ELECTROMAGNETICS I



Steven W. Ellingson
Virginia Tech



Virginia Tech

Electromagnetics I

Steven W. Ellingson

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

This text was compiled on 06/09/2026

TABLE OF CONTENTS

Licensing

1: Preliminary Concepts

- 1.1: What is Electromagnetics?
- 1.2: Electromagnetic Spectrum
- 1.3: Fundamentals of Waves
- 1.4: Guided and Unguided Waves
- 1.5: Phasors
- 1.6: Units
- 1.7: Notation

2: Electric and Magnetic Fields

- 2.1: What is a Field?
- 2.2: Electric Field Intensity
- 2.3: Permittivity
- 2.4: Electric Flux Density
- 2.5: Magnetic Flux Density
- 2.6: Permeability
- 2.7: Magnetic Field Intensity
- 2.8: Electromagnetic Properties of Materials

3: Transmission Lines

- 3.1: Introduction to Transmission Lines
- 3.2: Types of Transmission Lines
- 3.3: Transmission Lines as Two-Port Devices
- 3.4: Lumped-Element Model
- 3.5: Telegrapher's Equations
- 3.6: Wave Equation for a TEM Transmission Line
- 3.7: Characteristic Impedance
- 3.8: Wave Propagation on a TEM Transmission Line
- 3.9: Lossless and Low-Loss Transmission Lines
- 3.10: Coaxial Line
- 3.11: Microstrip Line
- 3.12: Voltage Reflection Coefficient
- 3.13: Standing Waves
- 3.14: Standing Wave Ratio
- 3.15: Input Impedance of a Terminated Lossless Transmission Line
- 3.16: Input Impedance for Open- and Short-Circuit Terminations
- 3.17: Applications of Open- and Short-Circuited Transmission Line Stubs
- 3.18: Measurement of Transmission Line Characteristics
- 3.19: Quarter-Wavelength Transmission Line
- 3.20: Power Flow on Transmission Lines
- 3.21: Impedance Matching - General Considerations
- 3.22: Single-Reactance Matching
- 3.23: Single-Stub Matching

4: Vector Analysis

- 4.1: Vector Arithmetic
- 4.2: Cartesian Coordinates
- 4.3: Cylindrical Coordinates
- 4.4: Spherical Coordinates
- 4.5: Gradient
- 4.6: Divergence
- 4.7: Divergence Theorem
- 4.8: Curl
- 4.9: Stokes' Theorem
- 4.10: The Laplacian Operator

5: Electrostatics

- 5.1: Coulomb’s Law
- 5.2: Electric Field Due to Point Charges
- 5.3: Charge Distributions
- 5.4: Electric Field Due to a Continuous Distribution of Charge
- 5.5: Gauss’ Law - Integral Form
- 5.6: Electric Field Due to an Infinite Line Charge using Gauss’ Law
- 5.7: Gauss’ Law - Differential Form
- 5.8: Force, Energy, and Potential Difference
- 5.9: Independence of Path
- 5.10: Kirchoff’s Voltage Law for Electrostatics - Integral Form
- 5.11: Kirchoff’s Voltage Law for Electrostatics - Differential Form
- 5.12: Electric Potential Field Due to Point Charges
- 5.13: Electric Potential Field due to a Continuous Distribution of Charge
- 5.14: Electric Field as the Gradient of Potential
- 5.15: Poisson’s and Laplace’s Equations
- 5.16: Potential Field Within a Parallel Plate Capacitor
- 5.17: Boundary Conditions on the Electric Field Intensity (E)
- 5.18: Boundary Conditions on the Electric Flux Density (D)
- 5.19: Charge and Electric Field for a Perfectly Conducting Region
- 5.20: Dielectric Media
- 5.21: Dielectric Breakdown
- 5.22: Capacitance
- 5.23: The Thin Parallel Plate Capacitor
- 5.24: Capacitance of a Coaxial Structure
- 5.25: Electrostatic Energy

6: Steady Current and Conductivity

- 6.1: Convection and Conduction Currents
- 6.2: Current Distributions
- 6.3: Conductivity
- 6.4: Resistance
- 6.5: Conductance
- 6.6: Power Dissipation in Conducting Media

7: Magnetostatics

- 7.1: Comparison of Electrostatics and Magnetostatics
- 7.2: Gauss’ Law for Magnetic Fields - Integral Form

- 7.3: Gauss' Law for Magnetism - Differential Form
- 7.4: Ampere's Circuital Law (Magnetostatics) - Integral Form
- 7.5: Magnetic Field of an Infinitely-Long Straight Current-Bearing Wire
- 7.6: Magnetic Field Inside a Straight Coil
- 7.7: Magnetic Field of a Toroidal Coil
- 7.8: Magnetic Field of an Infinite Current Sheet
- 7.9: Ampere's Law (Magnetostatics) - Differential Form
- 7.10: Boundary Conditions on the Magnetic Flux Density (B)
- 7.11: Boundary Conditions on the Magnetic Field Intensity (H)
- 7.12: Inductance
- 7.13: Inductance of a Straight Coil
- 7.14: Inductance of a Coaxial Structure
- 7.15: Magnetic Energy
- 7.16: Magnetic Materials

8: Time-Varying Fields

- 8.1: Comparison of Static and Time-Varying Electromagnetics
- 8.2: Electromagnetic Induction
- 8.3: Faraday's Law
- 8.4: Induction in a Motionless Loop
- 8.5: Transformers - Principle of Operation
- 8.6: Transformers as Two-Port Devices
- 8.7: The Electric Generator
- 8.8: The Maxwell-Faraday Equation
- 8.9: Displacement Current and Ampere's Law

9: Plane Waves in Loseless Media

- 9.1: Maxwell's Equations in Differential Phasor Form
- 9.2: Wave Equations for Source-Free and Lossless Regions
- 9.3: Types of Waves
- 9.4: Uniform Plane Waves - Derivation
- 9.5: Uniform Plane Waves - Characteristics
- 9.6: Wave Polarization
- 9.7: Wave Power in a Lossless Medium

10: Appendices

- 10.1: Permittivity of Some Common Materials
- 10.2: Permeability of Some Common Materials
- 10.3: Conductivity of Some Common Materials
- 10.4: Mathematical Formulas - Trigonometry
- 10.5: Mathematical Formulas - Vector Operators
- 10.6: Mathematical Formulas - Vector Identities
- 10.7: Physical Constants

Index

Glossary

Detailed Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

CHAPTER OVERVIEW

1: Preliminary Concepts

A field is the continuum of values of a quantity as a function of position and time. The quantity that the field describes may be a scalar or a vector, and the scalar part may be either real- or complex-valued. In electromagnetics, the electric field intensity $\mathbf{E}$ is a real-valued vector field that may vary as a function of position and time, and so might be indicated as "$\mathbf{E}(x,y,z,t)$," "$\mathbf{E}(\mathbf{r},t)$," or simply "$\mathbf{E}$." When expressed as a phasor, this quantity is complex-valued but exhibits no time dependence, so we might say instead "$\widetilde{\mathbf{E}}(\mathbf{r})$" or simply "$\widetilde{\mathbf{E}}$." An example of a scalar field in electromagnetics is the electric potential, V; i.e., $V(\mathbf{r},t)$. A *wave* is a time-varying field that continues to exist in the absence of the source that created it and is therefore able to transport energy.

1.1: What is Electromagnetics?
1.2: Electromagnetic Spectrum
1.3: Fundamentals of Waves
1.4: Guided and Unguided Waves
1.5: Phasors
1.6: Units
1.7: Notation

Thumbnail: Examples of phasors, displayed here as points in the real-imaginary plane.

Contributors and Attributions

- Ellingson, Steven W. (2018) Electromagnetics, Vol. 1. Blacksburg, VA: VT Publishing. https://doi.org/10.21061/electromagnetics-vol-1 Licensed with CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0. Report adoption of this book here. If you are a professor reviewing, adopting, or adapting this textbook please help us understand a little more about your use by filling out this form.

This page titled 1: Preliminary Concepts is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

1.1: What is Electromagnetics?

The topic of this book is applied engineering electromagnetics. This topic is often described as “the theory of electromagnetic fields and waves,” which is both true and misleading. The truth is that electric fields, magnetic fields, their sources, waves, and the behavior these waves are all topics covered by this book. The misleading part is that our *principal* aim shall be to close the gap between basic electrical circuit theory and the more general theory that is required to address certain topics that are of broad and common interest in the field of electrical engineering. (For a preview of topics where these techniques are required, see the list at the end of this section.) In basic electrical circuit theory, the behavior of devices and systems is abstracted in such a way that the underlying electromagnetic principles do not need to be considered. Every student of electrical engineering encounters this, and is grateful since this greatly simplifies analysis and design. For example, a resistor is commonly defined as a device which exhibits a particular voltage

$$V = IR$$

in response to a current I, and the resistor is therefore *completely* described by the value R. This is an example of a “lumped element” abstraction of an electrical device. Much can be accomplished knowing nothing else about resistors; no particular knowledge of the physical concepts of electrical potential, conduction current, or resistance is required. However, this simplification makes it impossible to answer some frequently-encountered questions. Here are just a few:

- What determines R? How does one go about designing a resistor to have a particular resistance?
- Practical resistors are rated for power-handling capability; e.g., discrete resistors are frequently identified as “1/8-W,” “1/4-W,” and so on. How does one determine this, and how can this be adjusted in the design?
- Practical resistors exhibit significant reactance as well as resistance. Why? How is this determined? What can be done to mitigate this?
- Most things which are not resistors also exhibit significant resistance and reactance – for example, electrical pins and interconnects. Why? How is this determined? What can be done to mitigate this?

The answers to the these questions must involve *properties of materials and the geometry in which those materials are arranged*. These are precisely the things that disappear in lumped element device models, so it is not surprising that such models leave us in the dark on these issues. It should also be apparent that what is true for the resistor is also going to be true for other devices of practical interest, including capacitors (and devices unintentionally exhibiting capacitance), inductors (and devices unintentionally exhibiting inductance), transformers (and devices unintentionally exhibiting mutual impedance), and so on. From this perspective, electromagnetics may be viewed as a generalization of electrical circuit theory that addresses these considerations. Conversely basic electric circuit theory may be viewed a special case of electromagnetic theory that applies when these considerations are not important. Many instances of this “electromagnetics as generalization” vs. “lumped-element theory as special case” dichotomy appear in the study of electromagnetics.

There is more to the topic, however. There are many devices and applications in which electromagnetic fields and waves are *primary* engineering considerations that must be dealt with directly. Examples include electrical generators and motors; antennas; printed circuit board stackup and layout; persistent storage of data (e.g., hard drives); fiber optics; and systems for radio, radar, remote sensing, and medical imaging. Considerations such as signal integrity and electromagnetic compatibility (EMC) similarly require explicit consideration of electromagnetic principles.

Although electromagnetic considerations pertain to all frequencies, these considerations become increasingly difficult to avoid with increasing frequency. This is because the wavelength of an electromagnetic field decreases with increasing frequency. When wavelength is large compared to the size of the region of interest (e.g., a circuit), then analysis and design is not much different from zero-frequency (“DC”) analysis and design.

For example, the free space wavelength at 3 MHz is about 100 m, so a planar circuit having dimensions 10 cm × 10 cm is just 0.1% of a wavelength across at this frequency. Although an electromagnetic wave may be present, it has about the same value over the region of space occupied by the circuit. In contrast, the free space wavelength at 3 GHz is about 10 cm, so the same circuit is one full wavelength across at this frequency. In this case, different parts of this circuit observe the same signal with very different magnitude and phase.

Some of the behaviors associated with non-negligible dimensions are undesirable, especially if not taken into account in the design process. However, these behaviors can also be exploited to do some amazing and useful things – for example, to launch an

electromagnetic wave (i.e., an antenna) or to create filters and impedance matching devices consisting only of metallic shapes, free of discrete capacitors or inductors.

Electromagnetic considerations become not only unavoidable but central to analysis and design above a few hundred MHz, and especially in the millimeter-wave, infrared (IR), optical, and ultraviolet (UV) bands. The discipline of electrical engineering encompasses applications in these frequency ranges even though – ironically – such applications may not operate according to principles that can be considered “electrical”! Nevertheless, electromagnetic theory applies.

Another common way to answer the question “What is electromagnetics?” is to identify the topics that are commonly addressed within this discipline. Here’s a list of topics – some of which have already been mentioned – in which explicit consideration of electromagnetic principles is either important or essential:

- Antennas
- Coaxial cable
- Design and characterization of common discrete passive components including resistors, capacitors, inductors, and diodes
- Distributed (e.g., microstrip) filters
- Electromagnetic compatibility (EMC)
- Fiber optics
- Generators
- Magnetic resonance imaging (MRI)
- Magnetic storage (of data)
- Microstrip transmission lines
- Modeling of non-ideal behaviors of discrete components
- Motors
- Non-contact sensors
- Photonics
- Printed circuit board stackup and layout
- Radar
- Radio wave propagation
- Radio frequency electronics
- Signal integrity
- Transformers
- Waveguides

Summary

Applied engineering electromagnetics is the study of those aspects of electrical engineering in situations in which the electromagnetic properties of materials and the geometry in which those materials are arranged is important. This requires an understanding of electromagnetic fields and waves, which are of primary interest in some applications.

Finally, here are two broadly-defined learning objectives that should now be apparent: (1) Learn the techniques of engineering analysis and design that apply when electromagnetic principles are important, and (2) Better understand the physics underlying the operation of electrical devices and systems, so that when issues associated with these physical principles emerge one is prepared to recognize and grapple with them.

This page titled 1.1: What is Electromagnetics? is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **1.1: What is Electromagnetics?** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

1.2: Electromagnetic Spectrum

Electromagnetic fields exist at frequencies from DC (0 Hz) to at least 10^{20} Hz – that's at least 20 orders of magnitude! At DC, electromagnetics consists of two distinct disciplines: *electrostatics*, concerned with electric fields; and *magnetostatics*, concerned with magnetic fields. At higher frequencies, electric and magnetic fields interact to form propagating waves. Waves having frequencies within certain ranges are given names based on how they manifest as physical phenomena. These names are (in order of increasing frequency): radio, infrared (IR), optical (also known as "light"), ultraviolet (UV), X-rays, and gamma rays (γ-rays). See Table 1.2.1 and Figure 1.2.1 for frequency ranges and associated wavelengths.

Definition: Electromagnetic Spectrum

The term *electromagnetic spectrum* refers to the various forms of electromagnetic phenomena that exist over the continuum of frequencies

The speed (properly known as "phase velocity") at which electromagnetic fields propagate in free space is given the symbol c, and has the value $\cong 3.00 \times 10^8$ m/s. This value is often referred to as the "speed of light." While it is certainly the speed of light in free space, it is also the speed of *any* electromagnetic wave in free space. Given frequency f, wavelength is given by the expression

$$\underbrace{\lambda = \frac{c}{f}}_{\text{in free space}}$$

Table 1.2.1 shows the free space wavelengths associated with each of the regions of the electromagnetic spectrum. This book presents a version of electromagnetic theory that is based on classical physics. This approach works well for most practical problems. However, at very high frequencies, wavelengths become small enough that quantum mechanical effects may be important. This is usually the case in the X-ray band and above. In some applications, these effects become important at frequencies as low as the optical, IR, or radio bands. (A prime example is the *photoelectric effect*; see "Additional References" below.) Thus, caution is required when applying the classical version of electromagnetic theory presented here, especially at these higher frequencies.

Table 1.2.1: The electromagnetic spectrum. Note that the indicated ranges are arbitrary but consistent with common usage.

Regime	Frequency Range	Wavelength Range
γ-Ray	$> 3 \times 10^{19}$ Hz	< 0.01 nm
X-Ray	3×10^{16} Hz – 3×10^{19} Hz	10–0.01 nm
Ultraviolet (UV)	2.5×10^{15} – 3×10^{16} Hz	120–10 nm
Optical	4.3×10^{14} – 2.5×10^{15} Hz	700–120 nm
Infrared (IR)	300 GHz – 4.3×10^{14} Hz	1 mm – 700 nm
Radio	3 kHz–300 GHz	100 km – 1 mm

The radio portion of the electromagnetic spectrum alone spans 12 orders of magnitude in frequency (and wavelength), and so, not surprisingly, exhibits a broad range of phenomena. This is shown in Figure 1.2.1.

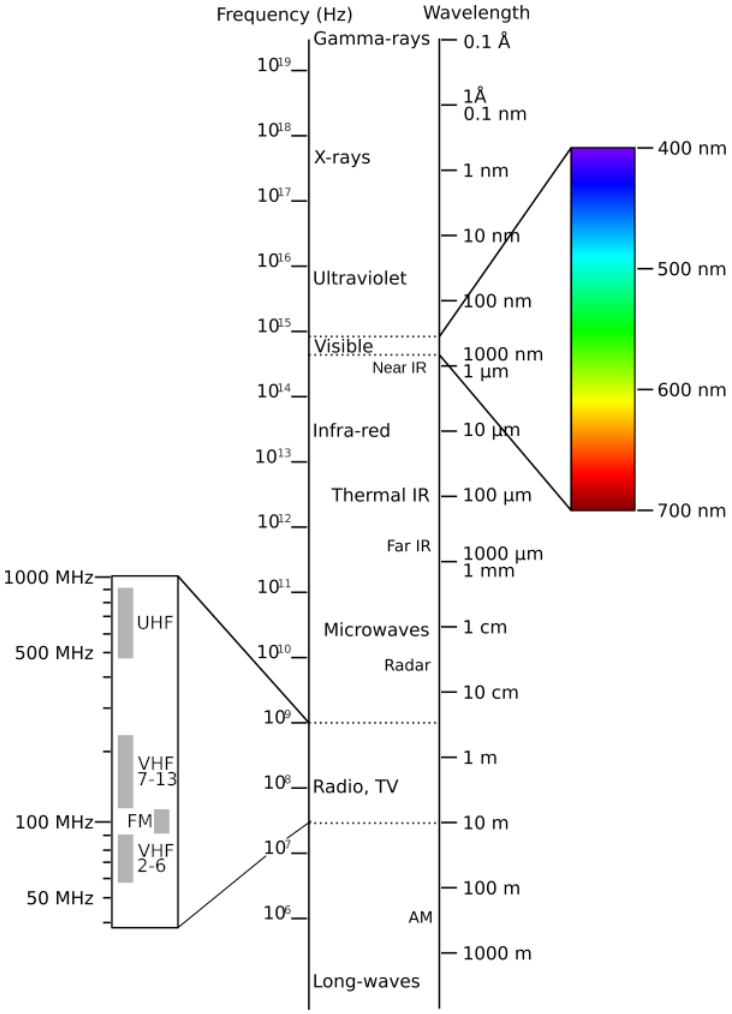


Figure 1.2.1: Electromagnetic Spectrum.

Table 1.2.2: The radio portion of the electromagnetic spectrum, according to a common scheme for naming ranges of radio frequencies. WLAN: Wireless local area network, LMR: Land mobile radio, RFID: Radio frequency identification

Band	Frequencies	Wavelengths	Typical Applications
EHF	30-300 GHz	10–1 mm	60 GHz WLAN, Point-to-point data links
SHF	3–30 GHz	10–1 cm	Terrestrial & Satellite data links, Radar
UHF	300–3000 MHz	1–0.1 m	TV broadcasting, Cellular, WLAN
VHF	30–300 MHz	10–1 m	FM & TV broadcasting, LMR
HF	3–30 MHz	100–10 m	Global terrestrial comm., CB Radio
MF	300–3000 kHz	1000–100 m	AM broadcasting
LF	30–300 kHz	10–1 km	Navigation, RFID
VLF	3–30 kHz	100–10 km	Navigation

Table 1.2.3: The optical portion of the electromagnetic spectrum.

Band	Frequencies	Wavelengths

Band	Frequencies	Wavelengths
Violet	668–789 THz	450–380 nm
Blue	606–668 THz	495–450 nm
Green	526–606 THz	570–495 nm
Yellow	508–526 THz	590–570 nm
Orange	484–508 THz	620–590 nm
Red	400–484 THz	750–620 nm

This page titled 1.2: Electromagnetic Spectrum is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **1.2: Electromagnetic Spectrum** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

1.3: Fundamentals of Waves

In this section, we formally introduce the concept of a wave and explain some basic characteristics.

To begin, let us consider not electromagnetic waves, but rather sound waves. To be clear, sound waves and electromagnetic waves are completely distinct phenomena. Sound waves are variations in pressure, whereas electromagnetic waves are variations in electric and magnetic fields. However, the mathematics that govern sound waves and electromagnetic waves are very similar, so the analogy provides useful insight. Furthermore, sound waves are intuitive for most people because they are readily observed. So, here we go:

Imagine standing in an open field and that it is completely quiet. In this case, the air pressure everywhere is about 101 kPa (101,000 N/m^2) at sealevel, and we refer to this as the *quiescent air pressure*. Sound can be described as the *differential air pressure* $p(x,y,z,t)$, which we define as the absolute air pressure at the spatial coordinates (x,y,z) minus the quiescent air pressure. So, when there is no sound, $p(x,y,z,t)=0$. The function p as an example of a *scalar field*.

Let's also say you are standing at $x=y=z=0$ and you have brought along a friend who is standing at $x=d$; i.e., a distance d from you along the x axis. Also, for simplicity, let us consider only what is happening along the x axis; i.e., $p(x,t)$.

At $t=0$, you clap your hands once. This forces the air between your hands to press outward, creating a region of increased pressure (i.e., $p>0$) that travels outward. As the region of increased pressure moves outward, it leaves behind a region of low pressure where $p<0$. Air molecules immediately move toward this region of lower pressure, and so the air pressure quickly returns to the quiescent value, $p=0$. The traveling disturbance in $p(x,t)$ is the sound of the clap. The disturbance continues to travel outward until it reaches your friend, who then hears the clap.

At each point in time, you can make a plot of $p(x,t)$ versus x for the current value of t. This is shown in Figure 1.3.1. At times $t<0$, we have simply $p(x,t)=0$. A short time after $t=0$, the peak pressure is located at slightly to the right of $x=0$. The pressure is not a simple impulse because interactions between air molecules constrain the pressure to be continuous over space. So instead, we see a rounded pulse representing the rapid build-up and similarly rapid decline in air pressure. A short time later $p(x,t)$ looks very similar, except the pulse is now further away.

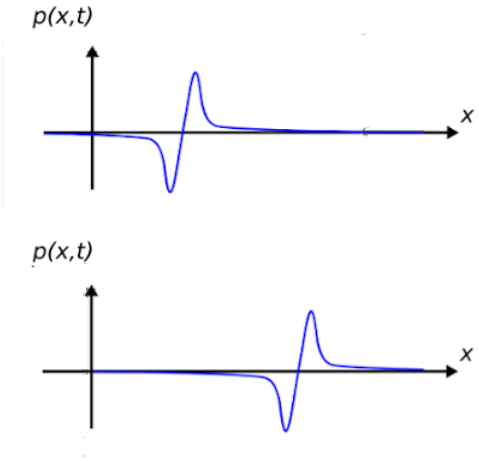


Figure 1.3.1: The differential pressure $p(x,t)$ (top) a short time after the clap and (bottom) a slightly longer time after the clap. (CC BY 4.0; Y. Qin).

Now: What precisely is $p(x,t)$? Completely skipping over the derivation, the answer is that $p(x,t)$ is the solution to the *acoustic wave equation* (see "Additional References" at the end of this section):

$$\frac{\partial^2 p}{\partial x^2}-\frac{1}{c_s^2}\frac{\partial^2 p}{\partial t^2}=0 \tag{1.3.1}$$

where c_s is the speed of sound, which is about 340 m/s at sea level. Just to emphasize the quality of the analogy between sound waves and electromagnetic waves, know that the acoustic wave equation is mathematically identical to equations that that govern electromagnetic waves

Although "transient" phenomena – analogous to a clap – are of interest in electromagnetics, an even more common case of interest is the wave resulting from a sinusoidally-varying source. We can demonstrate this kind of wave in the context of sound as well. Here we go:

In the previous scenario, you pick up a trumpet and blow a perfect A note. The A note is 440 Hz, meaning that the air pressure emerging from your trumpet is varying sinusoidally at a frequency of 440 Hz. Let's say you can continue to blow this note long enough for the entire field to be filled with the sound of your trumpet. Now what does the pressure-versus-distance curve look like? Two simple observations will settle that question:

- $p(x,t)$ at any *constant position* x is a sinusoid as a function of t. This is because the acoustic wave equation is linear and time invariant, so a sinusoidal excitation (i.e., your trumpet) results in a sinusoidal response at the same frequency (i.e., the sound heard by your friend).
- $p(x,t)$ at any *constant time* t is also a sinusoid as a function of x. This is because the sound is propagating away from the trumpet and toward your friend, and anyone in between will also hear the A note, but with a phase shift determined by the difference in distances.

This is enough information to know that the solution must have the form:

$$p(x,t) = A_m \cos(\omega t - \beta x + \psi) \tag{1.3.2}$$

where $\omega = 2\pi f, f = 440\ Hz$, and $A_m, \beta,$ and ψ remain to be determined. You can readily verify that Equation 1.3.2 satisfies the acoustic wave equation when

$$\beta = \frac{\omega}{c_s}$$

In this problem, we find $\beta \cong 8.13$ rad/m. This means that at any given time, the difference in phase measured between any two points separated by a distance of 1 m is 8.13 rad. The parameter β goes by at least three names: *phase propagation constant*, *wavenumber*, and *spatial frequency*. The term "spatial frequency" is particularly apt, since β plays precisely the same role for distance (x) as ω plays for time (t) – This is apparent from Equation 1.3.2. However, "wavenumber" is probably the more commonly-used term.

Definition: Wavenumber

The wavenumber β (rad/m) is the rate at which the phase of a sinusoidal wave progresses with distance.

Note that A_m and ψ are not determined by the wave equation, but instead are properties of the source. Specifically, A_m is determined by how hard we blow, and ψ is determined by the time at which we began to blow and the location of the trumpet. For simplicity, let us assume that we begin to blow at time $t \ll 0$; i.e., in the distant past so that the sound pressure field has achieved steady state by $t = 0$. Also, let us set $\psi = 0$ and set $A_m = 1$ in whatever units we choose to express $p(x,t)$. We then have:

$$p(x,t) = cos(\omega t - \beta x) \tag{1.3.3}$$

Now we have everything we need to make plots of $p(x)$ at various times. Figure 1.3.2*a* shows $p(x, t = 0)$. As expected, $p(x, t = 0)$ is periodic in x. The associated period is referred to as the *wavelength* λ. Since λ is the distance required for the phase of the wave to increase by 2π rad, and because phase is increasing at a rate of β rad/m, we find:

$$\lambda = \frac{2\pi}{\beta}$$

In the present example, we find $\lambda \cong 77.3$ cm.

> Wavelength $\lambda = 2\pi/\beta$ is the distance required for the phase of a sinusoidal wave to increase by one complete cycle (i.e., 2π rad) at any given time.

Now let us consider the situation at $t = +1/4f$, which is $t = 568\mu s$ and $\omega t = \pi/2$. We see in Figure 1.3.2(b) that the waveform has shifted a distance $\lambda/4$ to the right. It is in this sense that we say the wave is propagating in the $+x$ direction. Furthermore, we can now compute a *phase velocity* v_p: We see that a point of constant phase has shifted a distance $\lambda/4$ in time $1/4f$, so

$$v_p = \lambda f$$

In the present example, we find $v_p \cong 340$ m/s; i.e., we have found that the phase velocity is equal to the speed of sound c_s. It is in this sense that we say that the phase velocity is the speed at which the wave propagates.

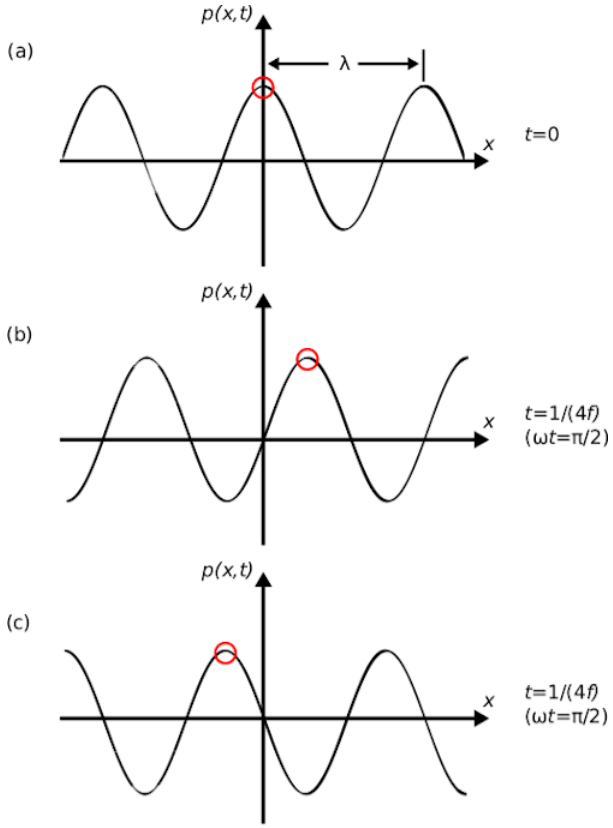


Figure 1.3.2: The differential pressure $p(x,t)$ for (a) $t=0$, (b) $t=1/4f$ for "$-\beta x$," as indicated in Equation 1.3.3 (wave traveling to right); and (c) $t=1/4f$ for "$+\beta x$" (wave traveling to left).

Definition: Phase velocity

Phase velocity $v_p = \lambda f$ is the speed at which a point of constant phase in a sinusoidal waveform travels.

Recall that in Equation 1.3.2 we declared that βx is subtracted from the argument of the sinusoidal function. To understand why, let's change the sign of βx and see if it still satisfies the wave equation – one finds that it does. Next, we repeat the previous experiment and see what happens. The result is shown in Figure 1.3.2*c*. Note that points of constant phase have traveled an equal distance, but now in the $-x$ direction. In other words, this alternative choice of sign for βx within the argument of the cosine function represents a wave that is propagating in the opposite direction. This leads us to the following realization:

> If the phase of the wave is decreasing with βx, then the wave is propagating in the $+x$ direction. If the phase of the wave is increasing with βx, then the wave is propagating in the $-x$ direction

Since the prospect of sound traveling toward the trumpet is clearly nonsense in the present situation, we may neglect the latter possibility. However, what happens if there is a wall located in the distance, behind your friend? Then, we expect an echo from the wall, which would be a second wave propagating in the reverse direction and for which the argument of the cosine function would contain the term "$+\beta x$."

Finally, let us return to electromagnetics. Electromagnetic waves satisfy precisely the same wave equation (i.e., Equation 1.3.1) as do sound waves, except that the phase velocity is much greater. Interestingly, though, the frequencies of electromagnetic waves are also much greater than those of sound waves, so we can end up with wavelengths having similar orders of magnitude. In particular, an electromagnetic wave with $\lambda = 77.3$ cm (the wavelength of the "A" note in the preceding example) lies in the radio portion of the electromagnetic spectrum.

An important difference between sound and electromagnetic waves is that electromagnetic waves are vectors; that is, they have *direction* as well as magnitude. Furthermore, we often need to consider multiple electromagnetic vector waves (in particular, both the *electric field* and the *magnetic field*) in order to completely understand the situation. Nevertheless the concepts of wavenumber,

wavelength, phase velocity, and direction of propagation apply in precisely the same manner to electromagnetic waves as they do to sound waves.

This page titled 1.3: Fundamentals of Waves is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **1.3: Fundamentals of Waves** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

1.4: Guided and Unguided Waves

Broadly speaking, waves may be either **guided** or **unguided**. Unguided waves include those that are radiated by antennas, as well as those that are unintentionally radiated. Once initiated, these waves propagate in an uncontrolled manner until they are redirected by scattering or dissipated by losses associated with materials. Examples of guided waves are those that exist within structures such as transmission lines, waveguides, and optical fibers. We refer to these as guided because they are constrained to follow the path defined by the structure.

> Antennas and unintentional radiators emit *unguided waves*. Transmission lines, waveguides, and optical fibers carry *guided waves*.

This page titled 1.4: Guided and Unguided Waves is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **1.4: Guided and Unguided Waves** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

1.5: Phasors

In many areas of engineering, signals are well-modeled as sinusoids. Also, devices that process these signals are often well-modeled as *linear time-invariant (LTI)* systems. The response of an LTI system to any linear combination of sinusoids is another linear combination of sinusoids having the same frequencies. In other words, (1) sinusoidal signals processed by LTI systems remain sinusoids and are not somehow transformed into square waves or some other waveform; and (2) we may calculate the response of the system for one sinusoid at a time, and then add the results to find the response of the system when multiple sinusoids are applied simultaneously. This property of LTI systems is known as *superposition*.

The analysis of systems that process sinusoidal waveforms is greatly simplified when the sinusoids are represented as phasors. Here is the key idea:

Definition: phasor

A *phasor* is a complex-valued number that represents a real-valued sinusoidal waveform. Specifically, a phasor has the magnitude and phase of the sinusoid it represents

Figure 1.5.1 and 1.5.2 show some examples of phasors and the associated sinusoids. It is important to note that a phasor by itself is not the signal. A phasor is merely a simplified mathematical representation in which the actual, real-valued physical signal is represented as a complex-valued constant.

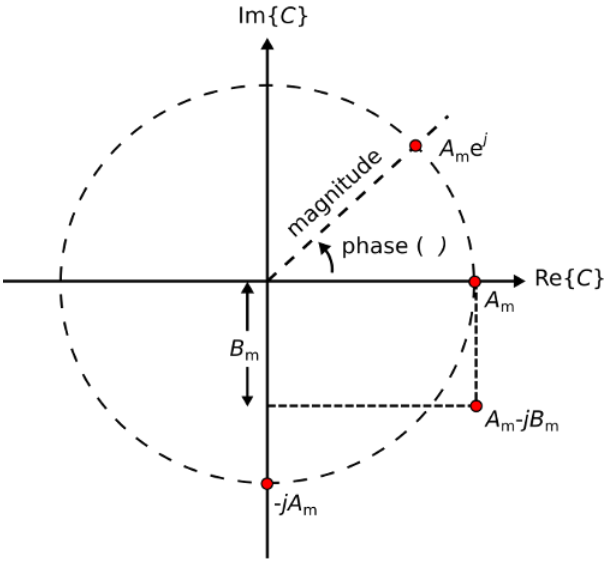


Figure 1.5.1: Examples of phasors, displayed here as points in the real-imaginary plane.

Equation 1.5.1 is a completely general form for a physical (hence, real-valued) quantity varying sinusoidally with angular frequency $\omega = 2\pi f$

$$A(t;\omega) = A_m(\omega)\cos(\omega t + \psi(\omega)) \tag{1.5.1}$$

where $A_m(\omega)$ is magnitude at the specified frequency, $\psi(\omega)$ is phase at the specified frequency, and t is time. Also, we require $\partial Am/\partial t = 0$; that is, that the time variation of $A(t)$ is completely represented by the cosine function alone. Now we can equivalently express $A(t;\omega)$ as a phasor $C(\omega)$:

$$C(\omega) = A_m(\omega)e^{j\psi(\omega)} \tag{1.5.2}$$

To convert this phasor back to the physical signal it represents, we (1) restore the time dependence by multiplying by $e^{j\omega t}$, and then (2) take the real part of the result. In mathematical notation:

$$\boxed{A(t;\omega) = \mathrm{Re}\left\{C(\omega)e^{j\omega t}\right\}} \tag{1.5.3}$$

To see why this works, simply substitute the right hand side of Equation 1.5.2 into Equation 1.5.3. Then

$$\begin{aligned} A(t) &= \mathrm{Re}\left\{A_m(\omega)e^{j\psi(\omega)}e^{j\omega t}\right\} \\ &= \mathrm{Re}\left\{A_m(\omega)e^{j(\omega t+\psi(\omega))}\right\} \\ &= \mathrm{Re}\left\{A_m(\omega)[\cos(\omega t+\psi(\omega))+j\sin(\omega t+\psi(\omega))]\right\} \\ &= A_m(\omega)\cos(\omega t+\psi(\omega)) \end{aligned}$$

as expected.

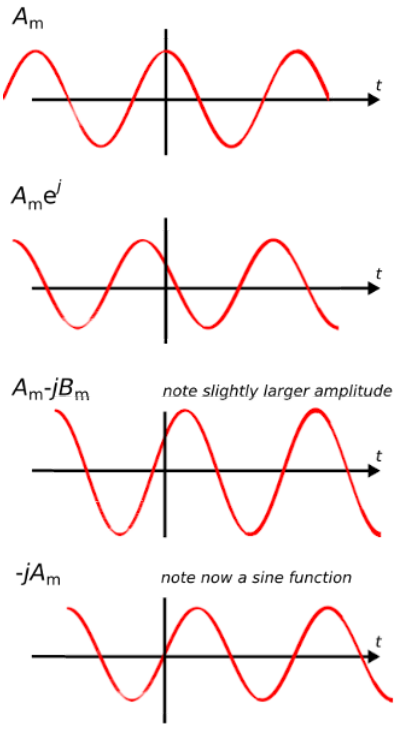


Figure 1.5.2: Sinusoids corresponding to the phasors shown in Figure 1.5.1

It is common to write Equation 1.5.3 as follows, dropping the explicit indication of frequency dependence:

$$C = A_m e^{j\psi}$$

This does not normally cause any confusion since the definition of a phasor requires that values of C and ψ are those that apply at whatever frequency is represented by the suppressed sinusoidal dependence $e^{j\omega t}$.

Table 1.5.1 shows mathematical representations of the same phasors demonstrated in Figure 1.5.1 (and their associated sinusoidal waveforms in Figure 1.5.2). It is a good exercise is to confirm each row in the table, transforming from left to right and vice-versa.

Table 1.5.1: Some examples of physical (real-valued) sinusoidal signals and the corresponding phasors. A_m and B_m are real-valued and constant with respect to t

$A(t)$	C
$A_m\cos(\omega t)$	A_m
$A_m\cos(\omega t+\psi)$	$A_m e^{j\psi}$
$A_m\sin(\omega t) = A_m(cos\omega t-\frac{\pi}{2})$	$-jA_m$
$A_m\cos(\omega t)+B_m\sin(\omega t) = A_m\cos(\omega t)+B_m\cos(\omega t-\frac{\pi}{2}$	$A_m - jB_m$

It is not necessary to use a phasor to represent a sinusoidal signal. We choose to do so because phasor representation leads to dramatic simplifications. For example:

- Calculation of the peak value from data representing $A(t;\omega)$ requires a time-domain search over one period of the sinusoid. However, if you know C, the peak value of $A(t)$ is simply $|C|$, and no search is required.

- Calculation of ψ from data representing $A(t;\omega)$ requires correlation (essentially, integration) over one period of the sinusoid. However, if you know C, then ψ is simply the phase of C, and no integration is required.

Furthermore, mathematical operations applied to $A(t;\omega)$ can be equivalently performed as operations on C, and the latter are typically much easier than the former. To demonstrate this, we first make two important claims and show that they are true.

Example 1.5.1: Claim 1

Let C_1 and C_2 be two complex-valued constants (independent of t). Also, $\mathrm{Re}\left\{C_1 e^{j\omega t}\right\} = \mathrm{Re}\left\{C_2 e^{j\omega t}\right\}$ for all t. Then, $C_1 = C_2$.

Proof

Evaluating at $t = 0$ we find $\mathrm{Re}\{C_1\} = \mathrm{Re}\{C_2\}$. Since C_1 and C_2 are constant with respect to time, this must be true for all t. At $t = \pi/(2\omega)$ we find

$$\mathrm{Re}\left\{C_1 e^{j\omega t}\right\} = \mathrm{Re}\{C_1 \cdot j\} = -\mathrm{Im}\{C_1\}$$

and similarly

$$\mathrm{Re}\left\{C_2 e^{j\omega t}\right\} = \mathrm{Re}\{C_2 \cdot j\} = -\mathrm{Im}\{C_2\}$$

therefore $\mathrm{Im}\{C_1\} = \mathrm{Im}\{C_2\}$. Once again: Since C_1 and C_2 are constant with respect to time, this must be true for all t. Since the real and imaginary parts of C_1 and C_2 are equal, $C_1 = C_2$.

What does this mean?

We have just shown that if two phasors are equal, then the sinusoidal waveforms that they represent are also equal.

Example 1.5.2: Claim 2

For any real-valued linear operator $\mathcal{T}$ and complex-valued quantity C,

$$\mathcal{T}(\mathrm{Re}\{C\}) = \mathrm{Re}\{\mathcal{T}(C)\}. \tag{1.5.4}$$

Proof

Let $C = c_r + jc_i$ where c_r and c_i are real-valued quantities, and evaluate the right side of Equation 1.5.4:

$$\begin{aligned}\mathrm{Re}\{\mathcal{T}(C)\} &= \mathrm{Re}\left\{\mathcal{T}\left(c_r + jc_i\right)\right\} \\ &= \mathrm{Re}\left\{\mathcal{T}\left(c_r\right) + j\mathcal{T}\left(c_i\right)\right\} \\ &= \mathcal{T}\left(c_r\right) \\ &= \mathcal{T}(\mathrm{Re}\{C\})\end{aligned}$$

What does this mean?

The operators that we have in mind for $\mathcal{T}$ include addition, multiplication by a constant, differentiation, integration, and so on. Here's an example with differentiation:

$$\mathrm{Re}\left\{\frac{\partial}{\partial\omega}C\right\} = \mathrm{Re}\left\{\frac{\partial}{\partial\omega}\left(c_r + jc_i\right)\right\} = \frac{\partial}{\partial\omega}c_r$$

$$\frac{\partial}{\partial\omega}\mathrm{Re}\{C\} = \frac{\partial}{\partial\omega}\mathrm{Re}\left\{\left(c_r + jc_i\right)\right\} = \frac{\partial}{\partial\omega}c_r$$

In other words, differentiation of a sinusoidal signal can be accomplished by differentiating the associated phasor, so there is no need to transform a phasor back into its associated real-valued signal in order to perform this operation.

Summary

Claims 1 and 2 together entitle us to perform operations on phasors as surrogates for the physical, real-valued, sinusoidal waveforms they represent. Once we are done, we can transform the resulting phasor back into the physical waveform it represents using Equation 1.5.3, if desired

However, a final transformation back to the time domain is usually *not* desired, since the phasor tells us everything we can know about the corresponding sinusoid

A skeptical student might question the value of phasor analysis on the basis that signals of practical interest are sometimes not sinusoidally-varying, and therefore phasor analysis seems not to apply generally. It is certainly true that many signals of practical interest are not sinusoidal, and many are far from it. Nevertheless, phasor analysis is broadly applicable. There are basically two reasons why this is so:

- Many signals, although not strictly sinusoidal, are "narrowband" and therefore well-modeled as sinusoidal. For example, a cellular telecommunications signal might have a bandwidth on the order of 10 MHz and a center frequency of about 2 GHz. This means the difference in frequency between the band edges of this signal is just 0.5% of the center frequency. The frequency response associated with signal propagation or with hardware can often be assumed to be constant over this range of frequencies. With some caveats, doing phasor analysis at the center frequency and assuming the results apply equally well over the bandwidth of interest is often a pretty good approximation.
- It turns out that phasor analysis is easily extensible to any physical signal, regardless of bandwidth. This is so because any physical signal can be decomposed into a linear combination of sinusoids – this is known as Fourier analysis. The way to find this linear combination of sinusoids is by computing the Fourier series, if the signal is periodic, or the Fourier Transform, otherwise. Phasor analysis applies to each frequency independently, and (invoking superposition) the results can be added together to obtain the result for the complete signal. The process of combining results after phasor analysis results is nothing more than integration over frequency; i.e..:

$$\int_{-\infty}^{+\infty} A(t;\omega)d\omega$$

Using Equation 1.5.3, this can be rewritten:

$$\int_{-\infty}^{+\infty} \mathrm{Re}\left\{C(\omega)e^{j\omega t}\right\} d\omega$$

We can go one step further using Claim 2:

$$\mathrm{Re}\left\{\int_{-\infty}^{+\infty} C(\omega)e^{j\omega t}d\omega\right\}$$

The quantity in the curly braces is simply the Fourier transform of $C(\omega)$. Thus, we see that we can analyze a signal of arbitrarily-large bandwidth simply by keeping ω as an independent variable while we are doing phasor analysis, and if we ever need the physical signal, we just take the real part of the Fourier transform of the phasor. So not only is it possible to analyze any time-domain signal using phasor analysis, it is also often far easier than doing the same analysis on the time-domain signal direct

summary

Phasor analysis does not limit us to sinusoidal waveforms. Phasor analysis is not only applicable to sinusoids and signals that are sufficiently narrowband, but is also applicable to signals of arbitrary bandwidth via Fourier analysis.

This page titled 1.5: Phasors is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **1.5: Phasors** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

1.6: Units

The term “unit” refers to the measure used to express a physical quantity. For example, the mean radius of the Earth is about 6,371,000 meters; in this case the unit is the meter.

A number like “6,371,000” becomes a bit cumbersome to write, so it is common to use a prefix to modify the unit. For example, the radius of the Earth is more commonly said to be 6371 kilometers, where one kilometer is understood to mean 1000 meters. It is common practice to use prefixes, such as “kilo-,” that yield values in the range 0.001 to 10, 000. A list of standard prefixes appears in Table 1.6.1.

Table 1.6.1: Prefixes used to modify units.

Prefix	Multiply by:	Multiply by:
exa	E	10^{18}
peta	P	10^{15}
tera	T	10^{12}
giga	G	10^{9}
mega	M	10^{6}
kil	k	10^{3}
milli	m	10^{-3}
micro	μ	10^{-6}
nano	n	10^{-9}
pico	p	10^{-12}
femto	f	10^{-15}
atto	a	10^{-18}

Writing out the names of units can also become tedious. For this reason, it is common to use standard abbreviations; e.g., “6731 km” as opposed to “6371 kilometers,” where “k” is the standard abbreviation for the prefix “kilo” and “m” is the standard abbreviation for “meter.” A list of commonly-used base units and their abbreviations are shown in Table 1.6.2.

Table 1.6.2: Some units that are commonly used in electromagnetics.

Unit	Abbreviation	Quantifies:
ampere	A	electric current
coulomb	C	electric charge
farad	F	capacitance
henry	H	inductance
hertz	Hz	frequency
joule	J	energy
meter	m	distance
newton	N	force
ohm	Ω	resistance
second	s	time
tesla	T	magnetic flux density

Unit	Abbreviation	Quantifies:
volt	V	electric potential
watt	W	power
weber	Wb	magnetic flux

To avoid ambiguity, it is important to always indicate the units of a quantity; e.g., writing “6371 km” as opposed to “6371.” Failure to do so is a common source of error and misunderstandings. An example is the expression:

$$l = 3t$$

where l is length and t is time. It could be that l is in meters and t is in seconds, in which case “3” really means “3 m/s”. However, if it is intended that l is in kilometers and t is in hours, then “3” really means “3 km/h” and the equation is literally different. To patch this up, one might write “$l = 3t$ m/s”; however, note that this does does not resolve the ambiguity we just identified – i.e., we still don’t know the units of the constant “3.” Alternatively, one might write “$l = 3t$ where l is in meters and t is in seconds,” which is unambiguous but becomes quite awkward for more complicated expressions. A better solution is to write instead:

$$l = (3\ \mathrm{m/s})t$$

or even better:

$$l = at$$

where $a = 3\ \mathrm{m/s}$ since this separates this issue of units from the perhaps more important fact that l is proportional to t and the constant of proportionality (a) is known.

The meter is the fundamental unit of length in the International System of Units, known by its French acronym “SI” and sometimes informally referred to as the “metric system.”

> In this work, we will use SI units exclusively.

Although SI is probably the most popular for engineering use overall, other systems remain in common use. For example, the English system, where the radius of the Earth might alternatively be said to be about 3959 miles, continues to be used in various applications and to a lesser or greater extent in various regions of the world. An alternative system in common use in physics and material science applications is the CGS (“centimeter-gram-second”) system. The CGS system is similar to SI, but with some significant differences. For example, the base unit of energy is the CGS system is not the “joule” but rather the “erg,” and the values of some physical constants become unitless. Therefore – once again – it is very important to include units whenever values are stated.

SI defines seven fundamental units from which all other units can be derived. These fundamental units are distance in meters (m), time in seconds (s), current in amperes (A), mass in kilograms (kg), temperature in kelvin (K), particle count in moles (mol), and luminosity in candela (cd). SI units for electromagnetic quantities such as coulombs (C) for charge and volts (V) for electric potential are derived from these fundamental units.

A frequently-overlooked feature of units is their ability to assist in error-checking mathematical expressions. For example, the electric field intensity may be specified in volts per meter (V/m), so an expression for the electric field intensity that yields units of V/m is said to be “dimensionally correct” (but not necessarily correct), whereas an expression that cannot be reduced to units of V/m *cannot* be correct.

This page titled 1.6: Units is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **1.6: Units** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

1.7: Notation

The list below describes notation used in this book.

- *Vectors:* Boldface is used to indicate a vector; e.g., the electric field intensity vector will typically appear as $\mathbf{E}$. Quantities not in boldface are scalars. When writing by hand, it is common to write "$\overline{E}$" or "$\vec{E}$ " in lieu of "$\mathbf{E}$."
- *Unit vectors:* A circumflex is used to indicate a unit vector; i.e., a vector having magnitude equal to one. For example, the unit vector pointing in the $+x$ direction will be indicated as $\hat{\mathbf{x}}$. In discussion, the quantity " $\hat{\mathbf{x}}$" is typically spoken "x hat."
- *Time*: The symbol t is used to indicate time.
- *Position*: The symbols $(x, y, z), (\rho, \phi, z)$, and (r, θ, ϕ) indicate positions using the Cartesian, cylindrical, and polar coordinate systems, respectively. It is sometimes convenient to express position in a manner which is independent of a coordinate system; in this case, we typically use the symbol $\mathbf{r}$. For example, $\mathbf{r} = \hat{\mathbf{x}}x + \hat{\mathbf{y}}y + \hat{\mathbf{z}}z$ in the Cartesian coordinate system.
- *Phasors:* A tilde is used to indicate a phasor quantity; e.g., a voltage phasor might be indicated as $\tilde{V}$, and the phasor representation of $\mathbf{E}$ will be indicated as $\tilde{\mathbf{E}}$.
- *Curves, surfaces, and volumes*: These geometrical entities will usually be indicated in script; e.g., an open surface might be indicated as $\mathcal{S}$ and the curve bounding this surface might be indicated as $\mathcal{C}$. Similarly, the volume enclosed by a closed surface $\mathcal{S}$ may be indicated as $\mathcal{V}$.
- *Integrations over curves, surfaces, and volumes* will usually be indicated using a single integral sign with the appropriate subscript. For example:

$$\int_{\mathcal{C}} \cdots dl \text{ is an integral over the curve } \mathcal{C}$$

$$\int_{\mathcal{S}} \cdots ds \text{ is an integral over the surface } \mathcal{S}$$

$$\int_{\mathcal{V}} \cdots ds \text{ is an integral over the volume } \mathcal{V}.$$

- *Integrations over closed curves and surfaces* will be indicated using a circle superimposed on the integral sign. For example:

$$\oint_{\mathcal{C}} \ldots dl \text{ is an integral over the closed curve } \mathcal{C}$$

$$\oint_{\mathcal{S}} \ldots ds \text{ is an integral over the closed surface } \mathcal{S}$$

  A "closed curve" is one which forms an unbroken loop; e.g., a circle. A "closed surface" is one which encloses a volume with no openings; e.g., a sphere.
- The symbol "$\cong$" means "approximately equal to." This symbol is used when equality exists, but is not being expressed with exact numerical precision. For example, the ratio of the circumference of a circle to its diameter is π, where $\pi \cong 3.14$.
- The symbol "$\approx$" also indicates "approximately equal to," but in this case the two quantities are unequal even if expressed with exact numerical precision. For example, $e^x = 1 + x + x^2/2 + \ldots$ as a infinite series, but $e^x \approx 1 + x$ for $x \ll 1$. Using this approximation $e^{0.1} \approx 1.1$, which is in good agreement with the actual value $e^{0.1} \cong 1.1052$.
- The symbol "$\sim$" indicates "on the order of," which is a relatively weak statement of equality indicating that the indicated quantity is within a factor of 10 or so the indicated value. For example, $\mu \sim 10^5$ for a class of iron alloys, with exact values being being larger or smaller by a factor of 5 or so.
- The symbol "$\triangleq$" means "is defined as" or "is equal as the result of a definition."
- Complex numbers: $j \triangleq \sqrt{-1}$.
- See Appendix C for notation used to identify commonly-used physical constants.

This page titled 1.7: Notation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **1.7: Notation** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

CHAPTER OVERVIEW

2: Electric and Magnetic Fields

2.1: What is a Field?
2.2: Electric Field Intensity
2.3: Permittivity
2.4: Electric Flux Density
2.5: Magnetic Flux Density
2.6: Permeability
2.7: Magnetic Field Intensity
2.8: Electromagnetic Properties of Materials

Thumbnail: The magnetic field of a current-bearing coil, illustrating field lines. (CC BY 4.0; Y. Qing).

Contributors and Attributions

- Ellingson, Steven W. (2018) Electromagnetics, Vol. 1. Blacksburg, VA: VT Publishing. https://doi.org/10.21061/electromagnetics-vol-1 Licensed with CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0. Report adoption of this book here. If you are a professor reviewing, adopting, or adapting this textbook please help us understand a little more about your use by filling out this form.

This page titled 2: Electric and Magnetic Fields is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

2.1: What is a Field?

The quantity that the field describes may be a scalar or a vector, and the scalar part may be either real- or complex-valued.

Definition: Field

A field is the continuum of values of a quantity as a function of position and time.

In electromagnetics, the electric field intensity $\mathbf{E}$ is a real-valued vector field that may vary as a function of position and time, and so might be indicated as "$\mathbf{E}(x,y,z,t),$" "$\mathbf{E}(\mathbf{r},t),$" or simply "$\mathbf{E}$." When expressed as a phasor, this quantity is complex-valued but exhibits no time dependence, so we might say instead "$\widetilde{\mathbf{E}}(\mathbf{r})$" or simply "$\widetilde{\mathbf{E}}$."

An example of a scalar field in electromagnetics is the electric potential, V; i.e., $\mathrm{V}(\mathbf{r},\mathrm{t})$.

A wave is a time-varying field that continues to exist in the absence of the source that created it and is therefore able to transport energy.

This page titled 2.1: What is a Field? is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **2.1: What is a Field?** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

2.2: Electric Field Intensity

Electric field intensity is a vector field we assign the symbol $\mathbf{E}$ and has units of electrical potential per distance; in SI units, volts per meter (V/m). Before offering a formal definition, it is useful to consider the broader concept of the electric field.

Imagine that the universe is empty except for a single particle of positive charge. Next, imagine that a second positively-charged particle appears; the situation is now as shown in Figure 2.2.1. Since like charges repel, the second particle will be repelled by the first particle and vice versa. Specifically, the first particle is *exerting force* on the second particle. If the second particle is free to move, it will do so; this is an expression of kinetic energy. If the second particle is somehow held in place, we say the second particle possesses an equal amount of potential energy. This potential energy is no less "real," since we can convert it to kinetic energy simply by releasing the particle, thereby allowing it to move.

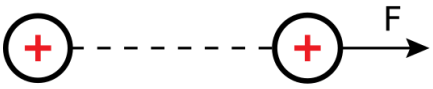


Figure 2.2.1: A positively-charged particle experiences a repulsive force $\mathbf{F}$ in the presence of another particle which is also positively-charged. (CC BY SA 4.0; M. Goldammer).

Now let us revisit the original one particle scenario. In that scenario, we could make a map in which every position in space is assigned a vector that describes the force that a particle having a specified charge q would experience if it were to appear there. The result looks something like Figure 2.2.2. This map of force vectors is essentially a description of the electric field associated with the first particle.

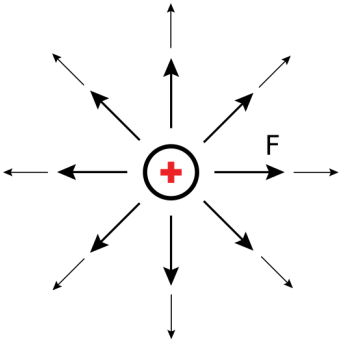


Figure 2.2.2: A map of the force that would be experienced by a second particle having a positive charge. Here, the magnitude and direction of the force is indicated by the size and direction of the arrow. (CC BY SA 4.0; M. Goldammer).

There are many ways in which the electric field may be quantified. Electric field intensity $\mathbf{E}$ is simply one of these ways. We define $\mathbf{E}(\mathbf{r})$ to be the force $\mathbf{F}(\mathbf{r})$ experienced by a test particle having charge q, divided by q; i.e.,

$$\mathbf{E}(\mathbf{r}) \triangleq \lim_{q \to 0} \frac{\mathbf{F}(\mathbf{r})}{q} \tag{2.2.1}$$

Note that it is required for the charge to become vanishingly small (as indicated by taking the limit) in order for this definition to work. This is because the source of the electric field is charge, so the test particle contributes to the total electric field. To accurately measure the field of interest, the test charge must be small enough not to significantly perturb the field. This makes Equation 2.2.1 awkward from an engineering perspective, and we'll address that later in this section.

According the definition of Equation 2.2.1, the units of $\mathbf{E}$ are those of force divided by charge. The SI units for force and charge are the newton (N) and coulomb (C) respectively, so $\mathbf{E}$ has units of N/C. However, we typically express $\mathbf{E}$ in units of V/m, not N/C. What's going on? The short answer is that 1 V/m = 1 N/C:

$$\frac{\mathrm{N}}{\mathrm{C}} = \frac{\mathrm{N \cdot m}}{\mathrm{C \cdot m}} = \frac{\mathrm{J}}{\mathrm{C \cdot m}} = \frac{\mathrm{V}}{\mathrm{m}}$$

where we have used the fact that 1 N·m = 1 joule (J) of energy and 1 J/C = 1 V.

Electric field intensity is a Vector Field

Electric field intensity (**E**, N/C or V/m) is a vector field that quantifies the force experienced by a charged particle due to the influence of charge not associated with that particle.

The analysis of units doesn't do much to answer the question of why we should prefer to express **E** in V/m as opposed to N/C. Let us now tackle that question.

Figure 2.2.3 shows a simple thought experiment that demonstrates the concept of electric field intensity in terms of an electric circuit. This circuit consists of a parallel-plate capacitor in series with a 9 V battery. The effect of the battery, connected as shown, is to force an accumulation of positive charge on the upper plate, and an accumulation of negative charge on the lower plate. If we consider the path from the position labeled "A," along the wire and through the battery to the position labeled "B," the change in electric potential is +9 V. It must also be true that the change in electric potential as we travel from B to A through the capacitor is −9 V, since the sum of voltages over any closed loop in a circuit is zero. Said differently, the change in electric potential between the plates of the capacitor, starting from node A and ending at node B, is +9 V. Now, note that the spacing between the plates in the capacitor is 1 mm. Thus, the rate of change of the potential between the plates is 9 V divided by 1 mm, which is 9000 V/m. This is literally the electric field intensity between the plates. That is, if one places a particle with an infinitesimally-small charge between the plates (point "C"), and then measures the ratio of force to charge, one finds it is 9000 N/C pointing toward A. We come to the following remarkable conclusion:

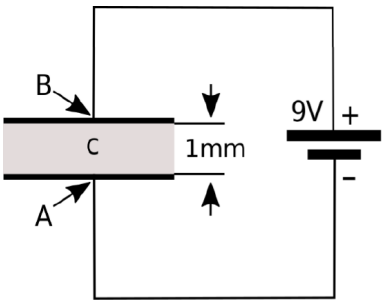


Figure 2.2.3: A simple circuit used to describe the concept of electric field intensity. In this example, E at point C is 9000 V/m directed from B toward A. (CC BY 3.0; Y. Qin).

E points in the direction in which electric potential is most rapidly decreasing, and the magnitude of **E** is the rate of change in electric potential with distance in this direction.

The reader may have noticed that we have defined the electric field in terms of what it *does*. We have have not directly addressed the question of what the electric field *is*. This is the best we can do using classical physics, and fortunately, this is completely adequate for the most engineering applications. However, a deeper understanding is possible using quantum mechanics, where we find that the electric field and the magnetic field are in fact manifestations of the same fundamental force, aptly named the *electromagnetic force*. (In fact, the electromagnetic force is found to be one of just four fundamental forces, the others being *gravity*, the *strong nuclear force*, and the *weak nuclear force*.) Quantum mechanics also facilitates greater insight into the nature of electric charge and of the photon, which is the fundamental constituent of electromagnetic waves. For more information on this topic, an excellent starting point is the video "Quantum Invariance & The Origin of The Standard Model" referenced at the end of this section.

This page titled 2.2: Electric Field Intensity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **2.2: Electric Field Intensity** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

2.3: Permittivity

Permittivity describes the effect of material in determining the electric field in response to electric charge. For example, one can observe from laboratory experiments that a particle having charge q gives rise to the electric field

$$\mathbf{E} = \hat{\mathbf{R}}\, q\, \frac{1}{4\pi R^2}\, \frac{1}{\epsilon}$$

where R is distance from the charge, $\hat{\mathbf{R}}$ is a unit vector pointing away from the charge, and ϵ is a constant that depends on the material. Note that $\mathbf{E}$ increases with q, which makes sense since electric charge is the source of $\mathbf{E}$. Also note that $\mathbf{E}$ is inversely proportional to $4\pi R^2$, indicating that $\mathbf{E}$ decreases in proportion to the area of a sphere surrounding the charge – a principle commonly known as the *inverse square law*. The remaining factor $1/\epsilon$ is the constant of proportionality, which captures the effect of material. Given units of V/m for $\mathbf{E}$ and C for Q, we find that ϵ must have units of farads per meter (F/m). (To see this, note that 1 F = 1 C/V.)

Permittivity

Permittivity (ϵ, F/m) describes the effect of material in determining the electric field intensity in response to charge.

In free space (that is, a perfect vacuum), we find that $\epsilon = \epsilon_0$ where:

$$\epsilon_0 \cong 8.854 \times 10^{-12}\ \mathrm{F/m}$$

The permittivity of air is only slightly greater, and usually can be assumed to be equal to that of free space. In most other materials, the permittivity is significantly greater; that is, the same charge results in a weaker electric field intensity.

It is common practice to describe the permittivity of materials relative to the permittivity of free space. This *relative permittivity* is given by:

$$\epsilon_r \triangleq \frac{\epsilon}{\epsilon_0}$$

Values of ϵ_r for a few representative materials is given in Appendix A1. Note that ϵ_r ranges from 1 (corresponding to a perfect vacuum) to about 60 or so in common engineering applications. Also note that relative permittivity is sometimes referred to as *dielectric constant*. This term is a bit misleading, however, since permittivity is a meaningful concept for many materials that are not dielectrics.

This page titled 2.3: Permittivity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **2.3: Permittivity** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

2.4: Electric Flux Density

Electric flux density, assigned the symbol $\mathbf{D}$, is an alternative to electric field intensity ($\mathbf{E}$) as a way to quantify an electric field. This alternative description offers some actionable insight, as we shall point out at the end of this section.

First, what is electric flux density? Recall that a particle having charge q gives rise to the electric field intensity

$$\mathbf{E} = \hat{\mathbf{R}}\ q\ \frac{1}{4\pi R^2}\ \frac{1}{\epsilon} \tag{2.4.1}$$

where R is distance from the charge and $\hat{\mathbf{R}}$ points away from the charge. Note that $\mathbf{E}$ is inversely proportional to $4\pi R^2$, indicating that $\mathbf{E}$ decreases in proportion to the area of a sphere surrounding the charge. Now integrate both sides of Equation 2.4.1 over a sphere $\mathcal{S}$ of radius R:

$$\oint_{\mathcal{S}} \mathbf{E}(\mathbf{r})\cdot d\mathbf{s} = \oint_{\mathcal{S}} \left[\hat{\mathbf{R}}\ q\ \frac{1}{4\pi R^2}\ \frac{1}{\epsilon}\right]\cdot d\mathbf{s}$$

Factoring out constants that do not vary with the variables of integration, the right-hand side becomes:

$$q\ \frac{1}{4\pi R^2}\ \frac{1}{\epsilon}\ \oint_{\mathcal{S}} \hat{\mathbf{R}}\cdot d\mathbf{s}$$

Note that $d\mathbf{s} = \hat{\mathbf{R}} ds$ in this case, and also that $\hat{\mathbf{R}}\cdot\hat{\mathbf{R}} = 1$. Thus, the right-hand side simplifies to:

$$q\ \frac{1}{4\pi R^2}\ \frac{1}{\epsilon}\ \oint_{\mathcal{S}} ds$$

The remaining integral is simply the area of $\mathcal{S}$, which is $4\pi R^2$. Thus, we find:

$$\oint_{\mathcal{S}} \mathbf{E}(\mathbf{r})\cdot d\mathbf{s} = \frac{q}{\epsilon} \tag{2.4.2}$$

The integral of a vector field over a specified surface is known as *flux* (see "Additional Reading" at the end of this section). Thus, we have found that the flux of $\mathbf{E}$ through the sphere $\mathcal{S}$ is equal to a constant, namely q/ϵ. Furthermore, this constant is the same regardless of the radius R of the sphere. Said differently, the flux of $\mathbf{E}$ is constant with distance, and does not vary as $\mathbf{E}$ itself does. Let us capitalize on this observation by making the following small modification to Equation 2.4.2:

$$\oint_{\mathcal{S}} [\epsilon\mathbf{E}]\cdot d\mathbf{s} = q$$

In other words, the flux of the quantity $\epsilon\mathbf{E}$ is equal to the enclosed charge, regardless of the radius of the sphere over which we are doing the calculation. This leads to the following definition:

electric flux density

The electric flux density $\mathbf{D} = \epsilon\mathbf{E}$, having units of C/m^2, is a description of the electric field in terms of flux, as opposed to force or change in electric potential.

It may appear that $\mathbf{D}$ is redundant information given $\mathbf{E}$ and ϵ, but this is true only in homogeneous media. The concept of electric flux density becomes important – and decidedly not redundant – when we encounter boundaries between media having different permittivities. As we shall see in Section 5.18, boundary conditions on $\mathbf{D}$ constrain the component of the electric field that is perpendicular to the boundary separating two regions. If one ignores the "flux" character of the electric field represented by $\mathbf{D}$ and instead considers only $\mathbf{E}$, then only the *tangential* component of the electric field is constrained. In fact, when one of the two materials comprising the boundary between two material regions is a perfect conductor, then the electric field is *completely determined* by the boundary condition on $\mathbf{D}$. This greatly simplifies the problem of finding the electric field in a region bounded or partially bounded by materials that can be modeled as perfect conductors, including many metals. In particular, this principle makes it easy to analyze capacitors.

We conclude this section with a warning. Even though the SI units for $\mathbf{D}$ are C/m^2, $\mathbf{D}$ describes an electric field and *not* a surface charge density. It is certainly true that one may describe the amount of charge distributed over a surface using units of C/m^2. However, $\mathbf{D}$ is not necessarily a description of *actual* charge, and there is no implication that the source of the electric field is a

distribution of surface charge. On the other hand, it *is* true that **D** can be interpreted as an *equivalent* surface charge density that would give rise to the observed electric field, and in some cases, this equivalent charge density turns out to be the actual charge density.

This page titled 2.4: Electric Flux Density is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **2.4: Electric Flux Density** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

2.5: Magnetic Flux Density

Magnetic flux density is a vector field which we identify using the symbol $\mathbf{B}$ and which has SI units of tesla (T). Before offering a formal definition, it is useful to consider the broader concept of the *magnetic field*.

Magnetic fields are an intrinsic property of some materials, most notably permanent magnets. The basic phenomenon is probably familiar, and is shown in Figure 2.5.1. A bar magnet has "poles" identified as "N" ("north") and "S" ("south"). The N-end of one magnet attracts the S-end of another magnet but repels the N-end of the other magnet and so on. The existence of a vector field is apparent since the observed force acts at a distance and is asserted in a specific direction. In the case of a permanent magnet, the magnetic field arises from mechanisms occurring at the scale of the atoms and electrons comprising the material. These mechanisms require some additional explanation which we shall defer for now.

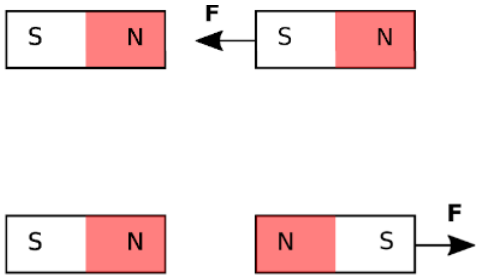


Figure 2.5.1: Evidence of a vector field from observations of the force perceived by the bar magnets on the right in the presence of the bar magnets on the left. (CC BY 3.0; Y. Qin).

Magnetic fields also appear in the presence of current. For example, a coil of wire bearing a current is found to influence permanent magnets (and vice versa) in the same way that permanent magnets affect each other. This is shown in Figure 2.5.2. From this, we infer that the underlying mechanism is the same – i.e., the vector field generated by a current-bearing coil is the same phenomenon as the vector field associated with a permanent magnet. Whatever the source, we are now interested in quantifying its behavior.

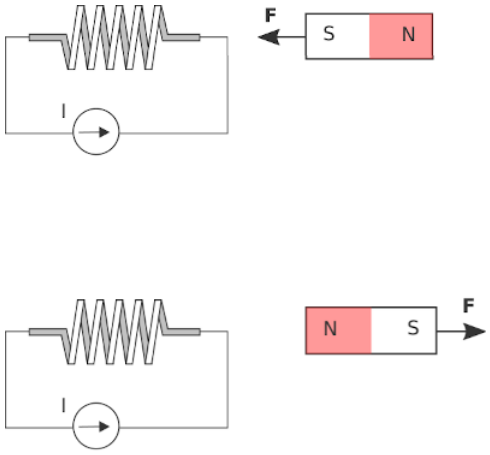


Figure 2.5.2: Evidence that current can also create a magnetic field. (CC BY 4.0; Y. Qin).

To begin, let us consider the effect of a magnetic field on a electrically-charged particle. First, imagine a region of free space with no electric or magnetic fields. Next, imagine that a charged particle appears. This particle will experience no force. Next, a magnetic field appears; perhaps this is due to a permanent magnet or a current in the vicinity. This situation is shown in Figure 2.5.2 (top). Still, no force is applied to the particle. In fact, nothing happens until the particle in set in motion. Figure 2.5.3 (bottom) shows an example. Suddenly, the particle perceives a force. We'll get to the details about direction and magnitude in a moment, but the main idea is now evident. A magnetic field is something that applies a force to a charged particle in motion, distinct from (in fact, in addition to) the force associated with an electric field.

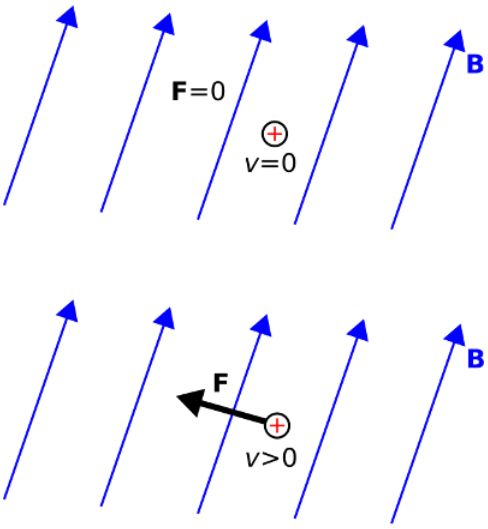


Figure 2.5.3: The force perceived by a charged particle that is (top) motionless and (bottom) moving with velocity $\mathbf{v} = \hat{z}v$, which is perpendicular to the plane of this document and toward the reader (CC BY 4.0; Y. Qing).

Now, it is worth noting that a single charged particle in motion is the simplest form of current. Remember also that motion is required for the magnetic field to influence the particle. Therefore, not only is current the source of the magnetic field, the magnetic field also exerts a force on current. Summarizing:

> The magnetic field describes the force exerted on permanent magnets and currents in the presence of other permanent magnets and currents.

So, how can we quantify a magnetic field? The answer from classical physics involves another experimentally-derived equation that predicts force as a function of charge, velocity, and a vector field $\mathbf{B}$ representing the magnetic field. Here it is: The force applied to a particle bearing charge q is

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B} \tag{2.5.1}$$

where $\mathbf{v}$ is the velocity of the particle and "$\times$" denotes the cross product. The cross product of two vectors is in a direction perpendicular to each of the two vectors, so the force exerted by the magnetic field is perpendicular to both the direction of motion and the direction in which the magnetic field points.

The reader would be well-justified in wondering why the force exerted by the magnetic field should *perpendicular* to $\mathbf{B}$. For that matter, why should the force depend on $\mathbf{v}$? These are questions for which classical physics provides no obvious answers. Effective answers to these questions require concepts from quantum mechanics, where we find that the magnetic field is a manifestation of the fundamental and aptly-named *electromagnetic force*. The electromagnetic force also gives rises to the electric field, and it is only limited intuition, grounded in classical physics, that leads us to perceive the electric and magnetic fields as distinct phenomena. For our present purposes – and for most commonly-encountered engineering applications – we do not require these concepts. It is sufficient to accept this apparent strangeness as fact and proceed accordingly.

Dimensional analysis of 2.5.1 reveals that $\mathbf{B}$ has units of (N·s)/(C·m). In SI, this combination of units is known as the tesla (T).

We refer to $\mathbf{B}$ as *magnetic flux density*, and therefore tesla is a unit of magnetic flux density. A fair question to ask at this point is: What makes this a flux density? The short answer is that this terminology is somewhat arbitrary, and in fact is not even uniformly accepted. In engineering electromagnetics, the preference for referring to $\mathbf{B}$ as a "flux density" is because we frequently find ourselves integrating $\mathbf{B}$ over a mathematical surface. Any quantity that is obtained by integration over a surface is referred to as "flux," and so it becomes natural to think of $\mathbf{B}$ as a flux density; i.e., as flux per unit area. The SI unit for magnetic flux is the weber (Wb). Therefore, $\mathbf{B}$ may alternatively be described as having units of Wb/m^2, and 1 Wb/m^2 = 1 T.

> Magnetic flux density ($\mathbf{B}$, T or Wb/m^2) is a description of the magnetic field that can be defined as the solution to Equation 2.5.1.

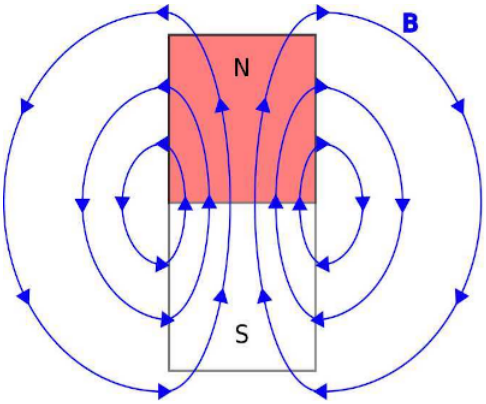


Figure 2.5.4: The magnetic field of a bar magnet, illustrating field lines. (CC BY 4.0; Y. Qing).

When describing magnetic fields, we occasionally refer to the concept of a *field line*, defined as follows:

> A magnetic field line is the curve in space traced out by following the direction in which the magnetic field vector points.

This concept is illustrated in Figure 2.5.4 for a permanent bar magnet and Figure 2.5.5 for a current-bearing coil.

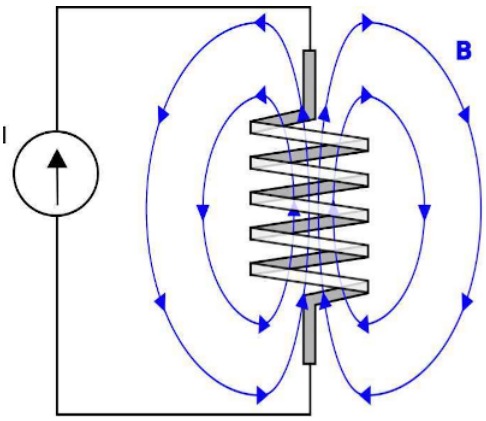


Figure 2.5.5: The magnetic field of a current-bearing coil, illustrating field lines. (CC BY 4.0; Y. Qing).

Magnetic field lines are remarkable for the following reason:

> A magnetic field line always forms a closed loop.

This is true in a sense even for field lines which seem to form straight lines (for example, those along the axis of the bar magnet and the coil in Figures 2.5.4 and 2.5.5, since a field line that travels to infinity in one direction reemerges from infinity in the opposite direction.

This page titled 2.5: Magnetic Flux Density is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **2.5: Magnetic Flux Density** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

2.6: Permeability

Permeability describes the effect of material in determining the magnetic flux density. All else being equal, magnetic flux density increases in proportion to permeability.

To illustrate the concept, consider that a particle bearing charge q moving at velocity $\mathbf{v}$ gives rise to a magnetic flux density:

$$\mathbf{B}(\mathbf{r}) = \mu \frac{q\mathbf{v}}{4\pi R^2} \times \hat{\mathbf{R}} \tag{2.6.1}$$

where $\hat{\mathbf{R}}$ is the unit vector pointing from the charged particle to the field point $\mathbf{r}$, R is this distance, and "$\times$" is the cross product. Note that $\mathbf{B}$ increases with charge and speed, which makes sense since moving charge is the source of the magnetic field. Also note that $\mathbf{B}$ is inversely proportional to $4\pi R^2$, indicating that $|\mathbf{B}|$ decreases in proportion to the area of a sphere surrounding the charge, also known as the *inverse square law*. The remaining factor, μ, is the constant of proportionality that captures the effect of material. We refer to μ as the *permeability* of the material. Since $\mathbf{B}$ can be expressed in units of Wb/m^2 and the units of $\mathbf{v}$ are m/s, we see that μ must have units of henries per meter (H/m). (To see this, note that 1 H $\triangleq$ 1 Wb/A.)

> Permeability (μ, H/m) describes the effect of material in determining the magnetic flux density.

In free space, we find that the permeability $\mu = \mu_0$ where:

$$\mu_0 = 4\pi \times 10^{-7} \text{ H/m}$$

It is common practice to describe the permeability of materials in terms of their *relative permeability*:

$$\mu_r \triangleq \frac{\mu}{\mu_0}$$

which gives the permeability relative to the minimum possible value; i.e., that of free space. Relative permeability for a few representative materials is given in Appendix A2.

Note that μ_r is approximately 1 for all but a small class of materials. These are known as *magnetic materials*, and may exhibit values of μ_r as large as $\sim 10^6$. A commonly-encountered category of magnetic materials is *ferromagnetic* material, of which the best-known example is iron.

This page titled 2.6: Permeability is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **2.6: Permeability** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

2.7: Magnetic Field Intensity

Magnetic field intensity $\mathbf{H}$ is an alternative description of the magnetic field in which the effect of material is factored out. For example, the magnetic flux density $\mathbf{B}$ (reminder: Section 2.5) due to a point charge q moving at velocity $\mathbf{v}$ can be written in terms of the *Biot-Savart Law*:

$$\mathbf{B} = \mu \frac{q\mathbf{v}}{4\pi R^2} \times \hat{\mathbf{R}} \tag{2.7.1}$$

where $\hat{\mathbf{R}}$ is the unit vector pointing from the charged particle to the field point $\mathbf{r}$, R is this distance, "$\times$" is the cross product, and μ is the permeability of the material. We can rewrite Equation 2.7.1 as:

$$\mathbf{B} \triangleq \mu\mathbf{H} \tag{2.7.2}$$

with:

$$\mathbf{H} = \frac{q\mathbf{v}}{4\pi R^2} \times \hat{\mathbf{R}} \tag{2.7.3}$$

so $\mathbf{H}$ in homogeneous media does not depend on μ.

Dimensional analysis of Equation 2.7.3 reveals that the units for $\mathbf{H}$ are amperes per meter (A/m). However, $\mathbf{H}$ does not represent surface current density, as the units might suggest. While it is certainly true that a distribution of current (A) over some linear cross-section (m) can be described as a current density having units of A/m, $\mathbf{H}$ is associated with the magnetic field and not a particular current distribution (the concept of *current density* is not essential to understand this section; however, a primer can be found in Section 6.2). Said differently, $\mathbf{H}$ can be viewed as a description of the magnetic field in terms of an equivalent (but not actual) current.

> The magnetic field intensity $\mathbf{H}$ (A/m), defined using Equation 2.7.2, is a description of the magnetic field independent from material properties.

It may appear that $\mathbf{H}$ is redundant information given $\mathbf{B}$ and μ, but this is true only in homogeneous media. The concept of magnetic field intensity becomes important – and decidedly not redundant – when we encounter boundaries between media having different permeabilities. As we shall see in Section 7.11, boundary conditions on $\mathbf{H}$ constrain the component of the magnetic field which is tangent to the boundary separating two otherwise-homogeneous regions. If one ignores the characteristics of the magnetic field represented by $\mathbf{H}$ and instead considers only $\mathbf{B}$, then only the perpendicular component of the magnetic field is constrained.

The concept of magnetic field intensity also turns out to be useful in a certain problems in which μ is not a constant, but rather is a function of magnetic field strength. In this case, the magnetic behavior of the material is said to be *nonlinear*. For more on this, see Section 7.16.

This page titled 2.7: Magnetic Field Intensity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **2.7: Magnetic Field Intensity** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

2.8: Electromagnetic Properties of Materials

In electromagnetic analysis, one is principally concerned with three properties of matter. These properties are quantified in terms of *constitutive parameters*, which describe the effect of material in determining an electromagnetic quantity in response to a source. Here are the three principal constitutive parameters:

- *Permittivity* (ϵ, F/m) quantifies the effect of matter in determining the electric field in response to electric charge. Permittivity is addressed in Section 2.3.
- *Permeability* (μ, H/m) quantifies the effect of matter in determining the magnetic field in response to current. Permeability is addressed in Section 2.6.
- *Conductivity* (σ, S/m) quantifies the effect of matter in determining the flow of current in response to an electric field. Conductivity is addressed in Section 6.3.

> The electromagnetic properties of most common materials in most common applications can be quantified in terms of the constitutive parameters ϵ, μ, and σ.

To keep electromagnetic theory from becoming too complex, we usually require the constitutive parameters to exhibit a few basic properties. These properties are as follows:

- *Homogeneity*. A material that is *homogeneous* is uniform over the space it occupies; that is, the values of its constitutive parameters are constant at all locations within the material. A counter-example would be a material that is composed of multiple chemically-distinct compounds that are not thoroughly mixed, such as soil.
- *Isotropy*. A material that is *isotropic* behaves in precisely the same way regardless of how it is oriented with respect to sources and fields occupying the same space. A counter-example is quartz, whose atoms are arranged in a uniformly-spaced crystalline lattice. As a result, the electromagnetic properties of quartz can be changed simply by rotating the material with respect to the applied sources and fields.
- *Linearity*. A material is said to be *linear* if its properties are constant and independent of the magnitude of the sources and fields applied to the material. For example, capacitors have capacitance, which is determined in part by the permittivity of the material separating the terminals (Section 5.23). This material is approximately linear when the applied voltage V is below the rated working voltage; i.e., ϵ is constant and so capacitance does not vary significantly with respect to V. When V is greater than the working voltage, the dependence of ϵ on V becomes more pronounced, and then capacitance becomes a function of V. In another practical example, it turns out that μ for ferromagnetic materials is nonlinear such that the precise value of μ depends on the magnitude of the magnetic field.
- *Time-invariance*. An example of a class of materials that is not necessarily time-invariant is piezoelectric materials, for which electromagnetic properties vary significantly depending on the mechanical forces applied to them – a property which can be exploited to make sensors and transducers.

Linearity and time-invariance (LTI) are particularly important properties to consider because they are requirements for *superposition*. For example, in a LTI material, we may calculate the field $\mathbf{E}_1$ due to a point charge q_1 at $\mathbf{r}_1$ and calculate the field $\mathbf{E}_2$ due to a point charge q_2 at $\mathbf{r}_2$. Then, when both charges are simultaneously present, the field is $\mathbf{E}_1 + \mathbf{E}_2$. The same is *not* necessarily true for materials that are not LTI. Devices that are nonlinear, and therefore not LTI, do not necessarily follow the rules of elementary circuit theory, which presume that superposition applies. This condition makes analysis and design much more difficult.

No practical material is truly homogeneous, isotropic, linear, and time-invariant. However, for most materials in most applications, the deviation from this ideal condition is not large enough to significantly affect engineering analysis and design. In other cases, materials may be significantly non-ideal in one of these respects, but may still be analyzed with appropriate modifications to the theory.

This page titled 2.8: Electromagnetic Properties of Materials is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **2.8: Electromagnetic Properties of Materials** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

CHAPTER OVERVIEW

3: Transmission Lines

3.1: Introduction to Transmission Lines
3.2: Types of Transmission Lines
3.3: Transmission Lines as Two-Port Devices
3.4: Lumped-Element Model
3.5: Telegrapher's Equations
3.6: Wave Equation for a TEM Transmission Line
3.7: Characteristic Impedance
3.8: Wave Propagation on a TEM Transmission Line
3.9: Lossless and Low-Loss Transmission Lines
3.10: Coaxial Line
3.11: Microstrip Line
3.12: Voltage Reflection Coefficient
3.13: Standing Waves
3.14: Standing Wave Ratio
3.15: Input Impedance of a Terminated Lossless Transmission Line
3.16: Input Impedance for Open- and Short-Circuit Terminations
3.17: Applications of Open- and Short-Circuited Transmission Line Stubs
3.18: Measurement of Transmission Line Characteristics
3.19: Quarter-Wavelength Transmission Line
3.20: Power Flow on Transmission Lines
3.21: Impedance Matching - General Considerations
3.22: Single-Reactance Matching
3.23: Single-Stub Matching

Contributors and Attributions

- Ellingson, Steven W. (2018) Electromagnetics, Vol. 1. Blacksburg, VA: VT Publishing. https://doi.org/10.21061/electromagnetics-vol-1 Licensed with CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0. Report adoption of this book here. If you are a professor reviewing, adopting, or adapting this textbook please help us understand a little more about your use by filling out this form.

[m0085_Reflection_Coefficient_for_Various_Terminations]

We now consider values of Γ that arise for commonly-encountered terminations.

Matched Load ($Z_L = Z_0$). In this case, the termination may be a device with impedance Z_0, or the termination may be another transmission line having the same characteristic impedance. When $Z_L = Z_0$, $\Gamma = 0$ and there is no reflection.

Open Circuit. An "open circuit" is the absence of a termination. This condition implies $Z_L \to \infty$, and subsequently $\Gamma \to +1$. Since the *current* reflection coefficient is $-\Gamma$, the reflected current wave is 180° out of phase with the incident current wave, making the total current at the open circuit equal to zero, as expected.

Short Circuit. "Short circuit" means $Z_L = 0$, and subsequently $\Gamma = -1$. In this case, the phase of Γ is 180°, and therefore, the potential of the reflected wave cancels the potential of the incident wave at the open circuit, making the total potential equal to zero, as it must be. Since the *current* reflection coefficient is $-\Gamma = +1$ in this case, the reflected current wave is in phase with the incident current wave, and the magnitude of the total current at the short circuit non-zero as expected.

Purely Reactive Load. A purely reactive load, including that presented by a capacitor or inductor, has $Z_L = jX$ where X is reactance. In particular, an inductor is represented by $X > 0$ and a capacitor is represented by $X < 0$. We find

$$\Gamma = \frac{-Z_0 + jX}{+Z_0 + jX} \tag{3.1}$$

The numerator and denominator have the same magnitude, so $|\Gamma| = 1$. Let ϕ be the phase of the denominator ($+Z_0 + jX$). Then, the phase of the numerator is $\pi - \phi$. Subsequently, the phase of Γ is $(\pi - \phi) - \phi = \pi - 2\phi$. Thus, we see that the phase of Γ is no longer limited to be 0° or 180°, but can be any value in between. The phase of reflected wave is subsequently shifted by this amount.

Other Terminations. Any other termination, including series and parallel combinations of any number of devices, can be expressed as a value of Z_L which is, in general, complex-valued. The associated value of $|\Gamma|$ is limited to the range 0 to 1. To see this, note:

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} = \frac{Z_L/Z_0 - 1}{Z_L/Z_0 + 1} \tag{3.2}$$

Note that the smallest possible value of $|\Gamma|$ occurs when the numerator is zero; i.e., when $Z_L = Z_0$. Therefore, the smallest value of $|\Gamma|$ is zero. The largest possible value of $|\Gamma|$ occurs when $Z_L/Z_0 \to \infty$ (i.e., an open circuit) or when $Z_L/Z_0 = 0$ (a short circuit); the result in either case is $|\Gamma| = 1$. Thus,

$$0 \leq |\Gamma| \leq 1 \tag{3.3}$$

This page titled 3: Transmission Lines is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

3.1: Introduction to Transmission Lines

A transmission line is a structure intended to transport electromagnetic signals or power.

