

University of New Mexico

UNM Digital Repository

Electrical and Computer Engineering ETDs

Engineering ETDs

Fall 12-4-2020

Theoretical and Experimental Studies of the Emission of Electromagnetic Radiation by Superluminal Polarization Currents

Andrea Caroline Schmidt-Zweifel
University of New Mexico

Follow this and additional works at: https://digitalrepository.unm.edu/ece_etds



Part of the [Electrical and Computer Engineering Commons](#)

Recommended Citation

Schmidt-Zweifel, Andrea Caroline. "Theoretical and Experimental Studies of the Emission of Electromagnetic Radiation by Superluminal Polarization Currents." (2020).
https://digitalrepository.unm.edu/ece_etds/497

This Dissertation is brought to you for free and open access by the Engineering ETDs at UNM Digital Repository. It has been accepted for inclusion in Electrical and Computer Engineering ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.

Andrea Caroline Schmidt-Zweifel

Candidate

Engineering

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Prof. Edl Schamiloglu

, Chairperson

Prof. Christos Christodoulou

Prof. Zhen Peng

Prof. John Singleton

Dr. Bruce Carlsten

Theoretical and Experimental Studies of the Emission of Electromagnetic Radiation by Superluminal Polarization Currents

by

Andrea Caroline Schmidt-Zweifel

B.S., Mathematics, University of New Mexico, 2006

M.S., Mathematics, University of New Mexico, 2012

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Engineering

The University of New Mexico

Albuquerque, New Mexico

December, 2020

©2020, Andrea Caroline Schmidt-Zweifel

Dedication

This work is dedicated to my mother, Ludmilla Zweifel-Thury, my brother, Martin Zweifel, and the memory of my father, Friedrich Zweifel.

I've seen things you people wouldn't believe. Attack ships on fire off the shoulder of Orion. I watched C-beams glitter in the dark near the Tannhäuser Gate. All those moments will be lost in time, like... tears in rain.

— Tears in Rain Monologue, *Blade Runner*

Acknowledgments

Firstly, I should like to thank my supervisor and committee chair Professor Edl Schamiloglu, who provided academic guidance and unwavering support for many years, during good times and difficult ones. I will always be deeply grateful for all that he has taught and given me. My thanks go also to the members who graciously agreed to serve on my committee: Professors Christos Christodoulou, Zhen Peng and John Singleton as well as Dr. Bruce Carlsten. I much appreciate their patience, humor and advice. John has been my mentor, guide star and, especially, friend ever since I started working at Los Alamos National Laboratory. Without his inspiration and innate enthusiasm for our research I would not have overcome the challenges of this long endeavor.

In my daily work I have been blessed with a group of outstandingly knowledgeable colleagues. Special thanks to Joe Fasel for his mentorship, counsel and insistence on sound programming habits, and to John Middleditch for teaching me the ways of the universe. Collective and individual acknowledgements are also due to my coworkers at UNM and Los Alamos National Laboratory, David Bizzozero, Frank Krawczyk, Quinn Marksteiner, Bill Romero, and Pinaki Sengupta. James Wigger, Connor Bailey, Simon Redman and Helen Lu deserve my thanks for assistance with the experiments and for carrying out calculations using my analytical and numerical routines. As Isaac Newton wrote in a 1675 letter to fellow scientist Robert Hooke, “If I have seen further, it is by standing on the shoulders of giants.”

Special thanks to Marie and Richard for their good-natured forbearance with the (very) long process of researching and writing this work, and for their pride in the finished product. And to Jürgen, who never stopped believing. I am also deeply grateful to Janine Tulenko and Ted Mockler, who aided this work by offering help, providing meals and humoring my eccentricities. Better friends cannot be found.

Lastly, I thank the Los Alamos LDRD program for funding almost all of the work reported in this dissertation.

Theoretical and Experimental Studies of the Emission of Electromagnetic Radiation by Superluminal Polarization Currents

by

Andrea Caroline Schmidt-Zweifel

B.S., Mathematics, University of New Mexico, 2006

M.S., Mathematics, University of New Mexico, 2012

PhD, Engineering, University of New Mexico, 2020

Abstract

Maxwell's equations establish that patterns of electric charges and currents can be animated to travel faster than the speed of light *in vacuo*, c , and that these superluminal distribution patterns emit tightly focused packets of electromagnetic radiation that are fundamentally different from the emissions by previously known sources – on Earth and in the Universe. This dissertation introduces a practical faster-than-light emitter: *The superluminal polarization current antenna*. Such devices use a polarization current that travels faster than c to give rise to electromagnetic radiation, a technique known as the *vacuum Čerenkov effect*. In what is to follow, the theoretical and mathematical foundations of this emission mechanism are laid, competing theories are critically examined, and experimental data from the antennas are compared with electromagnetic models.

The dissertation is organized as follows. After a brief introduction in Chapter 1, Chapter 2 chronicles the historical events that led to the realization of the vacuum

Čerenkov effect. Analogous phenomena observed in surface water waves and acoustics are described, culminating in the discovery and subsequent understanding of the sonic boom. In parallel, summaries are given of theoretical electromagnetic work on faster-than-light sources by luminaries such as Sommerfeld and Schott, along with an account of Sergey Vavilov and Pavel Čerenkov’s contributions. Pulsars may well be natural superluminal emitters, and so are briefly touched on. Chapter 3 covers the practical circular and linear machines (or “technology demonstrators”) that provide the experimental data against which models derived later in this work can be tested; an account is given of the measurement procedures followed and examples of typical results plotted. Chapter 4 contains a detailed examination and verification of the mathematical principles that underpin calculations of the electromagnetic emissions from superluminal polarization currents. The complications arising from a source that overtakes the waves that it generates are scrutinized. As a new discipline begins, unusual and erroneous ideas may well crop up; Chapter 5 examines why a competing theory of faster-than-light sources generates spurious results and provides a reasoned critique of the errors. As a stepping stone towards full theoretical models of some of the practical antennas, Chapter 6 treats small, point-like charges performing circular orbits at faster-than-light speeds. Potential pathological effects are found to be unimportant once finite wavelengths are introduced. The applicability of such “point source” models to pulsars is examined. Analytical electromagnetic frameworks of linear, circular and arced superluminal emitters are given in Chapter 7; the final expressions are evaluated numerically, with special attention being paid to differentiation routines that perform well out to extremely large distances from the source. The simulations are found to reproduce the experimental data well, and in a quantitative fashion. A desirable property of superluminal antennas is that they can potentially be made in almost any shape, uniquely optimized for a particular application; Chapter 8 describes a numerical paradigm – the discontinuous nodal Galerkin method – that will underpin the eventual topological optimization of the dielectric that hosts the polarization current.

Contents

1	Introduction	1
2	History	3
2.1	First Prologue	3
2.2	ἐπὶ οἴνοπα πόντον	4
2.3	The Sound Barrier: Man's Conquest of Heaven	7
2.3.1	Supersonic Ballistics	7
2.3.2	The Right Stuff	12
2.4	The Discovery of Light	20
2.4.1	Dr. Einstein's Universal Speed Limit	20
2.4.2	A Tale of Two Fields	27
2.4.3	A Necessary Word on Notation (Act I)	30
2.4.4	Ampère's Missing Term	32
2.5	Ad Astra	36
2.5.1	Pulsars 101	38

2.5.2	The Physics of Neutron Stars	40
2.5.3	Beyond the Light Cylinder	43
2.6	“... a waste of tax-payers’ money”	48
3	Meet the Lightslingers: Superluminal Sources in the Laboratory	50
3.1	Second Prologue	50
3.2	Fundamental Principles	53
3.3	Circular Accelerator TD 1	55
3.4	Linear Accelerator TD 2	58
3.5	Anechoic Chamber and Test Measurements	59
3.5.1	Angular Scans	59
3.5.2	Detecting the Radiation	60
3.5.3	Into the Wild	62
3.5.4	Power and Angle Corrections	66
3.5.5	Emitted Frequencies and Driving Voltages	69
3.6	Representative Results: Linear Antenna TD 2	71
3.6.1	Source Speed and Azimuthal Distribution of Radiation	71
3.6.2	A Qualitative Understanding of the Line Shape	73
3.6.3	Power Versus Speed	74
3.6.4	Polar Power Dependence	75
3.7	Range tests of TD 1	77

