plots, colorized: $\frac{1}{1-e^{-s\pi}}$, 306
poles, 286
poles, fractional $\mathbb{F}$, 286
positive definite operator, 326
positive-real, 326
postulates, 323
power, 326
Poynting vector, 268
prime, 309
propagation function, 248
propagation function, conical horn, 250
propagation function, exponential horn, 250
propagation function, uniform horn, 248
propagation function, with losses, 311, 315

quasistatics, 311, 327, 328
quasistatics: P10, 323

Ramo, Simon, 277
real: P4, 323
reciprocal system: eigenanalysis, 298
reciprocity, 298, 327
reciprocity: P6, 323
reversible system, eigenanalysis, 298
reversible systems, 327
reversible: P7, 323
Riemann zeta function, 301
roots of Bessel functions, 241
Rydberg atom, 240, 339

sampling theorem, 252
scalar product, 262
separation of variables, 251
set theory symbols defined, 282
sieve of Eratosthenes, 303
solenoidal, see incompressible, 261
solenoidal, see irrotational, 261
Sommerfeld, Arnold, 277
spatial invariant, 327
spatial invariant: P8, 323
speed of light, 265
speed of light, in vacuum c_o, 271
speed of sound, 313
speed of sound with losses, 311
step function, 305
Stokes theorem, 266
Stokes, George, 277
Sturm-Liouville, 243, 251
symmetry, conjugate, 326
symmetry, matrix, 329
symmetry, periodic, 329

Taylor series, 256
tesla, 265
Thévenin parameters, 266, 267, 298, 324
thermodynamics, speed of sound, 311
time-invariant, 326
time-invariant: P5, 323
train-mission-line problem, 252
transmission line, 241
transmission line, generalization, 240
transmission matrix, 248, 323, 324, 329
triangle inequality, 287

Units: $\boldsymbol{e}, \boldsymbol{d}, \boldsymbol{h}, \boldsymbol{b}$, 267

VC-1, 244
VC-2, 272
vector identity: $\nabla \times \nabla \times$, 262
vector identity: $\nabla \times \nabla$, 262
vector identity: $\nabla \cdot \nabla \times$, 262
vector identity: **gOd**, 264
vector identity: Laplacian, 262
vector operator mnemonics, definitions, 263
vector product, 262
vector space, 262
vector wedge-product, 294
visco-thermal loss, 241, 311, 315
viscosity of air, 313

wave equation, 240
wave filters, 312
webers, 265
Webster horn equation, derivation, 331
Webster horn equation, see *horn equation*, 240
Webster Laplacian, 240

zeros, 286
zeta function poles, 305
zeta function RoC, 305
zeta function, product formula, 304
zeta function: definition, 303
Zhang, Yiming, 346

Chapter 4

Stream 3A: Scalar Calculus

Stream 3 is ∞, a concept that typically means unbounded (immeasurably large), but in the case of calculus, ∞ means infinitesimal (immeasurably small), since taking a limit requires small numbers. Taking a limit means you may never reach the target, a concept that the Greeks called *Zeno's paradox* (Stillwell, 2010, p. 76).

When we speak of the class of *ordinary* (versus *vector*) differential equations, the term *scalar* is preferable, since the term ordinary is vague, if not meaningless. For scalar calculus, a special subset of fundamental theorems about integration are summarized in Table 4.1 (p. 179), starting with Leibniz's theorem.

Table 4.1: *The fundamental theorems of integral calculus, each of which deals with integration. At least two main theorems relate to scalar calculus, and three more to vector calculus.*

Name	Mapping	p.	Description
Leibniz (FTC)	$\mathbb{R}^1 \rightarrow \mathbb{R}^0$	179	Area under a real curve
Cauchy (FTCC)	$\mathbb{C}^1 \rightarrow \mathbb{R}^0$	179	Area under a complex curve
Cauchy's theorem	$\mathbb{C}^1 \rightarrow \mathbb{C}^0$	201	Close integral over analytic region is zero
Cauchy's integral formula	$\mathbb{C}^1 \rightarrow \mathbb{C}^0$	201	The fundamental theorem of complex calculus
Residue theorem	$\mathbb{C}^1 \rightarrow \mathbb{C}^0$	201	Residue integration
Helmholtz' decomposition	Sec. 5.8.3	270	A Cornerstone of Mathematical Physics

Following our discussion of the integral theorems on scalar calculus are those on vector calculus, without which there can be no understanding of Maxwell's equations. Of these, the fundamental theorem of vector calculus (also known as Helmholtz' decomposition), Gauss's law, and Stokes's theorem form the three cornerstones of modern vector field analysis. These theorems allow us to connect the differential (point) and macroscopic (integral) relationships. For example, we can write Maxwell's equations either as vector differential equations, as shown by Heaviside (along with Gibbs and Hertz), or in integral form. It is helpful to place these

two forms side by side to fully appreciate their significance. To understand the differential (microscopic) view, one must fully understand the integral (macroscopic) view (see Figs. 5.5 and 5.6 on pp. 257 and 258).

4.1 The beginning of modern mathematics

As shown in Fig. 1.2 (p. 13), mathematics as we know it today began in the 16th to 18th centuries, arguably starting with Galileo, Descartes, Fermat, Newton, the Bernoulli family, and most important Euler. Galileo was formidable because of his fame, fortune, and "successful" stance against the powerful Catholic establishment. His creativity in scientific circles was certainly well known due to his many skills and accomplishments. Descartes and Fermat were at the forefront of merging algebra and geometry. While Fermat kept meticulous notebooks, he did not publish and tended to be secretive. Thus Descartes's contributions were more widely acknowledged, though not necessarily deeper.

Regarding the development of calculus, much was yet to be developed by Newton and Leibniz using term-by-term integration of functions based on Taylor series representation. This was a powerful technique but, as stated earlier, incomplete because the Taylor series can represent only single-valued functions within the RoC. More important, Newton (and others) failed to recognize (i.e., rejected) the powerful generalization to complex–analytic functions. The first major breakthrough was Newton's publication of *Principia* (1687), and the second was by Riemann (1851), advised by Gauss but possibly more influenced by Cauchy.

Following both Galileo's and Newton's lead, the secretive and introverted behavior of the typical mathematician dramatically changed with the Bernoulli family (Fig. 3.1, p. 74). The oldest brother Jacob taught his much younger brother Johann, who then taught his son Daniel. But Johann's star pupil was Leonhard Euler. Euler first mastered all the tools and then published with a prolifically previously unknown.

Euler and the circular functions: Euler's first major task was to understand the family of analytic circular functions ($e^x, \sin(x), \cos(x)$, and $\log(x)$), a task begun by the Bernoulli family. Euler sought relationships among these many functions, some of which may not be thought of as being related, such as the log and sin functions. The connection that may "easily" be made is through their complex Taylor series representation (Eq. 3.2.9, p. 95). By the manipulation of the analytic series representations, the relationship between e^x and $\sin(x)$ and $\cos(x)$ was precisely captured with the equation

$$e^{\jmath\omega} = \cos(\omega) + \jmath\sin(\omega) \tag{4.1.1}$$

and its analytic inverse (Greenberg, 1988, p. 1135)

$$\tan^{-1}(z) = \frac{1}{2\jmath}\ln\left(\frac{1\jmath - z}{1\jmath + z}\right) = \frac{\jmath}{2}\ln\left(\frac{1 - z\jmath}{1 + z\jmath}\right). \tag{4.1.2}$$

Exercise #1

Starting from Eq. 4.1.1, derive Eq. 4.1.2.

Solution: We let $z(\omega) = \tan\omega$; then

$$z(\omega) = \tan(\omega) = \frac{\sin\omega}{\cos\omega} = -\jmath\frac{e^{\omega\jmath} - e^{-\omega\jmath}}{e^{\omega\jmath} + e^{-\omega\jmath}} = -\jmath\frac{e^{2\omega\jmath} - 1}{e^{2\omega\jmath} + 1}. \tag{4.1.3}$$

Euler's equation, 239
Euler's functional zeta equation, 307
evanescent modes, 251

Feshbach, Herman, 277
Feynman, Richard, 277
freight-train, 252
freight-train problem, 252
Fry, Thorton Carle, 277
FTVC, 261
fundamental theorem of arithmetic, 303
fundamental theorem: vector calculus, 259–261, 264

Gamma function, 307
Gauss's law, 257, 266
Gauss, CFA vs. GCD, 348
Gauss, GCD vs. CFA, 348
gOd, 264
gradient $\nabla()$, 262
greatest common divisor, 345
Green, 277
gyrator, 323

Hankel function, 247
Helmholtz, 260
Helmholtz portrait, 261
Helmholtz's decomposition theorem, 259–261
Helmholtz's decomposition: see FTVC , 259
Helmholtz, Hermann von, 277, 311
Hilbert space, see vector space, 241
horn equation, 240–242, 247, 251, 256
horn equation, derivation, 331
horn, conical, 250
horn, exponential, 250
horn, uniform, 247
Hunt, Fredrick, 277
Huygens, 240
Huygens principle, 254
Hydrogen spectrum, 339

impedance, light, 265
impulse response, 286
incompressible, 260
integral definition of $\nabla\times\boldsymbol{h}(\boldsymbol{x})$, 258
integral definition of $\nabla\cdot\boldsymbol{d}(\boldsymbol{x})$, 257
integral definition of $\nabla\phi(\boldsymbol{x})$, 255
integration by parts, 242
integration, half, 302
irrotational, 260, 261

Kirchhoff's portrait, 261
Kirchhoff, Gustav, 311

Lamb, Willis, 277
Laplace transforms: Table III, 302
Laplacian ∇^2, 240, 262
Laplacian, vector: $\boldsymbol{\nabla}^2$, 264
Lec20: zeta function, Euler, 309
Lec20: zeta function, Riemann, 309
light, speed of, 265
Lighthill, James, 277
linearity: P2, 323
Lord Rayleigh, 277
Lorentz force, 269

Maxwell's equation: matrix form, 266
Maxwell's equations, 264, 265, 267
Maxwell's equations, potential solutions, 270
Maxwell's equations: d'Alembert solution, 268
Maxwell's equations: variable names and units, 265
Maxwell's matrix wave equation, 268
ME: Maxwell's equations, 266
modal-analysis, 241
modulo, 285
monic polynomial, 285
Morse, Phillip, 277

network postulates, 323
networks, algebraic, 324
Newton, 240
Norton parameters, 267
NOVA, fusion experiment, 269
Number theory, 349, 351
Nyquist sampling, 252
Nyquist sampling of lumped-element models, 252
Nyquist, Harry, 252

Olson, Harry, 277
order, derivatives, 285

passive, 325
passive: P3, 323
Periodic ↔ discrete, 329
periodic ↔ discrete: P11, 323
Physical examples, 239, 247, 248
Pipes, L.A., 277

Index

$\nabla(\phi(\boldsymbol{x}))$, 255
$\mathbb{N}$: see counting numbers, 281
$\mathcal{Y}_r$: see characteristic admittance, 248
$\mathcal{Z}_r$: see characteristic impedance, 248
δ_k Dirac pulse, 305
$\boldsymbol{e}, \boldsymbol{d}, \boldsymbol{h}, \boldsymbol{b}$: Definitions, 265
$\boldsymbol{e}, \boldsymbol{d}, \boldsymbol{h}, \boldsymbol{b}$: Units, 267
$\ln \zeta(s)$, 303
$\rho_o = 1.2$ [kg/m^3] (air density), 242
$\zeta(s)$ function: See zeta function, 309
$c_o = 343$ [m/s] (air sound speed), 240
u_k step function, 305
gOd: ∇^2 (aka Little god), 264
`ceil(x)`, 284
`fix(x)`, 284
`floor(x)`, 284
`round(x)`, 284
`roundb(x)`, 284
DoG, 264
gOd, 264

acoustic properties of air (Mason, 1928), 313
adiabatic expansion, 311
adiabatic expansion, speed of sound, 239
analytic continuation and the $\mathcal{LT}^{-1}$, 307
Assignments: VC-1, 244
Assignments: VC-2, 272

Beranek, Leo, 277
Bessel function roots, 241
bilinear transformation, 286, 325
Boas, Ralph, 277
Bode, Hendrik, 277
branch cut, 269
branch point, 269
Brillouin, Léon, 242, 252, 277
Bull-Dog, 263

Campbell's electrical wave filters, 312
Campbell, George Ashley, 242, 252, 333
causality: P1, 323
CFA matrix, 346
characteristic admittance, 248
Characteristic impedance, in vacuum, r_o, 271
classification: poles and zeros, 307
classification: vector fields, 260
CoC, 263, 264
CoG, 263, 264
compressible, 260
conservative field, 260
curl $\nabla \times ()$, 262

d'Alembert, 248, 253, 254, 268
degree, fractional $\mathbb{F}$, 286
degree, irrational $\mathbb{I}$, 286
degree, polynomials, 285
degrees of freedom, 262
derivative, half, 302
determinant, 287, 289
Deterministic, 328
Deterministic: P9, 323
diffusion equation, 252
Dirac pulse, 305
divergence $\nabla \cdot ()$, 262
DoC, 263, 264
DoF: degrees of freedom, 262
DoG, 263, 264
driving-point impedance, 324

eigenanalysis, 293
eigenanalysis: 2x2 matrix, 297
eigenanalysis: reciprocal system, 298
eigenspace, see vector space, 241
eigenvalues, 293
eigenvalues, Pell's equation, 294
eigenvectors, 294
Einstein, Albert, 311
emf, 266
energy, 325, 326
Eratosthenes, 304
error function, 302
Euclid's formula, 339
Euler, 309

Solving for $e^{-2\omega\jmath}$, we get

$$e^{-2\omega\jmath} = \frac{1 - z\jmath}{1 + z\jmath}. \tag{4.1.4}$$

Taking $\ln()$ of both sides and using the definition of $z(\omega)$ give Eq. 4.1.2:

$$\omega = \tan^{-1}(z) = \frac{\jmath}{2} \ln \frac{1 - z\jmath}{1 + z\jmath}.$$

These equations are the basis of transmission lines (TL), and used throughout mathematical physics. They are related to the z (bilinear) transform, and Newton's method for finding roots of complex–analytic functions (Allen, 2024). Here $z(\omega)$ of Eq. 4.1.3 is the TL's input impedance and Eq. 4.1.4 is called the *reflectance*. ∎

Although many high school students memorize Euler's relationship, it seems unlikely that they appreciate the utility of the complex–analytic function. This possibility could change with a trivial teaching innovation (Use *OverLeaf* for written expression).

History of complex–analytic functions: Newton (ca.1650) famously ignored imaginary numbers and called them imaginary in a disparaging (pejorative) way. Given Newton's prominence, his view must have keenly attenuated interest in complex algebra, even though it had been described by Bombelli in 1526, likely based on his serendipitous finding of Diophantus's book *Arithmetic* in the Vatican library.

According to mathematical evidence, Euler did not appreciate the role of complex–analytic functions. They were first fully appreciated well after his death in 1785, by Augustin-Louis Cauchy (1789–1857), and then extended by Riemann in 1851.

Euler derived his relationships using real power series (i.e., real analytic functions). While Euler was fluent with $\jmath = \sqrt{-1}$, he did not consider functions to be complex–analytic. That concept was first explored by Cauchy almost a century later. The missing link to the concept of complex–analytic functions is the definition of the derivative with respect to the complex argument

$$F'(s) = \frac{dF(s)}{ds}, \tag{4.1.5}$$

where $s = \sigma + \omega\jmath$, without which the complex–analytic Taylor coefficients are not defined.

4.2 Fundamental Theorems of Scalar Calculus

History of scalar calculus: It some sense the story of calculus begins with the fundamental theorem of calculus (Table 4.1), also known generically as *Leibniz's formula*. The simplest integral is the length of a line $L = \int_0^L dx$. If we label a point on a line as $x = 0$ and wish to measure the distance to any other point x, we form the line integral between the two points. If the line is straight, this integral is the Euclidean length given by the difference between the two ends (Eq. 3.5.3, p. 120).

If $F(\chi) \in \mathbb{R}$ describes a height above the line $\chi \in \mathbb{R}$, then

$$f(x) - f(0) = \int_{x=0}^{x} F(\chi) d\chi \tag{4.2.1}$$

may be viewed as the *antiderivative* of $F(\chi)$. Here χ is a dummy variable of integration. Thus the area under $F(\chi)$ depends on only the difference in the area evaluated at the end points.

This property of the area as an integral over an interval, depending on only the end points, has important consequences in physics in terms of conservation of energy, allowing for important generalizations. For example, as long as $\chi \in \mathbb{R}$, we can let $F(\chi) \in \mathbb{C}$ with no loss of generality, due to the linear Postulate P1 of the integral (see p. 162).

4.2.1 The fundamental theorem of real calculus

If $f(x)$ is analytic (Eq. 3.2.8, p. 94), then

$$F(x) = \frac{d}{dx} f(x) \tag{4.2.2}$$

is an exact real differential. It follows that $F(x)$ is analytic. This is known as the *fundamental theorem of (real) calculus* (FTC). Thus Eq. 4.2.2 may be viewed as an exact real differential. This is easily shown by evaluating

$$\frac{d}{dx} f(x) = \lim_{\delta \to 0} \frac{f(x+\delta) - f(x)}{\delta} = F(x)$$

starting from the antiderivative, Eq. 4.2.1. If $f(x)$ is not analytic then the limit may not exist, so this is a necessary condition.

There are many important variations on this very basic theorem (see Table 4.1). For example, the limits could depend on time. Also when we take Fourier transforms, the integrand depends on both time $t \in \mathbb{R}$ and frequency $\omega \in \mathbb{R}$ via a complex exponential "kernel" function $e^{\pm \jmath \omega t} \in \mathbb{C}$, which is analytic in both t and ω.

4.2.2 The fundamental theorem of complex calculus (Sec. 4.2.1)

The fundamental theorem of complex calculus (FTCC) states (Greenberg, 1988, p. 1197) that for any complex–analytic function $F(s) \in \mathbb{C}$ with $s = \sigma + \omega\jmath \in \mathbb{C}$,

$$f(s) - f(s_o) = \int_{s_o}^{s} F(\zeta) d\zeta. \tag{4.2.3}$$

Equations 4.2.1 and 4.2.3 differ because the path of the integral is complex. Thus the line integral is over $s \in \mathbb{C}$ rather than a real integral over $\chi \in \mathbb{R}$. The FTCC states that the integral depends on only the end points, since

$$F(s) = \frac{d}{ds} f(s). \tag{4.2.4}$$

Comparing exact differentials, Eq. 4.1.5 (FTCC) and Eq. 4.2.2 (FTC), we see that $f(s) \in \mathbb{C}$ must be complex–analytic and have a Taylor series in powers in $s \in \mathbb{C}$. It follows that $F(s)$ is also complex–analytic.

Complex–analytic functions: The definition of a *complex–analytic function* $F(s)$ of $s \in \mathbb{C}$ is that the function may be expanded in a Taylor series (Eq. 3.2.20, p. 102) about an expansion point $s_o \in \mathbb{C}$. This definition follows the same logic as the FTC. Thus we need a definition for the coefficients $c_n \in \mathbb{C}$, which most naturally follow from Taylor's formula

$$c_n = \frac{1}{n!} \frac{d^n}{ds^n} F(s) \bigg|_{s=s_o}. \tag{4.2.5}$$

Winchester, S. (2009). *The man who loved China: The fantastic story of the eccentric scientist who unlocked the mysteries of the middle kingdom*. Harper Collins.

Youla, D. (1964). Analysis and synthesis of arbitrarily terminated lossless nonuniform lines. *IEEE Transactions on Circuit Theory*, 11(3):363–372.

Sommerfeld, A. (1949). *Partial Differential Equations in Physics, Lectures on Theoretical Physics, Vol. I*. Academic Press INC., New York.

Sommerfeld, A. (1952). *Electrodynamics, Lectures on Theoretical Physics, Vol. III*. Academic Press INC., New York.

Sondhi, M. and Gopinath, B. (1971). Determination of vocal-tract shape from impulse responses at the lips. *Journal of the Acoustical Society of America*, 49(6:2):1867–73.

Stewart, G. W. (1922). Acoustic wave filters. *Physical Review*, 20(6):528.

Stewart, G. W. and Lindsay, R. B. (1930). *Acoustics: A text on theory and applications*. D. Van Nostrand Company, Incorporated, New Jersey, NJ.

Stewart, J. (2012). *Essential Calculus: Early Transcendentals*. Cengage Learning, Boston, MA.
https://books.google.com/books?id=AcQJAAAAQBAJ.

Stillwell, J. (2002). *Mathematics and Its History; Undergraduate Texts in Mathematics; 2d edition*. Springer, New York.

Stillwell, J. (2010). *Mathematics and Its History; Undergraduate Texts in Mathematics; 3d edition*. Springer, New York.

Strang, G., Strang, G., Strang, G., and Strang, G. (1993). *Introduction to linear algebra*, volume 3. Wellesley-Cambridge Press Wellesley, MA.

Tisza, L. (1966). *Generalized Thermodynamics*, volume 1. MIT press Cambridge.

Tolman, R. C. (1948). Consideration of the Gibbs theory of surface tension. *The journal of chemical physics*, 16(8):758–774.

Tran-Cong, T. (1993). *On Helmholtz's decomposition theorem and Poisson's equaton with an infinite domain*, volume LI. Quarterly of Applied Mathematics.

Van Valkenburg, M. (1960). *Introduction to Modern Network Synthesis*. John Wiley and Sons, Inc., New York.

Van Valkenburg, M. (1964a). *Network Analysis second edition*. Prentice-Hall, Englewood Cliffs, N.J.

Van Valkenburg, M. E. (1964b). *Modern Network Synthesis*. John Wiley & Sons, Inc., New York, NY.

Vanderkooy, J. (1989). A model of loudspeaker driver impedance incorporating eddy currents in the pole structure. *JAES*, 37(3):119–128.

Webster, A. G. (1919). Acoustical impedance, and the theory of horns and of the phonograph. *Proceedings National Academe of Science*, 5:275–282.

White, M. (1999). *Isaac Newton: The Last Sorcerer*. Basic Books (AZ).

Wikipedia: Speed of Light (2019). Wikipedia citation. *Wikipedia*. How speed of light was first measured: https://en.wikipedia.org/wiki/Speed_of_light.

The requirement that $F(s)$ have a Taylor series naturally follows by taking derivatives with respect to s at s_o. The problem is that both integration and differentiation of functions of complex Laplace frequency $s = \sigma + \omega\jmath$ have not yet been defined.

Thus the question: What does it mean to take the derivative of a function $F(s) \in \mathbb{C}$, $s = \sigma + \omega\jmath \in \mathbb{C}$, with respect to s, where s defines a plane rather than a real line? We learned how to form the derivative on the real line. Can the same derivative concept be extended to the complex plane?

The answer is affirmative. The question may be resolved by applying the rules of the real derivative when defining the derivative in the complex plane. However, for the complex case, there is an issue regarding direction. Given any analytic function $F(s)$, is the partial derivative with respect to σ different from the partial derivative with respect to $\omega\jmath$? For complex–analytic functions, the FTCC states that the integral is independent of the path in the s plane. Based on the chain rule, the derivative must also be independent of the direction at s_o. This directly follows from the FTCC. If the integral of a function of a complex variable is to be independent of the path, then the derivative of a function with respect to a complex variable must be independent of the direction. This follows from Taylor's formula for the coefficients of the complex–analytic formula (Eq. 4.2.5).

The Cauchy-Riemann conditions: The FTC defines the area as an integral over a real differential ($dx \in \mathbb{R}$), while the FTCC relates an integral over a complex function $F(s) \in \mathbb{C}$ along a complex interval (i.e., path) ($ds \in \mathbb{C}$). For the FTC the area under the curve depends on only the end points of the antiderivative $f(x)$. But what is the meaning of an "area" along a complex path? The Cauchy-Riemann conditions provide the answer.

4.2.3 Cauchy-Riemann conditions

For the integral of $Z(s) = R(\sigma,\omega) + \jmath X(\sigma,\omega)$ to be independent of the path, the derivative of $Z(s)$ must also be independent of the path. This requirement leads to a pair of equations known as the *Cauchy-Riemann conditions*, described next.

To define

$$\frac{d}{ds}Z(s) = \frac{d}{ds}\left[R(\sigma,\omega) + \jmath X(\sigma,\omega)\right],$$

we take partial derivatives of $Z(s)$ with respect to σ and $\jmath\omega$, and equate them:

$$\frac{\partial Z}{\partial \sigma} = \frac{\partial R}{\partial \sigma} + \jmath\frac{\partial X}{\partial \sigma} \qquad \equiv \qquad \frac{\partial Z}{\partial \jmath\omega} = \frac{\partial R}{\partial \jmath\omega} + \jmath\frac{\partial X}{\partial \jmath\omega}.$$

This says that a horizontal derivative, with respect to σ, is equivalent to a vertical derivative, with respect to $\omega\jmath$. Taking the real and imaginary parts gives the two equations

$$\frac{\partial R(\sigma,\omega)}{\partial \sigma} = \not{\jmath}\frac{\partial X(\sigma,\omega)}{\partial \omega\not{\jmath}} \quad \text{(CR-1)} \quad \text{and} \quad \frac{\partial R(\sigma,\omega)}{\partial \omega\not{\jmath}} = -\not{\jmath}\frac{\partial X(\sigma,\omega)}{\partial \sigma} \quad \text{(CR-2)}, \qquad (4.2.6)$$

known as the *Cauchy-Riemann (CR) conditions*. The $\jmath$ cancels in CR-1 but introduces a $\jmath^2 = -1$ in CR-2. They may also be written in polar coordinates ($s = re^{\theta\jmath}$) as

$$\frac{\partial R}{\partial r} = \frac{1}{r}\frac{\partial X}{\partial \theta} \qquad \text{and} \qquad \frac{\partial X}{\partial r} = -\frac{1}{r}\frac{\partial R}{\partial \theta}.$$

The FTCC (Eq. 4.2.3) follows from CR-1 and CR-2 (Eq. 4.2.6).

You may wonder what would happen if we took a derivative at 45°. To do this we only need to multiply the function by $e^{j\pi/4}$. But doing so will not change the derivative. Thus we may take the derivative in any direction by multiplying by $e^{\theta j}$, and the CR conditions will not change.

The CR conditions are necessary so that the integral of $Z(s)$, and thus its derivative, is independent of the path, expressed in terms of conditions on the real and imaginary parts of Z. This is a very strong condition on $Z(s)$, which follows assuming that $Z(s)$ may be written as a Taylor series in s:

$$Z(s) = Z_0 + Z_1 s + \frac{1}{2} Z_2 s^2 + \cdots , \tag{4.2.7}$$

where $Z_n \in \mathbb{C}$ are complex constants given by the Taylor series formula (Eq. 4.2.5). As with the real Taylor series, there is the convergence condition that $|s| < 1$, called the *radius of convergence* (RoC). This is an important generalization of the region of convergence (RoC) for real $s = x$.

Every function that may be expressed as a Taylor series in $s - s_o$ about point $s_o \in \mathbb{C}$ is said to be *complex–analytic* at s_o. This series, which is single-valued, is said to converge within a radius of convergence (RoC). This highly restrictive condition has significant physical consequences. For example, every impedance function $Z(s)$ obeys the CR conditions over large regions of the s plane, including the entire *right half s plane* (RHP) ($\sigma > 0$). This condition is summarized by the Brune condition $\Re\{Z(\sigma > 0)\} \geq 0$, or alternatively $\angle Z(s) < \angle s$

When the CR conditions are generalized to volume integrals, this is called either Gauss's Law or Green's theorem, which is used in the solution of boundary value problems in engineering and physics (Kusse and Westwig, 2010).

We may merge these equations into a pair of second-order equations by taking a second round of partials. Specifically, eliminating the real part $R(\sigma, \omega)$ of Eq. 4.2.6 gives

$$\frac{\partial^2 R(\sigma,\omega)}{\partial\sigma\partial\omega} = \frac{\partial^2 X(\sigma,\omega)}{\partial^2\omega} = -\frac{\partial^2 X(\sigma,\omega)}{\partial^2\sigma}, \quad \text{(CR-3)} \tag{4.2.8}$$

which may be written compactly as $\nabla^2 X(\sigma, \omega) = 0$. Eliminating the imaginary part gives

$$\frac{\partial^2 X(\sigma,\omega)}{\partial\omega\partial\sigma} = \frac{\partial^2 R(\sigma,\omega)}{\partial^2\sigma} = -\frac{\partial^2 R(\sigma,\omega)}{\partial^2\omega}, \quad \text{(CR-4)} \tag{4.2.9}$$

which may be written as $\nabla^2 R(\sigma, \omega) = 0$.

In summary, for a function $Z(s)$ to be complex–analytic, the derivative dZ/ds must be independent of direction (path), which requires that the real and imaginary parts of the function obey Laplace's equation; that is,

$$\nabla^2 R(\sigma,\omega) = 0 \quad \text{and} \quad \nabla^2 X(\sigma,\omega) = 0. \tag{4.2.10}$$

Equations CR-1 and CR-2 are easy to work with because they are first-order, but the intuition behind them best follows from the properties of Laplace's equation (Eq. 4.2.10). Note two facts: (1) the derivative of a complex–analytic function is independent of its direction, and (2) the real and imaginary parts of the function obey Laplace's equation. Such relationships are known as *harmonic functions*.

As we shall see in the next few sections, complex–analytic functions must be smooth, since every analytic function may be differentiated an infinite number of times (the complex Taylor series) within the RoC. The magnitude must attain its maximum and minimum on the boundary. For example, when you stretch a rubber sheet over a jagged frame, the height of the rubber sheet

Olson, H. F. (1947). *Elements of Acoustical Engineering (2^n Edition)*. D. Van Nostrand Company, 539 pages.

O'Neill, M. E. (2009). The genuine sieve of eratosthenes. *Journal of Functional Programming*, 19(1):95–106.

Palmerino, C. (1999). Infinite degrees of speed: Marin mersenne and the debate over galileo's law of free fall. *Early Science and Medicine*, 4(4):269–328. 'www.jstor/stable/4130144'.

Papasimakis, N., Raybould, T., Fedotov, V. A., Tsai, D. P., Youngs, I., and Zheludev, N. I. (2018). Pulse generation scheme for flying electromagnetic doughnuts. *Phys. Rev. B*, 97:201409.

Papoulis, A. (1962). *The Fourier Integral and its Applications*. McGraw–Hill Book Co., New York.

Papoulis, A. and Pillai, S. U. (2002). *Probability, Random Variables, and Stochastic Processes*. Tata McGraw-Hill Education.

Pierce, A. D. (1981). *Acoustics: An Introduction to its Physical Principles and Applications*. McGraw-Hill, 678 pages, New York.

Pipes, L. A. (1958). *Applied Mathematics for Engineers and Physicists*. McGraw-Hill, NYC, NY.

Ramo, S., Whinnery, J. R., and Van Duzer, T. (1965). *Fields and Waves in Communication Electronics*. John Wiley & Sons, Inc., New York.

Rasetshwane, D. M., Neely, S. T., Allen, J. B., and Shera, C. A. (2012). Reflectance of acoustic horns and solution of the inverse problem. *The Journal of the Acoustical Society of America*, 131(3):1863–1873.

Rayleigh, J. W. (1896). *Theory of Sound, Vol. I 480 pages, Vol. II 504 pages*. Dover, New York.

Riemann, B. (1859). On the number of primes less than a given magnitude. *Monatsberichte der Berliner Akademie*, pages 1–10,. Translated by David R. Wilkins-1998.

Rotman, J. J. (1996). A first course in abstract algebra.

Salmon, V. (1946). Generalized plane wave horn theory. *Journal of the Acoustical Society of America*, 17(3):199–211.

Schelkunoff, S. (1943). *Electromagnetic Waves*. Van Nostrand Co., Inc., Toronto, New York and London. 6th edition.

Schwinger, J. S. and Saxon, D. S. (1968). *Discontinuities in waveguides: notes on lectures by Julian Schwinger*. Gordon and Breach, New York.

Scott, A. (2002). *Neuroscience: A Mathematical Primer*. Springer Science & Business Media.

Shaw, E. (1970). Acoustic horns with spatially varying density or elasticity. *Journal of the Acoustical Society of America*, 50(3):830–840.

Lin, J. Y. (1995). The Needham Puzzle: Why the Industrial Revolution did not originate in China. *Economic development and cultural change*, 43(2):269–292.

Lundberg, K. H., Miller, H. R., and Trumper, R. L. (2007). Initial conditions, generalized functions, and the laplace transform: Troubles at the origin. *IEEE Control Systems Magazine*, 27(1):22–35.

Mason, W. (1927). A study of the regular combination of acoustic elements with applications to recurrent acoustic filters, tapered acoustic filters, and horns,. *Bell System Technical Journal*, 6(2):258–294.

Mason, W. (1928). The propagation characteristics of sound tubes and acoustic filters. *Physical Review*, 31:283–295.

Mawardi, O. K. (1949). Generalized Solutions of Webster's Horn Theory. *Journal of the Acoustical Society of America*, 21(4):323–30.

Maxwell, J. C. (1865). Viii. a dynamical theory of the electromagnetic field. *Philosophical transactions of the Royal Society of London*, 32(155):459–512.

Mazur, J. (2014). *Enlightening Symbols: A Short History of Mathematical Notation and Its Hidden Powers*. Princeton University Press, Princeton, NJ.

McMillan, E. M. (1946). Violation of the reciprocity theorem in linear passive electromechanical systems. *Journal of the Acoustical Society of America*, 18:344–347.

Medicus, H. (1974). *Fifty years of matter waves*. Physics Today, AIP.

Menzies, G. (2004). *1421: The Year China Discovered America*. HarperCollins, NYC.

Menzies, G. (2008). *1434: The year a Magnificent Chinese Fleet Sailed to Italy and Ignited the Renaissance*. HarperCollins, NYC.

Miles, J. W. (1948). The coupling of a cylindrical tube to a half-infinite space. *Journal of the Acoustical Society of America*, 20(5):652–664.

Miller, D. A. (1991). Huygens's wave propagation principle corrected. *Optics Letters*, 16(18):1370–2.

Montgomery, C., Dicke, R., and Purcell, E. (1948). *Principles of Microwave Circuits*. McGraw-Hill, Inc., New York.

Moralee, D. (1974, 1980, 1995). *The First Ten Years*. Cambridge University Physics Society, Cambridge, UK.

Morse, P. M. (1948). *Vibration and Sound*. McGraw-Hill, 468 pages, NYC, NY.

Morse, P. M. and Feshbach, H. (1953). *Methods of Theoretical Physics. Vol. I. II*. McGraw-Hill, Inc., New York.

Nyquist, H. (1924). Certain factors affecting telegraph speed. *Bell System Technical Journal*, 3:324.

obeys Laplace's equation. Nowhere can the height of the sheet rise above or below its value at the boundary.

Harmonic functions define *conservative fields*, which means that energy (like a volume or area) is conserved. The work done in moving a mass from a to b in such a field is conserved. If you return the mass from b back to a, the stored energy is retrieved, thus zero net work is done.

4.3 Problems DE-1

4.3.1 Topics of this homework:

Complex numbers and functions (ordering and algebra), complex power series, fundamental theorem of calculus (real and complex); Cauchy-Riemann conditions, multivalued functions (branch cuts and Riemann sheets)

4.3.2 Complex Power Series

Problem # *1: In each case derive (e.g., using Taylor's formula) the power series of* $w(s)$ *about* $s = 0$ *and give the RoC of your series. If the power series doesn't exist, state why!* Hint: In some cases, you can derive the series by relating the function to another function for which you already know the power series at $s = 0$.

– *1.1:* $1/(1 - s)$

– *1.2:* $1/(1 - s^2)$

– *1.3:* $1/(1 + s^2)$.

– *1.4:* $1/s$

– *1.5:* $1/(1 - |s|^2)$

Problem # *2: Consider the function* $w(s) = 1/s$

– *2.1: Expand this function as a power series about* $s = 1$. Hint: Let $1/s = 1/(1 - 1 + s) = 1/(1 - (1 - s))$.

– *2.2: What is the RoC?*

– *2.3: Expand* $w(s) = 1/s$ *as a power series in* $s^{-1} = 1/s$ *about* $s^{-1} = 1$.

– *2.4: What is the RoC?*

Hill, J. M. (2007). Laplace transforms and the Riemann zeta function. *Integral Transforms and Special Functions*, 18(3):193–205. jhill@uow.edu.au.

Horn, R. and Johnson, C. (1988). *Matrix Analysis*. Cambridge Univ. Press, Cambridge, England.

Hunt, F. V. (1952). *Electroacoustics, The Analysis of Transduction, and Its Historical Background*. The Acoustical Society of America, 260 pages, Woodbury, NY 11797.

Jaynes, E. (1991). Scattering of light by free electrons as a test of quantum theory. In *The Electron*, pages 1–20. Springer.

Johnson, D. H. (2003). Origins of the equivalent circuit concept: the voltage-source equivalent. *Proceedings of the IEEE*, 91(4):636–640.
https://www.ece.rice.edu/~dhj/paper1.pdf.

Johnson, F. S., Cragin, B. L., and Hodges, R. R. (1994). Electromagnetic momentum density and the poynting vector in static fields. *American journal of physics*, 62(1):33–41.

Karal, F. C. (1953). The analogous acoustical impedance for discontinuities and constrictions of circular cross section. *Journal of the Acoustical Society of America*, 25(2):327–334.

Kelley, C. T. (2003). *Solving Nonlinear Equations with Newton's Method*, volume 1. SIAM.

Kennelly, A. E. (1893). Impedance. *Transactions of the American Institute of Electrical Engineers*, 10:172–232.

Kim, N. and Allen, J. B. (2013). Two-port network analysis and modeling of a balanced armature receiver. *Hearing Research*, 301:156–167.

Kim, N., Yoon, Y.-J., and Allen, J. B. (2016). Generalized metamaterials: Definitions and taxonomy. *Journal of the Acoustical Society of America*, 139:3412–3418.

Kirchhoff, G. (1868). On the influence of heat conduction in a gas on sound propagation. *Ann. Phys. Chem.*, 134:177–193.

Kirchhoff, G. (1974). On the influence of heat conduction in a gas on sound propagation (*English*). In Lindsay, R. B., editor, *Physical Acoustics, Benchmark Papers in Acoustics Volume 4*, pages 7–19. Dowden, Hutchinson & Ross, Inc., Stroudsburg, Pennsylvania.

Kleiner, M. (2013). *Electroacoustics*. CRC Press, Taylor & Francis Group, Boca Raton, FL, USA.
https://www.crcpress.com/authors/i546-mendel-kleiner.

Kuhn, T. (1978). *Black-Body Theory and the Quantum Discontinuity, 1894-1912; (371 pages)*. Oxford University Press (1978) & University of Chicago Press (1987), Oxford, Oxfordshire; Chicago, IL.

Kusse, B. R. and Westwig, E. A. (2010). *Mathematical physics: applied mathematics for scientists and engineers*. John Wiley & Sons.

Lamb, H. (1932). *Hydrodynamics*. Dover Publications, New York.

Lighthill, S. M. J. (1978). *Waves in fluids*. Cambridge University Press, England.

Goldstein, L. (1973). A history of the prime number theorem. *American Mathematical Monthly*, 80(6):599–615.

Graham, R. L., Knuth, D. E., and Patashnik, O. (1994). *Concrete Mathematics: A Foundation for Computer Science, 2n ed.* Addison-Wesley Longman, Boston, MA, 2nd edition.

Graham, R. L., Knuth, D. E., Patashnik, O., and Liu, S. (1989). Concrete mathematics: a foundation for computer science. *Computers in Physics*, 3(5):106–107.

Greenberg, M. D. (1988). *Advanced Engineering Mathematics*. Prentice Hall, Upper Saddle River, NJ.

Grosswiler, P. (2004). Dispelling the alphabet effect. *Canadian Journal of Communication*, 29(2).

Hahn, W. (1941). A new method for the calculation of cavity resonators. *Journal of Applied Physics*, 12:62–68.

Hamming, R. W. (2004). *Methods of Mathematics Applied to Calculus, Probability, and Statistics*. Courier Corporation, North Chelmsford, MA.

Hanna, C. R. and Slepian, J. (1924). The function and design of horns for loud speakers. *Transactions of the American Institute of Electrical Engineers*, XLIII:393–411.

Hardey, G. and Wright, E. (1938). *An Introduction to the Theory of Numbers*. Oxford Science Publications.

Heaviside, O. (1892). On resistance and conductance operators, and their derivatives inductance and permittance, especially in connection with electric and magnetic energy. *Electrical Papers*, 2:355–74.

Helmholtz, H. (1858). Ueber integrale der hydrodynamichen gleichungen welche den wirelbewegungen entsprechen (on integrals of the hydrodynamic equations that correspond to vortex motions). *Journal fur die reine und angewandte Mathematik*, 55:25–55.

Helmholtz, H. (1863a). *On the Sensations of Tone*. 1954, Dover, New York, 300 pages.

Helmholtz, H. (1863b). Ueber den einfluss der reibung in der luft auf die schallbewegung (on the influence of friction in the air on the sound movement). *Verhandlungen des naturhistorisch-medicinischen Vereins zu Heildelberg*, III(17):16–20. Sitzung vom 27.

Helmholtz, H. (1978). On integrals of the hydrodynamic equations that correspond to vortex motions. *International Journal of Fusion Energy*, 1(3-4):41–68. English translation of Helmholtz's 1858 paper, translated by Uwe Parpart (1978). The image of Helmholtz is on page 41.
`https://www.wlym.com/archive/fusion/fusionarchive_ijfe.html`.

Heras, R. (2016). The Helmholtz theorem and retarded fields. *European Journal of Physics*, 37(6):065204. `https://stacks.iop.org/0143-0807/37/i=6/a=065204`.

Hestenes, D. (2003). Spacetime physics with geometric algebra. *American Journal of Physics*, 71(7):691–714.

– 2.5: What is the residue of the pole?

Problem # *3: Consider the function* $w(s) = 1/(2-s)$

– 3.1: Expand $w(s)$ *as a power series in* $s^{-1} = 1/s$*. State the RoC as a condition on* $|s^{-1}|$. Hint: Multiply top and bottom by s^{-1}.

– 3.2: Find the inverse function $s(w)$*. Where are the poles and zeros of* $s(w)$*, and where is it analytic?*

Problem # *4:Summing the series*
The Taylor series of functions have more than one region of convergence.

– 4.1: Given some function $f(x)$*, if* $a = 0.1$*, what is the value of*

$$f(a) = 1 + a + a^2 + a^3 + \cdots?$$

Show your work.

– 4.2: Let $a = 10$*. What is the value of*

$$f(a) = 1 + a + a^2 + a^3 + \cdots?$$

4.3.3 Cauchy-Riemann Equations

Problem # *5: For this problem* $\jmath = \sqrt{-1}$*,* $s = \sigma + \omega\jmath$*, and* $F(s) = u(\sigma,\omega) + \jmath v(\sigma,\omega)$*. According to the fundamental theorem of complex calculus (FTCC), the integration of a complex–analytic function is independent of the path. It follows that the derivative of* $F(s)$ *is defined as*

$$\frac{dF}{ds} = \frac{d}{ds}\left[u(\sigma,\omega) + \jmath v(\sigma,\omega)\right]. \tag{DE-1.1}$$

If the integral is independent of the path, then the derivative must also be independent of the direction:

$$\frac{dF}{ds} = \frac{\partial F}{\partial \sigma} = \frac{\partial F}{\partial \jmath\omega}. \tag{DE-1.2}$$

The Cauchy-Riemann (CR) conditions

$$\frac{\partial u(\sigma,\omega)}{\partial \sigma} = \frac{\partial v(\sigma,\omega)}{\partial \omega} \quad \text{and} \quad \frac{\partial u(\sigma,\omega)}{\partial \omega} = -\frac{\partial v(\sigma,\omega)}{\partial \sigma}$$

may be used to show where Equation DE-1.2 holds.

– 5.1: Assuming Equation DE-1.2 is true, use it to derive the CR equations.

– 5.2: Merge the CR equations to show that u and v obey Laplace's equations.

$$\nabla^2 u(\sigma,\omega) = 0 \quad \text{and} \quad \nabla^2 v(\sigma,\omega) = 0.$$

– 5.3: What can you conclude?

Problem # *6: Apply the CR equations to the following functions. State for which values of $s = \sigma + i\omega$ the CR conditions do or do not hold (e.g., where the function $F(s)$ is or is not analytic).* Hint: Review where CR-1 and CR-2 hold.

– 6.1: $F(s) = e^s$

– 6.2: $F(s) = 1/s$

4.3.4 Branch cuts and Riemann sheets

Problem # *7: Consider the function $w^2(z) = z$. This function can also be written as $w_\pm(z) = \sqrt{z_\pm}$. Assume $z = re^{\phi j}$ and $w(z) = \rho e^{\theta j} = \sqrt{r}e^{\phi j/2}$.*

– 7.1: How many Riemann sheets do you need in the domain *(z) and the* range *(w) to fully represent this function as single-valued?*

– 7.2: Indicate (e.g., using a sketch) how the sheet(s) in the domain map to the sheet(s) in the range.

– 7.3: Use `zviz.m` *to plot the positive and negative square roots $+\sqrt{z}$ and $-\sqrt{z}$. Describe what you see.*

– 7.4: Where does `zviz.m` *place the branch cut for this function?*

– 7.5: Must the branch cut necessarily be in this location?

Problem # *8: Consider the function $w(z) = \log(z)$. As in Problem 7, let $z = re^{\phi j}$ and $w(z) = \rho e^{\theta j}$.*

– 8.1: Describe with a sketch and then discuss the branch cut *for $f(z)$.*

Crandall, I. B. (1926). *Theory of Vibrating Systems and Sound*. Van Nostrand, New York, 272 pages.

D'Angelo, J. P. (2017). *Linear and Complex Analysis for Applications*. CRC Press, Taylor & Francis Group, Boca Raton, FL.

Einstein, A. (1905). Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8):549–60.

Eisner, E. (1967). Complete solutions of the Webster horn equation. *Journal of the Acoustical Society of America*, 41(4B):1126–46.

Fagen, M., editor (1975). *A History of Engineering and Science in the Bell System – The Early Years (1875–1925)*. Bell Telephone Laboratories, Murray Hill, NJ, Cambridge, MA.

Fermi, E. (1936). *Thermodynamics*. Dover, New York, NY.

Fettweis, A. (1986). Wave digital filters: theory and practice. *Proc. IEEE*, 74:270–327.

Feynman, R. (1968). *The Character of Physical Law*. MIT press, Cambridge, MA.

Feynman, R. (1970a). *Feynman Lectures on Physics*. Addison-Wesley, Boston, MA.

Feynman, R. (1970b). *Feynman Lectures on Physics*, volume II. Addison-Wesley, Boston, MA.

Feynman, R. (1970c). *Feynman Lectures on Physics*, volume III. Addison-Wesley, Boston, MA.

Feynman, R. (1970d). *Feynman Lectures on Physics*, volume I. Addison-Wesley, Boston, MA.

Fick, A. (1855). Ueber diffusion. *Annalen der Physik*, 170(1):59–86.

Fitzpatrick, R. (2008). *Maxwell's Equations and the Principles of Electromagnetism*. Jones & Bartlett.

Flanders, H. (1973). Differentiation under the integral sign. *American Mathematical Monthly*, 80(6):615–27.

Fletcher, N. and Rossing, T. (2008). *The Physics of Musical Instruments*. Springer New York. https://books.google.com/books?id=9CRSRYQlRLkC.

Fry, T. (1928). *Probability and its Engineering Uses*. Van Nostrand, Princeton NJ.

Galileo (1638). *Two New Sciences*. 1914, Macmillan & Co, New York,Toronto. https://files.libertyfund.org/files/753/0416_Bk.pdf.

Gallagher, T. F. (2005). *Rydberg atoms*, volume 3. Cambridge University Press.

Garber, D. (2004). On the frontlines of the scientific revolution: How mersenne learned to love galileo. *Perspectives on science*, 12(2):135–63.

Goldsmith, A. N. and Minton, J. (1924). The performance and theory of loud speaker horns. *Proceedings of the Institute of Radio Engineers*, 12(4):423–78.

Bombieri, E. (2000). Problems of the millennium: The Riemann hypothesis. In *Publications of the Clay Institute*, pages 1–11. Clay institute. "RiemannZeta-ClayDesc.pdf".

Boyer, C. and Merzbach, U. (2011). *A History of Mathematics*. Wiley, Hoboken, NJ. https://books.google.com/books?id=BokVHiuIk9UC.

Brillouin, L. (1953). *Wave Propagation in Periodic Structures*. Dover, London. Updated 1946 edition with corrections and added appendix.

Brillouin, L. (1960). *Wave Propagation and Group Velocity*, volume 8 of *Pure and Applied Physics*. Academic Press, Cambridge, MA, 154 pages.

Britannica (2004). *Encyclopædia Britannica Online*. Britannica Online. https://en.wikipedia.org/wiki/Pierre-Simon_Laplace.

Brune, O. (1931a). Synthesis of a finite two-terminal network whose driving point impedance is a prescribed function of frequency. *Jol. Mathematics and Physics*, 10:191–236.

Brune, O. (1931b). *Synthesis of a finite two-terminal network whose driving point impedance is a prescribed function of frequency*. PhD thesis, MIT, Campridge, MA. https://dspace.mit.edu/handle/1721.1/10661.

Burton, D. M. (1985). The history of mathematics: An introduction. *Group*, 3(3):35.

Calinger, R. S. (2015). *Leonhard Euler: Mathematical Genius in the Enlightenment*. Princeton University Press, Princeton, NJ.

Campbell, G. (1937). *The collected papers of George Ashley Campbell*. AT&T, Archives, 5 Reiman Rd, Warren, NJ. Forward by Vannevar Bush; Introduction by E.H. Colpitts.

Campbell, G. A. (1903). On loaded lines in telephonic transmission. *Philosophical Magazine*, 5(27):313–30.

Campbell, G. A. (1910). Telephonic intelligibility. *Philosophical Magazine*, 19(6):152–59.

Campbell, G. A. (1922). Physical theory of the electric wave-filter. *Bell System Technical Journal*, 1(2):1–32.

Carlin, H. J. and Giordano, A. B. (1964). *Network Theory, An Introduction to Reciprocal and Nonreciprocal Circuits*. Prentice Hall, Englewood Cliffs, NJ.

Cauer, W. (1932). The Poisson integral for functions with positive real part. *Bulletin of the American Mathematical Society*, 38(1919):713–17. https://www.ams.org/bull/1932-38-10/S0002-9904-1932-05510-0/S0002-9904-1932-05510-0.pdf.

Cauer, W. (1958). *Synthesis of Linear Communication Networks*, volume 1 and 2. McGraw-Hill, New York.

Cauer, W., Klein, W., Pelz, F., Knausenberger, G., and Warfield, J. (1958). *Synthesis of Linear Communication Networks*. McGraw-Hill, New York.

Condon, E. and Morse, P. (1929). *Quantum Mechanics*. McGraw-Hill, New York.

– 8.2: What is the inverse of the function $z(f)$? Does this function have a branch cut? If so, where is it?

– 8.3: Using `zviz.m`*, show that*

$$\tan^{-1}(z) = -\frac{\jmath}{2}\log\frac{\jmath - z}{\jmath + z}. \tag{DE-1.3}$$

– 8.4: Algebraically justify Eq. DE-1.3. Hint: Let $w(z) = \tan^{-1}(z)$ and $z(w) = \tan w = \sin w/\cos w$; then solve for e^{wj}.

4.3.5 A Cauer synthesis of any Brune impedance

Problem # *9: One may synthesize a transmission line (ladder network) from a positive real impedance $Z(s)$ by using the continued fraction method. To obtain the series and shunt impedance values, we can use a residue expansion. Here we shall explore this method.*

– 9.1: Starting from the Brune impedance $Z(s) = \frac{1}{s+1}$, find the impedance network as a ladder network.

– 9.2: Use a residue expansion in place of the CFA floor function (Sec. 2.4.4, p. 50) for polynomial expansions. Find the residue expansion of $H(s) = s^2/(s+1)$ and express it as a ladder network.

– 9.3: Discuss how the series impedance $Z(s,x)$ and shunt admittance $Y(s,x)$ determine the wave velocity $\kappa(s,x)$ and the characteristic impedance $z_o(s,x)$ when (1) $Z(s)$ and $Y(s)$ are both independent of x; (2) $Y(s)$ is independent of x, $Z(s,x)$ depends on x; (3) $Z(s)$ is independent of x, $Y(s,x)$ depends on x; and (4) both $Y(s,x), Z(s,x)$ depend on x.

4.4 Complex–analytic Brune admittance

It is rarely stated that the variable that we are integrating over, either x (space) or t (time), is real ($x, t \in R$), since that fact is implicit, due to the physical nature of the formulation of the integral. But this intuition must be refined once complex numbers are included with $s \in \mathbb{C}$, where $s = \sigma + \omega\jmath$.

That time and space are real variables is more than an assumption: it is a requirement that follows from the *order property*. Real numbers have order. For example, if $t = 0$ is now (the present), then $t < 0$ is the past and $t > 0$ is the future. Since time and space are real ($t, \boldsymbol{x} \in \mathbb{R}$), they obey this order property. To have time travel, time and space would need to be complex (they are not), since if the space axis were complex the order property would be invalid.

Interestingly, it was shown by d'Alembert (1747) that time and space are related by the pure delay. To obtain a solution to the governing wave equation, which d'Alembert first proposed for sound waves, $x, t \in \mathbb{R}$ may be functionally combined as

$$\zeta_{\pm} = t \pm x/c_o,$$

where $c_o \in \mathbb{R}$ [m/s] is the wave phase velocity. The d'Alembert solution to the wave equation, describing waves on a string under tension, is

$$u(x,t) = f(t - x/c_o) + g(t + x/c_o), \tag{4.4.1}$$

which describes the transverse velocity (or displacement) of two independent waves $f(\zeta_-)$, $g(\zeta_+) \in \mathbb{R}$ on the string, which represent forward and backward traveling waves.[1] For example, starting with a string at rest, if one displaces the left end, at $x = 0$, by a step function $u(t)$, then that step displacement will propagate to the right as $u(t - x/c_o)$, arriving at location x_o [m], at time x_o/c_o [s]. Before this time, the string will not move to the right of the wave-front, at x_o [m], and after t_o [s] it will have a non-zero displacement. Since the wave equation obeys superposition (postulate P2, p. 162), it follows that the "plane-wave" eigenfunction of the wave equation for $\mathbf{x}, \mathbf{k} \in \mathbb{R}^3$ are given by

$$\psi_{\pm}(\mathbf{x}, t) = \delta(t \mp \mathbf{k} \cdot \mathbf{x}) \leftrightarrow e^{st \pm \jmath \mathbf{k} \cdot \mathbf{x}}, \tag{4.4.2}$$

where $|\mathbf{k}| = 2\pi/|\boldsymbol{\lambda}| = \omega/c_o$ is the *wave number*, $|\boldsymbol{\lambda}| \in \mathbb{R}$ is the wavelength, and $s = \sigma + \omega\jmath$, is the Laplace frequency.

Complex propagation function $\boldsymbol{\kappa}(s)$: When propagation dispersion and losses are considered, we must replace the *wave number* $\jmath\mathbf{k} \in \mathbb{C}$ having a complex–analytic vector wave number $\boldsymbol{\kappa}(s) = \mathbf{k}_r(s) + \jmath\mathbf{k}(s)$. This is known by several names: (1) the *complex propagation function*, (2) the *dispersion relation*, (3) the *propagation function*. Function $\boldsymbol{\kappa}(s)$ is a subtle and important generalization of the scalar wave number $k = 2\pi/\lambda$.[2]

An interesting example is the exact solution to the acoustic wave equation, including viscous and thermal losses, as discussed in Appendix D, where it is show that the eigenvalues are

$$\kappa_{\pm}(s) = \frac{s \pm 2\beta_o\sqrt{s}}{c_o} = \frac{(\beta_o \pm \sqrt{s})^2 - (\beta_o)^2}{c_o},$$

with $\beta_o \in \mathbb{R} \geq 0$.

[1] d'Alembert's solution is valid for functions that are not differentiable, such as $\delta(t - x/c_o)$.

[2] Recall that for lossless plane waves $\lambda f = c_o$, and $k = 2\pi/\lambda$.

Bibliography

Allen, J. B. (1994). How do humans process and recognize speech? *IEEE Transactions on Speech and Audio*, 2(4):567–577.

Allen, J. B. (1996). Harvey Fletcher's role in the creation of communication acoustics. *Journal of the Acoustical Society of America*, 99(4):1825–1839.

Allen, J. B. (2024). Chaotic convergence of Newton's Method. *Transactions on Signal Processing*, 72:1683–1690.

Allen, J. B. and Berkley, D. A. (1979). Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65:943–950. https://jontalle.web.engr.illinois.edu/Public/AllenBerkley79.pdf.

Ambaum, M. H. (2010). *Thermal physics of the atmosphere*. Wiley Online Library.

Apostol, T. M. (2013). *Introduction to analytic number theory*. Springer Science & Business Media.

Apte, S. (2009). The *genius germs* hypothesis: Were epidemics of leprosy and tuberculosis responsible in part for the great divergence? *Hypothesis*, 7(1):e3. https://www.hypothesisjournal.com/wp-content/uploads/2011/09/vol7no1-hj003.pdf.

Arnold, D. and Rogness, J. (2019). Möbius transformations revealed. *Youtube.com*. video explaining the Riemann sphere: https://www.youtube.com/watch?v=0z1fIsUNhO4.

Batchelor, G. (1967). *An Introduction to Fluid Dynamics*. Cambridge University Press, Cambridge, UK.

Beranek, L. L. (1954). *Acoustics*. McGraw-Hill, New York, 451 pages.

Beranek, L. L. and Mellow, T. J. (2012). *Acoustics: Sound Fields and Transducers*. Academic Press, Elsevier, Waltham, MA.

Bilbao, S. (2004). *Wave and Scattering Methods for Numerical Simulation*. Wiley, 398 pages, London.

Boas, R. (1987). *Invitation to Complex Analysis*. The Random House, New York.

Bode, H. (1945). *Network Analysis and Feedback Amplifier Design*. Van Nostrand, New York.

Bohr, N. (1954). *Rydberg's discovery of the spectral laws*. C.W.K. Gleerup, Sweden. BohrRydbergLaw.54.pdf.

Forms of energy loss, which include viscosity and radiation, require $\boldsymbol{\kappa}(s) \in \mathbb{C}$. Physical examples include acoustic plane waves, electromagnetic wave propagation, antenna theory, and one of the most difficult cases, that of 3D electron wave propagating in crystals (e.g., silicon), where electrons and electro-magnetic (EM) waves are in a state of quantum mechanical equilibrium.

Even when we cannot solve these more difficult problems, we can still appreciate their qualitative solutions. One of the principles that allows us to do that is the causal nature of $\boldsymbol{\kappa}(s)$. Namely the $\mathcal{LT}^{-1}$ of $\boldsymbol{\kappa}(s)$ must be causal, thus Eq. 4.4.2 must be causal. The *group delay* then describes the nature of the frequency dependent causal delay. For example, if the group delay is large at some frequency, then the solutions will have the largest causal delay at that frequency (Brillouin, 1953; Papoulis, 1962). Qualitatively this can give a deep insight into the solution, even if we cannot compute it.

Electrons and photons are simply different EM states, and $\boldsymbol{\kappa}(\boldsymbol{x}, s)$ describes the crystal's dispersion relations as functions of both frequency and direction, famously known as *Brillouin zones*. Dispersion is a property of the medium such that the wave velocity is a function of frequency and direction, as in silicon.[3] Highly readable discussions on the history of this topic may be found in Brillouin (1953).

4.4.1 Generalized admittance/impedance

The most elementary examples of Brune admittance and impedance are those made of resistors, capacitors, and inductors. Such discrete element circuits arise not only in electrical networks but in mechanical, acoustical, and thermal networks as well (Table 3.2, p. 149). These lumped-element networks can always be represented by ratios of polynomials in s. This gives them a similar structure, with easily classified properties. Such circuits are called *Brune admittances* (or impedances).[4] An example of a special symmetry is when the degrees of the numerator and denominator polynomials cannot differ by more than one. This restriction on the degrees comes about because the real part of the admittance/impedance must be positive, due to physical constraints.

But there is a much broader class of admittances that come from transmission lines and other physical structures, which we refer to as *generalized admittances*. An interesting example is an admittance of the form $1/\sqrt{s}$, called a *semicapacitor*, and $\sqrt{s}$, called a *semiinductor*. Generalized admittance/impedance is not the ratio of two polynomials. As a result, they are more difficult to characterize.

When a generalized admittance $Y(s)$ or its impedance $Z(s) = 1/Y(s)$ is transformed into the time domain, it must have a real and positive surge admittance $\mathcal{Y}_r \in \mathbb{R}$ or surge impedance $\mathcal{Z}_r \in \mathbb{R}$, followed by the residual response $\upsilon(t), \zeta(t)$. We define the following notation for the admittance (both frequency and time responses),

$$Y(s) = \mathcal{Y}_r + \Upsilon(s) \leftrightarrow y(t) = \mathcal{Y}_r\delta(t) + \upsilon(t), \tag{4.4.3}$$

and the impedance,

$$Z(s) = \mathcal{Z}_r + \mathcal{Z}_i(s) \leftrightarrow z(t) = \mathcal{Z}_r\delta(t) + \zeta(t). \tag{4.4.4}$$

The complexity of the notation is necessary and follows from the fact that $z(t) \leftrightarrow Z(s)$ and $y(t) \leftrightarrow Y(s)$ are positive-real and thus minimum phase.

[3] In case you missed it, what I'm suggesting is that photons (propagating waves) and electrons (evanescent waves) are different EM wave "states" (Jaynes, 1991). This difference depends on the medium, which determines the dispersion relation (Papasimakis et al., 2018).

[4] Some texts prefer the term *immittance* to include both admittance and impedance.

When we are dealing with a transmission line (i.e., wave guides), the generalized admittance is defined as the ratio of the flow to the force. For an electrical system (voltage Φ, current I), the input admittance looking to the right from location x is

$$Y_{in}^{+}(x>0,s)=\frac{I^{+}(x,\omega)}{\Phi^{+}(x,\omega)},$$

and looking to the left is

$$Y_{in}^{-}(x<0,s)=\frac{I^{-}(x,\omega)}{\Phi^{-}(x,\omega)}.$$

In general these two admittances $Y_{in}^{\pm}(x,s)$ are different.

Generalized reflectance: A function related to the generalized impedance is the *reflectance* $\Gamma(s)$, defined as the ratio of a reflected wave to the incident wave. For the case of acoustics (pressure $\mathcal{P}$, volume velocity $\mathcal{V}$),

$$Y_{in}(x,s)\equiv\frac{\mathcal{V}(\omega)}{\mathcal{P}(\omega)}=\frac{\mathcal{V}^{+}-\mathcal{V}^{-}}{\mathcal{P}^{+}+\mathcal{P}^{-}} \tag{4.4.5}$$

$$=\frac{\mathcal{V}^{+}}{\mathcal{P}^{+}}\,\frac{1-\mathcal{V}^{-}/\mathcal{V}_{+}}{1+\mathcal{P}^{-}/\mathcal{P}^{+}} \tag{4.4.6}$$

$$=\mathcal{Y}_{r}^{+}\,\frac{1-\Gamma(x,s)}{1+\Gamma(x,s)}. \tag{4.4.7}$$

When the physical system is continuous at the measurement point x, $\mathcal{Y}_{r}^{+}(x)=\mathcal{Y}_{r}^{-}(x)\in\mathbb{R}$. The reflectance $\Gamma(x,s)$ depends on the area function, the boundary conditions, or both.

There is a direct relationship between a transmission line's area function $A(x)\in\mathbb{R}$, its characteristic impedance $\mathcal{Y}_{r}(x)\in\mathbb{R}$, and its eigenfunctions. We shall provide specific examples as they arise in our analysis of transmission lines (e.g., Fig. 5.3, p. 247). This was previously discussed in Sec. 41 (p. 116).

A few papers that deal with the relationship between $Y_{in}(s)$ and the area function $A(x)$ are Youla (1964); Sondhi and Gopinath (1971); Rasetshwane et al. (2012). However, the general theory of this important and interesting problem is beyond the scope of this text (See homework DE-3, Problem # 2), as well as Appendix D.

Complex–analytic $\Gamma(s)$ and $Y_{in}(s)=Z_{in}^{-1}(s)$

When we define the complex reflectance $\Gamma(s)$, we make a key assumption: Even though $\Gamma(s)$ is defined by the ratio of two functions of real (radian) frequency ω, like the impedance, the reflectance must be causal (Postulate P1, p. 162). That $\gamma(t)\leftrightarrow\Gamma(s)$ and $\zeta(t)\leftrightarrow Z_{in}(s)=1/Y_{in}(s)$ are causal is required by the physics.

