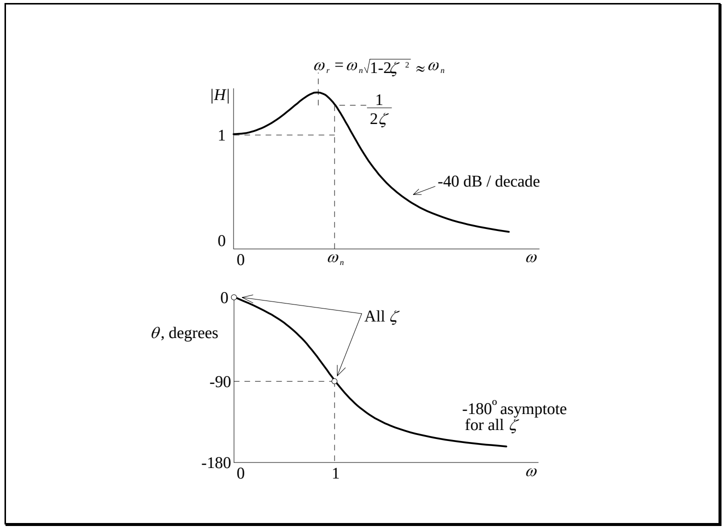


Figure 4B.8

For the magnitude function, from Eq. (4B.19) we see that:

$$|H(0)| = 1, \quad |H(\omega_n)| = 1/2\zeta, \quad |H(\infty)| \to 0 \tag{4B.21}$$

and for large ω, the magnitude decreases at a rate of -40 dB per decade, which is sometimes described as *two-pole rolloff*.

For the phase function, we see that:

$$\theta(0) = 0°, \quad \theta(\omega_n) = -90°, \quad \theta(\infty) \to -180° \tag{4B.22}$$

Visualization of the Frequency Response from a Pole-Zero Plot

The frequency response can be visualised in terms of the pole locations of the transfer function. For example, for a second-order lowpass system:

Standard form for a lowpass second order transfer function

$$\boxed{H(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}} \tag{4B.23}$$

the poles are located on a circle of radius ω_n and at an angle with respect to the negative real axis of $\psi = \cos^{-1}(\zeta)$. These complex conjugate pole locations are shown below:

Pole locations for a lowpass second order transfer function

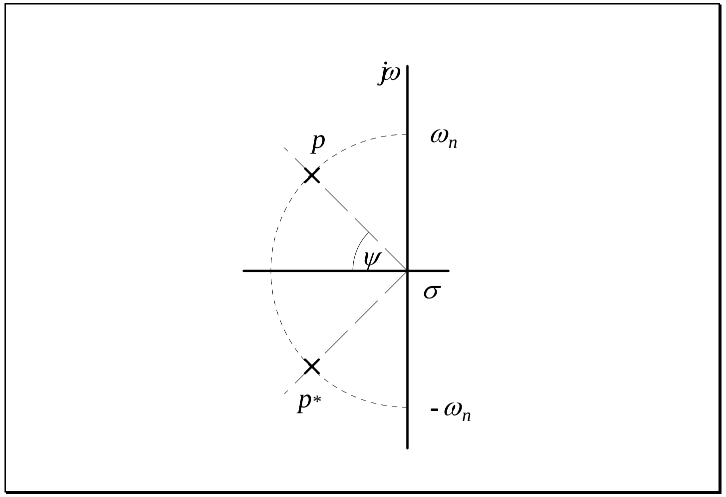


Figure 4B.9

In terms of the poles shown in Figure 4B.9, the transfer function is:

Lowpass second order transfer function using pole factors

$$H(s) = \frac{\omega_n^2}{(s-p)(s-p^*)} \tag{4B.24}$$

With $s = j\omega$, the two factors in this equation become:

Polar representation of the pole factors

$$j\omega - p = m_1 \angle \phi_1 \quad \text{and} \quad j\omega - p^* = m_2 \angle \phi_2 \tag{4B.25}$$

In terms of these quantities, the magnitude and phase are:

Magnitude function written using the polar representation of the pole factors

$$\left|H(\omega)\right| = \frac{1}{m_1 m_2} \tag{4B.26}$$

and:

Phase function written using the polar representation of the pole factors

$$\theta = -\left(\phi_1 + \phi_2\right) \tag{4B.27}$$

Vectors representing Eq. (4B.25) are shown below:

Determining the magnitude and phase response from the s plane

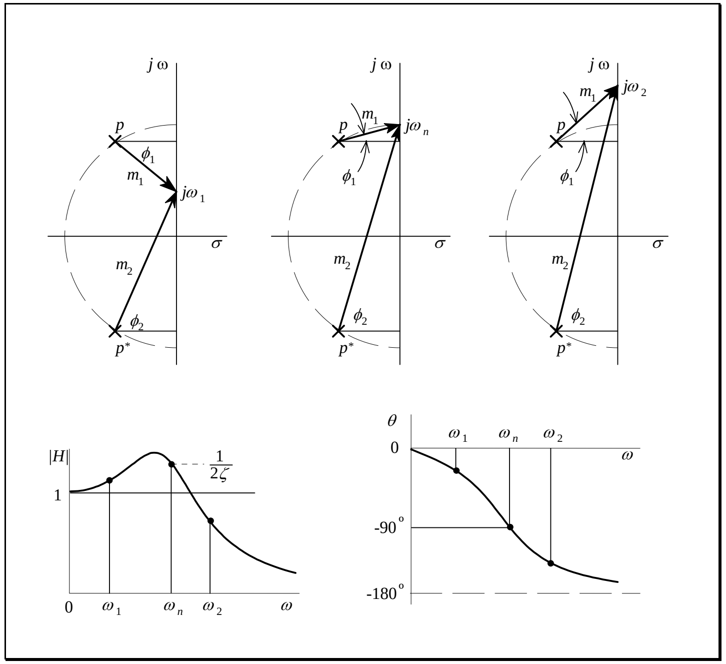


Figure 4B.10

Figure 4B.10 shows three different frequencies - one below ω_n, one at ω_n, and one above ω_n. From this construction we can see that the short length of m_1 near the frequency ω_n is the reason why the magnitude function reaches a peak near ω_n. These plots are useful in visualising the frequency response of the circuit.

Bode Plots

Bode* plots are plots of the magnitude function $\left|H(\omega)\right|_{\text{dB}} = 20\log\left|H(\omega)\right|$ and the phase function $\angle H(\omega)$, where the scale of the frequency variable (usually ω) is logarithmic. The use of logarithmic scales has several desirable properties:

The advantages of using Bode plots

- we can *approximate* a frequency response with straight lines. This is called an *approximate* Bode plot.
- the shape of a Bode plot is preserved if we decide to scale the frequency – this makes design easy.
- we add and subtract individual factors in a transfer function, rather than multiplying and dividing.
- the slope of all lines in a magnitude plot is $\pm 20n\,\text{dB/decade}$, and $\pm n45°/\text{decade}$ for phase plots, where n is any integer.
- by examining a few features of a Bode plot, we can readily determine the transfer function (for simple systems).

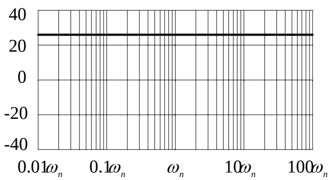
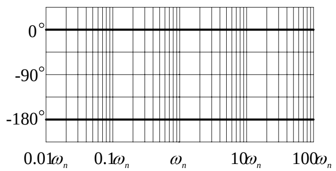
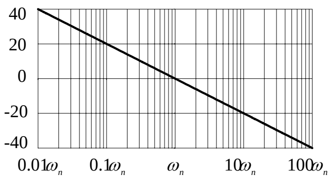
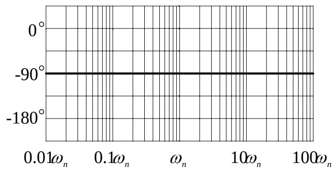
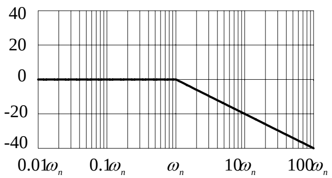
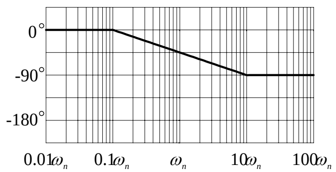
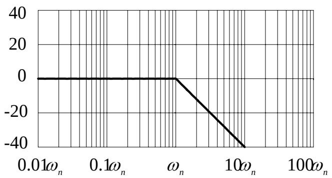
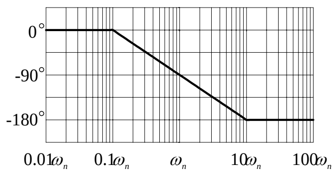
We normally don’t deal with equations when drawing Bode plots – we rely on our knowledge of the asymptotic approximations for the handful of factors that go to make up a transfer function.

---

* Dr. Hendrik Bode grew up in Urbana, Illinois, USA, where his name is pronounced *boh-dee*. Purists insist on the original Dutch *boh-dah*. No one uses *bohd*.

Approximating Bode Plots using Transfer Function Factors

The table below gives transfer function factors and their corresponding magnitude asymptotic plots and phase linear approximations:

Transfer Function Factor	Magnitude Asymptote $\lvert H\rvert$, dB	Phase Linear Approximation $\angle H$, °
K		
$\dfrac{1}{(s/\omega_n)}$		
$\dfrac{1}{(s/\omega_n+1)}$		
$\dfrac{1}{\left(\dfrac{s^2}{\omega_n^2}+\dfrac{2\zeta}{\omega_n}s+1\right)}$		

The corresponding numerator factors are obtained by "mirroring" the above plots about the 0 dB line and 0° line.

Transfer Function Synthesis

One of the main reasons for using Bode plots is that we can synthesise a desired frequency response by placing poles and zeros appropriately. This is easy to do asymptotically, and the results can be checked using MATLAB®.

A band enhancement filter

Example – Band-Enhancement Filter

The asymptotic Bode plot shown below is for a band-enhancement filter:

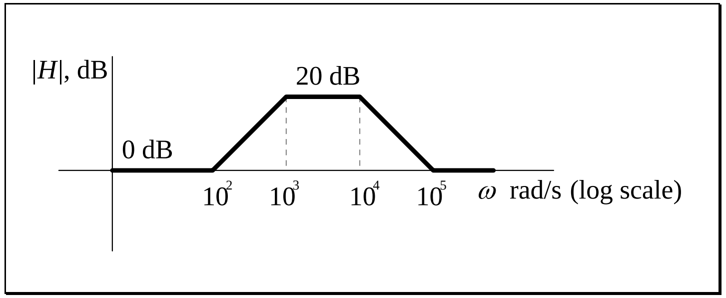


Figure 4B.11

We wish to provide additional gain over a narrow band of frequencies, leaving the gain at higher and lower frequencies unchanged. We wish to design a filter to these specifications and the additional requirement that all capacitors have the value $C = 10$ nF.

The composite plot may be decomposed into four first-order factors as shown below:

Decomposing a Bode plot into first-order factors

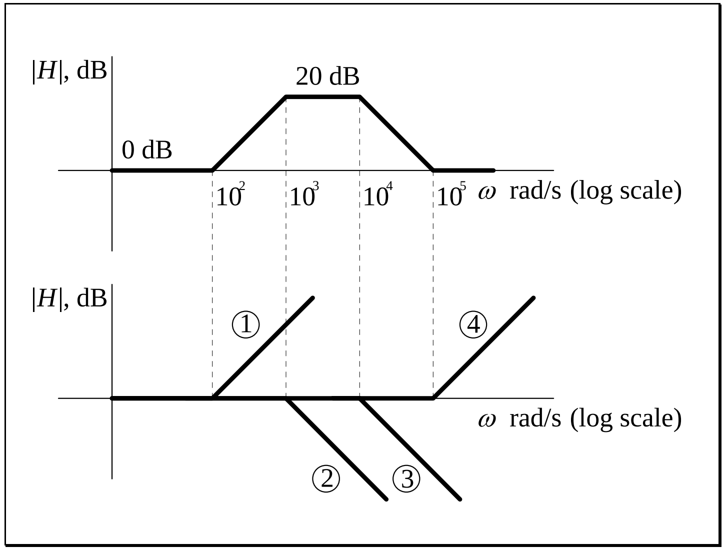


Figure 4B.12

Those marked 1 and 4 represent zero factors, while those marked 2 and 3 are pole factors. The pole-zero plot corresponding to these factors is shown below:

The pole-zero plot corresponding to the Bode plot

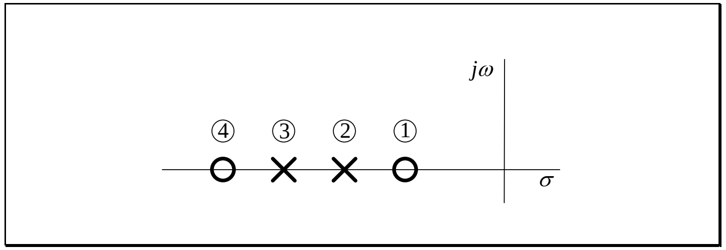


Figure 4B.13

From the break frequencies given, we have:

$$H(\omega)=\frac{\left(1+j\omega/10^2\right)\left(1+j\omega/10^5\right)}{\left(1+j\omega/10^3\right)\left(1+j\omega/10^4\right)} \tag{4B.28}$$

Substituting s for $j\omega$ gives the transfer function:

The transfer function corresponding to the Bode plot

$$H(s)=\frac{\left(s+10^2\right)\left(s+10^5\right)}{\left(s+10^3\right)\left(s+10^4\right)} \tag{4B.29}$$

We next write $H(s)$ as a product of bilinear functions. The choice is arbitrary, but one possibility is:

The transfer function as a cascade of bilinear functions

$$H(s)=H_1(s)H_2(s)=\frac{s+10^2}{s+10^3}\times\frac{s+10^5}{s+10^4} \tag{4B.30}$$

For a circuit realisation of H_1 and H_2 we decide to use an inverting op-amp circuit that implements a bilinear transfer function:

A realisation of the specifications

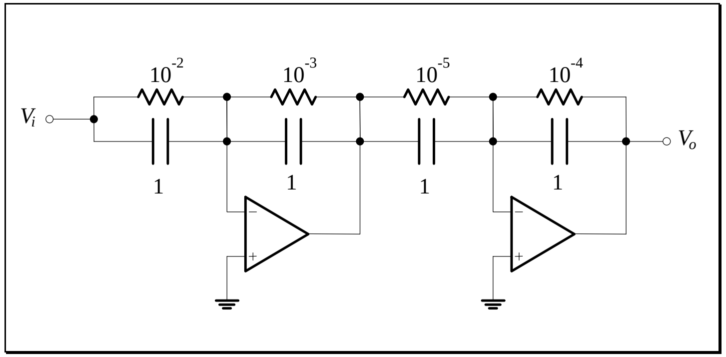


Figure 4B.14

To obtain realistic element values, we need to scale the components so that the transfer function remains unaltered. This is accomplished with the equations:

Magnitude scaling is required to get realistic element values

$$C_{\mathrm{new}} = \frac{1}{k_m} C_{\mathrm{old}} \quad \text{and} \quad R_{\mathrm{new}} = k_m R_{\mathrm{old}} \tag{4B.31}$$

Since the capacitors are to have the value 10 nF, this means $k_m = 10^8$. The element values that result are shown below and the design is complete:

A realistic implementation of the specifications

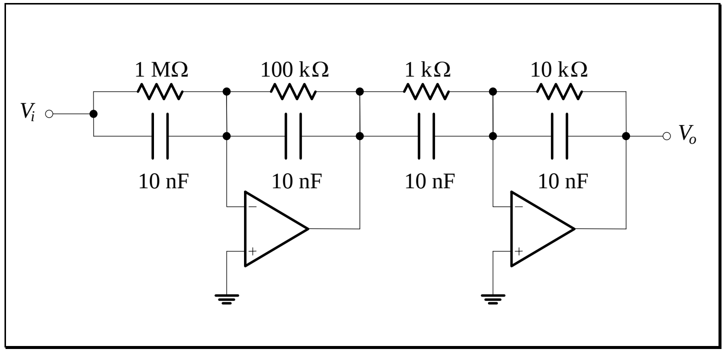


Figure 4B.15

In this simple example, the response only required placement of the poles and zeros on the real axis. However, complex pole-pair placement is not unusual in design problems.

Digital Filters

Digital filters use analog filters

Digital filtering involves sampling, quantising and coding of the input analog signal (using an analog to digital converter, or ADC for short). Once we have converted voltages to mere numbers, we are free to do any processing on them that we desire. Usually, the signal's spectrum is found using a *fast Fourier transform*, or FFT. The spectrum can then be modified by scaling the amplitudes and adjusting the phase of each sinusoid. An inverse FFT can then be performed, and the processed numbers are converted back into analog form (using a digital to analog converter, or DAC). In modern digital signal processors, an operation corresponding to a fast convolution is also sometimes employed – that is, the signal is convolved in the time-domain in real-time.

The components of a digital filter are shown below:

The components of a digital filter

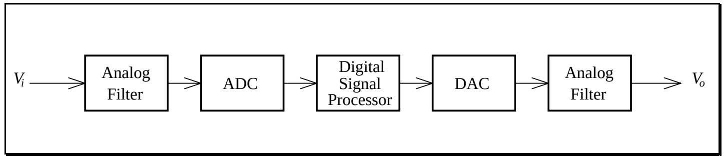


Figure 4B.16

Digital filter advantages

The digital signal processor can be custom built digital circuitry, or it can be a general purpose computer. There are many advantages of digitally processing analog signals:

1. A digital filter may be just a small part of a larger system, so it makes sense to implement it in software rather than hardware.

2. The cost of digital implementation is often considerably lower than that of its analog counterpart (and it is falling all the time).

3. The accuracy of a digital filter is dependent only on the processor word length, the quantising error in the ADC and the sampling rate.

4. Digital filters are generally unaffected by such factors as component accuracy, temperature stability, long-term drift, etc. that affect analog filter circuits.

5. Many circuit restrictions imposed by physical limitations of analog devices can be circumvented in a digital processor.

6. Filters of high order can be realised directly and easily.

7. Digital filters can be modified easily by changing the algorithm of the computer.

8. Digital filters can be designed that are always stable.

9. Filter responses can be made which always have linear phase (constant delay), regardless of the magnitude response.

Digital filter disadvantages

Some disadvantages are:

1. Processor speed limits the frequency range over which digital filters can be used (although this limit is continuously being pushed back with ever faster processors).

2. Analog filters (and signal conditioning) are still necessary to convert the analog signal to digital form and back again.

Summary

- A frequency response consists of two parts – a magnitude response and a phase response. It tells us the change in the *magnitude* and *phase* of a sinusoid at any frequency, in the steady-state.

- Bode plots are magnitude (dB) and phase responses drawn on a semi-log scale, enabling the easy analysis or design of high-order systems.

References

Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB®*, Prentice-Hall, 1997.

Exercises

1.

With respect to a reference frequency $f_0 = 20\text{ Hz}$, find the frequency which is (a) 2 decades above f_0 and (b) 3 octaves below f_0.

2.

Express the following magnitude ratios in dB: (a) 1, (b) 40, (c) 0.5

3.

Draw the approximate Bode plots (both magnitude and phase) for the transfer functions shown. Use MATLAB® to draw the exact Bode plots and compare.

(a) $G(s)=10$ (b) $G(s)=\dfrac{4}{s}$ (c) $G(s)=\dfrac{1}{10s+1}$

(d) $G(s)=\dfrac{1}{10s-1}$ (e) $G(s)=5(s+1)$ (f) $G(s)=5(s-1)$

Note that the magnitude plots for the transfer functions (c) and (d); (e) and (f) are the same. Why?

4.

Prove that if $G(s)$ has a single pole at $s=-1/\tau$ the asymptotes of the log magnitude response versus log frequency intersect at $\omega=1/\tau$. Prove this not only analytically but also graphically using MATLAB®.

5.

Make use of the property that the logarithm converts multiplication and division into addition and subtraction, respectively, to draw the Bode plot for:

$$G(s)=\frac{100(s+1)}{s(0.01s+1)}$$

Use asymptotes for the magnitude response and a linear approximation for the phase response.

6.

Draw the exact Bode plot using MATLAB® (magnitude and phase) for

$$G(s)=\frac{100s^2}{s^2+2\zeta\omega_n s+\omega_n^2}$$

when

$$\omega_n=10\text{ rads}^{-1}\text{ and }\zeta=0.3$$

Compare this plot with the approximate Bode plot (which ignores the value of ζ).

7.

Given

(a) $G(s)=\dfrac{20(6.66s+1)}{s(0.1s+1)^2(94.6s+1)}$

(b) $G(s)=\dfrac{4(-s+1.5+s^2)}{s^2(10-s)}$

Draw the approximate Bode plots and from these graphs find $|G|$ and $\angle G$ at (i) $\omega=0.1\text{ rads}^{-1}$, (ii) $\omega=1\text{ rads}^{-1}$, (iii) $\omega=10\text{ rads}^{-1}$, (iv) $\omega=100\text{ rads}^{-1}$.

8.

The experimental responses of two systems are given below. Plot the Bode diagrams and identify the transfer functions.

(a)

ω (rads^{-1})	$\lvert G_1\rvert$ (dB)	$\angle G_1$ (°)
0.1	40	-92
0.2	34	-95
0.5	25	-100
1	20	-108
2	14	-126
3	10	-138
5	2	-160
10	-9	-190
20	-23	-220
30	-32	-235
40	-40	-243
50	-46	-248
100	-64	-258

(b)

ω (rads^{-1})	$\lvert G_2\rvert$ (dB)	$\angle G_2$ (°)
0.01	-26	87
0.02	-20	84
0.04	-14	79
0.07	-10	70
0.1	-7	61
0.2	-3	46
0.4	-1	29
0.7	-0.3	20
1	0	17
2	0	17
4	0	25
7	-2	36
10	-3	46
20	-7	64
40	-12	76
100	-20	84
500	-34	89
1000	-40	89

9.

Given $G(s)=K/s(s+5)$

(a) Plot the closed loop frequency response of this system using unity feedback when $K=1$. What is the –3 dB bandwidth of the system?

(b) Plot the closed loop frequency response when K is increased to $K=100$. What is the effect on the frequency response?

10.

The following measurements were taken for an open-loop system:

(i) $\omega=\omega_1$, $|G|=6\text{ dB}$, $\angle G=25°$

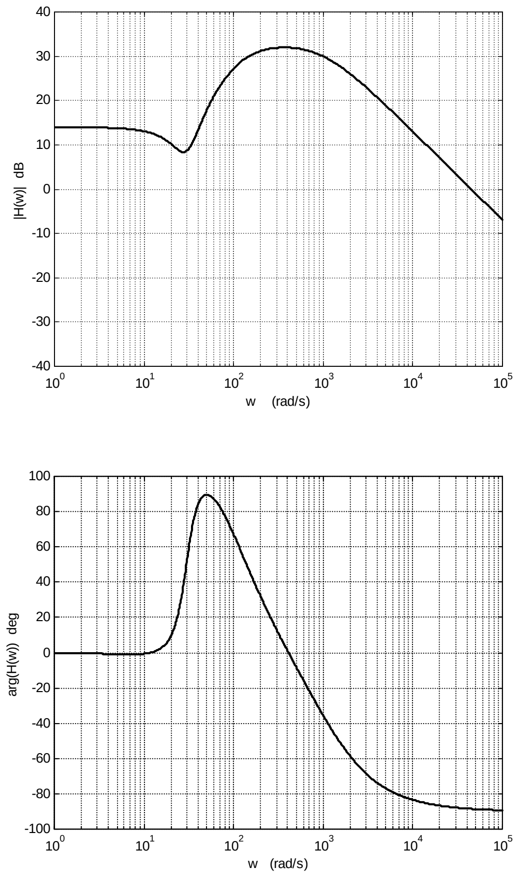
(ii) $\omega=\omega_2$, $|G|=-18\text{ dB}$, $\angle G=127°$

Find $|G|$ and $\angle G$ at ω_1 and ω_2 when the system is connected in a unity feedback arrangement.

4B.24

11.

An amplifier has the following frequency response. Find the transfer function.



Lecture 5A – Time-Domain Response

Steady-state error. Transient response. Second-order step response. Damping ratio. Damped frequency. Percent overshoot. Rise time. Delay time. Settling time.

Overview

Control systems employing feedback usually operate to bring the output of the system being controlled "in line" with the reference input. For example, a maze rover may receive a command to go forward 4 units – how does it respond? Can we control the dynamic behaviour of the rover, and if we can, what are the limits of the control? Obviously we cannot get a rover to move infinitely fast, so it will never follow a step input exactly. It must undergo a transient – just like an electric circuit with storage elements. However, with feedback, we may be able to change the transient response to suit particular requirements, like time taken to get to a certain position within a small tolerance, not "overshooting" the mark and hitting walls, etc.

Steady-State Error

One of the main objectives of control is for the system output to "follow" the system input. The difference between the input and output, in the steady-state, is termed the steady-state error:

Steady-state error defined

$$\boxed{\begin{aligned} e(t) &= r(t) - c(t) \\ e_{ss} &= \lim_{t \to \infty} e(t) \end{aligned}} \tag{5A.1}$$

Consider the unity-feedback system:

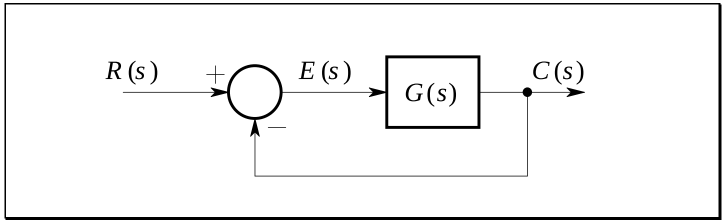


Figure 5A.1

System type defined

The *type* of the control system, or simply *system type*, is the number of poles that $G(s)$ has at $s=0$. For example:

$$G(s)=\frac{10(1+3s)}{s(s^2+2s+2)} \quad \text{type}\,1$$
$$G(s)=\frac{4}{s^3(s+2)} \quad \text{type}\,3 \tag{5A.2}$$

When the input to the control system in Figure 5A.1 is a step function with magnitude *R*, then $R(s)=R/s$ and the steady-state error is:

$$\begin{aligned} e_{ss} &= \lim_{t\to\infty} e(t) = \lim_{s\to 0} sE(s) \\ &= \lim_{s\to 0}\frac{sR(s)}{1+G(s)} = \lim_{s\to 0}\frac{R}{1+G(s)} \\ &= \frac{R}{1+\lim_{s\to 0} G(s)} \end{aligned} \tag{5A.3}$$

For convenience, we define the *step-error constant*, K_P, as:

Step-error constant – only defined for a step-input

$$\boxed{K_P = \lim_{s\to 0} G(s)} \tag{5A.4}$$

so that Eq. (5A.3) becomes:

$$\boxed{e_{ss} = \frac{R}{1+K_P}} \tag{5A.5}$$

We see that for the steady-state error to be zero, we require $K_P \to \infty$. This will only occur if there is *at least* one pole of $G(s)$ at the origin. We can summarise the errors of a unity-feedback system to a step-input as:

Steady-state error to a step-input for a unity-feedback system

$$\boxed{\begin{aligned} &\text{type}\,0\,\text{system}: && e_{ss}=\frac{R}{1+K_P}=\text{constant} \\ &\text{type}\,1\,\text{or higher system}: && e_{ss}=0 \end{aligned}} \tag{5A.6}$$

Transient Response

Consider a maze rover (MR) described by the differential equation:

Maze rover force / velocity differential equation

$$\frac{dv(t)}{dt}+\frac{k_f}{M}v(t)=\frac{1}{M}x(t) \tag{5A.7}$$

where $v(t)$ is the velocity, $x(t)$ is the driving force, *M* is the mass and k_f represents frictional losses. We may represent the MR by the following block diagram:

MR transfer function for velocity output

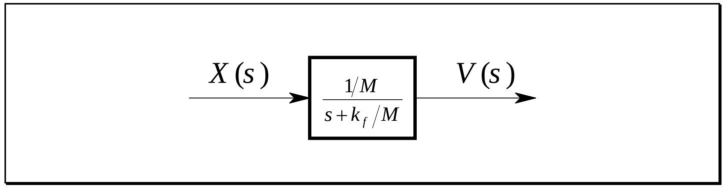


Figure 5A.2

Now, from the diagram above, it appears that our input to the rover affects the velocity in some way. But we need to control the output *position*, not the output *velocity*.

We're therefore actually interested in the following model of the MR:

MR transfer function for position output

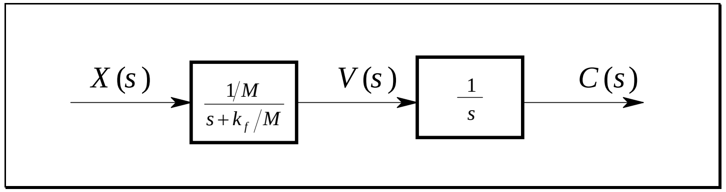


Figure 5A.3

This should be obvious, since position $c(t)$ is given by:

$$c(t) = \int_0^t v(\tau)d\tau \tag{5A.8}$$

Using block diagram reduction, our *position* model of the MR is:

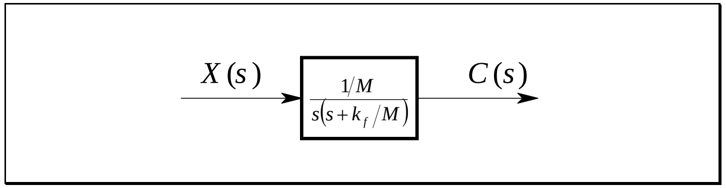


Figure 5A.4

The whole point of modelling the rover is so that we can control it. Suppose we wish to build a maze rover *position* control system. We will choose for simplicity a unity-feedback system, and place a "controller" in the feed-forward path in front of the MR's input. Such a control strategy is termed *series compensation*.

A block diagram of the proposed feedback system, with it's unity-feedback and series compensation controller is:

Simple MR position control scheme

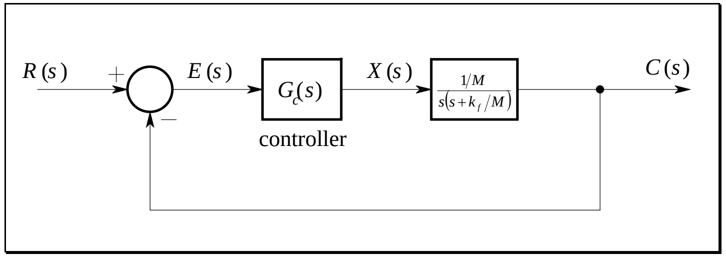


Figure 5A.5

Let the controller be the simplest type possible – a *proportional* controller which is just a gain K_P. Then the closed-loop transfer function is given by:

$$\frac{C(s)}{R(s)} = \frac{K_P/M}{s(s+k_f/M)+K_P/M} \tag{5A.9}$$

which can be manipulated into the standard form for a second-order transfer function:

Second-order control system

$$\boxed{\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}} \tag{5A.10}$$

Note: For a more complicated controller, we will not in general obtain a second-order transfer function. The reason we examine second-order systems is because they are amenable to analytical techniques – the concepts remain the same though for higher-order systems.

Controllers change the transfer function – and therefore the time-domain response

Although this simple controller can only vary ω_n^2, with $\zeta\omega_n$ fixed (*why?*), we can still see what sort of performance this controller has in the time-domain as K_P is varied. In fact, the goal of the controller design is to choose a suitable value of K_P to achieve certain criteria.

For a unit-step function input, $R(s)=1/s$, and the output response of the system is obtained by taking the inverse Laplace transform of the output transform:

$$C(s)=\frac{\omega_n^2}{s\left(s^2+2\zeta\omega_n s+\omega_n^2\right)} \tag{5A.11}$$

We have seen previously that the result for an underdamped system is:

$$c(t)=1-\frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\sin\left(\omega_n\sqrt{1-\zeta^2}\,t+\cos^{-1}\zeta\right) \tag{5A.12}$$

We normally desire the output to be slightly underdamped due to it's "fast" characteristics – it rises to near the steady-state output much quicker than an overdamped or critically damped system.

Although second-order control systems are rare in practice, their analysis generally helps to form a basis for the understanding of analysis and design for higher-order systems, especially ones that can be approximated by second-order systems. Also, time-domain specifications of systems can be directly related to an underdamped second-order response via simple formula.

Second-order Step Response

The time-domain unit-step response to the system described by Eq. (5A.10) has been solved previously (see Lecture 4A). There we found that there were three distinct solutions that depended upon the pole locations. We termed the responses overdamped (two real poles), critically damped (repeated real poles) and underdamped (complex conjugate poles).

Some time-domain criteria only apply to certain types of response. For example, percent overshoot only exists for the underdamped case.

We will now examine what sort of criteria we usually specify, with respect to the following diagram:

Step response definitions

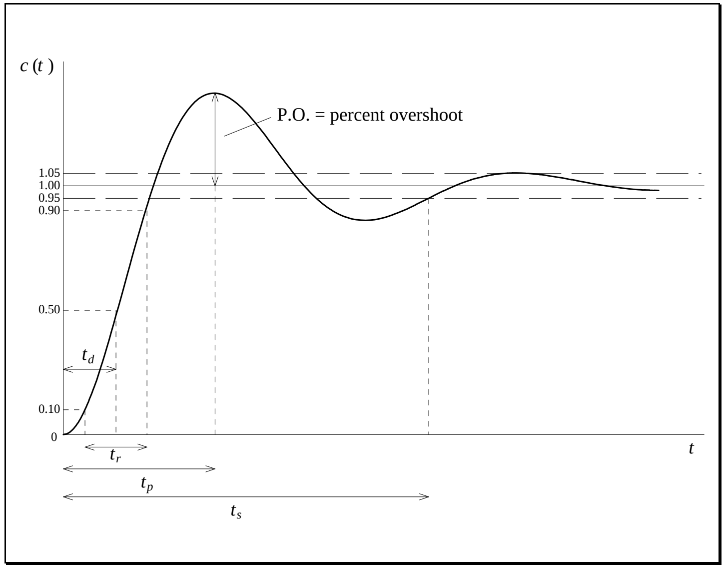


Figure 5A.6

The following definitions are made:

$$\text{P.O.} = \text{percent overshoot} = \frac{c_{\max}-c_{ss}}{c_{ss}}\times 100\% \tag{5A.13a}$$

Percent overshoot

$$\text{delay time} = t_d \qquad c(t_d)=c_{ss}\times 50\% \tag{5A.13b}$$

Delay time

$$\text{rise time} = t_r = t_{90\%}-t_{10\%} \qquad c(t_{90\%})=c_{ss}\times 90\%,\quad c(t_{10\%})=c_{ss}\times 10\% \tag{5A.13c}$$

Rise time

$$\text{settling time} = t_s \qquad \frac{c(t>t_s)-c_{ss}}{c_{ss}} <= 5\% \tag{5A.13d}$$

Settling time

The poles of Eq. (5A.10) are given by:

$$\begin{aligned} p_1,\, p_2 &= -\zeta\omega_n \pm j\omega_n\sqrt{1-\zeta^2} \\ &= -\alpha \pm j\omega_d \end{aligned} \tag{5A.14}$$

The natural frequency is:

Natural frequency defined

$$\boxed{\omega_n = \text{radial distance from origin}} \tag{5A.15}$$

The damping factor is defined as:

Damping factor defined

$$\boxed{\alpha = \text{real part of the poles}} \tag{5A.16}$$

and the damped frequency is:

Damped frequency defined

$$\boxed{\omega_d = \text{imaginary part of the poles}} \tag{5A.17}$$

The damping ratio is defined as:

Damping ratio defined

$$\boxed{\zeta = \frac{\alpha}{\omega_n}} \tag{5A.18}$$

All of the above are illustrated in the following diagram:

Second-order complex pole locations

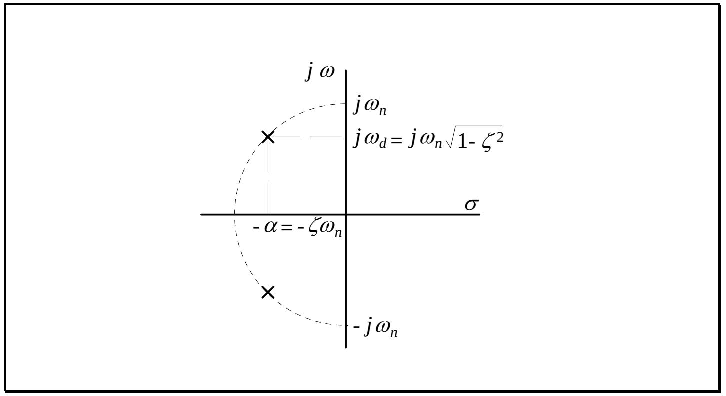


Figure 5A.7

The effect of ζ is readily apparent from the following graph of the step-input response:

Second-order step response for varying damping ratio

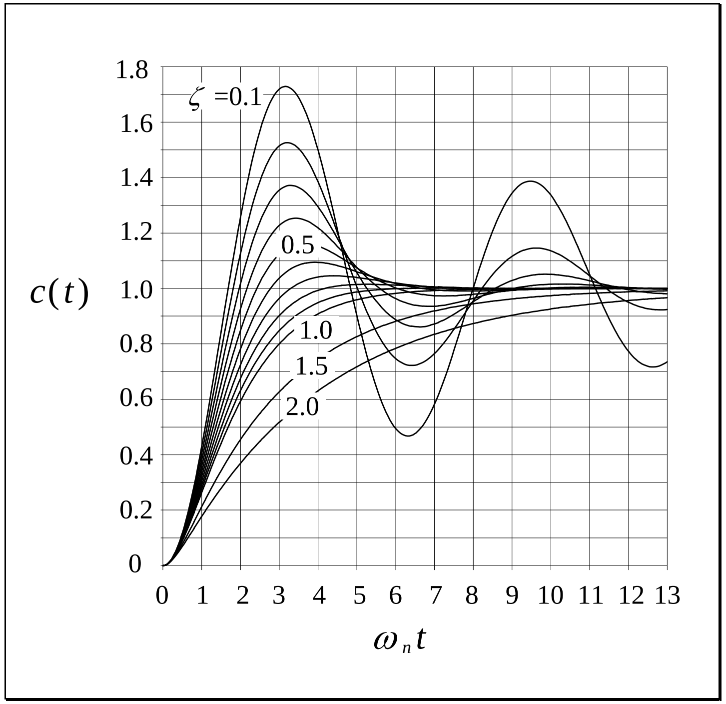


Figure 5A.8

Settling Time

The *settling time*, t_s, is the time required for the output to come *within and stay within* a given band about the actual steady-state value. This band is usually expressed as a percentage p of the steady-state value. To derive an estimate for the settling time, we need to examine the step-response more closely.

The standard, second-order lowpass transfer function of Eq. (5A.10) has the s-plane plot shown below if $0 < \zeta < 1$:

Pole locations showing definition of the angle θ

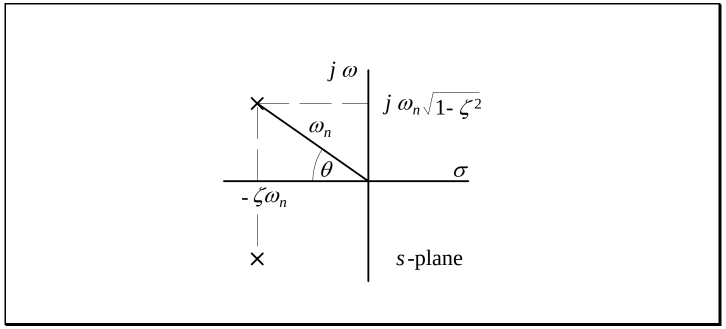


Figure 5A.9

It can be seen that the angle θ is given by:

$$\cos\theta = \frac{\zeta\omega_n}{\omega_n} = \zeta \tag{5A.19}$$

The step-response, as given by Eq. (5A.12), is then:

$$c(t) = 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\sin\left(\omega_n\sqrt{1-\zeta^2}\,t + \theta\right) \tag{5A.20}$$

The curves $1 \pm \left(e^{-\zeta\omega_n t}/\sqrt{1-\zeta^2}\right)$ are the envelope curves of the transient response for a unit-step input. The response curve $c(t)$ always remains within a pair of the envelope curves, as shown below:

Pair of envelope curves for the unit-step response of a lowpass second-order underdamped system

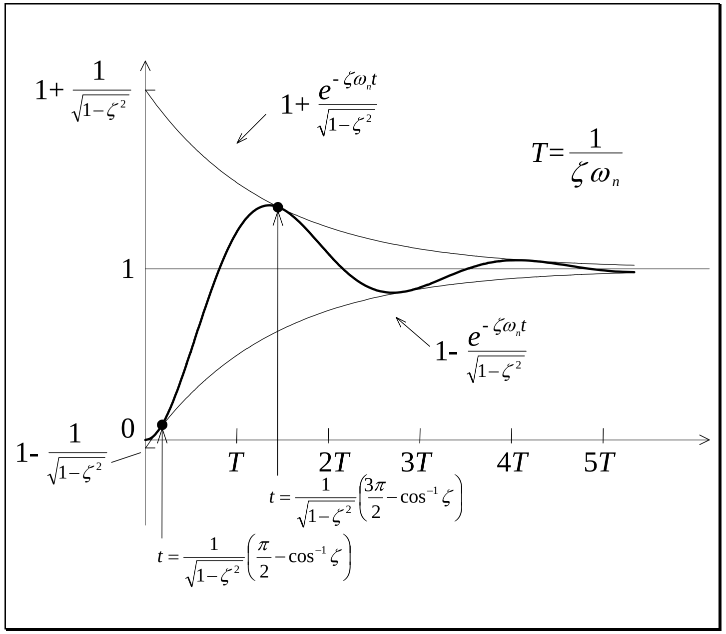


Figure 5A.10

To determine the settling time, we need to find the time it takes for the response to fall within and stay within a certain band about the steady-state value. This time depends on ζ and ω_n in a non-linear fashion, because of the oscillatory response. It can be obtained numerically from the responses shown in Figure 5A.8.

One way of analytically estimating the settling time with a simple equation is to consider only the minima and maxima of the step-response. For $0<\zeta<1$, the step-response is the damped sinusoid shown below:

Exponential curves intersecting the maxima and minima of the step-response

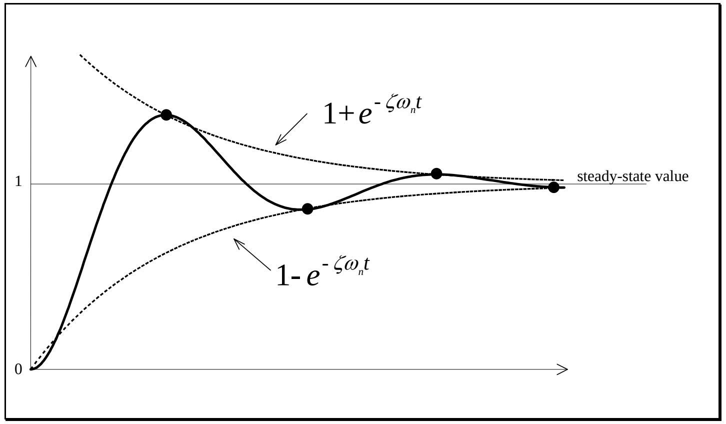


Figure 5A.11

The dashed curves in Figure 5A.11 represent the loci of maxima and minima of the step-response. The maxima and minima are found by differentiating the time response, Eq. (5A.20), and equating to zero:

$$\begin{aligned}\frac{dc(t)}{dt} &= \frac{-e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\cos\left(\omega_n\sqrt{1-\zeta^2}\,t+\theta\right)\omega_n\sqrt{1-\zeta^2} \\ &\quad -\frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\sin\left(\omega_n\sqrt{1-\zeta^2}\,t+\theta\right)\left(-\zeta\omega_n\right) \\ &= 0\end{aligned} \tag{5A.21}$$

Dividing through by the common term and rearranging, we get:

$$\begin{aligned}\sqrt{1-\zeta^2}\cos\left(\omega_n\sqrt{1-\zeta^2}\,t+\theta\right) &= \zeta\sin\left(\omega_n\sqrt{1-\zeta^2}\,t+\theta\right) \\ \frac{\sqrt{1-\zeta^2}}{\zeta} &= \tan\left(\omega_n\sqrt{1-\zeta^2}\,t+\theta\right) \\ \frac{\sqrt{1-\zeta^2}}{\zeta} &= \tan\left(\omega_n\sqrt{1-\zeta^2}\,t+\theta\pm n\pi\right)\end{aligned} \tag{5A.22}$$

where we have used the fact that $\tan\alpha=\tan(\alpha\pm n\pi)$, $n=0,1,2,\ldots$

Then, taking the arctangent of both sides of Eq. (5A.22), we have:

$$\tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}=\omega_n\sqrt{1-\zeta^2}\,t+\theta\pm n\pi \tag{5A.23}$$

From the *s*-plane plot of Figure 5A.9, we have:

$$\theta=\tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta} \tag{5A.24}$$

Substituting Eq. (5A.24) into Eq. (5A.23), and solving for t, we obtain:

Times at which the maxima and minima of the step-response occur

$$\boxed{t=\frac{\pm n\pi}{\omega_n\sqrt{1-\zeta^2}},\quad n=0,1,2,\ldots} \tag{5A.25}$$

Eq. (5A.25) gives the time at which the maxima and minima of the step-response occur. Since $c(t)$ in Eq. (5A.20) is only defined for $t\geq 0$, Eq. (5A.25) only gives valid results for $t\geq 0$.

Substituting Eq. (5A.25) into Eq. (5A.20), we get:

$$c[n] = 1 - \frac{1}{\sqrt{1-\zeta^2}} e^{\frac{-\zeta\omega_n n\pi}{\omega_n\sqrt{1-\zeta^2}}} \sin\left(\frac{\omega_n\sqrt{1-\zeta^2}\,n\pi}{\omega_n\sqrt{1-\zeta^2}} + \theta\right) \tag{5A.26}$$

Cancelling common terms, we have:

$$c[n] = 1 - \frac{1}{\sqrt{1-\zeta^2}} e^{\frac{-\zeta n\pi}{\sqrt{1-\zeta^2}}} \sin(n\pi + \theta) \tag{5A.27}$$

Since:

$$\begin{aligned} \sin(n\pi+\theta) &= -\sin\theta, \quad n \text{ odd} \\ \sin(n\pi+\theta) &= \sin\theta, \quad n \text{ even} \end{aligned} \tag{5A.28}$$

then Eq. (5A.27) defines two curves:

$$c_1[n] = 1 + \frac{1}{\sqrt{1-\zeta^2}} e^{\frac{-\zeta n\pi}{\sqrt{1-\zeta^2}}} \sin(\theta), \quad n \text{ odd} \tag{5A.29a}$$

$$c_2[n] = 1 - \frac{1}{\sqrt{1-\zeta^2}} e^{\frac{-\zeta n\pi}{\sqrt{1-\zeta^2}}} \sin(\theta), \quad n \text{ even} \tag{5A.29b}$$

From Figure 5A.9, we have:

$$\sin\theta = \sqrt{1-\zeta^2} \tag{5A.30}$$

Substituting this equation for $\sin\theta$ into Eqs. (5A.29), we get:

$$c_1[n] = 1 + e^{\frac{-\zeta n\pi}{\sqrt{1-\zeta^2}}}, \quad n \text{ odd} \tag{5A.31a}$$

$$c_2[n] = 1 - e^{\frac{-\zeta n\pi}{\sqrt{1-\zeta^2}}}, \quad n \text{ even} \tag{5A.31b}$$

Eq. (5A.31a) and Eq. (5A.31b) are, respectively, the relative maximum and minimum values of the step-response, with the times for the maxima and minima given by Eq. (5A.25). But these values will be exactly the same as those given by the following exponential curves:

$$c_1(t) = 1 + e^{-\zeta\omega_n t} \tag{5A.32a}$$

$$c_2(t) = 1 - e^{-\zeta\omega_n t} \tag{5A.32b}$$

Exponential curves bounding the maxima and minima of the step-response

evaluated at the times for the maxima and minima:

$$t = \frac{n\pi}{\omega_n\sqrt{1-\zeta^2}}, \; n = 0, 1, 2, \ldots \tag{5A.33}$$

Since the exponential curves, Eqs (5A.32), pass through the maxima and minima of the step-response, they can be used to approximate the extreme bounds of the step-response (note that the response actually goes slightly outside the exponential curves, especially after the first peak – the exponential curves are only an *estimate* of the bounds).

We can make an estimate of the settling time by simply determining the time at which $c_1(t)$ [or $c_2(t)$] enters the band $1-\delta < c(t) < 1+\delta$ about the steady-state value, as indicated graphically below:

Graph of an underdamped step-response showing exponential curves bounding the maxima and minima

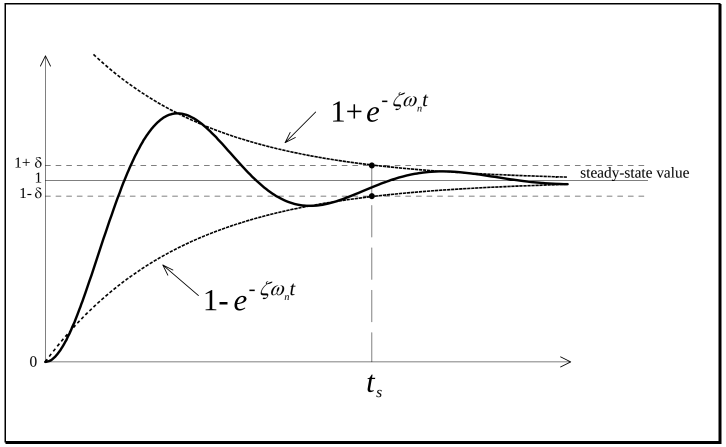


Figure 5A.12

The exponential terms in Eqs. (5A.32) represent the deviation from the steady-state value. Since the exponential response is monotonic, it is sufficient to calculate the time when the magnitude of the exponential is equal to the required error δ. This time is the settling time, t_s:

$$\delta = e^{-\zeta\omega_n t_s} \tag{5A.34}$$

Taking the natural logarithm of both sides and solving for t_s gives the "*p*-percent" settling time for a step-input:

Settling time for a second-order system

$$\boxed{t_s = -\frac{\ln\delta}{\zeta\omega_n}} \tag{5A.35}$$

where $\delta = \dfrac{p}{100}$.

