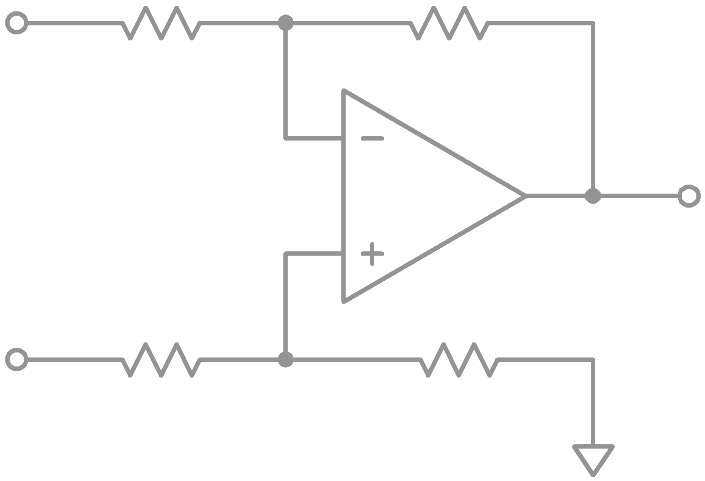
Caltech

Physics 5/105

Introductory Electronics Laboratory

2015-2016



Frank Rice

Contents

Copyright © Frank Rice 2013–2015
Pasadena, CA, USA
All rights reserved.

Introduction

Each week's experiment will require ***at least two hours of student preparation*** prior to coming to lab. Make sure you budget your time each week wisely so that you are prepared for lab! Make sure that you have ***studied the lab procedure section*** of the experiment notes before coming to lab!

The purpose of this course is to introduce you to the analysis and design of simple electronic circuits using commonly available components. We will primarily emphasize designs using analog integrated circuits (especially operational amplifiers), but you will get some exposure to simple digital logic and even designs using discrete transistors. The experiments will lead you through progressively more sophisticated concepts and techniques. They are intended to be open-ended, so the suggested procedures need not be followed exactly, and you should explore additional circuits of your own design based on the ones in the lab procedure. The course culminates in a two-week, end-of-term project of your own design which you will present to the class during finals week.

We want you to proceed at a pace which is appropriate for you, since you and your fellow students have a wide variety of backgrounds and prior experience with electronics hardware and circuits. Feel free to discuss alternative procedures and designs with the course instructor or with your TA during recitation. Please spend some time playing with circuit ideas and having fun with the devices and equipment in the lab!

Your goal should be to begin to develop a comfort with prototyping and testing your circuits and operating the sophisticated electronic test equipment used in the lab. You should start to develop an intuition concerning the art of electronic circuit analysis, design, and testing which you can continue to refine on you own after you've completed the course. If you are a graduate student taking Physics 105, then you may be mostly concerned with the design of experiment control or signal conditioning circuits you can use in your research, so we've included some particular circuits and exercises in the experiments which you may find useful.

One efficient approach to studying this material is to first examine the figures and carefully read their captions. The boxed paragraphs give important summary information useful as you start to design circuits. The text can then be used to further your understanding of the various ideas covered.

Course Structure and Conduct

Students will be assigned to lab sections of no more than 10 students each at the course organizational meeting. Each lab section will meet once a week for 3 ½ hours on a day and time to be determined at the organizational meeting (we will try to set up sections so that each meets in the afternoon, 1:00 – 4:30PM or thereabouts, but we may need a morning section).

The first 45 minutes of the lab section will be a discussion ("recitation") session led by the course instructor. During this time everyone will go over the course material for that week's lab work and review the prelab exercises found in the course notes. Students will hand in their solutions to these exercises during the discussion session.

Following a short break, each student will then perform the in-lab portion of the section which will involve the construction and testing of various circuits, including some circuits of the student's own design. Each student will work independently with occasional help from the course instructor and the section TA and assistant TA. Each student will be required to keep brief notes of the circuits constructed and the tests conducted on them, including oscilloscope screen captures, frequency response measurements, and possibly photos of the setup. Each student will maintain a lab notebook to organize these results and will turn in the notebook to the section TA for grading a couple of days following the section. Most of the lab results and conclusions should be recorded in your notebook during lab, so you should not need much time to finish them before handing in your work for grading. Your TA will provide specific instructions regarding when lab notebooks will be due and where they should be turned in.

Experiments 1 through 7 will each require one lab session (week) of work. The lab sessions begin the second week of the term (see the schedule on the first page of this handout). Experiment 8, work on which begins the last week of November, is a final project of the student's own choosing and design. All student projects will be presented during a "marathon" final lab session (combining all sections) at the beginning of finals.

Grading

Graduate students are encouraged to take the course (Ph-105) Pass-Fail!

The course grade will be based on the number of points a student achieves out of a maximum total of 100 for the entire suite of 8 experiments. The final course letter grade assignment will be on a curve, but will roughly follow the traditional scheme that 93 or above is an A, 90-92 A-, 87-89 B+, etc. The curves for the various sections will differ so that all letter grades are assigned equitably. An A+ may be awarded to one or two students whose performance is truly exceptional.

Each of the Experiments 1 through 7 is worth 10 points: 2 for the student's preparation before lab, including the solutions to the prelab exercises, performance during the recitation portion of the lab, and preparedness to conduct the in-lab experiments; the other 8 points are awarded based on the student's progress during the experiment portion of the lab, effectiveness at completing the lab tasks and measurements, and understanding of the circuits and their results as indicated by their lab notebook record.

Experiment 8 (the final project) will be worth a maximum of 30 points. The more that the circuit's design includes original work done by you, the more points you may expect; the greater the variety of concepts adapted from the previous weeks of experiments, the more points you may expect. More details concerning project grading will be provided as the time for it approaches.

Because each student starts the course with a different background and preparation in electronics and circuit design, ***the assignment of numerical grades for each experiment (including the project) will be tailored to each individual***. Hard work and notable improvement in your skills with circuit design and lab work will ensure that you get a good grade regardless of how advanced other students may appear to be.

Following the project presentations (Experiment 8), the course instructor and TAs will meet to discuss each student's final grade; at this meeting ***letter grades will be balanced and normalized among the various sections*** so that everyone is evaluated fairly and assigned the appropriate letter grade.

Late Policy

Students are required to successfully complete all 8 experiments to complete the course. If a lab session must be missed, then the student should inform his or her TA beforehand and work out a schedule for completing the missed experiment. If at all possible, the student should arrange to attend a different lab session the same week, or, as a last resort, the recitation portion of such a lab session.

You must show up to lab session on time — failure to do so will result in a loss of some credit for that experiment. Repeated late arrivals will result in an ever-increasing penalty. If you cannot turn in you lab notebook write-up for an experiment on time, you must arrange this beforehand with your TA (an email to your TA in the wee morning hours doesn't count!). Unexcused failure to turn in your lab write-up will result in a loss of some credit. Repeated failures to submit work on time will result in an ever-increasing penalty.

Collaboration Policy

Students are encouraged to study the experiment materials together in preparation for lab. Help each other to understand the concepts and circuits being presented, and discuss how to approach the prelab exercises. Each student must write up solutions to the prelab exercises independently, however, and may not consult another student's solutions when attempting to solve them. You may not refer to solutions from a previous year or solutions posted online.

During lab you are encouraged to occasionally assist your fellow students with use of the lab equipment (such as the oscilloscopes) or software. Give them a chance to use the equipment themselves, however, and to learn from their mistakes. Do not fail to complete your own work because you are "babysitting" a fellow student's efforts!

Students must complete their lab notebook write-ups independently, except for possible brief discussions of general topics to be included or general interpretation of lab data. Any questions concerning this policy should be directed to your TA or the course instructor.

The Experiments

Experiment 1: Introduction to analog circuits and operational amplifiers

Introduction to some pretty basic concepts concerning what sorts of elements make up electronic circuits and how they work together in an analog circuit design. The behavior of an ideal operational amplifier and an introduction to its applications using negative feedback.

Experiment 2: Impedance and frequency response

Introduction to the concepts of impedance and frequency response. Build simple RC filters and investigate the relationship between frequency response and transient response. Introduction to real operational amplifier limitations. Integrators and differentiators.

Experiment 3: Nonlinear circuits: diodes and analog multipliers

Introduction to nonlinear circuit elements. PN junction diodes: rectifiers, exponential and log amplifiers, temperature sensors. Applications of analog multiplier ICs including variable-gain amplifiers, frequency multipliers, analog RMS circuits.

Experiment 4: Comparators, positive feedback, and relaxation oscillators

More nonlinear circuit concepts: the comparator, a 1-bit analog-to-digital converter; The Schmitt trigger, using positive feedback to add hysteresis and improve comparator performance; the relaxation oscillator (or astable multivibrator); using the 555 timer IC to generate clocks and pulses.

Experiment 5: Resonant circuits and active filters

Using positive feedback with linear analog circuits this time — creating negative impedance and gyrator circuits and various forms of active second-order filters; investigation of a couple of common filter circuit topologies; choice of Q for use in a filter; Butterworth and Bessel filters.

Experiment 6: Transistors as amplifiers and switches

Introduction to the bipolar junction transistor as a discrete circuit element. How it works, how to properly bias the transistor and design basic amplifier and switch circuits using it as the active element. Introduction to the venerable LM311 comparator, which incorporates a transistor switch as its output element.

Experiment 7: Putting it all together: some mini-design projects

A menu of design problems from which the student must select and develop a circuit to perform the task; some additional, useful circuit fragments. Warm-up for the final project exercise.

Experiment 8: Final project

Your final project is the culmination of your efforts to learn some analog electronics and provides you the opportunity to demonstrate your new-found talents by creating something fun of your own design. You will have about ten days to build, test, and refine your circuit before presenting it to the judges — the course instructor and TAs.

Experiment 1

Introduction to analog circuits and operational amplifiers

Copyright © Frank Rice 2013–2015
Pasadena, CA, USA
All rights reserved.

Experiment 1

Introduction to analog circuits and operational amplifiers

Electronic circuit design falls generally into two broad categories: *analog* and *digital* (a third category, *interface* circuitry, includes hardware to join these two major circuit realms). Digital circuitry, as you probably already know, uses electronic components and systems to represent and store numerical data and to perform algebraic and logical operations on the data. Incredibly complicated digital structures are created by combining a few simple circuit building blocks (such as *registers*, *gates*, and *clocks*) into vast networks of components.

Analog circuitry, in contrast, is used to respond to continuously-variable electrical signals from sensors (such as microphones, thermistors, antennas, and accelerometers) or to provide continuously-variable control signals to actuators (such as loudspeakers, heaters, antennas, or motors). Analog circuitry is used, for example, to connect digital computers and circuits to many of the physical devices they use for data I/O and storage. Our focus in this course will be mainly on analog designs.

Probably the most important application of analog electronic circuitry (especially in the sciences) is to *amplify* and *filter* the power of a minute signal so that it can be accurately measured or used to control something interesting. In this experiment you'll jump right in and start designing and building simple, useful amplifier circuits using that truly marvelous, tiny building-block of modern analog electronics: the *operational amplifier*.

Before you can begin to understand how to construct such amplifiers, you must understand some pretty basic concepts concerning what sorts of elements make up electronic circuits and how they work together in a design. The first several pages that follow cover these basic ideas; the information may be dry and dense, but it is important that you read it! Hopefully much of the following section is a review of what you already know, but, if not, familiarize yourself with the content so you can quickly refer back to it when the time comes.

Next the text covers the behavior of an *ideal operational amplifier* (op-amp) and why this behavior makes it so versatile and easy to use by incorporating *negative feedback*. Finally we cover some additional concepts which will prove useful as you analyze and design circuits. The *Prelab Exercises* will test your understanding of this material and give you some practice in preparation for your lab session. Complete them and come to your recitation session ready to ask (and answer!) questions about the material.

One auxiliary, but very important, objective of this experiment is to give you some initial experience using the lab electronic equipment: signal generator, oscilloscope, cabling, and, of course, the analog trainer/breadboard you use to construct your circuits.

CIRCUIT BASICS

Current, voltage, power

Current is the flow of electric charge from place to place. Electronic circuitry employs networks of narrow, highly conductive elements (copper wiring or traces on printed circuit boards) to effectively confine the flow of charge to well-defined paths. Current is defined as the measure of the rate of charge flow through a surface (typically the cross section of a wire) and is measured in the SI unit *ampere* (amp, or A). The SI unit of charge is the *coulomb*, which is defined such that $1\,\mathrm{A} = 1\,\mathrm{Coulomb/second}$. An amp is a very large current for small, table-top circuit designs; our electronic circuits will have currents of about 10^{-7} amp to 10^{-2} amp, so we'll most often be dealing with currents of *microamps* (uA or μA) to *milliamps* (mA).

Currents are generated by the motions of charge carriers in the circuit in response to electromotive forces induced by electromagnetic fields. As a charge carrier moves about, its potential energy due to the fields varies. The work done on the charge by the fields is equal to the reduction in potential energy of the charge as it changes position. The potential energy per unit charge due to an electric field is called the *electrostatic potential* (or just *potential*) and is measured in the SI unit *volt* ($= 1\,\mathrm{Joule/Coulomb}$). We have various ways of establishing either steady or time-varying potentials in our circuits: *power supplies*, *signal generators*, and *batteries*. These devices also serve as sources and receivers of charge carriers, so that the circuit to which they are connected remains electrically neutral (no net charge). In our circuits maximum voltages are no more than ~12V, and signals have amplitudes of a fraction of a *millivolt* (mV) to a few volts.

Assume a circuit has charge carriers flowing steadily from a point A to a point B at potentials v_A and v_B. If the current flowing is i_{AB}, then the power being expended by the source of the potential difference must be $P = (v_A - v_B)\,i_{AB}$. If the potentials and the currents are time-varying, but the instantaneous current out of A remains equal to that arriving at B, then the current between the points remains a well-defined function of time, and we have the instantaneous power: $P(t) = (v_A(t) - v_B(t))\,i_{AB}(t)$. Our circuits will have power flows on the order of a few tenths to about a hundred *milliwatts* (mW).

A component which can continually add power to a circuit is called an *active element*. Other components (most of which dissipate power or otherwise remove it from the circuit) are called *passive elements*.[1]

[1] Electrical engineers often refer to semiconductor devices such as transistors and integrated circuits as *active elements*, because these devices can transfer power from a steady power source into a signal circuit.

Frequency, wavelength, lumped circuit elements

Electromagnetic fields propagate at the speed of light, 30 cm/nanosecond (or about a foot/nanosecond). One nanosecond is the period of a signal oscillating at a gigahertz (GHz, 10^9 hertz). The maximum frequencies we'll be using in our circuits are no more than a few megahertz (MHz) or kilohertz (kHz), so the wavelengths of these fields will usually exceed hundreds of meters, hundreds to thousands of times bigger than the physical sizes of our circuits. Consequently, but maybe not so obviously, each individual element in a circuit will have, to a high degree of accuracy, no change in its total net charge as the fields oscillate. Thus we may safely assume that there is *zero net total current flow into or out of all of an element's connections* to the circuit at any instant. Such a component is called a *lumped element*. Examples of lumped elements are the resistors, capacitors, LEDs, and integrated circuits (ICs) we'll be using.

In the high-frequency case, where wavelengths become comparable to the size of a component, the fields and currents may vary across it, making it a *distributed element*, and our assumption above is no longer valid. Examples of distributed elements include antennas, microwave waveguides, the motherboard in your computer or tablet, and the national electrical power grid.

Most of the elements we use in our circuits — resistors, capacitors, inductors, diodes, batteries, etc. — have two terminals for connections to circuit conductors, making them *lumped, two-terminal elements* (note that any lumped element will have *at least* two terminals, since the total current flow into the element's connections must vanish, as described above). Pictured in Figure 1-1 is a selection of symbols for typical two-terminal elements as used in an electrical circuit drawing, called a *circuit schematic*. Because the currents at the two terminals must be equal and opposite (flowing in at one terminal and out at the other, so the total net current into the element is zero) we can simply refer to *the current flowing through the element and the potential difference (voltage) across it*.

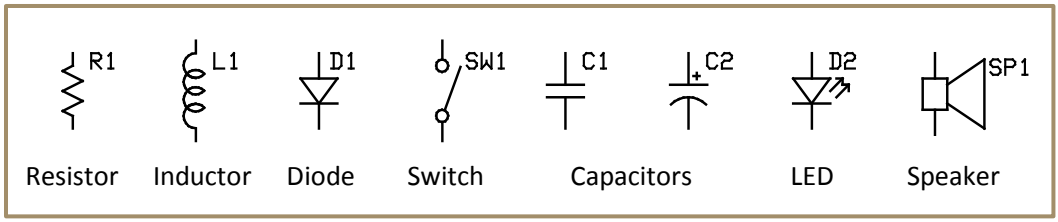


Figure 1-1: Schematic symbols for a selection of two-terminal, passive circuit elements. Also shown with each symbol is an example of an associated *reference designator* to uniquely identify that component in a circuit.

Sources and signals

In most cases the independent variables in the set of equations we will write out to describe the behavior of a circuit are a few voltages or currents we control as *inputs* to the circuit; the

equations then allow us to determine the circuit's *outputs* in response to the inputs. The independent inputs are termed *sources* or *signals*, and may be generated by batteries, power supplies, signal generators, microphones, antennas, thermocouples, or whatever else we can think of which produces a potential (*voltage source*) or injects a current (*current source*) into our circuit. Sources are usually described as two-terminal devices in our circuits and are *active elements* since they inject power into the circuit. If a source produces a constant output (such as a battery or power supply), then it is called a *DC* source (for "direct current"). If its output is sinusoidal, then it is an *AC* source or signal (for "alternating current").

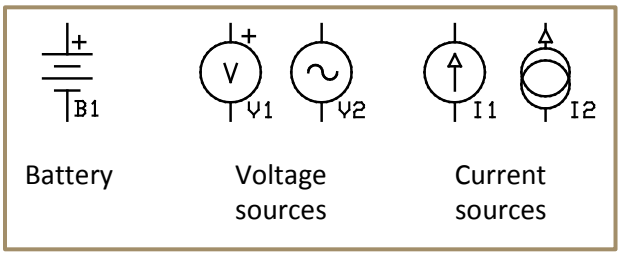


Figure 1-2: Schematic symbols for common voltage and current sources; each is a two-terminal, active element. Polarity or current flow has the direction indicated when the source output is positive. The battery symbol will often be used to represent any constant-voltage (DC) source such as a power supply. The other sources may be constant or time-varying (an "AC source" usually means that its output varies sinusoidally – V2 in the figure represents an AC voltage source).

The schematic symbols for some common active sources are shown in Figure 1-2. These sources are considered to be *ideal*, in the sense that each can maintain its specified output voltage or current regardless of what they may be connected to in the circuit and how much power they must supply. Of course, real sources are not quite so capable! The polarity or current direction included with the schematic symbol shows the relative potential or current flow when the source output has a positive value.

Grounds and power supply terminals

The fields in our circuits produce potential differences and corresponding current flows. Because only potential differences are significant, we will pick one convenient point in a circuit and *define it as having zero potential* (0 Volts); all other voltages in the circuit will be measured or specified with respect to this point. The symbol used in this text for the 0-Volt reference point is a triangle: ⏚ . We refer to this point as the *circuit ground*, and a terminal connected to this point is said to be *at ground potential* or to be *grounded*.

We will use a couple of conventions when drawing our circuit schematics which will considerably reduce the clutter in them. The amplifiers you will build will require DC power from a power supply or batteries in order to operate properly, so these constant-voltage power sources must be indicated in our circuit drawings. Additionally, often several components in the circuit (including the power supply) will have terminals connected together and to the 0-Volt reference point (ground). A power supply schematic drawing

simplification is shown in Figure 1-3; instead of explicitly showing the power supply source symbols, we'll put little arrow symbols with their associated voltages at whatever points need to be connected to the power supply (as shown on the right in Figure 1-3). Wherever these symbols appear, you must remember that physical wiring connects all symbols with the same voltage to the appropriate terminal of the power supply source. Similarly:

> Multiple ground symbols appearing in a schematic *are all implicitly connected together* and are connected to the appropriate power supply return terminal, which also serves as the 0-Volt reference point for the circuit (as shown in Figure 1-3).

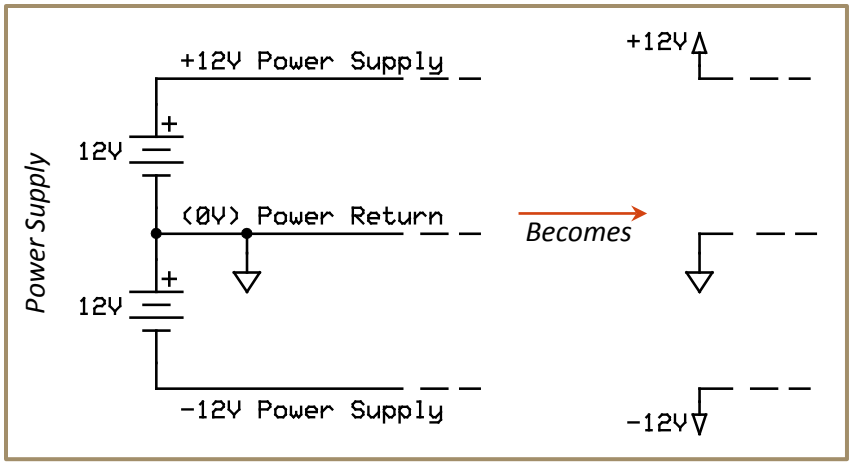


Figure 1-3: Simplified schematic diagram depiction of a power supply. Multiple arrow symbols labeled with the same voltage (such as +12V) may appear in a diagram. All must be considered to be connected together and to the appropriate power supply source terminal. Similar considerations apply to the appearance of multiple ground symbols in a diagram. One power supply terminal is connected to ground as well, as shown above. The voltage labels always give the voltage values with respect to ground, which is the 0-Volt reference point.

Resistors

The two most common two-terminal, passive elements we will use are the *resistor* and the *capacitor* (another common element, the *inductor*, is mostly used in radio-frequency circuits; we'll discuss the capacitor and inductor in later experiments). These elements are important to study first, because each of them has a simple, linear relationship between the voltage across and current through it (at least to a very good approximation, for low frequencies and voltages). You will generally find that your designs will contain more resistors and capacitors than the combined total of all of the rest of the components in the circuits.

Let's start with the resistor and its famous *Ohm's law*. Figure 1-5 on page 1-6 shows a variety of resistors with different physical sizes, shapes, uses, and power ratings; the ones you will use when "bread-boarding" prototype circuitry will look like the small, light-brown

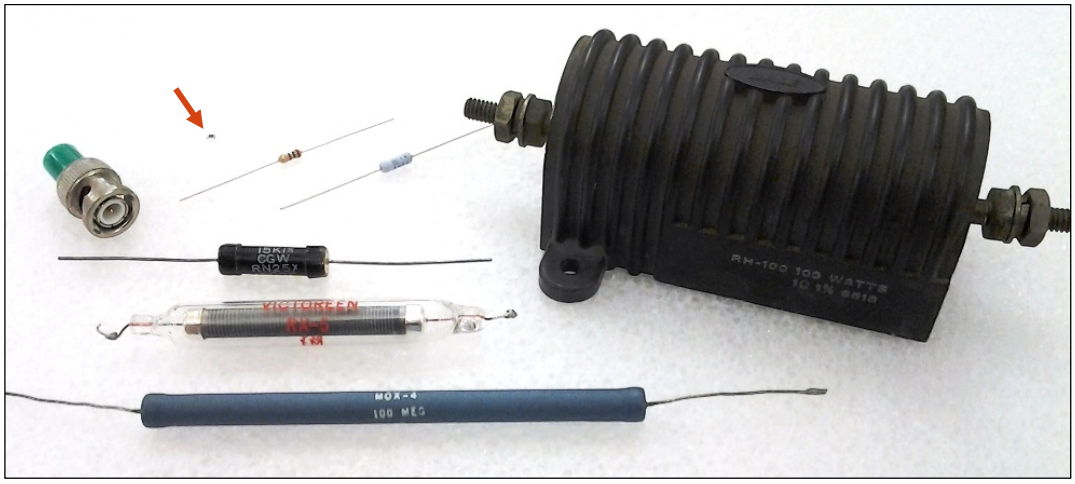


Figure 1-5: A variety of resistors, from the tiny (.08 inch × .05 inch) surface-mount component barely visible near the upper left corner of the photo (see arrow) to the beefy, 100 Watt power resistor to the upper right. The long resistor in front is designed to be used with very high voltages.

resistor towards the upper left in the photo. If we connect a resistor to a voltage source (a power supply or signal generator) as shown in Figure 1-4, then, obviously, there will be a voltage across the resistor and a current flowing through it. Figure 1-4 shows the usual conventions for the polarity of the voltage and direction of the current flow when the potential $v(t) > 0$.

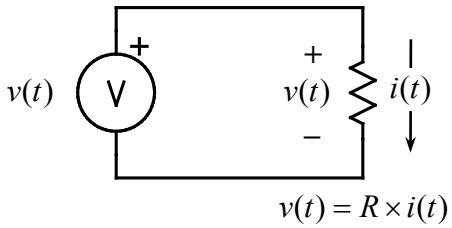


Figure 1-4: A simple circuit illustrating Ohm's Law for a resistor with value R.

Since the lines connecting the elements' terminals in a schematic diagram are considered to be perfectly-conducting wires, terminals connected by a wire in the diagram must have identical voltages, and current will flow through the connection without any change. Thus, if the output voltage of the source in Figure 1-4 is $v(t)$, then that is also the voltage across the resistor R. Ohm's law states that the voltage $v(t)$ across a resistor and the current $i(t)$ through it are strictly proportional at any instant:

Ohm's law for a resistor

1.1

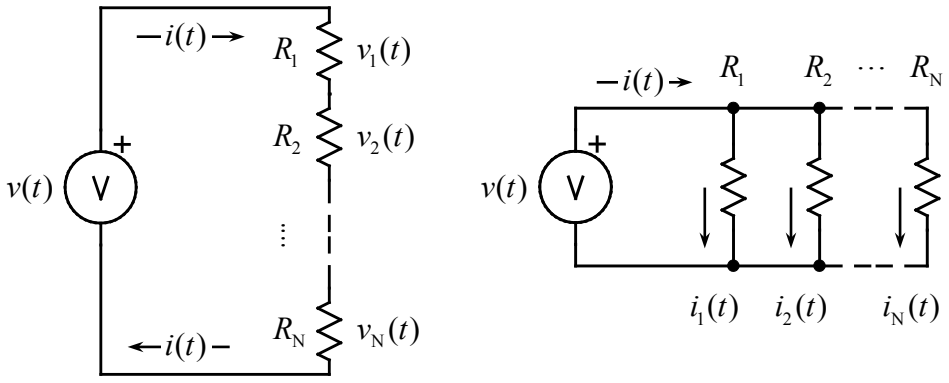
$$v(t) = R\, i(t)$$

The constant of proportionality, R, is called the *resistance* of the resistor and has SI units of *Ohms*: 1 Ohm = 1 Volt/Ampere. For an ideal resistor R is a real, constant number

(independent of voltage, current, or frequency) with $R \geq 0$. Equation 1.1 is the *defining relation* for an ideal resistor; the actual resistors you will use behave in a very nearly ideal manner. The circuits we build will have voltages mostly in the range of a few tenths to a few volts, and currents of a few microamps to a few milliamps. Thus the useful range of resistor values is a few hundred to a few million Ohms (MΩ). Commonly used values should be 10^3 to 10^5 Ohms (1 to 100's of kΩ).

The elegance of the lowly resistor lies in Ohm's Law, equation 1.1: a resistor is a current to voltage converter (and vice versa)! If we have a wire with, say, a few milliamps of current flowing in it, then inserting a 1.0 kΩ resistor (kΩ: kilo-ohm =10^3 Ohms) in the circuit *will produce a voltage across it of 1.0 V/mA, faithfully following any variation in the current flow*. Conversely, if we put a known voltage across a resistor, then we can immediately calculate the resulting current flow through it. These facts will be of enormous importance when you design and analyze various amplifiers during this course!

Series and parallel resistors



(Series) (Parallel)

Figure 1-6: Series and parallel resistor combinations. For the series case, the same current flows through all resistors, and the total voltage is the sum of the individual resistor voltages. For the parallel case, the voltage is the same across all resistors, whereas the total current from the source is the sum of the individual resistor currents. The results are that series resistances add; parallel conductances (1/resistance) add.

A common situation when analyzing part of a circuit is to find that two or more resistors (or other components) are connected in *series* or in *parallel*, as illustrated in Figure 1-6. The problem is to determine the *equivalent resistance* of such a combination, that is, the relationship between the voltage and current across the entire array of components. The solution is straightforward once you recognize that:

- *Series combination*: the same current must flow through all resistors, and the total voltage across them must be the sum of the individual resistor voltages.
- *Parallel combination*: the same voltage appears across all resistors, and the total current must be the sum of the individual resistor currents.

Consider the series combination first. Referring to Figure 1-6, the voltage across the k^{th} resistor is $v_k(t) = R_k\, i(t)$, so the total voltage is $v(t) = i(t)\sum R_k = R_{\text{series}}\, i(t)$, where R_{series} is the equivalent series resistance we seek. In the parallel case, $v(t) = v_k(t) = R_k\, i_k(t)$, so $i_k(t) = v(t) / R_k$. The total current is therefore $i(t) = v(t)\sum(1/R_k) = v(t) / R_{\text{parallel}}$, where R_{parallel} is the equivalent parallel resistance. The equivalent resistance (resistance seen by the source) for each case is then:

1.2

$$\text{Series combination:} \quad R_{\text{series}} = \sum_{k=1}^{N} R_k$$

$$\text{Parallel combination:} \quad \frac{1}{R_{\text{parallel}}} = \sum_{k=1}^{N} \frac{1}{R_k}$$

The reciprocal of resistance is called conductance, and has the SI unit *Siemens* ($= \text{Ohm}^{-1}$).

Series resistances add. Parallel conductances add.

Using the second of equations 1.2 for the case of only two resistors, we can derive the more familiar equation for two parallel resistors,

$$R_{\text{parallel}} = \frac{R_1 R_2}{R_1 + R_2}$$

which is less elegant than the more general expression in 1.2.

Now let's ask what the voltage drop across any particular resistor would be for our set of series resistors in Figure 1-6. The voltage across the k^{th} resistor is $v_k(t) = R_k\, i(t)$, and $i(t) = v(t)/R_{\text{series}}$, so we get expression 1.3, which may seem pretty obvious:

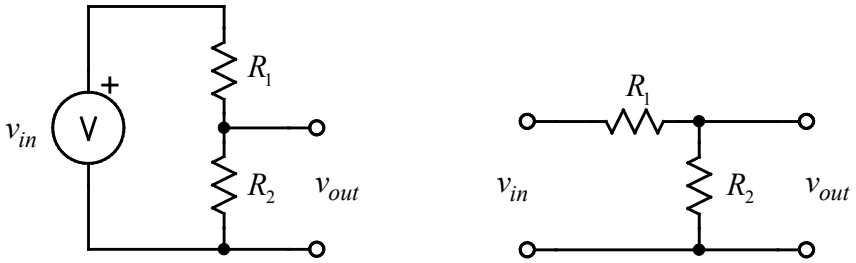
1.3

$$\frac{v_k(t)}{v(t)} = \frac{R_k}{R_{\text{series}}} = \frac{R_k}{\sum R}$$

Thus the ratio of the voltage across one resistor to the total voltage is just the ratio of the resistor's value to the total resistance in the series. This result segues nicely into our next topic.

Voltage dividers and the potentiometer

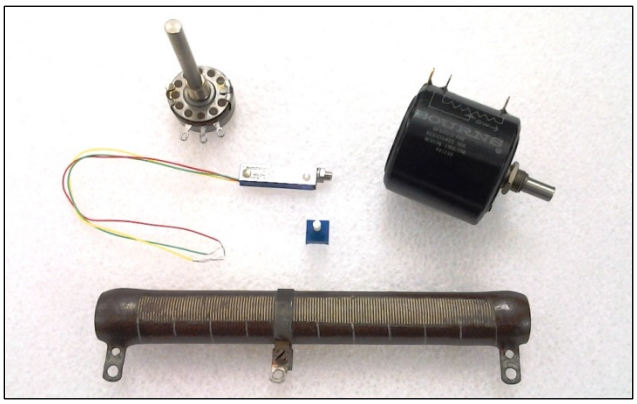
The result 1.3 for a series-connected pair of resistors provides the solution for the behavior of the very common *voltage divider* circuit, Figure 1-7.



$$\frac{v_{out}}{v_{in}} = \frac{R_2}{R_1 + R_2}$$

Figure 1-7: The *voltage divider* is just a series combination of two resistors. The *input* voltage, v_{in}, is the source voltage applied across both resistors; the *output* voltage, v_{out}, is the voltage across the bottom resistor of the pair. The gain of the divider is defined as $G = v_{out}/v_{in}$, and is trivial to derive from equation 1.3. The circuit on the right is the same voltage divider, but assumes that a source of voltage is applied between the two input terminals rather than explicitly showing the source.

The voltage divider configuration of two resistors (or other types of elements) will show up again and again in the circuits we will design and build. It would be wise for you to commit the schematics and the formula in Figure 1-7 to memory!



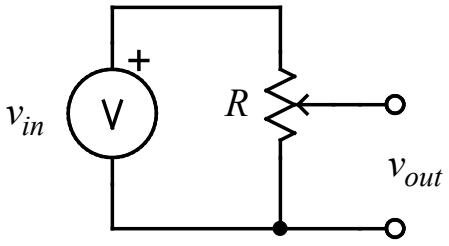


Figure 1-8 (left): A selection of potentiometers. All except the one at the bottom of the photo are adjusted by turning a shaft. The bottom device clearly shows how a potentiometer is constructed, in this case from a long coil of high-resistivity wire, such as nichrome. The two ends of the resistive element are attached to terminals on the device; the adjustable wiper contact is attached to the remaining terminal.

Figure 1-9 (right): A potentiometer with total resistance R used as a variable voltage divider. As the wiper position is adjusted from bottom to top, v_{out} varies from 0 to v_{in}.

Often we will need a voltage divider with a gain (v_{out}/v_{in}) which is easy to adjust. The *potentiometer* is a circuit element designed for just this job! A potentiometer (or, sometimes, rheostat or variable resistor) is a resistor made from a relatively long, partially-exposed resistive element. An electrical contact (wiper) may be moved along the exposed resistive element from one end to the other, varying the ratio of the resistance between the wiper and one end of the resistor to the total resistance (see Figure 1-8).

Figure 1-9 shows how a potentiometer may be used as a variable voltage divider. In this circuit the output voltage may be varied from 0 all the way up to the input voltage. This is how a volume control is implemented in many audio devices, with v_{in} representing an audio signal (the "wiper position" in a modern device is usually controlled using a digital logic circuit). Other useful potentiometer circuits are shown in Figure 1-10; more examples will come up throughout the course.

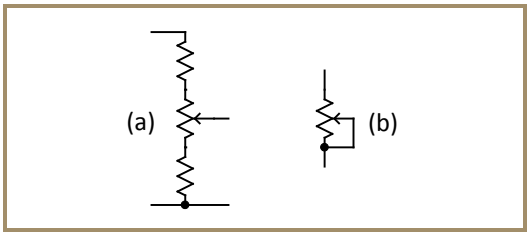


Figure 1-10: Other potentiometer circuits.
(a) adding a resistor in series at either end so that v_{out} varies over a smaller range for finer control
(b) using a potentiometer as a variable resistor

Networks, ports, gain

In Figure 1-7 (on page 1-9) the right-hand voltage divider sub-circuit (portion of a larger circuit) is our first example of a *two-port network*, a common general construction useful for analyzing circuits. Other important examples of two-port networks include amplifiers and filters. Consider Figure 1-11, which explicitly shows how we identify our voltage divider as a network with a single *input port* and a single *output port*. Each "port" of a network is comprised of two terminals which are meant to be connected into some larger, surrounding circuit. This surrounding circuit will, in general, inject signals (voltages and currents) at our network's input ports, and will it respond to signals emitted from the network's output ports.

Figure 1-11: The voltage divider is a form of *two-port network*. It has a pair of input terminals, the *input port*, and two output terminals: the *output port*. In this example, the network's transfer function (or gain function) is the ratio v_{out}/v_{in}, as already mentioned in Figure 1-7, page 1-9.

This network concept is useful because we can often describe its behavior (as far as the external, surrounding circuit is concerned) with just a few equations or parameters and

otherwise ignore its detailed internal construction. The network becomes a "black box" with inputs and outputs whose relationships are known, but we don't have to bother with the messy details of what's inside the box. This is exactly how we are going to handle operational amplifiers in the next section!

The most important parameter we will use to describe a "generic" two-port network is its *gain*, or, more generally, its *transfer function*, which describes the functional relationship between its output and its input.

If the network is *linear* (as is the voltage divider) the gain function becomes a simple, fixed ratio independent of the size of the input. For example, the gain of the voltage divider (v_{out}/v_{in}) depends only on the resistors' values and is therefore independent of the magnitude of v_{in} (see Figure 1-7). Other important network parameters (discussed in a later section) are its *input impedance* and *output impedance*.

One final comment about the depiction of a network and its ports: often, one terminal of a port will be directly connected to *ground* (the 0-Volt reference) somewhere inside the network's circuitry. In this case it is common to show only one terminal for that port, the ground connection being understood as the port's other terminal, as shown in Figure 1-12. Since ground is the 0-Volt reference, the voltage of the one explicitly-depicted terminal is also the voltage present at the port (right-hand circuit in Figure 1-12). If neither terminal is actually connected internally to ground, as in Figure 1-11, then the voltage at the port is the difference in the voltages of its two terminals.

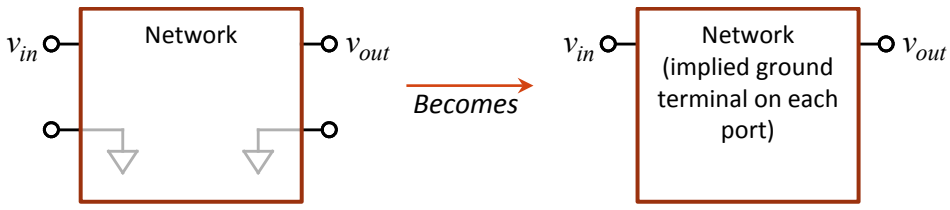


Figure 1-12: A network whose ports each use *ground* as one terminal. In this case the ground terminal is often not shown explicitly, so only a single terminal is shown for such a port. Since ground is our circuit's 0-Volt reference, the voltage of the single terminal is sufficient to determine the voltage of the port.

THE OPERATIONAL AMPLIFIER

The ideal op-amp

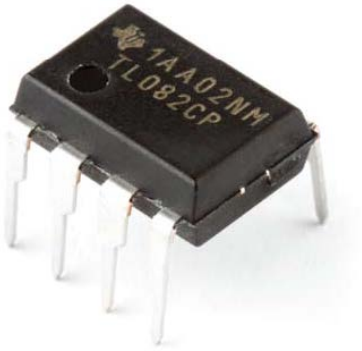


Figure 1-13: A photo of an IC operational amplifier, the Texas Instruments Inc.'s TL082 device. This integrated circuit is the type you'll use for your circuits in this experiment; it actually contains two independent op-amps in the package shown (called an 8-pin *DIP*, for "dual-inline package"). The spacing of the connector pins in one row is 0.10 inch (2.54 mm); the two rows are 0.30 inch apart. (Photo courtesy Texas Instruments Inc., ©2012)

The most important single element we'll use for our analog circuit designs is the *operational amplifier* (op-amp). Modern operational amplifiers are examples of analog *integrated circuits* (ICs), wherein an entire network of dozens (or even hundreds) of transistors, resistors, and even capacitors is created on a single small silicon wafer. The wafer is then mounted inside a (usually) plastic package with several external metal pins used to make electrical connections to the wafer's circuitry (Figure 1-13).

Modern IC op-amps are the culmination of decades of improvements and innovations by hundreds of electrical engineers at dozens of companies; they have outstanding linearity, gain, bandwidth, and noise performance (these terms will mean more to you as the course goes on). Because of this spectacular performance available for our designs, we will first learn how to design amplifier circuits using an *ideal operational amplifier*, which is, naturally, an idealization of an actual op-amp's behavior. As you'll discover when you get to the lab, your real op-amps will perform so well that your results will approach very closely to this ideal!

You should envision our ideal operational amplifier to operate like the "cartoon" diagram in Figure 1-14; right now we will consider the op-amp to be a 4-port network. There is an input port whose two terminals are labeled *+Input* and *–Input* (with voltages of v_+ and v_-, respectively). You should think of these two terminals as being connected to a *perfect voltmeter* inside the op-amp; this voltmeter measures the voltage difference between the two terminals as indicated in Figure 1-14. By "perfect" we mean that the voltmeter is sensitive to any tiny difference in the two terminals' voltages, and that *it does not draw any current at all*

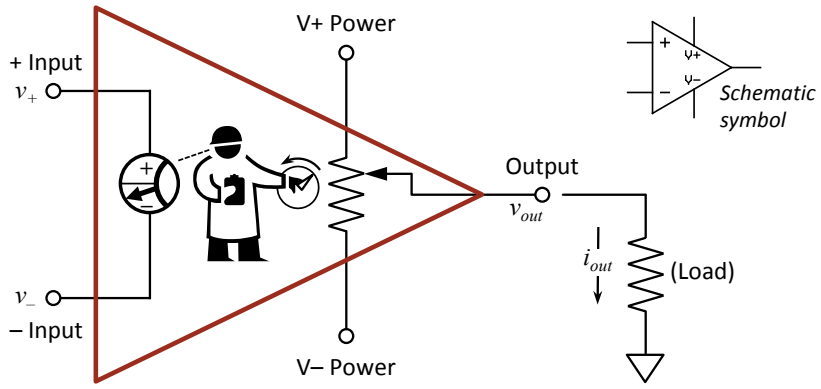


Figure 1-14: Cartoon illustrating the ideal op-amp's behavior. A "voltmeter" monitors the potential difference between the + and – inputs; if the two input voltages match, then the output voltage remains unchanged. If there is a voltage difference between the two inputs, then the op-amp *very quickly* changes the output voltage v_{out}: increasing v_{out} if $v_+ > v_-$, decreasing it if $v_+ < v_-$. The output stops changing only when the two input voltages again match (or when the "output potentiometer" has been adjusted all the way to one of its limits: a power supply terminal voltage). Since the source of the output comes from the op-amp's two power supply terminals, electrical power required by the output load comes from the op-amp's power supply, *not the inputs*. Output current supplied by the op-amp is returned to the power supply through the load, as shown (via the ground attached to other terminal of the load).

from whatever circuit is connected to an input terminal — each input has *infinite input impedance*.

The ideal op-amp also has two power supply ports: the *V+ Power* and the *V– Power* terminals and their associated, *implicit* ground terminals (implicit ground terminals like those in Figure 1-12 on page 1-11). The *V+ Power* and *V– Power* terminals will always be connected to a DC power supply for our circuits; the lab's circuit design trainer has *+12V and –12V power with a common ground connection* as shown way back in Figure 1-3 on page 1-5. You should think that inside the op-amp these two power terminals are connected to either end of a potentiometer as shown in Figure 1-14; the potentiometer's wiper is connected to the op-amp's *Output* terminal (the *Output* port also has an implicit ground terminal).

Now think of some little "technician" ensconced inside our ideal op-amp whose only job is to watch the input voltmeter and move the potentiometer's wiper depending on what the meter shows. If the *+Input* and *–Input* terminals have exactly the same voltage (so the voltmeter reads 0), then the technician stops moving the wiper or leaves it where it is; *the output terminal is thus going to be at some constant voltage between those of the V+ Power and the V– Power terminals*. If there is a voltage difference shown by the input voltmeter, then the technician starts moving the potentiometer wiper *very rapidly* — toward *V+ Power* if $v_+ > v_-$, or toward *V– Power* if $v_+ < v_-$ (see Figure 1-14). For the ideal op-amp, *the output voltage will change infinitely quickly as long as the two input terminals are at different voltages*. And that's it! That's all that the ideal op-amp is supposed to do!

THE IDEAL OPERATIONAL AMPLIFIER'S CHARACTERISTICS AND BEHAVIOR

- The two input terminals (*+Input* and *–Input*) draw 0 current from the external circuit (they each have infinite input impedance).
- The Output voltage is constant whenever $v_+ = v_-$.
- The Output voltage *increases* infinitely quickly whenever $v_+ > v_-$.
- The Output voltage *decreases* infinitely quickly whenever $v_+ < v_-$.
- The power for the Output comes from the *V+ Power* and *V– Power* terminals.

The voltage follower and negative feedback

So what can we do with such a thing as our ideal op-amp? First consider a very simple circuit known as the *voltage follower* (Figure 1-15). It is seemingly trivial — the op-amp's output is connected back to its *–Input* terminal (so that v_- will always equal v_{out}), and some sort of *input signal* is connected to the op-amp's *+Input*, so $v_+ = v_{in}$.

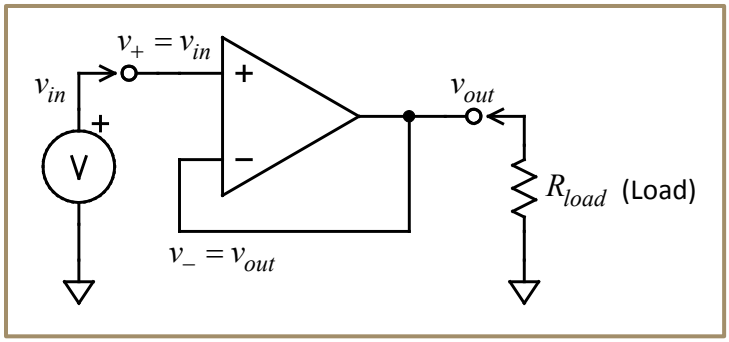


Figure 1-15: The *voltage follower* amplifier circuit. Note that the op-amp's *–Input* terminal is connected directly to its *Output* terminal, whereas the input voltage source is connected directly to the *+Input* terminal (the op-amp's two power supply terminals are not shown, but they still must be connected to a power supply!). If $v_{out} = v_{in}$, then $v_+ = v_-$, and the op-amp maintains the output voltage, v_{out}. If the input voltage v_{in} changes, then momentarily $v_+ \neq v_-$, and the op-amp rapidly changes v_{out} *in the same direction as the change in* v_{in} until the condition $v_{out} = v_{in}$, is restored. Thus the amplifier always keeps $v_{out} = v_{in}$, so its gain $G = 1$.

If $v_{in}(t)$ is actually constant, then clearly an equilibrium condition for the op-amp would be $v_{out} = v_{in}$, because then $v_- = v_+$, and v_{out} would remain constant. But what if $v_{in}(t)$ changes or there is some perturbation in v_{out} so that, momentarily at least, $v_{out} \neq v_{in}$? Because we have connected v_{out} to the *–Input* terminal (not *+Input*!), then if, for example, $v_{out} < v_{in}$, we would also have $v_+ > v_-$, so the op-amp would quickly increase v_{out} until the condition $v_{out} = v_{in}$ is restored. Similarly, the op-amp would correct the opposite condition, $v_{out} > v_{in}$.

Thus it would always be the case that the voltage follower circuit (Figure 1-15) will maintain $v_{out}(t) = v_{in}(t)$, so the voltage gain of this simple amplifier is $G = v_{out}/v_{in} = 1$.

Make sure you study Figure 1-15 in light of the ideal op-amp behavior (box on page 1-14) until you have convinced yourself that the voltage follower will always maintain $v_{out}(t) = v_{in}(t)$ (unless v_{in} exceeds the limits set by the op-amp's power supply voltages, which determine the maximum range of v_{out}).

The voltage follower circuit has a stable, linear relationship between v_{out} and v_{in} because the op-amp's output voltage is connected back to its *–Input* terminal. This arrangement is an example of *Negative Feedback*, which is the secret to the versatility of the op-amp.

How could an amplifier with a gain of 1 add any value to a system? Actually, you will find this circuit to be very useful and will possibly include one or more voltage followers in your final project design. The reason it is so useful is because the ideal voltage follower amplifier has *infinite power gain*: Since the input source is connected only to the op-amp's *+Input* terminal, which draws no current, the power required from the input source is $P_{in} = v_{in}\, i_{in} = v_{in} \cdot 0 = 0$, whereas the power delivered to the load attached to the amplifier's output is $P_{out} = v_{out}\, i_{out} = v_{out}^2 / R_{load} > 0$: the power gain $P_{out}/P_{in} \to \infty$. The current drawn by the amplifier output's load is supplied by the op-amp's power supplies, *so no power is required from the input source*.

The noninverting op-amp amplifier

Now that you understand how the voltage follower works, let's design a more general and flexible negative feedback setup using our ideal op-amp. Consider the circuit in Figure 1-16, where we now use a voltage divider consisting of resistors R_f and R_i to feed back only a fraction of v_{out} to the *–Input* terminal (assume that a source voltage and a load resistor are again connected to our amplifier like those in the voltage follower circuit, Figure 1-15). The

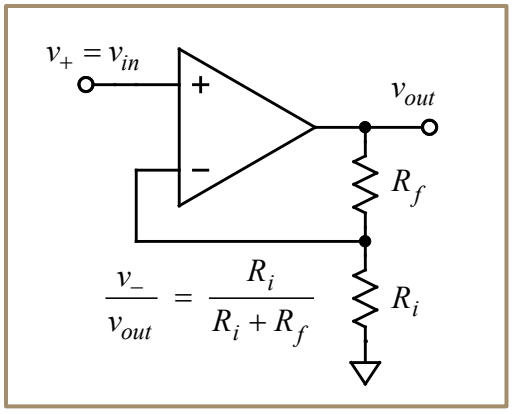


Figure 1-16: The general *noninverting amplifier* circuit. Now the op-amp's *–Input* terminal is connected to the output via a simple voltage divider circuit, so only a fraction of v_{out} is used as the *negative feedback* signal. As with the voltage follower (Figure 1-15), the input voltage source is connected directly to the *+Input* terminal. When the equilibrium condition $v_+ = v_-$ obtains, v_{out} will be larger than v_{in} by the factor $G = (R_i + R_f)/R_i = 1+(R_f/R_i)$, which is thus the *gain* of this amplifier.

op-amp's output voltage, v_{out}, will be stable when $v_- = v_+$, as before. We still have $v_+ = v_{in}$, but now we must use the voltage divider equation (see Figure 1-7 on page 1-9) to determine v_- from v_{out}; the resulting relation is shown in Figure 1-16. So at equilibrium v_- is smaller than v_{out} by a ratio determined by the voltage divider, and, since $v_- = v_+ = v_{in}$, we see that v_{in} must be smaller than v_{out} by this same ratio. Thus we now have an amplifier with a voltage gain G of whatever we want it to be (although $G \geq 1$): we just choose an appropriate pair of values for the resistors R_f and R_i (see equation 1.4) The amplifier is referred to as *noninverting* because v_{out} has the same sign as v_{in}.

Ideal, noninverting amplifier gain

1.4

$$G = \frac{v_{out}}{v_{in}} = \frac{R_i + R_f}{R_i} = 1 + \frac{R_f}{R_i}$$

Note that some current from the op-amp's output must flow to ground through the noninverting amplifier's voltage divider. *This current demand upon the output will add to the current required by the amplifier's load.* In this case the current through our feedback network will be $I_f = v_{out}/(R_f + R_i)$, since the two resistors are in series (remember, 0 current flows to the ideal op-amp's *–Input* terminal); using larger values for these resistors will reduce the required feedback current.

The inverting op-amp amplifier

Let's start with the noninverting amplifier of Figure 1-16, but instead of connecting the input signal to the op-amp's *+Input* terminal, let's connect it to the bottom end of the feedback voltage divider; we then ground the *+Input* terminal. The result is the *inverting amplifier* circuit of Figure 1-17 (note that we've rearranged the locations of the components and

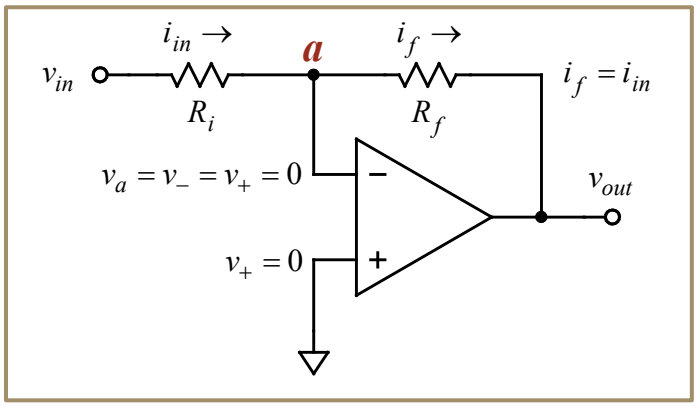


Figure 1-17: The general *inverting amplifier* circuit. The input signal is attached to one end of the negative feedback voltage divider, and the *+Input* terminal is connected to ground. At equilibrium $v_+ = v_-$, and the point *a* becomes a *virtual ground*, since its voltage will be 0 as well. Since no current flows into the op-amp's *–Input*, the current from the input source (i_{in}) must also flow through R_f to the op-amp's output terminal. The voltage across R_i is just v_{in}; the voltage across R_f must be (R_f / R_i) times larger, because the currents are the same. Thus the gain $G = -(R_f / R_i)$.

flipped the op-amp symbol so that the *–Input* terminal is above the *+Input*). You should closely compare the two circuits to convince yourself that the only change has been to swap the ground and source input connections. The inverting amplifier is a little harder to analyze, but there are a couple of clever shortcuts you can use to quickly derive the gain. These "tricks" are quite useful when analyzing op-amp circuits, so let's carefully consider them.

First, we assume that the negative feedback works effectively, so that the equilibrium output condition $v_- = v_+$ is maintained. This implies, as shown in the figure, that the voltage at *node* (connection) *a*, where the two resistors are joined to the *–Input*, must be 0 (ground), *even though there is no direct connection of this point to ground*. For obvious reasons, therefore, the node *a* is called a *virtual ground*. Now we know the potential across resistor R_i: $v_{in} - 0 = v_{in}$, so we immediately know that $i_{in} = v_{in} / R_i$.

Now comes the second trick: since no currents flow into an ideal op-amp's input terminals, the only place for the current i_{in} to go is to continue on through R_f, so $i_f = i_{in}$, and we now know the potential across R_f: $R_f i_f = 0 - v_{out} = -v_{out}$. Thus we have derived the gain for the *ideal, inverting amplifier*:

Ideal, inverting amplifier gain

1.5

$$G = \frac{v_{out}}{v_{in}} = -\frac{R_f}{R_i}$$

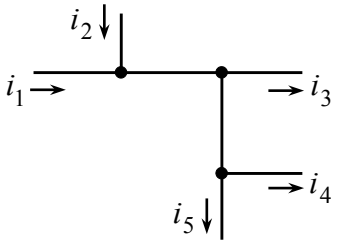
The amplifier is call *inverting* because the sign of the output voltage is the opposite of the sign of the input voltage. The magnitude of this circuit's gain may be chosen to be anything by picking values for R_f and R_i, whereas the gain of the noninverting amplifier must be at least 1 (compare equation 1.4). There is one significant drawback of the inverting amplifier circuit, however: *the current drawn from the input source is not zero*. In other words, the *input impedance* (v_{in} / i_{in}) of this amplifier is finite — in fact, it is equal to R_i. The concept of input impedance will be discussed in the next section, although we'll forego the general definition of *impedance* until Experiment 2.

You should again convince yourself that the negative feedback from the op-amp output to the *–Input* is such that the ideal op-amp behavior will keep $v_- = v_+ = 0$ as the input voltage changes. Remember, the condition $R_f i_{in} = -v_{out}$ is satisfied because the op-amp adjusts v_{out} to make it so — it adjusts v_{out} until $v_- = 0$, and this will be the case only when $v_{out} = -(R_f / R_i)\, v_{in}$. When you are confident that this is how the circuit works, you will have learned what you need to know about op-amps for now.

TECHNIQUES FOR ANALYZING CIRCUITS

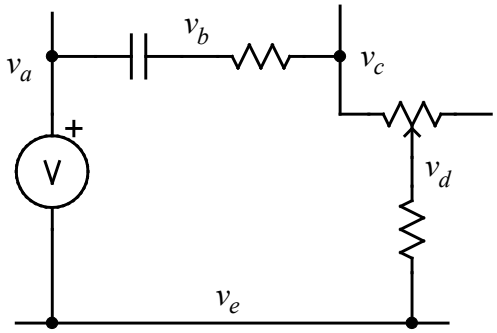
Circuit nodes and loops; Kirchhoff's laws

Now is probably the appropriate time to explicitly state the rules we've been using to determine the relationships between the voltages and currents in our circuits. We've already defined what we mean by a *lumped circuit element*, of which our resistors and op-amps are examples: a lumped element always has *zero net current* flowing into (or out of) it. It may be obvious, but let's state it anyway: the same consideration applies to the connections (called *nodes*) between the terminals of our various elements. For example, the node illustrated on the right connects 5 terminals of some assortment of elements (the placement of the "wires" and their connections at the dots are arbitrary and are chosen to make the schematic as readable as possible — all that matters is that this node makes a common connection to 5 different terminals). Our rule about currents states that, given the arbitrary way we've picked the directions for the current flows toward or away from each terminal, it must be true that $i_1 + i_2 = i_3 + i_4 + i_5$. This rule is commonly known as *Kirchhoff's Current Law*, named for the Prussian physicist Gustav Kirchhoff (1824–1887). If, when we use this rule and solve for the currents, we find that one or more of the currents has a negative value, this result just means, of course, that the actual current flow is opposite to the way we've drawn the arrow.



It also must be noted that for any given node (such as that pictured above), the voltage is the same everywhere along it: i.e., *all terminals connected together by a node are at the same voltage*. In other words, the lines connecting terminals in a schematic are not supposed to represent any sort of "physical" representation of real wires with some nonzero resistance. In our real circuits, though, if the physical distance between a pair of elements is large, and the current flow between them is substantial, then the wire you use to connect them may have enough resistance to introduce a noticeable voltage drop; in this case it might be wise to include the wire itself as another element in your schematic and your calculations.

Another rule relates the voltages across elements whose terminal connections form a closed loop in the circuit (of course, there must be *at least one closed loop* in our circuit, which is why it is called a *circuit*!). Consider a circuit fragment containing a loop like the one at right, where we've also labeled the voltage at each node of the loop. The *loop voltage rule* seems trivial when one labels the node voltages; it says that if we pick any node and add up the voltage differences across the elements' terminals as we go



around the loop, the total voltage change must be zero. In other words, if we start at, say, node *e* (at voltage v_e) and then proceed clockwise around the loop, the voltage across the signal source takes us from v_e to v_a; but if we proceed the other way, using the current-voltage relationship appropriate for each of the elements, the calculated voltages across the elements must take us from v_e to v_d to v_c to v_b and, finally, again to v_a, the same as before. This obtains because *the potentials of the nodes are all well-defined and single-valued*, which will be true as long as our circuit is small compared to the wavelengths of the electromagnetic fields (no magnetically-induced EMFs allowed around our circuit loops!). This rule is known as *Kirchhoff's Voltage Law*. Using these current and voltage laws for our nodes and loops along with the current-voltage laws for the various elements (like Ohm's law, equation 1.1) will give equations relating the various currents in the circuit and the voltages at the circuit's nodes.

A SIMPLE EXAMPLE OF HOW TO USE THE VOLTAGE AND CURRENT RULES

Let's illustrate the use of Kirchoff's laws to solve for the unknown voltages and currents in the very simple circuit shown in Figure 1-18. In this circuit a voltage source with voltage v_s drives a network of three resistors; we wish to find the values of all of the various voltages and currents in the circuit shown in the figure in terms of the specified resistor values (R_1, R_2, and R_3) and the source voltage (v_s).

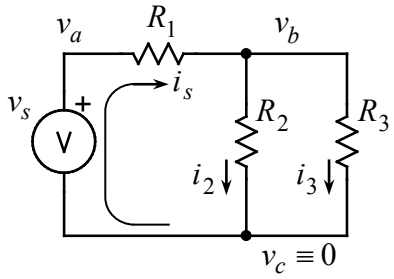


Figure 1-18: A simple example illustrating how Kirchhoff's laws are used to solve for the voltages and currents in a circuit.

The circuit in Figure 1-18 has three nodes: *node a* connects the source and R_1; *node b* connects R_1 to R_2 and R_3; and *node c* connects R_2 and R_3 back to the source. The voltages at these nodes will be designated as v_a, v_b, and v_c, respectively. We'll choose to use v_c as our voltage reference (*ground*), so *by definition* $v_c \equiv 0$, as shown in the figure. Thus we have two unknown voltages: v_a and v_b. Also shown in the figure are three unknown currents: the current supplied to the circuit by the source, i_s (which also must flow through resistor R_1, as shown); and the currents i_2 (through R_2) and i_3 (through R_3).

These five unknowns are connected to the source voltage v_s and to each other by Kirchhoff's voltage and current laws:

(1) Clearly, by going up the left side through the voltage source: $v_a = v_c + v_s = v_s$
(2) Following the current through R_1, the voltage will drop across it: $v_b = v_a - R_1 i_s$
(3) The voltage drop across R_2 takes us back to ground: $v_c = 0 = v_b - R_2 i_2$
(4) Ditto for the voltage drop across R_3: $v_c = 0 = v_b - R_3 i_3$
(5) The sum of the currents into *node b* must vanish, so: $i_s = i_2 + i_3$

Considering the sum of the currents into *node c* gives the same equation as (5), so this additional equation would be redundant.

> When writing down loop equations we must be careful about the direction of a current through an element and the sign of its associated voltage difference: if current flow through a resistor is positive (in the chosen direction of the arrow), then the voltage at the arrow's tail must be greater than the voltage at its head, as shown in Figure 1-4 (on page 1-6). Note that equations (2)–(4) comply with this condition.
>
> *The opposite is true for a power source*: current leaves at the terminal with the more positive voltage and enters at the other.

We have five independent equations for the five unknowns (v_a, v_b, i_s, i_2, and i_3); these equations are straightforward to solve. If we define $R_{\|}$ as the equivalent resistance of the three resistors in parallel:

$$\frac{1}{R_{\|}} \equiv \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

then a convenient way to express the solutions for the five unknowns is:

1.6

$$v_a = v_s \qquad i_s = \frac{v_s}{R_1}\left(1 - \frac{R_{\|}}{R_1}\right)$$

$$v_b = \frac{R_{\|}}{R_1} v_s \qquad i_2 = \frac{v_s}{R_2}\frac{R_{\|}}{R_1} \qquad i_3 = \frac{v_s}{R_3}\frac{R_{\|}}{R_1}$$

You should take a few minutes to show that this solution satisfies condition (5): $i_s = i_2 + i_3$.

Linear circuits and superposition

The node and loop equations used in the previous example were *linear*, so that all of the various currents and voltages were proportional to the single source signal, v_s. This is generally the case for circuits constructed from linear components like our ideal resistors. It

is also true for properly-designed, ideal op-amp amplifier circuits using negative feedback (it is not the case, however, for important nonlinear systems like digital circuits).

Linear circuits are particularly straightforward to analyze, even if there are multiple, independent sources of voltages and currents embedded in them. In the case of a linear circuit excited by multiple sources, the solution for each unknown voltage or current will be given by a sum of terms, each term *proportional to only one source*:

1.7
$$v = \sum_k a_k v_k + \sum_m b_m i_m$$

where v is the unknown response and the sums are over the various source voltages and source currents. The coefficients a_k and b_m of the various terms are derived from the circuit's component values and are independent of the values of the sources.

Because of this simple, linear structure of the solution, we can sometimes determine the various coefficients a_k and b_m quickly by a simple procedure: if all source voltages and/or currents but one were to vanish, then each unknown voltage and current is simply proportional to the one remaining, nonzero source – the ratio of the unknown response to that source is then the value of the corresponding coefficient (a_k or b_k) of its term in equation 1.7. By cycling through each source in turn you can then determine each of the coefficients, thus finding the general solution for the unknown voltage or current. This is the *Principle of Linear Superposition*.

IDEAL VOLTAGE AND CURRENT SOURCES SET TO 0

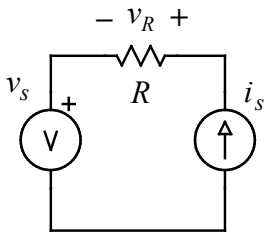
Setting a *voltage source* to 0 is the same as replacing it with a *short-circuit* (wire) connecting its two terminals.

Setting a *current source* to 0 is the same as replacing it with an *open circuit* (no connection at all) between its two terminals.

Time for a few examples…

A REALLY TRIVIAL EXAMPLE USING LINEAR SUPERPOSITION

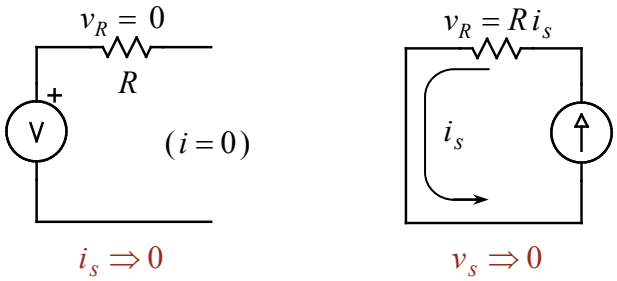
Consider the simple circuit shown at right, with a voltage and a current source joined by a single resistor. We want to know the voltage drop across the resistor (v_R), when the two sources are v_s and i_s (with the polarities shown). According to the general expression 1.7, we can write v_R as:



$$v_R = a\,v_s + b\,i_s$$

for some unknown parameters a and b (which should only depend on R, since that's the only other quantity in the circuit in this example). We can calculate the value of a by setting i_s to 0 (so $v_R = a\,v_s$), and then determining v_R in terms of v_s. Similarly, we get b by setting v_s to

0. Following the directions in the box concerning the sources on the previous page, the equivent circuits for the two cases are:



Setting i_s to 0 (left diagram above) opens the circuit, so no current can flow through R. Thus $v_R = R\,i = 0$, and therefore $(0 = v_R = a\,v_s) \Rightarrow (a = 0)$. Setting v_s to 0 (right diagram) simply connects resistor R across the current source, so $v_R = R\,i_s$, and $(R\,i_s = v_R = b\,i_s) \Rightarrow (b = R)$. This gives us the solution for the total voltage across R due to both sources v_s and i_s:

$$v_R = 0\,v_s + R\,i_s = R\,i_s$$

THE DIFFERENTIAL AMPLIFIER

Now for a more important example of linear superposition: consider an ideal op-amp amplifier, but this time we attach two independent voltage source inputs as shown in Figure 1-19. We now want to know what the output voltage v_{out} will be for any combination of values for the two input voltages v_{in-} and v_{in+}.

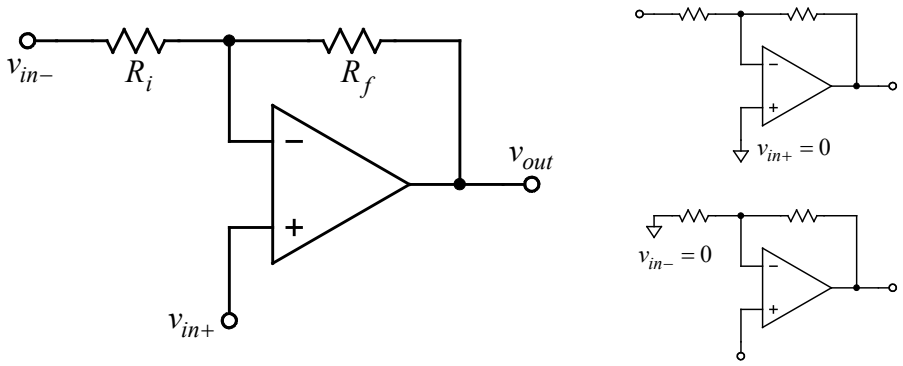


Figure 1-19: A combination of a noninverting and an inverting amplifier. If input signal v_{in+} = 0, then the circuit is an inverting amplifier for signal v_{in-} (top right diagram), and $v_{out} = -(R_f / R_i)\, v_{in-}$. If, instead, v_{in-} = 0, then the circuit is a noninverting amplifier for signal v_{in+} (bottom right diagram), and $v_{out} = (1 + R_f / R_i)\, v_{in+}$. By the *principle of linear superposition*, the response at v_{out} must be the sum of these two expressions: $v_{out} = (1 + R_f / R_i)\, v_{in+} - (R_f / R_i)\, v_{in-}$.

Because the circuit is linear, the solution for v_{out} will again be a sum of terms like equation 1.7. We will use the principle of linear superposition to obtain this solution by setting v_{in+}

and v_{in-} to 0 in turn. As we've said before, setting a voltage source to 0 is nothing more than replacing it with a short circuit (a wire) connecting its two terminals. The input voltage sources in Figure 1-19 each have one terminal connected to ground (that's why we show only a single terminal for each source), *so setting one of them to 0 is the same as connecting that circuit input to ground*.

Thus if $v_{in+} = 0$, the circuit becomes identical to the inverting amplifier in Figure 1-17, and the inverting amplifier gain formula in equation 1.5 will describe how v_{out} depends on v_{in-}. Conversely, $v_{in-} = 0$ results in a noninverting amplifier like Figure 1-16 with gain given by equation 1.4. Linear superposition then implies that v_{out} will vary as the sum of these two expressions, so that for arbitrary v_{in+} and v_{in-} we will get:

$$v_{out} = \left(\frac{R_i + R_f}{R_i}\right) v_{in+} - \left(\frac{R_f}{R_i}\right) v_{in-}$$

This result is almost proportional to the difference in the two input voltages $(v_{in+} - v_{in-})$, but not quite. If we were to first scale v_{in+} by $R_f/(R_i + R_f)$, then the output would indeed be proportional to the input voltage difference, and we would have designed a *differential amplifier*. But this correction factor is just what we would get if we were to add a voltage divider with resistors R_i and R_f between v_{in+} and the op-amp's *+Input*, as shown in Figure 1-20.

