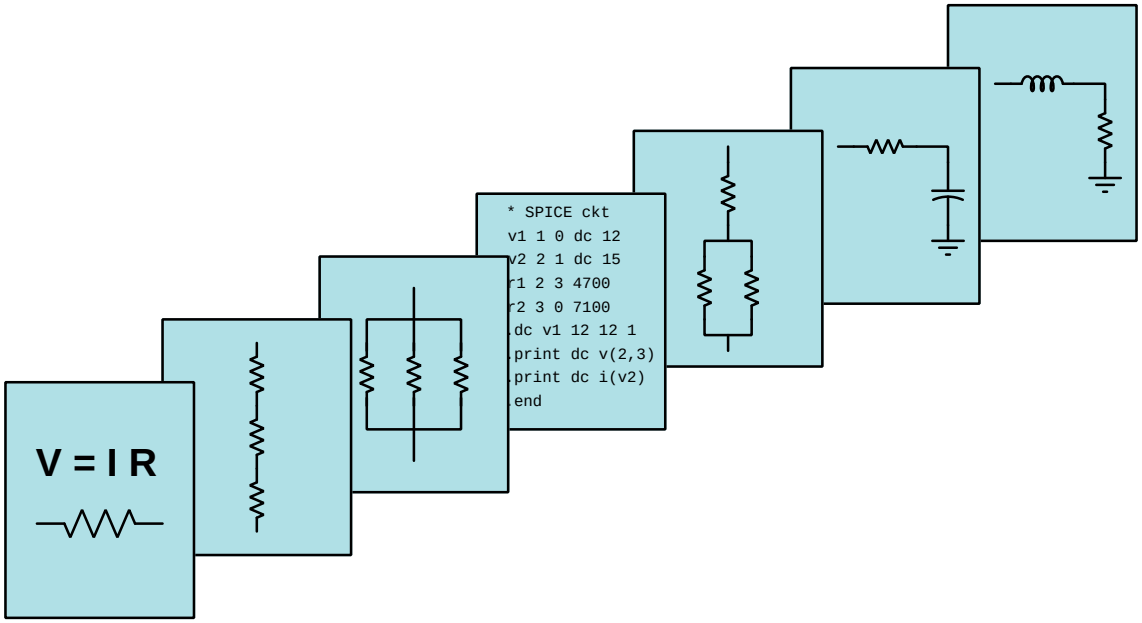
Modular Electronics Learning (ModEL) project



Analog Instrumentation Circuits

© 2023-2026 by Tony R. Kuphaldt – under the terms and conditions of the Creative Commons Attribution 4.0 International Public License

Last update = 10 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 analog instrumentation circuits . . . 5
1.3 Recommendations for instructors . . . 6

2 Tutorial **7**

2.1 4 to 20 mA analog current signals . . . 7
2.2 Relating 4 to 20 mA signals to instrument variables . . . 11
2.2.1 Example calculation: controller output to valve . . . 14
2.2.2 Example calculation: flow transmitter . . . 15
2.2.3 Example calculation: temperature transmitter . . . 18
2.2.4 Example calculation: pH transmitter . . . 21
2.2.5 Example calculation: reverse-acting I/P transducer signal . . . 23
2.2.6 Example calculation: PLC analog input scaling . . . 25
2.2.7 Graphical interpretation of signal ranges . . . 28
2.2.8 Thinking in terms of per unit quantities . . . 30
2.3 Instrument error documentation . . . 33
2.4 Controller output current loops . . . 35
2.5 4-wire ("self-powered") transmitter current loops . . . 38
2.6 2-wire ("loop-powered") transmitter current loops . . . 40
2.7 4-wire "passive" versus "active" output transmitters . . . 42
2.8 Instrument loop diagrams . . . 43
2.9 Troubleshooting current loops . . . 46
2.9.1 Using a standard milliammeter to measure loop current . . . 48
2.9.2 Using a clamp-on milliammeter to measure loop current . . . 50
2.9.3 Using "test" diodes to measure loop current . . . 51
2.9.4 Using shunt resistors to measure loop current . . . 53
2.9.5 Troubleshooting current loops with voltage measurements . . . 54
2.9.6 Using loop calibrators . . . 58
2.9.7 NAMUR signal levels . . . 65

3 Questions **67**
3.1 Conceptual reasoning . . . 71
3.1.1 Reading outline and reflections . . . 72
3.1.2 Foundational concepts . . . 73
3.1.3 Loop-powered transmitter circuit sketching . . . 75
3.1.4 1-5 VDC versus 4-20 mA signaling . . . 76
3.2 Quantitative reasoning . . . 78
3.2.1 Miscellaneous physical constants . . . 79
3.2.2 Introduction to spreadsheets . . . 80
3.2.3 Instrument range tables . . . 83
3.2.4 3-wire RTD temperature transmitter circuit . . . 85
3.3 Diagnostic reasoning . . . 86
3.3.1 Faults in a pressure control system . . . 87
3.3.2 Failed flow control system . . . 88
3.3.3 Faults in an air regulating system . . . 89
3.3.4 Failed reactor temperature control system . . . 91
3.3.5 Faulty pH measurement system . . . 93
3.3.6 Failed maple syrup process . . . 94

A Problem-Solving Strategies **97**

B Instructional philosophy **99**
B.1 First principles of learning . . . 100
B.2 Proven strategies for instructors . . . 101
B.3 Proven strategies for students . . . 103
B.4 Design of these learning modules . . . 104

C Tools used **107**

D Creative Commons License **111**

E References **119**

F Version history **121**

Index **122**

Chapter 1

Introduction

1.1 Recommendations for students

Many industrial sensing and control instruments use analog DC current signaling to convey information between the instrument and the control system it connects to. Learning how these signals – usually 4 to 20 milliAmperes DC – work is important for anyone aspiring to this field of employment.

Important concepts related to analog instrument signals and circuits include signal **range**, **span**, and **zero**, **Ohm's Law**, **process variable**, **setpoint**, **manipulated variable**, **linear equation**, **PLC**, **error** calculations, current **loop**, **negative feedback loop**, control valve **action**, electrical **sources** versus **loads**, and **loop diagrams**.

Due to the strong application of math to instrument signal ranges, the reader is urged to apply the mathematical principles as soon as possible. The Example Calculations contained in the tutorial serves this purpose well, as they allow you to apply the formulae discussed previously to an example and then check your work against the completed example to see if your application was correct. Do not simply read a quantitative example and assume you understand it just because nothing in the presentation seemed confusing. *Until you can perform the analysis yourself without assistance, you haven't mastered it!*

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

- How might an experiment be designed and conducted to explore the concept of negative versus positive feedback? What hypothesis (i.e. prediction) might you pose for that experiment, and what result(s) would either support or disprove that hypothesis?
- What are some practical applications of control systems?
- What does *slope* mean in the mathematical sense of the word, and how does this relate to analog signal ranges?
- Why are instrument calibrations usually checked with up-direction and down-direction tests?

- Identify some of the different types of common calibration errors, and how to discern each one from a table of data.
- How do the internal components of a two-wire transmitter work so that the resulting current both powers those components and conveys information about the variable being measured?
- What information is conveyed in a loop diagram that is specific/unique to the field of industrial measurement and control?
- What does it mean for a sensing instrument to be direct-acting versus reverse-acting?
- What purpose does a *loop calibrator* device serve?
- Why are common rectifying diodes often installed within current loop circuits?
- Why is it useful to know which devices function as sources versus which devices function as loads in a 4-20 mA signal circuit?
- What are some of the test modes that are usually offered by loop calibrator devices?
- What does the NAMUR standard specify, and how is this useful in industrial control systems?
- What does it mean for a final control element such as a control valve to be reverse-acting versus direct-acting?
- What does "per unit" mean with regard to instrument range calculations?
- What purpose does an *I/P transducer* serve in a control system?
- What is the difference between a *live-zero* signal versus a *dead-zero* signal?
- Why are *live-zero* signals generally preferred over *dead-zero* signals for conveying important process information?
- How might you alter one of the example analyses shown in the text, and then determine the behavior of that altered circuit?
- Devise your own question based on the text, suitable for posing to someone encountering this subject for the first time

1.2 Challenging concepts related to analog instrumentation circuits

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

- **Linear scaling** – the classical slope-intercept linear equation formula ($y = mx + b$) is quite useful in converting one signal range into another, and mastering this concept is key to being able to program digital devices (e.g. microcontrollers) to display measured variables represented by analog voltage signals in real units of measurement.
- **Voltage sources versus Current sources** – an voltage source maintains a constant output voltage regardless of current, while an ideal current source maintains a constant current regardless of voltage. Voltage sources are more common as sources of electrical power, and so they tend to be more familiar to students of electricity and electronics. 4-20 mA transmitters may act as current sources (4-wire) or as current regulators (2-wire), and in either case must be approached with the assumption of current being fixed (in proportion to the measured variable) and voltage being variable in accordance with Kirchhoff's Voltage Law.

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 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** – Mathematically relate instrument input/output values

  Assessment – Complete tables of input versus output quantities for instruments with given ranges; e.g. pose problems in the form of the "Instrument range tables" Quantitative Reasoning question.

- **Outcome** – Design a loop-powered transmitter circuit

  Assessment – Sketch wire placement in a pictorial diagram for a circuit employing a loop-powered 4-20 mA transmitter; e.g. pose problems in the form of the "Loop-powered transmitter circuit sketching" Conceptual Reasoning question.

- **Outcome** – Diagnose a faulted control system

  Assessment – Predict the effect(s) of a single component fault or a single condition change in a feedback control system; e.g. pose problems in the form of the "Faults in an air regulating system" Diagnostic Reasoning question.

  Assessment – Determine the probability of various component faults in a feedback control system given symptoms and measured values; e.g. pose problems in the form of the "Failed maple syrup process" Diagnostic Reasoning question.

  Assessment – Determine the value of various diagnostic tests in identifying the location and nature of a fault in a feedback control system given symptoms and measured values; e.g. pose problems in the form of the "Failed reactor temperature control system" Diagnostic Reasoning question.

- **Outcome** – Independent research

  Assessment – Locate 4-20 mA transmitter datasheets and properly interpret some of the information contained in those documents including measurement range, power supply voltage range, accuracy, etc.

Chapter 2

Tutorial

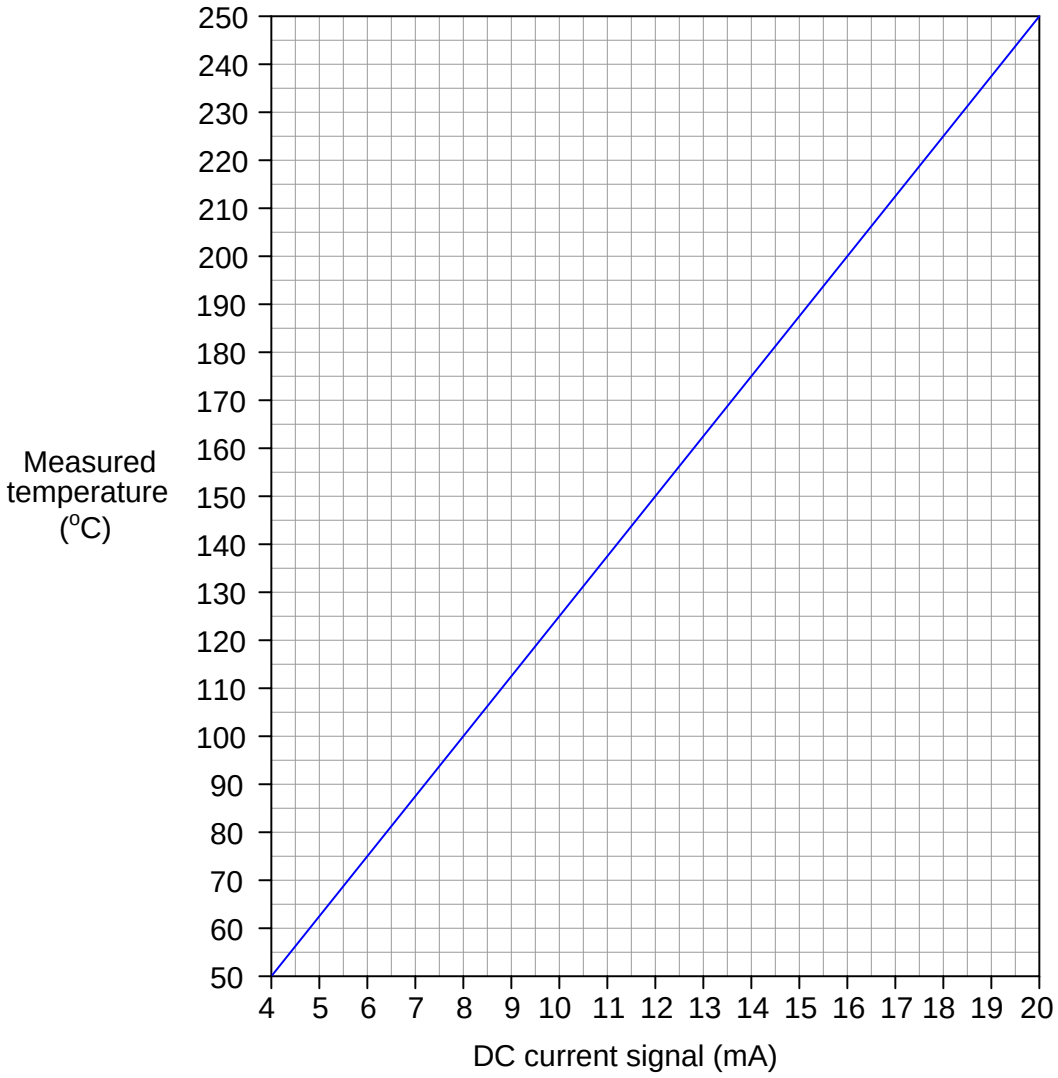
An "analog" electronic signal is a voltage or current proportionate to the value of some physical measurement or control quantity. An electronic-based measuring instrument is often classified as being "analog" simply by virtue of using an analog signal standard to communicate information, even if the internal construction and design of the instrument may be mostly digital in nature. This is to distinguish such instruments from those making use of no analog electronic signals at all (e.g. radio "wireless" or digital Fieldbus instruments).

2.1 4 to 20 mA analog current signals

The most popular form of signal transmission used in modern industrial instrumentation systems (as of this writing) is the 4 to 20 milliAmpere DC standard. This is an *analog* signal standard, meaning that the electric current is used to proportionately represent measurements or command signals. Typically, a 4 milliAmpere current value represents 0% of scale, a 20 milliAmpere current value represents 100% of scale, and any current value in between 4 and 20 milliAmperes represents a commensurate percentage in between 0% and 100%. The following table shows the corresponding current and percentage values for each 25% increment between 0% and 100%. Every technician tasked with maintaining 4-20 mA instruments commits these values to memory, because they are referenced so often:

Current value	**% of scale**
4 mA	0%
8 mA	25%
12 mA	50%
16 mA	75%
20 mA	100%

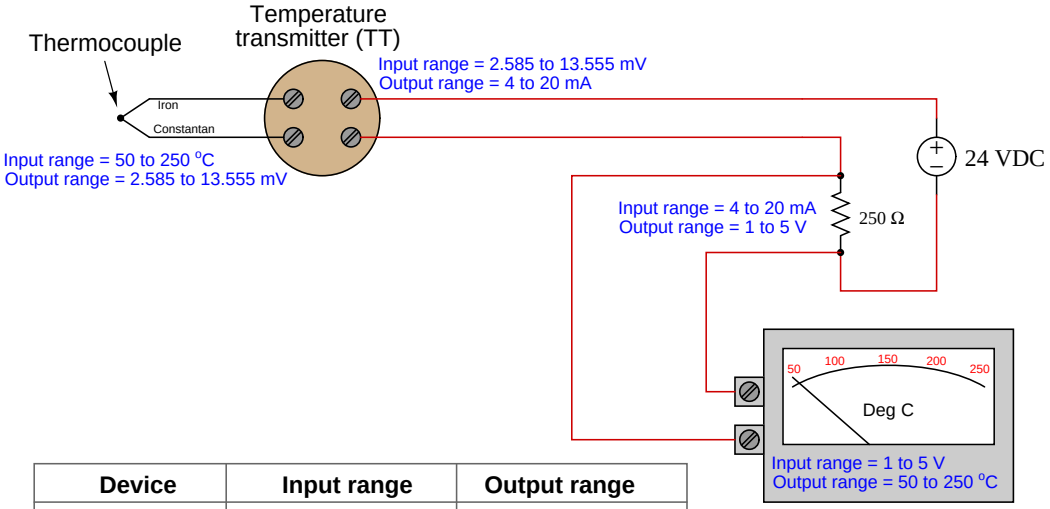
For example, if we were to calibrate a 4-20 mA temperature transmitter for a measurement range of 50 to 250 degrees C, we could relate the current and measured temperature values on a graph like this:



The 4-20 mA signal standard is referred to as *live zero* because its range begins with a non-zero value. This "live" zero provides a simple means of discriminating between a legitimate 0% signal value and a failed signal (e.g. leaking tube or severed cable)[1].

[1] Not all industrial measurement and control signals are "live zero" like the 4-20 mA standard. 0 to 10 Volts DC is a common "dead zero" signal standard, although far more common in environmental (building heating and cooling) control systems than industrial control systems. I once encountered an old analog control system using −10 Volts to

An important concept to grasp with all analog instrumentation is that instruments sending and receiving analog signals must be compatibly ranged in order to properly represent the desired variable. To illustrate, let us consider a temperature measurement system consisting of a thermocouple[2], a temperature transmitter whose function it is to regulate circuit current at some value proportional to the milliVoltage signal output by the thermocouple, a 250 Ohm resistor to convert the 4-20 mA analog signal into a 1-5 Volt analog signal, and a special voltmeter serving as a temperature indicator:



Device	Input range	Output range
Thermocouple	50 to 250 °C	2.585 to 13.555 mV
Transmitter	2.585 to 13.555 mV	4 to 20 mA
Resistor	4 to 20 mA	1 to 5 V
Indicator	1 to 5 V	50 to 250 °C

Note how the output range of each sending device matches the input range of its corresponding receiving device. If we view this system as a path for information to flow from the thermocouple's tip to the transmitter to the resistor and finally to the voltmeter/indicator, we see that the analog output range of each device must correspond to the analog input range of the *next* device, or else the real-world meaning of the analog signal will be lost.

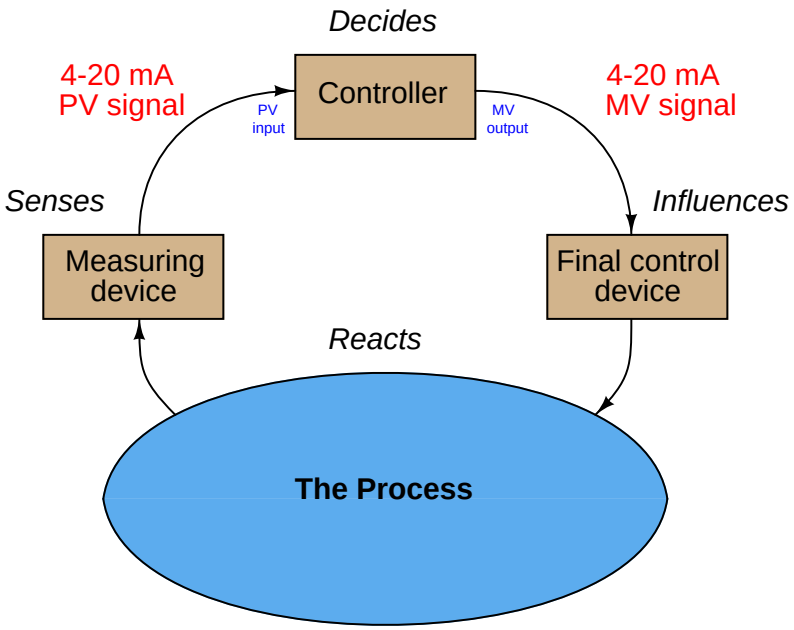
This correspondence does not happen automatically, but must be established by the technician building the system. In this case, it would be the technician's responsibility to properly adjust the range of the temperature transmitter, and also to ensure the indicator's display scale was properly labeled. Both the thermocouple and the resistor are non-adjustable devices, their input/output characteristics being fixed by physical laws.

+10 Volts as its analog signal range, which meant 0 Volts represented a 50% signal! A failed signal path in such a system could have been very misleading indeed, as a 50% signal value is not suspicious in the least.

[2]This is a temperature sensing element consisting of two different metal wires joined together, which generate a small voltage proportional to temperature. The correspondence between junction temperature and DC millivoltage is very well established by scientific testing, and so we may use this principle to sense process temperature.

DC current signals are also used in feedback control systems to command the positioning of a final control element, such as a control valve or a variable-speed motor drive (VSD or VFD). In these cases, the milliAmpere value does not directly represent a process measurement, but rather how the degree to which the final control element influences the process. Typically (but not always!), 4 milliAmperes commands a closed (shut) control valve or a stopped motor, while 20 milliAmperes commands a wide-open valve or a motor running at full speed. Final control elements often are equipped with adjustable ranges so that an accurate correspondence between the analog signal and the desired control action may be ensured.

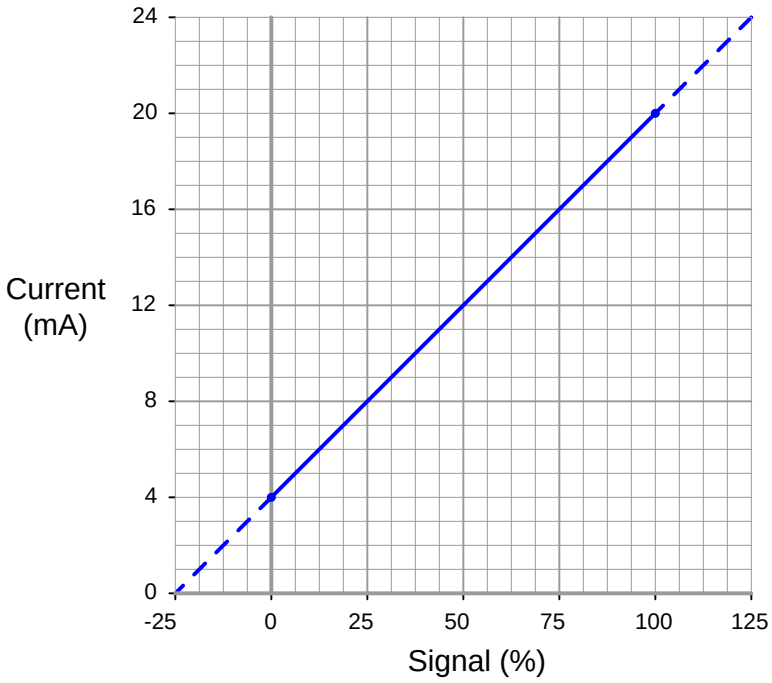
Thus, most industrial control systems use at least *two* different 4-20 mA signals: one to represent the process variable (PV) and one to represent the command signal to the final control element (the "manipulated variable" or MV):



The relationship between these two signals depends entirely on the response of the controller. There is no reason to ever expect the PV and MV current signals to be equal to each other except by chance, for they represent entirely different variables. In fact, if the controller is reverse-acting, it is entirely normal for the two current signals to be inversely related: as the PV signal increases going to a reverse-acting controller, the output signal will decrease. If the controller is placed into "manual" mode by a human operator, the output signal will have no automatic relation to the PV signal at all, instead being entirely determined by the operator's whim.

