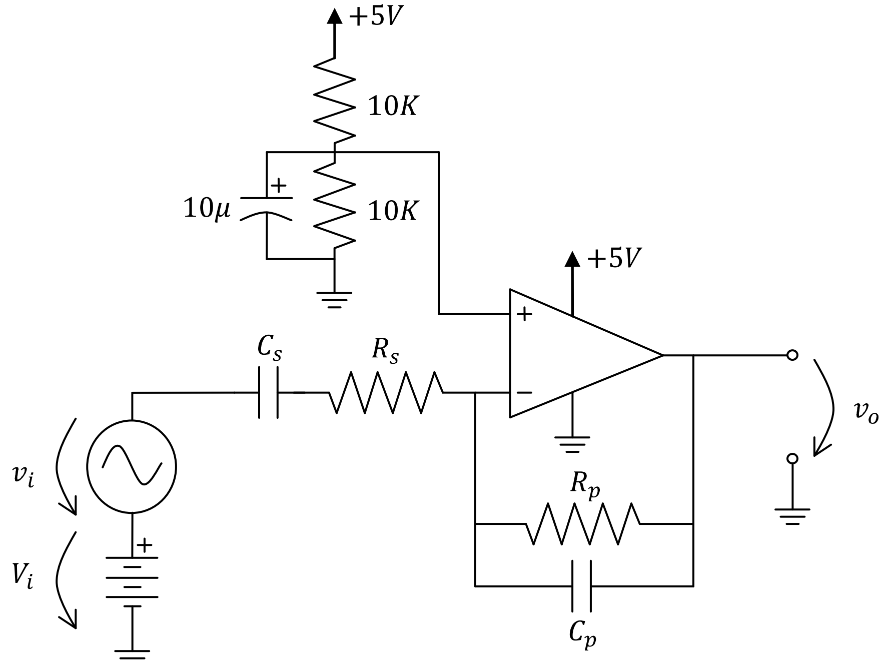
Frequency response of a general purpose single-sided OpAmp amplifier

One configuration for a general purpose amplifier using an operational amplifier is the following.



The circuit is characterized by:

- the op amp is powered by a single dc voltage, hence, single-sided;
- the non-inverting terminal, labeled with a "+" sign, is wired to a voltage divider which provides one-half the dc power supply voltage;
- the signal voltage is the sum of a constant "average" voltage V_i and a variable "small-signal" voltage v_i;
- the output voltage consists of a variable voltage v_o only;
- the input circuit leading to the inverting terminal, labeled with a "−" sign, consists of capacitor C_s in series with resistor R_s; and
- feedback is through capacitor C_p in parallel with resistor R_p.

The four principal assumptions are:

- The input impedance of the op amp is infinite. That is, the resistance between the inverting and non-inverting terminals of the op amp is sufficiently high that the two terminals can be considered to be unconnected. This is a good assumption for modern op amps, whose input impedance is usually $10^{13}\Omega$ or more;

- Neither the inverting terminal nor the non-inverting terminal sink or source any current. This is a good assumption for modern op amps, whose leakage currents are usually 10pA or less;

- The open-loop gain A_{OL} is infinite. The voltage at the output terminal of an ideal op amp is the open-loop gain multiplied by the difference in voltage between the inverting and non-inverting terminals. This is a pretty good assumption for modern op amps, whose open-loop gain is usually 100dB or more.

(Aside: Decibels)

When a number is expressed in dB (decibels), it is a ratio of two similar quantities – usually, but not always, voltages. The open-loop gain measures the ratio of the output voltage divided by the difference between the voltages at the two input terminals, as shown in the following formula:

$$A_{OL} = 20\log_{10}\left(\frac{V_{out}}{V_+ - V_-}\right)$$

An open-loop gain of 100dB is equivalent to amplification by a factor of 100,000. To see this, follow:

$$\begin{array}{lcl} & V_{out} & = 100{,}000 \times (V_+ - V_-) \\ \rightarrow & \dfrac{V_{out}}{V_+ - V_-} & = 100{,}000 = 10^5 \\ \rightarrow & \log_{10}\left(\dfrac{V_{out}}{V_+ - V_-}\right) & = \log_{10}(10^5) = 5 \\ \rightarrow & 20\log_{10}\left(\dfrac{V_{out}}{V_+ - V_-}\right) & = 20 \times 5 = 100 \\ \rightarrow & A_{OL} & = 100 \end{array}$$

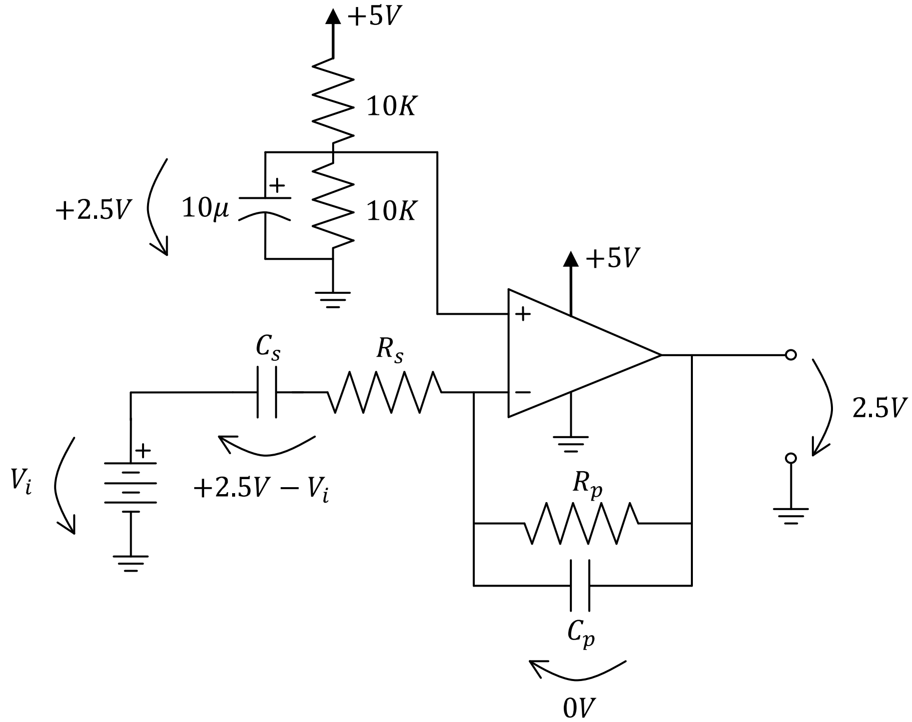
- The open-loop gain applies to any difference in voltage between the inverting and non-inverting terminals, no matter how small. This assumption is not so good, even for modern op amps. In practice, a small voltage difference between the inverting and non-inverting terminals will not be sensed by an op amp. This small voltage difference is called the “input offset voltage”. A typical input offset voltage is 150μV, but special purpose op amps are available with smaller offsets.

The dc analysis

It is useful to divide the analysis of the circuit into two parts: (i) the condition of the circuit when the driving voltage is constant (in which case, $v_i = 0$); and (ii) the variation from analysis (i) as the circuit responds to a small amplitude single frequency sinusoidal voltage. Many circuits, including this one, are “linear” in the sense that the two responses can be added together. The first analysis is called the “dc analysis” and the solution it gives is called the “operating point” of the circuit. The second analysis is called the “ac analysis”.

