16. Frequency response

In Section 3 we discussed the frequency response of a first order LTI operator. In Section 10 we used the Exponential Response Formula to understand the response of an LTI operator to a sinusoidal input signal. Here we will study this in more detail, and understand how gain and phase lag vary with the driving frequency. We'll focus on second order equations, but the methods are quite general.

A differential equation relates input signal to system response. What constitutes the "input signal" and what constitutes the "system response" are matters of convenience for the user, and are not determined by the differential equation. We saw this already in connection with Newtonian cooling in Section 2, and here we will illustrate this in a couple of second order examples. The case of sinusoidal input signal and system response are particularly important. The question is then: what is the ratio of amplitude of the system response to that of the input signal, and what is the phase lag of the system response relative to the input signal?

We will carry this analysis out three times: first for two specific examples of mechanical (or electrical) systems, and then in general using the notation of the damping ratio.

16.1. **Driving through the spring.** The Mathlet `Amplitude and Phase: Second order I` illustrates a spring/mass/dashpot system that is driven through the spring. Suppose that y denotes the displacement of the plunger at the top of the spring, and $x(t)$ denotes the position of the mass, arranged so that $x = y$ when the spring is unstretched and uncompressed. There are two forces acting on the mass: the spring exerts a force force given by $k(y - x)$ (where k is the spring constant), and the dashpot exerts a force given by $-b\dot{x}$ (against the motion of the mass, with damping coefficient b). Newton's law gives

$$m\ddot{x} = k(y - x) - b\dot{x}$$

or, putting the system on the left and the driving term on the right,

(1) $$m\ddot{x} + b\dot{x} + kx = ky\,.$$

In this example it is natural to regard y, rather than ky, as the input signal, and the mass position x as the system response.

Another system leading to the same equation is a series RLC circuit, discussed in Section 8 and illustrated in the Mathlet `Series RLC Circuit`. We consider the impressed voltage as the input signal, and
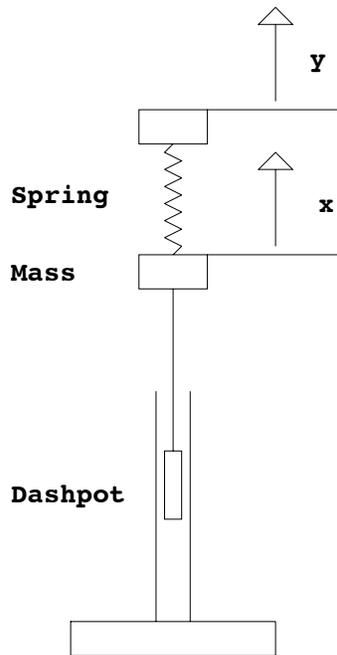
FIGURE 6. Spring-driven system

the voltage drop across the capacitor as the system response. The equation is then

$$L\ddot{V}_C + R\dot{V}_C + (1/C)V_C = (1/C)V$$

Again, we want to think of V as the input signal, but it is $(1/C)V$ that occurs on the right side of the equation. We will favor the mechanical system notation, but the mathematics is exactly the same in both systems.

When y is sinusoidal, say

$$y = A\cos(\omega t)\,,$$

then (putting aside the possibility of resonance) we expect a sinusoidal solution, one of the form

$$x = B\cos(\omega t - \phi)$$

The ratio of the amplitude of the system response to that of the input signal, B/A, is called the **gain** of the system. We think of the system as fixed, while the frequency ω of the input signal can be varied, so the gain is a function of ω, $g(\omega)$. Similarly, the **phase lag** ϕ is a function of ω. The entire story of the steady state system response to sinusoidal

input signals is encoded in those to functions of ω, the gain and the phase lag.

There is a systematic way to work out what g and ϕ are. The equation (1) is the real part of a complex-valued differential equation:

$$m\ddot{z} + b\dot{z} + kz = Ake^{st}$$

with $s = i\omega$. The Exponential Response Formula gives the solution

$$z_p = \frac{Ak}{p(s)} e^{st}$$

where

$$p(s) = ms^2 + bs + k$$

(as long as $p(s) \neq 0$).