A rudimentary transmission line is simply a pair of wires with one wire serving as a datum (i.e., a reference; e.g., “ground”) and the other wire bearing an electrical potential that is defined relative to that datum. Transmission lines having random geometry, such as the test leads shown in Figure 3.1.1, are useful only at very low frequencies and when loss, reactance, and immunity to electromagnetic interference (EMI) are not a concern.



Figure 3.1.1: These leads used to connect test equipment to circuits in a laboratory are a very rudimentary form of transmission line, suitable only for very low frequencies. (Public Domain; Dmitry G)

However, many circuits and systems operate at frequencies where the length or cross-sectional dimensions of the transmission line may be a significant fraction of a wavelength. In this case, the transmission line is no longer “transparent” to the circuits at either end. Furthermore, loss, reactance, and EMI *are* significant problems in many applications. These concerns motivate the use of particular types of transmission lines, and make it necessary to understand how to properly connect the transmission line to the rest of the system.

In electromagnetics, the term “transmission line” refers to a structure which is intended to support a *guided wave*. A guided wave is an electromagnetic wave that is contained within or bound to the line, and which does not radiate away from the line. This condition is normally met if the length and cross-sectional dimensions of the transmission line are small relative to a wavelength – say $\lambda/100$ (i.e., 1% of the wavelength). For example, two randomly-arranged wires might serve well enough to carry a signal at $f = 10$ MHz over a length $l = 3$ cm, since l is only 0.1% of the wavelength $\lambda = c/f = 30$ m. However, if l is increased to 3 m, *or* if f is increased to 1 GHz, then l is now 10% of the wavelength. In this case, one should consider using a transmission line that forms a proper guided wave.

Preventing unintended radiation is not the only concern. Once we have established a guided wave on a transmission line, it is important that power applied to the transmission line be delivered to the circuit or device at the other end and not reflected back into the source. For the random wire $f = 10$ MHz, $l = 3$ cm example above, there is little need for concern, since we expect a phase shift of roughly $0.001 \cdot 360^\circ = 0.36^\circ$ over the length of the transmission line, which is about 0.72° for a round trip. So, to a good approximation, the entire transmission line is at the same electrical potential and thus transparent to the source and destination. However, if l is increased to 3 m, *or* if f is increased to 1 GHz, then the associated round-trip phase shift becomes 72°. In this case, a reflected signal traveling in the opposite direction will add to create a total electrical potential, which varies in both magnitude and phase with position along the line. Thus, the impedance looking toward the destination via the transmission line will be different than the impedance looking toward the destination directly (Section 3.15 gives the details). The modified impedance will depend on the cross-sectional geometry, materials, and length of the line.

Cross-sectional geometry and materials also determine the loss and EMI immunity of the transmission line.

Summarizing:

> Transmission lines are designed to support guided waves with controlled impedance, low loss, and a degree of immunity from EMI.

This page titled 3.1: Introduction to Transmission Lines is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.1: Introduction to Transmission Lines** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.2: Types of Transmission Lines

Two common types of transmission line are **coaxial line** (Figure 3.2.1) and **microstrip line** (Figure 3.2.2). Both are examples of *transverse electromagnetic* (TEM) transmission lines. A TEM line employs a single electromagnetic wave "mode" having electric and magnetic field vectors in directions perpendicular to the axis of the line, as shown in Figures 3.2.3 and 3.2.4. TEM transmission lines appear primarily in radio frequency applications.

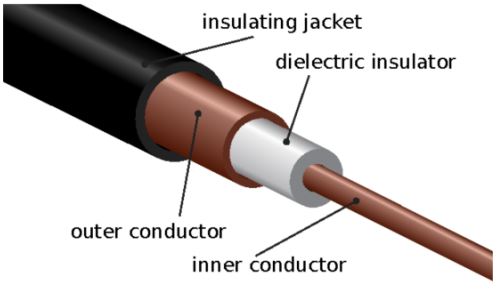


Figure 3.2.1: Structure of a coaxial transmission line. (CC BY 3.0 (modified)).

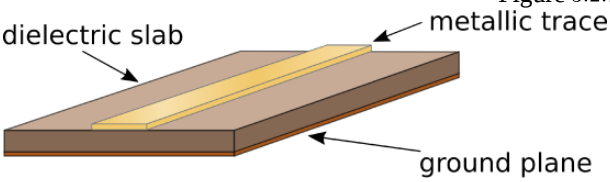


Figure 3.2.2: Structure of a microstrip transmission line. (CC BY SA 3.0 (modified))

> TEM transmission lines such as coaxial lines and microstrip lines are designed to support a single electromagnetic wave that propagates along the length of the transmission line with electric and magnetic field vectors perpendicular to the direction of propagation.

Not all transmission lines exhibit TEM field structure. In non-TEM transmission lines, the electric and magnetic field vectors that are not necessarily perpendicular to the axis of the line, and the structure of the fields is complex relative to the field structure of TEM lines. An example of a transmission line that exhibits non-TEM field structure is the waveguide (see example in Figure 3.2.5). Waveguides are most prevalent at radio frequencies, and tend to appear in applications where it is important to achieve very low loss or where power levels are very high. Another example is common "multimode" optical fiber (Figure 3.2.6). Optical fiber exhibits complex field structure because the wavelength of light is very small compared to the cross-section of the fiber, making the excitation and propagation of non-TEM waves difficult to avoid. (This issue is overcome in a different type of optical fiber, known as "single mode" fiber, which is much more difficult and expensive to manufacture.)

> Higher-order transmission lines, including radio-frequency waveguides and multimode optical fiber, are designed to guide waves that have relatively complex structure.

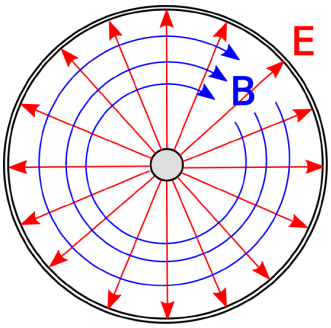


Figure 3.2.3: Structure of the electric and magnetic fields within coaxial line. In this case, the wave is propagating away from the viewer.

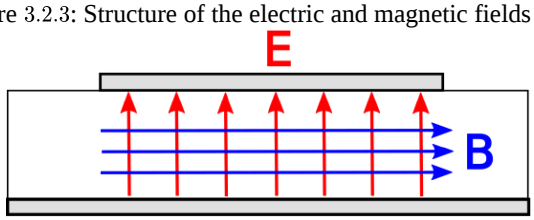


Figure 3.2.4: Structure of the electric and magnetic fields within microstrip line. (The fields *outside* the line are possibly significant, complicated, and not shown.) In this case, the wave is propagating away from the viewer. (CC BY SA 3.0 Unported).



Figure 3.2.5: A network of radio frequency waveguides in an air traffic control radar. (CC BY SA 2.0 Germany)

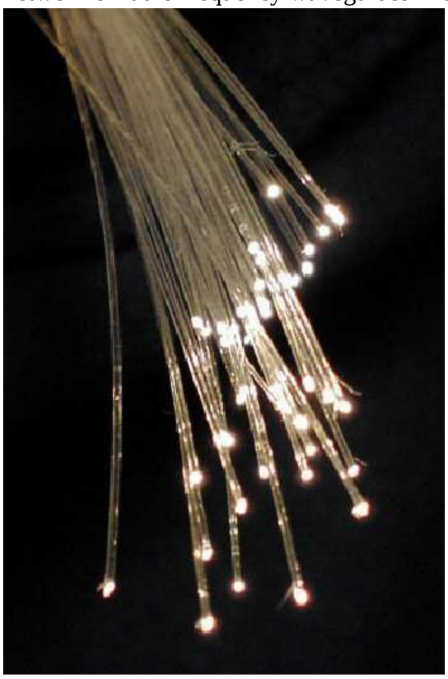


Figure 3.2.6: Strands of optical fiber.

This page titled 3.2: Types of Transmission Lines is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.2: Types of Transmission Lines** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.3: Transmission Lines as Two-Port Devices

Figure 3.3.1 shows common ways to represent transmission lines in circuit diagrams. In each case, the source is represented using a Thévenin equivalent circuit consisting of a voltage source V_S in series with an impedance Z_S.[1] In transmission line analysis, the source may also be referred to as the *generator*. The termination on the receiving end of the transmission line is represented, without loss of generality, as an impedance Z_L. This termination is often referred to as the *load*, although in practice it can be any circuit that exhibits an input impedance of Z_L.

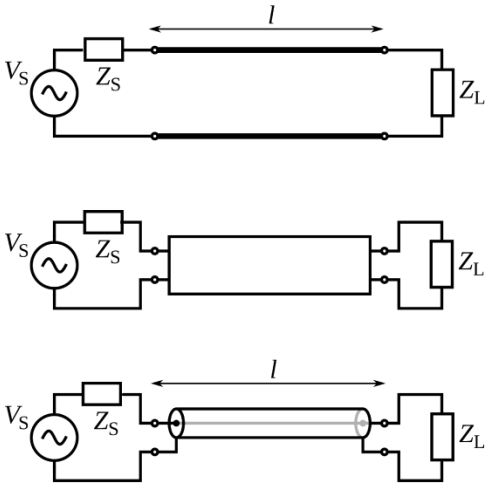


Figure 3.3.1: Symbols representing transmission lines: Top: As a generic two-conductor direct connection. Middle: As a generic two-port "black box." Bottom: As a coaxial cable. © CC BY SA 3.0 Unported (modified)

The two-port representation of a transmission line is completely described by its length l along with some combination of the following parameters:

- Phase propagation constant β, having units of rad/m. This parameter also represents the wavelength in the line through the relationship $\lambda = 2\pi/\beta$. (See Sections 1.3 and 3.8 for details.)
- Attenuation constant α, having units of 1/m. This parameter quantifies the effect of loss in the line. (See Section 3.8 for details.)
- Characteristic impedance Z_0, having units of Ω. This is the ratio of potential ("voltage") to current when the line is perfectly impedance-matched at both ends. (See Section 3.7 for details.)

These parameters depend on the materials and geometry of the line.

Note that a transmission line is typically not transparent to the source and load. In particular, the load impedance may be Z_L, but the impedance presented to the source may or may not be equal to Z_L. (See Section 3.15 for more on this concept.) Similarly, the source impedance may be Z_S, but the impedance presented to the load may or may not be equal to Z_S. The effect of the transmission line on the source and load impedances will depend on the parameters identified above.

1. For a refresher on this concept, see "Additional Reading" at the end of this section.↩

This page titled 3.3: Transmission Lines as Two-Port Devices is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.3: Transmission Lines as Two-Port Devices** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.4: Lumped-Element Model

It is possible to ascertain the relevant behaviors of a transmission line using elementary circuit theory applied to a differential-length lumped-element model of the transmission line. The concept is illustrated in Figure 3.4.1, which shows a generic transmission line aligned with its length along the z axis. The transmission line is divided into segments having small but finite length Δz. Each segment is modeled as an identical two-port having the equivalent circuit representation shown in Figure 3.4.2. The equivalent circuit consists of 4 components as follows:

- The resistance $R'\Delta z$ represents the series-combined ohmic resistance of the two conductors. This should account for *both* conductors since the current in the actual transmission line must flow through both conductors. The prime notation reminds us that R' is resistance *per unit length*; i.e., Ω/m, and it is only after multiplying by length that we get a resistance in Ω.
- The conductance $G'\Delta z$ represents the leakage of current directly from one conductor to the other. When $G'\Delta z > 0$, the resistance between the conductors is less than infinite, and therefore, current may flow between the conductors. This amounts to a loss of power separate from the loss associated with R' above. G' has units of S/m. Further note that G' is *not* equal to $1/R'$ as defined above. G' and R' are describing entirely different physical mechanisms (and in principle *either* could be defined as either a resistance or a conductance).
- The capacitance $C'\Delta z$ represents the capacitance of the transmission line structure. Capacitance is the tendency to store energy in electric fields and depends on the cross-sectional geometry and the media separating the conductors. C' has units of F/m.
- The inductance $L'\Delta z$ represents the inductance of the transmission line structure. Inductance is the tendency to store energy in magnetic fields, and (like capacitance) depends on the cross-sectional geometry and the media separating the conductors. L' has units of H/m.

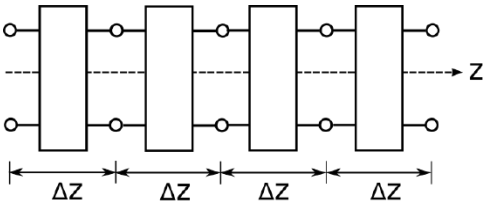


Figure 3.4.1: Interpretation of a transmission line as a cascade of discrete series-connected two-ports.

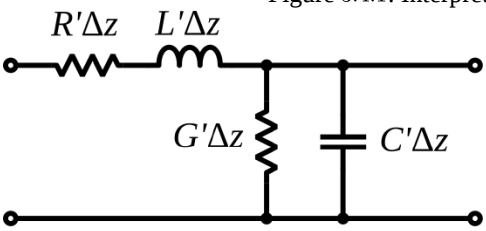


Figure 3.4.1: Lumped-element equivalent circuit model for each of the two-ports in Figure 3.4.2. (CC BY SA 3.0 Unported (modified))

In order to use the model, one must have values for R', G', C', and L'. Methods for computing these parameters are addressed elsewhere in this book.

This page titled 3.4: Lumped-Element Model is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.4: Lumped-Element Model** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.5: Telegrapher's Equations

In this section, we derive the equations that govern the potential $v(z,t)$ and current $i(z,t)$ along a transmission line that is oriented along the z axis. For this, we will employ the lumped-element model developed in Section 3.4.

To begin, we define voltages and currents as shown in Figure 3.5.1.

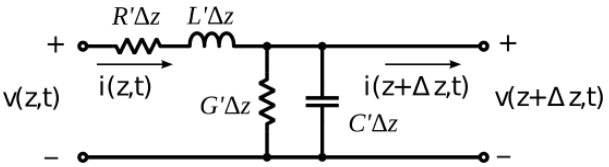


Figure 3.5.1: Lumped-element equivalent circuit transmission line model, annotated with sign conventions for potentials and currents. (© CC BY SA 3.0 Unported (modified))

We assign the variables $v(z,t)$ and $i(z,t)$ to represent the potential and current on the left side of the segment, with reference polarity and direction as shown in the figure. Similarly we assign the variables $v(z+\Delta z,t)$ and $i(z+\Delta z,t)$ to represent the potential and current on the right side of the segment, again with reference polarity and direction as shown in the figure. Applying Kirchoff's voltage law from the left port, through $R'\Delta z$ and $L'\Delta z$, and returning via the right port, we obtain:

$$\begin{gathered} v(z,t) - (R'\Delta z)\, i(z,t) - (L'\Delta z)\, \frac{\partial}{\partial t} i(z,t) \\ -v(z+\Delta z,t) = 0 \end{gathered}$$

Moving terms referring to current to the right side of the equation and then dividing through by Δz, we obtain

$$\begin{gathered} -\frac{v(z+\Delta z,t) - v(z,t)}{\Delta z} = \\ R'\ i(z,t) + L'\ \frac{\partial}{\partial t} i(z,t) \end{gathered}$$

Then taking the limit as $\Delta z \to 0$:

$$\boxed{-\frac{\partial}{\partial z} v(z,t) = R'\ i(z,t) + L'\ \frac{\partial}{\partial t} i(z,t)} \tag{3.5.1}$$

Applying Kirchoff's current law at the right port, we obtain:

$$\begin{gathered} i(z,t) - (G'\Delta z)\, v(z+\Delta z,t) - (C'\Delta z)\, \frac{\partial}{\partial t} v(z+\Delta z,t) \\ -i(z+\Delta z,t) = 0 \end{gathered}$$

Moving terms referring to potential to the right side of the equation and then dividing through by Δz, we obtain

$$\begin{gathered} -\frac{i(z+\Delta z,t) - i(z,t)}{\Delta z} = \\ G'\ v(z+\Delta z,t) + C'\ \frac{\partial}{\partial t} v(z+\Delta z,t) \end{gathered}$$

Taking the limit as $\Delta z \to 0$:

$$\boxed{-\frac{\partial}{\partial z} i(z,t) = G'\ v(z,t) + C'\ \frac{\partial}{\partial t} v(z,t)} \tag{3.5.2}$$

> Equations 3.5.1 and 3.5.2 are the *telegrapher's equations*. These coupled (simultaneous) differential equations can be solved for $v(z,t)$ and $i(z,t)$ given R', G', L', C' and suitable boundary conditions.

The time-domain telegrapher's equations are usually more than we need or want. If we are only interested in the response to a sinusoidal stimulus, then considerable simplification is possible using phasor representation.[1] First we define phasors $\widetilde{V}(z)$ and $\tilde{I}(z)$ through the usual relationship:

$$v(z,t) = \mathrm{Re}\left\{\widetilde{V}(z)\ e^{j\omega t}\right\}$$

$$i(z,t) = \text{Re}\left\{ I(z)\ e^{j\omega t} \right\}$$

Now we see:

$$\begin{aligned} \frac{\partial}{\partial z} v(z,t) &= \frac{\partial}{\partial z} \text{Re}\left\{ \widetilde{V}(z)\ e^{j\omega t} \right\} \\ &= \text{Re}\left\{ \left[\frac{\partial}{\partial z} \widetilde{V}(z) \right]\ e^{j\omega t} \right\} \end{aligned}$$

In other words, $\partial v(z,t)/\partial z$ expressed in phasor representation is simply $\partial \widetilde{V}(z)/\partial z$; and

$$\begin{aligned} \frac{\partial}{\partial t} i(z,t) &= \frac{\partial}{\partial t} \text{Re}\left\{ \tilde{I}(z)\ e^{j\omega t} \right\} \\ &= \text{Re}\left\{ \frac{\partial}{\partial t} \left[\tilde{I}(z) e^{j\omega t} \right] \right\} \\ &= \text{Re}\left\{ \left[j\omega \tilde{I}(z) \right] e^{j\omega t} \right\} \end{aligned}$$

In other words, $\partial i(z,t)/\partial t$ expressed in phasor representation is $j\omega \tilde{I}(z)$. Therefore, Equation 3.5.1 expressed in phasor representation is:

$$\boxed{-\frac{\partial}{\partial z} \widetilde{V}(z) = \left[R' + j\omega L' \right]\ \tilde{I}(z)} \tag{3.5.3}$$

Following the same procedure, Equation 3.5.2 expressed in phasor representation is found to be:

$$\boxed{-\frac{\partial}{\partial z} \tilde{I}(z) = \left[G' + j\omega C' \right]\ \widetilde{V}(z)} \tag{3.5.4}$$

> Equations 3.5.3 and 3.5.4 are the telegrapher's equations in phasor representation.

The principal advantage of these equations over the time-domain versions is that we no longer need to contend with derivatives with respect to time – only derivatives with respect to distance remain. This considerably simplifies the equations.

---

1. For a refresher on phasor analysis, see Section 1.5.↩

This page titled 3.5: Telegrapher's Equations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.5: Telegrapher's Equations** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.6: Wave Equation for a TEM Transmission Line

Consider a TEM transmission line aligned along the z axis. The phasor form of the Telegrapher's Equations (Section 3.5) relate the potential phasor $\widetilde{V}(z)$ and the current phasor $\tilde{I}(z)$ to each other and to the lumped-element model equivalent circuit parameters R', G', C', and L'. These equations are

$$-\frac{\partial}{\partial z}\widetilde{V}(z) = [R' + j\omega L']\ \tilde{I}(z) \tag{3.6.1}$$

$$-\frac{\partial}{\partial z}\tilde{I}(z) = [G' + j\omega C']\ \widetilde{V}(z) \tag{3.6.2}$$

An obstacle to using these equations is that we require both equations to solve for either the potential or the current. In this section, we reduce these equations to a single equation – a *wave equation* – that is more convenient to use and provides some additional physical insight.

We begin by differentiating both sides of Equation 3.6.1 with respect to z, yielding:

$$-\frac{\partial^2}{\partial z^2}\widetilde{V}(z) = [R' + j\omega L']\ \frac{\partial}{\partial z}\tilde{I}(z)$$

Then using Equation 3.6.2 to eliminate $\tilde{I}(z)$, we obtain

$$-\frac{\partial^2}{\partial z^2}\widetilde{V}(z) = -[R' + j\omega L'][G' + j\omega C']\ \widetilde{V}(z)$$

This equation is normally written as follows:

$$\boxed{\frac{\partial^2}{\partial z^2}\widetilde{V}(z) - \gamma^2\ \widetilde{V}(z) = 0} \tag{3.6.3}$$

where we have made the substitution:

$$\gamma^2 = (R' + j\omega L')(G' + j\omega C')$$

The principal square root of γ^2 is known as the *propagation constant*:

$$\gamma \triangleq \sqrt{(R' + j\omega L')(G' + j\omega C')} \tag{3.6.4}$$

> The *propagation constant* γ (units of m^{-1}) captures the effect of materials, geometry, and frequency in determining the variation in potential and current with distance on a TEM transmission line.

Following essentially the same procedure but beginning with Equation 3.6.2, we obtain

$$\boxed{\frac{\partial^2}{\partial z^2}\tilde{I}(z) - \gamma^2\ \tilde{I}(z) = 0} \tag{3.6.5}$$

> Equations 3.6.3 and 3.6.5 are the *wave equations* for $\widetilde{V}(z)$ and $\tilde{I}(z)$, respectively.

Note that both $\widetilde{V}(z)$ and $\tilde{I}(z)$ satisfy the *same* linear homogeneous differential equation. This does *not* mean that $\widetilde{V}(z)$ and $\tilde{I}(z)$ are equal. Rather, it means that $\widetilde{V}(z)$ and $\tilde{I}(z)$ can differ by no more than a multiplicative constant. Since $\widetilde{V}(z)$ is potential and $\tilde{I}(z)$ is current, that constant must be an impedance. This impedance is known as the *characteristic impedance* and is determined in Section 3.7.

The general solutions to Equations 3.6.3 and 3.6.5 are

$$\widetilde{V}(z) = V_0^+ e^{-\gamma z} + V_0^- e^{+\gamma z} \tag{3.6.6}$$

$$\tilde{I}(z) = I_0^+ e^{-\gamma z} + I_0^- e^{+\gamma z} \tag{3.6.7}$$

where V_0^+, V_0^-, I_0^+, and I_0^- are complex-valued constants. It is shown in Section 3.8 that Equations 3.6.6 and 3.6.7 represent sinusoidal waves propagating in the $+z$ and $-z$ directions along the length of the line. The constants may represent sources, loads, or simply discontinuities in the materials and/or geometry of the line. The values of the constants are determined by boundary conditions; i.e., constraints on $\widetilde{V}(z)$ and $\tilde{I}(z)$ at some position(s) along the line.

The reader is encouraged to verify that the Equations 3.6.6 and 3.6.7 are in fact solutions to Equations 3.6.3 and 3.6.5, respectively, for any values of the constants V_0^+, V_0^-, I_0^+, and I_0^-.

This page titled 3.6: Wave Equation for a TEM Transmission Line is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.6: Wave Equation for a TEM Transmission Line** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.7: Characteristic Impedance

Characteristic impedance is the ratio of voltage to current for a wave that is propagating in single direction on a transmission line. This is an important parameter in the analysis and design of circuits and systems using transmission lines. In this section, we formally define this parameter and derive an expression for this parameter in terms of the equivalent circuit model introduced in Section 3.4.

Consider a transmission line aligned along the z axis. Employing some results from Section 3.6, recall that the phasor form of the wave equation in this case is

$$\frac{\partial^2}{\partial z^2}\widetilde{V}(z) - \gamma^2\ \widetilde{V}(z) = 0 \tag{3.7.1}$$

where

$$\gamma \triangleq \sqrt{(R' + j\omega L')\,(G' + j\omega C')} \tag{3.7.2}$$

Equation 3.7.1 relates the potential phasor $\widetilde{V}(z)$ to the equivalent circuit parameters R', G', C', and L'. An equation of the same form relates the current phasor $\tilde{I}(z)$ to the equivalent circuit parameters:

$$\frac{\partial^2}{\partial z^2}\tilde{I}(z) - \gamma^2\ \tilde{I}(z) = 0 \tag{3.7.3}$$

Since both $\widetilde{V}(z)$ and $\tilde{I}(z)$ satisfy the *same* linear homogeneous differential equation, they may differ by no more than a multiplicative constant. Since $\widetilde{V}(z)$ is potential and $\tilde{I}(z)$ is current, that constant can be expressed in units of impedance. Specifically, this is the *characteristic impedance*, so-named because it depends only on the materials and cross-sectional geometry of the transmission line – i.e., things which determine γ – and not length, excitation, termination, or position along the line.

To derive the characteristic impedance, first recall that the general solutions to Equations 3.7.1 and 3.7.3 are

$$\widetilde{V}(z) = V_0^+ e^{-\gamma z} + V_0^- e^{+\gamma z} \tag{3.7.4}$$

$$\tilde{I}(z) = I_0^+ e^{-\gamma z} + I_0^- e^{+\gamma z} \tag{3.7.5}$$

where V_0^+, V_0^-, I_0^+, and I_0^- are complex-valued constants whose values are determined by boundary conditions; i.e., constraints on $\widetilde{V}(z)$ and $\tilde{I}(z)$ at some position(s) along the line. Also, we will make use of the telegrapher's equations (Section 3.5):

$$-\frac{\partial}{\partial z}\widetilde{V}(z) = [R' + j\omega L']\ \tilde{I}(z) \tag{3.7.6}$$

$$-\frac{\partial}{\partial z}\tilde{I}(z) = [G' + j\omega C']\ \widetilde{V}(z) \tag{3.7.7}$$

We begin by differentiating Equation 3.7.4 with respect to z, which yields

$$\frac{\partial}{\partial z}\widetilde{V}(z) = -\gamma\,[V_0^+ e^{-\gamma z} - V_0^- e^{+\gamma z}]$$

Now we use this to eliminate $\partial\widetilde{V}(z)/\partial z$ in Equation 3.7.6, yielding

$$\gamma\,[V_0^+ e^{-\gamma z} - V_0^- e^{+\gamma z}] = [R' + j\omega L']\ \tilde{I}(z)$$

Solving the above equation for $\tilde{I}(z)$ yields:

$$\tilde{I}(z) = \frac{\gamma}{R' + j\omega L'}\,[V_0^+ e^{-\gamma z} - V_0^- e^{+\gamma z}]$$

Comparing this to Equation 3.7.5, we note

$$I_0^+ = \frac{\gamma}{R' + j\omega L'}\ V_0^+$$

$$I_0^- = \frac{-\gamma}{R' + j\omega L'}\ V_0^-$$

We now make the substitution

$$Z_0 = \frac{R' + j\omega L'}{\gamma} \tag{3.7.8}$$

and observe

$$\boxed{\frac{V_0^+}{I_0^+} = \frac{-V_0^-}{I_0^-} \triangleq Z_0}$$

As anticipated, we have found that coefficients in the equations for potentials and currents are related by an impedance, namely, Z_0.Characteristic impedance can be written entirely in terms of the equivalent circuit parameters by substituting Equation 3.7.2 into Equation 3.7.8, yielding:

$$\boxed{Z_0 = \sqrt{\frac{R' + j\omega L'}{G' + j\omega C'}}}$$

> The characteristic impedance Z_0 (Ω) is the ratio of potential to current in a wave traveling in a single direction along the transmission line.

Take care to note that Z_0 is *not* the ratio of $\widetilde{V}(z)$ to $\tilde{I}(z)$ in general; rather, Z_0 relates only the potential and current waves traveling in the *same* direction.

Finally, note that transmission lines are normally designed to have a characteristic impedance that is completely real-valued – that is, with no imaginary component. This is because the imaginary component of an impedance represents energy *storage* (think of capacitors and inductors), whereas the purpose of a transmission line is energy *transfer*.

This page titled 3.7: Characteristic Impedance is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.7: Characteristic Impedance** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.8: Wave Propagation on a TEM Transmission Line

In Section 3.6, it is shown that expressions for the phasor representations of the potential and current along a transmission line are

$$\widetilde{V}(z) = V_0^+ e^{-\gamma z} + V_0^- e^{+\gamma z} \tag{3.8.1}$$

$$\tilde{I}(z) = I_0^+ e^{-\gamma z} + I_0^- e^{+\gamma z} \tag{3.8.2}$$

where γ is the propagation constant and it assumed that the transmission line is aligned along the z axis. In this section, we demonstrate that these expressions represent sinusoidal waves, and point out some important features. Before attempting this section, the reader should be familiar with the contents of Sections 3.4, 3.6 and 3.7. A refresher on fundamental wave concepts (Section 1.3) may also be helpful.

We first define real-valued quantities α and β to be the real and imaginary components of γ; i.e.,

$$\alpha \triangleq \mathrm{Re}\{\gamma\}$$

$$\beta \triangleq \mathrm{Im}\{\gamma\}$$

and subsequently

$$\gamma = \alpha + j\beta$$

Then we observe

$$e^{\pm\gamma z} = e^{\pm(\alpha+j\beta)z} = e^{\pm\alpha z}\, e^{\pm j\beta z}$$

It may be easier to interpret this expression by reverting to the time domain:

$$\mathrm{Re}\left\{e^{\pm\gamma z} e^{j\omega t}\right\} = e^{\pm\alpha z}\cos(\omega t \pm \beta z)$$

Thus, $e^{-\gamma z}$ represents a damped sinusoidal wave traveling in the $+z$ direction, and $e^{+\gamma z}$ represents a damped sinusoidal wave traveling in the $-z$ direction.

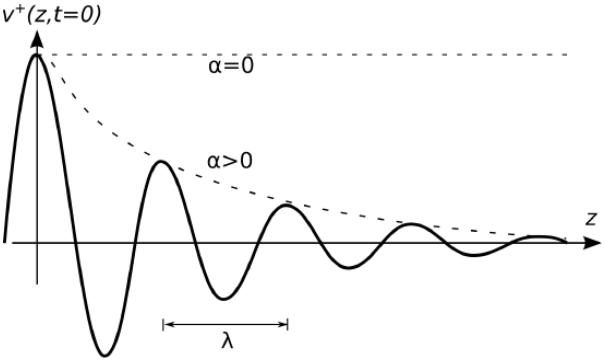


Figure 3.8.1: The potential $v^+(z,t)$ of the wave traveling in the $+z$ direction at $t = 0$ for $\psi = 0$.

Let's define $\widetilde{V}^+(z)$ and $\tilde{I}^+(z)$ to be the potential and current associated with a wave propagating in the $+z$ direction. Then:

$$\widetilde{V}^+(z) \triangleq V_0^+ e^{-\gamma z}$$

or equivalently in the time domain:

$$\begin{aligned} v^+(z,t) &= \mathrm{Re}\left\{\widetilde{V}^+(z)e^{j\omega t}\right\} \\ &= \mathrm{Re}\left\{V_0^+ e^{-\gamma z} e^{j\omega t}\right\} \\ &= \left|V_0^+\right| e^{-\alpha z}\cos(\omega t - \beta z + \psi) \end{aligned} \tag{3.8.3}$$

where ψ is the phase of V_0^+. Figure 3.8.1 shows $v^+(z,t)$. From fundamental wave theory we recognize

> $\beta \triangleq \mathrm{Im}\{\gamma\}$ (rad/m) is the *phase propagation constant*, which is the rate at which phase changes as a function of distance.

Subsequently the wavelength in the line is

$$\boxed{\lambda = \frac{2\pi}{\beta}}$$

Also we recognize:

> $\alpha \triangleq \mathrm{Re}\{\gamma\}$ (1/m) is the *attenuation constant*, which is the rate at which magnitude diminishes as a function of distance.

Sometimes the units of α are indicated as "Np/m" ("nepers" per meter), where the term "neper" is used to indicate the units of the otherwise unitless real-valued exponent of the constant e.

Note that $\alpha = 0$ for a wave that does not diminish in magnitude with increasing distance, in which case the transmission line is said to be *lossless*. If $\alpha > 0$ then the line is said to be *lossy* (or possibly "low loss" if the loss can be neglected), and in this case the rate at which the magnitude decreases with distance increases with α.

Next let us consider the speed of the wave. To answer this question, we need to be a bit more specific about what we mean by "speed." At the moment, we mean phase velocity; that is, the speed at which a point of constant phase seems to move through space. In other words, what distance Δz does a point of constant phase traverse in time Δt? To answer this question, we first note that the phase of $v^+(z,t)$ can be written generally as

$$\omega t - \beta z + \phi$$

where ϕ is some constant. Similarly, the phase at some time Δt later and some point Δz further along can be written as

$$\omega\left(t + \Delta t\right) - \beta\left(z + \Delta z\right) + \phi$$

The phase velocity v_p is $\Delta z/\Delta t$ when these two phases are equal; i.e., when

$$\omega t - \beta z + \phi = \omega\left(t + \Delta t\right) - \beta\left(z + \Delta z\right) + \phi$$

Solving for $v_p = \Delta z/\Delta t$, we obtain:

$$\boxed{v_p = \frac{\omega}{\beta}}$$

Having previously noted that $\beta = 2\pi/\lambda$, the above expression also yields the expected result

$$\boxed{v_p = \lambda f}$$

> The phase velocity $v_p = \omega/\beta = \lambda f$ is the speed at which a point of constant phase travels along the line.

Returning now to consider the current associated with the wave traveling in the $+z$ direction:

$$\tilde{I}^+(z) = I_0^+ e^{-\gamma z}$$

We can rewrite this expression in terms of the characteristic impedance Z_0, as follows:

$$\tilde{I}^+(z) = \frac{V_0^+}{Z_0} e^{-\gamma z}$$

Similarly, we find that the current $\tilde{I}^-(z)$ associated with $\widetilde{V}^-(z)$ for the wave traveling in the $-z$ direction is

$$\tilde{I}^-(z) = \frac{-V_0^-}{Z_0} e^{+\gamma z}$$

The negative sign appearing in the above expression emerges as a result of the sign conventions used for potential and current in the derivation of the telegrapher's equations (Section 3.5). The physical significance of this change of sign is that wherever the potential of the wave traveling in the $-z$ direction is positive, then the current at the same point is flowing in the $-z$ direction.

It is frequently necessary to consider the possibility that waves travel in both directions simultaneously. A very important case where this arises is when there is reflection from a discontinuity of some kind; e.g., from a termination which is not perfectly impedance-matched. In this case, the total potential $\widetilde{V}(z)$ and total current $\tilde{I}(z)$ can be expressed as the general solution to the wave equation; i.e., as the sum of the "incident" ($+z$-traveling) wave and the reflected ($-z$-traveling) waves:

$$\widetilde{V}(z) = \widetilde{V}^{+}(z) + \widetilde{V}^{-}(z)$$

$$\tilde{I}(z) = \tilde{I}^{+}(z) + \tilde{I}^{-}(z)$$

The existence of waves propagating simultaneously in both directions gives rise to a phenomenon known as a *standing wave*. Standing waves and the calculation of the coefficients V_0^- and I_0^- due to reflection are addressed in Sections 3.13 and 3.12 respectively.

This page titled 3.8: Wave Propagation on a TEM Transmission Line is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.8: Wave Propagation on a TEM Transmission Line** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.9: Lossless and Low-Loss Transmission Lines

Quite often the loss in a transmission line is small enough that it may be neglected. In this case, several aspects of transmission line theory may be simplified. In this section, we present these simplifications.

First, recall that "loss" refers to the reduction of magnitude as a wave propagates through space. In the lumped-element equivalent circuit model (Section 3.4), the parameters R' and G' of the represent physical mechanisms associated with loss. Specifically, R' represents the resistance of conductors, whereas G' represents the undesirable current induced between conductors through the spacing material. Also recall that the propagation constant γ is, in general, given by

$$\gamma \triangleq \sqrt{(R' + j\omega L')\,(G' + j\omega C')}$$

With this in mind, we now define "low loss" as meeting the conditions:

$$R' \ll \omega L' \tag{3.9.1}$$

$$G' \ll \omega C' \tag{3.9.2}$$

When these conditions are met, the propagation constant simplifies as follows:

$$\begin{aligned} \gamma &\approx \sqrt{(j\omega L')\,(j\omega C')} \\ &= \sqrt{-\omega^2 L' C'} \\ &= j\omega\sqrt{L'C'} \end{aligned} \tag{3.9.3}$$

and subsequently

$$\alpha \triangleq \mathrm{Re}\{\gamma\} \approx 0 \qquad \text{(low-loss approx.)} \tag{3.9.4}$$

$$\beta \triangleq \mathrm{Im}\{\gamma\} \approx \omega\sqrt{L'C'} \qquad \text{(low-loss approx.)} \tag{3.9.5}$$

$$v_p = \omega/\beta \approx \frac{1}{\sqrt{L'C'}} \qquad \text{(low-loss approx.)} \tag{3.9.6}$$

Similarly:

$$Z_0 = \sqrt{\frac{R' + j\omega L'}{G' + j\omega C'}} \approx \sqrt{\frac{L'}{C'}} \quad \text{(low-loss approx.)}$$

Of course if the line is strictly lossless (i.e., $R' = G' = 0$) then these are not approximations, but rather the exact expressions.

In practice, these approximations are quite commonly used, since practical transmission lines typically meet the conditions expressed in Inequalities 3.9.1 and 3.9.2 and the resulting expressions are much simpler. We further observe that Z_0 and v_p are approximately independent of frequency when these conditions hold.

However, also note that "low loss" does not mean "no loss," and it is common to apply these expressions even when R' and/or G' is large enough to yield significant loss. For example, a coaxial cable used to connect an antenna on a tower to a radio near the ground typically has loss that is important to consider in the analysis and design process, but nevertheless satisfies Equations 3.9.1 and 3.9.2. In this case, the low-loss expression for β is used, but α might not be approximated as zero.

This page titled 3.9: Lossless and Low-Loss Transmission Lines is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.9: Lossless and Low-Loss Transmission Lines** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.10: Coaxial Line

Coaxial transmission lines consists of metallic inner and outer conductors separated by a spacer material as shown in Figure 3.10.1. The spacer material is typically a low-loss dielectric material having permeability approximately equal to that of free space ($\mu \approx \mu_0$) and permittivity ϵ_s that may range from very near ϵ_0 (e.g., air-filled line) to 2–3 times ϵ_0. The outer conductor is alternatively referred to as the "shield," since it typically provides a high degree of isolation from nearby objects and electromagnetic fields. Coaxial line is single-ended[1] in the sense that the conductor geometry is asymmetric and the shield is normally attached to ground at both ends. These characteristics make coaxial line attractive for connecting single-ended circuits in widely-separated locations and for connecting antennas to receivers and transmitters.

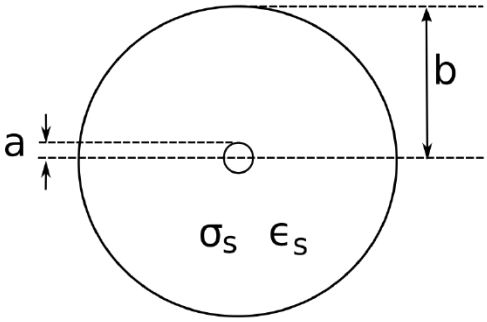


Figure 3.10.1: Cross-section of a coaxial transmission line, indicating design parameters.

Coaxial lines exhibit TEM field structure as shown in Figure 3.10.2.

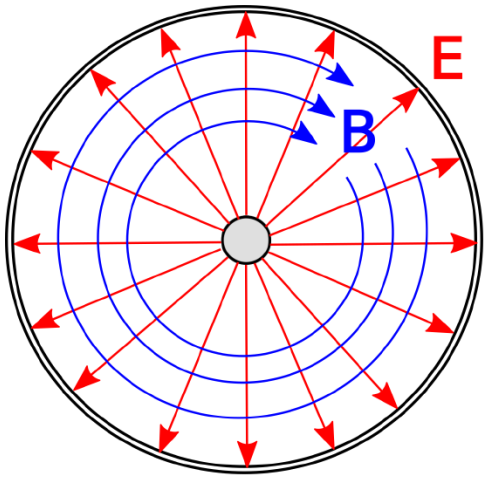


Figure 3.10.2: Structure of the electric and magnetic fields within coaxial line. In this case, the wave is propagating away from the viewer.

Expressions for the equivalent circuit parameters C' and L' for coaxial lines can be obtained from basic electromagnetic theory. It is shown in Section 5.24 that the capacitance per unit length is

$$C' = \frac{2\pi\epsilon_s}{\ln(b/a)} \tag{3.10.1}$$

where a and b are the radii of the inner and outer conductors, respectively. Using analysis shown in Section 7.14, the inductance per unit length is

$$L' = \frac{\mu_0}{2\pi}\ln\left(\frac{b}{a}\right) \tag{3.10.2}$$

The loss conductance G' depends on the conductance σ_s of the spacer material, and is given by

$$G' = \frac{2\pi\sigma_s}{\ln(b/a)}$$

This expression is derived in Section 6.5.

The resistance per unit length, R', is relatively difficult to quantify. One obstacle is that the inner and outer conductors typically consist of different materials or compositions of materials. The inner conductor is not necessarily a single homogeneous material; instead, the inner conductor may consist of a variety of materials selected by trading-off between conductivity, strength, weight, and cost. Similarly, the outer conductor is not necessarily homogeneous; for a variety of reasons, the outer conductor may instead be a metal mesh, a braid, or a composite of materials. Another complicating factor is that the resistance of the conductor varies significantly with frequency, whereas C', L', and G' exhibit relatively little variation from their electro- and magnetostatic values. These factors make it difficult to devise a single expression for R' that is both as simple as those shown above for the other parameters and generally applicable. Fortunately, it turns out that the low-loss conditions $R' \ll \omega L'$ and $G' \ll \omega C'$ are often applicable,[2] so that R' and G' are important only if it is necessary to compute loss.

Since the low-loss conditions are often met, a convenient expression for the characteristic impedance is obtained from Equations 3.10.1 and 3.10.2 for L' and C' respectively:

$$\begin{aligned} Z_0 &\approx \sqrt{\frac{L'}{C'}} (\text{ low-loss }) \\ &= \frac{1}{2\pi}\sqrt{\frac{\mu_0}{\epsilon_s}} \ln\frac{b}{a} \end{aligned} \tag{3.10.3}$$

The spacer permittivity can be expressed as $\epsilon_s = \epsilon_r \epsilon_0$ where ϵ_r is the relative permittivity of the spacer material. Since $\sqrt{\mu_0/\epsilon_0}$ is a constant, the above expression is commonly written

$$Z_0 \approx \frac{60\ \Omega}{\sqrt{\epsilon_r}} \ln\frac{b}{a} \quad \text{(low-loss)}$$

Thus, it is possible to express Z_0 directly in terms of parameters describing the geometry (a and b) and material (ϵ_r) used in the line, without the need to first compute the values of components in the lumped-element equivalent circuit model.

Similarly, the low-loss approximation makes it possible to express the phase velocity ν_p directly in terms the spacer permittivity:

$$\begin{aligned} \nu_p &\approx \frac{1}{\sqrt{L'C'}} \text{ (low-loss)} \\ &= \frac{c}{\sqrt{\epsilon_r}} \end{aligned} \tag{3.10.4}$$

since $c \triangleq 1/\sqrt{\mu_0\epsilon_0}$. In other words, the phase velocity in a low-loss coaxial line is approximately equal to the speed of electromagnetic propagation in free space, divided by the square root of the relative permittivity of the spacer material. Therefore, the phase velocity in an air-filled coaxial line is approximately equal to speed of propagation in free space, but is reduced in a coaxial line using a dielectric spacer.

Example 3.10.1: RG-59 Coaxial Cable

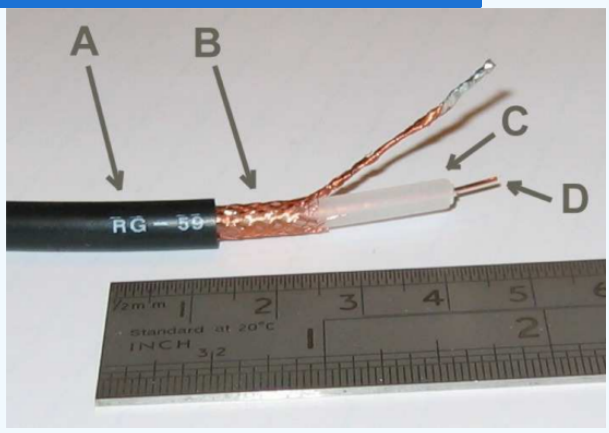


Figure 3.10.3: RG-59 coaxial line. A: Insulating jacket. B: Braided outer conductor. C: Dielectric spacer. D: Inner conductor (© CC BY SA 3.0; Arj)

RG-59 is a very common type of coaxial line. Figure 3.10.3 shows a section of RG-59 cut away so as to reveal its structure. The radii are $a \cong 0.292$ mm and $b \approx 1.855$ mm (mean), yielding $L' \approx 370$ nH/m. The spacer material is polyethylene having $\epsilon_r \cong 2.25$, yielding $C' \approx 67.7$ pF/m. The conductivity of polyethylene is $\sigma_s \cong 5.9 \times 10^{-5}$ S/m, yielding $G' \approx 200$ μS/m. Typical resistance per unit length R' is on the order of 0.1 Ω/m near DC, increasing approximately in proportion to the square root of frequency.

From the above values, we find that RG-59 satisfies the low-loss criteria $R' \ll \omega L'$ for $f \gg 43$ kHz and $G' \ll \omega C'$ for $f \gg 470$ kHz. Under these conditions, we find $Z_0 \approx \sqrt{L'/C'} \cong 74\ \Omega$. Thus, the ratio of the potential to the current in a wave traveling in a single direction on RG-59 is about $74\ \Omega$.

The phase velocity of RG-59 is found to be $v_p \approx 1/\sqrt{L'C'} \cong 2 \times 10^8$ m/s, which is about 67% of c. In other words, a signal that takes 1 ns to traverse a distance l in free space requires about 1.5 ns to traverse a length-l section of RG-59. Since $v_p = \lambda f$, a wavelength in RG-59 is 67% of a wavelength in free space.

Using the expression

$$\gamma = \sqrt{(R' + j\omega L')\,(G' + j\omega C')}$$

with $R' = 0.1\ \Omega$/m, and then taking the real part to obtain α, we find $\alpha \sim 0.01$ m^{-1}. So, for example, the magnitude of the potential or current is decreased by about 50% by traveling a distance of about 70 m. In other words, $e^{-\alpha l} = 0.5$ for $l \sim 70$ m at relatively low frequencies, and increases with increasing frequency.

Additional Reading:

- "Coaxial cable" on Wikipedia. *Includes descriptions and design parameters for a variety of commonly-encountered coaxial cables.*
- "Single-ended signaling" on Wikipedia.
- Sec. 8.7 ("Differential Circuits") in S.W. Ellingson, *Radio Systems Engineering*, Cambridge Univ. Press, 2016.

---

1. The references in "Additional Reading" at the end of this section may be helpful if you are not familiar with this concept.↩
2. See Section 3.9 for a reminder about this concept.↩

This page titled 3.10: Coaxial Line is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.10: Coaxial Line** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.11: Microstrip Line

A microstrip transmission line consists of a narrow metallic trace separated from a metallic ground plane by a slab of dielectric material, as shown in Figure 3.11.1. This is a natural way to implement a transmission line on a printed circuit board, and so accounts for an important and expansive range of applications. The reader should be aware that microstrip is distinct from *stripline*, which is a very different type of transmission line; see "Additional Reading" at the end of this section for disambiguation of these terms.

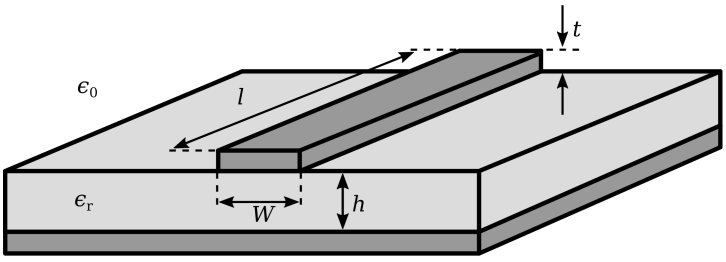


Figure 3.11.1: Microstrip transmission line structure and design parameters. (© CC BY SA 3.0 (modified))

A microstrip line is single-ended[1] in the sense that the conductor geometry is asymmetric and the one conductor – namely, the ground plane – also normally serves as ground for the source and load.

The spacer material is typically a low-loss dielectric material having permeability approximately equal to that of free space ($\mu \approx \mu_0$) and relative permittivity ϵ_r in the range 2 to about 10 or so.

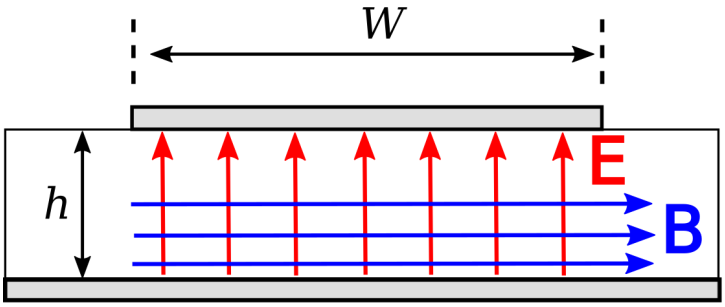


Figure 3.11.2: Structure of the electric and magnetic fields within microstrip line. (The fields *outside* the line are possibly significant, complicated, and not shown.) In this case, the wave is propagating away from the viewer.

A microstrip line nominally exhibits TEM field structure. This structure is shown in Figure 3.11.2. Note that electric and magnetic fields exist both in the dielectric and in the space above the dielectric, which is typically (but not always) air. This complex field structure makes it difficult to describe microstrip line concisely in terms of the equivalent circuit parameters of the lumped-element model. Instead, expressions for Z_0 directly in terms of h/W and ϵ_r are typically used instead. A variety of these expressions are in common use, representing different approximations and simplifications. A widely-accepted and broadly-applicable expression is:[2]

$$Z_0 \approx \frac{42.4\ \Omega}{\sqrt{\epsilon_r + 1}} \times \ln\left[1 + \frac{4h}{W'}\left(\Phi + \sqrt{\Phi^2 + \frac{1 + 1/\epsilon_r}{2}\pi^2}\right)\right] \tag{3.11.1}$$

where

$$\Phi \triangleq \frac{14 + 8/\epsilon_r}{11}\left(\frac{4h}{W'}\right)$$

and W' is W adjusted to account for the thickness t of the microstrip line. Typically $t \ll W$ and $t \ll h$, for which $W' \approx W$. Simpler approximations for Z_0 are also commonly employed in the design and analysis of microstrip lines. These expressions are limited in the range of h/W for which they are valid, and can usually be shown to be special cases or approximations of Equation 3.11.1. Nevertheless, they are sometimes useful for quick "back of the envelope" calculations.

Accurate expressions for wavelength λ, phase propagation constant β, and phase velocity v_p are similarly difficult to obtain for waves in microstrip line. An approximate technique employs a result from the theory of uniform plane waves in unbounded media (Equation 9.2.18 from Section 9.2):

$$\beta = \omega\sqrt{\mu\epsilon}$$

$$\gamma = \omega\sqrt{\mu\epsilon}$$

It turns out that the electromagnetic field structure in the space between the conductors is well-approximated as that of a uniform plane wave in unbounded media having the same permeability μ_0 but a different relative permittivity, which we shall assign the symbol $\epsilon_{r,eff}$ (for "*effective* relative permittivity"). Then

$$\begin{aligned}\beta &\approx \omega\sqrt{\mu_0\epsilon_{r,eff}\epsilon_0}\ (\text{ low-loss microstrip }) \\ &= \beta_0\sqrt{\epsilon_{r,eff}}\end{aligned} \tag{3.11.2}$$

In other words, the phase propagation constant in a microstrip line can be approximated as the free-space phase propagation $\beta_0 \triangleq \omega\sqrt{\mu_0\epsilon_0}$ times a correction factor $\sqrt{\epsilon_{r,eff}}$. Then $\epsilon_{r,eff}$ may be crudely approximated as follows:

$$\epsilon_{r,eff} \approx \frac{\epsilon_r + 1}{2}$$

i.e., $\epsilon_{r,eff}$ is roughly the average of the relative permittivity of the dielectric slab and the relative permittivity of free space. The assumption employed here is that $\epsilon_{r,eff}$ is approximately the average of these values because some fraction of the power in the guided wave is in the dielectric, and the rest is above the dielectric. Various approximations are available to improve on this approximation; however, in practice variations in the value of ϵ_r for the dielectric due to manufacturing processes typically make a more precise estimate irrelevant.

Using this concept, we obtain

$$\lambda = \frac{2\pi}{\beta} = \frac{2\pi}{\beta_0\sqrt{\epsilon_{r,eff}}} = \frac{\lambda_0}{\sqrt{\epsilon_{r,eff}}}$$

where λ_0 is the free-space wavelength c/f. Similarly the phase velocity v_p, can be estimated using the relationship

$$v_p = \frac{\omega}{\beta} = \frac{c}{\sqrt{\epsilon_{r,eff}}}$$

i.e., the phase velocity in microstrip is slower than c by a factor of $\sqrt{\epsilon_{r,eff}}$.

Example 3.11.1: 50 Ω Microstrip in FR4 Printed Circuit Boards.

FR4 is a low-loss fiberglass epoxy dielectric that is commonly used to make printed circuit boards (see "Additional Reading" at the end of this section). FR4 circuit board material is commonly sold in a slab having thickness $h \cong 1.575$ mm with $\epsilon_r \cong 4.5$. Let us consider how we might implement a microstrip line having $Z_0 = 50\ \Omega$ using this material. Since h and ϵ_r are fixed, the only parameter remaining to set Z_0 is W. A bit of experimentation with Equation 3.11.1 reveals that $h/W \approx 1/2$ yields $Z_0 \approx 50\ \Omega$ for $\epsilon_r = 4.5$. Thus, W should be about 3.15 mm. The effective relative permittivity is

$$\epsilon_{r,eff} \approx (4.5 + 1)/2 = 2.75$$

so the phase velocity for the wave guided by this line is about $c/\sqrt{2.75}$; i.e., 60% of c. Similarly, the wavelength of this wave is about 60% of the free space wavelength.

1. The reference in "Additional Reading" at the end of this section may be helpful if you are not familiar with this concept.↩
2. This is from Wheeler 1977, cited in "Additional Reading" at the end of this section.↩

This page titled 3.11: Microstrip Line is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.11: Microstrip Line** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.12: Voltage Reflection Coefficient

We now consider the scenario shown in Figure 3.12.1. Here a wave arriving from the left along a lossless transmission line having characteristic impedance Z_0 arrives at a termination located at $z = 0$. The impedance looking into the termination is Z_L, which may be real-, imaginary-, or complex-valued. The questions are: Under what circumstances is a reflection – i.e., a leftward traveling wave – expected, and what precisely is that wave?

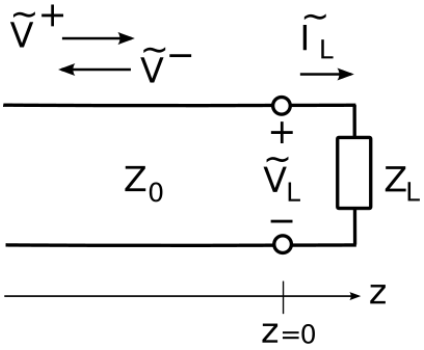


Figure 3.12.1: A wave arriving from the left incident on a termination located at $z = 0$.

The potential and current of the incident wave are related by the constant value of Z_0. Similarly, the potential and current of the reflected wave are related by Z_0. Therefore, it suffices to consider *either* potential or current. Choosing potential, we may express the incident wave as

$$\widetilde{V}^+(z) = V_0^+ e^{-j\beta z}$$

where V_0^+ is determined by the source of the wave, and so is effectively a "given." Any reflected wave must have the form

$$\widetilde{V}^-(z) = V_0^- e^{+j\beta z}$$

Therefore, the problem is solved by determining the value of V_0^- given V_0^+, Z_0, and Z_L.

Considering the situation at $z = 0$, note that by definition we have

$$Z_L \triangleq \frac{\widetilde{V}_L}{\tilde{I}_L} \tag{3.12.1}$$

where $\widetilde{V}_L$ and $\tilde{I}_L$ are the potential across and current through the termination, respectively. Also, the potential and current on either side of the $z = 0$ interface must be equal. Thus,

$$\widetilde{V}^+(0) + \widetilde{V}^-(0) = \widetilde{V}_L \tag{3.12.2}$$
$$\tilde{I}^+(0) + \tilde{I}^-(0) = \tilde{I}_L \tag{3.12.3}$$

where $\tilde{I}^+(z)$ and $\tilde{I}^-(z)$ are the currents associated with $\widetilde{V}^+(z)$ and $\widetilde{V}^-(z)$, respectively. Since the voltage and current are related by Z_0, Equation 3.12.3 may be rewritten as follows:

$$\frac{\widetilde{V}^+(0)}{Z_0} - \frac{\widetilde{V}^-(0)}{Z_0} = \tilde{I}_L \tag{3.12.4}$$

Evaluating the left sides of Equations 3.12.2 and 3.12.4 at $z = 0$, we find:

$$\begin{aligned} V_0^+ + V_0^- &= \widetilde{V}_L \\ \frac{V_0^+}{Z_0} - \frac{V_0^-}{Z_0} &= \tilde{I}_L \end{aligned} \tag{3.12.5}$$

Substituting these expressions into Equation 3.12.1 we obtain:

$$Z_L = \frac{V_0^+ + V_0^-}{V_0^+/Z_0 - V_0^-/Z_0}$$

Solving for V_0^- we obtain

$$V_0^- = \frac{Z_L - Z_0}{\quad}$$

$$V_0^- = \frac{\quad}{Z_L + Z_0}\, V_0^+$$

Thus, the answer to the question posed earlier is that

$$V_0^- = \Gamma V_0^+ \quad \text{, where}$$

$$\boxed{\Gamma \triangleq \frac{Z_L - Z_0}{Z_L + Z_0}} \tag{3.12.6}$$

The quantity Γ is known as the *voltage reflection coefficient*. Note that when $Z_L = Z_0$, $\Gamma = 0$ and therefore $V_0^- = 0$. In other words,

> If the terminating impedance is equal to the characteristic impedance of the transmission line, then there is no reflection.

If, on the other hand, $Z_L \neq Z_0$, then $|\Gamma| > 0$, $V_0^- = \Gamma V_0^+$, and a leftward-traveling reflected wave exists.

Since Z_L may be real-, imaginary-, or complex-valued, Γ too may be real-, imaginary-, or complex-valued. Therefore, V_0^- may be different from V_0^+ in magnitude, sign, or phase.

Note also that Γ is *not* the ratio of I_0^- to I_0^+. The ratio of the *current* coefficients is actually $-\Gamma$. It is quite simple to show this with a simple modification to the above procedure and is left as an exercise for the student.

Summarizing:

> The voltage reflection coefficient Γ, given by Equation 3.12.6, determines the magnitude and phase of the reflected wave given the incident wave, the characteristic impedance of the transmission line, and the terminating impedance.

We now consider values Γ that arise for commonly-encountered terminations.

Matched Load. ($Z_L = Z_0$). In this case, the termination may be a device with impedance Z_0, or the termination may be another transmission line having the same characteristic impedance. When $Z_L = Z_0$, $\Gamma = 0$ and there is no reflection.

Open Circuit. An "open circuit" is the absence of a termination. This condition implies $Z_L \to \infty$, and subsequently $\Gamma \to +1$. Since the *current* reflection coefficient is $-\Gamma$, the reflected current wave is 180° out of phase with the incident current wave, making the total current at the open circuit equal to zero, as expected.

Short Circuit. "Short circuit" means $Z_L = 0$, and subsequently $\Gamma = -1$. In this case, the phase of Γ is 180°, and therefore, the potential of the reflected wave cancels the potential of the incident wave at the open circuit, making the total potential equal to zero, as it must be. Since the *current* reflection coefficient is $-\Gamma = +1$ in this case, the reflected current wave is in phase with the incident current wave, and the magnitude of the total current at the short circuit non-zero as expected.

Purely Reactive Load. A purely reactive load, including that presented by a capacitor or inductor, has $Z_L = jX$ where X is reactance. In particular, an inductor is represented by $X > 0$ and a capacitor is represented by $X < 0$. We find

$$\Gamma = \frac{-Z_0 + jX}{+Z_0 + jX}$$

The numerator and denominator have the same magnitude, so $|\Gamma| = 1$. Let ϕ be the phase of the denominator ($+Z_0 + jX$). Then, the phase of the numerator is $\pi - \phi$. Subsequently, the phase of Γ is $(\pi - \phi) - \phi = \pi - 2\phi$. Thus, we see that the phase of Γ is no longer limited to be 0° or 180°, but can be any value in between. The phase of reflected wave is subsequently shifted by this amount.

Other Terminations. Any other termination, including series and parallel combinations of any number of devices, can be expressed as a value of Z_L which is, in general, complex-valued. The associated value of $|\Gamma|$ is limited to the range 0 to 1. To see this, note:

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} = \frac{Z_L/Z_0 - 1}{Z_L/Z_0 + 1}$$

Note that the smallest possible value of $|\Gamma|$ occurs when the numerator is zero; i.e., when $Z_L = Z_0$. Therefore, the smallest value of $|\Gamma|$ is zero. The largest possible value of $|\Gamma|$ occurs when $Z_L/Z_0 \to \infty$ (i.e., an open circuit) or when $Z_L/Z_0 = 0$ (a short circuit); the result in either case is $|\Gamma| = 1$. Thus,

$$\boxed{0 \leq |\Gamma| \leq 1}$$

This page titled 3.12: Voltage Reflection Coefficient is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.12: Voltage Reflection Coefficient** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.13: Standing Waves

A *standing wave* consists of waves moving in opposite directions. These waves add to make a distinct magnitude variation as a function of distance that does not vary in time.

To see how this can happen, first consider that an incident wave $V_0^+ e^{-j\beta z}$, which is traveling in the $+z$ axis along a lossless transmission line. Associated with this wave is a reflected wave $V_0^- e^{+j\beta z} = \Gamma V_0^+ e^{+j\beta z}$, where Γ is the voltage reflection coefficient. These waves add to make the total potential

$$\begin{aligned}\widetilde{V}(z) &= V_0^+ e^{-j\beta z} + \Gamma V_0^+ e^{+j\beta z} \\ &= V_0^+ \left(e^{-j\beta z} + \Gamma e^{+j\beta z}\right)\end{aligned}$$

The magnitude of $\widetilde{V}(z)$ is most easily found by first finding $|\widetilde{V}(z)|^2$, which is:

$$\begin{aligned}&\tilde{V}(z)\tilde{V}^*(z) \\ =&\left|V_0^+\right|^2 \left(e^{-j\beta z} + \Gamma e^{+j\beta z}\right)\left(e^{-j\beta z} + \Gamma e^{+j\beta z}\right)^* \\ =&\left|V_0^+\right|^2 \left(e^{-j\beta z} + \Gamma e^{+j\beta z}\right)\left(e^{+j\beta z} + \Gamma^* e^{-j\beta z}\right) \\ =&\left|V_0^+\right|^2 \left(1 + |\Gamma|^2 + \Gamma e^{+j2\beta z} + \Gamma^* e^{-j2\beta z}\right)\end{aligned} \tag{3.13.1}$$

Let ϕ be the phase of Γ; i.e.,

$$\Gamma = |\Gamma| e^{j\phi}$$

Then, continuing from the previous expression:

$$\begin{aligned}&\left|V_0^+\right|^2 \left(1 + |\Gamma|^2 + |\Gamma| e^{+j(2\beta z+\phi)} + |\Gamma| e^{-j(2\beta z+\phi)}\right) \\ =&\left|V_0^+\right|^2 \left(1 + |\Gamma|^2 + |\Gamma| \left[e^{+j(2\beta z+\phi)} + e^{-j(2\beta z+\phi)}\right]\right)\end{aligned} \tag{3.13.2}$$

The quantity in square brackets can be reduced to a cosine function using the identity

$$\cos\theta = \frac{1}{2}\left[e^{j\theta} + e^{-j\theta}\right]$$

yielding:

$$|V_0^+|^2 \left[1 + |\Gamma|^2 + 2|\Gamma| \cos\left(2\beta z + \phi\right)\right]$$

Recall that this is $|\widetilde{V}(z)|^2$. $|\widetilde{V}(z)|$ is therefore the square root of the above expression:

$$\left|\widetilde{V}(z)\right| = |V_0^+| \sqrt{1 + |\Gamma|^2 + 2|\Gamma| \cos\left(2\beta z + \phi\right)}$$

Thus, we have found that the magnitude of the resulting total potential varies sinusoidally along the line. This is referred to as a standing wave because the variation of the magnitude of the phasor resulting from the interference between the incident and reflected waves does not vary with time.

We may perform a similar analysis of the current, leading to:

$$\left|\tilde{I}(z)\right| = \frac{|V_0^+|}{Z_0} \sqrt{1 + |\Gamma|^2 - 2|\Gamma| \cos\left(2\beta z + \phi\right)}$$

Again we find the result is a standing wave.

Now let us consider the outcome for a few special cases.

Matched load. When the impedance of the termination of the transmission line, Z_L, is equal to the characteristic impedance of the transmission line, Z_0, $\Gamma = 0$ and there is no reflection. In this case, the above expressions reduce to $|\widetilde{V}(z)| = |V_0^+|$ and $|\tilde{I}(z)| = |V_0^+|/Z_0$, as expected.

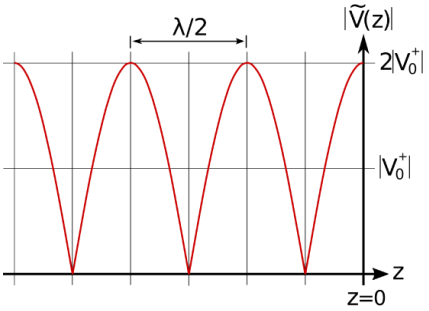
Open or Short-Circuit. In this case, $\Gamma = \pm 1$ and we find:

$$\left|\widetilde{V}(z)\right| = |V_0^+| \sqrt{2 + 2\cos\left(2\beta z + \phi\right)}$$

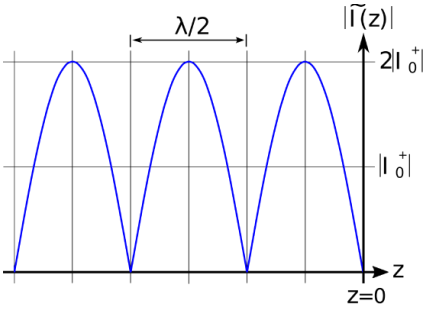
$$|V(z)| = |V_0^+|\sqrt{2 + 2\cos(2\beta z + \phi)}$$

$$\left|\tilde{I}(z)\right| = \frac{|V_0^+|}{Z_0}\sqrt{2 - 2\cos(2\beta z + \phi)}$$

where $\phi = 0$ for an open circuit and $\phi = \pi$ for a short circuit. The result for an open circuit termination is shown in Figure 3.13.1(a) (potential) and 3.13.1(b) (current). The result for a short circuit termination is identical except the roles of potential and current are reversed. In either case, note that voltage maxima correspond to current minima, and vice versa.



(a) Potential.



(b) Current.

Figure 3.13.1: Standing wave associated with an opencircuit termination at $z = 0$ (incident wave arrives from left).

Also note:

> The period of the standing wave is $\lambda/2$; i.e., one-half of a wavelength.

This can be confirmed as follows. First, note that the frequency argument of the cosine function of the standing wave is $2\beta z$. This can be rewritten as $2\pi\,(\beta/\pi)\,z$, so the frequency of variation is β/π and the period of the variation is π/β. Since $\beta = 2\pi/\lambda$, we see that the period of the variation is $\lambda/2$. Furthermore, this is true regardless of the value of Γ.

Mismatched loads. A common situation is that the termination is neither perfectly-matched ($\Gamma = 0$) nor an open/short circuit ($|\Gamma| = 1$). Examples of the resulting standing waves are shown in Figure 3.13.2.

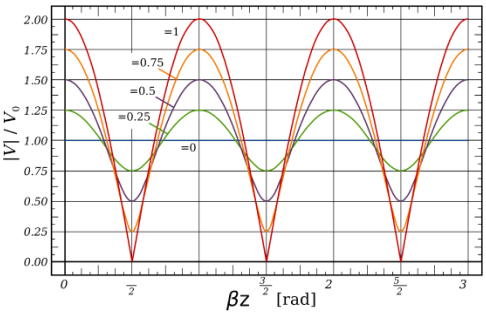


Figure 3.13.2: Standing waves associated with loads exhibiting various reflection coefficients. In this figure the incident wave arrives from the right.

This page titled 3.13: Standing Waves is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.13: Standing Waves** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.14: Standing Wave Ratio

Precise matching of transmission lines to terminations is often not practical or possible. Whenever a significant mismatch exists, a standing wave (Section 3.13) is apparent. The quality of the match is commonly expressed in terms of the *standing wave ratio* (SWR) of this standing wave.

> *Standing wave ratio* (SWR) is defined as the ratio of the maximum magnitude of the standing wave to minimum magnitude of the standing wave.

In terms of the potential:

$$\boxed{\text{SWR} \triangleq \frac{\text{maximum } |\widetilde{V}|}{\text{minimum } |\widetilde{V}|}}$$

SWR can be calculated using a simple expression, which we shall now derive. In Section 3.13, we found that:

$$\left|\widetilde{V}(z)\right| = |V_0^+|\sqrt{1 + |\Gamma|^2 + 2\,|\Gamma|\cos\left(2\beta z + \phi\right)}$$

The maximum value occurs when the cosine factor is equal to $+1$, yielding:

$$\max\ \left|\widetilde{V}\right| = |V_0^+|\sqrt{1 + |\Gamma|^2 + 2\,|\Gamma|}$$

Note that the argument of the square root operator is equal to $(1 + |\Gamma|)^2$; therefore:

$$\max\ \left|\widetilde{V}\right| = |V_0^+|\,(1 + |\Gamma|)$$

Similarly, the minimum value is achieved when the cosine factor is equal to -1, yielding:

$$\min\ \left|\widetilde{V}\right| = |V_0^+|\sqrt{1 + |\Gamma|^2 - 2\,|\Gamma|}$$

So:

$$\min\ \left|\widetilde{V}\right| = |V_0^+|\,(1 - |\Gamma|)$$

Therefore:

$$\boxed{\text{SWR} = \frac{1 + |\Gamma|}{1 - |\Gamma|}} \tag{3.14.1}$$

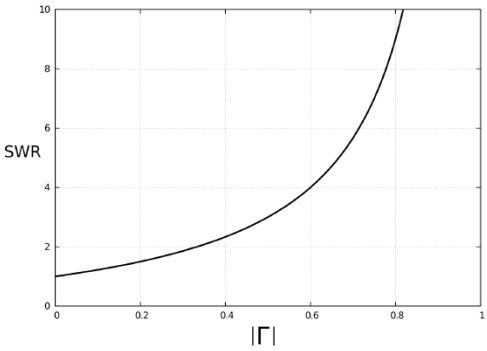


Figure 3.14.1: Relationship between SWR and $|\Gamma|$.

This relationship is shown graphically in Figure 3.14.1. Note that SWR ranges from 1 for perfectly-matched terminations ($\Gamma = 0$) to infinity for open- and short-circuit terminations ($|\Gamma| = 1$).

It is sometimes of interest to find the magnitude of the reflection coefficient given SWR. Solving Equation 3.14.1 for $|\Gamma|$ we find:

$$|\Gamma| = \frac{\text{SWR} - 1}{\text{SWR} + 1} \tag{3.14.2}$$

SWR is often referred to as the *voltage standing wave ratio* (VSWR), although repeating the analysis above for the current reveals that the current SWR is equal to potential SWR, so the term “SWR” suffices.

SWR < 2 or so is usually considered a “good match,” although some applications require SWR < 1.1 or better, and other applications are tolerant to SWR of 3 or greater.

Example 3.14.1: Reflection Coefficient for Various Values of SWR

What is the reflection coefficient for the above-cited values of SWR? Using Equation 3.14.2, we find:

- SWR = 1.1 corresponds to $|\Gamma| = 0.0476$.
- SWR = 2.0 corresponds to $|\Gamma| = 1/3$.
- SWR = 3.0 corresponds to $|\Gamma| = 1/2$.

This page titled 3.14: Standing Wave Ratio is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.14: Standing Wave Ratio** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.15: Input Impedance of a Terminated Lossless Transmission Line

Consider Figure 3.15.1, which shows a lossless transmission line being driven from the left and which is terminated by an impedance Z_L on the right. If Z_L is equal to the characteristic impedance Z_0 of the transmission line, then the input impedance Z_{in} will be equal to Z_L. Otherwise Z_{in} depends on both Z_L and the characteristics of the transmission line. In this section, we determine a general expression for Z_{in} in terms of Z_L, Z_0, the phase propagation constant β, and the length l of the line.

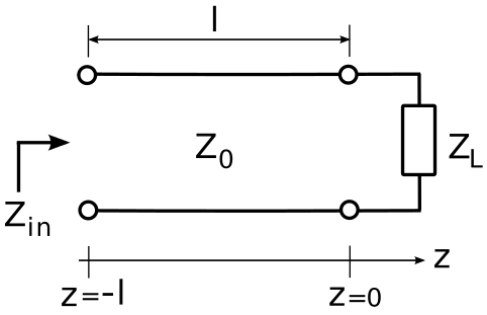


Figure 3.15.1: A transmission line driven by a source on the left and terminated by an impedance Z_L at $z = 0$ on the right

Using the coordinate system indicated in Figure 3.15.1, the interface between source and transmission line is located at $z = -l$. Impedance is defined at the ratio of potential to current, so:

$$Z_{in}(l) \triangleq \frac{\widetilde{V}(z=-l)}{\tilde{I}(z=-l)}$$

Now employing expressions for $\widetilde{V}(z)$ and $\tilde{I}(z)$ from Section 3.13 with $z = -l$, we find:

$$\begin{aligned} Z_{in}(l) &= \frac{V_0^+\left(e^{+j\beta l}+\Gamma e^{-j\beta l}\right)}{V_0^+\left(e^{+j\beta l}-\Gamma e^{-j\beta l}\right)/Z_0} \\ &= Z_0\frac{e^{+j\beta l}+\Gamma e^{-j\beta l}}{e^{+j\beta l}-\Gamma e^{-j\beta l}} \end{aligned} \tag{3.15.1}$$

Multiplying both numerator and denominator by $e^{-j\beta l}$:

$$\boxed{Z_{in}(l) = Z_0\frac{1+\Gamma e^{-j2\beta l}}{1-\Gamma e^{-j2\beta l}}} \tag{3.15.2}$$

Recall that Γ in the above expression is:

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{3.15.3}$$

Summarizing:

> Equation 3.15.2 is the input impedance of a lossless transmission line having characteristic impedance Z_0 and which is terminated into a load Z_L. The result also depends on the length and phase propagation constant of the line.

Note that $Z_{in}(l)$ is periodic in l. Since the argument of the complex exponential factors is $2\beta l$, the frequency at which $Z_{in}(l)$ varies is β/π; and since $\beta = 2\pi/\lambda$, the associated period is $\lambda/2$. This is very useful to keep in mind because it means that all possible values of $Z_{in}(l)$ are achieved by varying l over $\lambda/2$. In other words, changing l by more than $\lambda/2$ results in an impedance which could have been obtained by a smaller change in l. Summarizing to underscore this important idea:

> The input impedance of a terminated lossless transmission line is periodic in the length of the transmission line, with period $\lambda/2$.

Not surprisingly, $\lambda/2$ is also the period of the standing wave (Section 3.13). This is because – once again – the variation with length is due to the interference of incident and reflected waves.

Also worth noting is that Equation 3.15.2 can be written entirely in terms of Z_L and Z_0, since Γ depends only on these two parameters. Here's that version of the expression:

$$Z_{in}(l) = Z_0 \left[\frac{Z_L + jZ_0 \tan \beta l}{Z_0 + jZ_L \tan \beta l} \right] \tag{3.15.4}$$

This expression can be derived by substituting Equation 3.15.3 into Equation 3.15.2 and is left as an exercise for the student.

Finally, note that the argument βl appearing Equations 3.15.2 and 3.15.4 has units of radians and is referred to as *electrical length*. Electrical length can be interpreted as physical length expressed with respect to wavelength and has the advantage that analysis can be made independent of frequency.

This page titled 3.15: Input Impedance of a Terminated Lossless Transmission Line is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.15: Input Impedance of a Terminated Lossless Transmission Line** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.16: Input Impedance for Open- and Short-Circuit Terminations

Let us now consider the input impedance of a transmission line that is terminated in an open- or short-circuit. Such a transmission line is sometimes referred to as a *stub*. First, why consider such a thing? From a "lumped element" circuit theory perspective, this would not seem to have any particular application. However, the fact that this structure exhibits an input impedance that depends on length (Section 3.15) enables some very useful applications.

First, let us consider the question at hand: What is the input impedance when the transmission line is open- or short-circuited? For a short circuit, $Z_L = 0, \Gamma = -1$, so we find

$$\begin{aligned} Z_{in}(l) &= Z_0 \frac{1+\Gamma e^{-j2\beta l}}{1-\Gamma e^{-j2\beta l}} \\ &= Z_0 \frac{1-e^{-j2\beta l}}{1+e^{-j2\beta l}} \end{aligned} \tag{3.16.1}$$

Multiplying numerator and denominator by $e^{+j\beta l}$ we obtain

$$Z_{in}(l) = Z_0 \frac{e^{+j\beta l} - e^{-j\beta l}}{e^{+j\beta l} + e^{-j\beta l}}$$

Now we invoke the following trigonometric identities:

$$\begin{aligned} \cos\theta &= \tfrac{1}{2}\left[e^{+j\theta} + e^{-j\theta}\right] \\ \sin\theta &= \tfrac{1}{j2}\left[e^{+j\theta} - e^{-j\theta}\right] \end{aligned} \tag{3.16.2}$$

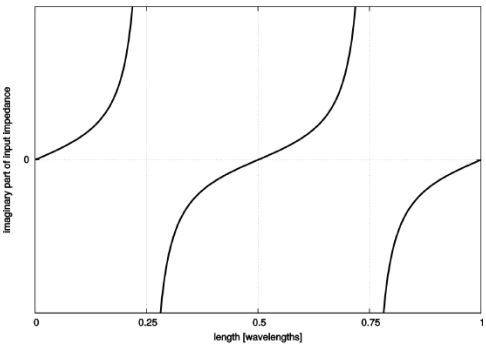
Employing these identities, we obtain:

$$Z_{in}(l) = Z_0 \frac{j2\,(\sin\beta l)}{2\,(\cos\beta l)}$$

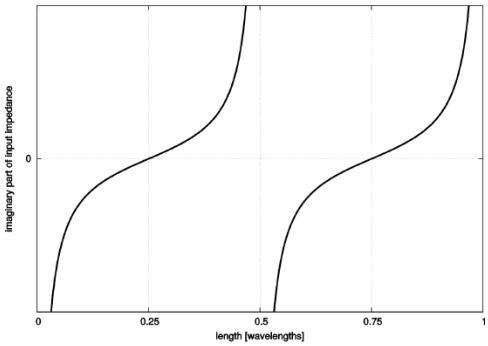
and finally:

$$\boxed{Z_{in}(l) = +jZ_0 \tan\beta l} \tag{3.16.3}$$

Figure 3.16.1(a) shows what's going on. As expected, $Z_{in} = 0$ when $l = 0$, since this amounts to a short circuit with no transmission line. Also, Z_{in} varies periodically with increasing length, with period $\lambda/2$. This is precisely as expected from standing wave theory (Section 3.13). What is of particular interest now is that as $l \to \lambda/4$, we see $Z_{in} \to \infty$. Remarkably, the transmission line has essentially transformed the short circuit termination into an open circuit!



(a) Short-ciruit termination ($Z_L = 0$.



(b) Open-circuit termination ($Z_L \to \infty$).

Figure 3.16.1: Input reactance ($\mathrm{Im}\{Z_{in}\}$) of a stub. $\mathrm{Re}\{Z_{in}\}$ is always zero.

For an open circuit termination, $Z_L \to \infty$, $\Gamma = +1$, and we find

$$\begin{aligned} Z_{in}(l) &= Z_0 \frac{1+\Gamma e^{-j2\beta l}}{1-\Gamma e^{-j2\beta l}} \\ &= Z_0 \frac{1+e^{-j2\beta l}}{1-e^{-j2\beta l}} \end{aligned} \tag{3.16.4}$$

Following the same procedure detailed above for the short-circuit case, we find

$$\boxed{Z_{in}(l) = -jZ_0 \cot \beta l} \tag{3.16.5}$$

Figure 3.16.1(b) shows the result for open-circuit termination. As expected, $Z_{in} \to \infty$ for $l = 0$, and the same $\lambda/2$ periodicity is observed. What is of particular interest now is that at $l = \lambda/4$ we see $Z_{in} = 0$. In this case, the transmission line has transformed the *open* circuit termination into a *short* circuit.

Now taking stock of what we have determined:

> The input impedance of a short- or open-circuited lossless transmission line is completely imaginary-valued and is given by Equations 3.16.3 and 3.16.5, respectively.

> The input impedance of a short- or open-circuited lossless transmission line alternates between open- ($Z_{in} \to \infty$) and short-circuit ($Z_{in} = 0$) conditions with each $\lambda/4$-increase in length.

This page titled 3.16: Input Impedance for Open- and Short-Circuit Terminations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.16: Input Impedance for Open- and Short-Circuit Terminations** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.17: Applications of Open- and Short-Circuited Transmission Line Stubs

The theory of open- and short-circuited transmission lines – often referred to as *stubs* – was addressed in Section 3.16. These structures have important and wide-ranging applications.

In particular, these structures can be used to replace discrete inductors and capacitors in certain applications. To see this, consider the short-circuited line (Figure 3.16.1(a) of Section 3.16). Note that each value of l that is less than $\lambda/4$ corresponds to a particular positive reactance; i.e., the transmission line "looks" like an inductor. Also note that lengths between $\lambda/4$ and $\lambda/2$ result in reactances that are negative; i.e., the transmission line "looks" like a capacitor. Thus, it is possible to replace an inductor or capacitor with a short-circuited transmission line of the appropriate length. The input impedance of such a transmission line is identical to that of the inductor or capacitor at the design frequency. The variation of reactance with respect to frequency will *not* be identical, which may or may not be a concern depending on the bandwidth and frequency response requirements of the application. Open-circuited lines may be used in a similar way.

This property of open- and short-circuited transmission lines makes it possible to implement impedance matching circuits (see Section 3.16 a), filters, and other devices entirely from transmission lines, with fewer or no discrete inductors or capacitors required. Transmission lines do not suffer the performance limitations of discrete devices at high frequencies and are less expensive. A drawback of transmission line stubs in this application is that the lines are typically much larger than the discrete devices they are intended to replace.

Example 3.17.1: Emitter Induction Using Short-Circuited Line

In the design of low-noise amplifiers using bipolar transistors in common-emitter configuration, it is often useful to introduce a little inductance between the emitter and ground. This is known as "inductive degeneration," "emitter induction," or sometimes by other names. It can be difficult to find suitable inductors, especially for operation in the UHF band and higher. However, a microstrip line can be used to achieve the desired inductive impedance. Determine the length of a stub that implements a 2.2 nH inductance at 6 GHz using microstrip line with characteristic impedance $50\ \Omega$ and phase velocity $0.6c$.

Solution

At the design frequency, the impedance looking into this section of line from the emitter should be equal to that of a 2.2 nH inductor, which is $+j\omega L = +j2\pi fL = +j82.9\ \Omega$. The input impedance of a short-circuited stub of length l which is grounded (thus, short-circuited) at the opposite end is $+jZ_0 \tan \beta l$ (Section 3.16). Setting this equal to $+j82.9\ \Omega$ and noting that $Z_0 = 50\ \Omega$, we find that $\beta l \cong 1.028$ rad. The phase propagation constant is (Section 3.8):

$$\beta = \frac{\omega}{v_p} = \frac{2\pi f}{0.6c} \cong 209.4\ \text{rad/m}$$

Therefore, the length of the microstrip line is $l = (\beta l)\,/\beta \cong 4.9$ mm.

This page titled 3.17: Applications of Open- and Short-Circuited Transmission Line Stubs is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.17: Applications of Open- and Short-Circuited Transmission Line Stubs** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.18: Measurement of Transmission Line Characteristics

This section presents a simple technique for measuring the characteristic impedance Z_0, electrical length βl, and phase velocity v_p of a lossless transmission line. This technique requires two measurements: the input impedance Z_{in} when the transmission line is short-circuited and Z_{in} when the transmission line is open-circuited.

In Section 3.16, it is shown that the input impedance Z_{in} of a short-circuited transmission line is

$$Z_{in}^{(SC)} = +jZ_0 \tan \beta l$$

and when a transmission line is terminated in an open circuit, the input impedance is

$$Z_{in}^{(OC)} = -jZ_0 \cot \beta l$$

Observe what happens when we multiply these results together:

$$Z_{in}^{(SC)} \cdot Z_{in}^{(OC)} = Z_0^2$$

that is, the product of the measurements $Z_{in}^{(OC)}$ and $Z_{in}^{(SC)}$ is simply the square of the characteristic impedance. Therefore

$$Z_0 = \sqrt{Z_{in}^{(SC)} \cdot Z_{in}^{(OC)}}$$

If we instead divide these measurements, we find

$$\frac{Z_{in}^{(SC)}}{Z_{in}^{(OC)}} = -\tan^2 \beta l$$

Therefore:

$$\tan \beta l = \left[-\frac{Z_{in}^{(SC)}}{Z_{in}^{(OC)}} \right]^{1/2}$$

If l is known in advance to be less than $\lambda/2$, then the electrical length βl can be determined by taking the inverse tangent. If l is of unknown length and longer than $\lambda/2$, one must take care to account for the periodicity of tangent function; in this case, it may not be possible to unambiguously determine βl. Although we shall not present the method here, it is possible to resolve this ambiguity by making multiple measurements over a range of frequencies.

Once βl is determined, it is simple to determine l given β, β given l, and then v_p. For example, the phase velocity may be determined by first finding βl for a known length using the above procedure, calculating $\beta = (\beta l)/l$, and then $v_p = \omega/\beta$.

This page titled 3.18: Measurement of Transmission Line Characteristics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.18: Measurement of Transmission Line Characteristics** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.19: Quarter-Wavelength Transmission Line

Quarter-wavelength sections of transmission line play an important role in many systems at radio and optical frequencies. The remarkable properties of open- and short-circuited quarter-wave line are presented in Section 3.16 and should be reviewed before reading further. In this section, we perform a more general analysis, considering not just open- and short-circuit terminations but any terminating impedance, and then we address some applications.

The general expression for the input impedance of a lossless transmission line is (Section 3.15):

$$Z_{in}(l) = Z_0 \frac{1+\Gamma e^{-j2\beta l}}{1-\Gamma e^{-j2\beta l}} \tag{3.19.1}$$

Note that when $l = \lambda/4$:

$$2\beta l = 2 \cdot \frac{2\pi}{\lambda} \cdot \frac{\lambda}{4} = \pi$$

Subsequently:

$$\begin{aligned} Z_{in}(\lambda/4) &= Z_0 \frac{1+\Gamma e^{-j\pi}}{1-\Gamma e^{-j\pi}} \\ &= Z_0 \frac{1-\Gamma}{1+\Gamma} \end{aligned}$$

Recall that (Section 3.15):

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0}$$

Substituting this expression and then multiplying numerator and denominator by $Z_L + Z_0$, one obtains

$$\begin{aligned} Z_{in}(\lambda/4) &= Z_0 \frac{(Z_L+Z_0)-(Z_L-Z_0)}{(Z_L+Z_0)+(Z_L-Z_0)} \\ &= Z_0 \frac{2Z_0}{2Z_L} \end{aligned}$$

Thus,

$$\boxed{Z_{in}(\lambda/4) = \frac{Z_0^2}{Z_L}} \tag{3.19.2}$$

Note that the input impedance is inversely proportional to the load impedance. For this reason, a transmission line of length $\lambda/4$ is sometimes referred to as a *quarter-wave inverter* or simply as a *impedance inverter*.