In many cases, including this one, the objective is to amplify only the varying part of the input signal. In the specified circuit, two resistor-capacitor pairs have been introduced in order that certain frequencies in the input signal are treated differently.

When power is initially applied to the specified circuit, there will be a number of short-lived, or “transient” currents, while the capacitors charge up and the internal circuitry of the op amp stabilizes. It is likely that the input voltages v_i and V_i are produced by other circuitry which will also require some amount of time to reach a steady, non-changing condition. If there is no varying component in the input voltage waveform, then, once the circuit stabilizes, the voltages at all points in the circuit will be constant and, if any currents flow, they will be constant and consistent with the constant voltages. These dc voltages and currents can be written down by inspection, as follows.



The two 10K resistors divide the +5V power supply voltage in half. Since negligible current flows into the non-inverting terminal of the op amp, the two 10K resistors form a perfect voltage-divider. The 10μ capacitor guards this reference voltage against rapid changes in the supply voltage, but serves no useful purpose in the dc case.

The use of +5V for the power supply is for illustration only and does not limit the generality of the analysis. Nor is there any magic to using 10K resistors. Other values can be used, so long as: (i) they are equal; and (ii) they are neither so small nor so large that the current which does flow into the non-inverting terminal (albeit very small) cannot be ignored.

When power is first applied to the circuit, the op amp conducts for a short period, and charges up capacitor C_s. After that short burst of current, no further current flows. Capacitor C_s is charged up to a voltage equal to 2.5V less the dc component V_i of the input voltage waveform.

Since the inverting and non-inverting terminals have the same voltage (2.5V), the op amp does not perform any "amplification" *per se*. It simply maintains the voltage at its output terminal at 2.5V. If there is a load attached to the output terminal, the op amp will deliver into the load whatever current is needed to keep the output voltage at 2.5V.

One could ask: why is the output voltage not zero? After all, since the difference between the voltages at the two input terminals is zero, and since the op amp "multiplies" this difference, the output voltage should also be zero, right?

No. The "multiplier" of the op amp is its <u>open-loop</u> gain. Open-loop means that nothing is connected between the output terminal and the two input terminals. In our case, there is a connection, by means of resistor R_p and capacitor C_p.

Imagine for a moment that the output voltage was higher than the voltage at the inverting terminal. Current would flow through resistor R_p. Since the inverting terminal neither sources nor sinks any current, this current would have nowhere to go except through resistor R_s and, then, into capacitor C_s. The voltage over capacitor C_s would increase. The net effect is that the voltage at the inverting terminal would rise. There would then be a difference in voltage between the two input terminals. The op amp would react to the voltage difference. Because this connection is at the inverting terminal (and not the non-inverting terminal), the op amp would react by lowering the voltage at its output terminal. The feedback is negative, in the direction that will reduce the voltage at the inverting terminal. The higher the open-loop gain of the op amp, the more strongly / quickly this correction would take place.

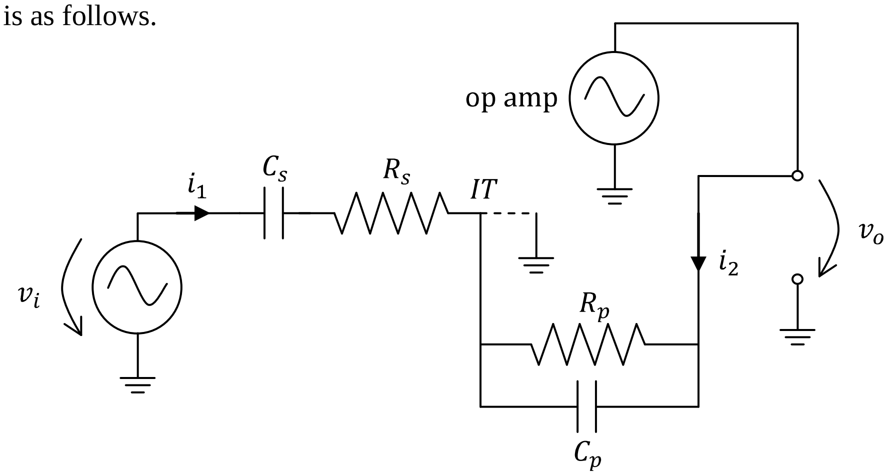
The only condition which is stable is the one where all three terminals of the op amp are at the same voltage potential.

In this condition, no current flows through the feedback resistor R_p. The voltage drops over both R_p and the feedback capacitor C_p are zero.

The ac circuit

We can call the voltages and currents in the dc analysis the "operating point" of the circuit. Now, let us set aside these steady voltages and currents and consider only variations from them which arise when the varying component v_i of the input voltage waveform is non-zero.

The ac circuit is as follows.



The op amp has been replaced by two components: (i) its inverting terminal, labeled IT; and (ii) a voltage source at its output terminal. The output voltage of that voltage source is v_o which is therefore equal to the voltage over the load. The objective of the ac analysis is to determine the output voltage waveform v_o as a function of the signal input voltage waveform v_i.

Note that the details of the op amp's voltage source have not been specified. It is enough for our purposes to assume that the op amp will respond strongly / quickly enough that the following characteristic of the inverting terminal holds true.

The inverting terminal is a "virtual ground". When the op amp is in operation, it will produce whatever output voltage v_o is required to keep the voltage at the inverting terminal equal to zero. In this sense, the inverting terminal is at ground potential. However, no current flows into or out of the inverting terminal, so it is not a "galvanic" ground.

It is the phrase "produce whatever" that allows us to avoid inquiring into the details of the op amp's circuitry. In reality, the op amp's response is neither instantaneous nor unlimited. Both of these practical limitations mean that the voltage at the inverting terminal will likely never be exactly at ground potential. If this lag at the inverting terminal becomes important to the performance of the circuit, then a more complicated model of the op amp's output voltage source would be needed. We will not delve into this.

Because the inverting terminal is at ground potential, the voltage and current on the input side of the op amp are related by Ohm's Law, which we can write as follows:

$$v_i = i_1 \times (R_s \text{ in series with } C_s) \quad (1)$$

Similarly, the voltage and current in the feedback branch around the op amp are also related by Ohm's Law, as follows:

$$v_o = i_2 \times \left(R_p \text{ in parallel with } C_p\right) \quad (2)$$

Because the inverting terminal is a virtual ground, neither taking nor supplying current, the currents in the two branches must be equal and opposite:

$$i_1 = -i_2 \quad (3)$$

We can combine the three previous equations into one:

$$\frac{v_o}{v_i} = -\frac{R_p \text{ in parallel with } C_p}{R_s \text{ in series with } C_s} \quad (4)$$

(Aside: Capacitive reactance as a differential equation)

The expressions "R_p in parallel with C_p" and "R_s in series with C_s" need to be interpreted in terms of reactance and, specifically, capacitive reactance. With the right adjustments, Ohm's Law can be extended to apply to capacitors as well as resistors.