Peak Time

The *peak time*, t_p, at which the response has a maximum overshoot is given by Eq. (5A.25), with $n=1$ (the first local maxima):

Peak time for a second-order system

$$\boxed{t_p = \frac{\pi}{\omega_n\sqrt{1-\zeta^2}} = \frac{\pi}{\omega_d}} \tag{5A.36}$$

This formula is only applicable if $0<\zeta<1$, otherwise the peak time is $t_p = \infty$.

Percent Overshoot

The magnitudes of the overshoots can be determined using Eq. (5A.31a). The maximum value is obtained by letting $n=1$. Therefore, the maximum value is:

$$c(t)\big|_{\max} = 1 + e^{\frac{-\zeta\pi}{\sqrt{1-\zeta^2}}} \tag{5A.37}$$

Hence, the maximum overshoot is:

$$\text{maximum overshoot} = e^{\frac{-\pi\zeta}{\sqrt{1-\zeta^2}}} \tag{5A.38}$$

and the maximum *percent overshoot* is:

Percent overshoot for a second-order system

$$\boxed{\text{P.O.} = e^{-\pi\zeta/\sqrt{1-\zeta^2}} \times 100\%} \tag{5A.39}$$

Rise Time and Delay Time

To determine rise times and delay times, we usually don't resort to solving the non-linear equation that results in substitution of the 10%, 50% and 90% values of the steady-state response and solving for *t*. We use a normalised delay time graph, or solve the resulting equations numerically using MATLAB®, or measure from a graph of the response (or on the DSO).

Summary

- The time-domain response of a system is important in control systems. In "set point" control systems, the time-domain response is the step response of the system. We usually employ feedback in a system so that the output tracks the input with some acceptable steady-state error.

- The transient part of a time-domain response is important. Control systems usually specify acceptable system behaviour with regards to percent overshoot, settling time, rise time, etc. For second-order systems, most of these quantities can be obtained from simple formula – in general they cannot.

- For second-order all-pole systems, we can directly relate pole locations (ζ and ω_n) to transient behaviour.

References

Kuo, B.: *Automatic Control Systems 7th ed.*, Prentice-Hall, 1995.

Exercises

1.

A second-order all-pole system has roots at $-2 \pm j3\,\text{rads}^{-1}$. If the input to the system is a step of 10 units, determine:

(a) the P.O. of the output

(b) the peak time of the output

(c) the damping ratio

(d) the natural frequency of the system

(e) the actual frequency of oscillation of the output

(f) the 0-100% rise time

(g) the 5% settling time

(h) the 2% settling time

2.

Determine a second-order, all-pole transfer function which will meet the following specifications for a step input:

(a) 10%-90% rise time ≤ 150 ms

(b) ≤ 5% overshoot

(c) 1% settling time ≤ 1 s

3.

Given:

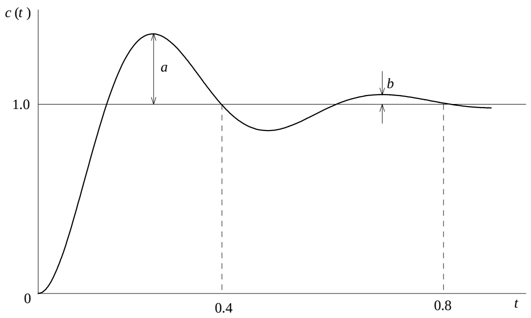
$$C(s)=\frac{\omega_n^2}{s\left(s^2+2\zeta\omega_n s+\omega_n^2\right)}$$

find $c(t)$ using the residue approach.

[Hint: Expand the denominator into the form $s(s+\alpha+j\omega_d)(s+\alpha-j\omega_d)$]

4.

The experimental zero-state response to a unit-step input of a second-order all-pole system is shown below:



(a) Derive an expression (in terms of *a* and *b*) for the damping ratio.

(b) Determine values for the natural frequency and damping ratio, given $a=0.4$ and $b=0.08$.

5.

Given:

(i) $G_1(s)=\dfrac{10}{s+10}$

(ii) $G_2(s)=\dfrac{1}{s+1}$

(iii) $G_3(s)=\dfrac{10}{(s+1)(s+10)}$

(a) Sketch the poles of each system on the *s*-plane.

(b) Sketch the time responses of each system to a unit-step input.

(c) Which pole dominates the time response of $G_3(s)$?

6.

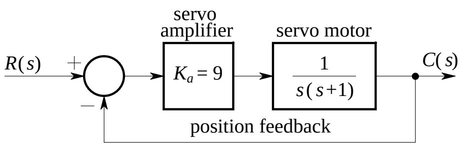
Find an approximate first-order model for the transfer function:

$$G(s)=\frac{4}{s^2+3s+2}$$

Sketch and compare the two time responses.

7.

Automatically controlled machine-tools form an important aspect of control system application. The major trend has been towards the use of automatic numerically controlled machine tools using direct digital inputs. Many CAD/CAM tools produce numeric output for the direct control of these tools, eliminating the tedium of repetitive operations required of human operators, and the possibility of human error. The figure below illustrates the block diagram of an automatic numerically controlled machine-tool position control system, using a computer to supply the reference signal.



(a) What is the undamped natural frequency ω_n and damping factor ζ?

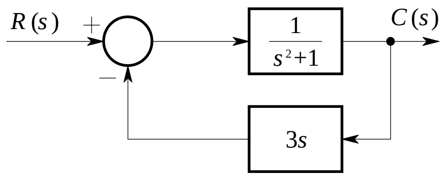
(b) What is the percent overshoot and time to peak resulting from the application of a unit-step input?

(c) What is the steady-state error resulting from the application of a unit-step input?

(d) What is the steady-state error resulting from the application of a unit-ramp $r(t) = tu(t)$ input?

8.

Find the steady-state errors to (a) a unit-step, and (b) a unit-ramp input, for the following feedback system:



Note that the error is defined as the difference between the actual input $r(t)$ and the actual output $c(t)$.

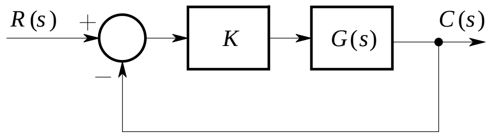
9.

Given

$$G(s) = \frac{0.1}{s + 0.1}$$

a) Find an expression for the step-response of this system. Sketch this response. What is the system time constant?

b) A unity feedback loop is to be connected around the system as shown:



Sketch the time responses of the closed-loop system and find the system time constants when (i) $K = 0.1$, (ii) $K = 1$ and (iii) $K = 10$.

What affect does feedback have on the time response of a *first-order* system?

Lecture 5B – Effects of Feedback

Modification of transient response. Closed-loop control. Disturbance rejection. System sensitivity.

Overview

We apply feedback in control systems for a variety of reasons. The primary purpose of feedback is to more accurately control the output - we wish to reduce the difference between a reference input and the actual output. When the input signal is a step, this is called *set-point* control.

Reduction of the system error is only one advantage of feedback. Feedback also affects the transient response, stability, bandwidth, disturbance rejection and sensitivity to system parameters.

Recall that the basic feedback system was described by the block diagram:

General feedback control system

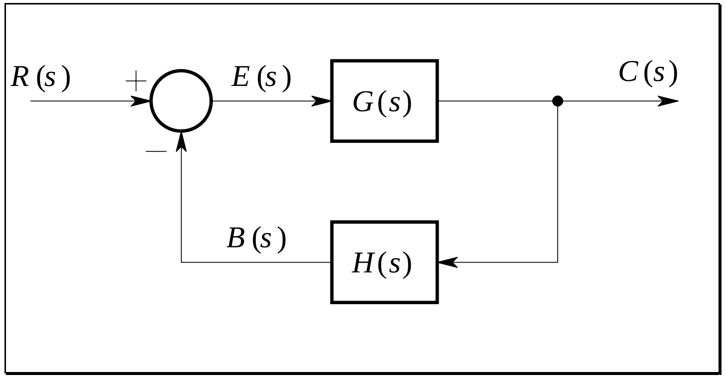


Figure 5B.1

The system is described by the following transfer function:

General feedback control system transfer function

$$\boxed{\frac{C(s)}{R(s)} = \frac{G(s)}{1+G(s)H(s)}} \tag{5B.1}$$

The only way we can improve "system performance" – whatever that may be – is by choosing a suitable $H(s)$ or $G(s)$. Some of the criteria for choosing $G(s)$, with $H(s)=1$, will be given in the following sections.

Transient Response

One of the most important characteristics of control systems is their transient response. We might desire a "speedy" response, or a response without an "overshoot" which may be physically impossible or cause damage to the system being controlled (eg. Maze rover hitting a wall!).

We can modify the response of a system by cascading the system with a transfer function which has been designed so that the overall transfer function achieves some design objective. This is termed *open-loop* control.

Open-loop control system

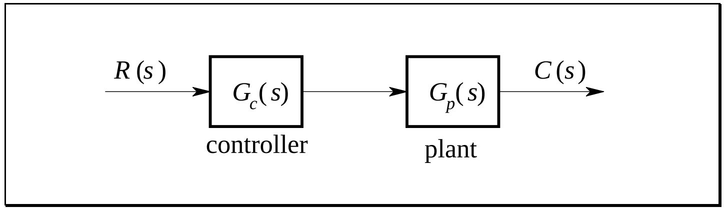


Figure 5B.2

A better way of modifying the response of a system is to apply feedback. This is termed *closed-loop* control. By adjusting the loop feedback parameters, we can control the transient response (within limits). A typical control system for set-point control simply derives the error signal by comparing the output directly with the input. Such a system is called a *unity-feedback* system.

Unity-feedback closed-loop control system

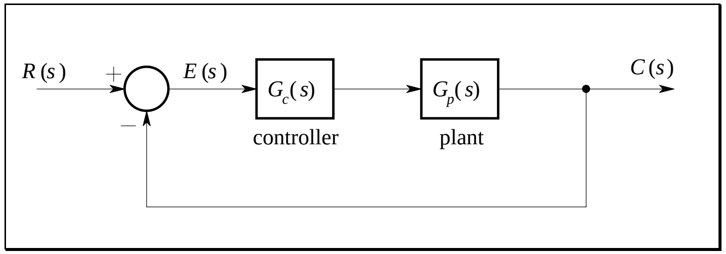


Figure 5B.3

Example

We have already seen that the MR can be described by the following block diagram (remember it's just a differential equation!):

MR transfer function for velocity output

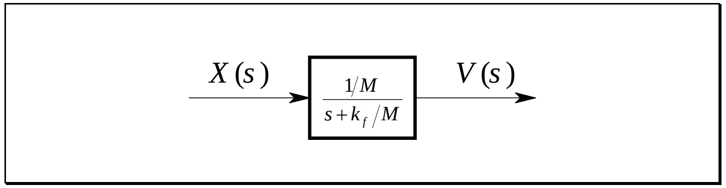


Figure 5B.4

In this example, the objective is to force $v(t)$ to be equal to a desired constant speed, v_0. As we have seen, this implies a partial fraction term v_0/s in the expression for $V(s)$. Since the MR's transfer function does not contain a K/s term, then the only way it can appear in the expression for $V(s)$ is if $X(s)=K/s$. (You might like to review Lecture 4A).

Assuming the input is a step function, then we have:

$$\begin{aligned} V(s) &= G(s)X(s) \\ &= \frac{1/M}{s+k_f/M}\frac{K}{s} \\ &= \frac{K/k_f}{s} - \frac{K/k_f}{s+k_f/M} \end{aligned} \tag{5B.2}$$

Inverse transforming yields:

$$v(t) = \frac{K}{k_f}\left[1-e^{-(k_f/M)t}\right], \quad t \geq 0 \tag{5B.3}$$

If K is set to $v_0 k_f$ then:

Step response of MR velocity

$$v(t) = v_0\left[1 - e^{-(k_f/M)t}\right], \quad t \geq 0 \tag{5B.4}$$

and since $k_f/M > 0$, then $v(t) \to v_0$ as $t \to \infty$. If our reference signal were $r(t) = v_0 u(t)$ then the system we have described is:

Simple open-loop MR controller

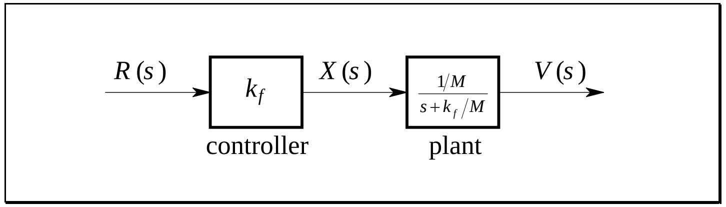


Figure 5B.5

This is referred to as *open-loop* control, since it depends only on the reference signal and not on the output. This type of control is deficient in several aspects. Note that the reference input has to be converted to the MR input through a gain stage equal to k_f, which must be known. Also, by examining Eq. (5B.4), we can see that we have no control over how fast the velocity converges to v_0.

Open-loop control has many disadvantages

Closed-Loop Control

To better control the output, we'll implement a *closed-loop* control system. For simplicity, we'll use a *unity-feedback* system, with our controller placed in the *feed-forward* path:

Closed-loop MR controller

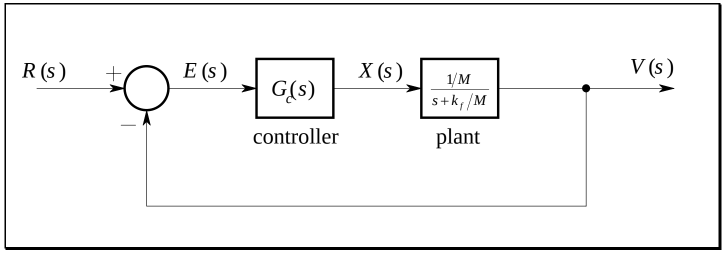


Figure 5B.6

Proportional Control (P Controller)

The simplest type of controller has transfer function $G_c(s) = K_P$. This is called *proportional* control since the control signal $x(t)$ is directly proportional to the error signal $e(t)$.

Proportional controller (P controller)

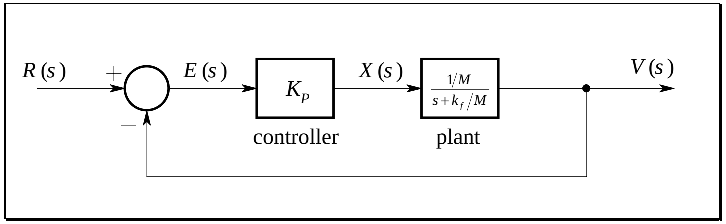


Figure 5B.7

With this type of control, the transform of the output is:

$$\begin{aligned} V(s) &= \frac{G_c(s)G(s)}{1+G_c(s)G(s)}R(s) \\ &= \frac{K_P/M}{(s+k_f/M+K_P/M)}\frac{v_0}{s} \\ &= \frac{K_P v_0/(k_f+K_P)}{s} - \frac{K_P v_0/(k_f+K_P)}{(s+k_f/M+K_P/M)} \end{aligned} \tag{5B.5}$$

Inverse transforming yields the response:

MR velocity step-response using a proportional controller

$$v(t) = \frac{K_P v_0}{k_f + K_P}\left[1 - e^{-[(k_f+K_P)/M]t}\right], \quad t \geq 0 \tag{5B.6}$$

Now the velocity converges to the value $K_P v_0/(k_f + K_P)$. Since it is now impossible for $v(t) = v_0$, the proportional controller will always result in a *steady-state tracking error* equal to $k_f v_0/(k_f + K_P)$. However, we are free to make this error as small as desired by choosing a suitably large value for K_P.

Sometimes a closed-loop system exhibits a steady-state error

Also, from Eq. (5B.6), we can see that the rate at which $v(t)$ converges to the steady-state value can be made as fast as desired by again taking K_P to be suitably large.

Integral Control (I Controller)

One of the deficiencies of the simple P controller in controlling the maze rover was that it did not have a zero steady-state error. This was due to the fact that the overall feedforward system was Type 0, instead of Type 1. We can easily make the overall feedforward system Type 1 by changing the controller so that it has a pole at the origin:

Integral controller (I controller)

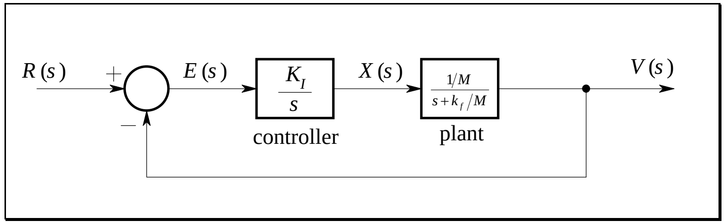


Figure 5B.8

We recognise that the controller transfer function is just an integrator, so this form of control is called *integral* control. With this type of control, the overall transfer function of the closed-loop system is:

$$\begin{aligned} T(s) &= \frac{K_I/sM}{\left(s+k_f/M+K_I/sM\right)} \\ &= \frac{K_I/M}{s^2+k_f/M\,s+K_I/M} \end{aligned} \tag{5B.7}$$

We can see straight away that the transfer function is 1 at DC (set $s=0$ in the transfer function). This means that the output will follow the input in the steady-state (zero steady-state error). By comparing this second-order transfer function with the standard form:

$$T(s) = \frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2} \tag{5B.8}$$

we can see that the controller is only able to adjust the natural frequency, or the distance of the poles from the origin, ω_n. This may be good enough, but we would prefer to be able to control the damping ratio ζ as well.

Proportional Plus Integral Control (PI Controller)

If we combine the two previous controllers we have what is known as a PI controller.

Proportional plus Integral controller (PI controller)

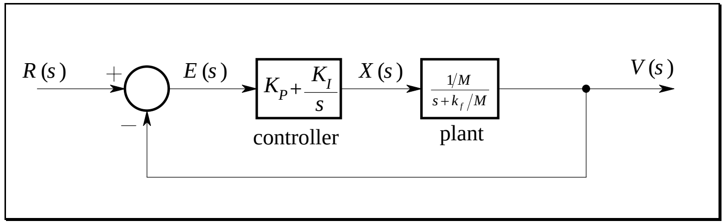


Figure 5B.9

The controller in this case causes the plant to respond to both the error and the integral of the error. With this type of control, the overall transfer function of the closed-loop system is:

$$\begin{aligned} T(s) &= \frac{\left(K_P+K_I/s\right)1/M}{s+k_f/M+\left(K_P+K_I/s\right)1/M} \\ &= \frac{K_I/M+K_P/M\,s}{s^2+\left(k_f/M+K_P/M\right)s+K_I/M} \end{aligned} \tag{5B.9}$$

Again, we can see straight away that the transfer function is 1 at DC (set $s=0$ in the transfer function), and again that the output will follow the input in the steady-state (zero steady-state error). We can now control the damping ratio ζ, as well as the natural frequency ω_n, *independently of each other*. But we also have a zero in the numerator. Intuitively we can conclude that the response will be similar to the response with integral control, but will also contain a term which is the derivative of this response (we see multiplication by s – the zero – as a derivative). We can analyse the response by rewriting Eq. (5B.9) as:

$$T(s) = \frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2} + \frac{K_P/K_I\,\omega_n^2 s}{s^2+2\zeta\omega_n s+\omega_n^2} \tag{5B.10}$$

For a unit-step input, let the output response that is due to the first term on the right-hand side of Eq. (5B.10) be $c_I(t)$. Then the total unit-step response is:

$$c(t) = c_I(t) + \frac{K_P}{K_I}\frac{dc_I(t)}{dt} \tag{5B.11}$$

The figure below shows that the addition of the zero at $s = -K_I/K_P$ reduces the rise time and increases the maximum overshoot, compared to the I controller step response:

PI controller response to a unit-step input

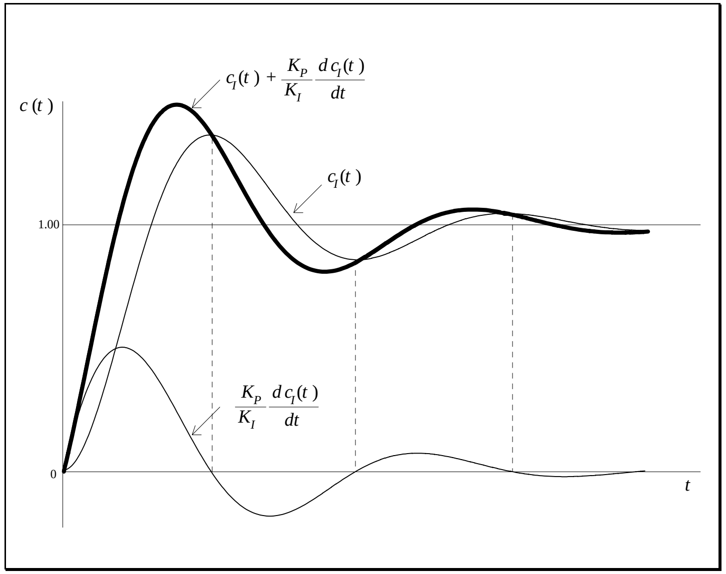


Figure 5B.10

The response is easy to sketch by drawing the derivative of $c_I(t)$ and adding a scaled version of this to $c_I(t)$. The derivative is sketched by noting that the derivative of $c_I(t)$ is zero when the tangent to $c_I(t)$ is horizontal, and the slope of $c_I(t)$ oscillates between positive and negative values between these points. The total response can then be sketched in, ensuring that the total response goes through the points where the $c_I(t)$ slope is zero.

Proportional, Integral, Derivative Control (PID Controller)

One of the best known controllers used in practice is the PID controller, where the letters stand for proportional, integral, and *derivative*. The addition of a derivative to the PI controller means that PID control contains *anticipatory* control. That is, by knowing the slope of the error, the controller can anticipate the direction of the error and use it to better control the process. The PID controller transfer function is:

$$G_c(s) = K_P + K_D s + \frac{K_I}{s} \tag{5B.12}$$

There are established procedures for designing control systems with PID controllers, in both the time and frequency-domains.

Disturbance Rejection

A major problem with open-loop control is that the output $c(t)$ of the plant will be "perturbed" by a disturbance input $d(t)$. Since the control signal $r(t)$ does not depend on the plant output $c(t)$ in open-loop control, the control signal cannot compensate for the disturbance $d(t)$. Closed-loop control can compensate to some degree for disturbance inputs.

Feedback minimizes the effect of disturbance inputs

Example

Consider the MR in open-loop:

Open-loop MR system with disturbance inputs

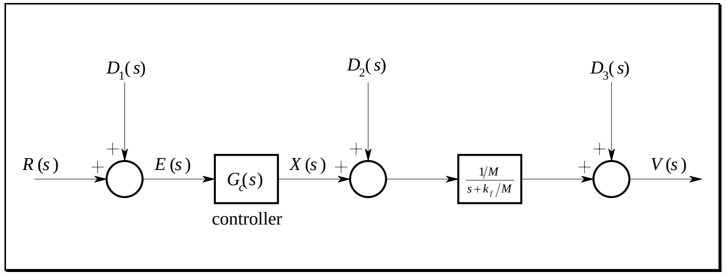


Figure 5B.11

The disturbance inputs could be modelling electronic noise in amplifiers, or a sudden increase in velocity due to an incline, or any other unwanted signal. In open-loop control, a disturbance $d_2(t)$ has just as much "control" as our controller!

Closed-loop feedback reduces the response of the system to disturbance inputs:

Closed-loop MR system with disturbance inputs

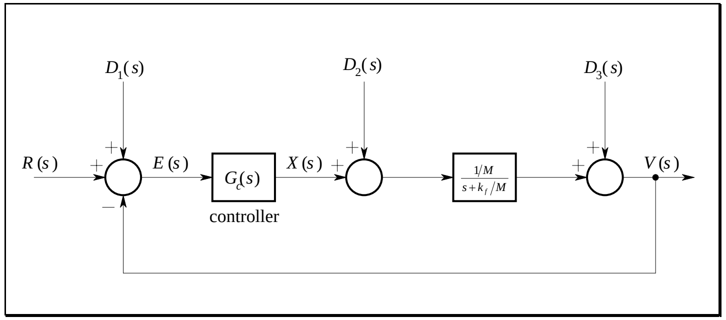


Figure 5B.12

Using superposition, the output is given by:

$$\begin{aligned} V(s) &= \frac{G_c(s)G(s)}{1+G_c(s)G(s)}\left[R(s)+D_1(s)\right] \\ &\quad + \frac{G(s)}{1+G_c(s)G(s)} D_2(s) \\ &\quad + \frac{1}{1+G_c(s)G(s)} D_3(s) \end{aligned} \tag{5B.13}$$

Most disturbance inputs are minimized when using feedback

Thus, we cannot eliminate noise at the input $d_1(t)$. The system cannot discriminate between $d_1(t)$ and $r(t)$. To minimise the other disturbances on the output, we need to make the *loop-gain* $1+G_c(s)G(s)$ large.

Sensitivity

Sensitivity is a measure of how the characteristics of a system depend on the variations of some component (or parameter) of the system. The effect of the parameter change can be expressed *quantitatively* in terms of a *sensitivity function*.

Sensitivity defines how one element of a system affects a characteristic of the system

System Sensitivity

This is a general formulation which applies to any type of system, open- or closed-loop. Let the system transfer function be expressed as $T(s,\alpha)$, where α is some parameter in the transfer function. Then:

System sensitivity defined

$$\boxed{S_\alpha^T = \frac{\partial T/T}{\partial \alpha/\alpha} = \frac{\alpha}{T}\frac{\partial T}{\partial \alpha}} \tag{5B.14}$$

is called the *system sensitivity* (with respect to α). It represents the fractional change in the system transfer function due to a fractional change in some parameter.

If $\left|S_\alpha^T\right|$ is small, the effect on T of changes in α is small. For *small changes* in α:

$$S_\alpha^T \approx \frac{\Delta T/T_0}{\Delta \alpha/\alpha_0} \tag{5B.15}$$

where T_0 and α_0 are the nominal or design values.

Example

How does $T(s)$ depend on changes in $G(s)$ for:

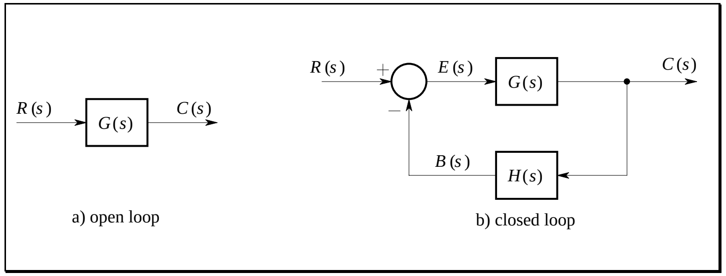


Figure 5B.13

Intuitively for open-loop, the output must depend directly on $G(s)$, so $S_G^T = 1$.

Intuitively for closed-loop:

$$T(s) = \frac{G(s)}{1+G(s)H(s)} \approx \frac{1}{H(s)} \tag{5B.16}$$

if $|G(s)H(s)| >> 1$. Therefore changes in $G(s)$ "don't matter", or the system is *not sensitive* to changes in $G(s)$.

Analytically, using system sensitivity for the open-loop case:

$$\begin{aligned} T(s) &= G(s) \\ \frac{\partial T}{\partial G} &= 1 \end{aligned} \tag{5B.17}$$

(here G corresponds to the parameter α). Therefore:

System sensitivity to *G* for open-loop systems

$$S_G^T = \frac{G}{T}\frac{\partial T}{\partial G} = 1 \tag{5B.18}$$

The transfer function is therefore directly sensitive to changes in *G*, as we thought.

Analytically, for the closed-loop case:

$$\begin{aligned} T &= \frac{G(s)}{1+G(s)H(s)} \\ \frac{\partial T}{\partial G} &= \frac{1+GH-GH}{(1+GH)^2} \end{aligned} \tag{5B.19}$$

Then:

System sensitivity to *G* for closed-loop systems

$$S_G^T = \frac{G}{G/(1+GH)} \frac{1}{(1+GH)^2} = \frac{1}{1+GH} \tag{5B.20}$$

Therefore with feedback, the effect of a percentage change in *G* is *reduced* by the factor $1/(1+GH)$. If the input *R* is held constant, the effect on the output *C* of a change in *G* is $1/(1+GH)$ less than it would have been without feedback.

Feedback elements must be accurate and stable

Show that $S_H^T \approx -1$ *for this example. This result means that* stable *feedback components must be used in order to receive the full benefits of feedback.*

Summary

- Various types of controllers, such as P, PI, PID are used to compensate the transfer function of the plant in unity-feedback systems.
- Feedback systems can minimise the effect of disturbance inputs or system parameter variations.
- Sensitivity is a measure of the dependence of a system's characteristics with respect to variations of a particular element (or parameter). Feedback can reduce the sensitivity of forward-path elements, but input and feedback elements must be highly stable because they have a much greater effect on the output.

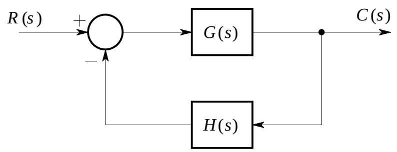
References

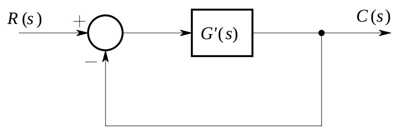
Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB®*, Prentice-Hall, 1997.

Exercises

1.

Show that the following block diagrams are equivalent:



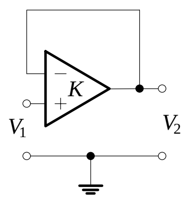


where $G'(s)=\dfrac{G(s)}{1+G(s)[H(s)-1]}$

2.

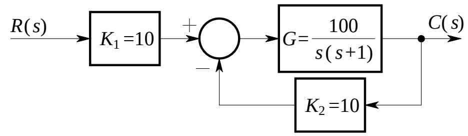
Assume that an operational amplifier has an infinite input impedance, zero output impedance and a very large gain *K*.

Show for the *feedback* configuration shown that $V_o/V_i = K/(K+1) \approx 1$ if K is large.



3.

For the system shown:



a) Determine the sensitivity of the system's transfer function T with respect to the input transducer, K_1.

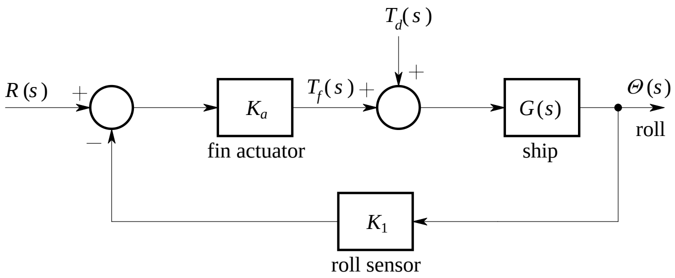
b) Determine the sensitivity of the system's transfer function T with respect to the output transducer, K_2.

c) Determine the sensitivity of the system's transfer function T with respect to the plant, G.

d) Indicate qualitatively the frequency dependency of S_G^T.

4.

It is important to ensure passenger comfort on ships by stabilizing the ship's oscillations due to waves. Most ship stabilization systems use fins or hydrofoils projecting into the water in order to generate a stabilization torque on the ship. A simple diagram of a ship stabilization system is shown below:



The rolling motion of a ship can be regarded as an oscillating pendulum with a deviation from the vertical of θ degrees and a typical period of 3 seconds. The transfer function of a typical ship is:

$$G(s)=\frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}$$

where $\omega_n = 2\pi/T = 2$, $T = 3.14\,\text{s}$, and $\zeta = 0.1$. With this low damping factor ζ, the oscillations continue for several cycles and the rolling amplitude can reach 18° for the expected amplitude of waves in a normal sea. Determine and compare the open-loop and closed-loop system for:

a) sensitivity to changes in the actuator constant K_a and the roll sensor K_1

b) the ability to reduce the effects of the disturbance of the waves.

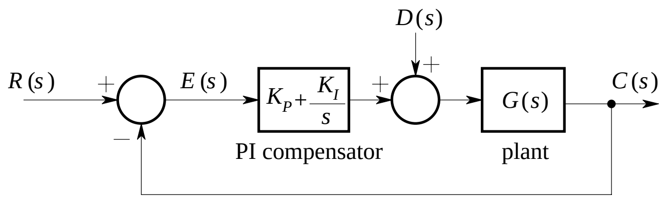
Note: 1. The desired roll angle $\theta(t)$ is zero degrees.

2. This regulating system is only effective for disturbances (waves) with frequencies $\omega_d \ll \omega_n$, the natural frequency of the ship. Can you show this?

Comment on these results with respect to the effect of feedback on the sensitivity of the system to parameter variations.

5.

The system shown uses a unity feedback loop and a PI compensator to control the plant.



Find the steady-state error, $e(\infty)$, for the following conditions [note that the error is always the difference between the reference input $r(t)$ and the plant output $c(t)$]:

(a) $G(s)=\dfrac{K_1}{1+sT_1}$, $K_I = 0$

(b) $G(s)=\dfrac{K_1}{1+sT_1}$, $K_I \neq 0$

(c) $G(s)=\dfrac{K_1}{s(1+sT_1)}$, $K_I = 0$

(d) $G(s)=\dfrac{K_1}{s(1+sT_1)}$, $K_I \neq 0$

when:

(i) $d(t)=0$, $r(t)=$unit-step

(ii) $d(t)=0$, $r(t)=$unit-ramp

(iii) $d(t)=$ unit-step, $r(t)=0$

(iv) $d(t)=$ unit-step, $r(t)=$unit-step

How does the addition of the integral term in the compensator affect the steady-state errors of the controlled system?

Lecture 6A – The *z*-Transform

z-transform. Properties of the z-transform. Evaluation of inverse z-transforms. System transfer function.

Overview

Digital control of continuous-time systems has become common thanks to the ever-increasing performance/price ratio of digital signal processors (microcontrollers, DSPs, gate arrays etc). Once we convert an analog signal to a sequence of numbers, we are free to do anything we like to them. Complicated control structures can be implemented easily in a computer, and even things which are impossible using analog components (such as median filtering). We can even create systems which "learn" or adapt to changing systems, something rather difficult to do with analog circuitry.

The side-effect of all these wonderful benefits is the fact that we have to learn a new (yet analogous) body of theory to handle what are essentially *discrete-time* systems. It will be seen that many of the techniques we use in the continuous-time domain can be applied to the discrete-time case, so in some cases we will design a system using continuous-time techniques and then simply discretize it.

To handle signals in the discrete-time domain, we'll need something akin to the Laplace transform in the continuous-time domain. That thing is the *z*-transform.

The *z*-Transform

We'll start with a discrete-time signal which is obtained by ideally and uniformly sampling a continuous-time signal:

We'll treat a discrete-time signal as the weights of an ideally sampled continuous-time signal

$$x_s(t) = \sum_{n=-\infty}^{\infty} x(t)\delta(t - nT_s) \tag{6A.1}$$

Now since the function $x_s(t)$ is zero everywhere except at the sampling instants, we replace the $x(t)$ with the *value* at the sample instant:

$$x_s(t) = \sum_{n=-\infty}^{\infty} x(nT_s)\delta(t - nT_s) \tag{6A.2}$$

Also, if we establish the time reference so that $x(t) = 0,\ \ t < 0$, then we can say:

$$x_s(t) = \sum_{n=0}^{\infty} x(nT_s)\delta(t - nT_s) \tag{6A.3}$$

This is the discrete-time signal in the time-domain. The *value* of each sample is represented by the *area* of an impulse, as shown below:

A discrete-time signal made by ideally sampling a continuous-time signal

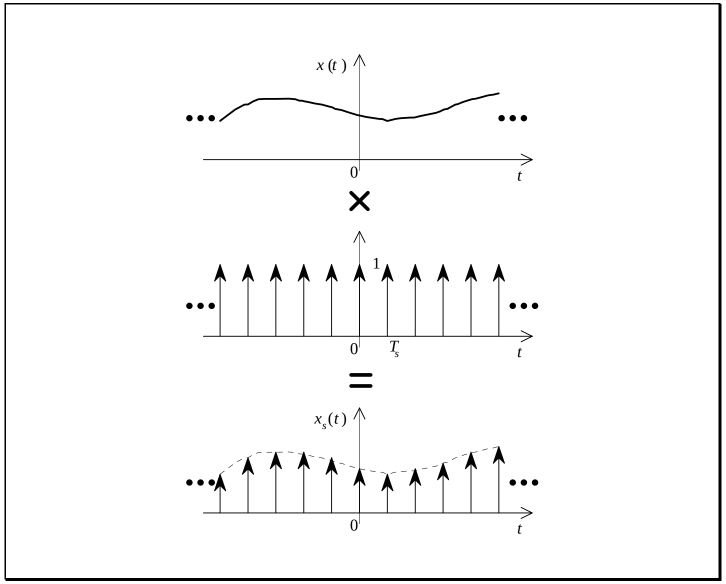


Figure 6A.1

What if we were to analyse this signal in the frequency domain? Taking the Laplace transform yields:

$$X_s(s) = \int_0^{\infty} \sum_{n=0}^{\infty} x(nT_s)\delta(t - nT_s)e^{-st}dt \tag{6A.4}$$

Since summation and integration are linear, we'll change the order to give:

$$X_s(s) = \sum_{n=0}^{\infty} x(nT_s)\int_0^{\infty} \delta(t - nT_s)e^{-st}dt \tag{6A.5}$$

Using the sifting property of the delta function, this integrates to:

$$X_s(s) = \sum_{n=0}^{\infty} x(nT_s)e^{-snT_s} \tag{6A.6}$$

Therefore, we can see that the Laplace transform of a sampled signal, $x_s(t)$, involves a summation of a series of functions, e^{-snT_s}. This is a problem because one of the advantages of transforming to the *s*-domain was to be able to work with algebraic polynomial equations. Fortunately, a very simple idea transforms the Laplace transform of a sampled signal into one which is an algebraic polynomial equation.

The idea is to define a complex variable *z* as:

$$\boxed{z = e^{sT_s}} \tag{6A.7}$$

Definition of *z*

Notice that this is a non-linear transformation. This definition gives us a new transform, called the *z*-transform, with independent variable *z*:

$$\boxed{X(z) = \sum_{n=0}^{\infty} x(nT_s)z^{-n}} \tag{6A.8}$$

Since $x(nT_s)$ is just a sequence of sample values, $x[n]$, we normally write the z-transform as:

Definition of z-transform

$$\boxed{X(z) = \sum_{n=0}^{\infty} x[n] z^{-n}} \tag{6A.9}$$

Thus, if we know the sample values of a signal, $x[n]$, it is a relatively easy step to write down the z-transform for the sampled signal:

$$X(z) = x[0] + x[1]z^{-1} + x[2]z^{-2} + \ldots \tag{6A.10}$$

Note that Eq. (6A.9) is the one-sided z-transform. We use this for the same reasons that we used the one-sided Laplace transform.

Mapping Between s-Domain and z-Domain

The mapping from the continuous-time s-plane into the discrete-time z-plane defined by $z = e^{sT_s}$ leads to some interesting observations. Since $s = \sigma + j\omega$, we have:

$$z = e^{\sigma T_s} e^{j\omega T_s} \tag{6A.11}$$

so that the magnitude and phase of z are, respectively:

$$|z| = e^{\sigma T_s} \tag{6A.12a}$$

$$\angle z = \omega T_s \tag{6A.12b}$$

Therefore, we can draw the mapping between the s-domain and the z-domain as follows:

The mapping between s-plane and z-plane

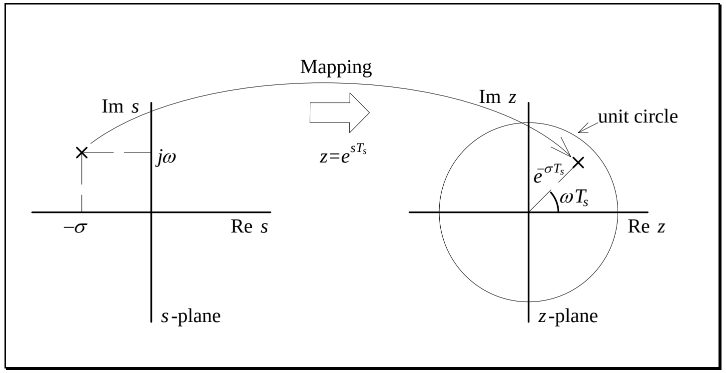


Figure 6A.2

The mapping of the s-domain to the z-domain depends on the sample interval, T_s. Therefore, the choice of sampling interval is crucial when designing a digital system.

Mapping the s-Plane Imaginary Axis

The $j\omega$-axis in the s-plane is when $\sigma = 0$. In the z-domain, this corresponds to a magnitude of $|z| = e^{\sigma T_s} = e^0 = 1$. Therefore, the frequency ω in the s-domain maps linearly on to the *unit-circle* in the z-domain with a phase angle $\angle z = \omega T_s$. In other words, distance along the $j\omega$-axis in the s-domain maps linearly onto angular displacement around the unit-circle in the z-domain.

The mapping between the $j\omega$-axis in the s-plane and the unit-circle in the z-plane

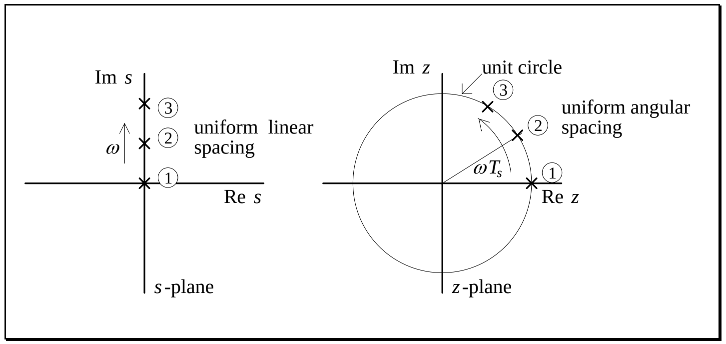


Figure 6A.3

Aliasing

With a sample period T_s, the angular sample rate is given by:

$$\omega_s = 2\pi f_s = \frac{2\pi}{T_s} \tag{6A.13}$$

When the frequency is equal to the foldover frequency (half the sample rate), $s = j\omega_s/2$ and:

$$z = e^{sT_s} = e^{j\frac{\omega_s}{2}T_s} = e^{j\pi} = -1 \tag{6A.14}$$

Thus, as we increase the frequency from 0 to half the sampling frequency along the $j\omega$-axis in the *s*-plane, there is a mapping to the *z*-domain in an anticlockwise direction around the unit-circle from $z = 1\angle 0$ to $z = 1\angle \pi$. In a similar manner, if we decrease the frequency along the $j\omega$-axis from 0 to $s = -j\omega_s/2$, the mapping to the *z*-plane is in a clockwise direction from $z = 1\angle 0$ to $z = 1\angle -\pi$. Thus, between the foldover frequencies $s = \pm j\omega_s/2$, there is a unique one-to-one mapping from the *s*-plane to the *z*-plane.

The mapping of the $j\omega$-axis in the *s*-plane up to the foldover frequency, and the unit-circle in the *z*-plane, is unique

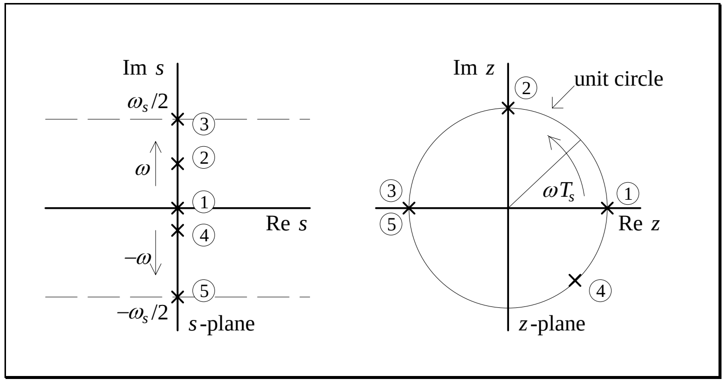


Figure 6A.4

However, if the frequency is increased beyond the foldover frequency, the mapping just continues to go around the unit-circle. This means that aliasing occurs – higher frequencies in the *s*-domain are mapped to lower frequencies in the *z*-domain.

The mapping of the $j\omega$-axis higher than the foldover frequency in the *s*-plane, and the unit-circle in the *z*-plane, causes aliasing

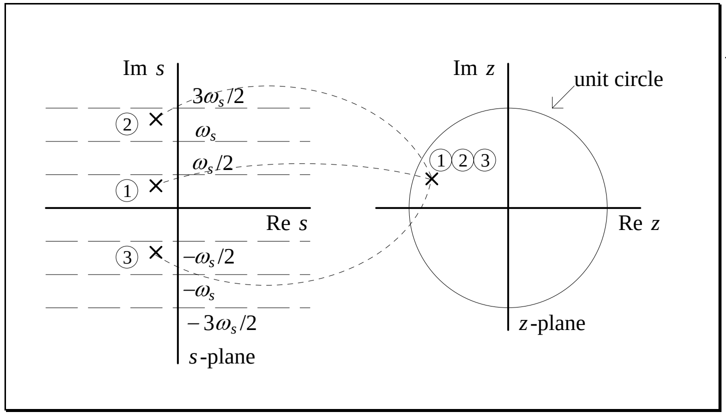


Figure 6A.5

Thus, absolute frequencies greater than the foldover frequency in the *s*-plane are mapped on to the same point as frequencies less than the foldover frequency in the *z*-plane. That is, they assume the alias of a lower frequency. The energies of frequencies higher than the foldover frequency add to the energy of frequencies less than the foldover frequency and this is referred to as frequency folding.

Finding z-Transforms

We will use the same strategy for finding *z*-transforms of a signal as we did for the other transforms – start with a known standard transform and successively apply transform properties. We first need a few standard transforms.

Example

To find the *z*-transform of a signal $x[n]=\delta[n]$, we substitute into the *z*- transform definition:

$$\begin{aligned} X(z) &= \sum_{n=0}^{\infty} \delta[n] z^{-n} \\ &= \delta[0] + \delta[1] z^{-1} + \delta[2] z^{-1} + \cdots \end{aligned} \tag{6A.15}$$

Since $\delta[0]=1$, and $\delta[n]=0$ for $n \neq 0$, we get a standard transform pair:

The *z*-transform of a unit-pulse

$$\boxed{\delta[n] \leftrightarrow 1} \tag{6A.16}$$

Thus, the *z*-transform of a unit-pulse is 1. Note that there is no such thing as an impulse for discrete-time systems – this transform pair is therefore similar to, but not identical to, the transform of an impulse for the Fourier and Laplace transforms.

Example

To find the *z*-transform of a signal $x[n]=a^n u[n]$, we substitute into the definition of the *z*-transform:

$$X(z) = \sum_{n=0}^{\infty} a^n u[n] z^{-n} \tag{6A.17}$$

Since $u[n]=1$ for all $n \geq 0$,

$$\begin{aligned} X(z) &= \sum_{n=0}^{\infty} \left(\frac{a}{z}\right)^n \\ &= 1 + \left(\frac{a}{z}\right) + \left(\frac{a}{z}\right)^2 + \left(\frac{a}{z}\right)^3 + \cdots \end{aligned} \tag{6A.18}$$

To convert this geometric progression into closed-form, let the sum of the first *k* terms of a general geometric progression be written as S_k:

$$S_k = \sum_{n=0}^{k-1} x^n = 1 + x + x^2 + \cdots + x^n + \cdots + x^{k-1} \tag{6A.19}$$

Then, multiplying both sides by x gives:

$$xS_k = x + x^2 + x^3 + \cdots + x^k \tag{6A.20}$$

Subtracting Eqs. (6A.19) and (6A.20) gives:

$$\begin{aligned} S_k(1-x) &= 1 + x + x^2 + x^3 + \cdots + x^{k-1} \\ &\quad - x - x^2 - x^3 - \cdots - x^k \end{aligned} \tag{6A.21}$$

This is a *telescoping sum*, where we can see that on the right-hand side only the first and last terms remain after performing the subtraction.

Dividing both sides by $(1-x)$ then results in:

$$S_k = \sum_{n=0}^{k-1} x^n = \frac{1-x^k}{1-x} \tag{6A.22}$$

If $k \to \infty$ then the series will only converge if $|x|<1$, because then $x^k \to 0$. For this special case, we have:

Closed-form expression for an infinite geometric progression

$$\boxed{\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \cdots = \frac{1}{1-x}, \quad |x|<1} \tag{6A.23}$$

Using this result to express Eq. (6A.18) results in:

$$X(z) = \frac{1}{1-\left(\dfrac{a}{z}\right)}, \quad \left|\frac{a}{z}\right| < 1 \tag{6A.24}$$

This can be rewritten as:

The z-transform of a geometric progression

$$\boxed{X(z) = \frac{z}{z-a}, \quad |z| > |a|} \tag{6A.25}$$

The ROC of $X(z)$ is $|z|>|a|$, as shown in the shaded area below:

A signal and the region of convergence of its z-transform in the z-plane

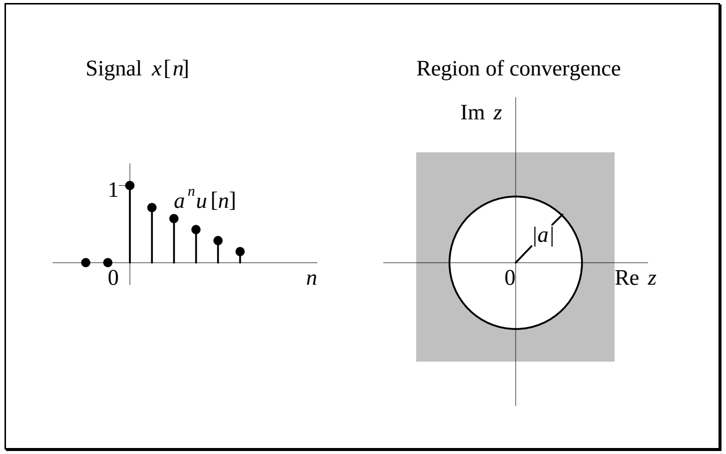


Figure 6A.6

As was the case with the Laplace transform, if we restrict the z-transform to causal signals, then we do not need to worry about the ROC.