2.2 Relating 4 to 20 mA signals to instrument variables

A 4 to 20 mA current signal represents some signal along a 0 to 100 percent scale. Usually, this scale is linear as shown by this graph:



Being a linear function, we may use the standard slope-intercept linear equation to relate signal percentage to current values:

$$y = mx + b$$

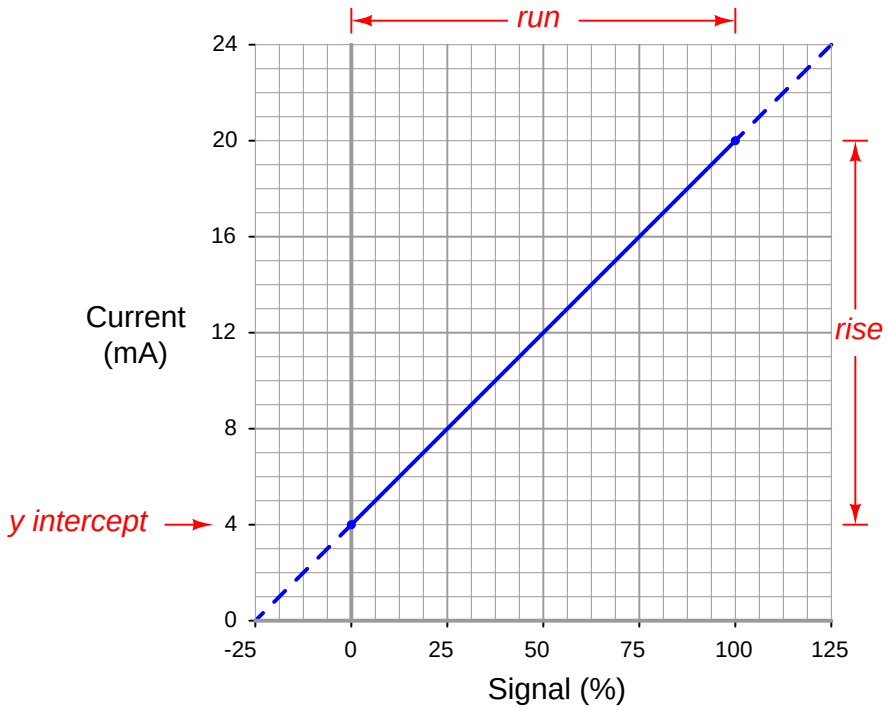
Where,

y = Output from instrument
x = Input to instrument
m = Slope
b = y-intercept point (i.e. the *live zero* of the instrument's range)

Once we determine suitable values for m and b, we may then use this linear equation to predict any value for y given x, and vice-versa. This is very useful for predicting the 4-20 mA signal output of a process transmitter, or the expected stem position of a 4-20 mA controlled valve, or any other correspondence between a 4-20 mA signal and some physical variable.

Before we may use this equation for any practical purpose, we must determine the slope (m) and intercept (b) values appropriate for the instrument we wish to apply the equation to. Next, we will see some examples of how to do this.

For the linear function shown, we may determine the slope value (m) by dividing the line's *rise* by its *run*. Two sets of convenient points we may use in calculating rise over run are 4 and 20 milliAmperes (for the rise), and 0 and 100 percent (for the run):



$$m = \frac{\text{Rise}}{\text{Run}} = \frac{(20 - 4)}{(100 - 0)} = \frac{16}{100}$$

$$y = \left(\frac{16}{100}\right) x + b$$

To calculate the y-intercept (b), all we need to do is solve for b at some known coordinate pair of x and y. Again, we find convenient points[3] for this task at 0 percent and 4 milliAmperes:

$$4 = \left(\frac{16}{100}\right) 0 + b$$

$$4 = 0 + b$$

$$b = 4$$

[3] We could have just as easily chosen 100 percent for x and 20 milliAmperes for y, for it would have yielded the same result of $b = 4$.

Now we have a complete formula for converting a percentage value into a milliAmpere value:

$$y = \left(\frac{16}{100}\right) x + 4$$

We may now use this formula to calculate how many milliAmperes represent any given percentage of signal. For example, suppose we needed to convert a percentage of 34.7% into a corresponding 4-20 mA current. We would do so like this:

$$y = \left(\frac{16}{100}\right) 34.7 + 4$$

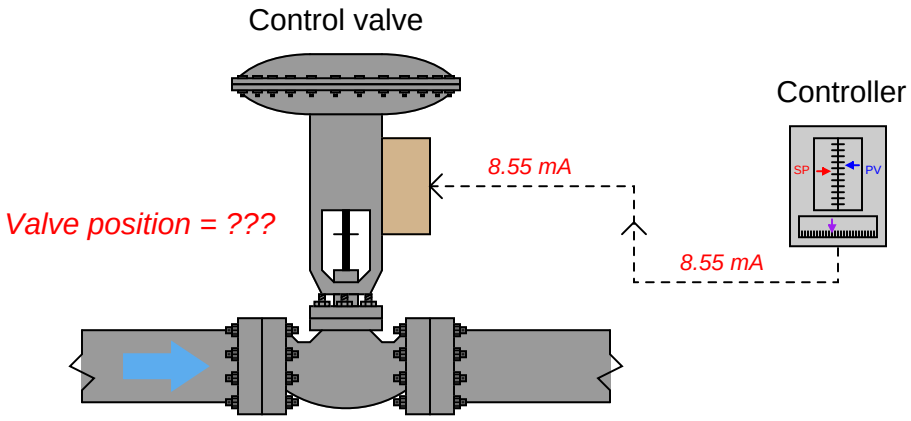
$$y = 5.552 + 4$$

$$y = 9.552$$

Thus, 34.7% is equivalent to 9.552 milliAmperes in a 4-20 mA signal range.

The slope-intercept formula for linear functions may be applied to *any* linear instrument, as illustrated in the following examples.

2.2.1 Example calculation: controller output to valve



An electronic loop controller outputs a signal of 8.55 mA to a direct-responding control valve (where 4 mA is shut and 20 mA is wide open). How far open should the control valve be at this MV signal level?

To solve for percentage of stem travel (x) at 8.55 milliAmperes of signal current (y), we may use the linear equation developed previously to predict current in milliAmperes (y) from signal value in percent (x):

$$y = \left(\frac{16}{100}\right) x + 4$$

$$8.55 = \left(\frac{16}{100}\right) x + 4$$

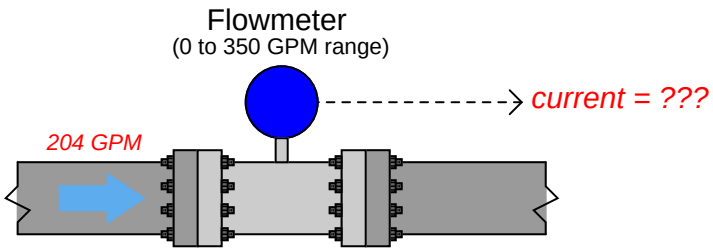
$$4.55 = \left(\frac{16}{100}\right) x$$

$$\left(\frac{100}{16}\right) 4.55 = x$$

$$x = 28.4$$

Therefore, we should expect the valve to be 28.4% open at an applied MV signal of 8.55 milliAmperes.

2.2.2 Example calculation: flow transmitter



A flow transmitter is ranged 0 to 350 gallons per minute, 4-20 mA output, direct-responding. Calculate the current signal value at a flow rate of 204 GPM.

One way we could solve for the amount of signal current is to convert the flow value of 204 GPM into a ratio of the flowmeter's full-flow value, then apply the same formula we used in the previous example relating percentage to milliAmperes. Converting the flow rate into a "per unit" ratio is a matter of simple division, since the flow measurement range is zero-based:

$$\frac{204 \text{ GPM}}{350 \text{ GPM}} = 0.583 \text{ per unit}$$

Converting a "per unit" ratio into percent merely requires multiplication by 100, since "percent" literally means "per 100":

$$0.583 \text{ per unit} \times 100\% = 58.3\%$$

Next, we plug this percentage value into the formula:

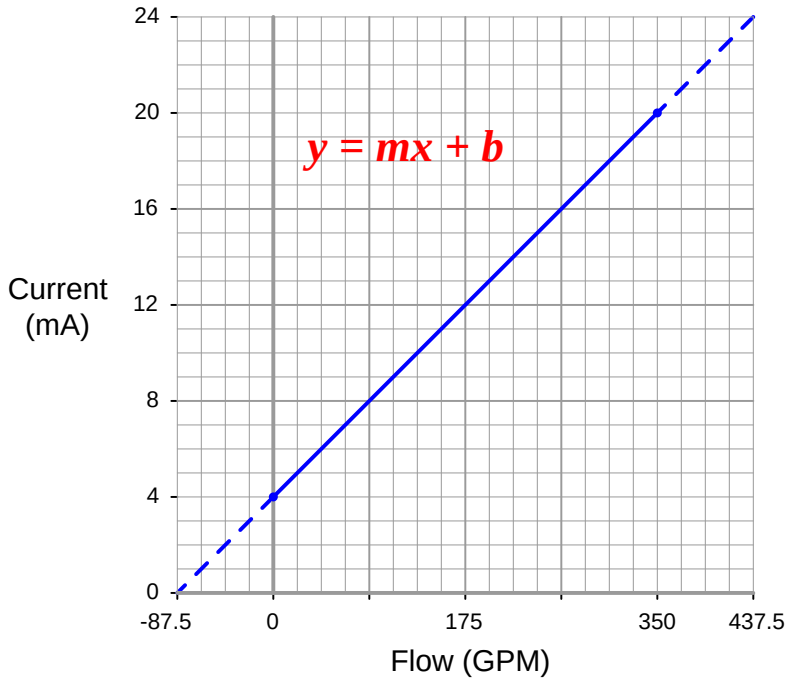
$$y = \left(\frac{16}{100}\right) 58.3 + 4$$

$$y = 9.33 + 4$$

$$y = 13.33$$

Therefore, the transmitter should output a PV signal of 13.3 mA at a flow rate of 204 GPM.

An alternative approach is to set up a linear equation specifically for this flowmeter given its measurement range (0 to 350 GPM) and output signal range (4 to 20 mA). We will begin this process by sketching a simple graph relating flow rate to current:



The slope (m) for this equation is rise over run, in this case 16 milliAmperes of rise for 350 GPM of run:

$$y = \left(\frac{20-4}{350-0}\right)x + b = \left(\frac{16}{350}\right)x + b$$

The y-intercept for this equation is 4, since the current output will be 4 milliAmperes at zero flow:

$$y = \left(\frac{16}{350}\right)x + 4$$

Now that the linear equation is set up for this particular flowmeter, we may plug in the 204 GPM value for x and solve for current:

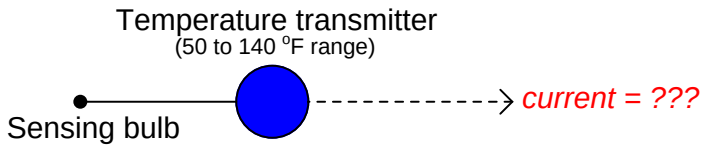
$$y = \left(\frac{16}{350}\right)204 + 4$$

$$y = 9.33 + 4$$

$$y = 13.33$$

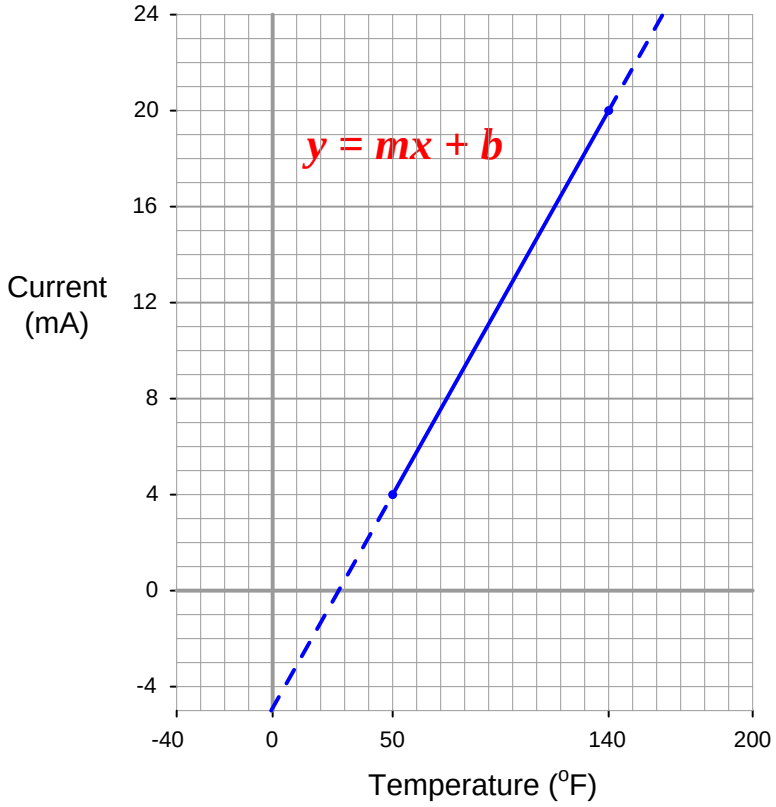
Just as before, we arrive at a current of 13.33 milliAmperes representing a flow rate of 204 GPM.

2.2.3 Example calculation: temperature transmitter



An electronic temperature transmitter is ranged 50 to 140 degrees Fahrenheit and has a 4-20 mA output signal. Calculate the current output by this transmitter if the measured temperature is 79 degrees Fahrenheit.

First, we will set up a linear equation describing this temperature transmitter's function:



Calculating and substituting the slope (m) value for this equation, using the full rise-over-run of the linear function:

$$y = \left(\frac{20 - 4}{140 - 50}\right) x + b = \left(\frac{16}{90}\right) x + b$$

The y-intercept value will be different[4] for this example than it was for previous examples, since the measurement range is not zero-based. However, the procedure for finding this value is the same – plug any corresponding x and y values into the equation and solve for b. In this case, I will use the values of 4 mA for y and 50 ^{o}F for x:

$$4 = \left(\frac{16}{90}\right) 50 + b$$

$$4 = 8.89 + b$$

$$b = -4.89$$

Therefore, our customized linear equation for this temperature transmitter is as follows:

$$y = \left(\frac{16}{90}\right) x - 4.89$$

At a sensed temperature of 79 ^{o}F, the transmitter's output current will be 9.16 mA:

$$y = \left(\frac{16}{90}\right) 79 - 4.89$$

$$y = 14.04 - 4.89$$

$$y = 9.16$$

[4] A common misconception for people learning to apply the slope-intercept formula to linear instrument ranges is that they tend to assume b will always be equal to the lower-range value (LRV) of the instrument's output range. For example, given a transmitter with a 4-20 mA output range, the false assumption is that $b = 4$. This does happen to be true *only if the instrument possesses a "dead-zero" input range*, but it will not be true for instruments with a live-zero input range such in this case here where the temperature input range is 50 to 140 degrees.

We may apply the same alternative method of solution to this problem as we did for the flowmeter example: first converting the process variable into a simple "per unit" ratio or percentage of measurement range, then using that percentage to calculate current in milliAmperes. The "tricky" aspect of this example is the fact the temperature measurement range does not begin at zero.

Converting 79 ^{o}F into a percentage of a 50-to-140 ^{o}F range requires that we first subtract the live-zero value, then divide by the span:

$$\text{Per unit ratio} = \frac{79 - 50}{140 - 50} = 0.3222$$

$$\text{Percentage} = 0.3222 \text{ per unit} \times 100\% = 32.22\%$$

Next, plugging this percentage value into our standard linear equation for 4-20 mA signals:

$$y = \left(\frac{16}{100}\right) 32.22 + 4$$

$$y = 5.16 + 4$$

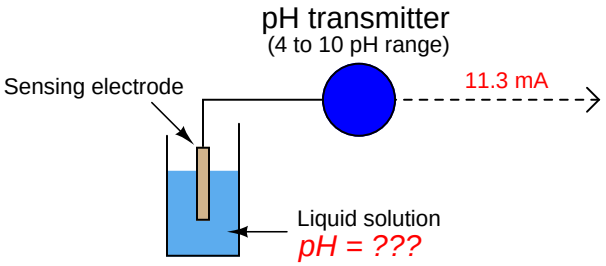
$$y = 9.16$$

Again, we arrive at the exact same figure for transmitter output current: 9.16 milliAmperes at a measured temperature of 79 ^{o}F.

The choice to calculate transmitter current by first setting up a "customized" linear equation for the transmitter in question or by converting the measured value into a percentage and using a "standard" linear equation for current is arbitrary. Either method will produce accurate results, although it could be argued that the "customized equation" approach may save time if many different current values must be calculated.

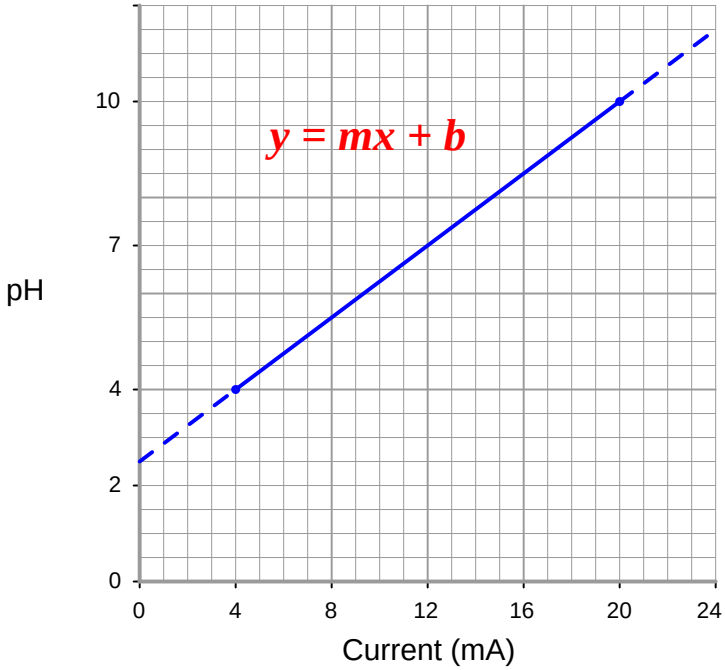
Certainly, if you are programming a computer to convert a received milliAmpere signal value into a measurement range (such as degrees Fahrenheit), it makes more sense to have the computer evaluate a single equation rather than perform multiple steps of calculations as we do when using percentage values as an intermediate step between the input and output value calculations. Evaluating one equation rather than two saves processing time.

2.2.4 Example calculation: pH transmitter



A pH transmitter has a calibrated range of 4 pH to 10 pH, with a 4-20 mA output signal. Calculate the pH sensed by the transmitter if its output signal is 11.3 mA.

First, we will set up a linear equation describing this temperature transmitter's function:



Note how we are free to set up 4-20 mA as the independent variable (x axis) and the pH as the dependent variable (y axis). We could arrange current on the y axis and the process measurement on the x axis as before, but this would force us to manipulate the linear equation to solve for x.

Calculating and substituting the slope (m) value for this equation, using the full rise-over-run of the linear function:

$$y = \left(\frac{10 - 4}{20 - 4}\right) x + b = \left(\frac{6}{16}\right) x + b$$

Solving for the y-intercept value using the coordinate values of 4 pH and 4 mA, we see again that this is an application where $b \neq 4$ mA. This is due to the fact that the instrument's input range (i.e. the domain of the $y = mx + b$ function) does *not* begin at zero:

$$4 = \left(\frac{6}{16}\right) 4 + b$$

$$4 = 1.5 + b$$

$$b = 2.5$$

Therefore, our customized linear equation for this pH transmitter is as follows:

$$y = \left(\frac{6}{16}\right) x + 2.5$$

Calculating the corresponding pH value for an output current signal of 11.3 mA now becomes a very simple matter:

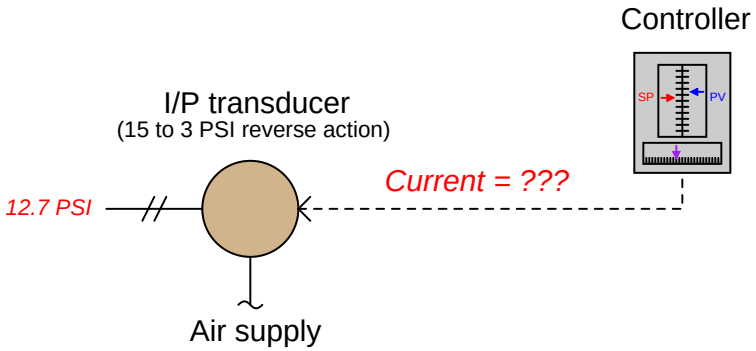
$$y = \left(\frac{6}{16}\right) 11.3 + 2.5$$

$$y = 4.24 + 2.5$$

$$y = 6.74$$

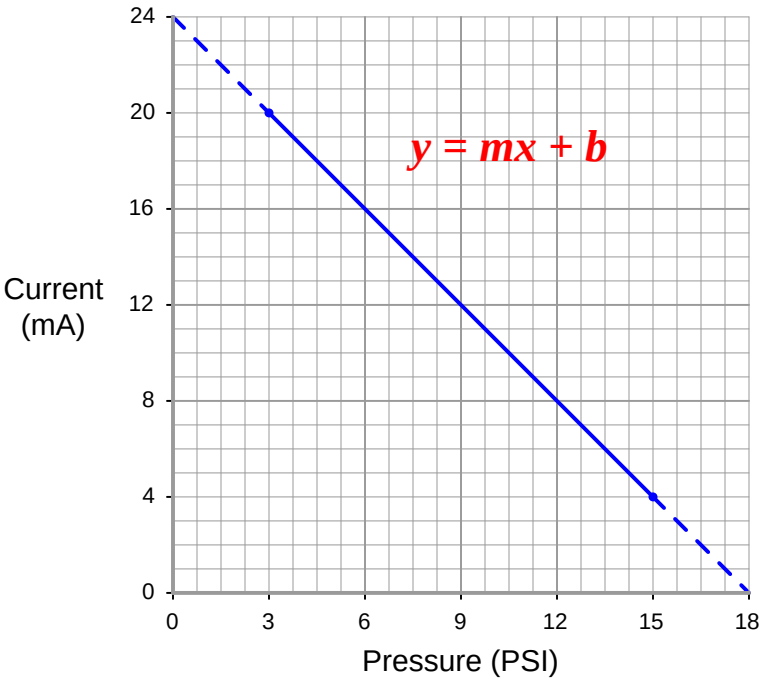
Therefore, the transmitter's 11.3 mA output signal reflects a measured pH value of 6.74 pH.