Resistors R_p and R_s impede, or resist, the flow of current in a way which is directly and proportionally related to the voltage over them. Capacitors C_p and C_s also resist the flow of current, but in a more complicated way. Their resistance to the flow of current, or "reactance", differs from the resistors' resistance in two important respects: (i) it depends on frequency; and (ii) it is "out-of-phase" with the applied voltage.

At any instant in time, the charge q stored in a capacitor is equal to its capacitance C (which is assumed in this analysis to be constant) multiplied by the voltage difference v between its plates. That is:

$$q(t) = Cv(t)$$

By definition, the current i flowing into the capacitor is the amount of charge which flows per unit of time. If the flow of charge is not constant, then the current changes with time and should be written as the derivative of the charge, thus:

$$i(t) = \frac{dq(t)}{dt}$$

If the capacitance C is constant, then the instantaneous current can be written in terms of the instantaneous voltage as:

$$\begin{aligned} i(t) &= \frac{d[Cv(t)]}{dt} \\ &= C\frac{dv(t)}{dt} \end{aligned}$$

This is the customary form for the instantaneous voltage-current characteristic of a capacitor. Note that there is an implied sign convention here, in which the current $i(t)$ is taken to be positive when it is flowing <u>into</u> the capacitor, making the voltage $v(t)$ more positive. Also note that this equation applies to all arbitrary waveforms, and not just to direct current or to sinusoidal waveforms at a single frequency.

As stated, this is a general equation. Nevertheless, it is useful to examine this equation for a very special case, where the voltage waveform is a sinusoid at some single frequency, say, f, and having amplitude P. In this special case, the voltage can be written using the imaginary number j as:

$$v(t) = Pe^{j2\pi ft}$$

Taking the derivative, we get:

$$i(t) = C\frac{dv(t)}{dt} = C\left(j2\pi fPe^{j2\pi ft}\right)$$

The ratio of the instantaneous voltage to the instantaneous current is therefore equal to:

$$\frac{v(t)}{i(t)} = \frac{Pe^{j2\pi ft}}{C(j2\pi fPe^{j2\pi ft})} = \frac{1}{j2\pi fC}$$

This ratio is a constant for any given frequency f. Furthermore, the ratio of voltage to current is what Ohm's Law equates to the resistance of a resistor. A resistor's resistance R is comparable to a capacitor's reactance $1/j2\pi fC$. The two quantities can be combined in series or in parallel, using imaginary numbers, in the same way as one combines resistors in series or in parallel. It is the imaginary number j which keeps track of the "phase" between the voltage waveform and the current waveform.

Of course, this handy relationship only applies when the voltage and the current can be expressed as single-frequency sinusoidal waveforms. Fortunately, things can very often be arranged so that this condition is met.

1. All practical waveforms can be written as the sum of sine and cosine waveforms, all of which are at multiples of some fundamental frequency f_0, in the following manner:

$$\begin{aligned} v_i = \; & C_0 + C_1\cos f_0t + C_2\cos 2f_0t + C_3\cos 3f_0t + C_4\cos 4f_0t + \cdots \\ & \cdots + D_1\sin f_0t + D_2\sin 2f_0t + D_3\sin 3f_0t + D_4\sin 4f_0t + \cdots \end{aligned}$$

   It is therefore possible to separate an arbitrary waveform into its sine and cosine components and then to examine each component separately. Dividing an arbitrary waveform up into its components is a "Fourier" analysis. This is particularly useful when the arbitrary waveform is periodic, and repeats itself over and over.

2. All practical circuits will show a "transient" response when there is any change in the driving voltage or current. Often, the magnitude of the transient will eventually decrease to zero, leaving a periodic, or "steady-state", component. This is not to say that the transient will die out quickly (how quickly it dies out, if at all, depends on the circuit and its components) or that the transient is not important (lots of circuits are designed to take advantage of the transient or fail if an unexpected transient persists). In many cases, the interesting activity of the circuit is what it does after the transient has died out.

The steady-state ac analysis

The transfer function v_o/v_i in Equation (4) lends itself to a steady-state analysis. Let us assume: (i) that the input voltage waveform v_i is a simple sinusoid at frequency f, say, with amplitude V_i; and (ii) that the transient response has died out. (Is it easy to assume that all transients eventually die out, but it is an assumption that must be confirmed before a circuit can be safely used. An unexpected transient response can cause unacceptable feedback or worse.)

The reactance "R_p in parallel with C_p" can be written using the same formula which combines two resistors in parallel, namely:

$$\begin{aligned} R_p \text{ in parallel with } C_p &= \frac{1}{\frac{1}{R_p} + \frac{1}{\left(\frac{1}{j2\pi f C_p}\right)}} \\ &= \frac{1}{\frac{1}{R_p} + j2\pi f C_p} \\ &= \frac{R_p}{1 + j2\pi f R_p C_p} \qquad (5) \end{aligned}$$

The reactance "R_s in series with C_s" can be written using the same formula which combines two resistors in series, namely:

$$\begin{aligned} R_s \text{ in series with } C_s &= R_s + \frac{1}{j2\pi f C_s} \\ &= \frac{1 + j2\pi f R_s C_s}{j2\pi f C_s} \qquad (6) \end{aligned}$$

Then, the transfer function in Equation (4) can be expanded as:

$$\begin{aligned} \frac{v_o}{v_i} &= -\frac{\frac{R_p}{1 + j2\pi f R_p C_p}}{\frac{1 + j2\pi f R_s C_s}{j2\pi f C_s}} \\ &= -\left(\frac{R_p}{1 + j2\pi f R_p C_p}\right)\left(\frac{j2\pi f C_s}{1 + j2\pi f R_s C_s}\right) \\ &= -\left(\frac{R_p}{1 + j2\pi f R_p C_p}\right)\left(\frac{1 - j2\pi f R_p C_p}{1 - j2\pi f R_p C_p}\right)\left(\frac{j2\pi f C_s}{1 + j2\pi f R_s C_s}\right)\left(\frac{1 - j2\pi f R_s C_s}{1 - j2\pi f R_s C_s}\right) \end{aligned}$$

Multiplying the numerator and denominator of each term by the conjugate of the denominator reduces the denominators to strictly real numbers. We get:

$$\begin{aligned}
\frac{v_o}{v_i} &= -j2\pi f R_p C_s \left(\frac{1 - j2\pi f R_p C_p}{1 + 4\pi^2 f^2 R_p^2 C_p^2}\right)\left(\frac{1 - j2\pi f R_s C_s}{1 + 4\pi^2 f^2 R_s^2 C_s^2}\right) \\
&= -j2\pi f R_p C_s \frac{1 - j2\pi f R_p C_p - j2\pi f R_s C_s - 4\pi^2 f^2 R_p C_p R_s C_s}{\left(1 + 4\pi^2 f^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f^2 R_s^2 C_s^2\right)} \\
&= -j2\pi f R_p C_s \frac{\left(1 - 4\pi^2 f^2 R_p C_p R_s C_s\right) - j2\pi f\left(R_p C_p + R_s C_s\right)}{\left(1 + 4\pi^2 f^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f^2 R_s^2 C_s^2\right)} \\
&= -2\pi f R_p C_s \frac{2\pi f\left(R_p C_p + R_s C_s\right) + j\left(1 - 4\pi^2 f^2 R_p C_p R_s C_s\right)}{\left(1 + 4\pi^2 f^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f^2 R_s^2 C_s^2\right)} \qquad (7)
\end{aligned}$$

