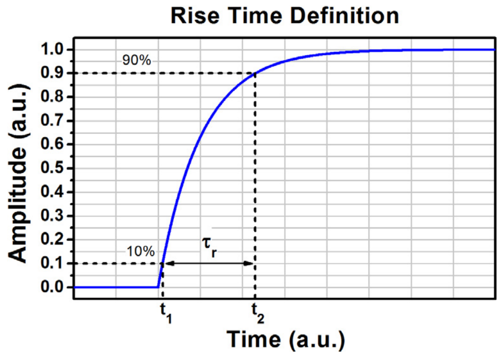
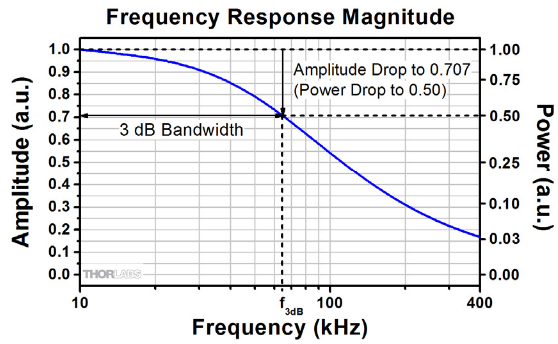
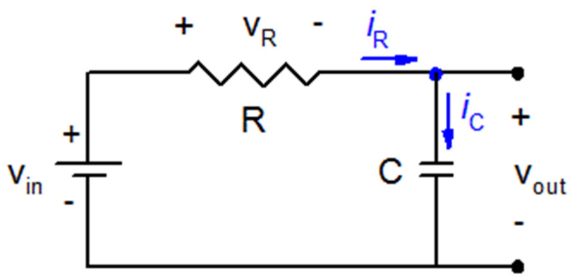
Lab Fact







$$\tau_r \cong \frac{0.35}{f_{3dB}}$$

Rise Time, Frequency Response, and 3 dB Bandwidth

Rise time and 3 dB bandwidth are parameters important for characterizing the performance of many electrical and electro-optical systems. These parameters are often discussed separately, but they have a close relationship that permits the value of one to be calculated when the other is known. This relationship can be advantageous for applications such as predicting and analyzing the distortion of output signals.

In this Lab Fact, general expressions for rise time (τ_r), frequency response, and 3 dB electrical bandwidth (f_{3dB}) were derived. The derivation was based on an RC low-pass filter circuit, which serves as a general model for systems exhibiting low-pass filter behavior. The derived expressions were then used to find a relationship between rise time and 3 dB electrical bandwidth: $\tau_r \cong 0.35/f_{3dB}$. Steps in the derivations included:

- Finding an expression for rise time by considering the dynamic movement of charge in the RC low-pass filter circuit.
- Determining an equation for the 3 dB electrical bandwidth using the transfer function of the circuit.
- Equating the two expressions to find the relationship between the two parameters.

Examples are provided in which rise time or 3 dB bandwidth was measured for photodiode-based systems. The unmeasured parameter was then calculated using the relationship between the two parameters.

April 1, 2019

Table of Contents

1 Introduction

Applications often require a system or device to output a scaled version of an input signal. An example is a current controller that accepts a modulated voltage signal from a function generator. The controller is expected to output a current signal whose modulated amplitude is proportional to that of the input voltage signal. The time-dependent differences between the shapes of the input and output signals are generally referred to as distortion. Distortion can be minimized by ensuring the input signal's parameters, including its electrical frequency range, are within the system's specifications.

A system's electrical frequency range is typically specified in terms of bandwidth, with units of hertz. When the lowest frequency in the range is 0 Hz, the values of the highest frequency and the bandwidth are the same. Assuming an input signal's frequency components are all within the system's bandwidth, the system should respond to changes in the input signal as quickly as they occur, resulting in low-distortion output signals.

If some fraction of the input signal's frequency components exceeds the system's bandwidth, the system's response to this portion of the signal will lag. As a result, the output signal will not accurately reproduce the fastest, most abrupt amplitude transitions of the input signal. This will result in output signal features that may have lower amplitudes, be wider, and have rounder edges than the corresponding features in the input signal.

Minimum rise time and 3 dB cutoff frequency are two parameters used to characterize the upper limits of the response a system can provide to a change in an input signal. This Lab Fact provides an overview of minimum rise time and 3 dB cutoff frequency, derives a mathematical relationship between them, and provides examples of their measurement.

It is important to note that read-out electronics and other equipment interfaced with a device under test contribute to the measured frequency response. Rather than just provide the isolated response of a device under test, the measured frequency response is always the response of the total system.

1.1 Rise Time

Rise time (τ_r) is the length of time (t) required for a signal to transition between two defined points on the rising edge of a curve. Rise time is frequently measured between points that are 10% and 90% up the rising edge of the curve, as indicated in Figure 1.

These points are chosen instead of points 0% and 100% up the rising edge of the curve for a number of reasons. The signal may asymptotically approach the 0% and 100% points, include noise that obscures these points, and/or vary around the points rather than stabilize at them. A measurement between the 10% and 90% points can be easier to make, yield more accurate results, and be more relevant to the application. The full transition time is also likely to be of less interest for applications that do not require the signal to stabilize at specific minimum and maximum amplitudes. For example, it may be sufficient to define a high-state amplitude as any value above a defined threshold.

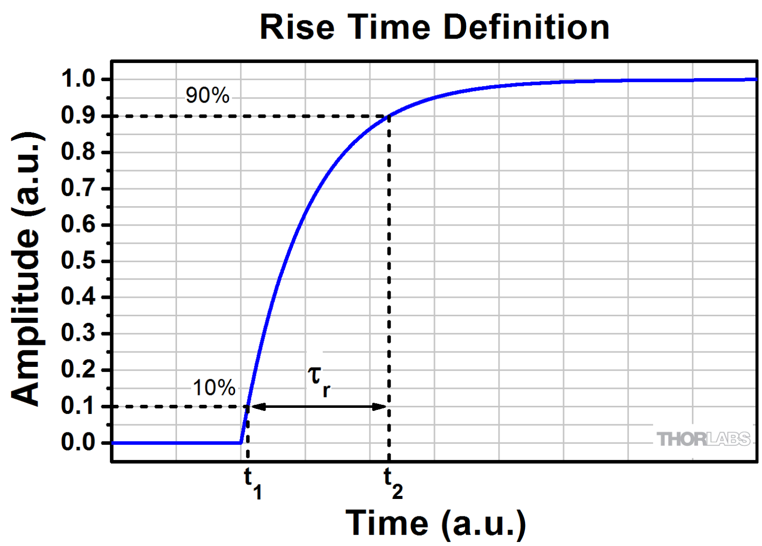


Figure 1 *Rise Time (τ_r) is defined as the time required for a signal to transition between points 10% and 90% up the rising edge of the curve.*

Minimum rise time is a specification of system performance. It is found by applying an input signal called a step function to the system and measuring the rise time of the output signal. A step function is characterized by an amplitude that transitions, more quickly than the system can respond, from being stable at one amplitude to being stable at a higher amplitude. In response to this type of input, the output signal transitions between its corresponding low and high amplitude values with the fastest speed achievable by the system.

1.2 Frequency Responses of Systems and Devices

While it is intuitive to work with signals expressed as functions of time, it is often easier to design systems and perform data analysis when working in the frequency domain. This requires expressing signals as and describing system response as functions of frequency.

A common way to find a frequency-domain representation of a time-domain signal is to perform a Fourier transform. This operation creates a frequency-dependent representation of the signal using a set of basis functions. This approach is similar to using $\hat{x}$, $\hat{y}$, and $\hat{z}$ basis vectors to represent a 3D vector in a Cartesian coordinate system. Like the basis vectors of the Cartesian coordinate system, each basis function in a set used to represent a signal must be orthogonal to the other functions in the set. The set of basis functions must also be complete, meaning that the functions in the set must be sufficient to perfectly represent any finite-energy signal. Since the basis functions are known, only their amplitudes need to be determined. The Fourier transform provides the signal-specific amplitude coefficients for each basis function. The expression for the signal is obtained by multiplying each basis function by the associated amplitude coefficient and then summing all of the resulting functions.

The Fourier transform can be applied to aperiodic, as well as periodic, signals and can be derived from the Fourier series, whose use is restricted to periodic signals. The basis functions for both Fourier transforms and Fourier series can be expressed as a set of sine and cosine functions.[1] Examples of using a truncated Fourier series, as a practical alternative to the infinitely long full representation, to represent periodic rectangular pulse trains are included in the *Pulse Distortion* Lab Fact[2].

