- 1 -
Basic Techniques

Background Techniques

1.1 A Little Mathematics

The study of acoustics is greatly simplified with the use of complex numbers. The complex exponential, an exponential raised to a complex power, is a periodic oscillatory function

$$e^{-i\omega t} \qquad (1.1.1)$$

where

$$e^{iz} = \cos(z) + i\sin(z) \qquad (1.1.2)$$

In this example the variable t represents time and the function, as shown in reference. It is a vector of unit amplitude, which rotates in the complex plane at a rate of ω radians per seconds. The value of ωt at any point in time yields the phase of the complex exponential at that moment. The real part of this vector is an oscillating function – a sine wave at the frequency $\omega/2\pi$

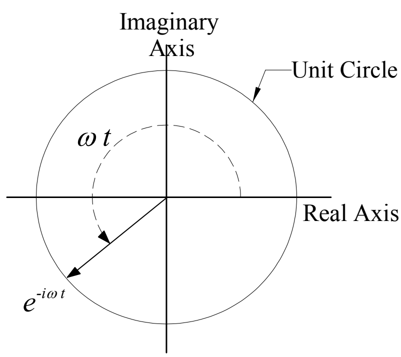


Figure 1-1 - The complex exponential at time t - rotate by ωt radians

There is no difference in the complex exponential for positive i or negative i other than the direction of rotation. For the time variable, however, we can choose only one or the other, but the choice is arbitrary. In this text, the convention for the complex exponential in time will be the negative sign. This is the physicist's convention, which is seemingly a different convention than that used by electronic engineers. However, since $i = -j$, they are, in fact, equivalent, although one must be careful not to read the equations as if i and j were identical.

Any linear signal can be described as a weighted sum of complex exponentials, each one at a different frequency and with different amplitudes. If the amplitude value is itself complex then this value can carry both the magnitude and phase information at any given frequency. Thus, a complex function times a complex exponential contains virtually all of the information that we need.

The complex exponential in time can also be combined with a complex exponential in space. The resulting forms

$$e^{-i(kx+\omega\, t)},\ e^{i(kx-\omega\, t)}$$

represent waves in the positive and negative x directions respectively. The variable $k=2\pi/\lambda$ is the wavenumber. Here, we must retain both signs since the waves are free to propagate in either direction.

Since virtually all of our equations will represents motions in time (static acoustics is not our interest) it will be easier to simply drop the time exponential. Throughout this text the time function will almost always be ignored as superfluous. It may appear at times for clarity such as in the case of a time derivative where we need to explicitly show a time dependence for the derivative to make sense.

1.2 The Fourier Transform and Fourier Series[1]

In the previous section we mentioned the fact that a function of a complex variable with a complex exponential time factor can represent nearly any signal of interest. The *Fourier Transform* formalizes this representation. It is important to distinguish between the subtle differences between the Fourier Transform, the *Fourier Series* and the *Fast Fourier Transform* (FFT).

The FFT is more closely related to the Fourier Series than the Fourier Transform. It is a computer implementation of the complex exponential Fourier Series. The FFT is sometimes thought of as a computer implementation of the Fourier Transform, but this is not exactly correct. Since the FFT deals only with discrete points and sets of data, the FFT is a discrete set of numbers – a series. The FFT is a set of discrete functions defined on a finite interval. The Fourier Series is a set of continuous functions defined on a finite interval, and the Fourier Transform is a continuous function defined on an infinite interval.

The definition of the Fourier Transform is

$$F(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)\cdot e^{i\omega t}\,dt \tag{1.2.3}$$

and

$$f(t)=\int_{-\infty}^{\infty} F(\omega)\cdot e^{-i\omega t}\,d\omega \tag{1.2.4}$$

1. See Churchill, *Operational Mathematics*

The Fourier Transform is a pair of equations that map functions between two domains. The domains in the equations above are the time domain and the frequency domain. The function $e^{-i\omega t}$ is called the kernel of the transform and contains the variables time and frequency – the two domains of the transform. The Fourier Transform can map between many different domains such as spatial coordinates and wavenumber. We will see several uses for this later transform.

It should be noted that when discussing systems the Fourier Transform can only be applied when the system is linear. There are generalizations of system theory that apply to nonlinear systems, but we will postpone that discussion until Chap. 10. The system referred to here is whatever means the signal uses to travel on its path from the input to the output. This means that in a very real sense the Fourier Transform should not be applied to the transducer problem. We will also discuss this limitation in Chap. 12 when we talk about measurements.

A direct result of the assumption of linearity in the application of the Fourier Transform is the complete equivalence of the time and frequency domains. This equivalence requires that the frequency domain response completely characterizes the transient response of the system. The two are not independent. As long as the Fourier Transform is used as the basis of any discussion it is an absolute requirement that no information is obtainable in one domain that is available in the other. These facts have wide ranging implications. So long as we use concepts such as impedance or frequency response we have assumed that the system is linear and that the Fourier Transform (as well as the Laplace Transform) is valid. This means, for instance, that the impulse response of a system or the systems time response to *any* signal, no matter how we may look at it, can contain no information that is not also available to us in the frequency domain. We must not lose sight of these fundamental facts in our study. In this text, we will not discuss transient response, etc., preferring to do the mathematics in the frequency domain knowing full well that any time domain signals have also been completely determined. These calculations would thus be redundant.

The Fourier Series is defined as

$$f(x) = B_0 + \sum_{m=1}^{\infty} B_m \cos(mx) + A_m \sin(mx) \tag{1.2.5}$$

where

$$\begin{aligned} B_m &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\cos(mx)dx \\ A_m &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\sin(mx)dx \end{aligned} \tag{1.2.6}$$

One could think of Eq. (1.2.5) and Eq. (1.2.6) as a transform pair mapping from m space to x space with real kernels $\cos(mx)$ and $\sin(mx)$ or a complex kernel e^{imx}. If

$f(x)$ is symmetric about some point and we are free to define the location of the origin of x to be that point, then the A_m's will disappear. We will use this principle to some advantage.

1.3 Orthogonal Functions

A key concept in our development will be that of orthogonality. There are many facets to this concept, such as orthogonal Eigenvectors, etc., but one of most interest to us is the concept of orthogonal sets of functions.

A set of functions S_n is *orthogonal* on the interval $a|b$ if and only if

$$\int_a^b S_n(x) S_m(x) r(x) dx \neq 0 \tag{1.3.7}$$

when $n = m$ for some $r(x)$ in $a|b$.

As an example consider the functions $\cos(mx)$

$$\int_{-\pi}^{\pi} \cos(mx)\cos(nx) \cdot dx = \left\langle \begin{matrix} 0 \; if \; m \neq n \\ 2\pi \, m = n \end{matrix} \right. \tag{1.3.8}$$

Thus the functions $\cos(mx)$ are orthogonal on the interval $x = \langle \pi, -\pi \rangle$.