The transfer function has been written as an imaginary number in standard form, that is, as $A + jB$, where both A and B are strictly real numbers. This form can be converted into exponential form, as:

$$\frac{v_o}{v_i} = \sqrt{A^2 + B^2}\, e^{j \tan^{-1}\left(\frac{B}{A}\right)}$$

where

$$\begin{aligned}
\sqrt{A^2 + B^2} &= \frac{2\pi f R_p C_s}{\left(1 + 4\pi^2 f^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f^2 R_s^2 C_s^2\right)} \sqrt{\left[2\pi f\left(R_p C_p + R_s C_s\right)\right]^2 + \left[1 - 4\pi^2 f^2 R_p C_p R_s C_s\right]^2} \\
&= \frac{2\pi f R_p C_s}{\left(1 + 4\pi^2 f^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f^2 R_s^2 C_s^2\right)} \sqrt{\begin{array}{c} 4\pi^2 f^2\left(R_p^2 C_p^2 + 2R_p C_p R_s C_s + R_s^2 C_s^2\right) + \cdots \\ \cdots + 1 - 8\pi^2 f^2 R_p C_p R_s C_s + 16\pi^4 f^4 R_p^2 C_p^2 R_s^2 C_s^2 \end{array}} \\
&= \frac{2\pi f R_p C_s}{\left(1 + 4\pi^2 f^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f^2 R_s^2 C_s^2\right)} \sqrt{1 + 4\pi^2 f^2\left(R_p^2 C_p^2 + R_s^2 C_s^2\right) + 16\pi^4 f^4 R_p^2 C_p^2 R_s^2 C_s^2} \\
&= \frac{2\pi f R_p C_s}{\left(1 + 4\pi^2 f^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f^2 R_s^2 C_s^2\right)} \sqrt{\left(1 + 4\pi^2 f^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f^2 R_s^2 C_s^2\right)} \\
&= \frac{2\pi f R_p C_s}{\sqrt{\left(1 + 4\pi^2 f^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f^2 R_s^2 C_s^2\right)}} \qquad (8)
\end{aligned}$$

and

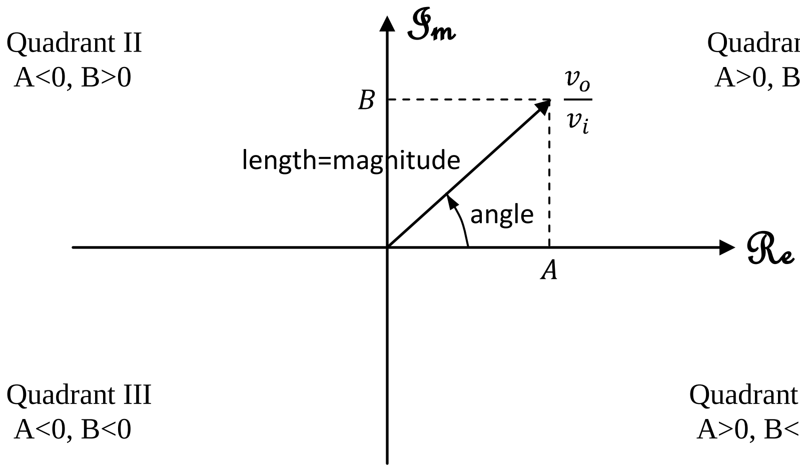
$$\tan^{-1}\left(\frac{B}{A}\right) = \tan^{-1}\left[\frac{-\left(1 - 4\pi^2 f^2 R_p C_p R_s C_s\right)}{-2\pi f\left(R_p C_p + R_s C_s\right)}\right] \qquad (9)$$

Note that both A and B are negative numbers. The minus signs in the argument in Equation (9) have not been cancelled out. Why they have not been cancelled out is explained in the next section.

Physical interpretation of the transfer function

The quantity $\sqrt{A^2 + B^2}$ is the magnitude of the ratio v_o/v_i. If the input voltage waveform v_i is a sinusoid at frequency f then, after any transient has died out, the output voltage waveform will be a sinusoid at the same frequency but with an amplitude $\sqrt{A^2 + B^2}$ times that of the input voltage waveform.

The quantity $\tan^{-1}(B/A)$ is an angle. It is the angle by which the output voltage waveform leads the input voltage waveform. The angle is expressed in radians, which can be converted to degrees by recalling that π radians is equal to $180°$. There is no mathematical restriction on the angle, which can fall anywhere in the circle. The following figure shows a sample transfer function plotted on the imaginary plane.



When calculating the inverse tangent function, it is necessary to check the algebraic sign of A and B to determine exactly which quadrant the angle falls in. If the angle formed by A and B falls in quadrants I or II, then the angle is in the range from $0°$ to $180°$. If the angle formed by A and B falls in quadrants III or IV, then the angle is in the range from $180°$ to $360°$. The algebraic signs of A and B are important because the tangent function is not unique for all angles.

It is customary to express the phase angle in the range from $-180°$ to $+180°$, and then to say that the output voltage waveform "leads" the input voltage waveform when the angle is positive and "lags" the input voltage waveform when the angle is negative. I do not like this custom. It is not physically possible for the output to anticipate the input and so to "lead" it. The output waveform can only follow the input waveform. My preference is always to express the angle as a negative number, which better describes how far the output voltage waveform is behind the input voltage waveform.

The period of the waveform (in seconds) at frequency f (in Hertz) is $1/f$. If the phase angle of the transfer function at this frequency is θ (in the range from $-360°$ to $0°$), then the output voltage waveform is delayed with respect to the input voltage waveform by a time interval of $\theta/360f$ seconds.

The magnitude of the transfer function at extreme values of frequency

Before proceeding, it is useful to examine the magnitude of the transfer function at very low and very high frequencies.

For very small frequencies, when $4\pi^2 f^2 R_p^2 C_p^2 \ll 1$ and $4\pi^2 f^2 R_s^2 C_s^2 \ll 1$, the denominator of the magnitude $\sqrt{A^2 + B^2}$ approaches unity and the magnitude itself approaches $2\pi f R_p C_s$. As the frequency gets even smaller, the numerator will also get smaller, reaching zero as the frequency reaches direct current. Physically, the series capacitor C_s blocks the input waveform at low frequencies.

On the other hand, for frequencies high enough that $4\pi^2 f^2 R_p^2 C_p^2 \gg 1$ and $4\pi^2 f^2 R_s^2 C_s^2 \gg 1$, the denominator of the magnitude $\sqrt{A^2 + B^2}$ approaches $4\pi^2 f^2 R_p C_p R_s C_s$ and the magnitude itself approaches $1/2\pi f C_p R_s$. As the frequency continues to rise, the magnitude will continue to fall, eventually reaching zero. Physically, the parallel capacitor C_p negatively feeds high frequencies back into the op amp's input terminals, cancelling out the input waveform.