4	Sources That Move Faster Than Light and How to Treat Them	81
4.1	Third Prologue	81
4.1.1	The Consequences of a Vexing Problem	81
4.1.2	Some Points of Principle	86
4.2	The Beauty of Properly Posed Conditions	96
4.2.1	On Characteristics, Singularities and the Classification of Second Order Partial Differential Equations	96
4.2.2	Sofia Kovalevskaya's Solitary Quest	110
4.2.3	Time- and Space-like Solutions	112
4.3	Mr. Green's Solution to Cauchy's Problem	120
4.3.1	A Necessary Word on Notation (Act II)	120
4.3.2	The Fundamental Properties of the Green's Function	124
4.3.3	Apropos the Wave Equation I: General Case	130
4.3.4	Apropos the Wave Equation II: Free Space	135
4.3.5	At Last: The Explicit Form of the Green's Function	139
5	The Curious Case of the Nonspherically Decaying Radiation Field	142
5.1	The Case	142
5.2	The Evidence	144
5.3	The Verdict	157
6	Point Source Models: The Little Engine That Could	158

6.1	Fourth Prologue	158
6.2	On the Anatomy of a Charge in Superluminal Rotation	160
6.2.1	Geometrical Considerations and the Function $h(t)$	160
6.2.2	Graphing $h(t)$ in Three Dimensions	168
6.2.3	Relationship of $h(t)$ to Catastrophe Theory and Optics	170
6.2.4	In the Far-field	175
6.2.5	Wavelength Scaling Behavior of the $h(t)$ “Catastrophe”	179
6.3	All About Temporal Focusing	181
6.3.1	General Idea	181
6.3.2	Definition of a Practical “Temporal Focusing Factor”	181
6.4	In Search of the Point-Source Response	186
6.5	The Relation between Emission and Observer Time: An Old Problem Revisited	191
6.5.1	An Iterative Solution to Kepler’s Equation	193
6.6	From Point Source to Wavepacket	198
6.6.1	Concept	198
6.6.2	Experimental Demonstration	205
6.6.3	Experimental Results	208
6.6.4	Implications for Pulsars	210
6.7	Results from this Chapter Applied to Pulsars	212
6.7.1	Fields Predicted by the Liénard-Wiechert Expressions	212

6.7.2	Stokes Parameters	214
6.7.3	Intrinsic Characteristics of the Radiation	218
6.7.4	Frequency Content of the Radiation	219
6.7.5	Summary	223
7	Green's Functions at Work	225
7.1	Fifth Prologue	225
7.2	The Secret Lives of Potentials	226
7.3	Magnetic Vector Potentials: Linear and Quasi-Linear Superluminal Antennas	230
7.3.1	Constant Motion	230
7.3.2	Implementing Acceleration I	233
7.3.3	Implementing Acceleration II	235
7.3.4	Arced Antenna	236
7.3.5	Other Non-Circular Cases	237
7.4	Magnetic Vector Potentials: Circular Superluminal Antenna	238
7.4.1	Simulation of TD 1	238
7.4.2	Circular Antenna with Unidirectional Polarization	240
7.5	Turning Potentials into Fields	241
7.5.1	Much Needed Vector Algebra	241
7.5.2	Numerical Methods for the Derivatives	243
7.6	Results from Calculations	245

7.6.1	Input Parameters	245
7.6.2	Linear Antenna: Superluminal Source Speed $v/c > 2$	246
7.6.3	Linear Antenna: Subluminal Source Speeds	248
7.6.4	Circular Superluminal Antenna	250
7.6.5	Using the Model to Explore Antenna Design	251
7.6.6	The Model as a Diagnostic Tool	255
7.6.7	Non-Spherical Decay Exponents	257
7.7	Focusing Effects: Linear Antenna	260
7.7.1	Formalism	260
7.7.2	Results and Potential Applications	262
7.8	Focusing Effects: Circular Superluminal Sources	264
7.8.1	Formalism	264
7.8.2	Results	266
7.8.3	Relevance to Pulsars and Speckle Patterns	267
8	Epilogue	269
8.1	Current Work	269
8.2	Back to the Future...	273
	Appendices	275
A	Equations (6.4) <i>et seq.</i> in Metrical Units	276

B	Circular Antenna TD 1: Design Details	277
B.1	Dielectric Element Design for TD 1	277
B.2	Performance of Antenna Element Wedges.	282
B.3	Control Electronics for TD 1	283
B.4	Labview Control Software.	287
B.5	Circuit Boards and Final Assembly	289
C	Linear Superluminal Antenna TD 2	293
C.1	Design Summary	293
C.2	Retardation and Field Confinement for $v/c < 2$	294
C.2.1	The Problem	295
C.2.2	Retardation	296
C.2.3	Confinement	297
C.2.4	Active Phase Control at Slow Speeds.	298
C.2.5	Modelling the Effects	300
	References	301

Chapter 1

Introduction

“Would you tell me, please, which way I ought to go from here?”

“That depends a great deal on where you want to get to,” said the cat.

“I don’t much care where –” said Alice.

“Then it doesn’t matter which way you go,” said the Cat.

“– so long as I get somewhere,” Alice added as an explanation.

“Oh, you’re sure to do that,” said the Cat, “if you only walk long enough.”

— Lewis Carroll, *Alice in Wonderland*

This is believed to be the first doctoral dissertation — certainly the first in the Western world – to focus on *superluminal polarization-current antennas*. As their name suggests, these antennas employ a polarization current that travels faster than the speed of light *in vacuo* to emit electromagnetic radiation; this mechanism is known as the *vacuum Čerenkov effect*.

As is the case in any new field, there is precious little precedent and much ground must be covered. Historical inquiries such as remarkable phenomena in water waves and acoustics caused by sources that travel faster than the wave speed are to be studied. Then the unconventional aspects of the new antennas themselves

must be treated; periodic structures of electrical conductors acting as point- or line-like sources in conventional phased arrays are replaced by an extended *volume source* that inhabits *an electrical insulator*. As there are no textbooks on such devices and only a handful of (somewhat lackluster) published accounts, there is much to describe.

Furthermore, it is necessary to establish whether the same mathematical techniques that underpin calculations of the electromagnetic emissions from conventional, subluminal emitters can be applied to sources that overtake the light that they give out? To answer this, one needs to re-examine the very foundations of 19th Century electromagnetism that are accepted without question or brushed under the carpet in modern texts. Canonical approaches such as Green's function treatments and the evaluation of Liénard-Wiechert potentials must be scrutinized carefully and, where necessary, re-derived and generalized. Assuming that such techniques *can* be adapted to the new antennas, do the resulting models provide quantitative descriptions of experimental data? After much work, cautious optimism prevails.

In the early days of a new discipline, unusual and sometimes erroneous ideas inevitably crop up; these should be tested and, if need be, discredited very carefully. Acrimonious disputes amongst others must be sifted to extract the valid points. However, as this field represents almost unexplored scientific territory, there may also be unexpected rewards; for instance vital clues towards an understanding of astronomical objects such as pulsars are slowly emerging as a byproduct of modeling superluminal antennas.

Finally, do these antennas have a future? One of their attractive properties is that they can potentially be shaped into almost any form, uniquely optimized for a particular application. Some first steps towards eventual topological optimization of the new antennas are examined.

All of this – and more! – is covered in the pages that follow.

Chapter 2

History

2.1 First Prologue

Radiation sources that travel faster than the electromagnetic waves they emit are a comparatively recent discovery. However, human knowledge of similar effects involving water and sound waves extends back thousands of years, from the battle ships of the Bronze-age Greeks to Lord Kelvin's studies of wake patterns and from Mach's ballistics experiments to the test pilots who conquered the sound barrier. (The latter serve as a stark reminder that scientific advances are oftentimes not born on the desks of theoreticians but, quite literally, from the blood and sweat of those who are willing to give their lives for them.) Our understanding of the novel antennas to be described later in this work is underpinned by the above historical advances, so we shall recount them in this chapter.