To see the usefulness of this concept consider an arbitrary function $f(x)$ defined in $x = \langle \pi, -\pi \rangle$. Assume that this function can be represented by a set of functions $\cos(mx)$

$$f(x) = \sum_{m=0}^{\infty} B_m \cos(mx) \tag{1.3.9}$$

How do we now find the values for the coefficients B_m?

Multiplying both sides by $\cos(mx)$ and integrating over the interval yields

$$\int_{-\pi}^{\pi} f(x)\cos(nx)dx = \int_{-\pi}^{\pi} \sum_m B_m \cos(nx)\cos(mx)dx = \sum_m B_m \int_{-\pi}^{\pi} \cos(nx)\cos(mx)dx \tag{1.3.10}$$

A direct result of orthogonality is that the right hand side equals 2π only when $n = m$ and zero otherwise. This allows us to replace n by m resulting in the definition of the B_m's

$$B_m = \frac{1}{2\pi} \int_0^{\pi} f(x)\cos(mx)dx \tag{1.3.11}$$

as in Eq. (1.2.6). The Fourier Series is but one orthogonal expansion. We will see many others.

An orthogonal set of functions is called *orthonormal* if Eq. (1.3.8) has the value of one. All of the orthogonal sets of functions described in this text are complete, which means that any function defined in the interval can be represented by a series of these functions, at any point, to any level of accuracy. This statement is made without proof.

It is also interesting to note that if we require that the series represents the expanded function in a least squares sense at *m* points, then it takes exactly $m + 1$ terms in the series to do that. This too is stated without proof. The FFT is a least squares approximation of the Fourier Transform at *n* points were *n* is the order of the FFT.

Orthogonality is a powerful tool that we shall often call upon in our studies.

1.4 Orthogonal Coordinate Systems[2]

The concepts of orthogonal coordinate systems have some intuitively obvious characteristics, but they also have some not so intuitive aspects. We will utilize many aspects of the orthogonal coordinate systems in future chapters, but a review of the background to this theory is appropriate here.

There are an infinite number of orthogonal coordinate systems; the only requirement being that each of the three coordinates (for three dimensional space) must be perpendicular to the other two at every point in space. Rectangular Coordinates obviously fit this requirement. If we further restrict ourselves to only those coordinate systems for which the Wave Equation is separable – they can be factored into three separate equations in the three dimensions – then we find that there are only eleven. Why are we so concerned with separable solutions? Because, if the Wave Equation is not separable, then there cannot be exact analytical solutions in terms of functions of a single spatial variable. As such there also cannot be solutions for which there is an orthogonal set in which to expand our solutions. Many of the more powerful techniques discussed in this book will depend on there being such a reasonably simple analytical solution.

Of these eleven coordinate systems, an even fewer number will be of much use to us here. The reason for this is simple. At very large distances from any finite source the wavefronts must be spherical or planar and there are only six Coordinates that meet these criteria. These are the Rectangular, Cylindrical, Spherical, Elliptic Cylinder, Oblate Spheroidal, Prolate Spheroidal and Ellipsoidal. It would appear that the two Cylindrical coordinate systems would not evolve into spherical waves at large distances, but we will see that they do, in fact, do so for finite sources.

Probably the most interesting for us would be the Ellipsoidal, since the other six coordinate systems are subsets of this single more general system. Unfortunately, there are no know solutions in this coordinate system which have the simplicity that we seek. Indeed, any solution that was general enough to include all other solutions would be very powerful. It is further interesting to note that the equations for all three dimensions of the Ellipsoidal Coordinates are the same,

2. See Morse and Feshbach, *Methods of Theoretical Physics*

but each has different boundary conditions and eigenvalues. If we had such a solution available to us, we would have an extraordinarily powerful capability in out hands. Alas, we have never found such a solution, although we have looked. For this reason, we will have to ignore the Ellipsoidal Coordinates for the present. We will however use all of the other five.

There are three basic Cylindrical coordinate systems that make up the five of interest. They all have the same z coordinate but differ in the two-dimensional representations that are orthogonal to the z coordinate. The two-dimensional representations are the Rectangular, which is very familiar to us; the polar, which is also familiar to us; and the Elliptical, which tends not to be so familiar. Fig. 1-2 shows the two-dimensional Elliptical coordinate system. Note that the origin of this coordinate system is not a point as we are used to thinking of it, but in this plane it is a short line segment of length $2d$. Extending this system into and out of the plane as shown generates the Elliptic Cylinder Coordinates, which have an origin that is a strip of width $2d$. Rotating this figure about the x-axis will generate the Oblate Spheroidal Coordinates with an origin that is a disk of radius d. Finally, the Prolate Spheroidal Coordinates are generated by rotating the figure about the y-axis, which has an origin that is a short line segment of length $2d$. Understanding the coordinate system shown in reference is fundamental to much of what we will discuss in latter chapters. Note that if we let $d = 0$, the above coordinate system will degenerate into polar coordinates and likewise a value of $d = \infty$ generates Cartesian Coordinates.

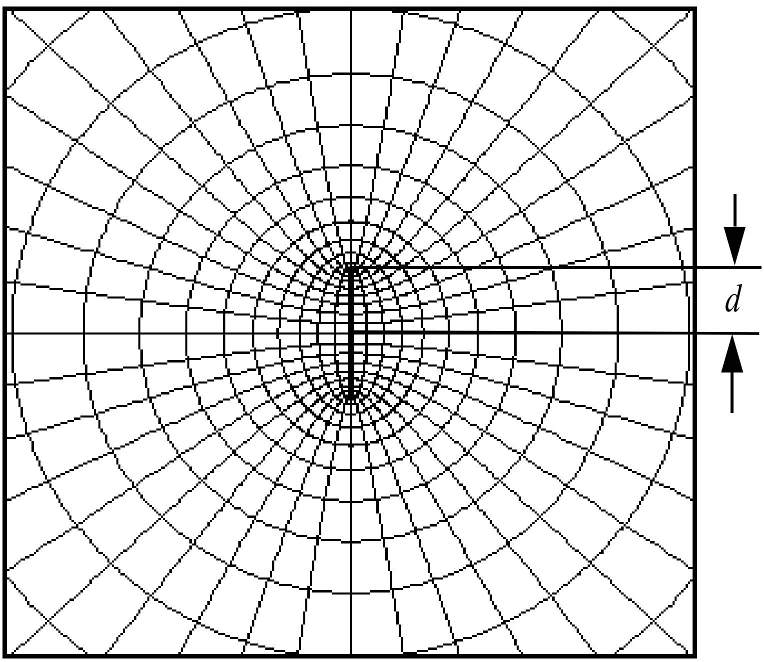


Figure 1-2 - Two dimensional elliptic coordinate system

A LITTLE ELECTRONICS

1.5 Electronic Basics

The field of audio is heavily linked with the field of electronics. However, there is really very little electronics involved with the study of transducers. The principal exception is the design of crossovers and of secondary signal processing techniques. Linear passive circuit theories are also used in the lumped parameter analysis of transducers. This section will introduce some fundamental electronics concepts as they apply to the transducer problem, but in general we would refer the interested reader to one of the multitude of alternative texts.