Example

To find the z-transform of the unit-step, just substitute $a=1$ into Eq. (6A.25). The result is:

The z-transform of a unit-step

$$\boxed{u[n] \leftrightarrow \frac{z}{z-1}} \tag{6A.26}$$

This is a frequently used transform in the study of control systems.

Example

To find the z-transform of $\cos(\Omega n)u[n]$, we recognise that:

$$\cos(\Omega n) = \left(e^{j\Omega n} + e^{-j\Omega n}\right)/2 \tag{6A.27}$$

According to Eq. (6A.25), it follows that:

$$e^{\pm j\Omega n} \leftrightarrow \frac{z}{z - e^{\pm j\Omega}} \tag{6A.28}$$

Therefore:

$$\cos(\Omega n) \leftrightarrow \frac{1}{2}\left(\frac{z}{z - e^{+j\Omega}} + \frac{z}{z - e^{-j\Omega}}\right), \tag{6A.29}$$

and so we have another standard transform:

$$\boxed{\cos(\Omega n)u[n] \leftrightarrow \frac{z(z - \cos\Omega)}{z^2 - 2z\cos\Omega + 1}} \tag{6A.30}$$

A similar derivation can be used to find the z-transform of $\sin(\Omega n)u[n]$.

Most of the z-transform properties are inherited Laplace transform properties.

Right Shift (Delay) Property

One of the most important properties of the z-transform is the right shift property. It enables us to directly transform a *difference* equation into an *algebraic* equation in the complex variable z.

The z-transform of a function shifted to the right by one unit is given by:

$$Z\{x[n-1]\} = \sum_{n=0}^{\infty} x[n-1]z^{-n} \tag{6A.31}$$

Letting $r = n - 1$ yields:

$$\begin{aligned} Z\{x[n-1]\} &= \sum_{r=-1}^{\infty} x[r]z^{-(r+1)} \\ &= x[-1] + z^{-1}\sum_{r=0}^{\infty} x[r]z^{-r} \\ &= z^{-1}X(z) + x[-1] \end{aligned} \tag{6A.32}$$

Thus:

$$\boxed{x[n-1] \leftrightarrow z^{-1}X(z) + x[-1]} \tag{6A.33}$$

The z-transform right shift property

Standard *z*-Transforms

$$u[n] \leftrightarrow \frac{z}{z-1} \tag{Z.1}$$

$$\delta[n] \leftrightarrow 1 \tag{Z.2}$$

$$a^n u[n] \leftrightarrow \frac{z}{z-a} \tag{Z.3}$$

$$a^n(\cos\Omega n)u[n] \leftrightarrow \frac{z^2 - (a\cos\Omega)z}{z^2 - (2a\cos\Omega)z + a^2} \tag{Z.4}$$

$$a^n(\sin\Omega n)u[n] \leftrightarrow \frac{(a\sin\Omega)z}{z^2 - (2a\cos\Omega)z + a^2} \tag{Z.5}$$

z-Transform Properties

Assuming $x[n] \leftrightarrow X(z)$.

Linearity
$$ax[n] \leftrightarrow aX(z) \tag{Z.6}$$

Multiplication by a^n
$$a^n x[n] \leftrightarrow X\left(\frac{z}{a}\right) \tag{Z.7}$$

Right shifting
$$x[n-q] \leftrightarrow z^{-q}X(z) + \sum_{k=0}^{q-1} x[k-q]z^{-k} \tag{Z.8}$$

no corresponding transform (Z.9)

Multiplication by n
$$nx[n] \leftrightarrow -z\frac{d}{dz}X(z) \tag{Z.10}$$

Left shifting
$$x[n+q] \leftrightarrow z^{q}X(z) - \sum_{k=0}^{q-1} x[k]z^{q-k} \tag{Z.11}$$

Summation
$$\sum_{k=0}^{n} x[k] \leftrightarrow \frac{z}{z-1}X(z) \tag{Z.12}$$

Convolution
$$x_1[n] * x_2[n] \leftrightarrow X_1(z)X_2(z) \tag{Z.13}$$

Initial-value theorem
$$x[0] = \lim_{z\to\infty} X(z) \tag{Z.14}$$

Final-value theorem
$$\lim_{n\to\infty} x[n] = \lim_{z\to 1}\left[(z-1)X(z)\right] \tag{Z.15}$$

Evaluation of Inverse z-Transforms

From complex variable theory, the definition of the inverse z-transform is:

$$\boxed{x[n] = \frac{1}{2\pi j}\oint_C X(z)z^{n-1}dz} \tag{6A.34}$$

Inverse z-transform defined – but too hard to apply!

You won't need to evaluate this integral to determine the inverse z-transform, just like we hardly ever use the definition of the inverse Laplace transform. We manipulate $X(z)$ into a form where we can simply identify sums of standard transforms that may have had a few properties applied to them.

Given:

$$F(z) = \frac{b_0 z^n + b_1 z^{n-1} + \cdots + b_n}{(z-p_1)(z-p_2)\cdots(z-p_n)} \tag{6A.35}$$

we want to put $F(z)$ in the form:

$$F(z) = k_0 + k_1\frac{z}{z-p_1} + k_2\frac{z}{z-p_2} + \cdots + k_n\frac{z}{z-p_n} \tag{6A.36}$$

Expand functions of z into partial fractions – then find the inverse z-transform

The approach we will use is to firstly expand $F(z)/z$ into partial fractions, then multiply through by z.

Example

$$F(z) = \frac{z^3 + 2z^2 + z + 1}{(z-2)^2(z+3)}$$

Therefore:

$$\begin{aligned}\frac{F(z)}{z} &= \frac{z^3 + 2z^2 + z + 1}{z(z-2)^2(z+3)} \\ &= \frac{k_0}{z} + \frac{k_1}{z-2} + \frac{k_2}{(z-2)^2} + \frac{k_3}{z+3}\end{aligned}$$

Now we evaluate the residues:

$$k_0 = z\frac{F(z)}{z}\bigg|_{z=0} = \frac{z^3+2z^2+z+1}{(z-2)^2(z+3)}\bigg|_{z=0} = \frac{1}{12}$$

$$k_1 = \frac{d}{dz}(z-2)^2\frac{F(z)}{z}\bigg|_{z=2} = \frac{d}{dz}\left[\frac{z^3+2z^2+z+1}{z(z+3)}\right]\bigg|_{z=2}$$
$$= \frac{z(z+3)(3z^2+4z+1)-(z^3+2z^2+z+1)(2z+3)}{z^2(z+3)^2}\bigg|_{z=2} = \frac{77}{100}$$

$$k_2 = (z-2)^2\frac{F(z)}{z}\bigg|_{z=2} = \frac{z^3+2z^2+z+1}{z(z+3)}\bigg|_{z=2} = \frac{19}{10}$$

$$k_3 = (z+3)\frac{F(z)}{z}\bigg|_{z=-3} = \frac{z^3+2z^2+z+1}{z(z-2)^2}\bigg|_{z=-3} = \frac{11}{75}$$

Therefore:

$$F(z) = \frac{1}{12} + \frac{77}{100}\frac{z}{z-2} + \frac{19}{10}\frac{z}{(z-2)^2} + \frac{11}{75}\frac{z}{z+3}$$

From our standard transforms:

$$f[n] = \frac{1}{12}\delta[n] + \frac{77}{100}2^n + \frac{19}{10}\left(\frac{n}{2}\right)2^n + \frac{11}{75}(-3)^n, \quad n \ge 0$$

Note: For $k_2\dfrac{z}{(z-p_2)^2}$ use $na^n \leftrightarrow \dfrac{za}{(z-a)^2}$ (Standard transform Z.3 and property Z.10). Then with linearity (property Z.6) we have: $\dfrac{n}{a}a^n \leftrightarrow \dfrac{z}{(z-a)^2}$.

Transforms of Difference Equations

The right shift property of the z-transform sets the stage for solving linear difference equations with constant coefficients. Because $y[n-k] \leftrightarrow z^{-k}Y(z)$, the z-transform of a difference equation is an algebraic equation that can be readily solved for $Y(z)$. Next we take the inverse z-transform of $Y(z)$ to find the desired solution $y[n]$.

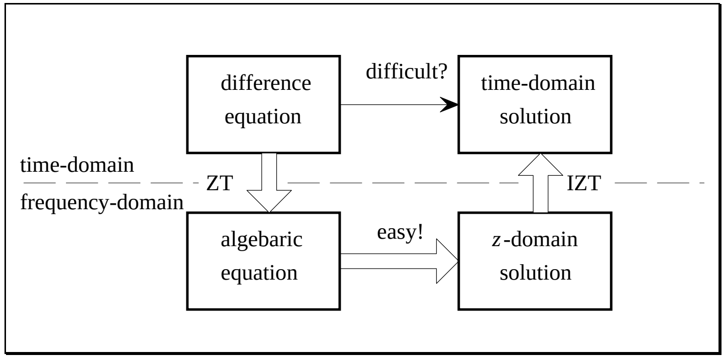


Figure 6A.7

Example

Solve the second-order linear difference equation:

$$y[n] - 5y[n-1] + 6y[n-2] = 3x[n-1] + 5x[n-2] \tag{6A.37}$$

if the initial conditions are $y[-1] = 11/6$, $y[-2] = 37/36$ and the input $x[n] = (2)^{-n}u[n]$.

Now:

$$y[n]u[n] \leftrightarrow Y(z)$$
$$y[n-1]u[n] \leftrightarrow z^{-1}Y(z) + y[-1] = z^{-1}Y(z) + \frac{11}{6}$$
$$y[n-2]u[n] \leftrightarrow z^{-2}Y(z) + z^{-1}y[-1] + y[-2] = z^{-2}Y(z) + \frac{11}{6}z^{-1} + \frac{37}{36} \tag{6A.38}$$

For the input, $x[-1]=x[-2]=0$. Then:

$$x[n]=(2)^{-n}u[n]=(0.5)^n u[n]\leftrightarrow \frac{z}{z-0.5}$$
$$x[n-1]u[n]\leftrightarrow z^{-1}X(z)+x[-1]=z^{-1}\frac{z}{z-0.5}=\frac{1}{z-0.5}$$
$$x[n-2]u[n]\leftrightarrow z^{-2}X(z)+z^{-1}x[-1]+x[-2]=z^{-2}X(z)=\frac{1}{z(z-0.5)} \tag{6A.39}$$

Taking the z-transform of Eq. (6A.37) and substituting the foregoing results, we obtain:

$$\begin{aligned}&Y(z)-5\left[z^{-1}Y(z)+\frac{11}{6}\right]+6\left[z^{-2}Y(z)+z^{-1}\frac{11}{6}+\frac{37}{36}\right]\\&=\frac{3}{z-0.5}+\frac{5}{z(z-0.5)}\end{aligned} \tag{6A.40}$$

or:

$$\left(1-5z^{-1}+6z^{-2}\right)Y(z)-\left(3-11z^{-1}\right)=\frac{3z+5}{z(z-0.5)} \tag{6A.41}$$

from which we obtain:

$$Y(z)=\underbrace{\frac{z(3z-11)}{(z^2-5z+6)}}_{\text{zero-input component}}+\underbrace{\frac{z(3z+5)}{(z-0.5)(z^2-5z+6)}}_{\text{zero-state component}} \tag{6A.42}$$

and:

$$\begin{aligned}\frac{Y(z)}{z}&=\frac{3z-11}{(z-2)(z-3)}+\frac{(3z+5)}{(z-0.5)(z-2)(z-3)}\\&=\underbrace{\frac{5}{z-2}-\frac{2}{z-3}}_{\text{zero-input component}}+\underbrace{\frac{26/15}{z-0.5}-\frac{22/3}{z-2}+\frac{28/5}{z-3}}_{\text{zero-state component}}\end{aligned} \tag{6A.43}$$

Therefore:

$$Y(z)=\underbrace{5\frac{z}{z-2}-2\frac{z}{z-3}}_{\text{zero-input component}}+\underbrace{\frac{26}{15}\frac{z}{z-0.5}-\frac{22}{3}\frac{z}{z-2}+\frac{28}{5}\frac{z}{z-3}}_{\text{zero-state component}} \tag{6A.44}$$

and:

$$\begin{aligned}y[n]&=\underbrace{5(2)^n-2(3)^n}_{\text{zero-input response}}+\underbrace{\tfrac{26}{15}(0.5)^n-\tfrac{22}{3}(2)^n+\tfrac{28}{5}(3)^n}_{\text{zero-state response}}\\&=\tfrac{26}{15}(0.5)^n-\tfrac{7}{3}(2)^n+\tfrac{18}{5}(3)^n,\quad n\geq 0\end{aligned} \tag{6A.45}$$

As can be seen, the z-transform method gives the total response, which includes zero-input and zero-state components. The initial condition terms give rise to the zero-input response. The zero-state response terms are exclusively due to the input.

System Transfer Function

Consider the simple first-order discrete-time system described by the difference equation:

$$y[n]+ay[n-1]=bx[n] \tag{6A.46}$$

First-order difference equation

Taking the z-transform of both sides and using the right-shift property gives:

$$Y(z)+a\left(z^{-1}Y(z)+y[-1]\right)=bX(z) \tag{6A.47}$$

Solving for $Y(z)$ gives:

$$Y(z)=\frac{-ay[-1]}{1+az^{-1}}+\frac{b}{1+az^{-1}}X(z) \tag{6A.48}$$

and corresponding z-transform

which can be written:

$$Y(z)=\frac{-ay[-1]z}{z+a}+\frac{bz}{z+a}X(z) \tag{6A.49}$$

The first part of the response results from the initial conditions, the second part results from the input.

If the system has no initial energy (zero initial conditions) then:

$$Y(z)=\frac{bz}{z+a}X(z) \tag{6A.50}$$

We now define the *transfer function* for this system as:

$$H(z)=\frac{bz}{z+a} \tag{6A.51}$$

so that:

Discrete-time transfer function defined

$$\boxed{Y(z)=H(z)X(z)} \tag{6A.52}$$

This is the *transfer function representation* of the system. To determine the output $y[n]$ we simply evaluate Eq. (6A.52) and take the inverse *z*-transform.

For a general n^{th} order system described by the difference equation:

$$y[n]+\sum_{i=1}^{N}a_i y[n-i]=\sum_{i=0}^{M}b_i x[n-i] \tag{6A.53}$$

and if the system has zero initial conditions, then taking the *z*-transform of both sides results in:

$$Y(z)=\frac{b_0 z^N+b_1 z^{N-1}+\cdots+b_M z^{N-M}}{z^N+a_1 z^{N-1}+\cdots+a_{N-1}z+a_N}X(z) \tag{6A.54}$$

so that the transfer function is:

Transfer function derived directly from difference equation

$$\boxed{H(z)=\frac{b_0 z^N+b_1 z^{N-1}+\cdots+b_M z^{N-M}}{z^N+a_1 z^{N-1}+\cdots+a_{N-1}z+a_N}} \tag{6A.55}$$

We can show that the convolution relationship of a linear discrete-time system:

$$y[n]=h[n]*x[n]=\sum_{i=0}^{\infty}h[i]x[n-i], \quad \text{n}\geq 0 \tag{6A.56}$$

when transformed gives Eq. (6A.52). We therefore have:

$$\boxed{h[n]\leftrightarrow H(z)} \tag{6A.57}$$

The unit-pulse response and transfer function form a *z*-transform pair

That is, the unit-pulse response and the transfer function form a *z*-transform pair.

Stability

The left-half *s*-plane, where $\sigma<0$, corresponds to $|z|=e^{\sigma T_s}<1$, which is *inside* the unit-circle. The right-half *s*-plane maps outside the unit circle.

Recall from Lecture 4A that functions of the Laplace variable *s* having poles with negative real parts decay to zero as $t\to\infty$. In a similar manner, transfer functions of *z* having poles with magnitudes less than one decay to zero in the time-domain as $n\to\infty$. Therefore, for a stable system, we must have:

$$\boxed{|p_i|<1} \tag{6A.58a}$$

where p_i are the poles of $H(z)$. This is equivalent to saying:

A system is stable if all the poles of the transfer function lie inside the unit-circle

(6A.58b)

Stability defined for a discrete-time system

Transfer Function Interconnections

The transfer function of an LTI discrete-time system can be computed from a block diagram of the system, just like for continuous-time systems.

Recall that an LTI discrete-time system is composed of elements such as adders, gains (which multiply the input by a constant), and the *unit-delay* element, which is shown below:

A discrete-time unit-delay element

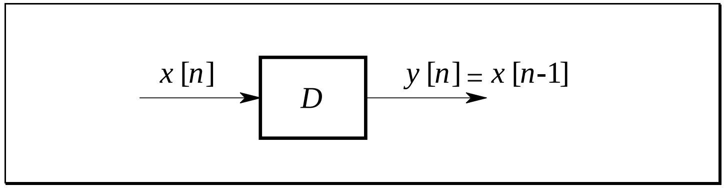


Figure 6A.8

By taking the z-transform of the input and output, we can see that we should represent a delay in the z-domain by:

Delay element in block diagram form in the *z*-domain

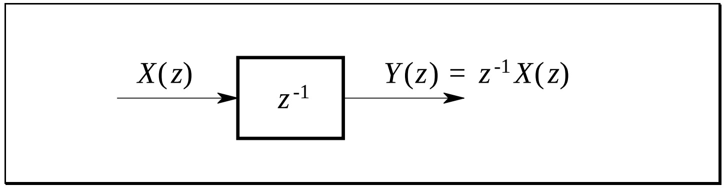


Figure 6A.9

Example

We saw in Lecture 1B that the discrete-time approximation to continuous-time integration was given by the difference equation:

$$y[n] = y[n-1] + Tx[n-1] \tag{6A.59}$$

The block diagram of the system is:

Time-domain block diagram of a numeric integrator

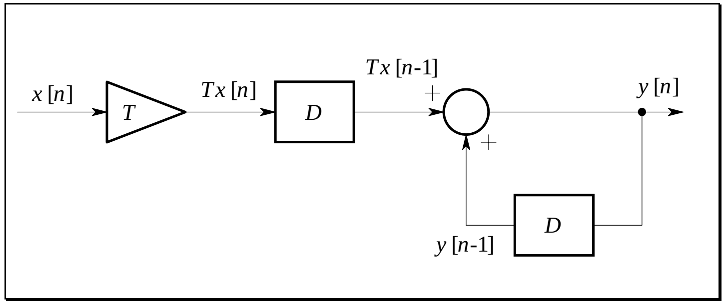


Figure 6A.10

The system in the z-domain is:

z-domain block diagram of a numeric integrator

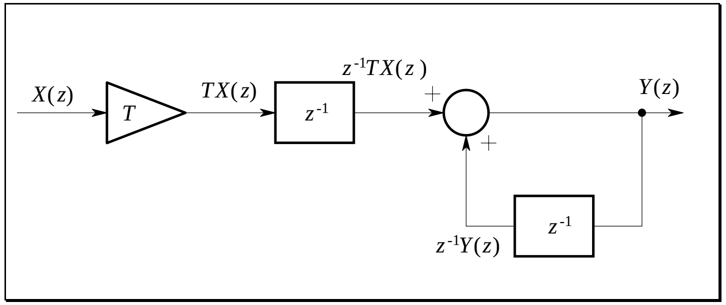


Figure 6A.11

Through the standard block diagram reduction techniques, we get:

Transfer function of a numeric integrator

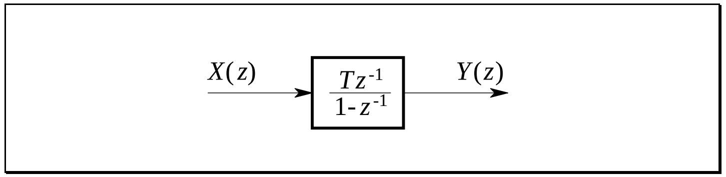


Figure 6A.12

You should confirm that this transfer function obtained using block diagram reduction methods is the same as that found by taking the z-transform of Eq. (6A.59).

Summary

- The z-transform is the discrete-time counterpart of the Laplace transform. There is a mapping from the s-domain to the z-domain given by $z = e^{sT_s}$. The mapping is not unique, and for frequencies above the foldover frequency, aliasing occurs.

- We evaluate inverse z-transforms using partial fractions, standard transforms and the z-transform properties.

- Systems described by difference equations have rational z-transforms. The z-transforms of the input signal and output signal are related by the *transfer function* of the system: $Y(z) = H(z)X(z)$. There is a one-to-one correspondence between the coefficients in the difference equation and the coefficients in the transfer function.

- We can use the z-transform to express discrete-time systems in transfer function (block diagram) form.

- The unit-pulse response and the transfer function form a z-transform pair: $h[n] \leftrightarrow H(z)$.

- The transfer function of a system can be obtained by performing analysis in the z-domain.

References

Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB®*, Prentice-Hall, 1997.

Exercises

1.

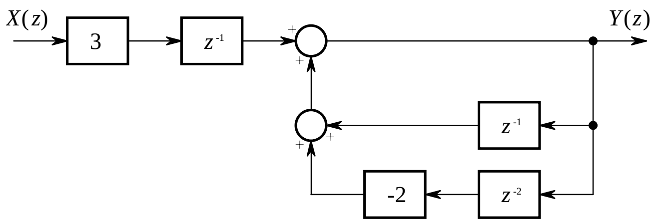
Construct z-domain block diagrams for the following difference equations:

(i) $y[n] = y[n-2] + x[n] + x[n-1]$

(ii) $y[n] = 2y[n-1] - y[n-2] + 3x[n-4]$

2.

(i) Construct a difference equation from the following block diagram:



(ii) From your solution calculate $y[n]$ for n = 0, 1, 2 and 3 given $y[-2] = -2$, $y[-1] = -1$, $x[n] = 0$ for $n < 0$ and $x[n] = (-1)^n$ for n = 0, 1, 2 …

3.

Using z-transforms:

(a) Find the unit-pulse response for the system given by:

$$y[n] = x[n] + \frac{1}{3}y[n-1]$$

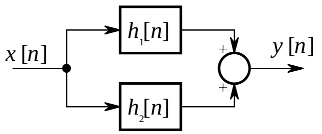
(b) Find the response of this system to the input

$$x[n] = \begin{cases} 0, & n = -1, -2, -3, \ldots \\ 2, & n = 0, 1 \\ 1, & n = 2, 3, 4, \ldots \end{cases}$$

Hint: $x[n]$ can be written as the sum of a unit step and two unit-pulses, or as the subtraction of two unit-steps.

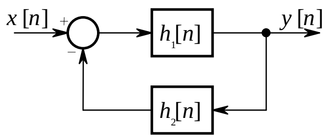
4.

Determine the weighting sequence for the system shown below in terms of the individual weighting sequences $h_1[n]$ and $h_2[n]$.



5.

For the feedback configuration shown below, determine the first three terms of the weighting sequence of the overall system by applying a unit-pulse input and calculating the resultant response. Express your results as a function of the weighting sequence elements $h_1[n]$ and $h_2[n]$.



6.

Using the definition of the z-transform

(a) Find $Y(z)$ when:

(i) $y[n]=0$ for $n<0$, $y[n]=(1/2)^n$ for $n=0,1,2,\ldots$

(ii) $y[n]=0$ for $n\le 0$, $y[n]=a^{n-1}$ for $n=1,2,3,\ldots$

(iii) $y[n]=0$ for $n\le 0$, $y[n]=na^{n-1}$ for $n=1,2,3,\ldots$

(iv) $y[n]=0$ for $n\le 0$, $y[n]=n^2a^{n-1}$ for $n=1,2,3,\ldots$

(b) Determine the z-transform of the sequence

$$x[n]=\begin{cases}2, & n=0,2,4,\ldots\\ 0, & \text{all other } n\end{cases}$$

by noting that $x[n]=x_1[n]+x_2[n]$ where $x_1[n]$ is the unit-step sequence and $x_2[n]$ is the unit-alternating sequence. Verify your result by directly determining the z-transform of $x[n]$.

(c) Use the linearity property of the z-transform and the z-transform of e^{anT} (from tables) to find $Z[\cos(n\omega T)]$. Check the result using a table of z-transforms.

7.

Poles and zeros are defined for the z-transform in exactly the same manner as for the Laplace transform. For each of the z-transforms given below, find the poles and zeros and plot the locations in the z-plane. Which of these systems are stable and unstable?

(a) $H(z)=\dfrac{1+2z^{-1}}{3+4z^{-1}+z^{-2}}$

(b) $H(z)=\dfrac{1}{1+3/4\,z^{-2}+1/8\,z^{-4}}$

(c) $H(z)=\dfrac{5+2z^{-2}}{1+6z^{-1}+3z^{-2}}$

Note: for a discrete-time system to be stable all the *poles* must lie *inside* the unit-circle.

8.

Given:

(a) $y[n]=3x[n]+x[n-1]-2x[n-4]+y[n-1]-y[n-2]$

(b) $y[n+4]-y[n+3]+y[n+2]=3x[n+4]+x[n+3]-2x[n]$

Find the transfer function of the systems

(i) by first finding the unit-pulse response

(ii) by directly taking the *z*-transform of the difference equation

9.

Use the direct division method to find the first four terms of the data sequence $x[n]$, given:

$$X(z)=\frac{14z^2-14z+3}{(z-1/4)(z-1/2)(z-1)}$$

10.

Use the partial fraction expansion method to find a general expression for $x[n]$ in Question 9. Confirm that the first four terms are the same as those obtained by direct division.

11.

Determine the inverse *z*-transforms of:

(a) $X(z)=\dfrac{z^2}{(z-1)(z-a)}$ (b) $X(z)=3+2z^{-1}+6z^{-4}$

(c) $X(z)=\dfrac{(1-e^{-aT})z}{(z-1)(z-e^{-aT})}$ (d) $X(z)=\dfrac{z(z+1)}{(z-1)(z^2-z+1/4)}$

(e) $X(z)=\dfrac{4}{z^3(2z-1)}$ (f) $X(z)=\dfrac{z}{z^2+z+1}$

12.

Given:

$$8y[n]-6y[n-1]+y[n-2]=x[n]$$

(a) Find the unit-pulse response $h[n]$ using time-domain methods.

(b) Find $H(z)$

(i) by directly taking the *z*-transform of the difference equation

(ii) from your answer in (a)

(c) From your solution in (b) find the unit-step response of the system. Check your solution using a convolution method on the original difference equation.

(d) Find the zero-state response if $x[n]=n$.

13.

Given $y[n]=x[n]+y[n-1]+y[n-2]$, find the unit-step response of this system using the transfer function method, assuming zero initial conditions.

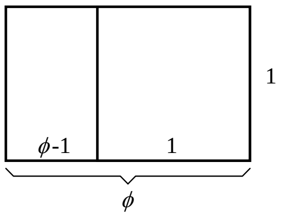
14.

Use *z*-transform techniques to find a closed-form expression for the Fibonacci sequence:

0, 1, 1, 2, 3, 5, 8, 13, 21, 34…

Hint: To use transfer function techniques, the initial conditions must be zero. Construct an input so that the ZSR of the system described by the difference equation gives the above response.

The ancient Greeks considered a rectangle to be perfectly proportioned (saying that the lengths of its sides were in a *golden ratio* to each other) if the ratio of the length to the width of the outer rectangle equalled the ratio of the length to the width of the inner rectangle:



That is:

$$\frac{\phi}{1}=\frac{1}{\phi-1}$$

Find the two values of ϕ that satisfy the golden ratio. Are they familiar values?

15.

Given $F(s)=\frac{1}{(s+a)^2}$, find $F(z)$.

Hint: 1. First find $f(t)$. 2. Then put $t=nT$ to obtain $f(nT)$. 3. Take the z-transform of $f(nT)$.

16.

Given:

$$X_2(z)=\frac{X_1(z)}{z+1}-\frac{z+6}{z+1}X_3(z)+D(z)$$

$$X_3(z)=X_1(z)+kX_2(z)$$

$$X_4(z)=\frac{10}{z-2}X_2(z)-\frac{3}{z+A}X_3(z)$$

(a) Draw a block diagram. Use block diagram reduction to find $X_3(z)$.

Are there any differences between the operations which apply to discrete-time and continuous-time transfer functions?

17.

Using the initial and final value theorems, find $f(0)$ and $f(\infty)$ of the following functions:

(a) $F(z)=\frac{1}{z-0.3}$

(b) $F(z)=1+5z^{-3}+2z^{-2}$

(c) $F(z)=\frac{z^2}{(z-1)(4z-a)}$

(d) $F(z)=\frac{14z^2-14z+3}{(z-1/4)(z-1/2)(z-1)}$

18.

Perform the convolution $y[n] * y[n]$ when

(i) $y[n] = u[n]$

(ii) $y[n] = \delta[n] - \delta[n-1] + \delta[n-2]$

using

(a) the property that convolution in the time-domain is equivalent to multiplication in the frequency domain, and

(b) using any other convolution technique.

Compare answers.

19.

Determine the inverse *z*-transform of:

$$F(z) = \frac{z(z+1)}{(z-1)(z^2 - z + 1/4)}$$

HINT: This has a <u>repeated</u> root. Use techniques analogous to those for the Laplace transform when multiple roots are present.

Lecture 6B – Discretization

Signal discretization. Signal reconstruction. System discretization. Bilinear transformation. Frequency response. Response matching.

Overview

Digital signal processing is now the preferred method of signal processing. Communication schemes and controllers implemented digitally have inherent advantages over their analog counterparts: reduced cost, repeatability, stability with aging, flexibility (one H/W design, different S/W), in-system programming, adaptability (ability to track changes in the environment), and this will no doubt continue into the future. However, much of how we think, analyse and design systems still uses analog concepts, and ultimately most embedded systems eventually interface to a continuous-time world.

It's therefore important that we now take all that we know about continuous-time systems and "transfer" or "map" it into the discrete-time domain.

Signal Discretization

We have already seen how to discretize a signal. An ideal sampler produces a weighted train of impulses:

Sampling a continuous-time signal produces a discrete-time signal (a train of impulses)

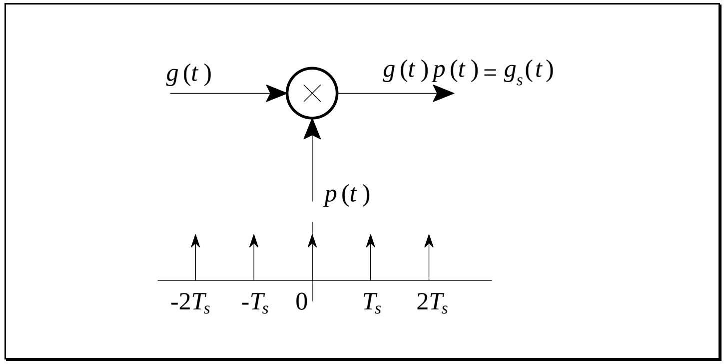


Figure 6B.1

This was how we approached the topic of z-transforms. Of course an ideal sampler does not exist, but a real system can come close to the ideal. We saw in the lab that it didn't matter if we used a rectangular pulse train instead of an impulse train – the only effect was that repeats of the spectrum were weighted by a sinc function. This didn't matter since reconstruction of the original continuous-time signal was accomplished by lowpass filtering the baseband spectrum which was not affected by the sinc function.

Digital signal processing also quantizes the discrete-time signal

In a computer, values can only be stored as discrete values, not only at discrete times. Thus, in a digital system, the output of a sampler is *quantized* so that we have a *digital* representation of the signal. The effects of quantization will be ignored for now – be aware that they exist, and are a source of errors for digital signal processors.

Signal Reconstruction

The reconstruction of a signal from ideal samples was accomplished by an ideal filter:

Lowpass filtering a discrete-time signal (train of impulses) produces a continuous-time signal

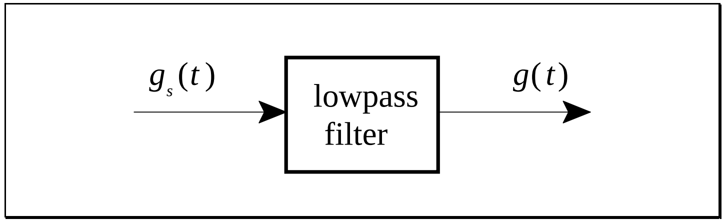


Figure 6B.2

We then showed that so long as the Nyquist criterion was met, we could reconstruct the signal perfectly from its samples (if the lowpass filter is ideal). To ensure the Nyquist criterion is met, we normally place an *anti-alias* filter before the sampler.

Hold Operation

The output from a digital signal processor is obviously digital – we need a way to convert a discrete-time signal back into a continuous-time signal. One way is by using a lowpass filter on the sampled signal, as above. But digital systems have to first convert their digital data into analog data, and this is accomplished with a DAC.

To model a DAC, we note that there is always some output (it never turns off), and that the values it produces are quantized:

The output from a DAC looks like a train of impulses convolved with a rectangle

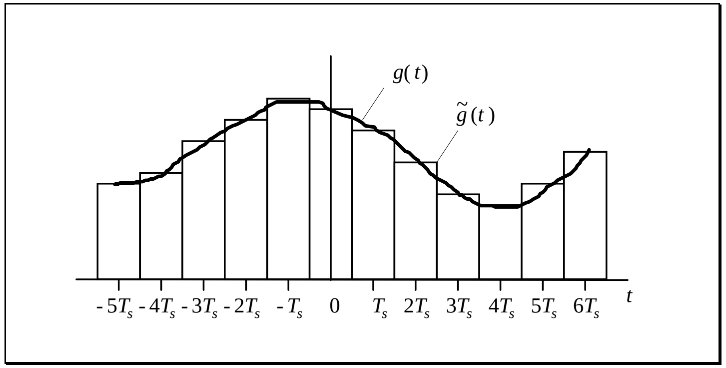


Figure 6B.3

The mathematical model we use for the output of a DAC is:

A DAC is modelled as a zero-order hold device

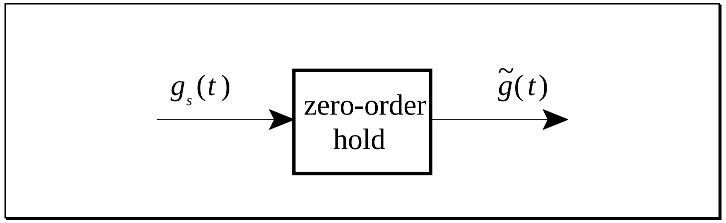


Figure 6B.4

The output of the zero-order hold device is:

$$\tilde{g}(t) = g(nT_s), \quad nT_s \le t \le nT_s + T_s \tag{6B.1}$$

where T_s is the sample period.

The operation of the zero-order hold device in terms of frequency response shows us that it acts like a lowpass filter (but not an ideal one):

Zero-order hold frequency response

$$\boxed{\tilde{G}(f) = T_s \operatorname{sinc}(fT_s) e^{-j\pi f T_s} G_s(f)} \tag{6B.2}$$

Show that the above is true by taking the Fourier transform of $\tilde{g}(t)$.

System Discretization

Suppose we wish to discretize a continuous-time LTI system. We would expect the input/output values of the discrete-time system to be:

$$x[n] = x(nT), \quad y[n] = y(nT) \tag{6B.3}$$

In the frequency-domain, we want an equivalent relationship by taking the Laplace transform of the continuous time system and the *z*-transform of the discrete-time system, while still maintaining Eq. (6B.3):

Discretizing a system should produce the same (sampled) signals

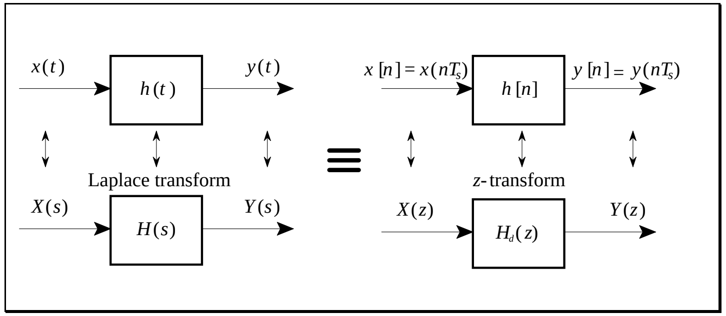


Figure 6B.5

We now have to determine the discrete-time transfer function $H_d(z)$ so that the relationship Eq. (6B.3) holds true. One way is to match the inputs and outputs in the frequency-domain. You would expect that since $z = e^{sT_s}$, then we can simply do:

An exact match of the two systems

$$H_d(z) = H(s)\big|_{s=(1/T_s)\ln z} \tag{6B.4}$$

Unfortunately, this leads to a *z*-domain expression which is not rational in *z*. We want rational *z* because eventually the system is implemented as a difference equation, which needs delays to be integers (how would you shift an array of numbers by a half?). If we knew complex variable theory we'd now use a Laurent's series to easily approximate $\ln z$.

We can't implement the exact match because the transfer function is not rational

However, we can get by using Taylor's theorem:

$$f(x) = f(a) + (x-a) f'(a) + \frac{(x-a)^2}{2!} f''(a) + \cdots \tag{6B.5}$$

Now set $f(x) = \ln x$, $a = 1$ and let $x = 1 + y$:

$$\ln(1+y) = y - \frac{y^2}{2} + \frac{y^3}{3} - \frac{y^4}{4} + \cdots \tag{6B.6}$$

Now, from inspection, we can also have:

$$\ln(1-y) = -y - \frac{y^2}{2} - \frac{y^3}{3} - \frac{y^4}{4} + \cdots \tag{6B.7}$$

Subtracting Eq. (6B.7) from Eq. (6B.6), and dividing by 2 gives us:

$$\frac{1}{2}\ln\left(\frac{1+y}{1-y}\right) = y + \frac{y^3}{3} + \frac{y^5}{5} + \frac{y^7}{7} + \cdots \tag{6B.8}$$

Now let:

$$z = \frac{1+y}{1-y} \tag{6B.9}$$

or, rearranging:

$$y = \frac{z-1}{z+1} \tag{6B.10}$$

Substituting Eqs. (6B.9) and (6B.10) into Eq. (6B.8) yields:

$$\ln z = 2\left\{ \left(\frac{z-1}{z+1}\right) + \frac{1}{3}\left(\frac{z-1}{z+1}\right)^3 + \frac{1}{5}\left(\frac{z-1}{z+1}\right)^5 + \cdots \right\} \tag{6B.11}$$

Now if $z \approx 1$, then we can truncate higher than first-order terms to get the approximation:

$$\ln z \approx 2\left(\frac{z-1}{z+1}\right) \tag{6B.12}$$

We can now use this as an approximate value for *s*. This is called the *bilinear transformation* (since it has 2 (bi) linear terms):

Bilinear transformation defined - an approximate mapping between s and *z*

$$s = \frac{1}{T_s}\ln z \approx \frac{2}{T_s}\left(\frac{z-1}{z+1}\right) \tag{6B.13}$$

This transformation has several desirable properties:

- The open LHP in the *s*-domain maps onto the open unit disk in the *z*-domain (thus the bilinear transformation preserves the stability condition).
- The $j\omega$-axis maps onto the unit circle in the *z*-domain. This will be used later when we look at frequency response of discrete-time systems.

So, an approximate mapping from a continuous-time system to a discrete-time system, from Eq. (6B.4) is:

Discretizing a system using the bilinear transformation

$$H_d(z) \approx H\left(\frac{2}{T_s}\frac{z-1}{z+1}\right) \tag{6B.14}$$

Frequency Response

Since $z = e^{sT_s}$, then letting $s = j\omega$ gives $z = e^{j\omega T_s}$. If we define:

The relationship between discrete-time frequency and continuous-time frequency

$$\Omega = \omega T_s \tag{6B.15}$$

then the frequency response in the *z*-domain is found by setting:

Value of *z* to determine discrete-time frequency response

$$z = e^{j\Omega} \tag{6B.16}$$

This value of *z* is found on the unit-circle and has an angle of $\Omega = \omega T_s$:

Frequency response mapping

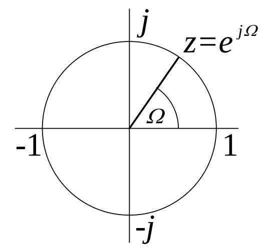


Figure 6B.6

But this point is also given by the angle $\Omega + 2n\pi$, so the mapping from the *z*-plane to the *s*-domain is not unique. In fact any point on the unit circle maps to the *s*-domain frequency $\omega T_s = \Omega + 2n\pi$, or:

Frequency response mapping is not unique – aliasing

$$\begin{aligned}\omega &= \frac{\Omega}{T_s} + n\frac{2\pi}{T_s} \\ &= \frac{\Omega}{T_s} + n\omega_s\end{aligned} \tag{6B.17}$$

where ω_s is the sampling frequency in radians.

This shows us two things:

- The mapping from the *z*-domain to the *s*-domain is not unique. Conversely, the mapping from the *s*-domain to the *z*-domain is not unique. This means that frequencies spaced at ω_s map to the *same* frequency in the *z*-domain. We already know this! It's called *aliasing*.

Discrete-time frequency response is periodic

- The frequency response of a discrete-time system is periodic with period 2π, which means it can be completely characterised by restricting Ω so that $-\pi \le \Omega \le \pi$.

Just like the continuous-time case where we set $s = j\omega$ in $H(s)$ to give the frequency response $H(\omega)$, we can set $z = e^{j\Omega}$ in $H_d(z)$ to give the frequency response $H_d(\Omega)$. Doing this with Eq. (6B.14) yields the approximation:

$$H_d(\Omega) \approx H\left(\frac{2}{T_s}\frac{e^{j\Omega}-1}{e^{j\Omega}+1}\right) \tag{6B.18}$$

The frequency that corresponds to Ω in the *s*-domain is approximately given by Eq. (6B.13):

$$j\omega \approx \frac{2}{T_s}\frac{e^{j\Omega}-1}{e^{j\Omega}+1} \tag{6B.19}$$

Using Euler's identity show that this can be manipulated into the form:

Mapping continuous-time frequency to discrete-time frequency using the bilinear transformation

$$\boxed{\omega = \frac{2}{T_s}\tan\frac{\Omega}{2}} \tag{6B.20}$$

The inverse relationship is:

$$\boxed{\Omega = 2\tan^{-1}\frac{\omega T_s}{2}} \tag{6B.21}$$

So now Eq. (6B.18), the approximate frequency response of an equivalent discrete-time system is given by:

Discretizing a system to get a similar frequency response

$$\boxed{H_d(\Omega) \approx H\left(\frac{2}{T_s}\tan\frac{\Omega}{2}\right)} \tag{6B.22}$$

The distortion caused by the approximation Eq. (6B.21) is called frequency *warping*, since the relationship is non-linear. If there is some critical frequency (like the cutoff frequency of a filter) that must be preserved in the transformation, then we can *pre-warp* the continuous-time frequency response before applying Eq. (6B.21).

Note that we may be able to select the sample period *T* so that all our "critical" frequencies will be mapped by:

$$\begin{aligned}\Omega_c &\approx 2\tan^{-1}\frac{\omega_c T_s}{2} \approx 2\left(\frac{\omega_c T_s}{2}\right) \\ &\approx \omega_c T_s\end{aligned} \tag{6B.23}$$

which is a small deviation from the real mapping $\Omega = \omega T_s$ given by Eq. (6B.15).

Response Matching

Time-domain (step response) matching

In control systems, it is usual to consider a mapping from continuous-time to discrete-time in terms of the time-domain response instead of the frequency response. For *set-point* control, this mapping is best performed as *step invariance synthesis*, although other mappings can be made (like impulse invariance).

We saw before that we want:

$$x[n] = x(nT_s), \quad y[n] = y(nT_s) \tag{6B.24}$$

where the output of the discrete-time system is obtained by sampling the step-response of the continuous-time system.

Since for a unit-step input we have:

$$X(z) = \frac{z}{z-1} \tag{6B.25}$$

then we want:

Step response matching to get the discrete-time transfer function

$$\boxed{H_d(z) = Y(z)\frac{z-1}{z}} \tag{6B.26}$$

where $Y(z)$ is the z-transform of the step-response of the continuous-time system.

Example

Suppose we design a maze rover velocity controller in the continuous-time domain and we are now considering its implementation in the rover's microcontroller. We might come up with a continuous-time controller transfer function such as:

$$G_c(s) = 500 + \frac{5}{s} \tag{6B.27}$$

Our closed loop system is therefore:

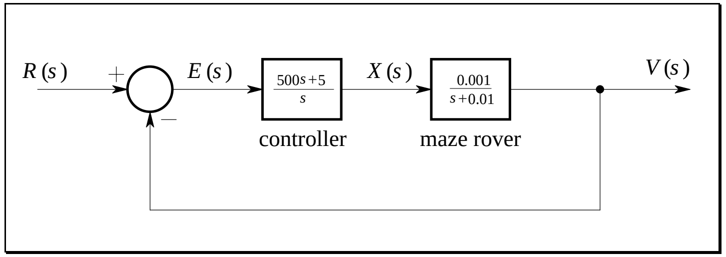


Figure 6B.7

Show that the block diagram can be reduced to the transfer function:

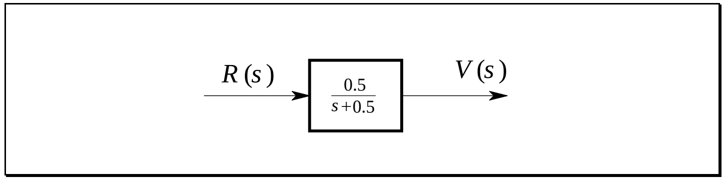


Figure 6B.8

The transform of the step response is then:

$$\begin{aligned} Y(s) &= \frac{0.5}{s+0.5}\frac{1}{s} \\ &= \frac{1}{s} - \frac{1}{s+0.5} \end{aligned} \tag{6B.28}$$

and taking the inverse Laplace transform gives the step response:

Continuous-time step response

$$y(t)=\left(1-e^{-0.5t}\right)u(t) \tag{6B.29}$$

The discretized version of the step response is:

and desired discrete-time step response

$$y[n]=\left(1-e^{-0.5nT_s}\right)u[n] \tag{6B.30}$$

and taking the z-transform gives:

$$Y(z)=\frac{z}{z-1}-\frac{z}{z-e^{-0.5T_s}} \tag{6B.31}$$

Hence, using Eq. (6B.26) yields the following transfer function for the corresponding discrete-time system:

Equivalent discrete-time transfer function using step response

$$H_d(z)=1-\frac{z-1}{z-e^{-0.5T_s}}=\frac{1-e^{-0.5T_s}}{z-e^{-0.5T_s}}=\frac{bz^{-1}}{1-az^{-1}} \tag{6B.32}$$

We would therefore implement the difference equation:

Equivalent discrete-time difference equation using step response

$$y[n]=ay[n-1]+bx[n-1] \tag{6B.33}$$

You should confirm that this difference equation gives an equivalent step-response at the sample instants using MATLAB® for various values of T.

Summary

- We can discretize a continuous-time signal by sampling. If we meet the Nyquist criterion for sampling, then all the information will be contained in the resulting discrete-time signal. We can then process the signal digitally.
- We can reconstruct a continuous-time signal from mere numbers by using a DAC. We model this as passing our discrete-time signal (a weighted impulse train) through a lowpass filter with a rectangular impulse response.
- A system (or signal) may be discretized using the bilinear transform. This maps the LHP in the s-domain into the open unit disk in the z-domain. It is an approximation only, and introduces warping when examining the frequency response.
- Response matching derives an equivalent discrete-time transfer function so that the signals in the discrete-time system exactly match samples of the continuous-time system's input and output signals.

References

Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB®*, Prentice-Hall, 1997.

6B.14

Exercises

1.

A Maze Rover phase-lead compensator has the transfer function:

$$H(s) = \frac{20(s+1)}{(s+4)}$$

Determine a difference equation that approximates this continuous-time system using the method of step-response matching. The sample period is 64 ms.

2.

Repeat Exercise 1, but this time use the bilinear transform.

3.

A continuous-time system has a transfer function $G(s) = \frac{-s+2}{s^2+3s+2}$ and it is required to find the transfer function of an equivalent discrete-time system $H(z)$ whose unit-step response consists of samples of the continuous-time system's unit-step response. Find $H(z)$ assuming a sample time of $1\,\text{s}$. Compare the time solutions at $t = 0, 1, 10, 100\,\text{s}$ to verify your answer.

4.

Compare the step responses of each answer in Exercises 1 and 2 using MATLAB®.

Lecture 7A – System Design

Design criteria for continuous-time systems. Design criteria for discrete-time systems.

Overview

The design of a control system centres around two important specifications – the steady-state performance and the transient performance. The steady-state performance is relatively simple. Given a reference input, we want the output to be exactly, or nearly, equal to the input. The way to achieve this steady-state error for a step-input was mentioned in Lecture 5A. The more complex specification of transient behaviour is tackled by an examination of the system's poles. We will examine transient performance as specified for an *all-pole second-order system* (the transient performance criteria still exist for higher-order systems, but the formula shown here only apply to all-pole second-order systems).

The transient specification for a control system will usually consist of times taken to reach certain values, and allowable deviations from the final steady-state value. For example, we might specify percent overshoot, peak time, and 5% settling time for a control system. Our task is to find suitable pole locations for the system.

Design Criteria for Continuous-Time Systems

Percent Overshoot

The percent overshoot for a second-order all pole step-response is given by (see Lecture 5A):

$$P.O. = e^{\frac{-\pi\zeta}{\sqrt{1-\zeta^2}}} \tag{7A.1}$$

Therefore, for a given *maximum* overshoot specification, we can evaluate the *minimum* allowable ζ. But what *specific* ζ do we choose? We don't know until we take into consideration other specifications! What we do is simply define a *region* in the *s*-plane where we can meet this specification.

