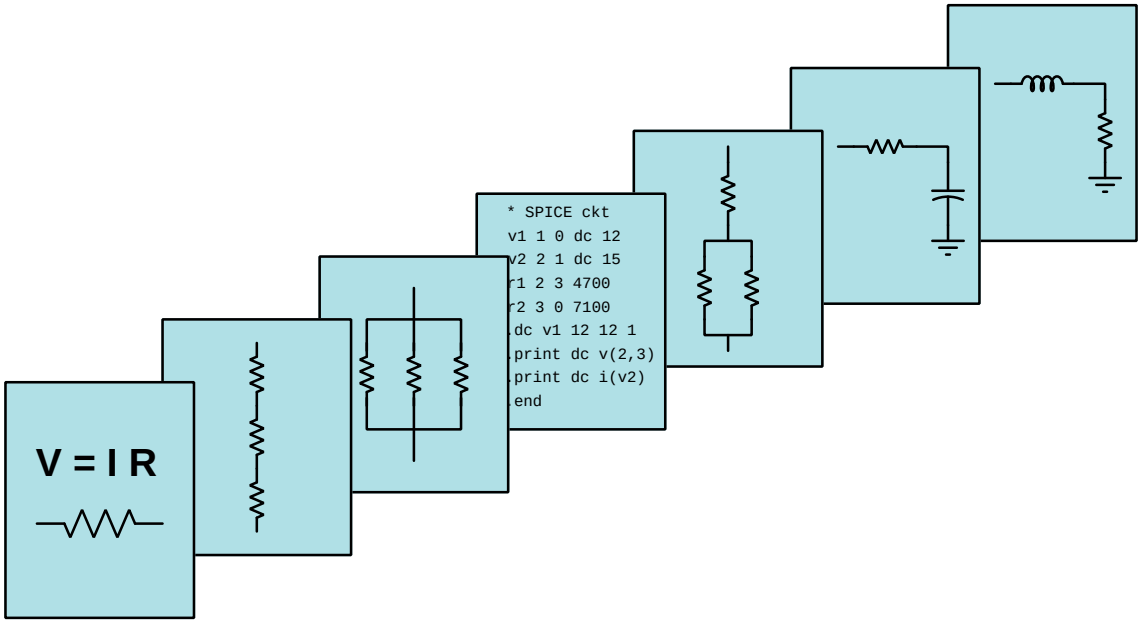
Modular Electronics Learning (ModEL) project



Operational Amplifiers

© 2019-2026 by Tony R. Kuphaldt – under the terms and conditions of the Creative Commons Attribution 4.0 International Public License

Last update = 23 September 2026

This is a copyrighted work, but licensed under the Creative Commons Attribution 4.0 International Public License. A copy of this license is found in the last Appendix of this document. Alternatively, you may visit `http://creativecommons.org/licenses/by/4.0/` or send a letter to Creative Commons: 171 Second Street, Suite 300, San Francisco, California, 94105, USA. The terms and conditions of this license allow for free copying, distribution, and/or modification of all licensed works by the general public.

Contents

1 Introduction **3**
- 1.1 Recommendations for students . . . 3
- 1.2 Challenging concepts related to operational amplifiers . . . 6
- 1.3 Recommendations for instructors . . . 8

2 Case Tutorial **11**
- 2.1 Example: Kirchhoff's Voltage Law in action . . . 12
- 2.2 Example: open-loop opamp response to various inputs . . . 13
- 2.3 Example: multi-stage transistor amplifier with negative feedback . . . 14
- 2.4 Example: opamp networks responding to various inputs . . . 16
- 2.5 Example: simulating opamp circuits using SPICE . . . 17
- 2.6 Example: simple opamp experiments . . . 20
- 2.7 Example: DC power supply rail splitter . . . 24
- 2.8 Example: simple triangle/square/PWM oscillator . . . 26
- 2.9 Example: measuring signal rates of change . . . 31

3 Tutorial **35**
- 3.1 Electronic amplifiers . . . 35
- 3.2 Differential inputs and outputs . . . 36
- 3.3 Integrated circuit amplifiers . . . 38
- 3.4 Amplifier split power supplies . . . 40
- 3.5 Signal comparison . . . 41
- 3.6 Negative feedback . . . 42
- 3.7 Regulation . . . 46
- 3.8 Servos and control systems . . . 49
- 3.9 Attenuated and offset feedback . . . 51
- 3.10 Inverting and noninverting amplification . . . 54
- 3.11 Amplification with offset . . . 60
- 3.12 Differential amplification . . . 65
- 3.13 Instrumentation amplifiers . . . 73
- 3.14 Summation . . . 80
- 3.15 Precision current circuits . . . 84
- 3.16 Boosted output . . . 88
- 3.17 Precision rectification . . . 92

3.18 Integrators and differentiators . . . 100
3.19 Analog computing . . . 109
3.20 Oscillator circuits . . . 115
3.20.1 Relaxation oscillators . . . 115
3.20.2 Integrating oscillators . . . 119
3.20.3 Phase shift oscillators . . . 121
3.20.4 Resonant oscillators . . . 125

4 Historical References **129**
4.1 Platform balance scale operation . . . 130
4.2 Harold Black's feedback patent . . . 133
4.3 George Philbrick's work on gun tracking systems . . . 142
4.4 Philbrick Researches model K2-W operational amplifier . . . 146
4.5 William Hewlett's oscillator . . . 149

5 Derivations and Technical References **153**
5.1 IEC standard component values . . . 154
5.2 Common-mode voltages and currents . . . 156
5.3 General-purpose opamp recommendations . . . 163
5.3.1 TLV27x series . . . 163
5.3.2 TLE214x series . . . 164
5.3.3 TL07x / TL08x series . . . 164
5.3.4 MCP600x series . . . 165
5.3.5 LM158/258/358 . . . 165
5.3.6 LM324 . . . 165
5.3.7 LM741 and LM1458 . . . 166
5.4 Decibels . . . 167
5.5 Split DC power supplies . . . 177
5.6 Derivation of Millman's Theorem . . . 179
5.7 Operational amplifier imperfections . . . 180
5.7.1 Output voltage limits . . . 181
5.7.2 Input voltage limits and latch-up . . . 182
5.7.3 Input bias current . . . 184
5.7.4 Input offset voltage . . . 188
5.7.5 Frequency compensation . . . 189
5.7.6 Gain-bandwidth product . . . 189
5.7.7 Slew rate . . . 190
5.7.8 Common-mode rejection . . . 191
5.8 Protecting opamp inputs from over-voltage . . . 192
5.9 Latch-up in integrated circuits . . . 197

6 Questions **207**
6.1 Conceptual reasoning . . . 211
6.1.1 Reading outline and reflections . . . 212
6.1.2 Foundational concepts . . . 213
6.1.3 Determining output polarities . . . 215

6.1.4 Fooling a voltage regulator . . . 219
6.1.5 Forming a split power supply . . . 221
6.1.6 Load current path . . . 222
6.1.7 Model 324 opamp qualitative analysis . . . 223
6.1.8 TL08x opamp qualitative analysis . . . 224
6.1.9 Philbrick K2-W opamp . . . 225
6.1.10 Voltage follower analysis . . . 227
6.1.11 Voltage follower versus a wire . . . 228
6.1.12 Unusual current-regulating circuit . . . 229
6.1.13 Push-pull voltage follower . . . 230
6.1.14 Electronic levers . . . 232
6.1.15 Comparison of input impedances . . . 233
6.1.16 Cable guarding . . . 234
6.1.17 Bias current compensation . . . 236
6.1.18 Differential signaling . . . 238
6.1.19 Vocal eliminator circuit . . . 240
6.1.20 Phototube amplifier . . . 241
6.1.21 Opamp power rectifier? . . . 243
6.1.22 Peak follower-and-hold circuit . . . 244
6.1.23 Integrator with DPDT switch . . . 245
6.1.24 Position, velocity, and acceleration signals . . . 246
6.1.25 Athletic accelerometer . . . 247
6.1.26 Rogowski coil power meter . . . 248
6.1.27 Adjustable oscillator . . . 250
6.1.28 Limiting network . . . 251
6.2 Quantitative reasoning . . . 253
6.2.1 Miscellaneous physical constants . . . 254
6.2.2 Introduction to spreadsheets . . . 255
6.2.3 Empirically determining open-loop gain . . . 258
6.2.4 Voltage follower simulation program . . . 259
6.2.5 Calculating output voltages and gains . . . 261
6.2.6 Voltage gain calculations . . . 262
6.2.7 Necessary resistor values for specified voltage gains . . . 263
6.2.8 Modifying voltage gain . . . 265
6.2.9 Multi-stage amplifier gains . . . 266
6.2.10 Calculating input and output voltages . . . 268
6.2.11 High-resistance analog voltmeter . . . 269
6.2.12 Opamp signal scaling circuits . . . 270
6.2.13 Current regulator limits . . . 272
6.2.14 Current regulator simulation program . . . 273
6.2.15 Constant-current electronic load circuit . . . 276
6.2.16 Constant-voltage electronic load circuit . . . 277
6.2.17 Subtractor calculations . . . 278
6.2.18 Summer circuit calculations . . . 280
6.2.19 Summer circuit simulation program . . . 282
6.2.20 Adding diodes to an inverting amplifier . . . 284

6.2.21 Precision rectifier analysis . . . 286
6.2.22 Polarity separator analysis . . . 287
6.2.23 Precision rectifier simulation program . . . 288
6.2.24 Signal selector circuit . . . 290
6.2.25 Integrator calculations . . . 291
6.2.26 Integrator and differentiator simulation program . . . 292
6.2.27 Three-phase oscillator . . . 293
6.2.28 Three-phase oscillator simulation program . . . 294
6.2.29 Capacitor voltage simulation program . . . 295
6.3 Diagnostic reasoning . . . 298
6.3.1 Internal faults in a model 324 . . . 299
6.3.2 Internal faults in a model TL082 . . . 300
6.3.3 Pushbutton-adjustable voltage . . . 302
6.3.4 PCB measurements . . . 304
6.3.5 Effects of faults on a simple amplifier circuit . . . 306
6.3.6 Faults in a power supply voltage regulator . . . 307
6.3.7 Voltage regulator with zero output . . . 308
6.3.8 Clipped output waveform . . . 309
6.3.9 Audio amplifier design improvement . . . 310
6.3.10 Effect of faults on amplifier input impedance . . . 311
6.3.11 Troubleshooting a microphone amplifier . . . 312
6.3.12 No longer virtual ground . . . 313
6.3.13 Testing common-mode rejection . . . 314
6.3.14 Design flaws in a peak follower-and-hold circuit . . . 315
6.3.15 Incorrect full-wave rectifier analysis . . . 316
6.3.16 Effects of faults in a relaxation oscillator . . . 317
6.3.17 Failed *LC* oscillator . . . 318

A Problem-Solving Strategies **319**

B Instructional philosophy **321**
B.1 First principles of learning . . . 322
B.2 Proven strategies for instructors . . . 323
B.3 Proven strategies for students . . . 325
B.4 Design of these learning modules . . . 326

C Tools used **329**

D Creative Commons License **333**

E References **341**

F Version history **343**

Index **348**

Chapter 1

Introduction

1.1 Recommendations for students

An *operational amplifier* is a high-gain differential amplifier intended to be used as an integral part of precision analog signal processing and computing circuits. Designed to be "building block" modules for analog circuits, operational amplifiers are extremely versatile devices. This module explores basic principles of operational amplifiers (often called *op-amps*) and some of their applications. Forged in the crucible of war and perfected in peacetime applications, the operational amplifier remains today an essential building-block of analog circuit design. Though the technology has changed quite a bit, from bulky vacuum tubes to microscopic semiconductor circuitry etched on chips of silicon, the underlying principles remain the same. This module explores the basic principles of these highly useful circuits.

Important concepts related to operational amplifiers include **single-ended** versus **differential** voltage signals, **amplification**, **gain**, **integrated circuits**, **bridge** networks, **feedback**, **block diagrams**, the **Cardinal Rule of Negative Feedback**, **voltage regulation**, **Kirchhoff's Voltage Law**, **Kirchhoff's Current Law**, **Ohm's Law**, **servo** systems, **voltage divider** networks, **sensors**, **functional inversion**, behavior of **sources** versus **loads**, **input impedance**, **Norton's Theorem**, **current sources**, **compliance voltage**, **insertion resistance**, **differentiation**, **integration**, **Fourier's Theorem**, **noise**, and **drift**.

Here are some good questions to ask of yourself while studying this subject:

- How might an experiment be designed and conducted to measure the offset voltage for an operational amplifier IC? What hypothesis (i.e. prediction) might you pose for that experiment, and what result(s) would either support or disprove that hypothesis?
- How might an experiment be designed and conducted to measure the slew rate for an operational amplifier IC? What hypothesis (i.e. prediction) might you pose for that experiment, and what result(s) would either support or disprove that hypothesis?
- How might an experiment be designed and conducted to determine whether or not an operational amplifier IC was capable of rail-to-rail output voltage swing? What hypothesis

(i.e. prediction) might you pose for that experiment, and what result(s) would either support or disprove that hypothesis?

- What are some practical applications of operational amplifier ICs?
- If energy is always conserved, how is an amplifier able to output more energy than that of its input signal?
- How may we reliably predict the output signal polarity of a comparator circuit?
- What is *feedback* in any system?
- What are the specific characteristics of *negative feedback*?
- What is the difference between an opamp's open-loop gain and its closed-loop gain?
- Why are the power supply "rail" voltages for an amplifier important to its performance?
- How do block diagrams differ in form and function from schematic diagrams?
- How may we make a voltage regulator circuit using an opamp?
- How does a *series* voltage regulator work?
- How does a *parallel* voltage regulator work?
- What is a *dump load* and where might we find one applied to practical benefit?
- How may we make a current regulator circuit using an opamp?
- What is a servo system, and how may an opamp form the heart of one?
- What does it mean to "scale" one signal voltage range into another?
- What is a control system, and how may an opamp form the heart of one?
- What is the "Cardinal Rule of Negative Feedback" and why is it important to us?
- Under what condition(s) does the "Cardinal Rule of Negative Feedback" no longer hold true?
- How may we tell from the schematic diagram of a servo or control system circuit that negative feedback is in effect?
- What is *functional inversion* and how is this achieved through negative feedback?
- What must be done in an opamp circuit to make it have a specified gain?
- How do Kirchhoff's Laws and Ohm's Law apply to the analysis of an inverting amplifier circuit?
- How do Kirchhoff's Laws and Ohm's Law apply to the analysis of a non-inverting amplifier circuit?
- What is a *virtual ground*?
- Why is it important to master fundamental circuit laws with regard to opamp circuits?

- How does the Superposition Theorem apply to the analysis of differential amplifier circuits?
- How does Norton's Theorem apply to the analysis of summer circuits?
- What is a *voltage buffer*, and what practical purpose(s) does it serve?
- Why does the connection of a voltmeter to a circuit affect the voltage between those two points?
- What limits the ability of a precision current amplifier circuit to regulate current through a load?
- What is *differentiation*, and where are some applications of this mathematical concept in the physical world?
- What is *integration*, and where are some applications of this mathematical concept in the physical world?
- What is the *Barkhausen criterion*, and to what type of circuit does it apply?
- How may an operational amplifier be used to convert a square wave into a triangle wave?
- What is the purpose of the incandescent lamp in certain Wein bridge oscillator circuit designs?
- How does Fourier's Theorem relate to oscillator circuits?
- How does training contrast against education, and how do these concepts specifically apply to the analysis of opamp circuits?
- How is the slope-intercept equation $y = mx + b$ capable of describing any straight-line graph?

1.2 Challenging concepts related to operational amplifiers

The following list cites concepts related to this module's topic that are easily misunderstood, along with suggestions for properly understanding them:

- **Voltage as a relative quantity** – this is arguably one of the most difficult concepts students first encounter when studying electricity: understanding that voltage is a *difference* in energy levels between two locations, and not something that can exist at any single point. In this respect, voltage is similar to *distance* which is always something existing *between* two specified points. Even when voltage is specified at a single point (as is commonplace for opamp circuits) there must always be a point of reference (called *ground*) serving to give that single-point potential some necessary context.

- **Relevance of foundational concepts** – this is not a new concept in itself, but rather an approach to learning and to problem-solving that many students struggle with. Students must be encouraged to continually apply trusted concepts such as Ohm's Law, Kirchhoff's Voltage Law, and Kirchhoff's Current Law to every new subject they encounter in Electronics, rather than to ignore these first principles in favor of memorizing new rules. Memorization without understanding is a recipe for failure. Logically reasoning from trusted laws and principles, however, strengthens recall and hones the ability to tackle novel problems. Within the subject of operational amplifiers, Ohm's Law, KVL, and KCL are ever-relevant and need only be augmented by a few new facts (e.g. the effects of negative feedback, and the high input impedance of opamps) to fully equip oneself for analyzing most any opamp circuit.

- **Negative feedback** – this non-intuitive concept is difficult to grasp due its recursive nature: the output of an amplifier is fed back to its input, which affects its output, which affects its input, *ad infinitum*. A particular challenge with negative feedback is comprehending why this does not irresistibly lead to oscillations, given the fed-back signal's tendency to counter whatever direction the output signal is changing in. One rationale for the oscillation hypothesis is that we as human beings tend to imagine changes in discrete steps rather than as continuous trends, and so it is understandable why a new student might assume negative feedback causes the amplifier to leap back and forth between the (only) two extremes of signal range they might be envisioning. In reality, as the output of an amplifier with negative feedback drifts in one direction, the fed-back signal has the effect of slowing that drift down until it settles at some value rather than oscillate: as the "error" approaches zero, so does the rate of drift.

- **Output current direction** – just because a device pin is called the "output" does not necessarily mean that current must always exit that pin! Here, *output* refers to the directional flow of *information* rather than the directional flow of electric charge carriers. Current flow through the output terminal of an operational amplifier is best understood in terms of the connection formed between the output terminal and one of the power "rail" terminals by the device's final-stage transistors, which depending on the device and on its operating conditions may source out through the output terminal or sink in through it.

- **Differential versus Ground-referenced signals** – voltage can only exist between two points, and so any voltage-based signal may either be referenced to ground or expressed between two ungrounded points (differential).

- **Common-mode voltage** – "common-mode" voltage is the mathematical average of ground-referenced voltages among multiple points. Literally, it is the amount of voltage, measured in reference to ground, that is shared in "common" with all of those terminals. Operational amplifiers, being differential-input devices, ideally ignore common-mode voltage and only respond to differential voltage.

1.3 Recommendations for instructors

This section lists realistic student learning outcomes supported by the content of the module as well as suggested means of assessing (measuring) student learning. The outcomes state what learners should be able to do, and the assessments are specific challenges to prove students have learned.

- **Outcome** – Demonstrate effective technical reading and writing

  Assessment – Students present their outlines of this module's instructional chapters (e.g. Case Tutorial, Tutorial, Historical References, etc.) ideally as an entry to a larger Journal document chronicling their learning. These outlines should exhibit good-faith effort at summarizing major concepts explained in the text.

  Assessment – Students show how quantitative results were obtained by the author in the Tutorial chapter's examples.

  Assessment – Students present their "Conceptual Toolbox" Theory Journal entries listing important concepts applied in the Tutorial text(s). A "Conceptual Toolbox" serves as inspiration when solving problems, each concept suggesting a potential problem-solving action. Challenge students to collaboratively generate a Conceptual Toolbox for the day's topic during the class session and then use it when solving problems together, thus fostering the habit of reasoning from principles rather than recalling procedures!

- **Outcome** – Analyze voltages and currents in DC opamp circuits from first principles

  Assessment – Calculate all voltages and currents for inverting opamp circuits given component values and input signal conditions

  Assessment – Calculate all voltages and currents for inverting opamp circuits given component values and either one voltage or one current not at the input

  Assessment – Calculate all voltages and currents for non-inverting opamp circuits given component values and input signal conditions

  Assessment – Calculate all voltages and currents for non-inverting opamp circuits given component values and either one voltage or one current not at the input

  Assessment – Calculate all voltages and currents for summer opamp circuits given component values and input signal conditions

  Assessment – Calculate all voltages and currents for subtractor opamp circuits given component values and input signal conditions

  Assessment – Calculate all voltages and currents for current-regulating opamp circuits given component values and input signal conditions

  Assessment – Calculate all voltages and currents for current-regulating opamp circuits given component values and either one voltage or one current not at the input

  Assessment – Calculate all voltages and currents for precision rectifier opamp circuits given component values and input signal conditions

- **Outcome** – Design practical circuits using opamps

Assessment – Choose components and sketch accurate schematic diagrams for opamp amplification circuits providing a specified amount of voltage gain.

Assessment – Choose components and sketch accurate schematic diagrams for opamp summer circuits given a specified number of input channels.

Assessment – Choose components and sketch accurate schematic diagrams for opamp subtractor circuits providing a specified amount of voltage gain.

Assessment – Choose components and sketch accurate schematic diagrams for opamp current-regulator circuits providing a specified amount of current through a load of specified (minimum and maximum) resistance.

- **Outcome** – Diagnose a faulted opamp circuit

  Assessment – Predict the effect(s) of a single component failing either open or shorted in a negative-feedback opamp circuit.

  Assessment – Determine the probability of various component faults in a negative-feedback opamp circuit given symptoms and measured values.

  Assessment – Determine the appropriateness of various diagnostic tests in identifying the location and nature of a fault in an opamp circuit given symptoms and measured values.

- **Outcome** – Independent research

  Assessment – Read and summarize in your own words reliable source documents on the subject of operational amplifiers. Recommended readings include anything written by the early opamp researcher George A. Philbrick, as well as patents on opamp design innovations.

  Assessment – Read and summarize in your own words Harold Black's 1937 patent on negative feedback (US Patent 2,102,671 "Wave Translation System").

  Assessment – Locate operational amplifier datasheets and properly interpret some of the information contained in those documents including open-loop voltage gain, input impedance, output voltage limits, input voltage limits, input bias currents, gain-bandwidth product, output slew rate, etc.

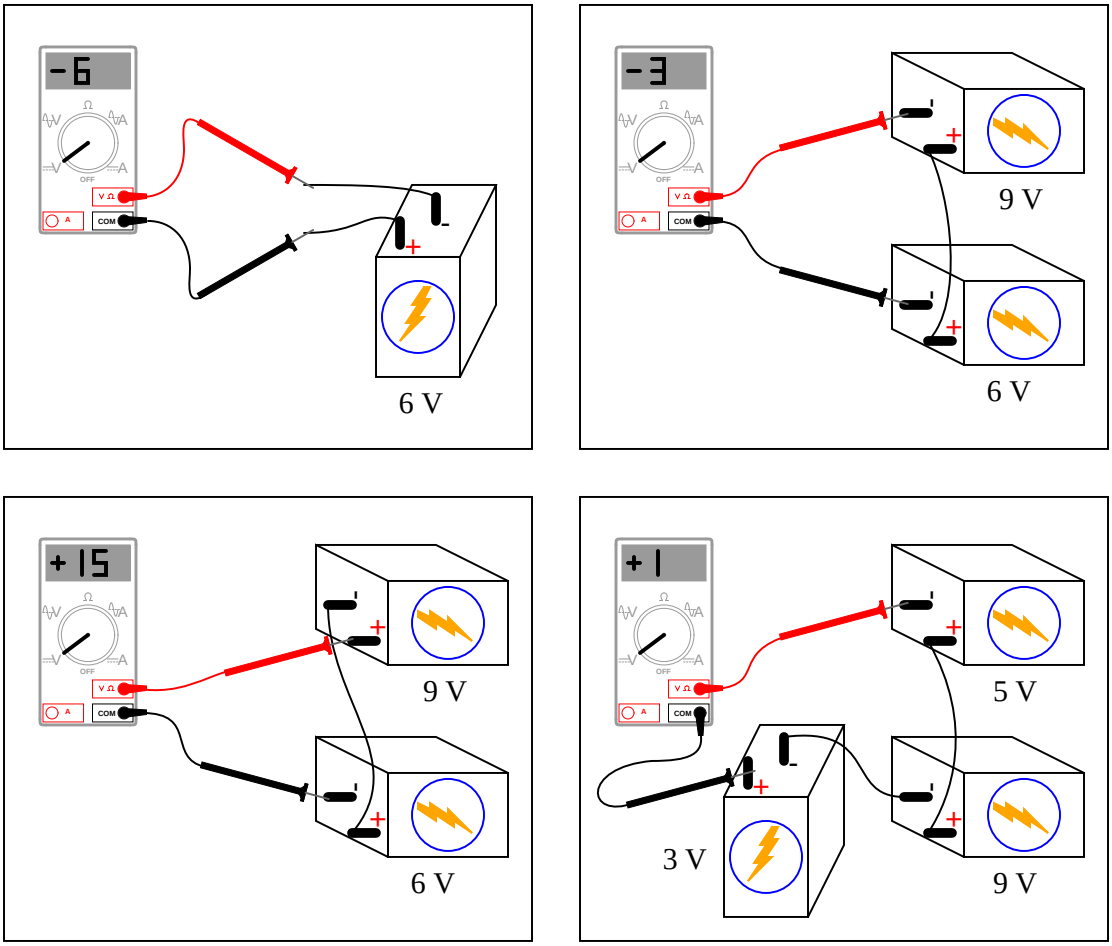
Chapter 2

Case Tutorial

The idea behind a *Case Tutorial* is to explore new concepts by way of example. In this chapter you will read less presentation of theory compared to other Tutorial chapters, but by close observation and comparison of the given examples be able to discern patterns and principles much the same way as a scientific experimenter. Hopefully you will find these cases illuminating, and a good supplement to text-based tutorials.

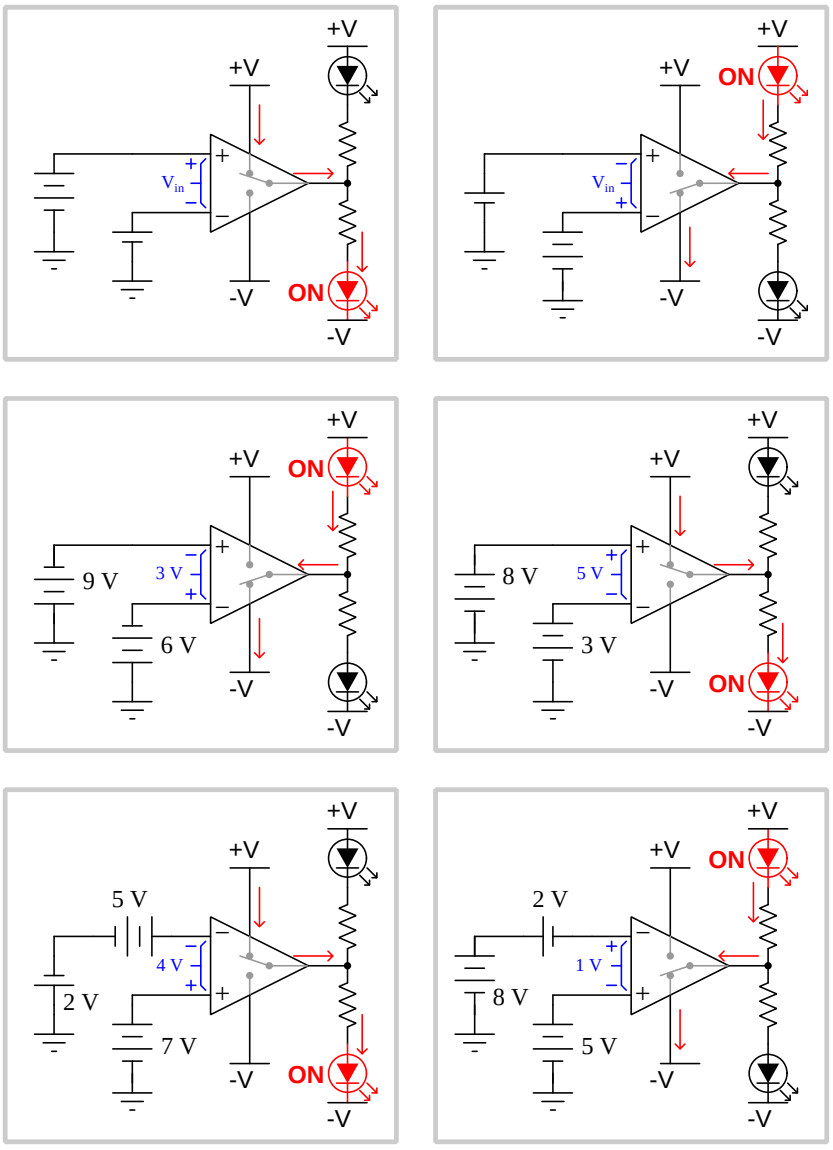
These examples also serve well as challenges following your reading of the other Tutorial(s) in this module – can you explain *why* the circuits behave as they do?

2.1 Example: Kirchhoff's Voltage Law in action



2.2 Example: open-loop opamp response to various inputs

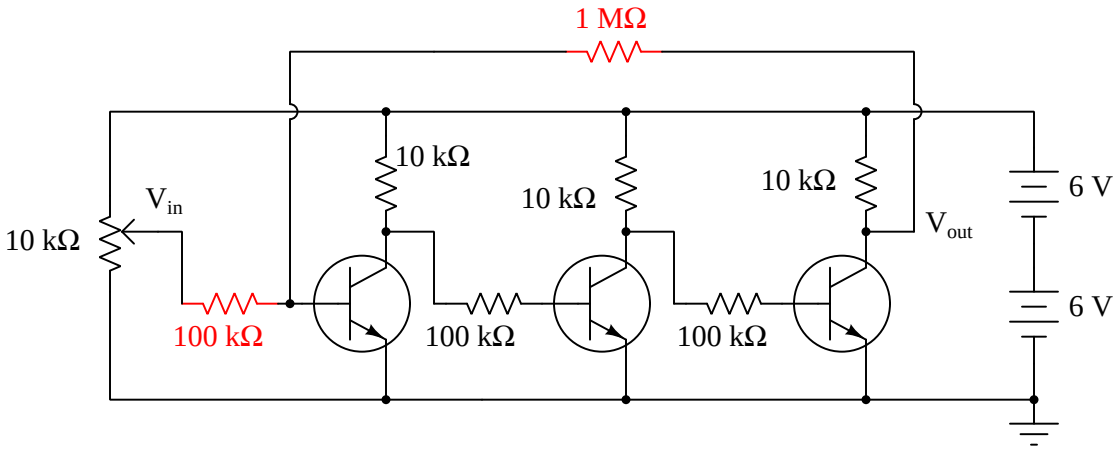
In these illustrations, I have likened the opamp's action to that of a single-pole, double-throw switch, showing the "connection" made between power supply terminals and the output terminal:



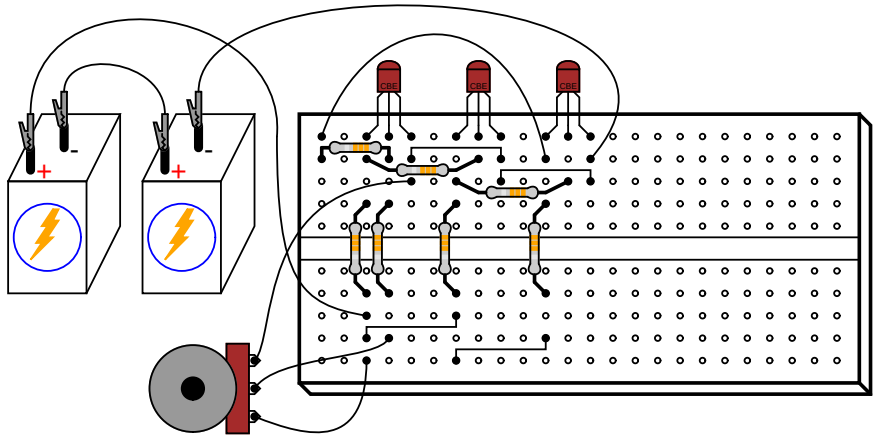
2.3 Example: multi-stage transistor amplifier with negative feedback

Single-stage transistor amplifiers are well-known for their somewhat unpredictable voltage gain behavior due to uncontrollable variations in transistor gain (beta for BJTs, and transconductance for FETs). However, if multiple stages are cascaded to produce a very high voltage gain, then negative feedback may be used to diminish that high voltage gain to a more reasonable (and much more stable!) value.

The following circuit shows how this may be done, using three grounded-emitter BJT amplifier stages to provide very high voltage gain and the inverting characteristic necessary for negative feedback:



The 1 MΩ feedback resistor in conjunction with the 100 kΩ "input" resistor yield a 10:1 ratio, translating into a voltage gain value that is nearly exactly 10.



Each stage of this amplifier circuit is *inverting*, in that a rising input signal causes the output signal to fall. Over-all, the entire amplifier is also inverting because there is an odd number of these inverting stages cascaded – we could have just as readily coupled five common-emitter stages together, or seven, or nine, and still maintained the inverting characteristic we need. This characteristic is important because we want the feedback to be *negative* in sign: the output signal's direction must fight against the causal input signal's direction in order to bring stability to the circuit.

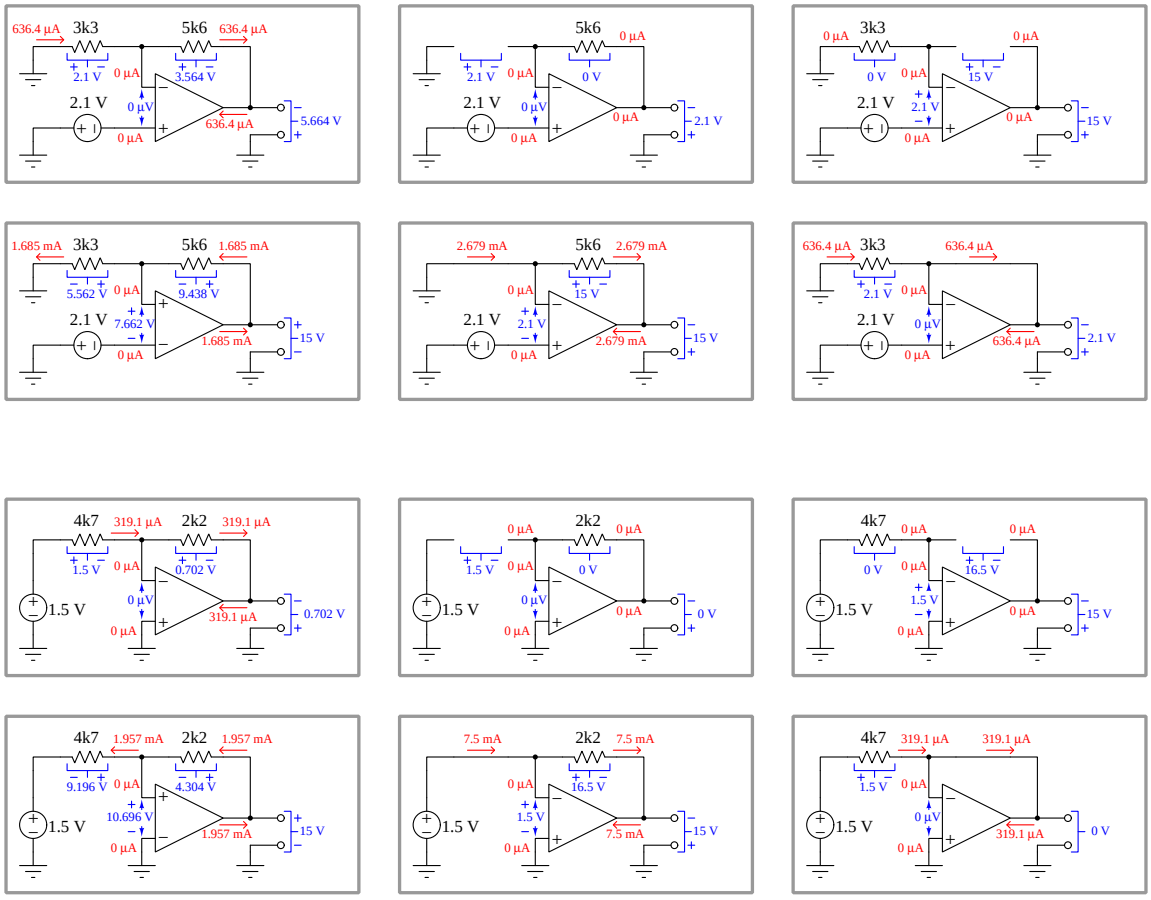
Although the notion of building an amplifier circuit to have an extremely high gain and then intentionally reducing that gain by way of negative feedback may seem counter-productive, the result is that the over-all gain becomes much more stable and the amplification becomes much more linear than with no feedback at all. As the transistors heat and cool, or if one must be replaced with another having slightly different characteristics, the strategy of arbitrarily high voltage gain intentionally scaled down via negative feedback results in those individual transistor variations having very little effect on the whole circuit's idealized performance.

This non-intuitive concept lies at the heart of Harold Black's 1920's era innovation, developed at a time when telephone repeater amplifiers used vacuum tubes as amplifying elements. These early tube circuits were notoriously unstable, the tubes' operating characteristics changing significantly as they aged. Black's application of negative feedback to the problem of electronic amplifier stability was revolutionary, and still holds merit with today's solid-state amplifier circuitry.

Operational amplifiers are perhaps the best and most common electronic application of this concept. Being integrate-circuit (IC) modules designed to provide extremely high voltage gain on their own, "opamps" are designed to be used in conjunction with negative feedback resistor networks like this to provide highly accurate and stable gains. This discrete transistor amplifier demonstrates the fundamental concept, but without the use of any integrated circuitry.

2.4 Example: opamp networks responding to various inputs

The following circuit examples assume each operational amplifier is powered by a split-voltage DC supply of ± 15 Volts (connections omitted for simplicity), and that the amplifier is capable of rail-to-rail output. Some of these circuit configurations represent practical amplifier circuits, while others represent *faulted* circuits (e.g. opened or shorted resistors):



2.5 Example: simulating opamp circuits using SPICE

One of the idealized components offered in SPICE is a *voltage-controlled voltage source*, which serves well to model the behavior of an ideal differential-voltage input amplifier. This particular SPICE component, identified by the label `E`, is specified by its two input terminals, two output terminals, and voltage gain value. The general format for this component is shown below:

[**E***name*] [+output_node_ID] [−output_node_ID] [+input_node_ID] [−input_node_ID] [*gain*]

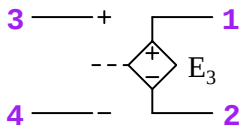
An example of a voltage-controlled voltage source is shown below:

- Name = *3*
- Polarity = *input on nodes 3 (+) and 4 (−) ; output on nodes 1 (+) and 2 (−)*
- Gain = *0.5 Volts output per Volt input*

Spice element description

```
E3 1 2 3 4 0.5
```

Schematic representation



$V_{out} = (0.5)(V_{in})$

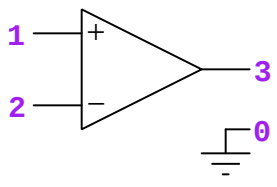
Operational amplifiers, of course, have extremely high differential voltage gains. An example of a voltage-controlled voltage source being used to model an opamp is shown below:

- Name = *opamp*
- Polarity = *input on nodes 1 (+) and 2 (−) ; output on nodes 3 with reference to ground (0)*
- Gain = 1×10^6

Spice element description

```
Eopamp 3 0 1 2 1e6
```

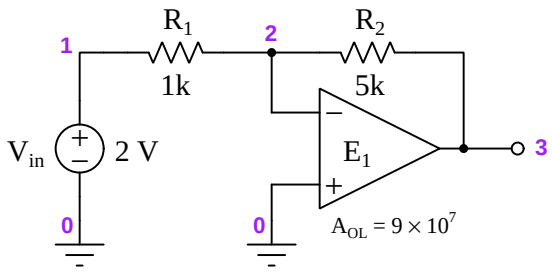
Schematic representation



$V_{out} = (1 \times 10^6)(V_{in})$

In the following SPICE simulation we will simulate a simple inverting amplifier:

Circuit schematic diagram (with node numbers listed):



SPICE netlist:

```
* SPICE circuit
vin 1 0
r1 1 2 1000
r2 2 3 5000
e1 3 0 0 2 9e7
.dc vin 2 2 1
.print dc v(2) v(3)
.width out=80
.end
```

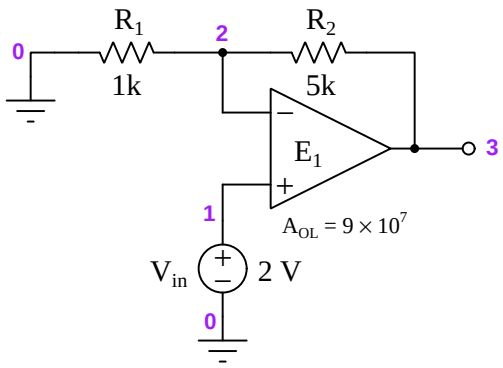
NGSPICE version 26 analysis (using batch mode on Linux console):

```
--------------------------------------------------
Index   v-sweep         v(2)            v(3)
--------------------------------------------------
0       2.000000e+00    1.11111e-07    -1.00000e+01
```

With an R_2/R_1 ratio of 5:1, this inverting amplifier circuit should have a gain of -5 ($-\frac{R_2}{R_1}$). This is proven by the fact it outputs -10 Volts for a V_{in} value of 2 Volts. Node 2 is the "virtual ground" point within this circuit, with negligible potential compared to the real ground (node 0).

Next, we will simulate a simple noninverting amplifier:

Circuit schematic diagram (with node numbers listed):



SPICE netlist:

```
* SPICE circuit
vin 1 0
r1 2 0 1000
r2 2 3 5000
e1 3 0 1 2 9e7
.dc vin 2 2 1
.print dc v(2) v(3)
.width out=80
.end
```

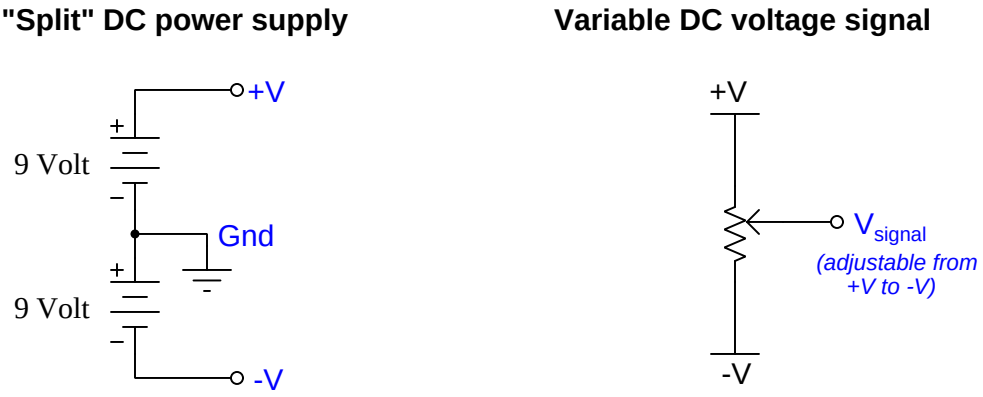
NGSPICE version 26 analysis (using batch mode on Linux console):

```
--------------------------------------------------
Index   v-sweep         v(2)            v(3)
--------------------------------------------------
0       2.000000e+00    2.000000e+00    1.200000e+01
```

With an R_2/R_1 ratio of 5:1, this inverting amplifier circuit should have a gain of 6 ($\frac{R_2}{R_1} + 1$). This is proven by the fact it outputs 12 Volts for a V_{in} value of 2 Volts. The voltage measured at node 2 with respect to ground matches V_{in} with negligible error.

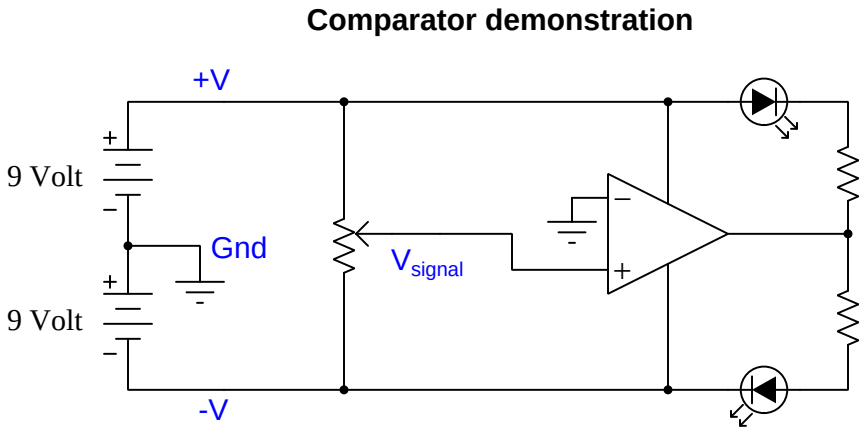
2.6 Example: simple opamp experiments

A very simple arrangement for exploring fundamental comparator and operational amplifier circuit behavior may be fashioned using a pair of 9-Volt batteries, a potentiometer, and a single operational amplifier as such:

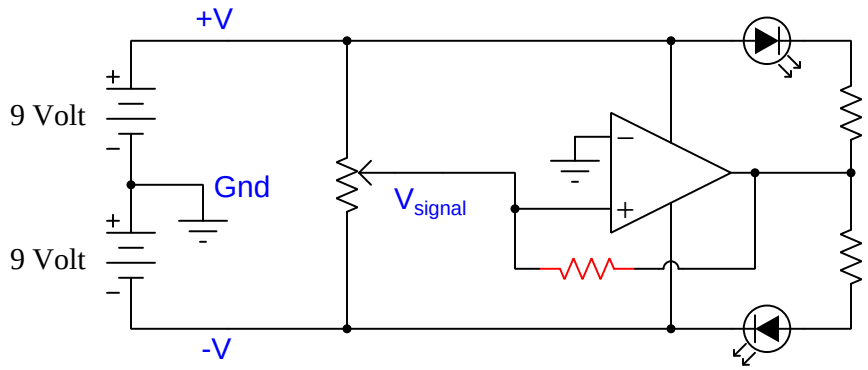


The "split" DC power supply provides a center connection (called Ground) and two "rail" connections between which we may measure voltages spanning +9 Volts (from +V to Ground) to −9 Volts (from -V to ground). The potentiometer divides the total 18 Volt differential into proportions adjustable by its wiper position, so that a voltmeter measuring between the wiper terminal (V_{signal}) and Ground may experience anything between +9 Volts and −9 Volts inclusive.

With this split DC power supply and adjustable signal source, we may build a range of different comparator and opamp circuits useful for self-instruction. In all cases I recommend choosing resistor values between 1 kΩ and 100 kΩ, and using a TL082 or comparable operational amplifier.

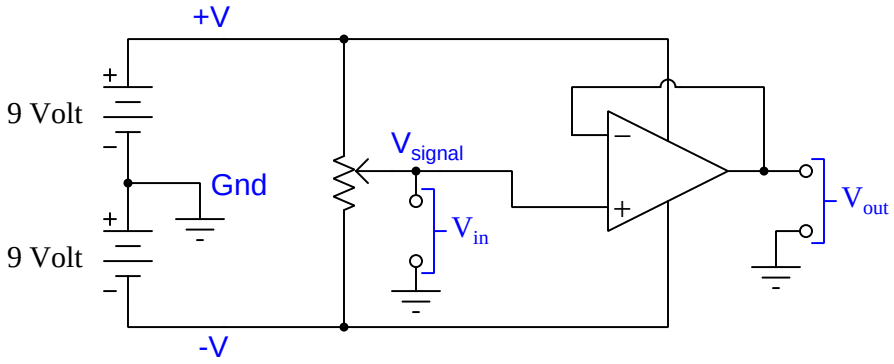


Comparator demonstration (with hysteresis)



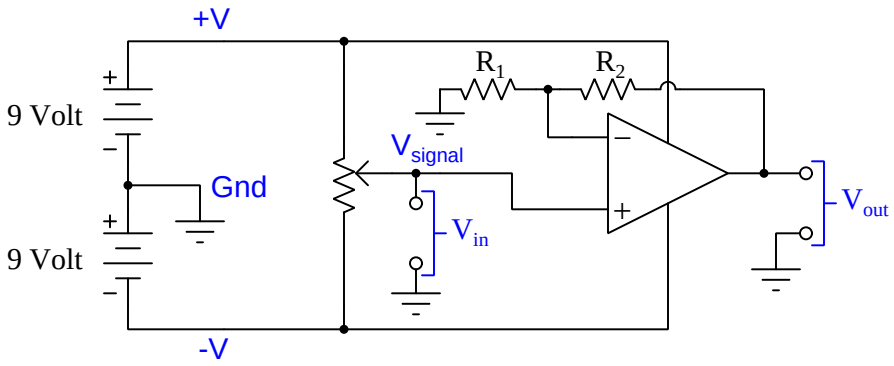
Positive-feedback resistor should be much larger in value than the potentiometer

Voltage follower demonstration



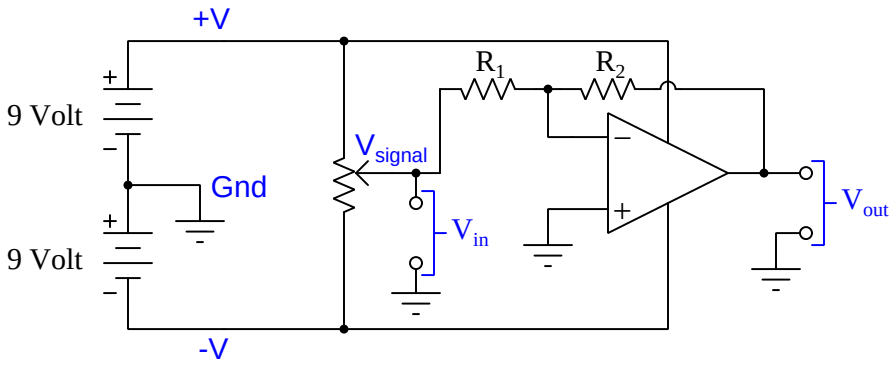
V_{out} should equal V_{in}

Non-inverting amplifier demonstration



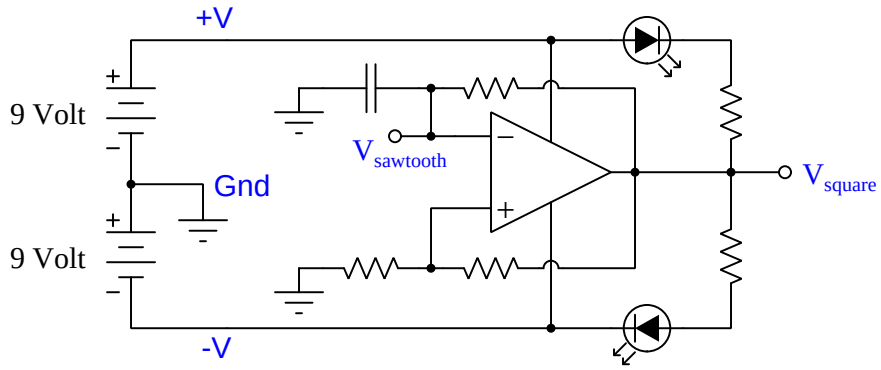
V_{out} should be equal to $V_{in}(R_2/R_1 + 1)$

Inverting amplifier demonstration

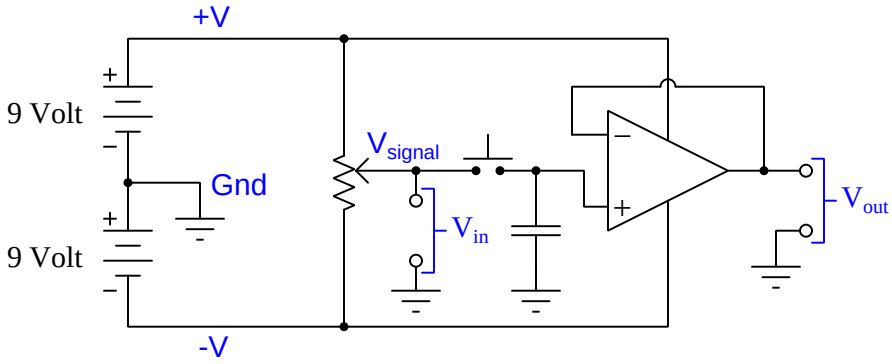


V_{out} should be equal to $V_{in}(-R_2/R_1)$

Oscillator demonstration



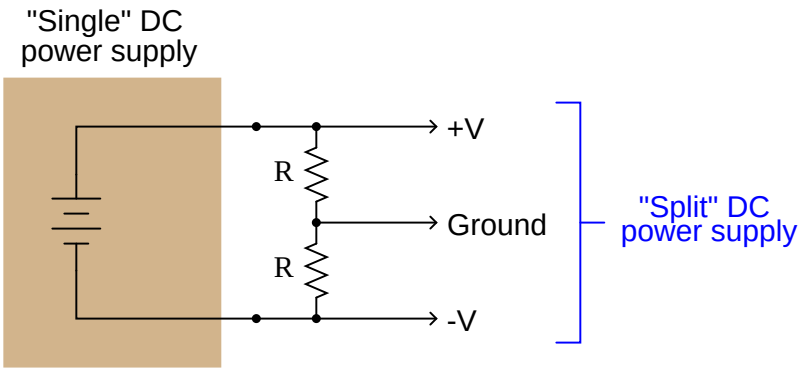
Sample-and-hold voltage follower demonstration



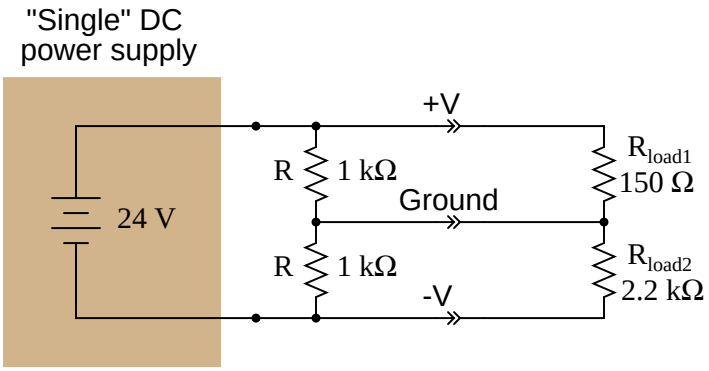
V_{out} should equal last V_{in} value when button was pressed

2.7 Example: DC power supply rail splitter

"Split" or "dual" DC power supplies are essential for powering many types of electronic circuits, especially certain types of operational amplifier circuits. If only a "single" DC power supply is not available, a "split" power supply may be roughly simulated through the use of a resistive voltage divider:



The problem with doing this is loading: if more current is drawn from one of the power supply rails than from the other, the "split" of voltage will become uneven. The only way that +V and −V will have the same (absolute) voltage value at the load is if the load impedance is balanced evenly between those rails and ground. This scenario is unlikely. Take for instance this example:

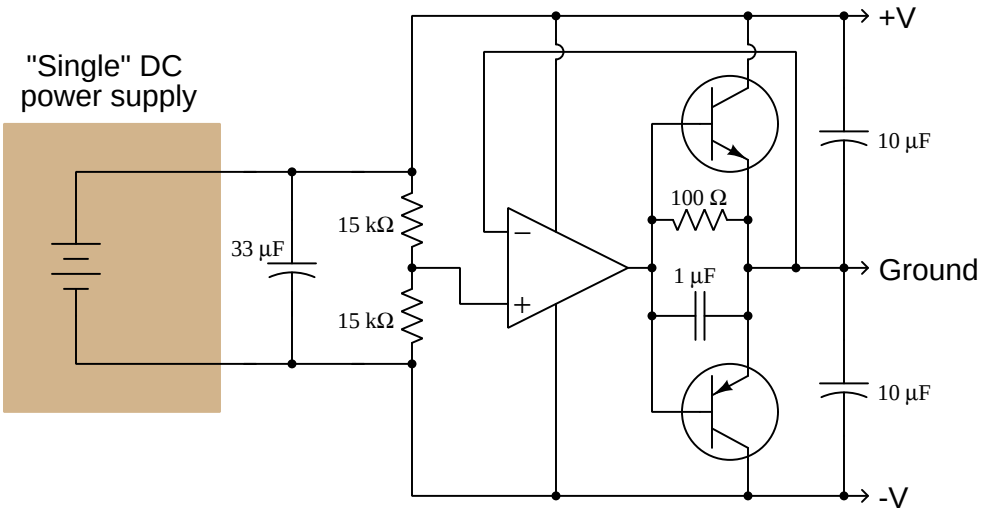


Voltage from +V to Ground = 3.827 V

Voltage from -V to Ground = -20.173 V

One way to improve the balancing of +V and −V potentials in this voltage divider circuit is to use smaller-resistance R resistors. However, this solution has the disadvantage of wasting more energy in the form of heat (dissipated by those two R resistors).

A simple opamp circuit, though, can correct this problem and maintain an even "split" of voltage between +V, Ground, and −V:

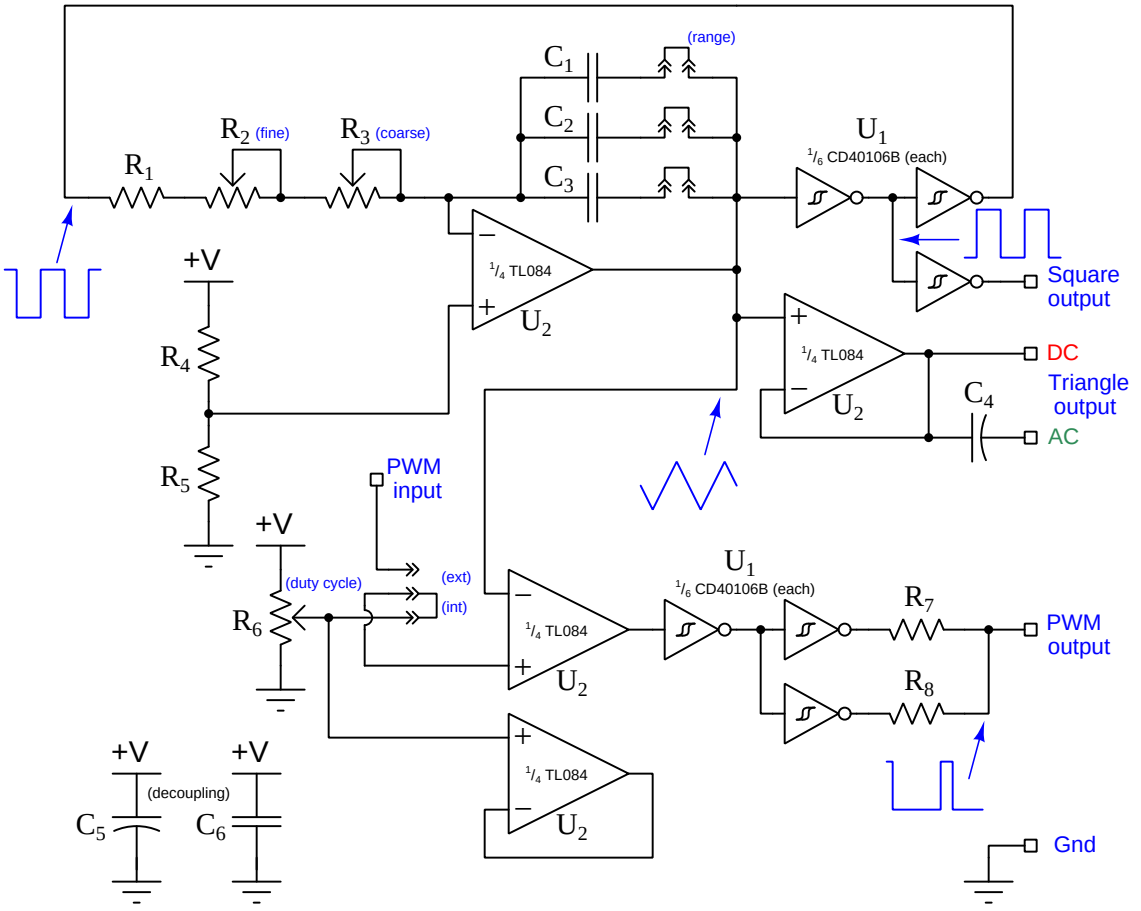


The PNP and NPN power transistors serve as a "push-pull" amplifier stage boosting the opamp's ability to source and sink current from any load(s) connected to the Ground terminal, allowing it to maintain a "stiffer" ground potential exactly mid-way between the two power supply rails than before. The two 15 kΩ resistors are large enough to keep wasted power at a minimum because the current at the opamp's noninverting input will be negligible, meaning that voltage divider will experience practically zero loading.

The 100 Ω resistor and 1 μF capacitor help this circuit achieve stability, as it would otherwise be prone to oscillation with capacitance connected at the output terminal(s). Note how the opamp's output must swing up to +0.7 Volts before the upper transistor turns on, and swing down to −0.7 Volts before the lower transistor turns on. This means the opamp will be prone to swinging its output back and forth between these two values in an attempt to stabilize ground voltage, with both transistors being off between these output voltage values. Placing a low-value resistor directly from the opamp's output terminal to the ground terminal allows the opamp to source and sink current *immediately* as necessary, with the transistors turning on only if the voltage drop across that resistor becomes large enough. The 1 μF capacitor also connecting the opamp's output to ground provides an even more "direct" connection for signals with fast rise/fall times, the capacitor effectively acting as a "short" to any rapidly rising or falling signal. Thus, the opamp is able to maintain a stable ground potential on its own, and only "calls for help" from either of the two power transistors if ground current rises to a high enough level from powering any load(s) connected between either +V or −V rail and ground. The resistor should be sized such that an opamp output current less than its maximum rating results in a voltage drop large enough to guarantee transistor saturation. For a 100 Ω resistor and a standard 0.7 Volt drop between base-emitter to turn on either transistor, the opamp will source or sink as much as 7 mA of current before either transistor begins to conduct and relieve the opamp of any further current burden.

2.8 Example: simple triangle/square/PWM oscillator

This simple oscillator circuit provides triangle wave output (both DC- and AC-coupled) as well as square wave output and pulse-width-modulated ("PWM") square wave output with an adjustable duty cycle. At the heart of the circuit is an operational amplifier wired as an *integrator*, sending a constant current through one or more capacitors to create a linearly-ramping voltage. A set of digital inverter gates with Schmitt-trigger inputs senses this triangle wave signal and feeds a square-wave signal back to the integrator's input to continue the cycle. Jumpers select capacitor values for frequency range, while two potentiometers provide coarse and fine frequency adjustment:



DC power source voltage may range widely, given the advertised operating voltage ranges for the CD40106B and TL084 integrated circuits (3 to 18 Volts for the hex inverter and 10 to 30 Volts for the opamp). I have successfully tested this circuit at a supply voltage as low as 5 Volts and it seems to work well.

It is worth noting that the fairly close similarity between triangle and sine wave-shapes means students may use the triangle-wave output of this oscillator circuit as an AC voltage source to

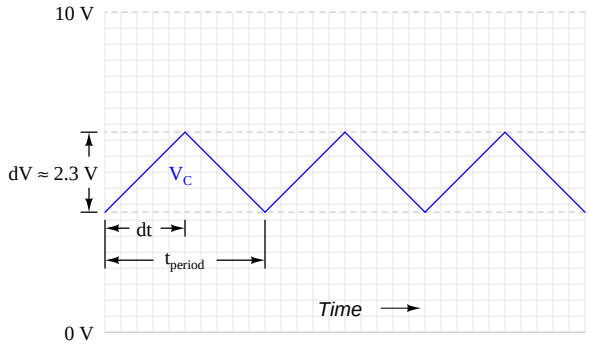
energize simple RLC networks rather than using a true sine-wave signal generator signal. The error between measured and calculated values of voltage and current are generally within the ± 5% tolerance of common resistors, capacitors, and inductors when using a true-RMS multimeter to take the measurements.

Parts list:

- U_1 = CD40106B CMOS hex inverter with Schmitt trigger inputs
- U_2 = TL084 quad operational amplifier
- R_1 = 10 kΩ – *this sets the maximum frequency when both pots are at minimum resistance*
- R_2 = 10 kΩ potentiometer – *this is the "fine" frequency adjustment*
- R_3 = 100 kΩ potentiometer – *this is the "coarse" frequency adjustment*
- R_4 = 10 kΩ – *the values of R_4 and R_5 are not critical, so long as they are equal*
- R_5 = 10 kΩ
- R_6 = 10 kΩ potentiometer – *the values of R_6 is not critical*
- R_7 = 100 Ω – *R_7 and R_8 equalize load-sharing between inverters, the values being non-critical*
- R_8 = 100 Ω
- C_1 = 0.0047 μF ceramic (for oscillator timing)
- C_2 = 0.047 μF ceramic (for oscillator timing)
- C_3 = 0.47 μF ceramic (for oscillator timing)
- C_4 = 4.7 μF electrolytic (for output coupling – value not critical)
- C_5 = 47 μF electrolytic (for power supply decoupling – value not critical)
- C_6 = 1 μF ceramic (for power supply decoupling – value not critical)

The triangle wave signal begins to exhibit "ringing" at the positive and negative peaks if the total oscillator network resistance (i.e. $R_1 + R_2 + R_3$) becomes too low (approximately 20 kΩ or less). This is the major function of the fixed resistor R_1, to make sure this resistance never becomes unreasonably low.

The upper-most opamp functions as a current source for capacitors C_1 through C_3, driving a constant current through the capacitor(s) in order to produce a ramping voltage across them. The amount of current driven through the capacitor(s) is equal to the difference between the hex inverter's output voltage (square wave) and one-half supply voltage (set by divider R_4/R_5) divided by the resistance of $R_1 + R_2 + R_3$, which should be approximately $\pm\frac{V_{supply}}{2(R_1+R_2+R_3)}$. The capacitor voltage will ramp at a rate $\frac{dV}{dt}$ equal to that current divided by capacitance C. The CD40106B Schmitt-trigger inverter has a typical hysteresis value of approximately[1] 23% (e.g. 2.3 Volts for a 10-Volt V_{supply}), so the time duration of each half-cycle of the waveform is the amount of time it takes the capacitor voltage to ramp that 23% of supply voltage as opamp U_2 forces a constant current:



Here is a mathematical derivation for output frequency based on these conditions, where C represents the capacitance offered by C_1, C_2, and/or C_3 (depending on which jumpers are connected) and R represents the total series resistance of R_1, R_2, and R_3:

$$I_C = C\frac{dV}{dt} \text{ ("Ohm's Law" for any capacitor)} \qquad I_C = \frac{V_{supply}}{2R} \text{ (Current forced by opamp)}$$

$$C\frac{dV}{dt} = \frac{V_{supply}}{2R}$$

$$\frac{dV}{dt} = \frac{V_{supply}}{2RC}$$

$$dV = \frac{V_{supply}}{2RC}dt$$

$$0.23V_{supply} \approx \frac{V_{supply}}{2RC}dt$$

$$(0.23)(2RC) \approx dt$$

[1] This figure of 23% is merely a rough estimate, and should not be relied upon for accuracy. Schmitt-trigger logic gates are designed for noise tolerance, not to function as precision comparators, and so it is unwise to assume this value will be stable from IC to IC or even between different operating frequencies in this circuit!

$$t_{period} \approx (2)(0.23)(2RC)$$

$$f \approx \frac{1}{(2)(0.23)(2RC)} \approx \frac{1.09}{RC}$$

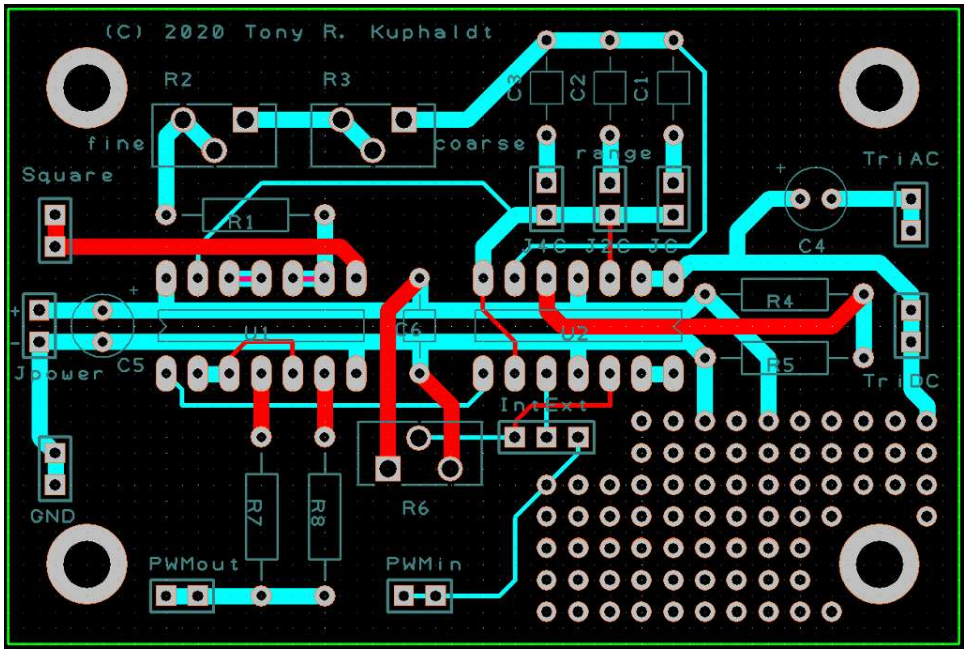
Assuming the component values shown in the parts list, with both potentiometers set to mid-position and only the middle capacitor (C_2) selected by jumper, the frequency should be approximately:

$$f \approx \frac{1.09}{(65 \times 10^3)(0.047 \times 10^{-6})} = 356 \text{ Hz}$$

I recommend choosing capacitor values of octave multiples (i.e. C, $2C$, $4C$) or decade multiples (i.e. C, $10C$, $100C$). With a large "coarse" potentiometer value (R_3), decade multiples make the most sense. Lower-valued frequency-adjust potentiometers would better suit octave-multiples of capacitor value, but then the total adjustment range of the oscillator will be more limited.

The int/ext jumper sets the signal source for the PWM signal's adjustable duty cycle. When set for "internal" the duty cycle is adjusted by potentiometer R_6; when set for "external" a signal voltage input at the "PWM input" terminal controls the duty cycle.

One possible PCB layout for this circuit is shown below, the dimensions of the board being 3 inches by 2 inches. Red traces are on the top copper layer of the board, while blue traces are on the bottom. DC supply power comes into the board at the left-hand edge, at the `Jpower` terminals:



Note how two pads are provided for each of the I/O terminals. This was done to facilitate a small loop of bare wire soldered between the two pads, which makes an easy connection point for "grabber" hooks or "alligator" clips. The four large pads at the board's corners are intended for nylon stand-offs, for mounting this PCB to a larger panel. These mounting pads are spaced 2.5 inches horizontally and 1.5 inches vertically from each other.

An important detail to note on this layout is the orientation of the two integrated circuits. U_1 (the CMOS hex inverter) has its pin 1 located on the left side, while U_2 (the quad opamp) is "upside-down" with its pin 1 located on the right. This is due to the locations of the DC power pins on each of these integrated circuits. In order to avoid crossing the +V and Ground power traces on the board, these two chips had to be positioned opposite-facing.

A "prototyping" area with an array of unallocated pads exists at the lower-right corner of the PCB, provided so that you may add other components to the board (e.g. power amplifier, wave-shaping networks) to enhance its capabilities.

2.9 Example: measuring signal rates of change

Capacitors and inductors relate voltage to current by rates of change. For a capacitor, the amount of current is proportional to how quickly voltage across that capacitor either rises or falls over time ($I = C\frac{dV}{dt}$). For an inductor, the amount of voltage is proportional to how quickly current through that inductor rises or falls over time ($V = L\frac{dI}{dt}$).

For example, a 330 microFarad capacitor experiencing a voltage increasing at a rate of 45 Volts per second will pass 14.85 milliAmperes. If the voltage happens to decrease at an equivalent rate (i.e. $\frac{dV}{dt} = -45$ Volts per second) then the 14.85 milliAmpere current will reverse direction through the capacitor compared to how it flowed with the increasing voltage.

Similarly, a 100 milliHenry inductor experiencing a current increasing at a rate of 5 Amperes per second will induce a voltage of 500 milliVolts. If the current happens to decrease by an equivalent rate (i.e. $\frac{dI}{dt} = -5$ Amperes per second) then the 500 milliVolt voltage induced across the inductor will reverse polarity from what it was during the period of increasing current.

Not only are rates-of-change important in determining how energy-storing devices such as capacitors and inductors will respond in circuits, but rates-of-change are also important for determining how *parasitic* capacitances and inductances will affect intended circuit behavior. Parasitic capacitance exists between any two conducting surfaces separated by an electrically insulating medium, and parasitic inductance exists along any length of conductor. This means any rate-of-change of voltage over time between two separated conductors will cause some amount of current to "pass"[2] between them, and that any rate-of-change of current over time through any single conductor will cause some amount of voltage to drop across its length. In many circuits these parasitic effects are negligible, but in circuits experiencing extremely fast rates of change for voltage and/or current the effects can be significant or even severe.

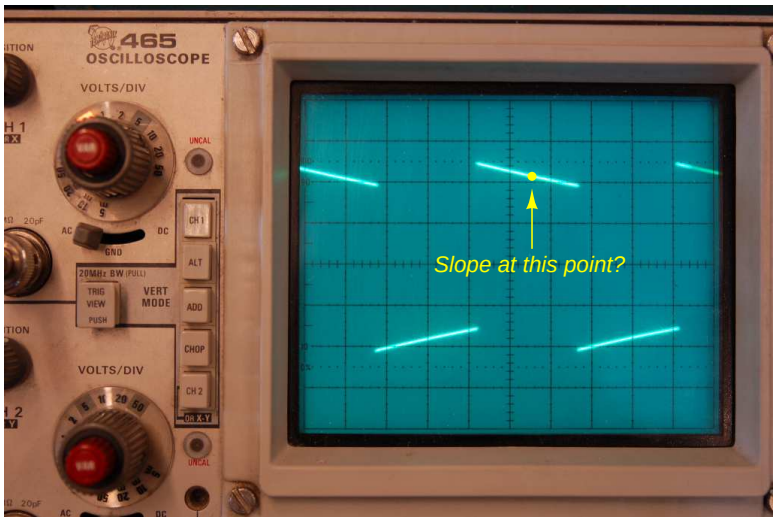
Oscilloscopes are ideal for performing empirical measurements of voltage rates-of-change, and of current rates-of-change given the proper accessories[3]. Some skill is required to do this, though, and here we will explore practical examples to show how it is done.

For any signal plotted in the time domain, where the horizontal axis of the plot is expressed in units of seconds, milliseconds, microseconds, etc., the signal's rate of change at any given point will be the *slope* or *pitch* of the waveform, mathematically defined as its *rise over run*. A great aid to discerning slope at any location on a waveform is to sketch a straight line visually matching the wave's slope at that point, then use locations along that straight line to more easily discern how far it rises (or falls) over some "run" of time. We call this straight line a *tangent line*.

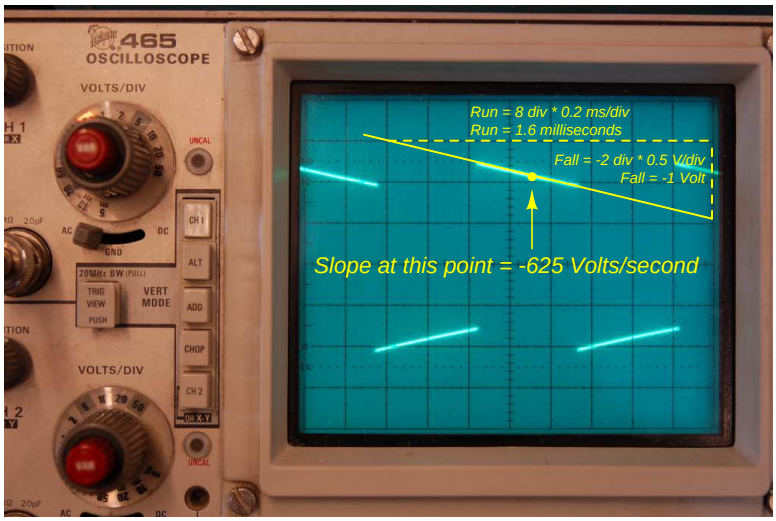
[2] The Scottish physicist James Clerk Maxwell referred to this current as *displacement current*.

[3] For example, a *current probe* converting a sensed current into a voltage the oscilloscope may directly sense, or a *shunt resistor* placed in the circuit developing an oscilloscope-measurable voltage drop for any current passing through.

Here we see an example of a waveform with sloping sections. In the first image we see a specific location on the waveform where we wish to measure voltage rate-of-change ($\frac{dV}{dt}$):



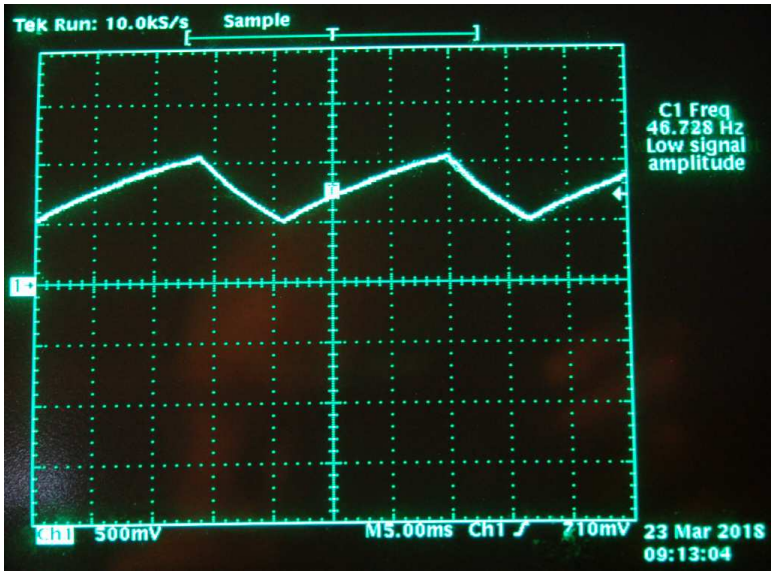
Next we see a tangent line drawn to match the slope of the waveform at the specified location, with convenient points for fall/run measurements taken on that line against the oscilloscope grid's major divisions. In this example, the oscilloscope's vertical sensitivity has been set for 0.5 Volts per division, and the horizontal timebase for 0.2 milliseconds per division:



As we can see, the tangent line falls 2 vertical divisions (-1 Volt) over a timespan of 8 horizontal divisions (1.6 milliseconds), yielding a $\frac{dV}{dt}$ quotient of -625 Volts per second, which may also be expressed as -0.625 Volts per millisecond. The negative sign is important, as it distinguishes this particular rate-of-change as *falling* rather than rising over time.

Such rate-of-change measurements are necessarily approximate, as they require us to visually gauge where a tangent line may be overlaid on the waveform's oscillograph, and also to visually assess the slope of that tangent line using the grid provided on the instrument's display screen. However, in most applications extremely precise rate-of-changes are not necessary, and such techniques suffice quite well.

Below is another example of a waveform with sloping sections, the oscilloscope configured for 500 milliVolts per division on the vertical axis and 5 milliseconds per division on the horizontal:

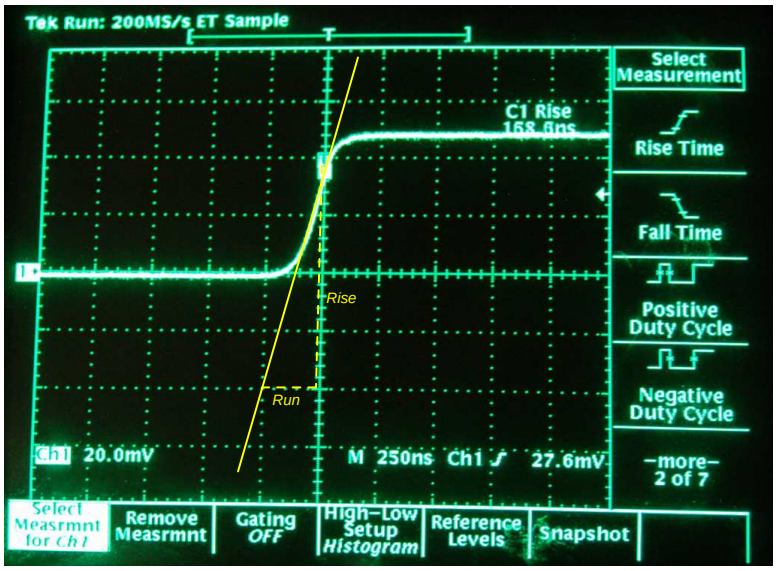


Approximating the slope for each rising section of this wave, we count one division of rise over 2.8 divisions of run, or 500 milliVolts rise over 14 milliseconds of run. This is a rate-of-rise of +35.7 Volts per second.

Approximating the slope for each falling section of this wave, we count one division of fall over 1.4 divisions of run, or −500 milliVolts fall over 7 milliseconds of run. This is a rate-of-fall of −71.4 Volts per second.

If these rates-of-change appear suspiciously large compared to the actual amplitude of the waveform, which barely crests over +1 Volt on the oscillograph, bear in mind that we are calculating *rates of change* for voltage and not absolute values of voltage itself. This is analogous to the distinction between *speed* and *distance*: traveling at a rate of 30 kilometers per hour does not necessarily mean you will travel 30 kilometers, as the actual distance traveled depends on how long that speed is sustained. A voltage rising at a rate of 35.7 Volts per second would indeed rise 35.7 Volts *if given a full second to do so*, but since each rising/falling portion of this waveform is so short in duration the actual amount of rise or fall in each case is only one-half of one Volt. It is therefore perfectly appropriate to consider any $\frac{dV}{dt}$ value as being the *speed* at which a voltage increases or decreases over time, distinct from the actual value of that voltage at any particular moment in time.

Here we see another oscillograph, this one zoomed into the rising edge of a square wave. For this measurement the oscilloscope was configured for 20 milliVolts per division of vertical sensitivity and a timebase of 250 nanoseconds per division on the horizontal:



The tangent line overlaid on this screenshot for the purpose of measuring the pulse edge's rate-of-change rises approximately 4 vertical divisions over a run of 1 division, which is 80 milliVolts of rise over 250 nanoseconds of run. The pulse edge's rate-of-change, therefore, is approximately +320,000 Volts per second, or +320 Volts per millisecond, or +0.32 Volts per microsecond (all equivalent expressions of $\frac{dV}{dt}$).

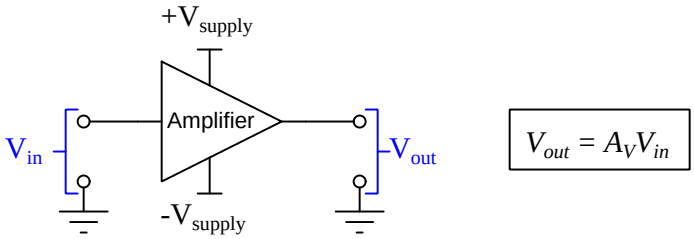
Chapter 3

Tutorial

3.1 Electronic amplifiers

An electronic *amplifier* is a circuit designed to input a relatively low-power signal and output a relatively high-power signal. In order to generate this higher-power signal, an amplifier circuit must be connected to an electrical power supply such as a DC voltage source. The "active" components inside the amplifier work to *control* the flow of energy from this external power supply to the output terminals where a load may connect.

The generic symbol for a simple electronic amplifier appears below, the complex internal circuitry of the amplifier represented in total by a simple triangle shape with the input terminal on the wide end and the output terminal at the far point:

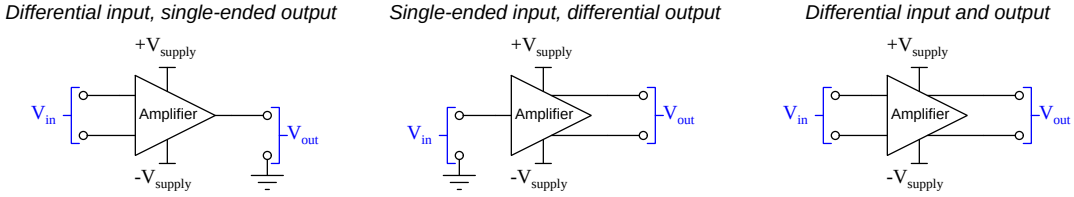


This simple amplifier has both a *single-ended* input and a *single-ended* output, meaning that both signals are conveyed by single terminals (with ground as a common reference point for each, since voltage is always defined between two points). The *gain* of the amplifier is shown here as a ratio of output voltage to input voltage, symbolized by A_V.

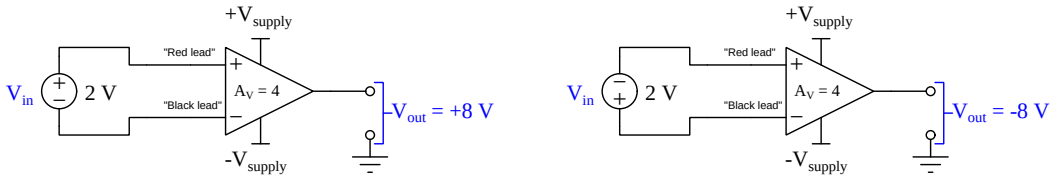
Amplifiers designed to input a voltage signal typically exhibit extremely high input impedance: that is to say, the input terminal current is practically zero. In fact, for most applications the input current is so vanishingly small that we may safely assume it to be zero without any noticeable effect in our calculations, much like we generally assume a voltmeter draws no current from the circuit it's measuring.

3.2 Differential inputs and outputs

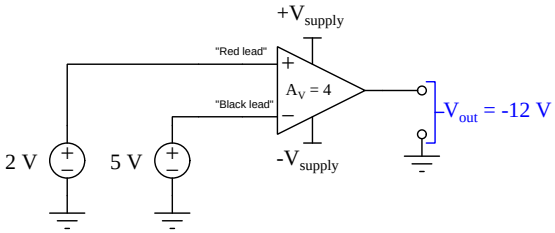
Other types of input and output exist for electronic amplifiers, though. Shown below are three other variations, each one sporting at least one *differential* signal where the voltage in question is not referenced to ground but rather between two dedicated terminals:



Of these, perhaps the most useful is the form having a differential input and a single-ended output. For the sake of clarity, the differential input terminals of such an amplifier are given + and − labels to distinguish their respective roles, roughly analogous to the red and black test leads on a DC voltmeter, respectively. Consider the behavior of such an amplifier connected in two different ways to a DC voltage signal source:



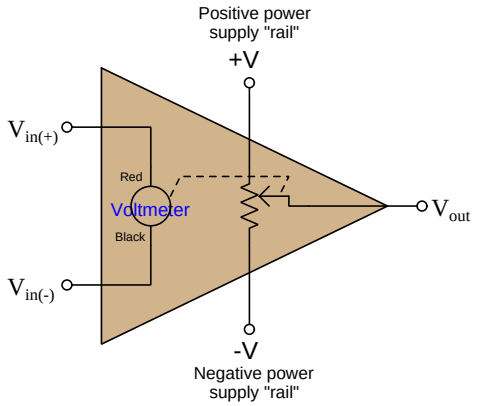
Differential-input amplifiers are useful for *comparing* two signals, the output being a multiple of the *difference* between those two signals' strengths at any given time. The following example shows this in action, with a +2 Volt signal connected to the "+" input terminal and a +5 Volt signal connected to the "−" terminal:



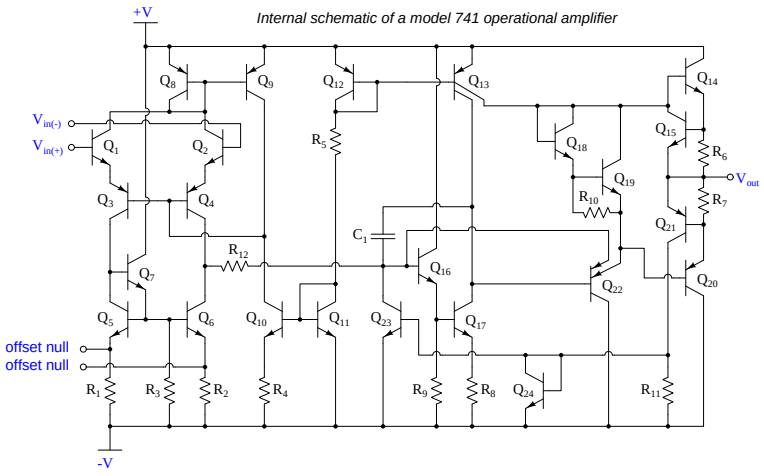
The difference between +2 Volts and +5 Volts, of course, is 3 Volts. The amplifier considers this a −3 Volt differential signal because the more positive of the two input signals connects to the amplifier's "−" terminal, just as a DC voltmeter would register −3 Volts if connected with its red lead at +2 Volts and its black lead at +5 Volts (both with respect to ground). With its voltage gain of 4, the output voltage signal of this amplifier circuit becomes four times the difference, or −12 Volts.

In honor of the respective effects of voltage signals applied to each input terminal, the "–" terminal of a differential amplifier is called *inverting* and the "+" terminal *non-inverting*.

A helpful model for envisioning the function of a differential amplifier is that of a voltmeter-driven potentiometer, the voltmeter being an electromagnetic type with its moving pointer linked to the wiper of a potentiometer. The output voltage of the amplifier is necessarily limited to the positive and negative power supply poles (commonly called *rails*):



Of course, a real differential amplifier is comprised solely of solid-state components. Our voltmeter-potentiometer model merely serves to demonstrate how the single-ended output voltage responds to a differential input voltage, and how the output signal is limited by the power supply. A schematic diagram showing the actual internal circuitry of a popular legacy differential amplifier, the model 741, is shown below for comparison:



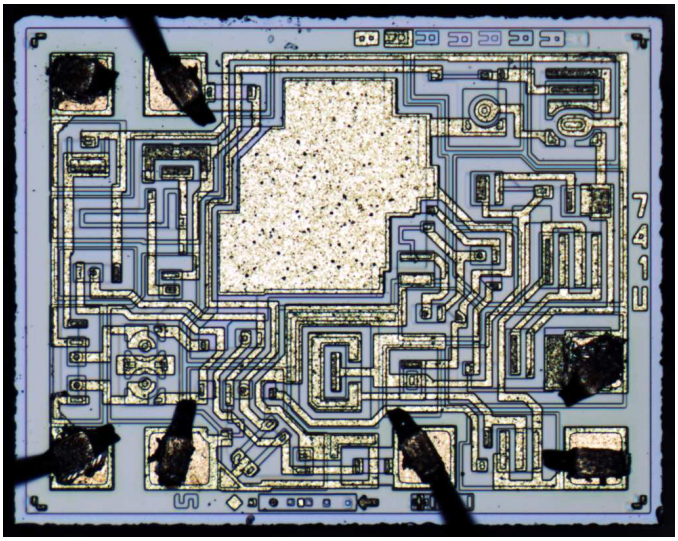
Differential inputs allow electronic amplifiers to be used as *comparators* to compare the magnitudes of two different signals and declare which one of those signals is larger by asserting either a "high" (saturated positive) or "low" (saturated negative) output state. Differential inputs

also allow amplifiers to be used as *operational amplifiers* (also called "opamps"), so named because these useful circuits are able to perform a variety of mathematical operations on analog signals.

3.3 Integrated circuit amplifiers

Electronic amplifier circuits may be constructed out of individual ("discrete") components, but for low-power applications it makes much more sense to manufacture amplifiers in the form of *integrated circuits* (ICs) with all internal components on a single wafer of silicon semiconductor material. A micro-photograph of a model LM741 operational amplifier manufactured by National Semiconductor appears below, this particular photo offered graciously by the creators of the website *zeptobars* (`https://zeptobars.com`) under the Creative Commons Attribution 3.0 Unported License:

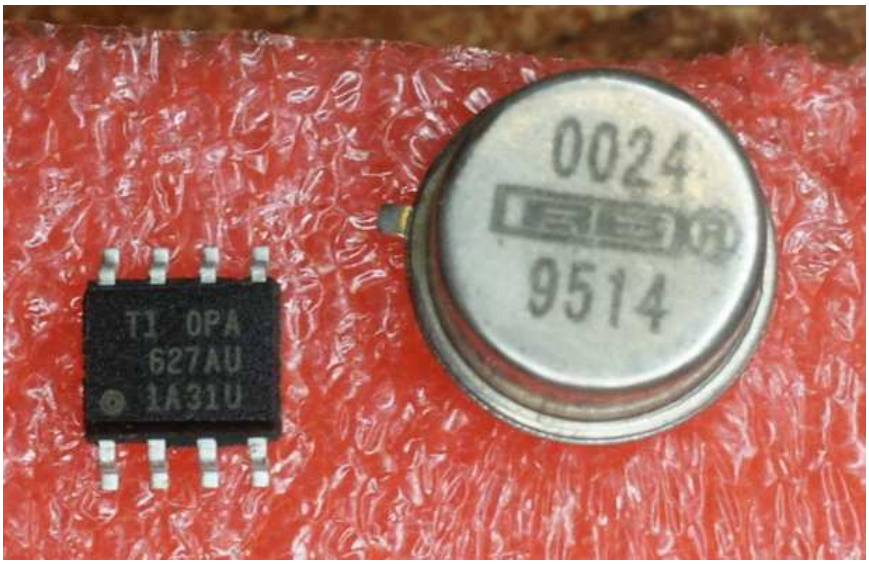
National LM741CH opamp, die size 1054 μm by 827 μm
Image source: `s.zeptobars.com/National-LM741CH-HD.jpg`



The black objects in this photograph are the metal *bond wires* connecting pads on the silicon die to metal pins on the IC package.

