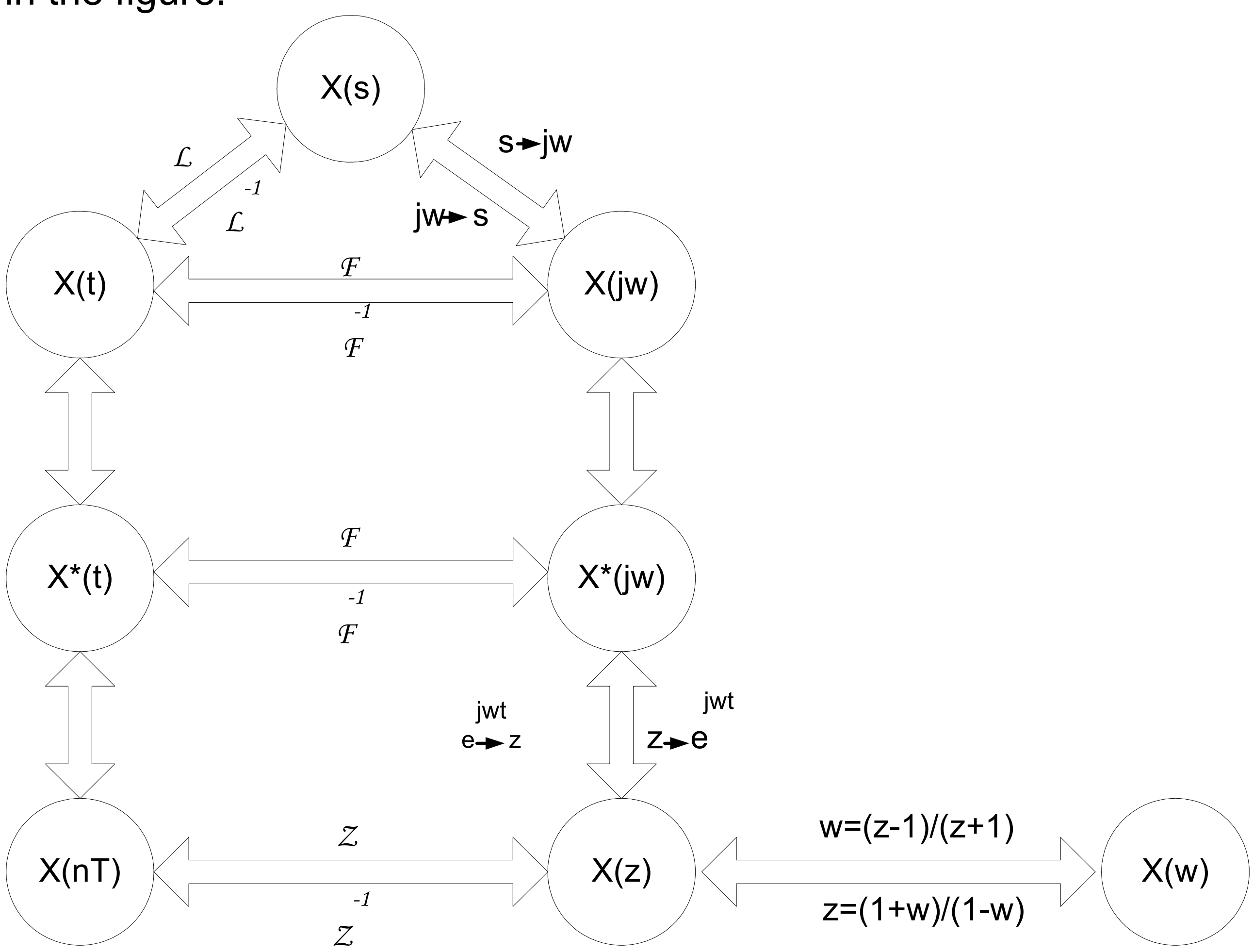
1

Transform methods

Some of the different forms of a signal, obtained by transformations, are shown in the figure.



We will very briefly review the use of Fourier and Laplace transforms, and then have a quick look at the DFT (Discrete Fourier Transform), the FFT (Fast Fourier Transform, a fast computational method for evaluating the DFT) and the z transform.

1.1 The s-plane and the Laplace domain

Introduction

You are familiar with the fundamental circuit laws and how they are applied to both direct current and alternating current circuits. You have also used the Laplace transform to help you with the solution of differential equations that you encounter in solving circuit problems. We will commence this course with a brief review of the s-plane, with the intention of looking at some of the characteristic properties of selected circuits from a slightly different angle. A formal presentation of the Laplace transform will not be attempted, rather, only its applications in the study of circuits (and later) and control systems.

We start by recognising the s-plane, a complex plane on which we can place the poles and zeros of a network function. (A network function is a function, which represents some characteristic, such as the input impedance, of the network. We will deal with each of these ideas later on.) The s-plane is a complex plane, whose coordinates are the real part σ and the so-called “imaginary” part $j\omega$

We can represent any function of the complex variable s (= $\sigma + j\omega$) by its poles and zeros, if we disregard any multiplying factors. The poles are those values of s for which the function goes to infinity, and zeros are those values of s for which the function vanishes.

The general complex exponential excitation function

You have studied the behaviour of circuits under different excitations, by dc, ac and special functions such as impulse and step functions (In practice, we have only approximations to these special functions, as sudden changes are not possible in real life.) We will now look at a more general excitation function that can be used to study the behaviour of circuits in a more convenient and uniform manner.

Consider the general complex exponential excitation function, $x(t) = X\, e^{s_0 t}\, u_{-1}(t)$.

$u_{-1}(t)$ represents a unit step function at the origin, and is introduced to ensure that the excitation is applied at time zero. We are only interested in causal systems, that is, where the response occurs only after the excitation. The complex variable s_0 (= $\sigma_0 + j\omega_0$) is general enough to represent most types of input that we are interested in, and in particular, sinusoidal inputs.

With σ_0 zero, we have a pure, steady, sinusoid; while with ω_0 zero, it is either an exponentially growing or an exponentially decaying quantity, depending on whether σ_0 is positive or negative. With complex s_0, we would have either growing or decaying sinusoids.

Network functions

We said earlier that a network function is a function, which represents some characteristic, such as the input impedance, of a network. Let us now look at this in a little more detail.

The simplest network that we can think of is a single-element, two-terminal network, such as a resistor. The characteristics of a resistor are described by the Ohm's Law equation:

$$v = iR$$

Where v is the voltage difference *across* the resistor, i is the current *through* it, and R is its resistance, which defines the relationship between v and i. With inductors (inductance L) and capacitors (capacitance C), we have relationships that are dependant on rates of change, giving rise to differential equations:

$v = L\, di / dt$ for an inductor
$i = C\, dv / dt$ for a capacitor

Taking Laplace transforms (assuming zero initial conditions), we get:

$V(s) = [Ls]\, I(s)$ for an inductor
$V(s) = [1/Cs]\, I(s)$ for a capacitor

We can now write the relationship between voltage and current for resistors, inductors or capacitors (or any combination of them) as:

$V(s) = Z(s)\, I(s)$
Or $Z(s) = V(s) / I(s)$

Z(s), a function of the complex variable s, is known as the impedance of the two-terminal network. We also have Y(s), the inverse of Z(s) defined by the relationship

$$Y(s) = I(s) / V(s)$$

Y(s) is known as the admittance of the network.

We can visualise a two-terminal network as a single-port network. If it is to be excited by a source, the source has to be connected across the two terminals, thus making it appear as a single port to the source. Most networks, if they are to accomplish any useful purpose, have to have more than one port. A network may be connected between (say) a source and a load to achieve some desired connectivity characteristic, a kind of transformation. We then have, at the simplest, two-port networks with just one pair of ports, which we naturally identify as the input port and the output port. A complex network is made up of interconnections of a large number of such networks, not necessarily two-port networks. However, multi-port networks can, in most cases, be analysed using two-port network theory, and we confine our attention to them in this course.

We saw earlier how impedance and admittance functions can describe a single-port network. In a two-port network, we would naturally have to extend these concepts to both ports, so that we will have:

Input impedance
Input admittance
Output impedance
Output admittance
Transfer impedance and admittance

as functions that characterise the network. There are other functions too which are of significance, such as:

Transfer function
Characteristic impedance
Scattering matrix, which defines the relationship between the incident power and the reflected power.

Some of these ideas will be discussed later.

We arrived at the concept of network function through the Laplace transform. *The system function of a network (a network function) is a function of the complex frequency s, representing the ratio of the Laplace transform of a response to the Laplace transform of the excitation causing the response*. We assume that the network is initially relaxed, that is the stored energy in the network is initially zero.

Pole-zero patterns

The most general form of a system function is given by a *real rational function of s*. It can be expressed as the quotient of two polynomials in s. Such a function is completely described by its poles, zeros and a multiplier factor.

Let us examine such a function, H(s).

$H(s) = q(s) / p(s)$, where

$q(s) = b_m(s-z_1)(s-z_2) \; . \; . \; . \; (s-z_m)$
$p(s) = (s-p_1)(s-p_2) \; . \; . \; . \; . \; (s-p_n)$

The finite zeros of H(s) are $z_1, z_2, \ldots, z_m$ while its finite poles are at $p_1, p_2, \ldots, p_n$. In addition to these finite poles and zeros, you can see that there would be additional poles or zeros at infinity, depending whether m>n or n>m, of order (m-n) or (n-m).

A study of the pole-zero pattern of a system (or network) function gives us an insight into its behaviour. For example, an examination of the driving point impedance function of a one-port network will define its impedance at all natural frequencies, and enable us to obtain a physical realisation of the impedance using actual components. Similarly, it is possible for us to obtain both the frequency characteristics and a physical realisation of a two-port filter network, given its pole zero pattern.

As a network function is completely defined by its poles and zeros, they are called the critical frequencies of the system function. The (finite) poles of the system function correspond to the natural frequencies of the network, that is the frequencies of natural oscillations. These are also known as the natural modes of the network.

The order of a system function is the highest order (in s) of the denominator or the numerator, and gives an indication of its complexity.

Properties of LC, RC & RLC network functions

Before considering the properties of passive network functions such as those of LC, RC and RLC networks, we need to be familiar with what are known as positive real functions.

LC, RC and RLC network functions all have certain common properties.

All these system functions, whether immitance (impedance or admittance) functions or transfer functions, are quotients of two polynomials in s, with real rational coefficients. They are thus *real rational functions*.

They are all passive functions, with no intrinsic energy sources. Hence, if viewed as response functions to an impulse excitation, they cannot diverge without limit. Therefore, they *cannot have poles (or zeros) on the right-half plane, nor can they have multiple-order poles on the j axes.*

They obey the reciprocity theorem. Thus, any two impedances obtained by interchanging the points of excitation and response are equal. This would mean that *the resulting matrices are symmetric*.

Let us now look at each of these types of networks.

The following properties of an LC network may be derived:

They are simple, that is there are no higher order poles or zeros.
They all lie on the j axis.
Poles and zeros alternate.

The origin and infinity are always critical frequencies, that is, there will be a pole or zero at both the origin and at infinity.
The multiplicative constant is positive.

RC (and RL) network functions have characteristics that are different from those of LC networks. Since an RC network has resistive components, it cannot have a zero-valued impedance at any real frequency. Its zeros (and poles) are on the non-positive σ axis of the s-plane. It should also be noted that the form of the impedance function is different from that of the admittance function, unlike in the case of the LC network. However, the impedance function of an RC network is similar in form to the admittance function of an RL network, and vice versa.

The poles and zeros of an RC driving point function lie on the non-positive real axis.
They are simple.
Poles and zeros alternate.
The slopes of impedance functions are negative, those of admittance functions are positive.

Energy functions

Energy functions or "energy-like-functions" may be used for the derivation of properties of driving point impedance and admittance functions of passive networks. When $s = j\omega$, these functions are directly related to the energy stored in the network, as currents through inductances or voltages across capacitances, that is, as electromagnetic or electrostatic energy.

These functions are positive semi-definite functions.

We define three functions T, F and V, corresponding to the kinetic, dissipation and potential energies as follows:

Consider the loop equations, in matrix form:

$$[Z]\,[I] = [E]$$

where $$[Z] = s[L] + [R] + [S] / s,$$

(S is the loop elastance, the reciprocal of loop capacitance)

Pre-multiplying both sides of the equation by the transpose of the complex conjugate of [I] will yield an energy-like expression with three terms (denoted by T, F and V) related to the inductance, resistance and elastance of the loop circuits.

We may define another set of parallel functions V*, F* and T* starting with the node-pair equations:

[Y] [E]   = [I]

Again, pre-multiplying both sides by the transpose of the complex conjugate of [E],we obtain the expressions for V*, F* and T*.

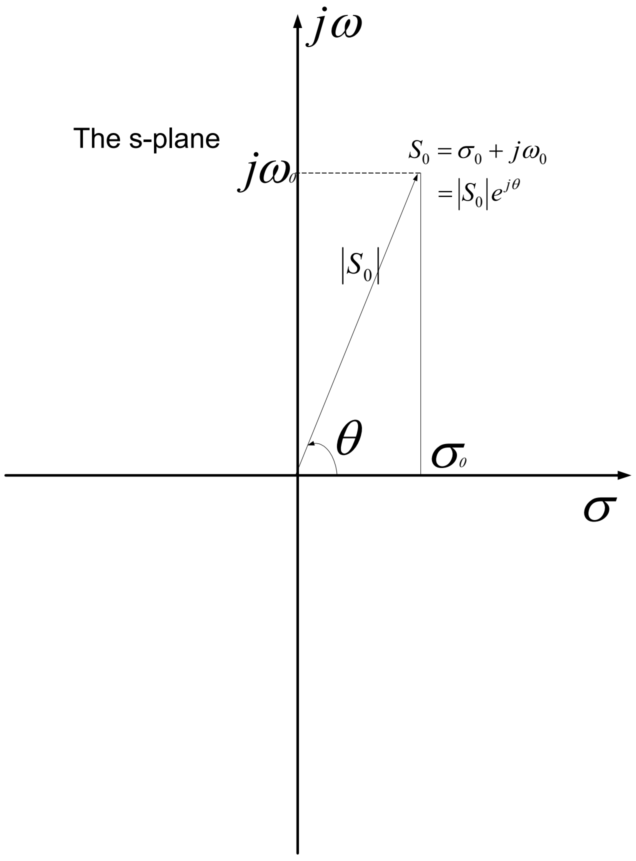
The two derivations give us two forms of the energy function as:

sT + F + V/s
sV* + F* + T*/s

These may be used to evaluate the driving point impedances and admittances by imposing suitable conditions.