2.2.5 Example calculation: reverse-acting I/P transducer signal



A current-to-pressure transducer is used to convert a 4-20 mA electronic signal into a 3-15 PSI pneumatic signal. This particular transducer is configured for ***reverse action*** *instead of direct, meaning that its pressure output at 4 mA should be 15 PSI and its pressure output at 20 mA should be 3 PSI. Calculate the necessary current signal value to produce an output pressure of 12.7 PSI.*

Reverse-acting instruments are still linear, and therefore still follow the slope-intercept line formula $y = mx + b$, albeit with a negative slope:



Calculating and substituting the slope (m) value for this equation, using the full rise-over-run of the linear function. Note how the "rise" is actually a "fall" from 20 milliAmperes down to 4 milliAmperes, yielding a negative value for m:

$$y = \left(\frac{4-20}{15-3}\right)x + b = \left(\frac{-16}{12}\right)x + b = \left(-\frac{16}{12}\right)x + b$$

Solving for the y-intercept value using the coordinate values of 3 PSI and 20 mA:

$$20 = \left(-\frac{16}{12}\right)3 + b$$

$$20 = -4 + b$$

$$b = 24$$

Therefore, our customized linear equation for this I/P transducer is as follows:

$$y = \left(-\frac{16}{12}\right)x + 24$$

Calculating the corresponding current signal for an output pressure of 12.7 PSI:

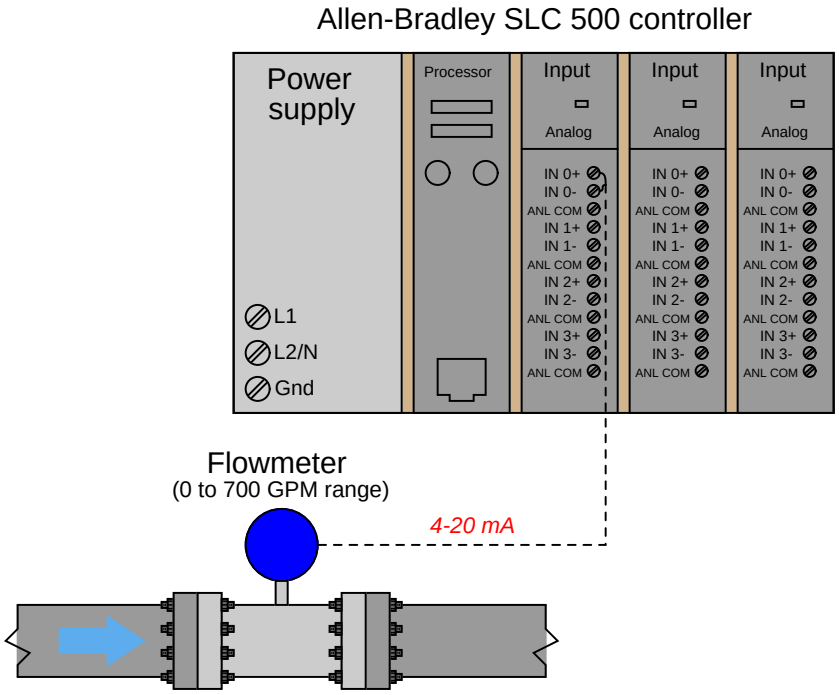
$$y = \left(-\frac{16}{12}\right)12.7 + 24$$

$$y = -16.93 + 24$$

$$y = 7.07$$

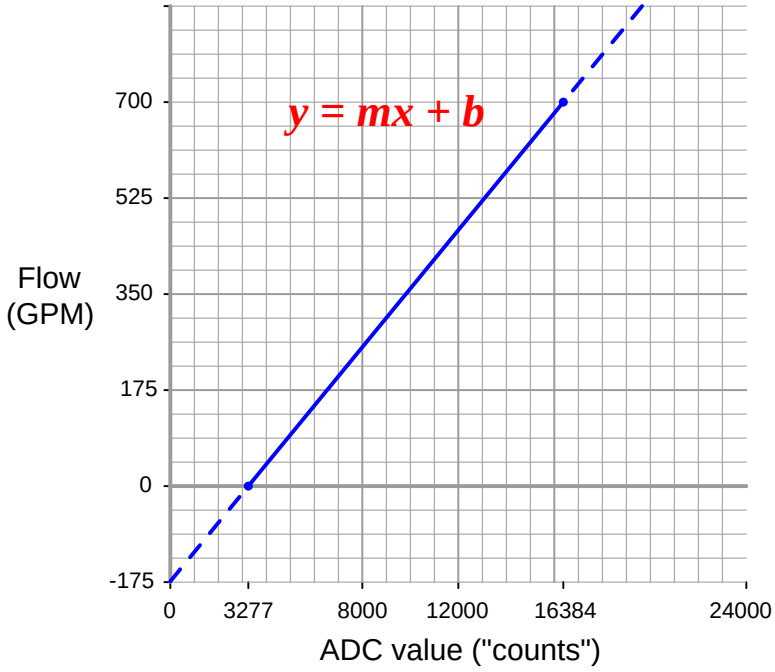
Therefore, a current signal of 7.07 mA is necessary to drive the output of this reverse-acting I/P transducer to a pressure of 12.7 PSI.

2.2.6 Example calculation: PLC analog input scaling



An Allen-Bradley SLC500 programmable logic controller (PLC) uses a 16-bit analog-to-digital converter in its model 1746-NI4 analog input card to convert 4-20 mA signals into digital number values ranging from 3277 (at 4 mA) to 16384 (at 20 mA). However, these raw numbers from the PLC's analog card must be mathematically scaled inside the PLC to represent real-world units of measurement, in this case 0 to 700 GPM of flow. Formulate a scaling equation to program into the PLC so that 4 mA of current registers as 0 GPM, and 20 mA of current registers as 700 GPM.

We are already given the raw number values from the analog card's analog-to-digital converter (ADC) circuit for 4 mA and 20 mA: 3277 and 16384, respectively. These values define the domain of our linear graph:



Calculating and substituting the slope (m) value for this equation, using the full rise-over-run of the linear function:

$$y = \left(\frac{700 - 0}{16384 - 3277}\right) x + b = \left(\frac{700}{13107}\right) x + b$$

Solving for the y-intercept value using the coordinate values of 0 GPM and 3277 ADC counts:

$$0 = \left(\frac{700}{13107}\right) 3277 + b$$

$$0 = 175 + b$$

$$b = -175$$

Therefore, our PLC scaling equation for this particular flowmeter is as follows:

$$y = \left(\frac{700}{13107}\right) x - 175$$

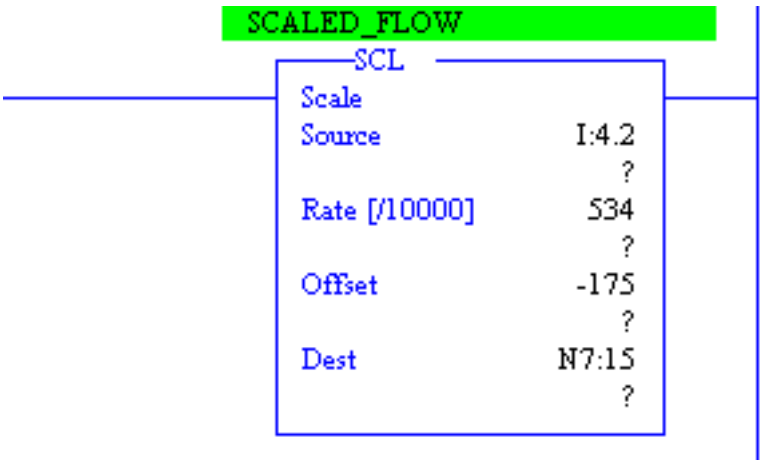
This type of scaling calculation is so common in PLC applications that Allen-Bradley has provided a special `SCL` ("scale") instruction just for this purpose. Instead of "slope" (m) and "intercept" (b), the instruction prompts the human programmer to enter "rate" and "offset" values, respectively. Furthermore, the rate in Allen-Bradley's `SCL` instruction is expressed as the numerator of a fraction where the denominator is fixed at 10000, allowing fractional (less than one) slope values to be specified using integer numbers. Aside from these details, the concept is exactly the same.

Expressing our slope of $\frac{700}{13107}$ as a fraction with 10000 as the denominator is a simple matter of solving for the numerator using cross-multiplication and division:

$$\frac{700}{13107} = \frac{r}{10000}$$

$$r = 534$$

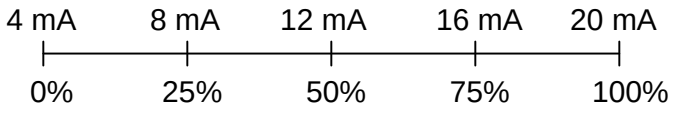
Thus, the `SCL` instruction would be configured as follows[5]



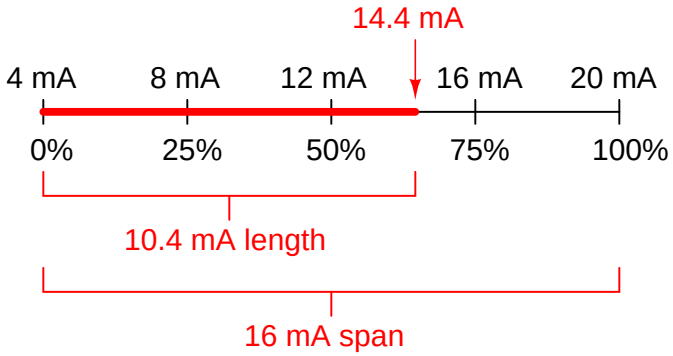
[5]The "Source" and "Dest" parameters shown in this instruction box refer to special addresses in the PLC's memory where the input (ADC count) and output (scaled flowrate) values will be found. You need not concern yourself with the meanings of `I:4.2` and `N7:15`, because these addresses are unimportant to the task of deriving a scaling formula.

2.2.7 Graphical interpretation of signal ranges

An illustration some students find helpful in understanding analog signal ranges is to consider the signal range as a *length* expressed on a number line. For example, the common 4-20 mA analog current signal range would appear as such:



If one were to ask the percentage corresponding to a 14.4 mA signal on a 4-20 mA range, it would be as simple as determining the length of a line segment stretching from the 4 mA mark to the 14.4 mA mark:



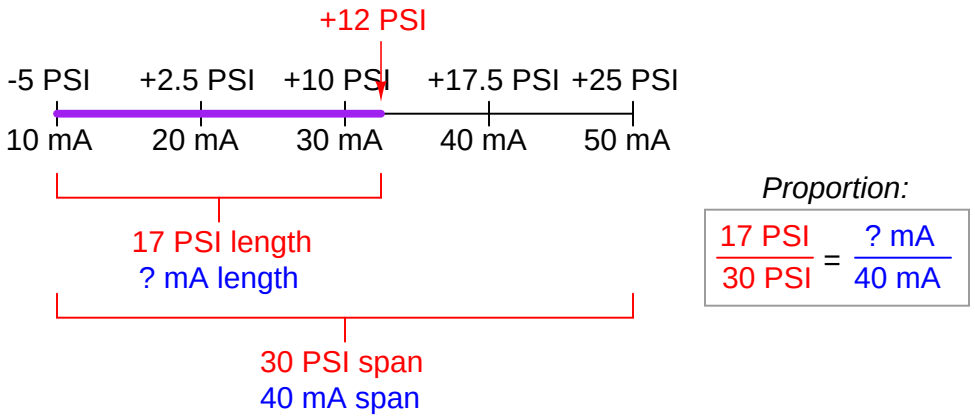
As a percentage, this thick line is 10.4 mA long (the distance between 14.4 mA and 4 mA) over a total (possible) length of 16 mA (the total span between 20 mA and 4 mA). Thus:

$$\text{Percentage} = \left(\frac{14.4 \text{ mA} - 4 \text{ mA}}{20 \text{ mA} - 4 \text{ mA}} \right) 100\%$$

$$\text{Percentage} = \left(\frac{10.4 \text{ mA}}{16 \text{ mA}} \right) 100\%$$

$$\text{Percentage} = 65\%$$

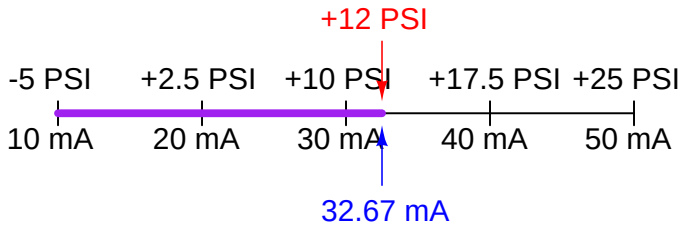
This same "number line" approach may be used to visualize any conversion from one analog scale to another. Consider the case of an electronic pressure transmitter calibrated to a pressure range of −5 to +25 PSI, having an (obsolete) current signal output range of 10 to 50 mA. The appropriate current signal value for an applied pressure of +12 PSI would be represented on the number line as such:



Finding the "length" of this line segment in units of milliAmperes is as simple as setting up a proportion between the length of the line in units of PSI over the total (span) in PSI, to the length of the line in units of mA over the total (span) in mA:

$$\frac{17 \text{ PSI}}{30 \text{ PSI}} = \frac{? \text{ mA}}{40 \text{ mA}}$$

Solving for the unknown (?) current by cross-multiplication and division yields a value of 22.67 mA. Of course, this value of 22.67 mA only tells us the length of the line segment on the number line; it does not directly tell us the current signal value. To find that, we must add the "live zero" offset of 10 mA, for a final result of 32.67 mA.

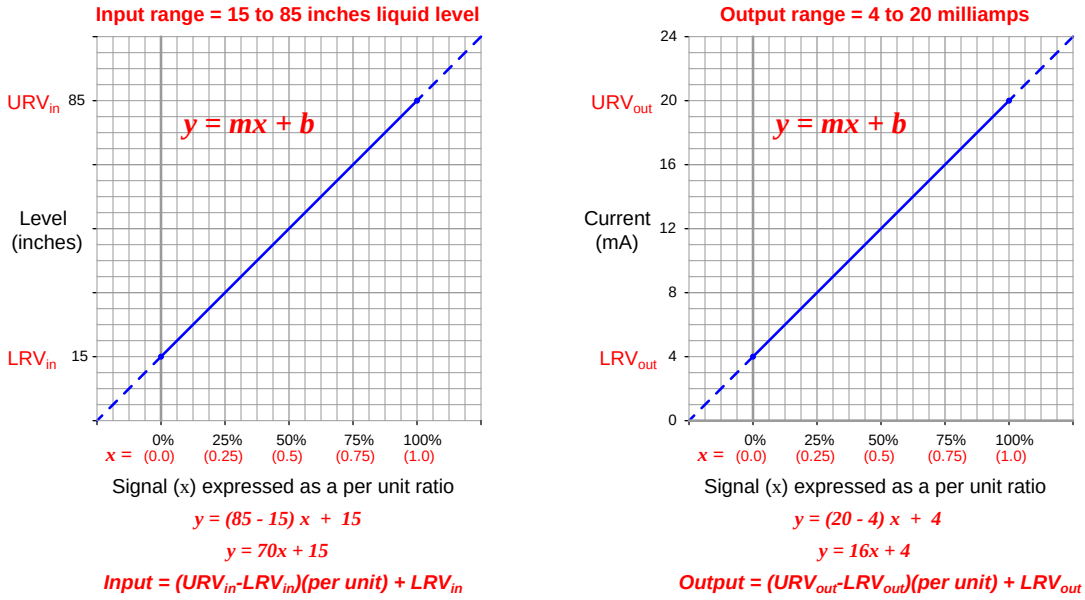


Thus, an applied pressure of +12 PSI to this transmitter should result in a 32.67 mA output signal.

2.2.8 Thinking in terms of per unit quantities

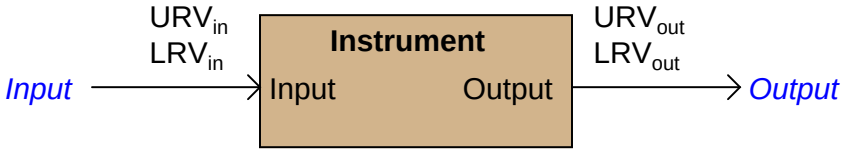
Although it is possible to generate a "custom" linear equation in the form of $y = mx + b$ for any linear-responding instrument relating input directly to output, a more general approach may be used to relate input to output values by translating all values into (and out of) *per unit* quantities. A "per unit" quantity is simply a ratio between a given quantity and its maximum value. A half-full glass of water could thus be described as having a fullness of *0.5 per unit*. The concept of percent ("per one hundred") is very similar, the only difference between per unit and percent being the base value of comparison: half-full glass of water has a fullness of 0.5 per unit (i.e. $\frac{1}{2}$ of the glass's full capacity), which is the same thing as 50 percent (i.e. 50 on a scale of 100, with 100 representing complete fullness).

Let's now apply this concept to a realistic 4-20 mA signal application. Suppose you were given a liquid level transmitter with an input measurement range of 15 to 85 inches and an output range of 4 to 20 milliAmperes, respectively, and you desired to know how many milliAmperes this transmitter should output at a measured liquid level of 32 inches. Both the measured level and the milliAmpere signal may be expressed in terms of *per unit* ratios, as shown by the following graphs:



So long as we choose to express process variable and analog signal values as a per unit ratios ranging from 0 to 1, we see how m (the slope of the line) is simply equal to the span of the process variable or analog signal range, and b is simply equal to the lower-range value (LRV) of the process variable or analog signal range. The advantage of thinking in terms of "per unit" is the ability to quickly and easily write linear equations for any given range. In fact, this is so easy that we don't even have to use a calculator to compute m in most cases, and we *never* have to calculate b because the LRV is explicitly given to us. The instrument's input equation is $y = 70x + 15$ because the span of the 15-to-85 inch range is 70, and the LRV is 15. The instrument's output equation is $y = 16x + 4$ because the span of the 4-to-20 milliAmpere range is 16, and the LRV is 4.

If we manipulate each of the $y = mx + b$ equations to solve for x (per unit of span), we may express the relationship between the input and output of any linear instrument as a pair of fractions with the per unit value serving as the proportional link between input and output:



$$\frac{\textit{Input} - \text{LRV}_{in}}{\text{URV}_{in} - \text{LRV}_{in}} = \text{Per unit of span (from 0 to 1 inclusive)} = \frac{\textit{Output} - \text{LRV}_{out}}{\text{URV}_{out} - \text{LRV}_{out}}$$

The question remains, how do we apply these equations to our example problem: calculating the milliAmpere value corresponding to a liquid level of 32 inches for this instrument? The answer to this question is that we must perform a *two-step* calculation: first, convert 32 inches into a per unit ratio, then convert that per unit ratio into a milliAmpere value.

First, the conversion of inches into a per unit ratio, knowing that 32 is the value of y and we need to solve for x:

$$32 = 70x + 15$$

$$32 - 15 = 70x$$

$$\frac{32 - 15}{70} = x$$

$$x = 0.2429 \text{ per unit} \quad \text{(i.e. } 24.29\%\text{)}$$

Next, converting this per unit ratio into a corresponding milliAmpere value, knowing that y will now be the current signal value using m and b constants appropriate for the 4-20 milliAmpere range:

$$y = 16x + 4$$

$$y = 16(0.2429) + 4$$

$$y = 3.886 + 4$$

$$y = 7.886 \text{ mA}$$

Instead of deriving a single custom $y = mx + b$ equation directly relating input (inches) to output (milliAmperes) for every instrument we encounter, we may use two simple and generic linear equations to do the calculation in two steps with "per unit" being the intermediate result. Expressed in general form, our linear equation is:

$$y = mx + b$$

$$\text{Value} = (\text{Span})(\text{Per unit}) + \text{LRV}$$

$$\text{Value} = (\text{URV} - \text{LRV})(\text{Per unit}) + \text{LRV}$$

Thus, to find the per unit ratio we simply take the value given to us, subtract the LRV of its range, and divide by the span of its range. To find the corresponding value we take this per unit ratio, multiply by the span of the other range, and then add the LRV of the other range.

Example: Given a pressure transmitter with a measurement range of 150 to 400 PSI and a signal range of 4 to 20 milliAmperes, calculate the applied pressure corresponding to a signal of 10.6 milliAmperes.

Solution: Take 10.6 milliAmperes and subtract the LRV (4 milliAmperes), then divide by the span (16 milliAmperes) to arrive at 41.25% (0.4125 per unit). Take this number and multiply by the span of the pressure range (400 PSI − 150 PSI, or 250 PSI) and lastly add the LRV of the pressure range (150 PSI) to arrive at a final answer of 253.125 PSI.

Example: Given a temperature transmitter with a measurement range of −88 degrees to +145 degrees and a signal range of 4 to 20 milliAmperes, calculate the proper signal output at an applied temperature of +41 degrees.

Solution: Take 41 degrees and subtract the LRV (−88 degrees) which is the same as *adding* 88 to 41, then divide by the span (145 degrees − (−88) degrees, or 233 degrees) to arrive at 55.36% (0.5536 per unit). Take this number and multiply by the span of the current signal range (16 milliAmperes) and lastly add the LRV of the current signal range (4 milliAmperes) to arrive at a final answer of 12.86 milliAmperes.

Example: Given a pH transmitter with a measurement range of 3 pH to 11 pH and a signal range of 4 to 20 milliAmperes, calculate the proper signal output at 9.32 pH.

Solution: Take 9.32 pH and subtract the LRV (3 pH), then divide by the span (11 pH − 3 pH, or 8 pH) to arrive at 79% (0.79 per unit). Take this number and multiply by the span of the current signal range (16 milliAmperes) and lastly add the LRV of the current signal range (4 milliAmperes) to arrive at a final answer of 16.64 milliAmperes.

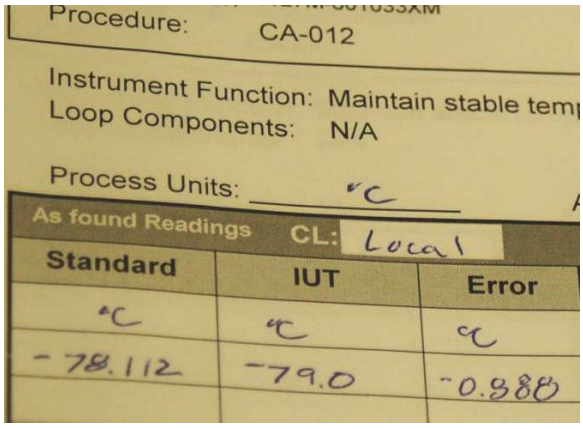
2.3 Instrument error documentation

Industrial instrument accuracy is often documented in the form of tables showing the instrument's response (i.e. output signal value) at various input (i.e. sensing) conditions. These calibration tables typically show the instrument's calibration as it was found *and* as it was left after adjustments were made. The purpose for documenting both conditions is to make data available for calculating instrument *drift* over time. If only one of these conditions is documented during each calibration event, it will be difficult to determine how well an instrument is holding its calibration over long periods of time. Excessive drift is often an indicator of impending failure, which is vital for any program of predictive maintenance or quality control.

