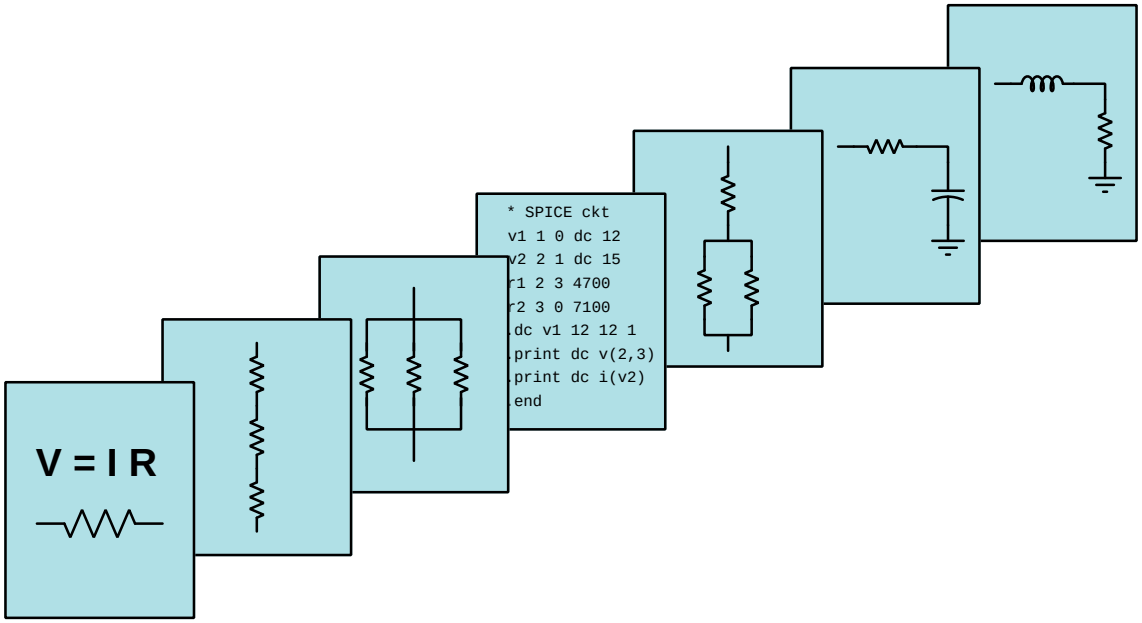
Modular Electronics Learning (ModEL) project



Electrical Metrology

© 2022-2026 by Tony R. Kuphaldt – under the terms and conditions of the Creative Commons Attribution 4.0 International Public License

Last update = 7 May 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 electrical metrology . . . 6
- 1.3 Recommendations for instructors . . . 7

2 Case Tutorial **9**
- 2.1 Example: sensitive audio detector circuit . . . 10
- 2.2 Example: electronic null detector circuit . . . 13
- 2.3 Example: using Python to control a LabJack model U3 DAQ . . . 15
- 2.4 Example: SCPI control of digital power supply and multimeter . . . 24
  - 2.4.1 Introduction and set-up . . . 24
  - 2.4.2 Python/VISA/SCPI script: user-specified voltages . . . 30
  - 2.4.3 Python/VISA/SCPI script: DC voltage sweep . . . 31
- 2.5 Example: SCPI control of digital power supply and oscilloscope . . . 33
  - 2.5.1 Introduction and set-up . . . 33
  - 2.5.2 Python/VISA/SCPI script: DC voltage sweep . . . 34
- 2.6 Example: SCPI control of power supply via GUI . . . 36
- 2.7 Example: damage caused by an errant test script . . . 38
- 2.8 Example: GPS receiver data stream . . . 41

3 Tutorial **43**
- 3.1 What is metrology? . . . 44
- 3.2 Metrological standards . . . 46
- 3.3 Measurement terminology . . . 48
- 3.4 Measurement uncertainty . . . 50
- 3.5 Proportionate expressions . . . 54
- 3.6 Typical instrument errors . . . 55
- 3.7 Drift and instability . . . 60
- 3.8 Intrinsic noise . . . 61
  - 3.8.1 Thermal noise . . . 62
  - 3.8.2 Shot noise . . . 63
  - 3.8.3 Flicker noise . . . 63
  - 3.8.4 Burst noise . . . 64
  - 3.8.5 Avalanche noise . . . 64

3.8.6 Intrinsic noise mitigation . . . 64
3.9 Electrical metrology standards . . . 64
3.9.1 DC voltage standards . . . 65
3.9.2 Resistance standards . . . 73
3.9.3 DC current standards . . . 77
3.9.4 Frequency and time standards . . . 78
3.10 Precision voltage measurement techniques . . . 85
3.10.1 Null-balance voltage measurement . . . 85
3.10.2 Null-balance meter calibration . . . 87
3.10.3 Thermal voltage errors . . . 90
3.11 Precision resistance measurement techniques . . . 92
3.11.1 Two-wire resistance measurement . . . 92
3.11.2 Four-wire resistance measurement . . . 93
3.11.3 Wheatstone bridge resistance measurement . . . 95
3.12 Guarded measurements . . . 97
3.13 Electrical metrology tools . . . 106
3.13.1 Multifunction calibrators . . . 106
3.13.2 Kelvin-Varley voltage dividers . . . 107
3.13.3 Hamon voltage dividers . . . 113
3.14 Automated electrical testing . . . 115
3.14.1 Vector Network Analyzer calibration . . . 115
3.14.2 Industrial instrument calibration . . . 118
3.14.3 Manufactured product testing . . . 119
3.14.4 Component characterization testing . . . 121
3.14.5 Accelerated stress testing . . . 126
3.14.6 Programmable test instruments . . . 130

4 Historical References **133**
4.1 Platform balance scale operation . . . 134
4.2 Standard voltage cells . . . 137
4.3 Standard resistors . . . 139
4.4 A binary resistance box . . . 143
4.5 Early references to Wheatstone bridges . . . 145

5 Derivations and Technical References **149**
5.1 IEC standard component values . . . 150
5.2 Capacitor parasitics . . . 152
5.2.1 Model of a real capacitor . . . 152
5.2.2 Parasitic resistance in capacitors . . . 153
5.2.3 Parasitic inductance in capacitors . . . 153
5.2.4 Other parasitic effects in capacitors . . . 154
5.3 Inductor parasitics . . . 155
5.3.1 Model of a real inductor . . . 155
5.3.2 Parasitic resistance in inductors . . . 156
5.3.3 Parasitic capacitance in inductors . . . 157
5.3.4 Other parasitic effects in inductors . . . 157

5.3.5 Inductor testing . . . 158
5.4 Transformer parasitics . . . 159
5.4.1 Parasitic resistance in transformers . . . 160
5.4.2 Leakage inductance in transformers . . . 160
5.4.3 Parasitic capacitance in transformers . . . 161
5.4.4 Excitation and magnetization currents . . . 162
5.4.5 Transformer testing . . . 163
5.5 Electric field quantities . . . 165
5.6 Magnetic field quantities . . . 167
5.7 The OSI Reference Model . . . 177
5.8 Statistics . . . 179

6 Questions **183**
6.1 Conceptual reasoning . . . 187
6.1.1 Reading outline and reflections . . . 188
6.1.2 Foundational concepts . . . 189
6.1.3 Analog multimeter calibration . . . 192
6.1.4 Potentiometric voltmeter . . . 193
6.1.5 Test uncertainty ratio . . . 193
6.1.6 Clock accuracy . . . 194
6.1.7 EKG monitor . . . 195
6.1.8 Testing intrinsic standards . . . 196
6.2 Quantitative reasoning . . . 197
6.2.1 Miscellaneous physical constants . . . 198
6.2.2 Introduction to spreadsheets . . . 199
6.2.3 Kelvin-Varley voltage divider . . . 202
6.2.4 Hamon voltage divider simulation program . . . 203
6.2.5 Testing a 250 V meter with a 10 V standard . . . 204
6.2.6 Measuring time in Python . . . 205
6.3 Diagnostic reasoning . . . 206
6.3.1 Thermometer calibration results . . . 207
6.3.2 Voltmeter calibration results . . . 208
6.3.3 Ammeter calibration results . . . 209

A Problem-Solving Strategies **211**

B Instructional philosophy **213**
B.1 First principles of learning . . . 214
B.2 Proven strategies for instructors . . . 215
B.3 Proven strategies for students . . . 217
B.4 Design of these learning modules . . . 218

C Tools used **221**

D Creative Commons License **225**

E References **233**

F Version history **237**

Index **239**

Chapter 1

Introduction

1.1 Recommendations for students

Metrology is the science of precision measurement, and it is absolutely essential for commerce and for scientific investigations. In this module we will explore basic principles of metrology as well as practical realizations of it for electrical quantities such as voltage, current, resistance, and frequency. Additionally, we will explore techniques useful for minimizing error in precision measurements of voltage, current, and resistance.

Important concepts related to metrology include **uncertainty**, **intrinsic standard**, **artifact standard**, **traceability**, **NMI**, **primary standard**, **secondary standard**, **working standard**, **transfer standard**, **accuracy**, **error**, **zero**, **span**, **range**, **resolution**, **calibration**, **standardization**, **Gaussian curve**, **standard deviation**, **confidence interval**, **percent**, **ppm**, **ppb**, **ppt**, **linearity**, **monotonicity**, **hysteresis**, **drift**, **coefficient**, **noise**, **harmonic** frequency, **noise floor**, **standard cell**, **Josephson effect**, **quantum Hall effect**, **Zener diode**, **GPS**, **atomic clock**, **quartz crystal**, **insertion resistance**, **null-balance**, **galvanometer**, **Seebeck effect**, **Kelvin 4-wire measurement**, **bridge** network, and magnetic **permeability**.

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

- Why is metrology important to modern life?
- Which NMI serves as the authority on weights and measures for the Unites States?
- What differentiates primary standards from secondary standards?
- What are some examples of *artifact* standards?
- What are some examples of *intrinsic* standards?
- How do artifact versus intrinsic standards compare in terms of practicality, portability, accuracy, etc.?
- How do the terms *zero*, *span*, and *range* relate to one another?

- What does *uncertainty* mean in the metrological sense of the word?
- How is uncertainty quantitatively expressed?
- What are some of the different types of calibration errors commonly seen in instrument responses?
- What does *parts per million* mean, and how might one apply this term to some practical measurement?
- What are some of the external factors that may cause a standard's true value to drift?
- How does a mathematical *coefficient* help us predict the extent to which some standard's value may be affected by such an external factor?
- What is a *random* error, and how does it contrast against a *systematic* error?
- What is a "noise floor"?
- What does it mean to specify the "color" of a particular noise?
- What are some of the different types of electrical noise intrinsic to components?
- To what extent are these different types of electrical noise controllable through judicious circuit design?
- How has the *Volt* been defined, historically?
- How do we define the *Volt* today?
- How may Zener diodes be made less prone to drift when used as voltage standards?
- How does *standardization* differ from *calibration*?
- How has the *Ohm* been defined, historically?
- How do we define the *Ohm* today?
- How has the *Ampere* been defined, historically?
- How do we define the *Ampere* today?
- How has the *second* been defined, historically?
- How do we define the *second* today?
- What method(s) exist(s) for mitigating the effects of noise on precision measurements?
- How might an experiment be designed and conducted to demonstrate the *drift* of a digital multimeter over time?
- How may computer programs be used to automate electronic device testing?
- What does it mean to *characterize* an electronic component?

- How do "bed of nails" test fixtures differ from "flying probe" testers?
- Why are Vector Network Analyzers (VNAs) so labor-intensive to calibrate?
- What is *SCPI*, and why is it relevant to automated testing?
- If the energy emitted by an excited cesium atom is fixed and not liable to variation, then how it is possible to design "better" atomic clocks based on the same cesium-atom standard?

1.2 Challenging concepts related to electrical metrology

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

- **Error versus Uncertainty** – error is a naive concept because it implies a perfect and perfectly-known standard against which we may compare something that is less perfect. The problem here is that no physical measurement can ever be known with perfect accuracy and perfect precision, inescapable *noise* being one of the confounding factors. Therefore, in metrology we speak of a measurement's *uncertainty* rather than *error*, knowing that the degree of uncertainty never will be and can never be zero.

- **Standardization versus Calibration** – to standardize an instrument means to check its operation against a trusted standard. Calibration, on the other hand, implies corrective action to make the instrument register more accurately than it did before.

- **Kelvin 4-wire resistance measurement** – the four-wire method of measuring resistance is a very clever solution to the problem of eliminating resistance-measurement errors stemming from inevitable lead wire resistance connecting the resistance specimen to the ohmmeter. One strategy to understand this concept better is to add example quantities to the circuit diagram (e.g. 1 mA current source, 1 Ohm for each lead wire resistance, 700 Ohm specimen, etc.) and then use Ohm's Law and Kirchhoff's Voltage Law to calculate voltage sensed by the instrument.

- **Kelvin-Varley voltage dividers** – an ingenious method for creating a voltage divider adjustable in precise decades (powers of ten) involves cascading multiple voltage divider networks such that each successive network is powered by the reduced output of the previous divider. Normally we try to avoid loading resistive voltage dividers because the load causes the voltage divider's output voltage to "sag" below its intended ratio, but with Kelvin-Varley voltage dividers this loading effect is anticipated in the design such that the ratios remain correct. Careful series-parallel analysis of each divider stage proves how this works.

- **Hamon voltage dividers** – an even more ingenious method for creating precise and self-calibrating voltage dividers is to use groups of resistors alternately connected in parallel (for 1:2 self-calibration) and then in series (for a $1:(n^2+1)$ ratio where n is the number of resistors). Such is called a *Hamon* divider, and its real genius consists of minimization of error by selecting resistors arbitrarily close in value for the series/parallel switchable network. An understanding and appreciation of this ingenuity is found by performing calculations on a divider's ratio given imperfect resistance values, seeing for oneself how small the division ratio error becomes compared to individual resistance errors.

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** – Apply foundational circuit concepts to the analysis of biomedical instrumentation circuits

  Assessment – Determine how parasitic electrical properties of electrode-skin contacts may be overcome in order to achieve accurate measurements; e.g. pose problems in the form of the "EKG monitor" Conceptual Reasoning question.

- **Outcome** – Design an electrical testing circuit

  Assessment – Explain how to connect a Kelvin-Varley voltage divider to a precision voltage reference in order to standardize or calibrate a voltmeter; e.g. pose problems in the form of the "Testing a 250 V meter with a 10 V standard" Conceptual Reasoning question.

- **Outcome** – Diagnose a miscalibrated instrument

  Assessment – Identify the nature of a calibration error based on given "As-Found" test data; e.g. pose problems in the form of the "Thermometer calibration results" and "Voltmeter calibration results" and "Ammeter calibration results" Diagnostic Reasoning questions.

- **Outcome** – Independent research

  Assessment – Locate laboratory-grade standard documentation (e.g. manuals, datasheets) and properly interpret some of the information contained in those documents including drift specifications, uncertainty, calibration ranges, etc.

Chapter 2

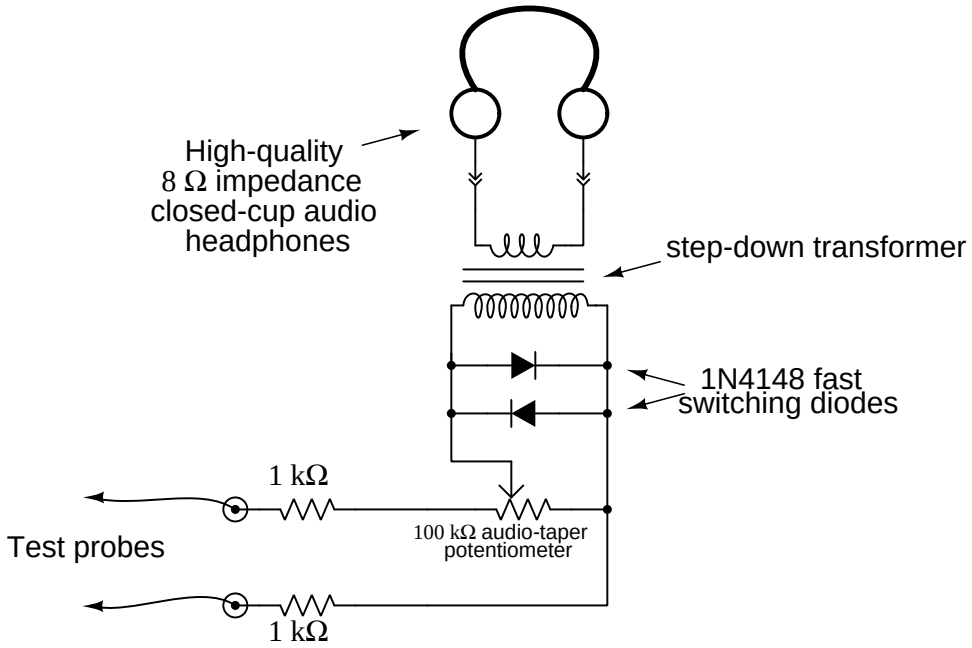
Case Tutorial

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: sensitive audio detector circuit

A useful project for sensing low-voltage signals is this "sensitive audio detector" circuit using headphones to convert electrical energy into acoustic energy you can hear:



Any AC signal with a frequency between approximately 20 Hz and 20 kHz (depending on the range of your own hearing) may be heard using this test instrument. Such an audio detector is extremely useful for probing audio amplifier circuits, allowing you to hear the signal at different points along the amplification circuitry. With good-quality headphones or earbuds, the sensitivity is great enough to allow the user to hear 60 Hz electrical "noise" even without direct contact of the probes (e.g. grounding one probe and holding the other *near* the noise source, or connecting both probes to an inductive pick-up coil). Even DC signals may be sensed by listening for a "click" or "scratching" noise whenever the test probes come into contact with a DC voltage source.

An interesting application of this test instrument is as a *null detector* in precision electrical measurement circuits such as Wheatstone bridges, where its purpose is to aurally indicate a *lack* of electrical potential between two points. To use this as a null detector in a DC circuit, you must connect one of the test probes in series with a pushbutton switch or some other convenient means of disconnection and reconnection, in order to make small DC voltages/currents audible. For example, when used to sense voltage between the measurement points of a Wheatstone bridge circuit it will indicate a state of balance when it produces no sound at all in response to being connected and disconnected from those points. Any time the bridge is not perfectly balanced, there will be an audible "click" in the headphones/earbuds when connected and disconnected from the circuit even with currents less than one microAmpere!

The transformer acts to match the low impedance of the headphones (typically 8 Ohms) to a much greater impedance presented at the high-turns winding, this impedance transformation described by the following formula:

$$\frac{Z_P}{Z_S} = \left(\frac{N_P}{N_S}\right)^2$$

Where,

Z_P = impedance at transformer's primary winding, in Ohms
Z_S = impedance at transformer's secondary winding, in Ohms
N_P = number of turns within transformer's primary winding
N_S = number of turns within transformer's secondary winding

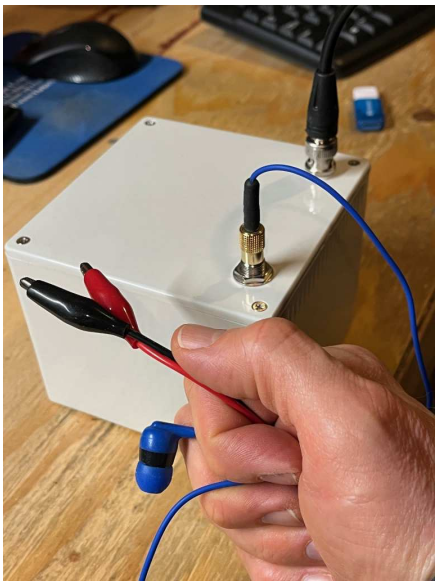
Another benefit of a transformer in this circuit design is that it also helps make very low-voltage DC signals audible by its ability to *store energy*. When connecting the detector circuit to a DC source, the transformer will build up a magnetic field in a direction corresponding to the polarity of the applied voltage. When the connection to the source is broken (by lifting one or both of the test probes away from the point(s) of contact), this magnetic field rapidly collapses and in doing so delivers all that stored energy over a brief time interval to the only load still connected: the headphones. Energy, of course, is always conserved, but it is possible for the transformer in this context to release more *power* (i.e. energy transfer per unit time) than it receives. Thus, in addition to providing safe electrical isolation between the circuit under test and the headphones, as well as impedance transformation for weak AC signals, the transformer also augments weak DC signals by its ability to rapidly release to the headphones any energy stored from the measured signal.

The two parallel-connected diodes serve to limit the maximum amount of voltage across the transformer's primary winding. These silicon model 1N4148 diodes have a typical forward voltage of approximately 0.7 Volts, meaning neither is able to conduct electricity unless the voltage across it reaches this value (in the correct polarity). Thus, any voltage less than 0.7 Volts (peak) will pass to the transformer's primary winding unattenuated, but any signal voltage greater than this will become "clipped" to a maximum value of 0.7 Volts. If the perceived volume at this peak signal amplitude is still too loud, one may insert a series resistance between the diode pair and the transformer's primary winding, experimenting with this resistor's value until a suitable maximum volume is found. Schottky diodes with lower forward-voltage drops (0.4 Volts rather than 0.7 Volts for PN-junction silicon diodes) may also be substituted.

A potentiometer provides simple volume control, useful for attenuating volume when sensing stronger signals. The two resistors in series with the test probes provide a minimum resistance between the test probes and transformer winding to prevent the transformer from excessively loading the circuit under test, as well as provide a maximum volume limit for listening. You may find it possible to use series resistors much greater than 1 kiloOhm each if your headphones are sensitive and your sense of hearing is acute. A good practice is to set the potentiometer to a position of minimal volume when first testing a signal of unknown magnitude, so that the resulting volume produced by the headphones will not be too loud. This is analogous to setting a manually-ranged multimeter to the highest range when testing a voltage or current, so that if the signal happens to be stronger than expected the meter's pointer will not be violently thrown to maximum position on

its scale. This same practice also minimizes energy draw from the circuit under test, by applying the greatest instrument impedance to any sensed voltage.

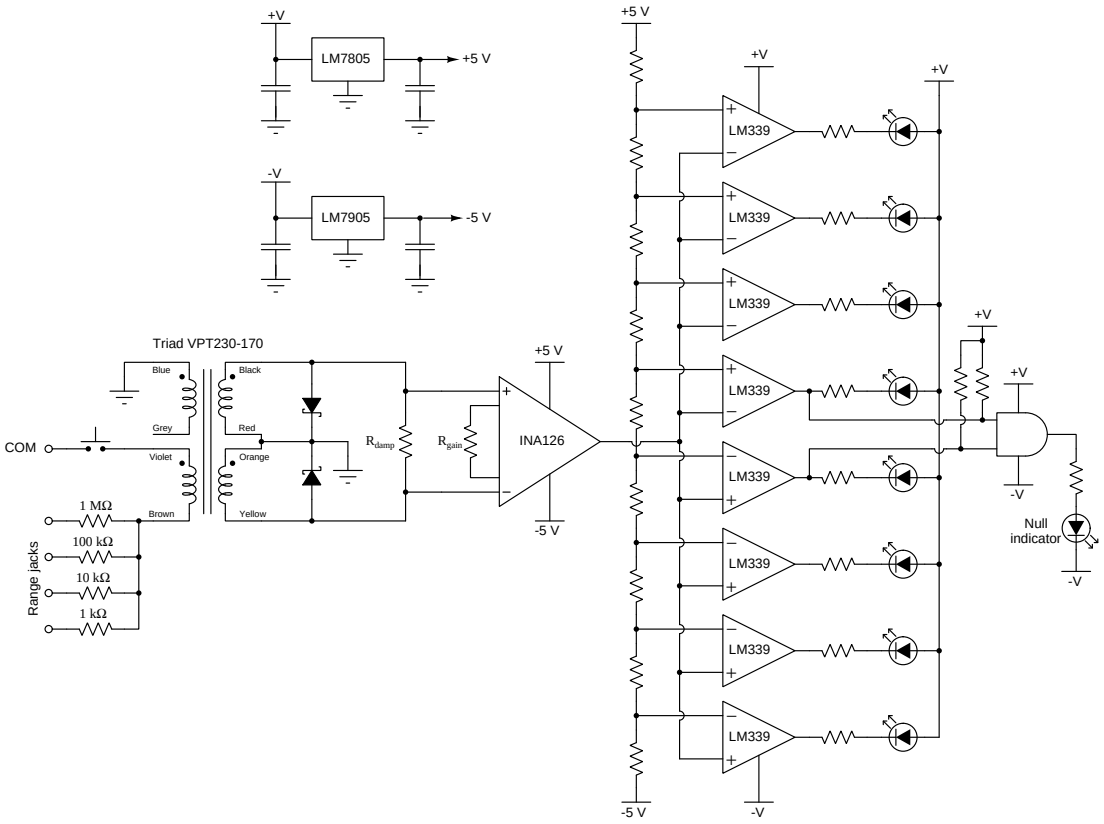
Below we see two photographs of constructed audio detectors. On the left is a unit enclosed in a plastic case (its potentiometer knob located on the far side of the enclosure, not visible in this photo), and on the right is a unit enclosed in a painted steel case with its volume-control potentiometer knob clearly visible:



A BNC connector serves as the input connection point to the plastic-enclosed unit, while a 1/4 inch audio jack serves as the output. In that left-hand photograph a standard BNC-to-alligator test cable serves as the test leads while "earbud" style earphones plug into the 1/4 inch audio jack using a 1/4 to 1/8 inch audio jack adapter. In the right-hand photograph we see a pair of banana-jack connectors serving as the input terminals while a 1/4 inch audio jack provides connection to either earphones or headphones. Steel is a superior enclosure material to plastic for this application due to its natural ability to shunt external magnetic fields around the unit's internal transformer, thereby making the steel-enclosed unit less susceptible to magnetic interference. The steel enclosure used to construct the right-hand unit also happened to have a 10-24 tapped mounting hole in the bottom perfect for mounting a toroidal transformer (Triad Magnetics model VPT24-1040 with both primary and secondary windings in series configuration, for a manufacturer-intended step-down ratio of 10:1, rated 240 VAC to 24 VAC). Toroidal transformers are also preferred over E-I core transformers for their superior rejection of external magnetic fields.

2.2 Example: electronic null detector circuit

This circuit utilizes a transformer to provide galvanic isolation between the electronic amplifier and display circuitry versus the circuit under test, in order to completely avoid any possibility of amplifier bias current entering the metrological test circuit. Pressing the pushbutton switch "charges" the transformer's primary winding, and releasing the pushbutton allows that primary winding to "discharge" very rapidly with a corresponding boost in signal voltage.

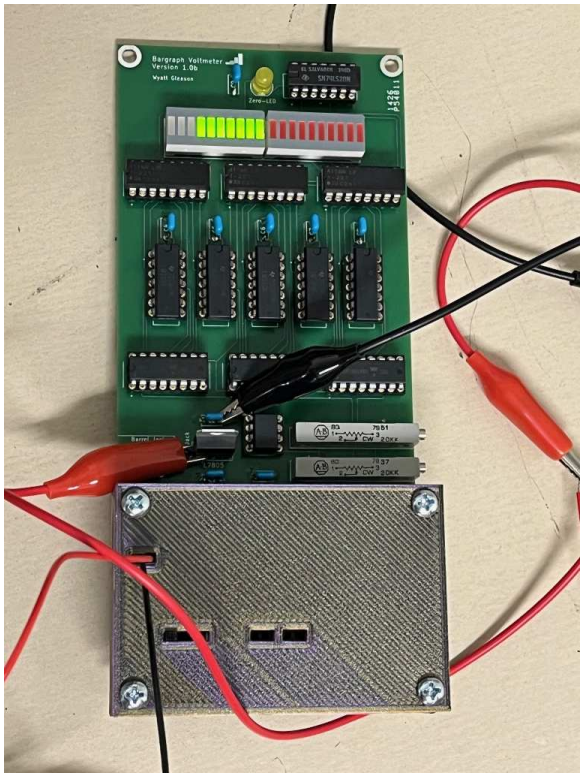


Only one of the two 120 Volt rated primary windings of this transformer is used to receive an input signal, and this is the one demonstrated to have less capacitive coupling to the secondary windings than the other. The other primary winding (Blue and Grey wires) connects to ground to function as an electrostatic shield between primary and secondary windings. The transformer's secondary windings (rated at 120 Volts each) provide isolated signal coupling to an instrumentation amplifier (INA126) with voltage gain suitably set by the value of R_{gain}. The two diodes serve to clamp the secondary windings' pulse voltage to acceptable limits for the instrumentation amplifier's inputs. The R_{damp} resistor is sized to provide critical damping for the transformer secondary windings for the purpose of damping oscillations that otherwise result from those windings' inductance combined with their parasitic capacitance. With proper damping there will only be a single-polarity pulse delivered to the amplifier with every pushbutton switch release.

After the instrumentation amplifier a series of LM339 comparators provide a simple bargraph display (eight LEDs) indicating the relative strength and polarity of each voltage pulse. A ninth LED serves as a "null indicator" to show when a condition of near-zero voltage exists. The series string of nine resistors need not be identical – in fact, it is useful to graduate the resistance of these so that the center resistor is lowest in value while the outer (+V and −V) resistors are highest, in order to provide a non-linear bargraph scale well-suited to representing a wide maximum voltage range while providing increased sensitivity near zero Volts.

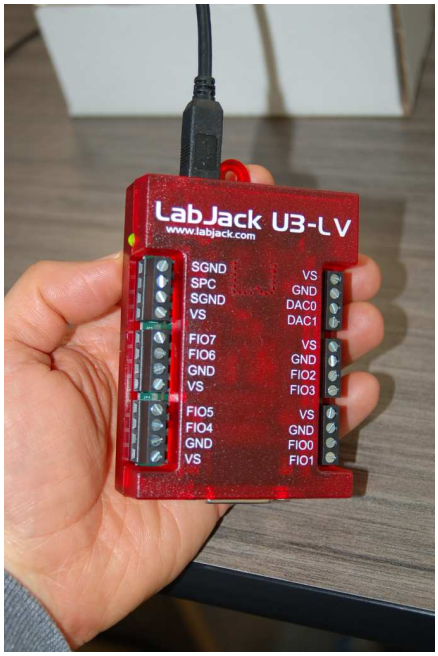
Note how all analog ICs receive power from regulated +5/−5 Volt DC rails, while digital ICs receive unregulated DC power from the batteries (two 9 Volt batteries recommended).

One student's implementation (and extension) of the basic bargraph voltmeter circuit is shown below. As you can see, this meter uses *ten* LEDs to either side of zero as well as the "null" indicating LED, necessitating a greater number of comparators than shown in the previous schematic diagram:



2.3 Example: using Python to control a LabJack model U3 DAQ

A popular manufacturer of low-cost data acquisition (DAQ) hardware is *LabJack*, with their model U3 DAQ being a good entry-level device. This particular model interfaces with a personal computer via a USB cable and is also powered by the computer's USB port 5 Volt DC source:



Several "Flexible I/O" ports (FIO0 through FIO7) are provided which may be configured for either discrete ("digital") or analog input usage. A Python package called `LabJackPython` provides a cross-platform Python programming language library with built-in functions and methods enabling low-level control of LabJack U3, U6, UE9, and U12 DAQ devices. A free software package provided by LabJack called `UD Library Installer` provides multiple applications[1] and software drivers for quick and easy configuration and control of LabJack devices with little or no programming required. The USB drivers contained in LabJack's `UD Library` are necessary in order to use the `LabJackPython` Python package, so you will need to perform *two* software installations: the `UD Library Installer` from LabJack, and the `LabJackPython` using the Python `pip` installation utility operated from the computer's command line.

Follow the instructions from LabJack on how to install this software on your computer before attempting to run the Python examples shown in this section. The Windows command-line instruction I used to install `LabJackPython` package on my computer for these demos is as follows:

```
py -m pip install LabJackPython
```

[1] Among these applications are `Kipling`, `LJControlPanel`, and `LJStreamUD`.

From a Python interpreter shell, you may enter the following commands to control the LED visible on the outside of the U3 unit:

```
>>> import u3
>>> d = u3.U3()
>>> d.toggleLED()
```

The first line (`import u3`) instructs Python to import the U3 device library previously installed on your personal computer. The next line constructs a new object named `d` tied to the first U3 device the computer finds plugged in to its USB port. Note that this simple discovery technique only works when you have *one* LabJack device plugged in at a time! The third and final instruction toggles the binary state of the DAQ's single external LED. Executing that line more than once continues to toggle the LED, first off, then on, then off again, etc. This is a simple and effective test to check that your computer is actually able to communicate with the DAQ via Python commands.

This next test uses the Python interpreter to manually read the analog voltage applied to input FIO0:

```
>>> import u3
>>> d = u3.U3()
>>> d.configIO(FIOAnalog = 0x01)
>>> d.getAIN(0)
```

The `configIO()` method[2] sets the functionality of the FIO inputs using a hexadecimal value whose eight bits relate to FIO0 through FIO7, respectively. In this case we are configuring FIO0 to be an analog input and the rest digital, since `0x01` is equal to a binary value of `0b 0000 0001`. If we had wished to configure the first three FIO inputs to be analog, we would have used the hexadecimal value `0x07` which is binary `0b 0000 0111`. It is important we specify the analog nature of any FIOs before we attempt to read their analog voltage values, otherwise the next instruction will return an error message!

The `getAIN(0)` method then reads the analog voltage applied to FIO0 and returns a floating-point value scaled in actual Volts according to the calibration data stored within the U3 DAQ's memory. Incidentally, if you wish to view this calibration data, you may do so using the following command:

```
>>> d.getCalibrationData()
```

[2] In object-oriented programming languages such as Python, a *method* is a function attached to an object. In this case, the object is `d` which is the particular U3 DAQ we're communicating with, and `configIO()` is one of the methods associated with that object.

An extremely simple Python program using the `toggleLED()` method is shown below, also using Python's `time` library to create a two-second pause in the `while` loop's execution:

```
import u3
import time
d = u3.U3()
while (1):
  d.toggleLED()
  time.sleep(2)
```

Writing this short program using a text editor program, then saving it to a filename ending with the extension `.py` makes it immediately recognizable to the computer's operating system as a Python source code file. Invoking that source file, either by double-clicking on its icon or by specifying it by name at a command-line interface (e.g. `py myfile.py`), will cause Python to execute these instructions.

A simple Python program reading the first three FIO analog voltage inputs (and re-reading them with every press of the "Enter" key on the controlling computer) is shown here ready to save to a file ending in `.py` and executed by double-clicking on the Windows icon for that file:

```
import u3
d = u3.U3()
d.configIO(FIOAnalog = 0x07) # Sets first three FIOs to analog mode

while (1):
  a = d.getAIN(0)  # Reads FIO0 as analog input and stores in "a"
  b = d.getAIN(1)  # Reads FIO0 as analog input and stores in "b"
  c = d.getAIN(2)  # Reads FIO0 as analog input and stores in "c"
  print("FIO0 =", a, "Volts")
  print("FIO1 =", b, "Volts")
  print("FIO2 =", c, "Volts")
  input()  # Pauses until user presses Enter
```

An example of this program's output when run is shown here, pressing the "Enter" key three times to get three readings of the analog inputs:

```
FIO0 = 0.438432256 Volts
FIO1 = 0.37528848000000004 Volts
FIO2 = 0.35026924800000003 Volts

FIO0 = 0.41698720000000006 Volts
FIO1 = 0.36397025600000005 Volts
FIO2 = 0.34490798400000006 Volts

FIO0 = 0.41758289600000004 Volts
FIO1 = 0.36397025600000005 Volts
FIO2 = 0.344312288 Volts
```

A more detailed and sophisticated way to read analog inputs on the model U3 DAQ makes use of the `getFeedback()` method. This method reads the raw analogto-digital converter (ADC) count value as an integer number.

```
import u3
d = u3.U3()
d.configIO(FIOAnalog = 0x07) # Sets first three FIOs to analog mode

while (1):
  a_raw, = d.getFeedback(u3.AIN(0))  # Reads FIO0 counts and stores in "a_raw"
  b_raw, = d.getFeedback(u3.AIN(1))  # Reads FIO1 counts and stores in "b_raw"
  c_raw, = d.getFeedback(u3.AIN(2))  # Reads FIO2 counts and stores in "c_raw"
  print("FIO0 =", a_raw, "counts")
  print("FIO1 =", b_raw, "counts")
  print("FIO2 =", c_raw, "counts")
  input()  # Pauses until user presses Enter
```

Note the comma symbols prior to the = assignment used to send the results of the `getFeedback()` method to the variables `a_raw`, `b_raw`, and `c_raw`. This is because the `getFeedback()` method returns a *list*[3] rather than a single variable. The list in this case happens to contain only one entry, and placing a comma before the "equals" assignment operator extracts that one entry's value as a regular integer value. If not for the commas, variables `a_raw`, `b_raw`, and `c_raw` would all become lists themselves.

An example of this program's output when run is shown here, pressing the "Enter" key three times to get three readings of the analog inputs:

```
FIO0 = 12048 counts
FIO1 = 10064 counts
FIO2 = 9520 counts

FIO0 = 11584 counts
FIO1 = 9808 counts
FIO2 = 9408 counts

FIO0 = 11568 counts
FIO1 = 9904 counts
FIO2 = 9440 counts
```

[3]Lists in Python are similar to arrays or structures in C or C++, namely collections of variables under a single name for easy access.

LabJackPython also supports discrete, or digital, I/O control for any FIO channels in digital mode rather than analog. For example, we may run the following commands manually from a Python interpreter to read the status of FIO6:

```
>>> import u3
>>> d = u3.U3()
>>> d.configIO(FIOAnalog = 0x01)
>>> d.getFeedback(u3.BitDirWrite(6,0))   # Sets the direction of FIO6 as input
>>> d.getFIOState(6)   # Reads the digital logic state of FIO6
```

In this example, the method `configIO(FIOAnalog = 0x01)` setting FIO0 to analog input mode is really just ensuring all the other FIO channels are set to digital. Next we have the `getFeedback(u3.BitDirWrite(6,0))` instruction setting (writing) the direction bit to a 0 value for FIO6 which configures that FIO to be an input rather than an output. The `getFIOState()` method reads the logical state of input FIO6 and displays it as either a 1 or a 0 depending on what the FIO6 input terminal is connected to.

Next, we will examine the Python instructions needed to configure FIO6 as a digital output rather than a digital input, and then setting its output value to be 1 and 0 in turn:

```
>>> import u3
>>> d = u3.U3()
>>> d.configIO(FIOAnalog = 0x01)
>>> d.getFeedback(u3.BitDirWrite(6,1))   # Sets the direction of FIO6 as output
>>> d.setFIOState(6, 1)   # Writes the digital logic state of FIO6 to a 1 value
>>> d.setFIOState(6, 0)   # Writes the digital logic state of FIO6 to a 0 value
```

Note that in all these examples it is not necessary to re-execute the `import u3` or `d = u3.U3()` or `d.configIO(FIOAnalog = 0x01)` instructions so long as they have already been executed during the same Python interpreter session. They are shown in these test examples assuming a fresh start of Python each time.

LabJack model U3 DAQs also contain 32-bit hardware counters which may be read via Python. These counters may be configured to read the digital (high/low) status of any FIO input terminal, incrementing by one every time the logic level transitions from a "high" state to a "low" state (i.e. a *negative-edge* clocked counter). The following Python program demonstrates how to read a counter inside of a model U3 DAQ:

```
import u3
d = u3.U3()

# Enables Counter 0 and sets FIO0-FIO3 to analog mode (0x0F = 0b 0000 1111)
d.configIO(EnableCounter0 = True, FIOAnalog = 0x0F)

while (1):
  count, = d.getFeedback(u3.Counter0(Reset = False))
  print("Counter value =", count)
  count, = d.getFeedback(u3.Counter0(Reset = True))
  input()  # Pauses until user presses Enter
```

Note how the first `getFeedback()` method does not reset the counter, but the second one does. This causes the counter's value to be re-set to zero just prior to the program waiting for the user's "Enter" keystroke. If no negative-edge pulses are detected before pressing "Enter", the next displayed counter value will be zero. However, if the input terminal receives some pulses prior to the "Enter" keystroke then the next value printed to the console after pressing "Enter" will register the number of pulses accumulated.

An example of this program's output when run is shown here, pressing the "Enter" key three times to get display four readings of the counter:

```
Counter value = 0

Counter value = 318

Counter value = 0

Counter value = 771
```

While it would be possible to write a Python program to repeatedly read the digital status of any FIO terminal and increment a Python variable to count how many times that digital signal has pulsed, this strategy would miss pulses whose frequency exceeded the Python program's loop execution rate. The beauty of a hardware counter is that the DAQ keeps its own count of pulses regardless of whether that counter value is read by a computer or not. Thus, when the Python program reads the counter's value, it is sure to read the true value of pulses captured in real time.

The model U3's counter is assigned to any FIO terminal other than FIO0 through FIO3 by means of an *offset* value which defaults to four (4). So, by default, FIO4 is the terminal driving the enabled counter. However, if we wish to assign the counter to sense logic states at a different terminal, all we have to do is change the offset value. The following program shows how this is done within the `configIO()` method, in this particular case setting that offset value to five (5) which makes FIO5 the terminal where any negative pulse edges will increment the counter:

```
import u3
d = u3.U3()

# Enables Counter 0 and sets FIO0-FIO3 to analog mode (0x0F = 0b 0000 1111)
d.configIO(EnableCounter0 = True, TimerCounterPinOffset = 5, FIOAnalog = 0x0F)

while (1):
  count, = d.getFeedback(u3.Counter0(Reset = False))
  print("Counter value =", count)
  count, = d.getFeedback(u3.Counter0(Reset = True))
  input()  # Pauses until user presses Enter
```

LabJack model U3 DAQs support up to *two* hardware counters, called counter 0 and counter 1. In the following program[4] we see both counters being read, with the offset value set to five (5) which means counter 0 is reading the digital status of FIO5 and counter 1 is reading the digital status of FIO6:

```
import u3
d = u3.U3()

# Enables Counters 0 & 1 and sets FIO0-FIO3 to analog mode (0x0F = 0b 0000 1111)
d.configIO(EnableCounter0 = True, EnableCounter1 = True,
           TimerCounterPinOffset = 5, FIOAnalog = 0x0F)

while (1):
  count0, = d.getFeedback(u3.Counter0(Reset = False))
  count1, = d.getFeedback(u3.Counter1(Reset = False))
  print("Counter 0 value =", count0, "  Counter 1 value =", count1)
  count0, = d.getFeedback(u3.Counter0(Reset = True))
  input()  # Pauses until user presses Enter
```

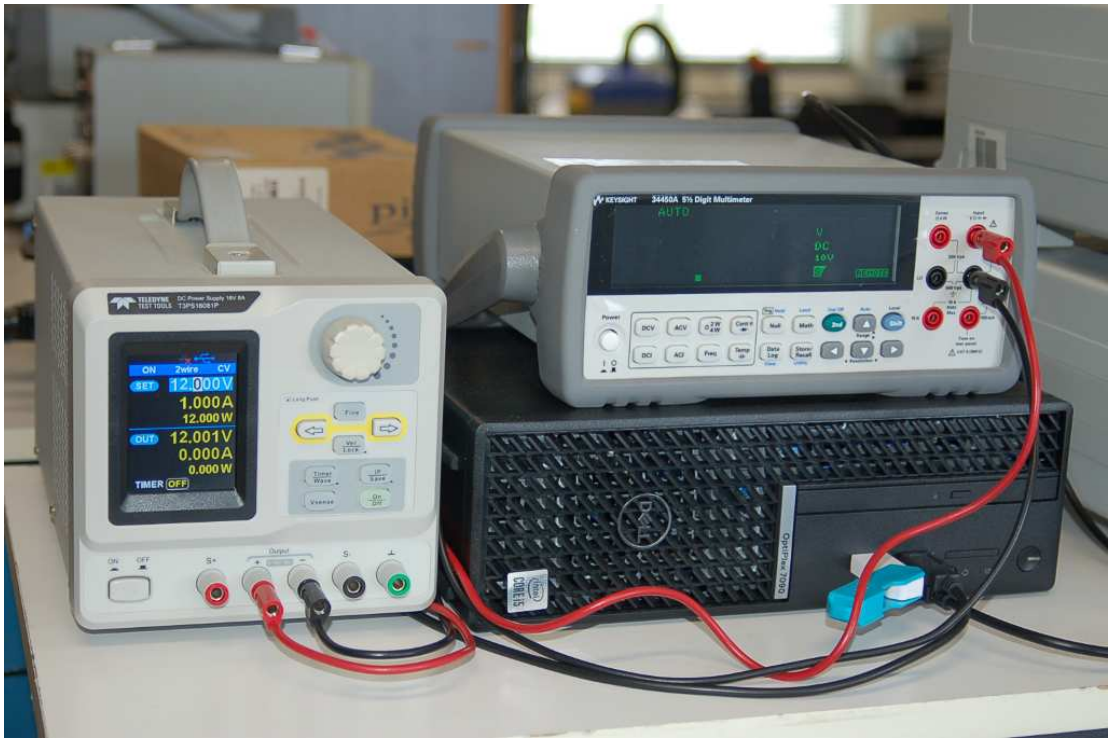
This program resets the value of counter 0 before every "Enter" keystroke from the user, but allows counter 1 to continue accumulating.

[4] Here the `configIO()` method's line is broken into two simply for readability on the printed page of this document.

2.4 Example: SCPI control of digital power supply and multimeter

2.4.1 Introduction and set-up

A common method of automating electronic tests is to use digital electronic test instruments equipped with communications ports and enabled for remote-control using SCPI ("Standard Commands for Programmable Instruments") commands, then using Python or some similar programming language to run "scripts" instructing these instruments to output certain electrical signals and measure the results. In this example we will perform possibly the simplest type of SCPI-based test one could imagine: controlling a DC power supply to output a voltage and then controlling a multimeter to measure that voltage. For this test we used a Teledyne-LeCroy model T3PS16081P digital DC power supply and a Keysight model 34450A digital multimeter, both connected to a personal computer via USB cables, and using the NI-VISA driver library freely available from National Instruments to enable simple Python scripting for SCPI-based remote control:



The following commands[5] were manually entered into a Python-language interpreter in the order shown, the typed commands each being preceded by the >>> Python command prompt, and all responses appearing without that prompt. Normally we would write the Python commands into a file with a name ending in .py and then invoke a Python interpreter to execute that file as a *script*, but here we are entering each command manually at the Python interactive interpreter's prompt just as a proof-of-concept demonstration:

```
>>> import pyvisa
>>> rm = pyvisa.ResourceManager()
>>> rm.list_resources()
('USB0::0xF4EC::0x1410::T0103C22420252::INSTR',
 'USB0::0x2A8D::0xB318::MY59100108::INSTR', 'ASRL1::INSTR')
>>> my_DMM = rm.open_resource('USB0::0x2A8D::0xB318::MY59100108::INSTR')
>>> my_PS = rm.open_resource('USB0::0xF4EC::0x1410::T0103C22420252::INSTR')
>>> my_PS.query_delay = 0.25
>>> my_PS.write_termination = '\n'
>>> print(my_DMM.query('*IDN?'))
Keysight Technologies,34450A,MY59100108,01.02-01.00
>>> print(my_PS.query('*IDN?'))
Siglent Technologies,T3PS16081P,T0103C22420252,2.1.1.8,V1.0
>>> print(my_PS.write('CH1:VOLTage 6.500'))
18
>>> print(my_PS.write('OUTPut CH1,ON'))
14
>>> print(my_DMM.write('CONF:VOLTage:DC'))
14
>>> print(my_DMM.query('READ?'))
6.49895216E+00
```

This proof-of-concept demonstration also serves the practical purpose of obtaining each instrument's resource identifier, which we will need in order to write any automated Python script(s) to perform tests using these devices. If you intend to control instruments via USB using the pyvisa library, these manual tests and queries will be an important first step in your work.

[5] Interestingly, the way that single-quote characters are rendered in the PDF version of this document may not be proper if you attempt to copy-and-paste what you see here on these pages into a Python interpreter or into a text editor to write Python source code. Be sure to use the single-quote character commonly appearing on the same keyboard button as the double-quote character instead!

Next we will explore each of the Python commands used in this example:

- `import pyvisa` – this instructs Python to import the NI-VISA software library so those commands will be available for us to use in this script.

- `rm = pyvisa.ResourceManager()` – this establishes a Python object named `rm`[6] which we'll use to establish communications with the two test instruments.

- `rm.list_resources()` – this instructs Python to list all the "resources" (i.e. test instruments) that the personal computer is able to communicate with. In response, it tells us there are two USB-based instruments ready to communicate, with identifiers `'USB0::0xF4EC::0x1410::T0103C22420252::INSTR'` and `'USB0::0x2A8D::0xB318::MY59100108::INSTR'`. The last of these identification strings before `::INSTR` are the serial numbers of the two test instruments, which means you may anticipate at least some of the identification data simply by examining the physical instruments. The identifier `ASRL1::INSTR` refers to non-existent instrument(s) connected to the computer's RS-232 serial port.

- `my_DMM = rm.open_resource('USB0::0x2A8D::0xB318::MY59100108::INSTR')` – this command instructs Python to open the "resource" (test instrument) having the identifier `USB0::0x2A8D::0xB318::MY59100108::INSTR` which happens to be the Keysight 34450A digital multimeter. Note that the `my_DMM` object name is arbitrary.

- `my_PS = rm.open_resource('USB0::0xF4EC::0x1410::T0103C22420252::INSTR')` – this command instructs Python to open the "resource" (test instrument) having the identifier `USB0::0xF4EC::0x1410::T0103C22420252::INSTR` which happens to be the Teledyne-LeCroy T3PS16081P power supply. Note that the `my_PS` object name is arbitrary.

- `my_PS.query_delay = 0.25` – by default the delay time for a SCPI "query" command was set for 0.0 seconds which was causing time-out errors for the Teledyne-LeCroy power supply. Through experimentation I found that one-quarter of a second was more than sufficient.

- `my_PS.write_termination = '\n'` – by default the line-terminating character sequence for all "write" SCPI commands was `\r\n` which is a "carriage-return" character followed by a "linefeed" character. This was also causing an error, so here we re-initialize this attribute to simply be a linefeed character.

- `print(my_DMM.query('*IDN?'))` – one of the most elementary SCPI commands is a "query" with the string `*IDN?` which prompts the test instrument to report its identity. Here we're querying the Keysight digital multimeter, and it responds in kind.

- `print(my_PS.query('*IDN?'))` – here we do the same for the Teledyne-LeCroy power supply. Interestingly, it responds by identifying itself as a Siglent power supply rather than a Teledyne-LeCroy. This is an important lesson to modern electronics professionals, that often test instruments with different manufacturers' labels on their front panels are actually identical underneath! In this particular case, the equivalent Siglent digital DC power supply costs approximately half that of the Teledyne-LeCroy!!

[6] The name `rm` is arbitrary, whereas `pyvisa.ResourceManager()` is not.

- `print(my_PS.write('CH1:VOLTage 6.500'))` – here we write a SCPI command to set channel 1 of the DC power supply to 6.5 Volts.
- `print(my_PS.write('OUTPut CH1,ON'))` – next, we write a different SCPI command to the power supply, enabling its channel 1 output so that the 6.5 Volts will become available at the output terminals.
- `print(my_DMM.write('CONF:VOLTage:DC'))` – here we write a SCPI command configuring the digital multimeter for DC voltage measurement.
- `print(my_DMM.query('READ?'))` – this SCPI query prompts the digital multimeter to report the results of the prior configuration, in this case the DC voltage read at its input terminals. Here, the displayed result is 6.49895216 Volts, which is very nearly equal to the 6.5 Volt command we gave the DC power supply.