1.1.1 The s-plane

The s-plane is a complex plane (axes σ and $j\omega$) as shown in the figure.



Any point s_0 on the plane will have two coordinates σ_0 and $j\omega_0$ as shown. They are known as the “real” part and the “imaginary” part of the complex number s_0 (= $\sigma_0 + j\omega_0$). However, there is nothing more real in the real part than in the imaginary, both are real enough. This terminology has arisen due to historical reasons, and should be treated as mere names, with no significance in the meaning.

s_0 can also be represented in polar form, as having a magnitude $|s_0|$, and angle θ. This of course further illustrates that all quantities are real in the normal meaning of the word.

The relationships among these various quantities are obvious from the diagram:

$$s_0 = \sigma_0 + j\omega_0 = |s_0| e^{j\theta};$$

$$|s_0| = \sqrt{\left(\sigma_0^{\ 2} + \omega_0^{\ 2}\right)}$$

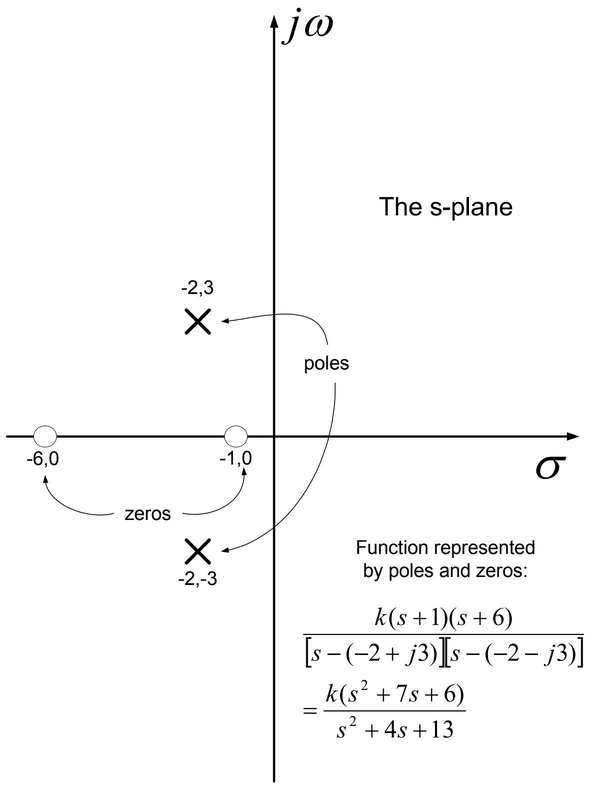
$$\theta = \tan^{-1} \frac{\omega_0}{\sigma_0}$$

$$\sigma_0 = |s_0| \cos\theta$$

$$\omega_0 = |s_0| \sin\theta$$

$$s_0 = |s_0|\ (\cos\theta + j\sin\theta)$$

We can represent a function of a complex variable s (= $\sigma + j\omega$). by its poles and zeros. By poles we mean those values of s for which the function becomes infinite, that is, its denominator becomes zero; and by zeros we mean those values of s for which the function vanishes, that is, the numerator becomes zero.



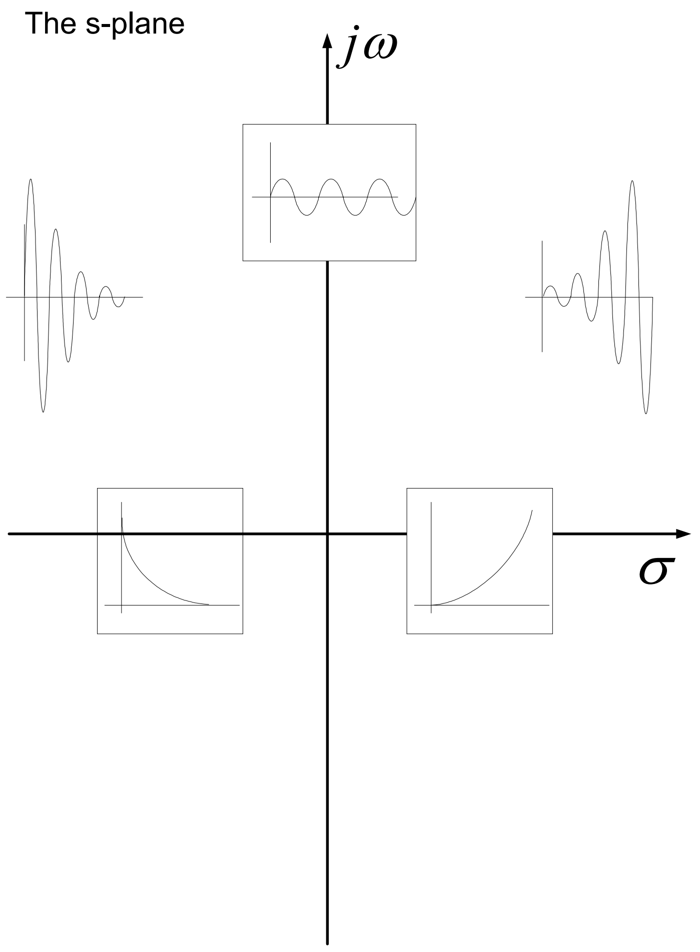
The poles are marked by crosses (✕) and zeros by noughts (O) as shown. Note that any multiplying factor (k, in the illustration) is not represented on the s-plane plot

1.1.2 The general complex exponential excitation function

The behaviour of the general exponential excitation function

$x(t) = X e^{s_0 t} u_{-1}(t).$

for s_0 lying in different parts of the s-plane is shown in the figure.

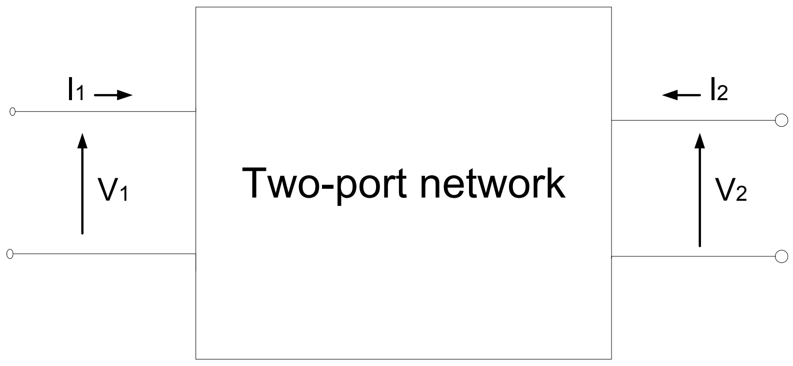


With σ_0 zero, we have a pure, steady, sinusoid; while with ω_0 zero, it is either an exponentially growing or an exponentially decaying function, depending on whether σ_0 is positive or negative. With complex s_0, we would have either growing or decaying sinusoids.

Fourier analysis enables us to express any repetitive waveform as the summation of a series of sinusoids, with arbitrary accuracy. Thus, any required repetitive excitation pattern might be implemented using the general complex exponential excitation. This also applies to non-repetitive waveforms, as they can be treated as repetitive waveforms with an infinite period. The use of the complex excitation function further enable us to decay any component of the excitation, using the real part of the complex exponential.

With s_0 equal to zero and X equal to one, we can obtain the unit step function.

1.1.3 Two-port networks



Consider a two-port network as shown.

The driving point impedance (or the input impedance) looking in at port 1 would obviously depend on the type of termination at port 2. If we recognise port 1 as the input and port 2 as the output, it is usual to compute the driving point impedance at port 1 with port 2 open circuited, that is, with I_2 equal to zero.

We can describe the behaviour of the two-port network by the following set of two equations:

$$V_1 = z_{11}I_1 + z_{12}I_2$$
$$V_2 = z_{21}I_1 + z_{22}I_2$$

The z's are called the open circuit impedances. As noted earlier, when I_2 is zero, we have, from the first equation:

$$z_{11} = \frac{V_1}{I_1},\ z_{21} = \frac{V_2}{I_1}$$

z_{11} is the driving point impedance at port 1, or the input impedance of the network, looking at port 1. z_{21} is a transfer impedance.

Similarly, with I_1 zero, we obtain:

$$z_{22} = \frac{V_2}{I_2},\ z_{12} = \frac{V_1}{I_2}$$

We could have described the network in terms of admittances, instead of impedances, as given below:

$$I_1 = y_{11}V_1 + y_{12}V_2$$
$$I_2 = y_{21}V_1 + y_{22}V_2$$

We saw earlier why the impedances z_{11}, z_{12}, z_{21} and z_{22} are known as open circuit impedances, for we obtained them by setting I_1 and I_2 equal to zero. Similarly, the admittance functions turn out to be short circuit admittances.

With V_2 equal to zero, we have:

$$y_{11} = \frac{I_1}{V_1},\ y_{21} = \frac{I_2}{V_1}$$

and setting V_1 to zero gives:

$$y_{22} = \frac{I_2}{V_2},\ y_{12} = \frac{I_1}{V_2}$$

Other representations of two-port networks are also in use. One such is the hybrid parameter representation used in electronic circuit analysis.

$$V_1 = h_{11}I_1 + h_{12}V_2$$
$$I_2 = h_{21}I_1 + h_{22}V_2$$

ABCD representation is commonly used in the study of transmission lines. Here, the current I_2 is, by convention, shown as leaving the network rather than as entering the network.

$$V_1 = AV_2 + BI_2$$
$$I_1 = CV_2 + DI_2$$

1.1.4 Positive real functions

A complex function G(s) is said to be positive real when:

G(s) is real for all real s
Re (G(s)) ≥ 0, for Re(s) ≥ 0

The driving point function of any physical network is positive real, and every rational function that is positive real can be realised as the driving point function of a network. Hence, the study of such functions is important in the study of networks.

If G is positive real, then 1/G is also positive real.

It can be shown that G(s) is positive real if it satisfies the following conditions:

1. G(s) is real for all real s
2. G(s) is analytic in the right-half plane.
3. Any poles on the j axis are simple, with real positive residues.
4. Re $(G(j\omega)) \geq 0$ for all ω

All conditions relating to poles also apply to zeros, as 1/G also has to be positive real.

The following properties of positive real functions may be deduced from these conditions:

No negative coefficients occur in either the numerator or denominator.

The highest (and lowest) powers of the numerator and denominator cannot differ by more than unity.

Poles (and zeros) on the j axes occur in conjugate pairs.

1.1.5 Positive semi-definite functions

We define a positive semi-definite function as one which is always either positive or zero. We will consider only quadratic forms in the present situation.

All the terms of a quadratic form are of the second order. In matrix form, we can represent a quadratic form as:

$$X^T A X$$

where X is a column vector and A is a symmetric matrix.

(Note: We will confine our attention to symmetric matrices, and both X and A are real)

The following definitions and properties apply:

1. A real matrix A is said to be positive semi-definite if and only if

$$X^T A X \geq 0$$

for all real and finite X

2. A is said to be positive definite if and only if it is positive semi-definite, and in addition

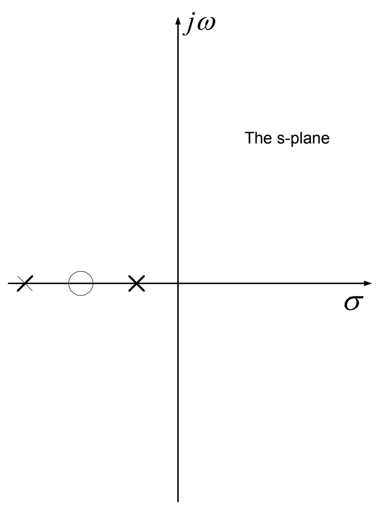
$$X^T A X = 0$$

only if X=0

3. For A to be positive definite, each of its principal minors should be positive. (This may be tested by testing a set of determinants obtained by deleting successive rows and columns from A for positiveness.)

4. For A to be positive semi-definite, each of its principal minors should be non-negative.

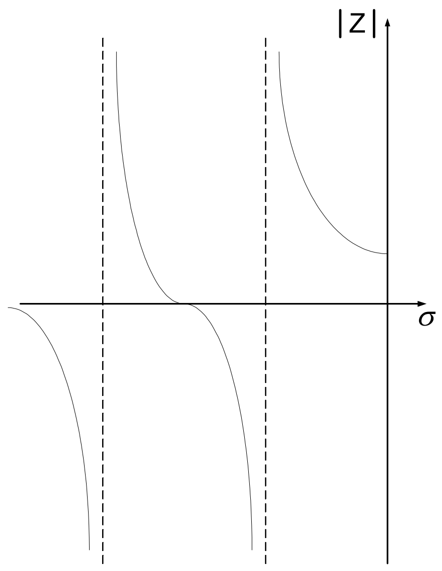
1.1.6 Properties of RC networks



The figure shows a typical pole-zero pattern of the impedance function of an RC network. Note that:

> The poles and zeros are simple.
> They all lie on the non-positive σ- axis.
> Poles and zeros alternate.

The plot of impedance vs' σ, along the non-positive real axis is shown below.

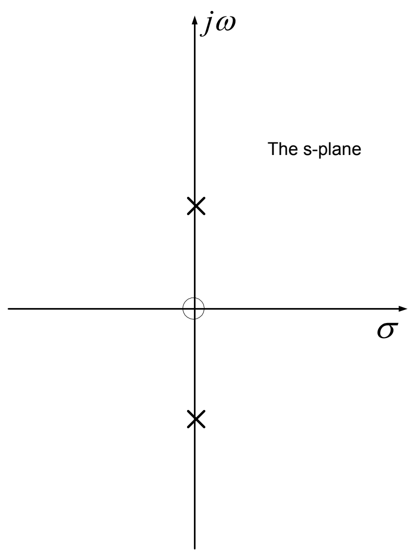


Note that there is a critical frequency at infinity, but not at the origin, in this particular case

In general, the critical frequency of the smallest magnitude is a pole, lying on the non-positive (that is, at the origin or on the negative) real axis.

When the number of finite poles is greater than the number of finite zeros, there will be a zero at infinity, as is the case in the example (two finite poles and only one finite zero). When the degrees of the numerator and denominator polynomials are equal, there will be no critical frequency at infinity. Driving point impedance functions of RC networks have negative gradients along the σ axis.