Since $\zeta = \cos\theta$, where θ is the angle with respect to the negative real axis, then we can shade the region of the *s*-plane where the *P.O.* specification is satisfied:

PO specification region in the *s*-plane

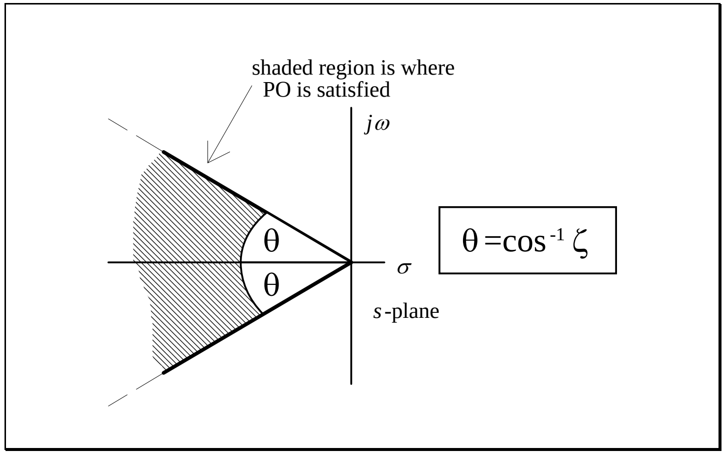


Figure 7A.1

Peak Time

The peak time for a second-order all-pole step-response is given by:

$$t_p = \frac{\pi}{\omega_d} \tag{7A.2}$$

Since a specification will specify a *maximum* peak time, then we can find the *minimum* $\omega_d = \pi/t_p$ required to meet the specification.

Again, we define the region in the *s*-plane where this specification is satisfied:

Peak time specification region in the *s*-plane

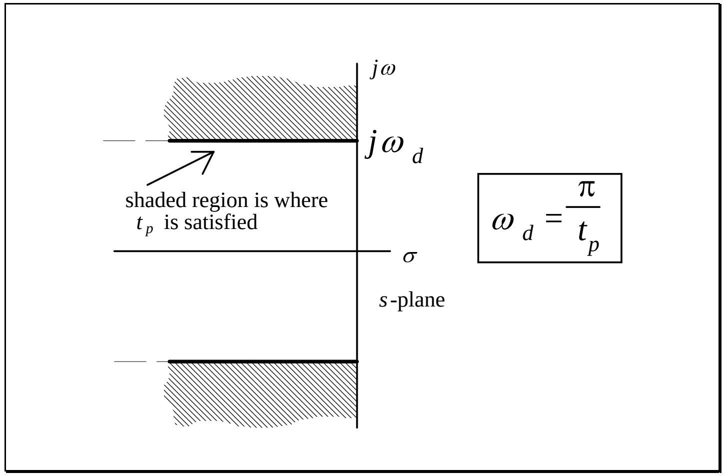


Figure 7A.2

Settling Time

The settling time for a second-order all-pole step-response is given by:

$$t_s = -\frac{\ln \delta}{\alpha} \tag{7A.3}$$

Since a specification will specify a *maximum* settling time, then we can find the *minimum* $\alpha = -\ln\delta / t_s$ required to meet the specification.

We define the region in the *s*-plane where this specification is satisfied:

Settling time specification region in the *s*-plane

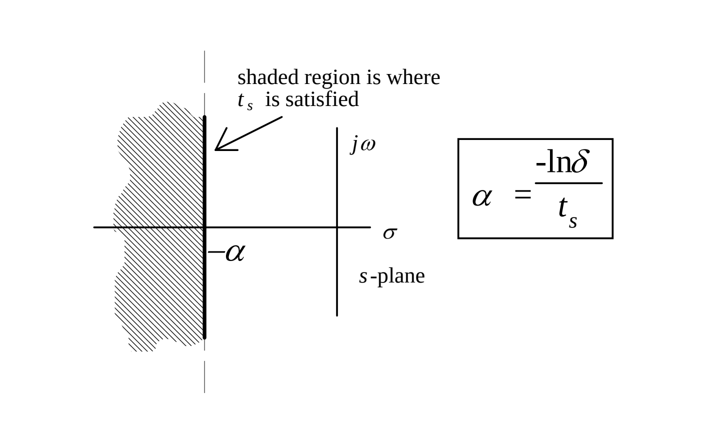


Figure 7A.3

Combined Specifications

Since we have now specified simple regions in the *s*-plane that satisfy each specification, all we have to do is *combine* all the regions to meet *every* specification.

Combined specifications region in the *s*-plane

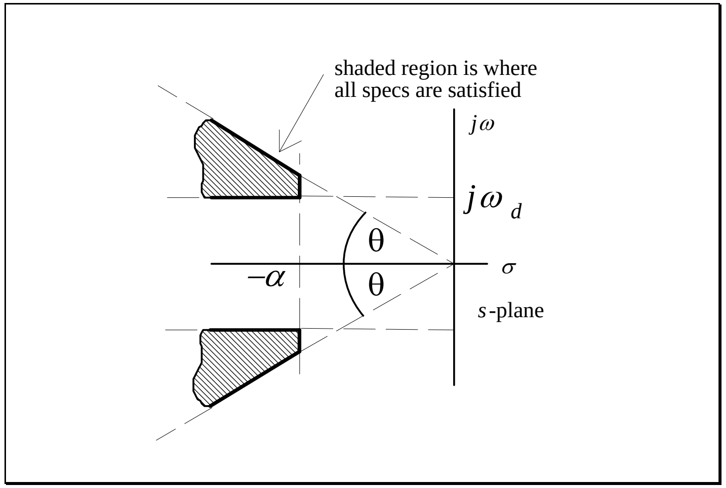


Figure 7A.4

Sometimes a specification is automatically met by meeting the other specifications – this is clear once the regions are drawn on the *s*-plane.

It is now up to us to *choose*, within reasonable limits, the *desired closed-loop* pole locations. The region we have drawn on the *s*-plane is an *output* specification – we must give consideration to the inputs of the system! For example, there is nothing theoretically wrong with choosing closed-loop pole locations out near infinity – we'd achieve a very nice, sharp, almost step-like response! In practice, we can't do this, because the inputs to our system would exceed the allowable linear range. Our analysis has considered the system to be linear – we know it is not in practice! Op amps saturate; motors cannot have megavolts applied to their windings without breaking down; and we can't put our foot on the accelerator past the floor of the car (no matter how much we try)!

We choose *s*-plane poles close to the origin so as not to exceed the linear range of systems

Practical considerations therefore mean we should choose pole locations that are close to the origin – this will mean we meet the specifications, and hopefully we won't be exceeding the linear bounds of our system. We normally indicate the *desired* pole locations by placing a square around them:

Desired closed-loop pole locations are represented with a square around them

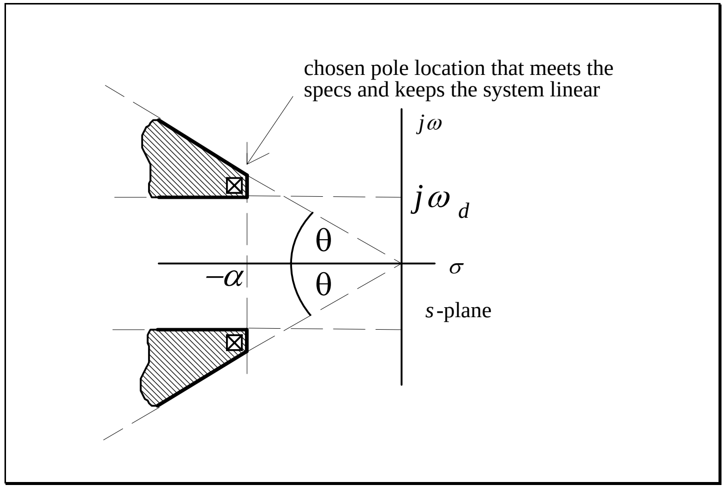


Figure 7A.5

Can we apply these specifications to real (higher-order) systems? Yes – if it's a dominant second-order all-pole system:

Dominant second-order all-pole system

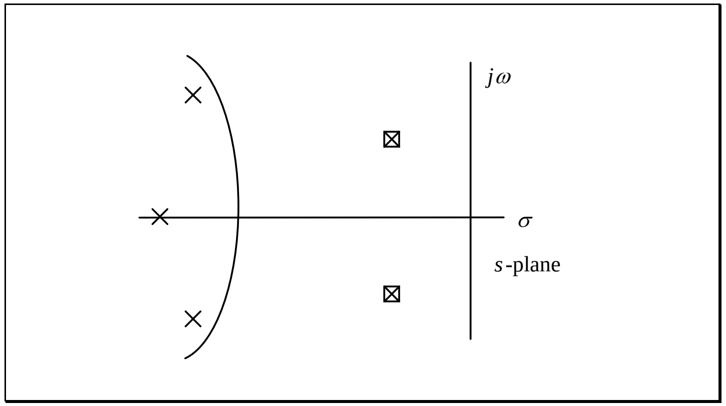


Figure 7A.6

or we introduce a minor-loop to move a dominant first-order pole away:

Dominant second-order all-pole system made by moving a pole using a minor-feedback loop

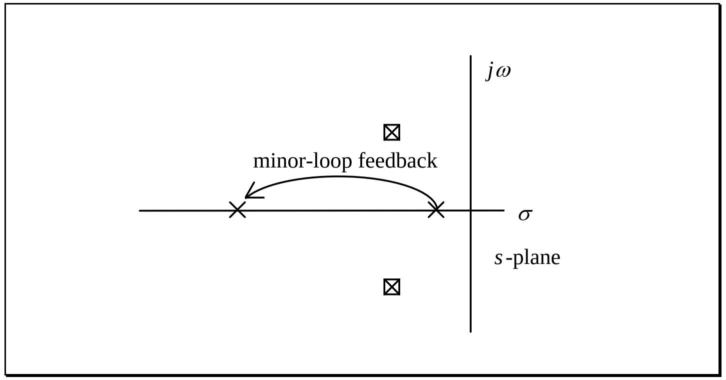


Figure 7A.7

or we achieve a pole-zero cancellation by placing a zero very close to the dominant pole:

Dominant second-order all-pole system made by pole-zero cancellation

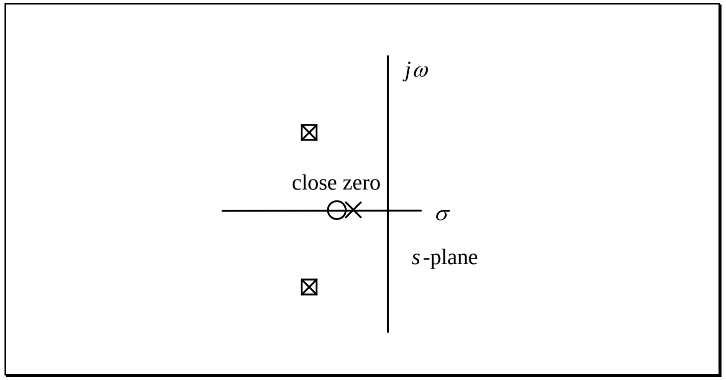


Figure 7A.8

We should be careful when implementing pole-zero cancellation. Normally we do not know *exactly* the pole locations of a system, so the "cancellation" is only approximate. In some cases this inexact cancellation can cause the opposite of the desired effect – the system will have a dominant first-order pole and the response will be "sluggish".

Pole-zero cancellation is inexact

Example

We need to design a maze rover position controller to achieve the following specifications:

(a) PO < 10%

(b) $t_p < 1.57\,\text{s}$

(c) 2% settling time $t_s < 9\,\text{s}$

We evaluate the regions in the order they are given. For the PO spec, we have:

$$0.1 = e^{\frac{-\pi\zeta}{\sqrt{1-\zeta^2}}}$$
$$\zeta = 0.591$$

so that:

$$\theta = \cos^{-1}\zeta = 53°$$

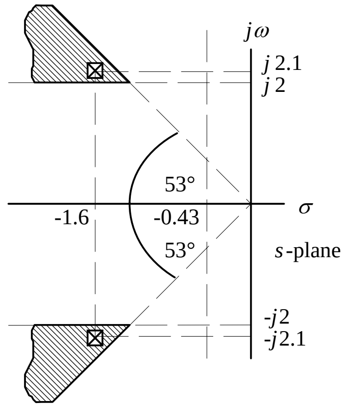
For the peak time spec, we have:

$$\omega_d = \frac{\pi}{1.57} \approx 2\,\text{rads}^{-1}$$

For the settling time spec, we get:

$$\alpha = \frac{-\ln 0.02}{9} = \frac{39}{9} = 0.43\,\text{nepers}$$

The regions can now be drawn:



We choose the desired closed-loop poles to be at $s = -1.6 \pm j2.1$ to meet all our specifications.

Design Criteria for Discrete-Time Systems

Specification regions in the *z*-plane

A second-order continuous-time system can be "translated" into a second-order discrete-time system using either the bilinear (Tustin) transform, zero-order holds, or a variety of other methods. Discretizing a continuous-time system inevitably involves a sample time – this sample time affects the pole locations, as will be shown in Lecture 7B. We therefore would like to see where our poles *should be* to satisfy the initial design specifications. In transforming the performance specifications to the *z*-plane, it is first convenient to see how a single point of the *s*-plane maps into the *z*-plane.

Mapping of a Point from the *s*-plane to the *z*-plane

By definition, $z = e^{sT_s}$, and $s = \sigma + j\omega$. Therefore:

$$\begin{aligned} z &= e^{\sigma T_s} e^{j\omega T_s} \\ &= e^{\sigma T_s} \angle \omega T_s \\ &= |z| \angle z \end{aligned} \tag{7A.4}$$

That is, we have:

Mapping a point from the *s*-plane to the *z*-plane

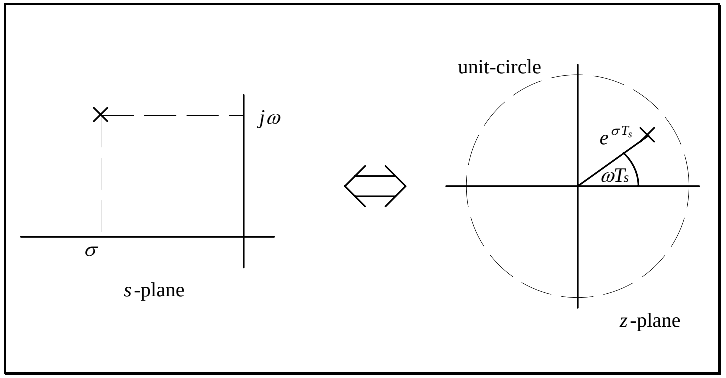


Figure 7A.9

We can see that if:

$$\begin{aligned} \sigma > 0 \quad & |z| > 1 \\ \sigma = 0 \quad & |z| = 1 \\ \sigma < 0 \quad & |z| < 1 \end{aligned} \tag{7A.5}$$

We will now translate the performance specification criteria areas from the *s*-plane to the *z*-plane.

Percent Overshoot

In the *s*-plane, the PO specification is given by $s = -\omega + j\omega \tan\left(\cos^{-1}\zeta\right)$. This line in the *z*-plane must be:

$$\begin{aligned} z = e^{sT_s} &= e^{-\omega T_s} e^{j\omega T_s \tan\left(\cos^{-1}\zeta\right)} \\ &= e^{-\omega T_s} \angle \omega T_s \tan\left(\cos^{-1}\zeta\right) \end{aligned} \tag{7A.6}$$

For a given ζ, this locus is a logarithmic spiral:

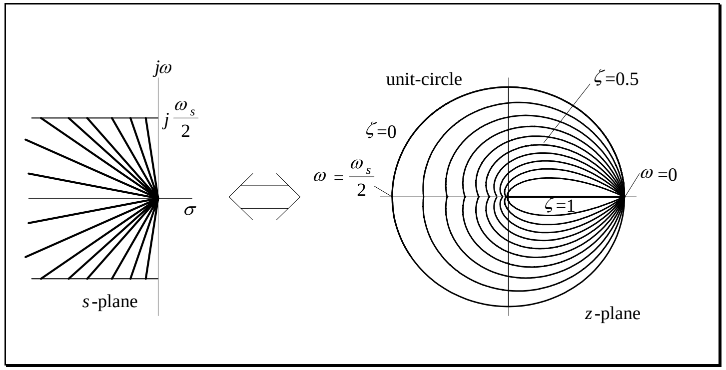


Figure 7A.10

The *region* in the *z*-plane that corresponds to the *region* in the *s*-plane where the PO specification is satisfied is shown below:

PO specification region in the *z*-plane

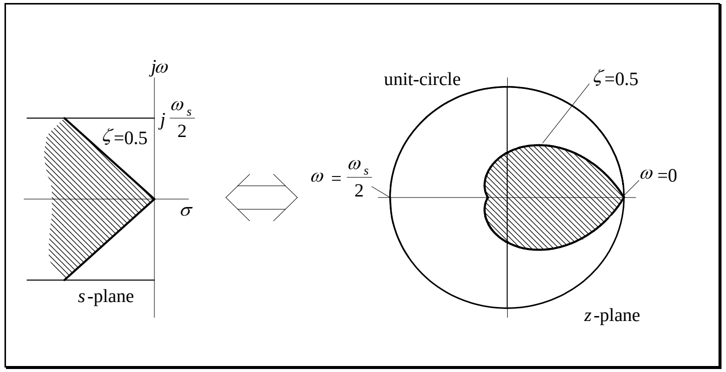


Figure 7A.11

Peak Time

In the s-plane, the peak time specification is given by $s = \sigma + j\omega_d$. A line of constant frequency in the z-plane must be:

$$\begin{aligned} z = e^{sT_s} &= e^{\sigma T_s} e^{j\omega_d T_s} \\ &= e^{\sigma T_s} \angle \omega_d T_s \end{aligned} \tag{7A.7}$$

For a given ω_d, this locus is a straight line between the origin and the unit-circle, at an angle $\Omega = \omega_d T_s$ (remember we only consider the LHP of the s-plane so that $-\infty < \sigma < 0$):

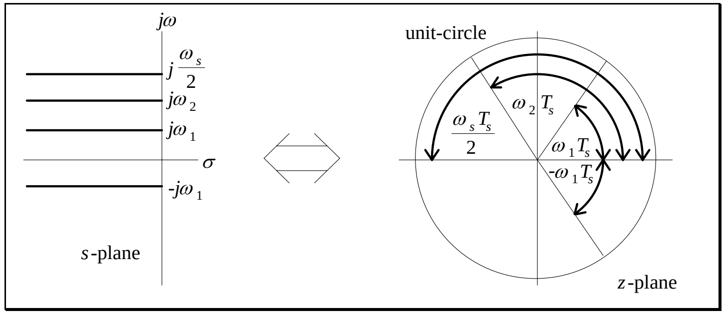


Figure 7A.12

The *region* in the z-plane that corresponds to the *region* in the s-plane where the peak time specification is satisfied is shown below:

Peak time specification region in the z-plane

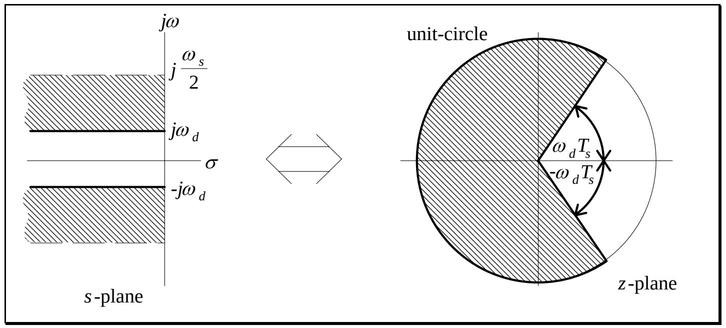


Figure 7A.13

Settling Time

In the s-plane, the settling time specification is given by $s = -\alpha + j\omega$. The corresponding locus in the z-plane must be:

$$\begin{aligned} z = e^{sT_s} &= e^{-\alpha T_s} e^{j\omega T_s} \\ &= e^{-\alpha T_s} \angle \omega T_s \end{aligned} \tag{7A.8}$$

For a given α, this locus is a circle centred at the origin with a radius $e^{-\alpha T_s}$:

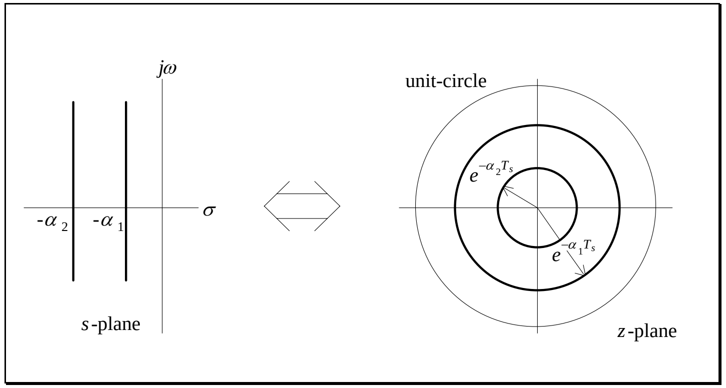


Figure 7A.14

The *region* in the z-plane that corresponds to the *region* in the s-plane where the settling time specification is satisfied is shown below:

Settling time specification region in the z-plane

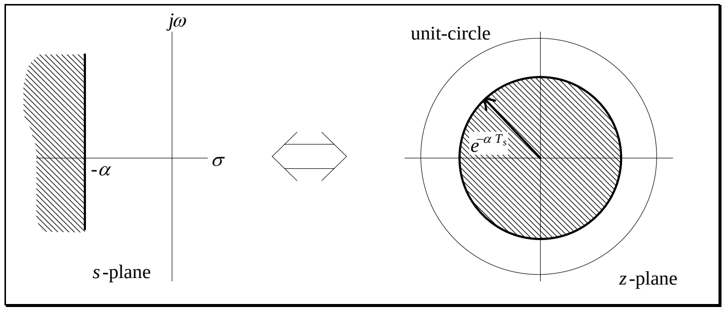


Figure 7A.15

Combined Specifications

We choose z-plane poles close to the unit-circle so as not to exceed the linear range of systems

A typical specification will mean we have to combine the PO, peak time and settling time regions in the z-plane. For the s-plane we chose pole locations close to the origin. Since $|z| = e^{\sigma T_s}$, and σ is a small negative number near the origin, then we need to maximize $|z|$. We therefore choose pole locations in the z-plane which satisfy all the criteria, and we choose them as far away from the origin as possible:

Combined specifications region in the z-plane

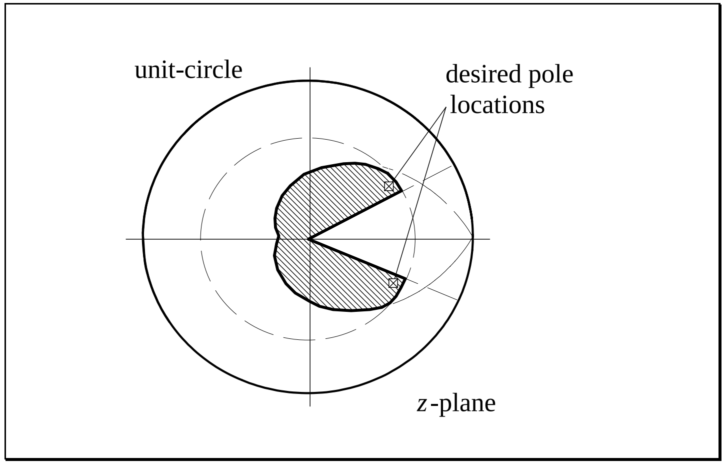


Figure 7A.16

Sometimes we perform a design in the s-plane, then discretize it using a method such as the bilinear transform or step-response matching. We should *always* check whether the resulting discrete-time system will meet the specifications – it may not, due to the imperfections of the discretization!

Summary

- The percent overshoot, peak time and settling time specifications for an all-pole second-order system can easily be found in the s-plane. Desired pole locations can then be chosen that satisfy all the specifications, and are usually chosen close to the origin so that the system remains in its linear region.

- We can apply the all-pole second-order specification regions to systems that are dominant second-order systems.

- The percent overshoot, peak time and settling time specifications for an all-pole second-order system can be found in the z-plane. The desired pole locations are chosen as close to the unit-circle as possible so the system stays within its linear bounds.

References

Kuo, B*: Automatic Control Systems*, Prentice-Hall, 1995.

Nicol, J.: *Circuits and Systems 2 Notes*, NSWIT, 1986.

Lecture 7B – Root Locus

Root locus. Root locus rules. MATLAB®'s RLTool. Root loci of discrete-time systems. Time-response of discrete-time control systems.

Overview

The roots of a system's characteristic equation are important in two respects – they determine whether the system is stable, and they play a part in determining the transient response (along with any zeros of course). A "root locus" is a graphical way of showing how the roots of a characteristic equation in the complex (s or z) plane vary as some parameter is varied. It is an extremely valuable aid in the analysis and design of control systems, and was developed by W. R. Evans in his 1948 paper "Graphical Analysis of Control Systems," *Trans. AIEE*, vol. 67, pt. II, pp. 547-551, 1948.

Root Locus

As an example of the root locus technique, we will consider a simple unity-feedback control system:

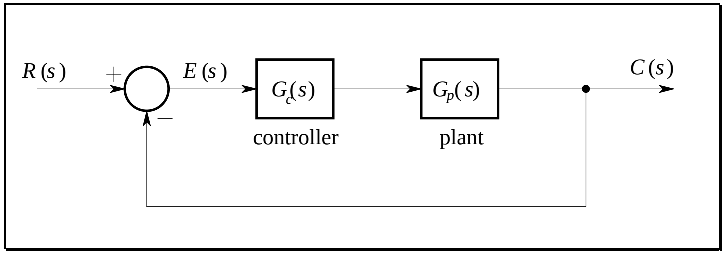


Figure 7B.1

We know that the closed-loop transfer function is just:

$$\frac{C(s)}{R(s)} = \frac{G_c(s)G_p(s)}{1+G_c(s)G_p(s)} \tag{7B.1}$$

This transfer function has the characteristic equation:

$$1+G_c(s)G_p(s)=0 \tag{7B.2}$$

Now suppose that we can "separate out" the parameter of interest, K, in the characteristic equation – it may be the gain of an amplifier, or a sampling rate, or some other parameter that we have control over. It could be part of the controller, or part of the plant. Then the characteristic equation can be written:

Characteristic equation of a unity-feedback system

$$\boxed{1+KP(s)=0} \tag{7B.3}$$

where $P(s)$ does *not* depend on K. The graph of the roots of this equation, as the parameter K is varied, gives the root locus. In general:

$$KP(s)=K\frac{(s-z_1)(s-z_2)\ldots(s-z_m)}{(s-p_1)(s-p_2)\ldots(s-p_n)} \tag{7B.4}$$

where z_i are the m open-loop zeros and p_i are the n open-loop poles of the system. Also, rearranging Eq. (7B.3) gives us:

$$KP(s)=-1 \tag{7B.5}$$

Taking magnitudes of both sides leads to the *magnitude criterion* for a root locus:

Root locus magnitude criterion

$$\boxed{|P(s)|=1/|K|} \tag{7B.6a}$$

Similarly, taking angles of both sides of Eq. (7B.5) gives the *angle criterion* for a root locus:

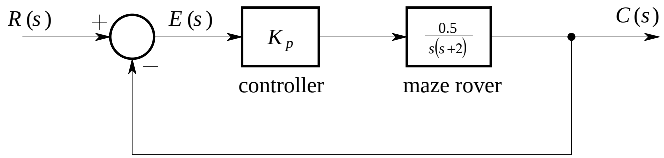
Root locus angle criterion

$$\boxed{\angle K+\angle P(s)=180^\circ \pm 360^\circ r \quad r=0,1,2,\ldots} \tag{7B.6b}$$

To construct a root locus we can just apply the angle criterion. To find a particular *point* on the root locus, we need to know the magnitude of K.

Example

A simple maze rover positioning scheme is shown below:



We are trying to see the effect that the controller parameter K_p has on the closed-loop system. First of all, we can make the following assignments:

$$G_p(s)=\frac{0.5}{s(s+2)} \text{ and } G_c(s)=K_p$$

Putting into the form of Eq. (7B.3), we then have:

$$K=K_p \text{ and } P(s)=\frac{0.5}{s(s+2)}$$

For such a simple system, it is easier to derive the root locus algebraically rather than use the angle criterion. The characteristic equation of the system is:

$$1+KP(s)=1+K_p\frac{0.5}{s(s+2)}=0$$

or just:

$$s^2+2s+0.5K_p=0$$

The roots are then given by:

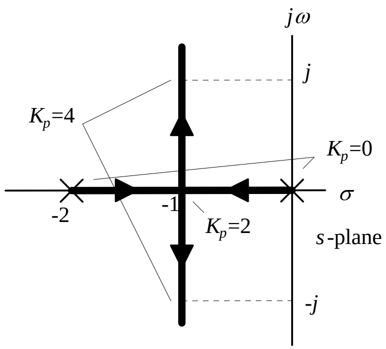
$$s=-1\pm\sqrt{1-0.5K_p}$$

We will now evaluate the roots for various values of the parameter K_p:

$$\begin{array}{ll} K_p = 0 & s_{1,2} = 0, -2 \\ K_p = 1 & s_{1,2} = -1+1/\sqrt{2}, -1-1/\sqrt{2} \\ K_p = 2 & s_{1,2} = -1, -1 \\ K_p = 3 & s_{1,2} = -1+j/\sqrt{2}, -1-j/\sqrt{2} \\ K_p = 4 & s_{1,2} = -1+j, -1-j \end{array}$$

The root locus is thus:

Root locus for a simple two-pole system



What may have not been obvious before is now readily revealed: the system is unconditionally stable (for positive K_p) since the poles always lie in the LHP; and the K_p parameter can be used to position the poles for an overdamped, critically damped, or underdamped response. Also note that we can't arbitrarily position the poles anywhere on the *s*-plane – we are restricted to the root locus. This means, for example, that we cannot increase the damping of our underdamped response – it will always be e^{-t}.

Root Locus Rules

We will now examine a few important "rules" about a root locus construction that can give us insight into how a system behaves as the parameter *K* is varied.

Root locus construction rules

1. Number of Branches

If $P(s)$ has *n* poles and *m* zeros then there will be *n* branches.

2. Locus End Points

The root locus starts (i.e. $K = 0$) at the poles of $P(s)$. This can be seen by substituting Eq. (7B.4) into Eq. (7B.3) and rearranging:

$$(s - p_1)(s - p_2)\ldots(s - p_n) + K(s - z_1)(s - z_2)\ldots(s - z_m) = 0 \qquad (7B.7)$$

This shows us that the roots of Eq. (7B.3), when $K = 0$, are just the open-loop poles of $P(s)$, which are also the poles of $G_c(s)G_p(s)$.

As $|K| \to \infty$, the root locus branches terminate at the zeros of $P(s)$. For $n > m$ then $n - m$ branches go to infinity.

3. Real Axis Symmetry

The root locus is symmetrical with respect to the real axis.

4. Real Axis Sections

Any portion of the real axis forms part of the root locus for $K > 0$ if the total number of real poles and zeros to the right of an exploratory point along the real axis is an odd integer. For $K < 0$, the number is zero or even.

5. Asymptote Angles

The $(n-m)$ branches of the root locus going to infinity have asymptotes given by:

$$\boxed{\phi_A = \frac{r}{n-m}180^\circ} \tag{7B.8}$$

For $K>0$, r is odd (1, 3, 5, …, $n-m$).

For $K<0$, r is even (0, 2, 4, …, $n-m$).

This can be shown by considering the root locus as it is mapped far away from the group of open-loop poles and zeros. In this area, all the poles and zeros contribute about the same angular component. Since the total angular component must add up to 180° or some odd multiple, Eq. (7B.8) follows.

6. Asymptotic Intercept (Centroid)

The asymptotes *all* intercept at *one* point on the real axis given by:

$$\boxed{\sigma_A = \frac{\sum_n \text{poles of } P(s) - \sum_m \text{zeros of } P(s)}{n-m}} \tag{7B.9}$$

The value of σ_A is the *centroid* of the open-loop pole and zero configuration.

7. Real Axis Breakaway and Break-In Points

A *breakaway* point is where a section of the root locus branches from the real axis and enters the complex region of the *s*-plane in order to approach zeros which are finite or are located at infinity. Similarly, there are branches of the root locus which must *break-in* onto the real axis in order to terminate on zeros.

Examples of breakaway and break-in points are shown below:

Breakaway and break-in points of a root locus

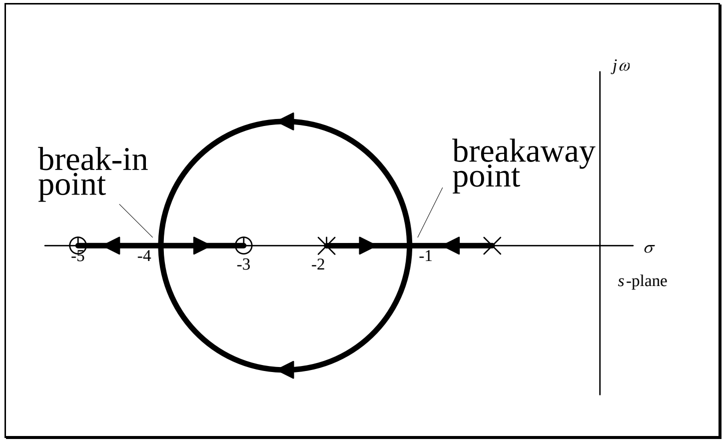


Figure 7B.2

The breakaway (and break-in) points correspond to points in the *s*-plane where multiple roots of the characteristic equation occur. A simple method for finding the breakaway points is available. Taking a lead from Eq. (7B.7), we can write the characteristic equation $1+KP(s)=0$ as:

$$f(s) = B(s) + KA(s) = 0 \tag{7B.10}$$

where $A(s)$ and $B(s)$ do not contain *K*. Suppose that $f(s)$ has multiple roots of order *r*. Then $f(s)$ may be written as:

$$f(s) = (s-p_1)^r (s-p_2)\cdots(s-p_k) \tag{7B.11}$$

If we differentiate this equation with respect to *s* and set $s=p_1$, then we get:

$$\left.\frac{df(s)}{ds}\right|_{s=p_1} = 0 \tag{7B.12}$$

This means that multiple roots will satisfy Eq. (7B.12). From Eq. (7B.10) we obtain:

$$\frac{df(s)}{ds} = B'(s) + KA'(s) = 0 \tag{7B.13}$$

The particular value of *K* that will yield multiple roots of the characteristic equation is obtained from Eq. (7B.13) as:

$$K = -\frac{B'(s)}{A'(s)} \tag{7B.14}$$

If we substitute this value of *K* into Eq. (7B.10), we get:

$$f(s) = B(s) - \frac{B'(s)}{A'(s)} A(s) = 0 \tag{7B.15}$$

or

$$B(s)A'(s) - B'(s)A(s) = 0 \tag{7B.16}$$

On the other hand, from Eq. (7B.10) we obtain:

$$K = -\frac{B(s)}{A(s)} \tag{7B.17}$$

and:

$$\frac{dK}{ds} = -\frac{B'(s)A(s) - B(s)A'(s)}{A^2(s)} \tag{7B.18}$$

If dK/ds is set equal to zero, we get the same equation as Eq. (7B.16). Therefore, the breakaway points can be determined from the roots of:

Equation to find breakaway points

$$\boxed{\frac{dK}{ds} = 0} \tag{7B.19}$$

It should be noted that not all points that satisfy Eq. (7B.19) correspond to actual breakaway points, if $K > 0$ (those that do not, satisfy $K < 0$ instead). Also, valid solutions must lie on the real-axis.

8. Imaginary Axis Crossing Points

The intersection of the root locus and the imaginary axis can be found by solving the characteristic equation whilst restricting solution points to $s = j\omega$. This is useful to analytically determine the value of *K* that causes the system to become unstable (or stable if the roots are entering from the right-half plane).

9. Effect of Poles and Zeros

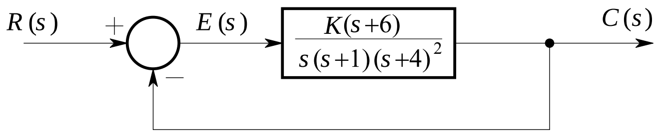
Zeros tend to “attract” the locus, while poles tend to “repel” it.

10. Use a computer

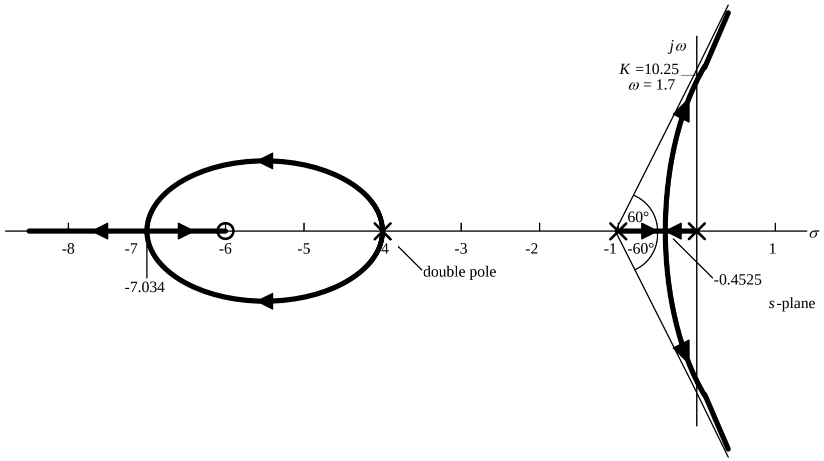
Use a computer to plot the root locus! The other rules provide intuition in shaping the root locus, and are also used to derive analytical quantities for the gain *K*, such as accurately evaluating stability.

Example

Consider a unity negative-feedback system:



This system has four poles (two being a double pole) and one zero, all on the negative real axis. In addition, it has three zeros at infinity. The root locus of this system, illustrated below, can be drawn on the basis of the rules presented.



Rule 1 There are four separate loci since the characteristic equation, $1+G(s)=0$, is a fourth-order equation.

Rule 2 The root locus starts ($K=0$) from the poles located at 0, -1, and a double pole located at -4. One pole terminates ($K=\infty$) at the zero located at -6 and three branches terminate at zeros which are located at infinity.

Rule 3 Complex portions of the root locus occur in complex-conjugate pairs.

Rule 4 The portions of the real axis between the origin and -1, the double poles at -4, and between -6 and $-\infty$ are part of the root locus.

Rule 5 The loci approach infinity as K becomes large at angles given by:

$$\phi_1 = \frac{1}{4-1}180° = 60°,$$

$$\phi_2 = \frac{3}{4-1}180° = 180°,$$

$$\phi_3 = \frac{5}{4-1}180° = 300°$$

Rule 6 The intersection of the asymptotic lines and the real axis occurs at:

$$\sigma_A = \frac{-9-(-6)}{4-1} = -1$$

Rule 7 The point of breakaway from the real axis is determined as follows. From the relation:

$$1+G(s)H(s) = 1+\frac{K(s+6)}{s(s+1)(s+4)^2} = 0$$

we have:

$$K = -\frac{s(s+1)(s+4)^2}{(s+6)}$$

Taking the derivative we get:

$$\frac{dK}{ds} = -\frac{(s+6)\left[2s(s+1)(s+4)+(2s+1)(s+4)^2\right]-s(s+1)(s+4)^2}{(s+6)^2} = 0$$

Therefore:

$$\begin{aligned}(s+6)[2s(s+1)+(2s+1)(s+4)]-s(s+1)(s+4)&=0\\ (s+6)\left[4s^2+11s+4\right]-s(s+1)(s+4)&=0\\ 3s^3+30s^2+66s+24&=0\\ s^3+10s^2+22s+8&=0\end{aligned}$$

Using MATLAB® with the command `roots([1 10 22 8])`, the roots of the equation are -7.034, -2.5135 and -0.4525. The root at -2.5135 is impossible for the negative feedback case, since the root locus doesn't lie there (-2.5135 is the breakaway point for a positive feedback system, i.e. when $K < 0$).

Rule 8 The intersection of the root locus and the imaginary axis can be determined by solving the characteristic equation when $s = j\omega$. The characteristic equation becomes:

$$
\begin{aligned}
s(s+1)(s+4)^2 + K(s+6) &= 0 \\
s^4 + 9s^3 + 24s^2 + (16+K)s + 6K &= 0 \\
\omega^4 - j9\omega^3 - 24\omega^2 + j(16+K)\omega + 6K &= 0
\end{aligned}
$$

Letting the real and imaginary parts go to zero, we have:

$$\omega^4 - 24\omega^2 + 6K = 0 \quad \text{and} \quad -9\omega^3 + (16+K)\omega = 0$$

Solving the second equation gives:

$$\omega = \frac{\sqrt{16+K}}{3}$$

Substituting this value into the first equation, we obtain:

$$\left(\frac{16+K}{9}\right)^2 - 24\left(\frac{16+K}{9}\right) + 6K = 0$$

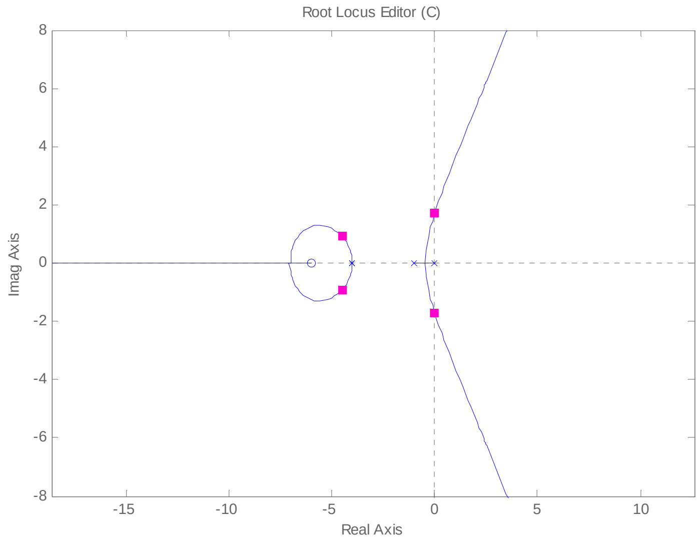
Solving this quadratic, we finally get $K = 10.2483$. Substituting this into the preceding equation, we obtain:

$$\omega = 1.7 \text{ rads}^{-1}$$

as the frequency of crossover.

Rule 9 follows by observation of the resulting sketch.

Rule 10 is shown below:



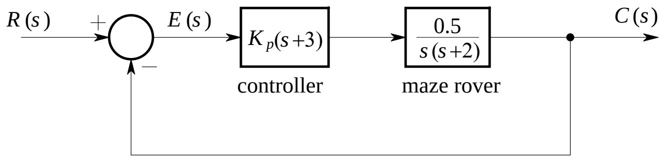
Thus, the computer solution matches the analytical solution, although the analytical solution provides more accuracy for points such as the crossover point and breakaway points; and it also provides insight into the system's behaviour.

MATLAB®'s RLTool

We can use MATLAB® to graph our root loci. From the command window, just type "rltool" to start the root locus user interface.

Example

The previous maze rover positioning scheme has a zero introduced in the controller:



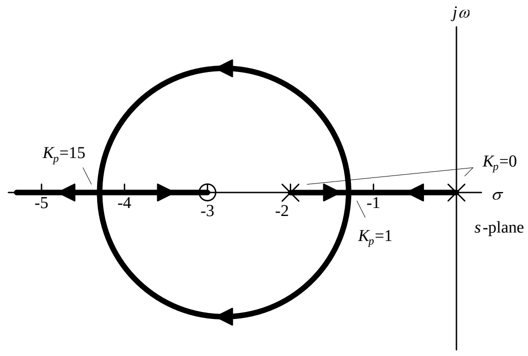
Using MATLAB®'s rltool, we would enter the following:

```
» Gp=tf(0.5, [1 2 0]);
» Gc=tf([1 3], 1);
» rltool
```

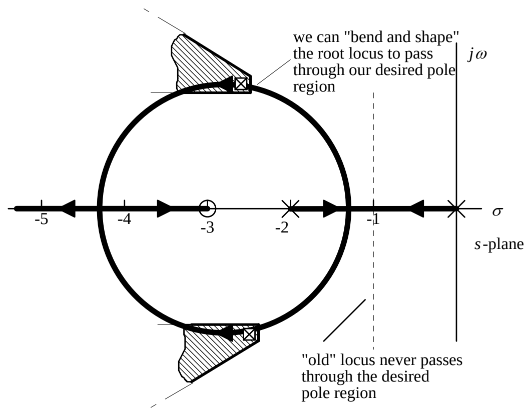
We choose "File|Import…" from the main menu. We place our transfer function G_p into the "G" position of MATLAB®'s control system model. We then place our G_c transfer function into the "C" position and press OK.

MATLAB® will draw the root locus and choose appropriate graphing quantities. The closed-loop poles are shown by red squares, and can be dragged by the mouse. The status line gives the closed-loop pole locations, and the gain K_p that was needed to put them there can be observed at the top of the user interface.

The locus for this case looks like:



We can see how the rules help us to "bend and shape" the root locus for our purposes. For example, we may want to increase the damping (move the poles to the left) which we have now done using a zero on the real axis. We couldn't do this for the case of a straight gain in the previous example:

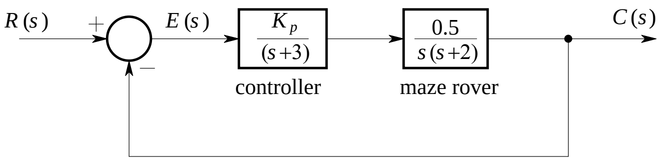


Root locus "bent and shaped" to pass through desired pole region

This system is no longer an "all-pole system", so any parameters such as rise time etc. that we were aiming for in the design must be checked by observing the step-response. This is done in MATLAB® by choosing "Analysis|Response to Step Command" from the main menu.

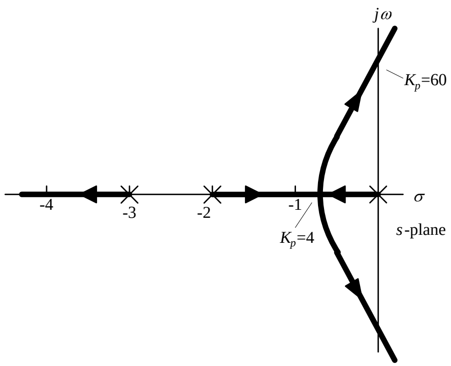
Example

We will now see what happens if we place a pole in the controller instead of a zero:



MATLAB® gives us the following root locus:

A root locus can clearly show the limits of gain for a stable system



We see that the pole "repels" the root locus. Also, unfortunately in this case, the root locus heads off into the RHP. If the parameter K_p is increased to over 60, then our system will be unstable!

Root Loci of Discrete-Time Systems

We can perform the same analysis for discrete-time systems, but of course the interpretation of the pole locations in the *z*-plane is different to the *s*-plane. The characteristic equation in this case is just:

Root locus applies to discrete-time systems also

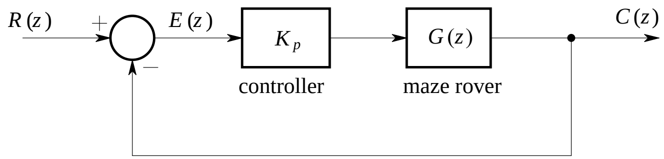
$$\boxed{1+G_c(z)G_p(z)=1+KP(z)=0} \qquad (7B.20)$$

Time Response of Discrete-Time Systems

We have already seen in Lecture 6B how to discretize a continuous-time system – there are methods such as the bilinear transform and step response matching. One parameter of extreme importance in discretization is the sample period, T_s. We need to choose it to be "sufficiently small" so that the discrete-time system approximates the continuous-time system closely. If we don't, then the pole and zero locations in the *z*-plane may lie out of the specification area, and in extreme cases can even lead to instability!

Example

We decide to simulate a maze rover by replacing it's continuous-time model with a discrete-time model found using the bilinear transform:



If the original maze rover continuous-time transfer function was:

$$G(s)=\frac{1}{s(s+1)}$$

then application of the bilinear transform gives:

$$G(z)=\frac{T_s^2(z+1)^2}{(4+2T_s)z^2-8z+4-2T_s}$$

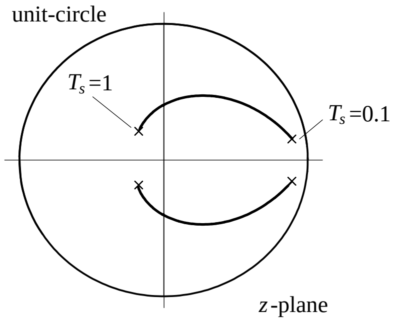
The closed-loop transfer function is now found to be:

$$T(z)=\frac{K_pT_s^2(z+1)^2}{\left(4+2T_s+T_s^2\right)z^2+\left(2T_s^2-8\right)z+4-2T_s+T_s^2}$$

Clearly, the pole locations depend upon the sample period T_s.

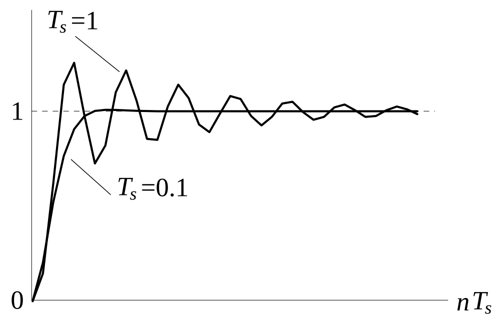
A locus of the closed-loop poles as T_s varies can be constructed. The figure below shows the root locus as T_s is varied from 0.1 s to 1 s:

Root locus for a discrete-time system as sample period is varied



The corresponding step-responses for the two extreme cases are shown below:

Step response due to two different sample periods



We can see that the sample time affects the control system's ability to respond to changes in the output. If the sample period is small relative to the time constants in the system, then the output will be a good approximation to the continuous-time case. If the sample period is much larger, then we inhibit the ability of the feedback to correct for errors at the output - causing oscillations, increased overshoot, and sometimes even instability.

Summary

- The root locus for a unity-feedback system is a graph of the closed-loop pole locations as a system parameter, such as controller gain, is varied from 0 to infinity.
- There are various rules for drawing the root locus that help us to analytically derive various values, such as gains that cause instability. A computer is normally used to graph the root locus, but understanding the root locus rules provides insight into the design of compensators in feedback systems.
- The root locus can tell us why and when a system becomes unstable.
- We can bend and shape a root locus by the addition of poles and zeros so that it passes through a desired location of the complex plane.
- Root locus techniques can be applied to discrete-time systems.
- The root locus of a discrete-time system as the sample period is varied gives us insight into how close an approximation we have of a continuous-time system, and whether the chosen sample period can meet the specifications.