Quarter-wave lines play a very important role in RF engineering. As impedance inverters, they have the useful attribute of transforming small impedances into large impedances, and vice-versa – we'll come back to this idea later in this section. First, let's consider how quarter-wave lines are used for impedance matching. Look what happens when we solve Equation 3.19.2 for Z_0:

$$Z_0 = \sqrt{Z_{in}(\lambda/4) \cdot Z_L} \tag{3.19.3}$$

This equation indicates that we may match the load Z_L to a source impedance (represented by $Z_{in}(\lambda/4)$) simply by making the characteristic impedance equal to the value given by the above expression and setting the length to $\lambda/4$. The scheme is shown in Figure 3.19.1.

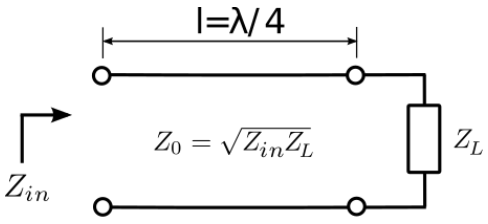


Figure 3.19.1: Impedance-matching using a quarter-wavelength transmission line.

Example 3.19.1: 300-to-50 Ω match using an quarter-wave section of line

Design a transmission line segment that matches 300 Ω to 50 Ω at 10 GHz using a quarter-wave match. Assume microstrip line for which propagation occurs with wavelength 60% that of free space.

Solution

The line is completely specified given its characteristic impedance Z_0 and length l. The length should be one-quarter wavelength with respect to the signal propagating in the line. The free-space wavelength $\lambda_0 = c/f$ at 10 GHz is $\cong 3$ cm. Therefore, the wavelength of the signal in the line is $\lambda = 0.6\lambda_0 \cong 1.8$ cm, and the length of the line should be $l = \lambda/4 \cong 4.5$ mm.

The characteristic impedance is given by Equation ref{m0091_eQWZ0}:

$$Z_0 = \sqrt{300\ \Omega \cdot 50\ \Omega} \cong 122.5\ \Omega$$

This value would be used to determine the width of the microstrip line, as discussed in Section 3.11.

It should be noted that for this scheme to yield a real-valued characteristic impedance, the product of the source and load impedances must be a real-valued number. In particular, this method is *not* suitable if Z_L has a significant imaginary-valued component and matching to a real-valued source impedance is desired. One possible workaround in this case is the two-stage strategy shown in Figure 3.19.2. In this scheme, the load impedance is first transformed to a real-valued impedance using a length l_1 of transmission line. This is accomplished using Equation 3.19.1 (quite simple using a numerical search) or using the Smith chart (see "Additional Reading" at the end of this section). The characteristic impedance Z_{01} of this transmission line is not critical and can be selected for convenience. Normally, the smallest value of l_1 is desired. This value will always be less than $\lambda/4$ since $Z_{in}(l_1)$ is periodic in l_1 with period $\lambda/2$; i.e., there are two changes in the sign of the imaginary component of $Z_{in}(l_1)$ as l_1 is increased from zero to $\lambda/2$. After eliminating the imaginary component of Z_L in this manner, the real component of the resulting impedance may then be transformed using the quarter-wave matching technique described earlier in this section.

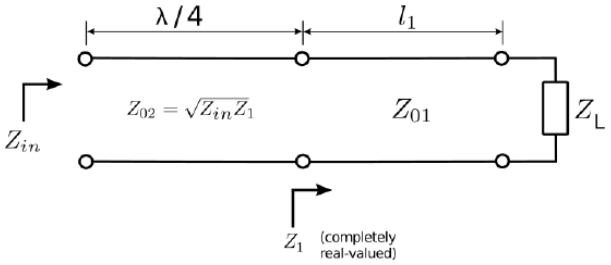


Figure 3.19.2: Impedance-matching a complex-valued load impedance using quarter-wavelength transmission line.

Example 3.19.2: Matching a patch Antenna to 50 Ω

A particular patch antenna exhibits a source impedance of $Z_A = 35 + j35\ \Omega$. (See "Microstrip antenna" in "Additional Reading" at the end of this section for some optional reading on patch antennas.) Interface this antenna to 50 Ω using the technique described above. For the section of transmission line adjacent to the patch antenna, use characteristic impedance $Z_{01} = 50\ \Omega$. Determine the lengths l_1 and l_2 of the two segments of transmission line, and the characteristic impedance Z_{02} of the second (quarter-wave) segment.

Solution

The length of the first section of the transmission line (adjacent to the antenna) is determined using Equation 3.19.1:

$$Z_1(l_1) = Z_{01}\frac{1 + \Gamma e^{-j2\beta_1 l_1}}{1 - \Gamma e^{-j2\beta_1 l_1}}$$

where β_1 is the phase propagation constant for this section of transmission line and

$$\Gamma \triangleq \frac{Z_A - Z_{01}}{Z_A + Z_{01}} \cong -0.0059 + j0.4142$$

We seek the value of smallest positive value of $\beta_1 l_1$ for which the imaginary part of $Z_1(l_1)$ is zero. This can determined using a Smith chart (see "Additional Reading" at the end of this section) or simply by a few iterations of trial-and-error. Either way we

find $Z_1(\beta_1 l_1 = 0.793 \text{ rad}) \cong 120.719 - j0.111\ \Omega$, which we deem to be close enough to be acceptable. Note that $\beta_1 = 2\pi/\lambda$, where λ is the wavelength of the signal in the transmission line. Therefore

$$l_1 = \frac{\beta_1 l_1}{\beta_1} = \frac{\beta_1 l_1}{2\pi}\lambda \cong 0.126\lambda$$

The length of the second section of the transmission line, being a quarter-wavelength transformer, should be $l_2 = 0.25\lambda$. Using Equation 3.19.3, the characteristic impedance Z_{02} of this section of line should be

$$Z_{02} \cong \sqrt{(120.719\ \Omega)\,(50\ \Omega)} \cong 77.7\ \Omega$$

Discussion. The total length of the matching structure is $l_1 + l_2 \cong 0.376\lambda$. A patch antenna would typically have sides of length about $\lambda/2 = 0.5\lambda$, so the matching structure is nearly as big as the antenna itself. At frequencies where patch antennas are commonly used, and especially at frequencies in the UHF (300–3000 MHz) band, patch antennas are often comparable to the size of the system, so it is not attractive to have the matching structure also require a similar amount of space. Thus, we would be motivated to find a smaller matching structure.

Although quarter-wave matching techniques are generally effective and commonly used, they have one important contraindication, noted above – They often result in structures that are large. That is, any structure which employs a quarter-wave match will be at least $\lambda/4$ long, and $\lambda/4$ is typically large compared to the associated electronics. Other transmission line matching techniques – and in particular, single stub matching (Section 3.23) – typically result in structures which are significantly smaller.

The impedance inversion property of quarter-wavelength lines has applications beyond impedance matching. The following example demonstrates one such application:

Example 3.19.3: RF/DC decoupling in transistor amplifiers

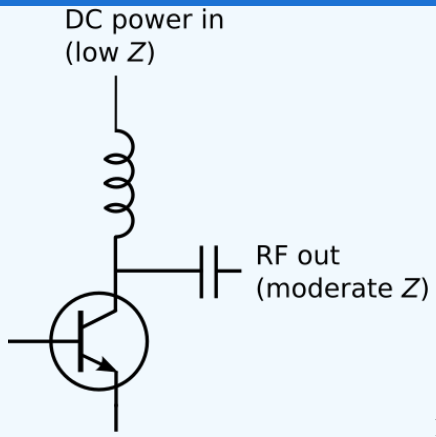


Figure 3.19.3: Use an inductor to decouple the DC input power from the RF output signal at the output of a common-emitter RF amplifier

Transistor amplifiers for RF applications often receive DC current at the same terminal which delivers the amplified RF signal, as shown in Figure 3.19.3. The power supply typically has a low output impedance. If the power supply is directly connected to the transistor, then the RF will flow predominantly in the direction of the power supply as opposed to following the desired path, which exhibits a higher impedance. This can be addressed using an inductor in series with the power supply output. This works because the inductor exhibits low impedance at DC and high impedance at RF. Unfortunately, discrete inductors are often not practical at high RF frequencies. This is because practical inductors also exhibit parallel capacitance, which tends to *decrease* impedance.

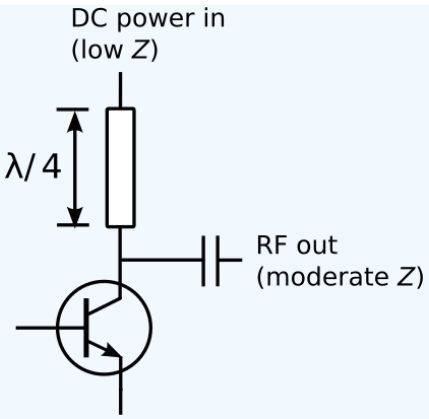


Figure 3.19.4: Decoupling of DC input power and RF output signal at the output of a common-emitter RF amplifier, using a quarter-wavelength transmission line.

A solution is to replace the inductor with a transmission line having length $\lambda/4$ as shown in Figure 3.19.4. A wavelength at DC is infinite, so the transmission line is essentially transparent to the power supply. At radio frequencies, the line transforms the low impedance of the power supply to an impedance that is very large relative to the impedance of the desired RF path. Furthermore, transmission lines on printed circuit boards are much cheaper than discrete inductors (and are always in stock!).

Additional Reading:

- "Quarter-wavelength impedance transformer" on Wikipedia.
- "Smith chart" on Wikipedia.
- "Microstrip antenna" on Wikipedia.

This page titled 3.19: Quarter-Wavelength Transmission Line is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.19: Quarter-Wavelength Transmission Line** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.20: Power Flow on Transmission Lines

It is often important to know the power associated with a wave on a transmission line. The power of the waves incident upon, reflected by, and absorbed by a load are each of interest. In this section we shall work out expressions for these powers and consider some implications in terms of the voltage reflection coefficient (Γ) and standing wave ratio (SWR).

Let's begin by considering a lossless transmission line that is oriented along the z axis. The time-average power associated with a sinusoidal wave having potential $v(z,t)$ and current $i(z,t)$ is

$$P_{av}(z) \triangleq \frac{1}{T}\int_{t_0}^{t_0+T} v(z,t)\ i(z,t)\ dt$$

where $T \triangleq 2\pi/f$ is one period of the wave and t_0 is the start time for the integration. Since the time-average power of a sinusoidal signal does not change with time, t_0 may be set equal to zero without loss of generality.

Let us now calculate the power of a wave incident from $z<0$ on a load impedance Z_L at $z=0$. We may express the associated potential and current as follows:

$$v^+(z,t) = \left|V_0^+\right|\cos\left(\omega t - \beta z + \phi\right)$$

$$i^+(z,t) = \frac{\left|V_0^+\right|}{Z_0}\cos\left(\omega t - \beta z + \phi\right)$$

And so the associated time-average power is

$$\begin{aligned} P_{av}^+(z) &= \frac{1}{T}\int_0^T v^+(z,t)\ i^+(z,t)\ dt \\ &= \frac{\left|V_0^+\right|^2}{Z_0}\cdot\frac{1}{T}\int_0^T \cos^2\left(\omega t - \beta z + \phi\right)\ dt \end{aligned} \tag{3.20.1}$$

Employing a well-known trigonometric identity:

$$\cos^2\theta = \frac{1}{2} + \frac{1}{2}\cos 2\theta$$

we may rewrite the integrand as follows

$$\cos^2\left(\omega t - \beta z + \phi\right) = \frac{1}{2} + \frac{1}{2}\cos\left(2\left[\omega t - \beta z + \phi\right]\right)$$

Then integrating over both sides of this quantity

$$\int_0^T \cos^2\left(\omega t - \beta z + \phi\right)\ dt = \frac{T}{2} + 0$$

The second term of the integral is zero because it is the integral of cosine over two complete periods. Subsequently, we see that the position dependence (here, the dependence on z) is eliminated. In other words, the power associated with the incident wave is the same for all $z<0$, as expected. Substituting into Equation ??? we obtain:

$$\boxed{P_{av}^+ = \frac{\left|V_0^+\right|^2}{2Z_0}} \tag{3.20.2}$$

This is the time-average power associated with the incident wave, measured at any point $z<0$ along the line.

> Equation 3.20.2 gives the time-average power associated with a wave traveling in a single direction along a lossless transmission line.

Using precisely the same procedure, we find that the power associated with the *reflected* wave is

$$P_{av}^- = \frac{\left|\Gamma V_0^+\right|^2}{2Z_0} = \left|\Gamma\right|^2\frac{\left|V_0^+\right|^2}{2Z_0}$$

or simply

$$\boxed{P_{av}^{-} = |\Gamma|^2 \, P_{av}^{+}} \tag{3.20.3}$$

Equation 3.20.3 gives the time-average power associated with the wave reflected from an impedance mismatch.

Now, what is the power P_L delivered to the load impedance Z_L? The simplest way to calculate this power is to use the principle of *conservation of power*. Applied to the present problem, this principle asserts that the power incident on the load must equal the power reflected plus the power absorbed; i.e.,

$$P_{av}^{+} = P_{av}^{-} + P_L$$

Applying the previous equations we obtain:

$$\boxed{P_L = \left(1 - |\Gamma|^2\right) P_{av}^{+}} \tag{3.20.4}$$

Equations 3.20.4 gives the time-average power transferred to a load impedance, and is equal to the difference between the powers of the incident and reflected waves.

Example 3.20.1: How important is it to match $50\ \Omega$ to $75\ \Omega$?

Two impedances which commonly appear in radio engineering are $50\ \Omega$ and $75\ \Omega$. It is not uncommon to find that it is necessary to connect a transmission line having a $50\ \Omega$ characteristic impedance to a device, circuit, or system having a $75\ \Omega$ input impedance, or vice-versa. If no attempt is made to match these impedances, what fraction of the power will be delivered to the termination, and what fraction of power will be reflected? What is the SWR?

Solution

The voltage reflection coefficient going from $50\ \Omega$ transmission line to a $75\ \Omega$ load is

$$\Gamma = \frac{75 - 50}{75 + 50} = +0.2$$

The fraction of power reflected is $|\Gamma|^2 = 0.04$, which is 4%. The fraction of power transmitted is $1 - |\Gamma|^2$, which is 96%. Going from a $50\ \Omega$ transmission line to a $75\ \Omega$ termination changes only the sign of Γ, and therefore, the fractions of reflected and transmitted power remain 4% and 96%, respectively. In either case (from Section 3.14):

$$\text{SWR} = \frac{1 + |\Gamma|}{1 - |\Gamma|} = 1.5$$

This is often acceptable, but may not be good enough in some particular applications. Suffice it to say that it is not necessarily required to use an impedance matching device to connect $50\ \Omega$ to $75\ \Omega$ devices.

This page titled 3.20: Power Flow on Transmission Lines is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.20: Power Flow on Transmission Lines** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.21: Impedance Matching - General Considerations

"Impedance matching" refers to the problem of transforming a particular impedance Z_L into a modified impedance Z_{in}. The problem of impedance matching arises because it is not convenient, practical, or desirable to have all devices in a system operate at the same input and output impedances. Here are just a few of the issues:

- It is not convenient or practical to market coaxial cables having characteristic impedance equal to every terminating impedance that might be encountered.
- Different types of antennas operate at different impedances, and the impedance of most antennas vary significantly with frequency.
- Different types of amplifiers operate most effectively at different output impedances. For example, amplifiers operating as current sources operate most effectively with low output impedance, whereas amplifiers operating as voltage sources operate most effectively with high output impedances.
- Independently of the above issue, techniques for the design of transistor amplifiers rely on intentionally *mismatching* impedances; i.e., matching to an impedance different than that which maximizes power transfer or minimizes reflection. In other words, various design goals are met by applying particular impedances to the input and output ports of the transistor.[1]

For all of these reasons, electrical engineers frequently find themselves with the task of transforming a particular impedance Z_L into a modified impedance Z_{in}.

The reader is probably already familiar with many approaches to the impedance matching problem that employ discrete components and do not require knowledge of electromagnetics.[2] To list just a few of these approaches: transformers, resistive (lossy) matching, single-reactance matching, and two-reactance ("L" network) matching. However, all of these have limitations. Perhaps the most serious limitations pertain to the performance of discrete components at high frequencies. Here are just a few of the most common problems:

- Practical resistors actually behave as ideal resistors in series with ideal inductors
- Practical capacitors actually behave as ideal capacitors in series with ideal resistors
- Practical inductors behave as ideal inductors in parallel with ideal capacitors, and in series with ideal resistors.

All of this makes the use of discrete components increasingly difficult with increasing frequency.

One possible solution to these types of problems is to more precisely model each component, and then to account for the non-ideal behavior by incorporating the appropriate models in the analysis and design process. Alternatively, one may consider ways to replace particular troublesome components – or, in some cases, all discrete components – with transmission line devices. The latter approach is particularly convenient in circuits implemented on printed circuit boards at frequencies in the UHF band and higher, since the necessary transmission line structures are easy to implement as microstrip lines and are relatively compact since the wavelength is relatively small. However, applications employing transmission lines as components in impedance matching devices can be found at lower frequencies as well.

1. For a concise introduction to this concept, see Chapter 10 of S.W. Ellingson, *Radio Systems Engineering*, Cambridge Univ. Press, 2016.↩
2. For an overview, see Chapter 9 of S.W. Ellingson, *Radio Systems Engineering*, Cambridge Univ. Press, 2016.↩

This page titled 3.21: Impedance Matching - General Considerations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.21: Impedance Matching - General Considerations** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.22: Single-Reactance Matching

An impedance matching structure can be designed using a section of transmission line combined with a discrete reactance, such as a capacitor or an inductor. In the strategy presented here, the transmission line is used to transform the real part of the load impedance or admittance to the desired value, and then the reactance is used to modify the imaginary part to the desired value. (Note the difference between this approach and the quarter-wave technique described in Section 3.19. In that approach, the first transmission line is used to zero the imaginary part.) There are two versions of this strategy, which we will now consider separately.

The first version is shown in Figure 3.22.1. The purpose of the transmission line is to transform the load impedance Z_L into a new impedance Z_1 for which $\text{Re}\{Z_1\} = \text{Re}\{Z_{in}\}$. This can be done by solving the equation (from Section 3.15)

$$\text{Re}\{Z_1\} = \text{Re}\left\{Z_0 \frac{1+\Gamma e^{-j2\beta l}}{1-\Gamma e^{-j2\beta l}}\right\} \tag{3.22.1}$$

for l, using a numerical search, or using the Smith chart.[1] The characteristic impedance Z_0 and phase propagation constant β of the transmission line are independent variables and can be selected for convenience. Normally, the smallest value of l that satisfies Equation 3.22.1 is desired. This value will be $\leq \lambda/4$ because the real part of Z_1 spans all possible values every $\lambda/4$.

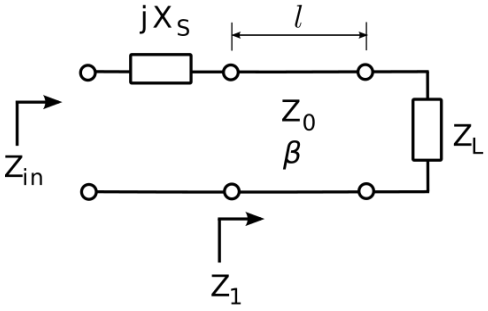


Figure 3.22.1: Single-reactance matching with a series reactance.

After matching the real component of the impedance in this manner, the imaginary component of Z_1 may then be transformed to the desired value ($\text{Im}\{Z_{in}\}$) by attaching a reactance X_s in series with the transmission line input, yielding $Z_{in} = Z_1 + jX_S$. Therefore, we choose

$$X_s = \text{Im}\{Z_{in} - Z_1\}$$

The sign of X_s determines whether this reactance is a capacitor ($X_s < 0$) or inductor ($X_s > 0$), and the value of this component is determined from X_s and the design frequency.

Example 3.22.1: Single reactance in series

Design a match consisting of a transmission line in series with a single capacitor or inductor that matches a source impedance of 50Ω to a load impedance of $33.9 + j17.6\ \Omega$ at 1.5 GHz. The characteristic impedance and phase velocity of the transmission line are 50Ω and $0.6c$ respectively.

Solution

From the problem statement: $Z_{in} \triangleq Z_S = 50\ \Omega$ and $Z_L = 33.9 + j17.6\ \Omega$ are the source and load impedances respectively at $f = 1.5$ GHz. The characteristic impedance and phase velocity of the transmission line are $Z_0 = 50\ \Omega$ and $v_p = 0.6c$ respectively.

The reflection coefficient Γ (i.e., Z_L with respect to the characteristic impedance of the transmission line) is

$$\Gamma \triangleq \frac{Z_L - Z_0}{Z_L + Z_0} \cong -0.142 + j0.239$$

The length l of the primary line (that is, the one that connects the two ports of the matching structure) is determined using the equation:

$$\text{Re}\{Z_1\} = \text{Re}\left\{Z_0 \frac{1+\Gamma e^{-j2\beta l}}{1-\Gamma e^{-j2\beta l}}\right\}$$

where here $\text{Re}\{Z_1\} = \text{Re}\{Z_S\} = 50\ \Omega$. So a more-specific form of the equation that can be solved for βl (as a step toward finding l) is:

$$1 = \text{Re}\left\{\frac{1+\Gamma e^{-j2\beta l}}{1-\Gamma e^{-j2\beta l}}\right\}$$

By trial and error (or using the Smith chart if you prefer) we find $\beta l \cong 0.408$ rad for the primary line, yielding $Z_1 \cong 50.0 + j29.0\ \Omega$ for the input impedance after attaching the primary line.

We may now solve for l as follows: Since $v_p = \omega/\beta$ (Section 3.8), we find

$$\beta = \frac{\omega}{v_p} = \frac{2\pi f}{0.6c} \cong 52.360\ \text{rad/m}$$

Therefore $l = (\beta l)/\beta \cong 7.8$ mm.

The impedance of the series reactance should be $jX_s \cong -j29.0\ \Omega$ to cancel the imaginary part of Z_1. Since the sign of this impedance is negative, it must be a capacitor. The reactance of a capacitor is $-1/\omega C$, so it must be true that

$$-\frac{1}{2\pi f C} \cong -29.0\ \Omega$$

Thus, we find the series reactance is a capacitor of value $C \cong 3.7pF$.

The second version of the single-reactance strategy is shown in Figure 3.22.2. The difference in this scheme is that the reactance is attached in parallel. In this case, it is easier to work the problem using admittance (i.e., reciprocal impedance) as opposed to impedance; this is because the admittance of parallel reactances is simply the sum of the associated admittances; i.e.,

$$Y_{in} = Y_1 + jB_p$$

where $Y_{in} = 1/Z_{in}$, $Y_1 = 1/Z_1$, and B_p is the discrete parallel susceptance; i.e., the imaginary part of the discrete parallel admittance.

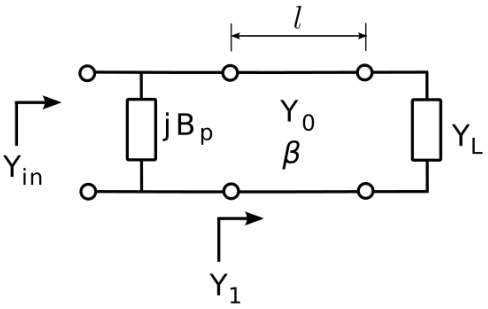


Figure 3.22.2: Single-reactance matching with a parallel reactance

So, the procedure is as follows. The transmission line is used to transform Y_L into a new admittance Y_1 for which $\text{Re}\{Y_1\} = \text{Re}\{Y_{in}\}$. First, we note that

$$Y_1 \triangleq \frac{1}{Z_1} = Y_0\frac{1-\Gamma e^{-j2\beta l}}{1+\Gamma e^{-j2\beta l}}$$

where $Y_0 \triangleq 1/Z_0$ is characteristic *admittance*. Again, the characteristic impedance Z_0 and phase propagation constant β of the transmission line are independent variables and can be selected for convenience. In the present problem, we aim to solve the equation

$$\text{Re}\{Y_1\} = \text{Re}\left\{Y_0\frac{1-\Gamma e^{-j2\beta l}}{1+\Gamma e^{-j2\beta l}}\right\}$$

for the smallest value of *l*, using a numerical search or using the Smith chart. After matching the real component of the admittances in this manner, the imaginary component of the resulting admittance may then be transformed to the desired value by attaching the susceptance B_p in parallel with the transmission line input. Since we desire jB_p in parallel with Y_1 to be Y_{in}, the desired value is

$$B_p = \text{Im}\{Y_{in} - Y_1\}$$

The sign of B_p determines whether this is a capacitor ($B_p > 0$) or inductor ($B_p < 0$), and the value of this component is determined from B_p and the design frequency.

In the following example, we address the same problem raised in Example 3.22.1, now using the parallel reactance approach:

Example 3.22.2: Single reactance in parallel

Design a match consisting of a transmission line in parallel with a single capacitor or inductor that matches a source impedance of 50Ω to a load impedance of $33.9 + j17.6\ \Omega$ at 1.5 GHz. The characteristic impedance and phase velocity of the transmission line are 50Ω and $0.6c$ respectively.

Solution

From the problem statement: $Z_{in} \triangleq Z_S = 50\ \Omega$ and $Z_L = 33.9 + j17.6\ \Omega$ are the source and load impedances respectively at $f = 1.5$ GHz. The characteristic impedance and phase velocity of the transmission line are $Z_0 = 50\ \Omega$ and $v_p = 0.6c$ respectively.

The reflection coefficient Γ (i.e., Z_L with respect to the characteristic impedance of the transmission line) is

$$\Gamma \triangleq \frac{Z_L - Z_0}{Z_L + Z_0} \cong -0.142 + j0.239$$

The length l of the primary line (that is, the one that connects the two ports of the matching structure) is the solution to:

$$\mathrm{Re}\{Y_1\} = \mathrm{Re}\left\{Y_0 \frac{1 - \Gamma e^{-j2\beta l}}{1 + \Gamma e^{-j2\beta l}}\right\}$$

where here $\mathrm{Re}\{Y_1\} = \mathrm{Re}\{1/Z_S\} = 0.02$ mho and $Y_0 = 1/Z_0 = 0.02$ mho. So the equation to be solved for βl (as a step toward finding l) is:

$$1 = \mathrm{Re}\left\{\frac{1 - \Gamma e^{-j2\beta l}}{1 + \Gamma e^{-j2\beta l}}\right\}$$

By trial and error (or the Smith chart) we find $\beta l \cong 0.126$ rad for the primary line, yielding $Y_1 \cong 0.0200 - j0.0116$ mho for the input admittance after attaching the primary line.

We may now solve for l as follows: Since $v_p = \omega/\beta$ (Section 3.8), we find

$$\beta = \frac{\omega}{v_p} = \frac{2\pi f}{0.6c} \cong 52.360\ \mathrm{rad/m}$$

Therefore, $l = (\beta l)/\beta \cong 2.4$ mm.

The admittance of the parallel reactance should be $jB_p \cong +j0.0116$ mho to cancel the imaginary part of Y_1. The associated impedance is $1/jB_p \cong -j86.3\ \Omega$. Since the sign of this impedance is negative, it must be a capacitor. The reactance of a capacitor is $-1/\omega C$, so it must be true that

$$-\frac{1}{2\pi f C} \cong -86.3\ \Omega$$

Thus, we find the parallel reactance is a capacitor of value $C \cong 1.2pF$.

Comparing this result to the result from the series reactance method (Example 3.22.1), we see that the necessary length of transmission line is much shorter, which is normally a compelling advantage. The tradeoff is that the parallel capacitance is much smaller and an accurate value may be more difficult to achieve.

---

1. For more about the Smith chart, see “Additional Reading” at the end of this section.↩

This page titled 3.22: Single-Reactance Matching is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.22: Single-Reactance Matching** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

3.23: Single-Stub Matching

In Section 3.22, we considered impedance matching schemes consisting of a transmission line combined with a reactance which is placed either in series or in parallel with the transmission line. In many problems, the required discrete reactance is not practical because it is not a standard value, or because of non-ideal behavior at the desired frequency (see Section 3.21 for more about this), or because one might simply wish to avoid the cost and logistical issues associated with an additional component. Whatever the reason, a possible solution is to replace the discrete reactance with a transmission line "stub" – that is, a transmission line which has been open- or short-circuited. Section 3.16 explains how a stub can replace a discrete reactance. Figure 3.23.1 shows a practical implementation of this idea implemented in microstrip. This section explains the theory, and we'll return to this implementation at the end of the section.

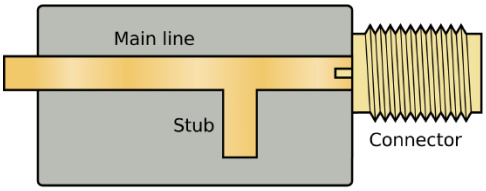


Figure 3.23.1: A practical implementation of a singlestub impedance match using microstrip transmission line. Here, the stub is open-circuited. (CC BY SA 3.0; Spinningspark)

Figure 3.23.2 shows the scheme. This scheme is usually implemented using the parallel reactance approach, as depicted in the figure. Although a series reactance scheme is also possible in principle, it is usually not as convenient. This is because most transmission lines use one of their two conductors as a local datum; e.g., the ground plane of a printed circuit board for microstrip line is tied to ground, and the outer conductor ("shield") of a coaxial cable is usually tied to ground. This is contrast to a discrete reactance (such as a capacitor or inductor), which does not require that either of its terminals be tied to ground. This issue is avoided in the parallel-attached stub because the parallel-attached stub and the transmission line to which it is attached *both* have one terminal at ground.

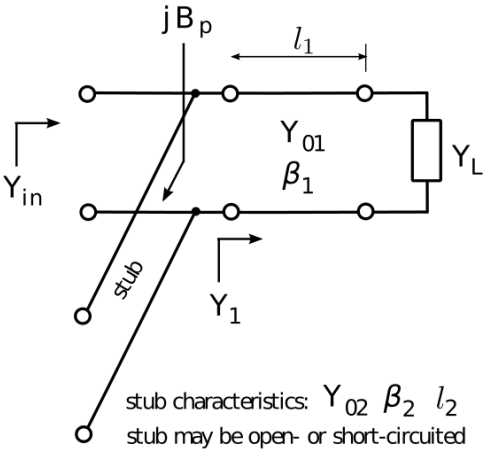


Figure 3.23.2: Single-stub matching

The single-stub matching procedure is essentially the same as the single parallel reactance method, except the parallel reactance is implemented using a short- or open-circuited stub as opposed a discrete inductor or capacitor. Since parallel reactance matching is most easily done using admittances, [admittance] it is useful to express $Z_{in}(l) = +jZ_0 \tan \beta l$ and $Z_{in}(l) = -jZ_0 \cot \beta l$ (input impedance of an open- and short-circuited stub, respectively, from Section 3.16) in terms of susceptance:

$$B_p = -Y_{02} \cot(\beta_2 l_2) \quad \text{short-circuited stub}$$

$$B_p = +Y_{02} \tan(\beta_2 l_2) \quad \text{open-circuited stub}$$

As in the main line, the characteristic impedance $Z_{02} = 1/Y_{02}$ is an independent variable and is chosen for convenience.

A final question is when should you use a short-circuited stub, and when should you use an open-circuited stub? Given no other basis for selection, the termination that yields the shortest stub is chosen. An example of an "other basis for selection" that frequently comes up is whether DC might be present on the line. If DC is present with the signal of interest, then a short circuit termination without some kind of remediation to prevent a short circuit for DC would certainly be a bad idea.

In the following example we address the same problem raised in Section 3.22 (Examples 3.22.1 and 3.22.2), now using the single-stub approach:

Example 3.23.1: Single stub matching.

Design a single-stub match that matches a source impedance of 50Ω to a load impedance of $33.9 + j17.6\ \Omega$. Use transmission lines having characteristic impedances of 50Ω throughout, and leave your answer in terms of wavelengths.

Solution

From the problem statement: $Z_{in} \triangleq Z_S = 50\ \Omega$ and $Z_L = 33.9 + j17.6\ \Omega$ are the source and load impedances respectively. $Z_0 = 50\ \Omega$ is the characteristic impedance of the transmission lines to be used. The reflection coefficient Γ (i.e., Z_L with respect to the characteristic impedance of the transmission line) is

$$\Gamma \triangleq \frac{Z_L - Z_0}{Z_L + Z_0} \cong -0.142 + j0.239$$

The length l_1 of the primary line (that is, the one that connects the two ports of the matching structure) is the solution to the equation (from Section 3.22):

$$\mathrm{Re}\{Y_1\} = \mathrm{Re}\left\{Y_{01}\frac{1-\Gamma e^{-j2\beta_1 l_1}}{1+\Gamma e^{-j2\beta_1 l_1}}\right\}$$

where here $\mathrm{Re}\{Y_1\} = \mathrm{Re}\{1/Z_S\} = 0.02$ mho and $Y_{01} = 1/Z_0 = 0.02$ mho. Also note

$$2\beta_1 l_1 = 2\left(\frac{2\pi}{\lambda}\right) l_1 = 4\pi\frac{l_1}{\lambda}$$

where λ is the wavelength in the transmission line. So the equation to be solved for l_1 is:

$$1 = \mathrm{Re}\left\{\frac{1-\Gamma e^{-j4\pi l_1/\lambda}}{1+\Gamma e^{-j4\pi l_1/\lambda}}\right\}$$

By trial and error (or using the Smith chart; see "Additional Reading" at the end of this section) we find for the primary line $l_1 \cong 0.020\lambda$, yielding $Y_1 \cong 0.0200 - j0.0116$ mho for the input admittance after attaching the primary line.

We now seek the shortest stub having an input admittance of $\cong +j0.0116$ mho to cancel the imaginary part of Y_1. For an open-circuited stub, we need

$$B_p = +Y_0 \tan 2\pi l_2/\lambda \cong +j0.0116\ \text{mho}$$

The smallest value of l_2 for which this is true is $\cong 0.084\lambda$. For a short-circuited stub, we need

$$B_p = -Y_0 \cot 2\pi l_2/\lambda \cong +j0.0116\ \text{mho}$$

The smallest positive value of l_2 for which this is true is $\cong 0.334\lambda$; i.e., much longer. Therefore, we choose the open-circuited stub with $l_2 \cong 0.084\lambda$. Note the stub is attached in parallel at the source end of the primary line.

Single-stub matching is a very common method for impedance matching using microstrip lines at frequences in the UHF band (300-3000 MHz) and above. In Figure 3.23.1, the top (visible) traces comprise one conductor, whereas the ground plane (underneath, so not visible) comprises the other conductor. The end of the stub is not connected to the ground plane, so the termination is an open circuit. A short circuit termination is accomplished by connecting the end of the stub to the ground plane using a *via*; that is, a plated-through that electrically connects the top and bottom layers.

This page titled 3.23: Single-Stub Matching is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **3.23: Single-Stub Matching** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

CHAPTER OVERVIEW

4: Vector Analysis

A *vector* is a mathematical object that has both a scalar part (i.e., a magnitude and possibly a phase), as well as a direction. Many physical quantities are best described as vectors. For example, the rate of movement through space can be described as *speed*; i.e., as a scalar having SI base units of m/s. However, this quantity is more completely described as *velocity*; i.e., as a vector whose scalar part is speed and direction indicates the direction of movement. Similarly, force is a vector whose scalar part indicates magnitude (SI base units of N), and direction indicates the direction in which the force is applied. Electric and magnetic fields are also best described as vectors.

4.1: Vector Arithmetic
4.2: Cartesian Coordinates
4.3: Cylindrical Coordinates
4.4: Spherical Coordinates
4.5: Gradient
4.6: Divergence
4.7: Divergence Theorem
4.8: Curl
4.9: Stokes' Theorem
4.10: The Laplacian Operator

Contributors and Attributions

- Ellingson, Steven W. (2018) Electromagnetics, Vol. 1. Blacksburg, VA: VT Publishing. https://doi.org/10.21061/electromagnetics-vol-1 Licensed with CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0. Report adoption of this book here. If you are a professor reviewing, adopting, or adapting this textbook please help us understand a little more about your use by filling out this form.

This page titled 4: Vector Analysis is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

4.1: Vector Arithmetic

In mathematical notation, a real-valued vector $\mathbf{A}$ is said to have a magnitude $A = |\mathbf{A}|$ and direction $\hat{\mathbf{a}}$ such that

$$\mathbf{A} = A\hat{\mathbf{a}}$$

where $\hat{\mathbf{a}}$ is a unit vector (i.e., a real-valued vector having magnitude equal to one) having the same direction as $\mathbf{A}$. If a vector is complex-valued, then A is similarly complex-valued.

Cartesian Coordinate System

Fundamentals of vector arithmetic are most easily grasped using the Cartesian coordinate system. This system is shown in Figure 4.1.1. Note carefully the relative orientation of the x, y, and z axes. This orientation is important. For example, there are two directions that are perpendicular to the $z = 0$ plane (in which the x- and y-axes lie), but the $+z$ axis is specified to be one of these in particular.

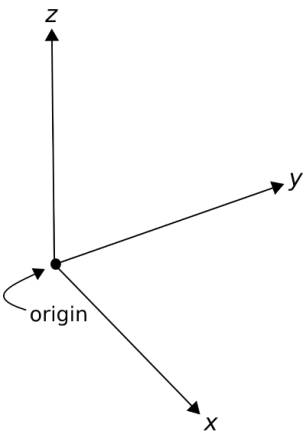


Figure 4.1.1: Cartesian coordinate system. (CC BY SA 4.0; K. Kikkeri).

Position-Fixed vs. Position-Free Vectors

It is often convenient to describe a position in space as a vector for which the magnitude is the distance from the origin of the coordinate system and for which the direction is measured from the origin toward the position of interest. This is shown in Figure 4.1.2. These *position vectors* are "position-fixed" in the sense that they are defined with respect to a single point in space, which in this case is the origin. Position vectors can also be defined as vectors that are defined with respect to some other point in space, in which case they are considered position-fixed to that position.

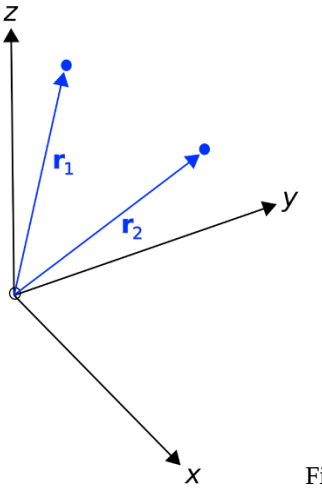


Figure 4.1.2: Position vectors. The vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ are position-fixed and refer to particular locations. (CC BY SA 4.0; K. Kikkeri).

Position-free vectors, on the other hand, are not defined with respect to a particular point in space. An example is shown in Figure 4.1.2. Particles 1 m apart may both be traveling at 2 m/s in the same direction. In this case, the velocity of each particle can be described using the same vector, even though the particles are located at different points in space.

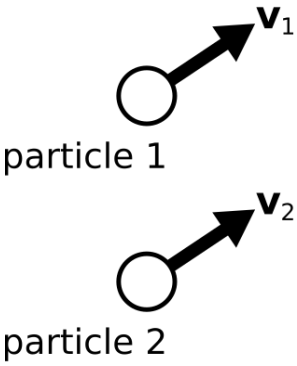


Figure 4.1.3: Two particles exhibiting the same velocity. In this case, the velocity vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are position-free and equal. (CC BY SA 4.0; K. Kikkeri).

Position-free vectors are said to be equal if they have the same magnitudes and directions. Position-fixed vectors, on the other hand, must also be referenced to the same position (e.g., the origin) to be considered equal.

Basis Vectors

Each coordinate system is defined in terms of three basis vectors which concisely describe all possible ways to traverse three-dimensional space. A basis vector is a position-free unit vector that is perpendicular to all other basis vectors for that coordinate system. The basis vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ of the Cartesian coordinate system are shown in Figure 4.1.4. In this notation, $\hat{\mathbf{x}}$ indicates the direction in which x increases most rapidly, $\hat{\mathbf{y}}$ indicates the direction in which y increases most rapidly, and $\hat{\mathbf{z}}$ indicates the direction in which z increases most rapidly. Alternatively, you might interpret $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ as unit vectors that are parallel to the x-, y-, and z-axes and point in the direction in which values along each axis increase.

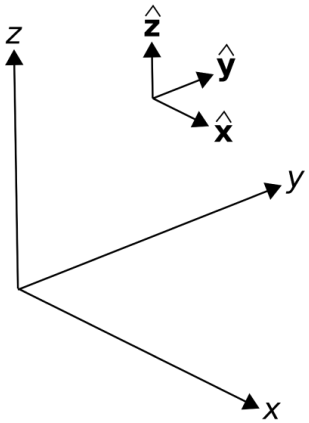


Figure 4.1.4: Basis vectors in the Cartesian coordinate system. (CC BY SA 4.0; K. Kikkeri).

Vectors in the Cartesian Coordinate System

In Cartesian coordinates, we may describe any vector $\mathbf{A}$ as follows:

$$\mathbf{A} = \hat{\mathbf{x}}A_x + \hat{\mathbf{y}}A_y + \hat{\mathbf{z}}A_z$$

where A_x, A_y, and A_z are scalar quantities describing the components of $\mathbf{A}$ in each of the associated directions, as shown in Figure 4.1.5. This description makes it clear that the magnitude of $\mathbf{A}$ is:

$$|\mathbf{A}| = \sqrt{A_x^2 + A_y^2 + A_z^2}$$

and therefore, we can calculate the associated unit vector as

$$\begin{aligned}\hat{\mathbf{a}} = \frac{\mathbf{A}}{|\mathbf{A}|} &= \frac{\mathbf{A}}{\sqrt{A_x^2 + A_y^2 + A_z^2}} \\ &= \hat{\mathbf{x}}A_x\left(A_x^2 + A_y^2 + A_z^2\right)^{-1/2} \\ &\quad + \hat{\mathbf{y}}A_y\left(A_x^2 + A_y^2 + A_z^2\right)^{-1/2} \\ &\quad + \hat{\mathbf{z}}A_z\left(A_x^2 + A_y^2 + A_z^2\right)^{-1/2}\end{aligned}$$

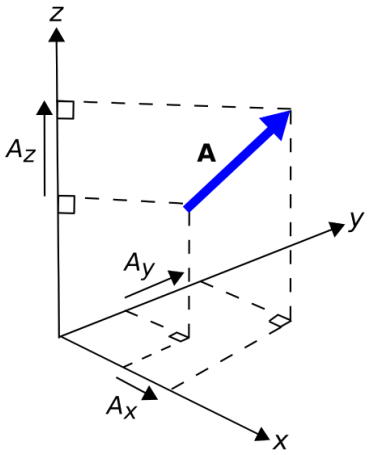


Figure 4.1.5: Components of a vector A in the Cartesian coordinate system. (CC BY SA 4.0; K. Kikkeri).

Vector Addition and Subtraction

It is common to add and subtract vectors. For example, vectors describing two forces $\mathbf{A}$ and $\mathbf{B}$ applied to the same point can be described as a single force vector $\mathbf{C}$ that is the sum of $\mathbf{A}$ and $\mathbf{B}$; i.e., $\mathbf{C} = \mathbf{A} + \mathbf{B}$. This addition is quite simple in the Cartesian coordinate system:

$$\begin{aligned}\mathbf{C} &= \mathbf{A} + \mathbf{B} \\ &= (\hat{\mathbf{x}}A_x + \hat{\mathbf{y}}A_y + \hat{\mathbf{z}}A_z) + (\hat{\mathbf{x}}B_x + \hat{\mathbf{y}}B_y + \hat{\mathbf{z}}B_z) \\ &= \hat{\mathbf{x}}\,(A_x + B_x) + \hat{\mathbf{y}}\,(A_y + B_y) + \hat{\mathbf{z}}\,(A_z + B_z)\end{aligned}$$

In other words, the $\hat{\mathbf{x}}$ component of $\mathbf{C}$ is the sum of the $\hat{\mathbf{x}}$ components of $\mathbf{A}$ and $\mathbf{B}$, and similarly for $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$. From the above example, it is clear that vector addition is commutative; i.e.,

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

In other words, vectors may be added in any order. Vector subtraction is defined similarly:

$$\begin{aligned}\mathbf{D} &= \mathbf{A} - \mathbf{B} \\ &= (\hat{\mathbf{x}}A_x + \hat{\mathbf{y}}A_y + \hat{\mathbf{z}}A_z) - (\hat{\mathbf{x}}B_x + \hat{\mathbf{y}}B_y + \hat{\mathbf{z}}B_z) \\ &= \hat{\mathbf{x}}\,(A_x - B_x) + \hat{\mathbf{y}}\,(A_y - B_y) + \hat{\mathbf{z}}\,(A_z - B_z)\end{aligned}$$

In other words, the $\hat{\mathbf{x}}$ component of $\mathbf{D}$ is the difference of the $\hat{\mathbf{x}}$ components of $\mathbf{A}$ and $\mathbf{B}$, and similarly for $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$. Like scalar subtraction, vector subtraction is *not* commutative.

Relative Positions and Distances

A common task in vector analysis is to describe the position of one point in space relative to a different point in space. Let us identify those two points using the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, as indicated in Figure 4.1.6. We may identify a third vector $\mathbf{r}_{12}$ as the position of $\mathbf{r}_2$ relative to $\mathbf{r}_1$:

$$\mathbf{r}_{12} \triangleq \mathbf{r}_2 - \mathbf{r}_1$$

Now $|\mathbf{r}_{12}|$ is the distance between these points, and $\mathbf{r}_{12}/\,|\mathbf{r}_{12}|$ is a unit vector indicating the direction to $\mathbf{r}_2$ from $\mathbf{r}_1$.

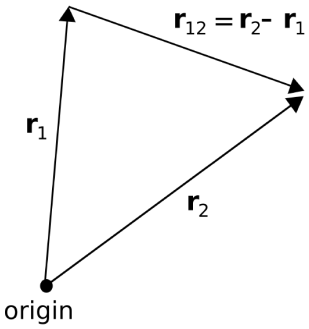


Figure 4.1.6: Relative position (distance and direction) between locations identified by their position vector. (CC BY SA 4.0; K. Kikkeri).

Example 4.1.1: Direction and distance between positions.

Consider two positions, identified using the position vectors $\mathbf{r}_1 = 2\hat{\mathbf{x}} + 3\hat{\mathbf{y}} + 1\hat{\mathbf{z}}$ and $\mathbf{r}_2 = 1\hat{\mathbf{x}} - 2\hat{\mathbf{y}} + 3\hat{\mathbf{z}}$, both expressed in units of meters. Find the direction vector that points from $\mathbf{r}_1$ to $\mathbf{r}_2$, the distance between these points, and the associated unit vector.

Solution

The vector that points from $\mathbf{r}_1$ to $\mathbf{r}_2$ is

$$\begin{aligned}\mathbf{R} &= \mathbf{r}_2 - \mathbf{r}_1 \\ &= (1-2)\hat{\mathbf{x}} + (-2-3)\hat{\mathbf{y}} + (3-1)\hat{\mathbf{z}} \\ &= -\hat{\mathbf{x}} - 5\hat{\mathbf{y}} + 2\hat{\mathbf{z}}\end{aligned}$$

The distance between $\mathbf{r}_1$ and $\mathbf{r}_2$ is simply the magnitude of this vector:

$$|\mathbf{R}| = \sqrt{(-1)^2 + (-5)^2 + (2)^2} \cong 5.48 \text{ m}$$

The unit vector $\hat{\mathbf{R}}$ is simply $\mathbf{R}$ normalized to have unit magnitude:

$$\begin{aligned}\hat{\mathbf{R}} &= \mathbf{R}/|\mathbf{R}| \\ &\cong (-\hat{\mathbf{x}} - 5\hat{\mathbf{y}} + 2\hat{\mathbf{z}})/5.48 \\ &\cong -0.182\hat{\mathbf{x}} - 0.913\hat{\mathbf{y}} + 0.365\hat{\mathbf{z}}\end{aligned}$$

Multiplication of a Vector by a Scalar

Let's say a particular force is specified by a vector $\mathbf{F}$. What is the new vector if this force is doubled? The answer is simply $2\mathbf{F}$ – that is, twice the magnitude applied in the same direction. This is an example of scalar multiplication of a vector. Generalizing, the product of the scalar α and the vector $\mathbf{A}$ is simply $\alpha\mathbf{A}$.

Scalar ("Dot") Product of Vectors

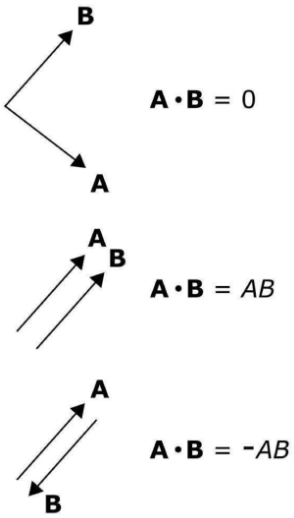


Figure 4.1.7: Special cases of the dot product. (CC BY SA 4.0; K. Kikkeri).

Another common task in vector analysis is to determine the similarity in the direction in which two vectors point. In particular, it is useful to have a metric which, when applied to the vectors $\mathbf{A} = \hat{\mathbf{a}}A$ and $\mathbf{B} = \hat{\mathbf{b}}B$, has the following properties (Figure 4.1.7):

- If $\mathbf{A}$ is perpendicular to $\mathbf{B}$, the result is zero.
- If $\mathbf{A}$ and $\mathbf{B}$ point in the same direction, the result is AB.
- If $\mathbf{A}$ and $\mathbf{B}$ point in opposite directions, the result is $-AB$.
- Results intermediate to these conditions depend on the angle ψ between $\mathbf{A}$ and $\mathbf{B}$, measured as if $\mathbf{A}$ and $\mathbf{B}$ were arranged "tail-to-tail" as shown in Figure 4.1.8.

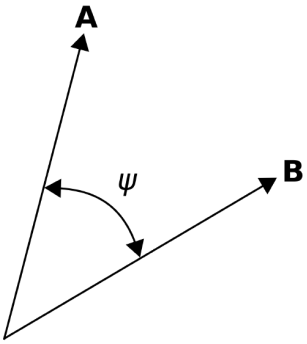


Figure 4.1.8: Calculation of the dot product. (CC BY SA 4.0; K. Kikkeri).

In vector analysis, this operator is known as the *scalar product* (not to be confused with scalar multiplication) or the *dot product*. The dot product is written $\mathbf{A} \cdot \mathbf{B}$ and is given in general by the expression:

$$\mathbf{A} \cdot \mathbf{B} = AB \cos \psi \tag{4.1.1}$$

Note that this expression yields the special cases previously identified, which are $\psi = \pi/2$, $\psi = 0$, and $\psi = \pi$, respectively. The dot product is commutative; i.e.,

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$$

The dot product is also distributive; i.e.,

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$$

The dot product has some other useful properties. For example, note:

$$\begin{aligned} \mathbf{A} \cdot \mathbf{A} &= (\hat{\mathbf{x}}A_x + \hat{\mathbf{y}}A_y + \hat{\mathbf{z}}A_z) \cdot (\hat{\mathbf{x}}A_x + \hat{\mathbf{y}}A_y + \hat{\mathbf{z}}A_z) \\ &= \hat{\mathbf{x}} \cdot \hat{\mathbf{x}}A_x^2 + \hat{\mathbf{x}} \cdot \hat{\mathbf{y}}A_xA_y + \hat{\mathbf{x}} \cdot \hat{\mathbf{z}}A_xA_z \\ &+ \hat{\mathbf{y}} \cdot \hat{\mathbf{x}}A_xA_y + \hat{\mathbf{y}} \cdot \hat{\mathbf{y}}A_y^2 + \hat{\mathbf{y}} \cdot \hat{\mathbf{z}}A_yA_z \\ &+ \hat{\mathbf{z}} \cdot \hat{\mathbf{x}}A_xA_z + \hat{\mathbf{z}} \cdot \hat{\mathbf{y}}A_yA_z + \hat{\mathbf{z}} \cdot \hat{\mathbf{z}}A_z^2 \end{aligned}$$

which looks pretty bad until you realize that

$$\hat{\mathbf{x}} \cdot \hat{\mathbf{x}} = \hat{\mathbf{y}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{z}} \cdot \hat{\mathbf{z}} = 1$$

and any other dot product of basis vectors is zero. Thus, the whole mess simplifies to:

$$\mathbf{A} \cdot \mathbf{A} = A_x^2 + A_y^2 + A_z^2$$

This is the square of the magnitude of $\mathbf{A}$, so we have discovered that

$$\mathbf{A} \cdot \mathbf{A} = |\mathbf{A}|^2 = A^2$$

Applying the same principles to the dot product of potentially different vectors $\mathbf{A}$ and $\mathbf{B}$, we find:

$$\mathbf{A} \cdot \mathbf{B} = (\hat{\mathbf{x}} A_x + \hat{\mathbf{y}} A_y + \hat{\mathbf{z}} A_z) \cdot (\hat{\mathbf{x}} B_x + \hat{\mathbf{y}} B_y + \hat{\mathbf{z}} B_z) \tag{4.1.2}$$

$$= A_x B_x + A_y B_y + A_z B_z \tag{4.1.3}$$

This is a particularly easy way to calculate the dot product, since it eliminates the problem of determining the angle ψ. In fact, an easy way to calculate ψ given $\mathbf{A}$ and $\mathbf{B}$ is to first calculate the dot product using Equation 4.1.3 and then use the result to solve Equation 4.1.1 for ψ.

Example 4.1.2: Angle between two vectors

Consider the position vectors $\mathbf{C} = 2\hat{\mathbf{x}} + 3\hat{\mathbf{y}} + 1\hat{\mathbf{z}}$ and $\mathbf{D} = 3\hat{\mathbf{x}} - 2\hat{\mathbf{y}} + 2\hat{\mathbf{z}}$, both expressed in units of meters. Find the angle between these vectors.

Solution

From Equation 4.1.1

$$\mathbf{C} \cdot \mathbf{D} = CD \cos\psi$$

where $C = |\mathbf{C}|$, $D = |\mathbf{D}|$, and ψ is the angle we seek. From Equation 4.1.3:

$$\begin{aligned} \mathbf{C} \cdot \mathbf{D} &= C_x D_x + C_y D_y + C_z D_z \\ &= 2 \cdot 3 + 3 \cdot (-2) + 1 \cdot 2\text{m}^2 \\ &= 2\ \text{m}^2 \end{aligned}$$

also

$$C = \sqrt{C_x^2 + C_y^2 + C_z^2} \cong 3.742\ \text{m}$$

$$D = \sqrt{D_x^2 + D_y^2 + D_z^2} \cong 4.123\ \text{m}$$

so

$$\cos\psi = \frac{\mathbf{C} \cdot \mathbf{D}}{CD} \cong 0.130$$

Taking the inverse cosine, we find $\psi = 82.6^\circ$.

Cross Product

The cross product is a form of vector multiplication that results in a vector that is perpendicular to both of the operands. The definition is as follows:

$$\mathbf{A} \times \mathbf{B} = \hat{\mathbf{n}} AB \sin\psi_{AB} \tag{4.1.4}$$

As shown in Figure 4.1.9, the unit vector $\hat{\mathbf{n}}$ is determined by the "right hand rule." Using your right hand, curl your fingers to traverse the angle ψ_{AB} beginning at $\mathbf{A}$ and ending at $\mathbf{B}$, and then $\hat{\mathbf{n}}$ points in the direction of your fully-extended thumb.

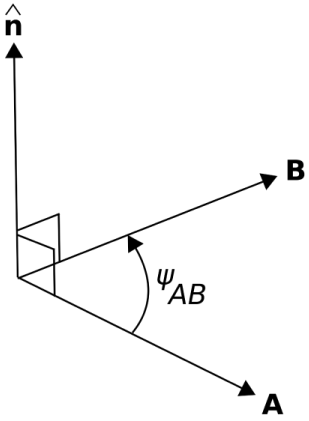


Figure 4.1.9: The cross product $\mathbf{A} \times \mathbf{B}$. (CC BY SA 4.0; K. Kikkeri).

It should be apparent that the cross product is not commutative but rather is *anticommutative*; that is,

$$\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}$$

You can confirm this for yourself using either Equation 4.1.4 or by applying the right-hand rule.

The cross product is distributive:

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$$

There are two useful special cases of the cross product that are worth memorizing. The first is the cross product of a vector with itself, which is zero:

$$\mathbf{A} \times \mathbf{A} = 0$$

The second is the cross product of vectors that are perpendicular; i.e., for which $\psi_{AB} = \pi/2$. In this case:

$$\mathbf{A} \times \mathbf{B} = \hat{\mathbf{n}} AB$$

Using these principles, note:

$$\hat{\mathbf{x}} \times \hat{\mathbf{x}} = \hat{\mathbf{y}} \times \hat{\mathbf{y}} = \hat{\mathbf{z}} \times \hat{\mathbf{z}} = 0$$

whereas

$$\hat{\mathbf{x}} \times \hat{\mathbf{y}} = \hat{\mathbf{z}}$$

$$\hat{\mathbf{y}} \times \hat{\mathbf{z}} = \hat{\mathbf{x}}$$

$$\hat{\mathbf{z}} \times \hat{\mathbf{x}} = \hat{\mathbf{y}}$$

A useful diagram that summarizes these relationships is shown in Figure 4.1.10.

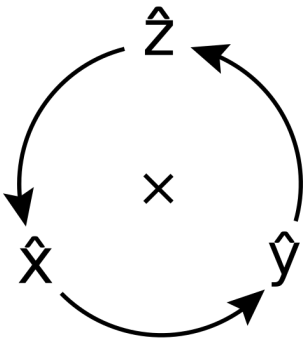


Figure 4.1.10: Cross products among basis vectors in the Cartesian system. The cross product of any two basis vectors is the third basis vector when the order of operands is counter-clockwise, as shown in the diagram, and is −1 times the third basis vector when the order of operands is clockwise with respect to the arrangement in the diagram. (CC BY SA 4.0; K. Kikkeri).

It is typically awkward to "manually" determine $\hat{\mathbf{n}}$ in Equation 4.1.4. However, in Cartesian coordinates the cross product may be calculated as:

$$\begin{aligned}\mathbf{A}\times\mathbf{B} = \quad & \hat{\mathbf{x}}\,(A_yB_z - A_zB_y) \\ & + \hat{\mathbf{y}}\,(A_zB_x - A_xB_z) \\ & + \hat{\mathbf{z}}\,(A_xB_y - A_yB_x)\end{aligned}$$

This may be easier to remember as a matrix determinant:

$$\mathbf{A}\times\mathbf{B} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}$$

Similar expressions are available for other coordinate systems.

Vector analysis routinely requires expressions involving both dot products and cross products in different combinations. Often, these expressions may be simplified, or otherwise made more convenient, using the vector identities listed in Appendix B3.

This page titled 4.1: Vector Arithmetic is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **4.1: Vector Arithmetic** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

4.2: Cartesian Coordinates

The Cartesian coordinate system is introduced in Section 4.1. Concepts described in that section – i.e., the dot product and cross product – are described in terms of the Cartesian system. In this section, we identify some additional features of this system that are useful in subsequent work and also set the stage for alternative systems; namely the cylindrical and spherical coordinate systems.

Integration Over Length

Consider a vector field $\mathbf{A} = \hat{\mathbf{x}}A(\mathbf{r})$, where $\mathbf{r}$ is a position vector. What is the integral of $\mathbf{A}$ over some curve $\mathcal{C}$ through space? To answer this question, we first identify a differential-length segment of the curve. Note that this segment of the curve can be described as

$$d\mathbf{l} = \hat{\mathbf{x}}dx + \hat{\mathbf{y}}dy + \hat{\mathbf{z}}dz$$

The contribution to the integral for that segment of the curve is simply $\mathbf{A} \cdot d\mathbf{l}$. We integrate to obtain the result; i.e.,

$$\int_{\mathcal{C}} \mathbf{A} \cdot d\mathbf{l}$$

For example, if $\mathbf{A} = \hat{\mathbf{x}}A_0$ (i.e., $A(\mathbf{r})$ is a constant) and if $\mathcal{C}$ is a straight line from $x = x_1$ and $x = x_2$ along some constant y and z, then $d\mathbf{l} = \hat{\mathbf{x}}dx$, $\mathbf{A} \cdot d\mathbf{l} = A_0 dx$, and subsequently the above integral is

$$\int_{x_1}^{x_2} A_0 dx = A_0 \left(x_2 - x_1\right)$$

In particular, notice that if $A_0 = 1$, then this integral gives the length of $\mathcal{C}$. Although the formalism seems unnecessary in this simple example, it becomes very useful when integrating over paths that vary in more than one direction and with more complicated integrands.

Note that the Cartesian system was an appropriate choice for preceding example because this allowed two of the three basis directions (i.e., y and z) to be immediately eliminated from the calculation. Said differently, the preceding example is expressed with the minimum number of varying coordinates in the Cartesian system. Here's a counter-example. If $\mathcal{C}$ had been a circle in the $z = 0$ plane, then the problem would have required two basis directions to be considered – namely, both x and y. In this case, another system – namely, cylindrical coordinates (Section 4.3) – minimizes the number of varying coordinates (to just one, which is ϕ).

Integration Over Area

Now we ask the question, what is the integral of some vector field $\mathbf{A}$ over some surface $\mathcal{S}$? The answer is

$$\int_{\mathcal{S}} \mathbf{A} \cdot d\mathbf{s} \tag{4.2.1}$$

We refer to $d\mathbf{s}$ as the differential surface element, which is a vector having magnitude equal to the differential area ds, and is normal (perpendicular) to each point on the surface. There are actually two such directions. We'll return to clear up the ambiguity in a moment. Now, as an example, if $\mathbf{A} = \hat{\mathbf{z}}$ and $\mathcal{S}$ is the surface bounded by $x_1 \leq x \leq x_2$, $y_1 \leq y \leq y_2$, then

$$d\mathbf{s} = \hat{\mathbf{z}}\ dx\ dy$$

since $dx\,dy$ is differential surface area in the $z = 0$ plane and $\hat{\mathbf{z}}$ is normal to the $z = 0$ plane. So $\mathbf{A} \cdot d\mathbf{s} = dxdy$, and subsequently the integral in Equation 4.2.1 becomes

$$\int_{x_1}^{x_2} \int_{y_1}^{y_2} dx\ dy = \left(x_2 - x_1\right)\left(y_2 - y_1\right)$$

Note that this has turned out to be a calculation of area.

Once again, we see the Cartesian system was an appropriate choice for this example because this choice minimizes the number of varying coordinates; in the above example, the surface of integration is described by a constant value of z with variable values of x and y. If the surface had instead been a cylinder or a sphere, not only would all three basis directions be variable, but also the surface normal would be variable, making the problem dramatically more complicated.

Now let's return to the issue of the direction of $d\mathbf{s}$. We chose $+\hat{\mathbf{z}}$, but why not choose $-\hat{\mathbf{z}}$ – also a normal to the surface – as this direction? To answer simply, the resulting area would be negative. "Negative area" is the expected ("positive") area, except with respect to the opposite-facing normal vector. In the present problem, the sign of the area is not important, but in some problems this sign becomes important. One example of a class of problems for which the sign of area is important is when the quantity of interest is a *flux*. If $\mathbf{A}$ were a flux density, then the integration over area that we just performed indicates the magnitude *and direction* of flux, and so the direction chosen for $d\mathbf{s}$ defines the direction of positive flux. Section 2.4 describes the electric field in terms of a flux (i.e., as electric flux density $\mathbf{D}$), in which case positive flux flows away from a positively-charged source.

Integration Over Volume

Another common task in vector analysis is integration of some quantity over a volume. Since the procedure is the same for scalar or vector quantities, we shall consider integration of a scalar quantity $A(\mathbf{r})$ for simplicity. First, we note that the contribution from a differential volume element

$$dv = dx\ dy\ dz$$

is $A(\mathbf{r})\ dv$, so the integral over the volume $\mathcal{V}$ is

$$\int_{\mathcal{V}} A(\mathbf{r})\ dv$$

For example, if $A(\mathbf{r}) = 1$ and $\mathcal{V}$ is a cube bounded by $x_1 \leq x \leq x_2$, $y_1 \leq y \leq y_2$, and $z_1 \leq z \leq z_2$, then the above integral becomes

$$\int_{x_1}^{x_2} \int_{y_1}^{y_2} \int_{z_1}^{z_2} dx\ dy\ dz = (x_2 - x_1)\,(y_2 - y_1)\,(z_2 - z_1)$$

i.e., this is a calculation of volume.

The Cartesian system was an appropriate choice for this example because $\mathcal{V}$ is a cube, which is easy to describe in Cartesian coordinates and relatively difficult to describe in any other coordinate system.

This page titled 4.2: Cartesian Coordinates is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **4.2: Cartesian Coordinates** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

4.3: Cylindrical Coordinates

Cartesian coordinates (Section 4.1) are not convenient in certain cases. One of these is when the problem has cylindrical symmetry. For example, in the Cartesian coordinate system, the cross-section of a cylinder concentric with the z-axis requires two coordinates to describe: x and y. However, this cross section can be described using a single parameter – namely the radius – which is ρ in the cylindrical coordinate system. This results in a dramatic simplification of the mathematics in some applications.

The cylindrical system is defined with respect to the Cartesian system in Figure 4.3.1. In lieu of x and y, the cylindrical system uses ρ, the distance measured from the closest point on the z axis, and ϕ, the angle measured in a plane of constant z, beginning at the $+x$ axis ($\phi = 0$) with ϕ increasing toward the $+y$ direction.

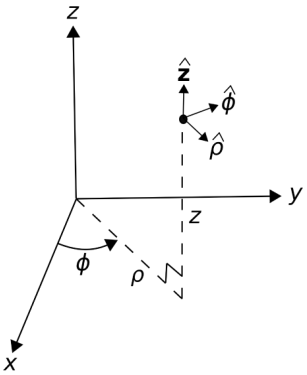


Figure 4.3.1: Cylindrical coordinate system and associated basis vectors. (CC BY SA 4.0; K. Kikkeri).

The basis vectors in the cylindrical system are $\hat{\rho}$, $\hat{\phi}$, and $\hat{\mathbf{z}}$. As in the Cartesian system, the dot product of like basis vectors is equal to one, and the dot product of unlike basis vectors is equal to zero. The cross products of basis vectors are as follows:

$$\hat{\rho} \times \hat{\phi} = \hat{\mathbf{z}}$$

$$\hat{\phi} \times \hat{\mathbf{z}} = \hat{\rho}$$

$$\hat{\mathbf{z}} \times \hat{\rho} = \hat{\phi}$$

A useful diagram that summarizes these relationships is shown in Figure 4.3.2.

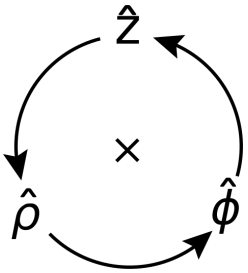


Figure 4.3.2: Cross products among basis vectors in the cylindrical system. (See Figure 4.1.10 for instructions on the use of this diagram.) (CC BY SA 4.0; K. Kikkeri).

The cylindrical system is usually less useful than the Cartesian system for identifying absolute and relative positions. This is because the basis directions depend on position. For example, $\hat{\rho}$ is directed radially outward from the $\hat{\mathbf{z}}$ axis, so $\hat{\rho} = \hat{\mathbf{x}}$ for locations along the x-axis but $\hat{\rho} = \hat{\mathbf{y}}$ for locations along the y axis. Similarly, the direction $\hat{\phi}$ varies as a function of position. To overcome this awkwardness, it is common to set up a problem in cylindrical coordinates in order to exploit cylindrical symmetry, but at some point to convert to Cartesian coordinates. Here are the conversions:

$$x = \rho \cos \phi$$

$$y = \rho \sin \phi$$

and z is identical in both systems. The conversion from Cartesian to cylindrical is as follows:

$$\rho = \sqrt{x^2 + y^2}$$

$$\phi = \arctan(y, x)$$

where arctan is the *four-quadrant* inverse tangent function; i.e., $\arctan(y/x)$ in the first quadrant ($x > 0$, $y > 0$), but possibly requiring an adjustment for the other quadrants because the signs of both x and y are individually significant.

Similarly, it is often necessary to represent basis vectors of the cylindrical system in terms of Cartesian basis vectors and vice-versa. Conversion of basis vectors is straightforward using dot products to determine the components of the basis vectors in the new system. For example, $\hat{\mathbf{x}}$ in terms of the basis vectors of the cylindrical system is

$$\hat{\mathbf{x}} = \hat{\rho}\,(\hat{\rho}\cdot\hat{\mathbf{x}}) + \hat{\phi}\left(\hat{\phi}\cdot\hat{\mathbf{x}}\right) + \hat{\mathbf{z}}\,(\hat{\mathbf{z}}\cdot\hat{\mathbf{x}})$$

The last term is of course zero since $\hat{\mathbf{z}}\cdot\hat{\mathbf{x}} = 0$. Calculation of the remaining terms requires dot products between basis vectors in the two systems, which are summarized in Table 4.3.1. Using this table, we find

$$\hat{\mathbf{x}} = \hat{\rho}\cos\phi - \hat{\phi}\sin\phi$$

$$\hat{\mathbf{y}} = \hat{\rho}\sin\phi + \hat{\phi}\cos\phi$$

and of course $\hat{\mathbf{z}}$ requires no conversion. Going from Cartesian to cylindrical, we find

$$\hat{\rho} = \hat{\mathbf{x}}\cos\phi + \hat{\mathbf{y}}\sin\phi$$

$$\hat{\phi} = -\hat{\mathbf{x}}\sin\phi + \hat{\mathbf{y}}\cos\phi$$

Table 4.3.1: Dot products between basis vectors in the cylindrical and Cartesian coordinate systems.

·	$\hat{\rho}$	$\hat{\phi}$	$\hat{\mathbf{z}}$
$\hat{\mathbf{x}}$	$\cos\phi$	$-\sin\phi$	0
$\hat{\mathbf{y}}$	$\sin\phi$	$\cos\phi$	0
$\hat{\mathbf{z}}$	0	0	1

Integration Over Length

A differential-length segment of a curve in the cylindrical system is described in general as

$$d\mathbf{l} = \hat{\rho}d\rho + \hat{\phi}\rho d\phi + \hat{\mathbf{z}}\;dz$$

Note that the contribution of the ϕ coordinate to differential length is $\rho d\phi$, not simply $d\phi$. This is because ϕ is an angle, not a distance. To see why the associated distance is $\rho d\phi$, consider the following. The circumference of a circle of radius ρ is $2\pi\rho$. If only a fraction of the circumference is traversed, the associated arclength is the circumference scaled by $\phi/2\pi$, where ϕ is the angle formed by the traversed circumference. Therefore, the distance is $2\pi\rho\cdot\phi/2\pi = \rho\phi$, and the differential distance is $\rho d\phi$.

As always, the integral of a vector field $\mathbf{A}(\mathbf{r})$ over a curve $\mathcal{C}$ is

$$\int_{\mathcal{C}} \mathbf{A}\cdot d\mathbf{l}$$

To demonstrate the cylindrical system, let us calculate the integral of $\mathbf{A}(\mathbf{r}) = \hat{\phi}$ when $\mathcal{C}$ is a circle of radius ρ_0 in the $z = 0$ plane, as shown in Figure 4.3.3. In this example, $d\mathbf{l} = \hat{\phi}\;\rho_0\;d\phi$ since $\rho = \rho_0$ and $z = 0$ are both constant along $\mathcal{C}$. Subsequently, $\mathbf{A}\cdot d\mathbf{l} = \rho_0 d\phi$ and the above integral is

$$\int_0^{2\pi} \rho_0\;d\phi = 2\pi\rho_0$$

i.e., this is a calculation of circumference.

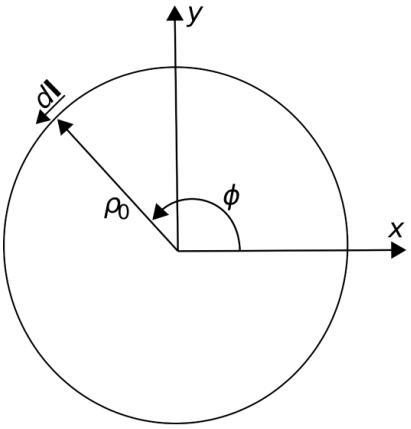


Figure 4.3.3: Example in cylindrical coordinates: The circumference of a circle. (CC BY SA 4.0; K. Kikkeri).

Note that the cylindrical system is an appropriate choice for the preceding example because the problem can be expressed with the minimum number of varying coordinates in the cylindrical system. If we had attempted this problem in the Cartesian system, we would find that both x and y vary over $\mathcal{C}$, and in a relatively complex way.

Integration Over Area

Now we ask the question, what is the integral of some vector field $\mathbf{A}$ over a circular surface $\mathcal{S}$ in the $z = 0$ plane having radius ρ_0? This is shown in Figure Figure 4.3.4. The differential surface vector in this case is

$$d\mathbf{s} = \hat{\mathbf{z}}\ (d\rho)\,(\rho d\phi) = \hat{\mathbf{z}}\ \rho\ d\rho\ d\phi \tag{4.3.1}$$

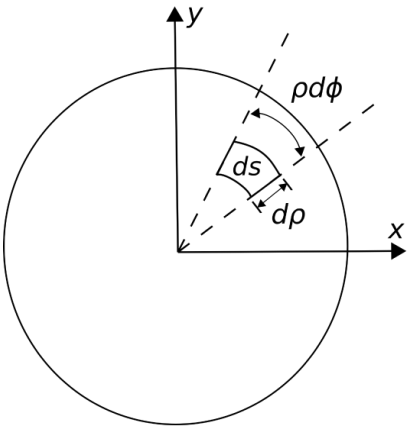


Figure 4.3.4: Example in cylindrical coordinates: The area of a circle. (CC BY SA 4.0; K. Kikkeri).

The quantities in parentheses of Equation 4.3.1 are the radial and angular dimensions, respectively. The direction of $d\mathbf{s}$ indicates the direction of positive flux – see the discussion in Section 4.2 for an explanation. In general, the integral over a surface is

$$\int_{\mathcal{S}} \mathbf{A}\cdot d\mathbf{s}$$

To demonstrate, let's consider $\mathbf{A} = \hat{\mathbf{z}}$; in this case $\mathbf{A}\cdot d\mathbf{s} = \rho\ d\rho\ d\phi$ and the integral becomes

$$\int_0^{\rho_0}\int_0^{2\pi} \rho\ d\rho\ d\phi = \left(\int_0^{\rho_0} \rho\ d\rho\right)\left(\int_0^{2\pi} d\phi\right) \tag{4.3.2}$$

$$= \left(\frac{1}{2}\rho_0^2\right)(2\pi) \tag{4.3.3}$$

$$= \pi\rho_0^2 \tag{4.3.4}$$

which we recognize as the area of the circle, as expected. The corresponding calculation in the Cartesian system is quite difficult in comparison.

Whereas the previous example considered a planar surface, we might consider instead a curved surface. Here we go. What is the integral of a vector field $\mathbf{A} = \hat{\rho}$ over a cylindrical surface $\mathcal{S}$ concentric with the z axis having radius ρ_0 and extending from $z = z_1$ to $z = z_2$? This is shown in Figure 4.3.5.

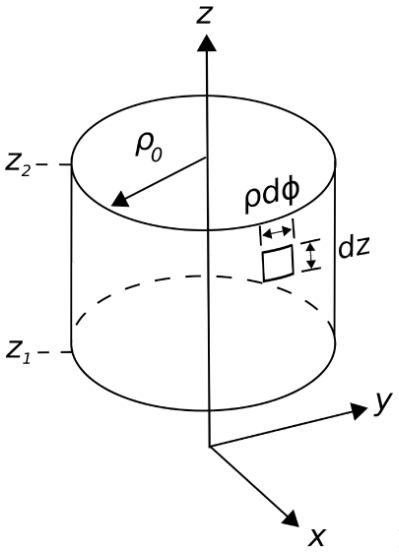


Figure 4.3.5: Example in cylindrical coordinates: The area of the curved surface of a cylinder. (CC BY SA 4.0; K. Kikkeri).

The differential surface vector in this case is

$$d\mathbf{s} = \hat{\rho}\,(\rho_0 d\phi)\,(dz) = \hat{\rho}\rho_0\ d\phi\ dz$$

The integral is

$$\int_{\mathcal{S}} \mathbf{A}\cdot d\mathbf{s} = \int_0^{2\pi}\int_{z_1}^{z_2} \rho_0\ d\phi dz \tag{4.3.5}$$

$$= \rho_0 \left(\int_0^{2\pi} d\phi\right)\left(\int_{z_1}^{z_2} dz\right) \tag{4.3.6}$$

$$= 2\pi\rho_0\,(z_2 - z_1) \tag{4.3.7}$$

which is the area of $\mathcal{S}$, as expected. Once again, the corresponding calculation in the Cartesian system is quite difficult in comparison.

Integration Over Volume

The differential volume element in the cylindrical system is

$$dv \ = \ d\rho\ \ (\rho d\phi)\ \ dz \ = \ \rho\ d\rho\ d\phi\ dz$$

For example, if $A(\mathbf{r}) = 1$ and the volume $\mathcal{V}$ is a cylinder bounded by $\rho \le \rho_0$ and $z_1 \le z \le z_2$, then

$$\int_{\mathcal{V}} A(\mathbf{r})\ dv = \int_0^{\rho_0}\int_0^{2\pi}\int_{z_1}^{z_2} \rho\ d\rho\ d\phi\ dz \tag{4.3.8}$$

$$= \left(\int_0^{\rho_0} \rho\ d\rho\right)\left(\int_0^{2\pi} d\phi\right)\left(\int_{z_1}^{z_2} dz\right) \tag{4.3.9}$$

$$= \pi\rho_0^2\,(z_2 - z_1) \tag{4.3.10}$$

i.e., area times length, which is volume.

Once again, the procedure above is clearly more complicated than is necessary if we are interested only in computing volume. However, if the integrand is not constant-valued then we are no longer simply computing volume. In this case, the formalism is appropriate and possibly necessary.

This page titled 4.3: Cylindrical Coordinates is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **4.3: Cylindrical Coordinates** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

4.4: Spherical Coordinates

The spherical coordinate system is defined with respect to the Cartesian system in Figure 4.4.1. The spherical system uses r, the distance measured from the origin; θ, the angle measured from the $+z$ axis toward the $z = 0$ plane; and ϕ, the angle measured in a plane of constant z, identical to ϕ in the cylindrical system.

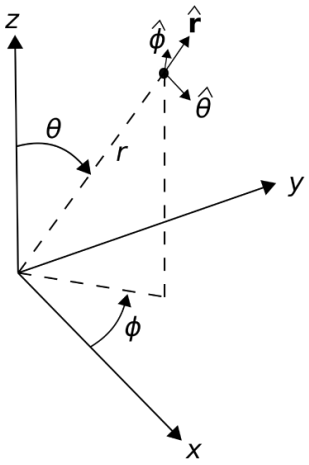


Figure 4.4.1: Spherical coordinate system and associated basis vectors. (CC BY SA 4.0; K. Kikkeri).

Spherical coordinates are preferred over Cartesian and cylindrical coordinates when the geometry of the problem exhibits spherical symmetry. For example, in the Cartesian coordinate system, the surface of a sphere concentric with the origin requires all three coordinates (x, y, and z) to describe. However, this surface can be described using a single constant parameter – the radius r – in the spherical coordinate system. This leads to a dramatic simplification in the mathematics in certain applications.

The basis vectors in the spherical system are $\hat{\mathbf{r}}$, $\hat{\theta}$, and $\hat{\phi}$. As always, the dot product of like basis vectors is equal to one, and the dot product of unlike basis vectors is equal to zero. For the cross-products, we find:

$$\hat{\mathbf{r}} \times \hat{\theta} = \hat{\phi}$$

$$\hat{\theta} \times \hat{\phi} = \hat{\mathbf{r}}$$

$$\hat{\phi} \times \hat{\mathbf{r}} = \hat{\theta}$$

A useful diagram that summarizes these relationships is shown in Figure 4.4.2.

Like the cylindrical system, the spherical system is often less useful than the Cartesian system for identifying absolute and relative positions. The reason is the same: Basis directions in the spherical system depend on position. For example, $\hat{\mathbf{r}}$ is directed radially outward from the origin, so $\hat{\mathbf{r}} = \hat{\mathbf{x}}$ for locations along the x-axis but $\hat{\mathbf{r}} = \hat{\mathbf{y}}$ for locations along the y axis and $\hat{\mathbf{r}} = \hat{\mathbf{z}}$ for locations along the z axis. Similarly, the directions of $\hat{\theta}$ and $\hat{\phi}$ vary as a function of position. To overcome this awkwardness, it is common to begin a problem in spherical coordinates, and then to convert to Cartesian coordinates at some later point in the analysis. Here are the conversions:

$$x = r\cos\phi\sin\theta \tag{4.4.1}$$

$$y = r\sin\phi\sin\theta \tag{4.4.2}$$

$$z = r\cos\theta \tag{4.4.3}$$

The conversion from Cartesian to spherical coordinates is as follows:

$$r = \sqrt{x^2 + y^2 + z^2} \tag{4.4.4}$$

$$\theta = \arccos\left(z/r\right) \tag{4.4.5}$$

$$\phi = \arctan\left(y, x\right) \tag{4.4.6}$$

where $\arctan$ is the four-quadrant inverse tangent function.

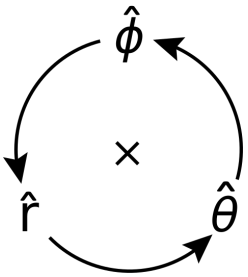


Figure 4.4.2 Cross products among basis vectors in the spherical system. (See Figure 4.1.10 for instructions on the use of this diagram.) (CC BY SA 4.0; K. Kikkeri).

Dot products between basis vectors in the spherical and Cartesian systems are summarized in Table 4.4.1. This information can be used to convert between basis vectors in the spherical and Cartesian systems, in the same manner described in Section 4.3; e.g.

$$\hat{\mathbf{x}} = \hat{\mathbf{r}}\,(\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}) + \hat{\theta}\left(\hat{\theta}\cdot\hat{\mathbf{x}}\right) + \hat{\phi}\left(\hat{\phi}\cdot\hat{\mathbf{x}}\right)$$

$$\hat{\mathbf{r}} = \hat{\mathbf{x}}\,(\hat{\mathbf{x}}\cdot\hat{\mathbf{r}}) + \hat{\mathbf{y}}\,(\hat{\mathbf{y}}\cdot\hat{\mathbf{r}}) + \hat{\mathbf{z}}\,(\hat{\mathbf{z}}\cdot\hat{\mathbf{r}})$$

and so on.

Table 4.4.1: Dot products between basis vectors in the spherical and Cartesian coordinate systems.

·	$\hat{\mathbf{r}}$	$\hat{\theta}$	$\hat{\phi}$
$\hat{\mathbf{x}}$	$\sin\theta\cos\phi$	$\cos\theta\cos\phi$	$-\sin\phi$
$\hat{\mathbf{y}}$	$\sin\theta\sin\phi$	$\cos\theta\sin\phi$	$\cos\phi$
$\hat{\mathbf{z}}$	$\cos\theta$	$-\sin\theta$	0

Example 4.4.1: Cartesian to spherical conversion

A vector field $\mathbf{G} = \hat{\mathbf{x}}\ xz/y$. Develop an expression for $\mathbf{G}$ in spherical coordinates.

Solution

Simply substitute expressions in terms of spherical coordinates for expressions in terms of Cartesian coordinates. Use Table 4.4.1 and Equations 4.4.1- 4.4.3. Making these substitutions and applying a bit of mathematical clean-up afterward, one obtains

$$\mathbf{G} = \left(\hat{\mathbf{r}}\sin\theta\cot\phi + \hat{\phi}\cos\theta\cot\phi - \hat{\phi}\right)\cdot r\cos\theta\cos\phi$$

Integration Over Length

A differential-length segment of a curve in the spherical system is

$$d\mathbf{l} = \hat{\mathbf{r}}\ dr + \hat{\theta}\ r\ d\theta + \hat{\phi}\ r\sin\theta\ d\phi$$

Note that θ is an angle, as opposed to a distance. The associated distance is $r\ d\theta$ in the θ direction. Note also that in the ϕ direction, distance is $r\ d\phi$ in the $z = 0$ plane and less by the factor $\sin\theta$ for $z <> 0$.

As always, the integral of a vector field $\mathbf{A}(\mathbf{r})$ over a curve $\mathcal{C}$ is

$$\int_{\mathcal{C}} \mathbf{A}\cdot d\mathbf{l}$$

To demonstrate line integration in the spherical system, imagine a sphere of radius a centered at the origin with "poles" at $z = +a$ and $z = -a$. Let us calculate the integral of $\mathbf{A}(\mathbf{r}) = \hat{\theta}$, where $\mathcal{C}$ is the arc drawn directly from pole to pole along the surface of the sphere, as shown in Figure 4.4.3. In this example, $d\mathbf{l} = \hat{\theta}\ a\ d\theta$ since $r = a$ and ϕ (which could be any value) are both constant along $\mathcal{C}$. Subsequently, $\mathbf{A}\cdot d\mathbf{l} = a\ d\theta$ and the above integral is

$$\int_0^{\pi} a\ d\theta = \pi a$$

i.e., half the circumference of the sphere, as expected.

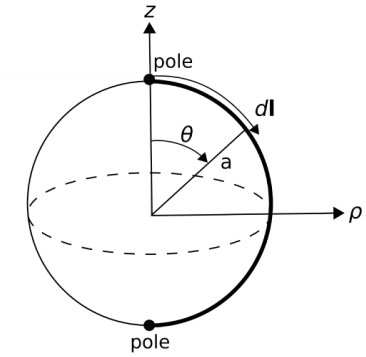


Figure 4.4.3: Example in spherical coordinates: Poleto-pole distance on a sphere. (CC BY SA 4.0; K. Kikkeri).

Note that the spherical system is an appropriate choice for this example because the problem can be expressed with the minimum number of varying coordinates in the spherical system. If we had attempted this problem in the Cartesian system, we would find that both z and either x or y (or all three) vary over $\mathcal{C}$ and in a relatively complex way.

Integration Over Area

Now we ask the question, what is the integral of some vector field $\mathbf{A}$ over the surface $\mathcal{S}$ of a sphere of radius a centered on the origin? This is shown in Figure 4.4.4. The differential surface vector in this case is

$$d\mathbf{s} = \hat{\mathbf{r}}\ \ (r\ d\theta)\,(r\sin\theta\ d\phi) = \hat{\mathbf{r}}\ r^2\sin\theta\ d\theta\ d\phi$$

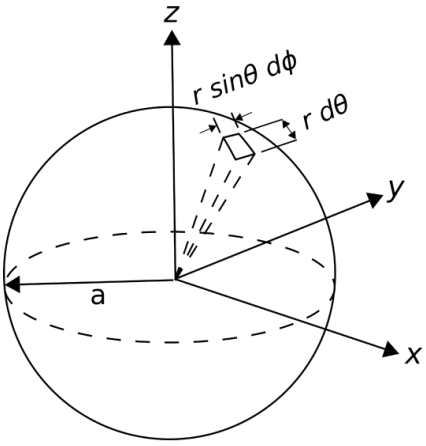


Figure 4.4.4: Example in spherical coordinates: The area of a sphere. (CC BY SA 4.0; K. Kikkeri).

As always, the direction is normal to the surface and in the direction associated with positive flux. The quantities in parentheses are the distances associated with varying θ and ϕ, respectively. In general, the integral over a surface is

$$\int_{\mathcal{S}} \mathbf{A}\cdot d\mathbf{s}$$

In this case, let's consider $\mathbf{A} = \hat{\mathbf{r}}$; in this case $\mathbf{A}\cdot d\mathbf{s} = a^2\ \sin\theta\ d\theta\ d\phi$ and the integral becomes

$$\begin{aligned}\int_0^{\pi}\int_0^{2\pi} a^2\ \sin\theta\ d\theta\ d\phi &= a^2\left(\int_0^{\pi}\sin\theta d\theta\right)\left(\int_0^{2\pi} d\phi\right)\\ &= a^2\,(2)\,(2\pi)\\ &= 4\pi a^2\end{aligned}$$

which we recognize as the area of the sphere, as expected. The corresponding calculation in the Cartesian or cylindrical systems is quite difficult in comparison.