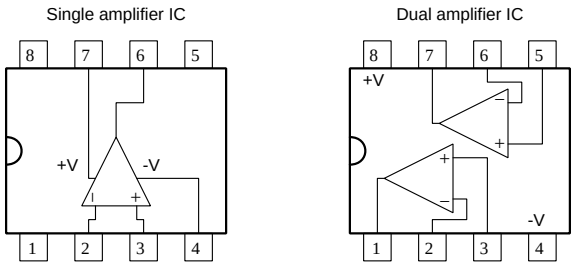
Some IC amplifiers are enclosed in plastic or ceramic packages, others in metal "can" packages. Examples of each package type are seen in the following photographs, also courtesy of *zeptobars*:

Image source: `zeptobars.com/en/read/OPA627-AD744-real-vs-fake-china-ebay/`



Here, we take a look inside the metal "can" package to see the silicon die and its bond wires connecting to the package's eight pins:

Due to the extreme compactness of ICs, it is common to find multiple amplifiers contained within a single IC package. Examples of single and dual differential amplifiers contained within 8-pin *DIP* (Dual Inline Package) integrated circuits appear in the following illustration:

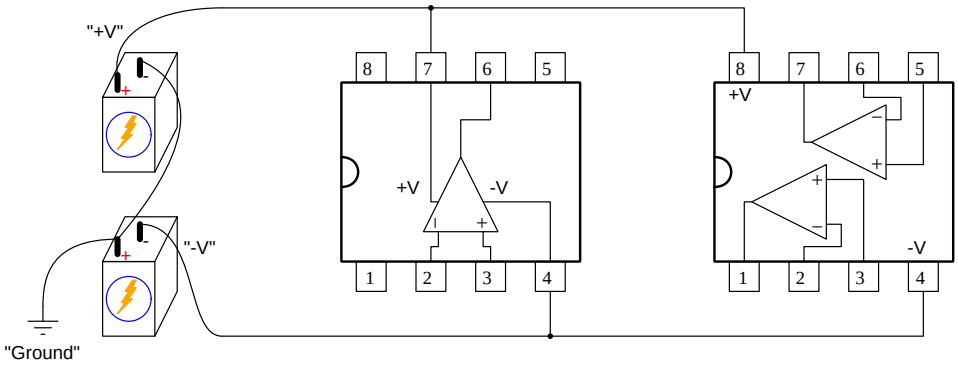


Quad amplifier ICs are also available, requiring more than eight pins of course.

3.4 Amplifier split power supplies

Many electronic amplifier circuits require a "split" DC power supply with a +V terminal, a −V terminal, and lastly a "ground" terminal serving as a reference point for voltage measurements. For experimental purposes you may create your own split DC power supply by connecting two batteries in series-aiding fashion, referring to the mid-point between them as "ground":

A "split" power supply made of two series-connected batteries

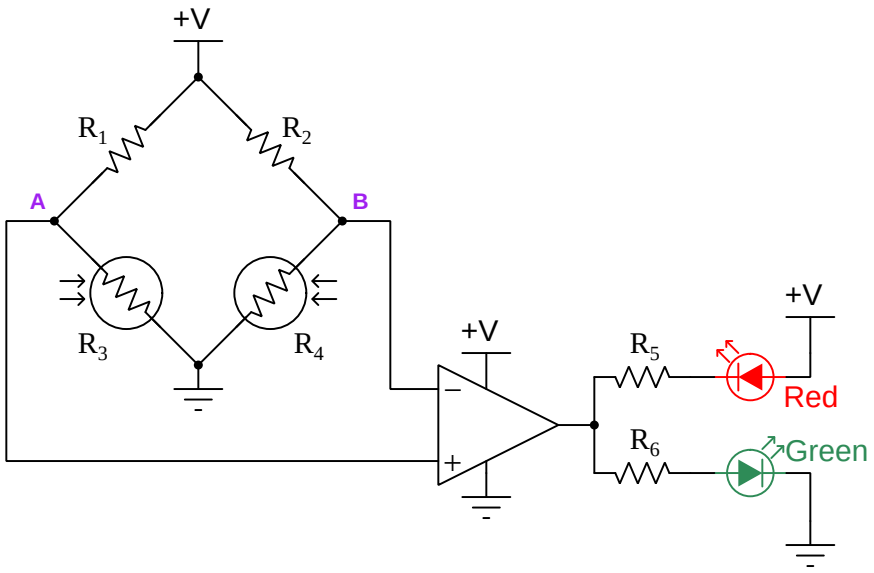


When this "ground" point is used as a reference for voltage measurements (i.e. connecting the black test lead of a voltmeter to that point) and the operational amplifier ICs are powered by these +V and −V "rails", it becomes possible to build circuits where signals may vary in magnitude from −V to +V and anywhere in between. This is most useful in circuits where AC voltage signals must be measured and/or output by the operational amplifier(s).

3.5 Signal comparison

If a differential amplifier is built to have a very large voltage gain, its output will saturate at or near one of the "rail" voltage limits even for small amounts of differential voltage between its input terminals. At first this may seem pointless, to build an amplifier such that it completely saturates its output in response to the slightest non-zero input, but this turns out to be incredibly useful. The first practical use we will explore for this ultra-high-gain behavior is *signal comparison*.

The following bridge circuit is an application for a differential amplifier as a signal comparator. The bridge network's balance is set by the relative light intensities detected by the two photocells, such[1] that the bridge is considered "balanced" when both photocells are exposed to the same amount of light and "imbalanced" when one photocell is more exposed than the other:



Photocells generally become more conductive when exposed to light. If photocell R_3 is exposed to more light than photocell R_4, test point A will sink to a lower potential than test point B. This will make the amplifier's inverting input more positive than its noninverting input, causing its output to saturate low (i.e. at or near ground potential), thus illuminating the red LED. If photocell R_4 is exposed to more light than photocell R_3, the reverse will be true and the green LED will illuminate. If the amplifier has a sufficiently high voltage gain, the slightest difference in light exposures will cause the LED states to switch. The amplifier's output signal will essentially be *discrete* (either "high" or "low", +V or Ground potential).

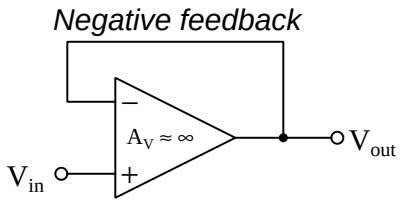
When used in such a manner, a differential amplifier is said to be in *open-loop* mode, functioning as a *comparator*. The output's state is purely a function of the differential voltage input and its own internal voltage gain (A_V), and with a sufficiently high voltage gain the amplifier essentially acts as a discrete switching device, saturating its output signal with any amount of applied differential voltage between the input terminals.

[1] If the two photocells are identical and $R_1 = R_2$ then the bridge circuit will be balanced with equal light exposure.

3.6 Negative feedback

In the late 1920's a Bell Labs researcher named Harold Black made a rather counter-intuitive discovery: by taking an electronic amplifier with very high gain and coupling part of its output signal back to its input in a degenerative manner (i.e. *negative feedback*) it became possible to achieve extremely precise values of gain and highly stable operating characteristics despite imperfections in the amplifier's internal circuitry. Black's US patent (filed in April of 1932, awarded December of 1937 – number 2,102,671) explains this discovery in fine detail, and his employer – the Bell Telephone Company – profitably used this design technique to improve the performance of "repeater" amplifiers used to boost telephone signals after suffering losses over many kilometers of cable. Although electronic amplifier technology has changed quite a bit since the 1920's, we may exploit this very same principle in a modern differential amplifier by building its internal circuitry to have an extremely high voltage gain, and then coupling its output terminal somehow to its *inverting* (−) input terminal for negative feedback.

The simplest implementation of negative feedback for a differential amplifier consists of directly connecting the output terminal to the inverting (−) input terminal as shown in the following diagram:



Note how the DC power supply connections have been omitted from the diagram for the sake of simplicity. The circuit's input now consists of a single terminal (connected in the noninverting, or + terminal of the amplifier) and the output also consists of a single terminal, both signals being ground-referenced voltages. The amplifier's voltage gain has been made extremely high, as indicated by $A_V \approx \infty$ (i.e. approximately equal to infinity).

As we saw with the photocell-bridge comparator circuit, if the inverting input is at a greater potential than the noninverting input, the amplifier's output signal moves aggressively toward its negative power supply rail limit. If the noninverting input is at the greater potential, the opposite occurs and the output signal swings toward the positive rail. Connecting the output terminal directly to the inverting input terminal, therefore, causes the amplifier to constantly *self-correct*: if the output signal is too high or too low, the amplifier forces itself to go the other way. What defines "too high" or "too low" is the other input's potential (the noninverting terminal, or V_{in}), since this is a *differential* amplifier and the output depends on the *difference* of potential between the + and − input terminals.

A common misconception is that aggressive negative feedback necessarily results in the amplifier breaking into oscillations, repeatedly *over*-correcting. While this can happen, it requires time lags in the feedback loop where the amplifier's output swings one way to correct, and then it senses *too late* that it has swung too far. If the feedback is immediate, what happens is that the amplifier quickly "finds" a point of equilibrium where its output signal stabilizes to a value just different enough from V_{in} to produce itself.

Finding that equilibrium value seems daunting at first inspection because it appears to be a logical paradox: how can we calculate V_{out} when its value *depends on itself* by way of feedback?

Applying some mathematical analysis will be helpful at this point. We have already seen how gain was defined as the relationship between an amplifier's input and output signals ($V_{out} = A_V V_{in}$), and in the case of a differential amplifier V_{in} is really the difference between the two input terminals' voltages measured with respect to ground; i.e. $V_{out} = A_V(V_{(+)} - V_{(-)})$. In a circuit where we directly connect the output terminal to the inverting input terminal to achieve negative feedback, $V_{(-)}$ is the same as V_{out}, and if we rename the noninverting input terminal's signal as V_{in} this formula condenses to $V_{out} = A_V(V_{in} - V_{out})$. Algebraically solving for V_{out} as a function of V_{in}:

$$V_{out} = A_V(V_{in} - V_{out})$$

$$V_{out} = A_V V_{in} - A_V V_{out}$$

$$A_V V_{out} + V_{out} = A_V V_{in}$$

$$V_{out}(A_V + 1) = A_V V_{in}$$

$$V_{out} = \frac{A_V}{A_V + 1} V_{in}$$

The ratio $\frac{A_V}{A_V+1}$ now becomes the effective voltage gain for the negative feedback circuit as a whole. If the amplifier's open-loop voltage gain is very large, then this ratio will be very nearly equal to one, making $V_{out} \approx V_{in}$. Using an assumed open-loop amplifier voltage gain (A_V) of one million, we may apply this equation to a table of values, showing how this negative-feedback amplifier circuit responds to a range of V_{in} signal values from -3 Volts to $+3$ Volts:

V_{in}	A_V	$\frac{A_V}{A_V+1}$	V_{out}
-3.0 V	1×10^6	0.999999	-2.999997 V
-2.0 V	1×10^6	0.999999	-1.999998 V
-1.0 V	1×10^6	0.999999	-0.999999 V
0.0 V	1×10^6	0.999999	0.0 V
$+1.0$ V	1×10^6	0.999999	$+0.999999$ V
$+2.0$ V	1×10^6	0.999999	$+1.999998$ V
$+3.0$ V	1×10^6	0.999999	$+2.999997$ V

As you can see, V_{out} comes very close to exactly equaling V_{in}, with the error magnitude being approximately -1 μV per Volt of input signal (i.e. error $\approx -\frac{V_{in}}{A_V}$). V_{out} cannot *exactly* equal V_{in} except at 0 Volts because then the amplifier would have no differential signal to amplify. Instead, V_{out} always settles at a value just far enough away from V_{in} where the difference multiplied by the extremely high gain is enough to generate that same V_{out}. For example, with V_{in} at a value of $+2$

Volts, V_{out} settles at about 2 μV shy of +2 Volts, so that the differential voltage of approximately +2 μV when multiplied by the A_V of one million will reproduce a signal that is approximately +2 Volts. For this reason, a circuit like this is commonly referred to as a *voltage follower*, since V_{out} very closely follows V_{in} for all voltage values within the limits of the amplifier circuit and its DC power supply.

What is more, this extremely close following of V_{out} to V_{in} created by negative feedback remains virtually unaffected by even large changes in the amplifier's internal voltage gain value. If, for example, we imagine the operational amplifier's A_V degrading from a value of one million to a mere one hundred thousand (i.e. a ten-fold reduction in gain), the ratio $\frac{A_V}{A_V+1}$ hardly changes at all: it goes from being 0.999999 to being 0.99999, still almost exactly equal to one. This was one of Harold Black's insights in the early 1930's: that adding negative feedback to an amplifier circuit causes it to become far more stable despite inevitable changes in the behavior of its active components, which in that day were vacuum tubes. Even today where vacuum tubes have long been replaced by solid-state transistors, the stability granted by negative feedback still has great value and benefit to circuit designs, allowing us to design and build highly precise amplifier circuits tolerant of variations in performance from one transistor to another as well as changes in environmental conditions such as ambient temperature, supply voltage variations, etc. that also affect transistor operation. In other words, combining a high-gain amplifier with negative feedback permits us to create a single circuit design that is exceptionally reliable and repeatable from unit to unit as well as for one unit over a wide range of operating conditions.

We must be careful when speaking of "gain" in such a circuit because there are really two different voltage gain values at play here. First is the differential amplifier's own internal gain which we have been careful to refer to as its *open-loop* gain (i.e. its voltage gain without feedback). Second is the circuit's "over-all" gain which we will now refer to as the *closed-loop* gain (i.e. the circuit's voltage gain with negative feedback in effect). The "loop" referred to here, of course, is negative feedback. Since the output gets fed back to the input, and this in turn affects the output, this circuit's causality is circular in form: that is to say, the information continually "loops" around.

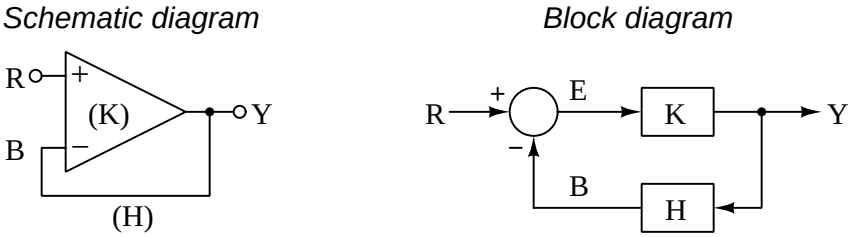
In conclusion, the effect of taking a differential amplifier with an extremely high open-loop voltage gain and connecting its output to its own inverting input for negative feedback is that the amplifier continually self-corrects to maintain the two input terminals at approximately the same electrical potential (i.e. $V_{diff} \approx 0$). This is the cardinal rule of negative feedback:

Cardinal Rule of Negative Feedback

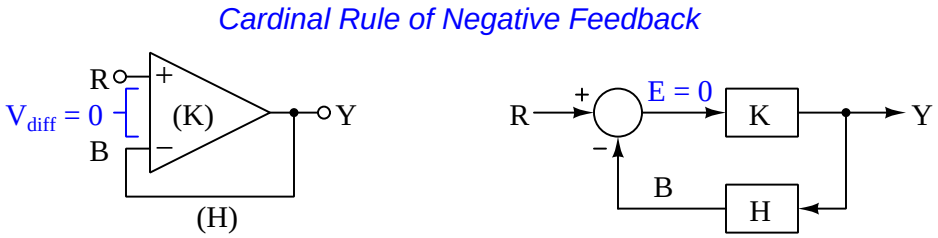
> *Given sufficiently high open-loop gain, the amplifier's output signal swings as far as it must within the limits of the power supply "rail" voltage to eliminate any "error" between its two input terminals.*

Any differential amplifier built with extremely high gain for the purpose of constructing a negative feedback system is called an *operational amplifier* or *opamp*, the name referencing a wide range of mathematical operations that may be performed in analog fashion with such a device. Subsequent sections of this tutorial will elaborate on how those functions are performed, and how the essential characteristic of self-correction and elimination of "error" is central to the performance of those functions.

A common engineering notation for feedback systems is the *block diagram*, representing signals as arrows and functions as rectangular "blocks". Block diagrams are useful for describing a great number of complex systems, but they work particularly well to describe operational amplifier circuits, as illustrated below. The letters R, B, E, K, H, and Y are fairly standard for block diagrams, representing *reference* signal, *back*-fed signal, *error*, *compensator* function, *feedback* function, and *output*, respectively:



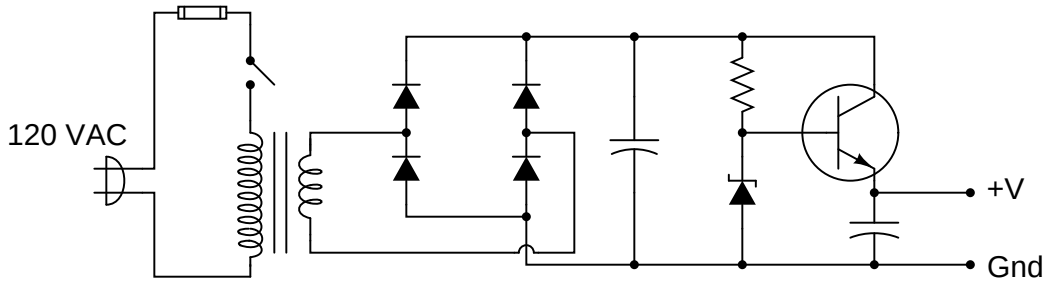
The high open-loop gain of the amplifier (K) works to minimize the difference (*error*, E) between the input signal (*reference*, R) and the back-fed signal (B). In an ideal system, E will be maintained at zero, making $B = R$. This, again, is our cardinal rule of negative feedback:



When the negative feedback is direct as in the case of a voltage follower, $Y = B = R$. A voltage follower, in fact, could be represented in block diagram form without a feedback block (H) at all, just a signal path straight from the output Y to the inverting input of the summation node (the circle). However, as we will see in subsequent sections, we can make operational amplifiers do many useful things besides replicating voltage signals 1:1 if we make the feedback function more interesting than a mere length of wire.

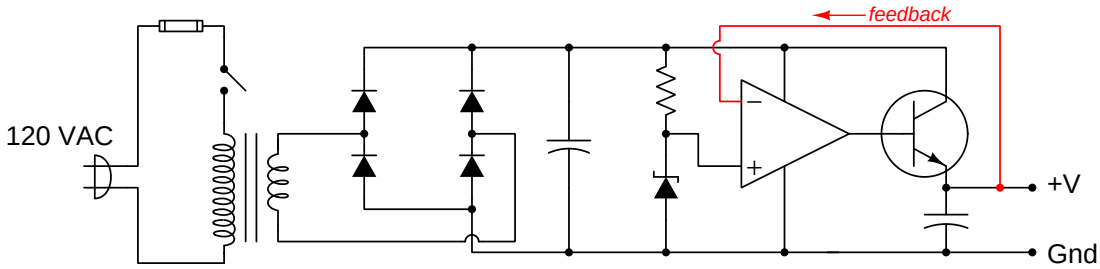
3.7 Regulation

The fact that an operational amplifier wired for negative feedback will act in such a way to self-correct and eliminate any detected error between two voltage signals makes it well-suited for voltage regulation: ensuring a steady voltage within a circuit under varying load and/or source conditions. Consider the following AC-DC power supply circuit with a simple *series* voltage regulation scheme:



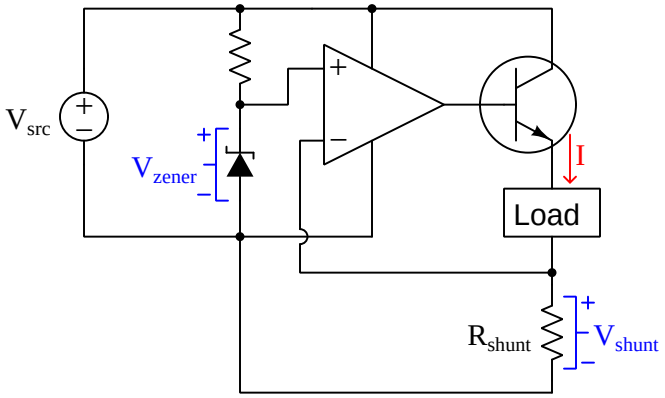
120 Volt AC line power is stepped down to a lower voltage by the transformer, rectified into DC by the bridge rectifier (four diodes), and filtered by the first capacitor. A resistor and Zener diode work to provide a stable reference voltage (dropped across the Zener), and this reference voltage signal drives the base of an NPN transistor which sends current through to the output terminals where a load may be connected. Unfortunately, though, the voltage at the output terminals will not be the same as the Zener diode's reference voltage due to the approximately 0.7 Volts of drop across the transistor's base-emitter PN junction. For example, using a 5.0 Volt Zener diode will result in an output voltage of approximately 4.3 Volts, and this will vary as the transistor warms and cools.

Installing an operational amplifier between the Zener diode and the transistor, with the inverting input connected to the output terminal so as to monitor the power supply's output voltage at all times, fixes this problem:



If the output voltage sags for any reason, the opamp outputs a stronger signal to the transistor's base to raise it back to its target value (set by the Zener diode's voltage drop). If the output voltage rises too high for any reason, the opamp drives a weaker signal to the transistor to correct. If the transistor's base-emitter junction voltage varies at all, the opamp automatically compensates. Under all operating conditions, the opamp varies its output as necessary to maintain the power supply's output terminal voltage equal to the Zener diode's voltage.

Opamps are also useful as *current sources*[2]. Consider the following circuit, also using a Zener diode as a voltage reference, designed to drive a constant current through a variable load:



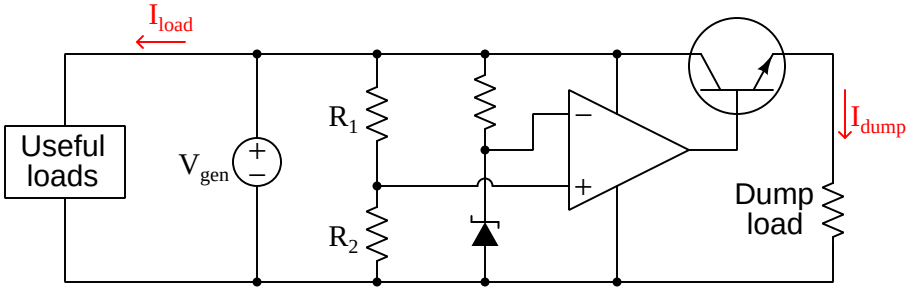
As in all negative feedback systems, the operational amplifier works to minimize error between its two input terminals. In this circuit the two input signals are V_{Zener} and V_{shunt}, with the opamp driving the transistor as hard as necessary to make these two voltage signals equal to each other.

The purpose of R_{shunt} is to drop the same voltage as the Zener diode at the desired amount of load current. For example, if we wished to regulate the load current's value at 20 mA, we might choose a 5.0 Volt Zener diode and a 250 Ω shunt, because 20 mA passing through 250 Ω should drop 5 V (the same as the Zener diode). If for any reason the load current exceeds the target value, the shunt resistor's voltage drop will rise, causing the opamp to decrease its output to re-establish a condition of equipotentiality between its two input terminals. If for any reason load current falls below target, the opamp will sense this deficiency and drive the BJT harder. As always, the combination of high open-loop gain and negative feedback causes the opamp to "fight" as hard as it must to maintain negligible differential input voltage.

This current-regulating circuit can only function well if we assume negligible current passes through the wire connecting the opamp's inverting input terminal to the top of the shunt resistor, because if substantial current ever did pass through that wire it would mean load current would not be equal to shunt resistor current (by Kirchhoff's Current Law). Differential voltage amplifiers in general exhibit very high input resistance (typically in the megaOhm range), and so this is a safe assumption. We will find this assumption very useful in the analysis of other feedback circuits. Applied in conjunction with our cardinal rule that negative feedback works to maintain zero "error" between the inverting and noninverting input terminals, we will find the analysis of opamp circuits a relatively straightforward application of Ohm's and Kirchhoff's Laws.

[2]Legacy analog literature often referred to these circuits as *current pumps*.

Another form of regulatory circuit is a *shunt regulator*, commonly used in certain renewable energy generation systems for the purpose of regulating line voltage by directing excess energy toward one or more *dump loads*[3]. Small-scale wind turbine generators benefit from this form of voltage regulation because during periods when generation exceeds demand (e.g. prolonged high-wind conditions) the turbine's speed will be more stable if placed under sufficient load. A simplified diagram of a shunt regulator on such a system appears here:



As usual, a resistor and Zener diode provides a stable reference voltage for the opamp to compare against the regulated voltage (V_{gen}). A voltage divider consisting of R_1 and R_2 provides the opamp with a scaled proportion[4] of the generator's voltage for feedback. If the generator's voltage rises too high, the opamp drives the transistor into conduction and passes current to the dump load to dissipate energy and thereby reduce the generator's terminal voltage.

A very interesting feature of this circuit becomes evident upon close inspection: the feedback signal is sensed by the operational amplifier via its *noninverting* (+) input terminal, yet this is still considered a *negative feedback* circuit. Compare this circuit with the previous regulatory circuits and you will see the difference for yourself: in the other circuits the Zener diode reference signal connects to the + terminal and the feedback signal to the − terminal, but here it is exactly opposite. How does this still result in *negative* feedback if the back-fed signal doesn't connect to the "negative" input terminal on the opamp? The answer is based on the fact that we are regulating the voltage of a *source* and not of a *load*. In previous regulator designs, activating the transistor resulted in a load voltage *increase*, but here the dump load's voltage is not what is being regulated, rather it is the generator we seek to regulate by loading it's output voltage down. In the other regulator circuits the appropriate opamp response to an excessive (load) voltage was to drive the transistor less; here the appropriate response to an excessive (source) voltage is to drive the transistor more.

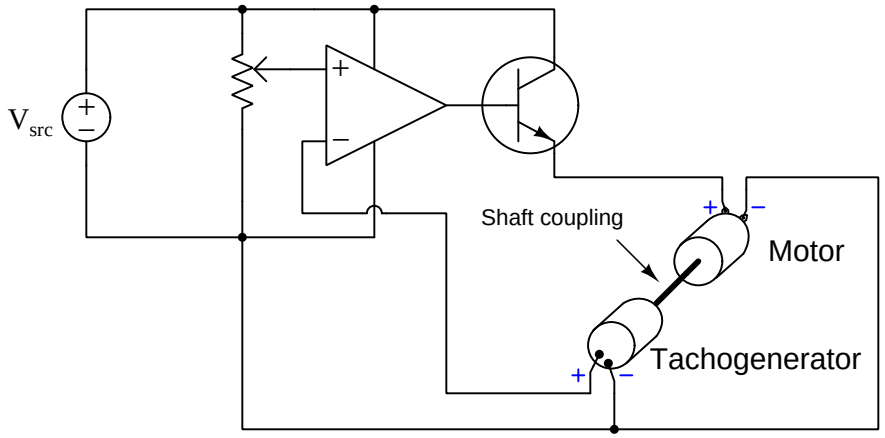
[3] This colorful term refers to an electrical load where excess energy is safe to "dump". The dump load may serve some practical purpose other than merely shedding excess energy, or it may be as crude as a resistive heating element exposed to outside air.

[4] A natural question at this point might be to ask "Why use a voltage divider, instead of just connecting the opamp directly to the generator's positive terminal to sense its full voltage for feedback?" The most important reason for this is related to the Zener diode reference network: in order for the resistor-diode network to provide a stable reference voltage, that reference voltage must be considerably less than the full "rail" voltage of the generator. If $V_{Zener} < V_{gen}$ by necessity, then we obviously cannot use V_{Zener} as a target value for V_{gen} to achieve. Instead, we must monitor a *fraction* of the generator's voltage for feedback and comparison against the Zener reference. For example, if the desired target voltage for the generator was 24 Volts and the Zener diode's reference voltage was 6 Volts, we would select R_1 and R_2 to have a voltage-division ratio of 4:1.

3.8 Servos and control systems

We have seen how operational amplifiers may serve to regulate voltage or current by comparing a reference signal against a measured signal and making corrections by driving a transistor with its output. The combination of high open-loop gain and negative feedback causes the opamp to function as an automatic control system, working tirelessly to maintain the regulated voltage or current at its target value. It stands to reason, then, that operational amplifiers may likewise regulate *any physical variable* representable as a voltage and controllable using electricity.

Consider the following *motor speed* control system as an example:

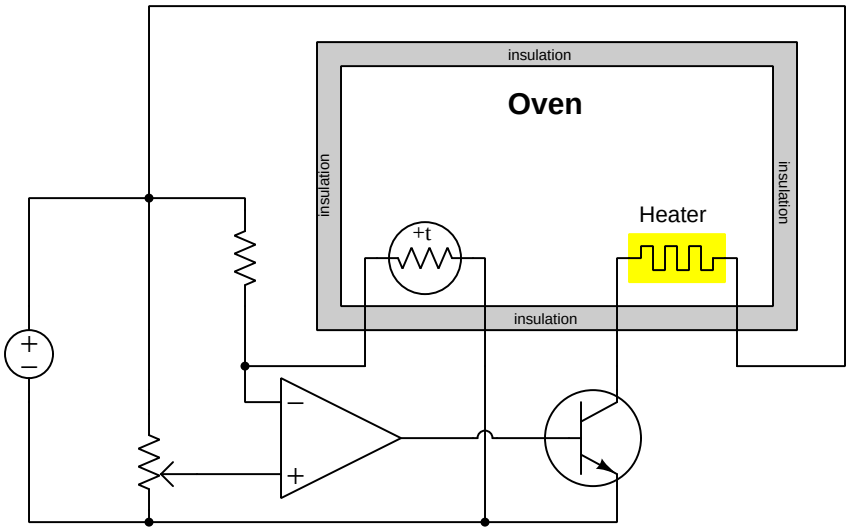


A DC motor receives power through the transistor, which in turn is driven by the opamp. A *tachogenerator* generates a DC voltage in direct proportion to its shaft speed, so that when the two are mechanically coupled the opamp receives continuous feedback on the motor's actual speed. A potentiometer connected as a voltage divider provides a "reference" or "target" speed signal to the opamp's noninverting input. From the opamp's perspective it "thinks" it is simply regulating voltage when in fact it is regulating motor speed. Voltage signals are used in this circuit as proportional representations of speed, and for this reason it is an *analog* circuit in the fullest sense[5].

If we were to replace the tachogenerator with a position sensor outputting a DC voltage, and replace the single power transistor with a transistor array capable of reversing polarity to the motor, we would have a full *servo* system where *position* and not speed would be the controlled variable. In such a system, the potentiometer would define the desired position of the mechanism, and the opamp would ceaselessly work to maintain the sensed position at that target or reference value. Servo systems are the basis of *robotics*, where a mechanism precisely follows position commands.

[5] Broadly defined, an "analog" system is one where the internal variables are continuously adjustable as opposed to a *digital* system where the variables are discrete. More specifically defined, a true analog system is one where the continuously-variable signals proportionately represent something else, serving as *analogues* of other quantities.

However, we are not limited to controlling just speed or position using opamps. Literally *any* physical variable representable as a voltage signal and controllable by some electrical element may be regulated in this manner. Consider the following oven temperature control system, which is very much like the servo system just described, just controlling temperature instead of mechanical position or speed:

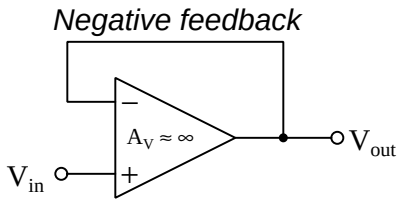


Here a *thermistor* with a positive temperature coefficient (i.e. increases resistance with increasing temperature) senses the oven's temperature, dropping a voltage representative of that temperature sensed at the opamp's inverting input. A potentiometer provides an adjustable reference voltage signal for the opamp at its noninverting input. A transistor serves to throttle power to the oven's heating element at the command of the opamp's output signal. If oven temperature deviates at all from the reference value, the opamp automatically corrects by re-adjusting power to the heating element: too hot results in less current to the heater; too cool results in more power sent to the heater.

It should be noted that these examples of automatic control systems based on operational amplifiers are all simplified. No thought has been given to noise filtering, power requirements, component ratings, precise calibration, or any number of practical considerations. The point here is to grasp the basic concept and wide applicability of negative feedback as a paradigm for automated control.

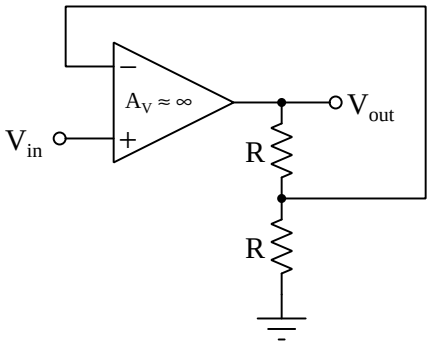
3.9 Attenuated and offset feedback

Most introductions to operational amplifiers with negative feedback begin with the humble voltage follower circuit, shown here[6] for reference:



The cardinal rule of negative feedback when implementing using an amplifier of suitably high gain is that the "error" signal (i.e. the differential voltage between the opamp's two input terminals) will be maintained at virtually zero. With this rule in hand, we may conclude that $V_{out} = V_{in}$ (or at least *close enough* for all practical purposes) for the voltage follower circuit. What happens, though, when we substitute a more complicated feedback network in place of the simple wire?

Consider the following modification, placing a 2:1 voltage divider network between the opamp's output and inverting input terminal:

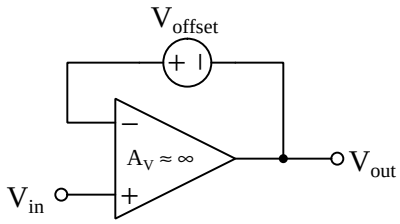


No longer is the opamp comparing V_{in} directly against V_{out}; now it is comparing V_{in} against *one-half* of V_{out}. Just a moment's thought should be necessary to conclude that V_{out} must rise to *twice* the value of V_{in} in order to eliminate error. Thus, by using a voltage *divider* network in the feedback signal path, we have prompted a closed-loop *multiplication* of voltage.

Practically any division ratio will work in this way: a 3:1 divider causes the closed-loop gain to be three; a 10:1 divider yields a closed-loop gain of ten. In all cases, the principle is the same: *attenuating* the feedback signal causes the circuit to *amplify* the input by that very same ratio.

[6] Note that as in many opamp schematic diagrams, the DC power supply terminal connections to the amplifier have been omitted for simplicity, so that the diagram focuses mostly or even exclusively on signal pathways. This is also commonly done in digital logic gate schematic diagrams, and for the exact same reason.

We see a similar phenomenon at work if we insert a voltage source within the feedback pathway to *offset* the signal:

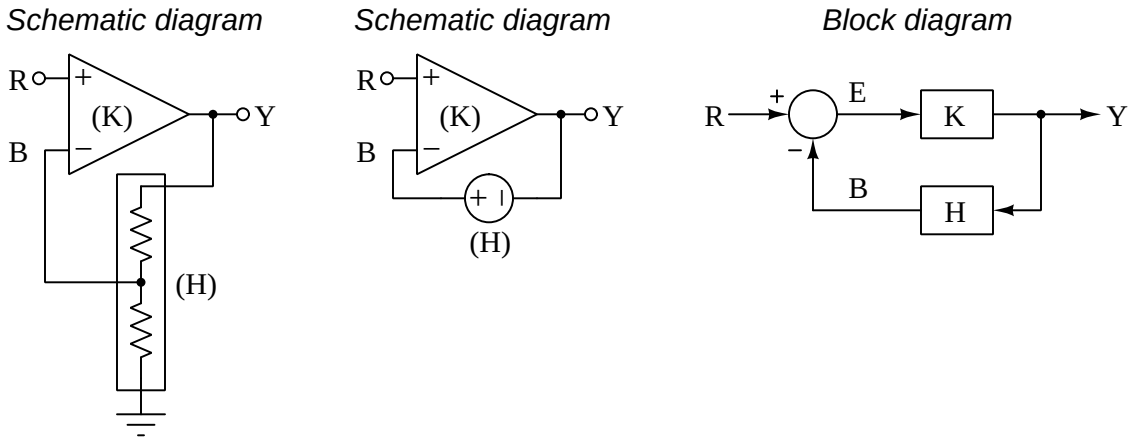


No longer is the opamp comparing V_{in} directly against V_{out}; now it is comparing V_{in} against $V_{out} + V_{offset}$. Just a moment's thought should be necessary to conclude that V_{out} must decrease to V_{offset} less than the value of V_{in} in order to eliminate error. Thus, by *adding* an offset voltage to the feedback signal, we have prompted a closed-loop *subtraction* of voltage.

Suppose the value of V_{offset} happened to be 2 Volts while V_{in} was +5 Volts (with reference to ground). As usual, the opamp drives its output as far as necessary to eliminate error (i.e. zero differential voltage between inverting and noninverting input terminals) through negative feedback and high internal gain. If we apply Kirchhoff's Voltage Law from ground to the output terminal past the offset source to the inverting terminal and back to ground, we see that the presence of the series-connected V_{offset} causes the inverting input to sense a voltage +2 Volts greater than V_{out}. This must mean that V_{out} will settle to a value *2 Volts less* than V_{in}. In this particular case, V_{out} = +3 Volts.

Take some time to note the general phenomenon seen in both the divided and offset feedback voltage signals: whatever mathematical function is placed in the feedback signal path becomes *inverted* to describe the over-all circuit's closed-loop behavior. This *functional inversion* offered by negative-feedback systems is an extremely useful property in electronics, and you will see it applied again and again in your future studies.

As mentioned in a previous section, *block diagrams* work well to describe operational amplifier circuits in abstract terms. This section explored the effects of attenuating or offsetting the output signal as it gets fed back negatively to the amplifier, and we can model these signal alterations in a block diagram by the function H:



The back-fed signal (B) is the result of the output signal (Y) passing through function H. Using mathematical *function notation* to show this relationship, $B = H(Y)$. If the amplifier is successful in eliminating error by making B equal to the input "reference" signal R, we can write $R = H(Y)$.

If R is the result of passing Y through function H, then Y may be thought of as the result of passing R *backward* through H. This does not always make literal sense in the schematic diagram[7], but it makes sense as a mathematical function. The way to write this using function notation is $Y = H^{-1}(R)$, and we would say that the *closed-loop function* of this system is the *inverse* of its feedback function.

Again, this is a general property of negative feedback: *whatever mathematical function we place in the feedback signal path becomes inverted from input to output.* If we were to somehow make the feedback network a square-root function, Y would become the *square* of R. If we were to somehow make the feedback network a power function, Y would become a *root* of R. If we were to somehow make the feedback network an exponential function, Y would become the logarithm of R. This is an incredibly useful phenomenon because it essentially doubles the number of mathematical functions we are able to perform using analog networks[8].

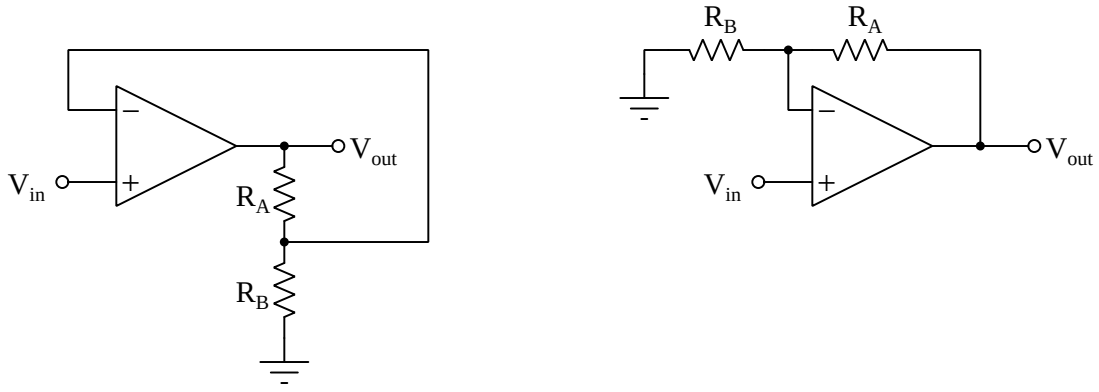
[7] For example, trying to pass a voltage signal backward through a voltage divider network certainly will *not* amplify that signal. However, this does actually work for the voltage-offset function.

[8] For example, we know that resistor networks cannot amplify signals, only attenuate. However, when included in a negative feedback system, now the *system* is able to amplify in inverse proportion to the resistor network's ability to attenuate. Negative feedback therefore extends the range of analog possibilities.

3.10 Inverting and noninverting amplification

As we have seen in the previous section, placing a voltage-dividing network in the feedback signal path causes the amplifier to exhibit a closed-loop voltage gain equal to that divider network's division ratio. This circuit configuration may be represented in slightly different ways – on the left we see an obvious voltage divider network consisting of resistors R_A and R_B feeding a divided V_{out} signal to the inverting input of the operational amplifier, and on the right we see the exact same circuit with the voltage divider drawn horizontally above the amplifier[9]. In this circuit the action of the opamp's high open-loop gain combined with negative feedback forces the divided signal to be equal to V_{in}:

Noninverting amplifier configuration



The division ratio of these two resistors follows the well-known voltage divider formula:

$$\frac{V_{in}}{V_{out}} = \frac{R_B}{R_B + R_A}$$

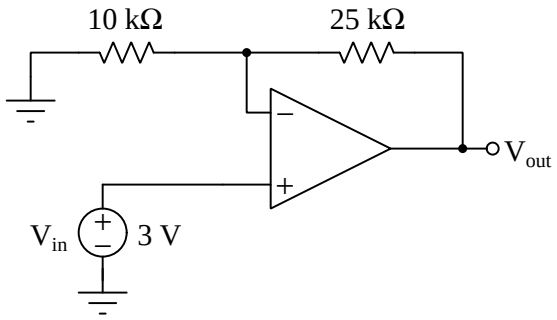
Since we always define *gain* for any amplifier as the ratio of output to input, the closed-loop voltage gain of this configuration is simply the reciprocal of the previous formula:

$$A_V \text{ (closed-loop)} = \frac{V_{out}}{V_{in}} = \frac{R_B + R_A}{R_B} = 1 + \frac{R_A}{R_B}$$

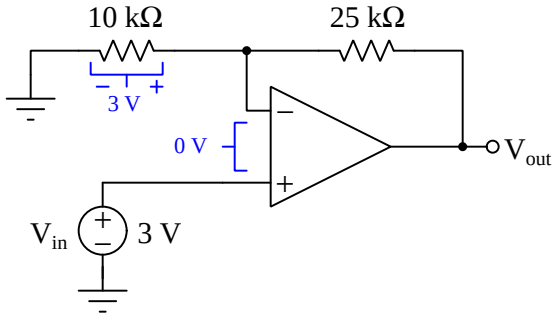
We refer to this as a *noninverting* amplifier because the output voltage has the same mathematical sign as the input voltage. For example, if R_A is twice the value of R_B and $V_{in} = +4$ Volts, then $V_{out} = +12$ Volts.

[9] Note that the necessary DC power supply connections for powering each of these operational amplifiers has been omitted from the schematic diagram for simplicity's sake. This permits us to focus on the signal pathways rather than the power which we already know must be present for any opamp to function.

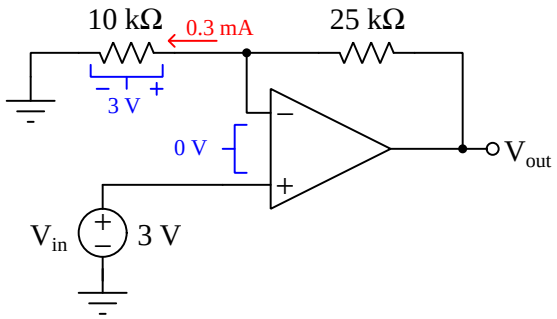
Gain formulae summarizing the behavior of an amplifier network are convenient, but it is instructive to analyze such a circuit with specific component values and signal voltages in order to understand how foundational principles such as Ohm's Law and Kirchhoff's Laws apply. Let us analyze this noninverting circuit given the parameters shown in the following diagram:



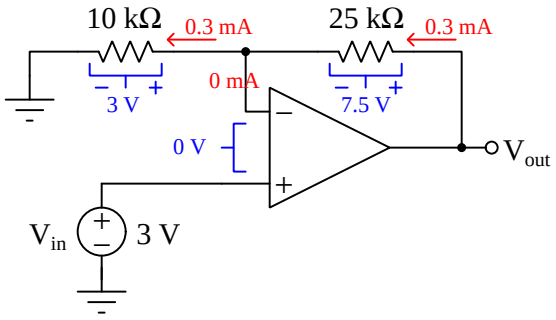
Recognizing the presence of negative feedback in this circuit, we may begin our analysis by assuming the opamp will output whatever voltage necessary in order to maintain zero "error" (i.e. zero differential voltage between its + and − input terminals). By Kirchhoff's Voltage Law, this means the voltage drop across the 10 kΩ resistor must be 3 Volts:



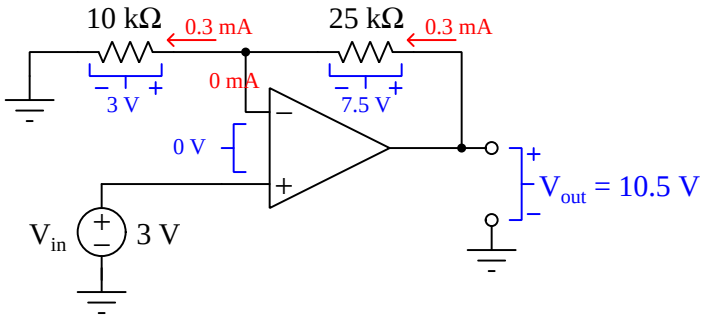
Applying Ohm's Law to the 10 kΩ resistor, we may conclude the current through it is 0.3 mA, and since resistors are always electrical *loads*, this current must be entering the + terminal and exiting the − terminal:



This 0.3 mA current must come from somewhere, and since we know the input terminals of the opamp itself pass negligible current we may apply Kirchhoff's Current Law to the node between the two resistors to conclude the entirety of that 0.3 mA must also pass through the 25 kΩ resistor. That current, as it passes through the 25 kΩ resistor, causes it to drop 7.5 Volts (by Ohm's Law), the polarity based on the fact that the resistor is a load:



V_{out} is ground-referenced, which means its value is the amount of voltage between the output terminal and ground. Kirchhoff's Voltage Law is helpful again, as we can see this will be the sum of the resistor voltage drops (since those two resistors together span the output terminal and ground:



This result agrees with our general gain formula for a noninverting configuration, where:

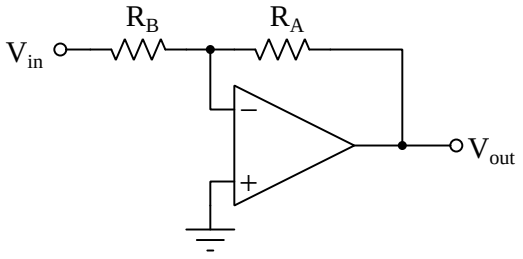
$$V_{out} = A_V V_{in}$$

$$V_{out} = \left(1 + \frac{25\ \text{k}\Omega}{10\ \text{k}\Omega}\right)(3\ \text{V})$$

$$V_{out} = 10.5\ \text{V}$$

An alternative configuration to make an operational amplifier exhibit a modest voltage gain is the so-called *inverting* circuit, which looks remarkably similar to the noninverting circuit except for the positions of the ground and V_{in} terminals:

Inverting amplifier configuration



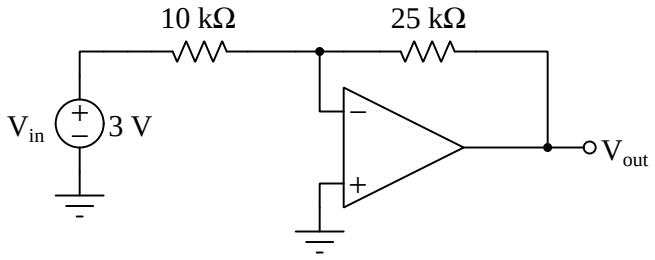
Note how the noninverting input connects directly to ground, while V_{in} connects to the left-hand side of R_B. As always, the action of negative feedback automatically works to eliminate error, and because the noninverting terminal is grounded the inverting terminal must be held at ground potential by the opamp's continual self-correcting action. For this reason the node between the two resistors and the inverting terminal is typically called a *virtual ground* because it is equipotential with ground yet not actually connected to ground.

V_{in}, being a ground-referenced voltage signal, is therefore impressed across R_B. The current $\frac{V_{in}}{R_B}$ passes through R_A as well, since it cannot enter or exit the opamp's input terminal. This results in a voltage drop across R_A in accordance with Ohm's Law, and this voltage drop must be equal to V_{out} since V_{out} is a ground-referenced signal and the left-hand side of R_A connects to virtual ground. If $R_A = R_B$ then V_{out} will have the same absolute value as V_{in}, i.e. a condition of no voltage amplification at all. If $R_A < R_B$ then V_{out} must be less than V_{in}; if $R_A > R_B$ then we will have true voltage amplification with the magnitude of V_{out} exceeding that of V_{in}. Formulating this mathematically:

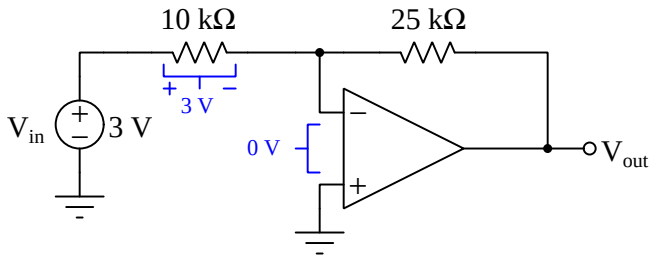
$$A_V \text{ (closed-loop)} = \frac{V_{out}}{V_{in}} = -\frac{R_A}{R_B}$$

The reason why closed-loop voltage gain is the *negative* ratio of R_A to R_B may be understood by running a simple thought experiment: imagining a positive value for V_{in} in the previous schematic diagram. This makes R_B's voltage drop + on left and − on right and the resulting current direction left-to-right through both resistors. Therefore, R_A's voltage drop must also be + on left and − on right, making V_{out} a negative quantity with reference to ground. This negative gain is the reason we call this circuit configuration an *inverting amplifier*.

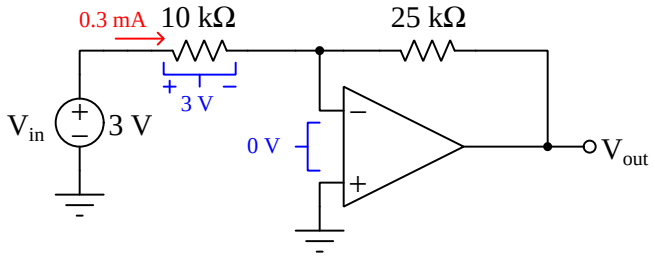
As with the noninverting amplifier configuration, it is instructive to analyze a circuit with given signal strengths and component values, noting where fundamental laws apply. We will begin with the following inverting amplifier circuit:



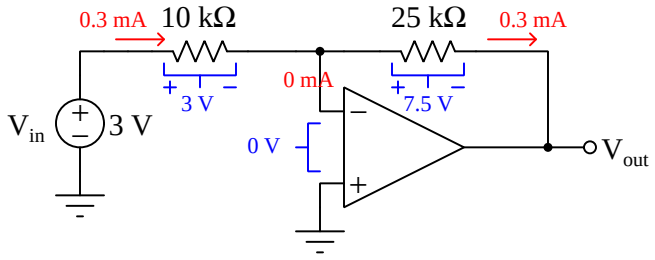
Negative feedback works to eliminate error, and so we will assume the operational amplifier is able to do this by holding the node between the two resistors at ground potential. This fact allows us to conclude, by Kirchhoff's Voltage Law, that the full input signal voltage of 3 Volts will be impressed across the 10 kΩ resistor:



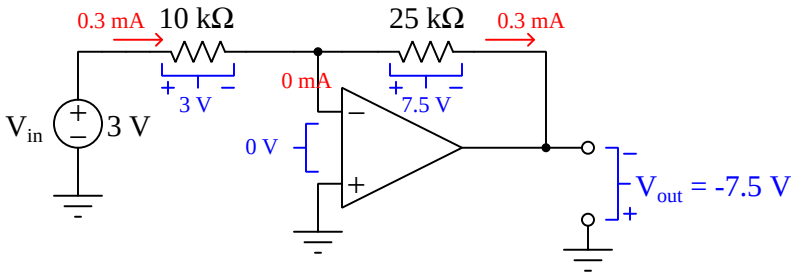
Ohm's Law allows us to calculate the current through the 10 kΩ resistor as 0.3 mA, and knowing resistors are always electrical loads allows us to conclude from the polarity of the 3 Volt drop that current must be flowing left-to-right:



Input terminal current for any differential voltage amplifier is negligible, and so by Kirchhoff's Current Law we may conclude that the entirety of this 0.3 mA current passes left-to-right through R_A, dropping 7.5 Volts with + on left and − on right:



Applying Kirchhoff's Voltage Law to ground, the V_{out} terminal, and the voltage drop across the 25 kΩ resistor, we see the output terminal must be at a potential of −7.5 Volts with respect to ground:



This agrees with our general gain formula for an inverting amplifier:

$$V_{out} = A_V V_{in}$$

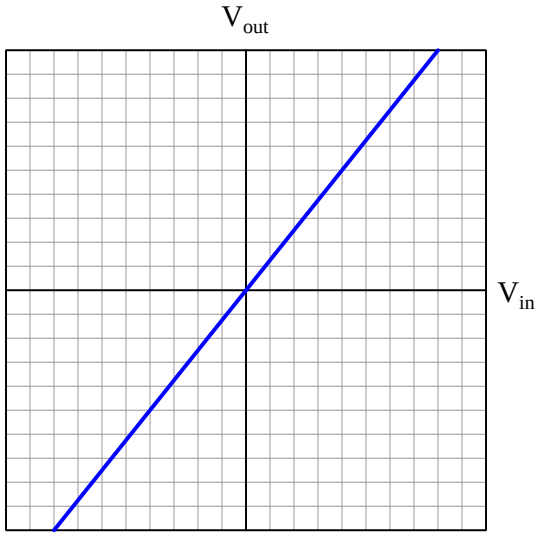
$$V_{out} = \left(-\frac{25\ \text{k}\Omega}{10\ \text{k}\Omega}\right)(3\ \text{V})$$

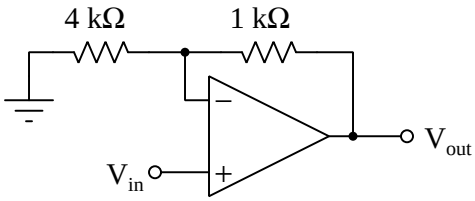
$$V_{out} = -7.5\ \text{V}$$

Both inverting and noninverting amplifier configurations have their practical uses and limitations. The most obvious difference between the two is the fact that one inverts the DC polarity of the voltage signal while the other does not. Beyond that, the inverting configuration is capable of a wide range of gain values, being able to achieve gains less than one while the noninverting configuration is limited to one as a minimum gain. Also, the amount of input resistance "seen" by the input signal source is markedly different between the two amplifiers: the noninverting configuration poses very little load on the signal source because the operational amplifier's input resistance is so very high, while the inverting amplifier forces V_{in} to bear the full burden of the resistor connected between the input terminal and virtual ground.

3.11 Amplification with offset

A common application for operational amplifiers is to *scale* one analog voltage signal range to another. Such tasks are simple if the two voltage ranges in question happen to share a common point at 0 Volts in and 0 Volts out, as shown here:





V_{in} = -8 Volts to +8 Volts

V_{out} = -10 Volts to +10 Volts

In this case, both the input and output signal range intersect at 0 Volts, which means a standard non-inverting amplifier will suffice. All we need is a voltage gain for that amplifier equal to $\frac{10}{8}$ or 1.25, which we may obtain easily enough using a $\frac{1}{4}$ feedback-to-grounding resistor ratio since we know $A_V = 1 + \frac{R_f}{R_g}$ for a non-inverting circuit.

We may describe any proportional relationship between two different signal ranges by using the standard *slope-intercept linear equation*:

$$y = mx + b$$

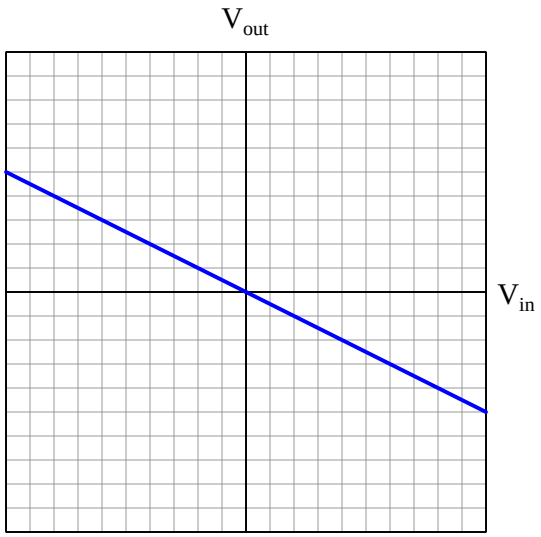
Where,

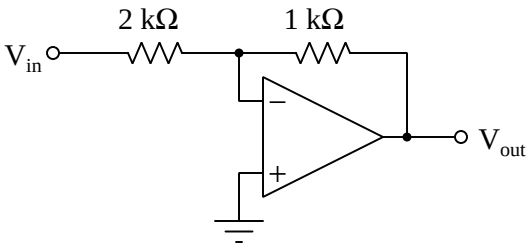
y = Output signal range
x = Input signal range
m = Slope
b = y-intercept point

In this example, the graph showing how -10 Volts to $+10$ Volts relates to -8 Volts to $+8$ Volts happens to cross the y axis at a y value of zero (i.e. $y = 0$ when $x = 0$), so $b = 0$. The slope of this linear equation, m, is simply the graph's "rise-over-run" pitch, in this case a total rise of 20 over a

total run of 16, which is $\frac{20}{16}$ or $\frac{10}{8}$ or 1.25. Our amplifier circuit's voltage gain value must be this slope value (m) in order to properly proportion the input signal to the output signal.

Devising a circuit to scale a set of voltage ranges with an inverse slope follows a similar pattern. So long as those two ranges intersect at zero Volts, a standard amplifier configuration will work just fine. In this case, the amplifier must be *inverting* so that V_{out} falls as V_{in} rises:





V_{in} = -10 Volts to +10 Volts

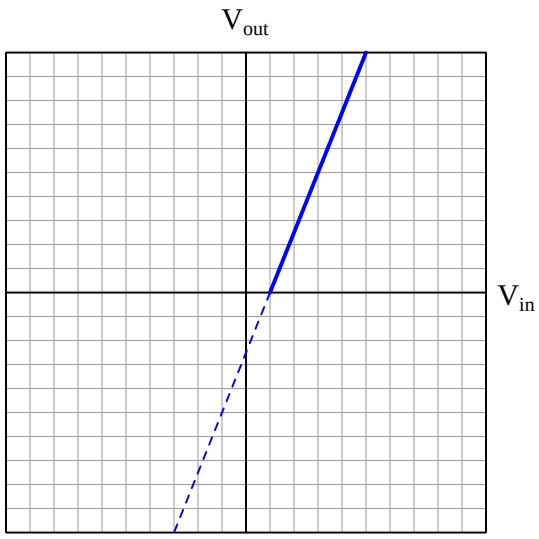
V_{out} = +5 Volts to -5 Volts

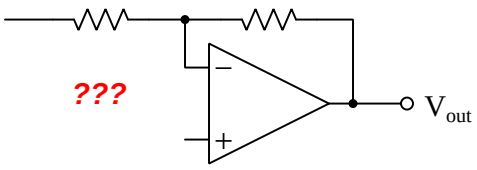
As before, the slope of this linear function is defined as its rise divided by its run. In this case the rise is actually a *fall* from +5 Volts to −5 Volts for a total fall of −10 Volts while the run is 20 Volts (from −10 Volts to +10 Volts). Thus, the function's slope is equal to $-\frac{5}{10}$ or −0.5, and since the graph once again crosses the y axis at the zero point, the y-intercept value will be zero:

$$y = -0.5x + 0$$

Achieving the requisite voltage gain of −0.5 means using a feedback-to-input resistor ratio of $\frac{1}{2}$ since we know $A_V = -\frac{R_f}{R_{in}}$ for an inverting opamp voltage amplifier circuit.

The preceding examples are both easy, involving only standard non-inverting and inverting operational amplifier circuit configurations. However, what can we do if the voltage ranges in question don't share an intersection at 0 Volts? For example, what if our input voltage range was 1 to 5 Volts while the output voltage range was 0 to 10 Volts as shown below?



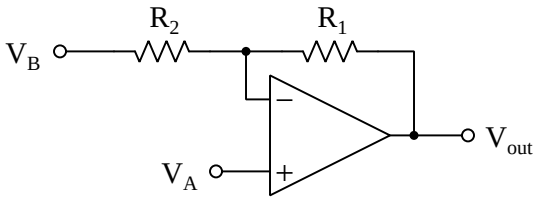


V_{in} = +1 Volt to +5 Volts

V_{out} = 0 Volts to +10 Volts

This function's rise-over-run ratio (i.e. slope) is $\frac{10}{4}$, which means the amplifier must have a voltage gain of 2.5. We also know the amplifier's characteristic must be non-inverting because we expect V_{out} to rise as V_{in} rises. However, an ordinary non-inverting amplifier configuration will not suffice because the linear function does not pass through the graph's origin $(0, 0)$ point. Instead, we should expect this amplifier to output a negative V_{out} signal if V_{in} were ever set to 0 Volts. In mathematical terms, this linear function's y *intercept* point (b) is a negative value rather than zero.

One way to "bias" or "offset" the output voltage of an amplifier is to replace the ground connection with another input terminal, connecting that new input terminal to some voltage source intended to offset the amplifier's output. In other words, our generic opamp-based amplifier circuit may be thought of as having *two* input terminals, which we will label V_A and V_B:



From the perspective of input V_A the circuit functions as a non-inverting amplifier, but from the perspective of V_B its characteristic is inverting. Whichever of these two input terminals we choose to use as V_{in} for our signal-scaling circuit, the remaining input will be used for the offset or bias function.

In fact, we may apply the *Superposition Theorem* to merge the characteristic equations for non-inverting and inverting amplifiers together into one equation describing this new amplifier-with-offset configuration. We know that $V_{out} = \left(1 + \frac{R_1}{R_2}\right) V_{in}$ for a non-inverting amplifier and that $V_{out} = \left(-\frac{R_1}{R_2}\right) V_{in}$ for an inverting amplifier. Therefore, the superposition of these two equations must be:

$$V_{out} = \left(1 + \frac{R_1}{R_2}\right) V_A - \left(\frac{R_1}{R_2}\right) V_B$$

We already know that the slope of our linear function is +2.5, and that V_A must be the signal input so as to have a non-inverting characteristic. This means $1 + \frac{R_1}{R_2}$ must be equal to 2.5, which means $\frac{R_1}{R_2}$ must be equal to 1.5 or $\frac{3}{2}$. Modifying the slope-intercept equation to reflect these realizations:

$$y = mx + b$$

$$V_{out} = mV_{in} - b$$

$$V_{out} = \left(1 + \frac{3}{2}\right) V_{in} - \left(\frac{3}{2}\right) V_B$$

The only unknown now is the necessary value of V_B to properly offset or bias this amplifier's output signal. This is easy to solve, however, by just substituting known corresponding values for V_{in} and V_{out} into the equation and then solving for V_B. Any pair of voltage values for V_{in} and V_{out} along the linear function will suffice, but here we will choose 1 Volt for V_{in} and 0 Volts for V_{out} simply to exploit the simplicity of those two numbers:

$$0 = \left(1 + \frac{3}{2}\right) 1 - \left(\frac{3}{2}\right) V_B$$

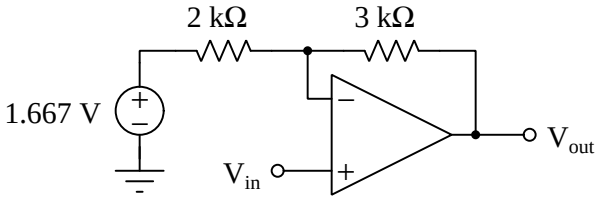
$$0 = \frac{5}{2} - \left(\frac{3}{2}\right) V_B$$

$$\left(\frac{3}{2}\right) V_B = \frac{5}{2}$$

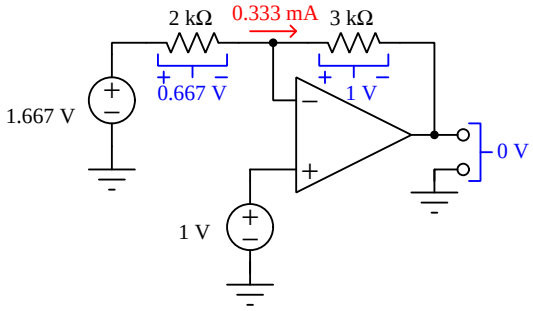
$$V_B = \left(\frac{5}{2}\right)\left(\frac{2}{3}\right)$$

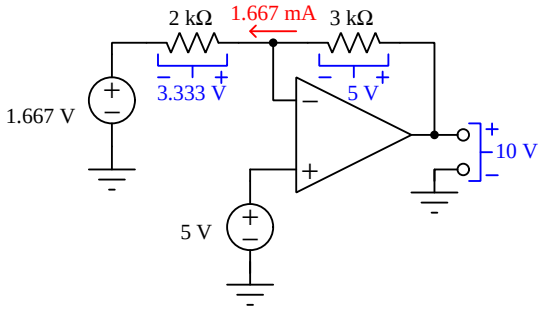
$$V_B = \frac{10}{6} = 1.667 \text{ Volts}$$

Therefore, an appropriate signal-scaling circuit for this application would be this:



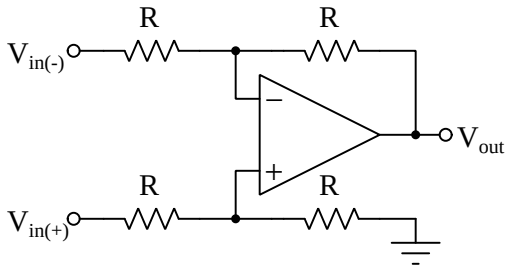
Showing this circuit amplifying an input of 1 Volt into an output of 0 Volts, and an input of 5 Volts into an output of 10 Volts:



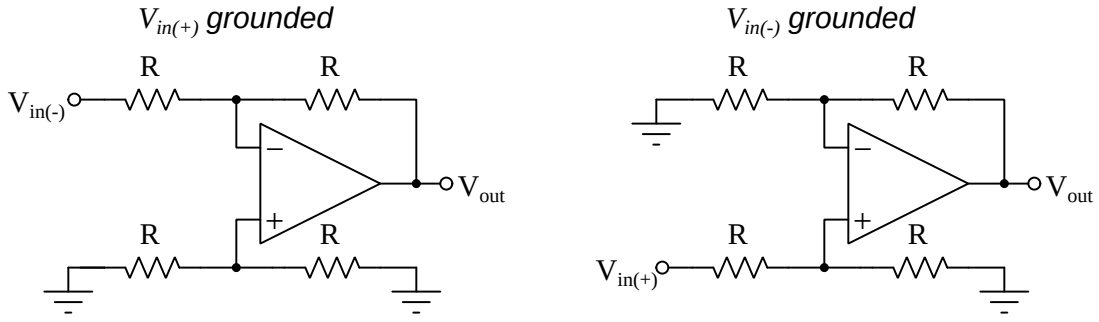


3.12 Differential amplification

Operational amplifiers are *differential-input* voltage amplifier circuits, but their voltage gain is optimized to be extremely high for use in negative-feedback networks. Some applications demand a differential-input voltage amplifier with a modest gain. Fortunately, this is relatively simple to achieve by adding more external resistors around an opamp. The following diagram shows a *differential* amplifier circuit sometimes referred to as a *subtractor*:



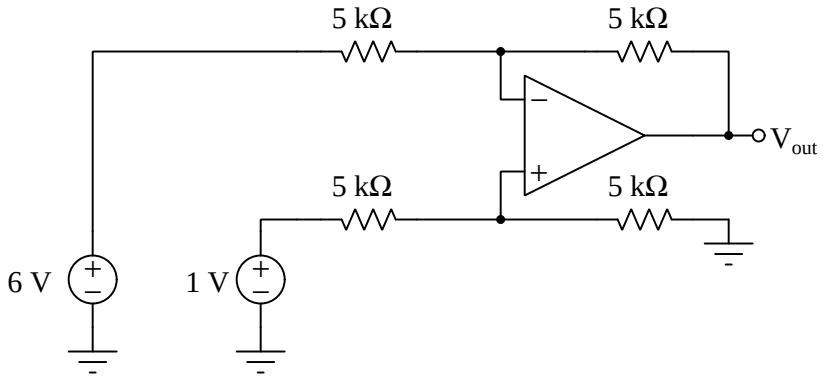
One way to analyze this circuit is to perform thought experiments where we set one of the input signals to zero (i.e. connect it to ground) and see how V_{out} responds to the remaining V_{in}. Consider the following modifications, each one with a different input terminal grounded:



In the left-hand schematic we see that the act of grounding the $V_{in(+)}$ terminal maintains a condition of ground potential on the operational amplifier's + input terminal, essentially eliminating the purpose of the lower resistors. This circuit is really the same as an inverting amplifier (seen in the previous section) with a voltage gain of *negative one*. In the right-hand schematic we see that the act of grounding the $V_{in(-)}$ terminal makes the circuit resemble a noninverting configuration. Normally, a noninverting amplifier circuit with equal resistor values in the feedback network would have a voltage gain of *positive two*, but notice how the remaining input terminal does not connect directly to the opamp but rather feeds through a voltage divider network with a 2:1 attenuation ratio. Dividing the input signal voltage by two, then boosting it by two, results in an closed-loop voltage gain of positive one. The *superposition* of these two inputs' responses is simply the *difference* between the two input signal voltages:

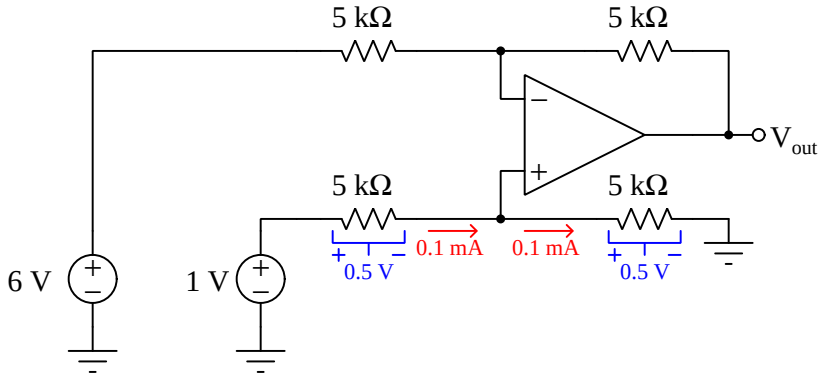
$$V_{out} = V_{in(+)} - V_{in(-)}$$

As usual it is instructive to analyze an example circuit with stated values to see how the fundamental laws of electric circuits apply. Let us consider the following differential amplifier circuit with specified resistor values and input signal voltages:

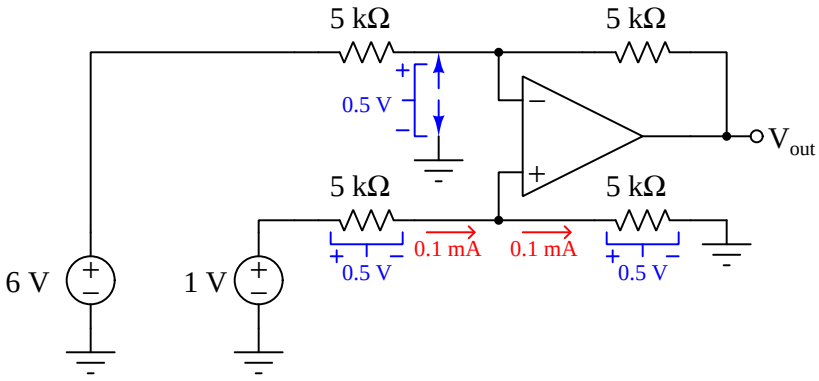


The only reasonable place to begin with our analysis is with the two lower resistors, being powered by the 1 Volt signal source. We cannot really analyze anything in the feedback network (yet) because we do not know the voltage between the opamp's inverting input and ground, since that network is being driven at the other end by the amplifier in an effort to eliminate error. The lower resistors, by contrast, are being powered only by the 1 Volt source and nothing else, so there are fewer unknowns.

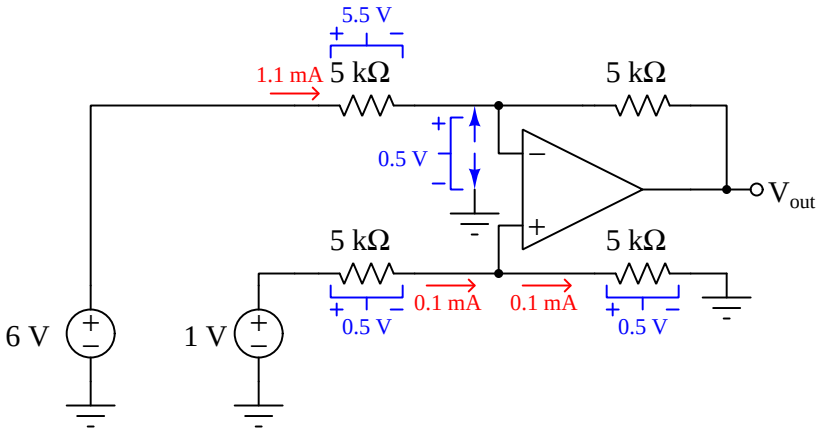
1 Volt impressed across a series total of 10 kΩ is 0.1 mA by Ohm's Law, and we know we may regard these resistors as being in series because the opamp's noninverting input is practically non-conducting with its very high input resistance. This 0.1 mA current drops 0.5 V across each of the two lower resistors:



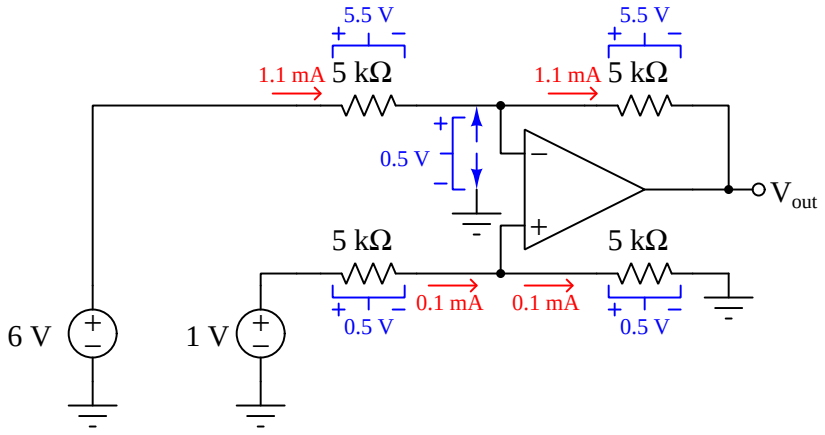
The opamp's noninverting input "sees" +0.5 Volts to ground, being the same as the voltage drop across the right-hand resistor in the lower network. Negative feedback will act to raise the other (inverting) input to that same potential as it strives to maintain zero error voltage between the two input terminals, so that the node between the two upper resistors will also have a potential of +0.5 Volts relative to ground:



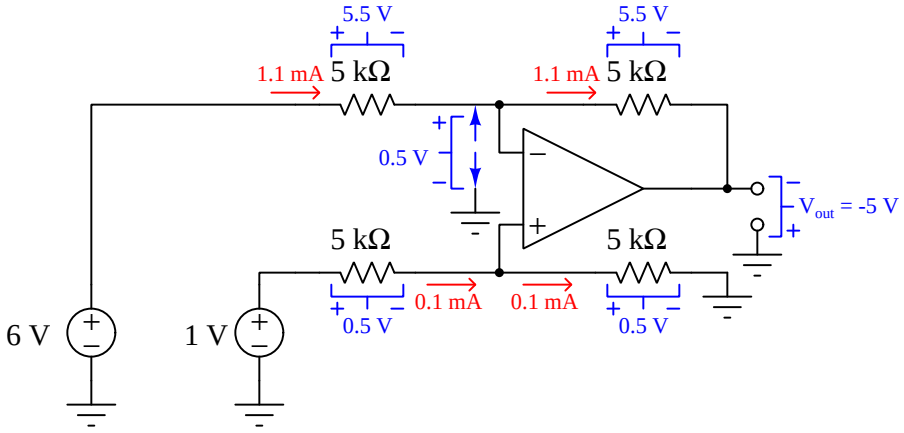
Applying Kirchhoff's Voltage Law to the upper-left resistor, we find it drops 5.5 Volts, and this in turn allows us to calculate current by applying Ohm's Law:



This 1.1 mA current must also pass through the upper-right resistor since we know neither input terminal of the opamp passes any substantial amount of current. With this, we may calculate that resistor's voltage drop (also 5.5 Volts) by Ohm's Law:

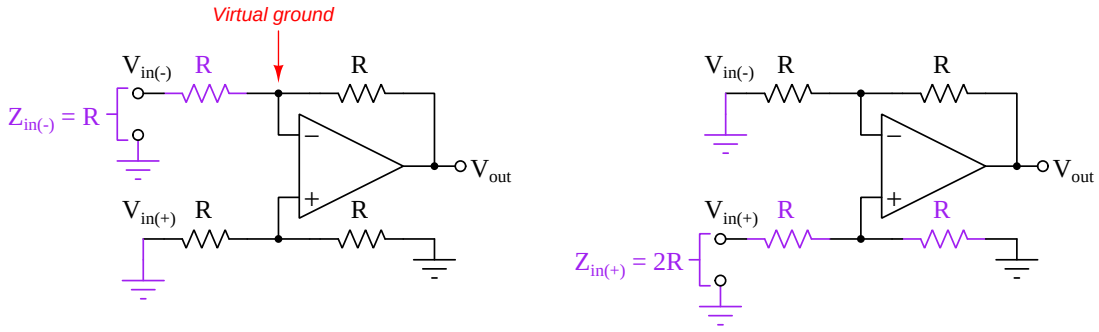


V_{out} is a single-ended voltage signal, which means it is measured in reference to ground. Using Kirchhoff's Voltage Law to step around a loop formed by ground, the V_{out} terminal, across the upper-right resistor, and back to ground, we find that V_{out} must be -5 Volts relative to ground:



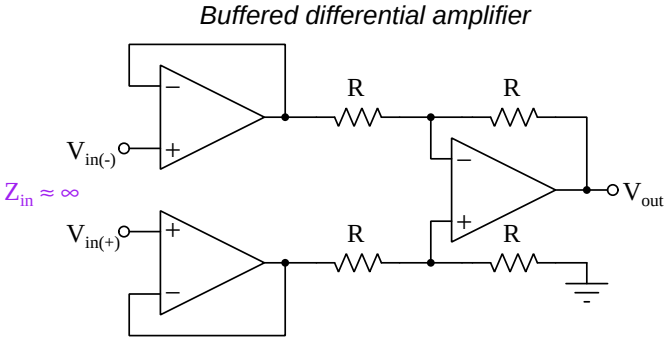
Of course, an output signal of -5 Volts is precisely what we would expect from a differential amplifier with a closed-loop gain of 1 and input signals of 6 Volts and 1 Volt on its inverting and noninverting terminals, respectively.

A characteristic of this differential amplifier circuit is that the resistor values directly define its input impedance. If we consider the basic differential amplifier circuit, looking at the input impedance of each input terminal as we assume the other input terminal is at ground potential, we see how the Z_{in} for the inverting input is simply equal to R while Z_{in} for the noninverting input is equal to $2R$:



Input impedance may be an important consideration of the signal sources for $V_{in(-)}$ and/or $V_{in(+)}$ have high internal impedance values themselves and therefore cannot source or sink any substantial amount of current to or from the amplifier input terminals without suffering signal voltage "sag". In such cases we require much more input impedance from our amplifier than the resistor values themselves provide.

A simple solution for boosting input impedance is to "buffer" the two inputs with additional operational amplifiers wired as voltage-followers:

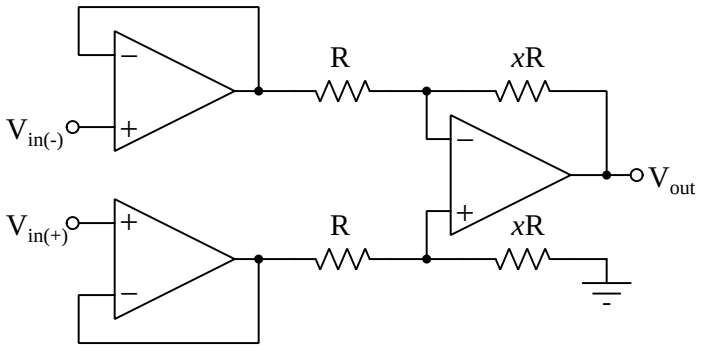


With this circuit we have the exact same differential amplification characteristics of the original circuit, but input impedance is now equal to that of the voltage-follower (buffer) amplifiers which is typically very high[10].

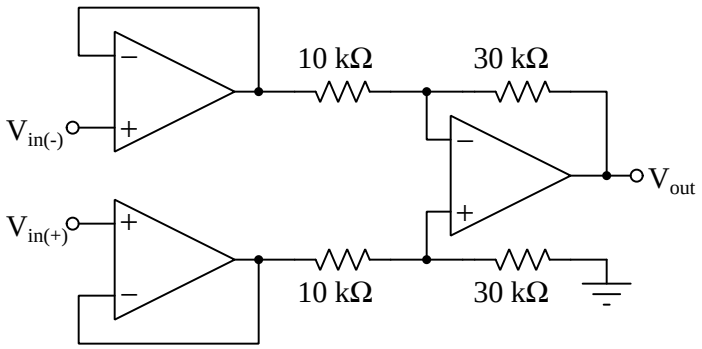
[10] With input bias current values typically in the *nano*-Ampere range for bipolar-input opamps, the effective input impedance is many millions of Ohms. For opamps designed with field-effect transistor input stages having bias current values in the *pico*-Ampere range the effective input impedance is orders of magnitude greater.

In all previous circuit examples the voltage gain of the differential amplifier was unity (1). However, if we desire a non-unity voltage gain, we may achieve that goal with little trouble simply by using *ratioed* resistor networks. Consider the following "buffered" differential amplifier circuit with resistor values of R and xR:

Differential amplifier with non-unity voltage gain



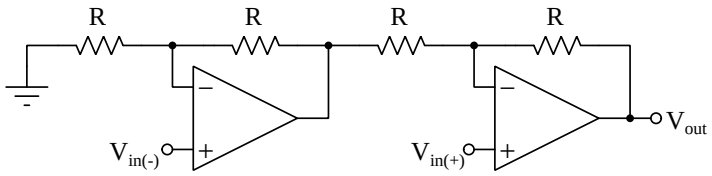
Here, x refers to a multiplication factor applied to the right-hand resistor values. For example, if we wish our differential amplifier to have a voltage gain of 3, we could use the following resistances:



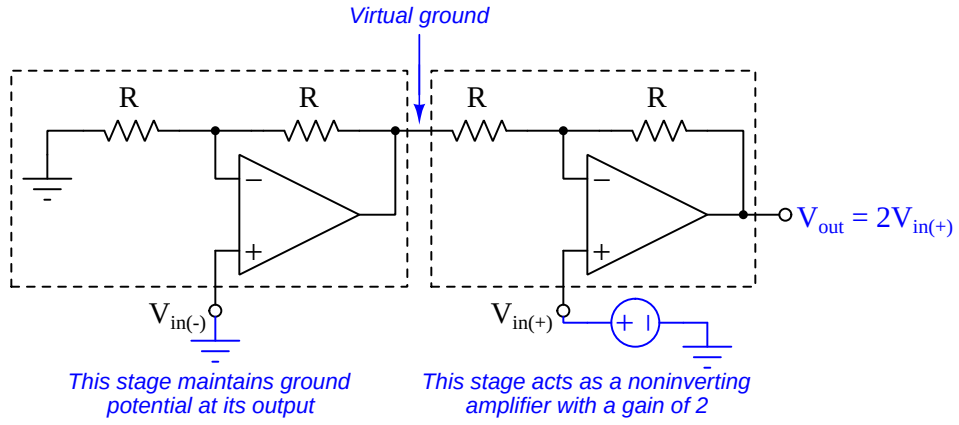
Values of x less than one are also valid if we wish to give the differential amplifier a voltage gain of less than one. For example, if we were to swap the 10 kΩ / 30 kΩ resistors' places, our differential amplifier would exhibit a voltage gain of only one-third instead of three.

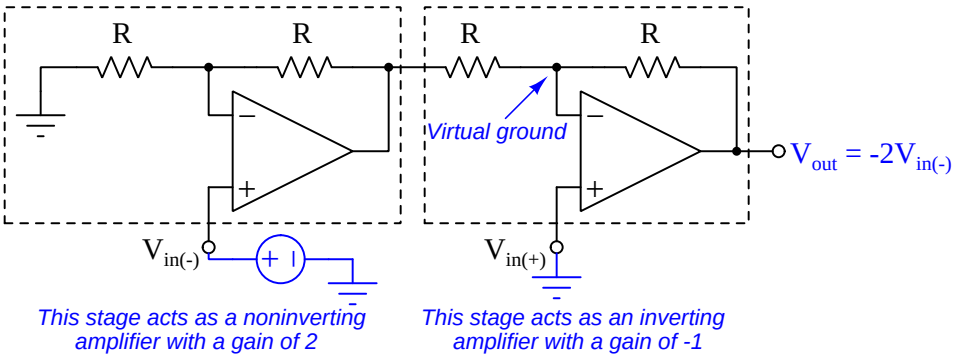
It is a worthwhile exercise to explore the operation of non-unity-gain differential amplifiers such as this by running "thought experiments" with simple input voltage values, determining the amplifier's output voltage strictly by Ohm's Law and Kirchhoff's Voltage Law.

An alternative design of differential amplifier offers the same high input impedance as the "buffered" amplifier design but uses only two opamps rather than three:



We will perform the same thought experiments applied to previous differential amplifier circuits to analyze this new circuit's behavior, namely applying a test signal to one input with the other input grounded to see how much effect each one has on V_{out}, knowing that the real amplifier's linear behavior will be the superposition of these two effects:





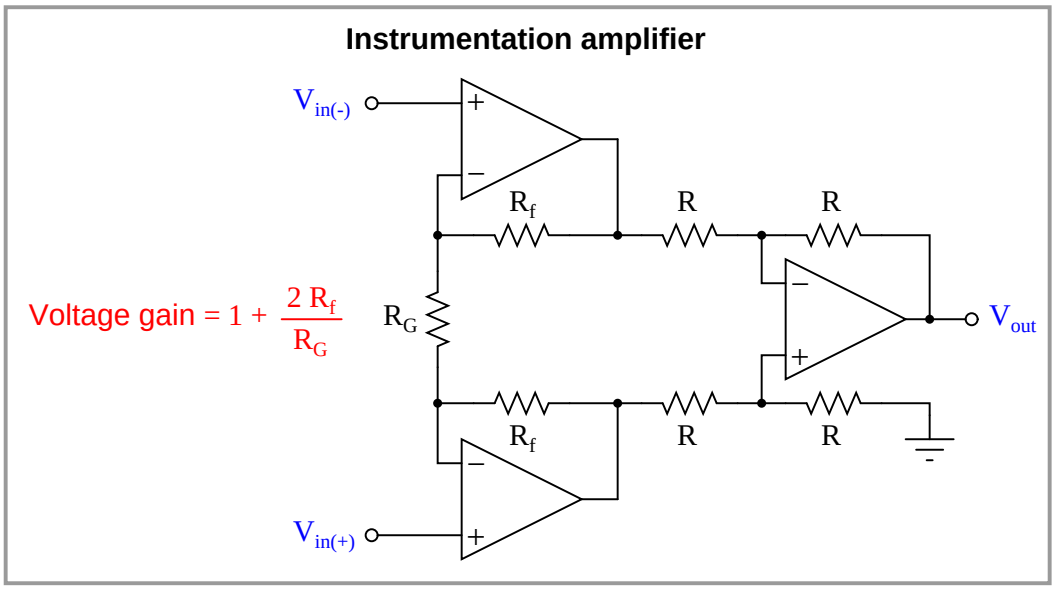
Our familiarity with inverting and noninverting opamp circuits proves useful here, allowing us to consider this differential amplifier circuit as having two stages with either inverting or noninverting

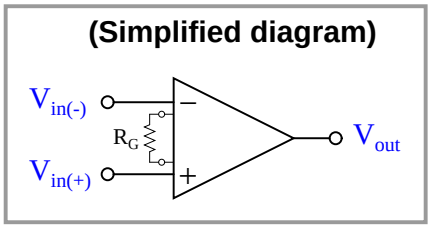
characteristics. In each case we see how V_{out} relates to the signal applied to either input by a factor of two, which means this differential amplifier circuit has a voltage gain of two when equipped with equal-valued resistors throughout.

One disadvantage of the two-opamp differential amplifier design is that it cannot provide any voltage gain *other* than two. Unlike the previous design which could be given any voltage gain desired simply by sizing the four resistors in specific ratios to each other, the two-opamp differential amplifier would amplify the two input signals unequally if we attempted using unequal-valued resistors. This fact becomes clear when we see how the inverting signal is amplified by *both* opamp stages, but the noninverting signal is only amplified by the last stage. This design amplifies both input signals equally (with opposite sign, of course) *only* if the two opamp stages have 1:1 resistor ratios.

3.13 Instrumentation amplifiers

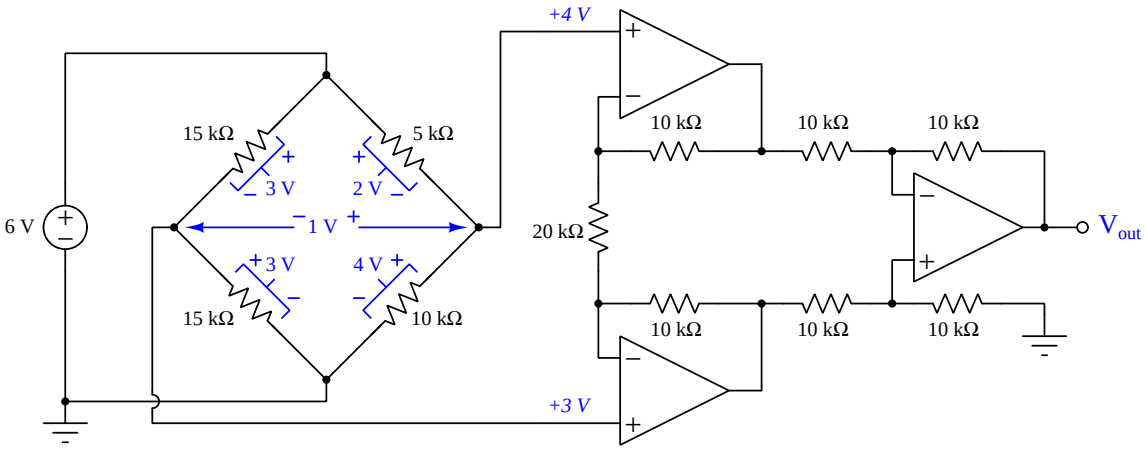
A special form of differential amplifier is the so-called *instrumentation amplifier*, noteworthy for its high input impedance as well as its ability to vary differential gain by changing the resistance of just one resistor:





Its name comes from the application of conditioning analog voltage signals from measurement sensors such as strain gauges, especially when those sensing elements are connected in a *bridge* network where differential signal amplification and common-mode rejection are necessary. Precision instrumentation amplifiers are manufactured in integrated-circuit (IC) form with all resistors precisely trimmed to proper values and two terminals made available for the end-user to connect their own gain resistor (R_G). Although it is entirely possible to construct your own instrumentation amplifier from a general-purpose opamp and some precision resistors, pre-packaged instrumentation amplifiers are well worth their cost in applications where measurement accuracy is a prime concern. The degree to which resistors may be matched to each other not only in base resistance, but also in temperature stability, is far greater when those resistors exist as devices sharing a common silicon die as opposed to being discrete components soldered into a printed circuit board.

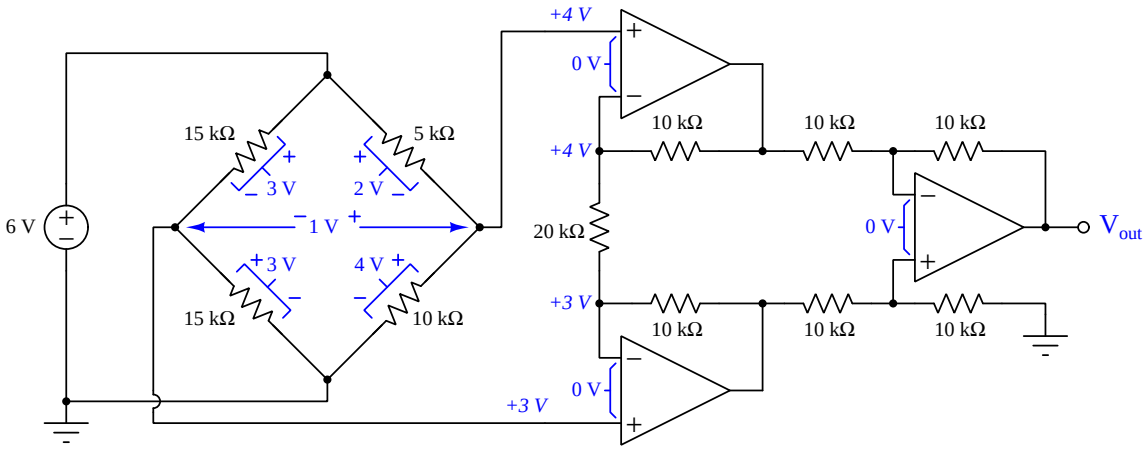
Let's analyze an instrumentation amplifier sensing voltage dropped across the span of a bridge network, where its input terminals experience both differential and common-mode signals. We will set R_G equal to twice the values of R_f in order to have a differential voltage gain of two:



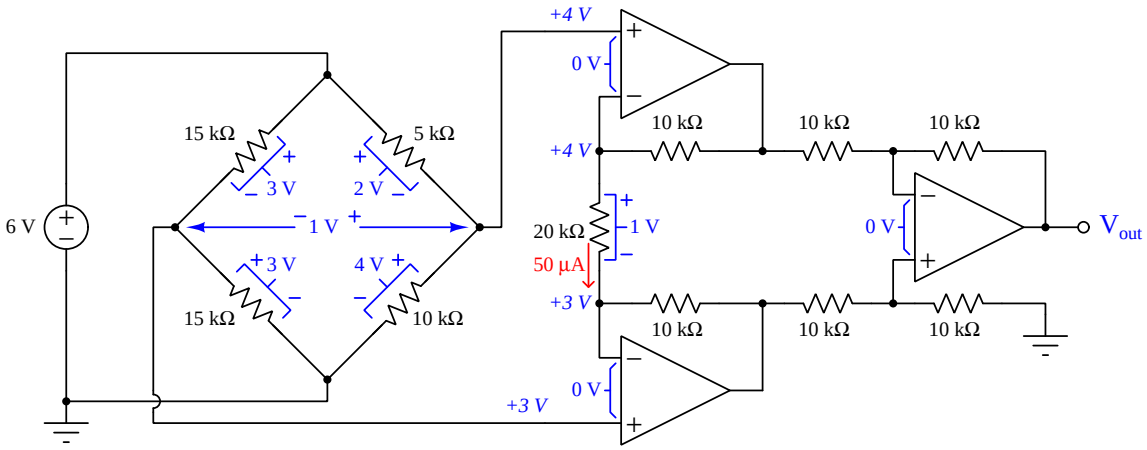
Here, the bridge network's four resistor values produce a differential voltage of 1 Volt with a common-mode voltage of +3.5 Volts[11]. We will find it useful in this circuit to label ground-referenced voltage values at certain points in the circuit, and in order to clearly distinguish a ground-referenced quantity from a differential quantity I will write all ground-referenced quantities in *italic* font. Note how the amplifier's upper input terminal has *+4 V* written by it while the lower input terminal bears the label *+3 V*. This is what a voltmeter would register if its red test lead were connected to the point in question and its black test lead were connected to ground. It is important to remember that voltage is *always relative between two points*, and that there really is no such thing as a voltage existing at any single point. So, whenever you see a voltage value specified at just one point, you know the *other* point must be whatever point in the circuit is designated as "ground".