References

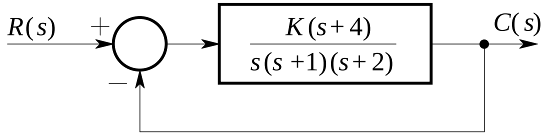
Kuo, B: *Automatic Control Systems*, Prentice-Hall, 1995.

Nicol, J.: *Circuits and Systems 2 Notes*, NSWIT, 1986.

Exercises

1.

For the system shown:



Use RLTool in MATLAB® for the following:

(a) Plot the root locus as K is varied from 0 to ∞.

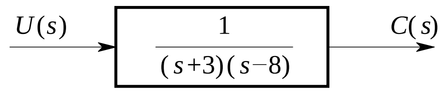
(b) Find the range of K for stability and the frequency of oscillation if unstable.

(c) Find the value of K for which the closed-loop system will have a 5% overshoot for a step input.

(d) Estimate the 1% settling time for the closed-loop pole locations found in part (c).

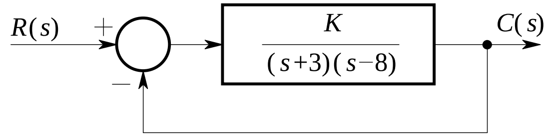
2.

The plant shown is open-loop unstable due to the right-half plane pole.

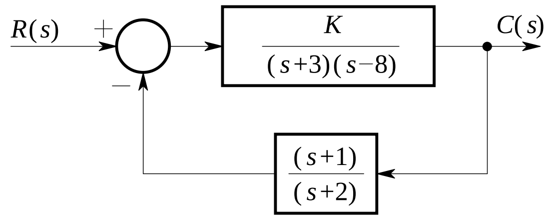


Use RLTool in MATLAB® for the following:

(a) Show by a plot of the root locus that the plant cannot be stabilized for any K, $-\infty < K < \infty$, if unity feedback is placed around it as shown.



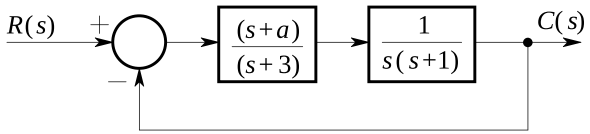
(b) An attempt is made to stabilize the plant using the feedback compensator shown:



Determine whether this design is successful by performing a root locus analysis for $0 < K < \infty$. (Explain, with the aid of a sketch, why $K < 0$ is not worth pursuing).

3.

For the system shown the required value of a is to be determined using the root-locus technique.



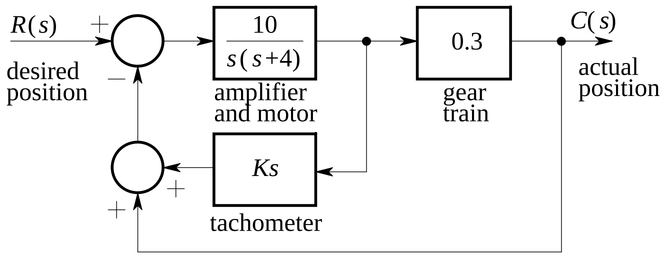
(a) Sketch the root-locus of $\frac{C(s)}{R(s)}$ as a varies from $-\infty$ to ∞.

(b) From the root-locus plot, find the value of a which gives both minimum overshoot and settling time when $r(t)$ is a step function.

(c) Find the maximum value of a which just gives instability and determine the frequency of oscillation for this value of a.

4.

The block diagram of a DC motor position control system is shown below.



The performance is adjusted by varying the tachometer gain K. K can vary from -100 to +100; 0 to +100 for the negative feedback configuration shown, and 0 to -100 if the electrical output connections from the tachometer are reversed (giving positive feedback).

(a) Sketch the root-locus of $\frac{C(s)}{R(s)}$ as K varies from $-\infty$ to ∞.

Use two plots: one for negative feedback and one for positive feedback. Find all important geometrical properties of the locus.

(b) Find the largest magnitude of K which just gives instability, and determine the frequency of oscillation of the system for this value of K.

(c) Find the steady-state error (as a function of K) when $r(t)$ is a step function.

(d) From the root locus plots, find the value of K which will give 10% overshoot when $r(t)$ is a step function, and determine the 10-90% rise time for this value of K.

<u>Note</u>: The closed-loop system has two poles (as found from the root locus) and <u>no zeros</u>. Verify this yourself using block diagram reduction.

James Clerk Maxwell (1831-1879)

Maxwell produced a most spectacular work of individual genius – he unified electricity and magnetism. Maxwell was able to summarize all observed phenomena of electrodynamics in a handful of partial differential equations known as *Maxwell's equations*:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{B} = \mu \mathbf{J} + \mu\varepsilon \frac{\partial \mathbf{E}}{\partial t}$$

From these he was able to predict that there should exist electromagnetic waves which could be transmitted through free space at the speed of light. The revolution in human affairs wrought by these equations and their experimental verification by Heinrich Hertz in 1888 is well known: wireless communication, control and measurement - so spectacularly demonstrated by television and radio transmissions across the globe, to the moon, and even to the edge of the solar system!

James Maxwell was born in Edinburgh, Scotland. His mother died when he was 8, but his childhood was something of a model for a future scientist. He was endowed with an exceptional memory, and had a fascination with mechanical toys which he retained all his life. At 14 he presented a paper to the Royal Society of Edinburgh on ovals. At 16 he attended the University of Edinburgh where the library still holds records of the books he borrowed while still an undergraduate – they include works by Cauchy on differential equations, Fourier on the theory of heat, Newton on optics, Poisson on mechanics and Taylor's scientific memoirs. In 1850 he moved to Trinity College, Cambridge, where he graduated with a degree in mathematics in 1854. Maxwell was edged out of first place in their final examinations by his classmate Edward Routh, who was also an excellent mathematician.

Maxwell stayed at Trinity where, in 1855, he formulated a "theory of three primary colour-perceptions" for the human perception of colour. In 1855 and 1856 he read papers to the Cambridge Philosophical Society "On Faraday's Lines of Force" in which he showed how a few relatively simple mathematical equations could express the behaviour of electric and magnetic fields.

In 1856 he became Professor of Natural Philosophy at Aberdeen, Scotland, and started to study the rings of Saturn. In 1857 he showed that stability could be achieved only if the rings consisted of numerous small solid particles, an explanation now confirmed by the Voyager spacecraft.

In 1860 Maxwell moved to King's College in London. In 1861 he created the first colour photograph – of a Scottish tartan ribbon – and was elected to the Royal Society. In 1862 he calculated that the speed of propagation of an electromagnetic wave is approximately that of the speed of light:

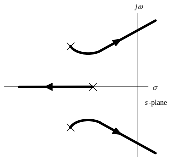
> *We can scarcely avoid the conclusion that light consists in the transverse undulations of the same medium which is the cause of electric and magnetic phenomena.*

Maxwell's famous account, "A Dynamical Theory of the Electromagnetic Field" was read before a largely perplexed Royal Society in 1864. Here he brought forth, for the first time, the equations which comprise the basic laws of electromagnetism.

All the mathematical sciences are founded on relations between physical laws and laws of numbers, so that the aim of exact science is to reduce the problems of nature to the determination of quantities by operations with numbers. – James Clerk Maxwell

Maxwell also continued work he had begun at Aberdeen, on the kinetic theory of gases (he had first considered the problem while studying the rings of Saturn). In 1866 he formulated, independently of Ludwig Boltzmann, the kinetic theory of gases, which showed that temperature and heat involved only molecular motion.

When creating his standard for electrical resistance, he wanted to design a governor to keep a coil spinning at a constant rate. He made the system stable by using the idea of negative feedback. It was known for some time that the governor was essentially a centrifugal pendulum, which sometimes exhibited "hunting" about a set point – that is, the governor would oscillate about an equilibrium position until limited in amplitude by the throttle valve or the travel allowed to the bobs. This problem was solved by Airy in 1840 by fitting a damping disc to the governor. It was then possible to minimize speed



fluctuations by adjusting the “controller gain”. But as the gain was increased, the governors would burst into oscillation again. In 1868, Maxwell published his paper “On Governors” in which he derived the equations of motion of engines fitted with governors of various types, damped in several ways, and explained in mathematical terms the source of the oscillation. He was also able to set bounds on the parameters of the system that would ensure stable operation. He posed the problem for more complicated control systems, but thought that a general solution was insoluble. It was left to Routh some years later to solve the general problem of linear system stability: “It has recently come to my attention that my good friend James Clerk Maxwell has had difficulty with a rather trivial problem…”.

In 1870 Maxwell published his textbook *Theory of Heat*. The following year he returned to Cambridge to be the first Cavendish Professor of Physics – he designed the Cavendish laboratory and helped set it up.

The four partial differential equations describing electromagnetism, now known as Maxwell’s equations, first appeared in fully developed form in his *Treatise on Electricity and Magnetism* in 1873. The significance of the work was not immediately grasped, mainly because an understanding of the atomic nature of electromagnetism was not yet at hand.

The Cavendish laboratory was opened in 1874, and Maxwell spent the next 5 years editing Henry Cavendish’s papers.

Maxwell died of the same disease, abdominal cancer, as his mother, in 1879, at the age of forty-eight. At his death, Maxwell’s reputation was uncertain. He was recognised to have been an exceptional scientist, but his theory of electromagnetism remained to be convincingly demonstrated. About 1880 Hermann von Helmholtz, an admirer of Maxwell, discussed the possibility of confirming his equations with a student, Heinrich Hertz. In 1888 Hertz performed a series of experiments which produced and measured electromagnetic waves and showed how they behaved like light. Thereafter, Maxwell’s reputation continued to grow, and he may be said to have prepared the way for twentieth-century physics.

Lecture 8A – State-Variables

State representation. Solution of the state equations. Transfer function. State-variable feedback.

Overview

The frequency-domain has dominated our analysis and design of signals and systems – up until now. Frequency-domain techniques are powerful tools, but they do have limitations. High-order systems are hard to analyse and design. Initial conditions are hard to incorporate into the analysis process (remember – the transfer function only gives the zero-state response). A time-domain approach, called the *state-space* approach, overcomes these deficiencies and also offers additional features of analysis and design that we have not yet considered.

State Representation

Consider the following simple electrical system:

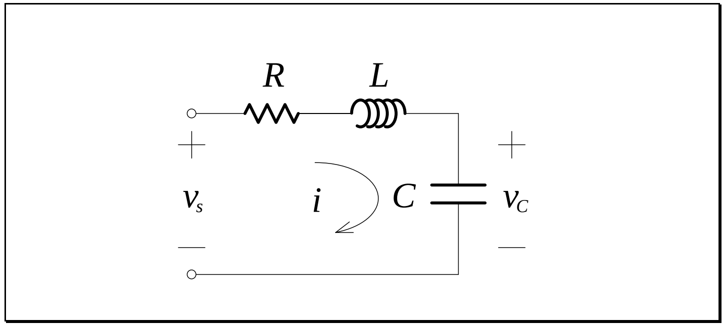


Figure 8A.1

In the analysis of a system via the state-space approach, the system is characterized by a set of first-order differential equations that describe its "state" variables. State variables are usually denoted by q_1, q_2, q_3, …, q_n. They characterize the future behaviour of a system once the inputs to a system are specified, together with a knowledge of the initial states.

States

For the system in Figure 8A.1, we can choose i and v_C as the state variables. Therefore, let:

State variables

$$q_1 = i \tag{8A.1a}$$

$$q_2 = v_C \tag{8A.1b}$$

From KVL, we get:

$$v_s = Ri + L\frac{di}{dt} + v_C \tag{8A.2}$$

Rearranging to get the derivative on the left-hand side, we get:

$$\frac{di}{dt} = -\frac{R}{L}i - \frac{1}{L}v_C + \frac{1}{L}v_s \tag{8A.3}$$

In terms of our state variables, given in Eqs. (8A.1), we can rewrite this as:

$$\frac{dq_1}{dt} = -\frac{R}{L}q_1 - \frac{1}{L}q_2 + \frac{1}{L}v_s \tag{8A.4}$$

Finally, we write Eq. (8A.4) in the standard nomenclature for state variable analysis – we use $\dot{q} = \frac{dq}{dt}$ and also let the input, v_s, be represented by the symbol x:

$$\dot{q}_1 = -\frac{R}{L}q_1 - \frac{1}{L}q_2 + \frac{1}{L}x \tag{8A.5}$$

Returning to the analysis of the circuit in Figure 8A.1, we have for the current through the capacitor:

$$i = C\frac{dv_C}{dt} \tag{8A.6}$$

Substituting our state variables, we have:

$$q_1 = C\dot{q}_2 \tag{8A.7}$$

Finally, rearranging to get the derivative on the left-hand side:

$$\dot{q}_2 = \frac{1}{C}q_1 \tag{8A.8}$$

Notice how, for a second-order system, we need to find two first-order differential equations to describe the system. The two equations can be written in matrix form:

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix} = \begin{bmatrix} -R/L & -1/L \\ 1/C & 0 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} + \begin{bmatrix} 1/L \\ 0 \end{bmatrix} x \tag{8A.9}$$

Using matrix symbols, this set of equations can be compactly written:

State equation

$$\boxed{\dot{\mathbf{q}} = \mathbf{A}\mathbf{q} + \mathbf{b}x} \tag{8A.10}$$

Matrix, vector and scalar notation

We will reserve small boldface letters for column vectors, such as $\mathbf{q}$ and $\mathbf{b}$ and capital boldface letters for matrices, such as $\mathbf{A}$. Scalar variables such as x are written in italics, as usual.

Eq. (8A.10) is *very* important – it tells us how the states of the system $\mathbf{q}$ change in time due to the input x.

Output

The system output can usually be expressed as a linear combination of all the state variables.

For example, if for the RLC circuit of Figure 8A.1 the output y is v_C then:

$$\begin{aligned} y &= v_C \\ &= q_2 \end{aligned} \tag{8A.11}$$

Therefore, in matrix notation, we write:

$$y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \tag{8A.12}$$

which is usually expressed as:

Output equation

$$\boxed{y = \mathbf{c}^T \mathbf{q}} \tag{8A.13}$$

Sometimes we also have:

$$\boxed{y = \mathbf{c}^T \mathbf{q} + dx} \tag{8A.14}$$

Multiple Input-Multiple Output Systems

State variable representation is good for multiple input – multiple output (MIMO) systems. All we have to do is generalise our input and output above to vector inputs and outputs:

Multiple input - multiple output (MIMO) state and output equations

$$\boxed{\begin{aligned} \dot{\mathbf{q}} &= \mathbf{A}\mathbf{q} + \mathbf{B}\mathbf{x} \\ \mathbf{y} &= \mathbf{C}^T \mathbf{q} + \mathbf{D}\mathbf{x} \end{aligned}} \tag{8A.15}$$

Solution of the State Equations

Once the state equations for a system have been obtained, it is usually necessary to find the output of a system for a given input (However, some parameters of the system can be directly determined by examining the **A** matrix, in which case we may not need to solve the state equations).

We can solve the state equations in the *s*-domain. Taking the Laplace Transform of Eq. (8A.10) gives:

$$s\mathbf{Q}(s) - \mathbf{q}(0^-) = \mathbf{A}\mathbf{Q}(s) + \mathbf{b}X(s) \tag{8A.16}$$

Notice how the initial conditions are automatically included by the Laplace transform of the derivative. The solution will be the *complete* response, not just the ZSR.

Making $\mathbf{Q}(s)$ the subject, we get:

$$\mathbf{Q}(s) = [s\mathbf{I} - \mathbf{A}]^{-1}\mathbf{q}(0^-) + [s\mathbf{I} - \mathbf{A}]^{-1}\mathbf{b}X(s) \tag{8A.17}$$

Because of its importance in state variable analysis, we define the following matrix:

Resolvent matrix defined

$$\boxed{\begin{aligned} \mathbf{\Phi}(s) &= [s\mathbf{I} - \mathbf{A}]^{-1} \\ &= \text{resolvent matrix} \end{aligned}} \tag{8A.18}$$

This simplifies Eq. (8A.17) to:

$$\mathbf{Q}(s) = \mathbf{\Phi}(s)\mathbf{q}(0^-) + \mathbf{\Phi}(s)\mathbf{b}X(s) \tag{8A.19}$$

Using Eq. (8A.13), the LT of the output (for $d = 0$) is then:

$$\mathbf{Y}(s) = \mathbf{c}^T\mathbf{\Phi}(s)\mathbf{q}(0^-) + \mathbf{c}^T\mathbf{\Phi}(s)\mathbf{b}X(s) \tag{8A.20}$$

All we have to do is take the inverse Laplace transform (ILT) to get the solution in the time-domain.

Before we do that, we also define the ILT of the resolvent matrix, called the transition matrix:

The transition matrix and resolvent matrix form a Laplace transform pair

$$\boxed{\begin{array}{ccc} \boldsymbol{\varphi}(t) & \leftrightarrow & \boldsymbol{\Phi}(s) \\ \text{transition} & & \text{resolvent} \\ \text{matrix} & & \text{matrix} \end{array}} \tag{8A.21}$$

The ILT of Eq. (8A.19) is just:

Complete solution of the state equation

$$\boxed{\begin{aligned} \mathbf{q}(t) &= \boldsymbol{\varphi}(t)\mathbf{q}(\mathbf{0}^-) + \int_0^t \boldsymbol{\varphi}(t-\tau)\mathbf{b}x(\tau)d\tau \\ &= \text{ZIR} + \text{ZSR} \end{aligned}} \tag{8A.22}$$

Notice how multiplication in the *s*-domain turned into convolution in the time-domain. The transition matrix is a generalisation of impulse response, but it applies to states – not the output!

We can get the output response of the system after solving for the states by direct substitution into Eq. (8A.14).

Transition Matrix

The transition matrix possesses two interesting properties that help it to be calculated by a digital computer:

$$\boldsymbol{\varphi}(0) = \mathbf{I} \tag{8A.23a}$$

$$\boldsymbol{\varphi}(t) = e^{\mathbf{A}t} \tag{8A.23b}$$

The first property is obvious by substituting $t = 0$ into Eq. (8A.22). The second relationship arises by observing that the solution to the state equation for the case of zero input, $\dot{\mathbf{q}} = \mathbf{A}\mathbf{q}$, is $\mathbf{q} = \mathbf{q}(0)e^{\mathbf{A}t}$. For zero input, Eq. (8A.22) gives

$\mathbf{q}(t) = \boldsymbol{\varphi}(t)\mathbf{q}(\mathbf{0}^-)$, so that we must have $\boldsymbol{\varphi}(t) = e^{\mathbf{A}t}$. The matrix $e^{\mathbf{A}t}$ is defined by:

How to raise *e* to a matrix power

$$e^{\mathbf{A}t} = \mathbf{I} + \frac{(\mathbf{A}t)}{1!} + \frac{(\mathbf{A}t)^2}{2!} + \frac{(\mathbf{A}t)^3}{3!} + \ldots \tag{8A.24}$$

This is easy to calculate on a digital computer, because it consists of matrix multiplication and addition. The series is truncated when the desired accuracy is reached.

Example

Suppose a system is described by the following differential equation:

$$\frac{d^2y}{dt^2} + 2\frac{dy}{dt} + y = \frac{dr}{dt} + r \tag{8A.25}$$

where the input and initial conditions are:

$$r = \sin t \quad y(0) = 1 \quad \dot{y}(0) = 0 \tag{8A.26}$$

Let:

$$q_1 = y, \quad q_2 = \dot{y}, \quad x = r + \dot{r} \tag{8A.27}$$

then:

$$\begin{aligned} \dot{q}_1 &= q_2 \\ \dot{q}_2 &= -q_1 - 2q_2 + x \end{aligned} \tag{8A.28}$$

or just:

$$\dot{\mathbf{q}} = \mathbf{A}\mathbf{q} + \mathbf{b}x \tag{8A.29}$$

with:

$$\mathbf{A}=\begin{bmatrix}0 & 1\\ -1 & -2\end{bmatrix},\quad \mathbf{b}=\begin{bmatrix}0\\1\end{bmatrix},\quad \mathbf{q}=\begin{bmatrix}q_1\\q_2\end{bmatrix},\quad \dot{\mathbf{q}}=\begin{bmatrix}\dot{q}_1\\ \dot{q}_2\end{bmatrix} \tag{8A.30}$$

We form the resolvent matrix by firstly finding $s\mathbf{I}-\mathbf{A}$:

$$s\mathbf{I}-\mathbf{A}=\begin{bmatrix}s & 0\\ 0 & s\end{bmatrix}-\begin{bmatrix}0 & 1\\ -1 & -2\end{bmatrix}=\begin{bmatrix}s & -1\\ 1 & s+2\end{bmatrix} \tag{8A.31}$$

Then remembering that the inverse of a matrix $\mathbf{B}$ is given by:

$$\mathbf{B}^{-1}=\frac{\operatorname{adj}\mathbf{B}}{|\mathbf{B}|} \tag{8A.32}$$

we get the resolvent matrix:

$$\mathbf{\Phi}(s)=[s\mathbf{I}-\mathbf{A}]^{-1}=\frac{\begin{bmatrix}s+2 & 1\\ -1 & s\end{bmatrix}}{(s+1)^2}=\begin{bmatrix}\frac{s+2}{(s+1)^2} & \frac{1}{(s+1)^2}\\ \frac{-1}{(s+1)^2} & \frac{s}{(s+1)^2}\end{bmatrix} \tag{8A.33}$$

The transition matrix is the inverse Laplace transform of the resolvent matrix:

$$\begin{aligned}\boldsymbol{\varphi}(t)&=L^{-1}\{\mathbf{\Phi}(s)\}\\ &=\begin{bmatrix}e^{-t}(t+1) & te^{-t}\\ -te^{-t} & e^{-t}(1-t)\end{bmatrix}\end{aligned} \tag{8A.34}$$

So, from Eq. (8A.22), the ZIR is given by:

$$\begin{aligned}\mathbf{q}_{\text{ZIR}}(t)&=\boldsymbol{\varphi}(t)\mathbf{q}(0^-)\\ \begin{bmatrix}q_1\\q_2\end{bmatrix}_{\text{ZIR}}&=\begin{bmatrix}e^{-t}(t+1) & te^{-t}\\ -te^{-t} & e^{-t}(1-t)\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}e^{-t}(t+1)\\ -te^{-t}\end{bmatrix}\end{aligned} \tag{8A.35}$$

Since we don't like performing convolution in the time-domain, we use Eq. (8A.19) to find the ZSR:

$$\mathbf{q}_{\text{ZSR}}(t)=L^{-1}\{\mathbf{\Phi}(s)\mathbf{b}X(s)\} \tag{8A.36}$$

The Laplace transform of the input is:

$$X(s)=\frac{1}{s^2+1}+\frac{s}{s^2+1} \tag{8A.37}$$

so the ZSR is:

$$\begin{aligned}\mathbf{q}_{\text{ZSR}}(t)&=L^{-1}\left\{\begin{bmatrix}\frac{s+2}{(s+1)^2} & \frac{1}{(s+1)^2}\\ \frac{-1}{(s+1)^2} & \frac{s}{(s+1)^2}\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix}\frac{s+1}{s^2+1}\right\}\\ &=L^{-1}\left\{\begin{bmatrix}\frac{1}{(s+1)(s^2+1)}\\ \frac{s}{(s+1)(s^2+1)}\end{bmatrix}\right\}\end{aligned} \tag{8A.38}$$

Use partial fractions to get:

$$\begin{bmatrix}q_1\\q_2\end{bmatrix}_{\text{ZSR}}=\begin{bmatrix}\frac{1}{2}\left(e^{-t}-\cos t+\sin t\right)\\ \frac{1}{2}\left(-e^{-t}+\cos t+\sin t\right)\end{bmatrix} \tag{8A.39}$$

The total response is the sum of the ZIR and the ZSR:

$$\begin{bmatrix}q_1\\q_2\end{bmatrix}=\begin{bmatrix}\frac{3}{2}e^{-t}+te^{-t}-\frac{1}{2}\cos t+\frac{1}{2}\sin t\\ -\frac{1}{2}e^{-t}-te^{-t}+\frac{1}{2}\cos t+\frac{1}{2}\sin t\end{bmatrix} \tag{8A.40}$$

This is just the solution for the states. To get the output, we use:

$$\begin{aligned}y&=\mathbf{c}^T\mathbf{q}+dx\\ \mathbf{c}^T&=[1\quad 0],\quad d=0\end{aligned} \tag{8A.41}$$

Therefore, the output is:

$$y=\tfrac{3}{2}e^{-t}+te^{-t}-\tfrac{1}{2}\cos t+\tfrac{1}{2}\sin t \tag{8A.42}$$

You should confirm this solution by solving the differential equation directly using your previous mathematical knowledge, eg. method of undetermined coefficients.

Transfer Function

The transfer function of a single input-single output (SISO) system can be obtained easily from the state variable equations. Since a transfer function only gives the ZSR (all initial conditions are zero), then Eq. (8A.19) becomes:

$$\mathbf{Q}(s) = \mathbf{\Phi}(s)\mathbf{b}X(s) \tag{8A.43}$$

The output in the s-domain, using the Laplace transform of Eq. (8A.13) and Eq. (8A.43), is just:

$$\begin{aligned} Y(s) &= \mathbf{c}^{\mathbf{T}}\mathbf{Q}(s) \\ &= \mathbf{c}^{\mathbf{T}}\mathbf{\Phi}(s)\mathbf{b}X(s) \end{aligned} \tag{8A.44}$$

Therefore, the transfer function (for $d = 0$) is given by:

Obtaining the transfer function from a state-variable description of a system

$$\boxed{H(s) = \mathbf{c}^{T}\mathbf{\Phi}(s)\mathbf{b}} \tag{8A.45}$$

Impulse Response

The impulse response is just the inverse Laplace transform of the transfer function:

$$\begin{aligned} h(t) &= \mathbf{c}^{T}\boldsymbol{\varphi}(t)\mathbf{b} \\ &= \mathbf{c}^{T}e^{\mathbf{A}t}\mathbf{b} \end{aligned} \tag{8A.46}$$

It is possible to compute the impulse response directly from the coefficient matrices of the state model of the system.

Example

Continuing the analysis of the system used in the previous example, we can find the transfer function using Eq. (8A.45):

$$\begin{aligned} H(s) &= \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix} \frac{s+2}{(s+1)^2} & \frac{1}{(s+1)^2} \\ \frac{-1}{(s+1)^2} & \frac{s}{(s+1)^2} \end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix} \\ &= \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix} \frac{1}{(s+1)^2} \\ \frac{s}{(s+1)^2} \end{bmatrix} \\ &= \frac{1}{(s+1)^2} \end{aligned} \tag{8A.47}$$

Compare with the Laplace transform of the original differential equation:

$$\begin{aligned} &\frac{d^2y}{dt^2} + 2\frac{dy}{dt} + y = x \\ &(s^2+2s+1)Y(s) = X(s) \end{aligned} \tag{8A.48}$$

from which the transfer function is:

$$\begin{aligned} H(s) &= \frac{Y(s)}{X(s)} \\ &= \frac{1}{(s^2+2s+1)} \\ &= \frac{1}{(s+1)^2} \end{aligned} \tag{8A.49}$$

Why would we use the state-variable approach to obtain the transfer function? For a simple system, we probably wouldn’t, but for multiple-input multiple-output systems, it is much easier using the state-variable approach.

Linear State-Variable Feedback

Block diagram of linear state-variable feedback

Consider the following system drawn using a state-variable approach:

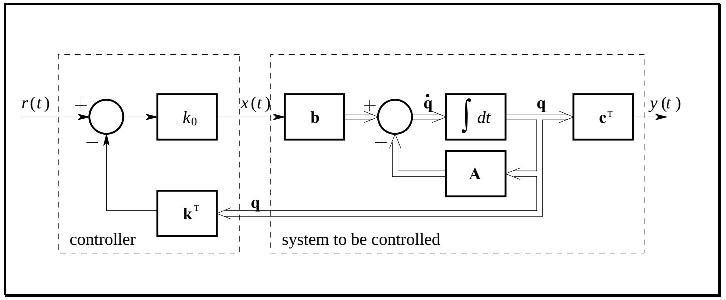


Figure 8A.2

The system has been characterised in terms of states – you should confirm that the above diagram of the "system to be controlled" is equivalent to the matrix formulation of Eqs. (8A.10) and (8A.13).

We have placed a controller in front of the system, and we desire the output y to follow the set point, or reference input, r. The design of the controller involves the determination of the controller variables k_0 and $\mathbf{k}$ to achieve a desired response from the system (The desired response could be a time-domain specification, such as rise time, or a frequency specification, such as bandwidth).

The controller just multiples each of the states q_i by a gain k_i, subtracts the sum of these from the input r, and multiplies the result by a gain k_0.

Now, the input x to the controlled system is:

The input to the open-loop system is modified by the feedback

$$x = k_0\left(r - \mathbf{k}^T\mathbf{q}\right) \tag{8A.50}$$

Therefore, the state equations are:

$$\dot{\mathbf{q}} = \mathbf{A}_k\mathbf{q} + \mathbf{b}k_0 r \tag{8A.51}$$

where:

State-variable feedback modifies the *A* matrix

$$\mathbf{A}_k = \mathbf{A} - k_0\mathbf{b}\mathbf{k}^T \tag{8A.52}$$

Eq. (8A.51) is the state-variable representation of the *overall* system (controller plus system to be controlled). The state equations Eq. (8A.51) still have the same form as Eq. (8A.10), but $\mathbf{A}$ changes to $\mathbf{A}_k$ and the input changes from x to $k_0 r$. Analysis of the *overall* system can now proceed as follows.

For the ZSR, the transfer function is given by Eq. (8A.45) with the above substitutions:

The closed-loop transfer function when linear state-variable feedback is applied

$$H(s) = k_0\mathbf{c}^T\mathbf{\Phi}_k(s)\mathbf{b} \tag{8A.53}$$

where:

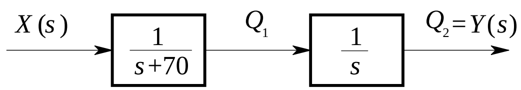
The modified resolvent matrix when linear state-variable feedback is applied

$$\mathbf{\Phi}_k(s) = \left[s\mathbf{I} - \mathbf{A}_k\right]^{-1} \tag{8A.54}$$

We choose the controller variables k_0 and $\mathbf{k}^T = \begin{bmatrix} k_1 & k_2 & \ldots & k_n \end{bmatrix}$ to create the transfer function obtained from the design criteria (easy for an $n = 2$ second-order system).

Example

Suppose that it is desired to control an open-loop process using state-variable control techniques. The open-loop system is shown below:



Suppose it is desired that the closed-loop second-order characteristics of the feedback control system have the following parameters:

$$\omega_n = 50\,\text{rads}^{-1}, \quad \zeta = 0.7071 \tag{8A.55}$$

Then the desired transfer function is:

$$\begin{aligned}\frac{C(s)}{R(s)} &= \frac{\omega_n^{\ 2}}{s^2 + 2\zeta\omega_n s + \omega_n^{\ 2}} \\ &= \frac{2500}{s^2 + 70.71s + 2500}\end{aligned} \tag{8A.56}$$

The first step is to formulate a state-variable representation of the system. The relationships for each state-variable and the output variable in the s-domain are obtained directly from the block diagram:

$$\begin{aligned}Q_1 &= \frac{1}{s+70}X \\ Q_2 &= \frac{1}{s}Q_1 \\ Y &= Q_2\end{aligned} \tag{8A.57}$$

Rearranging, we get:

$$\begin{aligned}sQ_1 &= -70Q_1 + X \\ sQ_2 &= Q_1 \\ Y &= Q_2\end{aligned} \tag{8A.58}$$

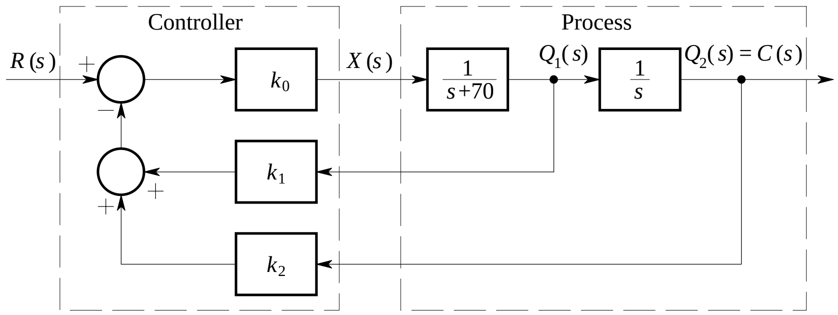
The corresponding state-variable representation is readily found to be:

$$\begin{aligned}\dot{\mathbf{q}} &= \begin{bmatrix} -70 & 0 \\ 1 & 0 \end{bmatrix}\mathbf{q} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}x \\ y &= \begin{bmatrix} 0 & 1 \end{bmatrix}\mathbf{q}\end{aligned} \tag{8A.59}$$

or just:

$$\begin{aligned}\dot{\mathbf{q}} &= \mathbf{A}\mathbf{q} + \mathbf{b}x \\ y &= \mathbf{c}^T\mathbf{q}\end{aligned} \tag{8A.60}$$

All linear, time-invariant systems have this state-variable representation. To implement state-variable feedback, we form the following system:



We see that the controller accepts a linear combination of the states, and compares this with the reference input. It then provides gain and applies the resulting signal as the “control effort”, $X(s)$, to the process.

The input signal to the process is therefore:

$$x = k_0\left(r - k_1q_1 - k_2q_2\right) \tag{8A.61}$$

or in matrix notation:

$$x = k_0\left(r - \mathbf{k}^T\mathbf{q}\right) \tag{8A.62}$$

Applying this as the input to the system changes the describing state equation of Eq. (8A.60) to:

$$\begin{aligned}\dot{\mathbf{q}} &= \mathbf{Aq} + \mathbf{b}\left[k_0\left(r - \mathbf{k}^T\mathbf{q}\right)\right] \\ \dot{\mathbf{q}} &= \left(\mathbf{A} - k_0\mathbf{bk}^T\right)\mathbf{q} + \mathbf{b}k_0 r \\ \dot{\mathbf{q}} &= \mathbf{A}_k\mathbf{q} + \mathbf{b}k_0 r\end{aligned} \tag{8A.63}$$

where:

$$\mathbf{A}_k = \mathbf{A} - k_0\mathbf{bk}^T \tag{8A.64}$$

If we let:

$$\mathbf{\Phi}_k(s) = \left[s\mathbf{I} - \mathbf{A}_k\right]^{-1} \tag{8A.65}$$

then the transfer function of the closed-loop system is given by:

$$H(s) = k_0\mathbf{c}^T\mathbf{\Phi}_k(s)\mathbf{b} \tag{8A.66}$$

We now need to evaluate a few matrices. First:

$$\begin{aligned}\mathbf{A}_k &= \mathbf{A} - k_0\mathbf{bk}^T \\ &= \begin{bmatrix}-70 & 0 \\ 1 & 0\end{bmatrix} - k_0\begin{bmatrix}1 \\ 0\end{bmatrix}\begin{bmatrix}k_1 & k_2\end{bmatrix} \\ &= \begin{bmatrix}-70 & 0 \\ 1 & 0\end{bmatrix} - k_0\begin{bmatrix}k_1 & k_2 \\ 0 & 0\end{bmatrix} \\ &= \begin{bmatrix}-70 - k_0k_1 & -k_0k_2 \\ 1 & 0\end{bmatrix}\end{aligned} \tag{8A.67}$$

Then:

$$\begin{aligned}\mathbf{\Phi}_k(s) &= \left[s\mathbf{I} - \mathbf{A}_k\right]^{-1} \\ &= \begin{bmatrix}s + 70 + k_0k_1 & k_0k_2 \\ -1 & s\end{bmatrix}^{-1} \\ &= \frac{1}{s^2 + (70 + k_0k_1)s + k_0k_2}\begin{bmatrix}s & -k_0k_2 \\ 1 & s + 70 + k_0k_2s\end{bmatrix}\end{aligned} \tag{8A.68}$$

The closed-loop transfer function is then found as:

$$\begin{aligned}H(s) &= k_0\mathbf{c}^T\mathbf{\Phi}_k(s)\mathbf{b} \\ &= \frac{k_0}{s^2 + (70 + k_0k_1)s + k_0k_2}\begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}s & -k_0k_2 \\ 1 & s + 70 + k_0k_1s\end{bmatrix}\begin{bmatrix}1 \\ 0\end{bmatrix} \\ &= \frac{k_0}{s^2 + (70 + k_0k_1)s + k_0k_2}\begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}s \\ 1\end{bmatrix} \\ &= \frac{k_0}{s^2 + (70 + k_0k_1)s + k_0k_2}\end{aligned} \tag{8A.69}$$

The values of k_0, k_1 and k_2 can be found from Eqs. (8A.56) and (8A.69). The following set of simultaneous equations result:

$$\begin{aligned}k_0 &= 2500 \\ k_0k_2 &= 2500 \\ 70 + k_0k_1 &= 70.71\end{aligned} \tag{8A.70}$$

We have three equations and three unknowns. Solving, we find that:

$$\begin{aligned}k_0 &= 2500 \\ k_1 &= 2.843 \times 10^{-4} \\ k_2 &= 1\end{aligned} \tag{8A.71}$$

This completes the controller design. The final step would be to draw the root locus and examine the relative stability, and the sensitivity of slight gain variations. For this simple system, the final step is not necessary.

Summary

- State-variables describe "internal states" of a system rather than just the input-output relationship.
- The state-variable approach is a time-domain approach. We can include initial conditions in the analysis of a system to obtain the *complete* response.
- The state-variable equations can be solved using Laplace transform techniques.
- We can derive the transfer function of a SISO system using state-variables.
- Linear state-variable feedback involves the design of gains for each of the states, plus the input.

References

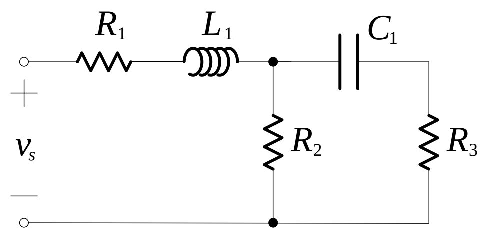
Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB*®, Prentice-Hall, 1997.

Shinners, S.: *Modern Control System Theory and Application*, Addison-Wesley, 1978.

Exercises

1.

Using capacitor voltages and inductor currents write a state-variable representation for the following circuit:



2.

Consider a linear system with input u and output y. Three experiments are performed on this system using the inputs $x_1(t)$, $x_2(t)$ and $x_3(t)$ for $t \geq 0$. In each case, the initial state at $t = 0$, $\mathbf{x}(0)$, is the same. The corresponding observed outputs are $y_1(t)$, $y_2(t)$ and $y_3(t)$. Which of the following three predictions are true if $\mathbf{x}(0) \neq \mathbf{0}$?

(a) If $x_3 = x_1 + x_2$, then $y_3 = y_1 + y_2$.

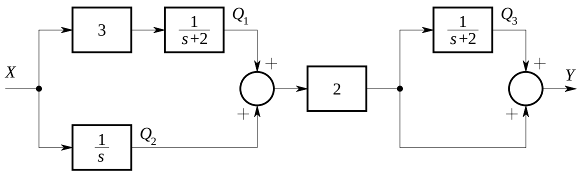
(b) If $x_3 = \frac{1}{2}(x_1 + x_2)$, then $y_3 = \frac{1}{2}(y_1 + y_2)$.

(c) If $x_3 = x_1 - x_2$, then $y_3 = y_1 - y_2$.

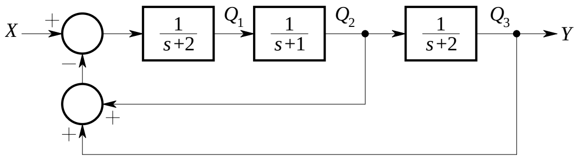
Which are true if $\mathbf{x}(0) = \mathbf{0}$?

3.

Write dynamical equation descriptions for the block diagrams shown below *with the chosen state variables*.



(a)



(b)

4.

Find the transfer function of the systems in Q3:

(i) by block diagram reduction

(ii) directly from your answers in Q3 (use the resolvent matrix etc)

5.

Given:

$$\dot{\mathbf{q}} = \begin{bmatrix} -1 & -1 \\ 4 & 1 \end{bmatrix}\mathbf{q} + \begin{bmatrix} 1 \\ 2 \end{bmatrix}x$$

and:

$$\mathbf{q}(0) = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \quad x(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

Find $\mathbf{q}(t)$.

6.

Write a simple MATLAB® script to evaluate the time response of the system described in state equation form in Q5 using the approximate relationship:

$$\dot{\mathbf{q}} = \frac{\mathbf{q}(t+T) - \mathbf{q}(t)}{T}$$

Use this script to plot $\mathbf{q}(t)$ using:

(a) $T = 0.01\,\mathrm{s}$, (b) $T = 0.1\,\mathrm{s}$, (c) $T = 1\,\mathrm{s}$

Compare these plots to the values given by the *exact* solution to Q5 (obtained by finding the inverse Laplace transforms of the answer given).

Comment on the effect of varying the time increment *T*.

Lecture 8B – State-Variables 2

Normal form. Similarity transform. Poles and eigenvalues. Solution of the state equations for ZIR. Discrete-time state-variables. Discrete-time response. Discrete-time transfer function.

Overview

State-variable analysis is useful for high-order systems and multiple-input multiple-output systems. It can also be used to find transfer functions between any output and any input of a system. It also gives us the complete response. The only drawback to all this analytical power is that solving the state-variable equations for high-order systems is difficult to do symbolically. Any computer solution also has to be thought about in terms of processing time and memory storage requirements.

The use of eigenvectors and eigenvalues solves this problem.

State-variable analysis can also be extended to discrete-time systems, producing exactly analogous equations as for continuous-time systems.

Normal Form

Solving matrix equations is hard...unless we have a “trivial” system:

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \vdots \\ \dot{z}_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \tag{8B.1}$$

which we would just write as:

$$\dot{\mathbf{z}} = \boldsymbol{\Lambda}\mathbf{z} \tag{8B.2}$$

This is useful for the ZIR of a state-space representation where:

$$\dot{\mathbf{q}} = \mathbf{A}\mathbf{q} \tag{8B.3}$$

We therefore seek a "similarity transform" to turn $\dot{\mathbf{q}} = \mathbf{A}\mathbf{q}$ into $\dot{\mathbf{z}} = \mathbf{\Lambda}\mathbf{z}$.

When solving differential equations, we know we should get solutions containing exponentials. Let us therefore try one possible "exponential trial solution" in Eq. (8B.3) by letting:

$$\mathbf{q} = \mathbf{q}(0)e^{\lambda t} \tag{8B.4}$$

where $\mathbf{q}(0)$ is a constant vector determined by the initial conditions of the states. Then $\dot{\mathbf{q}} = \lambda\mathbf{q}$ and, substituting into Eq. (8B.3), we get:

$$\mathbf{A}\mathbf{q} = \lambda\mathbf{q} \tag{8B.5}$$

Therefore:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{q} = \mathbf{0} \tag{8B.6}$$

For a non-trivial solution (one where $\mathbf{q} \neq \mathbf{0}$), we need to have:

Characteristic equation

$$\boxed{|\mathbf{A} - \lambda\mathbf{I}| = 0} \tag{8B.7}$$

This is called the *characteristic equation* of the system. The *eigenvalues* are the values of λ which satisfy $|\mathbf{A} - \lambda\mathbf{I}| = \mathbf{0}$. Once we have all the λ's, each column vector $\mathbf{q}_{(i)}$ which satisfies the original equation Eq. (8B.5) is called a *column eigenvector*.

Eigenvalues and eigenvectors defined

$$\boxed{\begin{array}{ll} |\mathbf{A} - \lambda\mathbf{I}| = 0 & \lambda_i = \text{eigenvalue} \\ \mathbf{A}\mathbf{q}_{(i)} = \lambda_i\mathbf{q}_{(i)} & \mathbf{q}_{(i)} = \text{eigenvector} \end{array}} \tag{8B.8}$$

An eigenvector corresponding to an eigenvalue is not unique – an eigenvector can be multiplied by any non-zero arbitrary constant. We therefore tend to choose the simplest eigenvectors to make the mathematics easy.

Example

Given a system's **A** matrix, we want to find the eigenvalues and eigenvectors.

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 2 \\ 10 & 3 & 4 \\ 3 & 6 & 1 \end{bmatrix} \tag{8B.9}$$

We find the eigenvalues first by solving:

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} (2-\lambda) & 3 & 2 \\ 10 & (3-\lambda) & 4 \\ 3 & 6 & (1-\lambda) \end{vmatrix} = 0 \tag{8B.10}$$

Evaluating the determinant, we get the characteristic equation:

$$\lambda^3 - 6\lambda^2 - 49\lambda - 66 = 0 \tag{8B.11}$$

Factorising the characteristic equation, we get:

$$(\lambda + 2)(\lambda + 3)(\lambda - 11) = 0 \tag{8B.12}$$

The solutions to the characteristic equation are:

$$\left.\begin{array}{l} \lambda_1 = -2 \\ \lambda_2 = -3 \\ \lambda_3 = 11 \end{array}\right\} \text{ eigenvalues} \tag{8B.13}$$

To find the eigenvectors, substitute each λ into $(\mathbf{A}-\lambda\mathbf{I})\mathbf{q}=\mathbf{0}$ and solve for $\mathbf{q}$. Take $\lambda_1=-2$:

$$\left\{\begin{bmatrix}2&3&2\\10&3&4\\3&6&1\end{bmatrix}-\begin{bmatrix}-2&0&0\\0&-2&0\\0&0&-2\end{bmatrix}\right\}\begin{bmatrix}q_1\\q_2\\q_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$
$$\begin{bmatrix}4&3&2\\10&5&4\\3&6&3\end{bmatrix}\begin{bmatrix}q_1\\q_2\\q_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix} \tag{8B.14}$$

Solve to get:

$$\mathbf{q}_{(1)}=\begin{bmatrix}1\\2\\-5\end{bmatrix}\quad \text{for } \lambda_1=-2 \tag{8B.15}$$

The other eigenvectors are:

$$\mathbf{q}_{(2)}=\begin{bmatrix}0\\2\\-3\end{bmatrix}\quad \text{for } \lambda_2=-3 \quad\text{and}\quad \mathbf{q}_{(3)}=\begin{bmatrix}2\\4\\3\end{bmatrix}\quad \text{for } \lambda_3=11 \tag{8B.16}$$

Similarity Transform

Similarity transform defined

The eigenvalues and eigenvectors that arise from Eq. (8B.5) are put to good use by transforming $\dot{\mathbf{q}}=\mathbf{A}\mathbf{q}$ into $\dot{\mathbf{z}}=\mathbf{\Lambda}\mathbf{z}$. First, form the square $n\times n$ matrix:

$$\mathbf{U}=\begin{bmatrix}\mathbf{q}_{(1)} & \mathbf{q}_{(2)} & \cdots & \mathbf{q}_{(n)}\end{bmatrix}=\begin{bmatrix}\begin{bmatrix}q_{(1)1}\\q_{(1)2}\\\vdots\\q_{(1)n}\end{bmatrix} & \begin{bmatrix}q_{(2)1}\\q_{(2)2}\\\vdots\\q_{(2)n}\end{bmatrix} & \cdots & \begin{bmatrix}q_{(n)1}\\q_{(n)2}\\\vdots\\q_{(n)n}\end{bmatrix}\end{bmatrix} \tag{8B.17}$$

The columns of the $\mathbf{U}$ matrix are the column eigenvectors corresponding to each of the n eigenvalues. Then since the columns of $\mathbf{U}$ are solutions to:

$$\mathbf{A}\mathbf{q}_{(i)}=\lambda_i\mathbf{q}_{(i)} \tag{8B.18}$$

then by some simple matrix manipulation, we get:

$$\mathbf{AU}=\mathbf{U\Lambda} \tag{8B.19}$$

where:

Diagonal matrix Λ defined

$$\mathbf{\Lambda}=\begin{bmatrix}\lambda_1&0&0&0\\0&\lambda_2&0&0\\0&0&\ddots&0\\0&0&0&\lambda_n\end{bmatrix} \tag{8B.20}$$

Example

From the previous example, we can confirm the following relationship.

$$\mathbf{AU}=\mathbf{U\Lambda}$$
$$\begin{bmatrix}2&3&2\\10&3&4\\3&6&1\end{bmatrix}\begin{bmatrix}1&0&2\\2&2&4\\-5&-3&3\end{bmatrix}=\begin{bmatrix}1&0&2\\2&2&4\\-5&-3&3\end{bmatrix}\begin{bmatrix}-2&0&0\\0&-3&0\\0&0&11\end{bmatrix} \tag{8B.21}$$

Perform the multiplication shown to verify the result.