This happens to be my first successful SCPI venture communicating with multiple pieces of test equipment, and I spent an embarrassingly long amount of time getting to this point. Here is a brief summary of what I had to do to get this to work:

- Download and install Python on the personal computer.
- Download and install the National Instruments "Measurement and Automation Explorer" (NI-MAX) as well as the "Virtual Instrument Software Architecture" (NI-VISA) runtime engine on the personal computer. Tested these using the NI-VISA "Interactive Control Application" which worked properly the first time.
- Using Windows `PowerShell` (as Administrator) I then had to run the instruction `py -3.11 -m pip install pyvisa` to install the pyvisa Python library under version 3.11 of Python which is what I installed in the previous step. Interestingly, I discovered that using the Windows command-line instructions `pip install pyvisa` as well as `python -m pip install pyvisa` did not actually work, but `py -3.11 -m pip install pyvisa` did.
- Doing all of this allowed me to query and control a Siglent model SDS1052DL+ oscilloscope and the Keysight 34450A multimeter mentioned earlier, but not the Teledyne-LeCroy T3PS16081P DC power supply mentioned earlier. All this power supply would do is make an audible "beep" in protest as I tried to query it using Python.
- Through a series of experiments I discovered that the Windows command-line instruction `pyvisa-info` is a very helpful diagnostic tool to determine which Python-specific VISA drivers, backends, and other software applications are installed on the computer.
- I also learned that the Python command `dir()` is extremely useful for identifying attributes (i.e. variables) within methods (i.e. functions). For example, issuing the Python command `dir(my_PS)` lists all the attributes found within the method instantiated as `my_PS` in the working example first shown in this Case Tutorial section. This is how I was able to find the `write_termination` attribute which instructs Python how to terminate each SCPI command line, as well as the `query_delay` attribute specifying how long to wait after issuing a SCPI query before declaring a time-out error.
- Finally, learned that the user manual for the Teledyne-LeCroy model T3PS16081P DC power supply is replete with typographical errors in its command summary, in the form of extra "space" characters shown between data fields of the SCPI commands. For example, the SCPI command to turn on the power supply's channel 1 output is shown as `OUTPut CH1, ON` in the manual, but the only way this command would actually work is if I omitted the "space" character between the comma and the letter "O" in "ON" like this: `OUTPut CH1,ON`.

After using a Python interpreter to test communication with the instruments connected to your personal computer, you are ready to write a complete Python script that will automatically control those instruments. Multiple test scripts are shown in the following subsections detailing control of the same power supply and digital multimeter used previously.

Some general tips for writing SCPI test scripts in Python using the VISA library:

- Test the instruments using Python commands typed into a Python interpreter first, to make sure you get the resource identifiers correct and that you are able to issue commands to those instruments without any errors.
- After opening each instrument "resource", reset it using the `*RST` SCPI command. This ensures the instrument will not begin in some unexpected mode prior to the start of your test.
- Close each resource at the end of your script using the VISA `close()` instruction.
- During the development cycle when you are trying different ideas and testing your Python script for proper operation, I recommend invoking your script from the computer's operating system console (command-line interface) window rather than simply double-clicking on the Python icon for your script source file. The reason for this recommendation is that if your script prematurely terminates due to some error, the viewing window you get from just double-clicking on the source file icon will close before you have an opportunity to read the error message(s). If started from a command-line console, however, all error messages will remain in view. To start a Python script from the command line, simply type `py` followed by your script file's name. For example, `py test.py` will invoke the test script named `test.py`.
- Release versions of your Python script should include an `input()` instruction at the end pausing the execution until the user presses the Enter key. This way any data displayed to the user will remain on the screen for viewing rather than disappearing when the window closes at program termination.
- Include the Python `time` library and use `sleep()` instructions following any SCPI commands telling sourcing instruments to output a new signal value, so that the computer will wait a specified time before you instruct any sensing instruments to measure the result(s). Instruments such as DC power supplies may require several seconds of time for their output voltages to settle immediately following a SCPI command to output some new voltage value. If your script doesn't contain these delays, you will be sure to experience measurement errors. For example, the Python instruction `time.sleep(8)` prompts the program to halt for eight seconds.
- Be sure to set all source instruments to zero at the conclusion of your script for physical safety, so that the test circuit is not left in an energized condition after the test concludes.

All of the example scripts shown in the following subsections apply these programming practices. They are all written for the specific Teledyne-LeCroy and Keysight instruments shown earlier, so of course the resource identifier strings will *not* work for your application! Lastly, beware if you attempt to copy-and-paste my Python code to your computer, as sometimes this may result in certain characters being modified, especially single-quote characters!

2.4.2 Python/VISA/SCPI script: user-specified voltages

```
import pyvisa
import time

rm = pyvisa.ResourceManager()
my_PS = rm.open_resource('USB0::0xF4EC::0x1410::T0103C22420252::INSTR')
my_PS.write("*RST")
my_DMM = rm.open_resource('USB0::0x2A8D::0xB318::MY59130016::INSTR')
my_DMM.write("*RST")

my_PS.query_delay = 0.25
my_PS.write_termination = '\n'

print("Enter desired output voltage from power supply: ", end='')
Voltage = float(input())

my_PS.write("CH1:VOLTage "+str(Voltage))
my_PS.write("OUTPut CH1,ON")

time.sleep(2)

Vmeter ,= my_DMM.query_ascii_values("MEASure:VOLT?")
print("The measured voltage is:", Vmeter, "Volts")

Voltage = 0.0
my_PS.write("CH1:VOLTage "+str(Voltage))
my_PS.write("OUTPut CH1,OFF")
print("Turning power supply off")

my_PS.close()
my_DMM.close()

input() # Pauses until user presses Enter
```

Note: the two-second pause provides time for the power supply's voltage to settle before the multimeter measures it. Also note the comma symbol prior to the `=` assigning the DMM query result to `Vmeter`. This is because the `query_ascii_values()` method returns a *list*[7] rather than a single variable. The list in this case happens to contain only one entry, and placing a comma before the "equals" assignment operator extracts that one entry's value as a simple numerical value. If not for the comma, the `Vmeter` variable would become a list itself.

[7]Lists in Python are similar to arrays or structures in C or C++, namely collections of variables under a single name for easy access.

2.4.3 Python/VISA/SCPI script: DC voltage sweep

```
import pyvisa
import time
rm = pyvisa.ResourceManager()
my_PS = rm.open_resource('USB0::0xF4EC::0x1410::T0103C22420252::INSTR')
my_PS.write("*RST")
my_DMM = rm.open_resource('USB0::0x2A8D::0xB318::MY59130016::INSTR')
my_DMM.write("*RST")

my_PS.query_delay = 0.25
my_PS.write_termination = '\n'

# Start with a fully-settled, 0-Volt output from the DC supply:
Voltage = 0.0
my_PS.write("CH1:VOLTage "+str(Voltage))
my_PS.write("OUTPut CH1,ON")
time.sleep(8)
print("Starting voltage sweep")

while (Voltage <= 5.0):
  my_PS.write("CH1:VOLTage "+str(Voltage))
  my_PS.write("OUTPut CH1,ON")
  time.sleep(2)
  Vmeter ,= my_DMM.query_ascii_values("MEASure:VOLT?")
  print("The measured voltage is:", Vmeter, "Volts")
  Voltage = Voltage + 0.5

Voltage = 0.0
my_PS.write("CH1:VOLTage "+str(Voltage))
my_PS.write("OUTPut CH1,OFF")
print("Turning power supply off")

my_PS.close()
my_DMM.close()
input() # Pauses until user presses Enter
```

Here, the two-second time delay allows the power supply's output voltage to settle before querying the digital multimeter for a voltage measurement. Note once again how we must place a comma before the = assignment operator to convert the *list* returned from the `query_ascii_values()` method to a regular numerical value stored in `Vmeter`.

Below we see the text output by this program:

```
Starting voltage sweep
The measured voltage is: -0.00326046485 Volts
The measured voltage is: 0.498196687 Volts
The measured voltage is: 0.998403729 Volts
The measured voltage is: 1.49861744 Volts
The measured voltage is: 1.99897391 Volts
The measured voltage is: 2.49920618 Volts
The measured voltage is: 2.99939074 Volts
The measured voltage is: 3.49966755 Volts
The measured voltage is: 3.99942833 Volts
The measured voltage is: 4.49940736 Volts
The measured voltage is: 4.99909598 Volts
Turning power supply off
```

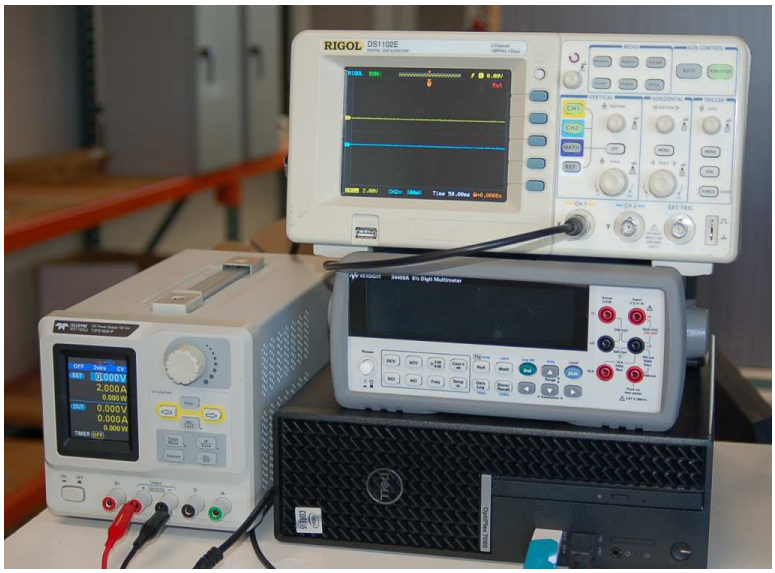
2.5 Example: SCPI control of digital power supply and oscilloscope

2.5.1 Introduction and set-up

In this example we will once again begin with Python commands[8] manually entered into a Python interpreter so we may interrogate the instruments to be used and obtain their resource identifiers:

```
>>> import pyvisa
>>> rm = pyvisa.ResourceManager()
>>> rm.list_resources()
('USB0::0xF4EC::0x1410::T0103C22420252::INSTR',
 'USB0::0x1AB1::0x0588::DS1ET180300904::INSTR', 'ASRL1::INSTR')
>>> my_PS = rm.open_resource('USB0::0xF4EC::0x1410::T0103C22420252::INSTR')
>>> my_OS = rm.open_resource('USB0::0x1AB1::0x0588::DS1ET180300904::INSTR')
>>> print(my_PS.query('*IDN?'))
Siglent Technologies,T3PS16081P,T0103C22420252,2.1.1.8,V1.0
>>> print(my_OS.query('*IDN?'))
Rigol Technologies,DS1102E,DS1ET180300904,00.04.02.01.00
```

In this particular experiment we are using a Teledyne-LeCroy[9] model T3PS16081P digital DC power supply connected to the first input channel of a Rigol model DS1102E digital oscilloscope, both of these instruments plugged into USB ports on a common personal computer with Python installed as well as the National Instruments `pyvisa` library:



[8] Again, please note that single-quote characters copied from the PDF version of this document may not properly paste into your Python interpreter or text editor. Be sure to use the *single-quote* character on your keyboard instead of the odd single-quote character rendered in the PDF!

[9] Actually, a Siglent brand power supply in disguise!

2.5.2 Python/VISA/SCPI script: DC voltage sweep

Next we see an automated test script written in the Python language, designed to sweep the DC power supply's voltage from 0 to 3 Volts in 0.5 Volt increments, pausing four seconds between increments. In addition to printing the supply output and oscilloscope measured voltages in comma-separated variable format, this script also calculates error in terms of percent of the three-Volt span:

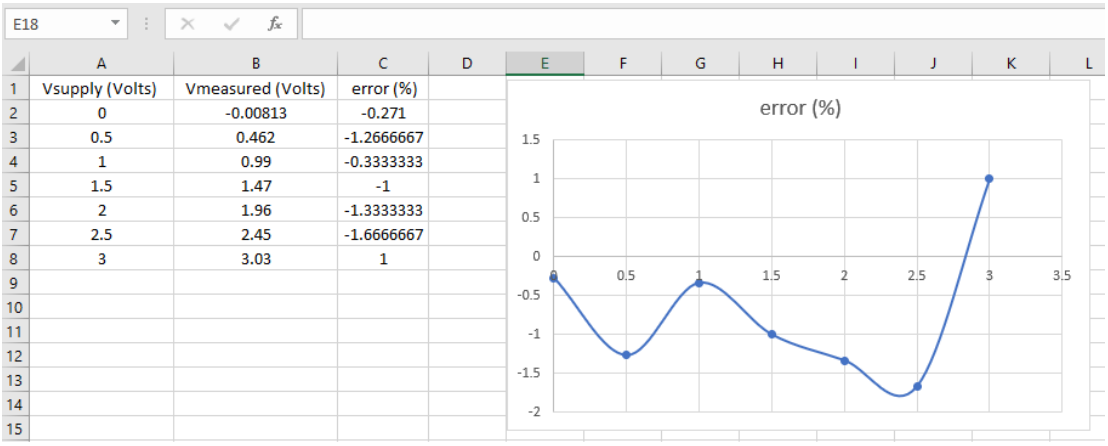
```
import pyvisa
import time
rm = pyvisa.ResourceManager()
my_PS = rm.open_resource('USB0::0xF4EC::0x1410::T0103C22420252::INSTR')
my_PS.write("*RST")
my_OS = rm.open_resource('USB0::0x1AB1::0x0588::DS1ET180300904::INSTR')
my_OS.write("*RST")
my_PS.query_delay = 0.25
my_PS.write_termination = '\n'
my_OS.write(":CHAN1:SCAL 2")
my_OS.write(":TIM:SCAL 0.05")
my_OS.write(":MEAS:TOT ON")
Voltage = 0.0
print("Vsupply (Volts) , Vmeasured (Volts) , error (%)")
while (Voltage <= 3.0):
  my_PS.write("CH1:VOLTage "+str(Voltage))
  my_PS.write("OUTPut CH1,ON")
  time.sleep(4)
  Vscope ,= my_OS.query_ascii_values(":MEASure:VAVerage? CH1")
  print(Voltage , "," , Vscope , "," , (Vscope - Voltage) / 3.0 * 100.0)
  Voltage = Voltage + 0.5
Voltage = 0.0
my_PS.write("CH1:VOLTage "+str(Voltage))
my_PS.write("OUTPut CH1,OFF")
my_PS.close()
my_OS.close()
print("Test complete!")
```

Here, the four-second time delay allows the power supply's output voltage to settle and the oscilloscope to plot multiple sweeps across its display before querying it for an average voltage measurement. Note once again how we must place a comma before the `=` assignment operator to convert the *list* returned from the `query_ascii_values()` method to a regular numerical value stored in `Vscope`.

Below we see the text output by this program:

```
Vsupply (Volts) , Vmeasured (Volts) , error (%)
0.0 , -0.00813 , -0.271
0.5 , 0.462 , -1.266666666666666
1.0 , 0.99 , -0.3333333333333336
1.5 , 1.47 , -1.0000000000000009
2.0 , 1.96 , -1.3333333333333344
2.5 , 2.45 , -1.6666666666666607
3.0 , 3.03 , 0.9999999999999934
Test complete!
```

Of course, the beauty of comma-separated variable data formatting is how easily it imports into graphical visualization tools such as spreadsheets. Below we see this tabular data imported and plotted in Microsoft Excel:



2.6 Example: SCPI control of power supply via GUI

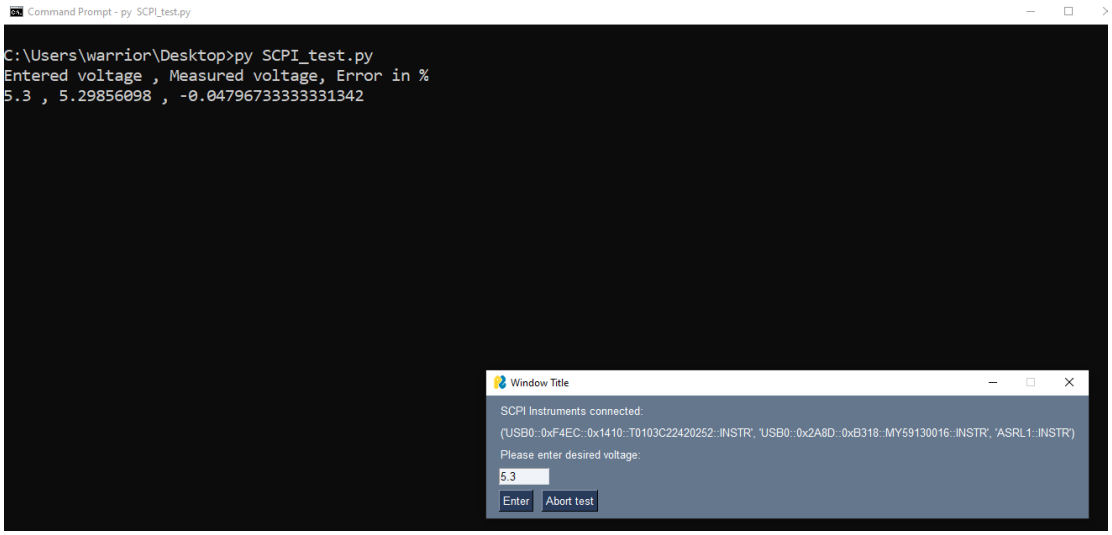
A *Graphic User Interface*, or *GUI* is a modern convenience for software users. When programming in the Python language a very easy-to-use GUI library called *PySimpleGUI* is available at low cost for adding user-interface windows to any Python program.

Using the same programmable DC power supply and multimeter shown earlier (Teledyne-LeCroy model T3PS16081P power supply and a Keysight model 34450A multimeter), connected directly to one another so that the meter should read the power supply's output voltage, we may write a Python script prompting the user to type in the desired voltage and then have the script wait a few seconds before querying the multimeter as to the value of that voltage.

```
import pyvisa
import PySimpleGUI as sg
import time
rm = pyvisa.ResourceManager()
my_PS = rm.open_resource('USB0::0xF4EC::0x1410::T0103C22420252::INSTR')
my_PS.write("*RST")
my_DMM = rm.open_resource('USB0::0x2A8D::0xB318::MY59130016::INSTR')
my_DMM.write("*RST")
my_PS.query_delay = 0.25
my_PS.write_termination = '\n'
Voltage = 0.0
my_PS.write("CH1:VOLTage "+str(Voltage))
my_PS.write("OUTPut CH1,ON")
layout = [ [sg.Text('SCPI Instruments connected:')],
           [sg.Text(rm.list_resources())],
           [sg.Text('Please enter desired voltage:')],
           [sg.In(size=(8,1), key='-VOLTAGE-')],
           [sg.Button('Enter'), sg.Button('Abort test')] ]
window = sg.Window('Window Title', layout)
print("Entered voltage , Measured voltage, Error in %")
while True:
  event, values = window.read()
  if event == sg.WIN_CLOSED or event == 'Abort test':
    break
  Voltage = float(values['-VOLTAGE-'])
  if Voltage < 0.0:
    Voltage = 0.0
  if Voltage > 12.0:
    Voltage = 12.0
  my_PS.write("CH1:VOLTage "+str(Voltage))
  my_PS.write("OUTPut CH1,ON")
  time.sleep(3)
```

```
  Vmeter ,= my_DMM.query_ascii_values("MEASure:VOLT?")
  print(Voltage , "," , Vmeter , "," , (Vmeter - Voltage) / 3.0 * 100.0)
window.close()
Voltage = 0.0
my_PS.write("CH1:VOLTage "+str(Voltage))
my_PS.write("OUTPut CH1,OFF")
my_PS.close()
my_DMM.close()
print("Test complete!")
```

In the following screenshot we can see this program running, the user having entered a desired voltage of 5.3 Volts in the pop-up window provided by `pysimplegui` while the Windows command-line console shows the test results:



Clicking on the "X" button of the GUI window or clicking the "Abort test" button will cause the program's execution to break out of the `while` loop which will close the GUI window, set the power supply to a safe (0 Volt, off) condition, and finally print "Test complete" to the console.

2.7 Example: damage caused by an errant test script

Modern research and development relies heavily on *automated* testing rather than having technicians manually conduct tests on prototype units. One such form of automated testing is to use personal computers running scripts (code) written in the Python programming language to send SCPI-format[10] commands to digital electronic test equipment such as power supplies, signal generators, multimeters, and oscilloscopes. A proper sequence of SCPI commands are able to trigger source devices such as power supplies and signal generators to output voltage and current to devices under test (DUT), and also able to digitally configure measurement devices such as digital multimeters and digital oscilloscopes to capture measurements of voltage and current synchronized in time with the source devices' output changes. These Python-based scripts alleviate a human technician from the drudgery of manually adjusting source output levels and recording the resulting measurements, especially when a large number of devices must be tested and/or a large number of data points must be collected on each device.

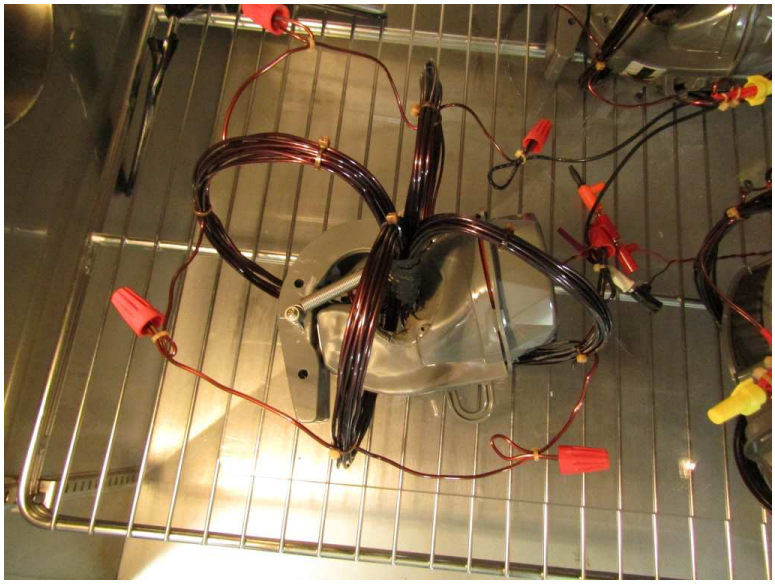
For all its utility and power, however, script-driven automated testing is not without its caveats. One such caveat is the potential for damage should the script be written incorrectly or incompletely. This next example is a cautionary one, showing the kind of damage that may occur with such oversights.

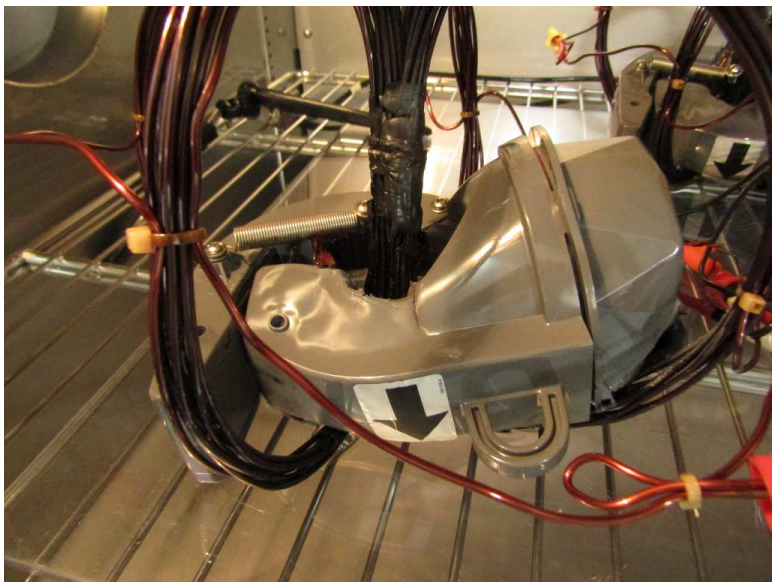
In the case we are about to examine, a technician was conducting automated testing on power-line current-sensing devices intended to be clamped around single high-voltage power conductors. These units were designed to sense current by measuring the magnetic field created around the power line conductor when it conducts electricity, upwards of 1200 Amperes of line current. In order to test these devices without the use of such extraordinarily high test current values, the technician wrapped coils of enameled-copper wire with multiple turns, then threaded four of these coil assemblies through the center of the device under test, wiring all four coils in series so that they shared the same current. For example, if the number of "turns" represented by four of these wire coils numbered 40 in total, only 30 Amperes of test current passing through these 40 turns would be necessary to generate a magnetic field precisely as strong as 1200 Amperes passing through a single conductor threaded through the center of the device under test.

The technician's Python script configured a digital AC current source to generate increasingly higher currents to be send through these wire coils, starting at zero and ultimately incrementing up to the equivalent of 1200 Ampere-turns of magneto-motive force (MMF). This same script also configured measurement instrumentation to sense the unit's response to these varying levels of current, collating all this data into a format easy to survey and assess.

[10] SCPI is a digital communications standard standing for Standard Commands for Programmable Instruments, intended to serve as a manufacturer-independent convention for commands useful in remotely configuring and acquiring data from electronic test equipment. Originally designed to be used with the GPIB (IEEE 488.1) bus network, SCPI is now commonly applied over RS-232 serial, USB serial, and Ethernet networks connecting personal computers to bench test equipment.

Unfortunately, the Python script was not robustly designed. It performed all the current-magnitude steps and measurement steps as it should, but upon termination the script left the AC current source at full output until the technician manually shut it off. This oversight on the part of the script writer may have been tolerable had the test been conducted during the technician's working shift, but unfortunately the technician initiated the automated test shortly before leaving the laboratory to go home. When the technician returned, this distressing sight awaited their arrival:





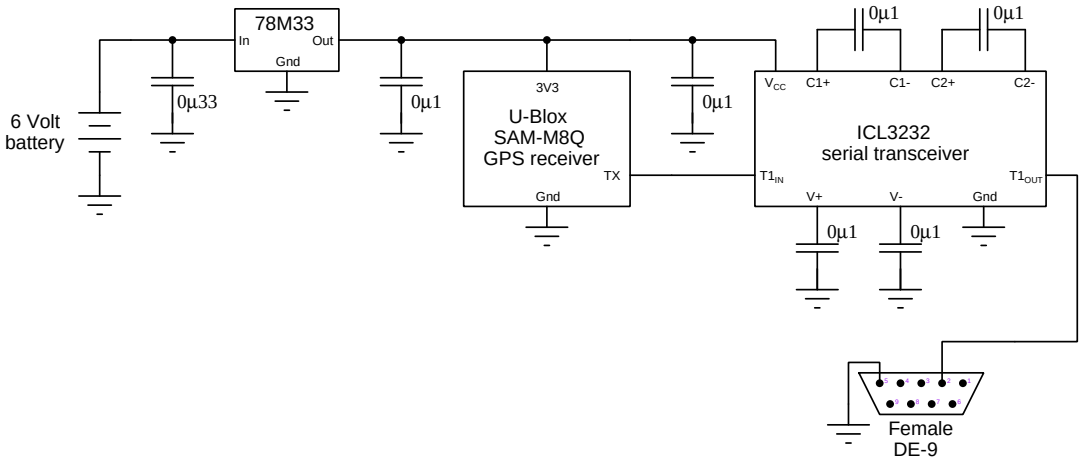
As you can see, the enameled copper wire is discolored from the excessive heat, turned from

a dull red (normal coloring) into a scorched brown. Electrical tape bundling the four wire coils together at the center of the DUT's magnetic aperture is clearly overheated as well. Finally, there is clearly-visible thermal damage to the plastic molding of the DUT.

Fortunately the stainless-steel test chamber prevented the excessive heat from this automated test from scorching or alighting any otherwise-nearby flammable objects, and it sustained no damage itself. The outcome of this poorly-scripted test could have been much worse than it was!

2.8 Example: GPS receiver data stream

An interesting circuit utilizing serial data communication is shown here, where a U-Blox model SAM-M8Q Global Positioning System (GPS) radio receiver module interprets radio signals from satellites in geosynchronous orbit and generates a stream of positioning data sent out of its UART "transmit" port. The GPS receiver is powered by 3.3 Volts DC, which means the serial data signals at its `TX` pin consist of voltage pulses with an approximate range of 0 to 3.3 Volts. This is not a strong enough signal to meet the EIA/TIA receiver standard of ± 3 Volts minimum. In order to properly interface the GPS unit to the serial port of a personal computer, a *transceiver* IC must be used to shift the voltage levels. In this case I'm using a Renesas model ICL3232 which can run as fast as 250 kbps and uses capacitive charge-pump[11] circuitry to boost the +3.3 Volt supply voltage to ± 5.5 Volts:



The female DE-9 connector then attaches to the COM1 port of a personal computer. With suitable serial terminal software[12] running on the computer and configured with the proper serial parameters for the SAM-M8Q module (9600 bps 8N1) what you see on the computer's screen is a burst of ASCII characters at a rate of one burst per second, the characters following a GPS data standard called *NMEA-0183*.

[11] A *charge pump* is a type of DC-DC power converter circuit using capacitors that are switched back and forth between a DC source and a DC load, in such a way as to boost, buck, and/or isolate one voltage level from another. Like most IC charge pump circuits, the ICL3232 contains all the oscillator and transistor switching circuitry inside the IC but leaves the capacitors as external components.

[12] For example, `Kermit`, `Termite`, `minicom`, or other programs designed to receive ASCII characters via the RS-232 serial standard and display those alphanumeric characters on the computer's monitor.

A sample of the GPS positioning data obtained from this circuit is shown here:

```
$GNRMC,183420.00,A,4623.57504,N,11658.41126,W,0.286,,020821,,,A*7E
$GNVTG,,T,,M,0.286,N,0.531,K,A*36
$GNGGA,183420.00,4623.57504,N,11658.41126,W,1,06,2.04,437.7,M,-18.2,M,,*76
$GNGSA,A,3,25,02,12,15,23,13,,,,,,,4.39,2.04,3.89*17
$GNGSA,A,3,,,,,,,,,,,,,4.39,2.04,3.89*16
$GPGSV,4,1,14,02,17,074,09,05,66,092,,11,10,070,,12,08,173,25*76
$GPGSV,4,2,14,13,10,114,10,15,12,145,22,16,03,329,,18,40,264,*7A
$GPGSV,4,3,14,20,37,054,,23,10,207,23,25,30,197,20,26,25,309,*78
$GPGSV,4,4,14,29,85,355,,31,02,271,*7D
$GLGSV,1,1,04,85,51,062,,86,56,323,,87,01,284,,,,19*5B
$GNGLL,4623.57504,N,11658.41126,W,183420.00,A,A*62
```

These characters represent *codes* defined by the NMEA-0183 standard. The first line (beginning with `$GNRMC`) may be interpreted as follows, each of the data fields separated by comma delimiting characters:

- `183420.00` represents the Coordinated Universal Time, read as 18 hours, 34 minutes, 20.00 seconds (the time at Greenwich, London, England).
- `A` represents the data status, with "A" meaning valid and "V" meaning a receiver warning.
- `4623.57504,N,11658.41126,W` represents the latitude and longitude, respectively. In this case we read it as 46 degrees and 23.57504 minutes north, 116 degrees and 58.41126 minutes west.
- `0.286` represents the ground speed in nautical miles per hour (knots).
- `020821` represents the Coordinated Universal Time date field, read as 2 August 2021.
- `A` represents the mode indicator, with "A" meaning autonomous, "N" meaning data invalid, "D" meaning differential mode, and "E" meaning dead reckoning mode.
- `7E` represents the checksum for this block of data, for error-detection purposes.

Chapter 3

Tutorial

3.1 What is metrology?

Metrology is the science of measurement, specifically the internationally coordinated study of standardized measurements around the world. Electrical metrology, naturally, deals with precise measurement of electrical quantities such as voltage, current, resistance, capacitance, inductance, and frequency.

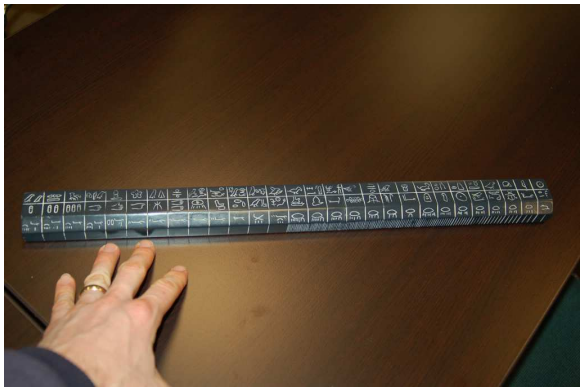
Since the dawn of civilization people have had to agree on what constitutes certain measures of important materials as well as dimensional measurements necessary for the construction of physical objects. The following quotation from the Old Testament criticizes those who seek to take advantage of others in market exchanges by "falsifying the balances":

> *Hear this, o ye that swallow up the needy, even to make the poor of the land to fail, saying When will the new moon be gone, so we may sell corn? and the sabbath, that we may set forth wheat, making the ephah small, and the shekel great, and falsifying the balances by deceit? that we may buy the poor for silver, and the needy for a pair of shoes, and sell the refuse of the wheat?* – **Amos 8:4-6, King James version**

While the importance of accurate, standardized measurements cannot be disputed for cooperation between people, agreement on the measurements themselves is a far less simple matter even in the absence of ill intent.

The *cubit* is an excellent example of an ancient measurement of length, supposedly the distance from an adult male's elbow to the tip of his middle finger. One need only reflect on this definition for a moment to perceive its major weakness: *adult humans do not all possess arms of the exact same length!* Standardizing length measurements by human-arm cubit would mean the same distance measured by different people would result in different values. For crude measurements such a definition may suffice, but certainly not for any application demanding precision.

Archaeologists have unearthed measuring sticks called *cubit rods*, these representing ancient attempts at standardizing length measurements. These wooden rods could be reproduced with much greater accuracy than the variations in arm length between different people, allowing anyone in possession of a standardized cubit rod to measure lengths with much more consistency than if using their own arm. A modern replica of an Egyptian cubit rod appears in the following photograph:

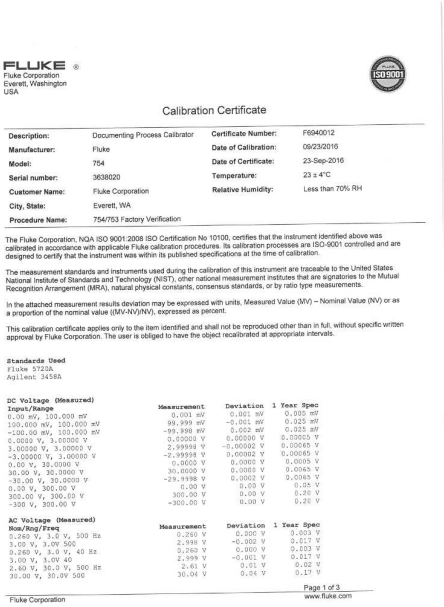


Presumably, all "working" cubit rods were patterned after some "master" cubit rod defining the cubit for that particular population. In a similar manner, all instruments we rely on for precision

measurements should be based on some trusted standard. Every modern nation has its own official bureau maintaining standards for weights and other measures, and for the United States this is the National Institute for Standards and Technology (NIST), formerly known as the National Bureau of Standards (NBS).

If you purchase a measuring rule at a store, you may be reasonably assured that the manufacturer of that rule laid out all the markings along its length at intervals that are not only *precise* (i.e. definite) but also that are *accurate* (i.e. corresponding faithfully to the lengths of the standard millimeter, centimeter, etc.). Somewhere in the factory where these rules are made, there is machinery responsible for placing those markings, and that machinery in turn was aligned against some standard of even greater precision and accuracy. In turn, that calibrating standard must be set (ultimately) to some *primary standard* maintained by the authoritative bureau. This pedigree, or calibration lineage, is referred to as *traceability*: namely, that the accuracy of any instrument's function has been indirectly verified against the highest standard of the nation or of the world.

For high-quality instruments, this traceability actually takes the form of a document declaring its last calibration test results as well as the standard(s) used to calibrate it. Here are photographs of a calibration certificate for a Fluke model 754 Documenting Process Calibrator used as a calibration standard for industrial sensing instruments, that calibrator instrument tested against a Fluke model 5720A Multifunction Calibrator and an Agilent model 3458A 8.5-digit laboratory multimeter:



FLUKE ®
Fluke Corporation
Everett, Washington
USA

Calibration Certificate

Description:	Documenting Process Calibrator	**Certificate Number:**	F6940012
Manufacturer:	Fluke	**Date of Calibration:**	09/23/2016
Model:	754	**Date of Certificate:**	23-Sep-2016
Serial number:	3638020	**Temperature:**	23 ± 4°C
Customer Name:	Fluke Corporation	**Relative Humidity:**	Less than 70% RH
City, State:	Everett, WA		
Procedure Name:	754/753 Factory Verification		

The Fluke Corporation, NQA ISO 9001:2008 ISO Certification No 10100, certifies that the instrument identified above was calibrated in accordance with applicable Fluke calibration procedures. Its calibration processes are ISO-9001 controlled and are designed to certify that the instrument was within its published specifications at the time of calibration.

The measurement standards and instruments used during the calibration of this instrument are traceable to the United States National Institute of Standards and Technology (NIST), other national measurement institutes that are signatories to the Mutual Recognition Arrangement (MRA), natural physical constants, consensus standards, or by ratio type measurements.

In the attached measurement results deviation may be expressed with units, Measured Value (MV) – Nominal Value (NV) or as a proportion of the nominal value ((MV-NV)/NV), expressed as percent.

This calibration certificate applies only to the item identified and shall not be reproduced other than in full, without specific written approval by Fluke Corporation. The user is obliged to have the object recalibrated at appropriate intervals.

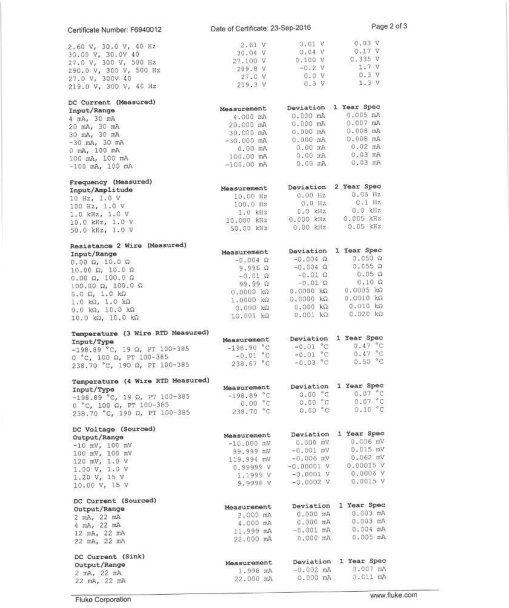
Standards Used
Fluke 5720A
Agilent 3458A

DC Voltage (Measured)

Input/Range	Measurement	Deviation	1 Year Spec
0.00 mV, 100.000 mV	0.001 mV	0.001 mV	0.005 mV
100.000 mV, 100.000 mV	99.999 mV	-0.001 mV	0.025 mV
-100.00 mV, 100.000 mV	-99.998 mV	0.002 mV	0.025 mV
0.00000 V, 3.00000 V	0.00000 V	0.00000 V	0.00005 V
3.00000 V, 3.00000 V	2.99998 V	-0.00002 V	0.00065 V
-3.00000 V, 3.00000 V	-2.99998 V	0.00002 V	0.00065 V
0.00 V, 30.0000 V	0.0000 V	0.0000 V	0.0005 V
30.00 V, 30.0000 V	30.0000 V	0.0000 V	0.0065 V
-30.00 V, 30.0000 V	-29.9998 V	0.0002 V	0.0065 V
0.00 V, 300.00 V	0.00 V	0.00 V	0.05 V
300.00 V, 300.00 V	300.00 V	0.00 V	0.26 V
-300 V, 300.00 V	-300.00 V	0.00 V	0.26 V

AC Voltage (Measured)

Nom/Rng/Freq	Measurement	Deviation	1 Year Spec
0.260 V, 3.0 V, 500 Hz	0.260 V	0.000 V	0.003 V
3.00 V, 3.0V 50Ω	2.998 V	-0.002 V	0.017 V
0.260 V, 3.0 V, 40 Hz	0.260 V	0.000 V	0.003 V
3.00 V, 3.0V 40	2.999 V	-0.001 V	0.017 V
2.60 V, 30.0 V, 500 Hz	2.61 V	0.01 V	0.02 V
30.00 V, 30.0V 50Ω	30.04 V	0.04 V	0.17 V



2.60 V, 30.0 V, 40 Hz	2.61 V	0.01 V	0.03 V
30.00 V, 30.0V 40	30.04 V	0.04 V	0.17 V
27.0 V, 300 V, 500 Hz	27.100 V	0.100 V	0.335 V
290.0 V, 300 V, 500 Hz	289.8 V	-0.2 V	1.7 V
27.0 V, 300V 40	27.0 V	0.0 V	0.3 V
219.0 V, 300 V, 40 Hz	219.3 V	0.3 V	1.3 V

DC Current (Measured)

Input/Range	Measurement	Deviation	1 Year Spec
4 mA, 30 mA	4.000 mA	0.000 mA	0.005 mA
20 mA, 30 mA	20.000 mA	0.000 mA	0.007 mA
30 mA, 30 mA	30.000 mA	0.000 mA	0.008 mA
-30 mA, 30 mA	-30.000 mA	0.000 mA	0.008 mA
0 mA, 100 mA	0.00 mA	0.00 mA	0.02 mA
100 mA, 100 mA	100.00 mA	0.00 mA	0.03 mA
-100 mA, 100 mA	-100.00 mA	0.00 mA	0.03 mA

Frequency (Measured)

Input/Amplitude	Measurement	Deviation	2 Year Spec
10 Hz, 1.0 V	10.00 Hz	0.00 Hz	0.05 Hz
100 Hz, 1.0 V	100.0 Hz	0.0 Hz	0.1 Hz
1.0 kHz, 1.0 V	1.0 kHz	0.0 kHz	0.0 kHz
10.0 kHz, 1.0 V	10.000 kHz	0.000 kHz	0.005 kHz
50.0 kHz, 1.0 V	50.00 kHz	0.00 kHz	0.05 kHz

Resistance 2 Wire (Measured)

Input/Range	Measurement	Deviation	1 Year Spec
0.00 Ω, 10.0 Ω	-0.004 Ω	-0.004 Ω	0.050 Ω
10.00 Ω, 10.0 Ω	9.996 Ω	-0.004 Ω	0.055 Ω
0.00 Ω, 100.0 Ω	-0.01 Ω	-0.01 Ω	0.05 Ω
100.00 Ω, 100.0 Ω	99.99 Ω	-0.01 Ω	0.10 Ω
0.0 Ω, 1.0 kΩ	0.0000 kΩ	0.0000 kΩ	0.0005 kΩ
1.0 kΩ, 1.0 kΩ	1.0000 kΩ	0.0000 kΩ	0.0010 kΩ
0.0 kΩ, 10.0 kΩ	0.000 kΩ	0.000 kΩ	0.010 kΩ
10.0 kΩ, 10.0 kΩ	10.001 kΩ	0.001 kΩ	0.020 kΩ

Temperature (3 Wire RTD Measured)

Input/Type	Measurement	Deviation	1 Year Spec
-198.90 °C, 19 Ω, PT 100-385	-198.90 °C	-0.01 °C	0.47 °C
0 °C, 100 Ω, PT 100-385	-0.01 °C	-0.01 °C	0.47 °C
238.70 °C, 190 Ω, PT 100-385	238.67 °C	-0.03 °C	0.50 °C

Temperature (4 Wire RTD Measured)

Input/Type	Measurement	Deviation	1 Year Spec
-198.89 °C, 19 Ω, PT 100-385	-198.89 °C	0.00 °C	0.07 °C
0 °C, 100 Ω, PT 100-385	0.00 °C	0.00 °C	0.07 °C
238.70 °C, 190 Ω, PT 100-385	238.70 °C	0.00 °C	0.10 °C

DC Voltage (Sourced)

Output/Range	Measurement	Deviation	1 Year Spec
-10.000 mV	-10.000 mV	0.000 mV	0.006 mV
99.999 mV	99.999 mV	+0.001 mV	0.015 mV
119.994 mV	119.994 mV	-0.006 mV	0.062 mV
2.99999 V	2.99999 V	-0.00001 V	0.00015 V
1.1999 V	1.1999 V	-0.0001 V	0.0006 V
9.9998 V	9.9998 V	-0.0002 V	0.0015 V

DC Current (Sourced)

Output/Range	Measurement	Deviation	1 Year Spec
2 mA, 22 mA	2.000 mA	0.000 mA	0.003 mA
4 mA, 22 mA	4.000 mA	0.000 mA	0.003 mA
12 mA, 22 mA	11.999 mA	-0.001 mA	0.004 mA
22 mA, 22 mA	22.000 mA	0.000 mA	0.005 mA

DC Current (Sink)

Output	Measurement	Deviation	1 Year Spec
2 nA, 22 mA	1.998 nA	-0.002 nA	0.007 nA
22 nA, 22 mA	22.000 nA	0.000 nA	0.011 nA

Note how this calibration certificate shows the actual readings given by the Fluke model 754 while being tested by the two superior instruments, as well as the 1-year or 2-year manufacturer's specification for calibration "drift".

3.2 Metrological standards

A *primary standard* is a standardized reference quantity maintained by an authority in weights and measures, usually a national government laboratory generally known as a *National Measurement Institute* or *NMI*, used to calibrate other quantities to be used as *secondary standards* outside of that laboratory. Standards calibrated against secondary standards are called *tertiary standards*.

Two major types of primary standards exist: *artifact standards* and *intrinsic standards*, with the latter of these being much highly favored over the former. An "artifact" standard is some constructed entity supposedly representing the desired quantity with a high degree of accuracy. Artifacts are things made either directly or indirectly by human hands, and whose value fundamentally depends on that craftwork. Thus, it is possible for an artifact standard to become damaged or inaccurate, or to have been made improperly with the same effect. By contrast, an "intrinsic" standard's value depends solely on physical constants, and as such are not liable to circumstantial errors.

A practical example helps make sense of this distinction. For a great many years the base-unit for distance in the metric system was defined by the length of a metal bar held in a French laboratory, the ultimate example of an artifact standard because it was based on a physical object which could have been crafted to be shorter or longer, and which in fact could *grow and shrink* under the influence of temperature. Years later the meter was re-defined in terms of the number of wavelengths of light emitted by atoms of Krypton-86, an example of an intrinsic standard based on the natural properties of Krypton-86 and not upon anyone's skill or circumstance.

The advantages of intrinsic over artifact standards are many, and are also fairly obvious. Any standard based on fundamental properties of the universe are not liable to drift over time, and they are also *reproduceable* by anyone having appropriate knowledge and test apparatus, whereas artifact standards are liable to physical decay and cannot be exactly reproduced but only approximated. That said, most intrinsic standards are currently the domain of well-funded metrology laboratories, priced out of reach of most private individuals and corporations, and so artifacts calibrated against intrinsic standards still have their place as "working standards" at the secondary level and below.

It is worth noting that the trend toward intrinsic standards is not just a practical matter for metrologists, but is also useful for the very definitions of the units comprising the metric system itself. Consider the following paragraphs taken from NIST Special Publication 330[1] wherein all units of measurement for the metric system are defined. The first of these paragraphs briefly outlines the history of scientific standards for units of measurement:

> Experimental methods used for the realization of units and which use equations of physics are known as primary methods. The essential characteristic of a primary method is that it allows a quantity to be measured in a particular unit directly from its definition by using only quantities and constants that themselves do not contain that unit.
>
> Traditionally, a unit for a given quantity was taken to be a particular example of that quantity, which was chosen to provide numerical values of common measurements of a convenient size. Before the rise of modern science, units were necessarily defined in terms of material artifacts, notably the meter and kilogram for length and mass, or the property of a particular object, namely the rotation of the earth for the second. Even at the origin of the metric system at the end of the 18th century it was recognized that a

[1] This is an English translation of the canonical French text *BIPM SI Brochure, 9th edition, published in 2019.*

> more desirable definition of a unit of length for example would be one based on a universal property of nature such as the length of a pendulum beating seconds. Such a definition would be independent of time and place and would in principle be accessible all over the world. At the time, practical considerations resulted in the simpler, artifact definitions for the meter and the kilogram and the second remained linked to the rotation of the Earth. It was only in 1960 that the first non-material definition was adopted, namely the wavelength of a specified optical radiation for the meter.
>
> Since then, definitions of the ampere, kelvin, mole and candela have been adopted that do not refer to material artifacts. In the case of the ampere it refers to a specified electric current required to produce a given electromagnetic force and, in the case of the kelvin, to a particular thermodynamic state, namely the triple point of water. Even the atomic definition of the second was in terms of a specified transition of the cesium atom. The kilogram has always stood out as the one unit that had resisted the transformation from an artifact. The definition that opened the way to real universality was that of the meter in 1983. This definition implied, although it did not state, a fixed numerical value for the speed of light. The definition was worded, however, in the traditional form and stated essentially that the meter was the distance travelled by light in a specified time. In this way it reflected the other definitions of the base units of the SI each of which has the same form, for example "the ampere is the current which . . ." and "the kelvin is a fraction of a specified temperature." Such definitions can be called explicit unit definitions.
>
> [page 107]

The next paragraph sampled from NIST Special Publication 330 summarizes the major change in the 2019 definition of metric units:

> This 9th edition of the SI brochure has been prepared following the adoption by the 26th meeting of the CGPM of a set of far-reaching changes. The meeting introduced a new approach to articulating the definitions of the units in general, and of the seven base units in particular, by fixing the numerical values of seven "defining" constants. Among them are fundamental constants of nature such as the Planck constant and the speed of light, so that the definitions are based on and represent our present understanding of the laws of physics. For the first time, a complete set of definitions is available that does not make reference to any artifact standards, material properties or measurement descriptions. These changes enable the realization of all units with an accuracy that is ultimately limited only by the quantum structure of nature and our technical abilities but not by the definitions themselves. Any valid equation of physics relating the defining constants to a unit can be used to realize the unit, thus creating opportunities for innovation, realization everywhere with increasing accuracy as technology proceeds. Thus, this redefinition marks a significant and historic step forward.
>
> [page xi]

3.3 Measurement terminology

When testing the response of any instrument against a trusted standard, the goal of that test is to see how faithful that instrument is in its measurement function – i.e. does it show the true value of the quantity being measured? Several common terms describe different facets of an instrument's fidelity to its sensed quantity.