1.1.7 Properties of LC networks.

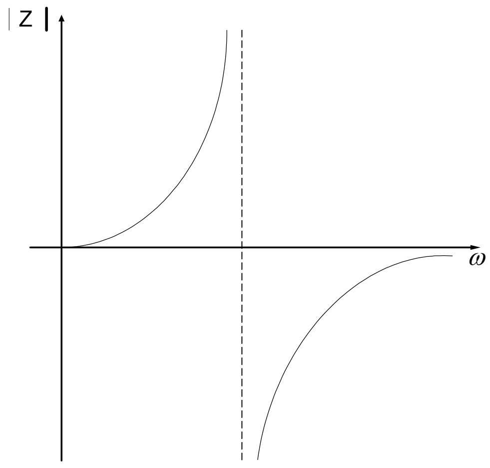


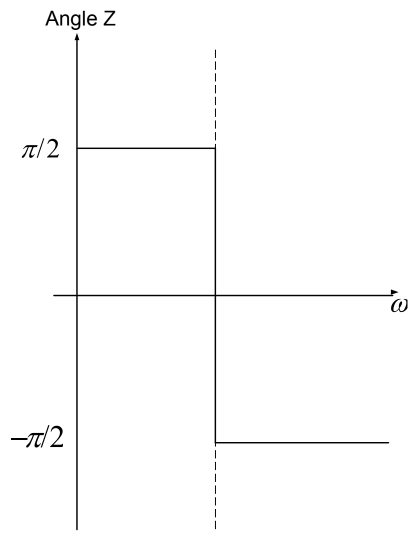
The figure shows a typical pole-zero pattern of the impedance function of an LC network

Note that

The poles and zeros are simple.
They all lie on the $j\omega$ axis.
Poles and zeros alternate.

The variation of impedance with frequency corresponding to this pole-zero pattern may be obtained by substituting jω for s, and is as shown below.

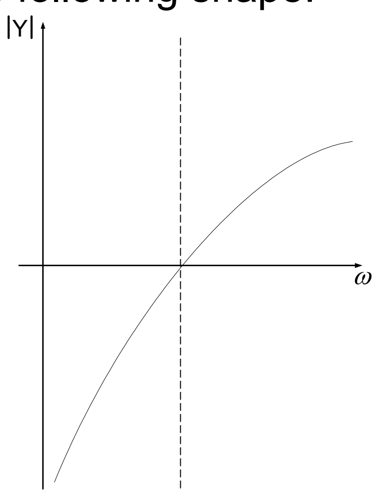




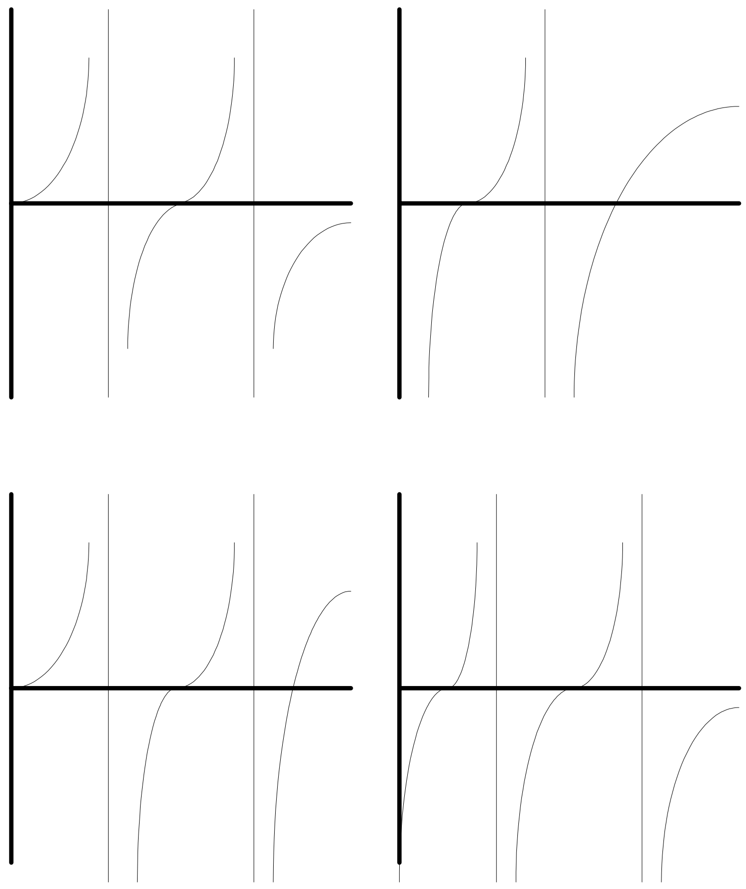
Note that there are critical frequencies at the origin and at infinity.

The characteristics of driving point admittance functions of LC networks are similar to those of impedance functions, and may be obtained by substituting 1/s for s.

For the impedance function considered earlier, the admittance function would have he following shape:



Plots of other typical immitance functions of LC networks are shown below.



In addition to the properties mentioned earlier, you may have noticed another property of LC functions.

They all have positive gradients, that is, the value of the function always increases with increase of frequency.

This arises from the fact that all the residues at the poles and zeros are positive, that is, the partial fraction expansion of the functions all have positive coefficients. Substitution of $s = j\omega$ and differentiation leads to this result.

1.2 The Fourier Transform Family

Introduction

We have noted that a signal can be transformed among various forms, and that it is advantageous to do so under certain circumstances. The information contained in the signal will manifest in different ways, in the different forms. Similarly, the characteristics of a system block in a signal flow path can also be represented in different ways. When dealing with time domain signals, it is convenient to represent a block by its impulse response, while in the frequency or Laplace domain; it will be represented by a transfer function.

The effect of a function block on a signal in the time domain is represented by the impulse response of the system block, and its effect is computed by a process known as convolution.

In this section, we will study about convolution and correlation and also about transformations in the discrete time domain, the DFT and its inverse IDFT. We will also see how it can be computed more easily using the FFT.

The Fourier transform

The frequency response of a system block is obtained by Fourier transformation of its impulse response, and the output (in the frequency domain) is obtained by multiplication of the Fourier transform of the input (frequency spectrum) and the frequency response of the device. Multiplication in the frequency domain is equivalent to convolution in the time domain, while multiplication in the time domain is equivalent to convolution in the frequency domain.

The Fourier transform is a process that transforms between a continuous time domain signal and the frequency domain (The reverse process is the inverse Fourier transform). It cannot handle discrete signals.

The Discrete Fourier Transform

We will begin the study of discrete signals by looking at the sampling process in somewhat more detail

The sampling of a continuous signal can be looked upon as multiplying it by a sequence of unit impulses. The result is their convolution. As noted earlier, convolution and multiplication are reciprocal processes in relation to Fourier transformation.

We have started this study of sampled signals on the basis of understanding the output obtained from a sampling-measuring device, such as a digital voltmeter. We now find that the technique of Fourier transformation (so useful in the study of continuous processes (and signals)) needs to be modified if we were to apply it to discrete systems.

The DFT (Discrete Fourier Transform) is the parallel process of transformation applicable to discrete signals. By discrete, we mean here signals that have been turned into a sequence of numerical values, suitable for processing on a digital computer. Hence, the DFT itself is a numerical algorithm that performs this function. The FFT is a modified algorithm that takes advantage of certain symmetries in the DFT to make it work much faster, and is now universally used in the computation of the DFT.

The z transform

The z transform is the only other transformation in our original diagram showing various forms in which a signal may be represented. Similar to the relationship between the Laplace transform and the Fourier transform, where we can obtain the Laplace transform by substituting s for $e^{j\omega}$, we can obtain the z transform from the discrete Fourier transform by substituting z for $e^{j\omega T}$,

1.2.1 The Fourier Transform

We are familiar with the Fourier series, where a periodic function of time may be expressed as a summation of sinusoidal functions, defined by:

$$\begin{aligned} f(t) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{2\pi nt}{T} + \sum_{n=1}^{\infty} b_n \sin\frac{2\pi nt}{T} \\ &= \sum_{n=-\infty}^{\infty} c_n e^{j\omega_n t} \end{aligned}$$

$$where\,T = period\ of\ the\ signal$$

$$and$$

$$a_0 = \frac{2}{T}\int_{t_0}^{t_0+T} f(t)\,dt$$

$$\omega_n = \frac{2\pi n}{T}$$

$$a_n = \frac{2}{T}\int_{t_0}^{t_0+T} f(t)\cos\left(\omega_n t\right)dt$$

$$b_n = \frac{2}{T}\int_{t_0}^{t_0+T} f(t)\sin\left(\omega_n t\right)dt$$

$$c_n = \frac{1}{2}(a_n - jb_n)$$

The Fourier Transform generalises this concept to non-periodic signals by extending the period T to infinity. This would mean that the fundamental frequency ω_n tends to zero. In effect, we end up with a continuous frequency spectrum

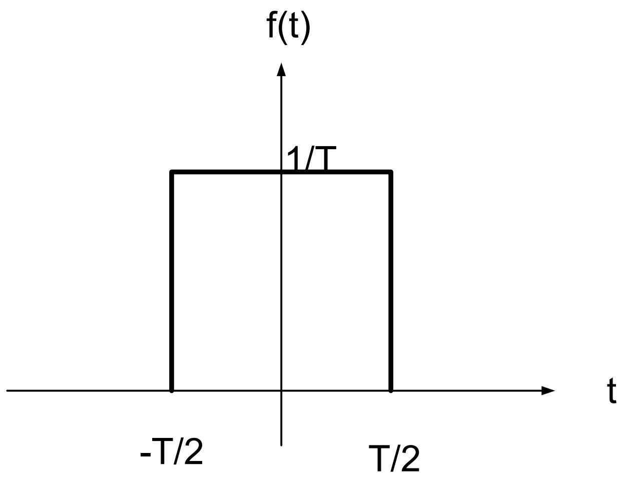
The Fourier Transform is defined as:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)\,e^{-j\omega t}\,dt\,.$$

and its inverse is defined as:

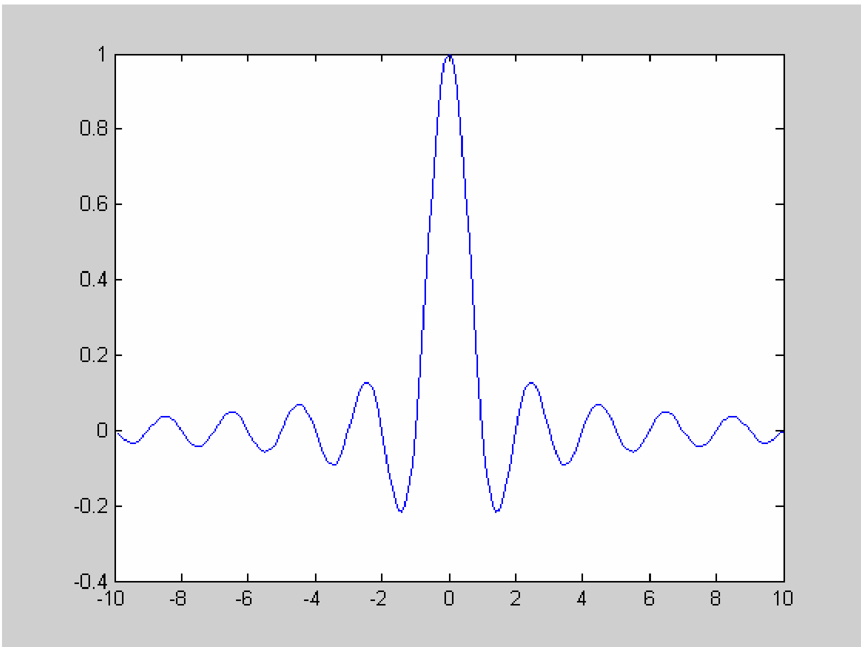
$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\,e^{j\omega t}\,d\omega$$

Let us evaluate the Fourier Transform of a square pulse of duration T and magnitude 1/T, centred on the origin.



$$
\begin{aligned}
F(\omega) &= \int_{-\infty}^{\infty} f(t)\, e^{-j\omega t}\, dt \\
&= \int_{-T/2}^{T/2} f(t)\, e^{-j\omega t}\, dt \\
&= \left| \frac{1}{T} \frac{1}{j\omega} e^{-j\omega t} \right|_{-T/2}^{T/2} \\
&= \frac{\sin \frac{\omega T}{2}}{\frac{\omega T}{2}} = sinc\,(x)
\end{aligned}
$$

Note that the dummy variable x has been introduced at the last stage, as the other variables no longer matter. The sinc function has the form shown below:



It is instructive to look at the Fourier Transformation as consisting of four different situations, namely,

Continuous time, repetitive
Continuous time, non-repetitive
Discrete time, repetitive
Discrete time, non-repetitive

As we have already seen, in the first case (continuous time, repetitive), we have the Fourier Series, where (time is continuous and) the frequency is discrete,

In the second case, as the signal is non-repetitive, it extends to infinity, and we have the Fourier Transform, where both time and frequency are continuous, extending to infinity.

In the third case, we have the Discrete Fourier Transform (DFT), where both time and frequency are discrete.

In the fourth case, the transformation again yields a continuous frequency spectrum extending to infinity. This is called the Discrete Time Fourier Transform (DTFT).

The four cases are shown in the following table:

Time	Repetitive	Non-repetitive	Frequency
Continuous	Fourier Series	Fourier transform	Discrete
Discrete	Discrete Fourier Transform (DFT)	Discrete Time Fourier Transform (DTFT)	Continuous

1.2.2 The Discrete Fourier Transform (DFT)

Before we consider the DFT, we will need to briefly look at operations on sequences. We will consider a sequence of N terms as follows:

0 $\rightarrow$ f(0)
1 $\rightarrow$ f(1)
2 $\rightarrow$ f(2)
3 $\rightarrow$ f(3)
.
k $\rightarrow$ f(k)
.
(N-1) $\rightarrow$ f(N-1)

We can define operations between two different sequences as:

Addition (You can only add sequences of the same order)

$$\{f(k)\}+\{g(k)\}=\{f(k)+g(k)\}$$

Multiplication $\{f(k)\}\{g(k)\}=\{f(k)g(k)\}$

Example: {0,1,2} . {4,2,1} = {0,2,2}

Division is similar, except that g(k) should be non-zero.