Our choice of input signal and system response correspond in the complex equation to regarding Ae^{st} as the input signal and z_p as the exponential system response. The **transfer function** is the ratio between the two:

$$W(s) = \frac{k}{p(s)}$$

so

$$z_p = W(s)Ae^{st}.$$

Now take $s = i\omega$. The **complex gain** for this system is

(2)
$$W(i\omega) = \frac{k}{k - m\omega^2 + ib\omega}.$$

The complex gain is a wonderful thing! I claim that

(3) $\qquad g(\omega) = |W(i\omega)| \qquad$ and $\qquad \phi(\omega) = -\mathrm{Arg}(W(i\omega))$

That is, the polar expression of $W(i\omega)$ contains both the gain g and the phase lag ϕ as follows:

(4)
$$W(i\omega) = ge^{-i\phi}$$

To verify this, substitute this expression into the formula for z_p—

$$z_p = g\,e^{-i\phi}Ae^{i\omega t} = gAe^{i(\omega t - \phi)}$$

—and extract the real part, to get the sinusoidal solution to (1):

$$y_p = gA\cos(\omega t - \phi).$$

This verifies (3).

The system response to a sinusoidal input signal is another sinusoidal signal, but the amplitude has been multiplied by the gain, which is in this case

$$(5) \qquad g(\omega) = |W(i\omega)| = \frac{k}{\sqrt{k^2 + (b^2 - 2mk)\omega^2 + m^2\omega^4}}$$

The phase lag of the system response, relative to the input signal, is $\phi = -\mathrm{Arg}(W(i\omega))$. In our spring system, ϕ is the argument of the denominator in (2), since $\mathrm{Arg}(1/z) = -\mathrm{Arg}(z)$, The tangent of the argument of a complex number is the ratio of the imaginary part by the real part, so

$$\tan \phi = \frac{b\omega}{k - m\omega^2}$$

The `Amplitude and Phase: Second order I` Mathlet shows how the gain varies with ω. Often there is a choice of frequency ω for which the gain is maximal: this is "near resonance." To compute what this frequency is, we can try to *minimize* the denominator in (5). That minimum occurs when $k^2 + (b^2 - 2mk)\omega^2 + m^2\omega^4$ is minimized. Setting the derivative of this expression to zero shows that the minima occur when ω is either 0 or the *resonant frequency*

$$\omega_r = \sqrt{\frac{k}{m} - \frac{b^2}{2m^2}}$$

When $b = 0$, this is the natural frequency $\omega_n = \sqrt{k/m}$ and we have true resonance; the gain becomes infinite. As we increase b, the resonant frequency decreases, till when $b = \sqrt{2mk}$ we find $\omega_r = 0$. For b less than this, practical resonance occurs only for $\omega = 0$.

16.2. **Driving through the dashpot.** Now suppose instead that we fix the top of the spring and drive the system by moving the bottom of the dashpot instead. This is illustrated in `Amplitude and Phase: Second order II`.

Suppose that the position of the bottom of the dashpot is given by $y(t)$, and again the mass is at $x(t)$, arranged so that $x = 0$ when the spring is relaxed. Then the force on the mass is given by

$$m\ddot{x} = -kx + b\frac{d}{dt}(y - x)$$

since the force exerted by a dashpot is supposed to be proportional to the speed of the piston moving through it. This can be rewritten

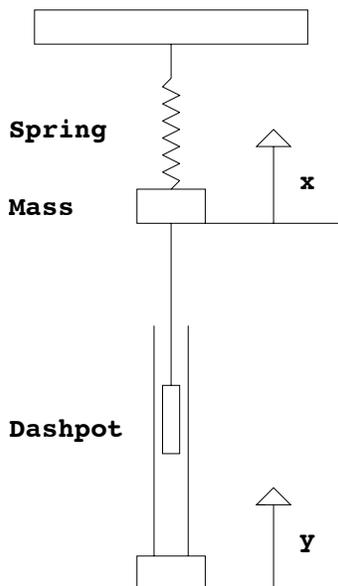$$(6) \qquad\qquad m\ddot{x} + b\dot{x} + kx = b\dot{y}\,.$$

FIGURE 7. Dashpot-driven system

Again we will consider x as the system response, and the position of the back end of the dashpot, y, as the input signal. Note that the *derivative* of the input signal (multiplied by b) occurs on the right hand side of the equation. Another system leading to the same mathematics is the series RLC circuit shown in the Mathlet `Series RLC Circuit`, in which the impressed voltage is the input variable and the voltage drop across the resistor is the system response. The equation is

$$L\ddot{V}_R + R\dot{V}_R + (1/C)V_R = R\dot{V}$$

Here's a frequency response analysis of this problem. We suppose that the input signal is sinusoidal:

$$y = B\cos(\omega t).$$

Then $\dot{y} = -\omega B\sin(\omega t)$ so our equation is

(7) $$m\ddot{x} + b\dot{x} + kx = -b\omega B\sin(\omega t).$$

The periodic system response will be of the form

$$x_p = gB\cos(\omega t - \phi)$$

for some gain g and phase lag ϕ, which we now determine by making a complex replacement.

