Fundamentals of Signal Analysis Series

Introduction to Time, Frequency and Modal Domains

Application Note 1405-1

Agilent Technologies

Table of Contents

Introduction

The analysis of electrical signals is a fundamental concern for many engineers and scientists. Even if the immediate problem is not electrical, the basic parameters of interest are often changed into electrical signals by means of transducers. Common transducers include accelerometers and load cells in mechanical work, EEG electrodes and blood pressure probes in biology and medicine, and pH and conductivity probes in chemistry. The rewards for transforming physical parameters to electrical signals are great, as many instruments are available for the analysis of electrical signals. The powerful measurement and analysis capabilities of these instruments can lead to rapid understanding of the system under study.

You can look at electrical signals from several different perspectives, and each of these different ways of looking at a problem often lends its own unique insights.

In this application note we introduce the concepts of the time, frequency and modal domains. These three ways of looking at a problem are interchangeable; that is, no information is lost in changing from one domain to another. By changing perspective, the solution to difficult problems can often become quite clear.

After developing the concepts of each domain, we will introduce the types of instrumentation available. The merits of each generic instrument type are discussed to give you an appreciation of the advantages and disadvantages of each approach.

Section 1: The Time Domain

The traditional way of observing signals is to view them in the time domain. The time domain is a record of what happens to a parameter of the system versus time. For instance, Figure 1.1 shows a simple spring-mass system where we have attached a pen to the mass and pulled a piece of paper past the pen at a constant rate. The resulting graph is a record of the displacement of the mass versus time, *a time-domain view of displacement.*

Such direct recording schemes are sometimes used, but usually it is much more practical to convert the parameter of interest to an electrical signal using a transducer. Transducers are commonly available to change a wide variety of parameters to electrical signals. Microphones, accelerometers, load cells, conductivity and pressure probes are just a few examples.

This electrical signal, which represents a parameter of the system, can be recorded on a strip chart recorder as in Figure 1.2. We can adjust the gain of the system to calibrate our measurement. Then we can reproduce exactly the results of our simple direct recording system in Figure 1.1.

Why should we use this indirect approach? One reason is that we are not always measuring displacement. We then must convert the desired parameter to the displacement of the recorder pen. Usually, the easiest way to do this is through the intermediary of electronics. However, even when measuring displacement, we would normally use an indirect approach. Why? Primarily because the system in Figure 1.1 is hopelessly ideal. The mass must be large enough and the spring stiff enough so that the pen's mass and drag on the paper will not affect the results appreciably. Also the deflection of the mass must be large enough to give a usable result, otherwise a mechanical lever system to amplify the motion would have to be added with its attendant mass and friction.

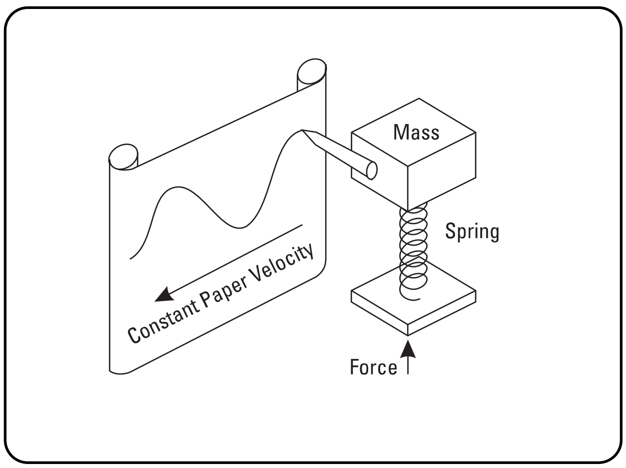

Figure 1.1. Direct recording of displacement - a time domain view

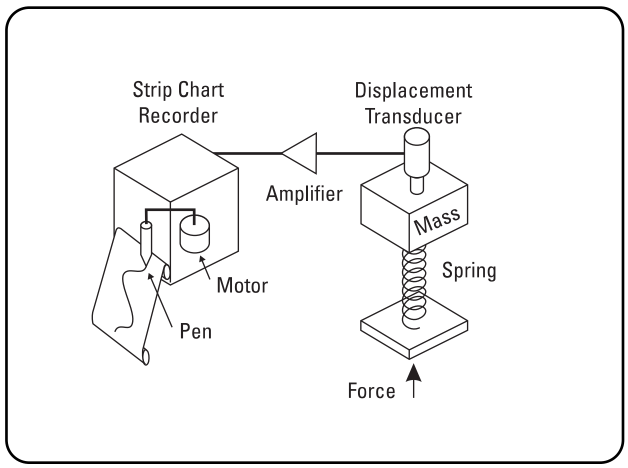

Figure 1.2. Indirect recording of displacement

With the indirect system, you can usually select a transducer that will not significantly affect the measurement. (This can go to the extreme of commercially available displacement transducers that do not even contact the mass.) You can easily set the pen deflection to any desired value by controlling the gain of the electronic amplifiers.

This indirect system works well until our measured parameter begins to change rapidly. Because of the mass of the pen and recorder mechanism and the power limitations of its drive, the pen can move only at finite velocity. If the measured parameter changes faster than the pen velocity, the output of the recorder will be in error. A common way to reduce this problem is to eliminate the pen and use a deflected light beam to record on photosensitive paper. Such a device is called an *oscillograph* (see Figure 1.3). Since it is only necessary to move a small, lightweight mirror through a very small angle, the oscillograph can respond much faster than a strip chart recorder.

Another common device for displaying signals in the time domain is the *oscilloscope* (see Figure 1.4). Here, an electron beam is moved using electric fields. The electron beam is made visible by a screen of phosphorescent material. An oscilloscope is capable of accurately displaying signals that vary even more rapidly than an oscillograph can handle. This is because it is only necessary to move an electron beam, not a mirror.

The strip chart, oscillograph and oscilloscope all show displacement versus time. We say that changes in this displacement represent the variation of some parameter versus time. We will now look at another way of representing the variation of a parameter.

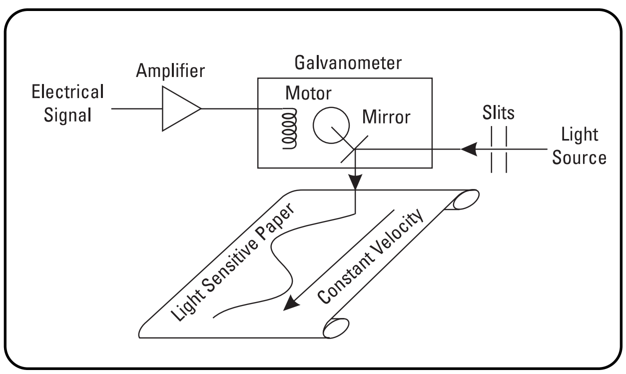

Figure 1.3. Simplified oscillograph operation

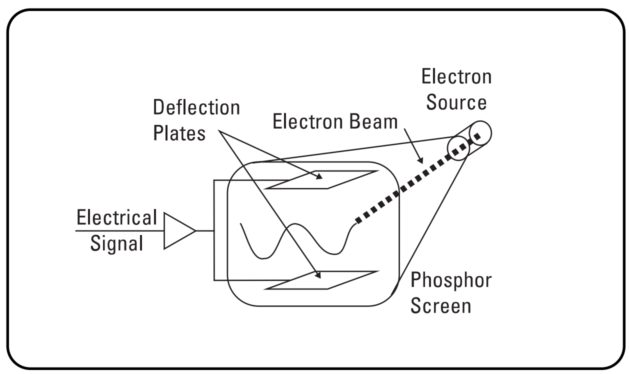

Figure 1.4. Simplified oscilloscope operation (Horizontal deflection circuits omitted for clarity)

Section 2: The Frequency Domain

Over one hundred years ago, Baron Jean Baptiste Fourier showed that any waveform that exists in the real world can be generated by adding up sine waves. We have illustrated this in Figure 2.1 for a simple waveform composed of two sine waves. By picking the amplitudes, frequencies and phases of these sine waves correctly, we can generate a waveform identical to our desired signal.

Conversely, we can break down our real world signal into these same sine waves. It can be shown that this combination of sine waves is unique; any real world signal can be represented by only one combination of sine waves.

Figure 2.2a is a three-dimensional graph of this addition of sine waves. Two of the axes are time and amplitude, familiar from the time domain. The third axis, frequency, allows us to visually separate the sine waves that add to give us our complex waveform. If we view this three-dimensional graph along the frequency axis we get the view in Figure 2.2b. This is the time-domain view of the sine waves. Adding them together at each instant of time gives the original waveform.

However, if we view our graph along the time axis as in Figure 2.2c, we get a totally different picture. Here we have axes of amplitude versus frequency, what is commonly called the frequency domain. Every sine wave we separated from the input appears as a vertical line. Its height represents its amplitude and its position represents its frequency. Since we know that each line represents a sine wave, we have uniquely characterized our input signal in the frequency domain*. This frequency domain representation of our signal is called the *spectrum* of the signal. Each sine wave line of the spectrum is called a *component* of the total signal.

It is very important to understand that *we have neither gained nor lost information, we are just representing it differently*. We are looking at the same three-dimensional graph from different angles. This different perspective can be very useful.

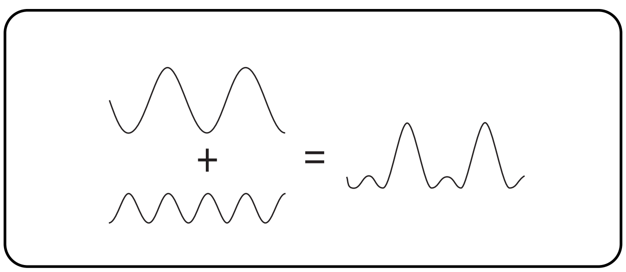

Figure 2.1. Any real waveform can be produced by adding sine waves together.

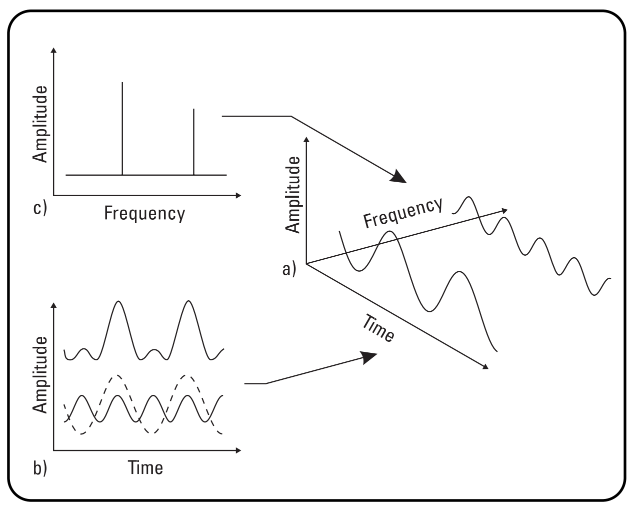

Figure 2.2. The relationship between the time and frequency domains
a) Three- dimensional coordinates showing time, frequency and amplitude
b) Time-domain view
c) Frequency-domain view

* Actually, we have lost the phase information of the sine waves. Agilent Application Note 1405-2 explains how we get this information.