We will now go on to study the DFT. We have the Fourier Transform pair

$$F(\omega) = \int_{-\infty}^{\infty} f(t)\, e^{-j\omega t}\, dt$$

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\, e^{j\omega t}\, d\omega$$

where f(t) is defined in the continuous time domain. Assume that the signal is sampled to yield a discrete time signal f(k). We can then write directly:

DFT Definition:

$$F(p) = \frac{1}{N}\sum_{k=0}^{N-1} f(k)\, e^{-2\pi\, j\, k\, p\, /\, N}, \;\; p \in [0, N-1]$$

Inverse DFT

$$f(k) = \sum_{p=0}^{N-1} F(p)\, e^{2\pi\, j\, k\, p\, /\, N}, \qquad k \in [0, N-1]$$

For convenience, we write $\omega_N = e^{2\pi\, j\, /\, N}$

Then

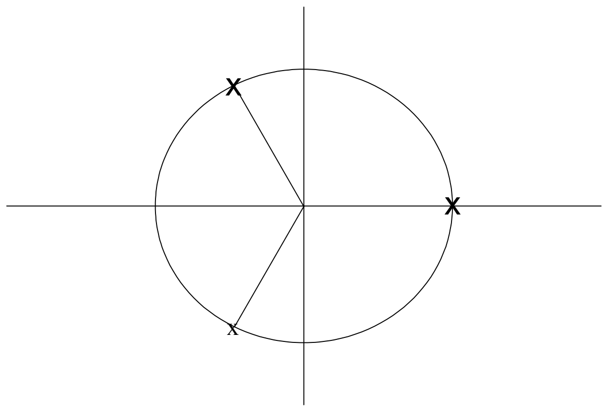
$$DFT: \qquad F(p) = \frac{1}{N}\sum_{k=0}^{N-1} f(k)\, {\omega_N}^{-k\, p}$$

$$IDFT: \qquad f(k) = \sum_{p=0}^{N-1} F(p)\, {\omega_N}^{kp}$$

Example

$$f(k) = \{0,1,2\}$$

$$\omega_3 = e^{2\pi\, j\, /\, 3}$$



$$F(p) = \frac{1}{3}\sum f(k)\,\omega_3^{\;-kp}$$

$$F(0) = \frac{1}{3}\left[f(0) + f(1) + f(2)\right] = \frac{1}{3}(0+1+2) = 1$$

$$F(1) = \frac{1}{3}\left[f(0) + f(1)\omega_3^{\;-1} + f(2)\omega_3^{\;-2}\right] = \frac{1}{3}\left[0 + 1(e^{-2\pi j/3}) + 2(e^{-4\pi j/3}\right]$$

$$F(2) = \frac{1}{3}\left[0 + 1(e^{-4\pi j/3}) + 2(e^{-8\pi j/3})\right]$$

$$F(p) = \left\{F(0), F(1), F(2)\right\}$$

We can use the MATLAB functions fft and ifft to verify our results as follows:

```
>> f=[0,1,2]

f =

    0     1     2

>> F=fft(f)

F =

   3.0000             -1.5000 + 0.8660i  -1.5000 - 0.8660i

>> f=ifft(F)

f =

    0     1     2
```

Note that we get back the original sequence when we obtain the inverse DFT of the transformed sequence. The discrepancy in the value obtained for the transformed sequence is due to slightly different conventions used by different algorithms for the evaluation of the transform. In our formulation, the factor 1/N appeared in the forward transformation while MATLAB has it in the inverse transformation. Still others use a factor of 1/√N in both directions.

Periodicity

If we extend the transformation beyond N, we get

$$\omega_N{}^{k+N} = e^{2\pi j(k+N)/N} = e^{2\pi jk/N} e^{2\pi j\,N/N} = e^{2\pi jk/N}.1$$

In general, if α is an integer, $\omega_N{}^{(k+\alpha N)} = \omega_N{}^{k}$

Hence, the transformations are periodic, with period N. We can also show that:

$$F(-p) = F(N-p)$$
$$f(-k) = f(N-k)$$

There are other properties of importance, but we will not go into their details here. In most aspects, they are parallel to the properties of the Laplace and Fourier Transforms. We will look at some of these in relation to the Z Transform later on.

1.2.3 The Fast Fourier Transform (FFT)

The FFT is an algorithm for the efficient computation of the DFT. Let us take another look at the definition of the DFT.

DFT Definition:

$$F(p) = \frac{1}{N}\sum_{k=0}^{N-1} f(k)\, e^{-2\pi jkp/N},\; p \in [0, N-1]$$

Inverse DFT

$$f(k) = \sum_{p=0}^{N-1} F(p)\, e^{2\pi jkp/N}, \qquad k \in [0, N-1]$$

With $\omega_N = e^{2\pi j/N}$

$$DFT: \qquad F(p) = \frac{1}{N}\sum_{k=0}^{N-1} f(k)\,\omega_N{}^{-kp}$$

$$IDFT: \qquad f(k) = \sum_{p=0}^{N-1} F(p)\,\omega_N{}^{kp}$$

[Sometimes the scaling factor 1/N is used with the DFT as shown above, and sometimes with the inverse transform. Some workers use a symmetrical scaling factor of 1/√N with both the DFT and the IDFT]

For any reasonably long sequence, this looks like a formidable task, even with the assistance of modern-day computers. If we attempt to implement the DFT by direct computation using the above definition, we will have to perform N complex multiplications for each term, giving a total of N^2 multiplications for the complete DFT.

The FFT attempts to simplify the computation of the DFT.

We have (assuming N is even)

$$\begin{aligned}
F(p) &= \frac{1}{N}\sum_{k=0}^{N-1} f(k)\omega_N^{-kp} \\
&= \frac{1}{N}\left[f(0)\omega_N^{0} + f(1)\omega_N^{-p} + f(2)\omega_N^{-2p} + \ldots + f(N-2)\omega_N^{-(N-2)p} + f(N-1)\omega_N^{-(N-1)p}\right] \\
&= \frac{1}{N}\left[f(0)\omega_N^{0} + f(2)\omega_N^{-2p} + \ldots. + f(N-2)\omega_N^{-(N-2)p}\right] + \frac{1}{N}\left[f(1)\omega_N^{-p} + f(3)\omega_N^{-3p} + \ldots. + f(N-1)\omega_N^{-(N-1)p}\right]
\end{aligned}$$

We will consider the sequence f(k) (of length N) as being composed of two sequences $f_1(k)$ and $f_2(k)$, each of length N/2, such that $f_1(k)$ contains all the even terms of f(k) while $f_2(k)$ contains all the odd terms. We also note that

$$\omega_N = e^{2\pi j/N}$$

$$\therefore \omega_{N/2} = e^{2\pi j/\frac{N}{2}} = e^{4\pi j/N} = \omega_N^{2}$$

We will now re-write the expression for F(p) in terms of $f_1(k)$ and $f_2(k)$:

$$\begin{aligned}
F(p) &= \frac{1}{N}\left[f_1(0)\omega_N^{0} + f_1(1)\omega_N^{-2p} + \ldots. + f_1(N/2-1)\,\omega_N^{-2(N/2-1)p}\right] \\
&\quad + \frac{1}{N}\left[f_2(0)\omega_N^{0} + f_2(1)\omega_N^{-2p} + \ldots. + f_2(N/2-1)\omega_N^{-2(N/2-1)p}\right]\omega_N^{-p} \\
&= \frac{1}{N}\left[f_1(0)\omega_{N/2}^{0} + f_1(1)\omega_{N/2}^{-p} + \ldots. + f_1(N/2-1)\omega^{-(N/2-1)p}\right] \\
&\quad + \frac{1}{N}\left[f_2(0)\omega_{N/2}^{0} + f_2(1)\omega_{N/2}^{-p} + \ldots. + f_2(N/2-1)\omega^{-(N/2-1)p}\right]\omega_N^{-p}
\end{aligned}$$

$$\therefore F(p) = \frac{1}{2}\left[F_1(p) + \omega_N{}^{-p}\, F_2(p)\right]$$

Each of the transforms F_1 and F_2 can now be further broken down into two shorter sequences to give (F_{11} and F_{12}) and (F_{21} and F_{22}), each of length N/4. We can proceed in this manner until we reach a stage where the sequence is of length one, when the transform can be written down by inspection, and is the same as the time sequence. We then proceed in the reverse direction to construct the transforms of length two, four, eight etc. until we arrive at the transform of the original sequence.

We have assumed that the sequence is of length 2^n, where n is an integer, to be able to continually divide the sequence into two halves. When this is not so, it is quite simple to fill-up the sequence by the addition of trailing zeros to make it up to a power of 2.

Example:

We will consider the same example that was used to illustrate the DFT, even though the power of the FFT becomes useful only with high order sequences.

$$f(k) = \{0,1,2\}$$

We will pad this up to make the sequence length a power of 2:

$$f(k) = \{0,1,2,0\}$$
$$f_1(k) = \{0,2\}, \quad f_2(k) = \{1,0\}$$
$$f_{11}(k) = \{0\}, f_{12}(k) = \{2\}; \qquad f_{21}(k) = \{1\}, f_{22}(k) = \{0\}$$
$$F_{11}(p) = \{0\}, F_{12}(p) = \{2\}; \qquad F_{21}(p) = \{1\}, F_{22}(p) = \{0\}$$

$$\omega_2 = e^{2\pi j/2} = e^{\pi j} = -1$$
$$\omega_2{}^0 = 1, \omega_2{}^{-1} = -1$$
$$F_1(p) = \frac{1}{2}\{(0+2),(0-2)\} = \{1,-1\}$$
$$F_2(p) = \frac{1}{2}\{(1+0),(1-0)\} = \left\{\frac{1}{2},\frac{1}{2}\right\}$$

$$\omega_4 = e^{2\pi j/4} = e^{\pi j/2} = j$$

$$\omega_4{}^0 = 1, \omega_4{}^{-1} = \frac{1}{j} = -j$$

$$F(p) = \frac{1}{2}\left\{(1+\frac{1}{2}),(-1-\frac{j}{2}),(1-\frac{1}{2}),(-1,\frac{j}{2}\right\} = \frac{1}{4}\{3,(-2-j),1,(-2+j)\}$$

Compare this with the results obtained using the MATLAB programme below:

```
>> f=[0 1 2 0]

f =

    0    1    2    0

>> F=fft(f)

F =

  Columns 1 through 3

   3.0000            -2.0000 - 1.0000i   1.0000

  Column 4

  -2.0000 + 1.0000i

>> f=ifft(F)

f =

    0    1    2    0
```

As before, they differ by a multiplying factor (4 in this case) due to the differences in the definitions used. I have obtained the inverse transform using MATLAB to confirm the correctness of the computation.

It differs from the DFT of {0, 1, 2} obtained earlier because of the differences in the discrete frequencies computed. This can be further illustrated by computing the FFT of the sequence obtained by extending the series to a length of 8 with additional zeros:

```
>> f=[0 1 2 0 0 0 0 0]
```

```
f =

    0     1     2     0     0     0     0     0

>> F=fft(f)

F =

  Columns 1 through 3

   3.0000             0.7071 - 2.7071i  -2.0000 - 1.0000i

  Columns 4 through 6

  -0.7071 + 1.2929i   1.0000            -0.7071 - 1.2929i

  Columns 7 through 8

  -2.0000 + 1.0000i   0.7071 + 2.7071i
```

The values for alternate frequencies coincide with those of the previous computation, as expected.

With the radix-two computation (that is, subdividing the sequence into two equal parts every time) the total amount of computation is drastically reduced. We have $N \log_2 N$ multiplications in place of the original N^2. For example, if N is 2^{10} (=1024), the number of multiplications comes down from 2^{20} to 10×2^{10}, a factor of more than a 100. For longer sequences, the saving is much more.

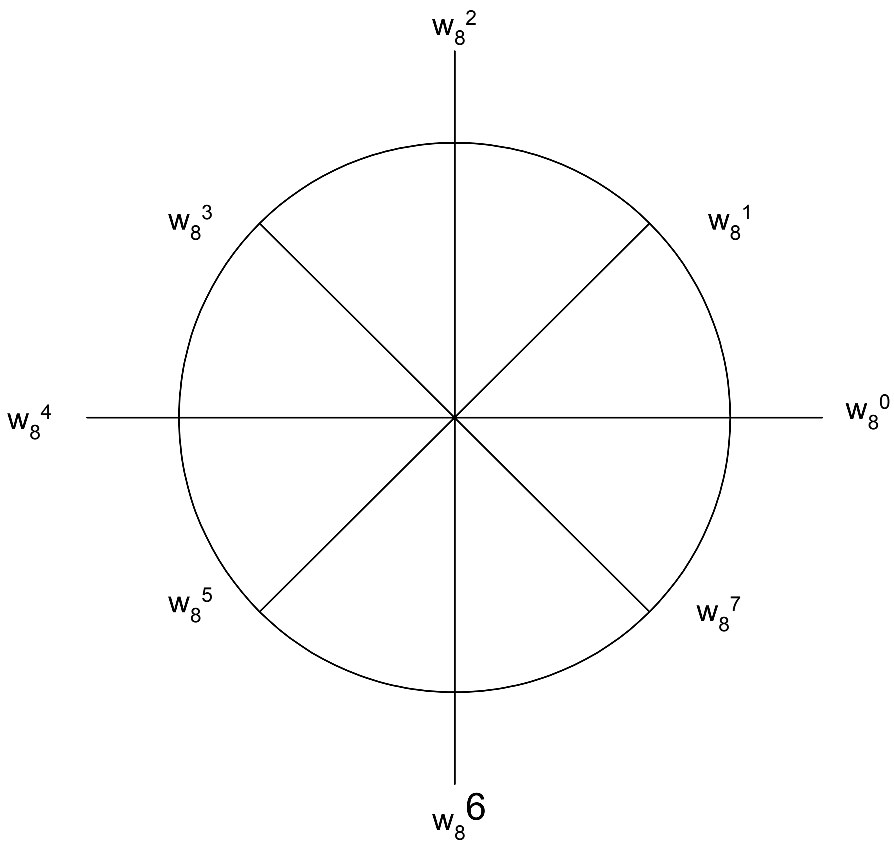
A diagrammatic representation of the FFT

The above algorithm may be better understood when it is presented in diagrammatic form, known as the Butterfly diagram. (What we have considered in the previous section is known as the decimation in time algorithm, and there is another parallel one known as decimation in frequency.) We will now consider a diagrammatic representation for the decimation in time algorithm of a sequence of length eight.

