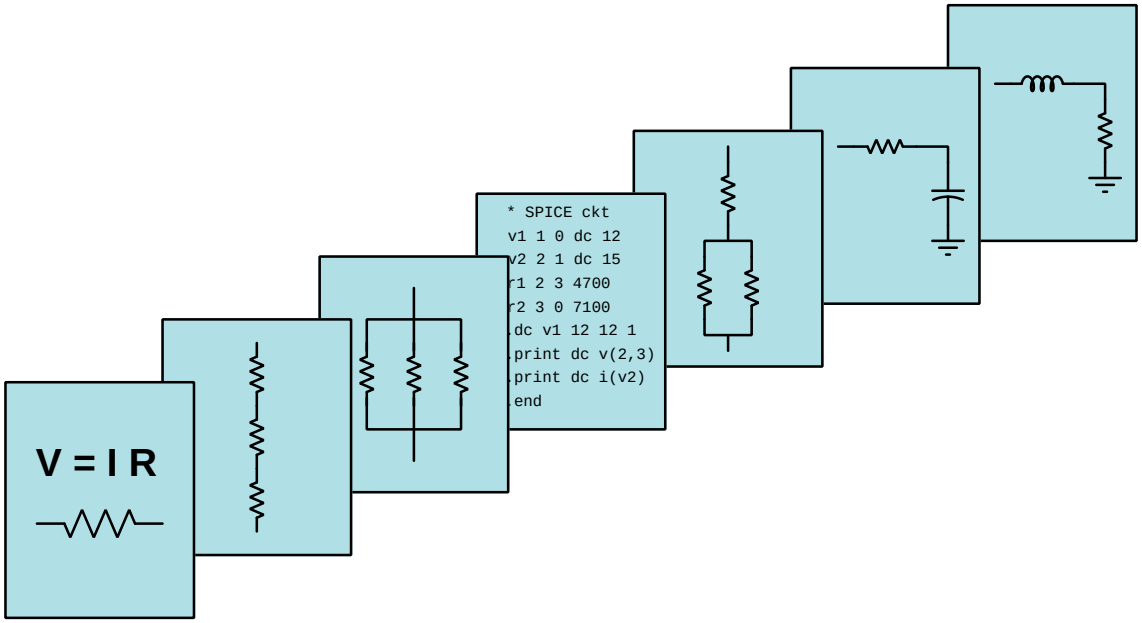
Modular Electronics Learning (ModEL) project



RF Link Budgets

© 2020-2026 by Tony R. Kuphaldt – under the terms and conditions of the Creative Commons Attribution 4.0 International Public License

Last update = 29 September 2026

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

Contents

1 Introduction **3**
- 1.1 Recommendations for students . . . 3
- 1.2 Challenging concepts related to RF link budgets . . . 5
- 1.3 Recommendations for instructors . . . 6

2 Case Tutorial **9**
- 2.1 Example: simple decibel values . . . 10
- 2.2 Example: cable loss decibel calculations . . . 11
- 2.3 Example: path loss calculations . . . 12
- 2.4 Examples: RF link budget calculations . . . 13

3 Tutorial **15**
- 3.1 Effective radiated power . . . 16
- 3.2 Link budget fundamentals . . . 18
- 3.3 Link budget graph . . . 26
- 3.4 Fresnel zones . . . 27

4 Derivations and Technical References **31**
- 4.1 Decibels . . . 32
- 4.2 Near-field versus far-field regions . . . 42
- 4.3 The Poynting vector . . . 46
- 4.4 Maxwell's equations of electromagnetism . . . 50
  - 4.4.1 Gauss's Law of electric fields . . . 50
  - 4.4.2 Gauss's Law of magnetic fields . . . 50
  - 4.4.3 Faraday's Law of electromagnetic induction, extended . . . 51
  - 4.4.4 Ampère's Circuital Law, extended . . . 51
  - 4.4.5 Proof of electromagnetic waves . . . 52
- 4.5 Estimating RF field strength . . . 54

5 Questions **55**
- 5.1 Conceptual reasoning . . . 59
  - 5.1.1 Reading outline and reflections . . . 60
  - 5.1.2 Foundational concepts . . . 61
  - 5.1.3 Link budget graph . . . 64

5.1.4 Cable loss and frequency . . . 65
5.2 Quantitative reasoning . . . 66
5.2.1 Miscellaneous physical constants . . . 67
5.2.2 Introduction to spreadsheets . . . 68
5.2.3 Power ratio and dB conversions . . . 71
5.2.4 Power and loss calculations . . . 72
5.2.5 Waveguide power loss . . . 74
5.2.6 Path loss estimation . . . 75
5.2.7 Fresnel zone interference . . . 76
5.2.8 Fresnel zone ground contact . . . 77
5.2.9 Estimating radio transmitter ERP . . . 78
5.2.10 Minimum transmitter power . . . 79
5.2.11 Link budget margin . . . 80
5.2.12 Maximum distance between antennas . . . 80
5.2.13 Using C to calculate receiver power . . . 81
5.3 Diagnostic reasoning . . . 82
5.3.1 Cable loss calculation error . . . 83
5.3.2 Possible faults in radio SCADA system . . . 84

A Problem-Solving Strategies **87**

B Instructional philosophy **89**
B.1 First principles of learning . . . 90
B.2 Proven strategies for instructors . . . 91
B.3 Proven strategies for students . . . 93
B.4 Design of these learning modules . . . 94

C Tools used **97**

D Creative Commons License **101**

E References **109**

F Version history **111**

Index **112**

Chapter 1

Introduction

1.1 Recommendations for students

Given the inherent power limitations of all radio transmitters, as well as the limited sensitivity of all radio receivers, the design of a robust radio communication "link" between transmitter and receiver must take into account all power gains and losses in between those devices. This module explores the basic principles involved with the calculation of "RF link budgets".

Important concepts related to RF link budgets include **decibels**, **effective radiated power** (ERP), **antenna gain**, **path loss**, **fade loss**, **noise figure**, **signal-to-noise ratio**, **receiver sensitivity**, **bit error rate** (BER), **inverse square law**, **fade margin**, **Fresnel zones**, **line-of-sight**, , , and .

Due to the heavy application of math to link budget analysis, the reader is urged to apply the mathematical principles as soon as possible. The link budget examples 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 the 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 site-test a radio transmitter/receiver pair for effective operation? What hypothesis (i.e. prediction) might you pose for that experiment, and what result(s) would either support or disprove that hypothesis?
- Why do we typically use decibels to express power gains and losses in RF systems?
- How can an antenna have a *gain* as a purely passive device, without violating the Law of Energy Conservation?
- What do the following terms mean: *dB*, *dBm*, *dBd*, and *dBi*?

- Where does noise originate in an RF system?
- What might happen if an RF link is not properly budgeted?
- How does the inverse square law apply to RF systems?
- What is "fade" and how do we compensate for it in an RF system?
- What accounts for path loss in an RF link?
- How do path loss and fade loss differ from one another?
- Can path loss ever be completely eliminated?
- Can fade loss ever be completely eliminated?
- What does BER represent, and why is it important in a digital radio system?
- What does it mean to have a "margin" built into a budget calculation of any type?
- What factors determine the dimensions of a Fresnel zone, and why is this concept important to RF links?
- What is a site test, and why is this a good idea to conduct for any RF system?

1.2 Challenging concepts related to RF link budgets

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

- **Decibels** – "decibels" are an attempt to express power ratios (i.e. power gains or attenuation factors) logarithmically rather than linearly, and as such they tend to generate confusion for students less familiar (or unfamiliar) with exponential and logarithmic functions.
- **Link budget calculations** – aside from logarithmic and decibel-related calculations, another complicating factor with RF link budgets is the sheer number of terms that must be accurately accounted for. I recommend writing all values in a table format (or in a link budget *graph*) to keep them neatly organized!

1.3 Recommendations for instructors

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

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

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

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

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

- **Outcome** – Apply the concept of decibels to RF signal attenuation

  Assessment – Convert between gain values (as ratios) and decibels; e.g. pose problems in the form of the "Power ratio and dB conversions" Quantitative Reasoning question.

  Assessment – Express RF signal powers in absolute decibel units (e.g. dBm, dBW, etc.); e.g. pose problems in the form of the "Power and loss calculations" Quantitative Reasoning question.

  Assessment – Calculate path loss value given transmitted and received powers; e.g. pose problems in the form of the "Path loss estimation" Quantitative Reasoning question.

- **Outcome** – Apply the concept of Fresnel zones to RF signal communication

  Assessment – Calculate Fresnel zone size; e.g. pose problems in the form of the "Fresnel zone interference" Quantitative Reasoning question.

- **Outcome** – Perform an entire RF link budget analysis for a system

  Assessment – Determine the power output necessary from a radio transmitter to be reliably detected by a receiver given all necessary physical parameters; e.g. pose problems in the form of the "Minimum transmitter power" Quantitative Reasoning question.

  Assessment – Determine the power margin afforded by a RF link budget given all equipment parameters; e.g. pose problems in the form of the "Link budget margin" Quantitative Reasoning question.

- **Outcome** – Independent research

Assessment – Read and summarize in your own words reliable documents on antenna feed systems. Publications by the American Radio Relay League (ARRL) are highly recommended!

Assessment – Locate antenna feed system component datasheets and properly interpret some of the information contained in those documents including impedance, RF connector types, physical dimensions, power rating, etc.

Chapter 2

Case Tutorial

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

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

2.1 Example: simple decibel values

Decibels are a logarithmic representation of power ratios, with negative decibel values representing power attenuation and positive decibel values representing power gain. Every addition or subtraction of 10 dB exactly represents a 10-fold multiplication or division of power ratio. A simple approximation that is useful in decibel calculations is that an addition or subtraction of 3 dB approximately represents a two-fold multiplication or division of power ratio.

Here are some numerical examples based on the equivalence of ± 10 dB with orders of magnitude:

- Power ratio gain of 10 = **10 dB**
- Power ratio gain of 100 = **20 dB**
- Power ratio gain of 1000 = **30 dB**
- Power ratio gain of 10000 = **40 dB**
- Power ratio reduction of $\frac{1}{10}$ = **−10 dB**
- Power ratio reduction of $\frac{1}{100}$ = **−20 dB**
- Power ratio reduction of $\frac{1}{1000}$ = **−30 dB**
- Power ratio reduction of $\frac{1}{10000}$ = **−40 dB**

Here are some numerical examples based on the approximation of ± 3 dB being equivalent to doubling or halving of power ratio:

- Power ratio gain of 2 ≈ **3 dB**
- Power ratio gain of 4 ≈ **6 dB**
- Power ratio gain of 8 ≈ **9 dB**
- Power ratio gain of 16 ≈ **12 dB**
- Power ratio reduction of $\frac{1}{2}$ ≈ **−3 dB**
- Power ratio reduction of $\frac{1}{4}$ ≈ **−6 dB**
- Power ratio reduction of $\frac{1}{8}$ ≈ **−9 dB**
- Power ratio reduction of $\frac{1}{16}$ ≈ **−12 dB**

Example approximations:

- Power ratio of $500 = 10 \times 10 \times 10 \times \frac{1}{2} \approx (10 + 10 + 10 - 3)$ dB = **27 dB**
- Power ratio of $80 = 10 \times 2 \times 2 \times 2 \approx (10 + 3 + 3 + 3)$ dB = **19 dB**
- Power ratio of $0.2 = \frac{1}{10} \times 2 \approx (-10 + 3)$ dB = **−7 dB**
- Power ratio of $0.005 = \frac{1}{10} \times \frac{1}{10} \times \frac{1}{2} \approx (-10 - 10 - 3)$ dB = **−23 dB**

2.2 Example: cable loss decibel calculations

If a radio transmitter sends 15 Watts of power into a coaxial cable having a loss rating of −0.16 dB per meter, and the cable is 20 meters in length, the power available at the far end of the cable must be **7.1795 Watts**.

If an RF signal source sends 10 mW into one end of a 5-meter cable and the other end of the cable outputs 8 mW of signal power, the cable must have a loss rating of **−0.194 dB per meter**.

An RF signal of 2 Watts exits a cable 3 meters in length having a loss rating of −0.14 dB per meter. The amount of power being input to this cable must be **2.203 Watts**.

An RF signal of 15 Watts enters one end of a cable having a loss rating of −0.21 dB per foot, and an RF signal of 10 Watts exits the other end of that same cable. The cable's length must be **8.38530 feet**.

A 50-Watt radio transmitter connected to a 12-foot cable having a loss rating of 0.041 dB/foot terminates at a lightning arrestor with a loss rating of 0.5 dB, and finally to an antenna. The amount of power received by this antenna will be **39.7896 Watts**.

2.3 Example: path loss calculations

A radio transmitter broadcasts a 800 kHz RF signal at a power level of 10 kW into the surrounding environment. At a distance of 40 kilometers away from the transmitting antenna, the signal power has spread out to just **0.0055657388 Watts**.

A radio transmitter broadcasts a 92 MHz RF signal such that the antenna's actual power output into the surrounding environment is 75 kW, and a radio receiver unit requires a signal that is at least −50 dBm at its antenna to reliably interpret this broadcast. The maximum distance away that this receiver can pick up the transmitter's signal is **710.647 kilometers**, assuming no fade loss.

A 900 MHz radio transmitter using a +6 dBi antenna is able to broadcast a signal to a single receiving station 50 km away with bare-minimum signal integrity. If the transmitter's antenna is upgraded to a more directional model having a gain of +18 dBi, that maximum distance now becomes **199.054 km**.

A 900 MHz radio transmitter is able to broadcast a signal to a single receiving station 10 km away with bare-minimum signal integrity. If the entire radio system is changed from 900 MHz to 2.4 GHz, the new maximum distance will be **3.75 km** assuming identical cable losses, identical antenna gains, identical transmitter power, identical receiver sensitivity, and identical operating conditions (weather, obstacles, no Fresnel zone interference in either case, etc.).

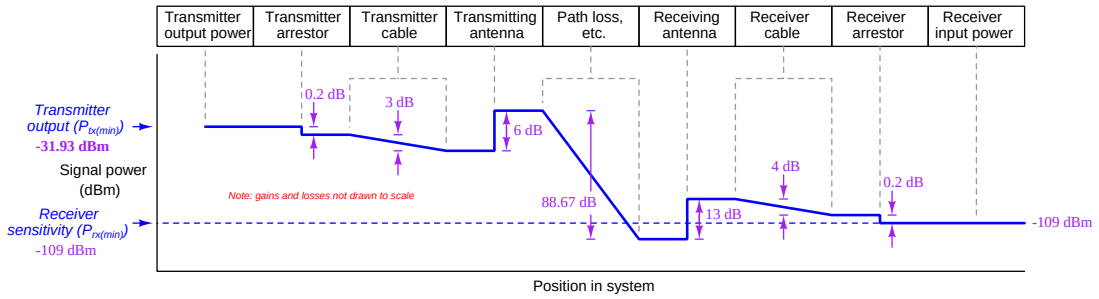
Suppose a radio transmitting antenna radiates 45 Watts of RF energy in a particular direction, and a receiving antenna intercepts 1.4 microWatts of RF energy from it. Assuming a frequency of 700 kHz and clear (uninterrupted) space in between with no Fresnel zone interference or fade loss, the distance between these two antennas must be **193.22 km**.