The matrix $\mathbf{U}$ is called the right-hand eigenvector matrix. If, instead of Eq. (8B.5), we were to solve the equation $\mathbf{q}^T\mathbf{A}=\lambda\mathbf{q}^T$ we get the same eigenvalues but different eigenvectors. Also, we can form an $n\times n$ matrix $\mathbf{V}$ so that:

$$\mathbf{VA}=\mathbf{\Lambda V} \tag{8B.22}$$

where $\mathbf{V}$ is made up of *rows* of *left-hand* eigenvectors:

$$\mathbf{V}=\begin{bmatrix}\mathbf{q}^{T(1)}\\\mathbf{q}^{T(2)}\\\vdots\\\mathbf{q}^{T(n)}\end{bmatrix}=\begin{bmatrix}\begin{bmatrix}q^{(1)}{}_1 & q^{(1)}{}_2 & \cdots & q^{(1)}{}_n\end{bmatrix}\\\begin{bmatrix}q^{(2)}{}_1 & q^{(2)}{}_2 & \cdots & q^{(2)}{}_n\end{bmatrix}\\\vdots\\\begin{bmatrix}q^{(n)}{}_1 & q^{(n)}{}_2 & \cdots & q^{(n)}{}_n\end{bmatrix}\end{bmatrix} \tag{8B.23}$$

Since eigenvectors can be arbitrarily scaled by any non-zero constant, it can be shown that we can choose $\mathbf{V}$ such that:

$$\mathbf{VU} = \mathbf{I} \tag{8B.24}$$

which implies:

Relationship between the two similarity transforms

$$\boxed{\mathbf{V} = \mathbf{U}^{-1}} \tag{8B.25}$$

Now, starting from Eq. (8B.19), pre-multiply by $\mathbf{V}$:

$$\begin{aligned}\mathbf{AU} &= \mathbf{U\Lambda} \\ \mathbf{VAU} &= \mathbf{VU\Lambda}\end{aligned} \tag{8B.26}$$

but using Eq. (8B.25), this turns into:

The similarity transform diagonalizes a matrix

$$\boxed{\mathbf{\Lambda} = \mathbf{U}^{-1}\mathbf{AU}} \tag{8B.27}$$

Solution of the State Equations for the ZIR

Given:

$$\dot{\mathbf{q}} = \mathbf{Aq} \tag{8B.28}$$

and $\mathbf{q}(0^-)$, we would like to determine the ZIR $\mathbf{q}_{\mathrm{ZIR}}(t)$.

First, put $\mathbf{q} = \mathbf{Uz}$ and substitute into Eq. (8B.28) to give:

$$\mathbf{U}\dot{\mathbf{z}} = \mathbf{AUz} \tag{8B.29}$$

Now pre-multiply by $\mathbf{U}^{-1}$:

$$\dot{\mathbf{z}} = \mathbf{U}^{-1}\mathbf{AUz} \tag{8B.30}$$

and using Eq. (8B.27), the end result of the change of variable is:

Diagonal form of state equations

$$\boxed{\dot{\mathbf{z}} = \mathbf{\Lambda z}} \tag{8B.31}$$

Written explicitly, this is:

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \vdots \\ \dot{z}_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \tag{8B.32}$$

This is just a set of n independent first-order differential equations:

$$\begin{aligned}\dot{z}_1 &= \lambda_1 z_1 \\ \dot{z}_2 &= \lambda_2 z_2 \\ &\vdots \\ \dot{z}_n &= \lambda_n z_n\end{aligned} \tag{8B.33}$$

Consider the first equation:

$$\frac{dz_1}{dt} = \lambda_1 z_1 \tag{8B.34}$$

The solution is easily seen to be:

$$z_1 = z_1(0)e^{\lambda_1 t} \tag{8B.35}$$

The solution to Eq. (8B.33) is therefore just:

$$\begin{aligned} z_1 &= z_1(0)e^{\lambda_1 t} \\ z_2 &= z_2(0)e^{\lambda_2 t} \\ &\vdots \\ z_n &= z_n(0)e^{\lambda_n t} \end{aligned} \tag{8B.36}$$

or written using matrix notation:

Solution to the diagonal form of state equations

$$\boxed{\mathbf{z} = e^{\mathbf{\Lambda}t}\mathbf{z}(0)} \tag{8B.37}$$

The matrix $e^{\mathbf{\Lambda}t}$ is defined by:

The matrix $e^{\mathbf{\Lambda}t}$ defined

$$\begin{aligned} e^{\mathbf{\Lambda}t} &= \mathbf{I} + \frac{(\mathbf{\Lambda}t)}{1!} + \frac{(\mathbf{\Lambda}t)^2}{2!} + \frac{(\mathbf{\Lambda}t)^3}{3!} + \cdots \\ &= \begin{bmatrix} e^{\lambda_1 t} & 0 & 0 & 0 \\ 0 & e^{\lambda_2 t} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & e^{\lambda_n t} \end{bmatrix} \end{aligned} \tag{8B.38}$$

We now revert back to our original variable:

$$\begin{aligned} \mathbf{q} &= \mathbf{U}\mathbf{z} \\ &= \mathbf{U}e^{\mathbf{\Lambda}t}\mathbf{z}(0) \\ &= \mathbf{U}e^{\mathbf{\Lambda}t}\mathbf{U}^{-1}\mathbf{q}(0) \end{aligned} \tag{8B.39}$$

and since we know the ZIR is $\mathbf{q}_{\text{ZIR}}(t) = \boldsymbol{\varphi}(t)\mathbf{q}(0)$ then the transition matrix is:

Transition matrix written in terms of eigenvalues and eigenvectors

$$\boxed{\boldsymbol{\varphi}(t) = \mathbf{U}e^{\mathbf{\Lambda}t}\mathbf{U}^{-1}} \tag{8B.40}$$

This is a quick way to find the transition matrix $\boldsymbol{\varphi}(t)$ for high-order systems.

The ZIR of the states is then just:

$$\boxed{\begin{aligned} \mathbf{q}(t) &= \boldsymbol{\varphi}(t)\mathbf{q}(0) \\ &= \mathbf{U}e^{\mathbf{\Lambda}t}\mathbf{U}^{-1}\mathbf{q}(0) \end{aligned}} \tag{8B.41}$$

The ZIR written in terms of eigenvalues and eigenvectors, and initial conditions

Example

Given:

$$\begin{aligned} \dot{\mathbf{q}}(t) &= \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}\mathbf{q}(t) + \begin{bmatrix} 0 \\ 2 \end{bmatrix}x(t) \\ y(t) &= \begin{bmatrix} 3 & 1 \end{bmatrix}\mathbf{q}(t) \\ \mathbf{q}(0) &= \begin{bmatrix} 2 \\ 3 \end{bmatrix} \end{aligned} \tag{8B.42}$$

find the ZIR.

We start by determining the eigenvalues:

$$\begin{aligned} |\mathbf{A} - \lambda\mathbf{I}| &= 0 \\ \begin{vmatrix} -\lambda & 1 \\ -2 & -3-\lambda \end{vmatrix} &= 0 \\ \lambda^2 + 3\lambda + 2 &= 0 \\ (\lambda+1)(\lambda+2) &= 0 \\ \lambda_1 = -1, \quad \lambda_2 &= -2 \end{aligned} \tag{8B.43}$$

Next we find the right-hand eigenvectors:

$$\mathbf{U}_{(1)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{U}_{(2)} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \therefore \mathbf{U} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \tag{8B.44}$$

and the left-hand eigenvectors:

$$\mathbf{V}^{(1)} = \begin{bmatrix} 2 & 1 \end{bmatrix}, \quad \mathbf{V}^{(2)} = \begin{bmatrix} -1 & -1 \end{bmatrix}, \quad \therefore \mathbf{V} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \tag{8B.45}$$

As a check, we can see if $\mathbf{UV}=\mathbf{I}$.

Forming $\mathbf{\Lambda}$ is easy:

$$\mathbf{\Lambda}=\begin{bmatrix}-1 & 0\\ 0 & -2\end{bmatrix} \tag{8B.46}$$

The transition matrix is also easy:

$$\begin{aligned}\boldsymbol{\varphi}(t)&=\mathbf{U}e^{\mathbf{\Lambda}t}\mathbf{U}^{-1}\\ &=\begin{bmatrix}1 & 1\\ -1 & -2\end{bmatrix}\begin{bmatrix}e^{-t} & 0\\ 0 & e^{-2t}\end{bmatrix}\begin{bmatrix}2 & 1\\ -1 & -1\end{bmatrix}\\ &=\begin{bmatrix}1 & 1\\ -1 & -2\end{bmatrix}\begin{bmatrix}2e^{-t} & e^{-t}\\ -e^{-2t} & -e^{-2t}\end{bmatrix}\\ &=\begin{bmatrix}2e^{-t}-e^{-2t} & e^{-t}-e^{-2t}\\ -2e^{-t}+2e^{-2t} & -e^{-t}+2e^{-2t}\end{bmatrix}\end{aligned} \tag{8B.47}$$

The ZIR of the states is now just:

$$\begin{aligned}\mathbf{q}_{\mathrm{ZIR}}(t)&=\boldsymbol{\varphi}(t)\mathbf{q}(0)\\ &=\begin{bmatrix}2e^{-t}-e^{-2t} & e^{-t}-e^{-2t}\\ -2e^{-t}+2e^{-2t} & -e^{-t}+2e^{-2t}\end{bmatrix}\begin{bmatrix}2\\ 3\end{bmatrix}\\ &=\begin{bmatrix}7e^{-t}-5e^{-2t}\\ -7e^{-t}+10e^{-2t}\end{bmatrix}\end{aligned} \tag{8B.48}$$

The ZIR of the output is then:

$$\begin{aligned}y_{\mathrm{ZIR}}(t)&=\begin{bmatrix}3 & 1\end{bmatrix}\mathbf{q}_{\mathrm{ZIR}}(t)\\ &=14e^{-t}-5e^{-2t}\end{aligned} \tag{8B.49}$$

For higher-order systems and computer analysis, this method results in considerable time and computational savings.

Poles and Repeated Eigenvalues

Poles

We have seen before that the transfer function of a system using state-variables is given by:

$$\begin{aligned}H(s)&=\mathbf{c}^T\mathbf{\Phi}(s)\mathbf{b}\\ &=\mathbf{c}^T(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}\\ &=\frac{\mathbf{c}^T\,\mathrm{adj}(s\mathbf{I}-\mathbf{A})\mathbf{b}}{|s\mathbf{I}-\mathbf{A}|}\end{aligned} \tag{8B.50}$$

where we have used the formula for the inverse $\mathbf{B}^{-1}=\mathrm{adj}\,\mathbf{B}/|\mathbf{B}|$. We can see that the poles of the system are formed directly by the characteristic equation $|s\mathbf{I}-\mathbf{A}|=0$. Thus, the *poles* of a system are given by the *eigenvalues* of the matrix $\mathbf{A}$.

$$\boxed{\text{poles} = \text{eigenvalues of } \mathbf{A}} \tag{8B.51}$$

Poles of a system and eigenvalues of $\mathbf{A}$ are the same!

There is one qualifier to this statement: there could be a pole-zero cancellation, in which case the corresponding pole will "disappear". These are special cases, and are termed *uncontrollable* and/or *unobservable*, depending on the state assignments.

Repeated Eigenvalues

When eigenvalues are repeated, we get repeated eigenvectors if we try to solve $\mathbf{Aq}=\lambda\mathbf{q}$. In these cases, it is not possible to diagonalize the original matrix (because the similarity matrix will have two or more repeated columns, and will be singular – hence no inverse).

In cases of repeated eigenvalues, the closest we can get to a diagonal form is a *Jordan canonical form*. Handling repeated eigenvalues and examining the Jordan form are topics for more advanced subjects in control theory.

Discrete-time State-Variables

The concepts of state, state vectors and state-variables can be extended to discrete-time systems.

A discrete-time SISO system is described by the following equations:

State and output equations for a discrete-time system

$$\boxed{\begin{aligned}\mathbf{q}[n+1] &= \mathbf{A}\mathbf{q}[n] + \mathbf{b}x[n] \\ y[n] &= \mathbf{c}^T\mathbf{q}[n] + dx[n]\end{aligned}} \tag{8B.52}$$

Example

Given the following second-order linear difference equation:

$$y[n] = y[n-1] + y[n-2] + x[n] \tag{8B.53}$$

we select:

$$\begin{aligned} q_1[n] &= y[n-1] \\ q_2[n] &= y[n-2] \end{aligned} \tag{8B.54}$$

We now want to write the given difference equation as a set of equations in state-variable form. Now:

$$q_1[n] = y[n-1] \tag{8B.55}$$

Therefore:

$$\begin{aligned} q_1[n+1] &= y[n] \\ &= y[n-1] + y[n-2] + x[n] \end{aligned} \tag{8B.56}$$

so that:

$$q_1[n+1] = q_1[n] + q_2[n] + x[n] \tag{8B.57}$$

Also:

$$q_2[n] = y[n-2] \tag{8B.58}$$

Therefore:

$$q_2[n+1] = y[n-1] \tag{8B.59}$$

so that, using Eq. (8B.55):

$$q_2[n+1] = q_1[n] \tag{8B.60}$$

From Eq. (8B.53), we also have:

$$y[n] = q_1[n] + q_2[n] + x[n] \tag{8B.61}$$

The equations are now in state variable form, and we can write:

$$\begin{aligned} \mathbf{q}[n+1] &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\mathbf{q}[n] + \begin{bmatrix} 1 \\ 0 \end{bmatrix}x[n] \\ y[n] &= \begin{bmatrix} 1 & 1 \end{bmatrix}\mathbf{q}[n] + [1]x[n] \end{aligned} \tag{8B.62}$$

Thus, we can proceed as above to convert any given difference equation to state-variable form.

Example

If we are given the transfer function:

$$H(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_N z^{-N}}{1 + a_1 z^{-1} + \cdots + a_N z^{-N}} \tag{8B.63}$$

then:

$$\frac{Y(z)}{b_0 + b_1 z^{-1} + \cdots + b_p z^{-p}} = \frac{X(z)}{1 + a_1 z^{-1} + \cdots + a_p z^{-p}} = P(z) \tag{8B.64}$$

From the right-hand side, we have:

$$P(z) = -a_1 z^{-1} P(z) - a_2 z^{-2} P(z) - \cdots - a_N z^{-N} P(z) + X(z) \tag{8B.65}$$

Now select the state variables as:

$$\begin{aligned} Q_1(z) &= z^{-1}P(z) \\ Q_2(z) &= z^{-2}P(z) \\ &\vdots \\ Q_N(z) &= z^{-N}P(z) \end{aligned} \tag{8B.66}$$

The state equations are then built up as follows. From the first equation in Eq. (8B.66):

$$\begin{aligned} Q_1(z) &= z^{-1}P(z) \\ zQ_1(z) &= P(z) \\ &= -a_1z^{-1}P(z) - a_2z^{-2}P(z) - \cdots - a_Nz^{-N}P(z) + X(z) \\ &= -a_1Q_1(z) - a_2Q_2(z) - \cdots - a_NQ_N(z) + X(z) \end{aligned} \tag{8B.67}$$

Taking the inverse *z*-transform, we get:

$$q_1[n+1] = -a_1q_1[n] - a_2q_2[n] - \cdots - a_Nq_N[n] + x[n] \tag{8B.68}$$

From the second equation in Eq. (8B.66), we have:

$$\begin{aligned} Q_2(z) &= z^{-2}P(z) \\ &= z^{-1}z^{-1}P(z) \\ &= z^{-1}Q_1(z) \\ zQ_2(z) &= Q_1(z) \end{aligned} \tag{8B.69}$$

Taking the inverse *z*-transform gives us:

$$q_2[n+1] = q_1[n] \tag{8B.70}$$

Similarly:

$$q_N[n+1] = q_{N-1}[n] \tag{8B.71}$$

We now have all the state equations. Returning to Eq. (8B.64), we have:

$$Y(z) = b_0P(z) + b_1z^{-1}P(z) + \cdots + b_Nz^{-N}P(z) \tag{8B.72}$$

Taking the inverse *z*-transform gives:

$$y[n] = b_0q_1[n+1] + b_1q_1[n] + b_2q_2[n] + \cdots + b_Nq_N[n] \tag{8B.73}$$

Eliminating the $q_1[n+1]$ term using Eq. (8B.68) and grouping like terms gives:

$$\begin{aligned} y[n] = {} & (b_1 - a_1b_0)q_1[n] + (b_2 - a_2b_0)q_2[n] + \\ & \cdots + (b_N - a_Nb_0)q_N[n] + b_0x[n] \end{aligned} \tag{8B.74}$$

which is in the required form.

Using matrix notation we therefore have:

$$\begin{aligned} \mathbf{q}[n+1] &= \begin{bmatrix} -a_1 & -a_2 & \cdots & -a_{N-1} & -a_N \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \mathbf{q}[n] + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} x[n] \\ y[n] &= \begin{bmatrix} c_1 & c_2 & \cdots & c_N \end{bmatrix} \mathbf{q}[n] + b_0x[n] \\ & \text{where } c_i = (b_i - a_ib_0) \quad i = 1, 2, \ldots, N \end{aligned} \tag{8B.75}$$

Thus we can convert $H(z)$ to state-variable form.

Discrete-time Response

Once we have the equations in state-variable form we can then obtain the discrete-time response.

For a SISO system we have:

$$\begin{aligned}\mathbf{q}[n+1] &= \mathbf{A}\mathbf{q}[n]+\mathbf{b}x[n]\\ y[n] &= \mathbf{c}^T\mathbf{q}[n]+dx[n]\end{aligned} \tag{8B.76}$$

First establish the states for the first few values of n:

$$\begin{aligned}\mathbf{q}[1] &= \mathbf{A}\mathbf{q}[0]+\mathbf{b}x[0]\\ \mathbf{q}[2] &= \mathbf{A}\mathbf{q}[1]+\mathbf{b}x[1]\\ &= \mathbf{A}(\mathbf{A}\mathbf{q}[0]+\mathbf{b}x[0])+\mathbf{b}x[1]\\ &= \mathbf{A}^2\mathbf{q}[0]+\mathbf{A}\mathbf{b}x[0]+\mathbf{b}x[1]\\ \mathbf{q}[3] &= \mathbf{A}\mathbf{q}[2]+\mathbf{b}x[2]\\ &= \mathbf{A}^3\mathbf{q}[0]+\mathbf{A}^2\mathbf{b}x[0]+\mathbf{A}\mathbf{b}x[1]+\mathbf{b}x[2]\end{aligned} \tag{8B.77}$$

The general formula can then be seen as:

$$\begin{aligned}\mathbf{q}[n] &= \mathbf{A}^n\mathbf{q}[0]+\mathbf{A}^{n-1}\mathbf{b}x[0]+\\ &\quad\cdots+\mathbf{A}\mathbf{b}x[n-2]+\mathbf{b}x[n-1] \quad n=1,2,\ldots\end{aligned} \tag{8B.78}$$

We now define:

Fundamental matrix defined

$$\boxed{\boldsymbol{\varphi}[n]=\mathbf{A}^n=\text{fundamental matrix}} \tag{8B.79}$$

From Eq. (8B.78) and the above definition, the response of the discrete-time system to any input is given by:

Solution to the discrete-time state equations in terms of convolution summation

$$\boxed{\begin{aligned}\mathbf{q}[n] &= \boldsymbol{\varphi}[n]\mathbf{q}[0]+\sum_{i=0}^{n-1}\boldsymbol{\varphi}[n-i-1]\mathbf{b}x[i]\\ \mathbf{y}[n] &= \mathbf{c}^T\mathbf{q}[n]+dx[n]\end{aligned}} \tag{8B.80}$$

This is the expected form of the output response. For the *states*, it can be seen that:

The ZIR and ZSR for a discrete-time state-variable system

$$\boxed{\begin{aligned}\mathbf{q}_{\text{ZIR}} &= \boldsymbol{\varphi}[n]\mathbf{q}[0]\\ \mathbf{q}_{\text{ZSR}} &= \sum_{i=0}^{n-1}\boldsymbol{\varphi}[n-i-1]\mathbf{b}x[i]\end{aligned}} \tag{8B.81}$$

Discrete-time Transfer Function

We can determine the transfer function from the state-variable representation in the same manner as we did for continuous-time systems.

Take the z-transform of Eq. (8B.76) to get:

$$\begin{aligned}z\mathbf{Q}(z)-z\mathbf{q}(0) &= \mathbf{A}\mathbf{Q}(z)+\mathbf{b}X(z)\\ (z\mathbf{I}-\mathbf{A})\mathbf{Q}(z) &= z\mathbf{q}(0)+\mathbf{b}X(z)\end{aligned} \tag{8B.82}$$

Therefore:

$$\boxed{\mathbf{Q}(z)=(z\mathbf{I}-\mathbf{A})^{-1}z\mathbf{q}(0)+(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}X(z)} \tag{8B.83}$$

Similarly:

$$\begin{aligned}Y(z) &= z\mathbf{c}^T(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{q}(0)\\ &\quad+\left\{\mathbf{c}^T(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}+d\right\}X(z)\end{aligned} \tag{8B.84}$$

For the transfer function, we put all initial conditions $\mathbf{q}(0)=\mathbf{0}$. Therefore, the transfer function is:

The discrete-time transfer function in terms of state-variable quantities

$$\boxed{H(z)=\mathbf{c}^T(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}+d} \tag{8B.85}$$

To get the unit-pulse response, we revert to Eq. (8B.80), set the initial conditions to zero and apply a unit-pulse response:

$$\begin{aligned} h[0]&=d \\ h[n]&=\mathbf{c}^T\boldsymbol{\varphi}[n-1]\mathbf{b} \quad n=1,2,\ldots \end{aligned} \tag{8B.86}$$

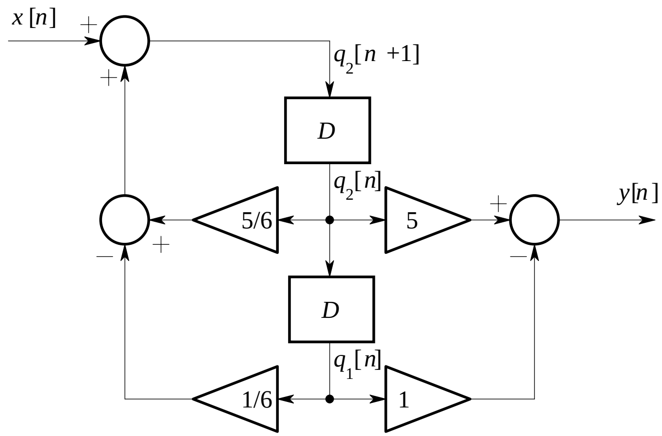
Using Eq. (8B.79), we get:

The unit-pulse response in terms of state-variable quantities

$$\boxed{\begin{aligned} h[0]&=d \\ h[n]&=\mathbf{c}^T\mathbf{A}^{n-1}\mathbf{b} \quad n=1,2,\ldots \end{aligned}} \tag{8B.87}$$

Example

A linear time-invariant discrete-time system is given by the figure shown:



If

We would like to find the output $y[n]$ if the input is $x[n]=u[n]$ and the initial conditions are $q_1[0]=2$ and $q_2[0]=3$.

Recognizing that $q_2[n]=q_1[n+1]$, the state equations are:

$$\begin{bmatrix} q_1[n+1] \\ q_2[n+1] \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ -\frac{1}{6} & \frac{5}{6} \end{bmatrix}\begin{bmatrix} q_1[n] \\ q_2[n] \end{bmatrix}+\begin{bmatrix} 0 \\ 1 \end{bmatrix}x \tag{8B.88}$$

and:

$$y[n]=\begin{bmatrix} -1 & 5 \end{bmatrix}\begin{bmatrix} q_1[n] \\ q_2[n] \end{bmatrix} \tag{8B.89}$$

To find the solution using time-domain techniques [Eq. (8B.81)], we must determine $\boldsymbol{\varphi}[n]=\mathbf{A}^n$. One way of finding this is to first determine a similarity transform to diagonalize $\mathbf{A}$.

If such a transform can be found, then $\mathbf{\Lambda} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$. Rearranging we then have:

$$\begin{aligned}
&\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\,\mathbf{U}^{-1} \\
&\mathbf{A}^2 = \mathbf{A}\cdot\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\,\mathbf{U}^{-1}\mathbf{U}\mathbf{\Lambda}\,\mathbf{U}^{-1} = \mathbf{U}\mathbf{\Lambda}^2\,\mathbf{U}^{-1} \\
&\vdots \\
&\mathbf{A}^n = \mathbf{U}\mathbf{\Lambda}^n\,\mathbf{U}^{-1}
\end{aligned} \tag{8B.90}$$

The characteristic equation of $\mathbf{A}$ is:

$$|\lambda\mathbf{I}-\mathbf{A}| = \begin{vmatrix} \lambda & -1 \\ \frac{1}{6} & \lambda - \frac{5}{6} \end{vmatrix} = \lambda^2 - \frac{5}{6}\lambda + \frac{1}{6} = \left(\lambda - \frac{1}{3}\right)\left(\lambda - \frac{1}{2}\right) = 0 \tag{8B.91}$$

Hence, $\lambda_1 = 1/3$ and $\lambda_2 = 1/2$ are the eigenvalues of $\mathbf{A}$, and:

$$\mathbf{\Lambda} = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/2 \end{bmatrix} \tag{8B.92}$$

The associated eigenvectors are, for $\lambda_1 = 1/3$:

$$\begin{bmatrix} 1/3 & -1 \\ 1/6 & -1/2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ so choose } \mathbf{u}_{(1)} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \tag{8B.93}$$

and for $\lambda_2 = 1/2$:

$$\begin{bmatrix} 1/2 & -1 \\ 1/6 & -1/3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ so choose } \mathbf{u}_{(2)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{8B.94}$$

Therefore:

$$\mathbf{U} = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \tag{8B.95}$$

The inverse of $\mathbf{U}$ is readily found to be:

$$\mathbf{U}^{-1} = \frac{\operatorname{adj}\mathbf{U}}{|\mathbf{U}|} = \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} \tag{8B.96}$$

Therefore $\boldsymbol{\varphi}[n] = \mathbf{A}^n = \mathbf{U}\mathbf{\Lambda}^n\,\mathbf{U}^{-1}$, and:

$$\begin{aligned}
\boldsymbol{\varphi}[n] &= \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} (1/3)^n & 0 \\ 0 & (1/2)^n \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} (1/3)^n & -2(1/3)^n \\ -(1/2)^n & 3(1/2)^n \end{bmatrix} \\
&= \begin{bmatrix} 3(1/3)^n - 2(1/2)^n & -6(1/3)^n + 6(1/2)^n \\ (1/3)^n - (1/2)^n & -2(1/3)^n + 3(1/2)^n \end{bmatrix}
\end{aligned} \tag{8B.97}$$

To solve for the ZIR, we have, using Eq. (8B.81):

$$\begin{aligned}
y_{ZIR}[n] &= \mathbf{c}^T\boldsymbol{\varphi}[n]\mathbf{q}[0] \\
&= \begin{bmatrix} -1 & 5 \end{bmatrix}\begin{bmatrix} 3(1/3)^n - 2(1/2)^n & -6(1/3)^n + 6(1/2)^n \\ (1/3)^n - (1/2)^n & -2(1/3)^n + 3(1/2)^n \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\
&= \begin{bmatrix} -1 & 5 \end{bmatrix}\begin{bmatrix} -12(1/3)^n + 14(1/2)^n \\ -4(1/3)^n + 7(1/2)^n \end{bmatrix} \\
&= -8(1/3)^n + 21(1/2)^n
\end{aligned} \tag{8B.98}$$

To solve for the ZSR, we have, using Eq. (8B.81):

$$y_{ZSR}[n] = \sum_{i=0}^{n-1}\mathbf{c}^T\boldsymbol{\varphi}[n-i-1]\mathbf{b}x[i] \tag{8B.99}$$

We have:

$$\begin{aligned}
&\mathbf{c}^T\boldsymbol{\varphi}[n-i-1]\mathbf{b} \\
&= \begin{bmatrix} -1 & 5 \end{bmatrix}\begin{bmatrix} 3(1/3)^{n-i-1} - 2(1/2)^{n-i-1} & -6(1/3)^{n-i-1} + 6(1/2)^{n-i-1} \\ (1/3)^{n-i-1} - (1/2)^{n-i-1} & -2(1/3)^{n-i-1} + 3(1/2)^{n-i-1} \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} -1 & 5 \end{bmatrix}\begin{bmatrix} -6(1/3)^{n-i-1} + 6(1/2)^{n-i-1} \\ -2(1/3)^{n-i-1} + 3(1/2)^{n-i-1} \end{bmatrix} \\
&= -4(1/3)^{n-i-1} + 9(1/2)^{n-i-1}
\end{aligned} \tag{8B.100}$$

so that:

$$
\begin{aligned}
y_{ZSR}[n] &= \sum_{i=0}^{n-1} -4(1/3)^{n-i-1} + 9(1/2)^{n-i-1} u[i] \\
&= -4(1/3)^{n-1} \sum_{i=0}^{n-1} (1/3)^{-i} + 9(1/2)^{n-1} \sum_{i=0}^{n-1} (1/2)^{-i} \\
&= -12(1/3)^n \frac{1-3^n}{1-3} + 18(1/2)^n \frac{1-2^n}{1-2} \\
&= 6(1/3)^n \left(1-3^n\right) - 18(1/2)^n \left(1-2^n\right) \\
&= 6(1/3)^n - 18(1/2)^n + 12
\end{aligned}
\tag{8B.101}
$$

for $n \geq 0$. Therefore the total response is:

$$
\begin{aligned}
y[n] &= -8(1/3)^n + 21(1/2)^n + 12 + 6(1/3)^n - 18(1/2)^n \\
&= 12 - 2(1/3)^n + 3(1/2)^n, \quad n \geq 0
\end{aligned}
\tag{8B.102}
$$

The solution using z-transforms follows directly from Eq. (8B.84):

$$
\begin{aligned}
Y(z) &= z\mathbf{c}^T (z\mathbf{I}-\mathbf{A})^{-1}\mathbf{q}(0) + \left\{\mathbf{c}^T (z\mathbf{I}-\mathbf{A})^{-1}\mathbf{b} + d\right\} X(z) \\
&= z\begin{bmatrix}-1 & 5\end{bmatrix}\begin{bmatrix} z & -1 \\ 1/6 & z-5/6 \end{bmatrix}^{-1}\begin{bmatrix}2\\3\end{bmatrix} + \begin{bmatrix}-1 & 5\end{bmatrix}\begin{bmatrix} z & -1 \\ 1/6 & z-5/6 \end{bmatrix}^{-1}\begin{bmatrix}0\\ \frac{z}{z-1}\end{bmatrix} \\
&= \frac{z}{z^2-5/6z+1/6}\begin{bmatrix}-1 & 5\end{bmatrix}\begin{bmatrix} z-5/6 & 1 \\ -1/6 & z \end{bmatrix}\begin{bmatrix}2\\3\end{bmatrix} + \frac{1}{z^2-5/6z+1/6}\begin{bmatrix}-1 & 5\end{bmatrix}\begin{bmatrix} z-5/6 & 1 \\ -1/6 & z \end{bmatrix}\begin{bmatrix}0\\ \frac{z}{z-1}\end{bmatrix} \\
&= \frac{z}{z^2-5/6z+1/6}\begin{bmatrix}-1 & 5\end{bmatrix}\begin{bmatrix} 2z+4/3 \\ 3z-1/3 \end{bmatrix} + \frac{1}{z^2-5/6z+1/6}\begin{bmatrix}-1 & 5\end{bmatrix}\begin{bmatrix}\frac{z}{z-1}\\ \frac{z^2}{z-1}\end{bmatrix} \\
&= \frac{13z^2-3z}{z^2-5/6z+1/6} + \frac{5z^2-z}{(z-1)\left(z^2-5/6z+1/6\right)} \\
&= -8\frac{z}{z-1/3} + 21\frac{z}{z-1/2} + 12\frac{z}{z-1} + 6\frac{z}{z-1/3} - 18\frac{z}{z-1/2}
\end{aligned}
\tag{8B.103}
$$

Therefore:

$$
y[n] = \underbrace{-8(1/3)^n + 21(1/2)^n}_{\text{zero-input response}} + \underbrace{12 + 6(1/3)^n - 18(1/2)^n}_{\text{zero-state response}} \quad n \geq 0 \tag{8B.104}
$$

Obviously, the two solutions obtained using different techniques are in perfect agreement.

Summary

- The similarity transform uses a knowledge of a system's eigenvalues and eigenvectors to reduce a high-order coupled system to a simple diagonal form.
- A diagonal form exists only when there are no repeated eigenvalues.
- A system's poles and eigenvalues are equal.
- Discrete-time state-variable representation can be used to derive the complete response of a system, as well as the transfer function and unit-pulse response.

References

Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB®*, Prentice-Hall, 1997.

Shinners, S.: *Modern Control System Theory and Application*, Addison-Wesley, 1978.

Exercises

1.

Find $\mathbf{\Lambda}$, $\mathbf{U}$ and $\mathbf{V}$ for the system described by:

$$\dot{\mathbf{q}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}\mathbf{q} + \begin{bmatrix} 0 & 1 \\ 2 & 7 \\ -20 & -5 \end{bmatrix}\mathbf{x}$$
$$y = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}\mathbf{q}$$

Note: **U** and **V** should be found directly (i.e. Do not find **V** by taking the inverse of **U**). You can then verify your solution by:

(i) checking that $\mathbf{VU} = \mathbf{I}$

(ii) checking that $\mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{V} = \mathbf{A}$

2.

Consider the system:

$$\dot{q}_1 = q_1 + x$$
$$\dot{q}_2 = q_1 + 2q_2 + x$$

with:

$$\mathbf{q}(0) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

(a) Find the system eigenvalues and find the zero input response of the system.

(b) If the system is given a unit-step input with the same initial conditions, find $q(t)$. (Use the resolvent matrix to obtain the time solution). What do you notice about the output?

3.

A system employing state-feedback is described by the following equations:

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -5 & -7 & -3 \end{bmatrix}\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}x(t)$$
$$y = \begin{bmatrix} -2 & 4 & 3 \end{bmatrix}\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}$$
$$x(t) = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} + r(t)$$

(i) Draw a block diagram of the state-feedback system.

(ii) Find the poles of the system without state feedback.

(iii) Obtain values for k_1, k_2 and k_3 to place the closed-loop system poles at -4, -4 and -5.

(iv) Find the steady-state value of the output due to a unit-step input.

(v) Comment upon the possible uses of this technique.

8B.26

4.

For the difference equation:

$$y[n]=3x[n]+\frac{5}{4}x[n-1]-\frac{5}{8}x[n-2]-\frac{1}{4}y[n-1]+\frac{1}{4}y[n-2]+\frac{1}{16}y[n-3]$$

(a) Show that $H(z)=\dfrac{3z^3+5/4z^2-5/8z}{z^3+1/4z^2-1/4z-1/16}$

(b) Form the state-variable description of the system.

(c) Find the transfer function $H(z)$ from your answer in (b).

(d) Draw a block diagram of the state-variable description found in (b), and use block-diagram reduction to find $H(z)$.

5.

Find $y[5]$ and $y[10]$ for the answer to Q4b given:

$$\mathbf{q}[0]=\begin{bmatrix}1\\-1\\0\end{bmatrix} \quad \text{and} \quad x[n]=\begin{cases}0 & n<0\\(-1)^n & n\geq 0\end{cases}$$

(a) by calculating $q[n]$ and $y[n]$ for $n=0,1,\ldots,10$ directly from the state equations.

(b) by using the fundamental matrix and finding $y[5]$ and $y[10]$ directly.

6.

A discrete-time system with:

$$H(z)=\frac{k(z+b)}{z(z+a)(z+c)}$$

is to be controlled using unity feedback. Find a state-space representation of the resulting closed-loop system.

Appendix A - The Fast Fourier Transform

Discrete-time Fourier transform. Discrete Fourier Transform. Fast Fourier transform.

Overview

The motivation behind developing the FFT

Digital signal processing is becoming prevalent throughout engineering. We have digital audio equipment (CDs, MP3s), digital video (MPEG2 and MPEG4, DVD, digital TV), digital phones (fixed and mobile). An increasing number (billions!) of embedded systems exist that rely on a digital computer (a microcontroller normally). They take input signals from the real analog world, convert them to digital signals, process them digitally, and produce outputs that are again suitable for the real analog world. (Think of the computers controlling any modern form of transport – car, plane, boat – or those that control nearly all industrial processes). Our motivation is to extend our existing frequency-domain analytical techniques to the "digital world".

Samples contain all the information of a signal

Our reason for hope that this can be accomplished is the fact that a signal's samples can convey the complete information about a signal (if Nyquist's criterion is met). We should be able to turn those troublesome continuous-time integrals into simple summations – a task easily carried out by a digital computer.

The Discrete-Time Fourier Transform (DTFT)

To illustrate the derivation of the discrete-time Fourier Transform, we will consider the signal and it's Fourier Transform below:

A strictly time-limited signal has an infinite bandwidth

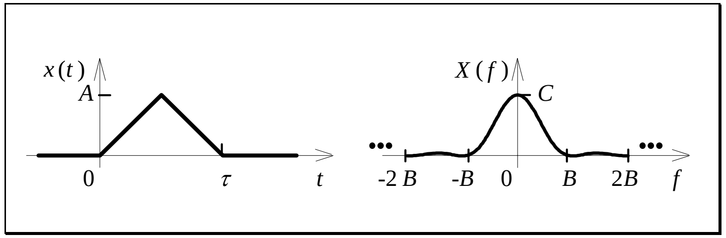


Figure F.1

Since the signal is *strictly* time-limited (it only exists for a finite amount of time), it's spectrum must be infinite in extent. We therefore cannot choose a sample rate high enough to satisfy Nyquist's criterion (and therefore prevent aliasing). However, in practice we normally find that the spectral content of signals drops off at high frequencies, so that the signal is *essentially* band-limited to *B*:

A time-limited signal which is also essentially band-limited

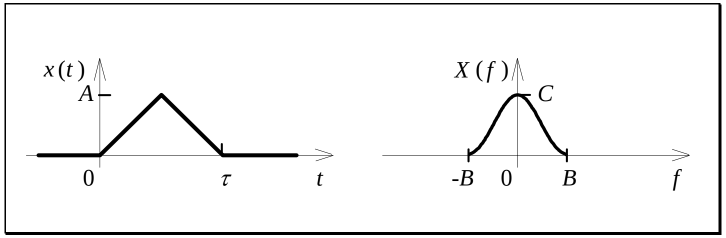


Figure F.2

We will now assume that Nyquist's criterion is met if we sample the time-domain signal at a sample rate $f_s = 2B$.

If we ideally sample using a uniform train of impulses (with spacing $T_s = 1/f_s$), the mathematical expression for the sampled time-domain waveform is:

$$\begin{aligned} x_s(t) &= x(t)\left[\sum_{k=-\infty}^{\infty}\delta(t-kT_s)\right] \\ &= \sum_{k=-\infty}^{\infty} x[k]\delta(t-kT_s) \end{aligned} \tag{F.1}$$

This corresponding operation in the frequency-domain gives:

$$\begin{aligned} X_s(f) &= X(f) * f_s\sum_{n=-\infty}^{\infty}\delta(f-nf_s) \\ &= \sum_{n=-\infty}^{\infty} f_s X(f-nf_s) \end{aligned} \tag{F.2}$$

The sampled waveform and its spectrum are shown graphically below:

A sampled signal and its spectrum

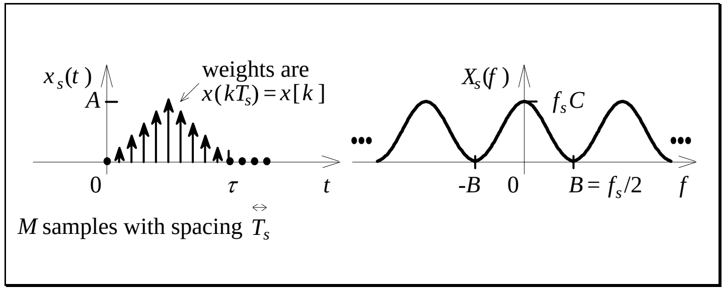


Figure F.3

We are free to take as many samples as we like, so long as $MT_s \geq \tau$. That is, we need to ensure that our samples will encompass the whole signal in the time-domain. For a one-off waveform, we can also sample past the extent of the signal – a process known as zero padding.

Substituting $x_s(t)$ into the definition of the Fourier transform, we get:

$$X_s(f) = \int_{-\infty}^{\infty} x_s(t) e^{-j2\pi ft} dt$$
$$= \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} x[k]\delta(t - kT_s) e^{-j2\pi ft} dt \quad \text{(F.3)}$$

Using the sifting property of the impulse function, this simplifies into the definition of the discrete-time Fourier transform (DTFT):

The discrete-time Fourier transform defined

$$\boxed{X_s(f) = \sum_{k=-\infty}^{\infty} x[k] e^{-j2\pi fkT_s}} \quad \text{(F.4)}$$

The DTFT is a *continuous* function of *f*. It is discrete in the sense that it operates on a discrete-time signal (in this case, the discrete-time signal corresponds to the weights of the impulses of the sampled signal).

As shown in Figure F.3, the DTFT is periodic with period $f_s = 2B$, i.e. the range of frequencies $-f_s/2 \le f < f_s/2$ uniquely specifies it.

The Discrete Fourier Transform (DFT)

The DTFT is a continuous function of *f*, but we need a discrete function of *f* to be able to store it in a computer. Our reasoning now is that since samples of a time-domain waveform (taken at the right rate) uniquely determine the original waveform, then samples of a spectrum (taken at the right rate) uniquely determine the original spectrum.

The discrete Fourier transform is just samples of the DTFT

One way to discretize the DTFT is to ideally sample it in the frequency-domain. Since the DTFT is periodic with period f_s, then we choose *N* samples per period where *N* is an integer. This yields *periodic* spectrum samples and we only need to compute *N* of them (the rest will be the same)!

The spacing between samples is then:

The frequency spacing of a DFT

$$f_0 = \frac{f_s}{N} \quad \text{(F.5)}$$

which gives a time-domain relationship:

The time-domain sample spacing of a DFT

$$T_0 = NT_s \quad \text{(F.6)}$$

The ideally sampled spectrum $X_{ss}(f)$ is then:

$$X_{ss}(f) = X_s(f)\left[\sum_{n=-\infty}^{\infty} \delta(f - nf_0)\right]$$
$$= \sum_{n=-\infty}^{\infty} X_s(nf_0)\delta(f - nf_0) \quad \text{(F.7)}$$

The ideally sampled spectrum is shown below:

Ideal samples of a DTFT spectrum

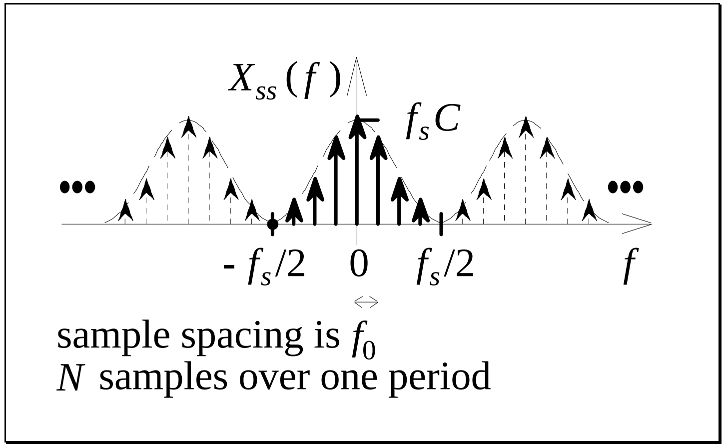


Figure F.4

What signal does this spectrum correspond to in the time-domain? The corresponding operation of Eq. (F.7) is shown below in the time-domain:

$$x_{ss}(t) = x_s(t) * \left[\frac{1}{f_0} \sum_{k=-\infty}^{\infty} \delta(t - kT_0) \right]$$
$$= \sum_{k=-\infty}^{\infty} T_0 x_s(t - kT_0) \tag{F.8}$$

This time-domain signal is shown below:

Periodic extension of the time-domain signal due to ideally sampling the DTFT spectrum

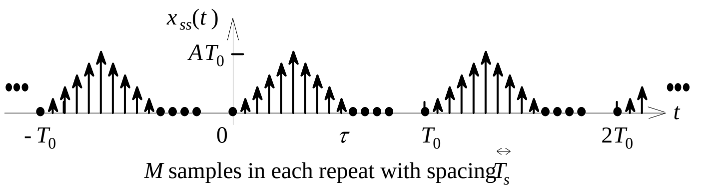


Figure F.5

We see that sampling in the frequency-domain causes periodicy in the time-domain. What we have created is a *periodic extension* of the original sampled signal, but scaled in amplitude. From Figure F.5, we can see that no time-domain aliasing occurs (no overlap of repeats of the original sampled signal) if:

$$T_0 \geq MT_s \tag{F.9}$$

Using Eq. (F.6), this means:

$$NT_s \geq MT_s$$

or

$$N \geq M \tag{F.10}$$

If the original time-domain waveform is periodic, and our samples represent *one* period in the time-domain, then we *must* choose the frequency sample spacing to be $f_0 = 1/T_0$, where T_0 is the period of the original waveform. The process of ideally sampling the spectrum at this spacing "creates" the original periodic waveform back in the time-domain. In this case, we must have $T_0 = MT_s$, and therefore $NT_s = MT_s$ so that $N = M$.

We *choose* the number of samples N in the frequency-domain so that Eq. (F.10) is satisfied, but we also choose N to handle the special case above. In any case, setting $N = M$ minimises the DFT calculations, and we therefore choose:

$$N = M \tag{F.11}$$

number of frequency-domain samples
= number of time-domain samples

The FFT produces the same number of samples in the frequency-domain as there are in the time-domain

If we do this, then the reconstructed sampled waveform has its repeats next to each other:

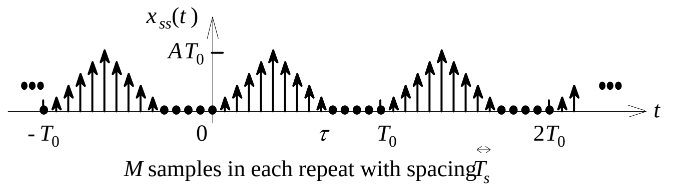


Figure F.6

Returning now to the spectrum in Figure F.4, we need to find the sampled spectrum impulse weights covering just one period, say from $0 \leq f < f_s$. These are given by the weights of the impulses in Eq. (F.7) where $0 \leq n \leq N-1$.

The DTFT evaluated at those frequencies is then obtained by substituting $f = nf_0$ into Eq. (F.4):

$$X_s(nf_0) = \sum_{k=0}^{N-1} x[k] e^{-j2\pi n f_0 k T_s} \tag{F.12}$$

Notice how the infinite summation has turned into a finite summation over $M = N$ samples of the waveform, since we know the waveform is zero for all the other values of *k*.

Now using Eq. (F.5), we have $f_0 T_s = 1/N$. We then have the definition of the discrete Fourier transform (DFT):

The discrete Fourier transform defined

$$\boxed{X[n] = \sum_{k=0}^{N-1} x[k] e^{-j\frac{2\pi nk}{N}}, \; 0 \le n \le N-1} \tag{F.13}$$

The DFT takes an input vector $x[n]$ (time-spacing unknown – just a function of *n*) and produces an output vector $X[n]$ (frequency-spacing unknown – just a function of *n*). It is up to us to interpret the input and output of the DFT.

The Fast Fourier Transform (FFT)

The FFT is a computationally fast version of the DFT

The Fast Fourier Transform is really a family of algorithms that are used to evaluate the DFT. They are optimised to take advantage of the periodicy inherent in the exponential term in the DFT definition. The roots of the FFT algorithm go back to the great German mathematician Gauss in the early 1800's, but was formally introduced by Cooley and Tukey in their paper "An Algorithm for the Machine Calculation of Complex Fourier Series," *Math. Comput.* 19, no. 2, April 1965:297-301. Most FFTs are designed so that the number of sample points, *N*, is a power of 2.

All we need to consider here is the *creation* and *interpretation* of FFT results, and not the algorithms behind them.

Creating FFTs

Knowing the background behind the DFT, we can now choose various parameters in the creation of the FFT to suit our purposes. For example, there may be a requirement for the FFT results to have a certain frequency resolution, or we may be restricted to a certain number of samples in the time-domain and we wish to know the frequency range and spacing of the FFT output.

The important relationships that combine all these parameters are:

The relationships between FFT sample parameters

$$\boxed{\begin{gathered} T_0 = NT_s \\ \text{or} \\ f_s = Nf_0 \\ \text{or} \\ N = T_0 f_s \end{gathered}} \tag{F.14}$$

Example

An analog signal with a known bandwidth of 2000 Hz is to be sampled at the minimum possible frequency, and the frequency resolution is to be 5 Hz. We need to find the sample rate, the time-domain window size, and the number of samples.

The minimum sampling frequency we can choose is the Nyquist rate of $f_s = 2B = 2 \cdot 2000 = 4000\,\text{Sa/s}$. To achieve a frequency resolution of 5 Hz requires a window size of $T_0 = 1/f_0 = 1/5 = 0.2\,\text{s}$. The resulting number of samples is then $N = T_0 f_s = 0.2 \cdot 4000 = 800$ samples.

Example

An analog signal is viewed on a DSO with a window size of 1 ms. The DSO takes 1024 samples. What is the frequency resolution of the spectrum, and what is the folding frequency (half the sample rate)?

The frequency resolution is $f_0 = 1/T_0 = 1/0.001 = 1\,\text{kHz}$.

The sample rate is $f_s = Nf_0 = 1024 \cdot 1000 = 1.024\,\text{MHz}$. The folding frequency is therefore $f_s/2 = 512\,\text{kHz}$, and this is the maximum frequency displayed on the DSO spectrum.

Example

A simulation of a system and its signals is being performed using MATLAB®. The following code shows how to set up the appropriate time and frequency vectors if the sample rate and number of samples are specified:

```
% Sample rate
fs=1e6;
Ts=1/fs;

% Number of samples
N=1000;

% Time window
T0=N*Ts;
f0=1/T0;

% Time vector of N time samples spaced Ts apart
t=0:Ts:T0-Ts;

% Frequency vector of N frequencies spaced f0 apart
f=-fs/2:f0:fs/2-f0;
```

The frequency vector is used to graph the shifted (double-sided) spectrum produced by the code below:

```
G=fftshift(fft(g));
```

The frequency resolution of the FFT in this case will be $f_0 = 1\,\text{kHz}$ and the output will range from –500 kHz to 499 kHz. Note carefully how the time and frequency vectors were specified so that the last value does not coincide with the first value of the second periodic extension or spectrum repeat.