Regardless of the medium, the fact that the source outdistances the generated disturbances leads to interesting (and sometimes quite dramatic) focusing and collimation effects. The latter arise when waves, emitted at different points along the source's path, intersect at a distant observer's position to create an intense burst. It

is not surprising that Huygens wavelet constructs – named after the Dutch physicist, mathematician, astronomer and inventor Christiaan Huygens (1629-1695) – were used by aeronautical engineers in the first descriptions of *temporal focusing* (that is focusing of waves in the time domain) in the 1950s. In the second half of this chapter, these ideas are applied to electromagnetic radiation. There, the initial discoveries were of Čerenkov radiation, emitted by charged particles that travel faster than the speed of light in a medium such as water. Colleagues of the Russian discoverers of this mechanism were soon predicting vacuum Čerenkov radiation, *i.e.*, radiation emitted by a source traveling faster than the speed of light *in a vacuum*. However, as we shall see briefly in Section 2.6 and much more fulsomely in Chapter 3, practical sources of this nature took several more decades to emerge. A pivotal ingredient was the discovery of *polarization currents*, and so a description of their part in Maxwell’s derivation of his famous quartet of equations will be given. Finally, the last section gives a brief description of natural superluminal antennas – pulsars! – to which some of the calculations in Chapters 4 – 7 are also relevant.

2.2 ἐπὶ οὔνοπα πόντον

A source that moves uniformly through a homogeneous medium emits focused¹ wave fronts if the velocity of the source exceeds that of the propagating disturbances themselves [1]. This fact – if not its theoretical underpinnings – has been known to mariners around the globe for more than five millennia: The invention of the sail made clear that a ship which travels across the sea faster than the wave speed excites easily observable patterns in the water (Fig. 2.1). That ancient seafaring nations were well aware of bow waves and wakes is evidenced by many historical accounts reaching back as far as Homer’s *Iliad* [2], whose written version is usually dated to around

¹We prefer the term “focused” over “directed” as used in [1] since the latter, although correct in principle, traditionally denotes “time-dependent wave propagation in a noisy environment”.

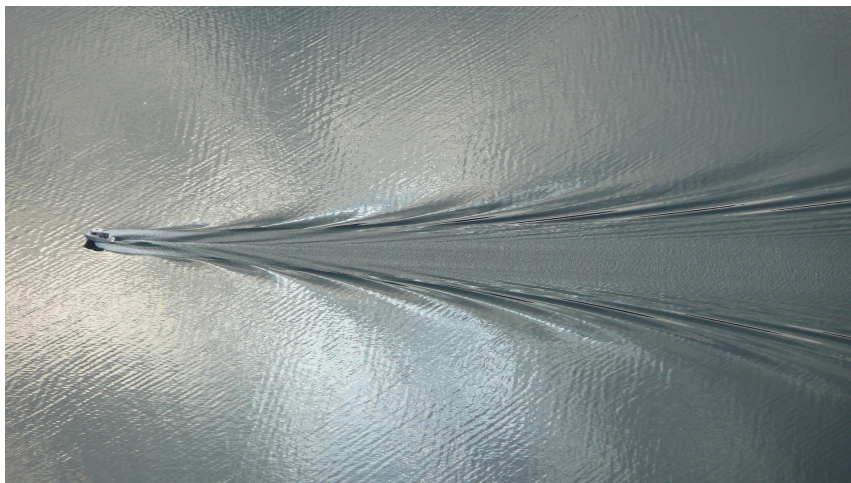


Figure 2.1: Kelvin wake pattern left by a small boat on the Lyse fjord in Norway. Each wake line is offset from the path of the boat by around $\sin^{-1}(1/3) = 19.47^\circ$ and made up of feathery wavelets angled at roughly 53° to the path. This pattern – first observed by Lord Kelvin – is independent of the speed and size of the wake source over a significant range of values. Credit: Wikipedia

the eighth century BC but describes an armed conflict carried on by the Bronze-age Greeks of four to five hundred years earlier. The blind bard tells of the assault on Ancient Troy – the legendarily wealthy city overlooking the Dardanelles – which was (supposedly) triggered by the abduction of an Aegean chieftain’s beguiling wife at the hand of an amorous Trojan prince². It represented a rare combined operation carried out by the major cities of Greece who, for the occasion, “forwent their traditional enthusiastic pastime of preying on one another” [3] to contribute ships and men to the great undertaking³.

A ship traveling swiftly – and Homer invariably calls his heroes’ galleys ὤκους, which is to say ‘swift’ or ‘quick’ – creates a bow wave which it has to climb, thereby

²The armed conflict described by Homer was but one of many sackings of Ancient Troy. The city’s riches and command over the sole entrance to the Hellespont proved too hard to resist for any enterprising chieftain of the day.

³Of particular interest to the nautically inclined might be the epic “Catalogue of the Ships” in Book II of the *Iliad*, which lists the Achaean forces engaged as well as their means of transportation: twenty-nine contingents under 46 captains accounting for a total of 1,186 vessels.

expending considerable energy that would be better spent to increase the vessel's speed⁴. The longer a ship is, the faster it can travel before it is hampered by this effect, but long vessels of any kind were difficult to construct with the limited technology available at the time [3,5]. Such obstacles notwithstanding, all Aegean galleys were long, narrow, and so low that, when the beached Greek battleships were under pressing Trojan attack, Hector could reach up to grab the ornament atop the stern post while Ajax was able to leap from the gunwale to the ground. Moreover, they showed gracefully rounded bilges and were light enough for the crews to run them up on the beach at night [5]. Book I of the Iliad describes a Greek contingent's return to camp after remitting Agamemnon's child bride, Chryseis, into the hands of her father:

As soon as early Dawn appeared, the rosy-fingered, then they set sail for the wide camp of the Achaeans. And Apollo, that worketh afar, sent them a favouring wind, and they set up the mast and spread the white sail. So the wind filled the belly of the sail, *and the dark wave sang loudly about the stern of the ship, as she went, and she sped over the wave*, accomplishing her way. [2],

Book I

Through a process of trial and error, the Homeric galley reached the peak of its development in the penteconter, a monoreme that was between 28 and 33 meters long, open and versatile; twenty-five oarsmen were seated on each side and a single sail placed amidships. It is generally assumed that Odysseus' "black ship"⁵, the fastest military vessel of its day, was a penteconter [5] whose (again!) "swift journey across the wine-dark" Ionian sea (Fig. 2.2) is immortalized in the *Odyssey*:

And as upon a plain four harnessed stallions spring forward all together at the crack of whip, and lifting high their feet speed swiftly on their way; *even*

⁴This happens invariably if the hull (or displacement) speed $- 1.34 \times \sqrt{L_{WL}}$ knots, where L_{WL} is the waterline length of the vessel in feet – is exceeded [4].

⁵In truth, most Greek galleys were painted black.

so the ship's stern lifted, while in her wake followed a huge upheaving wave of the resounding sea. Sagely and steadily she ran; no circling hawk, swiftest of winged things, could keep beside her. ([6], Book XIII)



Figure 2.2: Detail of an Attic black-figure krater with four penteconters on the inner rim. When the krater was full, it seemed as if the ships were sailing on a sea of wine. 6th century BC, Thira. Archaeological Museum, Thera, Santorini, Greece. Credit: <http://etc.ancient.eu/photos/10-greek-pottery-details>

2.3 The Sound Barrier: Man's Conquest of Heaven

Really it is not a question of flying faster than sound; it is a question of flying faster than any pressure effect produced in the air can be propagated.

— Theodore Von Kármán [7]

2.3.1 Supersonic Ballistics

It is perhaps not surprising that Austrian physicist Ernst Mach (1838-1916) recognized straight away the similarity between the ballistic pressure waves captured on his schlieren photographs⁶ (depicted in Fig. 2.3) and the bow shock of a – to speak

⁶Schlieren photography (or *method of striae*), invented by August Toepler in 1864, is a visual process used to capture the flow of fluids of varying densities. The basic optical

one last time with Homer – swiftly traveling ship when he investigated projectiles moving through the air at velocities exceeding the local speed of sound⁷:

If I did not tell what the image represents, you might believe it to be the picture of a boat that glides speedily across the water, taken from a bird's-eye view. In front, the bow wave *ww* can be seen, behind the body an appearance *kk* that is rather similar to the wake with its turbulences. In fact, the light, hyperbola-shaped bow at the apex of the projectile is an air-compression wave, which is very similar to the bow wave of a ship, just that the former is not a surface wave. It is formed in air space and surrounds the projectile like a bell from all sides.⁸

Ernst Mach, a polymath within science and, in his time, leading authority in philosophy, physiology, and physics, was the first to use various diagnostic methods in the systematic study of airborne projectiles [9]. In collaboration with his son Ludwig and Peter Salcher (1848-1928), a professor of physics and mechanics at the Royal and Imperial Naval Academy at Fiume, he succeeded in taking pictures of high-speed bullets, establishing that any body in flight has an ‘escort’ of compression and expansion waves. Moreover, he visually confirmed theoretical studies conducted half a century earlier by Christian Doppler, unfortunately without giving his compatriot due credit for his work.

schlieren system (from German “Schliere”, meaning “streak”) uses light from a single collimated source shining on, or from behind, a target object. Variations in the refractive index, caused by density gradients in the fluid, distort the light beam such that spatial variations in the intensity of the light can be visualized directly with a shadow graph camera.

