

Figure 2.16 Any arbitrary function, f(t), can be represented by summing up unit impulse functions, $\delta(t-\tau)$ applied at $t=\tau$ and multiplied by the area $f(\tau)\Delta\tau$ for all values of τ from $-\infty$ to t

used to describe any arbitrary shaped function as a sum of suitably scaled unit impulses, $\delta(t-a)$, applied at appropriate time, t=a. This fact is illustrated in Figure 2.16, where the function f(t) is represented by

$$f(t) = \sum_{\tau = -\infty}^{\infty} f(\tau) \Delta \tau \delta(t - \tau)$$
 (2.35)

or, in the limit $\Delta \tau \rightarrow 0$,

$$f(t) = \int_{-\infty}^{\infty} f(\tau)\delta(t - \tau) d\tau$$
 (2.36)

Equation (2.36) is one of the most important equations of modern control theory, because it lets us evaluate the response of a linear system to any arbitrary input, f(t), by the use of the superposition principle. We will see how this is done when we discuss the response to singularity functions in Section 2.5. While the singularity functions and their relatives are useful as test inputs for studying the behavior of control systems, we can also apply some well known continuous time functions as inputs to a control system. Examples of continuous time test functions are the harmonic functions $\sin(\omega t)$ and $\cos(\omega t)$, where ω is a frequency, called the excitation frequency. As an alternative to singularity inputs (which are often difficult to apply in practical cases), measuring the output of a linear system to harmonic inputs gives essential information about the system's behavior, which can be used to construct a model of the system that will be useful in designing a control system. We shall study next how such a model can be obtained.

2.3 Frequency Response

Frequency response is related to the steady-state response of a system when a harmonic function is applied as the input. Recall from Section 1.2 that steady-state response is the linear system's output after the transient response has decayed to zero. Of course, the requirement that the transient response should have decayed to zero after some time calls for the linear system to be stable. (An unstable system will have a transient response shooting to infinite magnitudes, irrespective of what input is applied.) The steady-state

response of a linear system is generally of the same *shape* as that of the applied input, e.g. a step input applied to a linear, stable system yields a steady-state output which is also a step function. Similarly, the steady-state response of a linear, stable system to a harmonic input is also harmonic. Studying a linear system's characteristics based upon the steady-state response to *harmonic* inputs constitutes a range of *classical control* methods called the *frequency response methods*. Such methods formed the backbone of the *classical control theory* developed between 1900–60, because the modern *state-space* methods (to be discussed in Chapter 3) were unavailable then to give the response of a linear system to any arbitrary input directly in the *time domain* (i.e. as a function of time). Modern control techniques still employ frequency response methods to shed light on some important characteristics of an unknown control system, such as the *robustness* of multi-variable (i.e. multi-input, multi-output) systems. For these reasons, we will discuss frequency response methods here.

A simple choice of the harmonic input, u(t), can be

$$u(t) = u_o \cos(\omega t)$$
 or $u(t) = u_o \sin(\omega t)$ (2.37)

where u_o is the constant amplitude and ω is the frequency of excitation (sometimes called the driving frequency). If we choose to write the input (and output) of a linear system as complex functions, the governing differential equation can be replaced by complex algebraic equations. This is an advantage, because complex algebra is easier to deal with than differential equations. Furthermore, there is a vast factory of analytical machinery for dealing with complex functions, as we will sample later in this chapter. For these powerful reasons, let us express the harmonic input in the complex space as

$$u(t) = u_0 e^{i\omega t} \tag{2.38}$$

where $i = \sqrt{-1}$ (a purely imaginary quantity), and

$$e^{i\omega t} = \cos(\omega t) + i\sin(\omega t) \tag{2.39}$$

Equation (2.39) is a *complex* representation in which $\cos(\omega t)$ is called the *real part* of $e^{i\omega t}$ and $\sin(\omega t)$ is called the *imaginary part* of $e^{i\omega t}$ (because it is multiplied by the imaginary number i). The complex space representation of the harmonic input given by Eq. (2.38) is shown in Figure 2.17. The two axes of the complex plane are called the *real*

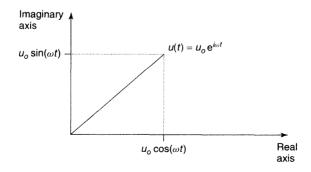


Figure 2.17 Complex space representation of a harmonic input, v(t)

and imaginary axis, respectively, as shown. Hence, complex space representation of a harmonic function is a device of representing both the possibilities of a simple harmonic input, namely $u_o \cos(\omega t)$ and $u_o \sin(\omega t)$, respectively, in one expression. By obtaining a steady-state response to the complex input given by Eq. (2.38), we will be obtaining simultaneously the steady-state responses of a linear, stable system to $u_o \cos(\omega t)$ and $u_o \sin(\omega t)$.