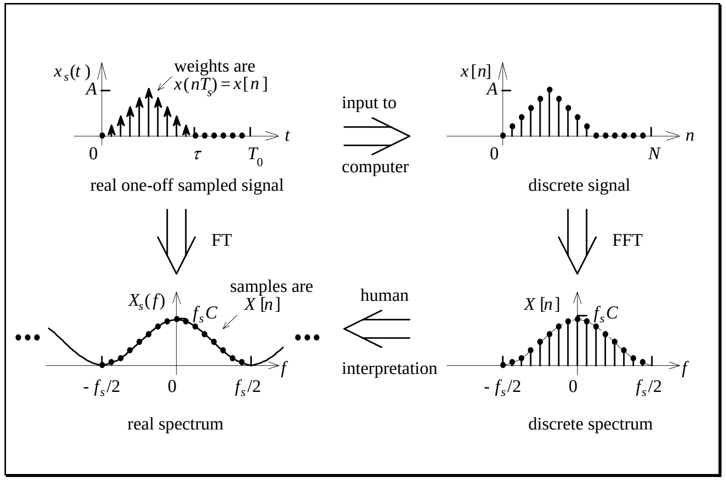
Interpreting FFTs

The output of the FFT can be interpreted in four ways, depending on how we interpret the time-domain values that we feed into it.

Case 1 – Ideally sampled one-off waveform

If the FFT input, $x[n]$, is the weights of the impulses of an ideally sampled time-limited "one-off" waveform, then we know the FT is a periodic repeat of the original unsampled waveform. The DFT gives the value of the FT at frequencies nf_0 for the first spectral repeat. This interpretation comes directly from Eq. (F.12).

With our example waveform, we would have:



Interpretation of FFT of ideally sampled one-off waveform

Figure F.7

Case 2 – Ideally sampled periodic waveform

Consider the case where the FFT input, $x[n]$, is the weights of the impulses of an ideally sampled periodic waveform over one period with period T_0. According to Eq. (F.8), the DFT gives the FT impulse weights for the first spectral repeat, if the time-domain waveform were scaled by T_0. To get the FT, we therefore have to scale the DFT by f_0 and recognise that the spectrum is periodic.

With our example waveform, we would have:

Interpretation of FFT of ideally sampled periodic waveform

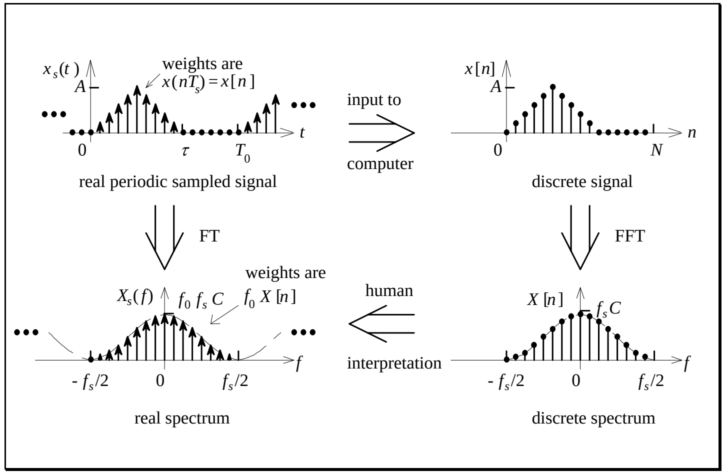


Figure F.8

Case 3 – Continuous time-limited waveform

If we ideally sample the "one-off" waveform at intervals T_s, we get a signal corresponding to Case 1. The sampling process creates periodic repeats of the original spectrum, scaled by f_s. To undo the scaling and spectral repeats caused by sampling, we should multiply the Case 1 spectrum by T_s and filter out all periodic repeats except for the first. The DFT gives the value of the FT of the sampled waveform at frequencies nf_0 for the first spectral repeat. This interpretation comes directly from Eq. (F.12). All we have to do is scale the DFT output by T_s to obtain the true FT at frequencies nf_0.

Interpretation of FFT of one-off waveform

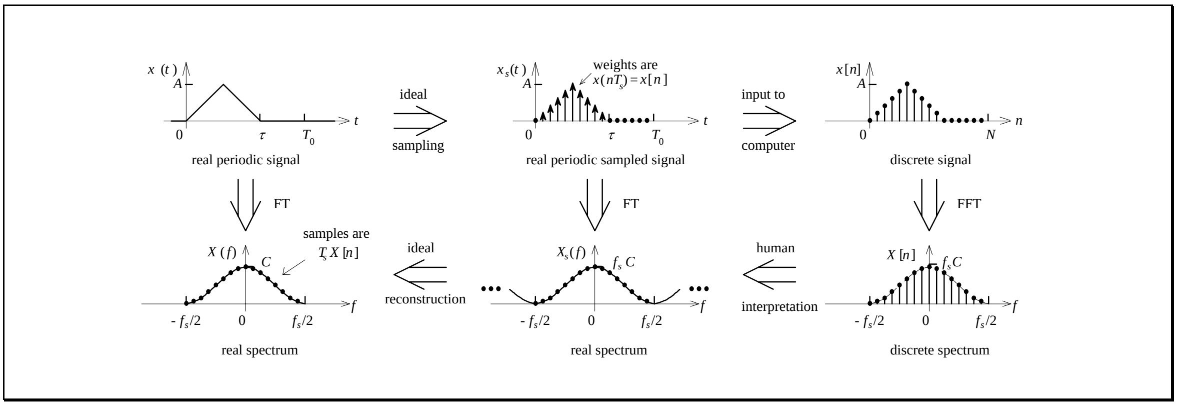


Figure F.9

Case 4 – Continuous periodic waveform

To create the FFT input for this case, we must ideally sample the continuous signal at intervals T_s to give $x[n]$. The sampling process creates periodic repeats of the original spectrum, scaled by f_s. We are now essentially equivalent to Case 2. To undo the scaling and spectral repeats caused by sampling, we should multiply the Case 2 spectrum by T_s and filter out all periodic repeats except for the first. According to Eq. (F.8), the DFT gives the FT impulse weights for the first spectral repeat. All we have to do is scale the DFT output by $f_0 T_s = 1/N$ to obtain the true FT impulse weights.

Interpretation of FFT of periodic waveform

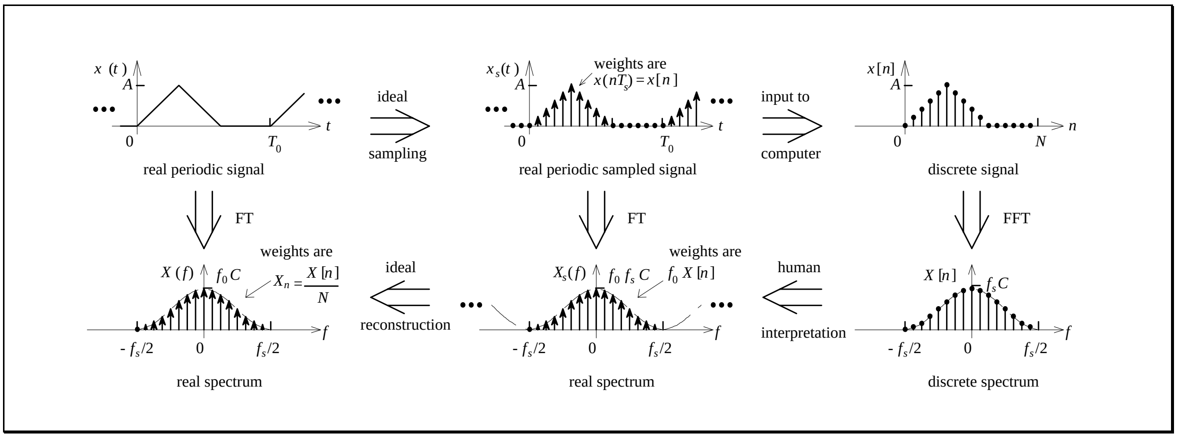


Figure F.10

Appendix B - The Phase-Locked Loop

Phase-locked loop. Voltage controlled oscillator. Phase Detector. Loop Filter..

Overview

The phase-locked loop (PLL) is an important building block for many electronic systems. PLLs are used in frequency synthesisers, demodulators, clock multipliers and many other communications and electronic applications.

Distortion in Synchronous AM Demodulation

In suppressed carrier amplitude modulation schemes, the receiver requires a local carrier for synchronous demodulation. Ideally, the local carrier must be in frequency and phase synchronism with the incoming carrier. Any discrepancy in the frequency or phase of the local carrier gives rise to distortion in the demodulator output.

For DSB-SC modulation, a constant *phase error* will cause attenuation of the output signal. Unfortunately, the phase error may vary randomly with time. A *frequency error* will cause a "beating" effect (the output of the demodulator is the original message multiplied by a low-frequency sinusoid). This is a serious type of distortion.

Frequency and phase errors cause distorted demodulator output

For SSB-SC modulation, a *phase error* in the local carrier gives rise to a phase distortion in the demodulator output. Phase distortion is generally not a problem with voice signals because the human ear is somewhat insensitive to phase distortion – it changes the quality of the speech but still remains intelligible. In video signals and data transmission, phase distortion is usually intolerable. A *frequency error* causes the output to have a component which is slightly shifted in frequency. For voice signals, a frequency shift of 20 Hz is tolerable.

Carrier Regeneration

A *pilot* is transmitted to enable the receiver to generate a local oscillator in frequency and phase synchronism with the transmitter

The human ear can tolerate a drift between the carriers of up to about 30 Hz. Quartz crystals can be cut for the same frequency at the transmitter and receiver, and are very stable. However, at high frequencies (> 1 MHz), even quartz-crystal performance may not be adequate. In such a case, a carrier, or *pilot*, is transmitted at a reduced level (usually about -20 dB) along with the sidebands.

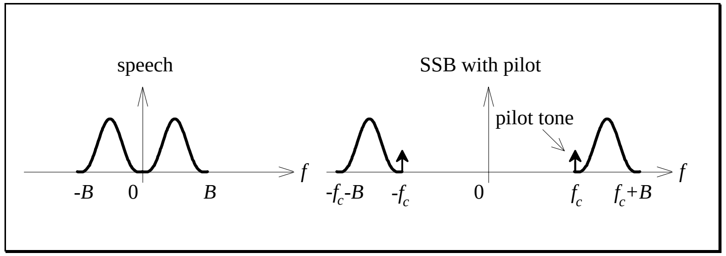


Figure P.1

One conventional technique to generate the receiver's local carrier is to separate the pilot at the receiver by a very narrowband filter tuned to the pilot frequency. The pilot is then amplified and used to synchronize the local oscillator.

A PLL is used to regenerate a local carrier

In demodulation applications, the PLL is primarily used in tracking the phase and frequency of the carrier of an incoming signal. It is therefore a useful device for synchronous demodulation of AM signals with a suppressed carrier or with a little carrier (the pilot). In the presence of strong noise, the PLL is more effective than conventional techniques.

For this reason, the PLL is used in such applications as space-vehicle-to-earth data links, where there is a premium on transmitter weight; or where the loss along the transmission path is very large; and, since the introduction of CMOS circuits that have entire PLLs built-in, commercial FM receivers.

The Phase-Locked Loop (PLL)

A block diagram of a PLL is shown below:

A phase-locked loop

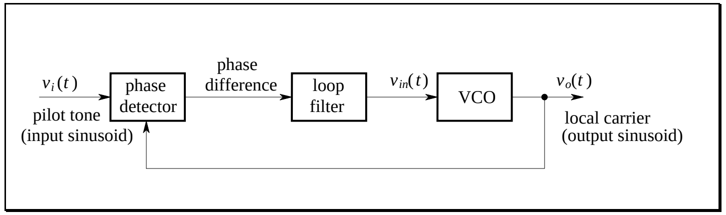


Figure P.2

It can be seen that the PLL is a feedback system. In a typical feedback system, the signal fed back tends to follow the input signal. If the signal fed back is not equal to the input signal, the difference (the error) will change the signal fed back until it is close to the input signal. A PLL operates on a similar principle, expect that the quantity fed back and compared is not the amplitude, but the *phase* of a sinusoid. The voltage controlled oscillator (VCO) adjusts its frequency until its phase angle comes close to the angle of the incoming signal. At this point, the frequency and phase of the two signals are in synchronism (expect for a difference of 90°, as will be seen later).

The three components of the PLL will now be examined in detail.

Voltage Controlled Oscillator (VCO)

The voltage controlled oscillator (VCO) is a device that produces a constant amplitude sinusoid at a frequency determined by its input voltage. For a fixed DC input voltage, the VCO will produce a sinusoid of a fixed frequency. The purpose of the control system built around it is to change the input to the VCO so that its output tracks the incoming signal's frequency and phase.

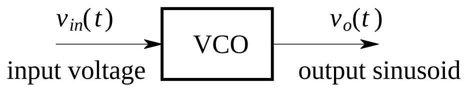


Figure P.3

The characteristic of a VCO is shown below:

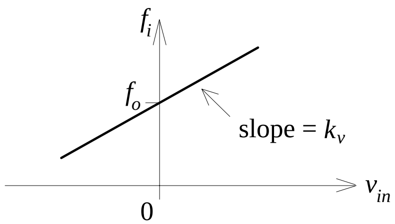


Figure P.4

The horizontal axis is the applied input voltage, and the vertical axis is the frequency of the output sinusoid. The amplitude of the output sinusoid is fixed. We now seek a model of the VCO that treats the output as the *phase* of the sinusoid rather than the sinusoid itself. The frequency f_o is the nominal frequency of the VCO (the frequency of the output for no applied input).

The instantaneous VCO frequency is given by:

$$\boxed{f_i(t) = f_o + k_v v_{in}(t)} \tag{P.1}$$

To relate this to the phase of the sinusoid, we need to generalise our definition of phase. In general, the frequency of a sinusoid $\cos(\theta) = \cos(2\pi f_1 t)$ is proportional to the rate of change of the phase angle:

$$\frac{d\theta}{dt} = 2\pi f_1$$

$$f_1 = \frac{1}{2\pi}\frac{d\theta}{dt} \tag{P.2}$$

We are used to having a constant f_1 for a sinusoid's frequency. The instantaneous phase of a sinusoid is given by:

$$\theta = \int_{-\infty}^{t} 2\pi f_1 d\tau \tag{P.3}$$

which reduces to:

$$\theta = 2\pi f_1 t \tag{P.4}$$

for f_1 a constant. For $f_i(t) = f_o + k_v v_{in}(t)$, the phase is:

$$\boxed{\begin{aligned}\theta(t) &= \int_{-\infty}^{t} \left[2\pi f_o + 2\pi k_v v_{in}(\tau)\right] d\tau \\ &= 2\pi f_o t + \int_{-\infty}^{t} 2\pi k_v v_{in}(\tau) d\tau\end{aligned}} \tag{P.5}$$

Therefore, the VCO output is:

$$v_o = A_o \cos\left(\omega_o t + \theta_o(t)\right) \tag{P.6}$$

where:

$$\boxed{\theta_o(t) = 2\pi k_v \int_{-\infty}^{t} v_{in}(\tau)\, d\tau} \tag{P.7}$$

This equation expresses the relationship between the input to the VCO and the phase of the resulting output sinusoid.

Example

Suppose a DC voltage of V_{DC} V is applied to the VCO input. Then the output phase of the VCO is given by:

$$\theta_o(t) = 2\pi k_v \int_{-\infty}^{t} V_{DC}\, d\tau = 2\pi k_v V_{DC} t$$

The resulting sinusoidal output of the VCO can then be written as:

$$\begin{aligned} v_o &= A_o \cos\left(\omega_o t + 2\pi k_v V_{DC} t\right) \\ &= A_o \cos\left(\left[\omega_o + 2\pi k_v V_{DC}\right] t\right) \\ &= A_o \cos\left(\omega_1 t\right) \end{aligned}$$

In other words, a constant DC voltage applied to the input of the VCO will produce a sinusoid of fixed frequency, $f_1 = f_0 + k_v V_{DC}$.

When used in a PLL, the VCO input should eventually be a constant voltage (the PLL has locked onto the phase, but the VCO needs a constant input voltage to output the tracked frequency).

Phase Detector

The phase detector is a device that produces the phase difference between two input sinusoids:

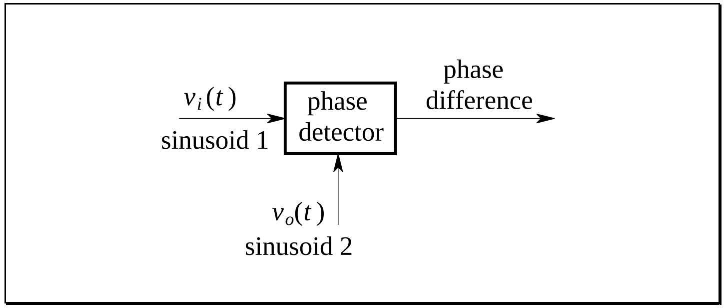


Figure P.5

A practical implementation of a phase detector is a *four-quadrant* multiplier:

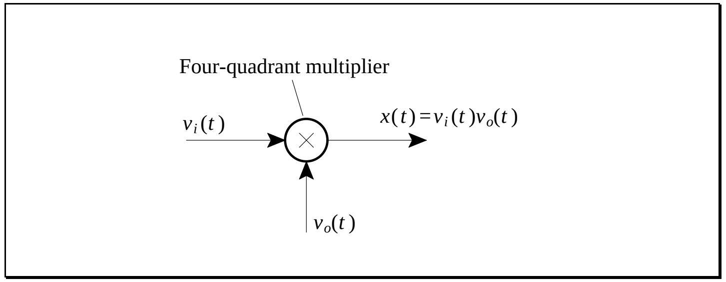


Figure P.6

To see why a multiplier can be used, let the incoming signal of the PLL, with constant frequency and phase, be:

$$v_i = A_i \sin\left(\omega_i t + \phi_i\right) \tag{P.8}$$

If we let:

$$\theta_i = \left(\omega_i - \omega_o\right) t + \phi_i \tag{P.9}$$

then we can write the input in terms of the nominal frequency of the VCO as:

$$v_i = A_i \sin\left(\omega_o t + \theta_i(t)\right) \tag{P.10}$$

Note that the incoming signal is in *phase quadrature* with the VCO output (i.e. one is a sine, the other a cosine). This comes about due to the way the multiplier works as a phase comparator, as will be shown shortly.

Thus, the PLL will lock onto the incoming signal but it will have a 90° phase difference.

The output of the multiplier is:

$$\begin{aligned} x &= v_i \times v_o \\ &= A_i \sin\left(\omega_o t + \theta_i\right) A_o \cos\left(\omega_o t + \theta_o\right) \\ &= \frac{A_i A_o}{2}\left[\sin\left(\theta_i - \theta_o\right) + \sin\left(2\omega_o t + \theta_i + \theta_o\right)\right] \end{aligned} \tag{P.11}$$

If we now look forward in the PLL block diagram, we can see that this signal passes through the "loop filter". If we assume that the loop filter is a lowpass filter that adequately suppresses the high frequency term of the above equation (not necessarily true in all cases!), then the phase detector output can be written as:

$$\boxed{x = \frac{A_i A_o}{2}\sin\left(\theta_i - \theta_o\right) = K_0 \sin\left(\theta_i - \theta_o\right)} \tag{P.12}$$

where $K_0 = A_i A_o / 2$. Thus, the multiplier used as a phase detector produces an output voltage that is proportional to the sine of the phase difference between the two input sinusoids.

PLL Model

A model of the PLL, in terms of *phase*, rather than *voltages*, is shown below:

A PLL model

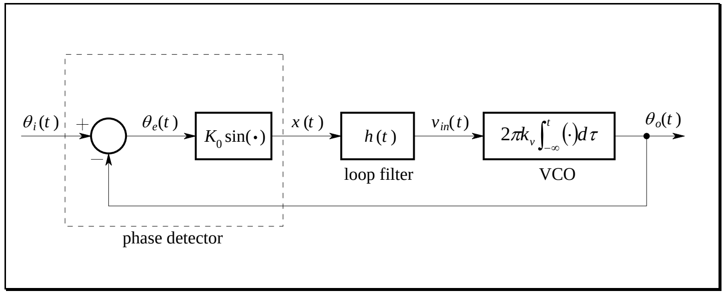


Figure P.7

Linear PLL Model

If the PLL is close to "lock", i.e. its frequency and phase are close to that of the incoming signal, then we can linearize the model above by making the approximation $\sin\theta_e \approx \theta_e$. With a linear model, we can convert to the *s*-domain and do our analysis with familiar block diagrams:

A linear PLL model in the *s*-domain

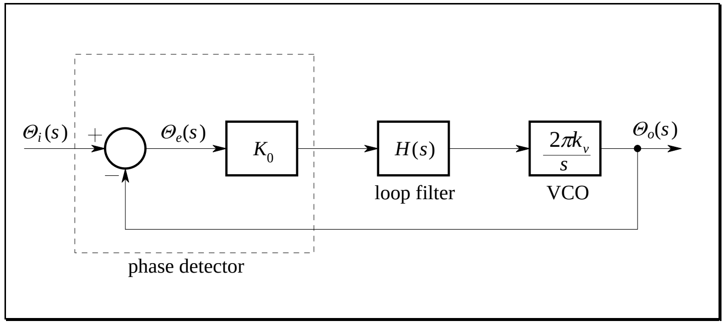


Figure P.8

Note that the integrator in the VCO in the time-domain becomes $1/s$ in the block diagram thanks to the integration property of the Laplace transform.

Reducing the block diagram gives:

The PLL transfer function for phase signals

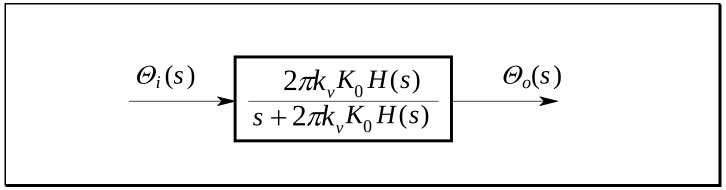


Figure P.9

This is the *closed-loop* transfer function relating the VCO output phase and the incoming signal's phase.

Loop Filter

The loop filter is designed to meet certain control system performance requirements. Once of those requirements is for the PLL to track input sinusoids with constant frequency and phase errors. That is, if the input phase is given by:

$$\theta_i(t) = (\omega_i - \omega_o)t + \phi_i \tag{P.13}$$

then we want:

$$\lim_{t \to \infty} \theta_e(t) = 0 \tag{P.14}$$

The analysis is best performed in the *s*-domain. The Laplace transform of the input signal is:

$$\Theta_i(s) = \frac{\omega_i - \omega_o}{s^2} + \frac{\phi_i}{s} \tag{P.15}$$

The Laplace transform of the error signal is:

$$\begin{aligned}\Theta_e(s) &= \Theta_i(s) - \Theta_o(s) \\ &= [1 - T(s)]\Theta_i(s) \\ &= \frac{s}{s + 2\pi k_v K_0 H(s)} \Theta_i(s) \\ &= \frac{s}{s + 2\pi k_v K_0 H(s)} \left[\frac{\omega_i - \omega_o}{s^2} + \frac{\phi_i}{s} \right]\end{aligned} \tag{P.16}$$

The final value theorem then gives:

$$\begin{aligned}&\lim_{t \to \infty} \theta_e(t) = \lim_{s \to 0} s\Theta_e(s) \\ &= \lim_{s \to 0} \frac{s^2}{s + 2\pi k_v K_0 H(s)} \left[\frac{\omega_i - \omega_o}{s^2} + \frac{\phi_i}{s} \right] \\ &= \left. \frac{\omega_i - \omega_o}{s + 2\pi k_v K_0 H(s)} \right|_{s=0} + \left. \frac{\phi_i s}{s + 2\pi k_v K_0 H(s)} \right|_{s=0}\end{aligned} \tag{P.17}$$

To satisfy Eq. (P.14) we must have:

$$\lim_{s \to 0} H(s) = \infty \tag{P.18}$$

We can therefore choose a simple PI loop filter to meet the steady-state constraints:

$$\boxed{H(s) = a + \frac{b}{s}} \tag{P.19}$$

Obviously, the constants a and b must be chosen to meet other system requirements, such as overshoot and bandwidth.

FFT - Quick Reference Guide

Definitions

Symbol	Description
T_0	time-window
$f_0 = 1/T_0$	discrete frequency spacing
f_s	sample rate
$T_s = 1/f_s$	sample period
N	number of samples

Creating FFTs

Given	Choose	Then
T_0 - time-window	f_s - sample rate	$N = T_0 f_s$
	N - number of samples	$f_s = Nf_0$
N - number of samples	f_s - sample rate	$T_0 = NT_s$
	T_0 - time-window	$f_s = Nf_0$
f_s - sample rate	N - number of samples	$T_0 = NT_s$
	T_0 - time-window	$N = T_0 f_s$

MATLAB® Code

Code	Description

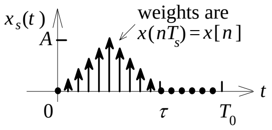
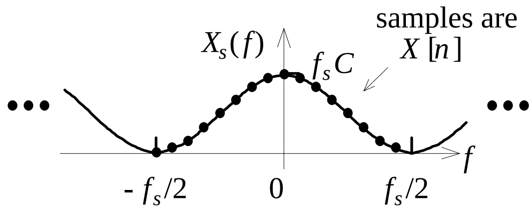
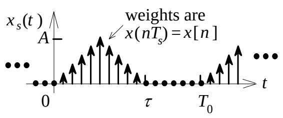
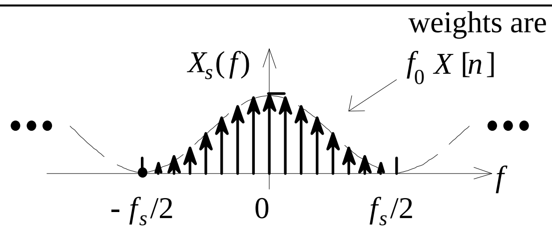
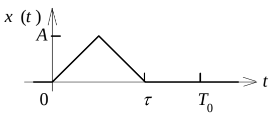
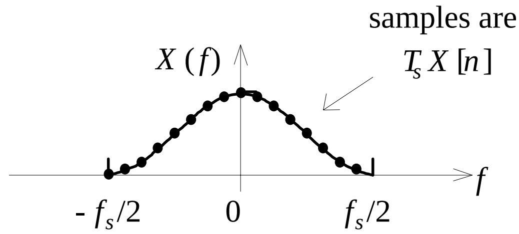
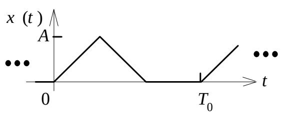
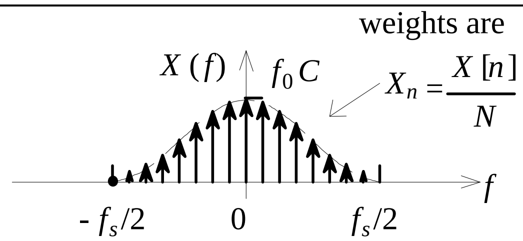
`% Sample rate`
`fs=1e6;`
`Ts=1/fs;`
 	
`% Number of samples`
`N=1024;`
	This code starts off with a given sample rate, and chooses to restrict the number of samples to 1024 for computational speed.

`% Time window and fundamental`
`T0=N*Ts;`
`f0=1/T0;`
 	
`% Time vector for specified DSO parameters`
`t=0:Ts:T0-Ts;`
	The time window takes samples up to **but not including** the point at $t = T_0$ (since this sample will be the same as the one at $t = 0$).

`% Frequency vector for specified DSO parameters` `f=-fs/2:f0:fs/2-f0;`	The corresponding frequency vector for the chosen sample rate and sample number. Note that because of the periodicy of the DFT, it **does not produce** a spectrum sample at $f = f_s/2$ since this corresponds to the sample at $f = -f_s/2$.

FFT.2

Interpreting FFTs

Case	Derivation of $x[n]$	Interpretation of $X[n]$	Action
1. "one-off" ideally sampled waveform 	Sample **<u>weights</u>**.	$X[n]$ gives **<u>values</u>** of one period of the true continuous FT at frequencies nf_0.	$X_s(f)=\operatorname{sinc}\left(\frac{f}{f_0}\right)*\sum_{n=-\infty}^{\infty}X[n]\delta(f-nf_0)$ 
2. periodic ideally sampled waveform (period T_0) 	Sample **<u>weights</u>** over **<u>one</u>** period.	Multiply $X[n]$ by f_0 to give **<u>weights</u>** of impulses at frequencies nf_0 in one period of true FT.	
3. "one-off" continuous waveform 	**<u>Values</u>** of waveform at intervals T_s.	Multiply $X[n]$ by T_s to give **<u>values</u>** of true continuous FT at nf_0.	$X(f)=\operatorname{sinc}\left(\frac{f}{f_0}\right)*\sum_{n=0}^{N-1}T_sX[n]\delta(f-nf_0)$ 
4. periodic continuous waveform (period T_0) 	**<u>Values</u>** of waveform over **<u>one</u>** period at intervals T_s.	Multiply $X[n]$ by $1/N$ to give **<u>weights</u>** of impulses at frequencies nf_0 in true FT.	

<u>Note</u>: Periodic waveforms should preferably be sampled so that an integral number of samples extend over an integral number of periods. If this condition is not met, then the *periodic extension* of the signal assumed by the DFT will have discontinuities and produce "spectral leakage". In this case, the waveform is normally *windowed* to ensure it goes to zero at the ends of the period – this will create a smooth periodic extension, but give rise to *windowing artefacts* in the spectrum.

MATLAB® - Quick Reference Guide

General

Code	Description
`Ts=0.256;`	Assigns the 1 x 1 matrix `Ts` with the value 0.256. The semicolon prevents the matrix being displayed after it is assigned.
`t=0:Ts:T;`	Assigns the vector `t` with values, starting from 0, incrementing by `Ts`, and stopping when `T` is reached or exceeded.
`N=length(t);`	`N` will equal the number of elements in the vector `t`.
`r=ones(1,N);`	`r` will be a `1` x `N` matrix filled with the value 1. Useful to make a step.
`n=45000*[1 18 900];`	Creates a vector (`[]`) with elements `1`, `18` and `900`, then scales all elements by `45000`. Useful for creating vectors of transfer function coefficients.
`wd=sqrt(1-zeta ^2)*wn;`	Typical formula, showing the use of `sqrt`; taking the square (`^2`); and multiplication with a scalar (`*`).
`gs=p.*g;`	Performs a vector multiplication (`.*`) on an element-by-element basis. Note that `p*g` will be undefined.

Graphing

Code	Description
`figure(1);`	Creates a new graph, titled "Figure 1".
`plot(t,y);`	Graphs `y` vs. `t` on the current figure.
`subplot(211);`	Creates a `2` x `1` matrix of graphs in the current figure. Makes the current figure the top one.
`title('Complex poles');`	Puts a title `'Complex poles'` on the current figure.
`xlabel('Time (s)');`	Puts the label `'Time (s)'` on the x-axis.
`ylabel('y(t)');`	Puts the label `'y(t)'` on the y-axis.
`semilogx(w,H,'k:');`	Makes a graph with a logarithmic x-axis, and uses a black dotted line (`'k:'`).
`plot(200,-2.38,'kx');`	Plots a point (`200,-2.38`) on the current figure, using a black cross (`'kx'`).
`axis([1 1e5 -40 40]);`	Sets the range of the x-axis from `1` to `1e5`, and the range of the y-axis from `-40` to `40`. Note all this information is stored in the vector `[1 1e5 -40 40]`.
`grid on;`	Displays a grid on the current graph.
`hold on;`	Next time you plot, it will appear on the current graph instead of a new graph.

M.2

Frequency-domain

Code	Description
`f=logspace(f1,f2,100);`	Creates a logarithmically spaced vector from `f1` to `f2` with 100 elements.
`H=freqs(n,d,w);`	`H` contains the frequency response of the transfer function defined by numerator vector `n` and denominator vector `d`, at frequency points `w`.
`Y=fft(y);`	Performs a fast Fourier transform (FFT) on `y`, and stores the result in `Y`.
`Hmag=abs(H);`	Takes the magnitude of a complex number.
`Hang=angle(H);`	Takes the angle of a complex number.
`X=fftshift(X);`	Swaps halves of the vector `X` - useful for displaying the spectrum with a negative frequency component.

Time-domain

Code	Description
`step(Gcl);`	Gives the step response of the system transfer function `Gcl`.
`y=conv(m2,Ts*h);`	Performs a convolution on `m2` and `Ts*h`, with the result stored in `y`.
`y=y(1:length(t));`	Reassigns the vector `y` by taking elements from position `1` to position `length(t)`. Normally used after a convolution, since convolution produces end effects which we usually wish to ignore for steady-state analysis.
`square(2*pi*fc*t,10)`	Creates a square wave from –1 to +1, of frequency `fc`, with a `10` percent duty cycle. Useful for generating a real sampling waveform.

Control

Code	Description
`Gmrv=tf(Kr,[Tr 1]);`	Creates the transfer function $G_{mrv} = \frac{K_r}{1+sT_r}$.
`I=tf(1,[1 0]);`	Creates the transfer function $I = \frac{1}{s}$, an integrator.
`Gcl=feedback(Gol,H);`	Creates the transfer function $G_{cl} = \frac{G_{ol}}{1+G_{ol}H}$.
`rlocus(Gol);`	Makes a root locus using the system `Gol`.
`K=rlocfind(Gol);`	An interactive command that allows you to position the closed-loop poles on the root locus. It returns the value of `K` that puts the roots at the chosen location.
`Gd=c2d(G,Ts,'tustin');`	Creates a discrete-time equivalent system `Gd` of the continuous-time system `G`, using a sample rate of `Ts` and the bilinear (`'tustin'`) method of discretization.

Matrices - Quick Reference Guide

Definitions

Symbol	Description
a_{ij}	Element of a matrix. i is the row, j is the column.
$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \left[a_{ij}\right]$	$\mathbf{A}$ is the representation of the matrix with elements a_{ij}.
$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$	$\mathbf{x}$ is a column vector with elements x_j.
$\mathbf{0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	Null matrix, every element is zero.
$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	Identity matrix, diagonal elements are one.
$\lambda\mathbf{I} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$	Scalar matrix.
$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$	Diagonal matrix, $a_{ij} = 0 \left(i \neq j\right)$.

Multiplication

Multiplication	Description
$\mathbf{Z} = k\mathbf{Y}$	Multiplication by a scalar: $z_{ij} = k y_{ij}$
$\mathbf{z} = \mathbf{A}\mathbf{x}$	Multiplication by a vector: $z_j = \sum_{k=1}^{n} a_{ik} x_k$
$\mathbf{Z} = \mathbf{A}\mathbf{B}$	Matrix multiplication: $z_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.
$\mathbf{AB} \neq \mathbf{BA}$	In general, matrix multiplication is not commutative.

B.2

Operations

Terminology	Description
$\mathbf{A}^t = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$	Transpose of **A** (interchange rows and columns): $a_{ij}^t = a_{ji}$.
$\lvert\mathbf{A}\rvert = \det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$	Determinant of **A**. If $\lvert\mathbf{A}\rvert = 0$, then **A** is *singular*. If $\lvert\mathbf{A}\rvert \neq 0$, then **A** is *non-singular*.
$\mathrm{a}_{ij} = \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ . & . & . & . & . \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ . & . & . & . & . \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$	Minor of a_{ij}. Delete the row and column containing the element a_{ij} and obtain a new determinant.
$A_{ij} = (-1)^{i+j}\, \mathrm{a}_{ij}$	Cofactor of a_{ij}.
$\operatorname{adj} \mathbf{A} = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$	Adjoint matrix of **A**. Replace every element a_{ij} by its cofactor in $\lvert\mathbf{A}\rvert$, and then transpose the resulting matrix.
$\mathbf{A}^{-1} = \dfrac{\operatorname{adj} \mathbf{A}}{\lvert\mathbf{A}\rvert}$	Reciprocal of **A**: $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Only exists if **A** is square and non-singular. Formula is only used for 3x3 matrices or smaller.

Linear Equations

Terminology	Description
$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$ $a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$ $a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$	Set of linear equations written explicitly.
$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$	Set of linear equations written using matrix elements.
$\mathbf{Ax} = \mathbf{b}$	Set of linear equations written using matrix notation.
$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$	Solution to set of linear equations.

Eigenvalues

Equations	Description
$\mathbf{Ax} = \lambda \mathbf{x}$	λ are the eigenvalues. **x** are the column eigenvectors.
$\lvert\mathbf{A} - \lambda \mathbf{I}\rvert = 0$	Finding eigenvalues.

Answers

1A.1

(a) $g(t) = \sin(t)\sum_{k=-\infty}^{\infty} \mathrm{rect}\left(\frac{t - \frac{\pi}{2} - 2\pi k}{\pi}\right)$, $T_0 = 2\pi$, $P = \frac{1}{4}$

(b) $g(t) = \sum_{k=-\infty}^{\infty}\left[2\,\mathrm{rect}(t-3k) + \left(\frac{t+1-3k}{\frac{1}{4}}\right)\mathrm{rect}\left(\frac{t+\frac{3}{4}-3k}{\frac{1}{2}}\right) + \left(\frac{t-1-3k}{-\frac{1}{4}}\right)\mathrm{rect}\left(\frac{t-\frac{3}{4}-3k}{\frac{1}{2}}\right)\right]$,

$T_0 = 3$, $P = \frac{16}{9}$

(c) $g(t) = \sum_{k=-\infty}^{\infty}\left[e^{-(t-10k)/10}(-1)^k\,\mathrm{rect}\left(\frac{t-5-10k}{10}\right)\right]$, $T_0 = 20$, $P = \frac{1}{2}\left(1-e^{-2}\right)$

(d) $g(t) = 2\cos(200\pi t)\sum_{k=-\infty}^{\infty}\mathrm{rect}\left(\frac{t-4k}{2}\right)$, $T_0 = 4$, $P = 1$

1A.2

(a) $g(t) = \left(\frac{t+2}{-\frac{1}{10}}\right)\mathrm{rect}\left(\frac{t+\frac{7}{4}}{\frac{1}{2}}\right) + \left(\frac{t+1}{\frac{1}{10}}\right)\mathrm{rect}\left(\frac{t+\frac{3}{4}}{\frac{3}{2}}\right) + \left(\frac{t-1}{-\frac{1}{10}}\right)\mathrm{rect}\left(\frac{t-\frac{3}{4}}{\frac{3}{2}}\right) + \left(\frac{t-2}{\frac{1}{10}}\right)\mathrm{rect}\left(\frac{t-\frac{7}{4}}{\frac{1}{2}}\right)$,

$E = 83\frac{1}{3}$

(b) $g(t) = \cos(t)\,\mathrm{rect}\left(\frac{t}{\pi}\right)$, $E = \frac{\pi}{2}$

(c) $g(t) = t^2\,\mathrm{rect}\left(t-\frac{1}{2}\right) + (t-2)^2\,\mathrm{rect}\left(t-\frac{5}{2}\right)$, $E = \frac{2}{5}$

(d) $g(t) = \mathrm{rect}\left(\frac{t-\frac{5}{2}}{5}\right) + \mathrm{rect}\left(\frac{t-\frac{5}{2}}{3}\right) + \mathrm{rect}\left(t-\frac{5}{2}\right)$, $E = 19$

1A.3

(i) 0 (ii) e^3 (iii) 5 (iv) $f(t_1 - t_2)$ (v) 1 (vi) $f(t-t_0)$

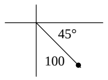
1A.4

Let $t = T\tau$. Then:

$$\begin{aligned}
\int_{-\infty}^{\infty} f(t)\,\delta\!\left(\frac{t-t_0}{T}\right)dt &= \int_{-\infty}^{\infty} f(T\tau)\,\delta(\tau - t_0/T)\,T\,d\tau \quad \text{if } T>0 \\
\text{or} \quad &= \int_{\infty}^{-\infty} f(T\tau)\,\delta(\tau - t_0/T)\,T\,d\tau \quad \text{if } T<0 \\
&= \int_{-\infty}^{\infty} f(T\tau)\,\delta(\tau - t_0/T)\,|T|\,d\tau \quad \text{all } T \\
&= |T|\,f(t_0) \quad \text{(using the sifting property)}
\end{aligned}$$

which is exactly the same result as if we started with $|T|\delta(t-t_0)$.

1A.5

(a) $\overline{X} = 27\angle 15°$ (b) $\overline{X} = 5\angle 10°$ (c) $\overline{X} =$ 

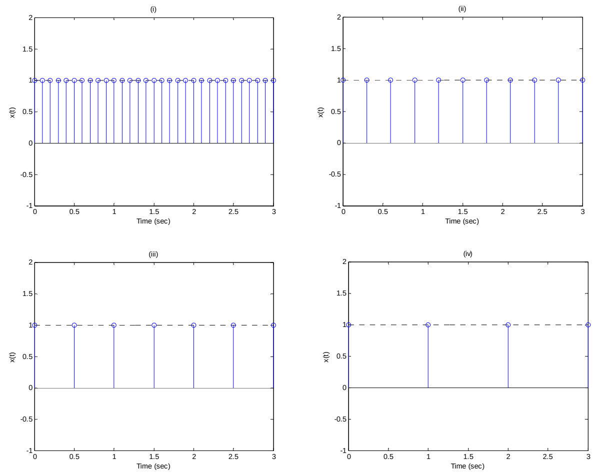
(d) $8\sqrt{3},\ 8$

(e) $x(t) = 29.15\cos(\omega t - 22.17°)$ (f) $x(t) = 100\cos(\omega t - 60°)$ (g) $\overline{X} = 5.596\angle 59.64°$

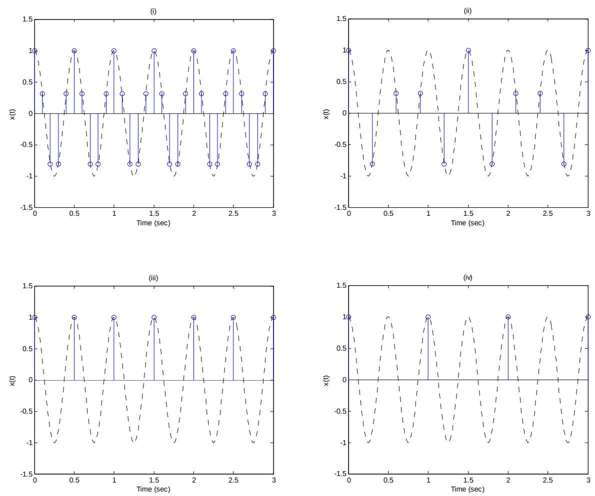
(h) $X = 1\angle 30°,\ X^* = 1\angle -30°$ (i) -2.445 (j) $X = 1.5\angle 30°,\ X^* = 1.5\angle -30°$

1B.1

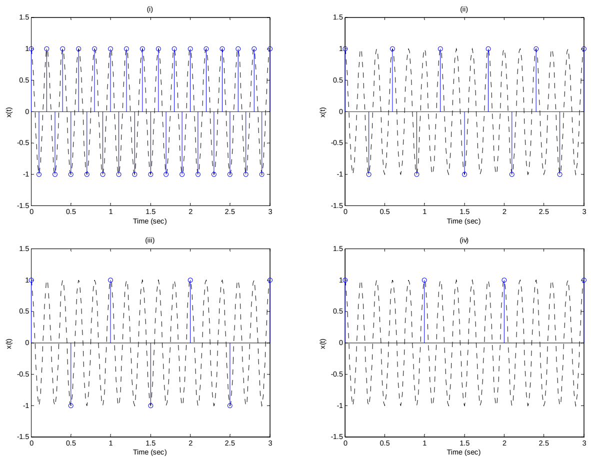
a)



b)



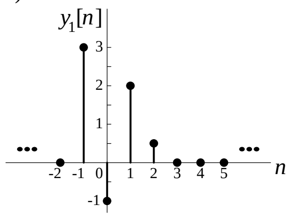
c)



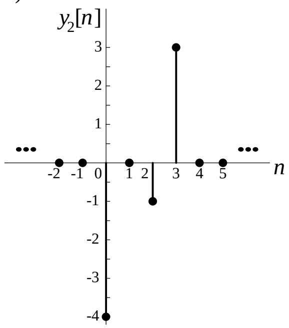
Sampling becomes more accurate as the sample time becomes smaller.

1B.2

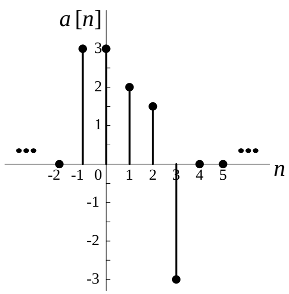
a)



b)



1B.3



1B.4

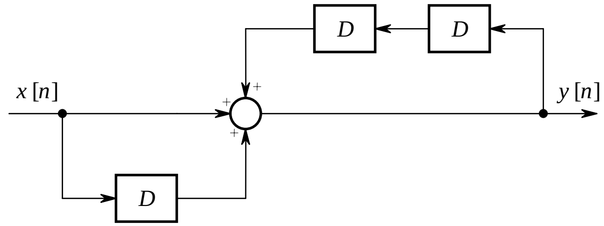
$$y[n]=\begin{cases}1, & n=1,3,5,\ldots\\ 0, & \text{all other } n\end{cases}$$

1B.5

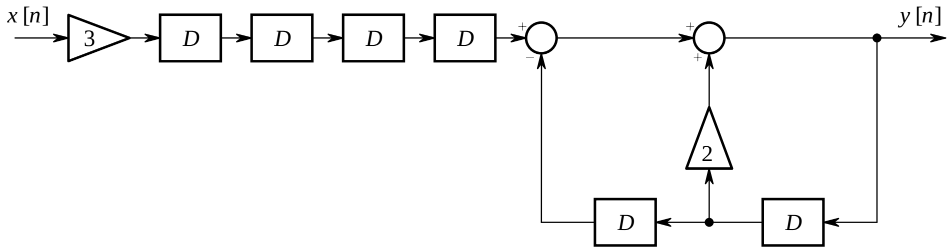
a) $y[n+2]=y[n+1]+y[n]$

1B.6

(i)



(ii)



1B.7

(i) $y[n]=y[n-1]-2y[n-2]+3x[n-1]$

(ii) $y[0]=3,\ y[1]=8,\ y[2]=-1,\ y[3]=-14$

1B.8

(a)

(i) $h[0]=\dfrac{T}{2}$, $h[n]=T$ for $n=1,2,3,\ldots$

(ii) $h[0]=1$, $h[n]=-0.25(0.5)^{n-1}$ for $n=1,2,3,\ldots$

(b) $y[-1]=1$, $y[0]=0.75$, $y[1]=-1.375$, $y[2]=1.0625$, $y[3]=-1.21875$

1B.9

(i) (a) $a_1=a_4=0$ (b) $a_5=0$

(ii) (a) $a_1=a_4=0$ (b) $a_5=0$

1B.10

$y_1[0]=1$, $y_1[1]=3$, $y_1[2]=7$, $y_1[3]=15$, $y_1[4]=31$, $y_1[5]=63$

$y[0]=4$, $y[1]=12$, $y[2]=26$, $y[3]=56$, $y[4]=116$, $y[5]=236$

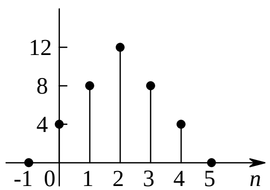
No, since they rely on superposition.

1B.11

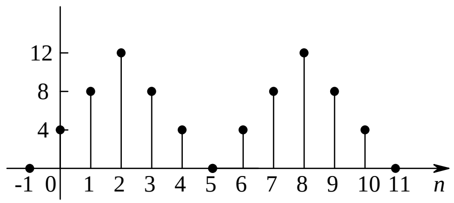
$y[n]=2\delta[n]+\delta[n-2]-2\delta[n-4]-\delta[n-6]$

1B.12

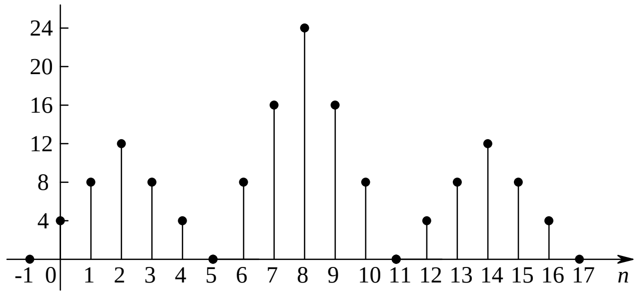
(i)



(ii)



(iii)



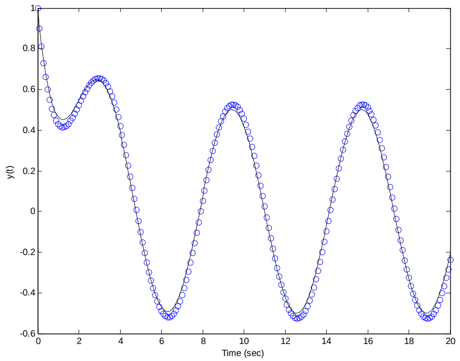
1B.14

The describing differential equation is:

$$\frac{d^2v_o}{dt^2}+\frac{R}{L}\frac{dv_o}{dt}+\frac{1}{LC}v_o=\frac{1}{LC}v_i\,.$$

The resulting discrete-time approximation is:

$$y[n]=(2-RT/L)y[n-1]-\left(1-RT/L+T^2/LC\right)y[n-2]+T^2/LC\,x[n-2]$$



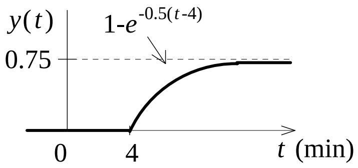
Use $T = 0.1$ so that the waveform appears smooth (any value smaller has a minimal effect in changing the solution).