Integration Over Volume

The differential volume element in the spherical system is

$$dv \;=\; dr\,(rd\theta)\,(r\sin\theta d\phi) \;=\; r^2 dr\;\sin\theta\;d\theta\;d\phi$$

For example, if $A(\mathbf{r}) = 1$ and the volume $\mathcal{V}$ is a sphere of radius a centered on the origin, then

$$\begin{aligned}\int_{\mathcal{V}} A(\mathbf{r})\;dv &= \int_0^a \int_0^\pi \int_0^{2\pi} r^2 dr\;\sin\theta\;d\theta\;d\phi \\ &= \left(\int_0^a r^2 dr\right)\left(\int_0^\pi \sin\theta\;d\theta\right)\left(\int_0^{2\pi} d\phi\right) \\ &= \left(\frac{1}{3}a^3\right)(2)\,(2\pi) \\ &= \frac{4}{3}\pi a^3\end{aligned}$$

which is the volume of a sphere.

This page titled 4.4: Spherical Coordinates is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **4.4: Spherical Coordinates** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

4.5: Gradient

The gradient operator is an important and useful tool in electromagnetic theory. Here's the main idea:

> The *gradient* of a scalar field is a vector that points in the direction in which the field is most rapidly increasing, with the scalar part equal to the rate of change.

A particularly important application of the gradient is that it relates the electric field intensity $\mathbf{E}(\mathbf{r})$ to the electric potential field $V(\mathbf{r})$. This is apparent from a review of Section 2.2; see in particular, the battery-charged capacitor example. In that example, it is demonstrated that:

- The *direction* of $\mathbf{E}(\mathbf{r})$ is the direction in which $V(\mathbf{r})$ decreases most quickly, and
- The *scalar part* of $\mathbf{E}(\mathbf{r})$ is the rate of change of $V(\mathbf{r})$ in that direction. Note that this is also implied by the units, since $V(\mathbf{r})$ has units of V whereas $\mathbf{E}(\mathbf{r})$ has units of V/m.

The gradient is the mathematical operation that relates the vector field $\mathbf{E}(\mathbf{r})$ to the scalar field $V(\mathbf{r})$ and is indicated by the symbol "∇" as follows:

$$\mathbf{E}(\mathbf{r}) = -\nabla V(\mathbf{r})$$

or, with the understanding that we are interested in the gradient as a function of position $\mathbf{r}$, simply

$$\mathbf{E} = -\nabla V$$

At this point we should note that the gradient is a very general concept, and that we have merely identified one application of the gradient above. In electromagnetics there are many situations in which we seek the gradient ∇f of some scalar field $f(\mathbf{r})$. Furthermore, we find that other differential operators that are important in electromagnetics can be interpreted in terms of the gradient operator ∇. These include *divergence* (Section 4.6), *curl* (Section 4.8), and the *Laplacian* (Section 4.10).

In the Cartesian system:

$$\nabla f = \hat{\mathbf{x}}\frac{\partial f}{\partial x} + \hat{\mathbf{y}}\frac{\partial f}{\partial y} + \hat{\mathbf{z}}\frac{\partial f}{\partial z} \tag{4.5.1}$$

Example 4.5.1: Gradient of a ramp function.

Find the gradient of $f = ax$ (a "ramp" having slope a along the x direction).

Solution

Here, $\partial f/\partial x = a$ and $\partial f/\partial y = \partial f/\partial z = 0$. Therefore $\nabla f = \hat{\mathbf{x}}a$. Note that ∇f points in the direction in which f most rapidly increases, and has magnitude equal to the slope of f in that direction.

The gradient operator in the cylindrical and spherical systems is given in Appendix B2.

This page titled 4.5: Gradient is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **4.5: Gradient** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

4.6: Divergence

In this section, we present the divergence operator, which provides a way to calculate the flux associated with a point in space. First, let us review the concept of flux.

The integral of a vector field over a surface is a scalar quantity known as *flux*. Specifically, the flux F of a vector field $\mathbf{A}(\mathbf{r})$ over a surface $\mathcal{S}$ is

$$\int_{\mathcal{S}} \mathbf{A} \cdot d\mathbf{s} = F$$

Note that $\mathbf{A}$ could be fairly described as a *flux density*; i.e., a quantity having units equal to the units of F, but divided by area (i.e., m^2). Also worth noting is that the flux of a vector field that has unit magnitude and is normal to all points on $\mathcal{S}$ is simply the area of $\mathcal{S}$.

It is quite useful to identify some electromagnetic quantities as either fluxes or flux densities. Here are two important examples:

- The electric flux density $\mathbf{D}$, having units of C/m^2, is a description of the electric field as a flux density. (See Section 2.4 for more about electric flux density.) The integral of $\mathbf{D}$ over a closed surface yields the enclosed charge Q_{encl}, having units of C. This relationship is known as Gauss' Law:

$$\oint_{\mathcal{S}} \mathbf{D} \cdot d\mathbf{s} = Q_{encl} \tag{4.6.1}$$

  (See Section 5.5 for more about Gauss' Law.)
- The magnetic flux density $\mathbf{B}$, having units of Wb/m^2, is a description of the magnetic field as a flux density. (See Section 2.5 for more about magnetic flux density.) The integral of $\mathbf{B}$ over a surface (open *or* closed) yields the magnetic flux Φ, having units of Wb:

$$\int_{\mathcal{S}} \mathbf{B} \cdot d\mathbf{s} = \Phi$$

  This is important because, for example, the time rate of change of Φ is proportional to electric potential. (See Section 8.3 for more about this principle, called *Faraday's Law*.)

Summarizing

> Flux is the scalar quantity obtained by integrating a vector field, interpreted in this case as a flux density, over a specified surface.

The concept of flux applies to a surface of finite size. However, what is frequently of interest is behavior at a single point, as opposed to the sum or average over a region of space. For example, returning to the idea of electric flux density ($\mathbf{D}$), perhaps we are not concerned about the total charge (units of C) enclosed by a surface, but rather the charge density (C/m^3) at a point. In this case, we could begin with Equation 4.6.1 and divide both sides of the equation by the volume V enclosed by $\mathcal{S}$:

$$\frac{\oint_{\mathcal{S}} \mathbf{D} \cdot d\mathbf{s}}{V} = \frac{Q_{encl}}{V}$$

Now we let V shrink to zero, giving us an expression that must be true at whatever point we decide to converge upon. Taking the limit as $V \to 0$:

$$\lim_{V \to 0} \frac{\oint_{\mathcal{S}} \mathbf{D} \cdot d\mathbf{s}}{V} = \lim_{V \to 0} \frac{Q_{encl}}{V}$$

The quantity on the right hand side is by definition the *volume charge density* ρ_v (units of C/m^3) at the point at which we converge. The left hand side is the *divergence* of $\mathbf{D}$, sometimes abbreviated "div $\mathbf{D}$." Thus, the above equation can be written

$$\text{div } \mathbf{D} = \rho_v$$

Summarizing

> Divergence is the flux per unit volume through an infinitesimally-small closed surface surrounding a point.

We will typically not actually want to integrate and take a limit in order to calculate divergence. Fortunately, we do not have to. It turns out that this operation can be expressed as the dot product $\nabla \cdot \mathbf{D}$; where, for example,

$$\nabla \triangleq \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}$$

in the Cartesian coordinate system. This is the *same* "∇" that appears in the definition of the gradient operator (Section 4.5) and is same operator that often arises when considering other differential operators. If we expand $\mathbf{D}$ in terms of its Cartesian components:

$$\mathbf{D} = \hat{\mathbf{x}}D_x + \hat{\mathbf{y}}D_y + \hat{\mathbf{z}}D_z$$

Then

$$\text{div } \mathbf{D} = \nabla \cdot \mathbf{D} = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}$$

This seems to make sense for two reasons. First, it is dimensionally correct. Taking the derivative of a quantity having units of C/m ^{2} with respect to distance yields a quantity having units of C/m^3. Second, it makes sense that flux from a point should be related to the sum of the rates of change of the flux density in each basis direction. Summarizing:

> The divergence of a vector field $\mathbf{A}$ is $\nabla \cdot \mathbf{A}$.

Example 4.6.1: Divergence of a uniform field

A field $\mathbf{A}$ that is constant with respect to position is said to be *uniform*. A completely general description of such a field is $\mathbf{A} = \hat{\mathbf{x}}A_x + \hat{\mathbf{y}}A_y + \hat{\mathbf{z}}A_z$ where A_x, A_y, and A_z are all constants. We see immediately that the divergence of such a field must be zero. That is, $\nabla \cdot \mathbf{A} = 0$ because each component of $\mathbf{A}$ is constant with respect to position. This also makes sense from the perspective of the "flux through an infinitesimally-small closed surface" interpretation of divergence. If the flux is uniform, the flux into the surface equals the flux out of the surface resulting in a net flux of zero.

Example 4.6.2: Divergence of a linearly-increasing field

Consider a field $\mathbf{A} = \hat{\mathbf{x}}A_0 x$ where A_0 is a constant. The divergence of $\mathbf{A}$ is $\nabla \cdot \mathbf{A} = A_0$. If we interpret $\mathbf{A}$ as a flux density, then we have found that the net flux per unit volume is simply the rate at which the flux density is increasing with distance.

To compute divergence in other coordinate systems, we merely need to know ∇ for those systems. In the cylindrical system:

$$\nabla = \hat{\rho}\frac{1}{\rho}\frac{\partial}{\partial \rho}\rho + \hat{\phi}\frac{1}{\rho}\frac{\partial}{\partial \phi}\rho + \hat{\mathbf{z}}\frac{\partial}{\partial z}$$

and in the spherical system:

$$\nabla = \hat{\mathbf{r}}\frac{1}{r^2}\frac{\partial}{\partial r}r^2 + \hat{\theta}\frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}\sin\theta + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial \phi} \tag{4.6.2}$$

Alternatively, one may use the explicit expressions for divergence given in Appendix B2.

Example 4.6.3: Divergence of a radially-decreasing field

Consider a vector field that is directed radially outward from a point and which decreases linearly with distance; i.e., $\mathbf{A} = \hat{\mathbf{r}}A_0/r$ where A_0 is a constant. In this case, the divergence is most easily computed in the spherical coordinate system since partial derivatives in all but one direction (r) equal zero. Neglecting terms that include these zero-valued partial derivatives, we find:

$$\nabla \cdot \mathbf{A} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \left[\frac{A_0}{r}\right]\right) = \frac{A_0}{r^2}$$

$$\nabla \cdot \mathbf{A} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\left[\frac{r}{r}\right]\right) = \frac{2}{r^2}$$

In other words, if we interpret $\mathbf{A}$ as a flux density, then the flux per unit volume is decreasing with as the square of distance from the origin.

It is useful to know that divergence, like ∇ itself, is a linear operator; that is, for any constant scalars a and b and vector fields $\mathbf{A}$ and $\mathbf{B}$:

$$\nabla \cdot (a\mathbf{A} + b\mathbf{B}) = a\nabla \cdot \mathbf{A} + b\nabla \cdot \mathbf{B}$$

This page titled 4.6: Divergence is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **4.6: Divergence** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

4.7: Divergence Theorem

The *Divergence Theorem* relates an integral over a volume to an integral over the surface bounding that volume. This is useful in a number of situations that arise in electromagnetic analysis. In this section, we derive this theorem.

Consider a vector field $\mathbf{A}$ representing a flux density, such as the electric flux density $\mathbf{D}$ or magnetic flux density $\mathbf{B}$. The divergence of $\mathbf{A}$ is

$$\nabla \cdot \mathbf{A} = f \tag{4.7.1}$$

where is $f(\mathbf{r})$ is the flux per unit volume through an infinitesimally-small closed surface surrounding the point at $\mathbf{r}$. Since f is flux per unit volume, we can obtain flux for any larger contiguous volume $\mathcal{V}$ by integrating over $\mathcal{V}$; i.e.,

$$\text{flux through } \mathcal{V} = \int_{\mathcal{V}} f \, dv$$

In the Cartesian system, $\mathcal{V}$ can be interpreted as a three-dimensional grid of infinitesimally-small cubes having side lengths dx, dy, and dz, respectively. Note that the flux *out of* any face of one of these cubes is equal to the flux *into* the cube that is adjacent through that face. That is, the portion of the total flux that flows between cubes cancels when added together. In fact, the only fluxes which do not cancel in the integration over $\mathcal{V}$ are those corresponding to faces which lie on the bounding surface $\mathcal{S}$, since the integration stops there. Stating this mathematically:

$$\int_{\mathcal{V}} f \, dv = \oint_{\mathcal{S}} \mathbf{A} \cdot d\mathbf{s} \tag{4.7.2}$$

Thus, we have converted a volume integral into a surface integral.

To obtain the Divergence Theorem, we return to Equation 4.7.1. Integrating both sides of that equation over $\mathcal{V}$, we obtain

$$\int_{\mathcal{V}} (\nabla \cdot \mathbf{A}) \, dv = \int_{\mathcal{V}} f \, dv$$

Now applying Equation 4.7.2 to the right hand side:

$$\boxed{\int_{\mathcal{V}} (\nabla \cdot \mathbf{A}) \, dv = \oint_{\mathcal{S}} \mathbf{A} \cdot d\mathbf{s}} \tag{4.7.3}$$

> The Divergence Theorem (Equation 4.7.3) states that the integral of the divergence of a vector field over a volume is equal to the flux of that field through the surface bounding that volume.

The principal utility of the Divergence Theorem is to convert problems that are defined in terms of quantities known throughout a volume into problems that are defined in terms of quantities known over the bounding surface and vice-versa.

This page titled 4.7: Divergence Theorem is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **4.7: Divergence Theorem** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

4.8: Curl

Curl is an operation, which when applied to a vector field, quantifies the *circulation* of that field. The concept of circulation has several applications in electromagnetics. Two of these applications correspond to directly to Maxwell's Equations:

- The circulation of an electric field is proportional to the rate of change of the magnetic field. This is a statement of the *Maxwell-Faraday Equation* (Section 8.8), which includes as a special case *Kirchoff's Voltage Law for electrostatics* (Section 5.11).
- The circulation of a magnetic field is proportional to the source current and the rate of change of the electric field. This is a statement of *Ampere's Law* (Sections 7.9 and 8.9)

Thus, we are motivated to formally define circulation and then to figure out how to most conveniently apply the concept in mathematical analysis. The result is the *curl* operator. So, we begin with the concept of circulation:

> "Circulation" is the integral of a vector field over a closed path.

Specifically, the circulation of the vector field $\mathbf{A}(\mathbf{r})$ over the closed path $\mathcal{C}$ is

$$\oint_{\mathcal{C}} \mathbf{A}\cdot d\mathbf{l}$$

The circulation of a uniform vector field is zero for any valid path. For example, the circulation of $\mathbf{A} = \hat{\mathbf{x}}A_0$ is zero because non-zero contributions at each point on $\mathcal{C}$ cancel out when summed over the closed path. On the other hand, the circulation of $\mathbf{A} = \hat{\phi}A_0$ over a circular path of constant ρ and z is a non-zero constant, since the non-zero contributions to the integral at each point on the curve are equal and accumulate when summed over the path.

Example 4.8.1: Circulation of the magnetic field intensity surrounding a line current

Consider a current I (units of A) flowing along the z axis in the $+z$ direction, as shown in Figure 4.8.1.

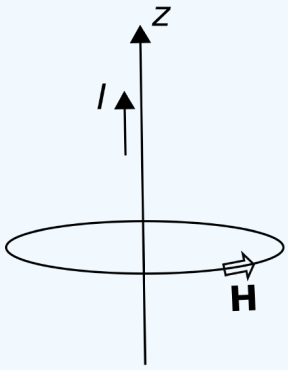


Figure 4.8.1: Magnetic field intensity due to a current flowing along the z axis. (CC BY SA 4.0; K. Klikkeri).

It is known that this current gives rise to a magnetic field intensity $\mathbf{H} = \hat{\phi}H_0/\rho$, where H_0 is a constant having units of A since the units of $\mathbf{H}$ are A/m. (Feel free to consult Section 7.5 for the details; however, no additional information is needed to follow the example being presented here.) The circulation of $\mathbf{H}$ along any circular path of radius a in a plane of constant z is therefore

$$\oint_{\mathcal{C}} \mathbf{H}\cdot d\mathbf{l} = \int_{\phi=0}^{2\pi} \left(\hat{\phi}\frac{H_0}{a}\right)\cdot\left(\hat{\phi}\ a\ d\phi\right) = 2\pi H_0$$

Note that the circulation of $\mathbf{H}$ in this case has two remarkable features: (1) It is independent of the radius of the path of integration; and (2) it has units of A, which suggests a current. In fact, it turns out that the circulation of $\mathbf{H}$ in this case is *equal* to the enclosed source current I. Furthermore, it turns out that the circulation of $\mathbf{H}$ along *any* path enclosing the source current is equal to the source current! These findings are consequences of *Ampere's Law*, as noted above.

Curl is, in part, an answer to the question of what the circulation *at a point in space* is. In other words, what is the circulation as $\mathcal{C}$ shrinks to it's smallest possible size. The answer in one sense is zero, since the arclength of $\mathcal{C}$ is zero in this limit – there is nothing to integrate over. However, if we ask instead what is the circulation *per unit area* in the limit, then the result should be the non-

trivial value of interest. To express this mathematically, we constrain $\mathcal{C}$ to lie in a plane, and define $\mathcal{S}$ to be the open surface bounded by $\mathcal{C}$ in this case. Then, we define the scalar part of the curl of $\mathbf{A}$ to be:

$$\lim_{\Delta s \to 0} \frac{\oint_{\mathcal{C}} \mathbf{A} \cdot d\mathbf{l}}{\Delta s}$$

where Δs is the area of $\mathcal{S}$, and (important!) we require $\mathcal{C}$ and $\mathcal{S}$ to lie in the plane that maximizes the above result.

Because $\mathcal{S}$ and it's boundary $\mathcal{C}$ lie in a plane, it is possible to assign a direction to the result. The chosen direction is the normal $\hat{\mathbf{n}}$ to the plane in which $\mathcal{C}$ and $\mathcal{S}$ lie. Because there are two normals at each point on a plane, we specify the one that satisfies the *right hand rule*. This rule, applied to the curl, states that the correct normal is the one which points through the plane in the same direction as the fingers of the right hand when the thumb of your right hand is aligned along $\mathcal{C}$ in the direction of integration. Why is this the correct orientation of $\hat{\mathbf{n}}$? See Section 4.9 for the answer to that question. For the purposes of this section, it suffices to consider this to be simply an arbitrary sign convention.

Now with the normal vector $\hat{\mathbf{n}}$ unambiguously defined, we can now formally define the curl operation as follows:

$$\text{curl } \mathbf{A} \triangleq \lim_{\Delta s \to 0} \frac{\hat{\mathbf{n}} \oint_{\mathcal{C}} \mathbf{A} \cdot d\mathbf{l}}{\Delta s} \tag{4.8.1}$$

where, once again, Δs is the area of $\mathcal{S}$, and we select $\mathcal{S}$ to lie in the plane that maximizes the magnitude of the above result. Summarizing:

> The curl operator quantifies the circulation of a vector field at a point.

> The magnitude of the curl of a vector field is the circulation, per unit area, at a point and such that the closed path of integration shrinks to enclose zero area while being constrained to lie in the plane that maximizes the magnitude of the result.

> The direction of the curl is determined by the right-hand rule applied to the associated path of integration.

Curl is a very important operator in electromagnetic analysis. However, the definition (Equation 4.8.1) is usually quite difficult to apply. Remarkably, however, it turns out that the curl operation can be defined in terms of the ∇ operator; that is, the same ∇ operator associated with the gradient, divergence, and Laplacian operators. Here is that definition:

$$\text{curl } \mathbf{A} \triangleq \nabla \times \mathbf{A}$$

For example: In Cartesian coordinates,

$$\nabla \triangleq \hat{\mathbf{x}} \frac{\partial}{\partial x} + \hat{\mathbf{y}} \frac{\partial}{\partial y} + \hat{\mathbf{z}} \frac{\partial}{\partial z}$$

and

$$\mathbf{A} = \hat{\mathbf{x}} A_x + \hat{\mathbf{y}} A_y + \hat{\mathbf{z}} A_z$$

so curl can be calculated as follows:

$$\nabla \times \mathbf{A} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix}$$

or, evaluating the determinant:

$$\begin{aligned} \nabla \times \mathbf{A} = \hat{\mathbf{x}} &\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \\ + \hat{\mathbf{y}} &\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \\ + \hat{\mathbf{z}} &\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \end{aligned}$$

Expressions for curl in each of the three major coordinate systems is provided in Appendix B2.

It is useful to know is that curl, like ∇ itself, is a linear operator; that is, for any constant scalars a and b and vector fields $\mathbf{A}$ and $\mathbf{B}$:

$$\nabla \times (a\mathbf{A} + b\mathbf{B}) = a\nabla \times \mathbf{A} + b\nabla \times \mathbf{B}$$

This page titled 4.8: Curl is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **4.8: Curl** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

4.9: Stokes' Theorem

Stokes' Theorem relates an integral over an open surface to an integral over the curve bounding that surface. This relationship has a number of applications in electromagnetic theory. Here is the theorem:

$$\boxed{\int_{\mathcal{S}} (\nabla \times \mathbf{A}) \cdot d\mathbf{s} = \oint_{\mathcal{C}} \mathbf{A} \cdot d\mathbf{l}}$$

where $\mathcal{S}$ is the open surface bounded by the closed path $\mathcal{C}$. The direction of the surface normal $d\mathbf{s} = \hat{\mathbf{n}} ds$ is related to the direction of integration along $\mathcal{C}$ by the *right-hand rule*, illustrated in Figure 4.9.1. In this case, the right-hand rule states that the correct normal is the one that points through the surface in the same direction as the fingers of the right hand when the thumb of your right hand is aligned along $\mathcal{C}$ in the direction of integration.

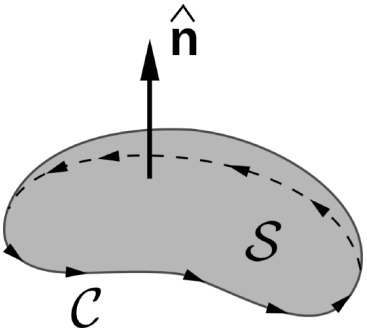


Figure 4.9.1: The relative orientations of the direction of integration $\mathcal{C}$ and surface normal $\vec{\mathbf{n}}$ in Stokes' Theorem. (CC BY SA 3.0; Cronholm144).

Stokes' Theorem is a purely mathematical result and not a principle of electromagnetics *per se*. The relevance of the theorem to electromagnetic theory is primarily as a tool in the associated mathematical analysis. Usually the theorem is employed to transform a problem expressed in terms of an integration over a surface into an integration over a closed path or vice-versa. For more information on the theorem and its derivation, see "Additional Reading" at the end of this section.

This page titled 4.9: Stokes' Theorem is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **4.9: Stokes' Theorem** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

4.10: The Laplacian Operator

The Laplacian $\nabla^2 f$ of a field $f(\mathbf{r})$ is the divergence of the gradient of that field:

$$\nabla^2 f \triangleq \nabla \cdot (\nabla f) \tag{4.10.1}$$

Note that the Laplacian is essentially a definition of the second derivative with respect to the three spatial dimensions. For example, in Cartesian coordinates,

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \tag{4.10.2}$$

as can be readily verified by applying the definitions of gradient and divergence in Cartesian coordinates to Equation 4.10.1.

The Laplacian relates the electric potential (i.e., V, units of V) to electric charge density (i.e., ρ_v, units of C/m^3). This relationship is known as *Poisson's Equation*:

$$\nabla^2 V = -\frac{\rho_v}{\epsilon}$$

where ϵ is the permittivity of the medium. The fact that V is related to ρ_v in this way should not be surprising, since electric field intensity ($\mathbf{E}$, units of V/m) is proportional to the derivative of V with respect to distance (via the gradient) and ρ_v is proportional to the derivative of $\mathbf{E}$ with respect to distance (via the divergence).

The Laplacian operator can also be applied to *vector* fields; for example, Equation 4.10.2 is valid even if the scalar field "f" is replaced with a vector field. In the Cartesian coordinate system, the Laplacian of the vector field $\mathbf{A} = \hat{\mathbf{x}} A_x + \hat{\mathbf{y}} A_y + \hat{\mathbf{z}} A_z$ is

$$\nabla^2 \mathbf{A} = \hat{\mathbf{x}} \nabla^2 A_x + \hat{\mathbf{y}} \nabla^2 A_y + \hat{\mathbf{z}} \nabla^2 A_z$$

An important application of the Laplacian operator of vector fields is the *wave equation*; e.g., the wave equation for $\mathbf{E}$ in a lossless and source-free region is

$$\nabla^2 \mathbf{E} + \beta^2 \mathbf{E} = 0$$

where β is the **phase propagation constant**.

It is sometimes useful to know that the Laplacian of a vector field can be expressed in terms of the gradient, divergence, and curl as follows:

$$\nabla^2 \mathbf{A} = \nabla (\nabla \cdot \mathbf{A}) - \nabla \times (\nabla \times \mathbf{A})$$

The Laplacian operator in the cylindrical and spherical coordinate systems is given in Appendix B2.

This page titled 4.10: The Laplacian Operator is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **4.10: The Laplacian Operator** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

CHAPTER OVERVIEW

5: Electrostatics

Electrostatics is the theory of the electric field in conditions in which its behavior is independent of magnetic fields, including

- The electric field associated with fixed distributions of electric charge
- *Capacitance* (the ability of a structure to store energy in an electric field)
- The *energy* associated with the electrostatic field
- *Steady current* induced in a conducting material in the presence of an electrostatic field (essentially, Ohm's Law)

The term "static" refers to the fact that these aspects of electromagnetic theory can be developed by assuming sources are time-invariant; we might say that electrostatics is the study of the electric field at DC. However, many aspects of electrostatics are relevant to AC, radio frequency, and higher-frequency applications as well.

5.1: Coulomb's Law
5.2: Electric Field Due to Point Charges
5.3: Charge Distributions
5.4: Electric Field Due to a Continuous Distribution of Charge
5.5: Gauss' Law - Integral Form
5.6: Electric Field Due to an Infinite Line Charge using Gauss' Law
5.7: Gauss' Law - Differential Form
5.8: Force, Energy, and Potential Difference
5.9: Independence of Path
5.10: Kirchoff's Voltage Law for Electrostatics - Integral Form
5.11: Kirchoff's Voltage Law for Electrostatics - Differential Form
5.12: Electric Potential Field Due to Point Charges
5.13: Electric Potential Field due to a Continuous Distribution of Charge
5.14: Electric Field as the Gradient of Potential
5.15: Poisson's and Laplace's Equations
5.16: Potential Field Within a Parallel Plate Capacitor
5.17: Boundary Conditions on the Electric Field Intensity (E)
5.18: Boundary Conditions on the Electric Flux Density (D)
5.19: Charge and Electric Field for a Perfectly Conducting Region
5.20: Dielectric Media
5.21: Dielectric Breakdown
5.22: Capacitance
5.23: The Thin Parallel Plate Capacitor
5.24: Capacitance of a Coaxial Structure
5.25: Electrostatic Energy

Thumbnail: Electric field lines due to a point charge in the vicinity of PEC regions (shaded) of various shapes. (CC BY SA 4.0; K. Kikkeri).

Contributors and Attributions

- Ellingson, Steven W. (2018) Electromagnetics, Vol. 1. Blacksburg, VA: VT Publishing. https://doi.org/10.21061/electromagnetics-vol-1 Licensed with CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0. Report adoption of this book here. If you are a professor reviewing, adopting, or adapting this textbook please help us understand a little more about your use by filling out this form.

This page titled 5: Electrostatics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

5.1: Coulomb's Law

Consider two charge-bearing particles in free space, identified as "particle 1" and "particle 2" in Figure 5.1.1. Let the charges borne by these particles be q_1 and q_2, and let R be the distance between them. If the particles bear charges of the same sign (i.e., if $q_1 q_2$ is positive), then the particles repel; otherwise, they attract. This repulsion or attraction can be quantified as a force experienced by each particle. Physical observations reveal that the magnitude of the force is proportional to $q_1 q_2$, and inversely proportional to R^2. For particle 2 we find:

$$\mathbf{F} = \hat{\mathbf{R}} F_0 \frac{q_1 q_2}{R^2}$$

where $\hat{\mathbf{R}}$ is the unit vector pointing from the particle 1 to the particle 2, and F_0 is a constant. The value of F_0 must have units of inverse permittivity; i.e., $(\mathrm{F/m})^{-1}$. This is most easily seen by dimensional analysis of the above relationship, including the suspected factor:

$$\frac{\mathrm{C}\cdot\mathrm{C}}{\mathrm{F/m}\cdot\mathrm{m}^2} = \frac{\mathrm{C}\cdot\mathrm{C}}{\mathrm{F}\cdot\mathrm{m}} = \frac{\mathrm{C}\cdot\mathrm{C}}{\mathrm{C/V}\cdot\mathrm{m}} = \frac{\mathrm{C}\cdot\mathrm{V}}{\mathrm{m}} = \frac{\mathrm{J}}{\mathrm{m}} = \mathrm{N}$$

where we have used the facts that 1 F = 1 C/V, 1 V = 1 J/C, and 1 N = 1 J/m. This finding suggests that $F_0 \propto \epsilon^{-1}$, where ϵ is the permittivity of the medium in which the particles exist. Observations confirm that the force is in fact inversely proportional to the permittivity, with an additional factor of $1/4\pi$ (unitless). Putting this all together we obtain what is commonly known as *Coulomb's Law*:

$$\boxed{\mathbf{F} = \hat{\mathbf{R}} \frac{q_1 q_2}{4\pi\epsilon R^2}}$$

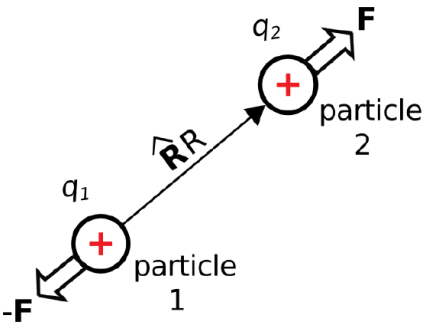


Figure 5.1.1: Coulomb's Law describes the force perceived by pairs of charged particles. (CC BY SA 4.0; K. Kikkeri)

Subsequently, the force perceived by particle 2 is equal and opposite; i.e., equal to $-\mathbf{F}$.

Separately, it is known that $\mathbf{F}$ can be described in terms of the electric field intensity $\mathbf{E}_1$ associated with particle 1:

$$\mathbf{F} = q_2 \mathbf{E}_1$$

This is essentially the *definition* of $\mathbf{E}_1$, as explained in Section 2.2. Combining this result with Coulomb's Law, we obtain a means to directly calculate the field associated with the first particle in the absence of the second particle:

$$\boxed{\mathbf{E}_1 = \hat{\mathbf{R}} \frac{q_1}{4\pi\epsilon R^2}} \tag{5.1.1}$$

where now $\hat{\mathbf{R}}R$ is the vector beginning at the particle 1 and ending at the point to be evaluated.

> The electric field intensity associated with a point charge (Equation 5.1.1) is (1) directed away from positive charge, (2) proportional to the magnitude of the charge, (3) inversely proportional to the permittivity of the medium, and (3) inversely proportional to distance squared.

We have described this result as originating from Coulomb's Law, which is based on physical observations. However, the same result may be obtained directly from Maxwell's Equations using Gauss' Law (Section 5.5).

Example 5.1.1: Electric Field of a Point Charge at the Origin

A common starting point in electrostatic analysis is the field associated with a particle bearing charge q at the origin of the coordinate system. Because the electric field is directed radially away from a positively-charged source particle in all directions, this field is most conveniently described in the spherical coordinate system. Thus, $\hat{\mathbf{R}}$ becomes $\hat{\mathbf{r}}$, R becomes r, and we have

$$\mathbf{E}(\mathbf{r}) = \hat{\mathbf{r}}\frac{q}{4\pi\epsilon r^2}$$

Here's a numerical example. What is the electric field intensity at a distance 1 μm from a single electron located at the origin, in free space? In this case, $q \cong -1.60 \times 10^{-19}$ C (don't forget that minus sign!), $\epsilon = \epsilon_0$, $r = 1$ μm, and we find:

$$\mathbf{E}(\mathbf{r}) = -\hat{\mathbf{r}}\,(1.44\ \mathrm{kV/m})$$

This is large relative to electric field strengths commonly encountered in engineering applications. The strong electric field of the electron is not readily apparent because electrons in common materials tend to be accompanied by roughly equal amounts of positive charge, such as the protons of atoms. Sometimes, however, the effect of individual electrons *does* become significant in practical electronics through a phenomenon known as *shot noise*.

This page titled 5.1: Coulomb's Law is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.1: Coulomb's Law** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.2: Electric Field Due to Point Charges

The electric field intensity associated with a single particle bearing charge q_1, located at the origin, is (Section 5.1)

$$\mathbf{E}(\mathbf{r}) = \hat{\mathbf{r}}\frac{q_1}{4\pi\epsilon r^2}$$

If this particle is instead located at some position $\mathbf{r}_1$, then the above expression may be written as follows:

$$\mathbf{E}(\mathbf{r};\mathbf{r}_1) = \frac{\mathbf{r}-\mathbf{r}_1}{|\mathbf{r}-\mathbf{r}_1|}\ \frac{q_1}{4\pi\epsilon\,|\mathbf{r}-\mathbf{r}_1|^2}$$

or, combining like terms in the denominator:

$$\mathbf{E}(\mathbf{r};\mathbf{r}_1) = \frac{\mathbf{r}-\mathbf{r}_1}{|\mathbf{r}-\mathbf{r}_1|^3}\ \frac{q_1}{4\pi\epsilon}$$

Now let us consider the field due to multiple such particles. Under the usual assumptions about the permittivity of the medium (Section 2.8), the property of superposition applies. Using this principle, we conclude:

> The electric field resulting from a set of charged particles is equal to the sum of the fields associated with the individual particles.

Stated mathematically:

$$\mathbf{E}(\mathbf{r}) = \sum_{n=1}^{N}\mathbf{E}(\mathbf{r};\mathbf{r}_n)$$

where N is the number of particles. Thus, we have

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon}\sum_{n=1}^{N}\frac{\mathbf{r}-\mathbf{r}_n}{|\mathbf{r}-\mathbf{r}_n|^3}\ q_n$$

This page titled 5.2: Electric Field Due to Point Charges is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.2: Electric Field Due to Point Charges** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.3: Charge Distributions

In principle, the smallest unit of electric charge that can be isolated is the charge of a single electron, which is $\cong -1.60 \times 10^{-19}$ C. This is very small, and we rarely deal with electrons one at a time, so it is usually more convenient to describe charge as a quantity that is continuous over some region of space. In particular, it is convenient to describe charge as being distributed in one of three ways: along a curve, over a surface, or within a volume.

Line Charge Distribution

Imagine that charge is distributed along a curve $\mathcal{C}$ through space. Let Δq be the total charge along a short segment of the curve, and let Δl be the length of this segment. The *line charge density* ρ_l at any point along the curve is defined as

$$\rho_l \triangleq \lim_{\Delta l \to 0} \frac{\Delta q}{\Delta l} = \frac{dq}{dl}$$

which has units of C/m. We may then define ρ_l to be a function of position along the curve, parameterized by l; e.g., $\rho_l(l)$. Then, the total charge Q along the curve is

$$Q = \int_{\mathcal{C}} \rho_l(l)\ dl$$

which has units of C. In other words, line charge density integrated over length yields total charge.

Surface Charge Distribution

Imagine that charge is distributed over a surface. Let Δq be the total charge on a small patch on this surface, and let Δs be the area of this patch. The *surface charge density* ρ_s at any point on the surface is defined as

$$\rho_s \triangleq \lim_{\Delta s \to 0} \frac{\Delta q}{\Delta s} = \frac{dq}{ds}$$

which has units of C/m^2. Let us define ρ_s to be a function of position on this surface. Then the total charge over a surface $\mathcal{S}$ is

$$Q = \int_{\mathcal{S}} \rho_s\ ds$$

In other words, surface charge density integrated over a surface yields total charge.

Volume Charge Distribution

Imagine that charge is distributed over a volume. Let Δq be the total charge in a small cell within this volume, and let Δv be the volume of this cell. The *volume charge density* ρ_v at any point in the volume is defined as

$$\rho_v \triangleq \lim_{\Delta v \to 0} \frac{\Delta q}{\Delta v} = \frac{dq}{dv}$$

which has units of C/m^3. Since ρ_v is a function of position within this volume, the total charge within a volume $\mathcal{V}$ is

$$Q = \int_{\mathcal{V}} \rho_v\ dv$$

In other words, volume charge density integrated over a volume yields total charge.

This page titled 5.3: Charge Distributions is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.3: Charge Distributions** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.4: Electric Field Due to a Continuous Distribution of Charge

The electric field intensity associated with N charged particles is (Section 5.2):

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon}\sum_{n=1}^{N}\frac{\mathbf{r}-\mathbf{r}_n}{|\mathbf{r}-\mathbf{r}_n|^3}\,q_n \tag{5.4.1}$$

where q_n and $\mathbf{r}_n$ are the charge and position of the n^{th} particle. However, it is common to have a continuous distribution of charge as opposed to a countable number of charged particles. In this section, we extend Equation 5.4.1 using the concept of continuous distribution of charge (Section 5.3) so that we may address this more general class of problems.

Distribution of Charge Along a Curve

Consider a continuous distribution of charge along a curve $\mathcal{C}$. The curve can be divided into short segments of length Δl. Then, the charge associated with the n^{th} segment, located at $\mathbf{r}_n$, is

$$q_n = \rho_l(\mathbf{r}_n)\;\Delta l$$

where ρ_l is charge density (units of C/m) at $\mathbf{r}_n$. Substituting this expression into Equation 5.4.1, we obtain

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon}\sum_{n=1}^{N}\frac{\mathbf{r}-\mathbf{r}_n}{|\mathbf{r}-\mathbf{r}_n|^3}\,\rho_l(\mathbf{r}_n)\;\Delta l$$

Taking the limit as $\Delta l \to 0$ yields:

$$\boxed{\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon}\int_{\mathcal{C}}\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3}\,\rho_l(\mathbf{r}')\;dl} \tag{5.4.2}$$

here $\mathbf{r}'$ represents the varying position along $\mathcal{C}$ with integration.

The simplest example of a curve is a straight line. It is straightforward to use Equation 5.4.2 to determine the electric field due to a distribution of charge along a straight line. However, it is much easier to analyze that particular distribution using Gauss' Law, as shown in Section 5.6. The following example addresses a charge distribution for which Equation 5.4.2 is more appropriate.

Example 5.4.1: Electric field along the axis of a ring of uniformly-distributed charge.

Consider a ring of radius a in the $z = 0$ plane, centered on the origin, as shown in Figure 5.4.1. Let the charge density along this ring be uniform and equal to ρ_l (C/m). Find the electric field along the z axis.

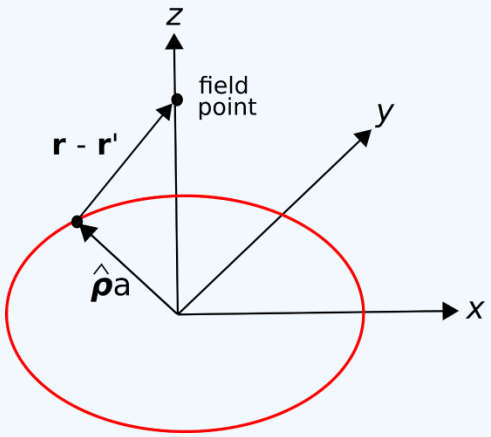


Figure 5.4.1: Calculating the electric field along the axis of a ring of charge. (CC BY-SA 4.0 K. Kikkeri).

Solution

The source charge position is given in cylindrical coordinates as

$$\mathbf{r}' = \hat{\rho}a$$

The position of a field point along the z axis is simply

$$\mathbf{r} = \hat{\mathbf{z}} z$$

Thus,

$$\mathbf{r} - \mathbf{r}' = -\hat{\rho} a + \hat{\mathbf{z}} z$$

and

$$|\mathbf{r} - \mathbf{r}'| = \sqrt{a^2 + z^2}$$

Equation 5.4.2 becomes:

$$\mathbf{E}(z) = \frac{1}{4\pi\epsilon} \int_0^{2\pi} \frac{-\hat{\rho} a + \hat{\mathbf{z}} z}{\left[a^2 + z^2\right]^{3/2}} \rho_l \ (a \ d\phi)$$

Pulling factors that do not vary with ϕ out of the integral and factoring into separate integrals for the $\hat{\phi}$ and $\hat{\mathbf{z}}$ components, we obtain:

$$\frac{\rho_l \ a}{4\pi\epsilon \left[a^2 + z^2\right]^{3/2}} \left[-a \int_0^{2\pi} \hat{\rho} \ d\phi + \hat{\mathbf{z}} z \int_0^{2\pi} d\phi \right]$$

The second integral is equal to 2π. The first integral is equal to zero. To see this, note that the integral is simply summing values of $\hat{\rho}$ for all possible values of ϕ. Since $\hat{\rho}(\phi + \pi) = -\hat{\rho}(\phi)$, the integrand for any given value of ϕ is equal and opposite the integrand π radians later. (This is one example of a *symmetry* argument.) Thus, we obtain

$$\mathbf{E}(z) = \hat{\mathbf{z}} \frac{\rho_l \ a}{2\epsilon} \frac{z}{\left[a^2 + z^2\right]^{3/2}}$$

It is a good exercise to confirm that this result is dimensionally correct. It is also recommended to confirm that when $z \gg a$, the result is approximately the same as that expected from a particle having the same total charge as the ring.

Distribution of Charge Over a Surface

Consider a continuous distribution of charge over a surface $\mathcal{S}$. The surface can be divided into small patches having area Δs. Then, the charge associated with the n^{th} patch, located at $\mathbf{r}_n$, is

$$q_n = \rho_s(\mathbf{r}_n) \ \Delta s$$

where ρ_s is the surface charge density (units of C/m^2) at $\mathbf{r}_n$. Substituting this expression into Equation 5.4.1, we obtain

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon} \sum_{n=1}^{N} \frac{\mathbf{r} - \mathbf{r}_n}{|\mathbf{r} - \mathbf{r}_n|^3} \rho_s(\mathbf{r}_n) \ \Delta s$$

Taking the limit as $\Delta s \to 0$ yields:

$$\boxed{\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon} \int_{\mathcal{S}} \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} \rho_s(\mathbf{r}') \ ds} \tag{5.4.3}$$

where $\mathbf{r}'$ represents the varying position over $\mathcal{S}$ with integration.

✔ Example 5.4.2: Electric field along the axis of a disk of uniformly-distributed charge.

Consider a circular disk of radius a in the $z = 0$ plane, centered on the origin, as shown in Figure 5.4.2. Let the charge density over this disk be uniform and equal to ρ_s (C/m^2). Find the electric field along the z axis.

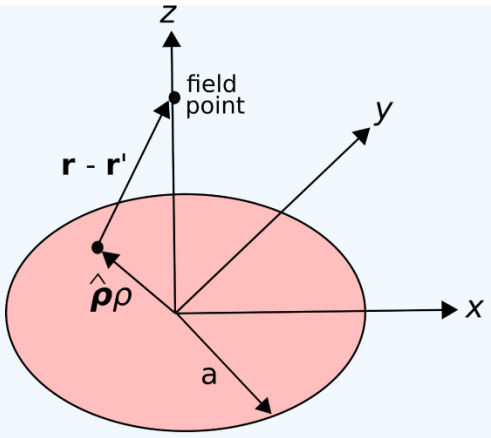


Figure 5.4.2: Calculating the electric field along the axis of a disk of charge. (CC BY-SA 4.0 K. Kikkeri).

Solution

The source charge position is given in cylindrical coordinates as

$$\mathbf{r}' = \hat{\rho}\rho$$

The position of a field point along the z axis is simply

$$\mathbf{r} = \hat{\mathbf{z}}z$$

Thus,

$$\mathbf{r} - \mathbf{r}' = -\hat{\rho}\rho + \hat{\mathbf{z}}z$$

and

$$|\mathbf{r} - \mathbf{r}'| = \sqrt{\rho^2 + z^2}$$

Equation 5.4.3 becomes:

$$\mathbf{E}(z) = \frac{1}{4\pi\epsilon}\int_{\rho=0}^{a}\int_{\phi=0}^{2\pi}\frac{-\hat{\rho}\rho + \hat{\mathbf{z}}z}{\left[\rho^2 + z^2\right]^{3/2}}\ \rho_s\ (\rho\ d\rho\ d\phi)$$

To solve this integral, first rearrange the double integral into a single integral over ϕ followed by integration over ρ:

$$\frac{\rho_s}{4\pi\epsilon}\int_{\rho=0}^{a}\frac{\rho}{\left[\rho^2 + z^2\right]^{3/2}}\left[\int_{\phi=0}^{2\pi}(-\hat{\rho}\rho + \hat{\mathbf{z}}z)\,d\phi\right]d\rho \tag{5.4.4}$$

Now we address the integration over ϕ shown in the square brackets in the above expression:

$$\int_{\phi=0}^{2\pi}(-\hat{\rho}\rho + \hat{\mathbf{z}}z)\,d\phi = -\rho\int_{\phi=0}^{2\pi}\hat{\rho}d\phi + \hat{\mathbf{z}}z\int_{\phi=0}^{2\pi}d\phi$$

The first integral on the right is zero for the following reason. As the integral progresses in ϕ, the vector $\hat{\rho}$ rotates. Because the integration is over a complete revolution (i.e., ϕ from 0 to 2π), the contribution from each pointing of $\hat{\rho}$ is canceled out by another pointing of $\hat{\rho}$ that is in the opposite direction. Since there is an equal number of these canceling pairs of pointings, the result is zero. Thus:

$$\int_{\phi=0}^{2\pi}(-\hat{\rho}\rho + \hat{\mathbf{z}}z)d\phi = 0 + \hat{\mathbf{z}}z\int_{\phi=0}^{2\pi}d\phi = \hat{\mathbf{z}}2\pi z$$

Substituting this into Expression 5.4.4 we obtain:

$$\frac{\rho_s}{4\pi\epsilon}\int_{\rho=0}^{a}\frac{\rho}{\left[\rho^2 + z^2\right]^{3/2}}[\hat{\mathbf{z}}2\pi z]d\rho$$
$$= \hat{\mathbf{z}}\frac{\rho_s z}{2\epsilon}\int_{\rho=0}^{a}\frac{\rho d\rho}{\left[\rho^2 + z^2\right]^{3/2}}$$

This integral can be solved using integration by parts and trigonometric substitution. Since the solution is tedious and there is no particular principle of electromagnetics demonstrated by this solution, we shall simply state the result:

$$\begin{aligned}\int_{\rho=0}^{a} \frac{\rho d\rho}{\left[\rho^2+z^2\right]^{3/2}} &= \left.\frac{-1}{\sqrt{\rho^2+z^2}}\right|_{\rho=0}^{a} \\ &= \frac{-1}{\sqrt{a^2+z^2}} + \frac{1}{|z|}\end{aligned}$$

Substituting this result:

$$\begin{aligned}\mathbf{E}(z) &= \hat{\mathbf{z}}\frac{\rho_s z}{2\epsilon}\left(\frac{-1}{\sqrt{a^2+z^2}} + \frac{1}{|z|}\right) \\ &= \hat{\mathbf{z}}\frac{\rho_s}{2\epsilon}\left(\frac{-z}{\sqrt{a^2+z^2}} + \frac{z}{|z|}\right) \\ &= \hat{\mathbf{z}}\frac{\rho_s}{2\epsilon}\left(\frac{-z}{\sqrt{a^2+z^2}} + \operatorname{sgn} z\right)\end{aligned}$$

where "sgn" is the "signum" function; i.e., $\operatorname{sgn}\ z = +1$ for $z > 0$ and $\operatorname{sgn}\ z = -1$ for $z < 0$.

Summarizing

$$\mathbf{E}(z) = \hat{\mathbf{z}}\frac{\rho_s}{2\epsilon}\left(\operatorname{sgn}\ z - \frac{z}{\sqrt{a^2+z^2}}\right) \tag{5.4.5}$$

It is a good exercise to confirm that this result is dimensionally correct and yields an electric field vector that points in the expected direction and with the expected dependence on a and z.

A special case of the "disk of charge" scenario considered in the preceding example is an *infinite sheet* of charge. The electric field from an infinite sheet of charge is a useful theoretical result. We get the field in this case simply by letting $a \to \infty$ in Equation 5.4.5, yielding:

$$\mathbf{E}(\mathbf{r}) = \hat{\mathbf{z}}\frac{\rho_s}{2\epsilon}\operatorname{sgn}\ z \tag{5.4.6}$$

Again, it is useful to confirm that this is dimensionally correct: C/m^2 divided by F/m yields V/m. Also, note that Equation 5.4.6 is the electric field at *any* point above or below the charge sheet – not just on z axis. This follows from symmetry. From the perspective of any point in space, the edges of the sheet are the same distance (i.e., infinitely far) away.

Distribution of Charge in a Volume

Consider a continuous distribution of charge within a volume $\mathcal{V}$. The volume can be divided into small cells (volume elements) having volume Δv. Then, the charge associated with the n^{th} cell, located at $\mathbf{r}_n$, is

$$q_n = \rho_v(\mathbf{r}_n)\ \Delta v$$

where ρ_v is volume charge density (units of C/m^3) at $\mathbf{r}_n$. Substituting this expression into Equation 5.4.1, we obtain

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon}\sum_{n=1}^{N}\frac{\mathbf{r}-\mathbf{r}_n}{|\mathbf{r}-\mathbf{r}_n|^3}\ \rho_v(\mathbf{r}_n)\ \Delta v$$

Taking the limit as $\Delta v \to 0$ yields:

$$\boxed{\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon}\int_{\mathcal{V}}\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3}\ \rho_v(\mathbf{r}')\ dv}$$

where $\mathbf{r}'$ represents the varying position over $\mathcal{V}$ with integration.

This page titled 5.4: Electric Field Due to a Continuous Distribution of Charge is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.4: Electric Field Due to a Continuous Distribution of Charge** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.5: Gauss' Law - Integral Form

Gauss' Law is one of the four fundamental laws of classical electromagnetics, collectively known as *Maxwell's Equations*. Before diving in, the reader is strongly encouraged to review Section 2.4. In that section, Gauss' Law emerges from the interpretation of the electric field as a flux density. Section 2.4 does not actually identify Gauss' Law, but here it is:

> *Gauss' Law* (Equation 5.5.1) states that the flux of the electric field through a closed surface is equal to the enclosed charge.

Gauss' Law is expressed mathematically as follows:

$$\boxed{\oint_{\mathcal{S}} \mathbf{D}\cdot d\mathbf{s} = Q_{encl}} \tag{5.5.1}$$

where $\mathbf{D}$ is the electric flux density $\epsilon\mathbf{E}$, $\mathcal{S}$ is a closed surface with differential surface normal $d\mathbf{s}$, and Q_{encl} is the enclosed charge. We can see the law is dimensionally correct; $\mathbf{D}$ has units of C/m^2, thus integrating $\mathbf{D}$ over a surface gives a quantity with units of C/m^2 $\cdot$ m^2 = C, which are the units of charge.

Gauss' Law has a number of applications in electromagnetic theory. One of them, as explored below, is as a method to compute the electric field in response to a distribution of electric charge. Note that a method to do this, based on Coulomb's Law, is described in Sections 5.1, 5.2, and 5.4. Gauss' Law provides an alternative method that is easier or more useful in certain applications.

✔ Example 5.5.1: Electric field associated with a charged particle, using Gauss' Law.

In this example, we demonstrate the ability of Gauss' Law to predict the field associated with a charge distribution. Let us do this for the simplest possible charge distribution. A particle of charge q located at the origin, for which we already have the answer (Section 5.1).

Solution

Gauss' Law applies to *any* surface that encloses the charge, so for simplicity we chose a sphere of radius r centered at the origin. Note that Q_{encl} on the right hand side is just q for any surface having $r > 0$. Gauss' Law in this case becomes

$$\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} \mathbf{D}\cdot\left(\hat{\mathbf{r}}r^2\sin\theta\ d\theta\ d\phi\right) = q$$

If we can solve for $\mathbf{D}$, we can get $\mathbf{E}$ using $\mathbf{D} = \epsilon\mathbf{E}$. The simplest way to solve for $\mathbf{D}$ is to use a symmetry argument, which proceeds as follows. In this problem, the *magnitude* of $\mathbf{D}$ can depend only on r, and not θ or ϕ. This is because the charge has no particular orientation, and the sphere is centered on the charge. Similarly, it is clear that $\mathbf{D}$ must point either directly toward or directly away from the charge. In other words, $\mathbf{D} = \hat{\mathbf{r}}D(r)$. Substituting this in the above equation, we encounter the dot product $\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}$, which is simply 1. Since $D(r)$ and r^2 are constants with respect to the integration, we obtain:

$$r^2D(r)\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi}\sin\theta\ d\theta d\phi = q$$

The remaining integral is simply 4π, thus we obtain:

$$D(r) = \frac{q}{4\pi r^2}$$

Bringing the known vector orientation of $\mathbf{D}$ back into the equation, we obtain

$$\mathbf{D} = \hat{\mathbf{r}}\frac{q}{4\pi r^2}$$

and finally using $\mathbf{D} = \epsilon\mathbf{E}$ we obtain the expected result

$$\mathbf{E} = \hat{\mathbf{r}}\frac{q}{4\pi\epsilon r^2}$$

Here's the point you should take away from the above example:

> Gauss' Law combined with a symmetry argument may be sufficient to determine the electric field due to a charge distribution. Thus, Gauss' Law may be an easier alternative to Coulomb's Law in some applications.

This page titled 5.5: Gauss' Law - Integral Form is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.5: Gauss' Law - Integral Form** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.6: Electric Field Due to an Infinite Line Charge using Gauss' Law

Section 5.5 explains one application of Gauss' Law, which is to find the electric field due to a charged particle. In this section, we present another application – the electric field due to an infinite line of charge. The result serves as a useful "building block" in a number of other problems, including determination of the capacitance of coaxial cable (Section 5.24). Although this problem can be solved using the "direct" approach described in Section 5.4 (and it is an excellent exercise to do so), the Gauss' Law approach demonstrated here turns out to be relatively simple.

Example 5.6.1: Electric field associated with an infinite line charge, using Gauss' Law.

Use Gauss' Law to determine the electric field intensity due to an infinite line of charge along the z axis, having charge density ρ_l (units of C/m), as shown in Figure 5.6.1.

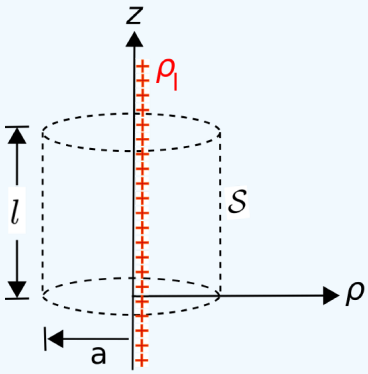


Figure 5.6.1: Finding the electric field of an infinite line of charge using Gauss' Law. (CC BY-SA 4.0; K. Kikkeri).

Solution

Gauss' Law requires integration over a surface that encloses the charge. So, our first problem is to determine a suitable surface. A cylinder of radius a that is concentric with the z axis, as shown in Figure 5.6.1, is maximally symmetric with the charge distribution and so is likely to yield the simplest possible analysis. At first glance, it seems that we may have a problem since the charge extends to infinity in the $+z$ and $-z$ directions, so it's not clear how to enclose all of the charge. Let's suppress that concern for a moment and simply choose a cylinder of finite length l. In principle, we can solve the problem first for this cylinder of finite size, which contains only a fraction of the charge, and then later let $l \to \infty$ to capture the rest of the charge. (In fact, we'll find when the time comes it will not be necessary to do that, but we shall prepare for it anyway.)

Here's Gauss' Law:

$$\oint_{\mathcal{S}} \mathbf{D} \cdot d\mathbf{s} = Q_{encl} \tag{5.6.1}$$

where $\mathbf{D}$ is the electric flux density $\epsilon\mathbf{E}$, $\mathcal{S}$ is a closed surface with outward-facing differential surface normal $d\mathbf{s}$, and Q_{encl} is the enclosed charge.

The first order of business is to constrain the form of $\mathbf{D}$ using a symmetry argument, as follows. Consider the field of a point charge q at the origin (Section 5.5):

$$\mathbf{D} = \hat{\mathbf{r}} \frac{q}{4\pi r^2}$$

We can "assemble" an infinite line of charge by adding particles in pairs. One pair is added at a time, with one particle on the $+z$ axis and the other on the $-z$ axis, with each located an equal distance from the origin. We continue to add particle pairs in this manner until the resulting charge extends continuously to infinity in both directions. The principle of superposition indicates that the resulting field will be the sum of the fields of the particles (Section 5.2). Thus, we see that $\mathbf{D}$ cannot have any component in the $\hat{\phi}$ direction because none of the fields of the constituent particles have a component in that direction. Similarly, we see that the magnitude of $\mathbf{D}$ cannot depend on ϕ because none of the fields of the constituent particles depends on ϕ and because the charge distribution is identical ("invariant") with rotation in ϕ. Also, note that for any choice of z the distribution of charge above and below that plane of constant z is identical; therefore, $\mathbf{D}$ cannot be a function of z *and* $\mathbf{D}$ cannot

have any component in the $\hat{\mathbf{z}}$ direction. Therefore, the direction of $\mathbf{D}$ must be radially outward; i.e., in the $\hat{\rho}$ direction, as follows:

$$\mathbf{D} = \hat{\rho} D_\rho(\rho)$$

Next, we observe that Q_{encl} on the right hand side of Equation 5.6.1 is equal to $\rho_l l$. Thus, we obtain

$$\oint_{\mathcal{S}} [\hat{\rho} D_\rho(\rho)] \cdot d\mathbf{s} = \rho_l l$$

The cylinder $\mathcal{S}$ consists of a flat top, curved side, and flat bottom. Expanding the above equation to reflect this, we obtain

$$\begin{aligned} \rho_l l = &\int_{top} [\hat{\rho} D_\rho(\rho)] \cdot (+\hat{\mathbf{z}} ds) \\ &+ \int_{side} [\hat{\rho} D_\rho(\rho)] \cdot (+\hat{\rho} ds) \\ &+ \int_{bottom} [\hat{\rho} D_\rho(\rho)] \cdot (-\hat{\mathbf{z}} ds) \end{aligned} \tag{5.6.2}$$

Examination of the dot products indicates that the integrals associated with the top and bottom surfaces must be zero. In other words, the flux through the top and bottom is zero because $\mathbf{D}$ is perpendicular to these surfaces. We are left with

$$\rho_l l = \int_{side} [D_\rho(\rho)]\, ds$$

The side surface is an open cylinder of radius $\rho = a$, so $D_\rho(\rho) = D_\rho(a)$, a constant over this surface. Thus:

$$\rho_l l = \int_{side} [D_\rho(a)]\, ds = [D_\rho(a)] \int_{side} ds$$

The remaining integral is simply the area of the side surface, which is $2\pi a \cdot l$. Solving for $D_\rho(a)$ we obtain

$$D_\rho(a) = \frac{\rho_l l}{2\pi a l} = \frac{\rho_l}{2\pi a}$$

Remarkably, we see $D_\rho(a)$ is independent of l, So the concern raised in the beginning of this solution – that we wouldn't be able to enclose all of the charge – doesn't matter.

Completing the solution, we note the result must be the same for any value of ρ (not just $\rho = a$), so

$$\mathbf{D} = \hat{\rho} D_\rho(\rho) = \hat{\rho} \frac{\rho_l}{2\pi\rho}$$

and since $\mathbf{D} = \epsilon \mathbf{E}$:

$$\boxed{\mathbf{E} = \hat{\rho} \frac{\rho_l}{2\pi\epsilon\rho}}$$

This completes the solution. We have found that the electric field is directed radially away from the line charge, and decreases in magnitude in inverse proportion to distance from the line charge.

Suggestion: Check to ensure that this solution is dimensionally correct.

This page titled 5.6: Electric Field Due to an Infinite Line Charge using Gauss' Law is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.6: Electric Field Due to an Infinite Line Charge using Gauss' Law** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.7: Gauss' Law - Differential Form

The integral form of Gauss' Law is a calculation of enclosed charge Q_{encl} using the surrounding density of electric flux:

$$\oint_{\mathcal{S}} \mathbf{D}\cdot d\mathbf{s} = Q_{encl} \tag{5.7.1}$$

where $\mathbf{D}$ is electric flux density and $\mathcal{S}$ is the enclosing surface. It is also sometimes necessary to do the inverse calculation (i.e., determine electric field associated with a charge distribution). This is sometimes possible using Equation 5.7.1 if the symmetry of the problem permits; see examples in Section 5.5 and 5.6. If the problem does not exhibit the necessary symmetry, then it seems that one must fall back to the family of techniques presented in Section 5.4 requiring direct integration over the charge, which is derived from Coulomb's Law.

However, even the Coulomb's Law / direct integration approach has a limitation that is very important to recognize: It does not account for the presence of structures that may influence the electric field. For example, the electric field due to a charge in free space is different from the electric field due to the same charge located near a perfectly-conducting surface. In fact, these approaches do not account for the possibility of *any* spatial variation in material composition, which rules out their use in many engineering applications.

To address this broader scope of problems, we require an alternative form of Gauss' Law that applies at individual points in space. That is, we require Gauss' Law expressed in the form of a differential equation, as opposed to an integral equation. This facilitates the use of Gauss' Law even in problems that do not exhibit sufficient symmetry and that involve material boundaries and spatial variations in material constitutive parameters. Given this differential equation and the boundary conditions imposed by structure and materials, we may then solve for the electric field in these more complicated scenarios. In this section, we derive the desired differential form of Gauss' Law. Elsewhere (in particular, in Section 5.15) we use this equation as a tool to find electric fields in problems involving material boundaries.

There are in fact two methods to develop the desired differential equation. One method is via the definition of divergence, whereas the other is via the divergence theorem. Both methods are presented below because each provides a different bit of insight. Let's explore the first method:

Derivation via the Definition of Divergence

Let the geometrical volume enclosed by $\mathcal{S}$ be $\mathcal{V}$, which has volume V (units of m^3). Dividing both sides of Equation 5.7.1 by V and taking the limit as $V \to 0$:

$$\lim_{V\to 0} \frac{\oint_{\mathcal{S}} \mathbf{D}\cdot d\mathbf{s}}{V} = \lim_{V\to 0} \frac{Q_{encl}}{V}$$

The quantity on the right hand side is the volume charge density ρ_v (units of C/m^3) at the point at which we converge after letting the volume go to zero. The left hand side is, by definition, the *divergence* of $\mathbf{D}$, indicated in mathematical notation as "$\nabla\cdot\mathbf{D}$" (Section 4.6). Thus, we have *Gauss' Law in differential form*:

$$\boxed{\nabla\cdot\mathbf{D} = \rho_v} \tag{5.7.2}$$

To interpret this equation, recall that divergence is simply the flux (in this case, *electric* flux) per unit volume.

> Gauss' Law in differential form (Equation 5.7.2) says that the electric flux per unit volume originating from a point in space is equal to the volume charge density at that point.

Derivation via the Divergence Theorem

Equation 5.7.2 may also be obtained from Equation 5.7.1 using the Divergence Theorem, which in the present case may be written:

$$\int_{\mathcal{V}} (\nabla\cdot\mathbf{D})\, dv = \oint_{\mathcal{S}} \mathbf{D}\cdot d\mathbf{s}$$

From Equation 5.7.1, we see that the right hand side of the equation may be replaced with the enclosed charge:

$$\int (\nabla\cdot\mathbf{D})\, dv = Q_{encl}$$

$J_{\mathcal{V}}$

Furthermore, the enclosed charge can be expressed as an integration of the volume charge density ρ_v over $\mathcal{V}$:

$$\int_{\mathcal{V}} (\nabla \cdot \mathbf{D})\, dv = \int_{\mathcal{V}} \rho_v dv$$

The above relationship must hold regardless of the specific location or shape of $\mathcal{V}$. The only way this is possible is if the integrands are equal. Thus, $\nabla \cdot \mathbf{D} = \rho_v$, and we have obtained Equation 5.7.2.

Example 5.7.1: Determining the charge density at a point, given the associated electric field

The electric field intensity in free space is

$$\mathbf{E}(\mathbf{r}) = \hat{\mathbf{x}} A x^2 + \hat{\mathbf{y}} B z + \hat{\mathbf{z}} C x^2 z$$

where $A = 3\ \text{V/m}^3$, $B = 2\ \text{V/m}^2$, and $C = 1\ \text{V/m}^4$. What is the charge density at $\mathbf{r} = \hat{\mathbf{x}}2 - \hat{\mathbf{y}}2$ m?

Solution

First, we use $\mathbf{D} = \epsilon \mathbf{E}$ to get $\mathbf{D}$. Since the problem is in free space, $\epsilon = \epsilon_0$. Thus we have that the volume charge density is

$$\begin{aligned}
\rho_v &= \nabla \cdot \mathbf{D} \\
&= \nabla \cdot (\epsilon_0 \mathbf{E}) = \epsilon_0 \nabla \cdot \mathbf{E} \\
&= \epsilon_0 \left[\frac{\partial}{\partial x} \left(A x^2 \right) + \frac{\partial}{\partial y} (B z) + \frac{\partial}{\partial z} \left(C x^2 z \right) \right] \\
&= \epsilon_0 \left[2Ax + 0 + C x^2 \right]
\end{aligned}$$

Now calculating the charge density at the specified location $\mathbf{r}$:

$$\begin{aligned}
&\epsilon_0 \left[2 \left(3\ \text{V/m}^3 \right) (2\ \text{m}) + 0 + \left(1\ \text{V/m}^4 \right) (2\ \text{m})^2 \right] \\
&= \epsilon_0 (16\ \text{V/m}) \\
&= 142\ \text{pC/m}^3
\end{aligned}$$

To obtain the electric field from the charge distribution in the presence of boundary conditions imposed by materials and structure, we must enforce the relevant boundary conditions. These boundary conditions are presented in Sections 5.17 and 5.18. Frequently, a simpler approach requiring only the boundary conditions on the electric potential ($V(\mathbf{r})$) is possible; this is presented in Section 5.15.

Furthermore, the reader should note the following. Gauss' Law does not always necessarily *fully* constrain possible solutions for the electric field. For that, we might also need Kirchoff's Voltage Law; see Section 5.11.

Before moving on, it is worth noting that Equation 5.7.2 can be solved in the special case in which there are no boundary conditions to satisfy; i.e., for charge only, in a uniform and unbounded medium. In fact, no additional electromagnetic theory is required to do this. Here's the solution:

$$\mathbf{D}(\mathbf{r}) = \frac{1}{4\pi} \int_{\mathcal{V}} \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3}\, \rho_v(\mathbf{r}')\, dv$$

which we recognize as one of the results obtained in Sections 5.4 (after dividing both sides by ϵ to get $\mathbf{E}$). It is reasonable to conclude that Gauss' Law (in either integral or differential form) is fundamental, whereas Coulomb's Law is merely a consequence of Gauss' Law.

This page titled 5.7: Gauss' Law - Differential Form is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.7: Gauss' Law - Differential Form** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.8: Force, Energy, and Potential Difference

The force $\mathbf{F}_e$ experienced by a particle at location $\mathbf{r}$ bearing charge q in an electric field intensity $\mathbf{E}$ is

$$\mathbf{F}_e = q\mathbf{E}(\mathbf{r}) \tag{5.8.1}$$

If left alone in free space, this particle would immediately begin to move. The resulting displacement represents a loss of potential energy. This loss can quantified using the concept of *work*, W. The incremental work ΔW done by moving the particle a short distance Δl, over which we assume the change in $\mathbf{F}_e$ is negligible, is

$$\Delta W \approx -\mathbf{F}_e \cdot \hat{\mathbf{l}} \Delta l \tag{5.8.2}$$

where in this case $\hat{\mathbf{l}}$ is the unit vector in the direction of the motion; i.e., the direction of $\mathbf{F}_e$. The minus sign indicates that potential energy of the system consisting of the electric field and the particle is being reduced. Like a spring that was previously compressed and is now released, the system is "relaxing."

To confirm that work defined in this way is an expression of energy, consider the units. The product of force (units of N) and distance (units of m) has units of N·m, and 1 N·m is 1 J of energy.

Now, what if the motion of the particle is due to factors other than the force associated with the electric field? For example, we might consider "resetting" the system to it's original condition by applying an external force to overcome $\mathbf{F}_e$. Equation 5.8.2 still represents the change in potential energy of the system, but now $\hat{\mathbf{l}}$ changes sign. The same magnitude of work is done, but now this work is positive. In other words, positive work requires the application of an *external* force that opposes and overcomes the force associated with the electric field, thereby increasing the potential energy of the system. With respect to the analogy of a mechanical spring used above, positive work is achieved by compressing the spring.

It is also worth noting that the purpose of the dot product in Equation 5.8.2 is to ensure that only the component of motion parallel to the direction of the electric field is included in the energy tally. This is simply because motion in any other direction cannot be due to $\mathbf{E}$, and therefore does not increase or decrease the associated potential energy.

We can make the relationship between work and the electric field explicit by substituting Equation 5.8.1 into Equation 5.8.2, yielding

$$\Delta W \approx -q\mathbf{E}(\mathbf{r}) \cdot \hat{\mathbf{l}} \Delta l \tag{5.8.3}$$

Equation 5.8.3 gives the work only for a short distance around $\mathbf{r}$. Now let us try to generalize this result. If we wish to know the work done over a larger distance, then we must account for the possibility that $\mathbf{E}$ varies along the path taken. To do this, we may sum contributions from points along the path traced out by the particle, i.e.,

$$W \approx \sum_{n=1}^{N} \Delta W(\mathbf{r}_n)$$

where $\mathbf{r}_n$ are positions defining the path. Substituting Equation 5.8.3, we have

$$W \approx -q \sum_{n=1}^{N} \mathbf{E}(\mathbf{r}_n) \cdot \hat{\mathbf{l}}(\mathbf{r}_n) \Delta l$$

Taking the limit as $\Delta l \to 0$ we obtain

$$W = -q \int_{\mathcal{C}} \mathbf{E}(\mathbf{r}) \cdot \hat{\mathbf{l}}(\mathbf{r}) dl$$

where $\mathcal{C}$ is the path (previously, the sequence of $\mathbf{r}_n$'s) followed. Now omitting the explicit dependence on $\mathbf{r}$ in the integrand for clarity:

$$W = -q \int_{\mathcal{C}} \mathbf{E} \cdot d\mathbf{l} \tag{5.8.4}$$

where $d\mathbf{l} = \hat{\mathbf{l}} dl$ as usual. Now, we are able to determine the change in potential energy for a charged particle moving along any path in space, given the electric field.

At this point, it is convenient to formally define the electric *potential difference* V_{21} between the start point (1) and end point (2) of $\mathcal{C}$. V_{21} is defined as the work done by traversing $\mathcal{C}$, per unit of charge:

$$V_{21} \triangleq \frac{W}{q}$$

This has units of J/C, which is volts (V). Substituting Equation 5.8.4, we obtain:

$$\boxed{V_{21} = -\int_{\mathcal{C}} \mathbf{E} \cdot d\mathbf{l}} \tag{5.8.5}$$

An advantage of analysis in terms of electrical potential as opposed to energy is that we will no longer have to explicitly state the value of the charge involved.

The potential difference V_{21} between two points in space, given by Equation 5.8.5, is the change in potential energy of a charged particle divided by the charge of the particle. Potential energy is also commonly known as "voltage" and has units of V.

Example 5.8.1: Potential difference in a uniform electric field

Consider an electric field $\mathbf{E}(\mathbf{r}) = \hat{\mathbf{z}} E_0$, which is constant in both magnitude and direction throughout the domain of the problem. The path of interest is a line beginning at $\hat{\mathbf{z}} z_1$ and ending at $\hat{\mathbf{z}} z_2$. What is V_{21}? (It's worth noting that the answer to this problem is a building block for a vast number of problems in electromagnetic analysis.)

Solution

From Equation 5.8.5 we have

$$V_{21} = -\int_{z_1}^{z_2} (\hat{\mathbf{z}} E_0) \cdot \hat{\mathbf{z}} dz = -E_0 (z_2 - z_1)$$

Note V_{21} is simply the electric field intensity times the distance between the points. This may seem familiar. For example, compare this to the findings of the battery-charged capacitor experiment described in Section 2.2. There too we find that potential difference equals electric field intensity times distance, and the signs agree.

The solution to the preceding example is simple because the direct path between the two points is parallel to the electric field. If the path between the points had been *perpendicular* to $\mathbf{E}$, then the solution is even easier – V_{21} is simply zero. In all other cases, V_{21} is proportional to the component of the direct path between the start and end points that is parallel to $\mathbf{E}$, as determined by the dot product.

This page titled 5.8: Force, Energy, and Potential Difference is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.8: Force, Energy, and Potential Difference** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.9: Independence of Path

In Section 5.8, we found that the potential difference ("voltage") associated with a path $\mathcal{C}$ in an electric field intensity $\mathbf{E}$ is

$$V_{21} = -\int_{\mathcal{C}} \mathbf{E}\cdot d\mathbf{l}$$

where the curve begins at point 1 and ends at point 2. Let these points be identified using the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively (see Section 4.1). Then:

$$V_{21} = -\int_{\mathbf{r}_1,\text{ along }\mathcal{C}}^{\mathbf{r}_2} \mathbf{E}\cdot d\mathbf{l}$$

The associated work done by a particle bearing charge q is

$$W_{21} = qV_{21}$$

This work represents the change in potential energy of the system consisting of the electric field and the charged particle. So, it must also be true that

$$W_{21} = W_2 - W_1$$

where W_2 and W_1 are the potential energies when the particle is at $\mathbf{r}_2$ and $\mathbf{r}_1$, respectively. It is clear from the above equation that W_{21} does not depend on $\mathcal{C}$; it depends only on the positions of the start and end points and not on any of the intermediate points along $\mathcal{C}$. That is,

$$\boxed{V_{21} = -\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{E}\cdot d\mathbf{l} \quad \text{, independent of } \mathcal{C}} \tag{5.9.1}$$

Since the result of the integration in Equation 5.9.1 is independent of the path of integration, *any* path that begins at $\mathbf{r}_1$ and ends at $\mathbf{r}_2$ yields the same value of W_{21} and V_{21}. We refer to this concept as *independence of path*.

> The integral of the electric field over a path between two points depends only on the locations of the start and end points and is independent of the path taken between those points.

A practical application of this concept is that some paths may be easier to use than others, so there may be an advantage in computing the integral in Equation 5.9.1 using some path other than the path actually traversed.

This page titled 5.9: Independence of Path is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.9: Independence of Path** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.10: Kirchoff's Voltage Law for Electrostatics - Integral Form

As explained in Section 5.9, the electrical potential at point $\mathbf{r}_2$ relative to $\mathbf{r}_1$ in an electric field $\mathbf{E}$ (V/m) is

$$V_{21} = -\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{E}\cdot d\mathbf{l}$$

where the path of integration may be any path that begins and ends at the specified points. Consider what happens if the selected path through space begins and ends at the *same* point; i.e., $\mathbf{r}_2 = \mathbf{r}_1$. In this case, the path of integration is a closed loop. Since V_{21} depends only on the positions of the start and end points and because the potential energy at those points is the same, we conclude:

$$\boxed{\oint \mathbf{E} \cdot d\mathbf{l} = 0} \tag{5.10.1}$$

This principle is known as *Kirchoff's Voltage Law for Electrostatics*.

> Kirchoff's Voltage Law for Electrostatics (Equation 5.10.1) states that the integral of the electric field over a closed path is zero.

It is worth noting that this law is a generalization of a principle of which the reader is likely already aware. In electric circuit theory, the sum of voltages over any closed loop in a circuit is zero. This is also known as Kirchoff's Voltage Law because it is precisely the same principle. To obtain Equation 5.10.1 for an electric circuit, simply partition the closed path into branches, with each branch representing one component. Then, the integral of $\mathbf{E}$ over each branch is the branch voltage; i.e., units of V/m times units of m yields units of V. Then, the sum of these branch voltages over any closed loop is zero, as dictated by Equation 5.10.1.

Finally, be advised that Equation 5.10.1 is specific to electrostatics. In electrostatics, it is assumed that the electric field is independent of the magnetic field. This is true if the magnetic field is either zero or not time-varying. If the magnetic field is time-varying, then Equation 5.10.1 must be modified to account for the effect of the magnetic field, which is to make the right hand size potentially different from zero. The generalized version of this expression that correctly accounts for that effect is known as the *Maxwell-Faraday Equation* (Section 8.8).

This page titled 5.10: Kirchoff's Voltage Law for Electrostatics - Integral Form is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.10: Kirchoff's Voltage Law for Electrostatics - Integral Form** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.11: Kirchoff's Voltage Law for Electrostatics - Differential Form

The integral form of Kirchoff's Voltage Law for electrostatics (KVL; Section 5.10) states that an integral of the electric field along a closed path is equal to zero:

$$\oint_{\mathcal{C}} \mathbf{E} \cdot d\mathbf{l} = 0$$

where $\mathbf{E}$ is electric field intensity and $\mathcal{C}$ is the closed curve. In this section, we derive the differential form of this equation. In some applications, this differential equation, combined with boundary conditions imposed by structure and materials (Sections 5.17 and 5.18), can be used to solve for the electric field in arbitrarily complicated scenarios. A more immediate reason for considering this differential equation is that we gain a little more insight into the behavior of the electric field, disclosed at the end of this section.

The equation we seek may be obtained using Stokes' Theorem (Section 4.9), which in the present case may be written:

$$\int_{\mathcal{S}} (\nabla \times \mathbf{E}) \cdot d\mathbf{s} = \oint_{\mathcal{C}} \mathbf{E} \cdot d\mathbf{l} \tag{5.11.1}$$

where $\mathcal{S}$ is any surface bounded by $\mathcal{C}$, and $d\mathbf{s}$ is the normal to that surface with direction determined by right-hand rule. The integral form of KVL tells us that the right hand side of the above equation is zero, so:

$$\int_{\mathcal{S}} (\nabla \times \mathbf{E}) \cdot d\mathbf{s} = 0$$

The above relationship must hold regardless of the specific location or shape of $\mathcal{S}$. The only way this is possible for all possible surfaces is if the integrand is zero at every point in space. Thus, we obtain the desired expression:

$$\boxed{\nabla \times \mathbf{E} = 0} \tag{5.11.2}$$

Summarizing:

> The differential form of Kirchoff's Voltage Law for electrostatics (Equation 5.11.2) states that the curl of the electrostatic field is zero.

Equation 5.11.2 is a partial differential equation. As noted above, this equation, combined with the appropriate boundary conditions, can be solved for the electric field in arbitrarily-complicated scenarios. Interestingly, it is not the only such equation available for this purpose – Gauss' Law (Section 5.7) also does this. Thus, we see a *system* of partial differential equations emerging, and one may correctly infer that that the electric field is not necessarily fully constrained by either equation alone.

This page titled 5.11: Kirchoff's Voltage Law for Electrostatics - Differential Form is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.11: Kirchoff's Voltage Law for Electrostatics - Differential Form** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.12: Electric Potential Field Due to Point Charges

The electric field intensity due to a point charge q at the origin is (see Section 5.1 or 5.5)

$$\mathbf{E} = \hat{\mathbf{r}}\frac{q}{4\pi\epsilon r^2} \tag{5.12.1}$$

In Sections 5.8 and 5.9, it was determined that the potential difference measured from position $\mathbf{r}_1$ to position $\mathbf{r}_2$ is

$$V_{21} = -\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{E}\cdot d\mathbf{l} \tag{5.12.2}$$

This method for calculating potential difference is often a bit awkward. To see why, consider an example from circuit theory, shown in Figure 5.12.1. In this example, consisting of a single resistor and a ground node, we've identified four quantities:

- The resistance R
- The current I through the resistor
- The node voltage V_1, which is the potential difference measured from ground to the left side of the resistor
- The node voltage V_2, which is the potential difference measured from ground to the right side of the resistor

Let's say we wish to calculate the potential difference V_{21} across the resistor. There are two ways this can be done:

- $V_{21} = -IR$
- $V_{21} = V_2 - V_1$

The advantage of the second method is that it is not necessary to know I, R, or indeed anything about what is happening between the nodes; it is only necessary to know the node voltages. The point is that it is often convenient to have a common *datum* – in this example, ground – with respect to which the potential differences at all other locations of interest can be defined. When we have this, calculating potential differences reduced to simply subtracting predetermined node potentials.

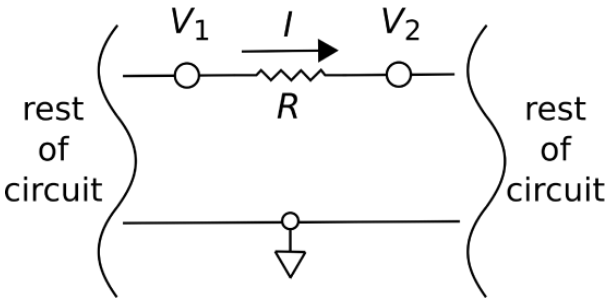


Figure 5.12.1: A resistor in a larger circuit, used as an example to demonstrate the concept of node voltages. (CC BY SA 4.0; K. Kikkeri).

So, can we establish a datum in general electrostatic problems that works the same way? The answer is yes. The datum is arbitrarily chosen to be a sphere that encompasses the universe; i.e., a sphere with radius $\to \infty$. Employing this choice of datum, we can use Equation 5.12.2 to define $V(\mathbf{r})$, the potential at point $\mathbf{r}$, as follows:

$$\boxed{V(\mathbf{r}) \triangleq -\int_{\infty}^{\mathbf{r}} \mathbf{E}\cdot d\mathbf{l}} \tag{5.12.3}$$

> The electrical potential at a point, given by Equation 5.12.3, is defined as the potential difference measured beginning at a sphere of infinite radius and ending at the point $\mathbf{r}$. The potential obtained in this manner is with respect to the potential infinitely far away.

In the particular case where $\mathbf{E}$ is due to the point charge at the origin:

$$V(\mathbf{r}) = -\int_{\infty}^{\mathbf{r}} \left[\hat{\mathbf{r}}\frac{q}{4\pi\epsilon r^2}\right]\cdot d\mathbf{l}$$

The principle of *independence of path* (Section 5.9) asserts that the path of integration doesn't matter as long as the path begins at the datum at infinity and ends at $\mathbf{r}$. So, we should choose the easiest such path. The radial symmetry of the problem indicates that the easiest path will be a line of constant θ and ϕ, so we choose $d\mathbf{l} = \hat{\mathbf{r}}dr$. Continuing:

$$\begin{aligned} V(\mathbf{r}) &= -\int_{\infty}^{r}\left[\hat{\mathbf{r}}\frac{q}{4\pi\epsilon r^2}\right]\cdot[\hat{\mathbf{r}}dr] \\ &= -\frac{q}{4\pi\epsilon}\int_{\infty}^{r}\frac{1}{r^2}dr \\ &= +\frac{q}{4\pi\epsilon}\frac{1}{r}\Big|_{\infty}^{r} \end{aligned} \tag{5.12.4}$$

so

$$\boxed{V(\mathbf{r}) = +\frac{q}{4\pi\epsilon r}} \tag{5.12.5}$$

(Suggestion: Confirm that Equation 5.12.5 is dimensionally correct.) In the context of the circuit theory example above, this is the "node voltage" at $\mathbf{r}$ when the datum is defined to be the surface of a sphere at infinity. Subsequently, we may calculate the potential difference from any point $\mathbf{r}_1$ to any other point $\mathbf{r}_2$ as

$$V_{21} = V(\mathbf{r}_2) - V(\mathbf{r}_1)$$

and that will typically be a *lot* easier than using Equation 5.12.2.

It is not often that one deals with systems consisting of a single charged particle. So, for the above technique to be truly useful, we need a straightforward way to determine the potential field $V(\mathbf{r})$ for arbitrary distributions of charge. The first step in developing a more general expression is to determine the result for a particle located at a point $\mathbf{r}'$ somewhere other than the origin. Since Equation 5.12.5 depends only on charge and the distance between the field point $\mathbf{r}$ and $\mathbf{r}'$, we have

$$V(\mathbf{r};\mathbf{r}') \triangleq +\frac{q'}{4\pi\epsilon\,|\mathbf{r}-\mathbf{r}'|} \tag{5.12.6}$$

where, for notational consistency, we use the symbol q' to indicate the charge. Now applying superposition, the potential field due to N charges is

$$V(\mathbf{r}) = \sum_{n=1}^{N} V(\mathbf{r};\mathbf{r}_n)$$

Substituting Equation 5.12.6 we obtain:

$$\boxed{V(\mathbf{r}) = \frac{1}{4\pi\epsilon}\sum_{n=1}^{N}\frac{q_n}{|\mathbf{r}-\mathbf{r}_n|}} \tag{5.12.7}$$

> Equation 5.12.7 gives the electric potential at a specified location due to a finite number of charged particles.

The potential field due to continuous distributions of charge is addressed in Section 5.13.

This page titled 5.12: Electric Potential Field Due to Point Charges is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.12: Electric Potential Field Due to Point Charges** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.13: Electric Potential Field due to a Continuous Distribution of Charge

The electrostatic potential field at $\mathbf{r}$ associated with N charged particles is

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon}\sum_{n=1}^{N}\frac{q_n}{|\mathbf{r}-\mathbf{r}_n|} \tag{5.13.1}$$

where q_n and $\mathbf{r_n}$ are the charge and position of the n^{th} particle. However, it is more common to have a continuous distribution of charge as opposed to a countable number of charged particles. We now consider how to compute $V(\mathbf{r})$ three types of these commonly-encountered distributions. Before beginning, it's worth noting that the methods will be essentially the same, from a mathematical viewpoint, as those developed in Section 5.4; therefore, a review of that section may be helpful before attempting this section.

Continuous Distribution of Charge Along a Curve

Consider a continuous distribution of charge along a curve $\mathcal{C}$. The curve can be divided into short segments of length Δl. Then, the charge associated with the n^{th} segment, located at $\mathbf{r}_n$, is

$$q_n = \rho_l(\mathbf{r}_n)\ \Delta l$$

where ρ_l is the line charge density (units of C/m) at $\mathbf{r}_n$. Substituting this expression into Equation 5.13.1, we obtain

$$\mathbf{V}(\mathbf{r}) = \frac{1}{4\pi\epsilon}\sum_{n=1}^{N}\frac{\rho_l(\mathbf{r}_n)}{|\mathbf{r}-\mathbf{r}_n|}\Delta l$$

Taking the limit as $\Delta l \to 0$ yields:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon}\int_{\mathcal{C}}\frac{\rho_l(l)}{|\mathbf{r}-\mathbf{r}'|}dl \tag{5.13.2}$$

where $\mathbf{r}'$ represents the varying position along $\mathcal{C}$ with integration along the length l.

Continuous Distribution of Charge Over a Surface

Consider a continuous distribution of charge over a surface $\mathcal{S}$. The surface can be divided into small patches having area Δs. Then, the charge associated with the n^{th} patch, located at $\mathbf{r}_n$, is

$$q_n = \rho_s(\mathbf{r}_n)\ \Delta s$$

where ρ_s is surface charge density (units of C/m^2) at $\mathbf{r}_n$. Substituting this expression into Equation 5.13.1, we obtain

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon}\sum_{n=1}^{N}\frac{\rho_s(\mathbf{r}_n)}{|\mathbf{r}-\mathbf{r}_n|}\ \Delta s$$

Taking the limit as $\Delta s \to 0$ yields:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon}\int_{\mathcal{S}}\frac{\rho_s(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\ ds \tag{5.13.3}$$

where $\mathbf{r}'$ represents the varying position over $\mathcal{S}$ with integration.

Continuous Distribution of Charge in a Volume

Consider a continuous distribution of charge within a volume $\mathcal{V}$. The volume can be divided into small cells (volume elements) having area Δv. Then, the charge associated with the n^{th} cell, located at $\mathbf{r}_n$, is

$$q_n = \rho_v(\mathbf{r}_n)\ \Delta v$$

where ρ_v is the volume charge density (units of C/m^3) at $\mathbf{r}_n$. Substituting this expression into Equation 5.13.1, we obtain

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon}\sum_{n=1}^{N}\frac{\rho_v(\mathbf{r}_n)}{|\mathbf{r}-\mathbf{r}_n|}\ \Delta v$$

Taking the limit as $\Delta v \to 0$ yields:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon} \int_{\mathcal{V}} \frac{\rho_v(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, dv$$

where r' represents the varying position over $\mathcal{V}$ with integration.