When you studied solution to ordinary differential equations, you learnt that their solution consists of two parts – the *complimentary solution* (or the solution to the *unforced* differential equation (Eq. (2.7)), and a *particular solution* which depends upon the input. While the transient response of a linear, stable system is *largely* described by the complimentary solution, the steady-state response is the same as the particular solution at large times. The particular solution is of the *same form* as the input, and must by itself satisfy the differential equation. Hence, you can verify that the steady-state responses to $u(t) = u_o \cos(\omega t)$ and $u(t) = u_o \sin(\omega t)$, are given by $y_{ss}(t) = y_o \cos(\omega t)$ and $y_{ss}(t) = y_o \sin(\omega t)$, respectively (where y_o is the amplitude of the resulting harmonic, steady-state output, $y_{ss}(t)$) by plugging the corresponding expressions of u(t) and $y_{ss}(t)$ into Eq. (2.4), which represents a general linear system. You will see that the equation is satisfied in each case. In the complex space, we can write the steady-state response to harmonic input as follows:

$$y_{ss}(t) = y_o(i\omega)e^{i\omega t} \tag{2.40}$$

Here, the steady-state response amplitude, y_o , is a *complex* function of the frequency of excitation, ω . We will shortly see the implications of a complex response amplitude. Consider a linear, lumped parameter, control system governed by Eq. (2.4) which can be re-written as follows

$$D_1\{y_{ss}(t)\} = D_2\{u(t)\} \tag{2.41}$$

where $D_1\{\cdot\}$ and $D_2\{\cdot\}$ are differential operators (i.e. they operate on the steady-state output, $y_{ss}(t)$, and the input, u(t), respectively, by differentiating them), given by

$$D_1\{\cdot\} = a_n d^n / dt^n + a_{n-1} d^{n-1} / dt^{n-1} + \dots + a_1 d / dt + a_0$$
 (2.42)

and

$$D_2\{\cdot\} = b_m d^m / dt^m + b_{m-1} d^{m-1} / dt^{m-1} + \dots + b_1 d / dt + b_0$$
 (2.43)

Then noting that

$$D_1(e^{i\omega t}) = [(i\omega)^n a_n d^n/dt^n + (i\omega)^{n-1} a_{n-1} d^{n-1}/dt^{n-1} + \dots + (i\omega)a_1 d/dt + a_0]e^{i\omega t}$$
(2.44)

and

$$D_2(e^{i\omega t}) = [(i\omega)^m b_m d^m / dt^m + (i\omega)^{m-1} b_{m-1} d^{m-1} / dt^{m-1} + \dots + (i\omega)b_1 d / dt + b_0]e^{i\omega t}$$
(2.45)

we can write, using Eq. (2.41),

$$y_o(i\omega) = G(i\omega)u_o \tag{2.46}$$

where $G(i\omega)$ is called the *frequency response* of the system, and is given by

$$G(i\omega) = [(i\omega)^m b_m + (i\omega)^{m-1} b_{m-1} + \dots + (i\omega)b_1 + b_o]/[(i\omega)^n a_n + (i\omega)^{n-1} a_{n-1} + \dots + (i\omega)a_1 + a_o]$$
(2.47)

Needless to say, the frequency response $G(i\omega)$ is also a complex quantity, consisting of both real and imaginary parts. Equations (2.46) and (2.47) describe how the steady-state output of a linear system is related to its input through the frequency response, $G(i\omega)$. Instead of the real and imaginary parts, an alternative description of a complex quantity is in terms of its magnitude and the phase, which can be thought of as a vector's length and direction, respectively. Representation of a complex quantity as a vector in the complex space is called a phasor. The length of the phasor in the complex space is called its magnitude, while the angle made by the phasor with the real axis is called its phase. The magnitude of a phasor represents the amplitude of a harmonic function, while the phase determines the value of the function at t = 0. The phasor description of the steady-state output amplitude is given by

$$y_o(i\omega) = |y_o(i\omega)|e^{i\alpha(\omega)}$$
 (2.48)

where $|y_o(i\omega)|$ is the magnitude and $\alpha(\omega)$ is the phase of $y_o(i\omega)$. It is easy to see that