Typically, the format for documenting both As-Found and As-Left data is a simple table showing the points of calibration, the ideal instrument responses, the actual instrument responses, and the calculated error at each point. The following table is an example for a pressure transmitter with a range of 0 to 200 PSI over a five-point scale:

Percent of range	**Input** pressure	**Output current** (ideal)	**Output current** (measured)	**Error** (percent of span)
0%	0 PSI	4.00 mA		
25%	50 PSI	8.00 mA		
50%	100 PSI	12.00 mA		
75%	150 PSI	16.00 mA		
100%	200 PSI	20.00 mA		

The following photograph shows a single-point "As-Found" calibration report on a temperature indicating controller, showing the temperature of the calibration standard (−78.112 degrees Celsius), the display of the instrument under test (IUT, −79 degrees Celsius), and the error between the two (−0.888 degrees Celsius):



Note that the mathematical sign of the error is important. An instrument that registers −79 degrees when it should register −78.112 degrees exhibits a *negative error*, since its response is lower (i.e. more negative) than it should be. Expressed mathematically: Error = IUT − Standard. When the error must be expressed in percentage of span, the formula becomes:

$$\text{Error} = \frac{\text{IUT} - \text{Standard}}{\text{Span}} \times 100\%$$

It is not uncommon for calibration tables to show multiple calibration points going *up* as well as going *down*, for the purpose of documenting hysteresis and deadband errors. Note the following example, showing a transmitter with a maximum hysteresis of 0.313 % (the offending data points are shown in bold-faced type):

Percent of range	**Input** pressure	**Output current** (ideal)	**Output current** (measured)	**Error** (percent of span)
0%	0 PSI	4.00 mA	3.99 mA	−0.0625 %
25% ↑	50 PSI	8.00 mA	**7.98 mA**	**−0.125 %**
50% ↑	100 PSI	12.00 mA	11.99 mA	−0.0625 %
75% ↑	150 PSI	16.00 mA	15.99 mA	−0.0625 %
100% ↑	200 PSI	20.00 mA	20.00 mA	0 %
75% ↓	150 PSI	16.00 mA	16.01 mA	+0.0625 %
50% ↓	100 PSI	12.00 mA	12.02 mA	+0.125 %
25% ↓	50 PSI	8.00 mA	**8.03 mA**	**+0.188 %**
0% ↓	0 PSI	4.00 mA	4.01 mA	+0.0625 %

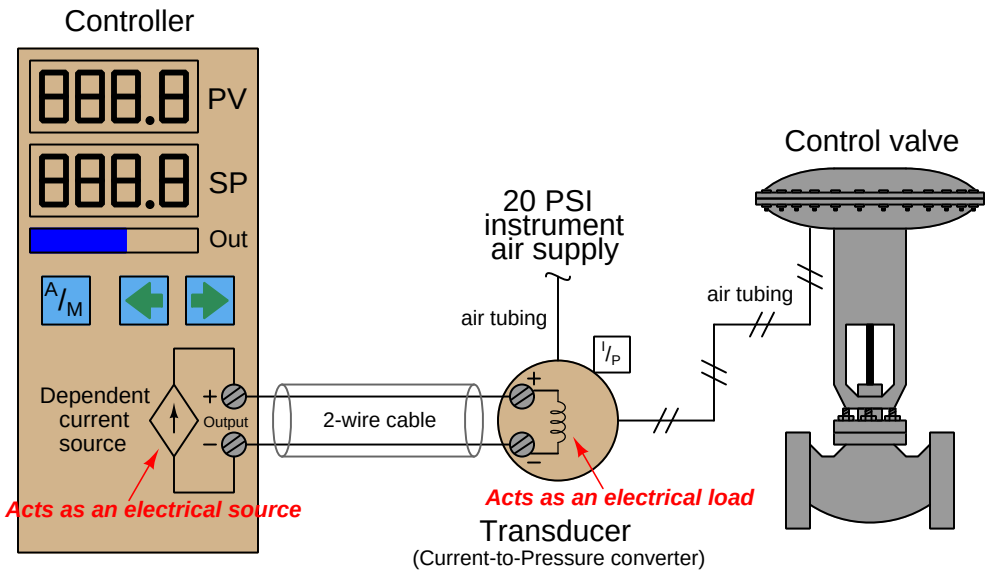
Note again how error is expressed as either a positive or a negative quantity depending on whether the instrument's measured response is above or below what it should be under each condition. The values of error appearing in this calibration table, expressed in percent of span, are all calculated by the following formula:

$$\text{Error} = \left(\frac{I_{measured} - I_{ideal}}{16 \text{ mA}} \right) (100\%)$$

In the course of performing such a directional calibration test, it is important not to overshoot any of the test points. If you do happen to overshoot a test point in setting up one of the input conditions for the instrument, simply "back up" the test stimulus and re-approach the test point from the same direction as before. Unless each test point's value is approached from the proper direction, the data cannot be used to determine hysteresis/deadband error.

2.4 Controller output current loops

The simplest form of 4-20 mA signaling circuit (commonly called a *current loop*) is the type used to represent the output of a process controller, sending a command signal to a final control element. Here, the controller supplies both the electrical power and signal information to the final control element, which acts as an electrical load. To illustrate, consider the example of a controller sending a 4-20 mA signal to an I/P (current-to-pressure) signal converter, which then pneumatically drives a control valve:



This particular controller has two digital displays, one for process variable (PV) and one for setpoint (SP), with a bargraph for displaying the output value (Out). One pushbutton provides the operator with a way to switch between Automatic and Manual modes (A/M), while two other pushbuttons provide means to decrement and increment either the setpoint value (in Automatic mode) or the Output value (in Manual mode).

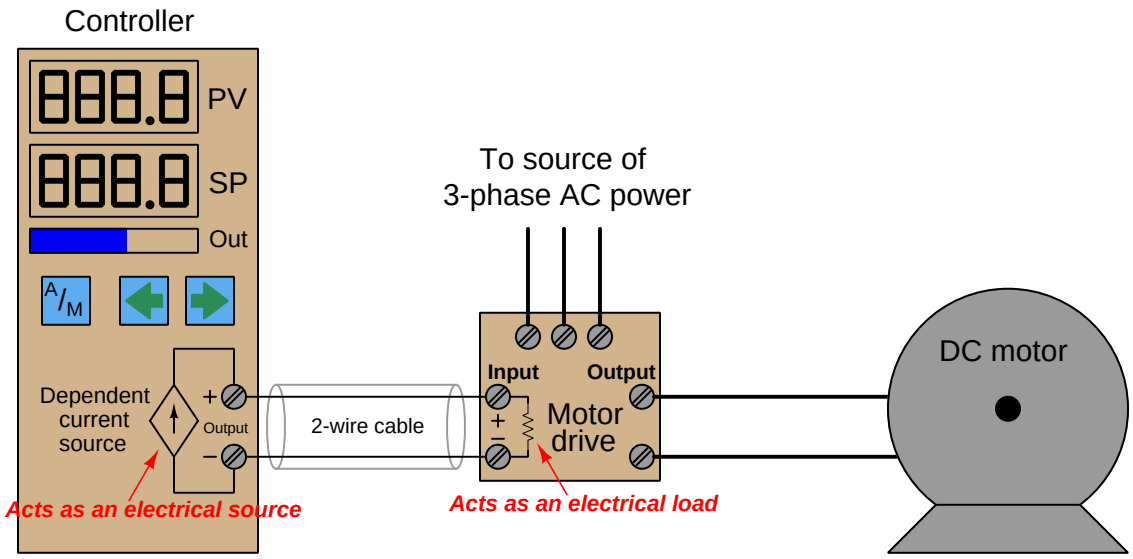
Inside the controller, a *dependent current source* provides the 4-20 mA DC current signal to the I/P transducer. Like all current sources, its purpose is to maintain current in the "loop" circuit regardless of circuit resistance or any external voltage sources. Unlike a constant current source, a "dependent" current source (represented by a diamond shape instead of a circle shape) varies its current value according to some external stimulus. In this case, either the mathematical function of the controller (Automatic mode) or the arbitrary action of the human operator (Manual mode) tells the current source how much DC current it should maintain in the circuit.

For example, if the operator happened to switch the controller into Manual mode and set the output value at 50%, the proper amount of DC current for this signal percentage would be 12 mA (exactly half-way between 4 mA and 20 mA). If everything is working properly, the current in the "loop" circuit to the I/P transducer should remain exactly at 12 mA regardless of slight changes in wire resistance, I/P coil resistance, or anything else: the current source inside the controller will

"fight" as hard as it has to in order to maintain this set amount of current. This current, as it flows through the wire coil of the I/P transducer mechanism, creates a magnetic field inside the I/P to actuate the pneumatic mechanism and produce a 9 PSI pressure signal output to the control valve (9 PSI being exactly half-way between 3 PSI and 15 PSI in the 3-15 PSI signal standard range). This should move the control valve to the half-way position.

The details of the controller's internal current source are not terribly important. Usually, it takes the form of an operational amplifier circuit driven by the voltage output of a DAC (Digital-to-Analog Converter). The DAC converts a binary number (either from the controller's automatic calculations, or from the human operator's manual setting) into a small DC voltage, which then commands the opamp circuit to regulate output current at a proportional value.

The scenario is much the same if we replace the I/P and control valve with a variable-speed motor drive. From the controller's perspective, the only difference it sees is a resistive load instead of an inductive load. The input resistance of the motor drive circuit converts the 4-20 mA signal into an analog voltage signal (typically 1-5 V, but not always). This voltage signal then commands the motor drive circuitry, telling it to modulate the power going to the electric motor in order to drive it at the desired speed:

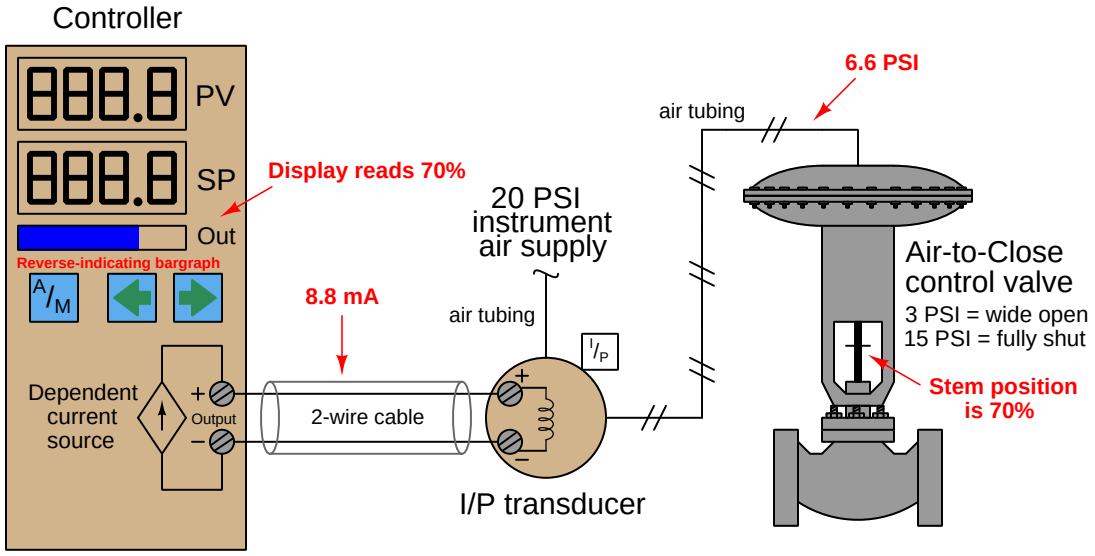


Here, the variable-speed motor drive is a high-power electronic circuit that takes 3-phase AC power in and converts it to DC power at a variable voltage (that voltage level controlled by the 4-20 mA signal from the controller output). Another style of motor "drive" is one that takes in AC power and outputs 3-phase AC power at variable voltage and frequency to power an AC induction motor. This latter style is usually called a *variable frequency drive* or *VFD*, but it looks the same to the controller output: a fixed resistive load accepting the 4-20 mA output signal.

In some process applications the final control element is "reverse-acting" in that a controller output current value of 4 mA tells the FCE to go to its "full" (100%) capacity and a controller output current value of 20 mA commands the FCE to go to its minimum (0%) capacity. Such is the case with an I/P transducer driving an "air-to-close" pneumatic control valve, where the valve's spring works to push the throttling mechanism open and air pressure on the valve diaphragm works to push the throttling mechanism shut. A practical reason for designing a control system like this is if the dictates of process safety require the valve go wide-open if ever the instrument air supply fails or the 4-20 mA output signal circuit fails.

In such applications, we need to configure the controller in such a way that the output display (digital read-out and/or bargraph) is *reverse-indicating* so as to avoid confusing any human operator using the controller. Since 4 mA represents a wide-open control valve and 20 mA represents a fully shut control valve, a reverse-indicating controller will display 0% output when the current signal is 20 mA and 100% output when the current signal is 4 mA (i.e. the controller display is a direct representation of *control valve stem position*, not of the current signal itself):

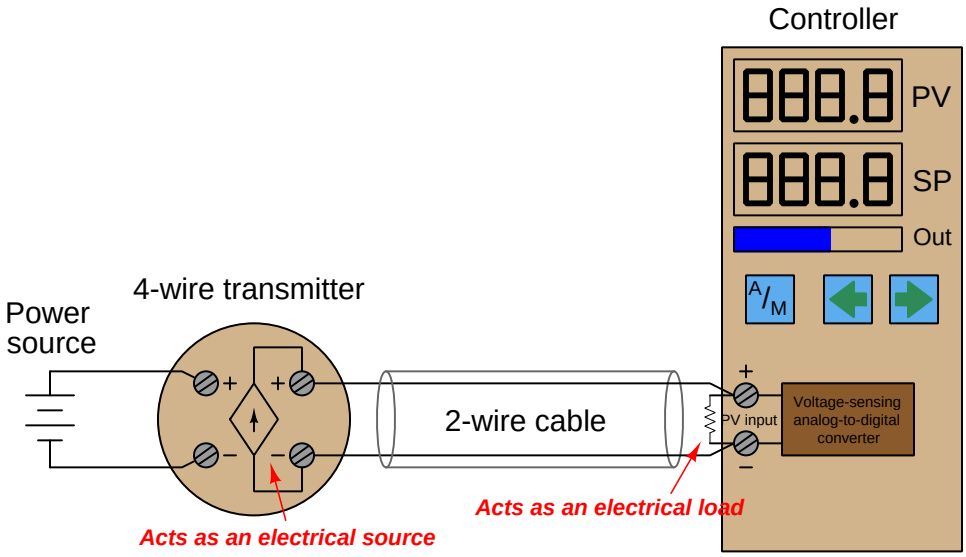
Air-to-Close valve with controller configured for "reverse indication" on its output display



It should be noted that direct or reverse *indication* on the controller's faceplate is unrelated to direct or reverse *control action* of the controller's algorithm. The former is merely the way in which the output signal is graphically represented to the human operator while the latter is the relationship between the process variable (PV) signal and the controller's output (MV) signal necessary for negative-feedback control. For the sake of argument, this controller's automatic action could still be *direct* (i.e. greater PV signal = greater output signal) even though its output bargraph indication is *reverse* in order to faithfully show valve stem position to the human operator.

2.5 4-wire ("self-powered") transmitter current loops

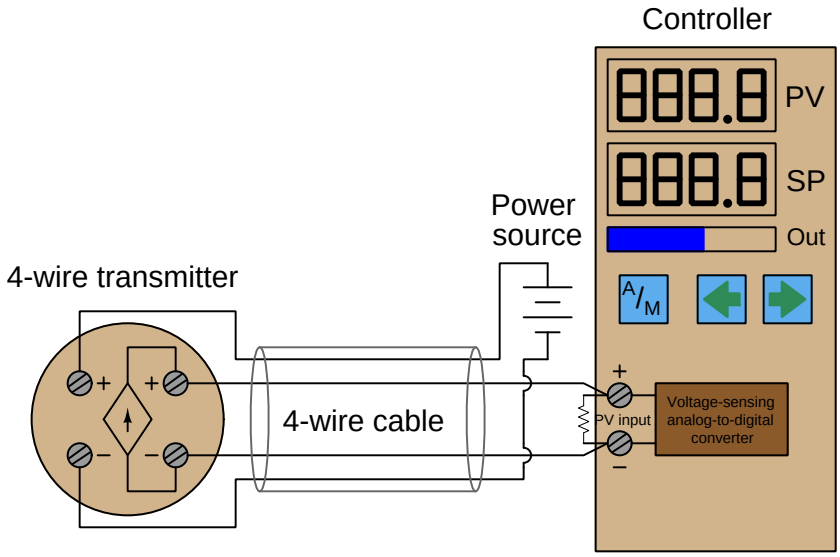
DC electric current signals may also be used to communicate process measurement information from transmitters to controllers, indicators, recorders, alarms, and other input devices. Recall that the purpose of a *transmitter* is to sense some physical variable (e.g. pressure, temperature, flow) and then report that quantity in the form of a signal, in this case a 4 to 20 milliAmpere DC current proportional to that measured quantity. The simplest form of 4-20 mA measurement loop is one where the transmitter has two terminals for the 4-20 mA signal wires to connect, and two more terminals where a power source connects. These transmitters are called "4-wire" or "self-powered" units. The current signal from the transmitter connects to the *process variable input* terminals of the controller to complete the loop:



Some process controllers are not equipped to directly accept milliAmpere input signals, but rather can only interpret DC voltage signals. In such cases we must connect a precision resistor across the controller's input terminals to convert the 4-20 mA transmitter signal into a standardized analog voltage signal the controller can understand. A voltage signal range of 1 to 5 Volts is standard, although some models of controller use different voltage ranges and therefore require different precision resistor values. If the voltage range is 1-5 Volts and the current range is 4-20 mA, the precision resistor value must be 250 Ohms according to Ohm's Law.

Since this is a digital controller, the input voltage at the controller terminals is interpreted by an analog-to-digital converter (ADC) circuit, which converts the measured voltage into a digital number the controller's microprocessor can interpret.

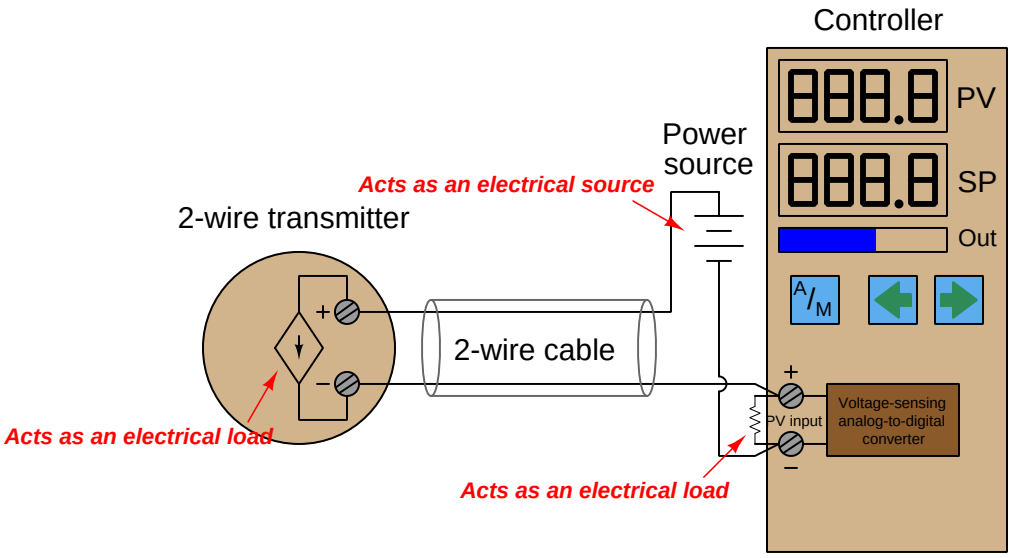
In some installations, transmitter power is supplied through additional wires in the cable from a power source located near the controller:



The obvious disadvantage of this scheme is the requirement of two more conductors in the cable. More conductors means the cable will be larger-diameter and more expensive for a given length. Cables with more conductors will require larger electrical conduit to fit in to, and all field wiring panels will have to contain more terminal blocks to marshal the additional conductors. If no suitable electrical power source exists at the transmitter location, though, a 4-wire cable is necessary to service a 4-wire transmitter.

2.6 2-wire ("loop-powered") transmitter current loops

It is possible to convey electrical power *and* communicate analog information over the same two wires using 4 to 20 milliAmperes DC, if we design the transmitter to be *loop-powered*. A loop-powered transmitter connects to a process controller with only two wires, which is why loop-powered transmitters are synonymously known as *2-wire transmitters*:

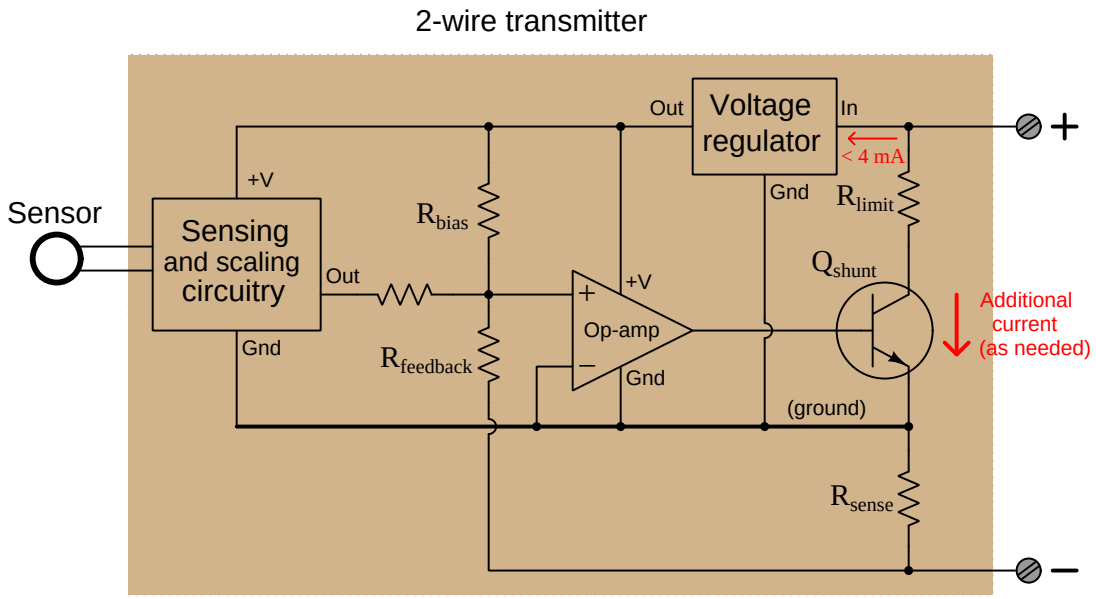


Here, the transmitter is not really a current *source* in the sense that a 4-wire transmitter is. Instead, a 2-wire transmitter's circuitry is designed to act as a current *regulator*, limiting current in the series loop to a value representing the process measurement, while relying on a remote source of power to motivate the electric current. Please note the direction of the arrow in the transmitter's dependent current source symbol, and how it relates to the voltage polarity marks. Refer back to the illustration of a 4-wire transmitter circuit for comparison. The current "source" in this loop-powered transmitter actually behaves as an electrical *load*[6], while the current source in the 4-wire transmitter functioned as a true electrical source.

A loop-powered transmitter gets its operating power from the minimum terminal voltage and current available at its two terminals. With the typical source voltage being 24 Volts DC, and the maximum voltage dropped across the controller's 250 Ohm resistor being 5 Volts DC, the transmitter should always have at least 19 Volts available at its terminals. Given the lower end of the 4-20 mA signal range, the transmitter should always have at least 4 mA of current to function on. Thus, the transmitter will always have a certain minimum amount of electrical power available on which to operate, while regulating current to signal the process measurement to the receiving instrument.

[6] Some of my students have referred to such a circuit as a *smart load*, since it functions as a load but nevertheless exerts control over the circuit current.

Internally, the electronic hardware of a 2-wire transmitter circuitry resembles the following (simplified) diagram. Note that everything shown within the shaded rectangle is represented by the "2-wire transmitter" circle in the previous diagram:



All sensing, scaling, and output conditioning circuitry inside the transmitter must be designed to operate on less than 4 mA of DC current, and at a modest terminal voltage. In order to create loop currents exceeding 4 mA – as the transmitter must do in order to span the entire 4 to 20 milliAmpere signal range – the transmitter circuitry uses a transistor to shunt (bypass) extra current from one terminal to the other as needed to make the total current indicative of the process measurement. For example, if the transmitter's internal operating current is only 3.8 mA, and it must regulate loop current at a value of 16 mA to represent a condition of 75% process measurement, the shunt transistor will be driven by the opamp to bypass exactly 12.2 mA of current (because 3.8 mA + 12.2 mA = 16.0 mA).

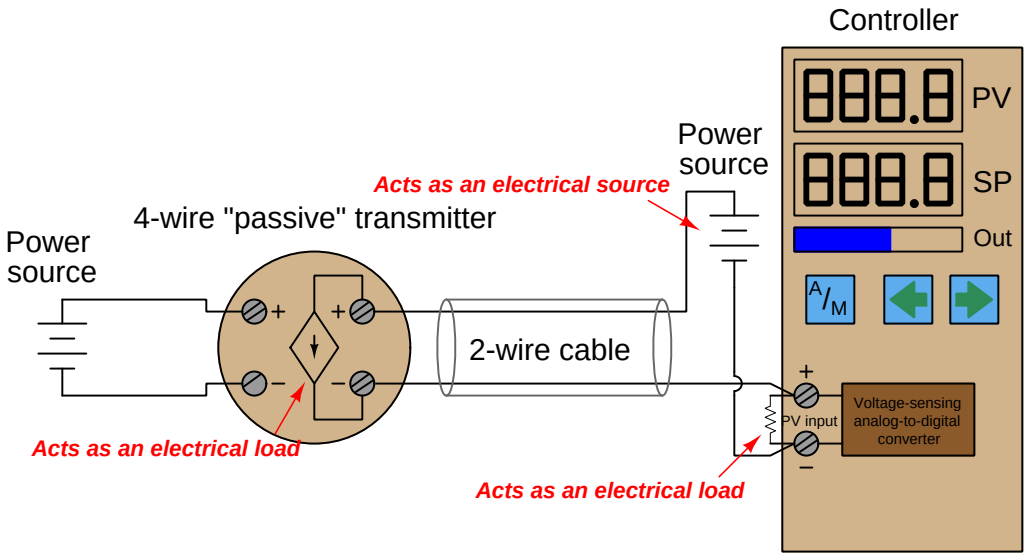
The very low amount of electrical power available at a 2-wire transmitter's terminals limits its functionality. If the transmitter requires more electrical power than can be delivered with 4 milliAmperes and 19 Volts (minimum each), the only solution is to go with a 4-wire transmitter where the power conductors are separate from the signal conductors. An example of a process transmitter that must be 4-wire is a chemical analyzer such as a gas chromatograph, requiring enough power to operate an electrical heater, solenoid valves, and an on-board computer to process the sensor data. There is simply no way to operate a machine as complex and power-hungry as a legacy gas chromatograph on 4 milliAmperes and 19 Volts!