The right hand side of (7) involves the sine function, so one natural choice would be to regard it as the imaginary part of a complex equation. This would work, but we should also keep in mind that the input signal is $B\cos(\omega t)$. For that reason, we will write (7) as the *real* part of a complex equation, using the identity $\mathrm{Re}\,(ie^{i\omega t}) = -\sin(\omega t)$. The equation (7) is thus the real part of

(8)
$$m\ddot{z} + b\dot{z} + kz = bi\omega Be^{i\omega t}\,.$$

and the complex input signal is $Be^{i\omega t}$ (since this has real part $B\cos(\omega t)$).

The sinusoidal system response x_p of (7) is the real part of the exponential system response z_p of (8). The Exponential Response Formula gives

$$z_p = \frac{bi\omega}{p(i\omega)}Be^{i\omega t}$$

where

$$p(s) = ms^2 + bs + k$$

is the characteristic polynomial.

The complex gain is the complex number $W(i\omega)$ by which you have to multiply the complex input signal to get the exponential system response. Comparing z_p with $Be^{i\omega t}$, we see that

$$W(i\omega) = \frac{bi\omega}{p(i\omega)}\,.$$

As usual, write

$$W(i\omega) = ge^{-i\phi}$$

so that

$$z_p = W(i\omega)Be^{i\omega t} = gBe^{i(\omega t - \phi)}$$

Thus

$$x_p = \mathrm{Re}\,(z_p) = gB\cos(\omega t - \phi)$$

—the amplitude of the sinusoidal system response is g times that of the input signal, and lags behind the input signal by ϕ radians.

We can make this more explicit. For notational convenience, let's use the natural frequency $\omega_n = \sqrt{k/m}$. Then

$$p(i\omega) = m(i\omega)^2 + bi\omega + m\omega_n^2 = m(\omega_n^2 - \omega^2) + bi\omega\,,$$

so

$$W(i\omega) = \frac{bi\omega}{m(\omega_n^2 - \omega^2) + bi\omega}\,.$$

Thus the gain $g(\omega) = |W(i\omega)|$ and the phase lag $\phi = -\mathrm{Arg}(W(i\omega))$ are determined as the polar coordinates of the complex function of ω

given by $W(i\omega)$. As ω varies, $W(i\omega)$ traces out a curve in the complex plane, shown by invoking the [Nyquist plot] in the applet. To understand this curve, divide numerator and denominator in the expression for $W(i\omega)$ by $bi\omega$, and rearrange:

$$W(i\omega) = \left(1 - \frac{i}{b/m} \frac{\omega_n^2 - \omega^2}{\omega}\right)^{-1}.$$

As ω goes from 0 to ∞, $(\omega_n^2 - \omega^2)/\omega$ goes from $+\infty$ to $-\infty$, so the expression inside the brackets follows the vertical straight line in the complex plane with real part 1, moving upwards. As z follows this line, $1/z$ follows a circle of radius $1/2$ and center $1/2$, traversed clockwise (exercise!). It crosses the real axis when $\omega = \omega_n$.

This circle is the "Nyquist plot." It shows that the gain starts small, grows to a maximum value of 1 exactly when $\omega = \omega_n$ (in contrast to the spring-driven situation, where the resonant peak is not exactly at ω_n and can be either very large or non-existent depending on the strength of the damping), and then falls back to zero. Near resonance occurs at $\omega_r = \omega_n$.

The Nyquist plot also shows that $-\phi = \text{Arg}(W(i\omega))$ moves from near $\pi/2$ when ω is small, through 0 when $\omega = \omega_n$, to near $-\pi/2$ when ω is large.

And it shows that these two effects are linked to each other. Thus a narrow resonant peak corresponds to a rapid sweep across the far edge of the circle, which in turn corresponds to an abrupt phase transition from $-\phi$ near $\pi/2$ to $-\phi$ near $-\pi/2$.