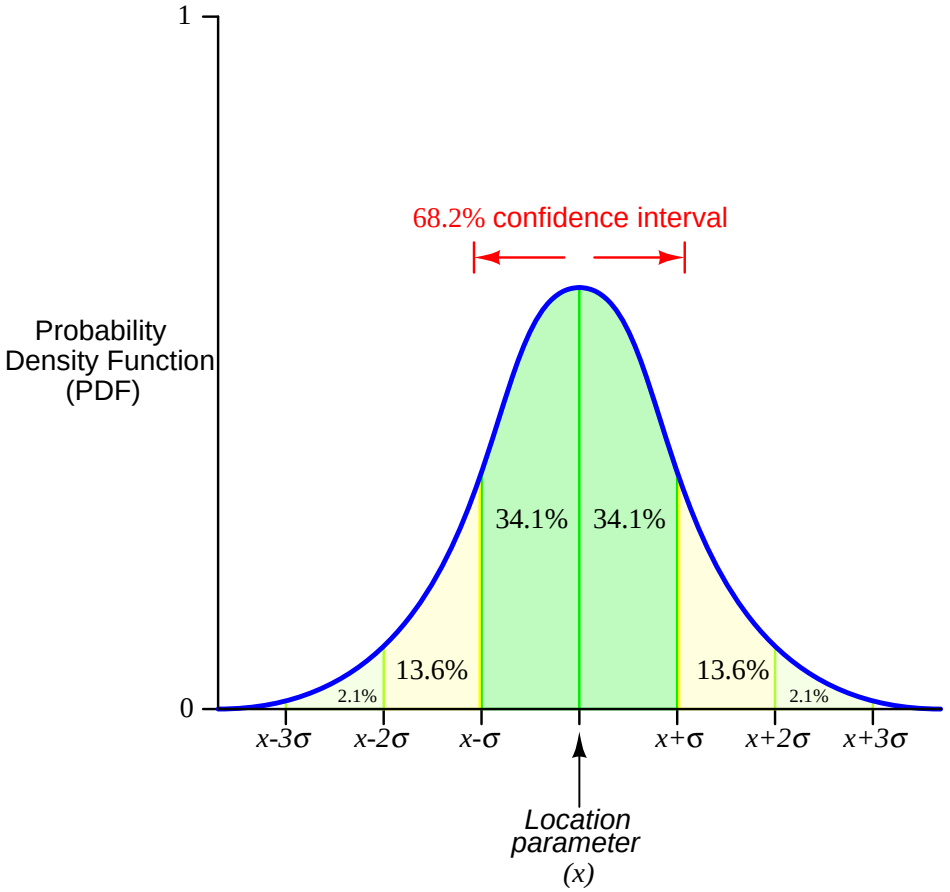
- *Accuracy* is the term we use to describe the instrument's degree of correspondence between its measured value versus the true value.
- *Error* or *Deviation* are the terms we use to describe any difference between an instrument's indicated value versus the true value of the measurement.
- *Repeatability* is the term we use to describe how consistent the instrument's response is to repeated applications of the same stimulus. An accurate instrument must be repeatable, but it's possible for an instrument to be repeatable without actually being accurate!
- *Resolution* is an expression of how finely an instrument is able to express its measurement – i.e. the degree to which it is capable of differentiating between two nearly-identical measured values. Resolution is usually a function of the instrument's design rather than its calibrated condition, for example the number of digits presented on the display of a digital multimeter, or the number of bits for an analog-to-digital converter.
- *Range* is a dual-valued expression of the highest and lowest measurements an instrument is designed to interpret. For example, a zero-centered analog voltmeter with -10 Volts on the far left of its scale and $+10$ Volts on the far right has a measurement range of -10 to $+10$ Volts (± 10 Volts).
- *Span* is the difference between the end-points of an instrument's range. For example, a zero-centered analog voltmeter with -10 Volts on the far left of its scale and $+10$ Volts on the far right has a measurement span of 20 Volts.
- *Calibration* is a process by which an instrument's measurement accuracy is checked against trusted standard(s), and corrected if significant errors are found. For example, checking the accuracy of a vehicle's speedometer and then adjusting it to register more accurately than it did before.
- *Standardization* is a process by which a standard's value is measured against trusted standard(s) and recorded for future reference. For example, checking the accuracy of a vehicle's speedometer and then merely noting how far off its indication is from the true speed so anyone reading it in the future can make the necessary correction(s) mentally.

When working at the level of a metrology laboratory, though, any quantitative notion of "accuracy" and/or "error" becomes quaint because it implies an absolutely certain comparison against some quantity whose value is also known with absolute certainty. In reality, nothing about the physical world is measurable with absolute, quantitative certainty. The inescapable presence of random noise is one reason for this uncertainty, but others exist as well[2], and so you will find metrologists prefer to quantify *measurement uncertainty* and use "accuracy" only in the qualitative sense. That is to say, we can speak of an instrument being "more" or "less" accurate, but any numerical expression of an instrument's performance should be expressed as an uncertainty. We will explore the concept of uncertainty further in the next section.

[2]Sources of uncertainty are often grouped into two categories: *random* and *systematic*. Noise is an example of a random contribution to uncertainty, while phenomena with a predictable effect on measurement (e.g. meter loading, contact resistance) would be classified as systematic. So-called *gross* sources of error such as improper procedures are also potential contributors to uncertainty, but are relatively easy to mitigate through diligence.

3.4 Measurement uncertainty

Real-world measurements are always affected to some degree by noise and other random influences. This means, among other things, that measuring the same quantity multiple times often yields slightly different results, but that the majority of those measurements lie near some central value. Therefore, ultra-precise measurements are always given in terms of both that most-likely value (called the *location parameter*) as well as the *uncertainty* of the distribution, the latter always being a positive quantity. Uncertainty is typically given as the value of one standard deviation (σ) for the bell-shaped "Gaussian" or "normal" curve describing[3] the range of measured values possible given the random influences.



The area encompassed by this curve represents 100% of all possible measurement values given the influences of random factors such as noise. The central location parameter (x) value is that measured value having the highest probability of occurrence. The area encompassed beneath the curve from x to $x + \sigma$ (the domain of one standard deviation from center) covers approximately

[3] Such a *distribution curve* is actually a type of *histogram*, where the height of the curve at any point is proportional to the number of identical measurements at that value on the horizontal axis.

34.1% of the possible measurement values, with another 34.1% of measured values encompassed from x to $x - \sigma$. The total width of that $\pm\sigma$ area represents a *confidence interval* of 68.2%, meaning that there is a 68.2% probability of the true value lying within that range. Conversely, we could say there is a 31.8% probability of the true value lying *outside* that specified range.

A practical example will help illustrate how these terms apply. Suppose we measured the height of all employees working for a particular company, and found all the measured heights fell along this Gaussian curve with an x height of 165 centimeters and a standard deviation of 6 centimeters. The x value of 165 cm simply means that this particular height measurement was the most common among all the employees. In this case, the source of height deviation is largely due to variations between the bodies of different individuals rather than random errors or noise, but the same statistical principles apply. This would mean 34.1% of the people working at this company measured between 165 and 171 centimeters (from x to $x + \sigma$), while another 34.1% measured between 159 and 165 centimeters (from $x - \sigma$ to x).

Applying this to physical constants, we may consult the *CODATA Recommended Values of the Fundamental Physical Constants* (NIST SP 961) published by the National Institute of Standards and Technology in May of 2019 for more examples. For instance, we find the measured mass ratio of one proton to one electron in this document has a value of 1836.15267343 with a standard deviation of 0.00000011. This means there is a 68.2% probability the true mass ratio value lies somewhere between 1836.15267332 and 1836.15267354 (i.e. from $x - \sigma$ to $x + \sigma$).

So far as anyone knows, all protons and all electrons are physically identical, and so the source of deviation must be random instrument errors and noise[4] rather than variations from one particle to another. It is also entirely possible (though unlikely) that the true mass ratio values is either higher than or lower than this $-\sigma$ to $+\sigma$ uncertainty interval.

A common format for expressing a measured value and its uncertainty is to append the significant digits of its uncertainty (over one standard deviation) in parentheses to the central value. In the case of the proton/electron mass ratio, we see it published by the NIST as 1836.15267343(11), and we take those uncertainty digits to apply to the final digits of the base number. In other words, the "11" uncertainty digits add to or subtract from the "43" digits at the end of the base value, representing the $x + \sigma$ and $x - \sigma$ values, respectively.

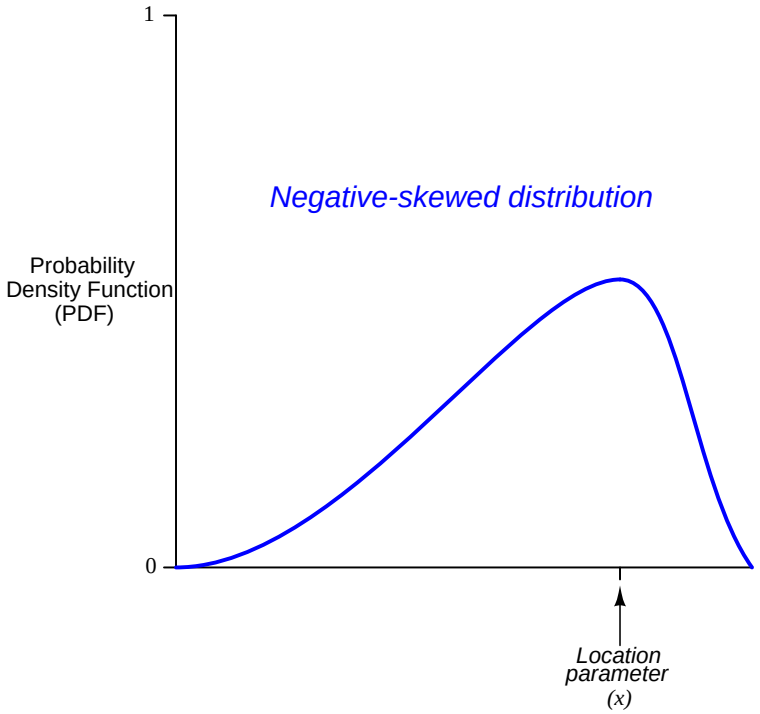
It is also common to find uncertainties specified for confidence intervals other than 68.2% (i.e. $\pm\sigma$). Many scientific studies and engineering specifications, for example, use 90%, 95%, or even 99% confidence intervals when specifying measured values and/or their uncertainties. Higher confidence intervals represent more conservative (i.e. "safer") claims of accuracy. The most important concept to understand here is that measured values and their uncertainties are meaningless quantities unless associated with a confidence interval to express their probability.

Confidence interval and uncertainty scale with one another – that is to say, the greater the confidence interval the larger the amount of uncertainty (i.e. the less certain) that measurement is. For example, the Fluke model 5730A multifunction calibrator has a 1-year DC voltage specification of $\pm$(9 ppm + 0.5 μV) within a 99% confidence interval and $\pm$(7.5 ppm + 0.4 μV) within a 95% confidence interval. The direct correlation of confidence interval and uncertainty strikes many people

[4] This suggests uncertainty values will likely decrease over time as scientific measurement technologies and techniques improve with continued research and development. Uncertainty can never reach zero for empirical measurements, however, due to the unavoidable existence of certain types of noise in the universe.

as counter-intuitive at first, but it makes sense when we think of confidence and precision in other contexts. Consider the case most people have experienced of losing some important personal item such as a comb, a set of keys, a mobile phone, etc., where we can be 100% confident that the item is located somewhere on planet Earth but less and less confident the more specific we get with predicting its location (e.g. 99% confident it's located in our hometown, 95% confident we left it somewhere in our home, 90% confident it's in the pocket of some garment in our closet, etc.). Another way to grasp the relationship between confidence and uncertainty is to consider *limiting cases*. First, how confident are we that the certainty of a measurement is absolutely perfect? The answer to this question must be *we have zero confidence in a perfect measurement; i.e. zero confidence in zero measurement uncertainty*. Second, how confident are we that a measurement is flawed? The answer to this question must be *we have absolute confidence in the measurement being imperfect*.

Probability distributions for real-world measurements do not always follow the symmetrical pattern shown by a Gaussian distribution curve where mean, median, and mode are co-equal. Sometimes the sources of uncertainty are asymmetrical, causing the probability distribution to become "lopsided". The direction and degree of this "lopsidedness" is called *skew*. An example of a probability distribution with a negative skew is shown below:



A good example of a negative-skewed probability distribution for a real-world measurement is the measurement of electrical resistance using a two-wire ohmmeter, in which the major source of uncertainty is extra electrical resistance stemming from the meter's test lead wires, contact resistance between the probe tips and the specimen being measured, and contact resistance between the test lead plugs and the meter's sockets (jacks). Since electrical resistances always *add* when in series

with each other, these unknown resistances must always work to offset the measurement in a positive direction away from the true resistance of the specimen being measured. This means the real value of the specimen resistance is far more likely to be *less* than the measured value than it is to be more than the measured value, skewing the probability distribution in the negative direction. Other sources of error such as electrical noise, meter calibration error, and parallax error (only with analog meters) may influence the measured value either positively or negatively, and the existence of these symmetrical error sources is why this negative-skewed distribution has any area at all lying to the right of the location parameter.

3.5 Proportionate expressions

Uncertainty and error may be represented many ways. One way is to simply state the difference between the measured value and the ideal (standard) value in the same units of measurement. However, it is more common to see uncertainty and error expressed as *proportions* devoid of any unit.

Proportionate uncertainty or error is often expressed using a special kind of fraction with a standardized denominator. A very common form of this is to represent a proportion as some fraction of one hundred, in which case we call it a *percentage*. For example, 5 percent (5%) means a proportion of $\frac{5}{100}$ or 0.05. "Percent" literally means "per cent" or "per hundred".

For applications where the degree of error or uncertainty is vanishingly small, percentages can become awkward. Alternatives to percentage include *parts per million* (ppm), *parts per billion* (ppb), and *parts per trillion* (ppt). These are exactly what they would seem to be: a fraction with the denominator being one million, one billion, or one trillion, respectively. For example, if a DC voltage standard had a long-term drift of ±3 ppm per year, it would mean that its output voltage could be expected to rise or fall by $\frac{3}{1{,}000{,}000}$ or $\frac{3}{10{,}000}\%$ of its nominal value per year.

Percentages and parts-per-x proportions are inherently unitless, though when used to express error or uncertainty for a given unit, sometimes you will see that unit expressed as a fraction where the units mathematically cancel. For example, a precision oscillator with an advertised frequency uncertainty of ±0.15 ppm could have that same uncertainty stated as ±0.15 ppm Hz/Hz, or even ±0.15 μHz/Hz. The Hz/Hz "unit" label clearly indicates this proportionality applies to the device's *frequency* and not some other measure of its performance.

Error is often calculated in terms of the measured value versus the ideal (standard) value, and this is called *percent of value*. A common formula for expressing this error in percent is as follows:

$$\text{Percent of ideal error} = \frac{\text{Measured} - \text{Ideal}}{\text{Ideal}} \times 100\%$$

Modifying this formula to express in parts per million, or parts per billion, or parts per trillion is as simple as modifying the multiplier at the end. For example:

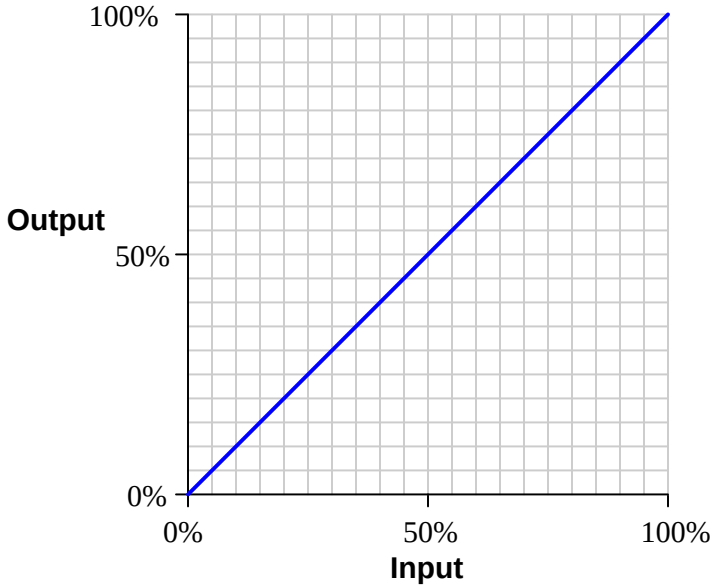
$$\text{ppm of ideal error} = \frac{\text{Measured} - \text{Ideal}}{\text{Ideal}} \times 10^6 \text{ ppm}$$

However, this is not the only valid representation of error. Another common way to quantify error is as a proportion of the instrument's *span* rather than as a proportion of the test-point value. This is particularly useful when the ideal (standard) value happens to be zero, where calculating as a percentage (or ppm, or ppb, etc.) of ideal would result in a divide-by-zero error:

$$\text{Percent of span error} = \frac{\text{Measured} - \text{Ideal}}{\text{Span}} \times 100\%$$

3.6 Typical instrument errors

The purpose of any sensor or measuring instrument is to convert some physical quantity into a corresponding signal or numerical value. Most instruments are designed to do this *linearly*, which means the mathematical relationship between the measured quantity (input) and the converted representation (output) should trace a straight line on a graph:



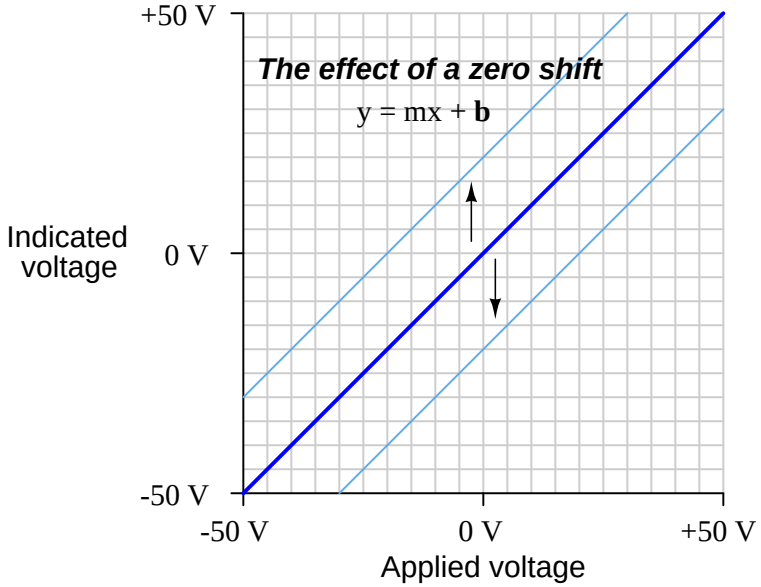
We know that any linear function may be described by the following *slope-intercept* form of linear equation:

$$y = mx + b$$

Where,
y = Output
m = Slope of line (instrument *span*)
x = Input
b = y-intercept of line (instrument *zero*)

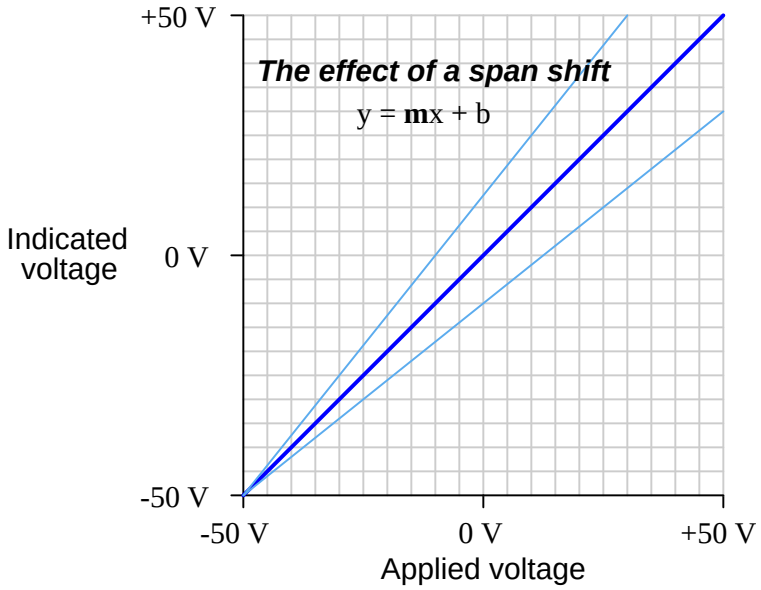
All calibration errors consist of deviations from this perfectly linear response. To illustrate, we will assume a scenario where we check the accuracy of a DC voltmeter over a range of -50 Volts to $+50$ Volts, graphing the indicated voltage versus applied voltage from a trusted voltage standard.

A *zero shift* calibration error shifts the function vertically on the graph, which is equivalent to altering the value of b in the slope-intercept equation. This error affects *all* calibration points equally, creating the same percentage of error across the entire range. Using the same example of a DC voltmeter tested over a range of −50 to +50 Volts:



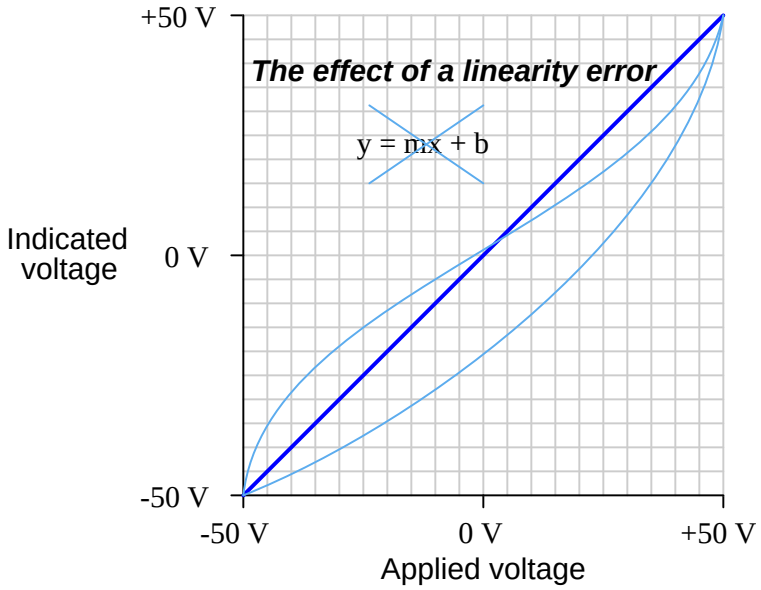
Zero errors in electronic instruments often originate from offset or bias voltages in their circuitry.

A *span shift* calibration error shifts the slope of the function, which is equivalent to altering the value of m in the slope-intercept equation. This error's effect is unequal at different points throughout the range:



Span errors in electronic instruments often originate from incorrect ratios such as what might occur if one resistor's value drifts in a multi-resistor divider network.

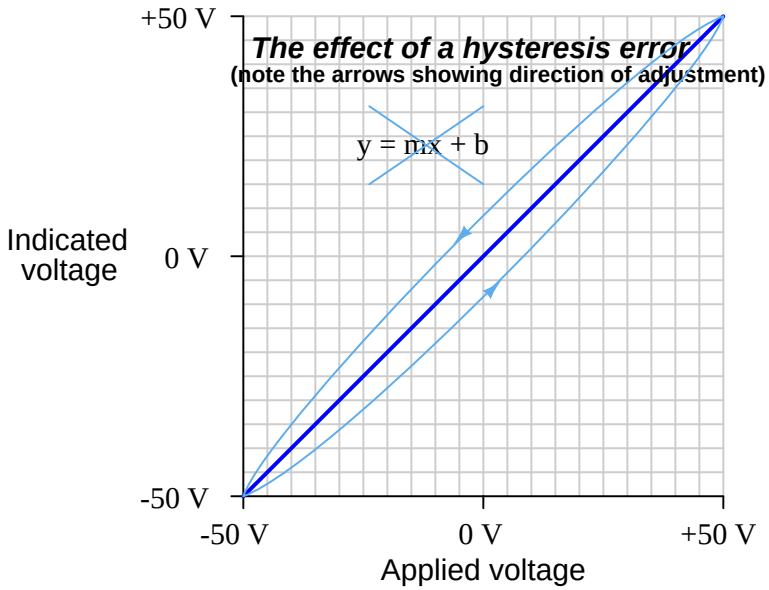
A *linearity* calibration error causes the instrument's response function to no longer be a straight line. This type of error does not directly relate to a shift in either zero (b) or span (m) because the slope-intercept equation only describes straight lines:



Linearity errors in electronic instruments are sometimes caused by component values that do not remain constant as other parameters vary, for example capacitors with a strong voltage coefficient where their capacitance (in Farads) actually varies as a function of applied voltage. Electromagnetic components such as inductors and transformers also exhibit some nonlinearity due to properties of ferromagnetic core materials such as *magnetic saturation*.

A special case of non-linearity possible in digital systems is *non-monotonicity* which is when the output *changes direction* for a given input change, for example a digital-to-analog converter (DAC) circuit where the output voltage increases, then decreases for a step, then continues to increase as the digital input is steadily increased.

A *hysteresis* calibration error occurs when the instrument responds differently to an increasing input compared to a decreasing input. The only way to detect this type of error is to do an *up-down* calibration test, checking for instrument response at the same calibration points going down as going up:



Mechanical instruments tend to suffer from much more severe hysteresis errors than electronic instruments, owing to the friction between the moving parts of a mechanism. An analog voltmeter with a moving pointer suspended on jeweled bearings, for example, may very well exhibit hysteresis when tested in this manner owing to the slight amount of friction in those bearings. However, hysteresis is not unknown in electronic systems, for example the hysteresis seen in solid-state semiconductor voltage references such as Zener diodes when the device's temperature changes and then returns to its original value. Electromagnetic components such as inductors and transformers may also exhibit hysteresis due to *magnetic remanence*[5] of their ferromagnetic core materials.

In practice, most calibration errors are some combination of zero, span, linearity, and hysteresis problems. An important point to remember is that with rare exceptions, zero errors always accompany other types of errors. In other words, it is extremely rare to find an instrument with a span, linearity, or hysteresis error that does not *also* exhibit a zero error. For this reason, technicians often perform a single-point calibration test of an instrument as a qualitative indication of its calibration health. If the instrument performs within specification at that one point, its calibration over the entire range is *probably* good. Conversely, if the instrument fails to meet specification at that one point, it definitely needs to be re-calibrated.

[5] Remanence is the tendency of a ferromagnetic material to retain some of its magnetism after the exciting field has been removed. In other words, some magnetism *remains* after the electric current through the device's winding(s) is stopped.

3.7 Drift and instability

Instrument errors are prone to change with external conditions and/or over long periods of time. For example, the calibration of an instrument may be nearly perfect on one day and then a year later be unacceptably poor. Such accumulation of error over time is often referred to as *drift*. Examples of external conditions potentially affecting an instrument's accuracy include the following:

- Changes in ambient temperature
- Changes in ambient relative humidity
- Changes in atmospheric pressure
- Vibration, including sound
- Changes in power supply voltage or current

Often the degree to which such external variables affect an instrument may be quantified as one or more mathematical *coefficients*. For example, the dependence on ambient temperature for the output voltage of a *Clark cell* (a now-obsolete chemical standard for voltage) followed this equation:

$$V_T = V_{15} - 0.0012(T - 15) - 0.0000062(T - 15)^2$$

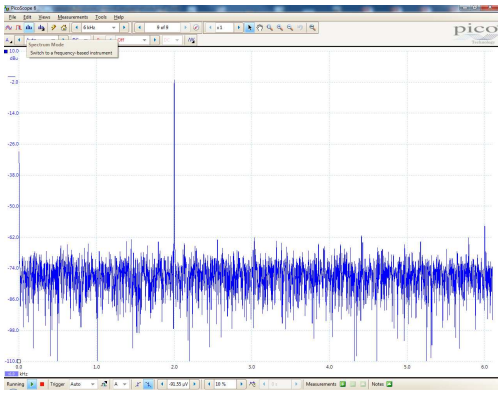
Where,
V_T = Output voltage of the Clark cell at any temperature T between 0 ^{o}C and 28 ^{o}C
V_{15} = Output voltage of the Clark cell at an ambient temperature of 15 ^{o}C

Note how this is a *second-order* mathematical function, with both a linear (-0.0012) and a quadratic (-0.0000062) coefficient. Many error coefficients are specified as simple first-order terms, such as +3 ppm per degree Celsius for the temperature coefficient of a Fluke model 732C solid-state DC voltage reference.

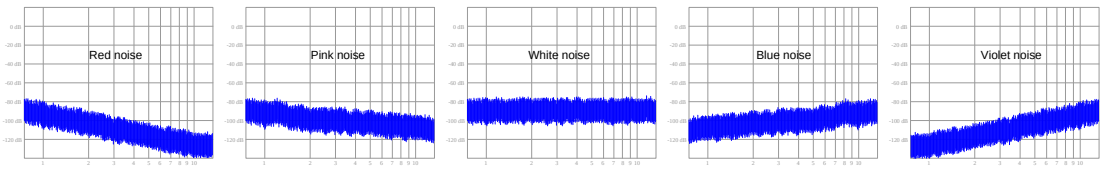
3.8 Intrinsic noise

Although noise may enter a circuit from some outside influence, such as nearby conductors belonging to another circuit, this is not the only source of noise in circuits. A variety of *intrinsic* (i.e. internal) noise sources exist as well. This is where the circuit's own components act to generate noise, even in perfect isolation from anything else. Intrinsic noise tends to be more random than extrinsic noise, and for this reason intrinsic noise often spans a range of frequencies. Recall that any periodic waveform is mathematically equivalent to a series of sinusoidal waves summed together, and that the frequencies of these summed sinusoids could be expressed as whole-number multiples of some fundamental frequency called *harmonics*. Noise that is random, however, is necessarily non-periodic, and so the relationship to definite frequencies is not so simple. Random noise generally manifests as continuous *bands* covering ranges of frequency.

A practical example of this phenomenon is the *noise floor* seen on the display of a spectrum analyzer. In the screenshot shown below, the strong "peak" represents a 2 kHz sinusoidal signal while the "fuzz" spanning the rest of the spectrum represents random noise coexisting with that pure 2 kHz signal:



Note how the noise floor appears to be more or less uniform across the entire width of the spectrum, from 0 kHz to 6 kHz. This is called *white noise* because the noise is equally spread across all frequencies just as white light is a combination of all colors. Not all noise is so evenly distributed across the frequency spectrum, though, and various color names describe noise with uneven spectra. *Red noise* (also called *Brownian noise*) and *pink noise* are both more heavily-weighted at the low-frequency end of the spectrum: that is to say, a "pink" or "red" noise source generates stronger signal variations at low frequency than at high frequency, like white noise after low-pass filtering. *Blue noise* and *violet noise* are both more heavily-weighted at the high-frequency end of the spectrum, like white noise following high-pass filtering. The disinction between pink/red and blue/violet is in the rolloff rate, as shown in the following spectra:



3.8.1 Thermal noise

All matter existing at temperatures greater than absolute zero possess motion on a molecular or atomic scale. In fact, temperature may be thought of in terms of this atomic or molecular motion, with absolute zero being that temperature where all motion ceases. These random motions of atoms and molecules is itself an intrinsic source of noise for electrical and electronic components, because electrical charge carriers are directly affected. As strange as it may seem, even a component such as a resistor which we typically consider a load, actually behaves as a *noise source* at all temperatures other than absolute zero:

The noise output by a resistor is random in all regards: varying in magnitude and polarity with no discernable pattern, all due to the influence of temperature on the electric charge carriers within. Its DC value is zero because all the random variations cancel out over time, but as an AC quantity it exists as an equal amount of voltage (open-circuit) or current (short-circuit) over equal ranges (i.e. bands) of frequency:

$$I_{noise} = \sqrt{\frac{4kT\Delta f}{R}} \qquad V_{noise} = \sqrt{4kT\Delta f R}$$

Where,

I_{noise} = Thermal noise current through short-circuited resistor (AC Amperes RMS)
V_{noise} = Thermal noise voltage across open-circuited resistor (AC Volts RMS)
k = Boltzmann's constant ($1.3806504 \times 10^{-23}$ J / K)
T = Absolute temperature (Kelvin), 273.15 more than degrees Celsius
Δf = Band-width of frequency over which noise is measured (Hertz)
R = Resistor value (Ohms)

All conductors exhibit thermal noise, and this includes the conductive channels of field-effect transistors (FETs).

Thermal noise is also known as *Johnson* noise, or *Johnson-Nyquist* noise after its discoverers John Johnson and Harry Nyquist at Bell Labs in 1926. Since it arises from thermally-motivated charge carriers passing through a resistance, it is fundamental and unavoidable, which is to say *all* resistances exhibit this phenomenon. It is also quite small, as one might conclude by the extremely small value of Boltzmann's constant. For example, the noise voltage produced by a 1000 Ohm resistor at 25 degrees Celcius (298.15 Kelvin) over a bandwidth of 15 kiloHertz (e.g. from 0 to 15 kHz, or from 40 kHz to 55 kHz) is only 0.496975 *micro*Volts.

The "color" of thermal noise is white, being constant in magnitude across all frequencies.

3.8.2 Shot noise

Another form of intrinsic and random noise is *shot noise*. Unlike thermal noise which depends only on temperature and resistance, shot noise is a function of current through any component where charge carriers traverse a gap, and it originates from the fact that electric current is not a continuous substance but rather a passage of discrete electrical charges. A good mental image of shot noise is that of dropping grains of sand onto a hard surface: each grain of sand makes an impact sound when it strikes the surface. Shot noise was investigated in vacuum tubes by Walter Schottky in 1918, arising from the effect of individual electrons reaching the tube's plate one at a time through the gap separating the plate from the tube's cathode. Semiconductor PN junctions (including the junctions of bipolar junction transistors and junction field-effect transistors) also exhibit shot noise, as do photoelectric devices where each incident photon liberates electron-hole pairs.

$$I_{noise} = \sqrt{2eI_{DC}\Delta f}$$

Where,

I_{noise} = Shot noise current through component (AC Amperes RMS)
e = Charge of a single electron ($1.602176487 \times 10^{-19}$ C)
I_{DC} = Constant current through component (DC Amperes)
Δf = Band-width of frequency over which noise is measured (Hertz)

Shot noise is also called *Poisson noise* in honor of its probability distribution. As the equation describes, shot noise is proportional to the intensity of the DC current passing through a component. That is to say, components generate more shot noise as more current passes through them. Shot noise is unrelated to device temperature and therefore is distinctly different from thermal (Johnson) noise which exists independent of circuit current.

The "color" of shot noise is white, being constant in magnitude across all frequencies.

3.8.3 Flicker noise

A type of intrinsic noise found in many types of electrical and electronic devices is *flicker noise*, which grows in proportion to the DC current through the device similar to shot noise. Unlike shot noise, though, flicker noise has a "pink" spectrum which means it is stronger at low frequencies than at high frequencies. The degree to which any device generates flicker noise depends on not only the amount of DC current passing through, but also on the type of device.

It is well known, for example, that carbon-composition resistors generate more flicker noise than either metal-film or metal-wire resistors. The tendency for some types of components to generate more flicker noise than others has led to flicker noise sometimes being referred to as *excess noise*. We know that thermal (Johnson) noise is present in all conductive components and cannot be avoided, but that flicker noise varies with component type. Interestingly, the metal alloy Manganin[6] which is often chosen for high-precision metal-wire resistor construction for its zero temperature coefficient (i.e. its resistivity remains constant over wide ranges of ambient temperature) happens to exhibit zero flicker noise as well.

[6] Manganin is a metal alloy consisting of 84% copper, 12% manganese, and 4% nickel.

3.8.4 Burst noise

A type of intrinsic noise unique to semiconductor devices is *burst noise*, also known as *popcorn noise*. As the name suggests, this type of noise takes the form of discrete steps or bursts in signal strength happening at random times, like the sound of popcorn popping in a hot pan. When this occurs in a device conducting an otherwise steady current, the current will be seen to step up and down at random intervals, making the graphed value look similar to a digital signal alternating[7] between high and low states. The frequency range at which burst noise typically occurs is well within the human audio range – usually 100 Hz or less – and indeed may sound like popcorn popping if heard through a loudspeaker.

Burst noise seems to be primarily the result of defects within the crystalline structure of semiconductor devices, and with improvements in semiconductor manufacturing this type of noise has become less and less of a problem for modern circuit designers.

3.8.5 Avalanche noise

When a PN semiconductor junction is operated with a reverse bias of sufficient voltage, minority charge carriers moving "backwards" through the junction may occasionally gain enough kinetic energy to liberate additional electron/hole pairs when they collide with stationary atoms. These liberated charge carriers then drift with the applied voltage and likewise gain kinetic energy. If the voltage is sufficient, these charge carriers may also liberate more charge carriers when they collide with stationary atoms, the result being an "avalanche" of mobile charge carriers which increases conductivity of the PN junction. The avalanche effect does not result in a constant current, though, but rather pulses of current that constitute noise which we call *avalanche noise*.

The "color" of avalanche noise is white, being constant in magnitude across all frequencies.

3.8.6 Intrinsic noise mitigation

Given the random nature of intrinsic noise, a simple and effective means to mitigate its effects is to *average* many samples of the signal over a period of time, the principle being that the random positive and negative variations of the noise will cancel out to yield a more steady signal. Of course, this incurs the price of having to wait longer periods of time to reliably measure the signal, and it also only works in cases where the signal of interest has a low enough frequency that the averaging process will not substantially corrupt it by filtering out (legitimate) fast rates-of-change.

3.9 Electrical metrology standards

All measured parameters associated with one or more units of measurement must have *standards* defining those units, either directly or indirectly. In this section we will explore the standards used to define important electrical parameters such as voltage, resistance, current, and frequency (and time).

[7] The actual magnitude of these steps will generally be quite small, much less than the steps separating valid digital "high" and "low" states. The point, however, is that burst or popcorn noise lends a randomly "stepped" profile to the graphed quantity.

3.9.1 DC voltage standards

Experimental apparatus for defining the Ohm as a unit of electrical resistance, and the Ampere as a unit of electrical current, came early in the field of electrical metrology. As a result, the Volt required no experimental definition of its own because it could be derived by Ohm's Law ($V = IR$) from precise determinations of the Ohm and of the Ampere. However, as a practical matter of metrology, since voltage is one of the most commonly-measured electrical quantities, a convenient standard for voltage was necessary from early on. It was simply not practical to set up experimental apparatus to independently establish a primary-standard resistance *and* a primary-standard current, then correlate the two to derive a primary-standard voltage every time a primary reference for voltage was needed.

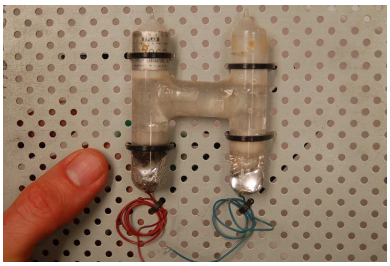
It was known for quite some time that chemical cells could be manufactured using particular combinations of metal electrodes and liquid solutions to produce very stable voltages. These became known as *standard cells* and were used for many decades by both primary and lower-level metrology laboratories as working voltage references, and toward the end of their relevance even inside field instruments. The earliest version was called a *Daniell cell* in honor of the British inventor John Frederic Daniell who made the first of its kind in 1836, using copper and zinc electrodes immersed in a liquid solution of sulfuric acid. Its output voltage was reliably between 1.04 and 1.14 Volts at room temperature, but it suffered from long-term drift and a fairly high temperature coefficient.

Later advances in standard cell technology included the Clark cell (invented 1872) at 1.433 Volts each, the De la Rue cell (invented 1878) at 1.03 Volts, the Helmholtz cell (invented 1882) at 1.00 Volts, the Guoy cell (invented 1888) at 1.390 Volts, and the Weston cell (invented 1892) at 1.01864 Volts. The Weston cell enjoyed the longest service of them all, being found in common use through the 1970's.

A photograph of a Weston-style standard cell appears in the following photograph:

Inside the plastic case of this device there was a glass tube blown in the shape of a letter "H"

containing the metallic electrodes and chemical electrolyte solution necessary for generating the precise voltage:



The "*Caution: do not measure with meter*" label affixed to the case of this standard cell warned the user that the cell could not supply even small amounts of current to any load without suffering error in the form of reduced voltage output. These cells had relatively high internal resistances (typically hundreds of Ohms), and their internal chemistry could be upset for long periods of time if loaded. Therefore, practical use of a standard cell required the application of the *null-balance* voltage measurement technique where negligible current would be drawn during any testing. For more information on this technique, refer to section 3.10.1 beginning on page 86.

Below you can see written on a paper tag affixed to the Weston standard cell its measured voltage, as compared against a trusted voltage standard, on two specific dates. The upper date code is the 24th of July 1974 and the lower date code is some day in 1975 (month value is missing). In 1974 this particular cell tested at 1.01913 Volts and in 1975 it tested at 1.01910 Volts, suggesting it drifted by −0.03 milliVolts over that time span. These dates, of course, suggest just how old this Weston cell is:

Another caveat of these electrochemical cells as voltage standards is their relative intolerance to physical vibration and physical shock. For this reason, they were best left undisturbed as a laboratory standard device.

A ground-breaking improvement in our ability to experimentally reproduce voltage came in 1962 with the prediction of the *Josephson effect*, named after British physicist Brian David Josephson. This is a quantum effect whereby electrons traveling within a superconductor in so-called *Cooper pairs* are, under certain conditions, able to "tunnel" through an insulating barrier separating one superconductor from another, the superconductor-insulator-superconductor sandwich being called a *Josephson junction*. The practical application of this phenomenon to metrology is that there is a quantized relationship between an AC current tunneling through a Josephson junction and a DC voltage across it:

$$V_J = \frac{nhf}{2e}$$

Where,

V_J = DC voltage developed across the Josephson junction, in Volts

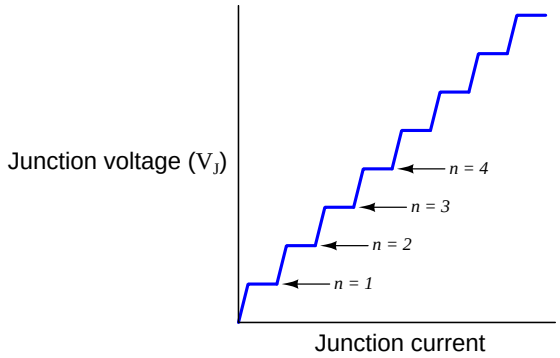
n = Integer (e.g. 1, 2, 3) based on the amount of current tunneling through the Josephson junction

h = Planck constant = $6.62607015 \times 10^{-34}$ joule-seconds (J-s)

f = Frequency of the AC tunneling current, in Hertz

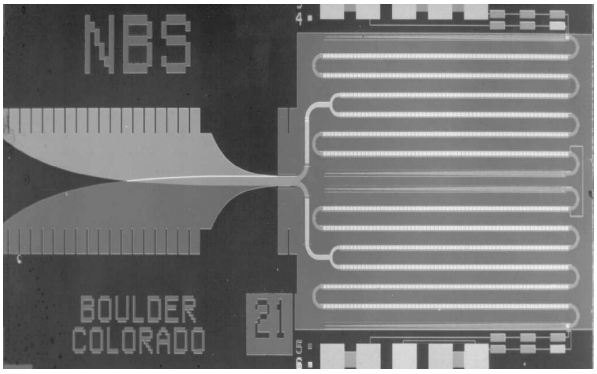
e = Elementary charge of an electron = $1.602176634 \times 10^{-19}$ Coulomb (C)

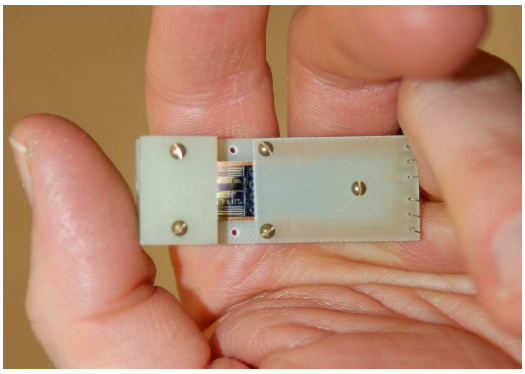
An increase in the AC frequency results in a linearly proportional increase in DC junction voltage, but an increase in the amount of *current* does not. Instead, gradually increasing current results in the voltage "stepping" from one discrete (quantized) value to another. This "stepping" is a manifestation of the junction's quantum behavior and makes the DC voltage independent of current over small ranges in current, the only dependency being the known value of the AC frequency. In other words, a properly biased Josephson junction is an ultra-precise *frequency-to-DC-voltage converter*. The following graph shows this "stepped" voltage generated by a Josephson junction for a fixed AC frequency:



For each of these plateaus on the voltage/current graph, the amount of voltage across the Josephson junction depends solely on the AC excitation frequency (f), Planck constant (h), and the elementary charge of the electron (e), the last two of these of course being immutable constants of nature. The fraction $\frac{2e}{h}$ represents the the proportionality between frequency and DC voltage for a Josephson junction, and is known as the *Josephson constant* K_J with a value of 483,597.8484 gigaHertz per Volt.

The DC voltage output by a single Josephson junction is quite small, typically fractions of a milliVolt. In order to exploit this phenomenon as a practical DC voltage reference we must connect many such junctions in series to form a *Josephson array* generating a larger (and therefore more practically-useful) voltage. A microphotograph (left photo) of a Josephson array built by the National Bureau of Standards (NBS, now known as the National Institute for Science and Technology) consists of 3020 series-connected Josephson junctions used to create a total voltage drop of one Volt, while in the right photo we see a packaged version held in hand:





On the left-hand side of this structure we see two "fin" shaped objects called *finlines* acting as an antenna for microwave radiation which is aimed at the Josephson array to create the AC excitation current. If you look closely at the set of rows on the right-hand side of this device you can see they are comprised of many small squares, those squares representing Josephson junctions, and those rows connected together to form a series network of Josephson junctions. Connections at the top and bottom of this device are where metal wires connect to sense the DC voltage and also to inject bias current necessary to control which integer "step" the array functions at.

Josephson junction technology made possible an intrinsic standard for defining the Volt as a unit of electrical potential, which was made official for the United States in 1990. However, one independent variable that must be precisely controlled in order to obtain accurate voltage from a Josephson junction is the *frequency* of the microwave radiation necessary to excite the junction with AC. Unlike the relationship between voltage and current which is quantized (i.e. the voltage increases in discrete "steps" as current increases), the relationship between voltage and frequency is linear. So, if the microwave radiation frequency is inaccurate, the voltage output by the Josephson junction will likewise be inaccurate.

Fortunately, methods pre-dating the Josephson junction already existed for defining frequency on an intrinsic, quantum basis. These take the form of *atomic clocks* which produce high-frequency signals based on the quantum jumps taken by electrons orbiting certain atoms such as Cesium. Every Josephson junction array must have access to such a precise frequency source in order to be useful at all, which means either the laboratory in question must maintain their own atomic clock or else they must receive a precise-frequency signal (via radio) from atomic clocks maintained by the country's *National Measurement Institute* or *NMI* (e.g. the National Institute of Science and Technology in the United States).

Below we see a photograph of a Josephson junction array maintained by the Fluke corporation in Everett, Washington. The Josephson junction array is buried within the stainless-steel Dewar vessel holding liquid helium coolant (labeled "Non-Magnetic"), while the equipment rack to the left of the Dewar holds the microwave signal and DC bias current apparatus for exciting the array:

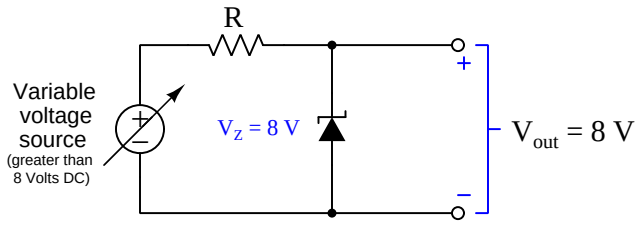
A Josephson junction array such as this is quite expensive. According to the book *Philosophy in Practice* (second edition) published by Fluke, the initial expense of their Josephson Array in 1992 was $85000 US dollars, with another $25000 budgeted for start-up costs. The annual operating cost of the array at that time was approximately $10000, mostly due to the cost of the liquid helium refrigerant necessary to keep the Josephson junction array at a superconducting temperature. This consumable cost does not include the salary of the personnel needed to maintain the system, either. Presumably, a metrology lab of this caliber would employ several engineers and scientists to maintain all standards in top condition and to perform continuing metrological research.

Situated behind this intrinsic voltage standard is a wooden rack holding multiple solid-state voltage references, used as "working standards" periodically calibrated against the Josephson array. These working standard units use internal Zener diodes to maintain stable voltages over time, to *transfer* the Josephson junction array's absolute voltage to other instruments for traceable voltage calibrations.

Electrochemical standard cells such as the Weston cell used to be commonplace in all levels of electrical metrology labs, from primary standards laboratories to industrial workshops, because of their excellent long-term drift characteristics. However, these devices were fragile and affected by vibration, which meant they could not be moved easily without affecting their output voltage. This relative immobility limited their used as field standards, and also complicated efforts to compare working voltage standards against higher-echelon metrology laboratory standards located elsewhere.

The advent of solid-state (semiconductor) devices with voltage-regulating characteristics, most notably Zener diodes, created new possibilities for rugged and transportable DC voltage standards. Zener diodes are PN-junction semiconductor devices specifically designed to "break down" in reverse-bias operation at specific voltage levels, creating a reliably consistent voltage drop. The following example schematic shows a Zener diode with an 8 Volt rating connected to a resistor in such a way as to create a nearly-constant 8 Volt drop despite variations in DC supply voltage:

Zener diode voltage regulator circuit



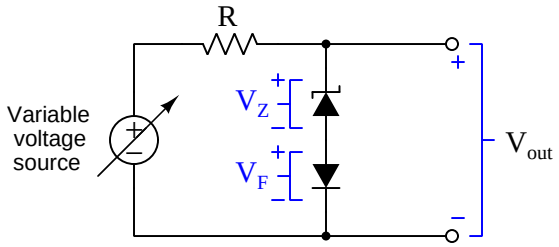
Here, "nearly-constant" is qualified largely by the effects of junction temperature on a Zener diode's V_Z voltage. Zener diodes rated at low voltages tend to have negative temperature coefficients (i.e. V_Z decreases with increasing temperature) while Zener diodes rated at higher voltages tend to have positive temperature coefficients (i.e. greater V_Z with higher temperature). The crossover point for most Zener diodes is somewhere around 6 Volts, and this is due to the fact that not all Zener diodes exploit the same mechanism of breakdown: Zener diodes with ratings less than 6 Volts primarily use *field emission* breakdown while Zener diodes with ratings greater than 6 Volts primarily use *avalanche breakdown*, these two breakdown phenomena exhibiting different temperature coefficients.

Typical temperature coefficients of voltage for discrete Zener diodes are significantly worse than for the legacy Weston standard cells they replace. Whereas saturated Weston cell voltage typically drifted 0.0041 percent per degree Celsius, and unsaturated Weston cell voltage even less at 0.001 percent per degree Celsius, Zener diode breakdown voltage may drift 0.03 percent per degree Celsius or more! For some applications where accuracy is not paramount, the Zener diode's relatively poor temperature coefficient of voltage may still be outweighed by its far superior tolerance to vibration and acceleration, but for demanding metrological applications the bare Zener diode is a step backward[8] from old Weston standard cells.

[8] Another parameter for which chemical standard cells outperformed Zener diodes is *electrical noise*. Shot noise and avalanche noise present in Zener diodes is completely absent from electrochemical cells.

Normal PN junctions operating in forward-bias mode always have negative temperature coefficients of voltage, and so one strategy to obtain better (i.e. closer to zero) temperature coefficients of voltage for Zener diodes is to connect a coefficient-matched pair of avalanche-type Zener and normal (rectifying) diodes together in series. If the two devices' temperature coefficients of voltage are complementary to each other (i.e. equal absolute value but opposite in sign) and the two device junctions are thermally bonded to each other so they always operate at the same temperature, then the over-all temperature coefficient will be nearly zero. As temperature changes, the avalanche-type Zener diode's voltage will increase but the regular forward-biased diode's voltage should decrease *by the same amount*, the result being that combined series voltage drop across the two should remain stable:

Temperature-compensated Zener reference



Another strategy for stabilizing the breakdown voltage of a Zener-based voltage reference circuit is to directly control the semiconducting junction temperature(s). This usually takes the form of a specialized integrated circuit (IC) with a resistive heating element built in to the silicon die along with the PN junction(s), known as *oven stabilization*, the heater providing a controllable junction temperature despite random changes in ambient temperature. Furthermore thermal resistance between the die and the device packaging may be intentionally increased to maximize thermal isolation so that die temperature is less dependent on ambient temperature than in normal ICs.