$$|y_o(i\omega)| = [\text{real } \{y_o(i\omega)\}^2 + \text{imag } \{y_o(i\omega)\}^2]^{1/2};$$

$$\alpha(\omega) = \tan^{-1}[\text{imag } \{y_o(i\omega)\}/\text{real } \{y_o(i\omega)\}]$$
(2.49)

where real $\{\cdot\}$ and imag $\{\cdot\}$ denote the real and imaginary parts of a complex number. We can also express the frequency response, $G(i\omega)$, in terms of its magnitude, $|G(i\omega)|$, and phase, $\phi(\omega)$, as follows:

$$G(i\omega) = |G(i\omega)|e^{i\phi(\omega)}$$
 (2.50)

Substituting Eqs. (2.48) and (2.50) into Eq. (2.46), it is clear that $|y_o(i\omega)| = |G(i\omega)|u_o$ and $\alpha(\omega) = \phi(\omega)$. Hence, the steady-state response of a linear system excited by a harmonic input of amplitude u_o and zero phase ($u_o = u_o e^{i0}$) is given through Eq. (2.40) by

$$y_{ss}(t) = y_o(i\omega)e^{i\omega t} = |G(i\omega)|u_o e^{i\phi(\omega)}e^{i\omega t} = |G(i\omega)|u_o e^{i[\omega t + \phi(\omega)]}$$
(2.51)

Thus, the steady-state response to a zero phase harmonic input acquires its phase from the frequency response, which is purely a characteristic of the linear system. You can easily show that if the harmonic input has a *non-zero* phase, then the phase of the steady-state response is the *sum* of the input phase and the phase of the frequency response, $\phi(\omega)$. The phasor representation of the steady-state response amplitude is depicted in Figure 2.18.

From Eq. (2.51), it is clear that the steady-state response is governed by the amplitude of the harmonic input, u_o , and magnitude and phase of the frequency response, $G(i\omega)$, which represent the characteristics of the system, and are functions of the frequency of excitation. If we excite the system at various frequencies, and measure the magnitude and phase of the steady-state response, we could obtain $G(i\omega)$ using Eq. (2.51), and consequently, crucial information about the system's characteristics (such as the coefficients a_k

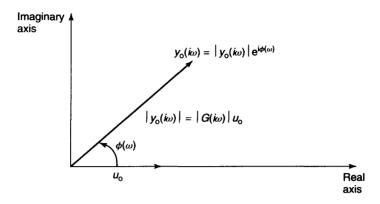


Figure 2.18 Phasor representations of a harmonic input, u(t), with zero phase and amplitude u_0 , and steady-state response amplitude, $y_0(i\omega)$, of a linear system with frequency response, $G(i\omega)$

and b_k , in Eq. (2.47)). In general, we would require $G(i\omega)$ at as many frequencies as are the number of unknowns, a_k and b_k , in Eq. (2.47). Conversely, if we know a system's parameters, we can study some of its properties, such as *stability* and *robustness*, using frequency response plots (as discussed later in this chapter). Therefore, plots of magnitude and phase of $G(i\omega)$ with frequency, ω , serve as important tools in the analysis and design of control systems. Alternatively, we could derive the same information as obtained from the magnitude and phase plots of $G(i\omega)$ from the *path* traced by the *tip* of the frequency response phasor in the complex space as the frequency of excitation is varied. Such a plot of $G(i\omega)$ in the complex space is called a *polar* plot (since it represents $G(i\omega)$ in terms of the *polar coordinates*, $|G(i\omega)|$ and $\phi(\omega)$). Polar plots have an advantage over the frequency plots of magnitude and phase in that both magnitude and phase can be seen in *one* (rather than two) plots. Referring to Figure 2.18, it is easily seen that a phase $\phi(\omega) = 0^{\circ}$ corresponds to the real part of $G(i\omega)$, while the phase $\phi(\omega) = 90^{\circ}$ corresponds to the imaginary part of $G(i\omega)$. When talking about stability and robustness properties, we will refer again to the polar plot.