Early current-based industrial transmitters were not capable of operating on such low levels of electrical power, and so used a different current signal standard: 10 to 50 milliAmperes DC. Loop power supplies for these transmitters ranged upwards of 90 Volts to provide enough power for the transmitter. Safety concerns made the 10-50 mA standard unsuitable for some industrial

installations, and modern microelectronic circuitry with its reduced power consumption made the 4-20 mA standard practical for nearly all types of process transmitters.

2.7 4-wire "passive" versus "active" output transmitters

Some self-powered (4-wire) analog electronic transmitters are designed to behave as electrical loads rather than as electrical sources. Such transmitters are commonly referred to as having *passive* or *sinking* 4-20 mA outputs, as opposed to the *active* or *sourcing* 4-wire transmitters previously described:

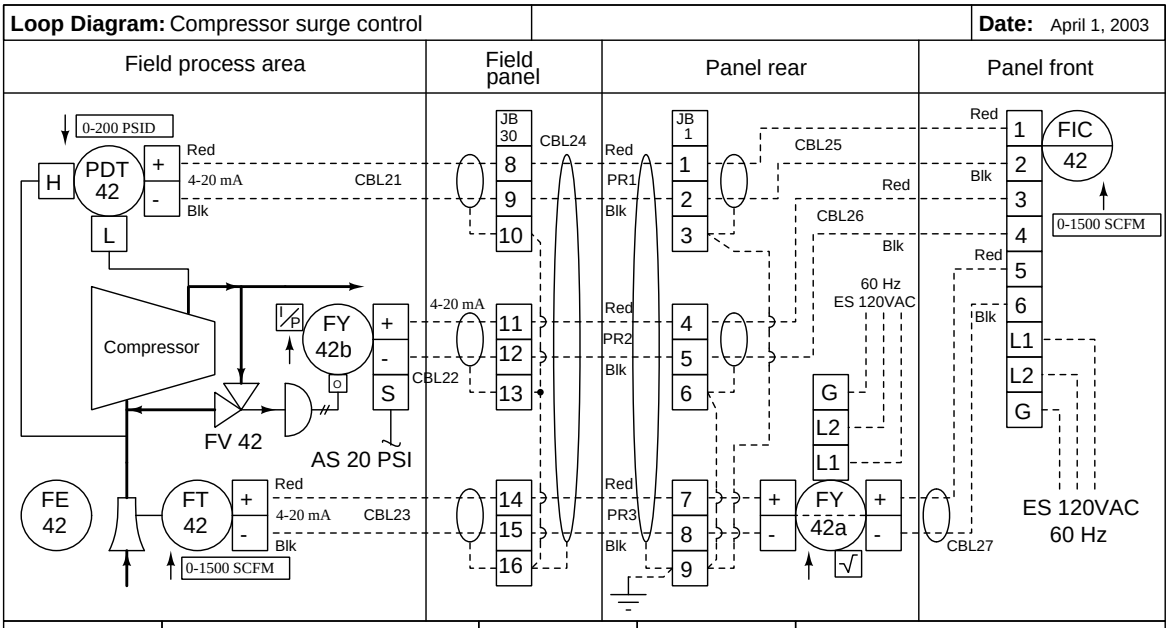


At first this seems needlessly confusing. Why build a self-powered transmitter requiring a second power supply in the circuit to drive the 4-20 mA signal? The reason for this type of transmitter's existence is the sheer popularity of loop-powered 2-wire 4-20 mA transmitters. Loop-powered field instruments have become so popular in industry that many control systems, PLCs, indicators, and other receiving devices have their own loop power supply built in so that these systems can *only* connect to loads and are therefore incompatible with current-sourcing 4-wire transmitters. Thus, "passive" or "sinking" 4-wire transmitters were developed to accommodate control systems designed to work (only) with loop-powered instruments.

Some 4-wire transmitters are configurable for either passive or active (i.e. sinking versus sourcing) operation, requiring the installing technician to pay close attention in order for the circuit to properly function.

2.8 Instrument loop diagrams

A style of electrical diagram unique to the industrial control field for documenting 4-20 mA signal circuits is called a *loop diagram* or *loop sheet*, following a standard codified by the ISA (Instrumentation, Systems, and Automation Society) as standard 5.1. Similar to electrical or electronic schematic diagrams, loop diagrams use dashed lines to represent electrical conductors, squares to represent terminal blocks, and incorporate ISA-standardized symbols for sensing and control devices as well as (occasionally) symbols for process piping and vessels too. An example of a loop diagram for a compressor anti-surge control system is shown here:



Tag number	Description	Input cal.	Output cal.	Notes
FE 42	Venturi tube	0-1500 SCFM	0-100 "WC	
FT 42	Suction flow transmitter	0-100 "WC	4-20 mA	
FY 42a	Square root extractor	4-20 mA	4-20 mA	
FY 42b	Current-to-pressure converter	4-20 mA	3-15 PSI	
FV 42	Anti-surge control valve	3-15 PSI	100%-0%	Air-to-close
PDT 42	Differential pressure transmitter	0-200 PSI	20-4 mA	Reverse action
FIC 42	Anti-surge controller	4-20 mA	4-20 mA	

Note how every instrument bears the "loop number" of 42, to distinguish it from other control systems within the same process or facility. It must be understood that the term "loop" is confusingly used for different things in instrumentation: it may refer to a signal current circuit (i.e. a *current loop*) or it may refer to a negative feedback loop (i.e. a *control loop*) as is the case here with loop #42. This particular control loop contains two transmitters, a controller, a valve, and two signal transducers. Transducer 42a modifies the flow transmitter's signal before it goes into the controller,

and transducer 42b converts the electronic 4 to 20 mA signal into a pneumatic 3 to 15 PSI air pressure signal. Each instrument "bubble" in a loop diagram represents an individual device, with its own terminals for connecting wires.

Dashed lines represent individual copper wires inside of multi-conductor cables. Electrical terminals where these wires connect to are represented by squares with numbers in them. Fluid ports on instruments are also represented by labeled squares. Cable numbers, wire colors, junction block numbers, panel identification, and even grounding points are all shown in loop diagrams.

The fine level of detail contained in a loop diagram such as wire colors and terminal numbers is highly useful to technician who must maintain and upgrade these critical control systems. The more detail contained in a loop diagram, the easier it the inevitable job of maintaining that system at some later date. When a loop diagram shows you exactly what wire color to expect at exactly what point in a loop, and exactly what terminal that wire should connect to, it becomes much easier to proceed with any troubleshooting, calibration, or upgrade task.

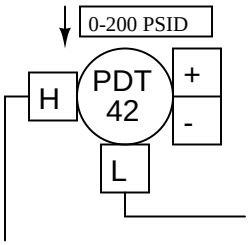
Loop diagrams are fairly constrained in their layout as per the ISA 5.1 standard. Instruments located near the process equipment ("in the field") are always placed on the left-hand side, while control-panel or control-room instruments must be located on the right-hand side. Text describing instrument tags, ranges, and notes are always placed on the bottom. Unlike electrical schematics and piping diagrams where component layout is largely left to the whim of whomever draws the diagram, loop sheets offer little room for creativity. This is intentional, as creativity and readability are mutually exclusive in cases where there is an immense amount of technical detail embedded in a diagram. It is simply easier to find details you're looking for when you know *exactly* where they ought to be.

An interesting detail seen on this loop diagram is an entry specifying "input calibration" and "output calibration" for each and every instrument in the system. This is actually a very important concept to keep in mind when troubleshooting a complex instrumentation system: every instrument has at least one input and at least one output, with some sort of mathematical relationship between the two. Diagnosing where a problem lies within a measurement or control system often means testing various instruments to see if their output responses appropriately match their input conditions, so it is important to document these input and output ranges.

For example, one way to test the flow transmitter in this system would be to subject it to a number of different pressures within its range (specified in the diagram as 0 to 100 inches of water column differential) and seeing whether or not the current signal output by the transmitter was consistently proportional to the applied pressure (e.g. 4 mA at 0 inches pressure, 20 mA at 100 inches pressure, 12 mA at 50 inches pressure, etc.).

Given the fact that a calibration error or malfunction in any one of these instruments can cause a problem for the control system as a whole, it is nice to know there is a way to determine which instrument is to blame and which instruments are not. This general principle holds true regardless of the instrument's type or technology. You can use the same input-versus-output test procedure to verify the proper operation of a pneumatic (3 to 15 PSI) level transmitter or an analog electronic (4 to 20 mA) flow transmitter or a digital (fieldbus) temperature transmitter alike. Each and every instrument has an input and an output, and there is always a predictable (and testable) correlation from one to the other.

Another interesting detail seen on this loop diagram is the *direction of action* of each instrument. You will notice a box and arrow (pointing either up or down) next to each instrument bubble. An "up" arrow (↑) represents a *direct-acting* instrument: one whose output signal increases as the input stimulus increases. A "down" arrow (↓) represents a *reverse-acting* instrument: one whose output signal decreases as the input stimulus increases. All the instruments in this loop are direct-acting with the exception of the pressure differential transmitter PDT-42:



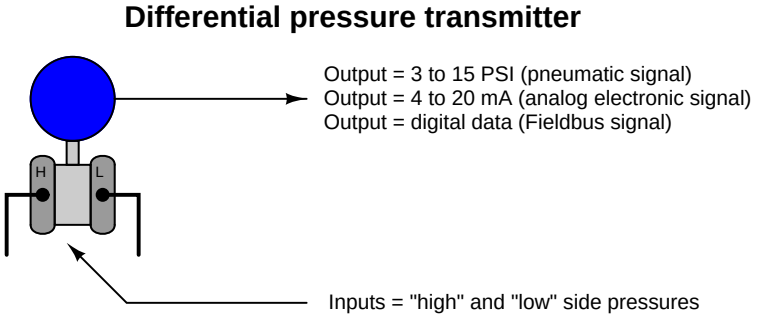
Here, the "down" arrow tells us the transmitter will output a full-range signal (20 mA) when it senses zero differential pressure, and a 0% signal (4 mA) when sensing a full 200 PSI differential. While this calibration may seem confusing and unwarranted, it serves a definite purpose in this particular control system. Since the transmitter's current signal decreases as pressure increases, and the controller must be correspondingly configured, a decreasing current signal will be interpreted by the controller as a high differential pressure. If any wire connection fails in the 4-20 mA current loop for that transmitter, the resulting 0 mA signal will be naturally "seen" by the controller as a pressure over-range condition. Excessive pressure drop across a compressor is considered dangerous because it may lead to the compressor surging, thus the controller will naturally take action to prevent surge by commanding the anti-surge control valve to open, because it "thinks" the compressor is about to surge. In other words, the transmitter is intentionally calibrated to be reverse-acting such that any break in the signal wiring will naturally bring the system to its safest condition.

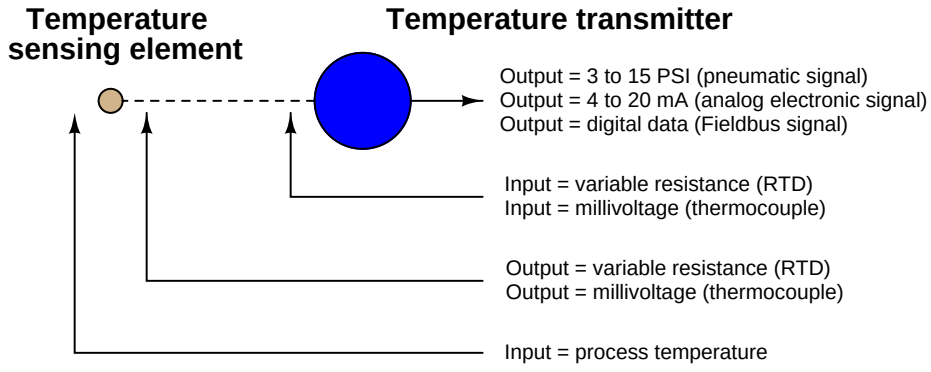
2.9 Troubleshooting current loops

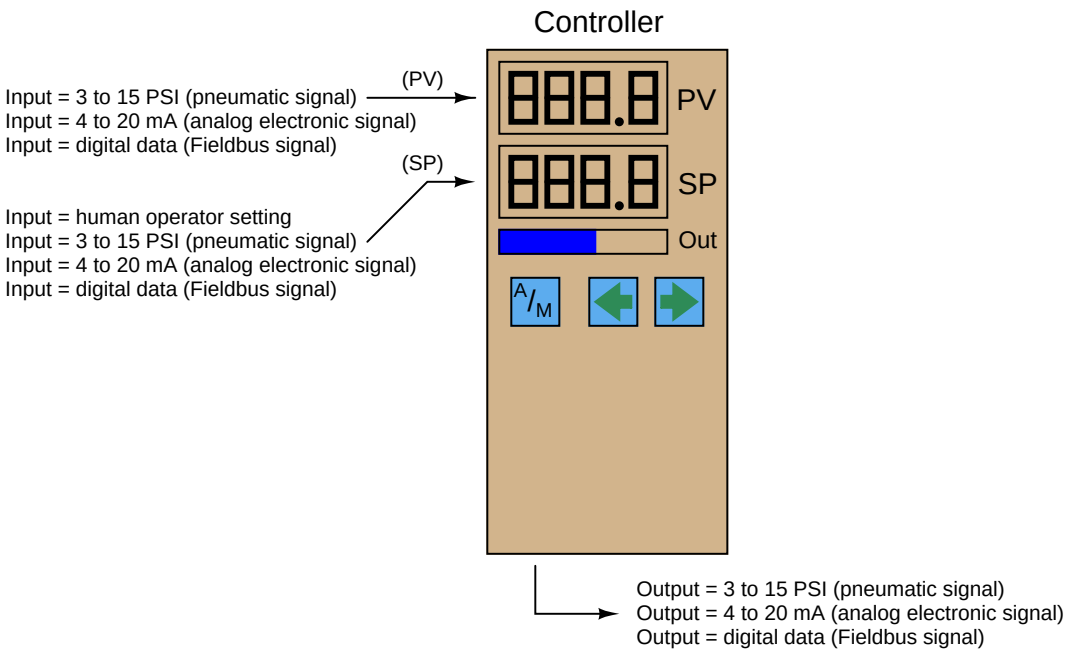
A fundamental principle in instrumentation system troubleshooting is that every instrument has at least one input and at least one output, and that the output(s) should accurately correspond to the input(s). If an instrument's output is not properly corresponding to its input according to the instrument's design function, there must be something wrong with that instrument.

Consider the inputs and outputs of several common instruments: transmitters, controllers, indicators, and control valves. Each of these instruments takes in (input) data in some form, and generates (output) data in some form. In any instrument "loop," the output of one instrument feeds into the input of the next, such that information passes from one instrument to another. By intercepting the data communicated between components of an instrument system, we are able to locate and isolate faults. In order to properly understand the intercepted data, we must understand the inputs and outputs of the respective instruments and the basic functions of those instruments.

The following illustrations highlight inputs and outputs for instruments commonly found in control systems:







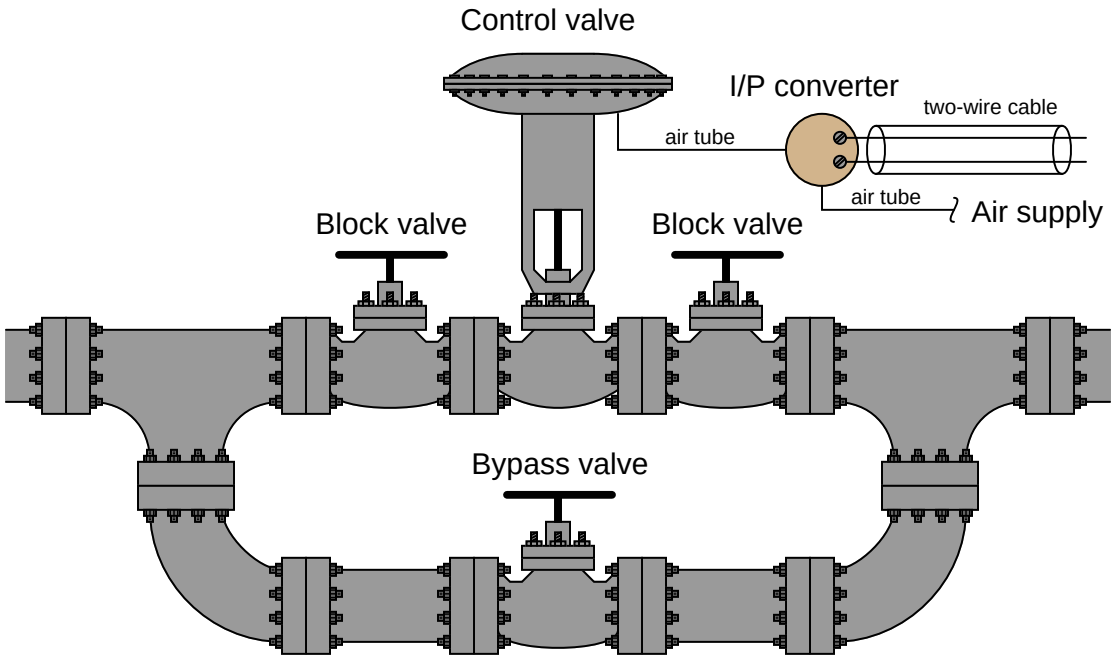
In order to check for proper correspondence between instrument inputs and outputs, we must be able to use appropriate test equipment to intercept the signals going into and out of those instruments. For 4-20 mA analog signal-based instruments, this means we must be able to use electrical meters capable of accurately measuring current and voltage.

2.9.1 Using a standard milliammeter to measure loop current

Since the signal of interest is represented by an electric current in an instrumentation current "loop" circuit, the obvious tool to use for troubleshooting is a multimeter capable of accurately measuring DC milliAmperes. Unfortunately, though, there is a major disadvantage to the use of a milliammeter: the circuit must be "broken" at some point to connect the meter in series with the current, and this means the current will fall to 0 mA until the meter is connected (then fall to 0 mA when the meter is removed from the circuit). Interrupting the current means interrupting the flow of information conveyed by that current, be it a process measurement or a command signal to a final control element. This *will* have adverse effects on a control system unless certain preparatory steps are taken.

Before "breaking the loop" to connect your meter, one must first warn all appropriate personnel that the signal will be interrupted at least twice, falling to a value of -25% each time. If the signal to be interrupted is coming from a process transmitter to a controller, the controller should be placed in Manual mode so it will not cause an upset in the process (by moving the final control element in response to the sudden loss of PV signal). Also, process alarms should be temporarily disabled so they do not cause panic. If this current signal also drives process shutdown alarms, these should be temporarily disabled so that nothing shuts down upon interruption of the signal.

If the current signal to be interrupted is a command signal from a controller to a final control element, the final control element either needs to be manually overridden so as to hold a fixed setting while the signal varies, or it needs to be bypasses completely by some other device(s). If the final control element is a control valve, this typically takes the form of opening a bypass valve and closing at least one block valve:



Since the manually-operated bypass valve now performs the job the automatic control valve used to do, a human operator must remain posted at the bypass valve to carefully throttle it and maintain control of the process.

Block and bypass valves for a large gas flow control valve may be seen in the following photograph:

In consideration of the labor necessary to safely interrupt the current signal to a control valve in a live process, we see that the seemingly simple task of connecting a milliammeter in series with a 4-20 mA current signal is not as easy as it may first appear. Better ways must exist, no?

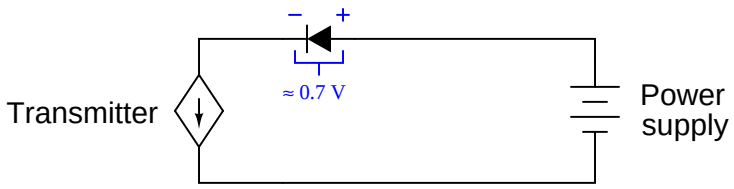
2.9.2 Using a clamp-on milliammeter to measure loop current

One better way to measure a 4-20 mA signal without interrupting it is to do so magnetically, using a clamp-on milliammeter. Modern Hall-effect sensors are sensitive and accurate enough to monitor the weak magnetic fields created by the passage of small DC currents in wires. Ammeters using Hall-effect sensors have are completely non-intrusive because they merely clamp around the wire, with no need to "break" the circuit. An example of a such a clamp-on current meter is the Fluke model 771, shown in this photograph:

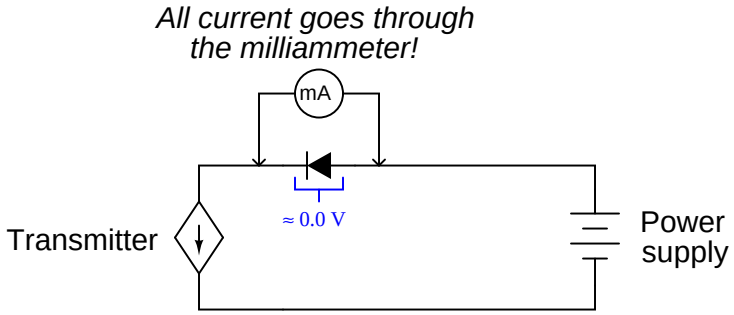
Note how this milliammeter not only registers loop current (3.98 mA as shown in the photograph), but it also converts the milliAmpere value into a percentage of range, following the 4 to 20 mA signal standard. One disadvantage to be aware of for clamp-on milliammeters is the susceptibility to error from strong external magnetic fields. Steady magnetic fields (from permanent magnets or DC-powered electromagnets) may be compensated for by performing a "zero" adjustment with the instrument held in a similar orientation prior to measuring loop current through a wire.

2.9.3 Using "test" diodes to measure loop current

Another way to measure a 4-20 mA signal without interrupting it involves the use of a rectifying diode, originally installed in the loop circuit when it was commissioned. A "test" diode may be placed anywhere in series within the loop in such a way that it will be forward-biased. During normal operation, the diode will drop approximately 0.7 Volts, as is typical for any silicon rectifying diode when forward biased. The following schematic diagram shows such a diode installed in a 2-wire transmitter loop circuit:



If someone connects a milliammeter in parallel with this diode, however, the very low input resistance of the ammeters "shorts past" the diode and prevents any substantial voltage drop from forming across it. Without the necessary forward voltage drop, the diode effectively turns off and conducts 0 mA, leaving the entire loop current to pass through the ammeter:



When the milliammeter is disconnected, the requisite 0.7 Volt drop appears to turn on the diode, and all loop current flows through the diode again. At no time is the loop current ever interrupted, which means a technician may take current measurements this way and never have to worry about generating false process variable indications, setting off alarms, or upsetting the process.

Such a diode may be installed at the nearest junction box, between terminals on a terminal strip, or even incorporated into the transmitter itself. Some process transmitters have an extra pair of terminals labeled "Test" for this exact purpose. A diode is already installed in the transmitter, and these "test" terminals serve as points to connect the milliammeter across.