As any finite-energy signal can be represented by the sum of a set of discrete sinusoids, each with its own frequency and amplitude, it is useful to describe the frequency response of a system in terms of a range of sinusoids with different oscillation frequencies. A system can alter both the phase and the magnitude of input sinusoids, as illustrated in Figure 2. In addition, each sinusoidal frequency component of the input signal can be affected differently.

A system's phase response describes the relative phase delay, as a function of frequency, that the system adds to the argument of each sinusoidal frequency component of the input signal. A phase delay creates an offset between the oscillations of the input and output sine waves, with the output signal's shifted later in time. When the phase delay provided by the system is not constant over the frequency range of the input signal, the output signal will be distorted. The magnitude response describes the frequency-dependent amplitude changes the system applies to each sinusoidal input signal frequency component. When the amplitude changes are not constant across the frequency range of the input signal, the output signal will be distorted.

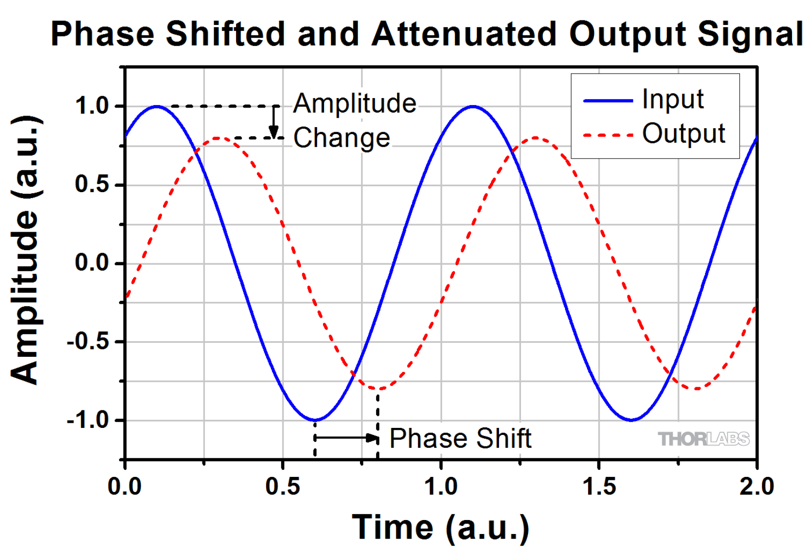


Figure 2 *Illustration of a signal output (red) by a system in response to the input (blue). In this example, points corresponding to particular radial values on the output signal occur at later times than the corresponding points on the input signal. The magnitude of the shift to later times is the phase delay. The system also attenuated the input signal, which reduced the comparative amplitude of the output signal.*

[1] F. Stremler, *Introduction to Communication Systems*, 3rd 3d., Addison-Wesley Publishing Company, New York, 1990, Chapter 2.
[2] Pulse Distortion Lab Fact: https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=11718.

Data describing the frequency response of a system can be used to predict the output signal, when the input signal is known. An instrument's frequency response data can be used to mathematically remove its contribution from the total measured frequency response[3], in order to better isolate the response of the device under test. Frequency response data is sometimes supplied with an instrument and can be found empirically. An example of measuring the magnitude of the frequency response is given in Section 1.2.1.

The frequency response of a system can also be modelled to obtain a single, frequency-dependent mathematical expression capable of relating any signal input to the system with the corresponding output signal. Section 2.2.1 describes an approach for finding a mathematical expression that models the behavior of the system.

1.2.1 Frequency Response Magnitude Measurement Example

Frequency response magnitude data for a system are the frequency-dependent amplitude scaling factors the system effectively applies to the individual frequency components of input signals. The frequency response magnitude can be found using a set of sinusoidal functions as input signals.

Each sine wave in the set of input signals should oscillate at a different constant frequency. The frequencies should span the frequency range of interest, and all sinusoids should have the same peak-to-peak amplitude. The peak-to-peak amplitudes of the output signals, recorded as a function of the input signals' frequencies, compose the system's magnitude response. Typically, the measured peak-to-peak amplitudes of the output signals are normalized with respect to a reference value. When the frequency response extends down to DC, the chosen reference value typically has a frequency near 0 Hz.

This approach was used to measure the frequency response magnitude of an experimental system. A function generator supplied the input signals, and an oscilloscope was used to measure the peak-to-peak amplitudes of the output signals. The measured frequency response data plotted in Figure 3 were normalized to the value measured at 10 Hz, the lowest-frequency input signal used for this test.

When the frequencies of the input sine waves were $<1.3 \times 10^3$ kHz, the peak-to-peak amplitudes of the output signals were approximately invariant. Over this lower frequency range, the amplitudes of the output sine waves increased and decreased with the amplitudes of the input sine waves. Both input and output sine waves achieved their respective maximum and minimum amplitudes during each cycle.

When the input sine waves oscillated with frequencies $>1.0 \times 10^4$ kHz, the output signals' peak-to-peak amplitudes were significantly attenuated. For these frequencies, the amplitudes of the input sine waves changed faster than the system could respond. Even though the input signals'

[3] For example, an application note for Thorlabs' MX40G E-O Converter describes the procedure for removing the instrument's frequency response (unit-specific frequency response data is supplied with each unit) from the output signal: https://www.thorlabs.com/images/TabImages/MX40G_De-Embed_Procedures.pdf

amplitudes changed at increasing rates, the rate at which the output signals' amplitudes changed did not increase. The amplitudes of the output signals still increased and decreased when the input signals' amplitudes increased and decreased. However, when the amplitude of the input signal reached a maximum, the amplitude of the output signal had reached only a fraction of its own full-scale maximum. The input signal then began decreasing, which forced the output signal to also decrease. A similar effect happened when the amplitude of the input signal reached a minimum. The result was an output signal whose peak-to peak amplitude decreased, and converged towards zero, as the frequencies of the input sine waves increased.

Input signal components with frequencies between approximately 1.3 x 10^4 kHz and 1.0 x 10^4 kHz occupy a transition region. Within this region, the values of the scaling factors transition from a maximum to a minimum. The amplitudes of input signal components with frequencies within this region are reduced, but not necessarily insignificant in the output signal.

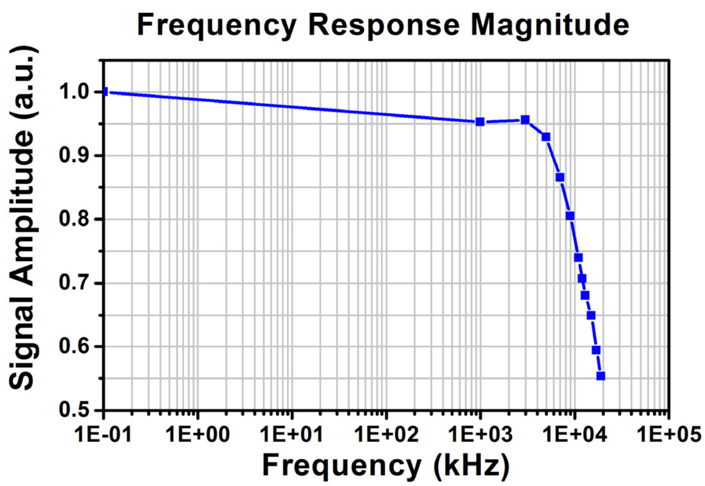


Figure 3 *The frequency response magnitude of an experimental system was found using a set of fixed-frequency input sine waves with equal amplitudes. The peak-to-peak amplitudes of the output signals were measured, normalized, and plotted as a function of the input signal frequency. Input signals with frequencies <1.3 x 10^3 kHz were all scaled by approximately the same high value. Input signals with frequencies >1 x 10^4 kHz were strongly attenuated by the system. The lines connecting the measured data points are guides for the eye.*

1.2.2 Frequency Response and 3 dB Bandwidth of a Low-Pass Filter

Low-pass filters have frequency response magnitudes with the general characteristics of the response shown in Figure 3. Ideal low-pass filters have normalized responses of one at low frequencies and zero at high frequencies, as shown in Figure 4. The responses of ideal filters instantaneously transition between one and zero at a frequency called the cutoff frequency. If a signal input to an ideal low pass filter included components with frequencies above and below the cutoff frequency, the output signal would consist entirely of input signal components whose frequencies were below the cutoff frequency. No higher-frequency component would be transmitted, since the filter would completely attenuate all frequency components above the cutoff frequency.