16.3. **Second order frequency response using damping ratio.** As explained in Section 15, it is useful to write a second order system with sinusoidal driving term as

(9) $$\ddot{x} + 2\zeta\omega_n\dot{x} + \omega_n^2 x = a\cos(\omega t).$$

The constant ω_n is the "natural frequency" of the system and ζ is the "damping ratio." In this abstract situation, we regard the right hand side, $a\cos(\omega t)$, as the input signal, and x as the system response. This corresponds to the spring-mass-dashpot system in which a force is applied directly to the mass, and is illustrated in the Mathlet `Amplitude and Phase: Second Order IV`.

The best path to the solution of (9) is to view it as the real part of the complex equation

(10) $$\ddot{z} + 2\zeta\omega_n\dot{z} + \omega_n^2 z = ae^{i\omega t}.$$

The Exponential Response Formula of Section 10 tells us that unless $\zeta = 0$ and $\omega = \omega_n$ (in which case the equation exhibits resonance, and has no periodic solutions), this has the particular solution

$$(11) \qquad z_p = a \, \frac{e^{i\omega t}}{p(i\omega)}$$

where $p(s) = s^2 + 2\zeta\omega_n s + \omega_n^2$ is the characteristic polynomial of the system. In Section 10 we wrote $W(s) = 1/p(s)$, so this solution can be written

$$z_p = aW(i\omega)e^{i\omega t} \, .$$

The complex valued function of ω given by $W(i\omega)$ is the **complex gain**. We will see now how, for fixed ω, this function contains exactly what is needed to write down a sinusoidal solution to (9).

As in Section 10 we can go directly to the expression in terms of amplitude and phase lag for the particular solution to (9) given by the real part of z_p as follows. Write the polar expression (as in Section 6) for the complex gain as

$$(12) \qquad W(i\omega) = \frac{1}{p(i\omega)} = ge^{-i\phi} \, .$$

so that

$$g(\omega) = |W(i\omega)| \, , \quad -\phi(\omega) = \mathrm{Arg}(W(i\omega))$$

Then

$$z_p = ag \, e^{i(\omega t - \phi)}, \qquad x_p = ag\cos(\omega t - \phi),$$

The particular solution x_p is the *only* periodic solution to (9), and, assuming $\zeta > 0$, any other solution differs from it by a transient. This solution is therefore the most important one; it is the "steady state" solution. It is sinusoidal, and hence determined by just a few parameters: its angular frequency, which is the angular frequency of the input signal; its amplitude, which is g times the amplitude a of the input signal; and its phase lag ϕ relative to the input signal.

We want to understand how g and ϕ depend upon the driving frequency ω. The gain is given by

$$(13) \qquad g(\omega) = \frac{1}{|p(i\omega)|} = \frac{1}{\sqrt{(\omega_n^2 - \omega^2)^2 + 4\zeta^2\omega_n^2\omega^2}}.$$

Figure 8 shows the graphs of gain against the angular frequency of the signal for $\omega_n = 1$ and several values of the damping ratio ζ (namely $\zeta = 1/(4\sqrt{2}), 1/4, 1/(2\sqrt{2}), 1/2, 1/\sqrt{2}, 1, \sqrt{2}, 2$.) As you can see, the gain may achieve a maximum. This occurs when the square of the denominator in (13) is minimal, and we can discover where this
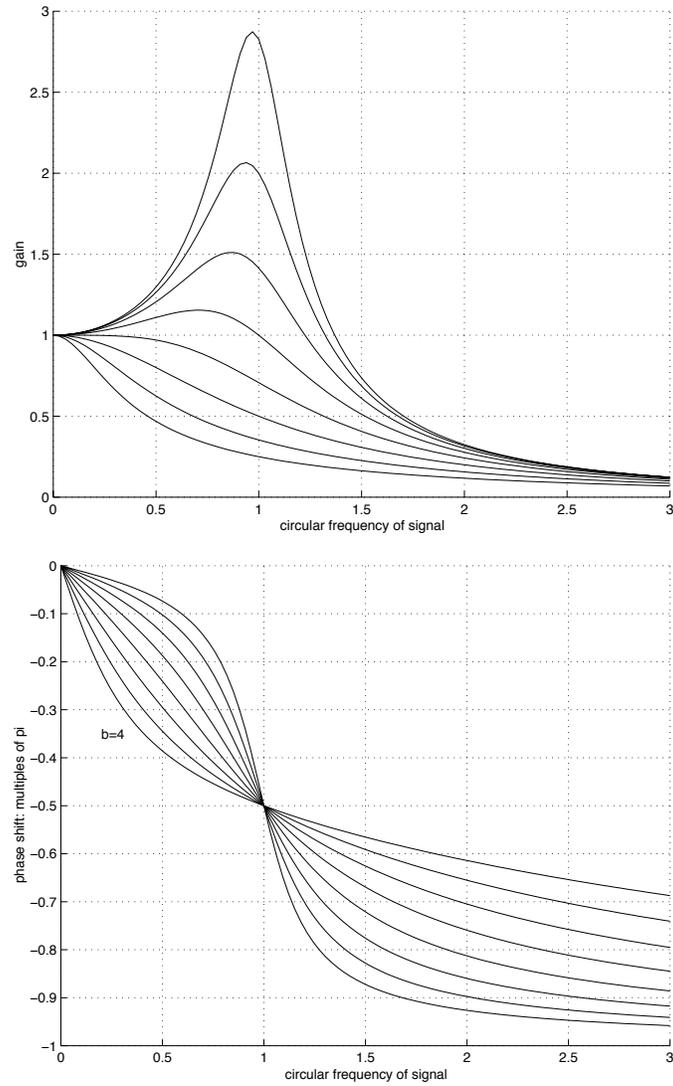
FIGURE 8. Second order frequency response curves