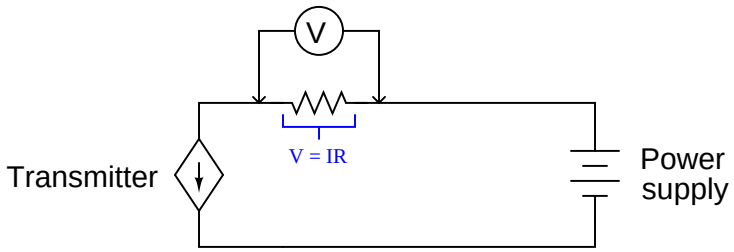
The following photograph shows an example of this on a Rosemount model 3051 differential pressure transmitter:

Note the two test points labeled "TEST" below and to the right of the main screw terminals where the loop wiring attaches. Connecting an ammeter to these two test points allows for direct measurement of the 4-20 mA current signal without having to un-do any wire connections in the circuit.

Transmitters equipped with analog meter movements for direct visual indication of the 4-20 mA signal usually connect the analog milliammeter in parallel with just such a diode. The reason for doing this is to maintain loop continuity in the event the fine-wire coil inside the milliammeter movement were to accidently break open.

2.9.4 Using shunt resistors to measure loop current

A similar method for non-invasively measuring current in a 4-20 mA instrumentation circuit is to install a precision resistor in series. If the resistance value is precisely known, the technician merely needs to measure voltage across it with a voltmeter and use Ohm's Law to calculate current:



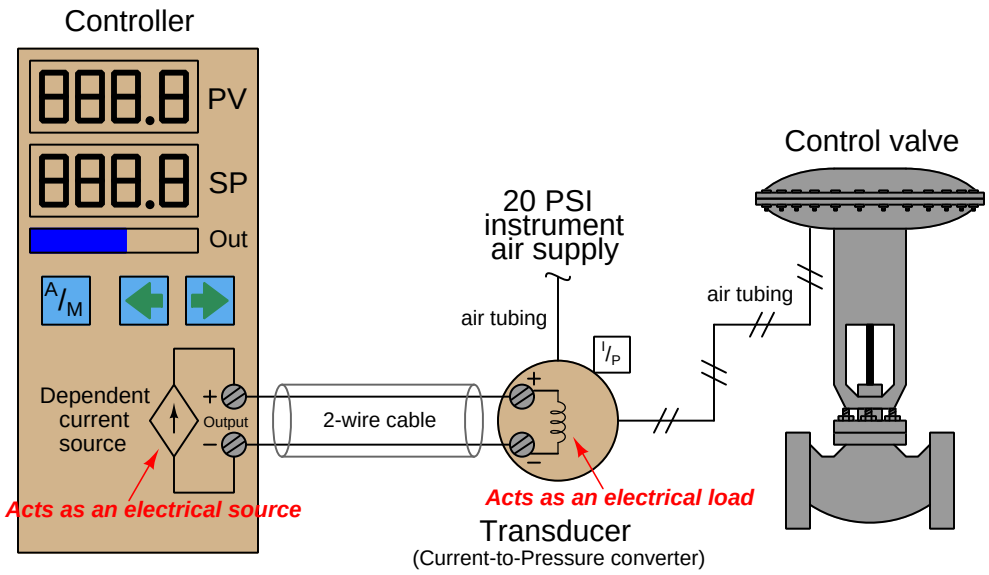
In electronics, such a precision resistor used for measuring current is often referred to as a *shunt* resistor. Shunt resistor values are commonly very small, for their purpose is to assist in current measurement without imposing undue voltage drop within a circuit. It is rare to find a 250 Ohm resistor used strictly as a diagnostic shunt resistor, because the extra voltage drop (1 to 5 Volts, depending on the current signal level) may "starve" loop-powered instruments of voltage necessary to operate. Shunt resistor values as low as 1 Ohm may be found installed in 4-20 mA current loops at strategic locations where technicians may need to measure loop current[7].

[7] Of course, a 1 Ohm resistor would drop 4 mV at 4 mA loop current, and drop 20 mV at 20 mA loop current. These small voltage values necessitate a highly accurate DC voltmeter for field measurement!

2.9.5 Troubleshooting current loops with voltage measurements

If neither component (diode nor shunt resistor) is pre-installed in the circuit, and if a Hall-effect (clamp-on) precision milliammeter is unavailable, a technician may still perform useful troubleshooting measurements using nothing but a DC voltmeter. Here, however, one must be careful of how to interpret these voltage measurements, for they may not directly correspond to the loop current as was the case with measurements taken in parallel with the precision resistor.

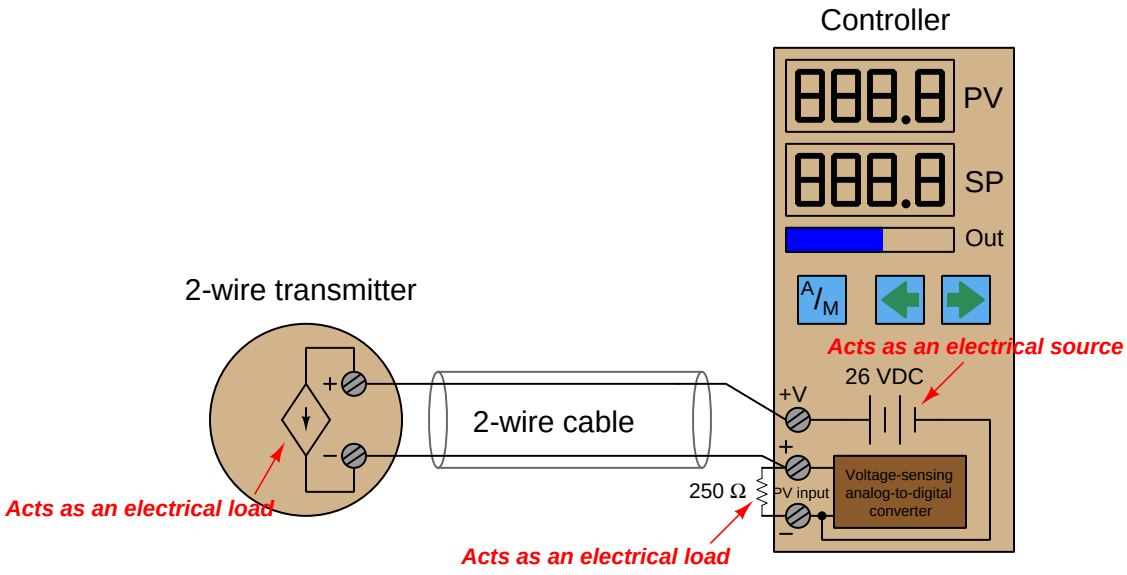
Take for example this 4-20 mA loop where a controller sends a command signal to an I/P transducer:



There is no standardized resistance value for I/P transducer coils, and so the amount of voltage dropped across the I/P terminals for any given amount of loop current will be unique for every different model of I/P. The Fisher model 567 I/P transducer built for 4-20 mA signals has a normal coil resistance of 176 Ohms. Thus, we would expect to see a voltage drop of approximately 0.7 Volts at 4 mA and a drop of approximately 3.5 Volts at 20 mA across the I/P terminals. Since the controller output terminals are directly in parallel with the I/P terminals, we would expect to see approximately the same voltage there as well (slightly greater due to wire resistance). The lack of known precision in the I/P coil resistance makes it difficult to tell exactly how much current is in the loop for any given voltage measurement we take with a voltmeter. However, if we do know the approximate coil resistance of the I/P, we can at least obtain an estimate of loop current, which is usually good enough for diagnostic purposes.

If the I/P coil resistance is completely unknown, voltage measurements become useless for quantitative determination of loop current. Voltage measurements would be useful only for qualitatively determining loop continuity (i.e. whether there is a break in the wiring between the controller and I/P).

Another example for consideration is this loop-powered 4-20 mA transmitter and controller circuit, where the controller supplies DC power for the loop:



It is very common to find controllers with their own built-in loop power supplies, due to the popularity of loop-powered (2-wire) 4-20 mA transmitters. If we know the transmitter requires a DC voltage source somewhere in the circuit to power it up, it makes sense to include one in the controller, right?

The only voltage measurement that directly and accurately corresponds to loop current is the voltage directly across the 250 Ohm precision resistor. A loop current of 4 mA will yield a voltage drop of 1 Volt, 12 mA will drop 3 Volts, 20 mA will drop 5 Volts, etc.

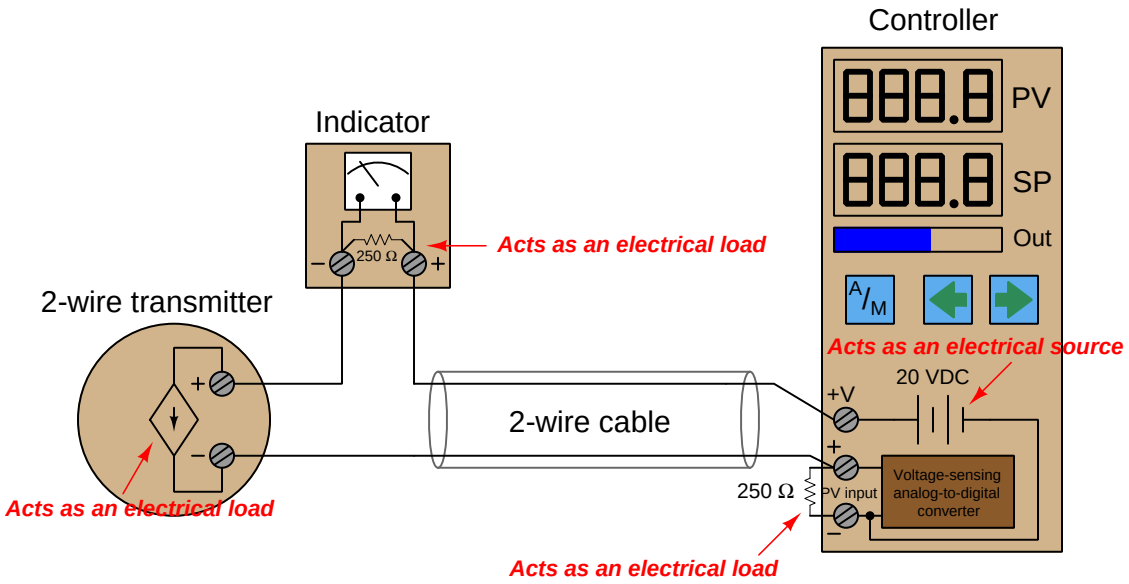
A voltage measurement across the transmitter terminals will show us the *difference* in voltage between the 26 Volt power supply and the voltage dropped across the 250 Ohm resistor. In other words, the transmitter's terminal voltage is simply what is left over from the source voltage of 26 Volts after subtracting the resistor's voltage drop. This makes the transmitter terminal voltage inversely proportional to loop current: the transmitter sees approximately 25 Volts at 4 mA loop current (0% signal) and approximately 21 Volts at 20 mA loop current (100% signal).

The use of the word "approximate" is very intentional here, for loop power supplies are usually non-regulated. In other words, the "26 Volt" rating is approximate and subject to change! One of the advantages of the loop-powered transmitter circuit is that the source voltage is largely irrelevant, so long as it exceeds the minimum value necessary to ensure adequate power to the transmitter. If the source voltage drifts for any reason, it will have no impact on the measurement signal at all, because the transmitter is built as a *current regulator*, regulating current in the loop to whatever value represents the process measurement, regardless of slight changes in loop source voltage, wire resistance, etc. This rejection of power supply voltage changes means the loop power supply need not be regulated, and so in practice it rarely is.

This brings us to a common problem in loop-powered 4-20 mA transmitter circuits: maintaining sufficient operating voltage at the transmitter terminals. Recall that a loop-powered transmitter

relies on the voltage dropped across its terminals (combined with a current of less than 4 mA) to power its internal workings. This means the terminal voltage must not be allowed to dip below a certain minimum value, or else the transmitter will not have enough electrical power to continue its normal operation. This makes it possible to "starve" the transmitter of voltage if the loop power supply voltage is insufficient, and/or if the loop resistance is excessive.

To illustrate how this can be a problem, consider the following 4-20 mA measurement loop, where the controller supplies only 20 Volts DC to power the loop, and an indicator is included in the circuit to provide operators with a field-mounted indication of the transmitter's measurement:



The indicator contains its own 250 Ohm resistor to provide a 1-5 Volt signal for the meter mechanism to sense. This means the total loop resistance has now risen from 250 Ohms to 500 Ohms (plus any wire resistance). At full current (20 mA), this total circuit resistance will drop (at least) 10 Volts, leaving 10 Volts or less at the transmitter terminals to power the transmitter's internal workings. 10 Volts may not be enough for the transmitter to successfully operate, though. The Rosemount model 3051 pressure transmitter, for example, requires a minimum of 10.5 Volts at the terminals to operate.

However, the transmitter *will* operate just fine at lower loop current levels. When the loop current is only 4 mA, for example, the combined voltage drop across the two 250 Ohm resistors will be only 2 Volts, leaving about 18 Volts at the transmitter terminals: more than enough for practically any model of 4-20 mA loop-powered transmitter to successfully operate. Thus, the problem of insufficient supply voltage only manifests itself when the process measurement nears 100% of range. This could be a difficult problem to diagnose, since it appears only during certain process conditions and not others. A technician looking only for wiring faults (loose connections, corroded terminals, etc.) would never find the problem.

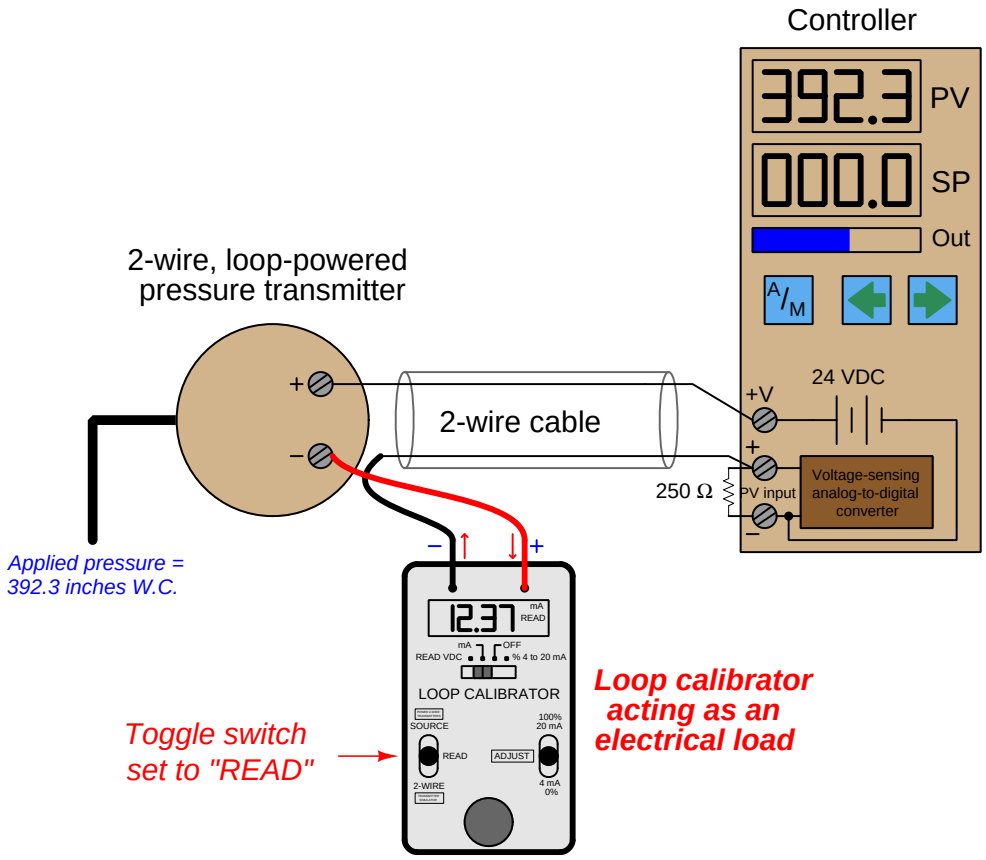
When a loop-powered transmitter is starved of voltage, its behavior becomes erratic. This is especially true of "smart" transmitters with built-in microprocessor circuitry. If the terminal voltage

dips below the required minimum, the microprocessor circuit shuts down. When the circuit shuts down, the current draw decreases accordingly. This causes the terminal voltage to rise again, at which point the microprocessor has enough voltage to start up. As the microprocessor "boots" back up again, it increases loop current to reflect the near-100% process measurement. This causes the terminal voltage to sag, which subsequently causes the microprocessor to shut down again. The result is a slow on/off cycling of the transmitter's current, which makes the process controller think the process variable is surging wildly. The problem disappears, though, as soon as the process measurement decreases enough that the transmitter is allowed enough terminal voltage to operate normally.

2.9.6 Using loop calibrators

Special-purpose electronic test instruments called *loop calibrators* exist for the express purpose of 4-20 mA current loop circuit troubleshooting. These versatile instruments are generally capable of not only measuring current, but also *sourcing* current to unpowered devices in a loop, and also *simulating* loop-powered 4-20 mA transmitters.

A very popular loop calibrator unit is the Altek model 334A, a battery-powered, hand-held unit with a rotary knob for current adjustment and toggle switches for mode setting. The following illustration shows how this calibrator would be used to measure current in a functioning input signal loop[8]:

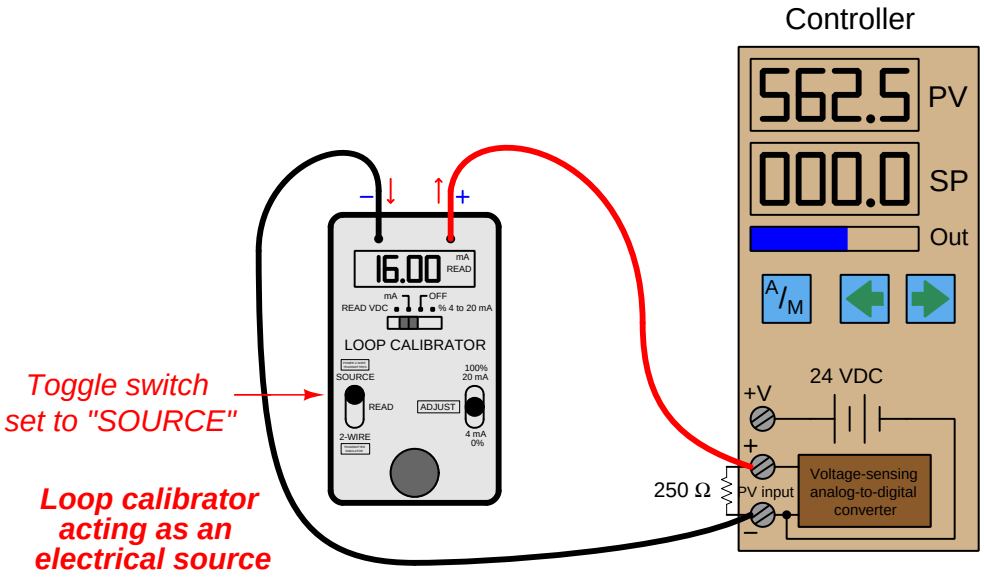


Here, the loop wiring is broken at the negative terminal of the loop-powered transmitter, and the calibrator connected in series to measure current. If this loop had a test diode installed, the calibrator could be connected in parallel with the diode to achieve the same function. Note the polarity of the calibrator's test leads in relation to the circuit being tested: the calibrator is acting

[8] In the following illustrated examples, the transmitter is assumed to be a pressure transmitter with a calibrated range of 0 to 750 inches of water column, 4-20 mA. The controller's PV (process variable) display is appropriately ranged to display 0 to 750 as well.

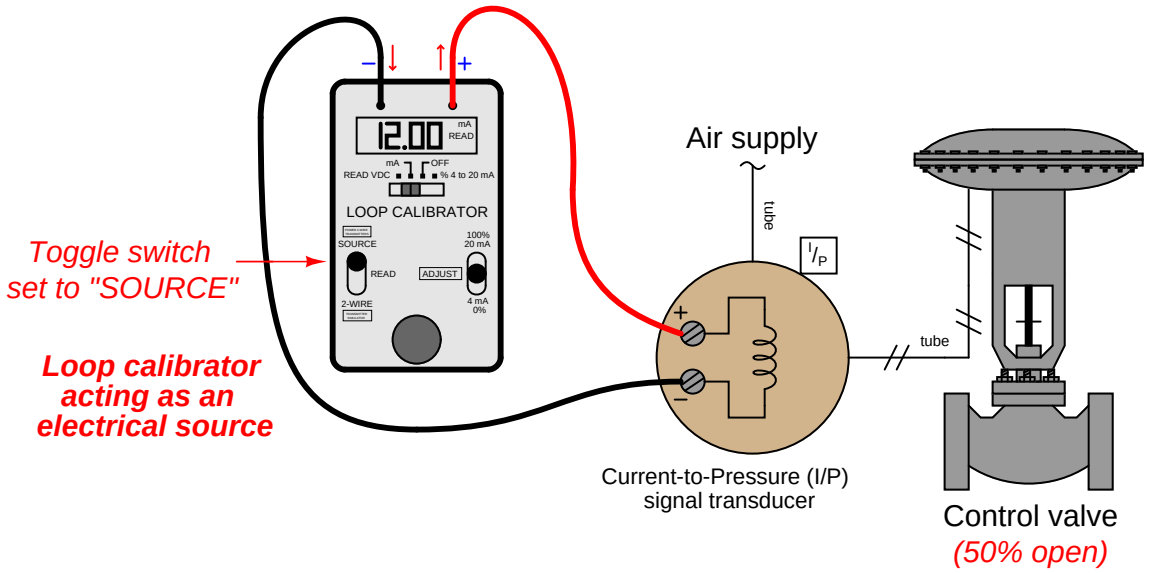
as a passive device (i.e. as a *load* rather than as a *source*), with the more positive loop terminal connected to the calibrator's red test lead and the more negative terminal connected to the black test lead.

The same loop calibrator may be used to *source* (or *drive*) a 4-20 mA signal into an indicating instrument to test the function of that instrument independently. Here, we see the Altek calibrator used as a current source to send a 16.00 mA signal to the PV (process variable) input of the controller, in order to check that the controller properly senses and displays the analog current signal:



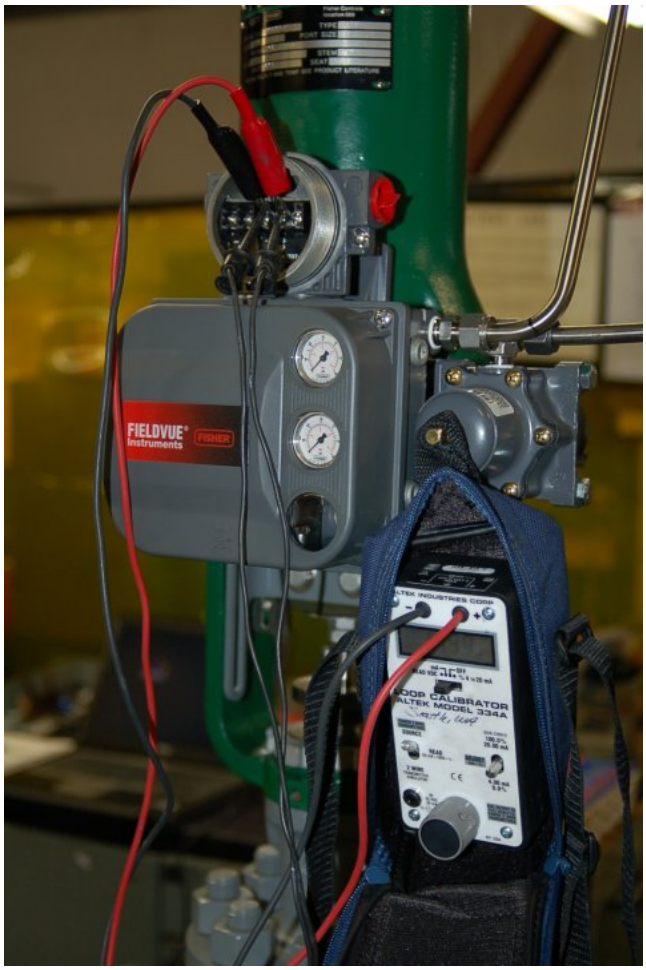
No transmitter need be included in this illustration, because the calibrator takes its place. Note how the calibrator functions here as an active *source* of current rather than a passive load as it was in the last example. Not only does it supply the information (i.e. regulate the current), but it also provides the energy in the circuit. The DC power source inside the controller is not used for loop power, because the calibrator in "source" mode provides the necessary power to drive current through the 250 Ohm resistor.