4.4.2 Complex–analytic impedance

Conservation of energy (or power) is a cornerstone of modern physics. It may first have been under consideration by Galileo Galilei (1564–1642) and Marin Mersenne (1588–1648). Today the question is not whether it is true, but why. Specifically, what is the physics behind conservation of energy? Surprisingly, the answer is straightforward, based on its definition and the properties of impedance. Recall that the power is the product of the force and the flow, and impedance is their ratio.

I.4.3 Gauss-Lucas theorem

Of special interest is the ratio of a Diophantine monic $M_K(s)$ and its derivative $M'_K(s)$, because the roots of $M'_K(s)$ must lie just inside of the circle of roots of the monic. Assuming that

$$M_K(s)=\sum_{k=0}^{K}s^k\quad\text{then}\quad M'_K(s)=\sum_{k=0}^{K-1}ks^{k-1}.$$

I.4.4 Roots of random events

Of particular interest is the case of Bernoulli trials having a fractional odds ration ($o\in\mathbb{F}$). An example would be $o=N/M$ with $N,M\in\mathbb{N}$ (positive integers), or even better $N,M\in\mathbb{P}$ (primes with $o\in\mathbb{I}$ having no common factors), thus with $p=N/(N+M)\in\mathbb{I}$ (irrational).

I.4.2 Monics having coefficients taken from {1,0} (Bernoulli trials)

A second interesting question is what are the roots of polynomials having coefficients of all $\{1\}$, for example the 4th degree monic

$$M_4(s) = s^4 + s^3 + s^2 + s + 1 \equiv [1, 1, 1, 1, 1]?$$

This case is easily dealt with because

$$M_K(s) = \sum_{k=0}^{K} s^k = \frac{1 - s^{K+1}}{1 - s}$$

thus the roots are $s_k = e^{j2\pi k/K}$ for $k = 1 : K$. Note that $s_0 = 1$ (i.e., $k = 0$) is not a zero because that obvious zero (root) is canceled (removed) by the pole at $s = 1$.

A second seemingly more difficult example comes from replacing one or more of the coefficients with 0, such as

$$M_4(s) = s^4 + 0s^3 + 0s^2 + s + 1 \equiv [1, 0, 0, 1, 1]?$$

The main question here is "What are the roots of such monics?" Also are there numerical errors in Matlab/Octave when the coefficients are integers, or perhaps even primes? Such problem arise when dealing the estimation of the odds ratio for a Bernoulli trial, or non-stationary Bernoulli trials (a long sequence of negatives trials, followed by a single (or a few) positives, such as

$$M(x) = x^{20} + x^{19} \cdots + x^2 + 0x^1 + 0 \equiv [1, 1, \cdots, 1, 0, 0].$$

or alternatively, the very different sequence, having the same number of pass/fail outcomes

$$M(x) = x^{20} + 0x^{19} + x^{18} \cdots + x^2 + 0x^1 + 1 \equiv [1, 0, 1, \cdots, 1, 0, 1].$$

Given a long string of negatives (1) followed by a few positives (0) is a clear indication that the subject has switched from negative to positive. A string of 18 negatives is strong evidence that the test has both high sensitive and specificity [i.e., the odds-ratio was negative (1) 18 times, and then on trial 19, switched to positive (0) twice].

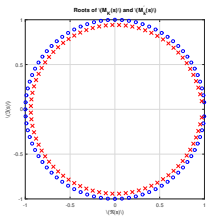


Figure I.1: *Plot of the zeros (o) of a monic $M_{80(s)}$ having coefficents $c_k = 1$ along with its derivative $M'_{80}(s)$ (x). The placement of the roots follows from the Gauss-Lucus theorem, which says that the roots of the Monic's derivative $M'_K(s)$ lie inside the convex hull of the roots of the monic $M_K(s)$.*

The power is given by the product of two variables, sometimes called *conjugate variables*, the force and the flow. In electrical terms, these are voltage (force) ($v(t) \leftrightarrow V(\omega)$) and current (flow) ($i(t) \leftrightarrow I(\omega)$); thus, the electrical power at any instant of time is[5]

$$\mathcal{P}(t) = v(t)i(t). \tag{4.4.8}$$

The total energy $\mathcal{E}(t)$ is the integral of the power, since $\mathcal{P}(t) = d\mathcal{E}/dt$. Thus if we start with all the elements at rest (no currents or voltages), then the energy as a function of time is always positive

$$\mathcal{E}(t) = \int_0^t \mathcal{P}(t)dt \geq 0 \tag{4.4.9}$$

and is simply the total energy applied to the network (Van Valkenburg, 1964a, p. 376). Since the voltage and current are related by either an impedance or an admittance, conservation of energy depends on the property of impedance. From Ohm's law and Postulate P1 (every impedance is causal), and we have

$$\begin{aligned} v(t) &= z(t) \star i(t) \\ &= \int_{\tau=0}^{t} z(\tau)i(t-\tau)d\tau \\ &\leftrightarrow V(s) = Z(s)I(s). \end{aligned}$$

Example: Let $i(t) = \delta(t)$ (a perfect impulse). Then

$$I_{xx}(t) = \int_{\tau=0}^{t} z(t-\tau)\delta(\tau)d\tau = \int_0^t z(-\tau)d\tau.$$

Every Brune impedance always has the form $z(t) = r_o\delta(t) + \zeta(t)$. The *characteristic impedance* r_o (also called *surge impedance*) may be defined as (Lundberg et al., 2007)

$$r_o = \int_{0-}^{\infty} z(t)dt.$$

This definition requires that the integral of $\zeta(t)$ is zero, a conclusion that warrants further investigation.

These ideas are perhaps easier to visualize if we work in the Laplace frequency domain. Then the total energy, equal to the integral of the real part of the power, is

$$\frac{1}{s}\Re VI = \frac{1}{2s}(V^*I + VI^*) = \frac{1}{2s}(Z^*I^*I + ZII^*) = \frac{1}{s}\Re Z(s)|I|^2 \geq 0.$$

Mathematically this is called a *positive definite operator*, since the positive and real resistance is sandwiched between the current, forcing the "definiteness."

In conclusion, conservation of energy is totally dependent on the properties of the impedance. Thus one of the most important and obvious applications of complex–analytic functions of a complex variable is the impedance function. This seems to be the ultimate example of the FTCC applied to $z(t)$.

[5]The voltage is sometimes called the Electromotive force (EMF). However $v(t)$ is relative to a reference ground. The actual EMF is $-\nabla v(t)$.

Every impedance must obey conservation of energy (Postulate P3): The impedance function $Z(s)$ has resistance R and reactance X as a function of complex frequency $s = \sigma + \jmath\omega$. From the causality postulate P1, $z(t < 0) = 0$. Every impedance is defined by a Laplace transform pair

$$z(t) \leftrightarrow Z(s) = R(\sigma, \omega) + \jmath X(\sigma, \omega),$$

with $R, X \in \mathbb{R}$.

According to postulate P3 a system is passive if it does not contain a power source. Drawing power from an impedance violates conservation of energy. This property is also called *positive-real*, which was defined by Brune (1931b,a)

$$\Re\{Z(s \geq 0)\} \geq 0. \tag{4.4.10}$$

Positive-real systems cannot draw more power than is stored in the impedance. The region $\sigma \leq 0$ is called the *left half s plane* (LHP), and the complementary region $\sigma > 0$ is called the *right half s plane* (RHP). According to the Brune condition, the real part of every impedance must be nonnegative (the RHP).

It is easy to construct examples of second-order poles or zeros in the RHP such that Postulate P3 is violated. Thus Postulate P3 implies that the impedance may *not* have more than simple (first-order) poles and zeros, strictly in the LHP. But there is more: These poles and zeros in the LHP must meet a *minimum phase condition*, a condition that is easily stated:

$$\angle Z(s) < \angle s \tag{4.4.11}$$

but difficult to prove. There seems to be no proof that second-order poles and zeros (e.g., second-order roots) are not allowed.[6] However, such roots must violate a requirement that the poles and zeros must alternate on the $\sigma = 0$ axis, which follows from Postulate P3. In the complex plane the concept of "alternate" is not defined (complex numbers cannot be ordered). What *has* been proved (i.e., Foster's reactance theorem), that says that if the poles are on the real or imaginary axis, they must alternate, leading to simple poles and zeros (Van Valkenburg, 1964a). The restriction on poles is sufficient but not necessary, as $Z(s) = 1/\sqrt{s}$ is a positive-real impedance but is less than a first-degree pole (Kim and Allen, 2013). The corresponding condition in the LHP, and its proof, remains elusive (Van Valkenburg, 1964a).

For example, a series resistor R_o and capacitor C_o have an impedance given by (Table 3.9, p. 160)

$$Z(s) = R_o + 1/sC_o \leftrightarrow R_o\delta(t) + \frac{1}{C_o}u(t) = z(t), \tag{4.4.12}$$

with constants $R_o, C_o \in \mathbb{R} \geq 0$. In mechanics, an impedance composed of a dashpot (damper) and a spring has the same mathematical form.

A full 2d order resonant system has an inductor, resistor, and capacitor, with an impedance given by

$$Z(s) = \frac{sC_o}{1 + sC_oR_o + s^2C_oM_o} \leftrightarrow C_o\frac{d}{dt}\left(c_+e^{s_+t} + c_-e^{s^-t}\right) = z(t), \tag{4.4.13}$$

which is a second-degree polynomial with two complex resonant frequencies $s_\pm$. When $R_o > 0$, these roots are in the LHP, with $z(t) \leftrightarrow Z(s)$.

Systems (networks) that contain many elements and transmission lines can be much more complicated, yet still have a simple frequency-domain representation. This is the key to understanding how these physical systems work, as we describe next.

[6] As best I know, this is an open problem in network theory (Brune, 1931a; Van Valkenburg, 1964a).

I.4 Properties of Diophantine polynomials

I.4.1 Polynomials having prime roots

There is an interesting debate regarding the convergence of Newton's method. My view is that it always converges (Sec. 3.1.2).

The following are the coefficients of a polynomial built from prime numbers

$$C = [950488110; -1696103097; 1255300345; -506088300; 123054080; -18928998; 1879134; -119820; 473$$

Since we know the roots are prime, we may find them by factoring $C(1)$, which gives

$$R = [2, 3, 3, 5, 11, 13, 13, 13, 19, 23],$$

where `R= factor(950488110).`

Directly finding the roots of the polynomial using Matlab/Octave's `roots(C)` gives

$$[0.5, 0.333333, 0.333333, 0.2, 0.090909, 0.076933, 0.076918, 0.076918, 0.052632, 0.043478].$$

These number are less than 1 because $C(10)s^{10}$ is normalized to 1, forcing it to be monic. Taking the reciprocal of the roots should return the prime roots. However note the two small errors ([12.9983, 13.0008])

$$[2.0, 3.0, 3.0, 5.0, 11.0, 12.9983, 13.0008, 13.0008, 19.0, 23.0],$$

due to floating point numerical issues.

There are several alternatives to this. One is to use the continued fraction algorithm (CFA) to find rational expansions of the roots. In Matlab the CFA is implemented with the command `rat()`. For example

```
rat(1/0.5) = 1/2
```

and the fifth root gives

```
rat(1/0.090909) = 11.
```

However the 6th root fails the rats expansion

```
rat(1/0.076933) = 13 + 1/(-596)
```

and for the 7th root,

```
rat(1/0.076918)=0 + 1/(13 + 1/(1165)).
```

An interesting question remains: "Can these errors be removed by factoring out a few of the smaller primes ($[2, 3, 5]$), or the larger primes ($[19, 23]$)?" Since there are 3 factors of 13, what is the effect of trying to remove that factor. Namely is there a remainder due to the near miss, or will it factor 13 with no remainder?

This question is interesting because Newton's method could be used to find roots, which could then be removed using Matlab/Octave's long division routine (`decon()`). Other potential options are a residue expansion, or its more power sister the *vector fitting* algorithm (Matlab's `vectorfit3()`).

I.3 Taking the inverse to get the `gcd`

Variables p and q are the remainders r_{n-1} and r_n, respectively. Using this notation with $n-1$ gives Eq. I.2.1. Inverting this gives the formula for the GCD:

$$\begin{bmatrix} r_{n-1} \\ r_n \end{bmatrix} = \begin{matrix} 0 & 1 \\ 1 & -\frac{r_{n-2}}{r_{n-1}} \end{matrix} \begin{bmatrix} r_{n-2} \\ r_{n-1} \end{bmatrix}.$$

This terminates when $r_n = 0$ and the `gcd(p,q)` is r_{n-1}. Not surprisingly these equations mirror Eq. 2.4.3 (p. 49), but with a different indexing scheme and interpretation of the variables.

This then explains why Gauss called the CFA the *Euclidean algorithm*. He was not confused. But since the equations have an inverse relationship, they are not strictly the same.

Poles and zeros of positive-real functions must be first-degree: The definition of *positive-real* (PR) functions requires that the poles and zeros of the impedance function be simple (only first-degree). Second-degree poles would have a reactive "secular" response of the form $h(t) = t\sin(\omega_k t + \phi)u(t)$, and these terms would not average to zero, depending on the phase, as is required of an impedance. As a result, only single-degree poles are possible.[7] I believe that no one has ever reported an impedance that has second-degree poles and zeros. Network analysis books never report second-degree poles and zeros in their impedance functions. Nor has there ever been any guidance about where the poles and zeros might lie in the LHP. Understanding the exact relationships between pairs of poles and zeros, to assure that the real part of the impedance is real, would resolve this longstanding unsolved problem (Van Valkenburg, 1964b). It is the residues that determine the LHP simple pole degree.

Calculus on Complex–analytic functions: To solve a differential equation or integrate a function, Newton used the Taylor series to integrate one term at a time. However, he used only real functions of a real variable due to his fundamental lack of appreciation of the complex analysis. This same method is how one finds solutions to scalar differential equations today, but using an approach that makes the solution method less obvious. Rather than working directly with the Taylor series, today we use the complex exponential, since the complex exponential is an eigenfunction of the derivative

$$\frac{d}{dt}e^{st} = se^{st}.$$

Since e^{st} may be expressed as a Taylor series that has coefficients $c_n = 1/n!$, in some real sense the modern approach is a compact way of doing what Newton did. Thus every linear constant coefficient differential equation in time may be simply transformed into a polynomial in complex Laplace frequency s, by looking for solutions of the form $F(s)e^{st}$ and transforming the differential equation into polynomial $F(s)$ in complex frequency.

For example

$$\frac{d}{dt}f(t) + af(t) = \delta(t) \leftrightarrow (s+a)F(s) = 1.$$

The pole of $F(s_r)$ is $s_r + a = 0$ is the eigenvalue of the differential equation. Thus a powerful tool for understanding the solutions of differential equations, both scalar and vector, is to work in the Laplace frequency domain using their eigenvalues (i.e., $s_r = -a$) and their eigenfunctions.

The Taylor series may be replaced by e^{st}, which transforms Newton's real Taylor series into the complex exponential eigenfunction. In some sense, these are the same methods, since

$$e^{s_r t} = \sum_{n=0}^{\infty} \frac{(s_r t)^n}{n!}. \tag{4.4.14}$$

Taking the derivative with respect to time gives

$$\frac{d}{dt}e^{st} = se^{st} = s\sum_{n=0}^{\infty} \frac{(st)^n}{n!}, \tag{4.4.15}$$

which is also complex–analytic. Thus if the series for $F(s)$ is valid (i.e., it converges), then its derivative is also valid. This was a very powerful concept, exploited by Newton for real functions of a real variable, and later by Cauchy and Riemann for complex functions of a

[7] Secular terms result from second-degree poles, since $u(t) \star u(t) = tu(t) \leftrightarrow 1/s^2$.

complex variable. The key question is "Where does the series fail to converge?" This is the main concept behind the FTCC (Eq. 4.2.3).

The FTCC (Eq. 4.2.3) is formally the same as the FTC (Eq. 4.2.2) (Leibniz's formula), the key (and significant) difference being that the argument of the integrand $s \in \mathbb{C}$. Thus this integration is a line integral in the complex plane. One would naturally assume that the value of the integral depends on the path of integration. And it does, but in a subtle way, as quantified by Cauchy's various theorems. If the path stays in the same RoC region, then the integral is independent of that path. If a path includes a different pole, then the integral depends on the path, as quantified by the Cauchy residue theorem. The test is to deform the path from the first to the second. If in that deformation the path crosses a pole, then the integral will change (namely it will depend on the path). All of this is follows from causality.

The FTC and FTCC are clearly distinguishable yet the reasoning is the same. If $F(s) = df(s)/ds$ is complex–analytic [i.e., has a power series $f(s) = \sum_k c_k s^k$, with $f(s), c_k, s \in \mathbb{C}$], then it may be integrated, term by term, and yet the integral does not depend on the path. At first blush, this is sort of amazing. The key is that $F(s)$ and $f(s)$ must be complex–analytic, which means they are differentiable. This all follows from the Taylor series formula Eq. 4.2.5 (p. 180) for the coefficients of the complex–analytic series. For Eq. 4.2.3 to hold, the derivatives must be independent of the direction (the path), as discussed on page 181. The concept of a complex–analytic function therefore has eminent consequences in the form of several key theorems on complex integration, as first discovered by Cauchy in about 1820.

Role of the Complex Taylor series: The complex Taylor series generalizes the functions it describes, with unpredictable consequences, as nicely shown by the domain-coloring diagrams in Fig. 3.12 (p. 167), where a simple translation of the s plane by a complex number can void the positive-real property ($s - \sqrt{j}$) cannot be a physical impedance).

Cauchy's of complex integration tools were first exploited in physics by Sommerfeld (1952), to explain the onset (e.g., causal) transients in waves, as he explained in detail in Brillouin (1960, Chap. 3).

The importance of causality: Up to 1910, when Sommerfeld first published his results using complex–analytic signals and saddle point integration in the complex plane, the implications of the causal wave-front were poorly understood. It would be reasonable to say that Sommerfeld's insight changed our understanding of wave propagation for both light and sound. Unfortunately, in my view, his insight has never been fully appreciated, perhaps even to this day. If you question this summary, please read Brillouin (1960, Chap. 1).

The full power of the complex–analytic function was first appreciated by Bernard Riemann (1826–1866) in his University of Göttingen PhD thesis of 1851, under the tutelage of Gauss (1777–1855), which drew heavily on the work of Cauchy (Fig. 3.1, page).

The key definition of a complex–analytic function is that it has a Taylor series representation, over a region of the complex frequency plane $s = \sigma + j\omega$, that converges in its RoC about the expansion point, with a radius determined by the nearest pole of the function. A further surprising feature of all analytic functions is that within the RoC, the inverse of that function also has a complex–analytic expansion. Thus given $w(s)$, in theory, we can determine $s(w)$ to any desired accuracy, critically depending on the RoC. As an example if $w(s) = e^s$ then its inverse is $s(w) = log(w)$. Given the right software (e.g., `zviz.m`), this relationship may be made precise for every complex–analytic function.

From the remainder formula, $u_1 = \lfloor p/q \rfloor$. Continuing with $n = 2$:

$$\begin{bmatrix} r_0 \\ r_1 \end{bmatrix} = \begin{bmatrix} u_2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix},$$

where $u_1 = \lfloor r_0/r_1 \rfloor$. Continuing with $n = 3$:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} u_3 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r_2 \\ r_3 \end{bmatrix},$$

where $u_2 = \lfloor r_1/r_2 \rfloor$.

For arbitrary n we find

$$\begin{bmatrix} r_{n-2} \\ r_{n-1} \end{bmatrix} = \begin{bmatrix} u_n & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r_{n-1} \\ r_n \end{bmatrix}, \tag{I.2.1}$$

where $u_n = \lfloor r_{n-1}/r_n \rfloor$. This terminates when $r_n = 0$ in the above nth step:

$$\begin{bmatrix} r_{n-2} \\ r_{n-1} \end{bmatrix} = \begin{bmatrix} u_n & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r_{n-1} \\ r_n = 0 \end{bmatrix}.$$

Example: We let $p = 355$ and $q = 113$, which are coprime, and set $n = 1$. Then Eq. I.2.1 becomes

$$\begin{bmatrix} 355 \\ 113 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix},$$

since $u_1 = \lfloor \frac{355}{113} \rfloor = 3$. Solving for the RHS gives $[r_0; r_1] = [113; 16]$ ($355 = 113 \cdot 3 + 16$). To find $[r_0; r_1]$, we take the inverse:

$$\begin{bmatrix} r_0 \\ r_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 355 \\ 113 \end{bmatrix}.$$

For $n = 2$, with the RHS from the previous step,

$$\begin{bmatrix} 113 \\ 16 \end{bmatrix} = \begin{bmatrix} u_2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix},$$

since $u_2 = \lfloor \frac{113}{16} \rfloor = 7$. Solving for the RHS gives $[r_1; r_2] = [16; 1]$ ($113 = 16 \cdot 7 + 1$). It seems we are done, but let's go one step further.

For $n = 3$ we now have

$$\begin{bmatrix} 16 \\ 1 \end{bmatrix} = \begin{bmatrix} u_3 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix},$$

since $u_3 = \lfloor \frac{16}{1} \rfloor = 16$. Solving for the RHS gives $[r_1; r_2] = [1; 0]$. This confirms that we are done, since $r_2 = 0$.

Derivation of Eq. I.2.1: Equation Eq. I.2.1 is derived as follows: Starting from the target $x \in \mathbb{R}$, we define

$$p = \lfloor x \rfloor \quad \texttt{and} \quad q = \frac{1}{x - p} \in \mathbb{R}.$$

These two relationships for truncation and remainder allow us to write the general matrix recursion relation for the CFA (Eq. I.2.1). Given $\{p, q\}$, we continue with the above CFA method.

One slight problem with the above is that the output is on the right and the input on the left. Thus we need to take the inverse of these relationships to turn this into a composition.

Note that the next value of $m = M(1)$ is always less than $n = M(2)$ and must remain greater than or equal to zero. This one-line vector operation is then repeated until the remainder $M(1)$ is 0. The gcd is then $n = M(2)$. When we use irrational numbers, the code still works except the error is never exactly zero due to IEEE 754 rounding. Thus the criterion must be that the error is within some small factor times the smallest number (which in Matlab/Octave is the number `eps` = 2.220446049250313 $\times 10^{-16}$, as defined in the IEEE 754 standard).

Thus, without factoring the two numbers, Eq. I.1.2 recursively finds the gcd. Perhaps this is best seen with some examples.

The GCD is an important and venerable method, useful in engineering and mathematics but, as best I know, not typically taught in the traditional engineering curriculum.

GCD applied to polynomials: An interesting generalization is to work with polynomials rather than numbers and apply the Euclidean algorithm.

How about trix case?

The GCD may be generalized in several significant ways. For example, what is the GCD of two polynomials? To answer this question, we must factor the two polynomials to identify possible common roots. For numerical reasons it is wise to represent the roots $s_k \in \mathbb{F}$.

I.2 Derivation of the CFA matrix

We can define the continued fraction algorithm (CFA) starting from the basic definitions of the floor and remainder formulas. Starting with a decimal number x, we split it into the decimal and remainder parts.[2] If we start with $n = 0$ and $x_o = x \in \mathbb{I}$, the integer part is

$$m_0 = \lfloor x \rfloor \in \mathbb{N}$$

and the remainder is

$$r_0 = x - m_0.$$

Corresponding to the CFA, the next target x_1 for $n = 1$ is

$$x_1 = r_0^{-1}$$

and the integer part is $m_1 = \lfloor x_1 \rfloor$. As in the case of $n = 0$, the integer part is

$$m_1 = \lfloor x_1 \rfloor$$

and the remainder is

$$r_1 = x_1 - m_1.$$

The recursion for $n = 2$ is similar.

For us to better appreciate what is happening, it is helpful to write these recursions in matrix format. Rewriting the case of $n = 1$ and using the remainder formula for the ratio of two numbers $p \geq q \in \mathbb{N}$ with $q \neq 0$, we have

$$\begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} u_1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}.$$

[2] The method presented here was developed by Yiming Zhang as a student project in 2019.

4.4.3 Multivalued functions

In the field of mathematics there seems to have been a tug-of-war regarding the basic definition of a function. The accepted definition today is a single-valued (i.e., complex–analytic) mapping from the domain to the codomain (or range). This makes the discussion of complex–analytic multivalued functions somewhat awkward. In 1851 Riemann (working with Gauss) resolved this problem for the complex–analytic set of multivalued functions by introducing the geometric concept of single-valued sheets, delineated by branch cuts.

Two simple yet important examples of multivalued functions are the circle $z^2 = x^2 + y^2$ and $w = \log(z)$. For example, if we assume $\rho = |z|$ is the radius of the circle, then solving for $y(x)$ gives the double-valued function

$$y(x) = \pm \sqrt{\rho^2 - x^2}.$$

The related function $z = \pm\sqrt{x}$, with $x \in \mathbb{C}$, is shown in Fig. 4.1 as a three-dimensional display in polar coordinates, with $z(r)$ as the vertical axis, as a function of the angle and radius of $x \in \mathbb{C}$.

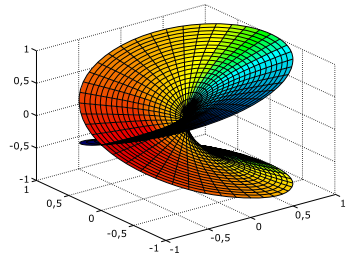


Figure 4.1: *The mapping for the square root function This function has two single-valued sheets of the x plane corresponding to the two signs of the square root. The best way to view this function is in polar coordinates, with $x = |x|e^{j\phi}$ and $z = \sqrt{|x|}e^{j\phi/2}$. p. 196, 197. See* `https://en.wikipedia.org/wiki/Riemann_surface`

If we accept the modern injective definition of a function, as the mapping from one set to a second set, then $y(x)$ is not a function, or even two functions. For example, what if $x > z$? Or worse, what if $z = 2\jmath$ with $|x| < 1$? Riemann's construction, using branch cuts for multivalued function, resolves all these difficulties (as best I know).

To proceed, we need definitions and classifications of the various types of complex singularities:

1. Poles of degree 1 are called *simple poles*. Their amplitude is called the *residue* (e.g., α/s has residue α). Simple poles are special (see Eq. 4.5.3, p. 201); they play a key role in mathematical physics, since their inverse Laplace transform defines a causal eigenfunction.

2. When the numerator and denominator of a rational function (i.e., ratio of two polynomials) have a common root (i.e., factor), that root is said to be *removable*.

3. A singularity that is not removable, a pole, or a branch point is called *essential*.

4. A complex–analytic function (except for isolated poles) is called *meromorphic* (Boas, 1987). Meromorphic functions can have any number of poles, even an infinite number. The poles need not be simple.

5. When the first derivative of a function $Z(s)$ has a simple pole at a, then a is said to be a *branch point* of $Z(s)$. An important example is the logarithmic derivative:

$$d\ln(s-a)^\alpha/ds = \alpha/(s-a), \ \alpha \in \mathbb{I}.$$

However, the converse does not necessarily hold.

6. I am not clear about the interesting case of an irrational pole ($\alpha \in \mathbb{I}$). In some cases (e.g., $\alpha \in \mathbb{F}$) this may be simplified with the logarithmic derivative operation, as mentioned on page 84.

More complex topologies are being researched today, and progress is expected to accelerate due to modern computing technology.[8] It is helpful to identify the physical meaning of these more complex surfaces, to guide us in their interpretation and possible applications.[9]

Branch cuts: Up to this point we have considered only poles of degree $\alpha \in \mathbb{N}$ of the form $1/s^\alpha$. The concept of a branch cut allows us to manipulate (and visualize) multivalued functions for which $\alpha \in \mathbb{F}$. This is done by breaking each region into single-valued sheets, as shown in Fig. 4.2 (right). The branch cut is a curve $\in \mathbb{C}$ that separates the various single-valued sheets of a multivalued function. The concepts of branch cuts, sheets, and the extended plane were first devised by Riemann, as described in his thesis of 1851. It was these three mathematical and geometrical constructions that provided deep insight into complex–analytic functions, greatly extending the important earlier work of Cauchy (1789–1857) on the calculus of complex–analytic functions. For an alternative helpful discussion of Riemann sheets and branch cuts, see Boas (1987, pp. 221–25), Kusse and Westwig (2010), and Greenberg (1988).

To study the properties of multivalued functions and branch cuts, we look at $w(s) = \sqrt{s}$ and $w(s) = \log(s)$, along with their inverse functions $w(s) = s^2$ and $w(s) = e^s$. For uniformity we refer to the *complex abscissa* ($s = \sigma + \omega\jmath$) and the *complex ordinate* ($w(s) = u + v\jmath$). When the complex domain and range are swapped, by taking the inverse of a function, multivalued functions are a common consequence. For example, $f(t) = \sin(t)$ is single-valued and analytic in t and thus has a Taylor series. The inverse function $t(f)$ is multivalued.

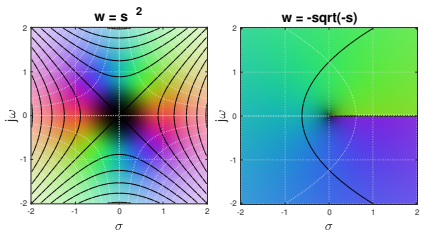


Figure 4.2: *Here we use Cartesian coordinates in the domain $s = \sigma + \omega\jmath$ and polar coordinates for the range $w(\sigma, \omega) = |w|e^{\psi\jmath}$. The color intensity indicates the magnitude $|s|$, with black being $|s = 0|$ and bright (eventually white) indicating $|s| \to \infty$.* **Left:** *Mapping: $w(s) = s^2$.* **Right:** *Mapping of the principal branch from $-s$ to $w(s) = -\sqrt{-s}$ (i.e., the rotated inverse of s^2). This sheet was generated by rotating w by 180°. The branch cut is on the $\psi = 0$ axis, with branch points at $|w| = 0$ and ∞. Neither of these functions are Brune impedances, since they violate the positive-real condition (Eq. 3.2.19, p. 102).*

The best way to explore the complex mapping from the complex planes $s \to w(s)$ is to master the single-valued function $s = w^2(s)$ and its double-valued inverse $w(s) = \sqrt{s}$.

Figure 4.2 (left) shows the single-valued function $w(s) = s^2$, and (right) one sheet of its inverse, the double-valued mapping of $w(s) = -\sqrt{-s}$. Single-valued functions such as $w(s) = s^2$ are relatively straightforward. Multivalued functions require the concept of a branch cut, defined in the image plane (also called the codomain or range). This is a technique to render the multiple values as single-valued on each of several sheets. Each sheet is labeled in the domain (s) plane by a sheet index $k \in \mathbb{Z}$, while branch points and cuts are defined in the image (w) plane.

[8] https://www.maths.ox.ac.uk/about-us/departmental-art/theory

[9] https://www.quantamagazine.org/secret-link-uncovered-between-pure-math-and-physics-20171201

Appendix I

Number theory applications

I.1 Division with rounding method

We want to show that the GCD for $m, n, k \in \mathbb{N}$ (Eq. 2.4.2, p. 47) may be written in matrix form as

$$\begin{bmatrix} m \\ n \end{bmatrix}_{k+1} = \begin{bmatrix} 0 & 1 \\ 1 & -\lfloor \frac{m}{n} \rfloor \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix}_k . \tag{I.1.1}$$

Eq.I.1.1 implements the $\gcd(m, n)$ for $m > n$.

This starts with $k = 0$, $m_0 = a$, and $n_0 = b$. With this method there is no need to test whether $n_n < m_n$, as it is built into the procedure. The method uses the floor function $\lfloor x \rfloor$, which finds the integer part of x ($\lfloor x \rfloor$ rounds toward $-\infty$). After each step we will see that the value $n_{k+1} < m_{k+1}$. The method terminates when $n_{k+1} = 0$ with $\gcd(a, b) = m_{k+1}$.

The following vectorized code is more efficient than the direct matrix method:

```
function n=gcd2(a,b)

M=[abs(a);abs(b)]; %Save (a,b) in array M(2,1)

% done when M(1) = 0
while M(1) ~= 0
disp(sprintf('M(1)=%g, M(2)=%g ',M(1),M(2)));
M=[M(2)-M(1)*floor(M(2)/M(1)); M(1)]; %automatically sorted
end  %done

n=M(2); %GCD is M(2)
```

With a minor extension in the test for "end," this code can be made to work with irrational inputs (e.g., $(n\pi, m\pi)$).

This method calculates the number of times $n < m$ must subtract from m using the floor function. This operation is the same as the mod function.[1] Specifically,

$$n_{k+1} = m_k - \frac{m}{n} \; n_k \tag{I.1.2}$$

so that the output is the definition of the remainder of modular arithmetic. This would have been obvious to anyone using an abacus, which explains why it was discovered so early.

[1] https://en.wikipedia.org/wiki/Modulo_operation

The sheets are indexed by a sheet index, and separated by the branch cut. It is important to understand that the path of every branch cut is not unique and may be moved. However, branch points are unique and thus not movable.

A function may be multivalued in both the domain and image planes. As an example consider $w(s) = s^{3/2}$.

The multivalued nature of $w(s) = \sqrt{s}$ is best understood by working with the function in polar coordinates. We let

$$s_k = re^{\theta_J} e^{2\pi k_J}, \tag{4.4.16}$$

where $r = |s|, \theta = \angle s, \in \mathbb{R}$, and $k \in \mathbb{Z}$ is the sheet index.

This concept of analytic inverses becomes important only when the function is multivalued. For example, since $w(s) = s^2$ has a period of 2, $s(w) = \pm\sqrt{w}$ is multivalued. Riemann dealt with such extensions using the concept of a branch cut with multiple sheets labeled by sheet numbers. Each sheet describes an analytic function (Taylor series) that converges within some RoC that has a radius out to the nearest pole. Thus Riemann's branch cuts and sheets explicitly deal with the need to define unique single-valued inverses of multivalued functions. Since the square root function has two overlapping regions corresponding to the $\pm$ due to the radical, there must be two connected region—sort of like mathematical Siamese twins: distinct, yet the same.

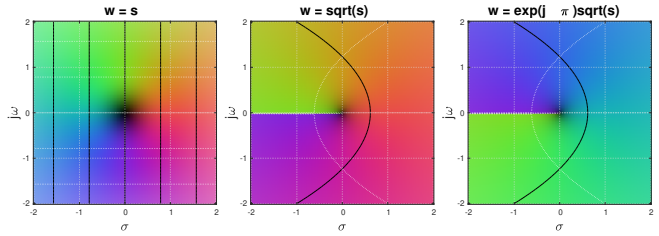


Figure 4.3: *Colorized plots of $s_k = |s|e^{\theta_J} e^{2\pi k_J}$ and $w_k(s) = \sqrt{s_k} = \sqrt{|s|}e^{\theta_J/2} e^{\pi k_J}$, as defined in polar coordinates by Eqs. 4.4.16 and 4.4.17.* **Left:** *Color map hue reference plane $s = |s|e^{\theta_J} e^{2\pi k_J}$. This function is a Brune impedance (it represents an inductor).* **Center:** *Sheet index $k = 0$, $w(s) = \sqrt{|s|}e^{j\theta/2}$ for $-\pi < \theta < +\pi$ and $\psi = \angle w(s) = \theta/2$ between $\pm\pi/2$ [rad]. This function is positive-real (Eq. 3.2.19, p. 102) but is not a Brune impedance, since it is not the ratio of polynomials in s. Thus it is the* generalized impedance, *known as a semiinductor.* **Right:** *Sheet index $k = 1$, $w(s) = e^{j\theta/2} e^{j\pi} \sqrt{|s|}$ for $\pi < \theta < 3\pi$ and $\pi/2 < \psi < 3\pi/2$ [rad]. The branch cut is at $\theta = \angle s = \pm 180°$ (π [rad]) where the hue of w changes abruptly from green to purple (center) and from purple back to green (right). Note how the hue matches between the center and right panels at the branch cut: in the center panel purple runs along -180° and along +180° on the right. Likewise, green runs along +180° in the center and along -180° on the right. Thus $w(s) = \sqrt{s}$ is analytic on the branch cut connecting the two sheets ($k = 0 \to k = 1$). This function is not an impedance, since it is negative in the RHP.*

Hue: By studying Fig. 4.3, we can appreciate domain-coloring. The angle-to-hue map is shown in the left panel of Fig. 4.3. The domain angles $\angle s$ go from $-90° < \theta < 90°$ (purple to red to green). For $w(s) = \sqrt{s}$, the $\angle s$ is *compressed* by a factor of 2 ($\psi = \theta/2$) with purple being $-180°$, and green being $+180°$.

Thus for Fig. 4.3 the principal ($k = 0$) angle $\angle s$ ($-180° < \theta < 180°$) maps to $w(s) = \sqrt{s}$ (middle panel) to half the w plane ($-90° < \psi < 90°$ (from purple to red to green).

The $k = 1$ branch, of $\angle s$ ($+180° < \theta < 180 + 360 = 520°$) maps to $\angle w$ (on the right) to green to blue to purple ($\psi = \theta/2$). Note how the panel on the left matches the right half of s (green = $+90°$, purple = $-90°$). The center panel $\angle w$ is green where $\angle s = 180°$. Thus $\angle w = \frac{1}{2}\angle s$. Going around the s plane one more time gives the right most figure. $w(s)$ is analytic everywhere except at the branch points $s = 0$ and $s = \infty$.

Moving the branch cut: Furthermore we can change $\angle s$ by $180°$ to move the branch cut

$$w = \rho e^{j\psi} = \sqrt{r} e^{j\theta/2} e^{j\pi k}, \tag{4.4.17}$$

where $\rho = |w|, \psi \angle w, \in \mathbb{R}$. The generic Cartesian coordinates are $s = \sigma + \omega j$ and $w(s) = u(\sigma, \omega) + v(\sigma, \omega) j$. For single-valued functions such as $w(s) = s^2$ on the left in Fig. 4.2 there is no branch cut, since $\psi = 2\theta$. Note how the red color ($\theta = 0°$) appears twice in this mapping. For multivalued functions, a branch cut is required, typically along the negative $v(\sigma, \omega)$ axis (i.e, $\psi = \pi$), but it may be freely distorted, as seen by comparing the right panel of Fig. 4.2 with the right panel of Fig. 4.3.

Properties of the branch cut: It is important to understand that the function is analytic on the branch cut but not at the branch points. One is free to move the branch cut (at will). It does not need to be on a line: it could be cut in almost any connected manner, such as a spiral. The only rule is that it must start and stop at the matching branch points, or at ∞, which must have the same degree.

The location of the branch cut may be moved by rotating the s coordinate system of Fig. 4.1. For example, $w(s) = \pm j\sqrt{s}$ and $w(s) = \pm\sqrt{-s}$ have different branch cuts, as may be verified using the Matlab/Octave commands `j*zviz(s)` and `zviz(-s)`. Every function is analytic on the branch cut (since moving it does not change the function). If a Taylor series is formed on the branch cut, it will describe the function on the two different sheets. Thus the complex–analytic series (i.e., the Taylor formula, Eq. 4.2.5) does not depend on the location of a branch cut, as it only describes the function uniquely (as a single-valued function), valid in its local region of convergence.

The second sheet ($k = 1$) in Fig. 4.3 picks up at $\theta = \pi$ [rads] and continues on to $\pi + 2\pi = 3\pi$. The first sheet maps the angle of w (i.e., $\phi = \angle w = \theta/2$) from $-\pi/2 < \phi < \pi/2$ ($w = \sqrt{r} e^{j\theta/2}$). This corresponds to $u = \Re\{w(s)\} > 0$. The second sheet maps $\pi/2 < \psi < 3\pi/2$ (i.e., 90° to 270°), which is $\Re\{w\} = u < 0$. In summary, twice around the s plane is once around the $w(s)$ plane because the angle is half due to the $\sqrt{s}$.

Branch cuts emanate and terminate at *branch points*, defined as singularities (poles) that can even have fractional degree, as for example $1/\sqrt{s}$, and terminate at one of the matching roots, which includes the possibility of ∞.[10] For example, suppose that in the neighborhood of the pole, at s_o the function is

$$f(s) = \frac{w(s)}{(s - s_o)^k}, \tag{4.4.18}$$

where $w, s, s_o \in \mathbb{C}$ and $k \in \mathbb{Q}$. When $k = 1$, $s_o = \sigma_o + \omega_o j$ is a first-degree "simple pole," having degree 1 in the s plane, with residue $w(s_o)$. Typically the order and degree are positive integers, but fractional degrees and orders are common in modern engineering applications (Kirchhoff, 1868; Lighthill, 1978). Here we allow both the degree and the order to be fractional ($\in \mathbb{F}$). When $k \in \mathbb{F} \subset \mathbb{R}$, $k = n/m$ is a real reduced fraction—namely, when `GCD(n,m)=1`,

[10] This presumes that poles and zeros appear in pairs, one of which may be at ∞.

Solution: Starting from Eq. H.1.2 with area $A(r) = A_o e^{2mr}$

$$\frac{1}{\cancel{A_o} e^{2mr}} \frac{\partial}{\partial r} \left(\cancel{A_o} e^{2mr} \frac{\partial \varrho}{\partial r} \right) = \frac{1}{c_o^2} \frac{\partial^2 \varrho}{\partial t^2}$$

$$\varrho_{rr}(r,t) + 2m\varrho_r(r,t) = \frac{1}{c_o^2} \frac{\partial^2 \varrho}{\partial t^2} \leftrightarrow \kappa^2 \mathcal{P}(r,\omega),$$

which is the time-domain version of Eq. H.4.6. ∎

Since this equation is second-order in time with constant coefficients, it has two closed-form solutions:

$$\begin{aligned} \mathcal{P}_c^{\pm}(r) &= \mathcal{P}_o^{\pm}(\omega)\, e^{-mr} e^{\mp\sqrt{m^2+\kappa^2}\, r} \\ &= \mathcal{P}_o^{\pm}(\omega)\, e^{-mr} e^{\mp j \frac{r}{c_o}\sqrt{\omega^2 - \omega_c^2}}, \end{aligned}$$

with $\omega_c = mc_o$. The two wave amplitudes $\mathcal{P}_0^{\pm}(\omega)$ must be determined from the boundary conditions.

Exercise #2

Shown that $\mathcal{P}^{\pm}(r,\omega)$ satisfy Eq. H.4.6.

Solution: Taking partials with respect to r,

$$\begin{aligned} \partial_r \mathcal{P}^{\pm}(r,\omega) &= \left(-m \mp \sqrt{m^2+\kappa^2} \right) \mathcal{P}^{\pm}(r,\omega) \\ \partial_{rr} \mathcal{P}^{\pm}(r,\omega) &= \left(-m \mp \sqrt{m^2+\kappa^2} \right)^2 \mathcal{P}^{\pm}(r,\omega) \\ &= \left(2m^2 + \kappa^2 \pm 2m\sqrt{m^2+\kappa^2} \right) \mathcal{P}^{\pm}(r,\omega). \end{aligned}$$

Thus Eq. H.4.6 reduces to

$$\left(2m^2 + \kappa^2 \pm 2m\sqrt{m^2+\kappa^2} \right) + 2m \left(-m \mp \sqrt{m^2+\kappa^2} \right) = \kappa^2,$$

which is an identity. ∎

Next consider the Fourier series (or Fourier transform) of the area function,

$$A(r) = \sum_k a_k e^{2m_k r}.$$

It follows from the linearity of the wave equation that the general solution of Eq. H.4.6 is

$$\mathcal{P}^{\pm}(r,\omega) = \sum_k a_k^{\pm}(\omega)\, e^{-m_k r}\, e^{\mp\sqrt{m_k^2+\kappa^2}\, r}.$$

Here we have combined $\mathcal{P}^{\pm}(\omega)$ and a_k as coefficients $a_k^{\pm}(\omega)$.

is

$$\imath\hbar\frac{\partial}{\partial t}\psi(x,t) = -\frac{\hbar^2}{2m_o}\nabla_r^2\psi(x,t) + V(r)\psi(x,t)$$
$$= -\frac{\hbar^2}{2m_o}\,\frac{1}{r^2}\frac{\partial}{\partial r}r^2\psi(x,t) + V(r)\;\psi(x,t) \tag{H.4.1}$$
$$= -\frac{\hbar^2}{2m_o}\left[\frac{2}{r}\frac{\partial}{\partial r}\psi(x,t) + \frac{\partial^2}{\partial r^2}\psi(x,t)\right] + V(r)\;\psi(x,t), \tag{H.4.2}$$

whereas the horn equation is given by Eq. H.1.2.

There are several obvious and disturbing differences between these two equations. First, the SE is, of course, first-order in time. Diffusion equations have no delay and thus cannot have traditional eigenmodes, which result from standing waves in a wave equation, due to boundary conditions. Second, the EM horn equation is of Sturm-Liouville (SL) form, which is a true wave equation (vs. the SE, which is a diffusion equation). The obvious question arises: Is there a potential V that would allow these two formulations to be equivalent? If so, then this would provide an explanation as to why the SE is successful in explaining the properties of Rydberg atoms.

To explore this possibility we may expand the two differential equations and directly compare them. Expanding Eq. H.1.2 gives

$$\frac{1}{c_o^2}\frac{\partial^2}{\partial t^2}\psi(r,t) = \frac{1}{A(r)}\frac{\partial}{\partial r}A(r)\frac{\partial}{\partial r}\psi(r,t) \tag{H.4.3}$$
$$= \frac{\partial^2}{\partial r^2}\psi(r,t) + \frac{1}{A(r)}\frac{\partial A(r)}{\partial r}\psi(r,t). \tag{H.4.4}$$

Between these two equations we may remove ψ'':

$$\imath\hbar\frac{\partial}{\partial t}\psi(x,t) = -\frac{\hbar^2}{2m_o}\left[\frac{2}{r}\frac{\partial}{\partial r}\psi(x,t) + \frac{1}{c_o^2}\frac{\partial^2}{\partial t^2}\psi(r,t) - \frac{1}{A(r)}\frac{\partial A(r)}{\partial r}\psi(r,t)\right] + V(r)\;\psi(x,t). \tag{H.4.5}$$

It seems that this may isolate the time and spatial parts (as in separation of variables).

H.4.1 The exponential horn

A relevant and motivational example is the solution of the exponential horn, having area function $A(r) = A_o e^{2mr}$. This case is interesting because it has a closed-form solution, which seems relevant and perhaps even related to the hydrogen atom.

Assuming that $\varrho(r,t) \leftrightarrow \mathcal{P}(r,\omega)$ are a Fourier transform pair, with $A(r) = A_o e^{2mr}$, Eq. H.1.2 reduces to

$$\frac{\partial^2\mathcal{P}(r,\omega)}{\partial r^2} + 2m\frac{\partial\mathcal{P}(r,\omega)}{\partial r} = \kappa^2\mathcal{P}(r,\omega) \leftrightarrow \frac{1}{c_o^2}\frac{\partial^2\varrho}{\partial t^2}, \tag{H.4.6}$$

with $\kappa(s) = s/c_o$.

Exercise #*1*
Show that Eq. H.4.6 follows from Eq. H.1.2.

$n \perp m$). This defines the *degree* of a fractional pole. In such cases there must be two sets of branch cuts of degrees n and m. For example, if $k = 1/2$, the singularity (branch cut) is of degree 1/2 and there are two Riemann sheets, as shown in Fig. 4.2.

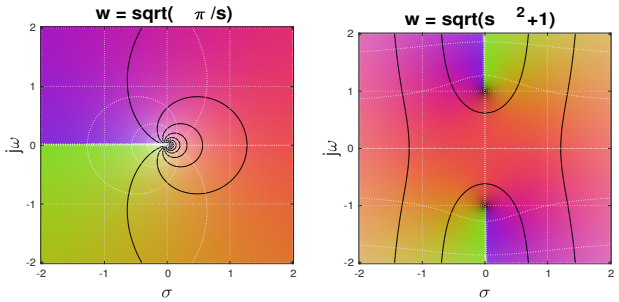


Figure 4.4: *Colorized plots of two $\mathcal{LT}$ pairs: Left:* $\sqrt{\pi/s} \leftrightarrow u(t)/\sqrt{t}$. *Right:* $\sqrt{s^2+1} \leftrightarrow \delta(t) + \frac{1}{t}J_1(t)u(t)$.

Fractional-order Bessel function: An important example is the Bessel function and its Laplace transform ($\mathcal{LT}$)

$$\delta(t) + \frac{1}{t}J_1(t)u(t) \leftrightarrow \sqrt{s^2+1},$$

as shown in Fig. 4.4, which is related to the solution to the wave equation in two-dimensional cylindrical coordinates (see Table C.2, p. 302). Bessel functions are the solutions (i.e., eigenfunctions) of guided acoustic waves in round pipes, or surface waves on the earth (seismic waves), or waves on the surface of a pond (Table 5.2, p. 247).

There are a limited number of possibilities for the degree, $k \in \mathbb{Z}$ or $\in \mathbb{F}$ of Eq. 4.4.18. If the degree is drawn from $\mathbb{R} \notin \mathbb{F}$, the pole can not have a residue. According to the definition of the residue, $k \in \mathbb{F}$ has no residue. But there remains open the possibility of generalizing the concept of the Riemann integral theorem to include $k \in \mathbb{F}$. One way to do this is to use the logarithmic derivative, which transforms fractional poles to simple poles with fractional residues. (See Ex.# 9, page 84).

If the singularity has an irrational degree ($k \in \mathbb{I}$), the branch point has the same irrational degree. Accordingly there are an infinite number of Riemann sheets, as in the case of the log function. An example is $k = \pi$, for which

$$F(s) = \frac{1}{s^\pi} = e^{-\log(s^\pi)} = e^{-\pi\log(s)} = e^{-\pi\log(\rho)}e^{-\pi\theta_J},$$

where the domain is expressed in polar coordinates $s = \rho e^{\theta_J}$. When $k \in \mathbb{F}$, it may be close (e.g., $k = \pi_{152}/\pi_{153} = \pi_{152}/(\pi_{152}+2) = 881/883 \approx 0.99883$, or its reciprocal ≈ 1.0023). The branch cut could be subtle (it could even go unnoticed), but it would have a significant impact on the function and on its inverse Laplace transform. More examples?

Exercise #*2*
Find the poles, zeros, and residues of $F(s)$.

1. $$F(s) = \frac{d}{ds} \ln \frac{s+e}{s+\pi}$$

   Solution:
   $$F(s) = \frac{d}{ds} \left[\ln(s+e) - \ln(s+\pi)\right] = \left(\frac{1}{s+e} - \frac{1}{s+\pi}\right)$$
   The poles are at $s_1 = -e$ and $s_2 = -\pi$ with respective residues of ± 1. ∎

2. $$F(s) = \frac{d}{ds} \ln \frac{(s+3)^e}{(s+\jmath)^{-\pi}}$$

   Solution:
   $$F(s) = \frac{d}{ds} \left(e \ln(s+3) + \pi \ln(s+\jmath)\right) = \frac{e}{s+3} + \frac{\pi}{s+\jmath}.$$
   There is a very important take-home message here regarding the utility of the logarithmic derivative, which "linearizes" the fractional pole. ∎

3. $$F(s) = e^{\pi \ln s}$$

   Solution: To simplify this expression take the log
   $$\ln F(s) = \ln e^{\pi \ln s} = \pi \ln s = \ln s^{\pi}$$
   Thus $F(s) = s^{\pi}$. The only pole is $s \to \infty$. Thus the definition of the residue is to multiply by the pole and take the limit as $s \to \infty$
   $$c_{-1} = \lim_{s \to \infty} \frac{s^{\pi}}{s} = 0^{\pi - 1} = 0.$$
   ∎

4. $$F(s) = \pi^{-s}$$

   Solution: Converting to exponential format: $F(s) = e^{-s \ln \pi} \leftrightarrow \delta(t - \ln \pi)$. I don't think the pure time delay has poles. I'm not sure what this tells us about the residue.∎

Log function: Next we discuss the multivalued nature of the log function. In this case there are an infinite number of Riemann sheets, not captured by Fig. 3.14 (p. 168), which displays only the principal sheet. However, if we look at the formula for the log function, the nature is easily discerned. The abscissa s may be defined as multivalued, since

$$s_k = re^{2\pi k \jmath} e^{\theta \jmath}.$$

Here we have extended the angle of s by $2\pi k$, where k is the sheet index $\in \mathbb{Z}$. Now we take the log:

$$\log(s) = \log(r) + (\theta + 2\pi k)\jmath.$$

When $k = 0$, we have the principal value sheet, which is zero when $s = 1$. For any other value of k, $w(s) \neq 0$, even when $r = 1$, since the angle is not zero, except for the $k = 0$ sheet.

H.3.1 Eigenmodes of the Rydberg atom

One way to think of eigenmodes is to make an analogy to a piano string or an organ pipe. In these much simpler systems, there is an almost constant delay, say τ, due to a characteristic length, say $L = \tau c_o$, such that the eigenmodes of a string are given by integer multiples of a half wavelength $\nu_n = nc_o/2L$, while the eigenmodes of the organ pipe are multiples of a quarter wavelength. The distinction is the boundary conditions. For the string the endpoint boundary conditions are pinned displacement (i.e., zero velocity). The organ pipe is closed at one end and open at the other, resulting in multiples of a quarter wavelength $\nu_n = nc_o/4L$. In each case $\nu = n/\tau$, where $\tau = 2L/c_o$ is the round-trip delay; thus $\nu = nc_o/2L$. We suggest looking at the Rydberg formula in the same way, but with very different eigenfrequencies (Eq. H.3.1). Sommerfeld (1949, p. 201) makes a very interesting comment regarding Eq. H.3.1:

> This equation reduces to a simple mathematical formula the enigma of the spectral lines, with their finite cumulation point, the behavior of which differs so fundamentally from that of all mechanical systems.

H.3.2 Discussion

The Rydberg frequencies f_{nl} ($n = 1, l = 1, \ldots, \infty$) has poles in the radiation impedance (Eq. H.1.3) when $\phi_l(f_{nl}) \in \mathbb{N}$. Working backwards from the Rydberg formula (Eq. H.2.1), we have solved for $\phi(f_{nl})$ indicating where this condition is valid (Eq. H.2.3). Since the reflectance and the impedance must be causal complex–analytic functions of Laplace frequency s, we must replace the discrete frequency f_{nl} with s:

$$\jmath 2\pi f_{nl} \to s = \sigma + \omega\jmath,$$

thereby forcing $l(s)$ to be a complex–analytic function of s. Then the poles of the radiation impedance must satisfy

$$\Gamma(s_{nl}) = e^{\jmath 2\pi l(f_{nl})} = 1,$$

resulting in eigenfrequencies at f_{nl}.

The next step in this analysis is to determine the area function $A(r)$ given Z_{rad} (Eq. H.1.3). To do this we must solve an integral equation for $A(r)$, as discussed by Sondhi and Gopinath (1971) and by Youla (1964).

Perhaps this could be done using an analytic representation for the area function,

$$A(r) = \sum_k a_k r^k.$$

H.4 Relations between Sturm-Liouville and quantum mechanics

If we compare the Schrödinger equation (SE) for hydrogen with the corresponding Sturm-Liouville equation we can begin to appreciate their differences. The QM equation for hydrogen

where $\nu_{n,m}$ are the possible eigenfrequencies, c_o is the speed of light, $R_\infty \approx 1.097 \times 10^7$ is the Rydberg constant, Z_n is the atomic number, along with positive integers $m > n \in \mathbb{N}$, which represent the *principal quantum numbers* that label all possible allowed atomic eigenstates. Integer n indicates the lowest (rest) atomic eigenstate while m labels the higher (excited) state.[4] When $n = 1$, the series is the Lyman series corresponding to hydrogen ($Z_1 = 1$). When $n = 1$, $m = 2$, and $Z_1 = 1$, the frequency is

$$\nu_{1,1} = c_o R_\infty \cancel{\left(\frac{1}{1^2} - \frac{1}{2^2}\right)^{\frac{3}{4}}} = 2.5 \times 10^{15} \text{ [Hz]}, \tag{H.3.2}$$

and the wavelength is $\lambda = c_o/\nu = 4 \times 10^8 / R_\infty = 36.36$ [m].

An open question in this model is: *Why are states either empty or filled?* The amplitudes of the modes of a string or organ pipe are not quantized. What is it about the atom that forces the energy to be quantized?

Given observed frequencies $\nu_{n,m}$ it is possible to find the area function that traps the photons into the Rydberg eigenstates. Eq. H.3.1 may be rewritten as

$$\nu_{n,m} = c_o R Z_n^2 4 \left(\frac{m^2 - n^2}{(2nm)^2}\right).$$

It is interesting to compare Eq. H.3.1 to Euclid's formula Eq. 2.5.1 (p. 61):

$$a = m^2 - n^2, \qquad b = 2mn, \qquad c = m^2 + n^2, \tag{H.3.3}$$

where $m > n \in \mathbb{N}$. Euclid's formula is equivalent to the Pythagorean theorem for integers, since

$$c^2 = a^2 + b^2, \tag{H.3.4}$$

with $\{a, b, c\} \in \mathbb{N}$. Here $a < b < c$.

If we interpret the quantum numbers as multiples of a quarter wavelength, then the Rydberg formula is congruent to the Pythagorean theorem. Given the symmetry, this cannot be an accident.

In terms of the lengths of the right triangle $\{a, b, c\}$, Rydberg's formula becomes

$$\nu_{n,m} = c_o R Z_n^2 4 \left(\frac{a}{b^2}\right).$$

But since $b^2 = c^2 - a^2$,

$$\begin{aligned}\nu_{n,m} &= c_o \frac{R Z_n^2}{a} 4 \left(\frac{a^2}{c^2 - a^2}\right) \\ &= c_o \frac{R Z_n^2}{a} 4 \frac{a^2}{c^2} \left(\frac{1}{1 - (a/c)^2}\right).\end{aligned}$$

In terms of quantized (discrete) angles, $\sin(\theta_{n,m}) = a/c$,

$$\begin{aligned}\nu_{n,m} &= c_o \frac{R Z_n^2}{a} 4 \left(\frac{\sin^2\theta}{1 - \sin^2\theta}\right) \\ &= c_o \frac{R Z_n^2}{a} 4 \left(\frac{\sin^2\theta}{\cos^2\theta}\right) \\ &= c_o \frac{R Z_n^2}{a} 4 \tan^2\theta_{n,m}.\end{aligned}$$

[4] http://en.wikipedia.org/w/index.php?title=Rydberg_formula

4.5 Three Cauchy integral theorems

4.5.1 Cauchy's theorems for integration in the complex plane

There are three basic definitions related to Cauchy's integral formula. They are closely related and can greatly simplify integration in the complex plane. The choice of names is unfortunate, if not totally confusing. Hence I call them CT-1, CT-2 and CT-3.

1. **Cauchy's (integral) theorem (CT-1):**

$$\oint_{\mathcal{C}} F(s)ds = 0 \tag{4.5.1}$$

   if and only if $F(s)$ is complex–analytic inside of a simple closed curve $\mathcal{C}$ (Stillwell, 2010; Boas, 1987, p. 45). The FTCC (Eq. 4.2.3) says that the integral depends on only the end points if $F(s)$ is complex–analytic. With the path (contour $\mathcal{C}$) closed, the end points are the same and thus the integral must be zero as long as $F(s)$ is complex–analytic.

2. **Cauchy's integral formula (CT-2):**

$$\frac{1}{2\pi j} \oint_{\mathcal{B}} \frac{F(s)}{s - s_o} ds = \begin{matrix} F(s_o), s_o \in \mathbb{C} < \mathcal{B} \text{ (inside)} \\ 0 \quad , s_o \in \mathbb{C} > \mathcal{B} \text{ (outside).} \end{matrix} \tag{4.5.2}$$

   Here $F(s)$ is required to be analytic everywhere within (and on) the boundary $\mathcal{B}$ of integration (Greenberg, 1988, p. 1200); (Boas, 1987, p. 51); (Stillwell, 2010, p. 220). When the point $s_o \in \mathbb{C}$ is within the boundary, the value $F(s_o) \in \mathbb{C}$ is the residue of the pole s_o of $F(s)/(s - s_o)$. When the point s_o lies outside the boundary, the integral is zero.

3. **The (Cauchy) residue theorem (CT-3):** (Greenberg, 1988, p. 1241); (Boas, 1987, p. 73)

$$\oint_{\mathcal{C}} f(s)ds = 2\pi j \sum_{k=1}^{K} c_k = \sum_{k=1}^{K} \frac{F(s)}{s - s_k} ds, \tag{4.5.3}$$

   where the residues $c_k \in \mathbb{C}$ correspond to the kth pole of $f(s)$ enclosed by the contour $\mathcal{C}$ [(Greenberg, 1988, p. 1241); (Boas, 1987, p. 73)]. Cauchy's integral formula (CT-2) is a special case of the residue theorem (CT-3).

How to calculate the residue: The case of first-degree poles has special significance because the Brune impedance allows only simple poles and zeros, thus increasing its utility. The residues for simple poles are $F(s_k)$, which is complex–analytic in the neighborhood of the pole, but not at the pole.

Consider the function $f(s) = F(s)/(s - s_k)$, where we have factored $f(s)$ to isolate the first-order pole at $s = s_k$, with $F(s)$ analytic at s_k. Then the residue of the poles at $c_k = F(s_k)$. This coefficient is computed by removing the singularity, placing a zero at the pole frequency, and taking the limit as $s \to s_k$—namely,

$$c_k = \lim_{s \to s_k} [(s - s_k) f(s)] \tag{4.5.4}$$

(Greenberg, 1988; Boas, 1987, p. 72).

When the pole is an Nth degree, the procedure is much more complicated and requires taking $N-1$ order derivatives of $f(s)$ followed by the limit process (Greenberg, 1988, p. 1242). Higher-degree poles are rarely encountered; thus it is good to know that this formula exists, but perhaps it is not worth the effort to learn (i.e., memorize) it.

Some open questions: Without the use of CT-3 it is hard to evaluate the inverse Laplace transform of $1/s$ directly. For example, how do we show that the integral (Eq. 4.5.2) is zero for negative time (or 1 for positive time)? CT-3 resolves this difficult problem by the convergence of the integral for negative and positive times. Clearly the convergence of the integral at $\omega \to \infty$ plays an important role.

4.5.2 Cauchy Integral Formula and Residue Theorem

CT-2 is an important extension of CT-1, in that a pole has been explicitly represented in the integrand at $s = s_o$. If the pole location is outside the curve $\mathcal{C}$, the result of the integral is zero, in keeping with CT-1. When the pole is inside $\mathcal{C}$, the integrand is no longer complex–analytic at the enclosed pole. When this pole is simple, the residue theorem applies. For the related CT-3 the same result holds, except it is assumed that there are K simple poles in the function $F(s)$. This requires K repeated applications of CT-2. Thus it represents a minor extension of CT-2. When the integrand is $f(s)/P_K(s)$ where $f(s)$ is analytic in $\mathcal{C}$ and $P_K(s)$ is a polynomial of degree K, with all of its roots $s_k \in \mathcal{C}$, then CT-3 applies.

Nonintegral degree singularities: A key point is that this theorem applies when $k \in \mathbb{I}$, including fractionals $k \in \mathbb{F}$, but for these cases the residue is always zero, since by definition the residue is the amplitude of the $1/s$ term (Boas, 1987, p. 73). Below are examples:

1. When $k \in \mathbb{F}$ (e.g., $k = 2/3$), the residue of s^k is zero, by definition.
2. The function $1/\sqrt{s}$ has a zero residue (we apply the definition of the residue, Eq. 4.5.4).
3. When $k \neq 1 \in \mathbb{I}$, the residue is, by definition, zero.
4. When $k = 1$, the residue is given by Eq. 4.5.4.
5. CT-1, CT-2, and CT-3 are essential when computing the inverse Laplace transform.

Summary and examples: These three CT theorems, all attributed to Cauchy, collectively are related to the fundamental theorems of complex calculus. The general principles are:

1. In general it makes no sense (nor is there any need) to integrate *through* a pole; thus the poles (or other singularities) must not lie on $\mathcal{C}$.
2. CT-1 (Eq. 4.5.1) follows trivially from the fundamental theorem of complex calculus (Eq. 4.2.3, p. 180), since if the integral is independent of the path and the path returns to the starting point, the closed integral must be zero. Thus Eq. 4.5.1 holds when $F(s)$ is complex–analytic within C.
3. Since the real and imaginary parts of every complex–analytic function obey Laplace's equation (Eq. 4.2.8, p. 182), it follows that every closed integral over a Laplace field—that is, those defined by Laplace's equation—must be zero. In fact, this is the property of a conservative system, corresponding to many physical systems. If a closed box has fixed potentials on the walls, with any distribution whatsoever, and a point charge (i.e., an electron) is placed in the box, then a force equal to $F = qE$ is required to move that charge, and thus work is done. However, if the point is returned to its starting location, the net work done is zero.