A premier example of a device exploiting all these strategies (and more!) is the model LTZ1000 integrated circuit, designed in the 1980's by analog engineers working for Linear Technology and still used to this day as a precision DC voltage reference. The LTZ1000 utilizes a 7 Volt "buried" Zener diode structure with a temperature coefficient of voltage of approximately +2 mV/^{o}C along with the base-emitter junction of a BJT having a temperature coefficient of voltage of approximately −2 mV/^{o}C. Direct temperature control of the silicon die helps compensate for the inevitably imperfect cancellation of these two device temperature coefficients, as well as inevitable thermal voltages resulting from dissimilar-material junctions (i.e. the Seebeck effect), so that the over-all device temperature coefficient of voltage can be minimized to less[9] than one part per million. Interesting, the LTZ1000 precision voltage reference IC is not designed to output any particular voltage with absolute accuracy, but rather to *maintain* its output voltage very well over long periods of time. Its datasheet shows approximately ± 0.25 Volt variation possible (e.g. as low as 7.0 Volts and as high as 7.5 Volts for 5 mA of Zener current). The notion here is that one would construct a voltage reference circuit using the LTZ1000, make external component adjustments as necessary to minimize voltage drift, and then compare the circuit's output voltage against a trusted voltage standard and record that measured voltage value for future reference. This act of measuring the output of a stable source

[9] Linear Technology's LTZ1000 datasheet claims 0.05 ppm/^{o}C temperature stability can be achieved!

of unknown quantity against a stable source of known quantity is known as *standardization*. This is different from *calibration* where one device is adjusted in value to match the other. In the case of standardization we simply make an accurate measurement of the previously unknown value based on some other trusted (known) value and then document what the previously unknown value is for future reference.

Semiconductor voltage references should remain powered as continuously as possible in order to maximize stability and accuracy. It is not unheard of for precision semiconductor references to suffer a few parts-per-million change in value simply from being powered down and powered back up! For this reason, solid-state metrology standards are typically left in an energized state for most of their service lives, and even have redundant power sources to maintain energization during inevitable interruptions in utility power service. This same policy applies to *DC voltage transfer standards* when transported from one laboratory to another: these use battery packs to maintain energization throughout travel to and from the other laboratory, so as not to drift unnecessarily.

3.9.2 Resistance standards

Resistance is friction opposing the motion of electric charge carriers through a conductor, usually taking the form of random collisions between moving charge carriers and stationary atoms within the conductive substance. Resistance (R) is expressed in the unit of the *Ohm* (Ω), and by Ohm's Law will permit one Ampere of current to pass through given one Volt of potential difference across its terminals.

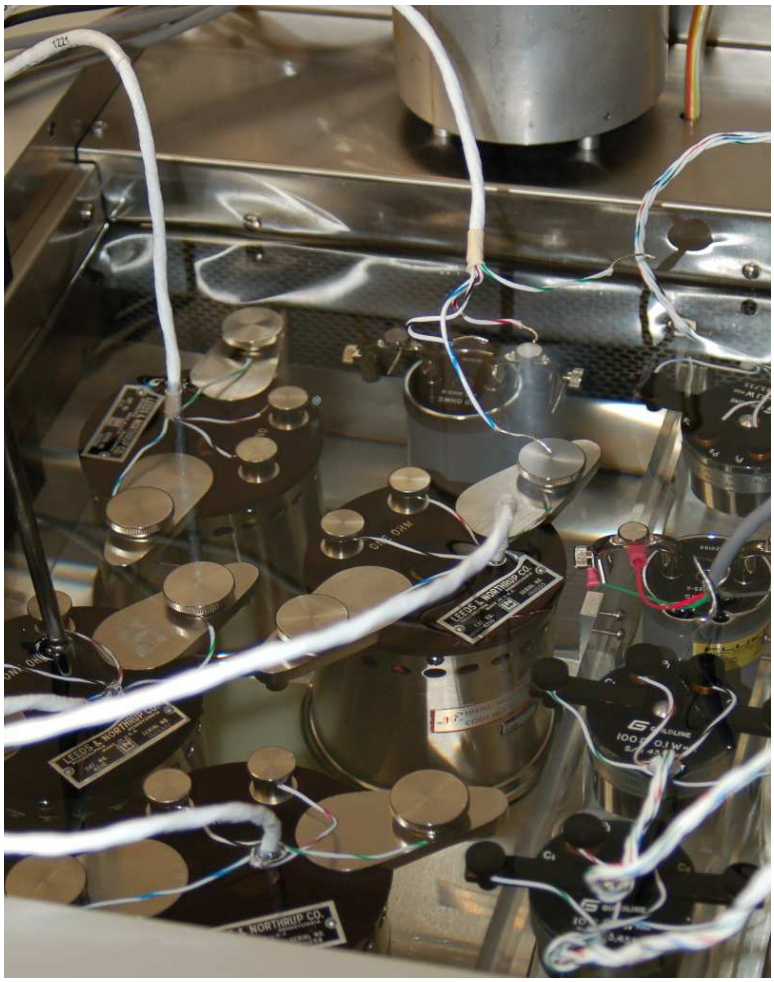
The Ohm was standardized as a unit of electrical resistance in the United States during the year 1893 by the Chicago Electrical Congress as the resistance exhibited by a column of pure mercury metal 1.063 meters in length with a cross-sectional area of 1 square millimeter, at the temperature of freezing water (0 ^{o}C). Later in the year 1894 this standard definition of the Ohm was modified to specify the total *mass* of this mercury column as 14.4521 grams (at the same length of 1.063 meters) rather than specifying the cross-sectional area of the column. In either case, this type of primary standardization was akin to defining the 0 and 100 degree points of the Celsius temperature scale as the freezing and boiling points of pure water at atmospheric pressure: a standard based on the unique properties of a particular substance easily reproduced by any capable laboratory.

A project started in the year 1929 by the National Bureau of Standards (now the National Institute for Standards and Technology) set out to standardize the Ohm in terms of other fundamental units within the metric system rather than in terms of an arbitrary metallic element (mercury). This project, finished in 1938, used a pair of wire coils set up as a transformer (i.e. mutual inductance[10]) with a pulsing current applied to one of the coils, and the average value of the rectified voltage induced across the other coil compared against the voltage dropped by a resistance element in the pulsed primary circuit. The voltage induced in the second coil was a function of how frequently the current changed direction ($V = L\frac{dI}{dt}$), and of course the voltage dropped by the resistor is a simple function of current through it ($V = IR$). When the average of these two voltages were made precisely equal to each other by adjusting the pulsing frequency, the resistance of the resistor could be precisely calculated based on that measured frequency and on the known mutual inductance of the two coils. Since that mutual inductance was a strict function of the coils' physical dimensions, it could be precisely fixed by careful dimensional controls during the construction of those coils. Ultimately, this apparatus tied the definition of the Ohm to the definitions of the meter (length), the second (time), and the magnetic permeability of free space (μ_0) which is a physical constant, rather than to a physical artifact of arbitrary dimensions.

[10] Self-inductance describes the production of an induced voltage in response to a change in current through a conductor. Mutual inductance describes the voltage induced along one conductor in response to a change in current through another conductor, the magnetic field of the current-carrying inductor linking with the first conductor.

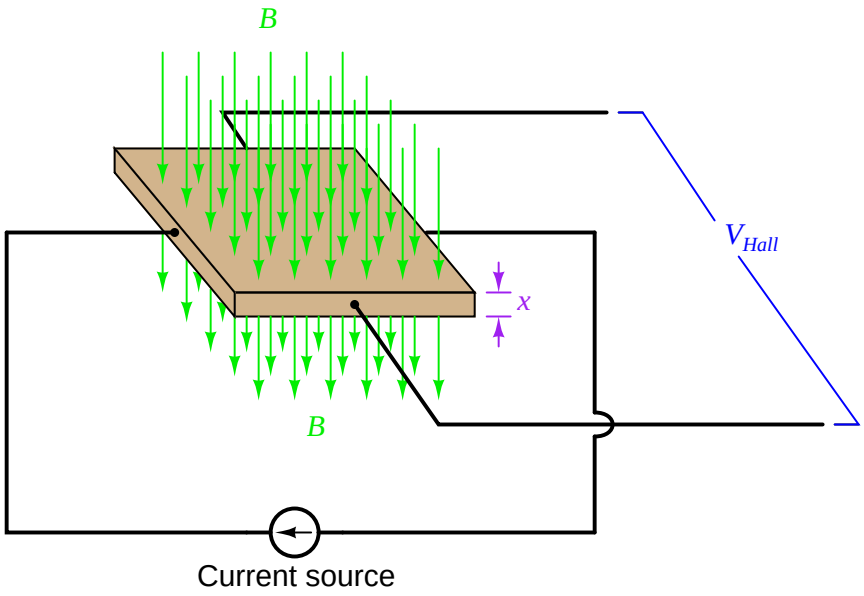
Special standard resistors made of annealed metal-alloy wire coils could be manufactured with great precision, these standard resistors measured by the mutual-inductance method previously described, and then disseminated to metrology laboratories as working standards for use calibrating other resistance-standard devices and instruments. To this date standard resistors are still used as working standards for electrical metrology given their simplicity and relative ruggedness.

Here we see a collection of standard resistors immersed in a heated oil bath at a primary standards laboratory:



In 1989 the decision was made to redefine the Ohm using a completely different method called the *Quantum Hall effect*, discovered by the German physicist Klaus von Klitzing in 1980. This phenomenon occurs as electric charge carriers flow in a two-dimensional formation (having length and width but negligible depth) within a transverse magnetic field at super-low temperatures.

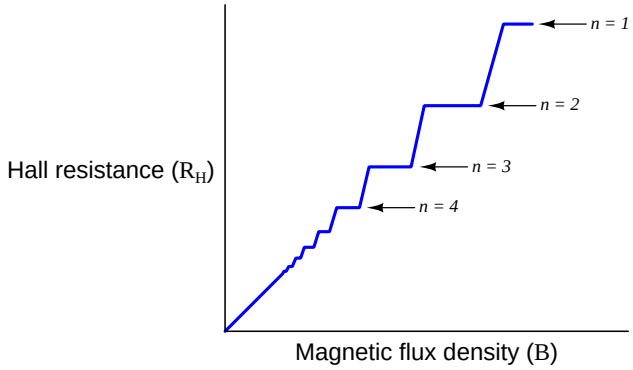
The *Hall effect* is a phenomenon where a current and a perpendicular magnetic field interact to produce a voltage perpendicular to both, as shown in the following illustration:



In normal Hall-effect devices this Hall voltage (V_{Hall}) is a linear function of current (I) and of magnetic flux density (B); that is to say, the density of the magnetic field establishes the ratio between V_{Hall} and I which is functionally equivalent to a resistance ($R = \frac{V}{I}$). The stronger the magnetic field, the more Hall voltage you get for any given amount of current, i.e. Hall resistance increase as the magnetic flux density increases. However, if the x dimension is reduced to a bare minimum and the material's temperature reduced to the point of superconductivity, this Hall resistance becomes *quantized* which means it "jumps" from one fixed value to the next as magnetic flux density (B) increases. These fixed Hall resistance values are based not on the dimensions of the Hall effect device itself, but rather on the value of two immutable constants of nature: the *Planck constant* (h) and the elementary charge of an electron (e)[11].

[11] Specifically, quantized Hall resistance is equal to $\frac{h}{e^2}$ which is 25812.807 Ohms.

Graphing Hall resistance as a function of the applied magnetic flux density, we see a "stepped" function not unlike that of the Josephson effect, except instead of voltage being a quantized function of current we see resistance (i.e. the ratio between Hall voltage and the DC excitation current) being a quantized function of magnetic flux density:



$$R_H = \frac{h}{e^2 n}$$

Where,

R_H = Hall resistance, in Ohms

h = Planck constant = $6.62607015 \times 10^{-34}$ joule-seconds (J-s)

e = Elementary charge of an electron = $1.602176634 \times 10^{-19}$ Coulomb (C)

n = Integer (e.g. 1, 2, 3) based on the amount of magnetic flux density (B) applied to the Hall effect device

Whether the magnetic flux density is manually adjusted or automatically regulated by a control system, it is easy to tell when any of these plateaus are reached, and in those states the quantum Hall effect device acts as an ultra-precision resistance. The fraction $\frac{h}{e^2}$ represents the the proportionality between current and voltage for a quantum Hall effect device, and is known as the *von Klitzing constant* R_K with a value of 25,812.80745 Ohms.

The discovery of the Quantum Hall effect provided a means to make the definition of Ohm *intrinsic*, and this became the officially recognized standard of the Ohm in 1989. According to the NIST, the standard deviation for uncertainty in this definition of the Ohm is approximately 0.2 parts per million which is far better than any previous method of defining the Ohm.

3.9.3 DC current standards

Current is the average rate of drift of electric charge carriers over a distance. Quantities of electric charge carriers are measured in the unit of the *Coulomb* (C), and the number of Coulombs of charge carriers drifting by any given location per second is the current (I) expressed in the unit of the *Ampere* (A). One Ampere of current (i.e. one Coulomb of charge motion per second) is what Ohm's Law predicts as the result of impressing one Volt of potential difference across the terminals of a one Ohm resistance.

One Ampere was originally defined by the British Association in the year 1892 as the amount of direct current (DC) that would deposit metallic silver at a rate of 0.001118 grams per second on an electrode immersed in a solution of silver nitrate. Interestingly, electrochemical cell apparatus such as the one used to define the Ampere were used to measure electricity consumption by customers of early DC electric power systems, the plated electrodes being periodically replaced and measured for the weight of the deposited metal, and then the power bill based on how much metal got deposited on the electrode during the time period it was in service!

In 1934 a different method of defining the Ampere was developed, based on the magnetic force generated between coils of current-carrying wire. Like the mutual-inductance method for standardizing the Ohm, this electromagnetic method similarly anchored the definition of the Ampere on standardized definitions of the meter (length) and the permeability of free space. In fact, this electromagnetic definition of the Ampere held for many years after that, being the amount of electric current that would produce a mutually-attractive force of 2×10^{-7} Newtons between two parallel wires of infinite length spaced one meter apart.

Presently, the Ampere is experimentally defined by the Volt standard (Josephson effect) and the Ohm standard (Quantum Hall effect) by means of Ohm's Law ($I = \frac{V}{R}$), although its official definition is based on a fixed number of charge carriers drifting by a given location per second of time, specifically $6.241509074 \times 10^{18}$ electrons per second.

3.9.4 Frequency and time standards

Accurate measurement of frequency and time is important for a great many applications. Just a few are listed here:

- Time is one of the fundamental measures of the SI metric system, codified in the unit of the *second*. Without accurate standards for time, we would be unable to accurately and consistently measure anything related to time, such as velocity, acceleration, growth and decay rates, and of course frequency which is the reciprocal of period (cycle time).
- Frequency is an important parameter in radio communication systems, since different "channels" of radio communication are often differentiated from one another by having different frequencies. We need stable frequency references to calibrate the oscillators used to build radio transmitters, and for accurately characterizing the filter networks and other frequency-selecting circuits necessary for radio receivers.
- The *Global Positioning Satellite* (GPS) network works by having receivers compare the delay times of radio signals broadcast by satellites in geosynchronous orbit around the earth, those slight differences in propagation time allowing the receiver to triangulate its location. Extreme precision in time measurement is necessary here, in order that the derived position measurement have any useful accuracy and resolution.
- Time signals received from GPS satellites are also useful for time-stamping any measured event anywhere on Earth, for example transient and fault events in electric power grids or pipeline systems. The time-stamped data may then be aligned and compared against similarly time-stamped data from any other location in order to discern causes and effects in large systems.
- For use in other intrinsic standards, such as the Josephson effect used to define the Volt.

Scientific measurement of time was originally defined in terms of planetary motion. In the case of the Earth's rotation defining the unit of the *day*, one could divide the amount of time of one day into 24 hours, then into 60 minutes per each of those hours, and then 60 seconds per each of those minutes. This calculation would define the second as $\frac{1}{24 \times 60 \times 60} = \frac{1}{86400}$ of one day. Other celestial object motions have been used to measure the passage of time as well, including the period of Earth's moon, the relative angle between Earth's axis and the sun (forming the basis of annual *seasons*), and the apparent motion of stars as viewed from Earth.

Defining units of time measurement by the motion of planets, moons, and/or stars has many disadvantages where metrological uncertainties are concerned. All of these motions change slightly over time. Even the Earth's rotation is not steady, thanks to its molten core, its oceanic mass, and the orbit of its moon.

Astronomical observations, useful though they may be as universal standards for time, become impractical as standards for measuring very brief intervals of time such as seconds, milliseconds, microseconds, and so on. The earliest metrological *working standards* for time and frequency used mechanical pendula[12] to produce repeatable oscillations. Later, electronic oscillator circuits

[12] "Pendula" is the plural of *pendulum*, a mass swinging from an elevated suspension pivot. The natural resonant frequency of a pendulum is directly proportional to gravitational force and inversely proportional to the square root of the length of the pendulum, and is independent of mass and of the pendulum's swing magnitude.

stabilized by the resonant frequency of quartz crystals replaced mechanical pendula as working time and frequency standards. Quartz-crystal oscillator circuits are extensively used in modern consumer and commercial electronic systems as local frequency references, and in metrology laboratories as secondary standards.

Fundamental definitions of time based on astronomical observations – in one form or another – held until the year 1967 when the decision was made[13] to internationally define the second in terms of radiation emitted by cesium-133 atoms. Specifically, the 1967 SI standard for one second of time was as follows:

> *The second is the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium 133 atom.*

What this cryptic sentence means is that if you take an atom of cesium-133 and "excite" it with sufficient energy to cause one of its electrons to "jump" from its unexcited ("ground") energy level to the next energy level above that, when the electron returns to its original energy level it must give up that invested energy (because energy is always conserved!), and it does so in the form of an electromagnetic wave having a frequency of precisely 9,192,631,770 Hertz. The reason this emitted wave's frequency is so exact is a function of the electron's quantum states within the cesium atom: energy levels of atomic electrons are *quantized*, which means they are unable to assume arbitrary energy levels but must instead conform to the limited energy-level allotments associated with shells and subshells around that atom. This standardization of the unit of the second held remarkably unchanged through 2019, when it was re-stated thusly:

> *The unperturbed ground state hyperfine transition frequency of the cesium-133 atom (Cs) is 9,192,631,770 Hz.*

Since frequency expressed in the unit of Hertz (Hz) is literally cycles per second, defining the radiation emitted by an atom of cesium-133 as having a frequency of 9,192,631,770 Hz is really the same thing as saying that one second of time is equivalent to that many periods of the same radiated wave.

Furthermore, these energy levels and their associated frequencies are an immutable property of cesium-133 atoms, and because cesium may be chemically separated and purified for laboratory use anywhere in the world, it was an intrinsic standard reproduceable anywhere given the requisite equipment sophistication and technical expertise. This fundamental concept is the basis of the so-called *atomic clock*, first prototyped in 1948 by the National Bureau of Standards (NBS, later re-named the National Institute of Standards and Technology, or NIST) in the United States in 1948, using an ammonia-based *maser*[14].

[13] Atomic clock technology was known before 1967, and in fact this definition for the second was publicly proposed in 1958. 1967 simply marks the *adoption* of the cesium-based standard into the official metric system.

[14] A "maser" is a device similar in function to a *laser*, except that instead of emitting light it emits microwave radiation. In both devices, a sample of material is subjected to external excitation, and is arranged in such a fashion that electromagnetic waves emitted by those excited atoms' electrons falling back to lower energy levels synchronize with each other to form *coherent* bursts of radiation.

From 1959 through 1998 the NBS/NIST refined atomic clock technology for greater measurement accuracy and practicality, creating eight different models of functional atomic clock useful as primary frequency/time standards: NBS-1, NBS-2, NBS-3, NBS-4, NBS-5, NBS-6, NIST-7, and NIST-F1, the latter two clocks' name reflecting the name change of the organization from NBS to NIST in 1988. The NIST-F1 clock completed in 1998 uses a "fountain" of cesium atoms ejected upward which gradually slow and then fall back down under the influence of gravity rather than a "beam" of fast-moving cesium atoms as used in earlier atomic clocks, the additional residence time those atoms spend in the atomic "fountain" making it easier to lock into those atoms' emission frequencies and therefore possible to measure those frequencies with greater precision.

Latest developments in intrinsic frequency standards include the NIST-F2, which is the NIST's next-generation atomic clock, as well as so-called *optical clocks* based on atomic emissions of *light* rather than radio-wave emissions. Since light frequencies are far greater than radio-wave frequencies ("RF") – i.e. light waves have far more cycles per second than radio waves – optical clocks promise the ability to measure time with far greater resolution than traditional atomic clocks.

It is interesting to note that time and frequency – which are essentially non-physical in nature – were the first primary standards to be based on quantum phenomenon, and at the time of this writing (2023) remain the least uncertain of all the measurement standards.

Atomic clocks based on rubidium metal vapor rather than cesium cost far less to manufacture, although they suffer from greater frequency drift caused by changes in gas pressure and in gas temperature inside the resonance cells containing the rubidium. At the time of this writing (2023), one can purchase a subminiature[15] rubidium atomic oscillator with an advertised output frequency of 10 MHz and a total frequency stability of ± 1.50 ppb for approximately $2200 US apiece!

For anyone requiring an ultra-accurate frequency standard but lacking the resources to build or purchase an atomic clock for themselves, some excellent alternatives exist. The first and oldest of these is to tune a shortwave radio to the time signals broadcast by the NIST in the United States over its radio transmitters WWV, WWVB, and WWVH; the first two located in Fort Collins, Colorado near the primary standard atomic clock frequency source and the latter located in Kauai, Hawaii. WWV and WWVH transmitter broadcasts may be heard as amplitude-modulated (AM) audio signals on high-frequency (HF) "shortwave" radio frequencies of 2.5 MHz, 5 MHz, 10 MHz, 15 MHz, and 20 MHz. WWVB is a special low-frequency (LF) transmitter broadcasting at precisely 60 kHz.

The WWV and WWVH shortwave radio broadcasts encode precise timing information in several different ways:

- The carrier frequency of each transmission (i.e. *precisely* 2.5, 5, 10, 15, and 20 MHz)
- Once-per-second "tick" tones
- Human voice announcements of hour and minute values of the *Coordinated Universal Time* just prior to the top of each minute (WWV uses a male voice and WWVH uses a female voice)
- An audible "beep" at the top of each minute
- Digitally-encoded data

[15] Package dimensions are 50.8 mm by 50.8 mm by 25 mm, with through-hole terminals for PCB mounting.

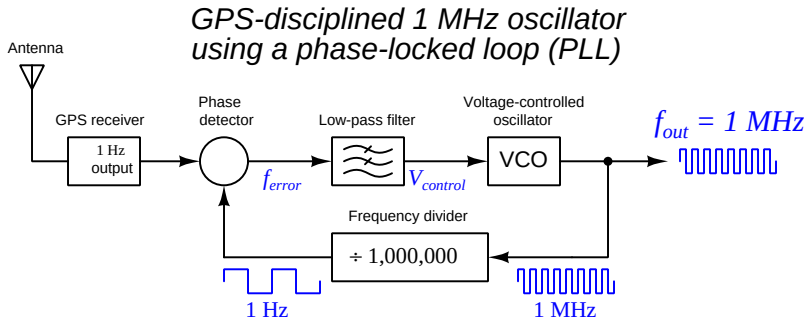
Station WWVB broadcasts a 60 kHz carrier signal synchronized with NIST atomic clocks, and does not encode any audio information as do WWV and WWVH. Instead, the output power for this radio transmitter steps down and up at 1-second intervals in order to slowly encode a BCD digital data stream using pulse-width modulation to differentiate bits: the width of the reduced-power pulse being 0.2 seconds for a binary "0", 0.5 seconds for a binary "1", and 0.8 seconds for a position marker separating BCD number fields. The entire code requires a full minute to send at this slow 1 Hz bit rate.

A newer method of receiving primary-standard time information disseminated by National Measurement Institutes such as the NIST is to use radio signals broadcast by Global Positioning System (GPS) satellites. Every GPS satellite contains its own atomic clock, updated periodically by signals broadcast from terrestrial atomic clocks of greater precision. An appropriate GPS radio receiver unit will be able to receive broadcasts from multiple GPS satellites and from that data report an extremely accurate time signal.

An example of an integrated-circuit GPS receiver unit soldered on to a red printed-circuit board appears here. At the time of its purchase (2022) the retail cost of the IC receiver on the assembled PCB was only about $40 US:

This particular GPS "chip" requires 3.3 Volts DC for its power source, and when locked in to a certain minimum number of GPS satellite signals outputs both a precise 1 Hz pulse as well as an EIA/TIA-232 serial data stream encoding time and GPS location data in *NMEA-0183* format.

If higher frequency than 1 Hz is desired, the 1 Hz GPS receiver pulse may be used to correct the pace of a crystal oscillator circuit. This arrangement is called a *GPS-disciplined oscillator* and usually takes the form of a megaHertz-range crystal oscillator with voltage-controlled fine-adjustment capability connected to a frequency divider circuit to divide the oscillator's high frequency down to a period compatible with the GPS receiver's:

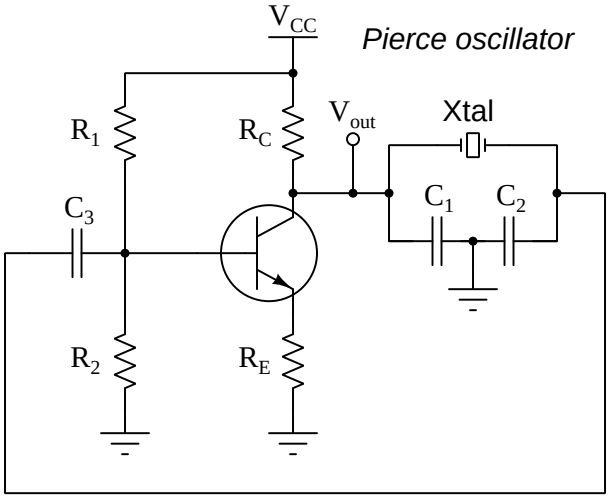


This strategy of adjusting one frequency source to match that of another is generally known as a *phase-locked loop* or *PLL*. If the phase detector senses any shift in phase between the GPS receiver's 1 Hz pulse signal and the digital counter/divider's pulse signal, its control voltage will cause the voltage-controlled oscillator (VCO) frequency to adjust slightly[16] until the two 1 Hz signals are brought back into phase (and therefore precisely matching in frequency) again.

Two examples of GPS-disciplined oscillator instruments appear in the following photographs, one manufactured by Fluke and the other by HP (Hewlett-Packard):

[16] This is referred to as *pulling* the oscillator's frequency.

If atomic-clock accuracy is not needed as a frequency standard, modern crystal oscillator circuits do a remarkably good job at maintaining steady frequency. One example of a crystal-controlled oscillator circuit – called a *Pierce* oscillator – appears in the following schematic diagram:



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 for a mechanical pendulum, helping to stabilize the electronic oscillator circuit's frequency much the same as a mechanical pendulum stabilizes the frequency of a clock mechanism.

Electronic oscillators controlled by the electro-mechanical resonance of a quartz crystal are an example of an *artifact standard* because their accuracy depends on the craftwork of a physical artifact (here, a quartz crystal), as opposed to atomic clocks which are *intrinsic standards* because their accuracy is based on the immutable characteristics of atoms. Crystals can be improperly manufactured, damaged, and influenced by external factors such as ambient temperature. Immediately prior to the adoption of the first atomic clock, NIST's primary frequency reference consisted of a bank of multiple crystal oscillators residing within temperature-controlled chambers.

A photograph of a modern quartz crystal oscillator module appears here:

This particular crystal oscillator module has a resonant frequency of 1.000 MHz (1 million cycles per second), output as a square-wave signal with a peak value of 5 Volts. It is powered by 5 Volts DC, which is a standard source voltage rating for many digital integrated circuits. Its frequency stability is $\pm$ 100 parts per million.

While typical quartz-crystal oscillator frequency is remarkably stable, it may be slightly adjusted by the capacitance connected to the crystal (e.g. C_1 and C_2 in the Pierce oscillator schematic). If one of these capacitances takes the form of a *varactor diode*, it may be adjusted by the application of a DC control voltage, making the circuit a *Voltage Controlled Crystal Oscillator* or *VCXO*. Since ambient temperature is the most significant environmental variable influencing a crystal's resonant frequency, it is possible to achieve outstanding frequency stability by supplying a control voltage to a VCXO from a temperature-sensing circuit, the idea being to compensate for the effects of changing ambient temperature on the crystal by adjusting the adjacent capacitance by just the right amount. When built as such, the circuit becomes a *Temperature-Compensated Crystal Oscillator* or *TCXO*.

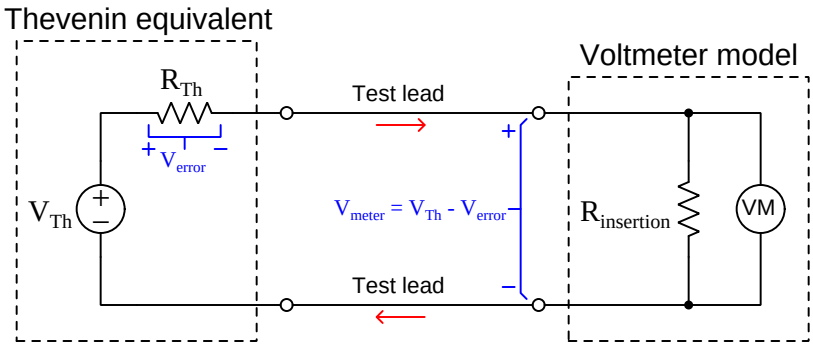
An even more direct method of stabilizing the resonant frequency of a quartz crystal is to control its temperature by placing it inside an electrically-heated oven, and thermostatically regulating that oven's temperature slightly above ambient. This arrangement is called an *Oven-Controlled Crystal Oscillator* or *OCXO*. For even greater frequency stability some crystal oscillators use a *double oven*, the crystal being inside of one tiny oven which itself is inside yet another tiny oven. Of course, providing a stable operating temperature for the crystal requires additional electric power to run the oven, but for some high-precision applications this extra power requirement is a small price to pay. These *Double Oven-Controlled Crystal Oscillator* (*DOCXO*) modules achieve both short- and long-term frequency stabilities less than $\pm$ 1 part per billion (ppb).

3.10 Precision voltage measurement techniques

3.10.1 Null-balance voltage measurement

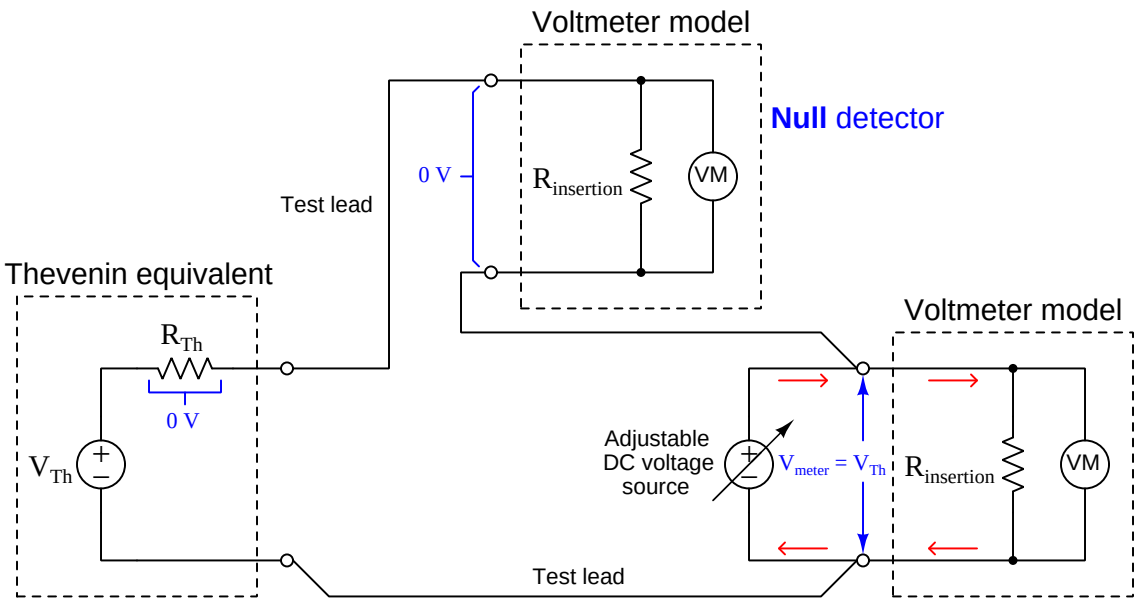
An ideal voltmeter has infinite internal resistance (often called *insertion resistance*), so that no current will be drawn by the meter from the circuit under test. Any current drawn by a voltmeter while measuring voltage constitutes energy removed from that circuit, Joule's Law predicting the rate of energy draw as the product of current and voltage ($P = IV$). Removing energy from a circuit affects the very voltage we are tasked to measure, and so reducing this energy loading to zero is our goal for precision voltage measurement.

A simple model of a voltmeter (shown as a high-valued resistance) connected to any energized electrical network (shown as a Thévenin equivalent network) demonstrates how voltmeter insertion resistance creates a systematic measurement error:



The small amount of current passing through the voltmeter's internal resistance ($R_{insertion}$) must also pass through the test circuit's Thévenin equivalent resistance (R_{Th}), resulting in the network's terminal voltage being slightly less than its internal Thévenin voltage (V_{Th}). It is as though the voltmeter forms a *voltage divider* with the test circuit's internal resistance. As such, the voltmeter cannot ever sense the circuit's true internal voltage, but always something less than that. In most voltage-measurement applications this error is so small it may be safely ignored, but in the field of metrology it is our goal to account for – and ideally to nullify – *every* source of error if possible.

An ingenious method to combat this error involves the addition of an adjustable DC voltage source as well as a second voltage-sensing meter which we call a *null detector*:



The sole purpose of the null detector is to indicate when we reach a condition of zero voltage between its terminals. If we imagine adjusting the variable-voltage DC source to the point where it exactly equals V_{Th} in the test circuit, we may see by applying Kirchhoff's Voltage Law that there must be zero voltage dropped across the test circuit's internal resistance (R_{Th}) as well as zero voltage across the null detector's terminals. In this "null-balanced" state, the right-hand voltmeter's current requirements are satisfied by the adjustable DC source and not by the circuit under test.

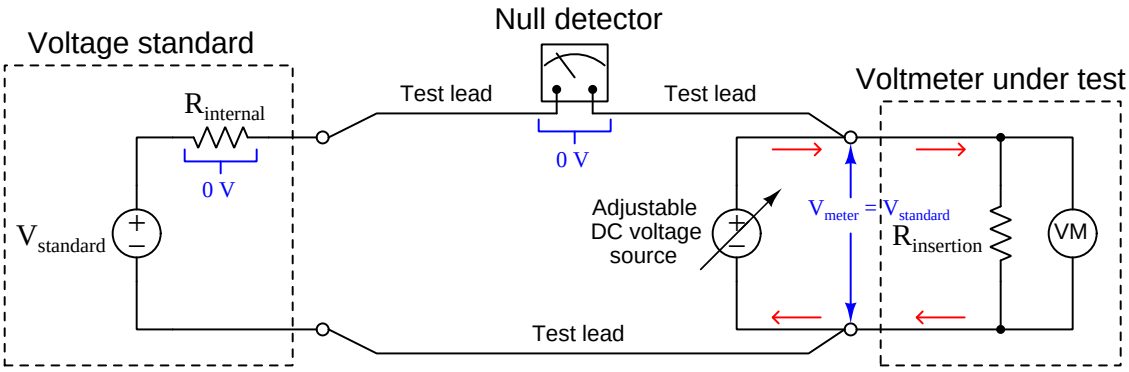
To operate this circuit, we simply adjust the variable DC voltage source until the null detector registers zero voltage, then we can read the other voltmeter's indication and interpret that as the test circuit's true internal voltage (V_{Th}) without any loading error. Thus, with this null-balance technique we are actually able to measure the internal voltage of the test circuit while truly drawing zero current, so long as our DC source adjustment is precise enough and our null detector is accurately registering zero voltage.

In the earliest days of electrical metrology, extremely sensitive *galvanometers* were used as null detectors, their only functional requirement being a sufficiently high degree of sensitivity so that the presence of any voltage across and any current through its terminals would be immediately apparent to the person operating the apparatus. This single-purpose functionality of a null detector meant the galvanometer need not have an accurate measurement range, or even linearity, but simply be able to indicate zero versus non-zero conditions. In fact, a common technique for field technicians before the advent of rugged meters having sufficient sensitivity was to employ audio headphones connected as null instruments through an on/off switch: the circuit was known to be in a balanced (null) state when the headphones no longer produced audible "click" noises while the switch was toggled on and off.

3.10.2 Null-balance meter calibration

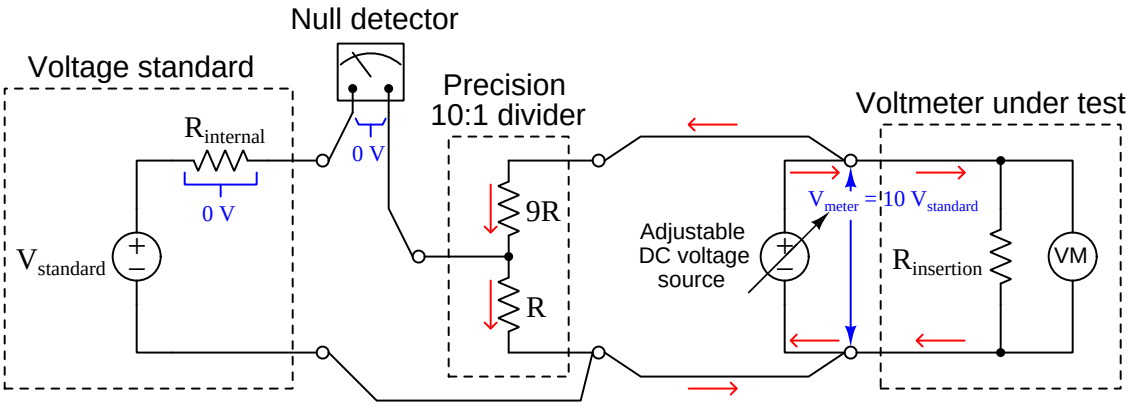
Null-balance voltage measurement is also useful when testing a voltage-measuring instrument against a voltage standard without drawing any current from that standard. This was an especially important technique back in the days of mercury standard cells which were fairly intolerant of loading without suffering some induced inaccuracy as a result of that loading.

This first circuit shows how to apply a standard voltage to a voltmeter under test without loading the voltage standard:

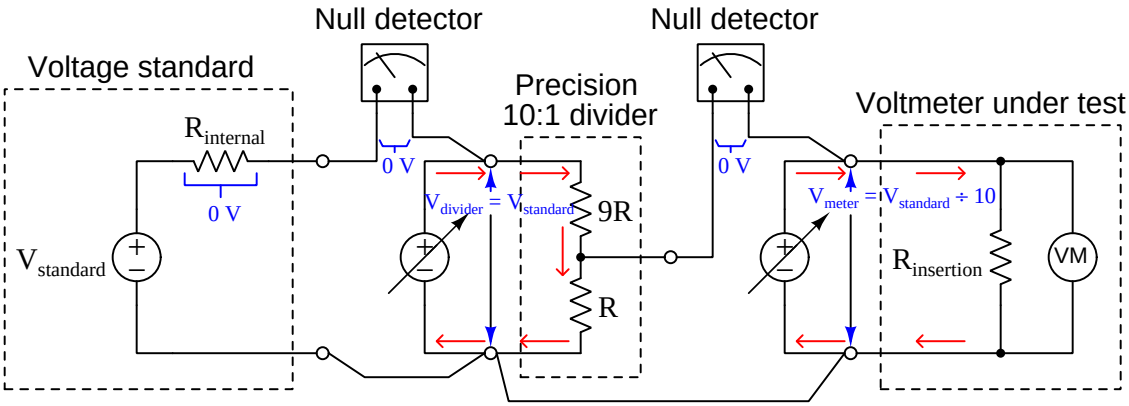


Here, when the null detector registers zero voltage we know the voltmeter under test will have precisely the same voltage across its terminals as the voltage standard. Furthermore, with the null detector registering zero we may confidently conclude that the voltage standard is not being loaded at all by the voltmeter under test, that voltmeter's insertion resistance being energized entirely by the adjustable (non-standard) voltage source.

Fixed-voltage standards are by definition non-adjustable, and so in order to use them to calibrate voltage-measuring instruments over wide ranges of voltage we must creatively employ precision voltage divider networks. Suppose we wished to use a voltage standard to generate ten times its rated voltage to calibrate a high-range voltmeter. By connecting a precision 10:1 voltage divider[17] as shown in the following diagram, we may trust that 10 times the standard's voltage will be applied to the voltmeter under test when the null detector registers zero:



Conversely, if we wished to test a low-range voltmeter at a voltage one-tenth that of the trusted standard, we could use the same 10:1 precision divider but with two adjustable sources and two null detectors like this:



The left-hand null detector indicates when the left-hand adjustable source's voltage precisely matches the voltage standard's while supplying the 10:1 divider network with the current it needs. The right-hand null detector indicates when the right-hand adjustable source's voltage

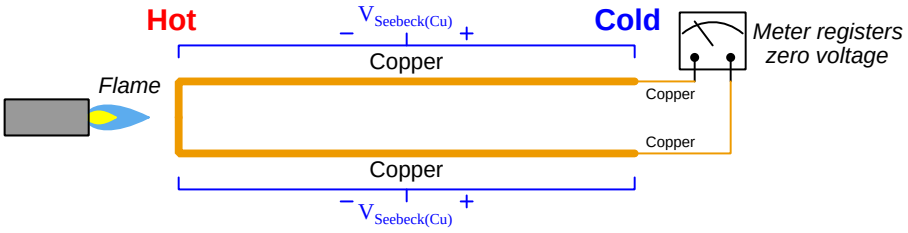
[17] A recommended resistor network for achieving a 10:1 division ration is the *Hamon voltage divider* consisting of three (nearly) equal-valued fixed resistors and an adjustable resistance which may be calibrated to achieve a nearly perfect 10:1 voltage division ratio using easily-obtained components.

precisely equals one-tenth of the divider's input voltage while supplying the test voltmeter's insertion resistance current.

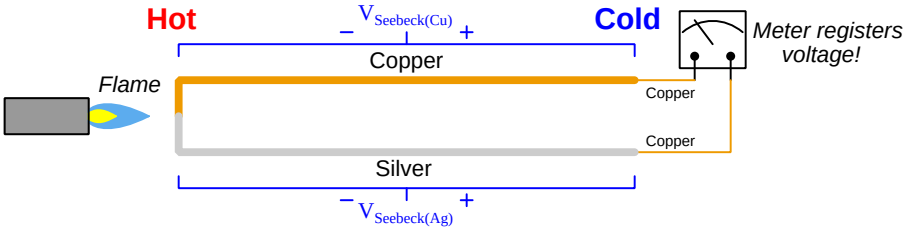
3.10.3 Thermal voltage errors

Another voltage-measurement error we need to consider stems from the *Seebeck effect*, a thermodynamic phenomenon by which a conductor develops a voltage along any temperature gradient, named after the German physicist Thomas Johann Seebeck who discovered this phenomenon in 1821. Any metal wire will produce a voltage (typically in the order of microVolts or milliVolts DC) if one end of that wire is hotter than the other, and in certain circumstances this thermo-electric voltage (often called *thermal EMF*) will constitute an error if it adds to or subtracts from some other voltage of interest.

Normally the Seebeck effect goes unnoticed in circuits comprised of a single conductive material because any closed circuit experiencing a difference of temperature from one location within it to another is guaranteed to have *two* opposing temperature gradients, and therefore two opposing Seebeck voltages that cancel out. Consider this all-copper circuit where all of the voltmeter's internal wiring is also made of copper:



Since all wires are made of the same material (copper metal) and they experience the same hot-to-cold temperature gradient, their Seebeck voltages will be exactly equal in magnitude and opposite in polarity, canceling at the meter. However, if we make one wire out of copper and the other out of a different metal such as silver having a different Seebeck coefficient than copper, the same temperature gradient produces a measurable voltage at the meter:



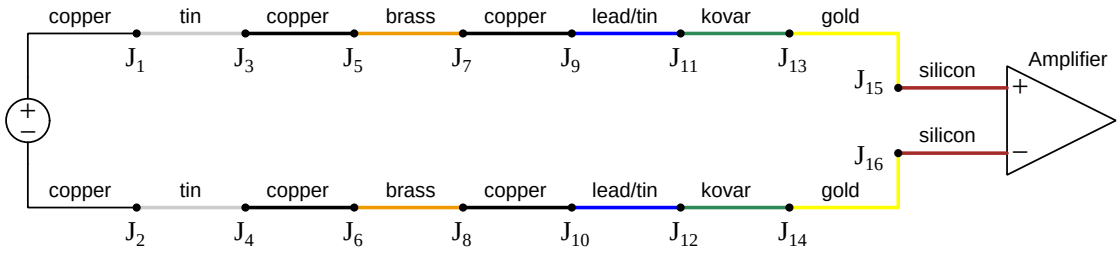
$$V_{Seebeck(Cu)} \neq V_{Seebeck(Ag)}$$

These *Seebeck coefficient* values are quite small for most materials – typically several microVolts per Kelvin[18] for common metals, and hundreds of microVolts per Kelvin for semiconductors – but

[18] The *Kelvin* temperature scale has the same interval value as the Celsius temperature scale, but is offset such that 0 K is *absolute zero* where all molecular motion ceases. This offset is equal to 273.15 degrees; e.g. 20 °C = 293.15 K. For this reason, it is valid to interpret Seebeck coefficients as microVolts *per degree Celsius of temperature change*. The reason Kelvin is preferred even though each degree Celsius of temperature change is the same thing as one Kelvin of temperature change, is that referencing the unit of Kelvin implies all Seebeck voltage vanishes at absolute zero.

remember that in the field of metrology we ultimately aim to account for *all* sources of error, and so for some applications even a few milliVolts of Seebeck voltage produced over a temperature gradient can wreak havoc with the voltage-measurement accuracy!

Given that fact that the Seebeck effect is only relevant as a source of voltage-measurement error in circuits consisting of dissimilar conductive materials, it may seem as though the solution to avoiding these errors is to simply avoid dissimilar-metal junctions in our metrology circuits. However, this is nearly impossible. Consider this example circuit where a voltage source connects to the input terminals of a differential amplifier integrated circuit:



From left to right, our circuit consists of copper wires within the voltage source, followed by tin plating on the source's external terminals, followed by more copper wire, followed by brass terminals inside a terminal block, followed by more copper wires, followed by the lead/tin solder joint to the IC's terminals, followed by kovar metal *bond wires* joining the IC's terminals to gold-plated pads on the silicon die, to (finally) the silicon semiconductor material within the IC.

The situation is not as hopeless as it might first appear from this complex collection of dissimilar metals joined together at sixteen places. In principle, the total Seebeck voltage generated by this circuit should be zero if every respective pair of junctions is at the same temperature (i.e. J_1 and J_2 at the same temperature, J_3 and J_4 at the same temperature, etc.), and if that is true then it means our differential amplifier will only see the intended source voltage. In practice, however, we generally see some differences in temperature[19] between these point-pairs which means Seebeck voltage remains a source of voltage-measurement error in any circuit comprised of dissimilar conductors.

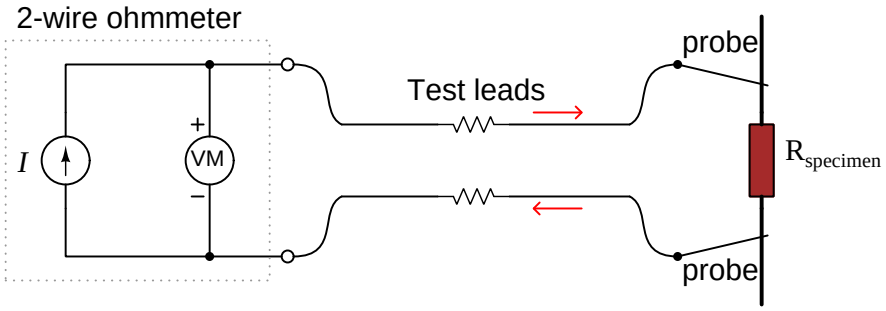
Practical mitigations for Seebeck voltage in precision voltage-measurement circuits include avoiding unnecessary dissimilar-metal junctions (e.g. using copper binding posts and bare copper test leads lacking any plating), using metal alloys exhibiting the lowest Seebeck coefficients compared with copper (e.g. tellurium-copper alloy rather than brass or other inexpensive copper alloys), and/or carefully controlling ambient temperature in the testing room (e.g. insulating the room from external temperature changes, using an airlock for entry to avoid air drafts whenever someone enters or exits, leaving the calibration apparatus alone for long periods of time to allow temperature gradients to decay).

[19] A good example of how this may occur in a real circuit is when certain portions of an IC dissipate more heat than others, and/or when certain portions of an IC enjoy better heat-sinking than others, causing differences of temperature between areas on the silicon die or between different IC terminals.

3.11 Precision resistance measurement techniques

3.11.1 Two-wire resistance measurement

Simple ohmmeters interpret resistance by Ohm's Law, passing a known current through the resistance under test, and inferring that resistance value by the voltage sensed at the meter's test jacks. While this arrangement works fine for most applications, when measuring resistances small enough that the wire resistance of the meter's test leads becomes significant by comparison that added wire resistance contributes a positive error to the measurement:



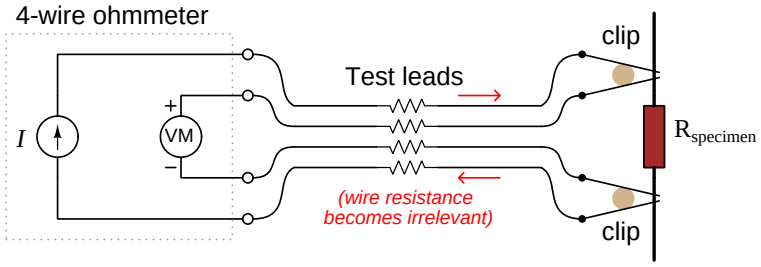
$$R_{specimen} + R_{leads} = \frac{\text{Voltmeter indication}}{\text{Current source}}$$

The internal voltage measurement taken between the two test jacks by the ohmmeter represents the voltage dropped by the test resistance ($R_{specimen}$) *and* the voltage dropped by the two test lead wires (R_{leads}). It is not uncommon for multimeter test leads to have a combined total resistance of a few tenths of an Ohm (especially if the plug/jack connections are not clean), adding this much error to the meter's resistance measurement.

One way to mitigate this error is to "zero" or "null" the ohmmeter while its test probes are connected directly together. This effectively subtracts the lead wire resistance from the final measurement, since defining "0 Ohms" with the probes directly touching means that any subsequent resistance measurement should be resistance above and beyond that of the test lead wires. However, this technique suffers from the uncertainties of varying contact resistance as the person holds the probes together.