A concept that is critical to the study of electronics (and to acoustics for that matter) is that of impedance. Impedance is defined as the ratio of the complex voltage (pressure) across a component to the complex current (volume velocity) through it

$$z(\omega) = \frac{e(\omega)}{i(\omega)} \tag{1.5.12}$$

Note that impedance is always a complex quantity and that this is a frequency domain definition. From our previous discussions of the Fourier Transform, we know that, working in this domain has certain implications. The first is that the impedance concept can only be applied to a linear system. This means that discussing the impedance of a transducer is not correct since, in general, we know that transducers are not linear devices. For small signals, however, we will find the concept of impedance to be indispensable, although we must always keep in mind that this concept is not valid when the signals are large enough to generate nonlinearity.

We also know that when we are using impedance concepts, there is absolutely no difference between a discussion of the time domain or the frequency domain. The analog of the impedance concept in the time domain is that of the impulse response. The impulse response of a system and its impedance are completely specified in either domain. As we stated before, they are not separate and independent concepts.

In electrical engineering, it is common to use $s = -i\omega = j\omega$ as the frequency variable. This is the sign convention that will be used here, but not the letter convention. Since most of this text is about acoustics, which virtually never uses the variable s in the common literature, we will also neglect to use it. Therefore, the equations shown here may look different from the ones an electrical engineer would be accustomed to.

1.6 Passive Circuits[3]

There are three principal components in passive electronics: the resistor, the inductor and the capacitor. The impedance of each of these elements, as a function of frequency, is shown in Table 1. Notice that with various combinations of

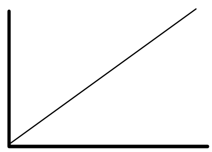
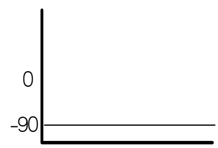
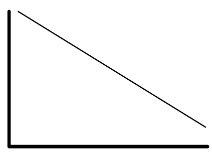
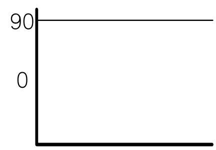
	$Z(\omega)$	Magnitude (log))	Phase (deg.))
Resistor	R		
Inductor	$i\omega L$		
Capacitor	$\frac{1}{i\omega C}$		

Table 1.1: Passive components characteristics

these three elements virtually any magnitude impedance could be obtained by the proper choice of the topology and the components. Also note that a choice of magnitude values will dictate a corresponding phase function. One is not free to choose the magnitude and phase independently for these components. This is known as the minimum phase requirement for a passive circuit.

These three circuit elements can be combined in an infinite number of ways, but there are a few fundamental arrangements that we should discuss. The first is a parallel combination of all three of these components, which generates the impedance shown in Fig. 1-3. This impedance is the parallel sum of the three

3. see Kuo, *Network Analysis and Synthesis*

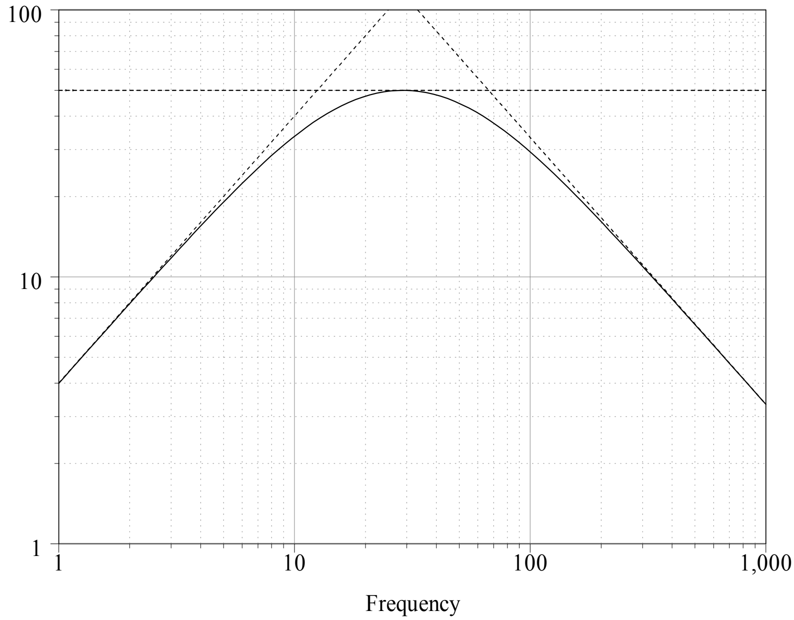


Figure 1-3 - Parallel RLC circuit impedance

impedances X_l, X_r and X_c. It has a peak value is at the point where $X_l = X_c$, also known as resonance.

The equation for the curve in Fig. 1-3 is simply

$$z(\omega) = \left(\frac{1}{X_r}+\frac{1}{X_l}+\frac{1}{X_c}\right)^{-1} = \frac{i\cdot\omega/C}{(\omega^2-\omega_0^2)+i2Q\omega\omega_0} \tag{1.6.13}$$

$\omega_0^2 = \frac{1}{LC}$ = *the resonance frequency*

$Q = R/2L$ = *the quality factor"- inverse of the damping*

The roots of the denominator in Eq. (1.6.13) are called the poles of the equation where

$$z(\omega) = \frac{i\omega/C}{(\omega-\omega_1)(\omega-\omega_2)} \tag{1.6.14}$$

The values ω_1 and ω_2 are the complex poles located at

$i\omega_0\left(\sqrt{Q^2-1}-Q\right)$ and $-i\omega_0\left(\sqrt{Q^2-1}+Q\right)$.

This equation also has a zero at $\omega=0$. The pole-zero form is convenient for a number of different reasons. When Q is large, this equation can be simplified into a set of symmetric – about the real axis – poles.

Another important combination of elements is a series LRC circuit. The impedance of this combination is shown in Fig. 1-4. In this configuration the impedance drops to a minimum at resonance, limited only by the resistance value. The impedance equation in pole-zero form is

$$z(\omega) = (\omega^2 - \omega_0^2) + i2Q\omega\omega_0 \tag{1.6.15}$$

where once again for large Q the equation simplifies considerably into a symmetric set of zeros. It is important to note that in our work high Q values are, in general, undesirable so we will usually have to retain the more complex form for these equations.