2.4 Examples: RF link budget calculations

The following parameters describe an RF communication system:

- Frequency = 900 MHz
- Path distance = 720 meters
- Receiver sensitivity = −109 dBm
- Transmitter cable = −0.25 dB per foot @ 900 MHz ; 12 feet long
- Transmitter lightning arrestor = −0.2 dB insertion loss
- Transmitter antenna = 6 dB gain
- Receiver lightning arrestor = −0.2 dB insertion loss
- Receiver cable = −0.25 dB per foot @ 900 MHz ; 16 feet long
- Receiver antenna = 13 dB gain

From these figures, we may calculate the bare-minimum minimum transmitter power required in this application (with no fade margin) to be −31.93 **dBm** or **0.643 microWatts**.



Lightning arrestors are installed at both the MTU and the RTU antennas in a two-station radio SCADA system. Each arrestor has an advertised power loss of -0.55 dB. A new transmitter with increased power output is specified to overcome these arrestor losses. Given that the old transmitter had an output power of 3.2 milliWatts, the new transmitter must have an output power of **4.122 milliWatts** to overcome these new losses.

Suppose an RTU in a radio SCADA system is able to receive an RF signal strength of 0.17 microWatts from the MTU using a standard "whip" antenna having a gain of 5.8 dBi. If we were to replace this whip antenna with a Yagi having a gain of 11.4 dBi, the RTU would now receive **0.617 microWatts** of RF power.

Suppose a radio receiver unit is equipped with a lightning arrestor that presents a signal power attenuation of -0.4 dB. During a storm, this arrestor is damaged. Lacking a replacement on hand, you decide to remove it from the system and re-connect the antenna directly to the receiver with no lighting arrestor at all, at least until you can purchase and install a new arrestor. If the receiver originally was able to intercept an RF signal of 0.8 microWatts with the arrestor in place, the new received signal power level without the arrestor in place will be **0.8772 microWatts**.

Chapter 3

Tutorial

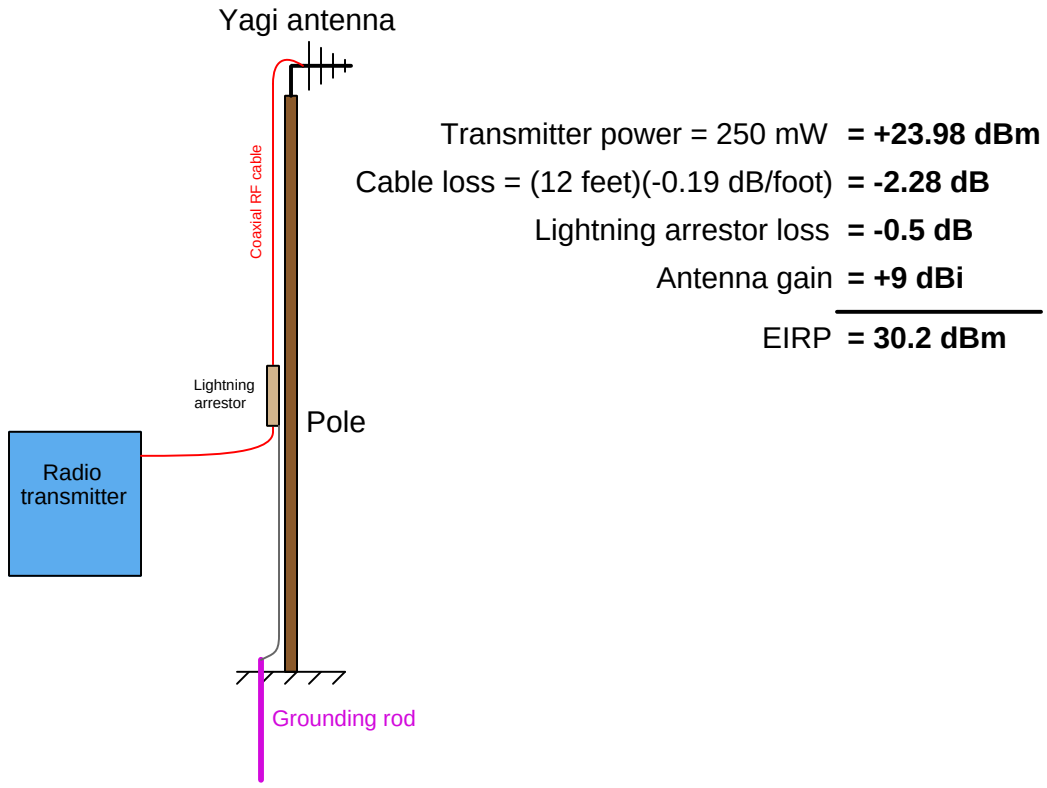
All radio transmitters are limited in power output. All radio receivers are limited in sensitivity to weak signals. Every device, every obstacle, and even empty space itself affects the attenuation of the radio-frequency (RF) signal from transmitter to receiver. Energy is always conserved, and so in order to design a reliable radio communication system one must carefully *budget* the available power in the same way one must budget financial resources when embarking on a project to ensure it may be successfully completed.

Each of the sections in this Tutorial outline a different aspect of *RF link budgeting*, and should be considered an introduction to the topic rather than a comprehensive tutorial.

3.1 Effective radiated power

The effectiveness of any radio system depends on more than just the amount of transmitted power and the sensitivity of the receiver(s). In addition to these important factors we must also consider dissipative power losses in cables, connectors, lightning arrestors, and other elements along the entire signal path, as well as any effective gain in power resulting from the antennas' directionality. A commonly accepted way to quantify the effectiveness of a radio transmitter is to rate it against a standard reference consisting of an ideal dipole antenna connected to a 1 milliWatt transmitter with no losses. Typically expressed in units of decibels, this is called the *Effective Radiated Power*, or *ERP*. If the ideal antenna model is isotropic instead of a dipole, then the calculation result is called the *Effective Isotropic Radiated Power*, or *EIRP*.

Let us consider the following example, where a 2.4 GHz radio transceiver outputs 250 milliWatts[1] of radio-frequency (RF) power to a Yagi antenna through a type LMR 195 coaxial cable 12 feet in length. A lightning arrestor with 0.5 dB loss is also part of the cable system. We will assume an antenna gain of 9 dBi for the Yagi and a loss of 0.19 dB per foot for the LMR 195 cable:



The EIRP for this radio system is 30.2 dBm: the sum of all gains and losses expressed in decibels.

[1] Actual signal power is typically expressed as a decibel ratio to a reference power of either 1 milliWatt (dBm) or 1 Watt (dBW). Thus, 250 mW of RF power may be expressed as $10 \log \frac{250}{1} = 23.98$ dBm or as $10 \log \frac{0.25}{1} = -6.02$ dBW. Power expressed in unit of dBm will always be 30 dB greater (1×10^3 greater) than power expressed in dBW.

This means our Yagi antenna will radiate 30.2 dB (1047 times) more RF power in its most effective direction than an isotropic antenna would radiate in the same direction powered by a 1 milliWatt transmitter. Note that if our hypothetical radio system also included an RF amplifier between the transceiver and the antenna, its gain would have to be included in the EIRP calculation as well.

An important concept here is that of *antenna gain* which is not quite the same conceptually as gain provided by an amplifier because an antenna is a passive component having no source of energy that might be used to boost the radio-frequency (RF) signal. Instead, the "gain" of any antenna is merely a measure of how much more focused its radiation pattern will be in its most-directed axis compared to a perfectly omnidirectional antenna. However, a highly-directional antenna will have the same effect (i.e. detected by any radio receiver) as a non-directional antenna utilizing a power amplifier to boost the signal strength. This is why we call it the *effective radiated power*, because a directional antenna focusing its radiated energy mostly in one direction is able to transmit to (or interfere with) other devices and systems as effectively as a non-focused antenna with a higher-power signal.

A practical application of EIRP is how the Federal Communications Commission (FCC) sets limits on radio transmitters within the United States. Not only is gross transmitter power limited by law within certain frequency ranges, but also the EIRP of a transmitting station. This makes sense, since a more directional transmitting antenna (i.e. one having a greater gain value) will make it appear as though the transmitter is more powerful than it would be radiating from a less-directional antenna. If FCC limits were based strictly on transmitter power output rather than EIRP, it might still be possible for an otherwise power-compliant transmitting station to generate excessive interference through the use of a highly directional antenna. This is why it is illegal, for example, to connect a large antenna to a low-power transmitting device such as a two-way ("walkie-talkie") radio unit: the two-way radio unit may be operated license-free only because its EIRP is low enough not to cause interference with other radio systems. If someone were to connect a more efficient antenna to this same two-way radio, its effective radiated power may increase to unacceptable levels (according to the FCC) even though the raw power output by the transmitter circuitry has not been boosted.

3.2 Link budget fundamentals

Electromagnetic radiation is a medium useful for conveying information, "linking" data from one physical location to another. In order for this to work, the amount of signal loss between transmitter and receiver must be small enough that the signal does not become lost in radio-frequency "noise" originating from external sources and from within the radio receiver itself. We may express radio-frequency (RF) power in terms of its comparison to 1 milliWatt: 0 dBm being 1 milliWatt, 3.01 dBm being 2 milliWatts, 20 dBm being 100 milliWatts, etc. We may use dBm as an absolute measurement scale for transmitted and received signal strengths, as well as for expressing how much ambient RF noise is present (called the "noise floor" for its appearance at the bottom of a spectrum analyzer display showing the frequency-domain view of the electromagnetic spectrum for a given test). We may use plain dB to express relative gains and losses along the signal path.

The basic idea behind an "RF link budget" is to add *all* gains and losses in an RF system – from transmitter to receiver with all intermediate elements accounted for – to ensure there is a large enough difference between signal and noise to ensure good data communication integrity. If we account all gains as positive decibel values and all losses as negative decibel values, the signal power at the receiver will be the simple sum of all the gains and losses:

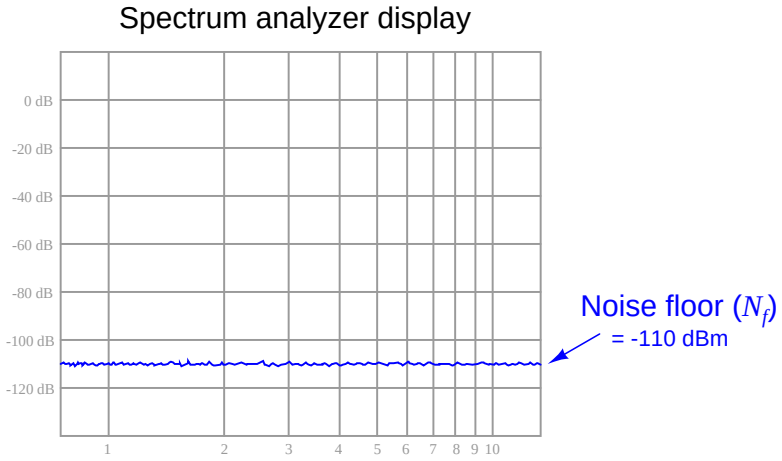
$$P_{rx} = P_{tx} + G_{total} + L_{total}$$

Where,

P_{rx} = Signal power delivered to receiver input (dBm)
P_{tx} = Transmitter output signal power (dBm)
G_{total} = Sum of all gains (amplifiers, antenna directionality, etc.), a positive dB value
L_{total} = Sum of all losses (cables, filters, path loss, fade, etc.), a negative dB value

This formula tells us how much signal power will be available at the radio receiver, but usually the purpose of calculating a link budget is to determine how much radio transmitter power will be *necessary* in order to have adequate signal strength at the receiver. More transmitter power adds expense, not only due to transmitter hardware cost but also to FCC licenses that are required if certain power limitations are exceeded. Excessive transmitter power may also create interference problems with other radio and electronic systems. Suffice it to say we wish to limit transmitter power to the minimum practical value.

In order for a radio receiver to reliably detect an incoming signal, that signal must be sufficiently greater than the ambient RF noise. All masses at temperatures above absolute zero radiate electromagnetic energy, with some of that energy falling within the RF spectrum. This *noise floor* value may be calculated[2] or empirically measured using an RF spectrum analyzer as shown in this simulated illustration:



On top of the ambient noise, we also have the *noise figure* of the receiver itself (N_{rx}): noise created by the electronic circuitry within the radio receiver. Thus, the minimum signal power necessary for the receiver to operate reliably ($P_{rx(min)}$) is equal to the decibel sum of the noise floor and the noise figure, by a margin called the minimum *signal-to-noise* ratio:

$$P_{rx(min)} = N_f + N_{rx} + S$$

Where,

$P_{rx(min)}$ = Minimum necessary signal power at the receiver input (dBm)
N_f = Noise floor value (dBm)
N_{rx} = Noise figure of radio receiver (dB)
S = Desired signal-to-noise ratio margin (dB)

Substituting this decibel sum into our original RF link budget formula and solving for the minimum necessary transmitter power output ($P_{tx(min)}$), we get the following result:

$$P_{tx(min)} = N_f + N_{rx} + S - (G_{total} + L_{total})$$

[2]Noise power may be calculated using the formula $P_n = kTB$, where P_n is the noise power in Watts, k is Boltzmann's constant (1.38×10^{-23} J/K), T is the absolute temperature in Kelvin, and B is the bandwidth of the noise in Hertz. Noise power usually expressed in units of dBm rather than Watts, because typical noise power values for ambient temperatures on Earth are so incredibly small.

It is common for radio receiver manufacturers to aggregate the noise floor, noise figure, and a reasonable signal-to-noise ratio into a single parameter called *receiver sensitivity*. The "sensitivity" of a radio receiver unit is the minimum amount of signal power (usually expressed in dBm) necessary at the input connector for reliable operation despite the inevitable presence of noise. If we simply express receiver sensitivity as $P_{rx(min)}$ and substitute this term for the sum of noise floor, noise figure, and signal-to-noise margin ($N_f + N_{rx} + S$) in the last formula, we see that the difference in receiver sensitivity (expressed in absolute decibels) and the sum of any gains and losses in the link (also expressed in decibels) tells us the minimum transmitter power required:

$$P_{tx(min)} = P_{rx(min)} - (G_{total} + L_{total})$$

Where,

$P_{tx(min)}$ = Minimum necessary transmitter output signal power, in dBm
$P_{rx(min)}$ = Receiver sensitivity (minimum necessary received signal power), in dBm
G_{total} = Sum of all gains (amplifiers, antenna directionality, etc.), a positive dB value
L_{total} = Sum of all losses (cables, filters, path loss, fade, etc.), a negative dB value

For digital radio receivers, sensitivity is a function of error rate: the fewer errors desired, the more signal power required. To give a practical example, the datasheet for a particular model of industrial 900 MHz radio transceiver shows a sensitivity of -110 dBm at a bit error rate (BER) of 10^{-4} bits (one error for every 10^4 bits received) and a sensitivity of -108 dBm at a BER of 10^{-6} bits. This relationship between signal power and error rate should make intuitive sense: the more powerful the signal compared to any background noise, the more reliably it will be received; the weaker the signal, the more it will become corrupted by noise and therefore the more errors we would expect to see over time.

Among the losses encompassed in L_{total} are *path loss* and *fade*. Path loss is the natural loss of signal strength with increasing distance from the radiation source. As electromagnetic waves propagate outward through space, they inevitably spread. The degradation in signal strength with increasing distance follows the *inverse square law*[3], where power decreases with the square of distance. Thus, doubling the distance from transmitting antenna to receiving antenna attenuates the signal by a factor of four ($\frac{1}{2^2}$, or -6.02 dB). Tripling the distance from transmitting antenna to receiving antenna attenuates the signal by a factor of nine ($\frac{1}{3^2}$, or -9.54 dB).