3.11.2 Four-wire resistance measurement

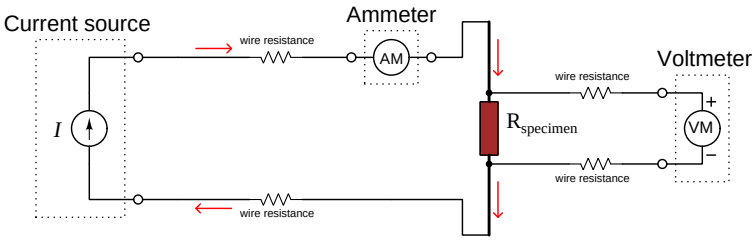
An ingenious solution to the problem of canceling out test lead wire resistance utilizes a slightly different design of ohmmeter with four jacks and four test leads rather than two. The meter's internal current source connects to two of those jacks, while its internal voltage-sensing circuitry connects to the remaining two jacks. This is known as the *Kelvin four-wire technique*:



$$R_{specimen} = \frac{\text{Voltmeter indication}}{\text{Current source}}$$

The meter's internal voltage-sensing circuitry senses *only* the voltage dropped by the resistance under test ($R_{specimen}$), and does not suffer any significant voltage drop along its two wires because that voltage-sensing circuit draws negligible current. The other two wires connecting the meter's current source to the resistance under test of course carry current, but their voltage drops aren't being sensed at all by the voltage-sensing circuit. Therefore, test lead wire resistance becomes irrelevant.

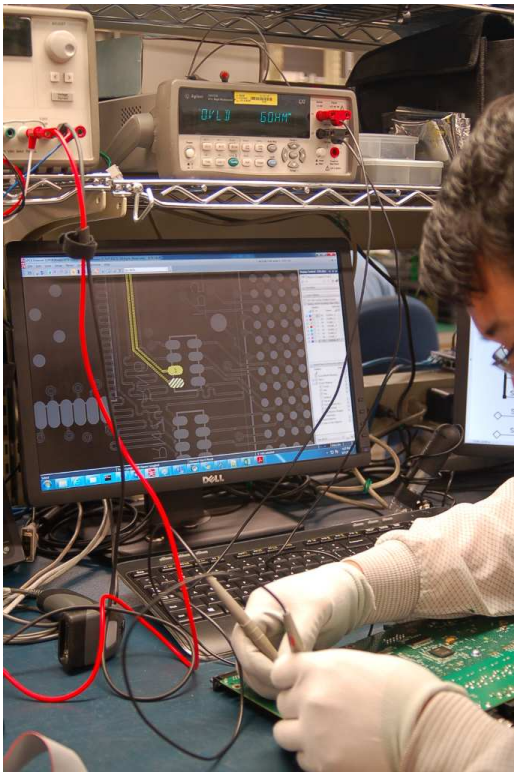
This precision resistance-measuring technique, of course, need not require a special four-wire ohmmeter, but in fact may be performed with a current source and two multimeters as shown below:



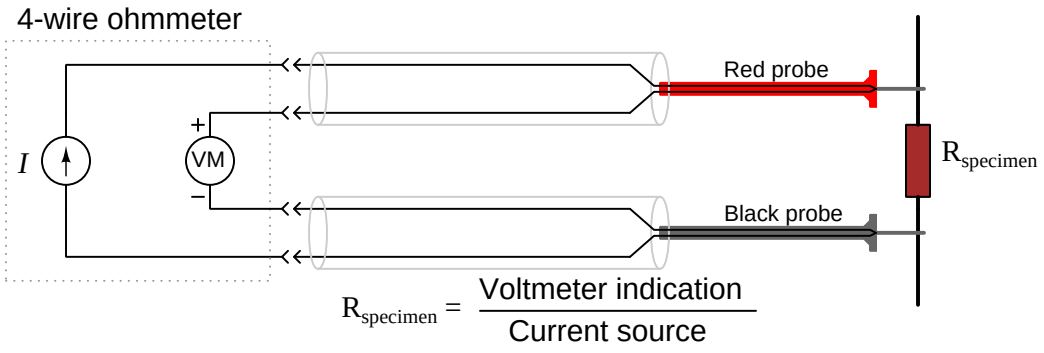
$$R_{specimen} = \frac{\text{Voltmeter indication}}{\text{Ammeter indication}}$$

All wire resistances are still irrelevant, which is the strength of this resistance-measuring technique. The voltmeter's wire resistances don't matter because they will carry no current and therefore drop no voltage. The current source's and ammeter's wire resistances don't matter because whatever voltages they will drop won't be sensed by the voltmeter. So long as both meters are accurate, the specimen's inferred resistance value will also be accurate and not include any extraneous resistances.

Here we see a photograph of an electronics technician diagnosing a short-circuit between two copper "traces" on a printed circuit board (PCB) using a digital multimeter with four-wire resistance measurement capability:

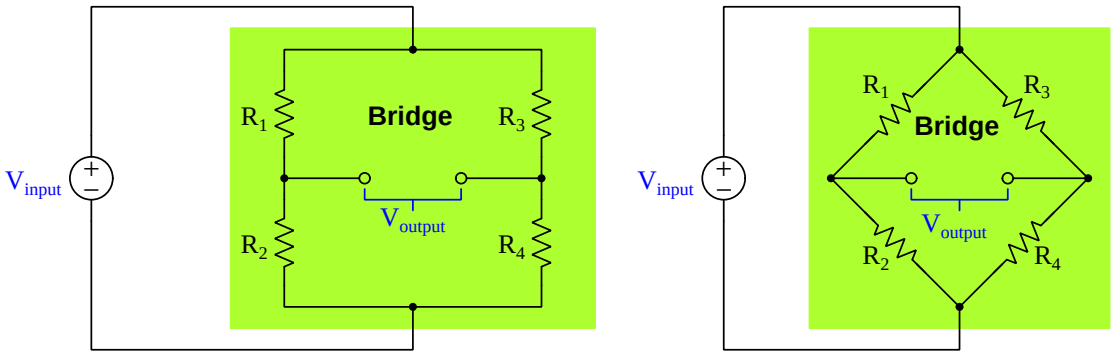


If you look closely at the multimeter positioned on the bench shelf (presently registering "`OVLD GOHM`" because one or both of the test probes is not touching anything else), you will see the two test leads (red and black) actually have *dual* banana-style plugs. Each "test lead" is actually a two-conductor *cable* with two banana plugs inserted into the multimeter, those two conductors joining together at the tip of each probe as shown in the following schematic:



3.11.3 Wheatstone bridge resistance measurement

Unknown resistance values may also be measured by comparison against resistors of precisely-known resistance. The most common way of doing this is to use a *Wheatstone bridge* network consisting of four resistors, three of them having known values and the fourth being unknown. These networks function as a pair of two-resistor voltage dividers, with test points provided to connect a sensitive voltmeter:



Note that these two networks are electrically identical. They only differ in orientation (i.e. angled resistors and connecting wires), not in the components or the connections. Although the "diamond" (right) depiction is more common, you should feel free to think of bridge circuits drawn rectangularly (left) if this is helpful.

A bridge circuit is said to be *balanced* when its two voltage divider sub-networks have equal voltage-division ratios, thereby causing V_{output} to be zero. With some algebraic manipulation, we may simplify the "balance equation" for this bridge circuit even more:

$$\frac{R_2}{R_1 + R_2} = \frac{R_4}{R_3 + R_4}$$

$$\frac{R_1 + R_2}{R_2} = \frac{R_3 + R_4}{R_4}$$

$$\frac{R_1}{R_2} + \frac{R_2}{R_2} = \frac{R_3}{R_4} + \frac{R_4}{R_4}$$

$$\frac{R_1}{R_2} + 1 = \frac{R_3}{R_4} + 1$$

$$\frac{R_1}{R_2} = \frac{R_3}{R_4}$$

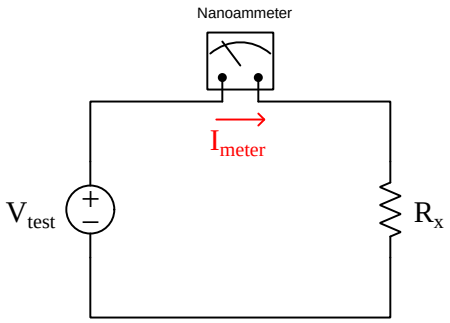
With this simple equation, it is possible to use a bridge circuit to compare the value of an unknown resistance against the values of three known resistances. If such a bridge circuit is constructed, energized, and then one of its known resistances varied until a "balanced" condition is achieved (as indicated by a *zero-voltage* or *null* condition sensed at the two output terminals), the value of the unknown resistance may be found by solving the balance equation for that one resistance.

Since all of these resistance-measurement techniques utilize voltage measurement in the determination of resistance, best-practices for precise voltage measurement (e.g. minimization of Seebeck voltages, or thermal EMFs) apply just as well to resistance measurement.

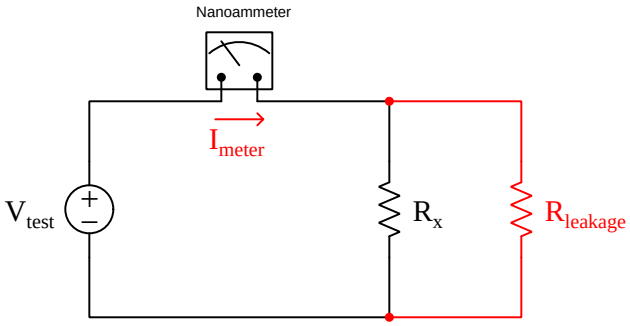
3.12 Guarded measurements

Parasitic resistance, capacitance, and inductance is unavoidable to some extent in all electrical test fixtures, and may significantly corrupt the accuracy of our measurements. One mitigation technique is called *guarding*, intended to minimize the effects of parasitic resistance and parasitic capacitance.

Perhaps the simplest and easiest-to-understand application of guarding is in the measurement of extremely high resistances, where any form of electrical insulation in parallel with the resistance under test comprises both parasitic resistance and parasitic capacitance that may interfere with proper measurement of the specimen resistance. Consider an elementary resistance-measuring circuit using a high-voltage DC power supply and a nanoammeter to measure the small current expected to pass through the specimen resistance R_x consisting of many GigaOhms:



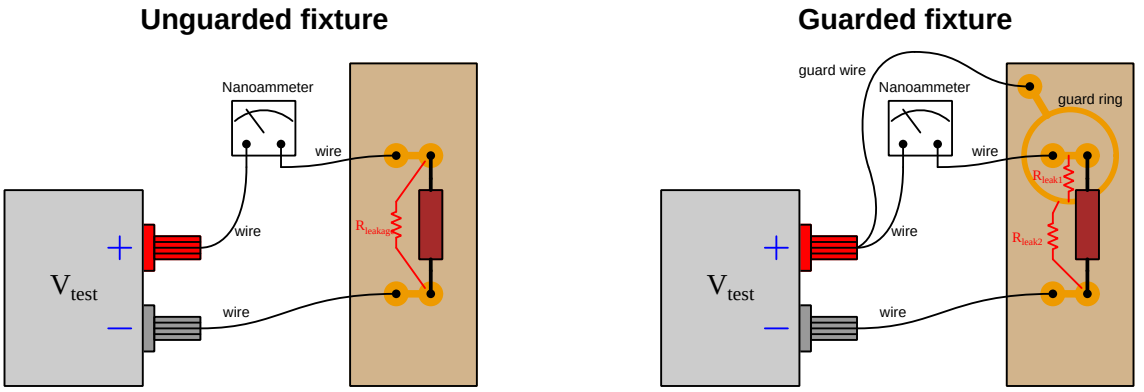
Determining the value of R_x should be as easy as dividing V_{test} by the current (I) measured by the nanoammeter. What could be simpler? However, if we consider the fact that the specimen resistance will be physically mounted on some form of insulating fixture (e.g. terminal blocks, printed circuit board) to hold it in place, and that this fixture will have its own high value of insulation resistance, we may re-draw an equivalent schematic diagram to show the fixture's *leakage resistance* in parallel with the specimen:



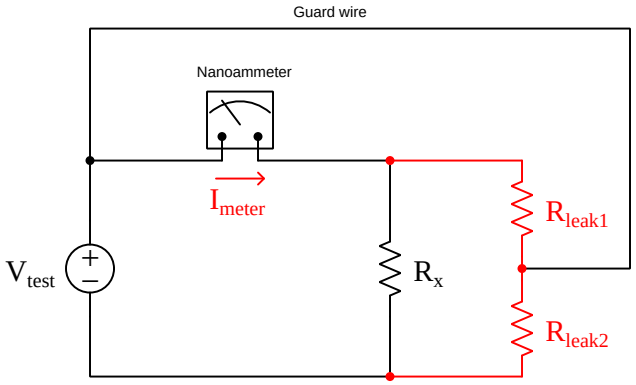
Now, it is plain to see that the nanoammeter's current will in fact be the result of the test voltage applied across *two* parallel resistances that are both unknown in value to us:

$$I_{meter} = \frac{V_{test}}{R_x || R_{leakage}}$$

A very clever way to render the leakage resistance mostly irrelevant is to equip the fixture with something called a *guard ring*. On a printed circuit board (PCB), this consists of a ring in the form of a circular copper trace surrounding the meter-side terminal where the resistance specimen attaches. The following illustration shows an unguarded versus a guarded PCB, with parasitic circuit-board resistances shown in red:

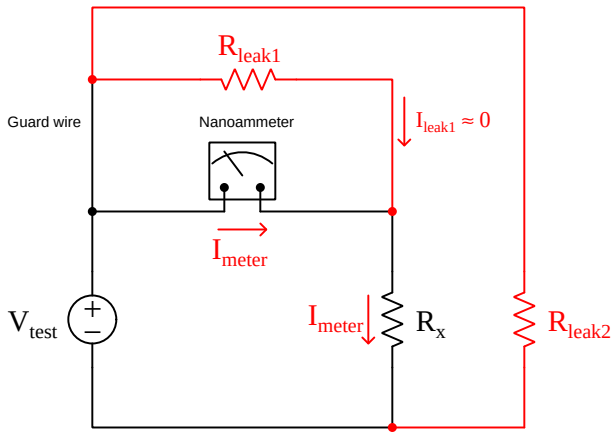


The equivalent electrical schematic of the guarded fixture appears below, showing how the PCB's $R_{leakage}$ resistance gets divided into two distinct parasitic resistances, R_{leak1} and R_{leak2}:



Note how the upper leakage resistance R_{leak1} is effectively placed in *parallel* with the nanoammeter by the inclusion of the guard wire and the PCB's guard ring. This is important because the nanoammeter has a relatively low insertion resistance like all ammeters – *much* less than R_{leak1} – and so the parallel equivalent resistance of R_{meter} and R_{leak1} is very nearly equal to R_{meter}, the result being that current passing through R_{leak1} is negligible. Some leakage current still passes through the remainder of the PCB's resistance, but this current is sourced through the guard wire directly from V_{test} and does not pass through the nanoammeter where it would otherwise inflate the meter's measurement.

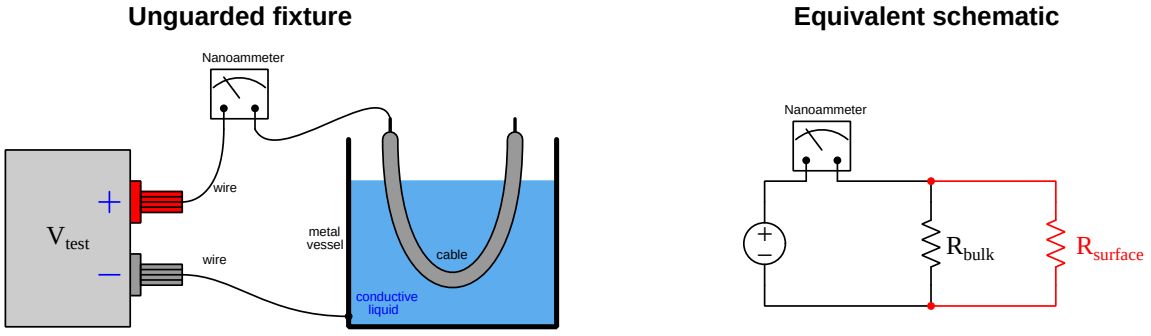
The next equivalent schematic has been re-drawn to more clearly show how these two parasitic resistances fit into the larger circuit:



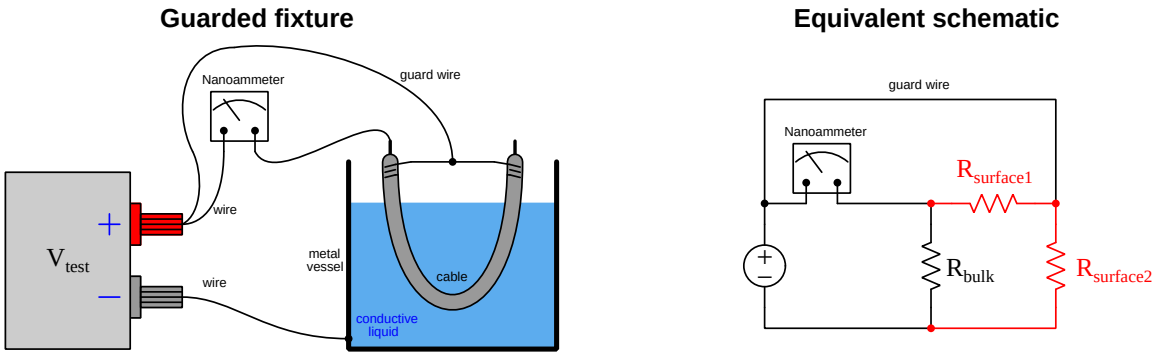
Note how the inclusion of the guard wire forces R_{leak1} to be in parallel with the nanoammeter, thereby ensuring negligible voltage drop across it and negligible current through it. With I_{leak1} essentially canceled, by Kirchhoff's Current Law we may conclude that current through the specimen resistance R_x must be entirely equal to the meter's current. Meanwhile, none of R_{leak2}'s current passes through that meter and therefore has no effect on the accuracy of our measurement. R_x may now be safely calculated using Ohm's Law on the basis of source voltage and meter current:

$$I_{meter} = \frac{V_{test}}{R_x} \qquad\qquad R_x = \frac{V_{test}}{I_{meter}}$$

Guarding is also useful for cable insulation resistance testing, where a section of insulated cable is immersed in a conductive liquid solution and high voltage impressed across the cable's insulation by connecting it between the cable conductor and the conductive solution. Here, the parasitic leakage resistance comes in the form of *surface* resistance along the outer surface area of the cable, resulting in excessive meter current which will be interpreted as falsely low insulation (bulk) resistance:



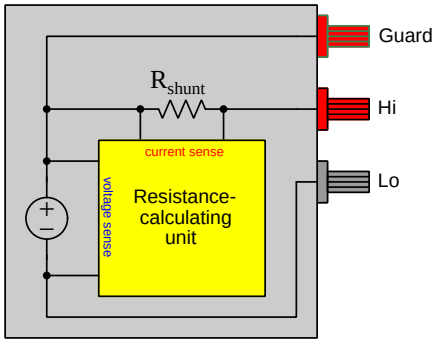
Since we only wish to measure the insulation's bulk resistance and not this surface resistance, guarding is necessary to mitigate the latter. This is achieved by wrapping a bare guard wire a few turns over the insulation of the cable at both ends as shown below:



Now, the surface resistance between the guard wire and the cable conductor ends ($R_{leakage1}$) has negligible voltage across it and therefore leaks negligible current, while the remainder of the surface resistance ($R_{leakage2}$) is supplied by current directly from the test voltage source that does not pass through the nanoammeter.

High-voltage insulation testing instruments[20] often provide a guard terminal for this purpose. A simplified schematic of such a source-measure unit (SMU) with guard terminal is shown below:

High-voltage SMU with guard terminal

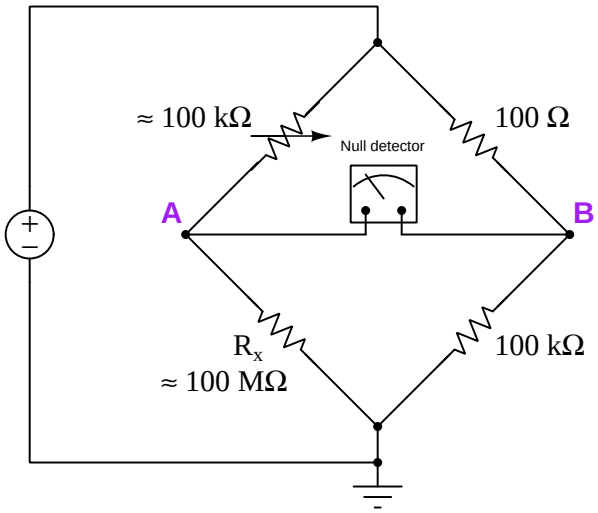


It is important to note which of the output terminals is at the same potential as the guard. The guard terminal may be held to the same potential as the positive output terminal ("Hi" as shown above) or as the negative output terminal ("Lo") depending on the design of the SMU instrument and its intended application. *What matters is that the guard is at the same potential as the output terminal through which current is measured.*

[20] A popular brand in the electrical industry is the *Megger*, and for this reason many electrical technicians refer to any high-voltage insulation testers as "Meggers" in the same way that other common tools are often colloquially referred to by quintessential brand names (e.g. "Vise-Grip" locking pliers, "Channellock" tongue-and-groove pliers, "Crescent" adjustable wrenches, "Fluke" multimeters).

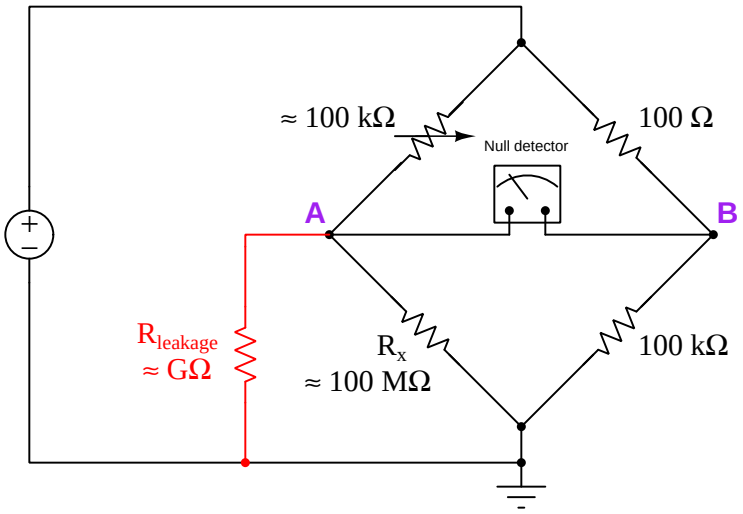
Guarding also finds application in high-resistance bridge circuits. Take for example a Wheatstone bridge configured to measure a very high specimen resistance (R_x) somewhere in the range of 100 MΩ, with other resistor values chosen to set up a 1000:1 resistance ratio:

High-resistance Wheatstone bridge

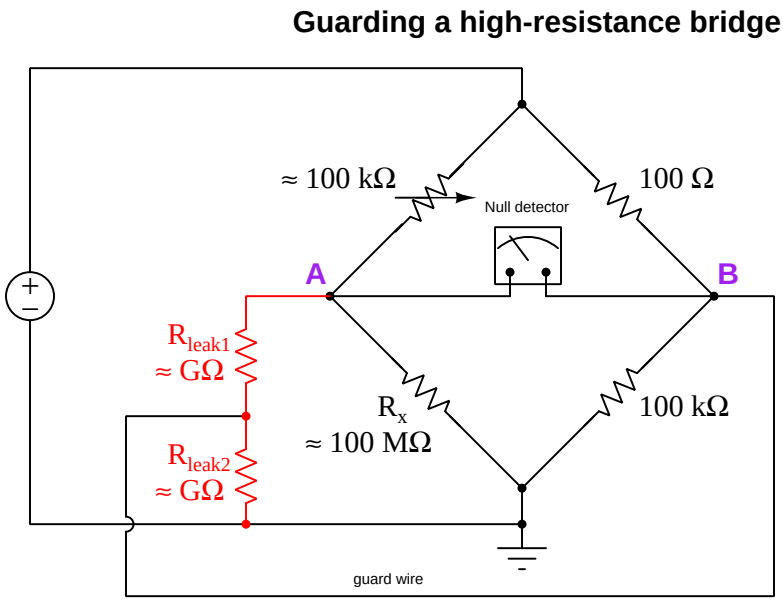


Leakage resistance parasitically exists across all four resistors in the bridge circuit, but it is only a concern in parallel with the specimen resistance given its exceptionally high value, and so it is there that we must apply guarding:

Leakage resistance shown in parallel with R_x

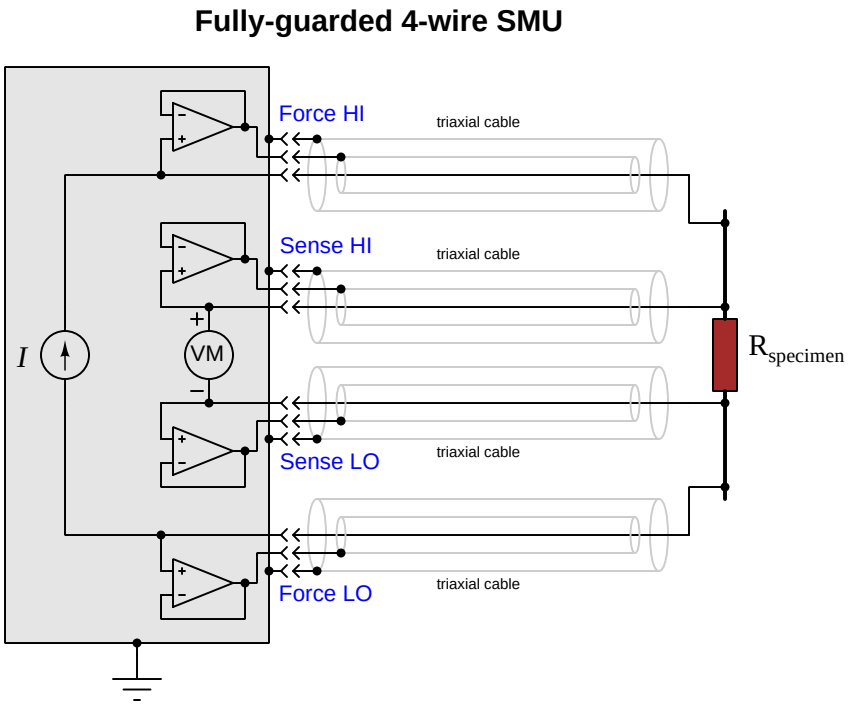


Knowing that points A and B in this bridge circuit will be equipotential once a condition of balance has been reached allows us to use point B as a guard potential for a conductive ring around terminal A:



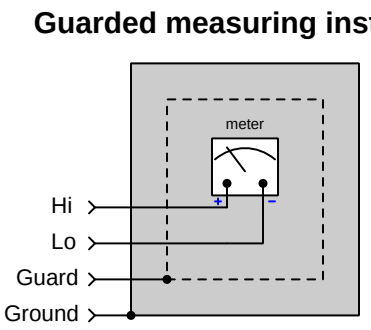
Leakage resistance R_{leak1} is now seen to be in parallel with the null detector where it will have zero voltage across it in a condition of balance, and therefore no current at all will pass through it. Leakage resistance R_{leak2} will still have voltage across it, but the inclusion of the guard wire places it in parallel with the 100 kΩ fixed resistor where its effect will be orders of magnitude less than if it were allowed to exist in parallel with R_x.

Laboratory-grade source-measure units (SMUs) may be equipped with *driven guards* and *triaxial cable* connectors to eliminate leakage currents that might otherwise pass through the cable insulation between the center conductor and grounded shields of ordinary coaxial cables. The following illustration shows a four-wire SMU suitable for Kelvin resistance measurements, each of the four conductors connected to the resistance specimen via triaxial cables with the inner shields driven by operational amplifier voltage buffers to maintain zero voltage difference between the center conductor of each cable and its inner shield. The outer shields are still connected to chassis ground for the purpose of electrostatic shielding, and any leakage currents passing between inner and outer shields of each cable are supplied by the opamps and not by the current source nor the voltmeter conductors:



As was mentioned at the beginning of this section, but is worth reiterating, guarding serves to mitigate the effects of both parasitic (leakage) resistance as well as *parasitic capacitance*. With the inner shields of the triaxial cables driven as guards, capacitance naturally existing between the inner shield and the signal conductor within each cable experiences zero voltage and therefore cannot store energy in the form of an electric field. This frees the signal conductors of having to convey energy to or from these capacitances with changing voltages, translating into quicker response from the instrument due to the effective elimination of parasitic RC time constants.

Some metrology equipment is build to have an internal *guard shield* which functions much the same as inner shields in triaxial cables: an enclosing surface maintained at a potential at or near the measurement conductors to minimize leakage current between those conductors and ground, itself enclosed within the *noise/safety shield* of the measuring instrument's grounded metal chassis:



The purpose of the guard shield, as always, is to redirect leakage currents away from the signal conductors (i.e. the test leads connected to the Hi and Lo terminals) so they do not cause measurement errors. Typically the guard terminal is connected to a source having potential equal to the common-mode voltage of the measured signal (usually the same potential as the Lo terminal). This guard source may be an operational amplifier configured to follow the common-mode potential, or else it may simply be the Lo-side connection on whatever device is driving the measuring instrument.

3.13 Electrical metrology tools

3.13.1 Multifunction calibrators

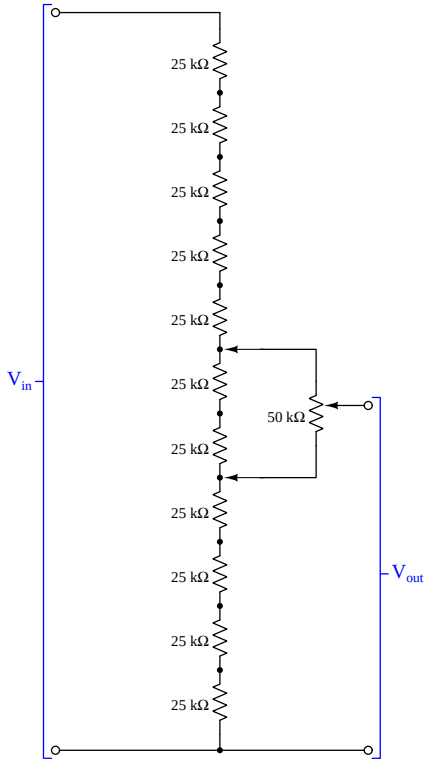
Perhaps the most versatile modern instrument for electrical metrology is the *multifunction calibrator*, an example below being a Fluke model 5730A:

These instruments are able to both source and measure electrical quantities, and for this reason are classified as *source-measure units* or *SMUs*[21]. They are ultra-precise in both their ability to accurately output DC and AC voltages and currents on command by entering the desired quantity on a keypad, as well as measure voltages and currents applied to them by external device. Multifunction calibrators are also able to simulate precise quantities of electrical resistance on demand by sinking current in precise proportion to an applied voltage.

[21] The lowly ohmmeter may be thought of as an elementary SMU because it measures resistance by simultaneously sourcing current and measuring voltage drop, although the term SMU is typically reserved for laboratory-grade instruments having much more sophisticated capability.

3.13.2 Kelvin-Varley voltage dividers

A special form of voltage divider network often used in precision-calibration work is the *Kelvin-Varley* voltage divider, constructed of eleven equal-rated precision resistors with a potentiometer having a total resistance of precisely *twice* that of any of the individual fixed-value resistors, the outer terminals of that potentiometer connected across any *two* of the fixed resistors in the network. An example is shown below:

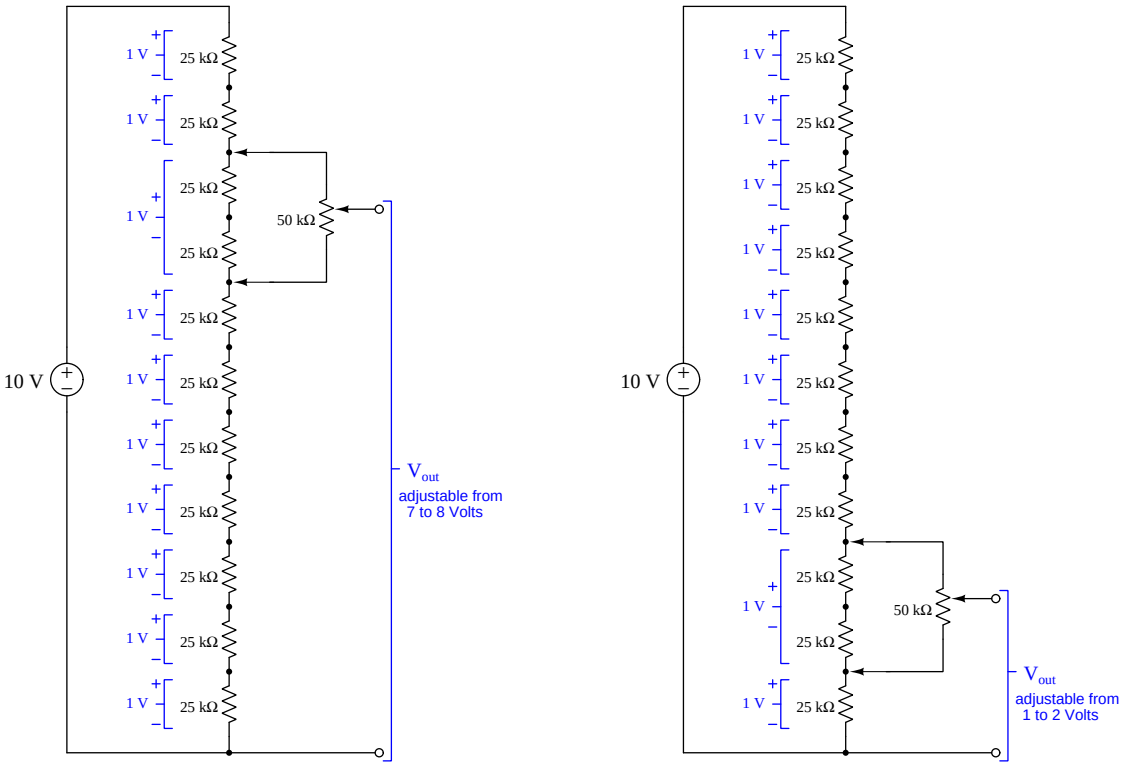


The 50 kΩ potentiometer connected in parallel with two of the series-connected 25 kΩ resistors forms an equivalent sub-network having a resistance of 25 kΩ. Essentially, connecting the potentiometer across any two of the eleven 25 kΩ resistors makes the circuit behave as if it were a series string of *ten* 25 kΩ resistors, dividing the input voltage (V_{in}) into ten equal proportions. Thus connected, moving the potentiometer's wiper position from one extreme to the other results in the output voltage (V_{out}) being adjustable between consecutive tenth-proportions of V_{in}. In the circuit shown above, if we had a 10 Volt source connected between the V_{in} terminals, moving the potentiometer wiper to the full-down position would give us an output voltage of precisely 4 Volts, and moving the wiper to the full-up position would output precisely 5 Volts.

Kelvin-Varley dividers are useful for dividing a given voltage precisely into *tenths* while providing an adjustment range also one-tenth that of the applied voltage. In the configuration shown above, moving the potentiometer's wiper from one extreme to the other yields a voltage range of 4 to 5

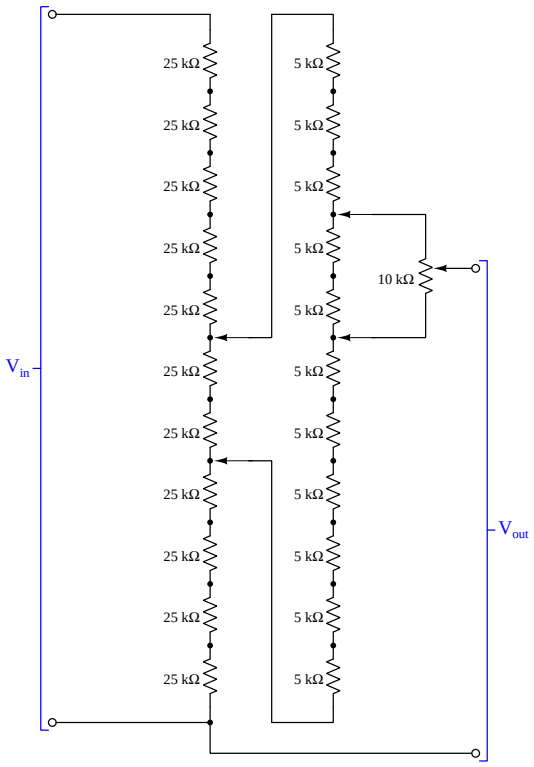
Volts (V_{out}) assuming a 10 Volt source (V_{in}). However, if the potentiometer's outer terminals are both moved one step up the series resistor chain the output voltage range becomes 5 to 6 Volts.

Just for clarity, we will examine two more examples using this same Kelvin-Varley voltage divider, the potentiometer terminals connected across different pairs of resistors in each case. The reader is encouraged to perform the necessary series-parallel analysis to confirm these voltage ranges:



One might think of the potentiometer's terminal positions along the eleven-resistor series network as being the "coarse" ratio adjustment while the wiper position of the potentiometer itself serves as the "fine" ratio adjustment.

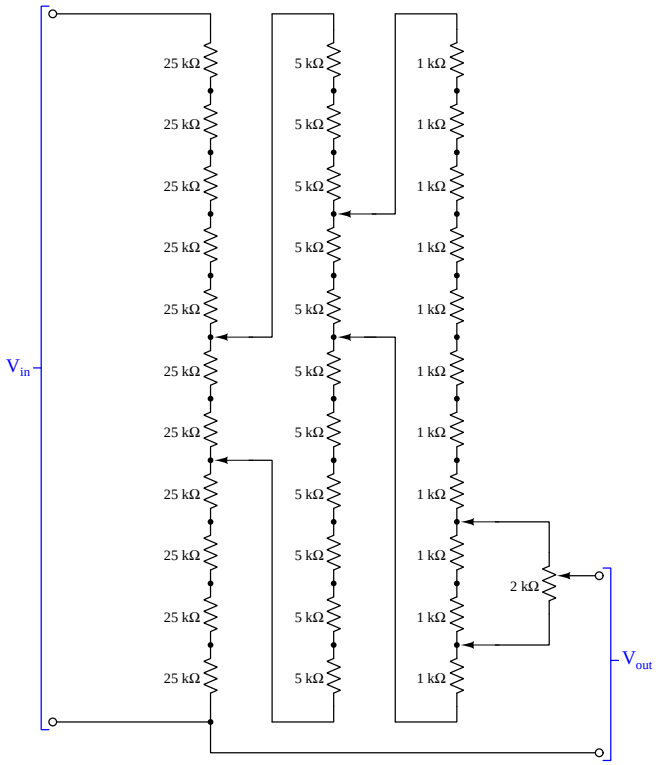
Kelvin-Varley voltage dividers used in electrical metrology[22] work consist of *cascaded* divider networks. We will gradually explore this concept by adding one more stage to the previous divider network. Just as with the single-stage divider network previously shown, where the potentiometer's total resistance value needed to be precisely *twice* that of any single fixed-resistor value, when we cascade divider stages we must ensure that the total end-to-end resistance of one divider is precisely twice that of any single fixed resistor in the other divider. Here is an example:



The eleven 5 kΩ precision resistors in tandem with the 10 kΩ potentiometer form a sub-network with a total resistance of 50 kΩ end-to-end, which in turn is connected across two of the 25 kΩ resistors in the first divider sub-network, replacing what used to be a single 50 kΩ potentiometer in the original Kelvin-Varley voltage divider circuit. In the configuration shown, V_{out} is adjustable between 46% and 47 % of the input voltage (e.g. between 4.6 Volts and 4.7 Volts if the whole network were powered by a 10 Volt source). The movable connections between the first and second divider sub-networks provides the coarsest adjustment in 10% increments, the movable connections between the potentiometer and the second divider sub-network provides less-coarse adjustment in 1% increments, and the potentiometer wiper provides fine adjustment. Those movable connections are said to be *decade adjustments* because each one provides one-tenth the adjustment range of the other.

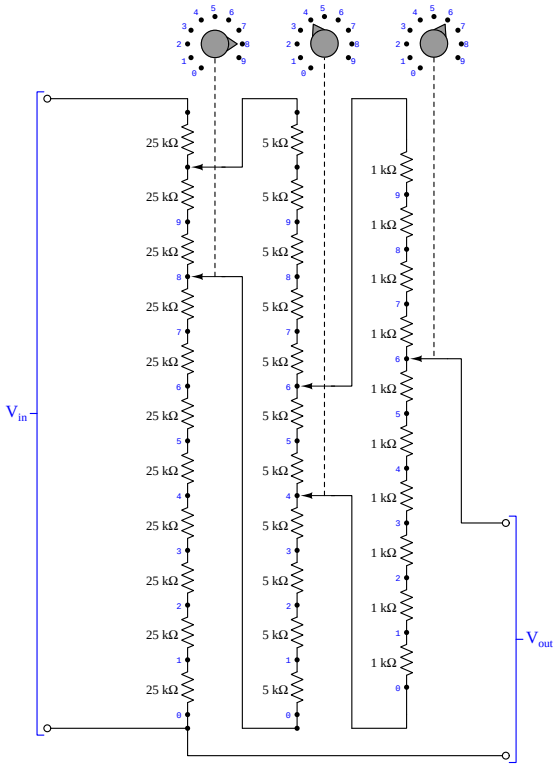
[22] Metrology is the science of precision measurement.

Extending this principle of cascaded voltage divider networks, we will now explore the behavior of a Kelvin-Varley voltage divider having *three* stages of "decade" division:



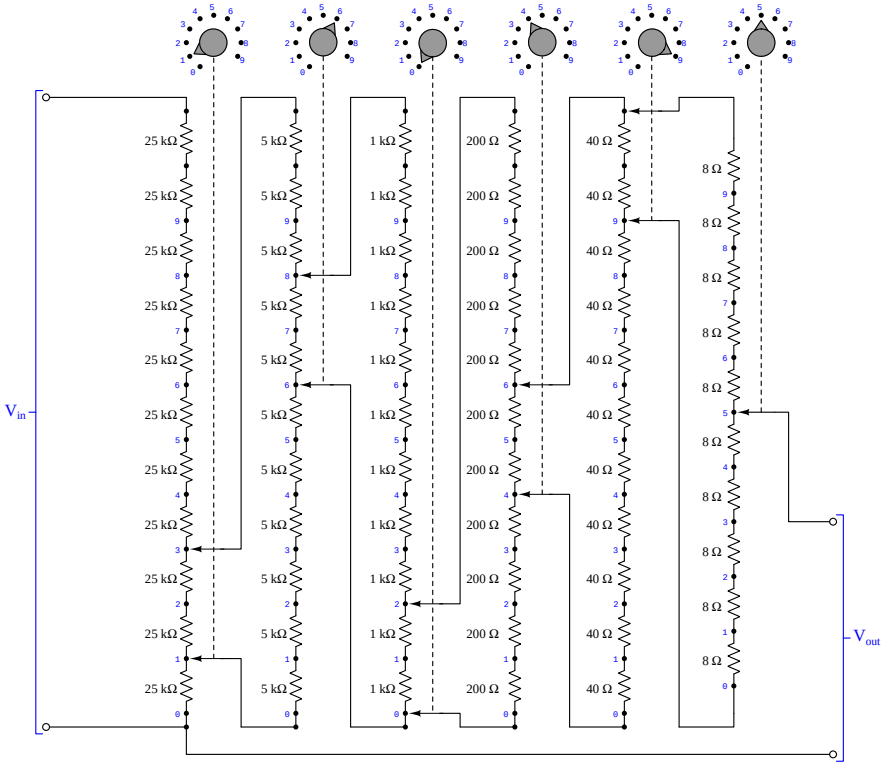
The eleven-resistor network comprised of fixed 1 kΩ resistors along with its 2 kΩ potentiometer connected across two of those forms a sub-network with an end-to-end resistance of 10 kΩ, suitable for connection across any two of the resistors in the middle sub-network (replacing the 10 kΩ potentiometer used previously). Those two sub-networks together form an end-to-end resistance of 50 kΩ suitable for connection across any two resistors of the first sub-network. Thus, the first set of movable contacts adjusts voltage in 10% increments, the second set adjusts in 1% increments, the third set adjusts in 0.1% increments, and finally the potentiometer's wiper provides fine adjustment.

As mentioned previously, Kelvin-Varley voltage divider networks are a common tool in the field of electrical metrology, where they are used to provide precise voltage-division ratios. For this type of application where we must precisely know the voltage-division ratio provided by the network in order for it to be of any practical use, it makes little sense to have a fine-adjustment potentiometer as the final stage unless that potentiometer had a graduated scale where we could accurately interpret its wiper position. Modern Kelvin-Varley voltage dividers are most often built with nothing but movable contacts forming multiple decades of discrete division ratios. Consider the following three-decade Kelvin-Varley voltage divider circuit:



Note how the first stage has eleven fixed resistors at 25 kΩ each, the second stage has eleven fixed resistors at 5 kΩ each, and the final stage has only *ten* fixed resistors at 1 kΩ each. As before, the first set of movable contacts provide adjustment in 10% increments, the second set of contacts provide 1% increments, and the final contact provides 0.1% increments. In the configuration shown above, the voltage-division ratio is 84.6%, which means if we were to power this entire network with a 10 Volt source our V_{out} voltage would be 8.46 Volts.

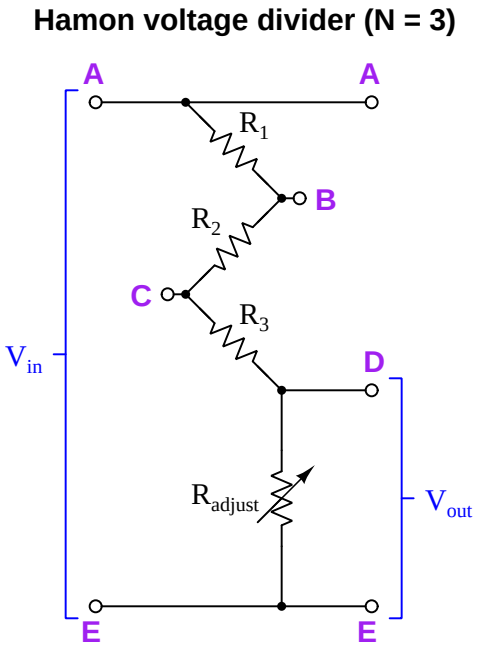
Here we see a six-decade Kelvin-Varley voltage divider, with its six rotary switches set to produce a voltage-division ratio of 16.0495% or 0.160495:



3.13.3 Hamon voltage dividers

Kelvin-Varley voltage dividers are extremely useful for their ability to set arbitrary voltage-division ratios with decimal precision, but in order to actually construct an accurate Kelvin-Varley divider one must begin with exceptionally precise and stable resistors to use for constructing all the series decade networks. If those resistors' values deviate from one another, it will corrupt the accuracy of any selected voltage-division ratio. If those resistors' values drift away from one another due to temperature changes, aging of the resistors, and/or wear, these errors will worsen as time goes on.

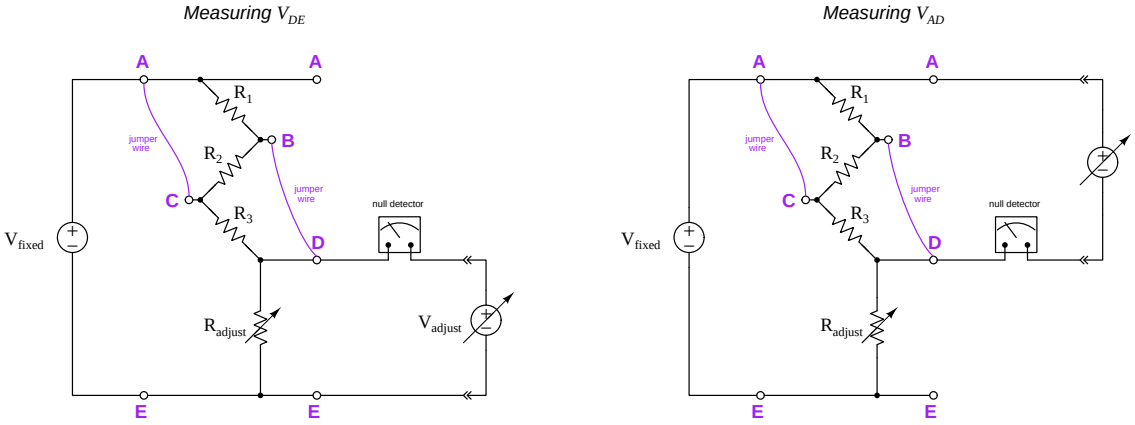
An ingenious solution to both of these problems is credited to B.V. Hamon, devising a voltage divider consisting of a set of equal-valued resistors which could be easily connected either in series or parallel, in series with a resistance that was adjustable. Here, we will explore a Hamon-style voltage divider consisting of three fixed resistors and one adjustable resistor:



The first step in using a Hamon voltage divider is to connect the fixed resistors in parallel with each other and then set the adjustable resistor to such a value that the entire network has a 1:2 voltage division ratio. This means connecting terminals A and C together, and also B and D together, applying a voltage between terminals A and E, and then setting R_{adjust} until voltage drop V_{AD} precisely equals voltage drop V_{DE}. When properly adjusted, R_{adjust} must be equal to $\frac{1}{\frac{1}{R_1}+\frac{1}{R_2}+\frac{1}{R_3}}$, and in this calibrated state the Hamon divider may be used productively as a voltage divider with a fixed 1:2 voltage-division ratio.

If then the jumper wires connecting terminals A and C together, as well as the wires connecting B and D together, are removed, the three fixed resistors will be in series with each other, their collective resistance now nine times greater than when connected in parallel. In this configuration, the Hamon voltage divider will have a voltage-division ratio of 1:10 and may be relied upon as such.

Self-calibration is easy to perform using nothing more precise than a null detector, which of course need only be accurate at one registered value: *zero Volts*. The Hamon divider simply needs to be tested in the following configurations, R_{adjust} moved in value until $V_{AD} = V_{DE}$ as judged by the null detector and an adjustable voltage source. The fixed voltage source value is largely irrelevant:



The user will know when R_{adjust} and V_{adjust} have been properly set when the null detector and adjustable voltage source may be moved between points A-D versus points D-E and the null detector registers zero with the adjustable source maintained at the same voltage. A regular DC voltmeter may be used at first to compare V_{AD} versus V_{DE}, trimming the value of R_{adjust} until approximately equal voltage readings are obtained, and after that the adjustable voltage source and null detector may be substituted for ultra-precise setting of R_{adjust}.

The practical value of the Hamon voltage divider design is that it is *self-calibrating*, in that the parallel resistance of R_1, R_2, and R_3 is first used as a standard for setting the value of R_{adjust}. After that, the other ratio (equal to $1:(n^2 + 1)$ where n is the number of fixed-value resistors in the network) may then be trusted on the basis of the prior configuration's self-calibration. This, of course, only offers *two* unique voltage-division ratios, but these are two ratios not dependent upon absolute resistor values which makes possible construction out of easier-to-obtain resistors. It also means that resistance drift over time is not a problem so long as the network is self-calibrated often enough. In other words, the Hamon voltage divider sacrifices the Kelvin-Varley's flexibility for gains in manufacturablity and accuracy.