A very common use of a loop calibrator in "source" mode is to test a control valve for proper calibration, quick response, and to measure friction. Here, the loop calibrator takes place of the loop controller output, serving as the sole source of current to the I/P transducer:



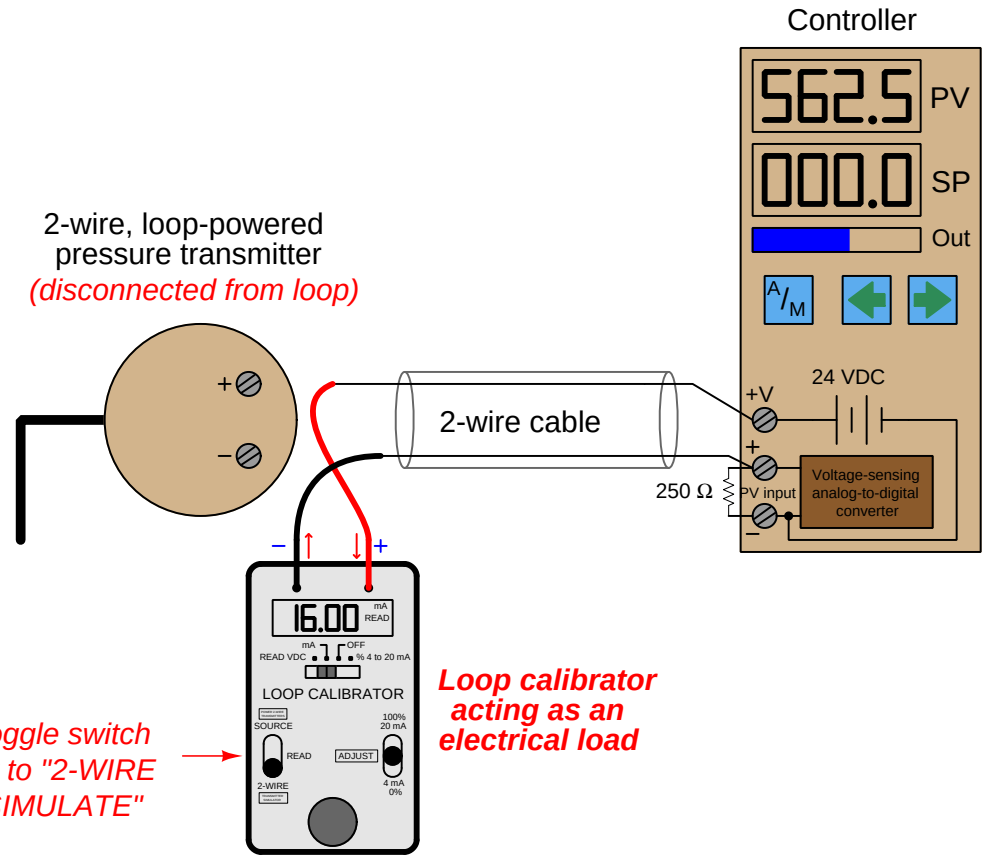
This circuit configuration is extremely useful to any technician testing the response of a control valve, because it allows the signal to be finely adjusted while in the direct presence of the valve to monitor its motion. If a control valve is suspected of having excessive friction in its moving parts, for instance, a technician may test the valve by incrementing and decrementing the loop calibrator's source current in progressively smaller steps. Large step-changes in current should cause the valve to overcome friction and move, but small step-changes will fail to move the valve mechanism when frictional forces exceed the incremental forces produced by the changing pressure.

A photograph showing this very use of a loop calibrator in a valve rebuild shop appears here:



In this particular example, the loop calibrator connects to a device on the control valve called a *positioner*, which is a more sophisticated device than an I/P transducer. In addition to converting a 4-20 mA signal into an air pressure, the positioner also actively monitors the valve stem's position to ensure it goes to the correct position for any given 4-20 mA command signal. Here, the technician is using the loop calibrator to verify the control valve faithfully obeys the command signal through the entire 4 to 20 milliAmpere signal range.

An alternative method of sending a known current signal into an indicating instrument providing loop power is to set the loop calibrator such that it mimics (or *simulates*) the behavior of a loop-powered (2-wire) transmitter. In this mode, the calibrator regulates loop current at a user-determined value, but provides no motivating voltage to drive this current. Instead, it passively relies on the loop's regular voltage source to provide the necessary power:



Note the polarity of the calibrator's test leads: current entering the red lead and exiting the black lead, behaving as an electrical *load* just the same as a loop-powered transmitter. Like a 2-wire transmitter, the calibrator in simulate mode regulates the circuit current while depending on an external voltage source for energy.

A loop calibrator's *simulate transmitter* mode is especially useful for testing the transmitter cable and controller input to ensure any 4-20 mA signal sent by a transmitter will be correctly received and displayed by the controller. This sort of test is commonly performed on newly-installed control systems as part of the commissioning procedure, prior to start-up of the controlled process, in order to verify the controller's process variable input, 24 VDC power supply, and transmitter wiring are all properly functioning. Typically a technician would simulate several different current values (e.g. 4 mA, 8 mA, 12 mA, 16 mA, 20 mA) with the calibrator in "simulate" mode while someone else monitors the controller's PV display and alarms to check for proper function.

A legacy loop calibrator still familiar to many technicians at the time of this writing is the classic Transmation model 1040:

Other examples of vintage loop calibrator technology include the Nassau model 8060 (left) and the Biddle Versa-Cal (right):

A modern loop calibrator manufactured by Fluke is the model 705:

With this calibrator, the *measure*, *source*, and *simulate* modes are accessed by repeatedly pushing a button, with the current mode displayed on the screen:

Note the dual-numeric display, showing both loop current and percentage (assuming a 4-20 mA range).

2.9.7 NAMUR signal levels

One of the intrinsic benefits of a "live zero" analog signal standard such as 4-20 mA is that a wire break (open fault) can immediately be detected by the absence of current in the circuit. If the signal scale started at zero (e.g. 0-20 mA), there would be no way to electrically distinguish between a broken wire and a legitimate 0% signal value. In other words, the "live" LRV point of a 4-20 mA signal provides us with a way to indicate a certain type of circuit fault in addition to indicating an analog measurement during normal operation.

The *NAMUR* signal standard takes this philosophy one step further by defining specific diagnostic meaning to values of current lying outside the 4-20 mA range:

Signal level	**Fault condition**
Output $\leq$ 3.6 mA	Sensing transducer failed low
3.6 mA $<$ Output $<$ 3.8 mA	Sensing transducer failed (detected) low
3.8 mA $\leq$ Output $<$ 4.0 mA	Measurement under-range
21.0 $>$ Output $\geq$ 20.5 mA	Measurement over-range
Output $\geq$ 21.0 mA	Sensing transducer failed high

NAMUR-compliant transmitters are designed to limit their output signals between 3.8 mA and less than 21 mA when functioning properly. Signals lying outside this range indicate some form of failure has occurred within the transmitter or the circuit wiring.

NAMUR-compliant control systems will recognize these errant milliAmpere values as fault states, and may be programmed to take specific action upon receiving these signal values. Such actions include forcing controllers into manual mode, initiating automatic shutdown procedures, or taking some other form of safe action appropriate to the knowledge of a failed process transmitter.

Chapter 3

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?

3.1 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.

3.1.1 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.

3.1.2 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

Voltage

Resistance

Current

Conservation of Electric Charge

Electric circuit

Electrical source

Electrical load

Voltmeter

Ammeter

Annotating diagrams as a problem-solving strategy

Graphing as a problem-solving strategy

Linear function

Process variable

Setpoint

Manipulated variable

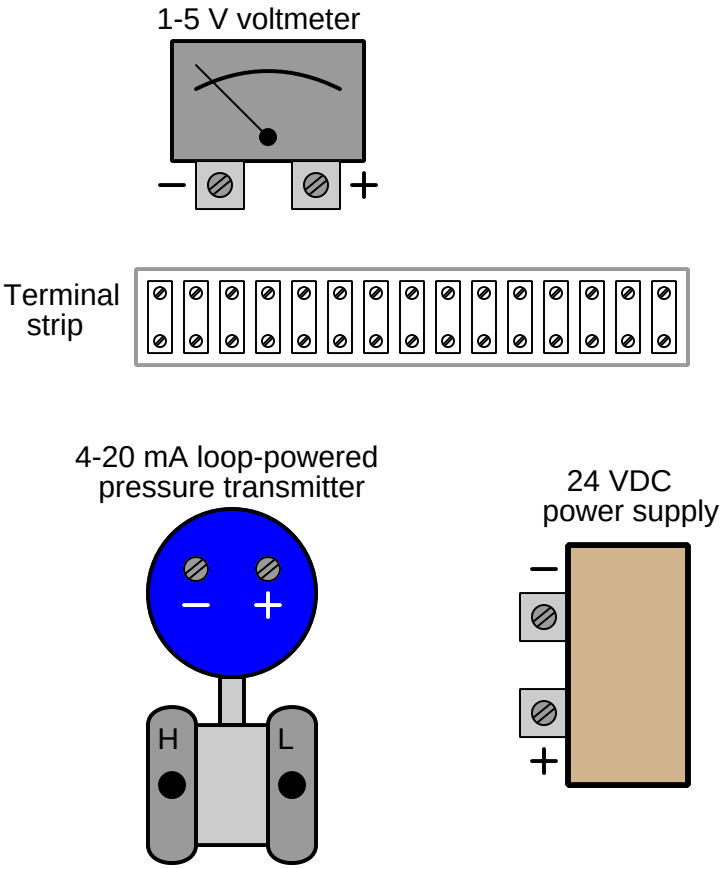
Calibration

Zero

Span

3.1.3 Loop-powered transmitter circuit sketching

Sketch a circuit whereby this loop-powered pressure transmitter sends a signal to an analog voltage meter (acting as a remote pressure gauge). Be sure to route all wiring and attach any necessary components to terminals on the terminal strip:



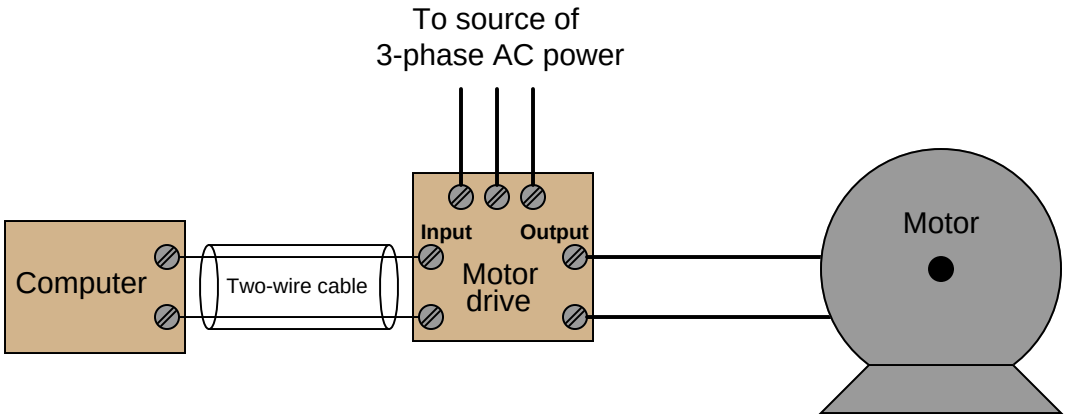
Note: avoid connecting more than two wires to each screw terminal on the terminal strip, to avoid "overcrowding" any connection points, and avoid crossing wires over each other.

Challenges

- What does it mean for a 4-20 mA transmitter to be "loop-powered"?

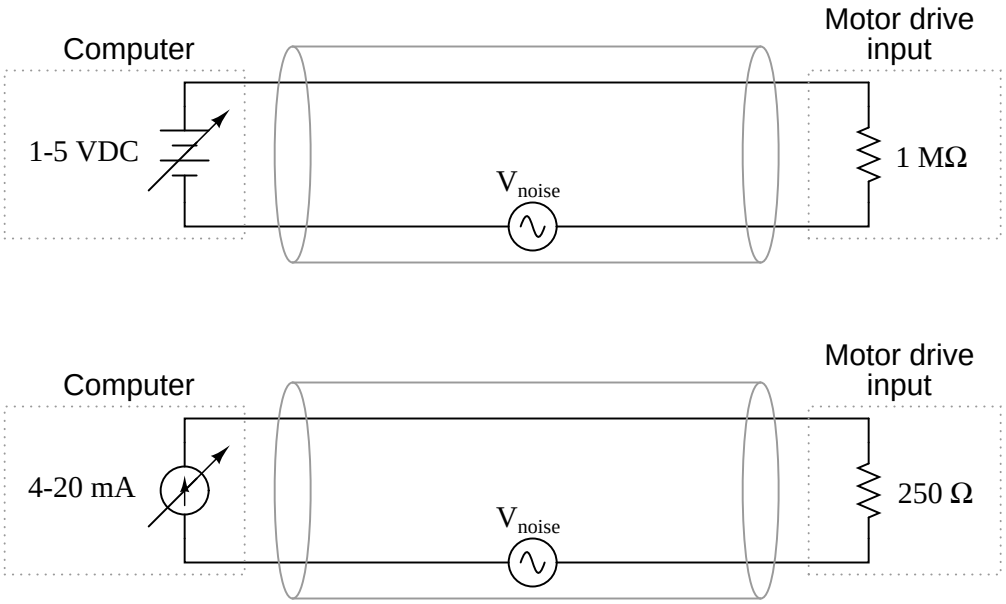
3.1.4 1-5 VDC versus 4-20 mA signaling

Electrical signals are frequently used in industrial control applications to communicate information from one device to another. An example of this is motor speed control, where a computer outputs a speed command signal to a motor "drive" circuit, which then provides metered power to an electric motor:



Two common standards for analog control signals are 1-5 Volts DC and 4-20 mA DC. In either case, the motor will spin faster when this signal from the computer grows in magnitude (1 Volt = motor stopped, 5 Volts = motor runs at full speed; or 4 mA = motor stopped, 20 mA = motor runs at full speed).

At first, it would seem as though the choice between 1-5 Volts and 4-20 mA as control signal standards is arbitrary. However, one of these standards exhibits much greater immunity to induced noise along the two-wire cable than the other. Shown here are two equivalent schematics for these signal standards, complete with an AC voltage source in series to represent the "noise" voltage picked up along the cable's length:



Use the Superposition Theorem to qualitatively determine which signal standard drops the greatest amount of noise voltage across the motor drive input's resistance, thereby most affecting the motor speed control.

Challenges

- What undesirable effects do you think noise superimposed on the DC signal cable would have on motor speed control?
- Why do you suppose the 1-5 Volt signal system requires a much greater input impedance (1 MΩ) than the 4-20 mA signal system? What might happen to the voltage signal received at the motor drive's input terminals if the input resistance were much less?

3.2 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.

3.2.1 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.

3.2.2 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.*

3.2.3 Instrument range tables

Example #1:

An electronic level transmitter has a calibrated range of 0 to 2 feet, and its output signal range is 4 to 20 mA. Complete the following table of values for this transmitter, assuming perfect calibration (no error):

Measured level	Percent of span	Output signal
1.6		
		7.1
	40	

Example #2:

A tachogenerator is used to measure the rotary speed of a machine. Its calibrated range is 0 to 1500 RPM (revolutions per minute) and its corresponding signal output is 0 to 10 Volts DC. Given these range values, calculate the output voltages for the following input shaft speeds, and then describe how you were able to correlate the different speeds to output voltage values:

Shaft speed (RPM)	Output voltage (V)
100	
350	
500	
750	
890	
975	
1230	
1410	
1500	

Example #3:

Suppose you wish to calibrate an electronic pressure transmitter to an input range of 0 to 50 PSI, with an output range of 4 to 20 mA. Complete the following calibration table showing the proper test pressures and the ideal output signal levels at those pressures:

Input pressure (PSI)	**Percent of span**	**Output signal (mA)**
	0	
	25	
	50	
	75	
	100	

Example #4:

An electronic pressure transmitter has a calibrated range of −10 to 60 PSI, and its output signal range is 4 to 20 mA. Complete the following table of values for this transmitter, assuming perfect calibration (no error):

Input pressure (PSI)	**Percent of span**	**Output signal (mA)**
0		
−5		
		12
		13.6
	40	
	22	

Challenges

- What is the purpose of designing a transmitter with a "live zero" range such as 4-20 mA as opposed to a "dead zero" range such as 0-16 mA?

3.2.4 3-wire RTD temperature transmitter circuit

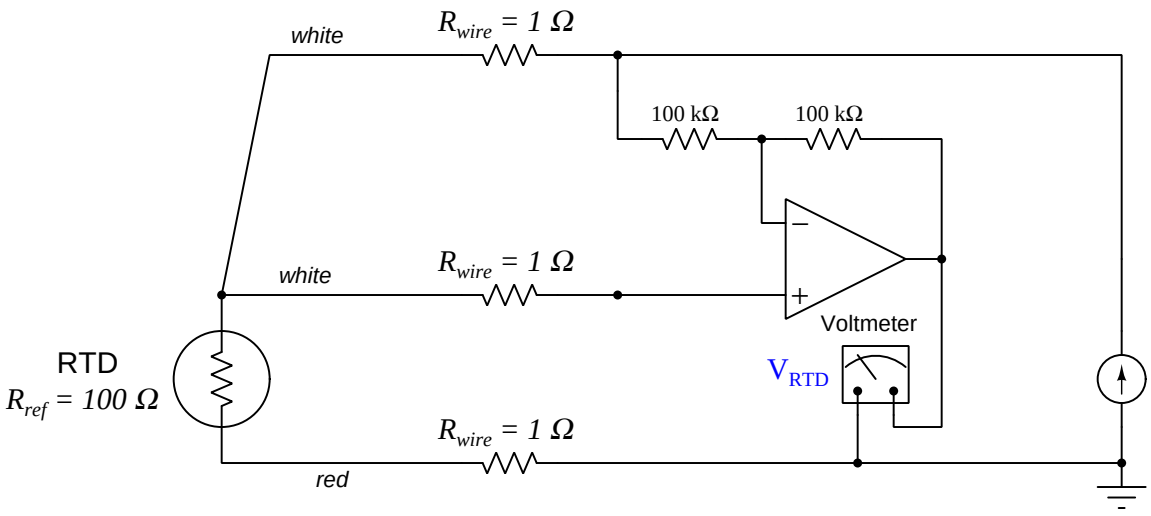
A *Resistive Temperature Detector* (RTD) is a special temperature-sensing element made of fine metal wire, the electrical resistance of which changes with temperature as approximated by the following formula:

$$R_T = R_{ref}[1 + \alpha(T - T_{ref})]$$

Where,

R_T = Resistance of RTD at given temperature T (Ohms)
R_{ref} = Resistance of RTD at the reference temperature T_{ref} of 0 degrees Celsius (Ohms)
α = Temperature coefficient of resistance (Ohms per Ohm per degree)

Assuming an RTD temperature of 35 ^{o}C, an RTD α value of 0.00385 $\Omega/\Omega/^o$C, and a current source output of 0.1 mA, determine the voltage registered by the voltmeter in the following circuit:



Finally, identify how some of the foundational concepts you've studied apply to this circuit: **Ohm's Law**, **Joule's Law**, **Kirchhoff's Voltage Law**, **Kirchhoff's Current Law**, properties of **series** and **parallel** networks, behavior of **PN junctions**, behavior of **BJTs**, the **Shockley diode equation**, etc. Feel free to include any other relevant foundational concepts not listed here.

Challenges

- Why not simply use two wires to sense the RTD's resistance?
- How might a four-wire RTD transmitter circuit be designed?

3.3 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.

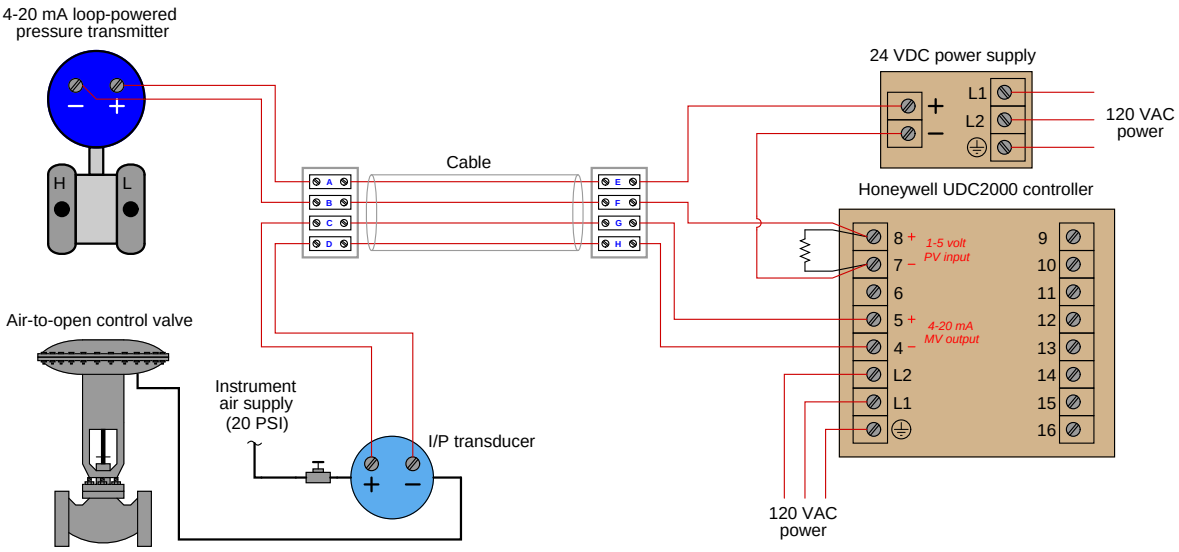
As always, your goal is to fully *explain* your analysis of each problem. Simply obtaining a correct answer is not good enough – you must also demonstrate sound reasoning in order to successfully complete the assignment. Your instructor's responsibility is to probe and challenge your understanding of 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.