This page titled 5.13: Electric Potential Field due to a Continuous Distribution of Charge is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.13: Electric Potential Field due to a Continuous Distribution of Charge** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.14: Electric Field as the Gradient of Potential

In Section 5.8, it was determined that the electrical potential difference V_{21} measured over a path $\mathcal{C}$ is given by

$$V_{21} = -\int_{\mathcal{C}} \mathbf{E}(\mathbf{r}) \cdot d\mathbf{l} \tag{5.14.1}$$

where $\mathbf{E}(\mathbf{r})$ is the electric field intensity at each point $\mathbf{r}$ along $\mathcal{C}$. In Section 5.12, we defined the scalar electric potential field $V(\mathbf{r})$ as the electric potential difference at $\mathbf{r}$ relative to a datum at infinity. In this section, we address the "inverse problem" – namely, how to calculate $\mathbf{E}(\mathbf{r})$ given $V(\mathbf{r})$. Specifically, we are interested in a direct "point-wise" mathematical transform from one to the other. Since Equation 5.14.1 is in the form of an integral, it should not come as a surprise that the desired expression will be in the form of a differential equation.

We begin by identifying the contribution of an infinitesimal length of the integral to the total integral in Equation 5.14.1. At point $\mathbf{r}$, this is

$$dV = -\mathbf{E}(\mathbf{r}) \cdot d\mathbf{l} \tag{5.14.2}$$

Although we can proceed using any coordinate system, the following derivation is particularly simple in Cartesian coordinates. In Cartesian coordinates,

$$d\mathbf{l} = \hat{\mathbf{x}}dx + \hat{\mathbf{y}}dy + \hat{\mathbf{z}}dz$$

We also note that for any scalar function of position, including $V(\mathbf{r})$, it is true that

$$dV = \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz$$

Note the above relationship is not specific to electromagnetics; it is simply mathematics. Also note that $dx = d\mathbf{l} \cdot \hat{\mathbf{x}}$ and so on for dy and dz. Making these substitutions into the above equation, we obtain:

$$dV = \frac{\partial V}{\partial x}\left(d\mathbf{l} \cdot \hat{\mathbf{x}}\right) + \frac{\partial V}{\partial y}\left(d\mathbf{l} \cdot \hat{\mathbf{y}}\right) + \frac{\partial V}{\partial z}\left(d\mathbf{l} \cdot \hat{\mathbf{z}}\right)$$

This equation may be rearranged as follows:

$$dV = \left(\left[\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}\right]V\right) \cdot d\mathbf{l}$$

Comparing the above equation to Equation 5.14.2, we find:

$$\mathbf{E}(\mathbf{r}) = -\left[\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}\right]V$$

Note that the quantity in square brackets is the gradient operator "∇" (Section 4.5). Thus, we may write

$$\boxed{\mathbf{E} = -\nabla V} \tag{5.14.3}$$

which is the relationship we seek.

> The electric field intensity at a point is the gradient of the electric potential at that point after a change of sign (Equation 5.14.3).

Using Equation 5.14.3, we can immediately find the electric field at any point $\mathbf{r}$ if we can describe V as a function of $\mathbf{r}$. Furthermore, this relationship between V and $\mathbf{E}$ has a useful physical interpretation. Recall that the gradient of a scalar field is a vector that points in the direction in which that field increases most quickly. Therefore:

> The electric field points in the direction in which the electric potential most rapidly decreases.

This result should not come as a complete surprise; for example, the reader should already be aware that the electric field points away from regions of net positive charge and toward regions of net negative charge (Sections 2.2 and/or 5.1). What is new here is

that both the magnitude and direction of the electric field may be determined given only the potential field, without having to consider the charge that is the physical source of the electrostatic field.

Example 5.14.1: Electric field of a charged particle, beginning with the potential field

In this example, we determine the electric field of a particle bearing charge q located at the origin. This may be done in a "direct" fashion using Coulomb's Law (Section 5.1). However, here we have the opportunity to find the electric field using a different method. In Section 5.12 we found the scalar potential for this source was:

$$V(\mathbf{r}) = \frac{q}{4\pi\epsilon r}$$

So, we may obtain the electric field using Equation 5.14.3:

$$\mathbf{E} = -\nabla V = -\nabla\left(\frac{q}{4\pi\epsilon r}\right)$$

Here $V(\mathbf{r})$ is expressed in spherical coordinates, so we have (Appendix 10.5):

$$\mathbf{E} = -\left[\hat{\mathbf{r}}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial\theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right]\left(\frac{q}{4\pi\epsilon r}\right)$$

In this case, $V(\mathbf{r})$ does not vary with ϕ or θ, so the second and third terms of the gradient are zero. This leaves

$$\begin{aligned}\mathbf{E} &= -\hat{\mathbf{r}}\frac{\partial}{\partial r}\left(\frac{q}{4\pi\epsilon r}\right) \\ &= -\hat{\mathbf{r}}\frac{q}{4\pi\epsilon}\frac{\partial}{\partial r}\frac{1}{r} \\ &= -\hat{\mathbf{r}}\frac{q}{4\pi\epsilon}\left(-\frac{1}{r^2}\right)\end{aligned} \tag{5.14.4}$$

So we find

$$\mathbf{E} = +\hat{\mathbf{r}}\frac{q}{4\pi\epsilon r^2}$$

as was determined in Section 5.1.

This page titled 5.14: Electric Field as the Gradient of Potential is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.14: Electric Field as the Gradient of Potential** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.15: Poisson's and Laplace's Equations

The electric scalar potential field $V(\mathbf{r})$, defined in Section 5.12, is useful for a number of reasons including the ability to conveniently compute potential differences (i.e., $V_{21} = V(\mathbf{r}_2) - V(\mathbf{r}_1)$) and the ability to conveniently determine the electric field by taking the gradient (i.e., $\mathbf{E} = -\nabla V$). One way to obtain $V(\mathbf{r})$ is by integration over the source charge distribution, as described in Section 5.13. This method is awkward in the presence of material interfaces, which impose boundary conditions on the solutions that must be satisfied simultaneously. For example, the electric potential on a perfectly conducting surface is constant[1] – a constraint which is not taken into account in any of the expressions in Section 5.13 (this fact is probably already known to the reader from past study of elementary circuit theory; however, this is established in the context of electromagnetics in Section 5.19.)

In this section, we develop an alternative approach to calculating $V(\mathbf{r})$ that accommodates these boundary conditions, and thereby facilitates the analysis of the scalar potential field in the vicinity of structures and spatially-varying material properties. This alternative approach is based on *Poisson's Equation*, which we now derive.

We begin with the differential form of Gauss' Law (Section 5.7):

$$\nabla \cdot \mathbf{D} = \rho_v$$

Using the relationship $\mathbf{D} = \epsilon \mathbf{E}$ (and keeping in mind our standard assumptions about material properties, summarized in Section 2.8) we obtain

$$\nabla \cdot \mathbf{E} = \frac{\rho_v}{\epsilon}$$

Next, we apply the relationship (Section 5.14):

$$\mathbf{E} = -\nabla V$$

yielding

$$\nabla \cdot \nabla V = -\frac{\rho_v}{\epsilon}$$

This is Poisson's Equation, but it is not in the form in which it is commonly employed. To obtain the alternative form, consider the operator $\nabla \cdot \nabla$ in Cartesian coordinates:

$$\begin{aligned} \nabla \cdot \nabla &= \left[\frac{\partial}{\partial x}\hat{\mathbf{x}} + \frac{\partial}{\partial y}\hat{\mathbf{y}} + \frac{\partial}{\partial z}\hat{\mathbf{z}}\right] \cdot \left[\frac{\partial}{\partial x}\hat{\mathbf{x}} + \frac{\partial}{\partial y}\hat{\mathbf{y}} + \frac{\partial}{\partial z}\hat{\mathbf{z}}\right] \\ &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \\ &= \nabla^2 \end{aligned}$$

i.e., the operator $\nabla \cdot \nabla$ is identically the Laplacian operator ∇^2 (Section 4.10). Furthermore, this is true regardless of the coordinate system employed. Thus, we obtain the following form of Poisson's Equation:

$$\boxed{\nabla^2 V = -\frac{\rho_v}{\epsilon}} \tag{5.15.1}$$

> Poisson's Equation (Equation 5.15.1) states that the Laplacian of the electric potential field is equal to the volume charge density divided by the permittivity, with a change of sign.

Note that Poisson's Equation is a partial differential equation, and therefore can be solved using well-known techniques already established for such equations. In fact, Poisson's Equation is an *inhomogeneous differential equation*, with the inhomogeneous part $-\rho_v/\epsilon$ representing the source of the field. In the presence of material structure, we identify the relevant boundary conditions at the interfaces between materials, and the task of finding $V(\mathbf{r})$ is reduced to the purely mathematical task of solving the associated boundary value problem (see "Additional Reading" at the end of this section). This approach is particularly effective when one of the materials is a perfect conductor or can be modeled as such a material. This is because – as noted at the beginning of this section – the electric potential at all points on the surface of a perfect conductor must be equal, resulting in a particularly simple boundary condition.

In many other applications, the charge responsible for the electric field lies outside the domain of the problem; i.e., we have non-zero electric field (hence, potentially non-zero electric potential) in a region that is free of charge. In this case, Poisson's Equation simplifies to *Laplace's Equation*:

$$\boxed{\nabla^2 V = 0 \quad \text{(source-free region)}} \tag{5.15.2}$$

> Laplace's Equation (Equation 5.15.2) states that the Laplacian of the electric potential field is zero in a source-free region.

Like Poisson's Equation, Laplace's Equation, combined with the relevant boundary conditions, can be used to solve for $V(\mathbf{r})$, but only in regions that contain no charge.

This page titled 5.15: Poisson's and Laplace's Equations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.15: Poisson's and Laplace's Equations** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.16: Potential Field Within a Parallel Plate Capacitor

This section presents a simple example that demonstrates the use of Laplace's Equation (Section 5.15) to determine the potential field in a source free region. The example, shown in Figure 5.16.1, pertains to an important structure in electromagnetic theory – the parallel plate capacitor. Here we are concerned only with the potential field $V(\mathbf{r})$ between the plates of the capacitor; you do not need to be familiar with capacitance or capacitors to follow this section (although you're welcome to look ahead to Section 5.22 for a preview, if desired). What *is* recommended before beginning is a review of the battery-charged capacitor experiment discussed in Section 2.2. In this section you'll see a rigorous derivation of what we figured out in an informal way in that section.

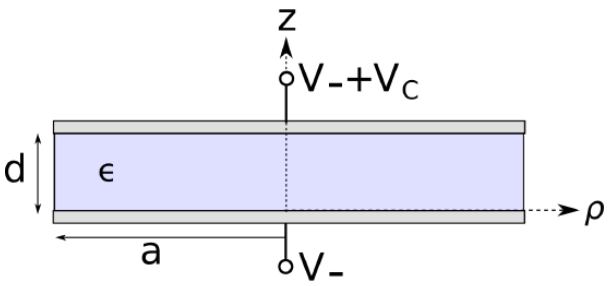


Figure 5.16.1: A parallel plate capacitor, as a demonstration of the use of Laplace's Equation.

The parallel-plate capacitor in Figure 5.16.1 consists of two perfectly-conducting circular disks separated by a distance d by a spacer material having permittivity ϵ. There is no charge present in the spacer material, so Laplace's Equation applies. That equation is (Section 5.15):

$$\nabla^2 V = 0 \quad \text{(source-free region)} \tag{5.16.1}$$

Let V_C be the potential difference between the plates, which would also be the potential difference across the terminals of the capacitor. The radius a of the plates is larger than d by enough that we may neglect what is going on at at the edges of the plates – more on this will be said as we work the problem. Under this assumption, what is the electric potential field $V(\mathbf{r})$ between the plates?

This problem has cylindrical symmetry, so it makes sense to continue to use cylindrical coordinates with the z axis being perpendicular to the plates. Equation 5.16.1 in cylindrical coordinates is:

$$\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right)+\frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2}+\frac{\partial^2}{\partial z^2}\right]V=0$$

or perhaps a little more clearly written as follows:

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial V}{\partial\rho}\right)+\frac{1}{\rho^2}\frac{\partial^2 V}{\partial\phi^2}+\frac{\partial^2 V}{\partial z^2}=0$$

Since the problem has radial symmetry, $\partial V/\partial\phi = 0$. Since $d \ll a$, we expect the fields to be approximately constant with ρ until we get close to the edge of the plates. Therefore, we assume $\partial V/\partial\rho$ is negligible and can be taken to be zero. Thus, we are left with

$$\frac{\partial^2 V}{\partial z^2} \approx 0 \quad \text{for } \rho \ll a \tag{5.16.2}$$

The general solution to Equation 5.16.2 is obtained simply by integrating both sides twice, yielding

$$V(z) = c_1 z + c_2 \tag{5.16.3}$$

where c_1 and c_2 are constants that must be consistent with the boundary conditions. Thus, we must develop appropriate boundary conditions. Let the node voltage at the negative ($z = 0$) terminal be V_-. Then the voltage at the positive ($z = +d$) terminal is $V_- + V_C$. Therefore: These are the relevant boundary conditions. Substituting $V(z = 0) = V_-$ into Equation 5.16.3 yields $c_2 = V_-$. Substituting $V(z = +d) = V_- + V_C$ into Equation 5.16.3 yields $c_1 = V_C/d$. Thus, the answer to the problem is

$$V(z) \approx \frac{V_C}{d} z + V_- \quad \text{for } \rho \ll a \tag{5.16.4}$$

Note that the above result is dimensionally correct and confirms that the potential deep inside a "thin" parallel plate capacitor changes linearly with distance between the plates.

Further, you should find that application of the equation $\mathbf{E} = -\nabla V$ (Section 5.14) to the solution above yields the expected result for the electric field intensity: $\mathbf{E} \approx -\hat{\mathbf{z}} V_C / d$. This is precisely the result that we arrived at (without the aid of Laplace's Equation) in Section 2.2.

A reasonable question to ask at this point would be, what about the potential field close to the edge of the plates, or, for that matter, beyond the plates? The field in this region is referred to as a *fringing field*. For the fringing field, $\partial V / \partial \rho$ is no longer negligible and must be taken into account. In addition, it is necessary to modify the boundary conditions to account for the outside surfaces of the plates (that is, the sides of the plates that face away from the dielectric) and to account for the effect of the boundary between the spacer material and free space. These issues make the problem much more difficult. When an accurate calculation of a fringing field is necessary, it is common to resort to a numerical solution of Laplace's Equation. Fortunately, accurate calculation of fringing fields is usually not required in practical engineering applications.

This page titled 5.16: Potential Field Within a Parallel Plate Capacitor is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.16: Potential Field Within a Parallel Plate Capacitor** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.17: Boundary Conditions on the Electric Field Intensity (E)

In homogeneous media, electromagnetic quantities vary smoothly and continuously. At an interface between dissimilar media, however, it is possible for electromagnetic quantities to be discontinuous. These discontinuities can be described mathematically as *boundary conditions* and used to to constrain solutions for the associated electromagnetic quantities. In this section, we derive boundary conditions on the electric field intensity $\mathbf{E}$.

To begin, consider a region consisting of only two media that meet at an interface defined by the mathematical surface $\mathcal{S}$, as shown in Figure 5.17.1.

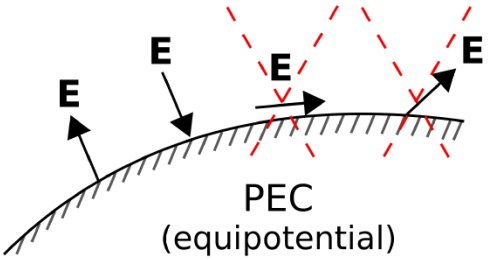


Figure 5.17.1: At the surface of a perfectly-conducting region, E may be perpendicular to the surface (two leftmost possibilities), but may not exhibit a component that is tangent to the surface (two rightmost possibilities). (© CC BY SA 4.0; K. Kikkeri)

If either one of the materials is a perfect electrical conductor (PEC), then $\mathcal{S}$ is an equipotential surface; i.e., the electric potential V is constant everywhere on $\mathcal{S}$. Since $\mathbf{E}$ is proportional to the spatial rate of change of potential (recall $\mathbf{E} = -\nabla V$; Section 5.14, we find:

> The component of $\mathbf{E}$ that is tangent to a perfectly-conducting surface is zero.

This is sometimes expressed informally as follows:

$$\boxed{E_{tan} = 0 \ \text{ on PEC surface}}$$

where "E_{tan}" is understood to be the component of $\mathbf{E}$ that is tangent to $\mathcal{S}$. Since the tangential component of $\mathbf{E}$ on the surface of a perfect conductor is zero, the electric field at the surface must be oriented entirely in the direction perpendicular to the surface, as shown in Figure 5.17.1.

The following equation expresses precisely the same idea, but includes the calculation of the tangential component as part of the statement:

$$\mathbf{E} \times \hat{\mathbf{n}} = 0 \ \ (\text{on PEC surface}) \tag{5.17.1}$$

where $\hat{\mathbf{n}}$ is either normal (i.e., unit vector perpendicular to the surface) to each point on $\mathcal{S}$. This expression works because the cross product of any two vectors is perpendicular to either vector (Section 4.1), and any vector which is perpendicular to $\hat{\mathbf{n}}$ is tangent to $\mathcal{S}$.

We now determine a more general boundary condition that applies even when neither of the media bordering $\mathcal{S}$ is a perfect conductor. The desired boundary condition can be obtained directly from Kirchoff's Voltage Law (KVL; Section 5.10):

$$\oint_{\mathcal{C}} \mathbf{E} \cdot d\mathbf{l} = 0 \tag{5.17.2}$$

Let the closed path of integration take the form of a rectangle centered on $\mathcal{S}$, as shown in Figure 5.17.2.

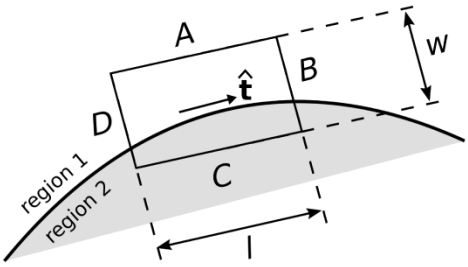


Figure 5.17.2: Use of KVL to determine the boundary condition on $\mathbf{E}$. (CC BY SA 4.0; K. Kikkeri)

Let the sides A, B, C, and D be perpendicular or parallel to the surface, respectively. Let the length of the perpendicular sides be w, and let the length of the parallel sides be l. From KVL we have

$$\begin{aligned}\oint_{\mathcal{C}} \mathbf{E}\cdot d\mathbf{l} = &\int_A \mathbf{E}\cdot d\mathbf{l} \\ &+ \int_B \mathbf{E}\cdot d\mathbf{l} \\ &+ \int_C \mathbf{E}\cdot d\mathbf{l} \\ &+ \int_D \mathbf{E}\cdot d\mathbf{l} \qquad = 0\end{aligned}$$

Now, let us reduce w and l together while (1) maintaining a constant ratio $w/l \ll 1$ and (2) keeping $\mathcal{C}$ centered on $\mathcal{S}$. In this process, the contributions from the B and D segments become equal in magnitude but opposite in sign; i.e.,

$$\int_B \mathbf{E}\cdot d\mathbf{l} + \int_D \mathbf{E}\cdot d\mathbf{l} \to 0$$

This leaves

$$\oint_{\mathcal{C}} \mathbf{E}\cdot d\mathbf{l} \to \int_A \mathbf{E}\cdot d\mathbf{l} + \int_C \mathbf{E}\cdot d\mathbf{l} \to 0$$

Let us define the unit vector $\hat{\mathbf{t}}$ ("tangent") as shown in Figure 5.17.2. When the lengths of sides A and C become sufficiently small, we can write the above expression as follows:

$$\mathbf{E}_1\cdot\hat{\mathbf{t}}\Delta l - \mathbf{E}_2\cdot\hat{\mathbf{t}}\Delta l \to 0 \tag{5.17.3}$$

where $\mathbf{E}_1$ and $\mathbf{E}_2$ are the fields evaluated on the two sides of the boundary and $\Delta l \to 0$ is the length of sides A and C while this is happening. Note that the only way Equation 5.17.3 can be true is if the tangential components of $\mathbf{E}_1$ and $\mathbf{E}_2$ are equal. In other words:

> The tangential component of $\mathbf{E}$ must be continuous across an interface between dissimilar media.

Note that this is a generalization of the result we obtained earlier for the case in which one of the media was a PEC – in that case, the tangent component of $\mathbf{E}$ on the other side of the interface must be zero because it is zero in the PEC medium.

As before, we can express this idea in compact mathematical notation. Using the same idea used to obtain Equation 5.17.1, we have found

$$\mathbf{E}_1\times\hat{\mathbf{n}} = \mathbf{E}_2\times\hat{\mathbf{n}} \quad \text{on } \mathcal{S}$$

or, as it is more commonly written:

$$\boxed{\hat{\mathbf{n}}\times(\mathbf{E}_1-\mathbf{E}_2)=0 \quad \text{on } \mathcal{S}} \tag{5.17.4}$$

We conclude this section with a note about the broader applicability of this boundary condition:

> Equation 5.17.4 is the boundary condition that applies to $\mathbf{E}$ for both the electrostatic *and the general (time-varying) case.*

Although a complete explanation is not possible without the use of the Maxwell-Faraday Equation (Section 8.8), the reason why this boundary condition applies in the time-varying case can be disclosed here. In the presence of time-varying magnetic fields, the right-hand side of Equation 5.17.2 may become non-zero and is proportional to the area defined by the closed loop. However, the above derivation requires the area of this loop to approach zero, in which case the possible difference from Equation 5.17.2 also converges to zero. Therefore, the boundary condition expressed in Equation 5.17.4 applies generally.

This page titled 5.17: Boundary Conditions on the Electric Field Intensity (E) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.17: Boundary Conditions on the Electric Field Intensity (E)** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.18: Boundary Conditions on the Electric Flux Density (D)

In this section, we derive boundary conditions on the electric flux density $\mathbf{D}$. The considerations are quite similar to those encountered in the development of boundary conditions on the electric *field intensity* ($\mathbf{E}$) in Section 5.17, so the reader may find it useful to review that section before attempting this section. This section also assumes familiarity with the concepts of electric flux, electric flux density, and Gauss' Law; for a refresher, Sections 2.4 and 5.5 are suggested.

To begin, consider a region at which two otherwise-homogeneous media meet at an interface defined by the mathematical surface $\mathcal{S}$, as shown in Figure 5.18.1. Let one of these regions be a perfect electrical conductor (PEC).

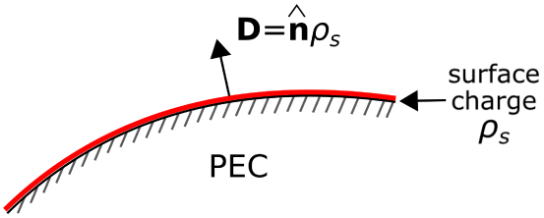


Figure 5.18.1: The component of $\mathbf{D}$ that is perpendicular to a perfectly-conducting surface is equal to the charge density on the surface. (CC BY SA 4.0; K. Kikkeri)

In Section 5.17, we established that the tangential component of the electric field must be zero, and therefore, the electric field is directed entirely in the direction perpendicular to the surface. We further know that the electric field within the conductor is identically zero. Therefore, $\mathbf{D}$ at any point on $\mathcal{S}$ is entirely in the direction perpendicular to the surface and pointing into the non-conducting medium. However, it is also possible to determine the magnitude of $\mathbf{D}$. We shall demonstrate in this section that

> At the surface of a perfect conductor, the magnitude of $\mathbf{D}$ is equal to the surface charge density ρ_s (units of C/m^2) at that point.

The following equation expresses precisely the same idea, but includes the calculation of the perpendicular component as part of the statement:

$$\mathbf{D} \cdot \hat{\mathbf{n}} = \rho_s \quad (\text{on PEC surface}) \tag{5.18.1}$$

where $\hat{\mathbf{n}}$ is the normal to $\mathcal{S}$ pointing into the non-conducting region. (Note that the orientation of $\hat{\mathbf{n}}$ is now important; we have assumed $\hat{\mathbf{n}}$ points into region 1, and we must now stick with this choice.) Before proceeding with the derivation, it may be useful to note that this result is not surprising. The very definition of electric flux (Section 2.4) indicates that $\mathbf{D}$ should correspond in the same way to a surface charge density. However, we can show this rigorously, and in the process we can generalize this result to the more-general case in which neither of the two materials are PEC.

The desired more-general boundary condition may be obtained from the integral form of Gauss' Law (Section 5.5), as illustrated in Figure 5.18.2.

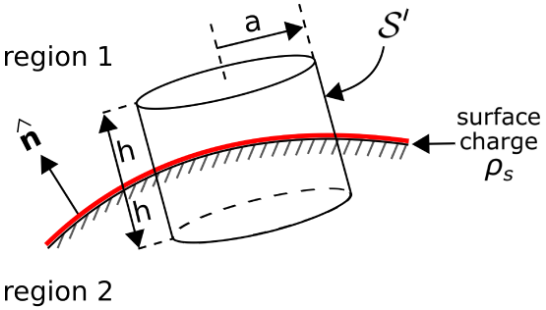


Figure 5.18.2: Use of Gauss' Law to determine the boundary condition on $\mathbf{D}$. (CC BY SA 4.0; K. Kikkeri)

Let the surface of integration $\mathcal{S}'$ take the form of closed cylinder centered at a point on the interface and for which the flat ends are parallel to the surface and perpendicular to $\hat{\mathbf{n}}$. Let the radius of this cylinder be a, and let the length of the cylinder be $2h$. From Gauss' Law we have

$$\oint_{\mathcal{S}'} \mathbf{D} \cdot d\mathbf{s} = \int_{top} \mathbf{D} \cdot d\mathbf{s} + \int \mathbf{D} \cdot d\mathbf{s}$$

$$+\int_{bottom} \mathbf{D}\cdot d\mathbf{s} \qquad = Q_{encl}$$

where the "top" and "bottom" are in Regions 1 and 2, respectively, and Q_{encl} is the charge enclosed by $\mathcal{S}'$. Now let us reduce h and a together while (1) maintaining a constant ratio $h/a \ll 1$ and (2) keeping $\mathcal{S}'$ centered on $\mathcal{S}$. Because $h \ll a$, the area of the side can be made negligible relative to the area of the top and bottom. Then as $h \to 0$ we are left with

$$\int_{top} \mathbf{D}\cdot d\mathbf{s} + \int_{bottom} \mathbf{D}\cdot d\mathbf{s} \to Q_{encl}$$

As the area of the top and bottom sides become infinitesimal, the variation in $\mathbf{D}$ over these areas becomes negligible. Now we have simply:

$$\mathbf{D}_1\cdot\hat{\mathbf{n}}\Delta A + \mathbf{D}_2\cdot(-\hat{\mathbf{n}})\,\Delta A \to Q_{encl}$$

where $\mathbf{D}_1$ and $\mathbf{D}_2$ are the electric flux density vectors in medium 1 and medium 2, respectively, and ΔA is the area of the top and bottom sides. The above expression can be rewritten

$$\hat{\mathbf{n}}\cdot(\mathbf{D}_1 - \mathbf{D}_2) \to \frac{Q_{encl}}{\Delta A}$$

Note that the left side of the equation must represent a actual, physical surface charge; this is apparent from dimensional analysis and the fact that h is now infinitesimally small. Therefore:

$$\boxed{\hat{\mathbf{n}}\cdot(\mathbf{D}_1 - \mathbf{D}_2) = \rho_s} \tag{5.18.2}$$

where, as noted above, $\hat{\mathbf{n}}$ points into region 1. Summarizing:

> Any discontinuity in the normal component of the electric flux density across the boundary between two material regions is equal to the surface charge.

Now let us verify that this is consistent with our preliminary finding, in which Region 2 was a PEC. In that case $\mathbf{D}_2 = 0$, so we see that Equation 5.18.1 is satisfied, as expected. If neither Region 1 nor Region 2 is PEC *and* there is no surface charge on the interface, then we find $\hat{\mathbf{n}}\cdot(\mathbf{D}_1 - \mathbf{D}_2) = 0$; i.e.,

> In the absence of surface charge, the normal component of the electric flux density must be continuous across the boundary.

Finally, we note that since $\mathbf{D} = \epsilon\mathbf{E}$, Equation 5.18.2 implies the following boundary condition on $\mathbf{E}$:

$$\hat{\mathbf{n}}\cdot(\epsilon_1\mathbf{E}_1 - \epsilon_2\mathbf{E}_2) = \rho_s \tag{5.18.3}$$

where ϵ_1 and ϵ_2 are the permittivities in Regions 1 and 2, respectively. The above equation illustrates one reason why we sometimes prefer the "flux" interpretation of the electric field to the "field intensity" interpretation of the electric field.

This page titled 5.18: Boundary Conditions on the Electric Flux Density (D) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.18: Boundary Conditions on the Electric Flux Density (D)** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.19: Charge and Electric Field for a Perfectly Conducting Region

In this section, we consider the behavior of charge and the electric field in the vicinity of a perfect electrical conductor (PEC).

First, note that the electric field – both the electric field intensity $\mathbf{E}$ and electric flux density $\mathbf{D}$ – throughout a PEC region is zero. This is because the electrical potential throughout a PEC region must be constant. (This idea is explored further in Section 6.3). Recall that the electric field is proportional to the spatial rate of change of electrical potential (i.e., $\mathbf{E} = -\nabla V$; Section 5.14). Thus, the electric field must be zero throughout a PEC region.

Second, the electric field is oriented directly away from (i.e., perpendicular to) the PEC surface, and the magnitude of $\mathbf{D}$ is equal to the surface charge density ρ_s (C/m^2) (Section 5.18).

Now we address the question of charge distribution; i.e., the location and density of charge. Consider the scenario shown in Figure 5.19.1.

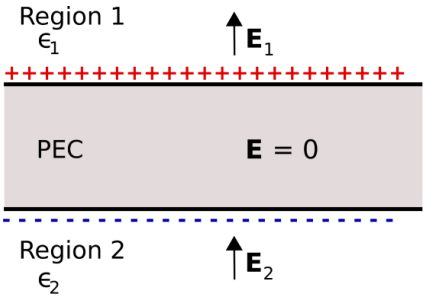


Figure 5.19.1: An infinite flat slab of PEC in the presence of an applied electric field. (CC BY SA 4.0; K. Kikkeri).

Here, a flat slab of PEC material is embedded in dielectric material.[1] The thickness of the slab is finite, whereas the length and width of the slab is infinite. The region above the slab is defined as Region 1 and has permittivity ϵ_1. The region below the slab is defined as Region 2 and has permittivity ϵ_2. Electric fields $\mathbf{E}_1$ and $\mathbf{E}_2$ are present in Regions 1 and 2, respectively, as shown in Figure 5.19.1. To begin, let us assume that these fields are the result of some external stimulus that results in the direction of these fields being generally upward, as shown in Figure 5.19.1.

Now, what do we know about $\mathbf{E}_1$ and $\mathbf{E}_2$? First, both fields must satisfy the relevant boundary conditions. That is, the component of $\mathbf{E}_1$ that is tangent to the upper PEC surface is zero, so that $\mathbf{E}_1$ is directed entirely in a direction perpendicular to the surface. Similarly, the component of $\mathbf{E}_2$ that is tangent to the lower PEC surface is zero, so that $\mathbf{E}_2$ is directed entirely in a direction perpendicular to the surface. At this point we have not determined the magnitudes or signs of $\mathbf{E}_1$ and $\mathbf{E}_2$; we have established only that there are no non-zero components tangential to (i.e., parallel to) the PEC surfaces.

Next, recall that the electric field must be zero within the slab. This means that there must be zero net charge within the slab, since any other distribution of charge will result in a non-zero electric field, and subsequently a potential difference between locations within the slab. Therefore:

> There can be no static charge *within* a PEC.

It follows that

> Charge associated with a PEC lies entirely on the surface.

Outside the slab, the boundary conditions on $\mathbf{D}_1$ and $\mathbf{D}_2$ in the dielectric regions require these fields to be non-zero when the surface charge density on the PEC is non-zero. The surface charge supports the discontinuity in the normal component of the electric fields. Specifically, $\mathbf{D}_1$ and $\mathbf{D}_2$ have the same magnitude $|\rho_s|$ because the surface charge densities on both sides of the slab have equal magnitude. However, the electric field intensity $\mathbf{E}_1 = \mathbf{D}_1/\epsilon_1$, whereas $\mathbf{E}_2 = \mathbf{D}_2/\epsilon_2$; i.e., these are different. That is, the electric field intensities are unequal unless the permittivities in each dielectric region are equal.

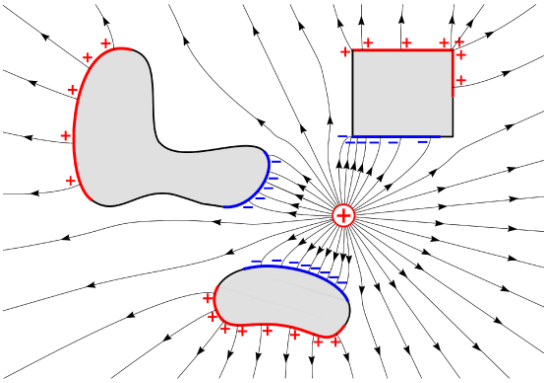


Figure 5.19.2: Electric field lines due to a point charge in the vicinity of PEC regions (shaded) of various shapes. (CC BY SA 4.0; K. Kikkeri).

Finally, let us consider the structure of the electric field in more general cases. Figure 5.19.2 shows field lines in a homogeneous dielectric material in which a point charge and PEC regions of various shapes are embedded. Note that electric field lines now bend in the dielectric so as to satisfy the requirement that the tangential component of the electric field be zero on PEC surfaces. Also note that the charge distribution arranges itself on the PEC surfaces so as to maintain zero electric field and constant potential within the cube.

---

1. For the purposes of this section, it suffices to interpret “dielectric” as a “nonconducting and well characterized entirely in terms of its permittivity.” For more, see Section 5.20.↩

This page titled 5.19: Charge and Electric Field for a Perfectly Conducting Region is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.19: Charge and Electric Field for a Perfectly Conducting Region** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.20: Dielectric Media

Dielectric is particular category of materials that exhibit low conductivity because their constituent molecules remain intact when exposed to an electric field, as opposed to shedding electrons as is the case in good conductors. Subsequently, dielectrics do not effectively pass current, and are therefore considered “good insulators” as well as “poor conductors.” An important application of dielectrics in electrical engineering is as a spacer material in printed circuit boards (PCBs), coaxial cables, and capacitors.

Examples of dielectrics include air, glass, teflon, and fiberglass epoxy (the material used in common “FR4” printed circuit boards). These and other dielectrics are listed along with values of their constitutive parameters in Section A1.

The electromagnetic properties of dielectric materials are quantified primarily by relative permittivity ϵ_r (Section 2.3), which ranges from very close to 1 upward to roughly 50, and is less than 6 or so for most commonly-encountered materials having low moisture content. The permeability of dielectric materials is approximately equal to the free-space value (i.e., $\mu \approx \mu_0$); therefore, these materials are said to be “non-magnetic.”

This page titled 5.20: Dielectric Media is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.20: Dielectric Media** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.21: Dielectric Breakdown

The permittivity of an ideal dielectric is independent of the magnitude of an applied electric field; the material is said to be "linear."[1] However, all practical dielectrics fail in this respect with sufficiently strong electric field. Typically, the failure is abrupt and is observed as a sudden, dramatic increase in conductivity, signaling that electrons are being successfully dislodged from their host molecules. The threshold value of the electric field intensity at which this occurs is known as the *dielectric strength*, and the sudden change in behavior observed in the presence of an electric field greater than this threshold value is known as *dielectric breakdown*.

Dielectric strength varies from about 3 MV/m for air to about 200 MV/m in mica (a dielectric commonly used in capacitors).

Dielectric breakdown is typically accompanied by "arcing," which is a sudden flow of current associated with the breakdown. A well known example of this phenomenon is lightning, which occurs when charge is exchanged between sky and ground when air (a dielectric) exhibits breakdown. Dielectric breakdown in solids typically damages the material.

1. See Section 2.8 for a review of this concept.↩

This page titled 5.21: Dielectric Breakdown is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.21: Dielectric Breakdown** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.22: Capacitance

When separate regions of positive and negative charge exist in proximity, Coulomb forces (Section 5.1) will attempt to decrease the separation between the charges. As noted in Section 5.8, this can be interpreted as a tendency of a system to reduce its potential energy. If the charges are fixed in place, then the potential energy remains constant. This potential energy is proportional to the Coulomb force. Referring back to Section 5.1, the Coulomb force is:

- Proportional to quantity of positive charge squared
- Inversely proportional to the separation between the charges squared
- Inversely proportional to the permittivity of the material separating the charges

Therefore, the potential energy of the system is likewise dependent on charge, separation, and permittivity. Furthermore, we see that the ability of a system to store energy in this manner depends on the geometry of the charge distribution and the permittivity of the intervening material.

Now recall that the electric field intensity $\mathbf{E}$ is essentially defined in terms of the Coulomb force; i.e., $\mathbf{F} = q\mathbf{E}$ (Section 2.2). So, rather than thinking of the potential energy of the system as being associated with the Coulomb force, it is equally valid to think of the potential energy as being stored in the electric field associated with the charge distribution. It follows from the previous paragraph that the energy stored in the electric field depends on the geometry of the charge distribution and the permittivity of the intervening media. This relationship is what we mean by *capacitance*. We summarize as follows:

> Capacitance is the ability of a structure to store energy in an electric field.

and

> The capacitance of a structure depends on its geometry and the permittivity of the medium separating regions of positive and negative charge.

Note that capacitance does *not* depend on charge, which we view as either a stimulus or response from this point of view. The corresponding response or stimulus, respectively, is the potential associated with this charge. This leads to the following definition:

$$\boxed{C \triangleq \frac{Q_+}{V}} \tag{5.22.1}$$

where Q_+ (units of C) is the total positive charge, V (units of V) is the potential associated with this charge (defined such that it is positive), and C (units of F) is the associated capacitance. So:

> In practice, capacitance is defined as the ratio of charge present on one conductor of a two-conductor system to the potential difference between the conductors (Equation 5.22.1).

In other words, a structure is said to have greater capacitance if it stores more charge – and therefore stores more energy – in response to a given potential difference.

Figure 5.22.1 shows the relevant features of this definition.

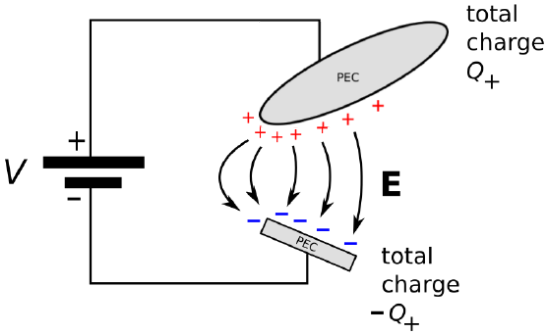


Figure 5.22.1: Electrostatic interpretation of capacitance.

Here, a battery imposes the potential difference V between two regions of perfectly-conducting material. Q_+ is the total charge on the surface of the PEC region attached to the positive terminal of the battery. An equal amount of negative charge appears on the surface of the PEC region attached to the negative terminal of the battery (Section 5.19). This charge distribution gives rise to an electric field. Assuming the two PEC regions are fixed in place, Q_+ will increase linearly with increasing V, at a rate determined by the capacitance C of the structure.

A *capacitor* is a device that is designed to exhibit a specified capacitance. We can now make the connection to the concept of the capacitor as it appears in elementary circuit theory. In circuit theory, the behavior of devices is characterized in terms of terminal voltage V_T in response to terminal current I_T, and vice versa. First, note that current does not normally flow through a capacitor,[1] so when we refer to "terminal current" for a capacitor, what we *really* mean is the flow of charge arriving or departing from one of the conductors via the circuit, which is equal to the flow of charge departing or arriving (respectively) at the other conductor. This gives the appearance of current flow through the capacitor when the current is examined from outside the capacitor. With that settled, we proceed as follows. Using Equation 5.22.1, we express the voltage V_T across the terminals of a capacitor having capacitance C:

$$V_T = \frac{Q_+}{C}$$

We seek a relationship between V_T and I_T. Current is charge per unit time, so the charge on either conductor is the integral of I_T over time; i.e.:

$$Q_+(t) = \int_{t_0}^{t} I_T(\tau)\ d\tau + Q_+(t_0)$$

where t_0 is an arbitrarily-selected start time. In other words, amps integrated over time is charge. If we define I_T as being positive in the direction of the flow of positive charge as is the usual convention, then we have:

$$V_T(t) = \frac{1}{C}\int_{t_0}^{t} I_T(\tau)\ d\tau + \frac{1}{C}Q_+(t_0)$$

Again applying Equation 5.22.1, we see that the second term is simply $V_T(t_0)$. This is the expected relationship from elementary circuit theory.

Finally, solving for I_T we obtain the differential form of this relationship:

$$I_T(t) = C\frac{d}{dt}V_T(t)$$

---

1. If it does, it's probably experiencing *dielectric breakdown*; see Section 5.21.↩

This page titled 5.22: Capacitance is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.22: Capacitance** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.23: The Thin Parallel Plate Capacitor

Let us now determine the capacitance of a common type of capacitor known as the *thin parallel plate capacitor*, shown in Figure 5.23.1. This capacitor consists of two flat plates, each having area A, separated by distance d. To facilitate discussion, let us place the origin of the coordinate system at the center of the lower plate, with the $+z$ axis directed toward the upper plate such that the upper plate lies in the $z = +d$ plane.

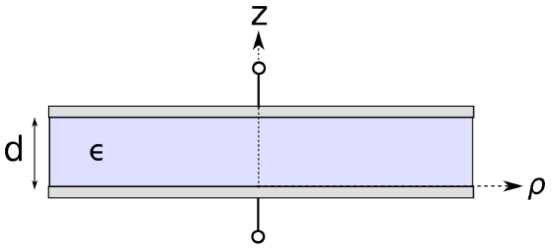


Figure 5.23.1: Thin parallel plate capacitor

Below we shall find the capacitance by assuming a particular charge on one plate, using the boundary condition on the electric flux density $\mathbf{D}$ to relate this charge density to the internal electric field, and then integrating over the electric field between the plates to obtain the potential difference. Then, capacitance is the ratio of the assumed charge to the resulting potential difference.

The principal difficulty in this approach is finding the electric field. To appreciate the problem, first consider that if the area of the plates was infinite, then the electric field would be very simple; it would begin at the positively-charged plate and extend in a perpendicular direction toward the negatively-charged plate (Section 5.19). Furthermore, the field would be constant everywhere between the plates. This much is apparent from symmetry alone. However, when the plate area is finite, then we expect a *fringing field* to emerge. "Fringing field" is simply a term applied to the non-uniform field that appears near the edge of the plates. The field is non-uniform in this region because the boundary conditions on the *outside* (outward-facing) surfaces of the plates have a significant effect in this region. In the central region of the capacitor, however, the field is not much different from the field that exists in the case of infinite plate area.

In any parallel plate capacitor having finite plate area, some fraction of the energy will be stored by the approximately uniform field of the central region, and the rest will be stored in the fringing field. We can make the latter negligible relative to the former by making the capacitor very "thin," in the sense that the smallest identifiable dimension of the plate is much greater than d. Under this condition, we may obtain a good approximation of the capacitance by simply neglecting the fringing field, since an insignificant fraction of the energy is stored there.

Imposing the "thin" condition leads to three additional simplifications. First, the surface charge distribution may be assumed to be approximately uniform over the plate, which greatly simplifies the analysis. Second, the shape of the plates becomes irrelevant; they might be circular, square, triangular, etc. When computing capacitance in the "thin" case, only the plate area A is important. Third, the thickness of each of the plates becomes irrelevant.

We are now ready to determine the capacitance of the thin parallel plate capacitor. Here are the steps:

1. Assume a total positive charge Q_+ on the upper plate.
2. Invoking the "thin" condition, we assume the charge density on the plates is uniform. Thus, the surface charge density on bottom side of the upper plate is $\rho_{s,+} = Q_+/A$ (C/m^2).
3. From the boundary condition on the bottom surface of the upper plate, $\mathbf{D}$ on this surface is $-\hat{\mathbf{z}}\rho_{s,+}$.
4. The total charge on the lower plate, Q_-, must be equal and opposite the total charge on the upper plate; i.e, $Q_- = -Q_+$. Similarly, the surface charge density on the upper surface of the lower plate, $\rho_{s,-}$, must be $-\rho_{s,+}$.
5. From the boundary condition on the top surface of the lower plate (Section 5.18), $\mathbf{D}$ on this surface is $+\hat{\mathbf{z}}\rho_{s,-}$. Since $+\hat{\mathbf{z}}\rho_{s,-} = -\hat{\mathbf{z}}\rho_{s,+}$, $\mathbf{D}$ on the facing sides of the plates is equal.
6. Again invoking the "thin" condition, we assume $\mathbf{D}$ between the plates has approximately the same structure as we would see if the plate area was infinite. Therefore, we are justified in assuming $\mathbf{D} \approx -\hat{\mathbf{z}}\rho_{s,+}$ *everywhere* between the plates. (You might also see that this is self-evident from the definition of $\mathbf{D}$ as the flux density of electric charge (Section 2.4).)
7. With an expression for the electric field in hand, we may now compute the potential difference V between the plates as follows (Section 5.8):

$$V = -\int_{\mathcal{C}} \mathbf{E}\cdot d\mathbf{l}$$

$$
\begin{aligned}
&= -\int_0 \left(\frac{1}{\epsilon}\mathbf{D}\right)\cdot(\hat{\mathbf{z}}dz) \\
&= -\int_0^d \left(-\hat{\mathbf{z}}\frac{\rho_{s,+}}{\epsilon}\right)\cdot(\hat{\mathbf{z}}dz) \\
&= +\frac{\rho_{s,+}d}{\epsilon}
\end{aligned}
\tag{5.23.1}
$$

8. Finally,

$$
C = \frac{Q_+}{V} = \frac{\rho_{s,+}\ A}{\rho_{s,+}\ d/\epsilon} = \frac{\epsilon A}{d}
$$

Summarizing:

$$
\boxed{C \approx \frac{\epsilon A}{d}} \tag{5.23.2}
$$

The capacitance of a parallel plate capacitor having plate separation much less than the size of the plate is given by Equation 5.23.2. This is an approximation because the fringing field is neglected.

It's worth noting that this is dimensionally correct; i.e., F/m times m^2 divided by m yields F. It's also worth noting the effect of the various parameters:

Capacitance increases in proportion to permittivity and plate area and decreases in proportion to distance between the plates.

Example 5.23.1: Printed circuit board capacitance

Printed circuit boards commonly include a "ground plane," which serves as the voltage datum for the board, and at least one "power plane," which is used to distribute a DC supply voltage (See "Additional Reading" at the end of this section). These planes are separated by a dielectric material, and the resulting structure exhibits capacitance. This capacitance may be viewed as an equivalent discrete capacitor in parallel with the power supply. The value of this equivalent capacitor may be either negligible, significant and beneficial, or significant and harmful. So, it is useful to know the value of this equivalent capacitor.

For a common type of circuit board, the dielectric thickness is about 1.6 mm and the relative permittivity of the material is about 4.5. If the area in common between the ground and power planes is 25 cm^2, what is the value of the equivalent capacitor?

Solution

From the problem statement, $\epsilon \cong 4.5\epsilon_0$, $A \cong 25\ \text{cm}^2 = 2.5\ \ \times 10^{-3}\ \text{m}^2$, and $d \cong 1.6$ mm. Using Equation 5.23.2, the value of the equivalent capacitor is 62.3 pF.

Additional Reading:

- "Printed circuit board" on Wikipedia.

This page titled 5.23: The Thin Parallel Plate Capacitor is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.23: The Thin Parallel Plate Capacitor** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.24: Capacitance of a Coaxial Structure

Let us now determine the capacitance of coaxially-arranged conductors, shown in Figure 5.24.1. Among other applications, this information is useful in the analysis of voltage and current waves on coaxial transmission line, as addressed in Sections 3.4 and 3.10.

For our present purposes, we may model the structure as consisting of two concentric perfectly-conducting cylinders of radii a and b, separated by an ideal dielectric having permittivity ϵ_s. We place the $+z$ axis along the common axis of the concentric cylinders so that the cylinders may be described as constant-coordinate surfaces $\rho = a$ and $\rho = b$.

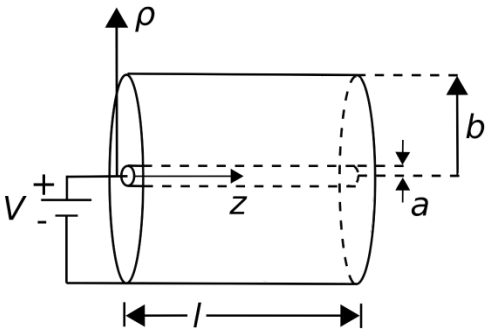


Figure 5.24.1: Determining the capacitance of a coaxial structure. (CC BY SA 4.0; K. Kikkeri)

In this section, we shall find the capacitance by assuming a total charge Q_+ on the inner conductor and integrating over the associated electric field to obtain the voltage between the conductors. Then, capacitance is computed as the ratio of the assumed charge to the resulting potential difference. This strategy is the same as that employed in Section 5.23 for the parallel plate capacitor, so it may be useful to review that section before attempting this derivation.

The first step is to find the electric field inside the structure. This is relatively simple if we assume that the structure has infinite length (i.e., $l \to \infty$), since then there are no fringing fields and the internal field will be utterly constant with respect to z. In the central region of a finite-length capacitor, however, the field is not much different from the field that exists in the case of infinite length, and if the energy storage in fringing fields is negligible compared to the energy storage in this central region then there is no harm in assuming the internal field is constant with z. Alternatively, we may think of the length l as pertaining to one short section of a much longer structure and thereby obtain the capacitance *per length* as opposed to the total capacitance. Note that the latter is exactly what we need for the transmission line lumped-element equivalent circuit model (Section 3.4).

To determine the capacitance, we invoke the definition (Section 5.22):

$$C \triangleq \frac{Q_+}{V} \tag{5.24.1}$$

where Q_+ is the charge on the positively-charged conductor and V is the potential measured from the negative conductor to the positive conductor. The charge on the inner conductor is uniformly-distributed with density

$$\rho_l = \frac{Q_+}{l}$$

which has units of C/m. Now we will determine the electric field intensity $\mathbf{E}$, integrate $\mathbf{E}$ over a path between conductors to get V, and then apply Equation 5.24.1 to obtain the capacitance.

The electric field intensity for this scenario was determined in Section 5.6, “Electric Field Due to an Infinite Line Charge using Gauss’ Law,” where we found

$$\mathbf{E} = \hat{\rho}\frac{\rho_l}{2\pi\epsilon_s\rho}$$

The reader should note that in that section we were considering merely a line of charge; not a coaxial structure. So, on what basis do we claim the field is the same? This is a consequence of Gauss’ Law (Section 5.5)

$$\oint_s \mathbf{D}\cdot d\mathbf{s} = Q_{encl}$$

which we used in Section 5.6 to find the field. If in this new problem we specify the same cylindrical surface $\mathcal{S}$ with radius $\rho < b$, then the enclosed charge is the same. Furthermore, the presence of the outer conductor does not change the radial symmetry of the problem, and nothing else remains that can change the outcome. This is worth noting for future reference:

> The electric field inside a coaxial structure comprised of concentric conductors and having uniform charge density on the inner conductor is identical to the electric field of a line charge in free space having the same charge density.

Next, we get V using (Section 5.8)

$$V = -\int_{\mathcal{C}} \mathbf{E} \cdot d\mathbf{l}$$

where $\mathcal{C}$ is any path from the negatively-charged outer conductor to the positively-charged inner conductor. Since this can be *any* such path (Section 5.9), we may as well choose the simplest one. This path is the one that traverses a radial of constant ϕ and z. Thus:

$$\begin{aligned} V &= -\int_{\rho=b}^{a} \left(\hat{\rho}\frac{\rho_l}{2\pi\epsilon_s\rho}\right) \cdot (\hat{\rho}d\rho) \\ &= -\frac{\rho_l}{2\pi\epsilon_s}\int_{\rho=b}^{a}\frac{d\rho}{\rho} \\ &= +\frac{\rho_l}{2\pi\epsilon_s}\int_{\rho=a}^{b}\frac{d\rho}{\rho} \\ &= +\frac{\rho_l}{2\pi\epsilon_s}\ln\left(\frac{b}{a}\right) \end{aligned} \tag{5.24.2}$$

Wrapping up:

$$C \triangleq \frac{Q_+}{V} = \frac{\rho_l l}{(\rho_l/2\pi\epsilon_s)\ln(b/a)}$$

Note that factors of ρ_l in the numerator and denominator cancel out, leaving:

$$\boxed{C = \frac{2\pi\epsilon_s l}{\ln(b/a)}}$$

Note that this expression is dimensionally correct, having units of F. Also note that the expression depends only on materials (through ϵ_s) and geometry (through l, a, and b). The expression does *not* depend on charge or voltage, which would imply non-linear behavior.

To make the connection back to lumped-element transmission line model parameters (Sections 3.4 and 3.10), we simply divide by l to get the per-unit length parameter:

$$\boxed{C' = \frac{2\pi\epsilon_s}{\ln(b/a)}} \tag{5.24.3}$$

✔ Example 5.24.1: Capacitance of RG-59 coaxial cable

RG-59 coaxial cable consists of an inner conductor having radius 0.292 mm, an outer conductor having radius 1.855 mm, and a polyethylene spacing material having relative permittivity 2.25. Estimate the capacitance per length of RG-59.

Solution

From the problem statement, $a = 0.292$ mm, $b = 1.855$ mm, and $\epsilon_s = 2.25\epsilon_0$. Using Equation 5.24.3 we find $C' = 67.7$ pF/m.

This page titled 5.24: Capacitance of a Coaxial Structure is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.24: Capacitance of a Coaxial Structure** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

5.25: Electrostatic Energy

Consider a structure consisting of two perfect conductors, both fixed in position and separated by an ideal dielectric. This could be a capacitor, or it could be one of a variety of capacitive structures that are not explicitly intended to be a capacitor – for example, a printed circuit board. When a potential difference is applied between the two conducting regions, a positive charge Q_+ will appear on the surface of the conductor at the higher potential, and a negative charge $Q_- = -Q_+$ will appear on the surface of the conductor at the lower potential (Section 5.19). Assuming the conductors are not free to move, potential energy is stored in the electric field associated with the surface charges (Section 5.22).

We now ask the question, what is the energy stored in this field? The answer to this question has relevance in several engineering applications. For example, when capacitors are used as batteries, it is useful to know to amount of energy that can be stored. Also, any system that includes capacitors or has unintended capacitance is using some fraction of the energy delivered by the power supply to charge the associated structures. In many electronic systems – and in digital systems in particular – capacitances are periodically charged and subsequently discharged at a regular rate. Since power is energy per unit time, this cyclic charging and discharging of capacitors consumes power. Therefore, energy storage in capacitors contributes to the power consumption of modern electronic systems. We'll delve into that topic in more detail in Example 5.25.1.

Since capacitance C relates the charge Q_+ to the potential difference V between the conductors, this is the natural place to start. From the definition of capacitance (Section 5.22):

$$V = \frac{Q_+}{C}$$

From Section 5.8, electric potential is defined as the work done (i.e., energy injected) by moving a charged particle, per unit of charge; i.e.,

$$V = \frac{W_e}{q}$$

where q is the charge borne by the particle and W_e (units of J) is the work done by moving this particle across the potential difference V. Since we are dealing with charge distributions as opposed to charged particles, it is useful to express this in terms of the contribution ΔW_e made to W_e by a small charge Δq. Letting Δq approach zero we have

$$dW_e = V dq$$

Now consider what must happen to transition the system from having zero charge ($q = 0$) to the fully-charged but static condition ($q = Q_+$). This requires moving the differential amount of charge dq across the potential difference between conductors, beginning with $q = 0$ and continuing until $q = Q_+$. Therefore, the total amount of work done in this process is:

$$\begin{aligned} W_e &= \int_{q=0}^{Q+} dW_e \\ &= \int_0^{Q+} V dq \\ &= \int_0^{Q+} \frac{q}{C} dq \\ &= \frac{1}{2}\frac{Q_+^2}{C} \end{aligned} \tag{5.25.1}$$

Equation 5.25.1 can be expressed entirely in terms of electrical potential by noting again that $C = Q_+/V$, so

$$\boxed{W_e = \frac{1}{2} C V^2} \tag{5.25.2}$$

Since there are no other processes to account for the injected energy, the energy stored in the electric field is equal to W_e. Summarizing:

> The energy stored in the electric field of a capacitor (or a capacitive structure) is given by Equation 5.25.2.

Example 5.25.1: Why multicore computing is power-neutral

Readers are likely aware that computers increasingly use multicore processors as opposed to single-core processors. For our present purposes, a "core" is defined as the smallest combination of circuitry that performs independent computation. A multicore processor consists of multiple identical cores that run in parallel. Since a multicore processor consists of N identical processors, you might expect power consumption to increase by N relative to a single-core processor. However, this is not the case. To see why, first realize that the power consumption of a modern computing core is dominated by the energy required to continuously charge and discharge the multitude of capacitances within the core. From Equation 5.25.2, the required energy is $\frac{1}{2}C_0V_0^2$ per clock cycle, where C_0 is the sum capacitance (remember, capacitors in parallel add) and V_0 is the supply voltage. Power is energy per unit time, so the power consumption for a single core is

$$P_0 = \frac{1}{2}C_0V_0^2f_0$$

where f_0 is the clock frequency. In a N-core processor, the sum capacitance is increased by N. However, the frequency is decreased by N since the same amount of computation is (nominally) distributed among the N cores. Therefore, the power consumed by an N-core processor is

$$P_N = \frac{1}{2}\left(NC_0\right)V_0^2\left(\frac{f_0}{N}\right) = P_0$$

In other words, the increase in power associated with replication of hardware is nominally offset by the decrease in power enabled by reducing the clock rate. In yet other words, the total energy of the N-core processor is N times the energy of the single core processor at any given time; however, the multicore processor needs to recharge capacitances $1/N$ times as often.

Before moving on, it should be noted that the usual reason for pursuing a multicore design is to increase the amount of computation that can be done; i.e., to increase the product f_0N. Nevertheless, it is extremely helpful that power consumption is proportional to f_0 only, and is independent of N.

The thin parallel plate capacitor (Section 5.23) is representative of a large number of practical applications, so it is instructive to consider the implications of Equation 5.25.2 for this structure in particular. For the thin parallel plate capacitor,

$$C \approx \frac{\epsilon A}{d}$$

where A is the plate area, d is the separation between the plates, and ϵ is the permittivity of the material between the plates. This is an approximation because the fringing field is neglected; we shall proceed as if this is an exact expression. Applying Equation 5.25.2:

$$W_e = \frac{1}{2}\left(\frac{\epsilon A}{d}\right)\left(Ed\right)^2$$

where E is the magnitude of the electric field intensity between the plates. Rearranging factors, we obtain:

$$W_e = \frac{1}{2}\epsilon E^2\left(Ad\right)$$

Recall that the electric field intensity in the thin parallel plate capacitor is approximately uniform. Therefore, the density of energy stored in the capacitor is also approximately uniform. Noting that the product Ad is the volume of the capacitor, we find that the energy density is

$$w_e = \frac{W_e}{Ad} = \frac{1}{2}\epsilon E^2 \tag{5.25.3}$$

which has units of energy per unit volume (J/m^3).

The above expression provides an alternative method to compute the total electrostatic energy. Within a mathematical volume $\mathcal{V}$, the total electrostatic energy is simply the integral of the energy density over $\mathcal{V}$; i.e.,

$$W_e = \int_{\mathcal{V}} w_e \, dv$$

This works even if E and ϵ vary with position. So, even though we arrived at this result using the example of the thin parallel-plate capacitor, our findings at this point apply generally. Substituting Equation 5.25.3 we obtain:

$$\boxed{W_e = \frac{1}{2}\int_v \epsilon E^2 dv} \tag{5.25.4}$$

Summarizing:

> The energy stored by the electric field present within a volume is given by Equation 5.25.4.

It's worth noting that this energy increases with the permittivity of the medium, which makes sense since capacitance is proportional to permittivity.

This page titled 5.25: Electrostatic Energy is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **5.25: Electrostatic Energy** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

CHAPTER OVERVIEW

6: Steady Current and Conductivity

In practice, we deal with two physical mechanisms for current: convection and conduction. The distinction between these types of current is important in electromagnetic analysis. **Convection current** consists of charged particles moving in response to mechanical forces, as opposed to being guided by the electric field (Sections 2.2 and/or 5.1). An example of a convection current is a cloud bearing free electrons that moves through the atmosphere driven by wind. **Conduction current** consists of charged particles moving in response to the electric field and not merely being carried by motion of the surrounding material. In some materials, the electric field is also able to dislodge weakly-bound electrons from atoms, which then subsequently travel some distance before reassociating with other atoms. For this reason, the individual electrons in a conduction current do not necessarily travel the full distance over which the current is perceived to exist. The distinction between convection and conduction is important because Ohm's Law (Section 6.3) – which specifies the relationship between electric field intensity and current – applies only to conduction current.

6.1: Convection and Conduction Currents

6.2: Current Distributions

6.3: Conductivity

6.4: Resistance

6.5: Conductance

6.6: Power Dissipation in Conducting Media

Contributors and Attributions

- Ellingson, Steven W. (2018) Electromagnetics, Vol. 1. Blacksburg, VA: VT Publishing. https://doi.org/10.21061/electromagnetics-vol-1 Licensed with CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0. Report adoption of this book here. If you are a professor reviewing, adopting, or adapting this textbook please help us understand a little more about your use by filling out this form.

This page titled 6: Steady Current and Conductivity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

6.1: Convection and Conduction Currents

In practice, we deal with two physical mechanisms for current: convection and conduction. The distinction between these types of current is important in electromagnetic analysis.

- **Convection current** consists of charged particles moving in response to mechanical forces, as opposed to being guided by the electric field (Sections 2.2 and/or 5.1). An example of a convection current is a cloud bearing free electrons that moves through the atmosphere driven by wind.
- **Conduction current** consists of charged particles moving in response to the electric field and not merely being carried by motion of the surrounding material. In some materials, the electric field is also able to dislodge weakly-bound electrons from atoms, which then subsequently travel some distance before reassociating with other atoms. For this reason, the individual electrons in a conduction current do not necessarily travel the full distance over which the current is perceived to exist.

The distinction between convection and conduction is important because *Ohm's Law* (Section 6.3) – which specifies the relationship between electric field intensity and current – applies only to conduction current.

This page titled 6.1: Convection and Conduction Currents is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **6.1: Convection and Conduction Currents** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

6.2: Current Distributions

In elementary electric circuit theory, current is the rate at which electric charge passes a particular point in a circuit. For example, 1 A is 1 C per second. In this view current is a scalar quantity, and there are only two possible directions because charge is constrained to flow along defined channels. Direction is identified by the sign of the current with respect to a reference direction, which is in turn defined with respect to a reference voltage polarity. The convention in electrical engineering defines positive current as the flow of positive charge into the positive voltage terminal of a passive device such as a resistor, capacitor, or inductor. For an active device such as a battery, positive current corresponds to the flow of positive charge *out of* the positive voltage terminal.

However, this kind of thinking only holds up when the systems being considered are well-described as "lumped" devices connected by infinitesimally thin, filament-like connections representing wires and circuit board traces. Many important problems in electrical engineering concern situations in which the flow of current is not limited in this way. Examples include wire- and pin-type interconnects at radio frequencies, circuit board and enclosure grounding, and physical phenomena such as lightning. To accommodate the more general class of problems, we must define current as a vector quantity. Furthermore, current in these problems can spread out over surfaces and within volumes, so we must also consider *spatial distributions* of current.

As noted above, if a current I is constrained to follow a particular path, then the only other consideration is direction. Thus, a line current is specified mathematically as $\hat{\mathbf{l}}I$, where the direction $\hat{\mathbf{l}}$ may vary with position along the path. For example, in a straight wire $\hat{\mathbf{l}}$ is constant, whereas in a coil $\hat{\mathbf{l}}$ varies with position along the coil.

Surface Current Distribution

In some cases, current may be distributed over a surface. For example, the radio-frequency current on a wire of radius a made from a metal with sufficiently high conductivity can be modeled as a uniform surface current existing on the wire surface. In this case, the current is best described as a surface current density $\mathbf{J}_s$, which is the total current I on the wire divided by the circumference $2\pi a$ of the wire:

$$\mathbf{J}_s = \hat{\mathbf{u}}\frac{I}{2\pi a} \quad (\text{units of A/m})$$

where $\hat{\mathbf{u}}$ is the direction of current flow.

Volume Current Distribution

Imagine that current is distributed within a volume. Let $\hat{\mathbf{u}}\Delta i$ be the current passing through a small open planar surface defined within this volume, and let Δs be the area of this surface. The *volume current density* $\mathbf{J}$ at any point in the volume is defined as

$$\mathbf{J} \triangleq \lim_{\Delta s \to 0} \frac{\hat{\mathbf{u}}\Delta i}{\Delta s} = \hat{\mathbf{u}}\frac{di}{ds} \quad (\text{units of A/m}^2)$$

In general, $\mathbf{J}$ is a function of position within this volume. Subsequently the total current passing through a surface $\mathcal{S}$ is

$$I = \int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{s} \quad (\text{units of A})$$

In other words, volume current density integrated over a surface yields total current through that surface. You might recognize this as a calculation of flux, and it is (but it is uncommon to refer to net current as a flux).

✔ Example 6.2.1: Current and current density in a wire of circular cross-section

Figure 6.2.1 shows a straight wire having cross-sectional radius $a = 3$ mm. A battery is connected across the two ends of the wire resulting in a volume current density $\mathbf{J} = \hat{\mathbf{z}}8$ A/m^2, which is uniform throughout the wire. Find the net current I through the wire.

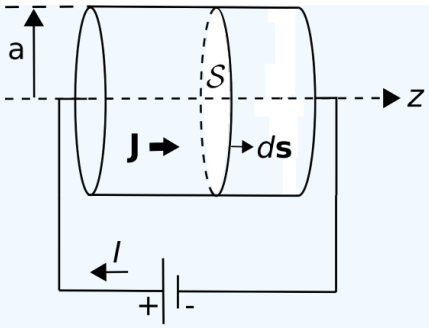


Figure 6.2.1: Net current I and current density J in a wire of circular cross-section. (CC BY SA 4.0; K. Kikkeri).

Solution

The net current is

$$I = \int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{s}$$

We choose $\mathcal{S}$ to be the cross-section perpendicular to the axis of the wire. Also, we choose $d\mathbf{s}$ to point such that I is positive with respect to the sign convention shown in Figure 6.2.1, which is the usual choice in electric circuit analysis. With these choices, we have

$$\begin{aligned} I &= \int_{\rho=0}^{a} \int_{\phi=0}^{2\pi} \left(\hat{\mathbf{z}} 8 \text{ A/m}^2\right) \cdot (\mathbf{z} \rho d\rho d\phi) \\ &= \left(8 \text{ A/m}^2\right) \int_{\rho=0}^{a} \int_{\phi=0}^{2\pi} \rho d\rho d\phi \\ &= \left(8 \text{ A/m}^2\right) \cdot \left(\pi a^2\right) \\ &= 226 \, \mu\mathbf{A} \end{aligned} \tag{6.2.1}$$

This answer is independent of the cross-sectional surface $\mathcal{S}$ used to do the calculation. For example, you could use an alternative surface that is tilted 45° away from the axis of the wire. In that case, the cross-sectional area would increase, but the dot product of $\mathbf{J}$ and $d\mathbf{s}$ would be proportionally less and the outcome would be the same. Since the choice of $\mathcal{S}$ is arbitrary – any surface with edges at the perimeter of the wire will do – you should make the choice that makes the problem as simple as possible, as we have done above.

This page titled 6.2: Current Distributions is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **6.2: Current Distributions** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

6.3: Conductivity

Conductivity is one of the three primary "constitutive parameters" that is commonly used to characterize the electromagnetic properties of materials (Section 2.8). The key idea is this:

> Conductivity is a property of materials that determines conduction current density in response to an applied electric field.

Recall that conduction current is the flow of charge in response to an electric field (Section 6.1). Although the associated force is straightforward to calculate (e.g., Section 5.1), the result is merely the force applied, not the speed at which charge moves in response. The latter is determined by the *mobility of charge*, which is in turn determined by the atomic and molecular structure of the material. Conductivity relates current density to the applied field directly, without requiring one to grapple separately with the issues of applied force and charge mobility.

In the absence of material – that is, in a true, perfect vacuum – conductivity is zero because there is no charge available to form current, and therefore the current is zero no matter what electric field is applied. At the other extreme, a *good conductor* is a material that contains a supply of charge that is able to move freely within the material. When an electric field is applied to a good conductor, charge-bearing material constituents move in the direction determined by the electric field, creating current flow in that direction. This relationship is summarized by *Ohm's Law for Electromagnetics*:

$$\boxed{\mathbf{J} = \sigma \mathbf{E}} \tag{6.3.1}$$

where $\mathbf{E}$ is electric field intensity (V/m); $\mathbf{J}$ is the volume current density, a vector describing the current flow, having units of A/m^2 (see Section 6.2); and σ is conductivity. Since $\mathbf{E}$ has units of V/m, we see σ has units of $\Omega^{-1}\mathrm{m}^{-1}$, which is more commonly expressed as S/m, where 1 S ("siemens") is defined as $1\ \Omega^{-1}$. Section A3 provides values of conductivity for a representative set of materials.

> Conductivity σ is expressed in units of S/m, where $1\ \mathrm{S} = 1\ \Omega^{-1}$.

It is important to note that the current being addressed here is *conduction current*, and not convection current, displacement current, or some other form of current – see Section 6.2 for elaboration. Summarizing:

Ohm's Law for Electromagnetics (Equation 6.3.1) states that volume density of conduction current (A/m^2) equals conductivity (S/m) times electric field intensity (V/m).

The reader is likely aware that there is also an "Ohm's Law" in electric circuit theory, which states that current I (units of A) is voltage V (units of V) divided by resistance R (units of Ω); i.e., $I = V/R$. This is in fact a special case of Equation 6.3.1; see Section 6.4 for more about this.

As mentioned above, σ depends on both the *availability* and *mobility* of charge within the material. At the two extremes, we have *perfect insulators*, for which $\sigma = 0$, and *perfect conductors*, for which $\sigma \to \infty$. Some materials approach these extremes, whereas others fall midway between these conditions. Here are a few classes of materials that are frequently encountered:

- A perfect vacuum – "free space" – contains no charge and therefore is a perfect insulator with $\sigma = 0$.
- *Good insulators* typically have conductivities $\ll 10^{-10}$ S/m, which is sufficiently low that the resulting currents can usually be ignored. The most important example is air, which has a conductivity only slightly greater than that of free space. An important class of good insulators is *lossless dielectrics*,[1] which are well-characterized in terms of permittivity (ϵ) alone, and for which $\mu = \mu_0$ and $\sigma = 0$ may usually be assumed.
- *Poor insulators* have conductivities that are low, but nevertheless sufficiently high that the resulting currents cannot be ignored. For example, the dielectric material that is used to separate the conductors in a transmission line must be considered a poor insulator as opposed to a good (effectively lossless) insulator in order to characterize loss per length along the transmission line, which can be significant.[2] These *lossy dielectrics* are well-characterized in terms of ϵ and σ, and typically $\mu = \mu_0$ can be assumed.
- *Semiconductors* such as those materials used in integrated circuits have intermediate conductivities, typically in the range 10^{-4} to 10^{+1} S/m.

- *Good conductors* are materials with very high conductivities, typically greater than 10^5 S/m. An important category of good conductors includes metals, with certain metals including alloys of aluminum, copper, and gold reaching conductivities on the order of 10^8 S/m. In such materials, minuscule electric fields give rise to large currents. There is no significant storage of energy in such materials, and so the concept of permittivity is not relevant for good conductors.

The reader should take care to note that terms such as *good conductor* and *poor insulator* are qualitative and subject to context. What may be considered a "good insulator" in one application may be considered to be a "poor insulator" in another.

One relevant category of material was not included in the above list – namely, *perfect conductors*. A perfect conductor is a material in which $\sigma \to \infty$. It is tempting to interpret this as meaning that any electric field gives rise to infinite current density; however, this is not plausible even in the ideal limit. Instead, this condition is interpreted as meaning that $\mathbf{E} = \mathbf{J}/\sigma \to 0$ throughout the material; i.e., $\mathbf{E}$ is zero independently of any current flow in the material. An important consequence is that the potential field V is equal to a constant value throughout the material. (Recall $\mathbf{E} = -\nabla V$ (Section 5.14), so constant V means $\mathbf{E} = 0$.) We refer to the volume occupied by such a material as an *equipotential volume*. This concept is useful as an approximation of the behavior of good conductors; for example, metals are often modeled as perfectly-conducting equipotential volumes in order to simplify analysis.

> A perfect conductor is a material for which $\sigma \to \infty$, $\mathbf{E} \to 0$, and subsequently V (the electric potential) is constant.

One final note: It is important to remain aware of the assumptions we have made about materials in this book, which are summarized in Section 2.8. In particular, we continue to assume that materials are linear unless otherwise indicated. For example, whereas air is normally considered to be a good insulator and therefore a poor conductor, anyone who has ever witnessed lightning has seen a demonstration that under the right conditions – i.e., sufficiently large potential difference between earth and sky – current will readily flow through air. This particular situation is known as *dielectric breakdown* (see Section 5.21). The non-linearity of materials can become evident before reaching the point of dielectric breakdown, so one must be careful to consider this possibility when dealing with strong electric fields.

This page titled 6.3: Conductivity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **6.3: Conductivity** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

6.4: Resistance

The concept of resistance is most likely familiar to readers via *Ohm's Law for Devices*; i.e., $V = IR$ where V is the potential difference associated with a current I. This is correct, but it is not the whole story. Let's begin with a statement of intent:

> Resistance R (Ω) is a characterization of the conductivity of a *device* (as opposed to a *material*) in terms of Ohms Law for Devices; i.e., $V = IR$.

Resistance is a property of devices such as resistors, which are intended to provide resistance, as well as being a property of most practical electronic devices, whether it is desired or not.

Resistance is a manifestation of the conductivity of the materials comprising the device, which subsequently leads to the "$V = IR$" relationship. This brings us to a very important point and a common source of confusion. *Resistance is not necessarily the real part of impedance.* Let's take a moment to elaborate. Impedance (Z) is defined as the ratio of voltage to current; i.e., V/I; or equivalently in the phasor domain as $\widetilde{V}/\tilde{I}$. Most devices – not just devices exhibiting resistance – can be characterized in terms of this ratio. Consider for example the input impedance of a terminated transmission line (Section 3.15). This impedance may have a non-zero real-valued component even when the transmission line and the terminating load are comprised of perfect conductors. Summarizing:

> Resistance results in a real-valued impedance. However, not all devices exhibiting a real-valued impedance exhibit resistance. Furthermore, the real component of a complex-valued impedance does not necessarily represent resistance.

Restating the main point in yet other words: *Resistance pertains to limited conductivity, not simply to voltage-current ratio.*

Also important to realize is that whereas conductivity σ (units of S/m) is a property of materials, resistance depends on *both* conductivity and the geometry of the device. In this section, we address the question of how the resistance of a device can be is determined. The following example serves this purpose.

Figure 6.4.1 shows a straight wire of length l centered on the z axis, forming a cylinder of material having conductivity σ and cross-sectional radius a. The ends of the cylinder are covered by perfectly-conducting plates to which the terminals are attached. A battery is connected to the terminals, resulting in a uniform internal electric field **E** between the plates.

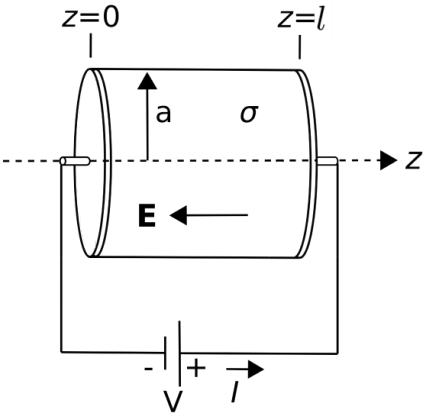


Figure 6.4.1: Analysis of the resistance of straight wire of circular cross-section. (CC BY SA 4.0; K.Kikkeri).

To calculate R for this device, let us first calculate V, then I, and finally $R = V/I$. First, we compute the potential difference using the following result from Section 5.8:

$$V = -\int_{\mathcal{C}} \mathbf{E}\cdot d\mathbf{l}$$

Here we can view V as a "given" (being the voltage of the battery), but wish to evaluate the right hand side so as to learn something about the effect of the conductivity and geometry of the wire. The appropriate choice of $\mathcal{C}$ begins at $z = 0$ and ends at $z = l$,

following the axis of the cylinder. (Remember: $\mathcal{C}$ defines the reference direction for increasing potential, so the resulting potential difference will be the node voltage at the end point minus the node voltage at the start point.) Thus, $d\mathbf{l} = \hat{\mathbf{z}}dz$ and we have

$$V = -\int_{z=0}^{l} \mathbf{E} \cdot (\hat{\mathbf{z}}dz)$$

We do not yet know $\mathbf{E}$; however, we know it is constant throughout the device and points in the $-\hat{\mathbf{z}}$ direction since this is the direction of current flow and Ohm's Law for Electromagnetics (Section 6.3) requires the electric field to point in the same direction. Thus, we may write $\mathbf{E} = -\hat{\mathbf{z}}E_z$ where E_z is a constant. We now find:

$$\begin{aligned} V &= -\int_{z=0}^{l} (-\hat{\mathbf{z}}E_z) \cdot (\hat{\mathbf{z}}dz) \\ &= +E_z \int_{z=0}^{l} dz \\ &= +E_z l \end{aligned} \tag{6.4.1}$$

The current I is given by (Section 6.2)

$$I = \int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{s}$$

We choose the surface $\mathcal{S}$ to be the cross-section perpendicular to the axis of the wire. Also, we choose $d\mathbf{s}$ to point such that I is positive with respect to the sign convention shown in Figure 6.4.1. With these choices, we have From Ohm's Law for Electromagnetics, we have

$$\mathbf{J} = \sigma\mathbf{E} = -\hat{\mathbf{z}}\sigma E_z$$

So now

$$\begin{aligned} I &= \int_{\rho=0}^{a} \int_{\phi=0}^{2\pi} (-\hat{\mathbf{z}}\sigma E_z) \cdot (-\hat{\mathbf{z}}\rho d\rho d\phi) \\ &= \sigma E_z \int_{\rho=0}^{a} \int_{\phi=0}^{2\pi} \rho d\rho d\phi \\ &= \sigma E_z \left(\pi a^2\right) \end{aligned}$$

Finally:

$$R = \frac{V}{I} = \frac{E_z l}{\sigma E_z \left(\pi a^2\right)} = \frac{l}{\sigma \left(\pi a^2\right)}$$

This is a good-enough answer for the problem posed, but it is easily generalized a bit further. Noting that πa^2 in the denominator is the cross-sectional area A of the wire, so we find:

$$\boxed{R = \frac{l}{\sigma A}} \tag{6.4.2}$$

i.e., the resistance of a wire having cross-sectional area A – regardless of the shape of the cross-section, is given by the above equation.

> The resistance of a right cylinder of material, given by Equation 6.4.2, is proportional to length and inversely proportional to cross-sectional area and conductivity.

It is important to remember that Equation 6.4.2 presumes that the volume current density $\mathbf{J}$ is uniform over the cross-section of the wire. This is an excellent approximation for thin wires at "low" frequencies including, of course, DC. At higher frequencies it may not be a good assumption that $\mathbf{J}$ is uniformly-distributed over the cross-section of the wire, and at sufficiently high frequencies one finds instead that the current is effectively limited to the exterior surface of the wire. In the "high frequency" case, A in Equation 6.4.2 is reduced from the physical area to a smaller value corresponding to the reduced area through which most of the current flows. Therefore, R increases with increasing frequency. To quantify the high frequency behavior of R (including determination of what constitutes "high frequency" in this context) one requires concepts beyond the theory of electrostatics, so this is addressed elsewhere.

Example 6.4.1:Resistance of 22 AWG hookup wire

A common type of wire found in DC applications is 22AWG (“American Wire Gauge”; see “Additional Resources” at the end of this section) copper solid-conductor hookup wire. This type of wire has circular cross-section with diameter 0.644 mm. What is the resistance of 3 m of this wire? Assume copper conductivity of 58 MS/m.

Solution

From the problem statement, the diameter $2a = 0.644$ mm, $\sigma = 58 \times 10^6$ S/m, and $l = 3$ m. The cross-sectional area is $A = \pi a^2 \cong 3.26 \times 10^{-7}$ m^2. Using Equation 6.4.2 we obtain $R = 159$ mΩ.

Example 6.4.2: Resistance of steel pipe

A pipe is 3 m long and has inner and outer radii of 5 mm and 7 mm respectively. It is made from steel having conductivity 4 MS/m. What is the DC resistance of this pipe?

Solution

We can use Equation 6.4.2 if we can determine the cross-sectional area A through which the current flows. This area is simply the area defined by the outer radius, πb^2, minus the area defined by the inner radius πa^2. Thus, $A = \pi b^2 - \pi a^2 \cong 7.54 \times 10^{-5}$ m^2. From the problem statement, we also determine that $\sigma = 4 \times 10^6$ S/m and $l = 3$ m. Using Equation 6.4.2 we obtain $R \cong 9.95$ mΩ.

This page titled 6.4: Resistance is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **6.4: Resistance** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

6.5: Conductance

Conductance, like resistance, is a property of devices. Specifically:

> Conductance G (Ω^{-1} or S) is the reciprocal of resistance R.

Therefore, conductance depends on both the conductivity of the materials used in the device, as well as the geometry of the device.

A natural question to ask is, why do we require the concept of conductance, if it simply the reciprocal of resistance? The short answer is that the concept of conductance is *not* required; or, rather, we need only resistance *or* conductance and not both. Nevertheless, the concept appears in engineering analysis for two reasons:

- Conductance is sometimes considered to be a more intuitive description of the underlying physics in cases where the applied voltage is considered to be the independent "stimulus" and current is considered to be the response. This is why conductance appears in the lumped element model for transmission lines (Section 3.4), for example.
- Characterization in terms of conductance may be preferred when considering the behavior of devices in parallel, since the conductance of a parallel combination is simply the sum of the conductances of the devices.

Example 6.5.1: Conductance of a coaxial structure

Let us now determine the conductance of a structure consisting of coaxially-arranged conductors separated by a lossy dielectric, as shown in Figure 6.5.1. The conductance per unit length G' (i.e., S/m) of this structure is of interest in determining the characteristic impedance of coaxial transmission line, as addressed in Sections 3.4 and Section 3.10.

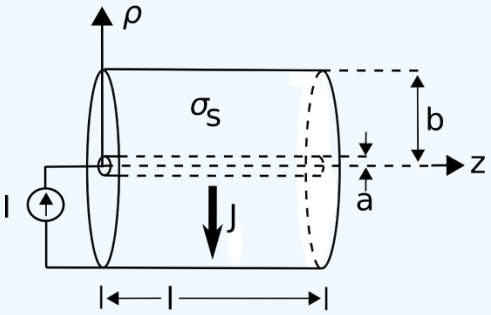


Figure 6.5.1: Determining the conductance of a structure consisting of coaxially-arranged conductors separated by a lossy dielectric.

For our present purposes, we may model the structure as two concentric perfectly-conducting cylinders of radii a and b, separated by a lossy dielectric having conductivity σ_s. We place the $+z$ axis along the common axis of the concentric cylinders so that the cylinders may be described as constant-coordinate surfaces $\rho = a$ and $\rho = b$.

There are at least 2 ways to solve this problem. One method is to follow the procedure that was used to find the capacitance of this structure in Section 5.24. Adapting that approach to the present problem, one would assume a potential difference V between the conductors, from that determine the resulting electric field intensity $\mathbf{E}$, and then using Ohm's Law for Electromagnetics (Section 6.3) determine the density $\mathbf{J} = \sigma_s \mathbf{E}$ of the current that leaks directly between conductors. From this, one is able to determine the total leakage current I, and subsequently the conductance $G \triangleq I/V$. Although highly recommended as an exercise for the student, in this section we take an alternative approach so as to demonstrate that there are a variety of approaches available for such problems.

The method we shall use below is as follows:

1. Assume a leakage current I between the conductors
2. Determine the associated current density $\mathbf{J}$, which is possible using only geometrical considerations
3. Determine the associated electric field intensity $\mathbf{E}$ using $\mathbf{J}/\sigma_s$
4. Integrate $\mathbf{E}$ over a path between the conductors to get V. Then, as before, conductance $G \triangleq I/V$.

The current I is defined as shown in Figure 6.5.1, with reference direction according to the engineering convention that positive current flows out of the positive terminal of a source. The associated current density must flow in the same direction, and the circular symmetry of the problem therefore constrains $\mathbf{J}$ to have the form

$$\mathbf{J} = \hat{\rho}\frac{I}{A}$$

where A is the area through which I flows. In other words, current flows radially outward from the inner conductor to the outer conductor, with density that diminishes inversely with the area through which the total current flows. (It may be helpful to view $\mathbf{J}$ as a flux density and I as a flux, as noted in Section 6.2.) This area is simply circumference $2\pi\rho$ times length l, so

$$\mathbf{J} = \hat{\rho}\frac{I}{2\pi\rho l}$$

which exhibits the correct units of A/m^2.

Now from Ohm's Law for Electromagnetics we find the electric field within the structure is

$$\mathbf{E} = \frac{\mathbf{J}}{\sigma_s} = \hat{\rho}\frac{I}{2\pi\rho l\sigma_s}$$

Next we get V using (Section 5.8)

$$V = -\int_{\mathcal{C}} \mathbf{E}\cdot d\mathbf{l}$$

where $\mathcal{C}$ is any path from the negatively-charged outer conductor to the positively-charged inner conductor. Since this can be *any* such path (Section 5.9), we should choose the simplest one. The simplest path is the one that traverses a radial of constant ϕ and z. Thus:

$$\begin{aligned} V &= -\int_{\rho=b}^{a}\left(\hat{\rho}\frac{I}{2\pi\rho l\sigma_s}\right)\cdot(\hat{\rho}d\rho) \\ &= -\frac{I}{2\pi l\sigma_s}\int_{\rho=b}^{a}\frac{d\rho}{\rho} \\ &= +\frac{I}{2\pi l\sigma_s}\int_{\rho=a}^{b}\frac{d\rho}{\rho} \\ &= +\frac{I}{2\pi l\sigma_s}\ln\left(\frac{b}{a}\right) \end{aligned}$$

Wrapping up:

$$G \triangleq \frac{I}{V} = \frac{I}{(I/2\pi l\sigma_s)\ln(b/a)}$$

Note that factors of I in the numerator and denominator cancel out, leaving:

$$\boxed{G = \frac{2\pi l\sigma_s}{\ln(b/a)}} \tag{6.5.1}$$

Note that Equation 6.5.1 is dimensionally correct, having units of $S = \Omega^{-1}$. Also note that this is expression depends only on materials (through σ_s) and geometry (through l, a, and b). Notably it does *not* depend on current or voltage, which would imply non-linear behavior.

To make the connection back to lumped-element transmission line model parameters (Section 3.4 and Section 3.10), we simply divide by l to get the per-length parameter:

$$\boxed{G' = \frac{2\pi\sigma_s}{\ln(b/a)}} \tag{6.5.2}$$

Example 6.5.2: Conductance of RG-59 coaxial cable

RG-59 coaxial cable consists of an inner conductor having radius 0.292 mm, an outer conductor having radius 1.855 mm, and a polyethylene spacing material exhibiting conductivity of about 5.9×10^{-5} S/m. Estimate the conductance per length of RG-59.

Solution

From the problem statement, $a = 0.292$ mm, $b = 1.855$ mm, and $\sigma_s \cong 5.9 \times 10^{-5}$ S/m. Using Equation 6.5.2, we find $G' \cong 200$ μS/m.