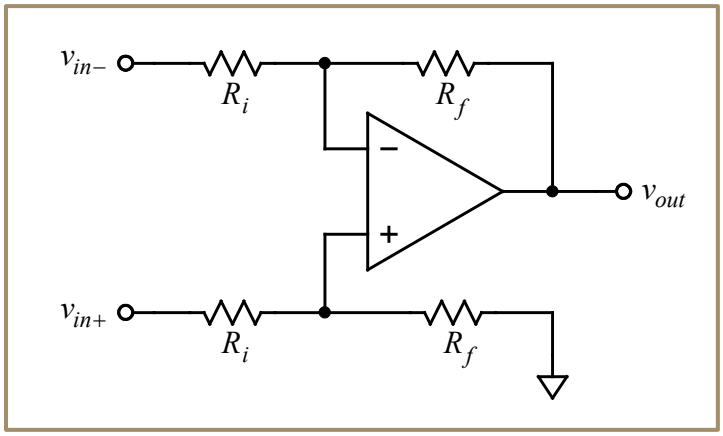


Figure 1-20: The *Differential Amplifier*. The added voltage divider on the v_{in+} input will divide it by just the right amount so that the output is proportional only to the *difference* between the two input signals, equation 1.8.

Because the ideal op-amp's *+Input* draws no current, the voltage at that input will be given using the basic voltage divider equation presented in Figure 1-7 on on page 1-9, reducing v_{in+} by the correct factor to give a purely *differential* gain, equation 1.8.

Ideal, differential amplifier

1.8

$$v_{out} = \frac{R_f}{R_i}\left(v_{in+} - v_{in-}\right)$$

Note one important point about the differential amplifier, however: the output will be given by (1.8) *only if the resistor values are exactly matched* so that the voltage divider on the v_{in+} input exactly compensates for the extra gain of the noninverting amplifier action. The resistors you will use have values whose tolerance is 5% (or maybe 1%), so their actual values may differ from the marked values by that percentage. In practice this implies that even though you may have $v_{in+} = v_{in-}$, you would have nonzero v_{out}. Thus inaccurate matching of the resistor values results in a nonzero *common mode gain* so that v_{out} will include a residual term proportional to the average value of the two input voltages. The ratio of the amplifier's *differential gain*, $G_{diff} = v_{out}/(v_{in+} - v_{in-})$, to its common mode gain, $G_{cm} = 2v_{out}/(v_{in+} + v_{in-})$, is known as its *Common Mode Rejection Ratio* and is an important specification when choosing or designing a differential amplifier.

THE SUMMING AMPLIFIER

For this example we again start with the inverting amplifier configuration, but this time with multiple input sources as shown below. We solve for the circuit's response to its various inputs by setting all but one to zero and using superposition.

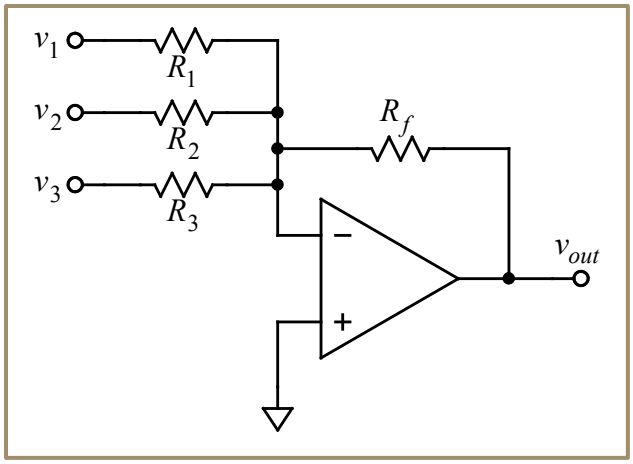


Figure 1-21: The inverting, *Summing Amplifier*. The node joining the resistors to the *–Input* is a *virtual ground*, so each input will be amplified independently of the others. The result is a weighted sum of the several inputs (inverted, of course) given in equation 1.9.

As discussed previously, the node joining the resistors to the op-amp's *–Input* is a *virtual ground*, with 0 voltage. Consequently, for each input set to 0 the associated input resistor (R_1, etc.) will have no voltage across it, *so the current through it must vanish as well*. Thus, each of these zeroed inputs can have no effect on the amplifier's behavior, and the amplifier will behave as a straightforward, inverting amplifier of the one active input, with gain given by equation 1.5. Adding such an inverting gain term for each input will give the result we seek, a weighted sum of the various inputs:

Ideal, inverting, summing amplifier

1.9

$$v_{out} = -\sum_k \frac{R_f}{R_k} v_k$$

GENERALIZING THE VOLTAGE DIVIDER

The ubiquitous voltage divider (Figure 1-7) often appears in circuits in a more general form: several voltages connected to a single node through resistors, as in the previous example, now shown explicitly in Figure 1-22. We would like a relatively simple, easy-to-remember formula for the resulting node voltage.

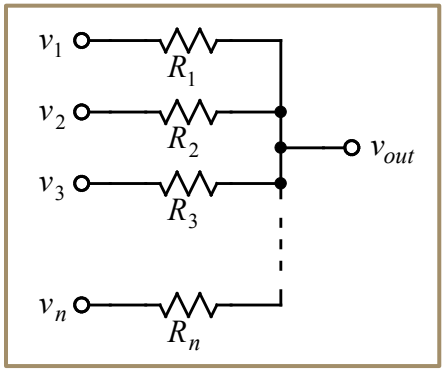


Figure 1-22: A generalization of the voltage divider circuit. Several voltage sources are connected to a single node via resistors, and we want to know the resulting voltage at that node (v_{out}).

The result, whose derivation we leave for the exercises, is a weighted average of the various input voltages:

Generalized voltage divider output

1.10

$$v_{out} = \frac{\sum\left(v_k / R_k\right)}{\sum\left(1 / R_k\right)}$$

Thus, the weight of each voltage source is simply the conductance ($1/R$) of its connection to the common node.

The powerful *principle of linear superposition* will be used throughout this course to solve circuit problems.

Input resistance

We've already mentioned the terms *input impedance* and *output impedance* in our discussion of the "black box" description of a network (or sub-circuit). Now is the time to start to understand what they mean and why they are important, at least in the context of the circuits we've analyzed so far. We can't actually define *impedance* yet (that will have to wait for Experiment 2), but we can talk about *input resistance* and *output resistance*, which are closely-related concepts (in order to restrict ourselves to *resistances*, we will have to assume that our power sources output constant [DC] voltages and currents, and that we wait long enough for any transient, time-varying behavior of the circuit to die away before we take any measurements). In this section we discuss *input resistance*; *output resistance* is a slightly more advanced topic and is addressed in the section starting on page 1-43.

When we connect a power source to an input of a network or circuit, the source will, in general, apply some voltage across the two input terminals, and some current will flow.

The *input resistance* of an input port of a network or device is *the ratio of the applied voltage and the resulting current flowing into that input*.

Input Resistance

1.11

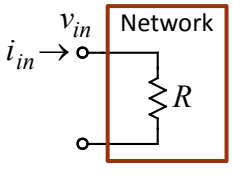
$$R_{in} \equiv \frac{v_{in}}{i_{in}}$$

THE IMPORTANCE OF INPUT RESISTANCE

Input resistance determines how much power must be supplied by an input source. For example, some voltage sources are very weak and cannot provide a significant amount of current. In this case, the input resistance of the circuit must be made high enough to not load down the source.

In other cases (usually involving high frequencies or short signal pulses), the circuit's input resistance must be chosen to match the *characteristic impedance* of the cable connecting the source. Otherwise, the signal will be reflected by the circuit input, wasting signal power and greatly distorting the signal shape. Typically, the input resistance in these cases should be 50Ω.

If our circuit is linear and if there is only one input power source, then R_{in} as defined in 1.11 will not vary with the magnitudes of v_{in} and i_{in}. For example, if our network consists of a single resistor R, as shown at right, then (trivially) connecting an input voltage source v_{in} to its terminals will result in current $i_{in} = v_{in}/R$, and, as expected, $R_{in} = R$.

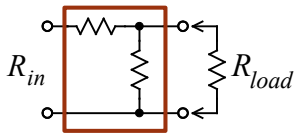


For a less trivial example, consider the input resistance of the 3-resistor network in Figure 1-18 on page 1-19. With the input voltage to that network $v_{in} = v_s$ and resulting current $i_{in} = i_s$ (refer to the figure for the definitions of v_s and i_s), then the solution to that example, equations 1.6 on page 1-20, shows that:

$$R_{in} = \frac{v_s}{i_s} = \frac{R_1}{1 - R_{\parallel}/R_1} = R_1 + \frac{R_2 R_3}{R_2 + R_3}$$

which we could have anticipated by noting that the network simply consists of resistor R_1 in series with the parallel combination of resistors R_2 and R_3.

In general, the input resistance R_{in} of a network port will depend upon what is connected to the network's other ports, including its output ports. For example, let's examine the simple voltage divider considered as a two-port network (at right). Clearly, its input resistance R_{in} will depend on the value of the load R_{load} connected to the divider's output port, since the load resistance is in parallel with one of the voltage divider's two resistors. In fact, this network (including the output load resistor R_{load}) is the same as that in Figure 1-18, with R_{load} assuming the role of that example's R_3. This particular situation (input resistance changing if the load changes) can be avoided by adding a voltage follower as in the third example in Figure 1-23 below: the op-amp *isolates* the input from the output load.



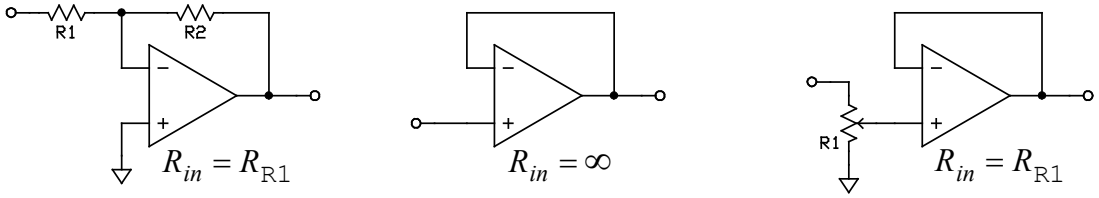


Figure 1-23: Amplifier circuit examples illustrating the *Input Resistance* concept. Inverting amplifier: since the *–Input* node of the op-amp is a virtual ground, the input resistance of the amplifier is just the value of R1, the input resistor. Voltage follower: since the op-amp inputs draw no current, the input resistance is infinite. Potentiometer as a variable voltage divider connected to a voltage follower: again, no current flows into the *+Input*, so the input resistance is given by the total potentiometer resistance R1. In each of these cases, the amplifier isolates the output load from its input, so the input impedance is unaffected by the current required by the load connected to its output.

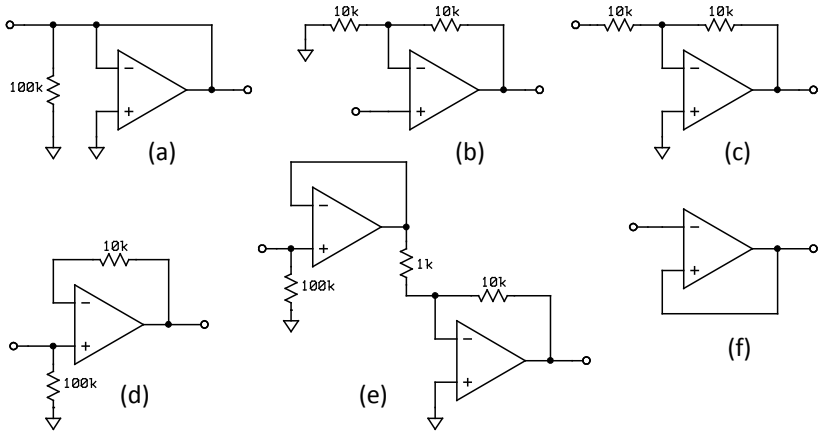
PRELAB EXERCISES

1. The lab power supplies you will use for your op-amp circuits supply $+12\,\text{V}$ and $-12\,\text{V}$ at up to $250\,\text{mA}$ each. How many total watts of power is this?

   If the resistors you use can absorb no more than $1/4\,\text{W}$ without damage, then what is the *minimum* resistor value which can be connected between the $+12\,\text{V}$ supply and ground without damage? between the $+12\,\text{V}$ supply and the $-12\,\text{V}$ supply?

2. If you take a resistor R and add another resistor with resistance $0.1\times R$ in series with it, the total resistance of the pair is obviously $1.1\times R$, a 10% increase. If instead you wish a combination which has a total effective resistance of only $0.9\times R$, you could do it by placing a second resistor in parallel with the original R. In this case, what should be the value of this second resistor (use the approximation $1/0.9\approx 1.1$ when calculating your answer)?

3. Which of the following amplifier circuits will work correctly (do something useful)? Which won't, and why not? For those that do work correctly, what is the amplifier gain (use a '−' sign for an inverting gain)? the input resistance? Look at them carefully!



4. Refer to Figure 1-22 on page 1-25 and derive the generalized voltage divider formula, equation 1.10 (repeated below).

1.10

$$v_{out} = \frac{\sum\left(v_k/R_k\right)}{\sum\left(1/R_k\right)}$$

Hint: a clever way to solve this problem is to convert each of the (v_k, R_k) sources to their *Norton equivalents* (see Figure 1-36 on page 1-45); this will give (current source + parallel resistance) pairs: $(i_k = v_k/R_k, R_k)$. Remember that each source has its other (implicit) terminal connected to ground ($\downarrow$), as does the output voltage node, v_{out}. Now use the principle of superposition, setting all but one current source to 0 in turn (see the rule on page 1-21 about this). Note that now all the source resistors are in parallel!

5. Consider the circuit below (Figure 1-24), which you will construct and evaluate during lab. What is the the circuit's gain (v_{out}/v_{in}) when the *gain adjust* potentiometer's wiper is set to the top end of its resistance element? Set to the bottom end? Centered?

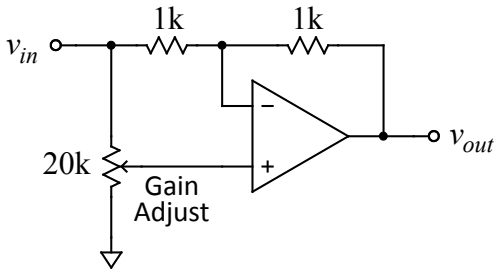


Figure 1-24: How does the gain of this circuit vary as the potentiometer is adjusted?

LAB PROCEDURE

> ***Ask questions during the lab!*** Don't just sit and stare helplessly at a circuit or piece of test equipment which stubbornly refuses to cooperate!

Overview

During lab you will experiment with various op-amp circuits and evaluate their performance. You will start to become familiar with the analog electronics lab trainer, the signal generator, the oscilloscope, and a couple of the data acquisition and control programs available on the lab computer workstations. For each circuit configuration you investigate, you should use the oscilloscope to measure the input and output peak-to-peak voltages (for oscillating signals) or the mean voltages (for DC signals). For oscillating signals you should try both sine and square waveforms (peak-to-peak: difference between maximum and minimum values).

> Interesting results may be recorded by saving oscilloscope screen-capture images. Record your findings in a bound notebook; insert and tape into place printed screenshots or other computer-generated graphs. Schematics (with component values and input and output ports labeled) are required for all circuits.

Figure 1-25 shows a typical lab station setup. The more quickly you become familiar with the oscilloscope's and signal generator's controls and menus, the more you will enjoy your time in the lab and the more productive you will become. This first experiment is a good one to spend time exploring the instruments' various modes and capabilities. Ask lots of questions!

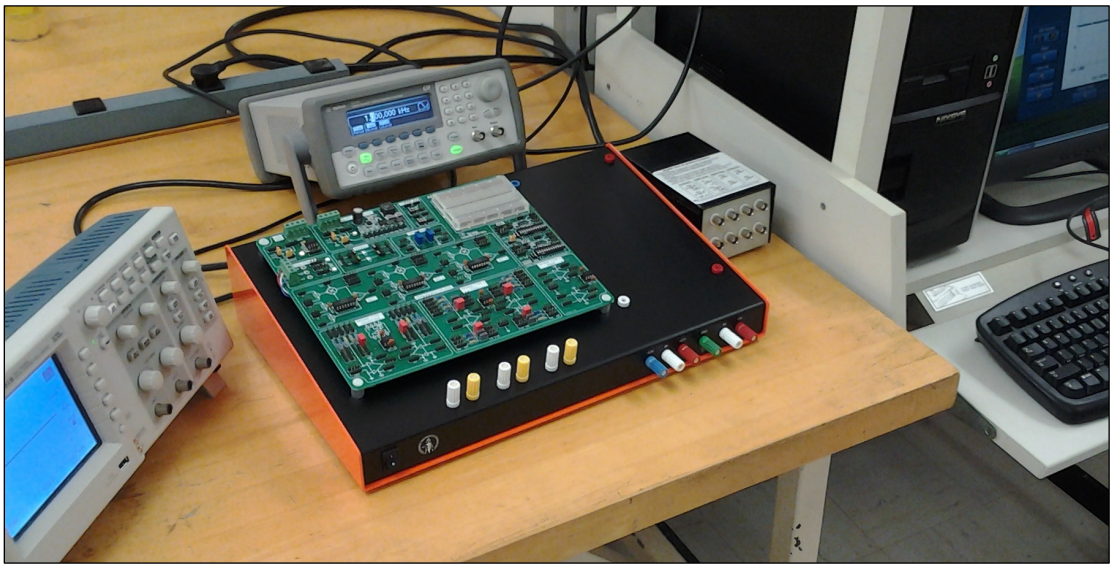


Figure 1-25: A typical lab station setup, with analog electronics trainer and breadboard, oscilloscope, signal generator, and computer with data acquisition and control software.

Using the analog trainer and breadboard

The analog breadboard we use is the Texas Instruments *ASLK PRO*, kindly donated by the company for our use. The manufacturer's web site supporting this system may be found at:
http://e2e.ti.com/group/universityprogram/educators/w/wiki/2047.analog-system-lab-kit-pro.aspx

The manufacturer's student manual for the breadboard is found here:
http://www.mikroe.com/downloads/get/1741/analog_system_lab_pro_manual.pdf

> The board shows power supply connections as +10V and −10V, but the actual power supply voltages are **+12V and −12V**.
>
> The board **Ground** connections are to the common power supply return as shown in Figure 1-3. *They are not connected to* **Earth Ground** *within the breadboard assembly*.

For this first experiment you will assemble various amplifier circuits by using jumper wires to connect components on the trainer. The photos in Figure 1-26 show the proper way to make connections among the components.

Note in the left photo of Figure 1-26 that several components (a bunch of resistors and capacitors) are installed on the breadboard along with each op-amp, ***each with one end already connected to one of the op-amp inputs***. The other end of each component has a set of pins to which you may connect jumper wires. Only connect a jumper wire to the component you wish to use; the others will then have no effect on your circuit. In the two right-hand photos in Figure 1-26 , for example, one yellow jumper wire connects the op-amp output to a 10k resistor on the op-amp *–Input*, so the 10k resistor is being used as R_f.

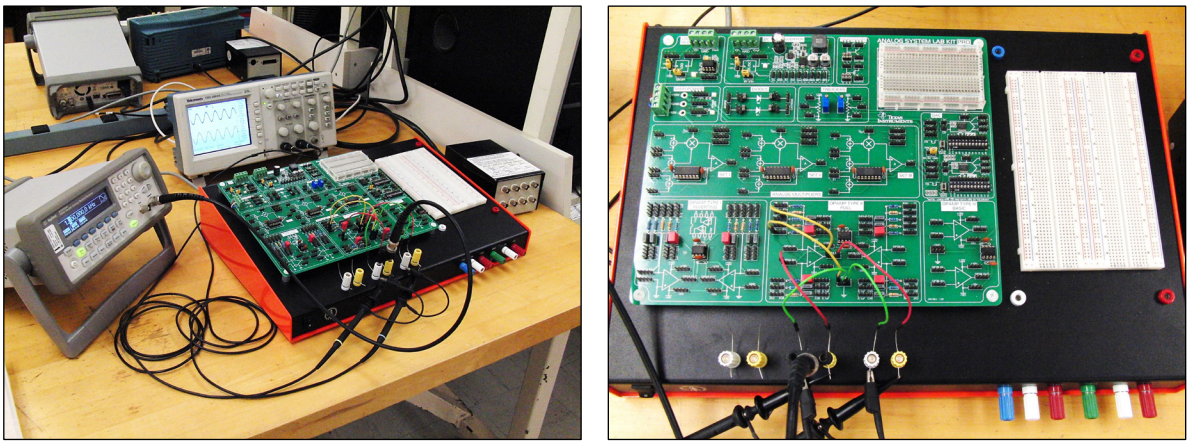


Figure 1-26: The proper way to build circuits using the breadboard's preinstalled component area.
***Left*: overall image of a setup showing the oscilloscope with its two *10× probes* and the signal generator using a *BNC cable* for its connection.**
***Right*: The breadboard area used for most amplifier designs has two op-amps and associated resistors and capacitors. In this image a ×11 noninverting amplifier is wired up in that area.**

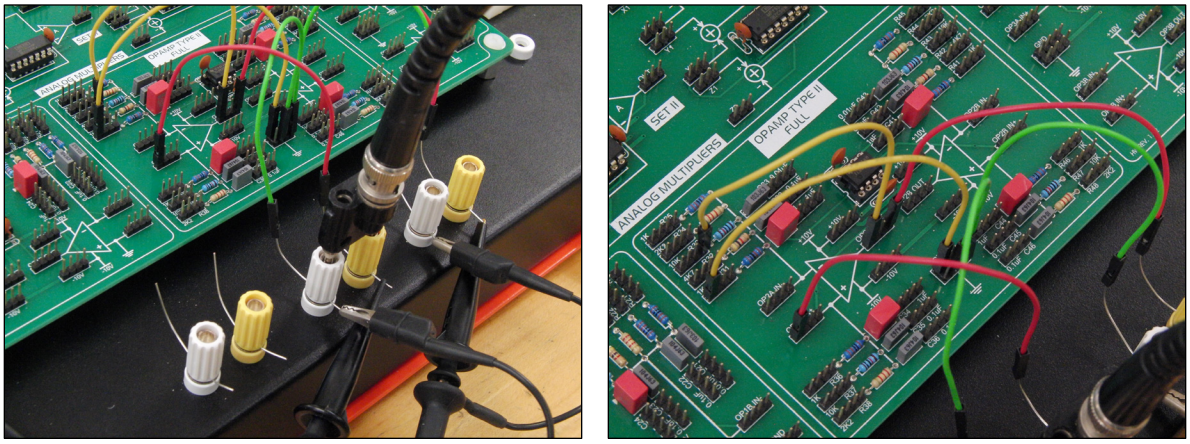


Figure 1-26 (continued).
***Left*: detail showing how the *5-way binding posts* are used to connect the circuit to the signal generator and the oscilloscope probes. 22 gauge solid wire is used as a "terminal" to which a circuit jumper and/or a probe clip may be attached. BNC cables are plugged into a pair of binding posts using an adapter (Figure 1-27).**
***Right*: detail showing how to construct a ×11 noninverting amplifier using jumper wires. The yellow wires connect a 10k resistor as R_f (to the op-amp output) and a 1k resistor as R_i (to ground). The red wires connect the amplifier input and output signals to the binding posts. Green wires connect the signal generator and oscilloscope grounds to the circuit ground, completing the circuit.**

Considerations when making BNC cable connections

Connections to external instrumentation (primarily the signal generator, oscilloscope, and the computer data acquisition system (DAQ) are made using coaxial cables with BNC connectors or, in the case of the oscilloscope, *10× probes*. The *BNC-banana interface adapter*, shown in Figure 1-27, lets you use a BNC cable with the breadboard assembly's 5-way binding posts, as shown in Figure 1-26. A BNC cable contains two conductors: an inner signal wire separated by an insulating sleeve from a surrounding braided shield. The outer metal shell of the BNC connector is connected to the cable shield, whereas the connector center pin goes to the cable's inner signal wire. The adapter's banana plug identified with the "GND" tab (as shown in the photo) connects to the BNC shell, the other banana plug (*opposite the "GND" tab*) goes to the BNC signal pin.

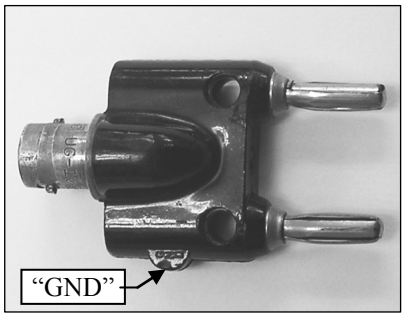


Figure 1-27: BNC adapter showing the tab identifying the connection to the BNC shell.

The two conductors of the BNC are not equivalent. For all the instruments in the lab the displayed voltage is that of the center pin with reference to the outer shell. In other words, *the BNC shell serves as the voltage reference for signals on the cable* (it is a *signal ground*).

The DAQ analog voltage input BNC connectors are isolated from the other connectors and from *Earth Ground*, so their two connections may be made anywhere on the breadboard to make a measurement. This is not the case for the oscilloscope inputs! *Both oscilloscope input BNC shells are connected to each other and to Earth Ground*. This is also the case for the other BNC connectors on the computer DAQ interface. The connector shells for any of these connections you make to the circuit should be connected to the breadboard ground.

> The two oscilloscope inputs have BNC shells connected to **Earth Ground**. Always connect these conductors (and the ground clips on the 10× probes) to the analog breadboard **Ground** as shown in the photos of the example circuit in Figure 1-26.

Detailed procedures

> **CAUTION**
>
> When assembling a circuit or making changes to it, turn off the signal generator output (using its OUTPUT button) and turn off the power switch on the breadboard.
>
> ***Never apply a signal to a circuit that is not powered up***. Otherwise the signal can damage the op-amps.
>
> The signal generator can supply up to 100mA to a circuit, which can cause some damage! The power supply terminals on the breadboard can source up to 250mA (at ±12V), which can make quite a spark!

Inverting and noninverting amplifiers

Assemble the ×11 noninverting amplifier circuit shown in Figure 1-26, including the connections to the signal generator and oscilloscope. A schematic of this complete setup is shown below.

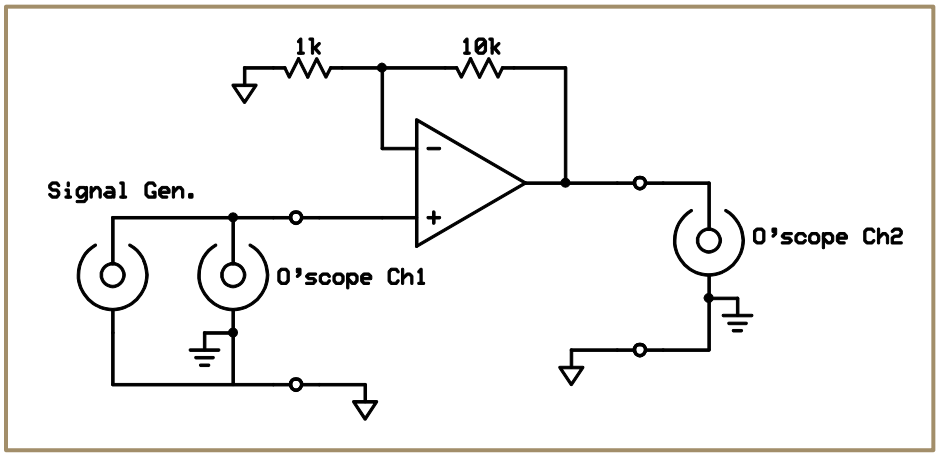


Figure 1-28: The first circuit you should build, which is shown in Figure 1-26 as well. Note that the oscilloscope's internal connections of its BNC shells to *Earth Ground* are also shown.

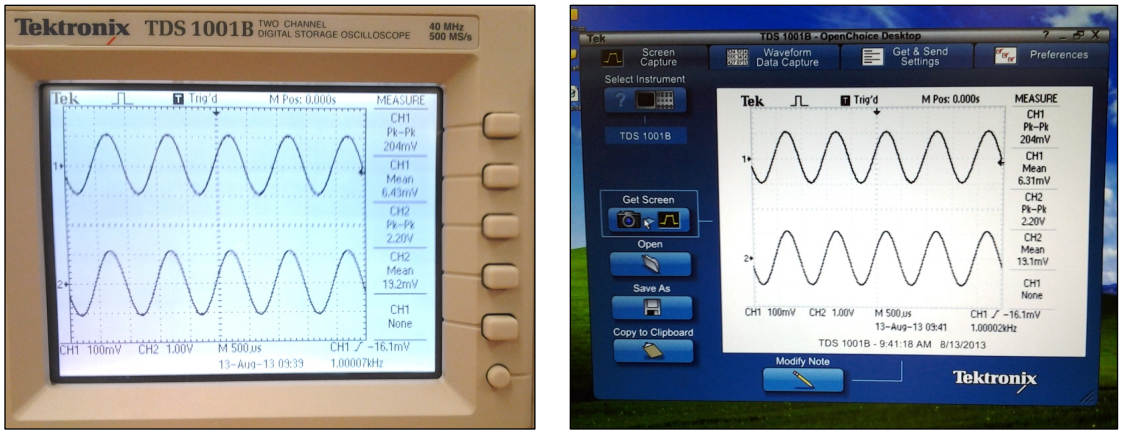


Figure 1-29: Waveform display using oscilloscope configured to measure input and output signal amplitudes and mean voltages, and a computer screen capture of similar data.

Investigate the behavior of this circuit using different waveforms, amplitudes, and frequencies. Determine the gain by comparing oscilloscope measurements of the peak-to-peak input and output amplitudes. Use the signal generator to input a constant (DC) voltage, and determine the gain by comparing measurements of the input and output mean voltages. The photos in Figure 1-29 show an oscilloscope configured to make these measurements. What happens to the output if the input amplitude is too large? Take a screen shot of this.

Next, reconfigure the input connection to the circuit to convert it into an inverting amplifier. What should be the gain for this amplifier? Confirm the operation of the amplifier in this configuration. Now try different resistor combinations to get various gains, both in noninverting and inverting amplifier configurations (don't forget to include a voltage follower). Fill out a table similar to this to summarize your results:

R_f	R_i	Noninverting Gain	Inverting Gain

Variable-gain amplifier

Construct the circuit presented in Figure 1-24 on page 1-29 (Prelab exercise 5). Use the 20k potentiometer included with your parts kit. The potentiometer should be installed in small the breadboard area in the upper right section of the circuit trainer — your TA or the laboratory instructor can show you how to properly connect it. How does the circuit gain vary as you adjust the wiper position? Does it behave as you predicted in your answers to the prelab problem?

Additional circuits

Try cascading two amplifiers [output of one goes to the input of another, like circuit (e) in Prelab exercise 3]. Finally, build and evaluate at least one of the circuits (your choice) from the **MORE CIRCUIT IDEAS** section.

What your experiment write-up should include

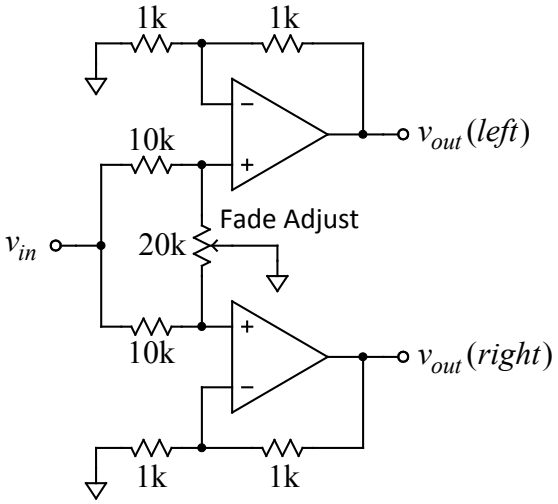
In addition to the specific requirements or questions to answer mentioned in the above procedures, your results should always include:

1. A schematic of each circuit you present results for showing all component values and how the signal generator and oscilloscope are connected to it.
2. What the expected performance of the circuit is predicted to be (gain, etc.).
3. Actual circuit performance, including interesting oscilloscope screen captures.
4. Comments about unexpected circuit behaviors and any useful "lessons learned" you ought to remember.
5. Any interesting ways you've discovered to configure the instruments to generate useful circuit inputs or make measurements.

MORE CIRCUIT IDEAS

Audio fader control

Consider the schematic at right, which demonstrates another interesting way to use a potentiometer. When the *fade adjust* wiper is in its center position, $v_{in}/2$ is presented to each opamp's +*Input*, which is configured as a ×2 noninverting amplifier. Thus both the left and right outputs equal v_{in}. As the wiper position is varied the input signal is distributed unequally to the two amplifiers, so that it may be smoothly panned from one output channel to the other. Note that with the resistor values shown, the total output power (which is proportional to the sum of the squares of the two output signals) stays approximately constant. If two different signals are applied to the pair of 10k resistors (rather than the single input as shown in the figure) then the circuit becomes a stereo balance control.



Current to voltage (transimpedance) amplifier

Sometimes the input signal to be amplified and measured is a current rather than a voltage. One popular way to accomplish this would be to let the current flow through a small resistor of known value and then to amplify the voltage drop across it (using an instrumentation amplifier, maybe). If the current flows into ground potential, then a *transimpedance amplifier* may be a better solution (figure at right). A current entering the amplifier's input port (which is, of course, a virtual ground) flows on through the feedback resistor R1. The voltage drop across this resistor appears at the op-amp output, so the gain is: $V_{out}/I_{in} = -R_{R1}$. The gain has units of impedance (resistance), which is called a *transimpedance* (or *transresistance*) because it is the ratio of a voltage and a current measured at two different places (*trans*: "transfer"). Because the input is a virtual ground, the voltage the current source must be able to supply at the amplifier input in order to maintain its current is tiny (this is called the required voltage *compliance* of the amplifier).

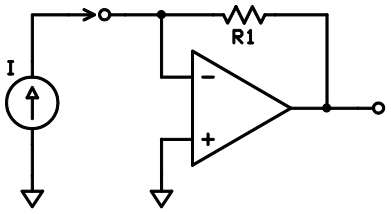


Figure 1-30 is an example of the usefulness of this amplifier. A *reverse-biased photodiode* is connected between a power supply and the amplifier input as shown (we'll learn more about diodes in a later experiment). Connected as shown, only a fraction of a microamp of current will flow through the diode when it is not illuminated (the diode's *dark current*). When illuminated the diode's current will be much larger — 10's to 100's of microamps; its current is very nearly proportional to the level of illumination, and so will be the amplifier's output voltage. Select the value of resistor R1 so that the output voltage range is appropriate for the light intensities you expect to experience.

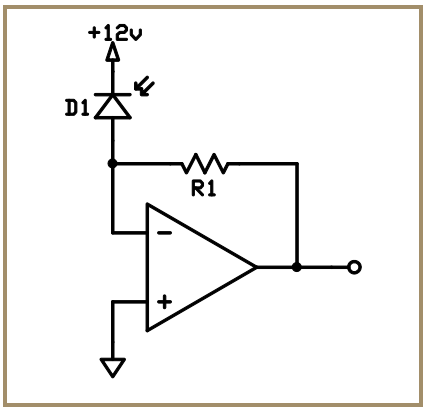
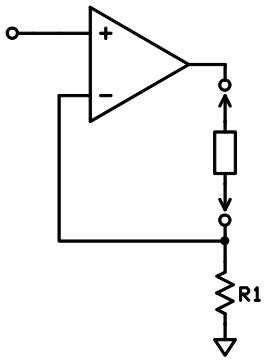


Figure 1-30: A photodiode amplifier. The output is proportional to the light intensity; *R1* sets the gain.

Voltage to current (transconductance) amplifier

Sometimes you would like to accurately set the current through a load by using a control voltage. Since a resistor is a "voltage to current converter" (Ohms' law, right?), we just need to apply the control voltage across a resistor and direct the resulting current through the load. A simple *transconductance amplifier*, shown at right, performs this trick. Essentially a noninverting amplifier configuration, the negative feedback ensures that the input voltage V_{in} will also appear at the *–Input* and thus across resistor R1. The current to establish this voltage drop comes from the op-amp output by passing through the load (shown connecting the output terminals). Thus the load current is given by: $I_{out} = V_{in} / R_{R1}$; the ratio of the output current to the input voltage (the amplifier gain) is the conductance $1/R_{R1}$. No current is required from the source of the control voltage V_{in}; all load current comes from the op-amp output.



One major drawback of this simple circuit is that the load must be an isolated, 2-terminal device (neither output terminal is a circuit ground, so the load must be able to *float*). We'll investigate transconductance amplifier circuits in a later experiment which relax this requirement.

A useful application of this circuit is to illuminate a light emitting diode (LED) with a current proportional to the input voltage (see Figure 1-31 on page 1-38). Since the output intensity of a LED is very nearly proportional to its current, we have a voltage to light intensity converter (the reverse of the circuit in Figure 1-30). LEDs require a minimum of about 1.5V to produce any light output, so using this amplifier can really simplify things. If R1 is a 1k resistor, then the LEDs will have 1mA current for every 1V input, a useful conversion factor for many applications.

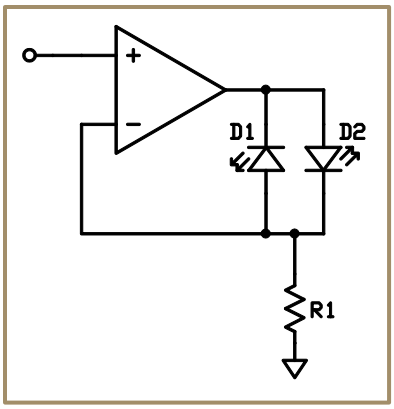
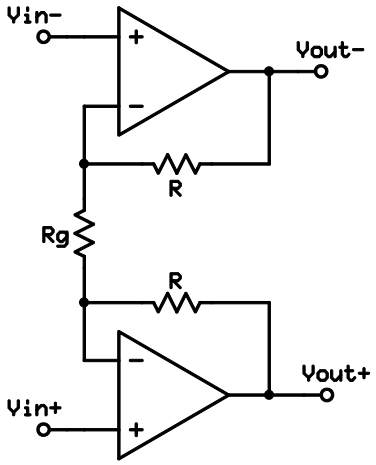


Figure 1-31: A LED driver using a transconductance amplifier. LED *D1* is illuminated for $V_{in} < 0$, *D2* whenever $V_{in} > 0$. Resistor *R1* sets the conversion from input voltage to LED intensity.

Instrumentation amplifier

Consider the circuit shown at right. How do we analyze its response to the two inputs? The answer is, naturally, to use linear superposition, but we're going to apply it in a clever way. Instead of setting V_{in+} or V_{in-} individually to zero, consider instead two combinations of them: their difference, $V_{diff} = V_{in+} - V_{in-}$, and their average, $V_{cm} = (V_{in+} + V_{in-})/2$. Then if we set $V_{diff} = 0$, we would have $V_{in+} = V_{in-} = V_{cm}$; if instead $V_{cm} = 0$, then $V_{in+} = -V_{in-}$. Now use linear superposition of these two new independent variables, V_{diff} and V_{cm}, to determine V_{out+} and V_{out-}, assuming that the two resistors labeled R are perfectly matched, and the resistor $R_g \neq 0$.



Consider the case $V_{diff} = 0$ first, so $V_{in+} = V_{in-}$. Assuming that the negative feedback does what it's supposed to, then the two op-amps' *–Inputs* are also equal, which means that the voltage drop across R_g vanishes, and thus so does the current through R_g. Since no current flows through it, resistor R_g could be removed from the circuit without affecting the circuit's operation! We're left with two independent voltage followers (the value of R doesn't matter), and $V_{out+} = V_{out-} = V_{cm}$. The circuit has a gain of 1 for a *common mode input*, V_{cm}: $G_{cm} = 1$.

Next set $V_{cm} = 0$, so $V_{in+} = -V_{in-}$ and ditto for the two op-amps' *–Inputs*. The resistor R_g has its two terminals with equal and opposite voltages, so *the center of this resistor is at ground potential*. We can split the resistor into two resistors in series, each with value $R_g/2$, and we know that their junction is at ground potential. This symmetry again lets us separate the circuit into two twins, but this time each half is a noninverting amplifier with $R_f = R$ and $R_i = R_g/2$, so each amplifier has a gain of $1 + 2R/R_g$. So:

$$V_{out+} - V_{out-} = \left(1 + 2R/R_g\right)\left(V_{in+} - V_{in-}\right) = \left(1 + 2R/R_g\right)V_{diff}$$

$$\boxed{G_{diff} = 1 + 2R/R_g}$$

We have a circuit with a high gain for a *differential input signal*, but a gain of 1 for a *common mode signal*. The differential gain, G_{diff}, can be adjusted by changing the value of only one resistor, R_g. Because each amplifier has a noninverting configuration, its *input resistance is very large*.

This clever circuit is meant to be combined with a traditional differential amplifier (Figure 1-20) to create the 3 op-amp *instrumentation amplifier*:

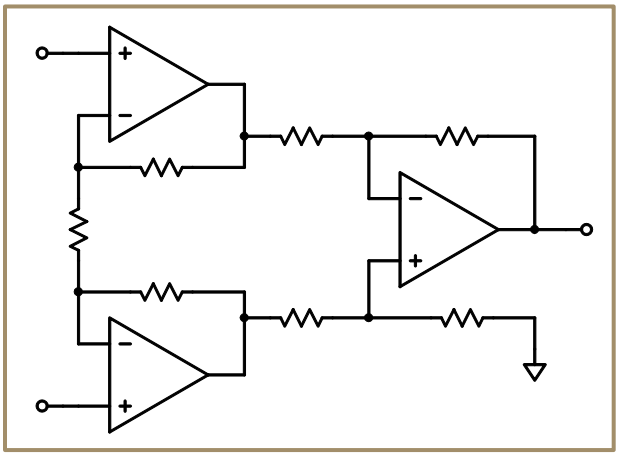


Figure 1-32: The *Instrumentation Amplifier*. This enhancement of the basic differential amplifier has a very high input impedance for both the + and – inputs, has a differential gain which can be changed by varying the value of one resistor, and has very high *common mode rejection*. It is so useful, especially for scientific applications, that many versions are available in the form of single integrated circuit devices.

ADDITIONAL INFORMATION ABOUT THE CIRCUITS

This section expands on some of the material presented earlier. It is better skipped during a first reading, but you may want to go over it after you thoroughly understand the concepts discussed in the first several sections.

Other sorts of circuit grounds

Often in commercial electronic equipment (but not always) the circuitry's 0-Volt reference point will be physically connected to the ground beneath the building using the "ground connection" in the device's 3-pin AC power-line cable; this reference point thus called *earth ground* and is said to be at *earth ground potential*.

If it is important to distinguish between a circuit's 0-Volt reference potential and earth ground (because the two may not be actually physically connected), the circuit reference point is then called *local ground* (or just *ground*), and *earth ground* would be, in general, a different reference point with a different potential. Other common reference points include *chassis ground* (the potential of the equipment's metal enclosure), *analog ground* and *digital ground* (if these different sections of a circuit do not share a common reference point), *signal ground* (the bottom terminal of an input signal source's circuit element), etc. The symbols used in this text for various grounds are shown in Figure 1-33.

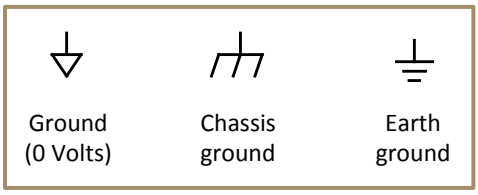
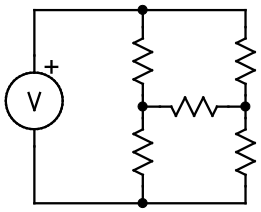


Figure 1-33: Ground symbols we might use. Usually, only ⏚ will represent the 0-Volt reference point.

A nontrivial example of circuit analysis using Kirchhoff's laws



Consider the *bridge circuit* of five resistors shown here, driven by a voltage source. We want to determine two things: the total equivalent resistance of the circuit, and the current through the center, horizontal resistor, assuming that the values of the various resistors are all different (the center resistor is called the "bridge resistor" because it "bridges" the outputs of two voltage dividers). You will quickly realize that this is a nontrivial problem; we can't use our series or parallel resistor formulas to simplify our analysis, so we will have to employ a more brute-force method.

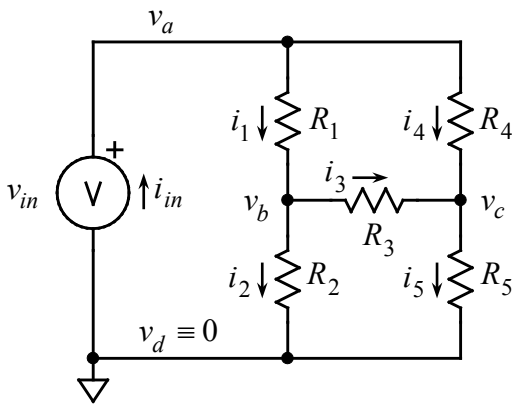


Figure 1-34: A nontrivial example to exercise the circuit voltage and current rules: a resistor "bridge" with voltages and currents identified and labeled for analysis. The bottom node is defined to be ground, so the voltage there is 0 by definition. The source is assumed to supply voltage v_{in}, and the circuit draws current i_{in} from it.

We start by labeling the currents and voltages at each node (there are 4 nodes in all). Figure 1-34 (below) gives a possible labeling which we'll use for this example. Notice that we have already used the rule concerning the total current flow into a lumped element's terminals: since each element has two terminals, we know that the various currents may be represented by a single current flowing through each of the elements. We have also chosen the bottom node to be *ground*, so the voltage there is 0, by definition (since there are no hidden or implied power supplies or other elements connected to ground for this example, we know that there are no other connections carrying current into or away from our ground node, and our list of currents is complete).

The example problem has one independent (driving) variable, v_{in}, and five fixed parameters, the resistor values R_1 through R_5. There are three unknown voltages, v_a through v_c, and six unknown currents, i_{in} and i_1 through i_5. So we need 9 independent equations to solve for the unknowns. Start with the loop voltage rule and use Ohm's law to relate the node voltages to the resistor currents:

(1) Clearly, by going up the left side through the voltage source: $v_a = v_d + v_{in} = v_{in}$
(2) Following the current through R_1, the voltage will drop across it: $v_b = v_a - R_1 i_1$
(3) Similarly for R_4: $v_c = v_a - R_4 i_4$
(4) The voltage drop across R_3: $v_c = v_b - R_3 i_3$
(5) The voltage drop across R_2 takes us back to ground: $v_d = 0 = v_b - R_2 i_2$
(6) Ditto for R_5: $v_d = 0 = v_c - R_5 i_5$

The node current rule gives:

(7) at node *a*: $i_1 + i_4 = i_{in}$
(8) at node *b*: $i_2 + i_3 = i_1$
(9) at node *c*: $i_3 + i_4 = i_5$
(10) at node *d* (ground): $i_2 + i_5 = i_{in}$

Now we have 10 equations for our 9 unknowns, but, of course, they are not all independent. As you might expect, the problem is with the node current equations; in fact the combination of equations $(9)+(10)-(8)$ yields equation (7), so we can discard one of these four equations and the remaining nine will form a complete, independent set. This will often be the case for the node current equations of a circuit.

We can solve this linear system of equations in any of the standard ways; this is the sort of problem computers were originally designed to solve. The solution is messy, and the answers to the original problem turn out to be:

The total equivalent resistance presented to the source:

$$R \equiv \frac{v_{in}}{i_{in}} = \frac{\left(R_1R_2R_4 + R_1R_2R_5 + R_1R_4R_5 + R_2R_4R_5\right) + R_3\left(R_1+R_2\right)\left(R_4+R_5\right)}{\left(R_1+R_4\right)\left(R_2+R_5\right) + R_3\left(R_1+R_2+R_4+R_5\right)}$$

The current through the bridge resistor, R3:

1.12
$$\frac{i_3}{v_{in}} = \frac{R_2R_4 - R_1R_5}{\left(R_1R_2R_4 + R_1R_2R_5 + R_1R_4R_5 + R_2R_4R_5\right) + R_3\left(R_1+R_2\right)\left(R_4+R_5\right)}$$

The bridge circuit is said to be *balanced* when the current through the bridge resistor, R_3, vanishes. We see that the requirement for balance is $R_2R_4 = R_1R_5$.

Incremental Input Resistance

If there are other sources providing inputs to a network, then at an input port in question it may no longer be the case that v_{in} and i_{in} are proportional; it may be, for example, that $i_{in} \neq 0$ even though $v_{in} = 0$. Consequently, we probably don't want to use equation 1.11 to calculate the port's input resistance. Consider, for example, the differential amplifier in Figure 1-20 (on page 1-23): the op-amp's *–Input* node *will not be a virtual ground* because the op-amp's *+Input* voltage depends on the source voltage v_{in+} (and the op-amp *–Input* will be at the same voltage as the *+Input*); therefore even if $v_{in-} = 0$, the current i_{in-} at that port will not vanish because of the nonzero voltage drop across the input resistor. To avoid this sort of problem with our definition of the input resistance, we instead define the circuit's *incremental input resistance*, r_{in}, *as a derivative* (equation 1.13).

Incremental Input Resistance

1.13
$$r_{in} \equiv \frac{\partial v_{in}}{\partial i_{in}}$$

We use a *partial* derivative in (1.13), meaning that all other *independent input sources* are held constant. For linear circuits, the incremental input resistance r_{in} will not vary with the magnitudes of v_{in} and i_{in}, even in the presence of other input sources. The incremental input resistance r_{in} also will not be affected by the amplitudes of the other input sources. The

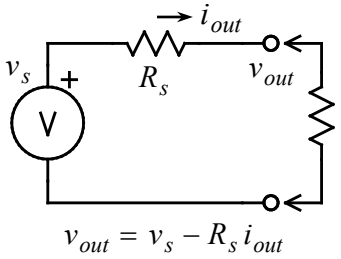
concept of an incremental resistance, however, is useful even for nonlinear circuits, although in this case you should expect that its value will vary with voltage and/or current.

CALCULATING A PORT'S INCREMENTAL INPUT RESISTANCE

In a linear circuit the incremental input resistance r_{in} of any individual input port is not affected by the magnitudes of any of the other independent input sources. Thus you can calculate its value by first setting all other input sources to 0 (using the rules on page 1-21) and then determining the ratio v_{in}/i_{in} for the single input port in question. Make sure all other ports of the network (sub-circuit) are first *terminated* (connected to whatever circuits to which they will interface).

Output resistance

The *output resistance* of a power source or a network's output port characterizes how its output varies with changing load resistance. For example, consider a power source which is made up of an ideal voltage source, v_s, in series with a nonzero source resistance, R_s, driving some load resistor as shown at right. The current through the load, i_{out}, also must flow through the source resistance, and the voltage drop across this resistance is then $R_s i_{out}$. Thus, v_{out}, the output voltage, is less than v_s, the voltage it would be if the load resistance were infinite ($i_{out} = 0$).



We define the output resistance of an output port in terms of the drop in the output voltage with increasing output current:

Output Resistance

1.14

$$r_{out} \equiv -\frac{\partial v_{out}}{\partial i_{out}}$$

We use a partial derivative in this expression to reflect the fact that the driving source v_s *must be held constant* as the output current is varied to determine r_{out}. An *ideal voltage source* is one whose output voltage is unaffected by the load current it must supply, so its output resistance is 0. An *ideal current source*, on the other hand, supplies a constant current regardless of the voltage required to push that current into its load, so $\Delta i_{out} \equiv 0$. Thus for a current source $r_{out} = \infty$.

Our old friend, the voltage divider, however, is not such an ideal character. Let's calculate r_{out} for the output of a voltage divider driven by an ideal voltage source at its input. We will solve this problem by using linear superposition, our powerful circuit analysis ally.

VOLTAGE DIVIDER OUTPUT RESISTANCE

The problem is to determine how the output voltage changes as we change the output current required by a load attached to a voltage divider. To do this using linear superposition, we perform a sort of "thought experiment": we replace the output load with a *current sink*, an ideal current source we can control to independently set the output current to mimic any load.

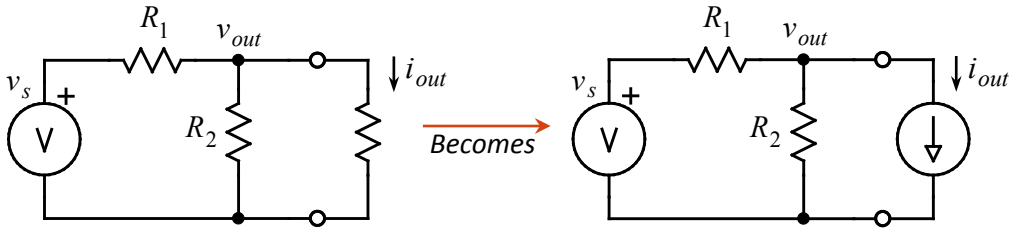


Figure 1-35: Determining the output resistance of a voltage divider driven by an ideal voltage source. Think of the load as a *current sink* (right schematic) and then set the source to 0; this puts the two resistors in parallel, and that parallel combination gives the output resistance, r_{out}.

With this change (Figure 1-35), we know that v_{out} will be a function of v_s and i_{out} with the form given by equation 1.7: $v_{out} = a\,v_s + b\,i_{out}$ for some constants a and b determined by the resistor values R_1 and R_2. This means that $r_{out} = -\partial v_{out} / \partial i_{out} = -b$, so we just need to determine the value of b. We do this by setting $v_s = 0$ and then determining $b = [v_{out} / i_{out}]_{v_s = 0}$. Setting the voltage source to 0 is the same as replacing it with a wire connecting its two terminals (see page 1-21). This places the two resistors *in parallel*, with total current i_{out} flowing *upward* through them. Consequently, $(-v_{out})$ is the voltage drop across them, so clearly r_{out} must be equal to the parallel resistor combination:

$$r_{out} = R_1 R_2 / (R_1 + R_2)$$

Note that the same result would obtain if we simply think of zeroing the embedded voltage source, then treating the output port as though it were an input and calculating its input resistance. This procedure is generally the correct way to calculate a linear circuit's output resistance (or, more generally, its output impedance). That this method works is yet another example of the principle of linear superposition. As with the determination of the input resistance of a network or sub-circuit, the answer will generally depend on how all its other ports are *terminated* (what sorts of circuits are attached to them).

CALCULATING A PORT'S OUTPUT RESISTANCE

To determine the output resistance of a network port, connect all other ports to whatever circuits they will interface to. Then set all independent driving voltage and current sources to 0, replacing them as described on page 1-21. Finally, treat the port in question as an input, and calculate its input resistance — this result will equal r_{out} (equation 1.14).

Thevenin and Norton models of power sources or outputs

We can model most real power sources or circuit output ports as ideal voltage or current sources combined with a finite resistance (or *impedance*, as we'll use in Experiment 2) so that their output resistance matches that of the real source, as shown in Figure 1-36. These are called *Thevenin* and *Norton* equivalent circuits (or *models*) of a real source. For example, many commercial signal generators (including the one you will use in lab) have a 50 Ohm output resistance and are best represented using a Thevenin model (voltage source in series with $R_s = 50\Omega$).

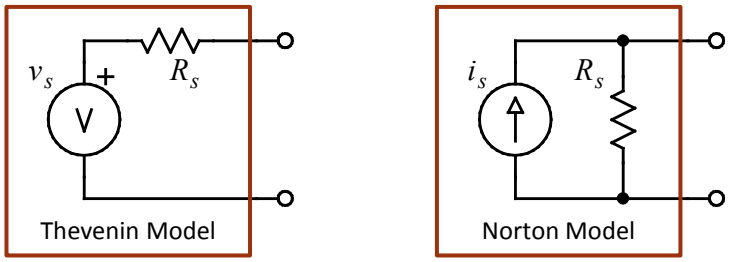


Figure 1-36: Thevenin and Norton models of power sources with finite output resistance. Which model you choose depends on your application, but if the output resistance R_s is relatively small, then the Thevenin (voltage source) model is probably the correct choice. Use the Norton (current source) model if R_s is large. The two models are completely equivalent (have the same output regardless of load) if the source resistance R_s is the same for both and the two source amplitudes are related by $v_s = R_s\, i_s$ (see also Prelab Exercise 4 on page 1-28). The laboratory signal generator has R_s = 50 Ohms.

Experiment 2

Impedance and frequency response

Copyright © Frank Rice 2014
Pasadena, CA, USA
All rights reserved.

Experiment 2

Impedance and frequency response

The first experiment has introduced you to some basic concepts of analog circuit analysis and amplifier design using the "ideal" operational amplifier. To make further progress we need to add a couple of powerful tools for understanding and describing the behavior of our analog circuits — the concepts of *impedance* and *frequency response*. These ideas live in the frequency domain: that world of "eternal" sinusoidal waves which is the Fourier transform space of our circuits' time-varying voltages and currents.

The amplitude and phase of any given sinusoid will be represented as a vector which takes the form of a complex number called a *phasor*. Linear differential equations of time become complex-valued algebraic expressions of frequency in this transformed space; the ubiquitous *capacitor* and *inductor* have voltage-current relationships in the frequency domain which look like Ohm's law. Thus we arrive at the complex-valued *impedance* as a generalization of resistance which includes these additional circuit elements.

Generalized voltage dividers which include capacitors and/or inductors as well as resistors have gains which usually vary with frequency — this sort of variation defines the *frequency response* of a circuit or network, which is displayed using a *Bode plot*. Designing a circuit to only pass signals within a specified frequency range leads to the concept of a *filter*. We can use a circuit's frequency response to predict how it will respond to abrupt changes in its inputs, which leads to the related concepts of *transient response* and *settling time*.

Finally, we have the tools we need to study the behavior of a real operational amplifier, which can only approximate the ideal op-amp. We see how our actual amplifier circuits have frequency limitations which depend on their gains; we also learn how to design amplifiers which can *integrate* or *differentiate* their time-varying inputs. By the end of this experiment you will be able to design circuits which can apply any linear integro-differential operator to an input signal — a basic building block of the *analog computer* and the *PID* (proportional-integral-differential) *feedback controller*.

CAPACITORS AND INDUCTORS

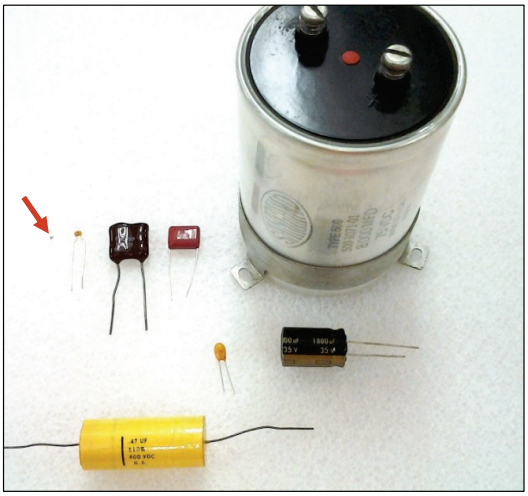
Capacitors

Several capacitors are shown in Figure 2-1. Conceptually, any capacitor consists of two relatively large conductive surfaces separated by a thin insulator; each conductive surface is connected to a terminal of the device (the schematic symbol for a capacitor is suggestive of this arrangement). As current flows through a capacitor, charge builds up on one surface while an equal and opposite charge accumulates on the other; the net charge of the element remains 0. The resultant electric field between the surfaces sets up a potential difference (voltage) between them and, consequently, the two terminals of the capacitor. This voltage difference is proportional to the charge of a surface, and the rate of change of this charge is proportional to the current flowing through the capacitor. Thus the relation between current and voltage for a capacitor is a differential one, as shown in Figure 2-2 and equation 2.1.

2.1

$$\text{Capacitor: } i(t) = C\frac{d}{dt}v(t)$$

The constant of proportionality, C, is called *capacitance* and has the SI unit *Farad*: 1 Farad = 1 Ampere/(Volt/sec) = 1 Coulomb/Volt = 1 sec/Ohm. For an ideal capacitor the capacitance C is a real, constant number (independent of voltage, current, or frequency) with $C \geq 0$. Equation 2.1 is the *defining relation* for an ideal capacitor; the actual capacitors you will use behave in a nearly ideal manner as long as you are careful with them — because of the thin insulating layer in a capacitor, it is subject to damage from internal electrical



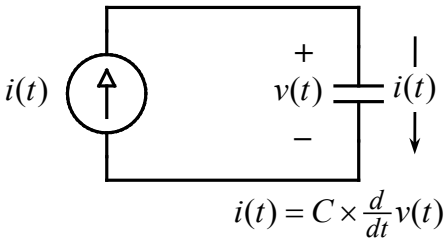


Figure 2-1 (left): A variety of capacitors, from the tiny (.08 inch × .05 inch) surface-mount component barely visible near the left side of the photo (see arrow) to the soda-can-sized power supply filter capacitor dominating the scene. The small, orange capacitor just to the right of the surface-mount device is a multi-layer ceramic type like you will mostly use for your designs.

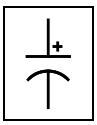
Figure 2-2 (right): A simple circuit illustrating the differential voltage-current relation for a capacitor with value *C*. Note the use of a *current source* to drive the capacitor in this case. Also shown are the usual conventions for the polarity of the *change in* voltage and direction of the current flow when the current $i(t) > 0$.

discharge (a spark) if too high a voltage is placed across it. *The higher the capacitance required, the lower the voltage limit for a capacitor of a given physical size*. This is the primary driver of the sizes of the capacitors shown in Figure 2-1. Each capacitor has a maximum voltage limit which is often written on the capacitor body. As the voltage across the capacitor approaches this limit, its behavior becomes much less ideal, and equation 2.1 could be quite inaccurate in this case.

> The relation between the current and voltage of a capacitor (equation 2.1) is the same whether we control the current through it and measure its resulting voltage, or control its voltage and ask what the current must be. Thus, if the voltage across a capacitor is constant, the current through it must vanish; conversely, rapid changes in a capacitor's voltage require large currents to change its charge: the higher the capacitance C, the higher the current required. As we will see in a later section, capacitors are very useful to separate high-frequency, rapidly changing signals from low-frequency signals and constant, DC voltages. Using a higher capacitance lowers the frequency around which this separation occurs.

Another consideration when choosing a capacitor is the nature of the insulator it employs. To get large capacitances, manufacturers have chosen to chemically deposit extremely thin insulating layers using electrolysis. Because this chemical process may be inadvertently reversed by you, the user, these *electrolytic* capacitors must be used carefully! Electrolytic capacitors have an inherent *polarity*; the voltage across such a capacitor must never have the opposite polarity, or it will be permanently damaged. If you look carefully at Figure 2-1 you can see the polarity markings on the electrolytic capacitors — the very large one in the figure and the one right below it.

A Farad (F) is a very large capacitance. The most useful capacitor range is within a couple of orders of magnitude of a *nanofarad* (nF, 10^{-9} F). Capacitances of more than about a *microfarad* (μF or uF, 10^{-6} F) usually require an electrolytic capacitor. An electrolytic capacitor is denoted in a schematic by a symbol with a polarity mark (+) as shown at right.



> Large capacitors (with values exceeding about 1μF) tend to behave in a less than ideal manner at high frequencies, especially if they are electrolytic. If you need a large value of capacitance but also need it to work well at high frequencies (above 1MHz), then you should place a smaller-valued capacitor (10–100nF) in parallel with the larger one. Capacitors with values of 10nF or less tend to behave very closely to the equation 2.1 ideal.

Inadvertent, stray capacitance exists anywhere in a circuit where two conductors (wires) come close to each other, and this fact has plagued many a beginning circuit designer and builder! Two nearby wires are coupled by a capacitance of about a *picofarad* (pF, 10^{-12} F)

per centimeter of wire length. This unfortunate fact is what will limit the useful frequencies of our circuits to no more than about 10 MHz. *To keep these stray capacitances from spoiling your project, you shouldn't design a circuit to require capacitor values of less than about 33 pF* (or even more, to stay on the safe side).

Inductors

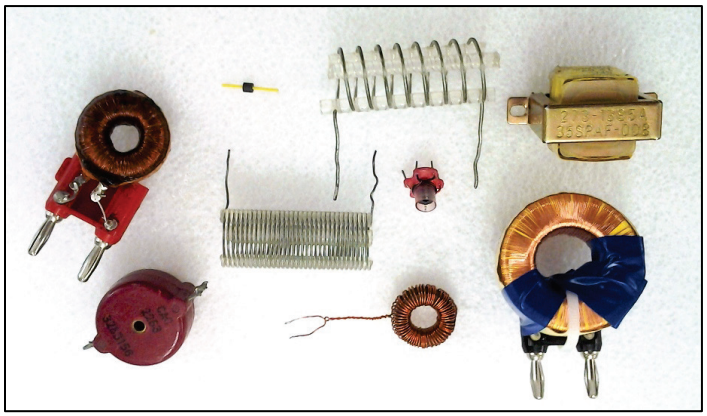
Any length of wire will act as an *inductor*, but unless the wire is formed into a coil, its inductance is likely to be very small (Figure 2-3). A wire carrying a current generates a surrounding magnetic field whose lines of flux encircle it; forming the wire into a coil or helix will intensify the field near the coil's axis. Any changes in the wire's current will change the total magnetic flux enclosed by the coil's loops, and this changing flux will induce an electric field along the wire. The field produces a potential drop between the ends of the wire which will oppose the change in current, resulting in a differential relationship between an inductor's voltage and current which is the reverse of that of a capacitor:

2.2

$$\text{Inductor: } v(t) = L\frac{d}{dt}i(t)$$

The constant of proportionality, L, is called *inductance* and has the SI unit *Henry*: 1 Henry = 1 Volt/(Ampere/sec) = 1 Ohm second. For an ideal inductor L is a real, constant number (independent of voltage, current, or frequency) with $L \geq 0$.

Equation 2.2 is the *defining relation* for an ideal inductor, but actual inductors may approach this ideal only for a disappointingly narrow range of frequencies and only for small currents,



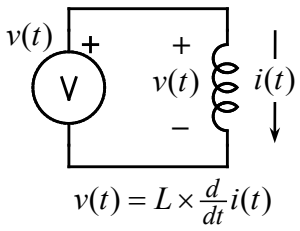


Figure 2-3 (left): A variety of inductors. The helical, air-filled coils are for high-frequency, tuned circuits and filters (above 100 MHz). The toroidal-shaped coils surround high-permeability cores to dramatically increase inductance and reduce the strength of stray magnetic fields. They are suitable for low frequency use. The power supply transformer at upper right is useable only at frequencies of 50–100 Hz. The inductances of these devices range from a few hundred *millihenries* (mH) to less than 100 *microhenries* (µH).

Figure 2-4 (right): A simple circuit illustrating the differential voltage-current relation for an inductor with value L. Also shown are the usual conventions for the direction of the *change in* the current flow and polarity of the induced voltage.

especially if they use a high-permeability core material or have very closely spaced turns of thin wire to increase inductance. As with the Farad, the Henry is a very large unit; reasonably-sized components rarely exceed more than a couple of *millihenries* (mH). With few exceptions, only large electromagnets, power transformers, motors, and solenoids have inductances exceeding 1 Henry.

The relation connecting the current and voltage of an inductor (equation 2.2) is the same whether we control the current through it and measure its resulting voltage, or control its voltage and ask what the current must be. Thus, a constant current through an inductor implies 0 voltage across it, whereas a constant voltage will result in an ever-increasing current.

Rapid changes in the current through an inductor can result in *extremely large induced voltages* as its magnetic field changes. This is especially worrisome should the current through an inductive device (such as an electromagnet, solenoid, or relay) be suddenly interrupted — induced voltages of hundreds and even thousands of volts may be generated, *causing extensive circuit damage and even lethal shocks*.

Practical problems with the construction of an effective inductor lead to the elements' many departures from ideal behavior. One of the most serious of these problems is that closely-spaced turns of wire have a stray capacitance coupling them — this capacitance is effectively in parallel with the element's inductance, so at high frequencies the oscillating voltage difference between adjacent windings will produce a current that bypasses the desired flow around each winding, thus reducing the magnetic field and making the device seem more like a capacitor than an inductor! This change in behavior from predominantly inductive to capacitive is usually quite abrupt as frequency is increased through the coil's *self-resonant frequency*. Because closely-spaced windings of thin wire have higher stray capacitance and result in lower self-resonant frequencies, inductors designed for high-frequency use avoid this design (look again at the high-frequency inductors in Figure 2-3). You will investigate this important phenomenon of *resonance* in a later experiment (although we'll soon see an early example when we examine the differentiator circuit).

Many real inductors can be quite lossy and nonlinear. The long, thin wire used to make an inductor may have noticeable resistance, resulting in an additional voltage drop caused by ohmic losses: this resistance is effectively in series with the inductance. Another source of power loss comes from nonlinear *hysteresis* in the ferromagnetic material used in the inductor's core as the magnetic field oscillates due to AC current flow — this effect is by far the dominant source of loss in many inductors. In addition, ferromagnetic materials are subject to *saturation*, a related nonlinear process which reduces their effectiveness in high magnetic fields: the inductance L changes in the presence of large currents, so that the simple, linear relationship (2.2) could be a poor model of an actual element's behavior.

THE FREQUENCY DOMAIN

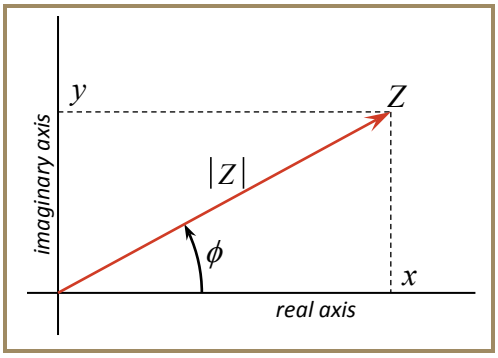
Quick review of complex numbers

Here is a brief review of the elementary mathematics of *complex numbers* which we will need to manipulate frequency-domain signals and impedances.

> Note that we use $j \equiv \sqrt{-1}$ rather than the mathematician's or physicist's "i." We do this to avoid confusion with a current $i(t)$ (this is standard electrical engineering practice, and is the convention used by nearly all electronics texts).
>
> We will also use "f" rather than the Greek "ν" to represent frequencies (in Hertz, i.e. cycles/second) so that there is no confusion with a voltage $v(t)$; the Greek "ω" will, as usual, represent *angular* frequencies (in radians/second, i.e. seconds^{-1}). We'll also generally use upper-case letters ("Z", "Y", "W") to represent complex-valued functions (usually of frequency), while lower-case letters shall be reserved for real-valued functions (usually of time). Obvious exceptions will be to use R, L, and C to represent resistance, inductance, and capacitance.