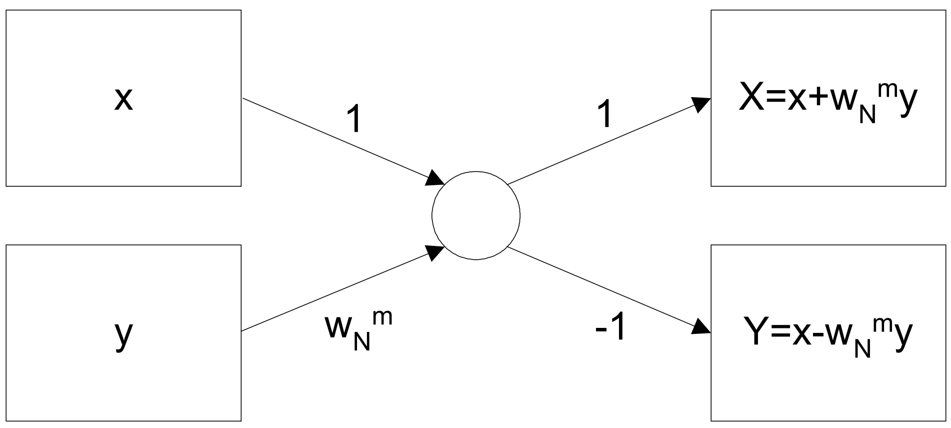
$$f(k) = \{f(0), f(1), f(2), f(3), f(4), f(5), f(6), f(7)\}$$
$$f_1(k) = \{f(0), f(2), f(4), f(6)\}, \qquad f_2(k) = \{f(1), f(3), f(5), f(7)\}$$
$$f_{11}(k) = \{f(0), f(4)\}, f_{12}(k) = \{f(2), f(6)\}, \quad f_{21}(k) = \{f(1), f(5)\}, f_{22}(k) = \{f(3), f(7)\}$$
$$f_{111} = \{f(0)\}, f_{112} = \{f(4)\}, f_{121} = \{f(2)\}, f_{122} = \{f(6)\}$$
$$f_{211} = \{f(1)\}, f_{212} = \{f(5)\}, f_{221} = \{f(3)\}, f_{222} = \{f(7)\}$$

The original time sequence {f(0), f(1), f(2), f(3), f(4), f(5), f(6),f(7)} of length eight has got “decimated in time” to eight sequences {f(0)}, {f(4)}, {f(2)}, {f(6)}, {f(1)}, {f(5)}, {f(3)}, and {f(7)} each of length one.

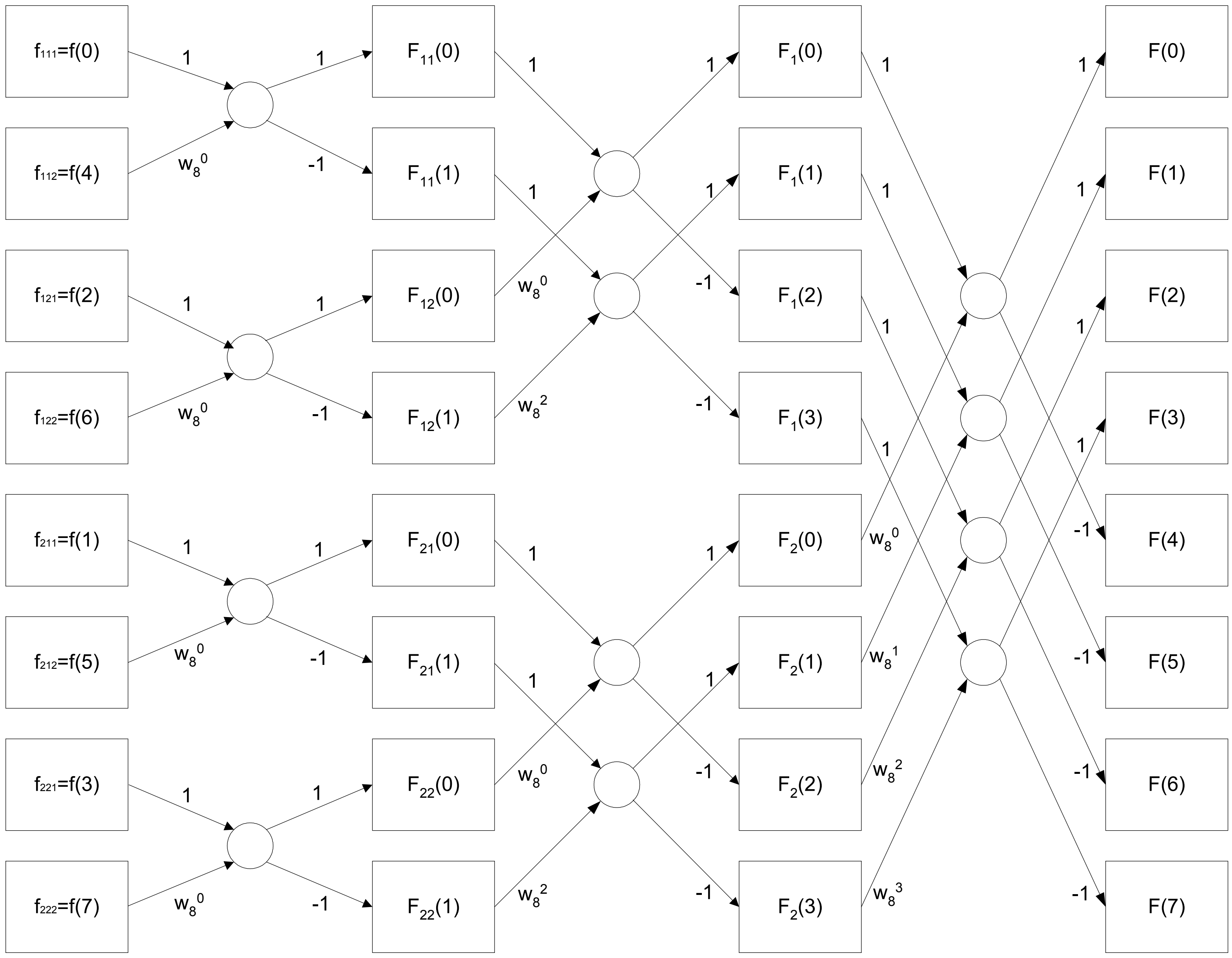
We will now examine the nature of the “ twiddle factor ω_8 ” which operates as a multiplying factor in generating the transformed sequences. $\omega_8 = e^{2\pi j/8}$ may be represented graphically on the complex plane as shown:



The weights, arrows and the central “butterfly” are interpreted as given in the following diagram:



We then form the successive transforms as shown below, where the definition used has the factor 1/N in the inverse transform and not the forward transform.



The reordering of the input sequence can be easily accomplished by using a trick known as “bit reversing”. For the sequence of length considered, the original order, in binary form is:

000, 001, 010, 011, 100, 101, 110, 111

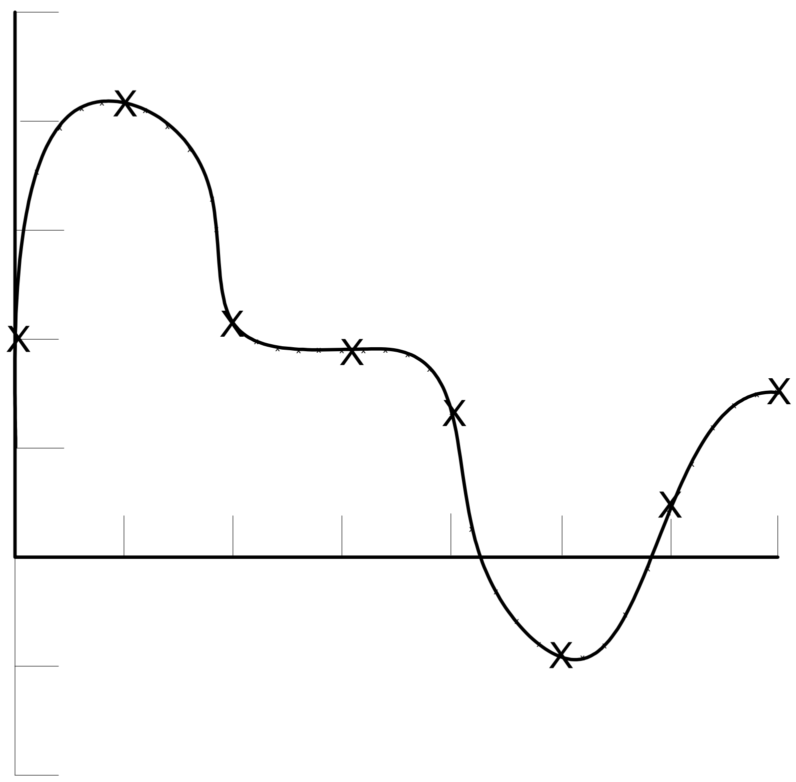
With the bits reversed, we end up with:

000, 100, 010, 110, 001, 101, 011, 111

This gives the required order 0, 4, 2, 6, 1, 5, 3, 7

1.3 The z operator, difference equations and the z transform

Let us look at a continuous signal (as a function of time t) that has been sampled at regular intervals T apart, that is, at time 0, T, 2T, 3T. . . etc.



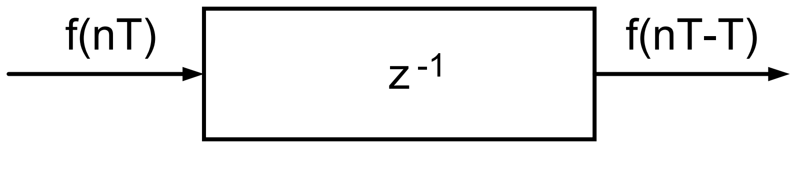
We can represent this by a sequence such as (say):

f(nT) = {20, 42, 21, 20, 15, -10, 5, 15}

What would be the sequence obtained if the original sequence were delayed by a single time interval T? Let us denote it by f(nT-T) or f((n-1)T):

f(nT-T)= (0, 20, 42, 21, 20, 15, -10, 5, 15}

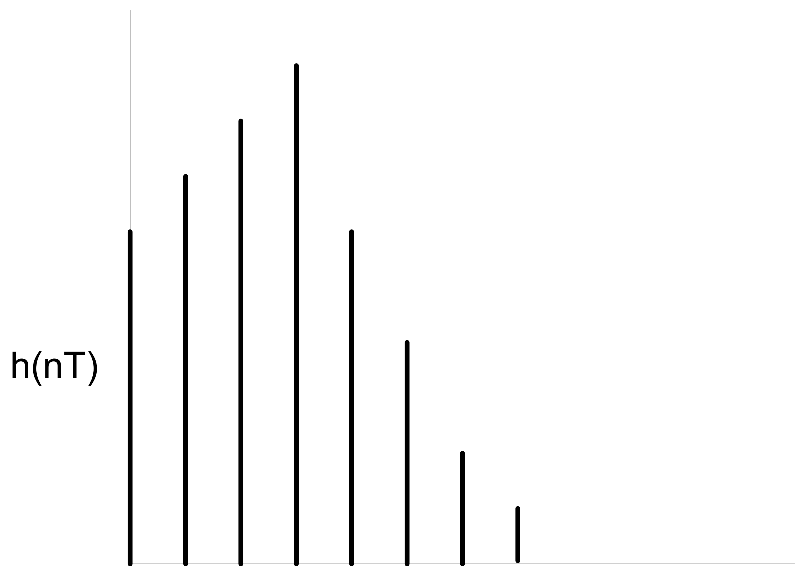
[The zero at the start is introduced on the assumption that we are only dealing with causal systems.]



We use the above representation for denoting a unit time delay. Here z^{-1} is simply an operator that delays a signal by one time step.

With this notation established, we will now look at difference equations describing the processing of sequences of discrete signals.

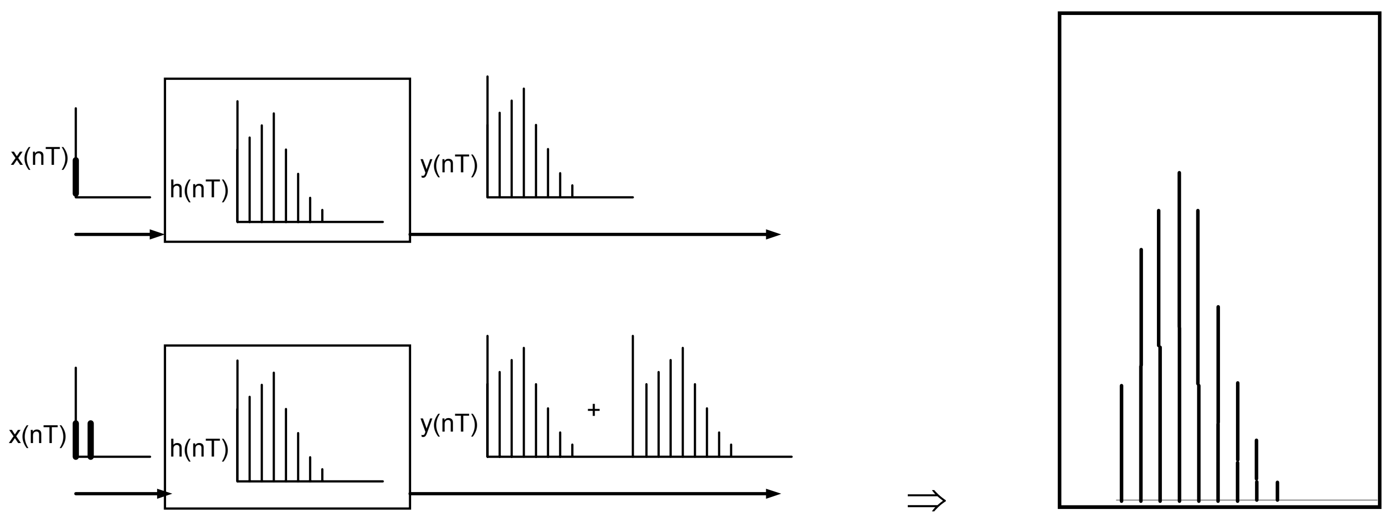
Let us consider a system with impulse response h(nT):



If a unit impulse is presented to this, we will get an output similar to the above sequence. Now, what would happen if we present two unit impulses, one after the other?

In response to the first impulse, we will get a response similar to what we got before. In response to the second impulse, we will get another sequence, similar to the first, but delayed by one period.

If the system is linear, the two responses will add up, and we get a sequence as shown:



Note that the sequence is now longer, as the two responses are displaced by one period. In general, if x is of length n and h is of length m, the response will be of length (n+m-1).