The Need for Decibels

Since one of the major uses of the frequency domain is to resolve small signals in the presence of large ones, let us now address the problem of how we can see both large and small signals on our display simultaneously.

Suppose we wish to measure a distortion component that is 0.1% of the signal. If we set the fundamental to full scale on a four-inch (10 cm) screen, the harmonic would be only four thousandths of an inch (0.1 mm) tall. Obviously, we could barely see such a signal, much less measure it accurately. Yet many analyzers are available with the ability to measure signals even smaller than this.

Since we want to be able to see all the components easily at the same time, the only answer is to change our amplitude scale. A logarithmic scale would compress our large signal amplitude and expand the small ones, allowing all components to be displayed at the same time.

db	Power Ratio	db	Voltage Ratio
+20	100	+40	100
+10	10	+20	10
+ 3	2	+ 6	2
0	1	0	1
- 3	1/2	- 6	1/2
- 10	1/10	- 20	1/10
- 20	1/100	- 40	1/100

db = 10 log (Power Ratio) = 20 log (Voltage Ratio)

Figure 2.3. The relationship between decibels, power and voltage

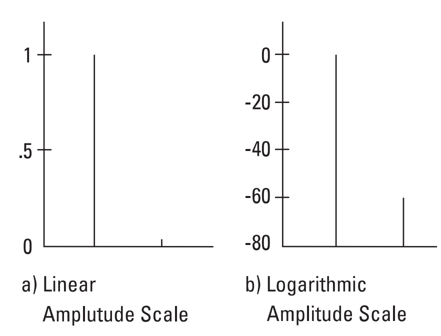

Figure 2.4. Small signals can be measured with a logarithmic amplitude scale

Alexander Graham Bell discovered that the human ear responded logarithmically to power difference and invented a unit, the Bel, to help him measure the ability of people to hear. One tenth of a Bel, the deciBel (dB) is the most common unit used in the frequency domain today. A table of the relationship between volts, power and dB is given in Figure 2.3. From the table we can see that our 0.1% distortion component example is 60 dB below the fundamental. If we had an 80 dB display as in Figure 2.4, the distortion component would occupy 1/4 of the screen, not 1/1000 as in a linear display.

Why the Frequency Domain?

Suppose we wish to measure the level of distortion in an audio oscillator. Or we might be trying to detect the first sounds of a bearing failing on a noisy machine. In each case, we are trying to detect a small sine wave in the presence of large signals. Figure 2.5a shows a time domain waveform that seems to be a single sine wave. But Figure 2.5b shows in the frequency domain that the same signal is composed of a large sine wave and significant other sine wave components (distortion components). When these components are separated in the frequency domain, the small components are easy to see because they are not masked by larger ones.

The frequency domain's usefulness is not restricted to electronics or mechanics. All fields of science and engineering have measurements like these where large signals mask others in the time domain. The frequency domain provides a useful tool for analyzing these small, but important, effects.

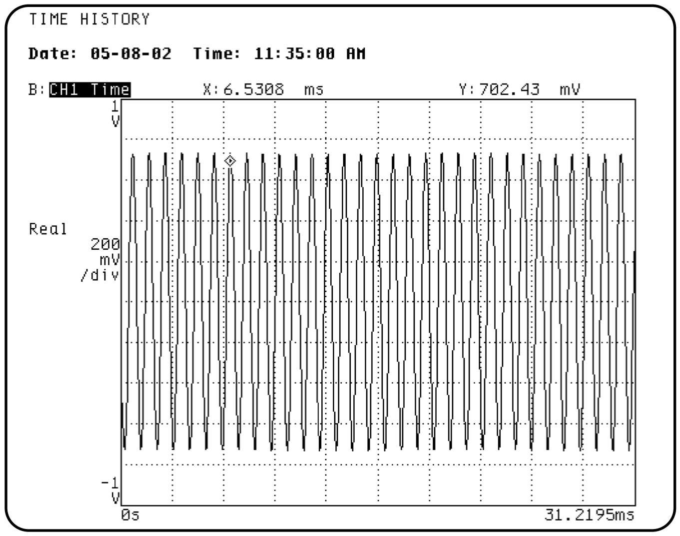

Figure 2.5.a Time Domain — small signal not visible

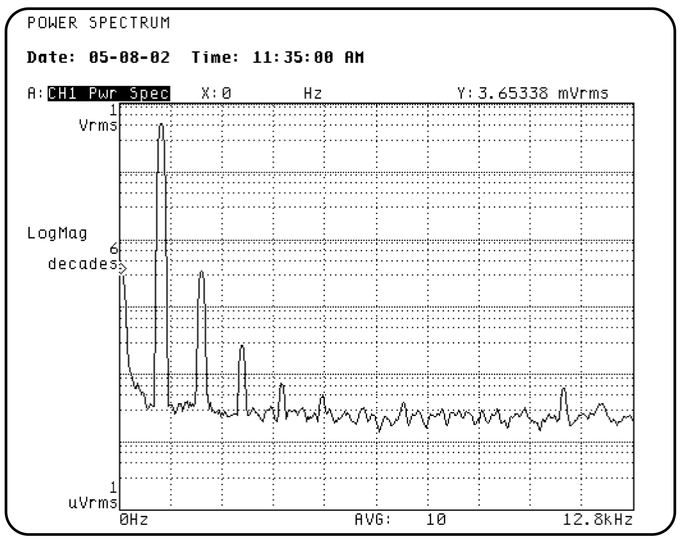

Figure 2.5.b Frequency Domain — small signal easily resolved

The Frequency Domain: A Natural Domain

At first the frequency domain may seem strange and unfamiliar, yet it is an important part of everyday life. Your ear-brain combination is an excellent frequency domain analyzer. The ear-brain splits the audio spectrum into many narrow bands and determines the power present in each band. It can easily pick small sounds out of loud background noise thanks in part to its frequency domain capability. A doctor listens to your heart and breathing for any unusual sounds. He is listening for frequencies that will tell him something is wrong. An experienced mechanic can do the same thing with a machine. Using a screwdriver as a stethoscope, he can hear when a bearing is failing because of the frequencies it produces.

So we see that the frequency domain is not at all uncommon. We are just not used to seeing it in graphical form. But this graphical presentation is really not any stranger than saying that the temperature changed with time, like the displacement of a line on a graph.

Spectrum Examples

Let us now look at a few common signals in both the time and frequency domains. In Figure 2.6a, we see that the spectrum of a sine wave is just a single line. We expect this from the way we constructed the frequency domain. The square wave in Figure 2.6b is made up of an infinite number of sine waves, all harmonically related. The lowest frequency present is the reciprocal of the square wave period. These two examples illustrate a property of the frequency transform: a signal that is periodic and exists for all time has a discrete frequency spectrum. This is in contrast to the transient signal in Figure 2.6c which has a continuous spectrum. This means that the sine waves that make up this signal are spaced infinitesimally close together.

Another signal of interest is the impulse shown in Figure 2.6d. The frequency spectrum of an impulse is flat, i.e., there is energy at all frequencies. It would, therefore, require infinite energy to generate a true impulse. Nevertheless, it is possible to generate an approximation to an impulse that has a fairly flat spectrum over the desired frequency range of interest. We will find signals with a flat spectrum useful in our next subject, network analysis.

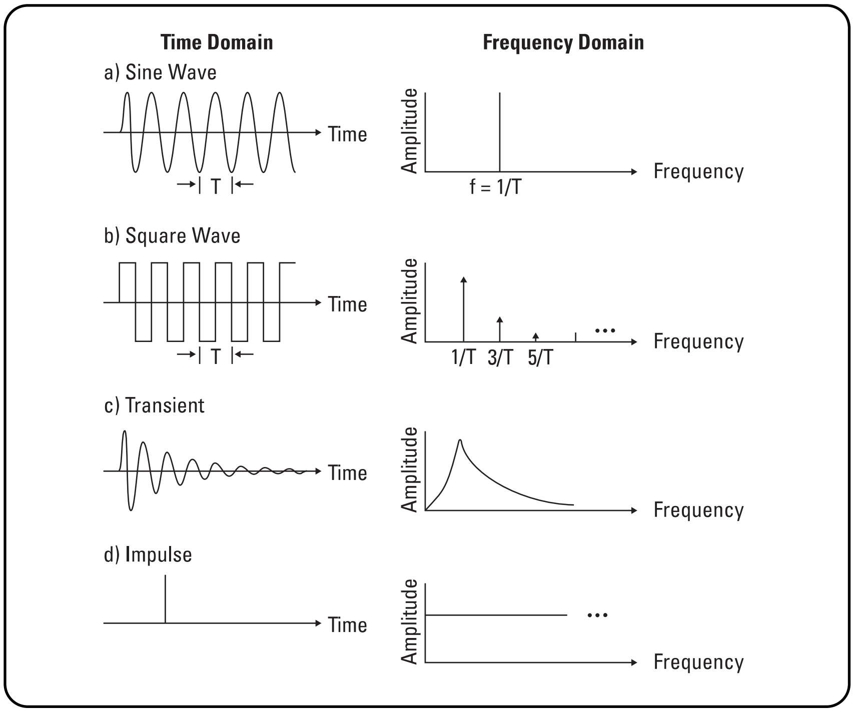

Figure 2.6. Frequency spectrum examples

Network Analysis

If the frequency domain were restricted to the analysis of signal spectrums, it would certainly not be such a common engineering tool. However, the frequency domain is also widely used in analyzing the behavior of networks (network analysis) and in design work.

Network analysis is the general engineering problem of determining how a network will respond to an input.* For instance, we might wish to determine how a structure will behave in high winds. Or we might want to know how effective a sound-absorbing wall we are planning to purchase would be in reducing machinery noise. Or perhaps we are interested in the effects of a tube of saline solution on the transmission of blood pressure waveforms from an artery to a monitor.

All of these problems and many more are examples of network analysis. As you can see a "network" can be any system at all. *One-port network* analysis is the variation of one parameter with respect to another, both measured at the same point (port) of the network. The impedance or compliance of the electronic or mechanical networks shown in Figure 2.7 are typical examples of one-port network analysis.