This page titled 6.5: Conductance is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **6.5: Conductance** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

6.6: Power Dissipation in Conducting Media

The displacement of charge in response to the force exerted by an electric field constitutes a reduction in the potential energy of the system (Section 5.8). If the charge is part of a steady current, there must be an associated loss of energy that occurs at a steady rate. Since power is energy per unit time, the loss of energy associated with current is expressible as power dissipation. In this section, we address two questions:

1. How much power is dissipated in this manner, and
2. What happens to the lost energy?

First, recall that *work* is force times distance traversed in response to that force (Section 5.8). Stated mathematically:

$$\Delta W = +\mathbf{F} \cdot \Delta \mathbf{l}$$

where the vector $\mathbf{F}$ is the force (units of N) exerted by the electric field, the vector $\Delta \mathbf{l}$ is the direction and distance (units of m) traversed, and ΔW is the work done (units of J) as a result. Note that a "+" has been explicitly indicated; this is to emphasize the distinction from the work being considered in Section 5.8. In that section, the work "ΔW" represented energy from an external source that was being used to increase the potential energy of the system by moving charge "upstream" relative to the electric field. Now, ΔW represents this internal energy as it is escaping from the system in the form of kinetic energy; therefore, positive ΔW now means a reduction in potential energy, hence the sign change (note: it is bad form to use the same variable to represent both tallies; nevertheless, it is common practice and so we simply remind the reader that it is important to be aware of the definitions of variables each time they are (re)introduced).

The associated power ΔP (units of W) is ΔW divided by the time Δt (units of s) required for the distance $\Delta \mathbf{l}$ to be traversed:

$$\Delta P = \frac{\Delta W}{\Delta t} = \mathbf{F} \cdot \frac{\Delta \mathbf{l}}{\Delta t}$$

Now we'd like to express force in terms of the electric field exerting this force. Recall that the force exerted by an electric field intensity $\mathbf{E}$ (units of V/m) on a particle bearing charge q (units of C) is $q\mathbf{E}$ (Section 2.2). However, we'd like to express this force in terms of a current, as opposed to a charge. An expression in terms of current can be constructed as follows. First, note that the total charge in a small volume "cell" is the volume charge density ρ_v (units of C/m^3) times the volume Δv of the cell; i.e., $q = \rho_v \Delta v$ (Section 5.3). Therefore:

$$\mathbf{F} = q\mathbf{E} = \rho_v \, \Delta v \, \mathbf{E}$$

and subsequently

$$\Delta P = \rho_v \, \Delta v \, \mathbf{E} \cdot \frac{\Delta \mathbf{l}}{\Delta t} = \mathbf{E} \cdot \left(\rho_v \frac{\Delta \mathbf{l}}{\Delta t} \right) \Delta v$$

The quantity in parentheses has units of C/m^3 $\cdot$ m $\cdot$ s^{-1}, which is A/m^2. Apparently this quantity is the volume current density $\mathbf{J}$, so we have

$$\Delta P = \mathbf{E} \cdot \mathbf{J} \, \Delta v$$

In the limit as $\Delta v \to 0$ we have

$$dP = \mathbf{E} \cdot \mathbf{J} \, dv$$

and integrating over the volume $\mathcal{V}$ of interest we obtain

$$P = \int_{\mathcal{V}} dP = \int_{\mathcal{V}} \mathbf{E} \cdot \mathbf{J} \, dv \tag{6.6.1}$$

The above expression is commonly known as *Joule's Law*. In our situation, it is convenient to use Ohm's Law for Electromagnetics ($\mathbf{J} = \sigma \mathbf{E}$; Section 6.3) to get everything in terms of materials properties (σ), geometry ($\mathcal{V}$), and the electric field:

$$P = \int_{\mathcal{V}} \mathbf{E} \cdot (\sigma \mathbf{E}) \, dv$$

which is simply

$$\boxed{P = \int \sigma \, |\mathbf{E}|^2 \, dv} \tag{6.6.2}$$

$J_\mathcal{V}$

Thus:

> The power dissipation associated with current is given by Equation 6.6.2. This power is proportional to conductivity and proportional to the electric field magnitude squared.

This result facilitates the analysis of power dissipation in materials exhibiting loss; i.e., having finite conductivity. But what is the power dissipation in a perfectly conducting material? For such a material, $\sigma \to \infty$ and $\mathbf{E} \to 0$ no matter how much current is applied (Section 6.3). In this case, Equation 6.6.2 is not very helpful. However, as we just noted, being a perfect conductor means $\mathbf{E} \to 0$ no matter how much current is applied, so from Equation 6.6.1 we have found that:

> No power is dissipated in a perfect conductor.

When conductivity is finite, Equation 6.6.2 serves as a more-general version of a concept from elementary circuit theory, as we shall now demonstrate. Let $\mathbf{E} = \hat{\mathbf{z}} E_z$, so $|\mathbf{E}|^2 = E_z^2$. Then Equation 6.6.2 becomes:

$$P = \int_\mathcal{V} \sigma E_z^2 \, dv = \sigma E_z^2 \int_\mathcal{V} dv \tag{6.6.3}$$

The second integral in Equation 6.6.3 is a calculation of volume. Let's assume $\mathcal{V}$ is a cylinder aligned along the z axis. The volume of this cylinder is the cross-sectional area A times the length l. Then the above equation becomes:

$$P = \sigma E_z^2 \, A \, l$$

For reasons that will become apparent very shortly, let's reorganize the above expression as follows:

$$P = (\sigma E_z A)\,(E_z l)$$

Note that σE_z is the current density in A/m^2, which when multiplied by A gives the total current. Therefore, the quantity in the first set of parentheses is simply the current I. Also note that $E_z l$ is the potential difference over the length l, which is simply the node-to-node voltage V (Section 5.8). Therefore, we have found:

$$P = IV$$

as expected from elementary circuit theory.

Now, what happens to the energy associated with this dissipation of power? The displacement of charge carriers in the material is limited by the conductivity, which itself is finite because, simply put, other constituents of the material get in the way. If charge is being displaced as described in this section, then energy is being used to displace the charge-bearing particles. The motion of constituent particles is observed as heat – in fact, this is essentially the definition of heat. Therefore:

> The power dissipation associated with the flow of current in any material that is not a perfect conductor manifests as heat.

This phenomenon is known as *joule heating*, *ohmic heating*, and by other names. This conversion of electrical energy to heat is the method of operation for toasters, electric space heaters, and many other devices that generate heat. It is of course also the reason that all practical electronic devices generate heat.

This page titled 6.6: Power Dissipation in Conducting Media is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **6.6: Power Dissipation in Conducting Media** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

CHAPTER OVERVIEW

7: Magnetostatics

Magnetostatics is the theory of the magnetic field in conditions in which its behavior is independent of electric fields, including

- The magnetic field associated with various spatial distributions of steady current
- The energy associated with the magnetic field
- Inductance, which is the ability of a structure to store energy in a magnetic field

The word ending "-statics" refers to the fact that these aspects of electromagnetic theory can be developed by assuming that the sources of the magnetic field are time-invariant; we might say that magnetostatics is the study of the magnetic field at DC. However, many aspects of magnetostatics are applicable at "AC" as well.

7.1: Comparison of Electrostatics and Magnetostatics
7.2: Gauss' Law for Magnetic Fields - Integral Form
7.3: Gauss' Law for Magnetism - Differential Form
7.4: Ampere's Circuital Law (Magnetostatics) - Integral Form
7.5: Magnetic Field of an Infinitely-Long Straight Current-Bearing Wire
7.6: Magnetic Field Inside a Straight Coil
7.7: Magnetic Field of a Toroidal Coil
7.8: Magnetic Field of an Infinite Current Sheet
7.9: Ampere's Law (Magnetostatics) - Differential Form
7.10: Boundary Conditions on the Magnetic Flux Density (B)
7.11: Boundary Conditions on the Magnetic Field Intensity (H)
7.12: Inductance
7.13: Inductance of a Straight Coil
7.14: Inductance of a Coaxial Structure
7.15: Magnetic Energy
7.16: Magnetic Materials

Contributors and Attributions

- Ellingson, Steven W. (2018) Electromagnetics, Vol. 1. Blacksburg, VA: VT Publishing. https://doi.org/10.21061/electromagnetics-vol-1 Licensed with CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0. Report adoption of this book here. If you are a professor reviewing, adopting, or adapting this textbook please help us understand a little more about your use by filling out this form.

This page titled 7: Magnetostatics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

7.1: Comparison of Electrostatics and Magnetostatics

Students encountering magnetostatics for the first time have usually been exposed to electrostatics already. Electrostatics and magnetostatics exhibit many similarities. These are summarized in Table 7.1.1. The elements of magnetostatics presented in this table are all formally introduced in other sections; the sole purpose of this table is to point out the similarities. The technical term for these similarities is *duality*. Duality also exists between voltage and current in electrical circuit theory.

Table 7.1.1: A summary of the duality between electrostatics and magnetostatics

	electrostatics	**magnetostatics**
Sources	static charge	steady current, magnetizable material
Field intensity	$\mathbf{E}$ (V/m)	$\mathbf{H}$ (A/m)
Flux density	$\mathbf{D}$ (C/m^2)	$\mathbf{B}$ (Wb/m^2=T)
Material relations	$\mathbf{D}=\epsilon\mathbf{E}$	$\mathbf{B}=\mu\mathbf{H}$
	$\mathbf{J}=\sigma\mathbf{E}$	
Force on charge q	$\mathbf{F}=q\mathbf{E}$	$\mathbf{F}=q\mathbf{v}\times\mathbf{B}$
Maxwell's Eqs. (integral)	$\oint_S \mathbf{D}\cdot d\mathbf{s}=Q_{encl}$	$\oint_S \mathbf{B}\cdot d\mathbf{s}=0$
	$\oint_C \mathbf{E}\cdot d\mathbf{l}=0$	$\oint_C \mathbf{H}\cdot d\mathbf{l}=I_{encl}$
Maxwell's Eqs. (differential)	$\nabla\cdot\mathbf{D}=\rho_v$	$\nabla\cdot\mathbf{B}=0$
	$\nabla\times\mathbf{E}=0$	$\nabla\times\mathbf{H}=\mathbf{J}$
Boundary Conditions	$\hat{\mathbf{n}}\times[\mathbf{E}_1-\mathbf{E}_2]=0$	$\hat{\mathbf{n}}\times[\mathbf{H}_1-\mathbf{H}_2]=\mathbf{J}_s$
	$\hat{\mathbf{n}}\cdot[\mathbf{D}_1-\mathbf{D}_2]=\rho_s$	$\hat{\mathbf{n}}\cdot[\mathbf{B}_1-\mathbf{B}_2]=0$
Energy storage	Capacitance	Inductance
Energy density	$w_e=\frac{1}{2}\epsilon E^2$	$w_m=\frac{1}{2}\mu H^2$
Energy dissipation	Resistance	

This page titled 7.1: Comparison of Electrostatics and Magnetostatics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.1: Comparison of Electrostatics and Magnetostatics** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.2: Gauss' Law for Magnetic Fields - Integral Form

Gauss' Law for Magnetic Fields (GLM) is one of the four fundamental laws of classical electromagnetics, collectively known as *Maxwell's Equations*. Before diving in, the reader is strongly encouraged to review Section Section 2.5. In that section, GLM emerges from the "flux density" interpretation of the magnetic field. GLM is not identified in that section, but now we are ready for an explicit statement:

> *Gauss' Law for Magnetic Fields* (Equation 7.2.1) states that the flux of the magnetic field through a closed surface is zero.

This is expressed mathematically as follows:

$$\boxed{\oint_{\mathcal{S}} \mathbf{B}\cdot d\mathbf{s} = 0} \tag{7.2.1}$$

where $\mathbf{B}$ is magnetic flux density and $\mathcal{S}$ is a closed surface with outward-pointing differential surface normal $d\mathbf{s}$. It may be useful to consider the units. $\mathbf{B}$ has units of Wb/m^2; therefore, integrating $\mathbf{B}$ over a surface gives a quantity with units of Wb, which is magnetic flux, as indicated above.

GLM can also be interpreted in terms of magnetic field lines. For the magnetic flux through a closed surface to be zero, every field line entering the volume enclosed by $\mathcal{S}$ must also exit this volume – field lines may not begin or end within the volume. The only way this can be true for every possible surface $\mathcal{S}$ is if magnetic field lines always form closed loops. This is in fact what we find in practice, as shown in Figure 7.2.1.

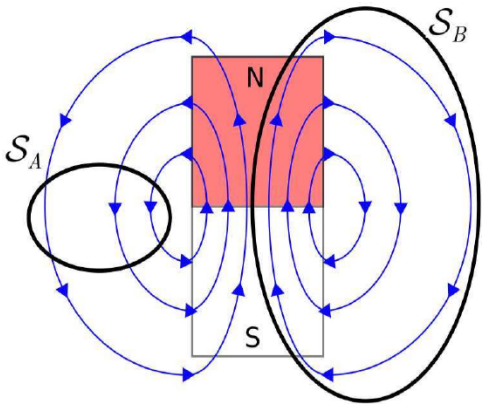


Figure 7.2.1: Gauss' Law for Magnetostatics applied to a two-dimensional bar magnet. For the surface $\mathcal{S} = \mathcal{S}_A$, every field line entering $\mathcal{S}$ also leaves $\mathcal{S}$, so the flux through $\mathcal{S}$ is zero. For the surface $\mathcal{S} = \mathcal{S}_B$, every field line within $\mathcal{S}$ remains in $\mathcal{S}$, so the flux through $\mathcal{S}$ is again zero. Image used wtih permission (CC BY SA 4.0; K. Kikkeri)

Following this argument one step further, GLM implies there can be no particular particle or structure that can be the source of the magnetic field (because then that would be a start point for field lines). This is one way in which the magnetic field is very different from the electrostatic field, for which every field line begins at a charged particle. So, when we say that current (for example) is the source of the magnetic field, we mean only that the field coexists with current, and *not* that the magnetic field is somehow attached to the current. Summarizing, there is no "localizable" quantity, analogous to charge for electric fields, associated with magnetic fields. This is just another way in which magnetic fields are weird!

Summarizing

> Gauss' Law for Magnetic Fields requires that magnetic field lines form closed loops. Furthermore, there is no particle that can be identified as the source of the magnetic field.

This page titled 7.2: Gauss' Law for Magnetic Fields - Integral Form is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.2: Gauss' Law for Magnetic Fields - Integral Form** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.3: Gauss' Law for Magnetism - Differential Form

The integral form of Gauss' Law states that the magnetic flux through a closed surface is zero. In mathematical form:

$$\oint_{\mathcal{S}} \mathbf{B}\cdot d\mathbf{s} = 0 \tag{7.3.1}$$

where $\mathbf{B}$ is magnetic flux density and $\mathcal{S}$ is the enclosing surface. Just as Gauss's Law for electrostatics has both integral and differential forms, so too does Gauss' Law for Magnetic Fields. Here we are interested in the differential form for the same reason. Given a differential equation and the boundary conditions imposed by structure and materials, we may then solve for the magnetic field in very complicated scenarios.

The equation we seek may be obtained from Equation 7.3.1 using the Divergence Theorem, which in the present case may be written:

$$\int_{\mathcal{V}} (\nabla \cdot \mathbf{B})\, dv = \oint_{\mathcal{S}} \mathbf{B}\cdot d\mathbf{s}$$

Where $\mathcal{V}$ is the mathematical volume bounded by the closed surface $\mathcal{S}$. From Equation 7.3.1 we see that the right hand side of the equation is zero, leaving:

$$\int_{\mathcal{V}} (\nabla \cdot \mathbf{B})\, dv = 0$$

The above relationship must hold regardless of the specific location or shape of $\mathcal{V}$. The only way this is possible is if the integrand is everywhere equal to zero. We conclude:

$$\boxed{\nabla \cdot \mathbf{B} = 0} \tag{7.3.2}$$

> The differential ("point") form of Gauss' Law for Magnetic Fields (Equation 7.3.2) states that the flux per unit volume of the magnetic field is always zero.

This is another way of saying that there is no point in space that can be considered to be the source of the magnetic field, for if it were, then the total flux through a bounding surface would be greater than zero. Said yet another way, the source of the magnetic field is not localizable.

This page titled 7.3: Gauss' Law for Magnetism - Differential Form is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.3: Gauss' Law for Magnetism - Differential Form** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.4: Ampere's Circuital Law (Magnetostatics) - Integral Form

Ampere's circuital law (ACL) relates current to the magnetic field associated with the current. In the magnetostatic regime, the law is (see also Figure 7.4.1):

$$\boxed{\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_{encl}} \tag{7.4.1}$$

That is, the integral of the magnetic field intensity $\mathbf{H}$ over a closed path $\mathcal{C}$ is equal to the current enclosed by that path, I_{encl}. Before proceeding to interpret this law, it is useful to see that it is dimensionally correct. That is, $\mathbf{H}$ (units of A/m) integrated over a distance (units of m) yields a quantity with units of current (i.e., A).

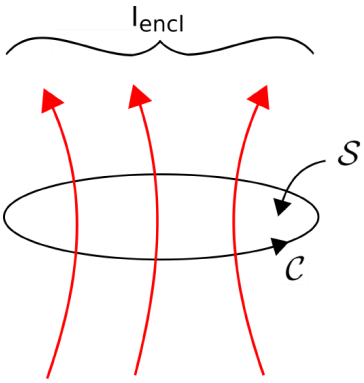


Figure 7.4.1: Reference directions for Ampere's Circuital Law (Equation 7.4.1). (CC BY SA 4.0; K. Kikkeri)

In general, I_{encl} may be either positive or negative. The direction corresponding to positive current flow must be correctly associated with $\mathcal{C}$. The relationship follows the *right-hand rule* of Stokes' Theorem summarized as follows. The direction of positive I_{encl} is the direction in which the fingers of the right hand intersect any surface $\mathcal{S}$ bordered by $\mathcal{C}$ when the thumb of the right hand points in the direction of integration. The connection to Stokes' Theorem is not a coincidence. See Section 7.9 for more about this.

Note that $\mathcal{S}$ can be *any* surface that is bounded by $\mathcal{C}$ – not just the taut surface implied in Figure 7.4.1.

> The integral form of Ampere's Circuital Law for magnetostatics (Equation 7.4.1) relates the magnetic field along a closed path to the total current flowing through any surface bounded by that path.

ACL plays a role in magnetostatics that is very similar to the role played by the integral form of Gauss' Law in electrostatics (Section 5.5). That is, ACL provides a means calculate the magnetic field given the source current. ACL also has a similar limitation. Generally, symmetry is required to simplify the problem sufficiently so that the integral equation may be solved. Fortunately, a number of important problems fall in this category. Examples include the problems addressed in Sections 7.5 (magnetic field of a line current), 7.6 (magnetic field inside a straight coil), 7.7 (magnetic field of a toroidal coil), 7.8 (magnetic field of a current sheet), and 7.11(boundary conditions on the magnetic field intensity). For problems in which the necessary symmetry is not available, the differential form of ACL may be required (Section 7.9).

Finally, be aware that the form of ACL addressed here applies to magnetostatics only. In the presence of a *time-varying* electric field, the right side of ACL includes an additional term known as the *displacement current* (Section 8.9).

This page titled 7.4: Ampere's Circuital Law (Magnetostatics) - Integral Form is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.4: Ampere's Circuital Law (Magnetostatics) - Integral Form** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.5: Magnetic Field of an Infinitely-Long Straight Current-Bearing Wire

In this section, we use the magnetostatic form of Ampere's Circuital Law (ACL) to determine the magnetic field due to a steady current I (units of A) in an infinitely-long straight wire. The problem is illustrated in Figure 7.5.1. The wire is an electrically-conducting circular cylinder of radius a. Since the wire is a cylinder, the problem is easiest to work in cylindrical coordinates with the wire aligned along the z axis.

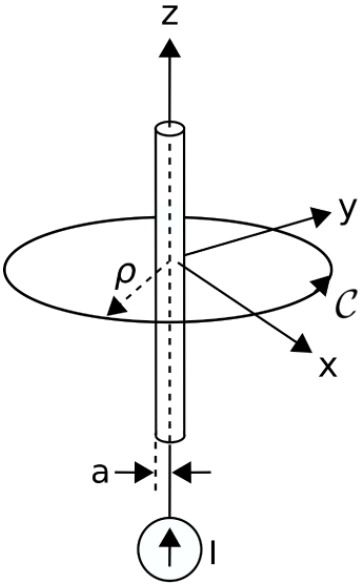


Figure 7.5.1: Determination of the magnetic field due to steady current in an infinitely-long straight wire. (CC BY SA 4.0; K. Kikkeri).

Here's the relevant form of ACL:

$$\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_{encl} \tag{7.5.1}$$

where I_{encl} is the current enclosed by the closed path $\mathcal{C}$. ACL works for *any* closed path, so to exploit the symmetry of the cylindrical coordinate system we choose a circular path of radius ρ in the $z = 0$ plane, centered at the origin. With this choice we have

$$I_{encl} = I \quad \text{for } \rho \geq a$$

For $\rho < a$, we see that $I_{encl} < I$. a steady (DC) current will be distributed uniformly throughout the wire (Section 6.4). Since the current is uniformly distributed over the cross section, I_{encl} is less than the total current I by the same factor that the area enclosed by $\mathcal{C}$ is less than πa^2, the cross-sectional area of the wire. The area enclosed by $\mathcal{C}$ is simply $\pi\rho^2$, so we have

$$I_{encl} = I\frac{\pi\rho^2}{\pi a^2} = I\frac{\rho^2}{a^2} \quad \text{for } \rho < a$$

For the choice of $\mathcal{C}$ made above, Equation 7.5.1 becomes

$$\int_{\phi=0}^{2\pi} \mathbf{H} \cdot \left(\hat{\phi}\ \rho\ d\phi\right) = I_{encl} \tag{7.5.2}$$

Note that we have chosen to integrate in the $+\phi$ direction. Therefore, the right-hand rule specifies that positive I_{encl} corresponds to current flowing in the $+z$ direction, which is consistent with the direction indicated in Figure 7.5.1. (Here's an excellent exercise to test your understanding. Change the direction of the path of integration and confirm that you get the same result obtained at the end of this section. Changing the direction of integration should not change the magnetic field associated with the current!)

The simplest way to solve for $\mathbf{H}$ from Equation 7.5.2 is to use a symmetry argument, which proceeds as follows:

- Since the distribution of current is uniform and infinite in the z-dimension, $\mathbf{H}$ can't depend on z, and so $\mathbf{H} \cdot \hat{\mathbf{z}}$ must be zero everywhere.

- The problem is identical after any amount of rotation in ϕ; therefore, the magnitude of $\mathbf{H}$ cannot depend on ϕ. This is a form of radial symmetry. Since we determined above that $\mathbf{H}$ can't depend on z either, it must be that the magnitude of $\mathbf{H}$ can depend only on ρ.
- The radial symmetry of the problem also requires that $\mathbf{H} \cdot \hat{\rho}$ be equal to zero. If this were not the case, then the field would not be radially symmetric. Since we determined above that $\mathbf{H} \cdot \hat{\mathbf{z}}$ is also zero, $\mathbf{H}$ must be entirely $\pm\hat{\phi}$-directed.

From the above considerations, the most general form of the magnetic field intensity can be written $\mathbf{H} = \hat{\phi}H(\rho)$. Substituting this into Equation 7.5.2, we obtain

$$\begin{aligned} I_{encl} &= \int_{\phi=0}^{2\pi} [\hat{\phi}H(\rho)] \cdot (\hat{\phi}\rho d\phi) \\ &= \rho H(\rho) \int_{\phi=0}^{2\pi} d\rho \\ &= 2\pi\rho H(\rho) \end{aligned} \tag{7.5.3}$$

Therefore, $H(\rho) = I_{encl}/2\pi\rho$. Reassociating the known direction, we obtain:

$$\mathbf{H} = \hat{\phi}\frac{I_{encl}}{2\pi\rho}$$

Therefore, the field outside of the wire is:

$$\boxed{\mathbf{H} = \hat{\phi}\frac{I}{2\pi\rho} \quad \text{for } \rho \geq a} \tag{7.5.4}$$

whereas the field inside the wire is:

$$\boxed{\mathbf{H} = \hat{\phi}\frac{I\rho}{2\pi a^2} \quad \text{for } \rho < a} \tag{7.5.5}$$

(By the way, this is a good time for a units check.)

Note that as ρ increases from zero to a (i.e., inside the wire), the magnetic field is proportional to ρ and therefore increases. However, as ρ continues to increase beyond a (i.e., outside the wire), the magnetic field is proportional to ρ^{-1} and therefore decreases.

If desired, the associated magnetic flux density can be obtained using $\mathbf{B} = \mu\mathbf{H}$.

Summarizing

The magnetic field due to current in an infinite straight wire is given by Equations [m0119_eACLLCe] (outside the wire) and [m0119_eACLLCi] (inside the wire). The magnetic field is $+\hat{\phi}$-directed for current flowing in the $+z$ direction, so the magnetic field lines form concentric circles perpendicular to and centered on the wire.

Finally, we point out another "right-hand rule" that emerges from this solution, shown in Figure 7.5.2 and summarized below:

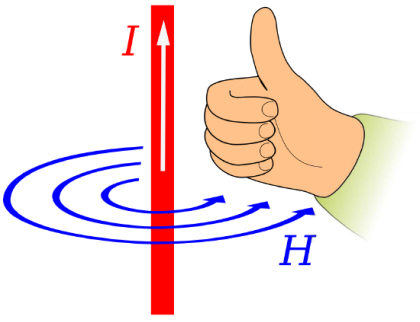


Figure 7.5.2: Right-hand rule for the relationship between the direction of current and the direction of the magnetic field. (CC BY SA 4.0 (modified); Jfmelero).

> The magnetic field due to current in an infinite straight wire points in the direction of the curled fingers of the right hand when the thumb of the right hand is aligned in the direction of current flow.

This simple rule turns out to be handy in quickly determining the relationship between the directions of the magnetic field and current flow in many other problems, and so is well worth committing to memory.

This page titled 7.5: Magnetic Field of an Infinitely-Long Straight Current-Bearing Wire is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.5: Magnetic Field of an Infinitely-Long Straight Current-Bearing Wire** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.6: Magnetic Field Inside a Straight Coil

In this section, we use the magnetostatic integral form of Ampere's Circuital Law (ACL) to determine the magnetic field inside a straight coil of the type shown in Figure 7.6.1 in response to a steady (i.e., DC) current. The result has a number of applications, including the analysis and design of inductors, solenoids (coils that are used as magnets, typically as part of an actuator), and as a building block and source of insight for more complex problems.

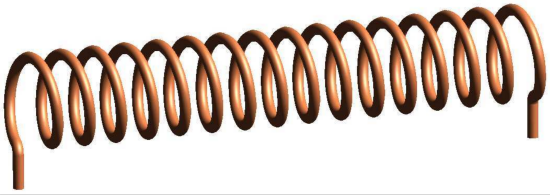


Figure 7.6.1: A straight coil. (public domain; Zureks).

The present problem is illustrated in Figure 7.6.2. The coil is circular with radius a and length l, and consists of N turns ("windings") of wire wound with uniform winding density. Since the coil forms a cylinder, the problem is easiest to work in cylindrical coordinates with the axis of the coil aligned along the z axis.

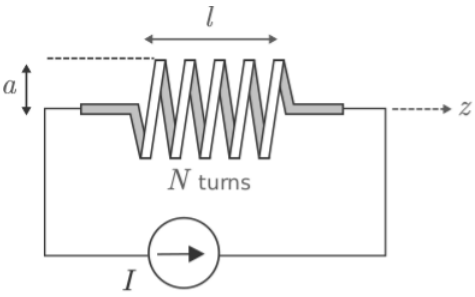


Figure 7.6.2: Determination of the magnetic field due to DC current in a coil.

To begin, let's take stock of what we already know about the answer, which is actually quite a bit. The magnetic field deep inside the coil is generally aligned with axis of the coil as shown in Figure 7.6.3. This can be explained using the result for the magnetic field due to a straight line current (Section 7.5), in which we found that the magnetic field follows a "right-hand rule." The magnetic field points in the direction of the fingers of the right hand when the thumb of the right hand is aligned in the direction of current flow. The wire comprising the coil is obviously *not* straight, but we can consider just one short segment of one turn and then sum the results for all such segments. When we consider this for a single turn of the coil, the situation is as shown in Figure 7.6.4. Summing the results for many loops, we see that the direction of the magnetic field inside the coil must be generally in the $+\hat{\mathbf{z}}$ direction when the current I is applied as shown in Figures 7.6.2 and 7.6.3. However, there is one caveat. The windings must be sufficiently closely-spaced that the magnetic field lines can only pass through the openings at the end of the coil and do not take any "shortcuts" between individual windings.

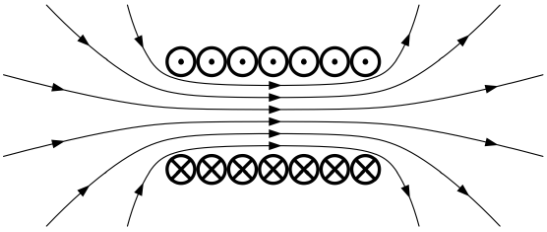


Figure 7.6.3: Magnetic field lines inside a straight coil with closely-spaced windings. (Dotted circles represent current flowing up/out from the page; crossed circles represent current flowing down/into the

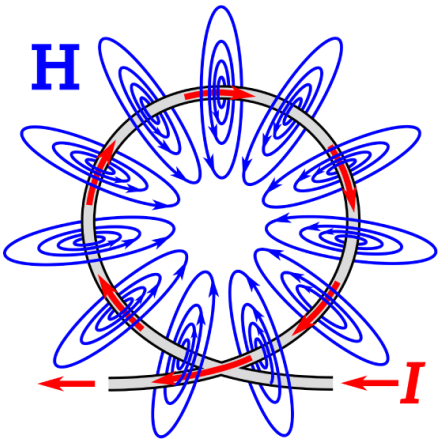


Figure 7.6.4: Magnetic field due to a single loop (© CC0 1.0 (modified); Chetvorno)

page.) (© CC BY SA 3.0; Geek3)

Figure 7.6.3 also indicates that the magnetic field lines near the ends of the coil diverge from the axis of the coil. This is understandable since magnetic field lines form closed loops. The relatively complex structure of the magnetic fields near the ends of the coil (the "fringing field") and outside of the coil make them relatively difficult to analyze. Therefore, here we shall restrict our attention to the magnetic field deep inside the coil. This restriction turns out to be of little consequence as long as $l \gg a$.

Also, it is apparent from the radial symmetry of the coil that the magnitude of the magnetic field cannot depend on ϕ. Putting these findings together, we find that the most general form for the magnetic field intensity deep inside the coil is $\mathbf{H} \approx \hat{\mathbf{z}} H(\rho)$. That is, the direction of $\mathbf{H}$ is $\pm\hat{\mathbf{z}}$ and the magnitude of $\mathbf{H}$ depends, at most, on ρ. In fact, we will soon find with the assistance of ACL that the magnitude of $\mathbf{H}$ doesn't depend on ρ either.

Here's the relevant form of ACL:

$$\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_{encl} \tag{7.6.1}$$

where I_{encl} is the current enclosed by the closed path $\mathcal{C}$. ACL works for *any* closed path that encloses the current of interest. Also, for simplicity, we prefer a path that lies on a constant-coordinate surface. The selected path is shown in Figure 7.6.5. The benefits of this particular path will soon become apparent. However, note for now that this particular choice is consistent with the right-hand rule relating the direction of $\mathcal{C}$ to the direction of positive I. That is, when I is positive, the current in the turns of the coil pass through the surface bounded by $\mathcal{C}$ in the same direction as the fingers of the right hand when the thumb is aligned in the indicated direction of $\mathcal{C}$.

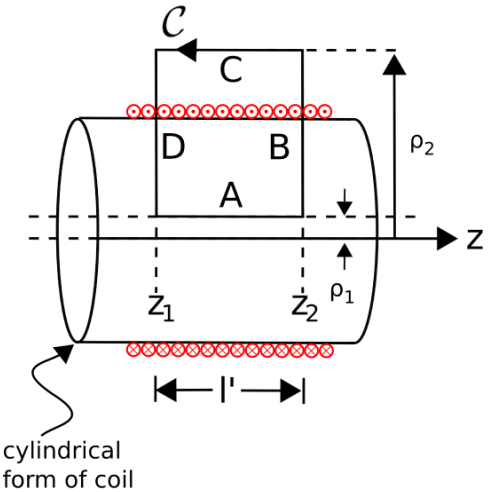


Figure 7.6.5: Selected path of integration. (© CC BY SA 4.0: K. Kikkeri)

Let's define N to be the number of windings in the coil. Then, the winding density of the coil is N/l (turns/m). Let the path length in the z direction be l', as indicated in Figure 7.6.5. Then the enclosed current is

$$I_{encl} = \frac{N}{l} l' I$$

That is, the number of turns per unit length times length gives number of turns, and this quantity times the current through the wire is the total amount of current crossing the surface bounded by $\mathcal{C}$.

For the choice of $\mathcal{C}$ made above, and taking our approximation for the form of $\mathbf{H}$ as exact, Equation 7.6.1 becomes

$$\oint_{\mathcal{C}} [\hat{\mathbf{z}} H(\rho)] \cdot d\mathbf{l} = \frac{N}{l} l' I \tag{7.6.2}$$

The integral consists of segments A, B, C, and D, as shown in Figure 7.6.5. Let us consider the result for each of these segments individually:

- The integral over segments B and D is zero because $d\mathbf{l} = \hat{\rho} d\rho$ for these segments, and so $\mathbf{H} \cdot d\mathbf{l} = 0$ for these segments.
- It is also possible to make the contribution from Segment C go to zero simply by letting $\rho_2 \to \infty$. The argument is as follows. The magnitude of $\mathbf{H}$ outside the coil must decrease with distance from the coil, so for ρ sufficiently large, $\mathbf{H}(\rho)$ becomes negligible. If that's the case, then the integral over Segment C also becomes negligible.

With $\rho_2 \to \infty$, only Segment A contributes significantly to the integral over $\mathcal{C}$ and Equation 7.6.2 becomes:

$$\begin{aligned} \frac{N}{l} l' I &= \int_{z_1}^{z_2} [\hat{\mathbf{z}} H(\rho_1)] \cdot (\hat{\mathbf{z}} dz) \\ &= H(\rho_1) \int_{z_1}^{z_2} dz \\ &= H(\rho_1) [z_2 - z_1] \end{aligned} \tag{7.6.3}$$

Note $z_2 - z_1$ is simply l'. Also, we have found that the result is independent of ρ_1, as anticipated earlier. Summarizing:

$$\boxed{\mathbf{H} \approx \hat{\mathbf{z}} \frac{NI}{l} \quad \text{inside coil}} \tag{7.6.4}$$

Let's take a moment to consider the implications of this remarkably simple result.

- Note that it is dimensionally correct; that is, current divided by length gives units of A/m, which are the units of $\mathbf{H}$.
- We have found that the magnetic field is simply winding density (N/l) times current. To increase the magnetic field, you can either use more turns per unit length or increase the current.
- We have found that the magnetic field is uniform inside the coil; that is, the magnetic field along the axis is equal to the magnetic field close to the cylinder wall formed by the coil. However, this does not apply close to ends of the coil, since we have neglected the fringing field.

These findings have useful applications in more complicated and practical problems, so it is worthwhile taking note of these now. Summarizing:

> The magnetic field deep inside an ideal straight coil (Equation 7.6.4) is uniform and proportional to winding density and current.

This page titled 7.6: Magnetic Field Inside a Straight Coil is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.6: Magnetic Field Inside a Straight Coil** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.7: Magnetic Field of a Toroidal Coil

A toroid is a cylinder in which the ends are joined to form a closed loop. An example of a toroidal coil is shown in Figure 7.7.1. Toroidal coils are commonly used to form inductors and transformers. The principal advantage of toroidal coils over straight coils in these applications is *magnetic field containment* – as we shall see in this section, the magnetic field outside of a toroidal coil can be made negligibly small. This reduces concern about interactions between this field and other fields and structures in the vicinity.

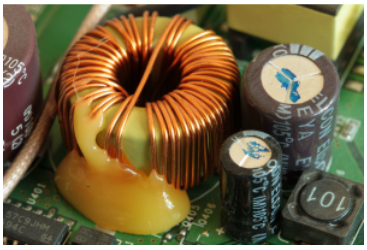


Figure 7.7.1: A toroidal coil used as a large-value inductor in the power supply of a wireless router. (© CC0 1.0; Slick)

In this section, we use the magnetostatic form of Ampere's Circuital Law (ACL) to determine the magnetic field due to a steady (DC) current flowing through a toroidal coil. The problem is illustrated in Figure 7.7.2. The toroid is circular with inner and outer radii a and b, respectively. The coil consists of N windings (turns) of wire wound with uniform winding density. This problem is easiest to work in cylindrical coordinates with the toroid centered on the origin in $z = 0$ plane.

To begin, let's take stock of what we already know about the answer, which is actually quite a bit. First, a review of Section 7.6 ("Magnetic Field Inside a Straight Coil") is recommended. There it is shown that the magnetic field deep inside a straight coil is aligned with axis of the coil. This can be explained using the result for the magnetic field due to a straight line current (Section 7.5), in which we found that the magnetic field follows a "right-hand rule" – The magnetic field points in the direction of the curled fingers of the right hand when the thumb of the right hand is aligned in the direction of current flow. The wire comprising the coil is obviously *not* straight, but we can consider just one short segment of one turn and then sum the results for all such segments. When we do this, we see that the direction of the magnetic field inside the coil must be in the $+\hat{\phi}$ direction when the current I is applied as shown in Figure 7.7.2. Also, because the problem is identical after any amount of rotation around the z axis, the magnitude of the magnetic field cannot depend on ϕ. Putting these findings together, we find that the most general form for the magnetic field intensity inside or outside the coil is $\mathbf{H} = \hat{\phi} H(\rho, z)$.

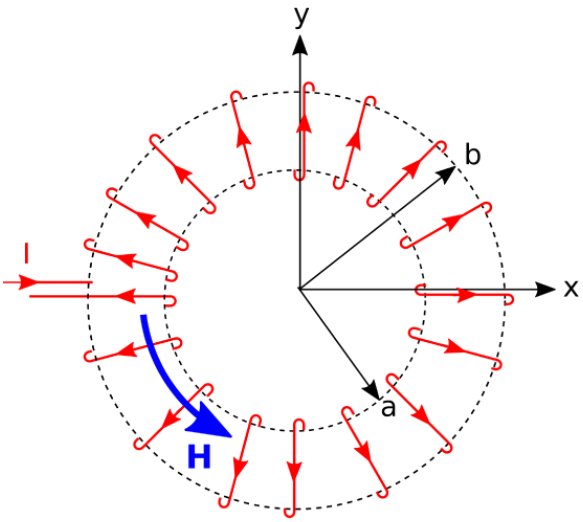


Figure 7.7.2: Geometry of a toroidal coil.

Here's the relevant form of ACL:

$$\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_{encl} \tag{7.7.1}$$

where I_{encl} is the current enclosed by the closed path $\mathcal{C}$. ACL works for *any* closed path, but we need one that encloses some current so as to obtain a relationship between I and $\mathbf{H}$. Also, for simplicity, we prefer a path that lies on a constant-coordinate surface. The selected path is a circle of radius ρ centered on the origin in the $z = z_0$ plane, as shown in Figure 7.7.3. We further require $\mathcal{C}$ to lie

entirely inside the coil, which ensures that the enclosed current includes the current of all the windings as they pass through the hole at the center of the toroid. We choose the direction of integration to be in the $+\phi$ direction, which is consistent with the indicated direction of positive I. That is, when I is positive, the current in the windings of the coil pass through the surface bounded by $\mathcal{C}$ in the same direction as the curled fingers of the right hand when the thumb is aligned in the indicated direction of $\mathcal{C}$.

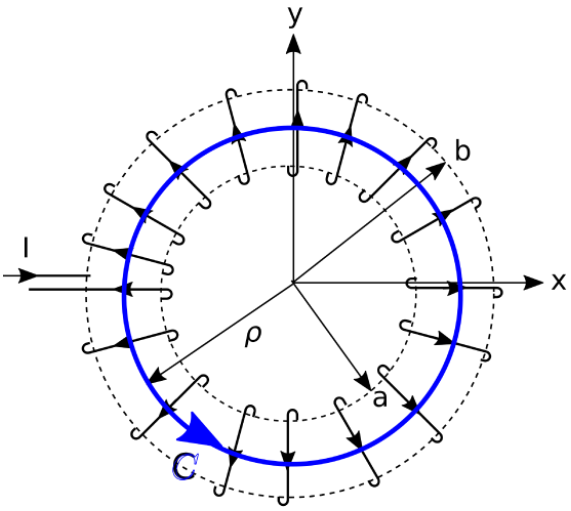


Figure 7.7.3: Selected path of integration.

In terms of the variables we have defined, the enclosed current is simply

$$I_{encl} = NI$$

Equation 7.7.1 becomes

$$\oint_{\mathcal{C}} \left[\hat{\phi} H(\rho, z_0)\right] \cdot d\mathbf{l} = NI \tag{7.7.2}$$

Now evaluating the integral:

$$\begin{aligned} NI &= \int_0^{2\pi} \left[\hat{\phi} H\left(\rho, z_0\right)\right] \cdot (\hat{\phi}\rho d\phi) \\ &= \rho H\left(\rho, z_0\right) \int_0^{2\pi} d\phi \\ &= 2\pi\rho H\left(\rho, z_0\right) \end{aligned}$$

It is now clear that the result is independent of z_0. Summarizing:

$$\boxed{\mathbf{H} = \hat{\phi}\frac{NI}{2\pi\rho} \quad \text{inside coil}} \tag{7.7.3}$$

Let's take a moment to consider the implications of this result.

- Note that it is dimensionally correct; that is, current divided by the circumference of $\mathcal{C}$ ($2\pi\rho$) gives units of A/m, which are the units of **H**.
- We have found that the magnetic field is proportional to winding density (i.e., number of windings divided by circumference) times current. To increase the magnetic field you can either use more windings or increase the current.
- Remarkably, we have found that the magnitude of the magnetic field inside the coil depends only on ρ; i.e., the distance from the central (here, z) axis. It is independent of z.

Summarizing:

> The magnetic field inside a toroidal coil (Equation 7.7.3) depends only on distance from the central axis and is proportional to winding density and current.

Now let us consider what happens outside the coil. For this, we consider any path of integration ($\mathcal{C}$) that lies completely outside the coil. Note that any such path encloses no current and therefore $I_{encl} = 0$ for any such path. In this case we have:

$$\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = 0$$

There are two ways this could be true. Either $\mathbf{H}$ could be zero everywhere along the path, or $\mathbf{H}$ could be non-zero along the path in such a way that the integral windings out to be zero. The radial symmetry of the problem rules out the second possibility – if $\mathbf{H}$ is radially symmetric and $\mathcal{C}$ is radially symmetric, then the sign of $\mathbf{H} \cdot d\mathbf{l}$ should not change over $\mathcal{C}$. Therefore:

The magnetic field everywhere outside an ideal toroidal coil is zero.

Note the caveat signaled by the use of the adjective “ideal.” In a practical toroidal coil, we expect there will be some leakage of magnetic flux between the windings. In practice, this leakage can be made negligibly small by using a sufficiently high winding density and winding the wire on material on a toroidal form (a “core”) having sufficiently large permeability. The use of a high-permeability core, as shown in Figure 7.7.1, will dramatically improve the already pretty-good containment. In fact, the use of such a core allows the spacing between windings to become quite large before leakage becomes significant.

One final observation about toroidal coils is that at no point in the derivation of the magnetic field did we need to consider the cross-sectional shape of the coil; we merely needed to know whether $\mathcal{C}$ was inside or outside the coil. Therefore:

> The magnetic field inside an ideal toroidal coil does not depend on the cross-sectional shape of the coil.

This page titled 7.7: Magnetic Field of a Toroidal Coil is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.7: Magnetic Field of a Toroidal Coil** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.8: Magnetic Field of an Infinite Current Sheet

We now consider the magnetic field due to an infinite sheet of current, shown in Figure 7.8.1. The solution to this problem is useful as a building block and source of insight in more complex problems, as well as being a useful approximation to some practical problems involving current sheets of finite extent including, for example, microstrip transmission line and ground plane currents in printed circuit boards.

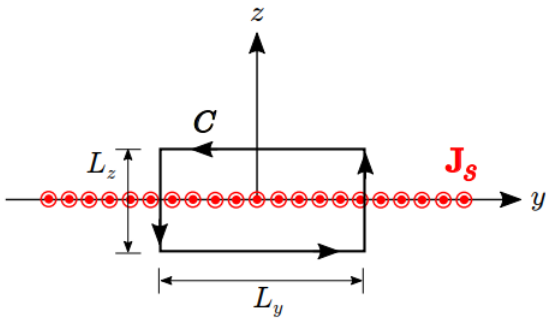


Figure 7.8.1: Analysis of the magnetic field due to an infinite thin sheet of current.

The current sheet in Figure 7.8.1 lies in the $z = 0$ plane and the current density is $\mathbf{J}_s = \hat{\mathbf{x}} J_s$ (units of A/m); i.e., the current is uniformly distributed such that the total current crossing any segment of width Δy along the y direction is $J_s \Delta y$.

To begin, let's take stock of what we already know about the answer, which is actually quite a bit. For example, imagine the current sheet as a continuum of thin strips parallel to the x axis and very thin in the y dimension. Each of these strips individually behaves like a straight line current $I = J_s \Delta y$ (units of A). The magnetic field due to each of these strips is determined by a "right-hand rule" – the magnetic field points in the direction of the curled fingers of the right hand when the thumb of the right hand is aligned in the direction of current flow. (Section 7.5). It is apparent from this much that $\mathbf{H}$ can have no $\hat{\mathbf{y}}$ component, since the field of each individual strip has no $\hat{\mathbf{y}}$ component. When the magnetic field due to each strip is added to that of all the other strips, the $\hat{\mathbf{z}}$ component of the sum field must be zero due to symmetry. It is also clear from symmetry considerations that the magnitude of $\mathbf{H}$ cannot depend on x or y. Summarizing, we have determined that the most general form for $\mathbf{H}$ is $\hat{\mathbf{y}} H(z)$, and furthermore, the sign of $H(z)$ must be positive for $z < 0$ and negative for $z > 0$.

It's possible to solve this problem by actually summing over the continuum of thin current strips as imagined above.[1] However, it's far easier to use Ampere's Circuital Law (ACL; Section 7.4). Here's the relevant form of ACL:

$$\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_{encl} \tag{7.8.1}$$

where I_{encl} is the current enclosed by a closed path $\mathcal{C}$. ACL works for *any* closed path, but we need one that encloses some current so as to obtain a relationship between $\mathbf{J}_s$ and $\mathbf{H}$. Also, for simplicity, we prefer a path that lies on a constant-coordinate surface. A convenient path in this problem is a rectangle lying in the $x = 0$ plane and centered on the origin, as shown in Figure 7.8.1. We choose the direction of integration to be counter-clockwise from the perspective shown in Figure 7.8.1, which is consistent with the indicated direction of positive J_s according to the applicable right-hand rule from Stokes' Theorem. That is, when J_s is positive (current flowing in the $+\hat{\mathbf{x}}$ direction), the current passes through the surface bounded by $\mathcal{C}$ in the same direction as the curled fingers of the right hand when the thumb is aligned in the indicated direction of $\mathcal{C}$.

Let us define L_y to be the width of the rectangular path of integration in the y dimension and L_z to be the width in the z dimension. In terms of the variables we have defined, the enclosed current is simply

$$I_{encl} = J_s L_y$$

Equation 7.8.1 becomes

$$\oint_{\mathcal{C}} [\hat{\mathbf{y}} H(z)] \cdot d\mathbf{l} = J_s L_y \tag{7.8.2}$$

Note that $\mathbf{H} \cdot d\mathbf{l} = 0$ for the vertical sides of the path, since $\mathbf{H}$ is $\hat{\mathbf{y}}$-directed and $d\mathbf{l} = \hat{\mathbf{z}} dz$ on those sides. Therefore, only the horizontal sides contribute to the integral and we have:

$$\int^{+L_w/2} \left[\hat{\mathbf{y}} H\left(-\frac{L_z}{}\right)\right] \cdot (\hat{\mathbf{y}} du)$$

$$\int_{-L_y/2}^{+L_y/2} \left[\hat{y}H\left(-\frac{L_z}{2}\right)\right] \cdot (\hat{y}dy) + \int_{+L_y/2}^{-L_y/2} \left[\hat{y}H\left(+\frac{L_z}{2}\right)\right] \cdot (\hat{y}dy) = J_s L_y \tag{7.8.3}$$

Now evaluating the integrals:

$$H\left(-\frac{L_z}{2}\right)\ L_y - H\left(+\frac{L_z}{2}\right)\ L_y = J_s L_y$$

Note that all factors of L_y cancel in the above equation. Furthermore, $H(-L_z/2) = -H(+L_z/2)$ due to (1) symmetry between the upper and lower half-spaces and (2) the change in sign between these half-spaces, noted earlier. We use this to eliminate $H(+L_z/2)$ and solve for $H(-L_z/2)$ as follows:

$$2H(-L_z/2) = J_s$$

yielding

$$H(-L_z/2) = +\frac{J_s}{2}$$

and therefore

$$H(+L_z/2) = -\frac{J_s}{2}$$

Furthermore, note that **H** is independent of L_z; for example, the result we just found indicates the same value of $H(+L_z/2)$ regardless of the value of L_z. Therefore, **H** is uniform throughout all space, except for the change of sign corresponding for the field above vs. below the sheet.

Summarizing

$$\boxed{\mathbf{H} = \pm\hat{y}\frac{J_s}{2} \quad \text{for } z \lessgtr 0} \tag{7.8.4}$$

> The magnetic field intensity due to an infinite sheet of current (Equation 7.8.4) is spatially uniform except for a change of sign corresponding for the field above vs. below the sheet.

1. In fact, this is pretty good thing to try, if for no other reason than to see how much simpler it is to use ACL instead.↩

This page titled 7.8: Magnetic Field of an Infinite Current Sheet is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.8: Magnetic Field of an Infinite Current Sheet** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.9: Ampere's Law (Magnetostatics) - Differential Form

The integral form of Amperes' Circuital Law (ACL) for magnetostatics relates the magnetic field along a closed path to the total current flowing through any surface bounded by that path. In mathematical form:

$$\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_{encl}$$

where $\mathbf{H}$ is magnetic field intensity, $\mathcal{C}$ is the closed curve, and I_{encl} is the total current flowing through any surface bounded by $\mathcal{C}$. In this section, we derive the differential form of this equation. In some applications, this differential equation, combined with boundary conditions associated with discontinuities in structure and materials, can be used to solve for the magnetic field in arbitrarily complicated scenarios. A more direct reason for seeking out this differential equation is that we gain a little more insight into the relationship between current and the magnetic field, disclosed at the end of this section.

The equation we seek may be obtained using Stokes' Theorem, which in the present case may be written:

$$\int_{\mathcal{S}} (\nabla \times \mathbf{H}) \cdot d\mathbf{s} = \oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} \tag{7.9.1}$$

where $\mathcal{S}$ is any surface bounded by $\mathcal{C}$, and $d\mathbf{s}$ is the differential surface area combined with the unit vector in the direction determined by the right-hand rule from Stokes' Theorem. ACL tells us that the right side of the above equation is simply I_{encl}. We may express I_{encl} as the integral of the volume current density $\mathbf{J}$ (units of A/m^2; Section 6.2) as follows:

$$I_{encl} = \int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{s}$$

so we may rewrite Equation 7.9.1 as follows:

$$\int_{\mathcal{S}} (\nabla \times \mathbf{H}) \cdot d\mathbf{s} = \int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{s}$$

The above relationship must hold regardless of the specific location or shape of $\mathcal{S}$. The only way this is possible for all possible surfaces in all applicable scenarios is if the integrands are equal. Thus, we obtain the desired expression:

$$\boxed{\nabla \times \mathbf{H} = \mathbf{J}} \tag{7.9.2}$$

That is, the curl of the magnetic field intensity at a point is equal to the volume current density at that point. Recalling the properties of the curl operator – in particular, that curl involves derivatives with respect to direction – we conclude:

> The differential form of Ampere's Circuital Law for magnetostatics (Equation 7.9.2) indicates that the volume current density at any point in space is proportional to the spatial rate of change of the magnetic field and is perpendicular to the magnetic field at that point.

This page titled 7.9: Ampere's Law (Magnetostatics) - Differential Form is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.9: Ampere's Law (Magnetostatics) - Differential Form** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.10: Boundary Conditions on the Magnetic Flux Density (B)

In homogeneous media, electromagnetic quantities vary smoothly and continuously. At an interface between dissimilar media, however, it is possible for electromagnetic quantities to be discontinuous. Continuities and discontinuities in fields can be described mathematically by *boundary conditions* and used to constrain solutions for fields away from these interfaces.

In this section, we derive the boundary condition on the magnetic flux density $\mathbf{B}$ at a smooth interface between two material regions, as shown in Figure 7.10.1.[1] The desired boundary condition may be obtained from Gauss' Law for Magnetic Fields (GLM):

$$\oint_{\mathcal{S}} \mathbf{B}\cdot d\mathbf{s} = 0$$

where $\mathcal{S}$ is any closed surface. Let $\mathcal{S}$ take the form of cylinder centered at a point on the interface, and for which the flat ends are parallel to the surface and perpendicular to $\hat{\mathbf{n}}$, as shown in Figure 7.10.1. Let the radius of this cylinder be a, and let the length of the cylinder be $2h$.

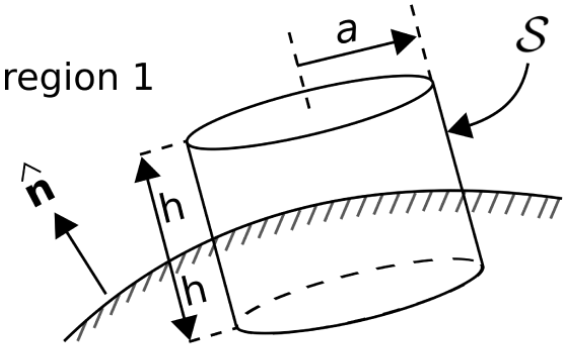


Figure 7.10.1: Determination of the boundary condition on B at the interface between material regions (© CC BY SA 4.0; K. Kikkeri)

From GLM, we have

$$\begin{aligned}\oint_{\mathcal{S}} \mathbf{B}\cdot d\mathbf{s} = &\int_{top} \mathbf{B}\cdot d\mathbf{s} \\ &+ \int_{side} \mathbf{B}\cdot d\mathbf{s} \\ &+ \int_{bottom} \mathbf{B}\cdot d\mathbf{s} \qquad = 0\end{aligned}$$

Now let us reduce h and a together while (1) maintaining a constant ratio $h/a \ll 1$ and (2) keeping $\mathcal{S}$ centered on the interface. Because $h \ll a$, the area of the side can be made negligible relative to the area of the top and bottom. Then, as $h \to 0$, we are left with

$$\int_{top} \mathbf{B}\cdot d\mathbf{s} + \int_{bottom} \mathbf{B}\cdot d\mathbf{s} \to 0$$

As the area of the top and bottom sides become infinitesimal, the variation in $\mathbf{B}$ over these areas becomes negligible. Now we have simply:

$$\mathbf{B}_1\cdot\hat{\mathbf{n}}\Delta A + \mathbf{B}_2\cdot(-\hat{\mathbf{n}})\,\Delta A \to 0$$

where $\mathbf{B}_1$ and $\mathbf{B}_2$ are the magnetic flux densities at the interface but in regions 1 and 2, respectively, and ΔA is the area of the top and bottom sides. Note that the orientation of $\hat{\mathbf{n}}$ is important – we have assumed $\hat{\mathbf{n}}$ points into region 1, and we must now stick with this choice. Thus, we obtain

$$\boxed{\hat{\mathbf{n}}\cdot(\mathbf{B}_1 - \mathbf{B}_2) = 0} \tag{7.10.1}$$

where, as noted above, $\hat{\mathbf{n}}$ points into region 1.

Summarizing

> The normal (perpendicular) component of $\mathbf{B}$ across the boundary between two material regions is continuous.

It is worth noting what this means for the magnetic field intensity $\mathbf{H}$. Since $\mathbf{B} = \mu\mathbf{H}$, it must be that

> The normal (perpendicular) component of $\mathbf{H}$ across the boundary between two material regions is discontinuous if the permeabilities are unequal.

---

1. It may be helpful to note the similarity (duality, in fact) between this derivation and the derivation of the associated boundary condition on $\mathbf{D}$ presented in Section 5.18.↩

This page titled 7.10: Boundary Conditions on the Magnetic Flux Density (B) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.10: Boundary Conditions on the Magnetic Flux Density (B)** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.11: Boundary Conditions on the Magnetic Field Intensity (H)

In homogeneous media, electromagnetic quantities vary smoothly and continuously. At a boundary between dissimilar media, however, it is possible for electromagnetic quantities to be discontinuous. Continuities and discontinuities in fields can be described mathematically by *boundary conditions* and used to constrain solutions for fields away from these boundaries. In this section, we derive boundary conditions on the magnetic field intensity $\mathbf{H}$.

To begin, consider a region consisting of only two media that meet at a smooth boundary as shown in Figure 7.11.1. The desired boundary condition can be obtained directly from Ampere's Circuital Law (ACL):

$$\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_{encl} \tag{7.11.1}$$

where $\mathcal{C}$ is any closed path and I_{encl} is the current that flows through the surface bounded by that path in the direction specified by the "right-hand rule" of Stokes' theorem.

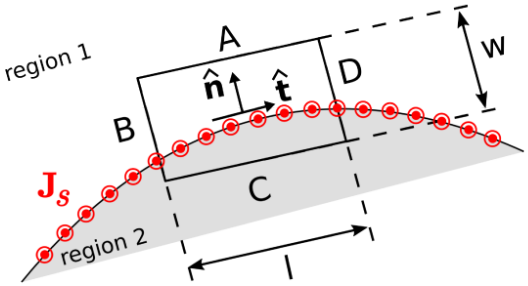


Figure 7.11.1: Determining the boundary condition on H at the smooth boundary between two material regions.

Let $\mathcal{C}$ take the form of a rectangle centered on a point on the boundary as shown in Figure 7.11.1, perpendicular to the direction of current flow at that location. Let the sides A, B, C, and D be perpendicular and parallel to the boundary. Let the length of the parallel sides be l, and let the length of the perpendicular sides be w. Now we apply ACL. We must integrate in a counter-clockwise direction in order to be consistent with the indicated reference direction for $\mathbf{J}_s$. Thus:

$$\begin{aligned} \oint \mathbf{H} \cdot d\mathbf{l} = &\int_A \mathbf{H} \cdot d\mathbf{l} \\ &+ \int_B \mathbf{H} \cdot d\mathbf{l} \\ &+ \int_C \mathbf{H} \cdot d\mathbf{l} \\ &+ \int_D \mathbf{H} \cdot d\mathbf{l} \qquad = I_{encl} \end{aligned}$$

Now we let w and l become vanishingly small while (1) maintaining the ratio l/w and (2) keeping $\mathcal{C}$ centered on the boundary. In this process, the contributions from the B and D segments become equal in magnitude but opposite in sign; i.e.,

$$\int_B \mathbf{H} \cdot d\mathbf{l} + \int_D \mathbf{H} \cdot d\mathbf{l} \to 0$$

This leaves

$$\int_A \mathbf{H} \cdot d\mathbf{l} + \int_C \mathbf{H} \cdot d\mathbf{l} \to I_{encl}$$

Let us define the unit vector $\hat{\mathbf{t}}$ ("tangent") as shown in Figure 7.11.1. Now we have simply:

$$\mathbf{H}_1 \cdot \left(-\hat{\mathbf{t}}\Delta l\right) + \mathbf{H}_2 \cdot \left(+\hat{\mathbf{t}}\Delta l\right) = I_{encl} \tag{7.11.2}$$

where $\mathbf{H}_1$ and $\mathbf{H}_2$ are the fields evaluated on the two sides of the boundary, and $\Delta l \to 0$ is the length of sides A and C. As always, I_{encl} (units of A) may be interpreted as the flux of the current density $\mathbf{J}_s$ (units of A/m) flowing past a line on the surface having length Δl (units of m) perpendicular to $\hat{\mathbf{t}} \times \hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is the normal to the surface, pointing into Region 1. Stated mathematically:

$$I_{encl} \to \mathbf{J}_s \cdot \left(\Delta l\, \hat{\mathbf{t}} \times \hat{\mathbf{n}}\right)$$

Before proceeding, note this is true regardless of the particular direction we selected for $\hat{\mathbf{t}}$; it is only necessary that $\hat{\mathbf{t}}$ be tangent to the boundary. Thus, $\hat{\mathbf{t}} \times \hat{\mathbf{n}}$ need not necessarily be in the same direction as $\mathbf{J}_s$. Now Equation 7.11.2 can be written:

$$\mathbf{H}_2 \cdot \hat{\mathbf{t}} \Delta l - \mathbf{H}_1 \cdot \hat{\mathbf{t}} \Delta l = \mathbf{J}_s \cdot \left(\hat{\mathbf{t}} \times \hat{\mathbf{n}} \right) \Delta l$$

Eliminating the common factor of Δl and arranging terms on the left:

$$(\mathbf{H}_2 - \mathbf{H}_1) \cdot \hat{\mathbf{t}} = \mathbf{J}_s \cdot \left(\hat{\mathbf{t}} \times \hat{\mathbf{n}} \right) \tag{7.11.3}$$

The right side may be transformed using a vector identity ($\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})$ of Section B3) to obtain:

$$(\mathbf{H}_2 - \mathbf{H}_1) \cdot \hat{\mathbf{t}} = \hat{\mathbf{t}} \cdot (\hat{\mathbf{n}} \times \mathbf{J}_s) \tag{7.11.4}$$

Equation 7.11.4 is the boundary condition we seek. We have found that the component of $\mathbf{H}_2 - \mathbf{H}_1$ (the difference between the magnetic field intensities at the boundary) in any direction tangent to the boundary is equal to the component of the current density flowing in the perpendicular direction $\hat{\mathbf{n}} \times \mathbf{J}_s$. Said differently:

> A discontinuity in the tangential component of the magnetic field intensity at the boundary must be supported by surface current flowing in a direction perpendicular to this component of the field.

An important consequence is that:

> If there is no surface current, then the tangential component of the magnetic field intensity is continuous across the boundary.

It is possible to obtain a mathematical form of the boundary condition that is more concise and often more useful than Equation 7.11.4. This form may be obtained as follows. First, we note that the dot product with respect to $\hat{\mathbf{t}}$ on both sides of Equation 7.11.4 means simply "any component that is tangent to the boundary." We need merely to make sure we are comparing the same tangential component on each side of the equation. For example $\hat{\mathbf{n}} \times (\mathbf{H}_2 - \mathbf{H}_1)$ is tangential to the boundary, since $\hat{\mathbf{n}}$ is perpendicular to the boundary and therefore any cross product involving $\hat{\mathbf{n}}$ will be perpendicular to $\hat{\mathbf{n}}$. The corresponding component of the current density is $\hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{J}_s)$, so Equation 7.11.4 may be equivalently written as follows:

$$\hat{\mathbf{n}} \times (\mathbf{H}_2 - \mathbf{H}_1) = \hat{\mathbf{n}} \times \hat{\mathbf{n}} \times \mathbf{J}_s \tag{7.11.5}$$

Applying a vector identity ($\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ of Section B3) to the right side of Equation 7.11.5 we obtain:

$$\begin{aligned} \hat{\mathbf{n}} \times \hat{\mathbf{n}} \times \mathbf{J}_s &= \hat{\mathbf{n}} \left(\hat{\mathbf{n}} \cdot \mathbf{J}_s \right) - \mathbf{J}_s \left(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}} \right) \\ &= \hat{\mathbf{n}} \left(0 \right) - \mathbf{J}_s \left(1 \right) \\ &= -\mathbf{J}_s \end{aligned}$$

Therefore:

$$\hat{\mathbf{n}} \times (\mathbf{H}_2 - \mathbf{H}_1) = -\mathbf{J}_s$$

The minus sign on the right can be eliminated by swapping $\mathbf{H}_2$ and $\mathbf{H}_1$ on the left, yielding

$$\boxed{\hat{\mathbf{n}} \times (\mathbf{H}_1 - \mathbf{H}_2) = \mathbf{J}_s}$$

This is the form in which the boundary condition is most commonly expressed.

It is worth noting what this means for the magnetic field intensity $\mathbf{B}$. Since $\mathbf{B} = \mu \mathbf{H}$:

> In the absence of surface current, the tangential component of $\mathbf{B}$ across the boundary between two material regions is discontinuous if the permeabilities are unequal.

This page titled 7.11: Boundary Conditions on the Magnetic Field Intensity (H) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.11: Boundary Conditions on the Magnetic Field Intensity (H)** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.12: Inductance

Current creates a magnetic field, which subsequently exerts force on other current-bearing structures. For example, the current in each winding of a coil exerts a force on every other winding of the coil. If the windings are fixed in place, then this force is unable to do work (i.e., move the windings), so instead the coil stores potential energy. This potential energy can be released by turning off the external source. When this happens, charge continues to flow, but is now propelled by the magnetic force. The potential energy that was stored in the coil is converted to kinetic energy and subsequently used to redistribute the charge until no current flows. At this point, the inductor has expended its stored energy. To restore energy, the external source must be turned back on, restoring the flow of charge and thereby restoring the magnetic field.

Now recall that the magnetic field is essentially defined in terms of the force associated with this potential energy; i.e., $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ where q is the charge of a particle comprising the current, $\mathbf{v}$ is the velocity of the particle, and $\mathbf{B}$ is magnetic flux density. So, rather than thinking of the potential energy of the system as being associated with the magnetic force applied to current, it is equally valid to think of the potential energy as being stored in the magnetic field associated with the current distribution. The energy stored in the magnetic field depends on the geometry of the current-bearing structure and the permeability of the intervening material because the magnetic field depends on these parameters.

The relationship between current applied to a structure and the energy stored in the associated magnetic field is what we mean by *inductance*. We may fairly summarize this insight as follows:

> Inductance is the ability of a structure to store energy in a magnetic field.

> The inductance of a structure depends on the geometry of its current-bearing structures and the permeability of the intervening medium.

Note that inductance does *not* depend on current, which we view as either a stimulus or response from this point of view. The corresponding response or stimulus, respectively, is the magnetic flux associated with this current. This leads to the following definition:

$$L = \frac{\Phi}{I} \quad \text{(single linkage)} \tag{7.12.1}$$

where Φ (units of Wb) is magnetic flux, I (units of A) is the current responsible for this flux, and L (units of H) is the associated inductance. (The “single linkage” caveat will be explained below.) In other words, a device with high inductance generates a large magnetic flux in response to a given current, and therefore stores more energy for a given current than a device with lower inductance.

To use Equation 7.12.1 we must carefully define what we mean by “magnetic flux” in this case. Generally, magnetic flux is magnetic flux density (again, $\mathbf{B}$, units of Wb/m^2) integrated over a specified surface $\mathcal{S}$, so

$$\Phi = \int_{\mathcal{S}} \mathbf{B} \cdot d\mathbf{s}$$

where $d\mathbf{s}$ is the differential surface area vector, with direction normal to $\mathcal{S}$. However, this leaves unanswered the following questions: Which $\mathcal{S}$, and which of the two possible normal directions of $d\mathbf{s}$? For a meaningful answer, $\mathcal{S}$ must uniquely associate the magnetic flux to the associated current. Such an association exists if we require the current to form a closed loop. This is shown in Figure 7.12.1. Here $\mathcal{C}$ is the closed loop along which the current flows, $\mathcal{S}$ is a surface bounded by $\mathcal{C}$, and the direction of $d\mathbf{s}$ is defined according to the right-hand rule of Stokes’ Theorem. Note that $\mathcal{C}$ can be a closed loop of *any* shape; i.e., not just circular, and not restricted to lying in a plane. Further note that $\mathcal{S}$ used in the calculation of Φ can be *any* surface bounded by $\mathcal{C}$. This is because magnetic field lines form closed loops such that any one magnetic field line intersects any open surface bounded by $\mathcal{C}$ exactly once. Such an intersection is sometimes called a “linkage.” So there we have it – we require the current I to form a closed loop, we measure the magnetic flux through this loop using the sign convention of the right-hand rule, and the ratio is the inductance.

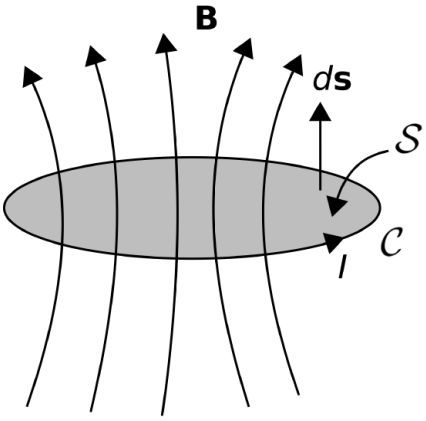


Figure 7.12.1: Association between a closed loop of current and the associated magnetic flux. (© CC BY SA 4.0; K. Kikkeri)

Many structures consist of multiple such loops – the coil is of course one of these. In a coil, each winding carries the same current, and the magnetic fields of the windings add to create a magnetic field, which grows in proportion to the winding density (Section 7.6). The magnetic flux density inside a coil is proportional to the number of windings, N, so the flux Φ in Equation 7.12.1 should properly be indicated as $N\Phi$. Another way to look at this is that we are counting the number of times the same current is able to generate a unique set of magnetic field lines that intersect $\mathcal{S}$.

Summarizing, our complete definition for inductance is

$$\boxed{L = \frac{N\Phi}{I} \quad \text{(identical linkages)}} \tag{7.12.2}$$

> An engineering definition of inductance is Equation 7.12.2, with the magnetic flux defined to be that associated with a single closed loop of current with sign convention as indicated in Figure 7.12.1, and N defined to be the number of times the same current I is able to create that flux.

What happens if the loops have different shapes? For example, what if the coil is not a cylinder, but rather cone-shaped? (Yes, there is such a thing – see "Additional Reading" at the end of this section.) In this case, one needs a better way to determine the factor $N\Phi$ since the flux associated with each loop of current will be different. However, this is beyond the scope of this section.

An *inductor* is a device that is designed to exhibit a specified inductance. We can now make the connection to the concept of the inductor as it appears in elementary circuit theory. First, we rewrite Equation 7.12.2 as follows:

$$I = \frac{N\Phi}{L}$$

Taking the derivative of both sides of this equation with respect to time, we obtain:

$$\frac{d}{dt}I = \frac{N}{L}\frac{d}{dt}\Phi \tag{7.12.3}$$

Now we need to reach beyond the realm of magnetostatics for just a moment. Section 8.3 ("Faraday's Law") shows that the change in Φ associated with a change in current results in the creation of an electrical potential equal to $-Nd\Phi/dt$ realized over the loop $\mathcal{C}$. In other words, the terminal voltage V is $+Nd\Phi/dt$, with the change of sign intended to keep the result consistent with the sign convention relating current and voltage in passive devices. Therefore, $d\Phi/dt$ in Equation 7.12.3 is equal to V/N. Making the substitution we find:

$$V = L\frac{d}{dt}I \tag{7.12.4}$$

This is the expected relationship from elementary circuit theory.

Another circuit theory concept related to inductance is *mutual inductance*. Whereas inductance relates changes in current to instantaneous voltage in the same device (Equation 7.12.4), mutual inductance relates changes in current in one device to instantaneous voltage in a *different* device. This can occur when the two devices are coupled ("linked") by the same magnetic field.

For example, *transformers* (Section 8.5) typically consist of separate coils that are linked by the same magnetic field lines. The voltage across one coil may be computed as the time-derivative of current on the other coil times the mutual inductance.

Let us conclude this section by taking a moment to dispel a common misconception about inductance. The misconception pertains to the following question. If the current does not form a closed loop, what is the inductance? For example, engineers sometimes refer to the inductance of a pin or lead of an electronic component. A pin or lead is not a closed loop, so the formal definition of inductance given above – ratio of magnetic flux to current – does not apply. The broader definition of inductance – the ability to store energy in a magnetic field – does apply, but this is *not* what is meant by "pin inductance" or "lead inductance." What is actually meant is the imaginary part of the impedance of the pin or lead – i.e., the reactance – expressed as an equivalent inductance. In other words, the reactance of an inductive device is positive, so any device that also exhibits a positive reactance can be viewed from a circuit theory perspective as an equivalent inductance. This is not referring to the storage of energy in a magnetic field; it merely means that the device can be modeled as an inductor in a circuit diagram. In the case of "pin inductance," the culprit is not actually inductance, but rather skin effect (see "Additional References" at the end of this section). Summarizing:

> Inductance implies positive reactance, but positive reactance does not imply the physical mechanism of inductance.

This page titled 7.12: Inductance is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.12: Inductance** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.13: Inductance of a Straight Coil

In this section, we determine the inductance of a straight coil, as shown in Figure 7.13.1. The coil is circular with radius a and length l and consists of N windings of wire wound with uniform winding density. Also, we assume the winding density N/l is large enough that magnetic field lines cannot enter or exit between windings but rather must traverse the entire length of the coil. Since the coil forms a cylinder, the problem is easiest to work in cylindrical coordinates with the axis of the coil aligned along the z axis.

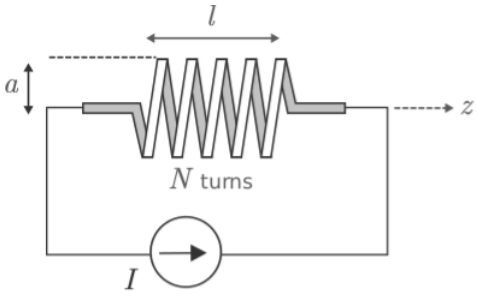


Figure 7.13.1: Determination of the inductance of a straight coil.

Inductance L in this case is given by (Section 7.12)

$$L = \frac{N\Phi}{I} \tag{7.13.1}$$

where I is current and Φ is the magnetic flux associated with one winding of the coil. Magnetic flux in this case is given by

$$\Phi = \int_{\mathcal{S}} \mathbf{B}\cdot d\mathbf{s}$$

where $\mathbf{B}$ is the magnetic flux density (units of T = Wb/m^2), $\mathcal{S}$ is the surface bounded by a single current loop, and $d\mathbf{s}$ points in the direction determined by the right hand rule with respect to the direction of positive current flow.

First, let's determine the magnetic field. The magnetic flux density deep inside the coil is (Section 7.6):

$$\mathbf{B} \approx \hat{\mathbf{z}}\frac{\mu NI}{l} \tag{7.13.2}$$

Is it reasonable to use this approximation here? Since inductance pertains to energy storage, the question is really what fraction of the energy is stored in a field that is well-described by this approximation, as opposed to energy stored in the "fringing field" close to the ends of the coil. If we make l sufficiently large relative to a, then presumably energy storage in the fringing field will be negligible in comparison. Since the alternative leads to a much more complicated problem, we shall assume that Equation 7.13.2 is valid for the interior of the coil.

Next, we determine Φ. In this case, a natural choice for $\mathcal{S}$ is the interior cross-section of the coil in a plane perpendicular to the axis. The direction of $d\mathbf{s}$ must be $+\hat{\mathbf{z}}$ since this is the direction in which the fingers of the right hand point when the current flows in the direction indicated in Figure 7.13.1. Thus, we have

$$\begin{aligned}\Phi &\approx \int_{\mathcal{S}} \left(\hat{\mathbf{z}}\frac{\mu NI}{l}\right)\cdot(\mathbf{z}ds) \\ &= \frac{\mu NI}{l}\int_{\mathcal{S}} ds \\ &= \frac{\mu NI}{l}A\end{aligned} \tag{7.13.3}$$

where A is the cross-sectional area of the coil.

Finally from Equation 7.13.1 we obtain

$$\boxed{L \approx \frac{\mu N^2 A}{l} \quad (l \gg a)} \tag{7.13.4}$$

Note that this is dimensionally correct; that is, permeability (units of H/m) times area (units of m^2) divided by length (units of m) gives units of H, as expected. Also, it is worth noting that inductance is proportional to permeability and cross-sectional area, and inversely proportional to length. Interestingly the inductance is proportional to N^2 as opposed to N; this is because field strength

increases with N, and independently there are N flux linkages. Finally, we note that the inductance does not depend on the *shape* of the coil cross-section, but only on the *area* of the cross-section. Summarizing:

The inductance of a long straight coil is given approximately by Equation 7.13.4.

Again, this result is approximate because it neglects the non-uniform fringing field near the ends of the coil and the possibility that magnetic field lines escape between windings due to inadequate winding density. Nevertheless, this result facilitates useful engineering analysis and design.

This page titled 7.13: Inductance of a Straight Coil is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.13: Inductance of a Straight Coil** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.14: Inductance of a Coaxial Structure

Let us now determine the inductance of coaxial structure, shown in Figure 7.14.1. The inductance of this structure is of interest for a number of reasons – in particular, for determining the characteristic impedance of coaxial transmission line, as addressed in Section 3.10.

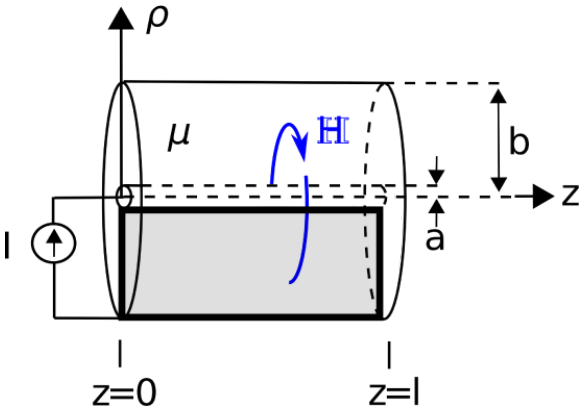


Figure 7.14.1: Determining the inductance of coaxial line.

For our present purpose, we may model the structure as shown in Figure 7.14.1. This model consists of two concentric perfectly-conducting cylinders of radii a and b, separated by a homogeneous material having permeability μ. To facilitate analysis, let us place the $+z$ axis along the common axis of the concentric cylinders, so that the cylinders may be described as the surfaces $\rho = a$ and $\rho = b$.

Below we shall find the inductance by assuming a current I on the inner conductor and integrating over the resulting magnetic field to obtain the magnetic flux Φ between the conductors. Then, inductance can be determined as the ratio of the response flux to the source current.

Before we get started, note the derivation we are about to do is similar to the derivation of the *capacitance* of a coaxial structure, addressed in Section 5.24. The reader may benefit from a review of that section before attempting this derivation.

The first step is to find the magnetic field inside the structure. This is relatively simple if we may neglect fringing fields, since then the internal field may be assumed to be constant with respect to z. This analysis will also apply to the case where the length l pertains to one short section of a much longer structure; in this case we will obtain the inductance *per length* as opposed to the total inductance for the structure. Note that the latter is exactly what we need for the transmission line lumped-element equivalent circuit model (Section 3.4).

To determine the inductance, we invoke the definition:

$$L \triangleq \frac{\Phi}{I} \tag{7.14.1}$$

A current I flowing in the $+z$ direction on the inner conductor gives rise to a magnetic field inside the coaxial structure. The magnetic field intensity for this scenario was determined in Section 7.5 where we found

$$\mathbf{H} = \hat{\phi}\frac{I}{2\pi\rho} \quad , \quad a \leq \rho \leq b$$

The reader should note that in that section we were considering merely a line of current; not a coaxial structure. So, on what basis do we claim the field for inside the coaxial structure is the same? This is a consequence of Ampere's Law (Section 7.14):

$$\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_{encl}$$

If in this new problem we specify the same circular path $\mathcal{C}$ with radius greater than a and less than b, then the enclosed current is simply I. The presence of the outer conductor does not change the radial symmetry of the problem, and nothing else remains that can change the outcome. This is worth noting for future reference:

The magnetic field inside a coaxial structure comprised of concentric conductors bearing current I is identical to the magnetic field of the line current I in free space.

We're going to need magnetic flux density (B) as opposed to $\mathbf{H}$ in order to get the magnetic flux. This is simple since they are related by the permeability of the medium; i.e., $\mathbf{B} = \mu\mathbf{H}$. Thus:

$$\mathbf{B} = \hat{\phi}\frac{\mu I}{2\pi\rho} \quad , \quad a \leq \rho \leq b$$

Next, we get Φ by integrating over the magnetic flux density

$$\Phi = \int_{\mathcal{S}} \mathbf{B} \cdot d\mathbf{s}$$

where $\mathcal{S}$ is any open surface through which all magnetic field lines within the structure must pass. Since this can be *any* such surface, we may as well choose the simplest one. The simplest such surface is a plane of constant ϕ, since such a plane is a constant-coordinate surface and perpendicular to the magnetic field lines. This surface is shown as the shaded area in Figure 7.14.1. Using this surface we find:

$$\begin{aligned} \Phi &= \int_{\rho=a}^{b}\int_{z=0}^{l}\left(\hat{\phi}\frac{\mu I}{2\pi\rho}\right)\cdot(\hat{\phi}d\rho dz) \\ &= \frac{\mu I}{2\pi}\left(\int_{z=0}^{l} dz\right)\left(\int_{\rho=a}^{b}\frac{d\rho}{\rho}\right) \\ &= \frac{\mu I l}{2\pi}\ln\left(\frac{b}{a}\right) \end{aligned} \tag{7.14.2}$$

Wrapping up:

$$L \triangleq \frac{\Phi}{I} = \frac{(\mu I l/2\pi)\ln(b/a)}{I}$$

Note that factors of I in the numerator and denominator cancel out, leaving:

$$\boxed{L = \frac{\mu l}{2\pi}\ln\left(\frac{b}{a}\right)}$$

Note that this is dimensionally correct, having units of H. Also note that this is expression depends only on materials (through μ) and geometry (through l, a, and b). Notably, it does *not* depend on current, which would imply non-linear behavior.

To make the connection back to lumped-element transmission line model parameters (Sections 3.4 and 3.10), we simply divide by l to get the per-unit length parameter:

$$\boxed{L' = \frac{\mu}{2\pi}\ln\left(\frac{b}{a}\right)} \tag{7.14.3}$$

which has the expected units of H/m.

Example 7.14.1: Inductance of RG-59 coaxial cable

RG-59 coaxial cable consists of an inner conductor having radius 0.292 mm, an outer conductor having radius 1.855 mm, and polyethylene (a non-magnetic dielectric) spacing material. Estimate the inductance per length of RG-59.

Solution

From the problem statement, $a = 0.292$ mm, $b = 1.855$ mm, and $\mu \cong \mu_0$ since the spacing material is non-magnetic. Using Equation 7.14.3, we find $L' \cong 370$ nH/m.

This page titled 7.14: Inductance of a Coaxial Structure is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.14: Inductance of a Coaxial Structure** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.15: Magnetic Energy

Consider a structure exhibiting inductance; i.e., one that is able to store energy in a magnetic field in response to an applied current. This structure could be a coil, or it could be one of a variety of inductive structures that are not explicitly intended to be an inductor; for example, a coaxial transmission line. When current is applied, the current-bearing elements of the structure exert forces on each other. Since these elements are not normally free to move, we may interpret this force as potential energy stored in the magnetic field associated with the current (Section 7.12).

We now want to know how much energy is stored in this field. The answer to this question has relevance in several engineering applications. One issue is that any system that includes inductance is using some fraction of the energy delivered by the power supply to energize this inductance. In many electronic systems – in power systems in particular – inductors are periodically energized and de-energized at a regular rate. Since power is energy per unit time, this consumes power. Therefore, energy storage in inductors contributes to the power consumption of electrical systems.

The stored energy is most easily determined using circuit theory concepts. First, we note that the electrical potential difference $v(t)$ (units of V) across an inductor is related to the current $i(t)$ (units of A) through the inductor as follows (Section 7.12):

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

where L (units of H) is the inductance. The instantaneous power associated with the device is

$$p(t) = v(t)i(t)$$

Energy (units of J) is power (units of J/s) integrated over time. Let W_m be the energy stored in the inductor. At some time t_0 in the past, $i(t_0) = 0$ and $W_m = 0$. As current is applied, W_m increases monotonically. At the present time t, $i(t) = I$. Thus, the present value of the magnetic energy is:

$$W_m = \int_{t_0}^{t_0+t} p(\tau)d\tau$$

Now evaluating this integral using the relationships established above:

$$\begin{aligned} W_m &= \int_{t_0}^{t+t_0} v(\tau)i(\tau)d\tau \\ &= \int_{t_0}^{t+t_0} \left[L\frac{d}{d\tau}i(\tau)\right] i(\tau)d\tau \\ &= L\int_{t_0}^{t+t_0} \left[\frac{d}{d\tau}i(\tau)\right] i(\tau)d\tau \end{aligned} \tag{7.15.1}$$

Changing the variable of integration from τ (and $d\tau$) to i (and di) we have

$$\begin{aligned} W_m &= L\int_{t_0}^{t+t_0} \frac{di}{d\tau} id\tau \\ &= L\int_0^I idi \end{aligned} \tag{7.15.2}$$

Evaluating the integral we obtain the desired expression

$$\boxed{W_m = \frac{1}{2}LI^2} \tag{7.15.3}$$

> The energy stored in an inductor in response to a steady current I is Equation 7.15.3. This energy increases in proportion to inductance and in proportion to the square of current.

The long straight coil (Section 7.13) is representative of a large number of practical applications, so it is useful to interpret the above findings in terms of this structure in particular. For this structure we found

$$L = \frac{\mu N^2 A}{l}$$

where μ is the permeability, N is the number of windings, A is cross-sectional area, and l is length. The magnetic field intensity inside this structure is related to I by (Section 7.6):

$$H = \frac{NI}{l}$$

Substituting these expressions into Equation 7.15.3, we obtain

$$\begin{aligned} W_m &= \frac{1}{2}\left[\frac{\mu N^2 A}{l}\right]\left[\frac{Hl}{N}\right]^2 \\ &= \frac{1}{2}\mu H^2 Al \end{aligned} \tag{7.15.4}$$

Recall that the magnetic field inside a long coil is approximately uniform. Therefore, the density of energy stored inside the coil is approximately uniform. Noting that the product Al is the volume inside the coil, we find that this energy density is W_m/Al; thus:

$$w_m = \frac{1}{2}\mu H^2 \tag{7.15.5}$$

which has the expected units of energy per unit volume (J/m^3).

The above expression provides an alternative method to compute the total magnetostatic energy in *any* structure. Within a mathematical volume $\mathcal{V}$, the total magnetostatic energy is simply the integral of the energy density over $\mathcal{V}$; i.e.,

$$W_m = \int_{\mathcal{V}} w_m \, dv$$

This works even if the magnetic field and the permeability vary with position. Substituting Equation 7.15.5 we obtain:

$$\boxed{W_m = \frac{1}{2}\int_{\mathcal{V}} \mu H^2 dv} \tag{7.15.6}$$

Summarizing:

> The energy stored by the magnetic field present within any defined volume is given by Equation 7.15.6.

It's worth noting that this energy increases with the permeability of the medium, which makes sense since inductance is proportional to permeability.

Finally, we reiterate that although we arrived at this result using the example of the long straight coil, Equations 7.15.5 and 7.15.6 are completely general.

This page titled 7.15: Magnetic Energy is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.15: Magnetic Energy** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

7.16: Magnetic Materials

As noted in Section 2.5, magnetic fields arise in the presence of moving charge (i.e., current) and in the presence of certain materials. In this section, we address these "magnetic materials."

A magnetic material may be defined as a substance that exhibits permeability μ (Section 2.6) that is significantly different from the permeability of free space μ_0. Since the magnetic flux density $\mathbf{B}$ is related to the magnetic field intensity $\mathbf{H}$ via $\mathbf{B} = \mu\mathbf{H}$, magnetic materials may exhibit magnetic flux density in response to a given magnetic field intensity that is significantly greater than that of other materials. Magnetic materials are also said to be "magnetizable," meaning that the application of a magnetic field causes the material itself to become a source of the magnetic field.