Here we considered x to be sequences of unit impulses. If they were of arbitrary magnitude, then, the corresponding responses would be obtained by multiplying the impulse response by these magnitudes. We can then write:

$$y(nT) = \sum_{k=-\infty}^{\infty} x(kT)\, h(nT - kT)$$

for linear, time-invariant systems. If in addition, the system is causal, we may restrict the summing interval to give:

$$y(nT) = \sum_{k=0}^{n} x(kT)\, h(nT - kT)$$

This is the convolution summation of x and h. By change of variable, it may also be written as:

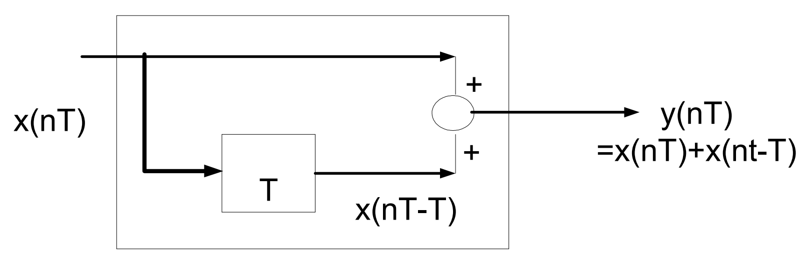
$$y(nT) = \sum_{k=0}^{n} x(nT - kT)\, h(kT)$$

Thus we will have:

y(0) = x(0) h(0)
y(T) = x(0) h(T)+x(T) h(0)
y(2T) = x(0) h(2T) + x(T) h(T) + x(2T) h(0)

etc.

The above convolution summation may also be seen as an equivalent multiplication in the transformed z plane. However before we look at that interpretation, we will look at a few extremely simple filters with our z-operator notation.



We can write:

$$y(nT) = x(nT) + x(nT - T)$$

If we now use the z-operator to denote the time delay, and if we represent x(nT) and y(nT) by their transforms, we can write:

$$Y(z) = X(z)(1 + z^{-1})$$
$$\frac{Y(z)}{X(z)} = H(z) = (1 + z^{-1}) = \frac{1+z}{z}$$

Here, H(z) corresponds to, and is the transform of the impulse function h(nT) which we met earlier.

The two relationships

$$y(nT) = \sum_{k=0}^{n} x(kT)\,h(nT - kT)$$

and

Y(z) = H(z) X(z)

are equivalent.

We will now look at what a time delay of T means in the Laplace domain.

Starting with the definition of the Laplace transform, we have:

$$L\{f(t-T)\,u(t-T)\} = \int_0^{\infty} t\; f(t-T)\,u(t-T)\,e^{-st} d$$

$$= \int_T^{\infty} f(t-T)\,e^{-st}\,dt \qquad = \int_0^{\infty} f(\tau)\,e^{-s(\tau+T)} d\tau$$

$$= e^{-Ts} \int_0^{\infty} f(\tau)\,e^{-s\tau} d\tau \quad = e^{-Ts} F(s)$$

A time delay of T is equivalent to multiplying by e^{-sT} in the Laplace domain. Since we defined the z^{-1} operator as a time delay of T, we will now define the z transform by replacing e^{-sT} by z^{-1} (or e^{sT} by z):

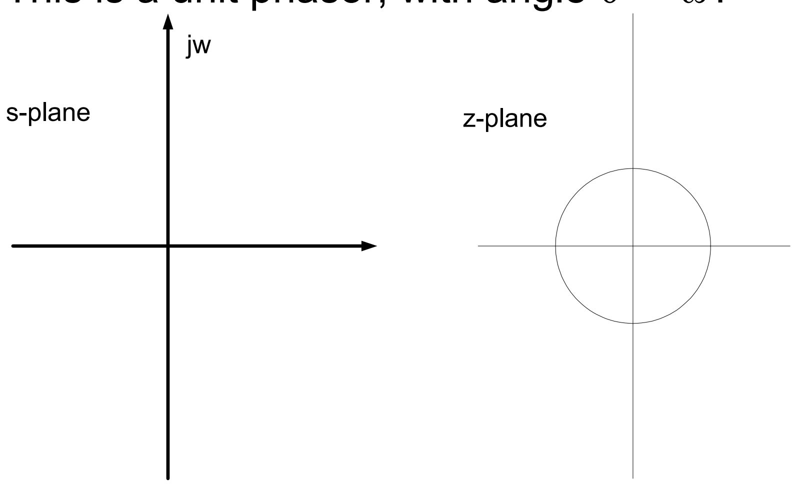
$$F(z) = \sum_{0}^{\infty} f(nT) z^{-n}$$

[The integral changes to summation as we are only interested in discrete time instances where t = nT.]

With $z = e^{sT}$, let us consider the mapping of the s-plane to the z-plane.

If s=jω, $z= e^{j\omega T} = \cos \omega T + j \sin \omega T$.

This is a unit phasor, with angle θ = ωT



Thus, the jω axis of the s-plane maps on to the unit circle of the z-plane.

When ω = 0, θ = 0.
When ω = π /T, θ = π
When ω = - π /T, θ = - π

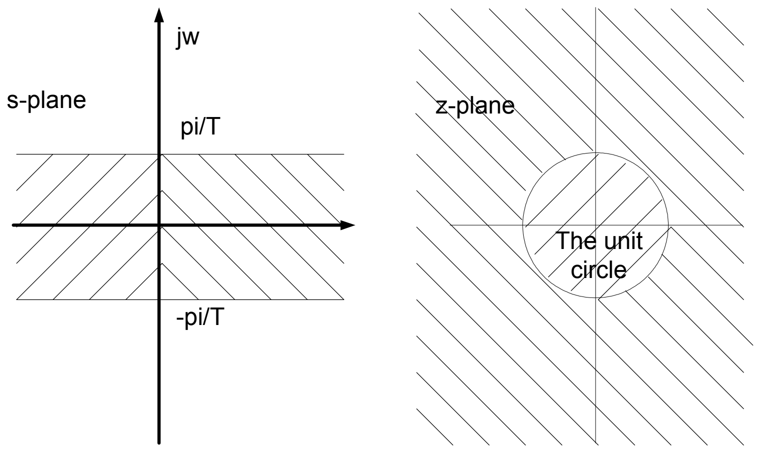
As ω goes from - π /T to π /T, the circle on the z-plane completes one cycle. As ω goes from -∞ to + ∞, the circle is traversed an infinite number of times, so that the mapping from the s-plane to the z-plane is obviously not one-to-one.

Let us now study the strip of the s-plane lying between - π /T and π /T. We have already seen that when the real part of s is zero, the imaginary axis of the s-p;ane is mapped on to the unit circle on the z-plane.

When s = σ+ jω,

$z= e^{(\sigma+j\omega)T} = e^{\sigma T}$ (cos ωT + j sin ωT).

On the left half of the s-plane (σ negative), we have the magnitude of the rotating phasor less than unity (going from zero at σ = - ∞ to one at σ = 0), and on the right half of the s-plane, we have its magnitude going from one (at σ = 0) to ∞ (at σ = ∞).



Each strip of height 2π/T on the s-plane maps on to the complete z-plane, with the left half mapping on to the inside of the unit circle, and the right half on to the outside of the unit circle.

1.3.1 Convolution

We will restate what was said about the output of a system block being the convolution of the input signal and the impulse response of the system block in mathematical form, with the unit delta function as the input signal δ(t), where. δ(t) is defined as:

$$\delta(t) = 0, \; for \; t \neq 0$$

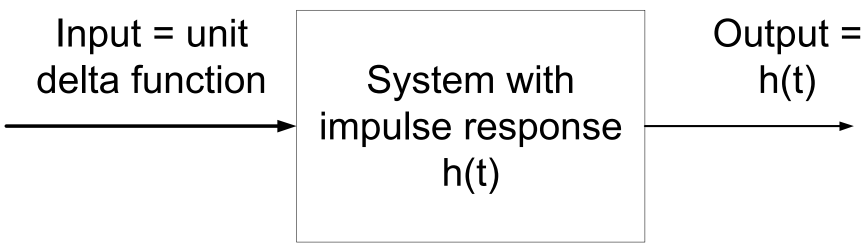
$$\int_{-\infty}^{\infty} \delta(t)\, dt = 1$$

Alternatively, we may define δ(t) by its property of extracting the value of a function at t=0:

$$x(t)\,\delta(t) = x(0)$$

or, more generally,

$$x(t)\,\delta(t - t_0) = x(t_0)$$



If the input to a system is the unit delta function, then its output is the impulse response of the system.

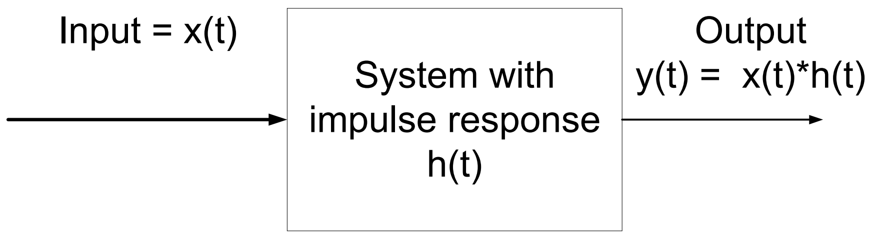
In practice, we are concerned only with causal signals, (a causal signal is one arising out of a cause, and so will be zero for all time less than zero), so that we can define:

$$x(t) = \int_0^\infty x(\tau)\,\delta(t-\tau)\,d\tau$$

The output y(t) for an arbitrary (but well-behaved) input x(t) is given by:

$$\begin{aligned} y(t) \quad &= \int_0^\infty x(\tau)\,h(t-\tau)\,d\tau \\ &= x(t) * h(t), \quad t \ge 0 \end{aligned}$$

This is known as the convolution of x(t) and h(t).



We will assume that the integrals exist, that is:

$$\int_0^t |x(\tau)|\,d\tau < \infty,$$

$$\int_0^t |h(\tau)|\,d\tau < \infty$$

Using the definition

$$\int_0^\infty x(\tau)\,h(t-\tau)\,d\tau = x(t) * h(t)$$

of convolution, it is easy to show that it is commutative, associative and distributive.

Using the notation:

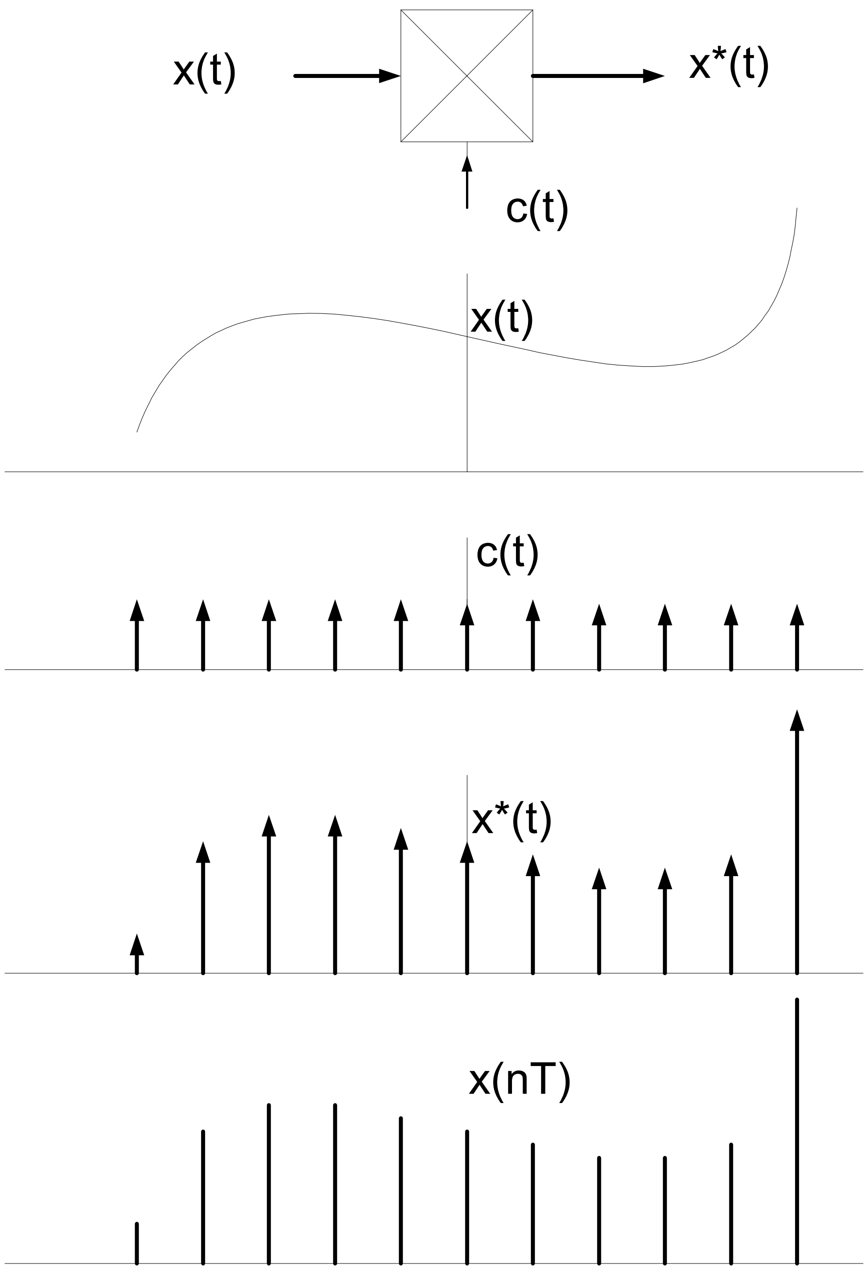
$$F[x(t)] = X(j\omega)$$
$$F^{-1}[X(j\omega)] = x(t)$$

to indicate the Fourier transform and the inverse Fourier transform, we can also show that:

$$x(t) * y(t) \leftrightarrow X(j\omega)\, Y(j\omega)$$

$$x(t)\ \ y(t) \leftrightarrow X(j\omega) * Y(j\omega)$$

With this background, we will now look at the sampling of a continuous signal x(t) as a process of multiplication (or modulation) by an ideal sampler.



The diagram represents an ideal sampler, a continuous time signal x(t), the sampling function c(t), the sampled signal x*(t) and the discrete signal x(nT). The interval between samples is T.