Since the range of frequencies required to study a linear system is usually very large, it is often useful to plot the magnitude, $|G(i\omega)|$, and phase, $\phi(\omega)$, with respect to the frequency, ω , on a *logarithmic scale* of frequency, called *Bode plots*. In Bode plots, the magnitude is usually converted to *gain* in *decibels* (dB) by taking the logarithm of $|G(i\omega)|$ to the base 10, and multiplying the result with 20 as follows:

$$Gain = 20 \log_{10} |G(i\omega)| \tag{2.52}$$

As we will see later in this chapter, important information about a linear, single-input, single-output system's behavior (such as *stability* and *robustness*) can be obtained from the Bode plots, which serve as a cornerstone of classical control design techniques. Factoring the polynomials in $G(i\omega)$ (Eq. (2.47)) just produces addition of terms in $\log_{10} |G(i\omega)|$, which enables us to construct Bode plots by log-paper and pencil. Despite this, Bode plots are cumbersome to construct by hand. With the availability of personal computers and software with mathematical functions and graphics capability – such as MATLAB – Bode plots can be plotted quite easily. In MATLAB, all you have to do is

specify a set of frequencies, ω , at which the gain and phase plots are desired, and use the intrinsic functions *abs* and *angle* which calculate the magnitude and phase (in radians), respectively, of a complex number. If you have the MATLAB's *Control System Toolbox* (CST), the task of obtaining a Bode plot becomes even simpler through the use of the command *bode* as follows:

```
>>G=tf(num,den); bode(G,w) <enter> %a Bode plot will appear on the screen
```

Here >> is the MATLAB prompt, <enter> denotes the pressing of the 'enter' (or 'return') key, and the % sign indicates that everything to its right is a comment. In the bode command, w is the specified frequency vector consisting of equally spaced frequency values at which the gain and phase are desired, G is the name given to the frequency response of the linear, time-invariant system created using the CST LTI object function f which requires f0 numerator and f1 denominator polynomials, respectively, of f2 (ig2) in (Eq. (2.47)) in decreasing powers of f3. These coefficients should be be specified as follows, before using the f3 and f4 and f5 and f6 commands:

```
>>num=[b<sub>m</sub> b<sub>m-1</sub> ... b<sub>0</sub>]; den=[a<sub>n</sub> a<sub>n-1</sub> ... a<sub>0</sub>]; <enter>
```

By using the MATLAB command logspace, the w vector can also be pre-specified as follows:

```
>>w=logspace(-2,3); <enter> %w consists of equally spaced frequencies in the
range 0.01-1000 rad/s.
```

(Using a semicolon after a MATLAB command suppresses the print-out of the result on the screen.)

Obviously, w must be specified *before* you use the *bode* command. If you don't specify w, MATLAB will automatically generate an appropriate w vector, and create the plot.

Instead of plotting the Bode plot, you may like to store the magnitude (mag), $|G(i\omega)|$, and the *phase*, $\phi(\omega)$, at given set of frequencies, w, for further processing by using the following MATLAB command:

```
>>[mag,phase,w]=bode(num,den,w); <enter>
```

For more information about Bode plots, do the following:

```
>>help bode <enter>
```

The same procedure can be used to get help on any other MATLAB command. The example given below will illustrate what Bode plots look like. Before we do that, let us try to understand in physical terms what a frequency response (given by the Bode plot) is.

Musical notes produced by a guitar are related to its frequency response. The guitar player makes each string vibrate at a particular frequency, and the notes produced by the various strings are the measure of whether the guitar is being played well or not. Each string of the guitar is capable of being excited at many frequencies, depending upon where

the string is struck, and where it is held. Just like the guitar, any system can be excited at a set of frequencies. When we use the word excited, it is quite in the literal sense, because it denotes the condition (called *resonance*) when the magnitude of the frequency response, $|G(i\omega)|$, becomes very large, or infinite. The frequencies at which a system can be excited are called its natural (or resonant) frequencies. High pitched voice of many a diva has shattered the opera-house window panes while accidently singing at one of the natural frequencies of the window! If a system contains energy dissipative processes (called *damping*), the frequency response magnitude at natural frequencies is large, but finite. An undamped system, however, has infinite response at each natural frequency. A natural frequency is indicated by a peak in the gain plot, or as the frequency where the phase changes by 180°. A practical limitation of Bode plots is that they show only an *inter*polation of the gain and phase through selected frequency points. The frequencies where $|G(i\omega)|$ becomes zero or infinite are excluded from the gain plot (since logarithm of zero is undefined, and an infinite gain cannot be shown on any scale). Instead, only frequency points located *close* to the zero magnitude frequency and the infinite gain frequencies of the system can be used in the gain plot. Thus, the Bode gain plot for a guitar will consist of several peaks, corresponding to the natural frequencies of the notes being struck. One could determine from the peaks the approximate values of the natural frequencies.