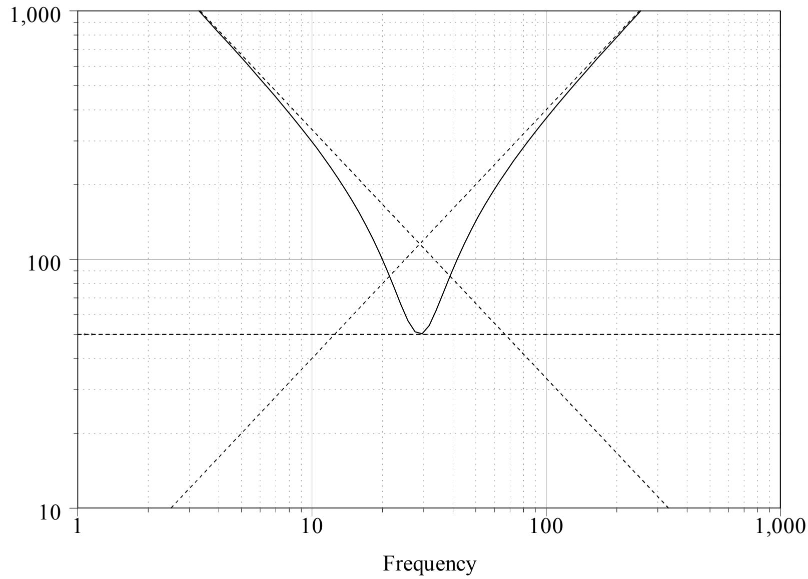


Figure 1-4 - Series RLC circuit impedance

Two more common passive element circuits are the low pass LR and the high pass RC combinations. These are simply a series inductance and a series capacitor loaded by a resistance. Fig. 1-5 shows the voltage transfer function for a low pass inductance circuit. Note that the transfer function has a value of .7 when the impedance of the inductor equals the impedance of the resistor at a frequency called cutoff. This filter passes frequencies below cutoff, and blocks frequencies above cutoff with an ever greater effectiveness – at the rate of 6 dB/octave.

Finally, the high pass transfer function of a RC circuit is also shown in Fig. 1-5. As in the previous example, the half power point, cutoff, is when the impedances

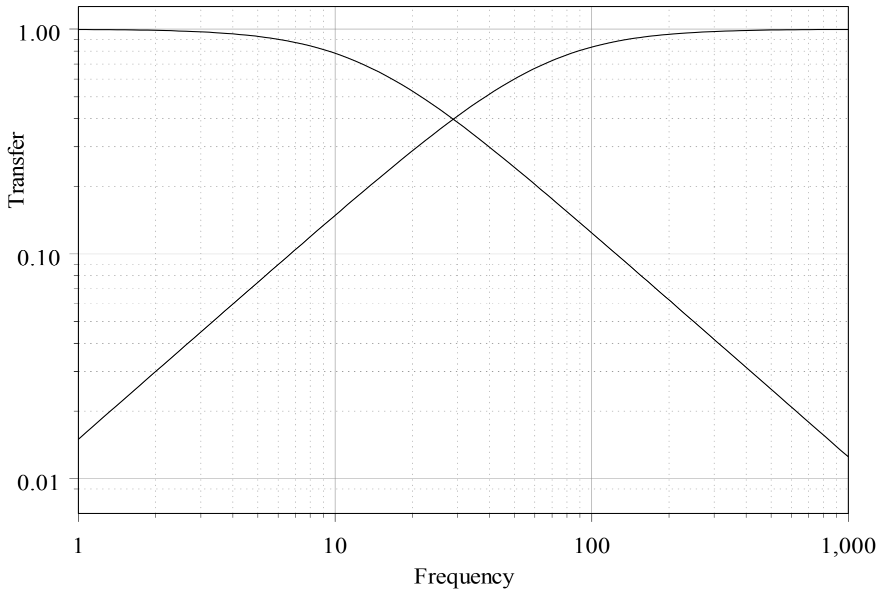


Figure 1-5 - Transfer function for an RC and LR circuit

of the two elements are equal. This combination of elements rejects low frequencies.

It is important to note that the curves shown in Fig. 1-5 have a pure resistance as the load of the reactive element. When this load is not a pure resistance, as is the case for a transducer load, then things are quite different. An example of what happens when the RC circuit has a moving coil loudspeaker load is shown in Fig. 1-6. The filters cutoff is below what we would like it to be. It is clear in this figure that loudspeaker has a substantial affect on the transfer function, basically moving the cutoff point well below what is predicted by a simple resistor load.

1.7 Active Circuits[4]

Active circuits are playing an ever increasingly important role in today's loudspeaker systems. They can be used to optimize the response or to dynamically modify amplifier characteristics for better reliability or control. However, these topics are beyond the scope of this text and we will only be discussing the most fundamental circuits.

Basically, there are two types of active circuits of interest to us – high-pass and low-pass. Active circuits are desirable, even necessary, for high slope crossovers

4. see Truxal, *Introductory Systems Analysis*

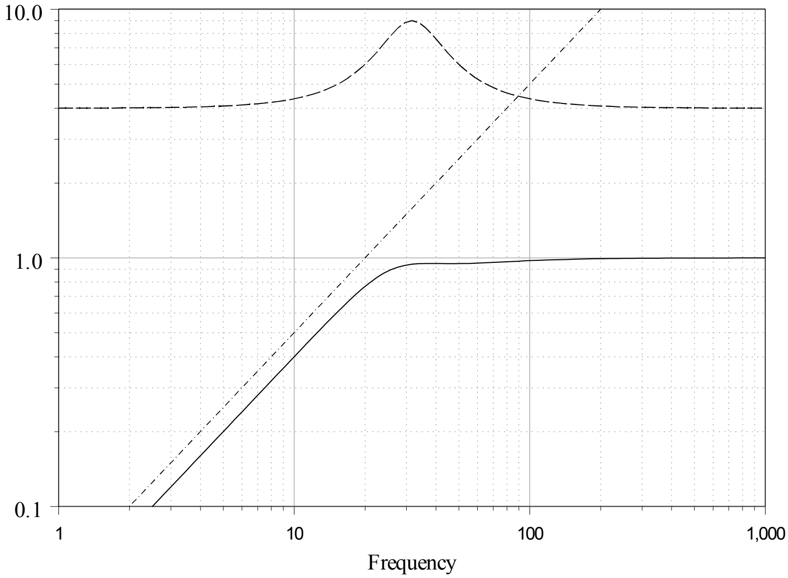


Figure 1-6 - High pass passive filter loaded by a complex impedance

since the components required for these higher order filters is impractical in a high power systems due to the large currents that the passive elements must withstand, preferably with very low loss. Active circuits become a better choice for filter sections above about two and sometimes even for a second order filter.

There is an extensive list of names by which specific filter shapes are known but it is not really that useful to learn these filters by name (although it is convenient to know them). Since any filter can be made up of a series combination of first and second order sections, a far better approach than learning filter names is to understand the pole-zero representation of a filter. In a pole-zero representation, all active filters are essentially the same, they simply have their poles and zeros in different places and sometimes there are a different number of them.

The simplest forms of active filters are the first order high-pass

$$T_h(\omega) = \frac{\omega}{\omega - i\omega_0} \tag{1.7.16}$$

and low-pass

$$T_l(\omega) = \frac{-i\omega_0}{\omega - i\omega_0}\,. \tag{1.7.17}$$

Both of these filters have a pole at ω_0, but the high-pass filter also has a zero at $\omega = 0$. In our study of active filters, we will look predominately at the denominators. The denominators for high and low pass sections of a given order are always the same, but the numerator are different. The form of the numerator will determine if it is high pass or low pass, or even bandpass.