Magnetic media are typically metals, semiconductors, or heterogeneous media containing such materials. An example is ferrite, which consists of iron particles suspended in a ceramic. Magnetic media are commonly classified according to the physical mechanism responsible for their magnetizability. These mechanisms include *paramagnetism*, *diamagnetism*, and *ferromagnetism*. All three of these mechanisms involve quantum mechanical processes operating at the atomic and subatomic level, and are not well-explained by classical physics. These processes are beyond the scope of this book (but information is available via "Additional References" at the end of this section). However, it is possible to identify some readily-observable differences between these categories of magnetic media.

Paramagnetic and diamagnetic materials exhibit permeability that is only very slightly different than μ_0 and typically by much less than 0.01%. These materials exhibit very weak and temporary magnetization. The principal distinction between paramagnetic and diamagnetic media is in the persistence and orientation of induced magnetic fields. Paramagnetic materials – including aluminum, magnesium, and platinum – exhibit a very weak persistent magnetic field, and the magnetic field induced in the material is aligned in the same direction as the impressed (external) magnetic field. Diamagnetic materials – including copper, gold, and silicon – do not exhibit a persistent magnetic field, and the magnetic field induced in the material is (counter to intuition!) aligned in the *opposite* direction as the impressed magnetic field. The magnetization of paramagnetic and diamagnetic media is typically so weak that it is not often a consideration in engineering analysis and design.

> Paramagnetic and diamagnetic media exhibit permeability only very slightly different than that of free space, with little or no magnetization.

Ferromagnetic materials, on the other hand, exhibit permeability that can be many orders of magnitude greater than μ_0. (See Appendix A.2 for some example values.) These materials can be readily and indefinitely magnetized, thus, permanent magnets are typically comprised of ferromagnetic materials. Commonly-encountered ferromagnetic materials include iron, nickel, and cobalt.

Ferromagnetic materials are significantly non-linear (see definition in Section 2.8), exhibiting *saturation* and *hysteresis*. This is illustrated in Figure 7.16.1. In this plot, the origin represents a ferromagnetic material that is unmagnetized and in a region free of an external magnetic field. The external magnetic field is quantified in terms of $\mathbf{H}$, plotted along the horizontal axis. As the external field is increased, so to is $\mathbf{B}$ in the material, according to the relationship $\mathbf{B} = \mu\mathbf{H}$. Right away we see the material is non-linear, since the slope of the curve – and hence μ – is not constant.

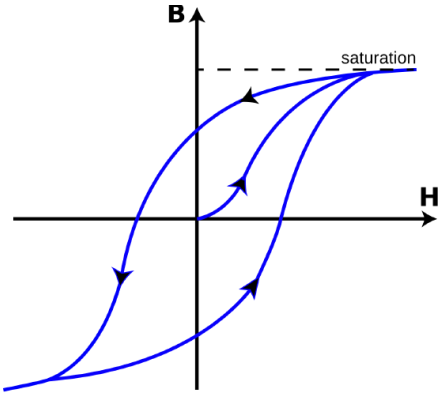


Figure 7.16.1: Non-linearity in a ferromagnetic material manifesting as saturation and hysteresis. ((modified) CC BY SA 3.0; Ndthe)

Once the external magnetizing field $\mathbf{H}$ exceeds a certain value, the response field $\mathbf{B}$ no longer significantly increases. This is saturation. Once saturated, further increases in the external field result do not significantly increase the magnetization of the material, so there is no significant increase in $\mathbf{B}$.

From this state of saturation, let us now reduce the external field. We find that the rate of decrease in $\mathbf{B}$ with respect to $\mathbf{H}$ is significantly less than the rate that $\mathbf{B}$ originally increased with respect to $\mathbf{H}$. In fact, $\mathbf{B}$ is still greater than zero even when $\mathbf{H}$ has been reduced to zero. At this point, the magnetization of the material is obvious, and a device comprised of this material could be used as a magnet.

If we now apply an external field in the reverse direction, we find that we are eventually able to zero and then redirect the response field. As we continue to decrease $\mathbf{H}$ (that is, increase the magnitude in the reverse direction), we once again reach saturation.

The same behavior is observed when we once again increase $\mathbf{H}$. The material is eventually demagnetized, remagnetized in the opposite direction and then saturated in that direction. At this point, it is apparent that a return to the start condition ($\mathbf{H} = \mathbf{B} = 0$; i.e., demagnetized when there is no external field) is not possible.

Hysteresis is the name that we apply to this particular form of non-linear behavior. Hysteresis has important implications in engineering applications. First, as identified above, it is an important consideration in the analysis and design of magnets. In applications where a ferromagnetic material is being used because high permeability is desired – e.g., in inductors (Section 7.12) and transformers (Section 8.5) – hysteresis complicates the design and imposes limits on the performance of the device.

Hysteresis may also be exploited as a form of memory. This is apparent from Figure 7.16.1. If $\mathbf{B} > 0$, then recent values of $\mathbf{H}$ must have been relatively large and positive. Similarly, If $\mathbf{B} < 0$, then recent values of $\mathbf{H}$ must have been relatively large and negative. Furthermore, the most recent sign of $\mathbf{H}$ can be inferred even if the present value of $\mathbf{H}$ is zero. In this sense, the material "remembers" the past history of its magnetization and thereby exhibits memory. This is the enabling principle for a number of digital data storage devices, including hard drives (see "Additional Reading" at the end of this section). Summarizing:

> Ferromagnetic media exhibit permeability μ that is orders of magnitude greater than that of free space and are readily magnetizable. These materials are also nonlinear in μ, which manifests as saturation and hysteresis.

This page titled 7.16: Magnetic Materials is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **7.16: Magnetic Materials** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

CHAPTER OVERVIEW

8: Time-Varying Fields

8.1: Comparison of Static and Time-Varying Electromagnetics
8.2: Electromagnetic Induction
8.3: Faraday's Law
8.4: Induction in a Motionless Loop
8.5: Transformers - Principle of Operation
8.6: Transformers as Two-Port Devices
8.7: The Electric Generator
8.8: The Maxwell-Faraday Equation
8.9: Displacement Current and Ampere's Law

Contributors and Attributions

- Ellingson, Steven W. (2018) Electromagnetics, Vol. 1. Blacksburg, VA: VT Publishing. https://doi.org/10.21061/electromagnetics-vol-1 Licensed with CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0. Report adoption of this book here. If you are a professor reviewing, adopting, or adapting this textbook please help us understand a little more about your use by filling out this form.

This page titled 8: Time-Varying Fields is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

8.1: Comparison of Static and Time-Varying Electromagnetics

Students encountering time-varying electromagnetic fields for the first time have usually been exposed to electrostatics and magnetostatics already. These disciplines exhibit many similarities as summarized in Table 8.1.1. The principles of time-varying electromagnetics presented in this table are all formally introduced in other sections; the sole purpose of this table is to point out the differences. We can summarize the differences as follows:

> Maxwell's Equations in the general (time-varying) case include extra terms that do not appear in the equations describing electrostatics and magnetostatics. These terms involve time derivatives of fields and describe coupling between electric and magnetic fields.

The coupling between electric and magnetic fields in the time-varying case has one profound consequence in particular. It becomes possible for fields to continue to exist even after their sources – i.e., charges and currents – are turned off. What kind of field can continue to exist in the absence of a source? Such a field is commonly called a wave. Examples of waves include signals in transmission lines and signals propagating away from an antenna.

Table 8.1.1: Comparison of principles governing static and time-varying electromagnetic fields. Differences in the time-varying case relative to the static case are highlighted in blue.

	Electrostatics / Magnetostatics	**Time-Varying (Dynamic)**
Electric & magnetic fields are...	*independent*	*possibly coupled*
Maxwell's Eqns. (integral)	$\oint_{\mathcal{S}} \mathbf{D}\cdot d\mathbf{s} = Q_{encl}$	$\oint_{\mathcal{S}} \mathbf{D}\cdot d\mathbf{s} = Q_{encl}$
	$\oint_{\mathcal{C}} \mathbf{E}\cdot d\mathbf{l} = 0$	$\oint_{\mathcal{C}} \mathbf{E}\cdot d\mathbf{l} = -\frac{\partial}{\partial t}\int_{\mathcal{S}} \mathbf{B}\cdot d\mathbf{s}$
	$\oint_{\mathcal{S}} \mathbf{B}\cdot d\mathbf{s} = 0$	$\oint_{\mathcal{S}} \mathbf{B}\cdot d\mathbf{s} = 0$
	$\oint_{\mathcal{C}} \mathbf{H}\cdot d\mathbf{l} = I_{encl}$	$\oint_{\mathcal{C}} \mathbf{H}\cdot d\mathbf{l} = I_{encl} + \int_{\mathcal{S}} \frac{\partial}{\partial t}\mathbf{D}\cdot d\mathbf{s}$
Maxwell's Eqns. (differential)	$\nabla\cdot\mathbf{D} = \rho_v$	$\nabla\cdot\mathbf{D} = \rho_v$
	$\nabla\times\mathbf{E} = 0$	$\nabla\times\mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B}$
	$\nabla\cdot\mathbf{B} = 0$	$\nabla\cdot\mathbf{B} = 0$
	$\nabla\times\mathbf{H} = \mathbf{J}$	$\nabla\times\mathbf{H} = \mathbf{J} + \frac{\partial}{\partial t}\mathbf{D}$

This page titled 8.1: Comparison of Static and Time-Varying Electromagnetics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **8.1: Comparison of Static and Time-Varying Electromagnetics** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

8.2: Electromagnetic Induction

When an electrically-conducting structure is exposed to a time-varying magnetic field, an electrical potential difference is induced across the structure. This phenomenon is known as *electromagnetic induction*. A convenient introduction to electromagnetic induction is provided by Lenz's Law. This section explains electromagnetic induction in the context of Lenz's Law and provides two examples.

Let us begin with the example depicted in Figure 8.2.1, involving a cylindrical coil. Attached to the terminals of the coil is a resistor for which we may identify an electric potential difference V and current I. In this particular case, the sign conventions indicated for V and I are arbitrary, but it is important to be consistent once they are established.

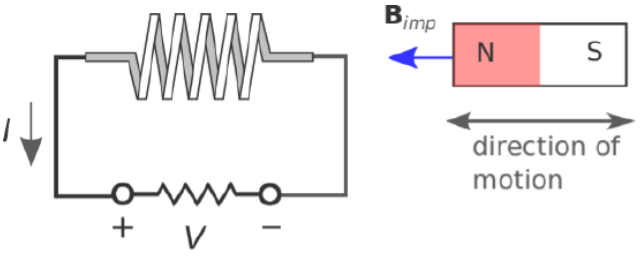


Figure 8.2.1: An experiment demonstrating electromagnetic induction and Lenz's Law. ((modified) CC BY 4.0; Y. Qin)

Now let us introduce a bar magnet as shown in Figure 8.2.1. The magnet is centered along the axis of coil, to the right of the coil, and with its north pole facing toward the coil. The magnet is responsible for the magnetic flux density $\mathbf{B}_{imp}$. We refer to $\mathbf{B}_{imp}$ as an *impressed* magnetic field because this field exists independently of any response that may be induced by interaction with the coil. Note that $\mathbf{B}_{imp}$ points to the left inside the coil.

The experiment consists of three tests. We will find in two of these tests that current flows (i.e., $|I| > 0$), and subsequently, there is an *induced* magnetic field $\mathbf{B}_{ind}$ due to this current. It is the direction of the current and subsequently the direction of $\mathbf{B}_{ind}$ inside the coil that we wish to observe. The findings are summarized below and in Table 8.2.1.

- When the magnet is motionless, we have the unsurprising result that there is no current in the coil. Therefore, no magnetic field is induced, and the total magnetic field is simply equal to $\mathbf{B}_{imp}$.
- When the magnet moves *toward* the coil, we observe current that is positive with respect to the reference direction indicated in Figure 8.2.1. This current creates an induced magnetic field $\mathbf{B}_{ind}$ that points to the *right*, as predicted by magnetostatic considerations from the right-hand rule. Since $\mathbf{B}_{imp}$ points to the left, it appears that the induced current is opposing the increase in the magnitude of the total magnetic field.
- When the magnet moves *away* from the coil, we observe current that is negative with respect to the reference direction indicated in Figure 8.2.1. This current yields $\mathbf{B}_{ind}$ that points to the *left*. Since $\mathbf{B}_{imp}$ points to the left, it appears that the induced current is opposing the decrease in the magnitude of the total magnetic field.

Table 8.2.1: Results of the experiment associated with Figure 8.2.1

Magnet is ...	$\lvert\mathbf{B}_{imp}\rvert$ in coil is ...	Circuit Response	$\mathbf{B}_{ind}$ inside coil
Motionless	constant	$V = 0$, $I = 0$	none
Moving toward coil	increasing	$V > 0$, $I > 0$	Pointing right
Moving away from coil	decreasing	$V < 0$, $I < 0$	Pointing left

The first conclusion one may draw from this experiment is that changes in the magnetic field can induce current. This was the claim made in the first paragraph of this section and is a consequence of Faraday's Law, which is tackled in detail in Section 8.3. The second conclusion – also associated with Faraday's Law – is the point of this section: The induced magnetic field – that is, the one due to the current induced in the coil – always opposes the change in the impressed magnetic field. Generalizing:

> Lenz's Law states that the current that is induced by a change in an impressed magnetic field creates an induced magnetic field that opposes (acts to reduce the effect of) the change in the total magnetic field.

When the magnet moves, three things happen: (1) A current is induced, (2) A magnetic field is induced (which adds to the impressed magnetic field), and (3) the value of V becomes non-zero. Lenz's Law does not address which of these are responding directly to the change in the impressed magnetic field, and which of these are simply responding to changes in the other quantities. Lenz's Law may leave you with the incorrect impression that it is I that is induced, and that $\mathbf{B}_{ind}$ and V are simply responding to this current. In truth, the quantity that is induced is actually V. This can be verified in the above experiment by replacing the resistor with a high-impedance voltmeter, which will indicate that V is changing even though there is negligible current flow. Current flow is simply a response to the induced potential. Nevertheless, it is common to say informally that "I is induced," even if it is only indirectly through V.

So, if Lenz's Law is simply an observation and not an explanation of the underlying physics, then what is it good for? Lenz's Law is often useful for quickly determining the direction of current flow in practical electromagnetic induction problems, without resorting to the mathematics associated with Faraday's Law. Here's an example:

Example 8.2.1: Electromagnetic induction through a transformer

Figure 8.2.2 shows a rudimentary circuit consisting of a battery and a switch on the left, a voltmeter on the right, and a transformer linking the two.

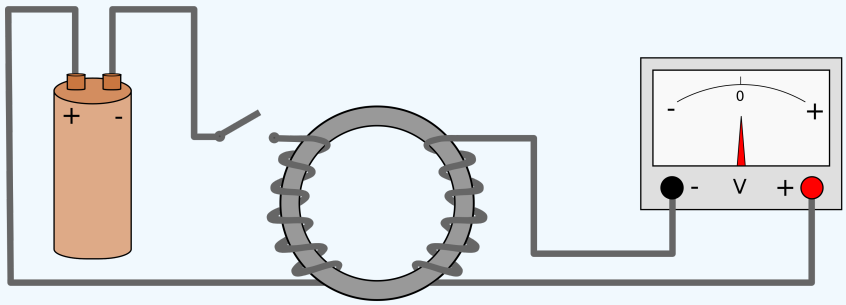


Figure 8.2.2: Electromagnetic induction through a transformer. (CC0 1.0 (modified))

It is not necessary to be familiar with transformers to follow this example; suffice it to say, the transformer considered here consists of two coils wound around a common toroidal core, which serves to contain magnetic flux. In this way, the flux generated by either coil is delivered to the other coil with negligible loss.

The experiment begins with the switch on the left in the open state. Thus, there is no current and no magnetic field apparent in the coil on the left. The voltmeter reads 0 V. When the switch is closed, what happens?

Solution

Closing the switch creates a current in the coil on the left. Given the indicated polarity of the battery, this current flows counter-clockwise through the circuit on the left, with current arriving in the left coil through the bottom terminal. Given the indicated direction of the winding in the left coil, the impressed magnetic field $\mathbf{B}_{imp}$ is oriented counter-clockwise through the toroidal core. The coil on the right "sees" $\mathbf{B}_{imp}$ increase from zero to some larger value. Since the voltmeter presumably has input high impedance, negligible current flows. However, if current were able to flow, Lenz's Law dictates that it would be induced to flow in a counter-clockwise direction around the circuit on the right, since the induced magnetic field $\mathbf{B}_{ind}$ would then be clockwise-directed so as to oppose the increase in $\mathbf{B}_{imp}$. Therefore, the potential measured at the bottom of the right coil would be higher than the potential at the top of the right coil. The figure indicates that the voltmeter measures the potential at its right terminal relative to its left terminal, so the needle will deflect to the right. This deflection will be temporary, since the current provided by the battery becomes constant at a new non-zero value and $\mathbf{B}_{ind}$ responds only to the *change* in $\mathbf{B}_{imp}$. The voltmeter reading will remain at zero for as long as the switch remains closed and the current remains steady.

Here are some follow up exercises to test your understanding of what is going on: (1) Now open the switch. What happens? (2) Repeat the original experiment, but before starting, swap the terminals on the battery.

Finally, it is worth noting that Lenz's law can also be deduced from the principle of conservation of energy. The argument is that if the induced magnetic field *reinforced* the change in the impressed magnetic field, then the sum magnetic field would increase. This would result in a further increase in the induced magnetic field, leading to a positive feedback situation. However, positive

feedback cannot be supported without an external source of energy, leading to a logical contradiction. In other words, the principle of conservation of energy requires the *negative* feedback described by Lenz's law.

This page titled 8.2: Electromagnetic Induction is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **8.2: Electromagnetic Induction** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

8.3: Faraday's Law

Faraday's Law describes the generation of electric potential by a time-varying magnetic flux. This is a form of *electromagnetic induction*.

To begin, consider the scenario shown in Figure 8.3.1. A single loop of wire in the presence of an impressed magnetic field $\mathbf{B}$. For reasons explained later, we introduce a small gap and define V_T to be the potential difference measured across the gap according to the sign convention indicated. The resistance R may be any value greater than zero, including infinity; i.e., a literal gap.

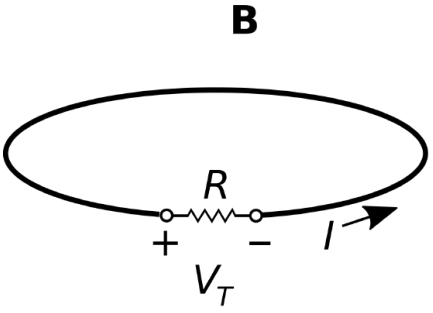


Figure 8.3.1: A single loop of wire in the presence of an impressed magnetic field.

As long as R is not infinite, we know from Lenz's Law (Section 8.2) to expect that a time-varying magnetic field will cause a current to flow in the wire. Lenz's Law also tells us the direction in which the current will flow. However, Lenz's Law does not tell us the magnitude of the current, and it sidesteps some important physics that has profound implications for the analysis and design of electrical devices, including generators and transformers.

The more complete picture is given by Faraday's Law. In terms of the scenario of Figure 8.3.1, Faraday's Law relates the potential V_T induced by the time variation of $\mathbf{B}$. V_T then gives rise to the current identified in Lenz's Law. The magnitude of this current is simply V_T/R. Without further ado, here's Faraday's Law for this single loop scenario:

$$V_T = -\frac{\partial}{\partial t}\Phi \quad \text{(single loop)} \tag{8.3.1}$$

Here Φ is the magnetic flux (units of Wb) associated with any open surface $\mathcal{S}$ bounded by the loop:

$$\Phi = \int_{\mathcal{S}} \mathbf{B} \cdot d\mathbf{s} \tag{8.3.2}$$

where $\mathbf{B}$ is magnetic flux density (units of T or Wb/m^2) and $d\mathbf{s}$ is the differential surface area vector.

To make headway with Faraday's Law, one must be clear about the meanings of $\mathcal{S}$ and $d\mathbf{s}$. If the wire loop in the present scenario lies in a plane, then a good choice for $\mathcal{S}$ is the simply the planar area bounded by the loop. However, *any* surface that is bounded by the loop will work, including non-planar surfaces that extend above and/or below the plane of the loop. All that is required is that every magnetic field line that passes through the loop also passes through $\mathcal{S}$. This happens automatically if the curve $\mathcal{C}$ defining the edge of the open surface $\mathcal{S}$ corresponds to the loop. Subsequently, the magnitude of $d\mathbf{s}$ is the differential surface element ds and the direction of $d\mathbf{s}$ is the unit vector $\hat{\mathbf{n}}$ perpendicular to each point on $\mathcal{S}$, so $d\mathbf{s} = \hat{\mathbf{n}}ds$.

This leaves just one issue remaining – the orientation of $\hat{\mathbf{n}}$. This is sorted out in Figure 8.3.2. There are two possible ways for a vector to be perpendicular to a surface, and the direction chosen for $\hat{\mathbf{n}}$ will affect the sign of V_T. Therefore, $\hat{\mathbf{n}}$ must be somehow related to the polarity chosen for V_T. Let's consider this relationship. Let $\mathcal{C}$ begin at the "$-$" terminal of V_T and follow the entire perimeter of the loop, ending at the "$+$" terminal. Then, $\hat{\mathbf{n}}$ is determined by the following "right-hand rule:" $\hat{\mathbf{n}}$ points in the direction of the curled fingers of the right hand when the thumb of the right hand is aligned in the direction of $\mathcal{C}$. It's worth noting that this convention is precisely the convention used to relate $\hat{\mathbf{n}}$ and $\mathcal{C}$ in Stokes' Theorem.

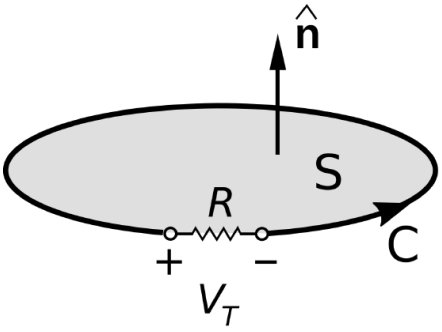


Figure 8.3.2: Relationship between the polarity of V_T and orientations of $\mathcal{C}$ and $\hat{\mathbf{n}}$ in the planar single-loop scenario.

Now let us recap how Faraday's Law is applied to the single loop scenario of Figure [m0055_fFaradaySingleLoop]:

1. Assign "+" and "−" terminals to the gap voltage V_T.
2. The orientation of $\hat{\mathbf{n}}$ is determined by the right hand rule, taking the direction of $\mathcal{C}$ to be the perimeter of the loop beginning at "−" and ending at "+"
3. $\mathbf{B}$ yields a magnetic flux Φ associated with the loop according to Equation 8.3.2. $\mathcal{S}$ is any open surface that intersects all the magnetic field lines that pass through the loop (so you might as well choose $\mathcal{S}$ in a way that results in the simplest possible integration).
4. By Faraday's Law (Equation 8.3.1), V_T is the time derivative of Φ with a change of sign.
5. The current I flowing in the loop is V_T/R, with the reference direction (i.e., direction of positive current) being from "+" to "−" through the resistor. Think of the loop as a voltage source, and you'll get the correct reference direction for I.

At this point, let us reiterate that the electromagnetic induction described by Faraday's Law induces *potential* (in this case, V_T) and not *current*. The current in the loop is simply the induced voltage divided by the resistance of the loop. This point is easily lost, especially in light of Lenz's Law, which seems to imply that current, as opposed to potential, is the thing being induced.

Wondering about the significance of the minus sign in Equation 8.3.1? That is specifically Lenz's Law: The current I that ends up circulating in the loop generates its own magnetic field ("$\mathbf{B}_{ind}$" in Section 8.2), which is distinct from the impressed magnetic field $\mathbf{B}$ and which tends to oppose change in $\mathbf{B}$. Thus, we see that Faraday's Law subsumes Lenz's Law.

Frequently one is interested in a structure that consists of multiple identical loops. We have in mind here something like a coil with $N \geq 1$ windings tightly packed together. In this case, Faraday's Law is

$$\boxed{V_T = -N\frac{\partial}{\partial t}\Phi} \tag{8.3.3}$$

Note that the difference is simply that the gap potential V_T is greater by N. In other words, each winding of the coil contributes a potential given by Equation 8.3.1, and these potentials add in series.

> Faraday's Law, given in general by Equation 8.3.3, states that the potential induced in a coil is proportional to the time derivative of the magnetic flux through the coil.

The induced potential V_T is often referred to as "emf," which is a contraction of the term *electromotive force* – a misnomer to be sure, since no actual force is implied, only potential. The term "emf" is nevertheless frequently used in the context of Faraday's Law applications for historical reasons.

Previously in this section, we considered the generation of emf by time variation of $\mathbf{B}$. However, Equation 8.3.1 indicates that what actually happens is that the emf is the result of time variation of the magnetic flux, Φ. Magnetic flux is magnetic flux density integrated over area, so it appears that emf can also be generated simply by varying $\mathcal{S}$, independently of any time variation of $\mathbf{B}$. In other words, emf may be generated even when $\mathbf{B}$ is constant, by instead varying the shape or orientation of the coil. So, we have a variety of schemes by which we can generate emf. Here they are:

1. Time-varying $\mathbf{B}$ (as we considered previously in this section). For example, $\mathbf{B}$ might be due to a permanent magnet that is moved (e.g., translated or rotated) in the vicinity of the coil, as described in Section 8.2. Or, more commonly, $\mathbf{B}$ might be due to

a different coil that bears a time-varying current. These mechanisms are collectively referred to as *transformer emf*. Transformer emf is the underlying principle of operation of transformers; for more on this see Section 8.5.

2. The perimeter $\mathcal{C}$ – and thus the surface $\mathcal{S}$ over which Φ is determined – can be time-varying. For example, a wire loop might be rotated or changed in shape in the presence of a constant magnetic field. This mechanism is referred to as *motional emf* and is the underlying principle of operation of generators (Section 8.7).
3. Transformer and motional emf can exist in various combinations. From the perspective of Faraday's Law, transformer and motional emf are the same in the sense in either case Φ is time-varying, which is all that is required to generate emf.

Finally, a comment on the generality of Faraday's Law. Above we have introduced Faraday's Law as if it were specific to loops and coils of wire. However, the truth of the matter is that Faraday's Law is fundamental physics. If you can define a closed path – current-bearing or not – then you can compute the potential difference achieved by traversing that path using Faraday's Law. The value you compute is the potential associated with electromagnetic induction, and exists independently and in addition to the potential difference associated with the static electric field (e.g., Section 5.12). In other words:

> Faraday's Law indicates the contribution of electromagnetic induction (the generation of emf by a time-varying magnetic flux) to the potential difference achieved by traversing a closed path.

In Section 8.8, this insight is used to transform the static form of Kirchoff's Voltage Law (Section 5.10) – which gives the potential difference associated with electric field only – into the *Maxwell-Faraday Equation* (Section 8.8), which is a general statement about the relationship between the instantaneous value of the electric field and the time derivative of the magnetic field.

This page titled 8.3: Faraday's Law is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **8.3: Faraday's Law** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

8.4: Induction in a Motionless Loop

In this section, we consider the problem depicted in Figure 8.4.1, which is a single motionless loop of wire in the presence of a spatially-uniform but time-varying magnetic field. A small gap is introduced in the loop, allowing us to measure the induced potential V_T. Additionally, a resistance R is connected across V_T in order to allow a current to flow. This problem was considered in Section 8.3 as an introduction to Faraday's Law; in this section, we shall actually work the problem and calculate some values. This is intended to serve as an example of the application of Faraday's Law, a demonstration of transformer emf, and will serve as a first step toward an understanding of transformers as devices.

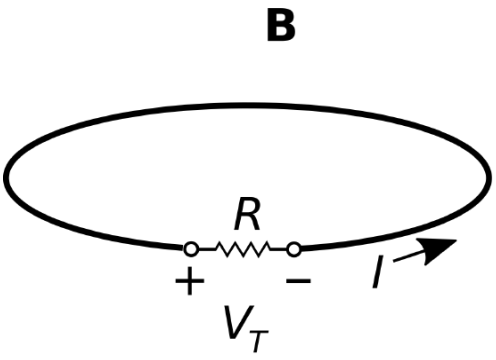


Figure 8.4.1: A single loop of wire in the presence of an impressed spatially-uniform but time-varying magnetic field.

In the present problem, the loop is centered in the $z = 0$ plane. The magnetic flux density is $\mathbf{B} = \hat{\mathbf{b}}B(t)$; i.e., time-varying magnitude $B(t)$ and a constant direction $\hat{\mathbf{b}}$. Because this magnetic field is spatially uniform (i.e., the same everywhere), we will find that only the area of the loop is important, and not it's specific shape. For this reason, it will not be necessary to specify the radius of the loop or even require that it be a circular loop. Our task is to find expressions for V_T and I.

To begin, remember that Faraday's Law is a calculation of electric potential and not current. So, the approach is to first find V_T, and then find the current I that flows through the gap resistance in response.

The sign convention for V_T is arbitrary; here, we have selected "+" and "−" terminals as indicated in Figure 8.4.1.[1] Following the standard convention for the reference direction of current through a passive device, I should be directed as shown in Figure 8.4.1. It is worth repeating that these conventions for the signs of V_T and I are merely *references*; for example, we may well find that I is negative, which means that current flows in a clockwise direction in the loop.

We now invoke Faraday's Law:

$$V_T = -N\frac{\partial}{\partial t}\Phi$$

The number of windings N in the loop is 1, and Φ is the magnetic flux through the loop. Thus:

$$V_T = -\frac{\partial}{\partial t}\int_{\mathcal{S}} \mathbf{B}\cdot d\mathbf{s}$$

where $\mathcal{S}$ is any open surface that intersects all of the magnetic field lines that pass through the loop. The simplest such surface is simply the planar surface defined by the perimeter of the loop. Then $d\mathbf{s} = \hat{\mathbf{n}}ds$, where ds is the differential surface element and $\hat{\mathbf{n}}$ is the normal to the plane of the loop. Which of the two possible normals to the loop? This is determined by the right-hand rule of Stokes' Theorem. From the "−" terminal, we point the thumb of the right hand in the direction that leads to the "+" terminal by traversing the perimeter of the loop. When we do this, the curled fingers of the right hand intersect $\mathcal{S}$ in the same direction as $\hat{\mathbf{n}}$. To maintain the generality of results derived below, we shall not make the substitution $\hat{\mathbf{n}} = +\hat{\mathbf{z}}$; nevertheless we see this is the case for a loop parallel to the $z = 0$ plane with the polarity of V_T indicated in Figure 8.4.1.

Taking this all into account, we have

$$\begin{aligned} V_T &= -\frac{\partial}{\partial t}\int_S (\hat{\mathbf{b}}B(t))\cdot(\hat{\mathbf{n}}ds) \\ &= -(\hat{\mathbf{b}}\cdot\hat{\mathbf{n}})\frac{\partial}{\partial t}\int_{\mathcal{S}} B(t)ds \end{aligned} \tag{8.4.1}$$

Since the magnetic field is uniform, $B(t)$ may be extracted from the integral. Furthermore, the shape and the orientation of the loop are time-invariant, so the remaining integral may be extracted from the time derivative operation. This leaves:

$$V_T = -(\hat{\mathbf{b}} \cdot \hat{\mathbf{n}}) \left(\frac{\partial}{\partial t} B(t) \right) \int_s ds$$

The integral in this expression is simply the area of the loop, which is a constant; let the symbol A represent this area. We obtain

$$\boxed{V_T = -\left(\hat{\mathbf{b}} \cdot \hat{\mathbf{n}} A \right) \frac{\partial}{\partial t} B(t)} \tag{8.4.2}$$

which is the expression we seek. Note that the quantity $\hat{\mathbf{b}} \cdot \hat{\mathbf{n}} A$ is the *projected area* of the loop. The projected area is equal to A when the the magnetic field lines are perpendicular to the loop (i.e., $\hat{\mathbf{b}} = \hat{\mathbf{n}}$), and decreases to zero as $\hat{\mathbf{b}} \cdot \hat{\mathbf{n}} \to 0$. Summarizing:

> The magnitude of the transformer emf induced by a spatially-uniform magnetic field is equal to the projected area times the time rate of change of the magnetic flux density, with a change of sign. (Equation 8.4.2).

A few observations about this result:

- As promised earlier, we have found that the shape of the loop is irrelevant; i.e., a *square* loop having the same area and planar orientation would result in the same V_T. This is because the magnetic field is spatially uniform, and because it is the magnetic flux (Φ) and not the magnetic field or shape of the loop alone that determines the induced potential.
- The induced potential is proportional to A; i.e., V_T can be increased by increasing the area of the loop.
- The peak magnitude of the induced potential is maximized when the plane of the loop is perpendicular to the magnetic field lines.
- The induced potential goes to zero when the plane of the loop is parallel to the magnetic field lines. Said another way, there is no induction unless magnetic field lines pass through the loop.
- The induced potential is proportional to the rate of change of $\mathbf{B}$. If $\mathbf{B}$ is constant in time, then there is no induction.

Finally, the current in the loop is simply

$$I = \frac{V_T}{R}$$

Again, electromagnetic induction induces potential, and the current flows only in response to the induced potential as determined by Ohm's Law. In particular, if the resistor is removed, then $R \to \infty$ and $I \to 0$, but V_T is unchanged.

One final comment is that even though the current I is not a *direct* result of electromagnetic induction, we can use I as a check of the result using Lenz's Law (Section 8.2). We'll demonstrate this in the example below.

Example 8.4.1: Induction in a motionless circular loop by a linearly-increasing magnetic field

Let the loop be planar in the $z = 0$ plane and circular with radius $a = 10$ cm. Let the magnetic field be $\hat{\mathbf{z}} B(t)$ where

$$\begin{aligned} B(t) &= 0 \quad , t < 0 \\ &= B_0 t / t_0 , 0 \le t \le t_0 \\ &= B_0 \quad , t > t_0 \end{aligned} \tag{8.4.3}$$

i.e., $B(t)$ begins at zero and increases linearly to B_0 at time t_0, after which it remains constant at B_0. Let $B_0 = 0.2$ T, $t_0 = 3$ s, and let the loop be closed by a resistor $R = 1$ kΩ. What current I flows in the loop?

Solution

Adopting the sign conventions of Figure 8.4.1 we first note that $\hat{\mathbf{n}} = +\hat{\mathbf{z}}$; this is determined by the right-hand rule with respect to the indicated polarity of V_T. Thus, Equation 8.4.2 becomes

$$V_T = -\left(\hat{\mathbf{b}} \cdot \hat{\mathbf{z}} A \right) \frac{\partial}{\partial t} B(t)$$

Note $\hat{\mathbf{b}} \cdot \hat{\mathbf{z}} A = A$ since $\hat{\mathbf{b}} = \hat{\mathbf{z}}$; i.e., because the plane of the loop is perpendicular to the magnetic field lines. Since the loop is circular, $A = \pi a^2$. Also

$$\frac{\partial}{}$$

$$\begin{aligned}\frac{\partial}{\partial t}B(t) &= 0 \quad , t < 0 \\ &= B_0/t_0, 0 \le t \le t_0 \\ &= 0 \quad , t > t_0\end{aligned} \tag{8.4.4}$$

Putting this all together:

$$V_T = -\pi a^2 \frac{B_0}{t_0} = -2.09 \text{ mV} \ , \ 0 \le t \le t_0$$

and $V_T = 0$ before and after this time interval, since $\mathbf{B}$ is constant during those times. Subsequently the induced current is

$$I = \frac{V_T}{R} = -2.09 \ \mu\text{A} \ , \ 0 \le t \le t_0$$

and $I = 0$ before and after this time interval. We have found that the induced current is a constant clockwise flow that exists only while $\mathbf{B}$ is increasing.

Finally, let's see if the result is consistent with Lenz's Law. The current induced while $\mathbf{B}$ is changing gives rise to an induced magnetic field $\mathbf{B}_{ind}$. From the right-hand rule that relates the direction of I to the direction of $\mathbf{B}_{ind}$ (Section 7.5), the direction of $\mathbf{B}_{ind}$ is generally $-\hat{\mathbf{z}}$ inside the loop. In other words, the magnetic field associated with the induced current opposes the increasing impressed magnetic field that induced the current, in accordance with Lenz's Law.

Example 8.4.3: Induction in a motionless circular loop by a sinusoidally-varying magnetic field.

Let us repeat the previous example, but now with

$$B(t) = B_0 \sin 2\pi f_0 t$$

with $f_0 = 1$ kHz.

Solution

Now

$$\frac{\partial}{\partial t}B(t) = 2\pi f_0 B_0 \cos 2\pi f_0 t$$

So

$$V_T = -2\pi^2 f_0 a^2 B_0 \cos 2\pi f_0 t$$

Subsequently

$$I = \frac{V_T}{R} = -\frac{2\pi^2 f_0 a^2 B_0}{R} \cos 2\pi f_0 t$$

Substituting values, we have:

$$I = -(39.5 \text{ mA}) \cos\left[(6.28 \text{ krad/s})t\right]$$

It should be no surprise that V_T and I vary sinusoidally, since the source ($\mathbf{B}$) varies sinusoidally. A bit of useful trivia here is that V_T and I are 90° out of phase with the source. It is also worth noting what happens when $\mathbf{B}(t) = 0$. This occurs twice per period, at $t = n/2f_0$ where n is any integer, including $t = 0$. At these times $\mathbf{B}(t)$ is zero, but V_T and hence I_R are decidedly non-zero; in fact, they are at their maximum magnitude. Again, it is the *change* in $\mathbf{B}$ that induces voltage and subsequently current, not $\mathbf{B}$ itself.

1. A good exercise for the student is to repeat this problem with the terminal polarity reversed; one should obtain the same answer.↩

This page titled 8.4: Induction in a Motionless Loop is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **8.4: Induction in a Motionless Loop** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

8.5: Transformers - Principle of Operation

A transformer is a device that connects two electrical circuits through a shared magnetic field. Transformers are used in impedance transformation, voltage level conversion, circuit isolation, conversion between single-ended and differential signal modes, and other applications.[1] The underlying electromagnetic principle is Faraday's Law – in particular, transformer emf.

The essential features of a transformer can be derived from the simple experiment shown in Figures 8.5.1 and 8.5.2. In this experiment, two coils are arranged along a common axis. The winding pitch is small, so that all magnetic field lines pass through the length of the coil, and no lines pass between the windings. To further contain the magnetic field, we assume both coils are wound on the same core, consisting of some material exhibiting high permeability. The upper coil has N_1 turns and the lower coil has N_2 turns.

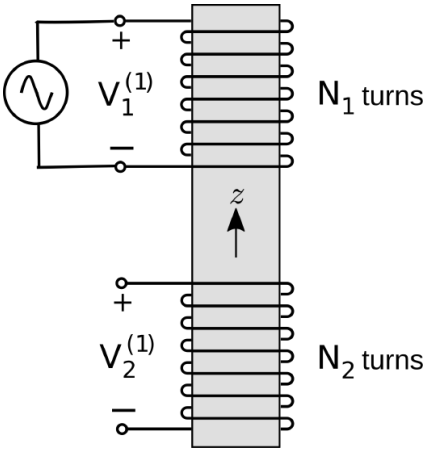


Figure 8.5.1: Part I of an experiment demonstrating the linking of electric circuits using a transformer.

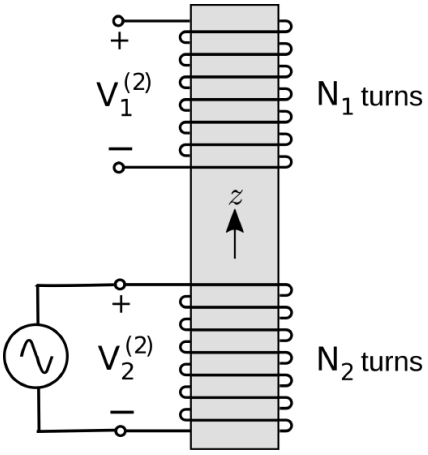


Figure 8.5.2: Part II of an experiment demonstrating the linking of electric circuits using a transformer.

In Part I of this experiment (Figure 8.5.1), the upper coil is connected to a sinusoidally-varying voltage source $V_1^{(1)}$ in which the subscript refers to the coil and the superscript refers to "Part I" of this experiment. The voltage source creates a current in the coil, which in turn creates a time-varying magnetic field $\mathbf{B}$ in the core.

The lower coil has N_2 turns wound in the *opposite direction* and is open-circuited. Given the closely-spaced windings and use of a high-permeability core, we assume that the magnetic field within the lower coil is equal to $\mathbf{B}$ generated in the upper coil. The potential induced in the lower coil is $V_2^{(1)}$ with reference polarity indicated in the figure. From Faraday's Law we have

$$V_2^{(1)} = -N_2 \frac{\partial}{\partial t} \Phi_2 \tag{8.5.1}$$

where Φ_2 is the flux through a single turn of the lower coil. Thus:

$$V_2^{(1)} = -N_2 \frac{\partial}{\partial t} \int_s \mathbf{B} \cdot (-\hat{\mathbf{z}} ds) \tag{8.5.2}$$

Note that the direction of $d\mathbf{s} = -\hat{\mathbf{z}}ds$ is determined by the polarity we have chosen for $V_2^{(1)}$.

In Part II of the experiment (Figure 8.5.2), we make changes as follows. We apply a voltage $V_2^{(2)}$ to the lower coil and open-circuit the upper coil. Further, we adjust $V_2^{(2)}$ such that the induced magnetic flux density is again $\mathbf{B}$ – that is, equal to the field in Part I of the experiment. Now we have

$$V_1^{(2)} = -N_1 \frac{\partial}{\partial t}\Phi_1$$

which is

$$V_1^{(2)} = -N_1 \frac{\partial}{\partial t}\int_{\mathcal{S}} \mathbf{B}\cdot(+\hat{\mathbf{z}}ds)$$

For reasons that will become apparent in a moment, let's shift the leading minus sign into the integral. We then have

$$V_1^{(2)} = +N_1 \frac{\partial}{\partial t}\int_{\mathcal{S}} \mathbf{B}\cdot(-\hat{\mathbf{z}}ds)$$

Comparing this to Equations 8.5.1 and 8.5.2, we see that we may rewrite this in terms of the flux in the lower coil in Part I of the experiment:

$$V_1^{(2)} = +N_1 \frac{\partial}{\partial t}\Phi_2$$

In fact, we can express this in terms of the *potential* in Part I of the experiment:

$$\begin{aligned} V_1^{(2)} &= \left(-\frac{N_1}{N_2}\right)\left(-N_2\frac{\partial}{\partial t}\Phi_2\right) \\ &= \left(-\frac{N_1}{N_2}\right)V_2^{(1)} \end{aligned} \tag{8.5.3}$$

We have found that the potential in the upper coil in Part II is related in a simple way to the potential in the lower coil in Part I of the experiment. If we had done Part II first, we would have obtained the same result but with the superscripts swapped. Therefore, it must be generally true – regardless of the arrangement of terminations – that

$$V_1 = -\frac{N_1}{N_2}V_2$$

This expression should be familiar from elementary circuit theory – except possibly for the minus sign. The minus sign is a consequence of the fact that the coils are wound in opposite directions. We can make the above expression a little more general as follows:

$$\boxed{\frac{V_1}{V_2} = p\frac{N_1}{N_2}} \tag{8.5.4}$$

where p is defined to $+1$ when the coils are wound in the same direction and -1 when coils are wound in opposite directions. (It is an excellent exercise to confirm that this is true by repeating the above analysis with winding direction changed for either the upper or lower coil, for which p will then turn out to be $+1$.) This is the "transformer law" of basic electric circuit theory, from which all other characteristics of transformers as two-port circuit devices can be obtained (See Section 8.6 for follow-up on that). Summarizing:

> The ratio of coil voltages in an ideal transformer is equal to the ratio of turns with sign determined by the relative directions of the windings, per Equation 8.5.4.

A more familiar transformer design is shown in Figure 8.5.3 – coils wound on a toroidal core as opposed to a cylindrical core. Why do this? This arrangement confines the magnetic field linking the two coils to the core, as opposed to allowing field lines to extend beyond the device. This confinement is important in order to prevent fields originating outside the transformer from interfering with the magnetic field linking the coils, which would lead to electromagnetic interference (EMI) and electromagnetic compatibility (EMC) problems. The principle of operation is in all other respects the same.

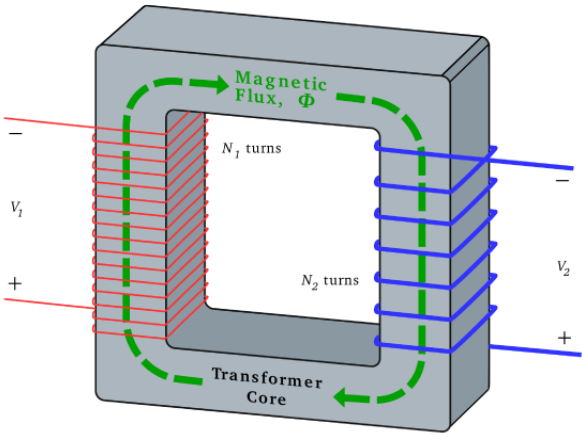


Figure 8.5.3: Transformer implemented as coils sharing a toroidal core. Here $p = +1$. (CC BY SA 3.0; BillC)

1. See “Additional Reading” at the end of this section for more on these applications.↩

This page titled 8.5: Transformers - Principle of Operation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **8.5: Transformers - Principle of Operation** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

8.6: Transformers as Two-Port Devices

Section 8.5 explains the principle of operation of the ideal transformer. The relationship governing the terminal voltages V_1 and V_2 was found to be

$$\frac{V_1}{V_2} = p\frac{N_1}{N_2}$$

where N_1 and N_2 are the number of turns in the associated coils and p is either $+1$ or -1 depending on the relative orientation of the windings; i.e., whether the reference direction of the associated fluxes is the same or opposite, respectively.

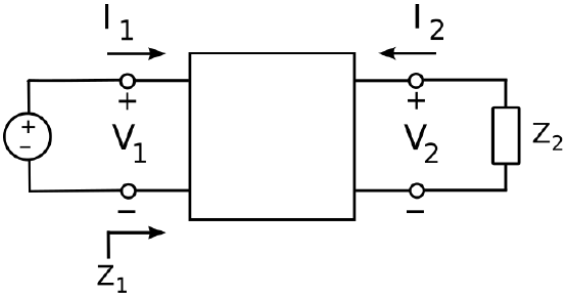


Figure 8.6.1: The transformer as a two-port circuit device.

We shall now consider ratios of current and impedance in ideal transformers, using the two-port model shown in Figure 8.6.1. By virtue of the reference current directions and polarities chosen in this figure, the power delivered by the source V_1 is $V_1 I_1$, and the power absorbed by the load Z_2 is $-V_2 I_2$. Assuming the transformer windings have no resistance, and assuming the magnetic flux is perfectly contained within the core, the power absorbed by the load must equal the power provided by the source; i.e., $V_1 I_1 = -V_2 I_2$. Thus, we have[1]

$$\boxed{\frac{I_1}{I_2} = -\frac{V_2}{V_1} = -p\frac{N_2}{N_1}}$$

We can develop an impedance relationship for ideal transformers as follows. Let Z_1 be the input impedance of the transformer; that is, the impedance looking into port 1 from the source. Thus, we have

$$\begin{aligned} Z_1 &\triangleq \frac{V_1}{I_1} \\ &= \frac{+p\,(N_1/N_2)\,V_2}{-p\,(N_2/N_1)\,I_2} \\ &= -\left(\frac{N_1}{N_2}\right)^2\left(\frac{V_2}{I_2}\right) \end{aligned} \tag{8.6.1}$$

In Figure 8.6.1, Z_2 is the the output impedance of port 2; that is, the impedance looking out port 2 into the load. Therefore, $Z_2 = -V_2/I_2$ (once again the minus sign is a result of our choice of sign conventions in Figure 8.6.1). Substitution of this result into the above equation yields

$$Z_1 = \left(\frac{N_1}{N_2}\right)^2 Z_2$$

and therefore

$$\boxed{\frac{Z_1}{Z_2} = \left(\frac{N_1}{N_2}\right)^2}$$

Thus, we have demonstrated that a transformer scales impedance in proportion to the square of the turns ratio N_1/N_2. Remarkably, the impedance transformation depends *only* on the turns ratio, and is independent of the relative direction of the windings (p).

The relationships developed above should be viewed as AC expressions, and are not normally valid at DC. This is because transformers exhibit a fundamental limitation in their low-frequency performance. To see this, first recall Faraday's Law:

$$V = -N\frac{\partial}{\partial t}\Phi$$

If the magnetic flux Φ is not time-varying, then there is no induced electric potential, and subsequently no linking of the signals associated with the coils. At very low but non-zero frequencies, we encounter another problem that gets in the way – *magnetic saturation*. To see this, note we can obtain an expression for Φ from Faraday's Law by integrating with respect to time, yielding

$$\Phi(t) = -\frac{1}{N}\int_{t_0}^{t} V(\tau)d\tau + \Phi(t_0)$$

where t_0 is some earlier time at which we know the value of $\Phi(t_0)$. Let us assume that $V(t)$ is sinusoidally-varying. Then the peak value of Φ after $t = t_0$ depends on the frequency of $V(t)$. If the frequency of $V(t)$ is very low, then Φ can become very large. Since the associated cross-sectional areas of the coils are presumably constant, this means that the magnetic field becomes very large. The problem with that is that most practical high-permeability materials suitable for use as transformer cores exhibit magnetic saturation; that is, the rate at which the magnetic field is able to increase decreases with increasing magnetic field magnitude (see Section 7.16). The result of all this is that a transformer may work fine at (say) 1 MHz, but at (say) 1 Hz the transformer may exhibit an apparent loss associated with this saturation. Thus, practical transformers exhibit highpass frequency response.

It should be noted that the highpass behavior of practical transistors can be useful. For example, a transformer can be used to isolate an undesired DC offset and/or low-frequency noise in the circuit attached to one coil from the circuit attached to the other coil.

The DC-isolating behavior of a transformer also allows the transformer to be used as a *balun*, as illustrated in Figure 8.6.2. A **balun** is a two-port device that transforms a single-ended ("unbalanced") signal – that is, one having an explicit connection to a datum (e.g., ground) – into a differential ("balanced") signal, for which there is no explicit connection to a datum. Differential signals have many benefits in circuit design, whereas inputs and outputs to devices must often be in single-ended form. Thus, a common use of transformers is to effect the conversion between single-ended and differential circuits. Although a transformer is certainly not the only device that can be used as a balun, it has one very big advantage, namely bandwidth.

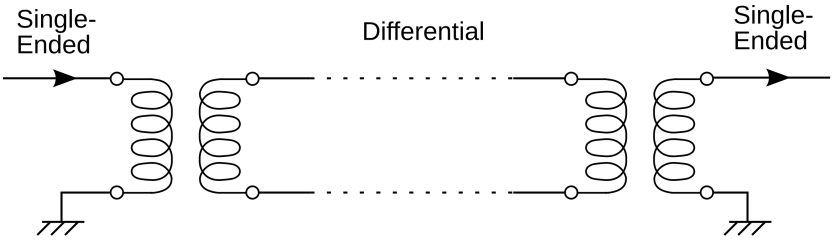


Figure 8.6.2: Transformers used to convert a singleended ("unbalanced") signal to a differential (balanced) signal, and back. (© CC BY SA 3.0 (modified); SpinningSpark)

Additional Reading:

- S.W. Ellingson, "Differential Circuits" (Sec. 8.7) in *Radio Systems Engineering*, Cambridge Univ. Press, 2016.

---

1. The minus signs in this equation are a result of the reference polarity and directions indicated in Figure 8.6.1. These are more-or-less standard in electrical two-port theory (see "Additional Reading" at the end of this section), but are certainly not the only reasonable choice. If you see these expressions without the minus signs, it probably means that a different combination of reference directions/polarities is in effect.↩

This page titled 8.6: Transformers as Two-Port Devices is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **8.6: Transformers as Two-Port Devices** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

8.7: The Electric Generator

A generator is a device that transforms mechanical energy into electrical energy, typically by electromagnetic induction via Faraday's Law. For example, a generator might consist of a gasoline engine that turns a crankshaft to which is attached a system of coils and/or magnets. This rotation changes the relative orientations of the coils with respect to the magnetic field in a time-varying manner, resulting in a time-varying magnetic flux and subsequently induced electric potential. In this case, the induced potential is said to be a form of *motional emf*, as it is due entirely to changes in geometry as opposed to changes in the magnitude of the magnetic field. Coal- and steam-fired generators, hydroelectric generators, wind turbines, and other energy generation devices operate using essentially this principle.

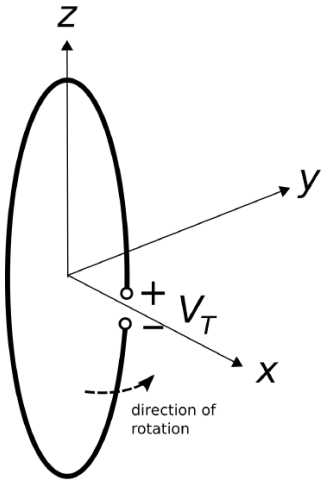


Figure 8.7.1: A rudimentary single-loop generator, shown at time $t = 0$.

Figure 8.7.1 shows a rudimentary generator, which serves as to illustrate the relevant points. This generator consists of a planar loop that rotates around the z axis; therefore, the rotation can be parameterized in ϕ. In this case, the direction of rotation is specified to be in the $+\phi$ direction. The frequency of rotation is f_0; that is, the time required for the loop to make one complete revolution is $1/f_0$. We assume a time-invariant and spatially-uniform magnetic flux density $\mathbf{B} = \hat{\mathbf{b}}B_0$ where $\hat{\mathbf{b}}$ and B_0 are both constants. For illustration purposes, the loop is indicated to be circular. However, because the magnetic field is time-invariant and spatially-uniform, the specific shape of the loop is not important, as we shall see in a moment. Only the area of the loop will matter.

The induced potential is indicated as V_T with reference polarity as indicated in the figure. This potential is given by Faraday's Law:

$$V_T = -\frac{\partial}{\partial t}\Phi$$

Here Φ is the magnetic flux associated with an open surface $\mathcal{S}$ bounded by the loop:

$$\Phi = \int_{\mathcal{S}} \mathbf{B} \cdot d\mathbf{s}$$

As usual, $\mathcal{S}$ can be any surface that intersects all magnetic field lines passing through the loop, and also as usual, the simplest choice is simply the planar area bounded by the loop. The differential surface element $d\mathbf{s}$ is $\hat{\mathbf{n}}ds$, where $\hat{\mathbf{n}}$ is determined by the reference polarity of V_T according to the "right hand rule" convention from Stokes' Theorem. Making substitutions, we have

$$\begin{aligned}\Phi &= \int_{\mathcal{S}} \left(\hat{\mathbf{b}}B_0\right) \cdot (\hat{\mathbf{n}}ds) \\ &= \left[\hat{\mathbf{b}} \cdot \hat{\mathbf{n}}\right] B_0 \int_{\mathcal{S}} ds \\ &= \left[\hat{\mathbf{b}} \cdot \hat{\mathbf{n}}A\right] B_0 \end{aligned} \tag{8.7.1}$$

where A is the area of the loop.

To make headway, an expression for $\hat{\mathbf{n}}$ is needed. The principal difficulty here is that $\hat{\mathbf{n}}$ is rotating with the loop, so it is time-varying. To sort this out, first consider the situation at time $t = 0$, which is the moment illustrated in Figure 8.7.1. Beginning at the "$-$" terminal, we point the thumb of our right hand in the direction that leads to the "$+$" terminal by traversing the loop; $\hat{\mathbf{n}}$ is then the direction perpendicular to the plane of the loop in which the fingers of our right hand pass through the loop. We see that at $t = 0$, $\hat{\mathbf{n}}$

$= +\hat{\mathbf{y}}$. A bit later, the loop will have rotated by one-fourth of a complete rotation, and at that time $\hat{\mathbf{n}} = -\hat{\mathbf{x}}$. This occurs at $t = 1/(4f_0)$. Later still, the loop will have rotated by one-half of a compete rotation, and then $\hat{\mathbf{n}} = -\hat{\mathbf{y}}$. This occurs at $t = 1/(2f_0)$. It is apparent that

$$\hat{\mathbf{n}}(t) = -\hat{\mathbf{x}} \sin 2\pi f_0 t + \hat{\mathbf{y}} \cos 2\pi f_0 t$$

as can be confirmed by checking the three special cases identified above. Now applying Faraday's Law, we find

$$\begin{aligned} V_T &= -\frac{\partial}{\partial t}\Phi \\ &= -\frac{\partial}{\partial t}\left[\hat{\mathbf{b}} \cdot \hat{\mathbf{n}}(t)\right] AB_0 \\ &= -\left[\hat{\mathbf{b}} \cdot \frac{\partial}{\partial t}\hat{\mathbf{n}}(t)\right] AB_0 \end{aligned} \tag{8.7.2}$$

For notational convenience we make the following definition:

$$\hat{\mathbf{n}}'(t) \triangleq -\frac{1}{2\pi f_0}\frac{\partial}{\partial t}\hat{\mathbf{n}}(t) \tag{8.7.3}$$

which is simply the time derivative of $\hat{\mathbf{n}}$ divided by $2\pi f_0$ so as to retain a unit vector. The reason for including a change of sign will become apparent in a moment. Applying this definition, we find

$$\hat{\mathbf{n}}'(t) = +\hat{\mathbf{x}} \cos 2\pi f_0 t + \hat{\mathbf{y}} \sin 2\pi f_0 t$$

Note that this is essentially the definition of the radial basis vector $\hat{\rho}$ from the cylindrical coordinate system (which is why we applied the minus sign in Equation 8.7.3). Recognizing this, we write

$$\hat{\rho}(t) = +\hat{\mathbf{x}} \cos 2\pi f_0 t + \hat{\mathbf{y}} \sin 2\pi f_0 t$$

and finally

$$\boxed{V_T = +2\pi f_0 AB_0 \hat{\mathbf{b}} \cdot \hat{\rho}(t)} \tag{8.7.4}$$

If the purpose of this device is to generate power, then presumably we would choose the magnetic field to be in a direction that maximizes the maximum value of $\hat{\mathbf{b}} \cdot \hat{\rho}(t)$. Therefore, power is optimized for $\mathbf{B}$ polarized entirely in some combination of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, and with $\mathbf{B} \cdot \hat{\mathbf{z}} = 0$. Under that constraint, we see that V_T varies sinusoidally with frequency f_0 and exhibits peak magnitude

$$\boxed{\max\ |V_T(t)| = 2\pi f_0 AB_0}$$

It's worth noting that the maximum voltage magnitude is achieved when the plane of the loop is parallel to $\mathbf{B}$; i.e., when $\hat{\mathbf{b}} \cdot \hat{\mathbf{n}}(t) = 0$ so that $\Phi(t) = 0$. Why is that? Because this is when $\Phi(t)$ is most rapidly increasing or decreasing. Conversely, when the plane of the loop is perpendicular to $\mathbf{B}$ (i.e., $\hat{\mathbf{b}} \cdot \hat{\mathbf{n}}(t) = 1$), $|\Phi(t)|$ is maximum but its time-derivative is zero, so $V_T(t) = 0$ at this instant.

Example 8.7.1: Rudimentary electric generator

The generator in Figure 8.7.1 consists of a circular loop of radius $a = 1$ cm rotating at 1000 revolutions per second in a static and spatially-uniform magnetic flux density of 1 mT in the $+\hat{\mathbf{x}}$ direction. What is the induced potential?

Solution

From the problem statement, $f_0 = 1$ kHz, $B_0 = 1$ mT, and $\hat{\mathbf{b}} = +\hat{\mathbf{x}}$. Therefore $\hat{\mathbf{b}} \cdot \hat{\rho}(t) = \hat{\mathbf{x}} \cdot \hat{\rho}(t) = \cos 2\pi f_0 t$. The area of the loop is $A = \pi a^2$. From Equation 8.7.4 we obtain

$$V_T(t) \cong (1.97\ \text{mV}) \cos\left[(6.28\ \text{krad/s})\, t\right]$$

Finally, we note that it is not really necessary for the loop to rotate in the presence of a magnetic field with constant $\hat{\mathbf{b}}$; it works equally well for the loop to be stationary and for $\hat{\mathbf{b}}$ to rotate – in fact, this is essentially the same problem. In some practical generators, both the potential-generating coils and fields (generated by some combination of magnets and coils) rotate.

This page titled 8.7: The Electric Generator is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **8.7: The Electric Generator** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

8.8: The Maxwell-Faraday Equation

In this section, we generalize Kirchoff's Voltage Law (KVL), previously encountered as a principle of electrostatics in Sections 5.10 and 5.11. KVL states that in the absence of a time-varying magnetic flux, the electric potential accumulated by traversing a closed path $\mathcal{C}$ is zero. Here is that idea in mathematical form:

$$V = \oint_{\mathcal{C}} \mathbf{E} \cdot d\mathbf{l} = 0$$

Now recall Faraday's Law (Section [m0055_Faradays_Law]):

$$V = -\frac{\partial}{\partial t}\Phi = -\frac{\partial}{\partial t}\int_{\mathcal{S}} \mathbf{B} \cdot d\mathbf{s}$$

Here, $\mathcal{S}$ is any open surface that intersects all magnetic field lines passing through $\mathcal{C}$, with the relative orientations of $\mathcal{C}$ and $d\mathbf{s}$ determined in the usual way by the Stokes' Theorem convention. Note that Faraday's Law agrees with KVL in the magnetostatic case. If magnetic flux is constant, then Faraday's Law says $V = 0$. However, Faraday's Law is very clearly *not* consistent with KVL if magnetic flux is time-varying. The correction is simple enough; we can simply set these expressions to be equal. Here we go:

$$\boxed{\oint_{\mathcal{C}} \mathbf{E} \cdot d\mathbf{l} = -\frac{\partial}{\partial t}\int_{\mathcal{S}} \mathbf{B} \cdot d\mathbf{s}} \tag{8.8.1}$$

This general form is known by a variety of names; here we refer to it as the *Maxwell-Faraday Equation* (MFE).

> The integral form of the Maxwell-Faraday Equation (Equation 8.8.1) states that the electric potential associated with a closed path $\mathcal{C}$ is due entirely to electromagnetic induction, via Faraday's Law.

Despite the great significance of this expression as one of Maxwell's Equations, one might argue that all we have done is simply to write Faraday's Law in a slightly more verbose way. This is true. The *real* power of the MFE is unleashed when it is expressed in differential, as opposed to integral form. Let us now do this.

We can transform the left-hand side of Equation 8.8.1 into a integral over $\mathcal{S}$ using Stokes' Theorem. Applying Stokes' theorem on the left, we obtain

$$\int_{\mathcal{S}} (\nabla \times \mathbf{E}) \cdot d\mathbf{s} = -\frac{\partial}{\partial t}\int_{\mathcal{S}} \mathbf{B} \cdot d\mathbf{s}$$

Now exchanging the order of integration and differentiation on the right hand side:

$$\int_{\mathcal{S}} (\nabla \times \mathbf{E}) \cdot d\mathbf{s} = \int_{\mathcal{S}} \left(-\frac{\partial}{\partial t}\mathbf{B}\right) \cdot d\mathbf{s}$$

The surface $\mathcal{S}$ on both sides is the same, and we have not constrained $\mathcal{S}$ in any way. $\mathcal{S}$ can be any mathematically-valid open surface anywhere in space, having any size and any orientation. The only way the above expression can be universally true under these conditions is if the integrands on each side are equal at every point in space. Therefore,

$$\boxed{\nabla \times \mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B}} \tag{8.8.2}$$

which is the MFE in differential form.

What does this mean? Recall that the curl of $\mathbf{E}$ is a way to take a derivative of $\mathbf{E}$ with respect to position (Section 4.8). Therefore the MFE constrains spatial derivatives of $\mathbf{E}$ to be simply related to the rate of change of $\mathbf{B}$. Said plainly:

> The differential form of the Maxwell-Faraday Equation (Equation 8.8.2) relates the change in the electric field with position to the change in the magnetic field with time.

Now that *is* arguably new and useful information. We now see that electric and magnetic fields are coupled not only for line integrals and fluxes, but also at each point in space.

This page titled 8.8: The Maxwell-Faraday Equation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **8.8: The Maxwell-Faraday Equation** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

8.9: Displacement Current and Ampere's Law

In this section, we generalize Ampere's Law, previously encountered as a principle of magnetostatics in Sections 7.4 and 7.9. Ampere's Law states that the current I_{encl} flowing through closed path $\mathcal{C}$ is equal to the line integral of the magnetic field intensity $\mathbf{H}$ along $\mathcal{C}$. That is:

$$\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_{encl} \tag{8.9.1}$$

We shall now demonstrate that this equation is unreliable if the current is not steady; i.e., not DC.

First, consider the situation shown in Figure 8.9.1. Here, a current I flows in the wire, subsequently generating a magnetic field $\mathbf{H}$ that circulates around the wire (Section 7.5). When we perform the integration in Ampere's Law along any path $\mathcal{C}$ enclosing the wire, the result is I, as expected. In this case, Ampere's Law is working even when I is time-varying.

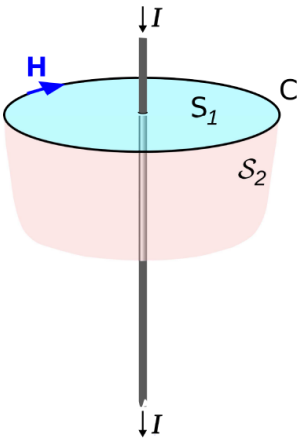


Figure 8.9.1: Ampere's Law applied to a continuous line of current (modified; C. Burks)

Now consider the situation shown in Figure 8.9.2, in which we have introduced a parallel-plate capacitor. In the DC case, this situation is simple. No current flows, so there is no magnetic field and Ampere's Law is trivially true. In the AC case, the current I *can* be non-zero, but we must be clear about the physical origin of this current. What is happening is that for one half of a period, a source elsewhere in the circuit is moving positive charge to one side of the capacitor and negative charge to the other side. For the other half-period, the source is exchanging the charge, so that negative charge appears on the previously positively-charged side and vice-versa. Note that at no point is current flowing directly from one side of the capacitor to the other; instead, all current must flow through the circuit in order to arrive at the other plate. Even though there is no current between the plates, there is current in the wire, and therefore there is also a magnetic field associated with that current.

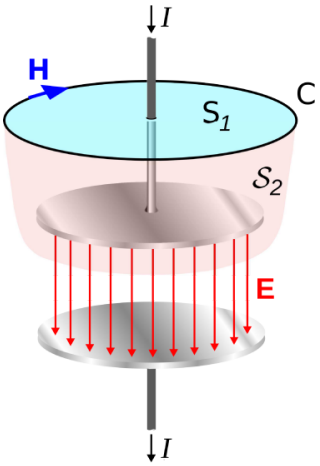


Figure 8.9.1: Ampere's Law applied to a parallel plate capacitor (modified; C. Burks)

Now we are ready to shine a light on the problem. Recall that from Stokes' Theorem, the line integral over $\mathcal{C}$ is mathematically equivalent to an integral over any open surface $\mathcal{S}$ that is bounded by $\mathcal{C}$. Two such surfaces are shown in Figure 8.9.1 and Figure 8.9.2, indicated as $\mathcal{S}_1$ and $\mathcal{S}_2$. In the wire-only scenario of Figure 8.9.1, the choice of $\mathcal{S}$ clearly doesn't matter; any valid surface intersects current equal to I. Similarly in the scenario of Figure 8.9.2, everything seems fine if we choose $\mathcal{S} = \mathcal{S}_1$. If, on the other

hand, we select $\mathcal{S}_2$ in the parallel-plate capacitor case, then we have a problem. There is no current flowing through $\mathcal{S}_2$, so the right side of Equation 8.9.1 is zero even though the left side is potentially non-zero. So, it appears that something necessary for the time-varying case is missing from Equation 8.9.1.

To resolve the problem, we postulate an additional term in Ampere's Law that is non-zero in the above scenario. Specifically, we propose:

$$\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_c + I_d$$

where I_c is the enclosed current (formerly identified as I_{encl}) and I_d is the proposed new term. If we are to accept this postulate, then here is a list of things we know about I_d:

- I_d has units of current (A).
- $I_d = 0$ in the DC case and is potentially non-zero in the AC case. This implies that I_d is the time derivative of some other quantity.
- I_d must be somehow related to the electric field.

How do we know I_d must be related to the electric field? This is because the Maxwell-Faraday Equation tells us that spatial derivatives of $\mathbf{E}$ are related to time derivatives of $\mathbf{H}$; i.e., $\mathbf{E}$ and $\mathbf{H}$ are coupled in the time-varying (here, AC) case. This coupling between $\mathbf{E}$ and $\mathbf{H}$ must also be at work here, but we have not yet seen $\mathbf{E}$ play a role. This is pretty strong evidence that I_d depends on the electric field.

Without further ado, here's I_d:

$$I_d = \int_{\mathcal{S}} \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{s}$$

where $\mathbf{D}$ is the electric flux density (units of C/m^2) and is equal to $\epsilon\mathbf{E}$ as usual, and $\mathcal{S}$ is the same open surface associated with $\mathcal{C}$ in Ampere's Law. Note that this expression meets our expectations: It is determined by the electric field, it is zero when the electric field is constant (i.e., not time varying), and has units of current.

The quantity I_d is commonly known as *displacement current*. It should be noted that this name is a bit misleading, since I_d is not a current in the conventional sense. Certainly, it is not a *conduction* current – conduction current is represented by I_c, and there is no current conducted through an ideal capacitor. It is not unreasonable to think of I_d as current in a more general sense, for the following reason. At one instant, charge is distributed one way and at another, it is distributed in another way. If you define current as a time variation in the charge distribution relative to $\mathcal{S}$ – regardless of the path taken by the charge – then I_d is a current. However, this distinction is a bit philosophical, so it may be less confusing to interpret "displacement current" instead as a separate electromagnetic quantity that just happens to have units of current.

Now we are able to write the general form of Ampere's Law that applies even when sources are time-varying. Here it is:

$$\boxed{\oint_{\mathcal{C}} \mathbf{H} \cdot d\mathbf{l} = I_c + \int_{\mathcal{S}} \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{s}} \tag{8.9.2}$$

As is the case in the Maxwell-Faraday Equation, most of the utility of Ampere's Law is unleashed when expressed in differential form. To obtain this form the first step is to write I_c as an integral of over $\mathcal{S}$; this is simply (see Section 6.2):

$$I_c = \int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{s}$$

where $\mathbf{J}$ is the volume current density (units of A/m^2). So now we have

$$\begin{aligned} \oint_C \mathbf{H} \cdot d\mathbf{l} &= \int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{s} + \int_S \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{s} \\ &= \int_{\mathcal{S}} \left(\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) \cdot d\mathbf{s} \end{aligned} \tag{8.9.3}$$

We can transform the left side of the above equation into a integral over $\mathcal{S}$ using Stokes' Theorem. We obtain

$$\int_{\mathcal{S}} (\nabla \times \mathbf{H}) \cdot d\mathbf{s} = \int_{\mathcal{S}} \left(\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) \cdot d\mathbf{s}$$

The surface $\mathcal{S}$ on both sides is the same, and we have not constrained $\mathcal{S}$ in any way. $\mathcal{S}$ can be any mathematically-valid open surface anywhere in space, having any size and any orientation. The only way the above expression can be universally true under these conditions is if the integrands on each side are equal at every point in space. Therefore:

$$\boxed{\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial}{\partial t}\mathbf{D}} \tag{8.9.4}$$

which is Ampere's Law in differential form.

What does Equation 8.9.4 mean? Recall that the curl of $\mathbf{H}$ is a way to describe the direction and rate of change of $\mathbf{H}$ with position. Therefore, this equation constrains spatial derivatives of $\mathbf{H}$ to be simply related to $\mathbf{J}$ and the time derivative of $\mathbf{D}$ (displacement current). Said plainly:

> The differential form of the general (time-varying) form of Ampere's Law (Equation 8.9.4) relates the change in the magnetic field with position to the change in the electric field with time, plus current.

As is the case in the Maxwell-Faraday Equation, we see that electric and magnetic fields become coupled at each point in space when sources are time-varying.

This page titled 8.9: Displacement Current and Ampere's Law is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **8.9: Displacement Current and Ampere's Law** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

CHAPTER OVERVIEW

9: Plane Waves in Loseless Media

9.1: Maxwell’s Equations in Differential Phasor Form
9.2: Wave Equations for Source-Free and Lossless Regions
9.3: Types of Waves
9.4: Uniform Plane Waves - Derivation
9.5: Uniform Plane Waves - Characteristics
9.6: Wave Polarization
9.7: Wave Power in a Lossless Medium

Contributors and Attributions

- Ellingson, Steven W. (2018) Electromagnetics, Vol. 1. Blacksburg, VA: VT Publishing. https://doi.org/10.21061/electromagnetics-vol-1 Licensed with CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0. Report adoption of this book here. If you are a professor reviewing, adopting, or adapting this textbook please help us understand a little more about your use by filling out this form.

This page titled 9: Plane Waves in Loseless Media is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

9.1: Maxwell's Equations in Differential Phasor Form

In this section, we derive the phasor form of Maxwell's Equations from the general time-varying form of these equations. Here we are interested exclusively in the differential ("point") form of these equations. It is assumed that the reader is comfortable with phasor representation and its benefits; if not, a review of Section 1.5 is recommended before attempting this section.

Maxwell's Equations in differential time-domain form are Gauss' Law:

$$\nabla \cdot \mathbf{D} = \rho_v \tag{9.1.1}$$

the Maxwell-Faraday Equation (MFE):

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B} \tag{9.1.2}$$

Gauss' Law for Magnetism (GSM):

$$\nabla \cdot \mathbf{B} = 0$$

and Ampere's Law:

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial}{\partial t}\mathbf{D}$$

We begin with Gauss's Law (Equation 9.1.1). We define $\widetilde{\mathbf{D}}$ and $\tilde{\rho}_v$ as phasor quantities through the usual relationship:

$$\mathbf{D} = \mathrm{Re}\left\{\widetilde{\mathbf{D}}e^{j\omega t}\right\}$$

and

$$\rho_v = \mathrm{Re}\left\{\tilde{\rho}_v e^{j\omega t}\right\}$$

Substituting these expressions into Equation 9.1.1:

$$\nabla \cdot \left[\mathrm{Re}\left\{\widetilde{\mathbf{D}}e^{j\omega t}\right\}\right] = \mathrm{Re}\left\{\tilde{\rho}_v e^{j\omega t}\right\}$$

Divergence is a real-valued linear operator. Therefore, we may exchange the order of the "Re" and "$\nabla\cdot$" operations (this is "Claim 2" from Section 1.5):

$$\mathrm{Re}\left\{\nabla \cdot \left[\widetilde{\mathbf{D}}e^{j\omega t}\right]\right\} = \mathrm{Re}\left\{\tilde{\rho}_v e^{j\omega t}\right\}$$

Next, we note that the differentiation associated with the divergence operator is with respect to position and not with respect to time, so the order of operations may be further rearranged as follows:

$$\mathrm{Re}\left\{\left[\nabla \cdot \widetilde{\mathbf{D}}\right] e^{j\omega t}\right\} = \mathrm{Re}\left\{\tilde{\rho}_v e^{j\omega t}\right\}$$

Finally, we note that the equality of the left and right sides of the above equation implies the equality of the associated phasors (this is "Claim 1" from Section 1.5); thus,

$$\boxed{\nabla \cdot \widetilde{\mathbf{D}} = \tilde{\rho}_v} \tag{9.1.3}$$

In other words, the differential form of Gauss' Law for phasors is identical to the differential form of Gauss' Law for physical time-domain quantities.

The same procedure applied to the MFE is only a little more complicated. First, we establish the phasor representations of the electric and magnetic fields:

$$\mathbf{E} = \mathrm{Re}\left\{\widetilde{\mathbf{E}}e^{j\omega t}\right\}$$

$$\mathbf{B} = \mathrm{Re}\left\{\widetilde{\mathbf{B}}e^{j\omega t}\right\}$$

After substitution into Equation 9.1.2:

$$\nabla \times \left[\mathrm{Re}\left\{\widetilde{\mathbf{E}}e^{j\omega t}\right\}\right] = -\frac{\partial}{\partial t}\left[\mathrm{Re}\left\{\widetilde{\mathbf{B}}e^{j\omega t}\right\}\right]$$

Both curl and time-differentiation are real-valued linear operations, so we are entitled to change the order of operations as follows:

$$\mathrm{Re}\left\{\nabla \times \left[\widetilde{\mathbf{E}}e^{j\omega t}\right]\right\} = -\mathrm{Re}\left\{\frac{\partial}{\partial t}\left[\widetilde{\mathbf{B}}e^{j\omega t}\right]\right\}$$

On the left, we note that the time dependence $e^{j\omega t}$ can be pulled out of the argument of the curl operator, since it does not depend on position:

$$\mathrm{Re}\left\{\left[\nabla \times \widetilde{\mathbf{E}}\right]e^{j\omega t}\right\} = -\mathrm{Re}\left\{\frac{\partial}{\partial t}\left[\widetilde{\mathbf{B}}e^{j\omega t}\right]\right\}$$

On the right, we note that $\widetilde{\mathbf{B}}$ is constant with respect to time (because it is a phasor), so:

$$\begin{aligned}\mathrm{Re}\left\{[\nabla \times \tilde{\mathbf{E}}]e^{j\omega t}\right\} &= -\mathrm{Re}\left\{\widetilde{\mathbf{B}}\frac{\partial}{\partial t}e^{j\omega t}\right\} \\ &= -\mathrm{Re}\left\{\widetilde{\mathbf{B}}j\omega e^{j\omega t}\right\} \\ &= \mathrm{Re}\left\{[-j\omega\widetilde{\mathbf{B}}]e^{j\omega t}\right\}\end{aligned} \tag{9.1.4}$$

And so we have found:

$$\boxed{\nabla \times \widetilde{\mathbf{E}} = -j\omega\widetilde{\mathbf{B}}} \tag{9.1.5}$$

Let's pause for a moment to consider the above result. In the general time-domain version of the MFE, we must take spatial derivatives of the electric field and time derivatives of the magnetic field. In the phasor version of the MFE, the time derivative operator has been replaced with multiplication by $j\omega$. This is a tremendous simplification since the equations now involve differentiation over position only. Furthermore, no information is lost in this simplification – for a reminder of why that is, see the discussion of Fourier Analysis at the end of Section 1.5. Without this kind of simplification, much of what is now considered "basic" engineering electromagnetics would be intractable.

The procedure for conversion of the remaining two equations is very similar, yielding:

$$\boxed{\nabla \cdot \widetilde{\mathbf{B}} = 0} \tag{9.1.6}$$

$$\boxed{\nabla \times \widetilde{\mathbf{H}} = \widetilde{\mathbf{J}} + j\omega\widetilde{\mathbf{D}}} \tag{9.1.7}$$

The details are left as an exercise for the reader.

> The differential form of Maxwell's Equations (Equations 9.1.3, 9.1.5, 9.1.6, and 9.1.7) involve operations on the phasor representations of the physical quantities. These equations have the advantage that differentiation with respect to time is replaced by multiplication by $j\omega$.

This page titled 9.1: Maxwell's Equations in Differential Phasor Form is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **9.1: Maxwell's Equations in Differential Phasor Form** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

9.2: Wave Equations for Source-Free and Lossless Regions

Electromagnetic waves are solutions to a set of coupled differential simultaneous equations – namely, Maxwell's Equations. The general solution to these equations includes constants whose values are determined by the applicable electromagnetic boundary conditions. However, this direct approach can be difficult and is often not necessary. In unbounded homogeneous regions that are "source free" (containing no charges or currents), a simpler approach is possible. In this section, we reduce Maxwell's Equations to *wave equations* that apply to the electric and magnetic fields in this simpler category of scenarios. Before reading further, the reader should consider a review of Section 1.3 (noting in particular Equation 1.3.1) and Section 3.6 (wave equations for voltage and current on a transmission line). This section seeks to develop the analogous equations for electric and magnetic waves.

We can get the job done using the differential "point" phasor form of Maxwell's Equations, developed in Section 9.1. Here they are:

$$\nabla \cdot \widetilde{\mathbf{D}} = \tilde{\rho}_v$$

$$\nabla \times \widetilde{\mathbf{E}} = -j\omega\widetilde{\mathbf{B}}$$

$$\nabla \cdot \widetilde{\mathbf{B}} = 0$$

$$\nabla \times \widetilde{\mathbf{H}} = \widetilde{\mathbf{J}} + j\omega\widetilde{\mathbf{D}}$$

In a source-free region, there is no net charge and no current, hence $\tilde{\rho}_v = 0$ and $\widetilde{\mathbf{J}} = 0$ in the present analysis. The above equations become

$$\nabla \cdot \widetilde{\mathbf{D}} = 0$$

$$\nabla \times \widetilde{\mathbf{E}} = -j\omega\widetilde{\mathbf{B}}$$

$$\nabla \cdot \widetilde{\mathbf{B}} = 0$$

$$\nabla \times \widetilde{\mathbf{H}} = +j\omega\widetilde{\mathbf{D}}$$

Next, we recall that $\widetilde{\mathbf{D}} = \epsilon\widetilde{\mathbf{E}}$ and that ϵ is a real-valued constant for a medium that is homogeneous, isotropic, and linear (Section 2.8). Similarly, $\widetilde{\mathbf{B}} = \mu\tilde{\mathbf{H}}$ and μ is a real-valued constant. Thus, under these conditions, it is sufficient to consider *either* $\widetilde{\mathbf{D}}$ or $\widetilde{\mathbf{E}}$ and *either* $\widetilde{\mathbf{B}}$ or $\widetilde{\mathbf{H}}$. The choice is arbitrary, but in engineering applications it is customary to use $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$. Eliminating the now-redundant quantities $\widetilde{\mathbf{D}}$ and $\widetilde{\mathbf{B}}$, the above equations become

$$\nabla \cdot \widetilde{\mathbf{E}} = 0 \tag{9.2.1}$$

$$\nabla \times \widetilde{\mathbf{E}} = -j\omega\mu\widetilde{\mathbf{H}} \tag{9.2.2}$$

$$\nabla \cdot \widetilde{\mathbf{H}} = 0$$

$$\nabla \times \widetilde{\mathbf{H}} = +j\omega\epsilon\widetilde{\mathbf{E}} \tag{9.2.3}$$

It is important to note that requiring the region of interest to be source-free precludes the possibility of loss in the medium. To see this, let's first be clear about what we mean by "loss." For an electromagnetic wave, loss is observed as a reduction in the magnitude of the electric and magnetic field with increasing distance. This reduction is due to the dissipation of power in the medium. This occurs when the conductivity σ is greater than zero because Ohm's Law for Electromagnetics ($\widetilde{\mathbf{J}} = \sigma\widetilde{\mathbf{E}}$; Section 6.3) requires that power in the electric field be transferred into conduction current, and is thereby lost to the wave (Section 6.6). When we required $\mathbf{J}$ to be zero above, we precluded this possibility; that is, we implicitly specified $\sigma = 0$. The fact that the constitutive parameters μ and ϵ appear in Equations 9.2.1 – 9.2.3, but σ does not, is further evidence of this.

> Equations 9.2.1–9.2.3 are Maxwell's Equations for a region comprised of isotropic, homogeneous, and source-free material. Because there can be no conduction current in a source-free region, these equations apply only to material that is lossless (i.e., having negligible σ).

Before moving on, one additional disclosure is appropriate. It turns out that there actually *is* a way to use Equations 9.2.1–9.2.3 for regions in which loss is significant. This requires a redefinition of ϵ as a complex-valued quantity. We shall not consider this technique in this section. We mention this simply because one should be aware that if permittivity appears as a complex-valued quantity, then the imaginary part represents loss.

To derive the wave equations we begin with the MFE, Equation 9.2.2. Taking the curl of both sides of the equation we obtain

$$\begin{aligned} \nabla \times (\nabla \times \widetilde{\mathbf{E}}) &= \nabla \times (-j\omega\mu\widetilde{\mathbf{H}}) \\ &= -j\omega\mu(\nabla \times \widetilde{\mathbf{H}}) \end{aligned} \tag{9.2.4}$$

On the right we can eliminate $\nabla \times \widetilde{\mathbf{H}}$ using Equation 9.2.3:

$$\begin{aligned} -j\omega\mu(\nabla \times \widetilde{\mathbf{H}}) &= -j\omega\mu(+j\omega e\widetilde{\mathbf{E}}) \\ &= +\omega^2\mu\epsilon\widetilde{\mathbf{E}} \end{aligned} \tag{9.2.5}$$

On the left side of Equation 9.2.4, we apply the vector identity

$$\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A}$$

which in this case is

$$\nabla \times \nabla \times \widetilde{\mathbf{E}} = \nabla\left(\nabla \cdot \widetilde{\mathbf{E}}\right) - \nabla^2\widetilde{\mathbf{E}}$$

We may eliminate the first term on the right using Equation 9.2.1, yielding

$$\nabla \times \nabla \times \widetilde{\mathbf{E}} = -\nabla^2\widetilde{\mathbf{E}}$$

Substituting these results back into Equation 9.2.4 and rearranging terms we have

$$\nabla^2\widetilde{\mathbf{E}} + \omega^2\mu\epsilon\widetilde{\mathbf{E}} = 0 \tag{9.2.6}$$

This is the wave equation for $\widetilde{\mathbf{E}}$. Note that it is a homogeneous (in the mathematical sense of the word) differential equation, which is expected since we have derived it for a source-free region.

It is common to make the following definition

$$\boxed{\beta \triangleq \omega\sqrt{\mu\epsilon}} \tag{9.2.7}$$

so that Equation 9.2.6 may be written

$$\boxed{\nabla^2\widetilde{\mathbf{E}} + \beta^2\widetilde{\mathbf{E}} = 0} \tag{9.2.8}$$

Why go the the trouble of defining β? One reason is that β conveniently captures the contribution of the frequency, permittivity, and permeability all in one constant. Another reason is to emphasize the connection to the parameter β appearing in transmission line theory (see Section 3.8 for a reminder). It should be clear that β is a *phase propagation constant*, having units of 1/m (or rad/m, if you prefer), and indicates the rate at which the phase of the propagating wave progresses with distance.

The wave equation for $\widetilde{\mathbf{H}}$ is obtained using essentially the same procedure, which is left as an exercise for the reader. It should be clear from the duality apparent in Equations 9.2.1-9.2.3 that the result will be very similar. One finds:

$$\boxed{\nabla^2\widetilde{\mathbf{H}} + \beta^2\widetilde{\mathbf{H}} = 0} \tag{9.2.9}$$

> Equations 9.2.8 and 9.2.9 are the wave equations for $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$, respectively, for a region comprised of isotropic, homogeneous, lossless, and source-free material.

Looking ahead, note that $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$ are solutions to the *same* homogeneous differential equation. Consequently, $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$ cannot be different by more than a constant factor and a direction. In fact, we can also determine something about the factor simply by examining the units involved: Since $\widetilde{\mathbf{E}}$ has units of V/m and $\widetilde{\mathbf{H}}$ has units of A/m, this factor will be expressible in units of the ratio

of V/m to A/m, which is Ω. This indicates that the factor will be an impedance. This factor is known as the *wave impedance* and will be addressed in Section 9.5. This impedance is analogous the characteristic impedance of a transmission line (Section 3.7).

Additional Reading:

- "Wave Equation" on Wikipedia.
- "Electromagnetic Wave Equation" on Wikipedia.

This page titled 9.2: Wave Equations for Source-Free and Lossless Regions is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **9.2: Wave Equations for Source-Free and Lossless Regions** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

9.3: Types of Waves

Solutions to the electromagnetic wave equations (Section 9.2) exist in a variety of forms, representing different types of waves. It is useful to identify three particular geometries for unguided waves. Each of these geometries is defined by the shape formed by surfaces of constant phase, which we refer to as *phasefronts*. (Keep in mind the analogy between electromagnetic waves and sound waves (described in Section 1.3), and note that sound waves also exhibit these geometries.)

A *spherical wave* has phasefronts that form concentric spheres, as shown in Figure 9.3.1. Waves are well-modeled as spherical when the dimensions of the source of the wave are small relative to the scale at which the wave is observed. For example, the wave radiated by an antenna having dimensions of 10 cm, when observed in free space over a scale of 10 km, appears to have phasefronts that are very nearly spherical. Note that the magnitude of the field on a phasefront of a spherical wave may vary significantly, but it is the shape of phasefronts that make it a spherical wave.