⁷In Chapter 4 we will show that a two-dimensional cut through a particular solution to the three-dimensional wave equation is *not* the same as solving the same problem in two dimensions. Although Mach was correct to point out the similarities between the pressure wave preceding a bullet and the bow wave of a boat, the physics of the two cases is notably different.

⁸The translation of this and all other quotes from [8] are due to the present author since none could be found in the published literature. Considering the relevance of Mach's writings to acoustics and aerodynamics, this omission seems in need of redemption.

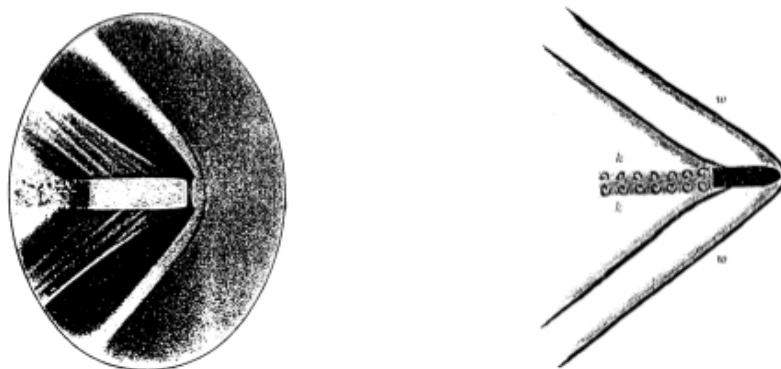


Figure 2.3: (left) Schlieren photograph by Ernst Mach depicting a projectile in supersonic flight. (right) Mach’s schematic representation of the same projectile. *ww* denotes the bow wave, *kk* the wake with its turbulences [8].

Letters exchanged weekly between Salcher and Mach reveal that the latter immediately (and correctly) identified the “head wave” shown in Fig. 2.3 as an organized *envelope of disturbances* that forms when a projectile moves faster than sound – a shock front created by a source whose speed exceeds that of the excited waves themselves [10, 11]:

Just as a slowly moving boat does not show a bow wave, and just as the latter only manifests itself when the boat moves at a velocity greater than the propagation speed of water waves, one cannot see a compression wave ahead of the projectile as long as the velocity of the projectile is less than the propagation speed of sound. If the speed of the projectile reaches and exceeds this value, however, the head wave, as we shall call it, increases visibly in magnitude, and, at the same time, is stretched out steadily, which is to say that the angle of the contours of the wave with the flight path decreases, just as something similar happens when the speed of the boat increases. In fact, from a moment picture taken in the manner described above, one is able to roughly estimate the speed of the projectile [8].

In his compilation of popular scientific lectures, *Populärwissenschaftliche Vorlesun-*

gen (1898), Mach visualized the sound waves emitted by a slender body in supersonic flight by means of the Huygens-Fresnel principle (Fig. 2.4), which stipulates that every point on a propagating wave front is itself a source of expanding wavelets [12]. Constantly emitted by the traveling body and growing at the speed of sound c_s , they are captured at a single instant in time during the source's journey along its trajectory. If the projectile remains subsonic, every wave created at some initial point \mathbf{r} and time t contains all subsequent ones (Fig. 2.5 (left)). If it 'breaks the sound barrier' (Fig. 2.5 (center)) and moves supersonically (Fig. 2.5 (right)), the wavelets intersect to form a conical shock wave front whose axis coincides with the line of flight and whose half angle α is given by $\sin \alpha = c_s/v = 1/M$, where v is the speed of the object and M a dimensionless constant named *Mach number* in honor of the notable physicist and philosopher [11].

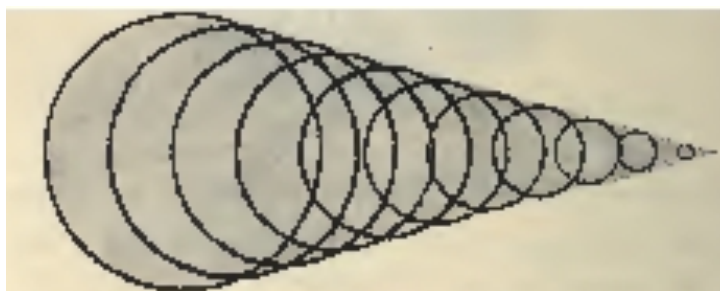


Figure 2.4: Sketch by Ernst Mach depicting Huygens wavelets that combine to form a *Mach cone*, a conical shock wave front with the line of flight as axis and half angle $\alpha = \sin^{-1}(c_s/v) = \sin^{-1}(1/M)$ [8].

Mach's discovery of the bow shock that precedes a projectile in supersonic flight did not only lay the foundation of modern aerodynamics, but also shed light on two unsolved ballistics problems of the day [11]. During the Franco-Prussian war of 1870, it was found that the new Chassepôt high-speed bullets caused big crater-shaped wounds. In consequence, the French were suspected of using explosive projectiles and, therefore, of violating the International Treaty of Petersburg which prohibited the use of such ammunition. Mach put an end to the controversy by giving the complete and correct explanation: The extensive injuries were, in fact, caused by the

high-pressure air between the bullet's bow wave and the projectile itself.

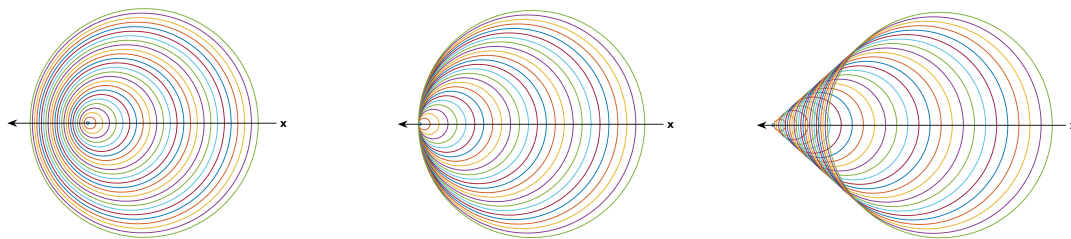


Figure 2.5: (left) Huygens wavelets emitted by a small body (such as a bullet) traveling along a rectilinear path at constant *subsonic* speed, $M = 0.5$. Since the source is moving, the center of each new wave front is slightly displaced to the left. As a result, waves accumulate ahead of its nose and spread out in its wake. This is known as the *Doppler effect*, first explained in 1842 by Christian Doppler. (center) The projectile is advancing *at the speed of sound*, which is to say $M = 1$. In consequence, the Mach cone – which depends on the ratio of source to sound speed alone – is a vertical line. (right) The same point-like body is now moving with constant *supersonic* speed, $M = 1.5$. Since the source is traveling faster than the sound waves it creates, it leads the advancing wave front and will pass by a stationary observer before he or she hears it.

Perhaps more importantly for our purposes, Mach's observations explained why artilleryists could hear two bangs downrange from a marksman when high-speed projectiles were fired, but only one from low-speed bullets⁹. It was realized that, in addition to the explosion from the muzzle, a distant observer would experience the arrival of a shock wave front. The precise nature of what would later be called a *sonic boom*, however, was not thoroughly investigated until the late 1940's and the testing of the first supersonic airplanes.