Given: $Z, W \in \mathbb{C}\,;\; x, y, \phi \in \mathbb{R}$

The complex number Z may be represented as a vector in a Cartesian plane as shown at right. We then have the following rectangular and polar coordinate representations of Z:



$$Z \;=\; x + j\,y \;=\; |Z|\,e^{j\phi}$$

$$\mathrm{Re}[Z] \equiv x\,; \quad \mathrm{Im}[Z] \equiv y$$

$$|Z| \;=\; \sqrt{x^2 + y^2}\,; \quad e^{j\phi} = \cos\phi + j\sin\phi\,; \quad x = |Z|\cos\phi\,; \quad y = |Z|\sin\phi\,; \quad \phi \;=\; \arctan(y/x)$$

Terminology:

x: real part of Z; y: imaginary part of Z; $|Z|$: magnitude of Z; ϕ: phase of Z

Conjugate, magnitude, reciprocal:

$$Z^* \;=\; \mathrm{conj}[Z] \equiv x - j\,y \;=\; |Z|\,e^{-j\phi}\,; \quad |Z|^2 \;=\; Z\,Z^* \;=\; Z^*Z; \quad |1/Z| \;=\; 1/|Z|$$

$$\mathrm{Re}[Z] \;=\; \tfrac{1}{2}(Z + Z^*); \quad \mathrm{Im}[Z] \;=\; \tfrac{1}{2j}(Z - Z^*); \quad \mathrm{Re}[Z^*] \;=\; \mathrm{Re}[Z]; \quad \mathrm{Im}[Z^*] \;=\; -\mathrm{Im}[Z]$$

$$\frac{1}{Z} \;=\; \frac{1}{x + j\,y} \;=\; \frac{x - j\,y}{x^2 + y^2} \;=\; \frac{1}{|Z|}\,e^{-j\phi}; \quad \left(\frac{1}{Z}\right)^* \;=\; \frac{1}{Z^*}\,; \qquad \boxed{\frac{1}{j} \;=\; -j}$$

Products of two complex numbers:

Let: $Z = x_Z + j\,y_Z = |Z|\,e^{j\phi_Z}$; $W = x_W + j\,y_W = |W|\,e^{j\phi_W}$

$ZW = |Z|\,|W|\,e^{j(\phi_Z+\phi_W)} = x_Z x_W - y_Z y_W + j(x_Z y_W + x_W y_Z)$; $(ZW)^* = Z^* W^*$

$|ZW| = |Z|\,|W|$; $|Z/W| = |Z|/|W|$

If we consider Z and W to be two *vectors* in the complex plane, then their products are:

$\vec{Z} \bullet \vec{W} = |Z|\,|W| \cos(\phi_W - \phi_Z) = \mathrm{Re}[Z^* W] = \mathrm{Re}[Z W^*] = x_Z x_W + y_Z y_W$

$\vec{Z} \times \vec{W} / \hat{x} \times \hat{y} = |Z|\,|W| \sin(\phi_W - \phi_Z) = \mathrm{Im}[Z^* W] = x_Z y_W - x_W y_Z$

Linear functions and time derivatives:

If $F(Z)$ is a *complex-valued* function of its complex argument Z, then we can write:
$F(Z) = u(Z) + j\,v(Z)$, where u and v return real values for every Z

If $F(Z)$ is a *linear* function of its complex argument Z, then, conversely, we can write:
$F(Z) = F(x + j\,y) = F(x) + j\,F(y)$ [Note: $F(x)$ may be complex-valued]

Thus, let $t \in \mathbb{R}$ (such as *time*), and let $Z(t) = x(t) + j\,y(t)$. Then:

$$\frac{d}{dt}x(t) + j\frac{d}{dt}y(t) = \frac{d}{dt}[x(t) + jy(t)] = \frac{d}{dt}Z(t);\ \therefore \frac{d}{dt}\mathrm{Re}[Z(t)] = \mathrm{Re}\left[\frac{d}{dt}Z(t)\right]$$

so we may exchange the order of time differentiation and taking the real part.

Frequency-domain representations and phasors

We can represent a sinusoidal function of time (such as an AC voltage or current) using complex numbers:

2.3

If: $y(t) = y_{\max}\cos(\omega t + \phi)$
then: $y(t) = \mathrm{Re}\left[Y\,e^{j\omega t}\right]$; where $Y = y_{\max}\,e^{j\phi}$

The complex "amplitude" Y is called a *phasor* since it determines both the amplitude $y_{\max}$ and phase ϕ of $y(t)$. If the phase ϕ of Y is greater than 0, then the phase of $y(t)$ *leads* the phase of $\cos\omega t$; if $\phi < 0$ the phase *lags* that of $\cos\omega t$.

For our electronic circuits *it will generally be the case* that a phasor $Y = Y(\omega)$, i.e. its magnitude and phase may be functions of frequency. It is not the case, however, that Y is a function of time t, because all the time dependence of $y(t)$ is captured in the $e^{j\omega t}$ term (implying that the sinusoid $y(t)$ has been and will be around forever!). What this implies is that when we consider complex-valued *phasors* as specifications of real-valued, sinusoidal functions of time we are using a *frequency-domain representation* of the functions, that is: we are working in the *Fourier transform space* of functions of time.

In many of the cases we will be interested in, our $y(t)$ will consist of a constant term (its DC component) along with a few, discrete sinusoidal terms (each like the first equation in 2.3). In this case $y(t)$ may be represented by a discrete sum over the sinusoids + the DC term; as a result the frequency-domain representation of the signal $Y(\omega)$ will be the sum of a few discrete phasor values $Y(\omega_k)$, one for each of the sinusoid terms, along with a constant term representing the signal's DC component.

The last (but very important) thing to consider about our frequency-domain representation is how the operation of a linear function on our $y(t)$ will affect its phasor representation $Y(\omega)$. The two prototypical linear operations on $y(t)$ are: (1) multiplying by a constant, and (2) taking a time derivative. *All* linear operations on $y(t)$ will consist of a combination of sums and compositions of these two prototypes and their inverses (note that adding a constant to $y(t)$ *is not a linear operation*). The Fourier transformation (2.3) between $y(t)$ and $Y(\omega)$ is itself a linear operation, as is a sum of such terms (one for each sinusoid in the signal). Since the composition of two linear operations results in another linear operation, we immediately know that a linear function $f(y(t))$ will transform to a corresponding linear function in the frequency-domain $F(Y(\omega))$. The problem at hand is to determine the correspondence between these two functions f and F.

The first linear operation, multiplying by a constant, is trivially easy to handle. A cursory glance at equation 2.3 should make it obvious that multiplying (scaling) $y(t)$ by some (real-valued) constant a will simply result in scaling $Y(\omega)$ by the same constant:

2.4
$$y(t) \rightarrow a\,y(t) \quad \Rightarrow \quad Y(\omega) \rightarrow a\,Y(\omega)$$

Multiplying $Y(\omega)$ by some complex-valued constant $Z = |Z|\,e^{j\phi_Z}$, on the other hand, corresponds to changing *both the amplitude and phase* of $y(t)$ (consider equations 2.3):

2.5
$$Y(\omega) \rightarrow Z\,Y(\omega) \quad \Rightarrow \quad y(t) = |Y|\cos(\omega t + \phi_Y) \rightarrow y(t) = |Z||Y|\cos(\omega t + \phi_Y + \phi_Z)$$

Taking a time derivative of $y(t)$ has a straightforward and quite simple effect on $Y(\omega)$. Consider equations 2.3, in which $y(t)$ is a simple sinusoid. In this case we can evaluate the derivative as it operates on both sides of the second equation of (2.3), keeping in mind that the order of time differentiation and taking the real part may be reversed:

$$\frac{d}{dt}y(t) = \frac{d}{dt}\mathrm{Re}\left[Y(\omega)e^{j\omega t}\right] = \mathrm{Re}\left[\frac{d}{dt}Y(\omega)e^{j\omega t}\right] = \mathrm{Re}\left[j\omega Y(\omega)e^{j\omega t}\right]$$

Thus, taking a time derivative of a function $y(t)$ corresponds to simply multiplying its frequency-domain representation $Y(\omega)$ by $j\omega$. This result obtains even for arbitrary $y(t)$ requiring a more complicated *Fourier integral* representation (see the text describing equations 2.25 on page 2-40).

2.6
$$\boxed{y(t) \rightarrow \frac{d}{dt}y(t) \quad \Leftrightarrow \quad Y(\omega) \rightarrow j\omega Y(\omega)}$$

The implication works both ways, as indicated by the "$\Leftrightarrow$" in (2.6). Similarly, dividing $Y(\omega)$ by $j\omega$ corresponds to an integration of $y(t)$. This result will be critical for our extension of Ohm's law to include capacitors and inductors.

Extending Ohm's law: impedance

Now that we have our powerful frequency-domain and phasor arsenal at our disposal, we consider again the voltage-current relationships for the resistor, capacitor, and inductor. We start by using (2.3) and/or (2.25) to define complex-valued voltage and current phasors $V(\omega)$ and $I(\omega)$ to represent the time-varying voltage across, $v(t)$, and current through, $i(t)$, a two-terminal circuit element. We then express the various voltage-current laws for our elements in terms of the phasors $V(\omega)$ and $I(\omega)$. Ohm's law is easy, since $v(t)$ and $i(t)$ are simply proportional to each other:

2.7 **Resistor:**

$$\boxed{\begin{aligned} v(t) &= R\, i(t) \\ V(\omega) &= R\, I(\omega) \end{aligned}}$$

For the capacitor and inductor, use (2.6) to transform the time derivatives in (2.1) and (2.2):

2.8 **Capacitor:**

$$\boxed{\begin{aligned} C\,\tfrac{d}{dt}v(t) &= i(t) \\ (j\omega C)\,V(\omega) &= I(\omega) \end{aligned}}$$

2.9 **Inductor:**

$$\boxed{\begin{aligned} v(t) &= L\,\tfrac{d}{dt}i(t) \\ V(\omega) &= (j\omega L)\,I(\omega) \end{aligned}}$$

Since ω, L, and C are all nonnegative, real numbers, and the imaginary j has a phase of +90° ($j = e^{j\pi/2}$), we see from (2.8) that for a capacitor, the phase of a sinusoidal current *leads* the voltage by 90°, whereas for the inductor, it's the other way around.

The frequency-domain (phasor) relationships in (2.7), (2.8), and (2.9) all look like Ohm's law: *the voltage and current phasors are simply proportional to each other* (except that in the cases of the inductor and capacitor the proportions are no longer real, and they depends on frequency). Thus we can generalize Ohm's law to include these more general phasor relationships by introducing the concept of *the complex-valued impedance*, $Z(\omega)$, and its reciprocal, the *admittance*, $Y(\omega)$.

Impedance and Admittance

2.10

$$\boxed{\begin{aligned} V(\omega) &= Z(\omega)\, I(\omega) \\ I(\omega) &= Y(\omega)\, V(\omega) \\ Z_R = R;\quad Z_C &= 1/(j\omega C);\quad Z_L = j\omega L \end{aligned}}$$

The real part of an impedance Z is called its *resistive* component. If Z happens to be a positive real number, it is said to be a *pure resistance*. The imaginary part of Z is its *reactive* component. An ideal capacitor or inductor is said to be a *pure reactance*. The real and imaginary parts of an admittance Y are called its *conductance* and *susceptance*, respectively.

> It is important to emphasize that impedance and admittance are *frequency-domain concepts* that describe an algebraic relationship between voltage and current *phasors*. In general, even linear relationships between voltage and current in the time domain are *differential* in character — in the time domain, Ohm's law applies only to a resistor.

With our formal definitions of voltage and current phasors and using our complex-valued impedance concept to extend Ohm's law, we can reexamine the various things we investigated in Experiment 1 to extend the time-domain relationships discovered there (using circuits with only resistors and the ideal op-amp) to more general cases involving capacitors and inductors as well. This extension is particularly simple, because nearly every equation we wrote down in the notes for that experiment expressed a linear relationship between the various combinations of $v(t)$ and $i(t)$ sources and responses. All of those equations remain valid if you make the simple substitutions:

2.11

$$\begin{array}{ll} v(t) \rightarrow V(\omega); & R \rightarrow Z(\omega) \\ i(t) \rightarrow I(\omega); & \dfrac{d}{dt} \rightarrow j\omega \end{array}$$

This mapping from the time domain to frequency-domain phasors includes expressions for series and parallel impedances, the voltage divider equation, Kirchhoff's voltage and current laws, input and output impedances, ideal amplifier gains, and Thevenin and Norton signal source models — just substitute "impedance" for "resistance" and speak of phasors rather than instantaneous values of the voltages and currents. The substitutions in (2.11) are not valid, however, for *nonlinear* expressions such as the power dissipated by a two-terminal element: $P(t) = v(t)i(t)$; we'll consider the correct phasor expression for power dissipation in a later experiment.

PARALLEL *RC* IMPEDANCE V. FREQUENCY

As a first example of a calculation of $Z(\omega)$, consider a parallel combination of a resistor R and capacitor C. We know that the *admittances of parallel elements add*, so the parallel impedance is:

$$\frac{1}{Z(\omega)} = \frac{1}{R} + j\omega C \quad \rightarrow \quad Z(\omega) = R\frac{1}{1 + j\omega RC}$$

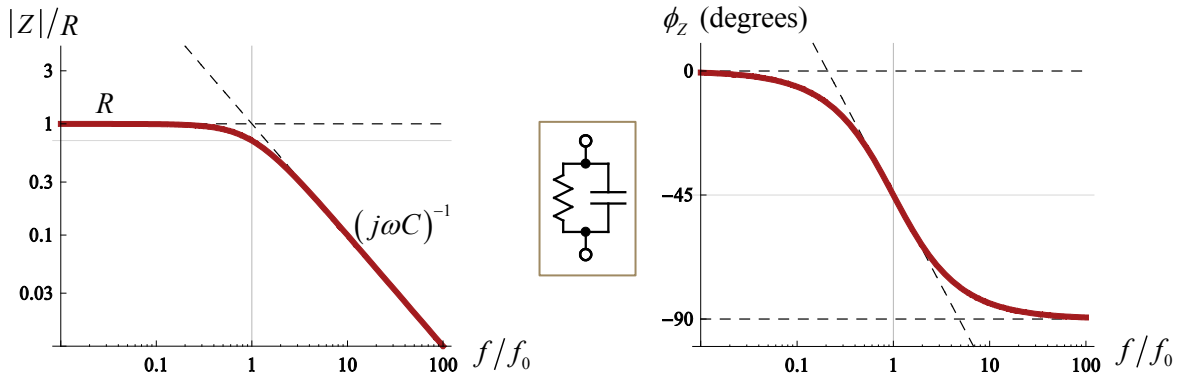


Figure 2-5: Bode plot (magnitude and phase v. frequency) of the parallel *RC* impedance phasor $Z(\omega)$. Frequency is relative to the *RC corner frequency*, $f_0 = 1/(2\pi RC)$. For low frequencies, the impedance is that of the resistor, R. At high frequency, the impedance looks like that of the capacitor, $-j/(2\pi fC)$. The asymptotic resistive and capacitive impedances are shown as dashed lines in the plots. Another dashed line, tangent to the phase response at frequency f_0 ($\phi = -45°$), intersects the two phase asymptotes at frequencies of approximately $f_0/5$ and $5f_0$. *Note the use of log scales for frequency and magnitude*.

The magnitude and phase v. frequency of this parallel *RC* impedance $Z(\omega)$ are plotted in Figure 2-5, called a *Bode plot*, named for Hendrik Bode (1905–1982) of Bell Laboratories. Note the log scales for frequency and magnitude. The *RC corner frequency*, $f_0 \equiv 1/(2\pi RC)$, is that frequency where the impedances of the R and C have the same magnitude, namely $R = 1/(\omega_0 C)$, with $\omega_0 = 2\pi f_0$. At low frequencies ($f \ll f_0$) the impedance approaches that of the resistor, R. At high frequencies ($f \gg f_0$) it approaches that of the capacitor, $1/(j\omega C)$. Note in the Bode plot (Figure 2-5) that these asymptotic behaviors plot as straight lines on the log-log magnitude plot and semi-log phase plot; *the magnitude plot asymptotes intersect at* f_0. At f_0 the magnitude of $Z(\omega)$ differs from the asymptotes by a factor of $\sqrt{2}$, and its phase is exactly half way between the asymptotic phase values.

Moving away from f_0, the magnitude of $Z(\omega)$ approaches its asymptotes relatively quickly — beyond a factor of 10 in frequency away from f_0 the magnitude is well within 1% of them. The phase of $Z(\omega)$ approaches its asymptotes much more gradually, however — a factor of 10 in frequency only brings the phase just within 6° of an asymptote; a factor of 60 is required to come to within 1°. This behavior is typical of simple circuits containing only a single capacitor or inductor.

Simple RC low-pass and high-pass filters

In the last section we considered the frequency dependence of the impedance phasor $Z(\omega)$ of an RC combination. Now we consider how a network's *gain* (or *transfer function*) may also be represented using a phasor, $G(\omega)$, called the *frequency response* of the network. We start with particularly simple but useful circuits, the *RC filters* shown in Figure 2-6. Clearly, these circuits are just voltage dividers with a capacitor replacing one of the resistors. Since the capacitor's impedance is inversely proportional to frequency (equation 2.10), we would expect that for the high-pass filter (top circuit) V_{out} will approach V_{in} at high frequencies but will approach 0 at low frequencies. The opposite would be the case for the low-pass filter (bottom circuit). Using the voltage divider equation gives:

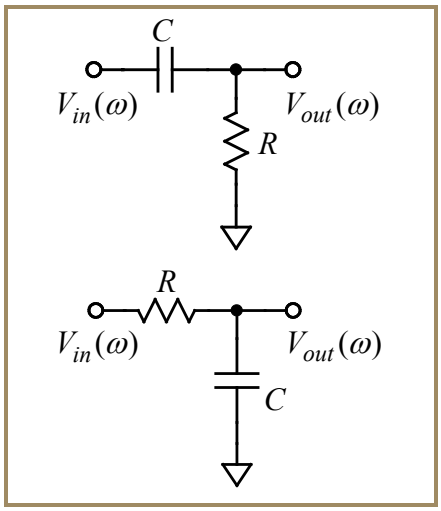


Figure 2-6: Simple *RC* high-pass (top) and low-pass (bottom) filters, which are just a voltage divider circuits.

$$G(\omega) \equiv \frac{V_{out}(\omega)}{V_{in}(\omega)} = \left\{ \frac{R}{R+Z_C} \text{ (high-pass)} ; \quad \frac{Z_C}{R+Z_C} \text{ (low-pass)} \right\}$$

2.12

***RC* high-pass filter response**

$$G(\omega) = \frac{1}{1 - j\,\omega_0/\omega}; \quad \omega_0 = \frac{1}{RC}$$

***RC* low-pass filter response**

$$G(\omega) = \frac{1}{1 + j\,\omega/\omega_0}; \quad \omega_0 = \frac{1}{RC}$$

The derivation of (2.12) from the equations above it is left to the exercises. Figure 2-7 presents a Bode plot of the high-pass filter's response; the low-pass filter's response is identical to the Bode plot of our parallel RC impedance in Figure 2-5 (on page 2-11), except that $G(\omega)$ replaces $Z(\omega)/R$ in that plot. As expected, the high-pass filter gain decreases as f/f_0 below the RC corner frequency, f_0; above this frequency it becomes constant (with $G = 1$). Also note the correlation between the asymptotic phase and the slope of the response. This correlation between response slope and phase is the consequence of the general rule stated in the highlighted box below the figure on page 2-13.

Note that in Figure 2-7 the gain is also plotted in *decibels* (dB), a logarithmic scale commonly used for expressing a network's gain. 10dB corresponds to a signal *power gain* of a factor of 10; since power is proportional to the square of a signal's amplitude, a factor of 10 in *amplitude gain* corresponds to a power gain of 100, which is 20dB. Decibels are defined in the other highlighted box on page 2-13.

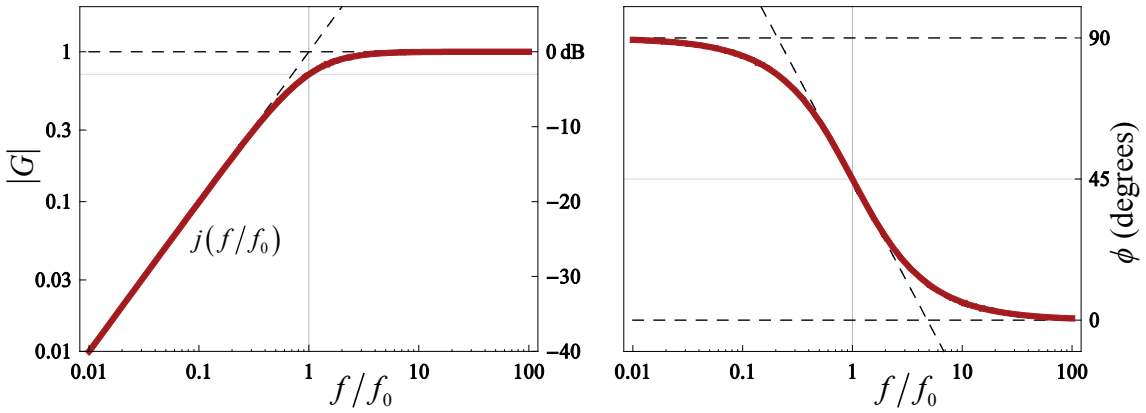


Figure 2-7: Bode plot of the RC high-pass filter response from equation 2.12. The asymptotes of the response are shown as dashed lines. Frequency is relative to the *RC corner frequency*, $f_0 = 1/(2\pi RC)$, as in Figure 2-5. Gain is also shown in *decibels* (dB): G(dB) = 20 $\log_{10}(G)$. This response is typical of that of a *single pole* filter.

RESPONSE SLOPE AND ASSOCIATED PHASE SHIFT IN BODE PLOTS

For filters constructed using only R's, C's, and L's, the slope of the asymptotic magnitude (on a log-log Bode plot) and the associated asymptotic phase obey a simple rule: if the magnitude goes as (*frequency*)^n for some integer *n*, then the associated phase shift is *n*×90°.

This rule obtains because the impedances of the C's and L's only involve factors of $j\omega$ (you never find ω without j multiplying it!), and the imaginary factor j changes the phase by 90°.

USING DECIBELS TO EXPRESS GAIN

Gain is often expressed in the logarithmic decibel (dB) scale. The conversion to dB for amplitude gain (such as the voltage gain of an amplifier) is:

$$\text{dB} = 20 \log_{10} \left| V_{out}/V_{in} \right|$$

Gains of less than 1 will have negative dB values: 1/10 → −20dB, etc. A change of 1dB corresponds to a 12% gain change; ±0.1dB corresponds to about a 1% gain variation. Other common dB values and their approximate conversions:

6dB → ×2 10dB → ×3 (actually closer to π) 14dB → ×5

Each simple *RC* filter has a *passband* which begins at f_0 and extends in frequency from there. At f_0 the gain is only $1/\sqrt{2}$ (−3dB) of its average value well into the passband; −3dB gain points are often used to specify the passband frequency limits of an amplifier or filter. Beyond the passband the asymptotic filter slope in this case is 20dB/decade (or 6dB/octave) for both filters, which is characteristic of what is called a *single pole response*: there is a

single value of $j\omega$ for which either of the denominators of the gain expressions (2.12) vanishes, namely $j\omega = -\omega_0$. The significance of this "pole" will be discussed shortly.

USING THE *RC* HIGH-PASS FILTER: BLOCKING A DC SIGNAL COMPONENT

As an important example of the use of the simple high-pass filter, Figure 2-6, consider this problem: a tiny time-varying signal must be amplified by at least a factor of 100 to be accurately measured, but this signal is a continual, small fluctuation of an otherwise constant (DC) voltage of about 5V. Obviously, amplifying the DC component by 100 as well would result in an output of 500V, which won't work. Thus we need to amplify the AC (fluctuating) signal while blocking the DC component. The solution to our problem is easy: just add a filter to the input of a noninverting op-amp amplifier circuit:

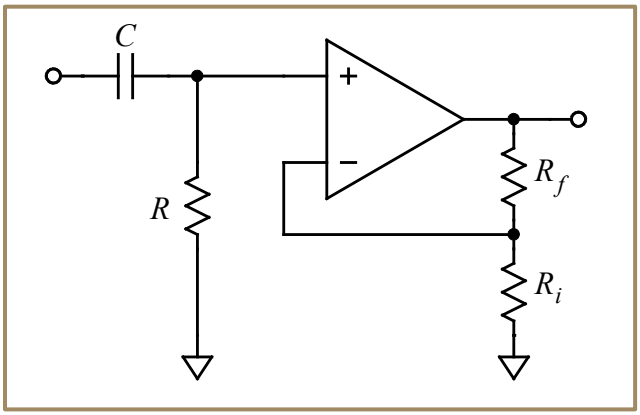


Figure 2-8: Using AC coupling to prevent unwanted amplification of the DC component of the input. The low-frequency cutoff of the amplifier is determined by the filter's *RC* corner frequency, $f_0 = 1/(2\pi RC)$. The gain within the passband is determined by the amplifier's feedback resistors: $G = 1+(R_f/R_i)$. The input impedance of the amplifier (in the passband) is *R*.

This technique is so common that it has a special name: *AC coupling* an amplifier. To choose proper values for the filter *R* and *C* requires some thought:

1. You must determine the lowest frequency component in the input signal you need to measure accurately so that you can determine the required *RC* corner frequency of the input filter.
2. The lower the *RC* corner frequency, $f_0 = 1/(2\pi RC)$, the longer the *settling time* of the filter output: sudden changes in the DC component will be passed through the filter until the capacitor can charge up to the new DC voltage value. The settling time is determined by the *RC time constant* $\tau_{RC} = RC = 1/\omega_0$ — the output *relaxes* back toward 0 as $e^{-t/\tau}$, so it takes almost $5\tau_{RC}$ for the output to settle back to less than 1% of the change in the input [remember, $\omega_0 \equiv 2\pi f_0 = 1/(RC)$].
3. Well within the filter's passband, the impedance of the capacitor is much smaller than that of the resistor, *R*, so *R* becomes the input impedance of the amplifier; *R* must be

chosen large enough to ensure that the amplifier does not draw too much current from the signal source.

4. The voltage rating of the capacitor *C* should be at least twice the input's maximum expected voltage magnitude (the DC voltage if the AC fluctuations are small). If the required capacitance is large, then an electrolytic (polarized) capacitor may be needed. *The polarity of an electrolytic capacitor should match the input signal's DC voltage polarity*. If good performance at high frequencies is important, then a small-valued capacitor (10–100nF) with the proper voltage rating should be placed in parallel with the electrolytic.
5. The resistor *R* must be included in the circuit! (you can't let $R = \infty$) The reason for this limitation is that *both op-amp inputs must have some DC connection to ground, no matter how roundabout or indirect*, or your real op-amp's output will immediately drift off to its voltage limit (caused by a small op-amp *input bias current*, discussed later).

Transient response of a single-pole filter

So far we have investigated the frequency responses of single-pole filters, both high-pass and low-pass. When discussing the design of a high-pass RC filter it was mentioned that the *settling time* of the filter's output was related to its corner frequency. To investigate this issue, we need to understand how a circuit's frequency-domain response is manifest in its time-domain behavior, or *transient response*. Consider our single-pole frequency responses (assume a gain of 1 in the amplifier or filter's passband; ω_0 is the corner frequency):

High-pass: $$\left(1 - j\frac{\omega_0}{\omega}\right)V_{out} = V_{in} \quad \Rightarrow \quad \left(\frac{j\omega}{\omega_0} + 1\right)V_{out} = \frac{j\omega}{\omega_0}V_{in}$$

Low-pass: $$\left(\frac{j\omega}{\omega_0} + 1\right)V_{out} = V_{in}$$

These equations are easy to derive from equations 2.12. Converting these equations to their time-domain equivalents is easy: just replace $j\omega$ with a time derivative, so:

2.13 **High-pass:** $$\frac{1}{\omega_0}\frac{d}{dt}v_{out}(t) + v_{out}(t) = \frac{1}{\omega_0}\frac{d}{dt}v_{in}(t)$$

2.14 **Low-pass:** $$\frac{1}{\omega_0}\frac{d}{dt}v_{out}(t) + v_{out}(t) = v_{in}(t)$$

The complementary solution is the same for these two differential equations: an exponential "relaxation" toward equilibrium with time constant $\tau = 1/\omega_0$:

2.15 $$\Delta v(t) \propto e^{-t/\tau}; \quad \tau = 1/\omega_0$$

In the case of an *RC* filter, $\tau = RC$, thus we speak of a circuit's *RC time constant*. If we suddenly change v_{in} and then hold it constant for a long time ($\gg \tau$), then the time derivatives in equations 2.13 and 2.14 will have vanished, and we see that $v_{out} \to 0$ for the high-pass filter, whereas $v_{out} \to v_{in}$ for the low-pass. Conversely, during the initial rapid change in v_{in} the high-pass filter passes this change through to the output, but the low-pass filter does not. These behaviors are illustrated in Figure 2-9. Note that the transient response time dependence (2.15) is also given by $e^{j\omega t}$ if we substitute the "pole" value $j\omega = -\omega_0$. This observation leads to the use of a *Laplace transform* rather than a *Fourier transform* to express our phasors, which is the method used in most electrical engineering or control systems texts (and is very briefly described in the section starting on page 2-40).

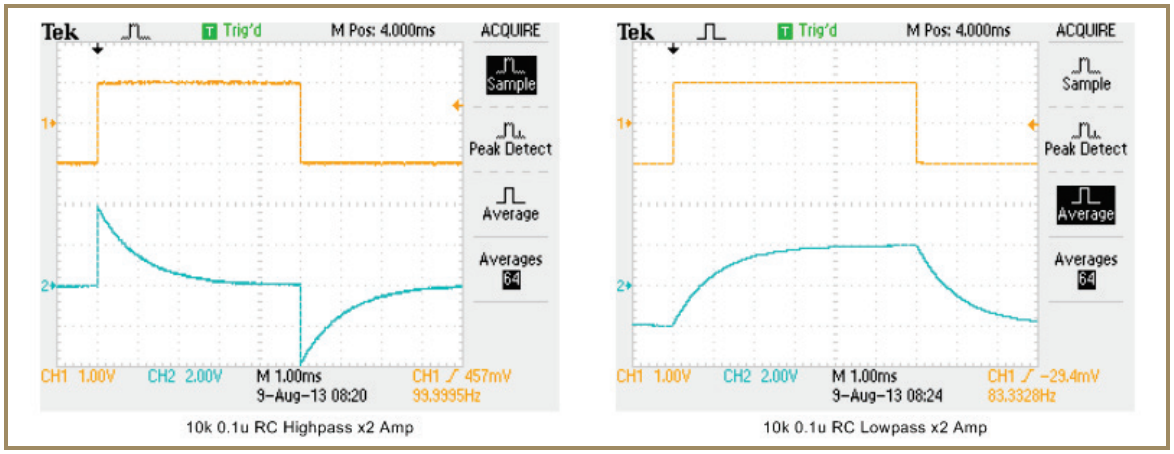


Figure 2-9: Oscilloscope recordings of the transient responses of single-pole high-pass (left) and low-pass (right) filters to a square wave input. The *RC* filters' outputs were isolated and amplified by a noninverting op-amp amplifier with gain = 2. The *RC* time constant for both filters was 1ms.

Look again at our frequency-domain expressions. For the high-pass filter, if $\omega \ll \omega_0$, then $V_{out} \propto j\omega V_{in}$ and the time-domain expression (2.13) simplifies to $v_{out} \propto dv_{in}/dt$. Thus at low frequencies the high-pass filter acts as a *differentiator*. The converse holds for the low-pass filter: if $\omega \gg \omega_0$, then $dv_{out}/dt \propto v_{in}$ and we have an *integrator*. We can do even better if we add an op-amp to the circuit, as will be explored in a later section.

THE REAL, FINITE-GAIN OP-AMP

Approximating the ideal

The operational amplifier model discussed in Experiment 1 is, of course, an idealization of an op-amp's real behavior. How do actual op-amps achieve this nearly ideal performance? Let's consider a more realistic model of the amplifier: it has a very large, but finite, differential gain which is a function of frequency: $g(\omega)$ (note that in this case we're violating our informal convention of using only upper-case letters for phasors). We assume that otherwise the amplifier remains ideal, in the sense that neither op-amp input draws current, and that the op-amp output voltage phasor is only dependent on the difference in its two input voltages:

$$V_{out} = g \times (V_+ - V_-)$$

Now we see from this equation that if the op-amp's *open-loop gain* g is large, then a very small difference in the two input voltages can result in a reasonably large value for V_{out}. Consider the effect this would have on our noninverting amplifier configuration, shown in Figure 2-10 using generic feedback elements with impedances Z_f and Z_i. An ideal op-amp in this configuration would provide an overall gain of $G(\omega) = 1 + Z_f / Z_i = V_{out} / V_-$ for the circuit, but the actual amplifier's finite gain changes this:

$$V_{out} = g(V_+ - V_-) = g(V_{in} - V_{out}/G) \quad \rightarrow \quad V_{out}(1 + g/G) = gV_{in}$$

$$G_f \equiv \frac{V_{out}}{V_{in}} = \frac{1}{1/g + 1/G}$$

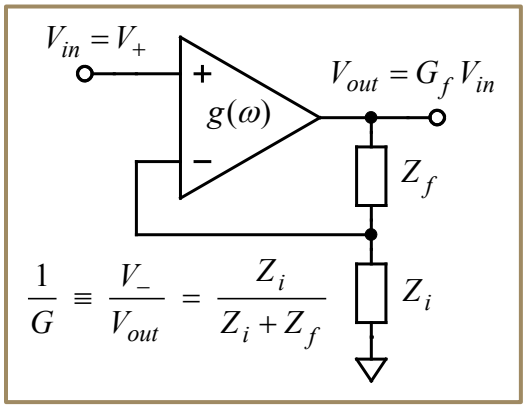


Figure 2-10: The generic noninverting amplifier configuration. The op-amp has finite *open-loop gain* $g(\omega)$. $G(\omega)$ is defined to be the gain of the circuit using an ideal op-amp; $G_f(\omega)$ is the actual *closed-loop* circuit gain.

So if we define G_f to be the amplifier circuit's gain when the finite-gain op-amp is used, whereas G is the circuit's gain with an ideal op-amp, then the noninverting amplifier's gain is given by:

Noninverting configuration gain (with finite op-amp open-loop gain)

$$\frac{1}{G_f(\omega)} = \frac{1}{G(\omega)} + \frac{1}{g(\omega)} \tag{2.16}$$

Equation 2.16 shows that the ideal amplifier circuit gain G combines with the op-amp's finite gain g like *two parallel impedances* to give the actual, resulting gain G_f. As the op-amp's gain $g \to \infty$, the circuit behaves like an ideal op-amp circuit would; for those frequencies where $|g(\omega)| \gg |G(\omega)|$, the amplifier circuit's performance is nearly ideal. The circuit gain G_f is called the *closed-loop gain*, because negative feedback "loops" the amplifier output back around to its input.

Examine expression (2.16) again, however, with the understanding that gain phasors are *complex numbers* and not simply real, positive values like a resistance is. For some range of frequencies it could be the case that $g(\omega) \approx -G(\omega)$, in which case the sum in (2.16) could get quite small, and the actual circuit gain would be quite a bit larger than expected: $|G_f(\omega)| > |G(\omega)|$. This phenomenon is referred to as *gain peaking* and can cause problems for the unwary.

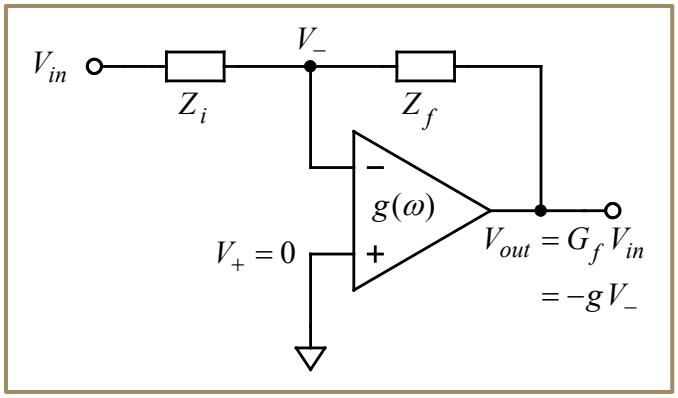


Figure 2-11: The generic inverting amplifier configuration. The op-amp has finite open-loop gain $g(\omega)$. $G(\omega)$ is defined to be the gain of the circuit using an ideal op-amp; $G_f(\omega)$ is the actual closed-loop gain.

Going through a similar calculation for the inverting amplifier configuration, Figure 2-11, results in a similar expression in the case of finite op-amp gain; in this case $G(\omega) \equiv -Z_f / Z_i$, appropriate for an ideal inverting amplifier:

Inverting configuration gain (with finite op-amp open-loop gain)

$$\frac{1}{G_f(\omega)} = \frac{1}{G(\omega)} - \frac{1}{g(\omega)} + \frac{1}{G(\omega)\,g(\omega)} \tag{2.17}$$

(Note the extra term in this expression and that we must be careful about the '−' signs) As expected, the closed-loop behavior becomes ideal as $g \to \infty$. Again, gain peaking will occur any time $g(\omega) \approx G(\omega)$. Otherwise, as long as the op-amp gain g is large, V_- will be quite small, and the node joining the *–Input* to the two impedances will still be a virtual ground (at least approximately); the amplifier circuit's input impedance will also be reasonably close to Z_i, as in the ideal op-amp case.

Frequency response of real amplifier circuits

Given these revised results which include the effect of finite op-amp open-loop gain, we now ask what the frequency response function of an inexpensive, reasonably well-designed op-amp looks like. Consider the Texas Instruments TL082, the one you used in Experiment 1, whose frequency response is shown in Figure 2-12. At very low frequencies its gain is typically 2×10^5, but its response rolls off as $1/j\omega$ over most of its useable frequency range because the op-amp contains a single-pole, low-pass *RC* filter as part of its internal circuitry. This so-called *dominant pole* frequency response keeps the amplifier from exhibiting high frequency instability or oscillations, even when it is configured with 100% feedback (as in a voltage follower circuit). This sort of frequency response is called *unity gain compensation*, and it is what you will want for the vast majority of op-amp applications.

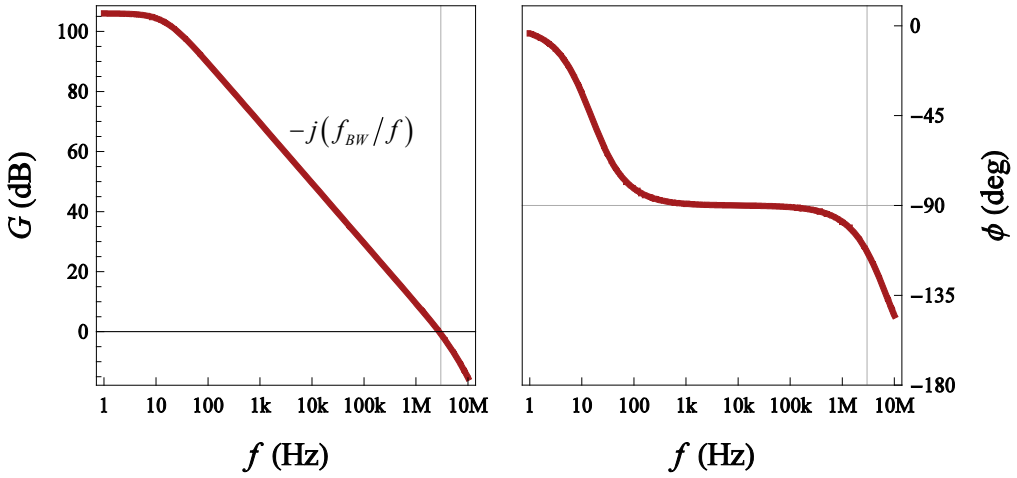


Figure 2-12: Bode plot of the TL082 op-amp open-loop gain, $g(\omega)$. The op-amp's *gain-bandwidth product*, f_{BW}, is 3MHz; the DC open-loop gain is 2×10^5.

The frequency at which $|g(\omega)| = 1$ is called the op-amp's *gain-bandwidth product*, f_{BW}. Except for very low frequencies and very high frequencies, the op-amp's open-loop gain is given by $-j(f_{BW}/f)$, as shown in the Bode plot. The gain-bandwidth product defines the upper frequency limit for the op-amp's usefulness, but, as we will soon see, circuits with substantial gain will have a much lower frequency limit.

The open-loop gain of a unity-gain compensated op-amp such as the TL082 may be approximated in the region where its $|g(\omega)| > 1$ as a parallel combination of two terms: the

very large, constant DC gain g_{DC} and the AC gain $g_{AC} = -j(f_{BW}/f)$. Thus the op-amp's open-loop gain may be modeled as in (2.18).

Simple op-amp open-loop gain model

2.18

$$\frac{1}{g(f)} = \frac{1}{g_{DC}} + j\frac{f}{f_{BW}}$$

In this model we have chosen to ignore the additional complication that the actual op-amp shows a more rapid gain roll-off and additional phase shift at high frequencies where $|g(\omega)| \lesssim 1$. The two parameters g_{DC} and f_{BW} are found in the op-amp device's *data sheet*; for the TL082: $g_{DC} = 2\times10^5$ and $f_{BW} = 3\,\text{MHz}$.

Consider a simple, noninverting amplifier configuration with closed-loop gain given by (2.16) and with a constant *ideal* closed-loop gain $G(\omega) = G_{DC}$. With our open-loop gain model (2.18), the closed-loop amplifier circuit gain is thus:

$$\frac{1}{G_f(f)} = \left(\frac{1}{G_{DC}} + \frac{1}{g_{DC}}\right) + j\frac{f}{f_{BW}} = \frac{1}{G'} + j\frac{f}{f_{BW}} = \frac{1}{G'}\left(1 + j\frac{f}{f_{BW}/G'}\right)$$

Note that in the final expression above G' is just the parallel combination of the op-amp's DC open-loop gain g_{DC} and the desired amplifier closed-loop gain G_{DC} (in most cases $g_{DC} \gg G_{DC}$, so we can just use G_{DC} instead of G'). The closed-loop gain has a simple, single-pole, low-pass frequency response similar to Figure 2-12, but with a −3dB corner frequency of $f_0 = f_{BW}/G'$, as illustrated in Figure 2-13 for closed-loop gains of 1, 10, and 100. Since $f_{BW} = G'f_0$, you see why f_{BW} is called the op-amp's *gain-bandwidth product*!

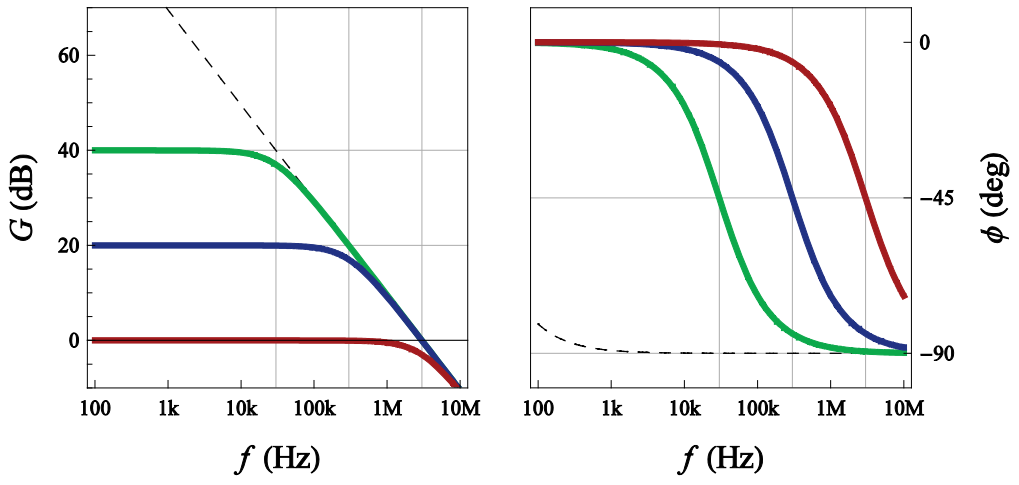


Figure 2-13: Bode plots of noninverting amplifiers with closed-loop gains of 100 (Green), 10 (Blue), and 1(Red). The –3dB corner frequencies are at f_{BW}/G_{DC}. Also shown (dashed) are the TL082 op-amp open-loop gain and phase using the simplified model given by (2.18).

If you need large bandwidth in a high-gain amplifier, you can achieve this by either: (1) cascading several lower-gain amplifier stages (the lower gain will give higher bandwidth), or (2) using a faster op-amp (one with a higher gain-bandwidth product). The first alternative is addressed in the exercises; which choice you make for you own designs may depend on other factors which you will learn to appreciate as you gain more experience with these devices.

Slew rate limitations

If you look back at our cartoon model of the ideal op-amp in Experiment 1 (Figure 1-15 on page 1-13), you can think of the frequency response limitation of a real op-amp as equivalent to a finite "reaction time" on the part of the little technician controlling the op-amp's output voltage — he can't quite keep up with rapid oscillations in $v_+ - v_-$, so the output gets smaller and is delayed in phase as the input frequency increases.

There is another limit to an op-amp's speed which may seem like another manifestation of its frequency response, but is actually an unrelated phenomenon: the output *slew rate* limit. To relate this limit to our little technician, slew rate describes how fast he can *change the output voltage* once he notices that it needs to be changed. Slew rate is specified in $\text{Volts}/\mu\text{sec}$ and comes into play when large swings in the output voltage are required. The slew rate of the TL082 is specified to be at least $8\,\text{V}/\mu\text{sec}$ and is typically $13\,\text{V}/\mu\text{sec}$. The effect of slew rate limiting of the TL082 is demonstrated in Figure 2-14.

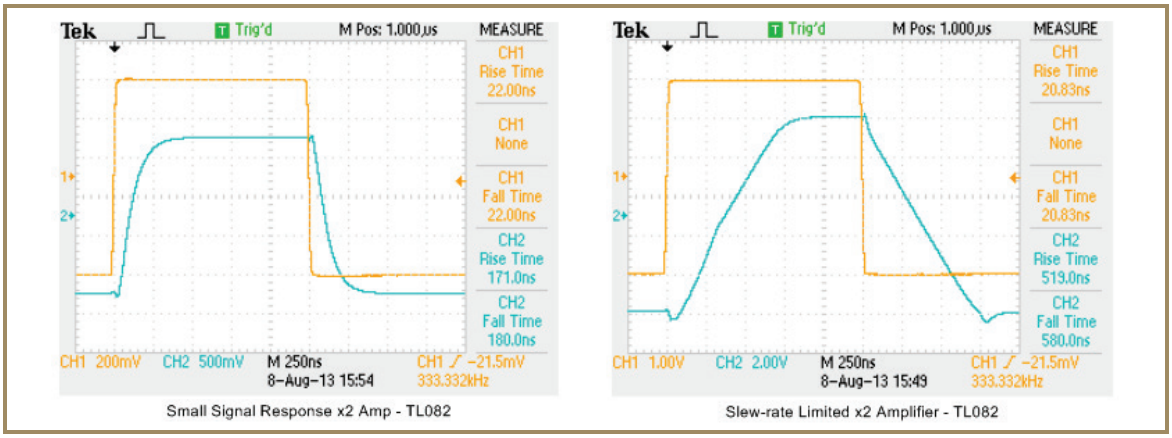


Figure 2-14: Oscilloscope recordings of the response of a gain = 2 noninverting TL082 op-amp amplifier to a square wave input. The output at left changes by 2V with each step and is limited only by the ×2 amplifier's bandwidth of about 1.5MHz. As you can see, it responds to an input step with an exponential relaxation; the time constant $\tau \approx 100$ns, as expected. For the right image the input is 5 times larger (5V v. 1V), and the output must change by 10V with every step. Note that the output changes at a nearly constant rate during these steps, and the output waveform shape is *qualitatively different* from the left-hand image. Thus, the right-hand output is *slew rate limited*: the slope is about 2V/150ns ≈ 13V/µs, as specified in the manufacturer's data sheet.

Unlike the effect of finite frequency response, slew rate limiting is an example of a *nonlinear effect* on the amplifier output. Linear effects aren't qualitatively affected by the size of the

input, but clearly, in this case, size *does* matter (the output waveforms in the right and left images in Figure 2-14 are *qualitatively* different).

Other imperfections of real operational amplifiers

We've discussed the two most important limitations of a real operational amplifier device: its finite (albeit quite large) DC gain (g_{DC}) and its finite bandwidth (f_{BW}). We've also discussed the limitation on large signal outputs caused by the amplifier's finite *slew rate*. There are several other limitations of real op-amps which can sometimes have significant effects on the performance of certain circuits — among these limitations are a device's input impedance, input bias current, input offset voltage, output impedance, output current limit, input noise voltage, and input noise current. As we explore more op-amp applications, we will discuss particular limitations whose presence can have a noticeable impact on our circuit design decisions.

THE INTEGRATOR AND DIFFERENTIATOR CIRCUITS

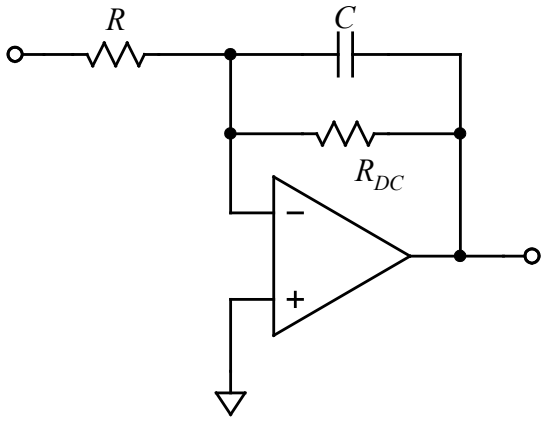


Figure 2-15: Integrator circuit. With an ideal op-amp and ignoring the action of resistor R_{DC}, then $G(\omega) = -1/(j\omega RC)$. In the time domain, $v_{in}(t) = -RC\,(d/dt)\,v_{out}(t)$, so v_{out} is proportional to the time integral of v_{in}. As explained in the text, the resistor R_{DC} limits the amplifier gain at DC so real op-amp imperfections won't cause the output to saturate.

The op-amp integrator circuit

Consider the inverting op-amp amplifier circuit in the figure above. For the moment assume that the op-amp is ideal and that $R_{DC} \to \infty$ (so that it doesn't affect the circuit). The inverting amplifier's gain is given by the impedance ratio Z_C/R, so:

$$G(\omega) = -1/(j\omega RC) = -\omega_0/(j\omega)$$

Thus $j\omega V_{out}(\omega) = -\omega_0 V_{in}(\omega)$, and we have an integrator:

Ideal, inverting integrator response (no R_{DC})

2.19

$$\frac{d}{dt}v_{out}(t) = -\frac{1}{RC}v_{in}(t)$$

$$G(f) = j\frac{f_0}{f};\quad f_0 = \frac{1}{2\pi RC}$$

So from the instant the circuit is put together and powered up, it will begin to integrate whatever voltage appears at its input and will continue to do so, indefinitely. This means that if the average input voltage is not precisely 0, then the average current through the capacitor will not be 0, and the op-amp will slowly (or maybe quickly) charge the capacitor until the output voltage has *saturated* (reached an output voltage limit).

Input offset voltage and bias current error sources

Even if the input is connected to ground (so $v_{in} \equiv 0$), slight imperfections in a real op-amp device will provide the equivalent of a small DC voltage input which will eventually cause output voltage saturation. The two most relevant op-amp imperfections you should consider are its *input offset voltage* and its *input bias current*. An ideal op-amp would produce no

change in its output whenever its two inputs exactly match, but tiny mismatches between the internal electronic components of the two input channels cause the op-amp to sense the equivalent of a small voltage difference *even when its two inputs are connected together* — the op-amp's *input offset voltage* specification gives an estimate of the magnitude of this sensed voltage error. For the TL082 this error is typically 3mV, but could be as high as 20mV. The op-amp will adjust its output to keep the voltage at the *–Input* equal to this offset error, since the *+Input* is grounded. Thus, even with $v_{in} \equiv 0$, the input offset voltage appears across the input resistor *R*, and the resulting DC current through it charges the capacitor *C*, ramping the output voltage v_{out} toward saturation.

The op-amp's input bias current is another potentially serious error source, especially if *C* is small. Real op-amps require some tiny DC *input bias current* to flow through each input in order for the internal electronic circuitry to operate properly. Let's estimate the impact of the *–Input* bias current (the bias current at the *+Input* will not affect this circuit, since that input is connected directly to ground). Ignoring the input offset error, the op-amp's *–Input* will be a virtual ground, so there is no voltage drop across the input resistor *R*. Thus the required input bias current can only flow through the feedback capacitor, slowly charging it until the output voltage saturates. The typical TL082 input bias current is only 30pA (flowing *into* each input), but it can reach 10nA at elevated ambient temperatures, doubling for about every 10°C increase. If *C* is 10nF, then v_{out} could ramp toward saturation at up to 1V/sec at high temperature (equation 2.8), but will typically change at a much slower rate (see Figure 2-16 for the proper way to model these op-amp error sources).

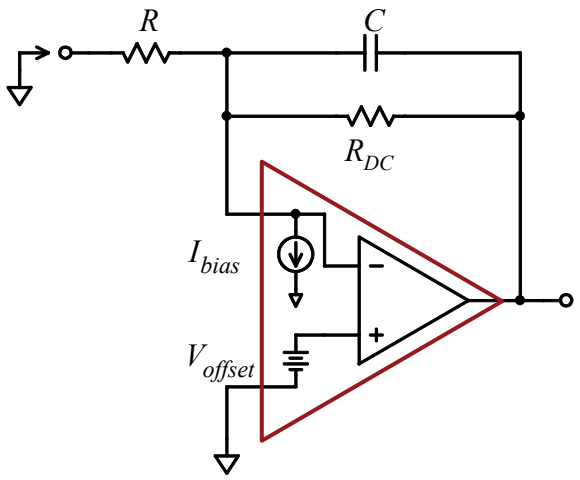


Figure 2-16: Effects of *input offset voltage* and *input bias current* errors on the Integrator circuit. These errors may be modeled as additional input sources to an ideal op-amp circuit, as shown. With the input grounded, the circuit is in a noninverting amplifier configuration as regards the offset voltage, V_{offset}, and it will be amplified by the noninverting circuit gain at DC. Without resistor R_{DC} the DC gain is the op-amp open-loop gain, which is extremely large! The input bias current, on the other hand, will flow through the parallel feedback components, and the resulting voltage drop across them will appear at the output. You should assume that each of the two error sources could be of either sign (+ or −), although I_{bias} has the direction shown for the TL082 op-amp.

The resistor R_{DC} is added to the circuit in Figure 2-15 in order to mitigate these DC error sources. For low frequencies (where the impedance $Z_C \to \infty$), the noninverting circuit gain is limited to $1 + R_{DC}/R$ (see Figure 2-16). The DC input offset voltage will be amplified by this gain to produce a resulting output offset voltage which you can limit to an acceptable level by properly choosing the ratio R_{DC}/R. The op-amp's input bias current will flow through R_{DC} (once the capacitor has reached its steady-state charge), so the output voltage error offset due to the bias current is $R_{DC}\, I_{bias}$.

OP-AMP INPUT BIAS CURRENTS REQUIRE A DC PATH

Because a real op-amp will always require some nonzero DC bias current to flow into (or out of) each of its two inputs, *there must be a path* for each of these DC currents to get to its input. If an op-amp input is left disconnected or is connected to only a capacitor terminal, then the op-amp's output voltage will soon *saturate* at one of its voltage limits, and your circuit will be useless.

Because the circuit is linear, the effects of these two error sources will add to produce an overall output voltage offset. The relative sign of the errors may be unknown for a particular op-amp, but there is a fifty-fifty chance that they will be in the same direction. You should assume that this will be the case for your circuit (Murphy's Law!), so choose a value for R_{DC} accordingly. The low frequency limit where R_{DC} begins to significantly affect the circuit's behavior as an integrator is given by the corner frequency $f_{DC} = 1/(2\pi R_{DC} C)$. The resulting circuit's response, including R_{DC}, is:

Inverting integrator response (ideal op-amp, Figure 2-15)

2.20
$$RC\frac{d}{dt}v_{out}(t) + \frac{R}{R_{DC}}v_{out}(t) = -v_{in}(t)$$

$$G(f) = j\frac{f_0}{f}\left(1 - j\frac{f_{DC}}{f}\right)^{-1}; \quad f_0 = \frac{1}{2\pi RC}, \quad f_{DC} = \frac{1}{2\pi R_{DC}C}$$

If $R_{DC} \gg R$, then the circuit still behaves as a good integrator (as in equation 2.19) for time intervals that aren't too long. In this case the equation relating $v_{in}(t)$ and $v_{out}(t)$ may be approximated as the integral below, if R_{DC} is large:

2.21 $\left(\text{for } R_{DC}\, C \gg t - t_0\right)$
$$v_{out}(t) = \frac{1}{RC}\int_{t_0}^{t} v_{in}(t)\,dt + v_{out}(t_0)$$

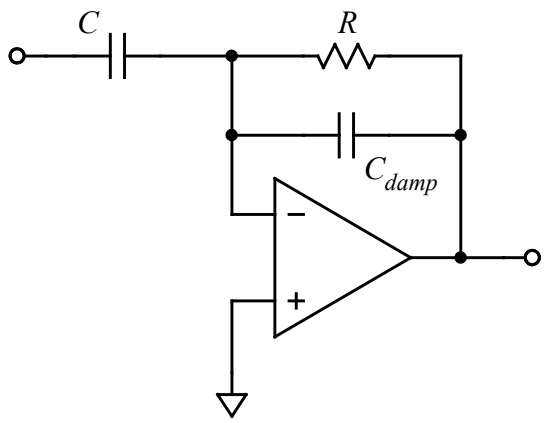


Figure 2-17: Differentiator circuit. With an ideal op-amp and ignoring the action of capacitor C_{damp}, then $G(\omega) = -j\omega RC$. In the time domain, $v_{out}(t) = -RC\,(d/dt)\,v_{in}(t)$. As explained in the text, the capacitor C_{damp} dampens response overshoot and "ringing" following a step input caused by a real op-amp's limited frequency response.

The op-amp differentiator

Another very useful application of the operational amplifier is to perform the inverse function of the integrator: take the time derivative of an input signal. A simple circuit to perform this function is shown in Figure 2-17; the positions of the integrator's R and C are simply exchanged (the added capacitor, C_{damp}, is needed to ensure that the circuit behaves itself, as did R_{DC} in Figure 2-15; it's effect will be discussed later).

Let's ignore C_{damp} for now (assume it is removed from the circuit). The ideal circuit gain is just $(-R/Z_C)$, so

$$G(\omega) = -j\omega RC = -j\omega/\omega_0$$

Thus $\omega_0 V_{out}(\omega) = -j\omega V_{in}(\omega)$, and, naturally, we now have a differentiator:

Ideal, inverting differentiator response (no C_{damp})

2.22

$$\frac{1}{RC}v_{out}(t) = -\frac{d}{dt}v_{in}(t)$$

$$G(f) = -j\frac{f}{f_0};\quad f_0 = \frac{1}{2\pi RC}$$

So the circuit's gain increases proportional to frequency, and its response to a rapid step input (such as the edge of a square wave signal) should be very large and very short. Unfortunately, the finite frequency response of our real op-amp limits this performance and adds a surprising extra effect to the circuit's response: *resonant oscillations*, or "ringing."

Ringing due to finite op-amp frequency response

Let's examine the effect of the op-amp's finite *gain-bandwidth product*, f_{BW}, on the circuit's behavior. To proceed we will use the inverting configuration gain expression (including

finite op-amp open-loop gain), equation 2.17, and the simplified op-amp open-loop gain model, equation 2.18, both repeated below:

2.17
$$\frac{1}{G_f(\omega)} = \frac{1}{G(\omega)} - \frac{1}{g(\omega)} + \frac{1}{G(\omega)\,g(\omega)}$$

2.18
$$\frac{1}{g(f)} = \frac{1}{g_{DC}} + j\frac{f}{f_{BW}}$$

Consider the first two terms on the right-hand side of equation 2.17 along with the differentiator gain equation 2.22 (now using frequency f rather than angular frequency ω):

$$\frac{1}{G(f)} - \frac{1}{g(f)} = j\frac{f_0}{f} - j\frac{f}{f_{BW}} - \frac{1}{g_{DC}}$$

At the frequency defined by the geometric mean of f_0 and f_{BW}, $f_{res} = \sqrt{f_0\, f_{BW}}$, the two imaginary parts cancel, and the gain is approximately:

2.23
$$\frac{1}{G_f(f_{res})} = \frac{1}{G(f_{res})g(f_{res})} - \frac{1}{g_{DC}} \approx \frac{-1}{\left|G(f_{res})\right|^2} \qquad \left(\text{for } g_{DC} \gg \left|G(f_{res})\right|^2\right)$$

So near frequency f_{res} the circuit gain jumps to approximately the square of what it is at frequencies further away, and the phase is very nearly real. At frequencies below f_{res}, where $|g(f)| \gg |G(f)|$, the response is close to the ideal, $G(f)$, with phase of −90°. Above f_{res}, the response is close to the (inverted) op-amp open-loop gain, $-g(f)$, with phase of +90°. Here is an extreme example of *gain peaking*, where very close to a single frequency (f_{res}) the gain may increase by orders of magnitude over what is expected (take a quick peek at Figure 2-19 on page 2-28 for an example).

In fact, this behavior is an example of *resonance*, where a narrow gain peak is observed as the response phase rapidly changes by 180°. We will study resonant behavior more thoroughly in a later experiment, but for now consider an example of this resonant effect on the transient response of the differentiator circuit. Let $R = 1\text{k}$ and $C = 1\mu\text{F}$ ($RC = 1\text{msec}$); then $f_0 \approx 160\text{Hz}$, and $f_{res} \approx 21\text{kHz}$ (for the TL082, with $f_{BW} = 3\text{MHz}$). The gain magnitude we expect at this frequency is $f_{res}/f_0 \approx 137$, but equation 2.17 tells us that *the actual gain will be much higher* (with the TL082 $g_{DC} = 2\times10^5$). As we will learn in a later experiment, the circuit exhibits resonant behavior with a *quality factor* $Q \gg 1$, implying that its output response to a sudden change in the input slope will be to overshoot and then "ring" at f_{res} for about Q cycles. This behavior is quite undesirable in our differentiator, so we must find some way to mitigate its effect (see the left-hand image in Figure 2-18 on page 2-28).

The answer is to add a capacitor in parallel with the feedback resistor R as shown in Figure 2-17, with $C/C_{damp} \approx \sqrt{f_{BW}/f_0}$. At high frequencies, the ideal gain function $G(f)$ will level off as the impedance of C_{damp} becomes small compared to R; the RC_{damp} corner

frequency is chosen to be close to f_{res}, which will limit Q to about 0.7. The result of this change on the circuit's transient behavior is captured in the right-hand image in Figure 2-18, showing the dramatic improvement in the differentiator's output. Figure 2-19 compares the measured frequency responses of a differentiator circuit with and without C_{damp}. As predicted, the undamped circuit shows a large gain peak; adding C_{damp} eliminates the gain peaking and its associated transient response ringing. See the section ***More about damping the differentiator*** for an alternate (often better) way to dampen differentiator ringing.

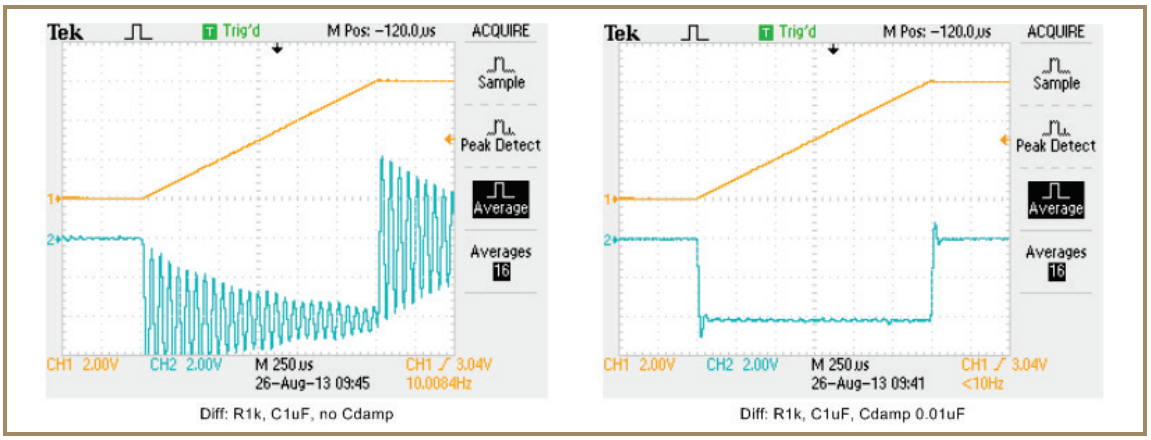


Figure 2-18: Undamped (left) and damped (right) differentiator responses to a sudden change in input slope. The input slope transitions between 0 and 4V/ms; the output should change between 0 and −4V (*RC* = 1ms). In the damped case C/C_{damp} = 100, resulting in a not quite *critically damped* response, with about a 10% *overshoot* and about 1 resonant cycle required to reach steady state.

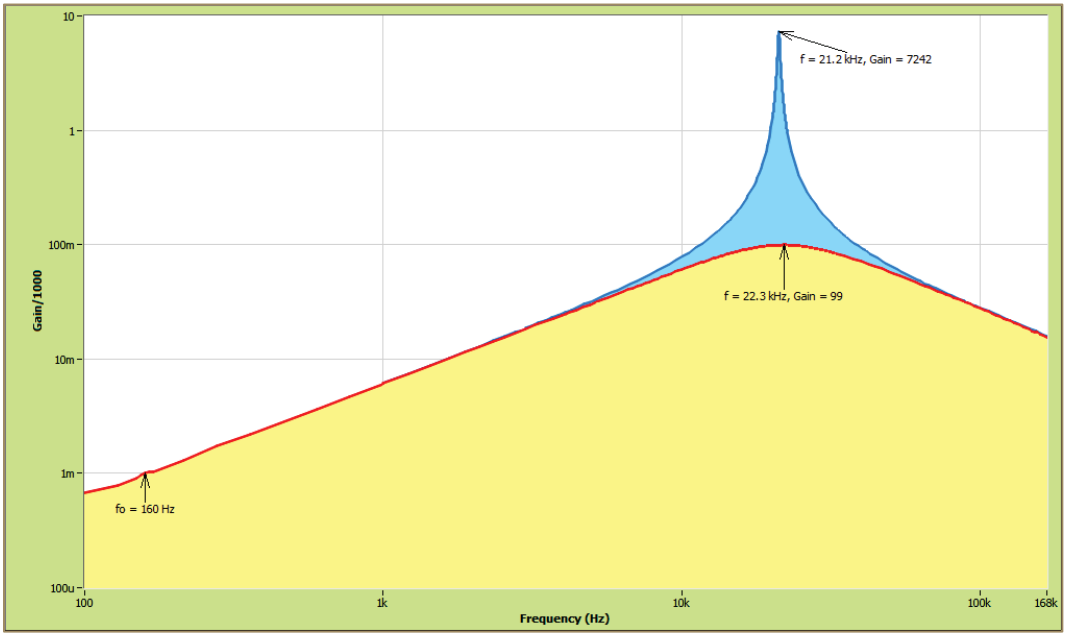


Figure 2-19: Comparison of the measured frequency responses of undamped and damped differentiator circuits (*RC* = 1ms, C/C_{damp} = 100). Data were taken using the *Frequency Response* data acquisition program. Adding the damping capacitor reduced the resonant *Q* from 55 to 0.75.

INSTRUMENT IMPEDANCE AND CABLE CAPACITANCE ISSUES

You want the instrumentation connected to your circuits to have negligible impact on the circuits' performance while still providing accurate information about them. Naturally, the impedance an instrument and its cabling present to your circuit is the major determiner of how it will affect the circuit's behavior, so we will devote some time to this important topic.

Input impedances of the oscilloscope and DAQ

The two input channels of a lab Tektronix oscilloscope maybe modeled as shown in the left-hand schematic in Figure 2-20. As shown, each input BNC shell is connected directly to *earth ground*; each input signal conductor is connected to ground via a parallel combination of a 1MΩ resistor and a 20pF capacitor. Changing the *channel coupling* to AC inserts an additional capacitor in series with the input signal to create an *RC* high-pass filter and block the DC component of the input signal; the *RC* corner frequency is a little less than 10Hz, and thus $\tau_{RC} \approx 20\text{ms}$.

> The oscilloscope forces a connection made to any of its BNC shells to *Earth Ground* potential, so you must always keep this in mind when making connections to your circuit. All of its BNC connector shells are internally connected together.

The two main Computer Data Acquisition (DAQ) analog channels, AI0 and AI1, on the other hand, are *fully differential* and have a relatively high input impedance. As shown in the right-

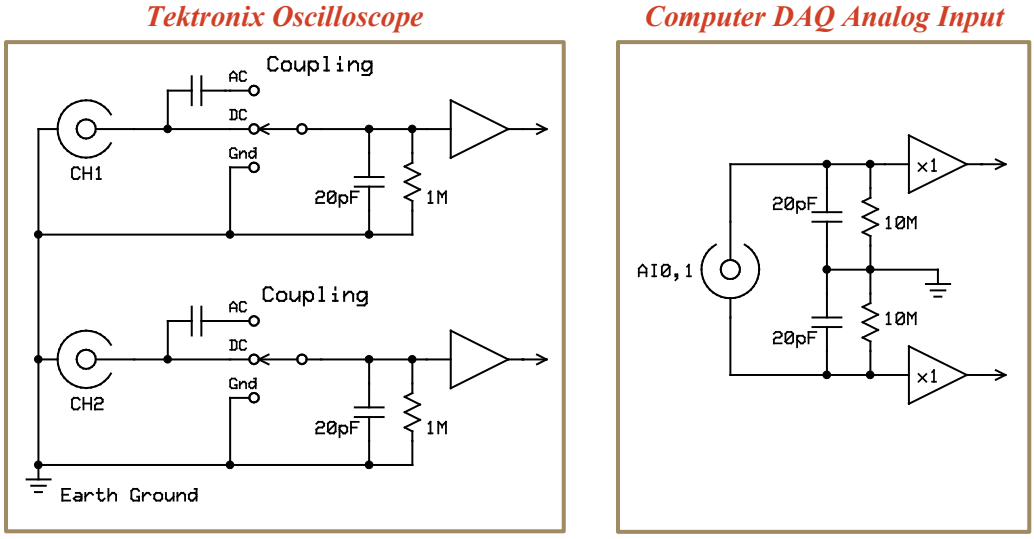


Figure 2-20: Input circuit models for the lab Tektronix oscilloscope and the DAQ interface. Each oscilloscope input channel has an impedance equivalent to 1MΩ in parallel with a 20pF capacitor, as shown. The BNC input connector shells are connected together and directly to *Earth Ground*, as shown. The inputs' *Coupling* settings are independently selected using each channel's input menu. The computer DAQ analog inputs, on the other hand, are *fully differential*, so each input's BNC conductors are isolated from each other and from Earth Ground. Each DAQ conductor has an input impedance of 10MΩ in parallel with a 20pF capacitor to Earth Ground.

hand schematic of Figure 2-20, each conductor of the BNC connector has an input impedance of 10MΩ in parallel with 20pF to earth ground, and none of them are directly connected together. What this means is that you may connect one of these DAQ channels across any component in your circuit with minimal effect (at least at low frequencies).