[11] Common-mode voltage for a differential amplifier is often defined as the average of its two input voltages with reference to ground. Here, the average of +3 Volts and +4 Volts is +3.5 Volts.

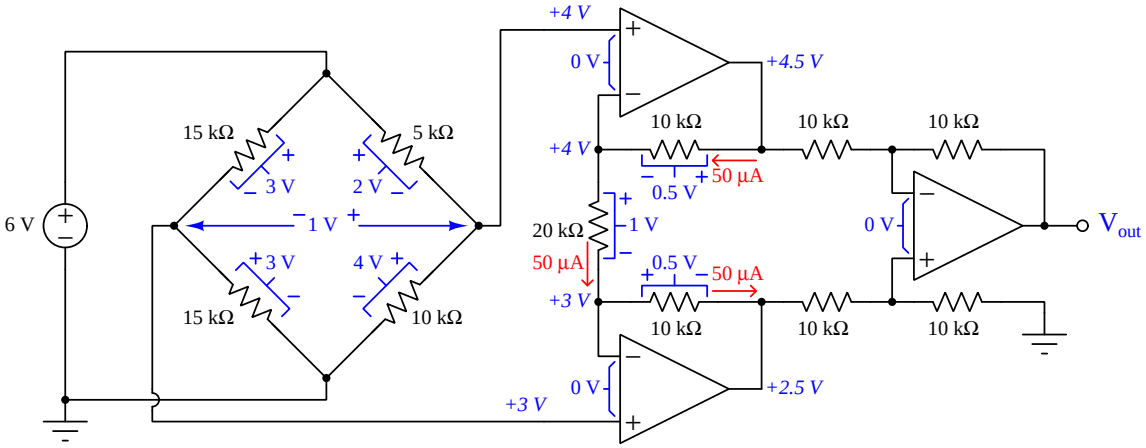
A good starting point as usual is to assume the cardinal rule of negative feedback is in effect, unless and until proven otherwise. This means zero differential voltage between input terminals at each opamp, and from this we may conclude each of the ground-referenced voltage values are replicated on either side of the 20 kΩ "gain" resistor:



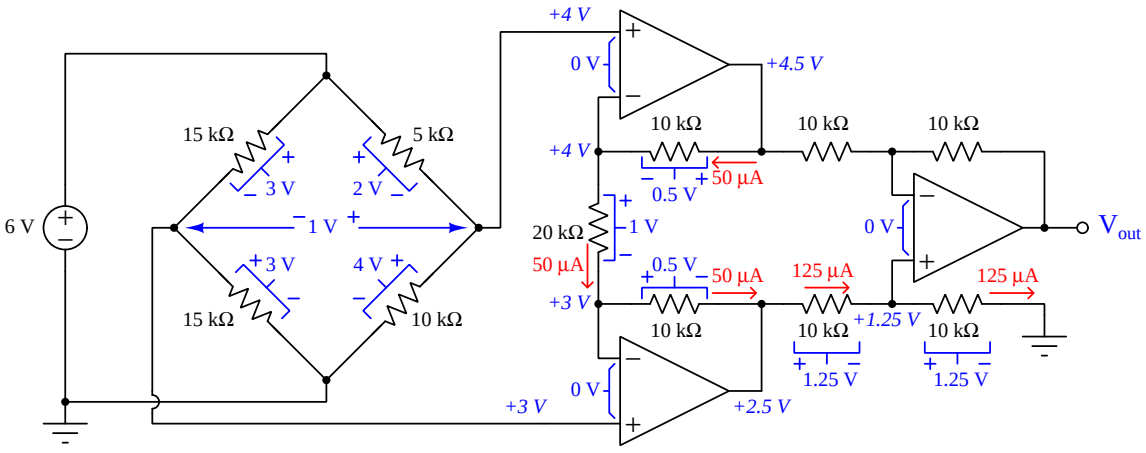
With these two ground-referenced potentials across the gain resistor, we may conclude that resistor has 1 Volt (differential) across its terminals, merely being the potential difference between +4 and +3, and from this we may also calculate current through it:



The same 50 μA current that passes through the 20 kΩ gain resistor must also pass through the two adjacent 10 kΩ resistors, because we know opamp inputs pass negligible current and therefore by Kirchhoff's Current Law we know these resistors' currents must be the same. This allows us to determine voltage drops across those two 10 kΩ resistors as well as ground-referenced voltage values at the output terminals of the first two opamps:

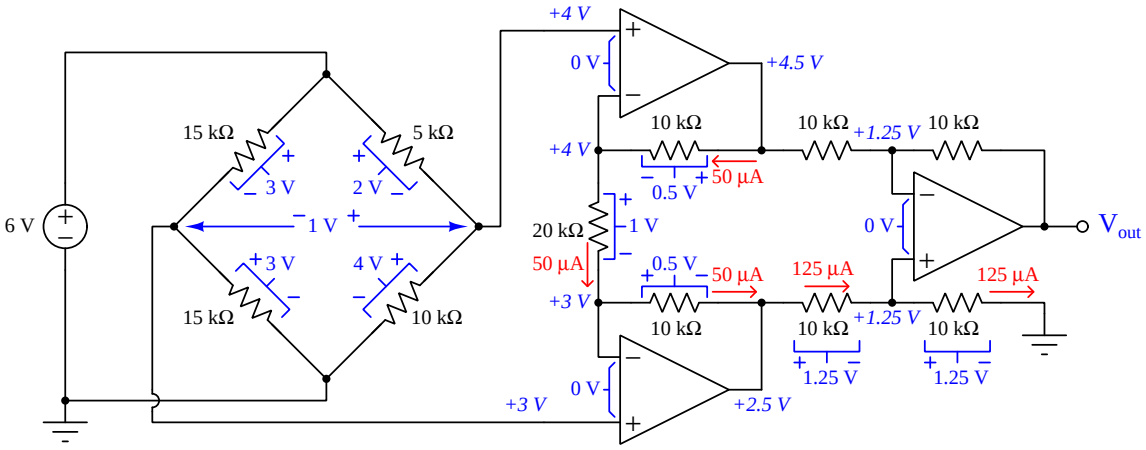


Knowing this, the two 10 kΩ resistors at the output of the lower opamp must drop 2.5 Volts, passing equal currents because (again) we know the last opamp's input terminals draw negligible current. Therefore, we may calculate voltage drops and currents as though those two resistors are in series with each other:

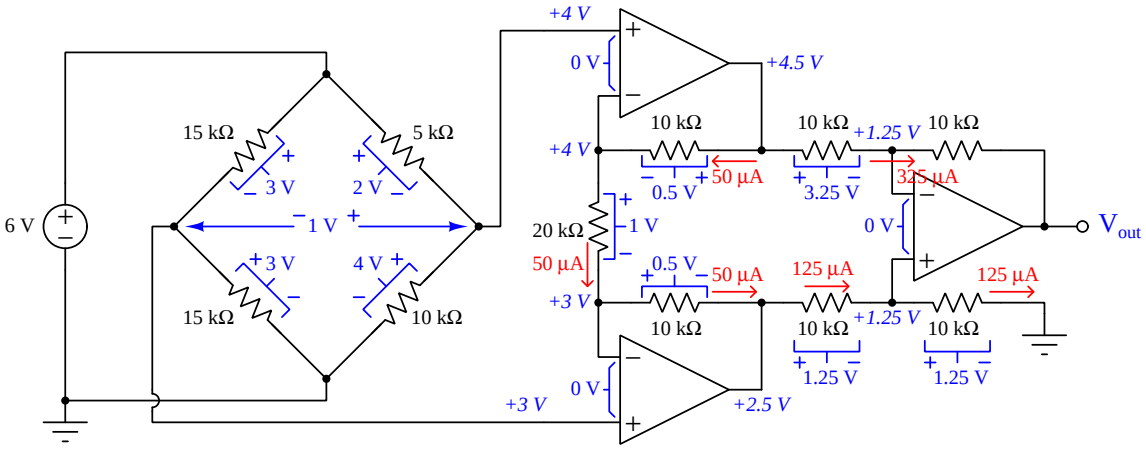


This places the last opamp's non-inverting input terminal at a ground-referenced potential of +1.25 Volts.

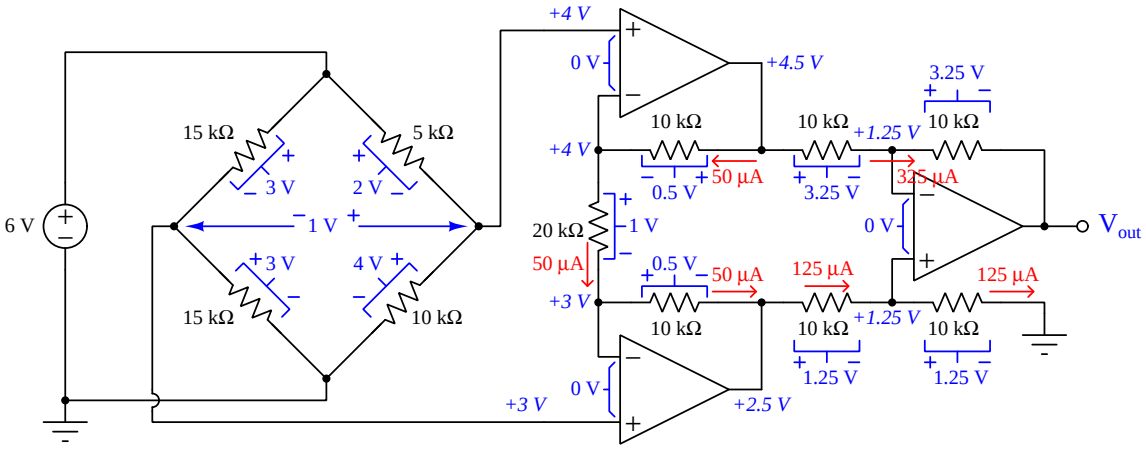
Applying the cardinal rule of negative feedback to the last opamp, that ground-referenced potential must be duplicated at its inverting terminal as well:



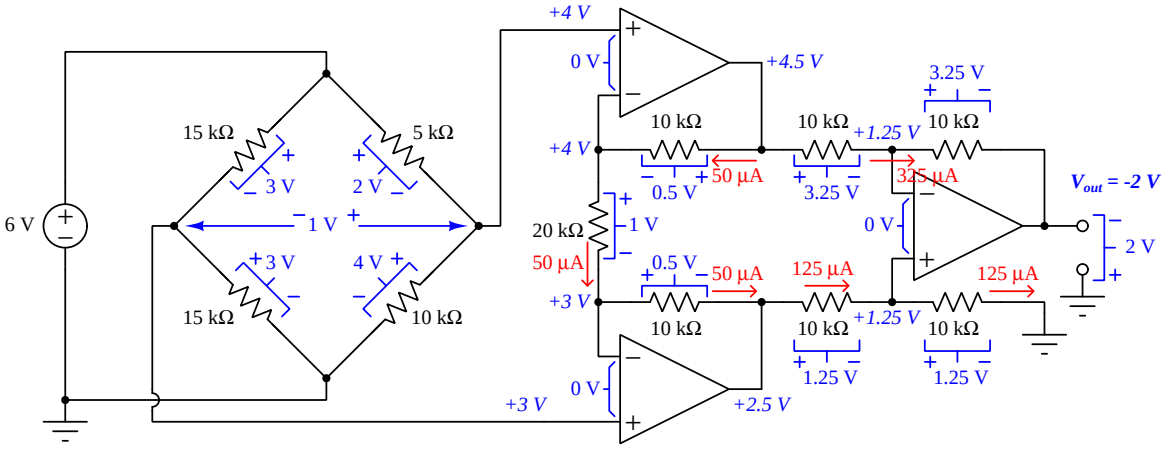
Examining the middle 10 kΩ resistor, we see it has (ground-referenced) potentials of +4.5 Volts and +1.25 Volts on left and right, respectively. Thus, it must drop a potential difference of 3.25 Volts and pass 325 microAmperes of current:



Knowing this 325 μA current must pass through the last 10 kΩ resistor as well, we know it must also 3.25 Volts:



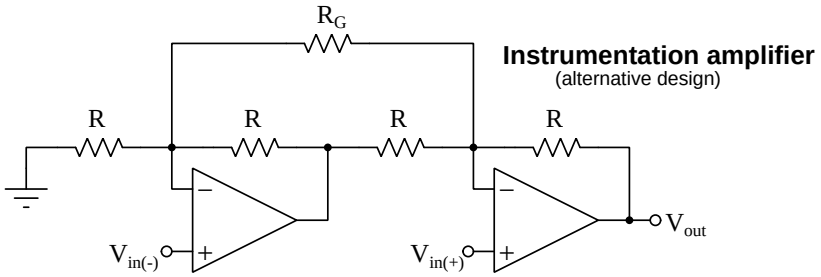
By applying Kirchhoff's Voltage Law, we may then determine that V_{out} must have a potential of −2 Volts with respect to ground:



Thus, we see this instrumentation amplifier amplifying the bridge network's 1 Volt differential signal by a gain of two to produce a −2 Volt output[12], while rejecting the 3.5 Volt common-mode voltage (averaged) between its two input terminals and ground. If a different amount of differential gain were desired, all we would have to do is replace the 20 kΩ resistor with one having another value.

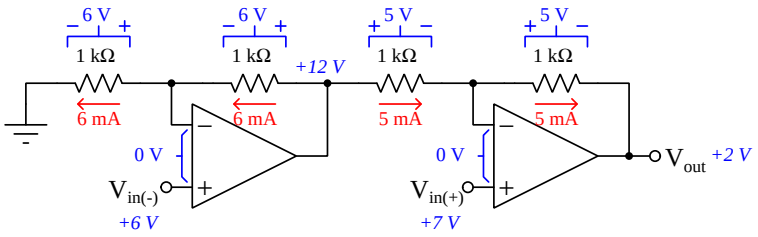
[12]The output voltage for this is negative in value (referenced to ground) only because of where the instrumentation amplifier's input terminals connect to the bridge network. If we were to swap these two inputs, the amplifier would output +2 Volts instead.

It is also possible to construct an instrumentation amplifier with single gain-setting resistance R_G using the two-opamp subtractor topology seen in the previous section:

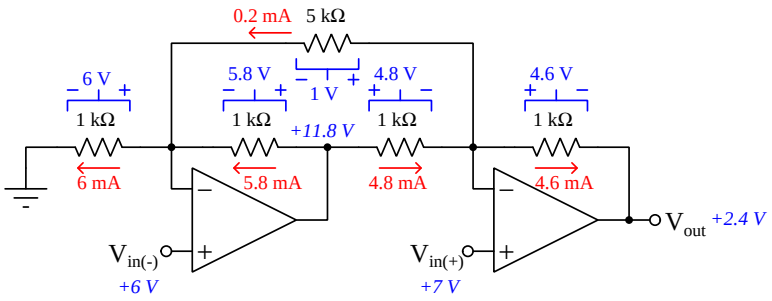


Just as with the three-opamp version the difference in the two input terminals' potentials is impressed across the gain resistor R_G, and this results in a current that influences voltages dropped across the other resistors. Without R_G in place the plain subtractor circuit will have a differential voltage gain of two. With R_G in place, this gain will be greater than two:

Without gain resistor in place



With gain resistor in place



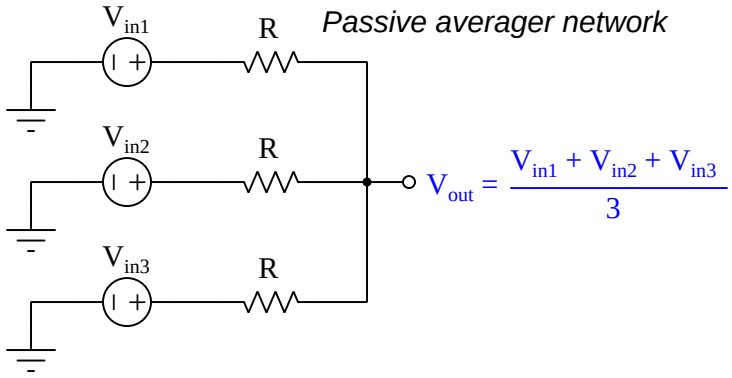
$$\text{Voltage gain} = 2 + \frac{2R_f}{R_G}$$

With an R_G value of 5 kΩ and all other resistors being 1 kΩ apiece, the voltage gain for this instrumentation amplifier will be 2.4 which is why we get +2.4 Volts output with a difference of 1 Volt between noninverting and inverting inputs.

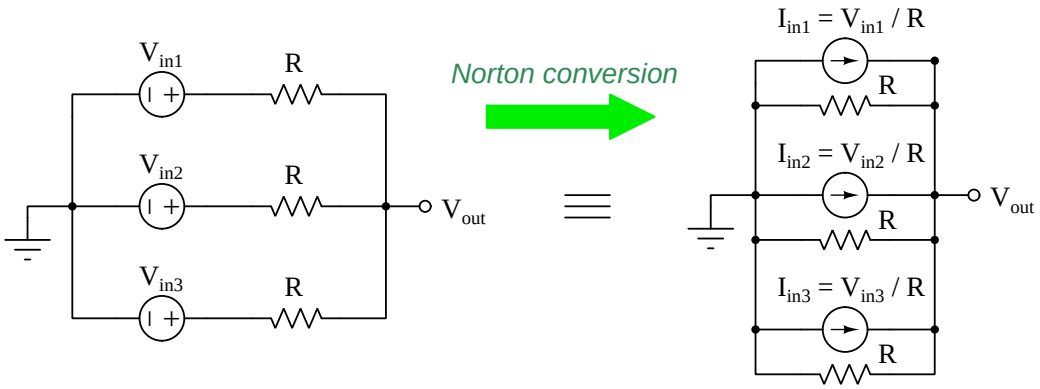
3.14 Summation

Another useful analog function implemented using operational amplifiers is *summation*, which simply means to *add* voltage signals together. As with simple signal amplification, we will find it is possible to construct both inverting and noninverting types of summer circuits using operational amplifiers.

First, the noninverting summer circuit. To begin, we will consider an extremely simple circuit called a *passive averager* consisting solely of equal-valued resistors:

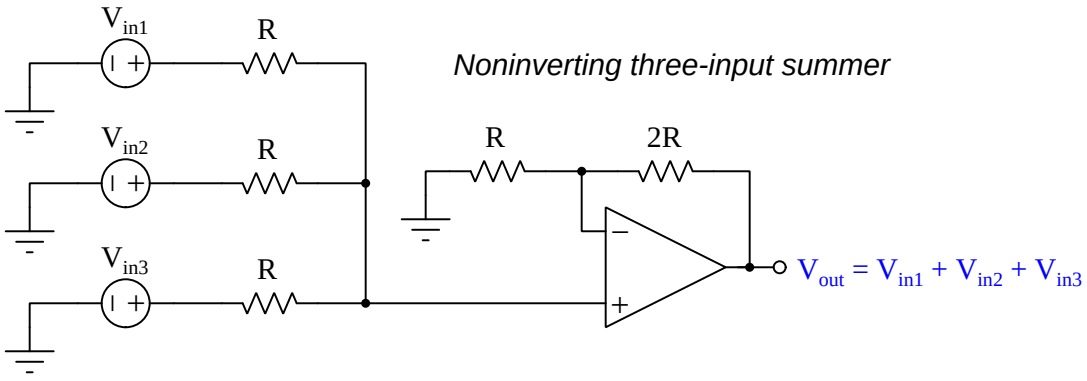


The passive averager network may be understood on the basis of *Norton's Theorem*, whereby any voltage source is equivalent to a current source in parallel with a resistance. If we consider the three voltage signal sources as being *current* sources, each current equal to the respective voltage divided by its series resistance, we may re-draw the network as three current sources in parallel with three resistances. Each source contributes $\frac{V_{in}}{R}$ current to the total network current, and when multiplied by total parallel resistance ($\frac{R}{3}$) to calculate total voltage (Ohm's Law, $V = IR$) the resulting effect on the output by each voltage source is $\frac{V_{in}}{3}$. Taken together, $\frac{V_{in1}}{3} + \frac{V_{in2}}{3} + \frac{V_{in3}}{3}$ is equal to $\frac{V_{in1}+V_{in2}+V_{in3}}{3}$ which is the mathematical definition of a three-term average[13].

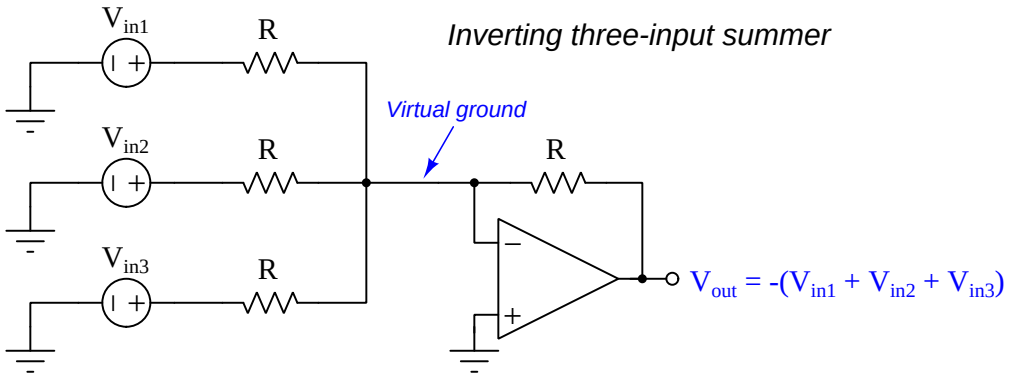


[13] Incidentally, if we build the passive averager network with *unequal* resistances we get a *weighted averager* network where the source having the lowest-valued resistance enjoys the greatest influence (i.e. highest weight) over V_{out}.

Converting this passive *averager* network into a *summer* circuit merely requires the addition of an inverting amplifier having a voltage gain equal to the number of signal inputs. In this case, since we have three input signals we need an amplifier stage with a voltage gain of three to turn the average into a sum:



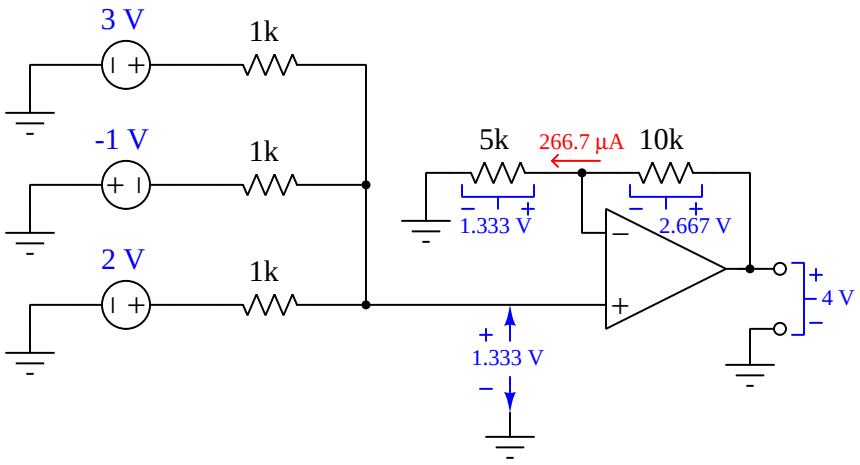
A clever design variation on this theme places the passive averager network on the inverting input of the opamp, grounding the noninverting input to establish a virtual ground point, then letting the passive network's current[14] pass through a feedback resistor:



One advantage of the inverting design is that it uses all equal-valued resistors, unlike the noninverting summer which required a different resistor value for the opamp's feedback in order to achieve the correct amount of gain.

[14] Norton's Theorem proves useful here as well. If we imagine replacing the three voltage sources with their equivalent current sources, connecting the three current sources and their resistors all in parallel, we will find that those paralleled resistors cannot pass any current whatsoever because the opamp's "virtual ground" makes both terminals of each resistor equipotential (i.e. 0 Volt drop across each resistor). The three Norton-equivalent current sources must then send their currents through to the opamp's feedback resistor where a voltage drop develops equal to the currents' sum multiplied by the feedback resistance R. If this feedback resistance is equal to the three sources' resistances, the resulting V_{out} will simply be the inverted sum of the three voltage signals.

In order to better understand each of these circuits, we will apply numerical values to the components and use fundamental principles of electric circuits to analyze the results. First, the noninverting summer:

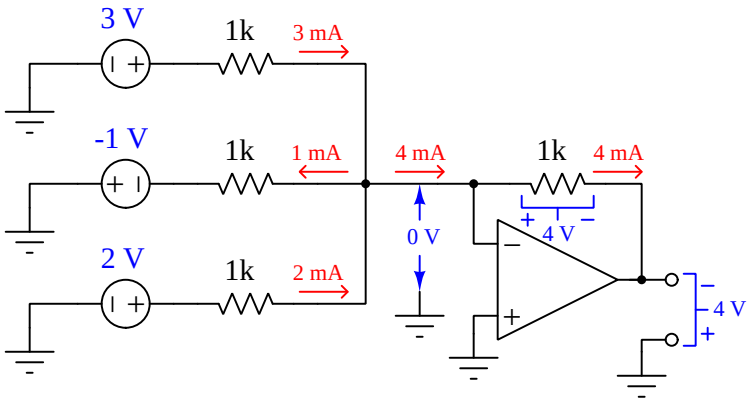


With inputs of +3, −1, and +2 Volts, the passive averager presents $\frac{3-1+2}{3} = 1.333$ Volts (relative to ground) to the opamp's noninverting terminal. We see the opamp wired for negative feedback, which means by the cardinal rule of negative feedback the opamp will do its best to maintain its other input terminal at that same potential. This places 1.333 Volts across the 5 kΩ resistor. Ohm's Law ($I = \frac{V}{R}$) then gives us the amount of current through that resistor, which must be the same as the amount of current through the 10 kΩ resistor (by Kirchhoff's Current Law) since we know opamp inputs have nearly-infinite impedance and pass negligible current. This current drops double the voltage across the 10 kΩ resistor as the 5 kΩ resistor – 2.667 Volts instead of 1.333 Volts. Looking at V_{out} as a ground-referenced signal, we see it must be equivalent to the combined voltage drops of the 5 kΩ and 10 kΩ resistors by Kirchhoff's Voltage Law, therefore V_{out} must be equal to 4 Volts which is the sum of 3 Volts, −1 Volt, and 2 Volts.

The importance of analyzing operational amplifier circuits by these fundamental principles cannot be over-emphasized as an important strategy for learning! Mastering the application of Kirchhoff's Laws, Ohm's Law, the cardinal rule of negative feedback, network theorems, etc. yields the enduring benefit of being able to analyze *any* circuit you encounter, as well as being able to more effectively diagnose malfunctioning circuits. Take this opportunity to practice applying these principles to every new operational amplifier circuit you study, and you will find yourself a much better-educated[15] person!

[15] Memorization of formulae without understanding is one of the hallmarks of technical *training*, where the goal is mere imitation and repetition. A trained person can only solve very specific types of problems they have seen before. The goal of *education*, on the other hand, is the ability to creatively apply first principles to *any* relevant problem one might encounter. Your value as a technical professional primarily lies in your ability to solve novel problems, and for this reason you should have a keen interest in and priority for education over training.

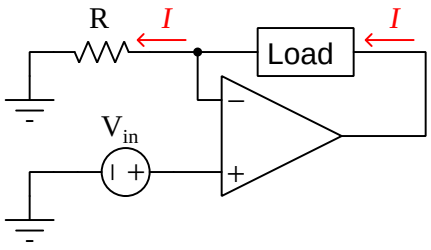
Next, let us consider an inverting summer circuit with the same input voltage values:



Here, the common node of the passive averager circuit is held at ground potential (i.e. "virtual ground") by the opamp as it tries to maintain its inverting input terminal at the same potential as its (grounded) noninverting input terminal. This is the cardinal rule of negative feedback in action once again. This means each of the input source's resistors experiences the full voltage of its respective source, making the current through each of those resistors a simple Ohm's Law calculation ($I = \frac{V}{R}$). These currents mesh together at the junction point to make a sum of 4 milliAmperes headed toward the opamp's feedback resistor. Negligible current enters or exits the inverting input terminal, and so by Kirchhoff's Current Law the entirely of this 4 mA current passes through the 1 kΩ feedback resistor, dropping 4 Volts. Given the virtual ground location, this means V_{out} with reference to ground must be the voltage dropped across the feedback resistor: -4 Volts. This output voltage value is simply the negative sum of the $+3$, -1, and 2 Volt input signals.

3.15 Precision current circuits

Another application of operational amplifier technology is the precise conversion of voltage signals into current signals. Consider the following circuit consisting of just[16] one opamp, one resistor, a variable-resistance load, and a voltage signal source:



By the cardinal rule of negative feedback, the operational amplifier will do its best to maintain the inverting input terminal's potential equal to that of its noninverting input. This means the voltage dropped across resistor R will be equal to V_{in} by Kirchhoff's Voltage Law. Ohm's Law then tells us current through that resistor must be equal to the input voltage divided by that resistance ($I = \frac{V}{R}$). Applying Kirchhoff's Current Law to the junction point near the inverting input, and knowing that opamp input terminals source and sink negligible current, leads us to conclude the load must experience the same amount of current as the resistor.

What is remarkable about this value of current is that it is defined entirely by V_{in} and R, and not by the load. So long as the opamp is able to drive its output with enough current to make V_R equal to V_{in}, the load *must* carry that current, and that current is set by $\frac{V_{in}}{R}$. This is true even if the load's resistance happens to vary: the opamp's output terminal voltage will change to follow suit, but the load current will remain constant. In summary, we have made the operational amplifier behave as a nearly perfect *current source* from the perspective of the load.

[16] Of course, a DC power supply is also necessary to energize the opamp, but it has been omitted for simplicity.

Let us consider such a circuit with $R = 2.2$ kΩ, $V_{in} = 3$ Volts, a DC supply voltage for the opamp of +12 Volts, an opamp capable of rail-to-rail output voltage swings, and a load with a varying resistance. Tabulating the resistor's voltage drop, load current, load resistance, the load's voltage drop, and the opamp's output terminal voltage as load resistance increases 1000 Ohms at a time, we see how this circuit acts to regulate current through the load to the extent that it can:

V_R	I_{load}	R_{load}	V_{load}	V_{out}
3 Volts	1.364 mA	1 kΩ	1.364 Volts	4.364 Volts
3 Volts	1.364 mA	2 kΩ	2.727 Volts	5.727 Volts
3 Volts	1.364 mA	3 kΩ	4.091 Volts	7.091 Volts
3 Volts	1.364 mA	4 kΩ	5.455 Volts	8.455 Volts
3 Volts	1.364 mA	5 kΩ	6.818 Volts	9.818 Volts
3 Volts	1.364 mA	6 kΩ	8.182 Volts	11.182 Volts
2.870 Volts	1.304 mA	7 kΩ	9.130 Volts	12.000 Volts
2.588 Volts	1.176 mA	8 kΩ	9.412 Volts	12.000 Volts
2.357 Volts	1.071 mA	9 kΩ	9.643 Volts	12.000 Volts

As the load's resistance increased from 1 kΩ all the way up through 6 kΩ the opamp was able to maintain a constant current of 1.364 milliAmperes through it, as though the opamp were a fixed-value current source. However, once the load's resistance climbed high enough to require more than 12 Volts at the opamp's output terminal – which of course the opamp cannot deliver when powered by a 12 Volt source – V_{out} became fixed at a constant value of 12 Volts. As load resistance continued to climb, the load and resistor merely behaved as a voltage divider with a fixed 12 Volt source, with current continuing to dwindle.

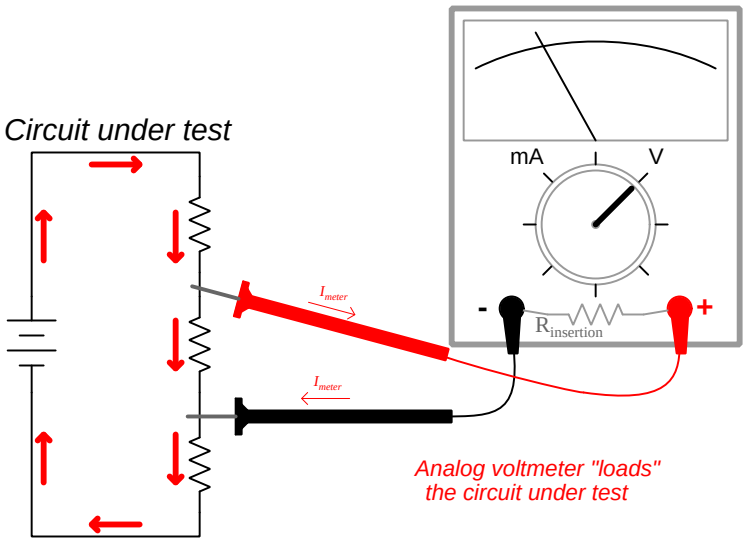
Here we see a case where the cardinal rule of negative feedback appears to fail: as V_R fell below 3 Volts while V_{in} remained at 3 Volts, the opamp's input terminals were no longer equipotential. *This tells us the cardinal rule of negative feedback has its limits.* To review:

Cardinal Rule of Negative Feedback

> *Given sufficiently high open-loop gain, the amplifier's output signal swings as far as it must within the limits of the power supply "rail" voltage to eliminate any "error" between its two input terminals.*

Note how the opamp's output voltage swings as far as it must *within the limits of the power supply "rail" voltage limits* to eliminate error between the two input terminals. Useful though they may be, operational amplifiers are not omnipotent. If the scenario demands a greater output voltage than the power supply and/or opamp is capable of delivering, the opamp will fail to eliminate that error. The maximum amount of voltage that any constant-current circuit is able to deliver to a load while maintaining proper current regulation is called its *compliance voltage*. Here, with a 3 Volt V_{in} signal and a 12 Volt supply, our circuit's compliance voltage is 9 Volts.

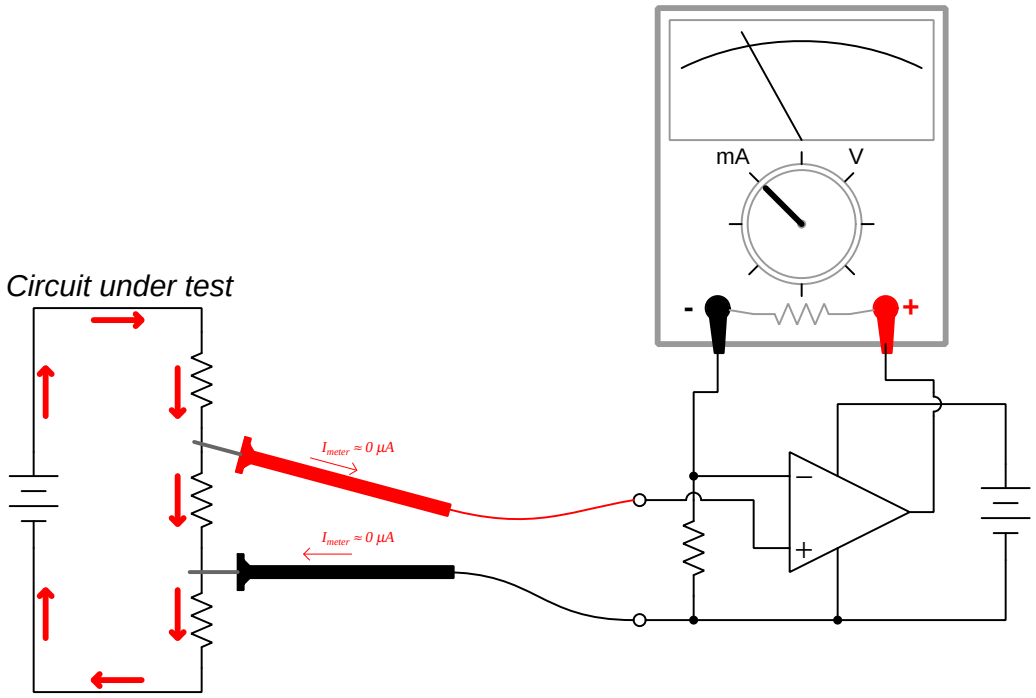
One practical application of such a circuit is improving the performance of an analog voltmeter, a type of electrical meter utilizing an electromechanical meter "movement" mechanism to indicate measured voltage by the motion of a pointer along a printed scale. Such electromechanical movements require small amounts of current to produce the magnetic field necessary to deflect the pointer and register a measurement, and in a regular "passive" meter this current must be drawn from the circuit being tested:



The small amount of current required by the analog meter movement's insertion resistance unfortunately "loads" the circuit being tested, and in so doing alters the voltage we wish to measure. Improving this state of affairs consists fundamentally of increasing the voltmeter's insertion resistance so that it draws less current from the circuit under test. We can do this by re-designing the movement mechanism to require less magnetic force and therefore less current, but this makes the movement more physically fragile. At some point we need a meter robust enough to carry around and use, and this means a movement requiring some finite (albeit small) amount of current to operate.

A legacy solution to this problem was the *vacuum-tube voltmeter* or *VTVM* which used a tube-based electronic amplifier to sense the measured voltage, and a battery to provide the driving power to the analog meter movement. We may use modern operational amplifier technology to do the same, and achieve insertion resistance values far superior to most digital multimeters!

Inserting a battery, resistor, and opamp into the circuit dramatically improves the "loading" effect, especially if we use an opamp with field-effect transistors in its input stage to minimize input bias currents:

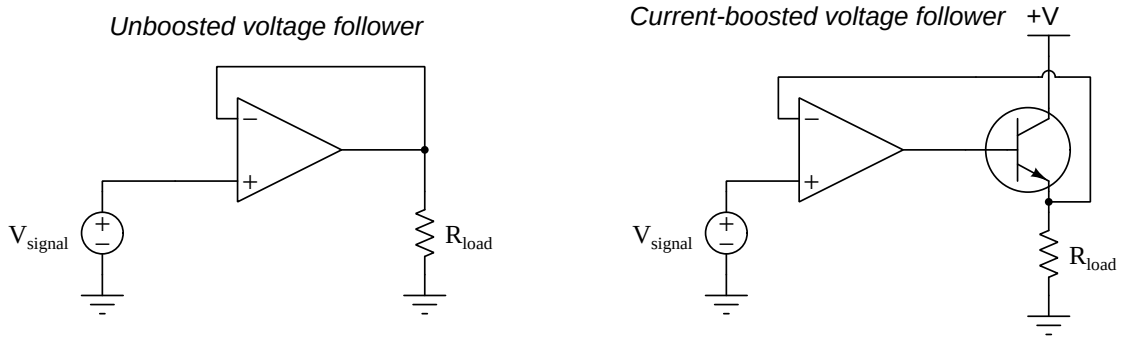


As with the current-source circuit seen earlier, this opamp does its best to maintain voltage drop across the resistor equal to voltage applied to the noninverting input terminal from the circuit under test, and drives an equivalent amount of current through the meter movement. Thus, the opamp serves as a current source to the meter, with the meter now set to respond to current rather than voltage (i.e. selecting the milliAmpere range on the meter). It is the opamp's own power-supply battery, however, providing energy to the analog meter to move its pointer, and in this configuration the meter's pointer deflection becomes an indication of voltage applied to the opamp. With this self-contained energy source, the meter movement need not be made as sensitive as before since it no longer demands energy from the circuit under test, making possible high-quality analog meters with cheaper meter movement mechanisms.

3.16 Boosted output

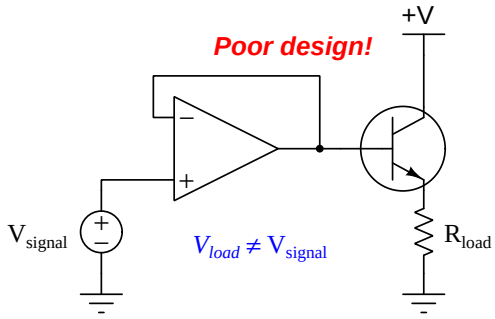
Operational amplifiers are incredibly versatile devices as far as analog signal processing is concerned, but in general they are quite limited in output power. Output voltage swing is limited, of course, by the power supply rail voltages. Output current is limited by the power dissipation ratings of the opamp's internal final-stage transistors, and being integrated circuits these transistors are usually quite small.

Fortunately, though, it is a fairly simple matter to augment an operational amplifier with discrete transistors on its output to provide a boost in output current and/or voltage. Take for example this common-collector booster stage:



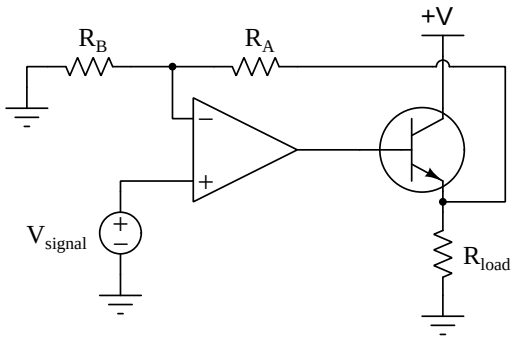
The voltage follower circuit on the left attempts to duplicate V_{signal} at R_{load}, using the opamp as an amplifier with a voltage gain of unity (1). For "heavy" (i.e. high-current) loads demanding more current than the opamp alone can provide, however, the current booster circuit is useful. The bipolar junction transistor only requires $\beta + 1$ times less current at its base terminal than the load requires in order to successfully drive that load.

Note how the negative-feedback wire connects from the opamp's inverting input directly to the load, so that the opamp is able to directly sense load voltage and drive the BJT as hard as it needs to be driven to make load voltage equal to V_{signal}. This is a very important feature of the design, in order to maintain fidelity of load voltage to V_{signal}. Had we simply left the feedback wire connected to the opamp's output terminal, the load voltage would surely be unequal to V_{signal} due to V_{BE} of the transistor:



This exact same principle works for voltage-multiplying opamp circuits as well. We simply use a discrete transistor as a "power stage" after the opamp's output, and connect the opamp's feedback to the load rather than to the opamp's own output terminal:

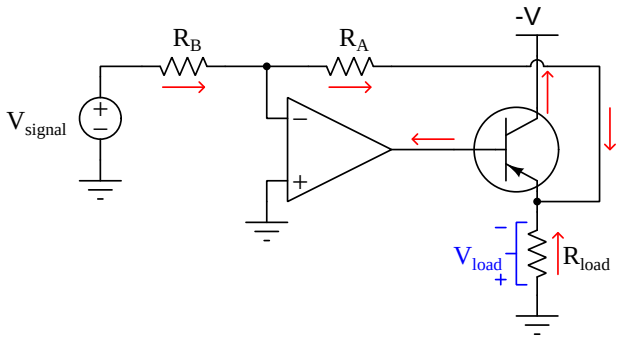
Boosted non-inverting amplifier



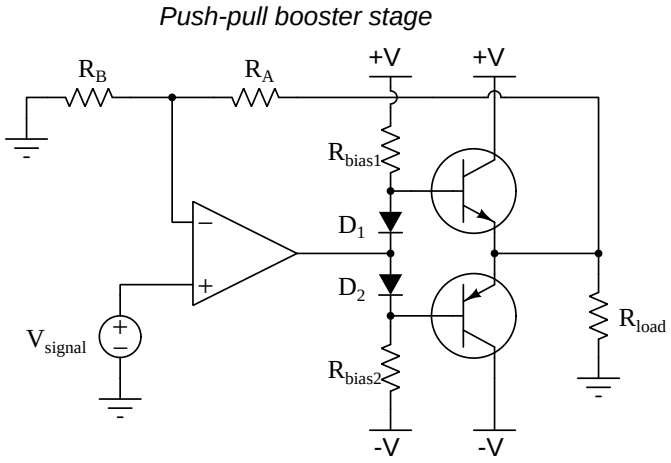
In this particular case, V_{load} will be equal to V_{signal} multiplied by $\frac{R_A}{R_B} + 1$, since the opamp is configured as a noninverting amplifier.

Inverting opamp circuits may also be boosted, but we must pay close attention to voltage polarities and current directions. The following circuit, for example, makes $V_{load} = -V_{signal} \left(\frac{R_A}{R_B}\right)$:

Boosted inverting amplifier



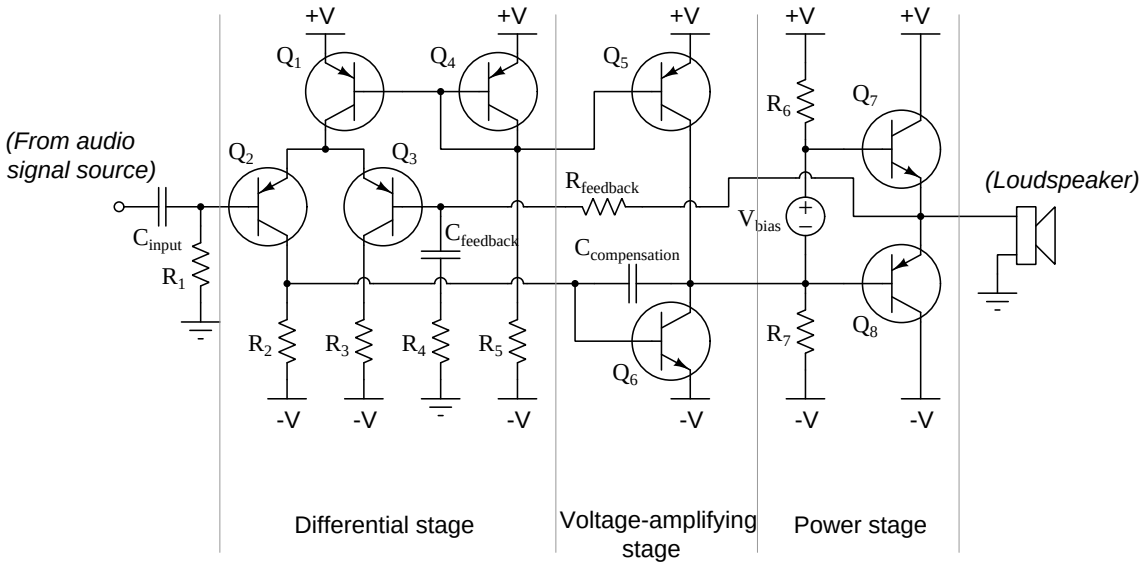
From these two boosted opamp circuits we may combine concepts to form an opamp with a push-pull booster stage of complementary bipolar transistors:



Two diodes and biasing resistors provide Class-B biasing for the complementary transistor pair, in order to minimize crossover distortion. For low-frequency applications this biasing may not be necessary because the opamp's slew rate will be high enough to "jump" quickly through the crossover point to provide faithful reproduction of the amplified V_{signal} at the load, but at higher frequencies this biasing "helps" the opamp's output voltage not have to rise or fall quite as rapidly to achieve the desired V_{load} value.

As with all these examples, we see the crucial design feature of the feedback wire connecting from the opamp's inverting input to the load itself rather than to the opamp's output terminal. This "load-sensing" feedback permits the opamp to control load voltage itself and thereby compensate for any V_{BE} voltage drops. Although the above example uses a non-inverting opamp configuration, the concept works just as well for inverting amplifiers.

The concept of negative feedback being used to control amplification through a push-pull transistor stage is commonly seen in audio power amplifier design, where sometimes instead of using an integrated circuit (IC) operational amplifier the designer(s) instead build their own differential amplifier stage using a differential pair of transistors:



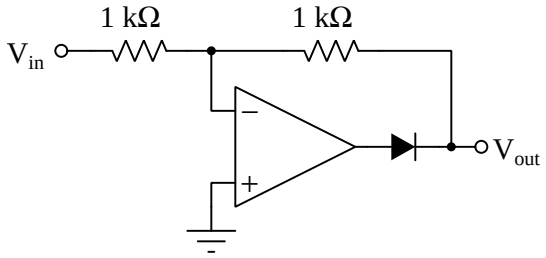
Active loading of the differential pair (using current-mirroring transistors Q_1 and Q_4 ensures the differential voltage amplification stage will have a high voltage gain. Negative feedback, however, will temper this gain to reasonable levels suitable to the amplification of music and other audio information.

Here, the negative feedback signal path is clear to see from the labeling of $R_{feedback}$. The series combination of $C_{feedback}$ and R_4 forms a low-pass filter network[17] designed to boost over-all voltage gain at high frequencies (by the reactance of $C_{feedback}$ decreasing as frequency increases). R_4 serves to establish a minimum impedance between the base of Q_3 and ground, so that voltage gain does not climb too high which could create high-frequency oscillations otherwise.

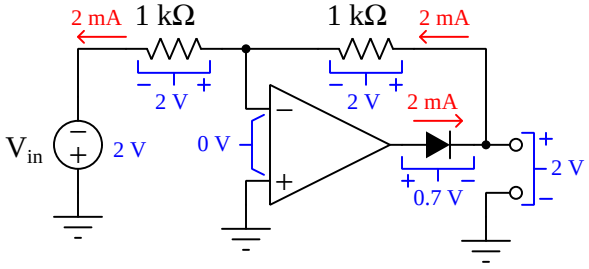
[17] Specifically, this would be a low-limited, low-pass filter network where the filtered output signal can never diminish to zero Volts no matter how high the frequency gets.

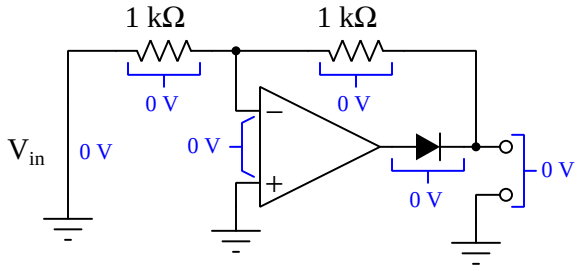
3.17 Precision rectification

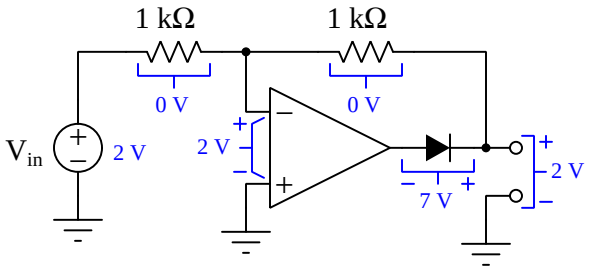
Suppose we added a diode to a simple inverting amplifier circuit with a voltage gain of negative 1:



One way to assess how this modified circuit will behave over a range of input signal voltages is to begin with negative, zero, and positive "test cases" (shown below). We will assume a split DC power supply of ±5 Volts and an opamp capable of rail-to-rail output:

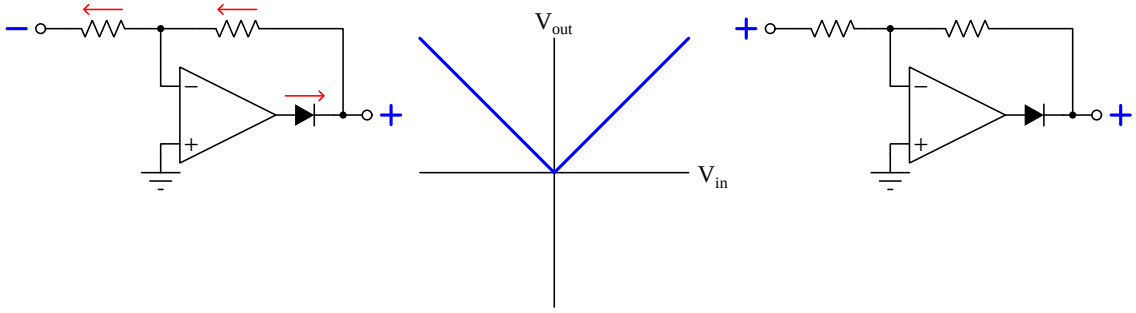






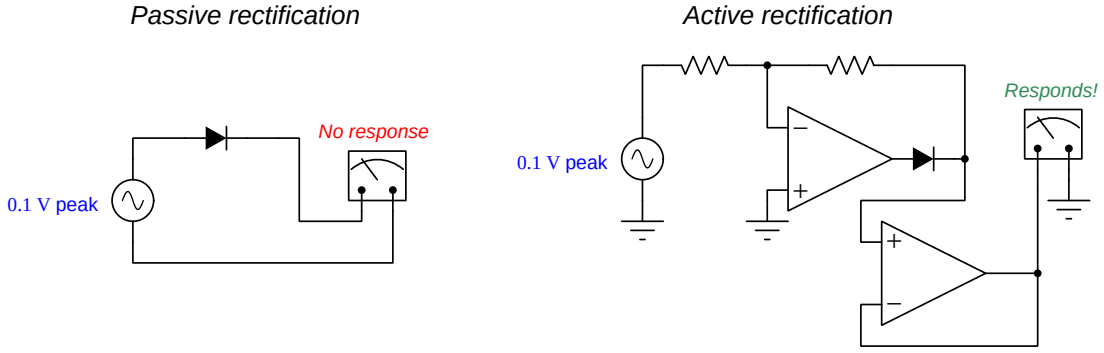
In the upper example circuit where the diode conducts, the output voltage is exactly what we would expect for an inverting amplifier with a gain of −1. The same is true for the middle example (0 Volts in and out). However, the lower circuit exhibits some unusual characteristics. First, note how the two inputs of the opamp are not equipotential with each other as we generally expect for any circuit having negative feedback. The reason for this is that there is actually no negative feedback at work in this case *because the diode is not conducting*: the opamp's fully-saturated-negative output signal reverse-biases the diode, which in turn causes it to act as an open which means the feedback loop is broken. Second, note how V_{out} becomes exactly equal to V_{in} with this broken feedback loop.

Thus, the circuit behaves in two fundamentally different ways depending on the polarity of V_{in}. With any negative value of V_{in} the circuit functions as a regular inverting amplifier, but with any positive value of V_{in} the opamp effectively loses control over anything happening in the rest of the circuit which means V_{in} gets transparently sent to V_{out}. The result is a circuit outputting the absolute value of the input signal ($V_{out} = |V_{in}|$). Another way of describing this circuit's behavior is to say it behaves as a *full-wave signal rectifier*.



Unlike an ordinary rectifier circuit comprised solely of one or more diodes, this precision rectifier does not suffer a forward voltage drop, but instead outputs the full magnitude of the input signal's absolute value.

This is an extremely useful property for precision analog measurement circuits such as AC voltmeters and AC ammeters, being able to rectify a small AC signal into a DC signal without losing any of that signal's strength from the forward voltage drop of any diode. Consider the following comparison of two AC voltmeter circuits, each one using a rectifier to convert the measured AC test voltage into DC that the permanent-magnet meter movement is able to display:



In the passively rectified voltmeter circuit (left) the meter movement is completely unresponsive to small AC voltages such as 0.1 Volt peak because that isn't enough to overcome the forward voltage potential of the rectifying diode. However, the actively rectified circuit (right) is able to drive the DC meter movement without trouble because the opamps provide[18] energy to overcome the forward voltage drop demanded by the diode as well as the energy demanded by the meter movement itself.

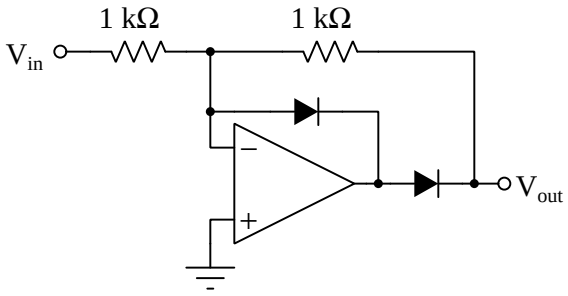
This is a remarkable and useful characteristic resulting from the addition of just a single component to an inverting opamp amplifier circuit! However, there is a problem with this circuit related to dynamic stability. For any positive V_{in} value the opamp is forced into a state of negative saturation. If the input signal oscillates or drifts between positive and negative values, the opamp's output must "snap" between a saturated condition ($+V_{in}$) and a controlled condition ($-V_{in}$). Opamp maximum slew rate becomes an issue here, as does the diode's *reverse recovery time*[19] and *junction capacitance*[20], which means the transition will not be seamless. The result will be "glitches" (i.e. transient voltage pulses) in the output signal at these transition points.

[18] It would be more accurate to say that the two opamps merely *direct* energy from the DC power supplied to them. However one wishes to say it, though, the point remains the same: the energy cost of using a semiconductor diode is paid for in the active circuit by a separate voltage source, not the AC source being tested.

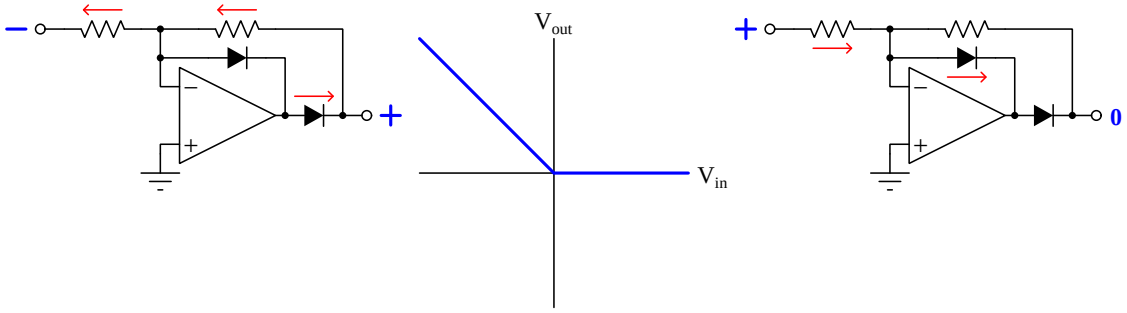
[19] This is the time required by the diode to switch from its conductive to its non-conductive state, based on the time required for charge carriers to be cleared out of the diode's depletion zone by the reverse-bias voltage.

[20] A reverse-biased diode's depletion region effectively acts as a small capacitance: two conductive regions sandwiching a non-conductive (depletion) region

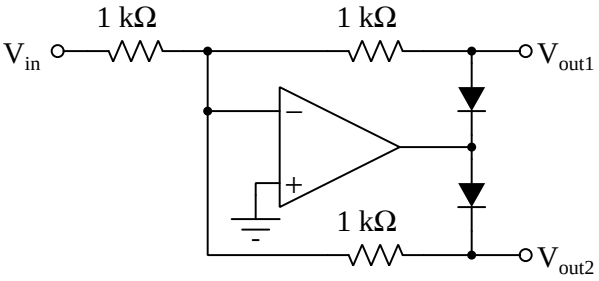
If we must have signal rectification but cannot tolerate "glitches" in the output signal, our only hope is to provide the opamp with a way to keep the feedback loop intact while the input signal changes polarity. If the negative feedback path from output to inverting input remains unbroken, the opamp's output will not saturate. One way to do this is to add another diode, this one bypassing the feedback resistor:



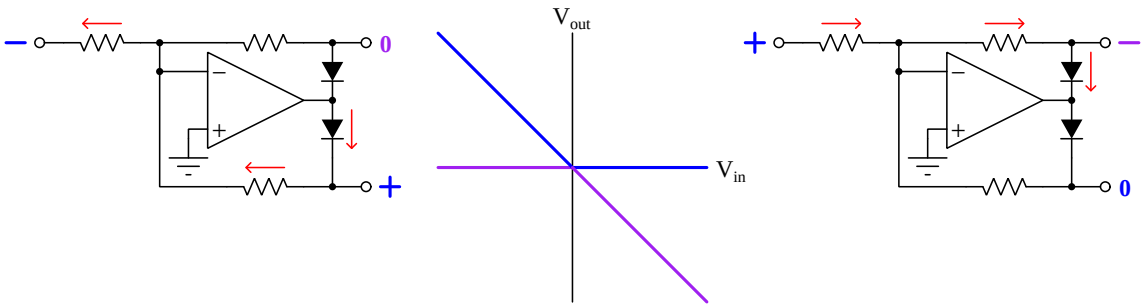
For negative V_{in} signal values the result is a positive V_{out} of the same magnitude, as before. For positive V_{in} signals, though, V_{out} simply remains at zero. Adding this second diode fixes the saturation problem, and consequently the transient "glitch" problem whenever the input signal goes from a positive value to a negative value, but it turns the circuit into a *half-wave signal rectifier*:



An interesting variation on this theme is to use two diodes to create two different negative feedback paths, each with its own feedback resistor. This preserves negative feedback for all input signal polarities while ensuring voltage gain at all times:

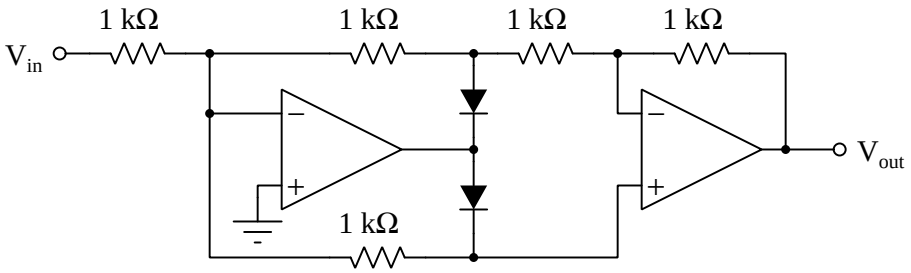


If we plot the characteristic of each output (in different colors) we see this circuit "separates" the signal polarities in the manner of two half-wave signal rectifiers:

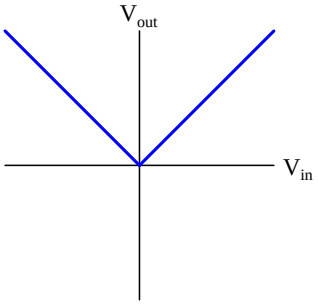


Positive V_{in} values are routed to the upper output terminal with inverted polarity while negative V_{in} values are inverted and routed to the lower output terminal, with each exhibiting a gain of $-\frac{R_f}{R_{in}}$. In each case negative feedback is preserved, with the opamp able to maintain its inverting input terminal equipotential with ground.

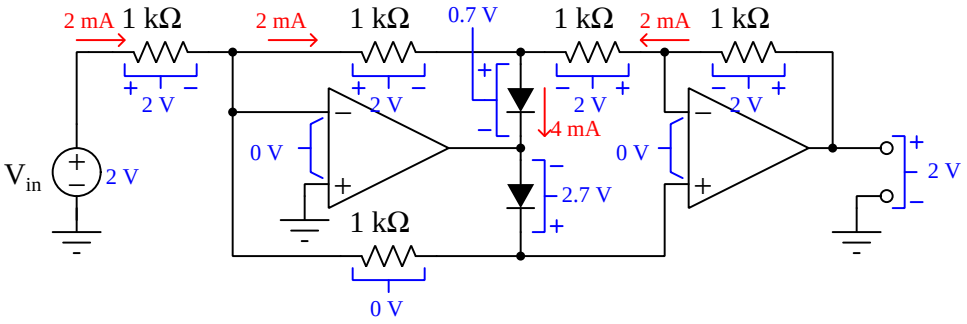
By adding another opamp stage we may combine these two signals into one for full-wave rectification:



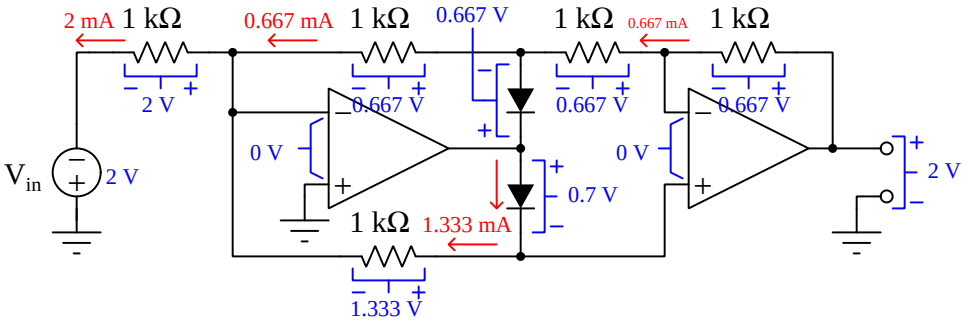
What used to be V_{out2} now connects to the non-inverting input of the final opamp, and since the signal polarity at that point (with respect to ground) was always positive it means it will drive the final output of this circuit positive. What used to be V_{out1} now connects to the input resistor of the final opamp, which from that perspective "looks" like an inverting amplifier. Since V_{out1}'s polarity was negative, this means it will also drive the final output positive. Thus, we have full-wave rectification but this time with no output "glitch" problem caused by the opamp having to drive its output into and out of saturation whenever V_{in}'s signal polarity goes positive and goes negative, respectively:



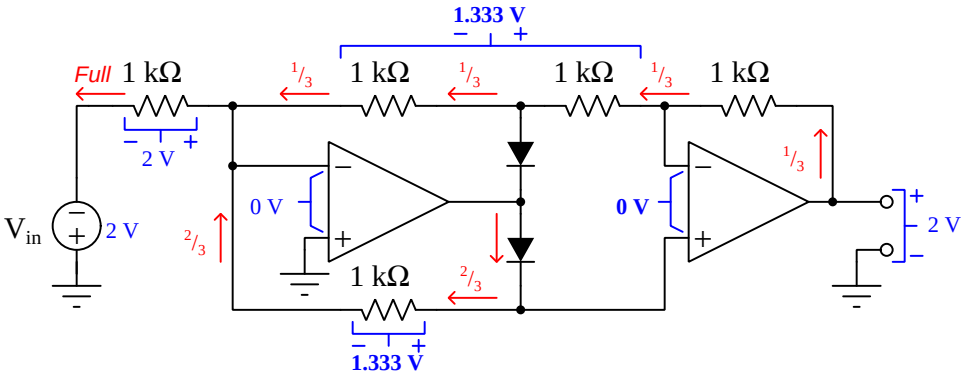
This full-wave precision rectifier circuit is rather complicated, so let's examine two "test case" examples. Beginning with $V_{in} = +2$ Volts, we see what is essentially a pair of inverting amplifiers (each with a voltage gain of -1) reproducing the +2 Volt input signal at the output:



In a case where $V_{in} = -2$ Volts, we see a significantly more complicated condition where the two opamps share the burden of providing 2 milliAmperes of feedback current to the input resistor. This burden is split between them at $\frac{1}{3}$ and $\frac{2}{3}$ ratios for reasons that may not be immediately apparent:



This next diagram shows why the two opamps work together to satisfy the input resistor's 2 milliAmpere current, and also why they perform that work unevenly. If we recognize that the second opamp must work to maintain its two input terminals in an equipotential condition, we see that the 1.333 Volts dropped by the single 1 kΩ resistor on the bottom must be dropped by the *two* 1 kΩ resistors above it[21]:



Therefore, however much current is carried by the bottom resistor, the two resistors above it must carry exactly half as much. This means two-thirds of the input resistor's current flows through the bottom resistor from the first opamp's output terminal while one-third flows through the upper resistors from the second opamp's output terminal.

A gentle reminder to the reader: it is well worth your time to take these circuit examples and perform your own "thought experiment" test-cases on them with different voltage and/or resistance values. In each case apply the foundational principles of operational amplifiers and of diodes, then apply Kirchhoff's and Ohm's Laws to solve for all voltages and currents. *See for yourself how these precision rectifier circuits work, rather than passively read the author's analysis.* This is the path of mastery.

[21] If this is not apparent, trace a Kirchhoff's Voltage Law loop around the path marked by the bottom resistor, the two resistors above it, and across the input terminals of the second opamp. If we trust that the second opamp does its job of maintaining those two input terminals at the same potential, we must conclude that voltage dropped by the bottom resistor must equal voltage dropped by the two resistors above it (in series). If each of those two upper resistors drops half as much voltage as the bottom resistor, and they all have the same resistance values, then current through the upper resistors must be one-half that of current through the bottom resistor by Ohm's Law.

3.18 Integrators and differentiators

Calculus is a branch of mathematics primarily concerned with two complementary concepts: *differentiation* and *integration*. Simply put, differentiation is the value of a variable's *rate-of-change* with relation to some other variable such as time. Simply put, integration is the value of a variable's accumulated product (multiplication) with some other variable such as time. These two operations – differentiation and integration – are inverse functions, in that one of them un-does the other, similar to addition and subtraction, or multiplication and vision, or powers and roots.

We see differentiation naturally at work in the characteristics of capacitance and inductance. Note the "Ohm's Law" equations relating voltage and current for each of these electrical functions:

$$I = C\frac{dV}{dt} \qquad\qquad V = L\frac{dI}{dt}$$

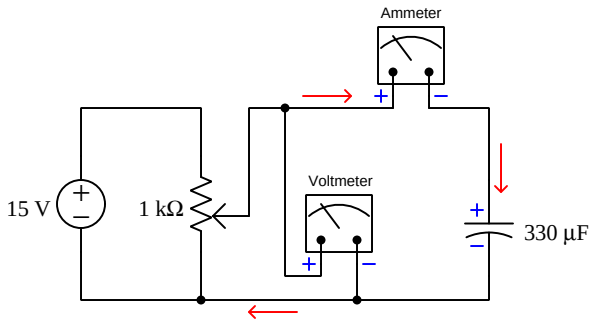
The "$\frac{d}{dt}$" expression in each of these formulae is the *derivative* of a specified variable in relation to time; $\frac{dV}{dt}$ is the rate-of-change of voltage over time, and $\frac{dI}{dt}$ is the rate-of-change of current over time. Thus, we can say that current "through" a capacitor is proportional to the time-derivative of voltage, and voltage dropped across an inductor is proportional to the time-derivative of current. Unlike resistors which follow the true Ohm's Law, where voltage and current are always directly proportional to each other ($V = IR$), with capacitors current is proportional to the *speed* at which voltage varies, and with inductors voltage is proportional to the *speed* at which current varies.

If we un-do these derivative expressions using integration ($\int$), we may write two new formulae for these components, each one solving for the other time-dependent variable:

$$V = \frac{1}{C}\int I\,dt + V_0 \qquad\qquad I = \frac{1}{L}\int V\,dt + I_0$$

Expressing these formulae in words, we may say that the amount of voltage accumulated by the capacitor is proportional to the amount of current "through" it as well as how long that current flows, plus whatever voltage value it started at time $t = 0$ (V_0); also, that the amount of current through an inductor is proportional to the amount of voltage impressed across it as well as how long that voltage was present, plus whatever current value was there at time $t = 0$ (I_0).

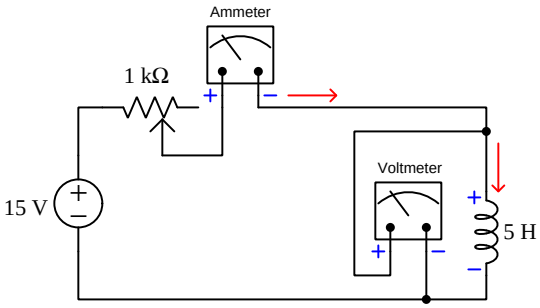
Here are some numerical examples to de-mystify the "derivative" expressions for capacitors and inductors:



If the potentiometer wiper moves steadily upward to make the voltmeter increase by 3 Volts every second, the ammeter will show 990 microAmperes of current

$$I = C\frac{dV}{dt}$$

$$I = (330\ \mu\text{F})\left(\frac{3\ \text{V}}{1\ \text{s}}\right) = 990\ \mu\text{A}$$



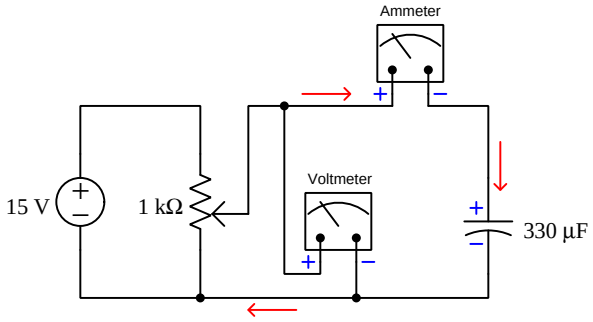
If the potentiometer wiper moves steadily left-ward to make the ammeter increase by 10 milliAmperes per second, the voltmeter will show 50 milliVolts of voltage

$$V = L\frac{dI}{dt}$$

$$V = (5\ \text{H})\left(\frac{10\ \text{mA}}{1\ \text{s}}\right) = 50\ \text{mV}$$

Of course, these examples using steady rates-of-change over time are given as simple illustrations for the calculus concept of differentiation. Those derivative expressions hold just as true for non-steady rates of change: the capacitor's current at any moment in time being equal to the capacitance times the voltage's rate-of-change at that instant; the inductor's voltage at any moment in time equal to the inductance times the current's instantaneous rate-of-change.

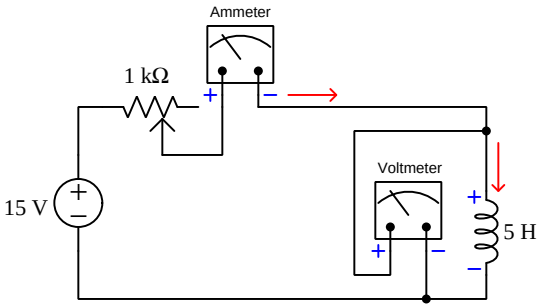
Here are some numerical examples to de-mystify the "integral" expressions for capacitors and inductors, using the same circuits show previously:



If the potentiometer wiper moves steadily upward to make the ammeter read a steady 150 microAmperes of current for 4 seconds, and the capacitor began in a pre-charged state of 3 Volts, its voltage at the end of those 4 seconds will be 4.8182 Volts

$$V = \frac{1}{C} \int I \, dt + V_0$$

$$V = \left[\frac{1}{330 \, \mu\text{F}} \int (150 \, \mu\text{A})(4 \text{ s}) \right] + 3 \text{ V} = 4.8182 \text{ V}$$



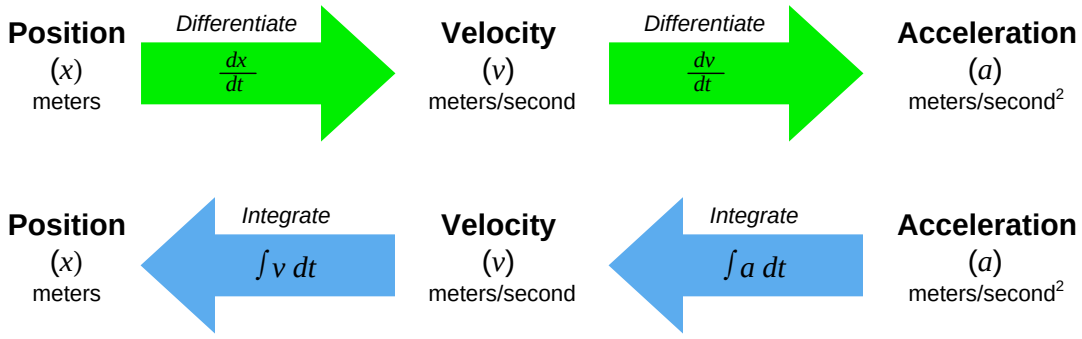
If the potentiometer wiper moves steadily left-ward to make the voltmeter read a steady 0.2 Volts for 8 seconds, and the inductor began in a pre-charged state of 5 milliAmperes, its current at the end of those 8 seconds will be 325 milliAmperes

$$I = \frac{1}{L} \int V \, dt + I_0$$

$$I = \left[\frac{1}{5 \text{ H}} \int (0.2 \text{ V})(8 \text{ s}) \right] + 5 \text{ mA} = 325 \text{ mA}$$

As before, these examples using steady applications of current and voltage are given as simple illustrations for the calculus concept of integration. Those integral expressions hold just as true even when the applied current or voltage is not steady over time: the capacitor's voltage rise over any time interval being equal to the accumulated product of current and time divided by the capacitance; the inductor's current rise over any interval equal to the product of voltage and time divided by the inductance.

A great many physical phenomena involve rates-of-change, especially rates of change over time, not just voltage and current for reactive components. The concepts of position, velocity, and acceleration, for example, are really expressions of differentiation and integration, as shown in this illustrative diagram:



Expressing the differentiation functions in words, we may say that velocity is the time-derivative of position (i.e. how rapidly an object's position changes over time) and that acceleration is the time-derivative of velocity (i.e. how rapidly an object's velocity changes over time). Expressing the integration functions in words, we may say that the amount of distance traversed by a moving object is the time-integral of its velocity (i.e. the accumulated product of velocity and time), and that the amount of velocity gained or lost by a moving object is the time-integral of acceleration (i.e. the accumulated product of acceleration and time).

Another application of these calculus concepts may be found in Isaac Newton's Second Law of Motion, which states that the acceleration of an object is proportional to the net force acting upon it, and inversely proportional to the object's mass (i.e. $F = ma$). Knowing that acceleration (a) is really just the time-derivative of velocity, we may re-write Newton's formula in terms of velocity (v) instead of acceleration:

$$F = m\frac{dv}{dt}$$

Expressing Newton's formula as an integral, we may solve for the velocity of an object given its mass (m), the force applied to it over time (F), and its initial velocity (v_0):

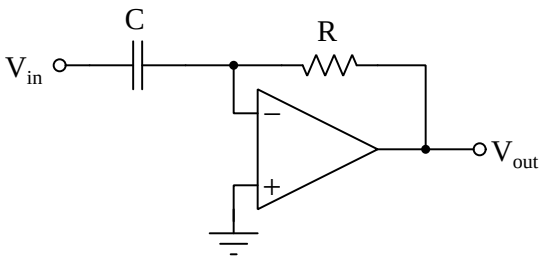
$$v = \frac{1}{m}\int F\,dt + v_0$$

As you can see, differentiation and integration apply to much more than just electrical quantities for capacitors and inductors. However, since we know that both capacitors and inductors naturally exhibit this time-derivative/integral characteristic, we may build operational amplifiers using these components to actually *perform* these calculus operations on analog signals representing other quantities (e.g. force, velocity). Just as we may use resistive operational amplifier networks to add and subtract analog signals, we may use capacitive[22] operational amplifier networks to integrate and differentiate analog signals for all kinds of different purposes.

[22] Technically, both capacitance and inductance are eligible in these circuits. However, capacitors are generally smaller, lighter weight, less expensive, and more "pure" in their reactive characteristics than inductors.

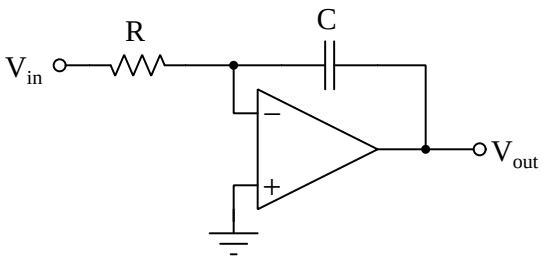
The following pair of opamp circuits show how differentiation and integration may be performed using a bare minimum of components. The differentiator generates an output signal voltage proportional to the rate-of-change over time of its input signal, while the integrator generates an output signal voltage proportional[23] to the time-product of its input signal.

Differentiator



$$V_{out} = -RC\frac{dV_{in}}{dt}$$

Integrator

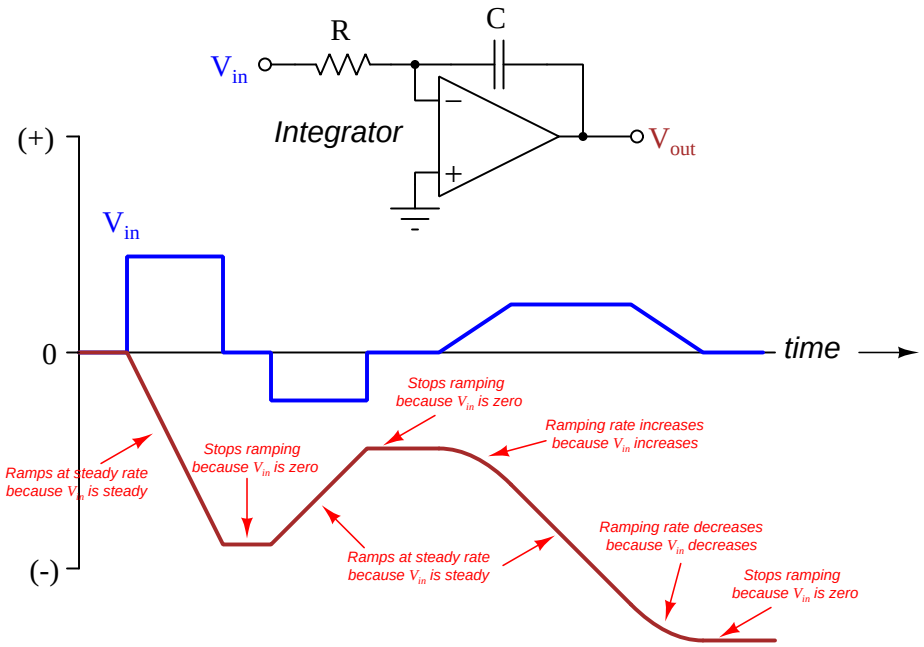


$$V_{out} = -\frac{1}{RC}\int V_{in}\,dt + V_0$$

What this means in a practical sense is that any constant (DC) voltage signal applied to the input of the integrator results in a *ramping* voltage signal at the output, the rate of rise or fall proportional to the magnitude of the input signal. The differentiator circuit does just the opposite: generating an output voltage signal proportional to how quickly its input signal ramps.

[23] The fact that a negative sign appears in each of these formulae is an artifact of the inverting nature of the operational amplifier circuit, with its virtual ground point at the inverting input terminal between the resistor and capacitor.

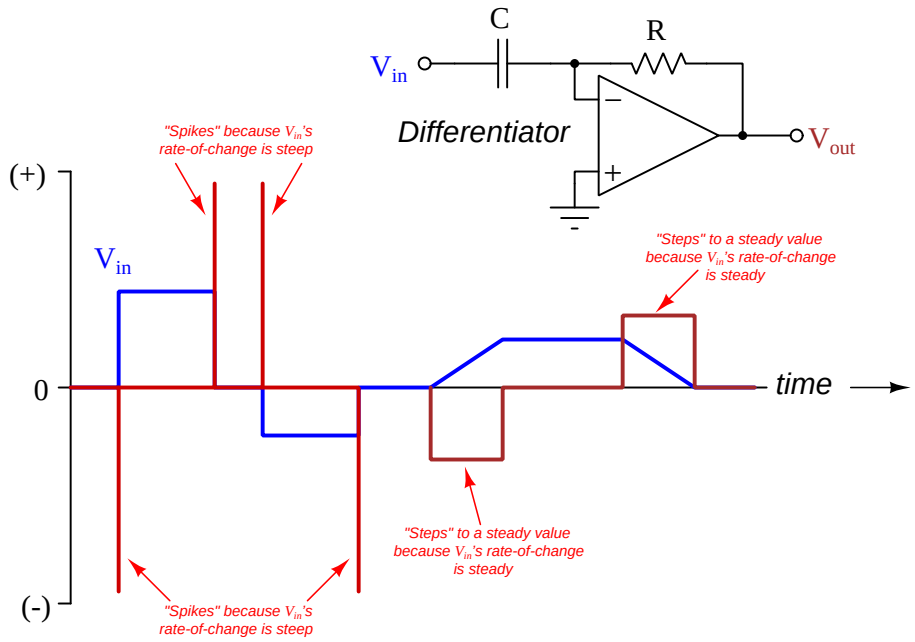
The concepts of integration and differentiation are easier to understand by "seeing it in action". Closely examine the following time-domain graph for an inverting integrator circuit to see the relationship between V_{in}'s magnitude and V_{out}'s rate-of-change ($\frac{dV_{out}}{dt}$). Note that the inverting nature of this circuit configuration explains why V_{out} ramps *negatively* whenever V_{in} is *positive*:



This circuit's time constant ($\tau = RC$) sets the proportionality between the output's rate of change over time versus the input signal's strength. If we wished to have an output signal that ramped slower for a given input voltage, we could choose a larger resistor and/or capacitor.

An integrator circuit could be used as an *analog computer*, integrating the input signal over time. A practical application might be calculating how far a vehicle has traveled (output signal) from a speed signal (input). Again, you are encouraged to look closely at the graph to see for yourself how this might work: imagine V_{in} is a signal from a vehicle's speed sensor (i.e. speedometer), while the integrator's V_{out} signal represents distance traveled (i.e. odometer). In fact, this very mathematical process is used in *inertial guidance systems*, where input signals are taken from *accelerometer* sensors and integrated to yield *velocity* signals, which in turn may be integrated again to obtain *position* signals. Thus, an aircraft or spacecraft may estimate its velocity and position in space from nothing but acceleration data.

Similarly, let us examine a time-domain graph showing V_{in} and V_{out} for an inverting differentiator circuit. Just for contrast, we will use the exact same V_{in} signal profile used in the previous (integrator) example:



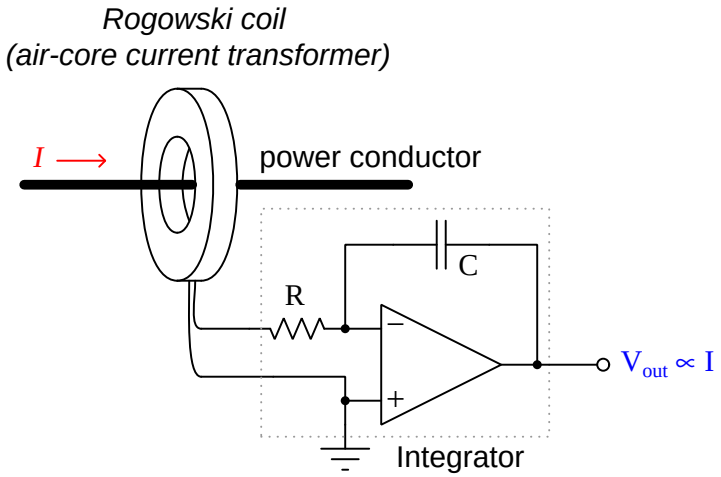
Just as with the integrator, this circuit's time constant ($\tau = RC$) sets the proportionality between the output's rate of change over time versus the input signal's strength. If we wished to have an output signal that responded less aggressively to input rates-of-change, we could choose a smaller resistor and/or capacitor.

Like integrator circuits, differentiators are also useful as analog computers. A practical application would be to sample the voltage signal from a robot's arm position sensor (input) to generate a signal representing that arm's velocity (output). Imagine the arm at a standstill (in any position), resulting in zero output voltage from the differentiator (i.e. zero velocity). However, once the arm begins to move the differentiator circuit will output a voltage proportional to the arm's speed of motion.

It is not difficult to conceive of other applications for integrator and differentiator circuits. Consider the following:

- **Passive infra-red (PIR) motion sensing** – a PIR sensor produces a weak voltage signal when struck by infra-red light emitted by warm objects. However, to sense *motion* we need a circuit detecting the rate-of-change of the measured signal (temperature) rather than the actual value of that temperature. We don't want a motion sensor to falsely trigger if an object in its view happens to be too warm, but we do want the sensor to trigger if it detects a *change* in temperature, signifying a warm object either entering or leaving its field of view. Thus, we may connect an infra-red sensor to the input of a differentiator, comparing the output of that differentiator against some reference voltage to trigger a comparator to signal when the rate-of-change of temperature is rapid enough to indicate a moving object.

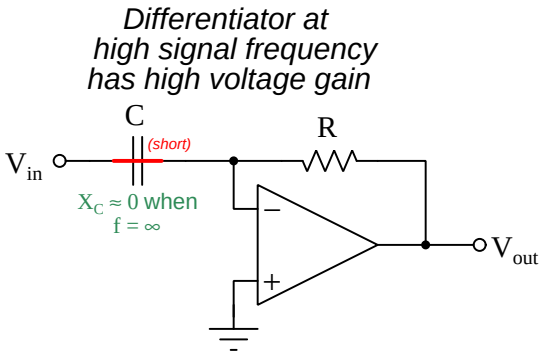
- **Rogowski coil signal conditioning** – a *Rogowski coil* is an air-core transformer used to magnetically sense current through a conductor. The current-carrying conductor passes through the center of the Rogowski coil, serving as the primary winding to the coil's secondary winding. The Rogowski coil's induced voltage is proportional to the rate-of-change of the current-carrying wire's magnetic field according to Faraday's Law of Electromagnetic Induction ($V = N\frac{d\Phi}{dt}$) which means the coil's voltage signal does not directly reflect current through the conductor, but only the time-derivative of that current. However, by integrating the Rogowski coil's voltage signal we "un-do" the Rogowski coil's natural derivative characteristic and end up with an analog signal proportional to the actual amount of current passing through the current-carrying conductor.

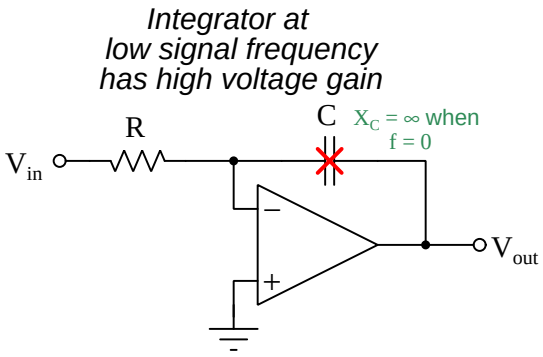


- **Triangle-wave oscillator** – If a square-wave signal is presented to the input of an integrator circuit, the integrator's output will be a triangle wave. Every time the square wave signal goes positive, the integrator's output ramps linearly in one direction; every time the square wave goes negative, the integrator's output ramps linearly in the other direction. If that triangle-wave signal connects to the input of a comparator with positive feedback (hysteresis), it will generate the necessary square-wave signal to perpetuate the cycle.

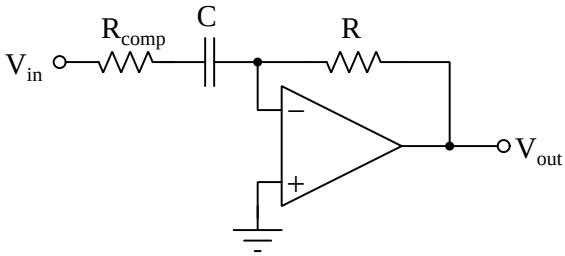
Analog differentiation and integration are not without certain caveats. With differentiation the major challenge is *noise*, as the rapid rise- and fall-times of any high-frequency noise in the input signal will become amplified by the differentiator. It isn't so much the amplitude of the noise that is problematic, as much as it is the *rates of change* which tend to be fast. Thus, any analog signal must be filtered of noise before it may be reliably differentiated. With integration the major challenge is *drift*, as any DC bias present at the input will cause the integrator to continually ramp its output until it reaches saturation. Even if the input signal is perfectly centered around 0 Volts, tiny bias currents present at the opamp input terminals will cause the capacitor to gradually accumulate a voltage that will eventually cause the integrator's output to saturate to its maximum limit.

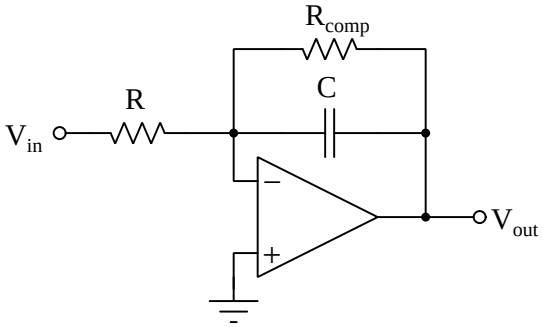
A practical problem facing the analog integrator and differentiator circuits we have explored so far is that their voltage gains approach infinity at limiting-case signal frequency values. Consider an integrator sensing an input signal with a frequency of 0 (i.e. a DC voltage signal), as well as a differentiator sensing an input signal of infinite frequency:





Extremely high gain means the operational amplifier's output is practically guaranteed to saturate even with small input signal levels. One solution to this problem is to add *compensation resistors* to the two circuits in order to provide a high (but not infinite) voltage gain under these conditions:



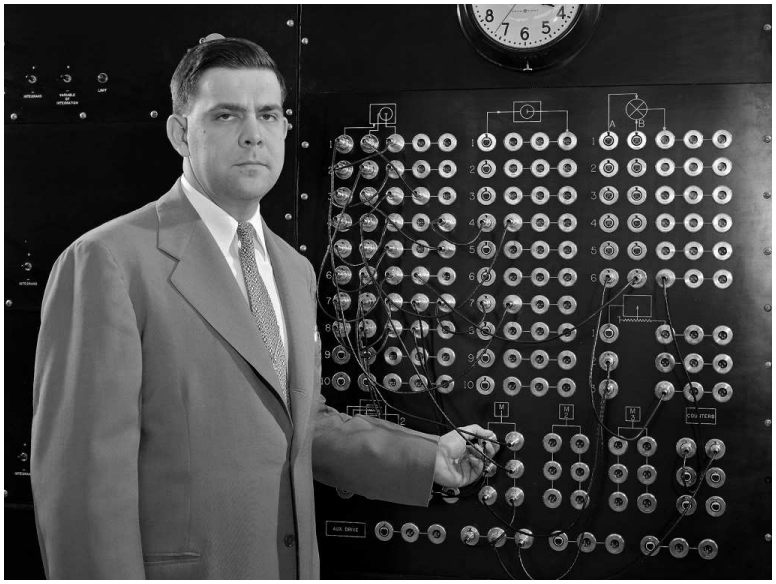


3.19 Analog computing

Operational amplifiers are so named because they were originally intended to perform *mathematical operations* using voltage signals as variables. In other words, opamps were invented to perform mathematical calculations in the analog domain where numerical values were represented as electrical voltages and/or currents. We have seen this in previous sections already, where opamps combined with resistors may form networks capable of adding, subtracting, multiplying, and dividing voltage signals. Combined with capacitors, opamps may even perform basic calculus functions of time-differentiation and time-integration. These capabilities made opamp circuits well suited for the computation of physical simulation problems such as projectile tracking, machine vibration simulations, aircraft and spacecraft flight simulations, chemical reaction simulations, etc.