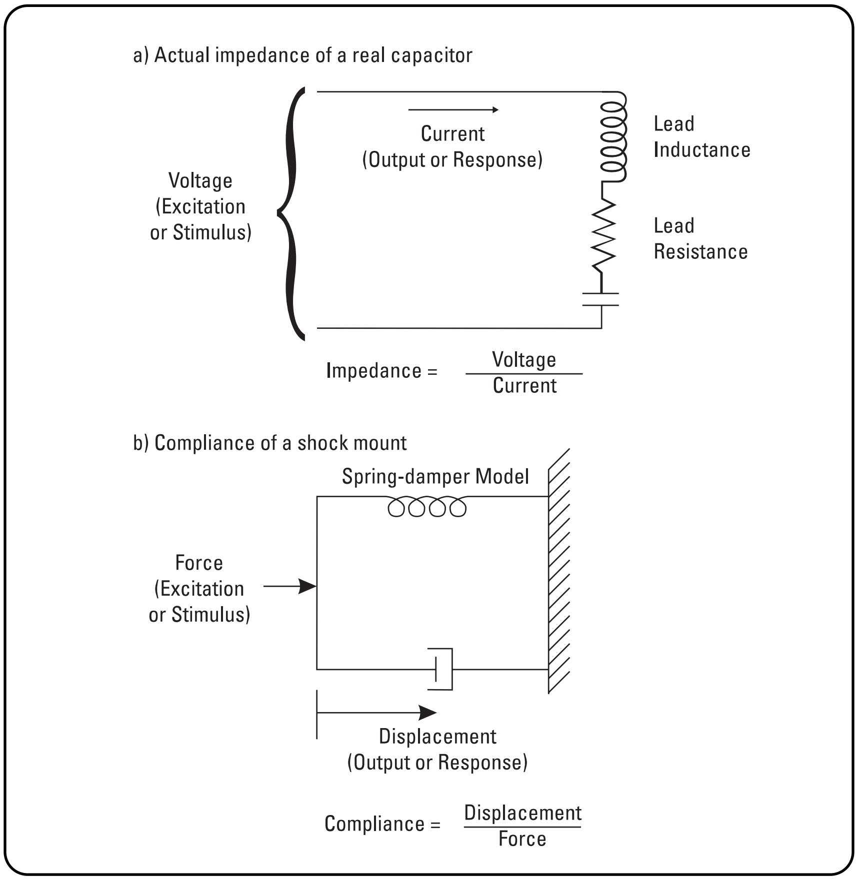

Figure 2.7. One-port network analysis examples

Two-port analysis gives the response at a second port due to an input at the first port. We are generally interested in the transmission and rejection of signals and in insuring the integrity of signal transmission. The concept of two-port analysis can be extended to any number of inputs and outputs. This is called *N-port analysis*, a subject we will use in modal analysis later in this application note.

We have deliberately defined network analysis in a very general way. It applies to all networks with no limitations. If we place one condition on our network, *linearity*, we find that network analysis becomes a very powerful tool.

* Network Analysis is sometimes called Stimulus/Response Testing. The input is then known as the stimulus or excitation and the output is called the response.

When we say a network is *linear*, we mean it behaves like the network in Figure 2.9. Suppose one input causes an output A and a second input applied at the same port causes an output B. If we apply both inputs at the same time to a linear network, the output will be the sum of the individual outputs, A + B.

At first glance it might seem that all networks would behave in this fashion. A counter example, a *non-linear* network, is shown in Figure 2.10. Suppose that the first input is a force that varies in a sinusoidal manner. We pick its amplitude to ensure that the displacement is small enough so that the oscillating mass does not quite hit the stops. If we add a second identical input, the mass would now hit the stops. Instead of a sine wave with twice the amplitude, the output is clipped as shown in Figure 2.10b.

This spring-mass system with stops illustrates an important principal: *no real system is completely linear*. A system may be approximately linear over a wide range of signals, but eventually the assumption of linearity breaks down. Our spring-mass system is linear before it hits the stops. Likewise a linear electronic amplifier clips when the output voltage approaches the internal supply voltage. A spring may compress linearly until the coils start pressing against each other.

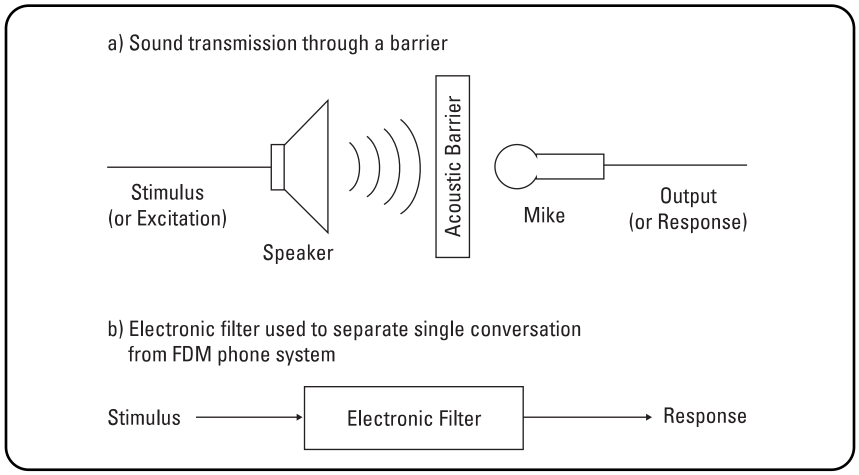

Figure 2.8. Two-port network analysis

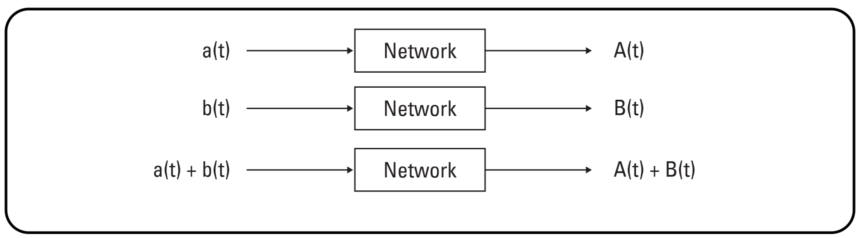

Figure 2.9. Linear network

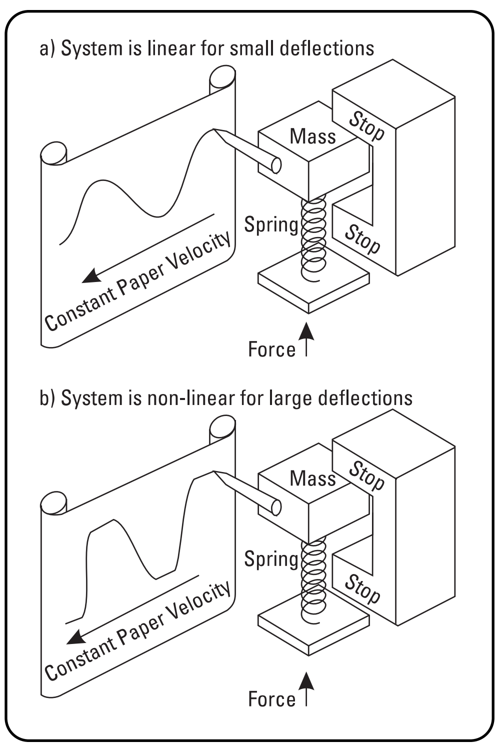

Figure 2.10. Non-linear system example

Other forms of non-linearities are also often present. Hysteresis (or backlash) is usually present in gear trains, loosely riveted joints and in magnetic devices. Sometimes the non-linearities are less abrupt and are smooth, but nonlinear, curves. The torque versus rpm of an engine or the operating curves of a transistor are two examples that can be considered linear over only small portions of their operating regions.

The important point is not that all systems are nonlinear; it is that *most systems can be approximated as linear systems*. Often a large engineering effort is spent in making the system as linear as practical. This is done for two reasons. First, it is often a design goal for the output of a network to be a scaled, linear version of the input. A strip chart recorder is a good example. The electronic amplifier and pen motor must both be designed to ensure that the deflection across the paper is linear with the applied voltage.

The second reason why systems are linearized is to reduce the problem of nonlinear instability. One example would be the positioning system shown in Figure 2.12. The actual position is compared to the desired position and the error is integrated and applied to the motor. If the gear train has no backlash, it is a straightforward problem to design this system to the desired specifications of positioning accuracy and response time.

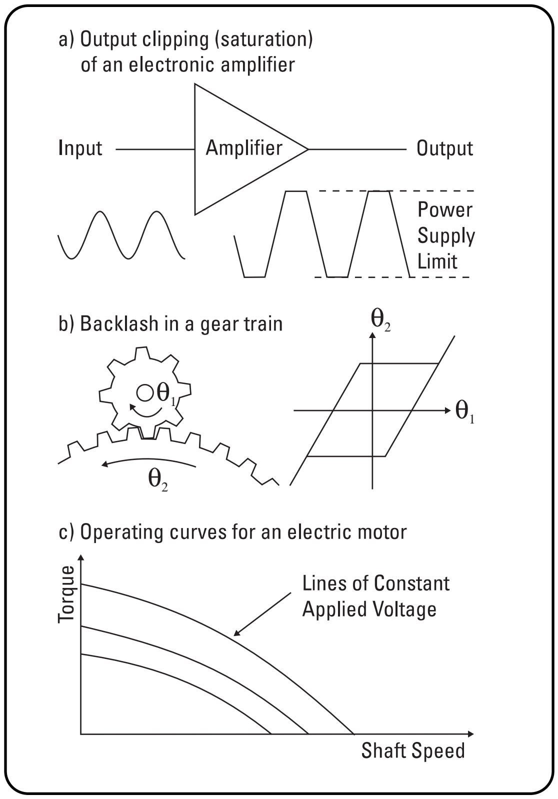

Figure 2.11. Examples of non-linearities

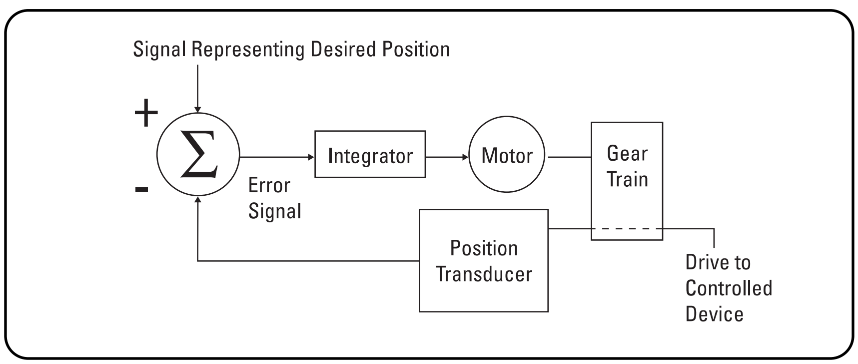

Figure 2.12. A positioning system

However, if the gear train has excessive backlash, the motor will “hunt,” causing the positioning system to oscillate around the desired position. The solution is either to reduce the loop gain and therefore reduce the overall performance of the system, or to reduce the backlash in the gear train. Often, reducing the backlash is the only way to meet the performance specifications.

Analysis of Linear Networks

As we have seen, many systems are designed to be reasonably linear to meet design specifications. This has a fortuitous side benefit when attempting to analyze networks*.

Recall that a real signal can be considered to be a sum of sine waves. Also, recall that the response of a linear network is the sum of the responses to each component of the input. Therefore, if we knew the response of the network to each of the sine wave components of the input spectrum, we could predict the output.