Fig. 1-7 shows the transfer functions for both a high-pass and low-pass first order filter with their “cutoffs”, ω_0, at 30 Hz. The final slope in the stop band for a first order section is always $\pm$ 20 dB / decade or $\pm$ 6 dB / octave. The choice of octaves or decades for plotting is arbitrary, but the convention in this text will be for decades, because it is simpler to plot.

Second order filters have the form

$$T(\omega) = \frac{\left\langle \omega^2 \middle| \omega_0^2 \right\rangle}{\omega_0^2 - \omega^2 - i\dfrac{\omega\omega_0}{Q}} \tag{1.7.18}$$

where $\langle \omega^2 \,|\, \omega_0{}^2 \rangle$ indicates the numerator for high-pass and low-pass filters respectively. The two transfer functions are shown in Fig. 1-8. The stop bands have slopes of $\pm$ 40 dB / decade. In general, the stop band slope is always the order of the filter times 20 dB/decade.

The equation for a completely general filter function can be written as

$$T(\omega) = \frac{1 + \sum_{n=1} MA_n \omega^n}{\sum_m AR_m \omega^m} \tag{1.7.19}$$

MA_n = the Moving Average coefficients

AR_m = the Auto Regressive coefficients.

All active filters are a subset of this equation.

Since a polynomial must always have roots (they may be complex) Eq. (1.7.19) can also be written as

$$T(\omega) = \frac{\prod(\omega - \omega_n)}{\prod(\omega - \omega_m)} \tag{1.7.20}$$

ω_n = the roots of the numerator polynomial also known as the zeros

ω_m = the roots of the denominator equation also known as the poles

Most filters, any which ultimately have stop band slopes, can be made up of a cascade of second order filters, with a single first order filter required if the order is odd. For completely arbitrary filters, some of the sections will have zeros that are not at the origin.

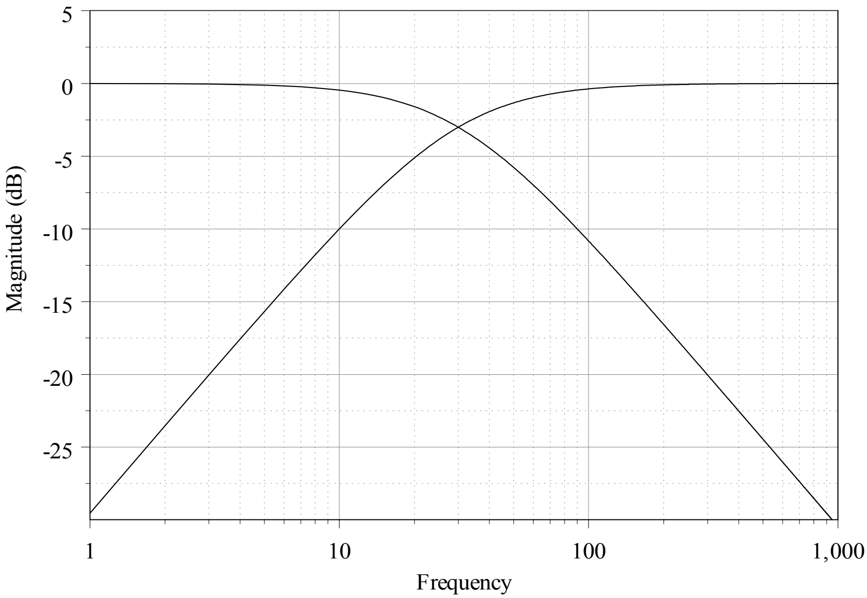


Figure 1-7 - High and low pass filters of first order

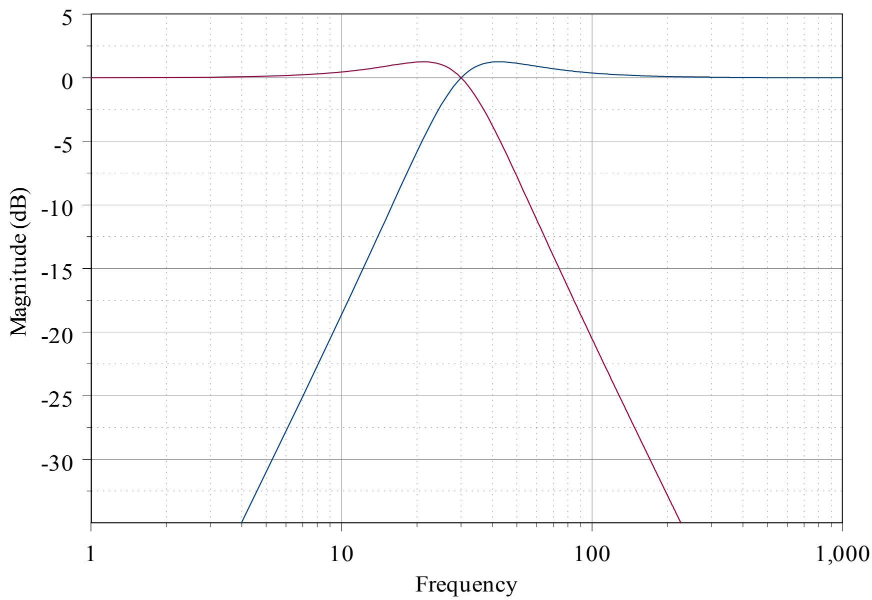


Figure 1-8 - Second order high and low pass filters with Q = 1

A LITTLE MECHANICS

1.8 Newtons Equations

Nearly all of the mechanics that we will encounter will rely on Newton's principle

$$m\,a = \sum Forces\ acting\ on\ the\ body \tag{1.8.21}$$

a = the acceleration

m = the mass.

This equation states that an object of mass *m* will accelerate based on all of the forces acting upon it at a given moment. The acceleration will be inversely proportional to the objects mass and directly proportional to the sum of the external forces. Of importance to note here is that internal forces do not play a role in the motion of the object, taken as a whole.

The forces that we will be concerned with will be of three basic types

1) The spring force: $F_c = -k\,x = \dfrac{-x}{C_m}$

k = the spring constant

$C_m = k^{-1}$ = *the mechanical compliance.*

This force tends to return the object back to the place of minimum x, usually $x = 0$.

2) The resistive force $F_R = -R_m\,v$

R_m = *the mechanical damping or resistive force,*

v = *the velocity.*

This force impedes all motion and is proportional to the velocity of the object. There are frictional or other loss type of forces which dissipate energy (notice that a spring cannot dissipate energy) which are not of the simple form shown in this. Viscous forces, for example, often depend on the square of the velocity and some internal frictional forces are sometimes constant.

3) The electro-magnetic force $F_{em} = Bl \cdot I$

Bl = the magnetic flux B in which a wire of length l is immersed

I = the current flow

This is the form of driving force found in most loudspeakers, although there are many types of driving forces. Several other varieties will be shown in Chap. 2.