Typically one uses the reflectance phase $2\pi\phi(f)$; thus the group delay is $\tau(f) = -\partial\phi(f)/\partial f$, which is physically interpreted here as the frequency-dependent delay from the proton to the radius of the electron's orbit. Thus this delay is given by

$$\tau(f) = n\frac{\partial}{\partial f}\left(1 - \frac{n^2}{c_o R}\, f\right)^{-1/2} = \frac{n^3}{2c_o R}\left(1 - \frac{n^2}{c_o R}\, f\right)^{-3/2},$$

which is constant for low frequencies and then rises to ∞ as frequency approaches the Rydberg frequency ($f \to c_o R/n^2$).

One may solve Eq. H.1.1 for m, for the case of the Lyman series ($n_o = 1$), by the use of the following identity for the Rydberg eigenfrequencies f_{nm}, which follow directly from Eq. H.1.1, with $m = n_o + l$ (with $n_o, m, l \in \mathbb{N}$)

$$\begin{aligned} f_{nm} = \frac{c_o}{\lambda_{nm}} &= c_o R\left(\frac{1}{n_o^2} - \frac{1}{(n_o + l)^2}\right) \\ &= \frac{c_o R}{n_o^2}\left(1 - \frac{1}{(1 + l/n_o)^2}\right). \end{aligned} \tag{H.2.2}$$

Note that as $l \to \infty$, $f_{n_o,l} \to = c_o R/n_o^2$, which is Sommerfeld's "finite cumulation point" [Hz] $f_{n_o,\infty}$ for the Lyman series ($n_o = 1$).

We can solve Eq. H.2.2 for the mode number $l/n < 1$ as a function of mode frequency:

$$\begin{aligned} n^2\frac{f_{nl}}{c_o R} &= 1 - \frac{1}{(1 + l/n)^2} && \text{Starting from Eq. H.2.2} \\ \frac{1}{(1 + l/n)^2} &= 1 - n^2\frac{f_{nl}}{c_o R} && \text{Solving for } l/n \\ (1 + l/n)^2 &= \frac{1}{1 - n^2\frac{f_{nl}}{c_o R}} \\ \frac{l}{n} &= \pm\frac{1}{\sqrt{1 - n^2\frac{f_{nl}}{c_o R}}} - 1 && \phi(f_{nl})/2\pi = l = m - n_o \in \mathbb{N}, \end{aligned} \tag{H.2.3}$$

as summarized in the lower panel of Fig. H.4.

H.2.2 Finding the area function

Once the phase has been determined, we can compute the impedance using Eq. H.1.3. We may then decompose the impedance by using the analytic continued fraction algorithm (or Cauer synthesis), discussed in Sec. 3.8, p. 144.

H.3 Euclid's formula and the Rydberg atom model

Fundamental to quantum mechanics is the Rydberg formula, which describes the quantized energy levels of atoms:[3]

$$\nu_{n,m} = c_o R_\infty Z_n^2\left(\frac{1}{n^2} - \frac{1}{m^2}\right), \tag{H.3.1}$$

[3] https://www.youtube.com/watch?v=e0IWPEhmMho

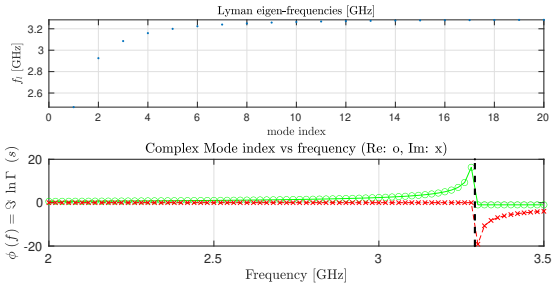


Figure H.4: *The top panel is a plot of Eq. H.1.1, showing how the eigenmode frequencies f_l depend on the eigen-number index l. As the mode number increases, the frequency reaches an asymptote at $c_o R \approx 3.29$ [GHz], with a wavelength limit near $1/R \approx 91.2$ [nm]. The lower panel shows the inverse mapping from frequency to the mode index number $\phi(f)$. This figure is for the Lyman series ($n = 1$ and $m = 1, \ldots, 20$). The inverse of this relationship is $l = \phi^{-1}(f_l)$ may be derived from Eq. H.1.1, which provides the pole frequencies required to satisfy Eq. H.1.3. Note that for frequencies greater than c_o/R the phase switches from purely real to imaginary, accounting for the free electrons that must exist above the upper accumulation frequency (i.e., 3.29 [GHz] for this example).*

H.2 Rydberg solution methods

The basic idea behind the method is to use Eq. H.1.3, by noting that the poles of the impedance are determined by the roots of the denominator of Z_{rad}. Specifically, if s_p is an impedance pole, then it must satisfy $\Gamma(s_p) \approx 1$. Except for losses due to radiation, the atom is lossless; thus $|\Gamma(s)| = 1$. Namely, it must be of the form

$$\Gamma(s) = e^{-\jmath\phi(f)}, \tag{H.2.1}$$

where the *phase* $\phi(f) \in \mathbb{R}$ and $s = \sigma + \omega\jmath$ is the complex Laplace radian frequency, with $\omega = 2\pi f$ [Hz]. Since we know the eigenmode frequencies, which obey $\phi(f_{n_o,m}) = 2\pi m$, we may find $\phi(f)$, as follows: For a given series index n_o, and given the eigenfrequencies f_m, we seek the phase mode function $\phi_{n_o}(f)$ that maps the eigenfrequencies to their mode index m, i.e.,

$$\phi_{n_o}(f_m) = 2\pi m.$$

H.2.1 Group delay $\tau(s)$

The phase $\phi(\omega)$ is related to the group delay $\tau(\omega)$ by the relation

$$\tau(\omega) = -\frac{\partial}{\partial\omega}\phi(\omega).$$

Here one may assume that the phase is complex–analytic,[2] thus allowing a causal damping term into the reflectance phase Eq. H.2.1. This follows naturally because the reflectance must be causal (Postulate 3.10.2, p. 162). In the time domain the delay may be written in terms of the inverse $\mathcal{LT}$ of the group delay,

$$\Gamma(s) = e^{-\jmath \int_o^s \tau(s)ds}.$$

[2] It follows that these relationships are related by a Hilbert transform.

4. Work is done in charging a capacitor, and energy is stored. However, when the capacitor is discharged, all of the energy is returned to the load.

5. Soap bubbles and rubber sheets on a wire frame obey Laplace's equation.

6. These are all cases where the fields are Laplacian, thus closed line integrals must be zero. Laplacian fields are commonly seen because they are so basic.

7. We have presented the impedance as the primary example of a complex–analytic function. Physically, every impedance has an associated stored energy, and every system having stored energy has an associated impedance. This impedance is usually defined in the frequency s domain, as a force *over* a flow (i.e., voltage over current). The power $\mathcal{P}(t)$ is defined as the force *times* the flow and the energy $\mathcal{E}(t)$ as the time integral of the power

$$\mathcal{E}(t) = \int_{-\infty}^{t} \mathcal{P}(t)dt, \tag{4.5.5}$$

which is similar to Eq. 4.2.1 (p. 179) [see Sec. 3.8.3, Eq. 3.8.9 (p. 148)]. In summary, impedance and power and energy are all fundamentally related.

4.6 Problems DE-2

4.6.1 Topics of this homework:

Integration of complex functions, Cauchy's theorem, integral formula, residue theorem, power series, Riemann sheets and branch cuts, inverse Laplace transforms, Quadratic forms.

Two fundamental theorems of calculus

Fundamental Theorem of Calculus (Leibniz):

According to the fundamental theorem of (real) calculus (FTC),

$$f(x) = f(a) + \int_a^x F(\xi)d\xi, \tag{DE-2.1}$$

where $x, a, \xi, F, f \in \mathbb{R}$. This is an indefinite integral (since the upper limit is unspecified). It follows that

$$\frac{df(x)}{dx} = \frac{d}{dx}\int_a^x F(x)dx = F(x).$$

This justifies also calling the indefinite integral the *antiderivative*.

For a closed interval $[a, b]$, the FTC is

$$\int_a^b F(x)dx = f(b) - f(a), \tag{DE-2.2}$$

thus the integral is independent of the path from $x = a$ to $x = b$.

Fundamental Theorem of Complex Calculus:

According to the fundamental theorem of complex calculus (FTCC),

$$f(z) = f(z_0) + \int_{z_0}^z F(\zeta)d\zeta, \tag{DE-2.3}$$

where $z_0, z, \zeta, f, F \in \mathbb{C}$. It follows that

$$\frac{df(z)}{dz} = \frac{d}{dz}\int_{z_0}^z F(\zeta)d\zeta = F(z). \tag{DE-2.4}$$

For a closed interval $[s, s_o]$, the FTCC is

$$\int_{s_o}^s F(\zeta)d\zeta = f(s) - f(s_o), \tag{DE-2.5}$$

umber thus the integral is independent of the path from $x = a$ to $x = b$.

mechanics, the history of which is nicely summarized in Condon and Morse (1929).

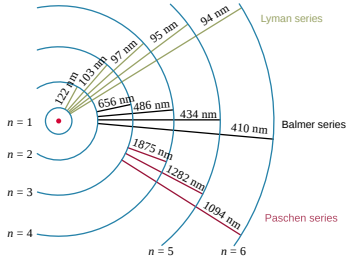


Figure H.3: *This diagram defines hydrogen's allowed electron transitions, defining the Lyman ($n = 1$), Balmer ($n = 2$), and Paschen ($n = 3$) series. The numbers represent the wavelengths λ [nm] of the photons having frequencies $f_{nm} = c_o/\lambda_{nm}$, following an electron transition from level n to m (taken from:* `https://en.wikipedia.org/wiki/Hydrogen_spectral_series`*).*

It was clear from the days of Bohr that the Rydberg formula did not follow the typical rules of eigenspectra, so much so that Arnold Sommerfeld wrote (Sommerfeld, 1949, p. 201):

> The lines of this spectrum cumulate at the limit given by the Rydberg constant R. The adjoining *continuum* lies in the near ultraviolet range. Both the discrete and the continuous spectrum are given by the Schrödinger equation. This equation reduces to a simple mathematical formula the enigma of the spectral lines, with their finite cumulation point, the behavior of which differs so fundamentally from that of all mechanical systems.

H.1.2 Rydberg wave equation

The objective of this analysis is to demonstrate that one can define a classical Sturm-Liouville model of the *enigmatic* Rydberg atom, by the use of the Webster horn equation

$$\frac{1}{A(r)}\frac{\partial}{\partial r}A(r)\frac{\partial}{\partial r}\psi(r,t) = \frac{1}{c_o^2}\frac{\partial^2}{\partial t^2}\psi(r,t), \tag{H.1.2}$$

which is a one-dimensional wave equation for the electric potential $\psi(r,t)$ propagating in a wave guide having area $A(r)$ as a function of the range, where r is the *range* variable (the axis of wave propagation).

We shall show that given the Rydberg spectrum (Eq. H.1.1), we may accurately estimate the electric reflectance $\Gamma(s)$ looking out from the origin (i.e., the proton location, as indicated by the small red dot in Fig. H.3). The radiation impedance $Z_{\text{rad}}(s)$ seen by the proton is related to the reflectance $\Gamma(s)$ by the relation

$$Z_{\text{rad}}(s) = r_o\frac{1+\Gamma(s)}{1-\Gamma(s)}. \tag{H.1.3}$$

This formula is the basis of the *Smith chart* used in both physics and engineering studies. It follows that once $\Gamma(s)$ is known (i.e., evaluated given Eq. H.1.1), the radiation impedance may be computed. It has been shown that the area function $A(r)$ may be found given the radiation impedance (Sondhi and Gopinath, 1971; Youla, 1964).

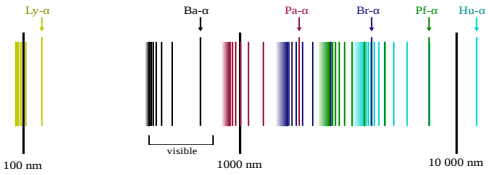


Figure H.1: *Diagram of the wavelength spectrum of hydrogen for the Lyman, Balmer, and Paschen series, as a function of each line's wavelength. The notation "Ly-α" indicates the longest wavelength $\lambda_{11} = 122$ [nm] (i.e., lowest frequency of 2.46 [GHz]) for the Lyman series. Figure citation:* `https://en.wikipedia.org/wiki/Hydrogen_spectral_series`

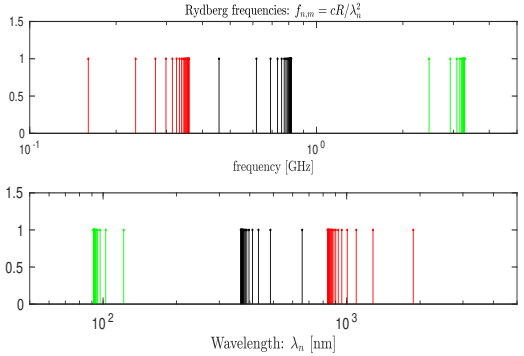


Figure H.2: *Rydberg frequencies in [GHz] and the corresponding wavelengths, computed from the Rydberg formula $\lambda_{nm}^{-1} = R\,(\frac{1}{n^2} - \frac{1}{m^2})$, where integer n defines the series (Lyman: $n = 1$, Balmer: $n = 2$, Paschen: $n = 3$, etc.) and integer $m > n$ defines the outer transition line (see Fig. H.3). For example, according to the lower panel (green series), the Lyman series line $\lambda_{1,2} = 122$ [nm] ($n = 1$ and $m = 2$), in agreement with the lower panel of this figure, Figs. H.3 and H.1. The frequency of the Paschen series line (3,6) is at 1.094 [μm] (0.3 [GHz]) (upper panel)* `(http://www.physics.drexel.edu/%7Etim/open/hydrofin/).`

all based on these simple observations. Here $R_\infty = 1.097 \times 10^7$ [m^{-1}]; $c_o = 3 \times 10^8$ [m/s] is the speed of light; f_{nm} are the dimensionless *Rydberg integer frequencies*; and $n, m \in \mathbb{N}$ are positive integers $\in \mathbb{N}$, where n labels the series and $m > n$ ($\lambda > 0$) describes the transition from orbit m to orbit n, as described in the caption of Fig. H.3.

H.1.1 The Rydberg atom model

In 1909 Rutherford demonstrated that the atom consisted of a dense core (the proton) surrounded by electrons. This view was supported by the spectrum of the atom, which allows for a radiation spectrum caused by electrons jumping from one energy level to another. It was then noted by Bohr in 1913 (Bohr, 1954) and others that the wavelengths of hydrogen, as described by Eq. H.1.1, are consistent with Fig. H.2, where the reciprocal wavelength [m^{-1}] is given by Eq. H.1.1, having frequencies $f_{nm} = c/\lambda_{nm}$ [Hz]. The challenge of the 1920s was to explain these intuitive and simple models of hydrogen. This gave rise to the birth of quantum

Problem # *1*

– 1.1: Consider Equation DE-2.1. What is the condition on $F(x)$ for which this formula is true?

– 1.2: Consider Equation DE-2.3. What is the condition on $F(z)$ for which this formula is true?

– 1.3: Let $F(z) = \sum_{k=0}^{\infty} c_k z^k$.

– 1.4: Let

$$F(z) = \frac{\sum_{k=0}^{\infty} c_k z^k}{z - \jmath}.$$

Problem # *2: In the following problems, solve the integral*

$$I = \oint_{\mathcal{C}} F(z)dz$$

for a given path $\mathcal{C} \in \mathbb{C}$.

– 2.1: Perform the following integrals ($z = x + iy \in \mathbb{C}$): $I = \int_0^{1+\jmath} z dz$

– 2.2: $I = \int_0^{1+j} z dz$, but this time make the path explicit: from 0 to 1, with $y = 0$, and then to $y = 1$, with $x = 1$.

– 2.3: Discuss whether your results agree with Eq. DE-2.4?

Problem # *3: Perform the following integrals on the closed path $\mathcal{C}$, which we define to be the unit circle. You should substitute $z = e^{i\theta}$ and $dz = ie^{i\theta}d\theta$, and integrate from $\{-\pi, \pi\}$ to go once around the unit circle.*

Discuss whether your results agree with Eq. DE-2.4?

– 3.1: $\oint_{\mathcal{C}} z dz$

– 3.2: $\oint_{\mathcal{C}} \frac{1}{z} dz$

– 3.3: $\oint_{\mathcal{C}} \frac{1}{z^2} dz$

– 3.4: $I = \int_{\mathcal{C}} \frac{1}{(z+2j)^2} dz$.

Recall that the path of integration is the unit circle, starting and ending at -1.

Problem # *4: FTCC and integration in the complex plane*
Let the function $F(z) = c^z$, where $c \in \mathbb{C}$ is given for each question. *Hint: Can you apply the FTCC?*

– 4.1: For the function $f(z) = c^z$*, where* $c \in \mathbb{C}$ *is an arbitrary complex constant, use the Cauchy-Riemann (CR) equations to show that* $f(z)$ *is analytic for all* $z \in \mathbb{C}$.

– 4.2: Find the antiderivative of $F(z)$.

– 4.3: $c = 1/e = 1/2.7183, \ldots$ *where* $\mathcal{C}$ *is* $\zeta = 0 \to i \to z$

– 4.4: $c = 2$*, where* $\mathcal{C}$ *is* $\zeta = 0 \to (1+i) \to z$

– 4.5: $c = i$*, where the path* $\mathcal{C}$ *is an inward spiral described by* $z(t) = 0.99^t e^{i2\pi t}$ *for* $t = 0 \to t_0 \to \infty$

– 4.6: $c = e^{t-\tau_0}$*, where* $\tau_0 > 0$ *is a real number and* $\mathcal{C}$ *is* $z = (1-i\infty) \to (1+i\infty)$.
Hint: Do you recognize this integral? If you do not, please do not spend a lot of time trying to solve it via the "brute force" method.

4.6.2 Cauchy's theorems CT-1, CT-2, CT-3

There are three basic definitions related to Cauchy's integral formula. They are all related and can greatly simplify integration in the complex plane. When a function depends on a complex variable, we use uppercase notation, consistent with the engineering literature for the Laplace transform.

Problem # *5: Describe the relationships between the theorems:*

– 5.1: CT-1 and CT-2

– 5.2: CT-1 and CT-3

– 5.3: CT-2 and CT-3

Appendix H

Quantum Mechanics and the WHEN

While it is clear that both Schrödinger's equation and Dirac's equations are highly accurate, after about 100 years, it is not clear why. Both of these theories seem to violate classical electromagnetics (EM), such as Ohm's law, since they are built on energy principles rather than electric and magnetic fields. One point I find disturbing is that QM defines a probabilistic wave function, and appears to not be causal. These two properties are in conflict. How can the past be probabilistic, and how can the future not be?

Here we delve into this question, by providing a classical (i.e., EM-based) derivation for the hydrogen atom, one of the most important and obvious successes of quantum mechanics (QM). The problem with QM is not that it fails—rather, it succeeds, without obvious basis. The problem is that we cannot understand the basic principles, and it seems to be in contradiction with any principles of a physical theory.

Based on the Rydberg formula, we determine the reflection coefficient, and thus the radiation impedance seen by the electron, in a radial coordinate system centered on the proton. Since the electron and proton both have spin $\frac{1}{2}$, their magnetic fields must attractively align, accounting for the near-field vector potential, and complementing the far-field attraction due to their opposite signs. As the electron and proton approach each other, due to their far-field potential attraction, the magnetic near field becomes more attractive at close range, due to the magnetic dipoles of the two "particles," causing them to merge with neutral net magnetic moment and neutral charge, giving a highly stable hydrogen atom. However, given a sufficiently strong distorting field, this highly symmetric state could be disturbed, leading to photon radiation, constrained by the radial eigenstates. It seems more clear than ever that photons and electrons are in a state of equilibrium at the outskirts of very large Rydberg atoms.[1]

H.1 Equation for Rydberg eigenmodes

Like every tuned resonant circuit, atoms have well-defined resonant frequencies, or eigenmodes. Figure H.1 shows the observed radiation spectra for hydrogen. From the very beginning, it has been clear that there is a pattern to these spectral lines. In 1880 Rydberg easily fitted a formula that quantifies the observed eigen spectral lines of hydrogen in terms of the reciprocals of the radiated wavelengths:

$$\frac{1}{\lambda_{nm}} = R_\infty \left(\frac{1}{n^2} - \frac{1}{m^2} \right), \qquad \frac{f_{nm}}{c_o R_\infty} = \frac{1}{n^2} - \frac{1}{m^2}, \tag{H.1.1}$$

[1] https://physics.aps.org/synopsis-for/10.1103/PhysRevLett.121.193401

– 5.4: Consider the function with poles at $z = \pm j$,

$$F(z) = \frac{1}{1+z^2} = \frac{1}{(z-j)(z+j)}.$$

Find the residue expansion.

Problem # *6: Apply Cauchy's theorems to solve the following integrals. State which theorem(s) you used and show your work.*

– 6.1: $\int_{\mathcal{C}} F(z)dz$*, where $\mathcal{C}$ is a circle centered at $z = 0$ with a radius of $\frac{1}{2}$*

– 6.2: $\int_{\mathcal{C}} F(z)dz$*, where $\mathcal{C}$ is a circle centered at $z = j$ with a radius of 1*

– 6.3: $\int_{\mathcal{C}} F(z)dz$*, where $\mathcal{C}$ is a circle centered at $z = 0$ with a radius of 2*

Integration of analytic functions

Problem # *7: In the following questions, you'll be asked to integrate $F(s) = u(\sigma, \omega) + iv(\sigma,$ around the contour $\mathcal{C}$ for complex $s = \sigma + i\omega$,*

$$\int_{\mathcal{C}} F(s)ds. \tag{DE-2.6}$$

Follow the directions carefully for each question. When asked to state where the function is and is not analytic, you are not required to use the Cauchy-Riemann equations

– 7.1: $F(s) = \sin(s)$

– 7.2: Given function $F(s) = \frac{1}{s}$ State where the function is and is not analytic.

– 7.3: Explicitly evaluate the integral when $\mathcal{C}$ is the unit circle, defined as $s = e^{i\theta}$, $0 \leq \theta \leq 2\pi$.

– 7.4: Evaluate the same integral using Cauchy's theorem and/or the residue theorem.

– 7.5: $F(s) = \frac{1}{s^2}$ *State where the function is and is not analytic.*

– 7.6: Explicitly evaluate the integral when $\mathcal{C}$ is the unit circle, defined as $s = e^{i\theta}$, $0 \leq \theta \leq 2\pi$.

– 7.7: What does your result imply about the residue of the second-order pole at $s = 0$?

– 7.8: $F(s) = e^{st}$: State where the function is and is not analytic.

– 7.9: Explicitly evaluate the integral when $\mathcal{C}$ is the square $(\sigma, \omega) = (1, 1) \to (-1, 1) \to (-1, -1) \to (1, -1) \to (1, 1)$.

– 7.10: Evaluate the same integral using Cauchy's theorem and/or the residue theorem.

– 7.11: $F(s) = \frac{1}{s+2}$: State where the function is and is not analytic.

– 7.12: Let $\mathcal{C}$ be the unit circle, defined as $s = e^{i\theta}$, $0 \leq \theta \leq 2\pi$. Evaluate the integral using Cauchy's theorem and/or the residue theorem.

– 7.13: Let $\mathcal{C}$ be a circle of radius 3, defined as $s = 3e^{i\theta}$, $0 \leq \theta \leq 2\pi$. Evaluate the integral using Cauchy's theorem and/or the residue theorem.

– 7.14: $F(s) = \frac{1}{2\pi i}\frac{e^{st}}{(s+4)}$ State where the function is and is not analytic.

– 7.15: Let $\mathcal{C}$ be a circle of radius 3, defined as $s = 3e^{i\theta}$, $0 \leq \theta \leq 2\pi$. Evaluate the integral using Cauchy's theorem and/or the residue theorem.

– 7.16: Let $\mathcal{C}$ contain the entire left half s plane. Evaluate the integral using Cauchy's theorem and/or the residue theorem. Do you recognize this integral?

– 7.17: $F(s) = \pm\frac{1}{\sqrt{s}}$ (e.g., $F^2 = \frac{1}{s}$) State where the function is and is not analytic.

– 7.18: This function is multivalued. How many Riemann sheets do you need in the domain (s) and the range (f) to fully represent this function? Indicate (e.g., using a sketch) how the sheet(s) in the domain map to the sheet(s) in the range.

– 7.19: Explicitly evaluate the integral $\int_{\mathcal{C}} \frac{1}{\sqrt{z}} dz$ when $\mathcal{C}$ is the unit circle, defined as $s = e^{i\theta}$, $0 \leq \theta \leq 2\pi$. Is this contour closed? State why or why not.

$M(x)$. The ratio C/M determines the area-dependent characteristic admittance $\mathcal{Y}_r(x)$ ($\in \mathbb{R}$):

$$\mathcal{Y}_r(x) = \frac{1}{\sqrt{\text{stiffness} \cdot \text{mass}}} = \frac{\overline{\mathcal{Y}(x,s)}}{\mathcal{Z}(x,s)} = \frac{\overline{C(x)}}{M(x)} = \frac{\overline{A(x)}\,\overline{\not{s}A(x)}}{\not{s}\rho_o \quad \eta_o \mathcal{P}_o} = \frac{A(x)}{\rho_o c_o} > 0 \quad \text{(G.1.9)}$$

(Campbell, 1903, 1910, 1922). The characteristic impedance is $\mathcal{Z}_r(x) = 1/\mathcal{Y}_r(x)$. Based on a physical argument, $\mathcal{Y}_r(x)$ must be positive and real; thus only the positive square root is allowed. As long as $A(x)$ has no jumps (is continuous), $\mathcal{Y}_r(x)$ must be the same in both directions. It is locally determined by the isopressure surface and its volume velocity.

Radiation admittance

The radiation admittance is defined looking into a horn with no termination (infinitely long) from the input at $x = 0$:

$$Y^{\pm}_{rad}(s) = \frac{\mathcal{V}^{\pm}}{\mathcal{P}^{\pm}} \in \mathbb{C}. \quad \text{(G.1.10)}$$

The impedance depends on the direction, with $+$ looking to the right and $-$ to the left.

The input admittance $Y^{\pm}_{in}(x,s)$ is computed using the upper equation of Eq. 5.2.11 (p. 242) for $\mathcal{V}(x,s)$ and then dividing by the pressure eigenfunction $\mathcal{P}^{\pm}$. This results in the logarithmic derivative of $\mathcal{P}^{\pm}(x,s)$:

$$Y^{\pm}_{in}(x,s) \equiv \frac{\mathcal{V}^{\pm}}{\mathcal{P}^{\pm}} = \frac{-1}{sM(x)}\frac{\partial}{\partial r} \ln \mathcal{P}^{\pm}(x,s).$$

For example, for the conical horn (last column of Table 5.2, p. 247)

$$Y^{\pm}_{in} = \mathcal{Y}_r(1 \pm c_o/sr_o). \quad \text{(G.1.11)}$$

Note that $Y^{+}_{in}(x,s) + Y^{-}_{in}(x,s) = 2\mathcal{Y}_r = 2A_0 r^2/\rho_o c_o \in \mathbb{R}$, which shows that the frequency-dependent parts of the two admittances, being equal and opposite in sign, exactly cancel.

As the wavefront travels down the variable-area horn, there is a mismatch in the characteristic admittance due to the change in area. This mismatch creates a reflected wave, which in the case of the conical horn is $= -c_o/sr_o$. Due to conservation of volume, there is a corresponding identical forward component that travels forward, equal to $+c_o/sr_o$. The sum of these two responses to the change in area must be zero in order to conserve volume velocity.

The resulting equation for the velocity eigenfunctions is therefore

$$\mathcal{V}^{\pm}(x,s) = Y^{\pm}_{in}(x,s)\,\mathcal{P}^{\pm}(x,s).$$

Propagation function $\kappa(s)$ The eigenfunctions of the lossless wave equation propagate as

$$\mathcal{P}^{\pm}(x,s) = \frac{e^{\mp\kappa(s)x}}{\overline{A(x)}},$$

where $\kappa(s) = \overline{\mathcal{Z}(x,s)\mathcal{Y}(x,s)} = \pm s\sqrt{MC}$. The velocity eigenfunctions $\mathcal{V}^{\pm}(x,s)$ may be computed from Eq. G.1.4.

From the above definitions,

$$\kappa(s) = \frac{\overline{s\rho_o \quad \cancel{sA(x)}}}{\cancel{A(x)} \quad \eta_o P_o} = \frac{s}{c_o}.$$

Thus $\kappa(s)$ and s are the eigenvalues of the differential operations $\partial/\partial x$ and $\partial/\partial t$ on the pressure $\mathcal{P}(x,s)$. See Appendix D for the inclusion of visco-thermal losses.

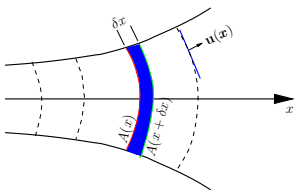


Figure G.1: *Derivation of the horn equation using Gauss's law: The divergence of the velocity $\nabla \cdot \mathbf{u}$ within δx, shown as the shaded region, is integrated over the enclosed volume. Next the divergence theorem is applied, transforming the integral to a surface integral normal to the surface of propagation. This results in the difference of the two volume velocities $\delta\nu = \nu(x+\delta x) - \nu(x) = [\mathbf{u}(x+\delta x)\cdot\mathbf{A}(x+\delta x) - \mathbf{u}(x)\cdot\mathbf{A}(x)]$. The flow is always perpendicular to the isopressure contours.*

G.1.2 Conservation of mass

Integrating Eq. 5.2.6 over the volume V gives

$$-\int_V \nabla\cdot\boldsymbol{u}\,dV = \frac{1}{\eta_o P_o}\frac{\partial}{\partial t}\int_V p(\boldsymbol{x},t)dV.$$

The volume V is defined by two isopressure surfaces between x and $x+\delta x$ (the shaded region of Fig. G.1). On the right-hand side we use the definition of the average pressure (i.e., Eq. G.1.1) integrated over the volume dV.

Applying Gauss's law to the left-hand side[1] and using the definition of ϱ (on the right) in the limit $\delta x \to 0$ give

$$-\frac{\partial}{\partial x}\nu(x,t) = \frac{A(x)}{\eta_o P_o}\frac{\partial}{\partial t}\varrho(x,t) \leftrightarrow \mathcal{Y}(x,s)\,\mathcal{P}(x,s), \tag{G.1.6}$$

where

$$\mathcal{Y}(x,s) = sA(x)/\eta_o P_o = sC(x). \tag{G.1.7}$$

$C(x) = A(x)/\eta_o P_o$ [m^4/N] is the per-unit-length compliance of the air. Equations G.1.4 and G.1.6 accurately characterize the Webster plane-wave mode up to the frequency where the higher-order eigenmodes begin to propagate (i.e., $f > f_c$).

G.1.3 Horn properties

Speed of sound c_o

In terms of $M(x)$ and $C(x)$, the speed of sound and the acoustic admittance are

$$c_o = \sqrt{\frac{\text{stiffness}}{\text{mass}}} = \frac{1}{\sqrt{C(x)M(x)}} = \sqrt{\frac{\eta_o P_o}{\rho_o}}. \tag{G.1.8}$$

This assumes the medium is lossless. For a discussion of lossy propagation, see Appendix D (p. 311).

G.1.4 Characteristic admittance $\mathcal{Y}_r(x)$:

Since the horn equation (Eq. 5.2.10) is second-order, it has two eigenfunction solutions $\mathcal{P}^{\pm}$. The ratios of Eq. G.1.7 to Eq. G.1.5 are determined by the local stiffness $1/C(x)$ and mass

[1] As shown in Fig. G.1, taking the limit of the difference between the two volume velocities $\nu(x+\delta x) - \nu(x)$ divided by δx results in $\partial\nu/\partial x$.

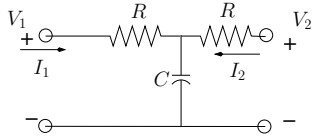
– 7.20: Explicitly evaluate the integral $\int_{\mathcal{C}} \frac{1}{\sqrt{z}}dz$ when $\mathcal{C}$ is twice *around the unit circle, defined as $s = e^{i\theta}$, $0 \le \theta \le 4\pi$. Is this contour closed? State why or why not. Hint: Note that $\sqrt{e^{i(\theta+2\pi)}} = \sqrt{e^{i2\pi}e^{i\theta}} = e^{i\pi}\sqrt{e^{i\theta}} = -1\sqrt{e^{i\theta}}$.*

– 7.21: What does your result imply about the residue of the (twice-around $\frac{1}{2}$ order) pole at $s = 0$?

– 7.22: Show that the residue is zero. Hint: Apply the definition of the residue.

4.6.3 Laplace transform applications

Problem # *8: A two-port network application for the Laplace transform*



This three-element electrical circuit is a system that acts to low-pass filter the signal voltage $V_1(\omega)$, to produce signal $V_2(\omega)$. It is convenient to define the dimensionless ratio $s/s_c = RCs$ in terms of a time constant $\tau = RC$ and cutoff frequency $s_c = 1/\tau$.

4.6.4 Computer exercises with Matlab/Octave

Problem # *9: With the help of a computer*

Now we look at a few important concepts using Matlab/Octave's `syms` commands or Wolfram Alpha's symbolic math toolbox.[11]

For example, to find the Taylor series expansion about $s = 0$ of

$$F(s) = -\log(1-s),$$

we first consider the derivative and its Taylor series (about $s = 0$)

$$F'(s) = \frac{1}{1-s} = \sum_{n=0}^{\infty} s^n.$$

Then, we integrate this series term by term:

$$F(s) = -\log(1-s) = \int_0^s F'(s)ds = \sum_{n=0}^{\infty}\frac{s^n}{n}.$$

Alternatively we can use Matlab/Octave commands:

[11] `https://www.wolframalpha.com/`

```
syms s
taylor(-log(1-s),'order',7)
```

– 9.1: Use Octave's `taylor(-log(1-s))` *to the seventh order, as in the example above. Try the above Matlab/Octave commands. Give the first seven terms of the Taylor series (confirm that Matlab/Octave agrees with the formula derived above).*

– 9.2: What is the inverse Laplace transform of this series? Consider the series term by term.

– 9.3: The function $1/\sqrt{z}$ *has a branch point at* $z = 0$*; thus it is singular there. Can you apply Cauchy's integral theorem when integrating around the unit circle?*

– 9.4: This Matlab/Octave code computes $\int_0^{4\pi} \frac{dz}{\sqrt{z}}$ *using Matlab's/Octave's symbolic analysis package. Run the following script:*

```
syms  z
I=int(1/sqrt(z))
J = int(1/sqrt(z),exp(-j*pi),exp(j*pi))
eval(J)
```

What answers do you get for I and J?

– 9.5: Modify this code to integrate $f(z) = 1/z^2$ once *around the unit circle. What answers do you get for* I *and* J*?*

– 9.6: Bessel functions can describe waves in a cylindrical geometry.
The Bessel function has a Laplace transform with a branch cut

$$J_0(t)u(t) \leftrightarrow \frac{1}{\sqrt{1+s^2}}.$$

Draw a hand sketch showing the nature of the branch cut. Hint: Use `zviz`.

Appendix G

Webster horn equation derivation

G.1 Overview

In this appendix we transform the acoustic equations, Eqs. 5.2.5 and 5.2.6 (p. 239), into their equivalent integral form, Eq. 5.2.10 (p. 241). This derivation is similar (but not identical) to that of Hanna and Slepian (1924) and Pierce (1981, p. 360).

G.1.1 Conservation of momentum

The first step is to integrate the normal component of Eq. 5.2.5 (p. 239) over the isopressure surface $\mathcal{S}$, defined by $\nabla p = 0$

$$-\int_{\mathcal{S}} \nabla p(\mathbf{x},t) \cdot d\mathbf{A} = \rho_o \frac{\partial}{\partial t} \int_{\mathcal{S}} \mathbf{u}(\mathbf{x},t) \cdot d\mathbf{A}.$$

The average pressure $\varrho(x,t)$ is defined by dividing by the total area:

$$\varrho(x,t) \equiv \frac{1}{A(x)} \int_{\mathcal{S}} p(x,t)\, \hat{\boldsymbol{n}} \cdot d\boldsymbol{A}. \tag{G.1.1}$$

From the definition of the gradient operator, we have

Is there an isopressure su given diffusion?

$$\nabla p = \frac{\partial p}{\partial x}\hat{\boldsymbol{n}}, \tag{G.1.2}$$

where $\hat{\boldsymbol{n}}$ is a unit vector perpendicular to the isopressure surface $\mathcal{S}$. Thus the left side of Eq. 5.2.5 reduces to $\partial\varrho(x,t)/\partial x$.

The integral on the right side defines the volume velocity,

$$\nu(x,t) \equiv \int_{\mathcal{S}} \mathbf{u}(x,t) \cdot d\boldsymbol{A}. \tag{G.1.3}$$

Thus the integral form of Eq. 5.2.5 becomes

$$-\frac{\partial}{\partial x}\varrho(x,t) = \frac{\rho_o}{A(x)}\frac{\partial}{\partial t}\nu(x,t) \leftrightarrow \mathcal{Z}(x,s)\,\mathcal{V}, \tag{G.1.4}$$

where

$$\mathcal{Z}(s,x) = s\rho_o/A(x) = sM(x) \tag{G.1.5}$$

and $M(x) = \rho_o/A(x)$ [kgm/m^5] is the per-unit-length mass density of air.

Problem # *10: Matlab/Octave exercises*

– 10.1: Try the following Matlab/Octave commands, and then comment on your findings

```
%Take the inverse LT of 1/sqrt(1+s^2)
syms s
I=ilaplace(1/(sqrt((1+s^2))));
disp(I)

%Find the Taylor series of the LT
T = taylor(1/sqrt(1+s^2),10); disp(T);
```

```
%Verify this
syms t
J=laplace(besselj(0,t));
disp(J);

%plot the Bessel function
t=0:0.1:10*pi;
b=besselj(0,t);
plot(t/pi,b);
grid on;
```

– 10.2: When did Friedrich Bessel live?

– 10.3: What did he use Bessel functions for?

Problem # *11: Colorized plots of analytic functions. Use* `zviz` *for each of the following.*

– 11.1: Describe the plot generated by `zviz S=Z`*.*

– 11.2: Describe the plot generated by `zviz` $1./sqrt(1+S.^2)$*.*

– 11.3: Describe the plot generated by `zviz` $1./sqrt(1-S.^2)$*. Is this function a Brune impedance (i.e., does this function obey*

– 11.4: `zviz` $1./(1+sqrt(S))$

4.6.5 Inverse of Riemann $\zeta(s)$ function

Problem # *12: Inverse zeta function (This problem is for extra credit).*

– 12.1: Find the $\mathcal{LT}^{-1}$ *of one factor of the Riemann zeta function* $\zeta_p(s)$*, where* $\zeta_p(s) \leftrightarrow z_p(t)$*. Describe your results in words. Hint: Consider the geometric series representation*

$$\zeta_p(s) = \frac{1}{1-e^{-sT_p}} = \sum_{k=0}^{\infty} e^{-skT_p}, \tag{DE-2.7}$$

for which you can look up the $\mathcal{LT}^{-1}$ of each term.

Problem # *13: Inverse transform of products:*
The time-domain version of Eq. DE-2.7 may be written as the convolution of all the $z_k(t)$ factors:

$$z(t) \equiv z_2(t) \star z_3(t) \star z_5(t) \star z_7(t) \star \cdots \star z_p(t) \star \cdots , \tag{DE-2.8}$$

where $\star$ represents time convolution.

Explain what this means in physical terms. Start with two terms (e.g., $z_1(t) \star z_2$). Hint: The input admittance of this cascade may be interpreted as the analytic continuation of $\zeta(s)$ by defining a cascade of eigenfunctions with eigenvalues derived from the primes. For a discussion of this idea see Sec. 3.2.3 and C.1.1.

Physical interpretation: Such functions may be generated in the time domain, as shown in Fig. 3.3, using a feedback delay of T_p seconds, described by the two equations in the Fig. 3.3 with a unity feedback gain $\alpha = -1$. Taking the Laplace transform of the system equation, we see that the transfer function between the state variable $q(t)$ and the input $x(t)$ is given by $\zeta_p(s)$, which is an all-pole function, since

$$Q(s) = e^{-sT_n}Q(s) + V(s), \text{ or } \zeta_p(s) \equiv \frac{Q(s)}{V(s)} = \frac{1}{1 - e^{-sT_p}}. \tag{DE-2.9}$$

Closing the feed-forward path gives a second transfer function $Y(s) = I(s)/V(s)$—namely,

$$Y(s) \equiv \frac{I(s)}{V(s)} = \frac{1 - e^{-sT_p}}{1 + e^{-sT_p}}. \tag{DE-2.10}$$

If we take $i(t)$ as the current and $v(t)$ as the voltage at the input to the transmission line, then $y_p(t) \leftrightarrow \zeta_p(s)$ represents the input impedance at the input to the line. The poles and zeros of the impedance interleave along the $j\omega$ axis. By a slight modification, $\zeta_p(s)$ may alternatively be written as

$$Y_p(s) = \frac{e^{sT_p/2} + e^{-sT_p/2}}{e^{sT_p/2} - e^{-sT_p/2}} = j\tan(sT_p/2). \tag{DE-2.11}$$

Every impedance $Z(s)$ has a corresponding *reflectance* function given by a Möbius transformation, which may be read off of Eq. DE-2.11 as

$$\Gamma(s) \equiv \frac{1 + Z(s)}{1 - Z(s)} = e^{-sT_p}, \tag{DE-2.12}$$

since impedance is also related to the round-trip delay T_p on the line. The inverse Laplace transform of $\Gamma(s)$ is the round-trip delay T_p on the line

$$\gamma(t) = \delta(t - T_p) \leftrightarrow e^{-sT_p}. \tag{DE-2.13}$$

Working in the time domain provides a key insight, as it allows us to parse out the best analytic continuation of the infinity of possible continuations that are not obvious in the frequency domain (See Sec. 3.2.3). Transforming to the time domain is a form of analytic continuation of $\zeta(s)$ that depends on the assumption that $Z^{eta}(t) \leftrightarrow \zeta(s)$ is one-sided in time (causal).

Sergei Schelkunoff may have been the first to formalize this concept (Schelkunoff, 1943), but he was not the first to use it, as exemplified by the Helmholtz resonator. George Ashley Campbell was the first to use the concept in the important application of a wave filter, some 30 years before Schelkunoff (Campbell, 1903). These two men were 40 years apart and both worked for the telephone company (after 1929, called AT&T Bell Labs) (Fagen, 1975).

There are alternative definitions of the QS approximation, depending on the geometric cell structure. The alternatives are listed in Table F.1.

The quasistatic approximation: Since the velocity perpendicular to the walls of a horn must be zero, any radial wave propagation is exponentially attenuated ($\kappa(s)$ is real and negative, i.e., the propagation function $\kappa(s)$ (Sec. 4.4, p. 189) will not describe radial wave propagation), with a space constant of about 1 diameter. The assumption that these radial waves can be ignored (i.e., more than 1 diameter from their source) is called the quasistatic approximation. As the frequency is increased and once $f \geq f_c = 2c_o/\lambda$, the radial wave can satisfy the zero normal velocity wall boundary condition and therefore will not be attenuated. Thus above this critical frequency, radial waves (also known as higher-order modes) are supported (κ becomes imaginary). Thus for Eq. 5.2.10 (p. 241) to describe guided wave propagation, $f < f_c$. But even under this condition, the solution is not precise within a diameter (or so) of any discontinuities (i.e., rapid variations) in the area.

Each horn, as determined by the area function $A(r)$, has a distinct wave equation and thus a distinct solution. Note that the area function determines the upper cutoff frequency via the quasistatic approximation, since $f_c = c_o/\lambda_c$, $\lambda_c/2 > d$, and $A(r) = \pi(d/2)^2$. Thus to satisfy the quasistatic approximation, the frequency f must be less than the cutoff frequency:

$$f < f_c(r) = \frac{c_o}{4}\sqrt{\frac{\pi}{A(r)}}. \tag{F.3.4}$$

We have discussed two alternative matrix formulations of these equations: the ABCD transmission matrix, used for computation, and the impedance matrix, used when working with experimental measurements (Pierce, 1981, Chap. 7). For each formulation, reciprocity and reversibility show up as different matrix symmetries, as addressed in Sec. 3.10 (p. 155) (Pierce, 1981, pp. 195–203).

Periodic $\leftrightarrow$ discrete: P11 As has been shown in the discussion on the Fourier transform, when the time (or frequency) domain response is periodic, the frequency (or time) domain is discrete. This is a fundamental symmetry property that must always be obeyed. This is closely related to the causal $\leftrightarrow$ complex–analytic property of the Laplace and z transforms.

Summary

A transducer converts between modalities. We propose the general definitions of the eleven system postulates that include all transduction modalities, such as electrical, mechanical, and acoustical. It is necessary to generalize the concept of the QS approximation (Postulate P9) to allow for guided waves.

Given the combination of the important QS approximation and these space-time, linearity, and reciprocity properties, a rigorous definition and characterization of a system can thus be established. It is based on a taxonomy of such materials formulated in terms of material and physical properties and extended network postulates.

different time scale. We first discuss generalized postulates, expanding on those of Carlin and Giordano.

Deterministicity (randomness): P9 When the media are uniform and time-invariant, the impedance and transfer functions will be deterministic. When the media are turbulent, the response will be random. When light propagates through the universe, it is strongly time-varying. Thus, astrophysics will be seen as random, whereas experiments in the shelter of the relatively uniform environment of the earth will be deterministic. Model calculation will also be deterministic unless one is trying to create a random, time-varying, turbulent medium.

The quasistatic constraint: P10 When a system is described by the wave equation, delay is introduced between two points in space and depends on the wave speed. When the wavelength is large compared to the delay, we can successfully apply the quasistatic (QS) approximation. This method has widespread application and is frequently used without mention of the assumption. This can lead to confusion, since the limitations of the approximation may not be appreciated. An example is the use of quasistatics in quantum mechanics. The QS approximation has widespread use when the signals may be accurately approximated by a band-limited signal. Examples include KCL, KVL, wave guides, transmission lines, and most important, impedance. The QS property is not mentioned in the six postulates of Carlin and Giordano (1964); thus they need to be extended in some fundamental ways.

When the dimensions of a cellular structure in the material are much smaller than the wavelength, can the QS approximation be valid? This effect can be viewed as a mode filter that suppresses unwanted (or conversely enhances the desired) modes (Ramo et al., 1965). QS may be applied to a three-dimensional specification, as in a semiconductor lattice. But such applications fall outside the scope of this text (Schwinger and Saxon, 1968).

Although I have never seen the idea discussed in the literature, the QS approximation is applied when Green's theorem is defined. For example, Gauss's law is not true when the volume of the container violates QS, since changes in the distribution of the charge have not reached the boundary, when doing the integral. Thus such integral relationships assume that the system is in quasi-steady-state (i.e., that QS holds).

Table F.1: *Several ways of indicating the quasistatic (QS) approximation. For network theory there is only one lattice constant a, which must be much less than the wavelength (wavelength constraint). These three constraints are not equivalent when the object may be a larger structured medium, spanning many wavelengths but with a cell structure size much smaller than the wavelength. For example, each cell could be a Helmholtz resonator or an electromagnetic transducer (e.g., an earphone).*

Measure	Domain
$ka < 1$	Wave number constraint
$\lambda > 2\pi a$	Wavelength constraint
$f_c < c/2\pi a$	Bandwidth constraint

Formally, QS is defined as $ka < 1$, where $k = 2\pi/\lambda = \omega/c$ and a is the cellular dimension or the size of the object. Other ways of expressing this include $\lambda/4 > a$, $\lambda/2\pi > a$, $\lambda > 4a$, or $\lambda > 2\pi a$. It is not clear whether it is better to normalize λ by 4 (quarter-wavelength constraint) or $2\pi \approx 6.28 > 4$, which is more conservative by a factor of $\pi/2 \approx 1.6$. Also k and a can be vectors (e.g., Eq. 3.1.5, p. 75).

4.6.6 Quadratic forms

A matrix that has positive eigenvalues is said to be positive-definite. The eigenvalues are real if the matrix is symmetric, so this is a necessary condition for the matrix to be positive-definite. This condition is related to conservation of energy, since the power is the voltage times the current. Given an impedance matrix

$$\mathbf{V} = \mathcal{Z}\mathbf{I},$$

the power $\mathcal{P}$ is

$$\mathcal{P} = \mathbf{I} \cdot \mathbf{V} = \mathbf{I} \cdot \mathcal{Z}\mathbf{I},$$

which must be positive-definite for the system to obey conservation of energy.

Problem # *14: In this problem, consider the 2×2 impedance matrix*

$$\mathcal{Z} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}.$$

– 14.1: Solve for the power $\mathcal{P}(i_1, i_2)$ by multiplying out this matrix equation (which is a quadratic form*):*

$$\mathcal{P}(i_1, i_2) = \mathbf{I}^T \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{I}.$$

– 14.2: Is the impedance matrix positive-definite? Show your work by finding the eigenvalues of the matrix **Z***.*

– 14.3: Should an impedance matrix always be positive-definite? Explain.

4.7 The $\mathcal{LT}$ and its inverse $\mathcal{LT}^{-1}$

The Laplace transform $\mathcal{LT}$ take causal time functions into the complex–analytic frequency domain s. The inverse Laplace transform $\mathcal{LT}^{-1}$ (Eq. 3.10.1, p. 156) transforms a function of complex frequency $F(s)$ and returns a causal function of time $f(t)$,

$$f(t)u(t) \leftrightarrow F(s),$$

where $u(t) = 0$ for $t < 0$. Examples are provided in Table 3.9 (p. 160). The forward transform is typically a relatively simple set of integrals to find $F(s)$. However the inverse transform is the key to understanding this powerful tool. Here we discuss the details of finding the inverse transform by using CT-3, and we see how the causal requirement $f(t < 0) = 0$ comes about.

As shown in Fig. 4.5, the integrand of the inverse transform is $F(s)e^{st}$ and the limits of integration are $\sigma_o \mp \omega\jmath$ with $\sigma_o > 0$. To find the inverse using CT-3 we must close the curve at $\omega \to \infty$, and specify that the integral converges. There are two ways to close the integral: to the right $\sigma > 0$ (RHP), and to the left $\sigma \leq 0$ (LHP). But there must be some logical reason for this choice. That logic is determined by the sign of t. For the integral to converge, the term $|e^{st}| = e^{\sigma t}$ must go to zero as $\omega \to \infty$. Note that both t and ω go to ∞. Thus it is the interaction between these two limits that determines how we pick the closure, RHP or LHP.

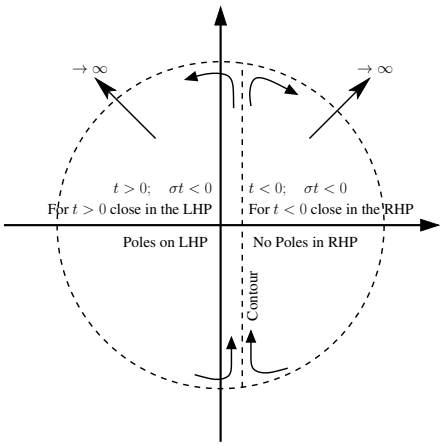


Figure 4.5: *When computing the $\mathcal{LT}^{-1}$ we must take advantage of the powerful Cauchy residue integral theorem (CT-3). To use CT-3 we must close the integral at $s \to \infty$. Furthermore this closure integral must go to zero as $\{t, \sigma\} \to \infty$. For $t < 0$, the convergence of the residue $e^{-\sigma t}$ depends on $\sigma t < 0$, since the product st must have a negative real part. For convergence for $t < 0$, $\sigma > 0$ so that $\sigma t < 0$. Thus the integral must be close in in the RHP. Likewise, for $t > 0$, $\sigma < 0$, so that $\sigma t < 0$. Thus the integral must be close in the LHP. Following these guidelines based on CT-3, poles in the LHP lead to causal stable solutions, since there $\sigma < 0$ and $t \to \infty$, while poles in the RHP will lead to causal unstable solutions, each the form $e^{\sigma t}u(t)$.*

Rayleigh reciprocity: P6 Reciprocity is defined in terms of the unloaded output voltage that results from an input current. Specifically (same as Eq. 3.8.6, p. 146),

$$\begin{bmatrix} z_{11}(s) & z_{12}(s) \\ z_{21}(s) & z_{22}(s) \end{bmatrix} = \frac{1}{C(s)} \begin{bmatrix} A(s) & \Delta_{\mathcal{T}} \\ 1 & D(s) \end{bmatrix}, \tag{F.3.3}$$

where $\Delta_{\mathcal{T}} = A(s)D(s) - B(s)C(s) = \pm 1$ for the reciprocal and antireciprocal systems, respectively. This is best understood in terms of Eq. F.2.1. The off-diagonal coefficients $z_{12}(s)$ and $z_{21}(s)$ are defined as

$$z_{12}(s) = \left.\frac{\Phi_i}{U_l}\right|_{I_i=0} \qquad z_{21}(s) = \left.\frac{F_l}{I_i}\right|_{U_l=0}.$$

When these off-diagonal elements are equal [$z_{12}(s) = z_{21}(s)$], the system is said to obey Rayleigh reciprocity. If they are opposite in sign [$z_{12}(s) = -z_{21}(s)$], the system is said to be antireciprocal. If a network has neither reciprocal nor antireciprocal characteristics, then we denote it nonreciprocal (McMillan, 1946). The most comprehensive discussion of reciprocity, even to this day, is that of Rayleigh (1896, Vol. I). The reciprocal case may be modeled as an ideal transformer (Van Valkenburg, 1964a), while for the antireciprocal case the generalized force and flow are swapped across the two-port. An electromagnetic transducer (e.g., a moving-coil loudspeaker or electrical motor) is antireciprocal (Kim and Allen, 2013; Beranek and Mellow, 2012); it requires a gyrator rather than a transformer, as shown in Fig. F.1.

Reversibility: P7 A second two-port property is the reversible/nonreversible postulate. A reversible system is invariant to the input and output impedances being swapped. This property is defined by the input and output impedances being equal.

Referring to Eq. F.3.3, when the system is reversible, $z_{11}(s) = z_{22}(s)$ or, in terms of the transmission matrix variables, $\frac{A(s)}{C(s)} = \frac{D(s)}{C(s)}$ or simply $A(s) = D(s)$, assuming $C(s) \neq 0$.

An example of a nonreversible system is a transformer with a turns ratio that is not 1. Also, an ideal operational amplifier (when the power is turned on) is nonreversible due to the large impedance difference between the input and output. Furthermore, it is active; it has a power gain due to the current gain at constant voltage (Van Valkenburg, 1964b).

Generalizations of this lead to group theory and Noether's theorem. These generalizations apply to systems that have many modes, whereas quasistatics holds when they operate below a cutoff frequency (Table F.1), meaning that, as in the case of the transmission line, there are no propagating transverse modes. While this assumption is never exact, it leads to highly accurate results because the nonpropagating evanescent transverse modes are attenuated over a short distance and thus, in practice, may be ignored (Montgomery et al., 1948; Schwinger and Saxon, 1968, Chaps. 9–11).

We extend the Carlin and Giordano postulate set to include Postulate P7, reversibility, which was refined by Van Valkenburg (1964a). To satisfy the reversibility condition, the diagonal components in a system's impedance matrix must be equal. In other words, the input force and flow are proportional to the output force and flow, respectively (i.e., $Z_e = z_m$).

Spatial invariance: P8 The characteristic impedance and wave number $\kappa(s, x)$ may be strongly frequency- and/or spatially dependent or even be negative over some limited frequency ranges. Due to causality, the concept of a negative group velocity must be restricted to a limited bandwidth (Brillouin, 1960). As Einstein's theory of relativity makes clear, all materials must be strictly causal (Postulate P1), a view that must therefore apply to acoustics but at a very

2. If $Z(s)$ is PR, then its reciprocal $Y(s) = 1/Z(s)$ is PR (the poles and zeros swap).

3. If the impedance can be written as the ratio of two polynomials (a limited case related to the quasistatics approximation, Postulate P9) that have degrees N and L, then $|N - L| \leq 1$.

4. The angle of the impedance $\theta \equiv \angle Z$ lies within $[-\pi \leq \theta \leq \pi]$.

5. The impedance and its reciprocal are complex–analytic in the right half-plane; thus each obeys the Cauchy–Riemann conditions there.

Energy and power: Since Postulate P3 requires the impedance PR condition, it ensures that every impedance is positive-definite (PD), thus guaranteeing that conservation of energy is obeyed (Schwinger and Saxon, 1968, p.17). This means that the total energy absorbed by any PR impedance must remain positive for all time. Mathematically we can state this as

$$\mathcal{E}(t) = \int_{-\infty}^{t} v(t)i(t)\,dt = \int_{-\infty}^{t} i(t)\star z(t)\;i(t)\,dt > 0,$$

where $i(t)$ is any current, $v(t) = z(t) \star i(t)$ is the corresponding voltage, and $z(t)$ is the real causal impulse response of the impedance [e.g., $z(t) \leftrightarrow Z(s)$ are a Laplace transform pair]. In summary, if $Z(s)$ is PR, then $\mathcal{E}(t)$ is PD.

As discussed in detail by Van Valkenburg, any rational PR impedance can be represented as a partial fraction expansion, which can be expanded into first-order poles as

$$Z(s) = K\frac{\Pi_{i=1}^{L}(s-n_i)}{\Pi_{k=1}^{N}(s-d_k)} = \sum_n \frac{\rho_n}{s-s_n}e^{j(\theta_n-\theta_d)}, \tag{F.3.2}$$

where ρ_n is a complex scale factor (residue). Every pole in a PR function has only simple poles and zeros, which requires that $|L - N| \leq 1$ (Van Valkenburg, 1964b).

Whereas the PD property clearly follows from Postulate P3 (conservation of energy), the physics is not so clear. Specifically, what is the physical meaning of the constraints on $Z(s)$? In many ways, the impedance concept is highly artificial, as expressed by Postulates P1–P7.

When the impedance is not rational, special care must be taken. An example of this is the semi-inductor $M\sqrt{s}$ and the semicapacitor $K/\sqrt{s}$ due, for example, to the skin effect in EM theory and viscous and thermal losses in acoustics, both of which are frequency-dependent boundary-layer diffusion losses (Vanderkooy, 1989). They remain positive-real but have a branch cut and thus are double-valued in frequency.

Real-time response: P4 The impulse response of every physical system is real, not complex. This requires that the Laplace transform have conjugate symmetry $H(s) = H^*(s^*)$, where the $*$ indicates conjugation [e.g., $R(\sigma,\omega) + X(\sigma,\omega) = R(\sigma,\omega) - X(\sigma,-\omega)$].

Time invariance: P5 The meaning of time-invariant requires that the impulse response of a system does not change over time. This requires that the system coefficients of the differential equation describing the system are constant (independent of time).

4.7.1 Case for negative time ($t < 0$) and causality:

Let us first consider negative time, including $t \to -\infty$. If we were to close $\mathcal{C}$ in the LHP ($\sigma < 0$), then the product σt is positive ($\sigma < 0, t < 0$, thus $\sigma t > 0$). In this case, as $\omega \to \infty$, the closure integral $|s| \to \infty$ will diverge. If we close in the RHP ($\sigma > 0$), then the product $\sigma t < 0$ and e^{st} will go to zero as $\omega \to \infty$. Thus we may not close in the LHP for negative time. This then justifies closing the contour, allowing for the use of Cauchy CT-3.

4.7.2 Case for zero time ($t = 0$):

When the time is zero, the integral does not, in general, converge, which leaves $f(t)$ undefined. This is most obvious in the case of the step function $u(t) \leftrightarrow 1/s$, where the integral may not be closed because the convergence factor $e^{st} = 1$ fails for $t = 0$.

The fact that $u(t)$ does not exist at $t = 0$ helps to explain the-Gibbs phenomenon in the inverse Fourier transform. At the time where a jump occurs, the derivative of the function does not exist, and thus the time response function is not analytic. The Fourier expansion does not pointwise converge where the function is not analytic. A low-pass filter may be used to smooth the function, but at the cost of temporal resolution.

4.7.3 Case for positive time ($t > 0$)

Next we investigate the convergence of the integral for positive time $t > 0$. In this case we must close the integral in the LHP ($\sigma < 0$) for convergence, so that $st < 0$ ($\sigma \leq 0$ and $t > 0$). When there are poles on the $\omega\jmath = 0$ axis, $\sigma_o > 0$ assures convergence by keeping the on-axis poles inside the contour. At this point, CT-3 is relevant. If we restrict ourselves to simple poles (as required for a Brune impedance), the residue theorem may be directly applied.

Unstable poles: An important but subtle point arises: If $F(s)$ has a pole in the RHP, then the above argument still applies if we pick σ_o to be to the right of the RHP pole. This means that the inverse transform may still be applied to unstable poles (those in the RHP). This then explains the need for the σ_o in the limits. If $F(s)$ has no RHP poles in the extended RHP ($\sigma \geq 0$), we may take $\sigma_o = 0$.

The simplest example is the step function, for which $F(s) = 1/s$, thus

$$u(t) = \int_{\text{LHP}} \frac{e^{st}}{s}\frac{ds}{2\pi\jmath} \leftrightarrow \frac{1}{s},$$

which is a direct application of CT-3. The forward transform of $u(t)$ is straightforward, as discussed in Appendix C. This is true of most of the elementary forward Laplace transforms. In these cases, causality may be built into the integral by the limits. An interesting problem is how to prove that $u(t)$ is not defined at $t = 0$.

The inverse Laplace transform of $F(s) = 1/(s+1)$ has a residue of 1 at $s = -1$, thus that is the only contribution to the integral. More demanding cases are Laplace transform pairs

$$\frac{1}{\sqrt{t}}u(t) \leftrightarrow \sqrt{\frac{\pi}{s}} \quad \text{and} \quad J_o(t)u(t) \leftrightarrow \frac{1}{\sqrt{s^2+1}},$$

as shown in Fig 4.6 (right), and more in Table C.2 (p. 302). Many of these are easily proved in the forward direction but are much more difficult in the inverse direction due to the properties at $t = 0$, unless CT-3 is invoked.

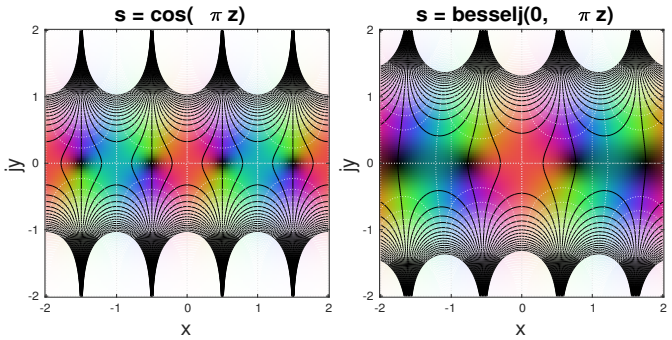


Figure 4.6: *Left: Colorized plot of $w(z) = sin(z)$. Right: Colorized plot of $w(z) = J_o(\pi z)$. Note the similarity of the two functions. The first Bessel zero is at 2.405 and thus appears at $0.7655 = 2.405/\pi$, about 1.53 times larger than the root of $\cos(\pi z)$ at $1/2$. Other than this minor distortion of the first few roots, the two functions are basically identical. It follows that their $\mathcal{LT}$s must have similar characteristics, as documented in Table C.2 p. 302. These colorized plots show that these two functions become the same for $x = \Re z > 0$. The black lines indicate where the function has a constant real part.*

Along the x-axis of Fig. 4.6, $\cos(\pi x)$ is periodic with a period of π. The dark spots are at the zeros at $\pm\pi/2, \pm 3\pi/2, \ldots$. Along the jy-axis, the function goes to either zero (black) or ∞ (white). This behavior carries the same π periodicity as on the $x = 0$ line.

The last $\mathcal{LT}$ example of Fig. 4.7 gives an important insight into the properties of the Hankel function $H_0^{(1)}(\pi z/2)$, which has a branch cut along the negative real axis. On the right is the Hankel function $H_0^{(1)}(\pi z/2)$, which is a mixed and distorted version of $\cos(\pi z)$ with the zeros pushed downward, and $e^{\pi z}$. Note how the white and black contour lines of the colorized maps are always perpendicular where they cross.

4.7.4 Properties of the $\mathcal{LT}$

As shown in Table 3.8 of Laplace transforms, there are integral (i.e., integration, not integer) relationships, or properties, that are helpful to identify. The first of these is a definition, not a property:

tation to Appdx A.

$$f(t) \leftrightarrow F(s).$$

Causality: When we take the $\mathcal{LT}$, the time-domain response is in lowercase (e.g., $f(t)$) and the frequency-domain transform is in uppercase (e.g., $F(s)$). It is required, but not always explicitly specified, that $f(t < 0) = 0$; that is, the time function must be causal, as stated by Postulate P1 (162).