BNC coaxial cable capacitance

Connecting the oscilloscope or DAQ inputs to the circuit requires the use of some sort of cabling, and a lab coaxial cable used to make a connection can have a significant effect on the behavior of your circuit. A coaxial cable is a *transmission line*, which, in general, may not be treated as a simple, lumped circuit element. Assume a single sinusoidal electromagnetic wave propagates along the cable. At any point in the cable, the ratio of the voltage phasor (between its center conductor and shield) and the current phasor (for current flowing in the center conductor) has a fixed value defined by the cable's *characteristic impedance*. For an ideal, lossless cable this impedance is real (i.e. a pure resistance). The characteristic impedance of nearly all standard laboratory coaxial cables is 50Ω (home analog video cables, on the other hand, are 75Ω); this impedance is determined by the ratio of the shield and center conductor radii along with the dielectric constant of the insulator separating them.

If a 50Ω BNC cable is attached to a device with a 50Ω terminal impedance, then the other end of the cable will behave as a perfect, lumped element with an impedance of 50Ω (i.e. a resistor) when connected to your circuit. Thus, since a lab signal generator has an output impedance of 50Ω, the end of the BNC cable used to connect it to your circuit will behave as a perfect, lumped voltage source with a 50Ω impedance, *no matter how long the BNC cable is between the generator and your circuit*. This only works if the terminating device's impedance matches the cable, as in the case of the signal generator.

If the impedance of the terminating device is much larger than the cable's characteristic impedance, however, things aren't so simple (the input impedances of the oscilloscope and DAQ are good examples). In this case, if the length of the cable is very short compared to the wavelengths of the signals propagating along it (much shorter than ¼ wavelength), then *the cable acts as an additional capacitance in parallel with the impedance of the terminating device* when you connect it to you your circuit. It should be easy to see how this comes about, since the cable consists of two closely-spaced conductors separated by a dielectric. The effective capacitance is proportional to the cable length; for the lab cables it is about $100\,\mathrm{pF/meter}$. An added $100\,\mathrm{pF}$ can have a significant impact on your circuit's performance; at 100kHz the magnitude of its impedance is only about 16kΩ, and its presence can cause a noticeable shift in the phase of a measured voltage or the corner frequency of a filter.

Connecting a BNC coaxial cable between a circuit and a device with a 50Ω impedance acts as if the 50Ω is connected directly into the circuit (by design!). Connecting a BNC cable between a circuit and a high impedance instrument adds a capacitance of 100pF/meter to the circuit in parallel with the device's input impedance.

So how can you mitigate the effect of this additional capacitance caused by adding a coaxial cable? One simple way is to *buffer* the cable connection using a voltage follower, as shown in Figure 2-21. The high input impedance of the voltage follower isolates the connection from your circuit, and a TL082 op-amp can successfully drive the capacitance of coaxial cables with lengths of several meters, especially if you add an output load resistor to the follower circuit as shown.

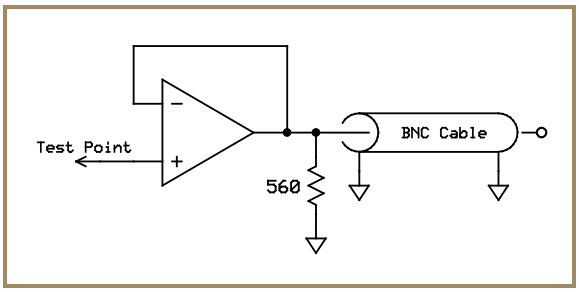


Figure 2-21: Using a voltage follower to *buffer* a coaxial cable connection. The high impedance of the follower's input has minimal effect on the circuit being monitored (the "test point"), and its low output impedance can drive the capacitance of a fairly long 50Ω coax connected to a high impedance instrument. The 560Ω load resistor on the follower reduces overshoot and ringing in its output when driving a long cable's capacitive load.

The 10x Oscilloscope probe

A straightforward, passive technique is usually employed to mitigate the effects of cable capacitance on a connection to an oscilloscope input: use of a so-called *10x probe*, Figure 2-22. To compensate for the coaxial cable capacitance between the probe tip and the oscilloscope input, the probe tip contains a 9MΩ resistor in parallel with a small, slightly adjustable capacitor of approximately 15pF. These components form a voltage divider with the oscilloscope's 1MΩ input resistance and the cable's capacitance, giving a 10:1 signal

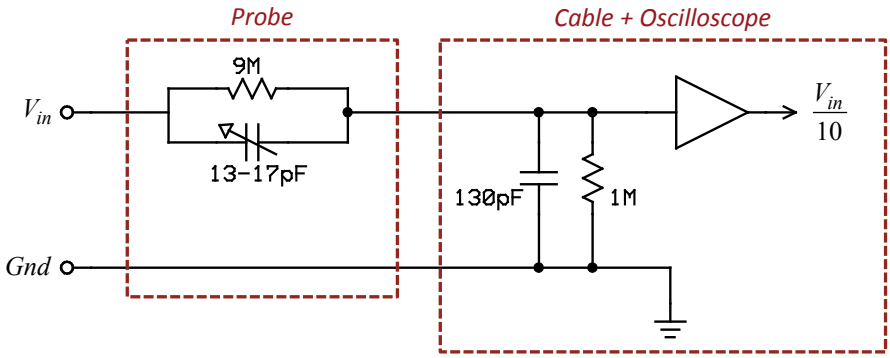


Figure 2-22: Schematic of a 10x oscilloscope probe connected to an oscilloscope input channel. The adjustable probe compensation capacitor keeps the voltage divider ratio at 10:1 at all frequencies, even in the presence of the large input + coaxial cable capacitance (130pF). The equivalent input impedance at the probe tip is 10MΩ in parallel with about 13pF.

reduction to the oscilloscope at DC while increasing the input resistance at the probe tip to 10MΩ. By setting the value of the probe's *compensation capacitor* to exactly 1/9 of the cable + oscilloscope capacitance, this 10:1 voltage division will be maintained at all input frequencies, *and the equivalent parallel input capacitance at the probe tip is reduced by a factor of 10 to about 13pF* (since the small capacitor and the cable capacitance are effectively in series). The section of the text starting on page 2-42 shows how to properly adjust the probe compensation capacitor so that it doesn't distort the input signal presented to the oscilloscope.

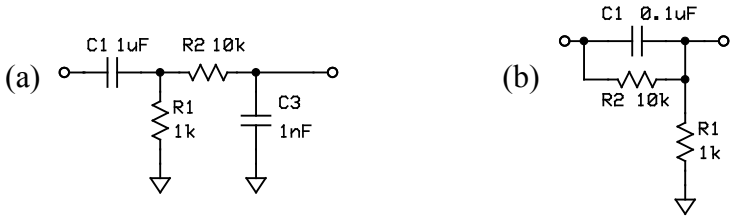
PRELAB EXERCISES

1. Show that the total equivalent capacitance of two or more ideal capacitors connected in parallel is equal to the sum of their individual capacitance values.

2. Derive equations 2.12 for the *RC* high-pass and low-pass filter frequency responses, $G(\omega)$, from the equations just above them on page 2-12.

3. Consider the AC coupled amplifier in the example shown in Figure 2-8 on page 2-14. What should be the value of *R* if the amplifier circuit's input impedance should be 100 kΩ? The audio frequency range is nominally taken to be 20Hz–20kHz. What should be the minimum value of *C* if you want no more than 3dB attenuation (reduction in gain) of a signal in this frequency range?

4. Sketch Bode plots of the gain for the following filters, both magnitude and phase v. frequency (input at left, output at right for each circuit). Use appropriate log scales for the plots and sketch the asymptotes (which should be straight lines). Calculate the corner frequencies (there are two for each circuit). Label each asymptote with its $G(f)$ function, including its phase (j for +90°, $-j$ for −90°, etc., as in Figure 2-7). Messy algebra shouldn't be required to complete this exercise.



5. You cascade two noninverting amplifier stages, each with a gain of 11 (cascade: connect the output of one to the input of the next). What is the resulting gain of the combined amplifiers? If both amplifiers use the TL082 op-amp ($f_{BW} = 3\,\text{MHz}$), what is the −3dB corner frequency for each amplifier stage? What will the gain of the amplifier combination be at this frequency? Sketch a schematic of this circuit: label feedback resistors with appropriate choices for their values; label the input and output terminals.

6. Sketch a Bode plot (magnitude and phase response) of the integrator circuit (Figure 2-15 on page 2-23), assuming that $R_{DC}/R = 100$ and $f_0 = 16\,\text{kHz}$ (equation 2.20). What is the circuit's gain at frequency f_0? at 1.6 kHz? What is the value of f_{DC}? What should be the value of *R* if $C = 0.01\,\mu\text{F}$ (round off *R* to only two significant figures)? What is the expected output offset voltage magnitude for a specified op-amp input offset voltage error of 3mV? Given the *R* value you've calculated and an op-amp input bias current of 30pA, what would be the input bias current's contribution to the output offset voltage?

LAB PROCEDURE

Overview

During lab you will investigate the frequency responses (magnitude and phase) of various simple filters and op-amp circuits. You will add another tool to your bag of tricks, the lab's *Frequency Response* program, which will take control of your lab signal generator and use the computer data acquisition system (DAQ) to generate Bode plots of the transfer functions of amplifiers and filters. In effect, this generator+DAQ system emulates the function of a *Network Analyzer*, a sophisticated laboratory instrument for measuring the transfer function and terminal impedances of a 2-port network as a function of frequency. Your lab TA or the course instructor will give a quick demo of the *Frequency Response* program at the start of your lab session.

> The Frequency Response program can save transfer function data as a text file which can be manipulated and plotted using various software (including spreadsheet programs). You can also capture images of plots the program generates and print them directly from the program.

Detailed procedures

Gain-bandwidth product limitations of op-amp amplifiers

Assemble a ×11 noninverting amplifier and connect the two DAQ analog inputs to the amplifier's input and output, as well as the oscilloscope (refer back to Figure 1-30 on page 1-34 of Experiment 1; connect the BNCs shown to the DAQ rather than the oscilloscope). Connect the signal generator's SYNC output to one of the DAQ digital PFI inputs (this digital signal from the generator will be used to trigger DAQ data acquisition events).

> The computer DAQ system requires a digital signal synchronized with the circuit analog signals to trigger its data acquisitions. Without this digital signal, the computer will never acquire data. There are several ways to generate this digital trigger signal — examine the DAQ interface description on the interface box or talk to the lab instructor about this issue.

Start the *Frequency Response* application. With the generator adjusted to give about a 0.5V output from the amplifier (at 1 kHz), use the program's manual mode to find the amplifier's −3dB bandwidth.

What should be the phase shift of the amplifier's transfer function at this frequency? How does this frequency compare with what you expect from the TL082 op-amp's *gain-bandwidth product*?

Input a square wave and use the oscilloscope to examine the steps in the output.

> Is the output step an exponential relaxation as in the right-hand image of Figure 2-9? What is the time constant? Does it agree with your estimated amplifier bandwidth? Increase the input square wave amplitude until you can recognize op-amp slew rate limiting in the output waveform. What is the observed slew rate? Save screen captures of your data.

Cascade another ×11 noninverting amplifier to give a 2-stage amplifier. Keep the output of the second stage to no more than 5 or 6 Volts so as to avoid the op-amp slew rate limit.

Use the *Frequency Response* program (back to sine wave output from the signal generator!) to sweep and plot the 2-stage amplifier's frequency response from about 1 kHz to 1 MHz.

> Consider your answers to Prelab exercise 5; does this 2-stage amplifier have the bandwidth you expect? Print copies of the gain magnitude and phase plots. Make sure to use Log scales appropriately! Save sweep data set to a file that you can then reload into the program to compare with another set of data.

> Make sure to select the **Save All** radio button in the data save dialog box. Otherwise, you won't be able to reload the saved data set into the *Frequency Response* program.

Now construct a single-stage noninverting amplifier with a gain of about 120. Using a 10k resistor for R_f, what should be the value for R_i? (note that 82Ω is a standard resistor value) Your TA will show you how to connect an odd resistor into the breadboard circuit. Sweep and plot this amplifier's frequency response from about 1 kHz to 1 MHz.

> How does the −3dB bandwidth compare to what you would expect for a single-stage amplifier? How does it compare to the 2-stage amplifier's bandwidth? Plot both frequency response curves on a single graph. With a square wave input, estimate the time constant in the output's step response.

Adding AC coupling to an amplifier input

Staying with the single-stage noninverting amplifier you've built, add a simple *RC* high-pass filter to its input using the resistors and capacitors integrated into the trainer breadboard, attached to the op-amp's *+Input* (Figure 2-8). Use a 10kΩ resistor and select a capacitor which will give a corner frequency of about 160Hz.

> Use the frequency response program to determine the corner frequency. Adjust the signal generator to add a significant DC offset voltage to a 1kHz input signal. Does the offset affect the amplifier output? Remember to reset the offset voltage to 0 when you've finished.

Integrator and differentiator

Assemble an integrator circuit (Figure 2-15) with $RC \sim 0.1\text{msec}$. At what frequency do you expect $|G| = 1$? Select a value for $R_{DC} \approx 100R$. What should be the circuit DC gain? If the op-amp input offset error is 3mV, what should be the DC output voltage with the input = 0?

With the input grounded, what is the measured DC output voltage? Watch the output on the oscilloscope as you disconnect one end of R_{DC} from the circuit. About how long does it take for the output voltage to saturate?

Input a square wave at about 1kHz and about 0.5Vpp. Compare the output and input waveforms. Does the relationship match the first of equations 2.19? The capacitors on the breadboard have a tolerance of about ±10%.

Sweep the circuit's frequency response between about 50Hz and 10kHz. What is your measured frequency f_0?

Assemble a differentiator (Figure 2-17) with $RC \sim 10\,\text{msec}$, and use a voltage follower to drive its input. Select a value for $C_{damp} \approx 0.01\,C$.

With one end of C_{damp} disconnected, what does the circuit's response to a 50Hz triangle wave input (ramp with symmetry = 50%) look like? Now connect C_{damp}; does the output waveform ringing disappear? Is the output waveform shape what you expect? Is its amplitude correctly predicted by the first of equations 2.22?

Drive the differentiator input directly from the signal generator (it has a 50Ω output impedance) and disconnect C_{damp}. Does the output ring in this case? What's going on? See the section ***More about damping the differentiator*** for an explanation.

Additional, self-directed investigations

Build one or more circuits from the **MORE CIRCUIT IDEAS** section or try some of your own design. Investigate their frequency responses and their transient behavior with a square-wave input. Look back at Experiment 1 to see if there are any other circuits from that chapter you would like to investigate.

Lab results write-up

Include a sketch of the schematic with component values for each circuit you investigate, along with appropriate oscilloscope screen shots and Bode plots. Make sure you've answered each of the questions posed in the ***Detailed procedures*** section.

MORE CIRCUIT IDEAS

Phase shifter (all-pass filter)

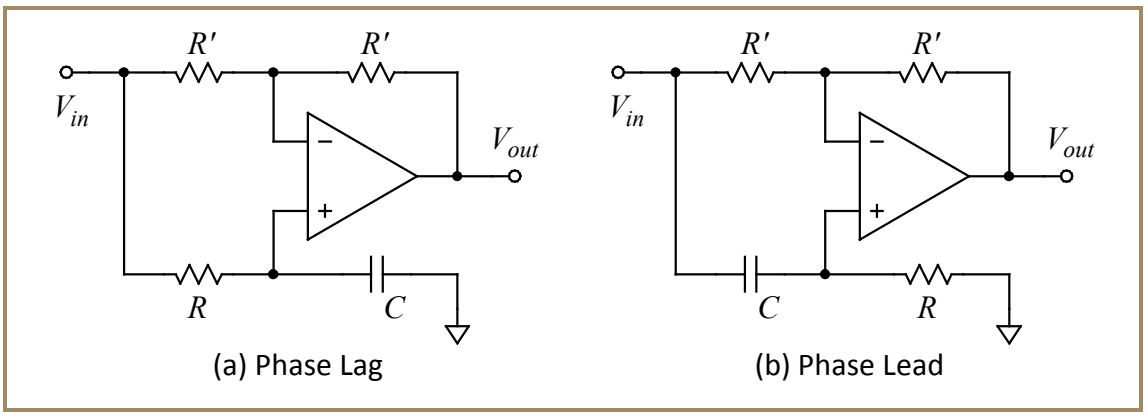


Figure 2-23: Unity-gain phase shifters (all-pass filters). Although $|V_{out}/V_{in}|$ = 1 at all frequencies, the phase of the output decreases through 180° as frequency goes from low to high. In the circuit (a) the phase at low frequency is 0°, but lags through −90° at f =1/(2πRC) toward −180° as frequency increases. In circuit (b) low frequency phase is +180°, +90° at f =1/(2πRC), and 0° at high frequency.

These interesting little circuits show that the rule of thumb relating phase shift and gain slope (i.e. Figure 2-7 on page 2-13) doesn't necessarily apply when we include an op-amp! The heart of each circuit is the combination inverting-noninverting amplifier discussed in Experiment 1 (Figure 1-20 on page 1-22). Using the formula in the figure's caption you should be able to show that the transfer functions of the two circuits are:

$$\frac{V_{out}}{V_{in}} = \frac{\pm(1 - j\omega RC)}{1 + j\omega RC} \qquad +\text{for circuit (a);}\ -\text{for circuit (b)}$$

The magnitudes of the numerator and denominator are equal, so $|G| = 1$ at all frequencies (assuming an ideal op-amp, of course). The phases of the numerator and denominator, however, each shift through 90° in opposite directions as frequency changes from low to high. In both circuits the resulting phase *decreases* (lags ever more) as frequency goes from low to high, changing by a total of 180°. This implies that in effect the circuits insert a *time delay* between input and output, $T = -d\phi/d\omega$. At midrange frequencies (near $1/RC$), this delay is equal to RC.

The circuits have interesting transient responses to a step input; input a low frequency ($\omega \ll 1/RC$) square wave and see if you can explain the shape of the output waveform.

More about damping the differentiator

Look at the differentiator circuit (Figure 2-17) again. Let's be sure you understand how the damping capacitor C_{damp} gets rid of the gain peaking. At high frequencies ($\omega \gg 1/RC_{damp}$) the impedance $Z_f \to Z_{C_{damp}}$; therefore the ideal inverting amplifier gain approaches $Z_f \to Z_{C_{damp}} / Z_C = C/C_{damp}$, independent of frequency. As long as this situation of

constant ideal amplifier gain is obtained before the ideal gain Bode curve intersects the op-amp's open-loop gain curve g_ω, then the differentiator won't ring.

There is a significant drawback to our differentiator: the input impedance of the differentiator circuit is Z_C, which becomes very small at high frequencies, potentially drawing a lot of current from the input source. If the source has a significant output resistance, then its voltage will sag at high frequencies, and so will the differentiator's output voltage. We should therefore always drive the differentiator in Figure 2-17 using a voltage follower to ensure that we have a nearly ideal voltage source supplying the differentiator's input — the 50Ω output impedance of the signal generator can have a significant effect on the circuit's high frequency performance.

This observation leads to different method to effectively damp the differentiator ringing: use a small series resistor in the input circuit rather than a parallel capacitor in the feedback, as shown at right. At high frequencies ($\omega \gg 1/R_{damp}C$) input impedance approaches R_{damp} and the ideal gain approaches R/R_{damp}, again damping the differentiator's ringing if you choose R_{damp} properly: $R/R_{damp} \approx \sqrt{f_{BW}/f_0}$. The signal generator's 50Ω output impedance may be sufficiently large to serve as an effective R_{damp}, so driving a differentiator directly from the signal generator may show little ringing, even without explicitly adding R_{damp} or C_{damp} to your circuit.

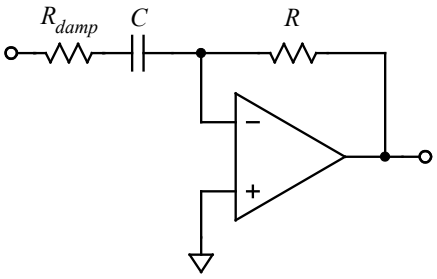


Figure 2-24: Using an input resistor to damp the differentiator.

PID operator

You can generate a general, *proportional+integral +differential* (*PID*) operator by generating a weighted sum of the outputs of amplifier, integrator, and differentiator circuits, as shown at right. Each term is generated by an appropriate circuit, and these outputs are summed. The choice of resistors in the inverting summer determines the weight of each term and the overall gain of the operator. An important part of an *analog computer* system, this circuit is also the heart of the ubiquitous *PID control loop* used for everything from industrial process control to missile guidance to home air conditioning temperature regulation (most modern systems use a digital controller implementation, however).

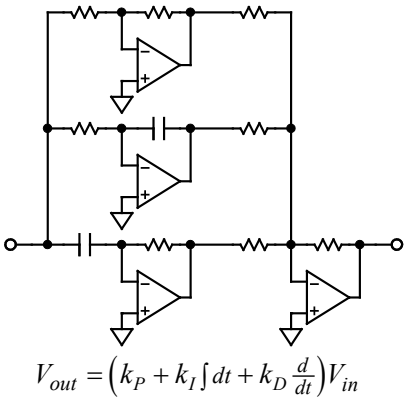


Figure 2-25: A circuit to perform a general *PID* operation.

High input impedance, high gain, inverting amplifier

Sometimes you need a high gain inverting amplifier stage, but you need an input impedance of, say, 100kΩ. Thus you would need an unreasonably large value for Z_f. The traditional way to get around this problem is to cascade a couple of voltage divider stages in the feedback loop, as shown in Figure 2-26.

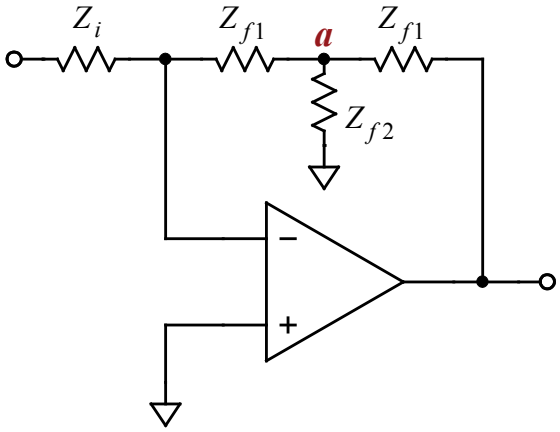


Figure 2-26: A high gain inverting stage which doesn't require a small value for the input impedance, Z_i. The current through Z_i also flows through the left-hand Z_{f1}, so the voltage at node a is $-(Z_{f1}/Z_i)\,V_{in}$. But the voltage at this node must be a weighted sum of the voltages at other ends of the impedances joined to it, two of which are 0, so also solving for V_a in terms of V_{out} gives the gain equation 2.24.

To derive the ideal closed-loop gain, $G(f)$, for this circuit is not too difficult if you consider the voltage at the node ***a*** in the figure. Since no current flows into the op-amp *–Input*, which is also a virtual ground, we know that the input current $I_{in} = V_{in}/Z_i$ also flows through the left-hand feedback resistor into node ***a***. Thus $V_a = -Z_{f1} I_{in} = -(Z_{f1}/Z_i)V_{in}$. But we can derive another expression for V_a from V_{out} using the generalized voltage divider expression for Experiment 1, equation 1.10 on page 1-25: $V_a = (V_{out}/Z_{f1})/(2/Z_{f1} + 1/Z_{f2})$. Equating these two expressions and solving for $G = V_{out}/V_{in}$:

2.24

$$G = -\frac{Z_{f1}}{Z_i}\left(2 + \frac{Z_{f1}}{Z_{f2}}\right)$$

Notice that we've chosen to set the two horizontal feedback impedances equal, which considerably simplifies the gain expression 2.24. Now we can pick a relatively large value for Z_i, so that we can keep the input impedance (which is, of course, also Z_i) reasonable. The feedback resistors need not be very large, because a large gain may be realized by making Z_{f2} small.

Combinations of "Tee" networks like the feedback circuit in Figure 2-26 are very useful in the design of filters.

ADDITIONAL INFORMATION ABOUT THE TEXT IDEAS AND CIRCUITS

This section expands on some of the material presented earlier. It is better skipped during a first reading, but you may want to go over it after you thoroughly understand the concepts discussed in the first several sections.

Fourier and Laplace transforms

General, real-valued functions of time (not necessarily sinusoids) may be represented as complex-valued functions of frequency by using the *Fourier transform* and its inverse (the pair given here is consistent with our sinusoid representation, equations 2.3, although more symmetric forms of the transform and its inverse are more common):

2.25

$$y(t) = \mathrm{Re}\left[\int_{0+}^{\infty} Y(\omega)\, e^{j\omega t}\, d\omega\right] + Y(0)$$

$$Y(\omega \neq 0) = \frac{1}{\pi}\int_{-\infty}^{\infty} y(t)\, e^{-j\omega t}\, dt; \quad Y(0) \equiv \overline{y(t)} = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} y(t)\, dt$$

So $y(t)$ and $Y(\omega)$ in equations 2.25 are called a *Fourier transform pair*. The first equation converts a phasor function $Y(\omega)$ to the actual, time-varying voltage or current $y(t)$ it represents. The second equation tells how to construct $Y(\omega)$ from a given $y(t)$ (note that $Y(0)$ is the *mean value* of $y(t)$, also called its *DC component*). Mathematicians have extensively studied just what sorts of functions can be represented this way, and what exactly is meant by the " =" in (2.25); we won't worry about such issues here.

Note that the second equation in (2.25) will give a valid result for $Y(\omega)$ even if $\omega < 0$ (negative frequencies) as well as for "normal" positive frequencies. What does it mean if $\omega < 0$, you may ask. There is nothing mysterious here: $\omega < 0$ just means that the phase of the complex number $Z(t) = Y(\omega)\, e^{j\omega t}$ decreases with increasing time: the arrow representing Z in the figure on page 2-6 rotates *clockwise* as t increases, rather than counterclockwise. Since our functions $y(t)$ are real-valued, it is easy to show that $Y(-\omega) \equiv Y^*(\omega)$, so we don't get any additional information about $y(t)$ from $Y(\omega)$ by including $\omega < 0$ in Y's domain of definition. We can just integrate Y over positive ω to get $y(t)$ as in the first expression in (2.25) (note further that the 0+ lower limit of the integral just means that $\omega = 0$ is not included in its domain).

The transformations (2.25) are linear in $y(t)$ and, so a linear function applied to $y(t)$ or $Y(\omega)$ will transform to some corresponding linear expression for $Y(\omega)$ or $y(t)$, respectively (with the caveat that multiplying $Y(\omega)$ by a *complex* scale factor also introduces a phase shift in that frequency component of $y(t)$). In particular, our very important result (2.6) that $dy(t)/dt \Leftrightarrow j\omega Y(\omega)$ remains valid. As you have already seen from the examples presented

earlier in this text, this concept is very useful when you analyze circuits (i.e., networks) with the tools we've developed: series and parallel impedances, the voltage divider, loop and node equations, the concepts of input and output resistance, transfer functions, etc. In particular, we have used the complex impedances and frequency responses of a circuit's various components to calculate a frequency-domain expression for its transfer function, as in equations (2.12), (2.16), and (2.17). In each case, the algebraic expression for the frequency-domain transfer function had the following form:

2.26
$$Y_{out}(\omega) = G(j\omega) \times Y_{in}(\omega) \qquad G(j\omega) = \frac{P(j\omega)}{Q(j\omega)}$$

where P and Q are polynomial functions of their arguments. Note that this relationship implies that $Y_{out}(\omega)$ is just proportional to $Y_{in}(\omega)$ with (generally) complex proportionality factor $G(j\omega)$; this result is a direct consequence of considering only *linear* circuits.

Note that we've written the transfer (gain) function G as a function of $j\omega$ rather than just ω in (2.26), which is suggestive of what we are going to do next: define the new variable $s = j\omega$, so that the transfer function becomes $G(s) = P(s)/Q(s)$. Using s as the independent variable, it must be the case that the polynomials $P(s)$ and $Q(s)$ each have only real coefficients (if any common factor of j is divided out of both), because our time-domain input and response functions $y_{in}(t)$ and $y_{out}(t)$ must both be real-valued functions of time (you may check that this is indeed the case for the transfer functions derived so far, including (2.12), (2.16), and (2.17)).

The zeroes of the two polynomials P and Q (i.e. the solutions s_k and s_m of $P(s_k) = 0$ and $Q(s_m) = 0$, respectively) are called the *zeroes* and *poles* of the transfer function G. These values provide important information about the resonant frequencies, transient behavior, and stability of the dynamical (time-varying) system being modeled. This sort of *pole-zero analysis* plays a prominent role in analog electrical and control system engineering, and naturally leads to the intense mathematical study of linear dynamical systems using an alternative to our Fourier analysis of equations (2.25): the *Laplace transform*,

$$Y(s) \;=\; \mathfrak{L}\{y(t)\} \;\equiv\; \int_0^\infty y(t)\, e^{-st}\, dt$$

where now s may take on any complex value (usually expressed as $s = \sigma + j\omega$), and it is assumed that $y(t < 0) \equiv 0$ (so the input signal to the system is "turned on" at time $t = 0$).

Oscilloscope 10x probe compensation adjustment

The oscilloscope provides a 1kHz square-wave output so that the 10x probe's compensation capacitor may be properly adjusted using a special tool, as shown in Figure 2-27.

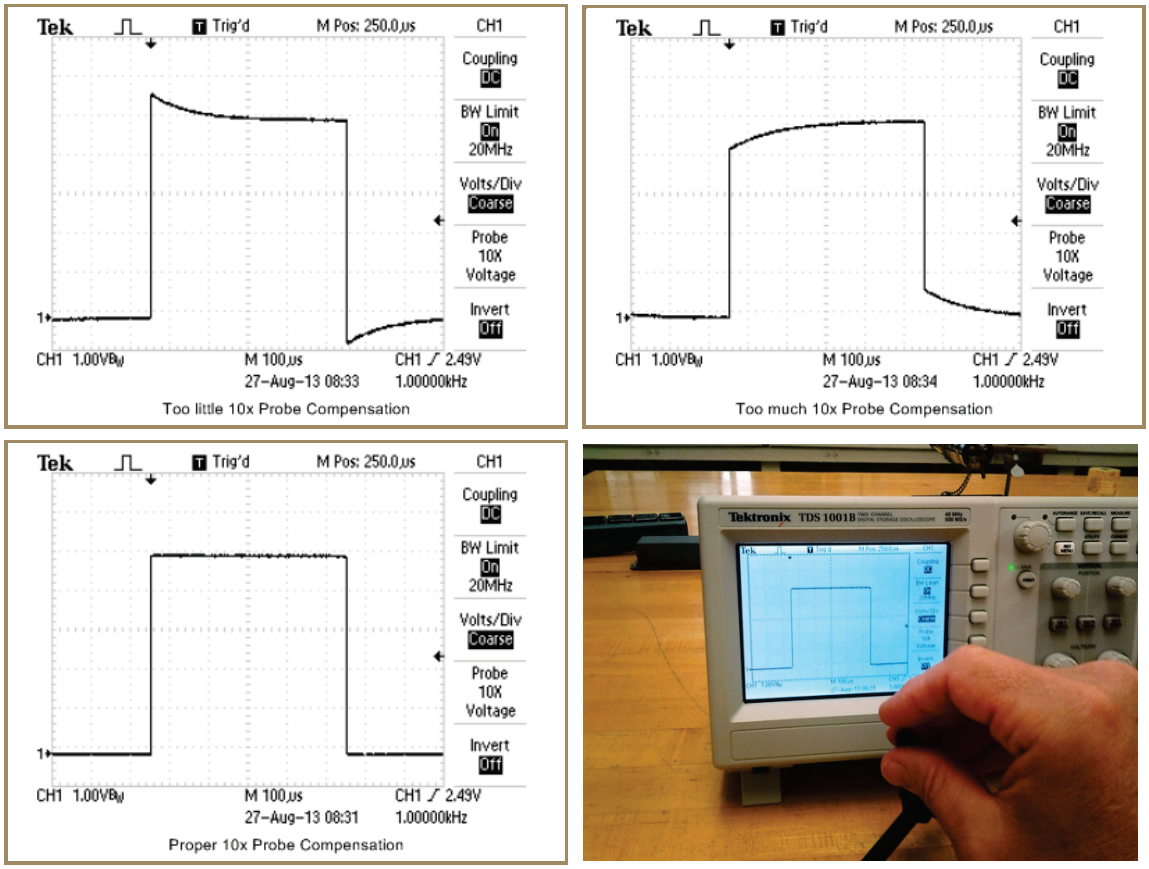


Figure 2-27: Properly adjusting the 10x probe compensation. The probe and its ground are clipped to the 1kHz signal output provided on the oscilloscope front panel. Using the special tool provided specifically for this purpose, the user carefully adjusts the probe's internal capacitor until the displayed waveform is as flat as possible, as shown in the lower left image. The upper images show outputs when the probe is not properly compensated.

> ***Caution***
>
> The compensation adjustment on the probe is easily damaged, so be careful with it! Do not use the adjustment tool for any other purpose except probe compensation adjustments.

Experiment 3

Nonlinear circuits: diodes and analog multipliers

Copyright © Frank Rice 2013
Pasadena, CA, USA
All rights reserved.

Experiment 3

Nonlinear circuits: diodes and analog multipliers

So far the analog circuits we have considered have all been linear, so that the output has been given by a sum of terms, each term strictly proportional to only one source value (see the section in Experiment 1: ***Linear circuits and superposition*** on page 1-20, and, in particular, equation 1.7 on 1-21). It's now time to extend our design toolbox to include *nonlinear* elements and networks, ones for which the *principle of linear superposition* no longer holds.

The simplest nonlinear circuit component is the *semiconductor diode*, which we consider first. This two-terminal element behaves in a most asymmetric manner: its resistance is very low for currents of more than a few milliamps flowing in one direction through the device, but it has an enormously high resistance to current flow in the opposite direction. This element is very useful for constructing absolute value, peak detection, overvoltage protection, and more general nonlinear resistance circuits. The diode's current-voltage relationship is actually exponential, so it is also useful for building exponential and logarithmic response amplifiers. Its characteristics are strongly temperature-dependent, so a diode also makes an excellent, accurate temperature sensor. Of course, some types of diodes also can emit and detects light (LEDs, laser diodes, and photodiodes), so the variety of applications of the seemingly simple semiconductor diode is nearly endless.

The other nonlinear element we'll consider in this experiment is the much more sophisticated *analog multiplier* integrated circuit, whose output is proportional to the *product* of two input voltages (making its transfer function a so-called *bilinear form* of its two inputs). This flexible device may be used in circuits which not only multiply but also divide, raise to powers, and take roots. You may use it to build voltage-controlled filters and variable-gain amplifiers, modulators and demodulators, phase detectors, automatic gain control and signal compression circuits, and oscillators — as well as its obvious applications in general *analog computing* circuits.

APPLICATIONS OF THE SEMICONDUCTOR DIODE

Diode rectifier circuits

A semiconductor *diode* is a two-terminal element which acts as a "one-way valve" for electrical current (i.e., it is a *rectifier*). The most common type of diode is made from a silicon crystal divided into two layers with different impurity atoms mixed into the silicon. The resulting structure creates a *PN junction* at the interface between the two layers which gives the diode its rectification property. A very simplistic, qualitative description of this rather complicated phenomenon of solid-state physics is provided in the section **THE PN JUNCTION DIODE** on page 3-26.

The schematic symbol of a diode is shown at right, along with a photo of a typical silicon signal diode. When the diode is *forward-biased* its resistance becomes very small, and current will flow through it in the direction shown (note that the schematic symbol includes an arrow (triangle) which points in the direction of the current flow). Forward-biasing is accomplished by applying a voltage so that the diode's *anode* is at a more positive (+) voltage than its *cathode*, as shown in the figure. As you can see in the figure, *the cathode is denoted by a line in the schematic symbol and is usually marked by a line or stripe on the physical diode's body*.

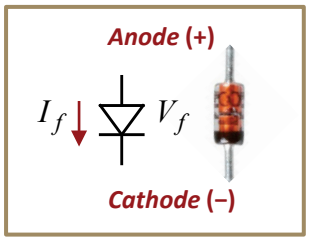


Figure 3-1: A typical silicon signal diode (type 1N4148) and its schematic symbol. The diode's glass case is actually only about 3 mm long.

When a diode is forward-biased and conducts a current of at least a few milliamps, the voltage drop across its two terminals (V_f in Figure 3-1) remains very nearly constant even for an increase in the forward current (I_f) of an order of magnitude or more. When a diode is *reverse-biased* (anode more negative than the cathode), on the other hand, only a very small *leakage current* flows through it. This leakage current is quite insensitive to the reverse voltage applied to the diode, at least until some critical reverse voltage is reached which causes the diode to suddenly *break down* (and, usually, catastrophically fail). Given this basic behavior of the diode as a rectifier, we can construct the following first approximation of its characteristics (good enough to use for many applications):

LOWEST-ORDER DIODE CHARACTERIZATION

A diode's basic behavior is characterized by the following two parameters:

V_f ***forward voltage drop***: the approximately constant voltage drop across the diode when it conducts current in its forward-biased direction.

I_R ***reverse leakage current***: the small current which flows through the diode when the applied voltage is less than V_f or whenever the diode is reverse-biased.

The forward voltage drop, V_f, of a *PN* junction diode is determined by the semiconductor material from which it is constructed; in the case of silicon (Si) this voltage is approximately 0.5–0.7V, whereas for germanium (Ge) V_f is 0.3–0.4V, and for LEDs V_f is >1V. The reverse leakage current, I_R, may vary by a couple of orders of magnitude depending on the type of diode, but is generally a few microamps or less (and may be much less). A *perfect diode* would be one with $V_f \equiv 0$ and $I_R \equiv 0$.

A *forward-biased* diode is said to be *on* or *conducting*. A *reverse-biased* diode is said to be *off*.

Basic diode rectifier circuits

Two common diode rectifier circuits are shown in Figure 3-2; we will analyze each using our simple diode model. The input source to each circuit provides an AC signal $v_{in}(t)$; the

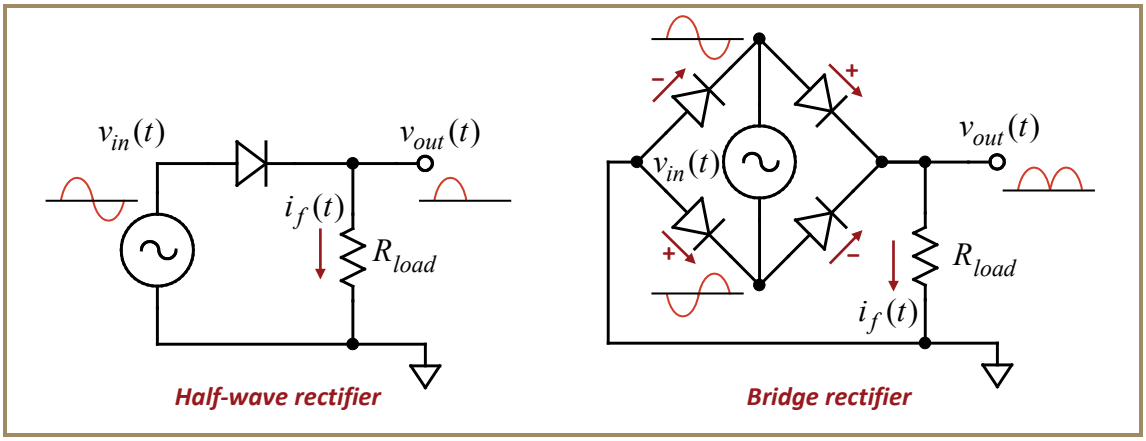


Figure 3-2: Diode rectifier circuits. The AC driving source is converted to a *rectified* output (always of the same polarity); with the chosen orientation of the diodes $v_{out}(t) \geq 0$. As shown by the waveform plots, only the positive half of an AC input cycle gets to the output of the *half-wave rectifier*, but the *bridge rectifier* routes both halves of an AC cycle to the output. A major drawback of the bridge, however, is that the input source must *float* (be isolated from ground potential), whereas the half-wave rectifier allows the source and the load (R_{load}) to share a ground connection.

circuits *rectify* the input to produce an output $v_{out}(t)$ across the load R_{load} which is always of the same polarity (in this case, $v_{out}(t) \geq 0$). We will analyze the action of the *half-wave rectifier* circuit first.

The half-wave rectifier (left-hand circuit in Figure 3-2) is simple to analyze, particularly if we make the assumption that the diode is perfect (both V_f and $I_R \equiv 0$). Whenever $v_{in}(t) > 0$, the diode's anode becomes positive, and it conducts current. The current from the source flows through R_{load}, and, since we assume $V_f = 0$, $v_{out} = v_{in}$. When $v_{in}(t) < 0$, the polarity across the diode reverses, and it turns off; thus $i_f = 0$, and therefore $v_{out} = R_{load}\, i_f = 0$. So the output waveform will appear as in the figure: only the positive half of an input cycle reaches the output.

The *bridge rectifier* (right-hand circuit in Figure 3-2) is more complicated but also more efficient. Assume the diodes are perfect, and consider first the case where the top terminal of the input source is positive with respect to its bottom terminal ($v_{in} > 0$). In this case the two diodes with the "**+**" current arrows can conduct (the other two will be off), and current from the top source terminal can flow through the upper-right diode, the load, and back through the lower-left diode to reach the bottom source terminal. Thus (since we assume the diodes' forward voltage drop vanishes), $v_{out} = v_{in}$. When the polarity of the source is reversed, the two diodes marked with "**–**" current arrows turn on instead, but the current from the source must still flow through R_{load} in the same direction as before, and now $v_{out} = -v_{in}$. Both halves of the source AC cycle are routed to the load, and the output waveform is the absolute value of the input waveform.

> Note the very specific orientation of the diodes in the bridge circuit in Figure 3-2; the roles of input and output may not be exchanged! In particular, the orientations of the pair of diodes at either of the two output terminals are the same: both cathodes attached to the positive output terminal and both anodes attached to the negative output terminal. This is an easy thing to get wrong when you attempt construct a diode bridge.

If the diodes aren't perfect in the two rectifier circuits, then a forward voltage drop V_f will be lost across each conducting diode (V_f is also commonly called a *diode drop*). Thus in the half-wave rectifier when the diode conducts, $v_{out} = v_{in} - V_f$ (see figure at right); in the bridge rectifier $v_{out} = |v_{in}| - 2V_f$. In both of these expressions, if the right-hand side ≤ 0, then $v_{out} = 0$: a diode remains off until the voltage drop across it has the correct polarity and reaches V_f. Similarly, nonzero I_R will flow in the opposite direction through a diode when it is reversed biased; this implies that v_{out} can become very slightly negative: $v_{out} = -R_{load} I_r$.

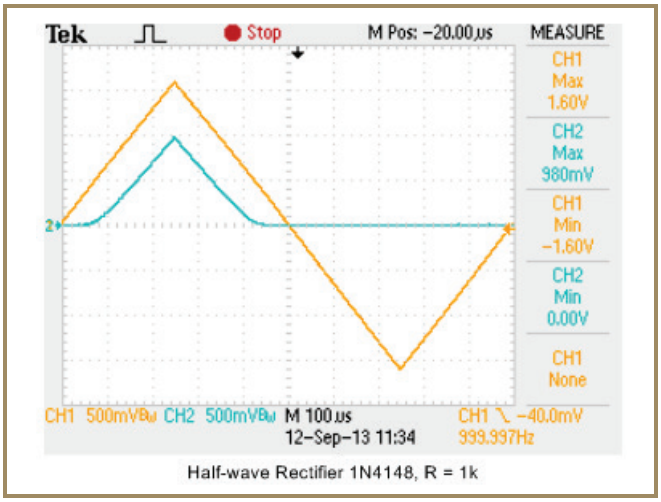


Figure 3-3: Response of the half-wave rectifier showing the effect of V_f, the diode's forward voltage drop. The diode is silicon, so V_f ≈ 0.6V (input is CH1, output CH2).

Precision rectifier circuits

A diode's forward voltage drop of a few tenths of a volt (Figure 3-3) means that the basic diode circuits of the previous section can hardly be considered to provide precision

rectification of an input signal, especially if the signal is small. As you might expect, adding the capabilities of an operational amplifier may be able to remedy this situation: consider the circuit in Figure 3-4.

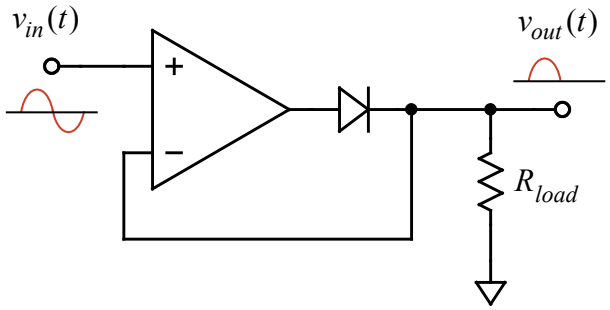


Figure 3-4: Simple, precision half-wave rectifier. Placing the diode inside the feedback loop ensures that the op-amp output will compensate for the diode's forward voltage drop. When v_{in} < 0, the diode turns off as the op-amp output goes negative; R_{load} then ensures that v_{out} = 0 in this case.

This circuit is essentially a voltage follower, but a diode has been added in series with the op-amp output *inside the feedback loop*. When $v_{in} > 0$, the op-amp will set its output one diode drop higher than v_{in} so that its *–Input* will equal v_{in}. Thus, for $v_{in} > 0$, $v_{out} = v_{in}$ even though the diode may have a significant forward voltage drop. When $v_{in} < 0$, the op-amp's output will go negative, and, because the diode's cathode is connected to ground through R_{load}, the diode will turn off, disconnecting the op-amp's output from v_{out}. Thus, for $v_{in} \le 0$, $v_{out} = 0$. The load resistor value should be around 1kΩ; if low output impedance is required, buffer the output with a voltage follower.

Adding another op-amp, we can construct a precision full-wave rectifier, Figure 3-5. The second amplifier (using op-amp U2) combines the original input signal with the half-wave rectifier output to form a fully rectified waveform (the absolute value of the input waveform).

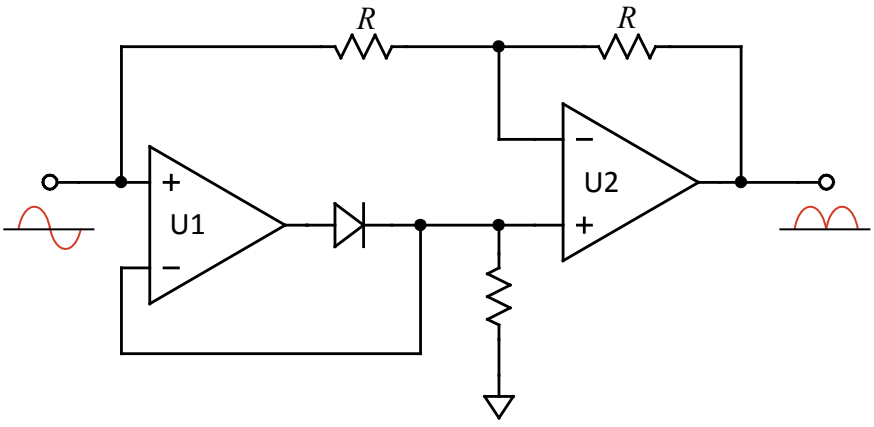


Figure 3-5: Precision full-wave rectifier. The output of a half-wave rectifier is combined with the original input signal by the amplifier U2. Its output is the absolute value of the input. The two resistors *R* should be well-matched in value (1% or better).

The two resistors used to form the feedback network for U2 should be well-matched in value; analysis of this circuit is left to the exercises. Typical outputs from the two rectifier circuits are shown in Figure 3-6; note that the effect of the forward diode drop has been eliminated (compare with Figure 3-3 on page 3-4).

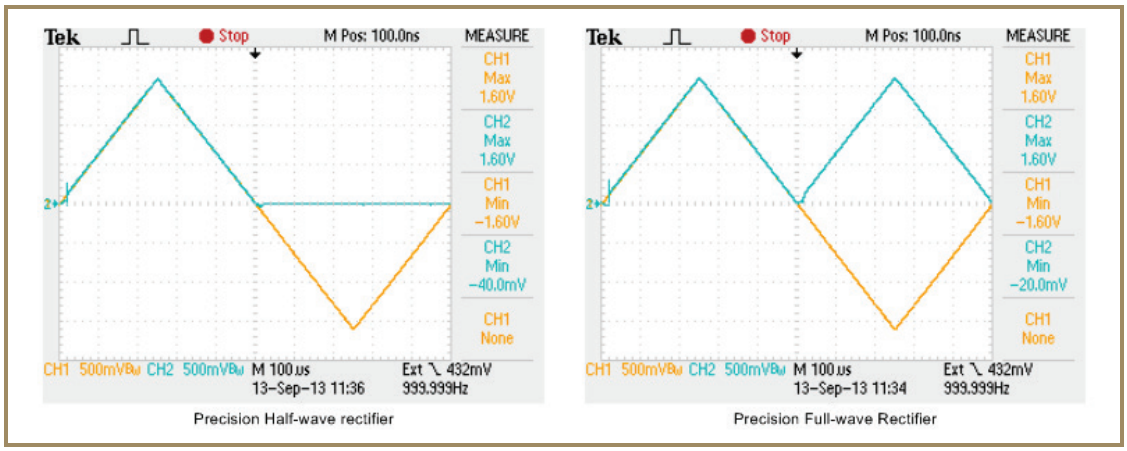


Figure 3-6: Precision rectifier outputs. Left: half-wave rectifier circuit of Figure 3-4. Right: full-wave rectifier of Figure 3-5. Input signal waveform and display scales are the same as in Figure 3-3 on page 3-4; note that the precision rectifier circuits eliminate the 0.6V forward voltage drop of the silicon diode used in the circuits. The small "glitches" in the output waveforms as the input goes through 0 are effects of the TL082 op-amp's finite slew rate.

Before proceeding further, some important diode limitations should be noted. For the 1N4148 silicon small-signal diodes (the type you will mostly use) the limits are:

$$V_R = 75\,\text{V} \qquad I_F = 300\,\text{mA} \qquad P_D = 0.3\,\text{W}$$

DIODE LIMITATIONS

Exceeding these limits could cause catastrophic diode and circuit failure:

V_R ***reverse breakdown voltage***: the maximum reverse-bias voltage which may be safely applied without the diode exhibiting avalanche or Zener breakdown.

I_F ***maximum forward current***: the maximum current the diode can tolerate when forward-biased.

P_D ***maximum power dissipation***: the maximum power dissipation the diode can tolerate without overheating and failing.

Warning

Because the forward voltage drop, V_f, of a diode is nearly independent of the current flowing through it, *the magnitude of the forward current must be limited by the external circuit*, or a forward-biased diode will quickly fail.

Exponential and logarithmic amplifiers

The exponential voltage-current relationship of the PN junction diode (equation 3.9 on page 3-32) may be exploited to build amplifiers with approximately exponential or logarithmic gains. If the diode current is much greater than I_R in equation 3.9, then the I-V relationship is approximately:

$$I = I_R e^{q_e V/\eta k_B T}$$

where, for a silicon diode, $I_R \sim 10^{-6}\text{mA}$ and $q_e/\eta k_B T \approx 20\text{V}^{-1}$ at room temperature. Thus, V will change by approximately 0.12V for a factor of 10 change in the current I, and $V \approx 0.6\text{V}$ for $I = 1\text{mA}$. Thus we can take the logarithm or the antilog (exponential) of an input voltage using the circuits in Figure 3-7. With resistor value R, the transfer functions are approximately:

$$|V_{out}| \approx 0.6\text{V} + 0.12\text{V} \times \log_{10}\left|\frac{V_{in}/\text{Volt}}{R/\text{k}\Omega}\right| \quad \text{(log amplifier)}$$

3.1

$$|V_{out}| \approx (1\text{mA} \times R)10^{|V_{in}/0.12\text{V}|-5.0} \quad \text{(exponential amplifier)}$$

The sign of the output voltage in each case is the opposite of the sign of the input, since the amplifiers are inverting. The gains and offsets of these circuits may be inconvenient, so more circuitry is usually added to scale and offset the output signals; the transfer function, however, is quite temperature-dependent.

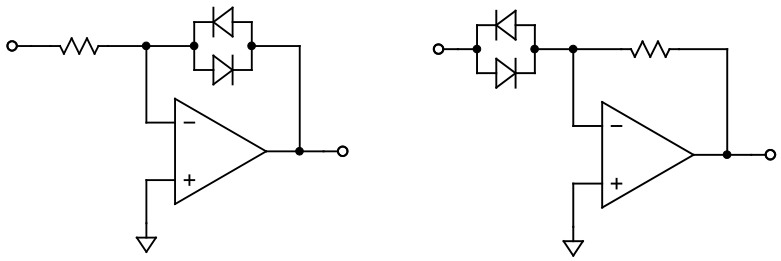


Figure 3-7: Inverting log (left) and exponential (right) circuits using diodes. These simple circuits are not very accurate and are very temperature-sensitive, but will work for noncritical applications. The paralleled back-to-back diodes will treat positive and negative input voltages the same, using whichever diode is forward-biased to dominate the circuit transfer function.

Temperature sensing

The PN junction's temperature sensitivity (equation 3.9 on page 3-32) makes the semiconductor diode a useful temperature sensor; in this section we present one example of a temperature monitor circuit. If a silicon diode is forward-biased so that its current is much greater than I_R in equation 3.9, then the I-V relationship is:

3.2

$$I \approx I_0 \exp\left[\frac{-q_e}{\eta k_B}\left(\frac{V_g - V}{T}\right)\right]$$

This implies that if a diode forward-bias current is held constant at, say, 0.1mA ($\sim 10^5 I_R$), then $V_g - V \propto T$. For a silicon diode at room temperature, $(V_g - V)/T \approx 2\,\text{mV/K}$, and this is the sensitivity of the forward-bias voltage to a change in junction temperature. Here is a circuit:

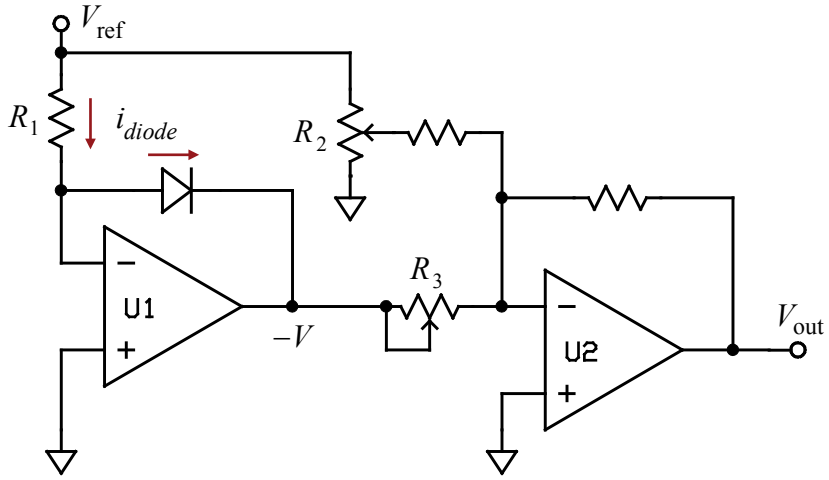


Figure 3-8: Temperature circuit using a diode sensor. Amplifier U1 maintains a constant diode forward-bias current of V_{ref}/R_1; the amplifier output is then $-V$, where V is the diode's forward-bias voltage. Amplifier U2 amplifies and offsets this voltage to provide the circuit's output (R_2 trims the offset, R_3 the gain).

To provide an output temperature response of 0.1V/°C, the gain of the U2 inverting amplifier stage should be ≈ 50 when R_3 is set to its center position; the summing amplifier's gain for the offset adjust input (R_2) should be 1 or less, so when R_2 is set near its center position the output corresponds to the diode temperature. There are several ways this circuit could be refined, but the Texas Instruments LM35 series of analog temperature sensor ICs implements a version of the circuit in Figure 3-8 in a single, small, calibrated device.

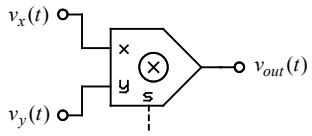
THE ANALOG MULTIPLIER

The ideal analog multiplier

Unlike the semiconductor diode, the modern analog multiplier is a sophisticated, complicated integrated circuit incorporating temperature-compensated voltage references, matched transistors, and laser-trimmed resistors. Much like the modern operational amplifier, this internal circuitry makes these devices particularly simple to utilize (although actual multipliers tend to behave in a less ideal manner than their op-amp cousins). In this section we consider the properties of an ideal analog multiplier and discuss applications of such a device; the next section will look at the use and limitations of an actual multiplier device.

An analog multiplier, of course, requires at least two user-controlled input signals, which are typically called the x and y inputs. The input and output analog values may be voltages or currents, depending on the particular device, but modern medium-speed, precision multipliers usually input and output *voltages*, so that is what we will assume here. The product of two voltages has, of course, units of Volt^2, but the output will be in Volts. Thus the product xy is divided by an additional voltage parameter, the *scale factor*, s, which converts the product to Volts (equation 3.3).

3.3 **Multiplier:** $v_{out} = v_x v_y / s$



The analog multiplier usually has a precision, built-in scale factor which is often $s = 10.0\text{V}$, but it may include an external terminal so that you can adjust the value of s (as shown in the schematic above). However the scale factor may be set, it is usually limited to positive values $(s > 0)$. Modern analog multipliers don't limit the signs of v_x and v_y (they support "four-quadrant multiplication"), but they may limit their magnitudes to $\leq s$.

Placing a multiplier in an op-amp's feedback loop provides for the calculation of the inverse operation, division, as shown in Figure 3-9. Assuming the negative feedback around the op-

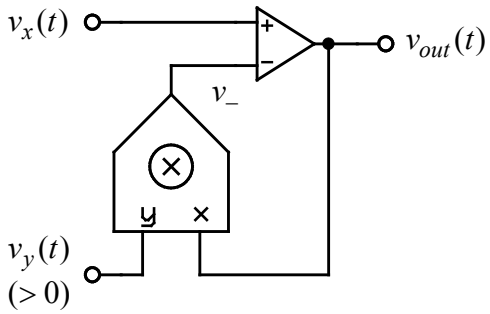


Figure 3-9: Basic divider circuit using an op-amp. Note that for the negative feedback to be effective, v_- must have the same sign as v_{out}, and this condition requires that $v_y > 0$.

amp is effective, then $v_- = v_+ = v_x$. But the multiplier output is $v_- = v_{out}\, v_y / s$, so the divider circuit output is:

3.4 **Divider:** $v_{out} = s\, v_x / v_y \qquad (v_y > 0)$

> The condition $v_y > 0$ in (3.4) comes from the requirement that v_- must have the same sign as v_{out} (see Figure 3-9), or the feedback will effectively change sign and become *positive*, which will cause the op-amp's output to proceed to one of its power supply limits (i.e. the output will *saturate*) until v_y becomes positive.

Since a typical multiplier will require that its inputs not exceed s, the gain of the circuit in Figure 3-9 with respect to the op-amp's *+Input* (v_x) will be greater than 1, and thus the circuit's bandwidth will be less than the op-amp's gain-bandwidth product f_{BW}; in fact, the bandwidth will be $f_{BW}(v_y / s)$.

You may calculate the square of a signal by connecting a multiplier's inputs together; the inverse operation, a square root circuit, again may be built by putting the multiplier in an op-amp feedback loop, but now circuit *latch-up* is a real possibility: the op-amp output will saturate (in this case at its negative limit) and *stay there* no matter what the input does. Using a precision rectifier configuration (as in Figure 3-4) is a popular way to avoid this latch-up problem (Figure 3-10).

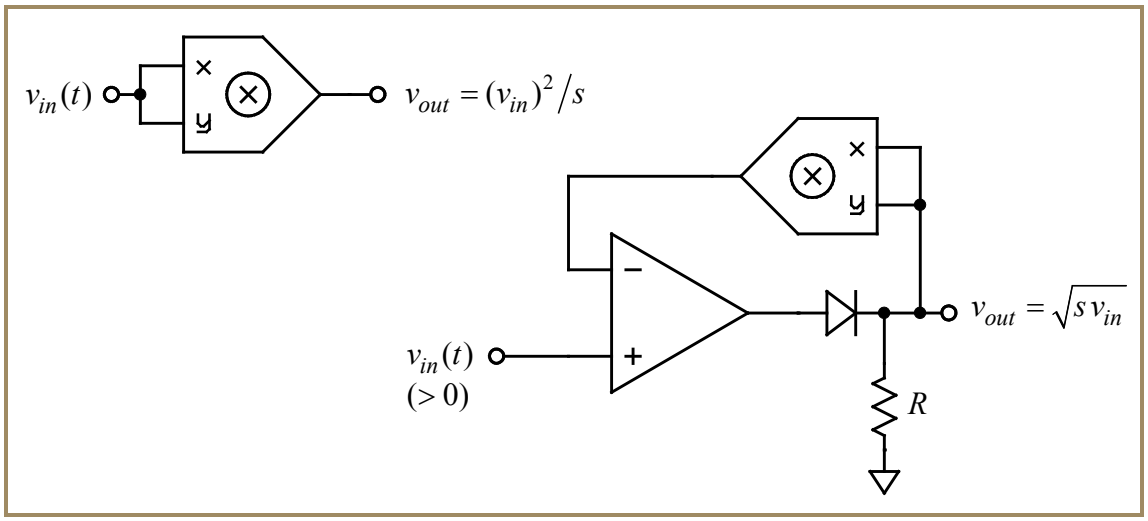


Figure 3-10: Square and square root circuits. The square root circuit is prone to *latch-up*: if the inputs to the multiplier were to go < 0, even momentarily, its positive voltage output would drive the op-amp into saturation at its negative voltage limit, permanently maintaining this undesirable state. The diode + resistor ensure that the multiplier inputs never go negative, avoiding latch-up. A value of about 1kΩ – 10kΩ is appropriate for *R*. Of course, *s* is the multiplier's *scale factor*.

A real analog multiplier IC

Because circuit designers often need to include op-amps and need to invert signals in their circuits, actual IC multiplier devices usually include extra circuitry to make the designer's job easier. In this section, we consider the Texas Instruments MPY634 device, which is included on the *ASLK PRO* breadboard; a similar, general purpose, relatively inexpensive multiplier is the Analog Devices AD633, which may be more readily available. In this section we discuss the use of the MPY634, a functional block diagram of which is shown below.

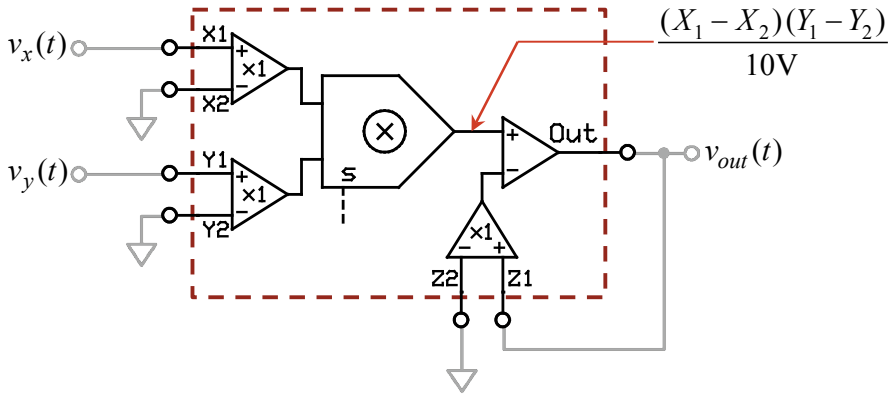


Figure 3-11: Functional block diagram of the Texas Instruments MPY634 analog multiplier. Besides the basic multiplier, the IC includes a general-purpose op-amp on the multiplier's output and has differential inputs for all parameters, including the *Z* inputs to the op-amp's feedback (*–Input*) terminal. Shown in light gray are the connections needed to emulate the basic generic multiplier discussed in the previous section.

The MPY634 has differential inputs for all input parameters and includes a general-purpose op-amp on the multiplier output so that it can be easily configured for different scale factors or for inverse operations. In Figure 3-11 the output op-amp has been configured as a voltage follower (output fed back through its Z1 input), and the negative differential input for each parameter has been grounded; the resulting circuit acts as the basic multiplier described by equation 3.3 in the previous section. Note that the default scale factor is 10V; this is the value if the IC's scale factor terminal (**S** in the figure) is left disconnected. The MPY634 basic multiplier accuracy in this configuration is approximately 2%. The op-amp $f_{BW} \geq 6\text{MHz}$, and it has a $20\text{V}/\mu\text{sec}$ slew rate, so it has similar performance to the TL082 op-amps on the breadboard.

Divider and square root configurations are shown in Figure 3-12 (on page 3-12). Because the multiplier output is connected to the internal op-amp's *+Input*, you have to *invert the multiplier's output* to provide negative feedback around the op-amp (as in Figure 3-9 and Figure 3-10). This is accomplished by using the (–) differential terminal for one of the arguments to the multiplier, so its output is the negative of the product of its two inputs. Similarly, the other op-amp input (to its *–Input* terminal) is also inverted by using the input's

(–) differential terminal. The MPY634 data sheet has more examples demonstrating this device's flexibility.

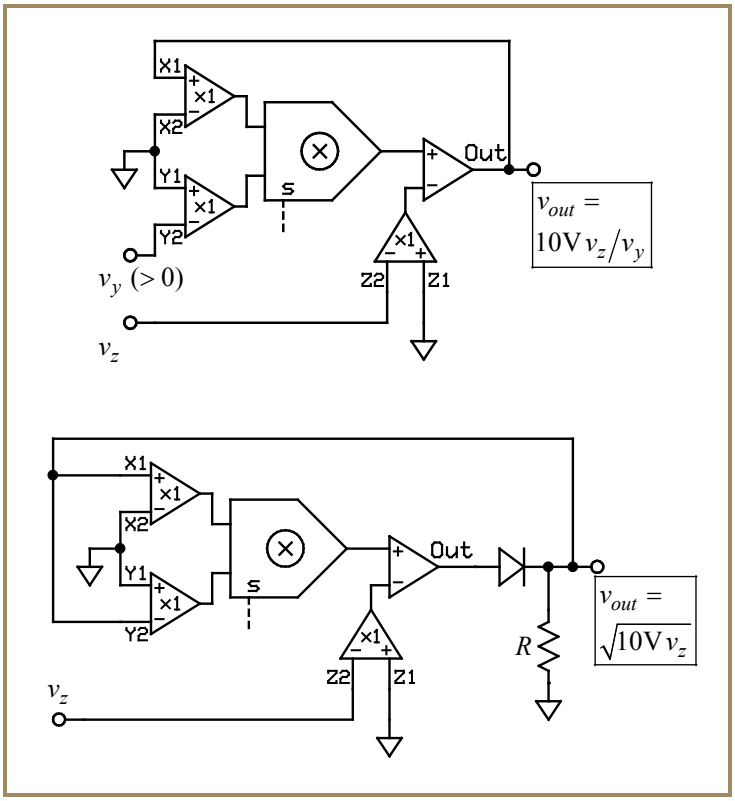


Figure 3-12: MPY634 divider (top) and square root (bottom) circuits.

In each case the multiplier sub-circuit must be in the *negative feedback loop* of the op-amp (see Figure 3-9 and Figure 3-10); to accomplish this using the MPY634, the multiplier output must be the *negative* of the product of its inputs, which is accomplished by *inverting only one* of its inputs. The v_z input to the op-amp is also inverted, so that it effectively becomes a *noninverting* op-amp input (–1 × –1 = +1).

A diode and resistor are still needed to prevent latch-up of the square root circuit output, as discussed in the previous section. The resistor value R should be about 1kΩ – 10kΩ, as before (Figure 3-10).

You must always provide negative feedback for the output op-amp of the MPY634.

Any unused differential input terminals must be connected to ground.

PRELAB EXERCISES

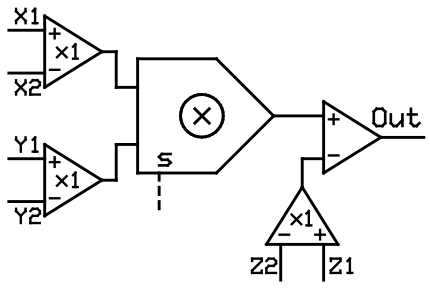
1. Consider the simple half-wave rectifier circuit (left-hand schematic in Figure 3-2 on page 3-3). If the diode is perfect (both its forward voltage and reverse leakage current are 0), then what is the circuit's *output resistance* during a positive half-cycle (when $v_{in}(t) > 0$)? What about during a negative half-cycle ($v_{in}(t) < 0$)? Include R_{load} as part of the circuit when calculating the output resistance.

   Hint: review the section on *output resistance* starting on page 1-43 of Experiment 1. Consider two distinct cases: one when the diode is forward-biased (and, since $V_f \equiv 0$, its resistance = 0), and the other when the diode is reverse-biased (so, since $I_R \equiv 0$, its resistance is infinite).