On a historical note that may resonate with many scientists today, we add that Ernst Mach remained a life long socialist, advocate for the working class, and peace activist [9]. It is unfortunate that his seminal research ultimately led to the development of more refined weapons, an irony which was not lost on the eminent

⁹This explains, incidentally, why sound suppressors are used mainly in conjunction with subsonic ammunition: Silencers can muffle the powder explosion but not the sonic boom that accompanies projectiles in supersonic flight.

scientist. He prefaced his lecture *Über Erscheinungen an fliegenden Projektilen* with the following words:

Humans nowadays feel compelled, sometimes for rather dubious goals and ideals, to shoot, within the shortest time, as many holes into each other's body as possible. And another ideal, which is most often in stark contrast to the one previously mentioned, demands that they simultaneously produce these holes with the smallest caliber possible, and to mend and heal the ones produced as swiftly as possible [8].

2.3.2 The Right Stuff



Figure 2.6: Chuck Yeager in front of his Bell X-1, named “Glamorous Glennis” in honor of his wife. Credit: U.S. Air Force

On October 14, 1947, engineers on the ground at Muroc field in the Mojave desert heard test pilot Chuck Yeager drawl laconically on the radio, “Say, Ridley¹⁰ ... make

¹⁰Colonel Jackie Lynwood “Jack” Ridley, an aeronautical engineer, test pilot and chief of the USAF’s Flight Test Engineering Laboratory. Ridley died on March 12, 1957 at age 41 when the C-47 transport plane that he co-piloted crashed into snow-covered Mt. Shirouma, northwest of Tokyo.

another note, will ya? ...there's somethin' wrong with this ol' machometer ..." (faint chuckle) "...it's gone kinda screwy on me...", code that the fabled pilot had broken the sound barrier and his airplane – an X-1 built by the Bell Aircraft Corporation under an Army contract – gone supersonic¹¹ [13]. At that very moment, a monumental boom rocked over the quonset-style hangars and the single concrete runway, just as predicted by physicists and engineers many years before. In fact, Theodore von Kármán, illustrious Hungarian physicist and, according to Wikipedia, “the outstanding aerodynamic theoretician of the twentieth century”, speculated as early as 1944 that “when an entire aircraft, not just the air accelerating over the thickest part of the wing, went supersonic, shock waves would be sent to the ground” [14]. It stood to reason that people nearby would hear and feel the passing of that pressure pulse.

Ever since the end of the Second World War, the American and British Air Forces had engaged in a furious race to achieve Mach 1 first; however, pilots reported that the controls locked, froze, or even drastically (and often fatally) altered their normal function [13] when the planes approached the speed of sound during dives. Moreover, airspeed gauges were known to ‘hang up’ at around 350 knots when the shock wave began to pass over the indicators’ static sources [14].

In 1946, Geoffrey de Havilland, son of the famous British aircraft designer and builder, tried to take one of his father’s DH 108 Swallows supersonic. The plane started buffeting and then disintegrated, killing the pilot¹² and leaving aeronautics engineers on both sides of the Atlantic ocean to speculate that the shock waves became too severe and unpredictable around Mach 1 for any aircraft to withstand.

¹¹Although Yeager’s X-1 undoubtedly exceeded Mach 1 on that day, fellow test pilot Al Blackburn maintains that George Welch, a World War II flying ace and one of the very few United States fighter pilots able to get airborne in the attack on Pearl Harbor, beat him – on the same airfield – by about two weeks [14]. Welch, who never made such a claim himself, died on Columbus Day 1954 when his Super Sabre disintegrated during a 7-G pullout at Mach 1.55.

¹²It is widely believed that de Havilland’s Swallow exceeded the speed of sound before the remnants of his plane crashed into the Thames estuary [15].



Figure 2.7: A Swiss pilot takes his F/A-18 Hornet supersonic while passing the Eigernordwand. Credit: Christian Merz for an article about the 2017 Swiss air force live fire event at Axalp, Switzerland. <http://www.bernerzeitung.ch/region/oberland/fliegerschiessen-lockt-tausende-auf-die-axalp/story/20165309>

Although mathematical calculations for bodies in supersonic motion were comparatively straightforward and, ever since Mach's *striae*, scientists had visible evidence of the bow shock, the *transonic* region (a small spread of velocity with the speed of sound at its center) remained a murky, ill-defined no man's land characterized by flows that were locally slower or faster than the aircraft's forward motion. Fig. 2.7 shows a Swiss fighter plane that – almost literally – ‘passes through’ Mach 1. The location of the pressure wave indicates that the leading edges of the Hornet's wings are traveling faster than sound while the trailing ones are still (barely) subsonic.

To add to the prevailing sense of doom, the behavior of the lift and drag forces around c_s remained shrouded in mystery, leading to the notion of a – purely hypothetical – *transonic gap* as shown in Fig. 2.8. (For a contemporary analysis of the aerodynamic force in steady transonic flow consult the – very impressive – figures in [16].) Talk about the ‘sonic wall’ and ‘sonic barrier’ emerged, implying that Mach 1 was a physical boundary, absolute and not to be ‘broken’ or ‘pierced’ by pilots. However, von Kármán, illustrious as ever, wrote quite cheerfully amidst the carnage,

“At present the problem of “piercing the sonic barrier” appears to be essentially one of propulsion. If enough propulsive force is available to overcome the drag increase occurring at and immediately before the sonic barrier, so that the airplane can pass quickly through the critical speed range, no specific difficulties need be expected” [7].

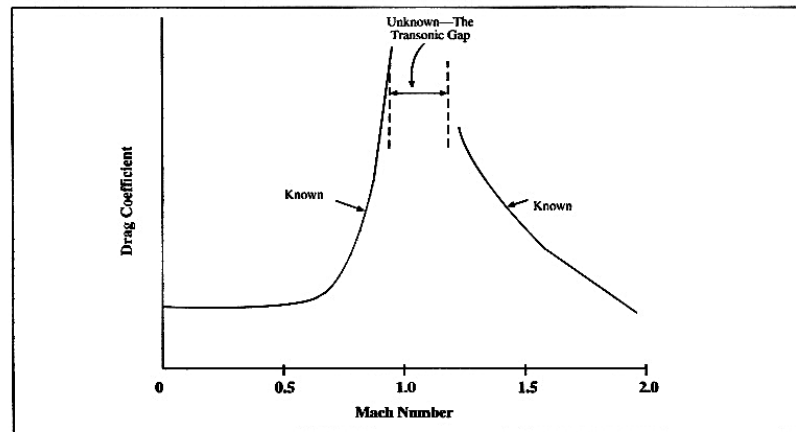


Figure 2.8: Simplified theories predict that the lift and drag coefficients are both infinite at precisely the local speed of sound which, in practice, is of course impossible. Consequently, scientists and engineers frustrated with the high death toll – in 1952, 62 USAF pilots died during 36 weeks of *training* alone – christened the region around Mach 1 the *transonic gap*. Credit: NASA

On September 6, 1948, John Derry became – quite serendipitously – “the first Briton to pass the speed of sound and live to tell about it”, and while Yeager’s tremendous achievement had become “a piece of thunder with no reverberation” [13] due to the highly classified nature of the USAF project, the “historic feat” of the “fine-looking” young British pilot immediately captured the imagination of the public and was rewarded with the “coveted R. Ae. C. Gold Medal” [17]. As a result, entire families flocked to events such as Britain’s annual Farnborough Air Show to witness novel airplanes swoop through the sound barrier¹³, generating shock waves

¹³In a cruel twist of fate, John Derry died four years, to a day, after passing the speed of sound in what became known as *the Farnborough Tragedy*. His plane, the first prototype de

that “sounded like cannon fire” and “ – palpable as ocean breakers – crashed against the crowd’s bodies and ears” [15].



Figure 2.9: John Derry’s Havilland 110 losing its tail during the *Show at Farnborough* [17].

It was not until these public demonstrations that physicists began to engage in an open discussion of the acoustic phenomena that accompany the dive of an aircraft to attain supersonic speed. After a period of much speculation and controversy, Gold, Warren and Rothwell [18–20] published a series of letters in *Nature*, pointing out that a shock will reach a stationary observer “when the wavelets emanating from successive elements of the path of the source (the aircraft) superpose at the observer” [18]. According to the authors, this happens whenever $-dr/dt = c_s$, where r is the scalar distance of the aircraft from the field point. We remark in passing – and, so far, without proof – that this is merely the condition for “pulse waves” [21] or the “boom carpet”, a hyperbola-shaped area delineated by the intersection of a Mach cone with the ground. The ear-shattering claps experienced by air show revelers require, in addition, that $-d^2r/dt^2 = 0$, *i.e.*, that the instantaneous distance to the observer change as c_s *without acceleration*.