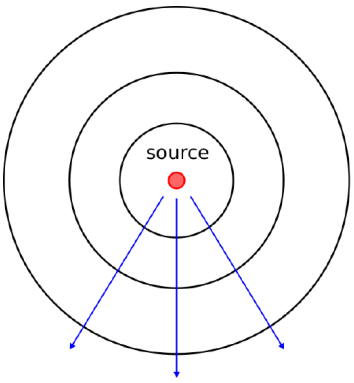


Figure 9.3.1: The phasefronts of a spherical wave form concentric spheres. (© CC BY 4.0; Y. Qin)

A *cylindrical wave* exhibits phasefronts that form concentric cylinders, as shown in Figure 9.3.2. Said differently, the phasefronts of a cylindrical wave are circular in one dimension, and planar in the perpendicular direction. A cylindrical wave is often a good description of the wave that emerges from a line-shaped source.

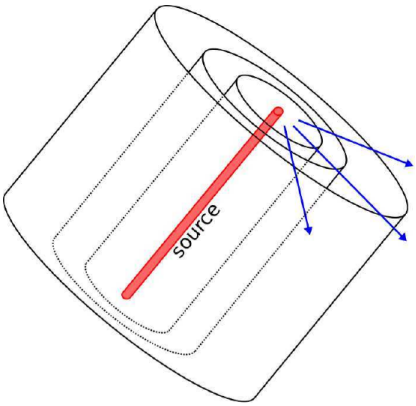


Figure 9.3.2: The phasefronts of a cylindrical wave form concentric cylinders. (© CC BY 4.0; Y. Qin)

A *plane wave* exhibits phasefronts that are planar, with planes that are parallel to each other as shown in Figure 9.3.3. There are two conditions in which waves are well-modeled as plane waves. First, some structures give rise to waves that appear to have planar phasefronts over a limited area; a good example is the wave radiated by a parabolic reflector, as shown in Figure 9.3.4. Second, all waves are well-modeled as plane waves when observed over a small region located sufficiently far from the source. In particular, spherical waves are “locally planar” in the sense that they are well-modeled as planar when observed over a small portion of the spherical phasefront, as shown in Figure 9.3.5. An analogy is that the Earth seems “locally flat” to an observer on the ground, even though it is clearly spherical to an observer in orbit. The “locally planar” approximation is often employed because it is broadly applicable and simplifies analysis.

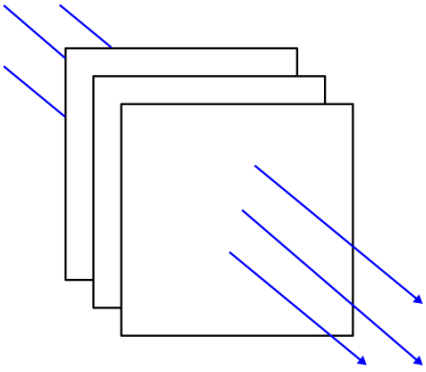


Figure 9.3.3: The phasefronts of a plane wave form parallel planes. (© CC BY 4.0; Y. Qin)

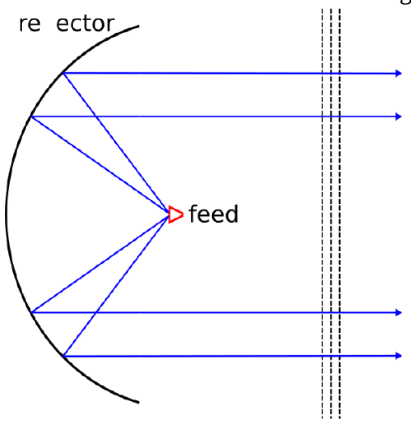


Figure 9.3.4: Plane waves formed in the region in front of a parabolic reflector antenna. (© CC BY 4.0; Y. Qin)

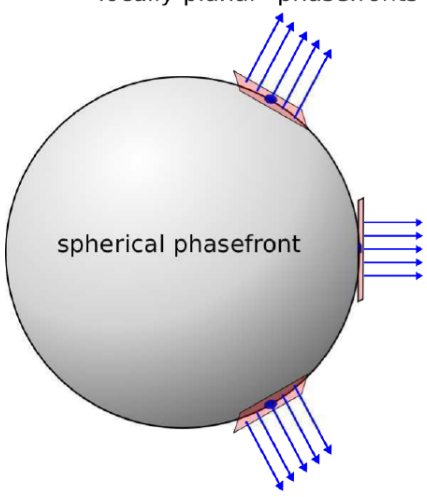


Figure 9.3.5: "Locally planar" approximation of a spherical wave over a limited area. (© CC BY 4.0; Y. Qin)

Most waves are well-modeled as spherical, cylindrical, or plane waves.

Plane waves (having planar phasefronts) are of particular importance due to wide applicability of the “locally planar” approximation.

This page titled 9.3: Types of Waves is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **9.3: Types of Waves** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

9.4: Uniform Plane Waves - Derivation

Section 9.2 showed how Maxwell's Equations could be reduced to a pair of phasor-domain "wave equations," namely:

$$\nabla^2\widetilde{\mathbf{E}} + \beta^2\widetilde{\mathbf{E}} = 0 \tag{9.4.1}$$

$$\nabla^2\widetilde{\mathbf{H}} + \beta^2\widetilde{\mathbf{H}} = 0 \tag{9.4.2}$$

where $\beta = \omega\sqrt{\mu\epsilon}$, assuming unbounded homogeneous, isotropic, lossless, and source-free media. In this section, we solve these equations for the special case of a uniform plane wave. A uniform plane wave is one for which both $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$ have constant magnitude and phase in a specified plane. Despite being a special case, the solution turns out to be broadly applicable, appearing as a common building block in many practical and theoretical problems in unguided propagation (as explained in Section 9.3), as well as in more than a few transmission line and waveguide problems.

To begin, let us assume that the plane over which $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$ have constant magnitude and phase is a plane of constant z. First, note that we may make this assumption with no loss of generality. For example, we could alternatively select a plane of constant y, solve the problem, and then simply exchange variables to get a solution for planes of constant z (or x).[1] Furthermore, the solution for any planar orientation not corresponding to a plane of constant x, y, or z may be similarly obtained by a rotation of coordinates, since the physics of the problem does not depend on the orientation of this plane – if it does, then the medium is not isotropic!

We may express the constraint that the magnitude and phase of $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$ are constant over a plane that is perpendicular to the z axis as follows:

$$\frac{\partial}{\partial x}\widetilde{\mathbf{E}} = \frac{\partial}{\partial y}\widetilde{\mathbf{E}} = \frac{\partial}{\partial x}\widetilde{\mathbf{H}} = \frac{\partial}{\partial y}\widetilde{\mathbf{H}} = 0 \tag{9.4.3}$$

Let us identify the Cartesian components of each of these fields as follows:

$$\begin{aligned}\tilde{\mathbf{E}} &= \hat{\mathbf{x}}\widetilde{E}_x + \hat{\mathbf{y}}\widetilde{E}_y + \hat{\mathbf{z}}\widetilde{E}_z, \text{ and}\\ \widetilde{\mathbf{H}} &= \hat{\mathbf{x}}\widetilde{H}_x + \hat{\mathbf{y}}\widetilde{H}_y + \hat{\mathbf{z}}\widetilde{H}_z\end{aligned} \tag{9.4.4}$$

Now Equation 9.4.3 may be interpreted in detail for $\widetilde{\mathbf{E}}$ as follows:

$$\begin{aligned}\tfrac{\partial}{\partial x}\widetilde{E}_x = \tfrac{\partial}{\partial x}\widetilde{E}_y = \tfrac{\partial}{\partial x}\widetilde{E}_z = 0\\ \tfrac{\partial}{\partial y}\widetilde{E}_x = \tfrac{\partial}{\partial y}\widetilde{E}_y = \tfrac{\partial}{\partial y}\widetilde{E}_z = 0\end{aligned} \tag{9.4.5}$$

and for $\widetilde{\mathbf{H}}$ as follows:

$$\begin{aligned}\frac{\partial}{\partial x}\widetilde{H}_x = \frac{\partial}{\partial x}\widetilde{H}_y = \frac{\partial}{\partial x}\widetilde{H}_z = 0\\ \frac{\partial}{\partial y}\widetilde{H}_x = \frac{\partial}{\partial y}\widetilde{H}_y = \frac{\partial}{\partial y}\widetilde{H}_z = 0\end{aligned} \tag{9.4.6}$$

The wave equation for $\widetilde{\mathbf{E}}$ (Equation 9.4.1 written explicitly in Cartesian coordinates is

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right]\left(\hat{\mathbf{x}}\widetilde{E}_x + \hat{\mathbf{y}}\widetilde{E}_y + \hat{\mathbf{z}}\widetilde{E}_z\right) + \beta^2\left(\hat{\mathbf{x}}\widetilde{E}_x + \hat{\mathbf{y}}\widetilde{E}_y + \hat{\mathbf{z}}\widetilde{E}_z\right) = 0 \tag{9.4.7}$$

Decomposing this equation into separate equations for each of the three coordinates, we obtain the following:

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right]\widetilde{E}_x + \beta^2\widetilde{E}_x = 0 \tag{9.4.8}$$

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right]\widetilde{E}_y + \beta^2\widetilde{E}_y = 0 \tag{9.4.9}$$

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right]\widetilde{E}_z + \beta^2\widetilde{E}_z = 0 \tag{9.4.10}$$

Applying the constraints of Equations 9.4.8 and 9.4.9, we note that many of the terms in Equations 9.4.8–9.4.10 are zero. We are left with:

$$\frac{\partial^2}{\partial z^2}\widetilde{E}_x + \beta^2 \widetilde{E}_x = 0 \tag{9.4.11}$$

$$\frac{\partial^2}{\partial z^2}\widetilde{E}_y + \beta^2 \widetilde{E}_y = 0 \tag{9.4.12}$$

$$\frac{\partial^2}{\partial z^2}\widetilde{E}_z + \beta^2 \widetilde{E}_z = 0 \tag{9.4.13}$$

Now we will show that Equation 9.4.3 also implies that $\widetilde{E}_z$ must be zero. To show this, we use Ampere's Law for a source-free region (Section 9.2):

$$\nabla \times \widetilde{\mathbf{H}} = +j\omega\epsilon\widetilde{\mathbf{E}}$$

and take the dot product with $\hat{\mathbf{z}}$ on both sides:

$$\hat{\mathbf{z}} \cdot (\nabla \times \widetilde{\mathbf{H}}) = +j\omega\epsilon\widetilde{E}_z$$

$$\frac{\partial}{\partial y}\widetilde{H}_x - \frac{\partial}{\partial x}\widetilde{H}_y = +j\omega\epsilon\widetilde{E}_z \tag{9.4.14}$$

Again applying the constraints of Equation 9.4.3, the left side of Equation 9.4.14 must be zero; therefore, $\widetilde{E}_z = 0$. The exact same procedure applied to $\widetilde{\mathbf{H}}$ (using the Maxwell-Faraday Equation; also given in Section 9.2) reveals that $\widetilde{H}_z$ is also zero.[2] Here is what we have found:

> If a wave is uniform over a plane, then the electric and magnetic field vectors must lie in this plane.

This conclusion is a direct consequence of the fact that Maxwell's Equations require the electric field to be proportional to the curl of the magnetic field and vice-versa.

The general solution to Equation 9.4.11 is:

$$\widetilde{E}_x = E_{x0}^{+} e^{-j\beta z} + E_{x0}^{-} e^{+j\beta z} \tag{9.4.15}$$

where E_{x0}^{+} and E_{x0}^{-} are complex-valued constants. The values of these constants are determined by boundary conditions – possibly sources – outside the region of interest. Since in this section we are limiting our scope to source-free and homogeneous regions, we may for the moment consider the values of E_{x0}^{+} and E_{x0}^{-} to be arbitrary, since any values will satisfy the associated wave equation.[3]

Similarly we have for $\widetilde{E}_y$:

$$\widetilde{E}_y = E_{y0}^{+} e^{-j\beta z} + E_{y0}^{-} e^{+j\beta z} \tag{9.4.16}$$

where E_{y0}^{+} and E_{y0}^{-} are again arbitrary constants. Summarizing, we have found

$$\widetilde{\mathbf{E}} = \hat{\mathbf{x}}\widetilde{E}_x + \hat{\mathbf{y}}\widetilde{E}_y$$

where $\widetilde{E}_x$ and $\widetilde{E}_y$ are given by Equations 9.4.15 and 9.4.16, respectively.

Note that Equations 9.4.15 and 9.4.16 are essentially the *same* equations encountered in the study of waves in lossless transmission lines; for a reminder, see Section 3.6. Specifically, factors containing $e^{-j\beta z}$ describe propagation in the $+z$ direction, whereas factors containing $e^{+j\beta z}$ describe propagation in the $-z$ direction. We conclude:

> If a wave is uniform over a plane, then possible directions of propagation are the two directions perpendicular to the plane.

Since we previously established that the electric and magnetic field vectors must lie in the plane, we also conclude:

> The direction of propagation is perpendicular to the electric and magnetic field vectors.

This conclusion turns out to be generally true; i.e., it is not limited to uniform plane waves. Although we will not provide a rigorous proof of this, one way to see that this is true is to imagine that any type of wave can be interpreted as the sum (formally, a *linear*

combination) of uniform plane waves, so perpendicular orientation of the field vectors with respect to the direction of propagation is inescapable.

The same procedure yields the uniform plane wave solution to the wave equation for $\widetilde{\mathbf{H}}$, which is

$$\widetilde{\mathbf{H}} = \hat{\mathbf{x}}\widetilde{H}_x + \hat{\mathbf{y}}\widetilde{H}_y$$

where

$$\widetilde{H}_x = H_{x0}^{+}e^{-j\beta z} + H_{x0}^{-}e^{+j\beta z} \tag{9.4.17}$$

$$\widetilde{H}_y = H_{y0}^{+}e^{-j\beta z} + H_{y0}^{-}e^{+j\beta z} \tag{9.4.18}$$

and where H_{x0}^{+}, H_{x0}^{-}, H_{y0}^{+} and H_{y0}^{-} are arbitrary constants. Note that the solution is essentially the same as that for $\widetilde{\mathbf{E}}$, with the sole difference being that the arbitrary constants may apparently have different values.

To this point, we have seen no particular relationship between the electric and magnetic fields, and it may appear that the electric and magnetic fields are independent of each other. However, Maxwell's Equations – specifically, the two curl equations – make it clear that there must be a relationship between these fields. Subsequently the arbitrary constants in the solutions for $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$ must also be related. In fact, there are two considerations here:

- The *magnitude and phase* of $\widetilde{\mathbf{E}}$ must be related to the magnitude and phase of $\widetilde{\mathbf{H}}$. Since both fields are solutions to the same differential (wave) equation, they may differ by no more than a multiplicative constant. Since the units of $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$ are V/m and A/m respectively, this constant must be expressible in units of V/m divided by A/m; i.e., in units of Ω, an impedance.
- The *direction* of $\widetilde{\mathbf{E}}$ must be related to direction of $\widetilde{\mathbf{H}}$.

Let us now address these considerations. Consider an electric field that points in one of the cardinal directions – let's say $+\hat{\mathbf{x}}$ – and make the definition $E_0 \triangleq E_{x0}^{+}$ for notational convenience. Then the electric field intensity may be written as follows:

$$\widetilde{\mathbf{E}} = \hat{\mathbf{x}}E_0 e^{-j\beta z} \tag{9.4.19}$$

Again, there is no loss of generality here since the coordinate system could be rotated in such a way that *any* uniform plane could be described in this way.

We may now determine $\widetilde{\mathbf{H}}$ from $\widetilde{\mathbf{E}}$ using the Maxwell-Faraday Equation (Section 9.2):

$$\nabla \times \widetilde{\mathbf{E}} = -j\omega\mu\widetilde{\mathbf{H}}$$

Solving this equation for $\widetilde{\mathbf{H}}$, we find:

$$\widetilde{\mathbf{H}} = \frac{\nabla \times \widetilde{\mathbf{E}}}{-j\omega\mu} = \frac{\nabla \times \left[\hat{\mathbf{x}}E_0 e^{-j\beta z}\right]}{-j\omega\mu} \tag{9.4.20}$$

Now let us apply the curl operator. The complete expression for the curl operator in Cartesian coordinates is given in Section B2. Here let us consider one component at a time, starting with the $\hat{\mathbf{x}}$ component:

$$\hat{\mathbf{x}} \cdot \left(\nabla \times \widetilde{\mathbf{E}}\right) = \frac{\partial \widetilde{E}_z}{\partial y} - \frac{\partial \widetilde{E}_y}{\partial z}$$

Since $\widetilde{E}_y = \widetilde{E}_z = 0$, the above expression is zero and subsequently $\widetilde{H}_x = 0$. Next, the $\hat{\mathbf{y}}$ component:

$$\hat{\mathbf{y}} \cdot \left(\nabla \times \widetilde{\mathbf{E}}\right) = \frac{\partial \widetilde{E}_x}{\partial z} - \frac{\partial \widetilde{E}_z}{\partial x}$$

Here $\widetilde{E}_z = 0$, so we have simply

$$\hat{\mathbf{y}} \cdot (\nabla \times \tilde{\mathbf{E}}) = \frac{\partial \widetilde{E}_x}{\partial z}$$

It is not necessary to repeat this procedure for $\widetilde{H}_z$, since we know in advance that $\widetilde{\mathbf{H}}$ must be perpendicular to the direction of propagation and subsequently $\widetilde{H}_z = 0$. Returning to Equation 9.4.20, we obtain:

$$\tilde{\mathbf{H}} = \hat{\mathbf{y}}\frac{1}{-i\omega\mu}\frac{\partial}{\partial z}\tilde{E}_x$$

$$= \hat{\mathbf{y}} \frac{1}{-j\omega\mu} \frac{\partial}{\partial z} \left(E_0 e^{-j\beta z} \right)$$
$$= \hat{\mathbf{y}} \frac{1}{-j\omega\mu} \left(-j\beta E_0 e^{-j\beta z} \right)$$
$$= \hat{\mathbf{y}} \frac{\beta}{\omega\mu} E_0 e^{-j\beta z} \tag{9.4.21}$$

Note that $\widetilde{\mathbf{H}}$ points in the $+\hat{\mathbf{y}}$ direction. So, as expected, both $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$ are perpendicular to the direction of propagation. However, we have now found a more specific relationship: $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$ are perpendicular *to each other*. Just as $\hat{\mathbf{x}} \times \hat{\mathbf{y}} = \hat{\mathbf{z}}$, we see that $\widetilde{\mathbf{E}} \times \widetilde{\mathbf{H}}$ points in the direction of propagation. This is illustrated in Figure 9.4.1. Summarizing:

> $\widetilde{\mathbf{E}}$, $\widetilde{\mathbf{H}}$, and the direction of propagation $\widetilde{\mathbf{E}} \times \widetilde{\mathbf{H}}$ are mutually perpendicular.

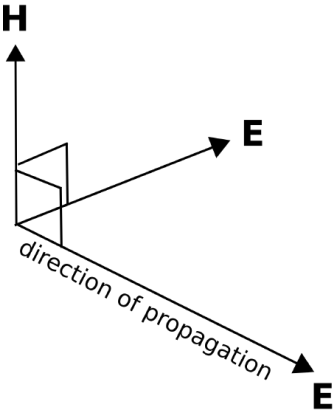


Figure 9.4.1: Relationship between the electric field direction, magnetic field direction, and direction of propagation

Now let us resolve the question of the factor relating $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$. The factor is now seen to be $\beta/\omega\mu$ in Equation 9.4.21, which can be simplified as follows:

$$\frac{\beta}{\omega\mu} = \frac{\omega\sqrt{\mu\epsilon}}{\omega\mu} = \frac{1}{\sqrt{\mu/\epsilon}}$$

The factor $\sqrt{\mu/\epsilon}$ appearing above has units of Ω is known variously as the *wave impedance* or the *intrinsic impedance of the medium*. Assigning this quantity the symbol "η," we have:

$$\boxed{\eta \triangleq \frac{\widetilde{E}_x}{\widetilde{H}_y} = \sqrt{\frac{\mu}{\epsilon}}} \tag{9.4.22}$$

> The ratio of the electric field intensity to the magnetic field intensity is the wave impedance η (Equation 9.4.22; units of Ω). In lossless media, η is determined by the ratio of permeability of the medium to the permittivity of the medium.

The wave impedance in free space, assigned the symbol η_0, is

$$\eta_0 \triangleq \sqrt{\frac{\mu_0}{\epsilon_0}} \cong 377 \ \Omega.$$

Wrapping up our solution, we find that if $\widetilde{\mathbf{E}}$ is as given by Equation 9.4.19, then

$$\widetilde{\mathbf{H}} = \hat{\mathbf{y}} \frac{E_0}{\eta} e^{-j\beta z}$$

---

1. By the way, this is a highly-recommended exercise for the student.↩
2. Showing this is a highly-recommended exercise for the reader.↩
3. The reader is encouraged to confirm that these are solutions by substitution into the associated wave equation.↩

This page titled 9.4: Uniform Plane Waves - Derivation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **9.4: Uniform Plane Waves - Derivation** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

9.5: Uniform Plane Waves - Characteristics

In Section 9.4, expressions for the electric and magnetic fields are determined for a uniform plane wave in lossless media. If the planar phasefront is perpendicular to the z axis, then waves may propagate in either the $+\hat{\mathbf{z}}$ direction of the $-\hat{\mathbf{z}}$ direction. If we consider only the former, and select $\widetilde{\mathbf{E}}$ to point in the $+\hat{\mathbf{x}}$ direction, then we find

$$\widetilde{\mathbf{E}} = +\hat{\mathbf{x}} E_0 e^{-j\beta z} \tag{9.5.1}$$

$$\widetilde{\mathbf{H}} = +\hat{\mathbf{y}} \frac{E_0}{\eta} e^{-j\beta z} \tag{9.5.2}$$

where $\beta = \omega\sqrt{\mu\epsilon}$ is the phase propagation constant, $\eta = \sqrt{\mu/\epsilon}$ is the wave impedance, and E_0 is a complex-valued constant associated with the magnitude and phase of the source. This result is in fact completely general for uniform plane waves, since any other possibility may be obtained by simply rotating coordinates. In fact, this is pretty easy because (as determined in Section 9.4) $\widetilde{\mathbf{E}}$, $\widetilde{\mathbf{H}}$, and the direction of propagation are mutually perpendicular, with the direction of propagation pointing in the same direction as $\widetilde{\mathbf{E}} \times \widetilde{\mathbf{H}}$.

In this section, we identify some important characteristics of uniform plane waves, including wavelength and phase velocity. Chances are that much of what appears here will be familiar to the reader; if not, a quick review of Sections 1.3 ("Fundamentals of Waves") and 3.8 ("Wave Propagation on a Transmission Line") are recommended.

First, recall that $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{H}}$ are phasors representing physical (real-valued) fields and are *not* the field values themselves. The actual, physical electric field intensity is

$$\begin{aligned} \mathbf{E} &= \mathrm{Re}\left\{\widetilde{\mathbf{E}} e^{j\omega t}\right\} \\ &= \mathrm{Re}\left\{\hat{\mathbf{x}} E_0 e^{-j\beta z} e^{j\omega t}\right\} \\ &= \hat{\mathbf{x}} \left|E_0\right| \cos(\omega t - \beta z + \psi) \end{aligned} \tag{9.5.3}$$

where ψ is the phase of E_0. Similarly:

$$\mathbf{H} = \hat{\mathbf{y}} \frac{|E_0|}{\eta} \cos\left(\omega t - \beta z + \psi\right) \tag{9.5.4}$$

This result is illustrated in Figure 9.5.1. Note that both $\mathbf{E}$ and $\mathbf{H}$ (as well as their phasor representations) have the same phase and have the same frequency and position dependence $\cos\left(\omega t - \beta z + \psi\right)$. Since β is a real-valued constant for lossless media, *only* the phase, and not the magnitude, varies with z. If $\omega \to 0$, then $\beta \to 0$ and the fields no longer depend on z; in this case, the field is not propagating. For $\omega > 0$, $\cos\left(\omega t - \beta z + \psi\right)$ is periodic in z; specifically, it has the same value each time z increases or decreases by $2\pi/\beta$. This is, by definition, the wavelength λ:

$$\lambda \triangleq \frac{2\pi}{\beta}$$

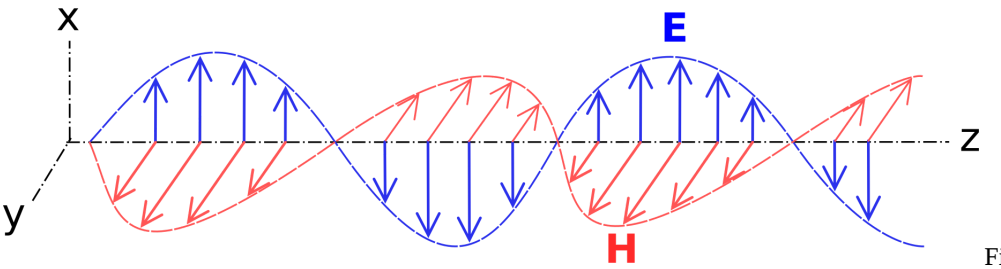


Figure 9.5.1: Relationship between the electric field direction, magnetic field direction, and direction of propogation (here, $+\hat{\mathbf{z}}$(© (Modified) CC BY SA 3.0; E. Boutet)

If we observe this wave at some fixed point (i.e., hold z constant), we find that the electric and magnetic fields are also periodic in time; specifically, they have the same value each time t increases by $2\pi/\omega$. We may characterize the speed at which the wave travels by comparing the distance required to experience 2π of phase rotation at a fixed time, which is $1/\beta$; to the time it takes to experience 2π of phase rotation at a fixed location, which is $1/\omega$. This is known as the *phase velocity*[1] v_p:

$$v_p \triangleq \frac{1/\beta}{1/\omega} = \frac{\omega}{\beta}$$

Note that v_p has the units expected from its definition, namely (rad/s)/(rad/m) = m/s. If we make the substitution $\beta = \omega\sqrt{\mu\epsilon}$, we find

$$v_p = \frac{\omega}{\omega\sqrt{\mu\epsilon}} = \frac{1}{\sqrt{\mu\epsilon}}$$

Note that v_p, like the wave impedance η, depends only on material properties. For example, the phase velocity of an electromagnetic wave in free space, given the special symbol c, is

$$c \triangleq v_p|_{\mu=\mu_0,\epsilon=\epsilon_0} = \frac{1}{\sqrt{\mu_0\epsilon_0}} \cong 3.00 \times 10^8 \text{ m/s}$$

This constant is commonly referred to as the *speed of light*, but in fact it is the phase velocity of an electromagnetic field at *any* frequency (not just optical frequencies) in free space. Since the permittivity ϵ and permeability μ of any material is greater than that of a vacuum, v_p in any material is less than the phase velocity in free space. Summarizing:

> Phase velocity is the speed at which any point of constant phase appears to travel along the direction of propagation.

> Phase velocity is maximum ($= c$) in free space, and slower by a factor of $1/\sqrt{\mu_r\epsilon_r}$ in any other lossless medium.

Finally, we note the following relationship between frequency f, wavelength, and phase velocity:

$$\lambda = \frac{2\pi}{\beta} = \frac{2\pi}{\omega/v_p} = \frac{2\pi}{2\pi f/v_p} = \frac{v_p}{f}$$

Thus, given any two of the parameters f, λ, and v_p, we may solve for the remaining quantity. Also note that as a consequence of the inverse-proportional relationship between λ and v_p, we find:

> At a given frequency, the wavelength in any material is shorter than the wavelength in free space.

Example 9.5.1: Wave propagation in a lossless dielectric

Polyethylene is a low-loss dielectric having $\epsilon_r \cong 2.3$. What is the phase velocity in polyethylene? What is wavelength in polyethylene? The frequency of interest is 1 GHz.

Solution

Low-loss dielectrics exhibit $\mu_r \cong 1$ and $\sigma \approx 0$. Therefore the phase velocity is

$$v_p = \frac{1}{\sqrt{\mu_0\epsilon_r\epsilon_0}} = \frac{c}{\sqrt{\epsilon_r}} \cong 1.98 \times 10^8 \text{ m/s}$$

i.e., very nearly two-thirds of the speed of light in free space. The wavelength at 1 GHz is

$$\lambda = \frac{v_p}{f} \cong 19.8 \text{ cm}$$

Again, about two-thirds of a wavelength at the same frequency in free space.

Returning to polarization and magnitude, it is useful to note that Equation 9.5.4 could be written in terms of $\mathbf{E}$ as follows:

$$\mathbf{H} = \frac{1}{\eta}\hat{\mathbf{z}} \times \mathbf{E}$$

i.e., $\mathbf{H}$ is perpendicular to both the direction of propagation (here, $+\hat{\mathbf{z}}$) and $\mathbf{E}$, with magnitude that is different from $\mathbf{E}$ by the wave impedance η. This simple expression is very useful for quickly determining $\mathbf{H}$ given $\mathbf{E}$ when the direction of propagation is known. In fact, it is so useful that it commonly explicitly generalized as follows:

$$\boxed{\mathbf{H} = \frac{1}{\eta}\hat{\mathbf{k}} \times \mathbf{E}} \tag{9.5.5}$$

where $\hat{\mathbf{k}}$ (here, $+\hat{\mathbf{z}}$) is the direction of propagation. Similarly, given $\mathbf{H}$, $\hat{\mathbf{k}}$, and η, we may find $\mathbf{E}$ using

$$\boxed{\mathbf{E} = -\eta\hat{\mathbf{k}} \times \mathbf{H}} \tag{9.5.6}$$

These spatial relationships can be readily verified using Figure 9.5.1.

> Equations 9.5.5 and 9.5.6 are known as the *plane wave relationships*.

The plane wave relationships apply equally well to the phasor representation of bfE and $\mathbf{H}$; i.e.,

$$\begin{aligned} \widetilde{\mathbf{E}} &= -\eta\hat{\mathbf{k}} \times \widetilde{\mathbf{H}} \\ \widetilde{\mathbf{H}} &= \frac{1}{\eta}\hat{\mathbf{k}} \times \widetilde{\mathbf{E}} \end{aligned} \tag{9.5.7}$$

These equations can be readily verified by applying the definition of the associated phasors (e.g., Equation 9.5.3). It also turns out that these relationships apply at each point in space, even if the waves are not planar, uniform, or in lossless media.

Example 9.5.2: Analysis of a radially-directed plane wave.

Consider the scenario illustrated in Figure 9.5.2.

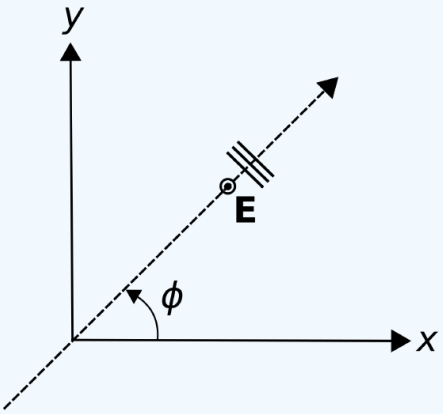


Figure 9.5.2: A radially-directed plane wave.

Here a uniform plane wave having frequency $f = 3$ GHz is propagating along a path of constant ϕ, where is ϕ is known but not specified. The phase of the electric field is $\pi/2$ radians at $\rho = 0$ (the origin) and $t = 0$. The material is an effectively unbounded region of free space. The electric field is oriented in the $+\hat{\mathbf{z}}$ direction and has peak magnitude of 1 mV/m. Find (a) the electric field intensity in phasor representation, (b) the magnetic field intensity in phasor representation, and (c) the actual, physical electric field along the radial path.

Solution

First, realize that this "radially-directed" plane wave is in fact a plane wave, and not a cylindrical wave. It may well we be that if we "zoom out" far enough, we are able to perceive a cylindrical wave (for more on this idea, see Section 9.3), or it might simply be that wave is *exactly* planar, and cylindrical coordinates just happen to be a convenient coordinate system for this application. In either case, the direction of propagation $\hat{\mathbf{k}} = +\hat{\rho}$ and the solution to this example will be the same.

Here's the phasor representation for the electric field intensity of a uniform plane wave in a lossless medium, analogous to Equation 9.5.1:

$$\widetilde{\mathbf{E}} = +\hat{\mathbf{z}}E_0 e^{-j\beta\rho}$$

From the problem statement, $|E_0| = 1$ mV/m. Also from the problem statement, the phase of E_0 is $\pi/2$ radians; in fact, we could just write $E_0 = +j|E_0|$. Thus, the answer to (a) is

$$\widetilde{\mathbf{E}} = +\hat{\mathbf{z}}j|E_0|\, e^{-j\beta\rho}$$

where the propagation constant $\beta = \omega\sqrt{\mu\epsilon} = 2\pi f\sqrt{\mu_0\epsilon_0} \cong 62.9$ rad/m.

The answer to (b) is easiest to obtain from the plane wave relationship:

$$\begin{aligned}\widetilde{\mathbf{H}} &= \frac{1}{\eta}\hat{\mathbf{k}} \times \mathbf{E} \\ &= \frac{1}{\eta}\hat{\rho} \times \left(+\hat{\mathbf{z}} j\,|E_0|\, e^{-j\beta\rho}\right) \\ &= -\hat{\phi}\frac{j\,|E_0|}{\eta} e^{-j\beta\rho}\end{aligned} \tag{9.5.8}$$

where η here is $\sqrt{\mu_0/\epsilon_0} \cong 377\ \Omega$. Thus, the answer to part (b) is

$$\widetilde{\mathbf{H}} = -\hat{\phi} j H_0 e^{-j\beta\rho}$$

where $H_0 \cong 2.65\ \mu$A/m. At this point you should check vector directions: $\widetilde{\mathbf{E}} \times \widetilde{\mathbf{H}}$ should point in the direction of propagation $+\hat{\rho}$. Here we find $\hat{\mathbf{z}} \times \left(-\hat{\phi}\right) = +\hat{\rho}$, as expected.

The answer to (c) is obtained by applying the defining relationship for phasors to the answer to part (a):

$$\begin{aligned}\mathbf{E} &= \mathrm{Re}\left\{\widetilde{\mathbf{E}} e^{j\omega t}\right\} \\ &= \mathrm{Re}\left\{\left(+\hat{\mathbf{z}} j\,|E_0|\, e^{-j\beta\rho}\right) e^{j\omega t}\right\} \\ &= +\hat{\mathbf{z}}\,|E_0|\,\mathrm{Re}\left\{e^{j\pi/2} e^{-j\beta\rho} e^{j\omega t}\right\} \\ &= +\hat{\mathbf{z}}\,|E_0|\cos\left(\omega t - \beta\rho + \frac{\pi}{2}\right)\end{aligned} \tag{9.5.9}$$

1. We acknowledge that this is a misnomer, since velocity is properly defined as speed in a *specified* direction, and v_p by itself does not specify direction. In fact, the phase velocity in this case is properly said to be $+\hat{\mathbf{z}} v_p$. Nevertheless, we adopt the prevailing terminology.↩

This page titled 9.5: Uniform Plane Waves - Characteristics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **9.5: Uniform Plane Waves - Characteristics** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

9.6: Wave Polarization

Polarization refers to the orientation of the electric field vector. For waves, the term "polarization" refers specifically to the orientation of this vector with increasing distance along the direction of propagation, or, equivalently, the orientation of this vector with increasing time at a fixed point in space. The relevant concepts are readily demonstrated for uniform plane waves, as shown in this section. A review of Section 9.5 ("Uniform Plane Waves: Characteristics") is recommended before reading further.

To begin, consider the following uniform plane wave, described in terms of the phasor representation of its electric field intensity:

$$\widetilde{\mathbf{E}}_x = \hat{\mathbf{x}} E_x e^{-j\beta z}$$

Here E_x is a complex-valued constant representing the magnitude and phase of the wave, and β is the positive real-valued propagation constant. Therefore, this wave is propagating in the $+\hat{\mathbf{z}}$ direction in lossless media. This wave is said to exhibit *linear* polarization (and "linearly polarized") because the electric field always points in the same direction, namely $+\hat{\mathbf{x}}$. Now consider the wave

$$\widetilde{\mathbf{E}}_y = \hat{\mathbf{y}} E_y e^{-j\beta z}$$

Note that $\widetilde{\mathbf{E}}_y$ is identical to $\widetilde{\mathbf{E}}_x$ except that the electric field vector now points in the $+\hat{\mathbf{y}}$ direction and has magnitude and phase that is different by the factor E_y/E_x. This wave too is said to exhibit linear polarization, because, again, the direction of the electric field is constant with both time and position. In fact, *all* linearly-polarized uniform plane waves propagating in the $+\hat{\mathbf{z}}$ direction in lossless media can be described as follows:

$$\widetilde{\mathbf{E}} = \hat{\rho} E_\rho e^{-j\beta z}$$

This is so because $\hat{\rho}$ could be $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, or any other direction that is perpendicular to $\hat{\mathbf{z}}$. If one is determined to use Cartesian coordinates, the above expression may be rewritten using (Section 4.3)

$$\hat{\rho} = \hat{\mathbf{x}} \cos\phi + \hat{\mathbf{y}} \cos\phi$$

yielding

$$\widetilde{\mathbf{E}} = \left(\hat{\mathbf{x}} \cos\phi + \hat{\mathbf{y}} \cos\phi\right) E_\rho e^{-j\beta z} \tag{9.6.1}$$

When written in this form, $\phi = 0$ corresponds to $\widetilde{\mathbf{E}} = \widetilde{\mathbf{E}}_x$, $\phi = \pi/2$ corresponds to $\widetilde{\mathbf{E}} = \widetilde{\mathbf{E}}_y$, and any other value of ϕ corresponds to some other constant orientation of the electric field vector; see Figure 9.6.1 for an example.

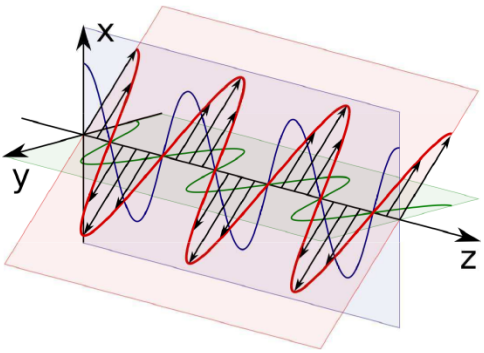


Figure 9.6.1: Linear polarization. Here $\mathbf{E_x}$ is shown in blue, $\mathbf{E_y}$ is shown in green, and bfE is shown in red with black vector symbols. In this example the phases of E_x and E_y are zero and $\phi = -\pi/4$. (Modified; RJB1)

> A wave is said to exhibit linear polarization if the direction of the electric field vector does not vary with either time or position.

Linear polarization arises when the source of the wave is linearly polarized. A common example is the wave radiated by a straight wire antenna, such as a dipole or a monopole. Linear polarization may also be created by passing a plane wave through a *polarizer*; this is particularly common at optical frequencies (see "Additional Reading" at the end of this section).

A commonly-encountered alternative to linear polarization is *circular polarization*. For an explanation, let us return to the linearly-polarized plane waves $\widetilde{\mathbf{E}}_x$ and $\widetilde{\mathbf{E}}_y$ defined earlier. If both of these waves exist simultaneously, then the total electric field intensity is simply the sum:

$$\begin{aligned}\tilde{\mathbf{E}} &= \tilde{\mathbf{E}}_x + \tilde{\mathbf{E}}_y \\ &= \left(\hat{\mathbf{x}}E_x + \hat{\mathbf{y}}E_y\right) e^{-j\beta z}\end{aligned} \tag{9.6.2}$$

If the phase of E_x and E_y is the same, then $E_x = E_\rho \cos\phi$, $E_y = E_\rho \sin\phi$, and the above expression is essentially the same as Equation 9.6.1. In this case, $\widetilde{\mathbf{E}}$ is linearly polarized. But what if the phases of E_x and E_y are different? In particular, let's consider the following case. Let $E_x = E_0$, some complex-valued constant, and let $E_y = +jE_0$, which is E_0 phase-shifted by $+\pi/2$ radians. With no further math, it is apparent that $\widetilde{\mathbf{E}}_x$ and $\widetilde{\mathbf{E}}_y$ are different only in that one is phase-shifted by $\pi/2$ radians relative to the other. For the physical (real-valued) fields, this means that $\mathbf{E}_x$ has maximum magnitude when $\mathbf{E}_y$ is zero and vice versa. As a result, the direction of $\mathbf{E} = \mathbf{E}_x + \mathbf{E}_y$ will rotate in the $x-y$ plane, as shown in Figure 9.6.2.

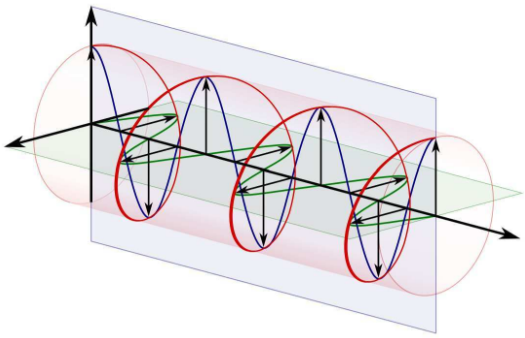


Figure 9.6.2: A circularly-polarized wave (in red, with black vector symbols) resulting from the addition of orthogonal linearly-polarized waves (shown in green and blue) that are phase-shifted by $\pi/2$ radians.

The rotation of the electric field vector can also be identified mathematically. When $E_x = E_0$ and $E_y = +jE_0$, Equation 9.6.2 can be written:

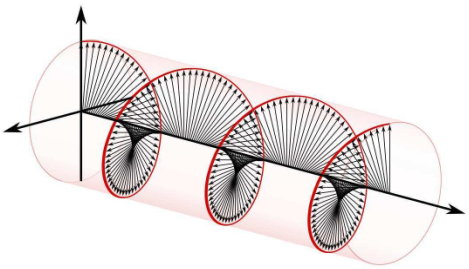
$$\widetilde{\mathbf{E}} = \left(\hat{\mathbf{x}} + j\hat{\mathbf{y}}\right) E_0 e^{-j\beta z}$$

Now reverting from phasor notation to the physical field:

$$\begin{aligned}\mathbf{E} &= \mathrm{Re}\left\{\widetilde{\mathbf{E}}e^{j\omega t}\right\} \\ &= \mathrm{Re}\left\{(\hat{\mathbf{x}} + j\hat{\mathbf{y}})E_0 e^{-j\beta z}e^{j\omega t}\right\} \\ &= \hat{\mathbf{x}}E_0\cos(\omega t - \beta z) + \hat{\mathbf{y}}E_0\cos\left(\omega t - \beta z + \frac{\pi}{2}\right)\end{aligned} \tag{9.6.3}$$

As anticipated, we see that both $\mathbf{E}_x$ and $\mathbf{E}_y$ vary sinusoidally, but are $\pi/2$ radians out of phase resulting in rotation in the plane perpendicular to the direction of propagation.

In the example above, the electric field vector rotates either clockwise or counter-clockwise relative to the direction of propagation. The direction of this rotation can be identified by pointing the thumb of the left hand in the direction of propagation; in this case, the fingers of the left hand curl in the direction of rotation. For this reason, this particular form of circular polarization is known as *left circular* (or "left-hand" circular) polarization (LCP). If we instead had chosen $E_y = -jE_0 = -jE_x$, then the direction of $\mathbf{E}$ rotates in the opposite direction, giving rise to *right circular* (or "right-hand" circular) polarization (RCP). These two conditions are illustrated in Figure 9.6.3.



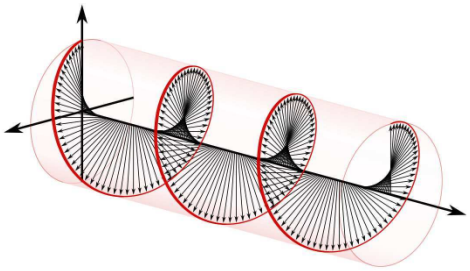


Figure 9.6.3: Left-circular polarization (LCP; top) and right-circular polarization (RCP; bottom)

> A wave is said to exhibit circular polarization if the electric field vector rotates with constant magnitude. Left- and right-circular polarizations may be identified by the direction of rotation with respect to the direction of propagation.

In engineering applications, circular polarization is useful when the relative orientations of transmit and receive equipment is variable and/or when the medium is able rotate the electric field vector. For example, radio communications involving satellites in non-geosynchronous orbits typically employ circular polarization. In particular, satellites of the U.S. Global Positioning System (GPS) transmit circular polarization because of the variable geometry of the space-to-earth radio link and the tendency of the Earth's ionosphere to rotate the electric field vector through a mechanism known *Faraday rotation* (sometimes called the "Faraday effect"). If GPS were instead to transmit using a linear polarization, then a receiver would need to continuously adjust the orientation of its antenna in order to optimally receive the signal. Circularly-polarized radio waves can be generated (or received) using pairs of perpendicularly-oriented dipoles that are fed the same signal but with a 90° phase shift, or alternatively by using an antenna that is intrinsically circularly-polarized, such as a helical antenna (see "Additional Reading" at the end of this section).

Linear and circular polarization are certainly not the only possibilities. *Elliptical polarization* results when E_x and E_y do not have equal magnitude. Elliptical polarization is typically not an intended condition, but rather is most commonly observed as a degradation in a system that is nominally linearly- or circularly-polarized. For example, most antennas that are said to be "circularly polarized" instead produce circular polarization only in one direction and various degrees of elliptical polarization in all other directions.

This page titled 9.6: Wave Polarization is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **9.6: Wave Polarization** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

9.7: Wave Power in a Lossless Medium

In many applications involving electromagnetic waves, one is less concerned with the instantaneous values of the electric and magnetic fields than the *power* associated with the wave. In this section, we address the issue of how much power is conveyed by an electromagnetic wave in a lossless medium. The relevant concepts are readily demonstrated in the context of uniform plane waves, as shown in this section. A review of Section 9.5 ("Uniform Plane Waves: Characteristics") is recommended before reading further.

Consider the following uniform plane wave, described in terms of the phasor representation of its electric field intensity:

$$\widetilde{\mathbf{E}} = \hat{\mathbf{x}} E_0 e^{-j\beta z} \tag{9.7.1}$$

Here E_0 is a complex-valued constant associated with the source of the wave, and β is the positive real-valued propagation constant. Therefore, the wave is propagating in the $+\hat{\mathbf{z}}$ direction in lossless media.

The first thing that should be apparent is that the amount of power conveyed by this wave is infinite. The reason is as follows. If the power passing through any finite area is greater than zero, then the total power must be infinite because, for a uniform plane wave, the electric and magnetic field intensities are constant over a plane of infinite area. In practice, we never encounter this situation because all practical plane waves are only "locally planar" (see Section 9.3 for a refresher on this idea). Nevertheless, we seek some way to express the power associated with such waves.

The solution is not to seek total power, but rather power per unit area. This quantity is known as the *spatial power density*, or simply "*power density*," and has units of W/m^2.[1] Then, if we are interested in total power passing through some finite area, then we may simply integrate the power density over this area. Let's skip to the answer, and then consider where this answer comes from. It turns out that the *instantaneous* power density of a uniform plane wave is the magnitude of the *Poynting vector*

$$\mathbf{S} \triangleq \mathbf{E} \times \mathbf{H} \tag{9.7.2}$$

Note that this equation is dimensionally correct; i.e. the units of $\mathbf{E}$ (V/m) times the units of $\mathbf{H}$ (A/m) yield the units of spatial power density (V·A/m^2, which is W/m^2). Also, the direction of $\mathbf{E} \times \mathbf{H}$ is in the direction of propagation (reminder: Section 9.5), which is the direction in which the power is expected to flow. Thus, we have some compelling evidence that $|\mathbf{S}|$ is the power density we seek. However, this is not proof – for that, we require the *Poynting Theorem*, which is a bit outside the scope of the present section, but is addressed in the "Additional Reading" at the end of this section.

A bit later we're going to need to know $\mathbf{S}$ for a uniform plane wave, so let's work that out now. From the plane wave relationships (Section 9.5) we find that the magnetic field intensity associated with the electric field in Equation 9.7.1 is

$$\widetilde{\mathbf{H}} = \hat{\mathbf{y}} \frac{E_0}{\eta} e^{-j\beta z} \tag{9.7.3}$$

where $\eta = \sqrt{\mu/\epsilon}$ is the real-valued impedance of the medium. Let ψ be the phase of E_0; i.e., $E_0 = |E_0| e^{j\psi}$. Then

$$\begin{aligned} \mathbf{E} &= \mathrm{Re}\left\{\widetilde{\mathbf{E}} e^{j\omega t}\right\} \\ &= \hat{\mathbf{x}} \, |E_0| \cos(\omega t - \beta z + \psi) \end{aligned} \tag{9.7.4}$$

and

$$\begin{aligned} \mathbf{H} &= \mathrm{Re}\left\{\tilde{\mathbf{H}} e^{j\omega t}\right\} \\ &= \hat{\mathbf{y}} \frac{|E_0|}{\eta} \cos(\omega t - \beta z + \psi) \end{aligned} \tag{9.7.5}$$

Now applying Equation 9.7.2,

$$\mathbf{S} = \hat{\mathbf{z}} \frac{|E_0|^2}{\eta} \cos^2(\omega t - \beta z + \psi)$$

As noted earlier, $|\mathbf{S}|$ is only the instantaneous power density, which is still not quite what we are looking for. What we are actually looking for is the *time-average* power density S_{ave} – that is, the average value of $|\mathbf{S}|$ over one period T of the wave. This may be calculated as follows:

$$S_{\text{ave}} = \frac{1}{T} \int^{t_0+T} |\mathbf{S}| dt$$

$$= \frac{|E_0|^2}{\eta} \frac{1}{T} \int_{t=t_0}^{t_0+T} \cos^2(\omega t - ks + \psi) dt \tag{9.7.6}$$

Since $\omega = 2\pi f = 2\pi/T$, the definite integral equals $T/2$. We obtain

$$\boxed{S_{ave} = \frac{|E_0|^2}{2\eta}} \tag{9.7.7}$$

It is useful to check units again at this point. Note $(\text{V/m})^2$ divided by Ω is W/m^2, as expected.

> Equation 9.7.7 is the time-average power density (units of W/m^2) associated with a sinusoidally-varying uniform plane wave in lossless media.

Note that Equation 9.7.7 is analogous to a well-known result from electric circuit theory. Recall the time-average power P_{ave} (units of W) associated with a voltage phasor $\widetilde{V}$ across a resistance R is

$$P_{ave} = \frac{|\widetilde{V}|^2}{2R} \tag{9.7.8}$$

which closely resembles Equation 9.7.7. The result is also analogous to the result for a voltage wave on a transmission line (Section 3.20), for which:

$$P_{ave} = \frac{|V_0^+|^2}{2Z_0}$$

where V_0^+ is a complex-valued constant representing the magnitude and phase of the voltage wave, and Z_0 is the characteristic impedance of the transmission line.

Here is a good point at which to identify a common pitfall. $|E_0|$ and $|\widetilde{V}|$ are the *peak magnitudes* of the associated real-valued physical quantities. However, these quantities are also routinely given as *root mean square* ("rms") quantities. Peak magnitudes are greater by a factor of $\sqrt{2}$, so Equation 9.7.7 expressed in terms of the rms quantity lacks the factor of $1/2$.

Example 9.7.1: Power density of a typical radio wave

A radio wave transmitted from a distant location may be perceived locally as a uniform plane wave if there is no nearby structure to scatter the wave; a good example of this is the wave arriving at the user of a cellular telephone in a rural area with no significant terrain scattering. The range of possible signal strengths varies widely, but a typical value of the electric field intensity arriving at the user's location is 10 μV/m rms. What is the corresponding power density?

Solution

From the problem statement, $|E_0| = 10\ \mu\text{V/m}$ rms. We assume propagation occurs in air, which is indistinguishable from free space at cellular frequencies. If we use Equation 9.7.7, then we must first convert $|E_0|$ from rms to peak magnitude, which is done by multiplying by $\sqrt{2}$. Thus:

$$S_{\text{are}} = \frac{|E_0|^2}{2\eta} \cong \frac{\left(\sqrt{2} \cdot 10 \times 10^{-6}\ \text{V/m}\right)^2}{2 \cdot 377\ \Omega} \cong 2.65 \times 10^{-13}\ \text{W/m}^2 \tag{9.7.9}$$

Alternatively, we can just use a version of Equation 9.7.7 which is appropriate for rms units:

$$S_{\text{ave}} = \frac{|E_{0,\text{ rms}}|^2}{\eta} \cong \frac{\left(10 \times 10^{-6}\ \text{V/m}\right)^2}{377\ \Omega} \cong 2.65 \times 10^{-13}\ \text{W/m}^2 \tag{9.7.10}$$

Either way, we obtain the correct answer, 0.265 pW/m^2 (that's *picowatts* per square meter).

Considering the prevalence of phasor representation, it is useful to have an alternative form of the Poynting vector which yields time-average power by operating directly on field quantities in phasor form. This is $\mathbf{S}_{ave}$, defined as:

$$\mathbf{S}_{ave} \triangleq \frac{1}{2}\mathrm{Re}\left\{\widetilde{\mathbf{E}} \times \widetilde{\mathbf{H}}^{*}\right\}$$

(Note that the magnetic field intensity phasor is conjugated.) The above expression gives the expected result for a uniform plane wave. Using Equations 9.7.1 and 9.7.3, we find

$$\mathbf{S}_{ave} = \frac{1}{2}\mathrm{Re}\left\{\left(\hat{\mathbf{x}}E_0 e^{-j\beta z}\right) \times \left(\hat{\mathbf{y}}\frac{E_0}{\eta}e^{-j\beta z}\right)^{*}\right\}$$

which yields

$$\mathbf{S}_{ave} = \hat{\mathbf{z}}\frac{|E_0|^2}{2\eta}$$

as expected.

---

1. Be careful: The quantities *power spectral density* (W/Hz) and *power flux density* (W/(m^2·Hz)) are also sometimes referred to as "power density." In this section, we will limit the scope to *spatial* power density (W/m^2).↩

This page titled 9.7: Wave Power in a Lossless Medium is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **9.7: Wave Power in a Lossless Medium** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

CHAPTER OVERVIEW

10: Appendices

10.1: Permittivity of Some Common Materials
10.2: Permeability of Some Common Materials
10.3: Conductivity of Some Common Materials
10.4: Mathematical Formulas - Trigonometry
10.5: Mathematical Formulas - Vector Operators
10.6: Mathematical Formulas - Vector Identities
10.7: Physical Constants

Contributors and Attributions

- Ellingson, Steven W. (2018) Electromagnetics, Vol. 1. Blacksburg, VA: VT Publishing. https://doi.org/10.21061/electromagnetics-vol-1 Licensed with CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0. Report adoption of this book here. If you are a professor reviewing, adopting, or adapting this textbook please help us understand a little more about your use by filling out this form.

This page titled 10: Appendices is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

10.1: Permittivity of Some Common Materials

The values below are relative permittivity $\epsilon_r \triangleq \epsilon/\epsilon_0$ for a few materials that are commonly encountered in electrical engineering applications, and for which permittivity emerges as a consideration. Note that "relative permittivity" is sometimes referred to as *dielectric constant*.

Here we consider only the physical (real-valued) permittivity, which is the real part of the complex permittivity (typically indicated as ϵ' or ϵ_r') for materials exhibiting significant loss.

Permittivity varies significantly as a function of frequency. The values below are representative of frequencies from a few kHz to about 1 GHz. The values given are also representative of optical frequencies for materials such as silica that are used in optical applications. Permittivity also varies as a function of temperature. In applications where precision better than about 10% is required, primary references accounting for frequency and temperature should be consulted. The values presented here are gathered from a variety of references, including those indicated in "Additional References."

Free Space (vacuum): $\epsilon_r \triangleq 1$

Table 10.1.1: Solid Dielectrics:

Material	ϵ_r	Common uses
Styrofoam[1]	1.1	
Teflon[2]	2.1	
Polyethylene	2.3	coaxial cable
Polypropylene	2.3	
Silica	2.4	optical fiber[3]
Polystyrene	2.6	
Polycarbonate	2.8	
Rogers RO3003	3.0	PCB substrate
FR4 (glass epoxy laminate)	4.5	PCB substrate

[1] Properly known as *extruded polystyrene foam* (XPS).
[2] Properly known as *polytetrafluoroethylene* (PTFE).
[3] Typically doped with small amounts of other materials to slightly raise or lower the index of refraction ($= \sqrt{\epsilon_r}$).

Non-conducting spacing materials used in discrete capacitors exhibit ϵ_r ranging from about 5 to 50.

- **Semiconductors** commonly appearing in electronics – including carbon, silicon, geranium, indium phosphide, and so on – typically exhibit ϵ_r in the range 5–15.
- **Glass** exhibits ϵ_r in the range 4–10, depending on composition.
- **Gasses**, including air, typically exhibit $\epsilon_r \cong 1$ to within a tiny fraction of a percent.
- **Liquid water** typically exhibits ϵ_r in the range 72–81. Distilled water exhibits $\epsilon_r \approx 81$ at room temperature, whereas sea water tends to be at the lower end of the range.
- **Other liquids** typically exhibit ϵ_r in the range 10–90, with considerable variation as a function of temperature and frequency. Animal flesh and blood consists primarily of liquid matter and so also exhibits permittivity in this range.
- **Soil** typically exhibits ϵ_r in the range 2.5–3.5 when dry and higher when wet. The permittivity of soil varies considerably depending on composition.

This page titled 10.1: Permittivity of Some Common Materials is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **10.1: Permittivity of Some Common Materials** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

10.2: Permeability of Some Common Materials

The values below are relative permeability $\mu_r \triangleq \mu/\mu_0$ for a few materials that are commonly encountered in electrical engineering applications, and for which μ_r is significantly different from 1. These materials are predominantly ferromagnetic metals and (in the case of ferrites) materials containing significant ferromagnetic metal content. Nearly all other materials exhibit μ_r that is not significantly different from that of free space.

The values presented here are gathered from a variety of references, including those indicated in "Additional References" at the end of this section. Be aware that permeability may vary significantly with frequency; values given here are applicable to the frequency ranges for applications in which these materials are typically used. Also be aware that materials exhibiting high permeability are also typically non-linear; that is, permeability depends on the magnitude of the magnetic field. Again, values reported here are those applicable to applications in which these materials are typically used.

- **Free Space** (vacuum): $\mu_r \triangleq 1$.
- **Iron** (also referred to by the chemical notation "Fe") appears as a principal ingredient in many materials and alloys employed in electrical structures and devices. Iron exhibits μ_r that is very high, but which decreases with decreasing purity. 99.95% pure iron exhibits $\mu_r \sim 200,000$. This decreases to ~ 5000 at 99.8% purity and is typically below 100 for purity less than 99%.
- **Steel** is an iron alloy that comes in many forms, with a correspondingly broad range of permeabilites. *Electrical steel*, commonly used in electrical machinery and transformers when high permeability is desired, exhibits $\mu_r \sim 4000$. *Stainless steel*, encompassing a broad range of alloys used in mechanical applications, exhibits μ_r in the range 750–1800. *Carbon steel*, including a broad class of alloys commonly used in structural applications, exhibits μ_r on the order of 100.
- **Ferrites** include a broad range of ceramic materials that are combined with iron and various combinations of other metals and are used as magnets and magnetic devices in various electrical systems. Common ferrites exhibit μ_r in the range 16–640.

This page titled 10.2: Permeability of Some Common Materials is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **10.2: Permeability of Some Common Materials** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

10.3: Conductivity of Some Common Materials

The values below are conductivity σ for a few materials that are commonly encountered in electrical engineering applications, and for which conductivity emerges as a consideration.

Note that materials in some applications are described instead in terms of *resistivity*, which is simply the reciprocal of conductivity.

Conductivity may vary significantly as a function of frequency. The values below are representative of frequencies from a few kHz to a few GHz. Conductivity also varies as a function of temperature. In applications where precise values are required, primary references accounting for frequency and temperature should be consulted. The values presented here are gathered from a variety of references, including those indicated in “Additional References” at the end of this section.

Free Space (vacuum): $\sigma \triangleq 0$.

Figure 10.3.1: Commonly encountered elements:

Material	σ (S/m)
Copper	5.8×10^7
Gold	4.4×10^7
Aluminum	3.7×10^7
Iron	1.0×10^7
Platinum	0.9×10^7
Carbon	1.3×10^5
Silicon	4.4×10^{-4}

Water exhibits σ ranging from about 6 μS/m for highly distilled water (thus, a very poor conductor) to about 5 S/m for seawater (thus, a relatively good conductor), varying also with temperature and pressure. Tap water is typically in the range 5–50 mS/m, depending on the level of impurities present.

Soil typically exhibits σ in the range 10^{-4} S/m for dry soil to about 10^{-1} S/m for wet soil, varying also due to chemical composition.

Non-conductors. Most other materials that are not well-described as conductors or semiconductors and are dry exhibit $\sigma < 10^{-12}$ S/m. Most materials that are considered to be *insulators*, including air and common dielectrics, exhibit $\sigma < 10^{-15}$ S/m, often by several orders of magnitude.

This page titled 10.3: Conductivity of Some Common Materials is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **10.3: Conductivity of Some Common Materials** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

10.4: Mathematical Formulas - Trigonometry

$$e^{j\theta} = \cos\theta + j\sin\theta$$

$$\cos\theta = \frac{1}{2}\left(e^{j\theta} + e^{-j\theta}\right)$$

$$\sin\theta = \frac{1}{j2}\left(e^{j\theta} - e^{-j\theta}\right)$$

$$\cos^2\theta = \frac{1}{2} + \frac{1}{2}\cos 2\theta$$

$$\sin^2\theta = \frac{1}{2} - \frac{1}{2}\cos 2\theta$$

This page titled 10.4: Mathematical Formulas - Trigonometry is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **10.4: Mathematical Formulas - Trigonometry** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

10.5: Mathematical Formulas - Vector Operators

This section contains a summary of vector operators expressed in each of the three major coordinate systems:

- Cartesian (x,y,z)
- cylindrical (ρ,ϕ,z)
- spherical (r,θ,ϕ)

Associated basis vectors are identified using a caret ($\hat{}$) over the symbol. The vector operand $\mathbf{A}$ is expressed in terms of components in the basis directions as follows:

- Cartesian: $\mathbf{A} = \hat{\mathbf{x}}A_x + \hat{\mathbf{y}}A_y + \hat{\mathbf{z}}A_z$
- cylindrical: $\mathbf{A} = \hat{\rho}A_\rho + \hat{\phi}A_\phi + \hat{\mathbf{z}}A_z$
- spherical: $\mathbf{A} = \hat{\mathbf{r}}A_r + \hat{\theta}A_\theta + \hat{\phi}A_\phi$

Gradient

Gradient in Cartesian coordinates:

$$\nabla f = \hat{\mathbf{x}}\frac{\partial f}{\partial x} + \hat{\mathbf{y}}\frac{\partial f}{\partial y} + \hat{\mathbf{z}}\frac{\partial f}{\partial z}$$

Gradient in cylindrical coordinates:

$$\nabla f = \hat{\rho}\frac{\partial f}{\partial \rho} + \hat{\phi}\frac{1}{\rho}\frac{\partial f}{\partial \phi} + \hat{\mathbf{z}}\frac{\partial f}{\partial z}$$

Gradient in spherical coordinates:

$$\nabla f = \hat{\mathbf{r}}\frac{\partial f}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial f}{\partial \theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}$$

Divergence

Divergence in Cartesian coordinates:

$$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

Divergence in cylindrical coordinates:

$$\nabla \cdot \mathbf{A} = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho A_\rho\right) + \frac{1}{\rho}\frac{\partial A_\phi}{\partial \phi} + \frac{\partial A_z}{\partial z}$$

Divergence in spherical coordinates:

$$\begin{aligned}\nabla \cdot \mathbf{A} = & \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 A_r\right) \\ & + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}\left(A_\theta \sin\theta\right) \\ & + \frac{1}{r\sin\theta}\frac{\partial A_\phi}{\partial \phi}\end{aligned}$$

Curl

Curl in Cartesian coordinates:

$$\begin{aligned}\nabla \times \mathbf{A} = & \hat{\mathbf{x}}\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) \\ & + \hat{\mathbf{y}}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) \\ & + \hat{\mathbf{z}}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\end{aligned}$$

Curl in cylindrical coordinates:

$$\begin{aligned}\nabla \times \mathbf{A} =& \hat{\rho}\left(\frac{1}{\rho}\frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z}\right) \\ &+ \hat{\phi}\left(\frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial \rho}\right) \\ &+ \hat{\mathbf{z}}\frac{1}{\rho}\left[\frac{\partial}{\partial \rho}\left(\rho A_\phi\right) - \frac{\partial A_\rho}{\partial \phi}\right]\end{aligned}$$

Curl in spherical coordinates:

$$\begin{aligned}\nabla \times \mathbf{A} =& \hat{\mathbf{r}}\frac{1}{r\sin\theta}\left[\frac{\partial}{\partial \theta}\left(A_\phi \sin\theta\right) - \frac{\partial A_\theta}{\partial \phi}\right] \\ &+ \hat{\theta}\frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial A_r}{\partial \phi} - \frac{\partial}{\partial r}\left(rA_\phi\right)\right] \\ &+ \hat{\phi}\frac{1}{r}\left[\frac{\partial}{\partial r}\left(rA_\theta\right) - \frac{\partial A_r}{\partial \theta}\right]\end{aligned}$$

Laplacian

Laplacian in Cartesian coordinates:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

Laplacian in cylindrical coordinates:

$$\nabla^2 f = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho\frac{\partial f}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}$$

Laplacian in spherical coordinates:

$$\begin{aligned}\nabla^2 f =& \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) \\ &+ \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial \theta}\sin\theta\right) \\ &+ \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \phi^2}\end{aligned}$$

This page titled 10.5: Mathematical Formulas - Vector Operators is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **10.5: Mathematical Formulas - Vector Operators** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

10.6: Mathematical Formulas - Vector Identities

Algebraic Identities

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})=\mathbf{B}\cdot(\mathbf{C}\times\mathbf{A})=\mathbf{C}\cdot(\mathbf{A}\times\mathbf{B}) \tag{10.6.1}$$

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C})=\mathbf{B}(\mathbf{A}\cdot\mathbf{C})-\mathbf{C}(\mathbf{A}\cdot\mathbf{B}) \tag{10.6.2}$$

Identities Involving Differential Operators

$$\nabla\cdot(\nabla\times\mathbf{A})=0 \tag{10.6.3}$$

$$\nabla\times(\nabla f)=0 \tag{10.6.4}$$

$$\nabla\times(f\mathbf{A})=f(\nabla\times\mathbf{A})+(\nabla f)\times\mathbf{A} \tag{10.6.5}$$

$$\nabla\cdot(\mathbf{A}\times\mathbf{B})=\mathbf{B}\cdot(\nabla\times\mathbf{A})-\mathbf{A}\cdot(\nabla\times\mathbf{B}) \tag{10.6.6}$$

$$\nabla\cdot(\nabla f)=\nabla^2 f \tag{10.6.7}$$

$$\nabla\times\nabla\times\mathbf{A}=\nabla(\nabla\cdot\mathbf{A})-\nabla^2\mathbf{A} \tag{10.6.8}$$

$$\nabla^2\mathbf{A}=\nabla(\nabla\cdot\mathbf{A})-\nabla\times(\nabla\times\mathbf{A}) \tag{10.6.9}$$

Divergence Theorem

Given a closed surface $\mathcal{S}$ enclosing a contiguous volume $\mathcal{V}$,

$$\int_{\mathcal{V}}(\nabla\cdot\mathbf{A})\,dv=\oint_{\mathcal{S}}\mathbf{A}\cdot d\mathbf{s}$$

where the surface normal $d\mathbf{s}$ is pointing out of the volume.

Stokes' Theorem

Given a closed curve $\mathcal{C}$ bounding a contiguous surface $\mathcal{S}$,

$$\int_{\mathcal{S}}(\nabla\times\mathbf{A})\cdot d\mathbf{s}=\oint_{\mathcal{C}}\mathbf{A}\cdot d\mathbf{l}$$

where the direction of the surface normal $d\mathbf{s}$ is related to the direction of integration along $\mathcal{C}$ by the "right hand rule."

This page titled 10.6: Mathematical Formulas - Vector Identities is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **10.6: Mathematical Formulas - Vector Identities** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

10.7: Physical Constants

The speed of light in free space (c), which is the phase velocity of any electromagnetic radiation in free space, is $\cong 2.9979 \times 10^8$ m/s. This is commonly rounded up to 3×10^8 m/s. This rounding incurs error of $\cong 0.07\%$, which is usually much less than other errors present in electrical engineering calculations.

The charge of an electron is $\cong -1.602 \times 10^{-19}$ C. The constant $e \triangleq +1.602176634 \times 10^{-19}$ C is known as the "elementary charge," so the charge of the electron is said to be $-e$.

The permittivity of free space (ϵ_0) is $\cong 8.854 \times 10^{-12}$ F/m.

The permeability of free space (μ_0) is $4\pi \times 10^{-7}$ H/m.

The wave impedance of free space (η_0) is the ratio of the magnitude of the electric field intensity to that of the magnetic field intensity in free space and is $\sqrt{\mu_0/\epsilon_0} \cong 376.7\ \Omega$. This is also sometimes referred to as the *intrinsic impedance of free space*.

This page titled 10.7: Physical Constants is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Steven W. Ellingson (Virginia Tech Libraries' Open Education Initiative) .

- **10.7: Physical Constants** by Steven W. Ellingson is licensed CC BY-SA 4.0. Original source: https://doi.org/10.21061/electromagnetics-vol-1.

Index

A

acoustic wave equation
1.3: Fundamentals of Waves

Ampère's law
4.8: Curl
7.4: Ampere's Circuital Law (Magnetostatics) - Integral Form
7.9: Ampere's Law (Magnetostatics) - Differential Form
8.9: Displacement Current and Ampere's Law

Amperes' Circuital Law for magnetostatics
7.9: Ampere's Law (Magnetostatics) - Differential Form

B

balun
8.6: Transformers as Two-Port Devices

boundary conditions
5.17: Boundary Conditions on the Electric Field Intensity (E)

C

capacitance
5.22: Capacitance

capacitor
5.16: Potential Field Within a Parallel Plate Capacitor

Cartesian Coordinate System
4.1: Vector Arithmetic

cartesian coordinates
4.2: Cartesian Coordinates

characteristic impedance
3.7: Characteristic Impedance

charge distribution
5.3: Charge Distributions

charge of an electron
10.7: Physical Constants

circular polarization
9.6: Wave Polarization

circulation
4.8: Curl

Coaxial Line
3.10: Coaxial Line
5.24: Capacitance of a Coaxial Structure

conductance
6.5: Conductance

Conduction current
6: Steady Current and Conductivity

conductivity
2.8: Electromagnetic Properties of Materials
6.3: Conductivity

Continuous Charge Distribution
5.4: Electric Field Due to a Continuous Distribution of Charge

Convection current
6: Steady Current and Conductivity

Coulomb's Law
5.1: Coulomb's Law

cross product
4.1: Vector Arithmetic

curl
4.8: Curl

Curves
1.7: Notation

cylindrical coordinates
4.3: Cylindrical Coordinates

D

Dielectric Breakdown
5.21: Dielectric Breakdown

dielectric constant
10.1: Permittivity of Some Common Materials

dielectrics
5.20: Dielectric Media

differential air pressure
1.3: Fundamentals of Waves

Displacement current
8.9: Displacement Current and Ampere's Law

divergence
4.6: Divergence

divergence theorem
4.7: Divergence Theorem

dot product
4.1: Vector Arithmetic

duality
7.1: Comparison of Electrostatics and Magnetostatics

E

electric field intensity
2.2: Electric Field Intensity
5.17: Boundary Conditions on the Electric Field Intensity (E)

Electric Flux Density
2.4: Electric Flux Density

Electric Generator
8.7: The Electric Generator

electromagnetic interference
3.1: Introduction to Transmission Lines

electromagnetic spectrum
1.2: Electromagnetic Spectrum

electromagnetic waves
1.1: What is Electromagnetics?

Electrostatic Energy
5.25: Electrostatic Energy

F

Faraday rotation
9.6: Wave Polarization

Faraday's Law
8.3: Faraday's Law

ferromagnetism
2.6: Permeability

field
2.1: What is a Field?

flux
2.4: Electric Flux Density

Fourier analysis
1.5: Phasors

frequency
1.2: Electromagnetic Spectrum

fringing field
5.23: The Thin Parallel Plate Capacitor

G

Gauss' Law
5.5: Gauss' Law - Integral Form
5.7: Gauss' Law - Differential Form

Gauss' Law for Magnetic Fields
7.2: Gauss' Law for Magnetic Fields - Integral Form
7.3: Gauss' Law for Magnetism - Differential Form

geometry
1.1: What is Electromagnetics?

gradient
4.5: Gradient

guided waves
1.4: Guided and Unguided Waves
3.1: Introduction to Transmission Lines

I

imaginary number
1.5: Phasors

impedance matching
3.21: Impedance Matching - General Considerations

inductance
7.12: Inductance

induction
8.2: Electromagnetic Induction
8.3: Faraday's Law

infinite line charge
5.6: Electric Field Due to an Infinite Line Charge using Gauss' Law

Input Impedance
3.15: Input Impedance of a Terminated Lossless Transmission Line

Integration Over Area
4.2: Cartesian Coordinates
4.3: Cylindrical Coordinates
4.4: Spherical Coordinates

Integration Over Length
4.2: Cartesian Coordinates
4.3: Cylindrical Coordinates
4.4: Spherical Coordinates

Integration Over Volume
4.2: Cartesian Coordinates
4.3: Cylindrical Coordinates
4.4: Spherical Coordinates

inverse square law
2.3: Permittivity

J

joule heating
6.6: Power Dissipation in Conducting Media

K

Kirchoff's Voltage Law
5.10: Kirchoff's Voltage Law for Electrostatics - Integral Form
5.11: Kirchoff's Voltage Law for Electrostatics - Differential Form

L

Laplace's equation
5.15: Poisson's and Laplace's Equations

Laplacian
4.10: The Laplacian Operator

Lenz's law
8.2: Electromagnetic Induction

Line Charge Distribution
5.3: Charge Distributions
Line Current
6.2: Current Distributions

M
magnetic field containment
7.7: Magnetic Field of a Toroidal Coil
magnetic flux
7.12: Inductance
Magnetic flux density
2.5: Magnetic Flux Density
magnetostatics
7: Magnetostatics
7.4: Ampere's Circuital Law (Magnetostatics) - Integral Form
7.9: Ampere's Law (Magnetostatics) - Differential Form
Maxwell's equations
8.1: Comparison of Static and Time-Varying Electromagnetics
Microstrip Line
3.11: Microstrip Line
mutual inductance
7.12: Inductance

N
notation
1.7: Notation

O
Ohm's Law for Electromagnetics
6.3: Conductivity
ohmic heating
6.6: Power Dissipation in Conducting Media

P
Parallel Plate Capacitor
5.16: Potential Field Within a Parallel Plate Capacitor
path function
5.9: Independence of Path
perfect electrical conductor
5.19: Charge and Electric Field for a Perfectly Conducting Region
permeability
2.6: Permeability
2.8: Electromagnetic Properties of Materials
10.2: Permeability of Some Common Materials
permeability of free space
10.7: Physical Constants
permittivity
2.3: Permittivity
2.8: Electromagnetic Properties of Materials
10.1: Permittivity of Some Common Materials
permittivity of free space
10.7: Physical Constants
phase propagation constant
4.10: The Laplacian Operator
phase velocity
1.3: Fundamentals of Waves
9.5: Uniform Plane Waves - Characteristics
phasefronts
9.3: Types of Waves
phasor
1.5: Phasors
1.7: Notation
photoelectric effect
1.2: Electromagnetic Spectrum
plane waves
9.4: Uniform Plane Waves - Derivation
Plate Capacitor
5.23: The Thin Parallel Plate Capacitor
Poisson equation
5.15: Poisson's and Laplace's Equations
Poisson's equation
4.10: The Laplacian Operator
polarization
9.6: Wave Polarization
position
1.7: Notation
poynting vector
9.7: Wave Power in a Lossless Medium
properties of materials
1.1: What is Electromagnetics?

R
resistance
6.4: Resistance
resistors
6.4: Resistance

S
SI units
1.6: Units
sound
1.3: Fundamentals of Waves
speed of light in free space
10.7: Physical Constants
spherical coordinates
4.4: Spherical Coordinates
standing waves
3.13: Standing Waves
Stokes' theorem
4.9: Stokes' Theorem
surface
1.7: Notation
Surface Charge Distribution
5.3: Charge Distributions
Surface Current Distribution
6.2: Current Distributions

T
Telegrapher's Equations
3.5: Telegrapher's Equations
thin parallel plate capacitor
5.23: The Thin Parallel Plate Capacitor
time
1.7: Notation
Toroidal Coil
7.7: Magnetic Field of a Toroidal Coil
Transformer
8.5: Transformers - Principle of Operation

U
unguided waves
1.4: Guided and Unguided Waves
Uniform Plane Waves
9.4: Uniform Plane Waves - Derivation
9.5: Uniform Plane Waves - Characteristics
unit vectors
1.7: Notation
Units
1.6: Units

V
Vector arithmetic
4.1: Vector Arithmetic
Vector Operators
10.5: Mathematical Formulas - Vector Operators
vectors
1.7: Notation
volume
1.7: Notation
Volume Charge Distribution
5.3: Charge Distributions
Volume Current Distribution
6.2: Current Distributions

W
wave
1.3: Fundamentals of Waves
wave impedance of free space
10.7: Physical Constants
wavelength
1.3: Fundamentals of Waves
wavenumber
1.3: Fundamentals of Waves

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Electromagnetics I (Ellingson)

Webpages: 138

All licenses found:

- CC BY-SA 4.0: 93.5% (129 pages)
- Undeclared: 6.5% (9 pages)

By Page

- Electromagnetics I (Ellingson) - *CC BY-SA 4.0*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Preliminary Concepts - *CC BY-SA 4.0*
    - 1.1: What is Electromagnetics? - *CC BY-SA 4.0*
    - 1.2: Electromagnetic Spectrum - *CC BY-SA 4.0*
    - 1.3: Fundamentals of Waves - *CC BY-SA 4.0*
    - 1.4: Guided and Unguided Waves - *CC BY-SA 4.0*
    - 1.5: Phasors - *CC BY-SA 4.0*
    - 1.6: Units - *CC BY-SA 4.0*
    - 1.7: Notation - *CC BY-SA 4.0*
  - 2: Electric and Magnetic Fields - *CC BY-SA 4.0*
    - 2.1: What is a Field? - *CC BY-SA 4.0*
    - 2.2: Electric Field Intensity - *CC BY-SA 4.0*
    - 2.3: Permittivity - *CC BY-SA 4.0*
    - 2.4: Electric Flux Density - *CC BY-SA 4.0*
    - 2.5: Magnetic Flux Density - *CC BY-SA 4.0*
    - 2.6: Permeability - *CC BY-SA 4.0*
    - 2.7: Magnetic Field Intensity - *CC BY-SA 4.0*
    - 2.8: Electromagnetic Properties of Materials - *CC BY-SA 4.0*
  - 3: Transmission Lines - *CC BY-SA 4.0*
    - 3.1: Introduction to Transmission Lines - *CC BY-SA 4.0*
    - 3.2: Types of Transmission Lines - *CC BY-SA 4.0*
    - 3.3: Transmission Lines as Two-Port Devices - *CC BY-SA 4.0*
    - 3.4: Lumped-Element Model - *CC BY-SA 4.0*
    - 3.5: Telegrapher's Equations - *CC BY-SA 4.0*
    - 3.6: Wave Equation for a TEM Transmission Line - *CC BY-SA 4.0*
    - 3.7: Characteristic Impedance - *CC BY-SA 4.0*
    - 3.8: Wave Propagation on a TEM Transmission Line - *CC BY-SA 4.0*
    - 3.9: Lossless and Low-Loss Transmission Lines - *CC BY-SA 4.0*
    - 3.10: Coaxial Line - *CC BY-SA 4.0*
    - 3.11: Microstrip Line - *CC BY-SA 4.0*
    - 3.12: Voltage Reflection Coefficient - *CC BY-SA 4.0*
    - 3.13: Standing Waves - *CC BY-SA 4.0*
    - 3.14: Standing Wave Ratio - *CC BY-SA 4.0*
    - 3.15: Input Impedance of a Terminated Lossless Transmission Line - *CC BY-SA 4.0*
    - 3.16: Input Impedance for Open- and Short-Circuit Terminations - *CC BY-SA 4.0*
    - 3.17: Applications of Open- and Short-Circuited Transmission Line Stubs - *CC BY-SA 4.0*
    - 3.18: Measurement of Transmission Line Characteristics - *CC BY-SA 4.0*
    - 3.19: Quarter-Wavelength Transmission Line - *CC BY-SA 4.0*
    - 3.20: Power Flow on Transmission Lines - *CC BY-SA 4.0*
    - 3.21: Impedance Matching - General Considerations - *CC BY-SA 4.0*
    - 3.22: Single-Reactance Matching - *CC BY-SA 4.0*
    - 3.23: Single-Stub Matching - *CC BY-SA 4.0*
  - 4: Vector Analysis - *CC BY-SA 4.0*
    - 4.1: Vector Arithmetic - *CC BY-SA 4.0*
    - 4.2: Cartesian Coordinates - *CC BY-SA 4.0*
    - 4.3: Cylindrical Coordinates - *CC BY-SA 4.0*
    - 4.4: Spherical Coordinates - *CC BY-SA 4.0*
    - 4.5: Gradient - *CC BY-SA 4.0*
    - 4.6: Divergence - *CC BY-SA 4.0*
    - 4.7: Divergence Theorem - *CC BY-SA 4.0*
    - 4.8: Curl - *CC BY-SA 4.0*
    - 4.9: Stokes' Theorem - *CC BY-SA 4.0*
    - 4.10: The Laplacian Operator - *CC BY-SA 4.0*
  - 5: Electrostatics - *CC BY-SA 4.0*
    - 5.1: Coulomb's Law - *CC BY-SA 4.0*
    - 5.2: Electric Field Due to Point Charges - *CC BY-SA 4.0*
    - 5.3: Charge Distributions - *CC BY-SA 4.0*

- 5.4: Electric Field Due to a Continuous Distribution of Charge - *CC BY-SA 4.0*
- 5.5: Gauss' Law - Integral Form - *CC BY-SA 4.0*
- 5.6: Electric Field Due to an Infinite Line Charge using Gauss' Law - *CC BY-SA 4.0*
- 5.7: Gauss' Law - Differential Form - *CC BY-SA 4.0*
- 5.8: Force, Energy, and Potential Difference - *CC BY-SA 4.0*
- 5.9: Independence of Path - *CC BY-SA 4.0*
- 5.10: Kirchoff's Voltage Law for Electrostatics - Integral Form - *CC BY-SA 4.0*
- 5.11: Kirchoff's Voltage Law for Electrostatics - Differential Form - *CC BY-SA 4.0*
- 5.12: Electric Potential Field Due to Point Charges - *CC BY-SA 4.0*
- 5.13: Electric Potential Field due to a Continuous Distribution of Charge - *CC BY-SA 4.0*
- 5.14: Electric Field as the Gradient of Potential - *CC BY-SA 4.0*
- 5.15: Poisson's and Laplace's Equations - *CC BY-SA 4.0*
- 5.16: Potential Field Within a Parallel Plate Capacitor - *CC BY-SA 4.0*
- 5.17: Boundary Conditions on the Electric Field Intensity (E) - *CC BY-SA 4.0*
- 5.18: Boundary Conditions on the Electric Flux Density (D) - *CC BY-SA 4.0*
- 5.19: Charge and Electric Field for a Perfectly Conducting Region - *CC BY-SA 4.0*
- 5.20: Dielectric Media - *CC BY-SA 4.0*
- 5.21: Dielectric Breakdown - *CC BY-SA 4.0*
- 5.22: Capacitance - *CC BY-SA 4.0*
- 5.23: The Thin Parallel Plate Capacitor - *CC BY-SA 4.0*
- 5.24: Capacitance of a Coaxial Structure - *CC BY-SA 4.0*
- 5.25: Electrostatic Energy - *CC BY-SA 4.0*

6: Steady Current and Conductivity - *CC BY-SA 4.0*

- 6.1: Convection and Conduction Currents - *CC BY-SA 4.0*
- 6.2: Current Distributions - *CC BY-SA 4.0*
- 6.3: Conductivity - *CC BY-SA 4.0*
- 6.4: Resistance - *CC BY-SA 4.0*
- 6.5: Conductance - *CC BY-SA 4.0*
- 6.6: Power Dissipation in Conducting Media - *CC BY-SA 4.0*

7: Magnetostatics - *CC BY-SA 4.0*

- 7.1: Comparison of Electrostatics and Magnetostatics - *CC BY-SA 4.0*
- 7.2: Gauss' Law for Magnetic Fields - Integral Form - *CC BY-SA 4.0*
- 7.3: Gauss' Law for Magnetism - Differential Form - *CC BY-SA 4.0*
- 7.4: Ampere's Circuital Law (Magnetostatics) - Integral Form - *CC BY-SA 4.0*
- 7.5: Magnetic Field of an Infinitely-Long Straight Current-Bearing Wire - *CC BY-SA 4.0*
- 7.6: Magnetic Field Inside a Straight Coil - *CC BY-SA 4.0*
- 7.7: Magnetic Field of a Toroidal Coil - *CC BY-SA 4.0*
- 7.8: Magnetic Field of an Infinite Current Sheet - *CC BY-SA 4.0*
- 7.9: Ampere's Law (Magnetostatics) - Differential Form - *CC BY-SA 4.0*
- 7.10: Boundary Conditions on the Magnetic Flux Density (B) - *CC BY-SA 4.0*
- 7.11: Boundary Conditions on the Magnetic Field Intensity (H) - *CC BY-SA 4.0*
- 7.12: Inductance - *CC BY-SA 4.0*
- 7.13: Inductance of a Straight Coil - *CC BY-SA 4.0*
- 7.14: Inductance of a Coaxial Structure - *CC BY-SA 4.0*
- 7.15: Magnetic Energy - *CC BY-SA 4.0*
- 7.16: Magnetic Materials - *CC BY-SA 4.0*

8: Time-Varying Fields - *CC BY-SA 4.0*

- 8.1: Comparison of Static and Time-Varying Electromagnetics - *CC BY-SA 4.0*
- 8.2: Electromagnetic Induction - *CC BY-SA 4.0*
- 8.3: Faraday's Law - *CC BY-SA 4.0*
- 8.4: Induction in a Motionless Loop - *CC BY-SA 4.0*
- 8.5: Transformers - Principle of Operation - *CC BY-SA 4.0*
- 8.6: Transformers as Two-Port Devices - *CC BY-SA 4.0*
- 8.7: The Electric Generator - *CC BY-SA 4.0*
- 8.8: The Maxwell-Faraday Equation - *CC BY-SA 4.0*
- 8.9: Displacement Current and Ampere's Law - *CC BY-SA 4.0*

9: Plane Waves in Loseless Media - *CC BY-SA 4.0*

- 9.1: Maxwell's Equations in Differential Phasor Form - *CC BY-SA 4.0*
- 9.2: Wave Equations for Source-Free and Lossless Regions - *CC BY-SA 4.0*
- 9.3: Types of Waves - *CC BY-SA 4.0*
- 9.4: Uniform Plane Waves - Derivation - *CC BY-SA 4.0*
- 9.5: Uniform Plane Waves - Characteristics - *CC BY-SA 4.0*
- 9.6: Wave Polarization - *CC BY-SA 4.0*
- 9.7: Wave Power in a Lossless Medium - *CC BY-SA 4.0*

- 10: Appendices - *CC BY-SA 4.0*
  - 10.1: Permittivity of Some Common Materials - *CC BY-SA 4.0*
  - 10.2: Permeability of Some Common Materials - *CC BY-SA 4.0*
  - 10.3: Conductivity of Some Common Materials - *CC BY-SA 4.0*
  - 10.4: Mathematical Formulas - Trigonometry - *CC BY-SA 4.0*
  - 10.5: Mathematical Formulas - Vector Operators - *CC BY-SA 4.0*
  - 10.6: Mathematical Formulas - Vector Identities - *CC BY-SA 4.0*
  - 10.7: Physical Constants - *CC BY-SA 4.0*
- Back Matter - *Undeclared*
  - Index - *Undeclared*
  - Glossary - *Undeclared*
  - Detailed Licensing - *Undeclared*