2. How does the full-wave rectifier circuit in Figure 3-5 (on page 3-5) work? Analyze the two circuit operating states separately: (1) the diode is forward-biased; (2) the diode is reverse biased. Assume the op-amps are ideal; for each of these states determine:
   a. the required input voltage (v_{in}) range for the circuit to be in that state
   b. the voltage at U2's +Input
   c. the circuit's transfer function v_{out}/v_{in}

3. Consider the square root circuit in Figure 3-10 on page 3-10. If the input $v_{in}(t) < 0$, then what is v_{out}? What is the output voltage of the op-amp? What happens to the op-amp output and v_{out} as the input rises through 0? Does the forward-bias diode voltage drop affect the accuracy of v_{out} (assuming the op-amp is ideal)? Why or why not?

4. By configuring the MPY634's output op-amp to provide gain, you can effectively adjust the device's multiplication scale factor to something less than its 10V default. Complete the schematic diagram below by adding a feedback network to the op-amp and properly connecting the input terminals so that you have a multiplier with a scale factor $s = 5\text{V}$ (assign appropriate values to any resistors you include).



Another problem on the next page…

5. Design a circuit using the MPY634 to create an amplifier with a *voltage-controlled gain*. The gain control input voltage range should be −7V to +7V (minimum), and the circuit gain should equal the control voltage in volts, e.g. −2V → gain = −2, etc. The allowable signal input and output voltage ranges (without *clipping* or distortion of the output) should be −7V to +7V as well (as long as the circuit gain is not set too high). Use additional op-amp amplifier stages as part of the circuit design if you need them (but you may not need them). Assume that the input signal is ground-referenced, so you don't need a fully differential input for the signal. The gain accuracy should be no worse than ±10%.

   Provide a full schematic of the circuit with all component values included. Use only these standard 5% tolerance resistor values: a power of 10 times…

   1.0, 1.1, 1.2, 1.3, 1.5, 1.6, 1.8, 2.0, 2.2, 2.4, 2.7, 3.0, 3.3, 3.6, 3.9, 4.3, 4.7, 5.1, 5.6, 6.2, 6.8, 7.5, 8.2, 9.1

   You will build and test this circuit during your lab session, so think carefully about it!

LAB PROCEDURE

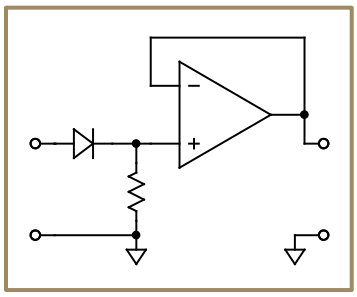
Overview

The semiconductor diode and the analog multiplier are two examples of nonlinear components, but they are nevertheless very different in their behaviors and their applications. Spend sufficient time during lab building circuits with each of them so that you become comfortable with using such elements in your designs. Spend some time investigating light emitting diodes (LEDs) and Zener diodes as well as the basic silicon diode. Combine circuits using them with the amplifier and filter designs you have already practiced with so that you start to think about and develop more complicated applications.

Detailed procedures

Diode rectifiers and an AM demodulator

Build a simple half-wave rectifier circuit using a 10k load resistor and a voltage follower to buffer the rectifier's output (figure at right). Using a 1kHz signal input of a couple of volts peak-peak amplitude, confirm that the diode's turn-on voltage is approximately 0.6V, as shown in Figure 3-3. Replace the diode with an LED. How to determine the LED polarity (anode v. cathode) is shown in Figure 3-13. Note that you must increase the signal generator output to at least 3V peak-peak to get the LED to conduct. What is its approximate turn-on voltage?



Again using a silicon diode, replace the simple half-wave rectifier circuit with a precision version (Figure 3-4). Examine the output and note that the 0.6V diode drop is no longer evident. Compare your results to Figure 3-6.

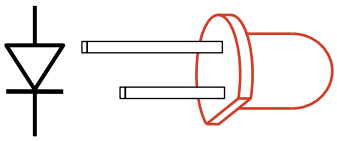


Figure 3-13: Determining LED polarity. The long lead is the anode (+); the short lead (nearest the flat section of the LED circumference) is the cathode (–).

Now filter the rectified output by connecting first the 0.1µF and then the 1µF capacitor in parallel with the 10k load resistor. Compare your results to the filtered half-wave rectifier output plotted in Figure 3-14 on page 3-17.

Using the 0.1µF capacitor to filter the rectifier output, next configure the signal generator to input an *amplitude modulated* (AM) signal to your rectifier. Use a *carrier frequency* of 100kHz, a *modulation frequency* of 200Hz, and a AM *modulation depth* of 30%. Get your TA to help you set up the signal generator; trigger the oscilloscope using the rectifier circuit output. Does the rectifier output *detect* the modulation (output mostly the 200Hz *envelope* of the 100kHz carrier)? Try different load resistor and filter capacitor combinations. Try varying the modulation frequency and modulation depth and note the effects on the output.

Using a multiplier as a frequency doubler

Construct a basic squarer circuit using one of the analog trainer's MPY634 multipliers; refer to Figure 3-11 on page 3-11 for the basic multiplier connections. Using a 5V peak-peak input signal, confirm that the squarer's output is given by $v_{out}(t) = v_{in}(t)^2 / 10\text{V}$.

Once the circuit is working and behaves as expected, note that the circuit's response to a sinusoid input is, of course, another sinusoid at twice the input frequency + a constant (DC) offset (since, of course, the output of the squarer is nonnegative). You may remove the DC component by *AC coupling* the output, as you learned in Experiment 2. Add an RC high-pass filter to couple the squarer output to a noninverting, gain 11 amplifier stage.

You have now constructed a *frequency doubler* — a sinusoid input produces a sinusoid output at twice the frequency. Using an input frequency of approximately 1kHz, determine what input signal amplitude is required to produce an output of the same amplitude. Increase the input frequency until you find the −3dB upper bandwidth limit of your doubler.

Amplifier with voltage-controlled gain

Build and test the circuit you designed for prelab exercise 5. The gain control voltage may be generated by a computer *DAQ* analog output port which you control using the National Instruments *Measurement & Automation Explorer* application; the lab instructor or your TA will show you how to accomplish this.

Additional, self-directed investigations

Maybe try building a cubing circuit: $v_{out} \propto v_{in}^3$; how many multipliers would this take? Consider a log or exponential amplifier, or one or more circuits from the **MORE CIRCUIT IDEAS** section (such as the true RMS circuit), or try a nontrivial one of your own design. Look back at earlier experiments to see if there are any other circuits you would like to investigate.

Lab results write-up

As always, include a sketch of the schematic with component values for each circuit you investigate, along with appropriate oscilloscope screen shots and, if appropriate, Bode plots. Make sure you've answered each of the questions posed in the ***Detailed procedures*** section.

MORE CIRCUIT IDEAS

Peak detectors

If you put a capacitor in parallel with the output of a simple, half-wave rectifier as shown in Figure 3-14 below, then whenever the diode conducts the AC voltage source will quickly charge the capacitor up to match its voltage (minus the diode's forward voltage drop). When the AC voltage source passes its peak positive output and starts to decrease, its voltage drops below the capacitor's voltage, and the diode turns off. The capacitor then begins to relatively slowly discharge through the load resistor until the AC source voltage output again becomes high enough to turn the diode back on and recharge the capacitor. As a result, the output voltage stays near the peak positive input voltage value; the smaller the capacitor or the smaller the load resistor, then the more the capacitor will discharge during a cycle. This simple circuit is useful as part of a *power supply* to turn the 60Hz power line AC voltage (usually coupled through a transformer) into a nearly constant DC voltage to power an electronic device. The amplitude of the capacitor charge-discharge output oscillation is called the power supply *ripple voltage*; small ripple voltage is desirable, so power supply filter capacitors tend to be large; they are usually electrolytic (refer again to the photo in Experiment 2, Figure 2-1 on page 2-6).

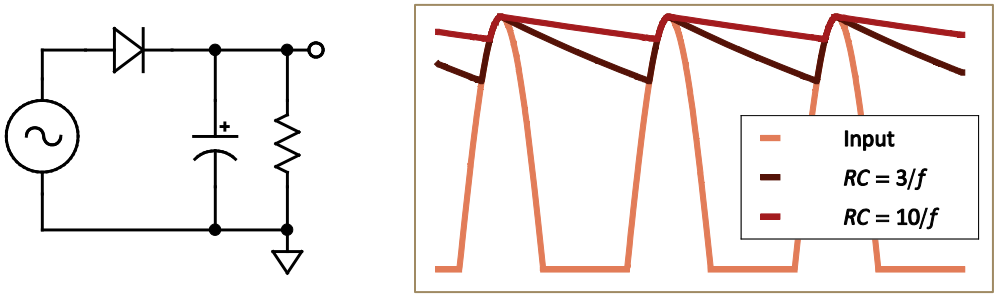


Figure 3-14: Using a filter capacitor to smooth the output of a simple, half-wave rectifier. The capacitor *C* is charged whenever the input voltage exceeds the output voltage so that the diode is forward-biased; when the input voltage drops and the diode becomes reverse-biased, the capacitor discharges through the load (the resistor *R*). The graph at right shows the resulting output voltage variation for various *RC* time constant values; *f* is the frequency of the input source. The effect of the diode's forward voltage drop is not included in the graph.

If we use this idea with a precision half-wave rectifier circuit, we get the *peak detector* shown in Figure 3-15, which works just the same as the simple circuit discussed above, except now an op-amp is used to correct for the diode forward voltage drop, and a voltage follower stage is added to the output so that the subsequent load will not discharge the capacitor. This circuit will hold the maximum positive input voltage encountered for quite a long time; if a resistor is not included and a high-quality capacitor is used, then the capacitor's discharge will only be because of the follower op-amp's input bias current and the diode's reverse leakage current, the sum of which can be kept to only a few nanoamps by

careful component selection. In this case, a 10μF capacitor would provide an output voltage which decayed at less than 1mV/sec. If you need to hold a peak voltage value for longer than this, then your best bet would be to record a *digitized* version of the output voltage value.

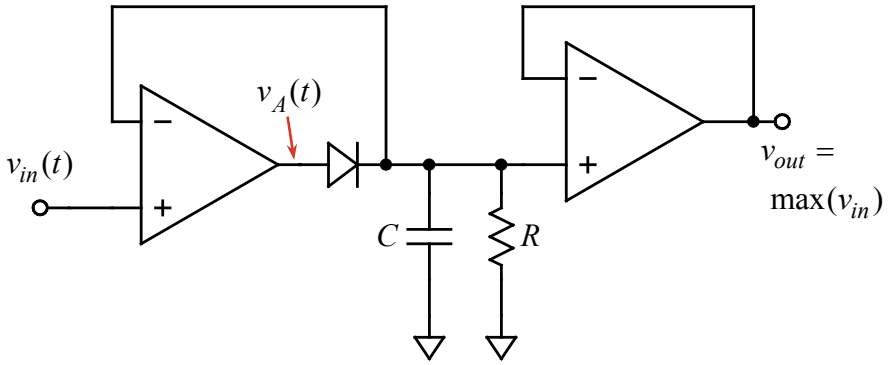


Figure 3-15: A peak detector circuit, which outputs the most positive value of the input signal seen so far; reversing the diode would output the most negative value of the input. The output voltage will exponentially decay toward 0 with time constant *RC*; because of the voltage follower on the output, *RC* may be made quite long by choosing a large value for *R*, or the resistor could even be replaced by a switch to reset the output to 0. The behavior of the first op-amp's output voltage, v_A, is discussed in the text.

One potential problem with the peak detector shown above is that it can output a voltage significantly lower than the peak amplitude of a high frequency input signal. This is a big problem for some applications, so now consider ways to improve the peak detector's speed. There are two major design issues which increase the time it takes to change the capacitor's voltage: (1) the amount of current the op-amp output can supply, and (2) the op-amp slew rate. The first issue is easy to deal with, so we consider it first. When the input exceeds the voltage currently stored by the capacitor, the diode becomes forward-biased, and the op-amp output can charge the capacitor toward the new maximum voltage. The rate that the capacitor's voltage will change is given by $i_{max} = C\,dv/dt$, where i_{max} is the maximum output current the op-amp is capable of supplying (usually specified in the op-amp's data sheet). The TL082, for example, can output up to about 40mA into a discharged capacitor, but its available output current decreases as its output voltage rises. If your application must track a signal whose peak amplitude changes rapidly, use as small a capacitor value *C* as you can (consistent with your required peak hold time) or get an op-amp with a larger current output capacity.

The second problem is more difficult to handle, because the design of the circuit in Figure 3-15 exacerbates the speed limitation imposed by the op-amp slew rate. Whenever the input is less than the voltage stored by the capacitor, the op-amp output voltage (v_A in Figure 3-15) goes all the way down to its negative limit, because the diode is reverse-biased (opening the feedback loop), and therefore $(v_{in} =) v_+ < v_- (= v_{out})$. If and when the input returns to above

v_{out} (so that $v_+ > v_-$), the op-amp output must change from its negative limit up to a diode-drop above v_{out} in order to forward-bias the diode and start to increase the capacitor's charge. The time it takes to change the output voltage is, of course, limited by the op-amp slew rate. Since the output starts at the op-amp's negative rail, the output must change by at least several volts, which will take about a microsecond or more for the TL082 ($13\,\mathrm{V/\mu sec}$ slew rate). One solution, obviously, would be to get a faster op-amp (the TI THS3201-EP, for example, has $9800\,\mathrm{V/\mu sec}$ slew rate, 1.8GHz f_{BW}, 100mA output current, and costs $5.54 each) — this is the only solution if your circuit must respond to very narrow, infrequent pulses. If, on the other hand, you must track the amplitude of a high-frequency sinusoid, a simple modification to the design of the peak detector circuit can greatly increase its frequency response (Figure 3-16).

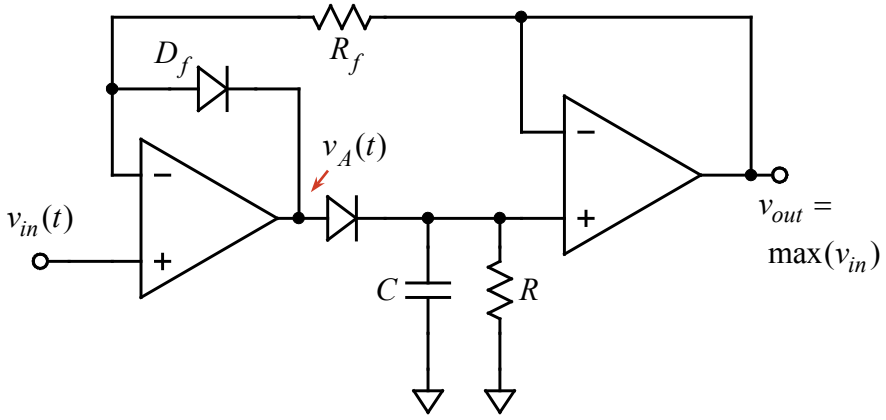


Figure 3-16: A modified peak detector with improved frequency response. The feedback loop now comes from the voltage follower output; the addition of R_f and D_f ensures that the loop stays closed when v_{in} drops below v_{out}. Now the first op-amp output (v_A) never gets further than a diode drop away from v_{in} while it is below v_{out}, rather than saturating at its negative limit as is the case in the Figure 3-15 design.

This modified design adds resistor R_f and diode D_f. First consider the state when v_{in} has been rising through the previously stored peak voltage, so that the output diode is conducting and C is being charged. Now $v_{out} = v_C = v_- = v_{in}$, and v_A is a diode drop higher, so the feedback diode D_f is reverse-biased and doesn't affect the circuit's operation (no current flows through R_f in this case, ensuring that $v_- = v_{out}$). When v_{in} passes its new peak and starts to drop, the op-amp output v_A rapidly decreases, turning off the output diode. Momentarily, $v_- = v_{out}$ stays constant, but once v_A has fallen a diode drop below v_- the feedback diode starts to conduct, and v_- will now follow v_A, remaining a diode drop above it. Since $v_{out} = v_C$, v_{out} remains constant, and current flows from the output op-amp through R_f and D_f to the first op-amp output (at voltage v_A). The first op-amp output voltage will slew down until $v_- = v_{in}$, and then will maintain that condition by keeping v_A a diode drop below v_{in}.

As v_{in} again rises and approaches the stored voltage v_{out}, v_A is only one diode drop away (unless the rate of change of v_{in} exceeds the op-amp slew rate). Thus as v_{in} passes v_{out}, the op-amp output only has to slew by two diode drops to turn on the output diode and start charging C, rather than slewing all the way up from its negative output rail. A two-diode-drop slew takes only about 0.1µsec for the TL082, an order of magnitude quicker than a slew of 13V up from its saturated negative output (at about −11V for our breadboards).

Zener diode regulator

Some diodes are designed to be used in their *reverse-bias breakdown region*: *Zener* and *avalanche* diodes. These types of diodes are very useful as voltage references, simple voltage regulators, and overvoltage protection devices. The Zener and avalanche breakdown effects are described very briefly in the section **THE PN JUNCTION DIODE** on page 3-26; in this section we address one common application of these devices.

Figure 3-17 shows a typical Zener reverse-bias *I-V characteristic curve* — a plot of the relationship between applied reverse-bias voltage and resulting current flow through the diode. The very steep portion of the curve corresponds to the diode's reverse-bias breakdown region; because the curve in this region is so steep, you can see that changes in the diode reverse current correspond to very small changes in reverse-bias voltage. Thus, *the voltage across the Zener diode in this breakdown region is very insensitive to changes in the current through it*. This characteristic makes the Zener diode useful as a simple *voltage regulator*.

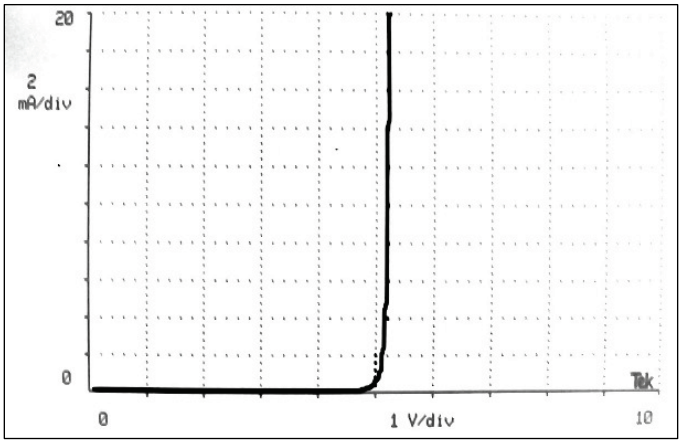


Figure 3-17: Measured Zener diode I-V curve. The reverse diode current is plotted as a function of the applied reverse-bias voltage. As the applied voltage exceeds 5V, the diode current dramatically increases as the diode suffers reverse breakdown. As can be seen from the plot, diode reverse currents above about 15 mA correspond to a reverse-bias voltage of 5.3V. A lab instrument called a *curve tracer* was used to perform this measurement (Tektronix 571).

Assume we have a power supply whose output voltage is greater than the maximum voltage allowed by some device you need to power, and this device needs a *well-regulated power supply voltage* (the voltage is largely unaffected by changes in load current or the input

voltage supplied to it). By using a Zener diode whose reverse breakdown voltage is the same as the voltage you need to power your device, you can build a simple voltage divider circuit which will satisfy the requirements for your device's power supply.

Consider the circuit in Figure 3-18, which is just a voltage divider with a reverse-biased Zener diode as the bottom element (note how the cathode end of the schematic symbol for a Zener diode differs from that for a regular diode). If the input voltage exceeds the Zener diode's breakdown voltage, then the diode may break down, with voltage V_R across it. Thus $v_{out} = V_R$. Now the voltage across the resistor R is determined: $v_{in} - V_R$, and the current through the resistor is also known: $i_{in} = (v_{in} - V_R)/R$. This current is divided between the current through the Zener diode, i_Z, and that through the load, i_{load}. Thus $i_Z = i_{in} - i_{load}$.

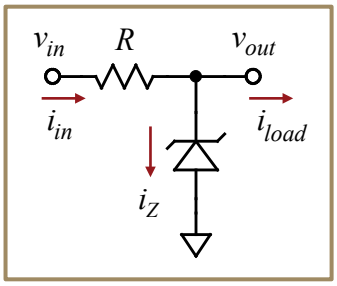


Figure 3-18: Zener diode voltage regulator circuit. The input voltage v_{in} must be greater than the Zener diode's reverse breakdown voltage, V_R. This large voltage causes the diode to suffer reverse breakdown, with the difference $v_{in} - V_R$ imposed across resistor R, establishing the input current i_{in}. This current is divided between that required by the load, i_{load}, and the reverse current through the diode, i_Z.

Note how this circuit holds the voltage applied to the load constant (equal to V_R) even if i_{load} or v_{in} varies. Changing i_{load} doesn't change v_{out} or i_{in} because the voltages across the diode and across R don't change; so for changes in i_{load}, $\Delta i_Z = -\Delta i_{load}$ (we assume here that the Zener's I-V characteristic curve is very steep in its breakdown region: $dV_R/di_Z \approx 0$). Similarly, changes in v_{in} don't affect v_{out} because, although i_{in} changes, $\Delta i_Z = \Delta i_{in} = \Delta v_{in}/R$, but the steep Zener I-V characteristic ensures that v_{out} is nearly unaffected by this change in i_Z.

Designing a Zener voltage regulator circuit

This example will illustrate how you would design a voltage regulator using the Zener diode circuit in Figure 3-18. Assume the load is a digital circuit which requires a stable voltage of no more than 5.5V in order to operate (the so-called CMOS *digital logic family* to be discussed in a later experiment would fall into this category), and you wish to power it from a 9V battery. The circuit will require a minimum of 10mA, but could draw as much as 20mA when an LED (light emitting diode), a part of the circuit, is illuminated. You want the circuit to work properly even if the battery has discharged to the point where it can only supply 7V.

Here are the basic Zener regulator design steps:

1. Choose a Zener diode breakdown voltage. The Zener diode whose characteristic curve is shown in Figure 3-17 should work, since its nominal breakdown voltage of 5.3V is near, but less than, the specified 5.5V limit.
2. Choose a *minimum* Zener current which will give good voltage regulation (steep I-V curve); according to Figure 3-17, a current of 10mA is well into the diode's breakdown region, so we will design the circuit for that minimum diode current.
3. The minimum Zener current + the maximum load current = the design target for the current i_{in} through the resistor R (Figure 3-18). Therefore, for this example, design $i_{in} = 10\text{mA} + 20\text{mA} = 30\text{mA}$. The resistor value is then chosen to give the correct voltage drop at this design current when the input voltage is at its *minimum* (7V for this example). Thus $R = (7\text{V} - 5.3\text{V})/30\text{mA} = 56.7\Omega$. Choose the closest standard resistor value, which, in this case, is $\boxed{R = 56\Omega}$.
4. Now consider what would happen when the circuit experiences the opposite extreme: maximum source voltage v_{in} (9V) along with minimum load current i_{load} (10mA). Since the Zener diode will keep the output voltage at 5.3V, the voltage drop across R is now $9\text{V} - 5.3\text{V} = 3.7\text{V}$, and with the chosen value for R, $i_{in} = 3.7\text{V}/56\Omega = 66\text{mA}$. Under these conditions the diode current will have increased to 56mA ($i_Z = i_{in} - i_{load}$).
5. Use the results from (4) to determine the worst-case power dissipations in the resistor and the diode:

   Resistor: $P = VI = 3.7\text{V} \times 66\text{mA} = 0.24\text{W}$
   Diode: $P = VI = 5.3\text{V} \times 56\text{mA} = 0.3\text{W}$

   These results specify the required minimum power dissipation capabilities of the components, which should be at least 150% of the calculated values (to be on the safe side).

Note that the power required from the 9V battery is $9\text{V} \times 66\text{mA} = 0.6\text{W}$, while the load may be consuming only $5.3\text{V} \times 10\text{mA} = 0.05\text{W}$, so the efficiency of our simple regulator is a measly $0.05/0.6 \approx 8\%$ (ouch!). This is typical for a Zener voltage regulator, which is called a *shunt regulator*: the current drawn from the source is always greater than the maximum required load current, even when the load is drawing little or no current most of the time.

> Zener diode voltage regulator circuits are really only practical when both the variation in the required load current (i_{load}) and the expected variation in the source voltage (v_{in}) are small.

For this particular problem a more efficient solution would be to use a *series regulator* (the standard type implemented by special-purpose voltage regulator ICs), or, even better, a *DC-*

DC converter, which is a type of *switching regulator* that can convert 9V power to 5V power with efficiencies exceeding 80%.

LEDs

A *light emitting diode* (LED) is a PN junction diode made from a semiconductor material with a relatively large energy gap. Consequently, when an electron and hole recombine the energy released may be carried away by a visible light (or near infrared) photon. Otherwise, LEDs behave much the same as any other semiconductor diode. Because of its larger gap voltage, the forward-bias voltage drop for an LED is significantly higher than for a silicon diode; typical forward voltages for various LEDs are:

Table 3-1

Typical LED Forward Voltage Drop (10mA forward current)

IR (950nm)	Red (630nm)	Yellow (590nm)	Green (565nm)	Blue (470nm)
1.2V	**1.8V**	**2.0V**	**2.2V**	**3.4V**

White LEDs are usually constructed using a blue emitter with a fluorescent coating, so their forward voltage drop is the same as for blue LEDs.

> ***Caution***
>
> *LED reverse breakdown voltage is typically only 5V!* Be careful to avoid using an LED in a circuit which could cause it to suffer reverse breakdown.
>
> Since the LED is a PN junction diode, *its forward current must be limited by the external circuit*; this is most often accomplished using a resistor in series with the diode.

The intensity of the light output of an LED is very nearly proportional to its forward-bias current; 10mA is usually more than sufficient to provide a nice, bright indicator light. Be careful to avoid excessive forward currents or the LED will quickly fail; if a very bright source is needed, specially designed and cooled LEDs are available (at a price). LED indicators are often controlled by digital logic circuitry, and are turned off or on depending on the logic circuit state; we'll discuss these sorts of circuits in a later experiment. In this section we consider some designs involving our analog circuits.

Often you will want an LED *pilot light* to illuminate whenever the power supply to a circuit is activated (like the LEDs on the circuit trainer breadboard). Some circuits to do this job are shown in Figure 3-19. For circuits (a) and (b) from that figure, the minimum voltage required for any significant illumination is the LED's turn-on voltage (Table 3-1). The resistor value is chosen so that the desired current will flow through the diode to give the desired intensity.

For example, illuminating a green LED with 10mA from a 12V source would require a resistor value of $R = (12\text{V} - 2.2\text{V})/10\text{mA} \approx 1\text{k}\Omega$; a 5V source would require $R \approx 270\Omega$.

The circuit in (c) is interesting because it uses a Zener diode to set the minimum voltage required for the LED to illuminate. This would be useful if you don't want the pilot light to illuminate if the power supply voltage is less than some threshold value. For example, assume you are using a 9V battery for your power supply, but the circuit needs at least 7V to operate properly. If the Zener breakdown voltage is 5.1V and you use a red LED (turn-on 1.8V), then at least $5.1\text{V} + 1.8\text{V} = 6.9\text{V}$ would be needed to illuminate the LED. With a current of 10mA at 9V, $R = (9\text{V} - 6.9\text{V})/10\text{mA} \approx 200\Omega$.

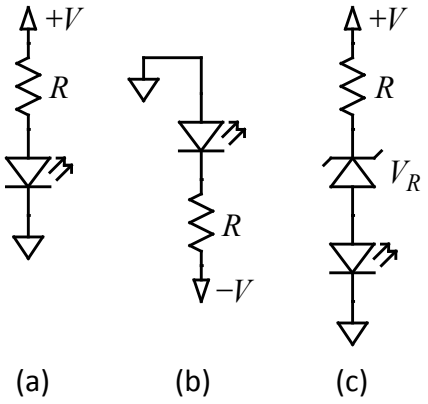


Figure 3-19: LED pilot light circuits. (a) positive supply voltage; (b) negative supply voltage; (c) positive voltage, but with a minimum voltage threshold for LED illumination using a Zener diode. See the text for details.

True RMS measurement using analog multipliers

The usual way to determine the power transmitted by an arbitrary, time-varying signal is to measure the average of its squared amplitude (where the average is over a time interval usually much longer than the period of the lowest frequency AC component in the signal). Consequently, the root-mean-square (*RMS*) amplitude of a waveform is typically used to characterize its magnitude:

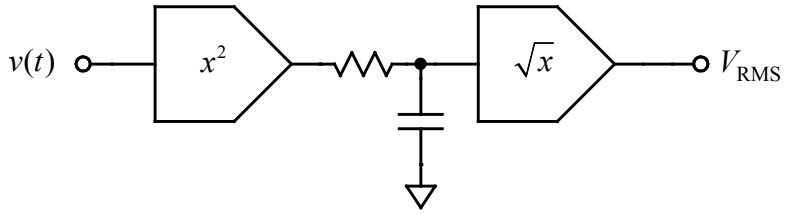
3.5

$$V_{\text{RMS}} \equiv \sqrt{\left(\overline{v(t)^2} \right)}$$

For a DC voltage, V_{RMS} is just the DC voltage value; for a sinusoid, it is $1/\sqrt{2}$ of the phasor magnitude, or about 35% of its peak-peak voltage. For a complicated signal composed of a DC and several AC components, V_{RMS} is the Pythagorean sum (square root of the sum of the squares) of the individual component *RMS* values.

Since the mean of a time-varying signal is just its DC component (see equation 2.2, page 2-4 of Experiment 2), we can extract the mean of a signal using a low-pass filter with a cutoff frequency well below the lowest frequency of any AC signal component (the effective

averaging time, for example, of a simple *RC* low-pass filter is approximately equal to its *RC* time constant). An analog circuit to output the RMS value of a time-varying input signal could then be constructed thusly:



The scale factors of the multipliers used for the square and square root circuits don't matter as long as they are the same; Figure 3-12 shows a square root circuit using the MPY634.

If you experiment with this circuit, use a variety of input waveforms with various frequencies and RMS amplitudes; compare the circuit's DC output voltage to the signal generator's RMS amplitude setting (make sure that the signal generator output setup is such that it reports the amplitude assuming a "High-Z" load). Estimate the circuit's accuracy and its high frequency response limit.

THE PN JUNCTION DIODE

Insulators, conductors, and semiconductors

The electrical conductivities of solid materials for the most part fall into one of two classes: conductors and insulators (metals make up most of the conductors, and nonmetals are usually insulators). Although all materials are very nearly electrically neutral (equal numbers of protons and electrons, so that they carry no net charge), the nature of the chemical bonds which bind the atoms or molecules of a solid to one another determines its class of electrical conductivity.

An atom's *valence electrons* — the outer, most weakly bound electrons — are the ones which participate in chemical bonding. The atomic nucleus along with the much more strongly bound inner electrons comprise a positively-charged *ion core* which remains intact and is surrounded by the interacting valence electrons. The chemical bonding process causes these many ion cores in a solid to arrange themselves in a mostly regular, crystalline structure. This regular, periodic array of positively-charged cores creates a similarly regular, periodic electrostatic field within which the myriad valence electrons move.

The quantum-mechanical nature of these microscopic, negatively-charged particles (the valence electrons) as they evolve in the periodic electrostatic potential of the ion cores requires that they each occupy a state of motion (and total energy) in one of several distinct *energy bands*, analogous to the quantized energy states an electron may occupy in a single atom or molecule. The width of a typical energy band is on the order of a few to several electron volts (same order of magnitude as the binding energy of a valence electron in one of the atoms), and adjacent energy bands are often separated by a similar energy, although they may also overlap. Each band has enough distinct quantum states to contain twice the number of electrons as there are molecules in the macroscopic solid crystal (i.e. $\sim 10^{23}$).

Room temperature ($\approx 290\,\mathrm{K}$) corresponds to random, thermal particle energies of $\sim 1/40\,\mathrm{eV}$ (electron volt), much smaller than the width of an energy band but much larger than the energy spacing between the individual states in a band. Because electrons are subject to *Pauli Exclusion* (each electron must be in a unique, distinct quantum state), the valence electrons of all the various atoms in a solid fill the available states starting with the lowest available energy. Because room temperature corresponds to a fairly small energy, the energy of the topmost filled states is fairly well-defined and is called the electrons' *Fermi energy*. Random thermal jostling can only affect the states of individual electrons with energies near the Fermi energy, because those with much lower energies are surrounded by quantum states already occupied by other electrons, so they're stuck in their current states.

Now, one of two situations can occur for our valence electrons in a solid:

(1) The number of electrons is such that they exactly fill all the available states in some number of energy bands, and higher energy bands are completely empty.
(2) One energy band (or possibly more, if some bands overlap) is only partially filled and has many unoccupied states still available; all other bands are either completely filled or completely empty.

Electrons occupying a completely filled energy band do not participate in electrical conduction. The reason for this is that such a band corresponds to all physically possible states of individual electron motion in all directions consistent with the energies of the electrons in that band. Applying an external electric field doesn't change this situation unless the field is so intense that it can cause electrons to transition to another (partially filled or empty) energy band. Thus, no new net motion of electrons can be induced by the presence of the field, so the electrical conductivity contributed by a completely full (or, of course, completely empty) energy band is zero.

This last result implies that solids with situation (1) above are *insulators* (or maybe semiconductors). Since each band has twice the number of states as there are molecules in the crystal, insulating materials most often arise when there is an even number of valence electrons participating in the chemical bonding forming the solid. Situation (2), on the other hand, allows electrical conduction to proceed using the electrons in the partially-filled energy band. Electrons near the Fermi energy in the band have a wide selection of nearby empty states, so an applied electric field can accelerate them, and their resulting motions can carry a net flow of charge (electric current) through the solid. These materials are *conductors*, and partially-filled energy bands are characteristic of the so-called *metallic bond*.

Semiconductors have valence electrons whose situation falls into category (1): bands containing electrons are completely filled, at least at cold temperatures. What makes them different from insulators, however, is that the bottom of the nearest empty energy band (called the *conduction band*) is only about an eV or so away from the top of the highest-energy filled band (the *valence band*). Consequently, random thermal jostling of the ions in the lattice can impart enough energy to a few electrons with energies near the top of the valence band to excite them into levels near the bottom of the conduction band. In this case both the valence band and the conduction band become *partially* occupied (although just barely), and the material becomes a poor conductor (poor because only a tiny fraction of the valence electrons get bumped up into the conduction band). The higher the temperature, the greater the number of valence electrons thermally excited into the conduction band — the number goes as:

3.6

$$n_i \propto T^{3/2} e^{-E_g/(2k_B T)}$$

where E_g is the magnitude of the energy gap between the valence and conduction bands and k_B is Boltzmann's constant. In the case of silicon, this amounts to $\sim 10^9$ electrons per cm^3 at room temperature (compare with copper's $\sim 10^{23}$ per cm^3).

The archetypical semiconductors are the elements silicon ($E_g = 1.12\,\text{eV}$), and germanium ($E_g = 0.67\,\text{eV}$), each of which form a diamond lattice with four covalent bonds per atom. Carbon in its diamond form ($E_g = 5.5\,\text{eV}$) is beginning to find applications in solid-state devices, but its large energy gap makes it more properly classified as an insulator. Several compounds and alloys form commercially important semiconductors, including GaAs, InP, GaAsP, and InGaN.

Electrons and holes; impurities and doping

The diagram at right illustrates the distribution of electrons between the top of the valence band and the bottom of the conduction band for a semiconductor at a fairly high temperature (so that there have been a considerable number of electrons excited into the conduction band). The density of quantum states grows as $\sqrt{\Delta E}$ as you move away from the band edges, as shown by the right-hand curves in the figure. The Boltzmann factor $\text{Exp}(-\Delta E/k_B T)$ gives the relative probability that any one state is occupied in the conduction band or unoccupied in the valence band. Because the number densities of the conduction electrons and the holes in their respective bands are low (much smaller than the crystal's atomic number density), the charge carriers will distribute themselves in accordance with the classical Maxwell distribution, so the overall occupation densities go as the left-hand curves in the figure.

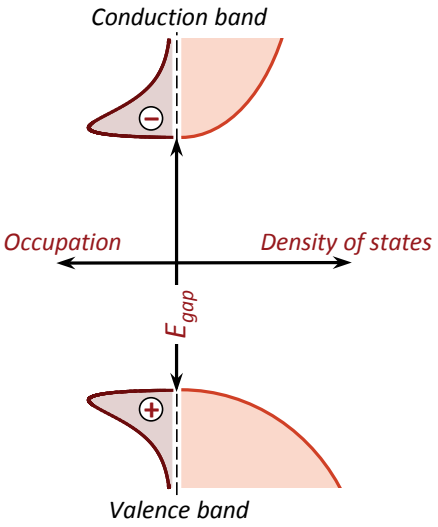


Figure 3-20: Densities of states and occupations by electrons (conduction band) and holes (valence band) for a pure semiconductor (only intrinsic charge carriers); energy increases in the vertical direction. The kinetic energy distribution of the charge carriers in each band is classical.

The dynamics of the relatively small number of electrons in the conduction band is very well approximated by treating them as classical particles (with a negative charge of $-q_e$, of course), but their *effective mass* is determined by the shape of the density of states curve near the bottom of the conduction band: the sharper the curve near the minimum, the lighter the effective mass. In the valence band, only a small fraction of the states near its top are unoccupied. Interestingly, the dynamics of the remaining electrons near the top of the valence band are such that they have a *negative effective mass*, since the density of states *decreases* with increasing energy near the top of the band.

The consequence of this unusual electron behavior near the top of the valence band is that the unoccupied quantum states evolve as though they were *positively charged particles* ($+q_e$)

with a positive effective mass and with energies increasing as they move further down from the band top! These "positive charge carriers" near the top of the valence band are called *holes*. Thus, when an electron is excited from the valence band to the conduction band, *two* charge carriers are created: the electron ($-q_e$) and the hole ($+q_e$) it left behind. Since the energy an electron must gain to cross the energy gap between valence and conduction bands is at least E_g, but two "particles" were created by this transition, *the required energy per particle* is $E_g/2$ — this explains the 2 in the Boltzmann factor in equation 3.6.

Thus a pure semiconductor has a conductivity which is a very strong function of temperature, rising rapidly as temperature increases (equation 3.6). This effect is used to make a *thermistor*: a resistor with a large, *negative temperature coefficient* (decreasing resistance as temperature rises) which acts as a very sensitive, fast-acting temperature sensor for the range of about −100°C to +150°C.

Semiconductor materials are custom-made to be much more flexible and useful through the process of *doping*: introducing various amounts of impurity atoms into the semiconductor crystal which have a different valence than the semiconductor. For example, mixing a small amount of phosphorous (valence 5) into a silicon crystal will introduce atoms each with an extra valence electron left over after it forms bonds with surrounding silicon atoms. What would be the consequences of these extra electrons to the physics of the material? It turns out that the energy of this extra valence electron is very close to the energy of the bottom of the conduction band (in the case of P in Si, the energy is only 0.044eV below the conduction band). If there are relatively few of these *donor* impurity atoms, then it is very likely that such electrons will eventually be thermally excited into the conduction band: once there they move away from the impurity atoms and are unlikely to recombine with them.

So even if the ambient temperature is cool enough that almost no electrons would be excited from the valence band to the conduction band, electrons from donor impurities will nearly all find their way into the conduction band, providing a largely temperature-independent cadre of negative charge carriers ($-q_e$) along with the same number of fixed, positively-charged ions ($+q_e$) distributed throughout the crystal lattice. Such a material is called an *N-type semiconductor*.

Similarly, introducing a valence 3 impurity atom (such as aluminum into silicon) will leave an unsatisfied bond because of the missing electron. Again, the energy required to promote a nearby valence electron into this spot is small compared to the semiconductor's energy gap (0.057eV for Al in Si). Thermal agitation will eventually do the trick, and the vacated valence state becomes a hole which quickly moves away from the impurity atom, trapping the promoted electron at the impurity site. Thus these *acceptor* impurity atoms become fixed, negatively-charged ions ($-q_e$) in the lattice, whereas an equal number of holes form a temperature-independent group of positive charge carriers ($+q_e$), creating a *P-type semi-conductor*.

Adding dopants to a semiconductor can not only introduce charge carriers (called *extrinsic* charge carriers), but will also suppress the thermal creation of electron-hole pairs described by equation 3.6 (called *intrinsic* charge carriers). This is because the product of the number of conduction electrons (n_c) and the number of holes (p_v) is related to the number of intrinsic charge carriers thermally created in a pure (undoped) semiconductor (n_i) by the laws of statistical mechanics:

3.7 $$n_c\, p_v \;=\; n_i^{\,2}$$

For example, the addition of 1 part per million phosphorous to a silicon crystal would introduce 5×10^{16} extrinsic conduction electrons per cm^3; with $n_i \sim 10^9$ electrons per cm^3, we see that there will be only $p_v \sim 100$ holes per cm^3! These holes are called *minority carriers* in the N-type silicon under discussion; the conduction electrons are the *majority carriers*. Since for this example $n_i \ll n_c$, the temperature dependence of n_c will be quite small, so, from equation 3.7, $p_v \propto n_i^{\,2}$. Thus *the temperature dependence of the minority carriers is very large*: from equation 3.7,

3.8 $$\text{<minority carrier density>} \;\propto\; T^3\, e^{-E_g/(k_B T)}$$

The equilibrium PN junction

Now consider the case of a semiconductor crystal with inhomogeneous doping. As a concrete (but quite artificial) example, assume that we take a single P-type crystal and a single N-type crystal and then join them along a planar boundary so as to form a single crystal with an abrupt change in doping at this boundary. The result is a *PN junction* (Figure 3-21) at the interface between the two semiconductor types.

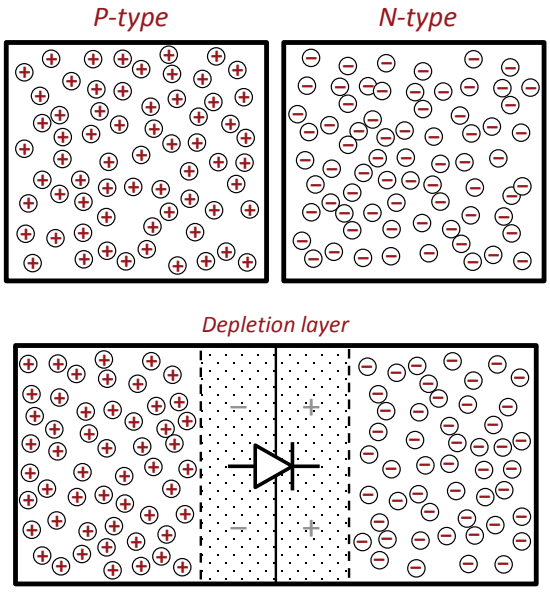


Figure 3-21: Formation of a PN junction diode and its depletion layer.

Far from the boundary the charge carrier densities must approach their homogeneous, thermal equilibrium values. Near the interface, on the other hand, the large gradients in the hole and conduction electron densities will drive diffusion of these charge carriers across the boundary, where they will eventually recombine with carriers of the opposite sign. The reduced majority carrier densities near the boundary induce a net charge density and resulting electric field near the interface because of the now unbalanced charge of the impurity ions in each semiconductor. This electric field will repel the majority charge carriers on either side of the boundary, and an

equilibrium condition is reached preventing further net diffusion of carriers across the boundary (bottom illustration in Figure 3-21). The electric field near the boundary generates a potential difference between the P-type and N-type sides of the junction, with the N-type material at the higher potential. This *contact potential* of the PN junction is very close to the *gap voltage*: $E_g/q_e \equiv V_g$ (1.12V for silicon). The result, as we shall see, is the creation of the *PN junction diode*.

It turns out that the equilibrium situation will be attained only when a region near the PN interface is almost completely devoid of charge carriers: the so-called *depletion layer*, as shown in Figure 3-21. The charge density in this region is then given by the number densities of the impurity ions on each side of the boundary, which are nearly equal to the corresponding majority carrier number densities far from the boundary. Since the potential changes by $\sim V_g$ across the depletion layer, it is straightforward to calculate its equilibrium width, which will typically be in the range of $10^2 - 10^4$ Å (about 2000 Å for silicon with a part per million doping), and the magnitude of the electric field at the interface is in the range of $10^5 - 10^7$ V/m.

The PN junction I-V characteristic curve

The equilibrium configuration (Figure 3-21) is maintained when the rate that the holes diffuse into the depletion layer from the P-type side (against the contact potential gradient) matches the small rate of hole diffusion from the N-type side (where holes are the minority carriers), so that the net flow of holes across the junction is 0; a similar condition holds for the conduction electron diffusion at the junction.

Holes entering the depletion layer from the N-type side are not impeded by the presence of the contact potential — on the contrary, the electric field in the depletion layer will accelerate them through it to the P-type side. This implies that the rate of the minority hole diffusion will simply be proportional to the hole density on the N-type side, which is given in equation 3.8 to be proportional to $e^{-q_e V_g/(k_B T)}$, and similarly for the minority electron diffusion from the P-type side. The majority carriers must cross the barrier imposed by the junction potential (V_j), so only those carriers with kinetic energies larger than $q_e V_j$ can cross to the other side; the number of such energetic carriers will be proportional to the Boltzmann factor $e^{-q_e V_j/(k_B T)}$ (because their kinetic energy distributions are classical, as mentioned before). At equilibrium, these two rates match, and V_j is equal to the junction contact potential, V_g.

When an external bias voltage V is applied across the PN junction, this applied potential will reduce the junction potential to $V_j = V_g - V$ ($V > 0$ is forward-biased). As a consequence, more majority charge carriers will have enough energy to diffuse through the depletion layer; the minority diffusion rate from the other side is unaffected. Thus there will be a net current flow across the junction given by the difference in these two diffusion rates:

$$I \propto e^{-q_e(V_g - V)/k_B T} - e^{-q_e V_g/k_B T} = e^{-q_e V_g/k_B T}\left(e^{q_e V/k_B T} - 1\right)$$

This simple result is known as the *ideal diode equation*:

$$I = I_R(e^{q_e V/k_B T} - 1); \quad I_R = I_0 e^{-q_e V_g/k_B T}$$

V is the applied bias voltage (+ for forward-bias), V_g is the semiconductor's gap voltage, I_0 is some constant, and I_R is the diode's reverse leakage current. Thus, the ideal diode's forward current rises exponentially with forward bias voltage (for voltages of more than a few tens of millivolts), and has some small, temperature-dependent leakage current when reverse-biased.

The above equation is not quite right, because its derivation ignores an effect which is especially important for the behavior of a silicon diode: generation and recombination of charge carrier pairs in the depletion layer. The assumption in the argument leading up to the diode equation was that the only charge carriers present in the depletion layer entered it through diffusion from the regions outside the layer, and that all of these carriers pass through the depletion layer. Actually, thermal excitation of electron-hole pairs will occur in the depletion layer, just as it would in a pure semiconductor; similarly, recombination of electrons and holes may also occur among those which diffuse into the depletion layer, so the number of charge carriers entering the depletion layer is larger than the number which escape, especially for small forward bias voltages in relatively large V_g diodes such a silicon.

The depletion layer generation and recombination processes depend exponentially on temperature, but the exponent goes as $q_e/2k_B T$ rather than as $q_e/k_B T$. The combination of this process with the ideal diode process leads to a "slight" modification of the ideal diode equation:

Diode equation

3.9

$$\boxed{I = I_R(e^{q_e V/\eta k_B T} - 1); \quad I_R = I_0 e^{-q_e V_g/\eta k_B T}}$$

The coefficient η depends on the importance of the depletion layer recombination process; it is a weak function of I and T and ranges between 1 and 2. For the small-signal silicon diodes you will use, $\eta \approx 1.9$ and $I_R \approx 5\,\text{nA}$; $q_e/\eta k_B T \approx 20\,\text{volt}^{-1}$ at 20°C. The exponential dependence of I on forward-bias voltage V (for $I \gg I_R$) is the basis for the exponential and logarithmic amplifiers presented earlier; unfortunately, this relationship is strongly temperature-dependent. When the junction is reverse-biased, the depletion layer generation rate depends on the size of the depletion layer, which grows as $\sqrt{1+V/V_g}$ (V is the reverse-bias voltage), so the reverse current does not "saturate" at the I_R value given by equation 3.9, but continues to grow slowly with increasing reverse-bias voltage as long as it remains well below the diode's breakdown voltage. Cartoons of these diffusion processes are shown in Figure 3-22; plots of the 1N4148 silicon small-signal diode I-V characteristic curves for both forward and reverse bias and at two temperatures are provided in Figure 3-23.

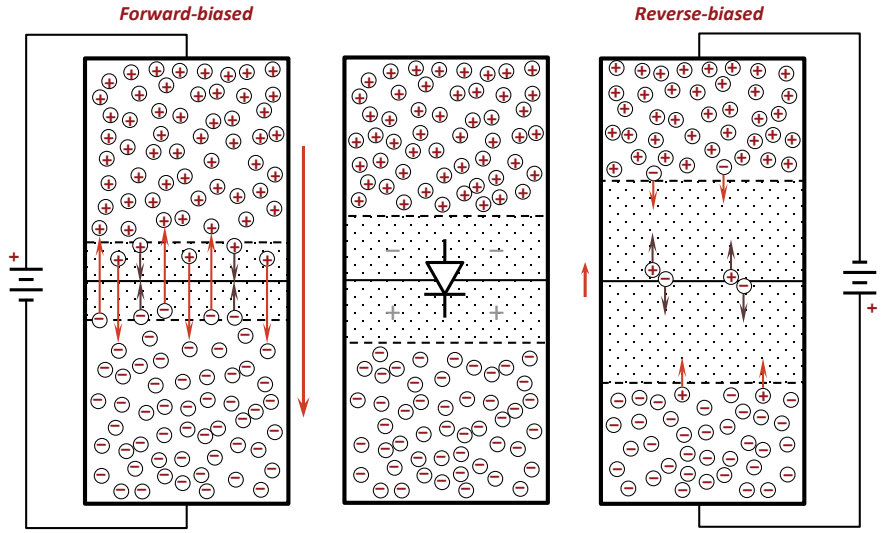


Figure 3-22: Depletion layer width and charge carrier diffusion of a PN junction as affected by applied bias voltage. Forward-bias (left) reduces the height of the potential barrier to majority carrier diffusion and decreases the depletion layer width, so many majority carriers can diffuse into and through the depletion layer; minority carrier diffusion is largely unaffected. Carriers that cross the depletion layer and recombine with majority carriers on the opposite side are indicated by the orange arrows; those that recombine inside the depletion layer are shown with gray arrows. Reverse-bias current is completely dominated by minority carrier diffusion: those that enter the depletion layer from the bulk semiconductor (orange arrows) and those pairs that are thermally generated within the depletion layer (gray arrows).

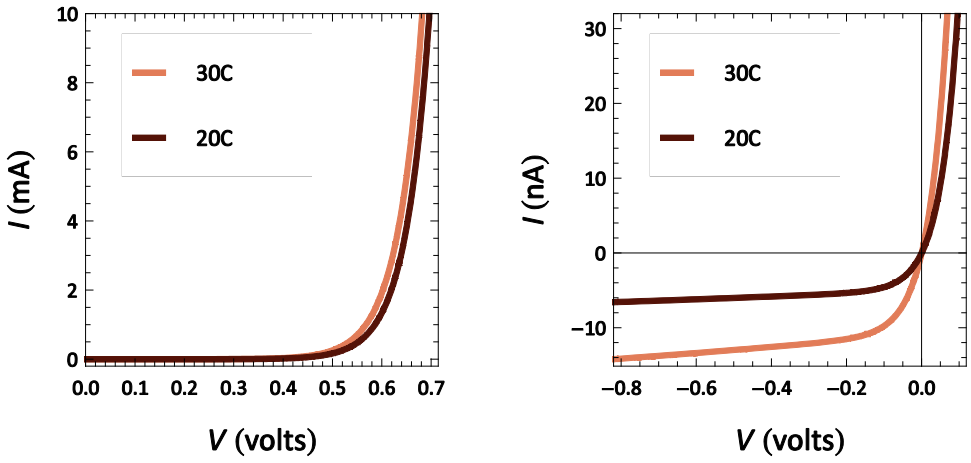


Figure 3-23: Forward-bias (left) and reverse-bias (right) I-V characteristic curves for the 1N4148 silicon diode. The forward-bias curves are exponential (equation 3.9), but they appear as though the diode suddenly "turns on" at a forward voltage of 0.6–0.7V; this voltage decreases slowly with rising diode temperature, as shown. The reverse-bias current is a much more sensitive function of temperature, however, approximately doubling for the same 10°C temperature increase. Note the different vertical scales for the two plots (a factor of 10^6).

Zener and avalanche breakdown

As the reverse-bias voltage on a PN junction is increased, the intensity of the electric field in the depletion layer rises; it is particularly intense at the interface between the P- and N-type areas. Minority carriers entering the depletion layer are accelerated by the field; when their kinetic energies reach a few eV or so, collisions with atoms in the lattice may knock valence electrons out of them, creating additional electron-hole pairs. These newly-created charge carriers are also accelerated by the field and can create even more carriers as they collide with lattice atoms.

At sufficiently high reverse voltages this collision-induced ionization process may lead to an *avalanche* of additional charge carriers, and the reverse current will grow exponentially with increasing voltage beyond some reverse-bias threshold. This is the *avalanche breakdown* process, and the reverse-bias voltage threshold for its action is the diode's *reverse breakdown voltage*. The electric field intensity for any particular applied reverse-bias voltage depends on the impurity concentrations and the abruptness with which these concentrations change near the P-type and N-type interface, so a target reverse breakdown voltage may be engineered into a particular diode type.

Another effect of a very intense electric field in the depletion layer is the large electric polarization of the atoms in the lattice it induces: at field strengths $\gtrsim 10^6\ \mathrm{V/m}$ the potential difference across a distance of about 100Å can exceed the semiconductor's gap voltage. In this case a valence electron may quantum mechanically *tunnel* across this distance into the conduction band, creating an electron-hole pair; because of this tunneling process the electric field required to ionize a lattice atom is much smaller than it would need to be to ionize a single, independent atom ($\sim 10^{10}$–$10^{11}\ \mathrm{V/m}$); this effect was first theorized by the American physicist Clarence Zener in 1934. The tunneling rate grows exponentially as the required tunneling distance (inversely proportional to electric field strength) decreases, again leading to a large increase in reverse current (breakdown) as applied reverse-bias voltage exceeds the tunneling threshold. The target reverse breakdown voltage may be engineered by adjusting a diode's impurity concentration and doping profile. See Figure 3-17 on page 3-20 for a typical reverse breakdown I-V characteristic.

Diodes with reverse-breakdown voltages exceeding 6V or so are dominated by the avalanche breakdown process; those below 5V are predominantly subject to Zener breakdown. Regardless of breakdown voltage, those diodes designed to be used as voltage regulators with precisely-tailored reverse breakdown voltages are collectively called *Zener diodes*; those with high current-handling capacity and extremely fast response to voltages exceeding their breakdown threshold are usually called *avalanche diodes* and are primarily used for *transient voltage suppression* (TVS) and overvoltage protection.

Experiment 4

Comparators, positive feedback, and relaxation oscillators

Copyright © Frank Rice 2013, 2014
Pasadena, CA, USA
All rights reserved.

Experiment 4

Comparators, positive feedback, and relaxation oscillators

This experiment will continue our investigations of nonlinear analog circuits. We consider first a simple op-amp application used to *interface* an analog signal to a digital device: the *Schmitt trigger*, a 1-bit *analog to digital converter* (in which the op-amp is used as a *comparator*). This circuit introduces us to the use of *positive feedback* in our op-amp designs, rather than the negative feedback we've used so far. In this case the positive feedback is used both to introduce *hysteresis* in the circuit's *state transition* trigger conditions and to speed up the op-amp's output state transitions.

Next we couple a Schmitt trigger with first an *RC* low-pass filter and then an op-amp integrator circuit to develop *relaxation oscillators*, simple signal generators which work much like a ticking clock to output a repetitive waveform. Spend some time studying this relaxation oscillator idea, because its feedback scheme is applicable to many types of simple analog signal generators, clocks, and timers (some of which could more correctly be considered to be simple digital circuits).

Finally, we introduce a special-purpose integrated circuit, the "555 timer," a versatile device we will use to build *astable* and *monostable multivibrator* circuits useful for a variety of applications. This device is our first true example of a *mixed-signal circuit* incorporating both analog and digital design concepts.

THE SCHMITT TRIGGER AND POSITIVE FEEDBACK

The op-amp as a "comparator"

Consider an op-amp used to amplify an input signal *without feedback* as shown in Figure 4-1. Because no feedback is used, the input signal is amplified by the op-amp's full open-loop gain, so even a tiny input voltage (on the order of a millivolt or less) will be enough to send the op-amp's output into saturation, as shown in the plots of v_{in} and v_{out}. Thus, in this case (since the op-amp's *+Input* is grounded), the output gives $-1 \times$ the sign of v_{in}, and the circuit is a one-bit *analog to digital converter* (*ADC*), also called a *comparator*.

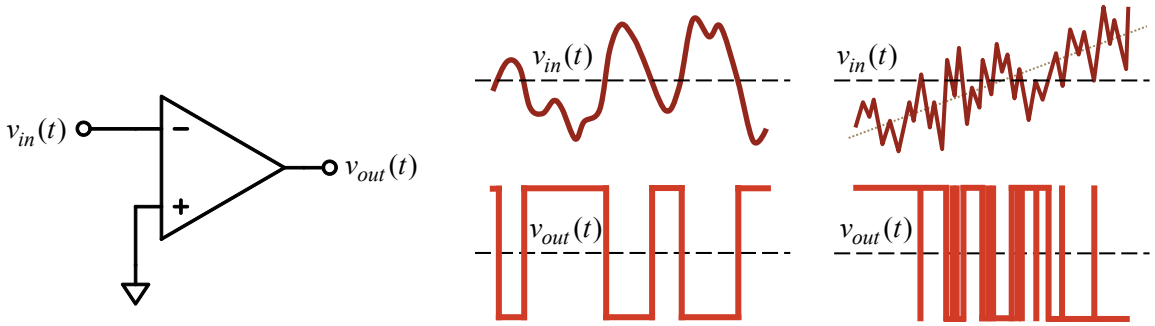


Figure 4-1: An op-amp used as a *comparator*. Whenever $v_{in} > 0$, the op-amp output v_{out} will go to its negative limit (saturation); when $v_{in} < 0$, v_{out} will go to its positive limit. This is an *inverting* comparator, since v_{in} is connected to the op-amp's *–Input*. One potential problem, however: a slowly rising, noisy input can cause many closely-spaced output transitions as it passes through 0, as shown by the right-hand graphs. This is undesirable if you need to use the circuit to accurately count 0-crossings of the underlying input signal.

Comparator-type circuits are useful in a variety of situations. For example, consider the circuit at right, where instead of grounding the *+Input*, it is connected to a constant *threshold* voltage source, V_{th} (shown here as a battery, but it could come from a user potentiometer setting or other voltage divider using the circuit power supplies, etc.). Whenever $v_{in} > V_{th}$, the op-amp output goes into negative saturation and the LED is illuminated; otherwise, the op-amp output is at positive saturation and the LED is off. Such a circuit could, for example, warn an operator of excessive temperature or pressure if v_{in} is generated by an appropriate sensor.

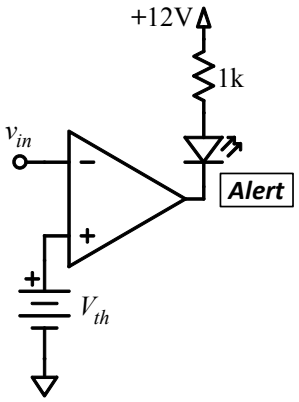


Figure 4-2: Inverting comparator used to illuminate a warning LED whenever $v_{in} > V_{th}$.

Another application could be to interface the comparator circuit output to a digital system in order to count zero crossings of the input signal to calculate its frequency or to count the number of events detected by a sensor. Unfortunately, if the input signal v_{in} rises through the threshold voltage slowly, but there is a significant amount of noise in the signal, many output transitions could

be generated by the noise while v_{in} is near V_{th}, as shown in the right-hand graphs in Figure 4-1. Such behavior would render the circuit useless for a counting application.

Using positive feedback to add hysteresis*: the Schmitt trigger*

A common solution to the problem just outlined is to add *noise immunity* to the comparator circuit by incorporating *hysteresis* into the transition threshold voltage V_{th}, as shown in Figure 4-3.

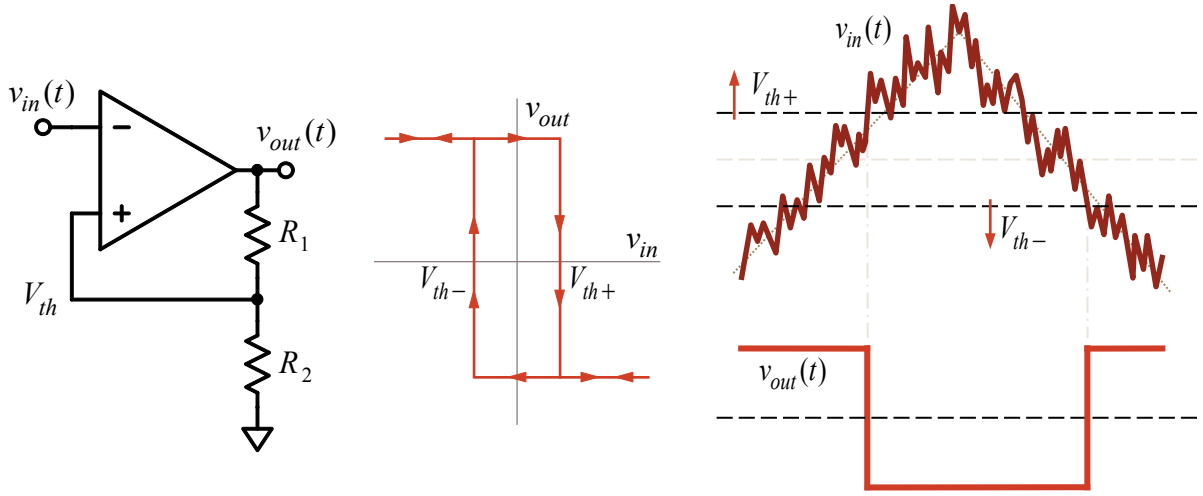


Figure 4-3: An inverting *Schmitt trigger* circuit. Positive feedback is used to add *hysteresis* to the transition threshold voltage. When the output is high the threshold voltage $V_{th+} > 0$, but when the output is low then $V_{th-} < 0$. If $(V_{th+} - V_{th-}) >$ [noise peak-peak amplitude], then the noise cannot trigger unwanted transitions in the output as v_{in} slowly passes through 0. The center plot shows the *hysteresis loop* defined by V_{th+} and V_{th-}; the right-hand plot shows how these differing thresholds provide some level of *noise immunity*. Of course, the underlying input signal variations must cross the V_{th+} and V_{th-} thresholds in order to generate changes in the circuit output.

By "hysteresis" we mean that the threshold voltage is a function of the system's current *operating state*, which is defined for this circuit by its output voltage: positive or negative saturation. Because V_{th} is determined by the voltage divider constructed from resistors R_1 and R_2, it changes in response to a change in the output voltage: once the output has gone high in response to an input which has passed below the threshold voltage, the threshold voltage is changed to a higher value (V_{th+}); conversely, an input voltage climbing through V_{th+} will change the output to its low state and cause the threshold voltage to be set to a lower value (V_{th-}), as illustrated in Figure 4-3.

As shown in the right-hand graphs in the figure, this difference in V_{th+} and V_{th-} means that once a transition is triggered by a change in v_{in}, small noise excursions in the input will not cause v_{in} to reverse its course enough to cross the *hysteresis gap* $(V_{th+} - V_{th-})$ and cause an undesired reversal of the output state. If the hysteresis gap is made large enough, then the system can be made completely impervious to the noise in the input signal, eliminating the spurious output transitions suffered by the basic comparator circuit (Figure 4-1).

There is another important advantage to the use of positive feedback in the comparator circuit (Figure 4-3): as the output changes, the feedback increases the difference between the op-amp input voltages, accelerating the change in the output even if the op-amp open-loop gain is relatively modest. Thus, because of the positive feedback, *the output voltage will change at an exponentially increasing rate* until the op-amp slew rate limit is reached, even if the initial difference between v_{in} and V_{th} is very small, or v_{in} is changing very slowly. This "pulling oneself up by one's own bootstraps" effect is why positive feedback is also referred to as *regenerative feedback*. This idea of using regenerative feedback to incorporate noise immunity and to vastly increase output transition (*switching*) speed was first developed by Otto Schmitt at Washington University (St. Louis, Missouri) in 1934; a circuit incorporating these two features (threshold hysteresis and positive feedback) is called a *Schmitt trigger*.

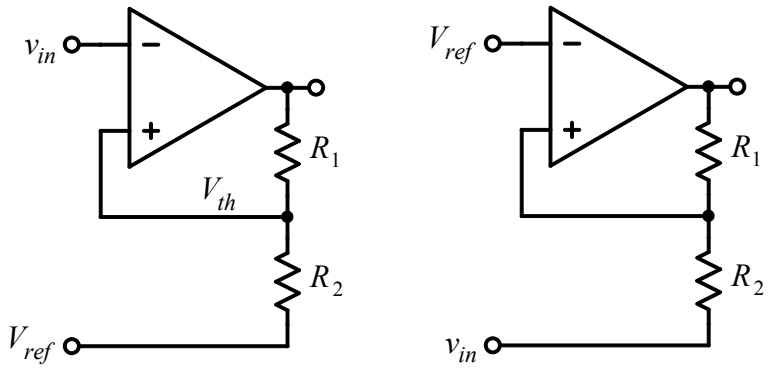
Schmitt trigger circuit variations and trigger point calculations

Call the op-amp positive and negative output saturation voltages V_{sat+} and V_{sat-}; the resulting hysteresis gap for the circuit of Figure 4-3 is:

4.1

$$V_{th+} - V_{th-} = \frac{R_2}{R_1 + R_2}\left(V_{sat+} - V_{sat-}\right) \qquad \textbf{(inverting)}$$

For the TL082 with ±12V power supplies, $V_{sat+} - V_{sat-} \approx 21\text{–}22\text{V}$. Because the other end of the voltage divider (bottom of R_2) is connected to ground, the threshold voltages V_{th+} and V_{th-} will be centered around 0V (assuming that $V_{sat-} = -V_{sat+}$). Connecting the bottom of R_2 to a voltage reference source rather than to ground *will not affect the hysteresis gap*, but it will center that gap around a nonzero mean threshold proportional to the reference V_{ref} (see Figure 4-4 and equation 4.2 on page 4-5).



Inverting *Noninverting*

Figure 4-4: Inverting and noninverting Schmitt triggers with a supplied reference voltage V_{ref} used to set the trigger thresholds. Note that the voltage source connected to the bottom of R_2 in each circuit must have an output impedance «R_2 or the trigger points will be affected because of the current flowing between the op-amp's output and that source.

4.2 $$\overline{V_{th}} = \tfrac{1}{2}\left(V_{th+} + V_{th-}\right) = \frac{R_1}{R_1 + R_2} V_{ref}$$ **(inverting)**

Note that a noninverting Schmitt trigger may be implemented by simply swapping the input and reference voltage connections, but now the trigger points are different from those for the inverting case, because now the voltage divider affects the input voltage rather than the reference.

4.3 $$V_{th+} - V_{th-} = \frac{R_2}{R_1}\left(V_{sat+} - V_{sat-}\right)$$ **(noninverting)**

4.4 $$\overline{V_{th}} = \tfrac{1}{2}\left(V_{th+} + V_{th-}\right) = \frac{R_1 + R_2}{R_1} V_{ref}$$ **(noninverting)**

Additional Schmitt trigger circuit design considerations

Note that equation 4.3 places an important restriction on the ratio R_2/R_1 for a noninverting Schmitt trigger: unless $R_2 < R_1$, the hysteresis gap $(V_{th+} - V_{th-})$ will exceed the output voltage swing range of the op-amp $(V_{sat+} - V_{sat-})$, and, depending on the reference voltage value V_{ref}, *one or both of the Schmitt trigger thresholds will be beyond the range of the op-amp output voltage*. Assuming the input signal voltage range is also limited to $V_{sat-} \le V_{in} \le V_{sat+}$, then the circuit's output could experience *lock-up* at V_{sat+} or V_{sat-}, rendering the circuit useless!

For either circuit in Figure 4-4, it is important to remember that the voltage source connected to R_2 must have a small output impedance, or its output impedance must be added to R_2 when calculating the trigger thresholds using equations 4.1 through 4.4. If necessary, use a voltage follower between the voltage input and R_2.

Another design consideration is the current required from the op-amp output to drive the voltage divider formed by resistors R_1 and R_2. If, say, both are chosen to be 1k, then their series resistance is 2k, and when the op-amp output is at saturation (approx. 11V), then over 5mA will flow through the resistors. This relatively large current draw will probably reduce the TL082 op-amp's saturation voltages. As the total resistance $R_2 + R_1$ is reduced, then the additional current drawn by them may cause significant changes in the op-amp's behavior, and the circuit will not work as you expect.

THE RELAXATION OSCILLATOR

Simple, one op-amp oscillator

If you feed the output of a Schmitt trigger back to its inverting input through a *RC* low-pass filter, you get a circuit whose output switches back and forth between the op-amp's two saturation limits: you have made a simple *relaxation oscillator* (Figure 4-5).