Example 2.8

Consider the electrical network shown in Figure 2.19 consisting of three resistances, R_1 , R_2 , and R_3 , a capacitor, C, and an inductor, L, connected to a voltage source, e(t), and a switch, S. When the switch, S, is closed at time t = 0, the current passing through the resistance R_1 is $i_1(t)$, and that passing through the inductor, L, is $i_2(t)$. The input to the system is the applied voltage, e(t), and the output is the current, $i_2(t)$.

The two governing equations of the network are

$$e(t) = R_1 i_1(t) + R_3 [i_1(t) - i_2(t)]$$
(2.53)

$$0 = R_2 i_2(t) + R_3 [i_2(t) - i_1(t)] + L i_2^{(1)}(t) + (1/C) \int_0^t i_2(\tau) d\tau$$
 (2.54)

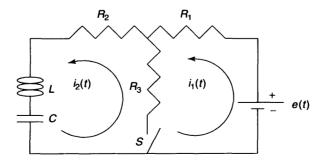


Figure 2.19 Electrical network for Example 2.8

Differentiating Eq. (2.54) and eliminating $i_1(t)$, we can write

$$Li_2^{(2)}(t) + [(R_1R_3 + R_1R_2 + R_2R_3)/(R_1 + R_3)]i_2^{(1)}(t)$$

+ $(1/C)i_2(t) = [R_3/(R_1 + R_3)]e^{(1)}(t)$ (2.55)

Comparing Eq. (2.55) with Eq. (2.4) we find that the system is linear and of second order, with $y(t) = i_2(t)$, u(t) = e(t), $a_0 = 1/C$, $a_1 = (R_1R_3 + R_1R_2 + R_2R_3)/(R_1 + R_3)$, $b_0 = 0$, and $b_1 = R_3/(R_1 + R_3)$. Hence, from Eq. (2.47), the frequency response of the system is given by

$$G(i\omega) = (i\omega)[R_3/(R_1 + R_3)]/[(i\omega)^2 L + (i\omega)(R_1 R_3 + R_1 R_2 + R_2 R_3)/(R_1 + R_3) + 1/C]$$
(2.56)

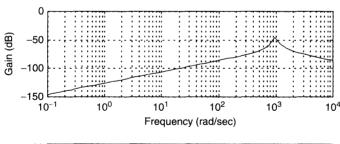
For $R_1 = R_3 = 10$ ohms, $R_2 = 25$ ohms, L = 1 henry, and $C = 10^{-6}$ farad, the frequency response is the following:

$$G(i\omega) = 0.5(i\omega)/[(i\omega)^2 + 30(i\omega) + 10^6]$$
 (2.57)

Bode gain and phase plots of frequency response given by Eq. (2.57) can be plotted in Figure 2.20 using the following MATLAB commands:

(This command produces equally spaced frequency points on logarithmic scale from 0.1 to $10\,000$ rad/s, and stores them in the vector w.)

$$>>G=i*w*0.5./(-w.*w+30*i*w+1e6);$$



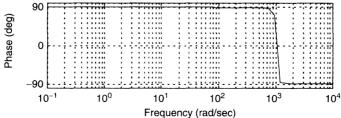


Figure 2.20 Bode plot for the electrical network in Example 2.8; a peak in the gain plot and the corresponding phase change of 180° denotes the natural frequency of the system

(This command calculates the value of $G(i\omega)$ by Eq. (2.57) at each of the specified frequency points in w, and stores them in the vector G. Note the MATLAB operations.* and J which allow element by element multiplication and division, respectively, of two arrays (see Appendix B).)

```
>>gain=20*log10(abs(G)); phase=180*angle(G)/pi; <enter>
```

(This command calculates the gain and phase of $G(i\omega)$ at each frequency point in w using the MATLAB intrinsic functions abs, angle, and log10, and stores them in the vectors gain and phase, respectively. We are assuming, however, that G does not become zero or infinite at any of the frequencies contained in w.)

(This command produces gain and phase Bode plots as two (unlabeled) subplots, as shown in Figure 2.20. Labels for the axes can be added using the MATLAB commands *xlabel* and *ylabel*.)