In most cases, we will be interested in a simple form of mechanical system – a simple harmonic oscillator. While real transducers do operate in regions of complex mechanical motion, for the most part they only work well when the motion is simple harmonic.

1.9 The Simple Harmonic Oscillator[5]

Using the above equations, we can write the differential equation which governs a freely suspended body (like a loudspeaker cone) driven by a non-contacting force (like a voice coil). The object sees a spring acting to return the object to its rest position; a damping force, which we will assume acts directly on the body and is of the internal dissipation variety; and an external force that acts directly on the body. From Newton's principle, we have

$$ma = External\ Force - R_m v - \frac{x}{C_m}$$

which leads to a differential equation of the form

$$m\frac{d^2x(t)}{dt^2} + R_m\frac{dx(t)}{dt} + \frac{1}{C_m}x(t) = F(t) \tag{1.9.22}$$

$F(t)$ = *the external driving force.*

We could solve this equation using standard differential equation methods but it will be more instructive for us to solve it in another way.

Take a Fourier Transform of Eq.(1.9.22) and noting the operator / transform pair which is central to the advantageous characteristics of this transform

$$\frac{d}{dt} \Leftrightarrow -i\omega$$

(which states that the derivative in the time domain is analogous to a multiplication by $i\omega$ in the frequency domain) results in

$$\left(m\omega^2 - i\omega\cdot R_m - \frac{1}{C_m}\right)\cdot X(\omega) = F(\omega) \tag{1.9.23}$$

which can now be solved as a simple algebraic equation to get

$$X(\omega) = \frac{F(\omega)/m}{\omega_0^2 - \omega^2 + i\omega\cdot\frac{R_m}{m}} \tag{1.9.24}$$

Fig.1-9 shows a typical displacement response for this system. The displacement function is the same as a classic low pass filter of second order (see Fig.1-7).

---

5. see Mirovitch, *Analytical Methods in Vibrations*

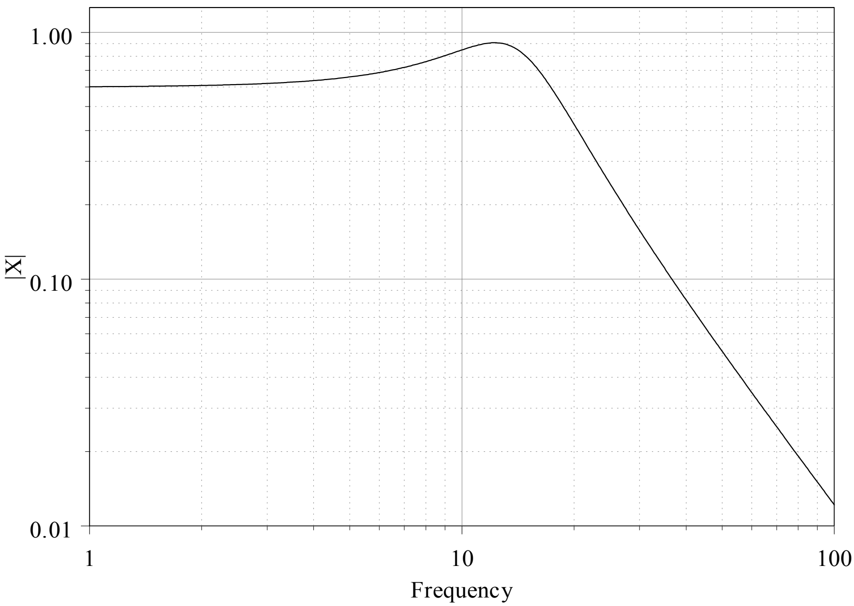


Figure 1-9 - Mechanical displacement example

We now want to look at the objects velocity instead of its displacement. Using $-i\omega$ for differentiation in the frequency domain, i.e. $V(\omega) = -i\omega X(\omega)$, Eq. (1.9.24) becomes

$$V(\omega) = \frac{F(\omega)/m}{-i\omega m + R_m + \frac{1}{-i\omega C_m}} \tag{1.9.25}$$

If we think of the velocity $V(\omega)$ as a current and the force $F(\omega)$ as a voltage then Eq. (1.9.25) is exactly analogous to that of a series electrical circuit with an impedance given by the denominator terms. In this analogy the mass has the analogous function of an inductor, the mechanical resistance that of an electrical resistor and the compliance that of a capacitor. The data in Table 1 then applies to these analogous mechanical components just as it does to the electrical ones. We will discuss these analogs further in Chap. 3.

1.10 Coupled Systems

Let us associate the force in the preceding section with the force of a voice coil, for example, such that

$$F(\omega) = Bl \cdot I(\omega) \tag{1.10.26}$$

Assume that we have a current that is independent of frequency or load – a constant current source. Applying the equations in the preceding section to this case, we would note that the mechanical object is uncoupled from the source – the two are not dependent upon one another. If, instead of a constant current source, we have the more common constant voltage source, then we must determine the current through the transducer because it is no longer constant. This current can be found by considering the voltage drops across the voice coil

$$E = I(\omega)z_e + E_{emf} = I(\omega)z_e + Bl\,V(\omega) \tag{1.10.27}$$

z_e = *the electrical impedance of the voice coil*

E_{emf} = $Bl\ V(\omega)$ *the back emf of the moving coil.*

Since this equation contains a term from the mechanical domain we must find an equation for the velocity in that domain accounting for the fact that the current is a variable. By summing the forces on the diaphragm and using Eq. (1.9.23), we get the coupled system

$$\begin{aligned} E(\omega) &= I(\omega)z_e - Bl\,V(\omega) \\ F(\omega) &= Bl\,I(\omega) - z_m V(\omega) \end{aligned} \tag{1.10.28}$$

z_m = *the mechanical impedance of the simple harmonic oscillator.*

A useful way of writing this equation is as a matrix

$$\begin{bmatrix} E(\omega) \\ F(\omega) \end{bmatrix} = \begin{bmatrix} z_e & -Bl \\ Bl & -z_m \end{bmatrix} \begin{bmatrix} I(\omega) \\ V(\omega) \end{bmatrix} \tag{1.10.29}$$

If there are no external forces on the object (it is in a vacuum), then $F(\omega)=0$. If the source is a constant voltage then the equations above simplify to

$$\begin{aligned} V(\omega) &= \frac{Bl\,I(\omega)}{z_m} = \frac{Bl}{z_m}\frac{(E - Bl\,V(\omega))}{z_e} \\ V(\omega) &= \frac{Bl\,E}{Bl^2 + z_m z_e} \end{aligned} \tag{1.10.30}$$

If the force is $-Bl\,E\,/\,R_e$, then this equation is identical with Eq. (1.9.25), except that the mechanical impedance in the denominator now contains an extra term $Bl\,/\,R_e$,. From the form of this term it appears as dampener of the motion – *electro-magnetic damping.* This new term is the result of the electro-magnetic coupling seen in the coupled equations above. In the uncoupled case of a current source, this extra damping does not occur.