Like so many other inventions born of necessity, the opamp owes much of its existence to the priority of combat. During World War Two researchers at Bell Labs in the United States developed high-gain vacuum-tube amplifier circuits for use in what later became the *M9 Gun Director*, an analog computer[24] used to automatically compute firing solutions for 90 millimeter anti-aircraft guns based on inputs from a radar unit as well as other variables such as air density and wind speed. Previous gun directors were mechanical in nature, unable to operate with precision necessary for reliable operation on the battlefield.

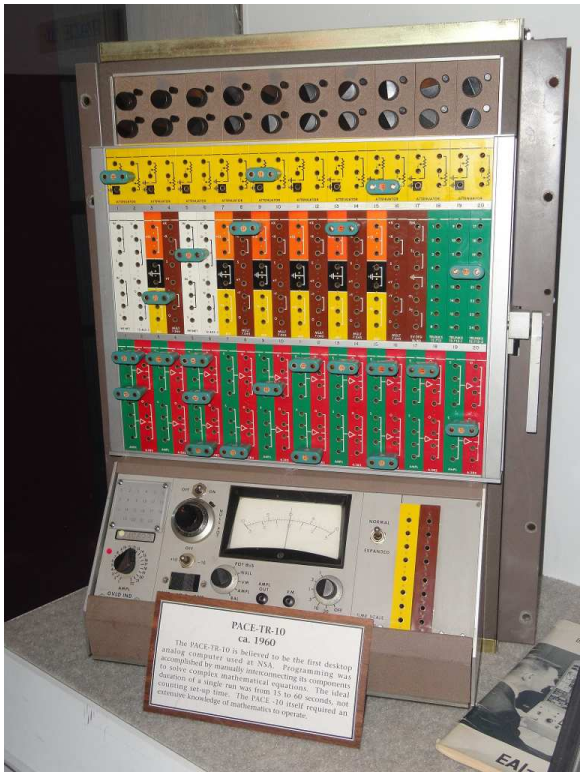
After that war, industrial and scientific applications for operational amplifiers flourished, with opamps being put to use within analog computer systems for automotive and aerospace vehicle development, industrial process simulation and control, and even some areas of mathematical research. The following photograph, courtesy of NASA, shows Harry Mergler standing before the patch-board of an analog electronic computer he helped develop, at that time called a *differential analyzer* because of its facility solving differential equations:



[24] In the *Bell Laboratories Record* journal article written in December of 1943 they described this electronic system as a *computor*.

Analog computers consisted of a collection of operational amplifier circuits complete with precision resistor and capacitor networks ready to interconnect in order to simulate mathematical equations. Programming such an analog computer consisted of converting equation coefficients into resistance and capacitance values, then forming connections between all those components using "patch cables" (i.e. jumper wires) to form a complete mathematical simulation. Input values often came from potentiometers wired as voltage dividers, so that twisting the knob of a potentiometer resulted in a new voltage value that could be coupled to some portion of an opamp circuit. Output values were also voltages, measurable using voltmeters and/or oscilloscopes.

A smaller analog computer known as the model PACE-TR-10 manufactured by Electronic Associates, Incorporated may be seen in the following photograph[25], with potentiometer and opamp symbols clearly visible on the patch panel. This particular analog computer was employed by the United States government in cryptological studies:



Knobs to control the potentiometers on this analog computer are visible on the top of its panel. Patch connections to those potentiometers are arranged in the yellow-colored area of that panel just below the knobs. Opamp terminal connections appear in the red/green colored area of the panel. The turquoise-colored plugs shown inserted into many of those jacks are actually precision resistors, their resistance values appearing in fine print on the front of each plug.

[25] Public-domain courtesy of Daderot, taken at the National Cryptologic Museum in March of 2013.

By the 1960's analog computers became reliable and affordable enough for personal and educational use. One example of an analog computer manufactured by Heathkit in 1960 and selling for around $400 US dollars is the model EC-1, shown in the following photograph[26]:

The fact that most practical analog computing applications revolve around differential equations meant that users of these machines required fairly advanced knowledge of calculus, and so these were never as popular for non-specialists as later digital computer technology would become. Still, a brief survey of the manual for this early educational-grade device shows opamp circuit configurations that readers of this Tutorial would feel quite comfortable interpreting: summers, subtractors, multipliers, integrators.

[26] Public-domain image courtesy of Michael Holley, taken at the Computer History Museum in Mountain View, California on 4 November 2007.

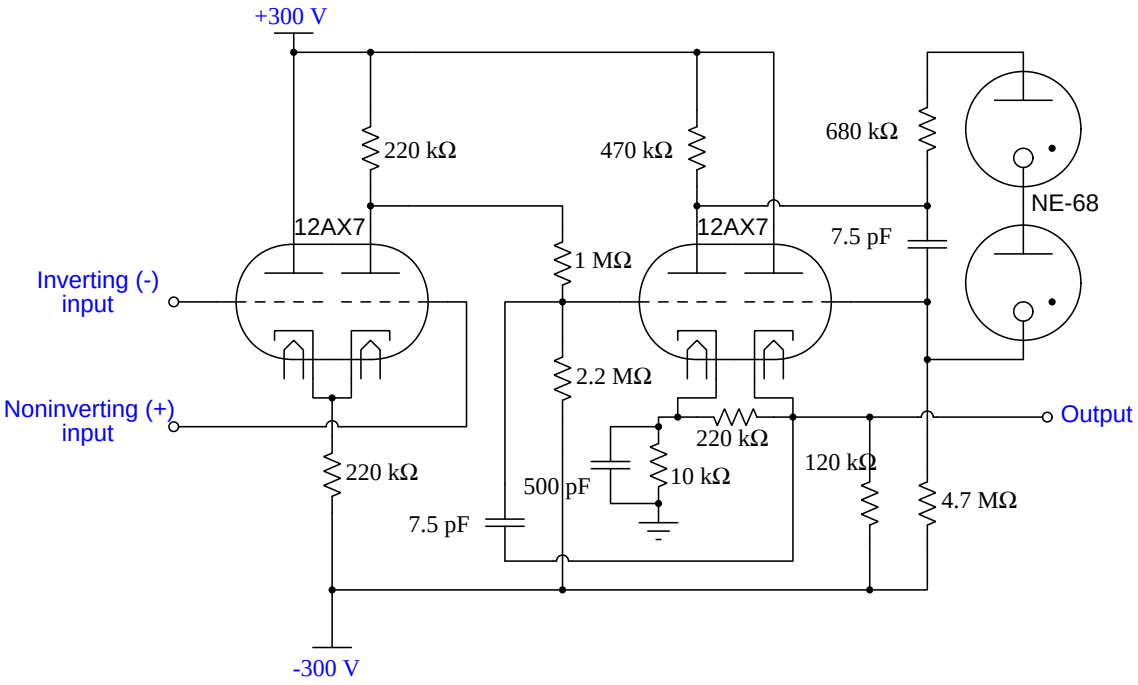
A rather large rack-mounted collection of opamps, function generators, miniature oscilloscopes, power supplies, and other components necessary for analog computing is shown in the following photograph[27]:

This particular rack and all its associated analog modules was manufactured by Tektronix, a major manufacturer of oscilloscopes and other electronic test equipment. Each model AM-501 opamp module was constructed of solid-state components (e.g. transistors), boasting a gain-bandwidth product of 5 MHz and an output range of ± 40 Volts DC at upwards of ± 50 milliAmperes DC.

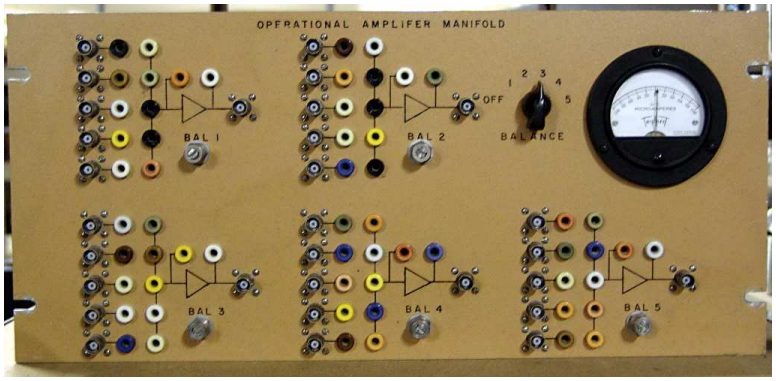
[27] This photo courtesy of Bob who hosts the website `dvq.com` containing photographs of legacy computers he has collected over the years.

One of the early researchers involved with early military applications of analog electronic computing was George Philbrick. After the end of World War 2, Philbrick formed a company called George A. Philbrick Researches Incorporated (GAP/R) which manufactured operational amplifiers and analog computing systems for decades thereafter, spurring much technological development of the operational amplifier and incentivizing its commercial applications with excellent product design and outstanding technical documentation. One of the flagship products GAP/R produced was the model K2-W opamp which used a pair of model 12AX7 dual-triode vacuum tubes for amplification and fit into a single 8-pin ("octal") socket. A schematic diagram of the K2-W opamp appears below:

Philbrick Researches model K2-W operational amplifier



Photographs[28] of a GAP/R analog computing "manifold" are shown here, the front patch-panel being where a person would plug either banana-jack-style jumper cables and/or BNC-style coaxial cables to form the interconnections between the unit's five opamps and other (external) components. The collection of components and their connecting wires formed the analog equivalent of a digital computer's *stored program* for solving some mathematical problem or simulating some physical system:



A rear view of this same unit shows five K2-W vacuum-tube opamps plugged into sockets for easy replacement in the event of failure. The two K2-W units on the right have missing vacuum tubes, those tubes also plugging into sockets in the K2-W base for easy replacement:

Note how the entire unit itself has a five-pin connector on its back, allowing it to be easily plugged into a larger analog computing system. This "manifold" only contained the five opamps and their associated power supply and balancing circuitry, making it incomplete as an analog computer.

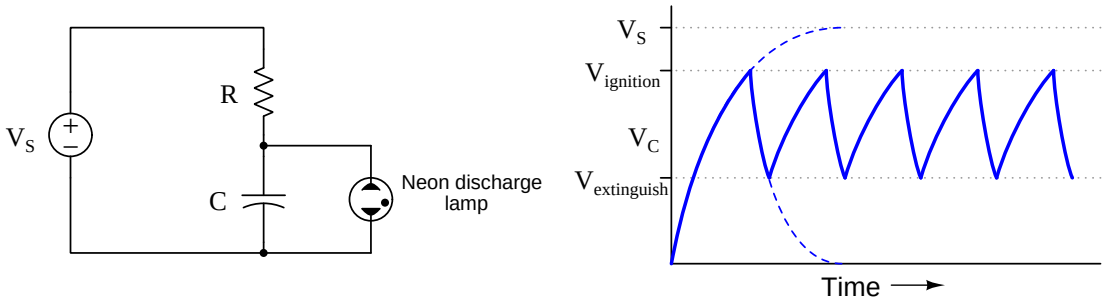
[28] These photographs courtesy of Bob who hosts the website `dvq.com` containing photographs of legacy computers he has collected over the years.

3.20 Oscillator circuits

Operational amplifiers may be used to fabricate different types of oscillator circuits. This section will explore some of the different types of opamp- and comparator-based oscillator circuits and how they work.

3.20.1 Relaxation oscillators

A *relaxation oscillator* is any circuit relying on the unregulated charging and/or discharging of a capacitor in conjunction with a hysteretic element to generate a sawtooth waveform. An elementary form of relaxation oscillator may be made from nothing but a capacitor, resistor, neon discharge tube, and of course a DC voltage source:



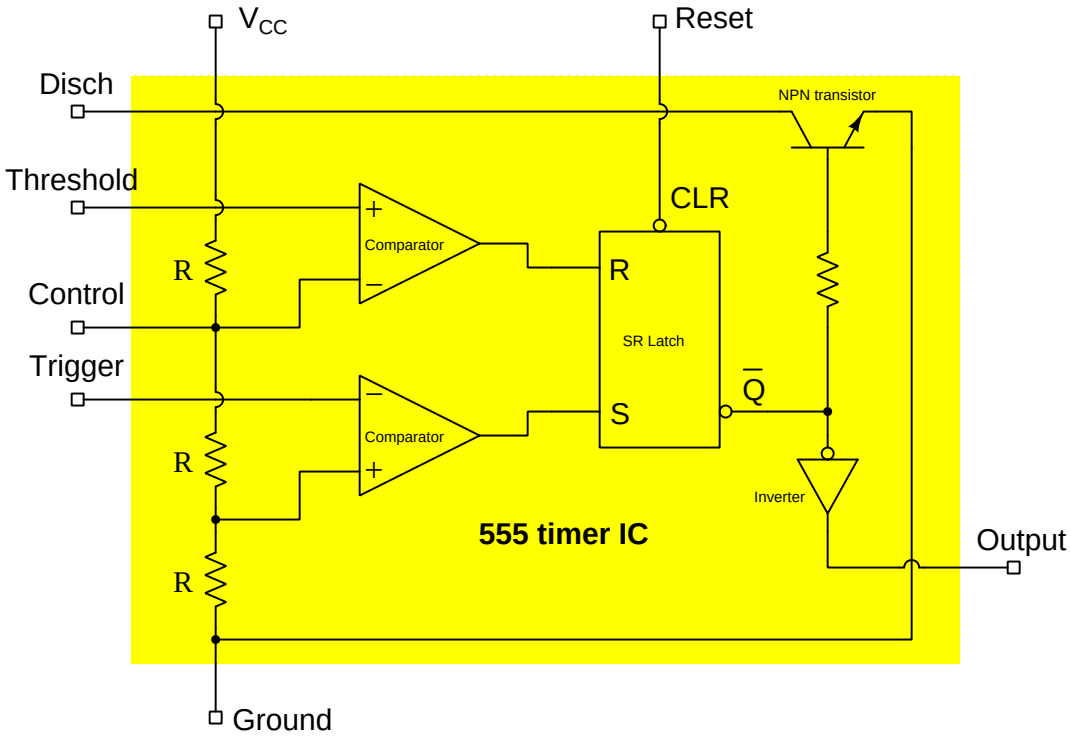
The "resting" state of a neon discharge tube is non-conducting, and so when this circuit is first energized the capacitor begins to absorb energy which results in a rising voltage across its terminals. The neon gas inside the discharge tube, however, ionizes and becomes electrically conductive at an applied potential of approximately 70 Volts. So, when the capacitor's voltage reaches this level the tube begins to conduct current, and will remain conductive until a much lower voltage is reached.

Hysteresis is that property of any component or system which fails to return to its original state after the removal of stimulus. Neon lamps are hysteretic, in that they remain non-conducting as the applied voltage rises to the point of ignition, and then once ionized they remain conductive until the applied voltage falls down to the extinguishing value. This is why the relaxation oscillator circuit requires a *gas discharge* lamp rather than something more conventional such as an incandescent lamp or LED[29].

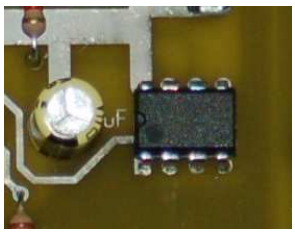
Operational amplifiers (and their closely-related electronic cousins, *comparators*) are able to compare voltage signals against each other, and as such they may be used to monitor voltage across a charging or discharging capacitor as part of an electronic relaxation oscillator circuit. Hysteresis – necessary for any relaxation oscillator's function – may be provided by a latching circuit driven by these opamp/comparator units or be made intrinsic to the opamp or comparator through the use of positive feedback.

[29] Even though both incandescent lamps and LEDs are nonlinear devices (i.e. they do not obey Ohm's Law), neither one is hysteretic. If one were to exchange the neon lamp for an incandescent lamp or LED in the oscillator circuit, the circuit would never oscillate. Instead, the capacitor voltage would stabilize at some constant value allowing the lamp to output some amount of steady light.

The venerable model 555 timer integrated circuit, for example, uses two comparators driving the inputs of a set-reset (SR) latch to control the charging and discharging of a capacitor:

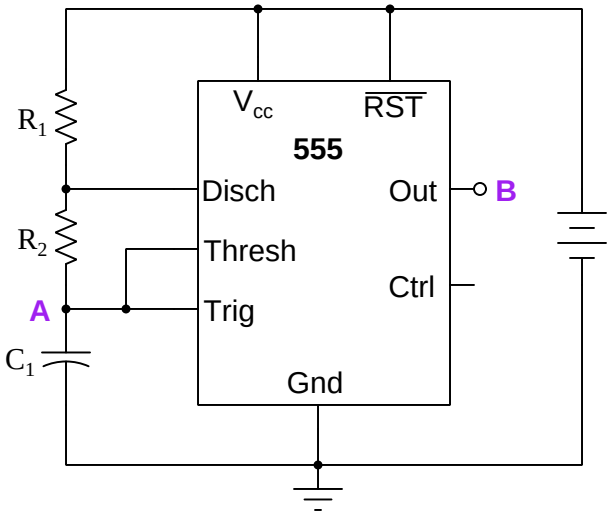


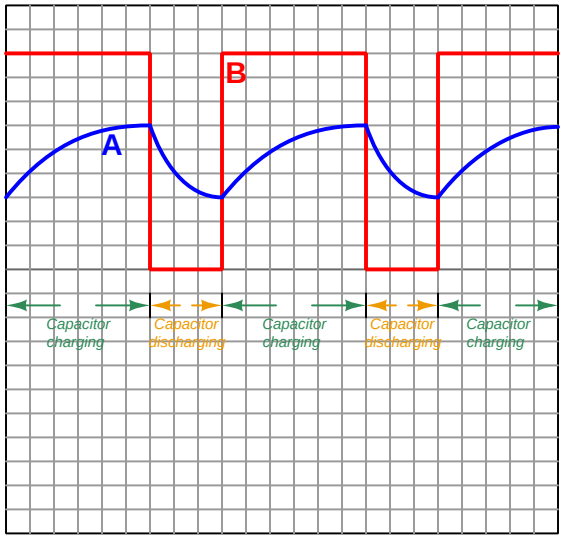
These integrated circuits are packaged as 8-terminal ICs, such as the 8-pin DIP[30] shown here:



[30] DIP = "Dual Inline Package" which is the most common arrangement of through-hole IC packages.

The standard "astable" configuration for the 555 timer uses two resistors and one capacitor as shown here, generating waveforms shown in the oscillograph illustration at test points **A** and **B**:



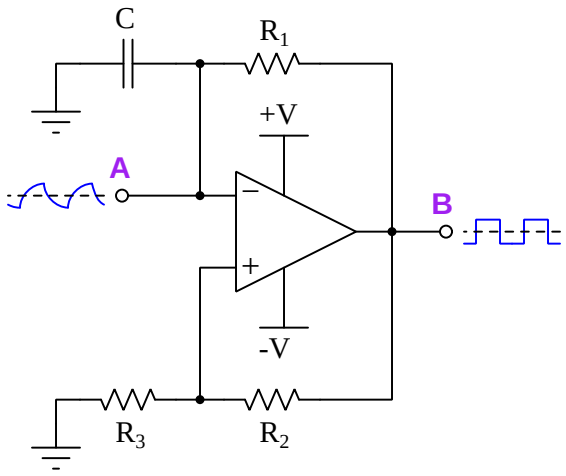


$$t_{charge} = 0.693(R_1 + R_2)C \qquad t_{discharge} = 0.693R_2C$$

$$t_{period} = 0.693(R_1 + R_2)C + 0.693R_2C = 0.693(R_1 + 2R_2)C$$

$$f = \frac{1.44}{(R_1 + 2R_2)C}$$

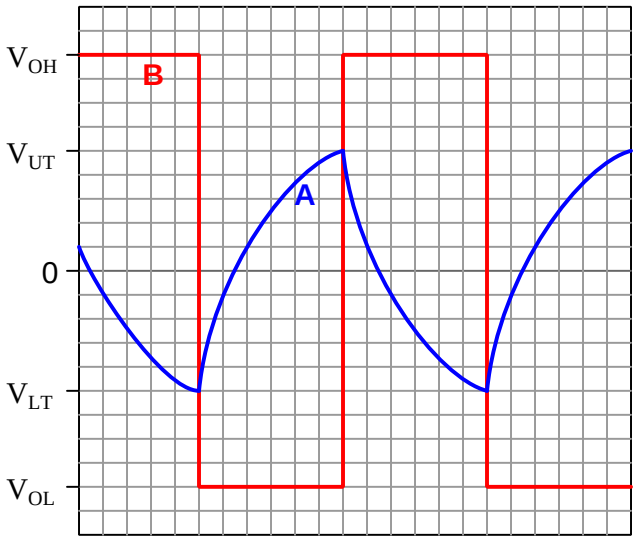
Here we see a comparator with positive feedback used to form a relaxation oscillator with a symmetrical (50%) duty cycle:



This circuit's positive feedback consists of the resistive voltage divider coupling the output signal to the non-inverting input, causing the comparator to have *two* threshold voltages, where V_{OH} is the comparator's "high" output voltage and V_{OL} is its "low" output voltage. The capacitor's voltage peaks at these two threshold voltage levels:

$$V_{UT} = (V_{OH})\frac{R_3}{R_2 + R_3}$$

$$V_{LT} = (V_{OL})\frac{R_3}{R_2 + R_3}$$

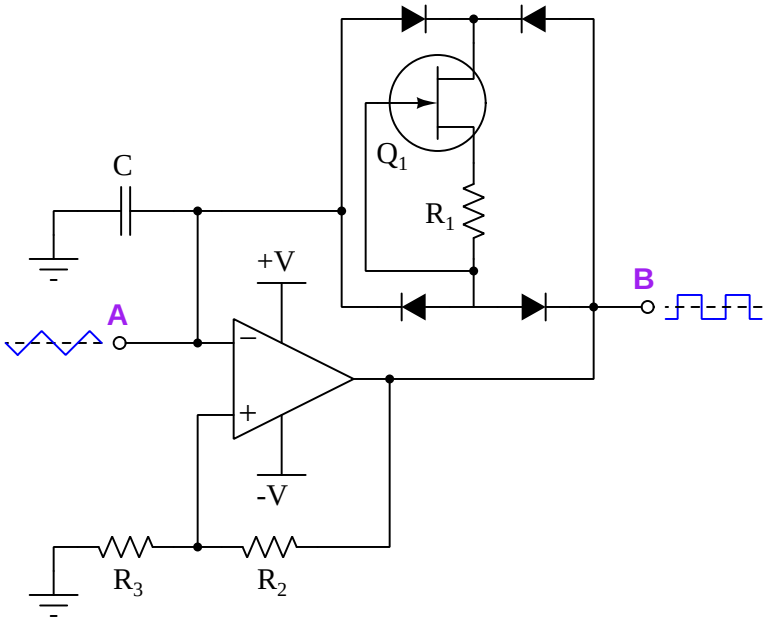


3.20.2 Integrating oscillators

If we somehow regulate the capacitor's current to some constant value, we may coax a triangle wave out of the relaxation oscillator based on the "Ohm's Law" equation for a capacitor, where capacitor current is proportional to the rate-of-change (i.e. the "speed") of its voltage:

$$I = C\frac{dV}{dt}$$

One simple way of doing this is to use a junction field-effect transistor[31] instead of just a current-limiting resistor as before. Since this combination of resistor and JFET is unidirectional, we must use a bridge rectifier network to "steer" the capacitor's alternating current the same direction through the JFET:

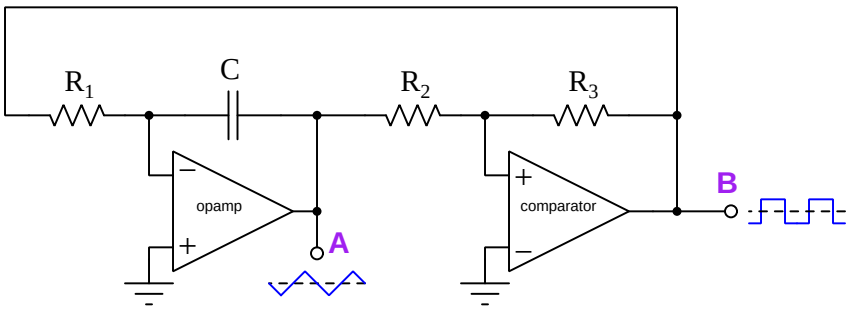


As we saw with our brief study of calculus earlier in this tutorial, the amount of voltage across a capacitor as it accumulates or releases energy over time is proportional to the *time-integral* of current through it. For example, if I is constant, then voltage will rise (or fall) the same amount over equal intervals of time. The value to which a capacitor's voltage will rise or fall depends on how much current has passed through it, and for how long of time, as well as the capacitor's initial voltage (V_0):

$$V = \frac{1}{C}\int I\,dt + V_0$$

[31] A more sophisticated solution, but just as simple to implement, is to use a "floating" three-terminal current regulator IC such as the LM334. With just a couple more components (another resistor and a diode) the LM334 is able to achieve a zero temperature coefficient which means it will regulate current at a constant value despite changes in operating temperature. Constant current means a constant slope (i.e. a constant $\frac{dV}{dt}$ rate for the capacitor's voltage) for our triangle-wave signal.

In the spirit of exploiting the capabilities of operational amplifiers, though, we may add another differential amplifier to the circuit to perform the task of current regulation:



As the comparator's output voltage alternately saturates to V_{OH} and V_{OL}, R_1 experiences that same voltage as the opamp uses negative feedback to maintain a virtual ground between R_1 and C. This constant voltage (either V_{OH} or V_{OL} depending on the comparator's output state) causes a constant current equal to $\frac{V_{OH}}{R_1}$ or $\frac{V_{OL}}{R_1}$ to pass through the capacitor. This in turn causes the opamp's output voltage to ramp linearly up and down in a nice "triangle" wave fashion as the capacitor alternately charges in one polarity and then the other, the rate of that ramping equal to $\frac{dV}{dt} = \frac{I}{C}$. Resistors R_2 and R_3 form a voltage divider that causes the comparator to switch output states (from "high" to "low" or vice-versa) whenever its noninverting input potential reaches zero, which is what will happen when the capacitor's voltage builds up to a sufficient value to "counterbalance" the comparator's output potential at the opposite polarity.

Any practical circuit of this design must use an R_3 resistance greater than R_2 in order to limit how high and how low the opamp's output voltage must swing in order to prompt the comparator to switch states. If $R_2 = R_3$, then the opamp's output would actually have to swing slightly higher and slightly lower than the comparator's output voltage in order to get the comparator to switch states, which is unlikely if both are powered by the same split-DC power supply rails. If, on the other hand, we were to choose $R_3 = 2R_2$, the opamp's output would only need to swing to peak values one-half[32] that of the comparator's V_{OH} and V_{OL}.

Anyone wishing to analyze this circuit is encouraged to do so with the following starting points:

- Assume the capacitor begins in a completely de-energized state (i.e. zero Volts) when the circuit is first powered up
- Since the comparator uses positive feedback, assume its output will always be either +V or −V and never any value in-between
- Arbitrarily choose to have the comparator output either full +V or full −V upon power-up

[32] Consider a case where R_3 is 10000 Ohms and R_2 is 5000 Ohms. If the supply rails are ±5 Volts, then the opamp's output voltage would only need to ramp down to −2.5 Volts in order to counter the comparator's +5 Volt "high" state output and achieve zero at the comparator's non-inverting input, after which the comparator would switch states to output −5 Volts and prompt the opamp to ramp up to +2.5 Volts over time, after which the comparator would switch states again and the cycle would continue.

3.20.3 Phase shift oscillators

All oscillator circuits rely on some form of *positive feedback* to function. For relaxation and integrating oscillators this positive feedback takes the form of a *hysteretic* element that tends to latch on once turned on and latch off once turned off. However, if we use analog rather than discrete (on/off) positive feedback in an amplifier circuit, we may create oscillators capable of outputting sinusoidal waves rather than square, sawtooth, and triangle waves.

If you have ever experienced the "howling" or "shrieking" sound of a public-address amplifier system whenever the microphone gets too close to the loudspeaker, you have heard positive-feedback oscillation in action. Oscillations may occur in any amplifier system if the following criteria[33] are met:

- The total phase-shift of the amplifier and feedback network must be 360^o or some whole multiple thereof
- The total gain of the amplifier and feedback network must be at least unity (1, or 0 dB)

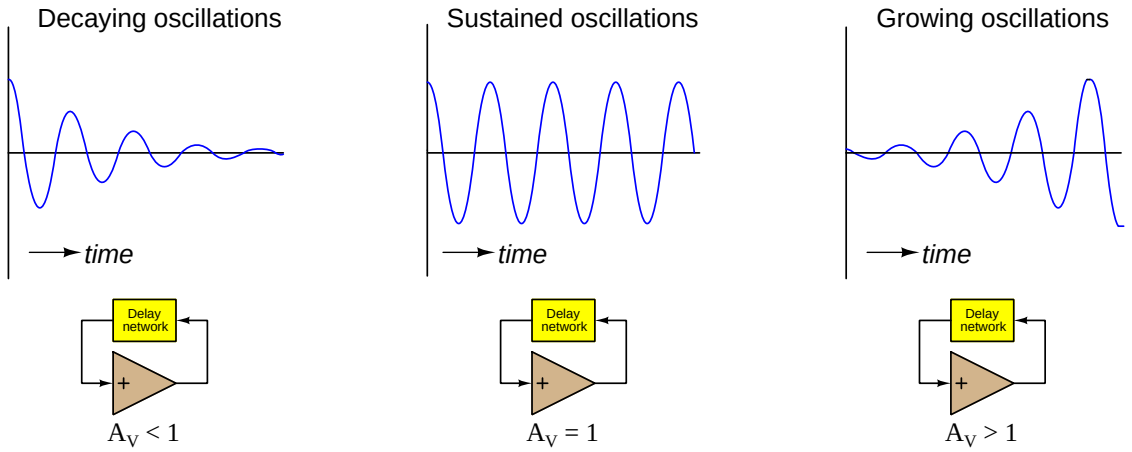
These criteria make sense once we grasp why positive feedback works to produce oscillations. In a positive-feedback system, any disturbance of the amplifier's output signal gets fed back to its input in such a way as to reinforce that original disturbance. If this positive feedback is instantaneous, the amplifier's output will simply saturate to one of its power supply rails (limits) and remain there. However, if the feedback is delayed by a full 360^o, it means the positive-feedback reinforcement will continue along all points of the waveform to reproduce that waveform in its entirety[34]. As the signal's voltage rises, the amplifier will act to make it rise even more; once the delayed signal begins to fall, the amplifier will act to make it fall faster.

Negative feedback also plays a role in these oscillator circuits by allowing us to control the voltage gain of the circuit to the desired ratio of unity (1).

[33] These conditions are often referred to as the *Barkhausen criterion* in honor of German physicist Georg Barkhausen.

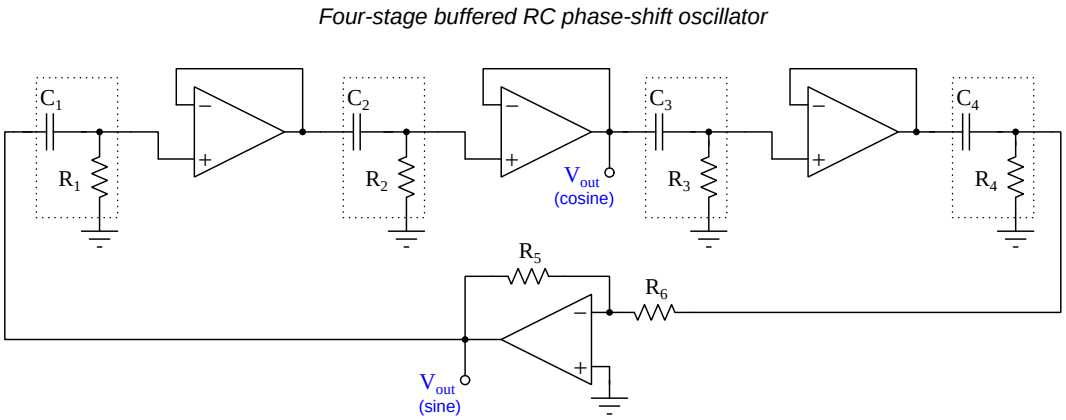
[34] This scenario evokes the image of a dog chasing its own tail: the fact that the dog must reach *back* to its own tail with its mouth leads it to spin in circles, that reach being the necessary "delay" to sustain a circular trajectory. A dog would never chase a stationary object lying straight ahead.

This cyclic process continues as long as the amplifier's gain is large enough to overcome all attenuation in the feedback-delay network. If the gain is not quite high enough, the oscillations will decay with every cycle until they die out completely. If the gain is exactly the right amount, the oscillations will continue with steady AC amplitude. If the gain is greater than necessary, the oscillations will grow in amplitude until the amplifier begins to "clip" and distort the wave-shape.



Self-oscillation by amplified positive feedback can be a useful phenomenon, or a problem to be avoided. If your design goal is to make a *stable* amplifier, then you want to avoid the conditions that will lead to self-oscillation. If your design goal is to make a sinusoidal oscillator from an amplifier circuit, then you want to create the necessary conditions that will cause that amplifier to continuously oscillate. The latter is our goal for the purposes of this section.

A rather "brute-force" method of making an opamp oscillator using phase-shift techniques is shown here:



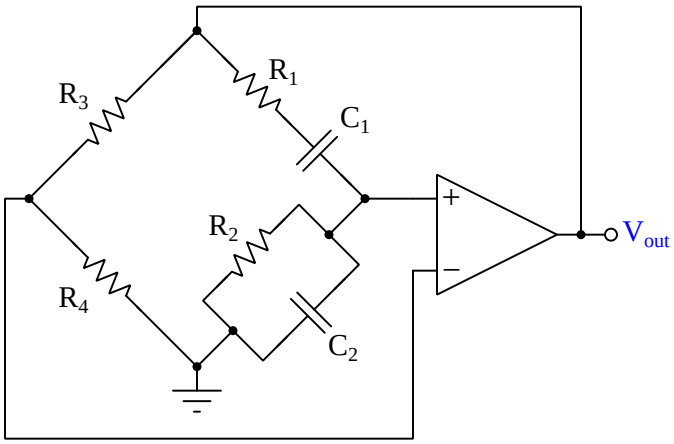
The purpose of each RC network is to produce a 45^o phase shift, so that four of those networks (each one "buffered" with its own opamp voltage follower) together result in 180^o of phase shift. The lower opamp is an inverting amplifier, providing the necessary voltage gain to overcome the attenuation of those four RC networks as well as introducing another 180^o of phase shift to keep the total loop shift at 360^o.

If we choose equal values for every R and every C in those RC phase-shifting networks, we find that the oscillation frequency must be that which results in $X_C = R$. In other words, $f = \frac{1}{2\pi RC}$ for this oscillator.

An interesting feature of this oscillator circuit is that each of the voltage-follower states provides its own sinusoidal output, meaning there are actually *four* sinusoidal voltage signals available in this circuit, all shifted 45^o apart from each other. Some electronic applications require multiple sinusoidal signals separated by a particular phase shift, and this circuit delivers. We could re-design it to have just three RC sections and then the phase shift would be 60^o between those stages, or re-design it to have more than four RC sections for less phase shift between outputs (albeit at a different frequency than $f = \frac{1}{2\pi RC}$ in each case).

A more practical phase-shift oscillator design utilizing an operational amplifier for voltage gain is the *Wein bridge* oscillator, shown in the following schematic diagram. This oscillator uses an RC bridge network to create a 360 degree phase shift as well as a non-shifted feedback signal. The phase-shifted signal gets sent to the opamp's "noninverting" (+) input terminal for positive feedback, while the non-shifted signal gets sent to the "inverting" (−) input terminal for negative feedback:

Wein-bridge oscillator



With equal resistor values $R_1 = R_2$ and equal capacitor values $C_1 = C_2$ in the phase-shifting side of the bridge network, that side of the bridge achieves 360 degrees of phase shift and an attenuation factor of $\frac{1}{3}$ when $X_C = R$. This means a Wein bridge oscillator circuit built with equal-value resistors and equal-value capacitors will oscillate at a frequency predicted by the formula $f = \frac{1}{2\pi RC}$, which happens to be the same formula used to predict the cutoff frequency for a simple RC filter network, and for the same reason[35]: a simple resistor-capacitor filter network is said to be in a state of "cutoff" when $X_C = R$. The non-shifting side of the bridge must be designed to feed back $\frac{1}{3}$ of the opamp's output voltage to the inverting input (i.e. R_3 must be twice the resistance of R_4), which is the necessary condition to prompt the opamp to exhibit the voltage gain of 3 necessary to overcome the shifted signal's attenuation and sustain oscillations.

The Wein bridge oscillator circuit requires fewer components than the staged RC-network oscillator, and therefore making its frequency adjustable by the user is an easier design feat. A weakness of all feedback oscillator circuits, though, is the necessity of maintaining *precisely* the right amount of amplifier gain to ensure proper oscillation. Too little gain, and any oscillations will decay over time until the circuit no longer oscillates; too much gain, and the oscillations grow in amplitude until the waveform "clips" at the amplifier's power supply rail limits. This "clipping" distorts the wave-shape from its ideal (sinusoidal) form, which may lead to poor results (e.g. unwanted harmonics[36]).

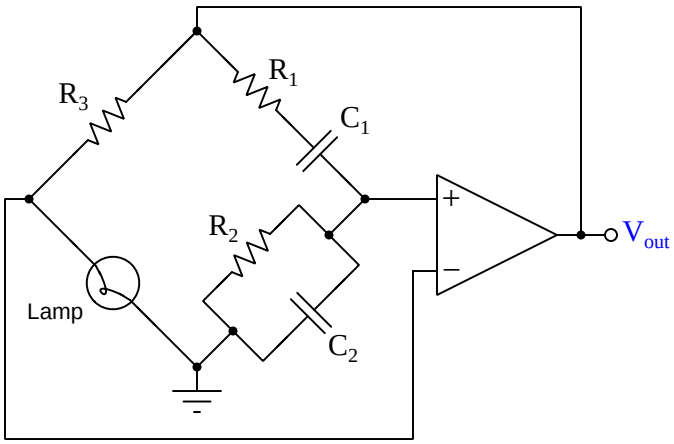
[35] We know that $X_C = \frac{1}{2\pi fC}$, so if a simple RC filter's cutoff point is defined by that frequency where $X_C = R$, then we may solve for f in the formula $R = \frac{1}{2\pi fC}$ to arrive at $f = \frac{1}{2\pi RC}$.

[36] Recall that any non-sinusoidal waveform is mathematically equivalent to multiple sinusoids at different frequencies (called *harmonics*) and amplitudes. Thus, an oscillator outputting a non-sinusoidal waveform is equivalent to

A practical solution to this problem is to equip the oscillator circuit with *automatic gain control* (abbreviated *AGC*). A feedback oscillator circuit able to automatically adjust its own voltage gain will not only produce sinusoidal signals of higher purity, but will also exhibit greater tolerance to loading when connected to other circuits.

An ingenious[37] implementation of AGC in the Wein bridge oscillator design was invented by William Hewlett (of Hewlett-Packard fame) in 1939. Instead of relying on fixed-value resistors to set his amplifier's voltage gain, he used an *incandescent lamp* as one of the resistive elements. The glowing metal filament of an incandescent lamp has a positive resistivity coefficient, which means its resistance rises as it becomes hotter. As V_{out} grows in magnitude, the lamp glows brighter (i.e. hotter filament) thereby increasing the voltage-division ratio $\frac{R_{lamp}}{R_{lamp}+R_3}$ and *automatically* reducing output voltage V_{out} through the action of negative feedback:

Wein-bridge oscillator with AGC



Modern variations on this theme use JFETs or other variable-resistance electronic devices in place of a lamp, the conductivity of that component usually adjusted by the output of another opamp.

3.20.4 Resonant oscillators

Another class of oscillator circuit used to generate sinusoidal waveforms is based on *resonance*: that phenomenon by which either a network of capacitors and inductors or a mechanical element with mass and elasticity cyclicly exchanges energy back and forth. Like phase-shift oscillator circuits, resonant oscillators exploit positive feedback but use the resonant network or element to introduce the 180^o phase shift necessary to combine with the opamp's 180^o phase shift to produce the 360^o of shift necessary to sustain oscillations.

outputting a *series* of sinusoidal signals, some of which may result in undesired effects in the circuit being excited by the oscillator.

[37] This really is an excellent example of creative engineering: exploiting the voltage/current characteristics of a component intended for a completely different purpose. For more detail on this design, refer to the Historical References chapter of this module.

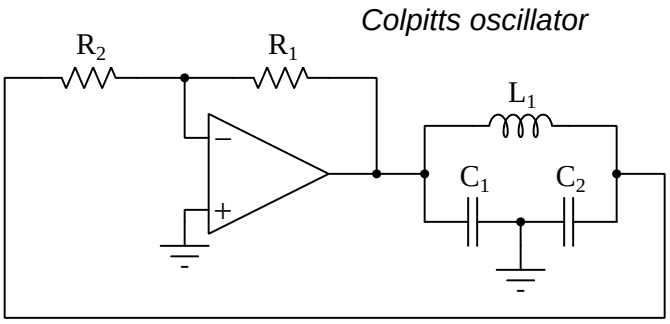
The resonant frequency of a simple LC "tank" circuit is that which causes inductive reactance and capacitive reactance to be equal ($X_L = X_C$), predicted by the following formula:

$$f = \frac{1}{2\pi\sqrt{LC}}$$

Where,

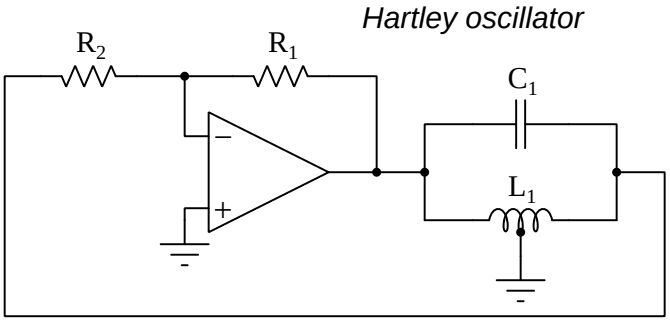
f = Resonant frequency, Hertz
L = Inductance, Henrys
C = Capacitance, Farads

The *Colpitts* oscillator design uses one inductor and two capacitors for its tank circuit:



In this tank circuit, the equivalent capacitance is the series combination of C_1 and C_2, which of course is $\frac{1}{\frac{1}{C_1}+\frac{1}{C_2}}$ because capacitances diminish rather than add in series.

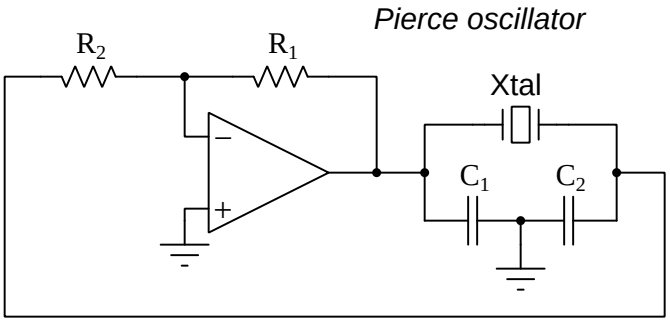
The *Hartley* oscillator is similar, using a tapped inductor rather than a pair of capacitors:



In both the Colpitts and Hartley oscillator designs, the "center-tap" feature of the tank circuit provides a 180^o phase shift[38] necessary to complement the inverting amplifier's characteristic to make 360^o of total phase shift. Resistors R_1 and R_2 establish the voltage gain necessary to overcome the tank circuit's natural attenuation of signal so that the total gain is unity (1).

[38] For any direction of current oscillating within the tank circuit, the voltage polarities across each half of the center-tapped element will be series-aiding, which means when the left-hand side of the tank circuit is positive with respect to ground the right-hand side will be negative, and vice-versa.

The *Pierce* oscillator design follows the Colpitts model, replacing the inductor with a *quartz crystal*:



Quartz is one of many different types of naturally-occurring crystalline materials exhibiting *piezoelectricity*: such a crystal produces voltage when physically strained (compressed or stretched), and also strains with the application of voltage. When cut into specific shapes, a quartz crystal will resonate mechanically, and its piezoelectric properties makes it a suitable analogue[39] for an LC tank circuit.

Crystals happen to have an extremely sharp frequency response – far more selective than most LC tank circuits – and therefore the frequency stability of the Pierce oscillator is exceptional.

As with RC phase-shift oscillators, resonant oscillators output sinusoidal wave-shapes for the same basic reason: only one frequency produces the necessary phase shift for the oscillator to function. Fourier's Theorem declares that all periodic wave-shapes are equivalent to a series of sinusoidal waves at different frequencies summed together, but that only a perfectly sinusoidal wave contains just one frequency. Therefore, if only one frequency produces the necessary phase shift to sustain oscillations in an oscillator circuit, those sustained oscillations – consisting of just one frequency – must be sinusoidal in shape.

[39] One might think of a quartz crystal as akin to a mechanical pendulum with a solenoid coil for magnetic "pickup" of motion as well as for producing a driving force on the pendulum. The crystal's mechanical resonance acts to stabilize the oscillator circuit at that same frequency, just as an electromagnetic pendulum would do the same. The idea of an electromagnetic resonator for an electronic oscillator circuit is not hypothetical either: a type of fluid flowmeter called a *Coriolis effect mass flowmeter* uses a positive-feedback oscillator to shake a steel tube filled with moving fluid, the tube's elasticity and the fluid's mass acting together to form a resonant mechanical assembly. As the fluid's density varies, the tube's effective mass varies with it, altering the assembly's resonant frequency and thereby shifting the frequency of the oscillator circuit driving and sensing the tube's motion. This frequency becomes an inverse-proportional representation of fluid density in the Coriolis flowmeter.

Chapter 4

Historical References

This chapter is where you will find references to historical texts and technologies related to the module's topic.

Readers may wonder why historical references might be included in any modern lesson on a subject. Why dwell on old ideas and obsolete technologies? One answer to this question is that the initial discoveries and early applications of scientific principles typically present those principles in forms that are unusually easy to grasp. Anyone who first discovers a new principle must necessarily do so from a perspective of ignorance (i.e. if you truly *discover* something yourself, it means you must have come to that discovery with no prior knowledge of it and no hints from others knowledgeable in it), and in so doing the discoverer lacks any hindsight or advantage that might have otherwise come from a more advanced perspective. Thus, discoverers are forced to think and express themselves in less-advanced terms, and this often makes their explanations more readily accessible to others who, like the discoverer, comes to this idea with no prior knowledge. Furthermore, early discoverers often faced the daunting challenge of explaining their new and complex ideas to a naturally skeptical scientific community, and this pressure incentivized clear and compelling communication. As James Clerk Maxwell eloquently stated in the Preface to his book *A Treatise on Electricity and Magnetism* written in 1873,

> It is of great advantage to the student of any subject to read the original memoirs on that subject, for science is always most completely assimilated when it is in its nascent state . . . `[page xi]`

Furthermore, grasping the historical context of technological discoveries is important for understanding how science intersects with culture and civilization, which is ever important because new discoveries and new applications of existing discoveries will always continue to impact our lives. One will often find themselves impressed by the ingenuity of previous generations, and by the high degree of refinement to which now-obsolete technologies were once raised. There is much to learn and much inspiration to be drawn from the technological past, and to the inquisitive mind these historical references are treasures waiting to be (re)-discovered.

4.1 Platform balance scale operation

The concept of balancing an unknown quantity against a known quantity to determine the value of the unknown quantity is an ancient one. An early application of this concept was for weighing objects of unknown mass: the unknown mass is placed on a pan or on a platform, and then a person places various masses of known quantity elsewhere on a mechanism until that mechanism shows a condition of balance. When that condition of balance is reached, the unknown mass's value is then determined from the sum total of the known masses used to achieve balance.

In the photograph below we see an antique *platform scale* employing this mass-balance principle. One or more objects of unknown mass are placed on the flat wooden platform, in this case a pair of wooden chests, the weight of that total mass pulling a linkage mechanism to unbalance a small metal beam near the top of the scale. At the far end of that beam is a hanger for cast-iron plates of precise mass to be stacked, as well as a sliding mass made of brass metal which may be positioned anywhere along a graduated scale spanning the beam's length:

The small metal beam enjoys a mechanical advantage over the wooden weighing platform, in that a relatively small mass positioned on the beam is able to counterbalance a much larger mass resting on the platform.

Below we see one of the cast-iron masses held in hand, rated at 100 pounds:

A slot cast through the radius of this mass disk allows it to be easily slipped over a metal rod supporting a stack of such masses. From the "100 lbs" rating cast into this metal disk we know its mass is capable of balancing 100 pounds' worth of mass on the platform, yet the disk itself weighs far less than a hundred pounds since we're able to easily hold it in our hand. This disparity between the disk's balancing capacity versus its actual mass attests to the mechanical advantage of the platform scale mechanism: that the balancing beam acts as a lever with unequal moment arms, the platform tugging on that beam over a short moment arm and the calibrated masses tugging on the other end of the same beam over a much larger moment arm.

To demonstrate how this scale would be adjusted to measure the amount of mass on the platform, I placed just enough of the cast-iron metal disks on the vertical rod where one more disk would tip the balance beam too far, then I slid the movable brass piece along the beam until it reached a balanced position. From left to right in the photographs below we see too little mass (left), just the right amount of mass (center), and too much mass (right):

When the brass piece was positioned too far toward the beam's fulcrum (left), the far-right end of the beam raised up until it rested against an upper stop on the scale's frame. When the brass piece was moved too far away from the beam's fulcrum (right), that end of the beam dropped until

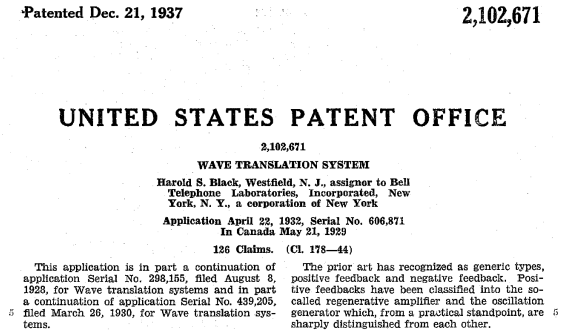
it rested on a lower stop. Only when the brass piece was moved to the correct position did the beam come into a state of balance with its far-right end hovering between the two frame stops. Determining the amount of mass placed on the platform is then a simple matter of adding up the values of the stacked disks plus the value indicated by the beam's graduated scale where the brass piece's pointer indicates.

This basic principle of balancing known quantities against an unknown quantity in order to determine the latter's value is still applied today in electrical signal measurement and amplification. The *null balance* technique of balancing a known voltage against an unknown voltage and using an extremely sensitive meter called a *null detector* to indicate when those two voltages are equal still finds application in electrical metrology. The *negative feedback* principle of an amplifier's output signal balancing against an input signal is widely used in *operational amplifier* circuit design. Null-balance voltage measurement is traditionally (and often still) performed with a human being making the adjustments until balance is achieved, much like operating the antique platform scale. Negative feedback amplifier circuits are *self-balancing* in that they sense the condition of balance and automatically adjust their output voltage signals until balance is achieved.

4.2 Harold Black's feedback patent

In April of 1932 an electrical engineer named Harold Black applied for a United States patent on a design technique useful for making exceptionally stable electronic amplifier circuits. His job working for Bell Labs permitted wide-ranging explorations in circuit theory while providing a very practical application – telephone circuits – to apply any discoveries and innovations toward. One of the problems at that time centered around vacuum tube-based "repeating" amplifiers used to boost telephone signals at certain points along long-distance cables. The gain values of these amplifiers were known to drift beyond acceptable limits on a regular basis, partly due to the electronic characteristics of the vacuum tubes changing as the tubes aged, and this (among other factors) posed problems for reliable communication in Bell's telephone system. Black's focus was to improve the gain-stability of these amplifiers, but to do so in a way that did not require improvements to the vacuum tubes themselves. He sought to find a more comprehensive design-based solution to the problem that would work despite irregularities in the electronic components.

Black's research culminated in a patent which will be the focus of this section. His notion of a "wave translation system" applied to electronic amplifiers to be sure, but had applications to any system involving waves, even non-electrical[1]:



Patented Dec. 21, 1937 2,102,671

UNITED STATES PATENT OFFICE

2,102,671

WAVE TRANSLATION SYSTEM

Harold S. Black, Westfield, N. J., assignor to Bell Telephone Laboratories, Incorporated, New York, N. Y., a corporation of New York

Application April 22, 1932, Serial No. 606,871
In Canada May 21, 1929

126 Claims. (Cl. 178—44)

This application is in part a continuation of application Serial No. 298,155, filed August 8, 1928, for Wave translation systems and in part a continuation of application Serial No. 439,205, filed March 26, 1930, for Wave translation systems.

The prior art has recognized as generic types, positive feedback and negative feedback. Positive feedbacks have been classified into the so-called regenerative amplifier and the oscillation generator which, from a practical standpoint, are sharply distinguished from each other.

A noteworthy feature of this patent is how it patiently explains the then-common understanding of feedback as applied to electronic circuits, and goes on to explain in great detail how Black's application of negative feedback was a true innovation. This clear exposition deserves the attention of all students of electronics.

[1]Page 1 of the patent contains this notable paragraph: *"The invention is applicable to any kind of wave transmission such as electrical, mechanical or acoustical, and thus far in the description the terms used have been generic to all such systems. The invention will be disclosed herein, however, as specifically applied to electrical systems, it being understood that the principles involved are equally applicable to other types of Wave transmission and that the generic claims are intended to include electrical and other than electrical wave systems and apparatus."*

Black begins his description of how feedback might be applied to an electronic amplifier circuit. Note his use of the word *grid*, referring to the "grid" structure of a vacuum tube, the metallic screen placed between cathode and anode ("plate") to control the flow of electrons streaming between those two metal surfaces. The modern equivalent of "grid" would be *base* for a bipolar transistor amplifier and *gate* for a field-effect transistor amplifier:

> Reverting now to the elemental feedback system above, described, and taking a typical electrical case, for illustration, the amplifying element may be a grid-controlled discharge tube of ordinary type. The input circuit leads up to the input side from a suitable source of input waves to be amplified, for example, signal waves. The feedback is connected from the output circuit to the input circuit. It will be noted that the waves actually applied to the grid come from two circuit branches, identified as the incoming circuit (from the signal or other wave source) and the feedback circuit. Of course, each of these two waves, the incoming or signal wave and the fed-back wave, could be thought of as being separately applied to the grid, but it will be simpler to think of a summation wave as the voltage actually effective on the grid. Thus, there are three waves to be considered in relation to the input side of the amplifier, (1) the incoming or original wave to be amplified, (2) the fed-back wave, and (3) the voltage wave effective on the grid, this latter being the resultant produced by the algebraic addition of the other two. [page 1]

Having described the electrical connections involved with adding feedback to a tube-type amplifier, Black proceeds to discuss some of the historical uses (i.e. *prior art*) of feedback in electronic amplifiers as a prelude to describing his own innovation:

> The prior art has recognized as generic types, positive feedback and negative feedback. Positive feedbacks have been classified into the so-called regenerative amplifier and the oscillation generator which, from a practical standpoint, are sharply distinguished from each other. [page 1]

As Black notes, the principle of "feeding back" part of an amplifier's output signal back to its own input was not unknown at his time, and in fact had been practically used for multiple purposes. The two purposes he notes are *regenerative amplifiers* and *oscillators*, the fundamental difference between these two circuits being the amount of positive feedback applied.

His next paragraphs describe existing uses of positive feedback very well. Large amounts of positive feedback is used to generate self-sustaining oscillations (making the amplifier circuit an *oscillator*) while small amounts of positive feedback work to boost the gain of an amplifier:

> If we picture the fed-back wave and the incoming wave as separately adjustable, we may imagine an increasing amount of fed-back waves in a direction to augment the effect of the incoming wave with a corresponding decrease of the incoming wave. The adjustments may continue until the input wave has been entirely replaced by the fed-back wave. That is, the incoming wave has been reduced to zero and the incoming circuit can be disconnected from the amplifier. The waves in the circuit, then become self-sustained and we have the familiar case of the wave generator developing and maintaining continuous oscillations, in a closed loop consisting of the input circuit, the amplifier, the output circuit and the feedback circuit back to the input circuit.
>
> If a smaller amount of positive feedback is employed than is required to produce sustained oscillations, we have the case of the regenerative amplifier. Some of the incoming signal wave after amplification is again put back on the input circuit in such manner as to augment the original incoming signal so that reamplification occurs in the same tube, causing the signal to build up to relatively higher value than could be done by traversing the amplifying element but once. The useful limit of amplification by such a circuit is reached when a further increase in the degree of feedback would result in producing self-oscillation or when the circuit is sufficiently close to the oscillating condition to cause an intolerable amount of signal distortion.
>
> A broad practical distinction between the oscillator and the regenerative amplifier is that in the latter the output current remains under the control of the incoming signal and must follow its variations of starting, stopping, growing stronger or weaker and exhibiting the characteristic quality of the Signal, whereas in the oscillator the wave circulating around the regenerative loop is self-determined and is beyond the control of any signal or other input wave. `[page 1]`

So far, Black has only given attention to *positive* feedback. With this, he then turns to an exposition of *negative* feedback. Again, this was nothing new in his time, and so he had to make a logical case that his invention was a true innovation and not just an existing design principle in disguise:

> Turning, now, from positive feedback to negative feedback we find that the principal field of use of negative feedback has been in radio frequency amplifiers.
>
> A radio frequency amplified is a familiar instance of a circuit in which there may exist an `[page 1]`
>
> inherent tendency toward self-oscillation because of a positive feedback produced by inductive or capacitive coupling between elements of the output and input circuits even where great care in design is exercised to reduce such coupling. These effects become more pronounced as the frequencies become higher and higher. The tendency toward self-oscillation in such circuits places a limit on the amount of amplification that can be used.
>
> Negative feedback has commonly been applied in the prior art in radio frequency amplifiers to oppose the inherent positive feedback. The negative feedback in all such

> cases has had as its purpose the reduction of the positive feedback. If we imagine that the negative feedback is increased from an infinitesimal amount in any given case, it reaches its optimum value in opposing positive feedback when it just equals the positive feedback. At that point, the net or resultant total feedback is zero and the effect is that of rendering the amplifier a strictly unilateral circuit or one having no feedback, either positive or negative. [page 2]

Black's description of negative feedback as a tool to cancel a radio-frequency amplifier's intrinsic positive feedback is commonly known as *neutralization*. The problem, as Black describes in 1932, is that unintentional "coupling" between different elements within a radio-frequency circuit biases it toward instability, causing RF amplifiers to self-oscillate. Small amounts of negative feedback added to an RF amplifier – just enough to exactly cancel out or "neutralize" the undesired positive feedback – thus serves a familiar and practical purpose. Black's claim, though, was that his use of negative feedback (i.e. the *invention*) was of a different scale and for a different purpose, which he begins to describe next:

> In contrast to the negative feedback of the prior art in which, as discussed above, the resultant total feedback is substantially zero, the invention uses negative feedback for an entirely different purpose and in very much larger amounts to achieve improved results in wave translation, e.g., amplification, not envisaged in the prior art.
>
> Applicant has discovered how to use larger amounts of negative feedback than were contemplated by prior art workers with a new and important kind of improvement in tube operation. One improvement is in lowered distortion arising in the amplifier. Another improvement is greater constancy of operation, in particular a more nearly constant gain despite variable factors such as ordinarily would influence the gain. Various other operating characteristics of the circuit are likewise rendered more nearly constant. Applicant has discovered that these improvements are attained in proportion to the sacrifice that is made in amplifier gain, and that by constructing a circuit with excess gain and reducing the gain by negative feedback, any desired degree of linearity between output and input and any desired degree of constancy or stability of operating characteristics can be realized, the limiting factor being in the amount of gain that can be attained rather than any limitation in the method of improvement provided by the invention. [page 2]

In this last paragraph, Black summarizes his discoveries about the use of large amounts of negative feedback in amplifier circuits: this causes the amplifier to exhibit lower amounts of signal distortion, as well as exhibit a more stable gain. Both of these improvements were of great interest to Black's employer – Bell Labs – for use in their telephone system. Perhaps the most counter-intuitive notion Black expresses in this paragraph is that it actually makes sense to build an amplifier with an abnormally high amount of gain ("excess gain") and then use negative feedback to globally reduce this gain in order to achieve the aforementioned benefits of lower distortion and greater stability. If one were to view negative feedback merely as a way to force an amplifier to have a reduced gain, then building a high-gain amplifier only to attenuate that gain with negative feedback seems utterly pointless. No, argues Black, the real point of reducing all that "excess" gain is to enjoy lower distortion and greater stability.

Having summarized the unique benefits of large amounts of negative feedback, the author goes on to explain practical problems alleviated by negative feedback. First, the problem of amplifier distortion:

> Also, from what has been said, it is apparent that applicant uses negative feedback for a purpose quite different from that of the prior art which was to prevent self-oscillation or "singing". To make this clearer, applicant's invention is not concerned, except in a very secondary way (to be explained later on), with the singing tendency of a circuit. Its primary purpose has no relation to the phenomenon of self-oscillation. If amplifiers could be built exactly like present day amplifiers in all respects except that they were absolutely free of any tendency of self oscillation regardless of how high their gains might be pushed, it is generally true that there would still be as great need for the present invention. The "perfect" amplifier is one in which the output wave is in all respects an exact replica of the input wave multiplied by some linear factor. The singing tendency is only one of several difficulties encountered in an amplifier. Actually, amplifiers produce distortion components along with the output fundamental components. These modulation or distortion components are mingled with the fundamental wave and detract from its purity. In multiplex carrier systems the distorting effect gives rise to cross-talk and, in any high quality system, it constitutes a limiting factor of design. `[page 2]`

Next, the problem of unstable amplifier gain:

> Another difficulty in amplifier operation is instability, not used here as meaning the singing tendency but rather signifying constancy of operation as an amplifier with changes in battery voltages, temperature, apparatus changes including changes in tubes, aging and kindred causes. Such instability is a limiting factor where, for example, a large number of repeaters are operated in tandem in a line. A simultaneous tendency toward increased gain by several or all of the repeaters might be disastrous. Without a means of preventing such changes, they become a limiting factor of design. `[page 2]`

Here it is easy to understand Bell Labs' interest in Black's discovery. By using negative feedback to reduce signal distortion and to tame unstable amplification, practical problems plaguing their repeater amplifiers could be solved. This would reduce channel-to-channel interference on multiplexed[2] telephone systems. By using negative feedback to stabilize each amplifier's gain, cascaded repeater stations could provide reliable end-to-end operation despite the inevitable aging of vacuum tubes, tube replacement, environmental changes, and supply voltage variations.

[2] *Multiplexing* refers to transmitting multiple streams of data along a common communication channel. There are many different ways to do this, but in Black's time it involved *modulating* high-frequency waveforms with information from audio-frequency voice signals, much in the same way as radio transmitting stations broadcast "modulated" RF waveforms from their antennas.

As is customary with patent applications, the author provides a series of illustrations and diagrams to augment the text's explanation of the invention. Black's patent is no exception, and includes several insightful diagrams and graphs.

Figure 1, shown below, gives a generalized schematic diagram of feedback applied to a multi-tube amplifier. To anyone unfamiliar with vacuum-tube circuit schematics, the symbols used here may look very strange. The grid appears as a kind of single-leg "resistor" oriented vertically, while the plate is a small rectangle. The cathode is a triangle located between the two. Grid, plate, and cathode are analogous to base, collector, and emitter in a BJT, or gate, drain, and source in a FET. The feedback is simply a resistor labeled f connecting the output terminal (plate) of the last tube with the input terminal (grid) of the first tube, with the size of that resistor determining the amount of negative feedback imposed on the amplifier:

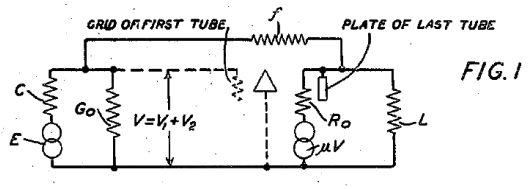
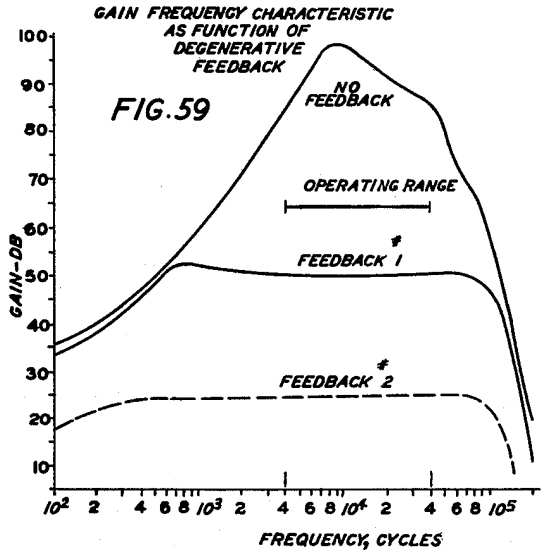
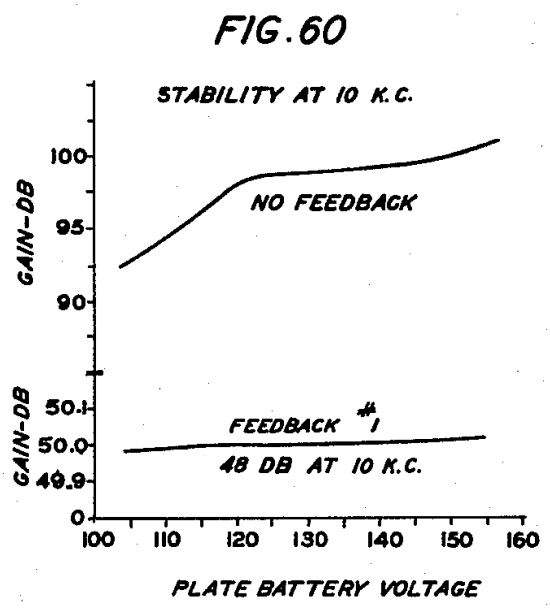


Figure 59, shown below, plots the gain of an amplifier as a function of signal frequency with varying degrees of negative (i.e. *degenerative*) feedback. As the graph shows, the amplifier's gain varies widely from 35 dB at the lowest frequency to nearly 100 dB at its peak gain without the benefit of any negative feedback. With some negative feedback added to the circuit, the gain appears much more stable, especially over the typical operating range from 4 kHz to 40 kHz. With even more feedback, the gain curve is almost perfectly "flat" at approximately 25 dB from 400 Hz all the way to 60 kHz. Bear in mind that the only modification made to the amplifier to achieve such stability was including a resistive feedback network from its output to its input:



The next graph, Figure 60, plots amplifier gain as a function of plate (DC power supply) voltage, given a constant 10 kHz signal to amplify. As in the previous graph, we see dramatic improvements in the stability of the gain due to the addition of negative feedback:



This last illustration, Figure 18, shows three different "views" of a three-stage amplifier with negative feedback. On top is a simplified schematic diagram of the amplifier circuit, using three "triode" vacuum tubes connected in common-cathode configuration (equivalent to common-emitter for BJT or common-source for FET). The odd number of common-cathode stages makes this amplifier *inverting* at the output with respect to the input, since each common-cathode stage inverts the polarity of its signal. The middle diagram represents the feedback "loop" which is even more simplified from the schematic. The bottom diagram is simplified even more, omitting the tubes entirely:

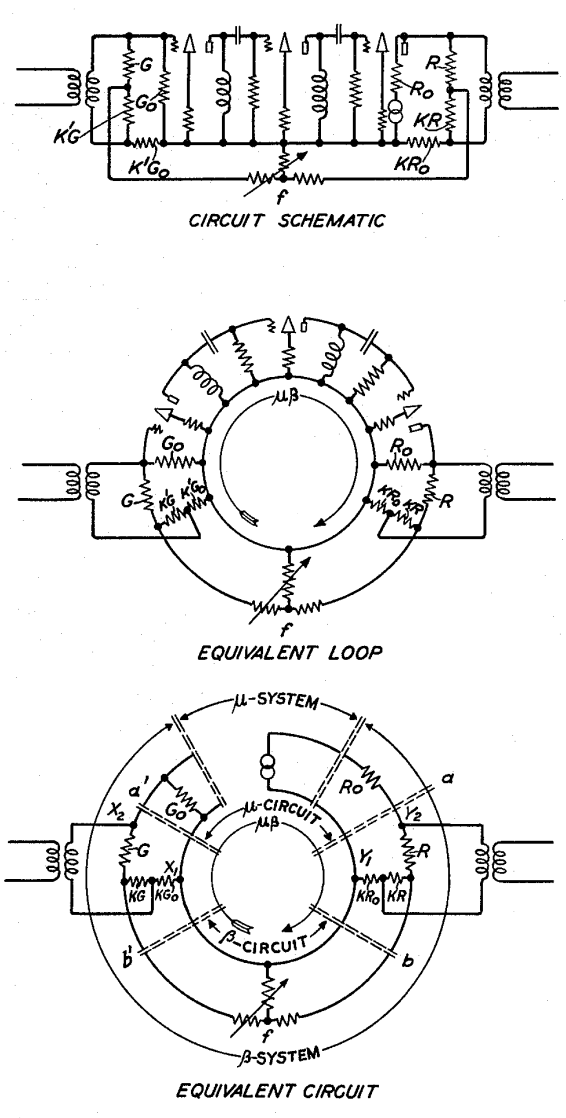


FIG. 18
ILLUSTRATIVE APPLICATION OF METHODS OF FIGS. 14, 15, 16, AND 17

Black's experimental feedback network consists of a simple "T-pad" attenuator, adjustable to

provide a convenient method for testing the amplifier with varying amounts of negative feedback. The feedback is inherently *negative* and not positive due to the inverting nature of the amplifier: with the output signal 180^o out of phase with the input signal, any feedback passed through a non-shifting (i.e. resistive) network must be negative feedback.

Throughout the patent's text, Black uses the lower-case Greek letter "mu" (μ) to symbolize the amplifier's gain without negative feedback (what we would call the *open-loop gain* of the amplifier). He also uses the Greek letter "beta" (β) to symbolize the feedback network's attenuation factor. If the amplifier's open-loop gain is arbitrarily large, such that the product $\mu\beta$ is much greater than unity, the system's gain with negative feedback in effect comes close to being equal to $-\frac{1}{\beta}$.

This conclusion is rather striking, as it tells us the system's gain may be set to a very stable value simply by adjusting a simple resistor network to the reciprocal of that desired gain. For example, if the feedback network has an attenuation of 5:1, then the amplifier's over-all (i.e. *closed-loop*) gain will approach 5 regardless of the internal (open-loop) gain of the amplifier without feedback. Modern operational amplifier circuits come very near this ideal realization, owing to the spectacular open-loop gain capabilities of modern semiconductor amplifiers.

4.3 George Philbrick's work on gun tracking systems

The *National Defense Research Committee* published a series of reports following the end of World War Two on their developmental work into automatic tracking systems for guns. One of these reports, "Summary Technical Report of Division 7, NDRC, Volume 3, Airborne Fire Control" contains writing of a certain George Philbrick who was to become a very important figure in the development of operational amplifiers and analog feedback systems in general. Philbrick was the founder and namesake of George Philbrick A. Researches, Inc., and his contributions to early fire-control system development were highly influential.

What follows are excerpts from this now-declassified military report.

Chapter 2
ON CERTAIN ASPECTS OF TRACKING

2.1 TRACKING AND OUR CONCERN THEREWITH

By "tracking" is here meant a continuous following or alignment procedure which is pursued through cognizance of its error or unbalance, as in the basic feedback or following operation described in Section 1.10. A good example, although somewhat abstractly disclosed, is given in Section 1.8, where an arbitrarily varying position vector is to be followed by an artificially generated position vector which is influenced by manipulation of an integration constant and a velocity.

Tracking pervades all branches of fire control, but it arises principally in aiming operations as such, whether these be manual or automatic. With regard to the latter, automatic tracking is essentially indistinguishable from other types of automatic control and will here be considered in that light. However, our major interest in the present chapter is in tracking by manual means, a human operator being involved. While the removal of human limitations gives the greater promise to automatic tracking devices and techniques, still there is a good bit to be learned, even for such unlimited future trends, from tracking of the human variety. Besides this, the human element can hardly be altogether eliminated, even ultimately, and until we give to our machines the final powers of judgment and choice we shall find it necessary personally to direct their efforts. It may also be doubted that the adaptability or educability of the brain, whereby it adjusts to altered circumstances, can soon be imparted to the automaton in any but the most trivial of cases. Or rather, if this is to be done, it will probably be through imitation of the processes by which such adaptations take place in the animal mechanism.

It is typical of the writer that he considers the entire subject of tracking to come within the broader boundaries of regulatory controls, which to him include all corrective apparatus having a closed causal loop. In this connection, and although the region has not been very far explored, there will be traced out briefly in what follows an analogy between the past development of automatic regulators and that of tracking aids. From this analogy a deduction will be made as to the future of the latter field of endeavor. Aside from this conjectural activity, we have been intimately concerned with tracking in a number of development programs where fire-control apparatus was being devised. Tracking in range, and we speak now of manual tracking, whether by stadiametric means or otherwise, has not entered particularly into our immediate sphere, and it is further

evident that this brand of tracking has been dealt with very thoroughly by other groups. Thus it may be said that tracking in angle, or more precisely in direction, is under discussion. The problems and technology of such tracking have appeared in gunnery systems, and in those for bombing, guided bombing, and for airborne rockets.

For flexible gunnery, an elaborate instrumental, statistical, and psychological study of the man-machine interactions which tracking comprises was undertaken and pursued by Section 7.2 at The Franklin Institute. This project (NO-268) dealt with standard types of computing dynamics and with known kinds of aiding controls. Many significant results were obtained; some unexpected. Other members of the section have been more directly involved, however, and their writings should be referred to rather than the present material. (See Part III, for example.) While allusions may be made here to this work, it will not be further discussed or described.

2.2 THE HUMAN FOLLOWER

People "track" during every conscious moment, unless their eyes be closed, their hands tightly tied, and their tongues clamped solidly between their teeth. Alignment processes, in which the alignment error serves as datum for its own annihilation, are forever being carried out in the familiar operations of living. We may thus expect that there is a rudimentary sort of tracking circumstance in which the human operator will be at home, and in which he will be found to be greatly and innately skilled. Pointing at a moving object with a pencil, or with a rifle under beneficial conditions of support and inertia, is a tracking operation which may be carried out relatively well. Provision of a reflecting sight, or of a nonmagnifying telescope with cross hairs, does not improve the operation extensively and has even been seen to impair it. Of course it is a different matter with magnification, since then visibility and visual resolution may be vastly increased. This is not to say that this "natural" type of tracking cannot be improved upon, for in tracking targets which are changing very slowly and uniformly an adjustable-rate device is superior, this being only one example, but on the whole for typical cases, and where scale factors in the presentation are reasonably adjusted, such natural tracking would suffice admirably. It is assumed here that closeness of tracking as such is desired, or simply smallness of tracking error. It is evident that for the larger purposes of aiming contol this is not the only index of excellence. The needs and nature of the interpretive and computing equipment cannot finally be separated from those of the tracking controls. We are thus discussing a subsidiary problem, but more on this again.

We repeat, those tracking arrangements would suffice in which the dynamic relationship between the immediate manual manipulation and the direction index were of the same character as in the arrangements referred to as natural. If this is an overstatement of the case, then we should like to assume that there might be discovered some even more ideal dynamic connection between the manual manipulation and, say, the visual index, and that the dynamic nature hereof might be expressible in quantitative form, even though the ideal differed from one individual to the next.

2.3 THE CAUSAL LOOP

It is frequently pointed out that the elements in a tracking system form a chain, and that consequently the whole assemblage perform as perfectly as each element must is intended

> to perform. But it has not been sufficiently reiterated that these elements form a *closed* chain, or a *complete loop*, together with the human operator, where one is present. This fact brings to the tracking process all the special characteristics, including the unique conditions for stability and periodicity, which are peculiar to that type of system. The formation of such a loop, in the causal sense, is particularly significant when there is included the dynamic characteristic of a "disturbed" or lead-computing sight, so-called. For in spite of the fact that the little understood dynamics of the human operator are included, many of the properties of those loops which are entirely automatic regulatory devices and in many automatic controls, are seen to be present in the tracking sequence. It seems evident that an approach to the tracking problem, thinking of this problem as requiring the provision of improved aiding equipment, would best be made a basis of the recognition that the operations involved are of this cyclic character. Operations in a closed loop are fundamentally different from those in an "open" or "straight-through" system. Both advantages and disadvantages follow from such an arrangement, and this will define our principal topic. We wish first to map out the causal circuit involved in a tracking system in operation.

Philbrick then goes on to diagram such feedback systems, beginning with a block diagram of a human-directed aiming "system" for a gun:

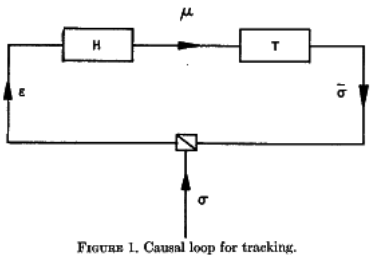


FIGURE 1. Causal loop for tracking.

In this system, variable μ is the position of the gun's handle, function H represents the nervous system of the human gunner, function T represents the physical gun, variable $\overline{\sigma}$ is the *sight index* which represents the position of the gunsight's crosshair center in the field of vision, variable σ is the "true sight direction" which represents the target's position in the field of vision, and variable ϵ is the *error* showing how far off the gun's aim is from the gunner's intended target.

Philbrick gives the feedback loop's error variable particular attention:

> Now the sight index $\overline{\sigma}$ is to be compared with the true sight direction or direction to the target, which may be given the symbol σ. Thus the difference between these two, or the error, which is visible to the human operator, is an important variable. We define the error as
>
> $$\epsilon = \sigma - \overline{\sigma}$$
>
> While under ideal tracking this difference will remain identically zero, still in real operation its value must be continually observed and interpreted to approximate this ideal.

This error between the gun's ideal and actual direction is an example of the alignment error serving as *datum for its own annihilation*, to borrow Philbrick's phrasing. The control system (in this case, the human gunner) constantly monitors this error and tries as best as possible to minimize it so that the gun points directly at its target.

Later in this same chapter, Philbrick notes the similiarities between this and standard "regulatory" control loops as applied in the process industries, no doubt informed by Philbrick's employment with The Foxboro Company which specialized in the design and manufacture of industrial instrumentation and control systems.

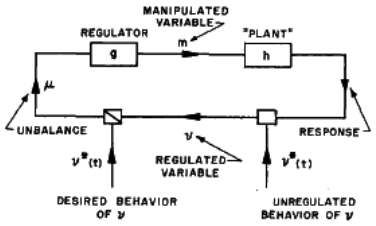


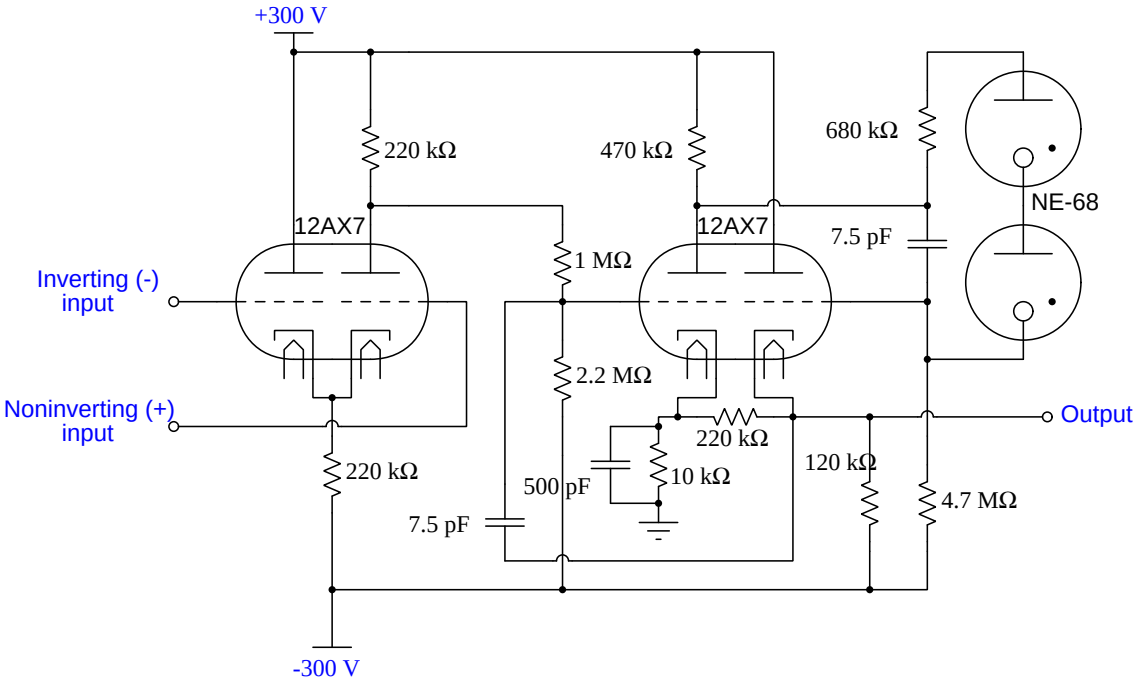
FIGURE 3. Symbolic representations for regulatory control loop.

Here, some of the labels have changed from the gun-aiming system, but the general concept of a feedback loop remains intact. A "regulator" device (g) replaces human judgment and action to influence (m) the physical process being controlled (the "plant" h). The process's response, in turn, is influenced by factors outside anyone's control to become the regulated variable (ν). This variable (known as the *process variable* or PV in modern control terms) gets compared against the desired behavior (known as the *setpoint* or SP in modern control terms), with the difference between those two variables known as the *unbalance* (or "error") μ. In an ideally-functioning control system this error has a value of zero.

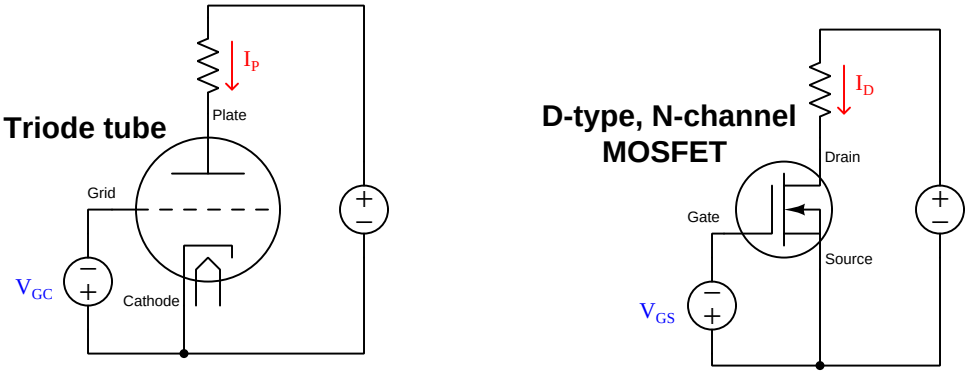
4.4 Philbrick Researches model K2-W operational amplifier

George A. Philbrick Researches, Incorporated, was a manufacturer of analog electronic computing systems founded in 1946. One of their products was a compact (for its time) operational amplifier based on two model 12AX7 triode vacuum tubes called the *K2-W*, a schematic diagram of which is shown here:

Philbrick Researches model K2-W operational amplifier

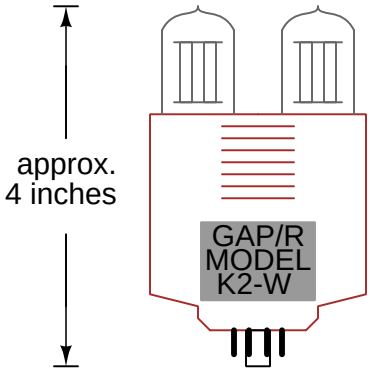


The vacuum tubes used in this circuit design are *dual triode tubes*, model 12AX7. Each tube contained two separate triode elements, each one behaving similar to a D-type N-channel MOSFET where conduction is controlled by an applied voltage between two of the device's terminals.

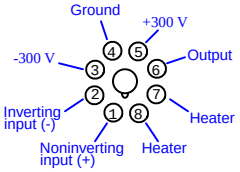


The model K2-W operational amplifier housed the two vacuum tubes along with all passive components shown in the schematic diagram neatly inside of a plastic case, with an eight-pin (octal) plug in the bottom. This octal-base pattern was the same used for some large power vacuum tubes as well as for electromechanical relays, and in fact this same octal pin pattern survives to this day for relays. Illustrations of a typical K2-W module and its pin assignments are shown below:

The Philbrick Researches op-amp, model K2-W



Octal-base "pinout" for model K2-W opamp



This amplifier unit was released for individual sale in 1952, for the purpose of allowing customers to build their own analog computing and signal-processing systems using general-purpose operational amplifier modules. A professor of electrical engineering at MIT, Henry M. Paynter, wrote a 28-page pamphlet titled *Applications Manual for Philbrick Octal Plug-In Computing Amplifiers* just a few years later in 1956, and with the publication of this document Philbrick Researches saw individual opamp sales exceed sales for their complete analog computing systems.

Technical specifications for the K2-W were as follows, taken from the manufacturer's datasheet:

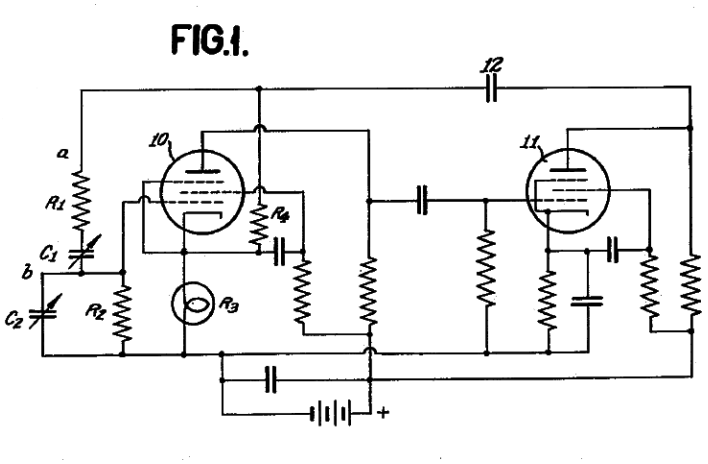
- $A_V(OL)$ = 15,000 open-loop voltage gain
- I_{supply} = 4.5 mA DC at ± 300 VDC
- Heater power = 6.3 Volts (AC or DC) at 0.6 Amperes
- R_{input} = at least 100 MΩ
- Input voltage range = ± 50 Volts
- Output voltage range = ± 50 Volts
- Input bias current = less than 0.1 μA
- Output load current = ± 1 mA driving 50 kΩ load resistance

For comparison, technical specifications for the legacy model 741 (solid-state) operational amplifier introduced in 1968 show dramatic improvements in most metrics:

- $A_V(OL)$ = 50,000 to 200,000 open-loop voltage gain
- I_{supply} = 1.7 mA DC at ± 15 VDC
- (no heaters required!)
- R_{input} = 300 kΩ to 2 MΩ
- Input voltage range = ± 13 Volts
- Output voltage range = ± 14 Volts
- Input bias current = less than 0.1 μA
- Output load current = ± 25 mA driving a short-circuit

4.5 William Hewlett's oscillator

One of the founders of the famous electronics manufacturer *Hewlett-Packard*, William Hewlett, filed for a US patent in early July of 1939 for a reliable sine-wave oscillator. This patent (number 2,268,872) was granted on the 6th of January 1942, and it shows a Wein-bridge design using vacuum tubes (*pentodes*, to be precise) as amplifying elements:



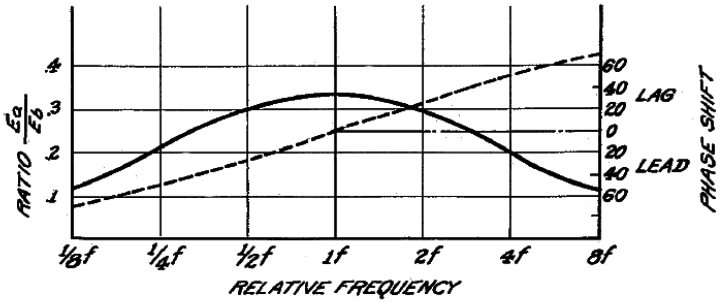
The text of Hewlett's patent begins as such on page 2 of the patent document:

> My invention relates to variable frequency oscillation generators and more particularly to oscillation generators wherein the frequency is determined by means of resistance capacity networks.
>
> Oscillation generators capable of tuning over a wide range of frequencies are useful for many purposes, for example, as test oscillation generators in place of beat frequency generators. In general such oscillation generators have been provided with a variably tuned resonant circuit to maintain the oscillations at a frequency to which the tuned circuit is adjusted.
>
> According to my invention, I provide an oscillation generator with a variable resistance capacity network forming a positive feedback path for the purpose of determining the frequency of oscillation.
>
> According to another feature of my invention an oscillation generator with resistance capacity frequency determining arrangements and with a negative feedback circuit for reducing distortion caused in the amplifier circuit, is provided.
>
> According to a further feature of my invention I provide an automatic amplitude control arrangement in a negative feedback path of an oscillation generator to maintain the amplitude of the generated oscillations constant despite changes in the supply voltages for the tubes. [page 2]

Hewlett continues, on the same page, making reference to the illustrations:

> Other advantages and features of my invention will be apparent from the particular description thereof made in connection with the accompanying drawing, in which
>
> Fig.1 shows a diagrammatic circuit of the preferred form of an oscillation generator in accordance with my invention, and
>
> Fig. 2 shows curves for the purpose of illustrating the operation of the circuit.
>
> In Fig. 1 is shown an oscillation generator comprising two amplifying tubes 10 and 11, coupled together by resistance capacity means in a known manner. The voltage supply sources and the cathode heating supply are omitted for simplifying the circuit. The anode of tube is coupled to the input grid of tube and the anode of tube is then connected to a circuit which provides a feedback path through a coupling condenser 12 and series connected resistance R_1 and variable capacity C_1, and a second circuit comprising R_2 and variable condenser C_2, to the input of tube 10. The circuit including R_1, C_1, R_2, C_2, provides a regenerative or positive coupling between the output and input circuits of the amplifiers 10 and 11, so that feedback occurs sufficient to maintain oscillations in the circuit. By properly proportioning the value of the resistance capacity networks R_1, C_1, R_2, C_2, the desired frequency may be obtained. If R_1, and C_1, respectively, equal R_2, and C_2, the ratio of the voltage E_b at point b to the voltage at point a varies with frequency in a manner similar to a resonance curve. In Fig. 2 this curve is shown in solid lines with voltage ratios of E_a and E_b plotted as ordinates against the relative frequency. At the maximum of this curve, the frequency
>
> $$f_o = \frac{1}{2\pi\sqrt{R_1, R_2, C_1, C_2}}$$
>
> and the voltages at points a and b have the same phase. Accordingly, oscillations tend to occur at the frequency f_o for this circuit. [page 2]

FIG. 2.



In Fig. 2 the broken line curve indicates the phase shifts as ordinates with the change in relative frequency. It can be seen that at zero phase shift, the maximum of the E_a, E_b ratio occurs. If the various condensers and resistances are not made equal as they were for plotting this curve, the maximum would occur at some other point and a symmetrical curve would not be obtained. However, the principles of my invention apply to the circuit regardless of equality of these elements. In practical construction, however, it is generally desirable to make the separate elements equal.

As shown in Fig. 1, frequency shift in the oscillation generator is accomplished by adjusting variable capacities C_1 and C_2. It is clear, however, that since the resistances as well as the capacities serve to control the frequency of the oscillator, either the resistances alone or both the resistances and capacities may be adjustable for varying the tuning.