The Bode plots shown in Figure 2.20 are obtained much more easily through the Control System Toolbox (CST) command *bode* as follows:

```
>>num=[0.5 0]; den=[1 30 1e6]; g=tf(num,den), bode(g,w) <enter>
```

Note the peak in the gain plot of Figure 2.20 at the frequency, $\omega = 1000$ rad/s. At the same frequency the phase changes by 180° . Hence, $\omega = 1000$ rad/s is the system's natural frequency. To verify whether this is the *exact* natural frequency, we can rationalize the denominator in Eq. (2.57) (i.e. make it a real number by multiplying both numerator and denominator by a suitable complex factor – in this case $(-\omega^2 + 10^6) - 30i\omega$ and express the magnitude and phase as follows:

$$|G(i\omega)| = [225\omega^4 + 0.25\omega^2(-\omega^2 + 10^6)^2]^{1/2} / [(-\omega^2 + 10^6)^2 + 900\omega^2];$$

$$\phi(\omega) = \tan^{-1}(-\omega^2 + 10^6) / (30\omega)$$
(2.58)

From Eq. (2.58), it is clear that $|G(i\omega)|$ has a maximum value (0.0167 or -35.547 dB) – and $\phi(\omega)$ jumps by 180° – at $\omega=1000$ rad/s. Hence, the natural frequency is exactly 1000 rad/s. Figure 2.20 also shows that the gain at $\omega=0.1$ rad/s is -150 dB, which corresponds to $|G(0.1i)|=10^{-7.5}=3.1623\times10^{-8}$, a small number. Equation (2.58) indicates that |G(0)|=0. Hence, $\omega=0.1$ rad/s approximates quite well the zero-frequency gain (called the DC gain) of the system. The frequency response is used to define a linear system's property called bandwidth defined as the range of frequencies from zero up to the frequency, ω_b , where $|G(i\omega_b)|=0.707|G(0)|$. Examining the numerator of $|G(i\omega)|$ in Eq. (2.58), we see that $|G(i\omega)|$ vanishes at $\omega=0$ and $\omega=1999\,100$ rad/s (the numerator roots can be obtained using the MATLAB intrinsic function roots). Since |G(0)|=0, the present system's bandwidth is $\omega_b=1999\,100$ rad/s (which lies beyond the frequency range of Figure 2.20). Since the degree of the denominator polynomial of $G(i\omega)$ in Eq. (2.47) is greater than that of the numerator polynomial, it follows

that $|G(i\omega)| \to 0$ as $\omega \to \infty$. Linear systems with $G(i\omega)$ having a higher degree denominator polynomial (than the numerator polynomial) in Eq. (2.47) are called *strictly proper* systems. Equation (2.58) also shows that $\phi(\omega) \to 90^{\circ}$ as $\omega \to 0$, and $\phi(\omega) \to -90^{\circ}$ as $\omega \to \infty$. For a general system, $\phi(\omega) \to -k90^{\circ}$ as $\omega \to \infty$, where k is the number by which the degree of the denominator polynomial of $G(i\omega)$ exceeds that of the numerator polynomial (in the present example, k=1).

Let us now draw a polar plot of $G(i\omega)$ as follows (note that we need more frequency points close to the natural frequency for a smooth polar plot, because of the 180° phase jump at the natural frequency):

```
>>w=[logspace(-1,2.5) 350:2:1500 logspace(3.18,5)]; <enter>
```

(This command creates a frequency vector, w, with more frequency points close to 1000 rad/s.)

```
>>G=i*w*0.5./(-w.*w+30*i*w+1e6); <enter>
>>polar(angle(G), abs(G)); <enter>
```

(This command for generating a polar plot requires phase angles in *radians*, but the plot *shows* the phase in *degrees*.)

The resulting polar plot is shown in Figure 2.21. The plot is in polar coordinates, $|G(i\omega)|$ and $\phi(\omega)$, with circles of constant radius, $|G(i\omega)|$, and radial lines of constant $\phi(\omega)$ overlaid on the plot. Conventionally, polar plots show either all positive, or all negative phase angles. In the present plot, the negative phase angles have been shown as positive angles using the transformation $\phi \to (\phi + 360^{\circ})$, which is acceptable since both sine and cosine functions are invariant under this transformation for $\phi < 0$ (e.g. $\phi = -90^{\circ}$ is the same as $\phi = 270^{\circ}$). Note that the 0° and

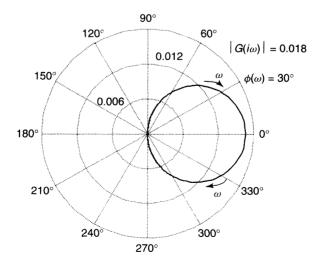


Figure 2.21 Polar plot of the frequency response, $G(i\omega)$, of the electrical system of Example 2.8