It is easy to show that the steady-state response of a linear network to a sine wave input is a sine wave of the same frequency. As shown in Figure 2.13, the amplitude of the output sine wave is proportional to the input amplitude. Its phase is shifted by an amount that depends only on the frequency of the sine wave. As we vary the frequency of the sine wave input, the amplitude proportionality factor (gain) changes, as does the phase of the output. *If we divide the output of the network by the input, we get a normalized result called the frequency response of the network*. As shown in Figure 2.14, the frequency response is the gain (or loss) and phase shift of the network as a function of frequency. Because the network is linear, the frequency response is independent of the input amplitude; *the frequency response is a property of a linear network*, not dependent on the stimulus.

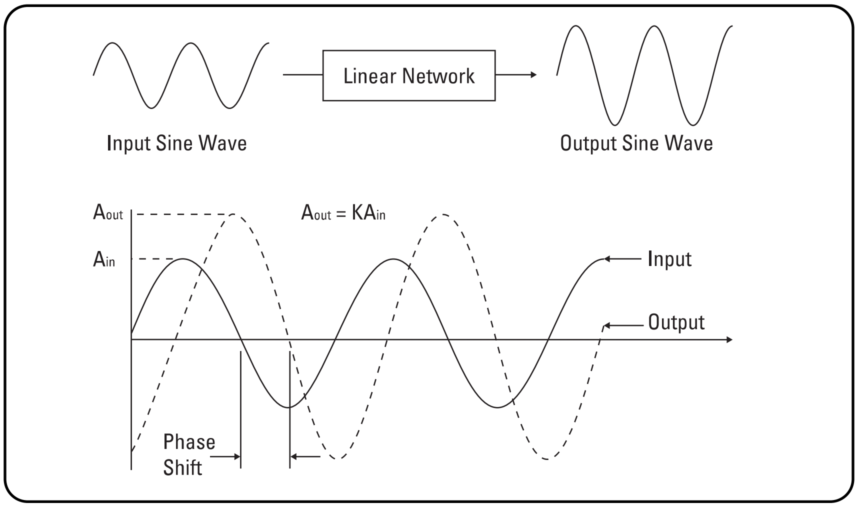

Figure 2.13. Linear network response to a sine wave input.

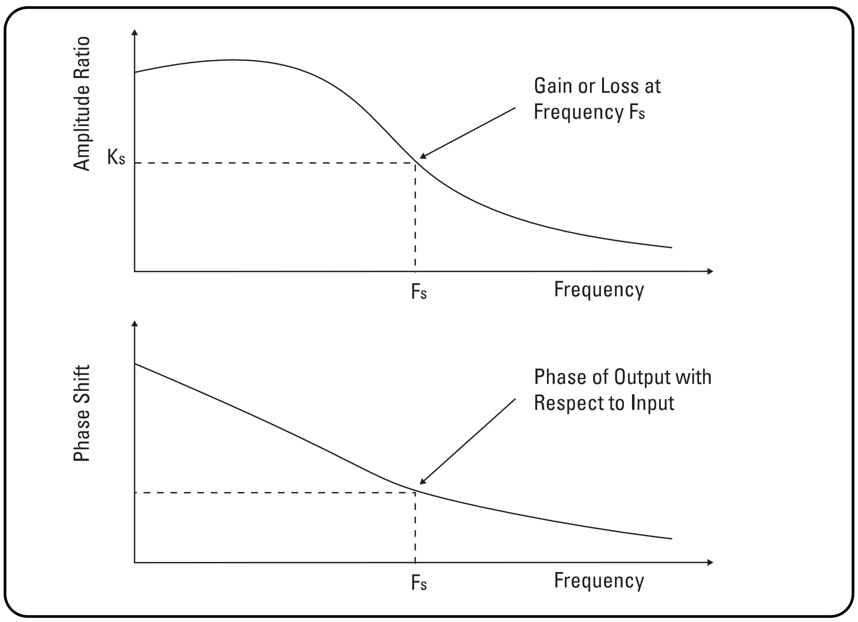

Figure 2.14. The frequency response of a network

* For a discussion of the analysis of networks that have not been linearized, see Agilent Application Note 1405-2.

The frequency response of a network will generally fall into one of three categories; low pass, high pass, bandpass or a combination of these. As the names suggest, their frequency responses have relatively high gain in a band of frequencies, allowing these frequencies to pass through the network. Other frequencies suffer a relatively high loss and are rejected by the network. To see what this means in terms of the response of a filter to an input, let us look at the bandpass filter case.

In Figure 2.16, we put a square wave into a bandpass filter. We recall from Figure 2.6 that a square wave is composed of harmonically related sine waves. The frequency response of our example network is shown in Figure 2.16b. Because the filter is narrow, it will pass only one component of the square wave. Therefore, the steady-state response of this bandpass filter is a sine wave.

Notice how easy it is to predict the output of any network from its frequency response. The spectrum of the input signal is multiplied by the frequency response of the network to determine the components that appear in the output spectrum. This frequency domain output can then be transformed back to the time domain.

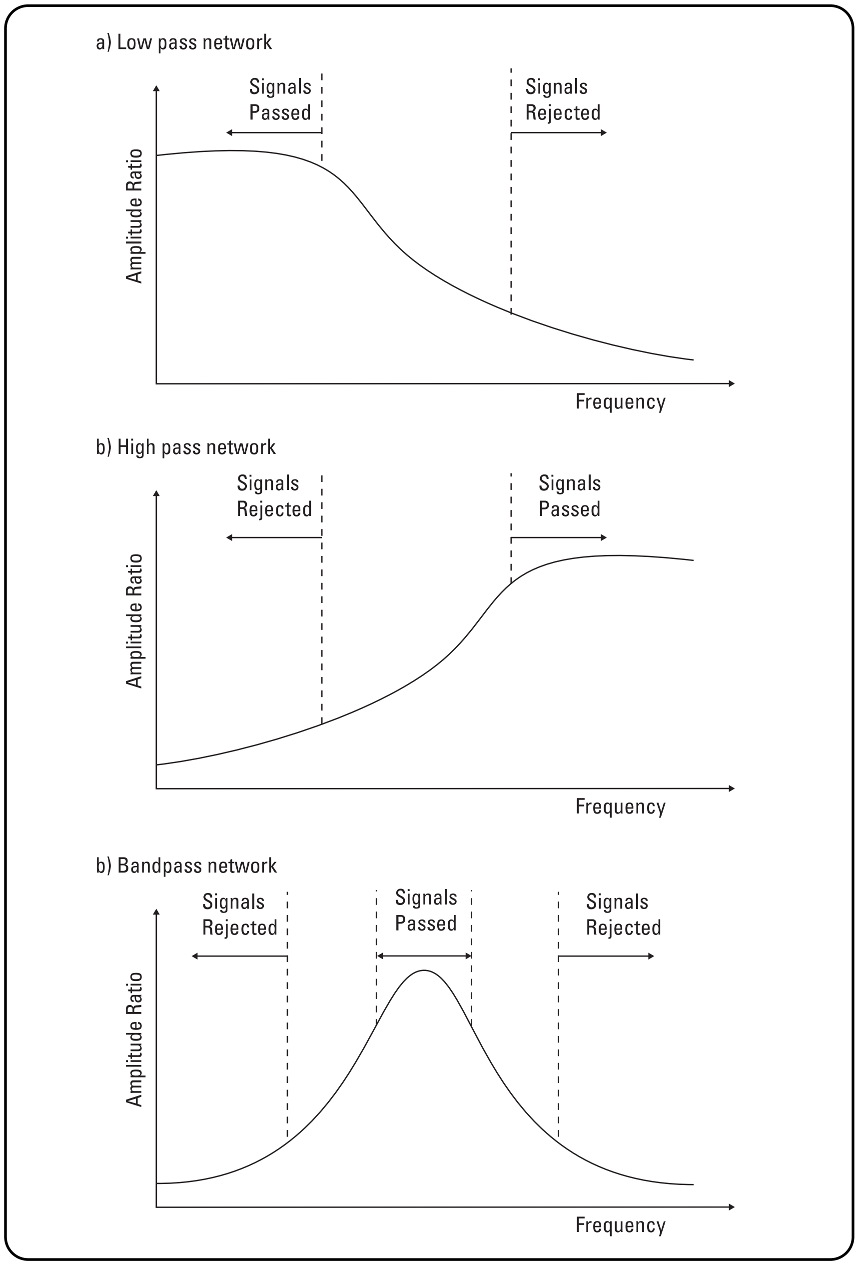

Figure 2.15. Three classes of frequency response

In contrast, it is very difficult to compute in the time domain the output of any but the simplest networks. A complicated integral must be evaluated, which often can be done only numerically on a computer*. If we computed the network response by both evaluating the time domain integral and by transforming to the frequency domain and back, we would get the same results. However, it is usually easier to compute the output by transforming to the frequency domain.

* This operation is called convolution.

Transient Response

Up to this point we have only discussed the steady-state response to a signal. By steady-state we mean the output after any transient responses caused by applying the input have died out. However, the frequency response of a network also contains all the information necessary to predict the transient response of the network to any signal.

Let us look qualitatively at the transient response of a bandpass filter. If a resonance is narrow compared to its frequency, then it is said to be a high-"Q" resonance.* Figure 2.17a shows a high-Q filter frequency response. It has a transient response that dies out very slowly. A time response that decays slowly is said to be "lightly damped." Figure 2.17b shows a low-Q resonance. It has a transient response that dies out quickly. This illustrates a general principle: *signals that are broad in one domain are narrow in the other*. Narrow, selective filters have very long response times, a fact we will find important in the next section.