The Bode plot

It is difficult to grapple with the magnitude of the transfer function as it is expressed in Equation (8). It is useful to follow in the footsteps of Messr. Bode, whose first recognized that taking the logarithm of both sides of Equation (8) provided useful insight. Recalling the identities that $\log(ST) = \log S + \log T$, that $\log(S/T) = \log S - \log T$ and that $\log(S^T) = T \log S$, we can re-arrange Equation (8) as:

$$\begin{aligned}
\log\left|\frac{v_o}{v_i}\right| &= \log \frac{2\pi f R_p C_s}{\sqrt{(1 + 4\pi^2 f^2 R_p^2 C_p^2)(1 + 4\pi^2 f^2 R_s^2 C_s^2)}} \\
&= \log\big(2\pi f R_p C_s\big) - \tfrac{1}{2}\log\big(1 + 4\pi^2 f^2 R_p^2 C_p^2\big) - \tfrac{1}{2}\log(1 + 4\pi^2 f^2 R_s^2 C_s^2)
\end{aligned}$$

It does not matter in what "base" the logarithms are taken; the relationship is the same in all bases. It is traditional in electronic work to use base ten. It is also traditional to multiply the logarithm of a ratio of voltages by a factor of 20 and to multiply the logarithm of a ratio of (electrical) powers by a factor of ten. Messr. Bode did both, but it is not necessary.

For frequencies which are low enough that both $4\pi^2 f^2 R_p^2 C_p^2 \ll 1$ and $4\pi^2 f^2 R_s^2 C_s^2 \ll 1$, this can be approximated as:

$$\begin{aligned}
\log\left|\frac{v_o}{v_i}\right| &\cong \log\big(2\pi f R_p C_s\big) - \tfrac{1}{2}\log(1) - \tfrac{1}{2}\log(1) \\
&= \log\big(2\pi f R_p C_s\big) - 0 - 0 \\
&= \log\big(R_p C_s\big) + \log(2\pi f) \\
&= Constant_{low} + \log(2\pi f)
\end{aligned}$$

At these low frequencies, there is a linear relationship between $\log|v_o/v_i|$ and $\log(2\pi f)$. If these two logarithms are plotted, the line will be straight with a slope of $+1$.

For frequencies which are high enough that both $4\pi^2 f^2 R_p^2 C_p^2 \gg 1$ and $4\pi^2 f^2 R_s^2 C_s^2 \gg 1$, the logarithm of the transfer function can be approximated as:

$$
\begin{aligned}
\log\left|\frac{v_o}{v_i}\right| &\cong \log\left(2\pi f R_p C_s\right) - \tfrac{1}{2}\log\left(4\pi^2 f^2 R_p^2 C_p^2\right) - \tfrac{1}{2}\log(4\pi^2 f^2 R_s^2 C_s^2) \\
&= \left[\log\left(R_p C_s\right) + \log(2\pi f)\right] - \tfrac{1}{2}\left[\log\left(R_p^2 C_p^2\right) + 2\log(2\pi f)\right] - \tfrac{1}{2}[\log(R_s^2 C_s^2) + 2\log(2\pi f)] \\
&= Constant_{high} + \log(2\pi f) - \log(2\pi f) - \log(2\pi f) \\
&= Constant_{high} - \log(2\pi f)
\end{aligned}
$$

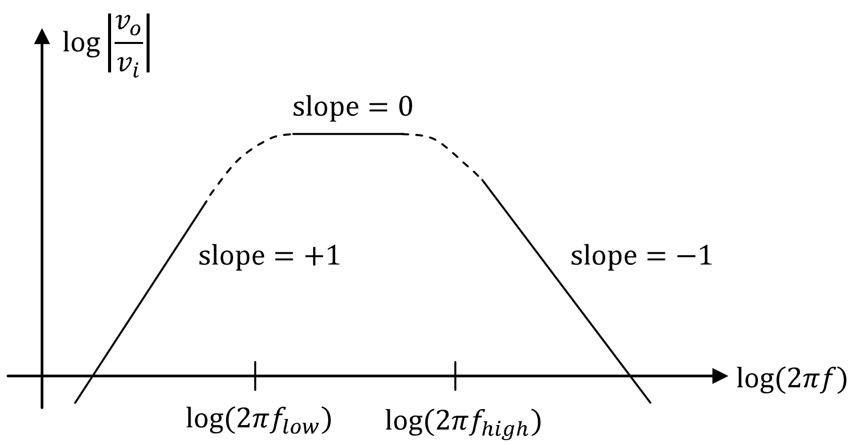
At these high frequencies, there is once again a linear relationship between $\log|v_o/v_i|$ and $\log(2\pi f)$. If these two logarithms are plotted, the line will be straight, but this time, with a slope of -1.

Now, <u>suppose</u> that there is some intermediate frequency or range of frequencies for which $4\pi^2 f^2 R_p^2 C_p^2 \ll 1$ but $4\pi^2 f^2 R_s^2 C_s^2 \gg 1$. In this case, the logarithm of the transfer function can be approximated as:

$$
\begin{aligned}
\log\left|\frac{v_o}{v_i}\right| &\cong \log\left(2\pi f R_p C_s\right) - \tfrac{1}{2}\log(4\pi^2 f^2 R_s^2 C_s^2) - \tfrac{1}{2}\log(1) \\
&= \left[\log\left(R_p C_s\right) + \log(2\pi f)\right] - \tfrac{1}{2}[\log(R_s^2 C_s^2) + 2\log(2\pi f)] - \tfrac{1}{2}\log(1) \\
&= Constant_{intermediate} + \log(2\pi f) - \log(2\pi f) - 0 \\
&= Constant_{intermediate}
\end{aligned}
$$

At these intermediate frequencies, $\log|v_o/v_i|$ is a constant with no dependence on frequency.

Based on these approximations, we can make a rough sketch of the magnitude of the logarithm of the transfer function as a function of the logarithm of the "angular frequency" $2\pi f$, as follows:



This figure identifies two frequencies f_{low} and f_{high} at the two points which divide the three linear regions of the curve. One says that the components of the input voltage waveform v_i which are at frequencies above f_{high} are "attenuated" to a greater extent that frequencies in the range where the slope $= 0$. By attenuated, we mean that those higher frequencies are not passed through into the output voltage waveform v_o. Furthermore, the extent to which the higher frequency components are blocked increases as the frequency increases above f_{high}.

Similarly, components of the input voltage waveform which are at frequencies below f_{low} are attenuated relative to those in the "pass band" and the degree of attenuation increases as the frequency decreases below f_{low}. As a whole, components of the input voltage waveform which have frequencies between f_{low} and f_{high} are selected and passed through to the output.