For the type of oscillators illustrated in Fig. 1, it is necessary in order to obtain satisfactory operation that the amplifiers comprising tubes 10 and 11 have phase shift independent of changes in supply, etc., and furthermore, there should be provided some means for controlling the amplitude of the oscillations so that they do not exceed the range over which tubes 10 and 11 will operate as class A amplifiers. A constant amplifier phase shift is necessary to insure a constant frequency, because the phase angle of the transfer impedance of the resistance capacity network from point a to point b varies only slowly with frequency. Accordingly, a small change in amplifier phase shift such as might be produced by a variation in voltage supply requires a comparatively large change in the frequency of opera- [page 2]

tion to produce a compensating phase shift in the resistance capacity Coupling System. [page 3]

After explaining the basic operation of the oscillator circuit, Hewlett begins to describe some of the more advanced features of his design. He begins by explaining how his circuit is stabilized against variations such as tube parameters, temperature, and supply voltage (using negative feedback):

> In the circuit of Fig. 1, I provide a negative feedback arrangement for stabilizing the amplifier. This feedback is made from the output of amplifier 11 over coupling condenser 12 and the resistance R_3, R_4 to the cathode of tube 10. The potential drop across the resistance produces a negative feedback to the input of amplifier 10. This negative feedback operates in a manner known per se to reduce the amplification of the system and at the same time to stabilize the amplifier gain in magnitude and phase with respect to variations occurring within the amplifier, such as tube changes, temperature variations and changes caused by fluctuation in the voltage supplies to the electrodes of the amplifier system. [page 3]

Next, he explains what is perhaps the most innovative feature of his oscillator's design, a method for automatically controlling gain for the purpose of reducing waveform distortion:

> Amplitude control to prevent the oscillations from building up to such a large value that distortion occurs, is obtained according to my invention by non-linear action in the amplifier circuit. In order to produce this non-linear variation, I provide for resistance R_3 a small incandescent lamp, or similar device in which the resistance increases rapidly with increased current flow, the lamp being heated by the plate current of the tubes 10 and 1l or by an auxiliary means, so such a temperature that its resistance will vary rapidly with a small change in current. Thus, when the oscillation amplitudes tend to increase, the temperature of the lamp R_3 increases with a resulting increase in resistance thereby causing a greater negative feedback, thus reducing the amplification. Similarly, as the oscillations decrease in amplitude, the current through the lamp is reduced permitting the lamp to cool with an accompanying decrease in resistance and reduction of the negative feedback, thus increasing the amplitude of the generated oscillations. As a result the system operates at substantially a constant amplitude which is preselected to be below the value at which grid current flows. As a result no distortion of the wave form takes place.
>
> While I have described the amplitude control system in connection with my particular resistance capacity tuned oscillation generator, it is clear that this feature of my invention is similarly applicable to various types of oscillation generators, wherein frequency stabilization is obtained by means of negative feedback. Since the lamp itself responds relatively slowly to current changes it tends to average out the voltage peaks and does not cut off the peaks of the generated waves, maintaining the amplitude constant without producing distortions in wave form. [page 3]

Chapter 5

Derivations and Technical References

This chapter is where you will find mathematical derivations too detailed to include in the tutorial, and/or tables and other technical reference material.

5.1 IEC standard component values

Components such as resistors, inductors, and capacitors are manufactured in several *standard values*, described by IEC standard 60063. Rather than having a single series of standard values, the IEC publishes lists called *E series* based on the number of unique values spanning a single *decade* (i.e. a 10:1 range).

The shortest of these series, called *E3* contains just three values: 10, 22, and 47. The next series is called *E6* with six unique values: 10, 15, 22, 33, 47, and 68. These values represent *significant values* for components, meaning the decimal point may be freely moved to create values spanning multiple decades. For example, "33" simply means one can expect to find components manufactured in values of 33, 3.3, 0.33, and 0.033 as well as 330, 3.3 k, 33 k, etc.

Although this may seem like a strange standard for component manufacturers to follow, there is a compelling logic to it. The terms of each series are closer-spaced at the low end than at the high end, and this allows for *series* and/or *parallel* combinations of components to achieve most any desired value. For example, in the E6 series we only have values with the significant figures 10, 15, 22, 33, 47, and 68, but this doesn't mean we are limited to *total* values with these significant figures. For example, if we needed 80 Ohms of resistance we could connect a 33 Ohm and 47 Ohm resistor together in series. 50 Ohms could be made from two 68 Ohm resistors in parallel (making 34 Ohms) plus a 15 Ohm and 1 Ohm resistor in series.

On the next page is a table showing the four most common E-series specified by IEC standard 60063.

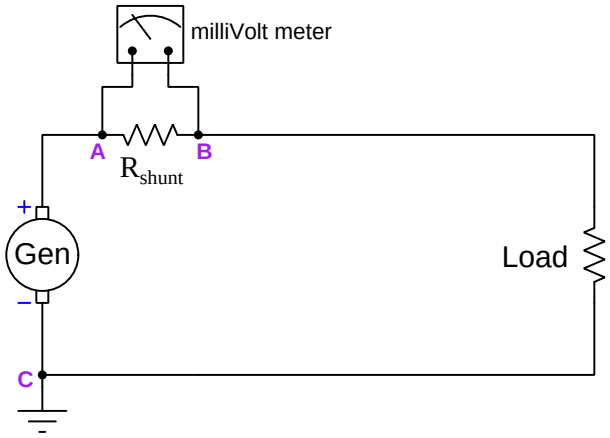
E3	**E6**	**E12**	**E24**
10	10	10	10
			11
		12	12
			13
	15	15	15
			16
		18	18
			20
22	22	22	22
			24
		27	27
			30
	33	33	33
			36
		39	39
			43
47	47	47	47
			51
		56	56
			62
	68	68	68
			75
		82	82
			91

E48, *E96*, and *E192* series are also found in the IEC 60063 standard, used for components with tighter tolerance ratings than typical.

5.2 Common-mode voltages and currents

A topic of much misunderstanding in electronic circuits, often relevant to precise measurements of voltage and current, is that of *common-mode signals*. The topic is especially confusing in regard to *voltage*, as voltage is fundamentally a differential quantity always existing *between two points* and never at a single location. We will explore this topic in some detail here, through practical examples.

Suppose we install a shunt resistor and a milliVolt meter on a power circuit to infer how much current a DC generator sends to a load. Given the existence of three electrically-distinct points in this circuit (**A**, **B**, and **C**) we have three unique voltages – the voltage registered by the milliVolt meter (V_{AB}), the voltage across the load (V_{BC}), and the voltage output by the generator (V_{AC}):

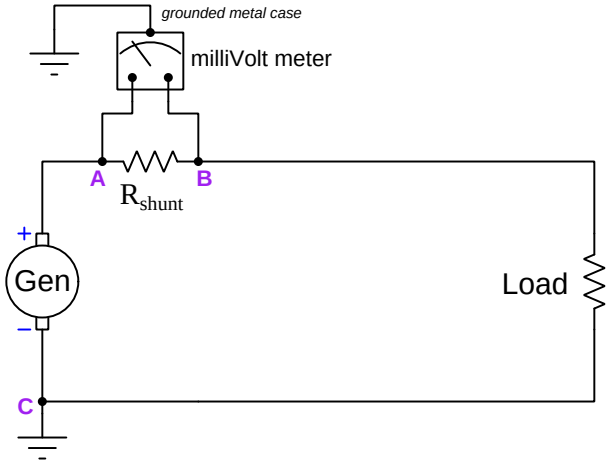


Supposing the generator output 300 Volts DC to a load that happens to be 40 Ohms at some given point in time, with the shunt resistor being a fixed value of 0.1 Ohms, these three voltages would be as follows:

- V_{AB} = 74.81 milliVolts DC
- V_{BC} = 299.25 Volts DC
- V_{AC} = 300 Volts DC

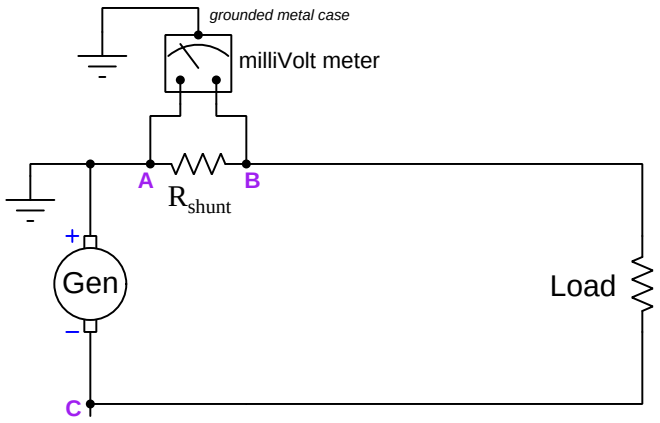
The milliVolt meter's purpose in this circuit, of course, is to infer load current by reporting the amount of potential difference dropped across R_{shunt}. In this case, the registered differential voltage between test points **A** and **B** is 74.81 milliVolts which equates to 7.481 Amperes of load current. The milliVolt meter should be completely immune to any changes in generator voltage, and only respond to current in this circuit.

Now suppose we add one more consideration to this seemingly simple circuit – we consider the milliVolt meter's metal casing attachment to a metal pole planted in the soil. This places the meter's case at ground potential (i.e. the potential of point **C**):



One would expect this to have no effect whatsoever on the meter's proper measurement of differential voltage between its test leads (points **A** and **B**), and in many cases this would be a correct assumption. However, there may be a *safety* concern if the milliVolt meter is not rated to handle such a high amount of voltage between either of its test leads and its metal case, in this instance 300 Volts between **A** and the case, and 299.25 Volts between **B** and the case. If the electrical insulation inside the milliVolt meter were insufficient to handle this much potential difference, it could break down over time and result in a fault where substantial current passes from either test lead to the case, through Earth ground, and back to the generator!

An interesting solution to this dilemma is to relocate the generator's ground connection from point **C** to point **A** as such:

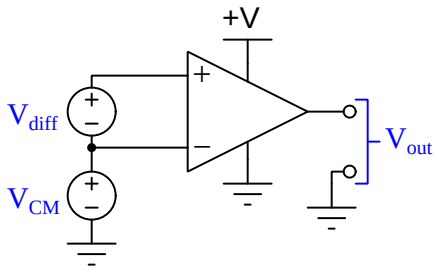


Now the voltage between point **A** and the meter's metal case is 0 Volts, and the potential between point **B** and the meter's metal case is −74.81 milliVolts. Clearly there no longer exists any hazard of electrical insulation breakdown inside the meter! This relocation of Earth-grounding point in the main power circuit has absolutely no effect whatsoever on the generator or on the load or even on the accurate reporting of current as inferred by the shunt resistor, but it certainly has an impact on how much potential difference the meter's internal insulation must withstand, and therefore on the meter's long-term reliability and safety.

By moving the circuit's Earth ground reference point we managed to vastly reduce the amount of electrical insulation stress inside the meter without affecting the generator or the load in the slightest. The meter still registers the shunt resistor voltage drop it's designed to measure, but without the hazard of insulation breakdown we faced before. It would be convenient to have a general term to refer to this change of stress on the meter, which is clearly related to voltage in a general sense but not to the differential voltage across the shunt resistor the meter was set up to read.

Such a term does exist, and it is called *common-mode voltage*. Since in this case we're concerned about the voltage impressed across the insulation protecting wires inside of the meter, the "common" element here would be the meter's metal case, since that is the other pole of any high voltage that any wire inside the meter would experience. We might interpret "common-mode voltage" to mean *any voltage shared in common between the two meter test leads and the metal case*, but this is a somewhat ambiguous definition. How much voltage, exactly is "common" between the 300 Volts (between **A** and case) and 299.29 Volts (between **B** and ground) when this circuit's Earth ground point is located on the generator's negative pole? We might define "common-mode" to be the lower of the two voltages (299.25 Volts), or perhaps the higher of the two (300 Volts), or maybe even the average of the two (299.63 Volts). A compelling argument could be made that the best way to define "common" in this context would be the greater of the two ground-referenced voltages (300 Volts) because what we're ultimately concerned about is over-stressing electrical insulation.

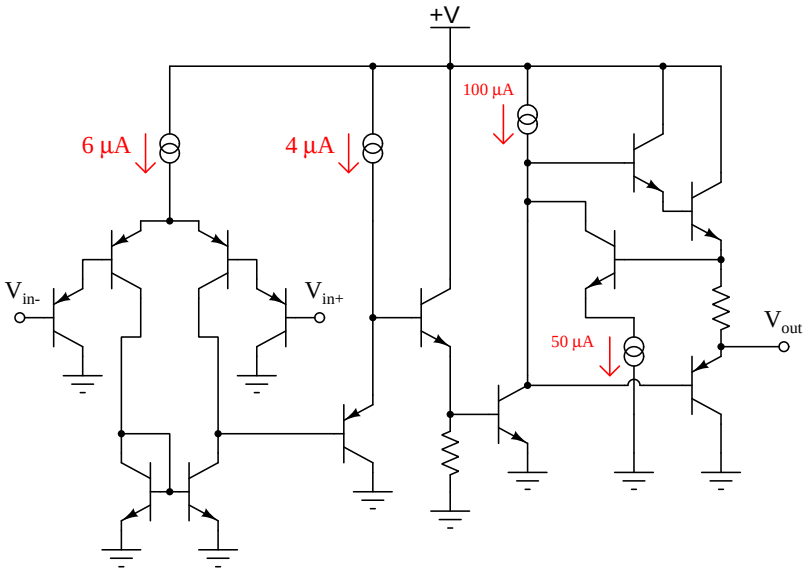
Another practical application of "common-mode voltage" is in differential amplifier circuits, which like the milliVolt meter in the prior example are designed to sense a difference of electrical potential between two input terminals but ignore any voltage with reference to ground or other common points in the circuit. Shown below is a typical illustration of common-mode voltage (V_{CM}) as applied to the input terminals of a differential amplifier (the triangular symbol):



Again, the purpose of any differential voltage amplifier is to respond to V_{diff} but ignore V_{CM}, the ground-referenced output signal V_{out} ideally being a function only of the differential input voltage V_{diff}. This is the ideal situation, but real differential voltage amplifiers do indeed respond (albeit slightly) to common-mode voltage, and if the common-mode voltage becomes too great the amplifier may cease to function entirely. For this reason we must be aware of how common-mode voltage is defined and why it matters to the internal operation of such amplifiers.

If we take this illustration literally, it would mean the lower input terminal of the amplifier experiences V_{CM} with reference to ground, while the upper input terminal experiences V_{diff} more than that, making "common-mode" voltage be the lesser of the two potentials with reference to ground. But, just as in the case of the milliVolt meter in the prior example what might actually *matter* in terms of real consequences might not be voltage measured with reference to ground. In other words, we might express V_{CM} in reference to ground, but there might be some other commonly-shared reference point that actually matters more to the proper operation of the amplifier.

Take for example the internal schematic diagram (simplified) for a model 324 operational amplifier:



Note how the transistors connected to the V_{in+} and V_{in-} input terminals of the amplifier are part of a Darlington pair array, with emitter arrows pointed in such a direction to indicate base currents must flow *out* from the amplifier to whatever device(s) provide the input terminal potentials. In order to ensure base current flows through each of these input transistors, the voltage as measured between either of these terminals and the amplifier's negative DC power terminal (i.e. in this case, ground) must never rise so high as to leave too little voltage across the base-emitter PN junctions of all four input transistors and the 6 μA current regulator for them to operate. This means the electrical potential of either input terminal must always be more than 1.4 Volts *lower* than the +V DC power supply "rail" potential in order to forward-bias two PN junctions (0.7 Volts each) plus provide enough voltage drop across the 6 μA current regulator for it to do its job properly.

One datasheet I consulted on the model LM324 operational amplifier states the common-mode input voltage range as such:

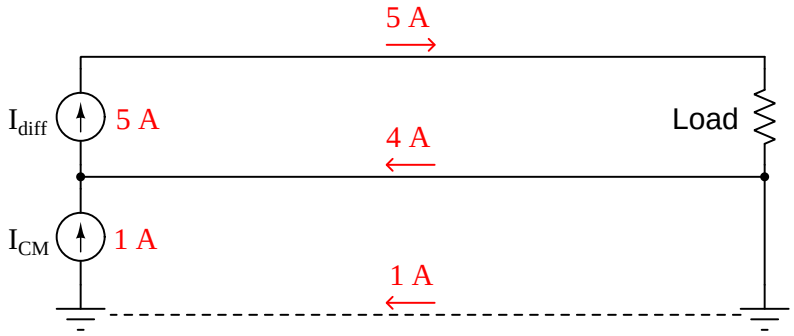
> Maximum V_{CM} common-mode voltage = V_{CC} − 2 Volts *(Where V_{CC} is the positive DC power rail)*

According to this datasheet, this model 324 amplifier circuit requires at least 2 Volts of drop spread between the two base-emitter PN junctions and the 6 μA current regulator in order to maintain a healthy amount of current through them all. So, even though we may define V_{CM} in terms of voltage measured between any amplifier input terminal and ground, what actually matters is maintaining a minimum common-mode voltage drop between any input terminal and the +V DC power supply terminal.

The moral of this story is that common-mode voltage may be defined in more than one way, and also that the way in which we define how it's measured may not necessarily directly relate

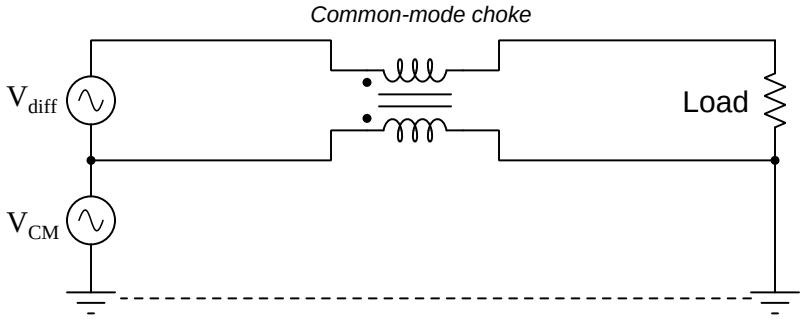
to the threat posed by having too much (or not enough) common-mode voltage! As usual, first principles matter: in the case of the milliVolt meter connected across a high-side shunt resistor we were concerned with possibly breaking down electrical insulation with too much potential difference between a test lead and the meter's metal case; with the model 324 operational amplifier we were concerned with keeping the input transistors in their active operational modes (conducting current).

In case you have been wondering, yes there is such a thing as common-mode *current* as well as common-mode voltage. In the case of common-mode current, it is defined as the net amount of electrical current among a multiple parallel conductors. An illustration is shown below:

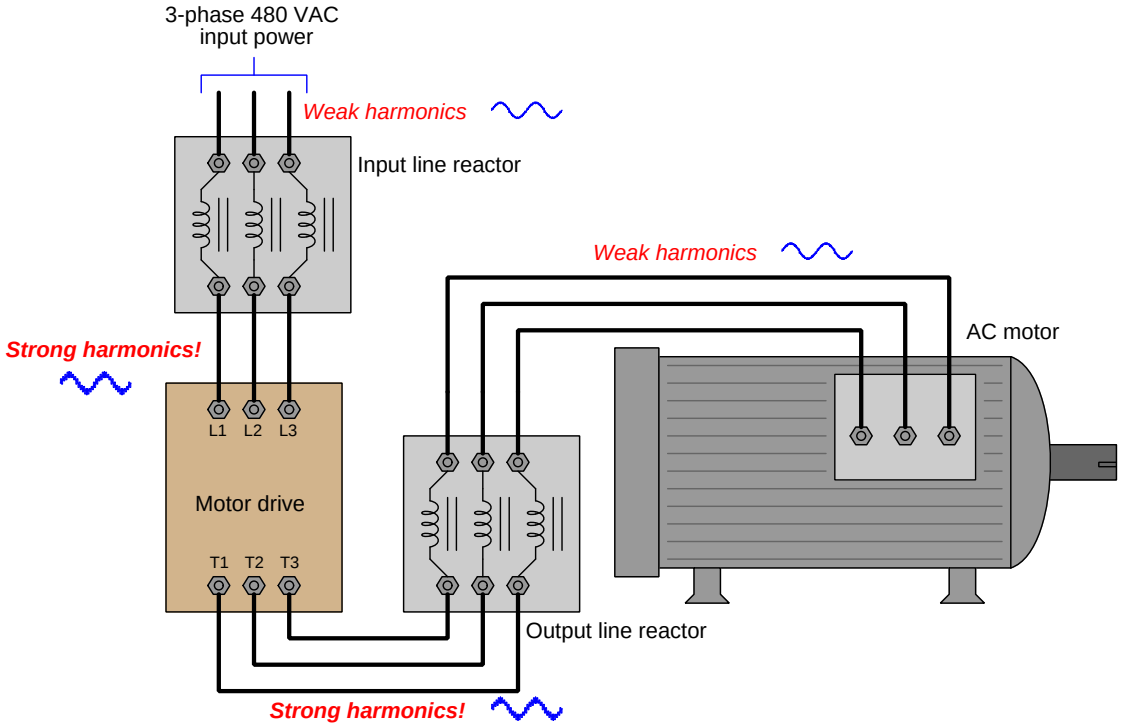


Here the differential current seen by the load is 5 Amperes, but the common-mode current through the two wires connected to that load is 1 Ampere.

In AC circuit applications where common-mode currents are undesirable, a popular way to attenuate them is to use a pair of magnetically-coupled inductors as a *common-mode choke*. The operating principle here is that the inductors' magnetic fields will be equal and opposite – and therefore completely cancel out – for differential current, but will aid each other to create a high impedance standing in the way of any common-mode current:



Common-mode chokes are often found in three-phase AC power systems where *triplen harmonics*[1] are found, as these harmonic currents exhibit zero phase shift between each other and so always flow in the same direction at any given time through the three power conductors. An example circuit shown below places two such three-phase common-mode chokes on either side of a variable-frequency motor drive (VFD):



[1] A "triplen" harmonic is any harmonic that is a multiple of three; e.g. 3rd, 6th, 9th, 12th, etc.

5.3 General-purpose opamp recommendations

Operational amplifiers come in many different sizes, models, and capabilities, making selection by hobbyists and students daunting if all they wish to do is build a simple DC or audio-frequency AC proof-of-concept amplifier circuit. The following is a list of popular general-purpose opamps available in DIP packaging[2] that I recommend for students and hobbyists just beginning their exploration of these amazingly capable circuits.

In all cases *please consult manufacturer datasheets for specific ratings* as the information I provide here is generic and may not cover the exact model you happen to have available to you! A good example of this caveat is output current limitations, which are usually dependent on the DC voltage powering the opamp at any given time, and which are typically specified in the form of graphs shown in the datasheet.

5.3.1 TLV27x series

These general-purpose MOSFET-input opamps are available in one, two, or four amplifiers per package. The last digit of the part number specifies that number, for example "TLV272" meaning two opamps in that package. They also offer full rail-to-rail output voltage range.

- *DC supply voltage range* = 2.7 to 16 Volts
- *Rail-to-rail output* = Yes
- *Rail-to-rail input* = No
- *Maximum output current* = 12 mA
- *Slew rate* = 2.4 Volts per microsecond
- *Special characteristics* = MOSFET inputs make these opamps susceptible to damage from static electricity (ESD)

[2] Modern mass-produced electronic circuits typically use surface-mount components for their small footprint and ease of automated assembly. However, SMD opamps are difficult to "breadboard" temporary circuits with, DIP-style packages being far superior in this regard.

5.3.2 TLE214x series

These general-purpose bipolar-transistor opamps have an extremely wide DC power supply voltage range and are available in one, two, or four amplifiers per package. The last digit of the part number specifies that number, for example "TLE2142" meaning two opamps in that package.

- *DC supply voltage range* = 4 to 44 Volts
- *Rail-to-rail output* = No
- *Rail-to-rail input* = No
- *Maximum output current* = 20 mA – short-circuit protected
- *Slew rate* = 27 Volts per microsecond
- *Special characteristics* = input offset adjustment capability for the model TLE2141 (two dedicated IC pins for this called "Offset N1" and "Offset N2")

5.3.3 TL07x / TL08x series

These are good general-purpose opamps with JFET inputs available in one, two, or four amplifiers per package. The last digit of the part number specifies that number, for example "TL082" meaning two opamps in that package.

- *DC supply voltage range* = 10 to 30 Volts
- *Rail-to-rail output* = No
- *Rail-to-rail input* = No
- *Maximum output current* = 26 mA – short-circuit protected
- *Slew rate* = 20 Volts per microsecond
- *Input noise voltage* = 9.2 microVolts peak-peak
- *Special characteristics* = input offset adjustment capability for the single-opamp models TL071 and TL081 (two dedicated IC pins for this called "Offset N1" and "Offset N2")

5.3.4 MCP600x series

These are excellent opamps for circuits operating at modest DC power supply voltage levels, such as TTL and low-voltage CMOS logic circuits where the opamp must operate on the same power supply rails as the digital logic. They have MOSFET inputs, and also offer full rail-to-rail output voltage range and rail-to-rail input voltage range which is important given their rather low power supply voltage range.

- *DC supply voltage range* = 1.8 to 6 Volts
- *Rail-to-rail output* = Yes
- *Rail-to-rail input* = Yes
- *Maximum output current* = 23 mA – short-circuit protected
- *Slew rate* = 0.6 Volts per microsecond
- *Input noise voltage* = 6.1 microVolts peak-peak
- *Special characteristics* = MOSFET inputs make these opamps susceptible to damage from static electricity (ESD)

5.3.5 LM158/258/358

This legacy bipolar-transistor opamp design offers two amplifiers in one package.

- *DC supply voltage range* = 3 to 30 Volts
- *Rail-to-rail output* = No
- *Rail-to-rail input* = No
- *Maximum output current* = 20 mA – short-circuit protected
- *Slew rate* = 0.5 Volts per microsecond
- *Input noise voltage* = 3 microVolts peak-peak

5.3.6 LM324

This legacy bipolar-transistor opamp design offers four amplifiers in one package.

- *DC supply voltage range* = 3 to 30 Volts
- *Rail-to-rail output* = No
- *Rail-to-rail input* = No
- *Maximum output current* = 20 mA – short-circuit protected
- *Slew rate* = 0.5 Volts per microsecond
- *Input noise voltage* = 3 microVolts peak-peak

5.3.7 LM741 and LM1458

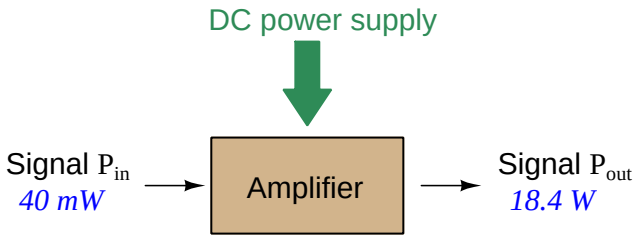
The LM741 has just one amplifier in an 8-pin package. The LM1458 is the same opamp design with two amplifiers in one 8-pin package. These are both legacy designs and should be considered a "last-resort" choice for most any circuit requiring an opamp. The high minimum DC supply voltage requirement and relatively slow slew rate specifications are disqualifying for many circuit applications.

- *DC supply voltage range* = 20 to 44 Volts
- *Rail-to-rail output* = No
- *Rail-to-rail input* = No
- *Maximum output current* = 25 mA – short-circuit protected
- *Slew rate* = 0.5 Volts per microsecond
- *Special characteristics* = input offset adjustment capability for the LM741 (two dedicated IC pins for this called "Offset Null")

5.4 Decibels

One of the mathematical tools popularly used to gauge increases and decreases of electrical power is the *common logarithm*, expressed as a measurement unit called the *decibel*. The basic idea of decibels is to express a ratio of two electrical power quantities in logarithmic terms. Every time you see the unit of "decibel" you can think: *this is an expression of how much greater (or how much smaller) one power is to another*. The only question is which two powers are being compared.

Electronic amplifiers are a type of electrical system where comparisons of power are useful. Students of electronics learn to compare the output power of an amplifier against the input power as a unitless ratio, called a *gain*. Take for example an electronic amplifier with a signal input of 40 milliWatts and a signal output of 18.4 Watts:



$$\text{Gain} = \frac{P_{out}}{P_{in}} = \frac{18.4\ W}{40\ mW} = 460$$

An alternative way to express the gain of this amplifier is to do so using the unit of the *Bel*, defined as the common logarithm of the gain ratio:

$$\log\left(\frac{P_{out}}{P_{in}}\right) = \log\left(\frac{18.4\ \text{W}}{40\ \text{mW}}\right) = 2.66276\ \text{B}$$

When you see an amplifier gain expressed in the unit of "Bel", it's really just a way of saying "The output signal coming from this amplifier is x powers of ten greater than the input signal." An amplifier exhibiting a gain of 1 Bel outputs 10 times as much power as the input signal. An amplifier with a gain of 2 Bels boosts the input signal by a factor of 100. The amplifier shown above, with a gain of 2.66276 Bels, boosts the input signal 460-fold.

At some point in technological history it was decided that the "Bel" (B) was too large and cumbersome, and so it became common to express powers in fractions of a Bel instead: the *deci*Bel (1 dB = $\frac{1}{10}$ of a Bel). Therefore, this is the form of formula you will commonly see for expressing electrical signal power gains or losses:

$$\text{dB} = 10 \log\left(\frac{P_{out}}{P_{in}}\right)$$

The gain of our hypothetical electronic amplifier, therefore, would be more commonly expressed as 26.6276 dB rather than 2.66276 B, although either expression is technically valid[3].

[3] It is interesting to note that although the "Bel" is a metric unit, it is seldom if ever used without the metric prefix

An operation students often struggle with is converting a decibel figure back into a ratio, since the concept of logarithms seems to be universally perplexing. Here I will demonstrate how to algebraically manipulate the decibel formula to solve for the power ratio given a dB figure.

First, we will begin with the decibel formula as given, solving for a value in decibels given a power ratio:

$$\text{dB} = 10 \log(\text{Ratio})$$

If we wish to solve for the ratio, we must "undo" all the mathematical operations surrounding that variable. One way to determine how to do this is to reverse the order of operations we would follow if we knew the ratio and were solving for the dB value. After calculating the ratio, we would then take the logarithm of that value, and then multiply that logarithm by 10: start with the ratio, then take the logarithm, then multiply last. To un-do these operations and solve for the ratio, we must un-do each of these operations in reverse order. First, we must un-do the multiplication (by dividing by 10):

$$\frac{\text{dB}}{10} = \frac{10 \log(\text{Ratio})}{10}$$

$$\frac{\text{dB}}{10} = \log(\text{Ratio})$$

Next, we must un-do the logarithm function by applying its mathematical inverse to both sides of the formula – making each expression a power of 10:

$$10^{\frac{\text{dB}}{10}} = 10^{\log(\text{Ratio})}$$

$$10^{\frac{\text{dB}}{10}} = \text{Ratio}$$

To test our algebra, we can take the previous decibel value for our hypothetical amplifier and see if this new formula yields the original gain ratio:

$$\text{Ratio} = 10^{\frac{26.6276 \text{ dB}}{10}}$$

$$\text{Ratio} = 10^{2.66276 \text{ B}}$$

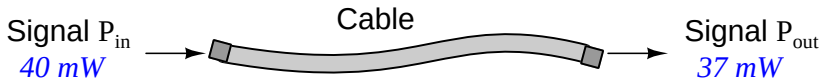
$$\text{Ratio} = 460$$

Sure enough, we arrive at the correct gain ratio of 460, starting with the decibel gain figure of 26.6276 dB.

"deci" ($\frac{1}{10}$). One could express powers in microbels, megabels, or any other metric prefix desired, but it is never done in industry: only the decibel is used.

We may also use decibels to express power *losses* in addition to power *gains*. There are many practical applications of this in signaling systems, both electronic and optical. One such application is *filtering*, where a "filter" circuit screens out certain components of the signal while letting others pass through (e.g. the bass or treble control for an audio system). Another application is *attenuation*, where the entirety of a signal is reduced in magnitude (e.g. the volume control for an audio system).

We will explore yet another application of signal power reduction as a case study for decibels: *cable loss*. Cables designed to convey signals over long distances are not perfect conduits of energy, as some of the signal's energy is inevitably lost along the way. This is true for different types of signals, electrical and optical being two popular examples. In the following illustration we see a signal cable losing power along its length[4], such that the power out is less than the power in:



$$\text{Gain} = \frac{P_{out}}{P_{in}} = \frac{37\ mW}{40\ mW} = 0.925$$

$$10 \log \left(\frac{P_{out}}{P_{in}} \right) = 10 \log \left(\frac{37 \text{ mW}}{40 \text{ mW}} \right) = -0.3386 \text{ dB}$$

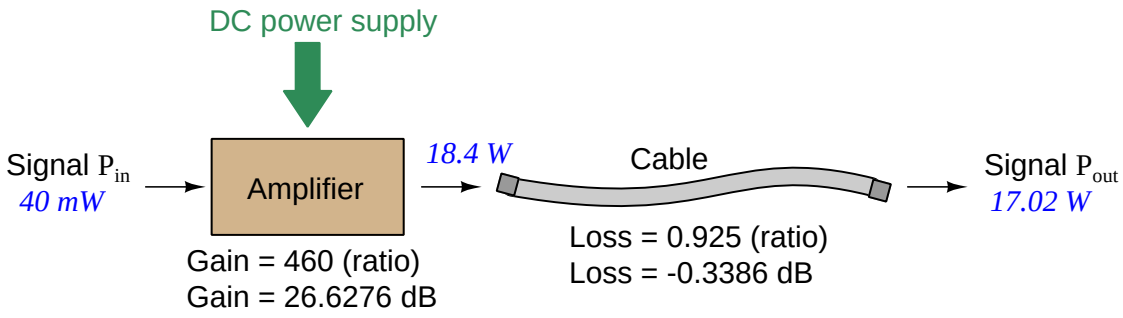
Contrasting this result against the previous result (with the amplifier) we see a very important property of decibel figures: any power *gain* is expressed as a *positive* decibel value, while any power *loss* is expressed as a *negative* decibel value. Any component outputting the exact same power as it takes in will exhibit a "gain" value of 0 dB (equivalent to a gain *ratio* of 1).

Remember that Bels and decibels are nothing more than logarithmic expressions of "greater than" and "less than". Positive values represent powers that are *greater* while negative values represent powers that are *lesser*. Zero Bel or decibel values represent *no change* (neither gain nor loss) in power.

A couple of simple decibel values are useful to remember for approximations, where you need to quickly estimate decibel values from power ratios (or vice-versa). Each addition or subtraction of 10 dB exactly represents a 10-fold multiplication or division of power ratio: e.g. +20 dB represents a power ratio gain of $10 \times 10 = 100$, whereas −30 dB represents a power ratio reduction of $\frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} = \frac{1}{1000}$. Each addition or subtraction of 3 dB approximately represents a 2-fold multiplication or division or power ratio: e.g. +6 dB is approximately equal to a power ratio gain of $2 \times 2 = 4$, whereas −12 dB is approximately equal to a power ratio reduction of $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16}$. We may combine ± 10 dB and ± 3 dB increments to come up with ratios that are products of 10 and 2: e.g. +26 dB is approximately equal to a power ratio gain of $10 \times 10 \times 2 \times 2 = 400$.

[4]For high-frequency signals such as those used in radio communications, the dominant mode of energy dissipation is *dielectric heating*, where the AC electric field between the cable conductors excites the molecules of the conductor insulation. This energy loss manifests as heat, which explains why there is less signal energy present at the load end of the cable than is input at the source end of the cable. For DC and low-frequency AC circuits the dominant mode of energy dissipation is cable conductor resistance, which is typically very small.

Observe what happens if we combine a "gain" component with a "loss" component and calculate the overall power out versus power in:



The overall gain of this amplifier and cable system expressed as a ratio is equal to the *product* of the individual component gain/loss ratios. That is, the gain ratio of the amplifier *multiplied* by the loss ratio of the cable yields the overall power ratio for the system:

$$\text{Overall gain} = \frac{17.02 \text{ W}}{40 \text{ mW}} = (460)(0.925) = 425.5$$

The overall gain may be alternatively expressed as a decibel figure, in which case it is equal to the *sum* of the individual component decibel values. That is, the decibel gain of the amplifier *added* to the decibel loss of the cable yields the overall decibel figure for the system:

$$\text{Overall gain} = 10 \log \left(\frac{17.02 \text{ W}}{40 \text{ mW}} \right) = 26.6276 \text{ dB} + (-0.3386 \text{ dB}) = 26.2890 \text{ dB}$$

It is often useful to be able to estimate decibel values from power ratios and vice-versa. If we take the gain ratio of this amplifier and cable system (425.5) and round it down to 400, we may easily express this gain ratio as an expanded product of 10 and 2:

$$425.5 \approx 400 = (10) \times (10) \times (2) \times (2)$$

Knowing that every 10-fold multiplication of power ratio is an addition of +10 dB, and that every 2-fold multiplication of power is an addition of +3 dB, we may express the expanded product as a sum of decibel values:

$$(10) \times (10) \times (2) \times (2) = (10 \text{ dB}) + (10 \text{ dB}) + (3 \text{ dB}) + (3 \text{ dB}) = 26 \text{ dB}$$

Therefore, our power ratio of 425.5 is approximately equal to +26 decibels.

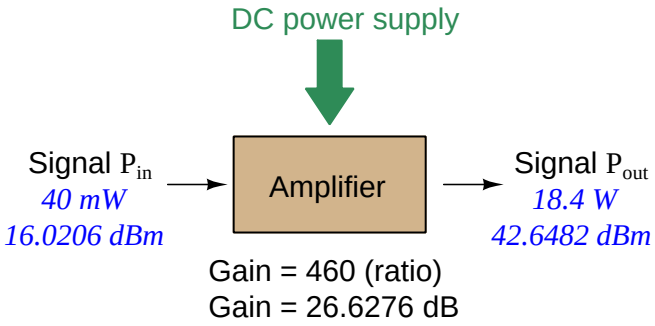
Decibels always represent comparisons of power, but that comparison need not always be P_{out}/P_{in} for a system component. We may also use decibels to express an amount of power compared to some standard reference. If, for example, we wished to express the input power to our hypothetical amplifier (40 milliWatts) using decibels, we could do so by comparing 40 mW against a standard "reference" power of exactly 1 milliWatt. The resulting decibel figure would be written as "dBm" in honor of the 1 *m*illiWatt reference:

$$P_{in} = 10 \log \left(\frac{40 \text{ mW}}{1 \text{ mW}} \right) = 16.0206 \text{ dBm}$$

The unit of "dBm" literally means the amount of dB "greater than" 1 milliWatt. In this case, our input signal of 40 milliWatts is 16.0206 dB greater than a standard reference power of exactly 1 milliWatt. The output power of that amplifier (18.4 Watts) may be expressed in dBm as well:

$$P_{out} = 10 \log \left(\frac{18.4 \text{ W}}{1 \text{ mW}} \right) = 42.6482 \text{ dBm}$$

A signal power of 18.4 Watts is 42.6482 dB greater than a standard reference power of exactly 1 milliWatt, and so it has a decibel value of 42.6482 dBm.



Notice how the output and input powers expressed in dBm relate to the power gain of the amplifier. Taking the input power and simply *adding* the amplifier's gain factor yields the amplifier's output power in dBm:

$$P_{in}(\text{dB}) + P_{gain}(\text{dB}) = P_{out}(\text{dB})$$

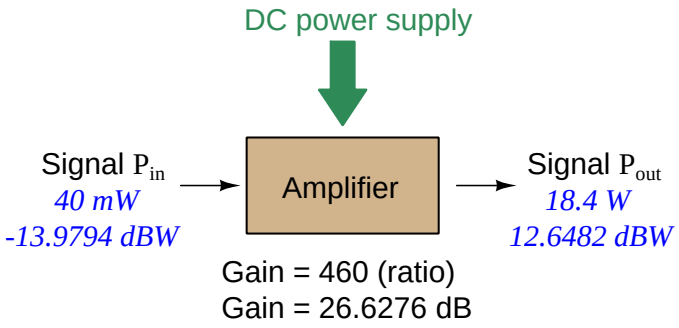
$$16.0206 \text{ dBm} + 26.6276 \text{ dB} = 42.6482 \text{ dBm}$$

An electronic signal that begins 16.0206 dB greater than 1 milliWatt, when boosted by an amplifier gain of 26.6276 dB, will become 42.6482 dB greater than the original reference power of 1 milliWatt.

We may alternatively express all powers in this hypothetical amplifier in reference to a 1-Watt standard power, with the resulting power expressed in units of "dBW" (decibels greater than 1 Watt):

$$P_{in} = 10 \log \left(\frac{40 \text{ mW}}{1 \text{ W}} \right) = -13.9794 \text{ dBW}$$

$$P_{out} = 10 \log \left(\frac{18.4 \text{ W}}{1 \text{ W}} \right) = 12.6482 \text{ dBW}$$



Note how the input power of 40 milliWatts equates to a negative dBW figure because 40 milliWatts is *less* than the 1 Watt reference, and how the output power of 18.4 Watts equates to a positive dBW figure because 18.4 Watts is *more* than the 1 Watt reference. A positive dB figure means "more than" while a negative dB figure means "less than."

Note also how the output and input powers expressed in dBW still relate to the power gain of the amplifier by simple addition, just as they did when previously expressed in units of dBm. Taking the input power in units of dBW and simply *adding* the amplifier's gain factor yields the amplifier's output power in dBW:

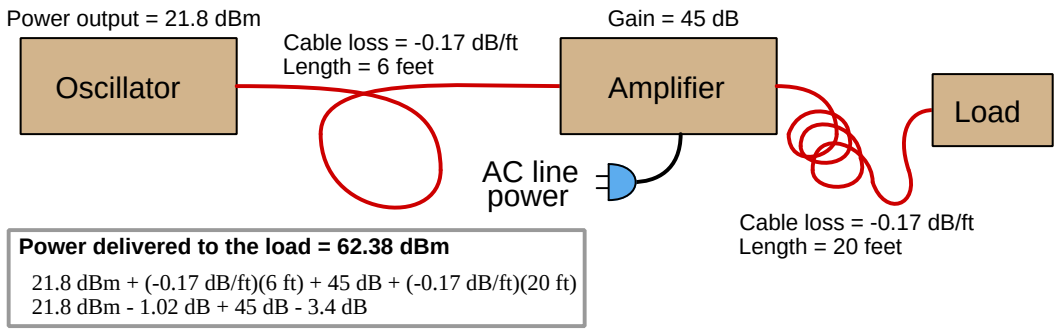
$$P_{in}(\text{dB}) + P_{gain}(\text{dB}) = P_{out}(\text{dB})$$

$$-13.9794 \text{ dBW} + 26.6276 \text{ dB} = 12.6482 \text{ dBW}$$

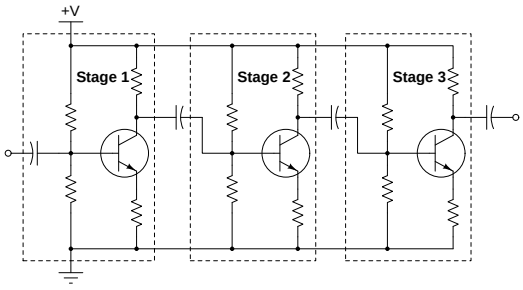
An electronic signal that begins 13.9794 dB less than 1 Watt, when boosted by an amplifier gain of 26.6276 dB, will become 12.6482 dB greater than the original reference power of 1 Watt.

This is one of the major benefits of using decibels to express powers: we may very easily calculate power gains and losses by summing a string of dB figures, each dB figure representing the power gain or power loss of a different system component. Normally, any compounding of *ratios* involves multiplication and/or division of those ratios, but with decibels we may simply add and subtract. One of the interesting mathematical properties of logarithms is that they "transform[5]" one type of problem into a simpler type: in this case, a problem of multiplying ratios into a (simpler) problem of adding decibel figures.

For example, we may express the power dissipated along a cable in terms of decibels per foot; the longer the cable, of course, the more power will be lost this way, all other factors being equal. For example, a radio-frequency signal cable having a loss figure of −0.15 decibels per foot at a signal frequency of 2.4 GHz will suffer −15 dB over 100 feet, and −150 dB over 1000 feet. To illustrate how decibels may be used to calculate power delivered to a load in such a system, accounting for various gains and losses along the way using decibel figures:



A similar application of decibels is found in multi-stage amplifier circuits, where one stage amplifies a signal to be fed into a successive stage to be amplified more. The power gains of these stages, each expressed as a ratio, *multiply* to make the over-all amplifier's power gain (ratio). The power gains of those same stages, each expressed as a decibel figure, *add* to make the over-all amplifier's power gain (dB):



[5] In fact, logarithms are one of the simplest examples of a *transform function*, converting one type of mathematical problem into another type. Other examples of mathematical transform functions used in engineering include the *Fourier transform* (converting a time-domain function into a frequency-domain function) and the *Laplace transform* (converting a differential equation into an algebraic equation).

Another common application of decibels is to express ratios of voltage and/or current rather than power. However, since the unit of the Bel has been defined as an expression of a *power* ratio, we cannot simply substitute V or I for P in any of the formulae we've seen so far.

Suppose an amplifier has a voltage gain of 2 (i.e. V_{out} is twice as large as V_{in}), and we would like to express this gain in decibels. Since decibels are intended to express power gain and not voltage gain, we must figure out how much power gain is equivalent to a voltage gain of two. Obviously, voltage and power are fundamentally different quantities, but if we imagine ourselves connecting a fixed load resistance to the input signal, and then to the output signal, we will realize that load's power dissipation will be more than double when energized by a voltage twice as large. Joule's Law is helpful to determine the exact ratio of power dissipation:

$$P = \frac{V^2}{R}$$

Doubling voltage for any given load resistance results in power *quadrupling* because power is proportional to the square of the voltage applied to a fixed resistance. Using this as the basis for applying decibels to a voltage ratio. Knowing that Joule's Law also declares power is proportional to the square of the current applied to a fixed resistance ($P = I^2R$) means this same mathematical relationship will apply to current gains and reductions as well as voltage gains and reductions:

$$\text{dB} = 10 \log \left(\frac{P_{out}}{P_{in}} \right) = 10 \log \left(\frac{V_{out}}{V_{in}} \right)^2 = 10 \log \left(\frac{I_{out}}{I_{in}} \right)^2$$

An algebraic identity of logarithms is that the logarithm of any quantity raised to a power is equal to that power multiplied by the logarithm of the quantity. Expressed in general terms:

$$\log x^y = y \log x$$

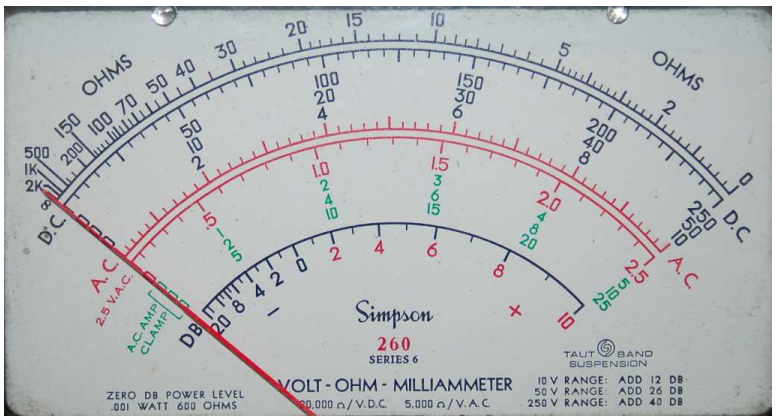
Therefore, we may simplify the decibel formula for voltage gain by removing the "2" power and making it a multiplier:

$$10 \log \left(\frac{V_{out}}{V_{in}} \right)^2 = (2)(10) \log \left(\frac{V_{out}}{V_{in}} \right) = 20 \log \left(\frac{V_{out}}{V_{in}} \right)$$

$$10 \log \left(\frac{I_{out}}{I_{in}} \right)^2 = (2)(10) \log \left(\frac{I_{out}}{I_{in}} \right) = 20 \log \left(\frac{I_{out}}{I_{in}} \right)$$

Thus, we may use decibels to express voltage or current ratios if we simply substitute 20 instead of 10 as the multiplier.

We can see the practicality of using decibels to represent something other than electrical *power* by examining this analog meter face, belonging to a Simpson model 260 VOM (Volt-Ohm-Milliammeter). Note the bottom scale on this meter's face, calibrated in decibels (DB):



Pay attention to the note on decibels written in the lower-left corner of the meter face, where 0 dB is defined as 0.001 Watt dissipated by 600 Ohms. The fact that 0 dB is defined as 1 milliWatt means it should (properly) be labeled dBm rather than dB[6]. A load resistance value is necessary as part of this definition for dB because this meter cannot measure power directly but must infer signal power from measurements of AC *voltage*. Without a specific load resistance, there is no clear relation between voltage and power. 600 Ohms is an old telecommunications standard for audio-frequency AC signals, and continues to be used today for voltage-based decibel measurements of audio-frequency AC signals.

The meter as shown is connected to nothing at all, and so registers 0 Volts AC. This, of course, corresponds to zero power, and it has no corresponding decibel value because the logarithm of zero is mathematically *undefined*[7]. Practically, it means $-\infty$ dB, which is why the needle at the 0 Volt position "falls off" the left-hand end of the dB scale.

Close inspection of the dB scale on this meter face reveals another interesting property of decibels, and that is the *nonlinear* nature of the dB scale. This contrasts starkly against all the voltage and current scales on this meter face which are linear. This nonlinearity is a fundamental property of decibels because it is based on the *logarithm* function which is nonlinear.

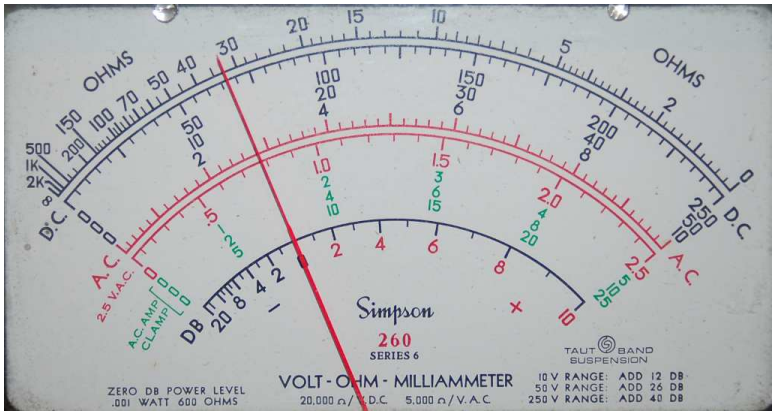
[6] Such mis-labeling is not that uncommon in the profession, the expectation being that the technician or engineer working with the instrument ought to be familiar enough with the concept of decibels to know when dB really means dBm, or dBW, etc.

[7] Your electronic calculator will complain if you attempt to take the logarithm of zero!

Now, we will explore what is necessary to make this meter register 0 dBm (i.e. 1 milliWatt) with an applied AC voltage. 1 milliWatt of power dissipated by 600 Ohms is equivalent to:

$$V = \sqrt{PR} = \sqrt{(0.001)(600)} = 0.7746 \text{ Volts}$$

Setting the VOM to the 2.5 VAC range and applying just enough AC voltage to bring the needle to the 0 dB mark allows us to verify that this is indeed equivalent to just under 0.8 Volts (read on the 2.5 VAC scale):



In the lower-right corner of the meter face we see some notes regarding correction values for decibel measurements when using different AC voltage ranges. The dB scale is read directly when the meter is set on the 2.5 VAC range. When set on the 10 VAC range (i.e. a range *four times* as great), the meter's needle will experience a deflection one-fourth as much as when set to the 2.5 VAC range, and therefore it will point to a lesser (or even negative) value on the dB scale. Converting a voltage ratio of 0.25 into a decibel figure shows us how much less the needle will register on the dB scale when the voltage range is quadrupled:

$$20 \log \left(\frac{2.5}{10} \right) = -12.04 \text{ dB}$$

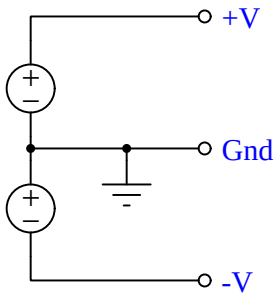
Therefore, when using the 10 VAC range instead of the 2.5 VAC range, one must add 12 dB to the reading. Likewise, we may prove each of the printed correction factors for the alternative voltage-measurement ranges listed (50 Volt AC range and 250 Volt AC range):

$$20 \log \left(\frac{2.5}{50} \right) = -26.02 \text{ dB}$$

$$20 \log \left(\frac{2.5}{250} \right) = -40.0 \text{ dB}$$

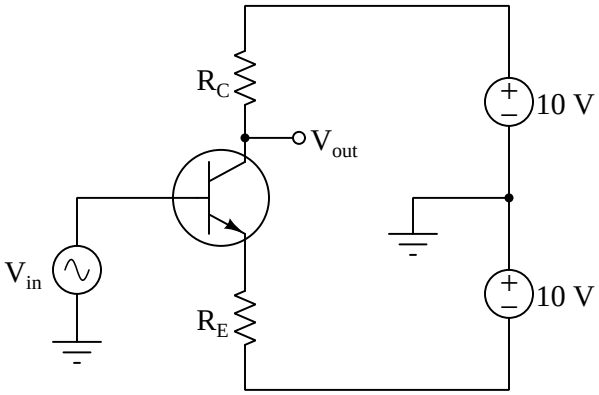
5.5 Split DC power supplies

A popular way to allow electronic amplifier circuits to input and output truly AC (alternating current) signals, as well as handle "bipolar" DC signals (i.e. signals having either positive or negative potential with respect to ground) is to power them using a *split* DC source. This is a DC power supply configuration providing three points of connection and three different voltage levels as shown below:

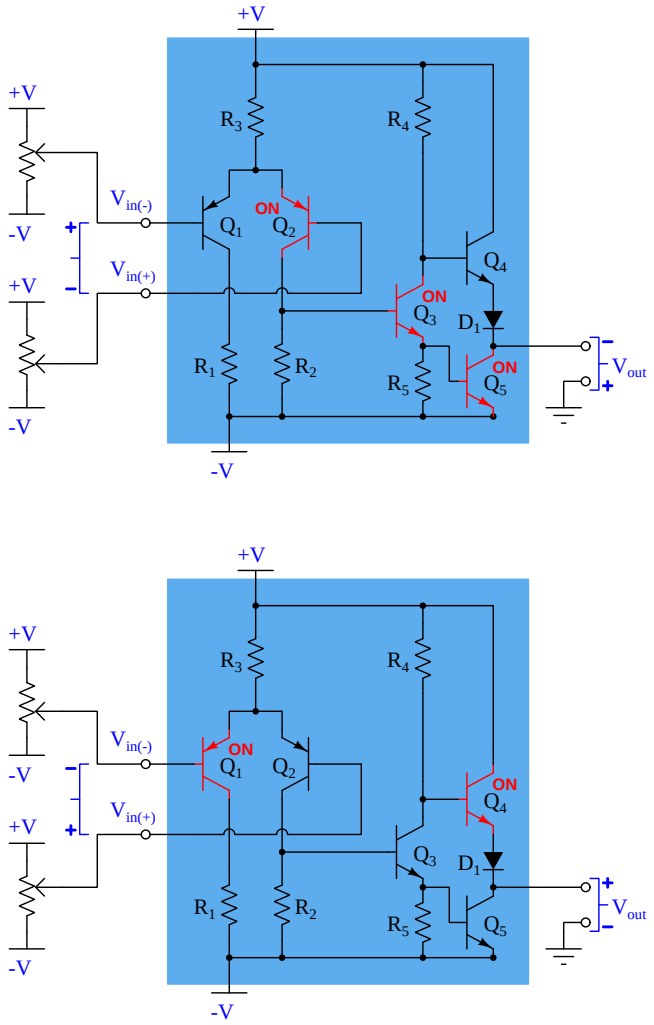


From +V to Ground we measure the voltage of the upper source; from −V to ground we measure the negative voltage of the lower source; between +V and −V we measure those sources' voltage sum. Typical split DC supplies are symmetrical, meaning that +V and −V have the same absolute values.

If we consider a common-emitter BJT amplifier energized by such a supply, we see it becomes possible to achieve Class-A amplification of an AC signal source without the need for biasing resistors anywhere. Moreover, the output voltage is capable of achieving true AC status as well, alternating positive and negative with respect to ground, all because of the split power supply:



In the following diagrams we show the simplified internal schematic diagram of a comparator powered by a split supply, in two "limiting case" conditions. In keeping with the problem-solving strategy of "limiting cases" we will assume all transistor states to be either fully on (saturated) or fully off, marking those saturated-on transistors in red:

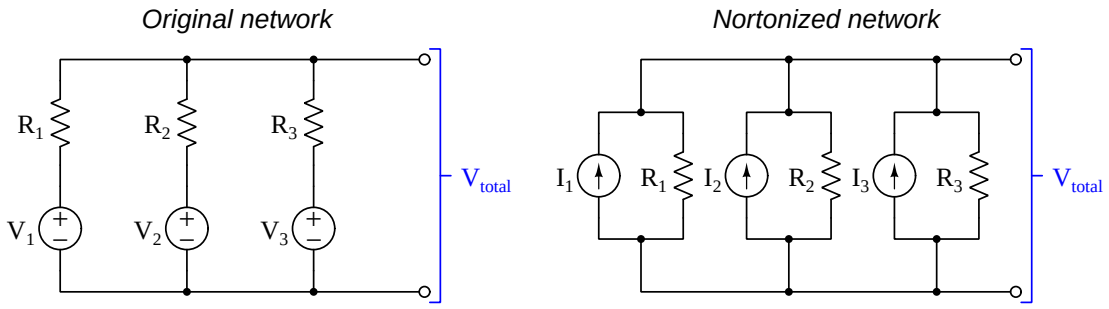


In the upper scenario we see the comparator subjected to a negative differential voltage input (i.e. inverting terminal more positive than noninverting), and this causes its output to saturate low by turning on transistor Q_5. With Q_5 on, the output terminal is nearly equipotential with the negative supply rail ($-$V), and so is negative with respect to ground.

In the lower scenario the comparator experiences a positive differential input (noninverting terminal more positive than inverting), causing Q_4 to saturate on. With Q_4 fully, on the output terminal is nearly equipotential with the positive rail (+V), and so is positive with respect to ground.

5.6 Derivation of Millman's Theorem

Whenever we encounter a network of multiple voltage sources connected together in parallel, each with its own associated series-connected resistance, there is an easy way to calculate that parallel network's voltage value. This method treats each source-resistor pair as a Thévenin voltage source and then converts each into a corresponding Norton current source (current source and parallel-connected resistance). In the converted network all sources and resistors are in parallel with each other, which means all the current source currents simply add together and pass through a parallel-equivalent resistance value determined by the individual resistors:



Voltage across the Nortonized network is fairly simple to calculate, multiplying total current by the total (parallel-diminished) resistance in accordance with Ohm's Law ($V = \frac{I}{R}$):

$$V_{total} = \frac{I_1 + I_2 + I_3}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}}$$

Expanding back to the original network with three voltage sources, we must re-write the numerator of the fraction to show each Norton-equivalent current as the quotient of a voltage source and its respective resistance:

$$V_{total} = \frac{\frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3}}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}}$$

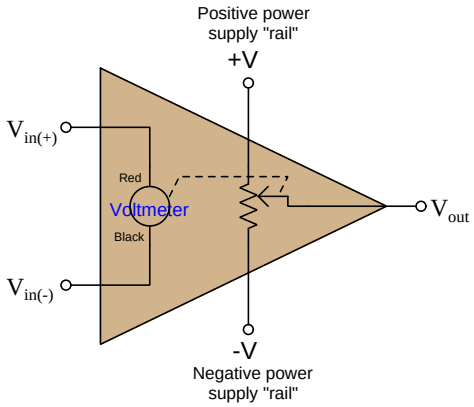
That last equation is the mathematical expression of *Millman's Theorem*, useful for predicting voltage in any network comprised of parallel-connected voltage sources (with individual resistances).

5.7 Operational amplifier imperfections

Comparators and operational amplifiers are incredibly useful and practical as electronic circuit building-blocks, but they are not perfect. For a great many applications these imperfections are insignificant, but for others they may be debilitating. Therefore, it is important to understand the ways in which operational amplifiers deviate from ideal behavior.

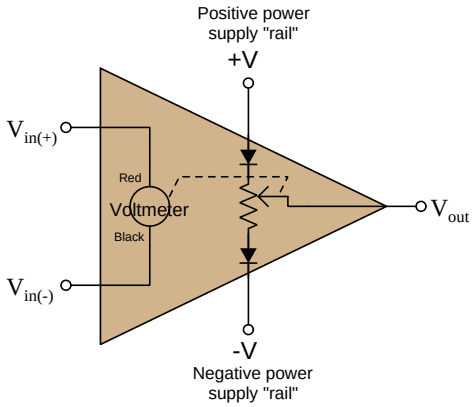
5.7.1 Output voltage limits

The output voltage of any amplifier circuit cannot exceed its DC supply "rail" voltages, because all an amplifier does is *modulate* its supply voltage. A useful model of a differential amplifier is that of a potentiometer with its wiper driven by the pointer of an analog voltmeter driven by the differential voltage applied to the input terminals:



No matter how much voltage is applied between the + and − input terminals, the potentiometer can only "swing" as far as the +V and −V "rails" of the DC power supply. This behavior is often referred to as *rail-to-rail output*, where the V_{out} limits are exactly (or very nearly) equal to the DC rail potentials.

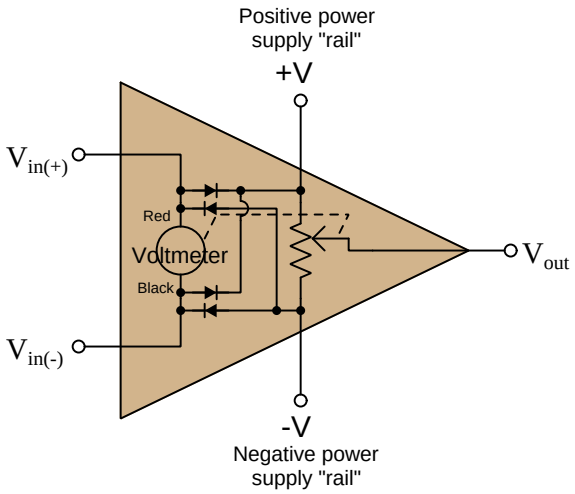
However, many practical amplifier models cannot even swing that far. A more accurate model for many is one where the potentiometer cannot quite reach the +V and −V rails:



With the diodes' voltage drops in effect, V_{out} is only able to "swing" to within approximately 0.7 Volts of each rail. Such internal voltage drops are typical of comparators and operational amplifiers using BJTs in the final output stage. Rail-to-rail capability demands FET output transistors which are capable of operating with mere milliVolts of drain-source voltage drop.

5.7.2 Input voltage limits and latch-up

As a general rule, input voltages for any integrated circuit should remain within the limits of its DC power supply "rail" potentials. Many integrated circuits including comparators and operational amplifiers include "protection" diodes at the input terminals effectively *clamping* input terminal potentials to the power supply rails (to within fractions of a Volt). Expanding upon our previous circuit model, we may show these protection diodes (two per input terminal) connected such that one will begin conducting if the potential at any input exceeds either power supply rail, essentially making that input terminal the new "rail" to power the amplifier:



If the common-mode[8] voltage applied to the inputs of a differential amplifier exceeds the point at which both transistors in the differential pair subcircuit turn off, the amplifier's output may saturate rather than respond properly to the differential input voltage. In some designs this may even lead to a condition where the output remains "latched" in this saturated mode even after the offending input condition ceases – a phenomenon known as *latch-up*[9].

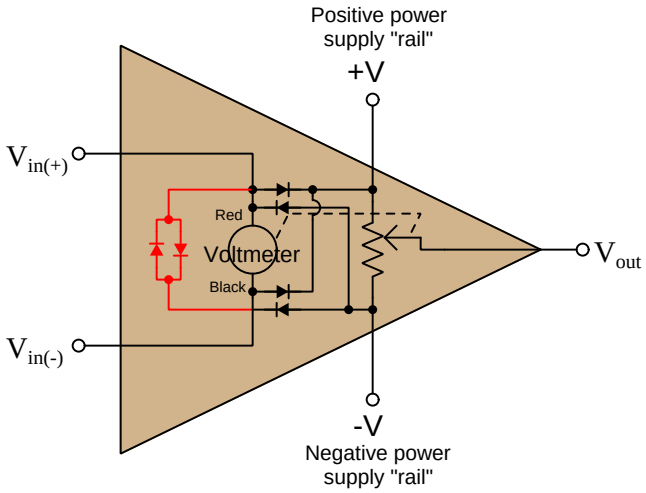
Amplifiers may fail to function properly even with input voltage values falling well within their supplied rail voltages. For example, the popular TL08x series of operational amplifiers specify common-mode voltage limits between +4 Volts above the negative rail voltage and −4 Volts below the positive rail voltage: if we happened to power one of these opamps on ± 12 Volt rails, the permissible input signal voltage range for either of its inputs would be −8 Volts to +8 Volts. In

[8] For a differential amplifier, the *common-mode voltage* is the amount of voltage with respect to ground that is shared among both input terminals, typically calculated as the average of the two inputs' ground-referenced voltages. In other words, if $V_{diff} = V_{in(+)} - V_{in(-)}$ then $V_{CM} = \frac{V_{in(+)} + V_{in(-)}}{2}$. However, for the sake of determining safe limits for amplifier input voltage we may more simply define "common-mode voltage" as being any voltage measured with respect to ground for either of a differential amplifier's two inputs that may be either high or low enough to cause a problem for that amplifier.

[9] Interestingly, the term "latch-up" more properly describes a condition in an integrated circuit where the particularly layering of semiconductor materials forms a PNPN (thyristor) structure accidently triggered by certain abnormal input conditions. Like the "latched" state just described, this triggered-thyristor state can only be reset by cycling power to the circuit.

contrast, other opamp models such as the MCP600x series offer *rail-to-rail inputs* capable of receiving any common-mode voltage up to and including the two DC supply rail voltages.

Some operational amplifiers additionally clamp *differential* input signals using a pair of paralleled protection diodes as shown in red below on the simplified opamp model:

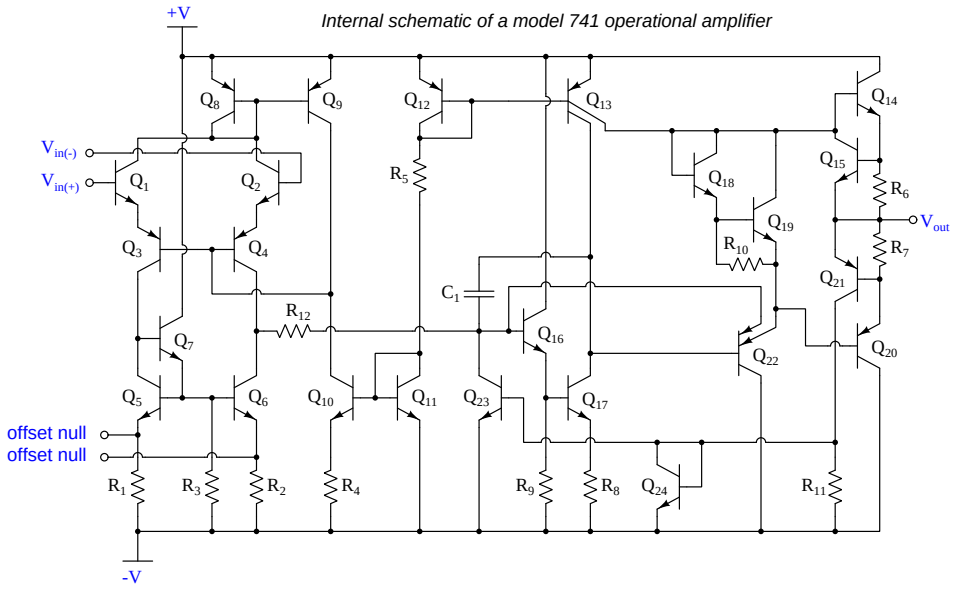


These differential protection diodes attempt to clamp differential voltage to a maximum of 0.7 Volts, which for a typical opamp employing negative feedback is perfectly acceptable because negative feedback should maintain the differential voltage at approximately zero anyway. However, if one attempts to use such a differential-clamped opamp as a *comparator* rather than as an operational amplifier with negative feedback to compare two input potentials differing by more than 0.7 Volts, improper circuit operation will result, with possible damage to the opamp as well!

5.7.3 Input bias current

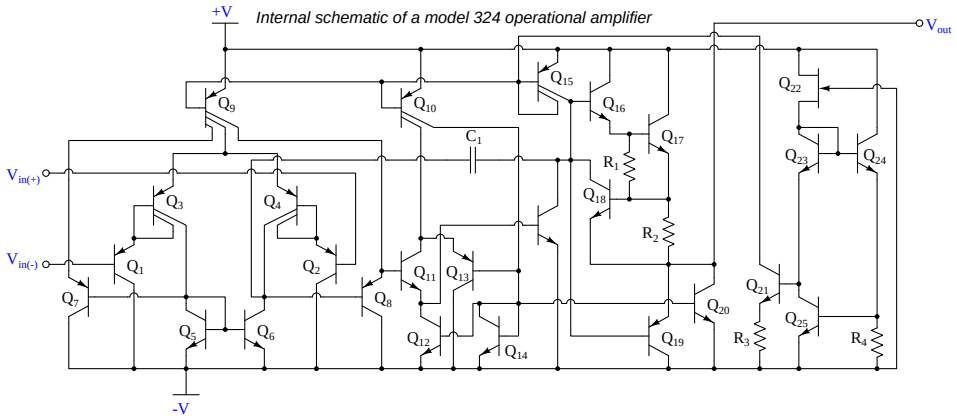
A generally safe assumption for most circuits is that the input terminals of a comparator or operational amplifier carry no current. Like the input of an ideal voltmeter, we assume these inputs have infinite input impedance and therefore pose no load on the sources they sense.

However, this is not entirely true. The input transistor stage of a differential amplifier represents a finite – albeit very large – impedance, and bipolar junction transistors do require *some* amount of base current to function. An inspection of the internal circuitry for the model 741 opamp reveals the source of these currents as well as their directions:



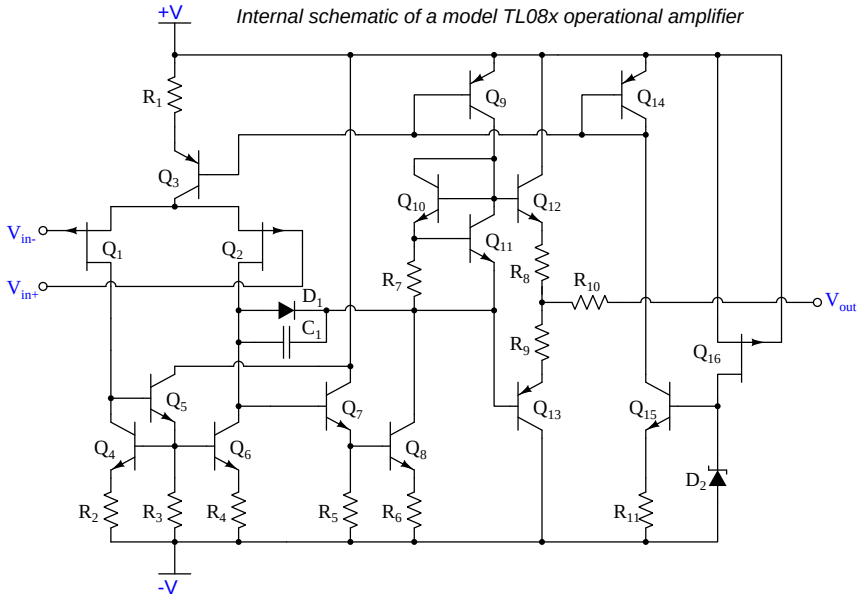
The two input terminals for the 741 connect to the bases of NPN transistors Q_1 and Q_2. In order for these transistors to turn on, current *must* enter the base terminals from outside the opamp. These *bias currents* are quite small – on the order of tens of *nano*Amperes, but they do indeed exist and must find a path through the external circuit.

Another legacy opamp model is the 324. Its internal circuitry is shown below:



Note the two input PNP transistors Q_1 and Q_2. In order for these transistors to turn on, current *must* exit their base terminals and find its way to the negative pole of the power supply through external circuitry.

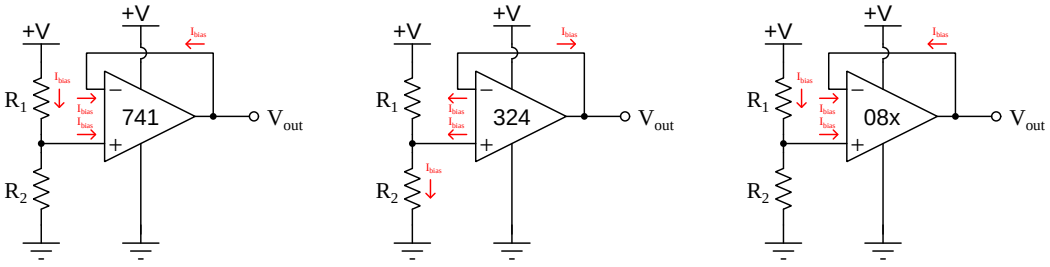
Some comparators and opamps have FET inputs rather than BJT inputs, usually in an attempt to increase input impedance (i.e. decrease bias current). The TL08x opamp series is an example of this design:



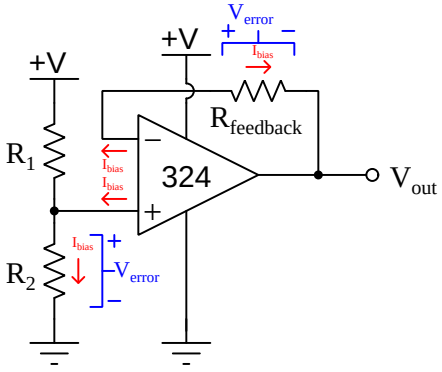
Knowing that JFETs are designed to have their gate-channel PN junctions reverse-biased, we may conclude the two P-channel JFETs Q_1 and Q_2 will exhibit extremely small bias currents *entering* their gate terminals from outside the opamp.

Some opamps utilize MOSFETs for their input transistors in order to maximize input impedance and minimize bias currents, the naturally high impedance of a MOSFET stemming from the layer of metal-oxide insulation separating each MOSFET's gate from the conducting channel beneath. In such opamps the main contribution to input terminal bias current actually comes from reverse-leakage current of *protection diodes* added to the circuit (connecting between each of those input terminals and the DC power rails) to protect the sensitive MOSFET gates from electrostatic discharge, rather than from the transistors themselves.

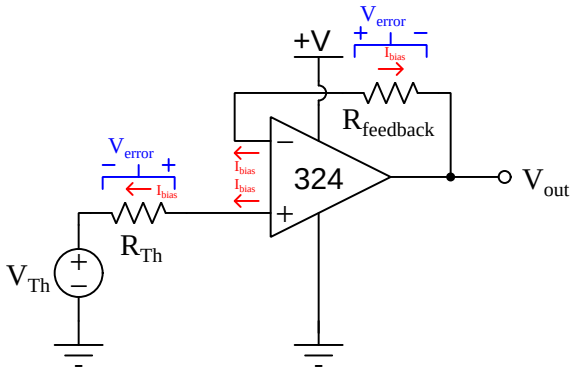
When we connect components to the input terminals of an opamp, we must be sure some pathway exists for these bias currents to flow in the proper directions. Consider the following examples, where a voltage divider provides a voltage signal which each opamp then "buffers" as a voltage follower:



Again, these currents are extremely small in magnitude, and their effects are usually negligible. However, if they must pass through a high impedance, they will cause an undesired voltage drop to develop. A common strategy to mitigate the effects of this voltage drop is to insert an equivalent amount of resistance within the feedback bias current path, assuming the two bias currents will be approximately equal in magnitude, and therefore the *two* unwanted voltage drops should be approximately equal in magnitude but opposite in polarity and so should cancel each other out as common-mode (not differential) voltage at the opamp inputs:



If we regard the voltage divider as a voltage *source* with its own internal Thévenin resistance, we see the bias current must pass through that resistance:



The value of the "source's" Thévenin resistance is equal to the parallel equivalent resistance of R_1 and R_2, and so our feedback resistance needs to be equal to the same:

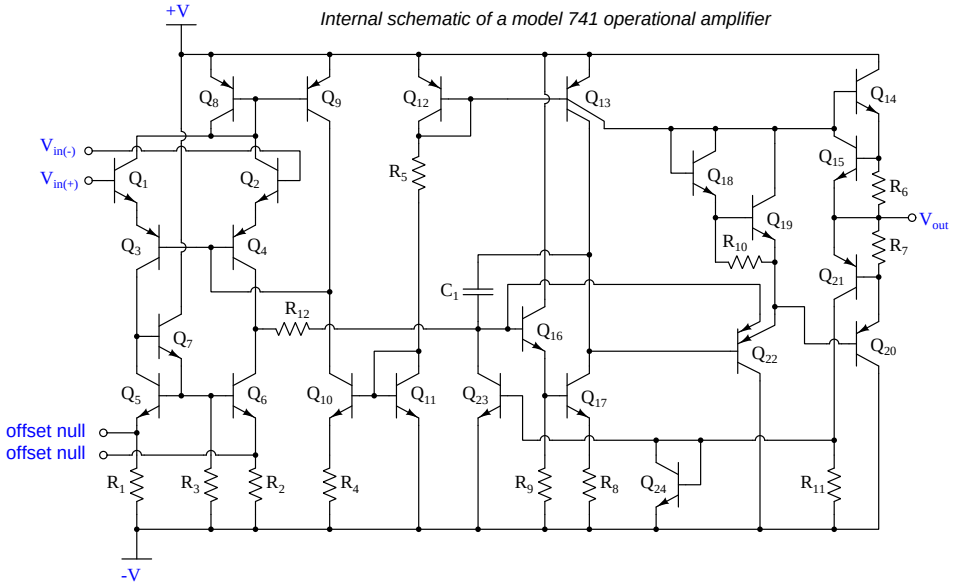
$$R_{feedback} = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2}}$$

It is important to note that the assumption of an opamp's two bias currents being equal is exactly that: an *assumption* and not necessarily reality. Small differences in bias current from one input to the other can exist, and this difference of bias current is called *offset current*.

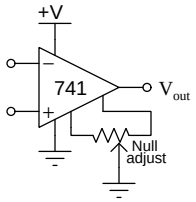
5.7.4 Input offset voltage

If the input terminals of a differential amplifier are shorted together to ensure zero differential input, the amplifier's output should ideally assume a state of zero Volts as well. However, this rarely happens in practice due to asymmetries in the internal circuitry of the amplifier. In practice, a very small amount of differential input voltage is necessary to achieve exactly zero Volts at the output, and this small differential voltage is called the *offset voltage* or *bias voltage* of the amplifier.

Some operational amplifier models provide external terminals for the user to connect a potentiometer, which may be used to "trim" the offset of the amplifier. The legacy 741 opamp is one of those designs, with *offset null* terminals labeled:



Externally, the "nulling" potentiometer connects to these two offset terminals and to the negative power supply terminal:



Essentially, the two portions of the potentiometer become connected in parallel with R_1 and R_2, respectively, inside the opamp. By externally skewing the equivalent resistances of R_1 and R_2, the bias currents through the two halves of the opamp's differential pair input stage may be likewise skewed to achieve the desired negation of offset voltage.

5.7.5 Frequency compensation

A problem common to high-gain amplifier circuits is undesired *oscillation*, and the root of this problem is *positive feedback*: where some of the amplifier's output signal becomes "fed back" to its input with a 360^o phase shift, so that it continually drives itself into oscillation. This principle finds constructive application in sinusoidal oscillator circuits, but for non-oscillator circuits it can be a serious problem.

Most operational amplifier circuits utilize *negative* feedback (i.e. 180^o phase shift) rather than positive feedback, and so it might seem at first oscillations would not be possible in these circuits. However, the 360^o of phase shift necessary for self-sustaining oscillations may occur in a negative-feedback circuit if a total of 180^o *additional* phase shift occurs elsewhere in the opamp circuit. One operating scenario inviting additional phase shift is when the opamp's output must drive a capacitive load. At high enough frequencies, the load's capacitance and stray capacitances in the circuit may conspire to provide the additional 180^o of phase shift necessary to make the total phase shift 360^o.

The criteria necessary for self-oscillation is signal feedback with 360^o of phase shift *and* a total signal gain of at least unity (1). If the 360^o of phase shift cannot be avoided, an effective solution to the problem of unintentional oscillation is to intentionally reduce the opamp's open-loop gain as frequency increases, so that at the frequency necessary to achieve 360^o phase shift the fed-back signal has a voltage gain less than one.

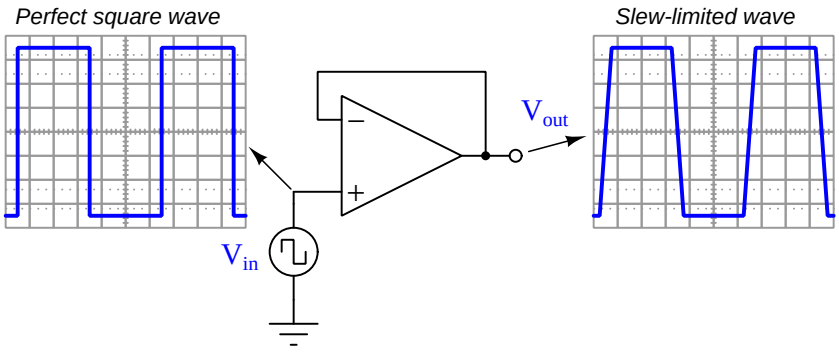
Many operational amplifiers are *internally compensated* for this by the intentional connection of a capacitor between amplification stages. The internal schematic diagrams of three models of opamp shown in previous subsections all contain compensation capacitors. They are easy to spot, being the *only* capacitor in the entire diagram!

5.7.6 Gain-bandwidth product

Operational amplifier open-loop voltage gain is very high, but degrades for multiple reasons as signal frequency increases. The mathematical product of amplifier gain and signal frequency is expressed as the amplifier's *Gain Bandwidth Product*, or *GBWP*.

5.7.7 Slew rate

The maximum rate at which the output voltage of a comparator or opamp is able to rise or fall is called the *slew rate*. If we drive a simple "voltage follower" opamp circuit with a perfect square-wave signal having instantaneous rise and fall times, what we will see on the output is a slew-limited waveform with definite rise and fall *slopes*:



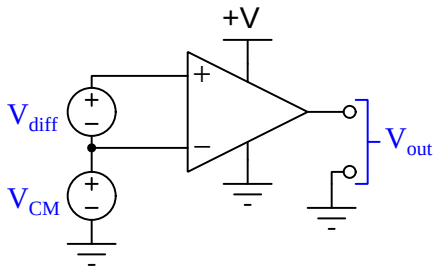
Slope is always measured as a *rise over run*, so for these voltage waveforms we may represent their rising- and falling-edge slopes as a rates-of-change of voltage over time ($\frac{dV}{dt}$), the input waveform having a *infinite* rates of change, and the output waveform having *finite* rates of change. The reason for an opamp's slew rate limitations is based on its internal current limitations and capacitances. Recall the relationship between voltage and current for any capacitance:

$$I_C = C\frac{dV_C}{dt}$$

If a current source drives a constant current through a capacitance, that voltage across that capacitance will either rise or fall (depending on current direction) at a constant *rate*. An opamp's internal transistors are current-limiting devices by nature, and when this limited current energizes or de-energizes capacitance the result is a linearly-sloping waveform rather than an instantaneous jump. A major source of capacitance for any opamp is the *frequency-compensating* capacitance previously discussed, but any capacitance connected to the opamp as part of the load of course also contributes to a limited slew rate. This is why opamp datasheets always specify load capacitance when declaring maximum slew rate (e.g. the model 741 opamp is typically limited to 0.5 Volts per microsecond, or 500,000 Volts per second, while driving a load having no more than 100 picoFarads of capacitance).

5.7.8 Common-mode rejection

Operational amplifiers are *differential* amplifiers, and as such they are only supposed to respond to *differences* in voltage between their two input terminals. Ideally, any common-mode[10] voltage is ignored (i.e. *rejected*) by the amplifier:



In other words, V_{out} for an ideal opamp is strictly a function of V_{diff} and is entirely independent of V_{CM}. However, real opamps are not ideal, and as such their output voltages will be slightly affected by V_{CM}. We may express the relationship between V_{out} and V_{CM} as a *gain*, just as we do for differential voltage:

$$A_{CM} = \frac{\Delta V_{out}}{\Delta V_{CM}} \qquad A_{diff} = \frac{\Delta V_{out}}{\Delta V_{diff}}$$

Differential voltage gain (A_{diff}) is typically referred to as open-loop voltage gain (A_{OL}) and for most operational amplifiers is an extremely large number. Common-mode voltage gain (A_{CM}) by contrast is much smaller (ideally being zero).

The degree to which a differential amplifier rejects a common-mode signal is typically gauged in comparison to its differential voltage gain. The ratio of A_{diff} to A_{CM} is called the *common-mode rejection ratio*, or *CMRR*:

$$\text{CMRR} = \frac{A_{diff}}{A_{CM}}$$

CMRR is more often found expressed as a decibel figure, so:

$$\text{CMRR(dB)} = 20 \log \left(\frac{A_{diff}}{A_{CM}} \right)$$

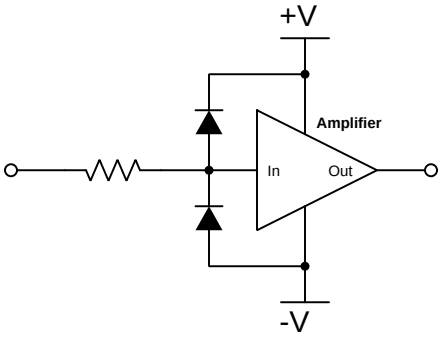
[10] Here, the term "common-mode" refers to a quantity shared in common by both input terminals of the amplifier. This is distinct from "differential" which by definition means something *different* or *distinct* between the input terminals.

5.8 Protecting opamp inputs from over-voltage

Electronic amplifier circuits are extremely useful devices, comprising portions of many practical integrated-circuit (IC) electronic components such as digital logic gates, comparators, operational amplifiers, signal mixers, etc. Solid-state amplifier circuits use *transistors* to allow one electrical signal to control another, and these transistors tend to be susceptible to damage from excessive applied signal voltage. In the case of MOSFET transistors, an excess of signal voltage may puncture the extremely thin layer of metal-oxide insulation separating the transistor's gate terminal from its current-carrying channel. In the case of bipolar (NPN, PNP) transistors, an excess of signal voltage may break down reverse-biased PN junctions inside the transistor, often causing them to fail in "shorted" states. Excessive applied signal voltage may also cause damage to bipolar transistors if forward-biasing PN junctions to the extent that they conduct high amounts of current which may lead to thermal damage.

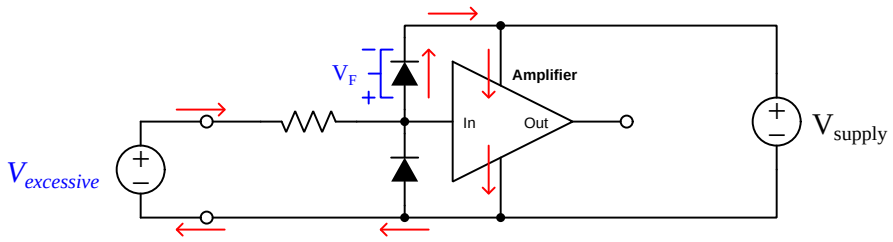
Sources of excessive signal voltage may be broadly categorized as *electro-static discharge* (ESD) or *electrical over-stress* (EOS). ESD happens when some object external to the circuit (including human bodies) accumulates an electro-static charge in its own capacitance, and this capacitively-stored charge is suddenly send to the circuit where it drops a voltage across circuit components high enough to cause damage. EOS is a more broad description of any over-voltage or over-current condition caused by one circuit sourcing energy to another, such as when an electrical test instrument is connected to a signal source that is too great for that instrument to handle.

A simple and effective way to limit both applied voltage and applied current to the input of any electronic amplifier is to connect a current-limiting resistor in series between the input terminal and the amplifier, as well as connect diodes between the input terminal and either DC power supply "rail" terminal:



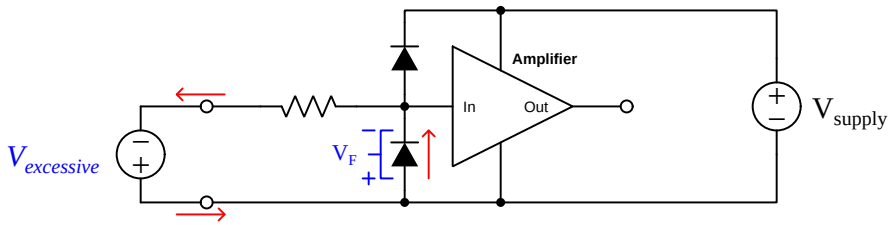
It matters not how the amplifier itself is internally constructed, nor what its larger purpose is – whether it be part of a digital logic gate, a comparator, an operational amplifier, or something even more complex. The purpose of these two diodes and the series resistor is to prevent any component within the amplifier from experiencing either too much voltage and/or too much current. Such resistor-diode networks are generally known as *clipping networks* because they truncate or "clip" an incoming signal's amplitude so as to never exceed some maximum value(s).

To understand how this protection network functions, consider the following circumstance where an excessive positive potential is applied to the input terminal by some external source. Here, "excessive" is defined as any voltage greater in magnitude than the DC power source energizing the amplifier:



The upper diode of the protection network forward-biases to permit current from the offending source to pass into the positive power bus of the amplifier. The resistor's function is to drop however much of the applied voltage exceeds the power supply rail potential plus the protection diode's forward voltage drop ($V_{supply} + V_F$), and to limit the current to a value acceptable to the protection diode and to the circuit's DC power supply so that neither of those is over-stressed. With a normal silicon diode this means a voltage no greater than 0.7 Volts plus the positive DC rail voltage. If Schottky diodes are used in the protection network, it means a voltage no greater than 0.4 Volts beyond the positive rail. Such mild over-voltage conditions are unlikely to cause damage to any of the amplifier's internal transistors.

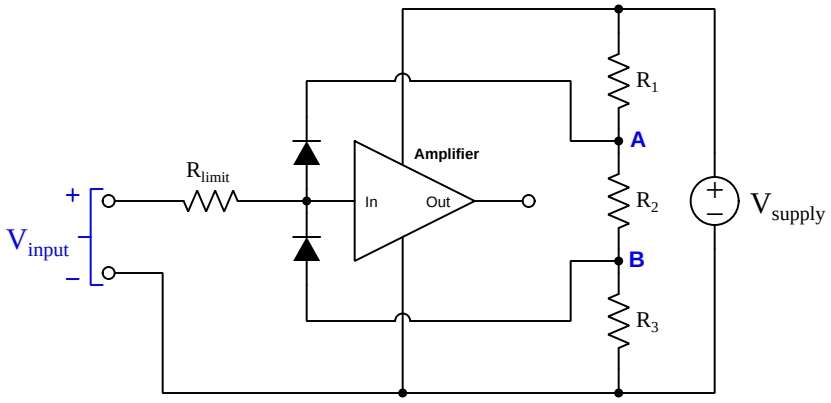
Similarly, if an excessive negative voltage is applied by an external source ("excessive" being any significant potential below that of the negative DC power supply rail):



Here the lower protection diode forward-biases and permits current from the external source to pass through the limiting resistor. While this occurs, the amplifier will never see more than that protection diode's forward voltage drop (V_F) between its input terminal and the negative DC rail bus, which again is unlikely to damage any transistor inside the amplifier. Again, the resistor should be sized to limit current to a value tolerable to the protection diode. However, this resistor should not be made so large that it might impair the normal operation of the amplifier when the protection diodes are not conducting: i.e. it should drop negligible voltage from the amplifier's normal input terminal current.

Such protection networks are often found inside of integrated circuits, sometimes without a series protection resistor in their most basic forms. Protection diodes are standard for CMOS digital logic gates to help protect their constituent MOSFET transistor gates from damage from ESD.

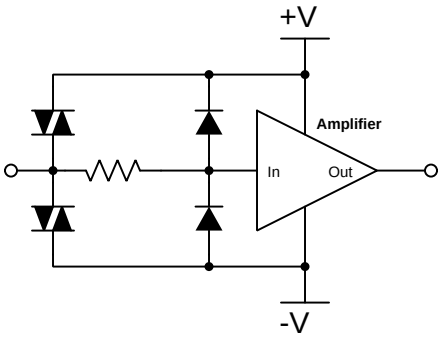
For some devices it is not sufficient for a diode network to clamp the input voltage to some value(s) slightly in excess of the device's power supply rail potential(s). For example, a device operating on a 5 Volt DC supply may tolerate input voltages no lower than zero and no greater than 5 Volts. A relatively simple solution to this problem is to provide the diode network with a set of "protection rails" just shy of the actual power supply rail potentials. Consider the following example as an illustration of this technique:



Here, resistors R_1, R_2, and R_3 form a voltage divider to produce two electrical potentials (labeled **A** and **B**), each one diode forward-voltage drop shy of the nearest rail potentials. For example, if the DC power supply output 5 Volts and the two protection diodes each dropped 0.7 Volts when forward-biased, the voltage divider would need to be designed such that point **A** was +4.3 Volts and point **B** was +0.7 Volts with respect to the DC supply's negative terminal. This would prompt the upper diode to turn on and clamp the amplifier's input potential to +5 Volts if ever the input potential exceeded +5 Volts, the limit resistor dropping the rest. Likewise, the lower diode would turn on if ever the input terminal's potential fell below the power supply's negative rail, clamping the amplifier's input terminal potential to exactly the same as that negative rail with R_{limit} dropping the rest.

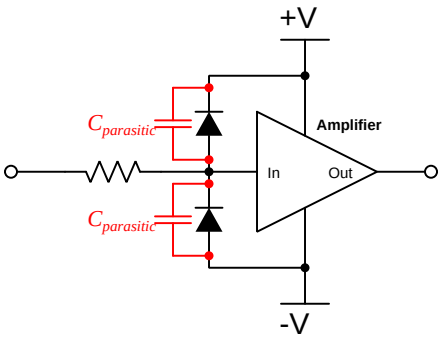
An important caveat to this strategy is that the voltage divider resistors R_1 and R_3 must be relatively small compared to the resistance of R_{limit} in order to ensure that those "protection rail" potentials **A** and **B** do not vary significantly when the protection diodes begin to conduct. A general engineering design principle here is to size R_{limit} at least ten times greater than either R_1 or R_3. If we size the resistors properly, this voltage-divider-based protection strategy may even be made *adjustable* by incorporating potentiometers into the voltage-divider network.

When protection must be provided against extraordinarily strong sources, additional protection may be added in their form of *DIACs* connected between each input terminal and power supply rail:



DIACs are classified as *thyristors*, because once triggered into an electrically conductive state by sufficiently high voltage (and/or sufficiently high *rate of change* of voltage, $\frac{dV}{dt}$) they will remain "on" so long as sufficient "holding" current passes through them even if voltage falls far below the initial triggering value. In summary, a thyristor acts as a very effective snubbing device to tame over-voltage conditions, essentially acting as a near-short to that offending source.