Path loss for free-space conditions is a rather simple function of distance and wavelength[4]:

$$L_p = -20 \log \left(\frac{4\pi D}{\lambda} \right)$$

Where,

L_p = Path loss, a negative dB value

D = Distance between transmitting and receiving antennas

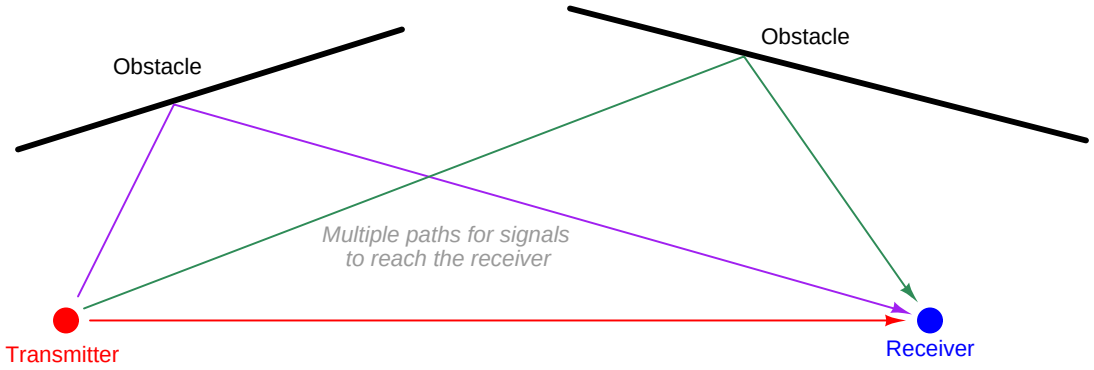
λ = Wavelength of transmitted RF field, in same physical unit as D

It should be emphasized that this simple path loss formula only applies to completely clear, empty space where the only mechanism of signal attenuation is the natural spreading of radio waves as they radiate away from the transmitting antenna. Path loss will be significantly greater if any objects or other obstructions lie between the transmitting and receiving antennas.

[3]The inverse square law applies to any form of radiation that spreads from a point-source. In any such scenario, the intensity of the radiation received by an object from the point-source diminishes with the square of the distance from that source, simply because the rest of the radiated energy misses that target and goes elsewhere in space. This is why the path loss formula begins with a -20 multiplier rather than -10 as is customary for decibel calculations: given the fact that the inverse square law tells us path loss is proportional to the square of distance (D^2), there is a "hidden" second power in the formula. Following the logarithmic identity that exponents may be moved to the front of the logarithm function as multipliers, this means what would normally be a -10 multiplier turns into -20 and we are left with D rather than D^2 in the fraction.

[4]In full disclosure, the degree to which electromagnetic radiation spreads out in space as a function of distance from the source – i.e. the inverse square law – has absolutely nothing to do with the frequency of that radiation. However, there is another factor called *antenna capture area* which does vary with the square of the wavelength, longer wavelengths resulting in a greater effective capture area for the receiving antenna than shorter wavelengths. In an effort to reduce the number of individual calculations needed for a link budget, wavelength is included in the path loss equation to account for this effect of wavelength on antenna capture area. One might choose to think of the "free-space path loss equation" as the "free-space *coupling* loss equation" because it considers not only the inverse-square law as a function of distance but also antenna capture as a function of wavelength.

This same spreading effect also accounts for "fade" – also known as *multipath-induced fading* – where radio waves taking different paths deconstructively interfere with each other, resulting in weakened signals at some locations but not in others equidistant from the transmitter. You may have personally experienced fade while driving a vehicle over long distances and listening to an analog (AM or FM) radio: sometimes a particular radio station's signal will "fade out" as you drive and then "fade in" again while driving in the same direction, for no obvious reason (e.g. no immediate obstructions to the signal). This is due to radio waves from the station emanating in all directions, then reflecting off of large objects and/or ionized regions high in the earth's atmosphere. There will inevitably be locations around that station where the incident wave from the transmitting antenna destructively interferes with those reflected waves, the result being regions of "dead" space where the signal is much weaker than one would expect from path loss alone:

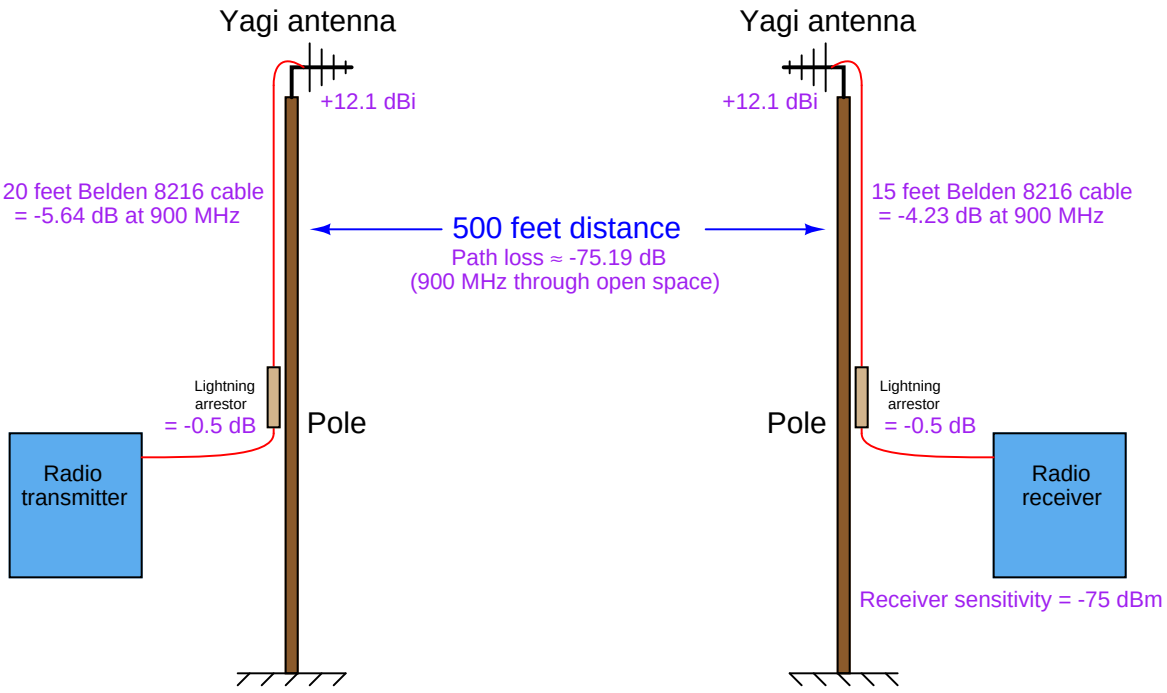


This is not unlike the phenomenon of certain seats in a concert hall having better acoustical properties than others, even when equidistant from the stage. Sound waves don't just travel straight from the performers to the audience members; they also reflect off of the concert hall's wall surfaces, and as a result there will be some seats where there is deconstructive interference between the direct sound wave (straight from the performer) and the reflected sound wave (bouncing off the wall) resulting in less volume at that particular seat location. Moving just a few seats away will avoid this "fade zone" and result in better sound quality.

Fade is a far more difficult factor to predict than path loss because of the complexity of modeling the interference patterns between all possible travel pathways from transmitter to receiver, and so generally radio system designers combat this loss by including a generous margin[5] of excess transmitter power to overcome the effects of fade. This *fade margin* is typically 20 dB to 30 dB, although it can be greater in scenarios where there are many signal paths due to reflections.

[5] "Margin" is the professionally accepted term to express extra allowance provided to compensate for unknowns. A more colorful phrase often used in the field to describe the same thing is *fudge factor*.

To illustrate, we will calculate the RF link budget for a 900 MHz radio transmitter/receiver pair directionally oriented toward each other with Yagi antennas. All sources of signal gain and loss will be accounted for, including the "path loss" of the RF energy as it travels through open air. The gains and losses of all elements are shown in the following illustration:



The path loss value shown in the illustration is a calculated function of the 900 MHz wavelength ($\lambda = \frac{c}{f} = 0.3331$ meters) and the distance between antennas (500 feet = 152.4 meters), assuming a completely obstruction-free path between antennas:

$$L_p = -20 \log \left(\frac{4\pi(152.4)}{0.3331} \right) = -75.19 \text{ dB}$$

According to the receiver manufacturer's specifications, the receiver in this system has a sensitivity of -75 dBm, which means our transmitter must be powerful enough to deliver an RF signal at least as strong as -75 dBm at the receiver's input connector in order to reliably communicate data. Inserting this receiver sensitivity figure into our RF link budget formula:

$$P_{tx(min)} = P_{rx(min)} - (G_{total} + L_{total})$$

$$P_{tx(min)} = -75 \text{ dBm} - (G_{total} + L_{total})$$

Now we need to tally all the gains and losses between the transmitter and the receiver. I strongly recommend making a table (as shown below) to hold all the different gain and loss values, as it becomes easy to miss or wrongly apply one of the many decibel values required for a full link budget calculation if these values are not neatly organized in one place. We will use a value of −20 dB for fade margin (i.e. our budget will leave room for up to 20 dB of power loss due to the effects of fade):

Gain or Loss	**Decibel value**
Transmitter cable loss	−5.64 dB
Transmitter arrestor loss	−0.5 dB
Transmitter antenna gain	+12.1 dBi
Path loss	−75.19 dB
Fade margin	−20 dB
Receiver antenna gain	+12.1 dBi
Receiver arrestor loss	−0.5 dB
Receiver cable loss	−4.23 dB
$G_{total} + L_{total}$	**−81.86 dB**

Inserting the decibel sum of all gains and losses into our RF link budget formula:

$$P_{tx(min)} = -75 \text{ dBm} - (-81.86 \text{ dB})$$

$$P_{tx(min)} = 6.86 \text{ dBm}$$

Converting a dBm value into milliWatts of RF power means we must manipulate the dBm power formula to solve for P_{mW}:

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

$$P_{mW} = 1 \text{ mW} \times 10^{\left(\frac{P_{dBm}}{10}\right)}$$

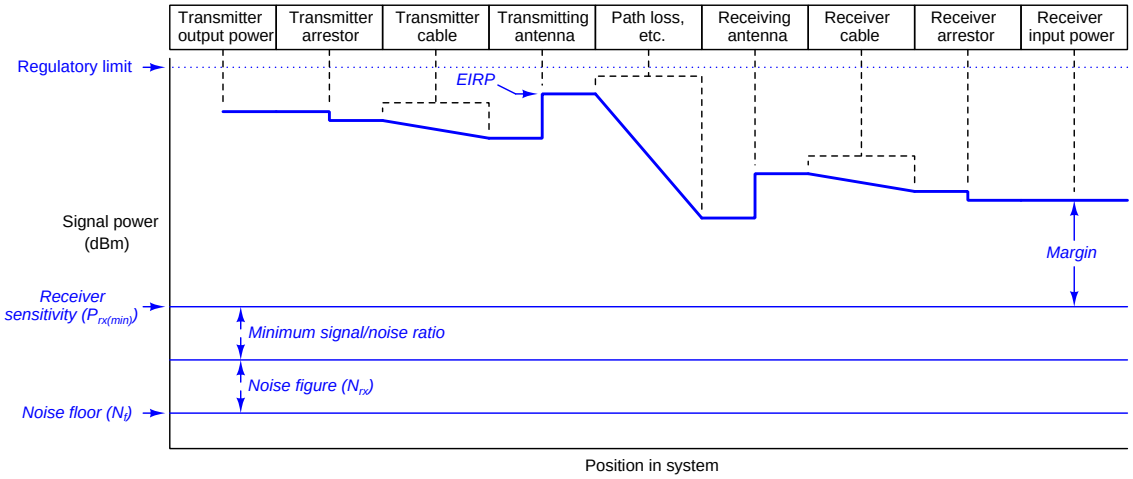
$$P_{tx} = 1 \text{ mW} \times 10^{\left(\frac{6.86}{10}\right)} = 4.85 \text{ milliWatts}$$

At this point we would do well to take stock of our assumptions. Power gains and losses inherent to the components (cables, arrestors, antennas) are quite certain because these are tested components, so we should not doubt these figures. Less certain are the environmental factors: noise floor can change, path loss *will* differ from our calculation if there is any obstruction near the signal path or under certain weather conditions (e.g. rain or snow dissipating RF energy), and fade loss is known to change dynamically as moving objects (people, vehicles) pass anywhere between the transmitter or receiver antennas. Our RF link budget calculation is really only a *rough estimate* of the transmitter power needed to get the job done.

How then do we improve our odds of building a reliable system? One way is to over-build it by equipping the transmitter with more power than the most pessimistic link budget predicts. However, this can cause other problems such as interference with nearby electronic systems if we are not careful. A preferable method is to conduct a *site test* where the real equipment is set up in the field and tested to ensure adequate received signal strength.

3.3 Link budget graph

Many of the concepts previously explored may be represented in a single graph[6], showing RF signal strength as a function of physical position within the RF link (from transmitter to receiver). The horizontal axis of this graph represents the position along the link path from transmitter to receiver, while the vertical axis represents RF signal strength in dBm:



At the very bottom of this graph we see the noise floor, representing natural RF noise that is unavoidable. Above that we see the added noise figure and signal/noise ratios summing up to a higher line representing the receiver's sensitivity, which is the minimum amount of signal power it must receive in order to reliably function.

Near the top of this graph we see a series of thicker line segments representing the various losses and gains within the system. Single-point losses such as lightning arrestors appear as vertical downward steps, while progressive losses such as cables and path loss appear as downward slopes. Single-point gains such as antennas appear as vertical upward steps. Note where EIRP appears on this graph: the amount of signal power at the transmitting antenna's output, factoring in the antenna's gain as well as any losses between the transmitter and the transmitting antenna. As previously mentioned, the Federal Communications Commission (FCC) sets regulatory limits for EIRP in the United States, because if only the transmitter's output power were regulated, it would be possible to thwart that limit using high-gain antennas.

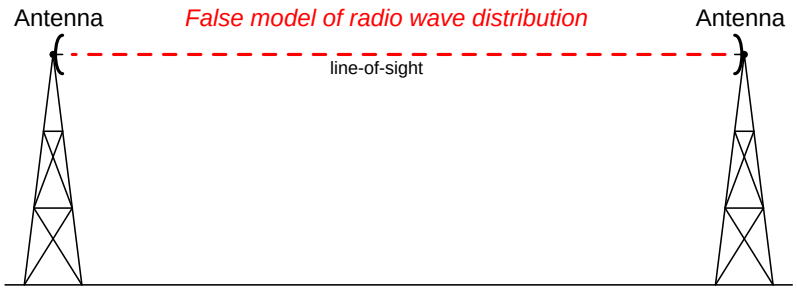
At the far right-hand end of the graph, we see the difference between the receiver's input signal power and the receiver's sensitivity as a *margin* for the link budget. This margin must exist in order to grant the system tolerance to unexpected power losses such as those resulting from inclement weather, interference from objects within the link path, cable deterioration, coupling corrosion, increases in noise floor, etc.

[6] I am indebted to Eric McCollum, Kei Hao, Shankar V. Achanta, Jeremy Blair, and David Kechalo for presenting this form of diagram in a technical paper presented at the 45th annual Western Protective Relay Conference in Spokane, Washington in October of 2018. I do not know if these authors are responsible for the invention of this form of graph, but it was certainly the first time I encountered one like it, and it so clearly showed all the fundamental quantities of an RF link budget that I had to include something similar in my tutorial!