is by differentiating with respect to ω and setting the result equal to zero:

$$(14) \qquad \frac{d}{d\omega}\left((\omega_n^2 - \omega^2)^2 + 4\zeta^2\omega_n^2\omega^2\right) = -2(\omega_n^2 - \omega^2)2\omega + 8\zeta^2\omega_n^2\omega,$$

and this becomes zero when ω equals the **resonant frequency**

$$(15) \qquad \qquad \omega_r = \omega_n\sqrt{1 - 2\zeta^2}.$$

When $\zeta = 0$ the gain becomes infinite at $\omega = \omega_n$: this is **true resonance**. As ζ increases from zero, the maximal gain of the system occurs at smaller and smaller frequencies, till when $\zeta = 1/\sqrt{2}$ the maximum occurs at $\omega = 0$. For still larger values of ζ, the only maximum in the gain curve occurs at $\omega = 0$. When ω takes on a value at which the gain is a local maximum we have **practical resonance**.

We also have the phase lag to consider: the periodic solution to (9) is

$$x_p = ga\cos(\omega t - \phi).$$

Returning to (12), ϕ is given by the argument of the complex number

$$p(i\omega) = (\omega_n^2 - \omega^2) + 2i\zeta\omega_n\omega\,.$$

This is the angle counterclockwise from the positive x axis of the ray through the point $(\omega_n^2 - \omega^2, 2\zeta\omega_n\omega)$. Since ζ and ω are nonnegative, this point is always in the upper half plane, and $0 \le \phi \le \pi$. The phase response graphs for $\omega_n = 1$ and several values of ζ are shown in the second figure.

When $\omega = 0$, there is no phase lag, and when ω is small, ϕ is approximately $2\zeta\omega/\omega_n$. $\phi = \pi/2$ when $\omega = \omega_n$, independent of the damping rato ζ: when the signal is tuned to the natural frequency of the system, the phase lag is $\pi/2$, which is to say that the time lag is one-quarter of a period. As ω gets large, the phase lag tends towards π: strange as it may seem, the sign of the system response tends to be opposite to the sign of the signal.

Engineers typically have to deal with a very wide range of frequencies. In order to accommodate this, and to show the behavior of the frequency response more clearly, they tend to plot $\log_{10}|1/p(i\omega)|$ and the argument of $1/p(i\omega)$ against $\log_{10}\omega$. These are the so-called **Bode plots**.

The expression $1/p(i\omega)$, as a complex-valued function of ω, contains complete information about the system response to periodic input signals. If you let ω run from $-\infty$ to ∞ you get a curve in the complex plane called the **Nyquist plot**. In cases that concern us we may restrict attention to the portion parametrized by $\omega > 0$. For one thing, the characteristic polynomial $p(s)$ has real coefficients, which means that $p(-i\omega) = p(\overline{i\omega}) = \overline{p(i\omega)}$ and so $1/p(-i\omega)$ is the complex conjugate of $1/p(i\omega)$. The curve parametrized by $\omega < 0$ is thus the reflection of the curve parametrized by $\omega > 0$ across the real axis.