Linearity: The most basic property is the linearity (superposition) property of the $\mathcal{LT}$, stated by Postulate P2 (p. 162).

property, we call it Einstein causality, which says that the delay must be proportional to the distance x, with impulse response $\delta(t - x/c)$.

Linearity: P2 The wave properties of a system may be nonlinear. This is not restrictive, as most physical systems are naturally nonlinear. For example, a capacitor is inherently nonlinear: As the charge builds up on the plates of the capacitor, a stress is applied to the intermediate dielectric due to the electrostatic force $F = qE$. In a similar manner, an inductor is nonlinear. Two wires carrying a current are attracted or repelled due to the force created by the flux. The net force is the product of the two fluxes due to each current.

In summary, most physical systems are naturally nonlinear; it's simply a matter of degree. An important counterexample is an amplifier with negative feedback and a very large open-loop gain. There are, therefore, many types of nonlinearity, both instantaneous types and those with memory (e.g., hysteresis). Given the nature of Postulate P1, even an instantaneous nonlinearity may be ruled out. The linear model is so critical for our analysis, providing fundamental understanding, that we frequently take Postulates P1 and P2 for granted.

Passive/Active impedances: P3 This postulate is about conservation of energy and Otto Brune's positive-real (PR, also called physically realizable) condition that every passive impedance must obey. Following on the work of his primary Ph.D. thesis advisor Wilhelm Cauer (1900–1945) and Ernst Guillemin, along with Norbert Wiener and Vannevar Bush at MIT, Brune mathematically characterized the properties of every PR one-port driving point impedance (Brune, 1931b).

When the input resistance of the impedance is real, the system is said to be passive, which means the system obeys conservation of energy. The real part of $Z(s)$ is positive if and only if the corresponding reflectance is less than 1 in magnitude. The reflectance of $Z(s)$ is defined as a bilinear transformation of the impedance, normalized by its surge resistance r_o (Campbell, 1903):

$$\Gamma(s) = \frac{Z(s) - r_o}{Z(s) + r_o} = \frac{\hat{Z} - 1}{\hat{Z} + 1},$$

where $\hat{Z} = Z/r_o$. The surge resistance is defined in terms of the inverse Laplace transform of $Z(s) \leftrightarrow z(t)$, which must have the form

$$z(t) = r_o\delta(t) + \rho(t),$$

where $\rho(t) = 0$ for $t < 0$. It naturally follows that $\gamma(t) \leftrightarrow \Gamma(s)$ is zero for negative and zero time—namely, $\gamma(0) = 0, t \le 0$.

Given any linear PR impedance $Z(s) = R(\sigma, \omega) + jX(\sigma, \omega)$ that has real part $R(\sigma, \omega)$ and imaginary part $X(\sigma, \omega)$, the impedance is defined as being PR (Brune, 1931b) if and only if

$$\Re Z(s) = R(\sigma \ge 0, \omega) \ge 0. \tag{F.3.1}$$

That is, the real part of any PR impedance is nonnegative everywhere in the right half-plane ($\sigma \ge 0$). This is a very strong condition on the complex–analytic function $Z(s)$ of a complex variable s. This condition is equivalent to any of the following statements (Van Valkenburg, 1964a):

1. There are no poles or zeros in the right half-plane ($Z(s)$ may have poles and zeros on the $\sigma = 0$ axis).

what is needed to fully characterize its operation, from electrical input to mechanical (acoustical) output.

The input is electrical (voltage and current) $[\Phi_i, I_i]$ and the output (load) is the mechanical (force and velocity) $[F_l, U_l]$. The first matrix is the general case, expressed in terms of four unspecified functions $A(s)$, $B(s)$, $C(s)$, and $D(s)$, while the second matrix is for the specific example of Fig. F.1. The three entries are the electrical driving point impedance $Z_e(s)$, the mechanical impedance $z_m(s)$, and the transduction $T = B_o l$, where B_o is the magnetic flux strength and l is the length of the wire crossing the flux. Since the transmission matrix is antireciprocal, its determinant $\Delta_{\mathcal{T}} = -1$, as is easily verified.

Other common examples of cross-modality transduction and current–thermal (thermoelectric effect) and force–voltage (piezoelectric effect). These systems are all reciprocal: thus the transduction has the same sign.

F.2 Impedance matrix

These eleven network postulates describe the properties of a system that has an input and an output. For an electromagnetic transducer (loudspeaker) the system is described by the two-port transmission matrix, as shown in Fig. F.1. The electrical input impedance of a loudspeaker is $Z_e(s)$, defined by

$$Z_e(s) = \left.\frac{V(\omega)}{I(\omega)}\right|_{U=0}.$$

Note that this driving-point impedance must be causal since it is a function of s; thus it has a Laplace transform. The corresponding two-port impedance matrix for Fig. F.1 is

$$\begin{bmatrix}\Phi_i\\ F_l\end{bmatrix} = \begin{bmatrix}z_{11}(s) & z_{12}(s)\\ z_{21}(s) & z_{22}(s)\end{bmatrix}\begin{bmatrix}I_i\\ U_l\end{bmatrix} = \begin{bmatrix}Z_e(s) & -T(s)\\ T(s) & z_m(s)\end{bmatrix}\begin{bmatrix}I_i\\ U_l\end{bmatrix}. \tag{F.2.1}$$

Such a description allows us to define Thévenin parameters, a concept used widely in circuit analysis and network models from other modalities.

The impedance matrix is an alternative description of the system but with generalized forces $[\Phi_i, F_l]$ on the left and generalized flows $[I_i, U_l]$ on the right. A rearrangement of terms allows us to go from the ABCD to the impedance parameters (Van Valkenburg, 1964b). The electromagnetic transducer is antireciprocal (Postulate P6), $z_{12} = -z_{21} = T = B_o l$.

F.3 Taxonomy of algebraic networks

The postulates are extended beyond those defined by Carlin and Giordano (Sec. 3.10, p. 155) when there is an interaction of waves and a structured medium, along with other properties not covered by classic network theory. Assuming quasistatics (QS), the wavelength must be large relative to the medium's lattice constants. Thus the QS property must be extended to three dimensions and possibly to the cases of anisotropic and random media.

Causality: P1 As we stated, due to causality the negative properties (e.g., negative refractive index) must be limited in bandwidth as a result of the Cauchy–Riemann conditions. However, even causality needs to be extended to include the delay, as quantified by the d'Alembert solution to the wave equation, which means that the delay is proportional to the distance. Thus we generalize Postulate P1 to include the space-dependent delay. When we wish to discuss this

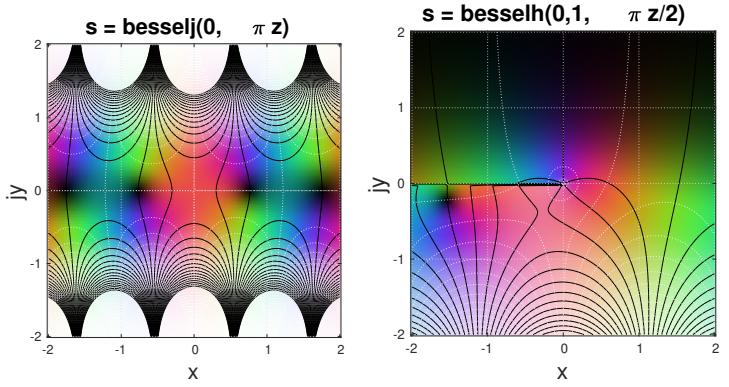


Figure 4.7: *Left: The Bessel function $J_0(\pi z)$, which is similar to $\cos(\pi z)$, except the zeros are distorted away from $s = 0$ by a small amount, due to the cylindrical geometry. Right: The related Hankel function $H_0^{(1)}(\pi z/2)$. The Hankel function $H_0^{(1)}(\pi z/2)$ has a branch cut and a complex zero at $z_{0,1}2/\pi = -1.5 - 0.1j$, as may be seen in the plot.*

Convolution property: The product of two $\mathcal{LT}$s in frequency results in convolution in time:

$$F(s)G(s) \leftrightarrow f(t) \star g(t) = \int_0^t f(\tau)g(t-\tau)d\tau,$$

where we use $\star$ to indicate the convolution of two time functions.

A key application of convolution is filtering, which takes many forms. The most basic filter is the moving average, the moving sum of data samples, normalized by the number of samples. Such a filter has very poor performance. It also introduces a delay of half the length of the average, which may or may not constitute a problem, depending on the application. Other important examples are a low-pass filter that removes high-frequency noise and a notch filter that removes line noise (i.e., 60 [Hz] in the United States, and its second and third harmonics, 120 and 180 [Hz]). Such noise is typically a result of poor grounding and ground loops. It is better to solve the problem at its root than to remove it with a notch filter. Still, filters are very important in engineering.

By taking the $\mathcal{LT}$ of the convolution we can derive this relationship:

$$\begin{aligned}
\int_0^\infty [f(t) \star g(t)]e^{-st}dt &= \int_{t=0}^\infty \left[\int_0^t f(\tau)g(t-\tau)d\tau\right] e^{-st}dt\\
&= \int_0^t f(\tau)\left(\int_{t=0}^\infty g(t-\tau)e^{-st}dt\right)d\tau\\
&= \int_0^t f(\tau)\left(e^{-s\tau}\int_{t'=0}^\infty g(t')e^{-st'}dt'\right)d\tau\\
&= G(s)\int_0^t f(\tau)e^{-s\tau}d\tau\\
&= G(s)F(s).
\end{aligned}$$

We first encountered this relationship on page 112 in the context of multiplying polynomials, which is the same as convolving their coefficients. The parallel should be obvious. In the case of polynomials, the convolution is discrete in the coefficients, and here it is continuous in time. But the relationships are the same.

Time-shift property: When a function is time-shifted by time T_o, the $\mathcal{LT}$ is modified by e^{-sT_o}, leading to the property

$$f(t - T_o) \leftrightarrow e^{-sT_o} F(s).$$

This is easily shown by applying the definition of the $\mathcal{LT}$ to a delayed time function.

Time derivative: The key to the eigenfunction analysis provided by the $\mathcal{LT}$ is the transformation of a time derivative on a time function—that is,

$$\frac{d}{dt} f(t) \leftrightarrow sF(s).$$

Here s is the eigenvalue corresponding to the time derivative of e^{st}. Given the definition of the derivative of e^{st} with respect to time, this definition seems trivial. Yet that definition was not obvious to Euler. It needed to be extended to the space of the complex–analytic function e^{st}, which did not happen before Cauchy's key results.

Given a differential equation of order K, the $\mathcal{LT}$ results in a polynomial in s of degree K. It follows that this $\mathcal{LT}$ property is the cornerstone of why the $\mathcal{LT}$ is so important to scalar differential equations, as it was to the early analysis of Pell's equation and the Fibonacci sequence, presented in chapter 2. While the relation $e^{\jmath\theta} = \cos\theta + \jmath \sin\theta$ was first uncovered by Euler. By the time of his death the formula's significance would have been clear to him. Who first coined the terms *eigenvalue* and *eigenfunction*? The word *eigen* is a German word meaning *of one*.

Initial and final value theorems: There are much more subtle relationships between $f(t)$ and $F(s)$ that characterize $f(0^+)$ and $f(t \to \infty)$, which are known as initial value theorems. If the system under investigation has potential energy at $t = 0$, then the voltage (velocity) need not be zero for negative time. An example is a charged capacitor or a moving mass. These are important concepts, but best explored in a more in-depth treatment. They are not violations of causality.

4.7.5 Solving differential equations:

Many differential equations may be solved by assuming a power series (i.e., Taylor series) solution of the form

$$y(x) = x^r \sum_{n=0}^{\infty} c_n x^n, \tag{4.7.1}$$

with $r \in \mathbb{Z}$ and coefficients $c_n \in \mathbb{C}$. This method of Frobenius is quite general (Greenberg, 1988, p. 193).

Example: When a solution of this form is substituted into the differential equation, a recursion relationship in the coefficients results. For example, if the equation is

$$y''(x) = \lambda^2 y(x),$$

Appendix F

Eleven postulates of systems of algebraic networks

Physical systems obey basic rules that follow from the physics. It is helpful to summarize these restrictions as postulates presented in terms of a taxonomy, or categorization method, of the fundamental properties of physical systems. Eleven of these are listed in this appendix from an article by Kim and Allen (2013).

F.1 Representative system

A taxonomy of physical systems comes from a systematic summary of the laws of physics, which includes at least the eleven basic network postulates, described in Sec. 3.10.

To describe the network postulates, it is helpful to start from a two-port matrix representation as discussed in Sec. 3.8 (p. 144).

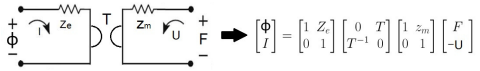


Figure F.1: *The schematic representation of an algebraic network, defined by its two-port ABCD transmission, that has three elements called the Hunt parameters (Hunt, 1952): $Z_e(s)$, the electrical impedance, $z_m(s)$, the mechanical impedance, and $T(s)$, the transduction coefficient matrix of an electromechanical transducer network. The port variables are $\Phi(f)$ and $I(f)$: the frequency domain voltage and current, and $F(f)$ and $U(f)$: the force and velocity (Hunt, 1952; Kim and Allen, 2013). This matrix factors the two-port model into three 2×2 matrices, separating the three physical elements as matrix algebra. It is a standard impedance convention that the flows $I(f) and U(f)$ are defined into each port. Thus it is necessary to apply a negative sign on the velocity $-U(f)$ so that it has an outward flow, as required to match the next cell with its inward flow.*

As shown in Fig. F.1, the two-port transmission matrix for an acoustic transducer (loudspeaker) is characterized by the equation

$$\begin{bmatrix}\Phi_i\\ I_i\end{bmatrix} = \begin{bmatrix}A(s) & B(s)\\ C(s) & D(s)\end{bmatrix}\begin{bmatrix}F_l\\ -U_l\end{bmatrix} = \frac{1}{T}\begin{bmatrix}z_m(s) & Z_e(s)z_m(s) + T^2\\ 1 & Z_e(s)\end{bmatrix}\begin{bmatrix}F_l\\ -U_l\end{bmatrix}, \tag{F.1.1}$$

shown as a product of three 2×2 matrices in the figure, with each factor representing one of the three Hunt parameters of the loudspeaker.

This figure represents the electromechanical motor of the loudspeaker and consists of three elements: the electrical input impedance $Z_e(s)$, a gyrator, which is similar to a transformer that relates current to force, and the output mechanical impedance $z_m(s)$. This circuit describes

the recursion is $c_n = c_{n-1}/n$. The resulting equation is

$$y(x) = e^{\lambda x} = x^0 \sum_{n=0}^{\infty} \frac{1}{n!} x^n,$$

namely, $c_n = 1/n!$, thus $nc_n = 1/(n-1)! = c_{n-1}$.

Exercise #3

Find the recursion relationship for $y(x) = J_\nu(x)$ of order ν that satisfies Bessel's equation

$$x^2 y''(x) + xy'(x) + (x^2 - \nu^2) y(x) = 0.$$

Solution: If we assume a complex–analytic solution of the form of Eq. 4.7.1, we find the Bessel recursion relationship for coefficients c_k (Greenberg, 1988, p. 231):

$$c_k = -\frac{1}{k(k+2\nu)} c_{k-2}.$$

■

4.8 Problems DE-3

4.8.1 Topics of this homework: Brune impedance

lattice transmission line analysis

4.8.2 Brune Impedance

Problem # *1: Residue form*

A Brune impedance is defined as the ratio of the force $F(s)$ to the flow $V(s)$ and may be expressed in residue form as

$$Z(s) = c_0 + \sum_{k=1}^{K} \frac{c_k}{s - s_k} = \frac{N(s)}{D(s)} \qquad \text{(DE-3.1)}$$

with

$$D(s) = \prod_{k=1}^{K} (s - s_k) \qquad \text{and} \qquad c_k = \lim_{s \to s_k} (s - s_k) D(s) = \prod_{n'=1}^{K-1} (s - s_n).$$

The prime on the index n' means that $n = k$ is not included in the product.

– 1.1: Find the Laplace transform ($\mathcal{LT}$) of a (1) spring, (2) dashpot, and (3) mass.

Express these in terms of the force $F(s)$ and the velocity $V(s)$, along with the electrical equivalent impedance: (1) Hooke's law $f(t) = Kx(t)$, (2) dashpot resistance $f(t) = Rv(t)$, and (3) Newton's law for mass $f(t) = Mdv(t)/dt$.

– 1.2: Take the Laplace transform ($\mathcal{LT}$) of Eq. DE-3.2 and find the total impedance $Z(s)$ of the mechanical circuit.

$$M\frac{d^2}{dt^2}x(t) + R\frac{d}{dt}x(t) + Kx(t) = f(t) \leftrightarrow (Ms^2 + Rs + K)X(s) = F(s). \qquad \text{(DE-3.2)}$$

– 1.3: What are $N(s)$ and $D(s)$ (see Eq. DE-3.1)?

– 1.4: Assume that $M = R = K = 1$ and find the residue form of the admittance $Y(s) = 1/Z(s)$ (see Eq. DE-3.1) in terms of the roots $s_\pm$. Hint: Check your answer with Octave's/Matlab's `residue` *command.*

– 1.5: By applying Eq. 4.5.3 (page 201), find the inverse Laplace transform ($\mathcal{LT}^{-1}$). Use the residue form of the expression that you derived in question 1.4.

SI Units

The pressure $\mathcal{P}$ [Pa] is a potential and the volume velocity $\mathcal{V}$ [kgm/m^2] is a flux having power $\mathcal{P}\mathcal{V}$. The corresponding thermal power is $Q = T\dot{\mathcal{S}}$ where $\dot{\mathcal{S}}$ is the entropy-rate (Q is the absorbed heat power) in Watts. Since the temperature T is a potential, analogous to a voltage, then $d\mathcal{S}/dt$ must be a flux, analogous to the current in an electrical system.

The total internal energy of a substance is defined by the enthalpy $H = PV - Q$, where $Q = T\mathcal{S}$ is the energy due to heat and PV is the mechanical energy. If H is constant and we vary the heat by a small amount, then $0 = \delta H \delta Q - \delta(PV) = \delta Q - P\delta V - V\delta P$. The thermodynamic constant c_p is defined as $\partial Q/\partial T|_{\mathcal{P}_o}$ (constant $\mathcal{P}$) and c_v is defined as $\partial Q/\partial T|_{\mathcal{V}_o}$ (constant $\mathcal{V}$).

Summary and conclusions

Since $\mathcal{P}\mathcal{V}$ is the acoustic power and $\mathcal{P}/\mathcal{V}$ the acoustic impedance, $\dot{Q} = T\dot{\mathcal{S}}$ is the thermal power, thus z_{22} is the associated thermal impedance. What is unique about this set of equations is that 1) they are linear in their variables, and that they must obey reciprocity, defined as either $|\mathcal{T}| = 1$ or $z_{12} = z_{21}$.

The transmission matrix plays a theoretical and modeling role. Thus $\mathcal{Z}$ is used for experiments and $\mathcal{T}$ for calculations. Other important matricides are the admittance matrix $\mathcal{Y}$ and $\mathcal{G}$. There are simple relationships between all these matrices (Van Valkenburg, 1964a, p. 310), three of which are:

$$\mathcal{T} = \frac{1}{z_{21}}\begin{bmatrix} z_{11} & \Delta_{\mathcal{Z}} \\ 1 & z_{22} \end{bmatrix}, \qquad \mathcal{Z} = \frac{1}{\mathcal{C}}\begin{bmatrix} \mathcal{A} & \Delta_{\mathcal{T}} \\ 1 & \mathcal{D} \end{bmatrix}, \qquad \mathcal{Y} = \frac{1}{\Delta_{\mathcal{T}}}\begin{bmatrix} z_{22} & -z_{12} \\ -z_{21} & z_{11} \end{bmatrix}.$$

Physical description of the 2x2 conjugate variables

To determine the physical meanings of the matrix elements we must start with the equation of state for the ideal gas $\mathcal{P}\,\mathcal{V} = RT$, where R is the specific gas constant, today defined in SI units as kA_v, where $k \equiv 1.380\,649 \times 10^{-23}$ [J/°K] is Boltzman's constant and A_v is Avogadro number $A_v = 8.314\,462\,618\,153\,24$ [J/°K mol]. The formal definition of the Avogadro number is the number of molecules in one gram-molecule of oxygen.

Two key thermodynamic constants c_p and c_v are

$$c_v \equiv \left.\frac{\partial u}{\partial T}\right|_{\mathcal{V}} = \left.\frac{\partial \mathcal{S}}{\partial T}\right|_{\mathcal{V}} \qquad c_p \equiv \left.\frac{\partial h}{\partial T}\right|_{\mathcal{P}} = \left.\frac{\partial \mathcal{S}}{\partial T}\right|_{\mathcal{P}}.$$

Note that $c_p - c_v \equiv R$ and c_p for H_2 at 300 [°K] is 14.31.

To identify c_p and c_v in terms of these matrix elements we need to do some algebra. The equations need to be reformulated as

$$\begin{bmatrix} \mathcal{P} \\ \dot{\mathcal{S}} \end{bmatrix} = \begin{bmatrix} z_{11} & \beta \\ \gamma & \delta \end{bmatrix}\begin{bmatrix} \mathcal{V} \\ T \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathcal{V} \\ \dot{\mathcal{S}} \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ sc_v & sc_p \end{bmatrix}\begin{bmatrix} \mathcal{P} \\ T \end{bmatrix}. \tag{E.1.1}$$

Here $s = \sigma + \omega j$ is called the Laplace frequency, common to the Laplace transform. Both c_p and c_v have units of Farads.

For incompressable liquids these are equal, thus we may define $c_l = c_p = c_v$. In this case the internal energy u and enthalpy h are given by

$$u \approx h = u_o + c_l T,$$

which means that any volume change is driven only by temperature rather than by work (Ambaum, 2010).

For water in the vapor state they are distinct with $c_p = c_v + R$ and $c_v = 5R/2$. It follows that a mixture of dry air with 5% water vapor is mainly diatomic

$$c_p/c_v = 1.4 = (5+2)/5.$$

From Eq. E.1.1, z_{11} is the reciprocal of the *bulk modulus*, defined as ratio of the volume over the pressure compliance for constant entropy (iso-thermal). Likewise, z_{22} is the ratio of $\mathcal{T}/\dot{\mathcal{S}}$ for constant volume (i.e., isobaric).

The Shockley diode equation in a semiconductor is the ratio of the electric current and the voltage across a p–n junction is $VT = k/q \approx 8.617333262 \times 10^{-5}$ [V/°K], is known as the *thermal voltage*.

Molecules possess additional internal degrees of freedom, as well as the three degrees of freedom for movement of the molecule as a whole. There are seven diatomic gas, the most important being Hydrogen H_2, Nitrogen N_2 and Oxygen O_2. These possess a total of six degrees of simple freedom per molecule due to the three translational, two rotational, and one vibrational mode. At lower temperatures, not all these degrees of freedom can fully participate in the gas' thermal heat capacity.

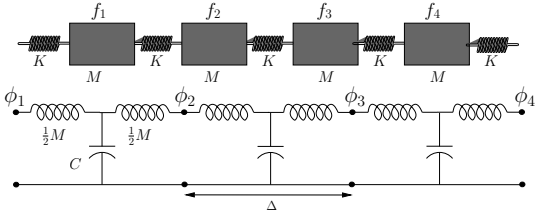


Figure 4.8: *Depiction of a train consisting of cars treated as masses M and linkages treated as springs of stiffness K or compliance $C = 1/K$. Below it is the electrical equivalent circuit for comparison. The masses are modeled as inductors and the springs as capacitors to ground. The velocity is analogous to a current and the force $f_n(t)$ to the voltage $\phi_n(t)$. The length of each cell is Δ [m]. The train may be accurately modeled as a transmission line (TL), since the equivalent electrical circuit is a lumped model of a TL. This method, called a* Cauer synthesis, *is based on the ABCD transmission line method of Sec. 3.8 (p. 144).*

4.8.3 Transmission-line analysis

Problem # *2:* **Train-mission-line** *We wish to model the dynamics of a freight train that has N such cars and study the* velocity transfer function *under various load conditions.*

As shown in Fig. 4.8, the train model consists of masses connected by springs.

Problem # *3: Transfer functions*

Use the ABCD method (see the discussion in Appendix B.3, p. 297) to find the matrix representation of the system of Fig. 4.8. Define the force on the nth train car $f_n(t) \leftrightarrow F_n(\omega)$ and the velocity $v_n(t) \leftrightarrow V_n(\omega)$.

Break the model into cells consisting of three elements: a series inductor representing half the mass ($M/2$), a shunt capacitor representing the spring ($C = 1/K$), and another series inductor representing half the mass ($L = M/2$), transforming the model into a cascade of symmetric ($\mathcal{A} = \mathcal{D}$) identical cell matrices $\mathcal{T}(s)$.

– 3.1: Find the elements of the ABCD matrix $\mathcal{T}$ for the single cell that relate the input node 1 to output node 2

$$\begin{bmatrix} F \\ V \end{bmatrix}_1 = \mathcal{T}\begin{bmatrix} F(\omega) \\ -V(\omega) \end{bmatrix}_2. \tag{DE-3.3}$$

– 3.2: Express each element of $\mathcal{T}(s)$ in terms of the complex Nyquist ratio $s/s_c < 1$ ($s = 2\pi j f$, $s_c = 2\pi j f_c$). The Nyquist wavelength sampling condition is $\lambda_c > 2\Delta$. It says the critical wavelength $\lambda_c > 2\Delta$. condition is $\lambda_c > 2\Delta$.[12] *It says the critical wavelength $\lambda_c > 2\Delta$. Namely it is defined in terms the minimum number of cells 2Δ, per minimum wavelength λ_c.*

The Nyquist wavelength sampling theorem says that there are at least two cars per wavelength.

Proof: From the figure, the distance between cars $\Delta = c_o T_o$ [m], where

$$c_o = \frac{1}{\sqrt{MC}} \quad \text{[m/s]}.$$

[12] The history of this relation has been traced back to 1841, as discussed by (Brillouin, 1953, Chap. I,II, Eq. 4.7).

The cutoff frequency obeys $f_c\lambda_c = c_o$. The Nyquist critical wavelength is $\lambda_c = c_o/f_c > 2\Delta$. Therefore the Nyquist sampling condition is

$$f < f_c \equiv \frac{c_o}{\lambda_c} = \frac{c_o}{2\Delta} = \frac{1}{2\Delta\sqrt{MC}} \quad \text{[rad/sec]}. \tag{DE-3.4}$$

Finally, $s_c = \jmath 2\pi f_c$.

– 3.3: Use the property of the Nyquist sampling frequency *$\omega < \omega_c$ (Eq. DE-3.4) to remove higher order powers of frequency*

$$1 + \cancel{\left(\frac{s}{s_c}\right)^2}^{\to 0} \approx 1 \tag{DE-3.5}$$

to determine a band-limited approximation of $\mathcal{T}(s)$.

Problem # *4: Now consider the cascade of N such $\mathcal{T}(s)$ matrices and perform an eigenanalysis.*

– 4.1: Find the eigenvalues and eigenvectors of $\mathcal{T}(s)$ as functions of s/s_c.

Problem # *5: Find the velocity transferfunction $H_{12}(s) = V_2/V_1|_{F_2=0}$.*

– 5.1: Assuming that $N = 2$ and $F_2 = 0$ (two half-mass problem), find the transfer function $H(s) \equiv V_2/V_1$. From the results of the $\mathcal{T}$ matrix, find

$$H_{21}(s) = \left.\frac{V_2}{V_1}\right|_{F_2=0}$$

Express H_{12} in terms of a residue expansion.

– 5.2: Find $h_{21}(t) \leftrightarrow H_{21}(s)$.

– 5.3: What is the input impedance $Z_2 = F_2/V_2$, assuming $F_3 = -r_0V_3$?

– 5.4: Simplify the expression for Z_2 as follows:

1. Assuming the *characteristic impedance* $r_0 = \sqrt{M/C}$,
2. terminate the system in r_0: $F_2 = -r_0V_2$ (i.e., $-V_2$ cancels).
3. Assume higher-order frequency terms are less than 1 ($|s/s_c| < 1$).
4. Let the number of cells $N \to \infty$. Thus $|s/s_c|^N = 0$.

When a transmission line is terminated in its characteristic impedance r_0, the input impedance $Z_1(s) = r_0$. Thus, when we simplify the expression for $\mathcal{T}(s)$, it should be equal to r_0. Show that this is true for this setup.

– 5.5: State the ABCD matrix relationship between the first and Nth nodes in terms of the cell matrix. Write out the transfer function for one cell, H_{21}.

– 5.6: What is the velocity transfer function $H_{N1} = \frac{V_N}{V_1}$?

fine by four differentials (Ambaum, 2010).

$$\begin{aligned} du &= \ TdS - pd\nu \\ dh &= \ TdS + \nu d\mathcal{P} \\ df &= -SdT - \mathcal{P}d\nu \\ dg &= -SdT + \nu d\mathcal{P}. \end{aligned}$$

Linear 2×2 systems of equations: The LCV thermodynamic variables are the mechanical (acoustical) pressure $\mathcal{P}$, volume velocity $\mathcal{V}$ and the heat variables, temperature T, and entropy rate $\dot{\mathcal{S}}$.

The linear transmission and impedance matrices $\mathcal{T}$ and $\mathcal{Z}$ relate the acoustic variables $\mathcal{P}$, $\mathcal{V}$ to the thermodynamic variables T, $\dot{\mathcal{S}}$ in two very different ways:

$$\begin{bmatrix} T \\ \dot{\mathcal{S}} \end{bmatrix} = \mathcal{T} \begin{bmatrix} \mathcal{P} \\ -\mathcal{V} \end{bmatrix} = \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \begin{bmatrix} \mathcal{P} \\ -\mathcal{V} \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} \mathcal{P} \\ T \end{bmatrix} = \mathcal{Z} \begin{bmatrix} \mathcal{V} \\ \dot{\mathcal{S}} \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} \mathcal{V} \\ \dot{\mathcal{S}} \end{bmatrix}.$$

On the left is the *transmission matrix* $\mathcal{T}$ and on the right is the corresponding *impedance matrix* $\mathcal{Z}$. By inspection, the $\mathcal{T}$ and $\mathcal{Z}$ matrix elements are

$$\mathcal{T} = \begin{matrix} \frac{T}{\mathcal{P}}\Big|_{\mathcal{V}=0} & \frac{T}{\mathcal{V}}\Big|_{\mathcal{P}=0} \\ \frac{\dot{\mathcal{S}}}{\mathcal{P}}\Big|_{\mathcal{V}=0} & \frac{\dot{\mathcal{S}}}{\mathcal{V}}\Big|_{\mathcal{P}=0} \end{matrix}, \mathcal{Z}_C = \begin{bmatrix} \frac{\mathcal{P}}{\mathcal{V}}\Big|_{\dot{\mathcal{S}}=0} & \frac{\mathcal{P}}{\dot{\mathcal{S}}}\Big|_{\mathcal{V}=0} \\ \frac{T}{\mathcal{V}}\Big|_{\dot{\mathcal{S}}=0} & \frac{T}{\dot{\mathcal{S}}}\Big|_{\mathcal{V}=0} \end{bmatrix}$$

Due to reciprocity, the determinant of $\mathcal{T} = 1$ and $z_{12} = z_{21}$ (Postulate P8). When using this matrix representation, any algebra is explicit. The transmission matrix $\mathcal{T}$ and impedance matrix $\mathcal{Z}$ are related by Eq. 3.8.6, as discussed on page 146.

Summary of Maxwell relations

There are six variants on the transmission matrix (Van Valkenburg, 1964a). The most common are the transmission $\mathcal{T}$ and the impedance $\mathcal{Z}$ matrices. Others, such as the admittance matrix $\mathcal{Y}$ is simply the inverse of $\mathcal{Z}$. But other variants are also used for special applications, such scattering matrices.

All the matrices used in thermodynamics are reciprocal, meaning $z_{12} = z_{21}$ or $\Delta_{\mathcal{T}} = 1$. Each matrix form show reciprocity in a different way. These expressions for reciprocity were first worked out by Maxwell, where they are called Maxwell relations (Ambaum, 2010). From the Maxwell relations, it may be shown that the specific heat capacities are equal to $c_v = y_{22} = z_{11}/\Delta_{\mathcal{Z}}$ and $c_p = g_{22} = 1/y_{22} = \Delta_{\mathcal{Z}}/z_{11}$.

Physical meaning of the matrix elements

The utility of the $\mathcal{T}$ and $\mathcal{Z}$ representations are very different. The impedance matrix $\mathcal{Z}$ always has a physically measurable interpretation. The diagonal elements z_{11}, z_{22} represent the input impedance looking into the two ports, while the off-diagonal elements are transfer impedances, related to the a transfer function between input and output fluxes and potentials. If the system is reciprocal they are equal ($z_{12} = z_{21}$).

2. *specific enthalpy* $h = u + \mathcal{P}\nu$,

3. *Helmholtz free energy* $f = u - TS$ and

4. *Gibbs function* $g = u - TS + \mathcal{P}\nu$.

The Gibbs function is a linear combination of the other three (Ambaum, 2010, Sec. 3.1). As explained by Tolman (1948), Gibbs viewed a phase transition as a membrane separating two states of a gas. These four relations could be combined into a 3×3 nonlinear energy matrix relating $[h, f, g]^T$ to $[u, \mathcal{P}\nu, TS]^T$. The entries of the matrix elements would all be either $0, 1, -1$.

When many variables are used to describe a system, as in the case of thermodynamics, a matrix formulation is more compact representation. Here traditional thermodynamic relations are reformulated using linear 2×2 matrix algebra.[a] Starting from the state equation and the Helmholtz expression for the free energy, we rewrite the system as a linear two-port network. By evaluating the impedance matrix elements, we may obtain the standard thermodynamic relations in terms of the two heat capacities c_p and c_v. Due to this unification of variables, the 2×2 formulation, such a formulation manipulation of thermodynamic relations becomes transparent.

[a] An alternative approach is to use a 3×3 scattering matrix in more than 2 variables (Bilbao, 2004).

In acoustics, the gradient of the pressure $-\nabla\,\mathcal{P}$ is the acoustic force density, which is proportional to the volume velocity $\mathcal{V}$ (the acoustic flux, or mass flow). Thus the potential and its corresponding mass flux are linear conjugate variables, the ratio of which defines the specific acoustic impedance, which for air is ρc_o. When working with impedances it is standard practice to work with the Laplace complex frequency domain s. The $\mathcal{LT}$ transforms differential equations in time into algebraic relations (i.e., polynomials). In thermodynamics the term *potential* has a very different meaning. Here we shall not refer to energies as potentials. When speaking of conjugate variables, pressure, voltage and temperature are potentials, and velocity, current and entropy are flows.

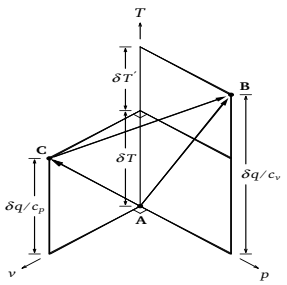
Using modern engineering terminology, the gradient of every potential is a generalized force F, which according to Ohm's law, is proportional to a generalized flow J. Working with LCV $\mathcal{P}$ and the volume velocity $s\,\mathcal{V}$ gives

$$-\nabla\,\mathcal{P} = \rho c_o s\,\mathcal{V}.$$

Here ρc_o is the impedance of air, defined as the ratio of the generalized force $F = -\nabla\,\mathcal{P}$ over the volume velocity (flow) $s\,\mathcal{V}$.

Ohm's law defines a linear relation, with the impedance being the complex proportional factor. The power is the product of LCV, which is a nonlinear relation, since it is quadratic in the conjugate variables. It follows from linear algebra that linear systems of equations are more transparent, compared, say, to the thermodynamic energy–based nonlinear of equations.

In terms of the four basic energy definitions, the force-flow relations for the four power relations are de-



Chapter 5

Stream 3B: Vector Calculus

5.1 Properties of fields and potentials

Before we can define the vector operations $\nabla(), \nabla\cdot(), \nabla\times(),$ and $\nabla^2()$, we must define the objects they operate on: scalar and vector fields. The word *field* has two very different meanings: a mathematical one, which defines an algebraic structure, and a physical one, discussed next.

Ultimately we wish to integrate in $\in \mathbb{R}^3, \mathbb{R}^n$ and $\mathbb{C}^n$. Integration is quantified by several fundamental theorems of calculus, each about integration (see pp. 179–181).

5.1.1 Scalar and vector fields: notation

Scalar fields:

We use the term *scalar field* interchangeably with *analytic* in a connected region of the spatial vector $\boldsymbol{x} = [x, y, z]^T \in \mathbb{R}^3$. In mathematics, functions that are piecewise differentiable are called *smooth*, which is different from *analytic*. A smooth function has at least one or more derivatives. Every analytic function is single-valued and is an infinitely differentiable power series (Sec. 3.2.4, p. 100).

Vector fields: A vector field is composed of three scalar fields. For example, the electric field used in Maxwell's equations, $\boldsymbol{e}(\boldsymbol{x},t) = [E_x, E_y, E_z]^T$ [V/m], has three components, each a scalar field. When the magnetic flux vector $\boldsymbol{b}(\boldsymbol{x})$ is static (Postulate P5, p. 163), the potential $\phi(\boldsymbol{x})$ [V] uniquely defines $\boldsymbol{e}(\boldsymbol{x},t)$ via the gradient,

$$\boldsymbol{e}(\boldsymbol{x},t) = -\nabla\phi(\boldsymbol{x},t) \quad \text{[V/m]}. \tag{5.1.1}$$

The electric force on a charge q is $\boldsymbol{F} = q\boldsymbol{e}$; thus $\boldsymbol{e}$ is proportional to the force, and when the medium is conductive, the current density (a flow) is $\boldsymbol{j_m} = \sigma_o \boldsymbol{e}$ [A/m^2]. The ratio of the potential to the flow is an impedance, so σ_o is a conductance.

Example: Suppose we are given the vector field in $\mathbb{R}^3$

$$\boldsymbol{a}(\boldsymbol{x}) = [\phi(\boldsymbol{x}), \psi(\boldsymbol{x}), \theta(\boldsymbol{x})]^T \quad \text{[Wb/m]},$$

where each of the three functions is a scalar field. Then $\boldsymbol{a}(\boldsymbol{x}) = [x, xy, xyz]^T$ is a legal vector field that has components analytic in $\boldsymbol{x}$.

Example: From Maxwell's equations, the magnetic flux vector is given by

$$\boldsymbol{b}(\boldsymbol{x},t) = \nabla \times \boldsymbol{a}(\boldsymbol{x},t) \quad \text{[Wb/m}^2\text{]}. \tag{5.1.2}$$

We shall see that this is always true because the magnetic charge $\nabla\cdot\boldsymbol{b}(\boldsymbol{x},t)$ must be 0, which is always true in-vacuo. Feynman (1970b, pp. 14-1 to 14-3) provides an extended and helpful tutorial on the vector potential, with many examples.

To verify that a field is a potential, we may check the units. However, a proper mathematical definition is that the potential must be an analytic function of $\boldsymbol{x}$ and t, so that we can operate on it with $\nabla()$ and $\nabla\times()$. Note that the divergence of a scalar field is not a legal vector operation.

Scalar potentials: The above discussion describes the utility of potentials for defining vector fields (e.g., Eqs. 5.1.1 and 5.1.2). The key distinction between a potential and a scalar field is that potentials have units and thus have a physical meaning. Scalar potentials (i.e., voltage $\phi(\boldsymbol{x},t)$ [V], temperature $T(\boldsymbol{x},t)$ [°C], and pressure $\varrho(\boldsymbol{x},t)$ [pascals]) are physical scalar fields. All potentials are composed of scalar fields, but not all scalar fields are potentials.

For example, the $\hat{\mathbf{y}}$ component of $\boldsymbol{e}$, $E_y(\boldsymbol{x},t) = \hat{\mathbf{y}}\cdot\boldsymbol{e}(\boldsymbol{x},t)$ [V/m], is not a potential. While $\nabla\boldsymbol{e}_y$ is mathematically defined as the gradient of one component of a vector field, it has no physical meaning (as best I know).

Vector potentials: Vector potentials, like scalar potentials, are vector fields with physically meaningful units. They are more complicated than scalar potentials because they are composed of three scalar fields. Vector fields are composed of laminar and rotational flow, which are mathematically described by the fundamental theorem of vector calculus (also called Helmholtz's decomposition theorem). One superficial but helpful comparison is the momentum of a mass, which may be decomposed into its forward (linear) and rotational momentum.

Since we find it is helpful to analyze problems using potentials (e.g., voltage) and then take the gradient (i.e., voltage difference) to find the flow (i.e., current density $\boldsymbol{j} = \sigma\boldsymbol{e}(\boldsymbol{x},t)$ $[A/m^2]$), where σ $[\mho/m^2]$ is the conductivity, the same logic and utility apply when we use the vector potential to describe the magnetic flux (flow) $\boldsymbol{b}(\boldsymbol{x},t)$ (Feynman, 1970c). When operating on a scalar potential, we use a gradient, whereas for the vector potential, we operate with the curl (Eq. 5.1.2).

Appendix E

Thermodynamic systems

Many people find thermodynamics difficult to understand. Here we explore the reasons behind this lack of transparency, and propose a solution. To understand the problem it is helpful to compare the nonlinear energy-equilibrium methods used in thermodynamic, to linear impedance methods, used in acoustics, mechanics and electrical circuit theory.

Acoustics, mechanics and electrical circuits are explained in terms of linear systems of equations, such as as Kirchhoff's and Ohm's laws, as discussed in Appendix D. The linear impedance formulation leads to a system of equations that is easily solved, using standard methods of linear algebra. Thermodynamics on the other hand is formulated in terms of equilibrium energies, resulting in a nonlinear systems of equations. This explains the difficulty in understanding the relationships. A second issue is that these equations are frequently overspecified.

Traditionally thermodynamics has been formulated with two types of variables, those that are proportional to the mass, called *extensive* such as volume and mass, and the those that are independent of mass, called *intensive* variables, such as temperature and density (Ambaum, 2010, p. 3).

A more modern and transparent notation is to work with two *conjugate* variables (CV),[1] *force density* and *flux*. Examples of the CV force density include pressure $\mathcal{P}$, temperature T, the Nernst potential, and electrical voltage. Fluxes CVs inclued mass flux and its area integral, volume velocity $\mathcal{V}$, heat flux, called entropy S, and electrical current.

Impedance is the ratio of CV, and the power is their product. For example, in electrical network theory, the impedance is the voltage over the current and the power is the product. Working with impedance always results in systems of linear equations. *linear conjugate variables* (LCV) are a generalization of force, the gradient of a potential, and flow. Namely the formulation in terms of LCV simplifies the thermodynamic system of nonlinear equations.

E.1 Summary of the thermodynamic relations

To reduce the number of variables we remove mass from the system by working with *specific* variables (those that are normalized by the mass). The ratio of the volume $\mathcal{V}$ and mass M defines the specific volume $\nu \equiv \mathcal{V}/M = 1/\rho$, where ρ is the density. Throughout this discussion we shall use SI units, with $\mathcal{V}$ [m^3], M [kgm], $\rho = 1/\nu$ [kg/m^3].

Four classical *thermodynamic energies* are defined:

1. *internal energy* u,

[1] https://en.wikipedia.org/wiki/Conjugate_variables_(thermodynamics)

In Eq. 5.1.1 we assumed that the magnetic flux vector $\boldsymbol{b}(\boldsymbol{x})$ was static, and thus $\boldsymbol{e}(\boldsymbol{x}, t)$ is the gradient of the time-dependent voltage $\phi(\boldsymbol{x}, t)$. However, when the magnetic field is dynamic (*not* static), Eq. 5.1.1 is not valid, due to magnetic induction: A voltage induced into a loop of wire is proportional to the time-varying flux cutting across that loop of wire. This is known as the *Ampere-Maxwell law*. In the static case the induced voltage is zero.

In summary, the electric field strength includes both scalar potential $\phi(\boldsymbol{x}, t)$ and magnetic flux vector potential $\boldsymbol{a}(\boldsymbol{x}, t)$ components, while the magnetic field strength depends only on the magnetic potential.

5.1.2 Gradient ∇, divergence $\nabla\cdot$, curl $\nabla\times$, and Laplacian ∇^2

Three key vector differential operators are used in linear partial differential equations, such as the wave and diffusion equations. All of these begin with the ∇ operator:

$$\nabla = \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}.$$

As outlined in Table 5.1, the official name of this operator is *nabla*. It has three basic uses: (1) the gradient of a scalar field, (2) the divergence of a vector field, and (3) the curl of a vector field. The shorthand notation $\nabla\phi(\boldsymbol{x}, t) = (\hat{\mathbf{x}}\partial_x + \hat{\mathbf{y}}\partial_x + \hat{\mathbf{z}}\partial_x)\phi(\boldsymbol{x}, t)$ is convenient.[1]

Table 5.1: *The three vector operators manipulate scalar and vector fields. The gradient converts scalar fields into vector fields. The divergence maps vector fields to scalar fields. The curl maps vector fields to vector fields. Four second-order operators (for example DoG and* **gOd***) are defined in Sec. 5.6.9 (p. 263). Bold mnemonics are reserved for vector-in, vector-out operators, with the curl being an exception.*
Several entries similar to **CoC***, seem feasible:*
Wedge of Wedge **WoW**,
Curl of Wedge **CoW**,
Wedge of Curl **WoC**=*0*,
Wedge of Div **WoD**,
Curl of Wedge **CoW =0**.

Name	Input	Output	Operator	Mnemonic
Gradient	Scalar	**Vector**	$\nabla(\,)$	grad
Divergence	**Vector**	Scalar	$\nabla\cdot(\,)$	div
Laplacian	Scalar	Scalar	$\nabla\cdot\nabla = \nabla^2(\,)$	DoG
Wedgie, CoW	**Vector**	Scalar	$\nabla\wedge(\,)$	CoW,**wedge**
Curl	**Vector**	**Vector**	$\nabla\times(\,)$	curl
God	**Vector**	**Vector**	$\boldsymbol{\nabla}^2(\,) = \nabla(\nabla\cdot(\,))$	**gOd**
Bull-DoG	**Vector**	**Vector**	$\boldsymbol{\nabla}^2(\,) = \nabla\cdot\nabla(\,)$	**DoG**
Curl of Curl	**Vector**	**Vector**	$\nabla\times\nabla\times = \boldsymbol{\nabla}^2(\,) - \boldsymbol{\nabla}^2(\,)$	**CoC**
Div of Curl	**Vector**	0	$\nabla\cdot\nabla\times()$	**DoC**
Curl of Grad	Scalar	0	$\nabla\times\nabla()$	**CoG**
Curl of Wedge	**Vector**	Scalar	$\approx$**COC**	CoW() or COW()

[1] https://en.wikipedia.org/wiki/Del_in_cylindrical_and_spherical_coordinates

Gradient $\nabla()$:

As shown in Fig. 5.1, the gradient transforms a complex scalar field $\Phi(\boldsymbol{x}, s) \in \mathbb{C}$ into a vector field ($\mathbb{C}^3$)

$$\begin{aligned}\nabla\Phi(\boldsymbol{x}, s) &= \left(\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}\right)\Phi(\boldsymbol{x}, s) \\ &= \hat{\mathbf{x}}\frac{\partial\Phi}{\partial x} + \hat{\mathbf{y}}\frac{\partial\Phi}{\partial y} + \hat{\mathbf{z}}\frac{\partial\Phi}{\partial z}.\end{aligned}$$

The gradient may also be factored into a unit vector $\hat{\mathbf{n}}$, as defined in Fig. 5.1, that gives the direction of the gradient, and the gradient's length $||\nabla()||$, defined in terms of the norm of the gradient. Thus the gradient of $\Phi(\boldsymbol{x})$ may be written in "polar coordinates" as $\nabla\Phi(\boldsymbol{x}) = ||\nabla\Phi||\,\hat{\mathbf{n}}$, which leads to the unit vector

$$\hat{\mathbf{n}} = \frac{\nabla(\Phi(\boldsymbol{x}))}{||\nabla\Phi||}.$$

Consider the paraboloid $z = 1 - (x^2 + y^2)$ as the potential, with isopotential circles of constant z that have a radius of zero at $z = 1$ and unit radius at $z = 0$. The negative gradient

$$\boldsymbol{e}(\boldsymbol{x}) = -\nabla z(x, y) = 2(x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + 0\hat{\mathbf{z}})$$

is $\perp$ to the circles of constant radius (constant z) and thus points in the direction of the radius.

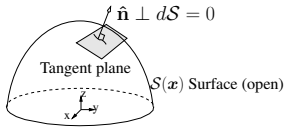


Figure 5.1: *Definition of the unit vector $\hat{\mathbf{n}}$ defined by the gradient $\nabla\mathcal{S} \perp$ to the tangent plane. A bifurcated volume defines surface $\mathcal{S}(\boldsymbol{x})$. At one point a tangent plane (shaded) touches the surface. At that point the gradient $\nabla\mathcal{S}(\boldsymbol{x})$ is normalized to unit length, defining $\hat{\mathbf{n}}$, which is perpendicular ($\perp$) to the shaded tangent plane.*

A skier in free fall on this surface would be the first one down the hill. Normally skiers try to stay close to the isoclines (not in the direction of the gradient) so they can stay in control. If you ski on an isocline, you must walk, since there is no pull due to gravity. The gravitational potential at the surface of the earth is

$$\phi = G\frac{mM}{r}.$$

Divergence $\nabla\cdot()$:

The divergence of a vector field results in a scalar field. For example, the divergence of the electric field flux vector $\boldsymbol{d}(x)$ [C/m^2] equals the scalar field charge density $\rho(\boldsymbol{x})$ [C/m^3]:

$$\nabla\cdot\boldsymbol{d}(\boldsymbol{x}) \equiv \left(\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}\right)\cdot\boldsymbol{d}(\boldsymbol{x}) = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} = \rho(\boldsymbol{x}). \tag{5.1.3}$$

Thus the divergence is analogous to the scalar (dot) product (e.g., $\boldsymbol{a}\cdot\boldsymbol{b}$) between two vectors.

Recall that the voltage is the line integral of the electric field,

$$V(a) - V(b) = \int_a^b \boldsymbol{e}(\boldsymbol{x})\cdot d\boldsymbol{x} = -\int_a^b \nabla V(\boldsymbol{x})\cdot d\boldsymbol{x} = -\int_a^b \frac{dV}{dx}dx, \tag{5.1.4}$$

Pressure Eigen-solutions: The forwarded P_- and backward P_+ pressure waves propagate as

$$P_\pm(s, x) = e^{-\kappa(s)x},\ \ e^{-\overline{\kappa}(s)x}, \tag{D.1.11}$$

where $\overline{\kappa}(s)$ the complex conjugate of $\kappa(s)$, such that $\Re\kappa(s) > 0$. The term $\beta_o\sqrt{s}$ affects both the real and imaginary parts of $\kappa(s)$. The real part is a frequency-dependent loss, and the imaginary part introduces a frequency-dependent speed of sound (Mason, 1928).

D.1.2 Impact of viscous and thermal losses

Equation D.1.2 and the measured data are compared in Fig. D.1, as reproduced from Mason's Fig. 4, which shows that the wave speed drops from 344 m/s at 2.6 kHz to 339 m/s at 0.4 kHz, a 1.5% reduction. At 1 kHz the loss is 1 dB/m for a 7.5-mm tube. Note that the loss and the speed of sound vary inversely with the radius. As the radius approaches the *boundary layer thickness* (i.e., the radial distance such that the loss is e^{-1}), the effect of the damping dominates the propagation.

Cut-off frequency s_o: The frequency where the lossless part equals the lossy part is defined as $\kappa(s_o) = 0$, namely, $\sqrt{s_o} = -\beta_o$, or $s_o = \beta_o^2$.

To get a feeling for the magnitude of s_o, let $R = 0.75/2$ [cm] (i.e., the average radius of the adult ear canal). Then for $R = 3.75 \times 10^{-3}$ [cm]

$$s_o = (1.9 \times 10^{-3}/3.75 \times 10^{-3})^2 = 1/4.$$

We conclude that the losses are insignificant in the audio range, since for the human ear canal, $f_o = \beta_o^2/\pi \approx 0.25/\pi = 0.08$ Hz.[3] This frequency represents the lower bound of the transition from adiabatic to iso-thermal equilibrium. It should be clear that acoustic frequencies do not actually obey the adiabatic approximation, due to the thin boundary layer. Both the real and imaginary part of propagation function $\kappa(s)$, the characteristic impedance $z_o(s)$ and the speed of sound $c_o'(s)$ all depended on frequency in the auditory range of human hearing.

Summary: The Helmholtz-Kirchhoff theory of viscous and thermal losses results in a frequency-dependent speed of sound that has a frequency dependence proportional to $1/\sqrt{s}$ rather than $1/\sqrt{\omega}$ (Mason, 1928, Eq. 4). This corresponds to a 2% change in the sound velocity over the decade from 0.2 to 2 kHz (Mason, 1928, Fig. 5), in agreement with Mason's experimental results.

[3] `/home/jba/Mimosa/2C-FindLengths.16/doc.2-c_calib.14/m/MasonKappa.m`

Mason (1928) assumed the viscosity to be $\mu_o = 18.6 \times 10^{-6}$ [Pa-s] (viscosity), thus the dynamic-to-adiabatic visco-thermal elastic ratio is

$$\frac{\eta'_o}{\eta_o} = \frac{1.5345}{1.4}\sqrt{\mu_o} = 1.0961\sqrt{\mu_o},$$

giving

$$2\beta_o = 1.0961\frac{P}{2S}\quad\overline{\frac{\mu_o}{\rho_o}}.$$

Here μ_o/ρ_o is known as the *kinematic viscosity* and μ_o as the *dynamic viscosity*.

Reduction of the lossy characteristic impedance: Rendering Eq. D.1.3 dimensionless, its complex–analytic expression greatly simplifies

$$\frac{z_o(\omega)}{\sqrt{P_o\eta_o\rho_o}} - 1 = \frac{2\beta_o}{\sqrt{\omega}}\frac{1-\jmath}{\sqrt{2}} = \frac{2\beta_o}{\sqrt{s}}\sqrt{\jmath}\sqrt{-\jmath}\,1 = \frac{2\beta_o}{\sqrt{s}}, \tag{D.1.9}$$

thus the lossy normalized characteristic impedance is $z_o(\omega)/r_o = 1+\frac{2\beta_o}{\sqrt{s}}$, where $r_o = \sqrt{P_o\eta_o\rho_o}$ is the lossless characteristic resistance, and $2\beta_o$ is defined in Eq. D.1.7.

Reduction of the lossy speed of sound: Finally, starting from Eq. D.1.4,

$$c'_o(\omega) = c_o\left\{1 - \frac{P\eta'_o}{2S\sqrt{2\omega\rho_o}}\right\}, \tag{D.1.10}$$

note that for $\omega > 0$, $\left(1 - \frac{c'_o}{c_o}\right) > 0$, thus

$$1 - \frac{c'_o(\omega)}{c_o} = \frac{P\eta'_o}{2S\sqrt{2\omega\rho_o}} = \frac{P}{2S}\frac{\eta'_o\sqrt{\jmath}}{\sqrt{2s\rho_o}} = \frac{P}{2S}1.0961\frac{\eta_o}{\sqrt{s}}\quad\overline{\frac{\mu_o}{\rho_o}}\quad\overline{\frac{\jmath}{2}} = 2\beta_o\frac{\eta_o}{\sqrt{s}}\quad\overline{\frac{\jmath}{2}}.$$

For $\omega < 0$, this will differ, and for $\omega = 0$ it is singular. This is because of the singular branch cut at $s = 0$, due to $1/\sqrt{s}$. The inverse $\mathcal{LT}$ of $1/\sqrt{s}$ is provided in Appendix C, Sect. C.2.

Case of the cylindrical guide: For the case of a cylindrical wave guide, $P/2S = 1/R$. Thus

$$2\beta_o R = \eta'_o/\sqrt{\rho_o} = 1.0961\,\eta_o\quad\overline{\mu_o/\rho_o} = 1.0961 \times 1.4\quad\overline{\frac{18.6\times10^{-6}}{1.2}} = 1.9105\times10^{-3}.$$

It is well documented in the literature that the boundary layer thickness is proportional to the square root of the kinematic viscosity (μ/ρ) over frequency. Our result, Eq. D.1.5, is a complex-analytic extension of this classic boundary layer equation, derived from the classic results of Kirchhoff (1868) and Rayleigh (1896), as verified by Mason (1928).

Ignores of the lossy wave equation: We may factor Eq. D.1.5 to reveal the mathematical impact of the damping on $\kappa(s)$

$$\sqrt{s_o} = -\beta_o \pm \beta_o,$$

namely $\sqrt{s_o} = \{0, -\beta_1\}$. In general there must be four roots, but since $\beta_1 \in \mathbb{R} > 0$, the roots in this case are degenerate, and in the left half plane.

The smaller the radius the greater the damping ($\beta_1 = 1.1 \times 10^{-3}/R$). Also note that the propagation function $\kappa(\omega)$ has a Helmholtz-Kirchhoff correction for both the real and imaginary parts. Thus both the speed of sound and the damping are dependent on frequency, and in a similar way.

which is simply the fundamental theorem of calculus (see p. 179). In a charge-free region, this integral is independent of the path from a to b, which is a property of a conservative system.

When we work with guided waves (narrow tubes of flux) having rigid walls that block the flow, such that the diameter is small compared with the wavelength (Postulate P10, p. 163), the divergence simplifies to

$$\nabla\cdot\boldsymbol{d}(\boldsymbol{x}) = \nabla_r D_r = \frac{1}{A(r)}\frac{\partial}{\partial r}A(r)D_r(r), \tag{5.1.5}$$

where r is the distance down the horn (range variable), $A(r)$ is the area of the isoresponse surface as a function of the range r, and $D_r(r)$ is the radial component of vector $\boldsymbol{d}$ as a function of the range r. In spherical, cylindrical, and rectangular coordinates, Eq. 5.1.5 provides the correct expression (Table 5.2, p. 247).

Properties of the divergence: The divergence is a direct measure of the flux density of the vector field. A vector field is said to be *incompressible* if the divergence of that field is zero. It is therefore *compressible* when the divergence is nonzero [e.g., $\nabla\cdot\boldsymbol{d}(x,s) = \rho(\boldsymbol{x},s)$] (see Table 5.3, page 260).

For example, compared to air, water is considered to be incompressible. The stiffness of a fluid (i.e., the bulk modulus) is a measure of its compressibility. At very low frequencies, air may be treated as incompressible (like water), since as $s \to 0$,

$$-\nabla\cdot\boldsymbol{u}(\boldsymbol{x},s) = \frac{s}{\eta_o P_o}\mathcal{P}(\boldsymbol{x},s) \to 0.$$

The definition of *compressible* depends on the wavelength in the medium, so the term must be used with some awareness of the frequencies being in the analysis. As a rule of thumb, if the wavelength $\lambda = c_o/f$ is much larger than the size of the system, the medium may be modeled as an incompressible fluid.

Curl $\nabla\times()$:

The curl ($\nabla\times()$) takes a vector in $\mathbb{C}^3$ into a second vector in $\mathbb{C}^3$. For example, in the case of fluids, the vorticity is defined as $\boldsymbol{\omega} = \nabla\times\nu$ and *rotation* as $\boldsymbol{\Omega} = \boldsymbol{\omega/2}$. The curl is a measure of the rotation of a vector field in a plane about the axis perpendicular to that plane. In the case of a liquid, it corresponds to the angular momentum, such as in a whirlpool in water, or a tornado in air. A massive top falls over when not spinning. But once spinning, it can stably stand on its pointed tip. These systems are stable due to conservation of angular momentum.

The curl and the divergence are both key operations when we working with Maxwell's four equations. For example, the curl transforms the vector field $\boldsymbol{h}(\boldsymbol{x},s) \in \mathbb{C}^3$ [A/m] into a complex vector current density $\boldsymbol{c}(\boldsymbol{x},s) \in \mathbb{C}^2$ [A/m^2]:

$$\nabla\times\boldsymbol{h}(\boldsymbol{x},s) \equiv \begin{vmatrix}\hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}}\\ \partial_x & \partial_y & \partial_z\\ H_x & H_y & H_z\end{vmatrix} = \boldsymbol{c}(\boldsymbol{x},s) \qquad [\text{A/m}^2]. \tag{5.1.6}$$

The notation $|\cdot|$ indicates the determinant (Appendix A.3.1, p. 289), ∂_x is shorthand for $\partial/\partial x$, and $\boldsymbol{h} = [H_x, H_y, H_z]^T$.

Exercise #1

If we let $\boldsymbol{h} = -y\hat{\mathbf{x}} + x\hat{\mathbf{y}} + 0\hat{\mathbf{z}}$, $\nabla \times \boldsymbol{h} = 2\hat{\mathbf{z}}$, thus $\boldsymbol{h}$ has a constant rotation; when $\boldsymbol{h} = 0\hat{\mathbf{x}} + 0\hat{\mathbf{y}} + z^2\hat{\mathbf{z}}$, $\nabla \times \boldsymbol{h} = \mathbf{0}$ has a curl of zero and thus is irrotational.

There are simple rules that precisely govern when a vector field is rotational versus irrotational, and compressible versus incompressible. These classifications are dictated by Helmholtz's theorem, the fundamental theorem of vector calculus (Eq. 5.6.11, p. 261).

Wedge $\nabla \wedge ()$:

A special case of the curl is the two-dimensional differential wedge products[2]

$$\nabla_x \wedge \boldsymbol{h}(\boldsymbol{x}, t) = \begin{vmatrix} \partial_y & \partial_z \\ H_y & H_z \end{vmatrix} = C_x(\boldsymbol{x}, s) \qquad [\text{A/m}^2].$$

The curl is made up of three such differential wedge products.[3]

Laplacian $\nabla^2()$:

The Laplacian operator $\nabla^2 \equiv \nabla \cdot \nabla$ (Table 5.1, page 225) is defined as the divergence of the gradient

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{5.1.7}$$

Since the Laplacian does so much common work, we nickname it DoG (Div of Grad).

Starting from a scalar field, the gradient produces a vector, which is then operated on by the divergence to take the output of the gradient back to a scalar field. Thus the Laplacian transforms a scalar field back to a scalar field. We have seen the Laplacian before when we defined complex–analytic functions (Eq. 4.2.8, p. 182).

A classic example of the Laplacian is a voltage scalar field $\Phi(\boldsymbol{x})$ [V], which results in the electric field vector

$$\boldsymbol{e}(\boldsymbol{x}) = [E_x(\boldsymbol{x}), E_y(\boldsymbol{x}), E_z(\boldsymbol{x})]^T = -\nabla\Phi(\boldsymbol{x}) \quad [\text{V/m}].$$

When this is scaled by the permittivity, we obtain the electric flux $\boldsymbol{d} = \epsilon_o \boldsymbol{e}$ [C/m^2], the charge density per unit area. Here ϵ_o [F/m] is the vacuum permittivity $\epsilon_o = 1/c_o r_o \approx 8.85 \times 10^{-12}$ [F/m].

Taking the divergence of $\boldsymbol{d}$ results in the charge density $\rho(\boldsymbol{x})$ [C/m^3] at $\boldsymbol{x}$:

$$\nabla \cdot \boldsymbol{d} = \nabla^2\Phi(\boldsymbol{x}) = \rho(\boldsymbol{x}).$$

Thus the Laplacian of the voltage, scaled by ϵ_o, results in the local charge density.

Another classic example of the Laplacian is an acoustic pressure field $\varrho(\boldsymbol{x}, t)$ [Pa], which defines a vector force density $\mathbf{f}(\boldsymbol{x}, t) = -\nabla\varrho(\boldsymbol{x}, t)$ [N/m^2] (Eq. 5.2.5, p. 239). When this force density [N/m^2] is integrated over an area, the net radial force [N] is

$$F_r = -\int_{\mathcal{S}} \nabla\varrho(\boldsymbol{x}) d\boldsymbol{x} \quad [\text{N}]. \tag{5.1.8}$$

[2] https://en.wikipedia.org/wiki/Triple_product#As_an_exterior_product

[3] This notation suggests that $||\nabla \cdot \boldsymbol{e} + \jmath \wedge \boldsymbol{e}||^2 = ||\nabla \cdot \boldsymbol{e}||^2 + ||\wedge \boldsymbol{e}||^2$, which is related to Helmholtz's theorem.