Protection networks, however, are not without their disadvantages. Chief among these is their tendency to adversely affect the signal being sensed by the amplifier being protected. Ideally a protection network should not corrupt the sensed signal at all, and only come into play if and when that signal becomes strong enough to pose threat of damage to the amplifier, but this would only be true if the protection diodes (and DIACs) had no parasitic properties when non-conducting. This is unfortunately untrue, as both diodes and DIACs exhibit *parasitic capacitance* which not only has the effect of storing electrical charge that the amplifier may interpret as a voltage that should not be present, but along with the protection resistor will form a *low-pass filter network* preventing the amplifier from being able to fully sense rapidly-changing input signals:



In other words, even with applied signals weak enough to pose absolutely no threat of harm to the amplifier, the capacitance inherent to the protection diodes will conspire with the series resistor

to "slow down" rates of rise and fall for any voltage arriving at the amplifier's input, thus making the amplifier "think" the signal isn't changing as rapidly as it really is.

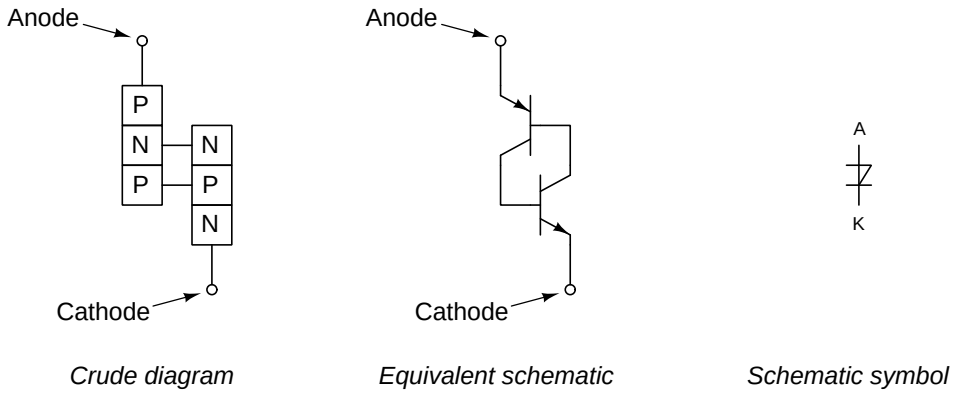
In order to minimize the effects of parasitic capacitance within the protection diodes, we must choose diodes with as little of that capacitance as possible and also select a protection resistor with as low a value as possible that still limits maximum current to a value safe for the amplifier.

5.9 Latch-up in integrated circuits

An often bewildering phenomenon for students of electronics is something called *latch-up*, known to plague certain forms of integrated circuit. "Latch-up" describes a condition where the IC suddenly ceases to function properly, and instead draws an unreasonably high amount of current from the DC power source, sometimes to the point of catastrophic failure. What is more, this sudden and unexpected high-current draw refuses to stop unless and until the DC power source is interrupted.

Close examination of the afflicted IC's internal schematic diagram is unhelpful, as no apparent cause for latch-up will ever be evident upon inspection. This is because latch-up is a phenomenon rooted in *parasitic* transistor structures within the IC itself resulting from the monolithic form of the circuit (i.e. being fabricated on a single wafer of semiconducting material). In fact, if the IC were constructed from a large collection of discrete components on a printed circuit board, latch-up could never occur!

In order to understand what latch-up is and why it occurs, we must first review a class of semiconductor devices called *thyristors*. A "thyristor" by definition is any semiconductor device manifesting *hysteresis*: the tendency to remain on once triggered on and to remain off once turned off. Thyristors include devices such as DIACs, TRIACs, silicon-controlled rectifiers (SCRs), Shockley diodes, and unijunction transistors. Among these various thyristor devices, the most elementary is the Shockley diode, illustrated below:



Crude diagram *Equivalent schematic* *Schematic symbol*

If we examine the equivalent Shockley diode schematic consisting of two interconnected bipolar junction transistors, one NPN and the other PNP, we find that it should not conduct current between the anode and cathode terminals. The only way for one of these transistors to ever turn on is to first initiate base current through it, but the interconnection with the other transistor makes base current for one contingent on collector current of the other, which cannot occur unless and until that other transistor's base is also activated by current. In other words, we have a "Catch-22" scenario where the NPN transistor cannot turn on until the PNP transistor turns on, and vice-versa. Therefore, Shockley diodes – and by extension all thyristors – naturally assume an "off" state and tend to maintain that state even as external voltage is applied across the two end-terminals.

If, however, we imagine *something* turning on one of these transistors, the resulting current would immediately turn on the other transistor, which in turn would hold the first transistor in its "on" state after the initiating event passed. Therefore, Shockley diodes – and by extension all

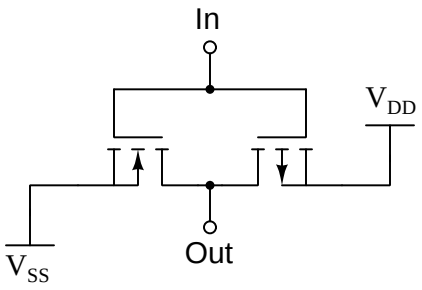
thyristors – tend to remain in their "on" states once triggered. Turning off a thyristor necessitates some external interruption of current. That "something" initially triggering the thyristor could be excessive applied voltage resulting in BJT breakdown, an extremely rapid rate-of-voltage change applied across the thyristor resulting in displacement current "through" one of the transistor's otherwise non-conducting depletion regions, excessive device temperature, a photon of light striking the depletion region and liberating an electron-hole pair, etc. Thyristors designed as controllable devices (e.g. SCRs, TRIACs, unijunction transistors) provide a third terminal to apply a triggering current.

Monolithic integrated circuit construction – where the entire circuit is built on a single wafer of semiconducting material – often results in parasitic PN junctions, and sometimes even parasitic NPN and PNP junction sets. If parasitic NPN and PNP junctions happen to coexist on the silicon die to form the equivalent of a Shockley diode thyristor structure, we will find that circuit susceptible to the unfortunate phenomenon of latch-up. These parasitic thyristors exist within many different types of integrated circuits, from logic gates to amplifiers to comparators to timers such as the popular model 555. They are guaranteed to exist within integrated circuits using complementary pairs of N-channel and P-channel MOSFETs as final "push-pull" output driver stages.

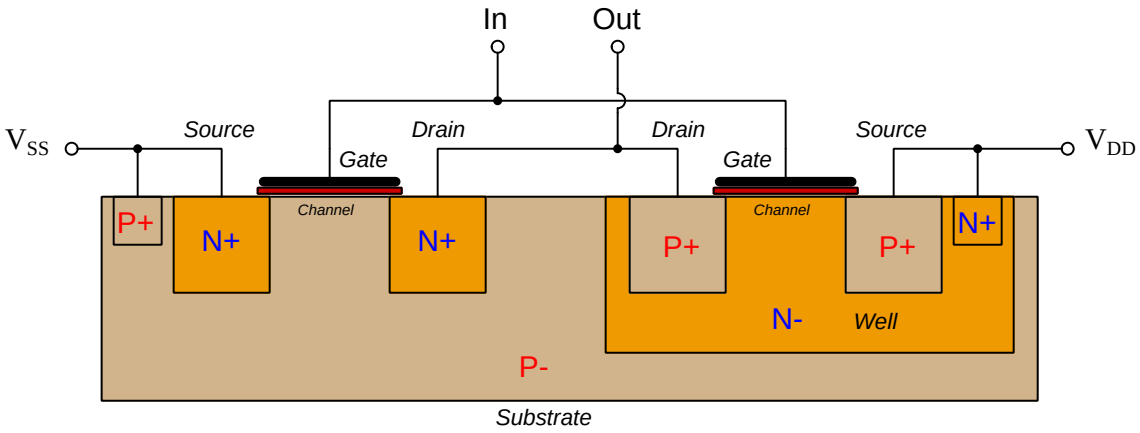
Sometimes the potential for latch-up is communicated in the form of a warning within IC datasheets, but other times not. Inconsistent and incomplete documentation, added to the fact that an examination of an IC's internal schematic does not necessarily reveal the presence of parasitic thyristors, makes latch-up risk a difficult thing to predict.

Let's begin our exploration of these parasitic structures by closely examining one of the most basic building-blocks of CMOS (Complementary Metal-Oxide Semiconductor) circuitry, the so-called *inverter*. A CMOS inverter network is nothing more than a P-channel enhancement-type MOSFET connected to an N-channel enhancement-type MOSFET, as shown below both in schematic form as well as semiconductor wafer cross-section. This complementary MOSFET pair acts as a "push-pull" transistor amplifier, capable of either sourcing current to a load or sinking current from one:

Schematic of an integrated CMOS inverter

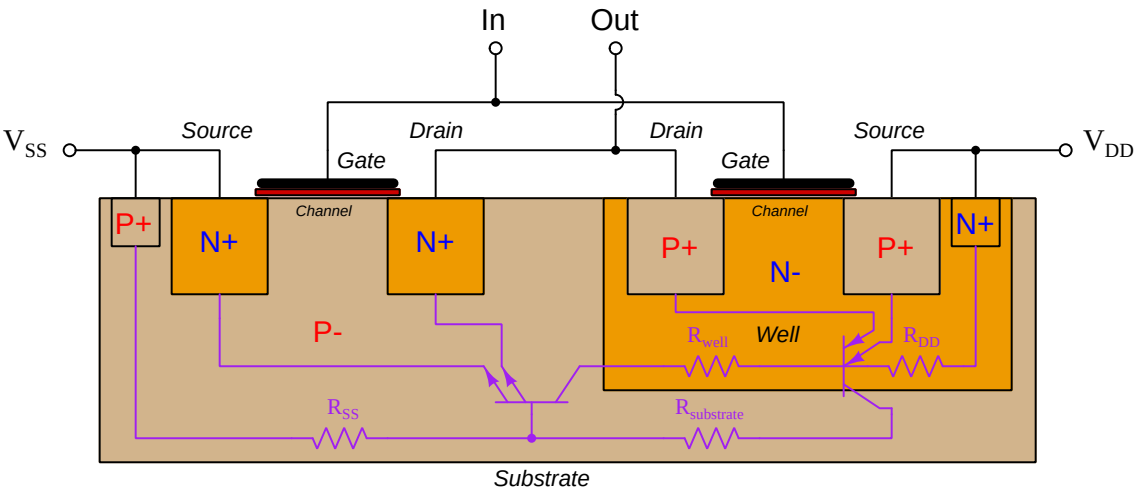


Cross-section of an integrated CMOS inverter

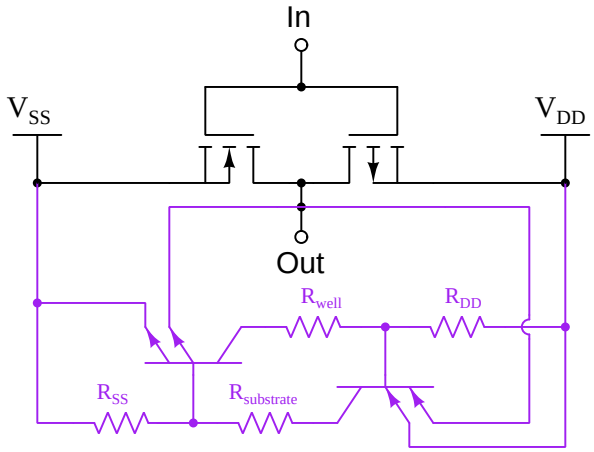


Each of the doped semiconductor layers in the cross-section are noted with their respective dopant concentration levels, a "+" meaning a high concentration of dopant and "−" meaning a low concentration of dopant. Strongly-doped P-type and N-type layers are shown as P+ and N+, respectively. Lightly-doped P-type and N-type layers are shown as P− and N−, respectively. Note how the source and drain regions are all heavily-doped in order to provide a rich reservoir of majority charge carriers to fill in each of the channels when prompted by a sufficient gate-to-substrate electric field.

Now we will include within this cross-sectional illustration some the parasitic transistors[11] and resistances formed by the P and N type doped regions:



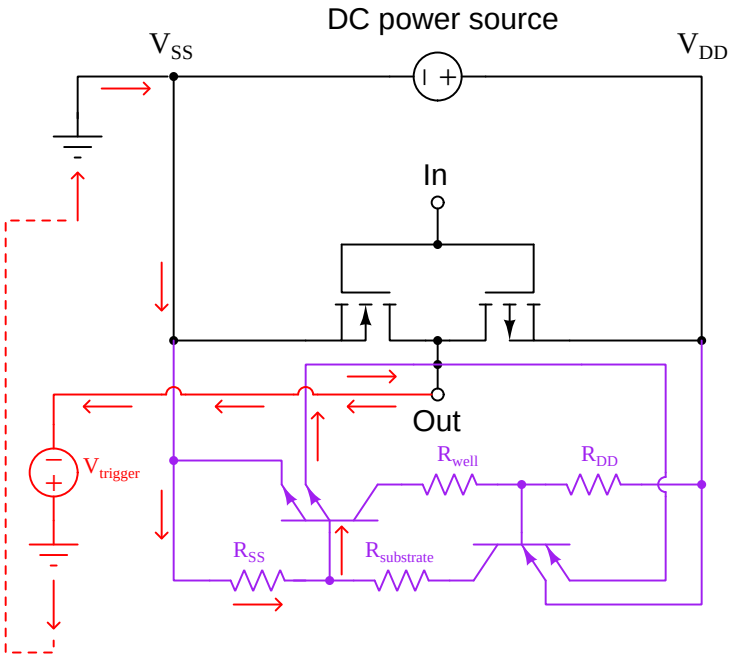
Including these parasitic components into the CMOS inverter schematic diagram:



Now it should be clear that a Shockley diode structure exists within the CMOS inverter's integrated circuit. Furthermore, with each bipolar transistor having two emitters, one each connected to the inverter's output terminal, that output terminal becomes a pathway to externally trigger this thyristor structure and thereby throw the IC into a state of latch-up.

[11] It is helpful to recall that the emitter region of any bipolar junction transistor is more heavily doped than the collector, in order to provide an abundance of charge carriers to inject into the base region as minority carriers. This is why the emitters of the PNP are P+ doped and the collector P− doped, and why the emitters of the NPN are N+ doped and the collector N− doped. Multiple emitters in a BJT simply mean more than one location from which charge carriers may be injected into the base and thereby render the transistor conductive between emitter and collector.

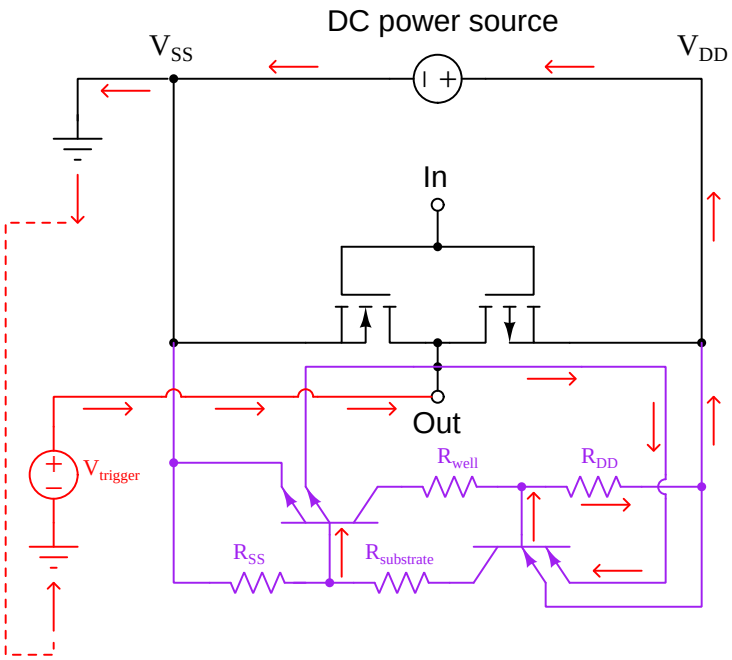
Triggering a latch-up condition is as simple as something external to the inverter IC driving its output terminal below V_{SS} potential or above V_{DD} potential. In the next example we will see a below-V_{SS} potential initiating latch-up:



In this scenario, external source $V_{trigger}$ forced the inverter's output terminal to a potential lower than ground (V_{SS}), and this in turn prompts a current through one of the emitters of the parasitic NPN transistor, injecting electrons into its P-type base region and turning it on. Once the NPN transistor turns on, its collector current permits base current through the PNP transistor, turning it on as well, and thus the two parasitic transistors latch each other into conductive states.

Once latched on, the parasitic thyristor creates a V_{DD}-to-V_{SS} short-circuit. Excessive current now passes through the DC power source and through the parasitic resistances of R_{well} and $R_{substrate}$, possibly damaging the entire integrated circuit by excessive heating. Only an interruption of the DC power source will halt the latch-up condition.

Next, we will see an above-V_{DD} potential initiating latch-up:



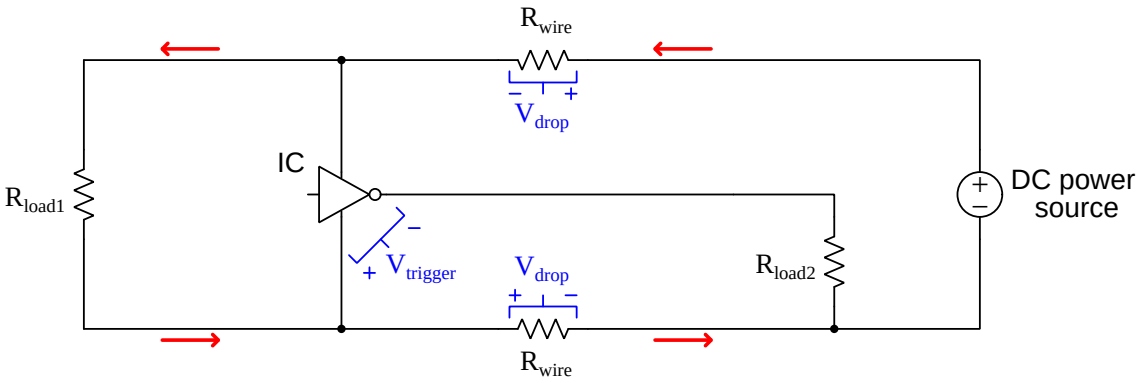
In this scenario, external source $V_{trigger}$ forced the inverter's output terminal to a potential higher than that of the DC source voltage (V_{DD}), and this in turn prompts a current through one of the emitters of the parasitic PNP transistor, injecting holes into its N-type base region and turning it on. Once the PNP transistor turns on, its collector current permits base current through the NPN transistor, turning it on as well, and thus the thyristor formed by the two parasitic transistors becomes latched in its "on" state.

Again, once the parasitic thyristor is latched in its "on" state it will create a V_{DD}-to-V_{SS} short-circuit. Excessive current will now pass through the DC power source and through the parasitic resistances of R_{well} and $R_{substrate}$, possibly damaging the entire integrated circuit by excessive heating. Only an interruption of the DC power source will abort the latch-up condition.

Clearly, one of the causes of latch-up is some external electrical source attempting to force the output terminal either above V_{DD} potential or below V_{SS} potential, and as such we should strive to avoid such possibilities in all our circuit designs.

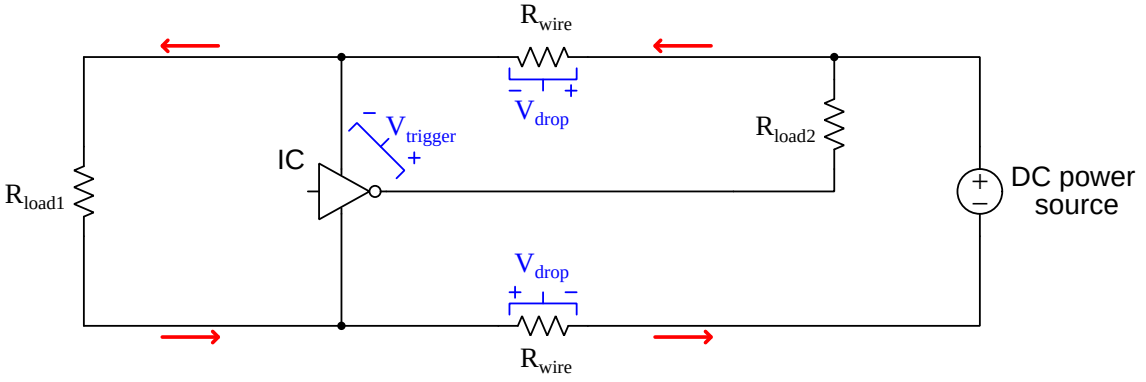
Another potential cause of latch-up is if the load driven by an IC output is too inductive. Inductance stores energy in magnetic fields, and the amount of energy is a function of the current through the inductance. If an inductive load is energized by the output of an IC and then the IC switches state to de-energize that load, the load inductance will become a temporary current source and try to force current through the IC's output, often generating surprisingly high voltages in the process. Such voltage transients may easily trigger latch-up in the IC.

Another possible cause of latch-up is power supply rail instability. That is, if something is causing an IC's positive and/or negative rail potentials to vary, and that same IC connects to some other device with higher positive or lower negative power supply rails, this may constitute a triggering mechanism for latch-up. Consider the following example:



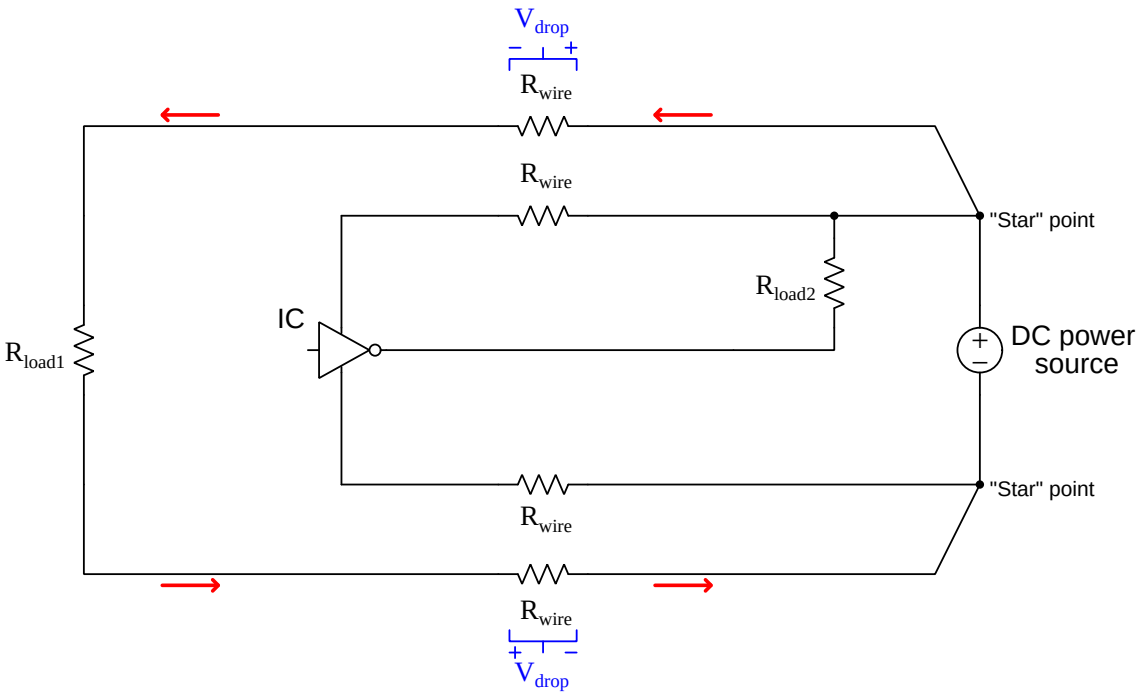
Here we see some voltage dropped along the conductors carrying current to a heavy load (R_{load1}), the effect of these voltage drops being to suppress the potential of the IC's positive rail terminal slightly below that of the power source's positive terminal, and to elevate the potential of IC's negative rail terminal slightly above that of the power source's negative terminal. In this case, with the IC's output connected to a second load (R_{load2}), we see that the voltage drop along the negative power conductor has the effect of making the IC's output terminal fall to a potential slightly less than its negative power terminal. As we have seen previously, this may be enough to trigger latch-up in that IC.

If R_{load2} happened to be connected to the positive conductor rather than the negative, we would similarly risk latch-up due to the load imposing a potential on the IC's output greater than that of its positive power terminal:



In either scenario the wiring voltage drop caused by the high current demand of R_{load1} risks latch-up in the IC by skewing its power-supply terminal potentials compared to the load it connects to. A

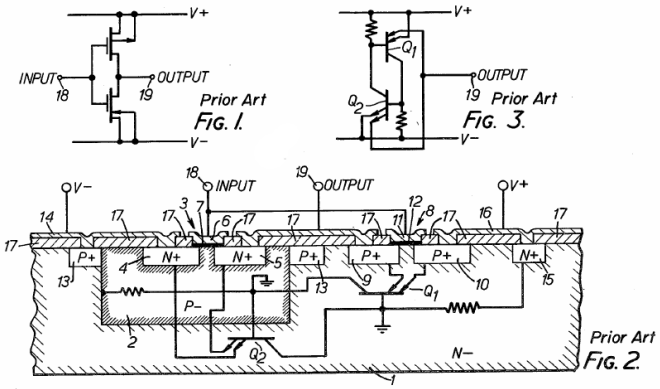
strategy for avoiding this problem is to provide high-current loads with their own dedicated power conductors, and not connect them to the same power conductors that sensitive ICs use for power delivery. This strategy is often referred to as *star-point grounding* or star-point power distribution:

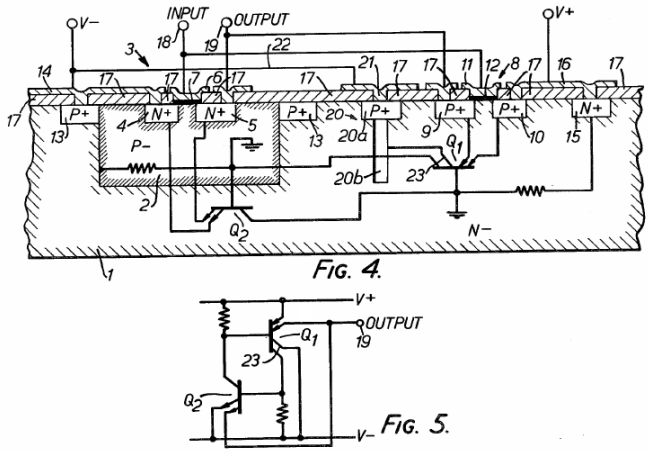


Star-point power connections do not eliminate voltage drop due to high load current, but they do isolate those voltage drops from affecting *other* loads such as ICs sensitive to latch-up.

Given the fact that Complementary MOSFET transistor pairs are so prevalent in modern integrated circuits (e.g. *all* CMOS digital logic) we must find mitigations for latch-up or else risk self-destruction of a great many ICs any time an output pin is subjected to even the briefest transient voltage outside of the DC power supply's rail potentials.

One such mitigation was invented by Alastair Stevenson of the Hughes Aircraft Company in 1978, consisting of an extra layer of implanted material in the semiconductor structure serving as a new collector of charge carriers. His patent ("CMOS Monolithic Integrated Circuit", US patent 4,173,767 filed 5 May 1978 and granted 6 November 1979) contrasts the "prior art" of CMOS inverter construction[12] (figures 1 through 3) against his innovation shown in figures 4 and 5:





In Stevenson's design we see the addition of a heavily-doped P+ region (20a in figure 4) near the surface of the semiconductor die connecting with a deeper region of lightly-doped P− material (20b in figure 4) inset into the N− substrate. This P− region serves to collect holes injected as minority

[12] Stevenson's prior art example happens to use an N-type substrate rather than a P-type substrate as I showed earlier, thus necessitating a P-type well surrounding the N-channel MOSFET rather than the N-type well surrounding the P-channel MOSFET in my illustrations.

carriers into the N− substrate that might otherwise travel to the P− well region of the N-channel MOSFET, essentially adding a second collector to the PNP parasitic transistor. The purpose for the P+ region interfacing the metal surface conductor to the deeper P− inset is to avoid "rectifying" contact between metal and semiconductor surfaces, the heavy doping of that P+ region (20a) serving instead to form an "ohmic" contact with the metal. This additional collector region serves to divert current away from the parasitic NPN transistor that would otherwise turn it on and ensure latch-up once the PNP transistor was triggered by an offending potential at the output terminal.

Stevenson describes how such current-collecting channels inserted into the substrate need only be placed between complementary MOSFETs at device input or output terminals, not between every pair of complementary MOSFETs, due to the fact that only those connecting to device pins are susceptible to the triggering that cause latch-up:

> It will be appreciated that the circuit arrangement shown in FIG. 4 represents only one small portion of the integrated circuit which will in fact include many more complementary MOS transistors. However, channels such as 20 need not necessarily be included between all adjacent complementary transistors in the circuit, and preferably channels such as 20 are only included between complementary transistors that constitute input or output stages of the circuit since it is only these transistors which will be subject to spurious voltage fluctuations which could trigger a bipolar conductive latch up in the circuit.

A final note on the topic of latch-up is that this term is unfortunately mis-used to describe a phenomenon plaguing certain comparators and operational amplifiers. In this other meaning of the term "latch-up", the comparator or opamp output terminal assumes an incorrect potential because one or more of the input terminal potentials has approached or exceeded a power supply rail potential. This kind of "latch-up" is sometimes referred to as *pseudo latch-up* because unlike the thyristor-based latch-up we've been exploring here, pseudo latch-up immediately ceases as soon as the offending input potential returns to normal. In the "real" form of latch-up, only interruption of DC power will cause the latching action to cease.

Chapter 6

Questions

This learning module, along with all others in the ModEL collection, is designed to be used in an inverted instructional environment where students independently read[1] the tutorials and attempt to answer questions on their own *prior* to the instructor's interaction with them. In place of lecture[2], the instructor engages with students in Socratic-style dialogue, probing and challenging their understanding of the subject matter through inquiry.

Answers are not provided for questions within this chapter, and this is by design. Solved problems may be found in the Tutorial and Derivation chapters, instead. The goal here is *independence*, and this requires students to be challenged in ways where others cannot think for them. Remember that you always have the tools of *experimentation* and *computer simulation* (e.g. SPICE) to explore concepts!

The following lists contain ideas for Socratic-style questions and challenges. Upon inspection, one will notice a strong theme of *metacognition* within these statements: they are designed to foster a regular habit of examining one's own thoughts as a means toward clearer thinking. As such these sample questions are useful both for instructor-led discussions as well as for self-study.

[1] Technical reading is an essential academic skill for any technical practitioner to possess for the simple reason that the most comprehensive, accurate, and useful information to be found for developing technical competence is in textual form. Technical careers in general are characterized by the need for continuous learning to remain current with standards and technology, and therefore any technical practitioner who cannot read well is handicapped in their professional development. An excellent resource for educators on improving students' reading prowess through intentional effort and strategy is the book textitReading For Understanding – How Reading Apprenticeship Improves Disciplinary Learning in Secondary and College Classrooms by Ruth Schoenbach, Cynthia Greenleaf, and Lynn Murphy.

[2] Lecture is popular as a teaching method because it is easy to implement: any reasonably articulate subject matter expert can talk to students, even with little preparation. However, it is also quite problematic. A good lecture always makes complicated concepts seem easier than they are, which is bad for students because it instills a false sense of confidence in their own understanding; reading and re-articulation requires more cognitive effort and serves to verify comprehension. A culture of teaching-by-lecture fosters a debilitating dependence upon direct personal instruction, whereas the challenges of modern life demand independent and critical thought made possible only by gathering information and perspectives from afar. Information presented in a lecture is ephemeral, easily lost to failures of memory and dictation; text is forever, and may be referenced at any time.

General challenges following tutorial reading

- Summarize as much of the text as you can in one paragraph of your own words. A helpful strategy is to explain ideas as you would for an intelligent child: as simple as you can without compromising too much accuracy.
- Simplify a particular section of the text, for example a paragraph or even a single sentence, so as to capture the same fundamental idea in fewer words.
- Where did the text make the most sense to you? What was it about the text's presentation that made it clear?
- Identify where it might be easy for someone to misunderstand the text, and explain why you think it could be confusing.
- Identify any new concept(s) presented in the text, and explain in your own words.
- Identify any familiar concept(s) such as physical laws or principles applied or referenced in the text.
- Devise a proof of concept experiment demonstrating an important principle, physical law, or technical innovation represented in the text.
- Devise an experiment to disprove a plausible misconception.
- Did the text reveal any misconceptions you might have harbored? If so, describe the misconception(s) and the reason(s) why you now know them to be incorrect.
- Describe any useful problem-solving strategies applied in the text.
- Devise a question of your own to challenge a reader's comprehension of the text.

General follow-up challenges for assigned problems

- Identify where any fundamental laws or principles apply to the solution of this problem, especially before applying any mathematical techniques.
- Devise a thought experiment to explore the characteristics of the problem scenario, applying known laws and principles to mentally model its behavior.
- Describe in detail your own strategy for solving this problem. How did you identify and organized the given information? Did you sketch any diagrams to help frame the problem?
- Is there more than one way to solve this problem? Which method seems best to you?
- Show the work you did in solving this problem, even if the solution is incomplete or incorrect.
- What would you say was the most challenging part of this problem, and why was it so?
- Was any important information missing from the problem which you had to research or recall?
- Was there any extraneous information presented within this problem? If so, what was it and why did it not matter?
- Examine someone else's solution to identify where they applied fundamental laws or principles.
- Simplify the problem from its given form and show how to solve this simpler version of it. Examples include eliminating certain variables or conditions, altering values to simpler (usually whole) numbers, applying a limiting case (i.e. altering a variable to some extreme or ultimate value).
- For quantitative problems, identify the real-world meaning of all intermediate calculations: their units of measurement, where they fit into the scenario at hand. Annotate any diagrams or illustrations with these calculated values.
- For quantitative problems, try approaching it qualitatively instead, thinking in terms of "increase" and "decrease" rather than definite values.
- For qualitative problems, try approaching it quantitatively instead, proposing simple numerical values for the variables.
- Were there any assumptions you made while solving this problem? Would your solution change if one of those assumptions were altered?
- Identify where it would be easy for someone to go astray in attempting to solve this problem.
- Formulate your own problem based on what you learned solving this one.

General follow-up challenges for experiments or projects

- In what way(s) was this experiment or project easy to complete?
- Identify some of the challenges you faced in completing this experiment or project.

- Show how thorough documentation assisted in the completion of this experiment or project.
- Which fundamental laws or principles are key to this system's function?
- Identify any way(s) in which one might obtain false or otherwise misleading measurements from test equipment in this system.
- What will happen if (component X) fails (open/shorted/etc.)?
- What would have to occur to make this system unsafe?

6.1 Conceptual reasoning

These questions are designed to stimulate your analytic and synthetic thinking[3]. In a Socratic discussion with your instructor, the goal is for these questions to prompt an extended dialogue where assumptions are revealed, conclusions are tested, and understanding is sharpened. Your instructor may also pose additional questions based on those assigned, in order to further probe and refine your conceptual understanding.

Questions that follow are presented to challenge and probe your understanding of various concepts presented in the tutorial. These questions are intended to serve as a guide for the Socratic dialogue between yourself and the instructor. Your instructor's task is to ensure you have a sound grasp of these concepts, and the questions contained in this document are merely a means to this end. Your instructor may, at his or her discretion, alter or substitute questions for the benefit of tailoring the discussion to each student's needs. The only absolute requirement is that each student is challenged and assessed at a level equal to or greater than that represented by the documented questions.

It is far more important that you convey your *reasoning* than it is to simply convey a correct answer. For this reason, you should refrain from researching other information sources to answer questions. What matters here is that *you* are doing the thinking. If the answer is incorrect, your instructor will work with you to correct it through proper reasoning. A correct answer without an adequate explanation of how you derived that answer is unacceptable, as it does not aid the learning or assessment process.

You will note a conspicuous lack of answers given for these conceptual questions. Unlike standard textbooks where answers to every other question are given somewhere toward the back of the book, here in these learning modules students must rely on other means to check their work. The best way by far is to debate the answers with fellow students and also with the instructor during the Socratic dialogue sessions intended to be used with these learning modules. Reasoning through challenging questions with other people is an excellent tool for developing strong reasoning skills.

Another means of checking your conceptual answers, where applicable, is to use circuit simulation software to explore the effects of changes made to circuits. For example, if one of these conceptual questions challenges you to predict the effects of altering some component parameter in a circuit, you may check the validity of your work by simulating that same parameter change within software and seeing if the results agree.

[3] *Analytical* thinking involves the "disassembly" of an idea into its constituent parts, analogous to dissection. *Synthetic* thinking involves the "assembly" of a new idea comprised of multiple concepts, analogous to construction. Both activities are high-level cognitive skills, extremely important for effective problem-solving, necessitating frequent challenge and regular practice to fully develop.

6.1.1 Reading outline and reflections

"Reading maketh a full man; conference a ready man; and writing an exact man" – Francis Bacon

Francis Bacon's advice is a blueprint for effective education: reading provides the learner with knowledge, writing focuses the learner's thoughts, and critical dialogue equips the learner to confidently communicate and apply their learning. Independent acquisition and application of knowledge is a powerful skill, well worth the effort to cultivate. To this end, students should read these educational resources closely, journal their own reflections on the reading, and discuss in detail their findings with classmates and instructor(s). You should be able to do all of the following after reading any instructional text:

☑ Briefly SUMMARIZE THE TEXT in the form of a journal entry documenting your learning as you progress through the course of study. Share this summary in dialogue with your classmates and instructor. Journaling is an excellent self-test of thorough reading because you cannot clearly express what you have not read or did not comprehend.

☑ Demonstrate ACTIVE READING STRATEGIES, including verbalizing your impressions as you read, simplifying long passages to convey the same ideas using fewer words, annotating text and illustrations with your own interpretations, working through mathematical examples shown in the text, cross-referencing passages with relevant illustrations and/or other passages, identifying problem-solving strategies applied by the author, etc. Technical reading is a special case of problem-solving, and so these strategies work precisely because they help solve any problem: paying attention to your own thoughts (metacognition), eliminating unnecessary complexities, identifying what makes sense, paying close attention to details, drawing connections between separated facts, and noting the successful strategies of others.

☑ Identify IMPORTANT THEMES, especially GENERAL LAWS and PRINCIPLES, expounded in the text and express them in the simplest of terms as though you were teaching an intelligent child. This emphasizes connections between related topics and develops your ability to communicate complex ideas to anyone.

☑ Form YOUR OWN QUESTIONS based on the reading, and then pose them to your instructor and classmates for their consideration. Anticipate both correct and incorrect answers, the incorrect answer(s) assuming one or more plausible misconceptions. This helps you view the subject from different perspectives to grasp it more fully.

☑ Devise EXPERIMENTS to test claims presented in the reading, or to disprove misconceptions. Predict possible outcomes of these experiments, and evaluate their meanings: what result(s) would confirm, and what would constitute disproof? Running mental simulations and evaluating results is essential to scientific and diagnostic reasoning.

☑ Specifically identify any points you found CONFUSING. The reason for doing this is to help diagnose misconceptions and overcome barriers to learning.

6.1.2 Foundational concepts

Correct analysis and diagnosis of electric circuits begins with a proper understanding of some basic concepts. The following is a list of some important concepts referenced in this module's full tutorial. Define each of them in your own words, and be prepared to illustrate each of these concepts with a description of a practical example and/or a live demonstration.

Thought experiments as a problem-solving strategy

Amplification

Gain

Single-ended voltage signal

Differential voltage signal

Rails

Inverting input

Noninverting input

Integrated circuit

Bridge network

Open loop

Closed loop

Negative feedback

Cardinal Rule of Negative Feedback

Series regulator

Shunt regulator

Servo

Functional inversion

Ohm's Law

Kirchhoff's Voltage Law

Kirchhoff's Current Law

Electrical source

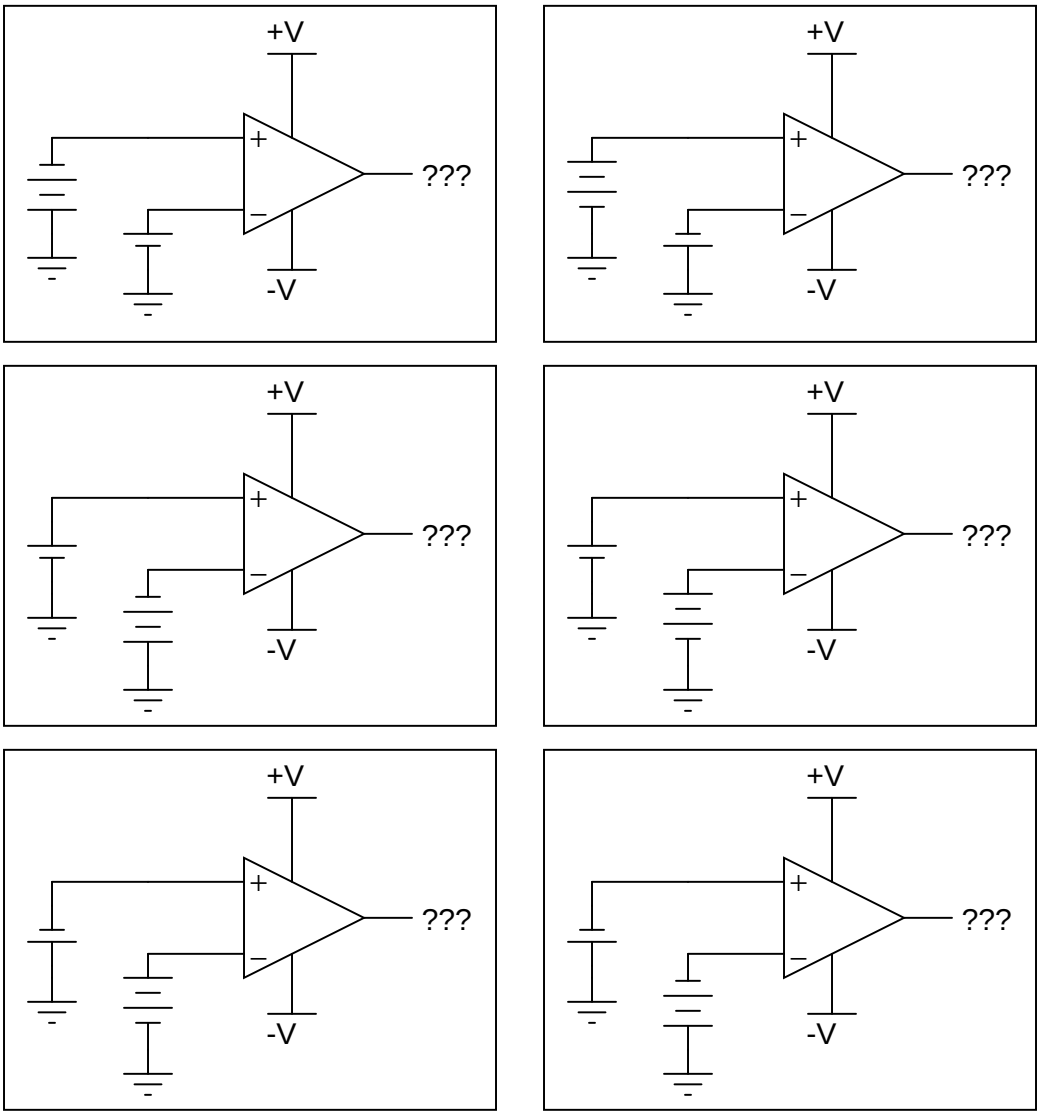
Electrical load

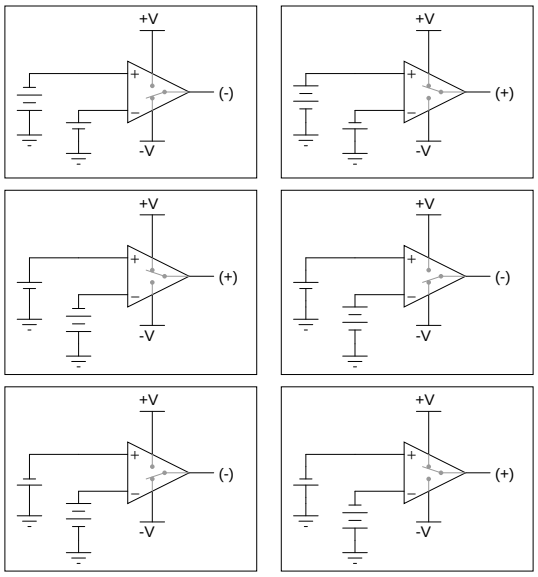
Decibels

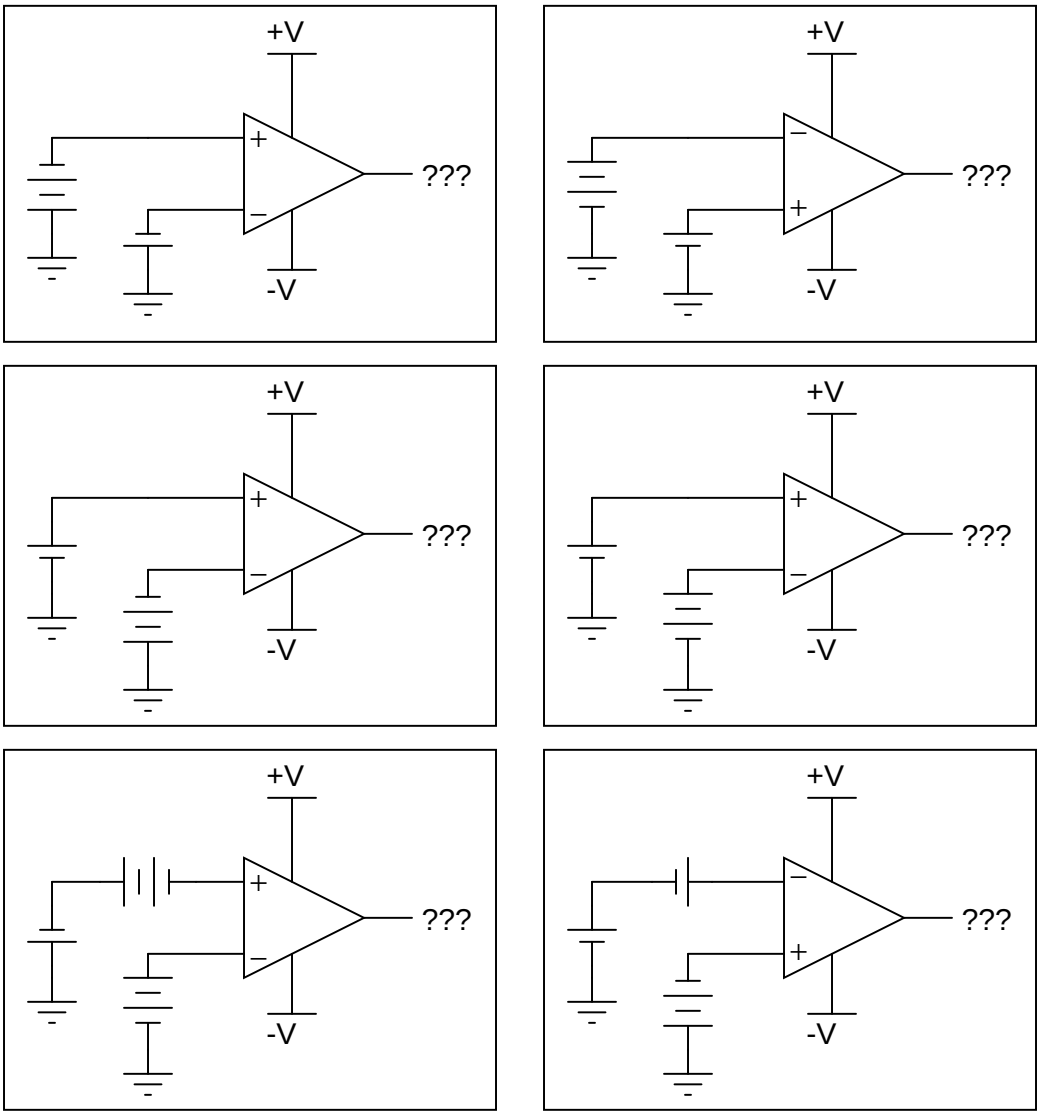
Sinusoidal decomposition (i.e. Fourier's Theorem)

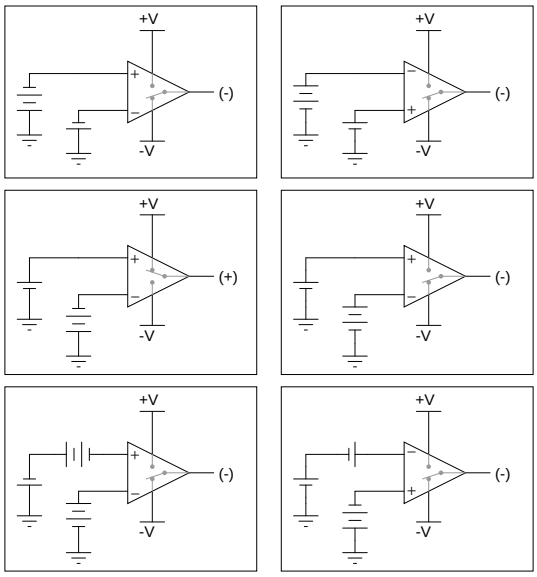
6.1.3 Determining output polarities

Determine the output voltage polarity of each opamp (with reference to ground), given the following input conditions:




+V
(-)
-V
+V
(+)
-V
+V
(+)
-V
+V
(-)
-V
+V
(-)
-V
+V
(+)
-V


+V
+
−
???
-V
+V
−
+
???
-V
+V
+
−
???
-V
+V
+
−
???
-V
+V
+
−
???
-V
+V
−
+
???
-V

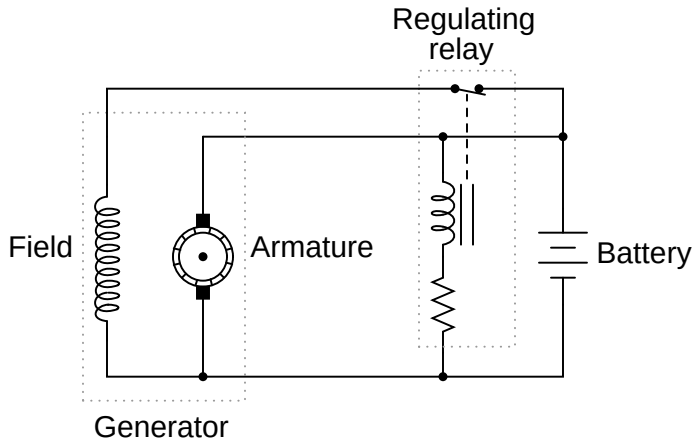


Challenges

- Are these opamps being operated in *open-loop* mode, or *closed-loop* mode.

6.1.4 Fooling a voltage regulator

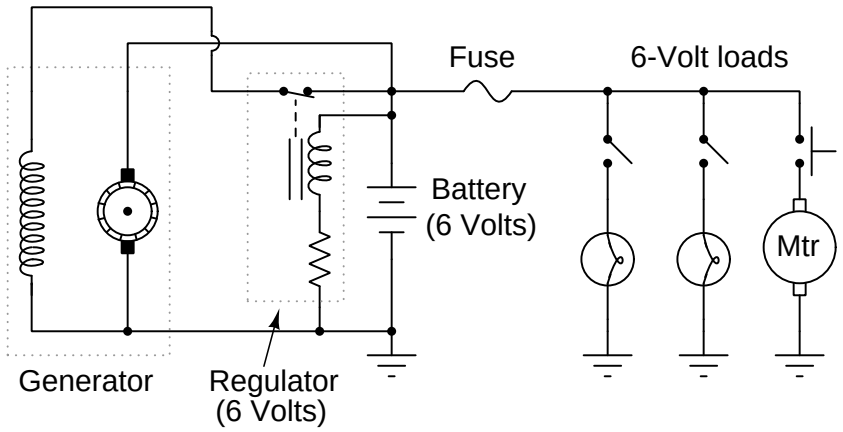
Generators used in battery-charging systems must be *regulated* so as to not over-charge the battery(ies) they are connected to. Here is a crude, relay-based voltage regulator for a DC generator:



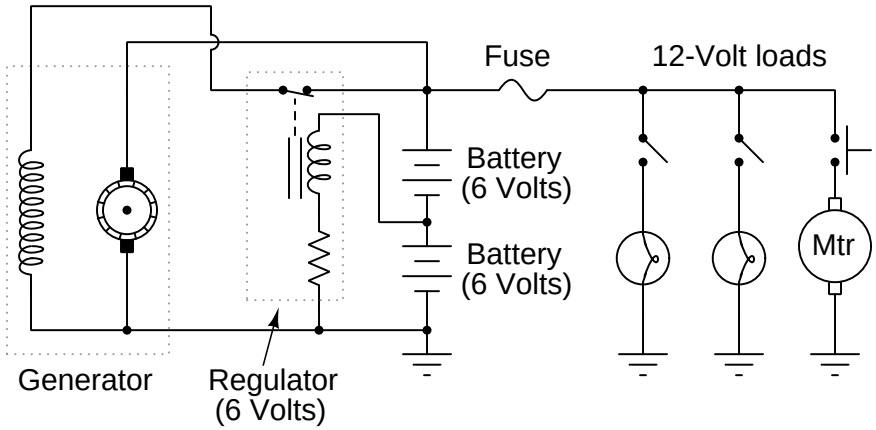
Simple electromechanical relay circuits such as this one were very common in automotive electrical systems during the 1950's, 1960's, and 1970's. The fundamental principle upon which their operation is based is called *negative feedback*: where a system takes action to *oppose* any change in a certain variable. In this case, the variable is generator output voltage. If the battery voltage becomes excessive, the relay opens and de-energizes the field winding. When the voltages sags back down to an acceptable level, the relay re-closes and re-energizes the field winding so that the generator can begin generating voltage again.

Identify what you could modify in this relay-based voltage regulator circuit to slightly increase the voltage output by the generator.

Suppose that a mechanic has an idea for upgrading the electrical system in an automobile originally designed for 6 Volt operation. He wants to upgrade the 6 Volt headlights, starter motor, battery, etc, to 12 Volts, but wishes to retain the original 6-Volt generator and regulator and somehow find a way to force the 6-Volt generator to output 12 Volts. Shown here is the original 6-Volt electrical system:



The mechanic's plan is to replace all the 6-Volt loads with 12-Volt loads, and use two 6-Volt batteries connected in series, with the original (6-Volt) regulator sensing voltage across only one of those batteries:



Explain how this system is supposed to work. Do you think the mechanic's plan is practical, or are there any problems with it?

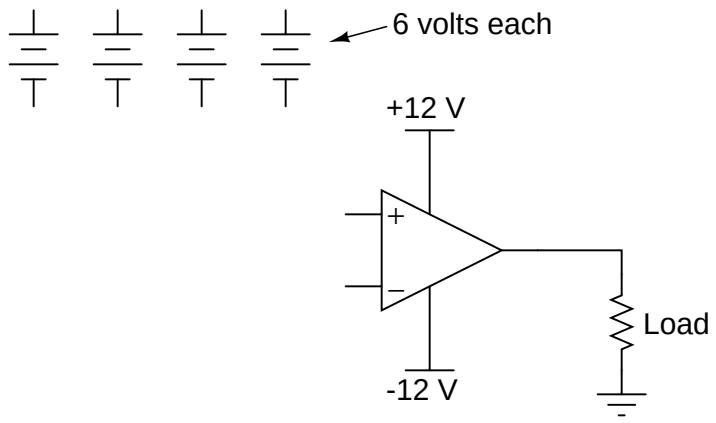
Challenges

- What would we have to change in the original circuit to alter the generator's voltage regulation set-point (the "target" voltage at which the generator's output is supposed to be regulated).

- Identify factors that may prevent the generator from outputting enough voltage with the regulator connected as shown in the last diagram.

6.1.5 Forming a split power supply

Many op-amp circuits require a *dual* or *split* power supply, consisting of three power terminals: +V, −V, and Ground. Draw the necessary connections between the 6 Volt batteries in this schematic diagram to provide +12 V, −12 V, and Ground to this op-amp:

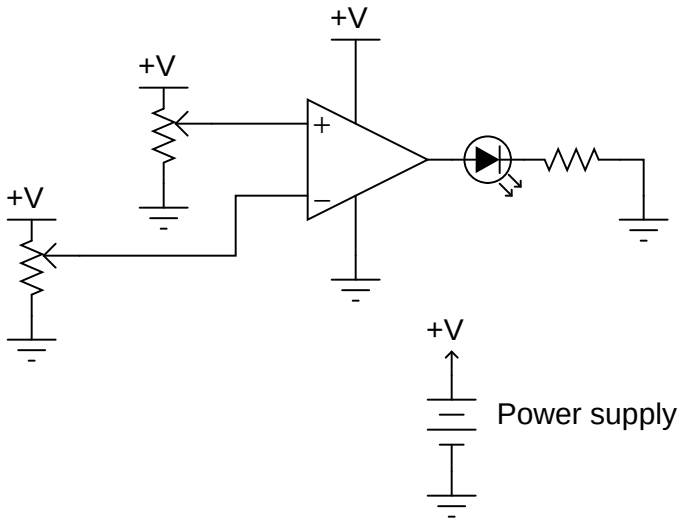


Challenges

- Describe how the opamp would be limited in its operation without a "split" power supply.

6.1.6 Load current path

In this circuit, an op-amp (functioning as a comparator) turns on an LED if the proper input voltage conditions are met:



Trace the complete path of current powering the LED. Where, exactly, does the LED get its power from?

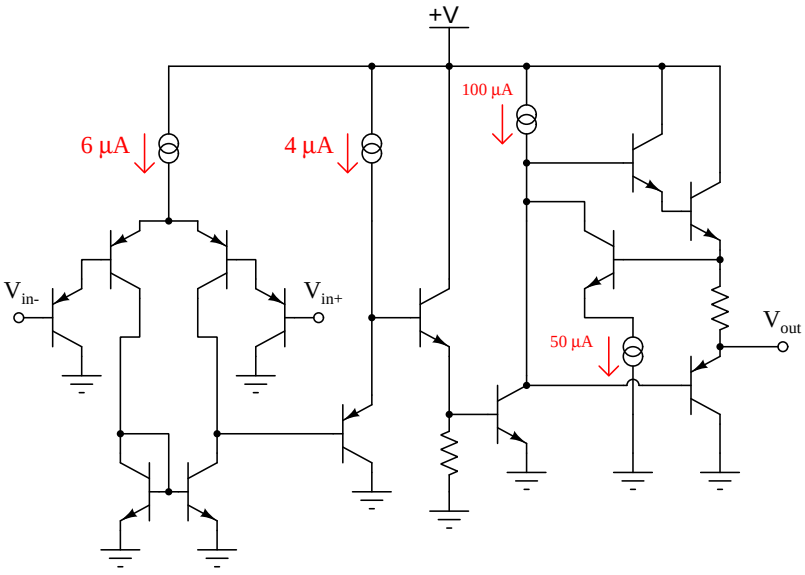
Now, suppose the LED were removed and then re-inserted into the circuit facing the other direction, with the other end of its resistor tied to the +V rail instead of to ground. Assuming proper input voltage conditions changed to energize the LED in its new orientation, trace the complete path of current powering the LED.

Challenges

- Identify wires in this circuit that do *not* bear any of this load current.
- Identify some wiper positions on the two potentiometers that would ensure an energized LED.

6.1.7 Model 324 opamp qualitative analysis

Shown here is a simplified schematic diagram of one of the operational amplifiers inside an 324 quad op-amp integrated circuit:



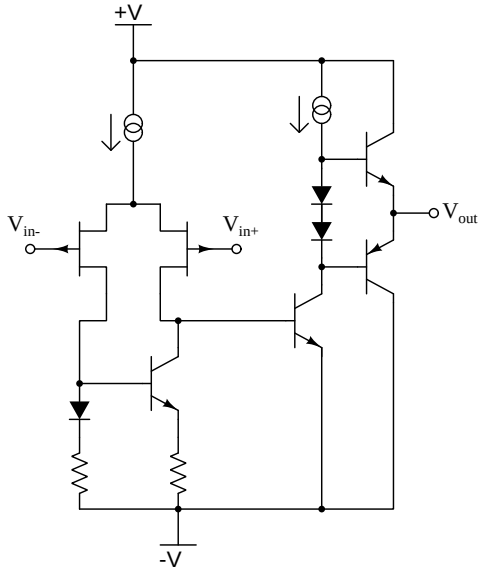
Qualitatively determine what will happen to the output voltage (V_{out}) if the voltage on the inverting input (V_{in-}) increases, and the voltage on the noninverting input (V_{in+}) remains the same (all voltages are positive quantities, referenced to ground). Explain what happens at every stage of the op-amp circuit (voltages increasing or decreasing, currents increasing or decreasing) with this change in input voltage.

Challenges

- Explain why the "limiting cases" problem-solving strategy would be useful in this analysis.
- Identify the function of the double-circle symbols in this simplified diagram.

6.1.8 TL08x opamp qualitative analysis

Shown here is a simplified schematic diagram of one of the operational amplifiers inside a TL08x (TL081, TL082, or TL084) op-amp integrated circuit:



Qualitatively determine what will happen to the output voltage (V_{out}) if the voltage on the noninverting input (V_{in+}) increases, and the voltage on the inverting input (V_{in-}) remains the same (all voltages are positive quantities, referenced to -V). Explain what happens at every stage of the op-amp circuit (voltages increasing or decreasing, currents increasing or decreasing) with this change in input voltage.

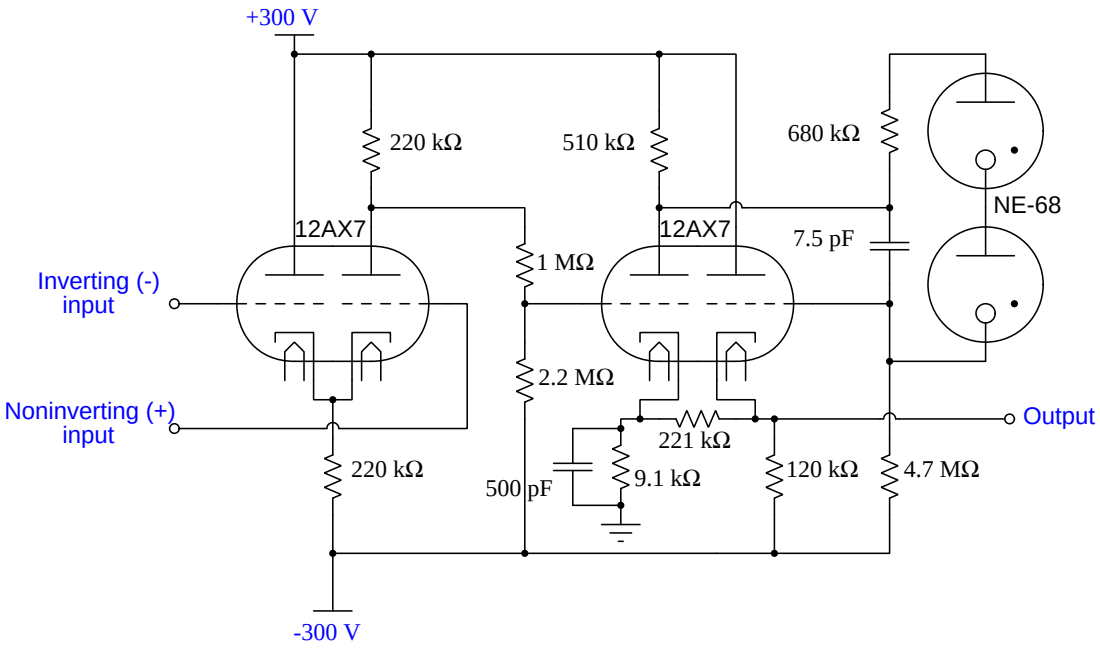
Challenges

- Explain why the "limiting cases" problem-solving strategy would be useful in this analysis.
- Identify the function of the double-circle symbols in this simplified diagram.

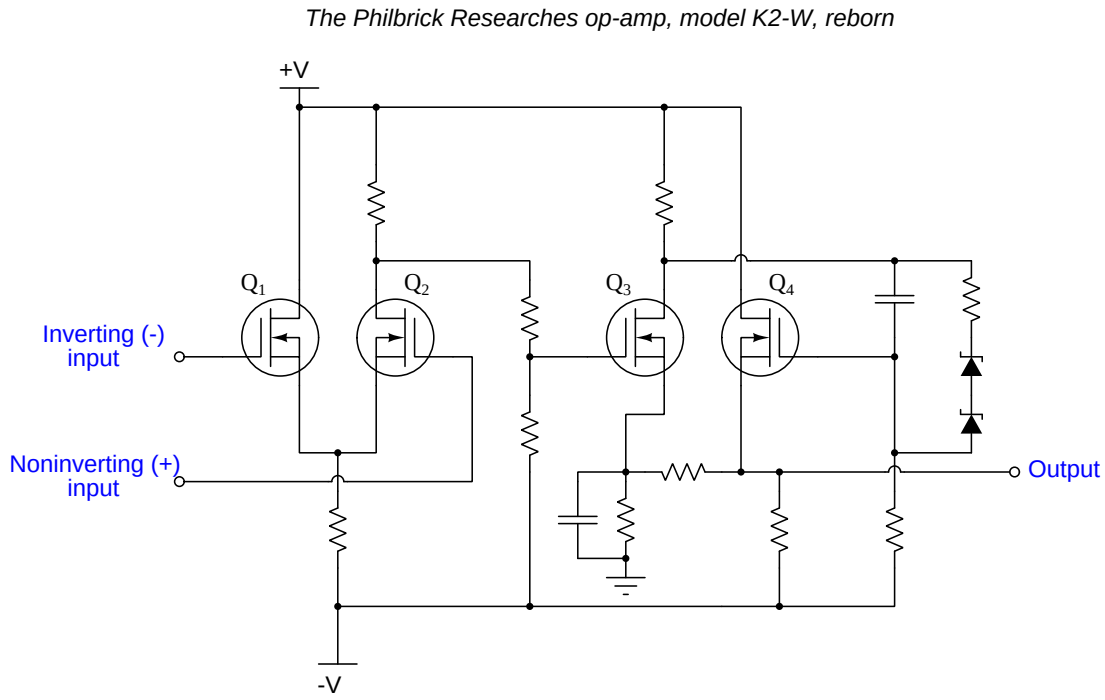
6.1.9 Philbrick K2-W opamp

One of the first popular operational amplifiers was manufactured by Philbrick Researches, and it was called the K2-W. Built with two dual-triode vacuum tubes, its original schematic diagram looked like this:

The Philbrick Researches op-amp, model K2-W



To make this opamp circuit easier for modern students to understand, I'll substitute equivalent solid-state components for all tubes in the original design:



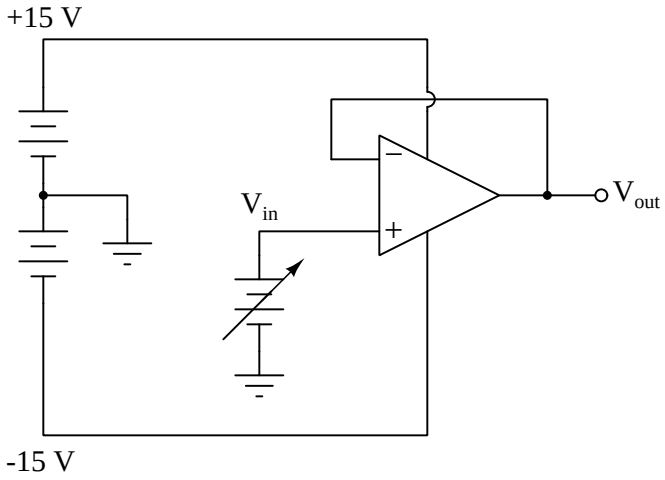
Explain the configuration (common-source, common-drain, or common-gate) of each transistor in the modernized schematic, identifying the function of each in the operational amplifier circuit.

Challenges

- Explain the function of the Zener diodes, and of their equivalent in the original K2-W circuit.
- Is a D-type MOSFET a normally-on or normally-off device?

6.1.10 Voltage follower analysis

Determine the output voltage of this "voltage follower" circuit given the following input voltage values:



- $V_{in} = -20$ V ; $V_{out} =$
- $V_{in} = -15$ V ; $V_{out} =$
- $V_{in} = -10$ V ; $V_{out} =$
- $V_{in} = -5$ V ; $V_{out} =$
- $V_{in} = 0$ V ; $V_{out} =$
- $V_{in} = 5$ V ; $V_{out} =$
- $V_{in} = 10$ V ; $V_{out} =$
- $V_{in} = 15$ V ; $V_{out} =$
- $V_{in} = 20$ V ; $V_{out} =$

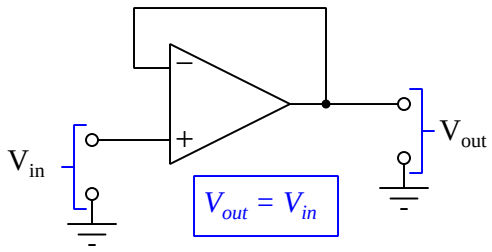
Challenges

- Explain what happens when $V_{in} = \pm$ 20 Volts.
- Some operational amplifiers have "rail-to-rail" output voltage capabilities, while others do not. What difference would this distinction make in this scenario?

6.1.11 Voltage follower versus a wire

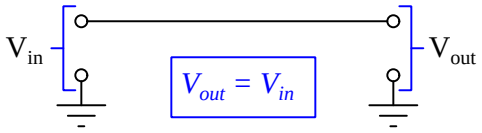
The following operational amplifier circuit is often referred to as a *voltage buffer*, because it has unity gain and therefore simply reproduces, or "buffers", the input voltage:

Voltage buffer circuit



What possible use is a circuit such as this, which offers no voltage gain or any other form of signal modification? Wouldn't a straight piece of wire do the same thing? Explain your answers.

A simpler voltage buffer?

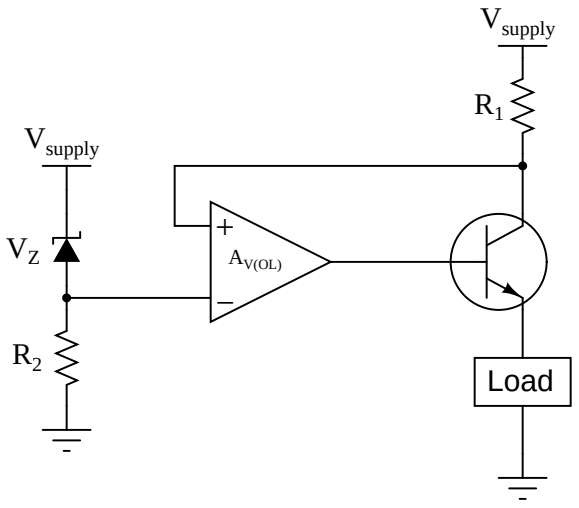


Challenges

- ???.
- ???.
- ???.

6.1.12 Unusual current-regulating circuit

At first appearance, the feedback appears to be wrong in this current-regulating circuit. Note how the feedback signal goes to the operational amplifier's *noninverting* (+) input, rather than the inverting input as one would normally expect for negative feedback:



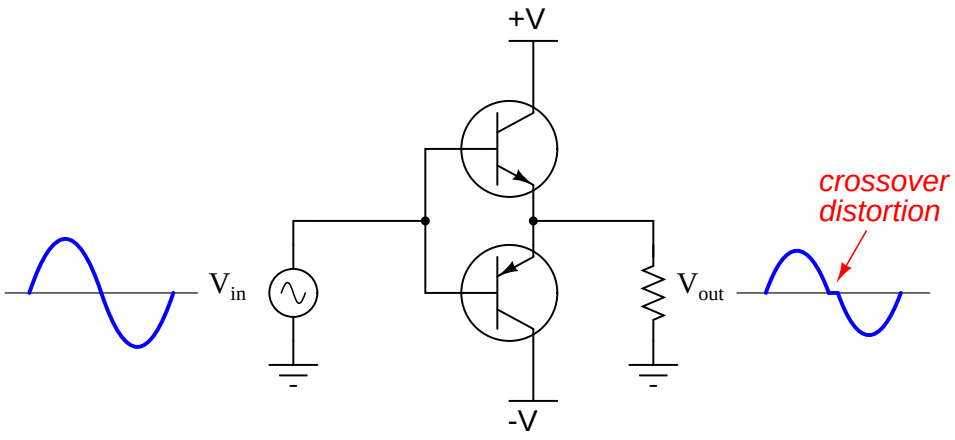
Explain how this op-amp really does provide *negative* feedback, which of course is necessary for stable current regulation, as positive feedback would be completely unstable.

Challenges

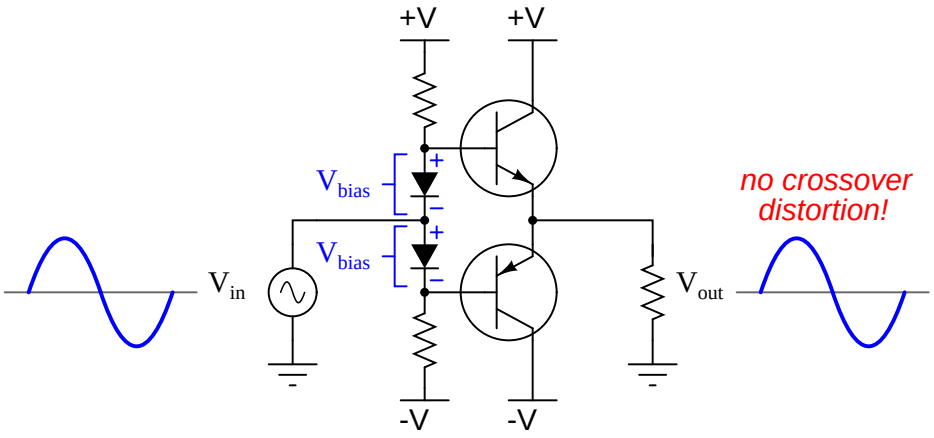
- What would happen if the Zener diode failed shorted?
- What would happen if resistor R_1 failed open?

6.1.13 Push-pull voltage follower

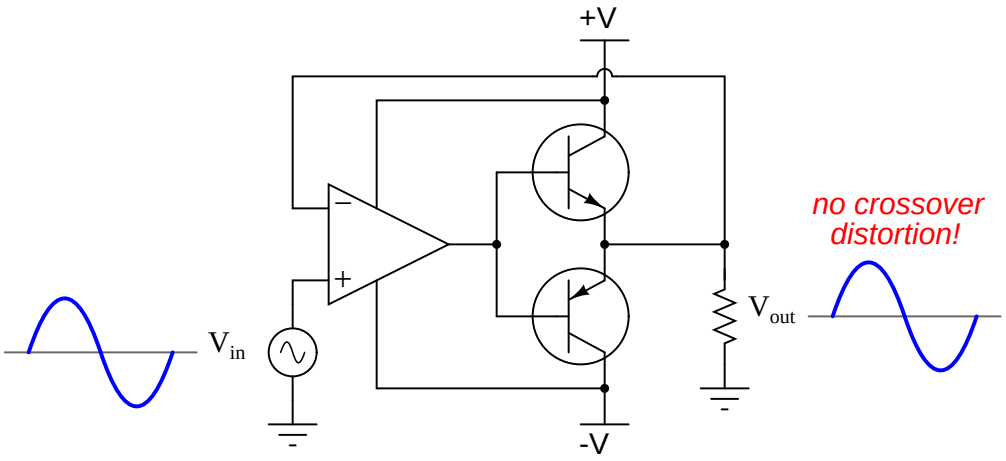
A complementary push-pull transistor amplifier built exactly as shown would perform rather poorly, exhibiting crossover distortion:



The simplest way to reduce or eliminate this distortion is by adding some bias voltage to each of the transistors' inputs, so there will never be a period of time when the two transistors are simultaneously cutoff:

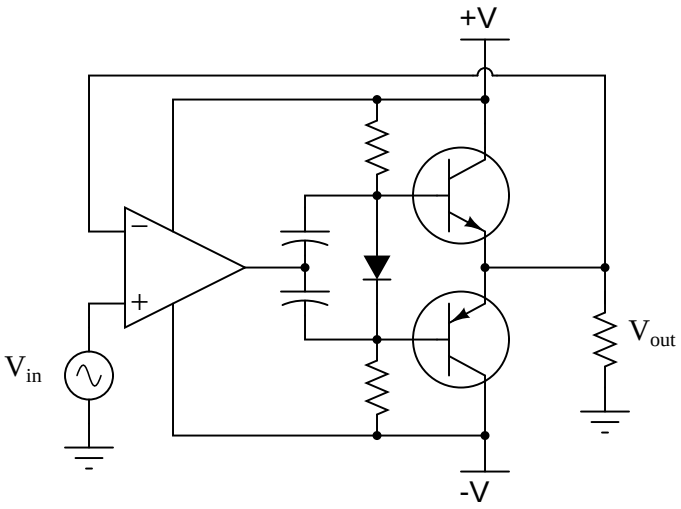


One problem with this solution is that just a little too much bias voltage will result in overheating of the transistors, as they simultaneously conduct current near the zero-crossing point of the AC signal. A more sophisticated method of mitigating crossover distortion is to use an opamp with negative feedback, like this:



Explain how the opamp is able to eliminate crossover distortion in this push-pull amplifier circuit without the need for biasing.

A more practical design blends the two strategies like this:



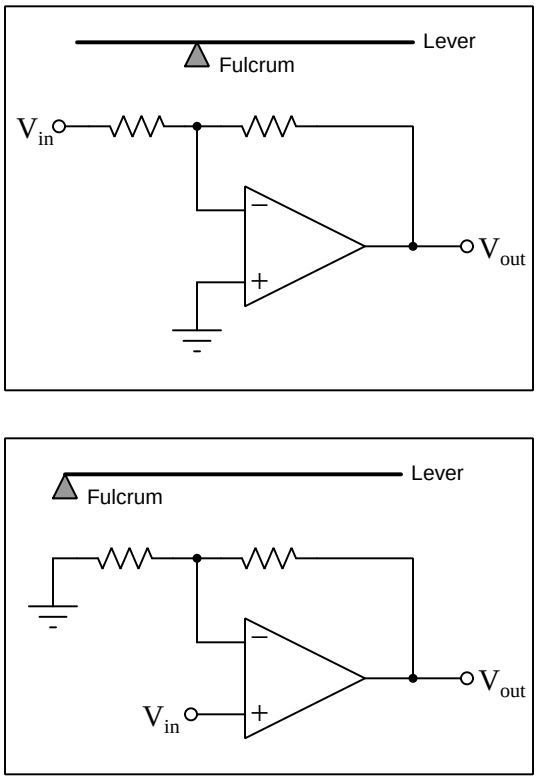
Explain why using less bias voltage *and* negative feedback with an opamp results in better performance than either method used alone.

Challenges

- It should be noted that this solution to crossover distortion in a push-pull amplifier circuit does not always yield perfect results. Identify at least one reason why the opamp cannot totally eliminate crossover distortion.

6.1.14 Electronic levers

Operational amplifier circuits employing negative feedback are sometimes referred to as "electronic levers" because their voltage gains may be understood through the mechanical analogy of a lever. Explain this analogy in your own words, identifying how the lengths and fulcrum location of a lever relate to the component values of an op-amp circuit:



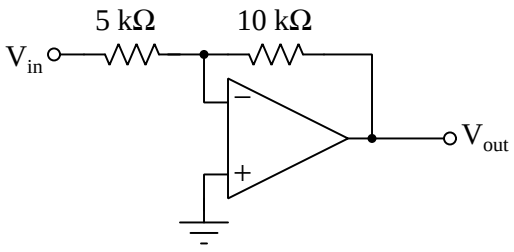
Challenges

- Levers are multipliers of force or motion, but they are not true *amplifiers*. Explain why not.
- In the lever analogy, is voltage analogous to *force* or to *motion*?

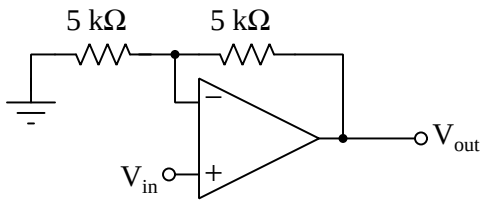
6.1.15 Comparison of input impedances

Shown here are two different voltage amplifier circuits with the same voltage gain. Which of them has greater input impedance, and why? Try to give as specific an answer for each circuit's input impedance as possible.

Inverting amplifier circuit



Noninverting amplifier circuit

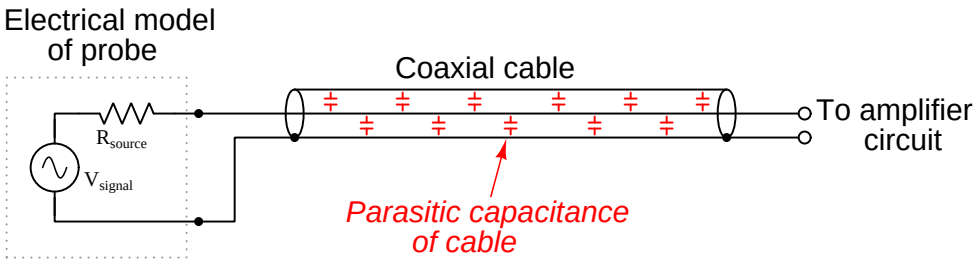


Challenges

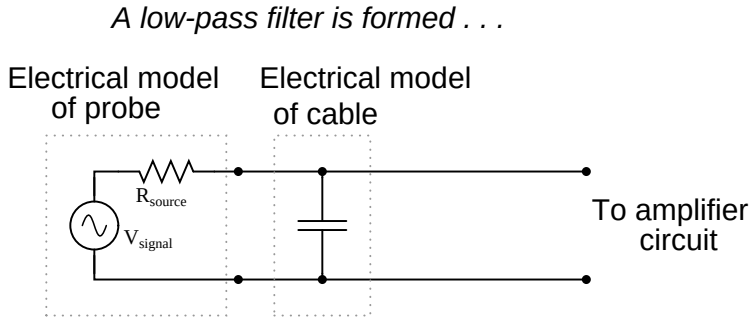
- Explain how to apply the definition of impedance ($Z_{in} = \frac{V_{in}}{I_{in}}$) to this circuit as an aid in answering the question.

6.1.16 Cable guarding

Parasitic capacitance naturally existing in two-wire cables can cause problems when connected to high-impedance electronic devices. Take for instance certain biomedical probes used to detect electrochemical events in living tissue. Such probes may be modeled as voltage sources in series with resistances, those resistances usually being rather large due to the probes' very small surface (contact) areas:

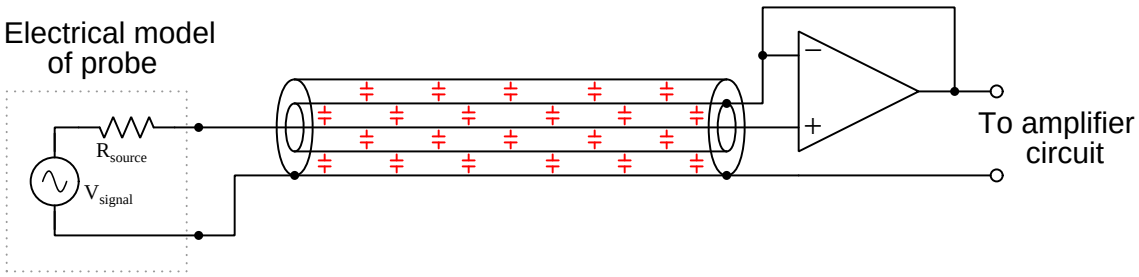


When connected to a cable with parasitic capacitance, a low-pass RC filter circuit is formed:

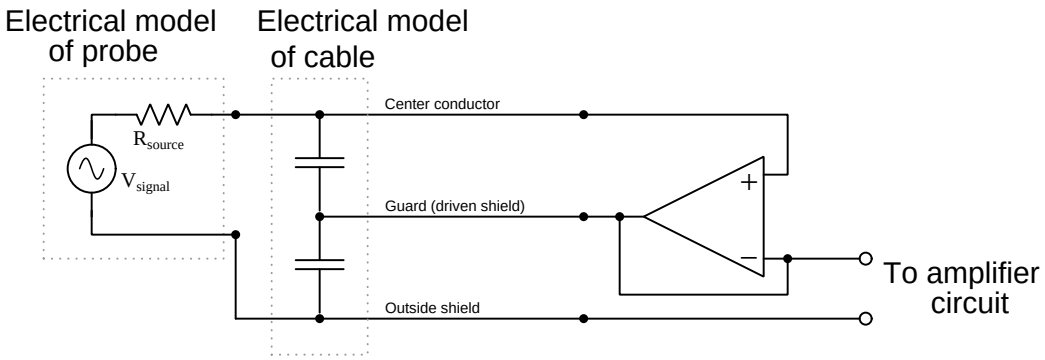


This low-pass filter (or passive integrator, if you wish) is purely unintentional. No one asked for it to be there, but it is there anyway just due to the natural resistance of the probe and the natural capacitance of the cable. Ideally, of course, we would like to be able to send the signal voltage (V_{signal}) straight to the amplifier with no interference or filtering of any kind so we can see exactly what it is we're trying to measure.

One clever way of practically eliminating the effects of cable capacitance is to encase the signal wire in its own shield, and then drive that shield with the exact same amount of voltage from a voltage follower at the other end of the cable. This is called *guarding*:



An equivalent schematic may make this technique more understandable:



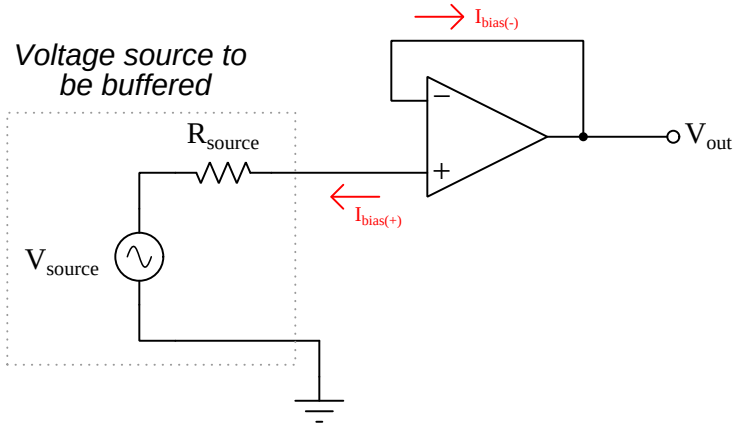
Explain why guarding the signal wire effectively eliminates the effects of the cable's capacitance. Certainly the capacitance is still present, so how can it not have any effect on the weak signal any more?

Challenges

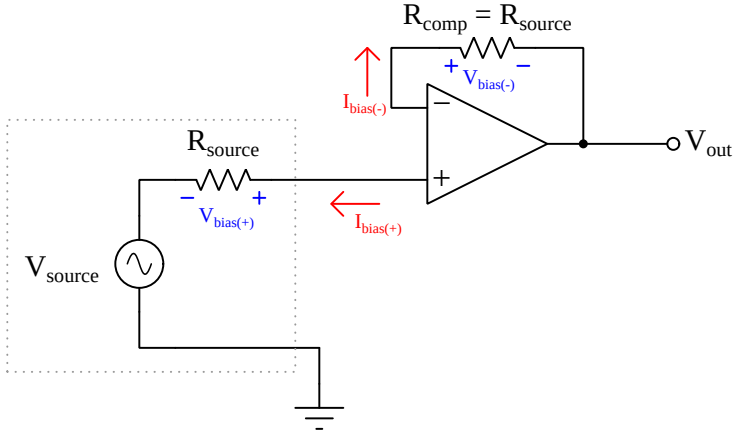
- Although the center-to-guard capacitance may have zero Volts across it at all times thanks to the opamp, the guard-to-(outer) shield capacitance still has the full signal voltage across it. Explain why this is of no concern to us, and why its presence does not form a low-pass filter as the original (unguarded) cable capacitance once did.

6.1.17 Bias current compensation

Ideally, the input terminals of an operational amplifier conduct zero current, allowing us to simplify the analysis of many opamp circuits. However, in actuality there is a very small amount of current going through each of the input terminals of any opamp with BJT input circuitry. This may cause unexpected voltage errors in circuits. Consider the following voltage buffer circuit:



$I_{bias(-)}$ does not cause any trouble for us, because it is completely supplied by the opamp's output. The other bias current, though, *does* cause trouble, because it must go through the source's Thévenin equivalent resistance. When it does, it drops some voltage across that intrinsic source resistance, skewing the amount of voltage actually seen at the noninverting terminal of the opamp. A common solution to this is to add another resistor to the circuit, like this:



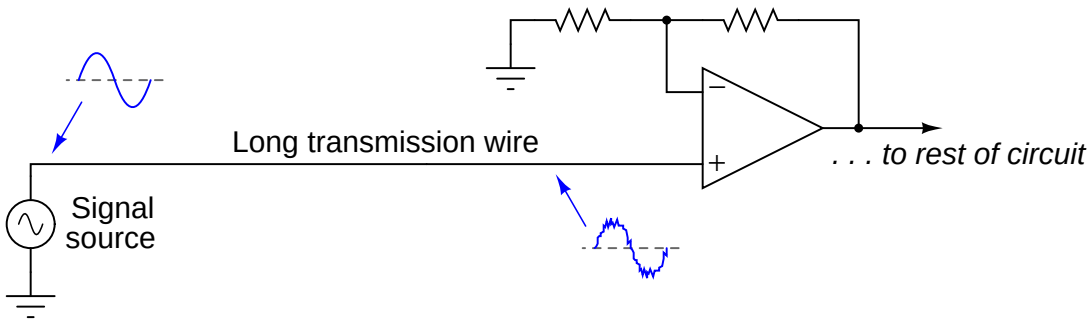
Explain why the addition of a resistor fixes the problem.

Challenges

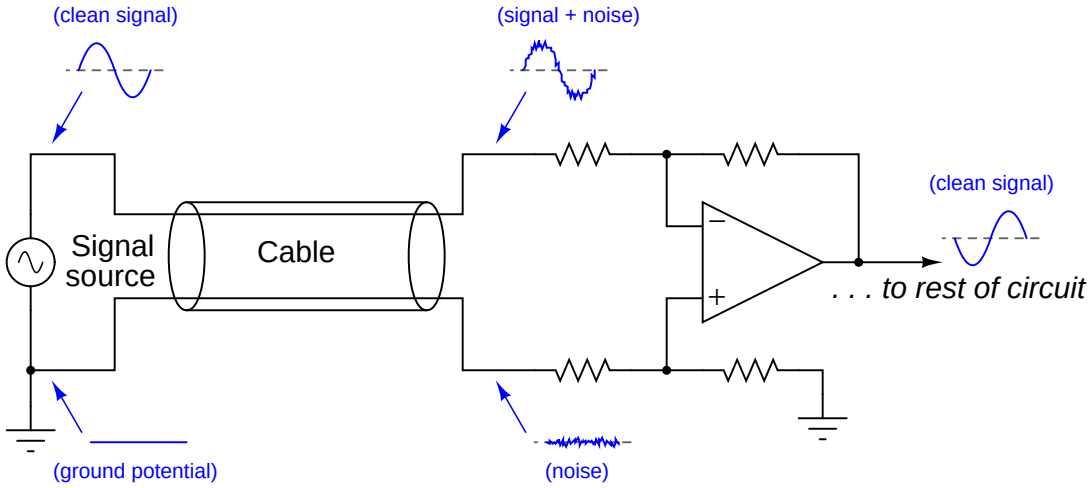
- This solution hinges on a critical assumption about bias currents for an operational amplifier. What is this assumption?

6.1.18 Differential signaling

If a weak voltage signal is conveyed from a source to an amplifier, the amplifier may detect more than just the desired signal. Along with the desired signal, external electronic "noise" may be coupled to the transmission wire from AC sources such as power line conductors, radio waves, and other electromagnetic interference sources. Note the two waveshapes, representing voltages along the transmission wire measured with reference to earth ground:



Shielding of the transmission wire is always a good idea in electrically noisy environments, but there is a more elegant solution than simply trying to shield interference from getting to the wire. Instead of using a single-ended amplifier to receive the signal, we can transmit the signal along *two* wires and use a *difference* amplifier at the receiving end. Note the four waveforms shown, representing voltages at those points measured with reference to earth ground:



If the two wires are run parallel to each other the whole distance, so as to be exposed to the exact same noise sources along that distance, the noise voltage at the end of the bottom wire will be the same noise voltage as that superimposed on the signal at the end of the top wire.

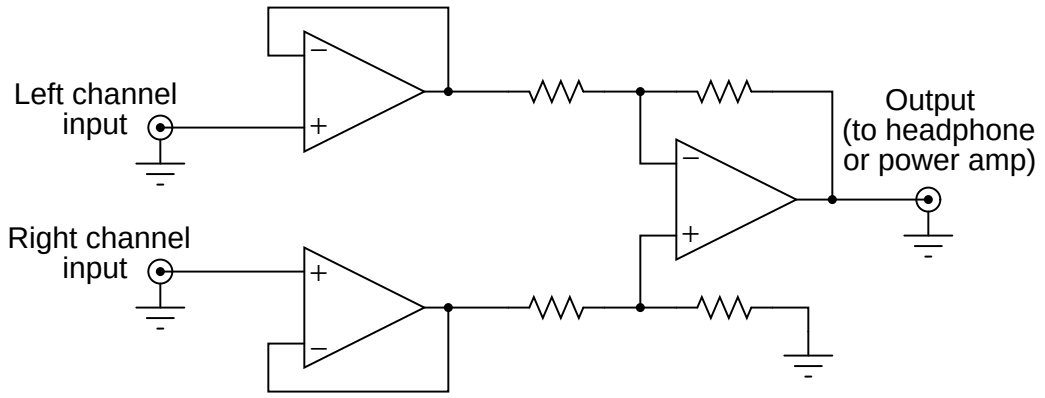
Explain how the difference amplifier is able to restore the original (clean) signal voltage from the two noise-ridden voltages seen at its inputs with respect to ground, and also how the phrase *common-mode voltage* applies to this scenario.

Challenges

- Re-draw the original (one wire plus ground) schematic to model the sources of interference and the wire's impedance, to show exactly how the signal could become mixed with noise from source to amplifier.