Havilland 110 supersonic fighter, disintegrated during the S.B.A.C. Show at Farnborough (Fig. 2.9), killing the young pilot, his co-pilot, and 29 spectators [17].

Rothwell [20] submitted that, in the case of a body which accelerates and then decelerates through the speed of sound, “the shock wave departs radically from the single conical surface and becomes a meniscus symmetrical about the line of flight, with two surfaces” as shown in Fig. 2.10. (Here, the duration of supersonic flight is 16 seconds and the Mach numbers are shown at the positions of the body at second intervals.) Using Huygens constructs and the method of *striae*, G. M. Lilley *et al.* [21] finally demonstrated that, for such a source, two curved wave fronts emerge, joined at their outer extremities by *cusps* (depicted in Fig. 2.11). These ‘bundles’ of pressure waves, born when the aircraft bursts through c_s , expand in a ring around that point in spacetime causing loud bangs as they pass the observer and, eventually, hit the ground.

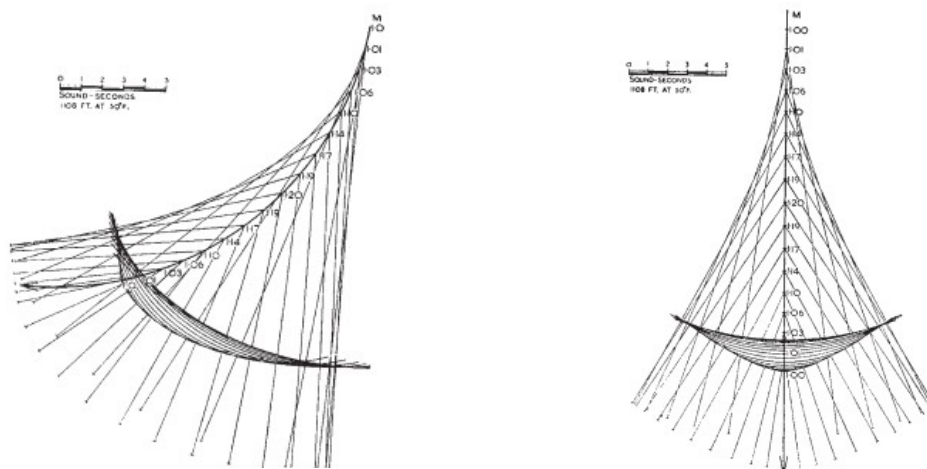


Figure 2.10: Shock wave formed by a body moving in a straight line at a speed which increases and then decreases through the speed of sound [20].

Meanwhile, von Kármán [7] along with Jakob Ackeret¹⁴ in Zürich [22] and Maurice Roy in Paris [23] reached similar conclusions, suggesting that the booms are due to the ‘piling up’ of sound impulses – such as the engine’s roar – that emanate from the flying body as it passes through the sonic speed [7]. In the aircraft’s frame of

¹⁴The present author’s mother, a Hungarian refugee, was Ackeret’s first female graduate student at the Swiss Polytechnic Institute (ETH Zürich).

reference, low energy noise is emitted over an extended time span, all of which arrives at a stationary spectator's frame of reference in a single instant, creating a large and concentrated boom¹⁵. Hence, prolonged periods of source time may contribute to a

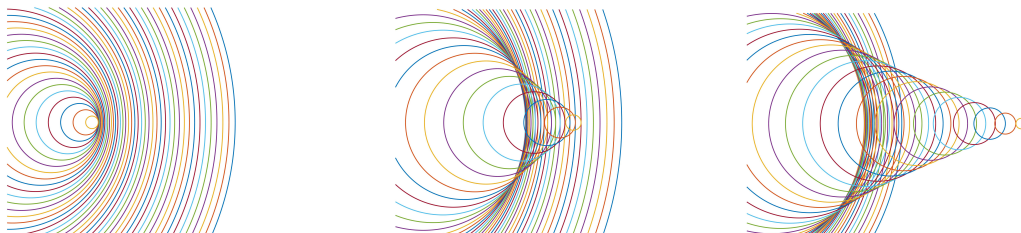


Figure 2.11: Huygens wavelets emitted by a point-like source accelerating along a straight line from zero through the sound barrier. (left) The instant velocity of the source remains below c_s . A distant observer will experience the Doppler effect as described in the caption to Fig. 2.5. (center) The small body in flight has broken the sound barrier and passes through pressure waves that were emitted earlier. (right) Free of bow waves, the source continues to accelerate, leaving a Mach cone in its wake that expands at the speed of sound. Notice the formation of cusps in the center and right-hand panels, bundles of pressure waves that begin to aggregate when the flying object accelerates through c_s .

single moment of reception, which is to say that multiple wave fronts with differing excitation times can pass through a distant point in spacetime almost simultaneously. The concept of *temporal focusing* – or *focusing in time* – was born, the concentration of energy carried by waves in the time domain.

As evidenced by a brief, anonymous communication in *Nature* on the impact of Concorde flights [24], the true nature of the “superboom” phenomenon (*i.e.*, the con-

¹⁵This becomes evident if one considers that all noise emitted by an aircraft that moves exactly at c_s straight toward a stationary observer reaches the latter in one short moment, namely when the plane arrives at the observation point (Fig. 2.5 (center)). The reason is that *sound* and *sound source* are traveling at the same speed. The present author witnessed this during a small air show in Subotica, Serbia, in 1984. A young and particularly adventurous fighter pilot ‘collected’ pressure waves over several dozen miles holding his MiG steady at Mach 1. Once he reached the gathered crowd at an altitude that can’t have exceeded 20 meters, he performed a barrel roll (!) through the sound barrier. The ensuing release of the pressure wave quite literally knocked the spectators off their feet while the associated boom rang in their ears for hours.

nection between acceleration, the formation of focused cusps as the aircraft bursts through the sound barrier, and the propagation of a monumental ring-shaped pressure wave) eluded the aeronautics community for many decades. Whereas it was well known that the bangs occur “when a supersonic aircraft changes speed or direction, with the result that propagated shock waves are momentarily focused at a single point”, theories of the day could not predict the strength of the latter which, by then, were deemed “undesirable both to animals as well as the stability of buildings” and “a cause of annoyance”. Although the unknown author suggested more flight and laboratory tests, he expressed his hope that with suitable planning the concentrated noise generated by commercial supersonic flights “can be dumped in the sea”. Of course, the well heeled passengers’ rest remained undisturbed. As a former Concorde pilot puts it, “You don’t actually hear anything on board. All we see is the pressure wave moving down the aeroplane – it gives an indication on the instruments. And that’s what we see of Mach 1. But we don’t hear the sonic boom or anything like that. That’s rather like the wake of a ship – it’s behind us” [25].

Following the pioneering work in Europe and the United States, the focusing of sonic booms was put on an evermore quantitative footing, culminating in detailed predictive studies ([26–28]). Of particular interest in the context of this work are texts that connect sonic boom modeling to catastrophe theory such as [29–32], a predilection which will not become fully transparent until the discussion of electromagnetic point sources in Chapter 6.

2.4 The Discovery of Light

Light thinks it travels faster than anything but it is wrong. No matter how fast light travels, it finds the darkness has always got there first, and is waiting for it.

— Terry Pratchett, *Reaper Man*

We have seen that a stationary observer may experience pressure waves emitted at different times along the flight path of a body in supersonic motion as a single burst of sound; elsewhere a sole contribution (or none at all) is received¹⁶. (More formally, one might say that the relation between source and reception time need not necessarily be monotonic and one-to-one.) Given this interesting attribute of supersonic emission along with the rather attractive ability of the sonic boom to deposit energy in a focused manner, it seems natural to ask whether *electromagnetic* waves can be manipulated in a similar fashion. If so, how would one achieve such effects in view of Einstein’s finding that nothing travels faster than light [33], or, rather, *nothing that possesses rest mass*¹⁷?

2.4.1 Dr. Einstein’s Universal Speed Limit

Although sources that travel faster than the characteristic wave speed have been studied extensively in hydrodynamics and acoustics, the same can, despite auspicious beginnings, not be said for electrodynamics. Modern-day scientific investigation of the potentials and fields generated by a charged particle that moves faster than light

¹⁶Consult the illustrations in Sections 2.3.1 and 2.3.2, in particular Figs. 2.5 and 2.11.