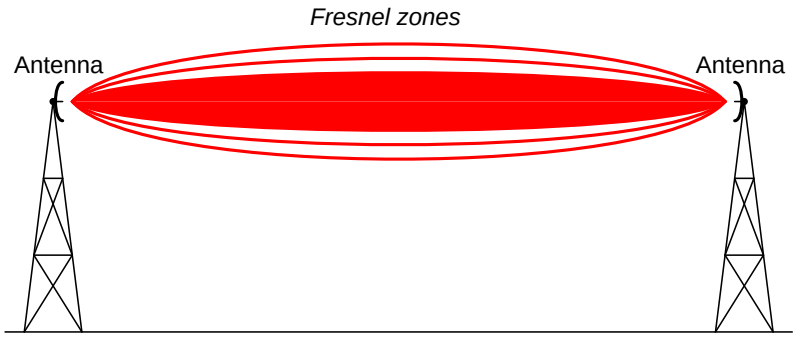
3.4 Fresnel zones

One of the many factors affecting RF link power transfer from transmitter to receiver is the openness of the signal path between the transmitting and receiving antennas. As shown in the previous section, *path loss* is relatively simple to calculate given the assumption of totally empty space between the two antennas. With no obstacles in between, path loss is simply a consequence of wave dispersion (spreading). Unless we are calculating an RF link budget between two airplanes or two spacecraft, though, there really is no such thing as totally empty space between the transmitting and receiving antennas. Buildings, trees, vehicles, and even the ground are all objects potentially disrupting what would otherwise be completely open space between transmitter and receiver.

A common expression in microwave radio communication is *line-of-sight*, or *LoS*. The very wording of this phrase evokes an image of being able to see a straight path between transmitting and receiving antennas. Obviously, if one does not have a clear line-of-sight path between antennas, the signal path loss will definitely be more severe than through open space. However, a common error is in thinking that the mere existence of an unobstructed straight line between antennas is all that is needed for unhindered communication, when in fact nothing could be further from the truth:



In fact, the free space necessary to convey energy in electromagnetic wave form takes on the form of football-shaped zones: the first one solid followed by annular (hollow) zones concentrically surrounding the first. These elliptical volumes are called *Fresnel zones*:



The precise shapes of these Fresnel zones are a function of wavelength and distance between antennas, *not* the size or configurations of the antennas themselves. In other words, you cannot change the Fresnel zone sizes or shapes by merely altering antenna types. This is because the

Fresnel zones do not actually map the distribution of electromagnetic fields, but rather map the free space we need to keep clear between antennas[7].

If any object protrudes at all into any Fresnel zone, it diminishes the signal power communicated between the two antennas. In microwave communication (GHz frequency range), the inner-most Fresnel zone carries most of the power, and is therefore the most important from the perspective of interference. Keeping the inner-most Fresnel zone absolutely clear of interference is essential for maintaining ideal path-loss characteristics (equivalent to open-space). A common rule followed by microwave system designers is to try to maintain an inner Fresnel zone that is at least 60% free of obstruction.

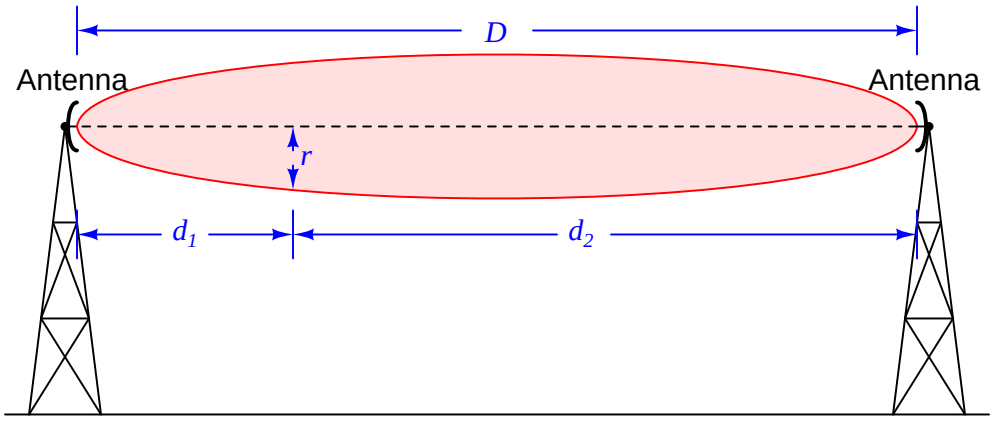
In order to design a system with this goal in mind, we need to have some way of calculating the width of that Fresnel zone. Fortunately, this is easily done with the following formula:

$$r = \sqrt{\frac{n\lambda d_1 d_2}{D}}$$

Where,

r = Radius of Fresnel zone at the point of interest
n = Fresnel zone number (an integer value, with 1 representing the first zone)
d_1 = Distance between one antenna and the point of interest
d_2 = Distance between the other antenna and the point of interest
D = Distance between both antennas
λ = Wavelength of transmitted RF field, in same physical unit as D, d_1, and d_2

Note: the units of measurement in this formula may be any unit of length, so long as they are all the same unit.



[7]The physics of Fresnel zones is highly non-intuitive, rooted in the wave-nature of electromagnetic radiation. It should be plain to see, though, that Fresnel zones cannot describe the actual electromagnetic field pattern between two antennas, because we know waves tend to spread out over space while Fresnel zones converge at each end. Likewise, Fresnel zones vary in size according to the distance between two antennas which we know radiation field patterns do not. It is more accurate to think of Fresnel zones as *keep-clear* areas necessary for reliable communication between two or more antennas rather than actual field patterns.

To illustrate by example, let us calculate the radius of the first Fresnel zone for two microwave antennas operating at 2.4 GHz, separated by one mile (1.609 km), at the widest portion of that zone. The wavelength of a 2.4 GHz signal is 0.1249 meters, from the formula $\lambda = \frac{c}{f}$. Distances d_1 and d_2 will both be equal to one-half of the total distance, since the widest portion of the Fresnel zone will be exactly at its mid-point between the two antennas ($d_1 = d_2 = 804.7$ meters). Here, we solve for r as follows:

$$r = \sqrt{\frac{n\lambda d_1 d_2}{D}}$$

$$r = \sqrt{\frac{(1)(0.1249)(804.7^2)}{1609}}$$

$$r = 7.089 \text{ meters}$$

Consider for a moment the significance of this dimension. At the very least, it means the antennas must be mounted this high off the ground in order to avoid having the most important Fresnel zone contact the earth itself (assuming level terrain), not to mention any objects above ground level such as buildings, vehicles, or trees between the two antennas. Consider also that the Fresnel zone is football-shaped, and therefore this 7.089 meter radius extends horizontally from the centerline connecting both antennas as well as vertically. This means that in order for this Fresnel zone to be untouched, there must be a clear path *14.18 meters wide* in addition to the antennas being at least 7.089 meters off the ground! If we were to consider a 900 MHz signal – another common frequency for industrial wireless devices – the minimum height above ground would be 11.58 meters, and the minimum clear path width 23.16 meters!

As you can see, "line of sight" is not as simple as it may first appear.

Chapter 4

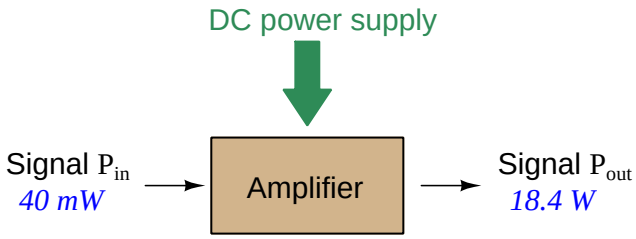
Derivations and Technical References

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

4.1 Decibels

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

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



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

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

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

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

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

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

The gain of our hypothetical electronic amplifier, therefore, would be more commonly expressed as 26.6276 dB rather than 2.66276 B, although either expression is technically valid[1].

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

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

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

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

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

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

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

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

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

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

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

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

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

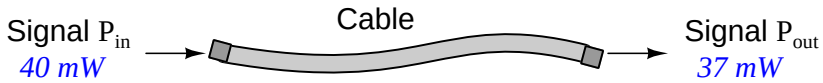
$$\text{Ratio} = 460$$

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

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

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

We will explore yet another application of signal power reduction as a case study for decibels: *cable loss*. Cables designed to convey signals over long distances are not perfect conduits of energy, as some of the signal's energy is inevitably lost along the way. This is true for different types of signals, electrical and optical being two popular examples. In the following illustration we see a signal cable losing power along its length[2], such that the power out is less than the power in:



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

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

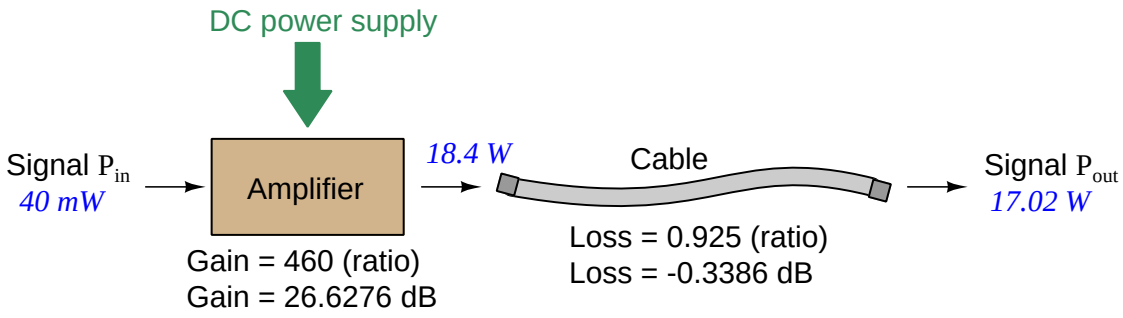
Contrasting this result against the previous result (with the amplifier) we see a very important property of decibel figures: any power *gain* is expressed as a *positive* decibel value, while any power *loss* is expressed as a *negative* decibel value. Any component outputting the exact same power as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

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

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

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

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



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

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

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

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

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

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

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

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

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

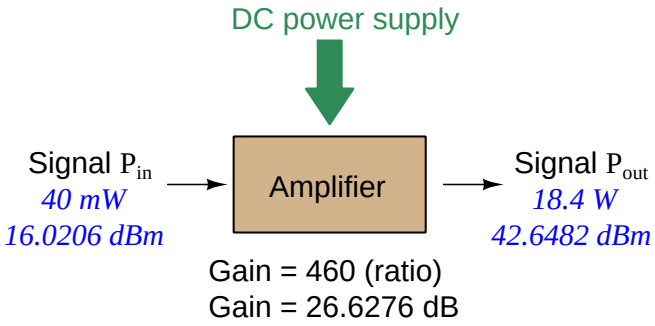
Decibels always represent comparisons of power, but that comparison need not always be P_{out}/P_{in} for a system component. We may also use decibels to express an amount of power compared to some standard reference. If, for example, we wished to express the input power to our hypothetical amplifier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

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

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

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

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



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

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

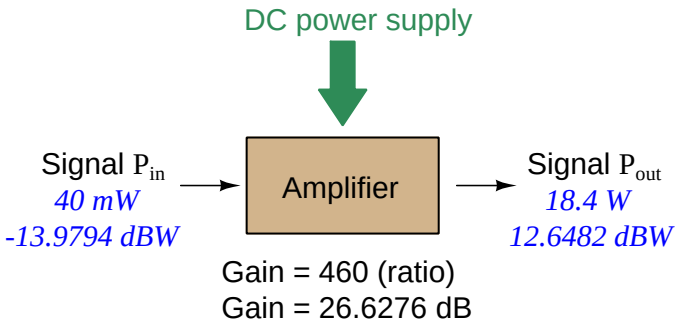
$$16.0206 \text{ dBm} + 26.6276 \text{ dB} = 42.6482 \text{ dBm}$$

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

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

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

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



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

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

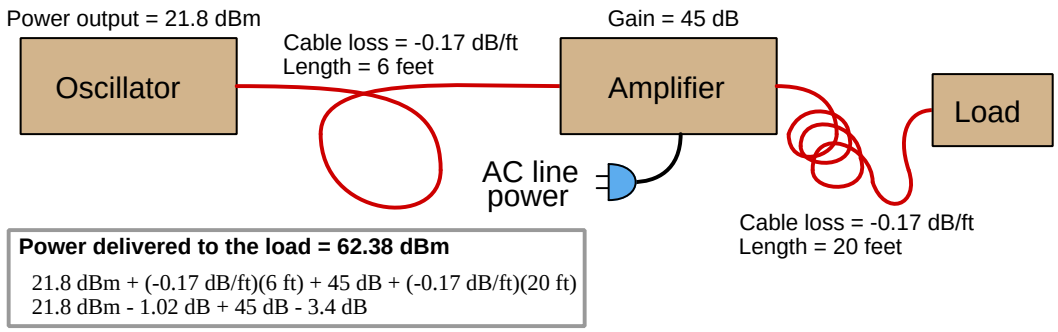
$$P_{in}(\text{dB}) + P_{gain}(\text{dB}) = P_{out}(\text{dB})$$

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

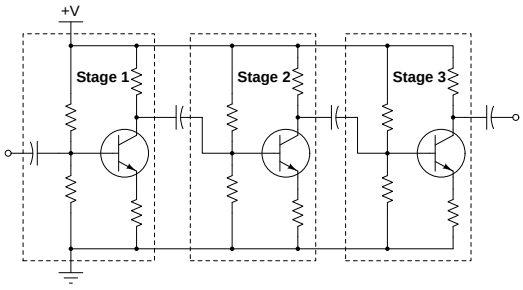
An electronic signal that begins 13.9794 dB less than 1 Watt, when boosted by an amplifier gain of 26.6276 dB, will become 12.6482 dB greater than the original reference power of 1 Watt.

This is one of the major benefits of using decibels to express powers: we may very easily calculate power gains and losses by summing a string of dB figures, each dB figure representing the power gain or power loss of a different system component. Normally, any compounding of *ratios* involves multiplication and/or division of those ratios, but with decibels we may simply add and subtract. One of the interesting mathematical properties of logarithms is that they "transform[3]" one type of problem into a simpler type: in this case, a problem of multiplying ratios into a (simpler) problem of adding decibel figures.

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



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



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

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

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

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

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

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

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

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

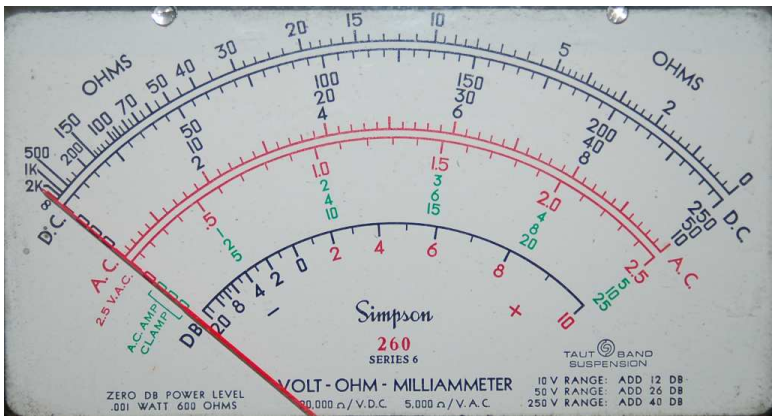
Therefore, we may simplify the decibel formula for voltage gain by removing the "2" power and making it a multiplier:

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

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

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

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



Pay attention to the note on decibels written in the lower-left corner of the meter face, where 0 dB is defined as 0.001 Watt dissipated by 600 Ohms. The fact that 0 dB is defined as 1 milliWatt means it should (properly) be labeled dBm rather than dB[4]. A load resistance value is necessary as part of this definition for dB because this meter cannot measure power directly but must infer signal power from measurements of AC *voltage*. Without a specific load resistance, there is no clear relation between voltage and power. 600 Ohms is an old telecommunications standard for audio-frequency AC signals, and continues to be used today for voltage-based decibel measurements of audio-frequency AC signals.