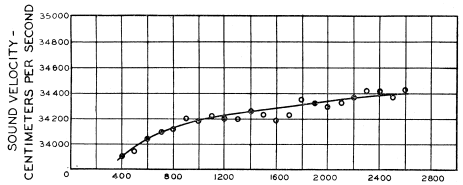


Figure D.2: *Figure 5 from Mason (1928) showing the velocity of sound as a function of frequency, when visco-thermal losses are included. Note the change in the slope around 0.8 [kHz].*

Reduction of Kirchhoff's equations to complex–analytic form: We may rewrite κ and z_o in terms of β_o

$$c_o\kappa_\pm(s) = \left(\beta_o \pm \sqrt{s}\right)^2 \tag{D.1.5}$$

Review again

The two boxed equations below provide the derivation for the reduction of Mason's formula, for $c_o\kappa(\omega)$, to its complex–analytic form (Eq. D.1.5). The first step is to define $2\beta_o$ and $c_o\kappa(\omega) - s$. The second box shows it is equal to $2\beta_o\sqrt{s}$. Thus $\kappa(s)$ is complex–analytic in the Laplace frequency s, and following a completion of squares in $\sqrt{s}$, is given by Eq. D.1.5. The inverse Laplace transforms of $\sqrt{s}$ and $1/\sqrt{s}$ are provided in Table C.2 of Appendix C and Table 3.9 of Sect. 3.10.1.

STARTING FROM MASON (1928) $c_o\kappa(\omega)$:

$$c_o\kappa(\omega) = \frac{P\eta_o'\sqrt{\omega}}{2S\sqrt{2\rho_o}} + \jmath\omega\left\{1 + \frac{P\eta_o'}{2S\sqrt{2\omega\rho_o}}\right\} \tag{D.1.6}$$

DEFINE VARIABLES $2\beta_o$ AND $s = \jmath\omega$:

$$c_o\kappa(\omega) = s + \frac{P\eta_o'}{2S\sqrt{\rho_o}}\left[\frac{2\beta_o}{\sqrt{2}} + \frac{s}{\sqrt{2\omega}}\right] \tag{D.1.7}$$

THUS

$$c_o\kappa(\omega) - s = \frac{2\beta_o}{\sqrt{2}}\left[\sqrt{\omega} + \frac{s}{\sqrt{\omega}}\right] \tag{D.1.8}$$

MULTIPLYING TOP AND BOTTOM ON RIGHT BY $\sqrt{\jmath}$:

$$\begin{aligned} c_o\kappa(\omega) - s &= \frac{2\beta_o}{\sqrt{2}}\left[\frac{\sqrt{\jmath\omega}}{\sqrt{\jmath}} + \frac{s\sqrt{\jmath}}{\sqrt{\jmath\omega}}\right] && \text{SET } \sqrt{\jmath\omega} = \sqrt{s} \\ &= \frac{2\beta_o}{\sqrt{2}}\left[\frac{\sqrt{s}}{\sqrt{\jmath}} + \frac{s\sqrt{\jmath}}{\sqrt{s}}\right], && \text{CROSS MULTIPLY} \\ &= \frac{2\beta_o}{\sqrt{2}}\left[\frac{s + \jmath s}{\sqrt{\jmath}\cdot\sqrt{s}}\right], && \text{FACTOR OUT } \frac{s}{\sqrt{s}} \\ &= 2\beta_o\left[\frac{1+\jmath}{\sqrt{2\jmath}}\right]\frac{s}{\sqrt{s}} && \text{REPLACE } \frac{s}{\sqrt{s}} = \sqrt{s} \\ &= 2\beta_o\sqrt{s} \end{aligned}$$

TO SHOW $1 + \jmath = \sqrt{2\jmath}$, SQUARE BOTH SIDES: $1 + 2\jmath = 2\jmath$.

Acoustic constants for air: Assuming $\eta_o = 1.4$ (ratio of specific heats), $\rho_o = 1.2$ [kgm/m^3] (density), a temperature of 23.5 [°C], and $P_o = 10^5$ [Pa] (atmospheric pressure), the lossless sound velocity is $c_o = \sqrt{P_o\eta_o/\rho_o} = 341.57$ [m/s]. By a comparison of this value of c_o to Fig. D.2, it is clear that this value does not agree with Mason's measurements. Thus to agree with his experimental results, either the ratio P_o/ρ_o or $\eta_o = 1.4$ need to be corrected. Since $2\beta_o$ depends on η_o', ρ_o and μ_o this is a good place to look for the discrepancy.

Correction for η_o': The dimensionless constant $\eta_o'/\sqrt{\mu}$ is defined as the *composite thermodynamic constant* (Kirchhoff, 1868; Rayleigh, 1896)

$$\frac{\eta_o'}{\sqrt{\mu_o}} = 1 + \sqrt{5/2}\left(\eta_o^{1/2} - \eta_o^{-1/2}\right) = 1.5345.$$

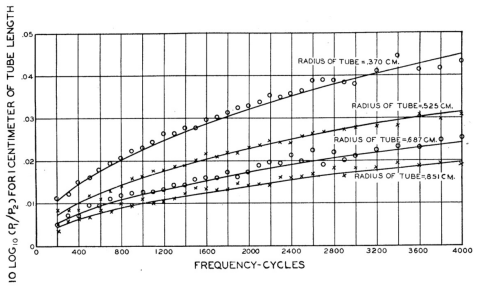


Figure D.1: *This figure, taken from Mason (1928), compares the Helmholtz-Kirchhoff theory for $|\kappa(f)|$ to Mason's 1928 experimental measurements of the loss. The ratio of two powers (P_1 and P_2) is plotted (see Mason's discussion immediately below his Fig. 4), and as indicated in the label: "$10\log_{10} P_1/P_2$ for 1 [cm] of tube length." This is a plot of the transmission power ratio in [dB/cm] which is $10\log|\Gamma(\omega)|^2$ where $\Gamma(i\omega)$ the reflection coefficient . For a discussion of the reflection coefficient and its properties, refer to Sec. 4.4.1, p. 189.*

1922; Stewart and Lindsay, 1930) was acoustic version of the work of George Campbell (1904-1923) on electrical wave filters. If today you design earphones and hearing aids, or otoacoustic research, the works of Steward and Mason are relevant.

Mason's specification of the propagation function

Mason's results are reproduced in Fig. D.1 for tubes of radii between 3.7 and 8.5 [mm] of lengths L, having power reflectance

$$|\Gamma_L(f)|^2 = \left|e^{-\kappa(f)\,L}\right|^2 \text{ [cm}^{-1}\text{]}. \tag{D.1.1}$$

The complex propagation function use by Mason (1928), as taken from (Rayleigh, 1896, p. 319), was

$$\kappa(\omega) = \frac{P\eta_o'\sqrt{\omega}}{2c_oS\sqrt{2\rho_o}} + \frac{i\omega}{c_o}\left\{1 + \frac{P\eta_o'}{2S\sqrt{2\omega\rho_o}}\right\}, \tag{D.1.2}$$

and the characteristic impedance is

$$z_o(\omega) = \quad \overline{P_o\eta_o\rho_o}\left\{1 + (1-j)\frac{P\eta_o'}{2S\sqrt{2\omega\rho_o}}\right\}, \tag{D.1.3}$$

where $S = \pi R^2$ is the tube area and $P = 2\pi R$ is its perimeter.

Following Mason (1928, Fig. 5), the measured speed of sound

$$c_o'(\omega) = c_o\left\{1 - \frac{P\eta_o'}{2S\sqrt{2\omega\rho_o}}\right\}. \tag{D.1.4}$$

depends on frequency ω, as derived from the imaginary part of Eq. D.1.2. Figure D.2 directly compares Mason's measured sound speed with this equation.

An inflated balloon with a static internal pressure of 3 [atm], in an ambient pressure of 1 [atm] (sea level), forms a sphere due to the elastic nature of the rubber, which acts as a stretched spring under its surface tension. The net normal force on the surface of the balloon is its area times the pressure drop of 2 [atm] across the surface. Thus the static pressure is

$$\varrho(\boldsymbol{x}) = 3u(r_o - r) + 1 \quad \text{[Pa]},$$

where $u(r)$ is a step function of the radius $r = ||\boldsymbol{x}|| > 0$, centered at the center of the balloon, having radius r_o.

Taking the gradient gives the negative[4] of the radial force density (i.e., perpendicular to the surface of the balloon):

$$-f_r(r) = \nabla\varrho(\boldsymbol{x}) = \frac{\partial}{\partial r}3u(r_o - r) + 1 = -2\delta(r_o - r) \quad \text{[Pa]}.$$

This equation describes a static pressure that is 1 [atm] (10^5 [Pa]) outside the balloon and 3 [atm] inside. The net positive force density is the negative of the gradient of the static pressure.

Finally, taking the divergence of the force produces a double delta function at the balloon's surface. Specifically, $\nabla^2\varrho(\boldsymbol{x}) = -2\delta^{(1)}(r_o - r)$, where 2 is the pressure drop across the balloon. If we take the thickness of the rubber (l [m]) into account, then $\nabla^2\varrho = -2(\delta(r_o) - \delta(r_o - l))$.

Vector Laplacian $\boldsymbol{\nabla}^2()$:

A second form of the Laplacian is the *vector Laplacian* $\boldsymbol{\nabla}^2()$, defined as the divergence of the gradient $\boldsymbol{\nabla}^2() \equiv \nabla\cdot\nabla()$, thus nicknamed **Bull-Dog**, operates on a vector to produce a vector (Table 5.1, page 225). We shall need this when working with Maxwell's equations.

5.1.3 Scalar Laplacian operator in N dimensions

In general, it may be shown that in $N = 1, 2, 3$ dimensions (Sommerfeld, 1949, p 227),

$$\nabla_r^2\,\mathcal{P} \equiv \frac{1}{r^{N-1}}\frac{\partial}{\partial r}\left(r^{N-1}\frac{\partial\,\mathcal{P}}{\partial r}\right). \tag{5.1.9}$$

For each value of N, the area $A(r) = A_o r^{N-1}$. This result will turn out to be useful when we work with the Laplacian in one, two, and three dimensions. This naturally follows from Eq. 5.2.10, p. 241:

$$\frac{1}{A(r)}\frac{\partial}{\partial r}\left[A(r)\frac{\partial}{\partial r}\right]\varrho(r,t) = \frac{1}{c_o^2}\frac{\partial^2}{\partial t^2}\varrho(r,t) \leftrightarrow \frac{s^2}{c_o^2}\,\mathcal{P}(r,s).$$

Example: When $N = 3$ (i.e., spherical geometry), $A(r) = A_o r^2$, thus

$$\nabla_r^2\,\mathcal{P} \equiv \frac{1}{r^2}\partial_r r^2\partial_r\,\mathcal{P} \tag{5.1.10}$$

$$= \frac{1}{r}\frac{\partial^2}{\partial r^2}r\,\mathcal{P}, \tag{5.1.11}$$

[4] The force is pointing out, stretching the balloon.

resulting in the general d'Alembert solutions (Eq. 4.4.1 p. 188) for the spherical wave equation,

$$\mathcal{P}^{\pm}(r,s) = \frac{1}{r}e^{\mp\kappa(s)r}$$

where $\kappa(s) = s/c_o$.

Exercise #2

Prove the result of the previous example by expanding Eqs. 5.1.10 and 5.1.11 using the chain rule.

Solution: Expanding Eq. 5.1.10:

$$\begin{aligned}\frac{1}{r^2}\partial_r r^2 \partial_r \mathcal{P} =& \frac{1}{r^2}\left(2r + r^2\partial_r\right)\partial_r \mathcal{P}\\ =& \frac{2}{r}\mathcal{P}_r + \mathcal{P}_{rr}.\end{aligned}$$

Expanding Eq. 5.1.11, we obtain

$$\begin{aligned}\frac{1}{r}\partial_{rr} r\,\mathcal{P} =& \frac{1}{r}\partial_r\left(\mathcal{P} + r\,\mathcal{P}_r\right)\\ =& \frac{1}{r}\left(\mathcal{P}_r + \mathcal{P}_r + r\,\mathcal{P}_{rr}\right)\\ =& \frac{2}{r}\mathcal{P}_r + \mathcal{P}_{rr}.\end{aligned}$$

Thus the two are equivalent. ■

Summary

The radial component of the Laplacian in spherical coordinates (Eq. 5.1.10) simplifies to

$$\nabla_r^2 \varrho(\boldsymbol{x}) = \frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\varrho(\boldsymbol{x}) = \frac{1}{r}\frac{\partial^2}{\partial r^2}r\varrho(\boldsymbol{x}).$$

Since DoG is $\nabla^2 = \nabla\cdot\nabla$, it follows that the net force $\boldsymbol{f}(\boldsymbol{x}) = [F_r, 0, 0]^T$, Eq. 5.1.8 in spherical coordinates has a radial component F_r and angular component of zero. Thus the force across a balloon may be approximated by a delta function across the thin sheet of stretched rubber.

We can extended the preceding example in an interesting way to the case of a rigid hose, a rigid tube, that terminates on the right in an elastic medium (the above example of a balloon), such as an automobile tire. On the far left let's assume there is a pump injecting the fluid into the rigid hose. Consider two different fluids: air and water. Air is treated as a compressible fluid, whereas water is incompressible. However, such a classification is relative, determined by the relative compliance of the balloon (i.e., tire) at the relatively rigid pump and hose.

This is a special case of a more general situation: When a fluid is treated as incompressible (rigid), the speed of sound becomes infinite, and the wave equation is an invalid description. In this case the motion is best approximated by Laplace's equation. This represents the transition from short to long wavelengths, from wave propagation having delay, to quasistatics, having no delay.

This example may be modeled as either an electrical or a mechanical system. While the two systems are very different in their physical realization, they are mathematically equivalent, forming a perfect analog. If we take the electrical analog, the pump is a current source, injecting

Appendix D

Visco-thermal losses

D.1 Adiabatic approximation at low frequencies

Newton's early development understandably ignored viscous and thermal losses, by assuming iso-thermal conditions. But starting at very low frequencies, the isothermal assumption breaks down. Modern theory, for audio frequencies, assumes the adiabatic approximation, and is thus described by the scalar wave equation (Pierce, 1981). But it turns out that even at audio frequencies, the adiabatic approximation is invalid. This was first shown by Kirchhoff, but was not fully appreciated for more than a century, due to mathematical difficulties, which we now believe can be explained, as discussed next.

Following Helmholtz (1858), as extended by Kirchhoff (1868), visco-thermal loss mechanisms are related. The full theory was first worked out by Kirchhoff (1868, 1974). To understand how the are related is complicated, due to both the history and the mathematics, as briefly discussed by Pierce (1981). Both forms of damping are caused by two different, but coupled, diffusion effects: (1) viscous effects, due to shear at the container walls, and (2) thermal effects, due to deviations from adiabatic expansion (Kirchhoff, 1868, 1974). I believe that Einstein was eventually involved, following his studies on Brownian motion (Einstein, 1905).[1,2]

These two loss mechanism are restricted to a thin region called the *boundary layer*, which critically depends on the square root of the Laplace frequency. A key transition region is where the acoustic wavelength approaches the complex boundary layer thickness. When the radius of the container (a horn) approaches the viscous boundary layer, the theory breaks down.

D.1.1 Lossy wave-guide propagation

The formulation of visco-thermal loss in air transmission was first worked out by Helmholtz (1863a) and then extended by Kirchhoff (1868) to include thermal damping (Rayleigh, 1896, Vol. II, p. 319). These losses are accurately represented by the complex–analytic propagation function $\kappa(s)$ (Eq. D.1.5). Following his review of these theories, Crandall (1926, Appendix A), the head of the 1926 Acoustic Research Department at that time, noted that the "Helmholtz-Kirchhoff" theory had never been experimentally verified. Acting on Crandall's suggestion, Physicist Warren Mason set out to experimentally verify Kirchhoff's 60 year old theory. Mason's analysis consumed several years (Mason, 1928).

This was a continued effort. Mason (1927) extended earlier work of Stewart's on acoustic transmission lines, by including viscous and thermal losses. Stewart's acoustic theory (Stewart,

[1] See Ch. 3 of `https://www.ks.uiuc.edu/Services/Class/PHYS498/`.
[2] `https://en.wikipedia.org/wiki/Einstein_relation_(kinetic_theory)`

charge Q_o into the hose, which being rigid cannot expand (has a fixed volume). The hose may be modeled as a resistor and the tire as a capacitor C_o, which fills with charge as it is delivered via the resistor, from the pump. The capacitor obeys the same equation as Hooke's law for a spring, $F = K_o\Delta$, where K_o is the stiffness of the spring, $C_o = 1/K_o$ is the spring's compliance, and Δ is the displacement. In electrical terms, $Q_o = C_o\Phi$ where Φ is the voltage, which acts like a force *F*; Q_o is the charge, which plays the role of the mass of the fluid. The charge Q is conserved, just as the mass of the fluid is conserved (they cannot be created or destroyed).

The flow of the fluid is called the ***flux***, which is the general term for the mass flow, heat or electrical current flow (Table 3.2, p. 149). The two equations may be rewritten directly in terms of the force (F, Φ) and flow (momentum or current flux). A third alternative is in terms of the electrical

$$I = C_o\frac{d}{dt}\Phi \quad \text{[A]} \tag{5.1.12}$$

and mechanical

$$J = C_o\frac{d}{dt}F \quad \text{[kgm-s/m]} \tag{5.1.13}$$

impedance/admittance relationship.

It is common to treat the stiffness of the balloon, which acts as a spring, as a compliance $C_o = 1/K_o$, in which case the impedance Z is defined in the frequency domain as the ratio of the generalized force over the generalized flow

$$Z(s) = \frac{1}{sC_o} \quad \text{[ohms]}.$$

In the case of mechanical systems $Z_m(s) \equiv F/J$, while for the electrical system, $Z_e(s) \equiv \Phi/I$. In thermodynamics the thermal compliance is $C_o = \mathcal{S}/T$, thus $Z_{themo} = T/s\mathcal{S} = 1/sC_o$. It is helpful to use the unit [ohms] when working with any impedance, allowing for a uniform terminology for the different physical situations and many forms of impedance. This greatly simplifies the notation.

In the time domain, Ohm's law becomes Eq. 5.1.13 for the case of a mechanical compliance $C_o = 1/K_o$ and Eq. 5.1.12 for the electrical capacitor C. As shown in Table 3.2, the formula for the generalized impedance is typically expressed in terms of the Laplace frequency s, which of course is the $\mathcal{LT}$ of the time variables.

The final solution of this system is solved in the frequency domain. The impedance seen by the source is the sum of the resistance R and the impedance of the load, giving

$$Z = R + \frac{1}{sC}.$$

This results in a simple relationship between the force and the flow, as determined by the action of the source on the load $Z(s)$. The results may be given in terms of the voltage across the compliance in terms of the voltage Φ_s (or current I_s) due to the source. Given some algebra, the voltage across the compliance Φ_c, divided by the voltage of the source, is

$$\frac{\Phi_c}{\Phi_{\text{source}}} = \frac{R}{R + 1/sC}.$$

Thus the calculus reduces to some algebra in the frequency domain, which in this case has a simple pole at $s_p = -1/RC$. The time domain response is then found by taking the inverse $\mathcal{LT}$.

Cauchy's residue theorem (p. 201) gives the final answer, which describes how the voltage across the compliance builds exponentially with time, from zero to the final value. Given the voltage, we can also compute the current, as a function of time. This then represents the entire process of either blowing up a balloon, charging a capacitor, or heating water on a stove, the difference being only the physical notation, as the math is identical.

Note that this differential equation is first-order in time, which in frequency means the impedance has a single pole. Thus the equation for charging a capacitor or pumping up a balloon describes a diffusion process. If we had taken the impedance of the mass of the fluid in the hose into account, we would have a lumped-parameter model of the wave equation with a second-order system. This is mathematically related to the homework assignment about train cars (masses) connected by springs (Fig. 5.4, Homework DE-3, problem 2).

Example: The voltage

$$\phi(\boldsymbol{x},t) = e^{-\boldsymbol{\kappa}\cdot\boldsymbol{x}}u(t-x/c) \leftrightarrow \frac{1}{s}e^{-\boldsymbol{\kappa}\cdot\boldsymbol{x}} \quad [\text{V}] \tag{5.1.14}$$

represents one of d'Alembert's solution (Eq. 4.4.1, p. 188) of the wave equation (Eq. 3.1.5, p. 75) as well as an eigenfunction of the gradient operator ∇. From the definition of the scalar (dot) product of two vectors (Fig. 3.4, p. 119),

$$\boldsymbol{\kappa}\cdot\boldsymbol{x} = \kappa_x x + \kappa_y y + \kappa_z z = ||\boldsymbol{\kappa}||\ ||\boldsymbol{x}|| \cos\theta_{\kappa x},$$

where $||\boldsymbol{\kappa}|| = \sqrt{\kappa_x^2+\kappa_y^2+\kappa_z^2}$ and $||\boldsymbol{x}|| = \sqrt{x^2+y^2+z^2}$ are the lengths of vectors $\boldsymbol{\kappa}$ and $\boldsymbol{x}$, and $\theta_{\kappa x}$ is the angle between them. As before, $s = \sigma + \omega\jmath$ is the Laplace frequency.

To keep things simple, we let $\boldsymbol{\kappa} = [\kappa_x, 0, 0]^T$ so that $\boldsymbol{\kappa}\cdot\boldsymbol{x} = \kappa_x x\hat{\mathbf{x}}$. We shall soon see that $||\boldsymbol{\kappa}|| = 2\pi/\lambda$ follows from the basic relationship between a wave's radian frequency $\omega = 2\pi f$ and its wavelength λ:

$$\omega\lambda = c_o. \tag{5.1.15}$$

As the frequency increases, the wavelength becomes shorter. This key relationship may have been first researched by Galileo in about 1564, followed by Mersenne[5] in about 1627 (Fig. 1.1, p. 8).

Exercise #3

Show that Eq. 5.1.14 is an eigenfunction of the gradient operator ∇.

Solution: Taking the gradient of $\phi(\boldsymbol{x},t)$ gives

$$\begin{aligned}\nabla e^{-\boldsymbol{\kappa}\cdot\boldsymbol{x}}u(t) &= -\nabla\boldsymbol{\kappa}\cdot\boldsymbol{x}\, e^{-\boldsymbol{\kappa}\cdot\boldsymbol{x}}u(t)\\ &= -\boldsymbol{\kappa}\, e^{-\boldsymbol{\kappa}\cdot\boldsymbol{x}}u(t),\end{aligned}$$

or in terms of $\phi(\boldsymbol{x},t)$,

$$\nabla\phi(\boldsymbol{x},t) = -\boldsymbol{\kappa}\,\phi(\boldsymbol{x},t) \leftrightarrow -\frac{s}{c}e^{-\boldsymbol{\kappa}\cdot\boldsymbol{x}}.$$

Thus $\phi(\boldsymbol{x},t)$ is an eigenfunction of ∇, having the vector eigenvalue $\boldsymbol{\kappa}$. As before, $\nabla\phi$ is proportional to the current, since ϕ is a voltage, and the ratio (i.e., the eigenvalue) may be thought of as a mass, analogous to the impedance of a mass (or inductor). In general, the units provide the physical interpretation of the eigenvalues and their spectra. ∎

[5] See https://www-history.mcs.st-and.ac.uk/Biographies/Mersenne.html;

"In the early 1620s, Mersenne listed Galileo among the innovators in natural philosophy whose views should be rejected. However, by the early 1630s, less than a decade later, Mersenne had become one of Galileo's most ardent supporters." (Garber, 2004)

Example:

Filter properties: Given the function

$$F(s) = \frac{(s+1)(s-1)}{(s+2)},$$

1. Find the minimum phase $M(s)$ and all-pass $A(s)$ parts. The minimum phase part has all of its poles and zeros in the left half-plane (LHP), while the all-pass part has its poles in the LHP and mirrored zeros in the RHP. Thus we place a removable pole zero pair symmetrically across from the RHP zero, and then write the expression as the product, that is $F(s) = M(s)\cdot A(s)$:

   $$F(s) = \frac{(s+1)(s-1)}{(s+2)}\cdot\frac{s+1}{s+1} = \frac{(s+1)^2}{s+2}\cdot\frac{s-1}{s+1}$$

   Thus $M(s) \equiv \frac{(s+1)^2}{s+2}$ and $A(s) \equiv \frac{s-1}{s+1}$

2. Find the magnitude of $M(s)$ Take the real part of the log of M and then the anti-log. Thus $|M| = e^{\Re\ln M(s)}$

3. Find the phase of $M(s)$ In this case we use the imaginary part: $\angle M = \Im\ln M(s)$

4. Find the magnitude of $A(s)$ 1, by definition.

5. Find the phase of $A(s)$ $\angle A = \Im\ln(A)$

More questions

There are a number of question to be addressed:

1. Can we interpret the zeta function as a frequency domain quantity, and then inverse transform it into the time domain?

   The answer to this is yes, and the results are quite interesting.

2. Make a histogram of the entropy for the first million integers.

   This is a 5 minute job in Matlab/Octave. For example:

```
K=1e5; N=1:K; F=zeros(K,10);
 for n=1:K;
  f=factor(n);
  F(n,1:length(f))=f;
 end;
hist(F);
```

This equation is of the form $F\left(\frac{s}{2}\right)\zeta(s) = F\left(\frac{1-s}{2}\right)\zeta(1-s)$, where $F(s) = \Gamma(s)/\pi^s$.

As shown in Table 3.8, the $\mathcal{LT}^{-1}$ of $f(-t) \leftrightarrow F(-s)$ represents time-reversal. This leads to causal and anticausal functions that are symmetric about $\Re\{s\} = 1/2$ (Riemann, 1859) leading to an interpretation of Euler's functional equation.

Riemann (1859, page 2) provides an alternative integral definition of $\zeta(s)$, based on the complex contour integration,[5]

$$2\sin(\pi s)\Gamma(s-1)\zeta(s) = \jmath \int_{x=-\infty}^{\infty} \frac{(-x)^{s-1}}{e^x - 1} dx \overset{-x \to y\jmath}{=} \int_{y=-\infty\jmath}^{\infty\jmath} \frac{(y\jmath)^{s-1}}{e^{-y\jmath} - 1} dx\jmath.$$

Given $\zeta_k(s)$, it seems important to look at the inverse $\mathcal{LT}$ of $\zeta_k(1-s)$ to gain insight into the analytically extended $\zeta(s)$.

What is the RoC of $\zeta(s)$? It is commonly stated that Euler's and thus Riemann's product formulas are valid only for $\Re s > 1$; however, this does not seem to be actually proved (I could be missing this proof). Here I argue that the product formula is entire except at the poles—namely, that the formula is analytic everywhere other than at the poles.

The argument goes as follows: Starting from the product formula (Eq. C.1.7, p. 304), we form the log-derivative and study the poles and residues:

$$\begin{aligned} D(s) &\equiv \frac{d}{ds} \ln \Pi_k \frac{1}{1 - e^{-sT_k}} \\ &= -\sum_{k=1}^{\infty} \frac{1}{1 - e^{-sT_k}} \frac{d}{ds} 1 - e^{-sT_k} \\ &= -\sum_k \frac{T_k e^{-sT_k}}{1 - e^{-sT_k}} \leftrightarrow \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \delta(t - nT_k). \end{aligned}$$

Here $T_k = \ln \pi_k$, as previously defined, and $\leftrightarrow$ denotes the inverse Laplace transform, transforming $D(s) \leftrightarrow d(t)$ into the time domain. Note that $d(t)$ is a causal function, composed of an infinite number of delta functions (i.e., time delays), as shown in Fig. C.2 (p. 307).

Zeros of $\zeta(s)$ We are still left with the most important question: Where are the zeros of $\zeta(s)$? Equation C.1.11 has no zeros; it is an all-pole system. The cascade of many such systems is also all-pole. As I see it, the issue is: What is the actual formula for $\zeta(s)$?

To answer this question, we need to study the properties of the reflectance function $\Gamma(s)$. Frequency-domain transfer functions having unity magnitude on the $j\omega$ axis are called *all-pass filters* in the engineering literature. When the reflectance is loss-less, it is therefore all-pass since $|\Gamma(j\omega| = 1$. An important property of all-pass filters is that they may be accurately approximated by pole-zero pairs straddling the $j\omega$ axis, with the poles to the left (as required by causality) and the zeros to the right. Given this placement, the phases of the poles and zeros add. The group delay gives the net delay of the all-pass filter, which is twice the delay of the poles alone. It would seem that this careful placement of the zeros exactly across from the poles provides the requirement that the zeros all line up parallel to the $j\omega$ axis, as deemed by the Riemann hypothesis. Could this be the resolution of this long-standing mystery? An alternative possibility is that the convergent product formula has zeros that are obscured by the lack of convergence of Eq. C.1.2.

[5] We can verify Riemann's use of x, which is taken to be real rather than complex. This could be more natural (i.e., modern Laplace transformation notation) if $-x \to y\jmath \to z$.

Exercise #4

Compute $\hat{\mathbf{n}}$ for $\phi(\boldsymbol{x}, s)$ as given by Eq. 5.1.14.

Solution: $\hat{\mathbf{n}} = \boldsymbol{\kappa}/||\boldsymbol{\kappa}||$ represents a unit vector in the direction of the wave propagation.■

Exercise #5

If the sign of $\boldsymbol{\kappa}$ is negative, what are the eigenvectors and eigenvalues of $\nabla\phi(\boldsymbol{x}, t)$?

Solution:

$$\begin{aligned} \nabla e^{-\boldsymbol{\kappa}\cdot\boldsymbol{x}} u(t) &= -\boldsymbol{\kappa} \cdot \nabla(\boldsymbol{x}) e^{-\boldsymbol{\kappa}\cdot\boldsymbol{x}} u(t) \\ &= -\boldsymbol{\kappa}\ e^{-\boldsymbol{\kappa}\cdot\boldsymbol{x}} u(t). \end{aligned}$$

Nothing changes other than the sign of $\boldsymbol{\kappa}$. Physically this means the wave is traveling in the opposite direction, corresponding to the forward and retrograde d'Alembert waves. ■

Prior to this section, we had considered the Taylor series in only one variable, such as for polynomials $P_N(x), x \in \mathbb{R}$ (Eq. 3.1.7, p. 77), and $P_N(s), s \in \mathbb{C}$ (Eq. 3.2.20, p. 102). The generalization from real to complex–analytic functions led to the $\mathcal{LT}$ and the host of integration theorems (FTCC, Cauchy CT-1, CT-2, CT-3). What is in store when we generalize from one spatial variable ($\mathbb{R}$) to three ($\mathbb{R}^3$)?

Exercise #6

Find the velocity $v(t)$ of an electron in a field $\boldsymbol{e}$.

Solution: From Newton's second law, $-qE = m_e \dot{v}(t)$ [N], where m_e is the mass of the electron. Thus we must solve this first-order differential equation to find $v(t)$. As before, this is best done in the frequency domain $v(t) \leftrightarrow V(s)$.■

Role of potentials: Note that the scalar fields (e.g., temperature, pressure, voltage) are all scalar potentials, summarized in Table 3.2 (p. 149). In each case the gradient of the potential results in a vector force field, just as in the electric case above (Eq. 5.1.1).

Table 3.2 is helpful in understanding the physical meaning of the gradient of a potential, which is typically a generalized force (electric field, acoustic force density, temperature flux), that in turn generates a flow (current, velocity, heat flux (entropy)). The ratio of the potential over flow determines the impedance. Four examples are

1. The voltage drop across a resistor causes a current to flow, as described by Ohm's law. The difference in voltage between two points is a crude form of gradient when the frequency f [Hz] is low, such that the wavelength is much larger than the distance between the two points. This is the essence of the quasistatic approximation (Postulate P10, p. 163).

2. The gradient of the pressure gives rise to a force density in the fluid medium (air, water, oil, etc.), that causes a flow (velocity vector) in the medium.

3. The gradient of the temperature also causes a flow of heat that is proportional to the thermal resistance, given Ohm's law for heat (Feynman, 1970b, p 3–7).

4. When a solution contains ions, it defines an *electrochemical Nernst potential* $N(\boldsymbol{x}, t)$ (Fermi, 1936; Scott, 2002). This electrochemical potential is similar to a voltage or temperature field, the gradient of which defines a virtual force on the ions, resulting in an ionic current.

Thus in the above examples there is a potential, the gradient of which is a force, that when applied to the medium (an impedance) causes a flow (flux or current) proportional to that impedance due to the medium, which is the ratio of the gradient of the potential to the current. The product of the force and flow is a power.

Exercise #7

Show that the integral of Eq. 5.1.1 is an antiderivative.

Solution: We use the definition of the antiderivative given by the FTC (Eq. 4.2.2, p. 180):

$$\begin{aligned}
\phi(\boldsymbol{x},t) - \phi(\boldsymbol{x}_o,t) &= \int_{\boldsymbol{x}_o}^{\boldsymbol{x}} \boldsymbol{e}(\boldsymbol{x},t)\cdot d\boldsymbol{x} \\
&= -\int_{\boldsymbol{x}_o}^{\boldsymbol{x}} \nabla\phi(\boldsymbol{x},t)\cdot d\boldsymbol{x} \\
&= -\int_{\boldsymbol{x}_o}^{\boldsymbol{x}} \left(\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}\right)\phi(\boldsymbol{x},t)\cdot d\boldsymbol{x} \\
&= -\int_{\boldsymbol{x}_o}^{\boldsymbol{x}} \left(\hat{\mathbf{x}}\frac{\partial\phi}{\partial x} + \hat{\mathbf{y}}\frac{\partial\phi}{\partial y} + \hat{\mathbf{z}}\frac{\partial\phi}{\partial z}\right)\cdot(\hat{\mathbf{x}}dx + \hat{\mathbf{y}}dy + \hat{\mathbf{z}}dz) \\
&= -\int_{x_o}^{x}\frac{\partial\phi}{\partial x}dx - \int_{y_o}^{y}\frac{\partial\phi}{\partial y}dy - \int_{z_o}^{z}\frac{\partial\phi}{\partial z}dz \\
&= -\int_{\boldsymbol{x}_o}^{\boldsymbol{x}} d\phi(\boldsymbol{x},t) \\
&= -\Big(\phi(\boldsymbol{x},t) - \phi(\boldsymbol{x}_o,t)\Big).
\end{aligned}$$

This may be verified by taking the gradient of both sides:

$$\nabla\phi(\boldsymbol{x},t) - \cancel{\nabla\phi(\boldsymbol{x}_o,t)}^{0} = -\nabla\int_{\boldsymbol{x}_o}^{\boldsymbol{x}}\boldsymbol{e}(\boldsymbol{x},t)\cdot d\boldsymbol{x} = \boldsymbol{e}(\boldsymbol{x},t).$$

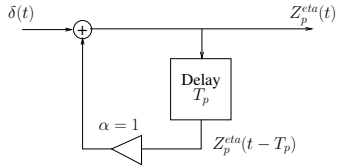
If we apply the FTC, the antiderivative must be $\phi(\boldsymbol{x},t) = E_x x\hat{\mathbf{x}} + 0\hat{\mathbf{y}} + 0\hat{\mathbf{z}}$. This same point is made by Feynman (1970b, p. 4-1, Eq. 4.28). ∎

Given that the force on a charge is proportional to the gradient of the potential, this exercise shows that the integral of the gradient depends on only the end points, the work done in moving a charge depends on only the limits of the integral, which is the definition of a *conservative field* but holds only in the ideal case where $\boldsymbol{e}$ is determined by Eq. 5.1.1—that is, the medium has no friction (there are no other forces on the charge).

The conservative field: An important question is: When is a field conservative? A field is conservative when the work done by the motion is independent of the path of the motion. Thus the conservative field is related to the FTC, which states that the integral of the work depends on only the end points.

A more complete answer must await the introduction of the fundamental theorem of vector calculus (Eq. 5.6.11, p. 261). A few examples provide insight:

Example: The gradient of a scalar potential, such as the voltage (Eq. 5.1.1), defines the electric field, which drives a current (flow) across a resistor (impedance). When the impedance is infinite, the flow is zero, leading to zero power dissipation. When the impedance is lossless, the system is conservative.



$$Z_p^{eta}(t) = \alpha Z_p^{eta}(n - T_p) + \delta(t)$$

Figure C.2: *This feedback network is described by a time-domain difference equation with delay $T_p = \ln\pi_k$. It has an all-pole transfer causal function given by Eq. C.1.11. Physically this delay corresponds to T_p [s].*

Such functions may be generated in the time domain as shown in Fig. C.2, using a feedback delay of T_p [s] as described by the equation in the figure, with a unity feedback gain $\alpha = 1$,

$$Z^{eta}(t) = Z^{eta}(t - T_p) + \delta(t).$$

Taking the Laplace transform of the system equation, we see that the transfer function between the state variable $q(t)$ and the input $x(t)$ is given by $Z_p^{eta}(t)$. Taking the $\mathcal{LT}$, we see that $\zeta(s)$ is a causal all-pole function,

$$\zeta_p(s) = e^{-sT_p}\zeta_p(s) + 1(t) \text{ or } \zeta_p(s) = \frac{1}{1 - e^{-sT_p}}. \tag{C.1.11}$$

This is a key expression in the theory of Black Body radiation where the Black Body radiation formula is (Allen, 2024) fix this

$$S(s) = \frac{-s^3}{1 - e^{-hs/kT}}$$

In terms of the physics, these transmission line equations are telling us that $\zeta(s)$ may be decomposed into an infinite cascade of transmission lines (Eq. C.1.10), each having a unique delay given by $T_k = \ln\pi_k$, $\pi_k \in \mathbb{P}$, the log of the primes. The input admittance of this cascade may be interpreted as an analytic continuation of $\zeta(s)$ that defines the eigenmodes of that cascaded impedance function.

Working in the time domain provides a key insight, as it allows us to determine the analytic continuation of the infinity of possible continuations, which may not be obvious in the frequency domain. Transforming to the time domain is a variant of analytic continuation of a function $Z(s) \leftrightarrow Z^{eta}(t)$ that depends on the assumption that $Z^{eta}(t)$ is one-sided in time (causal). It may be helpful to compare this variant to Euler's continuation of $\zeta(s)$, and later Riemann's classic 1851 definition of complex–analytic continuation $Z^{eta}(s)$.

Additional relationships: We need to know some important relationships provided by both Euler and Riemann (1859) when we study $\zeta(s)$. Move to Lec33

With the goal of generalizing his result, Euler extended the definition with the functional equation

$$\zeta(s) = 2^s\pi^{s-1}\sin\left(\frac{\pi s}{2}\right)\Gamma(1-s)\,\zeta(1-s). \tag{C.1.12}$$

This seems closely related to Riemann's time reversal symmetry properties (Bombieri, 2000),

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \pi^{-(1-s)/2}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s).$$

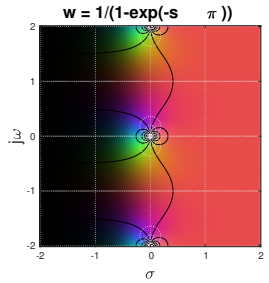


Figure C.1: *Plot of $w(s) = \frac{1}{1-e^{-s\pi}}$. Here $w(s)$ has poles where $e^{s_n \ln 2} = 1$—namely, where $\omega_n \ln 2 = n2\pi$, as seen in the colorized map ($s = \sigma + \omega\jmath$) is the Laplace frequency [rad]).*

We can now identify the poles of $\zeta_k(s)$ ($p \in \mathbb{N}$), which are required to determine the RoC. For example, the kth factor of Eq. C.1.7 expressed as an exponential, is

$$\zeta_k(s) \equiv \frac{1}{1-\pi_k^{-s}} = \frac{1}{1-e^{-sT_k}} = \sum_{k=0}^{\infty} e^{-skT_k}, \tag{C.1.8}$$

where $T_k \equiv \ln \pi_k$. Thus $\zeta_p(s)$ has poles at $-s_n T_p = 2\pi n\jmath$ (when $e^{-sT_p} = 1$), and

$$\omega_n = \frac{2\pi n}{T_k},$$

with $-\infty < n \in \mathbb{Z} < \infty$. These poles are the eigenmodes of the zeta function. Fig. C.1 is a domain-colorized plot of this function. It is clear that the RoC of ζ_k is > 0. It would be helpful to determine why $\zeta(s)$ has a more restrictive RoC than each of its factors.

Inverse Laplace transform

The inverse Laplace transform of Eq. C.1.8 is an infinite series of delays T_p (Table 3.9)[4]

$$Z_p^{eta}(t) = \delta(t))_{T_p} \equiv \sum_{k=0}^{\infty} \delta(t - kT_p) \leftrightarrow \frac{1}{1-e^{-sT_p}}. \tag{C.1.9}$$

Inverse transform of the product of factors

The time-domain version of Eq. C.1.7 may be written as the convolution of all the $Z_k^{eta}(t)$ factors

$$Z^{eta}(t) \equiv Z_2^{eta} \star Z_3^{eta}(t) \star Z_5^{eta}(t) \star Z_7^{eta}(t) \star \cdots \star Z_p^{eta}(t) \star \cdots, \tag{C.1.10}$$

where $\star$ represents time convolution (Table 3.8).

[4]Here we use a shorthand double-parentheses notation $f(t))_T \equiv \sum_{k=0}^{\infty} f(t - kT)$ to define the one-sided infinite sum.

Example: At audible frequencies the viscosity of air is quite small and thus, for simplicity, it may be taken as zero. However, when the wavelength is small (e.g., at 100 [kHz] $\lambda = c_o/f = 345/10^5 = 3.45$ [mm]) the lossless assumption breaks down, resulting in a significant propagation loss. When the viscosity is taken into account, the field is lossy and thus the field is no longer conservative. In narrow tubes, for example a flute, thermal loss plays a much larger role due to the walls (Appendix D).

Example: If a temperature field is a time-varying constant [i.e., $T(\boldsymbol{x}, t) = T_o(t)$], there is no "heat flux," since $\nabla T_o(t) = 0$. When there is no heat flux [i.e., flux, or flow], there is no heat power, since the power is the product of the force and the flow.

Example: The force of gravity is given by the gradient of Newton's gravitational potential (Eq. 3.1.1, p. 74):

$$F = -\nabla_r \phi_{\mathrm{N}}(r) = -\frac{\partial}{\partial r}\frac{1}{r} = \frac{1}{r^2}.$$

Historically speaking, $\phi_N(r)$ was the first conservative field, used to explain the elliptic orbits of the planets around the sun. Galileo's law says that bodies fall with constant acceleration, giving rise to a parabolic path and a time of fall proportional to t^2. This behavior of falling objects directly follows from the Galilean potential:

$$\phi_{\mathrm{G}}(r) = \frac{1}{(r - r_o)} = \frac{-r_o}{1 - r/r_o} \underset{r<r_o}{=} -r_o(1 - r/r_o + (r/r_o)^2 + \cdots) \underset{r \ll r_o}{\approx} r_o - r,$$

which, given the large radius r_o of the earth and the small distance of the object from the surface of the earth $r - r_o$, is equal to the distance above the ground. Thus Galileo's law says that the force a falling body sees is constant:

$$F_G = -G_o \nabla_r \phi_G(r) = G_o.$$

This can be scaled by G_o to account for the magnitude of the gravitational force.

Exercise #8

Galileo discovered that the height of a falling object is proportional to the square of the time it falls. Based on Newton's follow-up analysis, today we would say this height $h(t)$ is

$$h(t) = \frac{1}{2} m G_o (t - t_o)^2 \quad [\mathrm{m}],$$

where m is the object's mass and G_o is the gravitational constant for the earth at its surface r_o. Show that $h(t)$ directly follows from the potential $\phi_G = r_o - r$. This formula applies if you toss a ball into the air or if you drop it from a high place.

Solution: Given Galileo's potential $\phi_G(r) \underset{r \ll r_o}{\approx} m G_o (r_o - r)$, thus $\ddot{h}(t) = m G_o$. Given Galileo's formula for the height $h(t)$, the velocity is $v(t) = \dot{r}(t) = m G_o t$, and the acceleration is $\ddot{r}(t) = m G_o$. ∎

Exercise #9

Find the time that it takes to fall from a distance $r = L$. That is solve $h(t) = L$ for the time the object takes to fall the distance L.

Solution: Setting $t_o = 0$ gives $t^2 = 2L/mG_o$. Thus the time to fall is $T(L) = \sqrt{2L/mG_o}$. ∎

5.2 Partial differential equations and field evolution

The three main classes of partial differential equations (PDEs) are: elliptic, parabolic, and hyperbolic, distinguished by the order of the time derivative. These categories seem to have little mathematical utility (the categories are labels).

5.2.1 The Laplacian ∇^2:

In the most important case the space operator is the Laplacian ∇^2, the definition of which depends on the dimensionality of the waves—that is, the coordinate system being used. We first discussed the Laplacian as a 2D operator on p. 182 where we studied complex–analytic functions, and again on p. 225. An expression for ∇^2 for one, two, and three-dimensions was provided as Eq. 5.1.9 (p. 229). In three dimensional rectangular coordinates, it is defined as (see p. 228)

$$\nabla^2 T(\boldsymbol{x}) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) T(\boldsymbol{x}). \tag{5.2.1}$$

The Laplacian operator is ubiquitous in mathematical physics, starting with simple complex–analytic functions (Laplace's equation) and progressing to Poisson's equation, the diffusion equation, and finally the wave equation. Only the wave equation results in a delay. The diffusion equation "wave" has an instantaneous spread (the effective "wavefront" velocity is infinite, yet the wavelength is long; it's not a traveling wave).

Examples of elliptic, parabolic, and hyperbolic equations are

1. *Laplace's equation:* The equation

$$\nabla^2 \Phi(\boldsymbol{x}) = 0 \tag{5.2.2}$$

describes, for example, the voltage inside a closed chamber that has a given voltage on the walls or the steady-state temperature within a closed container having a specified temperature distribution on the walls. There are no dynamics to the potential, even when it is changing, since the potential instantaneously follows the potential on the walls.

2. *Poisson's equation:* In the steady state, the diffusion equation degenerates to either Poisson's or Laplace's equation; both are classified as *elliptic* equations (second-order in space, zero-order in time). As in the diffusion equation, the evolution has a wave velocity that is functionally infinite. For example,

$$\nabla^2 \Phi(\boldsymbol{x}, t) = \rho(x, t)$$

holds for gravitational fields or the voltage around a charge. It does not describe gravity waves, which travel at the speed of light.

3. *Fourier diffusion equation:* Equation 5.2.3 describes the evolution of the scalar temperature $T(\boldsymbol{x}, t)$ (a scalar potential), gradients of solution concentrations (i.e., ink in water), and Brownian motion. Diffusion is first-order in time, which is categorized as *parabolic* (first-order in time, second-order in space). When these equations are Laplace transformed, diffusion has a single real root, resulting in a real solution (e.g., $T \in \mathbb{R}$). There is no wavefront in the case of the diffusion equation. As soon as the source is turned on, the field is nonzero at every point in the bounded container. As an example,

$$\nabla^2 T(\boldsymbol{x}, t) = \kappa_o \frac{\partial T(\boldsymbol{x}, t)}{\partial t} \leftrightarrow s\kappa_o T(\boldsymbol{x}, s) \tag{5.2.3}$$

Finding the RoC of the product formula: It would be interesting to find the RoC for $\mathcal{P}_k(s)$, and for rigor, this question demands further investigation. To find the RoC, we need to evaluate

$$|\pi_k^{-s}| = |e^{-sT_k}| = |e^{-\sigma T_k}| = \left(\frac{1}{\pi_k}\right)^{\sigma} < 1 \quad \text{for} \quad \sigma > 0,$$

where $T_k = \ln \pi_k$. For example,

$$\frac{1}{\mathcal{P}_5(s)} = \frac{1}{1 - \left(\frac{1}{5}\right)^s} = 1 + \frac{1}{5^s} + \frac{1}{5^{2s}} + \frac{1}{5^{3s}} \cdots, \quad \Re\{s\} = \sigma > 0.$$

The RoC for each root is $\sigma \geq 0$ since when $\sigma < 0$.

Since $1/\pi_k < 1$ for all $k \in \mathbb{N}$, the Taylor series of $\zeta_k(x)$ is entire except at its poles. Note that the RoC of a Taylor series in powers of π_k^{-s} increases with k.

Exercise #*1*

Work out the RoC for $k = 2$.

Solution: The formula for the RoC is given above, which for $\pi_2 = 3$ is

$$|\pi_k^{-s_r}| = \begin{array}{ll} \left(\frac{1}{3}\right)^{\sigma_r} < 1 & \text{for} \quad \sigma_r > 0, \\ \left(\frac{1}{3}\right)^{-\sigma_r} < 1 & \text{for} \quad \sigma_r < 0, \end{array}$$

where σ_r is the boundary of the RoC.

■

Exercise #*2*

Show how to construct $Z_2^{eta}(t) \leftrightarrow \zeta_2(s)$ by working in the time domain.

Solution: The basic procedure for building a sieve is to sum the integers

$$S_1 = \sum_{n=1}^{\infty} n 2^{n-1} = 1 \cdot 2^0 + 2 \cdot 2^1 + 3 \cdot 2^2 + \cdots,$$

while the sieve for the kth prime π_k is

$$S_k = \sum_{n=1}^{\infty} n \pi_k^{n-1} = 1 \cdot \pi_k^0 + \pi_k \cdot 2^1 + \pi_k \cdot 2^2 + \cdots.$$

This sum may be written in terms of the convolution with the Heaviside step function u_k, since

$$u_k \star u_k = n u_k = 0 \cdot u_0 + 1 \cdot u_1 + 2 u_2 + \cdots + k u_k + \cdots.$$

■

Poles of $\zeta_k(s)$

Riemann proposed that Euler's zeta function $\zeta(s) \in \mathbb{C}$ has a complex argument [first explored by Chebyshev in 1850 (Bombieri, 2000)] that extends $\zeta(s)$ into the complex plane ($s \in \mathbb{C}$), thus making it a complex–analytic function. Thus we might presume that $\zeta(s)$ has an inverse Laplace transform. There seems to be very little written on this topic (Hill, 2007). We explore this question further here.

Euler product formula

As first published by Euler in 1737, we can recursively factor out the leading prime term, which results in Euler's product formula. Euler's procedure is an algebraic implementation of the sieve of Eratosthenes (Fig. 2.3, p. 45).

Multiplying $\zeta(s)$ by the factor $1/2^s$ and subtracting from $\zeta(s)$ remove all the powers of 2: $1/2^0 + 1/2^s + 1/2^{2s} + 1/2^{3s} + \cdots$

$$\left(1 - \frac{1}{2^s}\right)\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} + \cdots - \left(\frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{8^s} + \frac{1}{16^s} + \cdots\right), \quad \text{(C.1.3)}$$

oC? which results in

$$\zeta_1(s) = \left(1 - 2^{-s}\right)\zeta(s) = 1 + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{9^s} + \frac{1}{11^s} + \frac{1}{13^s} + \cdots. \quad \text{(C.1.4)}$$

Repeating this with a lead factor $1/3^s$ applied to Eq. C.1.4 gives[3]

$$\frac{1}{3^s}\left(1 - 2^{-s}\right)\zeta(s) = \frac{1}{3^s} + \frac{1}{9^s} + \frac{1}{15^s} + \frac{1}{21^s} + \frac{1}{27^s} + \frac{1}{33^s} + \cdots. \quad \text{(C.1.5)}$$

Subtracting Eq. C.1.5 from Eq. C.1.4 cancels the terms on the right-hand side of Eq. C.1.4, giving

$$\zeta_2(s) = \left(1 - 3^{-s}\right)\left(1 - 2^{-s}\right)\zeta(s) = 1 + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{11^s} + \frac{1}{13^s} + \frac{1}{17^s} + \frac{1}{19^s} + \cdots.$$

If we express this in terms of the primes π_k, we can better visualize the structure:

$$\zeta_2(s) = \left(1 - \pi_2^{-s}\right)\left(1 - \pi_1^{-s}\right)\zeta(s) = 1 + \frac{1}{\pi_3^s} + \frac{1}{\pi_4^s} + \frac{1}{\pi_5^s} + \frac{1}{\pi_6^s} + \frac{1}{\pi_7^s} + \frac{1}{\pi_8^s} + \cdots.$$

Thus ζ_2 has removed primes π_1, π_2, leaving π_3 as the lead term in the series on the right-hand side.

This leads to a recursion in ζ_k,

$$\zeta_k(s) = \zeta(s) \prod_{l=1}^{k} \zeta_l(s) = 1 + \sum_{l=k+1}^{\infty} \pi_l^{-s}.$$

The series on the right-hand side converges rapidly to 1 as each prime is removed, because the RoC is becoming much larger with each recursion. Each recursive step in this construction ensures that the lead term, along with all of its multiplicative factors, is subtracted out, just like the cancellations with the sieve of Eratosthenes. It is instructive to compare each iteration with that of the sieve (see Fig. 2.3, p. 45).

Repeating this process with the remaining primes removes all the terms on the right-hand side but the first (leaving 1), which results in *Euler's analytic product formula* ($s = x \in \mathbb{R}$), or *Riemann's complex–analytic product formula* ($s \in \mathbb{C}$):

$$\begin{aligned} 1 &= \zeta(s)(1 - 2^{-s}) \cdot (1 - 3^{-s}) \cdot (1 - 5^{-s}) \cdot (1 - 7^{-s}) \cdot \cdots \cdot (1 - \pi_n^{-s}) \cdot \cdots \\ &= \zeta(s) \prod_{k=1}^{\infty} (1 - \pi_k^{-s}) \end{aligned} \quad \text{(C.1.6)}$$

$$\zeta(s) = \frac{1}{\prod_k \mathcal{P}_k(s)}, \quad \Re\{s\} = \sigma > 0, \quad \text{(C.1.7)}$$

where the zeros of $\mathcal{P}_k(s) = 1 - \pi_k^{-s}$ define the poles of $\zeta(s)$ for prime π_k.

[3] This is known as *Euler's sieve*, as distinguished from the *Eratosthenes sieve*.

describes the temperature $T(\boldsymbol{x}, t) \leftrightarrow T(\boldsymbol{x}, \omega)$, as proposed by Fourier in 1822, or the diffusion of two miscible liquids (Fick, 1855) or Brownian motion (Einstein, 1905). The diffusion equation is not a wave equation, since the temperature wavefront propagates instantaneously. The diffusion equation does a poor job of representing the velocity of molecules banging into each other, since such collisions have a mean free path, and thus the velocity cannot be infinite.

4. *Wave equations:* There are scalar and vector forms of wave equations.

   (a) Scalar wave equations: Equation 3.1.3 (p. 75) describes the evolution of a scalar potential field, such as pressure $\varrho(\boldsymbol{x}, t)$ (sound) or the displacement of a string or membrane under tension. The wave equation is second-order in time. When transformed into the frequency domain, the solution has pairs of complex conjugate roots, leading to two real solutions. The wave equation is classified as *hyperbolic* (second-order in time and space).

   (b) Vector wave equations: Maxwell's equations describe the propagation of the electric $\boldsymbol{e}(\boldsymbol{x}, t)$ and magnetic $\boldsymbol{h}(\boldsymbol{x}, t)$ field strength vectors, as well as the electric $\boldsymbol{d}(\boldsymbol{x}, t) = \epsilon_o \boldsymbol{e}(\boldsymbol{x}, t)$ and magnetic $\boldsymbol{b}(\boldsymbol{x}, t) = \mu_o \boldsymbol{h}(\boldsymbol{x}, t)$ flux vectors. ME are antireciprocal (P6).

Solution evolution: The partial differential equation defines the evolution of the scalar field [pressure $\varrho(\boldsymbol{x}, t)$ and temperature $T(\boldsymbol{x}, t)$], or vector field ($\boldsymbol{e}, \boldsymbol{d}, \boldsymbol{b}, \boldsymbol{h}$), as functions of space $\boldsymbol{x}$ and time t. There are two basic categories of field evolution: diffusion and propagation.

1. *Diffusion*: The simplest and easiest PDE example, easily visualized, is a static[6] (time-invariant) scalar temperature field $T(\boldsymbol{x})$ [°C]. Just like an impedance or admittance, a field has regions where it is analytic, and for the same reasons, $T(\boldsymbol{x}, t)$ satisfies Laplace's equation

   $$\nabla^2 T(\boldsymbol{x}, t) = 0.$$

   Since there is no current when the field is static, such systems are lossless and thus are conservative.

   When $T(\boldsymbol{x}, t)$ depends on time (is not static), it is described by the *diffusion equation* (Eq. 5.2.3), a rule for how $T(\boldsymbol{x}, t)$ evolves with time from its initial state $T(\boldsymbol{x}, 0)$. The constant κ_o is called the *thermal conductivity*, which depends on the properties of the fluid in the container, with $s\kappa_o$ being the thermal admittance per unit area. The conductivity is a measure of how the heat gradients induce heat currents $\boldsymbol{j} = -\kappa_o \nabla T$, analogous to Ohm's law for electricity.

   Note that when $T(\boldsymbol{x}, t \to \infty)$ the temperature reaches a steady state, $\boldsymbol{j} = 0$ and $\nabla^2 T = 0$. This all depends on what is happening at the boundaries. When the wall temperature of a container is a function of time, the internal temperature $T(\boldsymbol{x}, t)$ will continue to change, but with a frequency-dependent delay that depends on the thermal conductivity κ_o.

   Such a system is analogous to an electrical resistor–capacitor series circuit connected to a battery. For example, the wall temperature (voltage across the battery) represents the potential driving the system. The thermal conductivity κ_o (the electrical resistor) is likewise analogous. The fluid (the electrical capacitor) is being heated (charged) by the

[6] Postulate P3, p. 163.

heat (charge) flux. In all cases Ohm's law defines the ratio of the potential (voltage) to the flux (current). How this happens can be understood only once the solution to the equations has been established. The fluid has a heat capacity analogous to that of an electrical capacitor (Kirchhoff, 1868, 1974). Sometimes the diffusion equation is called the telegraph equation.

2. *Propagation:* Pressure and electromagnetic waves are described by a scalar potential (pressure) (Eq. 3.1.3, p. 75) and a vector potential (Eq. 5.7.4, p. 268), leading to scalar and vector wave equations. Sometimes the wave equation is called the telephone equation.

The Taylor series of $f(\boldsymbol{x})$: Next we extend the concept of the Taylor series of one variable to $\boldsymbol{x} \in \mathbb{R}^3$. Just as we generalized the derivative with respect to a real frequency variable $\omega \in \mathbb{R}$ to complex frequency $s = \sigma + \omega\jmath \in \mathbb{C}$, here we generalize the derivative with respect to $x \in \mathbb{R}$ to the vector $\boldsymbol{x} \in \mathbb{R}^3$.

Since the scalar field is analytic in $\boldsymbol{x}$, it is a good place to start. Assuming we have carefully defined the Taylor series (Eq. 3.2.9, p. 95) in one and two (Eq. 4.2.7, p. 182) variables, the Taylor series of $f(\boldsymbol{x})$ in $\boldsymbol{x} \in \mathbb{R}^3$ about $\boldsymbol{x} = 0$ may be defined as

$$f(\boldsymbol{x} + \delta\boldsymbol{x}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x}) \cdot \delta\boldsymbol{x} + \frac{1}{2!} \sum_{k=1}^{3} \sum_{l=1}^{3} \frac{\partial^2 f(\boldsymbol{x})}{\partial x_k \partial x_l} \delta x_k \delta x_l + \text{HOT}, \tag{5.2.4}$$

where HOT stands for Higher Order Terms (Greenberg, 1988, p 639). From this definition, it is clear that the gradient is the generalization of the second term in the one-dimensional Taylor series expansion.

Summary: For every potential $\phi(\boldsymbol{x}, t)$ there exists a force density $\boldsymbol{f}(\boldsymbol{x}, t) = -\nabla\phi(\boldsymbol{x}, t)$, proportional to the potentials, that drives a generalized flow $\boldsymbol{u}(\boldsymbol{x}, t)$. If the normal components of the force and flow are averaged over a surface, the mean force and volume flow (i.e, volume velocity for the acoustic case) are defined. In such cases the impedance $Z(s)$ is the net force through the surface force over the net flow, and Gauss's law and quasistatics (Postulate P10, p. 163) come into play (Feynman, 1970a). We call this the *generalized impedance*. An example is $Z(s) = \sqrt{s}$.

Assuming linearity (Postulate P2, p. 162), the product of the force and flow is the power, and the ratio (force/flow) is an impedance (Table 3.2, p. 149). This impedance statement is called Ohm's law, Kirchhoff's laws, Laplace's law, or Newton's laws. In the simplest cases, they are all linearized (proportional) complex relationships between a force and a flow. Very few impedance relationships are inherently linear over a large range of force or current, but for physically useful levels, they are treated as linear. Nonlinear interactions require a more sophisticated approach, typically involving numerical methods, working in the time domain.

In electrical circuits it is traditional to define a zero potential *ground* point that all voltages use as the reference potential. The ground is a useful convention as a simplifying rule, but it obscures the physics and obscures the fact that the voltage is *not* the force. Rather, the force is the voltage difference, referenced to the ground, which is defined as zero volts. This results in abstracting away (i.e., hiding) the difference in voltage. It seems misleading (more precisely, it is wrong) to state Ohm's law as the voltage over the current, since Ohm's law actually says that the *voltage difference* (i.e., voltage gradient) over the current defines an impedance (Kennelly, 1893).

as described in his 1851 paper.[1] But is his definition correct? To resolve this one needs to improve our understanding of the core definition of analytic continuation, as it relates to the $\mathcal{LT}^{-1}$. In the case of the geometric series the analytic continuation is the closed form expression $f(s) = 1/(1-s)$ which is valid for all $s \neq 1$. This is not Riemann's definition of analytic continuation.

This section is a beginners review of $\zeta(s)$, building on the developments of analytic functions from Chapter 3, especially in Secs. 3.2.6, 5.4, and Fig. 5.3 p. 256. Well understood are the locations of the poles of zeta, which depend on the prime numbers. Not so well understood are the remaining analytic properties over the entire plane, such as the zeros of $\zeta(x)$, namely the poles of $1/\zeta(s)$. A key function is $\ln\zeta(s)$.

Consider $z \equiv e^{sT}$, where T is the sample period at which data are taken (every T seconds). For example, if $T = 22.676 = 10^6/44,100$ [μs], then the data is sampled at 44.10 [kHz]. This is how modern digital audio works, for CD-quality music. The unit-time delay time operator z^{-1} is

$$\delta(t - T) \leftrightarrow e^{-sT}.$$

When we deal with the Euler and Riemann zeta functions, the only sampling period that makes sense is $T = 1$ [s] or 1 [Hz] (i.e., $n \in \mathbb{Z}$). In this case, the samples of interest are $\text{mod}(n, \pi_k)$. Starting from the sieve of Eratosthenes, Euler showed that the counting numbers $n \in \mathbb{Z}$, presented at a rate of one per second [1-Hz], may be uniquely reduced to multiples of the primes. This is the basis for the fundamental theorem of arithmetic, the theorem of the concept of the prime number, which states that every integer may be uniquely factored into a product of prime numbers.

The zeta function The poles of the zeta function depend explicitly on the primes, which makes it a very special function. In 1737 Euler proposed the real-valued function $\zeta(x) \in \mathbb{R}$, and $x \in \mathbb{R}$ to prove that the number of primes is infinite (Goldstein, 1973). Euler's definition of $\zeta(x) \in \mathbb{R}$ is given by the analytic power series,

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x} \qquad \text{for } x > 1 \in \mathbb{R}. \tag{C.1.1}$$

This series converges for $x > 0$, since $R = n^{-x} < 1, n > 1 \in \mathbb{N}$.[2]

In 1860 Riemann extended the zeta function into the complex plane, resulting in $\zeta(s)$, defined by the complex–analytic power series, identical to the Euler formula except $x \in \mathbb{R}$ has been replaced by $s \in \mathbb{C}$;

$$Z^{eta}(t) \leftrightarrow \zeta(s) \equiv \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \cdots = \sum_{n=1}^{\infty} \frac{1}{n^s} = \sum_{n=1}^{\infty} n^{-s} \quad \text{for } \Re\{s\} = \sigma > 1. \tag{C.1.2}$$

This formula converges for $\Re\{s\} > 1$ (Goldstein, 1973). To determine the formula in other regions of the s plane, we need to extend the series via analytic continuation. As it turns out, Euler's formulation provided detailed information about the structure of primes, going far beyond his original goal.

[1] https://www.youtube.com/watch?v=v9nyNBLCPks

[2] Sanity check: For example, let $n = 2$ and $x > 0$. Then $R = 2^{-\epsilon} < 1$, where $\epsilon \equiv \lim x \to 0^+$. Taking the log gives $\log_2 R = -\epsilon \log_2 2 = -\epsilon < 0$. Since $\log R < 0$, $R < 1$.