The low cutoff frequency f_{low} is the "midpoint" of the low frequency transition region (shown as a dashed curve in the figure). Among the candidates which could be used to designate this specific frequency, it is customary to select as f_{low} the frequency defined by:

$$2\pi f_{low} = \frac{1}{R_s C_s} \quad (10)$$

Similarly, the "midpoint" of the high frequency transition region (also shown as a dashed curve in the figure) can be defined by:

$$2\pi f_{high} = \frac{1}{R_p C_p} \quad (11)$$

In a similar way, we could select a frequency $f_{midband}$ to represent the "midpoint" of the pass band between f_{low} and f_{high}. We could use the average of the logarithms, in the sense that:

$$\begin{aligned}
\log(2\pi f_{midband}) \quad &= \tfrac{1}{2}\left[\log(2\pi f_{low}) + \log\left(2\pi f_{high}\right)\right] \\
&= \tfrac{1}{2}\log\left(4\pi^2 f_{low} f_{high}\right) \\
&= \log\left(2\pi\sqrt{f_{low} f_{high}}\right) \\
\rightarrow \qquad 2\pi f_{midband} \quad &= 2\pi\sqrt{f_{low} f_{high}} \quad (12) \\
&= \sqrt{(2\pi f_{low})\left(2\pi f_{high}\right)} \\
&= \frac{1}{\sqrt{R_s C_s R_p C_p}}
\end{aligned}$$

Now, let's look at the (exact) value of the magnitude of the transfer function at these three frequencies:

$$\begin{aligned}
\left|\frac{v_o}{v_i}\right|_{low} \quad &= \frac{2\pi f_{low} R_p C_s}{\sqrt{\left(1 + 4\pi^2 f_{low}^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f_{low}^2 R_s^2 C_s^2\right)}} \\
&= \frac{\left(R_p C_s / R_s C_s\right)}{\sqrt{\left(1 + \frac{R_p^2 C_p^2}{R_s^2 C_s^2}\right)\left(1 + \frac{R_s^2 C_s^2}{R_s^2 C_s^2}\right)}} \\
&= \frac{R_p C_s}{\sqrt{2}\sqrt{R_s^2 C_s^2 + R_p^2 C_p^2}} \quad (13A)
\end{aligned}$$

$$\left|\frac{v_o}{v_i}\right|_{high} = \frac{2\pi f_{high} R_p C_s}{\sqrt{\left(1+4\pi^2 f_{high}^2 R_p^2 C_p^2\right)\left(1+4\pi^2 f_{high}^2 R_s^2 C_s^2\right)}}$$

$$= \frac{\left(R_p C_s / R_p C_p\right)}{\sqrt{\left(1+\frac{R_p^2 C_p^2}{R_p^2 C_p^2}\right)\left(1+\frac{R_s^2 C_s^2}{R_p^2 C_p^2}\right)}}$$

$$= \frac{R_p C_s}{\sqrt{2}\sqrt{R_s^2 C_s^2 + R_p^2 C_p^2}} \quad (13B)$$

So, the magnitude of the transfer function is the same at both cutoff frequencies f_{low} and f_{high}. At the midband frequency:

$$\left|\frac{v_o}{v_i}\right|_{midband} = \frac{2\pi f_{midband} R_p C_s}{\sqrt{\left(1+4\pi^2 f_{midband}^2 R_p^2 C_p^2\right)\left(1+4\pi^2 f_{midband}^2 R_s^2 C_s^2\right)}}$$

$$= \frac{\left(R_p C_s / \sqrt{R_s C_s R_p C_p}\right)}{\sqrt{\left(1+\frac{R_p^2 C_p^2}{R_s C_s R_p C_p}\right)\left(1+\frac{R_s^2 C_s^2}{R_s C_s R_p C_p}\right)}}$$

$$= \frac{\left(R_p C_s / \sqrt{R_s C_s R_p C_p}\right)}{\sqrt{\left(1+\frac{R_p C_p}{R_s C_s}\right)\left(1+\frac{R_s C_s}{R_p C_p}\right)}}$$

$$= \frac{R_p C_s}{\sqrt{\left(R_s C_s + R_p C_p\right)\left(R_p C_p + R_s C_s\right)}}$$

$$= \frac{R_p C_s}{R_s C_s + R_p C_p} \quad (13C)$$

The phase angle at extreme values of frequency

For very small frequencies, when $4\pi^2 f^2 R_p C_p R_s C_s \ll 1$, the numerator in Equation (9) approaches minus one and the phase angle approaches $\tan^{-1}\left[-1/-2\pi f\left(R_p C_p + R_s C_s\right)\right]$. As the frequency gets even smaller, the argument of the inverse tangent function will increase without bound, and the phase angle will asymptotically approach $\tan^{-1}[-1/-0]$, corresponding to $-90°$. For these small frequencies, therefore, the output voltage waveform will lag the input voltage waveform by one-quarter of a period.

On the other hand, for frequencies high enough that $4\pi^2 f^2 R_p C_p R_s C_s \gg 1$, the numerator in Equation (9) approaches $4\pi^2 f^2 R_p C_p R_s C_s$, and the argument of the inverse tangent function approaches $2\pi f R_p C_p R_s C_s / -\left(R_p C_p + R_s C_s\right)$. As the frequency continues to increase, the phase angle will asymptotically approach $\tan^{-1}[+\infty/-1]$, corresponding to $-270°$. For these high frequencies, therefore, the output voltage waveform will lag the input voltage waveform by three-quarters of a period.

It is also useful to look at the phase angle at the midband frequency. Equation (9) becomes:

$$
\begin{aligned}
\tan^{-1}\left(\frac{B}{A}\right)\bigg|_{midband} \quad &= \tan^{-1}\left[\frac{-\left(1-4\pi^2 f_{midband}^2 R_p C_p R_s C_s\right)}{-2\pi f_{midband}\left(R_p C_p + R_s C_s\right)}\right] \\
&= \tan^{-1}\left[\frac{-(1-1)}{-\left(R_p C_p + R_s C_s\right)/\sqrt{R_s C_s R_p C_p}}\right] \\
&= \tan^{-1}\left[\frac{-0}{-1}\right] \\
&= -180^\circ
\end{aligned}
$$

At the midband frequency, the output voltage waveform lags the input voltage waveform by one-half period.

<u>Specifying the midband "gain"</u>

Let's consider typical design objectives. Usually, a circuit like that shown at the outset of this document would be used when it is desirable to exclude certain low frequencies and certain high frequencies from the input voltage waveform. Stated differently, it is desirable to pass only a certain range of frequencies. These two goals help us pick the cutoff frequencies f_{low} and f_{high}. It should be noted from Equations $(13A)$ and $(13B)$ above, that the magnitude of the transfer function is the same at both of these frequencies.

Once the cutoff frequencies f_{low} and f_{high} have been chosen, the relationship between the resistors and their associated capacitors are no longer arbitrary. The sole remaining arbitrary parameter is the product $R_p C_s$. To see this, observe that, once f_{low} and f_{high} have been chosen, the transfer functions at the three key frequencies are:

$$
\begin{aligned}
\left|\frac{v_o}{v_i}\right|_{cutoff} \quad &= \left|\frac{v_o}{v_i}\right|_{low} = \left|\frac{v_o}{v_i}\right|_{high} \\
&= \frac{R_p C_s}{\sqrt{2}\sqrt{R_s^2 C_s^2 + R_p^2 C_p^2}} \\
&= \frac{R_p C_s}{\sqrt{2}\sqrt{\dfrac{1}{4\pi^2 f_{low}^2} + \dfrac{1}{4\pi^2 f_{high}^2}}} \\
&= \frac{\sqrt{2}\pi R_p C_s}{\sqrt{\dfrac{1}{f_{low}^2} + \dfrac{1}{f_{high}^2}}} \qquad (14A)
\end{aligned}
$$

and

$$\left|\frac{v_o}{v_i}\right|_{midband} = \frac{R_pC_s}{R_sC_s + R_pC_p} = \frac{R_pC_s}{\dfrac{1}{2\pi f_{low}} + \dfrac{1}{2\pi f_{high}}} = \frac{2\pi R_pC_s}{\dfrac{1}{f_{low}} + \dfrac{1}{f_{high}}} \quad (14B)$$

The only parameter in Equations $(14A)$ and $(14B)$ is the product R_pC_s.