By looking at the sampler as a modulator, we may write:

$$
\begin{aligned}
c(t) &= \sum_{n=-\infty}^{\infty} \delta(t-nT) \\
x^*(t) &= c(t)\,x(t) \\
&= \sum_{n=-\infty}^{\infty} x(nT)\,\delta(t-nT) \\
&= \sum_{n=-\infty}^{\infty} x(nT)
\end{aligned}
$$

Fourier transforming this we get:

$$
\begin{aligned}
X^*(j\omega) &= F[x^*(t)] \\
&= F\sum_{n=-\infty}^{\infty} x(nT)\ \delta(t-nT) \\
&= \sum_{n=-\infty}^{\infty} x(nT)\,e^{-j\omega nT}
\end{aligned}
$$

We will need the convolution integral (which we stated, without proof earlier) to continue. We will consider its proof first:

It was stated that convolution in the time domain is equivalent to multiplication in the Laplace domain.

$$
L\left[\int_0^t f(t-\tau)\,g(\tau)d\tau\right] = F(s)\,G(s)
$$

Starting from the definition:

$$
\begin{aligned}
H(s) &= L\left[\int_0^t f(t-\tau)\,g(\tau)\,d\tau\right] \\
&= \int_0^\infty \int_0^t \ f(t-\tau)\,g(\tau)\,d\tau\,e^{-st}dt \\
&= \int_0^\infty \int_0^\infty \ f(t-\tau)\,g(\tau)u(t-\tau)\,d\tau\,e^{-st}dt
\end{aligned}
$$

[Note: Limits of integration have been extended from t to ∝, after multiplying by delayed unit step.]

By interchanging the order of integration,

$$H(s) = \int_0^\infty g(\tau)\ \ d\tau \int_0^\infty \ f(t-\tau)u(t-\tau)\,e^{-st}dt$$

By changing the variable of integration, we can then write:

$$\begin{aligned} H(s) \quad &= \int_0^\infty g(\tau)\,e^{-s\tau}\,d\tau \int_0^\infty f(t)e^{-st}dt \\ &= G(s)F(s) \end{aligned}$$

1.3.2 Z Transforms

Continuous time ⇔ Laplace Transform

Discrete time ⇔ Z Transform

The dynamics of linear discrete time control systems are described by linear difference equations

Difference equations ⇒ Algebraic equations

Representation of discrete signals:

x(0), x(T), x(2T), . . . ,x(kT), . . .
x(0), x(1), x(2), . . . ,x(k), . . .

(Sampling period T)

The *bilateral z transform* of the discrete-time signal *x(n)* is defined to be

$$: X(z) \underset{=}{\Delta} \sum_{n=-\infty}^{\infty} x(n)\,z^{-n} \qquad - \qquad Bilateral\ \ z\ \ transform$$

For causal systems, we need to consider only Unilateral z transform:

$$X(z) \triangleq \sum_{n=0}^{\infty} x(n)\, z^{-n} \qquad - \qquad Unilateral\ z\ transform$$

Consider the sequence:.

$$x = [\ldots . 0,0,1,2,3,0,0, \ldots .]$$

$$\Updownarrow$$

$$X(z) = 1 + 2z^{-1} + 3z^{-2} = 1 + 2z^{-1} + 3[z^{-1}]^2$$

The *z* transform of a signal x will always *exist* provided

(1) the signal starts at a finite time and

(2) it is asymptotically exponentially bounded, i.e., there exists a finite integer n_f, and finite real numbers $A \geq 0$ and σ, such that

$$|x(n)| < Ae^{\sigma n} \ \ for\ all\ n \geq n_f .$$

The bounding exponential may be growing with n (σ>0). These are not the most general conditions for existence of the *z* transform, but they are sufficient for our purposes.

That is: There exists a finite integer n_f and finite real numbers A ≥ 0 and σ such that $|x[n]| < Ae^{\sigma n}$ for all $n \geq n_f$.

Evaluating z transforms

We will consider a few common examples. [We assume that if x(t) is discontinuous at any point, its value is taken to be that obtained by approaching from the right.]

Unit Impulse (Kronecker delta)

$$\delta_0(k) = \begin{Bmatrix} 1, & for\, k = 0 \\ 0, & for\, k \neq 0 \end{Bmatrix}$$

$$Z\{\delta_0(k)\} = 1$$

$$\delta_0(n-k) = \begin{Bmatrix} 1, & for\, n = k \\ 0, & for\, n \neq k \end{Bmatrix}$$

$$Z\{\delta_0(k)\} = z^{-k}$$

Unit step

$$x(t) = \begin{Bmatrix} 1, for \ \ 0 \le t \\ 0, for \ t < 0 \end{Bmatrix}$$

$$After\, sampling:$$

$$x(nT) = \begin{Bmatrix} 1, \ for \ 0 \le n \\ 0, \ for\, n < 0 \end{Bmatrix}$$

$$X(z) = \sum_0^{\infty} z^{-k} \ = 1 + z^{-1} + z^{-2} +, \ , \ , = \frac{1}{1 - z^{-1}} = \frac{z}{z-1}$$

Unit ramp

$$x(t) = \begin{Bmatrix} t, for \ \ 0 \le t \\ 0, for \ t < 0 \end{Bmatrix}$$

$$After\, sampling:$$

$$x(kT) = \begin{Bmatrix} kT, \ for \ 0 \le k \\ 0, \ for\, k < 0 \end{Bmatrix}$$

$$X(z) = \sum_0^{\infty} x(kT) z^{-k} \ = \sum_0^{\infty} kTz^{-k} \ = T\sum_0^{\infty} kz^{-k}$$

$$= T\left[z^{-1} + 2z^{-2} + 3z^{-3} +. \ . \ . \ .\right] = \frac{Tz^{-1}}{(1 - z^{-1})^2} = \frac{Tz}{(z-1)^2}$$

Polynomial function

$$x(k) = \begin{Bmatrix} a^k, \ for\, 0 \le k \\ 0, \ for \ 0 > k \end{Bmatrix}$$

$$X(z) = \sum_0^{\infty} x(k) z^{-k} \ = \sum_0^{\infty} a^k z^{-k} \ = 1 + az^{-1} + a^2 z^{-2} +. \ .. \ . = \frac{1}{1 - az^{-1}} = \frac{z}{z-a}$$

Exponential function

$$x(t) = \begin{Bmatrix} e^{-at}, & for\ 0 \le t \\ 0, & for\ 0 > t \end{Bmatrix}$$

$$x(kT) = \begin{Bmatrix} e^{-akT}, & for\ 0 \le k \\ 0, & for\ 0 > k \end{Bmatrix}$$

$$X(z) = \sum_{0}^{\infty} x(kT) z^{-k} = \sum_{0}^{\infty} e^{-akT} z^{-k} = 1 + e^{-aT} z^{-1} + e^{-2aT} z^{-2} + \ldots$$

$$= \frac{1}{1 - e^{-aT} z^{-1}} = \frac{z}{z - e^{-aT}}$$

Sinusoidal function

$$x(t) = \begin{Bmatrix} \sin \omega t, & for\ 0 \le t \\ 0, & for 0 > t \end{Bmatrix} = \begin{Bmatrix} \frac{1}{2j}(e^{j\omega t} - e^{-j\omega t}), & for\ 0 \le t \\ 0, & for 0 > t \end{Bmatrix}$$

$$x(kT) = \begin{Bmatrix} \frac{1}{2j}(e^{jk\omega T} - e^{-jk\omega T}), & for\ 0 \le k \\ 0, & for\ 0 > k \end{Bmatrix}$$

$$X(z) = \sum_{0}^{\infty} \frac{1}{2j}(e^{jk\omega T} - e^{-jk\omega T}) z^{-k} = \frac{1}{2j}\left[\frac{1}{1 - e^{j\omega T} z^{-1}} - \frac{1}{1 - e^{-j\omega T} z^{-1}}\right]$$

$$= \frac{1}{2j} \frac{(e^{j\omega T} - e^{-j\omega T}) z^{-1}}{1 - (e^{j\omega T} + e^{-j\omega T}) z^{-1} + z^{-2}} = \frac{(\sin \omega T) z^{-1}}{1 - 2(\cos \omega T) z^{-1} + z^{-2}} = \frac{z \sin \omega T}{z^2 - 2z \cos \omega T + 1}$$

Cosine function

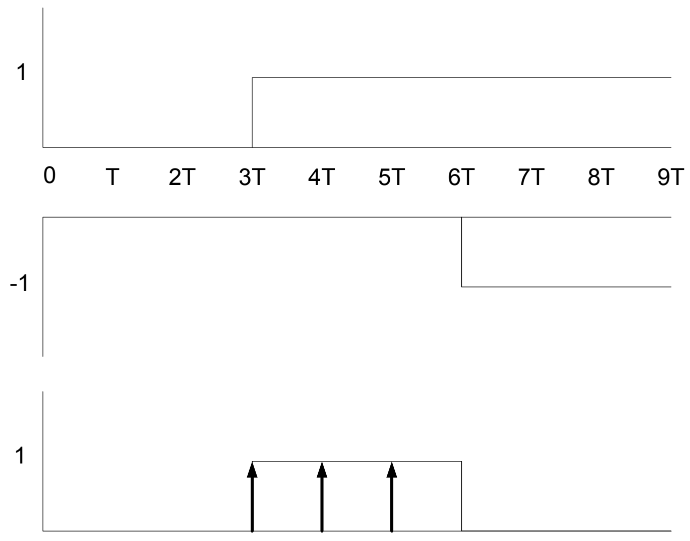
$$x(kT) = \begin{Bmatrix} \cos k\omega T = \frac{1}{2}(e^{jk\omega T} + e^{-jk\omega T}), & for\ 0 \le k \\ 0, & for\ 0 > k \end{Bmatrix}$$

$$X(z) = \frac{z^2 - z \cos \omega T}{z^2 - 2z \cos \omega T + 1}$$

Example

x(kT)={0,0,0,1,1,1,0,0,….]

This may be interpreted as the sum of a unit step function delayed by three sampling time periods and a negative unit step function delayed by six periods.



$$x(kT) = 1 \Leftrightarrow X(z) = \frac{1}{1 - z^{-1}}$$

$$x(kT - 3T) \Leftrightarrow z^{-3}\frac{1}{1 - z^{-1}} = \frac{z^{-3}}{1 - z^{-1}}$$

$$-x(kT - 6T) \Leftrightarrow \frac{-z^{-6}}{1 - z^{-1}}$$

$$x(kT - 3T) - x(kT - 6T) \Leftrightarrow \frac{z^{-3}}{1 - z^{-1}} - \frac{z^{-6}}{1 - z^{-1}} = \frac{z^{-3}(1 - z^{-3})}{(1 - z^{-1})}$$

$$= z^{-3}(1 + z^{-1} + z^{-2}) = z^{-4} + z^{-5} + z^{-6}$$

This could have been written down directly by inspection!

Properties

We did not consider the properties of the Fourier transform (or of the DFT) in detail, leaving it for consideration under the z transform. There are parallel properties among all these transforms, and here we will look at those of the z transform. The properties of other transforms may be deduced as and when necessary.

Multiplication by a constant

$$x(kT) \Leftrightarrow X(z)$$
$$a.x(kT) \Leftrightarrow aX(z)$$

Linearity

$$x(k) \Leftrightarrow X(z)$$
$$\alpha\, f(k) + \beta\, g(k) \Leftrightarrow \alpha\, F(z) + \beta\, G(z)$$

Multiplication by a^k

$$x(k) \Leftrightarrow X(z)$$
$$a^k\, x(k) \Leftrightarrow X(a^{-1}z)$$

Shifting Theorem (for n ≥ 0)

$$x(kT) \Leftrightarrow X(z)$$
$$x(kT - nT) \Leftrightarrow z^{-n} X(z)$$
$$x((kT + nT) \Leftrightarrow z^n X(z) - z^n \sum_{k=0}^{n-1} x(kT) z^{-k}$$

Complex translation

$$x(nT) \Leftrightarrow X(z)$$
$$e^{-anT} x(nT) \Leftrightarrow X(ze^{aT})$$

[*Transform of* $e^{-anT}x(nT)$ *is obtained by replacing z with* ze^{aT} *in the transform of* $x(nT)$]

Initial value Theorem

If $x(k) \Leftrightarrow X(z)$

Then $x(0) = \lim_{z\to\infty} X(z)$, *provided it exists.*

Proof:

$$X(z) = \sum_{0}^{\infty} x(k)z^{-k} = x(0) + x(1)z^{-1} + x(2)z^{-2} + \ldots$$

$$If \;\; you \; let \; z \to \infty, \; X(z) \to x(0)$$

Final value Theorem

$$If \;\; x(k) \Leftrightarrow X(z)$$

$$Then \; \lim_{k\to\infty} x(k) = \lim_{z\to 1}\left[(1 - z^{-1})X(z)\right]$$

$$provided \; that \; all \; poles \; of \; X(z) \; lie \; within \; the \; unit \; circle,$$

$$with \; the \; possible \; exception \; of \; a \; simple \; pole \; at \; z = 1$$

Proof:

$$x(k) \Leftrightarrow X(z)$$

$$\sum_{0}^{\infty} x(k)z^{-k} - \sum_{0}^{\infty} x(k-1)z^{-k} \Leftrightarrow X(z) - z^{-1}X(z)$$

$$As \; z \to 1$$

$$\lim_{z\to 1}\left[X(z) - z^{-1}X(z)\right] \to \left[\sum_{0}^{\infty} x(k) - x(k-1)\right]$$

$$= \left[\sum_{0}^{\infty} [x(0) - x(-1)] + [x(1) - x(0)] + ..\right] = \lim_{k\to\infty} x(k), \; as \; x(-1) \, is \; zero$$

Complex differentiation

$$If \;\; x(k) \Leftrightarrow X(z)$$

$$Then \; k^m x(k) \Leftrightarrow \left(-z\frac{d}{dz}\right)^m X(z)$$

$$\left[\left(-z\frac{d}{dz}\right)^m \; means \; that \; the \; operator \; \left(-z\frac{d}{dz}\right) is \; applied \; m \; times.\right]$$

Proof

$$X(z) = \sum_{0}^{\infty} x(k) z^{-k}$$

$$\frac{d}{dz} X(z) = \sum_{0}^{\infty} (-k) x(k) z^{-k-1}$$

$$\therefore -z \frac{d}{dz} X(z) = \sum_{0}^{\infty} k\, x(k) z^{-1}$$

This operation may be repeated

Complex integration

$$If \;\; x(k) \Leftrightarrow X(z),\; g(k) \Leftrightarrow G(z)$$

$$and \;\; g(k) = \frac{x(k)}{k}$$

$$[Assume \;\; that \; \frac{x(k)}{k} exists \;\; for \;\; k = 0]$$

$$Then \;\; G(z) = \int_{z}^{\infty} \frac{X(z_1)}{z_1} dz_1 + \lim_{k \to 0} \frac{x(k)}{k}$$

Proof

$$-\frac{d}{dz} G)z) = \frac{X(z)}{z} - from \;\; complex \;\; differentiation$$

Integrating both sides from z to ∞:

$$\left[G(z)\right]_{\infty}^{z} = \int_{z_1}^{\infty} \frac{X(z_1)}{z} dz_1$$

Invoking the initial value theorem:

$$G(\infty) = \lim_{z \to \infty} G(z) = g(0) = \lim_{k \to 0} \frac{x(k)}{k}$$

We have the result:

$$G(z) = \int_{z}^{\infty} \frac{X(z_1)}{z_1} dz_1 + \lim_{k \to 0} \frac{x(k)}{k}$$

Real Convolution Theorem

$$If\ \ x_x(kT) \Leftrightarrow X_1(z)\ \ and\ x_2(kT) \Leftrightarrow X_2(z)$$

$$Then\ \sum_{n=0}^{k} x_1(nT)x_2(kT-nT) \Leftrightarrow X_1(z)X_2(z)$$

Proof

$$Z\left[\sum_{n=0}^{k} x_1(nT)x_2(kT-nT)\right] = \sum_{k=0}^{\infty}\ \sum_{n=0}^{k} x_1(nT)x_2(kT-nT)z^{-k}$$

Inverse z transform

There are a number of techniques that can be used to obtain the inverse z transform. Some of these are:

Direct division
Cauchy product
Partial fraction expansion
Inversion

We will study some of them through examples.