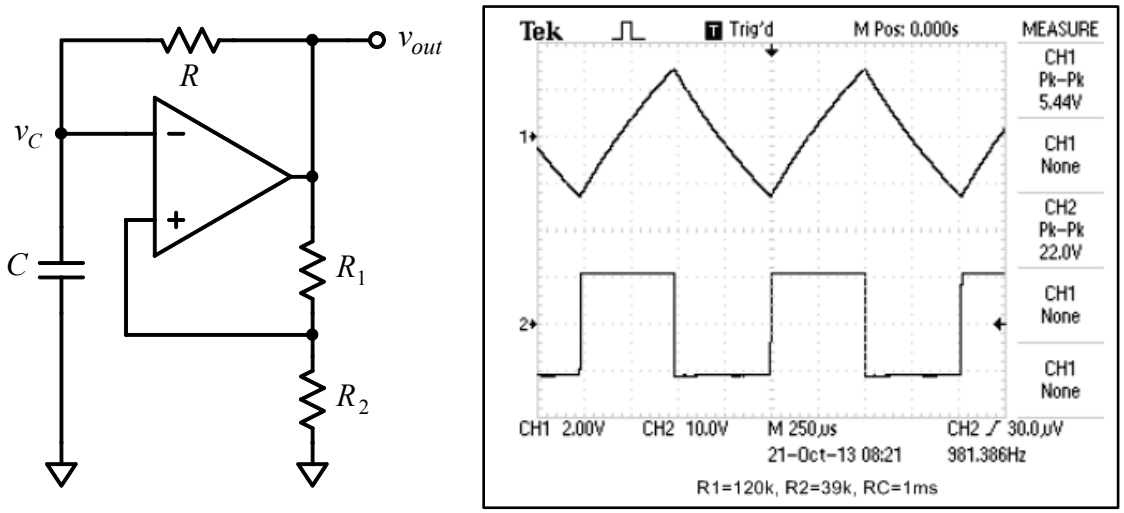


Figure 4-5: A simple *relaxation oscillator* using a Schmitt trigger to alternately charge and discharge the capacitor *C* through the resistor *R*. Whenever v_C, the voltage across *C*, reaches a trigger threshold, the op-amp output voltage reverses to its opposite saturation limit. Thus the current through *R* changes sign, and the capacitor voltage moves toward the opposite threshold. Consequently, v_C oscillates between the Schmitt trigger's two threshold voltages as the op-amp output switches back and forth between its two output saturation limits. The oscilloscope image shows v_C (CH1) and v_{out} (CH2) for trigger thresholds chosen so that $f = 1/RC$.

As should be clear from the figure, the op-amp's output charges the capacitor *C* via the resistor *R*. Because the capacitor's voltage is monitored by the op-amp's inverting input, every time it charges up to a trigger threshold, the op-amp output changes sign, and the capacitor voltage then begins to "relax" toward the opposite output saturation limit. The trigger threshold voltage at the op-amp's *+Input* has also changed sign, however, so that the op-amp output again changes state as the capacitor voltage reaches this opposite threshold; the process is then repeated.

The capacitor's voltage profile is an exponential *relaxation* toward an equilibrium voltage which will equal to the op-amp's output saturation voltage, V_{sat}, starting from the opposite trigger threshold voltage. If the + and – saturation voltages are assumed to be equal, then this exponential relaxation is described by:

4.5
$$v_C(t) = V_{sat} - (V_{sat} + V_{th})e^{-t/RC}$$

If the oscillation period is T, then after half a period the capacitor voltage reaches the next trigger threshold, so in equation 4.5 $v_C(T/2) = V_{th}$. With equation 4.1 relating V_{sat} and V_{th}, the relationship between the period T and the circuit's component values is:

$$\frac{R_1}{R_2} = \coth\left(\frac{T}{4RC}\right) - 1 \qquad \textbf{4.6}$$

where **coth** is the hyperbolic cotangent function. If you want the oscillator period to equal the filter's RC time constant $(T = RC)$, then

$$R_1 = 3.08R_2$$

$R_1 = 120\text{k}\Omega$, $R_2 = 39\text{k}\Omega$ provide a pair of standard resistor values which closely matches this ratio (within 0.2%).

Note that the current drawn by the *RC* feedback pair is as high as $(V_{sat} + V_{th})/R$ just after the op-amp output changes state — excessive current here will reduce V_{sat} as the op-amp tries to meet this output current requirement, distorting the output waveforms and lengthening *T*. Choosing $R \geq 10\text{k}\Omega$ should limit the capacitor charging current to a reasonable level.

Function generator

Using an op-amp integrator circuit rather than a simple RC pair would charge the capacitor as a constant rate, so the exponential relaxation in the last circuit would be replaced by a nice, linear ramp. The resulting circuit is shown in Figure 4-6; as mentioned in the caption, we must now use the noninverting form of the Schmitt trigger because the integrator is inverting.

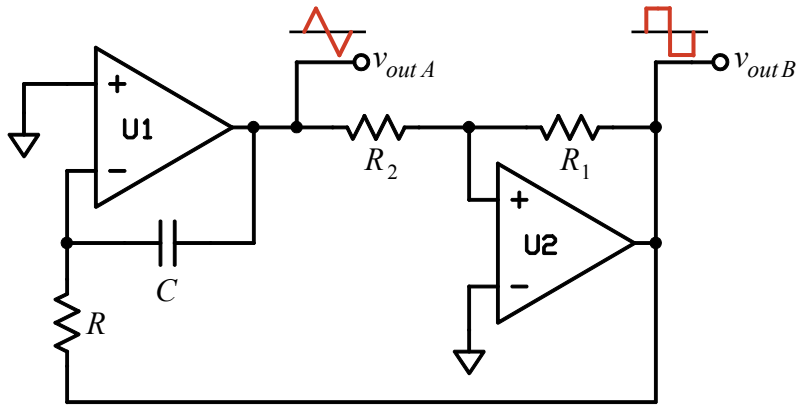


Figure 4-6: Integrator coupled to a Schmitt trigger to form a primitive *function generator*, outputting triangle and square waveforms. Since the integrator is inverting, its output must go to a noninverting Schmitt trigger, so that reaching a trigger point reverses the capacitor's charging current. This means that we must have $R_1 > R_2$, or the circuit won't work.

Since the input to the integrator is constant between triggers, its output will have a constant slope between triggers. For this reason the period of the output signals is much easier to calculate for this circuit; the formula is left to the exercises. To make the frequency variable, resistor R may be made variable; a switch could also be used to select one among a set of capacitors.

Below is a circuit which incorporates both variable frequency and symmetry adjustments of the output waveforms. Note how the diodes select which side of the symmetry potentiometer is used to set the current through the integrator's capacitor (depending on the sign of the voltage follower's output). The voltage follower (U3) isolates the Schmitt trigger's square wave output and the frequency adjust potentiometer from the current load required by the integrator, so changing the symmetry potentiometer setting will not affect the voltage divider ratio set by the frequency potentiometer or op-amp U2's output saturation voltages.

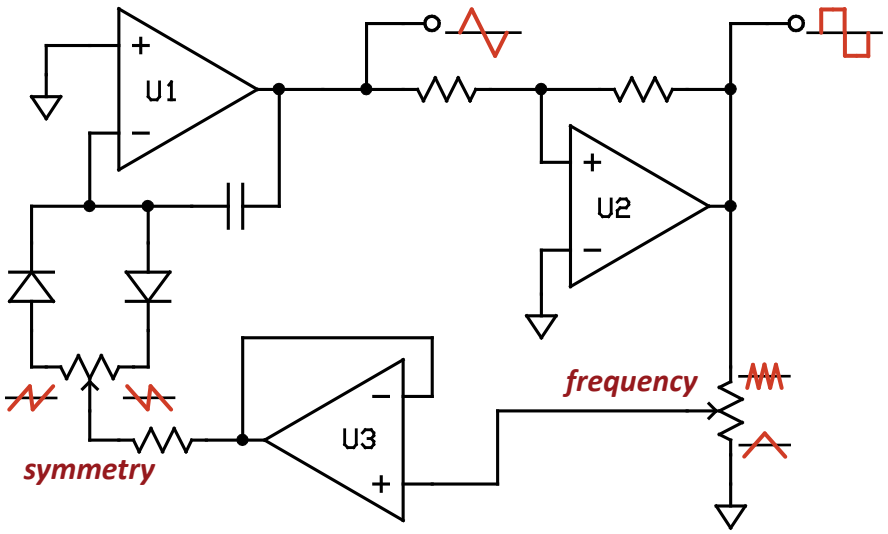


Figure 4-7: A function generator circuit with variable frequency and waveform symmetry. The voltage follower using op-amp U3 isolates the Schmitt trigger output and the frequency adjust potentiometer from the current load demands of the integrator, especially important when the symmetry potentiometer is set near one of its limits.

Figure 4-7 presents one of the most complicated circuits we've considered so far. You should spend some time studying this circuit so that you understand how it works and how you would select values for the components (the prelab exercises will help you focus on this task!). Why is the resistor in series with the output of op-amp U3 necessary?

Description

Next we consider a special integrated circuit designed specifically for timing and oscillator applications: the *555 timer IC*, originally invented in 1971 by engineers at *Signetics* (since absorbed into *NXP Semiconductors*). The version you will use for this experiment is the TLC555, an updated version manufactured by Texas Instruments. They and other companies also manufacture copies of the original version: for example, the LM555. This latter data sheet gives a few examples of the sorts of circuits you can build using this versatile device; a more thorough discussion of the device and its applications is provided in the original manufacturer's Application Note.

The 555 timer is an example of a *mixed signal* or *interface* IC, incorporating both analog and digital circuitry; we'll consider such circuits in more detail in a later experiment. Figure 4-8 shows the functional block diagram and the device *pinout* for the timer, which at first glance seems very complicated.

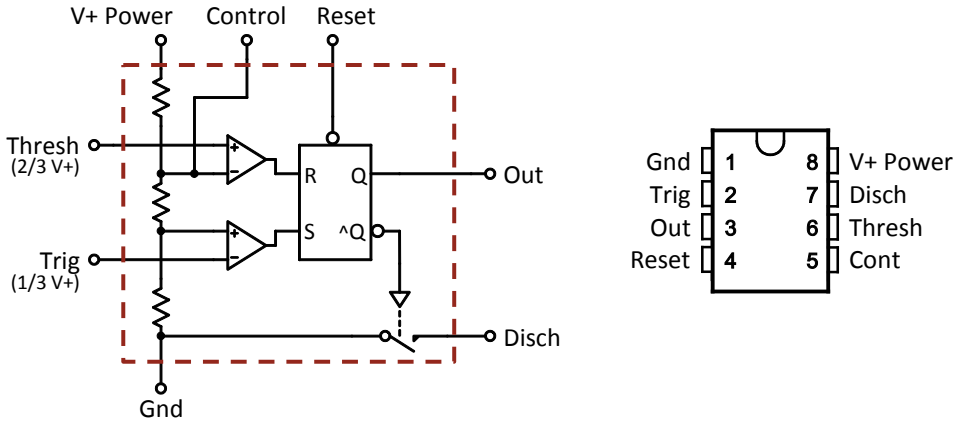


Figure 4-8: The *555 Timer* IC. Shown are its functional block diagram and the device *pinout*, or pin numbering scheme and identification. The heart of the circuit is an *RS flip-flop*; its operating state determines the outputs *Out* and *Disch*. The output (*Out*) is *LOW* and the discharge terminal (*Disch*) is shorted to ground (*Gnd*) whenever the flip-flop is in its *Reset* state; when the flip-flop is in its *Set* state, *Out* is *HIGH* and *Disch* is open (high impedance). The input voltages *Thresh* and *Trig* are used to trigger two analog comparators which control the flip-flop's state.

An *RS flip-flop* inside the 555 timer controls the device's two outputs: *Output* and *Discharge*. A *flip-flop* is the generic term for a two-state digital circuit which changes its operating state only when some particular sequence of its input signals is encountered; otherwise it remains in its current state — in other words, a flip-flop is an elementary, 1-bit *memory*. In this case, the flip-flop has two primary inputs: *Reset* (R) and *Set* (S). The inactive state for an input is *Low* (ground), whereas a *High* input (near the V+ power supply voltage) commands the flip-flop to its corresponding operating state: *Set* (Q = *High*; ^Q = *Low*) or *Reset* (Q = *Low*; ^Q = *High*). If both the R and S inputs are *High* concurrently, then the 555 gives priority to the S

input, driving the flip-flop to its *Set* state. The separate 555 *Reset* terminal input overrides any other command to its internal flip-flop and *clears* the flip-flop: drives it to the *Reset* state (the little circle on the wire from the *Reset* input at the top of the flip-flop means that it is active when *Low*: a 0V input on *Reset* commands the flip-flop to *clear*; otherwise the *Reset* pin should be connected to *V+ Power* to inactivate it.).

The operating state of the RS flip-flop determines the condition of the 555 terminals *Output* and *Discharge*. The *Output* terminal reflects the flip-flop's Q output: nearly equal to *V+ Power* when the flip-flop is *Set*, nearly equal to ground when the flip-flop is *Reset*. The *Discharge* terminal is connected via an *analog switch* to the ground terminal: When the flip-flop is *Set*, the switch is *open*, so the *Discharge* terminal is disconnected from ground; when the flip-flop is *Reset*, the switch closes, and the *Discharge* terminal is shorted to ground.

As shown in Figure 4-8 on page 4-9, the flip-flop R and S inputs are supplied by two comparators monitoring analog voltages on the 555's *Trigger* and *Threshold* inputs; the comparator reference voltages for these inputs are 1/3 and 2/3 of the power supply voltage applied to the *V+ Power* terminal. The following table itemizes the possible input combinations and how they affect the 555 output terminals.

Table 4-1
555 Timer State Table

Reset (pin 4)	Trigger (pin 2)	Threshold (pin 6)	Output (pin 3)	Discharge (pin 7)
LOW (< 0.4V)	-	-	Low (near Gnd)	Short to Gnd (on)
high	***LOW (< 1/3 V+)***	-	***HIGH (near V+)***	***OPEN (off)***
high	high	low	no change	no change
high	high	***HIGH (> 2/3 V+)***	Low (near Gnd)	Short to Gnd (on)

The ***ACTIVE*** state of each input is highlighted with italics, as shown. The *Reset* input (active when *Low*) overrides all other inputs; otherwise *Trigger* overrides *Threshold* when determining the flip-flop state. The active response state is *Output* ***HIGH***, *Discharge* ***OPEN***. Connect *Reset* to *V+* if it is not used.

The normal sequence of events when using the 555 is as follows (the *Reset* pin is kept *high*): the *Trigger* pin is brought *low*, setting *Output* to *high* and *Discharge* to *open.* Next, with the *Threshold* pin *low*, the *Trigger* is brought back to *high*; the outputs remain unchanged. After some time, the *Threshold* is brought *high*, which resets the *Output* to *low* and shorts *Discharge* to ground. Finally, *Threshold* is brought back *low*, returning the system to its initial state. Let's now see how to use this event sequence to do something interesting…

Astable multivibrator (relaxation oscillator)

The first application of the 555 IC we consider, Figure 4-9 on page 4-11, is as an *astable multivibrator* (which is the name used for a relaxation oscillator by digital electronics

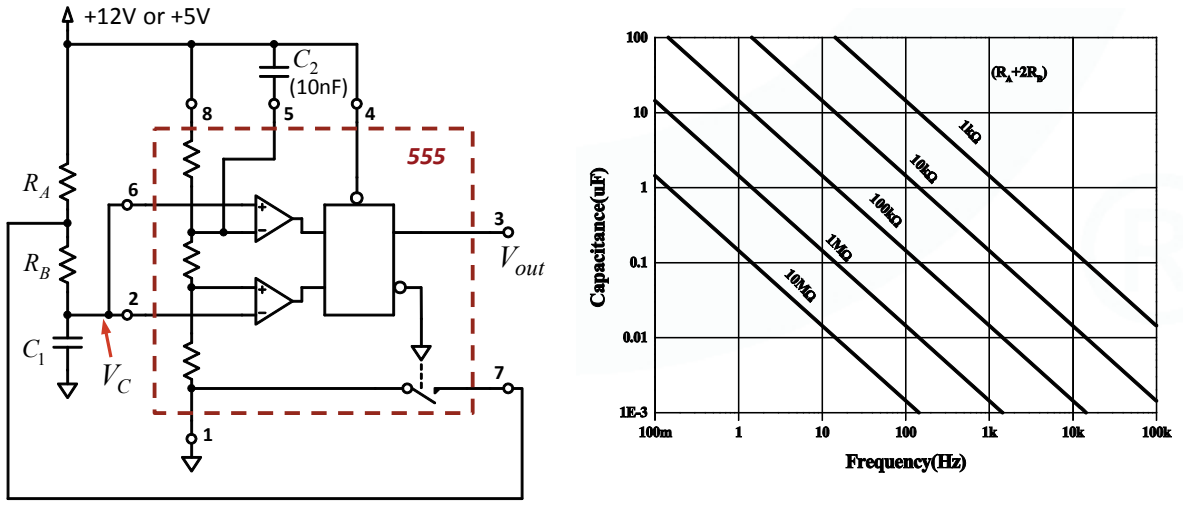


Figure 4-9: *Astable multivibrator* using the 555. The V_C and V_{out} waveforms are similar to those for the simple relaxation oscillator in Figure 4-5. Since the capacitor C_1 charges through R_A+R_B, but discharges through only R_B, the output waveform is not symmetrical. The graph at right is from the LM555 datasheet; it shows how resistor and capacitor selection affects the output frequency. IC pin numbers are shown next to the 555 terminals in the schematic.

engineers). The idea is to repeatedly charge and discharge a capacitor while monitoring its voltage using the 555's *Trigger* and *Threshold* inputs. The capacitor will be charged using the system power supply voltage (V_{supply}); when the capacitor voltage reaches $2/3\,V_{supply}$, the *Threshold* comparator will short the *Discharge* terminal to ground, and this can be used to discharge the capacitor. As it discharges to $1/3\,V_{supply}$, the 555's *Trigger* comparator opens the *Discharge* terminal connection, and the capacitor again starts to charge, repeating the process. Thus, the circuit's operation is similar to that of the simple Schmitt trigger relaxation oscillator (Figure 4-5).

Consider Figure 4-9; when the *Discharge* terminal is open (and the *Output* is high), the capacitor C_1 charges from the power supply through resistors R_A and R_B; Once V_C reaches $2/3\,V_{supply}$, the 555's flip-flop changes state, and the *Discharge* terminal is shorted to ground (and the *Output* goes low). Now C_1 discharges through R_B until V_C has dropped to 1/3 of the supply voltage. The flip-flop again changes state, the *Discharge* terminal is returned to its high impedance (open) state, and the capacitor again begins to charge.

Thus the capacitor voltage V_C relaxes back and forth between 1/3 and 2/3 of the supply voltage with time constants $(R_A + R_B)C_1$ and $R_B C_1$, so the oscillator period will be proportional to $(R_A + 2R_B)C_1$, as shown in the graph accompanying the schematic in Figure 4-9. Note that when the *Discharge* terminal is shorted to ground, the full supply voltage is applied across R_A, and this current will add to the capacitor discharge current flowing into the *Discharge* terminal. Clearly, the values of R_A and R_B should be large enough (at least a few kΩ) to keep these currents from becoming excessive.

The TLC555 has very high impedance inputs for its *Trigger* and *Threshold* terminals, so large resistor values may be used to achieve very long oscillator periods; only about 10pA is drawn by either input terminal, and the leakage current into the *Discharge* terminal is only about 100pA when it is open. Oscillator periods of a few hours or longer are easily achievable.

The 555's *Reset* and *Control* terminals aren't needed for this application; the *Reset* pin should be tied to the 555's *V+ Power* (IC pin 8) so that noise will not cause spurious resets of the 555 state; similarly, the *Control* terminal should be connected through a small capacitor $C_2 \sim 10\text{nF}$ to either ground (IC pin 1) or *V+ Power* (IC pin 8) to keep noise from affecting the comparators' trigger points (as in Figure 4-9 and Figure 4-10).

Monostable Multivibrator (one-shot)

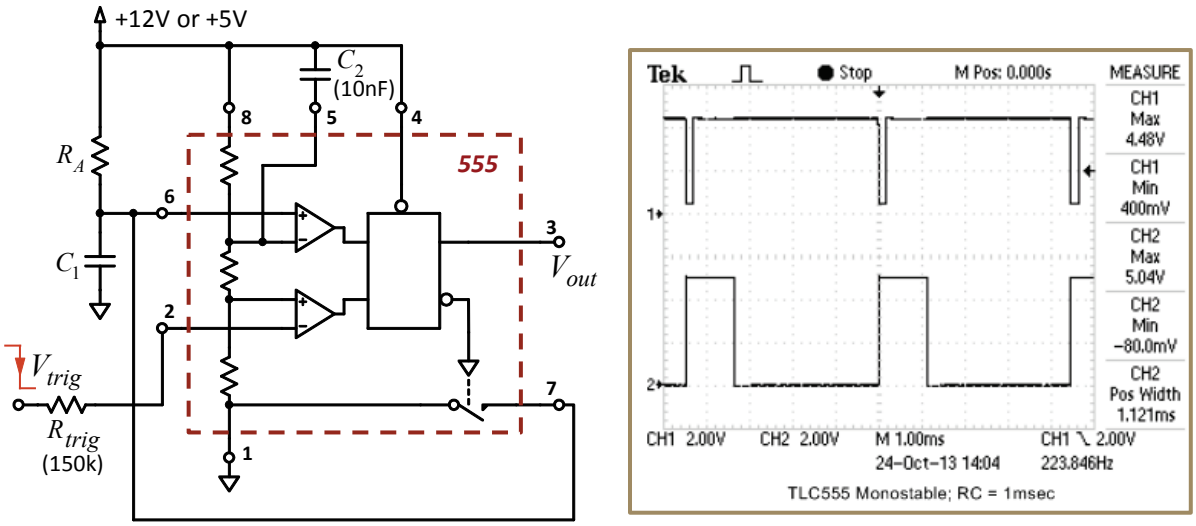


Figure 4-10: *Monostable multivibrator* using the 555. Whenever V_{trig} falls below 1/3 V_{supply} the output goes high and capacitor C_1 charges through resistor R_A. When the capacitor voltage reaches 2/3 V_{supply}, the output goes low, the *Discharge* terminal is shorted to ground, and C_1 is discharged back to 0. The oscilloscope shows trigger event inputs (CH1) and the resulting pulse outputs (CH2). Note that for proper operation, the trigger pulse input must be shorter than the output pulse.

Whereas neither the high nor the low output state of a relaxation oscillator is *stable* (because each state eventually changes to the other without any external input to the circuit), one of the states of a *monostable multivibrator* is stable — an external trigger input is needed to make the circuit transition to its other output state. The other state, however, is *unstable*: after some time the circuit will transition back to its stable state and remain there until another trigger event occurs. The idea is that a trigger event causes the circuit to emit a single pulse of a fixed width and then return to its original, quiescent state. For this reason the monostable multivibrator is also called a *one-shot*: one output pulse for each input trigger.

Figure 4-10 (on page 4-12) shows how to implement a simple one-shot using the 555 timer. Assume the capacitor C_1 is discharged and the 555's *Trigger* input voltage (V_{trig}) is greater than 1/3 of the power supply voltage (V_{supply}). Assume further that the 555 is in its *Reset* state ($V_{out} = 0$ and the *Discharge* terminal shorted to ground). This is the circuit's *stable* (quiescent) state.

A *trigger event* occurs when V_{trig} is momentarily taken well below $1/3\,V_{supply}$. Now the 555 transitions to its active state: the *Discharge* terminal opens and V_{out} goes high. The capacitor charges through R_A toward V_{supply}; when its charge reaches $2/3\,V_{supply}$, the 555 *Threshold* is triggered. If V_{trig} had returned to its quiescent state (well above $1/3\,V_{supply}$) before this happens, then the 555 will return to its *Reset* state, lowering V_{out} and discharging the capacitor back to ground. The time it takes the capacitor voltage to relax from 0V to 2/3 V_{supply} is just over one time constant, i.e. $\approx 1.1 R_A C_1$.

Note that if the trigger input voltage remains below $1/3\,V_{supply}$, then the 555 will remain in its active state (V_{out} high), since the *Trigger* comparator overrides the *Threshold* comparator (see Table 4-1 on page 4-10). If the capacitor voltage has exceeded $2/3\,V_{supply}$, then as soon as V_{trig} goes back above $1/3\,V_{supply}$ the 555 will change state, and its output will immediately return to 0.

The IC's *Trigger* input (pin 2) is protected against excessively low or high V_{trig} voltages ($V_{trig} < 0$ or $V_{trig} > V_{supply}$) by the resistor $R_{trig} = 150\text{k}\Omega$. This resistor limits the current flow within the 555 IC during these V_{trig} excursions so that the 555 isn't permanently damaged.

As with the astable multivibrator circuit, input signals to the 555's Reset and Control pins aren't required for this application, so properly connect them as described before (as in Figure 4-9 and Figure 4-10).

Additional 555 applications

So far we have just scratched the surface of the many applications of the 555 IC. Many more are described in the LM555 datasheet and the 555 Application Note. The web, of course, has sites with myriads of circuits; check out http://www.555-timer-circuits.com.

PRELAB EXERCISES

1. Consider the function generator circuit in Figure 4-6 on page 4-7. Sketch the waveform at the +*Input* of the Schmitt trigger op-amp (U2).

2. Use equation 4.3 on page 4-5 and equation 2.19 on page 2-23 of Experiment 2 to show that the oscillation period T of the simple function generator circuit in Figure 4-6 is given by:

4.7

$$T = 4RC\left(R_2/R_1\right)$$

Will the circuit work if $R_2 > R_1$ (consider equation 4.3)? What should the generator frequency f be if $R = 10\text{k}\Omega$, $C = 0.1\mu\text{F}$, $R_1 = 10\text{k}\Omega$, and $R_2 = 1\text{k}\Omega$? What is the output amplitude (peak-to-peak) of the triangle wave if the square wave amplitude is 22V peak-to-peak?

3. How does the symmetry control potentiometer in Figure 4-7 on page 4-8 affect the output waveform symmetry (how does this part of the circuit work)? Does it change the output frequency by any significant amount when it is adjusted? If the magnitude of the maximum output current available from an op-amp is 10mA and the magnitude of its saturation voltage is 11V, then what is the minimum allowable value for the resistor in series with the output of op-amp U3 for the circuit to work properly? How does the frequency control potentiometer affect the output frequency? Does it affect the waveform symmetry to any significant degree when it is adjusted?

4. Consider the monostable multivibrator circuit in Figure 4-10 on page 4-12. Sketch the voltage waveform at the 555 IC pin 6 (the *Threshold* terminal) to accompany the trigger voltage and output voltage waveforms shown in the oscilloscope screen shot.

5. Again consider the simple function generator circuit in Figure 4-6 on page 4-7. How could you add a multiplier (using the MPY634) to that circuit to provide an input so that an applied voltage will determine the function generator's frequency? Such a circuit is called a *VCO* for *voltage-controlled oscillator*.

   Design a circuit which will output a frequency proportional to the input control voltage applied to the circuit such that ½ the original circuit frequency ($1/T$, where T comes from equation 4.7 above) will be generated when the control voltage is ≈ 5V (it's ok if the circuit cannot quite generate a frequency as high as $1/T$, since the multiplier output will be < 10V). Provide a complete schematic of your circuit using the component values supplied in problem 2 above and assigning values to the additional components you add.

   Are there limits to the input control voltage beyond which the circuit stops working?

LAB PROCEDURE

Overview

During lab you will investigate the behavior of a Schmitt trigger circuit and measure its hysteresis. You will then look at a couple of relaxation oscillator circuits, including the circuit you designed in response to Prelab exercise problem 5.

Next you will build astable and monostable multivibrator circuits using the TLC555 timer IC. To accomplish this task you will need to construct the circuit in the breadboard area of the analog circuit trainer, including installing the integrated circuit and correctly wiring to its pins. This will help prepare you for your upcoming project work, which will be built completely on such a breadboard. ***Make sure you pay attention to the clock during lab and budget enough time to complete this work!***

The Schmitt trigger

Build an inverting Schmitt trigger (Figure 4-3 on page 4-3) using one of the op-amps on the analog trainer which have installed resistors available on its *+Input*. Using a triangle wave input signal, measure its trigger thresholds V_{th+} and V_{th-}. Note how these thresholds change when you change the feedback resistor values. See if you can use the **XY** display mode of the oscilloscope to generate a hysteresis plot like the center image in Figure 4-3.

Reconfigure the input and ground connections to convert your circuit to a noninverting Schmitt trigger (right-hand circuit in Figure 4-4 on page 4-4); use ground (0V) for V_{ref}. Again take an oscilloscope screen shot identifying the trigger thresholds using a triangle wave input for a least one combination of resistor values.

Simple relaxation oscillator

Use the installed components available on the op-amp's *–Input* to reconfigure your circuit into the simple relaxation oscillator shown in Figure 4-5 on page 4-6. Use the 10k and 2.2k resistors on the *+Input* for resistors R_1 and R_2, and choose some R and C pair from those components available on the *–Input*. Take an oscilloscope screen shot similar to that in Figure 4-5; does the formula given in equation 4.6 correctly relate your oscillator's period T to your component values?

Function generator

Construct the simple function generator circuit in Figure 4-6 on page 4-7. Using installed resistor and capacitors available on the trainer, use the component values listed in Prelab exercise problem 2. Take an oscilloscope screen shot showing the both the triangle and square waveform outputs. Does your oscillator's frequency f and the triangle wave amplitude agree with your solution to problem 2?

Voltage-controlled oscillator (VCO)

Using your solution to Prelab problem 5, add a multiplier to your function generator circuit to convert it to a *VCO*. Test its performance using the signal generator to input a constant DC voltage to your circuit's frequency control input.

Use a low-frequency sine wave signal generator output (with an appropriate amplitude and DC offset) to modulate the output frequency of your VCO. Trigger the oscilloscope from the signal generator's voltage to your VCO and take a screen shot showing this control voltage along with your VCO output. This sort of modulation is called *frequency modulation* (FM).

Building a TLC555 timer circuit

Warning

Do not connect the breadboard circuit to the analog trainer power supply until either your TA or the course instructor has looked over your circuit. Wiring the power supply incorrectly into an integrated circuit will often leave it fatally damaged.

Construct the astable multivibrator circuit (Figure 4-9 on page 4-11) in the breadboard area of the trainer. Your TA or the course instructor will give you a TLC555 IC and show you how to determine its pin numbers. Make sure you understand how the various contacts of the breadboard are interconnected and pay attention to the advice your TA and the course instructor give you regarding the layout and wiring of your circuitry on the breadboard.

Caution

555 Timer Voltage Limits

The maximum allowable voltage difference between the *V+ Power* and *Ground* terminals is ***less than 16V*** or so (depending on the specific IC version). Always connect the *Ground* terminal to the breadboard ground; you may then use +12V or +5V for *V+ Power*.

Never let a terminal voltage exceed the limits set by *Ground* and *V+ Power*! In particular, ***never let an input signal go < 0V (Ground)***. Violating this rule means nearly instant destruction of the IC. For the TLC555, putting a 150kΩ resistor in series with an input should protect the IC from inadvertent inputs between +12V and −12V, regardless of its power supply voltage values.

Pick component values which will give an output frequency of a couple of kilohertz. Use at least 10kΩ for resistor R_A to avoid excessive current draw when the capacitor C_1 is being discharged. Using the +12V power supply for V_{supply}, observe the capacitor voltage V_C and

the output voltage V_{out} . Does the output change state when $V_C = 4\text{V}$ and 8V ? Take a screen shot; compare your circuit's output frequency to the chart included with Figure 4-9.

Now use +5V as the power supply voltage. Does the output frequency change by very much?

Additional 555 circuits

If you have time, reconfigure your 555 circuit as a monostable multivibrator (Figure 4-10 on page 4-12). Don't forget the 150k resistor R_{trig} to protect the IC's *Trigger* input from errant trigger voltage inputs!

You can use the pulse output of the signal generator as a trigger source: set the pulse **HiLevel** to about 2/3 V_{supply} and its **LoLevel** to 0V. Set the **Dty Cyc** to ≈ 90% or more to generate a narrow, negative-going pulse as in the oscilloscope image in Figure 4-10. Try to capture a result similar to the result in that figure.

If you have time for more circuits, look through those in the 555 Application Note.

Lab results write-up

As always, include a sketch of the schematic with component values for each circuit you investigate, along with appropriate oscilloscope screen shots. Make sure you've answered each of the questions posed in the previous sections.

Experiment 5

Resonant circuits and active filters

Copyright © Frank Rice 2013 – 2015
Pasadena, CA, USA
All rights reserved.

Experiment 5

Resonant circuits and active filters

Now we return to the realm of linear, analog circuit design to consider the final major topic of the term regarding linear op-amp circuits: linear, second-order systems, which include *resonant circuits* and *active filters* — two-port networks in this category have transfer functions which are described by linear, second-order differential equations.

First we consider the archetypical resonant circuit: the *LC* resonator (inductor + capacitor). We use this circuit to define the *resonant frequency* and *quality factor* for a second-order system, and we investigate the frequency and transient responses of a high-*Q, tuned circuit*.

Next we switch topics to investigate how a bit of positive feedback may be added to our repertory of linear op-amp circuit design techniques. We start by considering a *negative impedance* circuit which employs mild amounts of positive feedback in conjunction with negative feedback. This sort of circuit is found in a wide variety of linear op-amp applications including amplifiers, *gyrators* (inductance emulators constructed from op-amps using only resistor and capacitor feedback networks), current sources, and, in particular, resonant circuits and sinusoidal oscillators.

Finally, we introduce the topic of second-order filters starting, naturally enough, with our friend the *LC* resonant circuit, but this time with an added resistor (a dissipative element) to lower the *Q* to around 1. In this way we can investigate the behavior of second-order low-pass, high-pass, and band-pass filters. We then implement such filters using linear op-amp circuits containing only *RC* combinations in their feedback networks, eliminating the need for costly and hard-to-find inductors. We discuss some of the tradeoffs when selecting the *Q* to use in a second-order filter, and look at Bessel and Butterworth designs in particular.

As a postscript, the *More circuit ideas* section presents a couple of sinusoidal oscillators constructed from resonant circuits.

SECOND-ORDER SYSTEMS AND RESONANCE

LC resonant amplifier

Consider the inverting amplifier circuit in Figure 5-1; assume for the moment that the op-amp is ideal. Note that the feedback impedance is given by the parallel combination of an inductor L and a capacitor C; in addition there is a resistance to which we assign the value QZ_{LC}, where Q is a dimensionless real number and Z_{LC} is an impedance (actually a resistance) which we will soon define. The gain of the amplifier will be, of course, given by the ratio of the feedback and input impedances, $G = -Z_f / R_i$. The parallel elements combine to create a feedback impedance of:

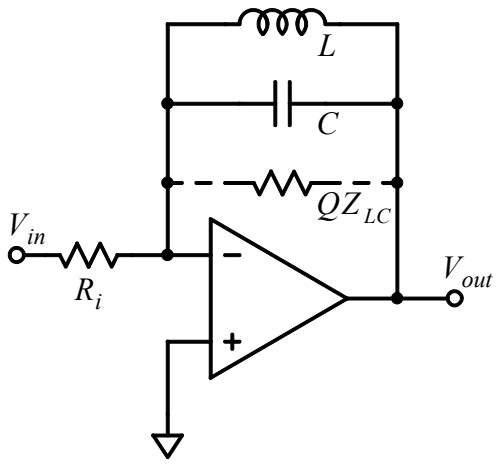


Figure 5-1: An inverting amplifier with a parallel *LC* combination forming the feedback impedance. In addition the circuit will include an equivalent parallel resistance along with the *L* and *C*. Its equivalent value QZ_{LC} is addressed in the text.

$$\frac{1}{Z_f} = \frac{1}{j\omega L} + j\omega C + \frac{1}{QZ_{LC}}$$

if $\omega^{-2} = LC$, the two imaginary terms in the expression for Z_f cancel, and at that frequency just the resistance remains, so $Z_f = QZ_{LC}$. This one frequency at which Z_f is real is called the *resonant frequency* of the LC combination, ω_{LC}. Although the impedances of the L and C cancel at ω_{LC}, neither one is equal to 0; in fact, the magnitude of each is equal to the *characteristic impedance* of the LC combination, $Z_{LC} = \sqrt{L/C}$. The *quality factor*, Q, is thus the ratio of the resistance in parallel with the L and C to the characteristic impedance Z_{LC}.

Using the LC resonant frequency ω_{LC} and characteristic impedance Z_{LC}, the gain of the inverting amplifier may be written as:

5.1

$$G = \frac{-Z_{LC}/R_i}{j\left(\frac{\omega}{\omega_{LC}} - \frac{\omega_{LC}}{\omega}\right) + \frac{1}{Q}}; \qquad Z_{LC} = \sqrt{\frac{L}{C}}, \quad \omega_{LC} = \frac{1}{\sqrt{LC}}$$

Figure 5-2 shows Bode plots of the magnitude and phase of equation 5.1 for $Q = 50$. If $Q > 1$ then far from the resonant frequency Z_f is dominated by either the impedance of the L or the C, and the asymptotic responses (shown by the diagonal dashed lines in Figure 5-2) intersect at ω_{LC} with gain $|G| = Z_{LC}/R_i$. The phases of the two asymptotes have opposite signs, however, and at ω_{LC} their contributions to the gain cancel, leaving only the parallel resistance to determine the gain. Thus the gain magnitude and phase change very rapidly near

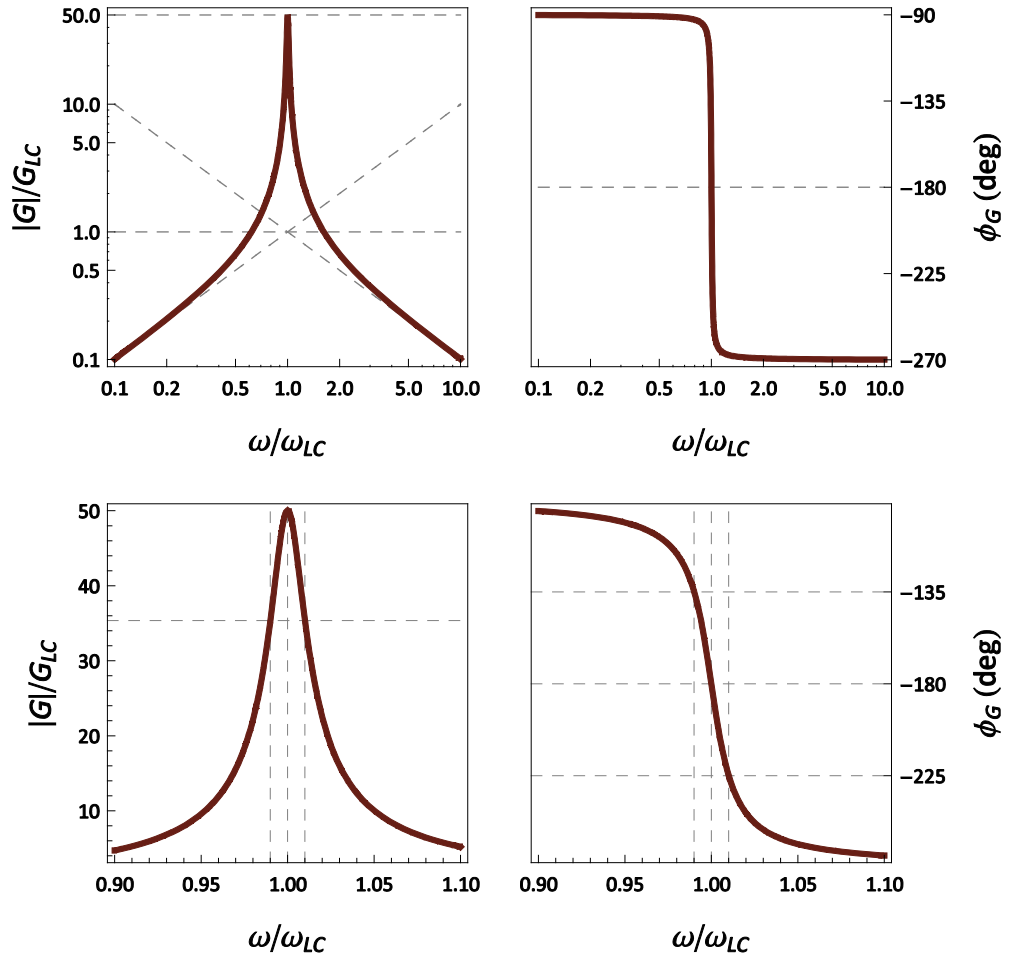


Figure 5-2: Bode plots of the inverting *LC* resonant amplifier for Q = 50. The gain magnitude is normalized to $G_{LC} = Z_{LC}/R_i$. Note that the asymptotic gain lines intersect at ω_{LC} with $G = G_{LC}$, but the actual gain at ω_{LC} is Q times larger. Note that the phase plots have been "unwrapped" so that they remain continuous as the phase passes through −180° at resonance.

The lower plots show details of the response near ω_{LC}. If Q is large, then the −3dB gain points are separated by $\Delta\omega_{-3dB}/\omega_{LC} = 1/Q$ (±1% around ω_{LC} in this example); the phase changes by 90° in this same small frequency range.

ω_{LC}, where the denominator of the gain function in expression 5.1 changes by a factor of Q within a narrow range of frequency of order ω_{LC}/Q (lower pair of plots in Figure 5-2). This behavior is characteristic of the phenomenon of *resonance*. Because the denominator of 5.1 is quadratic in the angular frequency ω, the circuit in Figure 5-1 is an example of a *second-order system*.

Transient response of a high-Q resonant circuit

Recall that because the gain is the ratio of the output and input voltages, equation 5.1 can be written as

$$\left(\frac{j\omega}{\omega_{LC}}\right)^2 V_{out} + \frac{1}{Q}\frac{j\omega}{\omega_{LC}} V_{out} + V_{out} = -\frac{j\omega}{\omega_{LC}}\frac{Z_{LC}}{R_i} V_{in}$$

This frequency-domain expression may be transformed to the time-domain by replacing $j\omega$ with a time derivative; thus the gain expression (5.1) corresponds to the second-order, linear differential equation

5.2
$$\frac{1}{\omega_{LC}^{\,2}}\ddot{v}_{out} + \frac{1}{Q\omega_{LC}}\dot{v}_{out} + v_{out} = \frac{-Z_{LC}}{R_i}\frac{1}{\omega_{LC}}\dot{v}_{in} = f(t)$$

If $Q > 1/2$, the homogeneous solution of (5.2) takes the form of a damped harmonic oscillation:

5.3
$$V_{out}(t) \propto e^{-\omega_{LC} t/2Q}\cos(\omega_T t + \phi); \qquad \omega_T = \omega_{LC}\sqrt{1 - 1/(4Q^2)}$$

This *ringing* following a sudden transient input will persist for many cycles if Q is large; after Q cycles the ringing amplitude will still be about 4% of its starting value (see Figure 5-3).

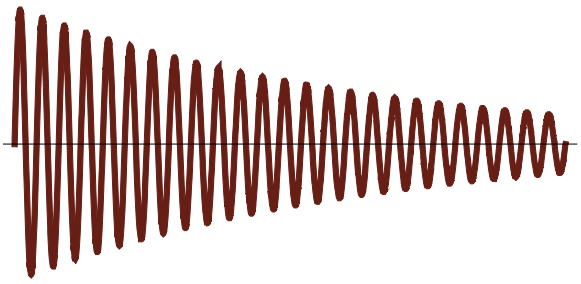


Figure 5-3: First 25 cycles of the output ringing of the circuit in Figure 5-1 for Q = 50. A negative step in the input would excite this response; the ringing initial amplitude would be approximately Z_{LC}/R_i times the input step amplitude.

High-Q resonant circuits such as this are useful to selectively amplify a very specific, narrow band of frequencies, such as when a radio receiver is tuned to a particular transmitter's frequency; for this reason they are sometimes called *tuned circuits*. Note that changing the input or the resonant frequency of the circuit must be followed by a waiting period of several Q's of cycles for the circuit's ringing in response to the change to largely dissipate — high-Q tuned circuits have a relatively long *settling time*.

Note that if a resonant circuit's Q were infinite, the ringing would persist indefinitely — we would have a sine-wave oscillator. This feat may be accomplished by a judicious use of positive feedback in our linear, resonant system to maintain the sinusoidal output. Sine-wave oscillators are discussed in **MORE CIRCUIT IDEAS**: *Sine-wave oscillators* on page 5-22.

POSITIVE FEEDBACK IN LINEAR APPLICATIONS

Stability of linear circuits which include positive feedback

The hallmark of linear op-amp circuits is their use of negative feedback to maintain the two op-amp inputs at the same voltage, as described all the way back in Experiment 1. When we used positive feedback in Experiment 4 it was to create a Schmitt trigger, a highly nonlinear circuit in which the op-amp output spends its time in either positive or negative saturation (except during those brief, slew-rate limited transitions from one saturation limit to the other), and the two op-amp inputs are generally at very different voltages.

In this section we consider linear, analog circuits which, nevertheless, include a bit of positive feedback along with a strong dose of negative feedback. By using this technique we can design analog circuits with a greatly expanded range of capabilities, as we shall soon see. Consider the generic analog op-amp circuit shown at right. The feedback network is constructed from of a set of four impedances forming voltage dividers from V_{out} back to both the op-amp's *+Input* and *–Input*. Call the fraction of V_{out} fed back to these two op-amp inputs f_+ and f_-, respectively; for this example

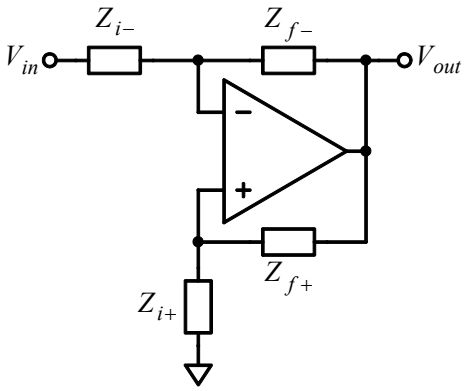


Figure 5-4: Generic inverting amplifier with some positive feedback as well.

$$f_+ = Z_{i+} \Big/ \left(Z_{i+} + Z_{f+} \right), \; f_- = Z_{i-} \Big/ \left(Z_{i-} + Z_{f-} \right).$$

At zero frequency (DC), both f_+ and f_- are nonnegative real numbers, so as long as $f_- > f_+$, the net feedback will be negative, and the circuit will remain stable and linear (note that we must include the output impedance of the source for V_{in} when calculating f_+ and f_-). At other frequencies, the stability criterion is a little more subtle: one must find the set of solutions $s = j\omega$ of the equation:

5.4

$$f_-(j\omega) \; - \; f_+(j\omega) \; + \; 1 \big/ g(j\omega) \; = \; 0$$

where $g(j\omega)$ is the open-loop gain function of the op-amp (as described in Experiment 2). We replace all occurrences of $j\omega$ in (5.4) with the complex variable s and then find the complex solutions s of (5.4). If the real parts of the various solutions of (5.4) are all negative, i.e. $\mathrm{Re}(s) < 0$ for all roots s of this equation, then the circuit will be stable. Proof of this criterion is left to the exercises (just kidding!). Actually, proof of this theorem belongs to the general field of *control system engineering*.

The criterion expressed above always gives a correct assessment of a feedback circuit's stability, and it is generally the best way to evaluate our systems. A simpler, more naïve

criterion is usually adequate for our elementary circuits, and it is more intuitive and easier to apply:

> **Stability criterion when including positive feedback**
>
> As long as the fraction of the output fed back through the *negative* feedback loop is greater than the fraction fed back through the *positive* feedback loop, then the circuit will probably remain stable and predictable.
>
> *This criterion must hold at all frequencies for which the circuit is capable and for all inputs and loads the circuit may encounter.*

A negative impedance circuit

With this stability criterion in mind, consider the circuit in Figure 5-5, a one-port network driven by a source with output impedance Z_s, as shown. What we are interested in is to determine the circuit's input impedance, $Z_{in} = V_{in}/I_{in}$, and to decide how the op-amp's + and – input terminals should be connected in order to ensure the circuit's stability.

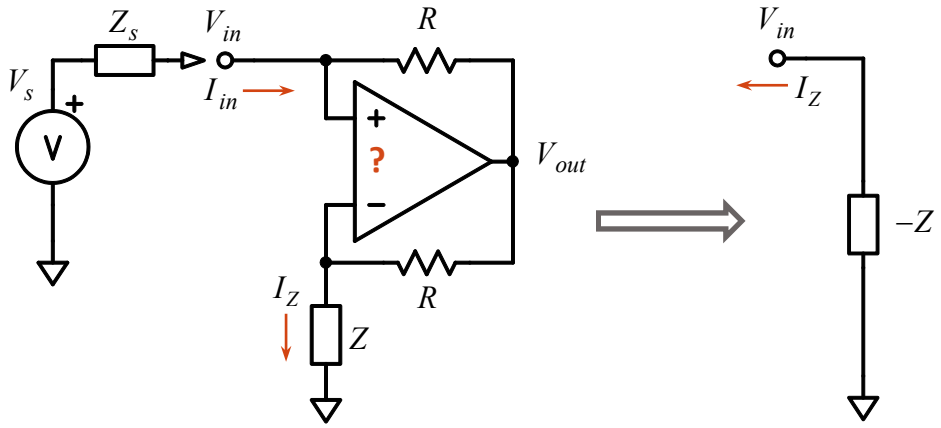


Figure 5-5: A *negative impedance* circuit. As explained in the text, the input impedance of the op-amp circuit, V_{in}/I_{in}, is $-Z$, the *negative* of the impedance of the element Z in the circuit. The "?" on the op-amp symbol implies that *we must carefully consider how the op-amp's inputs should be connected* — with *–Input* at top as shown, or should the op-amp's inputs be flipped? This issue is discussed in the text.

If we assume that the stability criterion is satisfied, i.e. that there is net negative feedback around the op-amp, then we can assume that the ideal op-amp's *+Input* and *–Input* voltages are equal. This means that the input voltage V_{in} appears at both op-amp inputs, and, therefore, that V_{in} must be the voltage drop across the impedance Z (since its other end is at ground). The resulting current through Z (I_Z in the figure) must come from the op-amp output by flowing through the feedback resistor. The voltage drops across the two resistors R in the figure are the same, i.e. $V_{out} - V_{in}$, so current I_Z must also flow through the top

resistor R from the op-amp output toward the input terminal of the circuit. Thus $I_{in} = -I_Z$! Therefore:

5.5

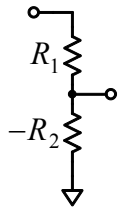
$$\boxed{Z_{in} = V_{in}/I_{in} = -V_{in}/I_Z = -Z}$$

The circuit's input impedance is the negative of the impedance Z — we have a device with a *negative impedance* to ground. Thus, for example, if the impedance Z is a simple resistor, *then when we apply a positive voltage to the circuit's input, the circuit will push current back into the source*; it would be a *source* of power rather than an *absorber* of power (of course, the power comes from the op-amp power supplies).

Before we consider how to use this circuit for something interesting, we must make sure that our original assumption is correct: is the stability criterion satisfied? We must check that there is less output feedback to the op-amp's *+Input* than to its *–Input*. Since the two feedback resistors each have value R in Figure 5-5, the feedback fraction will depend on the impedance to ground at each op-amp terminal, since that impedance will set the voltage divider ratio at that terminal. For the *–Input* terminal in Figure 5-5 this is the impedance Z; the impedance to ground at the op-amp's *+Input* is the *output impedance of the driving source*, Z_s.

Thus for the configuration shown in Figure 5-5, the stability criterion requires that $|Z_s| < |Z|$ at all frequencies the op-amp can amplify. If the opposite were true, i.e. $|Z_s| > |Z|$, then the inputs to the op-amp should be reversed.

Now let's consider a simple application of our negative impedance circuit: a voltage divider which includes a "resistor" with a *negative* value, as shown at right. The gain of this circuit is $V_{out}/V_{in} = -R_2/(R_1 - R_2)$, which has a magnitude greater than 1 if $R_1 < 2R_2$. We could build this sort of amplifier using our negative impedance circuit to realize the negative resistance R_2; we just have to be careful as we consider the stability criterion in this case.



The gyrator

An interesting application of the negative impedance circuit is to try to find a more complicated circuit using it which *inverts* an impedance: $Z \to 1/Z$; then we could emulate an inductor using a capacitor ($1/j\omega C \to j\omega C$). Such a device is called a *gyrator*. A circuit which accomplishes this task using voltage dividers and our negative impedance circuit is shown in Figure 5-6 on page 5-8. Working out that this circuit does indeed invert the impedance Z is left as an exercise to the reader. What is more problematic, however, is to decide how to orient the op-amps' inputs and apply the stability criterion to this circuit.

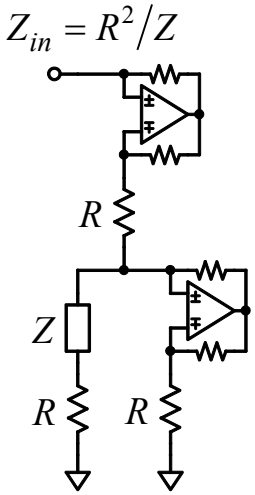


Figure 5-6: A simple *gyrator* built from negative impedance converter circuits. Deciding how to orient the op-amp inputs could be a problem, however.

A more sophisticated application of combined positive and negative feedback to implement a gyrator circuit is shown in Figure 5-7. This circuit is quite different from the negative impedance circuit, and it is quite stable, though we won't consider its stability in detail here.

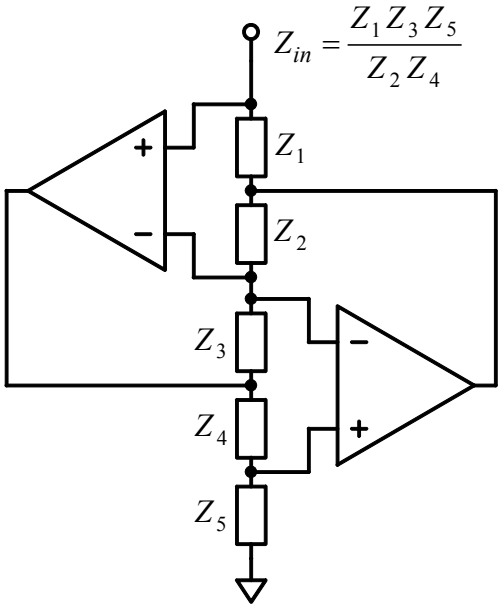


Figure 5-7: A practical gyrator circuit. The op-amps will maintain their inputs at the same potential, so the voltage across Z_5 and the voltage at the junction of Z_2 and Z_3 must both equal the input voltage. With this condition it is straightforward to derive the input impedance formula shown.

Assuming that the op-amps are ideal and working linearly and correctly, then their inputs must all be at the same potential, which must be the input voltage, since an op-amp input is connected there (Figure 5-7). With this constraint plus the observation that the current

through Z_2 must equal that through Z_3, and similarly for the currents through Z_4 and Z_5, it is straightforward to show that the circuit's input impedance is given by the expression included in the figure.

> The conventional way to emulate an inductor with the gyrator circuit in Figure 5-7 is to use a capacitor for Z_4 and resistors for the other impedance elements.

The Howland Current Pump

Another interesting linear op-amp circuit employing both positive and negative feedback is the *Howland Current Pump*, invented many years ago by Bradford Howland at MIT. This circuit is an example of *a voltage to current converter*: it establishes a constant current through a load which is proportional to a control voltage input to the circuit.

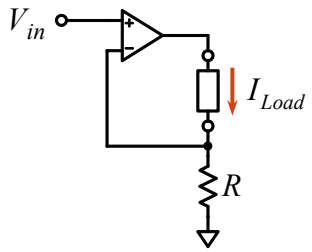


Figure 5-8: A simple voltage to current converter: $I_{Load} = V_{in}/R$.

First, however, consider the simple circuit in Figure 5-8, originally introduced in Experiment 1. In this case the load element is positioned in the op-amp's negative feedback loop. Because the two op-amp input voltages are equal, the supplied input voltage V_{in} also appears across the resistor R. The op-amp output current required to establish this voltage across R must pass through the load, so the load current is set to $I_{Load} = V_{in}/R$, and it will be independent of any variation in the load's impedance.

A significant drawback of the simple circuit of Figure 5-8, however, is that its load must *float*: the load's two terminals must both be independent of the circuit ground. A more common requirement which our simple circuit can't satisfy is to supply a fixed current into a single terminal of a load that itself might be part of a complicated sub-circuit; the injected current would then be returned through the system ground or power supplies. The Howland Current Pump shown in Figure 5-9 on page 5-10 provides an elegant solution to this design problem.

To understand how this circuit operates, we will analyze it in two stages: (1) determine the load current I_{Load} when the load impedance is 0, i.e. the circuit is connected directly to ground at that point; (2) show that the load current is independent of any variation in the load voltage $V_{Load} = Z_{Load}\, I_{Load}$. In any case, since the ideal op-amp inputs draw no current, it must be true that the load current is given by $I_{Load} = I_1 + I_2$ (see Figure 5-9).

Let us therefore first consider the case (1) in which the output of the circuit goes directly to ground $(Z_{Load} = 0)$ and determine the two currents I_1 and I_2. Since the output node is now a circuit ground, the op-amp *+Input* is grounded, and the op-amp circuit becomes a simple

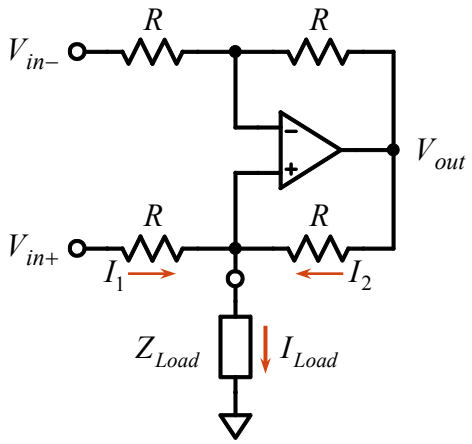


Figure 5-9: The *Howland Current Pump* supplies a current into a terminal of a grounded load which is independent of the load's impedance. The load current is given by equation 5.6; its operation is described in the text.

inverting amplifier with $V_{out} = -V_{in-}$. Now the voltages V_{in+} and V_{out} appear across the resistances *R*, determining currents I_1 and I_2. Thus the load current is given by:

5.6
$$I_{Load} = I_1 + I_2 = (V_{in+} - V_{in-})/R$$

Next consider the effect of a change in the load voltage on the load current. Because the circuit is linear, we can use the principle of linear superposition presented in Experiment 1. Thus we consider the case where $V_{in+} = V_{in-} = 0$, and ask how I_{Load} depends on V_{Load}, which is the voltage at the node connected to the op-amp's *+Input*. Since the terminal at V_{in-} is now ground, the op-amp circuit is a noninverting amplifier for the voltage at its *+Input*, and $V_{out} = 2V_{Load}$. Since the terminal at V_{in+} is also ground, the voltage drop across the bottom-left resistor is V_{Load}, and $I_1 = -V_{Load}/R$. The voltage drop across the bottom-right resistor is $V_{out} - V_{Load} = V_{Load}$, so $I_2 = V_{Load}/R$. This means that $I_{Load} = I_1 + I_2 = 0$, so the load current is unaffected by the load voltage. Therefore equation 5.6 is correct regardless of the load voltage and, thus, is independent of the load impedance Z_{Load}. The Howland Current Pump serves as a current source whose output current is programmed by the differential input voltage $V_{in+} - V_{in-}$ according to equation 5.6.

INTRODUCTION TO SECOND-ORDER ACTIVE FILTERS

Resonant circuits as filters

High-Q resonant circuits with responses similar to that shown in Figure 5-2 on page 5-3 are useful for the narrow-band, high gain filters needed in radio-frequency tuning circuits. For many applications, however, one needs low-pass and high-pass filters with a relatively flat pass-band response and a rapid reduction in gain (*roll-off*) as the signal frequency moves beyond the filter's cut-off frequency. A second-order resonant circuit with a modest Q ($Q \lesssim 1$) provides a solution which offers a flatter frequency response in the pass-band, a more dramatic out-of-band roll-off, and a faster settling time than you can achieve using the simple, first-order RC filters discussed in Experiment 2. As a set of simple examples reminiscent of our first-order RC filters, consider the series LCR circuits configured as voltage dividers in various ways shown in Figure 5-10.

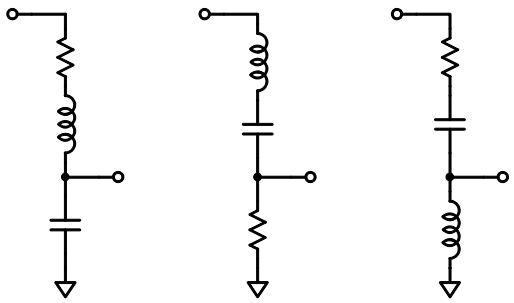


Figure 5-10: Series *LCR* circuits configured as simple filters. By forming a voltage divider and taking the output from the voltage across each of the various elements, various filter responses may be achieved. In each case the resonant frequency is, of course, $\omega_0 = 1/LC$; the quality factor $Q = Z_{LC}/R$.

For each of the circuits above, the transfer function will have a denominator (D) like that of the gain function in equation 5.1, namely:

5.7

$$D \equiv j\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right) + \frac{1}{Q}; \qquad \omega_0 = \frac{1}{\sqrt{LC}}, \; Q = \frac{\sqrt{L/C}}{R}$$

These are called *second-order* filters because $D(\omega)$ is quadratic. At ω_0, $1/D = Q$. The numerators for the various LCR filters depend on the bottom element used in the voltage divider (each numerator is given by $Z(\omega)/Z_{LC}$, where $Z(\omega)$ is the impedance of the bottom element and $Z_{LC} \equiv \sqrt{L/C}$). The resulting transfer functions for these three filters are plotted in Figure 5-11.

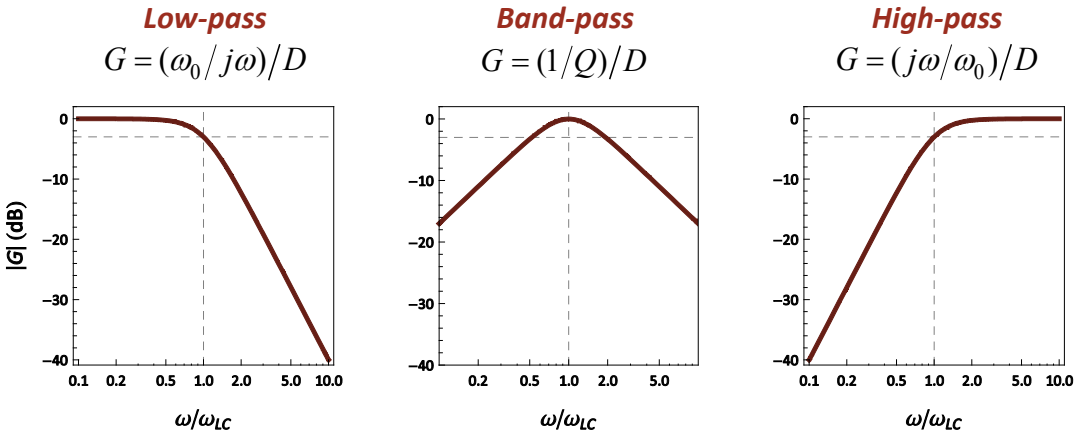


Figure 5-11: Transfer functions and Bode gain plots for the second-order filters of Figure 5-10 for Q^2 = 1/2. The −3dB gain points for the band-pass filter are at 0.52ω_0 and 1.93ω_0; for the others it is at ω_0. The out-of-band filter slopes are ω^{-2}, $\omega^{\pm 1}$, and ω^{+2}. The denominator D in the gain formulas is defined in equation 5.7.

A second-order filter with $Q = 1/\sqrt{2}$ (such as those plotted above) is an example of a *Butterworth filter*, a good choice for a general-purpose filter.

Filter types: choosing the right Q

If one were to simply cascade two first-order low-pass filters with the same cutoff frequency (as in the case of the two cascaded, gain-11 amplifiers evaluated in Experiment 2), then their combined response would be equivalent to a second-order filter with $Q = 1/2$. Such a response is characteristic of a *critically damped* second-order system, and this behavior is often sought for by control system engineers. For the purposes of signal filtering, however, such a low Q system may sacrifice a bit too much in the way of frequency response, because its attenuation (*roll-off*) at its −3dB corner frequency is quite gradual (*soft*), and the filter introduces significant phase shifts at frequencies far from the corner. A high-Q response, on the other hand, may have a steep roll-off at its −3dB corner frequency, but it will exhibit severe gain peaking near its resonant frequency and will show a lot of ringing in response to a transient input. Making a choice of Q which results in the best compromise of gain flatness and phase shift in the pass-band, steepness of the roll-off at the −3dB frequency, and transient response is often a difficult one and will depend on your specific application. In this section we give examples of some popular choices designers typically consider.

The most popular choice for a second-order filter is probably the *Butterworth filter*, named after the British physicist Stephen Butterworth and characterized by $Q = 1/\sqrt{2}$ (as with the filters in Figure 5-11). It is called the *maximally-flat* low-pass filter because the second derivative of its transfer function vanishes only at 0 frequency, making it the highest Q second-order low-pass filter with a monotonically decreasing gain as frequency increases (i.e. no gain peaking or pass-band gain *ripple*). Its −3dB corner frequency is also its resonant frequency, ω_0.

A less common choice is the *Bessel filter*, with $Q = 1/\sqrt{3}$. Its lower Q places its −3dB corner frequency is at $0.79\,\omega_0$, and the filter roll-off near its −3dB frequency is softer than the Butterworth (see left-hand graph in Figure 5-12 below). The advantage of the Bessel filter for certain applications, however, is that it has constant *group delay* for frequencies well within its pass-band, implying that it will cause the least distortion to the shape of a complicated waveform.

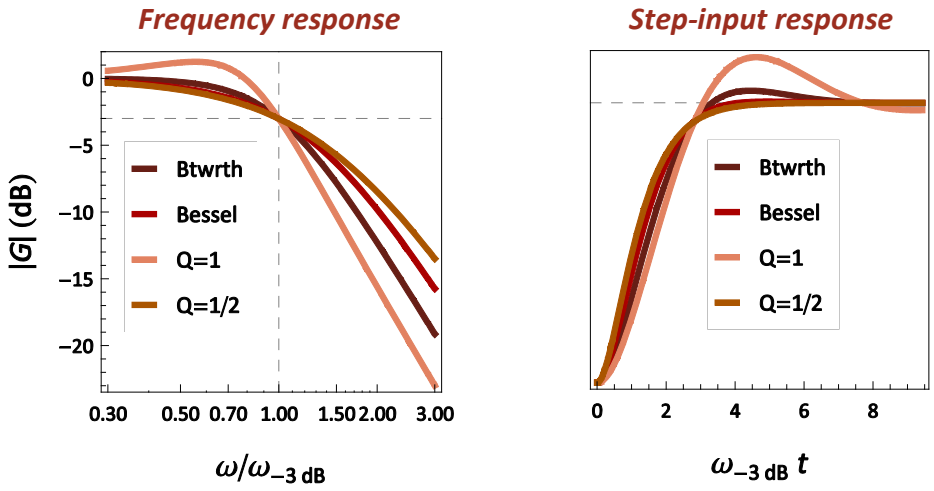


Figure 5-12: Frequency and transient responses of various 2nd order low-pass filters. All filters have the same −3dB frequency, but have different *Q*'s (the Butterworth filter has legend "Btwrth"). The critically damped filter has *Q* = 1/2. The *Q* = 1 filter has mild gain peaking (+1.25 dB), but significant step response overshoot; its resonant frequency is 0.79 of its −3dB corner frequency.

The critically damped filter ($Q = 1/2$) mentioned previously has its −3dB corner frequency at only $0.64\,\omega_0$, and the filter roll-off there is very soft. It has the distinction, however, of having the highest Q for which its transient response to a step input has no overshoot; its output settles following a step input more rapidly than for any other second-order filter with the same −3dB corner frequency (right-hand graph in Figure 5-12), although its resonant frequency must be made nearly 60% higher than that of an equivalent Butterworth filter.

A very simple VCVS active filter

The main drawback of an *LCR* filter is that it requires a high-quality inductor as one of the filter elements. This may not be a problem for a filter designed to operate at several MHz, but for frequencies below a few MHz your choice of suitable inductors may be limited (and such inductors are relatively expensive). Consequently, an *active filter* implemented using op-amps with *RC* feedback networks is the more practical solution for a low-frequency filter. A gyrator circuit (Figure 5-7) may be used to emulate the inductor, but only if the circuit has one terminal of the inductor connected to ground (only the high-pass series *LCR* filter discussed previously meets this requirement). Fortunately, electronics designers have

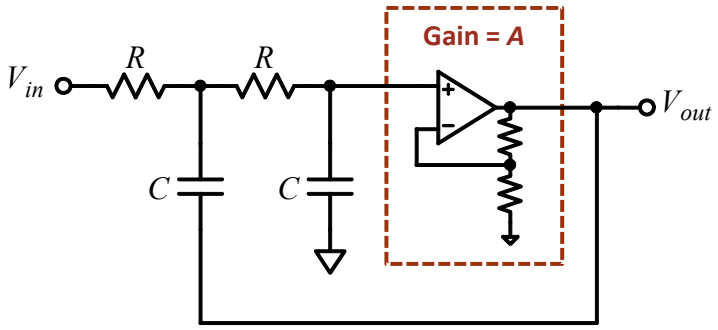


Figure 5-13: A simple version of the *VCVS* second-order, active low-pass filter, in which both the resistors and the capacitors of the two cascaded *RC* filters are chosen to be equal. The non-inverting op-amp amplifier stage has gain *A* (if a simple voltage follower is used, so *A* = 1, then the circuit is called a *Sallen-Key* filter). The circuit's resonant frequency is $\omega_0 = 1/RC$, and its $Q = 1/(3-A)$. The circuit's in-band gain is *A*.

invented several op-amp circuits employing both positive and negative feedback which implement quite effective second-order filters. In this section we discuss one of these, the *Voltage-Controlled Voltage Source* (*VCVS*) active filter.

A simple *VCVS* circuit for a second-order low-pass filter is shown in Figure 5-13. It is clearly a cascaded pair of simple *RC* filters followed by a noninverting op-amp gain stage. The wrinkle to this circuit, however, is that the first *RC* filter stage is terminated not by grounding its capacitor but by connecting it to the amplifier output, providing some positive feedback (also called *bootstrapping*). This clever change makes the filter's Q as well as its resonant frequency ω_0 adjustable by selecting the appropriate component values and amplifier gain A.