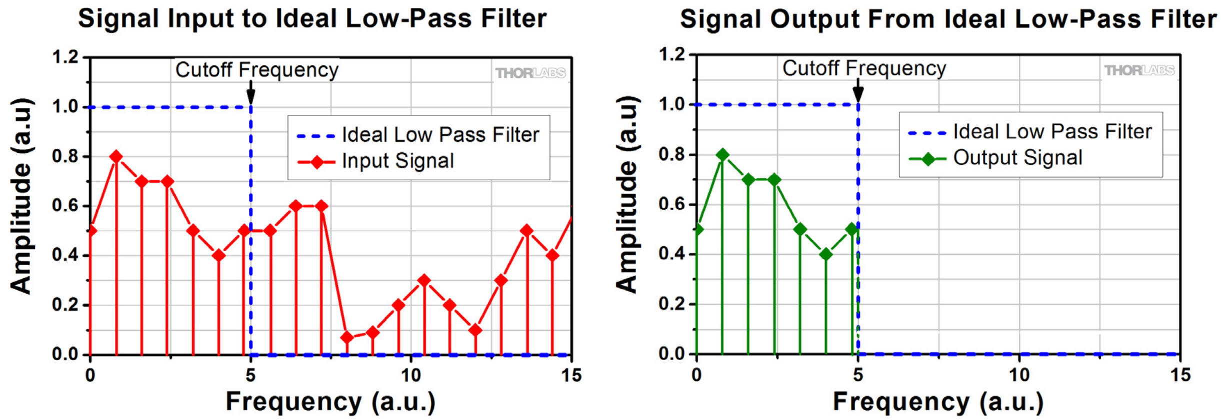


Figure 4 *When a signal input to an ideal low-pass filter contains components with frequencies above and below the cutoff frequency (left), the output signal will include only those components with frequencies less than or equal to the cutoff frequencies.*

Real low-pass filters have a finite-width transition region separating the highly transmissive and highly attenuating ranges of their frequency response magnitudes. The slope, which is often referred to as roll off, of the response in the transition region is one way of classifying the filter type. If the slope is proportional to approximately ω^{-N}, then the filter is an N^{th} order filter. Plots of idealized frequency response magnitudes, with orders of 1 through 5 are shown in Figure 5.

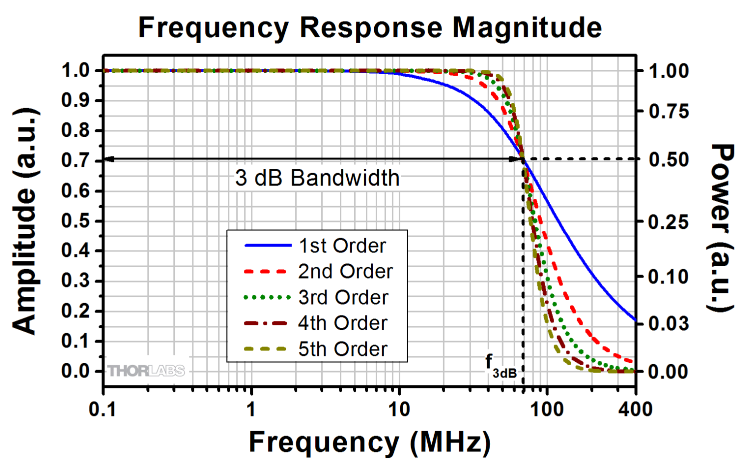


Figure 5 *The slope in the vicinity of the cutoff frequency (f_{3dB}) can be used to classify the filter type. Filters with slopes of ω^{-N} are N^{th} order low-pass filters. Note that all curves have a maximum value of one at DC and the same amplitude at f_{3dB}.*

The cutoff frequency of a real low pass filter is usually specified to be a frequency within its transition region. The frequency at which the magnitude of the response has decreased by (dropped to) a particular value is typically specified. An example of the frequency response magnitude measured for a low-pass filter with order 5 is plotted in Figure 6. As is discussed in the following, for electrical signals the 3 dB cutoff frequency is defined as the frequency at which the signal's voltage or current amplitude has dropped to 70.7% of a reference amplitude value.

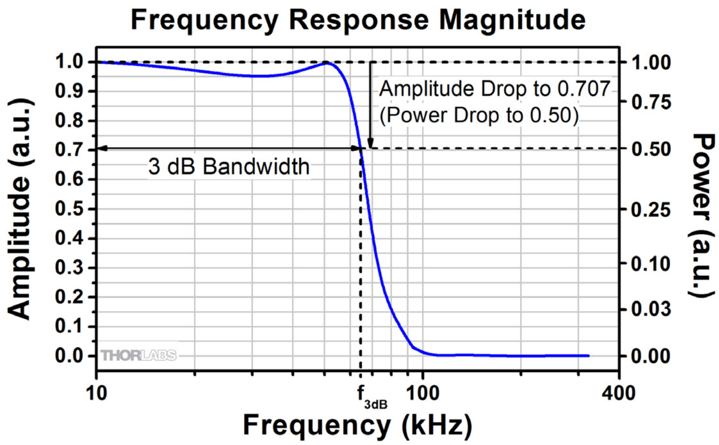


Figure 6 *Frequency magnitude response of the EF124 Low-Pass Electrical Filter showing a 3 dB cutoff frequency of approximately 68.4 kHz. Higher frequency components are strongly attenuated, while lower frequency components are preserved in the output signal. The cutoff frequency corresponds to a 0.707 multiplicative drop in voltage or current amplitude and a 0.5 multiplicative drop in power.*

An amplitude drop of 70.7% seems like an awkward value to reference, but it is equivalent to a more intuitive 50% drop in the signal's power. This can be shown using the relationship between the electrical signal's voltage amplitude (v), current amplitude (i), and power (P). As a first step, Ohm's Law,

$$\mathrm{v} = iR \quad , \tag{1}$$

which includes the value of resistance (R), is substituted into the equation for power,

$$P = \mathrm{v}i \quad . \tag{2}$$

This results in an expression for signal power,

$$P = i^2R = \frac{\mathrm{v}^2}{R} \quad , \tag{3}$$

as a function of current or voltage, which also includes the value of resistance. Eliminating the value of resistance from Eq. 3 can be done by dividing through by a reference value (P_o),

$$\frac{P}{P_o} = \frac{\mathrm{v}^2}{\mathrm{v}_o^2} = \frac{i^2}{i_o^2} \quad , \tag{4}$$

which results in unitless ratios of each parameter. From Eq. 4, it can be seen that when the ratio of voltage (v) to the reference voltage (v_o) is 70.7%,

$$\frac{\mathrm{P}}{\mathrm{P}_o} = \frac{\mathrm{v}^2}{\mathrm{v}_o^2} = \frac{(0.707\mathrm{v}_o)^2}{{\mathrm{v}_o}^2} = 0.5 \quad ,$$

the power (P) is 50% of the original power (P_o).

Signal power can vary, as both a function of time and frequency, over many orders of magnitude. When plotted on a linear scale, variations across a wide range of the lowest values can be obscured. This can hide important information about the performance of a system. Changing the scale of the y-axis from linear to logarithmic causes the evenly spaced major ticks of the y-axis to be incremented in powers of 10 (…0.001, 0.01, 0.1, 1, 10, 100 …). This reveals details about the lowest values, while allowing all data to be viewed on a single plot.

Alternatively, when the power is normalized using Eq. 4, the power ratio can be expressed in units of decibels (dB). The effect is similar to plotting the values on a logarithmic y-axis scale, since decibels are computed using the base-10 logarithm, as shown by Eq. 5,

$$P\ dB = 10\log_{10}\frac{P}{P_o} \quad . \tag{5}$$

The definition of the decibel includes a factor of 10 multiplying the logarithm for historical reasons. Only unitless ratios can be expressed as decibels.

Since power is often expressed in terms of decibels, equipment specifications such as the cutoff frequency of a low pass filter are also given in terms of decibels. As noted previously, the 3 dB cutoff frequency is equivalent to a 70.7% drop in signal amplitude, or a 50% drop in signal power. This decibel value can be calculated by substituting,

$$0.5 = \frac{P}{P_o} \quad ,$$

into Eq 5,

$$10\mathrm{Log}_{10}[0.5] = -3\ \mathrm{dB} \quad . \tag{6}$$

In the case of low-pass filters and systems with frequency response magnitudes similar to those of low-pass filters, the 3 dB bandwidth refers to the range of frequencies extending from DC to the 3 dB cutoff frequency (f_{3dB}), as indicated in Figure 6. In the phrases "3 dB bandwidth" and "3 dB cutoff frequency," the negative sign is typically omitted, as it is understood.

It is often useful to assume a low-pass filter behaves like an ideal low-pass filter when roughly modelling system behavior. This approximation scales all input signal components with frequencies below the 3 dB cutoff by the same maximum value. In other words, the filter is assumed to "pass" those frequency components relatively unchanged to the output signal. In addition, input signal components with frequencies above the 3 dB cutoff are assumed to be completely absent in the output signal, or to be so highly attenuated that they do not contribute significantly to the output signal. Note that the assumption of ideal low-pass filter behavior may not be appropriate. For example, contributions from signal components with frequency components above the cutoff frequency may not be insignificant to an application.[4]

When the input signal is known, the 3 dB bandwidth of a device or system can be used to model and analyze the distortion of the output signal. The *Pulse Distortion Lab Fact*[4] used this

[4] Pulse Distortion Lab Fact: https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=11718

approach to investigate the distortion of pulse trains output by a bandwidth limited system. The input signals were different rectangular pulse trains with a range of repetition rates and duty cycles.