Direct Division

Expand X(z) as a power series in $z^{-1.}$

$$X(z) = \frac{1}{1 - \frac{3}{2}z^{-1} + \frac{1}{2}z^{-2}}$$

$$= 1 + \frac{3}{2}z^{-1} + \frac{7}{4}z^{-2} + \frac{15}{8}z^{-3} + .\ .\ .\ .$$

$$x(k) = \partial(kT) + \frac{3}{2}\delta(kT-T) + \frac{7}{4}\delta(kT-2T) + .\ .$$

In general we have:

$$X(z) = \frac{b_0 + b_1 z^{-1} + \quad + b_m z^{-m}}{a_0 + a_1 z^{-1} + .\ .\ . + a_{mz^{-m}}}$$

$$x(n) = \frac{1}{a_0}\left[b_n - \sum_{0}^{k-1} x(kT)a_{n-k}\right]$$

Partial fraction expansion

$$X(z) = \frac{1}{1 - \frac{3}{2}z^{-1} + \frac{1}{2}z^{-2}}$$

$$= \frac{1}{(1 - z^{-1})(1 - \frac{1}{2}z^{-1})} .$$

$$= \frac{2}{1 - z^{-1}} - \frac{1}{1 - \frac{1}{2}z^{-1}}$$

$$x(kT) = 2u(kT) - \left(\frac{1}{2}\right)^k u(kT)$$

Inversion

Now let us attempt to obtain the inverse z transform. The z transform was defined as:

$$F(z) = \sum_0^{\infty} f(nT)z^{-n}$$

$$= f(0)z^0 + f(T)z^{-1} + f(2T)z^{-2} + \ldots + f(nT)z^{-n} + \ldots$$

We now invoke Cauchy's theorem on contour integrals:

$$I = \frac{1}{2\pi j}\oint_C z^k dz = \begin{Bmatrix} 1, for\, k = -1 \\ 0, for\, k \neq -1 \end{Bmatrix}$$

where C is any contour enclosing the origin of the z plane and where k is an integer.

To obtain f(nT), we multiply both sides of the equation defining F(z) by z^{n-1} to get:

$$z^{n-1}F(z) = f(0)z^{n-1} + f(T)z^{n-2} + f(2T)z^{n-3} + \ldots + f(nT)z^{-1} + \ldots$$

This gives:

$$f(nT) = \frac{1}{2\pi j}\oint_C z^{n-1}F(z)dz$$

The contour C has to be chosen so as to enclose all poles of F(z). For stable systems, where the poles are within the unit circle, w can use the unit circle as the contour of integration, and the integral may be evaluated by the method of residues.

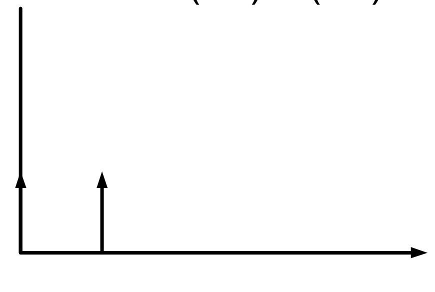
For practical applications, we usually obtain the inverse transforms from pre-computed tables. A short table of common transforms is given below:

Table of z Transforms

$X(s)$	$x(t)$	$x(kT)$ or $x(k)$	$X(z)$
–	–	Kronecker delta $\delta_0(k)$	1
–	–	$\delta_0(n-k)$	z^{-k}
$\frac{1}{s}$	$1(t)$	$1(k)$	$\frac{1}{1-z^{-1}}$
$\frac{1}{s+a}$	e^{-at}	e^{-akT}	$\frac{1}{1-e^{-aT}z^{-1}}$
$\frac{1}{s^2}$	t	kT	$\frac{Tz^{-1}}{(1-z^{-1})^2}$
$\frac{2}{s^3}$	t^2	$(kT)^2$	$\frac{T^2z^{-1}(1+z^{-1})}{(1-z^{-1})^3}$
$\frac{6}{s^4}$	t^3	$(kT)^3$	$\frac{T^3z^{-1}(1+4z^{-1}+z^{-2})}{(1-z^{-1})^4}$
$\frac{a}{s(s+a)}$	$1-e^{-at}$	$1-e^{-akT}$	$\frac{(1-e^{-aT})z^{-1}}{(1-z^{-1})(1-e^{-aT}z^{-1})}$

Let us consider the example that we took earlier:

$H(z)= 1+z^{-1}$
$h(nT)=\delta(nT)+\delta(nT-T)$



Let us consider another example:

$$H(z)=\frac{1}{1-\frac{3}{2}z^{-1}+\frac{1}{2}z^{-2}}$$

We will consider three different techniques for obtaining the inverse z transform of this function.

1, Long division

$$1+\frac{3}{2}z^{-1}+\frac{7}{4}z^{-2}+\frac{15}{8}z^{-3}+..$$

$$1-\frac{3}{2}z^{-1}+\frac{1}{2}z^{-2} \;\Big)\; 1$$

$$1-\frac{3}{2}z^{-1}+\frac{1}{2}z^{-2}$$

$$\frac{3}{2}z^{-1}-\frac{1}{2}z^{-2}$$

$$\frac{3}{2}z^{-1}-\frac{9}{4}z^{-2}+\frac{3}{4}z^{-3}$$

$$\frac{7}{4}z^{-2}-\frac{3}{4}z^{-3}$$

$$\frac{7}{4}z^{-1}-\frac{21}{8}z^{-3}+\frac{7}{8}z^{-4}$$

$$\frac{15}{8}z^{-3}+\frac{7}{8}z^{-4}$$

$$H(z)=1+\frac{3}{2}z^{-1}+\frac{7}{4}z^{-2}+\frac{15}{8}z^{-3}+\ldots$$

$$h(nT)=\delta(nT)+\frac{3}{2}\delta(nT-T)+\frac{7}{4}\delta(nT-2T)+\ldots$$

2. Partial fraction expansion

$$H(z)\quad=\frac{1}{1-\frac{3}{2}z^{-1}+\frac{1}{2}z^{-2}}=\frac{1}{(1-z^{-1})(1-\frac{1}{2}z^{-1})}=\frac{2}{1-z^{-1}}-\frac{1}{1-\frac{1}{2}z^{-2}}$$

$$h(nT)\quad=2u(nT)-\left(\frac{1}{2}\right)^{n}u(nT)$$

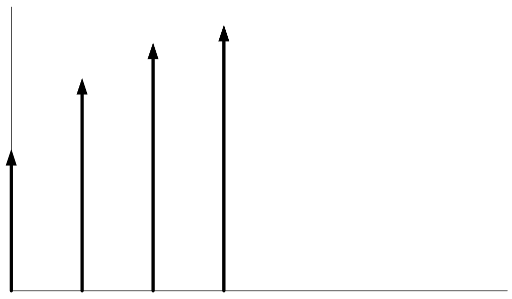
We can verify that these two results are identical:

n=0: h(0) = 2-1 = 1

n=1: h(T) =2 – ½ = 3/2

n=2: h(2T) = 2 – ¼ = 7/4

n=3: h(3T) = 2 – 1/8 =15/8 etc.



:

If

$$H(z) = \frac{b_0 + b_1 z^{-1} + . \ . \ . \ . + b_m z^{-m}}{a_0 + a_1 z^{-1} + . \ . \ . \ . + a_0 z^{-m}}$$

then,

$$h(nT) = \frac{1}{a_0}\left[b_n + \sum_{k=0}^{n-1} h(kT)\, a_{n-k} \right]$$

When

$$H(z) = \frac{1}{1 - \frac{3}{2} z^{-1} + \frac{1}{2} z^{-2}}$$

n=0: h(0) = 1/1(1) = 1

n=1: h(T) = 1/1 [0 – h(0) a_1] = 3/2

n=2: h(2T) = 1/1 [0 – {h(0) a2 + h(T) a1}]

= (-1) . ½ + 3/2 . 3/2 = 7/4

n=3: h(3T) = 1/1 [0 – {h(T) a2 + h(2T) a1}]

= -(3./2 . ½ - 7/4 . 3/2) = 15/8

etc.

1.3.3. The w-plane

The idea of a still another transformation will need some explanation.

The complex variable s (= σ + j ω) of the Laplace transform has a very natural place for the natural frequency ω, but this simplicity is lost with sampling and the resulting z transform. The relationship between z and s ($z = e^{sT}$) makes the meaning of frequency rather obscure in the z-plane. As frequency domain analysis has been very well developed over a long period of time, it is natural to want to extend these ideas and techniques to the study of sampled data systems.

We have already noted the mapping of the s-plane to the z-plane, where a strip of width $2\pi/T$ (where T is the sampling period) maps on to the entire z-plane. The $j\omega$ axis maps on to the unit circle, repeatedly. This does not facilitate the use of techniques such as the Bode plot, where ω is the independent variable, in the z-plane; and has provided the motivation for the new transformation.

Consider the transformation defined by:

$$\omega = u + jv = \frac{z-1}{z+1} = \frac{1-z^{-1}}{1+z^{-1}}$$

From the above, we may deduce the inverse transformation as:

$$z = \frac{1+\omega}{1-\omega}$$

[Note that, in common with the other transformations that we discussed, different proportionality constants are sometimes used in the definition of the transform. Another common form is the following:

$$z = \frac{1+(T/2)\omega}{1-(T/2)\omega}, \qquad \omega = \frac{2}{T}\frac{z-1}{z+1}$$

We will use the former version, without the T/2 factor.]

If we consider the range of real frequencies ω, corresponding to the imaginary axis of the s-plane, we have:

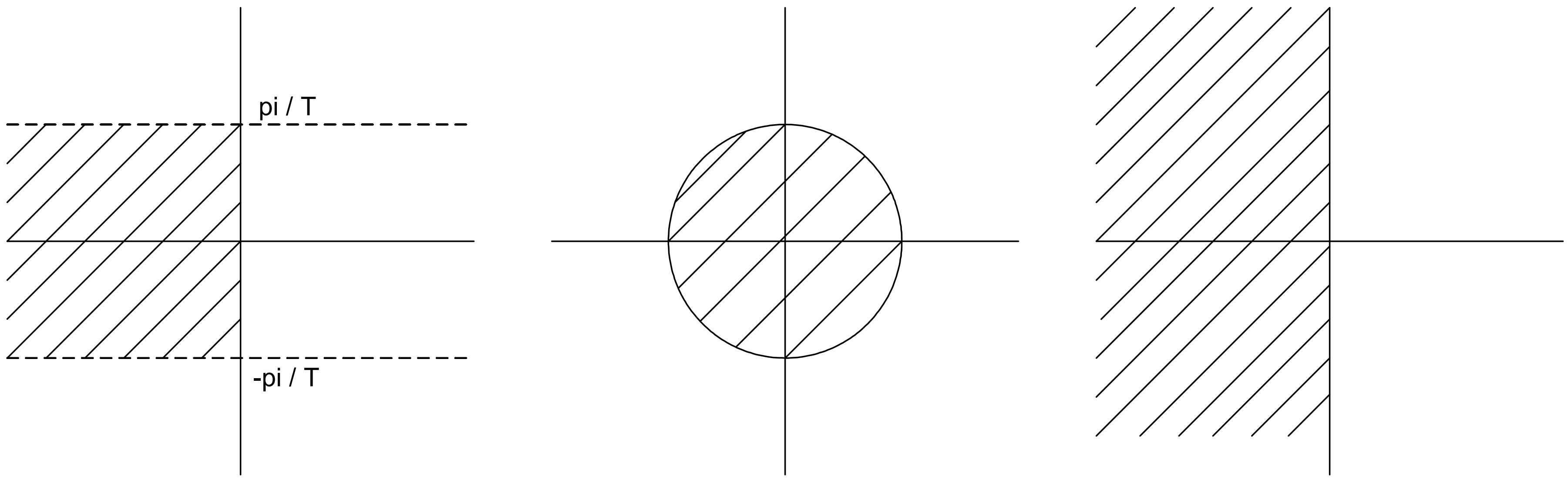
$$z = e^{sT} = e^{j\omega T}$$

$$\omega = u + jv = \frac{z-1}{z+1} = \frac{e^{j\omega T}-1}{e^{j\omega T}+1} = j\tan\left(\frac{\omega T}{2}\right)$$

$$v = \tan\left(\frac{\omega T}{2}\right)$$

$$\omega = \frac{2}{T}\tan^{-1} v$$

We can now attempt to map the z-plane to the w-plane:



The s-plane The z-plane The w-plane

The left half strip of height 2π / T on the s-plane maps on to the inside of the unit circle on the z-plane. This maps on to the *entire* left half of the w-plane. Thus we see that even though similar in some senses, the s-plane and the w-plane are not the same.

Some of the other points of interest are:

-∝ on the s-plane maps on the origin on the z-plane and to (-1,0) on the w-plane.
± j π / T on the s-plane maps on to (-1,0) on the z-plane and ± j ∝ on the w-plane.

Even though the two planes are different from each other, the similarities allow us to use the w-plane as a convenient tool to study sampled data control systems using the classical techniques.