In order to fully appreciate the insensitivity of the Hamon voltage divider to imperfectly-matched fixed resistors, the reader is recommended to perform their own calculations using values for R_1, R_2, and R_3 that are not *exactly* equal to each other, and not equal to any round-numbered resistance value at all. For example, if we were to choose 9.51 kΩ for R_1, 9.49 kΩ for R_2, and 9.52 kΩ for R_3 we would find after self-calibrating R_{adjust} for a perfect 1:2 voltage-division ratio that the nominal 1:10 voltage-division ratio is in fact 1:10.000015, constituting a mere 1.55 parts per million (ppm) deviation from the ideal 1:10 ratio. In other words, the final voltage-division ratio error for a Hamon voltage divider is *significantly* less than the deviation error of the fixed resistors when compared to their average. In this particular example with the 9.51, 9.49, and 9.52 kiloOhm resistors, the largest

deviation from average of the three three fixed-value resistors is −0.17532 percent, or −1753.2 ppm, while the calibrated voltage-division ratio error is only +1.55 ppm following the self-calibration: over a thousand times better accuracy than that of the worst individual fixed-value component!

3.14 Automated electrical testing

Nearly every electronic circuit requires functional testing at some point within its lifespan, and many additionally require calibration to ensure an adequate degree of functional accuracy. Manual testing or calibration of any electronic device is a matter of applying known input and/or power conditions and then measuring the output response(s), and this of course may be performed by a human technician adjusting signal sources and/or power supplies while monitoring the circuit outputs for proper response. However, some electronic circuits by their nature are laborious to test or calibration, while others may be laborious by dint of sheer volume as in the case of a modern high-volume assembly line. An extremely useful modern technique to ease the burden of this labor is to program another electronic device to perform the adjustments, monitoring, and data collection automatically. This is generally known as *automated testing*.

The following subsections describe different examples of automated testing.

3.14.1 Vector Network Analyzer calibration

Vector network analyzers or *VNAs* are specialized test devices used extensively for analysis of radio-frequency (RF) circuitry, outputting AC signals at precise voltage and frequencies while simultaneously measuring voltage and current in order to determine parameters of the circuitry being tested such as impedance, phase shift, and gain. Like all precision test instruments, VNAs are factory-calibrated and may be periodically re-calibrated to ensure their internal circuitry is accurately generating and accurately measuring RF signals. However, practical use of a VNA requires it be connected to any Device Under Test (DUT) through cabling and connectors custom to the test scenario, and the inclusion of these accessories necessarily adds potential error to the VNA's measurement. For this reason, anyone using a VNA must first perform a *user calibration* with all of those cables and connectors attached so their parasitic effects may be "calibrated out" and accurate results obtained.

VNAs are *multi-port* test instruments, which means they are capable of outputting a signal on one port (each port typically being a coaxial jack) and measuring the response signal on another. Each of these port's responses must be checked for accuracy, the most basic VNA calibration consisting of connecting a short, an open, and finally a resistive load to each port to test its ability to measure impedance over a specific range of frequencies. Multi-port testing requires one of the VNA's ports be connected directly to another with one port transmitting and the other port receiving.

The VNA itself analyzes its own state of error during a user calibration, but for a full multi-port calibration this still requires opens, shorts, resistive loads, and through-signal cables to be sequentially connected to and between the VNA's ports. To ease this multi-step calibration procedure there are VNA calibrator devices manufactured with ports of their own connected to each of the VNA's ports, and internal switches allowing those test ports to become opens, shorts, loads, and through-connections to other ports at the command of the VNA itself.

Below we see a two-port VNA, with ports labeled 1 and 2:

Note how a test cable has been connected to port 1 but not to port 2. In this particular case the VNA was only being used for a single-port test[23] where port 1 both provided the excitation signal to as well as measured the response from the DUT. Here, user calibration consisted solely of connecting a short, then an open, then a 50 Ω load to the end of that test cable while prompting the VNA to accept the results of each of those three tests. Such a three-point user calibration is easy enough to perform manually.

Consider, though, a two-port test where both ports 1 and 2 may alternately function as excitation sources and/or measurement inputs. A full user calibration for the VNA on both of its ports would require the following:

- Three S_{11} calibrations where port 1 is the signal source and also the measurement input: shorted, open, and resistively loaded at that port alone
- One S_{21} calibration where port 1 is the signal source and port 2 is the measurement input: a cable directly connecting one port to the other
- One S_{12} calibration where port 2 is the signal source and port 1 is the measurement input: a cable directly connecting one port to the other
- Three S_{22} calibrations where port 2 is the signal source and also the measurement input: shorted, open, and resistively loaded at that port alone

Such a user calibration would require no fewer than *seven* different cabling/termination configurations, and patience on the part of the technician to make those connections, prompt the

[23] In network theory we use so-called *S-parameters* ("scattering parameters") to identify the ports on a network where excitation signal is applied versus ports where some response is measured. For an antenna test where a VNA is used to check the impedance of a radio antenna over some range of frequency, the antenna only has one point of connection (port) and so we only use one port on the VNA. If that port is port number 1, then the S-parameter designation would be S_{11} which means the signal measured at port 1 and the excitation coming from port 1. For a filter network test where a VNA must inject a signal into one port of the filter and measure the output at another port, we would connect that filter network to the VNA's two ports and either perform an S_{21} test (filter output measured at port 2 while signal is input at port 1) or an S_{12} test (filter output at port 1 and excitation at port 2). If you closely examine the front panel of the VNA shown in the photograph you can see the S-parameter options helpfully diagrammed between the two ports.

VNA to run its own calibration test, and then wait for each test to complete before moving on to the next. A VNA with more than two ports[24] would of course have even more combinations and permutations of cabling and terminations to perform a full user calibration. Automated calibration of a VNA using a special electronic calibration standard would greatly reduce the amount of time and potential for technician error as opposed to manually terminating and un-terminating every cable all those times.

[24] A good test case necessitating a three-port VNA would be comprehensive testing of an RF *mixer* with two input ports (1 and 2) and one output port (3). A full test on a mixer would involve S_{11}, S_{22}, and S_{33} single-port impedance tests as well as S_{31}, S_{32}, S_{21}, and S_{12} "through signal" tests.

3.14.2 Industrial instrument calibration

Industrial control systems used extensively in heavy industry (e.g. chemical processing, power generation, general manufacturing, gas and liquid pipeline networks, etc.) require electronic sensors and signal-processing units to convert physical measurements of pressure, flow rate, level, temperature, and chemical composition into electronic signals that the control system is able to interpret and act upon. These sensors and signal converters must be routinely calibrated to ensure accurate measurement, a job typically performed by *instrumentation technicians*.

If there are only one or two devices in need of calibration, that task may be done well enough manually by a single technician. However, as often the case within heavy industry, opportunities to remove these devices from the functioning monitoring/control system are few, and this means routine calibration work is often done during times when the entire facility or system is shut down[25] for major maintenance. These infrequent shutdowns typically mean instrumentation technicians must calibrate large numbers of sensors and other instruments in a short period of time.

Special test instruments are made to automate calibration testing of industrial devices, one of them (a Fluke model 744 Documenting Process Calibrator) shown in the photograph below:

Calibrators such as this Fluke model have the ability to simultaneously output and input electronic signals to/from the DUT, incrementing the value of the output (stimulus) signal to the DUT while recording the resulting response and comparing that response against the pre-programmed expected response of the DUT. For example, in order to calibrate an electronic temperature "transmitter" device, the calibrator would output resistance or millivoltage signals to that transmitter mimicking the signals generated by RTD or thermocouple sensors, respectively, and then receive 4-20 milliAmpere signals from the transmitter that should linearly correspond to the simulated sensor signals. Once configured, the calibrator is able to do this repeatedly, ready to perform the same test on a set of identical DUTs, flagging any signal values falling outside of the tolerable range in order to let the technician know where DUT adjustments are required. The calibrator is also able to take this recorded data and upload it into a database of calibration records for archival. Any instrumentation technician faced with calibrating a dozen or more temperature transmitters will greatly appreciate the time saved by automating these calibrations in this manner.

[25] In the electric power industry these shutdowns are known as *planned outages*. In the oil industry these shutdowns are known as *turnarounds*.

3.14.3 Manufactured product testing

Modern electronics manufacturers use automated pick-and-place machines to place electronic components onto printed circuit boards (PCBs), wave-solder machines and reflow-solder ovens to solder those components on the PCBs, and a host of quality-control processes to ensure each of those populated PCBs functions properly. The high volume rate of production requires some form of automated testing of these electronic assemblies in order to keep up with the pace of the automated assembly processes.

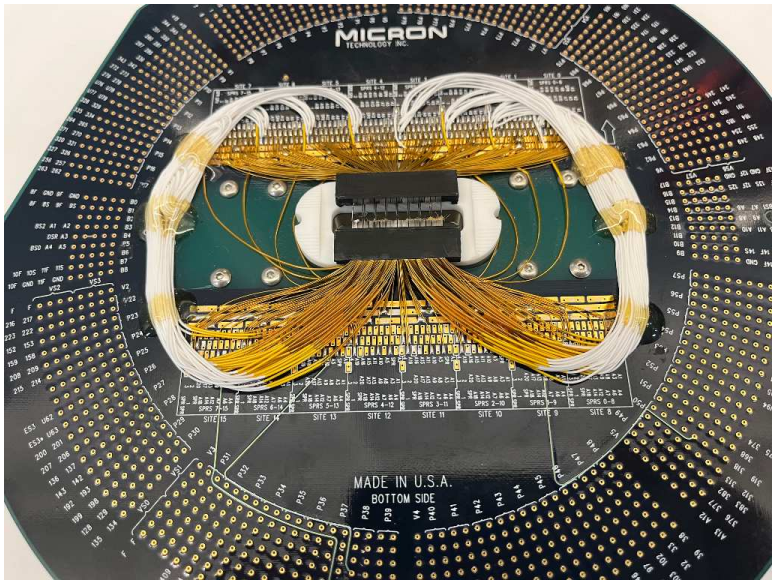
Such testing usually involves placing the finished PCB into a *test fixture* which then plugs into a *test station* where a set of test instruments apply electrical power and signals to the PCB while more test instruments monitor the PCB's output responses. These test instruments' functions are all controlled by a computer running a program, or *script*, directing each one what to do at the proper times, as well as collate the test data into a record and flag any points of non-conformance to functional criteria. The whole process may be made so automated, in fact, that even non-skilled personnel are able to execute the test, accepting PCBs that pass and rejecting those that don't.

One common style of test fixture is the so-called *bed-of-nails* where spring-loaded metal probes make contact with select test points on the manufactured PCB. The PCB is placed inside the fixture and clamped down to ensure adequate pressure on all the probe tips. Each probe connects to the test station with its own wire, enabling instruments in the test station to apply power and signals to the PCB while also measuring signals at other test points.

A more modern style of PCB testing dispenses with the fixture altogether in favor of robotic test probes moved into position by servo motors. This *flying probe* technique is able to rapidly inject and measure signals at individual component terminals as well as at test pads intentionally designed into the PCB, but suffers from the inability to make simultaneous contact with a multitude of test points as the bed-of-nails fixture is able to do, since space limits how many flying probes may hover over the PCB at any given time.

Semiconductor manufacturers use similar methods to test individual integrated-circuit *chips* etched onto a *wafer* of silicon prior to cutting those chips from the wafer and packaging them for final sale. Test fixtures with small wire probes are placed over the processed wafer, with those probes making electrical contact with wire-bonding pads on one or more of the individual chips. A test station applies appropriate power and signals to the IC while simultaneously measuring its responses (i.e. output signals), monitoring and recording any discrepancies. Defective IC chips are marked for rejection, while good chips are marked for acceptance. An automated diamond saw then cuts the individual IC chips from the wafer, sending the rejected chips to scrap and the accepted chips to final packaging and further testing.

Below we see a test fixture intended to automatically test digital memory ICs. In the center of this fixture you can see the fine-wire probes intended to make contact with the wire-bonding pads on the chips, with wires connecting these probe wires to pads on the circular printed circuit board, which in turn extends these electrical connections to other pads around the circumference of the PCB where the test station makes contact to the test instruments applying power and sending/receiving test signals:



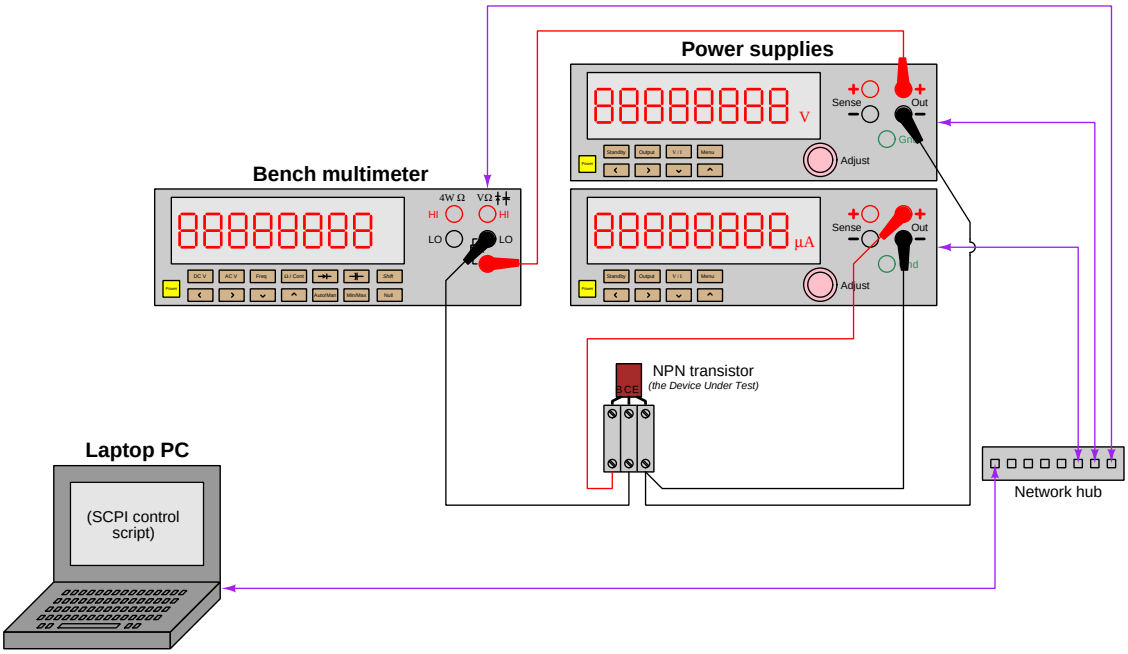
3.14.4 Component characterization testing

Another excellent application of automated electrical testing is component *characterization*. This is where we use test equipment to probe the characteristics of some electrical or electronic component. Some examples of characteristics we might wish to measure include:

- The forward voltage drop of a diode under different forward current values
- The collector current of a BJT for different base current values and different applied collector-emitter voltages, to produce a "family" of characteristic curves for that device
- The effect of power supply voltage on the output signal of an amplifier circuit (i.e. power supply rejection)
- The effect of common-mode input voltage on the output signal of a differential-input amplifier circuit (i.e. common-mode rejection)
- The effective turns ratio, polarity, phase shift, and/or magnetic saturation limits of a transformer for different amounts of excitation voltage, excitation frequency, etc.

Like all electronic testing, it is possible to perform component characterization in a manual fashion: a technician carefully adjusting input signals and/or power supply voltages and then measuring the results, many times over in order to gather enough data to plot a complete graph. However, it is generally more time-efficient to simply write a program (script) that a computer may execute to control electronic test equipment to do the same. That same script may also generate the graphical display to show that collected data!

Some specialized types of test equipment already offer this capability, and have for many years. A good example of this is a *curve tracer* which is useful for plotting the characteristic curves of semiconductor devices such as diodes and transistors. However, for custom testing applications it may be necessary to network multiple test instruments together and then sequence their operations using a computer program. The following illustration shows such an equipment assemblage consisting of two programmable DC power supplies (one to inject base current, the other to sweep collector-emitter voltage) and a digital benchtop multimeter for measuring the resulting transistor collector current:



The controlling script would set the microAmpere power supply to some small value of injected base current, then "sweep" the collector-emitter voltage power supply from 0 Volts to some pre-determined maximum value in a series of discrete steps, pausing for a short amount of time after each new V_{CE} value to let the multimeter's analog-to-digital converter settle to a stable value, then after that settling time capturing the meter's current indication and saving that to an array or list in the computer's memory along with the corresponding I_B and V_{CE} values. Then, the script would increment the base current value by some discrete step and re-sweep the V_{CE} power supply as before. This nested[26] sweep routine would continue until enough data points were collected to fully characterize the transistor.

Not only will automated testing remove the possibility of random errors introduced into the test results by human mistake(s), but it will also significantly speed up the testing process compared to doing everything manually.

[26] Implementing this in code would likely take the form of two `for()` loops, one nested inside the other. Each `for()` loop would start the swept variable at zero and increment in pre-determined steps as described.

A photograph of an automated test station for testing small electrical transformers emerging from an assembly line at a manufacturing facility appears below:

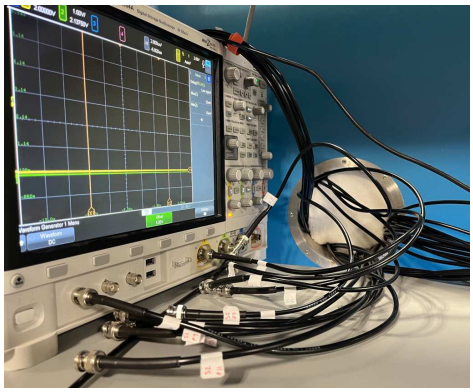
The aluminum box bearing the label "320-0509 FINAL TEST" is an example of a *test fixture* in which the transformer to be tested is placed. Inside this fixture is a holder for the transformer as well as electrical contacts to press against the metal terminals of the transformer. These contacts pass through a cable to the equipment located in the blue test station cabinet. A computer inside this cabinet runs custom software to sequence these test instruments and gather relevant data on the performance of each transformer before it is placed inside of a larger electronic assembly.

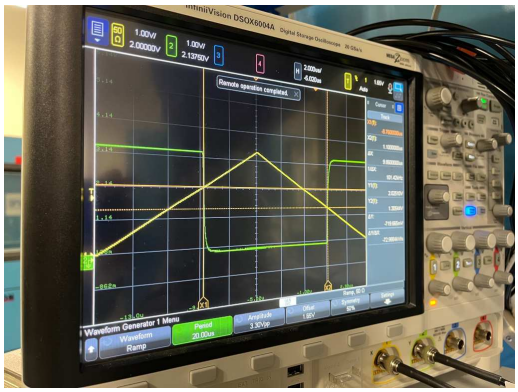
A hand-held barcode scanner is used to log each transformer's part number and associate that scanned part number with the test results, so that if in the future it were found to be faulty its pre-installed test results could be more easily located.

Some component characterizations must occur at different temperatures in order to ensure proper performance under extreme ambient conditions. For this purpose, *environmental test chambers* are made with built-in heating and/or cooling systems into which components and/or whole electronic assemblies are placed with cables extending to test instruments outside the chamber. Some environmental test chambers even have the ability to simulate wide ranges of humidity, which is important when component materials are permeable to water.

Shown below is a photograph of a heating/cooling test chamber with its access door partially opened to reveal the stainless-steel-lined interior:

Test equipment resides outside of the test chamber, in this case a programmable DC power supply and oscilloscope with built-in signal generator used to test the upper and lower threshold voltages of a Schmitt-trigger inverting ("NOT") logic gate by applying a triangle-shaped test input voltage to the gate and monitoring the gate's discrete output voltage signal. Coaxial cables conduct electrical signals and DC power to and from the logic gate test assembly inside the temperature-controlled chamber with the test equipment outside. A circular piece of thermal insulating foam blocks air from passing through the test chamber's access hole while still allowing the coaxial cables to pass through:





The oscilloscope tracks the correlation between input and output voltage signals while the test chamber varies in temperature between -45^{o}C and $+90^{o}$C, with the power supply energizing the gate at either 3.3 Volts or 5 Volts DC, so that the final test report shows the logic gate's threshold voltage values at these two power supply voltages for each of the three temperature values (hot, cold, and room temperature).

3.14.5 Accelerated stress testing

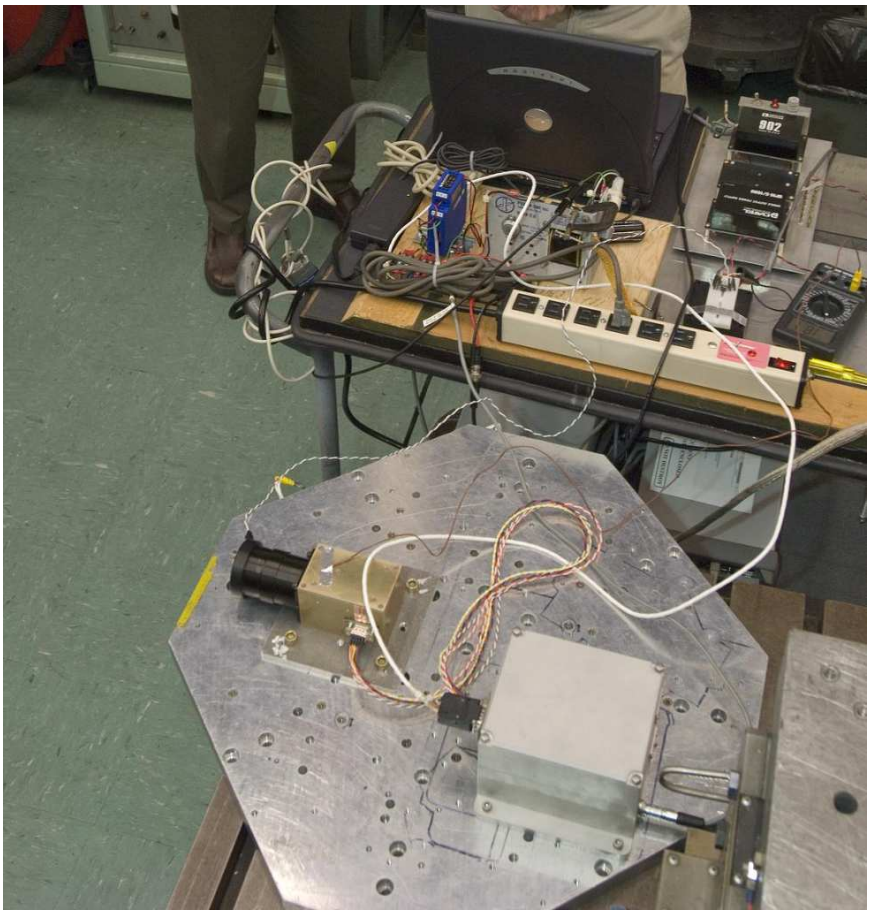
A common form of testing for electronic assemblies is *accelerated stress* or *accelerated life* testing, where the electronic assembly is intentionally subjected to a range of harsh conditions for the purpose of measuring how well it performs under unusually strenuous circumstances. Rapid cycling of these conditions may be performed in order to intentionally "age" the device, pushing it closer to end-of-life failure modes. Such conditions may include, but are not limited to:

- Mechanical vibration (while bolted to a "shaker table" apparatus)
- Temperature variations (very cold and/or very hot, inside of an "environmental test chamber")
- Humidity variations (very dry and/or very wet, inside of an "environmental test chamber")
- Ambient air pressure variations (vacuum to over-atmospheric pressure, inside of a "pressure test chamber")
- Electromagnetic field exposure (inside of an "RF test chamber")
- Nuclear radiation exposure (inside of a shielded "radiation test chamber")

As one might guess, even with accelerated aging caused by extreme exposures, the device under test (DUT) may still last days or weeks or even months before showing signs of age-related failure. Having human beings continually monitor an accelerated stress test is simply impractical for such lengthy test intervals, and so these sorts of tests are almost always automated in much the same manner as what we have explored previously: electronic test equipment connected to the DUT provides excitation signals, electrical power, and output signal measurements usually through networked connections so that a computer running some test program or script is able to control the test parameters. The major difference with accelerated stress testing is that the same computer is also controlling or at least monitoring the stress-inducing equipment as well.

Electronic test equipment controlling and monitoring any DUT inside of a test chamber of course will be located outside of that harsh environment. Cables passing through the test chamber's walls through *cable glands* connect the DUT inside the chamber to test equipment outside while maintaining a sealed environment for the test chamber itself.

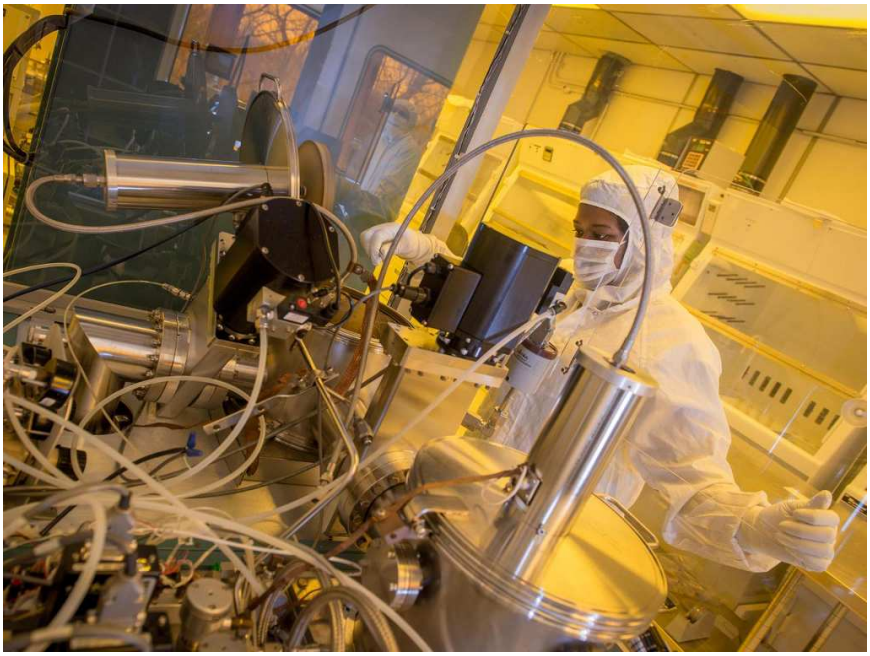
An example of an electronic system subjected to mechanical vibration is shown in the following image[27] taken at NASA's Ames Research Center near San Jose, California in April of 2007. The DUT is a total luminance photometer destined for the Lunar CRater Observation and Sensing Satellite (LCROSS). On the left we see the lens and photometer sensor unit, connected to an electronics package on the right. The eight-sided aluminum platform is the "shaker" table, attached to powerful actuators capable of shaking the unit in three dimensions under the control of a computer. Shake tests are necessary for any high-reliability system that will experience high-G forces in use, in this case including the launch procedure and the landing procedure:



An AC-DC power supply and various signal-interfacing hardware is visible on the workbench, along with a personal computer for collating the collected data. A shaker test is an example of an accelerated stress test that does *not* require a sealed chamber, as the only stresses involved in such a test are acceleration, not ambient temperature, pressure, humidity, or other factors requiring a sealed test environment.

[27] Public-domain image courtesy of NASA, photographer Dominic Hart.

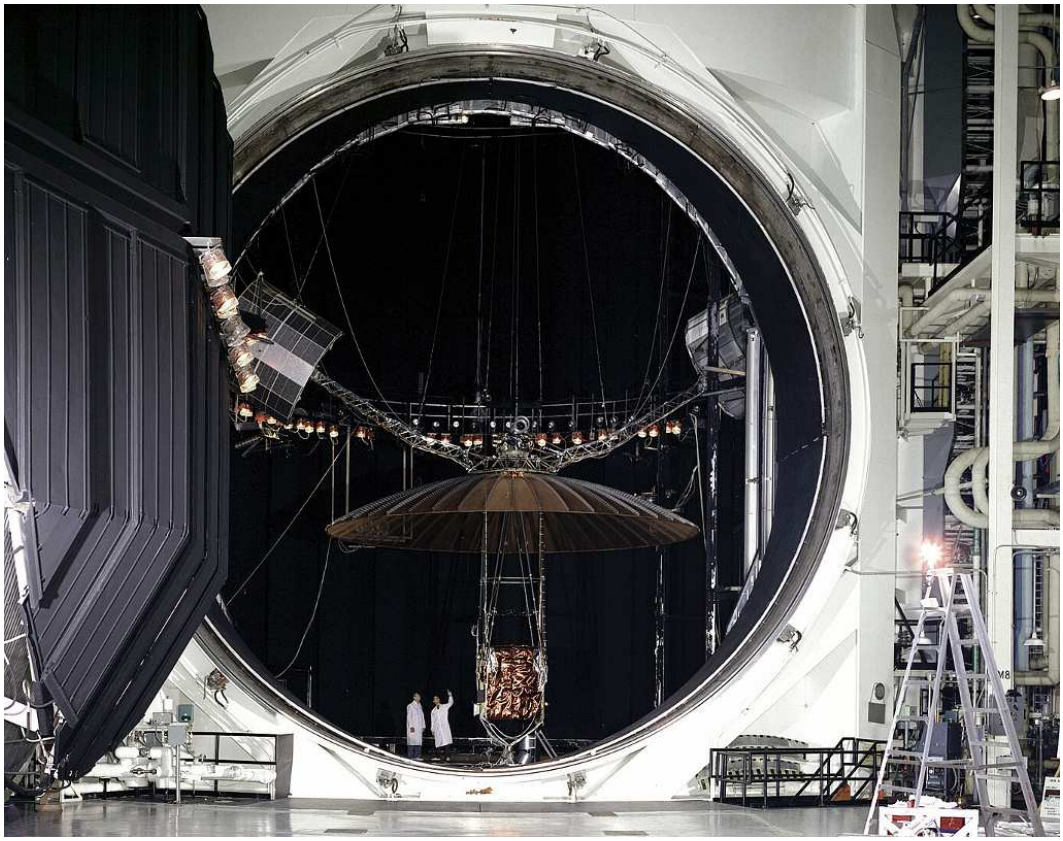
For tests subjecting the DUT to high or low pressures, high or low temperatures, etc. a sealed test chamber is absolutely necessary in order to prevent ambient atmospheric conditions from interfering. Below we see a photograph[28] of such a sealed test chamber being used at NASA's Glenn Research Center near Cleveland, Ohio in September of 2016. In this test chamber technician Ariana Miller is testing silicon carbide materials in ultra-high vacuum conditions simulating deep space:



Interestingly, the high-vacuum test chamber itself is located within a *clean room* where the air is ultra-purified of particulate matter. Clean rooms are standard manufacturing environments for semiconductor devices where dust and other particulate matter may cause catastrophic damage to the microscopic structures making up integrated circuits (ICs). This explains the smock and face mask Ariana is wearing in this photograph, intended to limit the release of skin particles, hairs, and other bodily matter that would otherwise contaminate the clean-room environment.

[28] Public-domain image courtesy of NASA.

A rather extreme example of an environmental test chamber appears in the following photograph[29] showing Chamber A of the Space Environment Simulation Laboratory at NASA's Johnson Space Center near Houston, Texas in September of 1973:



Environmental test chambers built for testing satellites and spacecraft subject those complex machines to high-vacuum conditions, high and low temperature extremes, nuclear radiation, etc. in order to simulate the rigors of space travel. Vacuum conditions are relevant to electronic system testing for multiple reasons, including lack of convective cooling for heat-producing components, changes in the maximum amount of voltage between separated conductors under low-pressure conditions, and trapped pressures inside electronic device packages that may impart physical stress to the packaging when the surrounding pressure falls to near-zero. Temperature cycling is relevant to spacecraft electronics because in the absence of surrounding air the craft's temperature while facing the sun will soar to extremely hot conditions while the side shaded from sunlight will become extremely cold. Any rotating or Earth-orbiting craft will therefore experience violent swings in temperature, causing mechanical stresses in electronic components as well as influencing electrical properties such as resistance, PN junction voltage drop, and transistor amplification parameters (e.g. beta for BJTs and transconductance for FETs).

[29] Public-domain image courtesy of NASA.

3.14.6 Programmable test instruments

Many of the performance and calibration tests described so far in this section rely on electronic test instruments that may be controlled remotely through digital networks by a computer executing a test program or test script, but so far we have not discussed these instruments or their programming in any detail. Here we will explore some of the network and programming technologies used.

For many years now professional-grade benchtop electronic test instruments such as power supplies, signal generators, oscilloscopes, impedance analyzers, etc. have offered digital networking capability allowing remote control by other digital electronic devices such as personal computers. This remote-control capability permits anyone to build their own test station and sequence the testing operations using simple programs (or "scripts") running on a personal computer.

The following photograph shows the rear panel of a Keysight model 34450A benchtop multimeter, equipped with two networking ports. Above my fingertip we see a GPIB (also known as HP-IB or IEEE-488) port while below and to the right of that port is a type-B USB port:

GPIB ports are definitely a legacy technology at the time of this writing (2024), requiring expensive adapters to connect to modern personal computing hardware. USB, however, is an extremely popular serial interface at this point in time and is the preferred medium for remote instrument control on lower-cost instruments. More expensive test instruments may also have Ethernet ports which of course offer far higher data communication rates than either GPIB or USB.

Some instrument manufacturers provide their own customized software for computer-control of the instruments they manufacture, making simple remote-control functions relatively easy to implement. However, for more sophisticated testing it is better to program a computer in some general-purpose programming language such as C, C++, Python, Java, Lua, etc. to remotely control instruments, and in order to simplify this task when multiple instruments (often of different model and manufacture) connect to a single controlling computer those instruments should all "speak" some common language.

One of these common languages was first introduced in 1987 as part of the IEEE 488.2 GPIB specification "Standard Codes, Formats, Protocols, and Common Commands". This command language exists at the 7th layer of the OSI model, complementing the IEEE-488 physical (layer 1) and data link (layer 2) layers, and is now popularly referred to as *SCPI*[30] ("Standard Commands for Programmable Instruments"). The SCPI standard defines common instrument-control commands as ASCII-encoded character strings such as `OUTP` (commanding an instrument to output a signal) and `CONF:VOLTage:DC` (commanding a measuring instrument to switch to the DC voltage mode). Being a layer-7 protocol, SCPI commands may be communicated over *any* kind of digital network, not just IEEE-488 (GPIB). At the time of this writing it is quite commonplace to issue SCPI commands over USB serial networks instead, or even via Ethernet.

A communications protocol suite known as *LXI* (LAN eXtensions for Instrumentation) was developed in 2005 and is based on the ubiquitous networking standards of Ethernet (layers 1 and 2), IPv4/IPv6 (layer 3), and TCP (layer 4) to exchange SCPI commands between a control computer and the test instrument(s). LXI seeks to avoid the problems[31] often encountered with establishing reliable USB connections between test instruments, as well as the expense and cumbersome connectors/cables associated with GPIB (IEEE-488) by leveraging the robust capabilities of Ethernet and TCP/IP. LXI devices also support web server technology, which means you are able to configure important instrument settings and even perform simple control functions using nothing more than standard web-browser software.

The exact commands (SCPI or otherwise) needed to instruct a programmable instrument to perform its various test functions are documented in a *programming guide* provided by that instrument's manufacturer. These documents show the proper command syntax, which often varies with the test functionality of each instrument.

[30] Pronounced "skippy".

[31] Anyone who has tried to establish communication between an electronic test instrument and a personal computer running a Python script can attest to how non-intuitive this process is! USB was developed for personal computers as "plug and play" technology, which is true when the USB device in question has a robust driver and interface software installed on the computer. There is no automated discovery process for USB devices here, though, which necessitates some "manual" prodding with Python commands simply to reveal the resource name of each connected device, and then getting all communication settings (e.g. `query_delay` timeout parameter, `write_termination` character) correct so as to avoid error messages is often a matter of trial and error.

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 Standard voltage cells

The earliest practical standards for voltage were chemical cells known to output extremely stable voltages. The book *Electrical Measurements In Practice* written by F. Malcolm Farmer in 1917 describes three different types of standard cells on pages 37-39:

> 43. Clark Cell.
>
> The Clark cell was the first form of voltaic cell to meet the requirements of a reliable standard of e.m.f., namely, reproductibility and reasonable permanence. The elements are: positive electrode, metallic mercury, negative electrode, amalgamated zinc; electrolyte, saturated solution of zinc sulphate and mercurous sulphate (Fig. 20). Its e.m.f. was originally taken as 1.434 international volts at 15^{o}C., which value was made legal in 1894 as stated above. Later and more careful determinations have shown, however, that 1.4328 is a more correct figure.

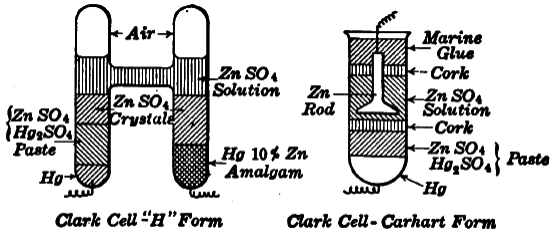


FIG. 20.

The e.m.f. of the Clark cell changes with temperature in accordance with the formula:

$$E = 1.4328[1 - 0.00077(t - 15)]$$

where E = international volts at any temperature, t, and t = deg. C.

With the development of the art, more precise measurements became necessary and certain characteristics of the Clark cell made it unsatisfactory as a standard for highly precise work. The principal defects are the relatively large temperature coefficient which necessitates the determination of the true temperature of the cell; and the time lag between the temperature and the e.m.f. Thus great care is necessary to insure that the e.m.f. has reached the value corresponding to the temperature of the cell. Because of these objectionable features, the Clark cell, while still the official standard, has been superseded in practice by the Weston cell which is practically free from these objections.

The elements of the Weston cell are essentially: positive electrode, metallic mercury; negative electrode, amalgamated cadmium; electrolyte, solution of cadmium sulphate and mercurous sulphate. There are two forms – in one the solution is saturated, and in the other it is unsaturated above 4^{o}C.

44. Weston Cell, Normal.

The form in which the electrolyte is saturated is called the Weston normal cell (Fig. 21), which is the form that has been adopted as the official standard cell of the Bureau of Standards[1] because it is more permanent and can be reproduced with greater accuracy than when the electrolyte is unsaturated. When carefully made according to the official specifications, cells will agree with each other within a few parts in 100,000. The e.m.f. changes slightly with temperature according to the following formula:

$$E = 1.0183[1 - 0.000041(t - 20)]$$

where E = international volts at any temperature, t, and t = deg. C.

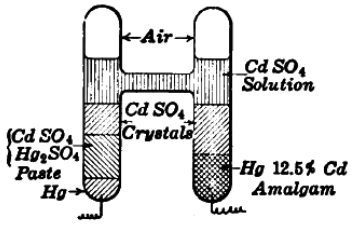


FIG. 21.

45. Weston Cell, Unsaturated.

In this form the electrolyte is saturated at 4^oC. and therefore it is unsaturated at all ordinary temperatures. The unsaturated cell is not so uniform in e.m.f. as the saturated form but the temperature coefficient is only 0.00001 per degree C. which is ordinarily negligible. Therefore, while the unsaturated cell has to be standardized, it is more convenient for general use as a secondary standard.

The cells furnished by the Weston Electrical Instrument Co. are of the unsaturated type. The manufacturers recommend that their cells be not subjected to temperatures below 4^oC. or above 40^oC. and that no current greater than 0.0001 amp. be passed through them.

[1] See Circular No. 29, Bureau of Standards. The value of the e.m.f. of the Weston normal cell was established by the Bureau on Jan. 1, 1911 as 1.01830 international volts at 20^oC.

4.3 Standard resistors

The book *Electrical Measurements In Practice* written by F. Malcolm Farmer in 1917 describes various standards for electrical resistance on pages 120-123:

> 138. Practical Unit and Standard
>
> The practical unit of resistance is the international ohm. The primary standard representing this unit is the resistance of a column of mercury of uniform cross-section, 106.300 cm. long, 14.4521 grams mass and at the temperature of melting ice. This standard is necessarily difficult to construct, maintain and use so that in actual measurements metallic resistors which have been standardized by comparison with the primary standard are employed. The primary standard will in fact be found only in the laboratories of the custodians of the electrical standards in the several countries.
>
> SECONDARY STANDARDS
>
> 139. General.
>
> Secondary standards are made with metal of high specific resistance in the form of wire or ribbon. Manganin, a copper-nickel-manganese alloy, is generally used because, when properly treated and aged, it meets the necessary requirements. These requirements are: permanent electrical and physical characteristics; low thermo e.m.f. against copper ; small temperature coefficient of resistance; relatively high specific resistance. The completed standard must, in addition, be unaffected by immersion in oil, or by changes in atmospheric conditions.
>
> In general, resistance standards may be divided into two classes: standards of resistance, or those used primarily for the measurement of resistance; and current standards or those intended primarily for the measurement of current.

140. Standards of Resistance.

Standards of resistance are usually made with very small current-carrying capacity. There are two forms in general use, the Reichsanstalt and the N.B.S. (National Bureau of Standards)[2].

The Reichsanstalt form is shown, partially in section in Fig. 77. The N. B. S. form is shown in Fig. 78. The distinctive features of the latter form are that it is immersed in oil and hermetically sealed. This prevents the absorption of moisture by the shellac, with which the resistance wire or strip is coated, and the consequent elongation of the fine wire used in the higher resistors[3]. Both forms are intended to be hung from mercury cups by means of lugs, and may be suspended in an oil bath in order to measure the temperature more accurately. The N.B.S. form is made only in sizes higher than 1 ohm.

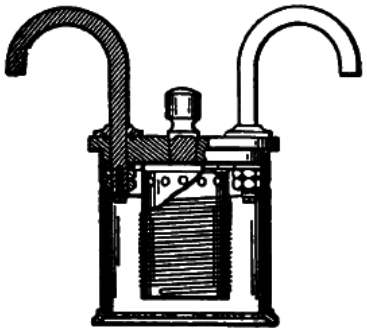


FIG. 77.

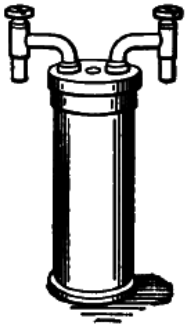


FIG. 78.

141. Current Standards.

Current standards are made in two types, the Reichsanstalt and air-cooled. The Reichsanstalt standards are made in two forms, the small pattern for moderate currents, and the large pattern (in low resistances) for large currents. The small-pattern form is similar in appearance to Fig. 77 except that, for 1 ohm and less, separate potential taps are provided. They are suspended from mercury cups in an oil bath for cooling purposes. Obviously, this type may be used also as a resistance standard. The following table shows the current rating assigned by Otto Wolff, and the Leeds and Northrup Co.

[2]"A New Form of Standard Resistance," E. B. Rosa, Bulletin, Bureau of Standards, vol . 5, p. 413 (1908).

[3]"The Variation of Resistances with Atmospheric Humidity," E. B. Rosa and H. D. BABCOCK, Bulletin, Bureau of Standards, vol. 4, p. 121 (1907-1908).

Current Rating of Standard Resistors, Reichsanstalt Type
Permissible Current Capacity in Amperes[1]

Nominal resistance, ohms	When used for resistance measurements (accuracy 0.01–0.02 per cent.)		When used for current measurements (accuracy 0.02–0.04 per cent.)	
	Still air	Still oil	Still air	Still oil
100,000.0	0.002	0.003	0.005	0.01
10,000.0	0.005	0.01	0.015	0.03
1,000.0	0.015	0.03	0.05	0.1
100.0	0.05	0.1	0.15	0.3
10.0	0.15	0.3	0.5	1.0
1.0	0.5	1.0	1.5	3.0
0.1	1.6	3.0	5.0	10.0
0.01	5.0	10.0	15.0	30.0
0.001	16.0	30.0	50.0	100.0
0.0001	55.0	110.0	165.0	330.0

The large pattern form as made by Wolff, in 0.001-ohm, 0.0001-ohm and 0.00001-ohm sizes is intended for very large currents. The resistance element is permanently mounted in a tank of oil, which is water-cooled. Connection with the source of current is made directly to heavy copper blocks. The potential terminals are separate and tap directly on to the resistance metal. Large pattern standards of low resistance have capacities from 100 watts to 2,500 watts and over.

Air-cooled current standards employ sufficient surface to permit of use in air without excessive temperature rise. While they are not as accurate or as reliable as the Reichsanstalt form, they are amply satisfactory for much commercial work and are especially convenient where oil baths would be inconvenient. Fig. 79 shows a Leeds and Northrup Co. 0.00002-ohm resistor of 2,000 amp . capacity for which an accuracy of 0.04 per cent. is claimed.

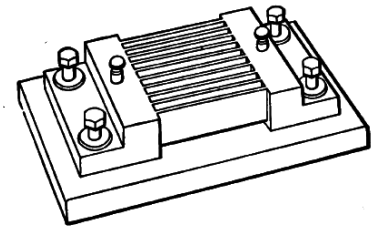


Fig. 79.

142. Precautions in the Use of Standard Resistors.

Care should be taken that standard resistors are not overheated and that the oil used is free from acid. Both the current terminals and the potential terminals should be kept perfectly clean. Poor contact at the current terminals of large-capacity standards will produce excessive heating which may permanently change the resistance. Poor contact at the potential terminals will introduce resistance in the measuring circuit.

In resistors smaller than 10 ohms used for current measurements, the potential drop should be measured between points some distance inside of the current terminals. This permits more convenient adjustment of the resistance, eliminates the joint resistance between the terminal lug and the resistance metal, and insures uniform distribution of the current at all values in that part of the resistance metal across which the potential is measured.

4.4 A binary resistance box

The Scottish physicist James Clerk Maxwell (1831-1879) wrote a book entitled *A Treatise on Electricity and Magnetism*, within which he describes the construction and use of a "resistance box" which could be used to simulate any number of different resistance values by shorting past certain resistive wire coils in particular patterns.

> There are various arrangements by which resistance coils may be easily introduced into a circuit.
>
> For instance, a series of coils of which the resistances are 1, 2, 4, 8, 16, &c, arranged according to the powers of 2, may be placed in a box in series.
>
> The electrodes consist of stout brass plates, so arranged on the outside of the box that by inserting a brass plug or wedge between two of them as a shunt, the resistance of the corresponding coil may be put out of the circuit. This arrangement was introduced by Siemens.
>
> Each interval between the electrodes is marked with the resistance of the corresponding coil, so that if we wish to make

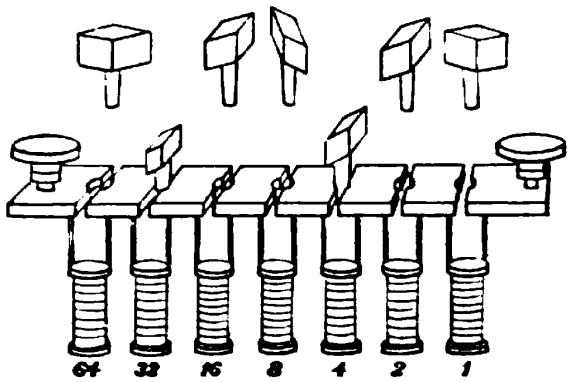


Fig. 29.

> the resistance in the box equal to 107 we express 107 in the binary scale as 64 + 32 + 8 + 2 + 1 or 1101011. We then take the plugs out of the holes corresponding to 64, 32, 8, 2 and 1, and leave the plugs in 16 and 4.
>
> This method, founded on the binary scale, is that in which the smallest number of separate coils is needed, and it is also that which can be most readily tested. For if we have another coil equal to 1 we can test the quality of 1 and 1', then that of 1 + 1 and 2, then that of 1 + 1' + 2 and 4, and so on.
>
> The only disadvantage of the arrangement is that it requires a familiarity with the binary scale of notation, which is not generally possessed by those accustomed to express every number in the decimal scale. [page 470]

Maxwell continues by giving an example of a different kind of resistance box also employing binary patterns. In this design, each resistance coil is labeled with a resistance value (in Ohms) as before, but with this design the insertion of brass plugs connects these resistances in parallel with each other to produce greater values of total *conductance* (measured in units of Siemens, or *Mhos*). Recall that conductance (G) is the reciprocal of resistance ($G = \frac{1}{R}$), so that a resistance of 2 Ohms would be a conductance of 0.5 Siemens, a resistance of 8 Ohms would be a conductance of 0.125 Siemens, etc. Just as resistances add together when connected in series, conductances add together when connected in parallel:

> A box of resistance coils may be arranged in a different way to the purpose of measuring conductivities instead of resistances. [page 470]
>
> The coils are placed so that one end of each is connected with a long thick piece of metal which forms one electrode of the box, and the other end is connected with a stout piece of brass plate as in the former case.
>
> The other electrode of the box is a long brass plate, such that by inserting brass plugs between it and the electrodes of the coils it may be connected to the first electrode through any given set of coils. The conductivity of the box is then the sum of the conductivities of the coils.

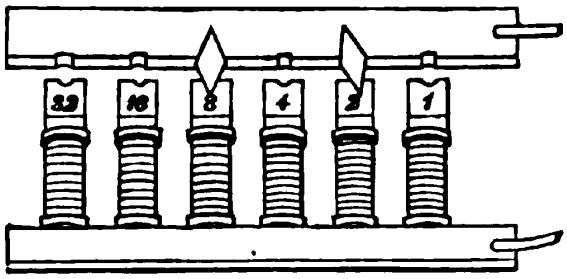


Fig. 30.

> In the figure, in which the resistances of the coils are 1, 2, 4, &c, and the plugs are inserted at 2 and 8, the conductivity of the box is $\frac{1}{2} + \frac{1}{8} = \frac{5}{8}$, and the resistance of the box is therefore $\frac{8}{5}$ or 1.6.
>
> This method of combining resistance coils for the measurement of fractional resistances was introduced by Sir W. Thomson under the name of the method of multiple arcs. See Art. 276. [page 471]

4.5 Early references to Wheatstone bridges

The Scottish physicist James Clerk Maxwell (1831-1879) wrote a book entitled *A Treatise on Electricity and Magnetism*, within which he describes two different methods of making comparative resistance measurements.

The first method used a special instrument called a *differential galvanometer*, consisting of two identical wire coils flanking a magnetized iron needle suspended on a thread. The needle will be deflected (i.e. turned to a different position) by the magnetic force generated by current passing through either wire coil. If equal currents are passed through the two coils, in different directions so as to produce opposing magnetic fields, the resultant force on the needle will be zero and no deflection will result. Thus, the differential galvanometer registers any *difference* in value between two electric currents: any inequality makes the needle move, while precisely equal currents makes no motion at all.

When the two coils of a differential galvanometer are connected in series with two resistances to be compared, and each of these series networks powered by a common voltage source (battery), the galvanometer needle will deflect if there is any difference between the two resistance values. If the two resistances are exactly equal and the two galvanometer coils are precisely identical to each other, then no deflection will result. The beauty of this technique is that it remains accurate even if the battery voltage is uncertain or even unstable over time. Maxwell describes the utility of measuring resistance in this manner:

> The merit of the method consists in the fact that the thing observed is the absence of any [sic] deflexion, or in other words, the method is a Null method, one in which the non-existence of a force is asserted from an observation in which the force, it it had been different from zero by more than a certain small amount, would have produced an observable effect. [page 473]
>
> Null methods are of great value where they can be employed, but they can only be employed where we can cause two equal and opposite quantities of the same kind to enter into the experiment together. [page 474]

By testing for the *absence* of galvanometer needle motion rather than some precise amount of needle motion, there is no need for absolute measurements of either voltage or current as there would be if we were inferring resistance from Ohm's Law ($R = \frac{V}{I}$).