¹⁷Malykin and Romanets provided a somewhat speculative historical review of superluminal sources in 2011 [34]. However, their material is presented in a rather non-linear fashion that is not particularly relevant to the topics covered in this work. We therefore choose to present here a history that is targeted towards the contents of Chapters 3 – 7 rather than a full review of the field.

began with a largely ignored article by physicist and mathematician Oliver Heaviside in 1887 [35]. Several monographs written by Sommerfeld followed in 1904 and early 1905 [36]. In his *Electrical Papers* [35], Heaviside describes the optical analogue to the Mach cone as follows:

Returning to the case of a charge q at a point moving through a dielectric, if the speed of motion exceeds that of light, the disturbances are wholly left behind the charge, and are confined within a cone, AqB . The charge is at the apex, moving from left to right along Cq . The semi-angle, θ , of the cone, or the angle AqC , is given by $\sin \theta = c/u$, where c is the speed of light, and u that of the charge¹⁸ [35].

A similar description is given by Sommerfeld [36], who illustrated his findings with what would now be referred to as a *Čerenkov cone* (see Fig. 2.12 and the discussion below), the envelope of wavelets emitted by a charge moving at a constant velocity that is higher than that of light.

The publication of Albert Einstein's theory of special relativity in June 1905 [33] brought much of this early work to an abrupt end, since one of its tenets is, of course, that nothing that has rest mass may travel faster than light. (Note, however, that it is the requirement of causality (*e.g.*, the grandfather paradox) rather than relativistic considerations that precludes massive or information-carrying bodies from moving superluminally [34].) All known charged particles, both then and now, have rest mass. Moreover, no source of electromagnetic radiation that exceeds the wave speed can be point-like, for this results in fields of infinite strength¹⁹ on the envelope of the

¹⁸The original document denotes the speed of light as v rather than the customary constant c . The denomination c , for the Latin *celeritas* – swiftness – was first introduced in 1856, when Wilhelm Eduard Weber and Rudolf Kohlrausch used it for a constant later shown to equal $\sqrt{2}$ times the speed of light *in vacuo*. In 1894 Paul Drude redefined c , giving it its modern meaning. Einstein used V in his original German-language papers on special relativity in 1905, but in 1907 he switched to c , which by then had become the standard symbol.

¹⁹As will be shown in coming sections, it is mathematically more accurate to describe the

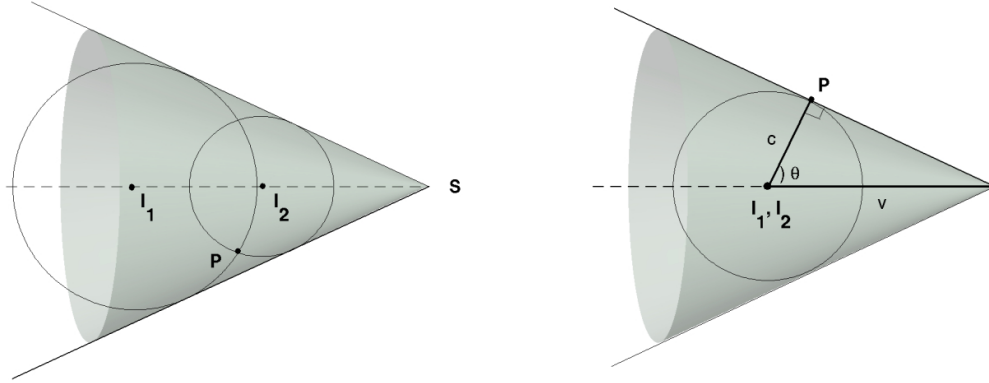


Figure 2.12: Huygens wavelets (light curves) and their envelope for a superluminal source moving at a constant velocity along a rectilinear trajectory. (left) An observer (P) inside the Čerenkov envelope samples two images of the source (I_1, I_2), one to his left, the other one to his right. (right) On the envelope of the wave fronts, the two contributions coalesce and a spectator will see one image only. Nothing is received outside the conical surface.

emitted wave fronts (see [37] and, more than 50 years later [1, 38, 39]).

These considerations did not deter the eccentric British mathematician G. A. Schott, one of the last ‘respectable’ scientists to reject the quantum formalism introduced by Niels Bohr, from considering the emission of radiation by superluminal charged particles; his calculations, illustrations and “solutions to problems” on this topic form part of a book that he published in 1912 [37]. While Schott recognized that “useful physical theories such as the Relatiftheorie²⁰ of Lorentz and Einstein are incompatible” with particle velocities greater than that of light, he considered the limitation to the subluminal regime as “undesirable from the standpoint of complete mathematical generality”. Acknowledging its many successes in explaining experimental observations, Schott nevertheless included a critique of the theory of special relativity based on commentaries and objections of the day. He concluded that “we must therefore be careful not to allow theoretical views ... unduly to influence our notions as to the possibility, or otherwise, of velocities greater than that of light”.

strength of the electromagnetic fields on the envelope as ‘undefined’ rather than ‘infinite’.

²⁰Theory of special relativity

Schott's classic publication, which was later honored with the Adams prize, gives credit to his predecessors by summarizing Heaviside and Sommerfeld's ideas on faster-than-light charges with constant velocities, but then goes much further by describing infinitesimal bodies undergoing linear and centripetal acceleration as well as oscillatory motion. (The envelope of a source in superluminal linear acceleration is shown in Fig. 2.13 and, looking ahead to Chapter 4, Fig. 4.6.) Some of this work predicts issues, such as multiple retarded times, raised forty years later in the acoustic studies of the sonic boom [7, 18–23] described in the preceding section. With an eye to the undefined fields produced by a charge without extent in superluminal motion, it also considers entities in which the source is distributed over surfaces or volumes.

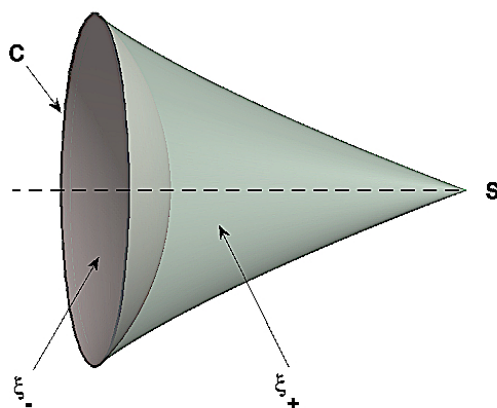


Figure 2.13: Superluminal source that is linearly accelerated. The wave fronts that are emitted after the source breaks the “light barrier” form a Čerenkov envelope consisting of two axisymmetric sheets (ξ_+ and ξ_-) which meet to form the cusp (C), a region of intense concentration of emitted radiation.

While European scientists universally abandoned their studies of bodies in faster-than-light motion in the wake of Einstein's findings, their counterparts in the Soviet Union were undeterred. In 1934, while working under Sergey Vavilov (1891-1951), Pavel Čerenkov (1904-1990) observed bright cerulean light emanating from a bottle of water that was subject to radioactive bombardment [40], a discovery which proved

to be of great consequence for subsequent experimental work in nuclear physics and the study of various cosmic phenomena [41]. (Marie and Pierre Curie observed the same effect as early as 1900 in solutions of radioactive salts but dismissed it as “luminescence”.)

Relativity holds that the speed of light *in vacuo* is a universal constant. However, the velocity at which electromagnetic waves propagate in a material may be significantly less: In water, for instance, they move at a mere $0.75c$ [41]. Whenever the velocity $v = \beta c$ of a charged particle, most commonly an electron, exceeds c/n , where n is the refractive index of the medium, this particle emits Čerenkov radiation [41], named after the man who discovered it.