6.1.19 Vocal eliminator circuit

Singers who wish to practice singing to popular music find that the following *vocal eliminator* circuit is useful:



The circuit works on the principle that vocal tracks are usually recorded through a single microphone at the recording studio, and thus are represented equally on each channel of a stereo sound system. This circuit effectively eliminates the vocal track from the song, leaving only the music to be heard through the headphone or speaker.

Explain how the operational amplifiers accomplish this task of vocal track elimination. What role does each opamp play in this circuit?

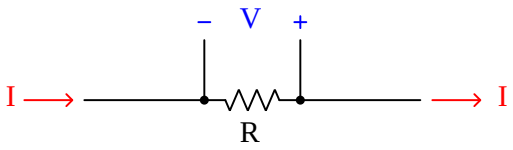
Unfortunately, the circuit as shown tends to eliminate bass tones as well as vocals, since the acoustic properties of bass tones make them represented nearly equally on both channels. Determine how the circuit may be expanded to include opamps that re-introduce bass tones to the "vocal-eliminated" output.

Challenges

- One of my students, when faced with the second question, suggested placing a high-pass filter before *one* of the subtractor's inputs, eliminating bass tones at one of the inputs and therefore reproducing bass tones at the subtractor output. This is a great idea, and shows what can happen when students are given a forum to think creatively and freely express ideas, but there are some practical reasons it would be difficult to implement. Identify at least one reason why this might not work.

6.1.20 Phototube amplifier

The simplest electronic device capable of converting a current signal into a voltage signal is a resistor:

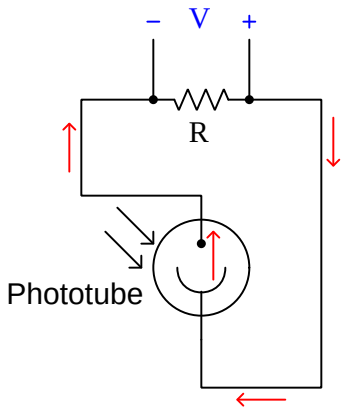


(electron flow notation used here)

Precision resistors typically work very well for this purpose, especially when the amount of voltage dropped across it is of little consequence. This is why *shunt resistors* are frequently used in power circuitry to measure current, a low-resistance "shunt" resistance element dropping voltage in precise proportion to the current going through it.

However, if we cannot afford to drop any voltage across a resistance in the circuit, this technique of current-to-voltage conversion will not be very practical. Consider the following scientific apparatus, used to measure the photoelectric effect (electrons emitted from a solid surface due to light striking it):

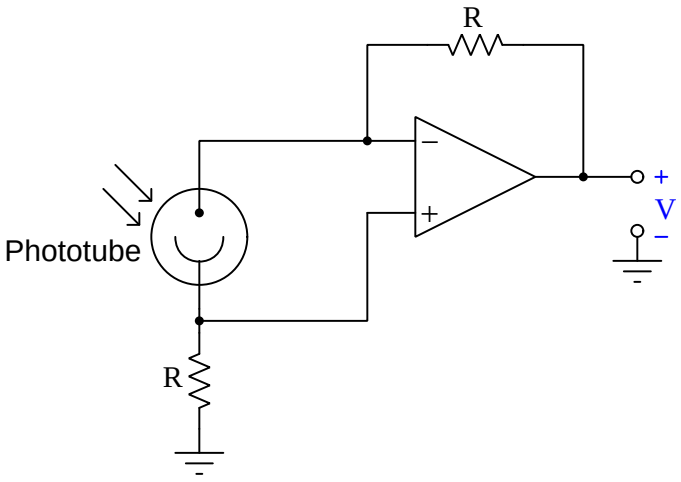
An impractical way to measure phototube current



(electron flow notation used here)

The current output by such a phototube is *very* small, and the voltage output by it is smaller yet. If we are to measure current through this device, we will have to find some way other than a shunt resistor to do it.

Enter the operational amplifier, to the rescue! Explain how the following opamp circuit is able to convert the phototube's weak current signal into a strong voltage signal, without imposing any significant resistance into the phototube circuit:

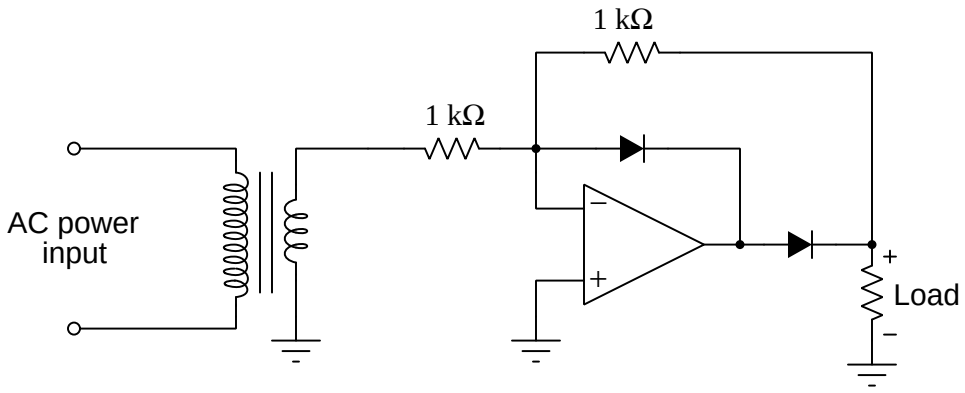


Challenges

- Why do you suppose we departed from our normal use of conventional flow notation to explain the action of a photocell?

6.1.21 Opamp power rectifier?

Explain why the following opamp circuit cannot be used as a power rectifier in an AC-DC power supply circuit:

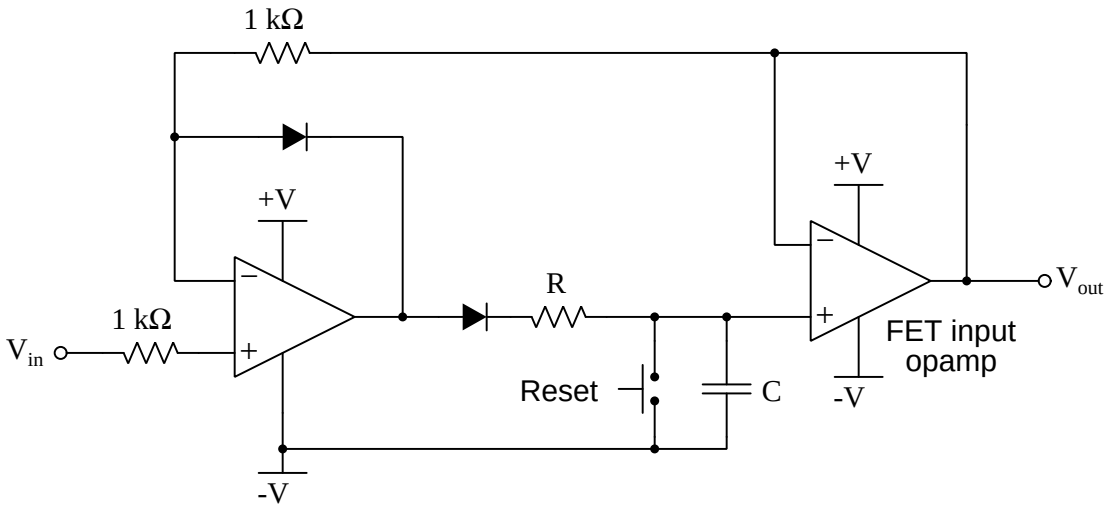


Challenges

- If this circuit is impractical as a power rectifier, what might it actually be used for?

6.1.22 Peak follower-and-hold circuit

This circuit is referred to as a *peak follower-and-hold*, taking the last greatest positive input voltage and "holding" that value at the output until a greater positive input voltage comes along:



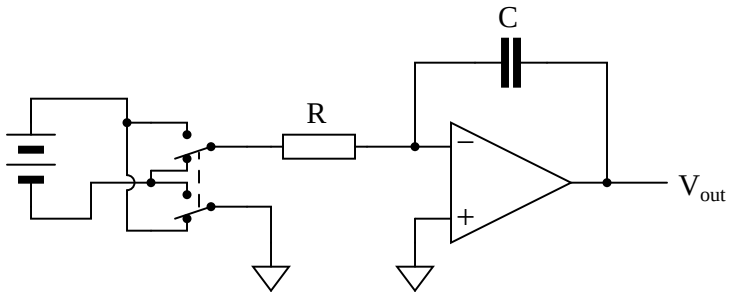
Give a brief explanation of how this circuit works, as well as the purpose and function of the "reset" switch. Also, explain why a FET input opamp is required for the last stage of amplification.

Challenges

- What purpose is served by resistor connected in series with the diode?

6.1.23 Integrator with DPDT switch

What will the output voltage of this integrator circuit do when the DPDT ("Double-Pole, Double-Throw") switch is flipped back and forth?



Be as specific as you can in your answer, explaining what happens in the switch's "up" position as well as in its "down" position.

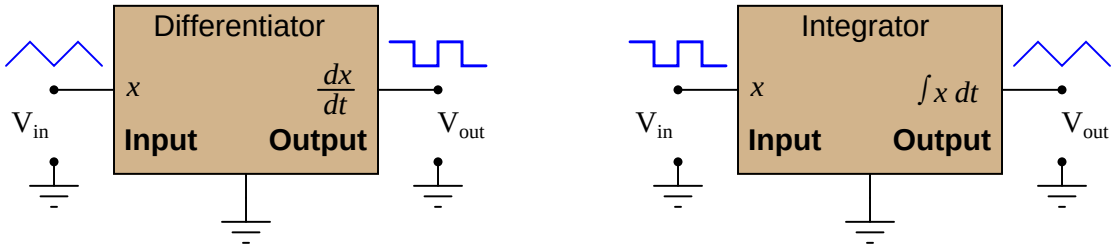
Challenges

- Identify ways we could modify this circuit to create a more aggressive response to the input signal.

6.1.24 Position, velocity, and acceleration signals

A familiar context in which to apply and understand basic principles of calculus is the motion of an object, in terms of *position* (x), *velocity* (v), and *acceleration* (a). We know that velocity is the time-derivative of position ($v = \frac{dx}{dt}$) and that acceleration is the time-derivative of velocity ($a = \frac{dv}{dt}$). Another way of saying this is that velocity is the rate of position change over time, and that acceleration is the rate of velocity change over time.

It is easy to construct circuits which input a voltage signal and output either the *time-derivative* or the *time-integral* (the opposite of the derivative) of that input signal. We call these circuits "differentiators" and "integrators", respectively.



Integrator and differentiator circuits are highly useful for motion signal processing, because they allow us to take voltage signals from motion sensors and convert them into signals representing other motion variables. For each of the following cases, determine whether we would need to use an integrator circuit or a differentiator circuit to convert the first type of motion signal into the second:

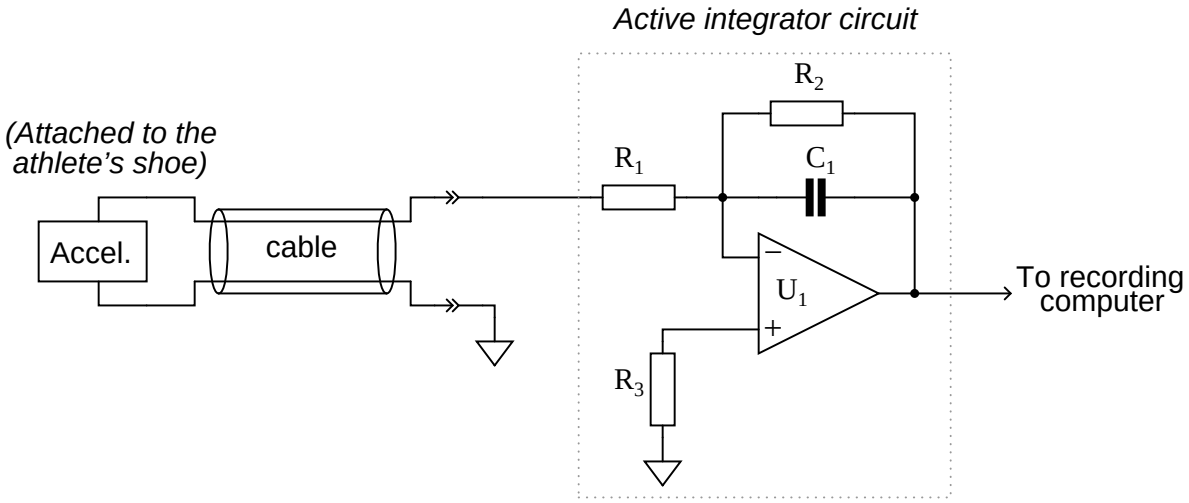
- Converting velocity signal to position signal: (*integrator* or *differentiator*?)
- Converting acceleration signal to velocity signal: (*integrator* or *differentiator*?)
- Converting position signal to velocity signal: (*integrator* or *differentiator*?)
- Converting velocity signal to acceleration signal: (*integrator* or *differentiator*?)
- Converting acceleration signal to position signal: (*integrator* or *differentiator*?)

Challenges

- Explain why a differentiator converts a triangle-wave input into a square-wave output.
- Explain why an integrator converts a square-wave input into a triangle-wave output.

6.1.25 Athletic accelerometer

This active integrator circuit processes the voltage signal from an *accelerometer*, a device that outputs a DC voltage proportional to its physical acceleration. The accelerometer is being used to measure the acceleration of an athlete's foot as she kicks a ball, and the job of the integrator is to convert that acceleration signal into a *velocity* signal so the researchers can record the velocity of the athlete's foot on a wearable computer that the athlete will wear during practice sessions and games:



During the set-up for this test, a radar gun is used to check the velocity of the athlete's foot as she does some practice kicks, and compare against the output of the integrator circuit. What the researchers find is that the integrator's output is reading a bit low. In other words, the integrator circuit is not integrating fast enough to provide an accurate representation of foot velocity.

Determine which component(s) in the integrator circuit may have been improperly sized to cause this calibration problem. Be as specific as you can in your answer(s).

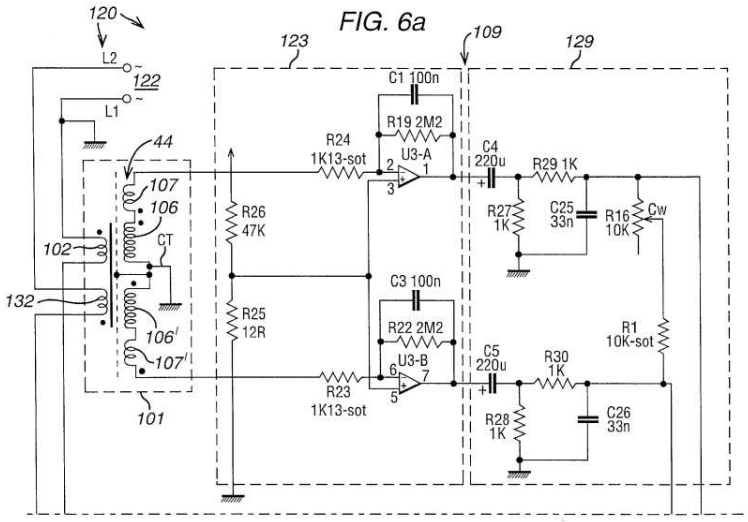
Challenges

- What is the purpose of resistor R_2?

6.1.26 Rogowski coil power meter

The following figures come from US Patent 6,414,475 ("Current Sensor") by inventors Andrew Nicholas Dames and Edward Crellier Colby (filed 6 November 2000, granted 2 July 2002), the two figures intended to go together to form a single schematic diagram showing circuitry for an electric power meter sensing residential AC voltage and current, and using a Rogowski coil to sense that line current:

FIG. 6a



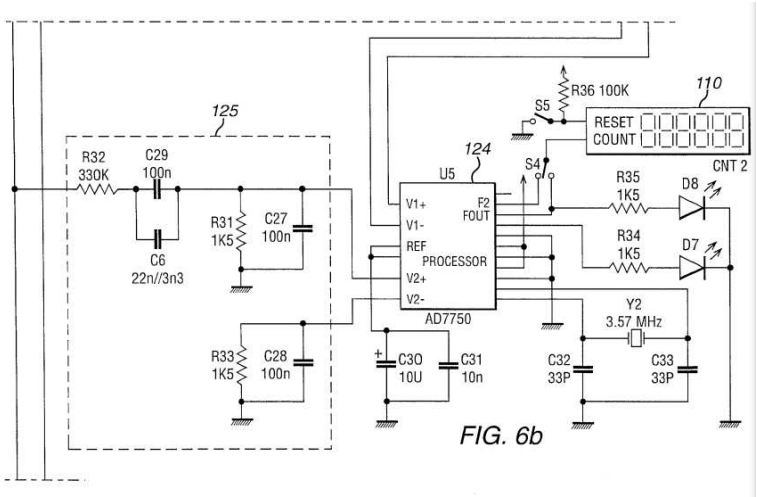


FIG. 6b

Identify the following from this schematic diagram:

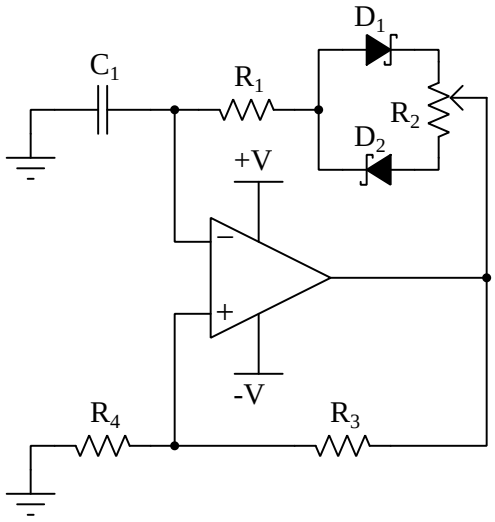
- The location of the Rogowski coil
- The purpose of the two operational amplifiers
- The purpose of capacitors C_4 and C_5
- The purpose of resistors R_{19} and R_{22}
- The purpose of the resistors and capacitors immediately to the right of C_4 and C_5
- The location of the voltage-conditioning network
- The type of signal output by the microcontroller ("processor")
- Where does the AC power source connect, and where does the customer's load(s) connect?

Challenges

- Which component value(s) might be adjusted to calibrate the power meter's voltage-sensing accuracy?
- Within the first paragraph of the patent's Abstract, the authors write, "The coils are arranged so that their magnetic axis are co-located and aligned together so that they also provide a null response to extraneous magnetic fields having a field gradient". Explain what this means.

6.1.27 Adjustable oscillator

Identify the purpose of the potentiometer in the following oscillator circuit:



Also, answer the following questions about this circuit:

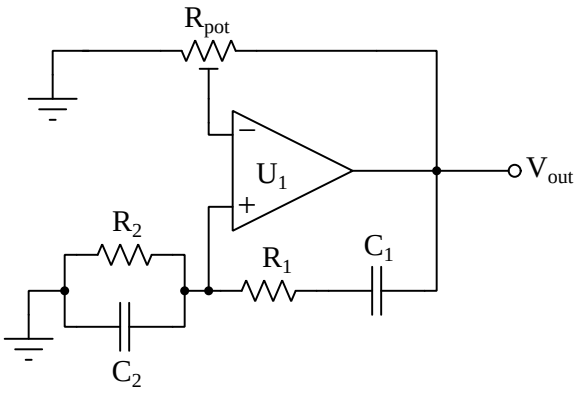
- Where in this circuit would you measure a square waveform?
- Where in this circuit would you measure a sawtooth waveform?
- What type of diodes are being used in this circuit, and why?

Challenges

- Will changes in DC supply voltage affect this oscillator's frequency?
- Will changes in DC supply voltage affect this oscillator's square wave duty cycle?

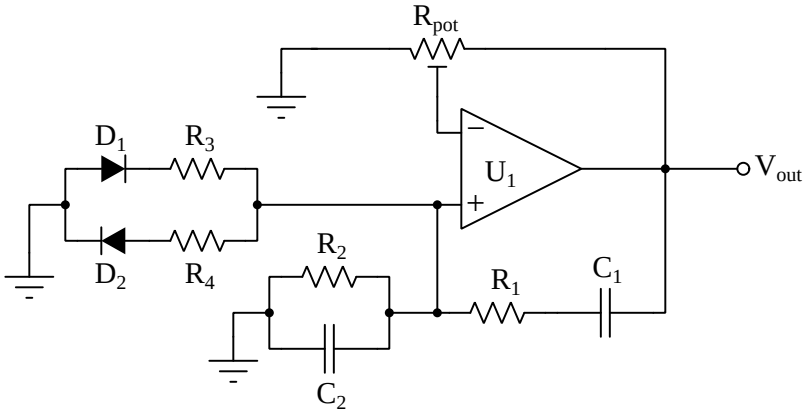
6.1.28 Limiting network

This Wein bridge oscillator circuit is very sensitive to changes in the gain. Note how the potentiometer used in this circuit is the "trimmer" variety, adjustable with a screwdriver rather than by a knob or other hand control:



The reason for this choice in potentiometers is to make accidental changes in circuit gain less probable. If you build this circuit, you will see that tiny changes in this potentiometer's setting make a huge difference in the quality of the output sine wave. A little too much gain, and the sine wave becomes noticeably distorted. Too little gain, and the circuit stops oscillating altogether!

Obviously, it is not good to have such sensitivity to minor changes in any practical circuit expected to reliably perform day after day. One solution to this problem is to add a *limiting network* to the circuit comprised of two diodes and two resistors:



With this network in place, the circuit gain may be adjusted well above the threshold for oscillation (Barkhausen criterion) without exhibiting excessive distortion as it would have without the limiting network. Explain why the limiting network makes this possible.

Challenges

- What effect does this "limiting network" have on the purity of the oscillator's output signal spectrum? In other words, does the limiting network increase or decrease the harmonic content of the output waveform?

6.2 Quantitative reasoning

These questions are designed to stimulate your computational thinking. In a Socratic discussion with your instructor, the goal is for these questions to reveal your mathematical approach(es) to problem-solving so that good technique and sound reasoning may be reinforced. Your instructor may also pose additional questions based on those assigned, in order to observe your problem-solving firsthand.

Mental arithmetic and estimations are strongly encouraged for all calculations, because without these abilities you will be unable to readily detect errors caused by calculator misuse (e.g. keystroke errors).

You will note a conspicuous lack of answers given for these quantitative questions. Unlike standard textbooks where answers to every other question are given somewhere toward the back of the book, here in these learning modules students must rely on other means to check their work. My advice is to use circuit simulation software such as SPICE to check the correctness of quantitative answers. Refer to those learning modules within this collection focusing on SPICE to see worked examples which you may use directly as practice problems for your own study, and/or as templates you may modify to run your own analyses and generate your own practice problems.

Completely worked example problems found in the Tutorial may also serve as "test cases[4]" for gaining proficiency in the use of circuit simulation software, and then once that proficiency is gained you will never need to rely[5] on an answer key!

[4] In other words, set up the circuit simulation software to analyze the same circuit examples found in the Tutorial. If the simulated results match the answers shown in the Tutorial, it confirms the simulation has properly run. If the simulated results disagree with the Tutorial's answers, something has been set up incorrectly in the simulation software. Using every Tutorial as practice in this way will quickly develop proficiency in the use of circuit simulation software.

[5] This approach is perfectly in keeping with the instructional philosophy of these learning modules: *teaching students to be self-sufficient thinkers*. Answer keys can be useful, but it is even more useful to your long-term success to have a set of tools on hand for checking your own work, because once you have left school and are on your own, there will no longer be "answer keys" available for the problems you will have to solve.

6.2.1 Miscellaneous physical constants

Note: constants shown in **bold** type are *exact*, not approximations. Values inside of parentheses show one standard deviation (σ) of uncertainty in the final digits: for example, the magnetic permeability of free space value given as $1.25663706212(19) \times 10^{-6}$ H/m represents a center value (i.e. the location parameter) of $1.25663706212 \times 10^{-6}$ Henrys per meter with one standard deviation of uncertainty equal to $0.0000000000019 \times 10^{-6}$ Henrys per meter.

Avogadro's number (N_A) = **6.02214076** $\times 10^{23}$ **per mole** (mol^{-1})

Boltzmann's constant (k) = **1.380649** $\times 10^{-23}$ **Joules per Kelvin** (J/K)

Electronic charge (e) = **1.602176634** $\times 10^{-19}$ **Coulomb** (C)

Faraday constant (F) = **96,485.33212...** $\times 10^{4}$ **Coulombs per mole** (C/mol)

Magnetic permeability of free space (μ_0) = $1.25663706212(19) \times 10^{-6}$ Henrys per meter (H/m)

Electric permittivity of free space (ϵ_0) = $8.8541878128(13) \times 10^{-12}$ Farads per meter (F/m)

Characteristic impedance of free space (Z_0) = 376.730313668(57) Ohms (Ω)

Gravitational constant (G) = $6.67430(15) \times 10^{-11}$ cubic meters per kilogram-seconds squared ($m^3/kg\text{-}s^2$)

Molar gas constant (R) = **8.314462618... Joules per mole-Kelvin** (J/mol-K) = 0.08205746(14) liters-atmospheres per mole-Kelvin

Planck constant (h) = **6.62607015** $\times 10^{-34}$ **joule-seconds** (J-s)

Stefan-Boltzmann constant (σ) = **5.670374419...** $\times 10^{-8}$ **Watts per square meter-Kelvin**^{4} ($W/m^2 \cdot K^4$)

Speed of light in a vacuum (c) = **299,792,458 meters per second** (m/s) = 186282.4 miles per second (mi/s)

Note: All constants taken from NIST data "Fundamental Physical Constants – Complete Listing", from `http://physics.nist.gov/constants`, National Institute of Standards and Technology (NIST), 2018 CODATA Adjustment.

6.2.2 Introduction to spreadsheets

A powerful computational tool you are encouraged to use in your work is a *spreadsheet*. Available on most personal computers (e.g. Microsoft Excel), *spreadsheet* software performs numerical calculations based on number values and formulae entered into cells of a grid. This grid is typically arranged as lettered columns and numbered rows, with each cell of the grid identified by its column/row coordinates (e.g. cell `B3`, cell `A8`). Each cell may contain a string of text, a number value, or a mathematical formula. The spreadsheet automatically updates the results of all mathematical formulae whenever the entered number values are changed. This means it is possible to set up a spreadsheet to perform a series of calculations on entered data, and those calculations will be re-done by the computer any time the data points are edited in any way.

For example, the following spreadsheet calculates average speed based on entered values of distance traveled and time elapsed:

	A	B	C	D
1	Distance traveled	46.9	Kilometers	
2	Time elapsed	1.18	Hours	
3	Average speed	= B1 / B2	km/h	
4				
5				

Text labels contained in cells `A1` through `A3` and cells `C1` through `C3` exist solely for readability and are not involved in any calculations. Cell `B1` contains a sample distance value while cell `B2` contains a sample time value. The formula for computing speed is contained in cell `B3`. Note how this formula begins with an "equals" symbol (`=`), references the values for distance and speed by lettered column and numbered row coordinates (`B1` and `B2`), and uses a forward slash symbol for division (`/`). The coordinates `B1` and `B2` function as *variables*[6] would in an algebraic formula.

When this spreadsheet is executed, the numerical value `39.74576` will appear in cell `B3` rather than the formula `= B1 / B2`, because 39.74576 is the computed speed value given 46.9 kilometers traveled over a period of 1.18 hours. If a different numerical value for distance is entered into cell `B1` or a different value for time is entered into cell `B2`, cell `B3`'s value will automatically update. All you need to do is set up the given values and any formulae into the spreadsheet, and the computer will do all the calculations for you.

Cell `B3` may be referenced by other formulae in the spreadsheet if desired, since it is a variable just like the given values contained in `B1` and `B2`. This means it is possible to set up an entire chain of calculations, one dependent on the result of another, in order to arrive at a final value. The arrangement of the given data and formulae need not follow any pattern on the grid, which means you may place them anywhere.

[6] Spreadsheets may also provide means to attach text labels to cells for use as variable names (Microsoft Excel simply calls these labels "names"), but for simple spreadsheets such as those shown here it's usually easier just to use the standard coordinate naming for each cell.

Common[7] arithmetic operations available for your use in a spreadsheet include the following:

- Addition (`+`)
- Subtraction (`-`)
- Multiplication (`*`)
- Division (`/`)
- Powers (`^`)
- Square roots (`sqrt()`)
- Logarithms (`ln()` , `log10()`)

Parentheses may be used to ensure[8] proper order of operations within a complex formula. Consider this example of a spreadsheet implementing the *quadratic formula*, used to solve for roots of a polynomial expression in the form of $ax^2 + bx + c$:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

	A	B
1	x_1	= (-B4 + sqrt((B4^2) - (4*B3*B5))) / (2*B3)
2	x_2	= (-B4 - sqrt((B4^2) - (4*B3*B5))) / (2*B3)
3	a =	9
4	b =	5
5	c =	-2

This example is configured to compute roots[9] of the polynomial $9x^2 + 5x - 2$ because the values of 9, 5, and -2 have been inserted into cells `B3`, `B4`, and `B5`, respectively. Once this spreadsheet has been built, though, it may be used to calculate the roots of *any* second-degree polynomial expression simply by entering the new a, b, and c coefficients into cells `B3` through `B5`. The numerical values appearing in cells `B1` and `B2` will be automatically updated by the computer immediately following any changes made to the coefficients.

[7] Modern spreadsheet software offers a bewildering array of mathematical functions you may use in your computations. I recommend you consult the documentation for your particular spreadsheet for information on operations other than those listed here.

[8] Spreadsheet programs, like text-based programming languages, are designed to follow standard order of operations by default. However, my personal preference is to use parentheses even where strictly unnecessary just to make it clear to any other person viewing the formula what the intended order of operations is.

[9] Reviewing some algebra here, a *root* is a value for x that yields an overall value of zero for the polynomial. For this polynomial ($9x^2 + 5x - 2$) the two roots happen to be $x = 0.269381$ and $x = -0.82494$, with these values displayed in cells `B1` and `B2`, respectively upon execution of the spreadsheet.

Alternatively, one could break up the long quadratic formula into smaller pieces like this:

$$y = \sqrt{b^2 - 4ac} \qquad z = 2a$$

$$x = \frac{-b \pm y}{z}$$

	A	B	C
1	x_1	= (-B4 + C1) / C2	= sqrt((B4^2) - (4*B3*B5))
2	x_2	= (-B4 - C1) / C2	= 2*B3
3	a =	9	
4	b =	5	
5	c =	-2	

Note how the square-root term (y) is calculated in cell C1, and the denominator term (z) in cell C2. This makes the two final formulae (in cells B1 and B2) simpler to interpret. The positioning of all these cells on the grid is completely arbitrary[10] – all that matters is that they properly reference each other in the formulae.

Spreadsheets are particularly useful for situations where the same set of calculations representing a circuit or other system must be repeated for different initial conditions. The power of a spreadsheet is that it automates what would otherwise be a tedious set of calculations. One specific application of this is to simulate the effects of various components within a circuit failing with abnormal values (e.g. a shorted resistor simulated by making its value nearly zero; an open resistor simulated by making its value extremely large). Another application is analyzing the behavior of a circuit design given new components that are out of specification, and/or aging components experiencing drift over time.

[10] My personal preference is to locate all the "given" data in the upper-left cells of the spreadsheet grid (each data point flanked by a sensible name in the cell to the left and units of measurement in the cell to the right as illustrated in the first distance/time spreadsheet example), sometimes coloring them in order to clearly distinguish which cells contain entered data versus which cells contain computed results from formulae. I like to place all formulae in cells below the given data, and try to arrange them in logical order so that anyone examining my spreadsheet will be able to figure out *how* I constructed a solution. This is a general principle I believe all computer programmers should follow: *document and arrange your code to make it easy for other people to learn from it.*

6.2.3 Empirically determining open-loop gain

Determine the open-loop gain of this operational amplifier based on the voltage measurements shown below:

$V_{in(+)}$	$V_{in(-)}$	V_{out}
1.0000 Volts	1.00003 Volts	1.5000 Volts
1.0000 Volts	1.00004 Volts	6.8000 Volts

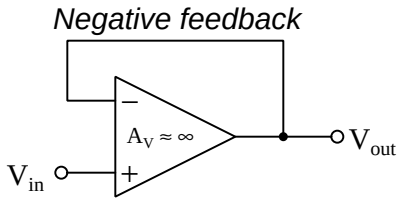
Also, write a mathematical formula solving for differential voltage gain (A_V) in terms of an op-amp's input and output voltages.

Challenges

- Convert this voltage gain figure (as a ratio) into a voltage gain figure in decibels.

6.2.4 Voltage follower simulation program

A *voltage follower* opamp circuit is formed by connecting the opamp's output terminal directly to its inverting input terminal as shown in the following schematic:

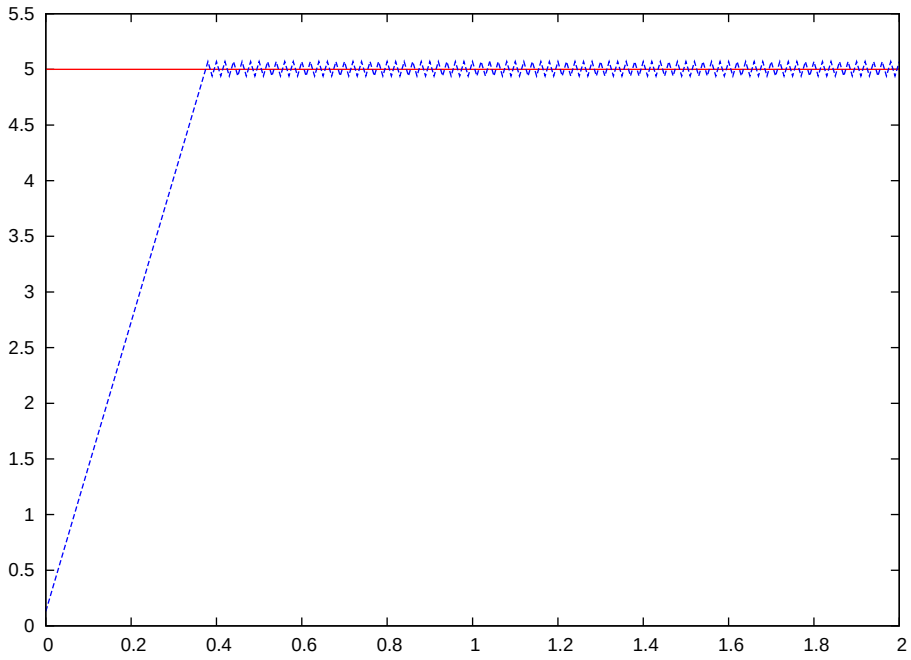


The following computer program (written in the C language) simulates such a voltage follower tracking a V_{in} signal of +5 Volts over a time period of 2 μs in 0.01 μs intervals:

```
#include <stdio.h>

int main (void)
{
  float Vin = 5.0, Vout = 0.0, time, slew = 13.0, interval = 0.01;
  // time and interval in microseconds, slew rate in Volts/microsecond

  printf("Time , Vin , Vout\n");

  for (time = 0 ; time < 2.0 ; time = time + interval)
  {
    if (Vin > Vout)
      Vout = Vout + (slew * interval);

    else if (Vin < Vout)
      Vout = Vout - (slew * interval);

    printf("%f , %f , %f\n", time, Vin, Vout);
  }

  return 0;
}
```

Note that *slew rate* (symbolized by the variable `slew` in this program) is a practical performance limitation of any opamp, referring to the maximum rate at which its output may rise or fall over time. This parameter is usually specified in units of *Volts per microsecond*, the value in this simulation being set to 13 V/μs.

When this program is compiled and executed, it outputs data in plain-text comma-separated variable format, which may be input to a graphical plotting program (e.g. spreadsheet) and displayed as a time-domain graph:



First, identify where on the graph the opamp's slew rate may be observed.

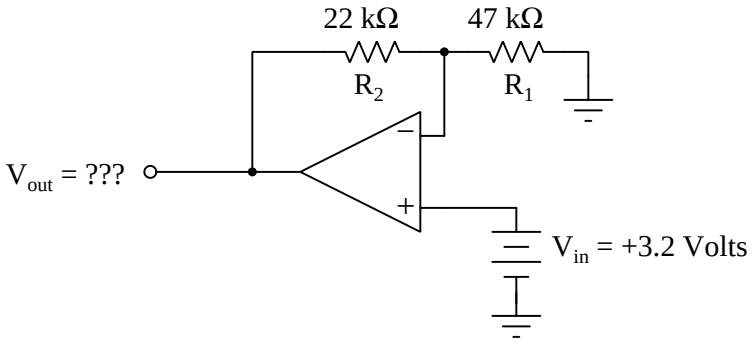
Next, explain why the graphed data shows the opamp's output voltage oscillating around +5 Volts rather than settling at that value. Then modify the simulation's code so that this oscillation is reduced or eliminated.

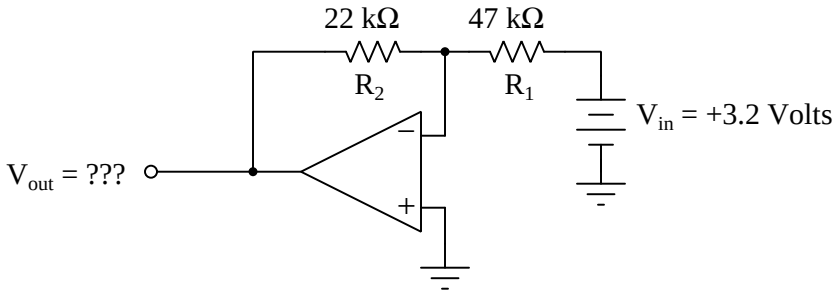
Challenges

- Modify the program so that the output voltage begins at some non-zero voltage value.

6.2.5 Calculating output voltages and gains

Calculate all voltage drops and currents in these circuits, complete with arrows for current direction and polarity markings for voltage polarity. Then, calculate their closed-loop voltage gains (A_V), both as a ratio and as a figure in units of decibels (dB):



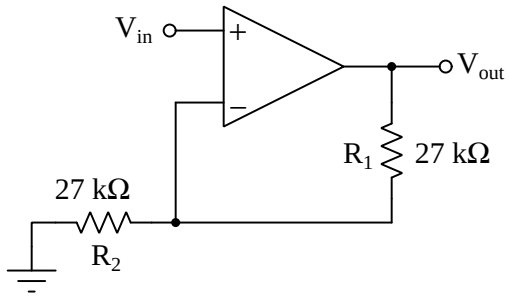


Challenges

- Determine resistor sizes which would give each of these amplifiers an (absolute) gain value of 2, or 6.021 dB.

6.2.6 Voltage gain calculations

Calculate the overall voltage gain of this amplifier circuit (A_V), both as a ratio and as a figure in units of decibels (dB). Also, write a general equation for calculating the voltage gain of such an amplifier, given the resistor values of R_1 and R_2:

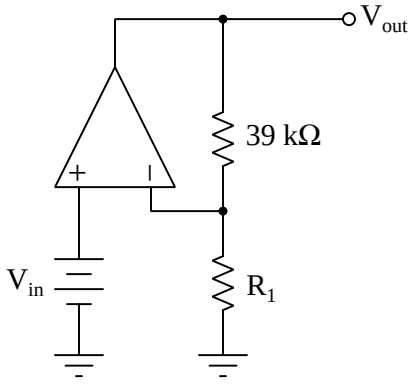


Challenges

- Explain how you could modify this particular circuit to have a voltage gain (ratio) of 7.

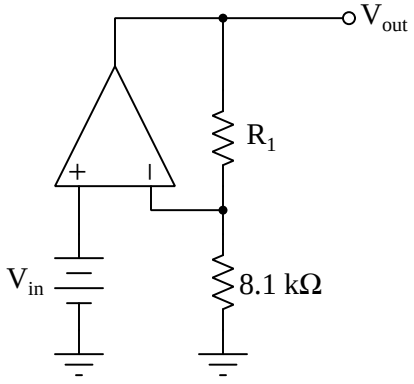
6.2.7 Necessary resistor values for specified voltage gains

Calculate the necessary resistor value (R_1) in this circuit to give it a voltage gain of 30:



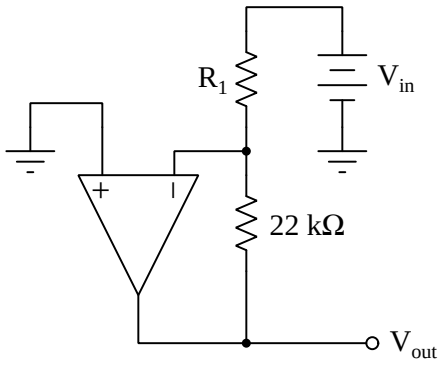
Now, choose the necessary size for R_1 to give this amplifier a voltage gain of 17.3 dB.

For this next amplifier circuit, calculate the necessary resistor value (R_1) in this circuit to give it a voltage gain of 10.5:



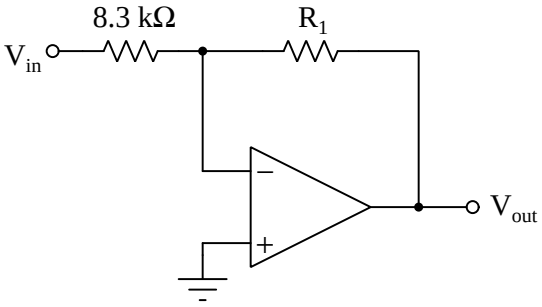
Now, choose the necessary size for R_1 to give this amplifier a voltage gain of 4.8 dB.

Calculate the necessary resistor value (R_1) in this circuit to give it a voltage gain of 15:



Now, choose the necessary size for R_1 to give this amplifier a voltage gain of 9.2 dB.

In this last circuit, calculate the necessary resistor value (R_1) to give it a voltage gain of 7.5:



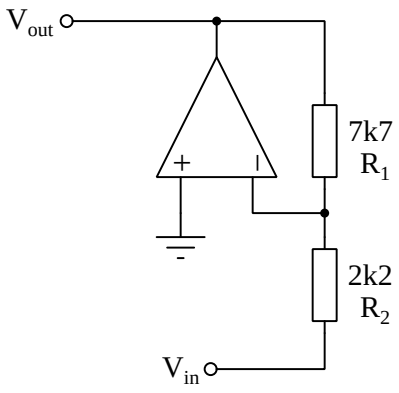
Now, choose the necessary size for R_1 to give this amplifier a voltage gain of 31 dB.

Challenges

- Suppose you were not aware of the voltage gain formulae for noninverting or inverting opamp circuits. Explain how you could still determine the proper resistor values in each case by assigning input and output voltage values corresponding to the desired voltage gain.
- Identify the largest and smallest voltage gain values possible with the first amplifier circuit, and the R_1 values necessary to achieve each of those limits.
- Identify the largest and smallest voltage gain values possible with the second amplifier circuit, and the R_1 values necessary to achieve each of those limits.
- Identify the largest and smallest voltage gain values possible with the third amplifier circuit, and the R_1 values necessary to achieve each of those limits.
- Identify the largest and smallest voltage gain values possible with the fourth amplifier circuit, and the R_1 values necessary to achieve each of those limits.

6.2.8 Modifying voltage gain

Determine the voltage gain of this circuit, and then determine the necessary value of R_2 to give it a new voltage gain value of -4.9:

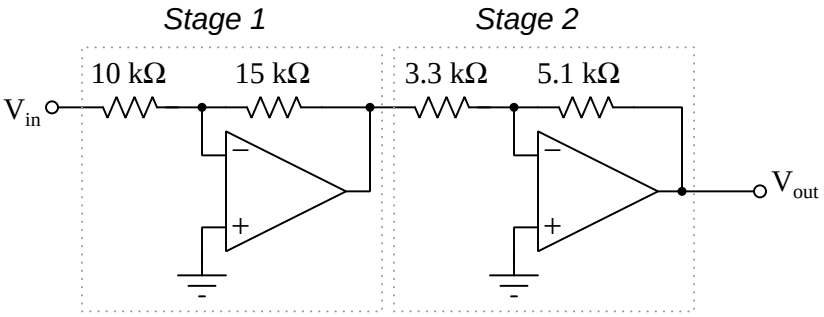


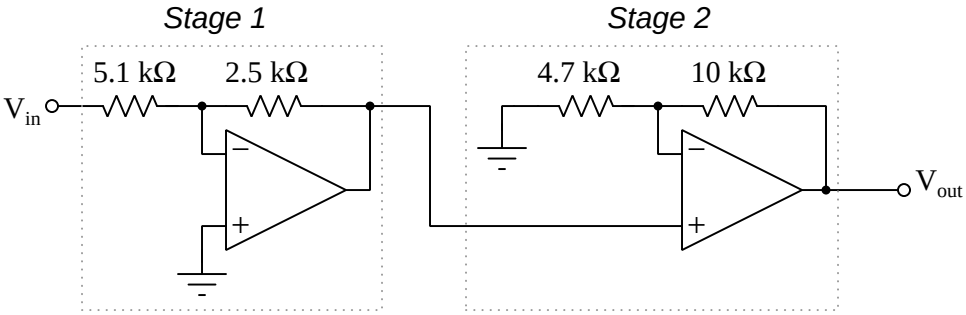
Challenges

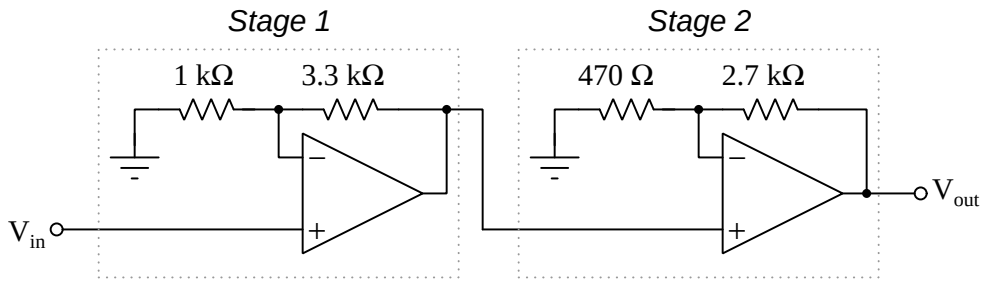
- Describe how to modify the gain of this circuit so that it is $+4.9$.
- Describe how to achieve the exact value of R_2 necessary to achieve the calculated gain, since 7.7 kΩ is not a common resistor size.

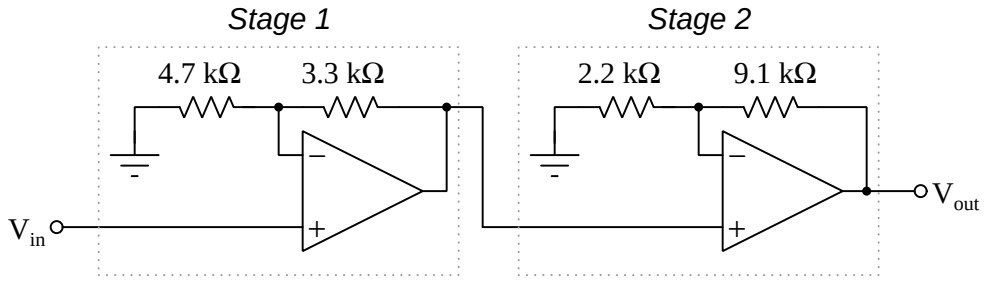
6.2.9 Multi-stage amplifier gains

Calculate the voltage gains for each stage of these amplifier circuits (both as a ratio and in units of decibels), then calculate the over-all voltage gain for each circuit:







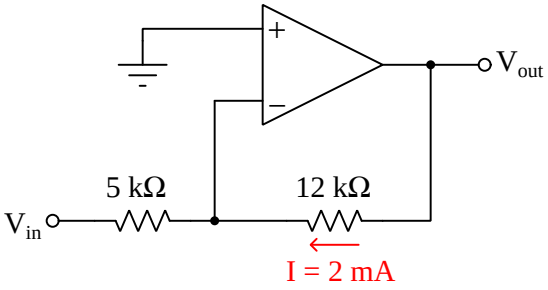


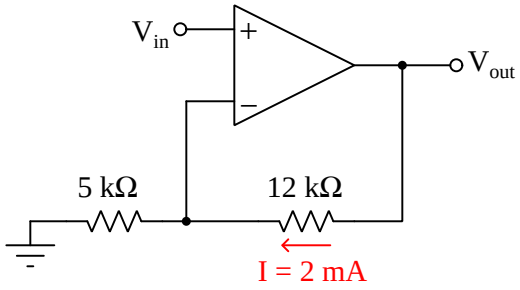
Challenges

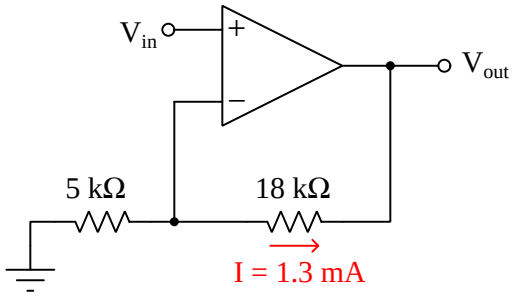
- What do you notice about the individual stage gains versus the over-all gains when expressed as ratios?
- What do you notice about the individual stage gains versus the over-all gains when expressed as decibels?

6.2.10 Calculating input and output voltages

Determine both the input and output voltage in these circuits:







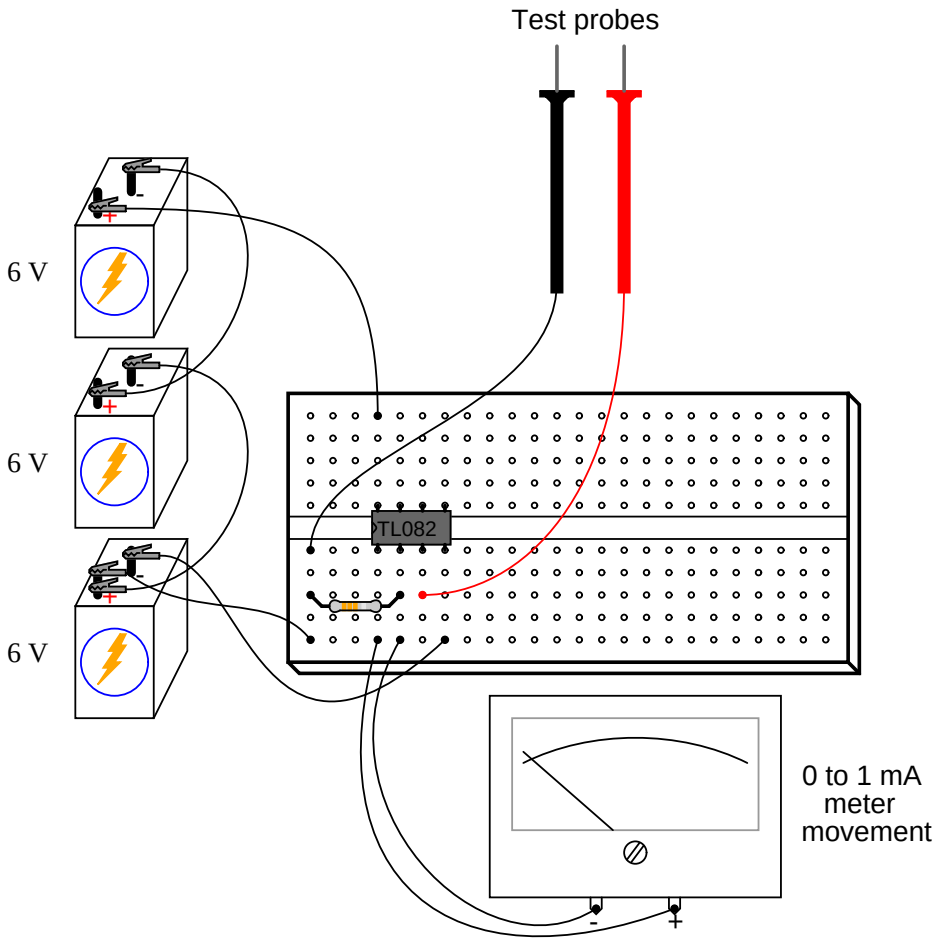
Finally, identify how some of the foundational concepts you've studied apply to these circuits: **Ohm's Law**, **Joule's Law**, **Kirchhoff's Voltage Law**, **Kirchhoff's Current Law**, properties of **series** and **parallel** networks, the **Cardinal Rule of Negative Feedback**, etc. Feel free to include any other relevant foundational concepts not listed here.

Challenges

- How are we able to discern the *polarities* of these two voltages?

6.2.11 High-resistance analog voltmeter

Shown here is a simple circuit for constructing an extremely high input impedance voltmeter on a wireless breadboard, using one half of a TL082 dual op-amp:



Draw a schematic diagram of this circuit, a calculate the resistor value necessary to give the meter a voltage measurement range of 0 to 5 Volts.

Challenges

- Determine the approximate input impedance of this voltmeter, and also the maximum voltage it is able to measure with *any* size resistor in the circuit.
- If the resistor's actual resistance happened to be slightly less than the calculated (ideal) value, what effect would that have on the meter's measurement accuracy, if any?

6.2.12 Opamp signal scaling circuits

Example #1

Design an operational amplifier circuit to scale an analog signal with a voltage range of 0 to +6 Volts to a new range of +4 Volts to −8 Volts, using a split DC power supply providing ±12 Volts to power the amplifier. Assume you are limited to using only E12 series resistor values, no potentiometers, with either LM7805 (+5 Volt) or LM7905 (−5 Volt) three-terminal regulator ICs for stable DC voltage references.

Example #2

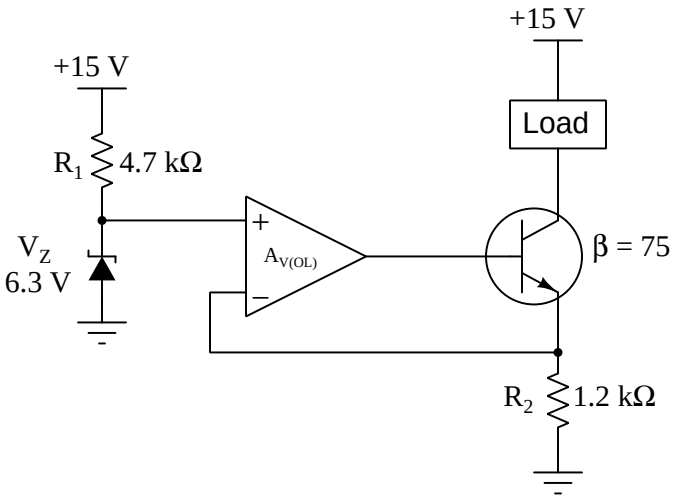
Design an operational amplifier circuit to read the voltage dropped across a current-sensing shunt resistor having a value of 50 milliOhms while passing current anywhere from 0 Amperes to 10 Amperes, the opamp circuit outputting a proportional signal to an analog-to-digital converter (ADC) circuit having a measurement range of 0 to +3.3 Volts. Assume the use of a split DC power supply providing ±12 Volts to power the amplifier, and that the shunt resistor is located on the "high-side" of a DC circuit where its common-mode potential is near +10 Volts at all times. Assume you are limited to using only E12 series resistor values, no potentiometers, with either LM7805 (+5 Volt) or LM7905 (−5 Volt) three-terminal regulator ICs for stable DC voltage references.

Challenges

- What is *common-mode* voltage, and why is it relevant in example #2?

6.2.13 Current regulator limits

Calculate the amount of current which will be regulated through the load in this circuit, then calculate the minimum and maximum load resistance values for which the circuit would be able to maintain that regulated current value. Assume the opamp is powered by the 15 Volt DC source and that it is capable of rail-to-rail output voltage swing:



Challenges

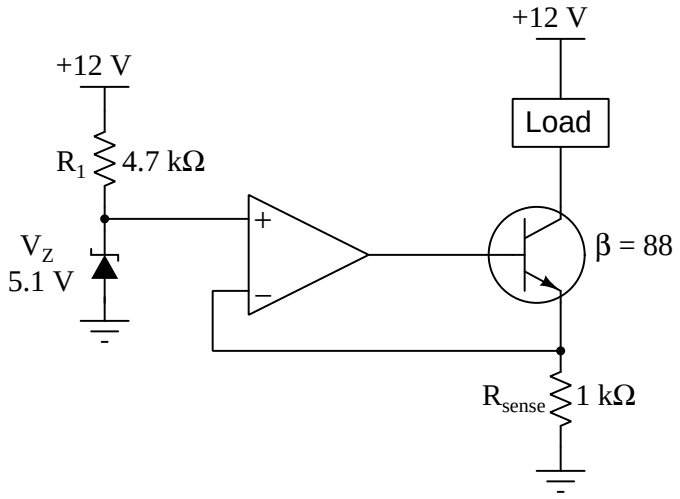
- A helpful problem-solving technique is the *thought experiment*, where we try to predict the effects of certain changes made to a system. In this case, we may imagine what will happen to all voltages and currents in the circuit as load resistance varies. Set up a thought experiment of your own, and try predicting the outcome of some load resistance change.
- Will the maximum load resistance depend at all on the opamp's ability to swing its output fully to the rail limits?
- Will the load's current be affected if the transistor's beta were to drift?

6.2.14 Current regulator simulation program

The following C-language source code simulates the behavior of a simple opamp-based current regulator circuit as it experiences different load resistances:

```
#include <stdio.h>

int main (void)
{
  float Rload, Iload, Ireg, Vce, Pq;
  float Vref = 5.1, Rsense = 1e3, beta = 88.0, Vsrc = 12.0;
  float Rloadmax = 2e3, Pqmax = 0.5, Vsat = 0.3;

  Ireg = Vref / Rsense;

  for (Rload = 0 ; Rload <= Rloadmax ; Rload = Rload + 200.0)
  {
    Iload = Ireg * beta / (beta + 1);
    Vce = Vsrc - Vref - (Iload * Rload);
    Pq = Iload * Vce + (Iload / beta * 0.7);

    if (Vce < Vsat)
      Iload = (Vsrc - Vref - Vsat) / Rload;

    printf("Rload = %.0f Ohms -- Iload = %0.3f mA", Rload, Iload*1e3);

    if (Iload < Ireg * beta / (beta + 1))
      printf(" -- Compliance voltage exceeded!");

    if (Pq > Pqmax)
      printf(" -- Transistor overloaded!");

    printf("\n");
  }

  return 0;
}
```

On the following page is the schematic diagram for this circuit, as well as a sample of this program's output when executed.




```
Rload = 0 Ohms -- Iload = 5.043 mA
Rload = 200 Ohms -- Iload = 5.043 mA
Rload = 400 Ohms -- Iload = 5.043 mA
Rload = 600 Ohms -- Iload = 5.043 mA
Rload = 800 Ohms -- Iload = 5.043 mA
Rload = 1000 Ohms -- Iload = 5.043 mA
Rload = 1200 Ohms -- Iload = 5.043 mA
Rload = 1400 Ohms -- Iload = 4.714 mA -- Compliance voltage exceeded!
Rload = 1600 Ohms -- Iload = 4.125 mA -- Compliance voltage exceeded!
Rload = 1800 Ohms -- Iload = 3.667 mA -- Compliance voltage exceeded!
Rload = 2000 Ohms -- Iload = 3.300 mA -- Compliance voltage exceeded!
```

Identify the foundational circuit concept(s) applied in each calculation represented in the source code.

Modify the parameters of this simulation so that the transistor actually over-loads for certain load resistance values.

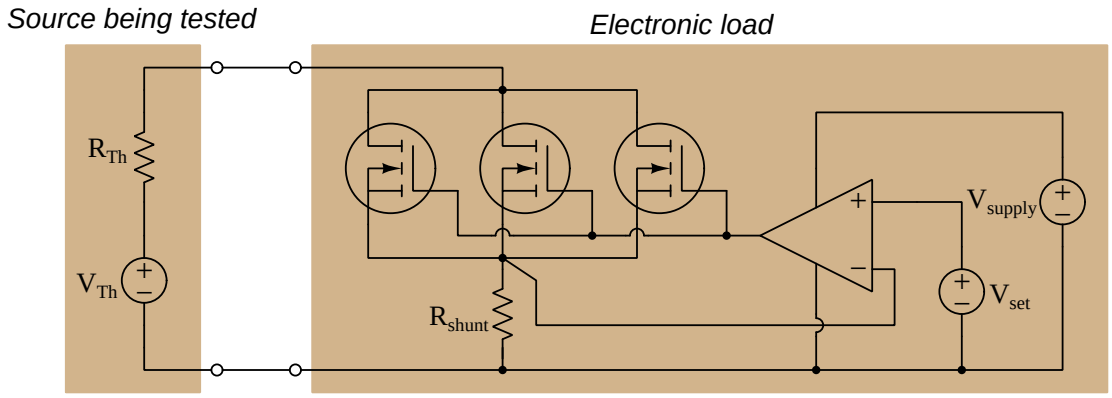
Challenges

- A helpful problem-solving technique is the *thought experiment*, where we try to predict the effects of certain changes made to a system. In this case, we may imagine what will happen to all voltages and currents in the circuit as load resistance varies. Set up a thought experiment of your own, and try predicting the outcome of some load resistance change.
- Will the maximum load resistance depend at all on the opamp's ability to swing its output fully to the rail limits?
- Will the load's current be affected if the transistor's beta were to drift?

- Why is the "new line" escape character located in a `printf` statement all on its own at the end of the program, rather than being included in any of the previous `printf` statements?

6.2.15 Constant-current electronic load circuit

The circuit on the right-hand side is called an *electronic load* because it functions as a load for testing the electrical source on the left-hand side:



In this case, the electronic load is designed to maintain a constant load current despite changes in source terminal voltage. Explain how the electronic load circuit accomplishes this task.

Supposing an R_{shunt} value of 250 milliOhms, a $V_{GS(on)}$ rating of 4.5 Volts for each MOSFET, a Thévenin source voltage of 15 Volts, and a Thévenin source resistance of 150 milliOhms, calculate the following parameters for a constant load current of 2 Amperes:

- $V_{terminals}$ =
- V_{set} =
- P_{shunt} =
- P_Q (power dissipated by all three MOSFETs) =
- V_{supply} assuming opamp has rail-to-rail output capability =

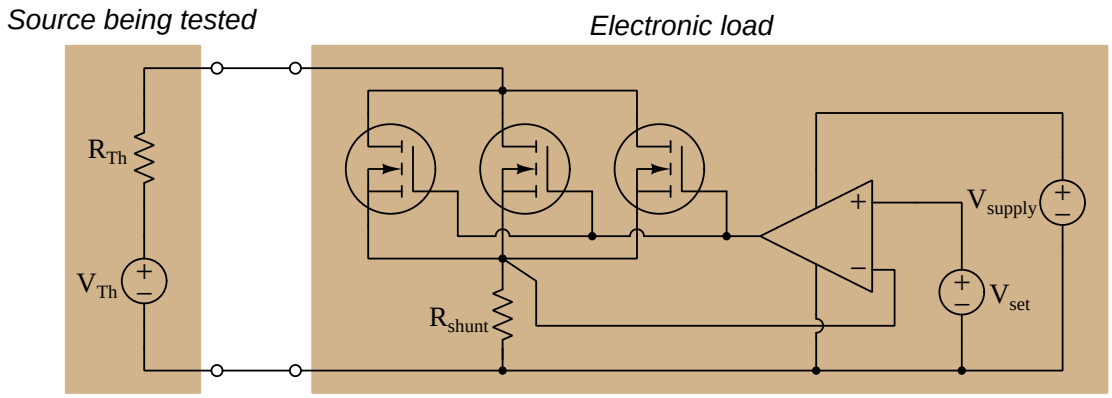
Finally, plot a *load line* for this electronic load circuit operating at 2 Amperes (constant), graphically showing the electronic load's terminal voltage (horizontal axis) versus its terminal current (vertical axis) supposing the source voltage were varied over a wide range.

Challenges

- Design a sub-circuit for the electronic load providing a V_{set} voltage using V_{supply} as the sole electrical source within the electronic load.
- What would happen if the source being tested were connected backwards to this electronic load circuit?
- Modify the electronic load circuit design such that it requires no internal V_{supply} at all.

6.2.16 Constant-voltage electronic load circuit

Sketch a schematic diagram for an opamp circuit to function as an electronic load regulating terminal *voltage* at some constant and pre-defined value. Shown below for reference is an electronic load circuit designed to regulate current connected to a Thévenized source:



Finally, plot a *load line* for this electronic load circuit operating at 10 Volts (constant), graphically showing the electronic load's terminal voltage (horizontal axis) versus its terminal current (vertical axis) supposing the source current were varied over a wide range.

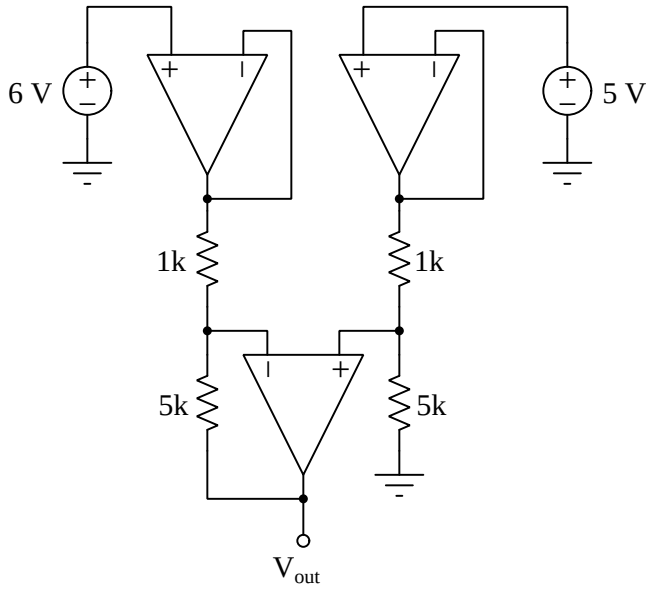
Challenges

- Design your electronic load circuit design in such a way that it will be able to operate on terminal voltage, requiring no internal V_{supply} at all.

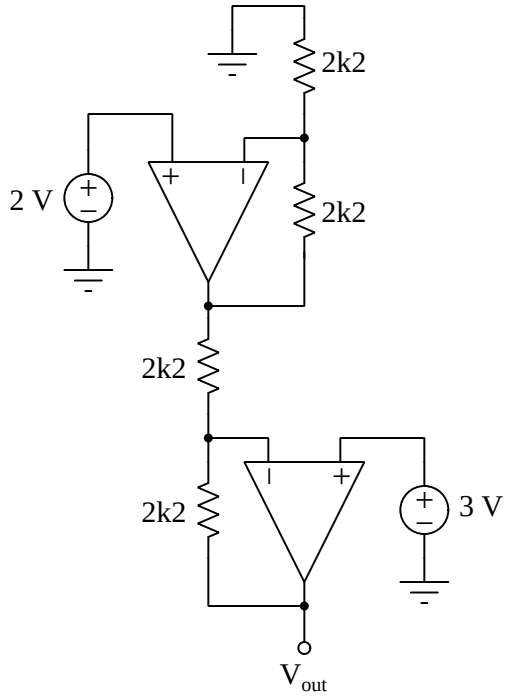
6.2.17 Subtractor calculations

Annotate all currents and voltage drops in the following circuits:

Example #1



Example #2



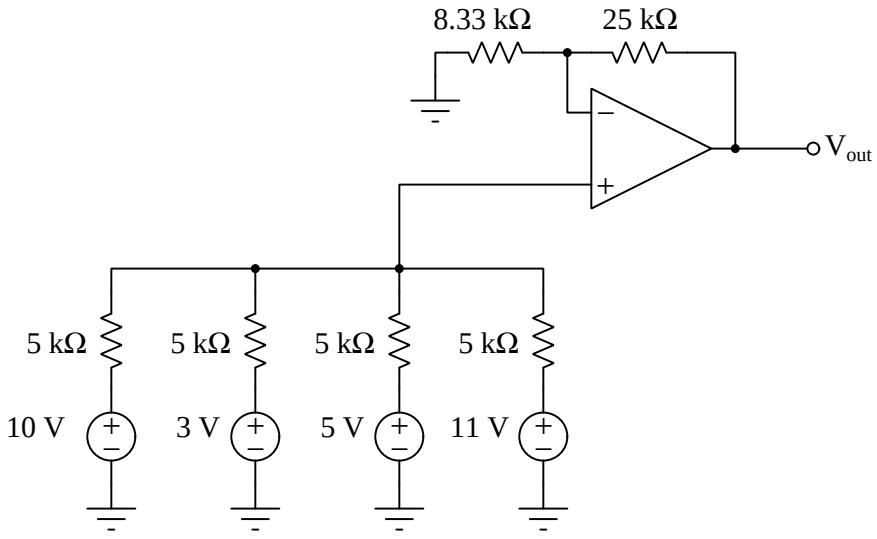
Challenges

- Calculate the voltage gain of each subtractor.
- Identify which opamps *source* current versus *sink* current in these examples.

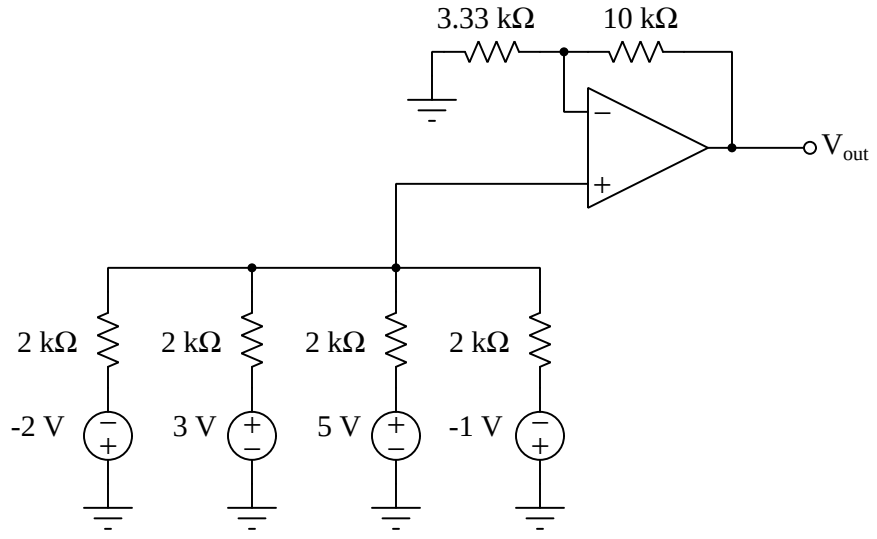
6.2.18 Summer circuit calculations

Annotate all currents and voltage drops in the following circuits:

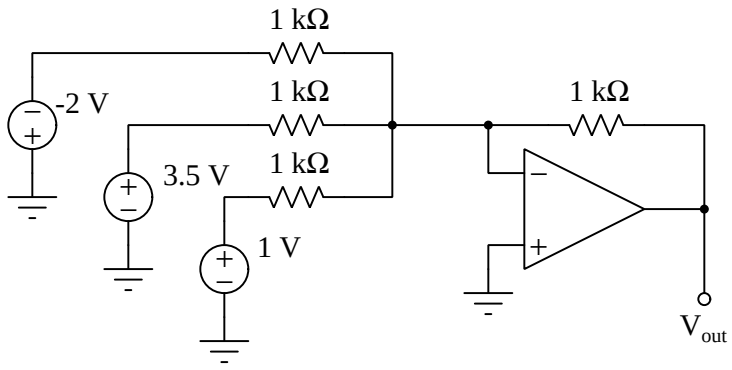
Example #1



Example #2



Example #3

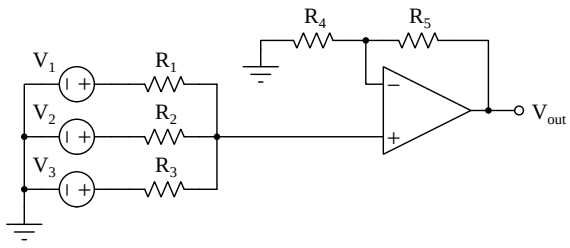


Challenges

- Which type of summer circuit, inverting or noninverting, has greater input impedance?

6.2.19 Summer circuit simulation program

Below we see a schematic diagram for a three-input non-inverting summer circuit followed by some C-language code simulating that circuit:




```
#include <stdio.h>

int main (void)
{
  float R[6]; // Array of resistor values R[1] through R[5]
  float V[4]; // Array of voltage source values V[1] through V[3]
  float Iavg, Vavg;
  int n;

  for (n = 1 ; n <= 3 ; ++n)
  {
    printf("Enter voltage of source #%i: ", n);
    scanf("%f", &V[n]);
  }

  for (n = 1 ; n <= 5 ; ++n)
  {
    printf("Enter resistance of resistor #%i: ", n);
    scanf("%f", &R[n]);
  }

  Iavg = V[1] / R[1] + V[2] / R[2] + V[3] / R[3];  // Total Norton current
  Vavg = Iavg * 1 / (1/R[1] + 1/R[2] + 1/R[3]);

  printf("Output voltage = %f Volts\n", Vavg * (R[5] / R[4] + 1));

  return 0;
}
```

Explain how the use of variable *arrays* (`R[1]`, `R[2]`, etc.) rather than distinct variables (`R1`, `R2`, etc.) for the user-entered component values helps make the program's source code more compact than it would be otherwise.

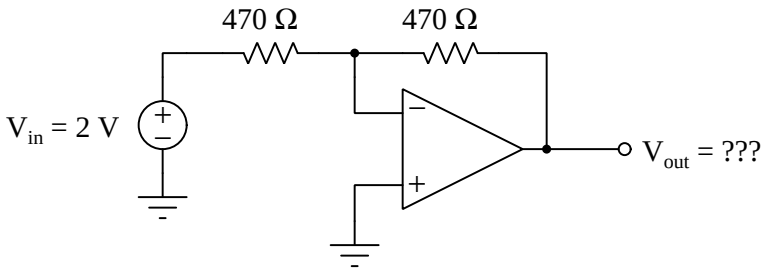
Also explain the comment "`Total Norton current`" and how its associated line of code works in the context of the summer circuit's simulation.

Challenges

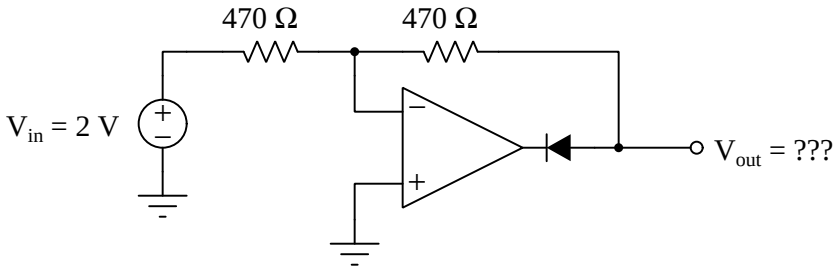
- What would happen in this circuit if resistors R_1 through R_3 were not equal in resistance as per usual?

6.2.20 Adding diodes to an inverting amplifier

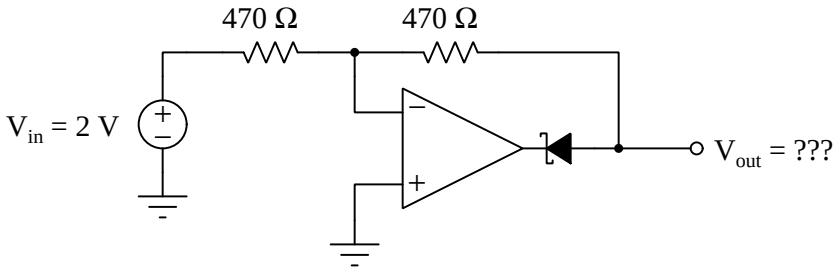
Determine the output voltage of this inverting amplifier circuit:



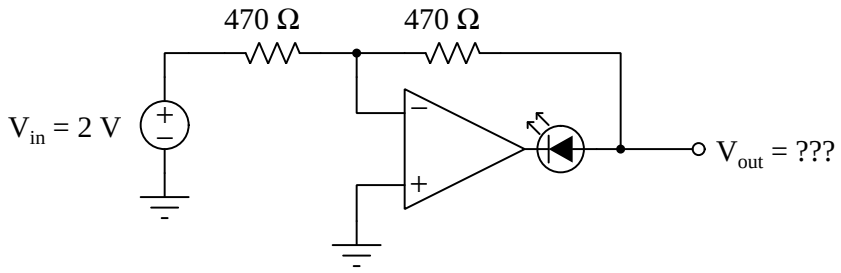
Determine the output voltage of this same circuit with an added diode, assuming the silicon diode has a typical 0.7 Volt forward drop:



Now, swap a Schottky diode (0.4 Volt typical forward drop) and re-calculate V_{out}:



Lastly, swap a light-emitting diode (1.7 Volts typical forward drop) and re-calculate V_{out}:



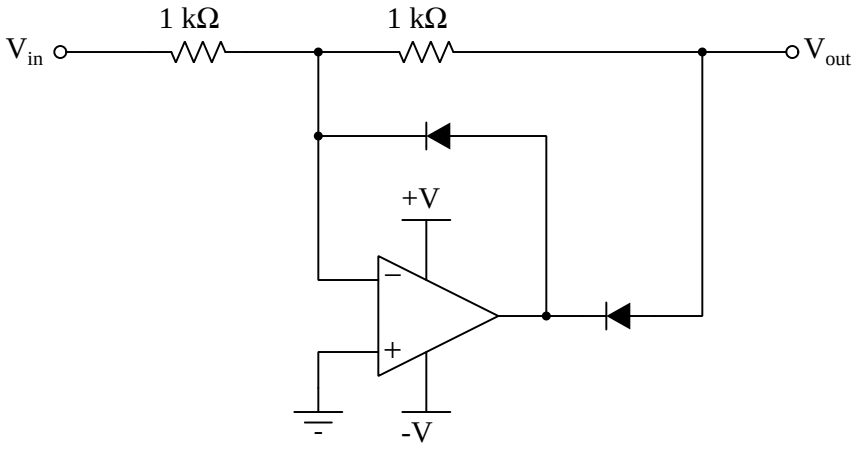
Comment on these three circuits' output voltages: what does this indicate about the effect of the diode's voltage drop on the opamp output?

Challenges

- What would happen if V_{in} were reversed in polarity, from +2 Volts to −2 Volts?

6.2.21 Precision rectifier analysis

This opamp circuit is called a *precision rectifier*. Analyze its output voltage as the input voltage smoothly increases from -5 volts to +5 volts, and explain why the circuit is worthy of its name:



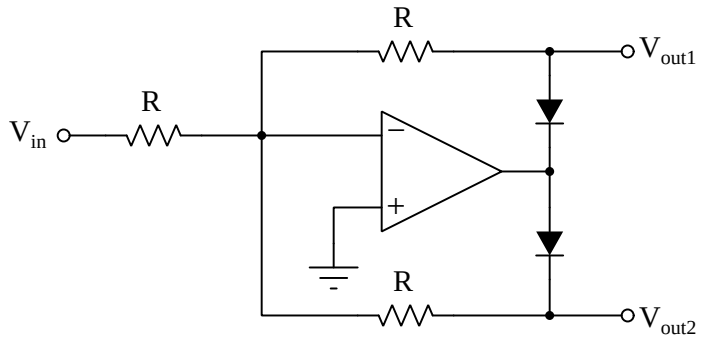
Assume that both diodes in this circuit are silicon switching diodes, with a nominal forward voltage drop of 0.7 Volts.

Challenges

- Would the behavior of this circuit change at all if the diodes were swapped out for others having a different forward voltage drop?

6.2.22 Polarity separator analysis

The following circuit is sometimes referred to as a *polarity separator*, constructed with all equal-value resistors. Invent some test conditions you would use to "prove" the operation of the circuit, then analyze the circuit under those imagined conditions and see what the results are:



Explain what each output does in this "polarity separator" circuit for any given input voltage.

Challenges

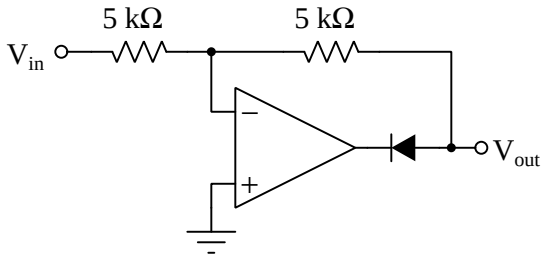
- How would this circuit behave if both diodes were reversed in direction?

6.2.23 Precision rectifier simulation program

Determine whether the following C-language computer program simulates the behavior of a *half-wave* precision rectifier circuit or a *full-wave* precision rectifier circuit, and also explain how the code works:

```
#include <stdio.h>

float rectify(float);

int main (void)
{
  float Vin;

  while(1)
  {
    printf("Enter input voltage in Volts: ");
    scanf("%f", &Vin);

    printf("Rectified output voltage = %f Volts\n\n", rectify(Vin));
  }

  return 0;
}

float rectify (float input)
{
  if (input > 0)
    return input;

  else if (input < 0)
    return -input;

  else
    return 0;
}
```

Next, modify the code to simulate the following rectifier circuit:

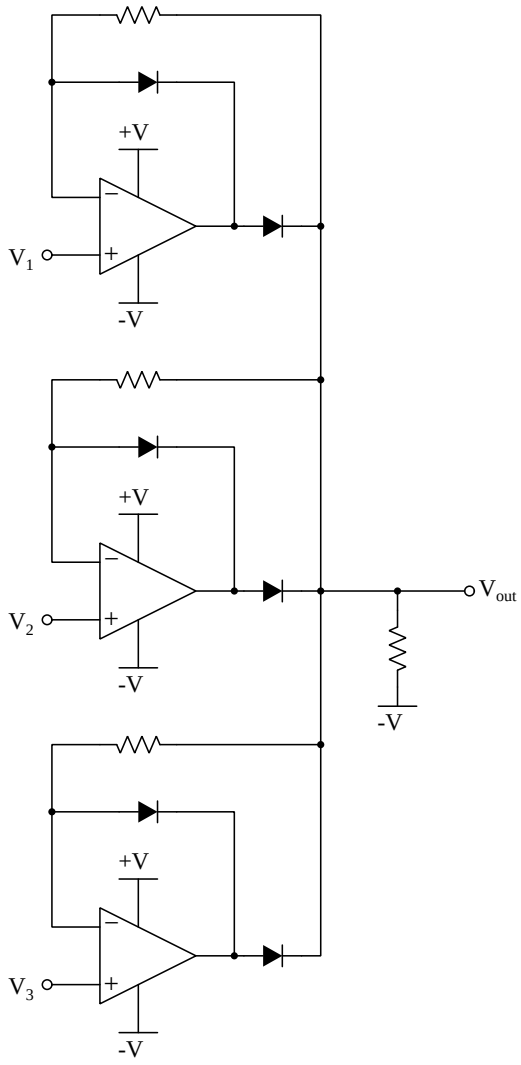


Challenges

- Where would a second diode be added to this rectifier circuit to avoid the saturation problem, and how else would the circuit's behavior be altered by this change?
- What would be the effect of making the two resistors unequal in resistance value?

6.2.24 Signal selector circuit

Determine the output voltage of this circuit for a scenario where $V_1 = +2$ Volts, $V_2 = -1.5$ Volts, $V_3 = +2.2$ Volts, and all resistors are equal in value:

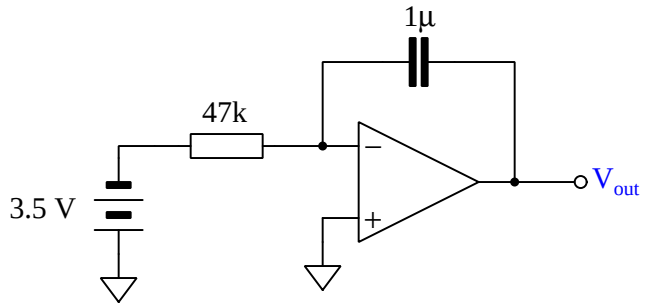


Challenges

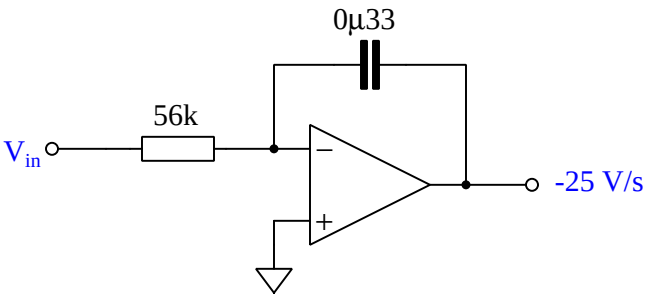
- Identify a practical application for this circuit.

6.2.25 Integrator calculations

Calculate the output voltage rate-of-change ($\frac{dV}{dt}$) for this active integrator circuit, being sure to explain all the steps involved in determining the answer:



Next, calculate the input voltage needed to produce an output voltage rate-of-change ($\frac{dV}{dt}$) of -25 Volts per second in this active integrator circuit:



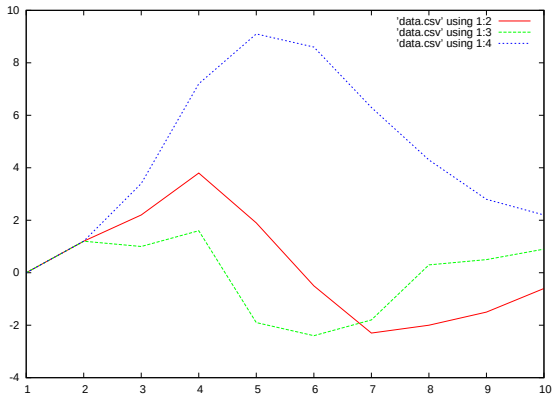
Challenges

- How could you modify either of these circuits to integrate at a less-aggressive rate for the same input voltage level?

6.2.26 Integrator and differentiator simulation program

Examine the following program written in C followed by a graph[11] of its output:

```
#include <stdio.h>

int main (void)
{
  int n;
  float data[10] = {0.0, 1.2, 2.2, 3.8, 1.9, -0.5, -2.3, -2.0, -1.5, -0.6};
  float out1[10] = {0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0};
  float out2[10] = {0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0};

  for (n = 1 ; n < 10 ; ++n)
    out1[n] = data[n] - data[n-1];

  for (n = 1 ; n < 10 ; ++n)
    out2[n] = data[n] + out2[n-1];

  for (n = 0 ; n < 10 ; ++n)
    printf("%i , %f , %f , %f\n", n+1, data[n], out1[n], out2[n]);

  return 0;
}
```



Which of these graphed functions represents the input data, the differentiated data, and the integrated data?

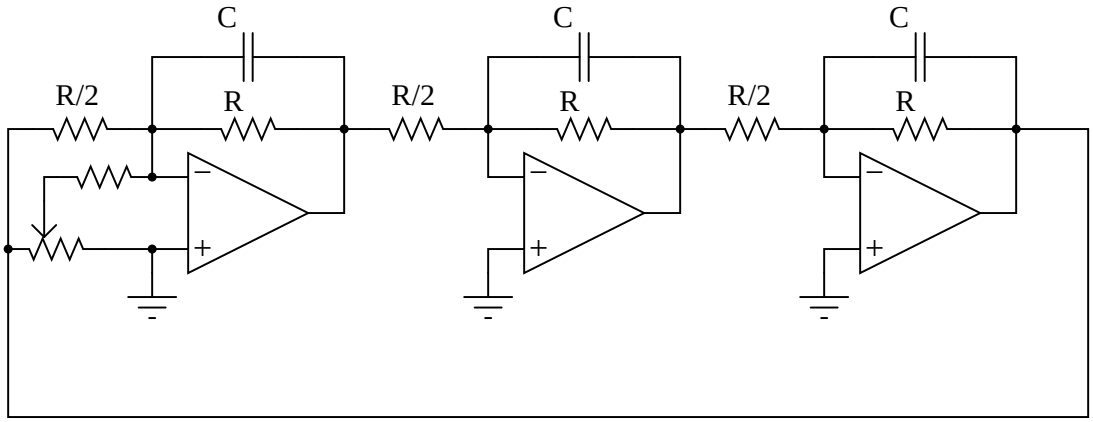
[11] The program itself simply outputs a comma-separated variable text listing. A spreadsheet or other mathematical visualization software is necessary to convert this text data listing into a two-dimensional graph.

Challenges

- The graphs as shown might not be the easiest to interpret, as far as determining which represents the derivative function and which the integral? Modify the input data set to make this determination easier.
- Why are no curly-braces necessary for any of the `for` loops or `if` conditionals in this program?
- Why do the first two `for` loops start with an `n` value of 1 rather than 0?

6.2.27 Three-phase oscillator

This interesting opamp circuit produces true sinusoidal voltage waveforms, three of them to be exact:



With all the resistors and capacitors, you might have guessed this to be a phase-shift type of oscillator circuit, and you would be correct. Here, each parallel RC network provides 60 degrees of lagging phase shift to combine with the 180 degrees of phase shift inherent to the inverting amplifier configurations, yielding 120 degrees of shift per opamp stage.

Derive a formula solving for the operating frequency of this oscillator circuit, knowing that the impedance of each parallel RC network will have a phase angle of -60^o. Also, determine where on this circuit you would obtain the three promised sine waves.

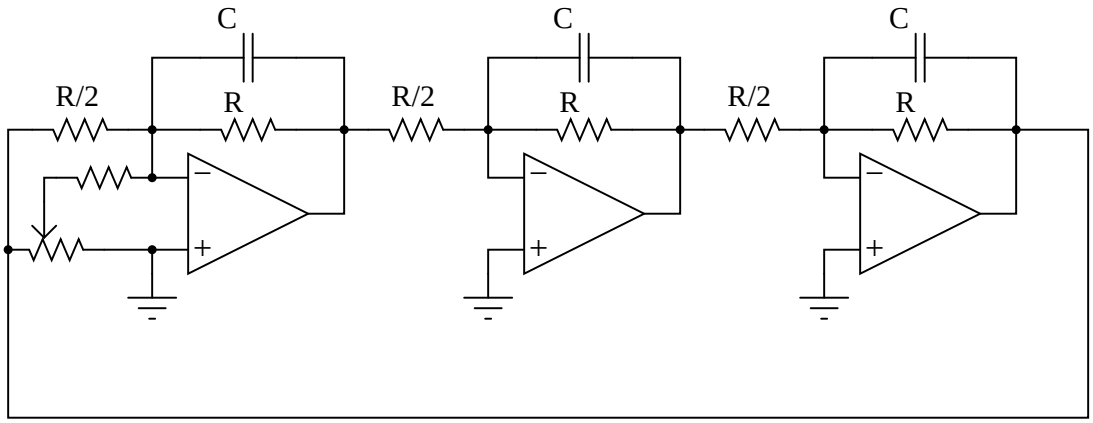
Now, determine the voltage gain of each opamp stage at the oscillating frequency.

Challenges

- Identify the purpose for the potentiometer.

6.2.28 Three-phase oscillator simulation program

This interesting opamp circuit produces true sinusoidal voltage waveforms, three of them to be exact:



With all the resistors and capacitors, you might have guessed this to be a phase-shift type of oscillator circuit, and you would be correct. Here, each parallel RC network provides 60 degrees of lagging phase shift to combine with the 180 degrees of phase shift inherent to the inverting amplifier configurations, yielding 120 degrees of shift per opamp stage.

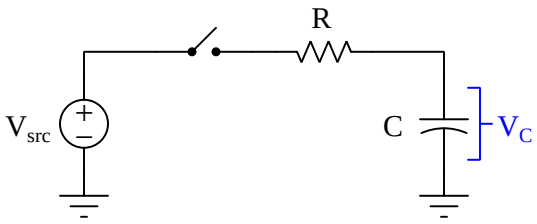
Write a program in a text-based language to simulate the parallel RC network to determine the proper amount of resistance to make the network exhibit a -60 degree impedance phase angle given a capacitor value and frequency value of your choice.

Challenges

- Identify the purpose for the potentiometer.

6.2.29 Capacitor voltage simulation program

When a capacitor is connected to a constant-voltage source through a series resistance, its voltage will gradually reach equilibrium with the constant-voltage source over time. The rate at which the capacitor's voltage approaches the constant voltage value depends on the resistor's resistance as well as the capacitor's capacitance:



As soon as the switch closes (at time = 0) the resistor's voltage drop will become equal to the difference between V_{src} and V_C. This potential difference results in a current (by Ohm's Law, $I = \frac{V}{R}$), and this current in turn determines how rapidly the capacitor's voltage will rise or fall over time ($I = C\frac{dV_c}{dt}$). Setting all this up in an equation:

$$\frac{V_{src} - V_C}{R} = C\frac{dV_C}{dt}$$

If we manipulate this equation to solve for the change in capacitor voltage (dV_C) for each interval of time (dt), we get this:

$$dV_C = \left(\frac{V_{src} - V_C}{RC}\right) dt$$

This type of equation, where a variable's value helps determine how rapidly that same variable's value changes, is known as a *differential equation*. In this particular case, the voltage for the capacitor (V_C) plays a role in determining how rapidly it will vary over time ($\frac{dV_C}{dt}$).

What this particular differential equation tells us is that over short intervals of time (dt), the capacitor's voltage will either increase or decrease by an amount equal to $\frac{V_{src}-V_C}{RC}$ over each time interval.

Differential equations are useful for describing many phenomena in the natural world, and techniques for solving them can become quite complex[12]. When electronics students begin their studies of RC networks without the benefit of calculus, they use inverse-exponential equations that are actually solutions to this differential equation. Having inverse-exponential equations ready to use on circuits like this saves us from having to know and apply calculus to solve the differential equation(s). However, if we know how to program a computer to perform a great many repeated calculations, it becomes possible to evaluate the differential equation in its raw form without having to know and apply calculus.

[12] Entire university-level mathematics courses exist to teach students how to solve differential equations, and in those courses you will also learn of many differential equations for which mathematicians have still not found formal solutions! However, all types of differential equations may be approximately solved by computer simulation as we are about to do here.