1B.15

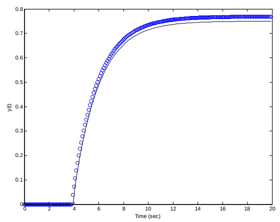
(a) $\dfrac{dy}{dt} + Ky = Kx$ (b) $h(t) = Ke^{-Kt}u(t)$

(c) $y(t) = 0.75\left[1 - e^{-0.5(t-4)}\right]u(t-4)$

(d)



(e)



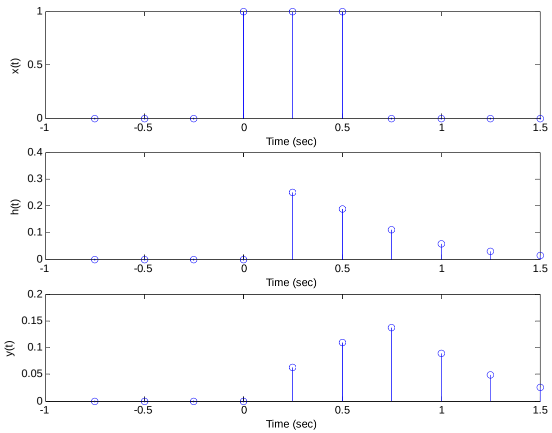
1B.16

The impulse response is $h(t) = te^{-t}u(t)$. Using numerical convolution will only give the ZSR. The ZIR needs to be obtained using other methods and is $e^{-t}u(t)$. The code snippet below shows relevant MATLAB® code:

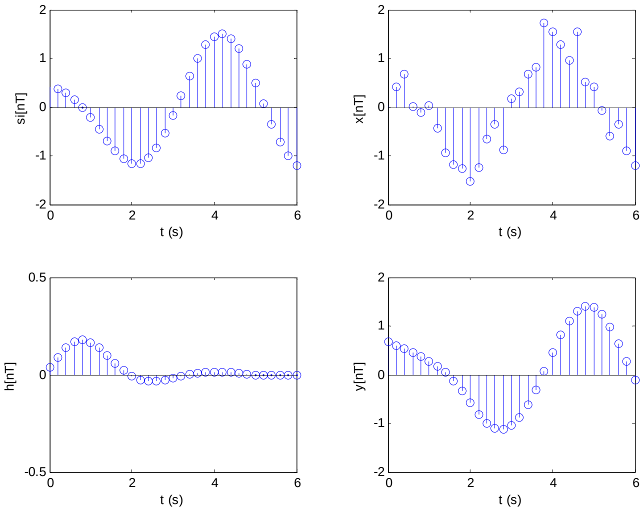
```
t=0:Ts:To-Ts;
h=t.*exp(-t);
x=sin(t);
yc=conv(x,h*Ts);        % ZSR
yc=yc(1:N)+exp(-t);     % total solution = ZSR + ZIR
```

1B.17

The output control signal is smoothed, since it is not changing as rapidly as the input control signal. It is also delayed in time.



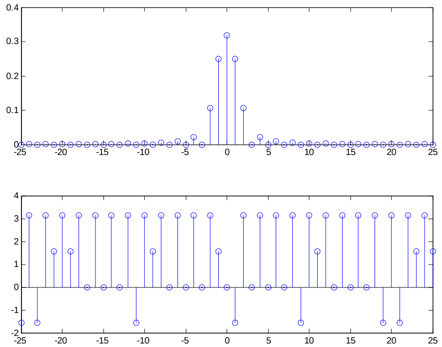
1B.18



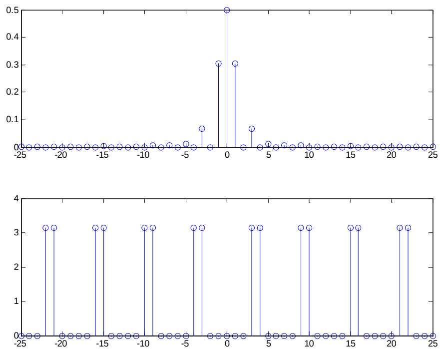
Note that the filter effectively removes the added noise; however, it also introduces a time delay of between three and four samples.

2A.1

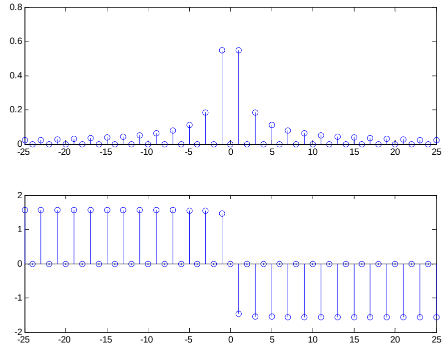
(a) $G_n = \frac{1}{4}(-j)^n \left[\operatorname{sinc}\left(\frac{n-1}{2}\right) + \operatorname{sinc}\left(\frac{n+1}{2}\right) \right]$



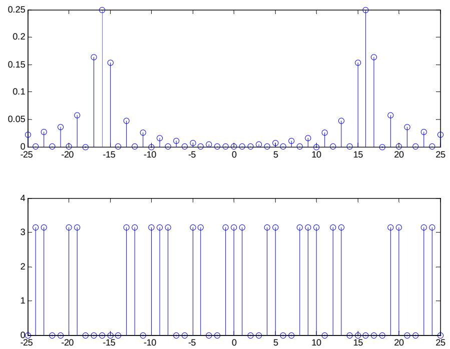
(b) $G_n = \frac{1}{2}\operatorname{sinc}(n/2)\operatorname{sinc}(n/6)$



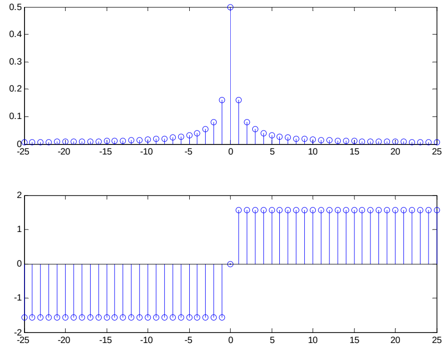
(c) $G_n = \frac{\left(e^{-\pi/10}+1\right)}{2\pi}\frac{\left(1-(-1)^n\right)}{0.1+jn}$



(d) $G_n = \frac{1}{4}\left[\operatorname{sinc}\left(\frac{n}{2}+8\right)+\operatorname{sinc}\left(\frac{n}{2}-8\right)\right]$

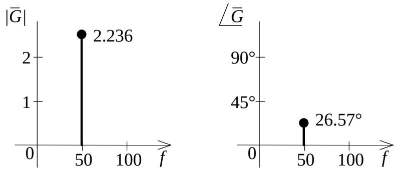


(e) $G_n = \frac{\operatorname{sinc}(n)(-1)^n - 1}{j2\pi n}$ and the special case of $n=0$ can be evaluated by applying l'Hospital's Rule or from first principles: $G_0 = \frac{1}{2}$.

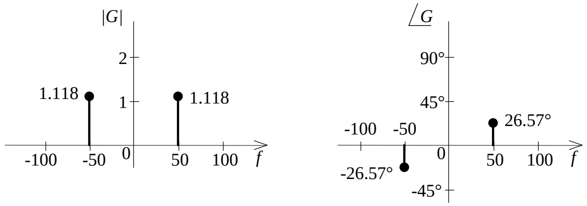


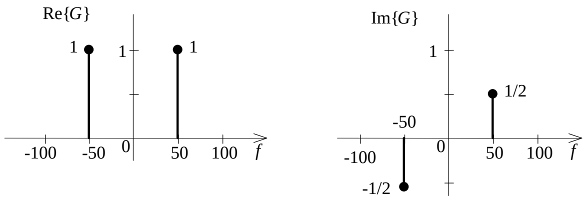
2A.2

(a)



(b)





(c)

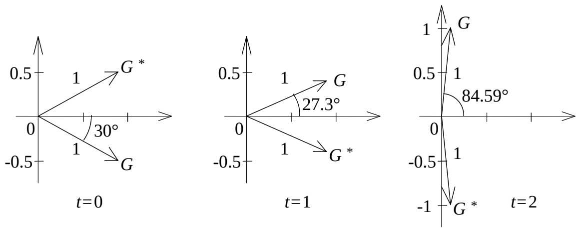
2A.3

$\overline{G} = 4\angle 45°$

2A.4

$g(t) = 4\cos(200\pi t)$

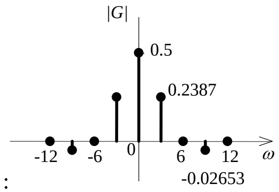
2A.5



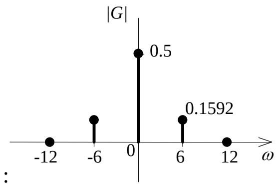
2A.6

$P = 4.294$ W

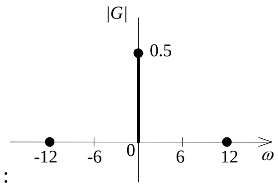
2A.7



(a) $T_0 = \dfrac{2\pi}{3}$: , $P = 73\%$



(b) $T_0 = \dfrac{\pi}{3}$: , $P = 60\%$



(c) $T_0 = \dfrac{\pi}{6}$: , $P = 50\%$

2A.8

$5/T_0$ Note: $G_n = A\operatorname{sinc}^2\left(\dfrac{n}{2}\right) - A\operatorname{sinc}(n)$

2B.1

Hint: $e^{-a|t|} = e^{at}u(-t) + e^{-at}u(t)$

2B.2

$3P(-1.5f)$

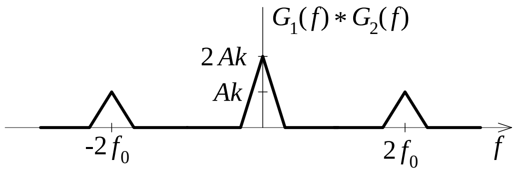
2B.3

This follows directly from the time shift property.

2B.4

$G_1(f) = \dfrac{a}{a + j2\pi f}$, $G_2(f) = \dfrac{1}{j2\pi f} + \dfrac{1}{2}\delta(f)$, $g_1(t) * g_2(t) = \left(1 - e^{-at}\right)u(t)$

2B.5



2B.6

(a) $\dfrac{5 - 10\cos(3\pi f) + 5\cos(4\pi f)}{\pi^2 f^2}$

(b) $\dfrac{\pi}{2}\operatorname{sinc}\left(\pi f - \tfrac{1}{2}\right) + \dfrac{\pi}{2}\operatorname{sinc}\left(\pi f + \tfrac{1}{2}\right)$

(c) $\dfrac{\left[(1 + j\pi f)e^{-j\pi f} - \operatorname{sinc}(f)\right]\cos(2\pi f)e^{-j3\pi f}}{\pi^2 f^2}$

(d) $\left[5\operatorname{sinc}(5f) + 3\operatorname{sinc}(3f) + \operatorname{sinc}(f)\right]e^{-j5\pi f}$

2B.7

(a) $2Af_0\operatorname{sinc}\left[2f_0(t - t_0)\right]$

(b) $2Af_0\operatorname{sinc}(f_0 t)\sin(\pi f_0 t)$

2B.8

$\dfrac{j2A\sin^2(\pi fT)}{\pi f}$

2B.9

$\dfrac{A\left[\operatorname{sinc}(4\mathrm{f}) - \cos(4\pi f)\right]}{j\pi f}$

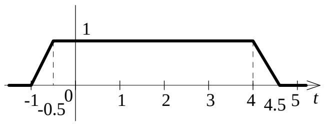
3A.1

$x(t) = 4\cos(2000\pi t - 30°)$, $P = 8$

3A.2

$$X(f) = \frac{\operatorname{sinc}(f)e^{j\pi f} - \operatorname{sinc}(2f)e^{-j6\pi f}}{j\pi f}$$

3A.3



3A.4

$0.2339\angle 0°$

3A.5

$X(f) = A\tau \operatorname{sinc}^2(f\tau)$

3A.6

$B/4$

3A.7

By passing the signal through a lowpass filter with 4 kHz cutoff - provided the original signal contained no spectral components above 4 kHz.

3A.8

Periodicy.

3A.9

(a) 20 Hz, 40 Hz, $P = 0.1325$ W

(b) $G_3 = 0.5e^{-j\frac{\pi}{3}}e^{j120\pi t} + 0.5e^{j\frac{\pi}{3}}e^{-j120\pi t}$, 0.9781

(c)

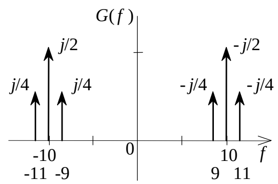
Harmonic #	Amplitude	Phase (°)
0	1	
1	3	-66
2	1	-102
3	0.5	-168
4	0.25	-234

Yes.

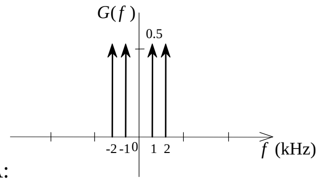
3A.10

Yes. The flat topped sampling pulses would simply reduce the amplitudes of the “repeats” of the baseband spectrum as one moved along the frequency axis. For ideal sampling, with impulses, all “repeats” have the same amplitude. Note that after sampling the pulses have tops which follow the original waveform.

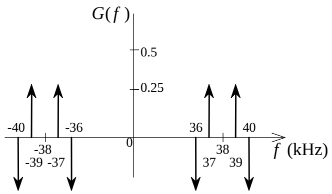
3A.11



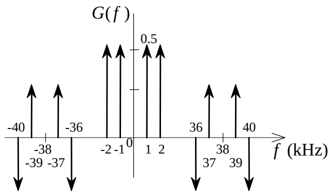
3A.12



A:

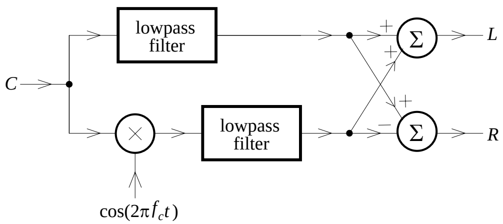


B:



C:

3A.13



3A.14

Truncating the 9.25 kHz sinusoid has the effect of convolving the impulse in the original transform with the transform of the window (a sinc function for a rectangular window). This introduces "leakage" which will give a spurious 9 kHz component.

3B.1

a) $\dfrac{1/RC}{s+1/RC}$ b) $\dfrac{(s+1/R_1C_1)(s+1/R_2C_2)}{s^2+(1/R_1C_1+1/R_2C_2+1/R_2C_1)s+1/R_1R_2C_1C_2}$

c) $\dfrac{1/2(s-R/L)}{s+R/L}$ d) $\dfrac{1/R_2C_1(s+R_1/L_1)}{s^2+(1/R_2C_1+R_1/L_1)s+((R_1+R_2)/R_2L_1C_1)}$

3B.2

a) $I(s)\left[sL+R+\dfrac{1}{sC}\right]-Li(0^-)=E(s)$

b) $X(s)\left[Ms^2+Bs+K\right]-M\left[sx(0^-)+\dfrac{dx(0^-)}{dt}\right]-Bx(0^-)=\dfrac{3}{s^2}$

c) $\Theta(s)\left[Js^2+Bs+K\right]-J\left[s\theta(0^-)+\dfrac{d\theta(0^-)}{dt}\right]-B\theta(0^-)=\dfrac{10\omega}{s^2+\omega^2}$

3B.3

a) $f(t)=\dfrac{5}{4}-\dfrac{\sqrt{7}}{2}e^{-t}\cos\left(\sqrt{3}t-\tan^{-1}\left(\dfrac{9}{\sqrt{3}}\right)\right)$

b) $f(t)=\dfrac{1}{3}(\cos(t)-\cos(2t))$

c) $f(t)=\dfrac{1}{40}\left[t^2+8t+\dfrac{15}{2}+e^{2t}\left(7t-\dfrac{15}{2}\right)\right]$

3B.4

a) $f(\infty)=5/4$ b), c) The final value theorem does not apply. Why?

3B.5

$$T(s)=\frac{-1/R_1C_2s}{(s+1/R_1C_1)(s+1/R_2C_2)}$$

3B.6

$$y(t) = \sqrt{5}e^{-t}\cos(2t + 26.6°)u(t)$$

3B.7

a) (i) $\dfrac{G}{1+GH}$ (ii) $\dfrac{1}{1+GH}$ (iii) $\dfrac{GH}{1+GH}$

b) All denominators are the same.

3B.8

a) $\dfrac{Y}{R} = \dfrac{G_1 - G_1G_3H_3 - G_1G_2H_2 + G_1G_2G_3H_2H_3 + G_1G_2 - G_1G_2G_3H_3}{1 - G_2H_2 + G_1G_3 - G_1G_2G_3H_2 + G_1G_2G_3 - G_3H_3 + G_2G_3H_2H_3}$

b) $X_5 = \left(\dfrac{ab}{1-bd-ac}\right)X_1 + \left(\dfrac{ac-1}{1-bd-ac}\right)Y$

3B.9

a) $\dfrac{C}{R} = \dfrac{G_1G_2G_3G_4}{1 - G_3G_4H_1 + G_2G_3H_2 + G_1G_2G_3G_4H_3}$

b) $\dfrac{Y}{X} = \dfrac{AE + CE - CD - AD}{1 - AB + EF - ABEF}$

4A.1

Yes, by examining the pole locations for R, L, $C > 0$.

4A.2

(i) $y(t) = 2\left[1 - e^{-4t}\right]u(t)$

(ii) $y(t) = \left[2t - \frac{1}{2} + \frac{1}{2}e^{-4t}\right]u(t)$

(iii) $y(t) = \left[\dfrac{8}{\sqrt{5}}\cos\left(2t + \tan^{-1}(2) + \pi\right) + \dfrac{8}{5}e^{-4t}\right]u(t)$

(iv) $y(t) = \left[\dfrac{8}{\sqrt{29}}\cos\left(10t + \tan^{-1}\left(\dfrac{2}{5}\right) + \pi\right) + \dfrac{40}{29}e^{-4t}\right]u(t)$

4A.3

(i) underdamped, $y(t) = \left[2 - \dfrac{4}{\sqrt{3}}e^{-2t}\sin\left(2\sqrt{3}t + \cos^{-1}(0.5)\right)\right]u(t)$,

$h(t) = \dfrac{16}{\sqrt{3}}e^{-2t}\sin\left(2\sqrt{3}t\right)$

(ii) critically damped, $y(t) = \left[2 - 2(1+4t)e^{-4t}\right]u(t)$, $h(t) = 32te^{-4t}$

(iii) overdamped, $y(t) = \left[2 - \frac{8}{3}e^{-2t} + \frac{2}{3}e^{-8t}\right]u(t)$, $h(t) = \frac{16}{3}\left(e^{-2t} - e^{-8t}\right)$

4A.4

(a) $y_{ss}(t) = u(t)$ (b) $y_{ss}(t) = \dfrac{10\sqrt{170}}{17}\cos\left(t + \tan^{-1}\left(\dfrac{1}{13}\right)\right)u(t)$

4B.1

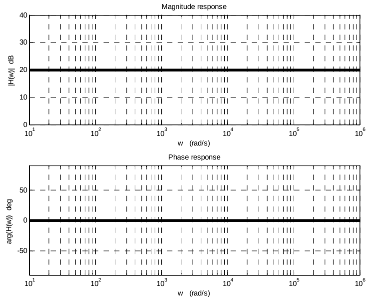
a) 2 kHz b) 2.5 Hz

4B.2

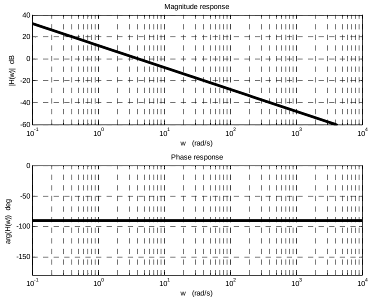
a) 0 dB b) 32 dB c) –6 dB

4B.3

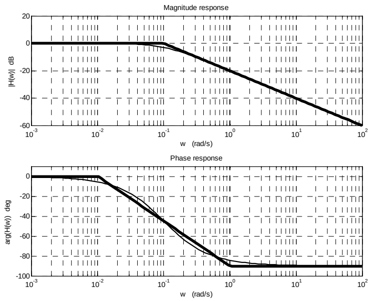
a)



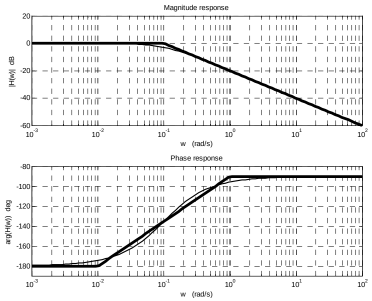
b)



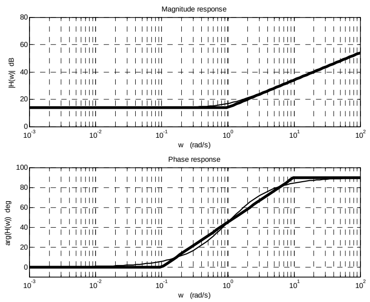
c)



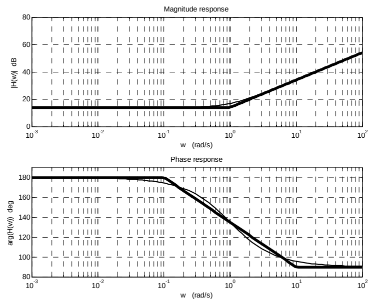
d)



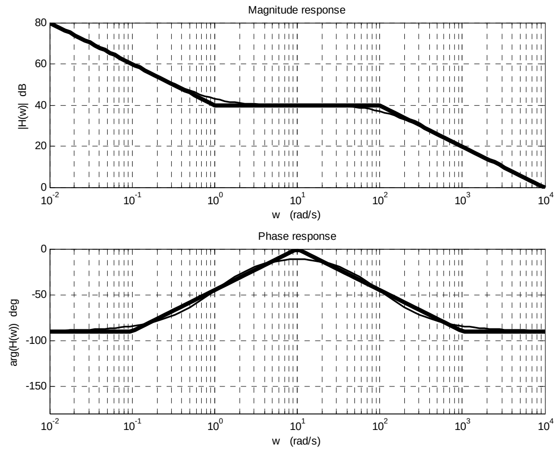
e)



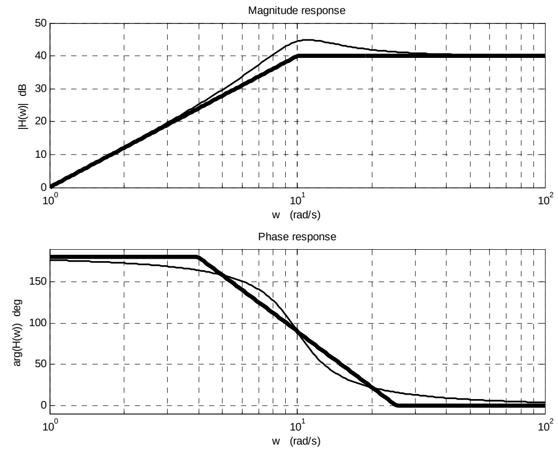
f)



4B.5

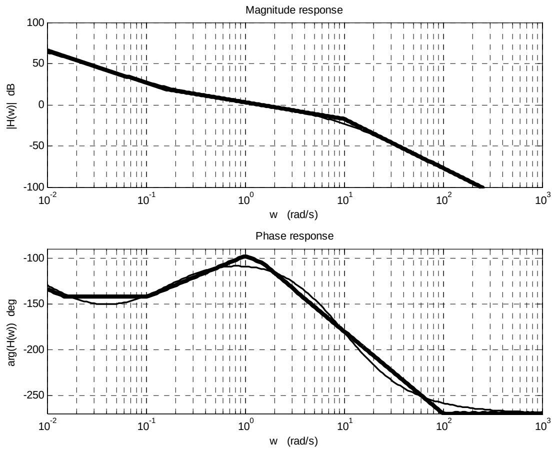


4B.6



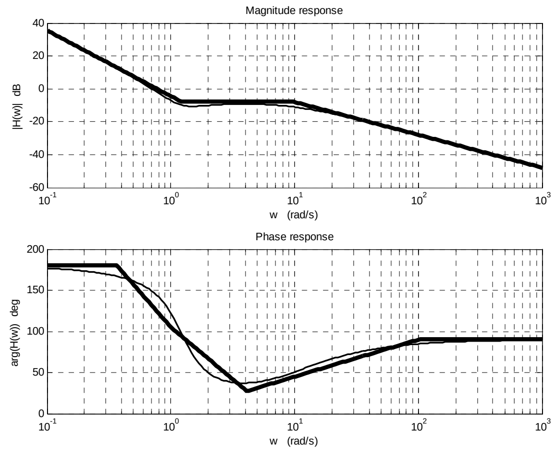
4B.7

a)



(i) +28 dB, -135° (ii) +5 dB, -105° (iii) –15 dB, -180° (iv) – 55 dB, -270°

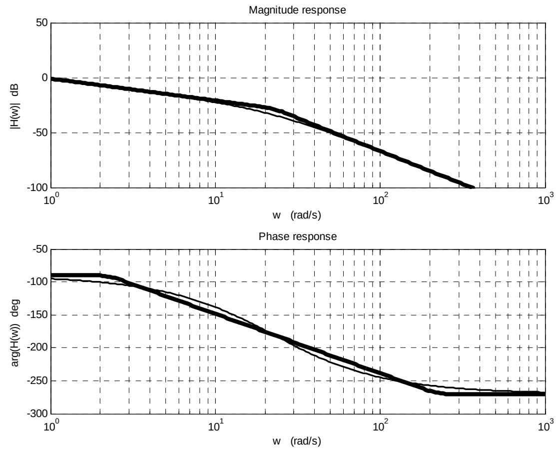
b)



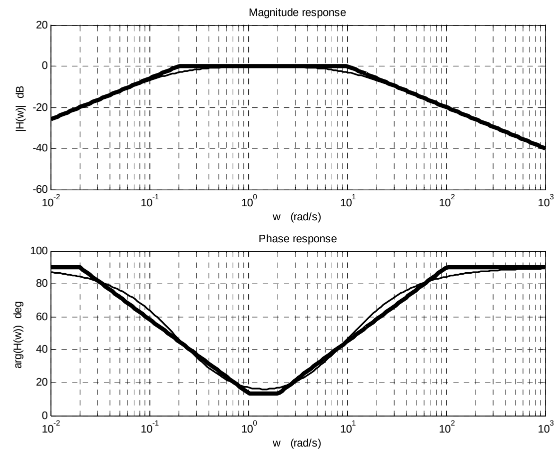
(i) +35 dB, +180° (ii) –5 dB, +100° (iii) –8 dB, +55° (iv) –28 dB, +90°

4B.8

a) $G_1(s) = \dfrac{450}{s(s+25)(s+20)}$

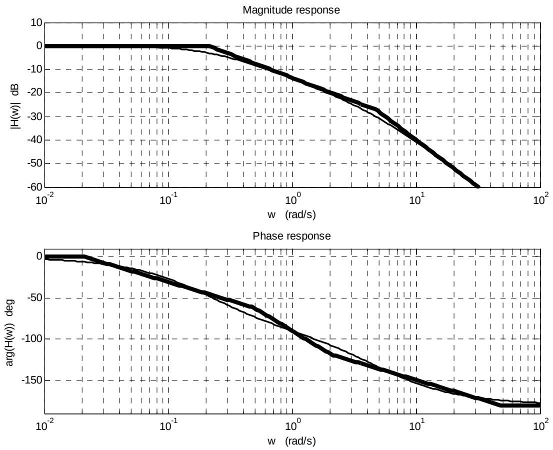


b) $G_2(s) = \dfrac{-10s}{(s+0.2)(s-10)}$



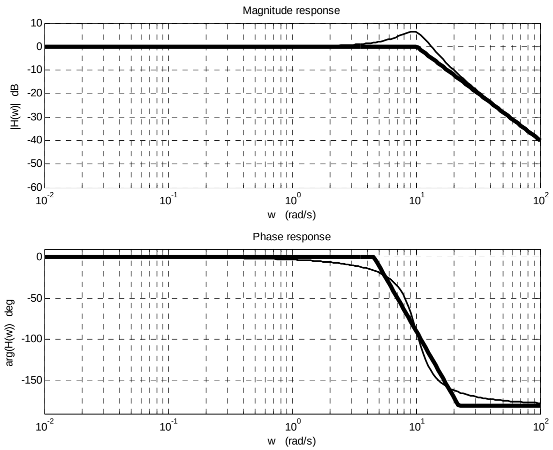
4B.9

a)



Bandwidth is 0 to 0.2 rads^{-1}.

b)

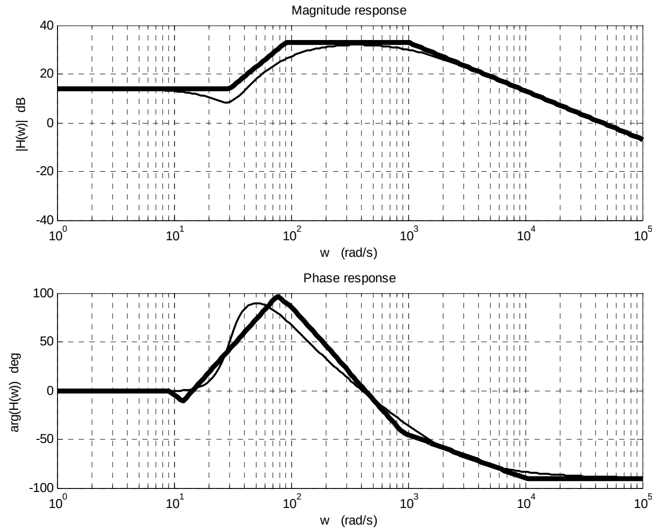


Bandwidth is increased to 15 rads^{-1} but is now peakier.

4B.10

(i) –3.3 dB, 8° (ii) –17.4 dB, 121°

4B.11



$$H(s)=\frac{45000\left(s^2+18s+900\right)}{(s+90)^2(s+1000)}$$

5A.1

a) 12.3% b) $t_p = 1.05\,\text{s}$ c) $\zeta = 0.555$ d) $\omega_n = \sqrt{13}\ \text{rads}^{-1}$

e) $\omega_d = 3\ \text{rads}^{-1}$ f) $t_r = 0.72\,\text{s}$ g) $t_s = 1.5\,\text{s}$ h) $t_s = 1.95\,\text{s}$

5A.2

$$G(s) = \frac{196}{s^2 + 19.6s + 196}$$

5A.3

$$c(t) = 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\sin\left(\sqrt{1-\zeta^2}\,\omega_n t + \cos^{-1}\zeta\right)$$

5A.4

a) $\zeta = \dfrac{\ln(a/b)}{\sqrt{4\pi^2 + [\ln(a/b)]^2}}$ b) $\zeta = 0.247$, $\omega_n = 16.21\ \text{rads}^{-1}$

5A.5

(a)

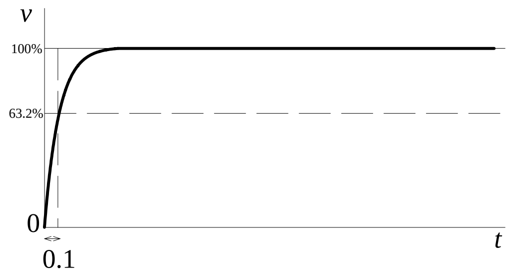
(i)

(ii)

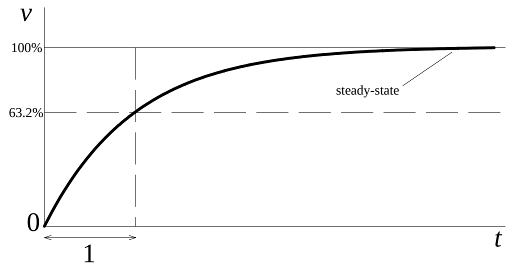
(iii)

(b)

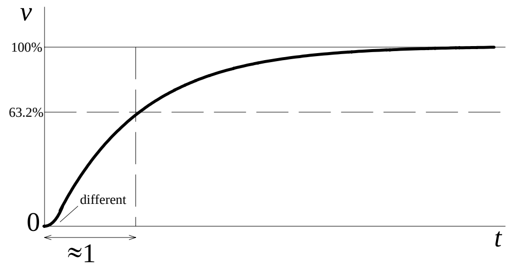
(i)



(ii)



(iii)



5A.7

a) $\omega_n = 3$, $\zeta = 1/6$ b) 58.8% c) 0 d) $1/9$

5A.8

a) 0 b) 3

5A.9

a) $T = 10$ sec b) (i) $T = 9.09$ sec (ii) $T = 5$ sec (iii) $T = 0.91$ sec

Feedback improves the time response.

5B.3

a) $S_{K_1}^T = 1$ b) $S_{K_2}^T = -\dfrac{1000}{s^2 + s + 1000}$ c) $S_G^T = \dfrac{s^2 + s}{s^2 + s + 1000}$

5B.4

a)

open-loop $S_{K_a}^T = 1$

open-loop $S_{K_1}^T = 0$

closed-loop $S_{K_a}^T = \dfrac{1}{1 + K_1 K_a G(s)} \approx \dfrac{1}{1 + K_1 K_a}$ for $\omega << \omega_n$

closed-loop $S_{K_1}^T \approx \dfrac{-K_1 K_a}{1 + K_1 K_a} \approx -1$ for $K_1 K_a >> 1$

b)

open-loop $\dfrac{\Theta(s)}{T_d(s)} = G(s) \approx 1$ for $\omega << \omega_n$

closed-loop $\dfrac{\Theta(s)}{T_d(s)} = \dfrac{G(s)}{1 + K_1 K_a G(s)} \approx \dfrac{1}{1 + K_1 K_a}$ for $\omega << \omega_n$

5B.5

a) (i) $\dfrac{1}{1 + K_P K_1}$ (ii) ∞ (iii) $\dfrac{-K_1}{1 + K_P K_1}$ (iv) $\dfrac{1 - K_1}{1 + K_P K_1}$

b) (i) 0 (ii) $\dfrac{1}{K_I K_1}$ (iii) 0 (iv) 0

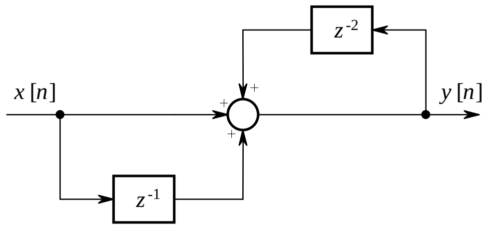
c) (i) 0 (ii) $\dfrac{1}{K_P K_1}$ (iii) $\dfrac{-1}{K_P}$ (iv) $\dfrac{-1}{K_P}$

d) (i) 0 (ii) 0 (iii) 0 (iv) 0

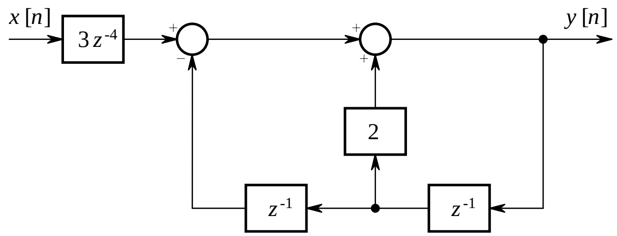
The integral term in the compensator reduces the "order" of the error, i.e. infinite values turn into finite values, and finite values become zero.

6A.1

(i)



(ii)



6A.2

(iii) $y[n] = y[n-1] - 2y[n-2] + 3x[n-1]$

(iv) $y[0] = 3,\ y[1] = 8,\ y[2] = -1,\ y[3] = -14$

6A.3

(a) $h[n] = (1/3)^n u[n]$

(b) $y[n] = -3\delta[n] + 3/2\,u[n] + 7/2\,(1/3)^n u[n]$

or $y[n] = 3/2\,u[n] + 1/2\,(1/3)^n u[n] + (1/3)^{n-1} u[n-1]$

or $y[n] = \{3 - (1/3)^n\} u[n] - 1/2\,\{3 - (1/3)^{n-2}\} u[n-2]$

(the responses above are equivalent – to see this, graph them or rearrange terms)

6A.4

$$h[n]=h_1[n]+h_2[n]$$

6A.5

$$h[0]=h_1[0]$$

$$h[1]=h_1[1]-h_2[1]h_1^2[0]$$

$$h[2]=h_1[2]+h_2^2[1]h_1^3[0]-2h_1[0]h_1[1]h_2[1]-h_1^2[0]h_2[2]$$

6A.6

(a)

(i) $F(z)=\dfrac{z}{z-1/2}$ (ii) $F(z)=\dfrac{1}{z-a}$

(iii) $F(z)=\dfrac{z}{(z-a)^2}$ (iv) $F(z)=\dfrac{z(z+a)}{(z-a)^3}$

(b) $X(z)=\dfrac{2z^2}{z^2-1}$

6A.7

(a) zeros at $z=0,-2$; poles at $z=-1/3,-1$; stable

(b) fourth-order zero at $z=0$; poles at $z=\pm j1/\sqrt{2},\pm j1/2$; stable

(c) zeros at $z=\pm j\sqrt{2/5}$; poles at $z=-3\pm\sqrt{6}$; unstable

6A.8

$$H(z)=\frac{3+z^{-1}-2z^{-4}}{1-z^{-1}+z^{-2}}$$

6A.9

$$x[n]=0,14,10.5,9.125,\ldots$$

6A.10

$$x[n]=8(1/4)^n u[n]+8(1/2)^n u[n]+8u[n]-24\delta[n]$$

6A.11

(a) $x[n]=\dfrac{1-a^{n+1}}{1-a}$ for $n=0,1,2,\ldots$

(b) $x[n]=3\delta[n]+2\delta[n-1]+6\delta[n-4]$ for $n=0,1,2,\ldots$

(c) $x[n]=1-e^{-anT}$ for $n=0,1,2,\ldots$

(d) $x[n]=8-8(1/2)^n-6n(1/2)^n$ for $n=0,1,2,\ldots$

(e) $x[n]=2(1/2)^{n-4}$ for $n=0,1,2,\ldots$

(f) $x[n]=2/\sqrt{3}\sin(2\pi n/3)$ for $n=0,1,2,\ldots$

6A.12

(a) $h[n]=\dfrac{2^{n+1}-1}{2^{2n+3}}=1/4(1/2)^n-1/8(1/4)^n$

(b) $H(z)=\dfrac{1/8\,z^2}{(z-1/2)(z-1/4)}$

(c) $y[n]=1/3+1/24(1/4)^n-1/4(1/2)^n$

(d) $y[n]=1/2(1/2)^n-1/18(1/4)^n+n/3-4/9$

6A.13

$$y[n]=-1+\frac{1}{\sqrt{5}}\left(0.5+\sqrt{1.25}\right)^{n+3}-\frac{1}{\sqrt{5}}\left(0.5-\sqrt{1.25}\right)^{n+3}\text{ for }n=0,1,2,\ldots$$

6A.14

$$y[n]=\frac{1}{\sqrt{5}}\left[\left(0.5+\sqrt{1.25}\right)^n-\left(0.5-\sqrt{1.25}\right)^n\right]$$

6A.15

$$F(z)=\frac{zTe^{-\alpha T}}{\left(z-e^{-\alpha T}\right)^2}$$

6A.16

(a) $X_3(z)=\left[\frac{z+1+k}{(k+1)z+6k+1}\right]X_1(z)+\left[\frac{k(z+1)}{(k+1)z+6k+1}\right]D(z)$

6A.17

(a) $f(0)=0$, $f(\infty)=0$ (b) $f(0)=1$, $f(\infty)=0$

(c) $f(0)=1/4$, $f(\infty)=\frac{1}{1-a}$ (d) $f(0)=0$, $f(\infty)=8$

6A.18

(i) $(n+1)u[n]$

(ii) $\delta[n]-2\delta[n-1]+3\delta[n-2]-2\delta[n-3]+\delta[n-4]$

6A.19

$$f[n]=8-6n(1/2)^n-8(1/2)^n,\quad n\geq 0$$

6B.1

$$y[n]=0.7741y[n-1]+20x[n]-18.87x[n-1]$$

6B.2

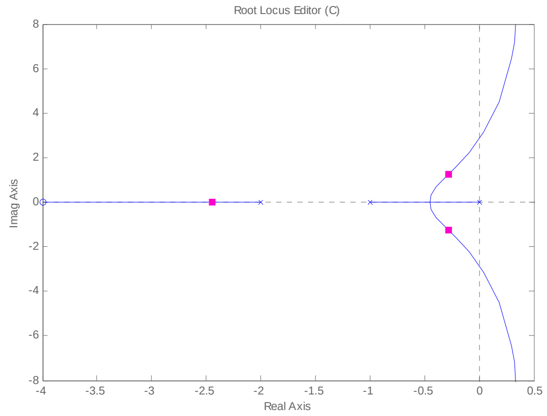
$$y[n]=0.7730y[n-1]+17.73x[n]-17.73x[n-1]$$

6B.3

$$H_d(z)\approx\frac{0.1670z+0.3795}{(z-0.3679)(z-0.1353)}$$

7B.1

(a)



(b) $0<K<6$, $\omega=2.82\text{ rads}^{-1}$

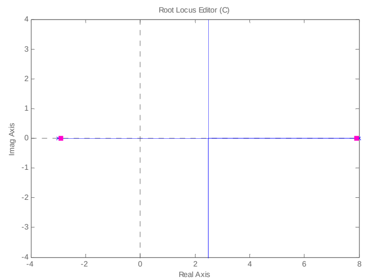
(c) $K\approx 0.206$

(d) $t_s\approx 10.7\text{ s}$

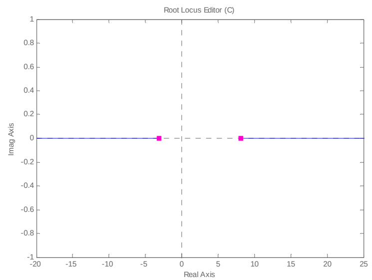
7B.2

(a) One pole is always in the right-half plane:

For $K\geq 0$:

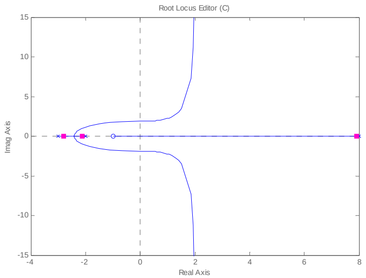


For $K<0$:

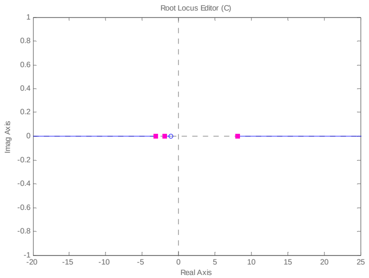


(b)

For $K \geq 0$:



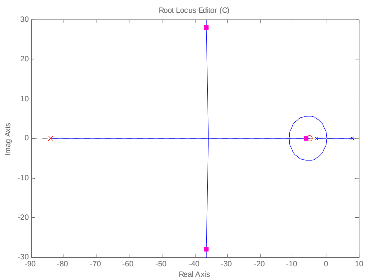
For $K < 0$:

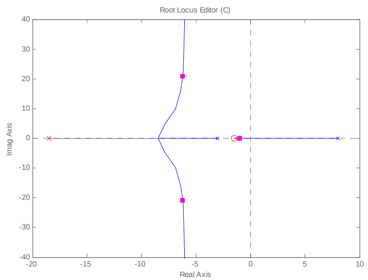


The system is always unstable because the pole pair moves into the right-half plane ($s = \pm j\sqrt{3.5}$) at a lower gain ($K = 37.5$) than that for which the right-half plane pole enters the left-half plane ($K = 48$). The <u>principle</u> is sound, however. A different choice of pole-zero locations for the feedback compensator is required in order to produce a stable feedback system.

For $K < 0$, the root in the right-half plane stays on the real-axis and moves to the right. Thus, negative values of K are not worth pursuing.

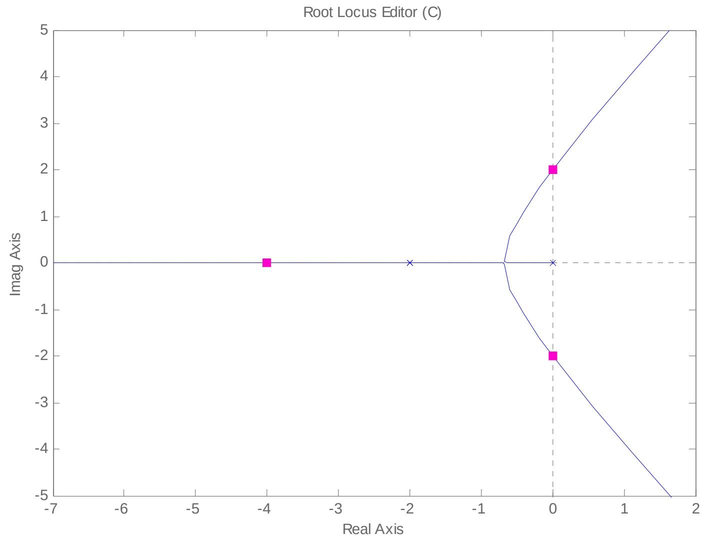
For the given open-loop pole-zero pattern, there are two different feasible locus paths, one of which includes a nearly circular segment from the left-half s-plane to the right-half s-plane. The relative magnitude of the poles and zeros determines which of the two paths occurs.





7B.3

(a)

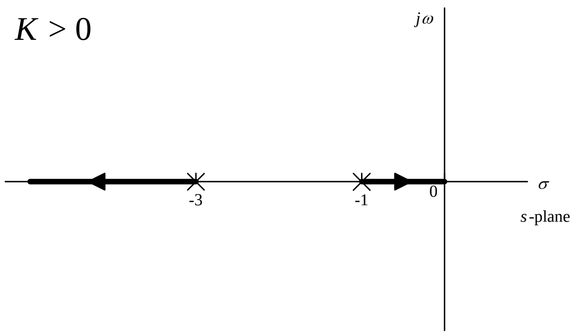


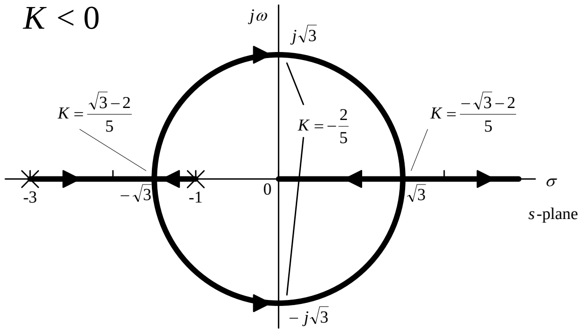
(b) $a = 32/27$

(c) $0 < a < 16$ for stability. For $a = 16$, the frequency of oscillation is $2\ \text{rads}^{-1}$.

7B.4

(a)





(b) $K=-2/5$, $\omega_n=\sqrt{3}\ \text{rads}^{-1}$

(c) Any value of $K>-2/5$ gives zero steady-state error. [The system is type 1. Therefore any value of K will give zero steady-state error provided the system is stable.]

(d) $K\approx -0.1952$, $t_r\approx 1.06\,\text{s}$

8A.1

$$\mathbf{q}=\begin{bmatrix} v_{C_1} \\ i_{L_1}\end{bmatrix},\ \mathbf{A}=\begin{bmatrix} \dfrac{-1}{C_1(R_2+R_3)} & \dfrac{R_2}{C_1(R_2+R_3)} \\ \dfrac{-R_2}{L_1(R_2+R_3)} & \dfrac{-1}{L_1}\left(\dfrac{R_2R_3}{R_2+R_3}+R_1\right)\end{bmatrix},\ \mathbf{b}=\begin{bmatrix} 0 \\ \dfrac{1}{L_1}\end{bmatrix}$$

8A.2

(a) false (b) false (c) false [all true for $\mathbf{x}(0)=\mathbf{0}$]

8A.3

(a) $$\dot{\mathbf{q}}=\begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & 2 & -2\end{bmatrix}\mathbf{q}+\begin{bmatrix} 3 \\ 1 \\ 0\end{bmatrix}x \qquad y=\begin{bmatrix} 2 & 2 & 1\end{bmatrix}\mathbf{q}+\begin{bmatrix} 0\end{bmatrix}x$$

(b) $$\dot{\mathbf{q}}=\begin{bmatrix} -2 & -1 & -1 \\ 1 & -1 & 0 \\ 0 & 1 & -2\end{bmatrix}\mathbf{q}+\begin{bmatrix} 1 \\ 0 \\ 0\end{bmatrix}x \qquad y=\begin{bmatrix} 0 & 0 & 1\end{bmatrix}\mathbf{q}+\begin{bmatrix} 0\end{bmatrix}x$$

8A.4

(a) $$H(s)=8\frac{\left(s+\frac{1}{2}\right)(s+3)}{s(s+2)^2}$$

(b) $$H(s)=\frac{1}{s^3+5s^2+9s+7}$$

8A.5

$$\begin{bmatrix} q_1(t) \\ q_2(t)\end{bmatrix}=L^{-1}\begin{bmatrix} \dfrac{3(s-1)}{s^2+3}+\dfrac{s-3}{s(s^2+3)} \\ \dfrac{12}{s^2+3}+\dfrac{2s+6}{s(s^2+3)}\end{bmatrix}=\begin{bmatrix} -1+4\cos\left(\sqrt{3}t\right)-\dfrac{2}{\sqrt{3}}\sin\left(\sqrt{3}t\right) \\ 2-2\cos\left(\sqrt{3}t\right)+\dfrac{14}{\sqrt{3}}\sin\left(\sqrt{3}t\right)\end{bmatrix}u(t)$$

A.42

8B.1

$\lambda_1 = -1$, $\lambda_2 = -2$, $\lambda_3 = -3$

8B.2

(a) 1, 2 (b) $\mathbf{q}(t) = \begin{bmatrix} -1+e^t \\ -e^t \end{bmatrix}$

8B.3

(ii) $-1,\ -1 \pm j2$ (iii) $k_1 = -75$, $k_2 = -49$, $k_3 = -10$ (iv) $y_{ss}(t) = -1/40$

(v) State feedback can place the poles of any system arbitrarily (if the system is controllable).

8B.4

$$\mathbf{q}[n+1] = \begin{bmatrix} 0 & 0 & 1/16 \\ 1 & 0 & 1/4 \\ 0 & 1 & -1/4 \end{bmatrix} \mathbf{q}[n] + \begin{bmatrix} 3/16 \\ 1/8 \\ 1/2 \end{bmatrix} x[n]$$
$$y[n] = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \mathbf{q}[n] + 3x[n]$$

8B.6

$$\mathbf{q}[n+1] = \begin{bmatrix} -a & 1 & 0 \\ -k & 0 & k \\ -(b-c) & 0 & -c \end{bmatrix} \mathbf{q}[n] + \begin{bmatrix} 0 \\ k \\ b-c \end{bmatrix} x[n]$$
$$y[n] = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \mathbf{q}[n]$$