The meter as shown is connected to nothing at all, and so registers 0 Volts AC. This, of course, corresponds to zero power, and it has no corresponding decibel value because the logarithm of zero is mathematically *undefined*[5]. Practically, it means $-\infty$ dB, which is why the needle at the 0 Volt position "falls off" the left-hand end of the dB scale.

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

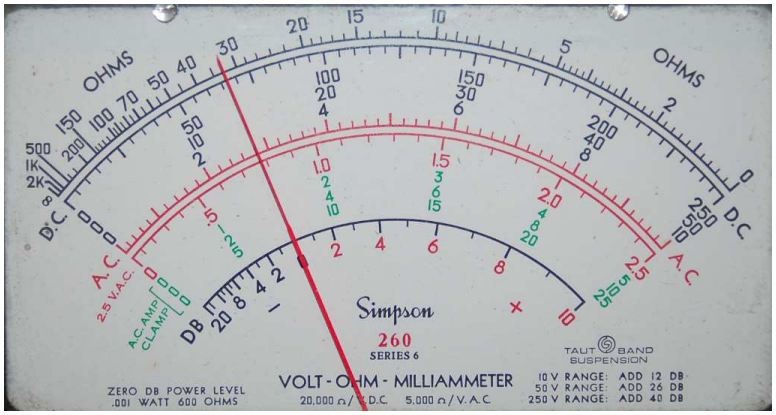
[4] Such mis-labeling is not that uncommon in the profession, the expectation being that the technician or engineer working with the instrument ought to be familiar enough with the concept of decibels to know when dB really means dBm, or dBW, etc.

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

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

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

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



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

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

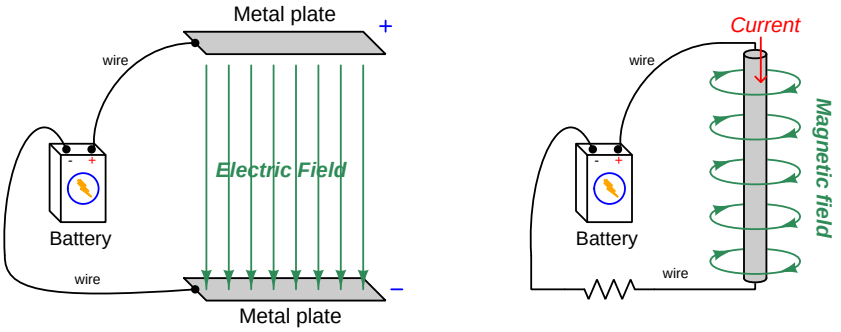
Therefore, when using the 10 VAC range instead of the 2.5 VAC range, one must add 12 dB to the reading. Likewise, we may prove each of the printed correction factors for the alternative voltage-measurement ranges listed (50 Volt AC range and 250 Volt AC range):

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

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

4.2 Near-field versus far-field regions

An electric field exists wherever voltage exists (i.e. wherever an imbalance exists between positive and negative electrical charges), parallel to the axis of that voltage. A magnetic field exists wherever an electric charge moves, perpendicular to the axis of that charge's motion. These phenomena exist for DC (direct) as well as AC (alternating) electricity, and are shown in the two following illustrations:



Electric and magnetic fields, however, are not simply manifestations of voltage and current, respectively. Mathematical discoveries made by the Scottish physicist James Clerk Maxwell (and later simplified by Oliver Heaviside) also relate electric fields and magnetic fields *directly to each other*. Consider two[6] of Maxwell's equations shown below:

$$\oint \vec{E} \cdot d\mathbf{s} = -\frac{d\Phi_B}{dt}$$

$$\oint \vec{B} \cdot d\mathbf{s} = \mu_0 I + \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}$$

The first equation describes how an electric field ($\vec{E}$) is created by a varying magnetic flux ($\frac{d\Phi_B}{dt}$), which is otherwise known as *Faraday's Law of Electromagnetic Induction*. This is the law exploited in the function of electro-mechanical generators whereby coils of wire are subjected to changing magnetic fields, creating voltage between the coils' end-points. The negative sign in this equation is an embodiment of *Lenz's Law*, which states that any current resulting from the induced voltage will produce its own magnetic field opposing the first magnetic field's direction of change.

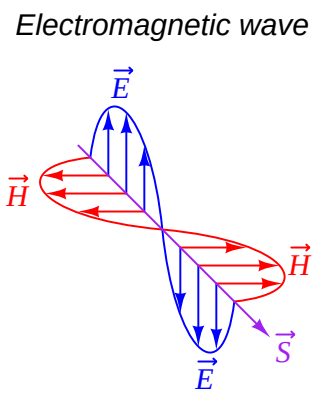
The second equation shown here describes *two* different ways to produce a magnetic field ($\vec{B}$). One way is to use a moving stream of electric charges known as a *current* (I), a fact also known as *Ampère's Law*. This is the law exploited in the function of electromagnets, where we produce a magnetic field by connecting a coil of wire to an electrical source. The second way is to vary an electric flux ($\frac{d\Phi_E}{dt}$) in empty space with surrounding magnetic permeability μ_0 and electric permittivity ϵ_0.

[6] Maxwell's equations of electromagnetism are typically presented as a set of *four*. Here we only list two of these, as the other two are not relevant to electromagnetic waves in particular.

Maxwell's equations suggest that in pulsed DC and AC circuits, where the resulting electric and/or magnetic fields vary over time ($\frac{d}{dt}$) rather than holding steady, those time-varying fields will mutually support one another over space and time. That is, a varying electric field will eventually produce a varying magnetic field, and vice-versa, those fields propagating one another as an *electromagnetic wave* radiating away from the circuit through space at the speed of light[7].

At close range, immediately adjacent to the circuit conductors, electric and magnetic fields are strictly functions of those conductors' voltages and currents, respectively. These fields' shapes follow the basic forms shown in the previous illustrations: magnetic field lines circling current-carrying conductors and electric field lines stretching between conductive surfaces. We generally refer to this region of space around an energized circuit as the *near-field region*, where electric and magnetic fields are distinct from one another in form and in relative magnitude.

At farther distances from the energized circuit, the electric and magnetic fields are found in pairs oscillating at right angles to one another in space, forming a single electromagnetic wave radiating away from the source circuit spherically in all directions. This form of wave is a direct consequence of Maxwell's equations describing how the rate-of-change of one field creates the other field, and vice-versa. Below we see a crude representation of an electromagnetic wave, with an electric $\vec{E}$ field, a magnetic $\vec{H}$ field, and the vector of propagation $\vec{S}$ called the *Poynting vector*:



This region where the electric and magnetic fields exist as coupled pairs comprising a well-formed wave is referred to as the *far-field region*. In this region the electric and magnetic fields are always perpendicular to one another and their relative magnitudes are in a fixed proportion governed by the permeability and permittivity of free space.

The demarcation between near- and far-field regions depends on the physical dimensions of the radiating circuit as well as the *wavelength* of the signal (λ), wavelength being calculated by dividing the speed of light (2.9979×10^8 meters per second in empty space) by the signal frequency in Hertz. There is no simple rule for predicting where the near-field region ends and the far-field region begins, and in fact there is a gradual morphing of one to the other, but one may reliably consider any distance from the circuit *in excess of multiple wavelengths* to be far-field.

[7] In fact, this is precisely what light is: an electromagnetic wave of exceptionally high frequency, well beyond the frequency range of contemporary electric circuits.

The following list compares and contrasts several near-field versus far-field characteristics:

- In the near-field region electric and magnetic field strengths depend greatly on the geometry of the energized circuit's conductors as well as on the specific voltage and current levels existing at each location. In the far-field region, however, the electric and magnetic field strengths are always in fixed proportion to one another.

- Far-field effects may be ignored for low-frequency AC circuits because these circuits' wavelengths are so long. At 60 Hz, for example, the wavelength is nearly 5 *million* meters. If we consider the far-field region of a circuit to begin at least multiple wavelengths from the source, this puts the far-field region of a 60 Hz circuit at least one-quarter of the Earth's circumference away from the circuit in question!

- In the near-field region the proportionality between electric and magnetic field strengths is a function of circuit impedance: high-impedance circuits will have stronger electric fields than magnetic fields, and low-impedance circuits will have stronger magnetic fields than electric fields, all other factors being equal. This is due to Ohm's Law establishing the relationship between voltage and current ($Z = \frac{V}{I}$), and how electric fields originate in a circuit from potential differences while magnetic fields originate in a circuit from currents. In the far-field region, however, where the electric and magnetic fields exist only through mutual support, their strengths are always in the same (fixed) proportion defined as the *characteristic impedance of free space* which is approximately 377 Ohms. This value is equal to the square root of the ratio of magnetic permeability to electric permittivity for free space ($Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$) because these parameters of space dictate just how strongly a varying electric field creates a magnetic field.

- The relationship between field strength and physical distance from the circuit is a very complex one within the near-field region, as this depends greatly on the geometry of the circuit conductors. In the far-field region, however, we find the *Inverse-Square Law* always holds true: as distance from the circuit increases, electromagnetic wave strength simply and reliably diminishes with the *square* of that distance as if the wave were radiating away in a spherical fashion from a point-source[8]. For example, if we double the distance from the radiating circuit (from a point in the far-field region to another point in the far-field region twice as far away in empty space), the power conveyed by the electromagnetic radiation will always be *four times* less. Tripling the far-field distance always weakens the wave's power by a factor of *nine*. This is simply because the area over which the radiated energy spreads increases with the square of the distance from any point-source, viewing that point-source as the center of a sphere.

- In the far-field region every radiating circuit has a definite *radiation pattern* where the strength of the radiated electromagnetic wave as a function of the *angle* from the geometric centerline follows a consistent pattern regardless of distance. This is a very important feature of *antennas*, where different antenna designs feature unique far-field radiation patterns. By contrast, in the near-field region of an antenna the radiation pattern varies significantly with distance.

[8] A "point source" is a hypothetical point in space having no height nor width nor depth (i.e. it has *zero* physical dimensions) that emits radiation. Point-sources are a theoretical concept only. Real sources have spatial dimensions which makes their near-field strength/distance relationships complex, but the farther away you get from them the more their radiative behavior approaches that of a theoretical point-source.

- The spatial *polarization* of the electromagnetic wave (i.e. the orientation of its perpendicular electric and magnetic fields) is always well-defined in the far-field region but often ill-defined in the near-field region. This means, among other things, that two or more antennas will exchange electromagnetic energy efficiently only if they are appropriately oriented to one another over far-field distances, but may exchange energy fairly well regardless of orientation over near-field distances.

- Any "gain" specifications for an antenna structure apply only to apparent gains in signal power over the far-field range, because they refer to the degree to which an antenna focuses its energy in one direction more than another (i.e. its directionality), implying a distance over which the electromagnetic wave has become well-formed and the antenna's radiation pattern is reliably established. At near-field distances these "gain" figures are meaningless.

4.3 The Poynting vector

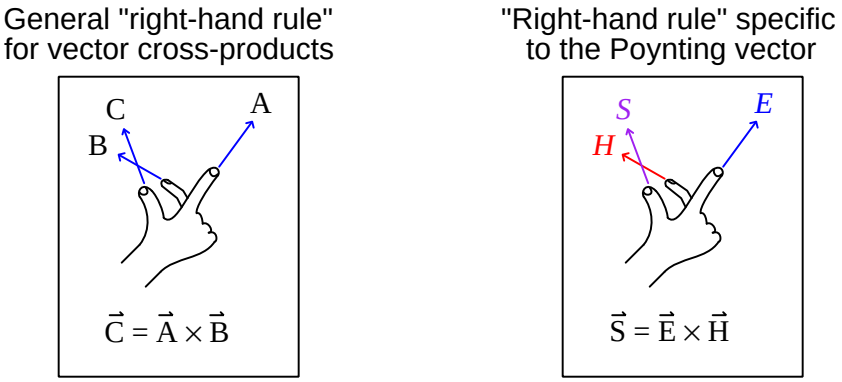
We know that electric fields (E) are always associated with a voltage (V), and magnetic fields (field strength B and field force H) with a current (I). Specifically, electric field strength is the quotient of voltage and distance between two points which is why electric fields are measured in units of *Volts per meter*. Similarly, magnetic field force around a straight current-carrying conductor is the quotient of current and circumference of the circular magnetic path (equal to 2π times the distance from the conductor) which is why magnetic field force is measured in units of *Amperes per meter*.

From Joule's Law we know that the rate of energy transfer (called power) whenever there is simultaneous voltage and current ($P = VI$). It stands to reason, then, that just as voltage multiplied by current yields power, an electric field multiplied by a magnetic field must similarly express power transferred by those two fields through space. That is, if $P = VI$, then the product of E and H should yield a power *intensity* which we shall label S:

$$P = VI \qquad \text{[Watts]} = \text{[Volts][Amperes]}$$

$$S = EH \qquad \text{[Watts/meter}^2\text{]} = \text{[Volts/meter][Amperes/meter]}$$

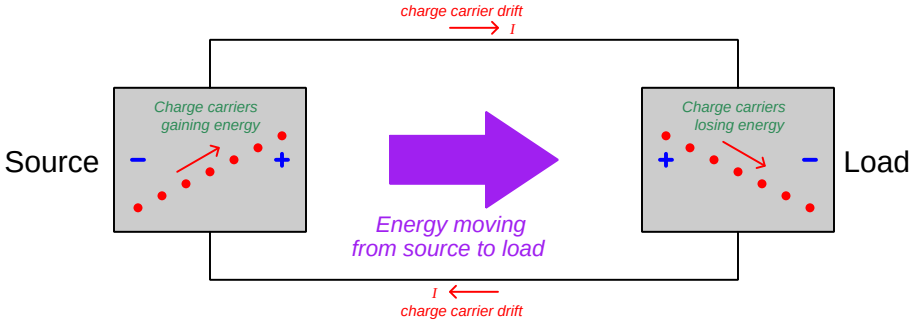
Notice how the Joule's Law equation $P = VI$ involves scalar quantities, while the $S = EH$ equation necessarily involves spatial dimensions of distance (meters) and area (square meters). This is because both electric and magnetic fields are *vector* quantities having both magnitude and spatial direction. Power intensity (S) is the measure of how much power is flowing per square meter of area in open space. As it so happens, power intensity in Watts per square meter relates to electric and magnetic fields by the right-hand rule of vector cross-products, the vector expression of that power intensity being called the *Poynting vector* ($\vec{S}$) after its formulator, the English physicist John H. Poynting:



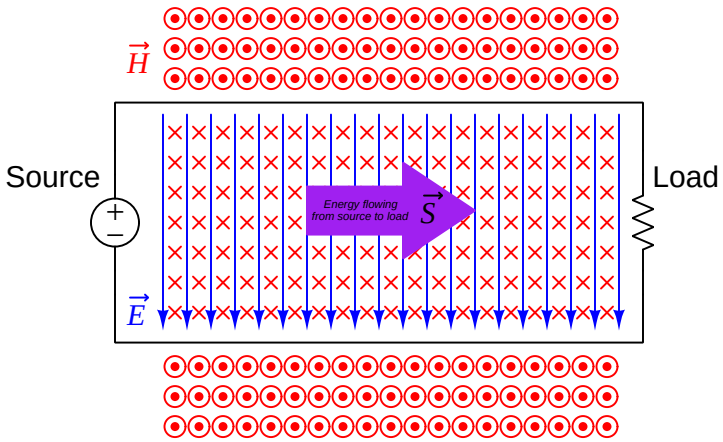
Poynting's contribution to the science of electromagnetism was expressing energy conservation and energy transfer in terms of electric and magnetic fields. That is, the electric and magnetic fields associated with any electrical *source* result in a flow of energy away from that source *through space*, and the electric and magnetic fields associated with any *load* result in a flow of energy toward that load *through space*. The direction of the Poynting vector shows the direction of that energy flow,

and the Poynting vector magnitude shows how many Watts of power flow per square meter of space.