1.3 Proportional Relationship for Low-Pass Filters

The minimum rise time and 3 dB cutoff frequency parameters provide related information about the performance of a system or device. Referencing the rise time may be most useful when considering the signal as a function of time, and referencing the 3 dB frequency may be preferred when working with the frequency content of the signal.

When a system has a frequency response magnitude resembling that of a low-pass filter,

$$\tau_{\mathrm{r}} \cong \frac{0.35}{f_{3dB}} \quad , \tag{7}$$

can be used to estimate the rise time from the 3 dB frequency, or *vice versa*. Section 2 details one approach to deriving Eq. 7.

2 Derivation of Response Equations for Low-Pass Filters[5,6]

Models of systems can serve a range of purposes, including serving as bases for deriving mathematical expressions that describe different aspects of system behavior. The ideal RC low pass filter circuit, which exhibits both time and frequency dependency, is one model. General expressions derived from it can be used to describe the time and frequency responses of systems exhibiting 1st order low-pass filter behavior (Section 1.2.2). This model can also be used to find equations relating system parameters, such as rise time and 3 dB bandwidth.

In this section, the ideal RC low-pass filter circuit model, which is shown in Figure 7 and consists of an ideal resistor (R) and capacitor (C) in series, is used to derive expressions for:

- The rise time, using the rate of charge accumulation on the parallel-plate capacitor under dynamic conditions.
- The 3 dB bandwidth, using the frequency transfer function of the RC circuit.
- The relationship between rise time and 3 dB bandwidth, given in Eq. 7.

2.1 Time-Dependent Response and Rise Time

The RC low-pass filter circuit used to derive the expression for rise time is shown in Figure 7. It includes a voltage source (v_o) and a switch between the resistor and capacitor. An expression for rise time can be derived by describing the dynamic movement of charge in the circuit

[5] Douglas. C. Giancoli, *Physics for Scientists and Engineers with Modern Physics, Second Edition* Prentice Hall, Englewood Cliff, New Jersey, 1989.
[6] Ferrel G. Stremler, *Introduction to Communications Systems, Third Edition*, Addison-Wesley Publishing Company, Inc. New York, 1990.

immediately after the switch closes. This derivation assumed that while the switch was open, no current flowed and the voltage drops across the resistor (v_R) and capacitor (v_C) were zero.

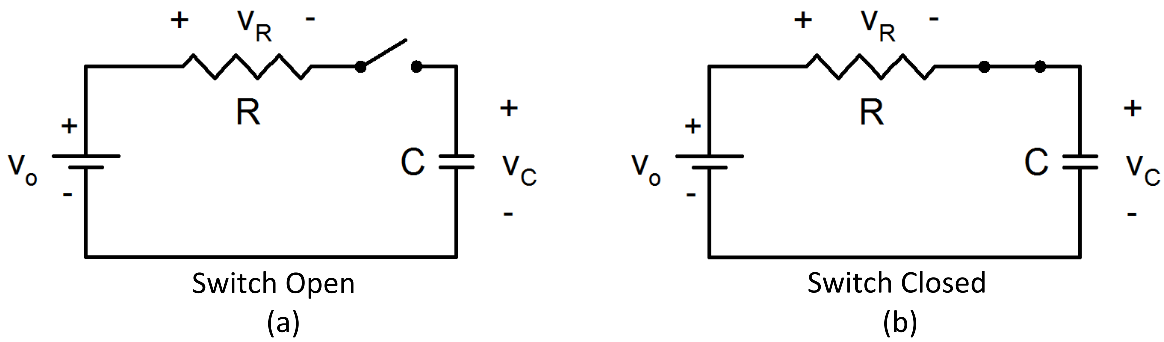


Figure 7 *An expression for rise time can be derived by considering the dynamic flow of charge in an RC low-pass filter circuit immediately after the open switch depicted in (a) is closed, as shown in (b). R, C,* v_R*, and* v_C *label the values of resistance, capacitance, and voltages. The voltage supplied by the source is* v_o*.*

Immediately after the switch closes, current begins to flow through the resistor. The current,

$$i(t) = \frac{dQ(t)}{dt} \quad , \tag{8}$$

equals the quantity of charge (Q) passing through the cross section of the wire during a unit time interval.

As illustrated in Figure 8, electrons flow from the negative terminal of the voltage source and accumulate on the bottom plate of the capacitor. Electrons also flow counter-clockwise from the upper plate of the capacitor towards the positive terminal of the voltage source.[7] A negative charge ($-Q$), builds on the bottom plate of the capacitor and a positive charge ($+Q$), accumulates on the top.

A voltage (v_C) proportional to the charge,

$$Q(t) = C\mathrm{v}_C(t) \quad , \tag{9}$$

develops across the capacitor, where the capacitance is the proportionality constant.

The voltage drop across the capacitor increases with time and asymptotically approaches the magnitude of the battery's voltage (v_o). Since v_C opposes the current, the flow of charge in the circuit is highest immediately after the switch closes and then decreases with time. When steady-state conditions are reached, v_C equals v_o and there is no net current flow.

The current flowing at a particular time,

$$i(t) = C\frac{d\mathrm{v}_C(t)}{dt} \quad , \tag{10}$$

[7] The flow of electrons is opposite the flow of conventional current, which flows clockwise through the circuit.

is directly proportional to how quickly the voltage across the capacitor is changing at that time. The capacitance is the proportionality constant that scales the magnitude of the current.

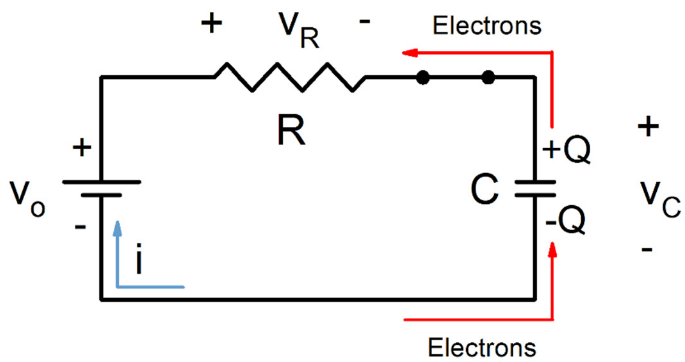


Figure 8 *After the switch closes, electrons flow to the bottom plate of the capacitor (C) and away from the top plate. The accumulation of negative charge (-Q) on the bottom plate and positive charge on the top plate (+Q) causes a voltage to develop across the capacitor (v_C). The labels R, i, v_R, and v_o indicate resistance, current*[7]*, voltage drops across the resistor, and the voltage supplied by the source, respectively.*

An intermediate step towards the goal of expressing v_C in terms of the known variables v_o, R, and C, is deriving an equation for the circuit in terms of a single unknown variable, $Q(t)$.

First the voltages around the circuit are summed,

$$\mathrm{v}_o = \mathrm{v}_\mathrm{R}(t) + \mathrm{v}_C(t) \quad . \tag{11}$$

Next, Ohm's law,

$$\mathrm{v}_R(t) = i(t)R \quad , \tag{12}$$

is used to remove v_R,

$$\mathrm{v}_o = i(t)R + \mathrm{v}_C(t) \quad , \tag{13}$$

and then Eq. 8 and Eq. 9 are used to remove i and v_C , respectively. The only unknown in the resulting equation,

$$\mathrm{v}_o = R\frac{dQ(t)}{dt} + \frac{Q(t)}{C} \quad , \tag{14}$$

is $Q(t)$.

After separating terms, Eq. 14 can be integrated,

$$\int \frac{dQ}{(Q - C\mathrm{v}_o)} = -\frac{1}{RC}\int dt \quad , \tag{15}$$

to eliminate the derivative.

The result of the integration,

$$ln(Q - C\mathrm{v}_o) = -\frac{t}{RC} + K \quad , \tag{16}$$

includes the constant K. Setting Q equal to zero at $\mathrm{t} = 0$ in Eq. 16 gives an expression for K,

$$ln(-C\mathrm{v}_o) = K \quad , \tag{17}$$

in terms of known quantities. The relationship given by Eq. 17 can be used to eliminate K from Eq. 16,

$$ln\left(1 - \frac{Q}{C\mathrm{v}_o}\right) = -\frac{t}{RC} \quad , \tag{18}$$

and then Eq. 18 can be converted into the equivalent exponential function,

$$\left(1 - \frac{Q}{C\mathrm{v}_o}\right) = e^{-t/RC} \quad , \tag{19}$$

for easier manipulation. After rearranging to isolate Q,

$$Q = C\mathrm{v}_o\left(1 - e^{-t/RC}\right) \quad , \tag{20}$$

Eq. 10 can be used to remove the variable Q and include the variable v_C. The resulting equation,

$$\mathrm{v}_C(t) = \mathrm{v}_o\left(1 - e^{-t/RC}\right) \quad , \tag{21}$$

can be used to derive an expression for the minimum rise time of the RC low-pass filter circuit.