Usually, a separate design objective is to set the "gain" at the midband frequency. The gain is the ratio of the amplitude of the output voltage waveform to the amplitude of the input voltage waveform. Let us use the symbol G for the desired midband gain. If $G = 1$, then the input and output amplitudes will be the same. If $G > 1$, then the output amplitude will be greater than the input amplitude and the op amp will act as an amplifier. In any event, if G is the desired midband gain, then we must have from Equation $(14B)$:

$$G = \frac{2\pi R_pC_s}{\dfrac{1}{f_{low}} + \dfrac{1}{f_{high}}}$$

$$\rightarrow \quad 2\pi R_pC_s = G\left(\frac{1}{f_{low}} + \frac{1}{f_{high}}\right)$$

$$\rightarrow \quad R_pC_s = \frac{G}{2\pi}\left(\frac{1}{f_{low}} + \frac{1}{f_{high}}\right) \quad (15)$$

Specifying the input impedance

There are four unknown component values in the circuit: R_s, C_s, R_p and C_p. So far, we have looked at three design objectives – f_{low}, f_{high} and G – which relate these component values. Sometimes, but not always, a fourth design objective is specified. If there is no fourth design objective, then the designer may be free to pick one of the components as a starting point.

On the other hand, it is sometimes important that the input impedance be a certain value. This could be the case, for example, if the input voltage waveform is produced by a microphone or other sensor which operates most effectively if the circuit it drives has a given input impedance. We will use the symbol $|Z_i|$ for the input impedance.

We have already looked at the ac circuit and seen that the input impedance is the series combination of resistor R_s and capacitor C_s. The input impedance has already been expressed in Equation (6) as:

$$Z_i = \frac{1 + j2\pi f R_sC_s}{j2\pi f C_s}$$

The magnitude of the impedance is therefore equal to:

$$
\begin{aligned}
|Z_i| &= \left|\frac{1 + j2\pi f R_s C_s}{j2\pi f C_s}\right| \\
&= \left|\frac{1 + j2\pi f R_s C_s}{2\pi f C_s}\right| \\
&= \frac{\sqrt{1 + (2\pi f R_s C_s)^2}}{2\pi f C_s}
\end{aligned}
$$

The input impedance will usually be specified for the midband frequency, so that:

$$
|Z_i| = \frac{\sqrt{1 + (2\pi f_{midband} R_s C_s)^2}}{2\pi f_{midband} C_s}
$$

Substituting the low cutoff frequency, we get:

$$
\begin{aligned}
|Z_i| &= \frac{\sqrt{1 + \left(\frac{2\pi f_{midband}}{2\pi f_{low}}\right)^2}}{2\pi f_{midband} C_s} \\
\rightarrow \quad 2\pi f_{midband} C_s |Z_i| &= \sqrt{1 + \left(\frac{f_{midband}}{f_{low}}\right)^2} \\
\rightarrow \quad C_s &= \frac{\sqrt{1 + \left(\frac{f_{midband}}{f_{low}}\right)^2}}{2\pi f_{midband} |Z_i|} \qquad (16)
\end{aligned}
$$

<u>**A numerical example**</u>

Let’s consider an audio amplifier with the following design objectives.

low cutoff frequency	$f_{low} = 120\text{Hz}$
high cutoff frequency	$f_{high} = 5000\text{Hz}$
midband gain	$G = 25$
input impedance	$\lvert Z_i \rvert = 6\text{K}\Omega$

Step 1: The midband frequency $f_{midband}$ is calculated from Equation (12) as:

$$
\begin{aligned}
f_{midband} &= \sqrt{f_{low} f_{high}} \\
&= \sqrt{120 \times 5000}\text{Hz} \\
&= 775\text{Hz}
\end{aligned}
$$

Step 2: The series capacitor's value C_s is calculated from Equation (16) as:

$$\begin{aligned} C_s &= \frac{\sqrt{1+\left(\frac{f_{midband}}{f_{low}}\right)^2}}{2\pi f_{midband}|Z_i|} \\ &= \frac{\sqrt{1+\left(\frac{775}{120}\right)^2}}{2\pi \times 775 \times 6{,}000}\,\mathrm{F} \\ &= 0.224\mu\mathrm{F} \end{aligned}$$

Step 3: The series resistor's value R_s is calculated from Equation (10) as:

$$\begin{aligned} R_s &= \frac{1}{2\pi f_{low} C_s} \\ &= \frac{1}{2\pi \times 120 \times 0.224 \times 10^{-6}}\,\Omega \\ &= 5.92\mathrm{K}\Omega \end{aligned}$$

Step 4: The parallel resistor's value R_p is calculated from Equation (15) as:

$$\begin{aligned} R_p &= \frac{G}{2\pi C_s}\left(\frac{1}{f_{low}} + \frac{1}{f_{high}}\right) \\ &= \frac{25}{2\pi \times 0.224 \times 10^{-6}}\left(\frac{1}{120} + \frac{1}{5000}\right)\Omega \\ &= 152\mathrm{K}\Omega \end{aligned}$$

Step 5: The parallel capacitor's value C_p is calculated from Equation (11) as:

$$\begin{aligned} C_p &= \frac{1}{2\pi f_{high} R_p} \\ &= \frac{1}{2\pi \times 5000 \times 152{,}000}\,\mathrm{F} \\ &= 209\mathrm{pF} \end{aligned}$$

None of these components is a standard value. Usually, the designer would iterate these calculations using standard component values to determine which combination best meets the design criteria. For our purposes, it will be sufficient to use the following standard values:

$$\begin{aligned} C_s &= 0.22\mu\mathrm{F} \\ R_s &= 6\mathrm{K}\Omega \\ C_p &= 200\mathrm{pF} \\ R_p &= 150\mathrm{K}\Omega \end{aligned}$$

Plots of the transfer function with respect to frequency

Using these standard component values, the magnitude of the transfer function [Equation (8)] becomes:

$$\left|\frac{v_o}{v_i}\right| = \frac{2\pi f R_p C_s}{\sqrt{\left(1 + 4\pi^2 f^2 R_p^2 C_p^2\right)\left(1 + 4\pi^2 f^2 R_s^2 C_s^2\right)}}$$

$$= \frac{0.207345f}{\sqrt{(1 + 3.55307 \times 10^{-8} \times f^2)(1 + 6.87872 \times 10^{-5} \times f^2)}}$$