Rather than analyze the general case where the component values and amplifier gain are all arbitrary, we consider only the special case where both resistors and both capacitors are chosen to have equal values R and C. In this special case the transfer function of the circuit is:

Simplified *VCVS* low-pass filter (Figure 5-13)

5.8

$$G = \frac{A}{1 - \frac{\omega^2}{\omega_0^2} + j\frac{\omega}{Q\omega_0}}; \qquad \omega_0 = \frac{1}{RC}, \qquad Q = \frac{1}{3-A}$$

The circuit is called a *VCVS* filter because the op-amp stage acts as a *voltage-controlled voltage source* (which is just another way of saying that its output voltage is a function of its input voltage, and that its output impedance is small). Examining the circuit should make it clear that its gain at very low frequency, in which case the capacitors' impedances get very large, should approach that of the noninverting amplifier stage, A. Not as clear, perhaps, is that the positive feedback stability criterion on page 5-6 will demand that $A < 3$ (note from

the equations 5.8 that $Q \to \infty$ as $A \to 3$). If the resistor and capacitor values are not well-matched, then the circuit's ω_0 and Q will vary a bit from the expressions in 5.8.

High-pass and band-pass versions of the simple *VCVS* filter are shown in Figure 5-14. If $A = 1$ (i.e. a voltage follower is used for the noninverting op-amp stage), then the circuit is called a *Sallen-Key filter* after its inventors at the MIT Lincoln Laboratory in 1955. The section **MORE CIRCUIT IDEAS:** *Sallen-Key low-pass filter* on page 5-20 shows how to design such a filter.

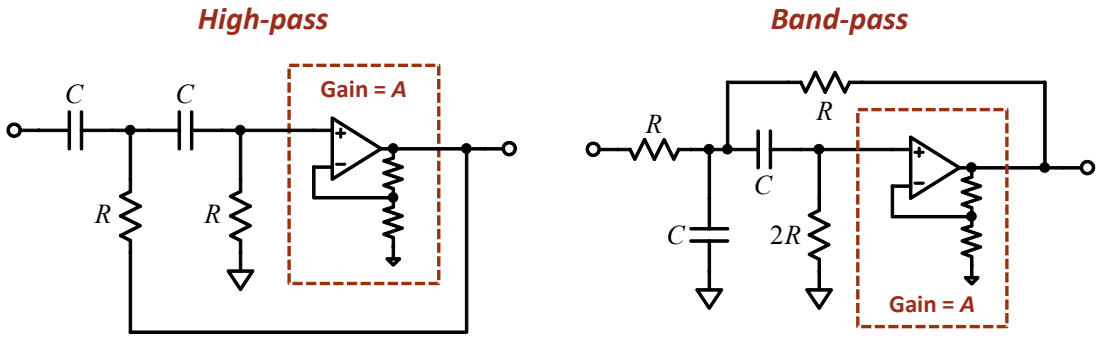


Figure 5-14: Other *VCVS* second-order filters. As with the low-pass filter, the filters' $\omega_0 = 1/RC$, and $Q = 1/(3-A)$. The in-band gain of the high-pass filter is A; the band-pass filter's gain at $\omega_0 = QA$. Note that one resistor in the band-pass version of the filter has value $2R$.

Other active filter circuit topologies; the state-variable filter

Another common second-order active filter which uses only a single op-amp is the *multiple-feedback* (*MFB*) design, which is most useful for high-Q or high-gain filter stages. A low-pass version of the circuit is shown in Figure 5-15; to get a high-pass design, simply swap the resistors and capacitors in the circuit. Note that this is an inverting amplifier (the pass-band gain is $-R_2/R_1$) and that the op-amp is configured as an integrator. We won't discuss this filter any further; you can find several references for designing this type of filter in the lab library or on the web.

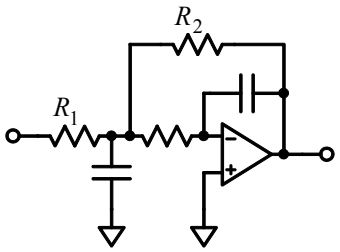


Figure 5-15: *Multiple-feedback* low-pass filter circuit.

A more flexible design than either the *VCVS* or *MFB* topologies is the second-order *state-variable filter* shown in Figure 5-16, which is from the Texas Instruments ASLK Pro Manual, page 32; that document's *Experiment 5* shows how to convert this circuit to a voltage-controlled filter. Although this filter uses 4 op-amps, it simultaneously provides low-pass, band-pass, high-pass, and *band-stop* filter outputs, and its pass-band gain A and its Q are easily adjustable. Note that the circuit topology is formed from a cascade of two inverting integrators and a gain −1 inverting summing amplifier whose output is fed back to become

the input to the first integrator. To this basic loop is added a feed-forward section around the second integrator — the input signal is summed with the output of the first integrator in this section.

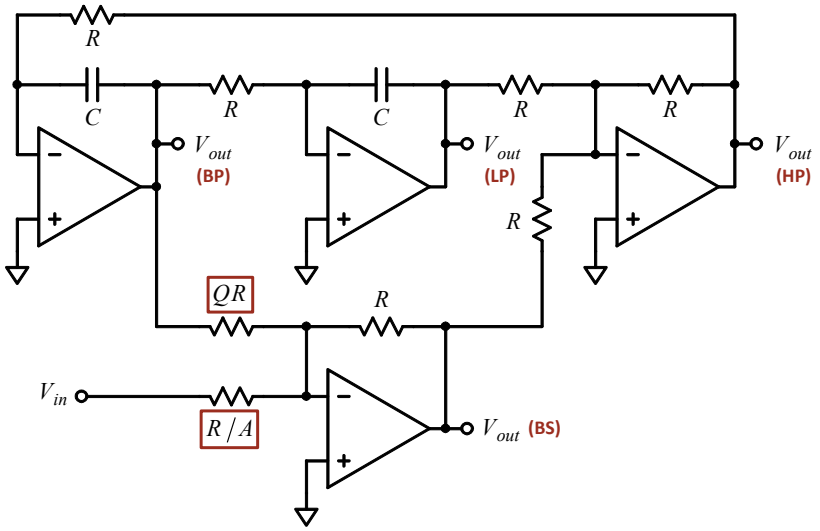


Figure 5-16: A *Universal Active Filter* (*UAF*), a form of second-order, state-variable active filter. This design is from the Texas Instruments ASLK Pro Manual, ©Texas Instruments, 2012. It simultaneously provides high-pass (HP), low-pass (LP), band-pass (BP), and band-stop (BS) outputs. The pass-band filter gain is *A*; its resonant frequency is $\omega_0 = 1/RC$.

The resonant frequency is, naturally, $\omega_0 = 1/RC$. The gain in the pass-band of the low-pass, high-pass, and band-stop filters is A, which from Figure 5-16 is just the gain for the input signal of the summing amplifier at the bottom of the figure; the gain of the band-pass filter at ω_0 is QA. Q is simply the reciprocal of the gain of the feed-forward signal from the first integrator through this same summing amplifier. The response of the band-stop filter is shown at right.

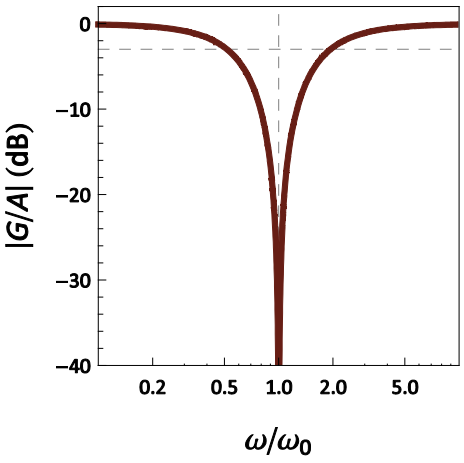


Figure 5-17: *Band stop* filter response.

You can see how the state-variable filter works by first reviewing Figure 5-11 on page 5-12. Starting with the high-pass filter's response function in that figure, note that you can obtain the response functions of the band-pass and low-pass filters by successive divisions by $j\omega/\omega_0$. But this is the same as integrating in the time-domain, so clearly the successive integrators in the state-variable filter loop (Figure 5-16) convert the high-pass output to band-pass and low-pass outputs, as long as $\omega_0 = 1/RC$. Now with a little bit of arithmetic and keeping in mind that the amplifiers are all inverting, you should be able to show that the two summing amplifiers combine the low-pass, band-pass, and input signals

in just the right way to generate the high-pass response. Note that the input impedance of the filter is equal to the single input resistor value R/A; you may need to add a voltage follower to the input to obtain a large input impedance.

If you need to be able to easily and independently adjust the gain or Q of a second-order active filter or if you need multiple filter outputs to split an input signal into frequency sections, then the state-variable filter is a good choice.

> The state-variable filter is really only suitable for filters with a modest *Q* of 2 or less (which is by far the most commonly encountered requirement); otherwise extremely good component matching of the various *R*s and *C*s may be necessary, especially for the band-stop filter output.

PRELAB EXERCISES

1. Consider the voltage divider with a negative impedance circuit as one "element" shown in Figure 5-18 at right.

   a. What is the gain (V_{out}/V_{in})?

   b. Which way should the op-amp's inputs be oriented (+*Input* at top or –*Input* at top)? (assume that the source of V_{in} is a perfect voltage source and the load at V_{out} has infinite input impedance)

   c. What are the divider's input and output impedances (assume that the source of V_{in} is a perfect voltage source and the load at V_{out} has infinite input impedance)?

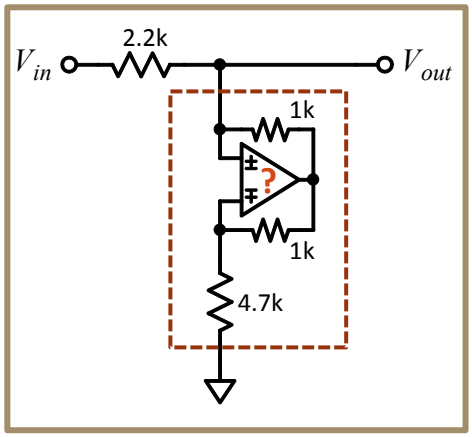


Figure 5-18: A voltage divider with a negative impedance (boxed circuit).

   d. If $V_{out} = 1.0\text{V}$, then what is the voltage at the op-amp output?

2. Consider the *Howland current pump* shown in Figure 5-9 on page 5-10. What is the op-amp output voltage V_{out} in terms of the input and load voltages V_{in-}, V_{in+}, and V_{Load}?

3. Sketch the phase response Bode plots to go with each of the gain magnitude plots shown in Figure 5-11 on page 5-12 for the *LCR* filters shown in Figure 5-10.

4. Consider the simple *VCVS* low-pass filter in Figure 5-13 on page 5-14. What should be the ratio of the two resistors of the noninverting amplifier (i.e. R_f/R_i) if you want a *Butterworth* filter response?

5. Design a cascaded amplifier-filter circuit which has a pass-band gain = 10 and has a *Butterworth* low-pass filter response with a –3dB corner frequency of 16kHz. Use a *VCVS* circuit for the low-pass filter (Figure 5-13 on page 5-14).

   Circuit specifications:

   Input impedance: 1MegΩ
   Pass-band gain: 10 (noninverting)
   Filter spec: Butterworth low-pass, –3dB corner frequency of 16kHz

   When you cascade the amplifier and filter, should you put the amplifier or the filter first? Which of the above specifications will most strongly influence your choice?

   Hints for resistor selections: $3.3\text{k}/5.6\text{k} = 0.59$; $27\text{k}/5.1\text{k} = 5.3$.

LAB PROCEDURE

Overview

During lab you will take a quick look at an amplifier using a negative impedance circuit as part of a voltage divider. Next, you will construct the amplifier-filter you designed as part of the prelab exercises. Finally, you will investigate the capabilities of the state-variable filter architecture.

Negative impedance circuit

Construct the amplifier with a negative impedance circuit shown in Figure 5-18 using your solution to Prelab Exercise 1 as a guide. Measure its gain and compare with your calculations.

Amplifier-filter using the VCVS architecture

Construct the amplifier-filter circuit you designed (Prelab exercise 5) in the breadboard area using a supplied TL082 dual op-amp IC (refer to the TL082 data sheet , page 3 for the IC pin-out — *make sure you look at the pin-out for the correct IC!*).

Use the *Frequency Response* application to plot the frequency response of your circuit. Next, input a square wave and use the oscilloscope to investigate the filter's transient response to a step input. Compare your results to the graphs in Figure 5-12 on page 5-13.

State-variable (UAF) filter

Using an additional TL082 IC, now construct a *UAF* with a 16kHz resonant frequency as in Figure 5-16 on page 5-16. Initially choose a Q of 1 and a pass-band gain A of 1 for component selections. You should be able to reuse the same RC pairs you used for the *VCVS* filter for the two integrator sections of the circuit. Check the frequency responses of each of the filter's four outputs. Check the transient response of the low-pass filter output.

Now, by changing the value of the appropriate resistor, increase the Q to approximately 10 and look again at the outputs' frequency responses, especially that of the band-stop (notch) filter. Check the low-pass filter's transient response.

Additional Circuits

If you have time, construct a circuit from the **MORE CIRCUIT IDEAS** section, from an earlier experiment, or one of you own design; the *Wien bridge oscillator* shown in Figure 5-22 on page 5-24 is an interesting choice.

Lab results write-up

As always, include a sketch of the schematic with component values for each circuit you investigate, along with appropriate oscilloscope screen shots.

MORE CIRCUIT IDEAS

Sallen-Key low-pass filter

The simplified VCVS filter presented earlier (Figure 5-13 on page 5-14) with equal *R*s and *C*s in the cascaded low-pass filters is rather inflexible, since the gain of the amplifier must be chosen to give the required *Q*. It is often desirable for the filter's pass-band gain to equal 1, so that the amplifier configuration used is just a simple voltage follower. In this case one gets the standard *Sallen-Key low-pass filter* shown in Figure 5-19. The resistors and capacitors now generally need to have unequal values as indicated in the figure by the ratios ρ and κ.

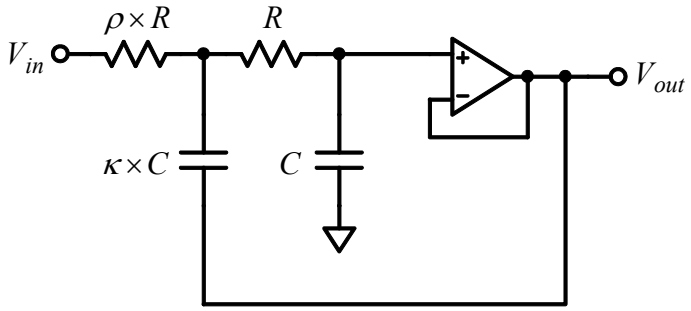


Figure 5-19: Sallen-Key low-pass filter. By using a voltage follower for the *VCVS* amplifier, the pass-band gain of the filter is unity; the ratios ρ of the resistors and κ of the two capacitors are chosen to select the filter's *Q* and resonant frequency ω_0.

The relationships between the filter's ω_0 and Q and these component value ratios are given by:

5.9

$$\omega_0 = \frac{1}{RC\sqrt{\kappa\rho}} \qquad Q = \frac{\sqrt{\kappa\rho}}{\rho+1}$$

Because the selection of capacitor values is generally more limited than that for resistors, one usually picks a convenient capacitor ratio such as $\kappa = 10$. With this choice the resistor value ratio is determined from the required *Q*:

$$\rho = (5/Q^2 - 1) \pm \sqrt{(5/Q^2 - 1)^2 - 1}$$

Clearly, the resistor value ratio ρ must be real and positive, so the argument of the square root must be nonnegative; thus the choice $\kappa = 10$ will be valid only for $Q \le \sqrt{5/2}$. This requirement is easily met by any reasonable filter *Q* (usually $Q < 1$). For a Butterworth filter response $(Q = 1/\sqrt{2})$,

Sallen-Key Butterworth filter component values

5.10

$$\kappa = 10, \quad \rho = 9 \pm \sqrt{80} \approx 18 \text{ or } \boxed{1/18} \quad \rightarrow \quad \omega_0 \approx \frac{1}{(0.75)RC}$$

By choosing $\rho = 1/18$, the filter's corner frequency is reasonably close to $1/RC$, and the impedance of the second RC filter is quite a bit larger than that of the first, so the first RC filter's response is only marginally affected by the load of the second RC filter. Once you have chosen appropriate values for the ratios κ and ρ, choose C and R to give the desired ω_0.

Many tools are available to help you design active filters. My favorite is this web-based application available from Analog Devices:

Analog Filter Wizard
http://www.analog.com/designtools/en/filterwizard/

A simple nonlinear amplifier

In the next section we discuss the design of sine-wave oscillators, which are essentially resonant circuits with a Q of "exactly" the point at complex ∞ when the resonator output amplitude takes on a certain target value; if its amplitude is too large, then the Q is finite and positive so that the output amplitude decays back to the target value, whereas if the amplitude is too small (as when the circuit is first turned on), then the Q becomes *negative*, so the amplitude *increases* toward the target value.

This seemingly remarkable behavior may be accomplished by using a simple *nonlinear amplifier whose gain decreases as its output amplitude rises*. A noninverting version of such an amplifier is shown in Figure 5-20.

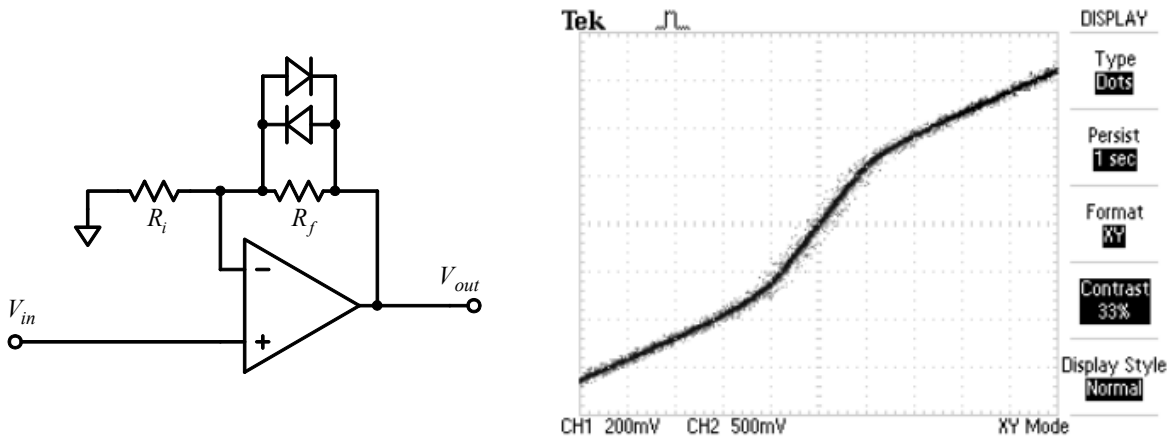


Figure 5-20: A simple nonlinear amplifier. For large signal inputs the amplifier's *dynamic gain* drops to approximately 1, as explained in the text and as illustrated by the V_{out} v. V_{in} data shown in the right-hand image (R_i = 1k, R_f = 2.2k for the data shown).

Here's how the circuit works:

If the feedback current (V_{in}/R_i) is small, then the amplifier's gain is given by the normal noninverting amplifier formula $1 + R_f/R_i$. As the signals get larger, the feedback current goes up; when the voltage drop across R_f approaches 0.6V, the silicon diodes begin to

conduct significant current, reducing the effective feedback resistance and thus reducing the amplifier's gain. Since the diodes will limit the voltage drop across R_f to about 0.6V as the input signal continues to rise, we see that for large inputs V_{out} will remain a diode drop above V_{in}, and the amplifier's *dynamic gain* (dV_{out}/dV_{in}) is reduced to ≈ 1. This behavior is illustrated by the V_{out} v. V_{in} data plotted in Figure 5-20, for $R_i = 1.0\text{k}$ and $R_f = 2.2\text{k}$. When the input voltage is small, the amplifier's gain is 3.2, as indicated by the slope of the curve near the origin. At an input voltage of $\approx 0.25\text{V}$, however, the slope of the curve changes abruptly as the diodes begin to conduct (the voltage across R_f is then $\sim 0.6\text{V}$). As V_{in} continues to increase, $V_{out}/V_{in} \to 1$ ($V_{out}/V_{in} \approx 1.6$ for $V_{in} = 1\text{V}$).

Sine-wave oscillators

As mentioned in the section on resonant circuits, if a second-order system's $Q \to \infty$, then its transient response ringing will continue indefinitely at its resonant frequency — it has become a sine-wave oscillator. In this section we briefly present a couple of simple sine-wave oscillator circuits. As a first example, consider the *LC* resonant oscillator in Figure 5-21.

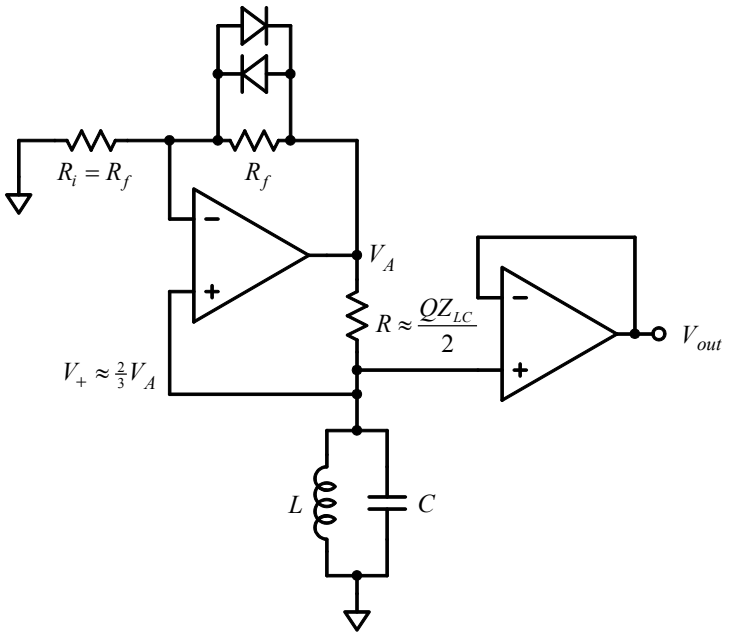


Figure 5-21: A simple *LC* resonant sine-wave oscillator. Its operation is explained in the text.

The oscillator output is taken (via the voltage follower) from the signal across the *LC* resonator, producing a low-distortion sine-wave output, even though the output of the nonlinear amplifier may be quite distorted.

Some positive feedback is required to sustain the circuit's oscillation; in this circuit it is generated by the voltage divider consisting of the resistor R and the parallel LC resonator. This feedback recirculates a fraction of the output of a noninverting nonlinear amplifier back to its input. At the resonant frequency $\omega_0 = 1/\sqrt{LC}$ the impedance of the LC resonator becomes very large: $QZ_{LC} = Q\sqrt{L/C}$; for a good-quality resonator, we would expect $Q \sim 150$ or more, and the LC component values should be chosen so that $Z_{LC} \sim$ 1k-10k. By choosing R to be about 1/2 of this resonant impedance, the positive feedback fraction is about

2/3 at frequency ω_0. At frequencies more than a couple ω_0/Q away from resonance, on the other hand, the LC impedance is $\sim \pm jZ_{LC}$ or less, so positive feedback at any other frequency is $\ll 1$ and has a 90° phase shift.

For sustained, constant amplitude oscillation, the gain around the feedback loop, including the amplifier, must be exactly 1. The nonlinear amplifier's gain for a small input signal is 2 when $R_f = R_i$, so for a small signal at ω_0 the loop gain is approximately $\frac{2}{3} \times 2 > 1$. Thus a small ω_0 signal will grow in amplitude exponentially quickly, whereas a signal at any other frequency will die away. As the ω_0 signal's amplitude grows, the nonlinear amplifier's gain will decrease until the amplitude is reached such that the loop gain is exactly unity, and the nonlinear amplifier will stabilize its output amplitude at that value. If the feedback resistors are each 1k, then the circuit's output amplitude will be $\sim$1V peak.

Because the diodes act to clamp the amplifier's peak output, the amplifier's output will take the form of a distorted sinusoid. The filtering action of the resonant RLC voltage divider, however, will provide a much more pure sinusoid across the LC pair, so that is where the oscillator's output is taken (using a voltage follower to isolate the resonator from its load). The distortion is minimized when the small-signal loop gain around the nonlinear amplifier at frequency ω_0 just barely exceeds 1, so trimming of the value of R may be necessary if really low distortion is required. Unfortunately, if the small-signal loop gain were to drop even a tiny bit below 1, then the circuit will cease to oscillate. If slightly higher distortion in the sine-wave output is tolerable, then the diodes in Figure 5-21 may be unnecessary: saturation of the amplifier op-amp output (V_A in Figure 5-21) will limit the oscillator's amplitude, but the signal at the LC resonator (where the output is taken) will still be a quite nice sinusoid.

Very similar in concept to the LC oscillator presented above, the *Wien bridge oscillator* uses RC pairs to form its resonant voltage divider as shown in Figure 5-22 on page 5-24 (a Wien bridge-type oscillator was the first product offered by the Hewlett-Packard company, back in 1939).

The two RC pairs in the Wien bridge circuit form a resonant voltage divider to provide some positive feedback to sustain oscillation. The resonant frequency is $\omega_0 = 1/RC$, but the filter's Q is only $1/3$ when the RC pairs are perfectly matched. The divider ratio is also $1/3$ at ω_0, so the nonlinear amplifier must have a small-signal gain of at least 3 for the circuit to oscillate. The diode clamping will severely distort the oscillator output if the amplifier's small-signal gain is even more than a few percent above 3, so some trimming of either R_f or R_i will be necessary to limit the distortion in the output waveform.

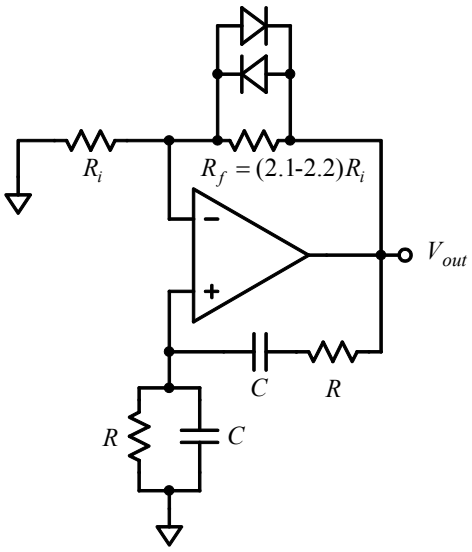


Figure 5-22: The *Wien bridge* oscillator uses two *RC* pairs to form its resonant voltage divider which provides positive feedback to sustain oscillation. The oscillation frequency is $\omega_0 = 1/RC$.

At ω_0 the positive feedback fraction is 1/3, so the small-signal gain of the nonlinear amplifier must be at least 3. The resonant divider's Q =1/3, so, because of the diode clamping, achieving a tolerable level of distortion will nearly always require trimming of the ratio R_f/R_i to just above 2.

An alternative (and much better) method of controlling the amplifier's gain is to use an incandescent light bulb (one with a glowing filament) in place of resistor R_i as shown at right; this is the method used in a truly low-distortion Wien bridge oscillator. This technique was invented by L. A. Meacham of Bell Laboratories in 1938 and was incorporated in the Hewlett-Packard product mentioned previously. As the bulb's filament heats up, its resistance increases, lowering the amplifier's gain. Because a change in the filament temperature is gradual (taking much longer than the oscillator's output period), it has essentially no effect on the output waveform shape. For our Wien bridge oscillator, the incandescent bulb used should have a resistance when cold of about 100Ω; the amplifier's feedback resistor should have a value about 2.5 times higher. With this version of the oscillator's amplifier, the output signal's distortion may be so low as to be very hard to measure.

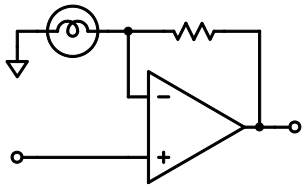


Figure 5-23: An amplifier whose gain is controlled by the temperature of the incandescent bulb used as the op-amp's R_i.

There are many other designs for op-amp sine-wave oscillators. The following article by Texas Instruments provides a nice summary and design procedures for several practical ones:

http://www.sophphx.caltech.edu/Physics_5/Useful_circuits/TI_Sine_Oscillators.pdf

Experiment 6

Transistors as amplifiers and switches

Copyright © Frank Rice 2013 – 2015
Pasadena, CA, USA
All rights reserved.

Experiment 6

Transistors as amplifiers and switches

Our final topic of the term is an introduction to the transistor as a discrete circuit element. Since an integrated circuit is constructed primarily from dozens to even millions of transistors formed from a single, thin silicon crystal, it might be interesting and instructive to spend a bit of time building some simple circuits directly from these fascinating devices.

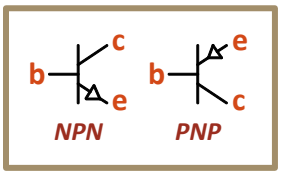
We start with an elementary description of how a particular type of transistor, the *bipolar junction transistor* (or *BJT*) works. Although nearly all modern digital ICs use a completely different type of transistor, the *metal-oxide-semiconductor field effect transistor* (*MOSFET*), most of the transistors in even modern analog ICs are still *BJT*s. With a basic understanding of the *BJT* in hand, we design simple amplifiers using this device. We spend a bit of time studying how to properly *bias* the transistor and how to calculate an amplifier's gain and input and output impedances.

Following our study of amplifiers, we turn to the use of the *BJT* as a *switch*, a fundamental element of a digital logic circuit. Single transistor switches are useful as a way to adapt a relatively low-power op-amp output in order to switch a high-current or high-voltage device on and off. These switches are also very useful to *translate* the output of an op-amp comparator to the proper 1 and 0 voltage levels of a digital logic circuit input. The final section of the text presents a few additional useful transistor applications including a discussion of a differential amplifier and a basic Class B power amplifier stage, both of which represent common building blocks of the operational amplifier.

THE BIPOLAR JUNCTION TRANSISTOR

What is a transistor and how does it work?

The *bipolar junction transistor* (or *BJT*) was invented at Bell Laboratories by William Shockley in 1948, the year after he, John Bardeen, and Walter Brattain invented the first working transistor (for which they were awarded the 1956 Nobel Prize in physics). It is constructed from a sandwich of three layers of doped semiconductor material, the thin middle layer being doped oppositely from the other two. Thus there exist two types of *BJT*: *NPN* and the *PNP*, whose schematic symbols are shown at right. The three layers are called the *emitter*, *base*, and *collector*, and their identification with the three schematic device terminals is also illustrated in the figure (note that the *emitter* is associated with the *arrow* in the schematic symbols). The base is the thin middle layer, and it forms one *PN* junction with the heavily-doped emitter and another with the lightly-doped collector. When used as an amplifier, the base-emitter junction is forward-biased,



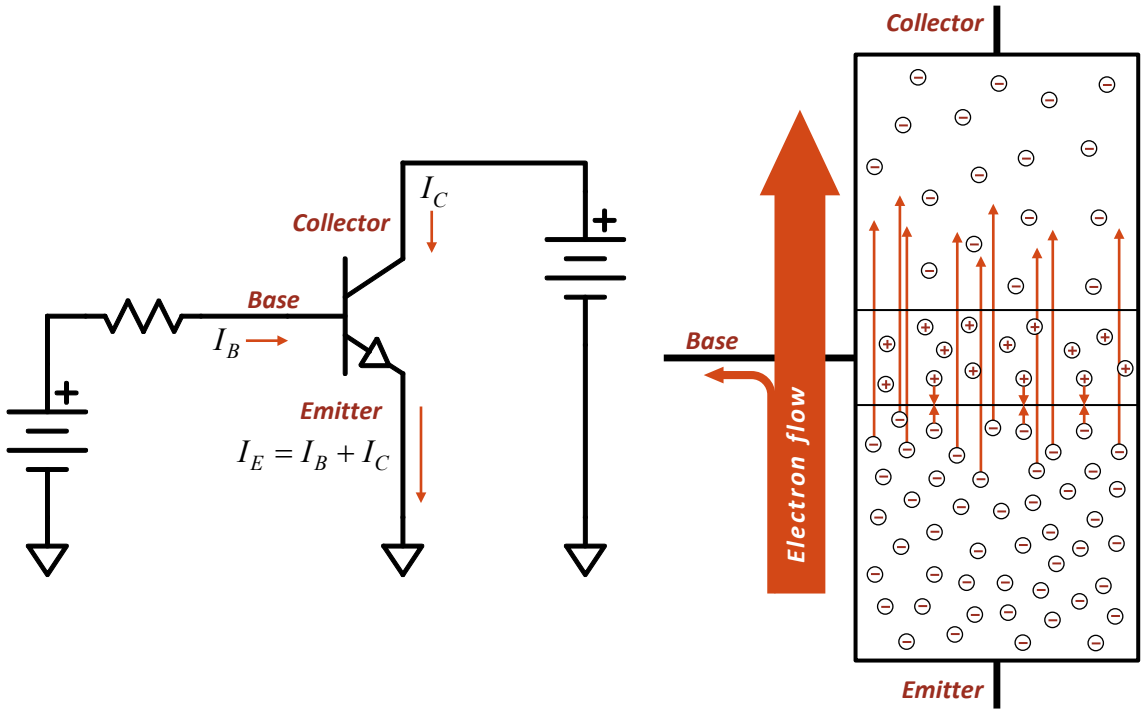


Figure 6-1: The inner workings of an *NPN* bipolar junction transistor. This transistor may be thought of as a "sandwich" of a thin *P*-type semiconductor layer (the *base*) between two *N*-type layers (the *emitter* and the *collector*). The emitter is very heavily doped with *N*-type charge carriers (conduction electrons). When the base-emitter *PN* junction is forward biased current flows from the base to the emitter. Because the base is very thin and the emitter is heavily doped, most of the emitter's charge carriers (electrons) which diffuse into the *PN* junction continue right on through it and into the collector. This results in a current flow from collector to emitter which can be much larger than the current flow from base to emitter.

whereas the collector-base junction is reverse-biased (in the case of an *NPN* transistor, the collector would be at the most positive voltage and the emitter at the most negative). The resultant charge carrier flows within the *NPN* transistor are illustrated in Figure 6-1.

Consider an *NPN* transistor's behavior as shown in Figure 6-1. When the base-emitter *PN* junction is forward-biased, a current flows into the base and out of the emitter, because this pair of terminals behaves like a typical *PN* junction semiconductor diode. The voltage drop from base to emitter is thus that of a typical semiconductor diode, or about 0.6–0.7V (nearly all transistors use silicon as their semiconductor). If the collector's potential is set more than a few 1/10ths of a volt higher than that of the base, however, an interesting effect occurs: nearly all of the majority charge carriers from the emitter which enter the base continue right on through it and into the collector!

The charge carriers from the emitter that enter the base become *minority* carriers once they arrive in the base's semiconductor material. Because the base is thin and the emitter has a large concentration of charge carriers (it is heavily doped), these many carriers from the emitter can approach the collector-base *PN* junction before they have a chance to recombine with the relatively few base majority charge carriers (holes in the case of an *NPN* transistor). Since the collector-base *PN* junction is reverse-biased, most of the charge carriers originally from the emitter can accelerate on through the base-collector junction and enter the collector, where they intermingle with the collector's lower concentration of similar charge carriers (Figure 6-1).

As a result, most of the emitter's charge carriers which enter the forward-biased base-emitter *PN* junction wind up passing through the base and entering the collector. This charge carrier flow out of the emitter is what comprises the current that flows through the emitter's device terminal on the transistor (I_E in Figure 6-1). The small fraction of these that recombine in the base then determines the current flow through the base terminal (I_B), whereas the much larger fraction passing through the base and entering the collector determines the collector current (I_C). If the base-emitter *PN* junction is not forward-biased, then majority carrier flow from the emitter across the junction does not occur, and the collector-emitter current vanishes (the only currents through the device are now the tiny reverse leakage currents from emitter and collector into the base).

The end result is that if we bias the transistor so that its base-emitter *PN* junction is forward-biased (base about 0.7V more positive than the emitter for an *NPN* transistor), then a small current flow in the base terminal can stimulate a much larger current flow in the collector terminal: the *BJT* transistor is a *current amplifier* (of course, the collector terminal must be connected to a power supply of some sort to complete an external circuit between collector and emitter as shown in Figure 6-1 — the power supply provides the energy required to move the current around this circuit).

The relationship between the base and collector currents: β (h_{fe})

It turns out that if the collector is at a potential of at least about 0.2V more than the base, then the ratio of the collector and base currents in a well-designed BJT transistor is remarkably independent of the magnitude of the base current and the collector-emitter voltage difference.

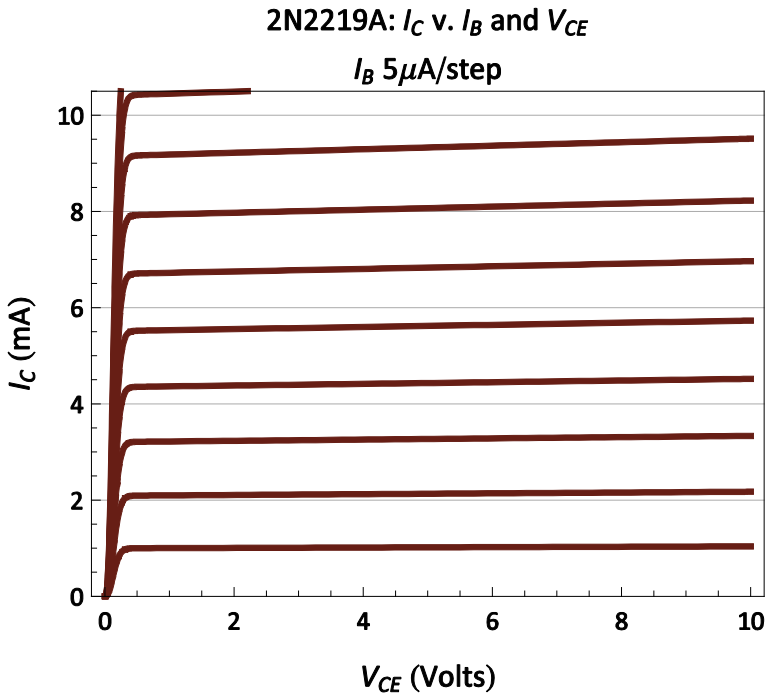


Figure 6-2: *Characteristic curves* of a typical *BJT*. The 2N2219A is a general-purpose, *NPN* transistor similar to the PN2222. Each curve shows the variation in collector current with the voltage between the collector and emitter for a fixed base current; base current was stepped through a range of values to generate the family of curves shown. The curves are nearly evenly-spaced and flat for V_{CE} > 0.8V, showing that the transistor's collector/base *current gain* (called β or h_{fe}) is nearly constant over the parameter space shown.

This behavior of a good *BJT* is strikingly illustrated by the family of collector *characteristic curves* plotted in Figure 6-2 for the *NPN* type 2N2219A transistor. Each curve in the family shows the collector current (I_C) as a function of collector-emitter voltage drop (V_{CE}) for a fixed value of base current (I_B); the base current is stepped by increments of 5μA to generate the various curves. Note how flat the various curves are for $V_{CE} > 0.8\,\text{V}$, showing that the collector current I_C is nearly independent of V_{CE} in this regime. Next note how evenly spaced the curves are (again, for $V_{CE} > 0.8\,\text{V}$), showing how very nearly proportional the collector and base currents are. The ratio $\Delta I_C / \Delta I_B$ defines the transistor's *current gain*, commonly designated variously by the symbols β (*beta*) and h_{fe} (the more "formal," engineering symbol for this parameter). Referring back to Figure 6-1, the current gain is clearly just the ratio of the number of emitter charge carriers which continue on to the collector to the number which recombine in the base. For the 2N2219A transistor shown in Figure 6-2, β averages about 250 (it rises a bit with increasing I_B); a well-designed,

general-purpose transistor will have $\beta \gg 1$. Unfortunately, for a given transistor β is a strong function of temperature, rising by about 10% for a 10°C rise in temperature; it also can vary by as much as 30% between transistors with the same model number. This possible variance in β should be considered when designing transistor amplifiers.

The transistor as an amplifier

Consider the circuit fragment shown at right, which includes an *NPN* transistor connected between two power supply "*rails*" V_{CC} and V_{EE} (with, naturally, $V_{CC} > V_{EE}$). Assume that some method has been used to *bias* the transistor's base terminal at the voltage $V_B > V_{EE}$ so that the transistor's base-emitter junction is forward-biased and conducting current I_B as shown (we'll discuss ways of biasing the transistor in a subsequent section). What we want to determine are the relationships between the various voltages and currents, the resistor values R_C and R_E, and the transistor's β.

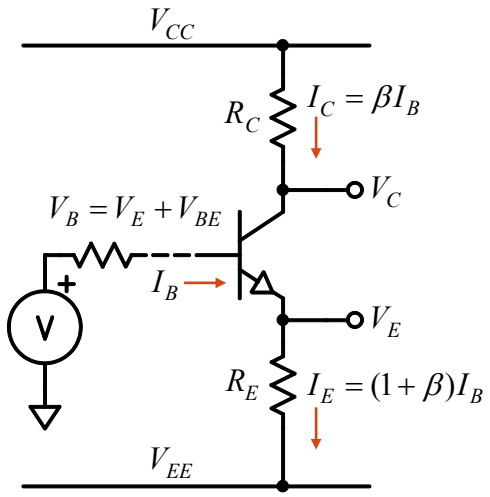


Figure 6-3: A collector-emitter circuit fragment of an *NPN* transistor used as an amplifier.

We start the analysis by connecting the base and emitter voltages using the forward-bias diode drop $V_{BE} \approx 0.6\text{V}$ across the base-emitter *PN* junction. As we know from our previous study of the semiconductor diode, this voltage will be a very weak function of the base current I_B, so we will use the working assumption that it is a fixed, constant value. Consequently, if we know V_B, then we also know V_E, and vice versa. Given V_E, we now know the voltage drop across the resistor R_E, so we also know the current through it: $I_E = (V_E - V_{EE})/R_E$.

Knowing I_E and the transistor's current gain β immediately tells us the other two transistor currents I_B and I_C, since the currents are related through β as shown in Figure 6-3. The value of R_C then gives us the collector voltage V_C, since $I_C = (V_{CC} - V_C)/R_C$. Thus we have succeeded in relating the circuit's state variables (currents and voltages), as we set out to do. Note that there are some conditions that must be met for our solution to be realistic: all the currents must flow in the directions shown by the arrows in Figure 6-3, and it must be the case that $V_{EE} < V_E < V_B \le V_C < V_{CC}$. If one or more of these conditions is violated by our solution, then our solution fails, and the transistor circuit is operating as a *switch* rather than as an *amplifier* (we'll discuss transistor switch circuits in the next section).

Now consider how V_E and V_C are affected by a small change dV_B in the transistor's base voltage. If the solution we have found is perturbed by changing the transistor base bias voltage from V_B to $V_B + dV_B$, then the emitter voltage will change by the same amount,

since we assume that V_{BE} is constant. Thus $dV_E = dV_B$, and the *small-signal voltage gain* of the circuit in Figure 6-3 from the base to the emitter of the transistor is:

6.1 Emitter follower voltage gain: $G \equiv dV_E / dV_B = 1$

A transistor amplifier circuit used in this way is called an *emitter follower*, analogous to the op-amp voltage follower circuit, since the voltage gain of the circuit is 1.

The change dV_E in the emitter voltage will change the emitter current as well because the voltage drop across R_E has changed, so $dI_E = dV_E / R_E$; this changes the collector current to $dI_C = dI_E / (1 + 1/\beta) \approx dI_E$, because $\beta \gg 1$ (the quantity $dI_C / dI_E = 1/(1 + 1/\beta)$ is sometimes referred to as the transistor's α; clearly for a good transistor $\alpha \approx 1$). Finally, dI_C changes the voltage drop across R_C; since the power supply voltage V_{CC} is constant, the change in voltage drop across R_C will change the collector voltage by $-R_C dI_C$, and the voltage gain of the circuit from the base to the collector of the transistor is:

6.2 Common-emitter voltage gain (β»1): $G \equiv dV_C / dV_B = -R_C / R_E$

This output configuration of the circuit in Figure 6-3 is called a *common-emitter amplifier*, an *inverting* amplifier whose gain formula resembles that of the op-amp inverting amplifier circuit. Note that just as in the case of the op-amp circuit, the ideal gain formula in (6.2) obtains as the transistor gain $\beta \to \infty$.

Next let us consider the input and output impedances of the transistor amplifier circuit fragment of Figure 6-3. The input impedance presented by the base terminal of the transistor is quite straightforward given the previous calculations. Since $Z_{in} = dV_B / dI_B$, $dV_E = dV_B$, and $(\beta + 1) dI_B = dI_E = dV_E / R_E$, we immediately see that the base input impedance is:

6.3 Base terminal input impedance: $Z_{in} = (\beta + 1) R_E \approx \beta R_E$

The output impedance at the terminal labeled V_C in Figure 6-3 is likewise easy to determine if we assume that the transistor's β is independent the collector-emitter voltage difference V_{CE} (in other words, if we assume that the characteristic curves shown in Figure 6-2 are flat and horizontal, as they very nearly are in that figure). In this case, since $Z_{out} = -dV_C / dI_{out}$ (where I_{out} is the change in the current drawn from the output terminal by some attached load), changing the load current cannot affect the current flow into the transistor's collector, because that current is determined solely by I_B. Thus any change in output current must pass through the collector resistor R_C, which therefore must be the circuit's output impedance (if the assumption regarding the characteristic curves is relaxed, then the reciprocal of the slope of the appropriate curve defines the *collector's dynamic output impedance*, and this impedance must be combined in parallel with R_C, slightly lowering the circuit's output impedance).

6.4 Common-emitter output impedance: $Z_{out} = R_C$

Finally, the emitter-follower output impedance at the terminal labeled V_E in Figure 6-3 must be determined: $Z_{out} = -dV_E/dI_{out}$. In this case, since $dV_E = dV_B$, a change in the emitter voltage will cause a corresponding change in the base voltage, which may in turn change the base bias current, I_B. Assume that the source of the bias for the transistor base has output impedance Z_S (the resistor shown in the base terminal circuit of Figure 6-3 between it and the constant bias voltage source V). The relationship between the base terminal current and voltage would then satisfy: $Z_S = -dV_B/dI_B$, so that the current out of the transistor's emitter would change as: $dI_E = (\beta+1)dI_B = -(\beta+1)dV_E/Z_S$. But this is not the whole story — dV_E will change the current through R_E by dV_E/R_E. Using Kirchhoff's current law, the sum of the change in current through R_E and at the output terminal dI_{out} must equal the change in current supplied by the emitter of the transistor, so:

$$-dV_E\frac{\beta+1}{Z_S} = dI_{out} + dV_E\frac{1}{R_E} \quad \rightarrow \quad -\frac{dV_E}{dI_{out}} = \left(\frac{1}{R_E} + \frac{1}{Z_S/(\beta+1)}\right)^{-1}$$

So the emitter follower's output impedance is the parallel combination of the circuit's *embedding impedance* seen by the transistor's base divided by $(\beta+1) \approx \beta$ and R_E, a combination which could result in a reasonably small output impedance:

6.5 Emitter follower output impedance: $Z_{out} = \left(\frac{1}{R_E} + \frac{1}{Z_S/\beta}\right)^{-1}$

Ways to design practical transistor amplifier circuits are discussed in a later section.

High gain amplifiers: the dynamic emitter resistance r_e

It seems from equation 6.2 that the common-emitter amplifier gain could be made arbitrarily large by reducing the emitter circuit resistor value R_E to zero (Figure 6-3). Of course, this turns out to be incorrect. One of the assumptions leading to (6.2) is that the base-emitter diode forward voltage drop V_{BE} is constant (independent of the transistor currents); this is, of course, inaccurate because it is an increase in V_{BE} applied to the base-emitter *PN* junction which drives an increase in the current through it. Returning to Experiment 3, page 3-32, it explained how the *diode equation* relates the voltage and current through a *PN* junction:

6.6 $$I_E = I_0(e^{q_e V_{BE}/k_B T} - 1)$$

We use the emitter current I_E in this equation because that quantity represents the total current through the base-emitter *PN* junction. The characteristic current $I_0 \sim 10^{-10}\,\text{mA}$; at room temperature and with $V_{BE} \sim 0.6\text{V}$, the exponential term is $\sim e^{24} > 10^{10}$, which is ridiculously larger than 1, so the −1 term in (6.6) may be very safely ignored! Thus the *dynamic emitter resistance*, $r_e \equiv dV_{BE}/dI_E$, near room temperature is:

6.7 Dynamic emitter resistance (300K): $$r_e = \frac{dV_{BE}}{dI_E} = \frac{k_B T}{q_e I_E} = \frac{26\Omega}{I_E/\text{mA}}$$

So if the emitter current is, for example, 2mA, then $r_e = 13\Omega$. The dynamic emitter resistance may be thought of as a small resistor added in series with the emitter terminal just inside the transistor so that the junction V_{BE} can still be treated as a constant — actual small variations in its value connected with changes in I_E are accounted for by a varying voltage drop across r_e. Consequently, the common-emitter amplifier gain when $R_E = 0$ is:

6.8 Common-emitter voltage gain limit (R_E = 0): $$G \equiv dV_C/dV_B = -R_C/r_e$$

An example of a high-gain common-emitter circuit will be presented later.

The transistor as a switch

If you apply a fairly large current to a transistor's base (a few mA, say), then $I_C = \beta I_B$ could be quite large. Consider the circuit in Figure 6-4, for example. The transistor's base-emitter junction is clearly forward-biased by the +5V V_{in} source; with $V_{BE} \approx 0.7\text{V}$, the base current would be

$$I_B = (5\text{V} - 0.7\text{V})/4.3\text{k}\Omega = 1\text{mA}$$

If the transistor $\beta \sim 150$, then we would expect a collector current of $I_C \sim 150\text{mA}$. But the collector power supply (V_{CC}) is only +5V, so with $R_C = 1\text{k}\Omega$ the most current that could possibly flow into the collector is $5\text{V}/1\text{k}\Omega = 5\text{mA}$. In this case the transistor will be driven into *saturation* by the large base current, and it will reduce its collector-emitter voltage drop (V_{CE}) to a fraction of a volt as the collector current is maximized (given the constraints imposed by V_{CC} and R_C). The transistor's operating state in this case is represented by the very right edge of the graph of the characteristic curves in Figure 6-2 on page 6-4; this region has been expanded in Figure 6-5. Note how a relatively small base current can drive this particular transistor's collector-emitter voltage drop to a very small value if the collector current is limited by external components to much less than βI_B.

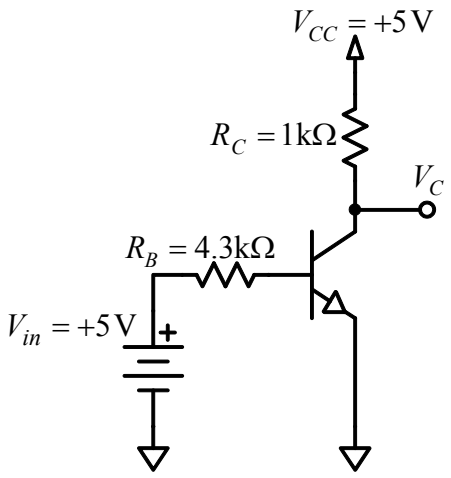


Figure 6-4: An *NPN switch* circuit.

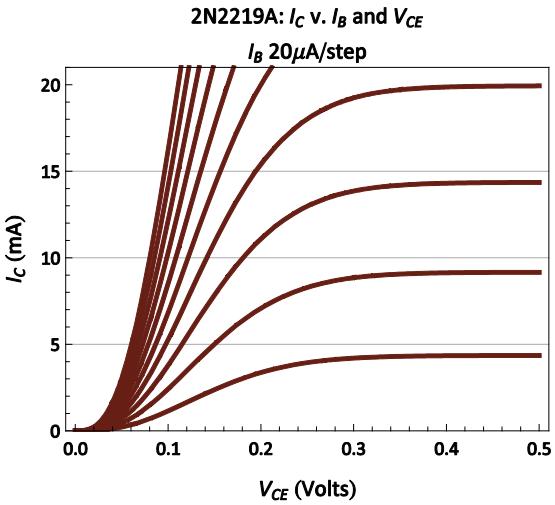


Figure 6-5: *Saturation region* of the transistor characteristic curves. The left-most curve corresponds to I_B = 0.2mA; note that even this relatively small base current will drive the collector-emitter voltage drop down to only about 0.1V even for collector currents as high as 20mA.

Ensuring transistor saturation

When using a transistor as a switch a good rule of thumb is to design the input circuit to the base so that I_B will be approximately 5% to 10% of the required I_C, ensuring that $I_C \ll \beta I_B$. This will drive the transistor well into its saturation region, minimizing V_C.

BASIC TRANSISTOR AMPLIFIER DESIGN

Designing a common-emitter amplifier stage

A simple and effective way to construct a transistor gain stage is to supply the transistor's base bias using a voltage divider and to *AC couple* the input and output signals as shown in Figure 6-6. The big advantage of this circuit is that it can be designed to work successfully almost completely independently of the transistor's gain β, so that it will work with nearly any available transistor and is very tolerant of circuit temperature and power supply voltage variations. In this section we will go through a design procedure for this circuit so that you can successfully assign the proper values to the resistors and capacitors in the circuit.

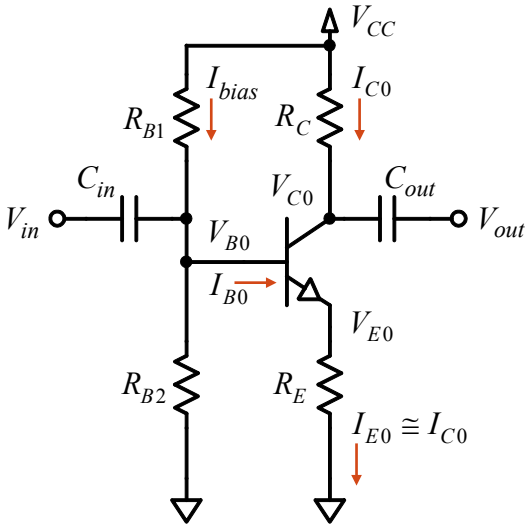


Figure 6-6: Basic *NPN* common-emitter amplifier stage. Component selection to establish the design stage gain and properly bias the transistor is discussed in the text.

Amplifier gain; the quiescent collector and emitter operating state

The first step is to set the required gain G of the stage. The circuit in Figure 6-6 is most effective for relatively modest gains, say $5 \le |G| \le 20$ or so; as the stage gain goes up, the circuit's input impedance will have to drop and the output impedance will rise, although these problems may be mitigated somewhat by a judicious choice of power supply voltage. If you need a large gain, it will probably require you to cascade several amplifier stages to achieve this result. In any event, the target gain for the amplifier stage sets the ratio of the collector and emitter resistor values through equation 6.2: $G = -R_C / R_E$.

The next step is to establish the value of V_{CC}, the power supply voltage you will use, and the transistor's *quiescent operating state*, target values of the collector and emitter DC voltages in the absence of an input signal: V_{C0} and V_{E0} (the subscript 0 implies that these are the quiescent, DC values of these circuit parameters). These choices will be driven by the circuit

gain G and the required peak-to-peak output voltage swing, V_{pk-pk}. To see how this works, the argument goes as follows:

(1) As described in the discussion leading up to the common-emitter gain equation (6.2), the ratio of the voltage drops across resistors R_C and R_E is equal to the magnitude of the gain $|G|$. Thus $|G| = (V_{CC} - V_C)/V_E$. This relationship should hold at all times, regardless of the signal output. When the output signal is at its positive peak, the voltages across the two resistors will be at their smallest values; when the output signal is at its negative peak, the difference $V_{CE} = V_C - V_E$ is at its minimum.

(2) To minimize distortion in the output of the amplifier, the design must keep the transistor from saturating; the voltage drop from collector to emitter V_{CE} should go no lower than about 1V (see Figure 6-2). Thus at the negative output peak $(V_C)_{\min} - (V_E)_{\max} = 1\text{V}$. Combining this with the above relation, we get:

$$\begin{aligned} (V_E)_{\max} &= \left(V_{CC} - 1\text{V}\right)/\left(|G| + 1\right) \\ (V_C)_{\min} &= \left(V_{CC} + |G| \times 1\text{V}\right)/\left(|G| + 1\right) \end{aligned}$$

(3) When the output is at its maximum, then the collector current I_C is at its minimum. To keep the range of collector current reasonable, this minimum current should be no less than about 5% of its maximum value, and this restriction would, of course, apply to the emitter current as well. This rule of thumb implies that $20 \times (V_E)_{\min} = (V_E)_{\max}$. The difference in these extremes times the gain G is the collector voltage range, which will determine the maximum V_{pk-pk}. Thus we have established a relationship between the gain G, power supply voltage V_{CC}, and the maximum output voltage swing V_{pk-pk}:

6.9

$$\left(1 + \frac{1}{|G|}\right) V_{pk-pk} = 0.95\left(V_{CC} - 1\text{V}\right)$$

The mean values of the extremes in V_C and V_E should be the target quiescent collector and emitter voltages, V_{C0} and V_{E0}, so that the full V_{pk-pk} output is available without hitting one of the voltage limits:

6.10

$$\begin{aligned} V_{E0} &= 0.525\left(V_{CC} - 1\text{V}\right)/\left(|G| + 1\right) \\ V_{C0} &= V_{CC} - |G| \times V_{E0} \end{aligned}$$

EXAMPLE COLLECTOR AND EMITTER VOLTAGE CALCULATIONS

Assume that you want a gain 5 common-emitter amplifier operating from a single +5V power supply. What would be the maximum available output voltage swing, and what should be the quiescent collector and emitter voltages V_{C0} and V_{E0} to achieve this?

Solution:

Solving equation 6.9 for V_{pk-pk} given the specified G and V_{CC}: $V_{pk-pk} = 3.17\text{V}$. The quiescent voltages are calculated using equations 6.10: $V_{C0} = 3.25\text{V}$; $V_{E0} = 0.35\text{V}$.

Quiescent operating currents and base bias voltage

As indicated by equation 6.4, the circuit in Figure 6-6 will have an output impedance equal to R_C. The value of this resistor and the quiescent collector voltage V_{C0} establish the quiescent collector and base bias currents, I_{C0} and I_{B0}:

$$\begin{aligned} I_{C0} &= \left(V_{CC} - V_{C0}\right)/R_C \\ I_{B0} &= I_{C0}/\beta \end{aligned} \qquad \textbf{6.11}$$

The base bias voltage, V_{B0}, must be established at 1 diode-drop (0.6V–0.7V) above the emitter voltage, V_{E0}. *This is the most important step*, since V_{B0} along with the values of R_C and R_E determine the transistor's quiescent operating state: V_{C0}, V_{E0}, and I_{C0}. Using equation 6.10,

$$V_{B0} = V_{E0} + 0.7\text{V} = 0.7\text{V} + 0.525\left(V_{CC} - 1\text{V}\right)/\left(\left|G\right| + 1\right) \qquad \textbf{6.12}$$

The base bias voltage divider; input impedance

To establish the critical parameter V_{B0}, the circuit in Figure 6-6 contains a voltage divider composed of resistors R_{B1} and R_{B2}. The quiescent base bias current I_{B0} will be supplied from the power supply V_{CC} through this divider, so we must carefully consider these resistor values. We want the circuit design to accommodate a fairly large variation in transistor β without significantly affecting the amplifier's performance; the only place β matters is in the second of equations 6.11. Calculate the maximum I_{B0} you might expect by picking a reasonable lower bound on β when using that equation. To keep this current from having a significant impact on the voltage divider, design it so that the current through R_{B2} is $\approx 10 I_{B0}$. By doing this you ensure that variations in β and thus I_{B0} will not significantly change V_{B0}, the amplifier's most important bias parameter. This consideration and the voltage divider equation determine the values of these resistors:

$$R_{B2} = \frac{V_{B0}}{10 I_{B0}}; \qquad R_{B1} = R_{B2}\left(\frac{V_{CC}}{V_{B0}} - 1\right) \qquad \textbf{6.13}$$

Finally, these values determine the amplifier's input impedance, which will be the parallel combination of R_{B1}, R_{B2}, and the transistor base input impedance, which from equation 6.3

is βR_E. Note that a large gain will tend to drive down the value of V_{B0} and thus R_{B2}, lowering the circuit's input impedance.

EXAMPLE BASE BIAS CALCULATIONS

Continuing our previous example calculations, if we choose $R_C = 750\Omega$, then that will also be the circuit's output impedance. For $G = -5$, the required $R_E = 150\Omega$. Assume that transistor $\beta \approx 150$ (minimum); the quiescent transistor currents will be (equations 6.11):

$$I_{C0} = (5\text{V} - 3.25\text{V})/750\Omega = 2.3\text{mA}$$
$$I_{B0} = 2.3\text{mA}/150 = 16\mu\text{A}$$

Now we can choose the base bias resistor values using equations 6.12 and 6.13:

$$V_{B0} = 0.7\text{V} + V_{E0} = 1.05\text{V}$$
$$R_{B2} = 1.05\text{V}/(10 \times 16\mu\text{A}) = 6.75\text{k}\Omega \rightarrow 6.2\text{k}\Omega$$
$$R_{B1} = 6.2\text{k}\Omega \times (5/1.05 - 1) = 23.3\text{k}\Omega \rightarrow 24\text{k}\Omega$$

Using 24kΩ and 6.2kΩ for R_{B1} and R_{B2} will introduce an error in the target V_{B0} of only 3%, which is less than the expected 5% resistor tolerances. The amplifier input impedance will be the parallel combination of the two bias resistors and βR_E, which will give only about 4kΩ.

Choosing the coupling capacitors

Since the DC quiescent bias voltages must be maintained at the transistor's three terminals for the amplifier in Figure 6-6 to operate properly, the input and output signals must be isolated from these DC voltages using the coupling capacitors C_{in} and C_{out}. Thus the amplifier will be *AC coupled* as originally described in Experiment 2 in the context of an op-amp amplifier. Consider the input capacitor C_{in} first. The low-frequency −3dB corner of the high-pass filter formed from C_{in} and the input impedance of the amplifier will be:

6.14
$$f_{in} = \frac{1}{2\pi Z_{in} C_{in}} = \frac{1}{2\pi C_{in}} \left(\frac{1}{R_{B1}} + \frac{1}{R_{B2}} + \frac{1}{\beta R_E} \right)$$

The output capacitor C_{out} will form another high-pass filter with the amplifier's output impedance and the input impedance of the load attached to the amplifier output:

6.15
$$f_{out} = \frac{1}{2\pi (Z_{out} + Z_{load}) C_{out}} = \frac{1}{2\pi (R_C + Z_{load}) C_{out}}$$

These two high-pass filters are cascaded, so the circuit's over-all low-frequency response will be determined by the product of the two filters' responses.

Simple high-gain amplifier stage

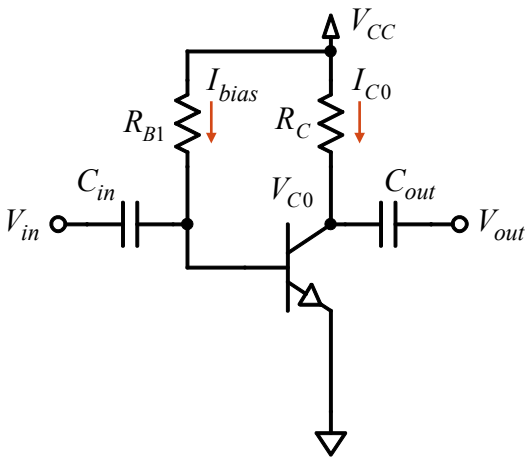


Figure 6-7: High-gain *NPN* common-emitter amplifier stage.

Although the high-gain amplifier circuit shown in Figure 6-7 is much simpler than the previous amplifier design, it will be quite dependent on the actual transistor β for its DC bias conditions and its resulting gain and output voltage range. The absence of an emitter resistor means that the transistor's dynamic emitter resistance r_e will determine the circuit gain along with the collector resistor R_C (using equations 6.7 and 6.8 on page 6-8).

The design process is similar to that used for the common-emitter amplifier of the last section, but is considerably simplified. Assuming that the transistor β is reasonably large, then we can substitute I_{C0} for I_E in equation 6.7; substituting for r_e in equation 6.8 results in an equivalent expression for the circuit gain:

6.16
$$-G = \frac{R_C}{26\Omega} \times \frac{I_{C0}}{\text{mA}}$$

The gain is completely determined by the voltage drop across R_C. Since $R_C I_{C0} = V_{CC} - V_{C0}$, we can express the gain in the equivalent form:

6.17
$$-G = \frac{V_{CC} - V_{C0}}{0.026\text{V}}$$

Again, the circuit output impedance is R_C; choosing this value and the gain G determines I_{C0}. Now use the transistor β to calculate $I_{B0} = I_{C0}/\beta$. The transistor emitter is connected directly to ground, so the base bias is one diode-drop higher: $V_{B0} = 0.7\text{V}$. These two values determine the base bias resistor: $R_{B1} = (V_{CC} - V_{B0})/I_{B0}$. Substituting from the above equations,

6.18
$$\frac{R_{B1}}{R_C} = \frac{\beta}{|G|} \times \frac{V_{CC} - 0.7\text{V}}{0.026\text{V}}$$

Note that equation 6.18 also shows that once you have chosen values for R_{B1} and R_C, the amplifier's actual gain will be proportional to the transistor β. If you build a high-gain amplifier, the actual transistor β will most likely deviate from the value assumed by your design; you may have to adjust the value of either R_{B1} or R_C to trim the measured V_{C0} to achieve your design target.

Finally, the input impedance will be the parallel combination of R_{B1} and βr_e; the coupling capacitors will then determine the circuit's low-frequency performance in the same way as was the case for the previous amplifier design (as with equations 6.14 and 6.15).

The input stage of an op-amp (in its most basic form) consists of a two-transistor differential amplifier stage; it acts as a low-gain common emitter stage for a common-mode input ($R_C / R_E \sim 1$ if the same voltage is applied to both the + and − inputs), but acts in a high-gain amplifier configuration for a differential voltage input. The design of a simple differential amplifier using two *NPN* transistors is presented in the section starting on page 6-24.

TRANSISTOR SWITCH CIRCUITS

Basic switching circuits

The use of a transistor as a switch was discussed earlier; the gist of that presentation is that when acting as a switch the transistor is either: (1) *off*, because the base-emitter junction does not have a sufficient forward-bias voltage applied to it; or (2) *saturated*, because a large base-emitter forward-bias current is applied.

Consider the use of an NPN transistor switch to apply power to a load such as an LED or relay as in Figure 6-8. By driving the transistor into saturation, nearly the entire power supply voltage V_{CC} is available to supply the load connected between the collector and the power supply; relatively large currents may be passed through the transistor collector without causing overheating because $V_{CE} < 1\text{V}$ when the transistor is saturated (the PN2222A transistor, for example, can easily handle collector currents of a few hundred milliamps).

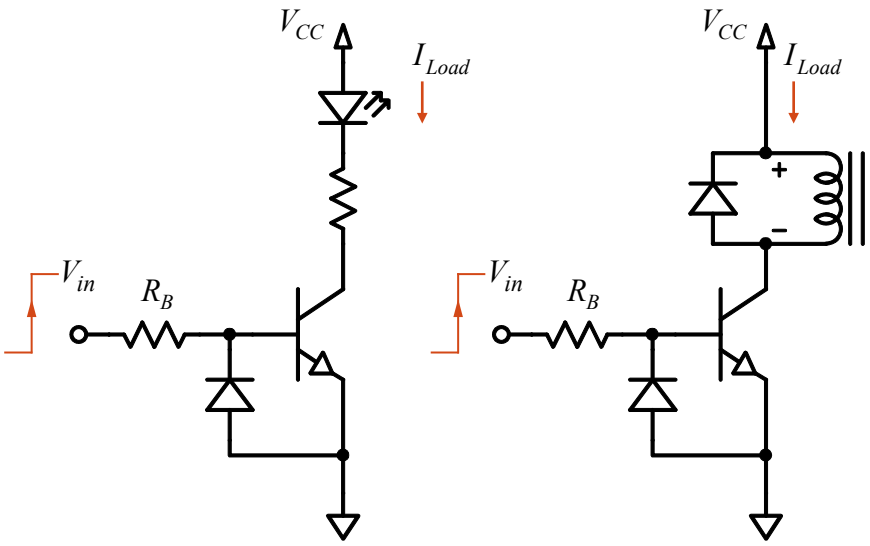


Figure 6-8: Examples of circuits using an *NPN* transistor switch: driving an *LED* or a relay coil. Bright *LED*s may require currents of several tens of mA at relatively low voltages; the same is true for some popular relays. The 10 mA output of the TL082 op-amp is not suitable for such loads, so using a transistor switch is a possible solution. The diodes across the transistors' base-emitter junctions prevent negative input voltages from causing reverse breakdown of the base-emitter *PN* junction. Note that these diodes are reverse-biased for positive input voltages, but conduct if the base voltage reaches –0.6V. The reverse-biased diode across the relay coil in the right-hand circuit prevents the large *back-EMF* from the coil's inductance from damaging the circuit when the transistor is turned off.

The base-emitter junction of the transistor has a relatively small reverse-breakdown voltage (only 6V for the PN2222A), so adding a reverse-biased diode across the transistor's base-emitter junction as shown in Figure 6-8 serves to protect the transistor from excessive negative input voltages (such as a saturated op-amp comparator output). The diode begins to

conduct as the transistor base voltage reaches –0.6V, keeping the base well away from its reverse breakdown voltage limit.

Using the ***Ensuring transistor saturation*** rule of thumb makes switch circuit design straightforward, as in the following example:

EXAMPLE: SWITCHING AN *LED*

Assume that you need to switch a 30mA green *LED* using the output of a TL082 op-amp Schmitt trigger circuit. The *LED* will have a forward voltage drop of about 2.2V when conducting 30mA, and you wish to use a 5V power supply for V_{CC}. According to the PN2222A transistor data sheet, the saturation $V_{CE} < 0.1\text{V}$ for $I_C = 30\text{mA}$ and $I_B = 3\text{mA}$ (10% of I_C as recommended by our design rule of thumb). Consider the left-hand circuit in Figure 6-8; assuming that the op-amp comparator's positive output voltage is 11V (V_{in}) when supplying a 3mA load (the switch transistor's base circuit is the comparator's load), then the base resistor value may be calculated to set this current to the desired 3mA:

$$R_B = (11\text{V} - 0.7\text{V})/3\text{mA} = 3.43\text{k} \rightarrow 3.3\text{k}$$

The resistor R_B will dissipate less than 0.03W when conducting 3mA, so a 1/4W resistor will work. The resistor in series with the *LED* load may also be calculated by considering the voltage drop it must have when conducting the 30mA *LED* current:

$$R_{LED} = (5\text{V} - 2.2\text{V})/30\text{mA} = 93\Omega \rightarrow 91\Omega$$

This resistor will dissipate under 0.1W when the *LED* is illuminated, so a 1/4W resistor will work fine here also.