2.1.1 Rise Time in Terms of the RC Product

As described in Section 1.1, a system's minimum rise time can be found using a step function as an input signal and measuring the duration of time separating two points on the rising edge of the output signal. This work follows the common convention of locating those points 10% and 90% up the rising edge of the curve. This is illustrated in Figure 9, with the first point occurring at time t_1 and the second at time t_2. The signal asymptotically approaches the maximum value and reaches it at infinite time ($t = \infty$).

Using Eq. 21 and the definition of rise time illustrated in Figure 9, a relationship can be found between t_1 and the RC constant,

$$\mathrm{v}_C(t_1) = 0.1 = \mathrm{v}_o\left(1 - e^{-t_1/RC}\right) \quad , \tag{22}$$

but an expression that does not include v_o is desired.

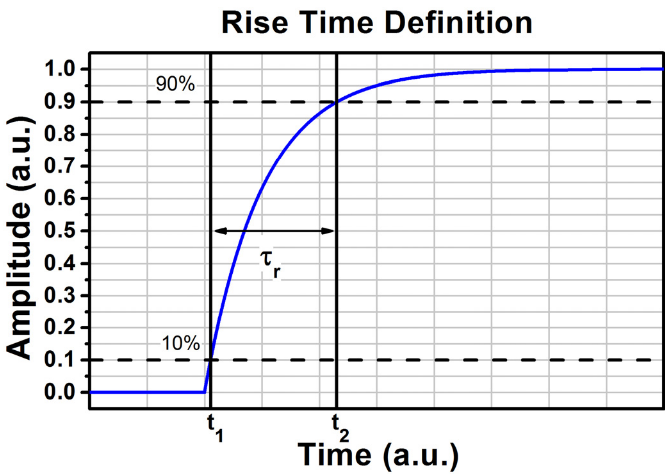


Figure 9 *The minimum rise time is the time required for the output signal to transition between points 10% and 90% up the rising edge of the curve generated in response to a step function input.*

A simplified equation, without the variable v_o, can be found by dividing Eq. 22 by the expression resulting from inserting an infinite time into Eq. 21,

$$\frac{\mathrm{v}_c(\tau_1)}{\mathrm{v}_c(\infty)} = \frac{0.1}{1} = \frac{\mathrm{v}_o\left(1 - e^{-t_1/RC}\right)}{\mathrm{v}_o(1-0)} \quad , \tag{23}$$

then simplifying,

$$0.1 = \left(1 - e^{-t_1/RC}\right) \quad , \tag{24}$$

isolating the exponential,

$$0.9 = e^{-t_1/RC} \quad , \tag{25}$$

and finally taking the natural logarithm of both sides,

$$ln(0.9) = -t_1/RC \quad . \tag{26}$$

The relationship for t_2,

$$ln(0.1) = -t_2/RC \quad , \tag{27}$$

was found using the same procedure. The rise time in terms of the RC product is then,

$$\tau_r = t_2 - t_1 = ln\left(\frac{0.9}{0.1}\right)[RC] \cong 2.2RC \quad . \tag{28}$$

2.2 Frequency Response and 3 dB Frequency[8]

Section 2.1 described the response of the RC low-pass filter circuit as a function of time. This section describes the response as a function of frequency. The RC low-pass filter circuit

[8] Ferrel G. Stremler, *Introduction to Communications Systems, Third Edition*, Addison-Wesley Publishing Company, Inc. New York, 1990.

illustrated in Figure 10 is used to derive mathematical expressions for a system's frequency response and 3 dB frequency. To obtain these equations, it is useful to view the circuit as a system relating input and output signals. The input signal (v_{in}) is the voltage source, and the output signal (v_{out}) is the voltage drop across the capacitor.

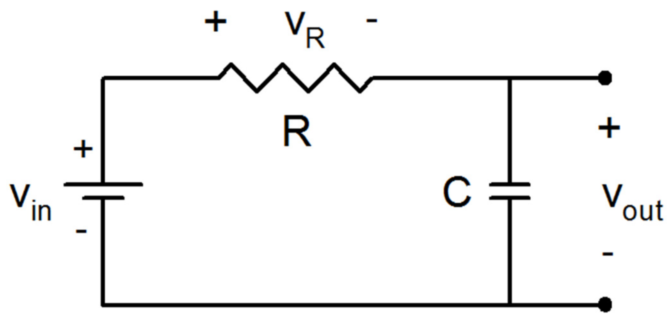


Figure 10 *An RC low-pass filter circuit can be seen as a system transferring an input, supplied by the source (v_{in}), to an output (v_{out}), measured across the capacitor (C). The resistor (R) is in series with the capacitor.*

2.2.1 Transfer Function Overview

A linear, time-invariant system like the RC low-pass circuit in Figure 10 can be described by the general diagram shown in Figure 11, which highlights the relationship among the input signal, system, and output signal. The source voltage (v_{in}) and the voltage measured across the capacitor (v_{out}) in Figure 10 become the input to and output from, respectively, the system block in Figure 11. The system block, which represents the frequency response of the circuit, converts the input signal to the output signal and is known as the transfer function ($H(\omega)$) of the system. The transfer function depends only on frequency and the parameters of the system and is valid for use with any physical input function.

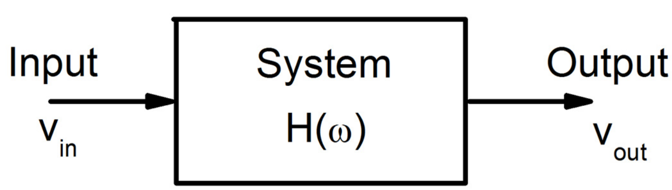


Figure 11 *A system provides the relationship between the input (v_{in}) and output (v_{out}) signals. A mathematical expression that relates any input signal with the resulting output signal is called a transfer function (H).*

The diagram in Figure 11 can be expressed as a mathematical equation,

$$\mathrm{v}_{out}(t) = H(\omega)\mathrm{v}_{in}(t) \quad . \tag{29}$$

The transfer function can be completely isolated by dividing both sides of Eq. 29 by the input function. To derive the transfer function, it must be possible to separate expressions for v_{in}, v_{out}, and $H(\omega)$. In addition, the expression for $H(\omega)$ must include only system parameters.

The relationships between input and output signals of electrical circuits, such as the low-pass filter in Figure 10, can be expressed mathematically using differential equations. Methods for solving these linear equations exist and can be used to find an expression in the form of Eq. 29. The differential expression can be written in a completely general way,

$$b_o \mathrm{v}_{out}(t) + b_1 \frac{d\mathrm{v}_{out}(t)}{dt} + b_2 \frac{d^2\mathrm{v}_{out}(t)}{dt^2} + \cdots = a_o \mathrm{v}_{in}(t) + a_1 \frac{d\mathrm{v}_{in}(t)}{dt} + a_2 \frac{d^2\mathrm{v}_{in}(t)}{dt^2} + \cdots \quad , \tag{30}$$

without specifying a particular v_{in} or v_{out}. The constant coefficients (a_m and b_n) are parameters that characterize the response of the system and do not depend on v_{in} or v_{out}. The generality of the expression in Eq. 30 makes it applicable to any linear, time-invariant system.

An established approach for isolating an expression for $H(\omega)$, which is dependent only on frequency and the system parameters represented by constant coefficients a_m and b_n, is to replace v_{in} with an exponential function,

$$\mathrm{v}_{in}(t) = e^{j\omega t} \quad , \tag{31}$$

which has a number of convenient properties that make it a favorite tool when working with differential equations.

One reason to use the exponential function, and to include the variable of differentiation (t) in its argument, is the form of its derivative,

$$\frac{d^n}{dt^n}\left(e^{j\omega t}\right) = (j\omega)^n e^{j\omega t} \quad , \tag{32}$$

which is the product of the original function with a constant. Angular frequency (ω) is included in the argument since an expression in terms of frequency is desired.

Another useful characteristic of Eq. 31 is its purely imaginary ($j = \sqrt{-1}$) argument. With an imaginary argument, the function will not become infinitely large at infinite time, and the function will remain valid at times before $t = 0$. Both mathematical properties are required to ensure physical results.