A weakness of this differential galvanometer method is the possibility that the two wire coils within the galvanometer might not be precisely identical to one another. If those two wire coils differed at all, either in the number of turns, the amount of wire resistance, their relative positions from the suspended needle, etc., the needle would in fact move slightly when exposed to two identical currents and therefore it would not be reliable as a "null" instrument.

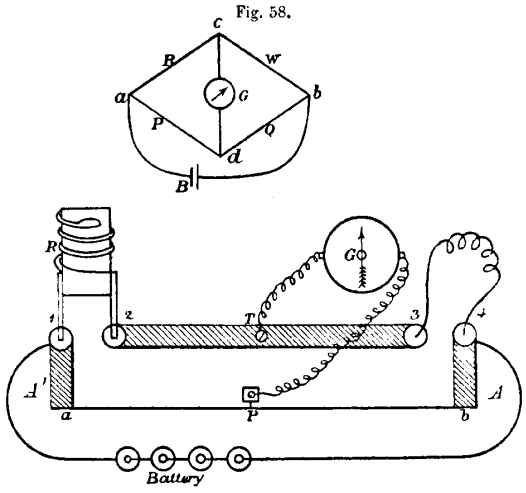
After describing the use of a differential galvanometer to make comparative resistance measurements, Maxwell then proceeds to describe the Wheatstone bridge method which escapes this weakness of differential galvanometers:

> The other null method, in which Wheatstone's Bridge is used, requires only an ordinary galvanometer, and the observed zero deflexion of the needle is due, not to the opposing action of two currents, but to the non-existence of a current in the wire. Hence we have not merely a null deflexion, but a null current as the phenomena observed, and no errors can arise from want of regularity or change of any kind in the coils of the galvanometer. The galvanometer is only required to be sensitive enough to detect the existence and direction of a current, without in any way determining its value or comparing its value with that of another current. [page 475]

It may seem strange to us living in the twenty-first century that so much trouble would be taken to perform a task as simple as measuring resistance, but this is only because we take for granted the existence of precise, accurate, and traceable standards for such things as voltage, current, and resistance measurement. In Maxwell's time none of these were a given, and the laboratory experimenter needed to devise clever means of making precise measurements using what we would consider primitive equipment. The Wheatstone bridge was a clever solution to the problem of precise resistance measurement, necessitating only an arbitrarily sensitive meter (galvanometer) and some trusted resistance specimen for comparison.

Another early reference to the use of Wheatstone bridge circuits for resistance measurement comes from John Ambrose Fleming's book *Short Lectures to Electrical Artisans* written in 1886. This text gives a fairly detailed explanation of the bridge circuit's function and operation. Bear in mind that this was in an ear before the advent of multimeters, when stable-voltage batteries were rare, when slide-wire potentiometers needed to be fabricated out of resistive wire, and when the most secure means of joining conductors together with a minimum of resistance and stray voltage was by dipping wires into cups of liquid mercury:

> . . . The well-known Wheatstone's bridge is the most usual appliance for comparing resistances.
>
> As is well known, the method was not invented by Wheatstone, but my Mr. Christie, in 1833[4]. It consists essentially of an arrangement of six conductors, joining four points. One of these conductors is a battery circuit, and the other a galvanometer. A Wheatstone bridge, suitable for workshop purposes, may be made as follows: – On a board about 4 feet long fasten down a paper scale of inches and tenths, and over the paper scale stretch a fine uniformly drawn German [page 142]



Wheatstone's Bridge.

> silver wire. The ends of the wire are soldered to thick copper strips A A' (see Fig. 58), which terminate in mercury cups 1, 4. Between these extreme cups there is a strip of copper B about 1 inch wide and one-sixteenth of an inch thick, having mercury cups 2, 3, at its ends. There is also a terminal T soldered in the middle.
>
> Prepare five resistance coils, as described above, of approximately 1, 5, 10, 50, 100 ohms, and let them, if possible, be adjusted by some one who possesses known standards. These coils are made so that their copper legs will dip into the mercury cups 1, 2. A battery of

[4] Footnote from Fleming's text: See Wheatstone's Scientific Papers, p. 129, published by the Physical Society of London.

> half-a-dozen Leclanché cells is also needed, and a delicate astatic galvanometer. [page 143]
>
> Insert the standard coil in the cups 1, 2, the ends of the wire whose resistance is to be measured in the cups 3, 4. Connect the battery to the cups 1, 4, one terminal of the galvanometer to the terminal T, and take the other end of the galvanometer wire in the hand, and just touch it to some point on the German silver wire. The galvanometer will in general make a deflection, but by sliding the galvanometer wire along the German silver wire a point P can be found at which there is no deflection right or left. When this is the case, note the lengths of the German silver wire Pa and Pb on either side of the point of no current. Then, as length Pa is to length Pb, so is the value of the standard of resistance R to the resistance of the wire to be measured to. Hence a simple rule of four sum gives us the required resistance.
>
> To understand the reason for this rule, we see, if we refer to the upper diagram in Fig. 58, that the battery current from B entering at a has two routes open to it to pass to b, one by the resistances R and W, and the other by the resistances P and Q. Down each of these paths there is a gradient of pressure or fall of potential, and when the "bridge is balanced," that is when a point is found on each path, such that no current will flow between them when they are connected by a galvanometer, it follows that the pressure at c is equal to that at d. Now the fall in potential in passing from a to c along R must be equal to the fall in potential in going from a to d along P, and similarly the fall of pressure from c to b is equal to that from d to b. Now just as in the first method given of measuring resistance we see that the ratio of resistance R to resistance W is equal to that ratio of fall of potential between a c and that between c b, and this ratio is equal to that of fall of potential between a d to that between d b, which in its turn is equal to the ratio of the resistance P to Q. Accordingly when the bridge is balanced we have the proportion – [page 144]
>
> As resistance R is to resistance W, so is resistance P to resistance Q, and any three of these being given we can find the fourth. [page 145]

Sliding the potentiometer's lead along the resistive German silver wire (i.e. a home-made *potentiometer*) allowed the user to adjust the bridge to obtain a condition of balance indicated by a zero ("null") reading on the galvanometer. Once balance was obtained, one could read the position of the sliding wire (at point "P") along the graduated scale laid down beside the wire, and read its position as a fraction, or ratio, of the scale's total length. This ratio would then indicate the ratio of the test-wire's resistance to the "standard coil" resistance. Fleming mentions the preparation of five such standard-resistance coils which would be used as standards of comparison against the wire of unknown resistance, implying that one might need to try several of these coils before finding a reasonable balance point on the bridge, and it is here we see the wisdom of using mercury-filled cups as points of contact for the standard and for the test-wire: the liquid mercury would make a very repeatable low-resistance connection with any wire dipped into it, even after repeated trials. Any attempt to use a spring clip, screw, or any other "modern" connection method to attach the standard coils to the bridge would have resulted in slight variations in connection resistance with each trial, whereas a mercury cup would be very consistent because liquid mercury is an excellent conductor of electricity and its ability to flow would guarantee optimum contact with the metal wire every time.

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.

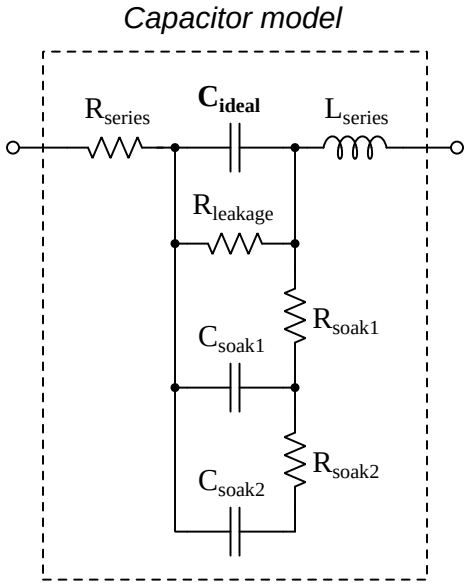
E3	**E6**	**E12**	**E24**
10	10	10	10
			11
		12	12
			13
	15	15	15
			16
		18	18
			20
22	22	22	22
			24
		27	27
			30
	33	33	33
			36
		39	39
			43
47	47	47	47
			51
		56	56
			62
	68	68	68
			75
		82	82
			91

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

5.2 Capacitor parasitics

5.2.1 Model of a real capacitor

An ideal capacitor exhibits only capacitance, with no inductance, resistance, or other characteristics to interfere. Real capacitors exhibit all these phenomena to varying degrees, and we collectively refer to these undesirable traits as *parasitic effects*. The following diagram models some of the parasitic effects observed in real capacitors:



In addition to the capacitance the capacitor is supposed to exhibit (C_{ideal}), the capacitor also has parasitic resistance (R_{series}, also known as *Equivalent Series Resistance*, or *ESR*), parasitic inductance (L_{series}), and additional energy storage in the form of *soakage* (also known as *dielectric absorption*) whereby the dielectric substance itself absorbs and releases energy after relatively long periods of time compared to the main (ideal) capacitance.

Some of these parasitic effects – such as leakage resistance and soakage – affect the capacitor's performance in DC applications. Most of the other parasitic effects cause problems in AC and pulsed applications. For example, the effective series capacitor-inductor combination formed by C_{ideal} and L_{series} will cause *resonance* to occur at a particular AC frequency, resulting in much less reactance at that frequency than what would be predicted by the capacitive reactance formula $X_C = \frac{1}{2\pi fC}$.

Next we will explore common mechanisms for each of these effects.

5.2.2 Parasitic resistance in capacitors

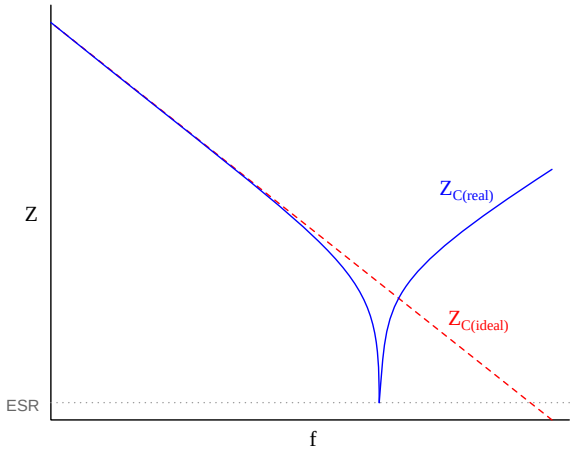
Wire resistance, of course, plays a part in this parasitic effect, but this is generally very small due to the short conductor lengths we typically see with capacitors. More significant is *dielectric losses* – energy dissipation caused by the stressing and relaxation of dipoles within the dielectric material – which act like resistance because energy ends up leaving the component (in the form of heat) and not returning to the circuit. Electrolytic capacitors have an additional source of parasitic resistance, in the form of the electrolytic gel substance used to make electrical contact from the metal-foil "plate" to the surface of the dielectric layer.

Another form of parasitic resistance within a capacitor behaves like a resistor connected in parallel with the ideal capacitance ($R_{leakage}$), resulting from the dielectric not being a perfectly insulating medium. This parasitic characteristic results in a small current passing through the capacitor even when the voltage across the capacitor is steady (i.e. $\frac{dV}{dt}$ is zero).

5.2.3 Parasitic inductance in capacitors

Any time a magnetic field forms around a current-carrying conductor, energy is stored in that magnetic field. We call this magnetic-based energy-storing capability *inductance*, and of course all capacitors must have some inductance due to the wire leads serving as connection points to the capacitor's metal plates. Much of a capacitor's parasitic inductance may be minimized by maintaining short lead lengths as it attaches to a printed-circuit board.

Parasitic inductance is a problem for capacitors in AC applications because inductive reactance (X_L) tends to cancel out capacitive reactance (X_C). If we plot the impedance of a capacitor as a function of frequency, we would expect an ideal capacitor to manifest a straight-line descent on a logarithmic plot. However, what we see is that at a certain frequency the series parasitic inductance resonates with the capacitance leaving only parasitic resistance (ESR), and then past that frequency the inductive effects overshadow the capacitance:



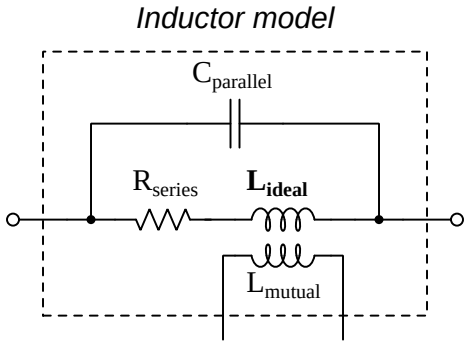
5.2.4 Other parasitic effects in capacitors

Soakage is an interesting effect resulting from dipole relaxation within the dielectric material itself, and may be modeled (as shown) by a series of resistor-capacitor subnetworks. This effect is especially prominent in aluminum electrolytic capacitors, and may be easily demonstrated by discharging a capacitor (by briefly connecting a shorting wire across the capacitor's terminals) and then monitoring the capacitor's DC voltage slowly "recover" with no connection to an external source.

5.3 Inductor parasitics

5.3.1 Model of a real inductor

An ideal inductor exhibits only inductance, with no capacitance, resistance, or other characteristics to interfere. Real inductors exhibit all these phenomena to varying degrees, and we collectively refer to these undesirable traits as *parasitic effects*. The following diagram models some of the parasitic effects observed in real inductors:



In addition to the inductance the inductor is supposed to exhibit (L_{ideal}), the inductor also has parasitic resistance (R_{series}, also known as *Equivalent Series Resistance*, or *ESR*), parasitic capacitance ($C_{parallel}$), and mutual inductance (L_{mutual}) with nearby wires and components.

Some of these parasitic effects – such as equivalent series resistance – affect the inductor's performance in DC applications. Most of the other parasitic effects cause problems in AC and pulsed applications. For example, the effective inductor-capacitor "tank circuit" formed by L_{ideal} and $C_{parallel}$ will cause *resonance* to occur at a particular AC frequency, resulting in much more reactance at that frequency than what would be predicted by the inductive reactance formula $X_L = 2\pi fL$.

Next we will explore common mechanisms for each of these effects.

5.3.2 Parasitic resistance in inductors

Wire resistance plays a dominant role in this parasitic effect due to the typically long lengths of wire necessary to wind the coil that forms most inductors. Wire resistance is not the only dissipative mechanism at work, though. Other losses include magnetic hysteresis of the iron core material as well as *eddy currents* induced in the iron core. An "eddy" current is a circulating electric current induced within the iron core of an inductor, made possible by the fact that iron is an electrically-conductive material as well as being ferromagnetic. These circulating currents do no useful work, and dissipate energy in the form of heating the iron. They may be minimized by forming the iron core from pieces of iron that are electrically insulated from one another, e.g. forming the iron core from *laminated sheets* or *powdered particles* of iron where each sheet or particle is electrically insulated from the next by a layer of non-conductive material on its outer surface.

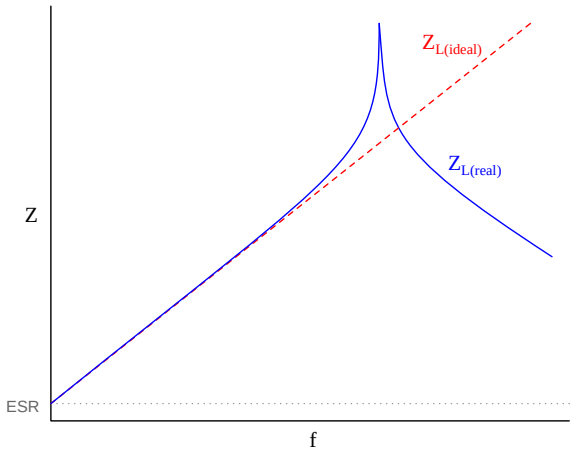
The series resistance of an inductor is always frequency-dependent. In DC conditions (i.e. frequency of zero Hertz) there will be the basic wire resistance of the coil at play. As frequency increases from zero, however, both the magnetic core losses from hysteresis and eddy currents also increase which add to the DC resistance to form a larger ESR. At extremely high frequencies the *skin effect*[1] further adds to the inductor's ESR.

[1] At high frequencies, electric current travels more toward the outer surface of a conductor rather than through the conductor's entire cross-section, effectively decreasing the conductor's cross-sectional area (gauge) as frequency rises.

5.3.3 Parasitic capacitance in inductors

Any time an electric field forms between two conductors, energy is stored in that electric field. We call this electric-based energy-storing capability *capacitance*, and of course all inductors must have some capacitance due to the insulating media between wire leads as well as between adjacent turns of wire within the coil (and between the wire turns and the iron core).

Parasitic capacitance is a problem for inductors in AC applications because capacitive reactance (X_C) tends to cancel out inductive reactance (X_L). If we plot the impedance of an inductor as a function of frequency, we would expect an ideal inductor to manifest a straight-line ascent on a logarithmic plot. However, what we see is that at a certain frequency the parallel parasitic capacitance resonates with the inductance to create a nearly-infinite impedance, and then past that frequency the capacitive effects overshadow the inductance:



Precious little may be done to eliminate parasitic capacitance within any inductor, whereas parasitic inductance is fairly easy to minimize within a capacitor. This explains why when faced with an equivalent choice between a circuit design using capacitors and a circuit design using inductors, capacitors nearly always win. Simply put, it is easier to make a nearly-ideal capacitor than it is to make a nearly-ideal inductor.

This also explains why the self-resonant frequency of most inductors is much lower than the self-resonant frequency of most capacitors: all other factors being equal, an inductor will have more parasitic capacitance in it than an equivalent capacitor will have parasitic inductance within it, making the LC product greater for the inductor than for the capacitor.

5.3.4 Other parasitic effects in inductors

Mutual inductance occurs whenever adjacent conductors' magnetic fields link with one another, which is difficult to avoid especially in physically dense circuit layouts. This parasitic effect may be minimized by proper placement of inductive components (e.g. keeping them spaced as far apart from each other as possible, orienting their axes perpendicular to each other rather than parallel) as well as by core designs with strong magnetic field containment (e.g. toroidal cores contain their magnetic fields better than rectangular cores).

5.3.5 Inductor testing

The following list shows typical tests for inductors at the time of manufacture:

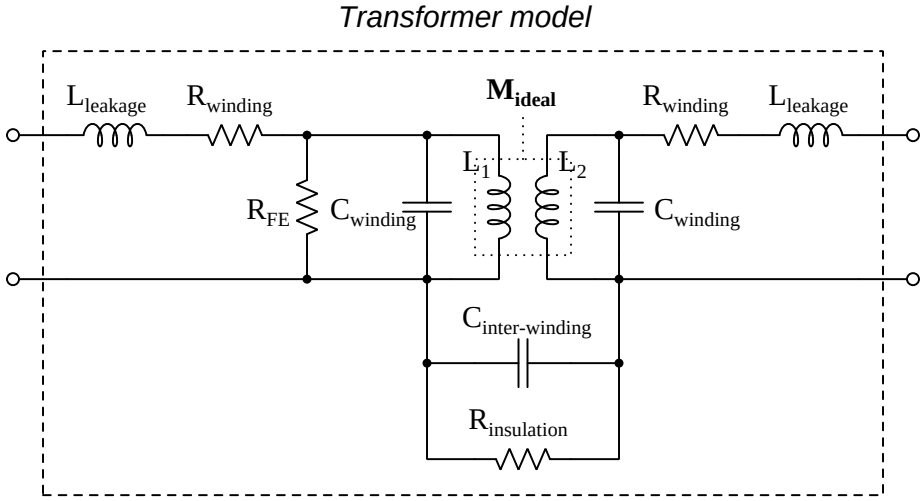
- **DC winding resistance** – the simplest of these to measure, using a multimeter
- **Winding inductance/impedance** – performed by exciting the winding with an AC source and comparing voltage/current, typically using either an LCR meter or an impedance analyzer
- **Equivalent winding resistance** – based on the ratio of AC voltage to AC current (only the portions of each in-phase with each other) representing winding resistance as well as energy losses in the magnetic core
- **Core saturation** – performed by exciting the winding with a sinusoidal AC current and measuring self-induced winding voltage until a non-sinusoidal waveshape is detected, or more crudely by determining when the self-induced voltage no longer increases proportionately to increased excitation current
- **Insulation resistance** – performed with a "Hi-Pot" (high-potential) insulation tester, generally connecting one terminal to either or both ends of the winding and the other tester terminal to the conductive core material, this infers insulation resistance by applying (typically) several thousand Volts DC while measuring the small current passing through the stressed insulation

A more thorough battery of tests used during transformer design development might include the following:

- **BH loop** – performed by exciting the winding with an AC voltage and measuring the self-induced voltage across it, the phasor portion of that induced voltage shifted by 90 degrees from the current waveform representing the rate-of-change of magnetic flux which may be mathematically integrated to infer the amount of magnetic flux, those two variables being plotted using an oscilloscope's X-Y mode much like "curve tracing" for a semiconductor component ; often performed at different temperatures to fully characterize the magnetic properties of the core material
- **High-voltage impulse** – similar to an insulation resistance test, this uses a high-voltage DC pulse applied across the winding insulation to stress it
- **Volt-second** – performed by applying a DC voltage pulse to the winding while measuring current through it, the ideal response being a linear-ramping current whose magnitude is proportional to the Volt-second product of the applied pulse
- **Overload** – subjecting a winding to dramatic overcurrent to thermally and mechanically stress the inductor as it might be in an overloaded power system scenario
- **External magnetic field rejection** – subjecting the entire inductor to an external (alternating) magnetic field and measuring any voltage induced across any winding, usually tested with the magnetic field applied across different physical axes
- **Vibration** – not an electrical test per se, but a useful stressor when followed by repeated electrical tests to ensure electrical/magnetic properties have not substantially changed following the vibration

5.4 Transformer parasitics

An ideal transformer couples electrical energy from its primary winding to its secondary winding(s) with no energy loss (i.e. dissipation) and with no additional impedance standing in the way of that energy transfer. When unloaded, an ideal transformer draws no current whatsoever from its connected source. Real transformers, however, fall quite short of these ideal characteristics. The following diagram models some of the parasitic effects observed in real transformers:



Each of the transformer's windings have resistance owing to the specific resistance of the metal making up those wire windings, this winding resistance effectively being in series with each winding. Additionally, each winding has *leakage inductance* which is the result of magnetic flux that does not couple between the two windings but instead merely stores and releases energy like the self-inductance of a single inductor. Together, winding resistance and leakage inductance contribute to the *impedance* of the transformer, limiting its ability to seamlessly transfer energy from primary to secondary circuits.

We also see each winding suffering from capacitance, resulting from the insulation between adjacent turns of wire. The presence of this parasitic capacitance, something found in every inductor as well as in transformers, means the primary and secondary windings are able to *self-resonate* at certain frequencies. Note also that there is capacitance coupling between the primary and secondary windings, which means no transformer can ever *perfectly* isolate the primary and secondary circuits from each other.

Lastly we have a resistance shown in the model as R_{FE} representing dissipative losses in the iron core material from effects such as hysteresis and magnetostriction. This is often referred to as *iron loss* (as opposed to the *copper loss* of winding resistance).

Next we will explore common mechanisms for each of these effects.

5.4.1 Parasitic resistance in transformers

Winding resistance depends on the size and composition of wire used to make each winding. Using low-resistivity metal and wire of sufficiently large gauge are really the only ways to minimize this parasitic effect. Using a transformer with over-sized wire (i.e. a transformer having current ratings in excess of the currents expected in your application) is one way to reduce this parasitic effect using commonly-available transformers.

So-called "iron loss" resistance is based on magnetic effects, primarily the hysteresis curve for the iron alloy. Modern silicon-steel alloys formulated for transformer applications enjoy very narrow B-H hysteresis curves which in turn translate to low levels of energy dissipation as the core magnetizes in one polarity, de-magnetizes, and re-magnetizes in the other polarity with each cycle of the energizing AC. Using a transformer with an over-sized magnetic core has the result of dispersing the magnetic flux over a larger cross-sectional area, resulting in less magnetic flux density (B) with all other factors being equal. One way to specify a transformer with an over-sized magnetic core is to select one with a higher voltage rating than your application strictly needs, as peak magnetic flux density for any given AC frequency is a function of the peak voltage experienced by each winding. If frequency is not variable, then the cycle time (period) of the AC waveform will likewise be fixed, and since total magnetic flux in any inductor is the integral of voltage and time (i.e. the *Volt-second product*) it means a transformer with a higher voltage rating must possess a larger magnetic core.

5.4.2 Leakage inductance in transformers

Leakage inductance, another contributing factor of transformer impedance, results from some of the core's magnetic flux not coupling perfectly between primary and secondary windings. Core shape and material choice affect this parameter, with toroidal transformer cores generally performing better than cores having abrupt corners.

5.4.3 Parasitic capacitance in transformers

Capacitance unavoidably exists between adjacent turns of wire in any inductor, and also between adjacent turns of primary and secondary windings in a transformer. Reducing turn-to-turn parasitic capacitance within any single winding requires those turns be spaced farther apart from each other, or wound in different patterns, or with an insulating material having a lower dielectric permittivity.

Primary-to-secondary winding capacitance can be mitigated to a large extent by placing an electrostatic *shield* between the two windings of the transformer and connecting that shield conductor to Earth ground. This works precisely the same way as wrapping a foil or braided shield conductor around a multi-conductor cable in order to shield those conductors from electric fields external to the cable. In this case, the shielding stops electric flux lines from linking from one winding to another. Shielded-winding transformers are more expensive to manufacture than their unshielded counterparts.

Another mitigation technique for primary-to-secondary parasitic capacitance is to use a *split-bobbin* transformer where the primary and secondary windings do not overlap each other, but rather are placed at different locations around the magnetic "circuit" formed by the iron core. This increases the separation distance between these two windings without significantly affecting other performance parameters of the transformer, but of course has the effect of increasing the over-all size of the transformer.

5.4.4 Excitation and magnetization currents

As described at the beginning of this section, an ideal transformer draws no current whatsoever from its energizing source. This, however, is possible only by meeting two idealized conditions:

- The primary winding inductance must be infinite
- Magnetic core losses must be zero (i.e. a lossless core)

An infinite primary winding inductance is simply impossible to achieve in any real transformer, though we may approach this ideal by building transformers whose primary inductance values are quite large. Large primary inductance means large values of inductive reactance ($X_L = 2\pi f L$), and so the larger the winding's inductance is the less *magnetization current* it will need to draw from an AC voltage source to create a rate-of-change of magnetic flux in the core sufficient to balance the applied source voltage as per Faraday's Law of Electromagnetic Induction ($V = N\frac{d\phi}{dt}$).

A lossless core material is similarly impossible to achieve in most real transformers, as all ferromagnetic materials invariably dissipate at least *some* energy when magnetized and demagnetized, and air/vacuum-core transformers become impractically large for typical AC power system frequencies (in order to achieve large amounts of inductance, as necessitated above).

Referencing our transformer model showing various parasitic and ideal effects, magnetization current is that amount of current drawn by ideal primary inductance L_1 when the secondary is unloaded (i.e. open-circuited). This, combined with the current drawn from the AC voltage source as a result of the iron losses (R_{FE}), is called *excitation current*. Thus, $I_{excite} = I_{L_1} + I_{R_{FE}}$ which is a phasor current sum, I_{L_1} lagging the applied voltage by 90^o and $I_{R_{FE}}$ being in-phase with the applied voltage. If the excitation current is measured using an instrument such as an oscilloscope capable of comparing phase angles between the measured current and the applied voltage, magnetization current will be the imaginary portion of the excitation current while core-loss current will be the real portion.

5.4.5 Transformer testing

The following list shows typical tests for transformers at the time of manufacture:

- **DC winding resistances** – the simplest of these to measure, using a multimeter
- **Winding inductances/impedances** – performed by exciting one winding with an AC source and comparing voltage/current, typically using either an LCR meter or an impedance analyzer
- **Equivalent winding resistances** – based on the ratio of AC voltage to AC current (only the portions of each in-phase with each other) representing winding resistance as well as energy losses in the magnetic core
- **Turns ratio** – performed by exciting one winding with an AC current source and measuring another winding's short-circuited current using an AC ammeter
- **Phase shift** – performed by exciting one winding with an AC current source and measuring the relative phase shift of another winding's induced voltage
- **Winding polarity/phase-shift** – performed by exciting one winding with an AC voltage and measuring another winding's induced voltage, an oscilloscope used to compare the phases of each voltage
- **Core saturation** – performed by exciting one winding with a sinusoidal AC current and measuring another winding's induced voltage until a non-sinusoidal waveshape is detected, or more crudely by determining when the induced voltage no longer increases proportionately to increased excitation current
- **Excitation current** – performed by exciting one winding with an AC voltage and measuring current drawn by that winding
- **Insulation resistance** – performed with a "Hi-Pot" (high-potential) insulation tester, generally connecting one terminal to both ends of one winding and the other tester terminal to both ends of another winding, this infers insulation resistance by applying (typically) several thousand Volts DC while measuring the small current passing through the stressed insulation

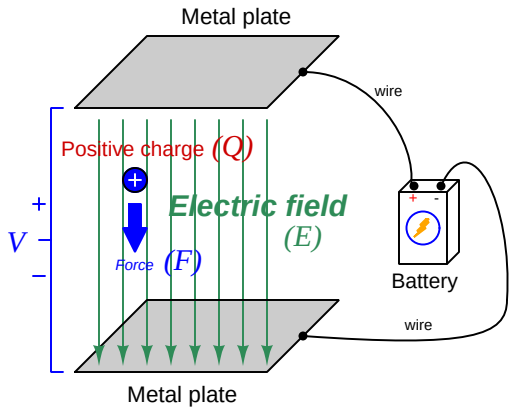
A more thorough battery of tests used during transformer design development might include the following:

- **Linearity** – turns ratio testing performed at several different amounts of excitation current, checking to see how consistent the effective turns ratio is over a wide range
- **BH loop** – performed by exciting one winding with an AC voltage and measuring the voltage induced in another winding, that induced voltage representing the rate-of-change of magnetic flux which may be mathematically integrated to infer the amount of magnetic flux, those two variables being plotted using an oscilloscope's X-Y mode much like "curve tracing" for a semiconductor component ; often performed at different temperatures to fully characterize the magnetic properties of the core material
- **High-voltage impulse** – similar to an insulation resistance test, this uses a high-voltage DC pulse applied across one winding to stress turn-to-turn insulation resistance

- **Volt-second** – performed by applying a DC voltage pulse to one winding while measuring current through it, the ideal response being a linear-ramping current whose magnitude is proportional to the Volt-second product of the applied pulse
- **Overload** – subjecting a winding to dramatic overcurrent to thermally and mechanically stress the transformer as it might be in an overloaded power system scenario
- **External magnetic field rejection** – subjecting the entire transformer to an external (alternating) magnetic field and measuring any voltage induced across any winding, usually tested with the magnetic field applied across different physical axes
- **Vibration** – not an electrical test per se, but a useful stressor when followed by repeated electrical tests to ensure electrical/magnetic properties have not substantially changed following the vibration

5.5 Electric field quantities

A useful definition of electric field (E) is in terms of the force (F) exerted on an electric charge (Q) influenced by that field:



$$\vec{F} = Q\vec{E}$$

Where,

$\vec{F}$ = Force exerted on the charge (Newtons)
Q = Charge quantity (Coulombs[2])
$\vec{E}$ = Electric field (Newtons per Coulomb)

The small "arrow" symbols above the variables for force and electric field in the equation denote those variables as *vector quantities*, having both magnitude and direction. Charge is a *scalar quantity* having only magnitude but no direction, and as a scalar quantity when multiplied by the electric field vector it simply magnifies the magnitude but does not alter the direction. Therefore, the force and electric field vectors always point in the same direction.

Alternatively electric field may be defined in terms of the voltage between the end-points and the distance separating them, in which case we may express the electric field in units of *Volts per meter* as an alternative to *Newtons per Coulomb*:

$$\vec{E} = \frac{V}{\vec{d}}$$

This measurement of electric field strength is very important for quantifying the *breakdown* of electrical insulators: the point at which the electric field becomes so powerful that otherwise immobile charges within the insulating substance are torn free to constitute a current and that substance is no longer an insulator. For rating the dielectric strength of insulating materials, we often see electric fields expressed in units of *kiloVolts per millimeter* rather than Volts per meter just to make the numerical quantities easier to manage (1 kV/mm = 1 *million* V/m).

[2] One Coulomb of electric charge is equal to 6.2415×10^{18} electrons.

The vector arrows shown in the previous illustration representing the electric field between two metal plates actually represent electric *flux* (Φ_E). The electric field (E) is related to electric flux by area (A), the field being a measurement of how densely-packed those flux lines are per unit area:

$$\vec{E} = \frac{\Phi_E}{\vec{A}}$$

Where,

$\vec{E}$ = Electric field, or electric flux density (Newtons per Coulomb)

Φ_E = Electric flux (Newton-meter squared per Coulomb)

$\vec{A}$ = Area over which flux is distributed (square meters)

The mere presence of an unbalanced electric charge at any point in space is sufficient to generate lines of electric flux, the total magnitude of that flux predicted by the following equation:

$$\Phi_E = \frac{Q}{\epsilon}$$

Where,

Φ_E = Electric flux (Newton-meter squared per Coulomb)

Q = Charge quantity (Coulombs)

ϵ = Electric permittivity of the surrounding space (Coulombs squared per Newton-meter squared, approximately 8.85×10^{-12} for empty space)

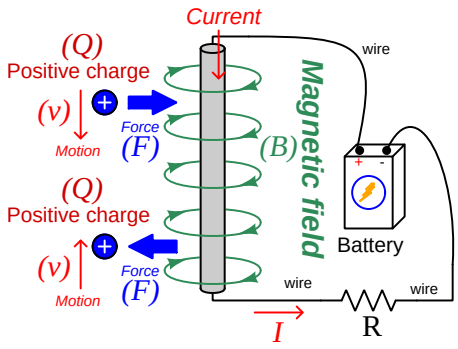
By convention, these flux vectors point *away* from positive charges and point *toward* negative charges, their direction indicating force exerted on any positively-charged particle influenced by that field[3]. As the equation states, the amount of flux depends on how much charge exists at each particle as well as the permittivity of the surrounding space:

For example, identical charges suspended in a vacuum versus in a substance such as oil will have different amounts of flux associated with them as a result of oil and vacuum having different permittivity values. Perfectly empty space has the least amount of permittivity, which means anything else (gas, liquid, or solid matter) has greater ϵ which acts to diminish the amount of electric flux surrounding any charged particle. Superconducting materials have infinite permittivity, which means they forbid the existence of any electric field inside their bulk.

[3] Conversely, the flux vectors point exactly opposite the direction of force applied to any *negatively*-charged particle within that field. This makes sense of the aphorism that *like charges repel and opposite charges attract*. If you consider the two charges shown in this illustration, the positive charge will be pulled in the direction of the flux vectors pointing toward the negative charge, as the negative charge will also be pulled opposite the direction of the flux vectors pointing away from the positive charge (i.e. the negative charge will be pulled *toward* the positive charge) – thus the positive and negative charges feel mutual attraction.

5.6 Magnetic field quantities

A useful definition of magnetic field (B) is in terms of the force (F, called the *Lorentz force*) exerted on a moving electric charge (Q) influenced by that field:



$$\vec{F} = Q\vec{v} \times \vec{B}$$

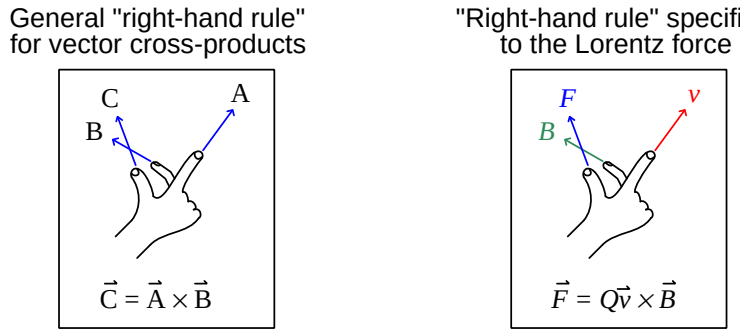
Where,

$\vec{F}$ = Force exerted on the charge (Newtons)
Q = Charge quantity (Coulombs[4])
$\vec{v}$ = Velocity of moving charge (meters per second)
$\vec{B}$ = Magnetic field (Tesla, Webers per square meter, or Newtons per Ampere-meter)

The small "arrow" symbols above the variables for force and velocity and magnetic field in the equation denote those variables as *vector quantities*, having both magnitude and direction. Charge is a *scalar quantity* having only magnitude but no direction, and as a scalar quantity when multiplied by the velocity vector it simply magnifies the magnitude but does not alter the direction. The "cross-product" ($\times$) is a specific form of vector multiplication, and it results in a product at right angles to the vector directions of both terms. Therefore, the force and velocity and electric field vectors never all point in the same direction.

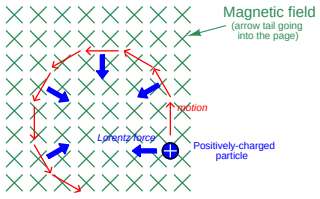
[4] One Coulomb of electric charge is equal to 6.2415×10^{18} electrons.

Vector cross-products conveniently relate to the fingers of the right hand, which is where the "right-hand rule" originates:



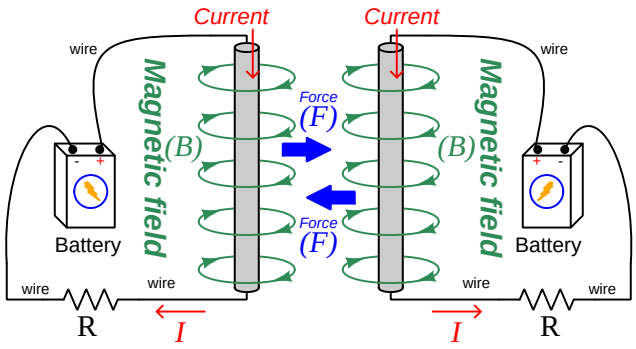
When holding the index finger, middle finger, and thumb of your right hand perpendicular to each other, your index finger points in the direction of the velocity vector ($\vec{v}$), your middle finger in the direction of the magnetic field vector ($\vec{B}$), and your thumb in the direction of the force vector ($\vec{F}$). A simple mnemonic I use to remember these relationships of fingers to vectors is that the **I**ndex finger points in the direction of current[5] (I), the **M**iddle finger points in the direction of the **m**agnetic field (B), and the **Th**umb points in the direction of the **th**rust (i.e. force) acting upon the moving charge.

The Lorentz force's effect on electrically-charged particles in motion has many applications, from redirecting the paths of charged-particle beams in particle accelerator machines, to bending the trajectory of electron beams in a cathode-ray tube (CRT), to forcing electrons to travel in spiral or circular trajectories inside of magnetron (microwave oscillator) tubes. An illustration of a positively-charged particle curving perpendicular to a magnetic field appears here:



[5] Some textbooks speak of a "left-hand rule" which is intended to make sense of electric charge motion (current) in terms of *electron flow*. As we know, electrons are the only real mobile charge carriers within metal conductors, and so technically "electron flow" notation is most physically accurate when describing the motion of electric charges in metallic circuits. However, the right-hand rule is a *mathematical definition* for vector cross products, the concept of the cross product arising in the late 18th century when electrical science was still in its infancy. Early explorers of electricity used the established mathematical tools of their time and applied it to their work with electric currents and magnetism. At that time, charge carriers in metal wires were assumed to be "positive" and this is how the motion of positively-charged carriers became associated with the first vector of the cross-product. As a result of this assumption which was later proven false, we have two different conventions for denoting the motion of electricity: electron-flow which is *physically* accurate (for metal wires, at least), and conventional flow which is *mathematically* accurate. This, perhaps more than any other reason, is why educational programs designed for mathematically rigorous fields (e.g. electrical engineering) exclusively use conventional flow notation rather than electron flow notation to denote the direction of current.

If the moving charge in question is not a single charged particle but rather part an electric *current* passing through a conductor parallel to the first, both conductors will experience a mutually-attracting force given by the following equation:



$$\vec{F} = I\vec{l} \times \vec{B}$$

Where,

$\vec{F}$ = Force exerted on both conductors (Newtons)

I = Current (Amperes)

$\vec{l}$ = Length of wire (meters)

$\vec{B}$ = Magnetic field (Tesla, or Webers per square meter, or Newtons per Ampere-meter)

The point-charge Lorentz force equation and the two-conductor Lorentz force equation are not that different from one another. Dimensional analysis validates this: the Lorentz force on a moving charge uses that charge quantity (Coulombs) multiplied by the point-charge's velocity in meters per second to give Coulomb-meters per second for the first term:

$$Q\vec{v} = [\mathrm{C}]\left[\frac{\mathrm{m}}{\mathrm{s}}\right] = \left[\frac{\mathrm{C} \cdot \mathrm{m}}{\mathrm{s}}\right]$$

The Lorentz force on a current-carrying conductor uses the current (Amperes, which is Coulombs per second) multiplied by length in meters, for the same composite units of Coulomb-meters per second:

$$I\vec{l} = \left[\frac{\mathrm{C}}{\mathrm{s}}\right][\mathrm{m}] = \left[\frac{\mathrm{C} \cdot \mathrm{m}}{\mathrm{s}}\right]$$

This dimensional equivalence makes conceptual sense as well: an electrically-charged particle moving through empty space is an electric current in its own right, and an electric current flowing through a conductor is just a collection of charged particles moving through space (just not *empty* space). In either case, the basis for the Lorentz force remains the same: the moving charge(s) create their own magnetic field, which reacts with the magnetic field of the original current-carrying wire to produce forces acting on both.

If the two currents flow in the same direction, their mutual forces *attract*. If the two currents flow in opposite directions, their mutual forces *repel*. This is the basis of electric motors: causing

mechanical motion by electro-magnetic attraction and repulsion. It also represents an interesting contrast with electric fields:

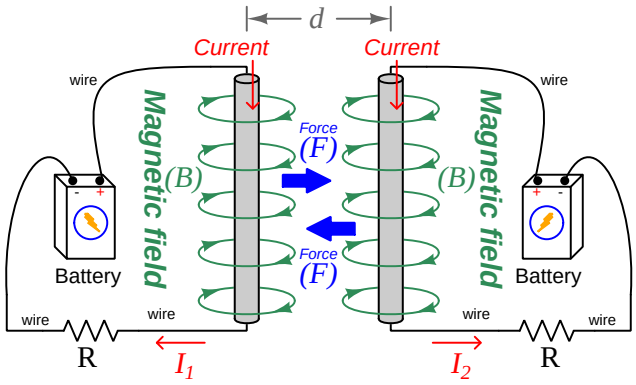
With electric fields, opposite *charges* attract and like *charges* repel.

With magnetic fields, opposite *poles* attract and like *poles* repel.

With parallel currents, opposite *directions* repel and like *directions* attract[6].

Two parallel current-carrying conductors of length l and separated by a distance d will generate a mutual force proportional to both their currents:

$$F = l\frac{\mu I_1 I_2}{2\pi d}$$



[6]That is, assuming it's *like* charges moving in these directions! If the charges in question are opposite each other – for example electrons in one circuit and holes in another – then like directions will repel and opposite directions will attract!

The circular loops surrounding the current-carrying conductors in the previous illustrations represent the magnetic lines of *flux* (Φ_B) surrounding each of those conductors. The magnetic field (B) is related to magnetic flux by area (A), the field being a measurement of how densely-packed those flux lines are per unit area. For this reason, magnetic field (B) is more properly known as *magnetic flux density*:

$$\vec{B} = \frac{\Phi_B}{\vec{A}}$$

Where,

$\vec{B}$ = Magnetic field or flux density (Tesla, Webers per square meter, or Newtons per Ampere-meter)

Φ_B = Magnetic flux (Webers)

$\vec{A}$ = Area over which flux is distributed (square meters)

An older unit of measurement for magnetic flux density B is the *Gauss* which is much smaller than a Tesla, with one Tesla equivalent to 10,000 Gauss. To put things into perspective, the Earth's natural magnetic field has a strength of approximately one-half of one Gauss[7].

Magnetic field strength is an inverse function of distance from any current-carrying wire, and also depends on the magnetic permeability of the space adjacent to the wire:

$$B = \frac{\mu I}{2\pi d}$$

Where,

B = Magnetic field or flux density (Tesla, Webers per square meter, or Newtons per Ampere-meter)

μ = Magnetic permeability of the surrounding space (Tesla-meters per Ampere, $4\pi \times 10^{-7}$ for empty space)

I = Current (Amperes)

d = Distance from conductor (meters)

[7] Using the online *Magnetic Field Calculator* application provided by NOAA (the National Oceanic and Atmospheric Administration) at `https://ngdc.noaa.gov/geomag/calculators/magcalc.shtml#igrfwmm`, applying the World Magnetic Model WMM modeling algorithm for years 2019-2024, the total magnetic field strength at my home is 53,584.4 nano-Tesla (53,584.4 nT or 0.535844 Gauss), and presently (May 2020) decaying at a rate of -104.1 nT per year.

The relation of magnetic *flux* to current through a conductor follows a similar equation:

$$\Phi = \frac{\mu A I}{2\pi d}$$

Where,

Φ = Magnetic flux (Webers)

μ = Magnetic permeability of the surrounding space (Tesla-meters per Ampere, $4\pi \times 10^{-7}$ for empty space)

A = Area over which flux is distributed (square meters)

I = Current (Amperes)

d = Distance from conductor (meters)

As this equation makes clear, the amount of magnetic flux surrounding a current-carrying conductor depends not only on the amount of current, but also on the sampled area, the distance from the wire, and also the surrounding material. Most[8] substances (gas, liquid, solid) have permeability values greater than that of empty space, and so this means magnetic flux is usually *enhanced* by the presence of matter around the current-carrying conductor.

The total magnetic flux enclosed by a circular wire loop follows a similar equation:

$$\Phi = \frac{\pi \mu I r}{2}$$

Where,

Φ = Magnetic flux (Webers)

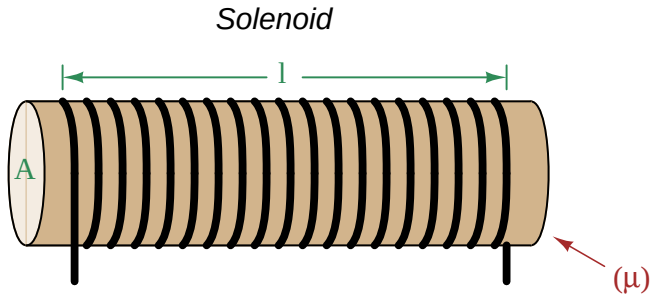
μ = Magnetic permeability of the surrounding space (Tesla-meters per Ampere, $4\pi \times 10^{-7}$ for empty space)

I = Current (Amperes)

r = Radius of circular loop (meters)

[8] Interestingly, superconducting materials *forbid* magnetic fields inside of their bulk, and so the permeability value of any superconductor must be zero!

A common form of electromagnet known as a *solenoid* takes the form of a wire coil wrapped in such a way as to form a long[9] cylinder, often wrapped around a plastic frame, and often with a ferromagnetic material such as iron in the center:



The amount of magnetic flux, and the flux density, within the interior of a current-carrying solenoid are given by the following formulae:

$$\Phi = \frac{\mu N A I}{l} \qquad\qquad B = \frac{\mu N I}{l}$$

Where,

Φ = Magnetic flux (Webers)

B = Magnetic field or flux density (Tesla, Webers per square meter, or Newtons per Ampere-meter)

μ = Magnetic permeability of the surrounding space (Tesla-meters per Ampere, $4\pi \times 10^{-7}$ for empty space)

N = Number of turns of wire in the coil

A = Cross-sectional area of solenoid coil (square meters)

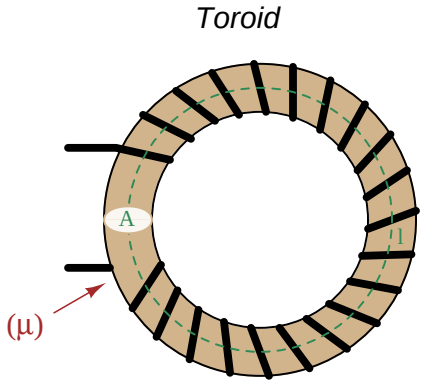
I = Current (Amperes)

l = Length of solenoid coil (meters)

These formulae have interesting implications for solenoid design. Note how a shorter (i.e. smaller length l) solenoid identical in all other respects will generate a stronger magnetic field for a given current. Note also how the flux density (B) remains constant with increasing cross-sectional area (A) if all other factors are equal, and that this necessarily means a greater amount of total magnetic flux (Φ) for a greater area A.

[9] These magnetic field formulae apply perfectly to a solenoid coil that is closely-packed (i.e. each turn adjacent to the next) and infinitely long. Therefore, they only approximate real solenoid behavior. This fact may be understood by performing a thought experiment where we decrease the solenoid coil's length to zero, in which case the formulae predict an *infinite* amount of magnetism for any amount of current at all, which of course cannot be true.

Another common form of electromagnet known as a *toroid* is really just a solenoid bent in a circle so that its two ends meet[10] cylinder, often wrapped around a plastic frame, and often with a ferromagnetic material such as iron in the center. Toroids have the unusual property of *containing* their magnetic flux lines extremely well, unlike solenoids, wires, and simple coils which all radiate magnetic fields. They find application as energy-storage devices, or as electromagnets suitable for applying magnetic fields to specimens placed *inside* the toroid's cross-section:



The amount of magnetic flux, and the flux density, within the interior of a current-carrying toroid are identical to that within an otherwise identical otherwise identical solenoid having a length (l) equal to the toroid's circumference:

$$\Phi = \frac{\mu N A I}{l} \qquad\qquad B = \frac{\mu N I}{l}$$

Where,

Φ = Magnetic flux (Webers)

B = Magnetic field or flux density (Tesla, Webers per square meter, or Newtons per Ampere-meter)

μ = Magnetic permeability of the surrounding space (Tesla-meters per Ampere, $4\pi \times 10^{-7}$ for empty space)

N = Number of turns of wire in the coil

A = Cross-sectional area of toroid (square meters)

I = Current (Amperes)

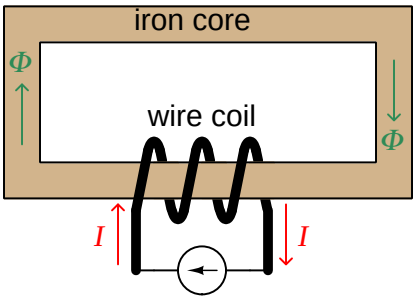
l = Circumference of toroid (meters)

If we wish to substitute toroid radius (r) for circumferential length (l), the formulae become the following:

$$\Phi = \frac{\mu N A I}{2\pi r} \qquad\qquad B = \frac{\mu N I}{2\pi r}$$

[10] Again, the magnetic field formulae are only accurate for a toroidal coil that is closely-packed (i.e. each turn adjacent to the next) and infinitely long, and therefore only approximate real toroid behavior. This fact may be understood by performing an equivalent thought experiment as before where we decrease the toroid's circumference to zero and absurdly end up with *infinite* magnetism for a finite current.

Many applications of electromagnetism involve conductive *coils* wrapped around some form of ferromagnetic core material, the purpose of that core being to provide a higher-permeability pathway for the magnetic flux than would exist otherwise through air, and the purpose of the wire coil being to intensify the amount of magnetism developed by the electric current beyond what would be possible with a straight current-carrying wire. These magnetic cores typically form a closed loop, or *magnetic circuit* for the lines of magnetic flux to naturally form a closed path. A simple example appears here:



The amount of magnetic flux (Φ) present in the magnetic "circuit" formed by the iron core depends on many factors. First and foremost is the amount of electric current (in Amperes) passing through the wire coil and the number of turns that coil makes around the iron core. The product of this current and the number of turns is called the *magnetomotive force* or *mmf* of the magnetic circuit, analogous to "electromotive force" or "emf" often used as a synonym for voltage in an electric circuit. Not surprisingly, the standard metric unit of measurement for magnetomotive force is the *Ampere-turn*.