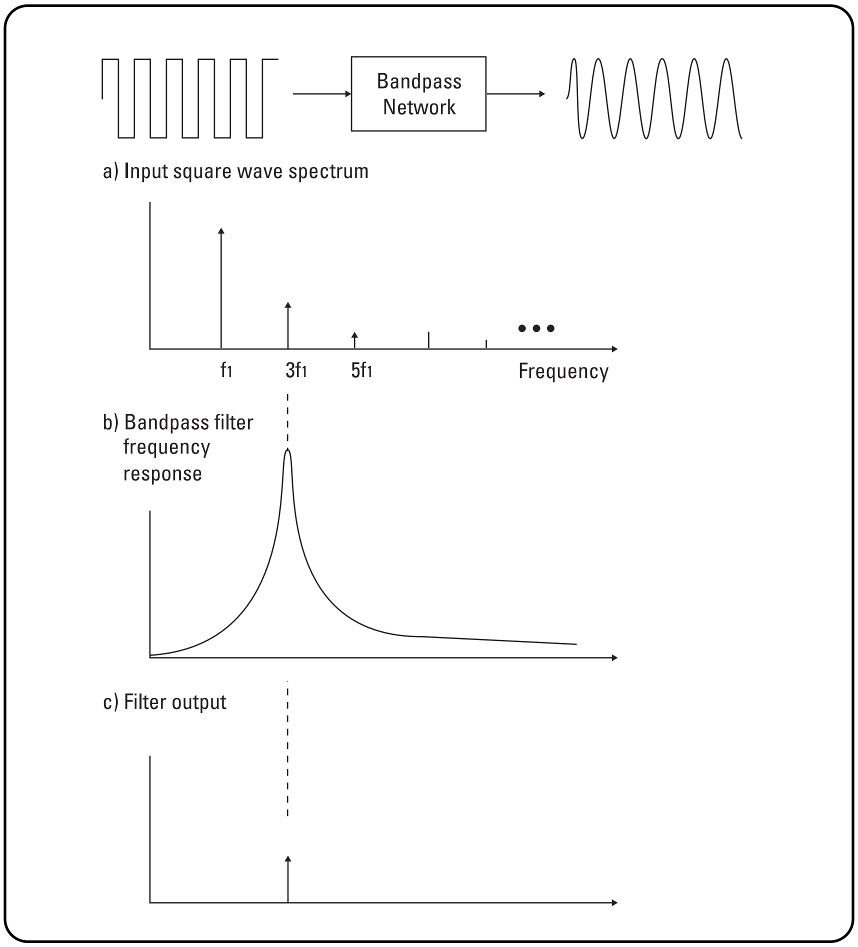

Figure 2.16. Bandpass filter response to a square wave input

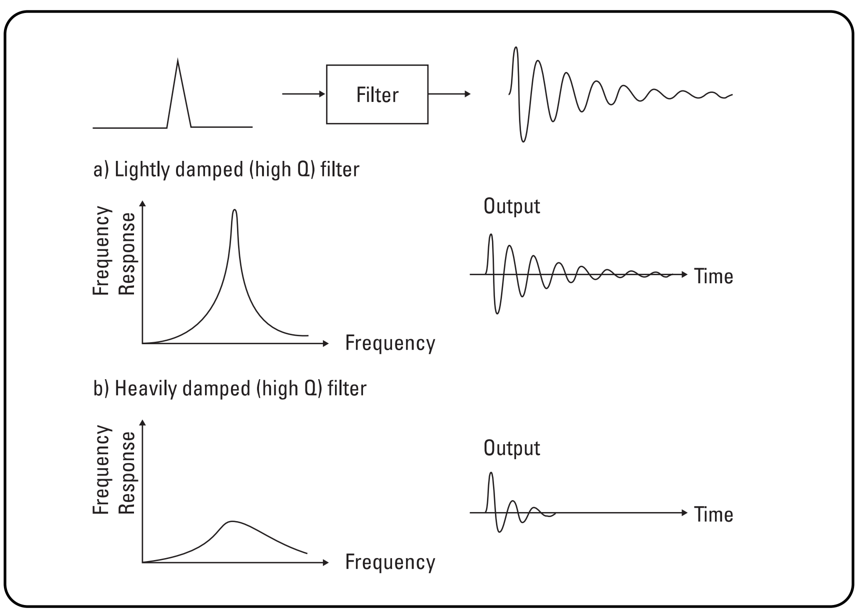

Figure 2.17. Time response of bandpass filters

* Q is usually defined as:

$$Q = \frac{\text{Center Frequency of Resonance}}{\text{Frequency Width of -3 dB Points}}$$

Section 3: Instrumentation for the Frequency Domain

Just as the time domain can be measured with strip chart recorders, oscillographs or oscilloscopes, the frequency domain is usually measured with spectrum and network analyzers.

Spectrum analyzers are instruments that are optimized to characterize signals. They introduce very little distortion and few spurious signals. This insures that the signals on the display are truly part of the input signal spectrum, not signals introduced by the analyzer.

Network analyzers are optimized to give accurate amplitude and phase measurements over a wide range of network gains and losses. This design difference means that these two traditional instrument families are not interchangeable.* A spectrum analyzer cannot be used as a network analyzer because it does not measure amplitude accurately and cannot measure phase. A network analyzer would make a very poor spectrum analyzer because spurious responses limit its dynamic range.

In this section we will discuss the properties of several types of analyzers in these two categories.

The Parallel-Filter Spectrum Analyzer

As we developed in Section 2 of this chapter, electronic filters can be built which pass a narrow band of frequencies. If we were to add a meter to the output of such a bandpass filter, we could measure the power in the portion of the spectrum passed by the filter. In Figure 3.1a we have done this for a bank of filters, each tuned to a different frequency. If the center frequencies of these filters are chosen so that the filters overlap properly, the spectrum covered by the filters can be completely characterized as in Figure 3.1b.

How many filters should we use to cover the desired spectrum? Here we have a trade-off. We would like to be able to see closely spaced spectral lines, so we should have a large number of filters. However, each filter is expensive and becomes more expensive as it becomes narrower, so the cost of the analyzer goes up as we improve its resolution. Typical audio parallel-filter analyzers balance these demands with 32 filters, each covering 1/3 of an octave.

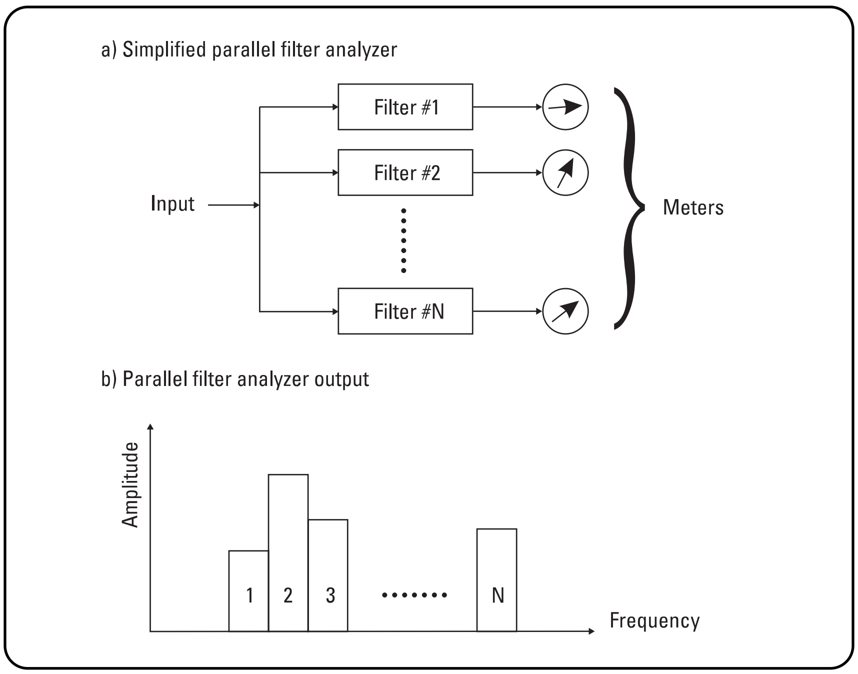

Figure 3.1. Parallel filter analyzer

* Dynamic signal analyzers are an exception to this rule. They can act as both network and spectrum analyzers.

Swept Spectrum Analyzer

One way to avoid the need for such a large number of expensive filters is to use only one filter and sweep it slowly through the frequency range of interest. If, as in Figure 3.2, we display the output of the filter versus the frequency to which it is tuned, we have the spectrum of the input signal. This swept analysis technique is commonly used in RF and microwave spectrum analysis.

We have, however, assumed the input signal hasn't changed in the time it takes to complete a sweep of our analyzer. If energy appears at some frequency at a moment when our filter is not tuned to that frequency, then we will not measure it.

One way to reduce this problem would be to speed up the sweep time of our analyzer. We could still miss an event, but the time in which this could happen would be shorter. Unfortunately though, we cannot make the sweep arbitrarily fast because of the response time of our filter.

To understand this problem, recall from Section 2 that a filter takes a finite time to respond to changes in its input. The narrower the filter, the longer it takes to respond. If we sweep the filter past a signal too quickly, the filter output will not have a chance to respond fully to the signal. As we show in Figure 3.3, the spectrum display will then be in error; our estimate of the signal level will be too low.

In a parallel-filter spectrum analyzer we do not have this problem. All the filters are connected to the input signal all the time. Once we have waited the initial settling time of a single filter, all the filters will be settled and the spectrum will be valid and not miss any transient events.

So there is a basic trade-off between parallel-filter and swept spectrum analyzers. The parallel-filter analyzer is fast, but has limited resolution and is expensive. The swept analyzer can be cheaper and have higher resolution, but the measurement takes longer (especially at high resolution), and it cannot analyze transient events*.

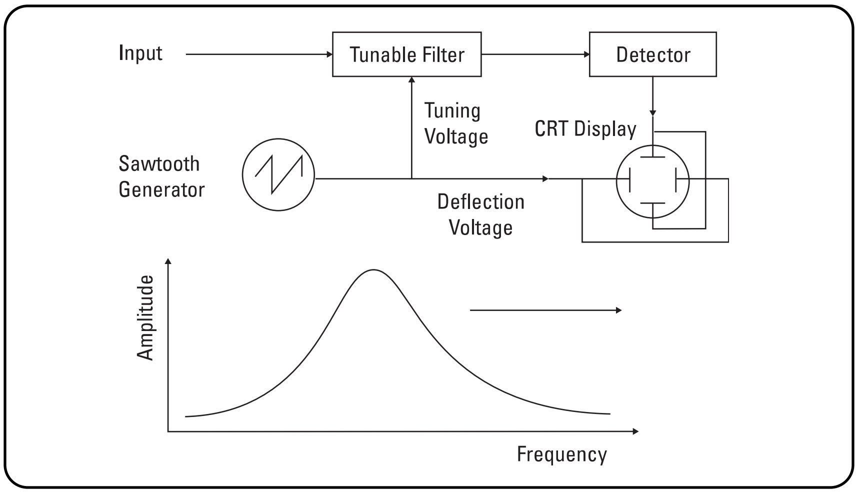

Figure 3.2. Simplified swept spectrum analyzer

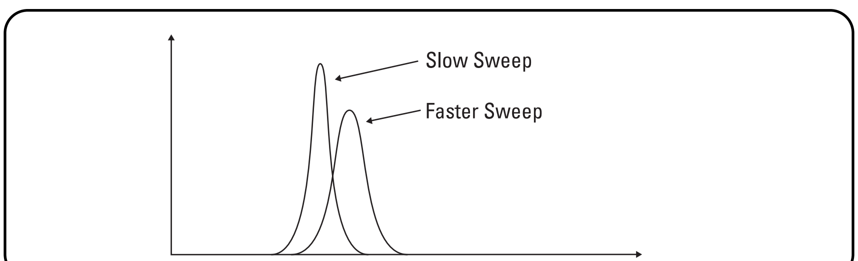

Figure 3.3. Amplitude error from sweeping too fast

* More information on the performance of swept spectrum analyzers can be found in Agilent Application Note Series 150.

Dynamic Signal Analyzer

In recent years, another kind of analyzer has been developed which offers the best features of the parallel-filter and swept spectrum analyzers. Dynamic signal analyzers are based on a high-speed calculation routine that acts like a parallel filter analyzer with hundreds of filters, yet they are cost competitive with swept spectrum analyzers. In addition, two-channel dynamic signal analyzers are in many ways better network analyzers than the ones we will introduce next.