Another significant benefit of using Eq. 31 is that with this input, the output signal provided by Eq. 30 is immediately known due to a property of linear systems. The output,

$$\mathrm{v}_{out}(t) = H(\omega) e^{j\omega t} \quad , \tag{33}$$

is a function equal to the product of the input and the transfer function ($H(\omega)$). The transfer function is complex valued and describes the difference between the input and output signals' amplitudes and phases.

An expression for $H(\omega)$, expressed solely in terms of system parameters, can be found by inserting the input signal (Eq. 31) and the output signal (Eq. 33) into the differential equation given in Eq. 30.

The infinite sum of derivatives on the right side of Eq. 30 becomes an infinite sum of terms,

$$a_0\left(e^{j\omega t}\right)+a_1(j\omega)\left(e^{j\omega t}\right)+a_2(j\omega)^2\left(e^{j\omega t}\right)+\cdots=e^{j\omega t}\sum_{m=0}a_m(j\omega)^m \quad , \qquad (34)$$

that are all multiplied by the input function. By substituting Eq. 33 into Eq. 30, and applying the property described in Eq. 32, the infinite sum of derivatives on the left side of Eq. 30,

$$\begin{aligned}b_o\left[H(\omega)e^{j\omega t}\right]+b_1(j\omega)\left[H(\omega)e^{j\omega t}\right]+b_2(j\omega)^2\left[H(\omega)e^{j\omega t}\right]+\cdots\\ =H(\omega)e^{j\omega t}\sum_{n=0}b_n(j\omega)^n \quad ,\end{aligned} \qquad (35)$$

produces a result similar to Eq. 34. Replacing the right and left sides of Eq. 30 by the right sides of Eq. 34 and Eq. 35, respectively,

$$H(\omega)e^{j\omega t}\sum_{n=0}b_n(j\omega)^n=e^{j\omega t}\sum_{m=0}a_m\,(j\omega)^n \quad , \qquad (36)$$

results in a frequency-dependent expression for the transfer function in terms of only system parameters, as desired.

Rearranging Eq. 36 results in an expression for the transfer function,

$$H(\omega)=\frac{\sum_{m=0}a_m(j\omega)^m}{\sum_{n=0}b_n(j\omega)^n} \quad , \qquad (37)$$

which is dependent only on the angular frequency and the constant coefficients (a_m and b_n). This equation is complex valued and describes how the system affects both the phase and the magnitude of every frequency component within the input signal.

2.2.2 Transfer Function of the RC Low-Pass Filter Circuit

The frequency transfer function for the RC low-pass filter circuit shown in Figure 12 can be found using circuit analysis techniques for linear, time-invariant systems. The following derivation assumes ideal conditions, in which the voltage source (v_{in}) has zero impedance while the output terminal used to measure v_{out} has an infinite load impedance.

The equation describing the currents (i_R and i_C) flowing through the node marked by the blue dot,

$$i_R=i_C \quad , \qquad (38)$$

provides the starting point for deriving the differential equation. This relationship becomes more useful when Eq. 9 and Eq. 11 are used to express Eq. 38 in terms of voltages,

$$\frac{v_R(t)}{R} = C\frac{dv_{out}(t)}{dt} \quad . \tag{39}$$

The next step towards obtaining the form of Eq. 30 is to express the voltage drop across the resistor (v_R) in terms of input and output voltages. The voltage drop across the resistor equals the difference of the voltages on either side ($v_{in} - v_{out}$),

$$\frac{v_{in}(t) - v_{out}(t)}{R} = C\frac{dv_{out}(t)}{dt} \quad . \tag{40}$$

The elements of Eq. 40 are then arranged,

$$v_{out}(t) + RC\frac{dv_{out}(t)}{dt} = v_{in}(t) \quad , \tag{41}$$

to group terms dependent on v_{out} and v_{in} on opposite sides of the equation. Substituting equations Eq. 31 and Eq. 33 in for v_{in} and v_{out}, respectively, and taking the time derivative provides an expression,

$$H(\omega)e^{j\omega t} + RC[j\omega]\mathrm{H}(\omega)e^{j\omega t} = e^{j\omega t} \quad , \tag{42}$$

in which each term is multiplied by the exponential input function, Eq. 31.

The complex-valued frequency transfer function,

$$H(\omega) = \frac{1}{j\omega RC + 1} \tag{43}$$

can be found by dividing through by the exponential function and separating terms.

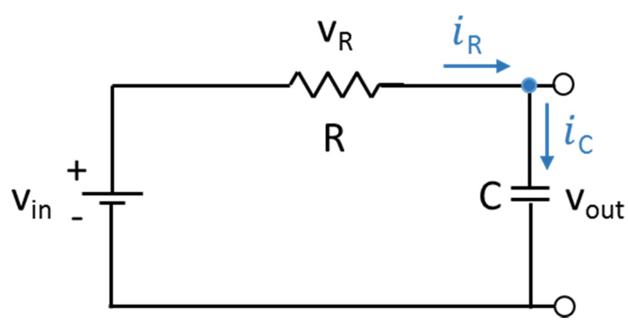


Figure 12 *The flow of current into and out of the node between the resistor and capacitor is noted for this RC low-pass filter circuit. Summing these currents is the first step in finding the frequency transfer function.*

2.2.3 The 3 dB Bandwidth of the RC Low-Pass Filter in Terms of the RC Constant

The frequency transfer function given by Eq. 43 is complex valued and provides information about both the frequency-dependent phase and magnitude responses of the RC low-pass filter. However, the 3 dB bandwidth is defined with respect to only the real-valued frequency response magnitude, which is the absolute value of Eq. 43 ($|H|$). The first step of finding the magnitude of the frequency response is multiplying Eq. 43 by its complex conjugate (H^*),

$$H(\omega)H^*(\omega) = \left[\frac{1}{j\omega RC + 1}\right]\left[\frac{1}{-j\omega RC + 1}\right] \quad , \tag{44}$$

which results in the absolute value of $H(\omega)$ squared,

$$|H(\omega)|^2 = \frac{1}{[\omega RC]^2 + 1} \quad . \tag{45}$$

The magnitude of the frequency response,

$$|H(\omega)| = \frac{1}{\sqrt{[\omega RC]^2 + 1}} \quad , \tag{46}$$

is found by taking the square root of both sides of Eq. 45. As Eq. 46 is approximately proportional to ω^{-1}, this circuit has the response of a 1st order low-pass filter (Figure 5).

As is discussed in Section 1.2.2, the 3 dB cutoff frequency is the lowest frequency for which the normalized frequency response magnitude is 0.707. Using Eq. 46 and the relationship between angular and temporal frequency ($\omega = 2\pi f$),

$$0.707 = \frac{1}{\sqrt{[2\pi f_{3dB} RC]^2 + 1}} \quad . \tag{47}$$

Solving Eq. 47 provides the 3 dB cutoff frequency for the RC low-pass filter,

$$f_{3dB} = \frac{1}{2\pi RC} \quad . \tag{48}$$

2.3 Rise Time Related to 3 dB Bandwidth

Both the minimum rise time (Eq. 28) and the 3 dB cutoff frequency (Eq. 48) derived for the RC low-pass filter circuit are functions of the RC product. These two equations can be combined to eliminate the RC constant and relate the minimum rise time and 3 dB bandwidth.

Rearranging the expressions for rise time (Eq. 28),

$$RC \cong \frac{\tau_r}{2.2} \quad ,$$

and the expression for the 3 dB cutoff frequency (Eq. 48),

$$RC = \frac{1}{2\pi f_{3dB}} \quad ,$$

and equating them,

$$\tau_r \cong \frac{0.35}{f_{3dB}} \quad , \tag{49}$$

results in an expression relating the two parameters through a proportionality constant. It is important to keep in mind that these equations were derived using an ideal RC low-pass filter circuit model. These equations can provide useful approximate values for systems that respond like low-pass filters but are not recommended for use during rigorous analysis.

3 Measuring and Calculating Minimum Rise Time

Referencing both rise time and 3 dB bandwidth parameters can facilitate the evaluation of system performance. However, both parameters are not always specified, and in some cases neither is provided. When this is the case, measuring these parameters is theoretically an option, but the available equipment may not support a direct measurement of both.

The equations derived in Section 2 are useful for estimating the minimum rise time and 3 dB electrical bandwidth from one another, assuming the system responds like an RC low-pass filter. When systems have responses similar to RC low-pass filters, Eq. 22 is a good model for the rising edge of the signal output in response to an input step function, and Eq. 46 is a good model for the system's frequency response magnitude. The following two sections provide examples of using Eq. 49, as well as directly measuring minimum rise time and 3 dB cutoff frequency.