Table C.2: *An extended table of Laplace transforms. J_0, K_1 are Bessel functions of the first and second kind.*

$f(t)$	$\leftrightarrow$	$F(s)$		Nar
$\frac{d^{1/2}}{dt^{1/2}} f(t)u(t)$	$\leftrightarrow$	$\sqrt{s}F(s)$		half derivat
$\frac{d^{1/2}}{dt^{1/2}} u(t)$	$\leftrightarrow$	$\sqrt{s}$	case of viscous damping	half-derivat
$\frac{d}{dt}\frac{1}{\sqrt{\pi t}}u(t)$	$\leftrightarrow$	$\frac{s}{\sqrt{s}} = \sqrt{s}$		semi-induc
$\frac{1}{\sqrt{\pi t}}u(t)$	$\leftrightarrow$	$\frac{1}{\sqrt{s}}$		half integrati
$\text{erfc}(\alpha\sqrt{t})$	$\leftrightarrow$	$\frac{1}{s}e^{-2\alpha\sqrt{s}}$	(Morse and Feshbach (1953), p. 1582)	$\alpha > 0$; e
$J_0(at)u(t)$	$\leftrightarrow$	$\frac{1}{\sqrt{s^2+a^2}}$		J-Bes
$J_n(\omega_o t)\,u(t)$	$\leftrightarrow$	$\frac{\left(\sqrt{s^2+\omega_o^2} - s\right)^n}{\omega_o^n \sqrt{s^2+\omega_o^2}}$		
$J_1(t)u(t)/t$	$\leftrightarrow$	$\sqrt{s^2+1} - s$		
$J_1(t)u(t)/t + 2u(t)$	$\leftrightarrow$	$\sqrt{s^2+1} + s = e^{\sinh^{-1}(s)}$		
$\delta(t) + J_1(t)u(t)/t$	$\leftrightarrow$	$\sqrt{s^2+1}$		
$I_0(t)u(t)$	$\leftrightarrow$	$1/\sqrt{s^2-1}$		I-Bes
$u(t)/\sqrt{t+1}$	$\leftrightarrow$	$e^s \quad \frac{\pi}{s}\text{erfc}(\sqrt{s})$		Error functi
$\sqrt{t}u(t) * \sqrt{1+t}u(t)$	$\leftrightarrow$	$e^{s/2}K_1(s/2)/2s$		K-Bes
$Z^{eta}(t)$	$\leftrightarrow$	$\zeta(s)$		Riemann zeta functi

When we measure the voltage between two points, it is a crude approximation to the gradient based on the quasistatic approximation (Postulate P10). The pressure is also a potential, the gradient of which is a force density, which drives the volume velocity (flow).

In Sec. 5.6.8 we introduce the fundamental theorem of vector calculus (otherwise known as Helmholtz's decomposition theorem), which generalizes Ohm's law to include circulation (e.g., angular momentum, vorticity, and the related magnetic effects). To understand these generalizations in flow, we need to understand compressible and rotational fields (Table 5.3, p. 260), complex–analytic functions, and more mathematical physics history.

It is the *difference* in the potential (i.e., voltage, temperature, pressure) that is proportional to the flux. This can be viewed as a major simplification of the gradient relationship, justified by the quasistatic assumption (Postulate P10, p. 163).

The roots of the impedance are related to the eigenmodes of the system equations. The solutions to the equations are the eigenfunctions, evaluated at the eigenvalues (p. 182)

5.2.2 Scalar wave equation (Acoustics)

In this section we discuss the general solution to the wave equation, which has two forms: scalar waves (acoustics) and vector waves (electromagnetics). These have an important mathematical distinction but a similar solution space, one scalar and the other vector. We start with the scalar wave equation. The vector case will be discussed Sec. 5.7.2 (p. 267).

A good starting point for understanding PDEs is to explore the scalar wave equation (Eq. 3.1.3, p. 75). Acoustic wave propagation was first analyzed mathematically by Isaac Newton in his famous book *Principia* (1687), in which he first calculated the speed of sound based on the conservation of mass and momentum.

Early history: The study of wave propagation begins at least as early as Huygens (ca. 1678) (Pierce, 1981, p 15). The acoustic variables are the *pressure*,

$$\varrho(\boldsymbol{x}, t) \leftrightarrow \mathcal{P}(\boldsymbol{x}, s),$$

and the particle velocity,

$$\nu(\boldsymbol{x}, t) \leftrightarrow \mathcal{U}(\boldsymbol{x}, s).$$

To obtain a wave, we must include two basic components: the stiffness of air and its mass. The two equations are called (1) Newton's second law ($F = ma$) and (2) Hooke's law ($F = kx$), respectively. In vector form these equations are (1) Euler's equation (i.e., conservation of momentum density),

$$-\nabla\varrho(\boldsymbol{x}, t) = \rho_o \frac{\partial}{\partial t}\nu(\boldsymbol{x}, t) \leftrightarrow \rho_o s\,\mathcal{U}(\boldsymbol{x}, s), \tag{5.2.5}$$

which assumes the time-average density ρ_o is independent of time and position $\boldsymbol{x}$, and (2) the continuity equation (i.e., conservation of mass density),

$$-\nabla\cdot\nu(\boldsymbol{x}, t) = \frac{1}{\eta_o P_o}\frac{\partial}{\partial t}\varrho(\boldsymbol{x}, t) \leftrightarrow \frac{s}{\eta_o P_o}\,\mathcal{P}(\boldsymbol{x}, s) \tag{5.2.6}$$

(Pierce, 1981; Morse, 1948, p 295). Here $P_o = 10^5$ [Pa] is the barometric pressure and $\eta_o P_o$ is the dynamic (adiabatic) stiffness, with $\eta_o = 1.4$ (See p. 313). Combining Eqs. 5.2.5 and 5.2.6 (removing $\nu(\boldsymbol{x}, t)$) results in the three-dimensional scalar pressure wave equation

$$\nabla^2\varrho(\boldsymbol{x}, t) = \frac{1}{c_o^2}\frac{\partial^2}{\partial t^2}\varrho(\boldsymbol{x}, t) \leftrightarrow \frac{s^2}{c_o^2}\,\mathcal{P}(\boldsymbol{x}, s) \tag{5.2.7}$$

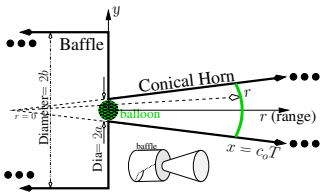


Figure 5.2: *Experimental setup showing a large pipe on the left terminating at the wall containing a small hole with a balloon, shown in green. At time $t = 0$ the balloon is pricked and a pressure pulse is released. The baffle on the left represents a semi-∞ long tube having a large radius compared to the horn input diameter $2a$. At time T the outbound pressure pulse $\varrho(r, T) = \delta(t - x/c_o)/r$ has reached a radius $x = r - r_o = c_o T$, where $r = x$ is the location of the source at the throat of the horn and r is measured from the vertex. At the throat of the horn $\mathcal{V}_+/A_+ = \mathcal{V}_-/A_-$. The term "horn" is used for the case of scalar waves, while the term "waveguide" is used when speaking of EM waves. When the propagation is constrained by a horn, or waveguide, the waves are "guided."*

with $c_o = \sqrt{\eta_o P_o/\rho_o}$ being the air sound velocity, and $\nu = \sqrt{\eta_o P_o \rho_o}$ being the characteristic resistance of air, assuming no visco-thermal losses.

Exercise #*10*

Show that Eqs. 5.2.5 and 5.2.6 can be reduced to Eq. 5.2.7.

Solution: Taking the divergence of Eq. 5.2.5 gives

$$-\nabla\cdot\nabla\varrho(\boldsymbol{x}, t) = \rho_o \frac{\partial}{\partial t}\nabla\cdot\nu(\boldsymbol{x}, t). \tag{5.2.8}$$

Note that $\nabla\cdot\nabla = \nabla^2$ (Table 5.1). Next, substituting Eq. 5.2.6 into the above equation results in the scalar wave equation, Eq. 5.2.7, since $c_o = \sqrt{\eta_o P_o/\rho_o}$. ■

5.2.3 The Webster horn equation (WHEN)

An important generalization of the problem of lossless plane-wave propagation in one-dimensional uniform tubes is known as *transmission line theory*. As depicted in Fig. 5.2, by allowing the area $A(r)$ [e.g., for the conical horn $A(r) = A_o(r/L)^2$ with $L = 1$ [m] and $A_0 \leq 4\pi$] of an acoustical waveguide (horn) to vary along the range axis r (the direction of wave propagation), we can explore general solutions to the wave equation. Classic applications of horns include vocal tract acoustics, loudspeaker design, cochlear mechanics, quantum mechanics (e.g., the hydrogen atom), and wave propagation in periodic media (Brillouin, 1953).

We must be more precise when defining the area $A(x)$: The area is *not* the cross-sectional area of the horn; rather it is the wavefront (isopressure) area, which is related to Gauss' law, since the gradient of the pressure defines the force that drives the mass flow (also called volume velocity).

For the scalar wave equation (Eq. 5.1.9, p. 229), the Webster Laplacian is

$$\nabla_r^2\,\varrho(r, t) = \frac{1}{A(r)}\frac{\partial}{\partial r}\left[A(r)\frac{\partial}{\partial r}\right]\varrho(r, t). \tag{5.2.9}$$

The Webster Laplacian is based on the quasistatic approximation (Postulate P10, p. 163), which requires that the frequency lie below the critical value $f_c = c_o/2d$—namely, that a half wavelength be greater than the horn diameter d (i.e., $d < \lambda/2$).[7] For the adult human ear canal,

[7] This condition may be written in several ways, the most common being $ka < 1$, where $k = 2\pi/\lambda$ and a is the horn radius. This may be expressed in terms of the diameter as $\frac{2\pi}{\lambda}\frac{d}{2} < 1$, or $d < \lambda/\pi < \lambda/2$. Thus $d < \lambda/2$ may be a more precise metric by the factor $\pi/2 \approx 1.6$. This is called the *half-wavelength assumption*, a synonym for the quasistatic approximation, and the Nyquist theorem (See DE-3, #2).

Appendix C

Laplace transforms $\mathcal{LT}$

Laplace transforms are discussed in Sec. 3.10 (p. 155), with the definition of the $\mathcal{LT}$ in Eq. 3.10.1 (p. 156). Level-I (basic) $\mathcal{LT}$s are listed in Table 3.6 (p. 156).

C.1 Tables of Laplace transforms

The following tables of $\mathcal{LT}$ and $\mathcal{LT}^{-1}$ are a convenient summary of their properties and evaluations for many different functions. Table 3.8 gives basic function properties such as convolution and the properties of step functions and frequency scaling. Table C.1 provides the commands for doing symbolic (computer algebra and calculus) transformations, which includes some unusual $\mathcal{LT}$s and Taylor series of the $\Gamma(s)$ function (Graham et al., 1994), a complex–analytic extension of the factorial. Table 3.9 gives the basic transforms typically used for more common calculations. Table C.2 provides extended less common transforms, such as the half-derivative and integration and Bessel functions.

These tables are available in most books on differential equations and remain a core technology for analytic methods for solving differential equations.

Table C.1: *Symbolic relationships among Laplace transforms. K_3 is a constant.*

syms	command	result
syms t s p	laplace$\left(t^{(p-1)}\right)$	$\Gamma(p)s^{-p}$
syms s	ilaplace(gamma(s))	$e^{e^{-t}}$
syms s t a	ilaplace(exp(-a*s)/s,s,t)	Heaviside$(t-a)$
syms Gamma s t	taylor(Gamma,s,t)	$\frac{1}{s} - \gamma + s\left(\frac{\gamma^2}{2} + \frac{\pi^2}{12}\right) +$ $s^2\left(+\frac{1}{6}\,\text{polygamma}\,(2,1) - \frac{\gamma\pi^2}{12} - \frac{\gamma^3}{6}\right) +$ $s^3 K_3 + \cdots$

C.1.1 $\mathcal{LT}^{-1}$ of the Riemann zeta function

The analytic properties of the zeta function $\zeta(s)$ have been a holy grail for mathematicians, starting with Euler, all of whom have made their reputation on that function. For the neophyte, $\zeta(s)$ is important because it is an analytic extension of the sieve, which is the prime identification method. Analytic continuation of the $\zeta(s)$ function was first stated by Riemann,

$d = 7.5$ [mm] and $f_c = (343/2 \cdot 7.5) \times 10^{-3} \approx 22.87$ [kHz], which is above the upper range of human hearing.

The term on the right of Eq. 5.2.9, which is identical to Eq. 5.1.9 (p. 229), is also the Laplacian for thin tubes (e.g., rectangular, spherical, and cylindrical coordinates). Thus the Webster horn "wave" equation is

$$\frac{1}{A(r)}\frac{\partial}{\partial r}\left[A(r)\frac{\partial}{\partial r}\right]\varrho(r,t) = \frac{1}{c_o^2}\frac{\partial^2}{\partial t^2}\varrho(r,t) \leftrightarrow \frac{s^2}{c_o^2}\mathcal{P}(r,s), \tag{5.2.10}$$

where $\varrho(r,t) \leftrightarrow \mathcal{P}(r,s)$ is the acoustic pressure in Pascals [Pa] (Hanna and Slepian, 1924; Mawardi, 1949; Eisner, 1967; Morse, 1948); Olson (1947, p. 101); Pierce (1981, p. 360). Extensive experimental analyses for various types of horns (conical, exponential, parabolic) along with a review of horn theory may be found in Goldsmith and Minton (1924). Of special interest is Eisner (1967) due to his history section and long list of relevant articles.

The limits of the Webster horn equation: It is commonly stated that the Webster horn equation (WHEN) is fundamentally limited and thus is an approximation that applies only to frequencies much lower than f_c (Morse, 1948; Shaw, 1970; Pierce, 1981). However, in all these discussions it is assumed that the area function $A(r)$ is the horn's cross-sectional area, not the area of the isopressure wavefront.

In the next section we show that this "limitation" may be avoided (subject to the $f < f_c$ quasistatic limit, Postulate P10, p. 163), making the Webster horn theory an "exact" solution for the lowest-order "plane-wave" eigenfunctions of Eq. 5.2.10. The limitation of the quasistatic approximation is that it "ignores" higher-order evanescent modes, which are naturally small since, being evanescent modes below their cutoff frequency, the wave number is real and thus they do not propagate (Hahn, 1941; Karal, 1953). This method is frequently called a *modal analysis* or *eigenanalysis*. This is the same approximation that is required to define an impedance, since every eigenmode has an impedance (Miles, 1948). These modes define a Hilbert "vector" space (also called an *eigenspace*).

As derived in Appendix G (p. 331), the acoustic variables (eigenfunctions) are redefined on the isopressure wavefront boundary for the pressure and the corresponding volume velocity (Hanna and Slepian, 1924; Morse, 1948; Pierce, 1981). The resulting acoustic impedance is then the ratio of the pressure to the volume velocity. This approximation is valid up to the frequency where the first cross-mode begins to propagate ($f > f_c$), which may be estimated given the roots of the Bessel eigenfunctions (Morse, 1948). Perhaps it should be noted that these ideas, which may come from acoustics, apply equally well to electromagnetics and quantum mechanics, and other wave phenomena.

Visco-thermal losses: When losses are to be included, the wave number $\kappa(s) = s/c_o$ must be replaced with Eq. D.1.5 (p. 313). This introduces dispersion in the wavefront due to the very small dispersive term $2\beta_0\sqrt{s}$, which contains a branch cut. When calculating the losses, we must be careful that they are always on the correct Riemann sheet. In cases where precise estimates of the wave properties and input impedance are required, this term is critical.

The best known examples of wave propagation are electrical and acoustic transmission lines. Such systems are loosely referred to as the *telegraph* or *telephone equations*, harking back to the early days of their discovery (Heaviside, 1892; Campbell, 1903; Brillouin, 1953; Feynman, 1970a). The telegraph equation characterizes the large resistance of the wire over long distances along with the stray capacitance of the wire to the ground (which at the time was

taken as the second conductor, to save wire). Thus the telegraph equation is best modeled by a diffusion line. The telephone equation included loading coils, consisting of inductors placed periodically in the wire, to increase the circuits inductance. This converted the circuit into a true transmission line. The loading coils were introduced by the AT&T engineer and mathematician George Ashley Campbell (Campbell, 1903, 1937), however they were first proposed and promoted by Heaviside.

In acoustics, waveguides are known as horns, such as the horn connected to the first phonographs from around the turn of the century (Webster, 1919). Thus the names reflect the historical development, back to a time when mathematics and its applications were related.

5.2.4 Matrix formulation of the WHEN

Newton's laws of conservation of momentum (Eq. 5.2.5) and mass (Eq. 5.2.6) are modern versions of Newton's starting point for calculating the horn lowest-order plane-wave eigenmode wave speed.

The acoustic equations for the average pressure $\mathcal{P}(r,\omega)$ and the volume velocity are derived in Appendix G, where the pressure and particle velocity equations (Eqs. G.1.4 and G.1.6) are transformed into a 2×2 matrix of acoustical variables, average pressure $\mathcal{P}(r,\omega)$ and volume velocity $\mathcal{V}(r,\omega)$:

$$-\frac{d}{dr}\begin{bmatrix}\mathcal{P}(r,\omega)\\ \mathcal{V}(r,\omega)\end{bmatrix} = \begin{matrix}0 & \frac{s\rho_0}{A(r)}\\ \frac{sA(r)}{\eta_o P_o} & 0\end{matrix}\begin{bmatrix}\mathcal{P}(r,\omega)\\ \mathcal{V}(r,\omega)\end{bmatrix}. \tag{5.2.11}$$

The equations

$$M(r)=\rho_o/A(r) \quad \text{and} \quad C(r)=A(r)/\eta_o P_o \tag{5.2.12}$$

define the per-unit-length mass and compliance of the horn (Ramo et al., 1965, p 213). The product of $\mathcal{P}(r,s)$ and $\mathcal{V}(r,s)$ defines the acoustic power [W/m^2], while their ratio defines the horn's admittance $Y_{in}^{\pm}(r,s)=\mathcal{V}^{\pm}/\mathcal{P}$, looking in the two directions (Pierce, 1981, p 37–41).

To obtain the Webster horn pressure equation Eq. 5.2.10 from Eq. 5.2.11, we take the partial derivative of the top equation

$$-\frac{\partial^2\mathcal{P}}{\partial r^2}=s\frac{\partial M(r)}{\partial r}\mathcal{V}+sM(r)\frac{\partial\mathcal{V}}{\partial r}$$

and then use the lower equation to remove $\partial\mathcal{V}/\partial r$,

$$\frac{\partial^2\mathcal{P}}{\partial r^2}-s\frac{\partial M(r)}{\partial r}\mathcal{V}=s^2M(r)C(r)\mathcal{P}=\frac{s^2}{c_o^2}\mathcal{P}.$$

Note that $c_o^2=MC=\left(\frac{\rho_o}{\cancel{A(r)}}\right)\cdot\left(\frac{\cancel{A(r)}}{\eta_o P_o}\right)$. In air $c_o=\sqrt{\eta_0 P_o/\rho_o}$.

We must then use the upper equation a second time to remove $\mathcal{V}$:

$$\frac{\partial^2}{\partial r^2}\mathcal{P}+\frac{1}{A(r)}\frac{\partial A(r)}{\partial r}\frac{\partial}{\partial r}\mathcal{P}=\frac{s^2}{c_o^2}\mathcal{P}(r,s). \tag{5.2.13}$$

By use of the chain rule, equations of this form may be directly integrated, since

$$\begin{aligned}\nabla_r\mathcal{P}&=\frac{1}{A(r)}\frac{\partial}{\partial r}\left[A(r)\frac{\partial}{\partial r}\right]\mathcal{P}(r,s)\\ &=\frac{\partial^2}{\partial r^2}\mathcal{P}(r,s)+\frac{1}{A(r)}\frac{\partial A(r)}{\partial r}\mathcal{P}(r,s).\end{aligned} \tag{5.2.14}$$

Impedance is not Hermitian: By definition, when a system is Hermitian its matrix is conjugate symmetric

$$\mathbf{Z}(\mathbf{s})=\mathbf{Z}^{\dagger}(\mathbf{s}),$$

a stronger condition than reciprocal, but not the symmetry of the Brune impedance matrix. A reciprocal Brune impedance is symmetric (not Hermitian).

B.3.2 Matrices with symmetry

Reversible systems:

When $\mathcal{A} = \mathcal{D}$

$$\boldsymbol{e} = \begin{matrix} - & \frac{\mathcal{B}}{\mathcal{C}} & + & \frac{\mathcal{B}}{\mathcal{C}} \\ 1 & & 1 & \end{matrix} \qquad \Lambda = \begin{bmatrix} \mathcal{A} - \sqrt{\mathcal{BC}} & 0 \\ 0 & \mathcal{A} + \sqrt{\mathcal{BC}} \end{bmatrix} \tag{B.3.5}$$

the transmission matrix is said to be *reversible,*, and the properties greatly simplify.

In this case note that $r_o = \sqrt{\mathcal{B}/\mathcal{C}}$ is the *characteristic impedance* and $\kappa(s) = 1/\sqrt{\mathcal{BC}}$ is the *propagation function*.

Reciprocal systems

For the case of the ABCD matrix the eigenvalues depend critically on reciprocity. A reciprocal matrix is when $\Delta_{\mathcal{T}} = 1$ and is anti-reciprocal for $\Delta_{\mathcal{T}} = -1$. It is helpful to display the eigenfunctions and values in terms of $\Delta_{\mathcal{T}}$ so this distinction is clear.

When the matrix is symmetric ($\mathcal{B} = \mathcal{C}$), thus $r_o = 1$ and the corresponding system is said to be *reciprocal*. Most physical systems are reciprocal. The determinant of the transmission matrix of a reciprocal network $\Delta_{\mathcal{T}} = \mathcal{AD} - \mathcal{BC} = 1$. For example, electrical networks composed of inductors, capacitors and resistors are always reciprocal. It follows that the complex impedance matrix is symmetric (Van Valkenburg, 1964a).

Magnetic systems such as dynamic loudspeakers are anti-reciprocal, and correspondingly $\Delta_{\mathcal{T}} = -1$. The impedance matrix of these loudspeakers is *skew symmetric* (Kim and Allen, 2013). All impedance matrices are either symmetric or anti-symmetric, depending on whether they are reciprocal (LRC networks) or anti-reciprocal (magnetic networks). These systems have complex eigenvalues with negative real parts, corresponding to damped (lossy) causal systems. This follows from conservation of energy.

The impedance matrix cannot be Hermitian, or the losses would be zero, because Hermitian matrices have real eigenvalues. A physical system having power losses cannot be Hermitian because the eigenvalues must have a negative real parts.

In summary, given a reciprocal system, the $\mathcal{T}$ matrix has $\Delta_{\mathcal{T}} = 1$, and the corresponding impedance matrix is symmetric (*not* Hermitian).

B.3.3 The impedance matrix

As previously discussed in Sec. 3.8 (p. 144), the impedance matrix $\mathcal{Z}$ is

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \boldsymbol{Z}(s) \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \frac{1}{\mathcal{C}} \begin{bmatrix} \mathcal{A} & \Delta_{\mathcal{T}} \\ 1 & \mathcal{D} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}.$$

Reciprocal systems have skew-symmetric impedance matrices, namely $z_{12} = z_{21}$ (i.e., $\Delta_{\mathcal{T}} = 1$). This condition is best understood using the T form of the impedance matrix, as shown in Fig. 3.9 (p. 146). When the system is both reversible $\mathcal{A} = \mathcal{D}$ and reciprocal ($\mathcal{B} = \mathcal{C}$), the impedance matrix simplifies to

$$\boldsymbol{Z}(s) = \frac{1}{\mathcal{C}} \begin{bmatrix} \mathcal{A} & 1 \\ 1 & \mathcal{A} \end{bmatrix}.$$

For such systems there are only two degrees of freedom, $\mathcal{A}$ and $\mathcal{C}$. As discussed previously in Sec. 3.8 (p. 144), each of these has a physical meaning: $1/\mathcal{A}$ is the *Thévenin source voltage* assuming a voltage drive, and $\mathcal{B}/\mathcal{A}$ is the *Thévenin impedance* (Sec. 3.8.3, p. 145).

This is equivalent to integration by parts, with integration factor $A(r)$. Finally we set $\kappa(s) \equiv s/c_o$, which later may be generalized to include visco-thermal losses (Eq. D.1.5, p. 313).

Merging Eqs. 5.2.13 and 5.2.14 results in the Webster horn equation (WHEN) (Eq. 5.2.10, p. 241):

$$\frac{1}{A(r)} \frac{\partial}{\partial r} A(r) \frac{\partial}{\partial r} \mathcal{P}(r,s) = \kappa^2(s)\, \mathcal{P}(r,s) \leftrightarrow \frac{1}{c_o^2} \frac{\partial^2}{\partial t^2} \varrho(r,t). \tag{5.2.15}$$

Equations having this form are known as *Sturm-Liouville equations*. This important class of ordinary differential equations follows from the use of separation of variables of the Laplacian in any (i.e., every) separable coordinate systems (Morse and Feshbach, 1953, p 494–523). The frequency domain eigensolutions are denoted $\mathcal{P}^{\pm}(r,s)$.

We transform the three-dimensional acoustic wave equation into acoustic variables (Eq. 5.2.7) in Appendix G by the application of Gauss's law, resulting in the one-dimensional WHEN (Eq. 5.2.10), which is a nonsingular Sturm-Liouville equation.[8] It seems significant that the integration factor corresponds to the horn's area function. Thus we have demonstrated that Eqs. 5.2.7 and 5.2.11 reduce to to Eq. 5.2.15 in a horn.

[8] The Webster horn equation is closely related to Schrödinger's equation (Salmon, 1946).

5.3 Problems VC-1

5.3.1 Topics of this homework:

Vector algebra and fields in $\mathbb{R}^3$, gradient and scalar Laplacian operators, definitions of divergence and curl, Gauss's (divergence) and Stokes's (curl) laws, system classification (postulates).

5.3.2 Scalar fields and the ∇ operator

Problem # *1: Let $T(x,y) = x^2 + y$ be an analytic scalar temperature field in two dimensions (single-valued $\in \mathbb{R}^2$).*

– 1.1: Find the gradient of $T(\boldsymbol{x})$ and make a sketch of T and the gradient.

– 1.2: Compute $\nabla^2 T(\boldsymbol{x})$ to determine whether $T(\boldsymbol{x})$ satisfies Laplace's equation.

– 1.3: Sketch the iso-temperature contours at $T = -10, 0, 10$ degrees.

– 1.4: The heat flux[9] *is defined as $\boldsymbol{J}(x,y) = -\kappa(x,y)\nabla T$, where $\kappa(x,y)$ is denotes the thermal conductivity at the point (x,y). Given that $\kappa = 1$ everywhere (the medium is homogeneous), plot the vector $\boldsymbol{J}(x,y) = -\nabla T$ at $x = 2$, $y = 1$. Be clear about the origin, direction, and length of your result.*

– 1.5: Find the vector $\perp$ to $\nabla T(x,y)$—that is, tangent to the iso-temperature contours. Hint: Sketch it for one (x,y) point (e.g., $2, 1$) and then generalize.

– 1.6: The thermal resistance R_T is defined as the potential drop ΔT over the magnitude of the heat flux $|\boldsymbol{J}|$. At a single point the thermal resistance is

$$R_T(x,y) = -\nabla T/|\boldsymbol{J}|.$$

How is $R_T(x,y)$ related to the thermal conductivity $\kappa(x,y)$?

Problem # *2: Acoustic wave equation*
Note: In this problem, we will work in the frequency domain.

[9] The heat flux is proportional to the change in temperature times the thermal conductivity κ of the medium.


```
A = [1 2; 1 1]; %define the matrix
[E,D] = eig(A); %compute the eigenvector and eigenvalue matrices
A*E-E*D %This should be $\approx 0$, within numerical error.
E*D-D*E %This is not zero
```

Summary:

Thus far we have shown that for the case of Pell matrix with $N = 2$, the normalized eigenmatrix and its inverse is

$$\boldsymbol{e} = [\boldsymbol{e}_+, \boldsymbol{e}_-] = \frac{1}{\sqrt{3}}\begin{bmatrix}\sqrt{2} & -\sqrt{2}\\ 1 & 1\end{bmatrix} \qquad \boldsymbol{e}^{-1} = \frac{\sqrt{6}}{4}\begin{bmatrix}1 & \sqrt{2}\\ -1 & \sqrt{2}\end{bmatrix}$$

and the eigenmatrix is

$$\Lambda = \begin{bmatrix}\lambda_+ & 0\\ 0 & \lambda_-\end{bmatrix} = \begin{bmatrix}1+\sqrt{2} & 0\\ 0 & 1-\sqrt{2}\end{bmatrix}.$$

B.3 Symbolic analysis of $\mathcal{T}\boldsymbol{e} = \boldsymbol{e}\Lambda$

B.3.1 The 2×2 transmission matrix

Here we assume

$$\mathcal{T} = \begin{bmatrix}\mathcal{A} & \mathcal{B}\\ \mathcal{C} & \mathcal{D}\end{bmatrix}$$

with $\Delta_{\mathcal{T}} = 1$.

The eigenvectors $\boldsymbol{e}_\pm$ of $\mathcal{T}$ are

$$\boldsymbol{e}_\pm = \begin{matrix}\frac{1}{2\mathcal{C}}\ (\mathcal{A}-\mathcal{D}) \mp \ \sqrt{(\mathcal{A}-\mathcal{D})^2 + 4\mathcal{B}\mathcal{C}}\\ 1\end{matrix} \tag{B.3.1}$$

and eigenvalues are

$$\lambda_\pm = \frac{1}{2}\ (\mathcal{A}+\mathcal{D}) \mp \ \sqrt{(\mathcal{A}-\mathcal{D})^2 + 4\mathcal{B}\mathcal{C}}\ . \tag{B.3.2}$$

Thus the expression under the radical may be rewritten in terms of the determinant of $\mathcal{T}$ (i.e., $\Delta_{\mathcal{T}} = \mathcal{A}\mathcal{D} - \mathcal{B}\mathcal{C}$) since

$$(\mathcal{A}-\mathcal{D})^2 - (\mathcal{A}+\mathcal{D})^2 = -4\mathcal{A}\mathcal{D}.$$

The for the ABCD matrix the expression under the radical becomes

$$\begin{aligned}(\mathcal{A}-\mathcal{D})^2 + 4\mathcal{B}\mathcal{C} &= \mathcal{A}^2 + \mathcal{D}^2 - 4\mathcal{A}\mathcal{D} + 4\mathcal{B}\mathcal{C}\\ &= \mathcal{A}^2 + \mathcal{D}^2 - 4\Delta_{\mathcal{T}}.\end{aligned}$$

Rewriting the eigenvectors and eigenvalues in terms of $\Delta_{\mathcal{T}} = \pm 1$, we find

$$\boldsymbol{e}_\pm = \begin{bmatrix}\frac{1}{\mathcal{C}}\left[\frac{\mathcal{A}-\mathcal{D}}{2} \mp \ \sqrt{\left(\frac{\mathcal{A}+\mathcal{D}}{2}\right)^2 \mp \Delta_{\mathcal{T}}}\right]\\ 1\end{bmatrix} \tag{B.3.3}$$

and

$$\lambda_\pm = \left[\frac{\mathcal{A}+\mathcal{D}}{2} \mp \ \sqrt{\left(\frac{\mathcal{A}+\mathcal{D}}{2}\right)^2 \mp \Delta_{\mathcal{T}}}\right]. \tag{B.3.4}$$

Note this may be further simplified since the radical is the same.

B.2.1 Pell equation eigenvalue-eigenvector analysis

Here we show how to compute the eigenvalues and eigenvectors for the 2×2 Pell matrix for $N = 2$

$$\boldsymbol{a} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

The Matlab/Octave command `[E,D]=eig(A)` returns the eigenvector matrix $\boldsymbol{e}$

$$\boldsymbol{e} = [\boldsymbol{e}_+, \boldsymbol{e}_-] = \frac{1}{\sqrt{3}} \begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0.8165 & -0.8165 \\ 0.5774 & 0.5774. \end{bmatrix}$$

and the eigenvalue matrix Λ (Matlab/Octave's D)

$$\Lambda \equiv \begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix} = \begin{bmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2.4142 & 0 \\ 0 & -0.4142 \end{bmatrix}.$$

Table B.1: *Summary of the solution of Pell's equation due to the Pythagoreans using matrix recursion, for the case of N=3. The integer solutions are shown on the right. Note that $x_n/y_n \to \sqrt{3}$, in agreement with the Euclidean algorithm. The Matlab/Octave program for generating this data is `PellSol3.m`. It seems likely that the powers of β_0 could be absorbed in the starting solution, and then be removed from the recursion.*

Pell's Equation for $N = 3$
Case of $N = 3$ & $[x_0, y_0]^T = [1, 0]^T$, $\beta_0 = \jmath/\sqrt{2}$
Note: $x_n^2 - 3y_n^2 = 1$, $\quad x_n/y_n \underset{\infty}{\longrightarrow} \sqrt{3}$

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \beta_0 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \beta_0 \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad (1\beta_0)^2 - 3(1\beta_0)^2 = 1$$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \beta_0^2 \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \beta_0^2 \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \left(4\beta_0^2\right)^2 - 3\left(2\beta_0^2\right)^2 = 1$$

$$\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \beta_0^3 \begin{bmatrix} 10 \\ 6 \end{bmatrix} = \beta_0^3 \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} \qquad \left(10\beta_0^3\right)^2 - 3\left(6\beta_0^3\right)^2 = 1$$

$$\begin{bmatrix} x_4 \\ y_4 \end{bmatrix} = \beta_0^4 \begin{bmatrix} 28 \\ 16 \end{bmatrix} = \beta_0^4 \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 10 \\ 6 \end{bmatrix} \qquad \left(28\beta_0^4\right)^2 - 3\left(16\beta_0^4\right)^2 = 1$$

$$\begin{bmatrix} x_5 \\ y_5 \end{bmatrix} = \beta_0^5 \begin{bmatrix} 76 \\ 44 \end{bmatrix} = \beta_0^5 \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 28 \\ 16 \end{bmatrix} \qquad \left(76\beta_0^5\right)^2 - 3\left(44\beta_0^5\right)^2 = 1$$

Pell's equation for $N = 3$

In Table B.1, Pell's equation for $N = 3$ is given, with $\beta_0 = \jmath/\sqrt{2}$. Perhaps try other trivial solutions such as $[-1, 0]^T$ and $[\pm\jmath, 0]^T$, to provide clues to the proper value of β_0 for cases where $N > 3$.[2]

Example: I suggest that you verify $\boldsymbol{e}\Lambda \neq \Lambda\boldsymbol{e}$ and $\boldsymbol{a}\boldsymbol{e} = \boldsymbol{e}\Lambda$ with Matlab/Octave. Here is the Matlab/Octave program which does this:

[2] My student Kehan found the general formula for β_o.

– 2.1: The basic equations of acoustics in one dimension are

$$-\frac{\partial}{\partial x}\mathcal{P} = \rho_o s\,\vec{\mathcal{V}} \quad \text{and} \quad -\frac{\partial}{\partial x}\,\vec{\mathcal{V}} = \frac{s}{\eta_o P_o}\mathcal{P}.$$

Here $\mathcal{P}(x,\omega)$ is the pressure (in the frequency domain), $\mathcal{V}(x,\omega)$ is the volume velocity (the integral of the velocity over the wavefront with area A), $s = \sigma + \omega\jmath$, $\rho_o = 1.2$ is the specific density of air, $\eta_o = 1.4$, and P_o is the atmospheric pressure (i.e., 10^5 Pa). Note that the pressure field $\mathcal{P}$ is a scalar (pressure does not have direction), while the volume velocity field $\vec{\mathcal{V}}$ is a vector (velocity has direction).

We can generalize these equations to three dimensions using the ∇ operator

$$-\nabla\mathcal{P} = \rho_o s\,\vec{\mathcal{V}} \quad \text{and} \quad -\nabla\cdot\,\vec{\mathcal{V}} = \frac{s}{\eta_o P_o}\mathcal{P}.$$

– 2.2: Starting from these two basic equations, derive the scalar wave equation in terms of the pressure $\mathcal{P}$,

$$\nabla^2\mathcal{P} = \frac{s^2}{c_0^2}\mathcal{P},$$

where c_0 is a constant representing the speed of sound.

– 2.3: What is c_0 in terms of η_0, ρ_0, and P_0?

– 2.4: Rewrite the pressure wave equation in the time domain using the time derivative property of the Laplace transform [e.g., $dx/dt \leftrightarrow sX(s)$]. For your notation, define the time-domain signal using a lowercase letter, $p(x,y,z,t) \leftrightarrow \mathcal{P}$.

5.3.3 Vector fields and the ∇ operator

5.3.4 Vector algebra

Problem # *3: Let $\boldsymbol{R}(x,y,z) \equiv x(t)\hat{\mathbf{x}} + y(t)\hat{\mathbf{y}} + z(t)\hat{\mathbf{z}}$.*

– 3.1: If a, b, and c are constants, what is $\boldsymbol{R}(x,y,z)\cdot\boldsymbol{R}(a,b,c)$?

– 3.2: If a, b, and c are constants, what is $\frac{d}{dt}\left(\boldsymbol{R}(x,y,z)\cdot\boldsymbol{R}(a,b,c)\right)$?

Problem # *4: Find the divergence and curl of the following vector fields:*

– 4.1: $\boldsymbol{v} = \hat{\mathbf{x}} + \hat{\mathbf{y}} + 2\hat{\mathbf{z}}$

– 4.2: $\boldsymbol{v}(x,y,z) = x\hat{\mathbf{x}} + xy\hat{\mathbf{y}} + z^2\hat{\mathbf{z}}$

– 4.3: $\boldsymbol{v}(x,y,z) = x\hat{\mathbf{x}} + xy\hat{\mathbf{y}} + \log(z)\hat{\mathbf{z}}$

– 4.4: $\boldsymbol{v}(x,y,z) = \nabla(1/x + 1/y + 1/z)$

5.3.5 Vector and scalar field identities

Problem # *5: Find the divergence and curl of the following vector fields:*

– 5.1: $\boldsymbol{v} = \nabla\phi$*, where* $\phi(x, y) = xe^y$

– 5.2: $\boldsymbol{v} = \nabla \times \boldsymbol{A}$*, where* $\boldsymbol{A} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$

– 5.3: $\boldsymbol{v} = \nabla \times \boldsymbol{A}$*, where* $\boldsymbol{A} = y\hat{\mathbf{x}} + x^2\hat{\mathbf{y}} + z\hat{\mathbf{z}}$

– 5.4: For any differentiable vector field $\boldsymbol{V}$*, write two vector calculus identities that are equal to zero.*

– 5.5: What is the most general form a vector field may be expressed in, in terms of scalar Φ *and vector* $\boldsymbol{A}$ *potentials?*

Problem # *6: Perform the following calculations. If you can state the answer without doing the calculation, explain why.*

– 6.1: Let $\boldsymbol{v} = \sin(x)\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$*. Find* $\nabla \cdot (\nabla \times \boldsymbol{v})$.

– 6.2: Let $\boldsymbol{v} = \sin(x)\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$*. Find* $\nabla \times (\nabla\sqrt{\boldsymbol{v} \cdot \boldsymbol{v}})$

– 6.3: Let $\boldsymbol{v}(x, y, z) = \nabla(x + y^2 + \sin(\log(z))$*. Find* $\nabla \times \boldsymbol{v}(x, y, z)$.

5.3.6 Integral theorems

Problem # *7: For each of the following problems, in a few words, identify either Gauss's or Stokes's law, define what it means, and explain the formula that follows the question.*

– 7.1: What is the name of this formula?

$$\int_{\mathcal{S}} \hat{\mathbf{n}} \cdot \boldsymbol{v}\, dA = \int_{\mathcal{V}} \nabla \cdot \boldsymbol{v}\, dV.$$

– 7.2: What is the name of this formula?

$$\int_{S} (\nabla \times \boldsymbol{V}) \cdot d\boldsymbol{S} = \int_{C} \boldsymbol{V} \cdot d\boldsymbol{R}$$

Give one important application.

– 7.3: Describe a key application of the vector identity

$$\nabla \times (\nabla \times \vec{V}) = \nabla(\nabla \cdot \vec{V}) - \boldsymbol{\nabla^2}\vec{V}.$$

If we divide the top row by $\sqrt{N}$ the two rows are identical, since the matrix must be singular. Thus this matrix equation gives two identical equations. This is the price of an over-specified equation (the singular matrix is degenerate).

We can determine each eigenvectors direction, but not their magnitudes.

Following the same procedure for $\lambda_- = -\sqrt{N}$, the equation for $\boldsymbol{e}_-$ is

$$\begin{bmatrix} -\sqrt{N} & N \\ 1 & -\sqrt{N} \end{bmatrix} \begin{bmatrix} e_1^- \\ e_2^- \end{bmatrix} = 0.$$

As before, this matrix is singular. Here $e_1^- - \sqrt{N}e_2^- = 0$, thus the eigenvector is $e^- = c\,[\sqrt{N}, 1]^T$ where c is a normalization constant.

Thus the *unnormalized* eigenmatrix is

$$\boldsymbol{e} = \begin{bmatrix} e_1^+ & e_2^- \\ e_2^+ & e_2^- \end{bmatrix} = \begin{bmatrix} \sqrt{N} & -\sqrt{N} \\ 1 & 1 \end{bmatrix}.$$

Normalization of the eigenvectors:

The constant c may be determined by normalizing the eigenvectors to have unit length. Since we cannot determine the length, we set it to 1. Thus the degeneracy may be resolved by the one degree of freedom normalization

$$\left(\pm\sqrt{N}\right)^2 + 1^2 = N + 1 = 1/c^2.$$

Thus the normalization factor to force each eigen vector to have length 1 is $c = 1/\sqrt{N+1}$.

B.2 Pell equation solution example

Sec. 2.5.2 (p. 62) showed that the solution $[x_n, y_n]^T$ to Pell's equation is given by powers of the Pell matrix $\boldsymbol{a}$. For $N = 2$, in Sec. 2.5.2 we found the explicit formula for $[x_n, y_n]^T$, based on powers of the Pell matrix

$$\boldsymbol{a} = 1\jmath \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}. \tag{B.2.1}$$

This recursive solution to Pell's equation (Eq. 2.5.2) is Eq. 2.5.4 (p. 62). Thus we need powers of A, that is A^n, which gives an explicit expression for $[x_n, y_n]^T$. By the diagonalization of A, its powers are simply the powers of its eigenvalues.

From Matlab/Octave with $N = 2$ the eigenvalues of Eq. B.2.1 are $\lambda_\pm \approx [2.4142\jmath, -0.4142\jmath]$ (i.e., $\lambda_\pm = 1\jmath(1 \pm \sqrt{2})$). The solution for $N = 3$ is shown on page 296.

Once the matrix has been diagonalized, one may compute powers of that matrix as powers of the eigenvalues. This results in the general solution given by

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = 1\jmath^n \boldsymbol{a}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1\jmath^n \boldsymbol{e}\Lambda^n \boldsymbol{E}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The eigenvalue matrix D is diagonal with the eigenvalues sorted, largest first. The Matlab/Octave command `[E,D]=eig(A)` is helpful to find D and E given any A. As we saw above,

$$\Lambda = 1\jmath \begin{bmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{bmatrix} \approx \begin{bmatrix} 2.414\jmath & 0 \\ 0 & -0.414\jmath \end{bmatrix}.$$

$\boldsymbol{a} - \lambda_\pm \boldsymbol{I}_2$ is singular:

If we restrict Eq. B.1.1 to a single eigenvector (one of $\boldsymbol{e}_\pm$), along with the corresponding eigenvalue $\lambda_\pm$, we obtain the two matrix equations

$$\boldsymbol{a}\boldsymbol{e}_\pm = \boldsymbol{e}_\pm \lambda_\pm = \lambda_\pm \boldsymbol{e}_\pm.$$

Note the swap in the order of $\boldsymbol{e}_\pm$ and $\lambda_\pm$. Since $\lambda_\pm$ is a scalar, this is legal (and critically important), since this allows us to factor out $\boldsymbol{e}_\pm$

$$(\boldsymbol{a} - \lambda_\pm \boldsymbol{I}_2)\boldsymbol{e}_\pm = 0. \tag{B.1.6}$$

The matrix $\boldsymbol{a} - \lambda_\pm \boldsymbol{I}_2$ must be singular because when it operates on $\boldsymbol{e}_\pm$, having nonzero norm, it must be zero.

It follows that its determinant (i.e., $|(\boldsymbol{a} - \lambda_\pm \boldsymbol{I}_2)| = 0$) must be zero. This equation uniquely determines the eigenvalues $\lambda_\pm$.

B.1.1 Calculating the eigenvalues $\lambda_\pm$

The eigenvalues $\lambda_\pm$ of $\boldsymbol{a}$ may be determined from $|(\boldsymbol{a} - \lambda_\pm \boldsymbol{I}_2)| = 0$. As an example we let $\boldsymbol{a}$ be Pell's equation (Eq. 2.5.4, p. 62). In this case the eigenvalues may be found from

$$\begin{vmatrix} 1-\lambda_\pm & N \\ 1 & 1-\lambda_\pm \end{vmatrix} = (1-\lambda_\pm)^2 - N = 0,$$

thus $\lambda_\pm = (1 \mp \sqrt{N})$.[1]

B.1.2 Calculating the eigenvectors $\boldsymbol{e}_\pm$

Once the eigenvalues have been determined, they are substituted into Eq. B.1.6, which determines the eigenvectors $\boldsymbol{e} = \left[\boldsymbol{e}_+, \boldsymbol{e}_-\right]$, by solving

$$(\boldsymbol{a} - \lambda_\pm)\mathbf{e}_\pm = \begin{bmatrix} 1-\lambda_\pm & 2 \\ 1 & 1-\lambda_\pm \end{bmatrix} \boldsymbol{e}_\pm = 0, \tag{B.1.7}$$

where $1 - \lambda_\pm = 1 - (1 \mp \sqrt{N}) = \pm\sqrt{N}$, thus the Pell equation eigenvalues are

$$\lambda_\pm = 1 \mp \sqrt{N}.$$

Recall that Eq. B.1.6 is singular because we are using an eigenvalue, and each eigenvector is pointing in a unique direction (this is why it is singular). You might expect that this equation has no solution. In some sense you would be correct. When we solve for $\boldsymbol{e}_\pm$, the two equations defined by Eq. B.1.6 are *co-linear* (the two equations describe parallel lines so their scalar product is one). This follows from the fact that there is only one eigenvector for each eigenvalue.

Since there is only one eigenvalue we are expecting trouble, yet we may proceeding to solve for $\boldsymbol{e}_+ = [e_1^+, e_2^+]^T$ with eigenvalue $+\sqrt{N}$

$$\begin{bmatrix} \sqrt{N} & N \\ 1 & \sqrt{N} \end{bmatrix} \begin{bmatrix} e_1^+ \\ e_2^+ \end{bmatrix} = 0.$$

[1] It is a convention to order the eigenvalues from largest to smallest.

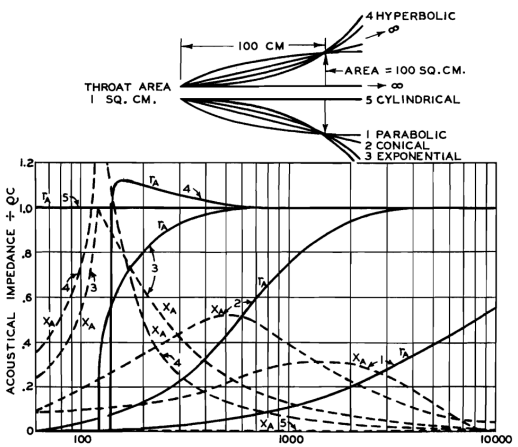


Figure 5.3: *Throat acoustical resistance r_A and acoustical reactance x_A, frequency characteristics of infinite eigenfunctions of the parabolic, conical, exponential, hyperbolic, and cylindrical horns, having a throat area of 1 [cm^2]. Note how the "critical" frequency (defined here as the frequency where the reactive and real parts of the radiation impedance are equal) of the horn reduces dramatically with the type of horn. For the uniform horn, the reactive component is zero, so there is no cutoff frequency. For the parabolic horn (#1), the cutoff is around 3 kHz. For the conical horn (#2), the cutoff is at 0.6 [kHz]. For the exponential horn (#3), the critical frequency is around 0.18 [kHz], which is one-16th that of the parabolic horn. For each horn the cross-sectional area is taken as 100 [cm^2] at a distance of $L = 1$ [m] from the throat (Olson, 1947, p 101); (Morse, 1948, p 283).*

5.4 Examples of finite-length horns

Figure 5.3 is taken from the classic book by Olson (1947, p. 101), showing the theoretical acoustical radiation impedance $Z_{rad}(r, \omega)$ for five horns having different area functions, resulting in frequence dependent impedance between 50 [Hz] to 10 [kHz]. The parabolic horn (#1) has a mild peak in the reactive input impedance just below 2 [kHz]. The conical (megaphone) horn (#2) has a maximum reactive peat at 5 [kHz], close to the frequency where the real and imaginary parts are equal. The most interesting horns are #3 and #4 which, have a razor sharp cutoff near 150 [Hz], where the impedance abruptly switches from resistive to reactive. The cylindrical horn (#5) has a constant impedance. All the impedances have been normalized to have their impedance of 1 at high frequncies.

Table 5.2: *Horns and their properties for $N = 1$, 2, and 3 dimensions, along with the exponential horn* (EXP)*. The range variable goes from $r_o \le r \le L$ [m] with area $1 \le A(r) \le 100$ [cm^2]. $F(r)$ is the coefficient on $\mathcal{P}_x$, $\kappa(s) \equiv s/c_o$, where c_o is the speed of sound and $s = \sigma + \omega\jmath$ is the Laplace frequency. The horn's eigenfunctions are $\mathcal{P}^\pm(\xi, s) \leftrightarrow \varrho^\pm(\xi, t)$. When $\pm$ is indicated, the outbound solution corresponds to the negative sign. Eigenfunctions $H_0^\pm(\xi, s)$ are outbound and inbound Hankel functions. The rightmost column is the input radiation admittance normalized by the* characteristic admittance *$\mathcal{Y}_r(r) = A(r)/\rho_o c_o$.*

N	Name	radius	Area/A_o	$F(r)$	$\mathcal{P}^\mp(r,s)$	$\varrho^\mp(r_o,t)$	$Y_{rad}^\mp/\mathcal{Y}_r$
1D	uniform	1	1	0	$e^{\mp\kappa(s)r}$	$\delta(t)$	1
2D	parabolic	$\sqrt{r}$	r	$1/r$	$H_0^\mp(-\jmath\kappa(s)r)$	—	$\frac{-\jmath r_o H_1^\mp}{H_0^\mp}$
3D	conical	r	r^2	$2/r$	$e^{\mp\kappa(s)r}/r$	$\delta(t) \pm \frac{c_o}{r_o}u(t)$	$1 \pm c_o/sr_o$
EXP	exponential	e^{mr}	e^{2mr}	$2m$	$e^{-\left(m\mp\sqrt{m^2+\kappa^2}\right)r}$	$e^{-mr}E(t)$	Eq. 5.4.11

Table 5.2 summarizes the properties of four of these horns: uniform (cylindrical) ($A = A_o$), parabolic ($A(r) = A_o r$), conical (spherical) ($A(r) = A_o r^2$), and exponential ($A(r) = A_o e^{2mr}$), where m is called the flair parameter.

5.4.1 Uniform Horn #5

The one-dimensional wave equation [$A(r) = A_o = 1$ [cm^2]] is

$$\frac{d^2}{dr^2}\,\mathcal{P} = \kappa^2(s)\,\mathcal{P},$$

thus the roots (eigenvalues) are $\kappa_\pm^2(s) \equiv \pm s^2/c_o^2$.

Solution: The two eigenfunctions of this equation are the two d'Alembert waves (Eq. 4.4.1, p. 188):

$$\varrho(x,t) = \mathcal{P}_0^+\varrho^+(t-x/c) + \mathcal{P}_0^-\varrho^-(t+(x-L)/c) \;\leftrightarrow\; \mathcal{P}_0^+e^{-\kappa(s)x} + \mathcal{P}_0^-e^{\kappa(s)(x-L)},$$

where $\mathcal{P}_0^\pm \in \mathbb{C}$ are wave amplitudes and $\kappa(s) = s/c_o = \jmath\omega/c$ is called the *propagation function* (also known as the wave-evolution function, propagation constant, and wave number).

Note that for the uniform lossless horn, $\omega/c_o = 2\pi/\lambda \geq 0$. It is convenient to normalize $\mathcal{P}_0^+ = 1$ and $\mathcal{P}_L^- = 1$. The characteristic admittance $\mathcal{Y}_r(x)$ (Table 5.2) is independent of direction. The signs must be physically chosen, with the velocity $\mathcal{V}^\pm$ into the port, to ensure that the real part $\mathcal{Y}_r \geq 0$, to assure the system remains stable.

Applying the boundary conditions: The general solution in terms of the eigenvector matrix, evaluated at $x = L$, is

$$\begin{bmatrix}\mathcal{P}(x)\\ \mathcal{V}(x)\end{bmatrix}_L = \begin{bmatrix}e^{-\kappa x} & e^{\kappa(x-L)}\\ \mathcal{Y}_r e^{-\kappa x} & -\mathcal{Y}_r e^{\kappa(x-L)}\end{bmatrix}_L \begin{bmatrix}\mathcal{P}_0^+\\ \mathcal{P}_0^-\end{bmatrix}_L = \begin{bmatrix}e^{-\kappa L} & 1\\ \mathcal{Y}_r e^{-\kappa L} & -\mathcal{Y}_r\end{bmatrix}\begin{bmatrix}\mathcal{P}_0^+\\ \mathcal{P}_0^-\end{bmatrix}_L, \tag{5.4.1}$$

where $\mathcal{P}_0^+$ and $\mathcal{P}_0^-$ are the relative amplitudes of the two unknown eigenfunctions to be determined by the boundary conditions at $x = 0, L$, $\kappa = s/c$, and $\mathcal{Y}_r = 1/\mathcal{Z}_r = A_o/\rho_o c$.

Solving Eq. 5.4.1 for $\mathcal{P}_0^+$ and $\mathcal{P}_0^-$ with determinant $\Delta = -2\mathcal{Y}_r e^{-\kappa L}$,we get

$$\begin{bmatrix}\mathcal{P}_L^+\\ \mathcal{P}_L^-\end{bmatrix} = \frac{-1}{2\mathcal{Y}_r e^{-\kappa L}}\begin{bmatrix}-\mathcal{Y}_r & -1\\ -\mathcal{Y}_r e^{-\kappa L} & e^{-\kappa L}\end{bmatrix}\begin{bmatrix}\mathcal{P}\\ \mathcal{V}\end{bmatrix}_L = \frac{1}{2}\begin{bmatrix}e^{\kappa L} & -\mathcal{Z}e^{\kappa L}\\ 1 & \mathcal{Z}\end{bmatrix}\begin{bmatrix}\mathcal{P}\\ -\mathcal{V}\end{bmatrix}_L. \tag{5.4.2}$$

In the final step we swapped all the signs, including those on $\mathcal{V}$, and moved $\mathcal{Z}_r = 1/\mathcal{Y}_r$ inside the matrix.

We can uniquely determine these two weights $\mathcal{P}_L^+$, $\mathcal{P}_L^-$ given the pressure and velocity at the boundary $x = L$, which is typically determined by the load impedance ($Z_L(s) = \mathcal{P}_L/\mathcal{V}_L$).

The weights may now be substituted back into Eq. 5.4.1 to determine the pressure and velocity amplitudes at any point $0 \leq x \leq L$:

$$\begin{bmatrix}\mathcal{P}\\ \mathcal{V}\end{bmatrix}_x = \frac{1}{2}\begin{bmatrix}e^{-\kappa x} & e^{\kappa(x-L)}\\ \mathcal{Y}_r e^{-\kappa x} & -\mathcal{Y}_r e^{\kappa(x-L)}\end{bmatrix}_x \begin{bmatrix}e^{\kappa L} & -\mathcal{Z}e^{\kappa L}\\ 1 & \mathcal{Z}\end{bmatrix}\begin{bmatrix}\mathcal{P}\\ -\mathcal{V}\end{bmatrix}_L. \tag{5.4.3}$$

Setting $x = 0$ and multiplying these out give the final transmission matrix:

$$\begin{bmatrix}\mathcal{P}\\ \mathcal{V}\end{bmatrix}_0 = \frac{1}{2}\begin{bmatrix}e^{\kappa L}+e^{-\kappa L} & \mathcal{Z}_r(e^{\kappa L}-e^{-\kappa L})\\ \mathcal{Y}_r(e^{\kappa L}-e^{-\kappa L}) & e^{\kappa L}+e^{-\kappa L}\end{bmatrix}_x \begin{bmatrix}\mathcal{P}\\ -\mathcal{V}\end{bmatrix}_L. \tag{5.4.4}$$

Note that the diagonal terms are $\cosh\kappa L$ and the off-diagonal terms are $\sinh\kappa L$.

Appendix B

Eigenanalysis

Eigenanalysis is ubiquitous in engineering applications. It is useful in solving differential and difference equations, data-science applications, numerical approximation and computing, and linear algebra applications. Typically one must take a course in linear algebra to become knowledgeable in the inner workings of this method. In this appendix we intend to provide sufficient basics to allow one to read the text.

B.1 The eigenvalue matrix (Λ)

Given 2×2 matrix $\boldsymbol{a}$, the related matrix eigen-equation is

$$\boldsymbol{a}\boldsymbol{e} = \boldsymbol{e}\Lambda. \tag{B.1.1}$$

Pre-multiplying by $\boldsymbol{e}^{-1}$ diagonalizes $\boldsymbol{a}$, resulting in the *eigenvalue matrix*

$$\Lambda = \boldsymbol{E}^{-1}\boldsymbol{a}\boldsymbol{e} \tag{B.1.2}$$

$$= \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}. \tag{B.1.3}$$

Post-multiplying by $\boldsymbol{e}^{-1}$ recovers $\boldsymbol{a}$

$$\boldsymbol{a} = \boldsymbol{e}\Lambda\boldsymbol{E}^{-1} = \begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}. \tag{B.1.4}$$

Matrix product formula:

This last relation is the entire point of the eigenvector analysis, since it shows that any power of $\boldsymbol{a}$ may be computed from powers of the eigenvalues. Specifically,

$$\boldsymbol{a}^n = \boldsymbol{e}\Lambda^n\boldsymbol{E}^{-1}. \tag{B.1.5}$$

For example, $\boldsymbol{a}^2 = \boldsymbol{a}\boldsymbol{a} = \boldsymbol{e}\Lambda\,(\boldsymbol{E}^{-1}\boldsymbol{e})\,\Lambda\boldsymbol{E}^{-1} = \boldsymbol{e}\Lambda^2\boldsymbol{E}^{-1}$.

Equations B.1.1, B.1.3 and B.1.4 are the key to eigenvector analysis, and you need to memorize them. You will use them repeatedly throughout this text.

Derivation of the inverse of a 2×2 matrix

1. Step 1: To derive (*ii*) starting from (*i*), normalize the first column to 1.

$$\begin{bmatrix} 1 & \frac{b}{a} \\ 1 & \frac{d}{c} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{c} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

2. Step 2: Subtract row (1) from row (2): (2) $\leftarrow$ (2)$-$(1)

$$\begin{bmatrix} 1 & \frac{b}{a} \\ 0 & \frac{d}{c} - \frac{b}{a} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{a} & 0 \\ -\frac{1}{a} & \frac{1}{c} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

3. Step 3: Multiply row (2) by ca and express result in terms of the determinate $\Delta = ad - bc$.

$$\begin{bmatrix} 1 & \frac{b}{a} \\ 0 & \Delta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{a} & 0 \\ -c & a \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

4. Step 4: Solve row (2) for x_2: $x_2 = -\frac{c}{\Delta} y_1 + \frac{a}{\Delta} y_2$.

5. Step 5: Solve row (1) for x_1:

$$x_1 = \frac{1}{a} y_1 - \frac{b}{a} x_2 = \left[\frac{1}{a} + \frac{b}{a} \frac{c}{\Delta} \right] y_1 - \frac{b}{a} \frac{a}{\Delta} y_2.$$

.

Rewriting in matrix format, in terms of $\Delta = ad - bc$, gives:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{a} + \frac{bc}{a\Delta} & -\frac{b}{\Delta} \\ -\frac{c}{\Delta} & \frac{a}{\Delta} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} \frac{\Delta + bc}{a} & -b \\ -c & a \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$

since $d = (\Delta + bc)/a$.

Summary: This is a lot of messy algebra, so it is essential that you memorize the final result: (1) swap the diagonal, (2) change the off-diagonal signs, and (3) normalize by the determinant Δ.

Applying the last boundary condition, we evaluate Eq. 5.4.2 to obtain the ABCD matrix at the input ($x = 0$) (Pipes, 1958),

$$\begin{bmatrix} \mathcal{P} \\ \mathcal{V} \end{bmatrix}_0 = \begin{bmatrix} \cosh \kappa L & \mathcal{Z}_r \sinh \kappa L \\ \mathcal{Y}_r \sinh \kappa L & \cosh \kappa L \end{bmatrix} \begin{bmatrix} \mathcal{P} \\ -\mathcal{V} \end{bmatrix}_L. \tag{5.4.5}$$

Note that the uniform horn is reversible (P7) and reciprocal (P6).

Exercise #*11*

Evaluate the expression in terms of the load impedance.

Solution: Since $Z_{\text{load}} = -\mathcal{P}_L / \mathcal{V}_L$, we have

$$\left. \frac{\mathcal{P}}{\mathcal{V}} \right|_0 = \frac{Z_{\text{load}} \cosh \kappa L - \mathcal{Z}_r \sinh \kappa L}{Z_{\text{load}} \mathcal{Y}_r \sinh \kappa L - \cosh \kappa L} \tag{5.4.6}$$

■

Impedance matrix: Equation 5.4.5 is an impedance matrix (algebra required)

$$\begin{bmatrix} \mathcal{P}_0 \\ \mathcal{P}_L \end{bmatrix} = \frac{\mathcal{Z}_r}{\sinh(\kappa L)} \begin{bmatrix} \cosh(\kappa L) & 1 \\ 1 & \cosh(\kappa L) \end{bmatrix} \begin{bmatrix} \mathcal{V}_0 \\ \mathcal{V}_L \end{bmatrix}.$$

Exercise #*12*

Write out the short-circuit ($\mathcal{V}_L = 0$) input impedance $Z_{in}(s)$ for the uniform horn.

Solution:

$$Z_{in}(s) = \left. \frac{\mathcal{P}}{\mathcal{V}} \right|_{\mathcal{V}_L = 0} = \mathcal{Z}_r \frac{\cosh \kappa L}{\sinh \kappa L} = \mathcal{Z}_r \tanh \kappa L |_{\mathcal{V}_L = 0}.$$

■

Input admittance Y_{in}: Given the input admittance of the horn, it is possible to determine whether it is uniform without further analysis. That is, if the horn is uniform and infinite in length, the input admittance at $x = 0$ is

$$Y_{in}(x = 0, s) \equiv \frac{\mathcal{V}(0, \omega)}{\mathcal{P}(0, \omega)} = \mathcal{Y}_r,$$

since $\mathcal{P}_0^+ = 1$ and $\mathcal{P}_L^- = 0$. For an infinite uniform horn, there are no reflections.

When the uniform horn is terminated with a fixed impedance $\mathcal{Z}_r$ at $x = L$, we can substitute pressure and velocity measurements into Eq. 5.4.2 to find $\mathcal{P}_0$ and $\mathcal{P}_0$, and given these, we can calculate the pressure reflectance at $x = L$ (Eq. 3.4.6, p. 116),

$$\Gamma_L(s) = \frac{\mathcal{P}_L}{\mathcal{P}_0} = \frac{\mathcal{P}(L, \omega) - \mathcal{Z}_r \mathcal{V}(L, \omega)}{\mathcal{P}(L, \omega) + \mathcal{Z}_r \mathcal{V}(L, \omega)} = \frac{Z_L - \mathcal{Z}_r}{Z_L + \mathcal{Z}_r}.$$

Given sufficiently accurate measurements of the throat pressure and assuming $\Gamma_L = 0$, the input impedance Z_{in} at the input $\boldsymbol{x} = 0$ is $\mathcal{Z}_r = \rho_o c / A_o$.