Network Analyzers

Network analysis requires measurements of both the input and output, so network analyzers are generally two-channel devices with the capability of measuring the amplitude ratio (gain or loss) and phase difference between the channels. All of the analyzers discussed here measure frequency response by using a sinusoidal input to the network and slowly changing its frequency. Dynamic signal analyzers use a different, much faster technique for network analysis. See Agilent Application Note 1405-2 for more information.

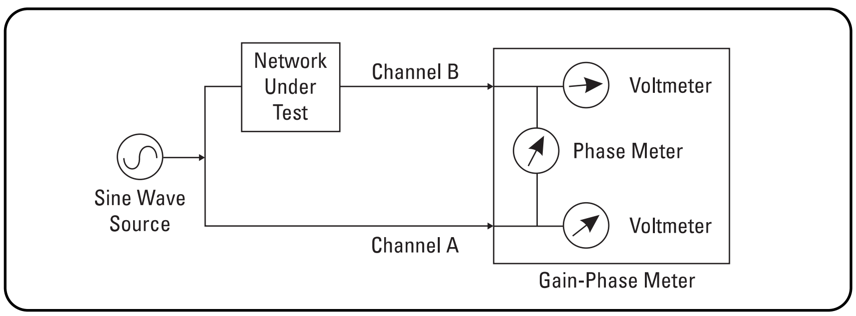

Figure 3.4. Gain-phase meter operation

Gain-phase meters are broadband devices that measure the amplitude and phase of the input and output sine waves of the network. A sinusoidal source must be supplied to stimulate the network when using a gain-phase meter as in Figure 3.4. The source can be tuned manually and the gain-phase plots done by hand or a sweeping source, and an x-y plotter can be used for automatic frequency response plots.

The primary attraction of gain-phase meters is their low price. If a sinusoidal source and a plotter are already available, frequency response measurements can be made for a very low investment. However, because gain-phase meters are broadband, they measure all the noise of the network as well as the desired sine wave. As the network attenuates the input, this noise eventually becomes a floor below which the meter cannot measure. This typically becomes a problem with attenuations of about 60 dB (1,000:1).

Tuned network analyzers minimize the noise floor problems of gain-phase meters by including a bandpass filter which tracks the source frequency. Figure 3.5 shows how this tracking filter virtually eliminates the noise and any harmonics to allow measurements of attenuation to 100 dB (100,000:1).

By minimizing the noise, it is also possible for tuned network analyzers to make more accurate measurements of amplitude and phase. These improvements do not come without their price, however, as tracking filters and a dedicated source must be added to the simpler and less costly gain-phase meter.

Tuned analyzers are available in the frequency range of a few Hertz to many Gigahertz (10^9 Hertz). If lower frequency analysis is desired, a frequency response analyzer is often used. To the operator, it behaves exactly like a tuned network analyzer. However, it is quite different inside. It integrates the signals in the time domain to effectively filter the signals at very low frequencies where it is not practical to make filters by more conventional techniques. Frequency response is limited to a range from 1 mHz to about 10 kHz.

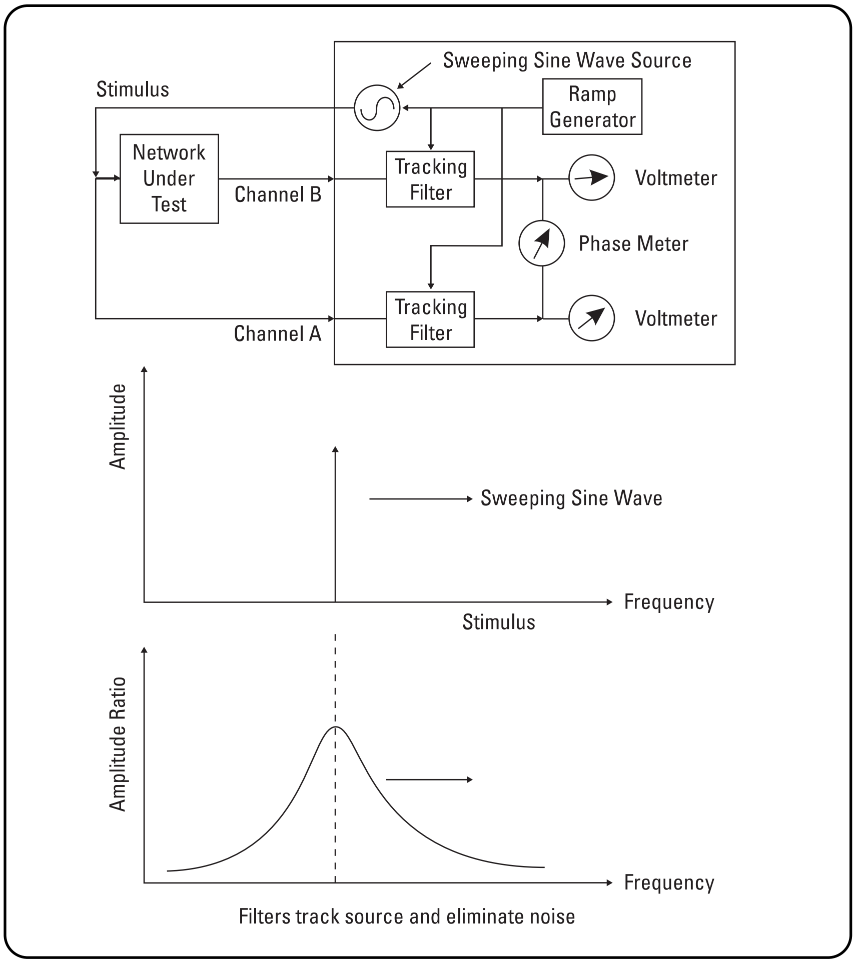

Figure 3.5. Tuned network analyzer operation

Section 4: The Modal Domain

In the preceding sections we discussed the properties of the time and frequency domains and the instrumentation used in these domains. In this section, we will delve into the properties of another domain, the modal domain. This change in perspective to a new domain is particularly useful if we are interested in analyzing the behavior of mechanical structures.

To understand the modal domain, let us begin by analyzing a simple mechanical structure, a tuning fork. If we strike a tuning fork, we easily conclude from its tone that it is primarily vibrating at a single frequency. We see that we have excited a network (tuning fork) with a force impulse (hitting the fork). The time domain view of the sound caused by the deformation of the fork is a lightly damped sine wave shown in Figure 4.1b.

In Figure 4.1c, we see in the frequency domain that the frequency response of the tuning fork has a major peak that is very lightly damped, which is the tone we hear. There are also several smaller peaks.

Each of these peaks, large and small, corresponds to a "vibration mode" of the tuning fork. For instance in this simple example, we might expect the major tone to be caused by the vibration mode shown in Figure 4.2a. The second harmonic might be caused by a vibration like Figure 4.2b.

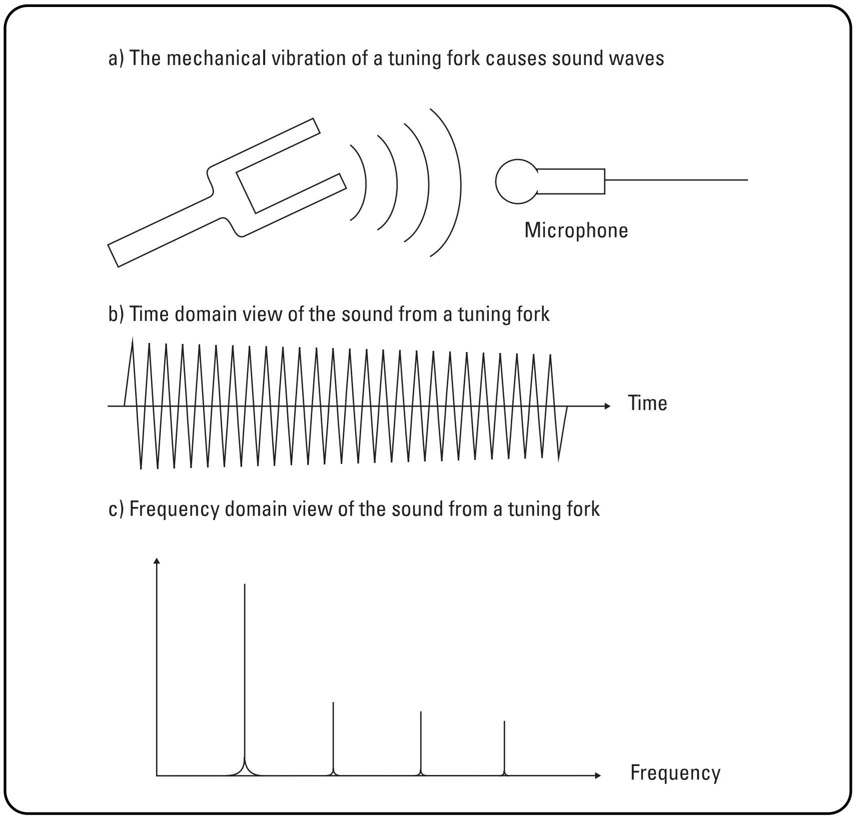

Figure 4.1. The vibration of a tuning fork

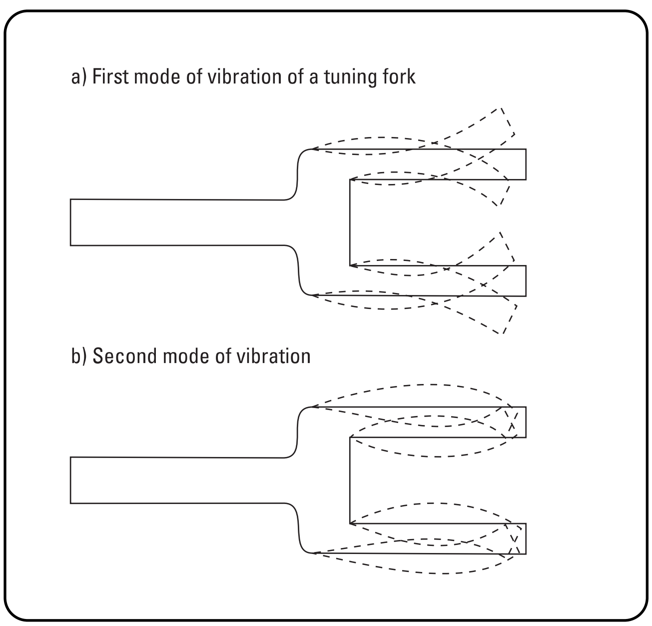

Figure 4.2. Example vibration modes of a tuning fork

We can express the vibration of any structure as a sum of its vibration modes. Just as we can represent a real waveform as a sum of much simpler sine waves, we can represent any vibration as a sum of much simpler vibration modes. The task of “modal” analysis is to determine the shape and the magnitude of the structural deformation in each vibration mode. Once these are known, it usually becomes apparent how to change the overall vibration.