This section does not consider possible complications arising from electrical feedback in the system. Reflections of the electromagnetic field occur at locations of impedance mismatch within the system and contribute to the feedback. Feedback can introduce non-negligible artifacts into both time-domain and frequency domain signal. These artifacts typically manifest as oscillations superposed on the signal in the time domain and peaks added to the signal in the frequency domain, with the frequencies of the artifacts being related to the resonance frequencies of the system. Accurately determining rise time and 3 dB cutoff frequency can require minimizing the artifacts by reducing the impedance mismatch(es) and / or post-processing the measured data.[9]

3.1 Measurement of Minimum Rise Time and Calculation of f_{3dB}

A single measurement of the output signal is sufficient to determine minimum rise time. However, an accurate measurement assumes certain conditions are met, including:

- An input signal that transitions between steady-state low and high amplitudes faster than the system under test can respond.
- An oscilloscope, or other instrument used to measure the output signal, that can respond quickly enough to obtain a high-fidelity measurement of the output signal.
- An output signal with well-defined minimum and maximum steady-state amplitudes.

Measurements of minimum rise time were made for a photodiode-based system using equipment and input signals that met these requirements. The input signal consisted of 3 ms duration pulses supplied by a pulsed laser diode, the photodiode was the sensor in a DET20C InGaAs photodetector, and a 1 GHz oscilloscope was used to measure the output signal.

[9] Alan V. Oppenheim, *et. al.*, *Discrete-Time Signal Processing*, Prentice-Hall, Inc., Upper Saddle River, NJ, 1989.

A load resistor (R_{Load}) was placed after the photodetector, in parallel with the oscilloscope's 1 MΩ internal impedance. The oscilloscope's internal impedance added to the load resistance,

$$R_{\text{Effective}} = \left[\frac{1}{R_{Load}} + \frac{1}{1\ \text{M}\Omega}\right]^{-1} \quad , \tag{50}$$

to provide an effective resistance ($R_{Effective}$) for the system. The value of the load resistor was varied to obtain different effective load resistances from 50 Ω to 1 MΩ.

Output signals were measured with respect to time, and representative examples are plotted in Figure 13. The RC low-pass filter circuit is a generally accepted model for the response of a photodiode-based linear system.[10] Two of the curves included in Figure 13 are plotted individually in Figure 14. The red curves included in the plots shown in Figure 14 were calculated using Eq. 22, which specifies the time-domain output signal of the RC low-pass filter circuit in response to a step function input. The good fit of the modelled to the measured data indicate this is an appropriate model for describing the response of this system.

According to Eq. 28, the rise time measured for the output signal should increase with the value of effective resistance, and the data support this relationship. The angle between the rising edge of the signal and the y-axis increased with the value of the load resistor. Shallower slopes on the rising edges of the traces correspond to higher rise times.

The rise times were measured directly from the oscilloscope traces, as shown in Figure 14. The 3 dB bandwidth values estimated using the measured rise times and Eq. 49 are listed in Table 1. Since the two parameters have a reciprocal relationship, an increase in the system's rise time corresponds to a decrease in the range of input signal frequencies to which the system is responsive. The system attenuates the frequency components outside its 3 dB bandwidth. However, higher frequency components in the signal are responsible for the abrupt and narrow time-domain features. It can be seen in Figure 13 that longer rise times result in more rounded edges in the time-domain pulse shapes. With higher rise times (reduced 3 dB bandwidths), finer features will be broader and more rounded in the time-domain output signal. This distortion of the output signal can be predicted and analyzed using knowledge of the input signal and the system's 3 dB bandwidth.

The data plotted in Figure 13 illustrate a challenge with this measurement approach: the duration of the input pulse must be long enough to allow the output signal to stabilize at the maximum value. A pulse duration of 1000 µs would have been sufficient to measure the rise times for all but the 1 MΩ case, which could not have been measured with confidence. The 1 MΩ curve reaches the maximum value after approximately 2000 µs, but a longer pulse duration is necessary to confirm the output signal's maximum amplitude.

These plots in Figure 14 illustrate another challenge of making time-domain measurements of rise time: the shorter the time scales, the fewer the data points collected during each

[10] Govind P. Agrawal, *Fiber-Optic Communication Systems, Second Edition*, John Wiley & Sons, Inc. New York, 1997.

measurement. In the case of these measurements, there were sufficient data points to determine the minimum rise times.

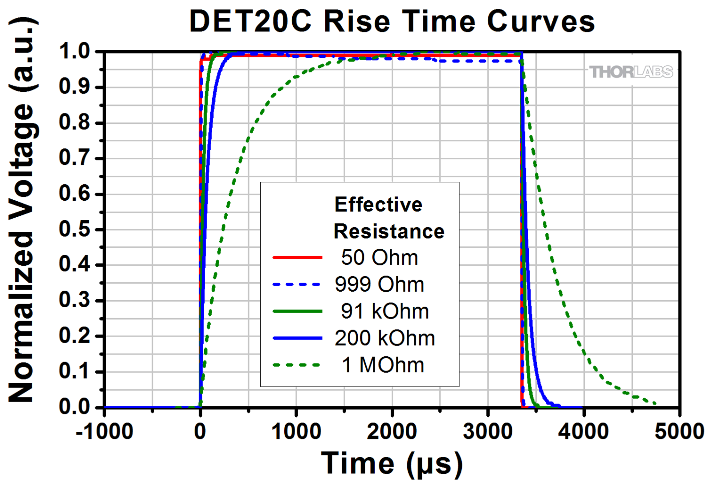


Figure 13 *Pulses output by a photodiode-based system were measured by an oscilloscope. The minimum rise times were measured between points 10% and 90% up the rising edges of the pulses, whose slopes varied with the effective load resistance.*

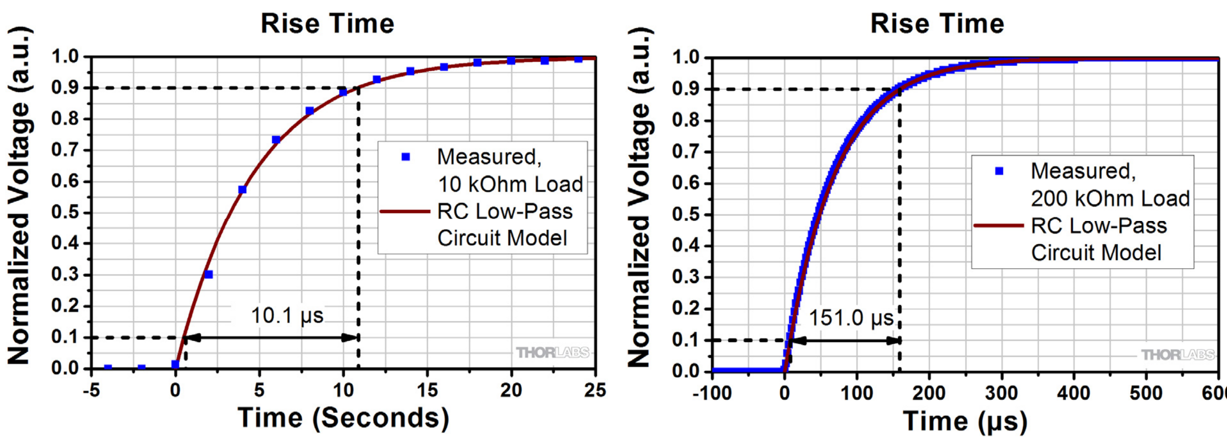


Figure 14 *Measurements of rise time, for two cases of effective load resistance, plotted with red curves calculated using the RC low-pass filter model (Eq. 22). The good fit of the curves to the data indicate the RC low-pass filter circuit is a good model for this system. Increasing the effective resistance increases the rise time. Shorter rise times may result in fewer data points measured along the rising edge.*

Table 1 The 3 dB bandwidths were estimated from the rise time measurements plotted in Figure 14.

Effective Resistance (kΩ)	10	200
Minimum Rise Time, τ_r (µs)	10.1	151.0
3 dB Bandwidth, f_{3dB} (kHz)	34.7	2.3

3.2 Measurement of 3 dB Bandwidth and Calculation of τ_r

The 3 dB bandwidth of a system can be obtained from its frequency response magnitude, as described in Section 1.2. The frequency response magnitude is obtained by making multiple individual measurements, while the rise time can be determined from a single measurement. However, equipment or other limitations may render a direct measurement of a system's minimum rise time impractical. When this is the case, and the response of the system resembles that of an RC low-pass filter circuit, rise time can be estimated from the system's 3 dB bandwidth using Eq. 49.

This approach was used during an investigation of the dependence of a photodiode's rise time on the reverse-bias voltage. Rise times were expected to occur on a nanosecond time scale, which could not be measured accurately by immediately available equipment.