This should be a rather startling conclusion to anyone familiar with voltage, current, and power in simple DC circuits. Beginning students of electricity regard energy transfer in simple circuits as the gain and loss of potential energy in mobile electric charges, those charges gaining energy as they pass through a source and losing energy as they pass through a load:

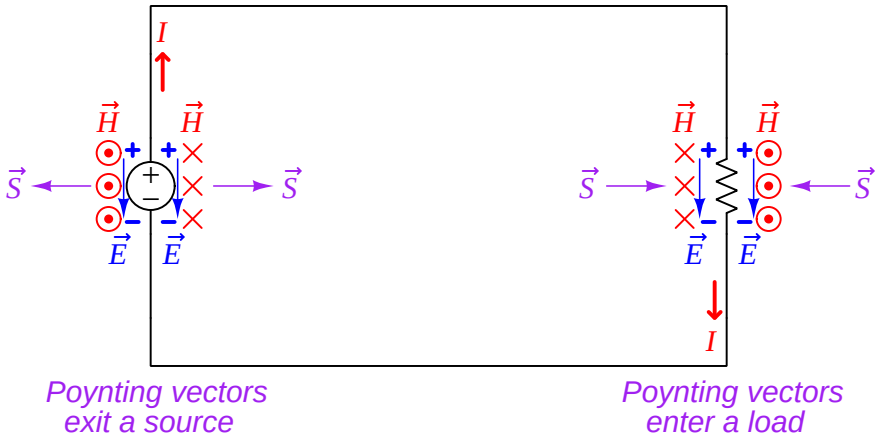


Though this model is conceptually correct, *so is Poynting's model of energy flowing through space*. In fact, if we re-draw this simple circuit illustration showing the electric and magnetic field vectors between the two wires, we can see the Poynting vector expressing that flow of energy from source to load[9]:



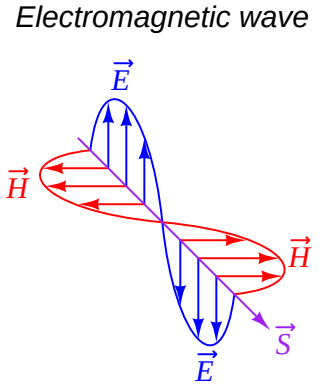
[9] Holding your right hand such that your index finger points downward ($\vec{E}$) and your middle finger into the page ($\vec{H}$), your thumb will point to the right ($\vec{S}$) from source to load.

If we examine the source and load separately using the right-hand rule, we see how the Poynting vector in each case shows the proper direction of energy transfer:



It is well-worth your time to apply the right-hand rule to each of the four Poynting vectors shown above, to familiarize yourself with the use of this tool as a means to determine vector cross-product directions. For example, to the left of the source we see the electric field vector pointing down which means your index finger should point downward as well. The magnetic field vector on that side is coming out of the page (middle finger pointing toward you), and holding your right hand in that manner points your thumb to the left.

For simple DC circuits and circuit elements, Poynting's theory of power transfer may seem like a mere curiosity. It fails to explain anything we could not already explain using the model of potential energy gained and lost by mobile charge carriers, and it requires contortions of your right hand to properly align all the vectors. However, in some applications such as *radio* where energy radiates away from structures called *antennas* in the form of electromagnetic waves (oscillating electric and magnetic fields), Poynting's theory is the *only* suitable explanation for power transfer because there are no wires to convey mobile electric charges from source (transmitting antenna) to load (receiving antenna). Below we see a simplified representation of an electromagnetic wave consisting of two oscillating fields (shown here as an electric field oscillating vertically and a magnetic field oscillating horizontally), and if you use the right-hand rule in each half-cycle you will find that the Poynting vector consistently points in the same direction at the speed of light:



Based on the direction of the Poynting vector we see in this illustration, the source (transmitting antenna) would be located somewhere above and to the left of the drawn wave. There is no definite location for a load, as electromagnetic waves spread out in *all directions* away from the source. Our simple illustration merely shows a "cross-section" of a wave radiating *cylindrically* away from the transmitting antenna. The real electromagnetic wave (of which this illustration shows just a portion) would have an electric field resembling ripples on a pond oscillating vertically while radiating away in all horizontal directions from a single point, and have a magnetic field resembling a set of concentric circles oscillating rotationally in the horizontal plane while also radiating in all directions from that same point.

4.4 Maxwell's equations of electromagnetism

The Scottish physicist James Clerk Maxwell mathematically proved the interrelationships of electric and magnetic fields, publishing his results in the book *A Treatise on Electricity and Magnetism* in 1873. His equations were later simplified and consolidated by the work of Oliver Heaviside, commonly expressed in the following forms describing electric fields (**E**) and magnetic fields (**B**) within a vacuum:

$$\oint \mathbf{E} \cdot d\mathbf{A} = \frac{Q}{\epsilon_0}$$

$$\oint \mathbf{B} \cdot d\mathbf{A} = 0$$

$$\oint \mathbf{E} \cdot d\mathbf{s} = -\frac{d\Phi_B}{dt}$$

$$\oint \mathbf{B} \cdot d\mathbf{s} = \mu_0 I + \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}$$

We will now explore each of these four equations, relating them to laws codified by other early electrical researchers, as well as describe how they relate to the phenomenon of *electromagnetic waves*.

4.4.1 Gauss's Law of electric fields

$$\oint \mathbf{E} \cdot d\mathbf{A} = \frac{Q}{\epsilon_0}$$

The first of these equations simply describes electric fields as spreading out from a quantity of electric charge (Q). If we imagine a three-dimensional surface completely enclosing an electric charge (e.g. a soap bubble surrounding a charge), the total quantity of electric flux passing through that surface will be the product of the electric field (also known as the electric flux density, **E**) and the enclosing area (**A**). This is also known as *Gauss's Law of electric fields*.

4.4.2 Gauss's Law of magnetic fields

$$\oint \mathbf{B} \cdot d\mathbf{A} = 0$$

The second of these equations describes magnetic fields passing through a three-dimensional enclosing surface. If we imagine a magnet completely enclosed within a soap bubble, the sum total of magnetic flux passing out through that soap bubble's surface must equal the sum total of magnetic flux re-entering it somewhere else. Literally, this equation says the product of magnetic field (also known as magnetic flux density **B**) and the enclosing area (**A**) must be algebraically equal to zero. This is to say, magnetic fields must always exist as *closed loops*, or alternatively we may say there is no such thing as a magnetic monopole (i.e. a solitary North or South pole without its complement, or a *magnetic charge*). In this way magnetic fields are quite different from electric fields which radiate away from discrete electric charges without ever returning. This is also known as *Gauss's Law of magnetic fields*.

4.4.3 Faraday's Law of electromagnetic induction, extended

$$\oint \mathbf{E} \cdot d\mathbf{s} = -\frac{d\Phi_B}{dt}$$

The third equation states that an electric field ($\mathbf{E}$) will be produced in open space by a varying magnetic flux $\left(\frac{d\Phi_B}{dt}\right)$. Michael Faraday expressed this phenomenon in terms of voltage induced along the length of a conductor by a time-varying magnetic field, but Maxwell demonstrated mathematically that this same thing happens even in empty space!

4.4.4 Ampère's Circuital Law, extended

$$\oint \mathbf{B} \cdot d\mathbf{s} = \mu_0 I + \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}$$

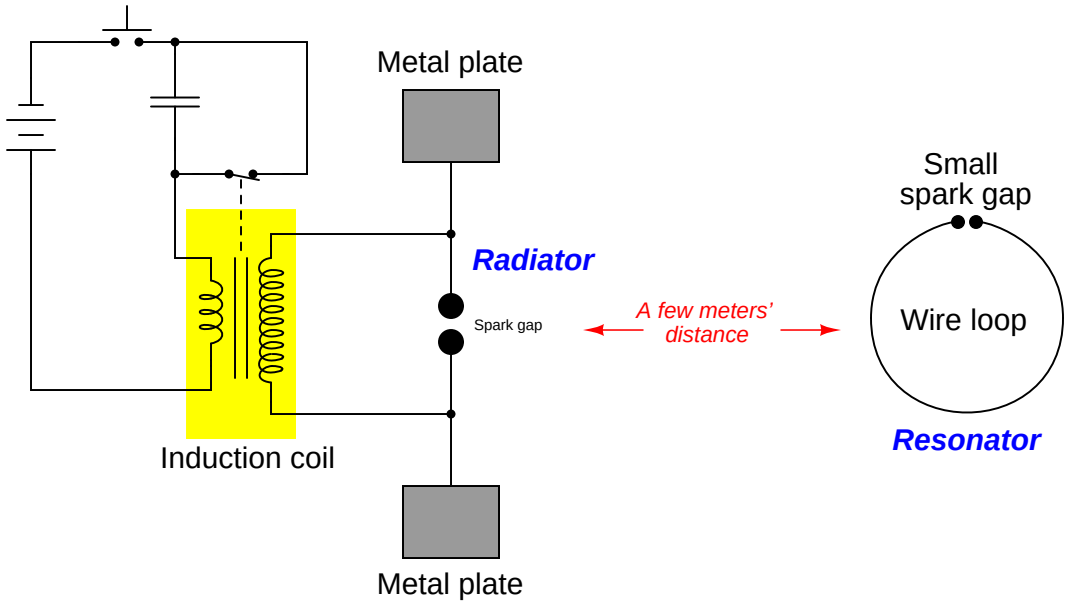
The fourth equation states than a magnetic field ($\mathbf{B}$) will be produced in open space either by an electric current (I) and/or by a varying electric flux $\left(\frac{d\Phi_E}{dt}\right)$. The French scientist André-Marie Ampère described the relationship between current passing through a conductor and the magnetic field created around that conductor by the current, but once again Maxwell extended this concept to include magnetic fields generated in empty space. In empty space, where there are no moving electric charges (i.e. current, I) to generate a magnetic field, only the second term of the equation applies with a time-varying electric field substituting for actual electric current.

4.4.5 Proof of electromagnetic waves

Given the complementary relationship between time-varying electric and magnetic fields, Maxwell reasoned, it was possible for a varying electric field to create a varying magnetic field which would then create another varying electric field, and so on. This cause-and-effect cycle could continue, ad infinitum, with fast-changing electric and magnetic fields radiating off into open space without needing wires to carry or guide them, or a tank circuit to sustain the oscillations. In other words, *these complementary fields self-sustain as they travel*, although they do become weaker due to their spreading out of a finite amount of energy over larger and larger volumes of space.

The Prussian Academy of Science offered a reward to anyone who could experimentally validate Maxwell's prediction, and this challenge was answered by Professor Heinrich Hertz at the Engineering College in Karlsruhe, Germany in 1887, eight years after Maxwell's death. Hertz constructed and tested a pair of devices: a "radiator" to produce the electromagnetic waves, and a "resonator" to receive them.

A simplified diagram showing Hertz's experimental device is shown here:



An "induction coil" (a buzzing device constructed of a self-interrupting relay and step-up transformer winding to generate a continuous pulsing waveform of high voltage) provided an extremely noisy (i.e. harmonically rich)[10] AC signal to the radiator, while a spark gap at the

[10] A fundamental property of waves is that any wave-shape, no matter how complex, is mathematically equivalent to a series of perfectly sinusoidal waves of differing frequencies added together. Waveforms with fast rise/fall times, such as the pulse waveforms produced by switched induction coils, contain many such high-frequency "harmonic" signals, and it is these high-frequency signals that created the radiated electromagnetic waves in Hertz's experiments. The induction coil's secondary winding inductance combined with the parasitic capacitance of the radiator's wires and plates provided the means to momentarily sustain some of these high-frequency oscillations long enough to sustain visible sparks at the resonator's gap.

resonator provided indication that the electromagnetic waves were captured and converted into voltage by the resonator wire.

Both the radiator and the resonator are antennas. The purpose of the transmitting antenna (radiator) is to take high-frequency AC power and radiate that power in the form of electromagnetic waves: self-sustaining electric and magnetic fields propagating outward into open space. The purpose of the receiving antenna is to capture those electromagnetic waves and convert them into an electrical signal (i.e. AC voltage and current). All antennas – from historical to the most modern – behave in fundamentally the same way: energize them with high-frequency AC electrical power, and they will radiate electromagnetic waves at that frequency; expose them to electromagnetic waves, and they will produce a very small AC electrical signal at the same frequency as the incident radiation.

4.5 Estimating RF field strength

An important measurement of radio-frequency (RF) source emissions is *electric field strength*, often specified at a standard distance of 3 meters from the radiating antenna. This field strength must not exceed certain regulatory limits set by the FCC (Federal Communications Commission) for devices operated within the United States. Electric field strength may be empirically determined using a field strength meter, but it may also be estimated mathematically based on transmitter power, distance, and antenna gain.

All electronics students should be familiar with Joule's Law relating voltage and resistance to power:

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

Where,
P = Power in Watts
V = Voltage in Volts
R = Resistance in Ohms

This same basic concept applies to RF power propagating through open space, where the impedance of free space is approximately 120π Ohms (roughly 377 Ohms) and E represents the electric field strength in Volts per meter rather than plain voltage in Volts:

$$P = \frac{E^2}{Z}$$

Where,
P = Power in Watts at a given location
E = Electric field strength at a given location in Volts per meter (V/m)
Z = Impedance of free space in Ohms ($\approx$ 377 Ohms)

We may expand this equation to include the gain of the antenna as a ratio (not as a dB figure), as well as distance between the antenna and the field measurement point in meters:

$$\frac{PG}{4\pi D^2} = \frac{E^2}{Z}$$

Where,
P = Power in Watts at a given location
G = Antenna gain, as a plain ratio (isotropic antenna = 1.0)
D = Distance between radiating antenna and electric field measurement point in meters
E = Electric field strength at distance D from antenna in Volts per meter (V/m)
Z = Impedance of free space in Ohms ($\approx$ 377 Ohms)

The expression $4\pi D^2$ comes from the formula for the surface area of a sphere. A (theoretical) isotropic antenna radiates spherically in all directions, and so the weakening of power at any given point on the surface of that sphere is inversely proportional to the square of the distance from the point-source of the RF energy.

Chapter 5

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?

5.1 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.

5.1.1 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.