It is desirable to put the electrical quantities on one side of the matrix in Eq. (1.10.29) and the mechanical terms on the other side. The resulting form

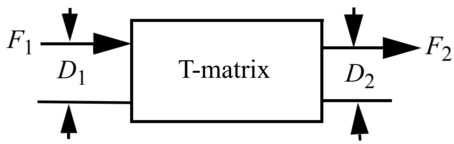
$$\begin{bmatrix} E(\omega) \\ I(\omega) \end{bmatrix} = \begin{bmatrix} \frac{z_e}{Bl} & \frac{z_m z_e}{Bl} - Bl \\ \frac{1}{Bl} & \frac{z_m}{Bl} \end{bmatrix} \begin{bmatrix} F(\omega) \\ V(\omega) \end{bmatrix} \tag{1.10.31}$$

is not nearly as succinct as Eq. (1.10.29). However, we will still find it to be useful in later chapters.

1.11 Basic T-Matrices

From the results of the previous section, we can see that matrices offer a convenient way of looking at a mechanical – in fact any – system. The entire field of Finite Element Analysis (FEA) is a concise but rigorous formulation for determining the matrix equations that best represent a complex system, given the numbers of degrees of freedom being allowed. For us, we shall see that a simple formulation will give us all of the degrees of freedom that we will require.

The T-matrix (for Transfer matrix) formulation is usually discussed as a two port system – an single input port and an single output port. It is drawn as



The subscript 1 refers to the input and subscript *2,* the output. The letter D refers to the "drop" quantity and the letter F refers to the "flow" quantity, the usual analogy being voltage and current respectively. The quantities can also be force and velocity or pressure and volume velocity. The black box system above can be written as

$$\begin{bmatrix} D_1 \\ F_1 \end{bmatrix} = \begin{bmatrix} T_{1,1} & T_{1,2} \\ T_{2,1} & T_{2,2} \end{bmatrix} \cdot \begin{bmatrix} D_2 \\ F_1 \end{bmatrix} \tag{1.11.32}$$

An example has already been shown in Eq. (1.10.31).

The rules for manipulating T-matrices are quite simple. A system is built up of elements as they appear in the physical system from the input through to the output. T-matrices, which represent each of the fundamental elements, are then substituted for the schematic elements. The matrices are then simply multiplied out to yield the equations for final system. For multiple inputs or multiple outputs, the concepts can be generalized to any number of degrees of freedom, as in FEA. For our purposes, most of what we will do is of the single input/single output variety. The open issue at this point is what are the T-matrices for typical elements found in a transducer? First, we need to discuss some fundamental rules for T-matrix elements.

1.12 Series and Parallel Elements in a T-matrix

We will need to know the forms for a T-matrix when elements are in either series or parallel. The formulations are quite simple. For a series connection, we have the fundamental relationship that

$$\begin{aligned} F_1 &= F_2 \\ D_1 &= z_{series} \cdot F_2 + D_2 \end{aligned} \qquad (1.12.33)$$

Written in matrix form these relationships become

$$\begin{pmatrix} D_1 \\ F_1 \end{pmatrix} = \begin{bmatrix} 1 & z_{series} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} D_2 \\ F_2 \end{pmatrix} \qquad (1.12.34)$$

where D and F can be any drop or flow variables.

For a parallel element we have the relationships that

$$\begin{aligned} D_1 &= D_2 \\ F_1 &= \frac{D_2}{z_{par}} + F_2 \end{aligned} \qquad (1.12.35)$$

from which we can immediately write the matrix form

$$\begin{pmatrix} D_1 \\ F_1 \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ (z_{par})^{-1} & 1 \end{bmatrix} \begin{pmatrix} D_2 \\ F_2 \end{pmatrix} \qquad (1.12.36)$$

These forms are most useful and we can derive nearly every case of interest from them.

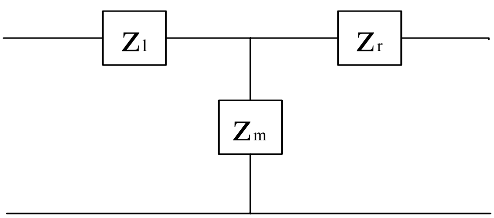


Figure 1-10 - Ladder network with three impedances

As an example consider the T-matrix of a T network ladder topology as shown in reference.

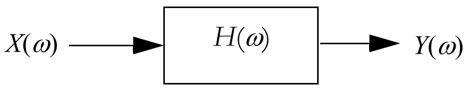
$$\begin{bmatrix} 1 & z_l \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ (z_m)^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 & z_r \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1+\dfrac{z_l}{z_m} & \left(1+\dfrac{z_l}{z_m}\right) z_r + z_l \\ \dfrac{1}{z_m} & 1+\dfrac{z_r}{z_m} \end{bmatrix} \qquad (1.12.37)$$

Nearly any level of complexity is possible in either the electrical or mechanical domains, but we will see the true value of the T-matrix when we get to the electro-acoustical problems.

A LITTLE SIGNAL PROCESSING

1.13 Correlation and Spectral Analysis[6]

A key concept in any filed which deals with systems is the use of transfer functions. Consider the system shown below



$X(\omega)$=the input signal at frequency $\omega=2\pi f$
$H(\omega)$=the complex multiplier of the input signal at ω
$Y(\omega)$= the output signal

The system is then defined by

$$Y(\omega) = X(\omega)H(\omega) \tag{1.13.38}$$

which leads to the definition of the transfer function

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} \tag{1.13.39}$$

This equation is valid for all ω. If we let the input signal contain all frequencies of interest, i.e. $X(\omega)=1.0$ (note that the phase is zero at all ω) then the transfer function is simply equal to the measured outputs complex spectrum.

Note here that we could define the transfer function to be a function of two frequencies ω_1 and ω_2 such that

$$Y(\omega_1,\omega_2) = X(\omega_1,\omega_2)H(\omega_1,\omega_2) \tag{1.13.40}$$

which is called the bispectrum and it quantifies the second order component of a nonlinear system. This concept can be generalized to any number of dimensions.

Taking a Fourier Transform of Eq. (1.13.38) we get

$$y(t) = x(t) \otimes h(t) \tag{1.13.41}$$

where $\otimes$ stands for convolution – a well know integration process. In this equation $h(t)$ is known as the systems impulse response. If the input, as above, was flat (i.e. 1.0) then $x(t)=\delta(0)$ (the Fourier Transform of a flat spectrum) and $h(t)=y(t)$. The output $y(t)$ is the response to an impulse – the impulse response. This means that the transfer function (frequency response) is the Fourier Transform of the impulse response and visa-versa.