For instance, let us look again at our tuning fork example. Suppose that we decided that the second harmonic tone was too loud. How should we change our tuning fork to reduce the harmonic? If we had measured the vibration of the fork and determined that the modes of vibration were those shown in Figure 4.2, the answer becomes clear. We might apply damping material at the center of the tines of the fork (see Figure 4.3). This would greatly affect the second mode that has maximum deflection at the center, while only slightly affecting the desired vibration of the first mode. Other solutions are possible, but all depend on knowing the geometry of each mode.

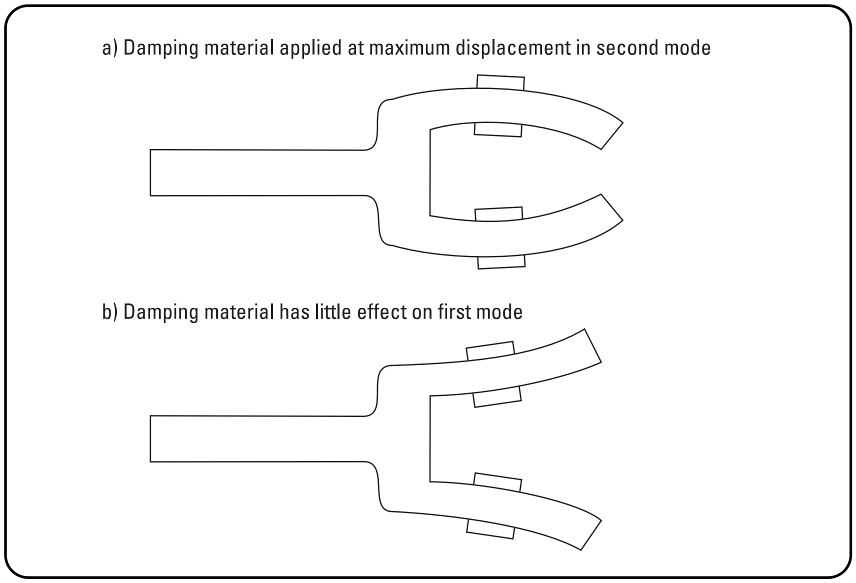

Figure 4.3. Reducing the second harmonic by damping the second vibration mode

The Relationship between the Time, Frequency and Modal Domain

To determine the total vibration of our tuning fork or any other structure, we have to measure the vibration at several points on the structure. Figure 4.4a shows some points we might pick. If we transformed this time domain data to the frequency domain, we would get results like Figure 4.4b. We measure frequency response because we want to measure the properties of the structure independent of the stimulus.*

We see that the sharp peaks (resonances) all occur at the same frequencies independent of where they are measured on the structure. Likewise we would find by measuring the width of each resonance that the damping (or Q) of each resonance is independent of position. The only parameter that varies as we move from point to point along the structure is the relative height of resonances.** By connecting the peaks of the resonances of a given mode, we trace out the mode shape of that mode.

Experimentally we have to measure only a few points on the structure to determine the mode shape. However, to clearly show the mode shape in our figure, we have drawn in the frequency response at many more points in Figure 4.5a. If we view this three-dimensional graph along the distance axis, as in Figure 4.5b, we get a combined frequency

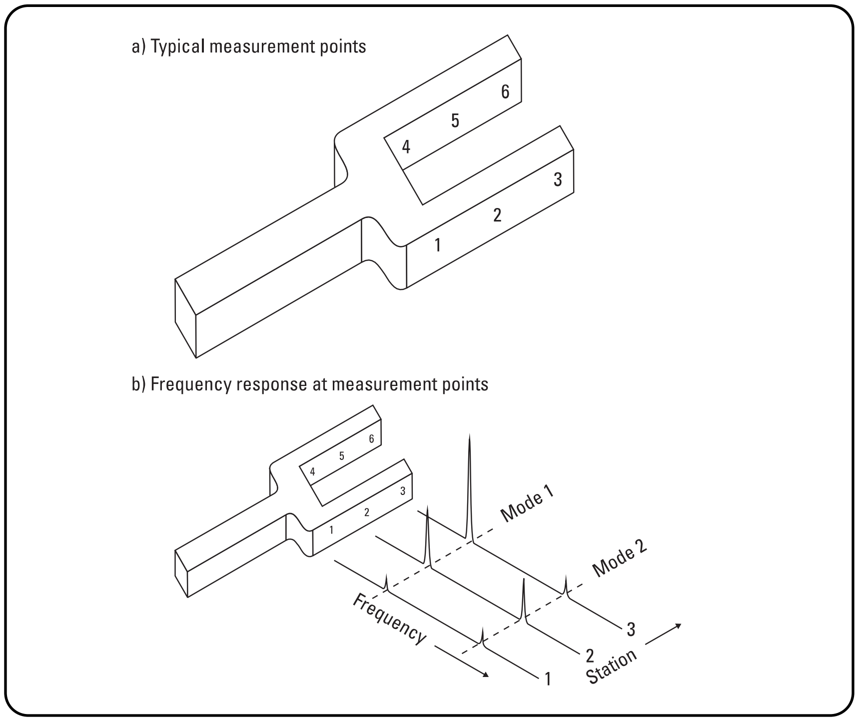

Figure 4.4. Modal analysis of a tuning fork

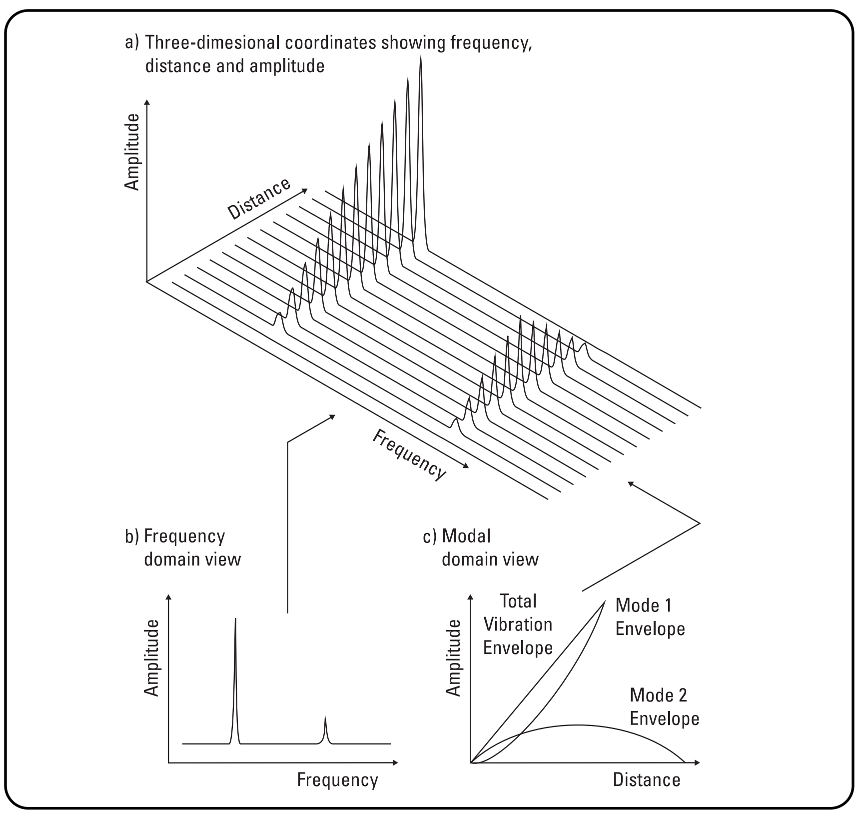

Figure 4.5. The relationship between the frequency and modal domains

* Those who are more familiar with electronics might note that we have measured the frequency response of a network (structure) at N points and thus have performed an N-port analysis.

** The phase of each resonance is not shown for clarity of the figures but it, too, is important in the mode shape. The magnitude of the frequency response gives the magnitude of the mode shape, while the phase gives the direction of the deflection.

response. Each resonance has a peak value corresponding to the peak displacement in that mode. If we view the graph along the frequency axis, as in Figure 4.5c, we can see the mode shapes of the structure.

We have not lost any information by this change of perspective. Each vibration mode is characterized by its mode shape, frequency and damping from which we can reconstruct the frequency domain view.

However, the equivalence between the modal, time and frequency domains is not quite as strong as that between the time and frequency domains. Because the modal domain portrays the properties of the network independent of the stimulus, transforming back to the time domain gives the impulse response of the structure, no matter what the stimulus. A more important limitation of this equivalence is that curve fitting is used in transforming from our frequency response measurements to the modal domain to minimize the effects of noise and small experimental errors. No information is lost in this curve fitting, so all three domains contain the same information, but not the same noise. Therefore, transforming from the frequency domain to the modal domain and back again will give results like those in Figure 4.6. The results are not exactly the same, yet in all the important features, the frequency responses are the same. This is also true of time domain data derived from the modal domain.

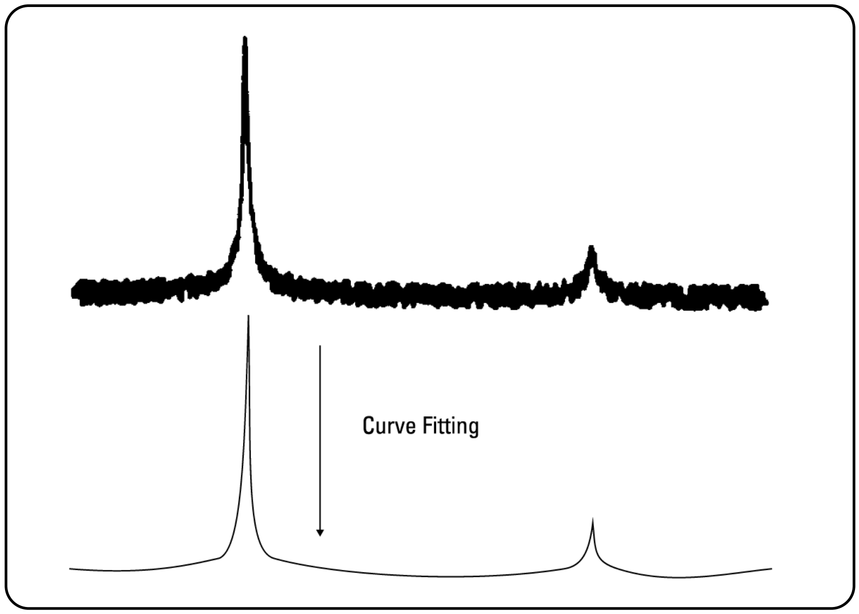

Figure 4.6. Curve fitting removes measurement noise.

There are many ways that the modes of vibration can be determined. In our simple tuning fork example, we could guess what the modes were. In simple structures like drums and plates it is possible to write an equation for the modes of vibration. However, in almost any real problem, the solution can neither be guessed nor solved analytically because the structure is too complicated. In these cases it is necessary to measure the response of the structure and determine the modes.

There are two basic techniques for determining the modes of vibration in complicated structures: 1) exciting only one mode at a time, and 2) computing the modes of vibration from the total vibration.