5.1.2 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

Simplification as a problem-solving strategy

Graphing as a problem-solving strategy

Algebraic substitution

Electromagnetic wave

Antenna

Frequency

Wavelength

Directionality

Dipole

Decibels

Isotropic

Gain

Margin

ERP

Noise

Frequency domain

Signal-to-noise ratio

BER

Path loss

Inverse square law

Fade loss

Site test

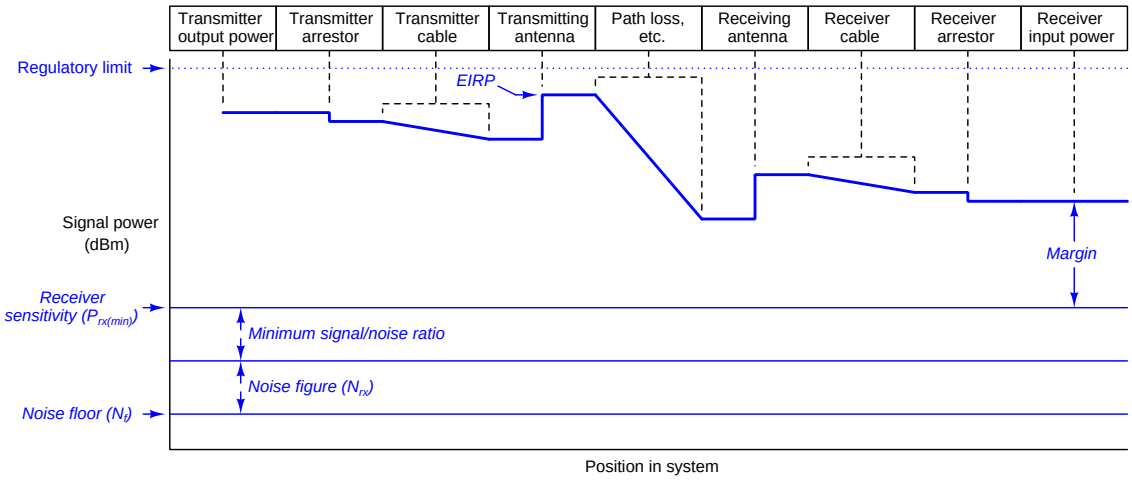
Fresnel zone

Line-of-sight

5.1.3 Link budget graph

Show where the following parameters (or parameters associated with these components) fit into the following link budget graph. Note that this is a generic graph, and is not necessarily tailored to this particular application:

- BER (bit error rate) = 10^{-6}
- Distance = 3 miles
- EAN2400NR 1/2 wave whip transmitting antenna
- ASC0032SF transmitting cable (3 feet long)
- EAN2414CR Corner Reflector receiving antenna
- ASC0102SN receiving cable (10 feet long)



Challenges

- Explain why the *bit error rate* is an important factor in this calculation. If we desire a lower (smaller) bit error rate, what must change?
- Define "sensitivity" for a radio receiver unit.
- Explain why one should always incorporate some *fade loss* into the minimum transmitter power calculation for any realistic radio application.
- After calculating the minimum transmitter power for this application, calculate the amount of margin (for fade and other losses) based on this particular transmitter's rated power output.

5.1.4 Cable loss and frequency

Why do RF signal cables experience energy dissipation (in the form of heat) when conducting RF signals? What is/are the fundamental principle(s) at work here?

What effect, if any, does a change in signal frequency have on power dissipation for a coaxial cable? Why?

Challenges

- What differences in cable design mitigate these losses?

5.2 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.

5.2.1 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.

5.2.2 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.*

5.2.3 Power ratio and dB conversions

Convert between the following power ratio and decibel values:

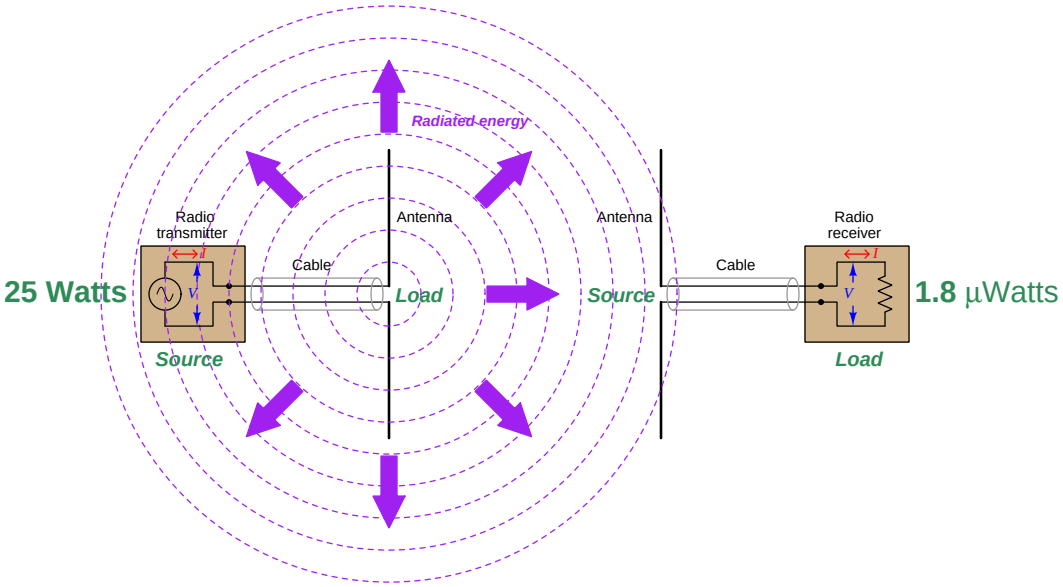
- -21 dB = power ratio gain of
- Power ratio gain of 20 =
- $+50$ dB = power ratio gain of
- Power ratio reduction of $\frac{1}{40}$ =
- Power ratio gain of 1600 =
- $+26$ dB = power ratio gain of
- Power ratio reduction of $\frac{1}{800}$ =
- -33 dB = power ratio gain of

Challenges

- What advantage(s) is/are there to using decibels to express power gains/losses?
- Estimate some of these answers using mental math, knowing that $+10$ dB = $\times$ 10 ratio, -10 dB = $\times$ $\frac{1}{10}$ ratio, $+3$ dB = $\times$ 2 ratio, and -3 dB = $\times$ $\frac{1}{2}$ ratio.
- What would it mean if any of the dB figures had an “m” or a “W” suffix attached?

5.2.4 Power and loss calculations

Suppose a radio transmitter unit outputs 25 Watts of RF power at a frequency of 900 MHz, into a half-wave dipole antenna. The radio energy radiates outward from the transmitting antenna, and is received by another antenna to send a signal of 1.8 microWatts into a radio receiver:



Calculate the transmitter's output power (P_{tx}) in both dBm and dBW, and also the receiver's input signal power (P_{rx}) in dBm and dBW. Then, calculate the total power loss between transmitter and receiver (P_{loss}), in dB.

P_{tx} = _____ dBm = _____ dBW

P_{rx} = _____ dBm = _____ dBW

P_{loss} = _____ dB

Explain in your own words the meaning of the following units: *dBm*, *dBW*, and *dB*. Also, explain why it is appropriate to use dBm or dBW to represent P_{tx} and/or P_{rx}, but not for P_{loss} (instead, P_{loss} must be represented in plain dB).

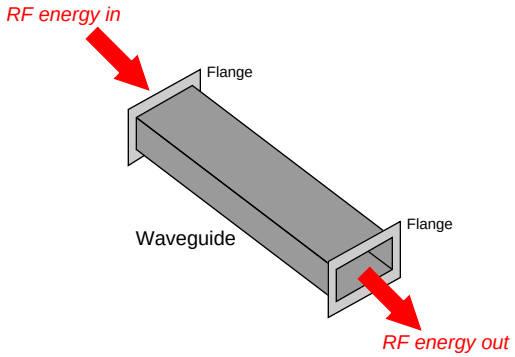
Challenges

- Explain why the "decibel" is such a commonly used unit for power, and also for power gains and losses, in RF communications work.
- Does it matter if P_{loss} is calculated from the transmit and receive powers in units of Watts, dBm, or dBW? Why or why not?

- If we wished to reduce the power loss from transmitter and receiver, how could we do so? Note: there are many different solutions to this problem!
- If the transmitting and receiving antennas were spaced farther apart, would P_{loss} increase, decrease, or remain the same?
- If the transmitting and receiving antennas were both located closer to ground level (i.e. not as high off the ground), would P_{loss} increase, decrease, or remain the same?
- Estimate these decibel values by first rounding the power ratio to the nearest product of 10 and/or 2, then applying the equivalence of 10 dB to a 10-fold ratio and 3 dB to a two-fold ratio.

5.2.5 Waveguide power loss

In many microwave radio systems, hollow rectangular tubes known as *waveguides* are often used in lieu of cables to serve as transmission lines between the transceiver and the antenna. These hollow metal tubes acts as "pipes" to convey GHz-range electromagnetic waves with great efficiency, far exceeding the efficiency of coaxial cables:



The following table shows a comparison between RG-49/U waveguide and RG-9/U flexible cable:

Parameter	RG-49/U waveguide	RG-9/U coaxial cable
External dimensions	2 inch $\times$ 1 inch	0.420 inch diameter
Dielectric	Air	Polystyrene plastic
Weight	1.4 pounds per foot	0.15 pound per foot
Attenuation at $f = 5$ GHz	0.011 dB per foot	0.23 dB per foot
Maximum power rating	1.2 MW	66 W

Suppose an RG-49/U waveguide is used to transfer RF energy from a 300 Watt radar transceiver to a dish antenna. Calculate the RF power (in Watts) available at the antenna, assuming the waveguide is 8 feet long and that there are no other power losses between the transceiver and the antenna.

Now suppose that radar transceiver's power is increased to 800 Watts. Calculate the RF power (in Watts) available at the antenna, assuming the same waveguide length.

Now suppose that radar transceiver's power is increased to 3 kiloWatts. Calculate the RF power (in Watts) available at the antenna, assuming the same waveguide length.

Does the waveguide power loss remain constant as transmitter power increases, or does it change with the amount of transmitted power? How does the "lost" RF energy manifest itself through the waveguide (given that the Law of Energy Conservation tells us energy cannot be created or destroyed)?

Challenges

- Explain why the waveguide shown has such a greater power rating than the coaxial cable, based on their respective dB/foot loss ratings.
- Why aren't waveguides used for all GHz-frequency radio applications, given that their power losses are so much less than the losses exhibited by coaxial cable?
- For those who have studied level measurement technologies, explain how the concept of a *waveguide* relates to guided-wave radar (GWR) level transmitters.
- Is is possible for a waveguide to experience *reflected signals* as in the case of improperly terminated transmission line cables? Why or why not?

5.2.6 Path loss estimation

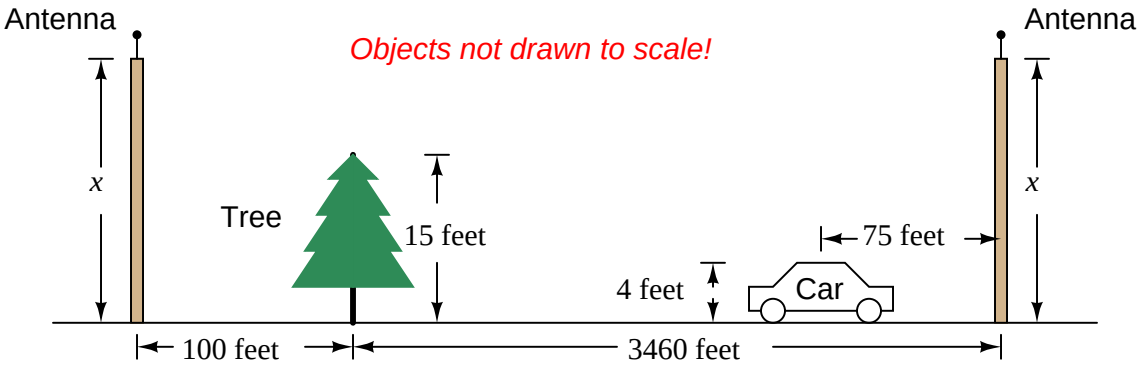
Suppose a radio transmitting antenna radiates 120 kW of RF power into space, and that only 30 milliWatts of this power reaches a receiving antenna some distance away. Calculate the path loss in decibels *without using a calculator*. Note that an addition or subtraction of 10 dB is exactly equal to a 10-fold multiplication or division ratio (respectively), and that an addition or subtraction of 3 dB is approximately equal to a two-fold multiplication or division ratio.

Challenges

- Explain what creates path loss. Is this something that can be overcome or at least mitigated by applying the correct technology? Why or why not?
- Explain what *fade loss* is, what creates it, and how it differs from path loss.
- Is path loss affected by the type of antenna used to transmit the RF energy? Explain why or why not.
- Is path loss affected by the type of antenna used to receive the RF energy? Explain why or why not.
- Is path loss affected by the frequency of the RF energy transmitted?

5.2.7 Fresnel zone interference

Calculate the necessary height the radio antennas (assume the same height for each) in order that the first Fresnel zone will have no interference. Assume a signal frequency of 900 MHz, and that the two objects shown lie exactly along the line-of-site path between the two antennas:



Challenges

- If you were the one designing this antenna system, what other concerns might you have about the antenna height in relation to objects possibly infringing on the Fresnel zone?
- Suppose zoning regulations prohibited you from installing an antenna tower any higher than 20 feet. Devise appropriate solution(s) for the radio link given this limitation.

5.2.8 Fresnel zone ground contact

Suppose two radio antennas are located on 12-foot towers, with free and open space in between. If their operating frequency is 900 MHz, calculate the distance between these towers where the first Fresnel zone contacts the ground in between. Is this tower-to-tower distance a *minimum* or a *maximum* value for practical radio operation?

Challenges

- Identify at least one practical factor which would change the distance value allowable between the two antennas. For example, what could change in the system to permit a longer distance between antennas?

5.2.9 Estimating radio transmitter ERP

A student constructs an experimental radio transmitter and needs to know if its radiated power emissions fall within FCC limits for that frequency. One way to determine this is to measure the transmitter's field strength at some standard distance away (e.g. 3 meters), or determining its effective radiated power (ERP) in Watts and then calculate field strength based on that radiated RF power output.

Unfortunately, the student has access to neither an RF field strength meter, nor an RF power meter. They *do* have access to a radio receiver unit equipped with a signal strength meter in decibels, but this dB meter on the radio receiver is simply marked in relative units of "dB" rather than absolute units of "dBm" or "dBW". In other words, the radio receiver cannot provide the quantitative measurement of absolute signal power necessary to determine FCC emissions compliance.

The instructor has an idea, though. She has a hand-held radio transmitter with a known ERP value of 1 Watt, transmitting at a frequency very close to that of the student's experimental transmitter. Setting up this "standard" transmitter some distance from the receiver and pressing the "transmit" button causes the receiver to register -40 dB. Setting up the student's experimental transmitter at the same location and activating it causes the receiver to register -73 dB.

Based on this information, calculate the ERP of the student's experimental transmitter. Then, estimate the electric field strength at a distance of 3 meters from this experimental transmitter.

Challenges

- Assuming a transmission frequency of 433 MHz, does this estimated field strength fall within the unlicensed FCC limits for periodic operation?

5.2.10 Minimum transmitter power

Calculate the minimum required transmitter power for an FM radio broadcast station, assuming we wish an FM radio receiver with a sensitivity of −90 dBm be able to receive the signal. Assume the following gains, losses and other parameters in the system:

- Frequency = 104.5 MHz
- Distance = 120 km
- Transmitter antenna gain = 6 dBi
- Receiver antenna gain = 0 dBi
- Transmitter cable loss = 0.09 dB per foot
- Receiver cable loss = 0.12 dB per foot
- Transmitter cable length = 300 feet
- Receiver cable length = 3 feet long

After calculating the required (minimum) transmitter power, re-calculate to include a *fade margin* of 20 dB for good measure. Explain what "fade margin" is, and why it might be a good idea to include it in an RF link budget calculation.

Challenges

- Identify at least one alteration to any of the given parameters that would result in more power required at the transmitter.

5.2.11 Link budget margin

Perform a link budget analysis of the following radio communication system, using a FreeWave FGR2 industrial data radio (900 MHz) transmitting at 5 mW with the following components and parameters:

- Receiver sensitivity = −110 dBm at BER of 10^{-4}
- Distance = 8.5 miles
- EAN0900YC Yagi transmitting antenna gain = +12.15 dBi
- ASC0202SN transmitting cable (20 feet long) = −1.7 dB
- EAN0906YA Yagi receiving antenna gain = +8 dBi
- ASC0302SN receiving cable (30 feet long) = −2.5 dB

Specifically, calculate the amount of *margin* (i.e. safety factor) between the 5 mW transmitter power and the bare-minimum transmitter power you calculate from the link budget.