3.3.1 Faults in a pressure control system

This pictorial diagram shows the wiring connections for a simple pressure control system, where a loop-powered 4-20 mA pressure transmitter sends a signal to a Honeywell controller, which in turn sends another 4-20 mA signal to a control valve:



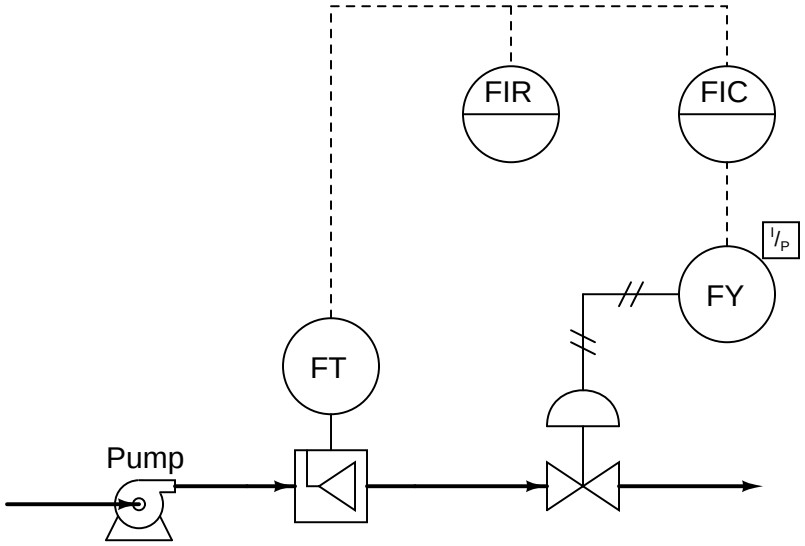
- Sketch all directions of current.
- Identify which electrical devices in this system act as *sources* and which act as *loads*.
- If an operator informs you that the pressure indicated by the Honeywell controller is below range ("pegged" full downscale, reading −25%), what types and locations of electrical faults might you suspect? Are there any non-electrical faults which might also cause this to happen?
- If an operator informs you that the control valve remains fully shut no matter the output value of the controller (even in "manual" mode), what types and locations of electrical faults might you suspect? Are there any non-electrical faults which might also cause this to happen?
- Suppose that a short-circuit developed between the transmitter wires in the four-conductor cable. Explain what effect this would have on the operation of the system, as well as how you could determine that this fault was in the cable (and not in the transmitter) with your only piece of test equipment being a voltmeter.

Challenges

- Should the controller be direct-acting or reverse-acting, or do we not have enough information to determine this?

3.3.2 Failed flow control system

There is a problem somewhere in this liquid flow control system. The controller is in automatic mode, with a setpoint of 65%, yet the flow indicator and the flow controller both register 0.3%: (nearly) zero flow. A piping and instrument diagram (P&ID) of the control system appears here:



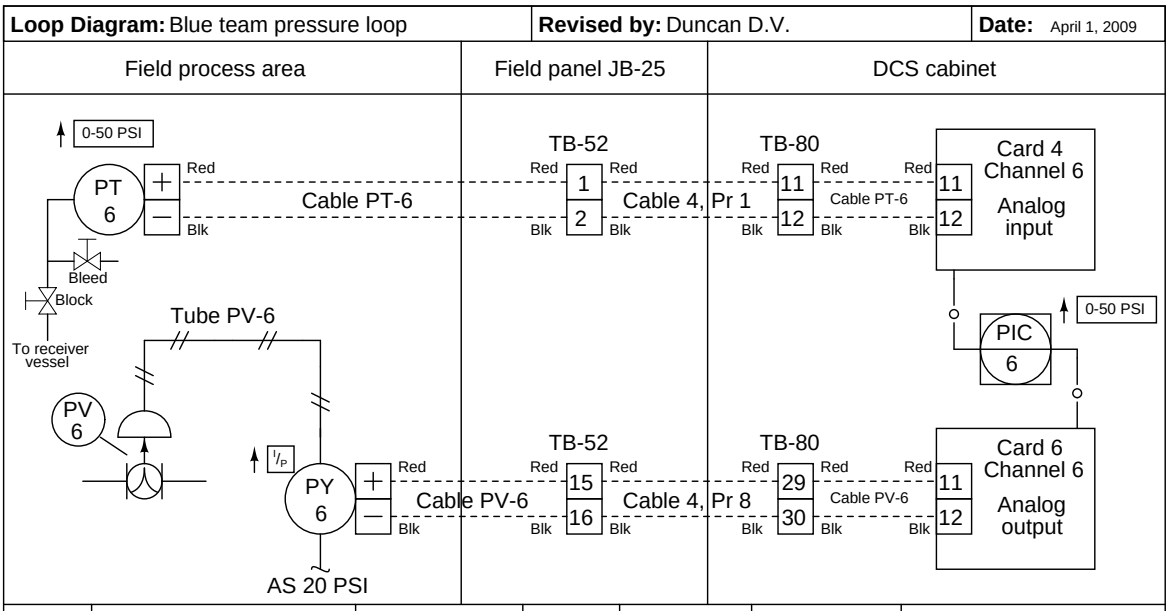
Explain how you would begin troubleshooting this system, and what possible faults could account for the controller not being able to maintain liquid flow at setpoint.

Challenges

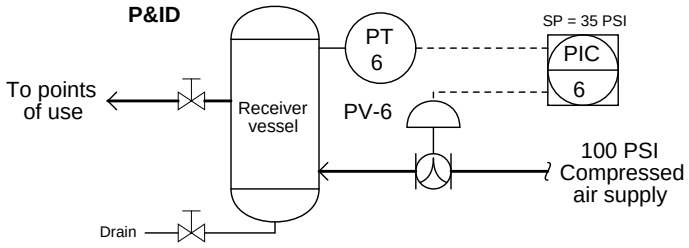
- Explain how you could divide this control system into distinct areas or zones which you may then begin to refer to when "dividing and conquering" the problem.
- Why is it important for technicians to be able to easily convert milliAmpere signal values into corresponding process variable (PV) values?

3.3.3 Faults in an air regulating system

Examine both the loop diagram and the process and instrumentation diagram (P&ID) shown below for an air-pressure regulating system:



Tag #	Description	Manufacturer	Model	Input range	Output range	Notes
PT-6	Pressure transmitter	Rosemount	3051CD	0-50 PSI	4-20 mA	
PIC-6	Controller	Emerson	DeltaV	4-20 mA	4-20 mA	HART-enabled input Direct-acting control
PY-6	I/P transducer	Fisher	846	4-20 mA	3-15 PSI	
PV-6	Control valve	Fisher	Vee-ball	3-15 PSI	0-100%	Fail-open



What kinds of details appear in a loop diagram that are absent in a P&ID?

What kinds of details appear in a P&ID that are absent in a loop diagram?

Next, determine whether or not the following faults (considered individually) could account for this air pressure regulating system failing with zero pressure in the receiver vessel. Answer either "yes" or "no" for each fault:

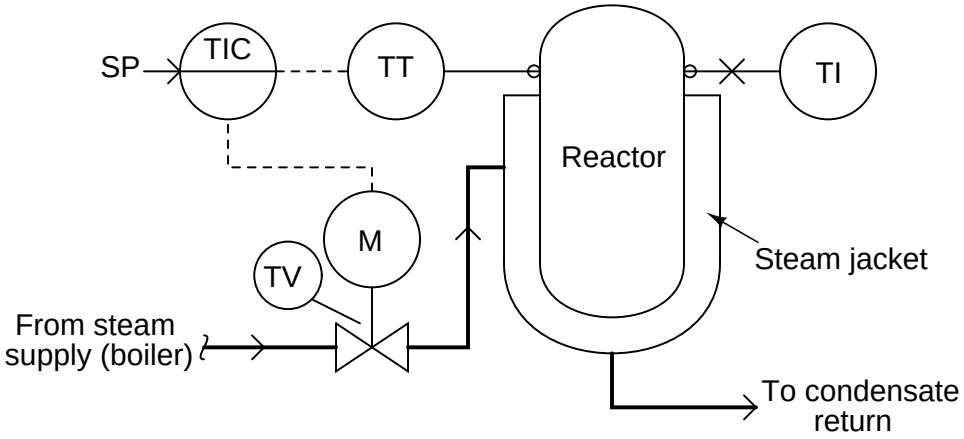
- Receiver vessel drain valve left open
- Transmitter block valve shut and bleed valve open
- Transmitter block valve open and bleed valve shut
- PIC left in manual mode, with output set to 20 mA
- Cable PV-6 severed (failed open)
- I/P air supply shut off
- Short between TB-52, terminals 1 and 2
- Short between TB-52, terminals 15 and 16
- PT-6 miscalibrated, registering 5 PSI too high
- PY-6 output failed high (15 PSI)

Challenges

- Where must the PT be receiving its loop power from?

3.3.4 Failed reactor temperature control system

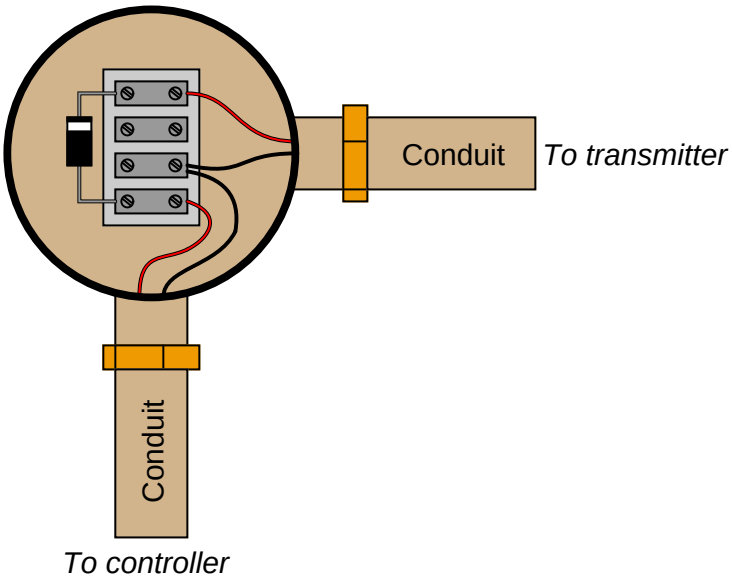
Consider this control system, set up to maintain the temperature of a chemical reactor vessel at a constant ("setpoint") value. The reactor's source of heat is a steam "jacket" where hot steam is admitted through a motor-operated (M) control valve (TV) according to the temperature inside the reactor sensed by the temperature transmitter (TT):



While doing some clean-up work near the reactor, you receive a frantic call from the operator on your two-way radio. He says that the controller (TIC) is registering a temperature of 186 ^{o}F, which is 11 degrees higher than the setpoint of 175 ^{o}F. A temperature this high could ruin the product inside the reactor. He wants you to check the temperature indicator on the side the reactor (TI) and let him know what it reads.

You look at the TI which is nothing more than a mechanical thermometer inserted into the reactor vessel, and see that it registers a temperature of 172 ^{o}F, which is a bit too cold if anything, not too hot. You immediately report this to the operator using your radio, who then asks you to check out the system to see why he's getting a false reading on the controller display.

Fortunately, you have your multimeter and tool set with you, so you proceed to the temperature transmitter to measure the milliAmpere signal it is outputting. Removing a cover from a round junction box on the conduit where the transmitter's wires are routed, you see a terminal block inside with a 1N4001 rectifying diode placed in series with the circuit:



Setting your multimeter to measure milliAmperes, you connect the red and black test leads across the diode. This shorts past the diode, forcing all the current to go through the meter instead of the diode, allowing you to "break in" to the 4-20 mA circuit without having to physically break a wire connection anywhere. Making a mental note to thank your instructor later for teaching you this trick, you see that your multimeter registers 15.683 mA.

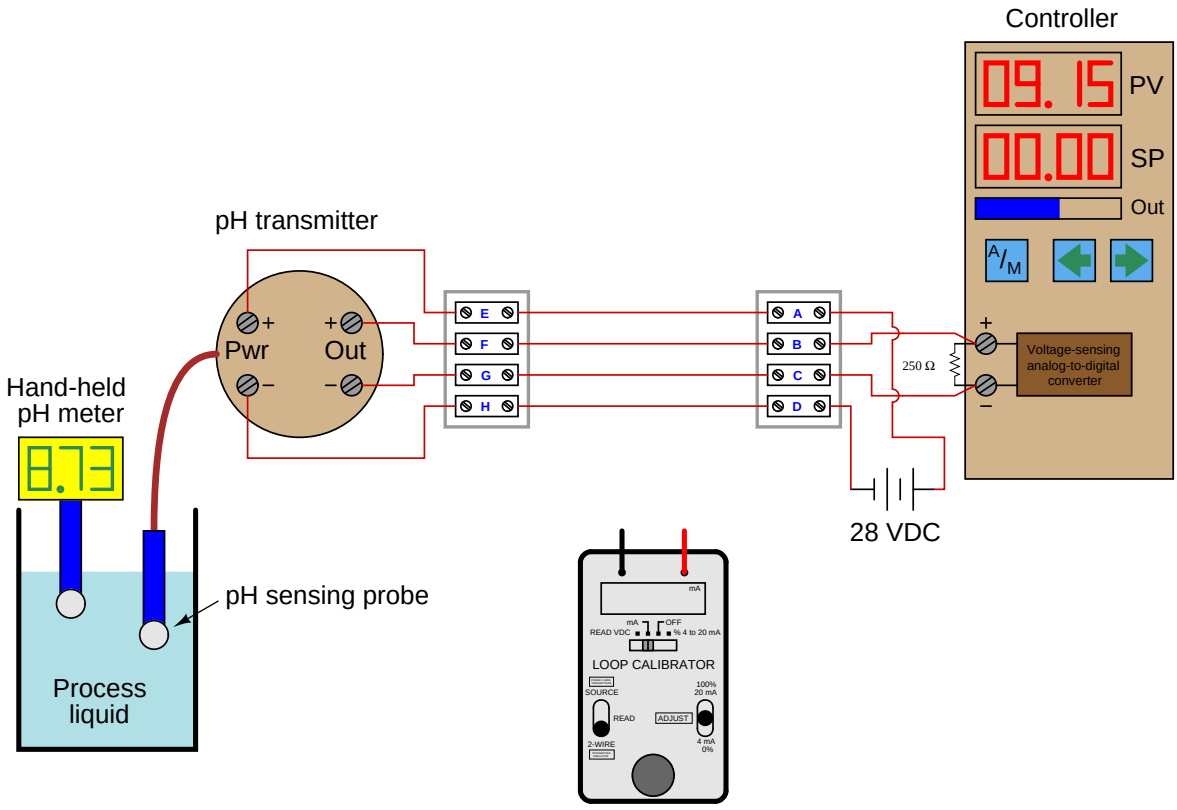
Given a calibrated temperature transmitter range of 100 to 200 degrees F, determine what this current measurement tells you about the location of the problem in this temperature control loop, and explain how you made that determination.

Challenges

- Why is it important for technicians to be able to easily convert milliAmpere signal values into corresponding process variable (PV) values?
- How does the diode perform this useful function of allowing current measurement without breaking the circuit?
- Supposing there were no diode in this loop circuit, how would you suggest we measure the transmitter's output current?
- Is it possible that the fault in this system could be something to do with the control valve? Why or why not?

3.3.5 Faulty pH measurement system

A newly-installed pH measurement system does not seem to be measuring the pH of the process liquid accurately. The indicating controller's display does not match the display of the hand-held pH meter used by an operator:



The calibrated range of the 4-wire pH transmitter is supposed to be 2 to 12 pH, with a 4 to 20 mA signal output range. A technician begins to diagnose the problem by using a *loop calibrator* as a milliammeter to measure the current signal being sent to the indicating controller. The loop calibrator registers 15.43 milliAmperes.

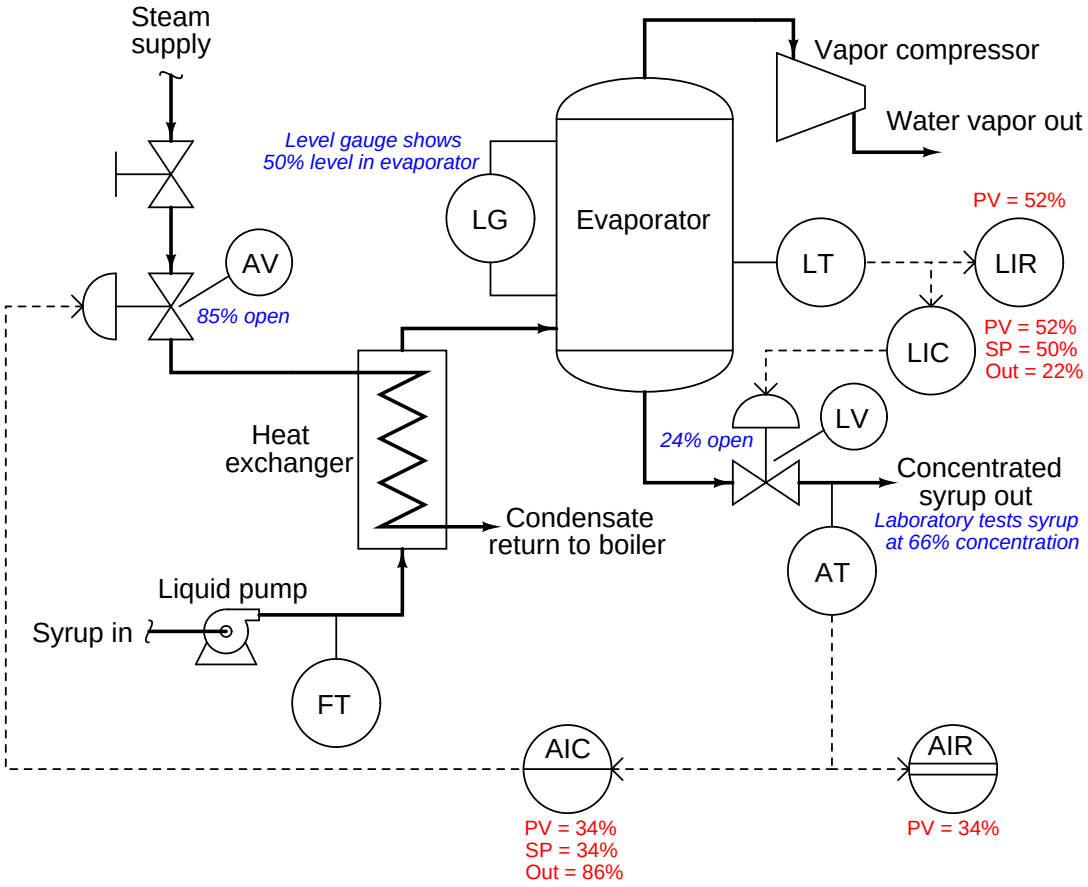
Based on this information, determine where the problem is in this system.

Challenges

- If the technician had no test equipment except for a voltmeter, could a good diagnostic test still be made in this system?
- Identify where you could install a rectifying diode in this circuit to allow convenient measurement of loop current.

3.3.6 Failed maple syrup process

In this process, maple syrup is heated as it passes through a steam heat exchanger, then enters an evaporator where the water boils off. The purpose of this is to raise the sugar concentration of the syrup, making it suitable for use as a food topping. A level control system (LT, LIR, LIC, and LV) maintains constant syrup level inside the evaporator, while an analytical control system (AT, AIR, AIC, and AV) monitors the sugar concentration of the syrup and adjusts steam flow to the heat exchanger accordingly.



Examine the live variable values shown in the above diagram, and then determine where any problems may exist in this syrup concentrating system.

Challenges

- A valuable principle to apply in a diagnostic scenario such as this is *correspondence*: identifying which variables correspond at different points within the system, and which do not. Apply

this comparative test to the variables scenario shown in the diagram, and use the results to defend your answer of where the problem is located and what type of problem it is.

- Explain why the AIR cannot be at fault in this scenario.
- Explain why the AIC cannot be at fault in this scenario.
- Explain why the steam supply cannot be at fault in this scenario.

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

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.

6-10 May 2026 – added comments about 4-20 mA transmitters being current regulators to the Tutorial's first section. Also deleted "instrument" from "instrument technician" throughout the module, as the target audience of this document is both instrument technicians *and* electronics technicians. Also corrected some typographical errors and made other minor edits including additions to some instructor notes.

30 April through 4 May 2025 – added instructor notes, and edited others.

4-6 December 2024 – added explanation of reference temperature to the "3-wire RTD temperature transmitter circuit" Quantitative Reasoning question. Also capitalized units of measurement such as "Ohms" that were lower-case before. Also edited some instructor notes. Also added another Conceptual Reasoning question, one applying Superposition Theorem to 1-5 VDC versus 4-20 mA signaling circuits.

18 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. Also commented out unpopulated sections in the Questions chapter.

2 May 2024 – edited image_6641 to move current arrow to better location.

30 April 2024 – corrected several instances of "Volt" being uncapitalized (i.e. "volt").

24 April 2024 – added Foundational Concepts to the Conceptual Reasoning section.

16-17 April 2024 – added more questions and added content to the Introduction chapter, as well as commented out the Case Tutorial chapter which to date has no content.

14 April 2024 – added a new Tutorial section on loop diagrams.

7-8 April 2024 – added some questions and fixed the size of image_6647.

31 March 2024 – added some questions.

29 April 2023 – document first created from content in the Lessons In Industrial Instrumentation textbook.

Index

10 to 50 mA, 42
2-wire transmitter, 40
4 to 20 mA, 7
4-wire transmitter, 38
4-wire transmitter, active output, 42
4-wire transmitter, passive output, 42
4-wire transmitter, sinking output, 42
4-wire transmitter, sourcing output, 42

Adding quantities to a qualitative problem, 98
Allen-Bradley SLC 500 PLC, 25
Altek model 334A loop calibrator, 58
Annotating diagrams, 97
As-found calibration, 33
As-left calibration, 33

Biddle Versa-Cal loop calibrator, 63

Calibrator, loop, 58
Checking for exceptions, 98
Checking your work, 98
Clamp-on milliammeter, 50
Code, computer, 107
Control loop, 44
Current loop, 35

Dependent current source, 35
Dimensional analysis, 97
Diode, in current loop circuit, 51
Direct-acting transmitter, 45
Drift, 33

Edwards, Tim, 108

Fluke model 771 clamp-on milliammeter, 50

Graph values to solve a problem, 98
Greenleaf, Cynthia, 67

How to teach with these modules, 105
Hwang, Andrew D., 109

I/P transducer, 54
Identify given data, 97
Identify relevant principles, 97
Intermediate results, 97
Inverted instruction, 105
ISA, 43

Knuth, Donald, 108

Lamport, Leslie, 108
Limiting cases, 98
Live zero, 8, 11, 29
Loop calibrator, 58
Loop diagram, 43
Loop sheet, 43
Loop, control, 44
Loop, current, 35
Loop-powered transmitter, 40

Metacognition, 72
Moolenaar, Bram, 107
Murphy, Lynn, 67

Nassau model 8060 loop calibrator, 63

Open-source, 107

Per unit, 15, 30
Percent, 15
PLC, 25
Predictive maintenance, 33
Problem-solving: annotate diagrams, 97
Problem-solving: check for exceptions, 98
Problem-solving: checking work, 98
Problem-solving: dimensional analysis, 97

Problem-solving: graph values, 98
Problem-solving: identify given data, 97
Problem-solving: identify relevant principles, 97
Problem-solving: interpret intermediate results, 97
Problem-solving: limiting cases, 98
Problem-solving: qualitative to quantitative, 98
Problem-solving: quantitative to qualitative, 98
Problem-solving: reductio ad absurdum, 98
Problem-solving: simplify the system, 97
Problem-solving: thought experiment, 97
Problem-solving: track units of measurement, 97
Problem-solving: visually represent the system, 97
Problem-solving: work in reverse, 98
Programmable Logic Controller, 25

Qualitatively approaching a quantitative problem, 98

Reading Apprenticeship, 67
Reductio ad absurdum, 98, 104, 105
Reverse indication, controller output, 37
Reverse-acting transmitter, 45
Rosemount model 3051 differential pressure transmitter, 56

Schoenbach, Ruth, 67
Scientific method, 72
Self-powered transmitter, 38
Shunt resistor, 53
Simplifying a system, 97
Smart transmitter, 57
Socrates, 104
Socratic dialogue, 105
SPICE, 67
Stallman, Richard, 107

Test diode, 51
Thought experiment, 97
Torvalds, Linus, 107
Transmation model 1040 loop calibrator, 63

Units of measurement, 97

Variable-frequency drive, 36
VFD, 10, 36
Visualizing a system, 97
VSD, 10

Work in reverse to solve a problem, 98
WYSIWYG, 107, 108