Single-Mode Excitation Modal Analysis

To illustrate single-mode excitation, let us look once again at our simple tuning fork example. To excite just the first mode, we need two shakers, driven by a sine wave and attached to the ends of the tines as in Figure 5.1a. Varying the frequency of the generator near the first mode resonance frequency would then give us its frequency, damping and mode shape.

In the second mode, the ends of the tines do not move, so to excite the second mode we must move the shakers to the center of the tines. If we anchor the ends of the tines, we will constrain the vibration to the second mode alone.

In more realistic, three-dimensional problems, it is necessary to add many more shakers to ensure that only one mode is excited. The difficulties and expense of testing with many shakers has limited the application of this traditional modal analysis technique.

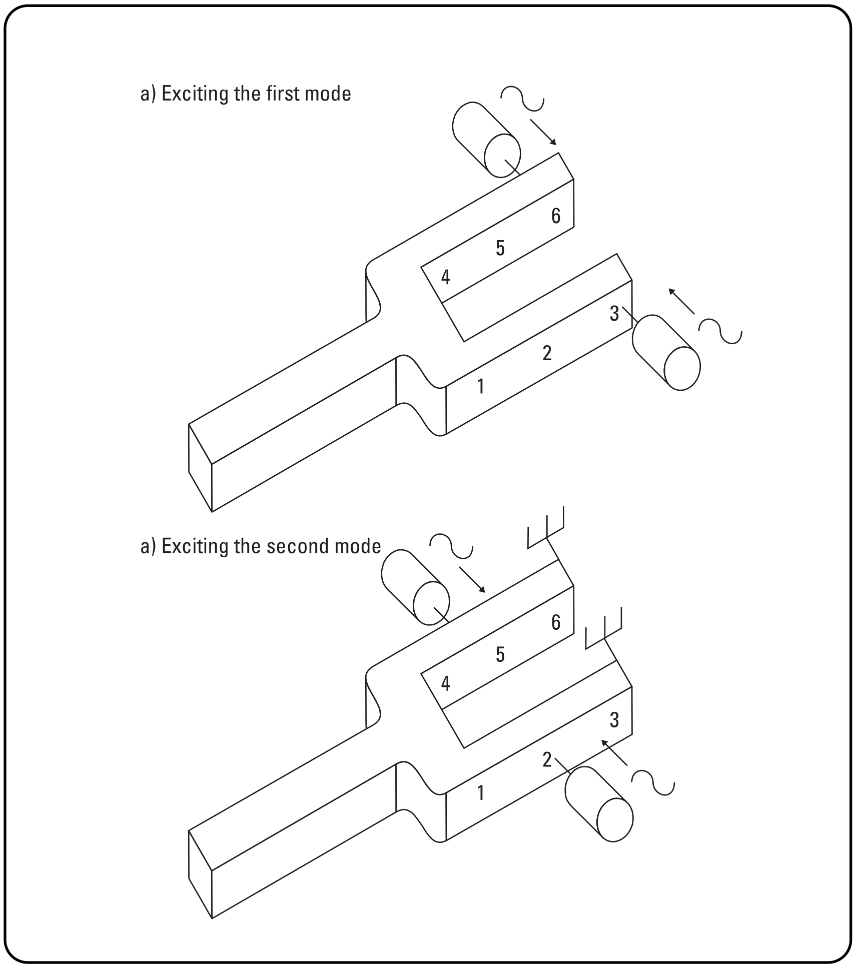

Figure 5.1. Single-mode excitation modal analysis

Modal Analysis from Total Vibration

To determine the modes of vibration from the total vibration of the structure, we use the techniques developed in the previous section. Basically, we determine the frequency response of the structure at several points and compute at each resonance the frequency, damping and what is called the residue (which represents the height of the resonance). This is done by a curve-fitting routine to smooth out any noise or small experimental errors. From these measurements and the geometry of the structure, the mode shapes are computed and drawn on a display or a plotter. You can animate these displays to help you understand the vibration mode.

From the above description, it is apparent that a modal analyzer requires some type of network analyzer to measure the frequency response of the structure and a computer to convert the frequency response to mode shapes. This can be accomplished by connecting a dynamic signal analyzer through a digital interface to a computer furnished with the appropriate software.

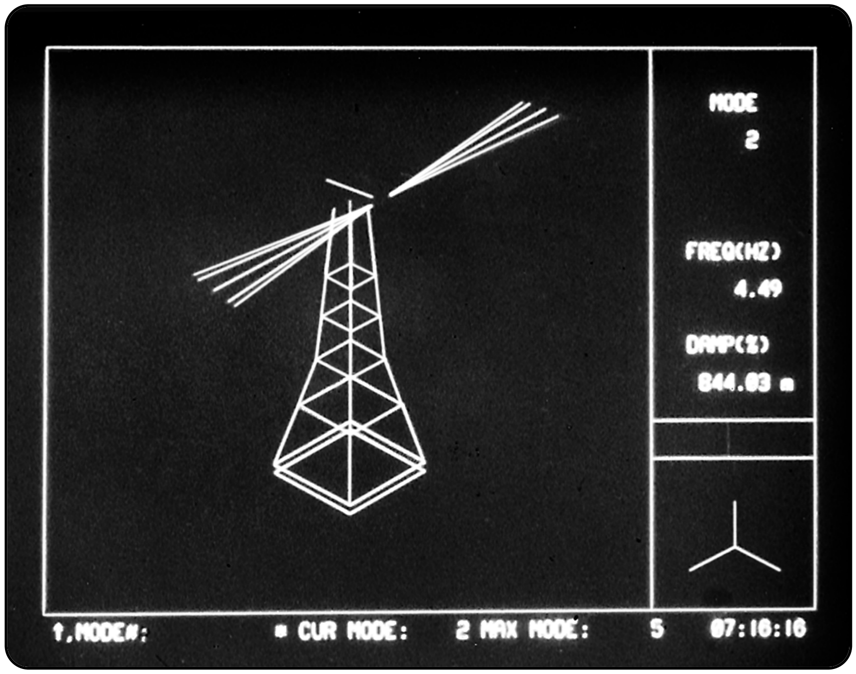

Figure 5.2. Measured mode shape

Summary

In this chapter we have developed the concept of looking at problems from different perspectives. These perspectives are the time, frequency and modal domains. Phenomena that are confusing in the time domain are often clarified by changing perspective to another domain. Small signals are easily resolved in the presence of large ones in the frequency domain. The frequency domain is also valuable for predicting the output of any kind of linear network. A change to the modal domain breaks down complicated structural vibration problems into simple vibration modes.

No one domain is always the best answer, so the ability to easily change domains is quite valuable. Of all the instrumentation available today, only dynamic signal analyzers can work in all three domains. See Agilent Application Note 1405-2 for a discussion of the properties of this important class of analyzers.

Related Agilent Literature

Agilent Application Note — *Understanding Dynamic Signal Analysis*, pub. no. 1405-2

Agilent Application Note —*Using Dynamic Signal Analysers*, pub. no. 1405-3

Agilent Application Note — *The Fourier Transform: A Mathematical Background*, pub. no. 1405-4

Product Overview — *Agilent 35670A Dynamic Signal Analyzer,* pub. no. 5966-3063E

Product Overview — *Agilent E1432/33/34 VXI Digitizers/Source,* pub. no. 5968-7086E

Product Overview — *Agilent E9801B Data Recorder/Logger,* pub. no. 5968-6132E

Glossary

Accelerometer — A transducer whose output is directly proportional to acceleration. Typically uses piezoelectric crystals to produce output.

Curve-fit — A method for creating a mathematical model that best fits a set of sampled data. The least square method is commonly used.

Damping — The dissipation of energy with time or distance

Decibels (dB) — A logarithmic representation of a ratio expressed as 10 or 20 times the log of the ratio

Distortion — An undesired change in waveform

Fourier — French mathematician Jean Baptiste Joseph Fourier (1768-1830)

Fourier transform — An algorithm used to transform time domain data into frequency domain

Frequency response — A ratio of the output over the input, both as a function of frequency

Gain-phase meter — A two-channel instrument that compares the amplitude levels and phases of two signals and displays the results

Linearity — The response of each element is proportional to the excitation

Load cell — A transducer whose output is directly proportional to force

Modal analysis — A method for characterizing the dynamic behavior of a structure in terms of natural frequencies, mode shapes and damping

Network analysis — The general engineering problem of determining how a network will respond to an input. Network analysis is sometimes called stimulus/response testing. The input is then known as the stimulus or excitation, and the output is called the response.

Network analyzer — An instrument used to characterize the frequency response of electronic networks

Oscilloscope — An instrument that displays voltage waveforms as a function of time

Q (of resonance) — A measure of the sharpness of resonance or frequency selectivity of a resonant vibratory system having a single degree of freedom. In a mechanical system, equal to _ the reciprocal of the damping ratio.

Resonance — Resonance of a system in forced vibration exists when any change in frequency, however small, causes a decrease in system response

Shaker — A device for subjecting a mechanical system to controlled and reproducible mechanical vibration

Spectrum — A frequency domain representation of the signal

Spectrum analyzer — An instrument for characterizing waveforms in the frequency domain

Steady-state — The condition that exists after all initial transients or fluctuating conditions have damped out, and all currents, voltages, or fields remain essentially constant, or oscillate uniformly

Stimulus/response testing — Another name for network analysis, or determining how a network will respond to an input. The input is known as the stimulus or excitation, and the output is called the response.

Strip chart recorders — A device that uses one or more pens to record data on a strip of paper moving at a constant speed. The device provides a permanent graphic record of a parameter (i.e. displacement) vs. time.

Transient response — The transitional period of a system's response to excitations until it reaches the steady state. Used to characterize the dynamic behavior of a system.

Vibration mode — A characteristic pattern assumed by a vibrating system in which the motion of every particle is a simple harmonic with the same frequency. Two or more modes may exist concurrently in a system with multiple degrees of freedom.

www.agilent.com

Agilent Email Updates

www.agilent.com/find/emailupdates
Get the latest information on the products and applications you select.

By internet, phone, or fax, get assistance with all your test & measurement needs
Online assistance:
www.agilent.com/find/assist

Phone or Fax

United States:
(tel) 1 800 452 4844

Canada:
(tel) 1 877 894 4414
(fax) (905) 282 6495

China:
(tel) 800 810 0189
(fax) 800 820 2816

Europe:
(tel) (31 20) 547 2323
(fax) (31 20) 547 2390

Japan:
(tel) (81) 426 56 7832
(fax) (81) 426 56 7840

Korea:
(tel) (82 2) 2004 5004
(fax) (82 2) 2004 5115

Latin America:
(tel) (305) 269 7500
(fax) (305) 269 7599

Taiwan:
(tel) 0800 047 866
(fax) 0800 286 331

Other Asia Pacific Countries:
(tel) (65) 6375 8100
(fax) (65) 6836 0252
(e-mail) tm_asia@agilent.com

Printed in the USA May 24, 2002
5988-6765EN

Agilent Technologies