5.4.2 Conical horn #2

Using the conical horn area $A(r) \propto r^2$ in Eq. 5.2.10, on page 241 (or Eq. 5.2.11 on page 242) results in the spherical wave equation

$$\mathcal{P}_{rr}(r,\omega) + \frac{2}{r}\mathcal{P}_r(r,\omega) = \kappa^2 \mathcal{P}(r,\omega), \tag{5.4.7}$$

where $\kappa(s) = \pm s/c_o$. The eigensolutions of Eq. 5.4.7 are

$$\mathcal{P}^{\pm}(r,s) = \frac{e^{\mp\kappa r}}{r} \leftrightarrow \frac{1}{r}\delta(t \mp r/c_o).$$

Radiation admittance for the conical horn: The conical horn's acoustic input admittance $Y_{in}(r,s)$ at any location r is found by dividing $\mathcal{V}(r,s)$ by $\mathcal{P}(r,s)$:

$$Y^{\pm}_{in}(r,s) = \frac{\mathcal{V}^{\pm}}{\mathcal{P}^{\pm}} = -\frac{A(r)}{s\rho_o}\frac{d}{dr}\ln \mathcal{P}^{\pm}(r,s) \tag{5.4.8}$$

$$= \mathcal{Y}_r(r)\ \ 1 \pm \frac{c_o}{sr} \ \ \leftrightarrow \frac{A(r)}{\rho_o c_o}\left(\delta(t-r/c_o) \pm \frac{c_o}{r}u(t-r/c_o)\right). \tag{5.4.9}$$

The pressure pulse is delayed by r/c_o due to $e^{-\kappa(s)r}$. As the area of the horn increases, the pressure decreases as $1/r = 1/\ \ \overline{A(r)}$. This results in the uniform backflow $c_o u(t)/r \leftrightarrow c_o/sr$ due to conservation of mass, and the characteristic admittance $\mathcal{Y}_r(r)$ variation with r.

5.4.3 Exponential horn #3

If we define the area as $A(r) = A_o e^{2mr}$, the eigenfunctions of the horn are

$$\mathcal{P}^{\pm}(r,\omega\jmath) = e^{-mr}e^{\mp\jmath\sqrt{\omega^2-\omega_c^2}\ r/c}, \tag{5.4.10}$$

which may be shown by the substitution of $\mathcal{P}^{\pm}_c(r,\omega\jmath)$ into Eq. 5.2.10 (p. 241).

This case is of special interest because the radiation impedance is purely reactive below the horn's cutoff frequency ($\omega < \omega_c = mc_o$), as may be seen from curves 3 and 4 of Fig. 5.3 (p. 247). As a result, no energy can radiate from an open horn for $\omega < \omega_c$ because the eigenvalues

$$\kappa(s) = -m \pm \frac{\jmath}{c_o}\ \ \overline{\omega^2-\omega_c^2} = -m \pm \ \ \overline{s^2-s_c^2}$$

are purely real (this is the case of nonpropagating evanescent waves).

If we use Eq. 4.4.7 (p. 190), the input admittance is

$$Y^{\pm}_{in}(x,s) = -\frac{A(x)}{s\rho_o}\left(m \pm \sqrt{m^2+\kappa^2}\right)x. \tag{5.4.11}$$

Kleiner (2013) gives an equivalent expression for $Y_{in}(x,\omega)$ given area $S(x) = e^{mx}$,

$$Y_{in}(x,\omega) = \frac{S(x)}{\jmath\omega\rho}\ \ \frac{m}{2} + \jmath\frac{\overline{4\omega^2-(mc)^2}}{2c}\ \ ,$$

and impedance

$$Z_{in}(r,\omega) = \frac{\rho c}{S_T}\ \ \jmath\frac{\omega_c}{\omega} + \ \ \overline{1-\left(\frac{\omega_c}{\omega}\right)^2}\ \ ,$$

whereas post-multiplication swaps the columns,

$$AP = \begin{bmatrix} a & b \\ \alpha & \beta \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ \beta & \alpha \end{bmatrix}.$$

For the 3×2 case there are $3 \cdot 2/2 = 3$ such matrices (swap a row with the other 2, then swap the remaining two rows).

6. *Gaussian elimination (GE) operations G_k:* There are three types of elementary row operations, which may be performed without fundamentally altering a system of equations (e.g. the resulting system of equations is *equivalent*). These operations are (1) swap rows (e.g. using a permutation matrix), (2) scale rows, or (3) perform addition/subtraction of two scaled rows. All such operations can be performed using matrices.

   For lack of a better term, we'll describe these as 'Gaussian elimination' or 'GE' matrices.[2] We will categorize any matrix that performs only elementary row operations (but any number of them) as a 'GE' matrix. Therefore, a cascade of GE matrices is also a GE matrix.

   Consider the GE matrix

   $$G = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}.$$

   (a) This pre-multiplication scales and subtracts row (1) from (2) and returns it to row (2).

   $$GA = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} a & b \\ \alpha & \beta \end{bmatrix} = \begin{bmatrix} a & b \\ a-\alpha & b-\beta \end{bmatrix}.$$

   The shorthand for this Gaussian elimination operation is $(1) \leftarrow (1)$ and $(2) \leftarrow (1)-(2)$.

   (b) Post-multiplication adds and scales *columns*.

   $$AG = \begin{bmatrix} a & b \\ \alpha & \beta \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} a-b & b \\ \alpha-\beta & \beta \end{bmatrix}.$$

   Here the second column is subtracted from the first, and placed in the first. The second column is untouched. This operation is *not* a Gaussian elimination. Therefore, to put Gaussian elimination operations in matrix form, we form a cascade of pre-multiply matrices.

   Here $\det(G) = 1$, $G^2 = I$, which won't always be true if we scale by a number greater than 1. For instance, if $G = \begin{bmatrix} 1 & 0 \\ m & 1 \end{bmatrix}$ (scale and add), then we have $\det(G) = 1$, $G^n = \begin{bmatrix} 1 & 0 \\ n \cdot m & 1 \end{bmatrix}$.

[2] The term 'elementary matrix' may also be used to refer to a matrix that performs an elementary row operation. Typically, each elementary matrix differs from the identity matrix by a single row operation. A cascade of elementary matrices could be used to perform Gaussian elimination.

row $[x_1, x_2, x_3, \ldots]$ or single column $[x_1, x_2, x_3 \ldots]^T$ (where T indicates the transpose). In this text we will typically use column vectors. The vector may also be written out using unit vector notation to indicate direction. For example: $\mathbf{x}_{3\times1} = x_1\hat{\mathbf{x}} + x_2\hat{\mathbf{y}} + x_3\hat{\mathbf{z}} = [x_1, x_2, x_3]^T$, where $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ are unit vectors in the x, y, z Cartesian directions (here the vector's subscript 3×1 indicates its dimensions). The type of notation used may depend on the engineering problem you are solving.

3. *Matrix:* $A = \left[\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \ldots, \mathbf{a}_M\right]_{N\times M} = \{a_{n,m}\}_{N\times M}$ can be a non-square matrix if the number of elements in each of the vectors (N) is not equal to the number of vectors (M). When $M = N$, the matrix is square. It may be inverted if its determinant $|A| = \quad \lambda_k \neq 0$ (where λ_k are the eigenvalues). In this text we work only with 2×2 and 3×3 square matrices.

4. *Linear system of equations*: $A\mathbf{x} = \mathbf{b}$ where $\mathbf{x}$ and $\mathbf{b}$ are vectors and matrix A is a square.

   (a) *Inverse:* The solution of this system of equations may be found by finding the inverse $\mathbf{x} = A^{-1}\mathbf{b}$.

   (b) *Equivalence:* If two systems of equations $A_0\mathbf{x} = \mathbf{b}_0$ and $A_1\mathbf{x} = \mathbf{b}_1$ have the same solution (i.e., $\mathbf{x} = A_0^{-1}\mathbf{b}_0 = A_1^{-1}\mathbf{b}_1$), they are said to be equivalent.

   (c) *Augmented matrix:* The first type of augmented matrix is defined by combining the matrix with the right-hand side. For example, given the linear system of equations of the form $A\mathbf{x} = \mathbf{y}$

   $$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$

   the augmented matrix is

   $$[A|y] = \left[\begin{array}{cc|c} a & b & y_1 \\ c & d & y_2 \end{array}\right].$$

   A second type of augmented matrix may be used for finding the inverse of a matrix (rather than solving a specific instance of linear equations $Ax = b$). In this case the augmented matrix is

   $$[A|I] = \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right].$$

   Performing Gaussian elimination on this matrix, until the left side becomes the identity matrix, yields A^{-1}. This is because multiplying both sides by A^{-1} gives $A^{-1}A|A^{-1}I = I|A^{-1}$.

5. *Permutation matrix* (P): A matrix that is equivalent to the identity matrix, but with scrambled rows (or columns). Such a matrix has the properties $\det(P) = \pm 1$ and $P^2 = I$. For the 2×2 case, there is only one permutation matrix:

   $$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad P^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

   A permutation matrix P swaps rows or columns of the matrix it operates on. For example, in the 2×2 case, pre-multiplication swaps the rows,

   $$PA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ \alpha & \beta \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ a & b \end{bmatrix},$$

where $\omega_c(r)$ is the cutoff frequency. Given this exact solution to the exponential horn, we could use a series expansion of the form

$$A(r) = \sum_k a_k e^{b_k x} \quad \in \mathbb{R} \tag{5.4.12}$$

to obtain the general solution, for an arbitrary analytic $A(r)$.

5.5 Solution methods

To model the wave equation, two distinct mathematical techniques are described. The first of these is called separation of variables. This method is limited to a small and restrictive number of separable coordinate systems (SCS). Once the SCS is chosen, the eigenfunctions are known as solutions to that specific Sturm-Liouville equation, which are always scalar (ordinary) differential equation (ODEs).

The second method uses the transmission matrix (a lumped-parameter method). This solution method assumes a limit on the upper frequency response. That is, quasistatics assumes the wavelength is larger than the size of the object being modeled. Thus other than the upper frequency limit, there are no limitations in the analytical and numerical solutions using the transmission matrix method. Furthermore, it may be computed in both the frequency and time domains (Fettweis, 1986).

1a. **Separable coordinate systems:** Classically PDEs are solved by *separation of variables*. This method is limited to a few ortho-normal coordinate systems, such as rectangular, cylindrical, and spherical coordinates (Morse, 1948, p 296–97). Even a slight deviation from separable specific coordinate systems represents a major barrier toward further analysis and understanding, blocking insight into more general cases. Separable coordinate systems have a high degree of symmetry. Note that the solution of the wave equation is not tied to a specific coordinate system.

1b. **Sturm-Liouville methods and eigenvectors:** When the coordinate system is separable, the resulting PDEs are always reduced to a system of Sturm-Liouville equations. The solutions of this important class of Sturm-Liouville eigenfunctions are all tabulated.

Webster horn theory (Webster, 1919; Morse, 1948; Pierce, 1981) is a generalized Sturm-Liouville equation that adds its physics in the form of the horn's area function. The Webster equation sidesteps the seriously limiting problem of separation of variables by using the alternative quasistatic solution, which ignores the non propagating high-frequency evanescent modes. This is essentially a one-dimensional low-pass approximation to the wave equation.

While mathematics provides rigor, physics provides understanding. The physical applications make the theory useful.

2. **Lumped-element method:** As described (see page 144) a system may be represented in terms of lumped elements, as either electrical inductors, capacitors, and resistors or their mechanical counterparts, masses, springs, and dashpots. Such systems are represented by 2×2 transmission matrices in the s (i.e., Laplace) domain (Ramo et al., 1965, Appendix IV).

When a system of lumped-element networks contains only resistors and capacitors or resistors and inductors, the solution is a diffusion equation, which does not support propagated waves. Depending on the elements in the system of equations, there can be an overlap between a diffusion process and scalar waves, represented as transmission lines, both modeled as lumped-element networks of 2×2 matrices (Eq. 3.8.1, p. 144) (Campbell, 1922; Brillouin, 1953; Ramo et al., 1965).

Nyquist sampling and quasistatics: Quasistatic methods provide band-limited solutions below a critical frequency f_c for a much wider class of geometries, by avoiding high-frequency cross-modes. The model of a train is depicted in Fig. 5.4, which is the same as Fig. 4.8 (p. 221).

Example: Train-mission-line problem. The mechanical mass-spring system of Fig. 5.4 is the electrical equivalent circuit. The mass is modeled as an inductor and the springs as capacitors to ground. The velocity is analogous to a current and the force $f_n(t)$ to the voltage $\phi_n(t)$. The length of each cell is Δ [m].

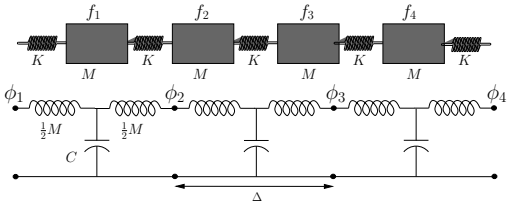


Figure 5.4: *Depiction of a train consisting of cars, treated as a mass M and linkages, treated as springs of stiffness K or compliance $C = 1/K$. The equivalent electrical circuit is shown below the mass-spring system, with the masses modeled by inductors (M) and springs modeled as capacitors (C). For this model to accurately represent a transmission line the frequency must be less than the equivalent Nyquist frequency f_c. The delay of one cell is $T_o = \Delta/c_o$. One can measure either T_o or Δ and $c_o = 1/\sqrt{MC}$. Since $\lambda_c = 2\Delta$, it follows that the cutoff frequency is $f_c = c_o/2\Delta = 1/2T_o$. For $f < f_c$, the frequency response is independent of frequency, thus acting as a pure delay. As the frequency is increased above f_c, the wavelength becomes shorter that the critical wavelength $\lambda_c = 2\Delta$, and the delay line becomes a lowpass filter, strongly departing from transmission line properties. For more detailed information see the Figure caption on p. 221.*

When the wavelength $\lambda = c_o/f_c$ is greater than twice the physical distance Δ between the elements

$$\lambda > \lambda_c = 2\Delta \quad \text{[m]},$$

the approximation is mathematically equivalent to a transmission line. As described in DE-3, problem #2, the velocity is $c_o = 1/\sqrt{MC}$ [m/s]. As the frequency increases, the wavelength becomes shorter. When the frequency is equal to the critical frequency f_c the critical wavelength $\lambda_c = c_o/f_c = 2\Delta$. Above the critical frequency the quasistatic (lumped-element) model breaks down and transitions from a delay line to a lowpass filter, as discussed in DE-3, problem #2.

The frequency is under the control of the modeling process, since more elements may be added to allow for higher frequencies (shorter wavelengths). If the nature of the solution at high frequencies ($f > f_c$) is desired, we may add more sections (make Δ smaller). For

Complex power: In this special case, the *complex power* $\mathcal{P}(s) \in \mathbb{R}(s)$ is defined, in the complex frequency domain (s), as

$$\mathcal{P}(s) = \boldsymbol{I}^\dagger(s)\boldsymbol{V}(s) = \boldsymbol{I}^\dagger(s)\boldsymbol{Z}(s)\boldsymbol{I}(s) \leftrightarrow p(t) \quad \text{[W]}.$$

The real part of the complex power must be positive. The imaginary part corresponds to available stored energy.

The case of three-dimensions is special, allowing definitions that are not easily defined in more than three dimensions. A vector in $\mathbb{R}^3$ labels the point having the coordinates of that vector.

A.3.4 Matrices

When working with matrices, the role of the weights and vectors can change, depending on the context. A useful way to view a matrix is as a set of column vectors, weighted by the elements of the column-vector of weights multiplied from the right. For example,

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1M} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2M} \\ & & \ddots & & \\ a_{N1} & a_{N2} & a_{N3} & \cdots & a_{NM} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_M \end{bmatrix} = w_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{N1} \end{bmatrix} + w_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{N2} \end{bmatrix} + \cdots + w_M \begin{bmatrix} a_{1M} \\ a_{2M} \\ a_{3M} \\ \vdots \\ a_{NM} \end{bmatrix},$$

where the weights are $\begin{bmatrix} w_1, w_2, \ldots, w_M \end{bmatrix}^T$. Alternatively, the matrix is a set of row vectors of weights, each of which is applied to the column vector on the right ($[w_1, w_2, \ldots, W_M]^T$).

The determinant of a matrix is denoted as either $\det \boldsymbol{A}$ or simply $|\boldsymbol{A}|$ (as in the absolute value). The inverse of a square matrix is $\boldsymbol{A}^{-1}$ or inv$\boldsymbol{A}$. If $|A| = 0$, the inverse does not exist. $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A}$.

Matlab/Octave's notional convention for a row-vector is $[a, b, c]$ and a column-vector is $[a; b; c]$. A prime on a vector takes the complex conjugate transpose. To suppress the conjugation, place a period before the prime. The : argument converts the array into a column vector, without conjugation. A tacit notation in Matlab is that *vectors* are columns and the index to a vector is a row vector. Matlab defines the notation $1:4$ as the "row-vector" $[1, 2, 3, 4]$, which is unfortunate as it leads users to assume that the default vector is a row. This can lead to serious confusion later, as Matlab's default vector is a column. I have not found the above convention explicitly stated, and it took me years to figure this out for myself.

When writing a complex number we shall adopt $1\jmath$ to indicate $\sqrt{-1}$. Matlab/Octave allows either $1\imath$ or $1\jmath$.

Units are SI; angles are in degrees [deg] unless otherwise noted. The units for π are always radians [rad]. For example $\sin(\pi), e^{\jmath 90^\circ} = e^{\jmath\pi/2}$.

A.3.5 2×2 complex matrices

Does this discussio long in Appendix A?

Here are some definitions to learn:

1. *Scalar*: A number – for example $\{a, b, c, \alpha, \beta, \ldots\} \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{I}, \mathbb{R}, \mathbb{C}\}$

2. *Vector*: A quantity having direction as well as magnitude, often denoted by a bold letter $\mathbf{x}$, or with an arrow over the top $\vec{x}$. In matrix notation, this is typically represented as a single

The triple product: This is defined between three vectors as

$$\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{C})=\det\begin{vmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}.$$

This may be indicated without the use of parentheses, since there can be no other meaningful interpretation. However for clarity, parentheses should be used. The triple product is the volume of the parallelepiped (3D-crystal shape) outlined by the three vectors, as shown in Fig. 3.4 p. 119.

Dialects of vector notation: Physical fields are, by definition, functions of space $\boldsymbol{x}$ [m], and in the most general case, time t[s]. When Laplace transformed, the fields become functions of space and complex frequency (e.g., $\boldsymbol{e}(\boldsymbol{x},t)\leftrightarrow\boldsymbol{e}(\boldsymbol{x},s)$). As before, there are several equivalent vector notations. For example, $\boldsymbol{e}(\boldsymbol{x},t)=\begin{bmatrix}E_x, & E_y, & E_z\end{bmatrix}^T=E_x(\boldsymbol{x},t)\hat{\mathbf{x}}+E_y(\boldsymbol{x},t)\hat{\mathbf{y}}+E_z(\boldsymbol{x},t)\hat{\mathbf{z}}$ is "in-line," to save space. The same equation may written in "displayed" notation as:

$$\boldsymbol{e}(\boldsymbol{x},t)=\begin{bmatrix}E_x(\boldsymbol{x},t)\\ E_y(\boldsymbol{x},t)\\ E_z(\boldsymbol{x},t)\end{bmatrix}=\begin{bmatrix}E_x\\ E_y\\ E_z\end{bmatrix}(\boldsymbol{x},t)\quad\begin{bmatrix}E_x, & E_y, & E_z\end{bmatrix}^T\equiv E_x\hat{\mathbf{x}}+E_y\hat{\mathbf{y}}+E_z\hat{\mathbf{z}}.$$

Note the three notations for vectors, bold font, element-wise columns, element-wise transposed rows and dyadic format. These are all shorthand notations for expressing the vector. Such usage is similar to a dialect in a language.

Complex elements: When the elements are complex ($\in\mathbb{C}$), the transpose is defined as the complex conjugate of the elements. In such complex cases the transpose conjugate may be denoted with a † rather than T

$$\begin{bmatrix}-2\jmath\\ 3\jmath\\ 1\end{bmatrix}^\dagger=\begin{bmatrix}2\jmath & -3\jmath & 1\end{bmatrix}\in\mathbb{C}.$$

For this case when the elements are complex, the dot product is a real number

$$\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{a}^\dagger\boldsymbol{b}=\begin{bmatrix}a_1^* & a_2^* & a_3^*\end{bmatrix}\begin{bmatrix}b_1\\ b_2\\ b_3\end{bmatrix}=a_1^*b_1+a_2^*b_2+a_3^*b_3\in\mathbb{R}.$$

Norm of a complex vector: The dot product of a vector with itself is called the *norm* of $\boldsymbol{a}$

$$\|\boldsymbol{a}\|=\sqrt{\boldsymbol{a}^\dagger\boldsymbol{a}}\geq 0.$$

which is always non-negative.

Such a construction is useful when $\boldsymbol{a}$ and $\boldsymbol{b}$ are related by an impedance matrix

$$\boldsymbol{V}(s)=\boldsymbol{Z}(s)\boldsymbol{I}(s)$$

and we wish to compute the power. For example, the impedance of a mass is ms and a capacitor is $1/sC$. When given a system of equations (a mechanical or electrical circuit) one may define an impedance matrix.

many (perhaps most) problems, lumped elements are easy to use and accurate, as long as we don't violate the Nyquist condition (Nyquist, 1924; Brillouin, 1953; Ramo et al., 1965).

5.5.1 Eigenfunctions $\varrho^{\pm}(r,t)$ of the WHEN

Because the wave equation (Eq. 5.2.7) is second-order in time, there are two causal independent eigenfunction solutions: an outbound (right–traveling) $\varrho^+(r,t)$ wave, and an inbound (left-traveling) $\varrho^-(r,t)$ wave.

Every eigenfunction depends on an area function $A(r)$ (Eq. 5.2.10, p. 241). In theory then, given an eigenfunction, it should be possible to find the area $A(r)$. This is known as the *inverse problem*, which is generally believed to be a difficult problem. Specifically, given the eigenvalues λ_k, how does one determine the corresponding area function $A(r)$?

Because the characteristic impedance $\mathcal{Y}_r(r)$ of the wave in the horn changes with location, there are local reflections due to these area variations. Thus there are fundamental relationships between the area change $dA(r)/dr$, the horn's eigenfunctions $\mathcal{P}^{\pm}(r,s)$, the eigenmodes, and the input impedance.

Complex vs. real frequency: We shall continue to maintain the distinction that functions of ω are Fourier transforms and causal functions of Laplace frequency s correspond to Laplace transforms, which are necessarily complex–analytic in s in the right half-plane (RHP) region of convergence (RoC). This distinction is critical, since we typically describe impedance $Z(s)$ and admittance $Y(s)$ as complex–analytic functions in s in terms of their poles and zeros. The eigenfunctions $\mathcal{P}^{\pm}(r,s)$ of Eq. 5.2.10 are also causal complex–analytic functions of s.

Plane-wave eigenfunction solutions: In 1690, nine years before Newton's publication of *Principia*, Christiaan Huygens was the first to gain insight into wave propagation, today known as *Huygens's principle*. While his concept showed a deep insight, we now know it was flawed, as it ignored the backward-traveling wave (Miller, 1991). In 1747 d'Alembert published the first correct solution for the plane-wave scalar wave equation,

$$\varrho(x,t)=f(t-x/c_o)+g(t+x/c_o),\tag{5.5.1}$$

where $f(\cdot)$ and $g(\cdot)$ are general functions of their argument. That this is the solution may be shown by use of the chain rule, by taking partials with respect to x and t.

In terms of physics, d'Alembert's general solution describes two arbitrary waveforms $f(\cdot)$ and $g(\cdot)$ traveling at a speed c_o, one forward and one reversed. This solution is quite easily visualized.

Exercise #*13*

By the use of the chain rule, prove that d'Alembert's formula satisfies the one-dimensional wave equation.

Solution: Taking a derivative with respect to t and r gives

$$\partial_t\varrho(r,t)=-c_of'(r-c_ot)+c_og'(r+c_ot)$$
$$\partial_r\varrho(r,t)=f'(r-c_ot)+g'(r+c_ot),$$

and a second derivative gives

$$\partial_{tt}\varrho(r,t) = c_o^2 f''(r - c_o t) + c_o^2 g''(r + c_o t)$$
$$\partial_{rr}\varrho(r,t) = f''(r - c_o t) + g''(r + c_o t).$$

From these last two equations we have the one-dimensional wave equation

$$\partial_{rr}\varrho(r,t) = \frac{1}{c_o^2}\partial_{tt}\varrho(r,t),$$

which has solutions in Eq. 5.5.1. ■

Exercise #*14*

Assuming $f(\cdot)$ and $g(\cdot)$ are $\delta(\cdot)$, find the Laplace transform of the solution corresponding to the uniform horn $A(x) = 1$.

Solution: Using Table 3.9 (p. 160) of Laplace transforms on Eq. 5.5.1 gives

$$\varrho(x,t) = \delta(t - x/c_o) + \delta(t + x/c_o) \leftrightarrow e^{-sx/c_o} + e^{sx/c_o}. \tag{5.5.2}$$

Note that the delay $T_o = \pm x/c_o$ depends on the range x. ■

Three-dimensional d'Alembert spherical eigenfunctions: We can generalize the d'Alembert solution to spherical waves by changing the area function of Eq. 5.2.10 to $A(r) = A_o r^2$ (see Eq. 5.1.9, p. 229 and Table 5.2, p. 247). The radial wave equation then becomes

$$\nabla_r^2 \varrho(r,t) = \frac{1}{c_o^2}\frac{\partial^2}{\partial t^2}\varrho(r,t). \tag{5.5.3}$$

where

$$\nabla_r^2\, p(r,t) \equiv \nabla \cdot \nabla\, p(r,t). \tag{5.5.4}$$

Multiplying by r (see the discussion of the conical horn in Sec. 5.4.7, p. 254) provides the d'Alembert wave equation solution for arbitrary forward $f(\cdot)$ and retrograde $g(\cdot)$ waveforms

$$r\varrho(r,t) = f(t - r/c_o) + g(t + r/c_o). \tag{5.5.5}$$

Finally dividing by r results in the well known spherical (3D) d'Alembert wave equation eigenfunctions (Eq. 5.5.3)

$$\varrho(r,t) = \frac{f(t - r/c_o)}{r} + \frac{g(t + r/c_o)}{r}. \tag{5.5.6}$$

This dependence on r is required for the conservation of energy for spherical waves (Allen and Berkley, 1979).

5.6 Integral definitions of $\nabla()$, $\nabla\cdot()$, $\nabla\times()$, and $\nabla\wedge()$

In Sec. 5.2, page 236, we described to two forms of wave equations, scalar and vector. Up to now we have only discussed the scalar case. The vector wave equation describes the evolution of a vector field, such as Maxwell's electric field vector $\boldsymbol{E}(\boldsymbol{x}, t)$.

There are two equivalent definitions for each of the four operators: differential and integral. The integral form provides a more intuitive view of the operator, which in the limit converges to the differential form. Following a discussion of the gradient, divergence, and curl integral operators, we discuss these two forms.

In addition there are three fundamental vector theorems: Gauss's law (divergence theorem), Stokes's law (curl theorem), and Helmholtz's decomposition theorem. Without the use of these fundamental vector calculus theorems, we could not understand Maxwell's equations.

Polar wedge product: The vector wedge product in polar coordinates is (Fig. 3.4, p. 119)

$$\boldsymbol{B} \wedge \boldsymbol{C} = \jmath \|\boldsymbol{B}\|\, \|\boldsymbol{C}\| \sin\theta \in \mathbb{R},$$

where $\sin\theta \in \mathbb{R}$ is therefore the *direction-sine* between $\boldsymbol{B}$ and $\boldsymbol{C}$.

Complex polar vector product: From these two polar definitions and $e^{\jmath\theta} = \cos\theta + \jmath\sin\theta$,

$$\boldsymbol{B}\cdot\boldsymbol{C} + \jmath\boldsymbol{B}\wedge\boldsymbol{C} = ||\boldsymbol{B}||\cdot||\boldsymbol{C}||\, e^{st}.$$

Hence

$$|\boldsymbol{B}\cdot\boldsymbol{C}|^2 + |\boldsymbol{B}\wedge\boldsymbol{C}|^2 = \|\boldsymbol{B}\|^2\,\|\boldsymbol{C}\|^2\,|\cos^2\theta| + \|\boldsymbol{B}\|^2\,\|\boldsymbol{C}\|^2\,|\sin^2\theta| = ||\boldsymbol{B}||^2\,||\boldsymbol{C}||^2.$$

As described in Fig. 3.4, this relationship holds true in any vector space, of any number of dimensions, containing vectors $\boldsymbol{B}$ and $\boldsymbol{C}$.

In this case $s = \sigma + \omega\jmath \in \mathbb{C}$ can be the Laplace frequency. Jaynes (1991) has a brilliant insight into this type of vector product. reread!

A.3.3 Norms of vectors

The norm of a vector is the scalar product of the vector with itself

$$\|\boldsymbol{A}\| = \sqrt{\boldsymbol{A}\cdot\boldsymbol{A}} \geq 0,$$

forming the Euclidean length of the vector.

Euclidean distance between two points in $\mathbb{R}^3$: The scalar product of the difference between two vectors $(\boldsymbol{A} - \boldsymbol{B})\cdot(\boldsymbol{A} - \boldsymbol{B})$ is the Euclidean distance between the points they define

$$\|\boldsymbol{A} - \boldsymbol{B}\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}.$$

Triangle inequality

$$\|\boldsymbol{A} + \boldsymbol{B}\| = \sqrt{(a_1 + b_1)^2 + (a_2 + b_2)^2 + (a_3 + b_3)^2} \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|.$$

In terms of a right triangle this says the the sum of the lengths of the two sides is greater to the length of the hypotenuse, and equal when the triangle degenerates into a line.

Vector cross product: The *vector product* (aka cross product) $\boldsymbol{A}\times\boldsymbol{B} = \|\boldsymbol{A}\|\,\|\boldsymbol{B}\|\sin\theta$ is defined between the two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$. In Cartesian coordinates

$$\boldsymbol{A}\times\boldsymbol{B} = \det\begin{vmatrix}\hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}}\\ a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\end{vmatrix}.$$

As with monics, the lead coefficient must always be 1. The complex variable $s \in \mathbb{C}$ is the *Laplace frequency*.

The ratio of the output $Y(s)$ over the input $X(s)$ is called the system *transfer function* $H(s)$. The coefficient $\alpha \in \mathbb{R}$ is called the *gain*. The roots of the numerator are called the *zeros* and those of the denominator, the *poles*. When $H(s)$ is the ratio of two degree 1 monics, the transfer function is said to be *bilinear*, since it is linear in both the input and output. In such cases $H(s)$ has only one pole and one zero. The inverse Laplace transform of the transfer function is called the system *impulse response*, which describes the system's output signal $y(t)$ for any given input signal $x(t)$, via convolution (i.e., $y(t) = h(t) \star x(t)$).

A.3 Matrix algebra: Systems

A.3.1 Vectors

Vectors as columns of ordered sets of scalars $\in \mathbb{C}$. When we write them out in text, we typically use row notation, with the *transpose* symbol:

$$[a, b, c]^T = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

This is strictly to save space on the page. The notation for *conjugate transpose is* $\dagger$, for example

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix}^\dagger = \begin{bmatrix} a^* & b^* & c^* \end{bmatrix}.$$

The above example is said to be a *3-dimensional* vector because it has three components.

A.3.2 Vector products

A *scalar product* (aka dot product) is defined to "weight" vector elements before summing them, resulting in a scalar. The transpose of a vector (a *row-vector*) is typically used as a *scale factor* (i.e., weights) on the elements of a vector. For example,

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 + 2 \cdot 2 - 3 = 2.$$

A more interesting example is

$$\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ s \\ s^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}^T \begin{bmatrix} 1 \\ s \\ s^2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ s \\ s^2 \end{bmatrix} = 1 + 2s + 4s^2.$$

Polar scalar product: The vector-scalar product in polar coordinates is (Fig. 3.4, p. 119)

$$\boldsymbol{B} \cdot \boldsymbol{C} = \|\boldsymbol{B}\| \, \|\boldsymbol{C}\| \cos\theta \in \mathbb{R},$$

where $\cos\theta \in \mathbb{R}$ is called the *direction-cosine* between $\boldsymbol{B}$ and $\boldsymbol{C}$.

5.6.1 Maxwell's equations

Maxwell's two curl electric equations, for the electrical and magnetic field strengths $\boldsymbol{e}, \boldsymbol{h}$ are the most fundamental

$$\nabla \times \boldsymbol{e} = -\partial_t \boldsymbol{b}. \tag{5.6.1}$$

and magnetic field strengths

$$\nabla \times \boldsymbol{h} = \boldsymbol{j} + \partial_t \boldsymbol{d}. \tag{5.6.2}$$

An alternate popular notation for $\partial_t \boldsymbol{b}$ is $\dot{\boldsymbol{b}}$, but the dot over $\boldsymbol{b}$ is easy to visually miss.

The units on $\boldsymbol{e}$ is in [*Volts/m*] while $\boldsymbol{h}$ has units of [Amps/m^2]. The *conduction current* $\boldsymbol{j}$ obeys Ohm's Law ($\boldsymbol{j} = \sigma \boldsymbol{e}$), where σ is the conductivity in $\mho/m^2$ and $\boldsymbol{d}$ is know as the *displacement current*, first identified and then named by Maxwell, as he first conceived these these four equations.

The last two equations are know as *constitutive relations*, define as $\boldsymbol{d} = \epsilon_o \boldsymbol{e}$ and $\boldsymbol{b} = \mu_o \boldsymbol{h}$.

$$\nabla \cdot \boldsymbol{d} = \rho \tag{5.6.3}$$

and

$$\nabla \times \boldsymbol{b} = 0. \tag{5.6.4}$$

Two additional important constants are the *characteristic impedance* of free space $z_o = \sqrt{\epsilon_o/\mu_o} = 377\ \Omega$, and the speed of light $c_o = 1/\sqrt{\epsilon_o \mu_o} \approx 3{\times}10^8$ [m/s]. These are know as the most significant equations in science, only rivaled by Einstein's famous energy-mass equivalence relation $E = mc_o^2$, which is easily derived from Maxwell first two equations in Eq. 3.5.10, p. 122.

5.6.2 Electric fields $\boldsymbol{e}, \boldsymbol{d}$

5.6.3 Gradient: $\boldsymbol{e} = -\nabla\phi(\boldsymbol{x})$

As shown in Fig. 5.1 (p. 226), the gradient maps $\mathbb{R}^1 \underset{\nabla}{\rightarrow} \mathbb{R}^3$. The gradient is defined as the unit-normal $\hat{\mathbf{n}}$ weighted by the potential $\phi(\boldsymbol{x})$ averaged over a closed surface $\mathcal{S}$,

$$\nabla\phi(\boldsymbol{x}) \equiv \lim_{\mathcal{S},\mathcal{V}\to 0} \left\{ \frac{\int_{\mathcal{S}} \phi(\boldsymbol{x})\, \hat{\mathbf{n}}\, d\mathcal{S}}{\mathcal{V}} \right\} \quad [\text{V/m}], \tag{5.6.5}$$

having area $\mathcal{S}$ and volume $\mathcal{V}$ and centered at $\boldsymbol{x}$ (Greenberg, 1988, p 773).[10] Here $\hat{\mathbf{n}}$ is a dimensionless unit vector perpendicular to the surface $\mathcal{S}$:

$$\hat{\mathbf{n}} = \frac{\nabla\phi}{\|\nabla\phi\|}. \tag{5.6.6}$$

The dimensions of Eq. 5.6.5 are in the units of the potential times the area, divided by the volume, as needed for a gradient (e.g., [V/m]). The units depend on the potential. If ϕ were temperature, the units would be [deg/m].

Exercise #*15*
Justify the units of Eq. 5.6.5.

Solution: The units depend on ϕ per unit length. If ϕ is voltage, then the gradient has units of [V/m]. Under the limit, $d|\mathcal{S}|/||\mathcal{S}||$ must have units of m^{-1}.■

[10] For further discussions, see Greenberg (1988, pp. 778, 791, 809).

The natural way to define the surface and volume is to place the surface on the isopotential surfaces, forming either a cube or a pill-box-shaped volume. As the volume $||\mathcal{S}||$ goes to zero, so must the area $|\mathcal{S}|$. One must avoid irregular volumes where the area is finite as the volume goes to zero (Greenberg, 1988, footnote p 762).

A well-known example is the potential

$$\phi(x,y,z) = \frac{Q}{\epsilon_o \sqrt{x^2+y^2+z^2}} = \frac{Q}{\epsilon_o R} \quad \text{[V]}$$

around a point charge Q [SI units of coulombs]. The constant ϵ_o is the permittivity [F/m^2]. A second well-known example is the acoustic pressure potential around an oscillating sphere, which has the same form (see Table 5.2, p. 247).

How does this work? To better understand what Eq. 5.6.5 means, consider a three-dimensional Taylor series expansion (See Eq. 5.2.4, p. 238) of the potential in $\boldsymbol{x}$ about the limit point $\boldsymbol{x}_o$:

$$\phi(\boldsymbol{x}) \approx \phi(\boldsymbol{x}_o) + \nabla\phi(\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{x}_o) + \text{HOT}.$$

We could define the gradient using this relationship as

$$\nabla\phi(\boldsymbol{x}_o) = \lim_{\boldsymbol{x}\to\boldsymbol{x}_o} \frac{\phi(\boldsymbol{x}) - \phi(\boldsymbol{x}_o)}{\boldsymbol{x} - \boldsymbol{x}_o}.$$

For this definition to apply, $\boldsymbol{x}$ must approach $\boldsymbol{x}_o$ along $\hat{\mathbf{n}}$. To compute the higher-order terms (HOT), we need the Hessian matrix.[11]

The natural way to define a surface $|\mathcal{S}|$ is to find the isopotential contours. The gradient is in the direction of maximum change in the potential, thus perpendicular to the isopotential surface. The key to the integral definition is in taking the limit. As the volume $||\mathcal{S}||$ shrinks to zero, the HOT are small and the integral reduces to the first-order term in the Taylor expansion, since the constant term integrates to zero. Such a construction was used in the proof of the Webster horn equation (Appendix G, p. 331; Fig. G.1, p. 332).

One major problem with Eq. 5.6.5 is that this definition is self-referencing, since $\hat{\mathbf{n}}$ is based on the gradient. Thus the integral definition of the gradient is based on the gradient. Equation 5.6.5 is similar to the mean value theorem for the gradient.

5.6.4 Divergence: $\nabla\cdot\boldsymbol{d} = \rho$ [C/m^3]

The definition of the divergence at $\boldsymbol{x} = [x,y,z]^T$ is (see Eq. 5.1.3)

$$\nabla\cdot\boldsymbol{d}(\boldsymbol{x},t) \equiv [\partial_x,\, \partial_y,\, \partial_z]\cdot\boldsymbol{d}(\boldsymbol{x},t) = \left[\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}\right](\boldsymbol{x},t) = \rho(\boldsymbol{x},t),$$

which maps $\mathbb{R}^3 \underset{\nabla\cdot}{\to} \mathbb{R}^1$.

5.6.5 Divergence and Gauss's law

Like the gradient, the divergence of a vector field may be defined as the surface integral of a compressible vector field in the limit as the volume enclosed by the surface goes to zero. As in

[11] $H_{i,j} = \partial^2(\boldsymbol{x})\phi/\partial x_i \partial x_j$, which exists if the potential is analytic in $\boldsymbol{x}$ at $\boldsymbol{x}_o$.

A.1.6 Periodic functions

Fourier series tells us that periodic functions are discrete in frequency, with frequencies given by nT_s, where T_s is the sample period. The discrete $\mathcal{FT}$ (DFT) is a good example. When using the DFT, the sample period is $T_s = 1/2F_{max}$ and the minimum and maximum frequencies are given by $F_{min} = F_{max}/\text{NFT}$ where NFT is the size of the DFT.

This concept is captured by the *Fourier series*, which is a frequency expansion of a periodic function. This concept is quite general. Periodic in frequency implies discrete in time. Periodic and discrete in time requires periodic and discrete in frequency (the case of the DFT). The modulo function $x = \mod(x,y)$ is periodic with period y ($x, y \in \mathbb{R}$).

A periodic function may be conveniently indicated using double-parentheses notation. This is sometimes known as modular arithmetic. For example,

$$f((t))_T = f(t) = f(t \pm kT)$$

is periodic on $t, T \in \mathbb{R}$ with a period of T and $k \in \mathbb{Z}$. This notation is useful when dealing with Fourier series of periodic functions such as $\sin(\theta)$ where $\sin(\theta) = \sin((\theta))_{2\pi} = \mod(\sin(\theta), 2\pi)$.

When a discrete valued (e.g., time $t \in \mathbb{N}$) sequence is periodic with period $N \in \mathbb{Z}$, we use square brackets

$$f[[n]]_N = f[n] = f[n \pm kN],$$

with $k \in \mathbb{Z}$. This notation will be used with discrete-time signals that are periodic, such as the case of the DFT.

It is common for fractions to repeat. For example 1/7=0.((142857)) where the double brackets indicates this number repeats. That is 1/7=0.142857,142857,142857,142857, ...

A.2 Differential equations vs. polynomials

A polynomial has *degree* N defined by the largest power. A quadratic equation is degree 2, and a cubic has degree 3. We shall indicate a polynomial by the notation

$$P_N(z) = z^N + a_{N-1}z^{N-1}\cdots a_0.$$

It is a good practice to normalize the polynomial so that $a_N = 1$. This will not change the roots, defined by Eq. 3.1.7 (p. 77). The coefficient on z^{N-1} is always the sum of the roots z_n ($a_{N-1} = \sum_n^N z_n$), and a_0 is always their product ($a_0 = \prod_n^N z_n$). Polynomials with $a_N = 1$ are denoted *monic polynomials*. Because that is a mouthful, it is helpful to call $M_N(s) = P_N(s)/a_N$ a monic.

Differential equations have *order* (polynomials have degree). If a second-order differential equation is Laplace transformed (Lec. 3.10, p. 155), one is left with a degree 2 polynomial. For example,

$$\frac{d^2}{dt^2}y(t) + b\frac{d}{dt}y(t) + cy(t) = \alpha\left(\frac{d}{dt}x(t) + \beta x(t)\right) \leftrightarrow \tag{A.2.1}$$

$$(s^2 + bs + c)Y(s) = \alpha(s+\beta)X(s). \tag{A.2.2}$$

$$\frac{Y(s)}{X(s)} = \alpha\frac{s+\beta}{s^2+bs+c} \equiv H(s) \leftrightarrow h(t). \tag{A.2.3}$$

which may be read as "the set '$\{\cdots\}$' of all p/q such that ':' $q \neq 0$, 'and' $p, q \subset \mathbb{Z}$. The translation of the symbols is in single ('$\cdots$') quotes.

Irrational numbers $\mathbb{I}$ are very special: They are formed by taking a limit of fractionals, as the numerator and denominator $\to \infty$, and approach a limit point. It follows that irrational numbers must be approximated by fractionals.

The reals ($\mathbb{R}$) include complex numbers ($\mathbb{C}$) having a zero imaginary part (i.e., $\mathbb{R} \subset \mathbb{C}$).

The *size* of a set is denoted by taking the absolute value (e.g., $|\mathbb{N}|$). Normally in mathematics this symbol indicates the cardinality, so we are defining it differently from the standard notation.

A.1.4 Classification of numbers:

From the above definitions there exists a natural hierarchical structure of numbers:

$$\mathbb{P} \subset \mathbb{N}, \quad \mathbb{Z} : \{\mathbb{N}, 0, -\mathbb{N}\}, \quad \mathbb{F} \perp \mathbb{Z}, \quad \mathbb{Q} : \mathbb{Z} \cup \mathbb{F}, \quad \mathbb{R} : \mathbb{Q} \cup \mathbb{I} \subset \mathbb{C}$$

1. The primes are a subset of the counting numbers: $\mathbb{P} \subset \mathbb{N}$.
2. The signed integers $\mathbb{Z}$ are composed of $\pm\mathbb{N}$ and 0, thus $\mathbb{N} \subset \mathbb{Z}$.
3. The fractionals $\mathbb{F}$ do not include of the signed integers $\mathbb{Z}$.
4. The rationals $\mathbb{Q} = \mathbb{Z} \cup \mathbb{F}$ are the union of the signed integers and fractionals.
5. Irrational numbers $\mathbb{I}$ have the special properties $\mathbb{I} \perp \mathbb{Q}$.
6. The reals $\mathbb{R} : \mathbb{Q}, \mathbb{I}$ are the union of rationals and irrationals $\mathbb{I}$.
7. Reals $\mathbb{R}$ may be defined as a subset of those complex numbers $\mathbb{C}$ having zero imaginary part.

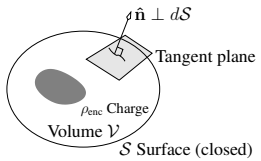
A.1.5 Rounding schemes

In Matlab/Octave there are five different rounding schemes (i.e., mappings): `round(x)`, `fix(x)`, `floor(x)`, `ceil(x)`, `roundb(x)`, with input $x \in \mathbb{R}$ and output $k \in \mathbb{N}$. For example $3 = \lfloor\pi\rceil, 3 = \lfloor e^1 \rceil = \lfloor 2.7183 \rceil$ rounds to the nearest integer, whereas 3=`floor(`π`)` rounds down while 3=`ceil(`e^1`)` = $\lceil e^1 \rceil$ rounds up. Rounding schemes are use for quantizing a number and generating a remainder. For example: `y=rem(x)` is equivalent to $y = x - \lfloor x \rceil$. Note `round(`π`)` $\equiv \lfloor\pi\rceil$ introduces negative remainders when ever a number rounds up ($\pi = \lceil\pi\rceil - 0.8541$).

The *continued fraction algorithm* (CFA), Sec. 2.4.4 (p. 50) is a recursive rounding scheme, operating on the reciprocal of the remainder. For example:

$$\begin{aligned}\exp(1) &= 3 + 1/(-4 + 1/(2 + 1/(5 + 1/(-2 + 1/(-7))))) - o\left(1.75 \times 10^{-6}\right), \\ &= [3; -4, 2, 5, -2, -7] - o(1.75 \times 10^{-6}).\end{aligned}$$

The expressions in square brackets is a notation for the CFA integer coefficients. The Octave/-Matlab function having $x \in \mathbb{R}$, is either `rat(x)` with output $\in \mathbb{N}$, or `rats(x)`, with output $\in \mathbb{F}$.



$$\nabla\cdot\boldsymbol{d} \equiv \lim_{\mathcal{V},\mathcal{S}\to 0}\left\{\frac{\oint_{\mathcal{S}} \hat{\mathbf{n}}\cdot\boldsymbol{d}\, d\mathcal{S}}{\mathcal{V}}\right\} \quad [\text{C/m}^3]$$

$$\begin{aligned}\mathcal{Q}_{\text{enc}} &= \oint_{\mathcal{S}} \hat{\mathbf{n}}\cdot\boldsymbol{d}\, d\mathcal{S} \\ &= \int_{\mathcal{V}} \rho_{\text{enc}}\, d\mathcal{V} \quad [\text{C}]\end{aligned}$$

Figure 5.5: *Left: On the left is the physical layout of the integral. Right: The top equation is the integral definition of the divergence of $\boldsymbol{d}$ as an integral over the closed surface $\mathcal{S}$ of the normal component of vector $\boldsymbol{d}$, given the limit as the surface and volume shrink to 0. The middle equation states the enclosed charge $\mathcal{Q}_{\text{enc}}$ in terms of the surface integral of the normal component of $\boldsymbol{d}$. The bottom equation gives the charge enclosed in terms of a volume integral over the enclosed charge density ρ_{enc}.*

the case of the gradient, if this definition is to make sense, the surface $\mathcal{S}$ must be closed, defining the volume $\mathcal{V}$. The difference here is that the surface integral is over the normal component of the vector field being operated on (Greenberg, 1988, p 762–763)

$$\nabla\cdot\boldsymbol{d} = \lim_{\mathcal{V},\mathcal{S}\to 0}\left\{\frac{\oint_{\mathcal{S}} \boldsymbol{d}\cdot\hat{\mathbf{n}}\, d\mathcal{S}}{\mathcal{V}}\right\} = \rho(x,y,z) \quad [\text{C/m}^3]. \tag{5.6.7}$$

As with the gradient, we have defined the surface with area $\mathcal{S}$, and volume $\mathcal{V}$. It is a necessary condition that as the area $\mathcal{S}$ goes to zero, so does the volume $\mathcal{V}$.

As defined previously (Eq. 5.6.6) and shown here in Fig. 5.5, $\hat{\mathbf{n}}$ is a unit vector normal to the closed iso-potential surface $\mathcal{S}$. The limit, as the volume and surface simultaneously go to zero, defines the total flux across the surface. Thus the surface integral is a measure of the total flux perpendicular to the surface. It is helpful to compare this formula with that for the gradient, Eq. 5.6.5.

Gauss's law: The definitions in Fig. 5.5 resulted in Gauss's law, a major breakthrough in vector calculus. For the electrical case, this is equivalent to the observation that the total flux across the surface is equal to the net charge enclosed by the surface. Since the volume integral over the charge density $\rho(x,y,z)$ is the total charge enclosed Q_{enc},

$$Q_{\text{enc}} = \int_{\mathcal{V}} \nabla\cdot\boldsymbol{D}\, d\mathcal{V} = \oint_{\mathcal{S}} \boldsymbol{d}\cdot\hat{\boldsymbol{n}}\, d\mathcal{S} \quad [\text{C}]. \tag{5.6.8}$$

When the surface integral over the normal component of $\boldsymbol{d}(\boldsymbol{x})$ is zero, the total charge enclosed is zero. If there is only positive (or negative) charge inside the surface, $\nabla\cdot\boldsymbol{d} = \rho(\boldsymbol{x}) = 0$. It is clear that this result only holds in the quasistatic limit, which is always satisfied because $\mathcal{S} \to 0$.

Taking the derivative with respect to time gives the total current normal to the surface:

$$\mathcal{I}_{\text{enc}} = \oint_{\mathcal{S}} \boldsymbol{d}\cdot\hat{\boldsymbol{n}}\, d\mathcal{S} = \dot{Q}_{\text{enc}} = \int_{\mathcal{V}} \dot{\rho}_{\text{enc}}\, d\mathcal{V} \quad [\text{A}]. \tag{5.6.9}$$

As summarized by Feynman (1970b, p. 13-2):

> The current leaving the closed surface $\mathcal{S}$ equals the rate of the charge leaving that volume $\mathcal{V}$, defined by that surface.

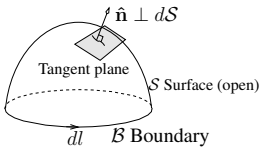
The integral definition reduces to a common-sense summary that can be grasped intuitively.

5.6.6 Magnetic fields h, b

5.6.7 Integral definition of the curl: $\nabla \times \boldsymbol{h} = \boldsymbol{c}$

The second of Maxwell's equations is As briefly summarized on page 227, the differential definition of the curl maps $\mathbb{R}^3 \underset{\nabla\times}{\to} \mathbb{R}^3$. The curl of the magnetic field strength $\boldsymbol{h}(\boldsymbol{x})$ is the current density $\boldsymbol{c} = \sigma\boldsymbol{e} + \dot{\boldsymbol{d}}$:

$$\nabla \times \boldsymbol{h} \equiv \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial_x & \partial_y & \partial_z \\ H_x & H_y & H_z \end{vmatrix} = \boldsymbol{c} \quad [\text{A/m}^2].$$



$$\nabla \times \boldsymbol{h} \equiv \lim_{\mathcal{B},\mathcal{S}\to 0} \left\{ \frac{\int_{\mathcal{S}} \hat{\mathbf{n}} \times \boldsymbol{h}\, d\mathcal{S}}{\mathcal{S}} \right\} \quad [\text{A/m}^2]$$

$$\mathcal{I}_{enc} = \int_{\mathcal{S}} (\nabla \times \boldsymbol{h}) \cdot \hat{\mathbf{n}}\, d\mathcal{S} = \int_{\mathcal{B}} \boldsymbol{h} \cdot d\boldsymbol{l} \quad [\text{A}]$$

Figure 5.6: *Left: The integral definition of the curl is related to that of the divergence (Eq. 5.6.7), as an integration over the tangent to the surface, except: (1) the curl is defined as the cross-product $\hat{\mathbf{n}} \times \boldsymbol{h}$ [A/m^2] of unit vector $\hat{\mathbf{n}}$ with the current density $\boldsymbol{h}$ (Greenberg, 1988, p 823), and (2) the surface is open, leaving a boundary $\mathcal{B}$ along the open edge. As with the divergence, which leads to Gauss's law, this definition leads to a second fundamental theorem of vector calculus:* Stokes's law *(also called the* curl theorem*). Right: Equations that summarize Stokes theorem (law).*

Curl and Stokes's law: Like the gradient and divergence, the curl may be written in integral form, allowing for the physical interpretation of its meaning:

> *The surface integral definition of $\nabla \times \boldsymbol{h} = \boldsymbol{c}$ [A/m^2], where the current density $\boldsymbol{c} = \sigma\boldsymbol{e} + \dot{\boldsymbol{d}}$ is perpendicular to the rotation plane of $\boldsymbol{h}$.*

Stokes's law states that the open surface integral over the normal component of the curl of the magnetic field strength ($\hat{\mathbf{n}} \cdot \nabla \times \boldsymbol{h}$ [A/m^2]) is equal to the line integral $\int_{\mathcal{B}} \boldsymbol{h} \cdot dl$ along the boundary $\mathcal{B}$. As summarized in Fig. 5.6, Stokes's law is

$$\begin{aligned} \mathcal{I}_{\text{enc}} &= \int_{\mathcal{S}} (\nabla \times \boldsymbol{h}) \cdot \hat{\mathbf{n}}\, d\mathcal{S} \\ &= \int_{\mathcal{S}} \boldsymbol{c} \cdot \hat{\mathbf{n}}\, d\mathcal{S} \\ &= \int_{\mathcal{B}} \boldsymbol{h} \cdot d\boldsymbol{l} \quad [\text{A}]. \end{aligned} \tag{5.6.10}$$

That is: *The line integral of $\boldsymbol{H}$ along the open surface's boundary $\mathcal{B}$ is equal to the total current enclosed $\mathcal{I}_{\text{enc}}$.*

In many texts the normalization (denominator under the integral) is a volume $\mathcal{V}$ (Greenberg, 1988, p 778,823–4). However, because the surface is open, this volume does not exist (when we define a volume, the surface must be closed). The definition must hold even in the limit

Greek letters

Frequently Greek letters, as provided in Fig. A.1, are associated in engineering and physics with a specific physical meaning. For example, ω [rad] is the radian frequency $2\pi f$, ρ [kgm/m^3] is commonly the density. ϕ, ψ are commonly used to indicate angles of a triangle, and $\zeta(s)$ is the Riemann zeta function. Many of these are so well established it makes no sense to define new terms, so we will adopt these common terms (and define them).

Likely you do not know all of these Greek letters, commonly used in mathematics. Some of them are pronounced in strange ways. The symbol ξ is pronounced "see," ζ is "zeta," β is "beta," and χ is "kie" (rhymes with pie and sky). I will assume you know how to pronounce the others, which are more phonetic in English. One advantage of learning LaTeX the powerful open-source math-oriented word-processing system used to write this book, is that math symbols are included, making then easily learned.

Table A.2: *Double-bold notation for the types of numbers. (#) is a page number. Symbol with an exponent denote the dimensionality. Thus $\mathbb{R}^2$ represents the real plane. An exponent of 0 denotes point, e.g., $\jmath \in \mathbb{C}^0$. It is reasonable to consider negative primes to be primes.*

Symbol (p. #)	Genus	Examples	Counter Examples
$\mathbb{N}$ (24)	Counting	1,2,17,3, 10^{20}	-5, 0, π, -10.3, 5j
$\mathbb{P}$ (24)	Prime	2, 3, 17, 199, 23993	0, 1, 4, 3^2, 12
$\mathbb{Z}$ (25)	Integer	-1, 0, 17, 5j, -10^{20}	1/2,π, $\sqrt{5}$
$\mathbb{Q}$ (25)	Rational	2/1, 3/2, 1.5, 1.14	$\sqrt{2}$, $3^{-1/3}$, π
$\mathbb{F}$ (26)	Fractional	1/2, 7/11·2	$1/\sqrt{2}$
$\mathbb{I}$ (26)	Irrational	$\sqrt{2}$, $3^{-1/3}$, π, e	Vectors
$\mathbb{R}$ (26)	Reals	$\sqrt{2}$, $3^{-1/3}$, π	$2\pi j$
$\mathbb{C}$ (27)	Complex	$1, \sqrt{2}j$, $3^{-j/3}$, π^j	Vectors
$\mathbb{G}$	Gaussian integers	$3 - 2\jmath \in \mathbb{Z} \cup \mathbb{C}$	complex integers

Double-Bold notation

Table A.2 indicates the symbol followed by a page number indication where it is discussed, and the Genus (class) of the number type. For example, $\mathbb{N} > 0$ indicates the infinite set of *counting numbers* $\{1, 2, 3, \cdots\}$, not including zero. Starting from any counting number, you get the next one by adding 1. Counting numbers are sometimes called the natural or cardinal numbers.

We say that a number is in the set with the notation $3 \in \mathbb{N} \subset \mathbb{R}$, which is read as "3 is in the set of counting numbers, which in turn is in the set of real numbers," or in vernacular language "3 is a real counting number."

Prime numbers ($\mathbb{P} \subset \mathbb{N}$) are taken from the counting numbers, but do not include 1.

The signed integers $\mathbb{Z}$ include 0 and negative integers. Rational numbers $\mathbb{Q}$ are historically defined to include $\mathbb{Z}$, a somewhat inconvenient definition, since the more interesting class are the *fractionals* $\mathbb{F}$, a subset of rationals $\mathbb{F} \in \mathbb{Q}$ that exclude the integers (i.e., $\mathbb{F} \perp \mathbb{Z}$). This is a useful definition because the rationals $\mathbb{Q} = \mathbb{Z} \cup \mathbb{F}$ are formed from the union of integers and fractionals.

The rationals may be defined, using set notation (a very sloppy notation with an incomprehensible syntax), as

$$\mathbb{Q} = \{p/q : q \neq 0 \;\&\; p, q \in \mathbb{Z}\},$$

index:	1	10	20	+1
+0	a	j	t	—
+5	e	o	y	z

This is helpful for quickly moving around the alphabet, when looking up words in the dictionary or an alphabetized list. If you forget how many letters there are in the English alphabet, this will help you recall it is 26.

When working with Fourier $\mathcal{FT}$ and Laplace $\mathcal{LT}$ transforms, lower case symbols are in the time domain while upper case indicates the frequency domain, as $f(t) \leftrightarrow F(\omega)$. An important exception is Maxwell's equations because they are so widely used as upper-case bold letters (e.g., $\boldsymbol{e}(\boldsymbol{x},\omega)$). It would seem logical to change this to $\boldsymbol{e}(\boldsymbol{x},t) \leftrightarrow \boldsymbol{E}(\boldsymbol{x},\omega)$, to conform.

A.1.3 Common mathematical symbols

There are many pre-defined symbols in mathematics, too many to summarize here. We shall only use a small subset, defined here.

- A *set* is a collection of objects that have a common property, defined by braces. For example, if set $P = \{a,b,c\}$ such that $a^2 + b^2 = c^2$, then members of P obey the Pythagorean theorem. Thus we could say that $\{1,1,\sqrt{2}\} \in P$.
- Number sets: $\mathbb{N},\mathbb{P},\mathbb{Z},\mathbb{Q},\mathbb{F},\mathbb{I},\mathbb{R},\mathbb{C}$ are briefly discussed below, and in greater detail in Sec. 2.1 (pp. 24–27).
- One can define sets of sets and subsets of sets, and this is prone (in my experience) to error. For example, what is the difference between the number 0 and the null set $\varnothing = \{0\}$? Is $0 \in \varnothing$? Ask a mathematician. This seems a lackluster construction in the world of engineering.
- A vector is a column n-tuple. For example $[3,5]^T = \begin{bmatrix}3\\5\end{bmatrix}$.
- The symbol $\perp$ is used in different ways to indicate two things are perpendicular, orthogonal, or in disjoint sets. In set theory $A \perp B$ is equivalent to $A \cap B = \varnothing$. If two vectors $\boldsymbol{e},\boldsymbol{h}$ are perpendicular $\boldsymbol{e} \perp \boldsymbol{h}$, then their inner product $\boldsymbol{e}\cdot\boldsymbol{h} = 0$ is zero. One must infer the meaning of $\perp$ from its context.

Table A.1: *List of all upper and lower case Greek letters used in the text.*

A	$\mathbf{A}$	α	$\boldsymbol{\alpha}$	alpha	N	$\mathbf{N}$	ν	$\boldsymbol{\nu}$	nu
B	$\mathbf{B}$	β	$\boldsymbol{\beta}$	beta	Ξ	$\mathbf{\Xi}$	ξ	$\boldsymbol{\xi}$	xi
Γ	$\mathbf{\Gamma}$	γ	$\boldsymbol{\gamma}$	gamma	O	$\mathbf{O}$	o	$\mathbf{o}$	omicron
Δ	$\mathbf{\Delta}$	δ	$\boldsymbol{\delta}$	delta	Π	$\mathbf{\Pi}$	π	$\boldsymbol{\pi}$	pi
E	$\mathbf{E}$	ε	$\boldsymbol{\varepsilon}$	epsilon	P	$\mathbf{P}$	ρ	$\boldsymbol{\rho}$	rho
Z	$\mathbf{Z}$	ζ	$\boldsymbol{\zeta}$	zeta	Σ	$\mathbf{\Sigma}$	σ	$\boldsymbol{\sigma}$	sigma
H	$\mathbf{H}$	η	$\boldsymbol{\eta}$	eta	T	$\mathbf{T}$	τ	$\boldsymbol{\tau}$	tau
Θ	$\mathbf{\Theta}$	θ	$\boldsymbol{\theta}$	theta	Y	$\mathbf{Y}$	υ	$\boldsymbol{\upsilon}$	upsilon
I	$\mathbf{I}$	ι	$\boldsymbol{\iota}$	iota	Φ	$\mathbf{\Phi}$	ϕ	$\boldsymbol{\phi}$	phi
K	$\mathbf{K}$	κ	$\boldsymbol{\kappa}$	kappa	X	$\mathbf{X}$	χ	$\boldsymbol{\chi}$	chi
Λ	$\mathbf{\Lambda}$	λ	$\boldsymbol{\lambda}$	lambda	Ψ	$\mathbf{\Psi}$	ψ	$\boldsymbol{\psi}$	psi
M	$\mathbf{M}$	μ	$\boldsymbol{\mu}$	mu	Ω	$\mathbf{\Omega}$	ω	$\boldsymbol{\omega}$	omega

when the curved surface $\mathcal{S}$ degenerates to a plane, with the boundary $\mathcal{B}$ enclosing $\mathcal{S}$. In this limit there is no volume.

To resolve this problem, we take the normalization to be the surface $\mathcal{S}$ of Fig. 5.6. Note that in the limit $\mathcal{B} \to 0$, the limiting definition is independent of any curvature, since the integral is over the normal component of $\boldsymbol{h}$ (i.e., $\hat{\mathbf{n}} \perp \boldsymbol{h}(\boldsymbol{x},t)$) at the limit point $\boldsymbol{x}_o$. The net flux is independent of the curvature of $\mathcal{S}$ as $\mathcal{B} \to 0$.

Finally, the curl of a vector is composed of three wedge products. However, the curl is defined only in $\mathbb{R}^3$. Since the wedge product is defined in $\mathbb{R}^2$, for vectors of any length, it can be used to generalize the curl to every vector space, independent of its dimension.

The wedge product (Fig. 3.4, p. 119) is also a natural tool for dealing with the Hall effect and superconductivity, since these are two-dimensional phenomena.

Summary: Since integration is a linear process (sums of smaller elements), we can tile (tessellate) the surface, breaking it up into smaller surfaces and their boundaries, the sum of which is equal to the integral over the original boundary. This is an important concept that leads to the proof of Stokes's law.

Table 5.1 (p. 225) provides a description of the three basic integration theorems along with their mapping domains. The integral formulations of Gauss's and Stokes's laws use $\hat{\mathbf{n}}\cdot\boldsymbol{d}$ and $\boldsymbol{h}\times\hat{\mathbf{n}}$ in the integrands. The key distinction between the two laws naturally follows from the properties of the scalar ($\boldsymbol{A}\cdot\boldsymbol{B}$) and vector ($\boldsymbol{A}\times\boldsymbol{B}$) products, as discussed in Fig. 3.4 (p. 119). To fully appreciate the differences between Gauss's and Stokes's laws, we must master these two types of vector products.

Paraphrasing Feynman (1970b, 3-12), we have

1. $\Phi_2 = \Phi_1 + \int_1^2 \nabla\Phi\cdot d\mathcal{S}$

   The line integral of an analytic function depends only on the end points.

2. $\boldsymbol{d}\cdot\hat{\mathbf{n}}\,d\mathcal{S} = \nabla\cdot\boldsymbol{d}\,d\mathcal{V} = \rho\,d\mathcal{V}$

   The normal component over a surface integral equal the divergence over the volume integral.