Challenges

- Explain the purpose for having a *budget* or *margin* of difference between the minimum required transmitting power and the actual transmitting power. Why should we need more than what is minimally calculated, especially when considering that excess transmitting power may cause problems for neighboring electronic systems?

5.2.12 Maximum distance between antennas

Suppose two radio transceivers must communicate with each other over a long distance. Each one uses a half-wave "whip" antenna with 6 dBi gain, a lightning arrestor at the transceiver (−0.5 dB loss at each transceiver), and each one has 25 feet of coaxial cable connecting the transceiver to the antenna (−0.2 dB loss per foot). Each transceiver transmits 350 mW of RF power, and has a receiving sensitivity of −103 dBm.

Assuming a 900 MHz signal frequency, no fade loss, and absolutely clear line-of-sight between the two antennas, how far away may these transceivers be separated and still fall within the RF budget for good communication?

Challenges

- If the signal frequency were higher, how would this affect the maximum distance?

5.2.13 Using C to calculate receiver power

Examine the following program written in the C programming language, for calculating the amount of RF power available at the input of a receiver:

```
#include <stdio.h>
#include <math.h>

int main (void)
{
  float dist, tgain, rgain, tcable, rcable, f, tpower, Lp, lambda,
        rpower, margin;

  printf("Enter xmtr power (W): ");         scanf("%f", &tpower);
  printf("Enter frequency (Hz): ");         scanf("%f", &f);
  printf("Enter xmtr feed losses (dB): ");  scanf("%f", &tcable);
  printf("Enter xmtr antenna gain (dB): "); scanf("%f", &tgain);
  printf("Enter distance (m): ");           scanf("%f", &dist);
  printf("Enter rcvr antenna gain (dB): "); scanf("%f", &rgain);
  printf("Enter rcvr feed losses (dB): ");  scanf("%f", &rcable);
  printf("Enter fade margin (dB): ");       scanf("%f", &margin);

  tpower = 10 * log10(tpower/1e-3);
  lambda = 2.99792458e8 / f;
  Lp = -20 * log10(4 * M_PI * dist / lambda);
  rpower = tpower + tgain + rgain + Lp + tcable + rcable - margin;

  printf("RF power at receiver input = %f dBm = %e Watts\n",
          rpower, pow(10,rpower/10)/1000);

  return 0;
}
```

Explain how this program calculates receiver power, and also verify its accuracy by simulating one of the link budget examples given in this module.

Challenges

- Add lines of code to ensure that even if the user mistakenly enters positive values for feed losses, the result will still be correct.
- How would the user enter feed losses for a receiver with a built-in antenna?

5.3 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.

5.3.1 Cable loss calculation error

A student is attempting to calculate how much RF power is available at the end of a long coaxial cable given the following parameters:

- RF source power = 50 Watts
- Cable length = 20 feet
- Cable loss rating = 0.11 dB per foot

Taking the 20 foot length and the loss rating of 0.11 dB per foot, the student calculates a total cable loss of 2.2 dB. Next, they convert this dB figure into Watts using the following equation:

$$10^{\frac{2.2}{10}} = 1.6596$$

Subtracting 1.6596 Watts from the cable's input power of 50 leaves 48.34 Watts left over. However, the student's instructor notes this answer is incorrect. Rather than giving away the answer, the instructor poses a question to the student: *"What answer would you get if the RF source only output 1 Watt of power?"*

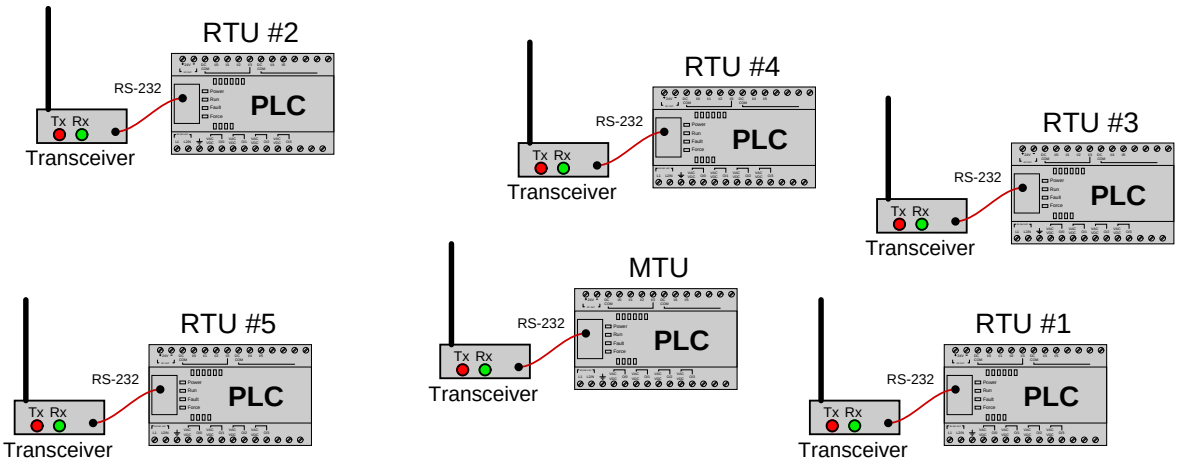
Explain what the student did wrong in their calculations, and also why the instructor's question reveals a conceptual error and not just a procedural error on the student's part.

Challenges

- Come up with a different question exposing the student's conceptual error even clearer than the instructor's question does.

5.3.2 Possible faults in radio SCADA system

In this SCADA system, multiple Programmable Logic Controllers (PLCs) in remote locations communicate to a centralized "master" PLC. This Master Terminal Unit (MTU) communicates with multiple "slave" Remote Terminal Units (RTUs) via radio using Modbus protocol. Each PLC communicates with its radio transceiver via a short RS-232 serial cable:



Unfortunately, there is a problem with this system. RTU number 4 does not communicate with the MTU. Upon inspection, you see that the "Tx" and "Rx" LED indicators are blinking on all the RTU radio transceivers except for RTU #4: there, only the "Rx" LED blinks.

Identify the likelihood of each specified fault for this system. Consider each fault one at a time (i.e. no coincidental faults), determining whether or not each fault is compatible with *all* measurements and symptoms in this SCADA network.

- Damaged antenna on MTU
- Damaged antenna on RTU #4
- PLC halted (not running) RTU #4
- Radio transceiver powered down on RTU #4
- "TD" conductor broken inside RS-232 cable on RTU #4
- "RD" conductor broken inside RS-232 cable on RTU #4
- PLC baud rate set incorrectly on RTU #4
- PLC baud rate set incorrectly on MTU
- Modbus mode (RTU/ASCII) set incorrectly on RTU #4
- Modbus mode (RTU/ASCII) set incorrectly on MTU

- Fresnel zone interference between MTU and RTU #4

Challenges

- A powerful troubleshooting technique is to *swap identical components* between elements of a system to see if and where the problem moves. Explain how this simple test could be applied to this SCADA system in order to diagnose the fault.
- An important question to answer is whether the "Tx" and "Rx" LED indicators on the radio transceivers refer to transmission and reception over the wireless network or over the RS-232 serial network. Explain how we may conclusively answer this question based on the information given and our knowledge of serial Modbus networks.

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

"FGR2 900 MHz Industrial Radio", FreeWave Technologies, Inc., Boulder, CO, 2009.

Goldsmith, Andrea, *Wireless Communications*, draft of second edition, 18 December 2019.

McCollum, Eric; Hao, Kei; Achanta, Shankar V.; Blair, Jeremy; Keckalo, David; "Low-Cost Fast Bus Tripping Scheme Using High-Speed Wireless Protection Sensors", presented at the 45th annual Western Protective Relay Conference, Spokane, WA, October 2018.

The ARRL Antenna Book, Eleventh Edition, The American Radio Relay League, Inc., Newington, CT, 1968.

The ARRL Handbook For Radio Amateurs, 2001 Edition, ARRL – the national association for Amateur Radio, Newington, CT, 2001.

Young, Michael F., "Planning a Microwave Radio Link", YDI Wireless, Falls Church, VA, 2002.

Zeluff, Vin and Markus, John, *Electronics Manual for Radio Engineers*, first edition, McGraw-Hill Book Company Incorporated, New York, 1949.

Zyren, Jim and Petrick, Al, "Tutorial on Basic Link Budget Analysis", Intersil Application Note AN9804.1, 1998.

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.

29 September 2026 – added a new Quantitative Reasoning question focused on estimating ERP for an experimental radio transmitter based on remote signal-strength measurements with that transmitter versus one of known ERP. Also added a new Derivations section on estimating RF field strength.

21 April 2026 – added a new Diagnostic Reasoning question.

12-14 October 2025 – added some instructor notes, and also added comments in the Tutorial about why free-space path loss is frequency-dependent.

9 March 2025 – added some comments on antenna gain to the "Effective radiated power" section of the Tutorial chapter, and also corrected some typographical errors.

30 September 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 added Technical Reference sections on the Poynting vector and Maxwell's equations.

12 March 2024 – added a more instructor notes.

20 February 2024 – added a new section to the Technical References chapter on near-field versus far-field effects.

17-18 October 2023 – minor edits to image_5227, and also refined the questions (and instructor notes) in the "Power and loss calculations" Quantitative Reasoning question. Also added some instructor notes for Challenge questions, and also clarified the verbiage in the Tutorial's bit error rate example.

28 November 2022 – placed questions at the top of the itemized list in the Introduction chapter prompting students to devise experiments related to the tutorial content.

14-15 March 2022 – corrected all instances where EIRP was expressed in relative dB rather than (absolute) dBm, and also added some more content to the Introduction chapter. Also clarified the "Link budget graph" Conceptual Reasoning question.

1-2 December 2021 – added some challenge questions to the "Power ratio and dB conversions" calculation problem, and made changes to the Case Tutorial chapter including the inclusion of a link budget graph to manage all the gains and losses of a link budget calculation problem. The graph makes it visually apparent which decibel figures must be added versus subtracted.

13 November 2021 – added "multipath-induced fading" as a synonym for "fade" loss.

6 September 2021 – added a Quantitative Reasoning question analyzing a C-language program to calculate RF power at a receiver's input given all the standard link budget parameters.

10 May 2021 – commented out or deleted empty chapters.

5 May 2021 – added several more examples to the Case Tutorial chapter.

3 May 2021 – added a Conceptual Reasoning question on the effects of frequency on cable loss.

26 April 2021 – added more examples to the Case Tutorial chapter.

20-21 April 2021 – deleted several "placeholder" lines left over from initial module development in the Questions chapter. Also added a graphic illustration to the "Power and loss calculations" Quantitative Reasoning question. Made minor edits to the Tutorial text. Added more content to the Case Tutorial and Introduction chapters.

2 December 2020 – minor edits to instructor notes.

18 November 2020 – significantly edited the Introduction chapter to make it more suitable as a pre-study guide and to provide cues useful to instructors leading "inverted" teaching sessions. Also made some minor edits to the Tutorial.

28 July 2020 – added questions.

27 July 2020 – document first created.

Index

Adding quantities to a qualitative problem, 88
Ampère's Circuital Law, 51
Ampère's Law, 42
Ampère, André-Marie, 51
Analyzer, spectrum, 18
Annotating diagrams, 87
Antenna, 53
Antenna capture area, 21
Antenna gain, 16, 17, 24
Arrestor, lightning, 16

Bel, 32
BER (bit error rate), 20
Bit error rate, 20
Budget, link, 18

Capture area, antenna, 21
Characteristic impedance of free space, 44
Checking for exceptions, 88
Checking your work, 88
Code, computer, 97
Common logarithm, 32
Conservation of Energy, 15, 47
Cross product, 46

dB, 32
dBm, 16, 36
dBW, 16, 37
Decibel, 32
Dimensional analysis, 87
Dipole antenna, 16

Edwards, Tim, 98
Effective Isotropic Radiated Power (EIRP), 16
Effective Radiated Power (ERP), 16
EIRP, 16
Electromagnetic wave, 43, 49
ERP, 16

Error rate, 20

Fade margin, RF, 22
Fade, RF, 21, 22
Far field region, 43
Faraday's Law of Electromagnetic Induction, 42
Faraday's Law of electromagnetic induction, 51
Faraday, Michael, 51
FCC, 17, 18, 26
Federal Communications Commission, 17, 18, 26
Fourier transform, 38
Frequency domain, 18
Fresnel zone, 27

Gain, amplifier, 32
Gain, antenna, 16, 17, 24
Gauss's Law of electric fields, 50
Gauss's Law of magnetic fields, 50
Graph values to solve a problem, 26, 88
Greenleaf, Cynthia, 55

Heaviside, Oliver, 42, 50
Hertz, Heinrich, 52
How to teach with these modules, 95
Hwang, Andrew D., 99

Identify given data, 87
Identify relevant principles, 87
Impedance of free space, 44
Induction coil, 53
Intermediate results, 87
Inverse square law of radiation, 21
Inverse-Square Law, 44
Inverted instruction, 95
Isotropic antenna, 16

Joule's Law, 39, 46

Knuth, Donald, 98

Lamport, Leslie, 98
Laplace transform, 38
Lenz's Law, 42
Lightning arrestor, 16
Limiting cases, 88
Line of Sight (RF), 27
Link budget, 18
Logarithm, common, 32

Margin, 22
Maxwell's electromagnetic equations, 42, 50
Maxwell, James Clerk, 42, 50
Metacognition, 60
Monopole, 50
Moolenaar, Bram, 97
Multipath, 22
Murphy, Lynn, 55

Near field region, 43
Noise floor, 18–20, 25

Open-source, 97

Path loss, RF, 21
Poynting vector, 43
Poynting, John H., 46
Problem-solving: annotate diagrams, 87
Problem-solving: check for exceptions, 88
Problem-solving: checking work, 88
Problem-solving: dimensional analysis, 87
Problem-solving: graph values, 26, 88
Problem-solving: identify given data, 87
Problem-solving: identify relevant principles, 87
Problem-solving: interpret intermediate results, 87
Problem-solving: limiting cases, 88
Problem-solving: qualitative to quantitative, 88
Problem-solving: quantitative to qualitative, 88
Problem-solving: reductio ad absurdum, 88
Problem-solving: simplify the system, 87
Problem-solving: thought experiment, 87
Problem-solving: track units of measurement, 87
Problem-solving: visually represent the system, 87
Problem-solving: work in reverse, 88
Qualitatively approaching a quantitative problem, 88

Radiator, 52, 53
Radio Frequency (RF), 16
Ratio, signal-to-noise, 19
Reading Apprenticeship, 55
Reductio ad absurdum, 88, 94, 95
Region, far-field, 43
Region, near-field, 43
Resonator, 52, 53
RF, 15, 16
RF link budget, 18
Right-hand rule, 46

Schoenbach, Ruth, 55
Scientific method, 60
Sensitivity, radio receiver, 20
Signal-to-noise ratio, 19
Simplifying a system, 87
Site test, 25
Socrates, 94
Socratic dialogue, 95
Spectrum analyzer, 18
SPICE, 55
Stallman, Richard, 97

Thought experiment, 87
Torvalds, Linus, 97
Transform function, 38

Units of measurement, 87

Vector cross-product, 46
Vector, Poynting, 43
Visualizing a system, 87

Wave, electromagnetic, 43
Wavelength, 21
Western Protective Relay Conference, 26
Work in reverse to solve a problem, 88
WYSIWYG, 97, 98

Yagi antenna, 16