6. See Bendat, *Engineering Applications of Correlation and Spectral Analysis.*

If the input function is a random signal, then all of the above equation hold except that we need to modify their interpretation. Since a random signal never repeats we will never get the same input and output twice. This means that we should expect different results in Eq. (1.13.39). The way around this is to define the expected value operator E[ ]

$$E[x(t)] = \int_{-\infty}^{\infty} x(\tau)\, p(\tau)\, d\tau \tag{1.13.42}$$

$p(\tau)$=the probability density function of the signal

This equation will yield a weighted average of values, weighted by the likeliness of occurrence. For our purposes $p(t)$ will always be Gaussian.
Several useful cases occur

$$E[x(t)] = \int_{-\infty}^{\infty} x(\tau)\, p(\tau)\, d\tau = \mu \tag{1.13.43}$$

u=the mean value of $x(t)$

$$E[x^2(t)] = \int_{-\infty}^{\infty} x^2(\tau)\, p(\tau)\, d\tau = \psi^2 \tag{1.13.44}$$

ψ^2=mean square value of $x(t)$

The square root of ψ is called the *rms* value. Higher order moments can be defined but are of far less useful.

If we define the *cross-correlation* function $R_{xy}(t)$

$$R_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} x(t)\, y(t+\tau)\, dt \tag{1.13.45}$$

$R_{xx}(t)$ is called the *autocorrelation* function.

Taking the Fourier Transform of $R_{xy}(t)$ yields the *cross-spectral density* function $S_{xy}(\omega)$ and the Fourier Transform of $R_{xx}(t)$ is called the *Power Spectral Density* function or PSD, $S_{xx}(w)$.

By using all of the above results we can find that

$$H(\omega) = \frac{S_{xy}(\omega)}{S_{xx}(\omega)} \tag{1.13.46}$$

which is a useful result, especially if $S_{xx}(\omega)$=1.0 – white noise. There is one distinct difference in the definitions in Eq. (1.13.39) and Eq. (1.13.46). The later equation considers only those input and output signals which are correlated, while the former equation does not. For uncontaminated measurements, the two will be identical, but for contaminated measurements only Eq. (1.13.46) will give the correct results. Contamination can occur as noise, reflections, or most importantly, nonlinearity.

It can be shown that the optimum system as shown by Eq. (1.13.46) is not even the optimum linear system for a nonlinear system. The two transfer functions are different.[7] This is an important result when making measurements of nonlinear systems. The odd order nonlinearities are correlated with the linear system and hence contaminate the measurement. It is often thought that the cross spectrum approach to measuring a system is immune to nonlinearity, but it is not. This measurement ignores even orders but not the odd orders – they are correlated with the input signal.

1.14 Modeling Transfer Functions

One of the common tools that we will use in later chapters is that of modeling a transfer function. There are numerous ways to do this and the whole subject would take a text to describe.[8] We will describe a technique that we have found useful which highlights the general approach and leave the alternative techniques to further reading.

Consider a system with a discretely sampled input $x[n]$ and a discreetly sampled output $y[n]$. The z transform $\mathbf{Z}\{\ \}$ is defined by

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} h[n]z^{-n} = \mathbf{Z}\left\{h[n]\right\} \tag{1.14.47}$$

In the z domain (the discrete equivalent of the frequency domain) we can represent the transfer function $\mathbf{H}(z)$ as

$$\mathbf{H}(z) = \frac{\mathbf{B}(z)}{\mathbf{A}(z)} \tag{1.14.48}$$

where $\mathbf{B}(z)$ and $\mathbf{A}(z)$ are the polynomials that we would like to determine. Note from Eq. (1.7.19) that $\mathbf{B}(z)$ are the MA coefficients and $\mathbf{A}(z)$ are the AR coefficients. The roots of $\mathbf{B}(z)$ are therefore the zeros of the model and the roots of $\mathbf{A}(z)$ the poles.

If we now take the inverse z transform of Eq. (1.14.48) we can write out a set of equations

---

7. See Bendat, *Nonlinear Systems Techniques and Applications*

8. See Marple, *Digital Spectral Analysis with Applications*

$$h[n] = \begin{cases} -\sum_{p=1}^{P} a[p]h[n-p] + b[n] & \text{for } 1 \le n \le Q \\ -\sum_{p=1}^{P} a[p]h[n-p] & \text{for } n > Q \\ 1 & \text{for } n = 0 \end{cases} \tag{1.14.49}$$

$a[n]$=*the AR coefficients*
$b[n]$=*the MA coefficients*

where we have assumed a value of $a[0]=1$ (without loss of generality since it only scales the results) and that $h[n]=0$ for $n < 0$ (from causality).

The equation for $n > Q$ can be solved directly for the a[n] coefficients from

$$\begin{pmatrix} h[Q+1] \\ h[Q+2] \\ \vdots \\ h[Q+P] \end{pmatrix} = -\begin{pmatrix} h[Q-1] & h[Q-2] & \cdots & h[Q-P] \\ h[Q] & h[Q-1] & \cdots & h[Q-P+1] \\ \vdots & \vdots & \ddots & \vdots \\ h[Q+P-2] & h[Q+P-3] & \cdots & h[Q-1] \end{pmatrix} \begin{pmatrix} a[1] \\ a[2] \\ \vdots \\ a[P] \end{pmatrix} \tag{1.14.50}$$

Thus for a given $Q \ge P$ we can find the $a[n]$ coefficients directly from the impulse response sequence by using any number of matrix techniques (although we have found that Singular Value Decomposition SVD avoids a lot of pitfalls[9]). While the optimum Q is not know a apriori, it is really not a problem to find several values of Q which work well. The results are not very sensitive to the value chosen so long as it is not too large.

Returning now to Eq. (1.14.49) we have for $n \le Q$

$$b[n] = h[n] + \sum_{p=1}^{P} a[p]h[n-p] \quad n = 1 \ldots Q \tag{1.14.51}$$

which can be solved directly for the $b[n]$'s.

We thus have procedure for finding a model of any transfer function in a minimum number of terms

- Find $H(\omega)$ via FFT techniques selecting a frequency range and number of points which matches the desire resolution of the model. Too many points will cause problem in fitting the parameters, less data is the goal so use as few points in the FFT as permissible.
- Find the impulse response sequence by taking the inverse FFT of $H(\omega)$ (which is really $\mathbf{H}(z)$). The impulse sequence are simply the data

9. See Press, *Numerical Recipes*

values. Smoothing $H(\omega)$ at the end points to force a real, causal impulse response may be required.

- Find the $a[n]$'s from Eq. (1.14.50).
- Find the $b[n]$'s from Eq. (1.14.51).
- Further data reduction can be achieved by finding the roots of $\mathbf{A}(z)$ (the poles) and $\mathbf{B}(z)$ (the zeros). Very close pole-zero pairs imply that there were too many coefficients in the expansion and these pairs can simple be discarded.

With the above techniques we can simplify any transfer function to any order (amount of data) desired. This data reduction will be used extensively in later chapters.

1.15 Summary

This chapter has introduced some of the techniques that we will use in the following chapters. It is not, by far, a complete discussion of these techniques, but it does cover the general topics and gives references to where someone could go to find out more information.