Generating digital logic levels from analog signals; simple logic circuits

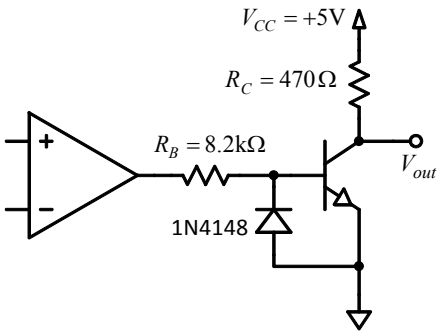


Figure 6-9: A transistor switch added to a TL082 op-amp comparator output may be used to translate the op-amp's saturated output levels of ±11V to the 0V-5V levels used by many logic circuits.

A common digital logic standard uses +5V for a logic "1" signal and 0V for a logic "0" (actually, "1" may be in the range 2.0V–5.0V and "0" is 0V–0.8V, depending on the

particular *logic family* to which you are interfacing). Using a transistor switch to generate these levels is straightforward (Figure 6-9). See if you can confirm that the circuit will work properly. Note that the output of the switch is inverted: positive op-amp saturation corresponds to $V_{out} \approx 0\text{V}$, negative op-amp saturation to $V_{out} = 5\text{V}$. With the selected value $R_C = 470\,\Omega$, the current through R_C and the transistor's collector will be approximately 10mA when the transistor is saturated (on).

R_C in Figure 6-9 is called a *pull-up resistor* because it "pulls" the output voltage up to V_{CC} when the transistor turns off. Since V_{CC} is supplied to the switch's output load through this pull-up resistor, you must make sure that the current drawn by the load does not cause the output voltage to *droop* below the required logic "1" voltage level (R_C is the circuit's *output impedance* when in the logic high state); this will usually not be a problem if the load is an input to a member of the *TTL* or *CMOS* logic families. If the load does require a significant amount of current when the switch output is high (more than 1 or 2 mA), then a simple solution is to use a switch constructed from a PNP transistor.

The requirement to interface an op-amp comparator's output to the voltage levels required by digital logic circuitry is so common that entire families of special integrated circuits to perform this task have been developed by nearly all analog device manufacturers: the *comparator* ICs. One of the first truly successful comparator ICs was introduced many years ago by National Semiconductor (now a part of Texas Instruments): the LM311 (data sheet) which incorporates a flexible transistor switch into its comparator output as shown in Figure 6-10. Although venerable, the LM311 is still a useful device to include in your design "toolbox."

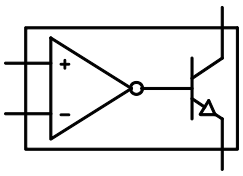


Figure 6-10: The LM311 Comparator IC (introduced in 1969) incorporates a transistor switch with an op-amp comparator. As can be seen in the simplified functional schematic, both the output transistor's collector and emitter are accessible for incorporation into a circuit design; the output transistor is turned on (transistor collector shorted to its emitter) whenever the *–Input* voltage exceeds the *+Input* voltage.

PRELAB EXERCISES

1. You are required to design a common-emitter amplifier (Figure 6-6 on page 6-10) with an inverting gain of 10 using a single +12V power supply ($V_{CC} = +12\text{V}$), and with an output impedance of 1kΩ.
   a. What should be the values of R_C and R_E?
   b. Using the design rules in the text, what is the maximum achievable output peak-to-peak voltage swing?
   c. What should be the design quiescent collector and emitter voltages (V_{C0} and V_{E0})?

2. Continuing the above common-emitter amplifier design problem,
   a. What will be the quiescent collector current I_{C0}? If the minimum expected transistor $\beta = 150$, then what will be I_{B0}, and what should be the minimum current through resistor R_{B2} (Figure 6-6)?
   b. What is the target value for V_{B0}? Of the following pairs of values, which pair would make the best choice for R_{B1} and R_{B2}:

      (130 kΩ and 150 kΩ), (91 kΩ and 10 kΩ), (30 kΩ and 3.3 kΩ), (11kΩ and 1.3 kΩ)?
   c. What will be the amplifier's input impedance?

3. Concluding the common-emitter amplifier design problem,
   a. What will be the low-frequency −3dB corner of the amplifier's input if $C_{in} = 0.1\mu\text{F}$?
   b. What will be the low-frequency −3dB corner of the amplifier's output if $C_{out} = 0.01\mu\text{F}$ and the load impedance $Z_{load} = 100\,\text{k}\Omega$?

4. Draw a complete schematic of the amplifier you've designed (Figure 6-6) with all component values specified. You will use a PN2222A transistor for the amplifier.

5. Consider a high-gain amplifier circuit (Figure 6-7 on page 6-14) with $R_C = 1\text{k}\Omega$ and $V_{CC} = +5\text{V}$. If the gain of the amplifier is to be $G = -100$ and the transistor $\beta = 200$, then what should be the value of R_{B1}? What will be the values of I_{C0} and V_{C0}?

LAB PROCEDURE

Common-emitter amplifier

Construct the common-emitter amplifier you've designed in Prelab exercises 1 through 4. Be sure you use the PN2222A transistor data sheet to properly identify the transistor's emitter, base, and collector leads.

This design calls for $V_{CC} = +12\text{V}$ and $V_{EE} = 0\text{V}$ (ground). With no input signal, measure the transistor's terminal bias voltages V_{C0}, V_{E0}, V_{B0}. How do they compare to your design values?

Apply an input signal and measure the amplifier's gain at 15 kHz. Increase the input amplitude until you find the amplifier's maximum output peak-to-peak voltage range. Find its low-frequency −3dB corner.

This amplifier has a $Z_{out} = R_C = 1\text{k}\Omega$. For a high-frequency application driving a BNC cable, the cable's $\sim 100\,\text{pF/meter}$ capacitance will form a low-pass filter with the amplifier's 1kΩ Z_{out}, limiting its high-frequency response. Check this statement by connecting the amplifier output to the DAQ using a BNC cable about 5 ft long. Use the DAQ to measure the amplifier's frequency response over the range 500Hz – 2MHz and determine its high-frequency −3dB corner. What is the amplifier's gain at 1 MHz? To correct this situation the amplifier needs a lower output impedance, which we can accomplish by adding an emitter follower output stage.

Adding an emitter follower output stage

An emitter follower amplifier stage has a gain of +1 and a low output impedance (equations 6.1 and 6.5). Add an emitter follower stage to the common-emitter amplifier to lower its output impedance and thus improve its high frequency performance (the modified circuit is shown in Figure 6-11 on page 6-21).

Consider the operation and biasing of the 2-stage amplifier circuit in Figure 6-11. The biasing and operation of the common emitter amplifier (Q_1 and its biasing resistors) is just the same as for the original single-transistor amplifier. The base bias of the emitter follower transistor (Q_2) is set by Q_1's collector circuit; the base bias current Q_2 draws from that circuit should be much smaller than Q_1's collector current (5 mA), so it should have only a small effect on the high-gain stage operation. Thus the base bias voltage of Q_2 is set to equal the collector bias voltage of Q_1, which should be about 6.75V according to the solution to prelab exercise 1.c.

The emitter voltage of Q_2 should be approximately 0.7V less than this, or about 6V. The Q_2 emitter bias current is then set by the value of R_E; choosing $R_E = 3.0\text{k}$ will establish a reasonable bias current of around 2 mA. Assuming the transistor $\beta \sim 150$, what is the base current Q_2 would then draw from Q_1's collector circuit? Should this current indeed have

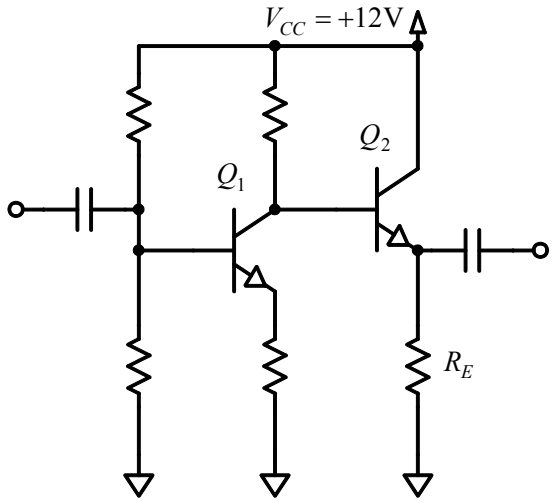


Figure 6-11: Adding an emitter follower output stage to the common emitter transistor amplifier circuit. The added components are transistor Q_2 and its emitter resistor R_E. The output coupling capacitor is moved from the collector of Q_1 to Q_2's emitter.

only a small effect on Q_1's biasing as we assumed? What should be the output impedance of this emitter follower stage? By about what factor should this raise the high-frequency cutoff of the (amplifier + BNC + DAQ) combination (actually, another factor also limits the high-frequency performance of the amplifier: the capacitance of the transistor's base-collector *PN* junction.

Build and test the operation of this 2-stage amplifier. Use the DAQ to measure the modified amplifier's frequency response over the range 500Hz – 2MHz. What is the amplifier's gain at 1 MHz? How does this compare to the original, single-stage amplifier's performance?

High-gain amplifier

Construct the high-gain, NPN transistor amplifier (Figure 6-7 on page 6-14) with the specifications provided by Prelab exercise 5. Remember, this new circuit calls for $V_{CC} = +5\text{V}$. Measure V_{C0} and adjust the value of R_C to better achieve your design value. Calculate a new gain estimate from your final V_{C0} value (see equation 6.17).

The gain of this amplifier is large (~ 100), but its maximum output voltage range is only $\sim 4\,\text{V}$ pk-pk ; make sure that you use an input signal which is small enough that the amplifier output is not saturated. Measure the amplifier's gain at 10 kHz and determine its low-frequency −3dB corner.

Increase the input amplitude until you find the amplifier's maximum output peak-to-peak voltage. Does the output waveform appear distorted even when you are below this maximum? What could be causing this behavior?

Transistor switch

Construct a transistor switch to illuminate an LED with a 30mA current from the +5V power supply (see Figure 6-8 on page 6-16 and the Example: Switching an *LED* on page 6-17) when driven by a saturated op-amp output such as from the Schmitt trigger circuit of Experiment 4, Figure 4-3 on page 4-3. Don't forget to include the base-emitter protection diode shown in Figure 6-8 – use a 1N4148 or 1N914 for this diode.

Additional circuits

If you have time, attempt to build one or more additional circuits and test their operation. Use examples from the ***More circuit ideas*** section, from previous experiments, or try one of your own design.

Lab results write-up

As always, include a sketch of the schematic with component values for each circuit you investigate, along with appropriate oscilloscope screen shots. Make sure you've answered each of the questions posed in the previous sections.

MORE CIRCUIT IDEAS

Phase Splitter

Figure 6-12 shows a simple circuit from Horowitz and Hill (see references) which combines common-emitter and emitter follower ideas to provide two equal-amplitude, opposite-phase outputs. By choosing $R_C = R_E$ (and both $\gg r_E$), the gain to each output is 1, but the collector and emitter voltages will vary out of phase with each other. Shown here using bipolar (±12V) power supplies, the voltage swing from either output can nearly reach ±6V as the transistor operating state varies from cutoff to saturation. Of course, the base bias resistors R_{B1} and R_{B2} must be chosen to set the base voltage to $V_{E0} + 0.7\text{V} = -5.3\text{V}$, and the coupling capacitors' values set the circuit's low-frequency cutoff.

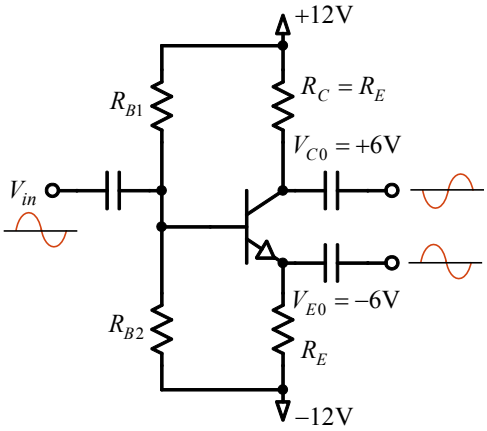


Figure 6-12: Phase splitter. Outputs are 180° out of phase.

Using PNP transistors

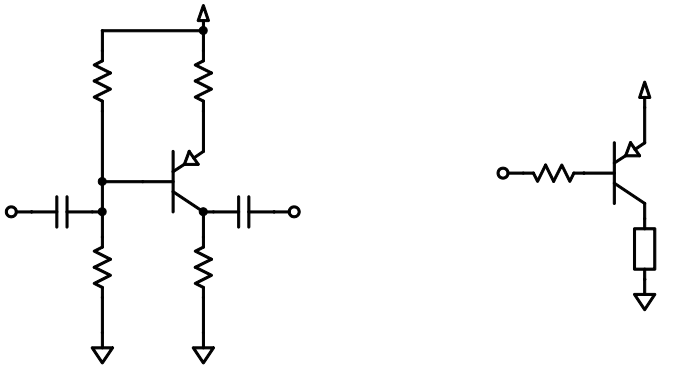


Figure 6-13: *PNP* versions of a common-emitter amplifier (left) and a transistor switch (right). The power supply voltage is >0 and is indicated by the upward-pointing arrows in the circuits. In the case of the switch, pulling the input *low* (to ground) will turn the transistor on and connect its load to the power supply. Note that unlike the *NPN* switches shown in Figure 6-8, the load of a *PNP* switch may have one terminal connected to ground, and a large current may then be supplied when the output is *high*.

In a *PNP* transistor the roles of conduction electrons and holes in the various regions of the transistor's semiconductor material are reversed; consequently the externally-applied bias voltages and resulting currents change sign from those for the *NPN* transistor. Figure 6-13 shows *PNP* versions of amplifier and switch circuits; note that they are derived from the corresponding *NPN* circuits by reflecting those circuits vertically (top for bottom) and

changing the transistor type. The type 2907 transistor may be used as a *PNP* counterpart of the type 2222 *NPN* transistor.

A differential amplifier stage

A differential amplifier suitable for use as the input stage of a simple op-amp is shown in Figure 6-14. Two matched transistors are arranged in identical common-emitter configurations which share the emitter resistor R_E. The output is taken from one of the two amplifiers, whose input then becomes the *–Input* of the differential stage. The object of the amplifier circuit is to have a high gain for a differential input signal, $V_{in+} - V_{in-}$, but have a much smaller output in response to a common-mode signal, $V_{in+} = V_{in-}$. In addition, the amplifier should be able to respond to DC (or very slowly changing) inputs, so it must be *directly coupled* — no capacitors may be used to isolate the amplifier stage's inputs from its required DC bias voltages.

Consider the design shown in Figure 6-14. Note that bipolar (both + and −) power supplies are used so that the transistor bases can still be biased properly when they are directly connected to 0V low-impedance inputs (grounds). This characteristic will satisfy the last criterion listed above: direct coupling of the input stage to the signal inputs. When both inputs are connected to ground, the resultant quiescent operating state currents and voltages are indicated in the right-hand diagram in Figure 6-14. Since both halves of the circuit are identical, so will be the corresponding currents and voltages on the two sides.

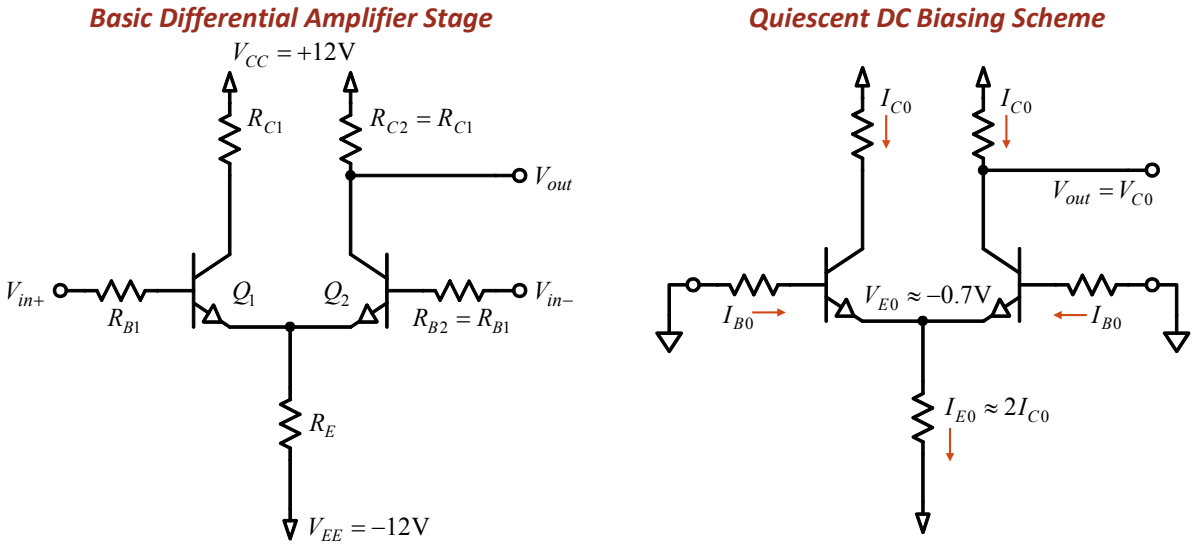


Figure 6-14: A basic *NPN* transistor differential amplifier stage, suitable for use in a simple operational amplifier design. Component and terminal identifications are shown on the left, quiescent bias currents and voltages on the right (for 0V DC inputs).

A good operational amplifier should have a high input impedance and small input bias currents; this may be accomplished by choosing a large value for R_E: a few MΩ will suffice for this simple circuit. Because the voltage drop across R_E is approximately equal to V_{EE}, the

bias current through I_{E0} will then be a few microamps; the currents into the two transistor bases will be $\sim\beta$ times smaller, so the input bias currents should be less than 10^{-7} amps. The base resistors R_{B1} and R_{B2} are required to provide some protection to the transistors in the event that the differential input voltage is large; they should have values of ~10k, so the voltage drops across them due to the input bias currents will be on the order of a millivolt or less. Thus we can ignore them for now. The collector resistors, R_{C1} and R_{C2}, set the collector bias voltages and will determine the circuit's differential and common-mode gains. Since the combined emitter bias current I_{E0} will be divided equally by the two identical halves of the circuit, we get $V_{C0} = V_{CC} - (I_{E0}/2)R_{C2}$: this will also be the quiescent output voltage with both inputs grounded.

To determine the behavior of the differential amplifier when nonzero input signals are present, we analyze its response to two different input conditions: $V_{in+} = V_{in-}$ (a common-mode input signal) and $-V_{in+} = V_{in-}$ (a differential input signal). By linear superposition of these two results we can predict the circuit's response to any arbitrary combination of V_{in+} and V_{in-} signals.

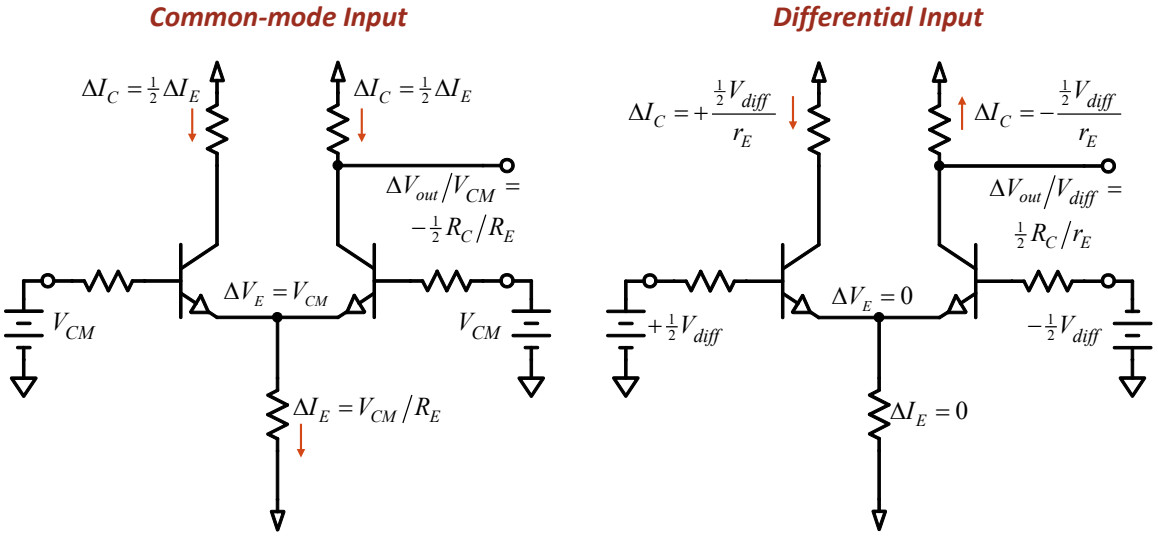


Figure 6-15: Responses of the differential amplifier stage to common-mode (left) and differential (right) inputs. The currents and voltages shown are the changes to the quiescent operating state values shown in Figure 6-14.

Consider the common-mode input first, shown in the left-hand diagram of Figure 6-15. In this case the same voltage is added to both transistor bases, and, since the base-emitter voltage drop of 0.7V is maintained, the voltage across the emitter resistor must change by the same amount. The resulting increase in the emitter resistor current, ΔI_E, is divided equally between the two transistors because of the symmetry in the circuit. Ignoring the small base currents, half of ΔI_E then appears across each of the collector resistors, and the resultant change in the voltage drop across R_{C2} appears at the circuit's output. Thus the common-mode voltage gain $g_{CM} = -\frac{1}{2}R_C/R_E$, as shown in Figure 6-15.

The gain of the circuit to a differential input may be similarly analyzed using the right-hand diagram in Figure 6-15. In this case equal and opposite voltages are applied to the inputs, resulting in equal and opposite changes in the transistors' base biases. Thus the emitter currents will also change in opposite directions, and the total current through R_E remains constant (to first order in the assumed small differential input voltage). This implies that the emitter voltage of each transistor doesn't change, so we have a situation identical to that of the high-gain common-emitter amplifier in Figure 6-7. Each transistor input sees one half of the total differential input voltage, and the gain of each is determined by the ratio of its collector resistor to its dynamic emitter resistance: $r_E = (26\,\Omega\,\text{mA})/I_{C0}$. Consequently, the differential voltage gain $g_{diff} = \frac{1}{2} R_C / r_E = \frac{1}{2} I_{C0} R_C / 26\text{mV} \approx 20 \times (V_{CC} - V_{C0})$ (voltages in volts). Typically, with $V_{CC} \sim 10\text{V}$, we would get $g_{diff} \sim 100$.

The *common-mode rejection ratio* (*CMRR*) of a differential amplifier stage is defined to be the ratio of these two gains: $CMRR \equiv \left| g_{diff} / g_{CM} \right|$. For the case of our simple differential amplifier stage, Figure 6-14 and Figure 6-15,

$$CMRR = \frac{R_E}{r_E} = \frac{I_{C0} R_E}{26\,\text{mV}} = \frac{\frac{1}{2} I_{E0} R_E}{26\,\text{mV}} \approx 20 \times (|V_{EE}| - 0.7)$$

where V_{EE} is measured in volts. For our example, $CMRR \approx 20 \times 11.3 \approx 226 = 47\,\text{dB}$, not bad for two transistors and a few resistors. This ratio can only be achieved in practice, however, if the two halves of the circuit are very well matched; if not, then the two collector currents will differ somewhat, and our assumption of perfect symmetry between the two halves of the circuit will be invalid: because of the mismatch, a common-mode input will generate a small differential signal in the two transistors' emitter circuits as well, and this error will reduce the *CMRR*.

Boosting op-amp output power

If you need your circuit output to supply more than about 10mA, the standard TL082 op-amp's output is insufficient. To quickly fashion together a "power op-amp" which can supply higher currents (100mA or so), you can add an emitter follower transistor amplifier to the op-amp's output. To enable this added amplifier to supply both positive and negative output voltages into a load, we can use both *PNP* and *NPN* transistors in a so-called *push-pull* emitter follower circuit as shown in Figure 6-16.

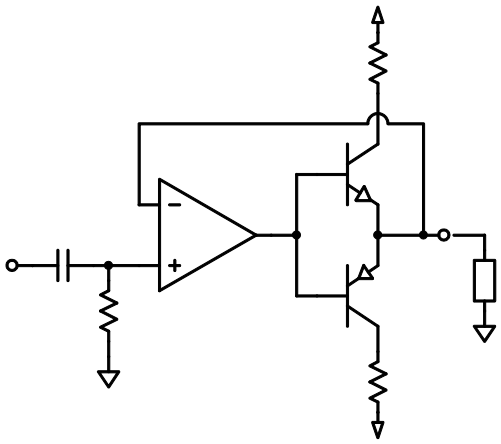


Figure 6-16: A simple *push-pull* emitter follower circuit uses a *complementary pair* (*NPN* and *PNP*) of transistors to boost the output current of an op-amp. Note that the negative feedback is from the transistors' output (the little hump in the feedback wire as it crosses below the resistor means that there is no connection between the crossing wires). The collector resistors reduce the transistors' power dissipation when driving a low-resistance load.

Because (1) the feedback loop includes the transistors and (2) the op-amp adjusts its output voltage to keep its two inputs equal, the output voltage applied to the load should follow the input voltage to the op-amp +*Input*. When the input is positive, then the op-amp raises its output voltage until the upper, *NPN* transistor turns on, supplying current to the load from the positive power supply. Since this requires that the base-emitter junction is forward-biased at about 0.7V, the op-amp output voltage must be maintained 0.7V higher than the voltage applied to the load. This reverse-biases the lower, *PNP* transistor base-emitter junction, so that transistor is turned off. Conversely, when the input voltage goes negative, the op-amp slews its output to 0.7V below the desired load voltage, activating the lower, *PNP* transistor and turning off the *NPN* transistor. Thus the two transistors alternate between inactive and conducting on alternate half-cycles of the input waveform, a mode of operation commonly referred to as *Class B* (the amplifiers presented earlier are all *Class A*: the transistors are biased on during all normal input and output conditions).

One significant drawback of this simple, Class B power amplifier is that it introduces *crossover distortion*: glitches in the output waveform as it passes through 0. Because each transistor requires that its base-emitter junction is forward-biased by approximately 0.7V for

it to operate, there is a "dead zone" in the circuit's output whenever the op-amp is within about ±0.7V of 0. The op-amp's output must pass through this zone as rapidly as possible whenever the input passes through 0, but it is limited by its slew rate (about 13 V/µs for the TL082). Consequently, the output waveform has little "flat spots" at 0V as the op-amp slews through the required 1.4V to activate the other output transistor (Figure 6-17). A more complicated design (as used in an actual op-amp's output stage) includes some base biasing so that both transistors are properly biased when the output is near zero, eliminating crossover distortion. This modified type of amplifier designated *Class AB*, because it is a hybrid combination of classes A and B.

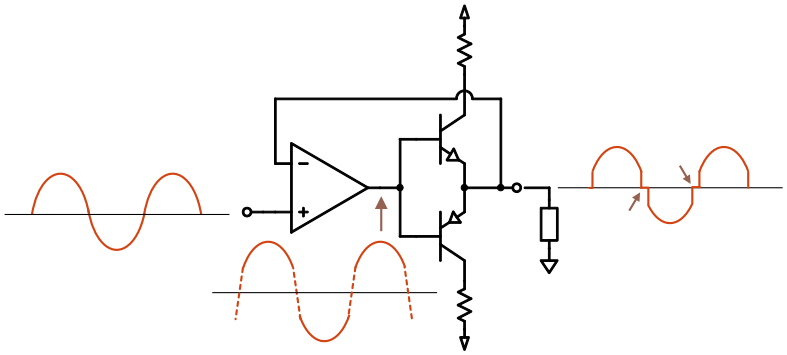


Figure 6-17: Generation of *crossover distortion* in the simple, *Class B* power-booster's output caused by the transistors' 0.7V base-emitter turn-on voltage coupled with the op-amp's finite slew rate.

If the load resistance is fairly high, so that high output currents will require that most of the transistors' power supply voltages be applied to the load, then the average power dissipated in the amplifier's transistors may be small, and the collector resistors may be omitted. If the load resistance is small, on the other hand (for example, an 8Ω speaker), and the power supply voltages are relatively high (such as the ±12V supplies on the breadboard), then the power dissipated by the transistors may be excessive.

The 250mA maximum available current from the ±12V supplies requires only 2V across an 8Ω speaker load, so the other 10V has to be dropped through the transistors and their associated collector resistors. The RMS power into the load in this case is only 0.25W, but *the average power drawn from each power supply line* is nearly 1W (for a sinusoid input only; a square wave input would draw 1.5W average). Where does all this extra power go? into heating of the transistors and their collector resistors, of course! On average each resistor + transistor combination will drop about 11V while the load draws about 100mA; to divide this equally between the two, each collector resistor should be about 60Ω: two 120Ω resistors in parallel will do the trick, and they together can handle 1/2W, leaving 1/2W for each transistor.

Note that the circuit has its input AC coupled (the high-pass *RC* filter). This is so that a steady, DC input will not cause the op-amp output to saturate, turning one transistor on and causing a steady, large current flow through one half of the circuit with its consequent large power dissipation.

If a truly versatile, low distortion, high frequency, large current output is needed for your design, then the simple transistor power booster will be inadequate, and a more complicated solution is called for. An interesting design choice is to use a special purpose buffer amplifier IC such as the Linear Technologies LT1010 which can be used with an op-amp to boost its current output to 150mA. Otherwise, you may choose a high-power op-amp — the Texas Instruments OPA541, for example, can output 10A peak using ±40V power supplies! A less dramatic example is the Analog Devices AD817: a 50MHz, 350 V/µs, 50mA op-amp.

Simple logic operations using transistor switches

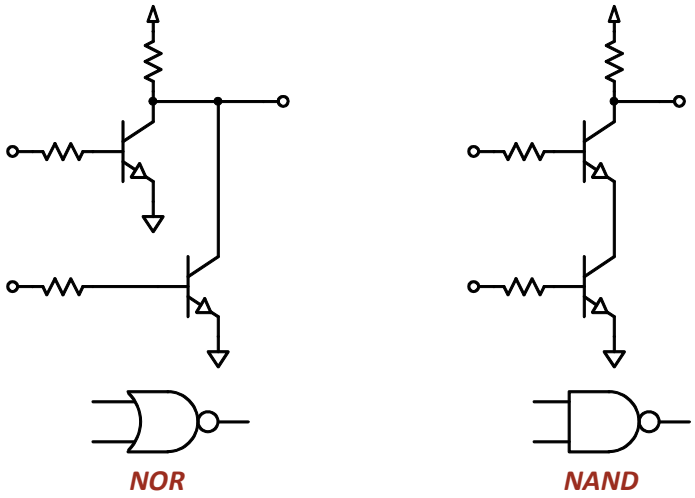


Figure 6-18: Simple logic operations implemented using basic *NPN* switch circuits.

Multiple transistor switch circuits using a common collector pull-up resistor may be used to perform simple logic operations as shown in Figure 6-18. In the case of the *NOR* configuration, if either of the transistor bases receives a *high* input, its transistor saturates and brings the output *low* (near ground potential). For the *NAND* circuit, both transistors must have their bases *high* for the output to be *low*, since the transistors are in series. More inputs could be added to either circuit. The collector resistor could be replaced by a load such as an indicator light, relay, or motor; in this case the logic operations would be *OR* and *AND*.

If more complicated logic operations are required, then the design should probably use digital logic ICs rather than logic constructed from discrete transistors and resistors. A final transistor switch could then be used if the load requires a high voltage or large current.

Experiment 7

Mini-Project Exercise

Copyright © Frank Rice 2013, 2014
Pasadena, CA, USA
All rights reserved.

Experiment 7

Mini-Project Exercise

Now it's time to apply what you've learned and attempt to design a circuit to perform some assigned function. What follows is a "menu" of design problems; your Prelab exercise will be to design a circuit to accomplish the task assignment you've chosen from the menu. During your lab period you will assemble and test your circuit, making any slight adjustments required for the circuit to be a success. You will demonstrate your circuit's performance at the end of the lab session, and you will provide a more detailed description of its performance in your lab write-up.

THE DESIGN PROBLEM MENU

Here they are! Each subsection here will provide a performance specification for an application whose circuit must be designed (with all component values assigned) which will satisfy the performance requirements. Examine them all, and then pick one to do for your lab.

1. *Swept-frequency function generator*

Implement a function generator (square and triangle outputs) that has an output frequency which increases linearly through a factor of ten, then rapidly resets to the original lower frequency, and then starts the linear increase again, repeating this process indefinitely. The circuit must also generate a "sync" signal which is high while the output frequency is sweeping upward and low which the frequency is being reset to its starting value (see Figure 7-1).

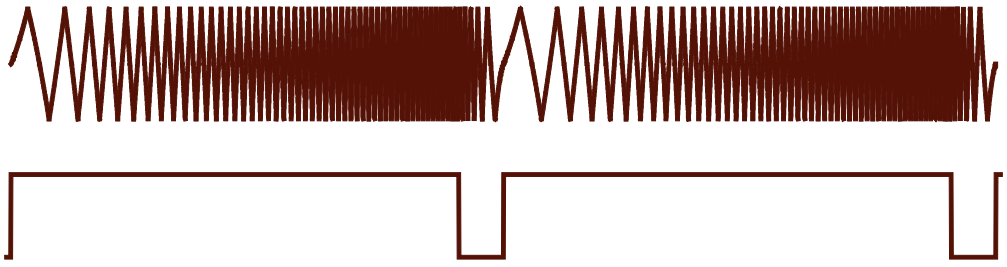


Figure 7-1: Sample output of the *swept-frequency function generator*. Top trace: generator output sweeping up through a 10-fold frequency range and then rapidly resetting to its original frequency and repeating (triangle output shown, but a sweeping square output should be available as well). Bottom trace: the accompanying "sync" signal which could be used to synchronize an oscilloscope to the output sweep.

Circuit specification details:

1. Both triangle and square outputs should be available, and their peak-to-peak amplitudes should be somewhere in the 2V to 10V range (these need not be adjustable, and the two output amplitudes need not match).
2. The sweep must cover a factor of ten in output frequency, frequency must increase linearly, and the sweep should take 1 to 2 seconds to complete. The starting (lowest) frequency of the sweep should be $\sim 100\,\text{Hz}$.
3. The time required for the circuit to reset its output frequency to its lowest value and begin another sweep should be $\ll 1$ second.
4. The "sync" output must be 4.5–5.0V when *high* (during each sweep) and 0.0–0.4V when *low* (while frequency is resetting for another sweep).

2. Amplifier with automatic gain control (AGC)

Implement an amplifier with a gain which is continuously, automatically adjusted so that its mean-square output voltage matches a control input DC voltage:

$$\overline{V_{out}{}^2}\,/10\text{V} = V_{control}$$

The amplifier's gain should be adjusted slowly, taking a few seconds to settle to a new gain following a change in the input signal amplitude or a change in the control voltage setting. Similarly, the time over which the amplifier's output is averaged should be on the order of a few seconds. Otherwise, the frequency response of the amplifier should cover the *audio* range: 20Hz to 20kHz.

The resulting behavior of the circuit is that it should amplify a voice or music input signal without distortion, but the gain will be such that the output will always have the same average power (over intervals of a few seconds) regardless of the input signal power (this behavior is called *automatic gain control* or *AGC*).

Circuit specification details:

1. A nominal input voltage of 8V peak-to-peak should be near the middle of the input voltage range for effective operation of the *AGC* amplifier.
2. The *AGC* should be effective over a range of about a factor of 5 variation in the mean squared input voltage (about a factor of 2 or 3 in the input peak-to-peak amplitude).
3. The range of $V_{control}$ for effective circuit operation should be around 3V to 5V.
4. The time constant both for averaging the squared output voltage and for adjusting the circuit gain should be 3sec to 5sec.
5. The *AGC*'s input impedance should be at least 100kΩ.

Note and hint: the most effective approach to solve this problem is to generate an error signal proportional to:

$$V_{control} - \left(\overline{V_{out}{}^2}\,/10\text{V}\right)$$

and then use this error signal to control the gain of the amplifier which generates V_{out}. See the section ***Circuit design hint: error signal generation and control loops*** for some comments about error signals and their uses in control loops such as the one needed for this circuit.

3. Audio tone controls

Consider the circuit shown below, an inverting amplifier whose gain is adjustable by changing the potentiometer wiper position. Moving the wiper through its range would result in an inverting gain range of $1/11 < -G < 11$; with the wiper in its center position the gain would be −1 (you should make sure you agree with these numbers).

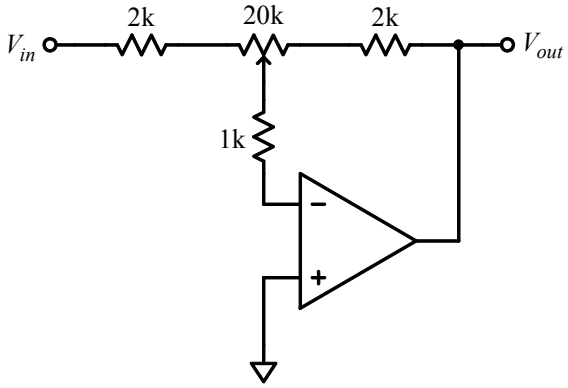


Figure 7-2: Inverting amplifier with an adjustable gain. With the potentiometer wiper all the way to the left, the gain would be −22k/2k = −11; wiper all the way to the right would result in a gain of −2k/22k = −1/11. This circuit will form the basis of a tone control; adding capacitors in strategic spots could make the gain adjustment only affect a subset of frequencies in the input signal.

By cleverly adding a couple of capacitors to this basic circuit (and maybe another resistor or two, you could use the resulting *RC* filtering to restrict the action of the potentiometer gain adjustment to only a certain range of frequencies, while for other frequencies the gain would stay at −1, independent of the wiper position. If the circuit is to be used at *audio* frequencies (20Hz to 20kHz), then the result is a tone control (e.g. *Bass* or *Treble*). An example of the frequency response of a modified circuit for a bass control is shown in Figure 7-3.

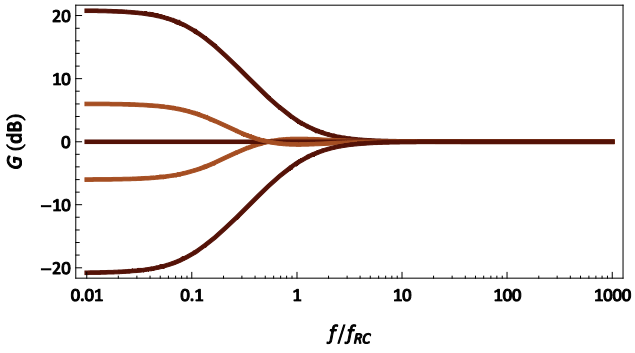


Figure 7-3: An example of a *Bass* (low frequency) *tone control* based on the circuit of Figure 7-2 with two capacitors added to restrict the action of the potentiometer to only affect a range of frequencies. For a typical bass tone control, f_{RC} would be 220Hz or so. A Treble control would be designed to affect frequencies above an f_{RC} of about 1800Hz. The curves shown are for *full boost* and *full cut*, 30% *boost* and *cut*, and potentiometer centered (no effect on the frequency response).

Your design task is to develop modifications to the circuit in Figure 7-2 to create a bass tone control or a treble tone control. A truly clever circuit could incorporate potentiometers for both functions into a single op-amp inverting amplifier circuit. The gain of the circuit when the potentiometer wiper is centered should be −1 at all frequencies in the audio range specified above. If you attempt a design which incorporates both bass and treble control in a single amplifier, then you may find it wise to use different resistance values for the two potentiometers. In this case the available potentiometer values are 1k, 20k, and 100k.

4. Window comparator

Design a comparator-type circuit which outputs a *high* level while its analog input signal lies between two adjustable threshold voltages (an *upper* and a *lower* threshold, which define the comparator's *window*), and outputs a *low* level when the analog input signal voltage lies outside the voltage window defined by the two thresholds.

The comparator should also respond to a *gate* input such that it will only output a high level signal in response to the analog input voltage level *while the gate input is high*; otherwise the output will remain low even if the analog input signal is within the comparator's voltage window. This gate input should have two additional features: it should incorporate hysteresis (implement using a Schmitt trigger), and it should act as though the gate were always high (the gate is *open*) if no external signal is attached to the circuit's gate input.

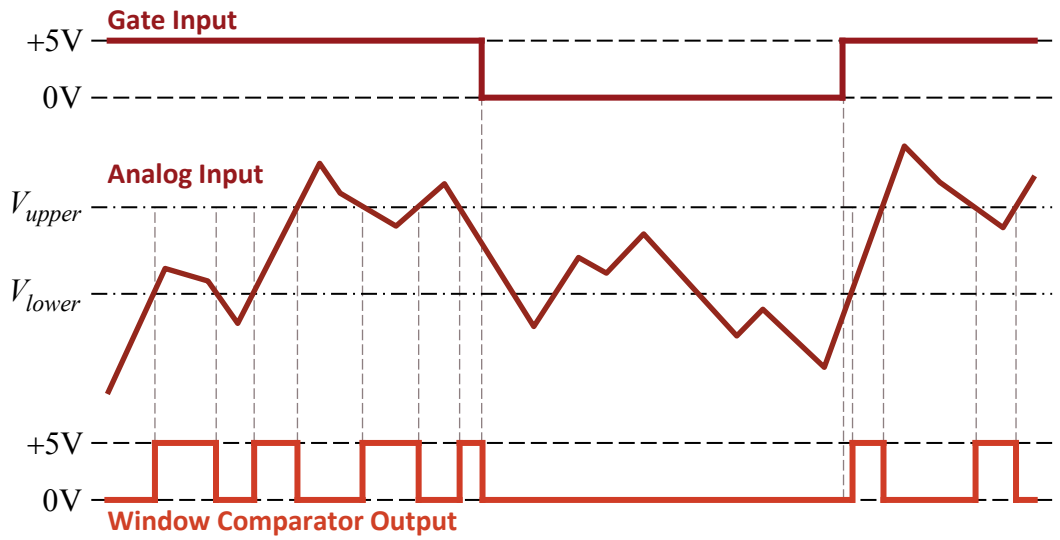


Figure 7-4: Sample *window comparator* behavior. The comparator's output is active only while its *Gate* input is high (or disconnected). When active, the comparator's output is high whenever its analog input signal voltage falls within the window defined by two adjustable thresholds, V_{lower} and V_{upper}.

Circuit specification details:

1. The valid analog signal input range should be at least −10V to +10V. The upper and lower thresholds should be adjustable to any two voltages within this range.
2. The comparator output must be 4.5–5.0V when *high* and 0.0–0.4V when *low*.
3. The gate input should incorporate hysteresis on its input such that the gate signal must go above +2V when going from *low* to *high* (to enable the comparator output), and it must go below +1V when going from *high* to *low* (to disable the comparator output).
4. The valid gate input voltage range should be at least −10V to +10V, but the hysteresis thresholds described above will determine what voltage levels define a high and low gate input (the 0V and +5V levels for the gate input shown in Figure 7-4 would meet the hysteresis threshold requirements, but are not mandatory).
5. Both the analog and gate inputs should have an input impedance of at least 100kΩ.

5. Pulse-width modulator and demodulator

Design a circuit which will accept an input signal and use its instantaneous voltage value to control the *duty cycle* (symmetry) of a high-frequency square-wave signal generated by your circuit. The duty cycle should be 50% when the input signal is 0V; the duty cycle should increase toward 100% (output always high) as the input signal becomes more positive and decrease toward 0% (output always low) as the input signal becomes more negative. This variation in the duty cycle of a high-frequency square-wave in response to an input signal is called *pulse-width modulation* (Figure 7-5).

You also must design a circuit which will accept a pulse-width modulated signal (such as the upper trace in the left-hand image in Figure 7-5) and *demodulate* it to reproduce the original

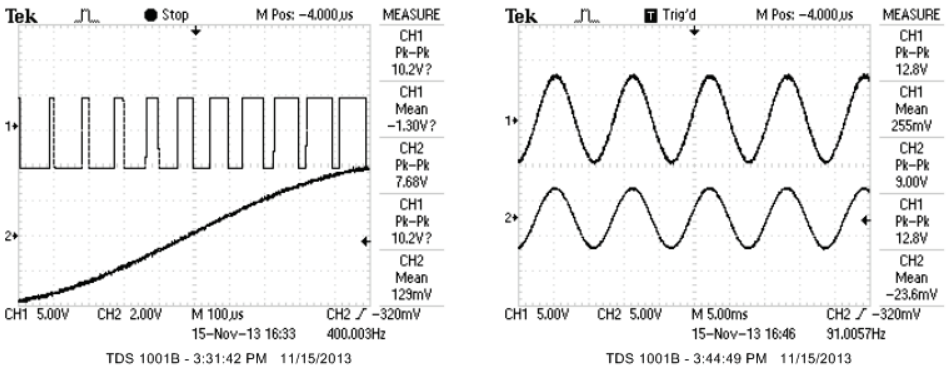


Figure 7-5: *Pulse-width modulation*. Left: the input signal level (CH2, lower trace) controls the *duty cycle* of a high-frequency, square-wave *carrier* output signal (CH1, upper trace); the *frequency* of the carrier is not affected by the input signal. Right: the *demodulated* carrier signal (CH1, upper trace) *recovers* (reproduces) the original input signal (CH2, lower trace).

input signal source (demodulated output shown in the upper trace of the right-hand image in Figure 7-5).

Circuit specification details:

1. The valid analog signal input range which will not *saturate* the modulated square wave (duty cycle reaches 100% or 0%) should be at least −8V to +8V.
2. The square-wave carrier frequency should not change as long as the input signal does not saturate it (see above requirement). Only the output duty cycle should change, not the output frequency.
3. The square wave frequency should be at a fixed value somewhere in the range of 20kHz to 80kHz.
4. The demodulated output should faithfully reproduce the input signal (not show significant distortion) for any input signal with a frequency of less than 1/20 of the carrier (square-wave) frequency and an amplitude of less than 90% of the saturation amplitude. A residual bit of the carrier frequency may show up in the demodulated output as a slight ripple in the output (less than 5%).

Note and hint: no multipliers, diodes, or transistors should be needed for this circuit; only op-amps, resistors, and capacitors. The demodulator is particularly simple.

Circuit design hint: error signal generation and control loops

A common circuit design problem, especially in such applications as industrial process control, automatic stability and guidance systems, or signal frequency or amplitude control, is to design a circuit so that some characteristic of a dynamic system output is automatically, quickly, and precisely adjusted by the circuit to match some value determined by an independent control parameter input to it. The field of study of the most effective way to implement systems to meet this requirement is part of the subject known as *control theory* (Caltech even offers an undergraduate minor in this subject: Control and Dynamical Systems). For example, the use of negative feedback to ensure that an op-amp's output is a linear function of its inputs is an implementation of a basic method studied by students of control theory.

Consider the abstract representation of a simple *control loop* shown in Figure 7-6. The control loop continually monitors the system's output by comparing it to a control *set-point* input. The difference in the set-point and the circuit output values becomes the *error signal* which is used to generate a *control command* to adjust the system's behavior. The object of the loop is to drive the error signal toward 0 in a *prompt and well-controlled manner* until the error is acceptably small.

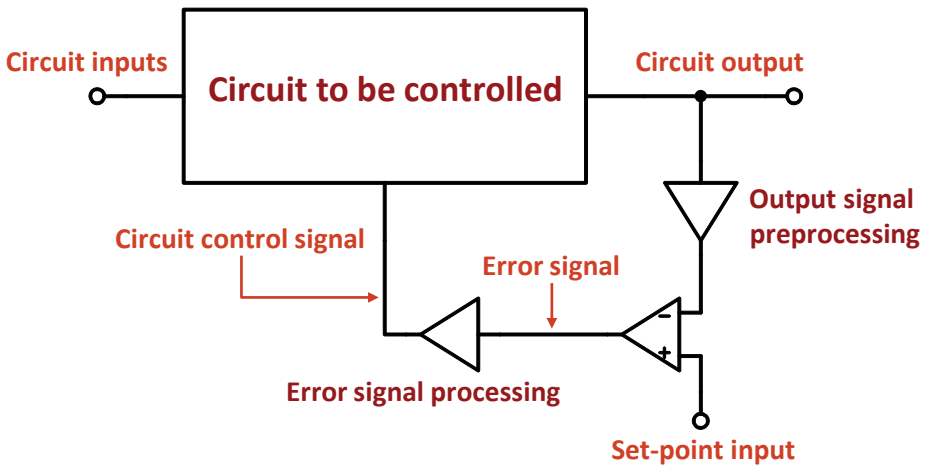


Figure 7-6: An abstract representation of a simple *control loop*. The main circuit or process to be controlled produces some output which is a function of its inputs; this function is adjusted based on the value of the loop's *circuit control signal*. The control loop monitors the circuit's output by comparing it to a *set-point*; the difference in the set-point value and the circuit output becomes an *error signal*. The error signal is processed by amplifiers and filters to adjust the circuit control signal until the error signal becomes acceptably small.

For the control loop to be effective, the error signal must be properly generated and carefully handled; even once it has succeeded in driving the error to a sufficiently small value, changes in the system's inputs, perturbations to and noise in the system itself, and changes in the set-

point input value will require the control loop to continue to adjust the circuit control signal. The control loop may fail by: (1) generating control signals of insufficient range or quickness to reduce the error signal enough to control the circuit's output error; (2) commanding the system's adjustments in the wrong direction, so that the error grows rather than shrinks; or (3) overcorrecting the system by making adjustments which are too large and too slow, so that the error overshoots and oscillates about 0 rather than converging on 0 with time.

Avoiding these errors in the control loop design requires careful consideration of how you handle the error signal and control signal calculations. As shown in Figure 7-6, some preprocessing of the circuit output monitor (such as scaling or filtering) may be required before it is subtracted from the set-point value to generate an error signal. This error signal in turn may require significant amplification, filtering, and other signal processing actions in order to convert it into the control signal required to adjust the system's performance.

A common error signal processing system is called a *PID* loop, for *Proportional-Integral-Differential* control: the error signal is amplified, integrated, and differentiated by parallel op-amp circuits. These circuits' outputs are then added by a summing amplifier; the weight of each term in the sum is then adjusted to provide the best control loop response to changes in the system or the set-point (Figure 7-7).

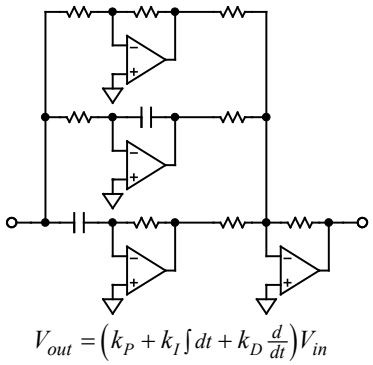


Figure 7-7: A PID operator circuit commonly used to process an error signal. The weights of the individual terms may be adjusted by selecting appropriate values for the input resistors in the final summing amplifier.

PRELAB EXERCISES

1. Carefully read the Experiment 8 notes, particularly the section titled **DEVELOPING THE SYSTEM DESIGN**. Come up with an idea for your final project and develop an initial performance specification for it. Draw a tentative system block diagram of the signal flow through your circuit and the operation each sub-circuit must perform. Use this to make a rough estimate of the numbers of op-amps, multipliers, timers, other ICs, and/or transistors you may need to complete the project.

2. Pick one of the design problems from the menu of choices presented and design a circuit that will perform that task. Provide a complete schematic with component values assigned, using only the types of op-amps, multipliers, and other circuit elements you've used in previous experiments, except that you may include potentiometers in your design to set the variable parameters the circuit may require (such as the audio tone control design task, or the adjustable thresholds in the window comparator design task).

Available potentiometer values are 1k, 20k, and 100k.

LAB PROCEDURE

Implement and test your circuit

Construct and test the circuit you've designed to satisfy the requirements of your chosen design problem. You may use the components preinstalled on the analog trainer circuit board in addition to components you need to install in the breadboard areas.

Practice good component layout techniques when you use the breadboard area: identify the rails you'll use for power supply voltages and ground; determine the most effective positioning of the ICs on the breadboard so that the wiring will be clean and easy to check; double and triple check the pin number assignments shown on the IC data sheets; make the power supply and ground connections to each IC before installing other components; add IC pin numbers to your schematic so that you can quickly add components and check you connections.

Lab results write-up

As always, include a sketch of the schematic with component values for each circuit you investigate, along with appropriate oscilloscope screen shots. Make sure you check the performance of your circuit against each of the design problem specifications.

Experiment 8

Final Project

Copyright © Frank Rice 2014
Pasadena, CA, USA
All rights reserved.

Experiment 8

Final Project

Your final project is the culmination of your efforts to learn some analog electronics and provides you the opportunity to demonstrate your new-found talents by creating something fun of your own design. You will have about ten days to build, test, and refine your circuit before presenting it to the judges — the course instructor and TAs.

Your project grade will be based on your circuit's originality, how well it meets your design specification, and how well you understand its operation and limitations.

PROJECT SCOPE AND SOME IDEAS

Your project's sophistication should be matched to how much you have come to understand the concepts of analog circuit design and testing: if you pick a project idea that is too simple for you, then you should expect a disappointing grade; if you pick a project that is too difficult, then you won't have enough time to complete it, and you won't have a good understanding of what is happening in the circuit to keep it from working the way you want. Your TA and the course instructor can work with you to decide on a project that you find interesting and that is suited to your current level of understanding.

You will construct the electronic circuit for your project on one or more solderless breadboards provided to you. The breadboard and its circuitry are yours to keep (except for some special components you may need to borrow from the lab). Identify any unique parts you will need for your project early so that the lab has time to acquire them for you — this is particularly important for any sensors, actuators, or special connectors you may want to use.

Sub-circuits of your project system may use designs you have found in books or on the internet, but they should generally involve mainly components and techniques to which you have been introduced as a part of your studies. In any case you must, however, thoroughly understand the design and operation of such circuits and, importantly, why the various resistors and capacitors in the circuit were assigned their particular values. You should always apply some customizations of your own design to the original circuit to complete your project.

You may design parts of your project using ICs which are more sophisticated than what you have studied during your course work, but they must not be *ASICs* (application-specific ICs) which require only a power supply and a speaker, for example, to complete the system. In any case, you must study the IC's data sheet and thoroughly understand what it does, its limitations, and how to use it.

Your circuit may not be connected to the building 120VAC power network unless you use a lab power supply or similar, approved, safe device. 9V batteries and their associated connectors will be available to use with your project, if you wish. Your circuit must not connect to a human or other animal body using a low-impedance contact.

You should not plan to design and build RF oscillators or amplifiers (frequencies over about 2MHz) because the solderless breadboards are not suitable for these high-frequency circuits. Antennas are generally difficult beasts to successfully tame, so make sure you thoroughly discuss your project and any antenna requirements with the course instructor.

Project ideas

Here is a selection of possible project ideas arranged roughly in order of increasing complexity or conceptual difficulty. In all cases they should be constructed only from op-amps, bipolar transistors, and the other component types to which you have been introduced in lab (e.g. no voltage regulator ICs for the voltage converter projects).

Fairly straightforward

Electronic stethoscope to amplify and listen to a heartbeat

A "clapper" for an LED ("clap on… clap off…")

Phase-locked oscillator outputting twice the frequency of an input signal

Touch or proximity sensor to turn an LED on and off

A gated function generator to produce a repeating tone burst of its output

A frequency to voltage converter (1V DC output for a 1kHz input frequency)

A capacitance meter (1V DC output for a 100nF capacitor)

A 4-quadrant analog divider: divide by negative voltages as well as positive

Moderately challenging

Color organ: 3 filter channels with each channel's LEDs' intensities varying with sound level

Generate a synthetic heartbeat waveform

A dual, tracking voltage regulator producing regulated ±5V from two 9V batteries

Intersection traffic lights (using red, yellow, green LEDs) with proper light sequencing

FM signal demodulator using a phase-locked loop

Send data or audio over an optical data link using an LED and a photodiode

Metal detector

A simple Theremin circuit (frequency control only), or include a volume control as well (harder)

Temperature controller of the glass envelope of a small incandescent light bulb using the bulb filament as the heater and a thermistor as the temperature sensor

Class-D power amplifier outputting 1W RMS into an 8 ohm speaker from a 5V power supply

An envelope generator to take a tone burst and vary its attack and decay profiles

Pulse-stretcher circuit using discrete transistors: 50ns, 5V input to a 10µs, 5V output

Plot any general 4th-order polynomial on an oscilloscope (using XY mode)

Generate 100V DC from a 9V battery

Diode temperature sensor with linear output of 2.5V for 25°C, etc.

Efficient Class AB transistor power amplifier for an op-amp output with no crossover distortion and over-current protection

More sophisticated

Generate ±12V DC from a single 9V battery

A simple op-amp constructed from discrete transistors (must be DC-coupled)

Add a sine-wave output to the function generator circuit of Experiment 4

Add voltage-controlled symmetry to the function generator circuit of Experiment 4

Chua's chaotic oscillator circuit showing its behavior on an oscilloscope

Analog computer: plot projectile trajectory on an oscilloscope (using XY mode) given initial velocity

Analog computer: Lorenz attractor showing chaotic behavior on an oscilloscope

A PID control loop to maintain the balance of an inverted pendulum

DEVELOPING THE SYSTEM DESIGN

In this section we work through an example project and its development phases, from initial specification to final, detailed schematics. First, start with a short description of the project. For this example:

> ***Design a preamplifier for a phonograph cartridge magnetic pickup that includes the proper RIAA frequency response and which has bass and treble tone controls as well.***

From this start, we expand and develop the project idea step by step. First, a more detailed specification of the circuit is required…

Detailed specification

Starting from the rough project description statement, outline the actual requirements and scope of the project, as was done for the mini-project tasks in Experiment 7. By specifying the actual, detailed purpose of the project and defining its scope, you have something which can serve to focus your effort.

> This specification will be the reference against which the level of success of your project will be measured.

For our example phono preamplifier project, we choose to meet the following specifications:

Phono sensor nominal output voltage:	3mV RMS at 1kHz
Preamplifier nominal output voltage:	0.1V RMS at 1kHz
Preamplifier peak output voltage:	10 V peak-peak into a 10kΩ load
Preamplifier output frequency response:	−3dB 10Hz–30kHz
Tone control boost and cut:	±10dB boost and cut (minimum)
Tone control 3dB corner frequencies:	300Hz (bass), 1.5kHz (treble)
Input impedance:	47kΩ

RIAA equalization (frequency response):
Poles at 50Hz and 2122Hz; zero at 500Hz

Power supply: ±12V

System block diagram

With the detailed specification in hand, next sketch out a system block diagram showing the major sub-circuits, how they will be interconnected, and how the signal is modified as it flows through each sub-circuit. You may later decide that you need to change this block diagram as you learn more about the design problems your project entails, but at least you'll have an initial plan from which to deviate! You should go over this block diagram and your

specifications thoroughly with your TA; once this review is finished both you and your TA should be reasonably satisfied that your final project is appropriate for your skill level and that you can successfully complete it on time.

For the example phono preamplifier project, the system block diagram is quite simple (Figure 8-1); for some projects it might be quite complicated.

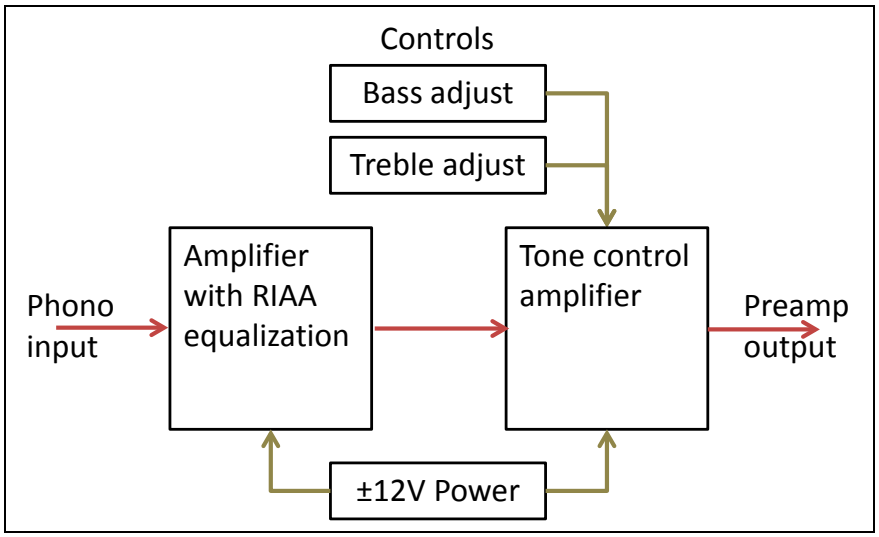


Figure 8-1: A system block diagram for the phono preamp example project.

Sub-circuit specifications

Using the system block diagram, you then define the detailed specifications of each sub-circuit by describing what its inputs and outputs are and how the signals there are related to each other: for a linear stage, what are its input and output impedance limits and what is its transfer function; for a nonlinear stage, what are its various operating states, what are the inputs required to cause various state transitions, and what are the outputs associated with each state. The required frequency response and/or state transition speed for each sub-circuit should also be determined, as well as its minimum and maximum signal voltage and current requirements.

Identify those blocks or sub-circuits which are likely to give you the most trouble, either because you are unsure how to design a circuit to perform that subtask, or because it involves the use of components with which you are unfamiliar. Make sure that you plan your development and testing schedule to include extra time to figure out and thoroughly test those parts of the project. If necessary, develop a fallback option in case you can't get some section of the project to perform as well as you would like.

Identify the ICs and other more complicated components that may be needed for each sub-circuit. Find and download the manufacturers' *data sheets* for these components and thoroughly read them to make sure that they will meet the requirements of your design. If

necessary, go over the data sheets with the course instructor to make sure you understand the many very technical specifications included in them.

When choosing an unfamiliar IC for your project make sure that it is really suitable! In particular, check that:

1. The IC's power supply voltages are compatible with your design.
2. It comes in a package that you can actually make connections to.
3. It is neither too fast (and therefore difficult to keep stable) nor too slow for your application.

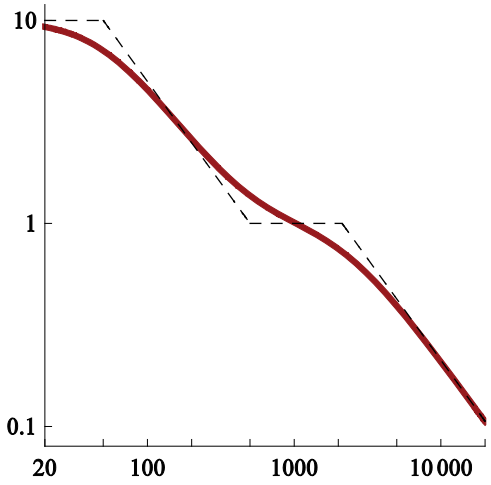
Here are sets of specifications for each of the two amplifier blocks for the example project:

b) Example: amplifier sub-circuit with RIAA equalization:

i) Gain at 1kHz: $G_0 = 33$
ii) Input impedance: 47kΩ
iii) RIAA equalization gain function and Bode plot:

Figure 8-2: Gain function and Bode magnitude plot of the RIAA equalization gain plot normalized to 1 at mid frequencies (near 1kHz). The asymptotic response is shown as a dashed line. The first-order corner frequencies are at 50Hz, 500Hz, and 2120Hz.

$$G(f) = G_0 \frac{(1 + j\,f/500\text{Hz}) \times (500\text{Hz}/50\text{Hz})}{(1 + j\,f/50\text{Hz}) \times (1 + j\,f/2120\text{Hz})}$$



iv) Frequency response modifications to the RIAA curve: first-order roll-off at low frequencies with −3dB corner at 10Hz.

c) Example: tone control amplifier sub-circuit:

i) Gain with tone controls centered (no boost or cut): 1
ii) Frequency response with tone controls centered (−3dB corners): 10Hz, 50kHz

iii) Nominal target max boost and cut response curves (note that we target ±20dB in order to have some leeway to achieve the ±10dB system specification):

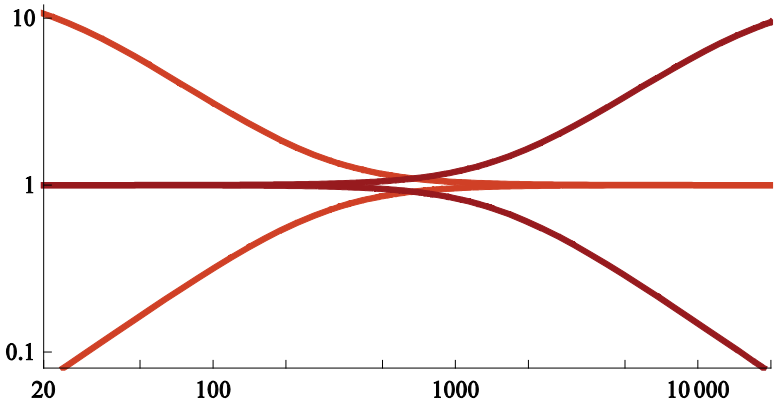


Figure 8-3: Target tone control amplifier gain response at max boost and cut for both bass and treble controls. ±3dB corners are 300Hz (bass) and 1.5kHz (treble).

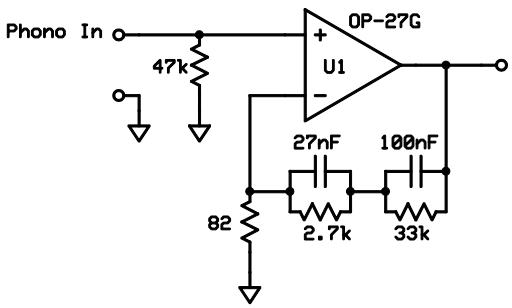
Detailed sub-circuit schematics

Now you can generate an initial schematic diagram for each sub-circuit and begin to determine its component values.

In some cases you may have to *prototype* a sub-circuit: assemble a preliminary version and check its performance against its specifications in order to finalize its design and determine all of its component values. This effort overlaps that of the next section: *construction and testing*, but it is vital for the success of your project that you use prototyping early on any complicated or subtle section of the design so that you can learn how to make that part work properly!

Continuing our example project, an initial schematic for the RIAA preamplifier sub-circuit is shown in Figure 8-4. The op-amp and one of the capacitor values are both unusual, so a list containing these components must be submitted for parts ordering.

Figure 8-4: Initial schematic for the input amplifier with RIAA equalization sub-circuit. The OP27GP op-amp is a low-noise, high-gain, precision unit particularly suited for this application. It and the 27nF capacitor may not be available in the lab, so these parts will need to be ordered.



Sensors and actuators

Many projects require some sort of sensor (such as a microphone) which should be mounted separately from the circuit breadboard. You will need to fashion some sort of cabling and connectors to attach such sensors to the circuit. Get advice and assistance from your TA and the course instructor with this problem. You will probably need to solder some of the connections; the course instructor can show you how to properly solder wires and connectors. Similar issues arise with output actuators such as loudspeakers.

> Make sure that you obtain the data sheet or other instructions for properly connecting and using sensors, actuators, or input-output devices; many are easily and permanently damaged if incorrectly connected into a circuit.

Construction and testing

Once you are fairly confident that your sub-circuit designs are successful, start building up the final, complete circuit by connecting sub-circuits together. If you have many sub-circuits, don't assemble the complete project circuit at first, but start by connecting up independent subsets of the project and determining whether they successfully perform their parts of the overall project task.

You may find that connecting together two sub-circuits substantially degrades the performance of one or the other, possibly because their input and output impedances or signal levels aren't compatible. Adding a voltage follower, voltage divider, or amplifier between them may correct the problem. Maybe the signs of the signals are incompatible; maybe one signal has a DC offset which is too large for a sub-circuit to handle. Recheck your sub-circuit design specifications and try to understand how this interfacing problem arose.

In some cases you may decide that you have misidentified the correct division of your project into sub-circuits, and some part of the system requires a redesign. Make sure you have allotted enough time for this construction and testing phase!

PRELAB EXERCISES

1. Finalize your project's circuit specification.

2. Develop a detailed system block diagram of the signal flow through your circuit and the operation on the signal each sub-circuit must perform. The sub-circuit breakdown should be at the level of individual op-amps or IC functions.

3. Develop a fairly complete schematic diagram of each of your sub-circuits. Label each sub-circuit's inputs and outputs and describe how they will be connected to the other sub-circuits. Present an initial, complete parts list for your circuit, especially of the required ICs, input-output devices, and other components not available in the lab.

LAB PROCEDURE

Construct and test the circuit you've designed to satisfy the requirements of your chosen project. As was discussed previously, you should build, test, analyze, and fix each sub-circuit independently before connecting various sub-circuits into sub-systems. These sub-systems should in turn be fully tested and working before being connected to complete the final project system. Following this bottom-up assembly and test procedure will improve your chances of success.

Double and triple check the pin number assignments shown on the IC data sheets, especially for the power supply and ground connections. Make sure that any unfamiliar ICs you use are rated for the power supply voltages before you install them!

Practice good component layout techniques when you assemble circuits on your breadboard: add IC pin numbers to your schematic so that you can quickly add components and check you connections; identify the rails you'll use for power supply voltages and ground; determine the most effective positioning of the ICs on the breadboard so that the wiring and component placement will be clean and easy to check; make the power supply and ground connections to each IC before installing other components.

Once the project is working, make sure you collect the data you need to fully document its performance: frequency response sweeps and oscilloscope screen captures, for example. This data will back up your demonstration of the circuit when you present your project.

Double check and update your circuit schematics to reflect the final component values and circuit topology you use during the final presentation.

PROJECT PRESENTATION

Your final project will be demonstrated in the lab to a committee consisting of the course TAs and the lab administrator. At the presentation you must submit the following project documents in support of it:

1. The final detailed specification of the project
2. A comprehensive system block diagram identifying all sub-circuits and their various purposes
3. A complete circuit schematic identifying all component values
4. Data (frequency response plots, o-scope screen shots, etc.) demonstrating that the circuit meets the project specification

You will set up your project circuit at one of the lab stations; you may use the analog trainer power supply, the signal generator, and the oscilloscope, if necessary, to provide a live demonstration of your system.