The following computer program written in the C language simulates the capacitor's voltage over a one-second span of time, assuming an initial capacitor voltage of 0 Volts, a source voltage value of 10 Volts, a 10 kΩ resistor, a 22 μF capacitor, and a computation time interval of 0.1 milliseconds:

```
#include <stdio.h>

int main (void)
{
  float t, dVc;
  float Vc = 0.0;
  float Vsrc = 10.0;
  float R = 10e3;
  float C = 22e-6;
  float dt = 0.1e-3;

  printf("Time (sec) , Capacitor voltage (Volts)\n");

  for (t = 0.0 ; t < 1.0 ; t = t + dt)
  {
    printf("%f , %f\n", t, Vc);

    dVc = (Vsrc - Vc) / (R * C) * dt;
    Vc = Vc + dVc;
  }

  return 0;
}
```

Answer the following questions about this simulation program:

- How could this simulation be modified for a longer span of time, perhaps 2 seconds instead of 1 second?
- What purpose is served by the calculation `Vc = Vc + dVc;`?
- Based on your understanding of RC circuits, how much voltage should the capacitor accumulate after the switch has been closed for exactly one time-constant's worth of time?
- Is the value of `dt` significant to the simulation? Why or why not?
- Will this program properly simulate capacitor voltage over time if the capacitor begins at a non-zero voltage value when the switch closes?

Challenges

- What is the justification for the `printf` occurring *before* the calculation of `dVc` and of `Vc` within the `for()` loop, rather than *after*?

6.3 Diagnostic reasoning

These questions are designed to stimulate your deductive and inductive thinking, where you must apply general principles to specific scenarios (deductive) and also derive conclusions about the failed circuit from specific details (inductive). In a Socratic discussion with your instructor, the goal is for these questions to reinforce your recall and use of general circuit principles and also challenge your ability to integrate multiple symptoms into a sensible explanation of what's wrong in a circuit. Your instructor may also pose additional questions based on those assigned, in order to further challenge and sharpen your diagnostic abilities.

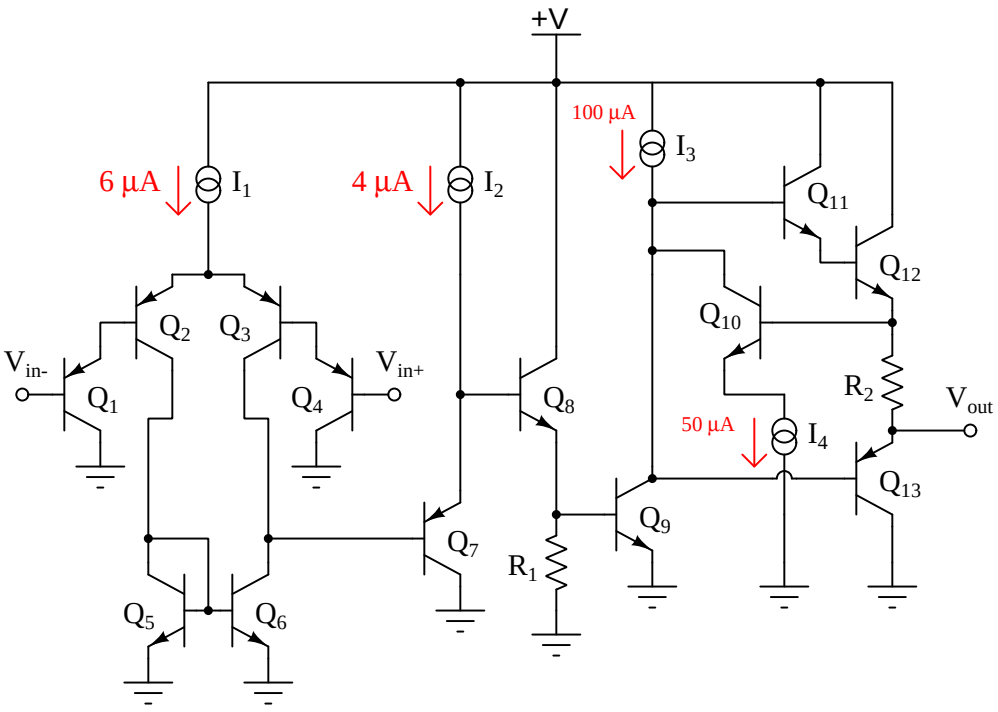
As always, your goal is to fully *explain* your analysis of each problem. Simply obtaining a correct answer is not good enough – you must also demonstrate sound reasoning in order to successfully complete the assignment. Your instructor's responsibility is to probe and challenge your understanding of the relevant principles and analytical processes in order to ensure you have a strong foundation upon which to build further understanding.

You will note a conspicuous lack of answers given for these diagnostic questions. Unlike standard textbooks where answers to every other question are given somewhere toward the back of the book, here in these learning modules students must rely on other means to check their work. The best way by far is to debate the answers with fellow students and also with the instructor during the Socratic dialogue sessions intended to be used with these learning modules. Reasoning through challenging questions with other people is an excellent tool for developing strong reasoning skills.

Another means of checking your diagnostic answers, where applicable, is to use circuit simulation software to explore the effects of faults placed in circuits. For example, if one of these diagnostic questions requires that you predict the effect of an open or a short in a circuit, you may check the validity of your work by simulating that same fault (substituting a very high resistance in place of that component for an open, and substituting a very low resistance for a short) within software and seeing if the results agree.

6.3.1 Internal faults in a model 324

Predict how the operation of this model 324 operational amplifier circuit will be affected as a result of the following faults. Specifically, identify whether the output voltage (V_{out}) will move in a positive direction (closer to the +V rail) or in a negative direction (closer to ground). Consider each fault independently (i.e. one at a time, no coincidental faults), and explain *why* the resulting effects will occur:



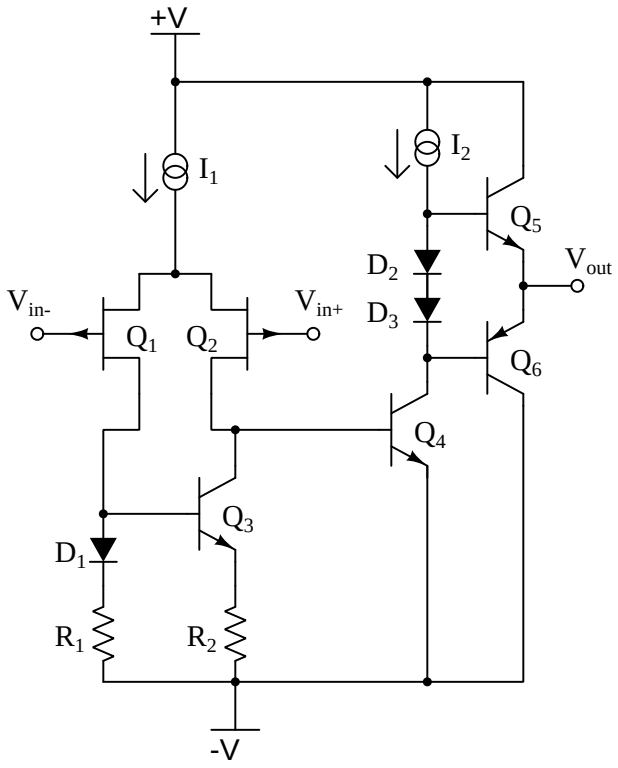
- Transistor Q_5 fails shorted (collector-to-emitter):
- Transistor Q_6 fails shorted (collector-to-emitter):
- Resistor R_1 fails open:
- Current source I_2 fails shorted:

Challenges

- Perform a "thought experiment" demonstrating the identities of the two inputs (i.e. inverting versus noninverting).

6.3.2 Internal faults in a model TL082

Predict how the operation of this operational amplifier circuit will be affected as a result of the following faults. Specifically, identify whether the output voltage (V_{out}) will move in a positive direction (closer to the +V rail) or in a negative direction (closer to the −V rail). Consider each fault independently (i.e. one at a time, no coincidental faults):



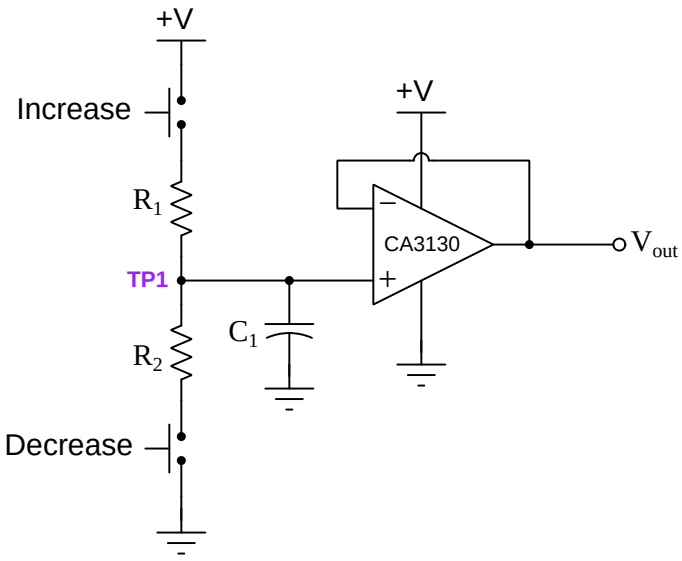
- Diode D_1 fails open:
- Resistor R_1 fails shorted:
- Transistor Q_2 fails shorted (drain-to-source):
- Transistor Q_5 fails shorted (collector-to-emitter):
- Resistor R_2 fails open:
- Current source I_2 fails open:

Challenges

- Perform a "thought experiment" demonstrating the identities of the two inputs (i.e. inverting versus noninverting).

6.3.3 Pushbutton-adjustable voltage

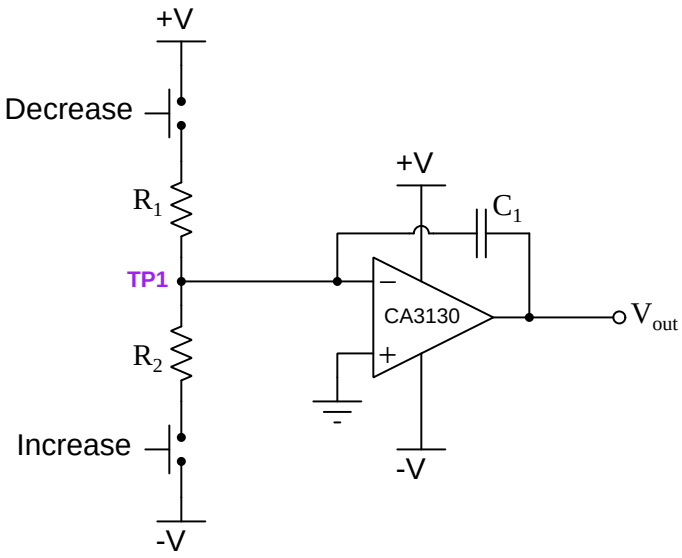
The purpose of this circuit is to provide a pushbutton-adjustable voltage. Pressing one button causes the output voltage to increase, while pressing the other button causes the output voltage to decrease. When neither button is pressed, the voltage remains stable:



After working just fine for quite a long while, the circuit suddenly fails: now it only outputs zero Volts DC all the time.

An experienced technician first checks the power supply voltage to see if it is within normal limits, and it is. Then, they measure voltage between test point TP1 and ground while pressing each of the pushbuttons one at a time. Explain why these are good diagnostic tests, and what the results of each test would tell the technician about the nature of the fault.

Next, consider this alternative design:



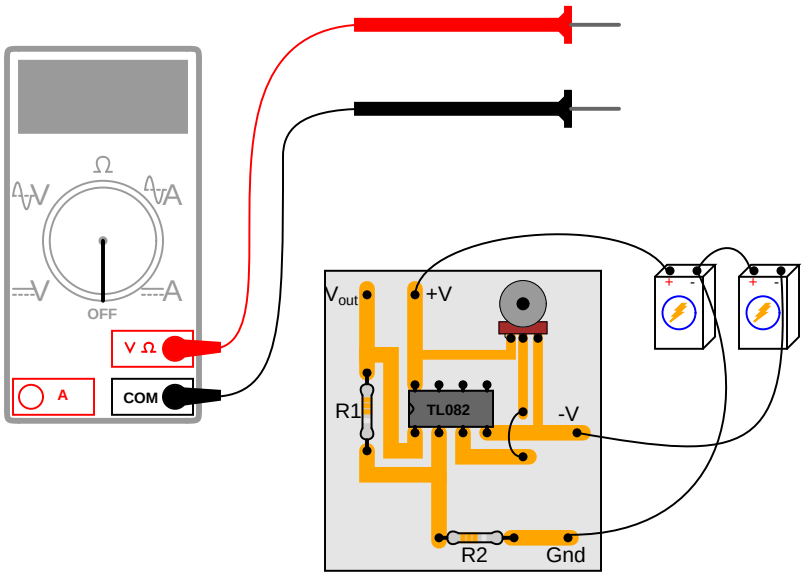
Explain how this circuit's behavior differs from the previous version, and then identify whether or not the same tests conducted on the previous circuit would be valid checks on this circuit given the same symptom (unchanging output voltage).

Challenges

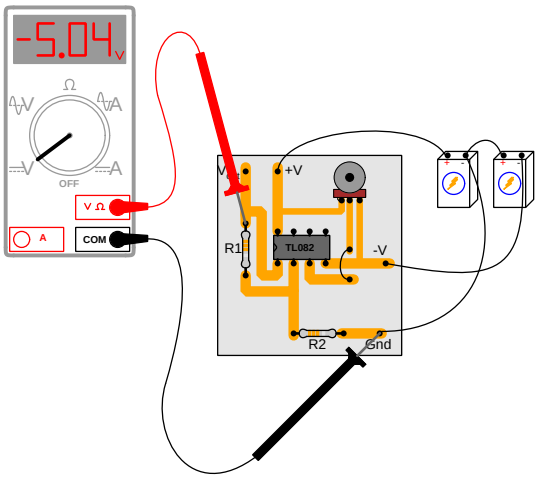
- Why do you suppose a model CA3130 operational amplifier is specified for this particular circuit? What is special about this opamp that qualifies it for the task?
- How would these circuits respond if both pushbuttons were pressed simultaneously?

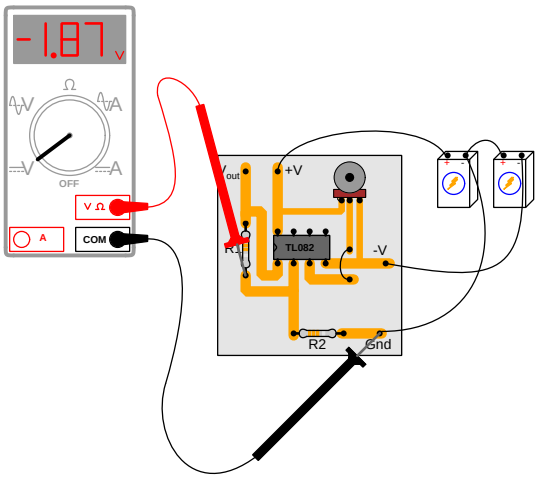
6.3.4 PCB measurements

Suppose a technician is checking the operation of the following electronic circuit:



She decides to measure the voltage on either side of resistor R_1 with reference to ground, and obtains these readings:





On the top side of R1, the voltage with reference to ground is -5.04 Volts. On the bottom side of R_1, the voltage with reference to ground is -1.87 Volts. The color code of resistor R_1 is Yellow, Violet, Orange, Gold. From this information, determine the following:

- Voltage across R_1 (between top to bottom):
- Polarity (+ and -) of voltage across R_1:
- Current (magnitude) through R_1:
- Direction of current through R_1:

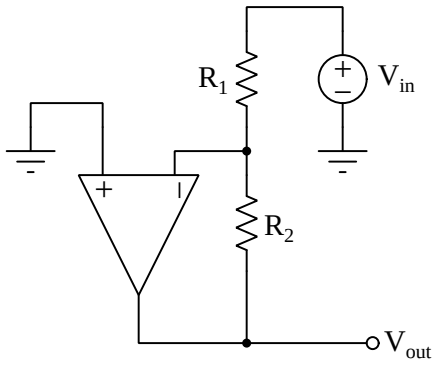
Additionally, explain how this technician would make each one of these determinations. What rules or laws of electric circuits would she apply?

Challenges

- Calculate the *range* of possible currents, given the specified tolerance of resistor R_1.
- Identify the voltage between pin 3 and ground.

6.3.5 Effects of faults on a simple amplifier circuit

Predict how the operation of this operational amplifier circuit will be affected as a result of the following faults. Consider each fault independently (i.e. one at a time, no coincidental faults):



- Resistor R_1 fails open:
- Solder bridge (short) across resistor R_1:
- Resistor R_2 fails open:
- Solder bridge (short) across resistor R_2:
- Broken wire between R_1/R_2 junction and inverting opamp input:

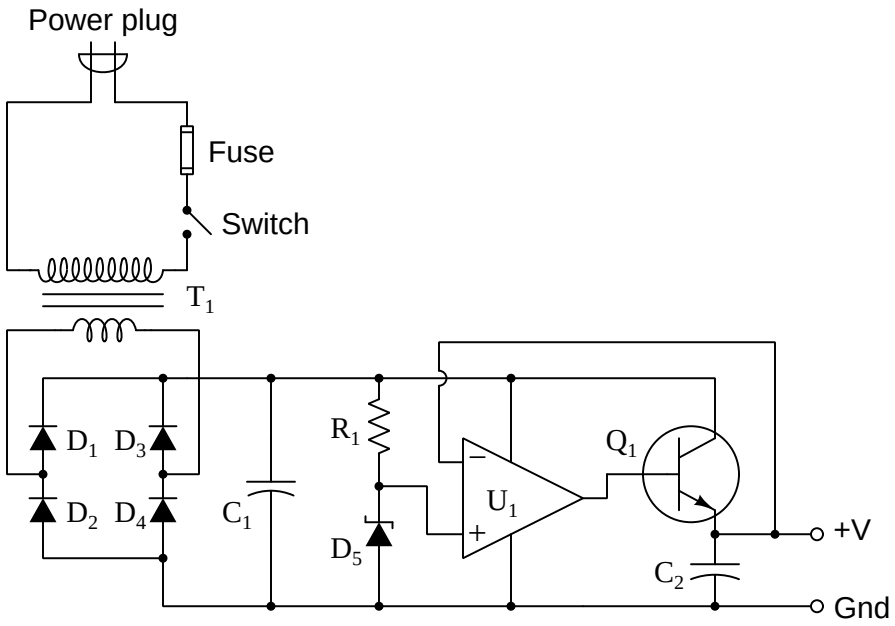
For each of these conditions, explain *why* the resulting effects will occur.

Challenges

- When an opamp's output *saturates*, what factor(s) limits the extent of its voltage swing?

6.3.6 Faults in a power supply voltage regulator

Predict how the operation of this regulated power supply circuit will be affected as a result of the following faults, and explain why for each case. Consider each fault independently (i.e. one at a time, no coincidental faults):



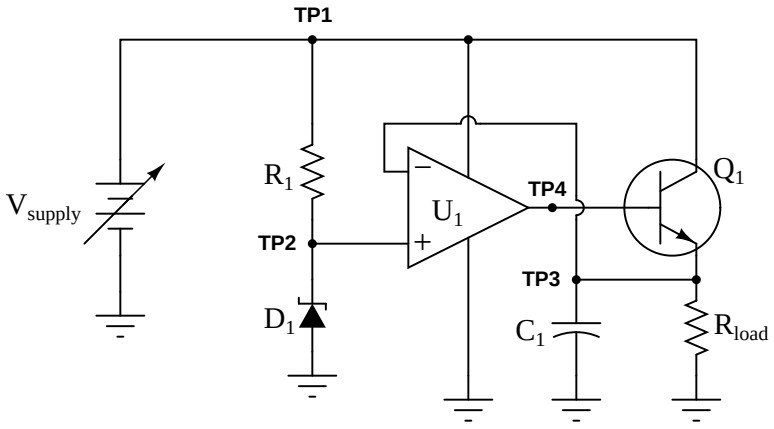
- Transformer T_1 primary winding fails open:
- Rectifying diode D_3 fails open:
- Rectifying diode D_4 fails shorted:
- Resistor R_1 fails open:
- Zener diode D_5 fails open:
- Operational amplifier U_1 fails with output saturated positive:
- Transistor Q_1 fails open (collector-to-emitter):

Challenges

- Identify how to modify this circuit for a higher regulated voltage.
- Identify what would need to be modified or replaced in this circuit to achieve a greater maximum load current rating.

6.3.7 Voltage regulator with zero output

This regulated power supply circuit has a problem. Instead of outputting 15 Volts DC (exactly) as it should, it is outputting 0 Volts DC to the load:



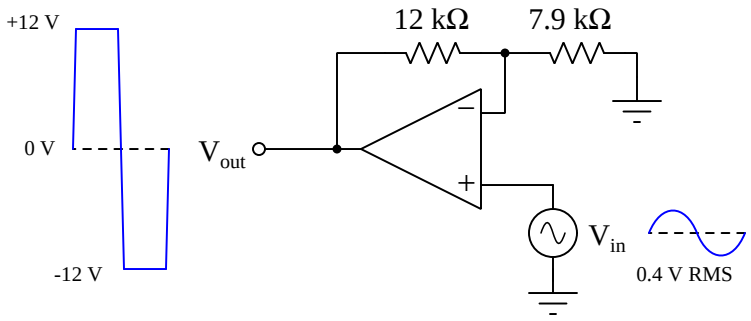
You measure 0.25 Volts DC between TP4 and ground, and 20 Volts between TP1 and ground, using your voltmeter. From this information, determine at least two independent faults that could cause this particular problem.

Challenges

- Modify this circuit to use a MOSFET rather than a BJT.

6.3.8 Clipped output waveform

There is something wrong with this amplifier circuit. Note the relative amplitudes of the input and output signals as measured by an oscilloscope:



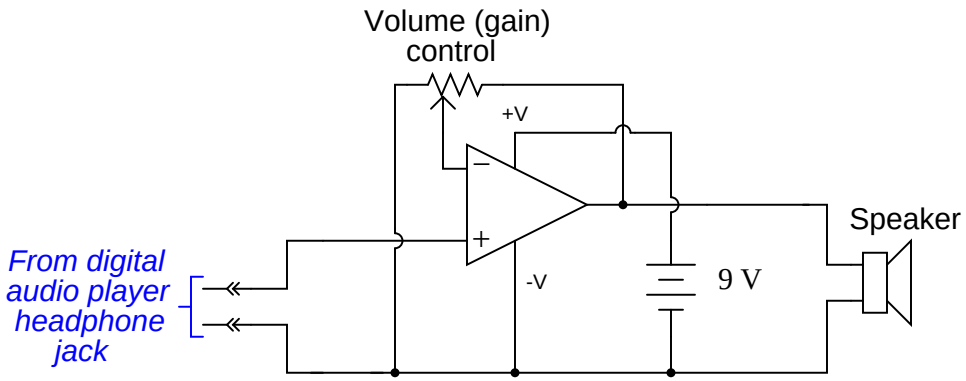
This circuit used to function perfectly, but then began to malfunction in this manner: producing a "clipped" output waveform of excessive amplitude. Determine the approximate amplitude that the output voltage waveform *should* be for the component values given in this circuit, and then identify possible causes of the problem and also elements of the circuit that you know cannot be at fault.

Challenges

- Suppose someone suggests to you that the power supply voltage was too great. Explain why this cannot be to blame.

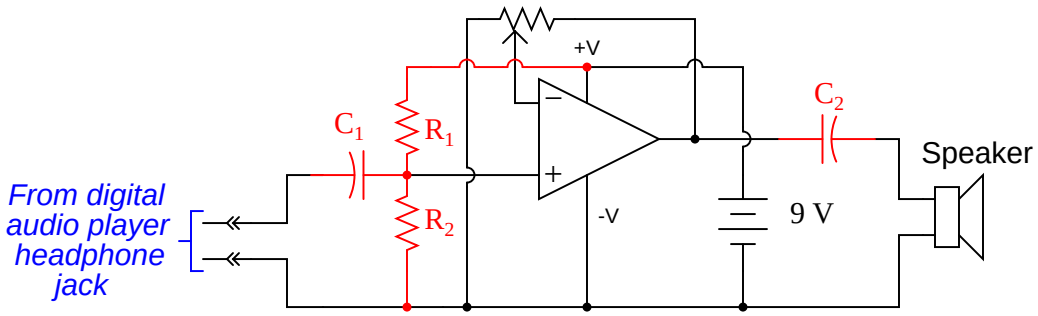
6.3.9 Audio amplifier design improvement

A student wishes to build a variable-gain audio amplifier circuit using an operational amplifier and a potentiometer, to drive a small loudspeaker so he can listen to the output of a digital audio player without having to use headphones:



Before building the project in a finalized form, the student prototypes it on a solderless breadboard to make sure it functions as intended. And it is a good thing he decided to do this before wasting time on a final version, because it sounds terrible!

When playing a song, the student can hear sound through the headphones, but it is terribly distorted. Taking the circuit to his instructor for help, the instructor suggests the following additions:



$C_1 = C_2 = 47\ \mu F$

$R_1 = R_2 = 33\ k\Omega$

After adding these components, the circuit works great. Now, music may be heard through the speaker with no noticeable distortion.

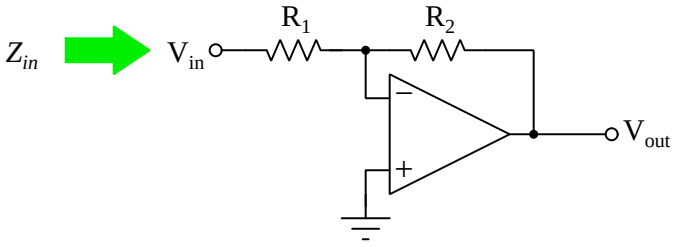
Explain what functions the extra components perform, and why the circuit did not work as originally built.

Challenges

- Another problem this student notes while prototyping his audio amplifier is that he can never turn the volume all the way off. Even at the lowest volume setting, you can still hear some sound from the loudspeaker. Identify the problem, and modify the circuit accordingly to fix it.

6.3.10 Effect of faults on amplifier input impedance

Predict how the input impedance (Z_{in}) of this inverting operational amplifier circuit will be affected as a result of the following faults. Consider each fault independently (i.e. one at a time, no multiple faults):



- Resistor R_1 fails open:
- Solder bridge (short) across resistor R_1:
- Resistor R_2 fails open:
- Solder bridge (short) across resistor R_2:
- Broken wire between R_1/R_2 junction and inverting opamp input:
- Operational amplifier loses power:

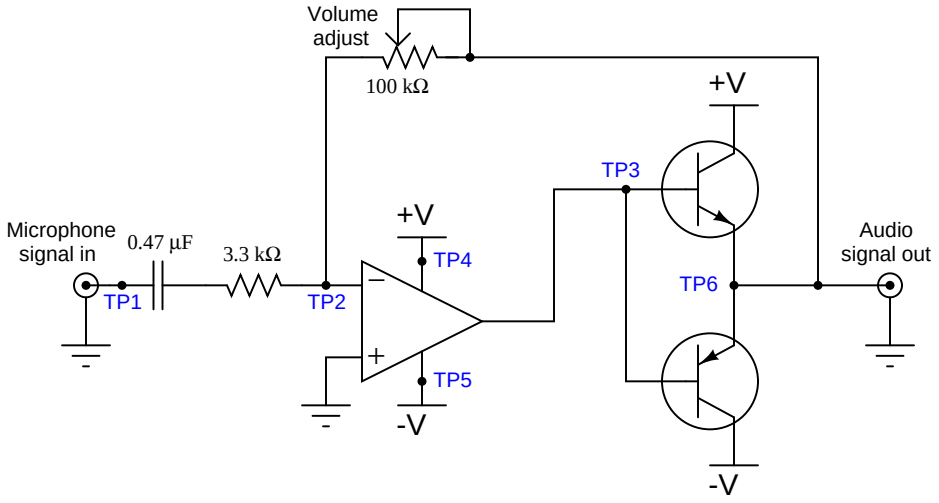
For each of these conditions, explain *why* the input impedance changes as it does.

Challenges

- Explain why the input impedance of an amplifier is a significant factor in circuit design.

6.3.11 Troubleshooting a microphone amplifier

There is something wrong with this amplifier circuit. Despite an audio signal of normal amplitude detected at test point 1 (TP1), there is no output measured at the "Audio signal out" jack:



Next, you decide to check for the presence of a good signal at test point 3 (TP3). There, you find 0 Volts AC and DC no matter where the volume control is set.

From this information, formulate a plan for troubleshooting this circuit, answering the following questions:

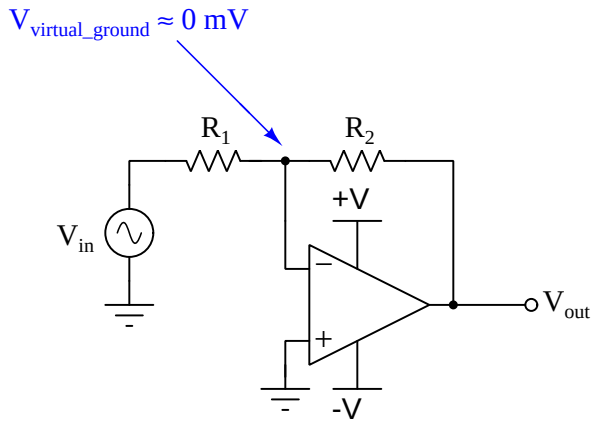
- What type of signal would you expect to measure at TP3?
- What would be your next step in troubleshooting this circuit?
- Are there any elements of this circuit you know to be working properly?
- What do you suppose would be the most *likely* failure, assuming this circuit once worked just fine and suddenly stopped working all on it's own?

Challenges

- Explain how and why the volume control is supposed to function.

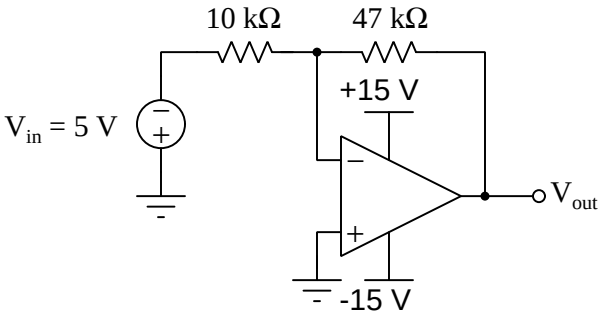
6.3.12 No longer virtual ground

The junction between the two resistors and the inverting input of the operational amplifier is often referred to as a *virtual ground*, the voltage between it and ground being (almost) zero over a wide range of circuit conditions:



If the operational amplifier is driven into saturation, though, the "virtual ground" will no longer be at ground potential. Explain why this is, and what condition(s) may cause this to happen.

Then, analyze all currents and voltage drops in the following circuit, assuming an opamp with the ability to swing its output voltage rail-to-rail:

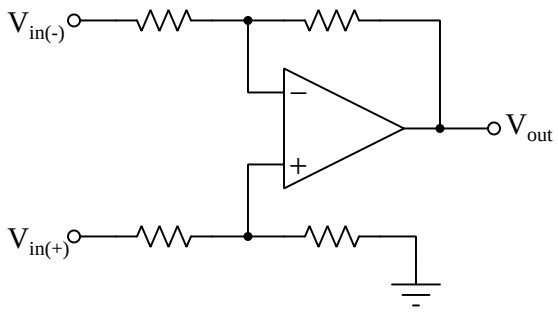


Challenges

- Explain what "saturation" means for an operational amplifier, and identify its cause(s).

6.3.13 Testing common-mode rejection

Devise a test for common-mode rejection in the following differential amplifier circuit:

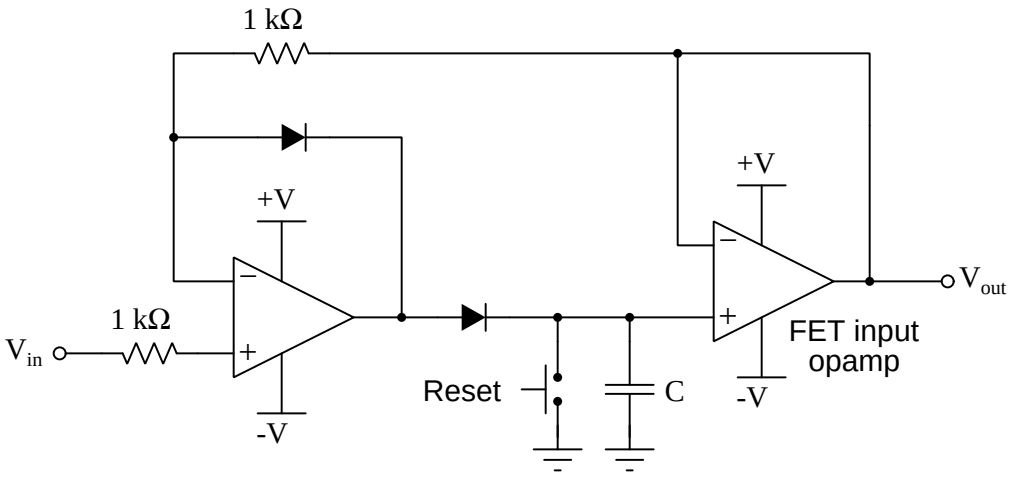


Challenges

- Explain why common-mode rejection is important in a differential amplifier.

6.3.14 Design flaws in a peak follower-and-hold circuit

Examine the following circuit and identify at least one mistake in it:

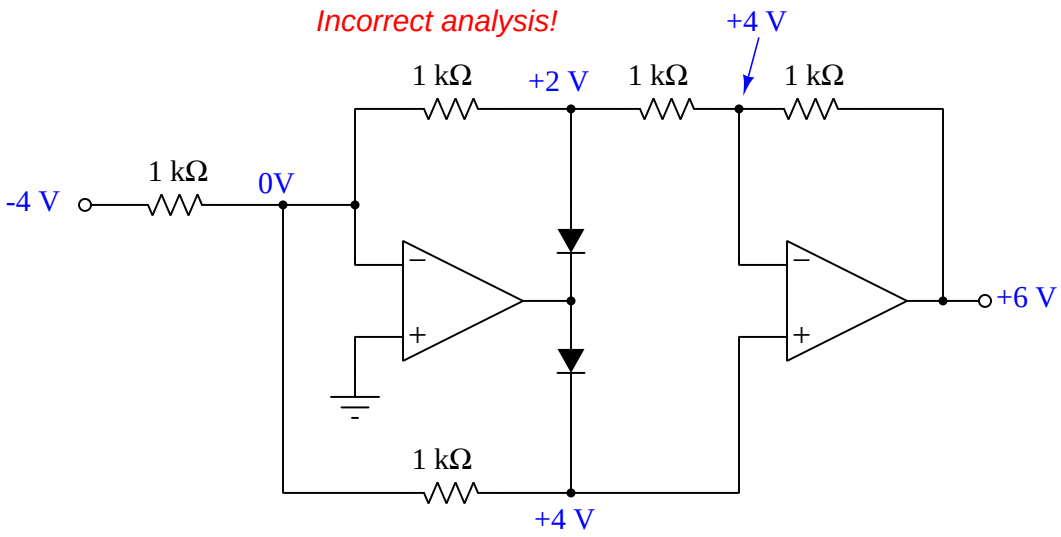


Challenges

- Identify a practical application for such a circuit.

6.3.15 Incorrect full-wave rectifier analysis

Suppose a student attempts to analyze the following full-wave precision rectifier circuit for an input signal voltage of −4 Volts, annotating ground-referenced voltages in the circuit as shown:



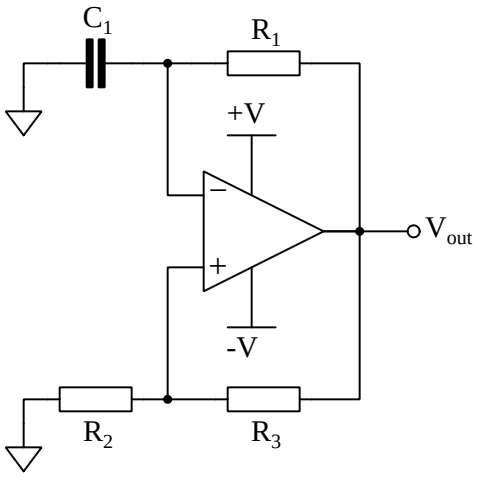
Explain why this analysis is incorrect, and then correct it.

Challenges

- Identify the consequence(s) of the input resistor failing open.

6.3.16 Effects of faults in a relaxation oscillator

Predict how the operation of this relaxation oscillator circuit will be affected as a result of the following faults. Consider each fault independently (i.e. one at a time, no multiple faults):



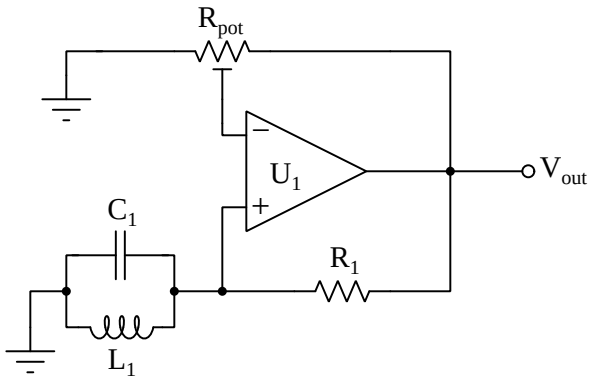
- Resistor R_1 fails open
- Solder bridge (short) across resistor R_1
- Capacitor C_1 fails shorted
- Solder bridge (short) across resistor R_2
- Resistor R_3 fails open

Challenges

- Identify at least two different ways to increase this oscillator's frequency.

6.3.17 Failed *LC* oscillator

Suppose this LC oscillator stopped working, and you suspected either the capacitor or the inductor as having failed. How could you check these two components *without* the use of an LCR meter?



Also, identify the purpose of the potentiometer.

Challenges

- Identify at least two different ways to increase this oscillator's frequency.
- What type of potentiometer is R_{pot}?

Appendix A

Problem-Solving Strategies

The ability to solve complex problems is arguably one of the most valuable skills one can possess, and this skill is particularly important in any science-based discipline.

- Study principles, not procedures. Don't be satisfied with merely knowing how to compute solutions – learn *why* those solutions work.
- Identify what it is you need to solve, identify all relevant data, identify all units of measurement, identify any general principles or formulae linking the given information to the solution, and then identify any "missing pieces" to a solution. Annotate all diagrams with this data.
- Sketch a diagram to help visualize the problem. When building a real system, always devise a plan for that system and analyze its function *before* constructing it.
- Follow the units of measurement and meaning of every calculation. If you are ever performing mathematical calculations as part of a problem-solving procedure, and you find yourself unable to apply each and every intermediate result to some aspect of the problem, it means you don't understand what you are doing. Properly done, every mathematical result should have practical meaning for the problem, and not just be an abstract number. You should be able to identify the proper units of measurement for each and every calculated result, and show where that result fits into the problem.
- Perform "thought experiments" to explore the effects of different conditions for theoretical problems. When troubleshooting real systems, perform *diagnostic tests* rather than visually inspecting for faults, the best diagnostic test being the one giving you the most information about the nature and/or location of the fault with the fewest steps.
- Simplify the problem until the solution becomes obvious, and then use that obvious case as a model to follow in solving the more complex version of the problem.
- Check for exceptions to see if your solution is incorrect or incomplete. A good solution will work for *all* known conditions and criteria. A good example of this is the process of testing scientific hypotheses: the task of a scientist is not to find support for a new idea, but rather to *challenge* that new idea to see if it holds up under a battery of tests. The philosophical

principle of *reductio ad absurdum* (i.e. disproving a general idea by finding a specific case where it fails) is useful here.

- Work "backward" from a hypothetical solution to a new set of given conditions.
- Add quantities to problems that are qualitative in nature, because sometimes a little math helps illuminate the scenario.
- Sketch graphs illustrating how variables relate to each other. These may be quantitative (i.e. with realistic number values) or qualitative (i.e. simply showing increases and decreases).
- Treat quantitative problems as qualitative in order to discern the relative magnitudes and/or directions of change of the relevant variables. For example, try determining what happens if a certain variable were to increase or decrease before attempting to precisely calculate quantities: how will each of the dependent variables respond, by increasing, decreasing, or remaining the same as before?
- Consider limiting cases. This works especially well for qualitative problems where you need to determine which direction a variable will change. Take the given condition and magnify that condition to an extreme degree as a way of simplifying the direction of the system's response.
- Check your work. This means regularly testing your conclusions to see if they make sense. This does *not* mean repeating the same steps originally used to obtain the conclusion(s), but rather to use some other means to check validity. Simply repeating procedures often leads to *repeating the same errors* if any were made, which is why alternative paths are better.

Appendix B

Instructional philosophy

B.1 First principles of learning

- **Anyone can learn anything** given appropriate time, effort, resources, challenges, encouragement, and expectations. Dedicating time and investing effort are the student's responsibility; providing resources, challenges, and encouragement are the teacher's responsibility; high expectations are a responsibility shared by both student and teacher.

- **Transfer is not automatic.** The human mind has a natural tendency to compartmentalize information, which means the process of taking knowledge learned in one context and applying it to another usually does not come easy and therefore should never be taken for granted.

- **Learning is iterative.** The human mind rarely learns anything perfectly on the first attempt. Anticipate mistakes and plan for multiple tries to achieve full understanding, using the lessons of those mistakes as feedback to guide future attempts.

- **Information is absorbed, but understanding is created.** Facts and procedures may be memorized easily enough by repeated exposure, but the ability to reliably apply principles to novel scenarios only comes through intense personal effort. This effort is fundamentally creative in nature: casting new concepts in one's own words, running mental experiments to perceive patterns and trends, etc. These acts of making knowledge "one's own" need not be perfect in order to be effective, as the value lies in the activity and not necessarily the finished product. Depth of understanding and permanence of memory are both directly related to mental effort during the learning process – i.e. *easy learning is shallow and fleeting.*

- **Education trumps training.** There is no such thing as an entirely isolated subject, as all fields of knowledge are connected. Training is narrowly-focused and task-oriented. Education is broad-based and principle-oriented. When preparing for a life-long technical career, education beats training every time.

- **Character matters.** Poor habits are more destructive than deficits of knowledge or skill. This is especially true in collective endeavors, where a team's ability to function depends on trust between its members. Simply put, no one wants an untrustworthy person on their team. An essential component of education then, is character development.

- **People learn to be responsible by bearing responsibility.** An irresponsible person is someone who has never *had* to be responsible for anything that mattered enough to them. Just as anyone can learn anything, anyone can become responsible if the personal cost of irresponsibility becomes high enough.

- **What gets measured, gets done.** Accurate and relevant assessment of learning is key to ensuring all students learn. Therefore, it is imperative to measure what matters.

- **Failure is nothing to fear.** Every human being fails, and fails in multiple ways at multiple times. Eventual success only happens when we don't stop trying.

B.2 Proven strategies for instructors

- Assume every student is capable of learning anything they desire given the proper conditions. Treat them as capable adults by granting real responsibility and avoiding artificial incentives such as merit or demerit points.
- Create a consistent culture of high expectations across the entire program of study. Demonstrate and encourage patience, persistence, and a healthy sense of self-skepticism. Anticipate and de-stigmatize error. Teach respect for the capabilities of others as well as respect for one's own fallibility.
- Replace lecture with "inverted" instruction, where students first encounter new concepts through reading and then spend class time in Socratic dialogue with the instructor exploring those concepts and solving problems individually. There is a world of difference between observing someone solve a problem versus actually solving a problem yourself, and so the point of this form of instruction is to place students in a position where they *cannot* passively observe.
- Require students to read extensively, write about what they learn, and dialogue with you and their peers to sharpen their understanding. Apply Francis Bacon's advice that "reading maketh a full man; conference a ready man; and writing an exact man". These are complementary activities helping students expand their confidence and abilities.
- Use artificial intelligence (AI) to challenge student understanding rather than merely provide information. Find productive ways for AI to critique students' clarity of thought and of expression, for example by employing AI as a Socratic-style interlocutor or as a reviewer of students' journals. Properly applied, AI has the ability to expand student access to critical review well outside the bounds of their instructor's reach.
- Build frequent and rapid feedback into the learning process so that students know at all times how well they are learning, to identify problems early and fix them before they grow. Model the intellectual habit of self-assessing and self-correcting your own understanding (i.e. a cognitive *feedback loop*), encouraging students to do the same.
- Use "mastery" as the standard for every assessment, which means the exam or experiment or project must be done with 100% competence in order to pass. Provide students with multiple opportunity for re-tries (different versions of the assessment every time).
- Require students to devise their own hypotheses and procedures on all experiments, so that the process is truly a scientific one. Have students assess their proposed experimental procedures for risk and devise mitigations for those risks. Let nothing be pre-designed about students' experiments other than a stated task (i.e. what principle the experiment shall test) at the start and a set of demonstrable knowledge and skill objectives at the end.
- Have students build as much of their lab equipment as possible: building power sources, building test assemblies[1], and building complete working systems (no kits!). In order to provide

[1] In the program I teach, every student builds their own "Development Board" consisting of a metal chassis with DIN rail, terminal blocks, and an AC-DC power supply of their own making which functions as a portable lab environment they can use at school as well as take home.

this same "ground-up" experience for every new student, this means either previous students take their creations with them, or the systems get disassembled in preparation for the new students, or the systems grow and evolve with each new student group.

- Incorporate external accountability for you and for your students, continuously improving the curriculum and your instructional methods based on proven results. Have students regularly network with active professionals through participation in advisory committee meetings, service projects, tours, jobshadows, internships, etc. Practical suggestions include requiring students to design and build projects for external clients (e.g. community groups, businesses, different departments within the institution), and also requiring students attend all technical advisory committee meetings and to dialogue with the industry representatives at those meetings.
- Repeatedly explore difficult-to-learn concepts across multiple courses, so that students have multiple opportunities to build their understanding.
- Relate all new concepts, whenever possible, to previous concepts and to relevant physical laws. Challenge each and every student, every day, to *reason* from concept to concept and to explain the logical connections between. Challenge students to verify their conclusions by multiple approaches (e.g. double-checking their work using different methods). Ask *"Why?"* often.
- Maintain detailed records on each student's performance and share these records privately with them. These records should include academic performance as well as professionally relevant behavioral tendencies.
- Hold mandatory "check-in" meetings between all program faculty and each new student during their first term. Offer these to all other students as an option, except for any students continuing to manifest unprofessional behaviors, poor academic performance, or who have some other need for a face-to-face meeting with faculty.
- Address problems while they are small, before they grow larger. This is equally true for tutoring technical concepts as it is for helping students build professional habits.
- Build rigorous quality control into the curriculum to ensure every student masters every important concept, and that the mastery is retained over time. This includes (1) review questions added to every exam to re-assess knowledge taught in previous terms, (2) cumulative exams at the end of every term to re-assess all important concepts back to the very beginning of the program, and (3) review assessments in practical (hands-on) coursework to ensure critically-important skills were indeed taught and are still retained. What you will find by doing this is that it actually boosts retention of students by ensuring that important knowledge gets taught and is retained over long spans of time. In the absence of such quality control, student learning and retention tends to be spotty and this contributes to drop-out and failure rates later in their education.
- Finally, *never rush learning.* Education is not a race. Give your students ample time to digest complex ideas, as you continually remind yourself of just how long it took you to achieve mastery! Long-term retention and the consistently correct application of concepts are always the result of *focused effort over long periods of time* which means there are no shortcuts to learning.

B.3 Proven strategies for students

The single most important piece of advice I have for any student of any subject is to take responsibility for your own development in all areas of life including mental development. Expecting others in your life to entirely guide your own development is a recipe for disappointment. This is just as true for students enrolled in formal learning institutions as it is for auto-didacts pursuing learning entirely on their own. Learning to think in new ways is key to being able to gainfully use information, to make informed decisions about your life, and to best serve those you care about. With this in mind, I offer the following advice to students:

- **Approach all learning as valuable**. No matter what course you take, no matter who you learn from, no matter the subject, there is something useful in every learning experience. If you don't see the value of every new experience, you are not looking closely enough!
- **Continually challenge yourself**. Let other people take shortcuts and find easy answers to easy problems. The purpose of education is to stretch your mind, in order to shape it into a more powerful tool. This doesn't come by taking the path of least resistance. An excellent analogy for an empowering education is productive physical exercise: becoming stronger, more flexible, and more persistent only comes through intense personal effort.
- **Master the use of language**. This includes reading extensively, writing every day, listening closely, and speaking articulately. To a great extent language channels and empowers thought, so the better you are at wielding language the better you will be at grasping abstract concepts and articulating them not only for your benefit but for others as well.
- **Do not limit yourself to the resources given to you**. Read books that are not on the reading list. Run experiments that aren't assigned to you. Form study groups outside of class. Take an entrepreneurial approach to your own education, as though it were a business you were building for your future benefit.
- **Express and share what you learn**. Take every opportunity to teach what you have learned to others, as this will not only help them but will also strengthen your own understanding[2].
- Realize that **no one can give you understanding**, just as no one can give you physical fitness. These both must be *built*.
- **Above all, recognize that learning is hard work, and that a certain level of frustration is unavoidable**. There are times when you will struggle to grasp some of these concepts, and that struggle is a natural thing. Take heart that it will yield with persistent and varied[3] effort, and never give up! That concepts don't immediately come to you is not a sign of something wrong, but rather of something right: that you have found a worthy challenge!

[2] On a personal note, I was surprised to learn just how much my own understanding of electronics and related subjects was strengthened by becoming a teacher. When you are tasked every day with helping other people grasp complex topics, it catalyzes your own learning by giving you powerful incentives to study, to articulate your thoughts, and to reflect deeply on the process of learning.

[3] As the old saying goes, "Insanity is trying the same thing over and over again, expecting different results." If you find yourself stumped by something in the text, you should attempt a different approach. Alter the thought experiment, change the mathematical parameters, do whatever you can to see the problem in a slightly different light, and then the solution will often present itself more readily.

B.4 Design of these learning modules

"The unexamined circuit is not worth energizing" – Socrates (if he had taught electricity)

These learning modules, although useful for self-study, were designed to be used in a formal learning environment where a subject-matter expert challenges students to digest the content and exercise their critical thinking abilities in the answering of questions and in the construction and testing of working circuits. Every effort has been made to embed the following instructional and assessment philosophies within:

- The first goal of education is to enhance clear and independent thought, in order that every student reach their fullest potential in a highly complex and inter-dependent world. Robust reasoning is *always* more important than particulars of any subject matter, because its application is universal.
- Literacy is fundamental to independent learning and thought because text continues to be the most efficient way to communicate complex ideas over space and time. Those who cannot read with ease are limited in their ability to acquire knowledge and perspective.
- Articulate communication is fundamental to work that is complex and interdisciplinary.
- Faulty assumptions and poor reasoning are best corrected through challenge, not presentation. The rhetorical technique of *reductio ad absurdum* (disproving an assertion by exposing an absurdity) works well to discipline student's minds, not only to correct the problem at hand but also to learn how to detect and correct future errors.
- Important principles should be repeatedly explored and widely applied throughout a course of study, not only to reinforce their importance and help ensure their mastery, but also to showcase the interconnectedness and utility of knowledge.

These learning modules were expressly designed to be used in an "inverted" teaching environment[4] where students first read the introductory and tutorial chapters on their own, then individually attempt to answer the questions and construct working circuits according to the experiment and project guidelines. The instructor never lectures, but instead meets regularly with each individual student to review their progress, answer questions, identify misconceptions, and challenge the student to new depths of understanding through further questioning. Regular meetings between instructor and student should resemble a Socratic[5] dialogue, where questions serve as scalpels to dissect topics and expose assumptions. The student passes each module only after consistently demonstrating their ability to logically analyze and correctly apply all major concepts in each question or project/experiment. The instructor must be vigilant in probing each student's understanding to ensure they are truly *reasoning* and not just *memorizing*. This is why "Challenge" points appear throughout, as prompts for students to think deeper about topics and as starting points for instructor queries. Sometimes these challenge points require additional knowledge that hasn't been covered in the series to answer in full. This is okay, as the major purpose of the Challenges is to stimulate analysis and synthesis on the part of each student.

The instructor must possess enough mastery of the subject matter and awareness of students' reasoning to generate their own follow-up questions to practically any student response. Even completely correct answers given by the student should be challenged by the instructor for the purpose of having students practice articulating their thoughts and defending their reasoning. Conceptual errors committed by the student should be exposed and corrected not by direct instruction, but rather by reducing the errors to an absurdity[6] through well-chosen questions and thought experiments posed by the instructor. Becoming proficient at this style of instruction requires time and dedication, but the positive effects on critical thinking for both student and instructor are spectacular.

An inspection of these learning modules reveals certain unique characteristics. One of these is a bias toward thorough explanations in the tutorial chapters. Without a live instructor to explain concepts and applications to students, the text itself must fulfill this role. This philosophy results in lengthier explanations than what you might typically find in a textbook, each step of the reasoning process fully explained, including footnotes addressing common questions and concerns students raise while learning these concepts. Each tutorial seeks to not only explain each major concept in sufficient detail, but also to explain the logic of each concept and how each may be developed

[4] In a traditional teaching environment, students first encounter new information via *lecture* from an expert, and then independently apply that information via *homework*. In an "inverted" course of study, students first encounter new information via *homework*, and then independently apply that information under the scrutiny of an expert. The expert's role in lecture is to simply *explain*, but the expert's role in an inverted session is to *challenge*, *critique*, and if necessary *explain* where gaps in understanding still exist.

[5] Socrates is a figure in ancient Greek philosophy famous for his unflinching style of questioning. Although he authored no texts, he appears as a character in Plato's many writings. The essence of Socratic philosophy is to leave no question unexamined and no point of view unchallenged. While purists may argue a topic such as electric circuits is too narrow for a true Socratic-style dialogue, I would argue that the essential thought processes involved with scientific reasoning on *any* topic are not far removed from the Socratic ideal, and that students of electricity and electronics would do very well to challenge assumptions, pose thought experiments, identify fallacies, and otherwise employ the arsenal of critical thinking skills modeled by Socrates.

[6] This rhetorical technique is known by the Latin phrase *reductio ad absurdum*. The concept is to expose errors by counter-example, since only one solid counter-example is necessary to disprove a universal claim. As an example of this, consider the common misconception among beginning students of electricity that voltage cannot exist without current. One way to apply *reductio ad absurdum* to this statement is to ask how much current passes through a fully-charged battery connected to nothing (i.e. a clear example of voltage existing without current).

from "first principles". Again, this reflects the goal of developing clear and independent thought in students' minds, by showing how clear and logical thought was used to forge each concept. Students benefit from witnessing a model of clear thinking in action, and these tutorials strive to be just that.

Another feature of these learning modules is that they do not treat topics in isolation. Rather, important concepts are introduced early in the series, and appear repeatedly as stepping-stones toward other concepts in subsequent modules. This helps to avoid the "compartmentalization" of knowledge, demonstrating the inter-connectedness of concepts and simultaneously reinforcing them. Each module is fairly complete in itself, reserving the beginning of its tutorial to a review of foundational concepts.

To high standards of education,

Tony R. Kuphaldt

Appendix C

Tools used

I am indebted to the developers of many open-source software applications in the creation of these learning modules. The following is a list of these applications with some commentary on each.

You will notice a theme common to many of these applications: a bias toward *code*. Although I am by no means an expert programmer in any computer language, I understand and appreciate the flexibility offered by code-based applications where the user (you) enters commands into a plain ASCII text file, which the software then reads and processes to create the final output. Code-based computer applications are by their very nature *extensible*, while WYSIWYG (What You See Is What You Get) applications are generally limited to whatever user interface the developer makes for you.

The `GNU/Linux` computer operating system

> There is so much to be said about Linus Torvalds' `Linux` and Richard Stallman's `GNU` project. First, to credit just these two individuals is to fail to do justice to the *mob* of passionate volunteers who contributed to make this amazing software a reality. I first learned of `Linux` back in 1996, and have been using this operating system on my personal computers almost exclusively since then. It is *free*, it is completely *configurable*, and it permits the continued use of highly efficient `Unix` applications and scripting languages (e.g. shell scripts, Makefiles, `sed`, `awk`) developed over many decades. `Linux` not only provided me with a powerful computing platform, but its open design served to inspire my life's work of creating open-source educational resources.

Bram Moolenaar's `Vim` text editor

> Writing code for any code-based computer application requires a *text editor*, which may be thought of as a word processor strictly limited to outputting plain-ASCII text files. Many good text editors exist, and one's choice of text editor seems to be a deeply personal matter within the programming world. I prefer `Vim` because it operates very similarly to `vi` which is ubiquitous on `Unix`/`Linux` operating systems, and because it may be entirely operated via keyboard (i.e. no mouse required) which makes it fast to use.

Donald Knuth's TEX typesetting system

Developed in the late 1970's and early 1980's by computer scientist extraordinaire Donald Knuth to typeset his multi-volume magnum opus *The Art of Computer Programming*, this software allows the production of formatted text for screen-viewing or paper printing, all by writing plain-text code to describe how the formatted text is supposed to appear. TEX is not just a markup language for documents, but it is also a Turing-complete programming language in and of itself, allowing useful algorithms to be created to control the production of documents. Simply put, TEX *is a programmer's approach to word processing*. Since TEX is controlled by code written in a plain-text file, this means anyone may read that plain-text file to see exactly how the document was created. This openness afforded by the code-based nature of TEX makes it relatively easy to learn how other people have created their own TEX documents. By contrast, examining a beautiful document created in a conventional WYSIWYG word processor such as Microsoft `Word` suggests nothing to the reader about *how* that document was created, or what the user might do to create something similar. As Mr. Knuth himself once quipped, conventional word processing applications should be called WYSIAYG (What You See Is *All* You Get).

Leslie Lamport's LATEX extensions to TEX

Like all true programming languages, TEX is inherently extensible. So, years after the release of TEX to the public, Leslie Lamport decided to create a massive extension allowing easier compilation of book-length documents. The result was LATEX, which is the markup language used to create all ModEL module documents. You could say that TEX is to LATEX as `C` is to `C++`. This means it is permissible to use any and all TEX commands within LATEX source code, and it all still works. Some of the features offered by LATEX that would be challenging to implement in TEX include automatic index and table-of-content creation.

Tim Edwards' `Xcircuit` drafting program

This wonderful program is what I use to create all the schematic diagrams and illustrations (but not photographic images or mathematical plots) throughout the ModEL project. It natively outputs PostScript format which is a true vector graphic format (this is why the images do not pixellate when you zoom in for a closer view), and it is so simple to use that I have never had to read the manual! Object libraries are easy to create for `Xcircuit`, being plain-text files using PostScript programming conventions. Over the years I have collected a large set of object libraries useful for drawing electrical and electronic schematics, pictorial diagrams, and other technical illustrations.

`Gimp` graphic image manipulation program

Essentially an open-source clone of Adobe's `PhotoShop`, I use `Gimp` to resize, crop, and convert file formats for all of the photographic images appearing in the ModEL modules. Although `Gimp` does offer its own scripting language (called `Script-Fu`), I have never had occasion to use it. Thus, my utilization of `Gimp` to merely crop, resize, and convert graphic images is akin to using a sword to slice bread.

`SPICE` circuit simulation program

`SPICE` is to circuit analysis as TEX is to document creation: it is a form of markup language designed to describe a certain object to be processed in plain-ASCII text. When the plain-text "source file" is compiled by the software, it outputs the final result. More modern circuit analysis tools certainly exist, but I prefer `SPICE` for the following reasons: it is *free*, it is *fast*, it is *reliable*, and it is a fantastic tool for *teaching* students of electricity and electronics how to write simple code. I happen to use rather old versions of `SPICE`, version 2g6 being my "go to" application when I only require text-based output. `NGSPICE` (version 26), which is based on Berkeley `SPICE` version 3f5, is used when I require graphical output for such things as time-domain waveforms and Bode plots. In all `SPICE` example netlists I strive to use coding conventions compatible with all `SPICE` versions.

Andrew D. Hwang's `ePiX` mathematical visualization programming library

This amazing project is a `C++` library you may link to any `C/C++` code for the purpose of generating PostScript graphic images of mathematical functions. As a completely free and open-source project, it does all the plotting I would otherwise use a Computer Algebra System (CAS) such as `Mathematica` or `Maple` to do. It should be said that `ePiX` is *not* a Computer Algebra System like `Mathematica` or `Maple`, but merely a mathematical *visualization* tool. In other words, it won't determine integrals for you (you'll have to implement that in your own `C/C++` code!), but it can graph the results, and it does so beautifully. What I really admire about `ePiX` is that it is a `C++` programming library, which means it builds on the existing power and toolset available with that programming language. Mr. Hwang could have probably developed his own stand-alone application for mathematical plotting, but by creating a `C++` library to do the same thing he accomplished something much greater.

`gnuplot` mathematical visualization software

Another open-source tool for mathematical visualization is `gnuplot`. Interestingly, this tool is *not* part of Richard Stallman's GNU project, its name being a coincidence. For this reason the authors prefer "gnu" *not* be capitalized at all to avoid confusion. This is a much "lighter-weight" alternative to a spreadsheet for plotting tabular data, and the fact that it easily outputs directly to an X11 console or a file in a number of different graphical formats (including PostScript) is very helpful. I typically set my `gnuplot` output format to default (X11 on my Linux PC) for quick viewing while I'm developing a visualization, then switch to PostScript file export once the visual is ready to include in the document(s) I'm writing. As with my use of `Gimp` to do rudimentary image editing, my use of `gnuplot` only scratches the surface of its capabilities, but the important points are that it's *free* and that it *works well*.

`Python` programming language

Both Python and C++ find extensive use in these modules as instructional aids and exercises, but I'm listing Python here as a *tool* for myself because I use it almost daily as a *calculator*. If you open a Python interpreter console and type `from math import *` you can type mathematical expressions and have it return results just as you would on a hand calculator. Complex-number (i.e. *phasor*) arithmetic is similarly supported if you include the complex-math library (`from cmath import *`). Examples of this are shown in the Programming References chapter (if included) in each module. Of course, being a fully-featured programming language, Python also supports conditionals, loops, and other structures useful for calculation of quantities. Also, running in a console environment where all entries and returned values show as text in a chronologically-ordered list makes it easy to copy-and-paste those calculations to document exactly how they were performed.

Appendix D

Creative Commons License

Creative Commons Attribution 4.0 International Public License

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution 4.0 International Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits the Licensor receives from making the Licensed Material available under these terms and conditions.

Section 1 – Definitions.

a. **Adapted Material** means material subject to Copyright and Similar Rights that is derived from or based upon the Licensed Material and in which the Licensed Material is translated, altered, arranged, transformed, or otherwise modified in a manner requiring permission under the Copyright and Similar Rights held by the Licensor. For purposes of this Public License, where the Licensed Material is a musical work, performance, or sound recording, Adapted Material is always produced where the Licensed Material is synched in timed relation with a moving image.

b. **Adapter's License** means the license You apply to Your Copyright and Similar Rights in Your contributions to Adapted Material in accordance with the terms and conditions of this Public License.

c. **Copyright and Similar Rights** means copyright and/or similar rights closely related to copyright including, without limitation, performance, broadcast, sound recording, and Sui Generis Database Rights, without regard to how the rights are labeled or categorized. For purposes of this Public License, the rights specified in Section 2(b)(1)-(2) are not Copyright and Similar Rights.

d. **Effective Technological Measures** means those measures that, in the absence of proper authority, may not be circumvented under laws fulfilling obligations under Article 11 of the WIPO Copyright Treaty adopted on December 20, 1996, and/or similar international agreements.

e. **Exceptions and Limitations** means fair use, fair dealing, and/or any other exception or

limitation to Copyright and Similar Rights that applies to Your use of the Licensed Material.

f. **Licensed Material** means the artistic or literary work, database, or other material to which the Licensor applied this Public License.

g. **Licensed Rights** means the rights granted to You subject to the terms and conditions of this Public License, which are limited to all Copyright and Similar Rights that apply to Your use of the Licensed Material and that the Licensor has authority to license.

h. **Licensor** means the individual(s) or entity(ies) granting rights under this Public License.

i. **Share** means to provide material to the public by any means or process that requires permission under the Licensed Rights, such as reproduction, public display, public performance, distribution, dissemination, communication, or importation, and to make material available to the public including in ways that members of the public may access the material from a place and at a time individually chosen by them.

j. **Sui Generis Database Rights** means rights other than copyright resulting from Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, as amended and/or succeeded, as well as other essentially equivalent rights anywhere in the world.

k. **You** means the individual or entity exercising the Licensed Rights under this Public License. **Your** has a corresponding meaning.

Section 2 – Scope.

a. License grant.

1. Subject to the terms and conditions of this Public License, the Licensor hereby grants You a worldwide, royalty-free, non-sublicensable, non-exclusive, irrevocable license to exercise the Licensed Rights in the Licensed Material to:

A. reproduce and Share the Licensed Material, in whole or in part; and

B. produce, reproduce, and Share Adapted Material.

2. Exceptions and Limitations. For the avoidance of doubt, where Exceptions and Limitations apply to Your use, this Public License does not apply, and You do not need to comply with its terms and conditions.

3. Term. The term of this Public License is specified in Section 6(a).

4. Media and formats; technical modifications allowed. The Licensor authorizes You to exercise the Licensed Rights in all media and formats whether now known or hereafter created, and to make technical modifications necessary to do so. The Licensor waives and/or agrees not to assert any right or authority to forbid You from making technical modifications necessary to exercise the Licensed Rights, including technical modifications necessary to circumvent Effective Technological Measures.

For purposes of this Public License, simply making modifications authorized by this Section 2(a)(4) never produces Adapted Material.

5. Downstream recipients.

A. Offer from the Licensor – Licensed Material. Every recipient of the Licensed Material automatically receives an offer from the Licensor to exercise the Licensed Rights under the terms and conditions of this Public License.

B. No downstream restrictions. You may not offer or impose any additional or different terms or conditions on, or apply any Effective Technological Measures to, the Licensed Material if doing so restricts exercise of the Licensed Rights by any recipient of the Licensed Material.

6. No endorsement. Nothing in this Public License constitutes or may be construed as permission to assert or imply that You are, or that Your use of the Licensed Material is, connected with, or sponsored, endorsed, or granted official status by, the Licensor or others designated to receive attribution as provided in Section 3(a)(1)(A)(i).

b. Other rights.

1. Moral rights, such as the right of integrity, are not licensed under this Public License, nor are publicity, privacy, and/or other similar personality rights; however, to the extent possible, the Licensor waives and/or agrees not to assert any such rights held by the Licensor to the limited extent necessary to allow You to exercise the Licensed Rights, but not otherwise.

2. Patent and trademark rights are not licensed under this Public License.

3. To the extent possible, the Licensor waives any right to collect royalties from You for the exercise of the Licensed Rights, whether directly or through a collecting society under any voluntary or waivable statutory or compulsory licensing scheme. In all other cases the Licensor expressly reserves any right to collect such royalties.

Section 3 – License Conditions.

Your exercise of the Licensed Rights is expressly made subject to the following conditions.

a. Attribution.

1. If You Share the Licensed Material (including in modified form), You must:

A. retain the following if it is supplied by the Licensor with the Licensed Material:

i. identification of the creator(s) of the Licensed Material and any others designated to receive attribution, in any reasonable manner requested by the Licensor (including by pseudonym if designated);

ii. a copyright notice;

iii. a notice that refers to this Public License;

iv. a notice that refers to the disclaimer of warranties;

v. a URI or hyperlink to the Licensed Material to the extent reasonably practicable;

B. indicate if You modified the Licensed Material and retain an indication of any previous modifications; and

C. indicate the Licensed Material is licensed under this Public License, and include the text of, or the URI or hyperlink to, this Public License.

2. You may satisfy the conditions in Section 3(a)(1) in any reasonable manner based on the medium, means, and context in which You Share the Licensed Material. For example, it may be reasonable to satisfy the conditions by providing a URI or hyperlink to a resource that includes the required information.

3. If requested by the Licensor, You must remove any of the information required by Section 3(a)(1)(A) to the extent reasonably practicable.

4. If You Share Adapted Material You produce, the Adapter's License You apply must not prevent recipients of the Adapted Material from complying with this Public License.

Section 4 – Sui Generis Database Rights.

Where the Licensed Rights include Sui Generis Database Rights that apply to Your use of the Licensed Material:

a. for the avoidance of doubt, Section 2(a)(1) grants You the right to extract, reuse, reproduce, and Share all or a substantial portion of the contents of the database;

b. if You include all or a substantial portion of the database contents in a database in which You have Sui Generis Database Rights, then the database in which You have Sui Generis Database Rights (but not its individual contents) is Adapted Material; and

c. You must comply with the conditions in Section 3(a) if You Share all or a substantial portion of the contents of the database.

For the avoidance of doubt, this Section 4 supplements and does not replace Your obligations under this Public License where the Licensed Rights include other Copyright and Similar Rights.

Section 5 – Disclaimer of Warranties and Limitation of Liability.

a. Unless otherwise separately undertaken by the Licensor, to the extent possible, the Licensor offers the Licensed Material as-is and as-available, and makes no representations or warranties of any kind concerning the Licensed Material, whether express, implied, statutory, or other. This includes, without limitation, warranties of title, merchantability, fitness for a particular purpose, non-infringement, absence of latent or other defects, accuracy, or the presence or absence of errors,

whether or not known or discoverable. Where disclaimers of warranties are not allowed in full or in part, this disclaimer may not apply to You.

b. To the extent possible, in no event will the Licensor be liable to You on any legal theory (including, without limitation, negligence) or otherwise for any direct, special, indirect, incidental, consequential, punitive, exemplary, or other losses, costs, expenses, or damages arising out of this Public License or use of the Licensed Material, even if the Licensor has been advised of the possibility of such losses, costs, expenses, or damages. Where a limitation of liability is not allowed in full or in part, this limitation may not apply to You.

c. The disclaimer of warranties and limitation of liability provided above shall be interpreted in a manner that, to the extent possible, most closely approximates an absolute disclaimer and waiver of all liability.

Section 6 – Term and Termination.

a. This Public License applies for the term of the Copyright and Similar Rights licensed here. However, if You fail to comply with this Public License, then Your rights under this Public License terminate automatically.

b. Where Your right to use the Licensed Material has terminated under Section 6(a), it reinstates:

1. automatically as of the date the violation is cured, provided it is cured within 30 days of Your discovery of the violation; or

2. upon express reinstatement by the Licensor.

For the avoidance of doubt, this Section 6(b) does not affect any right the Licensor may have to seek remedies for Your violations of this Public License.

c. For the avoidance of doubt, the Licensor may also offer the Licensed Material under separate terms or conditions or stop distributing the Licensed Material at any time; however, doing so will not terminate this Public License.

d. Sections 1, 5, 6, 7, and 8 survive termination of this Public License.

Section 7 – Other Terms and Conditions.

a. The Licensor shall not be bound by any additional or different terms or conditions communicated by You unless expressly agreed.

b. Any arrangements, understandings, or agreements regarding the Licensed Material not stated herein are separate from and independent of the terms and conditions of this Public License.

Section 8 – Interpretation.

a. For the avoidance of doubt, this Public License does not, and shall not be interpreted to, reduce, limit, restrict, or impose conditions on any use of the Licensed Material that could lawfully

be made without permission under this Public License.

b. To the extent possible, if any provision of this Public License is deemed unenforceable, it shall be automatically reformed to the minimum extent necessary to make it enforceable. If the provision cannot be reformed, it shall be severed from this Public License without affecting the enforceability of the remaining terms and conditions.

c. No term or condition of this Public License will be waived and no failure to comply consented to unless expressly agreed to by the Licensor.

d. Nothing in this Public License constitutes or may be interpreted as a limitation upon, or waiver of, any privileges and immunities that apply to the Licensor or You, including from the legal processes of any jurisdiction or authority.

Creative Commons is not a party to its public licenses. Notwithstanding, Creative Commons may elect to apply one of its public licenses to material it publishes and in those instances will be considered the "Licensor." Except for the limited purpose of indicating that material is shared under a Creative Commons public license or as otherwise permitted by the Creative Commons policies published at `creativecommons.org/policies`, Creative Commons does not authorize the use of the trademark "Creative Commons" or any other trademark or logo of Creative Commons without its prior written consent including, without limitation, in connection with any unauthorized modifications to any of its public licenses or any other arrangements, understandings, or agreements concerning use of licensed material. For the avoidance of doubt, this paragraph does not form part of the public licenses.

Creative Commons may be contacted at `creativecommons.org`.

Appendix E

References

"Applications Manual for Computing Amplifiers for Modelling, Measuring, Manipulating, & Much Else", Philbrick Researches, Inc., Analog Devices, 1998.

"Bell Laboratories Record", Volume XXII, Number IV, "Electrical Gun Director Demonstrated" article, pages 157-167, December 1943.

Black, Harold, *US Patent 2,102,671*, "Wave Translation System", application 22 April 1932, patent granted 21 Dec 1937.

Coughlin, Robert F. and Driscoll, Frederick F., *Operational Amplifiers and Linear Integrated Circuits*, Third Edition, Prentice-Hall Incorporated, Englewood Cliffs, NJ, 1987.

Dames, Andrew Nicholas and Colby, Edward Crellier, *US Patent 6,414,475*, "Current Sensor", application 6 November 2000, patent granted 2 July 2002.

Haseloff, Eilhard, "Latch-Up, ESD, and Other Phenomena", application report SLYA014A, Texas Instruments Incorporated, May 2000.

Johnson, Marty; Cline, Roger; Ward, Scott; Schichl, Joe; "Latch-Up", whitepaper SCAA124, Texas Instruments Incorporated, April 2015.

"LFx5x JFET Input Operational Amplifiers", datasheet SNOSBH0D, Texas Instruments Incorporated, original May 2000, revised November 2015.

"LM324, LM324A, LM324E, LM224, LM2902, LM2902E, LM2902V, NCV2902 Single Supply Quad Operational Amplifiers", datasheet LM324/D, revision 29, Semiconductor Component Industries, LLC, Aurora, CO, October 2016.

Mindell, David A. "Opening Black's Box – Rethinking Feedback's Myth of Origin", *Technology and Culture*, Volume 41, Society for the History of Technology, July 2000.
"Model K2-W Operational Amplifier", GAP/R Model K2-W, Electronic Analog Computing Instruments datasheet, George A. Philbrick Researches, Inc., Analog Devices, Boston, MA.

"Operational Manual for the Heath Educational Analog Computer model EC-1", datasheet LM324/D, revision 29, Semiconductor Component Industries, LLC, Aurora, CO, 20 March 1959.

Ragazzini, John R. and Randall, Robert H, and Russell, Frederick A. "Analysis of Problems in Dynamics by Electronic Circuits", *Proceedings of the I.R.E*, May 1947.

Rider, John F., *Vacuum-Tube Voltmeters*, Second Edition, John F. Rider Publisher, Inc, New York, 1951.

Sedra, Adel S. and Smith, Kenneth C., *Microelectronic Circuits*, Seventh Edition, Oxford University Press, New York, 2015.

Smith, John I., *Modern Operational Circuit Design*, Wiley-Interscience, John Wiley & Sons, Inc., New York, NY, 1971.

"TL08xx JFET-input Operational Amplifiers", datasheet SLOS0811, Texas Instruments Incorporated, Dallas, TX, original February 1977, revised May 2015.

"μA741 General-Purpose Operational Amplifiers", datasheet SLOS094G, Texas Instruments Incorporated, Dallas, TX, original November 1970, revised January 2018.

Appendix F

Version history

This is a list showing all significant additions, corrections, and other edits made to this learning module. Each entry is referenced by calendar date in reverse chronological order (newest version first), which appears on the front cover of every learning module for easy reference. Any contributors to this open-source document are listed here as well.

23 September 2026 – added a new Historical Reference section on George Philbrick's pioneering work on automatic gun tracking systems.

12 April 2026 – added a new Conceptual Reasoning question challenging students to design an electronic load circuit capable of both current regulation and voltage regulation.

25 March 2026 – added some instructor notes.

24 February 2026 – added a new Historical Reference section showing the operation of a balance scale.

29 November 2025 – added to the "Precision rectification" section of the Tutorial to show AC meter application. Also corrected a typographical error.

6 November 2025 – added a Quantitative Reasoning question based on the existing "Three-phase oscillator" question, the new one having students write a simulation program to determine good resistor values.

26-29 October 2025 – edited image_1921 to give a more realistic perspective on the tachogenerator. Also made minor edits to the wording of the "Fooling a voltage regulator" Conceptual Reasoning question. Also changed the title of the "Internal faults in a TL082" Diagnostic Reasoning question to "Internal faults in a model TL082". Also edited some instructor comments.

7 October 2025 – reorganized the first section of the Tutorial into multiple sections, and added micro-photographs of some opamp ICs from zeptobars.

20 September 2025 – added a new Derivations and Technical references section exploring the

phenomenon of latch-up.

6 May 2025 – added more content to the "Instrumentation amplifiers" section of the Tutorial chapter showing the two-amplifier design in addition to the three-amplifier design that was already discussed.

24 April 2025 – added another Technical References section, this one on common-mode signals.

8 April 2025 – edited some instructor notes.

30 March to 2 April 2025 – added some instructor notes and instructor images to questions. Also expanded upon the "Load current path" Conceptual Reasoning question.

7 January 2025 – changed the title of the "Testing common-mode rejection" Diagnostic Reasoning question to "Experimentally testing common-mode rejection".

13 November 2024 – added a new Technical Reference section on general opamp recommendations.

5 November 2024 – minor edits to the Tutorial section on amplification with offset.

4 November 2024 – added a pagebreak in the "Current regulator simulation program" Quantitative Reasoning question. Also corrected some typographical errors (including non-capitalized examples of *Zener*). Also added schematic diagram of an audio power amplifier circuit employing negative feedback to the "Boosted output" section of the Tutorial.

30-31 October 2024 – added a Challenge question to the "Necessary resistor values for specified voltage gains" Quantitative Reasoning question. Also added a heading to the last circuit example in the "Summer circuit calculations" Quantitative Reasoning question.

29 October 2024 – added some bullet-point items to the list of challenging concepts in the Introduction chapter.

2-3 October 2024 – divided the Introduction chapter into sections, one with recommendations for students, one with a listing of challenging concepts, and one with recommendations for instructors.

10 April 2024 – corrected image_5663 where I had an error on the square wave output waveform.

8 April 2024 – typo correction (LM7095 should have been LM7905). Also renamed Case Tutorial section from "Example: DC power supply splitter" to "Example: DC power supply rail splitter".

4 April 2024 – added more discussion of $y = mx + b$ to the "Amplification with offset" Tutorial section.

3 April 2024 – typographical error correction courtesy of Jacob Stormes. Also numbered the two sub-questions in the "Subtractor calculations" and "Summer circuit calculations" Quantitative Reasoning questions.

1-2 April 2024 – corrected a copy-and-paste error in one of the instructor notes. Also added "0

mA" labels to image_1936 and image_1937.

15 March 2024 – added some photos (courtesy of `dvq.com`) to the Tutorial section on analog computing, showing an old Philbrick analog computing manifold.

24 February 2024 – added more instructor notes to the "Three-phase oscillator" Quantitative Reasoning question. Also added another Quantitative Reasoning question called "Opamp signal scaling circuits". Also added the IEC60063 component series reference to the Technical References chapter, as this will be necessary to answer the signal scaling question. Also added another page of text to the "Differential amplification" Tutorial section discussing non-unity voltage gains.

20 January 2024 – added a new Technical Reference section on protecting amplifier inputs against overvoltage from electro-static discharge (ESD) as well as electrical over-stress (EOS) from improperly connected signal sources.

15 November 2023 – corrected error in image_5613 (output spiking wrong direction for two of the input's step-changes), courtesy of Joe Archer.

14 November 2023 – added a new Case Tutorial section on empirically determining signal rates of change.

26 October to 12 November 2023 – added more comments to the Tutorial regarding the simplifying omission of DC power supply connections in opamp schematic diagrams, as well as comments regarding how negative feedback improves amplifier stability and added questions to the Introduction chapter. Minor text edits throughout the Tutorial chapter. Corrected error in C code for the instructor notes of the "Precision rectifier simulation program" Quantitative Reasoning question. Added new Tutorial section on analog computing. Elaborated on common-mode voltage input limitations for opamps.

10 September 2023 – added Quantitative Reasoning question on electronic load circuitry, complete with sketching a load line for the electronic load circuit.

11 June to 6 July 2023 – added questions related to computer programming (for the simulation of opamp circuits).

28 November 2022 – placed questions at the top of the itemized list in the Introduction chapter prompting students to devise experiments related to the tutorial content.

20 July 2022 – corrected an error in image_2841 by swapping polarities of the 2 Volt, 5 Volt, and 7 Volt sources in the lowest-left example.

24 May 2022 – minor grammatical edit to footnote on the use of capacitors versus inductors in integrator circuits, courtesy of Ty Weich.

19 April 2022 – added a pictorial illustration of a split supply made from two batteries, to the Tutorial. Also added a new Tutorial section on instrumentation amplifiers.

13 April 2022 – corrected two rather embarrassing omissions today. The first was the text of the

last paragraph in the "Oscillators" section of the Tutorial was incomplete. The last sentence of this paragraph merely said "Fourier's Theorem" and nothing else(!). The next corrected omission was a reference in the Tutorial suggesting readers to go the Historical References chapter to read about William Hewlett's clever Wein bridge oscillator design, but in fact I had not included that historical reference in this module. After fixing these blunders, I did more prosaic work by adding an entry to the "Foundational Concepts" subsection as well as some more questions in the Introduction chapter, and editing some instructor notes.

12 April 2022 – added two more pages to the Tutorial section on integrators and differentiators showing numerical examples of the capacitor and inductor formulae in action, to make those calculus expressions less mysterious.

4-8 April 2022 – added more elaboration on functional inversion to the Tutorial chapter, and added more questions to the Introduction. Also edited the Case Tutorial section "Example: DC power supply splitter" schematic diagram to include a low-value resistor between the transistors' base and emitter terminals for improved stability.

8 March 2022 – added a Case Tutorial section showing how to simulate opamp circuits in SPICE using voltage-controlled voltage source elements.

4 March 2022 – added a Case Tutorial section on a DC power supply "splitting" circuit.

11 December 2021 – added a Case Tutorial section on a negative feedback with a multi-stage amplifier to achieve a modest and stable voltage gain.

16 November 2021 – added note on Wein bridge oscillators about sizing the gain resistors.

10 November 2021 – added another bullet-point question to the "Rogowski coil power meter" Conceptual Reasoning question, as well as added TP1 test point labels to image_5621 and image_1943.

2-4 November 2021 – added more content to the new section on oscillators and added new questions on oscillator.

1 November 2021 – added a Tutorial section discussing amplification with offset.

27-30 October 2021 – expanded the "Integrators and Differentiators" section of the Tutorial, included a Case Tutorial section on a simple triangle/square/PWM oscillator, and added many questions. Also added a new section on precision diode circuits.

24 October 2021 – minor edits to the Tutorial regarding shunt voltage regulation. Relocated the "Boosted output" section to precede integrators and differentiators. Added questions. Added details to the Historical Reference on the Philbrick K2-W vacuum tube opamp.

20 October 2021 – added some Challenge questions.

14 October 2021 – added a Case Tutorial section showing multiple operational amplifier experiments. Also colorized DMM digits in image_3741.

31 May 2021 – modified image_2841 to have more complex examples.

10 May 2021 – commented out or deleted empty chapters.

18 April 2021 – changed all lower-case Greek letter "phi" symbols (ϕ) to upper-case (Φ).

18 March 2021 – corrected multiple instances of "volts" that should have been capitalized "Volts".

15 March 2021 – edited the wording of the Cardinal Rule of Negative Feedback to eliminate one "limits".

11 March 2021 – added some instructor notes.

11 February 2021 – corrected instances where "Ohms" was uncapitalized.

28 December 2020 – added a new section to the Tutorial on differentiators and integrators.

26 December 2020 – added a new section to the Tutorial on boosting the output of an opamp.

8-9 December 2020 – minor edits to the Introduction, added a Challenge question, and edited some of the instructor notes.

19 October 2020 – corrected minor typographical error in two schematic diagrams in the Tutorial, where a resistor was labeled $R_{feeback}$ instead of $R_{feedback}$.

14-5 October 2020 – minor edits to illustrations in questions, and minor edits to the Tutorial, and minor additions to the Introduction chapter.

2 October 2020 – significantly edited the Introduction chapter to make it more suitable as a pre-study guide and to provide cues useful to instructors leading "inverted" teaching sessions.

20-21 May 2020 – added more content to the Tutorial (more on differential amplification, an additional section on precision current circuits, an additional section on summers) as well as additional quantitative problems.

11 May 2020 – corrected error in diagram within the Case Tutorial chapter, where I had all the current arrows pointing the wrong way on the positive feedback example with the 5.6 kΩ and 3.3 kΩ resistors. Thanks to Ty Weich for finding this error!

6-7 May 2020 – added analog examples to the Case Tutorial, and edited two of the Case Tutorial section titles.

4-5 April 2020 – minor edits.

30 April 2020 – added Case Tutorial, copied almost verbatim from the *Comparators* learning module.

29 March 2020 – added more Foundational Concepts to the list in the Conceptual Reasoning section.

20 January 2020 – corrected a couple of answers which should have been for instructors only.

12 January 2020 – minor correction to a schematic diagram.

27 December 2019 – reassigned some questions from this module to the Comparators module, made a minor edit to the Tutorial.

26 December 2019 – removed "LM" prefix on opamp part numbers.

19 December 2019 – added more questions and made minor edits to some illustrations.

18 December 2019 – added more questions; minor edits to diagnostic questions, replacing "no multiple faults" with "no coincidental faults".

17 December 2019 – finished first draft of Technical Reference section on opamp imperfections, added some Conceptual Reasoning questions.

16 December 2019 – minor edits to Tutorial to improve page formatting. Added comment on the first page about voltage amplifiers generally having extremely high impedance, and the assumption of zero input current. Added a Technical Reference section on opamp imperfections.

29 March 2020 – made minor changes to section on K2-W opamp in the Historical References chapter.

24 January 2020 – added some entries to the Foundational Concepts list.

20 October 2019 – added patent number and award date to the Historical Reference section on Harold Black's invention using negative feedback for improved amplifier stability.

19 June 2019 – minor edits and additions made to the Tutorial.

18 June 2019 – minor edits and additions made to the Tutorial.

17 June 2019 – continued writing Tutorial, also added and modified many questions and added a section in the Historical References chapter to the Philbrick Researches model K2-W.

14 June 2019 – continued writing Tutorial.

13 June 2019 – continued writing Tutorial, adding sections.

12 June 2019 – began writing Tutorial.

11 June 2019 – document first created.

Index

y intercept, 62
741 amplifier, 37

Absolute value, 93
Acceleration, 103
Accelerometer, 105
Adding quantities to a qualitative problem, 320
AGC, 125
Analog computer, 105
Annotating diagrams, 319
Array, C programming, 283
Artifact, 104
Automatic gain control, 125

Barkhausen criterion, 121
Barkhausen, Georg, 121
Bel, 167
Bias current, 184
Bias voltage, opamp, 188
Bipolar transistor, 192
Black, Harold, 15
Block diagram, 45, 53
Bridge circuit, 41
Bridge rectifier, 46

Calculus, 100
Capacitance, 100
Capacitance, parasitic, 195
Cardinal rule of negative feedback, 44, 45, 47, 51, 75, 82–85
Checking for exceptions, 320
Checking your work, 320
Choke, common-mode, 161
Clipping network, 192
Closed-loop gain, 44
CMOS, 193
CMRR, 191
Code, computer, 329
Colpitts oscillator, 126
Common logarithm, 167
Common-mode choke, 161
Common-mode current, 161
Common-mode rejection ratio, 191
Common-mode signal, 156
Common-mode voltage, 156, 183, 191
Compensation resistor, 108
Compliance voltage, 85
Component values, IEC standard, 154
Computer, analog, 105
Concentration, doping, 199
Contact, ohmic, 206
Contact, rectifying, 206
Control loop, 145
Coriolis flowmeter, 127
Crystal, 127
Current probe, 31
Current pump, 47
Current source, 47, 84
Current, bias, 184
Current, common-mode, 161

dB, 167
dBm, 171
dBW, 172
Decibel, 167
DIAC, 195
Differential equation, 295
Differentiation, 100
Dimensional analysis, 319
Diode, protection, 182, 186, 192
Diode, Schottky, 193
Displacement current, 31
Distance versus speed, 33
Drift, 108

Edwards, Tim, 330
Electrical load, 55, 56
Electrical over-stress, 192
Electro-static discharge, 192
EOS, 192
ESD, 192

Feedback, 42, 107, 115, 121
Feedback, negative, 14
Field-effect transistor, 192
Filter, low-pass, 195
Flowmeter, Coriolis, 127
Follower, voltage, 44, 69
Fourier transform, 173
Fourier's Theorem, 127
Full-wave rectifier, 93
Function notation, 53
Function, linear, 60–62
Functional inversion, 52

Gain control, automatic, 125
Gain, amplifier, 167
Gain, closed-loop, 44
Gain, open-loop, 44
Gain-bandwidth product, 189
GBWP, 189
George A. Philbrick Researches, Inc., 146
George A. Philbrick Researches, Incorporated, 113
Glitch, 94
Graph values to solve a problem, 320
Greenleaf, Cynthia, 207
Ground, virtual, 57

Half-wave rectifier, 95
Harmonic, 124
Hartley oscillator, 126
Hewlett, William, 149
Hewlett-Packard, 149
How to teach with these modules, 327
Hwang, Andrew D., 331
Hysteresis, 115, 121

Identify given data, 319
Identify relevant principles, 319
IEC 60063 standard, 154
IEC standard component values, 154
Impedance, input, 35, 69
Inductance, 100
Inertial guidance, 105
Infra-red light sensor, 107
Input impedance, 35, 69
Input protection diode, 182, 186, 192
Input, rail-to-rail, 183
Insertion resistance, 86
Instrumentation amplifier, 73
Integration, 100
Integrator, 26
Intermediate results, 319
Inversion, functional, 52
Inverted instruction, 327
Inverter gate, 199
Inverting input, 37

Joule's Law, 174

K2-W operational amplifier, 113, 146
Kirchhoff's Current Law, 47, 59, 76, 78, 82–84
Kirchhoff's Voltage Law, 47, 55, 56, 58, 67, 70, 78, 82, 84, 99
Knuth, Donald, 330

Lamport, Leslie, 330
Laplace transform, 173
Latch-up, 183, 197
Latch-up, pseudo, 206
Limiting cases, 178, 320
Linear function, 60–62
Load, 55, 56
Logarithm, common, 167
Low-pass filter, 195

Maxwell, James Clerk, 129
Mergler, Harry, 109
Metacognition, 212
Meter loading, 86
Metrology, 132
Millman's Theorem, 179
Moolenaar, Bram, 329
MOSFET, 192
Murphy, Lynn, 207

Negative feedback, 14, 42, 132

Neutralization, 136
Newton, Isaac, 103
Noise, 108
Non-inverting input, 37
Norton equivalent source, 179
Norton's Theorem, 80
Notation, function, 53
NPN, 192
Null balance technique, 132
Null detector, 132

Offset currrent, opamp, 187
Offset voltage, opamp, 188
Ohm's Law, 47, 55, 56, 58, 66–68, 70, 80, 82–84, 99, 179
Ohmic contact, 206
Opamp, 15, 38, 44
Open-loop, 41
Open-loop gain, 44
Open-source, 329
Operational amplifier, 15, 38, 44, 132
Output, rail-to-rail, 92, 181

Parasitic capacitance, 195
Parasitic effect, 31
Parasitic transistor, 197
Paynter, Henry M., 148
Philbrick Researches, Inc., 146
Philbrick, George A., 113
Photocell, 41
Pierce oscillator, 127
Piezoelectricity, 127
PIR sensor, 107
Pitch, 31
PNP, 192
Position, 103
Positive feedback, 107, 115, 121
Problem-solving: annotate diagrams, 319
Problem-solving: check for exceptions, 320
Problem-solving: checking work, 320
Problem-solving: dimensional analysis, 319
Problem-solving: graph values, 320
Problem-solving: identify given data, 319
Problem-solving: identify relevant principles, 319
Problem-solving: interpret intermediate results, 319
Problem-solving: limiting cases, 178, 320
Problem-solving: qualitative to quantitative, 320
Problem-solving: quantitative to qualitative, 320
Problem-solving: reductio ad absurdum, 320
Problem-solving: simplify the system, 319
Problem-solving: thought experiment, 57, 65, 71, 319
Problem-solving: track units of measurement, 319
Problem-solving: visually represent the system, 319
Problem-solving: work in reverse, 320
Process variable, 145
Protection diode, 182, 186, 192
Pseudo latch-up, 206
Pump, current, 47
Push-pull amplifier, 199

Qualitatively approaching a quantitative problem, 320
Quartz crystal, 127

Rail, DC power supply, 192
Rail-to-rail input, 183
Rail-to-rail output, 92, 181
Reading Apprenticeship, 207
Rectifier, bridge, 46
Rectifier, full-wave, 93
Rectifier, half-wave, 95
Rectifying contact, 206
Reductio ad absurdum, 320, 326, 327
Regulator, voltage, 46
Resistor, compensation, 108
Resonance, 125
Rise, 61, 62
Rise over run, 31
Robotics, 49
Rogowski coil, 107
Run, 61, 62

Schoenbach, Ruth, 207
Schottky diode, 193
Scientific method, 212
Second Law of Motion, 103
Series regulator, 46
Servo system, 49

Setpoint, 145
Shunt regulator, 48
Signal, common-mode, 156
Simplifying a system, 319
Sinking, 199
Slope, 31, 61, 62
Socrates, 326
Socratic dialogue, 327
Source, current, 47
Sourcing, 199
Speed versus distance, 33
SPICE, 207
Split DC power supply, 177
Stallman, Richard, 329
Standard component values, IEC, 154
Star grounding, 204
Star-point power distribution, 204
Subtractor, 65
Superposition Theorem, 63
Superposition theorem, 65, 71

Tangent line, 31
Temperature control system, 50
Thévenin resistance, 187
Thought experiment, 57, 65, 71, 319
Thyristor, 195, 197
Torvalds, Linus, 329
Transform function, 173
Transient, 94
Transistor, 192
Transistor, parasitic, 197

Units of measurement, 319

Vacuum tube voltmeter, 86
Velocity, 103
Virtual ground, 57, 104
Visualizing a system, 319
Voltage follower, 44, 51, 69, 186, 190
Voltage regulator, 46
Voltage regulator, series, 46
Voltage regulator, shunt, 48
Voltage, common-mode, 156, 183, 191
Voltage, compliance, 85
VTVM, 86

Wein bridge, 124, 149
Work in reverse to solve a problem, 320
WYSIWYG, 329, 330