3. $\oint_{\mathcal{B}} \boldsymbol{e}\cdot d\boldsymbol{l} = \int_{\mathcal{S}} (\nabla\times\boldsymbol{e})\cdot\hat{\mathbf{n}}\,d\mathcal{S} = -\dot{\boldsymbol{B}}\cdot\hat{\mathbf{n}}\,d\mathcal{S}$

   The line integral of the electric field (the induced EMF in any loop) equals the integral of the normal component of time-rate-of-change of the magnetic flux. The induced EMF is the Thévenin voltage in series with the loop. This explains the working of an electrical transformer.

5.6.8 Helmholtz's decomposition theorem

This is the time to rethink everything we have defined in terms of the two types of vector fields that decompose every analytic vector field as shown in Table 5.3. The *irrotational field* is defined as one that is curl free. The *incompressible field* is one that is divergence free.

According to Helmholtz's decomposition theorem, every analytic vector field may be decomposed into independent rotational and compressible components (Helmholtz, 1978). An alternative name for Helmholtz's decomposition theorem is the fundamental theorem of vector calculus (FTVC). Gauss's and Stokes's theorems,[12] along with Helmholtz's decomposition

[12] These theorems are mathematical relationships that follow from physical principles.

theorem, are the three fundamental theorems of vector calculus. Portraits of Helmholtz and Kirchhoff are provided in Fig. 5.8 (p. 261).

Table 5.3: *The four possible classifications of scalar and vector potential fields: rotational/irrotational and compressible/incompressible. Rotational fields are generated by the vector potential (e.g., $\boldsymbol{A}(\boldsymbol{x},t)$), while compressible fields are generated by the scalar potentials (e.g., voltage $\phi(\boldsymbol{x},t)$, velocity ψ, pressure $\varrho(\boldsymbol{x},t)$, or temperature $T(\boldsymbol{x},t)$).*

Field: $\nu(\boldsymbol{x},t)$	Compressible $\nabla\cdot\nu\neq 0$	Incompressible $\nabla\cdot\nu=0$
Rotational $\nabla\times\nu\neq 0$	$\nu=\nabla\phi+\nabla\times\boldsymbol{\omega}$ Vector wave Eq. $\boldsymbol{\nabla^2}\nu=\frac{1}{c^2}\ddot{\nu}$	$\nu=\nabla\times\boldsymbol{w}$ Lubrication theory Boundary layers
Irrotational conservative $\nabla\times\nu=0$	Acoustics $\nu=\nabla\psi$ $\nabla^2\varrho(\boldsymbol{x},t)=\frac{1}{c^2}\ddot{\varrho}(\boldsymbol{x},t)$	Statics $\nabla^2\phi=0$ Laplace's Eq. ($c\to\infty$)

The four categories of linear fluid flow: The following is a summary of the four cases for fluid flow, as shown in Table. 5.3:

1,1 Compressible, rotational fluid (general case): $\nabla\psi\neq 0$, $\nabla\times\boldsymbol{w}\neq 0$. This is wave propagation in a medium where viscosity cannot be ignored, as in acoustics close to the boundaries, where viscosity contributes to losses (Batchelor, 1967).

1,2 Incompressible, rotational fluid (lubrication theory): $\nu=\nabla\times\boldsymbol{w}\neq 0, \nabla\cdot\nu=0, \nabla^2\psi=0$. In this case, the flow is dominated by the walls, while the viscosity and heat transfer introduce shear. This is typical of lubrication theory (solenoidal fields).

2,1 Compressible, irrotational fluid (acoustics): $\nu=\nabla\psi$, $\nabla\times\boldsymbol{w}=0$. Here losses (viscosity and thermal diffusion) are small (assumed to be zero). We can define a velocity potential ψ, the gradient of which gives the air particle velocity, thus $\nu=-\nabla\psi$. For an irrotational fluid, $\nabla\times\nu=0$ (Greenberg, 1988, p 826). This is the case for the conservative field, where $\int_{\mathcal{B}}\nu\cdot\hat{\mathbf{n}}dR$ depends on only the end points and $\int_{\mathcal{B}}\nu\cdot\hat{\mathbf{n}}dR=0$. When a fluid may be treated as having no viscosity, it is typically assumed to be irrotational, since it is viscosity and shear that lead to rotation (Greenberg, 1988, p 814). A fluid's angular velocity is $\Omega=\frac{1}{2}\nabla\times\nu=0$; thus irrotational fluids have zero angular velocity ($\Omega=0$).

2,2 Incompressible, irrotational fluid (statics): $\nabla\cdot\nu=0$ and $\nabla\times\nu=0$; thus $\nu=\nabla\psi$ and $\nabla^2\psi=0$. An example is water in a small space at low frequencies, where the wavelength is long compared to the size of the container; the fluid may be treated as incompressible. When $\nabla\times\nu=0$, the effects of viscosity may be ignored, as it is the viscosity that creates the shear that leads to rotation. This is the case in modeling the cochlea, where fluid losses are ignored and the quasistatic limit is justified.

In summary, each of these cases is an approximation that best applies in the low-frequency limit. This is why it is called *quasistatic*, meaning low but not zero frequency, where the wavelength is more than twice the diameter.

A magnetic solenoidal field is a uniform-flux field $\boldsymbol{b}_z(\boldsymbol{x})$ that is generated by a solenoidal coil and, to an excellent approximation, is uniform inside the coil, making it similar to that of

Appendix A

Notation

A.1 Number systems

The notation used in this book is defined in this appendix so that it may be quickly accessed.[1] Where the definition is sketchy, page numbers are provided where these concepts are fully explained, along with many other important and useful definitions. For example a discussion of $\mathbb{N}$ may be found on page 24. Math symbols such as $\mathbb{N}$ may be found at the top of the index, since they are difficult to alphabetize.

A.1.1 Units

Strangely, or not, classical mathematics, as taught today in schools, does not seem to acknowledge the concept of physical units. Units seem to have been abstracted away. This makes mathematics distinct from physics, where almost everything has units. Presumably this makes mathematics more general (i.e., abstract). But for the engineering mind, this is not ideal, or worse, as it necessarily means that important physical meaning, by design, has been surgically removed. We shall use SI units whenever possible, which means this book is not not a typical book on mathematics. Spatial coordinates are quoted in meters [m], and time in seconds [s]. Angles in degrees have no units, whereas radians have units of inverse-seconds [s^{-1}]. A complete list of SI units may be found at https://physics.nist.gov/cuu/pdf/sp811.pdf and in Graham et al. (1994), for a discussion of basic math notation.

When writing a complex number we shall adopt $1\jmath$ to indicate $\sqrt{-1}$. Matlab/Octave allows either $1\imath$ or $1\jmath$.

Units are SI; angles are in degrees [deg] unless otherwise noted. The units for π are always radians [rad]. For example $\sin(\pi), e^{\jmath 90^\circ}=e^{\jmath\pi/2}$.

A.1.2 Notation

We use $\ln$ as the log function base e, $\log$ as base 2, and π_k to indicate the kth prime (e.g., $\pi_1=2,\pi_2=3$).

It is helpful to know where the letters of the alphabet are. Everyone knows the first letter is /a/ and the last /z/, but what is the 10th or 20th letter? The table below shows that /j/ is the 10th letter, /t/ the 20th, /o/ the 15th, and /z/ the 26th.

[1] https://en.wikipedia.org/wiki/List_of_mathematical_symbols_by_subject#Definition_symbols

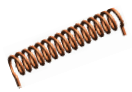


Figure 5.7: *A solenoid is a uniform coil of wire. The term comes from the Greek meaning "shaped like a pipe or channel." When a steady current is passed through the wire, a uniform magnetic field intensity $\boldsymbol{h}$ is created. Such a coil is indistinguishable from a permanent bar magnet with its north and south poles. Depending on the direction of the current, one end of a finite solenoidal coil is the north pole of the magnet, and the other end is the south pole. The uniform field inside the coil is called* solenoidal*, a confusing synonym for* rotational *(Figure from Wikipedia).*

a permanent magnet. As a result, the divergence of a solenoidal field is approximately zero, which makes it incompressible ($\nabla\cdot = 0$) and rotational ($\nabla\times \neq 0$).

You need to know the term *solenoidal* since it is widely used. However the preferred terms are *incompressible* and *rotational*. Strictly speaking, the term *solenoidal field* applies to only a magnetic field produced by a solenoid.

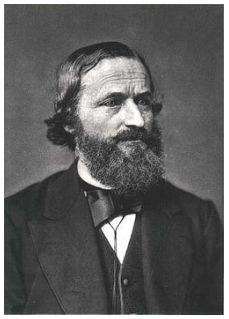


Figure 5.8: *Left: Hermann von Helmholtz (taken from Helmholtz (1978)), Right: Gustav Kirchhoff. Together they were the first to account for viscous (Helmholtz, 1858, 1978, 1863b) and thermal (Kirchhoff, 1868, 1974; Tran-Cong, 1993, English translation) losses in the acoustic propagation of airborne sound, as first experimentally verified by Mason (1928) p. 311).*

Application and derivation

Helmholtz's decomposition theorem is expressed as the linear sum of a scalar potential $\phi(x, y, z)$ (think voltage) and a vector potential (think magnetic vector potential). Specifically,

$$\boldsymbol{b}(\boldsymbol{x}, s) = \nabla \times A + \partial_t \boldsymbol{d} \tag{5.6.11}$$

where ϕ is the scalar and $\boldsymbol{a}$ is the vector potential as a function of the Laplace frequency s. Of course, this decomposition is general (not limited to the electromagnetic case). It applies to linear fluid vector fields, which include linear fluids (water, air). If rotational and dilation become coupled, this relationship will break down.

This theorem is easily stated (and proved) but less easily appreciated (Heras, 2016). A physical description is helpful: Every vector field may be split into two independent parts: translation and rotation. We have seen this same idea in vector algebra, where the scalar and wedge products of two vectors are perpendicular (Fig. 3.4, p. 119).

For example, think of linear versus angular momentum, which are independent in that they represent different ways of delivering kinetic energy via different modalities (degrees of freedom, DoF). Linear and rotational motions are a common theme in physics, rooted in geometry.

Thus it seems a natural extension to split a vector field into independent dilation and rotation parts.

Example: A fluid with mass and momentum can be moving along a path and independently be rotating. These independent modes of motion correspond to different types of kinetic energy (modes), such as translational, compressional, and rotational. Each eigenmode of vibration can be viewed as an independent degree of freedom (DoF).

Helmholtz's decomposition theorem (FTVC) quantifies these degrees of freedom. To prove the construction, second-order vector identities **DoC**: $\nabla\cdot\nabla\times() = 0$ and **CoG**: $\nabla\times\nabla() = 0$ may be used to verify the FTVC. The FTVC also plays an important role when applied to Maxwell's equations.

Helmholtz's decomposition theorem proof: To prove Eq. 5.6.11 we must understand how it splits the vector field into parts. The identities have a physical meaning: Every vector field may be split into its translational and rotational parts. If $\boldsymbol{e}$ is the electric field [V/m], ϕ is the voltage, and $\boldsymbol{a}$ is the current induced rotational part.

To do this we need the two key vector identities that are always zero for analytic fields: the curl of the gradient ($\nabla\times\nabla()$),

$$\nabla\times\nabla\phi(\boldsymbol{x}) = 0, \tag{5.6.12}$$

and the divergence of the curl[13] ($\nabla\cdot\nabla\times()$)

$$\nabla\cdot(\nabla\times\boldsymbol{a}) = 0. \tag{5.6.13}$$

These identities are easily verified by working out a few specific examples based on the definitions of the three operators, gradient, divergence, and curl, or in terms of the operator's integral definitions (see page 254).

By applying these two identities to Helmholtz's decomposition, we can appreciate the theorem's significance. We can work backward via a physical argument that rotational momentum (rotational energy) is independent of translational momentum. Once these forces are made clear, the vector operations take on meaning. One might conclude that the physics is simply related to geometry via the scalar and vector products.

Specifically, if we take the divergence of Eq. 5.6.11 and use the DoG, then

$$-\frac{1}{\epsilon_o}\rho = \nabla\cdot\boldsymbol{e} = \nabla\cdot\{-\nabla\phi + \cancelto{0}{\nabla\times\boldsymbol{a}}\} = -\nabla\cdot\nabla\phi = -\nabla^2\phi,$$

If we take the curl, then

$$-\dot{\boldsymbol{b}} = \nabla\times\boldsymbol{e} = \nabla\times\{\cancelto{0}{-\nabla\phi} + \nabla\times\boldsymbol{a}\} = \nabla\times\nabla\times\boldsymbol{a} = \boldsymbol{\nabla}^2\boldsymbol{a} - \boldsymbol{\nabla}^2\boldsymbol{a},$$

since the $\nabla\times\nabla$ zeros the scalar field $\phi(x, y, z)$.

5.6.9 Second-order operators

In addition to the first-order vector derivatives are second-order operators, the most important being the scalar Laplacian $\nabla^2() = \nabla\cdot\nabla()$ and vector Laplacian $\boldsymbol{\nabla}^2() = \nabla\cdot\nabla()$, which operates on vectors.[14]

[13] Helmholtz was the first person to apply mathematics in modeling the eye and the ear (Helmholtz, 1863a).

[14] https://en.wikipedia.org/wiki/Del_in_cylindrical_and_spherical_coordinates#Non-trivial_calculation_rules

Appendices

BAD ERRORS: Terminology:

There are six second-order combinations of ∇, requiring six mnemonics (Table 5.1, p. 225):

1. DoG: Divergence of the gradient (*scalar Laplacian* operates on scalar potentials (Greenberg, 1988, p. 779)):

$$\begin{aligned}\nabla^2\phi &= (\nabla\cdot\nabla)\phi \\ &= \frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}\end{aligned} \tag{5.6.14}$$

2. **DoG**: Divergence of the Gradient (Bull-Dog, the *vector Laplacian* $\boldsymbol{\nabla}^2$ (Sommerfeld, 1952, p 33)):

$$\begin{aligned}\boldsymbol{\nabla}^2\boldsymbol{a} &= (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\boldsymbol{a} \\ &= \frac{\partial^2\boldsymbol{a}}{\partial x^2}+\frac{\partial^2\boldsymbol{a}}{\partial y^2}+\frac{\partial^2\boldsymbol{a}}{\partial z^2} \\ &= \boldsymbol{\nabla}^2\boldsymbol{a}-\nabla\times\nabla\times\boldsymbol{a}\end{aligned} \tag{5.6.15}$$

3. **gOd**: Gradient of the Divergence ($\boldsymbol{\nabla}^2\boldsymbol{a}=\nabla(\nabla\cdot\boldsymbol{a})$)

$$\begin{aligned}\boldsymbol{\nabla}^2\boldsymbol{a} &= \nabla(\nabla\cdot\boldsymbol{a}) \\ &= \nabla\left(\frac{\partial a_x}{\partial x}+\frac{\partial a_y}{\partial y}+\frac{\partial a_z}{\partial z}\right) \\ &= \left(\hat{\mathbf{x}}\frac{\partial}{\partial x}+\hat{\mathbf{y}}\frac{\partial}{\partial y}+\hat{\mathbf{z}}\frac{\partial}{\partial z}\right)\left(\frac{\partial a_x}{\partial x}+\frac{\partial a_y}{\partial y}+\frac{\partial a_z}{\partial z}\right) \\ &= \hat{\mathbf{x}}\frac{\partial}{\partial x}\nabla\cdot A+\hat{\mathbf{y}}\frac{\partial}{\partial y}\nabla\cdot A+\hat{\mathbf{z}}\frac{\partial}{\partial z}\nabla\cdot A\end{aligned} \tag{5.6.16}$$

4. **CoC:** Curl of the curl ($\nabla\times\nabla\times$ **= gOd- DoG**)

$$\nabla\times\nabla\times\boldsymbol{a}=\nabla(\nabla\cdot\boldsymbol{a})-(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\boldsymbol{a}$$

   which is simply a reorganization of Eq. 5.6.15 (Sommerfeld, 1952, p 33, Eq. 2b).

5. DoC: Divergence of the curl ($\nabla\cdot\nabla\times=0$)

6. CoG: Curl of the gradient ($\nabla\times\nabla=0$)

DoC$(\cdot)$ and CoG$(\cdot)$ are special because they are always zero:

$$\nabla\cdot\nabla\times\boldsymbol{a}=0,\qquad\nabla\times\nabla\phi=0, \tag{5.6.17}$$

making them useful in proving the FTVC (Eq. 5.6.11, p. 261).

The third special vector identity

$$\nabla\times\nabla\times\boldsymbol{a} = \boldsymbol{\nabla}^2\boldsymbol{a} - \boldsymbol{\nabla}^2\boldsymbol{a}, \tag{5.6.18}$$

as the difference between **gOd** and **CoC** (i.e., **DoG** = **gOd** − **CoC**):

$$\boldsymbol{\nabla}^2() = \boldsymbol{\nabla}^2() - \nabla\times\nabla\times(). \tag{5.6.19}$$

The role of **gOd** ($\boldsymbol{\nabla}^2$) is commonly ignored because it is zero for the magnetic wave equation, due to there being no magnetic charge [$\nabla\cdot\boldsymbol{b}(x,t) = 0$; thus $\boldsymbol{\nabla}^2\boldsymbol{b}(x,t) \equiv 0$]. However for the electric vector wave equation it plays a role as a source term:

$$\boldsymbol{\nabla}^2\phi(\boldsymbol{x},t) = -\nabla\boldsymbol{e}(\boldsymbol{x},t) = -\frac{1}{\epsilon_o}\boldsymbol{\nabla}^2\boldsymbol{d}(x,t) = -\frac{1}{\epsilon_o}\nabla\rho(\boldsymbol{x},t),$$

or since $\nabla\cdot\boldsymbol{d} = \rho$,

$$\boldsymbol{\nabla}^2\boldsymbol{d}(\boldsymbol{x},t) = \nabla\nabla\cdot\boldsymbol{d} = -\nabla\rho(\boldsymbol{x},t).$$

When the charge density is inomogeneous, such as the case of a plasma (e.g., the sun), this term plays an important role as a source term in the electric wave equation. This case needs to be further explored through some physical examples.

Exercise #*16*
Show that **DoG**: $\boldsymbol{\nabla}^2$ and **gOd**: $\boldsymbol{\nabla}^2$ differ.

Solution: Use $\nabla\times\nabla\times()$ to explore this relationship. If **DoG** and **gOd** were the same, $\nabla\times\nabla\times()$ would be null.■

Exercise #*17*
What is the difference between **DoG**: ∇^2 and Bull-DoG: $\boldsymbol{\nabla}^2$?

Solution: DoG operates on scalar functions while Bull-DoG operates on vector functions.■

Discussion: It is helpful to view these two groups as playing fundamentally different roles:

utility operators $\nabla\times\nabla\times()$: **DoG**: $\nabla^2()$, and gOd: $\boldsymbol{\nabla}^2()$,

and

identity operators DoC: $DoC() = 0$ and CoG: $\nabla\times\nabla()$=0 (Eq. 5.6.17).

When using second-order differential operators, we must be careful with the order of operations, which can be subtle. Most of this is common sense. For example, don't operate on a scalar field with $\nabla\times$, and don't operate on a vector field with ∇.[15] **DoG** acts on on each vector component $\boldsymbol{\nabla}^2\boldsymbol{a} = \nabla^2 A_x\hat{\mathbf{x}} + \nabla^2 A_y\hat{\mathbf{y}} + \nabla^2 A_z\hat{\mathbf{z}}$ (Eq. 5.6.15), which is very different from the action of the Laplacian (DoG).

5.7 The unification of electricity and magnetism

Once you have mastered the three basic vector operations—gradient, divergence, and curl—we are ready to appreciate Maxwell's equations. Like the vector operations, these equations may be written in integral or differential form. An important difference is that with Maxwell's equations, we are dealing with well-defined physical quantities. The scalar and vector fields take on meaning and units. Thus, to understand these important equations, we must master

– 10.4: Find the charge density per unit area ($\sigma = Q/A$, where Q is charge and A is area) on the surface of each plate. Hint: $\boldsymbol{E} = -\nabla V$, and Gauss's law states that $\int_S \boldsymbol{D}\cdot\hat{\mathbf{n}}dS = Q_{\text{enc}}$.

– 10.5: Determine the per-unit-area capacitance C of the system.

5.9.9 Webster horn equation

Problem # *11: Horns illustrate an important generalization of the solution of the one dimensional wave equation in regions where the properties (i.e., area of the tube) vary along the axis of wave propagation. Classic applications of horns are in vocal tract acoustics, loudspeaker design, cochlear mechanics, and any case that has wave propagation. Write the formula for the Webster horn equation, and explain the variables.*

5.10 Further readings

The above concepts come straight from mathematical physics, as developed in the 17th–19th centuries. Much of this was first developed in acoustics by Helmholtz, Stokes, and Rayleigh, following through Green's footsteps, as described by Rayleigh (1896). When it comes to fully appreciating Green's theorem and reciprocity, I have found Rayleigh (1896) to be the key reference. To repeat my reading experience, start with Brillouin (1953, 1960), followed by Sommerfeld (1952) and Pipes (1958). Second-tier reading contains many items: Morse (1948); Sommerfeld (1949); Morse and Feshbach (1953); Ramo et al. (1965); Feynman (1970a); Boas (1987). A third tier might include Helmholtz (1863a); Fry (1928); Lamb (1932); Bode (1945); Montgomery et al. (1948); Beranek (1954); Fagen (1975); Lighthill (1978); Hunt (1952); Olson (1947). Other physics writings include the impressive series of mathematical-physics text books by authors J.C. Slater, and Landau and Lifshitz.[20]

Successful reading of these books critically depends on what you already know. It assumes a rudimentary (high school) level math has been mastered. You must enter the reading list at a level that assumes you to understand the previous material. Read in the order that helps you best understand the material. Pick topics you are intersted in, you will make progress faster. An example that almost everyone likes is acoustics, such as loudspeakers and microphones. Electronics is a good match for these acoustic electromechanical gagets.

[20] https://www.amazon.com/Mechanics-Third-Course-Theoretical-Physics/dp/0750628960

Problem # *8: Consider the next equation in three dimensions. Take the divergence of both sides and integrate over a volume V (closed surface S):*

$$\int_V \nabla \cdot [\nabla \times \mathbf{H}] dV = \int_V \nabla \cdot \mathbf{C} dV.$$

– 8.1: What happens to the left-hand side of this equation? Hint: Can you apply a vector identity? *Apply the divergence theorem (sometimes known as Gauss's theorem) to the right-hand side of the equation, and interpret your result.* Hint: Can you relate your result to one of Kirchhoff's laws?

5.9.7 Second-order differentials

Problem # *9: This problem is about second-order vector differentials.*

– 9.1: If $\boldsymbol{v}(x,y,z) = \nabla\phi(x,y,z)$, then what is $\nabla \cdot \boldsymbol{v}(x,y,z)$?

– 9.2: Evaluate $\nabla^2\phi$ and $\nabla \times \nabla\phi$ for $\phi(x,y) = xe^y$.

– 9.3: Evaluate $\nabla \cdot (\nabla \times \boldsymbol{v})$ and $\nabla \times (\nabla \times \boldsymbol{v})$ for $\boldsymbol{v} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$.

– 9.4: When $\boldsymbol{V}(x,y,z) = \nabla(1/x + 1/y + 1/z)$, what is $\nabla \times \boldsymbol{V}(x,y,z)$?

– 9.5: When was Maxwell born and when did he die? How long did he live (within $\pm$10 years)?

5.9.8 Capacitor analysis

Problem # *10: Find the solution to the Laplace equation between two infinite*[19] *parallel plates separated by a distance d. Assume that the left plate at $x = 0$ is at voltage $V(0) = 0$ and the right plate at $x = d$ is at voltage $V_d \equiv V(d)$.*

– 10.1: Write Laplace's equation in one dimension for $V(x)$.

– 10.2: Write the general solution to your differential equation for $V(x)$.

– 10.3: Apply the boundary conditions $V(0) = 0$ and $V(d) = V_d$ to determine the constants in your equation from question 10.2.

[19] We study plates that are infinite because this means the electric field lines are perpendicular to the plates, running directly from one plate to the other. However, we solve for per-unit-area characteristics of the capacitor.

Table 5.4: *This table defines the four electro-magnetic variables that define Maxwell's equations, $\boldsymbol{e}, \boldsymbol{h}, \boldsymbol{d}, \boldsymbol{b}$. They are easily understood by noting their units. A fifth is called the conduction current (CC), define by $\boldsymbol{j}_c = \sigma E$, which is the same current that is define by Ohm's Law. The variables of Maxwell's equations have names names shown in the table belo (e.g., EF, ED, MF, MI) and units (in square brackets [SI units]). The units are necessary to obtain a full understanding of each of the four variables and their corresponding equations. For example, Eq.* EF *has units [V/m]. By integrating $\boldsymbol{e}$ from $x = a, b$, one obtains the voltage difference between the two points. The* speed of light in-vacuo *is $c = 3 \times 10^8 = 1/\sqrt{\mu_o \epsilon_o}$ [m/s], and the* characteristic resistance *of light is $r_o = 377 = \sqrt{\mu_o/\epsilon_o}$ [Ω] (i.e., ohms).*

Symbol	Name	Units	Maxwell's Eq.
$\boldsymbol{e}$	EF: Electric field strength	[V/m]	$\nabla \times \boldsymbol{e} = -\partial_t \boldsymbol{b}$
$\boldsymbol{d} = \epsilon_o \boldsymbol{e}$	ED: Electric displacement (flux density)	[C/m^2]	$\nabla \cdot \boldsymbol{d} = \rho$
$\boldsymbol{h}$	MF: Magnetic field strength	[A/m]	$\nabla \times \boldsymbol{h} = \boldsymbol{j_m} + \partial_t \boldsymbol{d}$
$\boldsymbol{b} = \mu_o \boldsymbol{h}$	MI: Magnetic induction (flux density)	[Wb/m^2]	$\nabla \cdot \boldsymbol{b} = 0$
$\boldsymbol{j}_c$	CC: conduction current	[A/m^2]	$\sigma \boldsymbol{e}$

both the names and units of the four fields $\boldsymbol{e}, \boldsymbol{h}, \boldsymbol{b}, \boldsymbol{d}$, as described in Table 5.4. Note that $\boldsymbol{j}_c$ is not $\sqrt{-1}$.

If you wish to become an Electrical Engineer (there is an unlimited demand for well trained EEs), then the number one topic to master is ME.

Field strength $\boldsymbol{e}, \boldsymbol{h}$: As summarized by Eqs. 5.7.1, there are two field strengths: the electric $\boldsymbol{e}$ with units of [V/m] and the magnetic $\boldsymbol{h}$ with units of [A/m]. The ratio $\boldsymbol{j}_c = \boldsymbol{e}/\boldsymbol{h} = \sqrt{\mu_o/\epsilon_o} = 377$ [ohms] for in-vacuo plane-waves (μ_o, ϵ_o). It's a vector because of the direction of flow, as determined by a voltage drop. Once you grasp the concept of Ohm's Law, your there.

If two conducting plates are placed 1 [m] apart, with 1 [V] across them, the electric field is $\boldsymbol{e} = 1$ [V/m]. If a charge (i.e., an electron) is placed in an electric field, it feels a force $\boldsymbol{f} = q\boldsymbol{e}$ [N], where q is the magnitude of the charge [C].

To help us understand the meaning of $\boldsymbol{h}$, consider the solenoid made of wire, as shown in Fig. 5.7, that carries a current of I_θ [A]. The axial (along the long axis) magnetic field H_z inside such a solenoid is uniform, with a direction that depends on the polarity of I_θ, is

$$H_z = \frac{N}{L} I_\theta \quad [\text{Wb}],$$

where L is the length of the coil, N is the number of turns.

Flux $\boldsymbol{d}, \boldsymbol{b}$: Flux is a flow, such as the mass flux of water flowing in a pipe [kg/s] driven by a force (pressure drop) across the ends of the pipe, or the heat flux in a thermal conductor, that has a temperature drop across it (i.e., a window or a wall). The flux is the same as the flow, be it charge, mass, or heat (Table 3.2, p. 149). In Maxwell's equations there are also two fluxes: the electric flux $\boldsymbol{d}$ and the magnetic flux $\boldsymbol{b}$. The flux density units for $\boldsymbol{d}$ is [C/m^2], and the magnetic flux density $\boldsymbol{b}$ is measured in either weber [Wb/m^2]) or its SI equivalent [tesla] [T].

5.7.1 Maxwell's equations

Maxwell's equations (ME) consist of two curl equations (Eqs. 5.7.1) operating on the field strengths EF $\boldsymbol{e}$ [V/m^2] and MF $\boldsymbol{h}$ [A/m^2], and two divergence equations (Eq. 5.7.2) operating

[15] This operation defines a *dyadic tensor*, which I view as a obtuse generalization of a vector.

on the field fluxes ED $\boldsymbol{d}$ and MI $\boldsymbol{b}$. In matrix format, the ME are

$$\begin{aligned}\nabla\times\begin{bmatrix}\boldsymbol{e}(\boldsymbol{x},t)\\ \boldsymbol{h}(\boldsymbol{x},t)\end{bmatrix} &= \partial_t\begin{bmatrix}-\boldsymbol{b}(\boldsymbol{x},t)\\ \boldsymbol{d}(\boldsymbol{x},t)\end{bmatrix}\\ &= \begin{bmatrix}0 & -\mu_o\\ \epsilon_o & 0\end{bmatrix}\partial_t\begin{bmatrix}\boldsymbol{e}(\boldsymbol{x},t)\\ \boldsymbol{h}(\boldsymbol{x},t)\end{bmatrix} \qquad (5.7.1)\\ &\leftrightarrow \begin{bmatrix}0 & -s\mu_o\\ \sigma_o+s\epsilon_o & 0\end{bmatrix}\begin{bmatrix}\mathbf{E}(\boldsymbol{x},\omega)\\ \mathbf{H}(\boldsymbol{x},\omega)\end{bmatrix}.\end{aligned}$$

When the medium is conducting, $\partial_t\boldsymbol{d}$ must be replaced by $\boldsymbol{c}=\sigma_o E+\partial_t\boldsymbol{d}\leftrightarrow(\sigma_o+s\epsilon_o)\boldsymbol{e}(\boldsymbol{x},\omega)$ where $\sigma_o+s\epsilon_o$ is an admittance density [$\mho/\text{m}^2$].

There are also two auxiliary equations:

$$\nabla\cdot\begin{bmatrix}\boldsymbol{d}\\ \boldsymbol{b}\end{bmatrix} = -\partial_t\begin{bmatrix}\rho(\boldsymbol{x})\\ 0\end{bmatrix}. \qquad (5.7.2)$$

The top equation states conservation of charge, while the lower states that there is no magnetic charge. When expressed in integral format, Stokes's law follows from the curl equations and Gauss's law from the divergence equations.

Exercise #*18*

When a static current is flowing in a wire in the $\hat{\mathbf{z}}$ direction, the magnetic flux is determined by Stokes's theorem (Fig. 5.6). Thus, just outside the wire we have

$$\mathcal{I}_{\text{enc}} = \int_{\mathcal{S}}(\nabla\times\boldsymbol{h})\cdot\hat{\mathbf{n}}\,d|\mathcal{S}| = \int_{\mathcal{B}}\boldsymbol{h}\cdot d\boldsymbol{l} \quad [\text{A}].$$

For this simple geometry, the current in a wire is related to $\boldsymbol{h}(\boldsymbol{x},t)$ by

$$\mathcal{I}_{\text{enc}} = \int_{\mathcal{B}}\boldsymbol{h}\cdot d\boldsymbol{l} = H_\phi 2\pi r.$$

Here H_ϕ is perpendicular to both the radius r and the direction of the current $\hat{\mathbf{z}}$. Thus

$$\boldsymbol{h}_\phi = \frac{\mathcal{I}_{\text{enc}}}{2\pi r},$$

where $\boldsymbol{h}_\phi$ is attenuated by $1/r$ (Ramo et al., 1965, Eq. 9, page 244).

Exercise #*19*

Explain how Stokes's theorem may be applied to $\nabla\times\boldsymbol{e}=-\dot{\boldsymbol{b}}$, and explain what it means. Hint: This is the same argument given above for the current in a wire, but for the electric case.

Solution: Integrating the left side of equation Eq. 5.7.3 over an open surface results in a voltage (emf) induced in the loop closing the boundary $\mathcal{B}$ of the surface

$$\phi_{\text{induced}} = \int_{\mathcal{S}}(\nabla\times\boldsymbol{e})\cdot\hat{\mathbf{n}}\,d|\mathcal{S}| = \int_{\mathcal{B}}\boldsymbol{e}\cdot d\boldsymbol{l} \quad [\text{V}]. \qquad (5.7.3)$$

The emf (electromagnetic force) is the same as the Thévenin source voltage induced by the rate of change of the flux. Integrating Eq. 5.7.3 over the same open surface $\mathcal{S}$ results in the source

5.9.5 Speed of light

Problem # *5: The* speed of light *in vacuo is* $c_o = 1/\sqrt{\mu_o\epsilon_o}\approx 3\times10^8$ *[m/s]. The* characteristic resistance *in vacuo is* $r_o = \sqrt{\mu_o/\epsilon_o}\approx 377\,[\Omega]$.

– 5.1: Find a formula for the in-vacuo permittivity ϵ_o *and* permeability *in terms of* c_o *and* r_o. *Based on your formula, what are the numeric values of* ϵ_o *and* μ_o*?*

– 5.2: In a few words, identify the law given by this equation, define what it means, and explain the formula:

$$\int_{\mathcal{S}}\hat{\mathbf{n}}\cdot\boldsymbol{v}\,dA = \int_{\mathcal{V}}\nabla\cdot\boldsymbol{v}\,dV.$$

5.9.6 Application of Maxwell's equations

Problem # *6: The electric Maxwell equation is* $\nabla\times\boldsymbol{E}=-\dot{\boldsymbol{B}}$, *where* $\boldsymbol{E}$ *is the* electric field strength *and* $\dot{\boldsymbol{B}}$ *is the time rate of change of the* magnetic induction field, *or simply the magnetic flux density. Consider this equation integrated over a two-dimensional surface* S, *where* $\hat{\mathbf{n}}$ *is a unit vector normal to the surface (you may also find it useful to define the closed path* C *around the surface):*

$$\int_S[\nabla\times\boldsymbol{E}]\cdot\hat{\mathbf{n}}dS = -\frac{\partial}{\partial t}\int_S\boldsymbol{B}\cdot\hat{\mathbf{n}}dS.$$

– 6.1: Apply Stokes's theorem to the left-hand side of the equation.
Hint: view this relation in therms of the "integral forms" of the curl.

– 6.2: Consider the right-hand side of the equation. How is it related to the magnetic flux Ψ *through the surface* S*?*

– 6.3: Assume the right-hand side of the equation is zero. Can you relate your answer in question 6.1 to one of Kirchhoff's laws?

Problem # *7: The magnetic Maxwell equation is* $\nabla\times\boldsymbol{H}=\boldsymbol{C}\equiv\boldsymbol{J}+\dot{\boldsymbol{D}}$, *where* $\boldsymbol{H}$ *is the* magnetic field strength, $\boldsymbol{J}=\sigma\boldsymbol{E}$ *is the conductive (resistive) current density, and the* displacement current $\dot{\boldsymbol{D}}$ *is the time rate of change of the* electric flux density $\boldsymbol{D}$. *Here we defined a new variable* $\boldsymbol{C}$ *as the total current density.*

– 7.1: First consider the equation over a two-dimensional surface S*:*

$$\int_S[\nabla\times\boldsymbol{H}]\cdot\hat{\mathbf{n}}dS = \int_S[\boldsymbol{J}+\dot{\boldsymbol{D}}]\cdot\hat{\mathbf{n}}dS = \int_S\boldsymbol{C}\cdot\hat{\mathbf{n}}dS.$$

Then apply Stokes's theorem to the left-hand side of this equation. In a sentence or two, explain the meaning of the resulting equation. Hint: What is the right-hand side of the equation?

The definitions and generating potential functions of irrotational (conservative) and incompressible (solenoidal) fields naturally follow from two key vector identities: (1) $\nabla\cdot(\nabla\times\boldsymbol{w})=0$ and (2) $\nabla\times(\nabla\phi)=0$.

Problem # *3: Define the following:*

– 3.1: A conservative vector field

– 3.2: An irrotational vector field

– 3.3: An incompressible vector field

– 3.4: A solenoidal vector field

– 3.5: When is a conservative field irrotational?

– 3.6: When is an incompressible field irrotational?

Problem # *4: For each of the following, (i) compute $\nabla\cdot\boldsymbol{v}$, (ii) compute $\nabla\times\boldsymbol{v}$, and (iii) classify the vector field (e.g., conservative, irrotational, incompressible, etc.).*

– 4.1: $\boldsymbol{v}(x,y,z)=-\nabla(3yx^3+y\log(xy))$

– 4.2: $\boldsymbol{v}(x,y,z)=xy\hat{\mathbf{x}}-z\hat{\mathbf{y}}+f(z)\hat{\mathbf{z}}$

– 4.3: $\boldsymbol{v}(x,y,z)=\nabla\times(x\hat{\mathbf{x}}-z\hat{\mathbf{y}})$

5.9.4 Maxwell's Equations

The variables have the following names and defining equations (see Table 5.4, p. 265):

Symbol	Equation	Name	Units
$\boldsymbol{E}$	$\nabla\times\boldsymbol{E}=-\dot{\boldsymbol{B}}$	Electric field strength	[volts/m]
$\boldsymbol{D}=\epsilon_o\boldsymbol{E}$	$\nabla\cdot\boldsymbol{D}=\rho$	Electric displacement (flux density)	[coul/m^2]
$\boldsymbol{H}$	$\nabla\times\boldsymbol{H}=\boldsymbol{J}+\dot{\boldsymbol{D}}$	Magnetic field strength	[amps/m]
$\boldsymbol{B}=\mu_o\boldsymbol{H}$	$\nabla\cdot\boldsymbol{B}=0$	Magnetic induction (flux density)	[webers/m^2]

Note that $\boldsymbol{J}=\sigma\boldsymbol{E}$ is the *current density* (which has units of [amps/m^2]). Furthermore, the *speed of light in vacuo* is $c_o=3\times10^8=1/\sqrt{\mu_o\epsilon_o}$ [m/s], and the *characteristic resistance* of light $r_o=377=\sqrt{\mu_o/\epsilon_o}$ [Ω (i.e., ohms)].

of the induced voltage ϕ_{induced}, which is proportional to the rate of change of the flux [weber]:

$$\phi_{\text{induced}}=-\frac{\partial}{\partial t}\int_S \boldsymbol{b}\cdot\hat{\mathbf{n}}\,dA=L\dot{\psi}\quad\text{[Wb/s] or [V],}$$

where L [H] is the inductance of the wire, with impedance $Z_L(s)=sL$. The area integral on the right is the total flux crossing normal to the surface ψ [Wb]. The rate of change of the total flux [Wb/s] is the induced (Thévenin) voltage [V]. ∎

If we apply Gauss's theorem to the divergence equations, we find the total flux that crosses the closed surface.

Exercise #*20*

Apply Gauss's theorem to equation ED and explain what it means in physical terms.

Solution: The area of the normal component of $\boldsymbol{d}$ is equal to the volume integral over the charge density. Thus Gauss's theorem says that the total charge within the volume Q_{enc}, found by integrating the charge density $\rho(\boldsymbol{x})$ over the volume $\mathcal{V}$, is equal to the normal component of the flux $\boldsymbol{d}$ that crosses the surface $\mathcal{S}$:

$$Q_{\text{enc}}=\int_{\mathcal{V}}\nabla\cdot\boldsymbol{D}\,dV=\int_{\mathcal{S}}\boldsymbol{d}\cdot\hat{\boldsymbol{n}}\,dA.$$

When equal amounts of positive and negative charge exist within the volume, regardless of its distribution, the integral is zero. ∎

Summary: Maxwell's four equations relate the field strengths to the flux densities. There are two types of variables: field strengths $(\boldsymbol{e},\boldsymbol{h})$ and flux densities $(\boldsymbol{d},\boldsymbol{b})$. There are two classes: electric $(\boldsymbol{e},\boldsymbol{d})$ and magnetic $(\boldsymbol{h},\boldsymbol{b})$. This is a 2×2 matrix, with column being field strength and flux densities and rows being electric and magnetic variables.

	Strength	Flux density
Electric	$\boldsymbol{e}$ [V/m]	$\boldsymbol{d}$ [C/m^2]
Magnetic	$\boldsymbol{h}$ [A/m]	$\boldsymbol{b}$ [Wb/m^2]

Applying Stokes's curl theorem to the forces induces a Thévenin voltage (emf) or Norton current source. Applying Gauss's divergence theorem to the flows gives the total charge enclosed. The magnetic charge is zero ($\nabla\cdot\boldsymbol{b}=0$) because magnetic monopoles do not exist. However, magnetic dipoles do exist, as in the example of the electron, which contains a magnetic dipole.

5.7.2 Derivation of the vector wave equation

Next we provide the derivation of the vector wave equation starting from Maxwell's equations (Eq. 5.7.1), which is reminiscent of the derivation of the Webster horn equation (Eq. 5.2.2, p. 239).

Here we use the same 2x2 matix method we used in Eq 5.7.1, to both swap $\boldsymbol{b}=\mu_o\boldsymbol{h}$ and $\boldsymbol{d}=\epsilon_o\boldsymbol{e}$ and transform $\partial_t\leftrightarrow s$, due to the Laplace transform relation $\partial_t\leftrightarrow s=\sigma+j\omega$.

Working in the frequency domain and taking the curl of both sides gives

$$\begin{aligned}\nabla\times\nabla\times\begin{bmatrix}\boldsymbol{e}\\ \boldsymbol{h}\end{bmatrix}(\boldsymbol{x},t) &= \begin{bmatrix}0 & -s\mu_o\\ s\epsilon_o & 0\end{bmatrix}\nabla\times\begin{bmatrix}\boldsymbol{e}\\ \boldsymbol{h}\end{bmatrix}(\boldsymbol{x},t)\\ &= \begin{bmatrix}0 & -s\mu_o\\ s\epsilon_o & 0\end{bmatrix}\begin{bmatrix}0 & -s\mu_o\\ s\epsilon_o & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{e}\\ \boldsymbol{h}\end{bmatrix}\\ &= -\frac{s^2}{c_o^2}\begin{bmatrix}\boldsymbol{e}\\ \boldsymbol{h}\end{bmatrix}.\end{aligned}$$

Note that $s\mu_o s\epsilon_o = s^2/c_o^2$. If μ and ϵ were either functions of space or time, this transformation would not apply.

We must deal with the left hand side, which is $\nabla\times\nabla\times() = \boldsymbol{\nabla}^2() - \boldsymbol{\nabla}^2()$, which results in

$$\boldsymbol{\nabla}^2\begin{bmatrix}\mathbf{E}\\ \mathbf{H}\end{bmatrix} - \boldsymbol{\nabla}^2\begin{bmatrix}\mathbf{E}\\ \mathbf{H}\end{bmatrix} = \frac{s^2}{c_o^2}\begin{bmatrix}\mathbf{E}\\ \mathbf{H}\end{bmatrix}$$

Rearranging terms results in driven vector wave equations, for both $\boldsymbol{e}(\boldsymbol{x},t)$ and $\boldsymbol{h}(\boldsymbol{x},t)$

$$\begin{aligned}\boldsymbol{\nabla}^2\begin{bmatrix}\mathbf{E}\\ \mathbf{H}\end{bmatrix} - \frac{s^2}{c_o^2}\begin{bmatrix}\mathbf{E}\\ \mathbf{H}\end{bmatrix} &= \nabla\begin{matrix}\frac{1}{\epsilon_o}\nabla\cdot\mathbf{D}\\ \frac{1}{\mu_o}\cancelto{0}{\nabla\cdot\mathbf{B}}\end{matrix}\\ &= \frac{1}{\epsilon_o}\begin{bmatrix}\nabla\rho(\boldsymbol{x},s)\\ 0\end{bmatrix},\end{aligned} \tag{5.7.4}$$

with the charge term $\nabla\rho(\boldsymbol{x},s)$ being the sole source of excitation. Note that if μ and ϵ depended on $\boldsymbol{x}$, the terms on the right would not be zero.

In remote outer space, where black holes and hot plasmas frequently exist (e.g., inside the sun), this seems likely.

Recall the d'Alembert solutions of the scalar wave equation (Eq. 4.4.1, p. 188)

$$\boldsymbol{e}(\boldsymbol{x},t) = \boldsymbol{f}(\boldsymbol{x}-ct) + \boldsymbol{g}(\boldsymbol{x}+ct),$$

where $\boldsymbol{f}$ and $\boldsymbol{g}$ are arbitrary vector fields and c is the wave speed. As expected, the solutions are delayed by the speed of light.

Poynting vector: The EM power flux density $\boldsymbol{\mathcal{P}}$ [W/m^2] is perpendicular to $\boldsymbol{e}$ and $\boldsymbol{h}$, denoted as

$$\boldsymbol{\mathcal{P}} = \boldsymbol{e}\times\boldsymbol{h} \quad [\text{W/m}^2].$$

The corresponding EM momentum flux density $\boldsymbol{\mathcal{M}}$ (hence ME are related to mass, thus gravity) is

$$\boldsymbol{\mathcal{M}} = \boldsymbol{d}\times\boldsymbol{b} = \epsilon_o\mu_o\boldsymbol{e}\times\boldsymbol{h}.$$

Since the speed of light is $c_o = 1/\sqrt{\mu_o\epsilon_o}$, we see that

$$\boldsymbol{\mathcal{P}} = c_o^2\boldsymbol{\mathcal{M}} \quad [\text{W/m}^2],$$

which is related to Einstein's energy–mass equivalence formula $\mathcal{E} = mc_o^2$, as derived in a more general complex-analytic form on p. 122.

For example, the power emitted by the sun is about 1360 [W/m^2], with a radiation pressure of 4 [μN/m^2] (i.e., 4 [μPa]) (Fitzpatrick, 2008). By way of comparison, the threshold audible acoustic pressure at the human eardrum at 1 [kHz] is 20 [μPa], which is 14 [dB] (a factor of 5) the solar radiation pressure. Also:

– 2.3: Use the results from Eq. VC-2.1 to show that the solution to the spherical wave equation is

$$\nabla_r^2\,\varrho(r,t) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\,\varrho(r,t) \tag{VC-2.2}$$

$$\varrho(r,t) = \frac{f(t-r/c)}{r} + \frac{g(t+r/c)}{r}. \tag{VC-2.3}$$

– 2.4: Using $f(\xi) = \sin(\xi)u(\xi)$ and $g(\xi) = e^{\xi}u(\xi)$, write the solutions to the spherical wave equation, where $u(\xi)$ is the Heaviside step function.

– 2.5: Sketch this $f(\xi)$ and $g(\xi)$ for several times (e.g., 0, 1, and 2 seconds), and describe the behavior of the pressure $\varrho(r,t)$ as a function of time t and radius r.

– 2.6: What happens when the inbound wave reaches the center at $r = 0$?

5.9.3 Helmholtz's formula

Every differentiable vector field may be written as the sum of a scalar potential ϕ and a vector potential $\boldsymbol{w}$. This relationship is best known as the fundamental theorem of vector calculus (also called Helmholtz's formula):

$$\boldsymbol{v} = -\nabla\phi + \nabla\times\boldsymbol{w}. \tag{VC-2.4}$$

This formula seems to be a natural extension of the algebraic products $\boldsymbol{a}\cdot\boldsymbol{b} \perp \boldsymbol{a}\times\boldsymbol{b}$, since $\boldsymbol{a}\cdot\boldsymbol{b} \propto \|\boldsymbol{a}\|\|\boldsymbol{b}\|\cos(\theta)$ and $\boldsymbol{a}\times\boldsymbol{b} \propto \|\boldsymbol{a}\|\|\boldsymbol{b}\|\sin(\theta)$, as developed in Appendix A.3.1, page 286. Thus these orthogonal components have magnitude 1 when we take the norm, due to Euler's identity $(\cos^2(\theta)+\sin^2(\theta) = 1)$.

As shown in Table 5.1 (p. 225), Helmholtz's formula separates a vector field (i.e., $\mathbf{v}(\mathbf{x})$) into compressible and rotational parts:

1. The rotational (e.g., angular) part is defined by the vector potential $\boldsymbol{w}$, which requires that $\nabla\times\nabla\times\boldsymbol{w} \neq 0$. A field is irrotational (conservative) when $\nabla\times\boldsymbol{v} = 0$, meaning that the field $\boldsymbol{v}$ can be generated using only a scalar potential, $\boldsymbol{v} = \nabla\phi$ (note that this is how a conservative field is usually defined, by saying there exists some ϕ such that $\boldsymbol{v} = \nabla\phi$).[18]

2. The compressible (e.g., radial) part of a field is defined by the scalar potential ϕ, which requires that $\nabla\cdot\nabla\phi = \nabla^2\phi \neq 0$. A field is incompressible (solenoidal) when $\nabla\cdot\boldsymbol{v} = 0$, meaning that the field $\boldsymbol{v}$ can be generated using only a vector potential, $\boldsymbol{v} = \nabla\times\boldsymbol{w}$.

[18] A note about the relationship between the generating function and the test: You might imagine special cases where $\nabla\times\boldsymbol{w} \neq 0$ but $\nabla\times\nabla\times\boldsymbol{w} = 0$ (or $\nabla\phi \neq 0$ but $\nabla^2\phi = 0$). In these cases, the vector (or scalar) potential can be recast as a scalar (or vector) potential. For example, consider a field $\boldsymbol{v} = \nabla\phi_0 + \boldsymbol{b}$, where $\boldsymbol{b} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$. Note that $\boldsymbol{b}$ can actually be generated by either a scalar potential ($\phi_1 = \frac{1}{2}[x^2+y^2+z^2]$, such that $\nabla\phi_1 = \boldsymbol{b}$) or a vector potential ($\boldsymbol{w}_0 = \frac{1}{2}[z^2\hat{\mathbf{x}} + x^2\hat{\mathbf{y}} + y^2\hat{\mathbf{z}}]$, such that $\nabla\times\boldsymbol{w}_0 = \boldsymbol{b}$). We find that $\nabla\times\boldsymbol{v} = 0$; therefore $\boldsymbol{v}$ must be irrotational. We say this irrotational field is generated by $\nabla\phi = \nabla(\phi_0+\phi_1)$.

5.9 Problems VC-2

5.9.1 Topics of this homework:

Partial differential equations; fundamental theorem of vector calculus (Helmholtz's theorem); wave equation; Maxwell's equations (ME) and variables ($\boldsymbol{E}, \boldsymbol{D}; \boldsymbol{B}, \boldsymbol{H}$); Second-order vector differentials; Webster horn equation.

Notation: The following notation is used in this homework:

1. $s = \sigma + j\omega$ is the Laplace frequency, as used in the Laplace transform.
2. A Laplace transform pair is indicated by the symbol $\leftrightarrow$: for example, $f(t) \leftrightarrow F(s)$.
3. Here π_k is the kth prime: $\pi_k \in \mathbb{P}$, $\pi_k = [2, 3, 5, 7, 11, 13, \ldots]$.

5.9.2 Partial differential equations (PDEs): Wave equation

Problem # *1: Solve the wave equation in one dimension by defining $\xi = t \mp x/c$.*

– 1.1: Show that d'Alembert's solution, $\varrho(x,t) = f(t - x/c) + g(t + x/c)$, is a solution to the acoustic pressure wave equation in one dimension:

$$\frac{\partial^2 \varrho(x,t)}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 \varrho(x,t)}{\partial t^2},$$

where $f(\xi)$ and $g(\xi)$ are arbitrary functions.

Problem # *2: Solving the wave equation in spherical coordinates (i.e., three dimensions)*

– 2.1: Write the wave equation in spherical coordinates $\varrho(r, \theta, \phi, t)$. Consider only the radial term r (i.e., dependence on angles θ and ϕ is assumed to be zero). Hint: The form of the Laplacian as a function of the number of dimensions is given in Eq. 5.1.9 (page 229). Alternatively, look it up on the internet or in a calculus book.

– 2.2: Show that this equation is true:

$$\nabla_r^2 \varrho(r) \equiv \frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r}\varrho(r) = \frac{1}{r}\frac{\partial^2}{\partial r^2} r\varrho(r). \qquad \text{(VC-2.1)}$$

Hint: Expand both sides of the equation.

> *The lasers used in Inertial Confinement Fusion (e.g., the NOVA experiment in Lawrence Livermore National Laboratory) typically have energy fluxes of 10^{18} [W/m^2]. This translates to a radiation pressure of about 10^4 atmospheres!*
>
> *–Fitzpatrick (2008, p. 291)* One [atm] is 10^5 [Pa].

Electrical impedance seen by an electron: Up to now we have considered only the Brune impedance, which is a special case with no branch points or branch cuts. We can define *impedance* for diffusion, as in the diffusion of heat. There is also the diffusion of electrical and magnetic fields at the surface of a conductor, where the resistance of the conductor dominates the dielectric properties. This is called the electrical *skin effect*, where the conduction currents are dominated by the conductivity of the metal rather than the displacement currents. In such cases, the impedance is proportional to $\sqrt{s}$, which requires that it has a branch cut. Still, the real part of the impedance must be positive in the right s half-plane, the required condition of all impedances, such that Postulate P3 (p. 162) is satisfied. The same effect is observed in acoustics (see Appendix D).

When we deal with Maxwell's equations, the force is defined by the Lorentz force,

$$\boldsymbol{f} = q\boldsymbol{e} + q\nu \times \boldsymbol{b} = q\boldsymbol{e} + \boldsymbol{c} \times \boldsymbol{b},$$

which is the force on a charge (e.g., electron) due to the electric $\boldsymbol{e}$ and magnetic $\boldsymbol{b}$ fields. The magnetic field plays a role when the charge has a velocity ν. When a charge is moving with velocity ν, it may be viewed as a current $\boldsymbol{c} = q\nu$ (see the discussion on p. 161).

The complex admittance density is

$$Y(s) = \sigma_o + s\epsilon_o \quad [\text{℧/m}^2]$$

(Feynman, 1970b, p 13–1). Here σ_o is the electrical conductivity and ϵ_o is the electrical permittivity. Since $\omega\epsilon_o \ll \sigma_o$, the conductivity reduces to the resistance of the wire per unit length.[16]

5.8 Potential solutions of Maxwell's equations

One primary reason for using potentials is to generate solutions to Maxwell's equations. For example, if we extend Eq. 5.1.1 (p. 224), we can express Maxwell's equations in terms of scalar and vector potentials. These relationships are (Sommerfeld (1952, p. 146), Feynman (1970d, p. 18-10)):

$$\boldsymbol{e}(\boldsymbol{x},t) = -\nabla\phi(\boldsymbol{x},t) - \frac{\partial \boldsymbol{a}(\boldsymbol{x},t)}{\partial t} \quad \text{[V/m]} \qquad (5.8.1)$$

and

$$\boldsymbol{h}(\boldsymbol{x},t) = \frac{1}{\mu_o}\left[\nabla \times \boldsymbol{a}(\boldsymbol{x},t) + \frac{\partial \boldsymbol{d}(\boldsymbol{x},t)}{\partial t}\right] \quad \text{[A/m]}. \qquad (5.8.2)$$

We have extended $\boldsymbol{h}(\boldsymbol{x},t)$ to include the electric potential term

$$\boldsymbol{d}(\boldsymbol{x},t) = \epsilon(\boldsymbol{x},t)\boldsymbol{e}(\boldsymbol{x},t) = -\epsilon(\boldsymbol{x},t)\nabla\phi(\boldsymbol{x},t),$$

[16] For copper $\omega \ll \omega_c = \sigma_o/\epsilon_o \approx 6\times10^7/9\times10^{-12} \approx 6.66\times10^{18}$ [rad/s], or $f_c = 10^{18}$ [Hz]. This corresponds to a wavelength of $\lambda_o \approx c_o/f_c = 0.30\times10^{-9}$ [m]. For comparison, the Bohr radius (hydrogen) is $\approx 0.053\times10^{-9}$ [m] (5.66 times smaller) and the Lorentz radius (of the electron) is estimated to be $\mu_o = 2.8 \times 10^{-15}$ [m] (2.8 [femtometers]).

normally taken to be zero because taking the curl of $\boldsymbol{h}(t)$ naturally removes any electrical potential term due to CoG.[17] The extension makes the potential solutions symmetric so that $\boldsymbol{e}$ and $\boldsymbol{h}$ each have electrical and magnetic excitation.

Exercise #*21*

Explain why some dependence on $\phi(\boldsymbol{x},t)$ do not appear in Eq. 5.8.2 but do in 5.8.1.

Solution: For $\boldsymbol{h}(\boldsymbol{x},t)$ to depend on $\phi(x,t)$ it must appear through the electric strength, as $\boldsymbol{e}(x,t) = -\nabla\Phi(x,t)$. But then $\nabla\times\boldsymbol{h}(x,t)$ would mean applying CoG (i.e., $\nabla\times\nabla\phi = 0$) on the right side of the equation. Since this term would be zero; it is assumed to be zero, thus $\boldsymbol{h}(x,t)$ dependents on only $\boldsymbol{a}(x,t)$. To fill out the symmetry, we have added $\partial_t\boldsymbol{d}(\boldsymbol{x},t)$ to Eq. 5.8.2, to see what might happen in a more general case. ■

Use of Helmholtz's theorem on potential solutions: The generalized solutions to Maxwell's equations (Eqs. 5.8.1 and 5.8.2) have been expressed in terms of EM potentials $\phi(\boldsymbol{x})$ and $\boldsymbol{a}(\boldsymbol{x})$ and Helmholtz's theorem. These are solutions to Maxwell's equations expressed in terms of the potentials $\phi(\boldsymbol{x},s)$ and $\boldsymbol{a}(\boldsymbol{x},s)$, as determined at the boundaries (Sommerfeld, 1952, p 146). These relationships are invariant to certain functions added to each potential, as shown below. They are equivalent to Maxwell's equations following the application of $\nabla\cdot$ and $\nabla\times$.

Next we show that the potential equations (Eqs. 5.8.1 and 5.8.2) are consistent with Maxwell's equations (Eq. 5.7.1).

ME for $\boldsymbol{e}(\boldsymbol{x},t)$: Taking the curl of Eq. 5.8.1, applying CoG $= 0$, and using Eq. 5.8.2, we find that

$$\begin{aligned}\nabla\times\boldsymbol{e} &= -\cancel{\nabla\times\nabla}\cancel{\Phi}^{\,0} - \nabla\times\frac{\partial\boldsymbol{a}}{\partial t}\\ &= -\frac{\partial\boldsymbol{b}}{\partial t}\end{aligned} \tag{5.8.3}$$

recovers Maxwell's equation for $\boldsymbol{e}(\boldsymbol{x})$ (Eq. 5.7.1).

Taking the divergence of Eq. 5.8.2 and applying DoC $= 0$ give Eq. 5.7.2 for $\boldsymbol{b}(\boldsymbol{x})$:

$$\nabla\cdot\boldsymbol{b}(\boldsymbol{x}) = \cancel{\nabla\cdot\nabla\times\boldsymbol{a}(\boldsymbol{x})}^{\,0} = 0.$$

ME for $\boldsymbol{h}(\boldsymbol{x},t)$: To recover Maxwell's equation for $\boldsymbol{h}(\boldsymbol{x})$ (Eq. 5.7.1, $\nabla\times\boldsymbol{h} = \boldsymbol{c}$) from the potential equation (Eq. 5.8.2), we take the curl and use $\boldsymbol{b} = \epsilon_o\boldsymbol{h}$ (Table 5.4, p. 265):

$$\begin{aligned}\nabla\times\boldsymbol{b}(\boldsymbol{x}) &= \mu_o\nabla\times\boldsymbol{h}(\boldsymbol{x})\\ &= \nabla\times\nabla\times\boldsymbol{a}(\boldsymbol{x})\\ &= \nabla^2\boldsymbol{a}(\boldsymbol{x},t) - \boldsymbol{\nabla}^2\boldsymbol{a}(\boldsymbol{x},t)\\ &= \nabla\nabla\cdot\boldsymbol{a}(\boldsymbol{x},t) - \frac{1}{c_o^2}\frac{\partial^2}{\partial_t^2}\boldsymbol{a}(\boldsymbol{x},t)\\ &= -\frac{1}{c^2}\left(\ddot{\boldsymbol{a}} + \nabla\dot{\Phi}\right) + \mu_o\mathbf{J}.\end{aligned}$$

[17] In-vacuo $\epsilon_o = 8.85\times10^{-12}$ [F/m^2] is the capacitance, and $s\epsilon_o$ is the electric compliance-density of light. The related magnetic mass-density is the permeability $\mu_o = 4\pi\times10^{-7}$ [H/m^2] having an inductive impedance of $s\mu_o$ [Ω/m]. It is helpful to think of ϵ_o as a capacitance per unit area and μ_o as an inductance per unit area (consistent with their units). The in-vacuo speed of light is $c_o = 1/\sqrt{\epsilon_o\mu_o} = 3\times10^8$ [m/s], but is slower when traveling in matter (Brillouin, 1960).

This last equation may be split into two independent equations by the use of Helmholtz's theorem:

$$\boldsymbol{\nabla}^2\boldsymbol{a} - \frac{1}{c_o^2}\ddot{\boldsymbol{a}} = -\mu_o\mathbf{J} \quad\text{and}\quad \nabla\cdot\boldsymbol{a} + \frac{1}{c_o^2}\ddot{\Phi} = 0.$$

Taking the divergence of Eq. 5.8.2 and applying DoC = 0 gives Eq. 5.7.2 ($\nabla\cdot\boldsymbol{d} = -\dot{\rho}$). Alternatively,

$$\nabla^2\Phi - \frac{1}{c_o^2}\ddot{\Phi} = -\frac{\rho}{\epsilon_o},$$

which is the scalar potential wave equation driven by the charge (Sommerfeld, 1952, p 146).

Summary: In conclusion, Eq. 5.8.1, along with DoC $= 0$ and CoG $= 0$, gives Maxwell's Eqs. 5.7.1 and 5.7.2 for $\boldsymbol{e}$. Likewise, Eq. 5.8.2, along with DoC $= 0$ and CoG $= 0$, gives Maxwell's Eqs. 5.7.1 and 5.7.2 for $\boldsymbol{h}$. The above derivation for $H(x,t)$ from A and Φ derives the magnetic component of the field, expressed in terms of its vector potential, in the same way as Eq. 5.7.1 describes $\boldsymbol{e}(x,t)$ in terms of the potentials.

We may view the potential equations (Eqs 5.8.1 and 5.8.2) as equivalent to Maxwell's equations; thus they are the solutions to ME in terms of potentials.

Exercise #*22*

Starting with the values of the speed of light $c_o = 3\times10^8$ [m/s] and the characteristic resistance of light waves $r_o = 377$ [ohms], use the formulas $c_0 = 1/\sqrt{\mu_o\epsilon_o}$ and $r_o = \sqrt{\mu_o/\epsilon_o}$ to find values for ϵ_o and μ_o.

Solution: Squaring $c_o^2 = 1/\mu_o\epsilon_o$ and $r_o^2 = \mu_o/\epsilon_o$, we may solve for the two unknowns: $c_o^2 r_o^2 = \frac{1}{\cancel{\mu_o}\epsilon_o}\cancel{\mu_o}\epsilon_o = 1/\epsilon_o^2$; thus $\epsilon_o = 1/c_o r_o = 10^{-8}/2.998\cdot377 \approx 8.85\times10^{-12}$ [F/m]. Likewise, $\mu_o = r_o/c_o = (377/2.998)\times10^{-8} \approx 125.75\times10^{-8}$ [H/m]. The value of μ_o is defined in international SI units as $4\pi\times10^{-7} \approx 12.566\times10^{-7}$ [H/m].

In conclusion, it is more productive to memorize c_0 and r_0, from which ϵ_o and μ_o may be easily derived. ■

Exercise #*23*

Starting from Eq. 5.7.1, with $\boldsymbol{j_m} = \sigma_o\boldsymbol{e}$, Maxwell's equation, including the magnetic intensity, is

$$\nabla\times\boldsymbol{h}(\boldsymbol{x},t) = \boldsymbol{j_m}(\boldsymbol{x},t) + \frac{\partial}{\partial t}\boldsymbol{d}(\boldsymbol{x},t).$$

Find the equation for the magnetically induced current $\boldsymbol{j_m}(\boldsymbol{x},t)$.

Solution: The divergence of the curl is zero (DoC $= 0$),

$$\cancel{\nabla\cdot\nabla\times\boldsymbol{h}}^{\,0}(\boldsymbol{x},t) = \nabla\cdot\boldsymbol{j_m}(\boldsymbol{x},t) + \frac{\partial}{\partial t}\rho(\boldsymbol{x},t) = 0, \tag{5.8.4}$$

which is conservation of charge (i.e., Gauss's theorem). ■