Using the procedure described in Section 1.2, the frequency response magnitude of the system was measured between 0.1 MHz and 39 MHz for reverse-bias voltages between 0.1 V and 3 V applied across the photodiode. This frequency range included the 3 dB cutoff frequency for each reverse-bias test case, as shown in Figure 15. This system's response could be adequately modeled by an RC low-pass filter, since curves calculated using Eq. 46 provided good fits to the measured data. Examples are plotted in Figure 16.

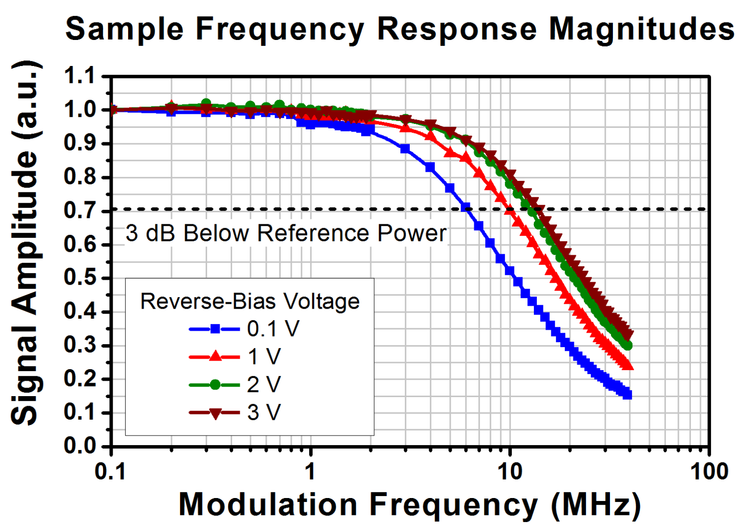


Figure 15 *Frequency response magnitude data were acquired using a photodiode-based system for cases in which four different reverse-bias voltages were applied across the photodiode. The 3 dB bandwidth was measured for each case. The lines connecting the data points are guides to the eye.*

The rise times for the photodiode-based system were estimated using the measured 3 dB bandwidth values and Eq. 49. The calculated rise time values are listed in Table 2 and plotted in Figure 17. The data show that increasing the reverse-bias voltage across the photodiode was effective in increasing the 3 dB bandwidth of the system. As the 3 dB bandwidth increased, the rise times decreased. This would increase the rate at which the system would respond to abrupt changes in the input signal, and reduce the distortion in the output signal. As an excessively high reverse-bias voltage will damage the photodiode, there is a limit to the improvement in response that can be obtained using this technique.

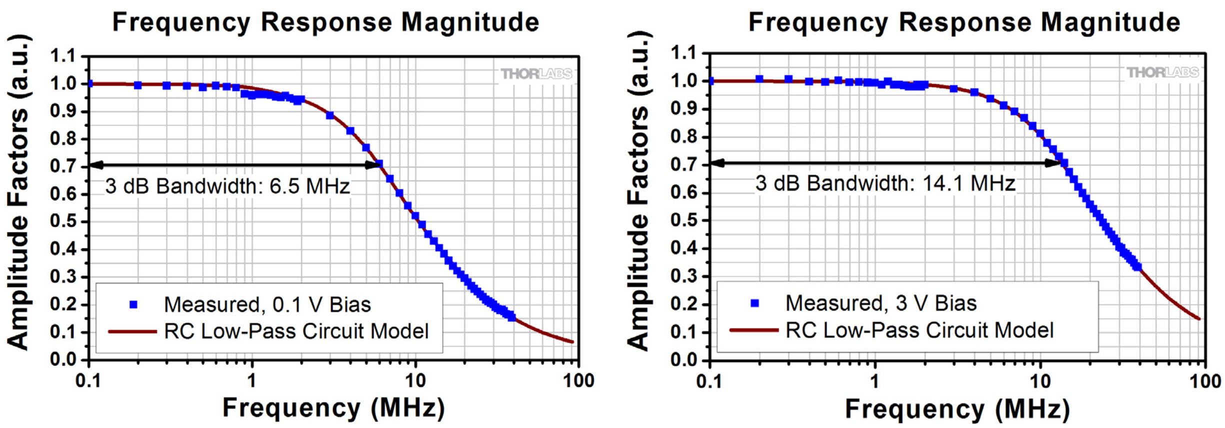


Figure 16 *Two frequency response magnitude data sets are shown, for cases of 0.1 and 3 V reverse-bias, with red curves calculated using the RC low-pass filter model (Eq. 46). The good fit of the curves to the data indicated the RC low-pass filter circuit is a good model for this system. Increasing the reverse bias increases the 3 dB bandwidth.*

Table 2 The rise times were estimated from the 3 dB bandwidths, which were obtained from the measured frequency response magnitudes, and Eq. 49.

Reverse Bias Voltage (V)	0.1	0.3	0.5	1.0	1.5	2.0	2.5	3.0
3 dB Bandwidth, f_{3dB} (MHz)	6.5	7.4	8.7	10.0	11.5	12.6	13.6	14.1
Minimum Rise Time, τ_r (ns)	54.1	47.4	40.2	34.9	30.4	27.8	25.8	24.7

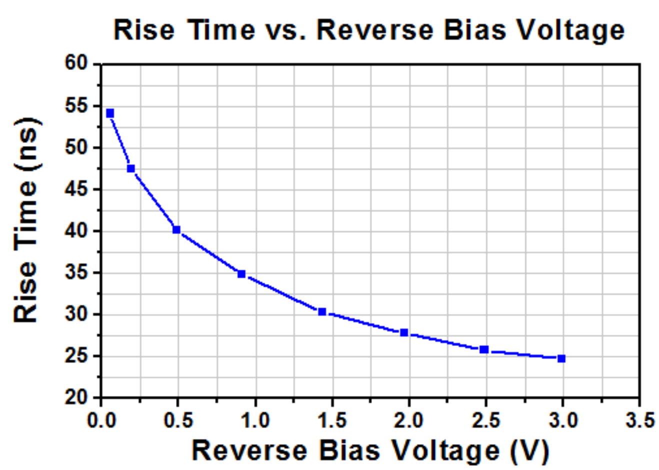


Figure 17 *Estimates of minimum rise time were calculated for each reverse bias voltage test case using the measured 3 dB bandwidth and Eq. 49.*

The quality of the frequency response data has a direct influence on the accuracy of the values of the 3 dB bandwidths found using this data. Variations in the peak-to-peak amplitudes of the input sinusoidal signals, differences between their recorded and actual fixed oscillation frequencies, and errors in the measurements of the peak-to-peak oscillation frequencies of the output signals would all have negatively impacted the quality of the data. In addition, errors in the 3 dB bandwidth values could arise from interpolating the value of the cutoff frequency.

4 Summary

An expression ($\tau_r \cong 0.35/f_{3dB}$) sometimes used to relate the minimum rise time and 3 dB bandwidth of electrical and electro-optical systems was considered. It was shown that this expression results when equations for the time- and frequency-dependent responses of an ideal RC low-pass filter circuit are combined. Since the mathematical expression relating rise time and 3 dB frequency is based on this ideal circuit, the accuracy of the values computed using it depend on the similarity of the response of the system under consideration to that of an ideal RC low-pass circuit.

The use and utility of the relation was demonstrated using experimental data acquired for two photodiode-based systems. It was determined that the response of these systems could be adequately modelled by an ideal RC low-pass-filter circuit, by observing good fits between the measured data and curves calculated using equations derived for the circuit.

For one system, rise times of the output signals were measured, and corresponding estimates of 3 dB bandwidth values were calculated. The input signal consisted of pulsed emission from a laser diode incident on a photodiode. The output signal was acquired by an oscilloscope, and the value of the load resistor was different for each rise time measurement. The data showed that the rise time of the system increased with load resistance. As the rise time increased, the 3 dB bandwidth decreased. As more higher-order frequencies were attenuated by the system, the output pulse shapes were increasingly distorted from the pulse shapes of the input signals.

For the second system, the 3 dB bandwidths were found from measurements of the system's frequency response magnitude, and estimates of the rise times were calculated using the relation. The input signals were sine waves with equal peak-to-peak amplitudes but different fixed oscillation frequencies. The reverse-bias voltage applied across the photodiode was varied for each set of measurements. The frequency response magnitude consisted of the measured peak-to-peak amplitudes of the output signals recorded as a function of input signal frequency. The data showed that as the reverse-bias voltage was increased, the 3 dB bandwidth also increased. As the 3 dB bandwidth of the system increased, the rise time decreased. The system's response to abrupt changes, with respect to time, in the amplitude of the input signal improved with reverse-bias voltage.

Each application has unique requirements with respect to maximum acceptable rise time and minimum acceptable 3 dB bandwidth, and it is important to remember these two parameters are related. The expressions derived in this Lab Fact provide one option for estimating the value of one parameter when the other is known.