However, magnetomotive force alone does not fully describe the current's effect on magnetism within the iron core. The total length of the magnetic circuit is also an important factor, since a longer path distributes that magnetomotive force over a greater distance. The quotient of magnetomotive force and magnetic circuit length is called *magnetic field intensity*, symbolized by the variable H and expressed in units of *Ampere-turns per meter*.

Magnetic permeability (μ) relates magnetic field intensity (H) to the magnetic flux density (B) within the core material, such that a greater permeability will result in higher flux density for any given amount of field intensity. Permeability is a property of the core material and not its geometry, mathematically defined as the ratio of flux density to field intensity: $\mu = \frac{B}{H}$

Magnetic reluctance ($\Re$) relates magnetomotive force (mmf) to magnetic flux (Φ), and is related not only to the core material's permeability but also its geometry. It is mathematically defined as the ratio of magnetomotive force to magnetic flux: $\Re = \frac{\text{mmf}}{\Phi}$

If all this seems confusing, you are in good company. Not only are there many magnetic variables, some related to physical geometry and others not, but there are *two different sets of metric units* appropriate for expressing each! The older units were based on the centimeter-gram-second (CGS) version of the metric system, while the newer units are based on the meter-kilogram-second or SI (Système International) version of the metric system.

Quantity	**Symbol**	**SI unit**	**CGS unit**
Magnetomotive force	mmf	Ampere-turn (A-t)	Gilbert (Gb)
Flux	Φ	Weber (Wb)	Maxwell (Mx)
Field intensity	H	Ampere-turns per meter (A-t/m)	Oersted (Oe)
Flux density	B	Tesla (T)	Gauss (G)
Permeability	μ	Tesla-meters per Ampere-turn (T-m/A-t)	Gauss per Oersted (G/Oe)
Reluctance	$\Re$	Ampere-turns per Weber (A-t/Wb)	Gilberts per Maxwell (G/Mx)

Magnetomotive force (mmf) and magnetic flux (Φ) may be thought of as the "raw" measures of magnetism, with Ampere-turns and Webers being their respective SI metric units. Reluctance ($\Re$) is the ratio of the two for any given magnetic circuit with known dimensions and core material. Simply put, reluctance tells you how many Ampere-turns of magnetomotive force will be necessary to create one Weber of magnetic flux in a given space.

Magnetic field intensity (H) and magnetic flux density (B) may be thought of as the "normalized" measures of magnetism, with Ampere-turns per meter and Tesla being their respective SI metric units. H and B relate to mmf and flux by the physical dimensions of the magnetic circuit (length and cross-sectional area, respectively). Permeability is the ratio of the two for any given magnetic core material. Simply put, permeability tells you how many Tesla of magnetic field (i.e. flux density, or Webers of flux per square meter or cross-sectional core area) you will obtain for one Ampere-turn per meter of magnetic field intensity applied to a given core material.

Conversion between the newer SI and the older CGS metric units are as follows:

Quantity	**Conversion equivalence**
Magnetomotive force (mmf)	1 Ampere-turn = $\frac{4\pi}{10}$ Gilberts
Magnetic flux (Φ)	1 Weber = 10^8 Maxwells
Magnetic field intensity (H)	1 Ampere-turn/meter = $\frac{4\pi}{1000}$ Oersteds
Magnetic flux density (B)	1 Tesla = 10^4 Gauss
Permeability (μ)	1 Tesla-meter/Ampere-turn = $\frac{10^7}{4\pi}$ Gauss/Oersteds
Reluctance ($\Re$)	1 Ampere-turn/Weber = $\frac{4\pi}{10^9}$ Gilberts/Maxwell

5.7 The OSI Reference Model

Digital communication is a fairly complex topic, and there are a great many ways in which one digital device might communicate information to another. In order to help describe how different methods (protocols) function in forming a complete communication system, the *OSI Reference Model* was developed. This is not a standard itself, but rather a taxonomical system for *describing* and *classifying* various aspects of any communications system.

For example, one communication system might use discrete voltage pulses in a non-return-to-zero (NRZ) format to represent 0 and 1 bit states, while another might use analog audio tone frequency voltage signals to do the same. According to the OSI model, the choice over electrical signal types is a *Layer 1* matter. The patterns in which bit states make up larger serial data frames is another matter which must be standardized in any communication system, and according to the OSI model that is the domain of *Layer 2*. At the highest level of the OSI model we have *Layer 7* which regards the practical meaning of the bits communicated, whether those be alphanumeric codes representing an internet web page's contents, Modbus control functions, or anything else.

Layer 7 **Application**	This is where digital data takes on practical meaning in the context of some human or overall system function. *Examples: HTTP, FTP, HART, Modbus*
Layer 6 **Presentation**	This is where data gets converted between different formats. *Examples: ASCII, EBCDIC, MPEG, JPG, MP3*
Layer 5 **Session**	This is where "conversations" between digital devices are opened, closed, and otherwise managed for reliable data flow. *Examples: Sockets, NetBIOS*
Layer 4 **Transport**	This is where complete data transfer is handled, ensuring all data gets put together and error-checked before use. *Examples: TCP, UDP*
Layer 3 **Network**	This is where the system determines network-wide addresses, ensuring a means for data to get from one node to another. *Examples: IP, ARP*
Layer 2 **Data link**	This is where basic data transfer methods and sequences (frames) are defined within the smallest segment(s) of a network. *Examples: CSMA/CD, Token passing, Master/Slave*
Layer 1 **Physical**	This is where data bits are equated to electrical, optical, or other signals. Other physical details such as cable and connector types are also specified here. *Examples: EIA/TIA-232, 422, 485, Bell 202*

Students of philosophy might see a parallel between the OSI model's description of digital communication versus the ancient Greek philosopher Aristotle's classification of *causes*. Aristotle proposed four different ways to describe what causes a thing to exist or to occur. I have arranged these four causes in an order analogous to the OSI model's layers:

Aristotle's Four Causes

Layer 4 **Final**	This is the end-purpose for which something exists or has happened.
Layer 3 **Efficient**	This is the means by which something came into existence or occurred.
Layer 2 **Formal**	This is the structure or form taken by something.
Layer 1 **Material**	This is the fundamental substance something is made of.

For example, we might classify the four causes of a textbook as follows:

- **Final cause** – to educate and enlighten the reader
- **Efficient cause** – a human author wrote the book, and it came into existence through the operation of a printing press
- **Formal cause** – a series of characters and images arranged into words, paragraphs, diagrams, graphs, equations, etc. following one or more standards of language (e.g. English, international math notation, Cartesian-coordinate graphs)
- **Material cause** – paper and ink

All of these causes describe different aspects of what was required for a book to come into existence. Similarly, with the OSI model, each level of the model is like one of Aristotle's causes in that it describes something necessary for the communication to take place.

5.8 Statistics

Statistics is the study of information collected from a multitude of real-world measurements or incidents. We refer to any such collection of information as a *data set*. For example, the collected resistance values from a particular batch of resistors would constitute a data set. The totality of information about a particular measure is called the *population*, while any subset of that population is called a *sample*. To illustrate, if we were to record the resistance value of each and every resistor in a manufactured lot, that data would represent the *population* of resistances produced. If we were to measure only some of the resistors from that lot, the collected data would constitute a *sample*.

Choosing which elements of a population to sample in order to learn important things about the population is not a trivial matter, since improper sample selection may influence the results and skew one's perception of the population. The greater the sample size, of course, the more accurate any conclusions will be about the population as a whole. One common method is to randomly sample elements from a population in the hope that the random sampling will tend to cancel out any variations that might otherwise make the sample deviate substantively from the population.

Perhaps the most widely-recognized measurements of any population or sample from a population are those known as *measures of central tendency*: the *average* (otherwise known as the *mean*), the *median*, and the *mode*. Calculating the mean consists of summing the values of the elements and then dividing that sum by the number of elements (N) contained within the data set. The *median* which is the middle value of a data set when all its elements are arranged in numerical order, or the average of the two middle values when the set contains an even number of elements. The *mode* is simply the most commonly-repeated value within a data set.

Another class of statistical measurements is known as *measures of variation*, used to express the degree to which values within a data set deviate from one another. The first of these measures is *deviation*, calculated simply as the difference between a given value from the set versus the mean (average) of the entire set. As such, deviation applies to individual elements of a data set and not to the data set as a whole. In order to express the degree of deviation for the whole set, we typically *square* all of the individual deviation values and then find the average of those squared deviations, the result called the *variance* of the set. If our data set represents the entire population of data we're interested in analyzing, the variance is averaged by dividing the sum of squares by N as we would with any typical mean, and we call it the *population variance*. However, if our data is merely a subset sampled from a larger population, the variance is averaged by dividing the sum of squares by $N - 1$, and we call the result the *sample variance*.

The square root of variance is known as *standard deviation* (σ): the square root of population variance being population standard deviation, and the square root of sample variance being sample standard deviation. Standard deviation has a decided advantage over variance in that the units of measurement for standard deviation are the same as for the raw data elements. For example, if we were to measure the electrical resistance of a multitude of resistors to populate a data set, all measures of central tendency (e.g. mean, median, mode) would be in units of Ohms, as would deviation for each resistance. However, variation would have to be expressed in units of Ohms *squared*, while standard deviation would be expressed in plain Ohms.

Some computer programming languages such as Python offer libraries of statistical functions. Below we see the source code listing for a Python script analyzing a set of one hundred 1 kΩ (nominal) resistor measurements taken with a multimeter reading in whole-numbered values of Ohms, this particular program making use of Python's `matplotlib` and `statistics` libraries[11]:

```
import matplotlib.pyplot as plt
import statistics as stat

data = [1012, 1008, 1006, 1017, 1006, 1003, 1005, 1011, 1004, 1006, \
        1006, 1017,  989, 1001,  997, 1008, 1019, 1018, 1006, 1006, \
        1004,  986, 1006, 1009, 1006, 1008, 1003, 1005, 1018, 1015, \
        1004, 1003, 1012, 1003,  988, 1007,  993, 1010, 1013, 1003, \
        1031, 1035, 1003, 1007, 1014,  999, 1015, 1006, 1006,  997, \
        1006, 1001, 1005, 1012,  996, 1012, 1015,  998, 1017, 1009, \
        1007, 1003, 1011, 1004, 1000, 1011, 1006, 1002, 1005, 1009, \
        1006, 1005, 1009, 1008, 1011,  997, 1007, 1002, 1010, 1018, \
        1004, 1002, 1002, 1016, 1000, 1006, 1007, 1001, 1006, 1001, \
        1006, 1013, 1011, 1011, 1011, 1010,  997, 1003, 1004, 1003]

print(" Sum =", sum(data))
print(" Mean =", stat.mean(data))
print(" Median =", stat.median(data))
print(" Mode =", stat.mode(data))
print(" Population variance =", stat.pvariance(data))
print(" Population standard deviation =", stat.pstdev(data))
print(" Sample variance =", stat.variance(data))
print(" Sample standard deviation =", stat.stdev(data))
plt.hist(data, 20, color='orange', edgecolor='black')
plt.show()
```

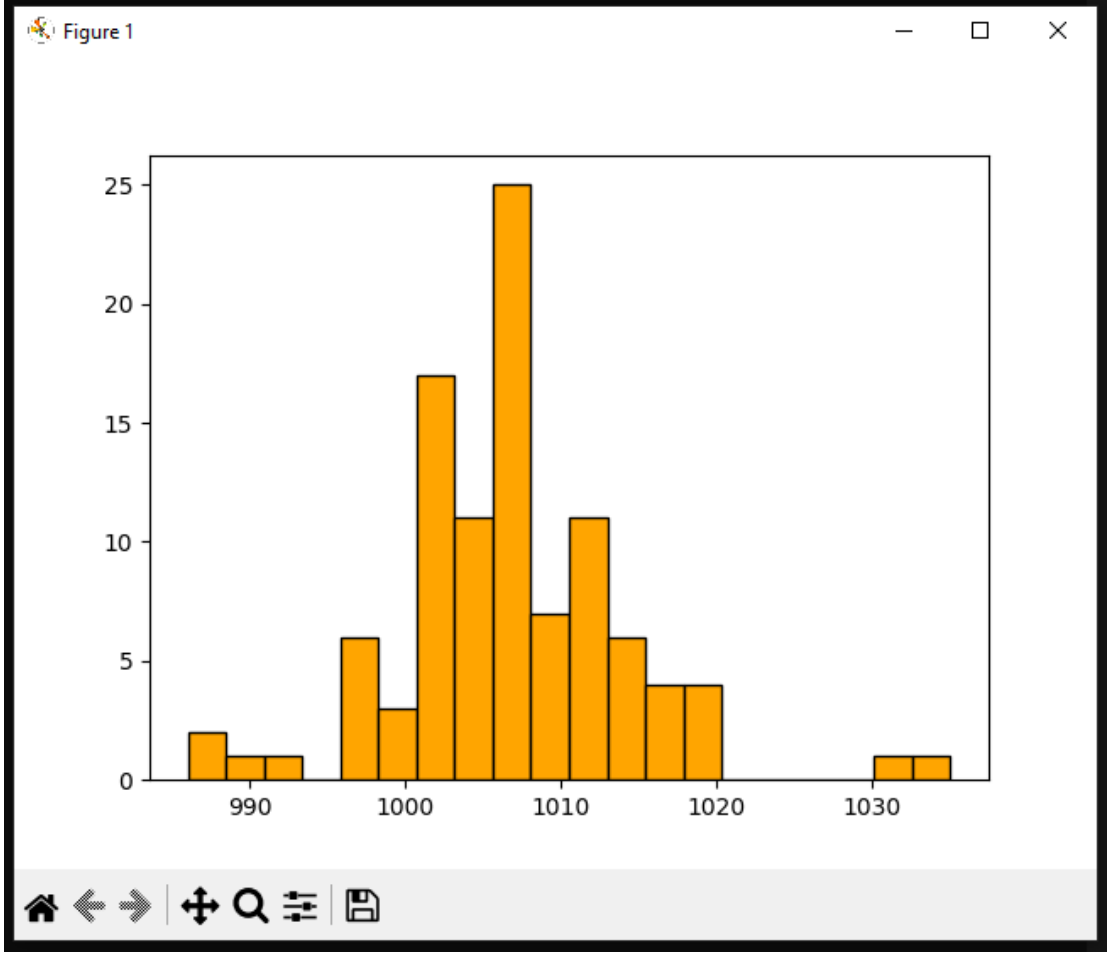
The text output by this program when run is as follows:

```
Sum = 100670
Mean = 1006.7
Median = 1006.0
Mode = 1006
Population variance = 53.75
Population standard deviation = 7.33143914930759
Sample variance = 54.292929292929294
Sample standard deviation = 7.368373585325957
```

[11] At the time of this writing, these libraries are not installed with the standard Python software package but must be installed separately using the *pip* utility. The command-line instuction to install the `matplotlib` library is `py -m pip install matplotlib`, and the instruction to install the `statistics` library works simiarly.

The `matplotlib` Python library invoked in this program is able to plot a *histogram* showing the frequency[12] of each value in the set by the height of vertical bars. For this particular analysis we have specified the data set be divided into twenty "bins" constituting the number of horizontal places in the histogram for bars:

```
Sum = 100670
Mean = 1006.7
Median = 1006.0
Mode = 1006
Population variance = 53.75
Population standard deviation = 7.33143914930759
Sample variance = 54.292929292929294
Sample standard deviation = 7.368373585325957
```



[12] In the context of statistics, "frequency" means how often the same value appears within a data set.

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.

Energy

Conservation of Energy

Mass

Electrons

Voltage

Resistance

Current

Ohm's Law

Frequency

Electrical source

Electrical load

Kirchhoff's Voltage Law

Electric field

Magnetic field

Hysteresis

Faraday's Law of electromagnetic induction

Annotating diagrams as a problem-solving strategy

Graphing as a problem-solving strategy

Parts per million (ppm)

Parts per billion (ppb)

Josephson effect

Hall effect

Quantum Hall Effect

Seebeck effect

Zener diode

Bridge network

Kelvin four-wire method

Null-balance method

Capacitance

Inductance

Mutual induction

Permeability

Sinusoidal decomposition (i.e. Fourier's Theorem)

Fundamental frequency

Harmonic frequency

Noise

???

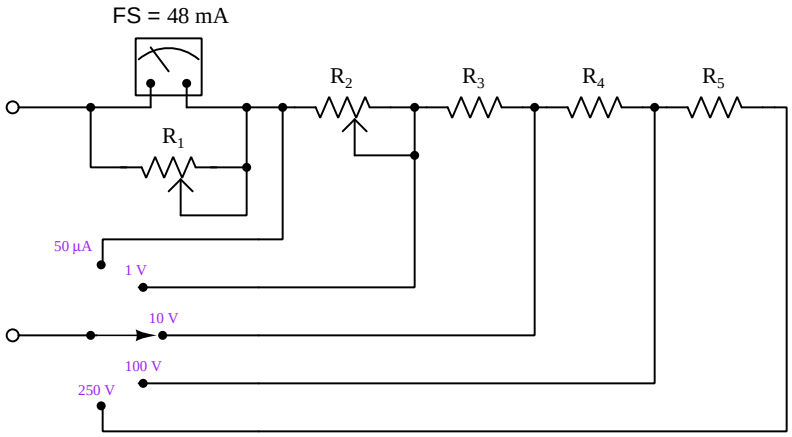
???

???

???

6.1.3 Analog multimeter calibration

Suppose you must calibrate this analog multimeter, the mechanism of which contains two variable resistances as well as a mechanical zero-adjust knob on the D'Arsonval meter movement itself:



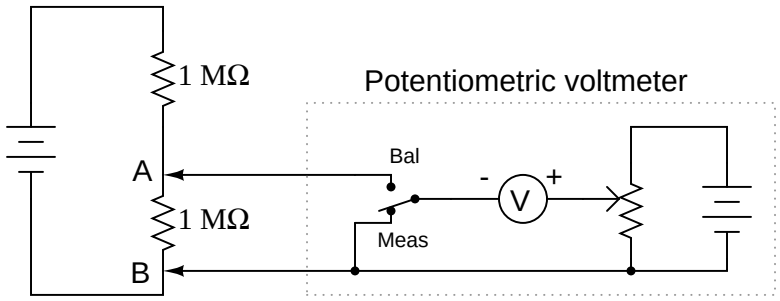
Devise a calibration procedure that makes the most sense for this meter, specifying the order in which any adjustments should be made.

Challenges

- Suppose you determined one of the fixed resistors had too little resistance. What could be done to remedy this?

6.1.4 Potentiometric voltmeter

The following voltmeter circuit avoids problems of "loading" when measuring high-resistance voltage sources. Describe how to operate this circuit, and how loading error is eliminated by using such a *null-balance* instrument:



Challenges

- Describe how this apparatus could be mis-used such that it actually did load down the circuit under test.

6.1.5 Test uncertainty ratio

Calibration laboratories often make reference to *Test Uncertainty Ratios* (TURs), also known as Test Specification Ratios (TSRs), normally insisting on a ratio of 4:1 or better when performing calibration work on instruments. What does this figure mean?

Challenges

- What ill effect(s) might result if you were to use a calibration apparatus with a TUR of less then 4:1?

6.1.6 Clock accuracy

Explain how the following metrological terms might relate to the function and/or calibration of a household clock, like the type that is battery-powered and hangs on a wall:

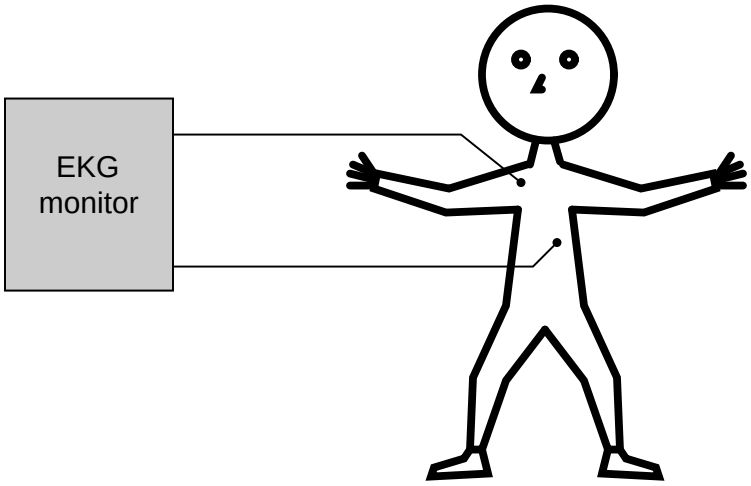
- **Error**
- **Resolution**
- **Primary standard**

Challenges

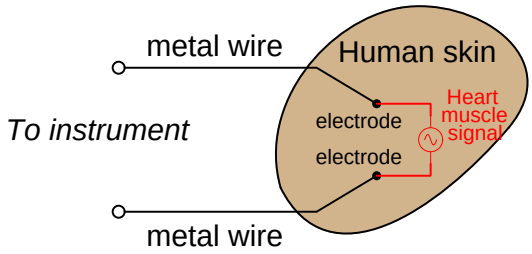
- Are digital clocks more accurate, less accurate, or the same accuracy as analog clocks?

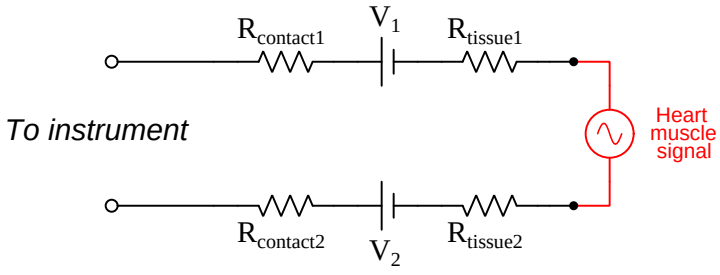
6.1.7 EKG monitor

An *electrocardiogram* or *EKG* is a time-domain measurement taken of electrical impulses from the heart as it pumps blood. Metal electrodes placed on the patient's skin intercept these tiny voltage signals from the contracting heart muscles:



A junction of metal wire to human skin is surprisingly complex from an electrical perspective, and may be approximately modeled by the following collection of idealized components:





Resistors $R_{contact1}$ and $R_{contact2}$ represent electrical resistance between the metal electrodes and skin at the point of contact. Resistors $R_{tissue1}$ and $R_{tissue2}$ represent electrical resistance of human

tissue between the points of electrode contact and the actual heart muscle. The two series-opposing potentials (V_1 and V_2) are not intentional, but rather the result of electrochemical action between the metal electrode surfaces and human skin and therefore cannot be eliminated.

Explain how these parasitic effects may (or may not!) compromise the integrity of the voltage signals necessary for a valid EKG test.

How are the challenges of accurately measuring heart muscle voltage signals similar to challenges faced when making ultra-precise voltage measurements in metrology applications?

Challenges

- Identify a scenario where the skin-contact voltages *could* become problematic.
- Note that the heart muscle's voltage is modeled here as an AC source. Does KVL apply to AC voltages as well as to DC voltages?

6.1.8 Testing intrinsic standards

Primary standards reside in metrological laboratories, and are used to calibrate lesser-tier standards which in turn are used to calibrate other instruments used for practical work. Many of these are *intrinsic* standards, their stability based on fundamental constants of the universe. How do you suppose such intrinsic, primary standards are compared against one another especially as they are rarely located next to one another? If their values were to differ, how do you suppose this is possible since they're supposedly based on immutable properties of physics?

Challenges

- Identify some of the intrinsic standards used in electrical metrology.

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	

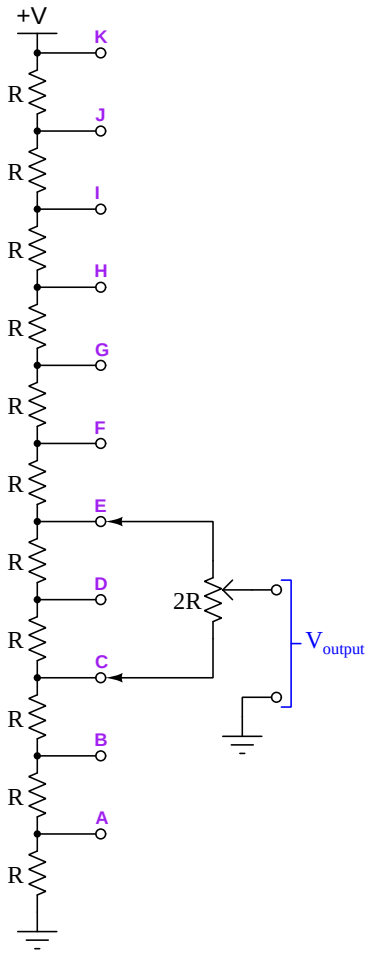
Note how the square-root term (y) is calculated in cell C1, and the denominator term (z) in cell C2. This makes the two final formulae (in cells B1 and B2) simpler to interpret. The positioning of all these cells on the grid is completely arbitrary[10] – all that matters is that they properly reference 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 Kelvin-Varley voltage divider

The following network is called a *Kelvin-Varley voltage divider*, this one consisting of eleven equally-sized precision resistors along with a potentiometer having a total resistance twice that of any of those fixed resistors, connected across any *two* of the fixed resistors:



Such a network is designed to achieve fine voltage adjustment from the potentiometer by constraining the potentiometer's adjustment range to just one-tenth of the total applied voltage. This is useful in electrical *metrology*, the science and practice of precise electrical measurements.

Calculate the amount of voltage between the "Output" and ground terminals of the divider given an R size of 10 kΩ (each), a potentiometer resistance ($2R$) of 20 kΩ, a wiper position that is exactly centered (50%), the potentiometer connected as shown (to terminals E and C), and a supply voltage of 50 Volts.

Determine where to connect the potentiometer, and where to set its wiper, to output 6.3 Volts

given a power supply voltage of 10 Volts.

Challenges

- Why not just use a plain and simple potentiometer, without the eleven-resistor network? What purpose does a voltage divider this complex serve?

6.2.4 Hamon voltage divider simulation program

Write a text-based computer program (e.g. C, C++, Python) to calculate the true voltage-division ratio of a Hamon voltage divider having three resistors of any values you choose that are relatively close to each other. Your program should calculate the error of each resistor compared to its nominal value, as well as the error of the final voltage-division ratio compared to the ideal voltage-division ratio.

Challenges

- How precise of a voltage-division ratio is possible using resistors varying from each other by $\pm$ 1%?
- Modify this program to simulate a Hamon divider network with *five* fixed resistors rather than three.

6.2.5 Testing a 250 V meter with a 10 V standard

Suppose you need to check the calibration of a DC voltmeter with a measurement range of 0 to 250 Volts. However, the only trusted voltage standard you happen to have in the laboratory only outputs 10.000 Volts DC. How could you use an adjustable DC voltage source, a six-decade Kelvin-Varley voltage divider, and a null detector in conjunction with the 10-Volt standard to thoroughly test the voltmeter throughout its 0-250 Volt measurement range? Describe the circuit you will build with these components in detail.

Next, determine which of these test voltages may be accurately generated to test the DC voltmeter and what the corresponding six-digit setting would be for the Kelvin-Varley divider, describing the procedure you would use to configure the connected components:

- 25 Volts
- 50 Volts
- 75 Volts
- 100 Volts
- 125 Volts
- 150 Volts
- 175 Volts
- 200 Volts
- 225 Volts
- 250 Volts

Challenges

- Would having a Kelvin-Varley divider with more digits (more decades) help make some of these tests more practical?
- For the test voltages in this list that cannot be accurately replicated, identify any alternative voltage values that are possible.

6.2.6 Measuring time in Python

Try running this computer program written in the Python language, and determine the unit of time measurement for one "tick":

```
import time

ticks = time.time()
last_ticks = ticks

while (1):
  ticks = time.time()
  print("Number of ticks:", ticks)
  print("Interval in ticks since last read:", ticks - last_ticks)
  last_ticks = ticks
  input()
```

Mathematically convert any of the "ticks" values into an approximate date.

Also, explain the purpose of the `last_ticks` variable in this program.

Finally, replace the user prompt `input()` with the instruction `time.sleep(1)` and comment on how this alteration affects the execution of the program.

Challenges

- Why must the time be read in "ticks" and the `last_ticks` variable initialized prior to entering the `while()` loop?
- What is the smallest interval of time usable by the `sleep()` method?

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 Thermometer calibration results

Suppose a thermometer has a measurement range of 0 to 400 degrees Celsius. When subjected to a series of known temperatures from a trusted temperature reference, it responds as shown in the following "As-Found" table. This table's data was collected over time from top to bottom, so that the instrument was tested with increasing stimulus and then with decreasing stimulus:

Reference value	**Instrument reading**
0	0
100	98
200	195
300	297
400	400
300	302
200	206
100	101
0	0

What type(s) of calibration error(s) is/are apparent from this "As-Found" test result?

Also, identify the test point in this "As-Found" table with the largest error, and express that error both as *percent of span* and as *percent of value*.

Challenges

- Identify what some other common calibration errors would appear as when represented as numerical values in the "As-Found" table.

6.3.2 Voltmeter calibration results

Suppose a voltmeter has a measurement range of −10 to +10 Volts DC. When subjected to a series of known voltages from a trusted voltage reference, it responds as shown in the following "As-Found" table. This table's data was collected over time from top to bottom, so that the instrument was tested with increasing stimulus and then with decreasing stimulus:

Reference value	**Instrument reading**
0.00000	0.00005
2.00000	2.00005
4.00000	4.00005
6.00000	6.00005
8.00000	8.00005
10.00000	10.00005
8.00000	8.00005
6.00000	6.00005
4.00000	4.00005
2.00000	2.00005
0.00000	0.00005

What type(s) of calibration error(s) is/are apparent from this "As-Found" test result?

Also, identify the test point in this "As-Found" table with the largest error, and express that error both as *ppm of span* and as *ppm of value*.

Challenges

- Identify what some other common calibration errors would appear as when represented as numerical values in the "As-Found" table.

6.3.3 Ammeter calibration results

Suppose an ammeter has a measurement range of 0 to +50 Amperes DC. When subjected to a series of known currents from a trusted current reference, it responds as shown in the following "As-Found" table. This table's data was collected over time from top to bottom, so that the instrument was tested with increasing stimulus and then with decreasing stimulus:

Reference value	**Instrument reading**
0.00000	0.00000
10.00000	9.99998
20.00000	19.99996
30.00000	29.99994
40.00000	39.99992
50.00000	49.99990
40.00000	39.99992
30.00000	29.99994
20.00000	19.99996
10.00000	9.99998
0.00000	0.00000

What type(s) of calibration error(s) is/are apparent from this "As-Found" test result?

Also, identify the test point in this "As-Found" table with the largest error, and express that error both as *ppm of span* and as *ppm of value*.

Challenges

- Identify what some other common calibration errors would appear as when represented as numerical values in the "As-Found" table.

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

"732C/734C DC Reference Standard Operators Manual", Fluke Corporation, Everett, WA, 2018.

Calibration: Philosophy In Practice, second edition, Fluke Corporation, Everett, WA, 1994.

Fundamental Physical Constants – Complete Listing, 2018 CODATA Adjustment, National Institute of Standards and Technology, `http://physics.nist.gov/constants`, accessed 2 April 2023.

Elmquist, Randolph E.; Cage, Marvin E.; Tang, Yi-hua; Jeffery, Anne-Marie; Kinard, Joseph R.; Dziuba, Ronald F.; Oldham, Nile M.; Williams, Edwin R.; "The Ampere and Electrical Standards", Journal of Research of the National Institute of Standards and Technology, Volume 106, Number 1, National Institute of Standards and Technology, Gaithersburg, MD, January-February 2001.

Farmer, F. Malcolm, *Electrical Measurements in Practice*, first edition, McGraw-Hill Book Company Inc., New York, NY, 1917.

"Floating Measurements and Guarding", Application Note 123, Hewlett-Packard, Palo Alto, CA.

Hamer, Walter J., *Standard Cells – Their Construction, Maintenance, and Characteristics*, National Bureau of Standards monograph 84, National Bureau of Standards, 15 January 1965.

Hamon, B.V., "a 1-100 Ohm build-up resistor for the calibration of standard resistors", Journal of Scientific Instruments, Volume 31, Number 12, 1954.

"Intro to VNA Basics" primer, document 70W-60918-0, Tektronix, March 2017.

Lombardi, Michael A., *NIST Time and Frequency Services*, NIST Special Publication 432 (2002 Edition), National Institute of Standards and Technology, January 2002.

Lombardi, Michael A.; Heavner, Thomas P.; Jefferts, Steven R.; "NIST Primary Frequency Standards and the Realization of the SI Second", NCSL International Measure – the journal of measurement science, Volume 2, Number 4, December 2007.

Lombardi, Michael A. and Nelson, Glenn K.; "WWVB: A Half Century of Delivering Accurate

Frequency and Time by Radio", Journal of Research of the National Institute of Science and Technology, Volume 119, http://dx.doi.org/10.6028/jres.119.004, published 12 March 2014.

"Low Level Measurements Handbook", seventh edition, document 1KW-1559-0, Tektronix, February 2016.

"LXI Reference Design Overview", LXI Consortium Incorporated, Niwot, CO, 16 March 2016.

"Model 2758 Low Thermal EMF Binding Posts", Low Thermal Electronics Inc., Itasca, TX.

Newell, David B. and Tiesinga, Eite, *The International System of Units*, NIST Special Publication 330, 2019 Edition, CODEN: NSPUE2, Physical Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, August 2019.

Riley, William J. "A History of the Rubidium Frequency Standard".

"Rubidium Oscillator Specification", part number LFRBXO059244Bulk, data sheet version 1.0018, IQD Frequency Products, 5 September 2021.

"Standard Commands for Programmable Instruments (SCPI) Volume 1: Syntax and Style", version 1999.0, SCPI Consortium, San Diego, CA, May 1999.

"Standard Commands for Programmable Instruments (SCPI) Volume 2: Command Reference", version 1999.0, SCPI Consortium, San Diego, CA, May 1999.

"Standard Commands for Programmable Instruments (SCPI) Volume 3: Data Interchange Format", version 1999.0, SCPI Consortium, San Diego, CA, May 1999.

"Standard Commands for Programmable Instruments (SCPI) Volume 4: Instrument Classes", version 1999.0, SCPI Consortium, San Diego, CA, May 1999.

Stout, Melville B., *Basic Electrical Measurements*, Prentice-Hall Incorporated, Englewood Cliffs, NJ, 1950.

Taylor, Barry N. and Kuyatt, Chris E., *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297, National Institute of Standards and Technology, Gaithersburg, MD, September 1994.

"The LXI Primer" version 1.0, LXI Consortium Incorporated, Niwot, CO, 2012.

Thomas, James L.; Peterson, Chester; Cooter, Irvin L.; Kotter, F. Ralph; "An Absolute Measurement of Resistance by the Wenner Method", Journal of Research of the National Bureau of Standards, Research paper RP2029, Volume 43, National Bureau of Standards, October 1949.

"TVS/Zener Theory and Design Considerations", revision 0, document HBD854/D, On Semiconductor, June 2005.

"Ultra High Precision Z Foil Through-Hole Resistor", document number 63187, Vishay Precision

Group Inc., 1 November 2016.

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.

7 May 2026 – added explanations to some of the Case Tutorial sections regarding the SCPI control Python scripts where we needed to use the `,=` operator to convert data *lists* returned from certain methods to single numerical values. This programming construct was explained in the Case Tutorial section on controlling a LabJack DAQ unit, but not in any of the SCPI control sections. Any reader going first to those SCPI sections might miss this detail, or wonder if it's a typographical error in the code!

22-29 April 2026 – minor edits to the Tutorial.

24 February 2026 – added a new Historical Reference section showing the operation of a balance scale.

17 November 2025 – minor edits to the Tutorial.

12 November 2025 – added a quote from the KJV Bible (Amos 8:4-6) as an historical reference of dishonest weights and measures, and made some other minor edits to the Tutorial. Also added more explanatory text regarding burst (popcorn) noise.

7 November 2025 – added "wire resistance" labels to image_7056.

5 October 2025 – added index entries for S-parameters and also a brief explanation of the name (scattering parameters).

20-23 April 2025 – added cross-references to Hamon voltage divider networks in places where a 10:1 precision voltage divider was useful or necessary. Also edited image_7418 and added more explanatory text to the end of the "Hamon voltage dividers" subsection of the "Electrical metrology tools" section of the Tutorial chapter. Also corrected another typographical error courtesy of Jacob

Stormes. Also changed "phase comparator" to "phase detector" with reference to phase-locked loop technology.

16-17 April 2025 – corrected multiple typographical errors. One of them in the "Ammeter calibration results" Diagnostic Reasoning question courtesy of Galen Bennett and another one in the "Clock accuracy" Conceptual Reasoning question courtesy of Jacob Stormes, and multiple in the Tutorial courtesy of Jacob Stormes.

21 December 2024 – added more photographs and explanations and index entries to the Tutorial section on automated testing. Also, changed all instances of *test jig* to *test fixture*.

11 December 2024 – added a new Conceptual Reasoning question on analog multimeter calibration.

1-3 December 2024 – added a new Tutorial section on Hamon voltage dividers.

29-30 November 2024 – added a Derivations and Technical References section on elementary statistics. Also added a new Tutorial section on guarded measurements.

21 November 2024 – expanded the "Measuring time in Python" Quantitative Reasoning question.

14 November 2024 – edited some of the instructor notes, and corrected errors in the table for the "Ammeter calibration results" Diagnostic Reasoning question.

9 November 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.

5 July 2024 – replaced some non-ASCII "degree" symbols that got copied and pasted from historical books into various Historical References sections.

22-25 April 2024 – minor edits to the Tutorial, as well as more Challenge questions.

20 April 2024 – added an introductory subsection to the Case Tutorial section on SCPI control of a power supply and multimeter. Also added other sections showing SCPI instrument control, including one with an oscilloscope and another using PySimpleGUI for an elementary graphic user interface.

18 April 2024 – renamed "Frequency and time metrology standards" subsection to "Frequency and time standards". Also corrected some typographical errors. Also elaborated on the distinction between calibration and standardization.

13 April 2024 – added more instructor notes as well as a new image (image_7213) to the Tutorial showing a Josephson Junction.

6-7 April 2024 – added more content to the Case Tutorial section on SCPI instrument control, and also moved multiple questions related to instrument calibration errors from the Quantitative Reasoning section to the Diagnostic Reasoning section.

26-29 March 2024 – added a new Tutorial section on automated electrical testing, and corrected a copy-and-paste error in the Case Tutorial section showcasing SCPI control of instruments.

9 March 2024 – added a new Case Tutorial section showing simple acquisition of analog voltage data from a LabJack model U3 DAQ using the Python programming language.

3 February 2024 – added a new Tutorial section specifically for the discussion of measurement uncertainty.

24 January 2024 – added some clarifying comments and graphics to the Tutorial section on four-wire resistance measurement, showing how this technique need not be performed with a special 4-wire instrument.

10 January 2024 – added a new Case Tutorial section showing a typical serial data stream collected from a GPS receiver. Also added comments in the "Frequency and time measurements" section of the Tutorial chapter explaining how phase-locked loop (PLL) technology is used within GPS-disciplined oscillators.

4 January 2024 – added TSR as an alternative expression for TUR.

2-3 July 2023 – added a Case Tutorial section demonstrating SCPI remote control of a programmable power supply and programmable multimeter.

23 June 2023 – added a photograph showing 4-wire resistance measurement in action.

15 June 2023 – added a new Case Tutorial section showing catastrophic results of an automated (script-driven) test of a current-sensing device.

4 May 2023 – added photographs of GPS-disciplined clock standards.

25 March to 29 April 2023 – added content throughout the module, including photographs from Fluke calibration lab tour in 2013 and text for the Tutorial sections, Historical References chapter, Introduction chapter, and the Derivations and Technical References chapter.

21 November 2022 – document first created.

Index

Absolute zero, 62
Absolute zero temperature, 91
Accelerated stress testing, 126
Accuracy, 48
Adding quantities to a qualitative problem, 212
Agilent model 3458A digital multimeter, 45
AM, 80
Ampere, definition of, 77
Amplifier, operational, 104
Amplitude modulation, 80
Annotating diagrams, 211
Application layer, 131
Artifact standard, 46, 83
ASCII, 41
Atomic clock, 68, 79
Automated testing, 115
Avalanche noise, 64
Average, 179

Bed-of-nails test fixture, 119
Beta, transistor, 129
Bipolar junction transistor, 63
BJT, 63
Blue noise, 61
Bond wire, 91
Breakdown, dielectric, 165
Bridge, Wheatstone, 145–147
Brownian noise, 61
Burst noise, 64

Cable gland, 126
Calibration, 48
Calibration drift, 45
Carbon composition resistor, 63
Cathode-ray tube, 168
CGS, 176
Charge pump, 41
Checking for exceptions, 212
Checking your work, 212
Clark cell, 60
Clark standard cell, 65, 137
Coaxial cable, 104
CODATA, 51
Code, computer, 221
Coefficient, temperature, 60, 65, 70
Coefficient, voltage, 58
Component values, IEC standard, 150
Confidence interval, 51
Conservation of Energy, 11
Conservation of energy, 79
Conventional flow notation, 168
Cooper pair, 67
Coriolis flowmeter, 83
Coulomb, 165–167
Cross product, 168
Cross-product, 167
CRT, 168
Crystal, 83
Current, excitation, 162
Current, magnetization, 162

Daniell standard cell, 65
DAQ, 15
Data acquisition unit (DAQ), 15
Data link layer, 131
Data set, 179
De la Rue standard cell, 65
Decade, 109
Device Under Test, 38, 40, 115
Dielectric breakdown, 165
Dielectric strength, 165
Differential galvanometer, 145
Dimensional analysis, 169, 211
Distribution curve, 50

Drift, 60
Drift, calibration, 45
Driven guard, 104
DUT, 38, 40, 115

Eddy current, 156
Edwards, Tim, 222
EKG, 195
Electrocardiogram, 195
Electromotive force, 175
Electron flow notation, 168
Elementary charge of an electron, 67, 75, 76
emf, 175
Energy, atomic electrons, 79
Environmental test chamber, 126
Equivalent Series Resistance, 152, 155
ESR, 152, 155
Excitation current, 162

Faraday's Law of Electromagnetic Induction, 162
Ferromagnetic material, 58
FET, 62
Field intensity, magnetic, 175
Field-effect transistor, 62
Filter, 61
Fixture, test, 97, 119, 120, 123
Flicker noise, 63
Flowmeter, Coriolis, 83
Fluke 744 Documenting Process Calibrator, 118
Fluke model 5720A Multifunction Calibrator, 45
Fluke model 5730A Multifunction Calibrator, 106
Fluke model 754 Documenting Process Calibrator, 45
Flux density, magnetic, 171
Flying probe tester, 119
Frequency, 78

Galvanometer, 145
Gauss, 171
Gaussian curve, 50
Gland, cable, 126
Global Positioning System, 78, 81
GPIB, 130
GPS, 78, 81
GPS-disciplined oscillator, 82
Graph values to solve a problem, 212
Graphic user interface, 36
Greenleaf, Cynthia, 183
Gross error, 49
Guard shield, 105
Guard terminal, 101
Guard, driven, 104
GUI, 36
Guoy standard cell, 65

Hall effect, 75
Hamon voltage divider, 88, 113
Harmonic frequency, 61
Helmholtz standard cell, 65
HF radio, 80
High-pass filter, 61
Histogram, 50
How to teach with these modules, 219
HP-IB, 130
Hwang, Andrew D., 223
Hysteresis error, 59

ICL3232 serial transceiver, 41
Identify given data, 211
Identify relevant principles, 211
IEC 60063 standard, 150
IEC standard component values, 150
IEEE-488, 130
Inductance, mutual, 73
Inductive reactance, 162
Insertion resistance, 85
Intermediate results, 211
Interval, confidence, 51
Intrinsic standard, 46, 83
Inverted instruction, 219

Java, 130
JFET, 63
Johnson, John, 62
Josephson array, 68
Josephson constant K_J, 67
Josephson effect, 67, 77
Josephson junction, 67
Josephson, Brian David, 67
Joule's Law, 85
Junction field-effect transistor, 63

Kelvin four-wire method, 93
Kelvin-Varley voltage divider, 107, 202
Kirchhoff's Current law, 99
Kirchhoff's Voltage Law, 86
Knuth, Donald, 222

LabJack, 15
Lamport, Leslie, 222
Laser, 79
Left-hand rule, 168
LF radio, 80
Life testing, 126
Limiting cases, 212
Linear Technology, 72
Location parameter, 50
Lorentz force, 167
Low-pass filter, 61
LTZ1000, 72
Lua, 130

Magnetic circuit, 175
Magnetic field intensity, 175
Magnetic flux density, 171
Magnetic remanence, 59
Magnetic saturation, 58
Magnetization current, 162
Magnetomotive force, 38, 175
Manganin alloy, 63
Maser, 79
Maxwell, James Clerk, 133
Mean, 179
Median, 179
Megger, 101
Metacognition, 188
Metal film resistor, 63
Metric system, CGS, 176
Metric system, SI, 176
Metrology, 44, 50, 109, 136
Mho, 144
Miller, Ariana, 128
mmf, 38, 175
Mode, 179
Moolenaar, Bram, 221
Multi-port instrument, 115
Multifunction calibrator, 106
Murphy, Lynn, 183
Mutual inductance, 73

National Bureau of Standards, 45
National Institute of Standards and Technology, 45
National Measurement Institute, 46
NBS, 45
Negative feedback, 136
Newton, 165–167, 169
NIST, 45, 51
NMEA-0183 data format, 41, 42, 81
NMI, 46
NOAA, 171
Noise shield, 105
Noise, avalanche, 64
Noise, blue, 61
Noise, Brownian, 61
Noise, burst, 64
Noise, flicker, 63
Noise, pink, 61
Noise, popcorn, 64
Noise, red, 61
Noise, shot, 63
Noise, thermal, 62
Noise, violet, 61
Noise, white, 61
Non-monotonicity, 58
Normal distribution, 50
Null, 95
Null balance technique, 136
Null detector, 10, 86, 136
Nyquist, Harry, 62

OCXO, 84
Ohm's Law, 73, 77, 99
Ohm, definition of, 73
Opamp, 104
Open-source, 221
Operational amplifier, 104, 136
OSI reference model, 131

Parasitic effect, 152, 155, 159
Particle accelerator, 168
Parts per billion, 84
Parts per billion (ppb), 54
Parts per million, 84

Parts per million (ppm), 54
Parts per trillion (ppt), 54
Percent, 54, 84
Periodic waveform, 61
Permeability, 171
Permeability of free space, 73, 77
Permittivity, 166
Phase-locked loop, 82
Physical layer, 131
Pierce oscillator, 83
Piezoelectricity, 83
Pink noise, 61
Planck constant, 67, 75, 76
PLL, 82
Popcorn noise, 64
Population, 179
Population standard deviation, 179
Population variance, 179
Port, instrument, 115
ppb, 84
ppm, 84
Primary standard, 46
Problem-solving: annotate diagrams, 211
Problem-solving: check for exceptions, 212
Problem-solving: checking work, 212
Problem-solving: dimensional analysis, 211
Problem-solving: graph values, 212
Problem-solving: identify given data, 211
Problem-solving: identify relevant principles, 211
Problem-solving: interpret intermediate results, 211
Problem-solving: limiting cases, 212
Problem-solving: qualitative to quantitative, 212
Problem-solving: quantitative to qualitative, 212
Problem-solving: reductio ad absurdum, 212
Problem-solving: simplify the system, 211
Problem-solving: thought experiment, 173, 174, 211
Problem-solving: track units of measurement, 211
Problem-solving: visually represent the system, 211
Problem-solving: work in reverse, 212
Pulling, oscillator frequency, 82
Python, 24, 130

Qualitatively approaching a quantitative problem, 212
Quantized energy levels, 79
Quantum Hall effect, 75, 77
Quartz crystal, 83

Radiation test chamber, 126
Random error, 49
Range, 48
Reactance, inductive, 162
Reading Apprenticeship, 183
Red noise, 61
Reductio ad absurdum, 212, 218, 219
Reichsanstalt standard resistance, 140
Remanence, magnetic, 59
Repeatability, 48
Resistor type, 63
Resolution, 48
RF test chamber, 126
Right-hand rule, 168
Rolloff, 61

S-parameters, 116
Safety shield, 105
SAM-M8Q GPS radio receiver, 41
Sample, 179
Sample standard deviation, 179
Sample variance, 179
Saturation, magnetic, 58
Scattering parameters, 116
Schoenbach, Ruth, 183
Scientific method, 188
SCPI, 38, 131
Script, programming, 24–26, 29, 33, 34, 38, 119, 121, 126, 130, 131
Secondary standard, 46
Seebeck coefficient, 91
Seebeck effect, 72, 90, 96
Seebeck, Thomas Johann, 90
Series, 38
Set, data, 179
Shield, guard, 105
Shield, noise, 105
Shield, safety, 105
Shielding, 104
Shortwave radio, 80

Shot noise, 63
SI, 176
SI Brochure, 9th edition, 46
Siemens, 144
Simplifying a system, 211
Skew, 52
SMU, 101, 104, 106
Soakage, capacitor, 152, 154
Socrates, 218
Socratic dialogue, 219
Solenoid, 173
Source-measure unit, 101, 104, 106
Span, 48
Span shift, 57
SPICE, 183
Stallman, Richard, 221
Standard component values, IEC, 150
Standard deviation, 50, 179
Standard, primary, 46
Standard, secondary, 46
Standard, tertiary, 46
Standard, working, 46, 65, 69, 74, 79
Standardization, 48, 72
Statistics, 179
Stress testing, 126
Superconductor, 166, 172
Systematic error, 49

Tank circuit, 155
TCXO, 84
Temperature, 62
Temperature coefficient, 63
Temperature coefficient of voltage, 60, 65, 70
Temperature transmitter, 118
Tertiary standard, 46
Tesla, 167, 169, 171, 173, 174
Test chamber, 126
Test fixture, 97, 119, 120, 123
Testing, automated, 115
Thévenin equivalent network, 85
Thermal EMF, 72, 90, 96
Thermal noise, 62
Thought experiment, 173, 174, 211
Time, 78
Toroid, 174
Torvalds, Linus, 221
Transconductance, transistor, 129
Transfer standard, 72
Transistor, bipolar junction, 63
Transistor, field-effect, 62
Transistor, junction field-effect, 63
Transmitter, temperature, 118
Triaxial cable, 104
Tube, vacuum, 63

Uncertainty, 49, 50
Units of measurement, 211

Vacuum conditions, 129
Vacuum tube, 63
Variance, 179
VCXO, 84
Vector, 167
Vector cross-product, 168
Vector network analyzer, 115
Violet noise, 61
Visualizing a system, 211
VNA, 115
Voltage, 175
Voltage coefficient, 58, 70
Voltage coefficient of capacitance, 58
von Klitzing constant R_K, 76
von Klitzing, Klaus, 75

Weber, 171
Weston standard cell, 65, 137
Wheatstone bridge, 95, 145–147
White noise, 61
Wire wound resistor, 63
WMM, 171
Wolff standard resistor, 140
Work in reverse to solve a problem, 212
Working standard, 46, 65, 69, 74, 79
World Magnetic Model, 171
WYSIWYG, 221, 222

Zener diode, 69, 70
Zener diode voltage reference, 59
Zero shift, 56