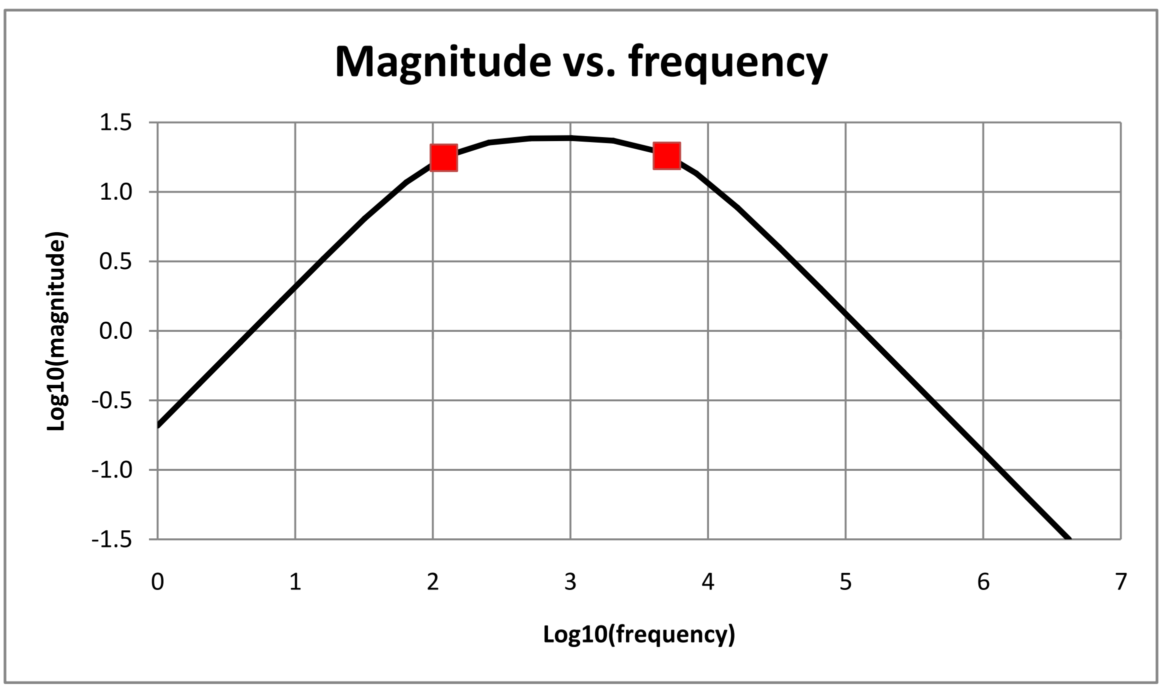
and the phase angle by which the output voltage waveform lags the input voltage waveform [Equation (9)] becomes:

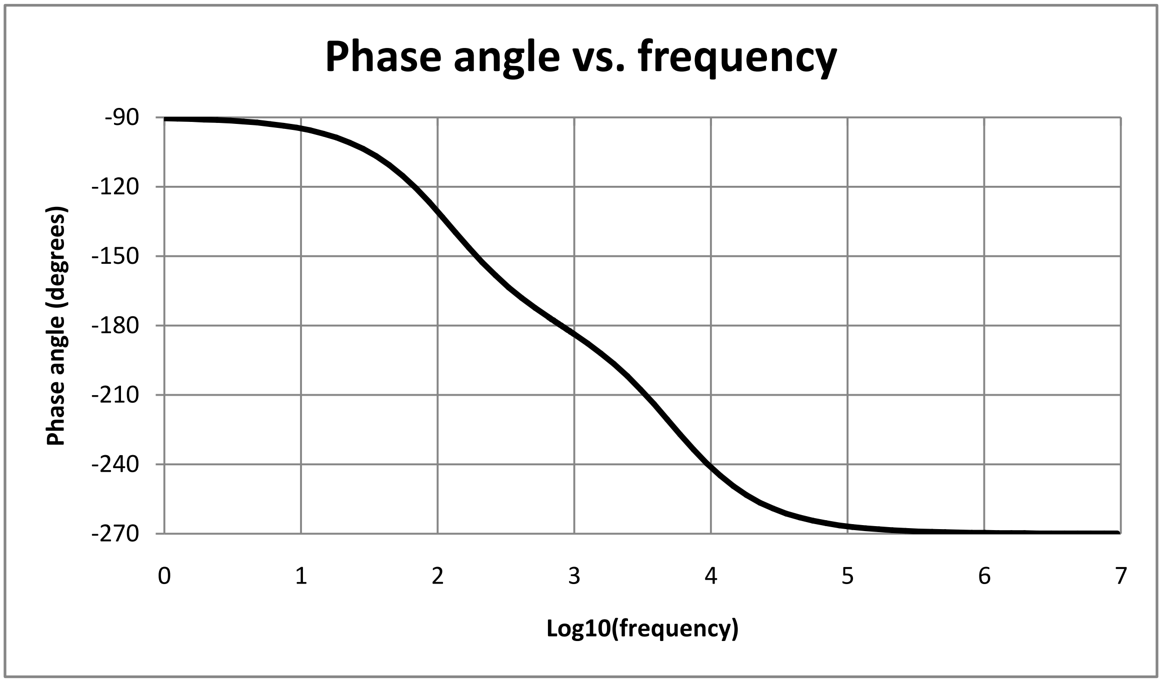
$$\text{angle}\left(\frac{v_o}{v_i}\right) = \tan^{-1}\left[\frac{-\left(1 - 4\pi^2 f^2 R_p C_p R_s C_s\right)}{-2\pi f\left(R_p C_p + R_s C_s\right)}\right]$$

$$= \tan^{-1}\left[\frac{-(1 - 1.56335 \times 10^{-6} \times f^2)}{-0.0084823f}\right]$$

The following two plots show the magnitude and phase, respectively, of the transfer function. In both plots, the horizontal axis is the base-10 logarithm of the frequency. For example, the abscissa 2 corresponds to a frequency of $10^2 = 100$Hz and the abscissa 4 corresponds to a frequency of $10^4 = 10{,}000$Hz. The two squares shown in the magnitude plot are the two cutoff frequencies: 120Hz and 5000Hz.

The vertical scale in the magnitude plot is the base-10 logarithm of the magnitude of the transfer function. For example, the ordinate 0 corresponds to an amplification of $10^0 = 1$ and the ordinate 1 corresponds to an amplification of $10^1 = 10$. The squares shown at the cutoff frequencies have an ordinate of about 1.25, corresponding to an amplification of $10^{1.25} = 18$. At midband, the ordinate is about 1.4, corresponding to an amplification of $10^{1.4} = 25$.





Summary of formulae

R_s and C_s in series at input; R_p and C_p in parallel for feedback

$$2\pi f_{low} = \frac{1}{R_s C_s}$$

$$2\pi f_{high} = \frac{1}{R_p C_p}$$

$$f_{midband} = \sqrt{f_{low} f_{high}}$$

$$Gain = \frac{2\pi R_p C_s}{\frac{1}{f_{low}} + \frac{1}{f_{high}}}$$

$$|Z_i| = \frac{\sqrt{1 + (2\pi f_{midband} R_s C_s)^2}}{2\pi f_{midband} C_s}$$

Jim Hawley
November 2011

An e-mail describing errors and omissions would be appreciated.