Čerenkov radiation is very similar in nature to a bow wave or the Mach cone: Radiation emitted by an excited ion or electron traveling along a straight path will form a coherent wave front – the Čerenkov envelope – at a specific angle [41]. This caustic constitutes the surface of a cone whose axis coincides with the particle’s trajectory and whose aperture is given by

$$\theta = \arccos\left(\frac{ct/n}{\beta ct}\right) = \arccos\left(\frac{1}{\beta n}\right), \quad \beta > \frac{1}{n}; \quad (2.1)$$

hence, the higher the source speed, the narrower the cone. For a genuine superluminal source, *i.e.*, one that gives rise to so-called *vacuum Čerenkov radiation* by moving faster than the speed of light *in vacuo*, expression Eq. (2.1) reduces to $\theta = \arccos(1/\beta)$ since $n = 1$.²¹

Čerenkov’s experiments gave visual proof of a quality unique to the radiation emitted by a charge that exceeds the speed of light as predicted by Schott decades earlier: The relation between source and observation time may, but need not, be monotonic and one-to-one. Of particular interest is the crossover between regions

²¹The first demonstration that faster-than-light technology demonstrators function as true superluminal sources involved the detection of vacuum Čerenkov radiation emitted at an angle that depends on the speed of the source alone [42–45].

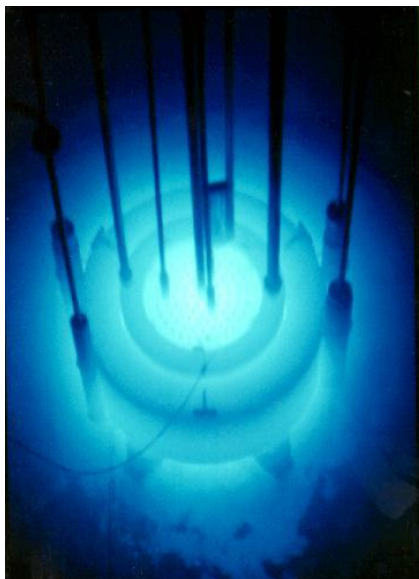


Figure 2.14: Blue Čerenkov light emanating from the water above the Reed Research Reactor in Portland, OR. Credit: <http://www.webexhibits.org/causesofcolor/4BA.html>

where the retarded time is multiple-valued and others where it is not [37]. As a visual aid, reconsider Fig. 2.12: An observer situated outside the Čerenkov envelope cannot see the source since none of the advancing waves have reached him yet, whereas one within it samples two distinct images from the source’s history (Fig. 2.12 (left)). On the envelope of the wave fronts the two contributions coalesce and one sole image is received (Fig. 2.12 (right))²².

²²Readers of a sci-fi bent will be amused to learn that this effect is known as the *Picard Maneuver* among followers of the American television series *Star Trek* and denotes a battle tactic invented by Starfleet Captain Jean-Luc Picard. In 2355, Picard was in command of the USS Stargazer on a routine mission in the Maxia Zeta system when the starship was attacked by an unknown alien vessel eventually determined to be of Ferengi origin. During the engagement – later to be known as the *Battle of Maxia* – the Stargazer, which was holding station several million kilometers distant from its adversary, suddenly accelerated to warp 9 (1516 times the speed of light) directly towards the Ferengi ship. Because the latter was only equipped with light speed sensors, its crew had no way of knowing that the Stargazer had changed position until it was too late. When the light from the moved vessel reached the Ferengi ship’s sensors, wave fronts from its previous position were still arriving. Alas, the Ferengi saw *two* images of the Stargazer concurrently and decided to fire at the wrong one. The Picard Maneuver later became required reading at Starfleet Academy.

Whilst Schott’s work was largely forgotten in the West (at least until the advent of infrared cloaking technologies in the early twenty-first century) Ilya Frank (1908-1990) and Igor Tamm (1895-1971) soon produced a convincing theory that describes the emission of Čerenkov radiation [46]. It is, perhaps, for this reason that interest in superluminal sources never quite died in the Soviet Union.

Beginning in the late 1950s Vitalii Ginzburg [47, 48] and, later, Boris Bolotovskii [1, 38, 39] – both of the Russian Academy of Sciences – developed an interest in vacuum Čerenkov emission. In [38] they boldly stated, “It has been long known that velocities exceeding that of light in vacuum are possible and are encountered in physics and astronomy”. Acknowledging that “superlight” sources must be treated differently than their “sublight” counterparts, they christened the volume within the Čerenkov cone “region of anomalous Doppler effect”. As Schott did decades earlier (and giving him credit for his discovery), the two Soviet scientists realized that “from the point of view of radiation theory, the essential difference between superluminal ($v > c$) and subluminal ($v < c$) sources is that *a superluminal source cannot comprise an individual “elementary” particle*, and has therefore always some size” (which is to say that it cannot be a point charge).

By the same token Ginzburg and Bolotovskii were the first to point out that a *polarization current* (or, in the words of [38]’s translator, a “bunch” of individual particles) fulfills the requirements of “some size” and zero rest mass. This allows it to function as a faster-than-light source of radiation without breaking any laws of physics, a proposal which they illustrated convincingly by deriving Snell’s laws of refraction and total internal reflection using superluminal polarization currents at the interface between two different dielectric media [38]. Cautioning that these distribution patterns, albeit capable of advancing faster than light, cannot be used to make *the emitted signal* move superluminally and thereby facilitate transtemporal travel, they nevertheless conclude that “interesting applications” may be found for this new and, at that point, purely theoretical technology. Rather unsurprisingly, a

“powerful generator of broadband directed radiation” was proposed in the wake of Ginzburg and Bolotovskii’s findings [34] and, in 2004, built [49–51].

To fully appreciate Bolotovskii and Ginzburg’s argument, however, it is necessary to step back about one hundred years and into the mind of Scottish physicist and mathematician James Clerk Maxwell (1831-1879), who was the first to characterize polarization currents and how they might serve as sources of electromagnetic radiation.

2.4.2 A Tale of Two Fields

*This crystal tube the electric ray
Shows optically clean,
No dust or haze within, but stay!
All has not yet been seen.
What gleams are these of heavenly blue?
What air-drawn form appearing,
What mystic fish, that, ghostlike, through
The empty space is steering?*

— James Clerk Maxwell

Maxwell’s equations serve as the most elegant and concise way to express how electricity and magnetism combine to form one of the fundamental forces of the universe – electromagnetism. The quartet of partial differential equations describes the electric and magnetic fields arising from varying distributions of charges and currents as well as how they change with time. Maxwell’s own contribution is just the last term of the last equation (he called it the *displacement current*), but realizing its necessity had dramatic consequences: It made evident for the first time that varying electric and magnetic fields can draw on one another, allowing them to travel through space without end, far from the charges and currents where they originated.

The four equations²³ that relate the electric and magnetic field intensities \mathbf{E} and \mathbf{H} to the electric and magnetic flux densities \mathbf{D} and \mathbf{B} are – in SI units – given by

$$\nabla \cdot \mathbf{D} = \rho \quad (2.2)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2.3)$$

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} \mathbf{B} \quad (2.4)$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial}{\partial t} \mathbf{D} \quad (2.5)$$

where ρ and \mathbf{J} denote the electric charge density and current density of free charges. (The significance of the designation “free” will become clear in the section to follow.) The first two equations are attributed to eminent²⁴ German mathematician Carl Friedrich Gauss (1777-1855); they state, in turn, that the electric flux through any closed surface is proportional to the total electric charge contained within and that the magnetic field is solenoidal (*e.g.*, divergence free). The third is Faraday’s law of induction, and the last Ampère’s circuital law including Maxwell’s ancillary term, the displacement current $\partial \mathbf{D}/\partial t$, which will feature prominently in the sections to come. In an insulating, uniform, and isotropic medium ρ and \mathbf{J} are zero whilst \mathbf{B} and \mathbf{D} are connected to \mathbf{H} and \mathbf{E} by the *constitutive relations*

$$\mathbf{B} = \mu \mathbf{H} \quad (2.6)$$

$$\mathbf{D} = \varepsilon \mathbf{E} \quad (2.7)$$

where μ and ε – both constants²⁵ – are the magnetic permeability and electric permittivity respectively. Hence, Eqs. 2.4 and 2.5 may be written as

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} \mathbf{B} = -\mu \frac{\partial}{\partial t} \mathbf{H} \quad (2.8)$$

²³ [52] is an excellent resource for novices to Maxwell’s equations, which were voted “the most important equations of all time” by readers of *Physics World*.

²⁴Such was his influence on mathematics and science that Gauss is, to this day, considered *princeps mathematicorum* - “the foremost among mathematicians”.

²⁵In an inhomogeneous, anisotropic medium, however, μ and ε are position-dependent tensors.

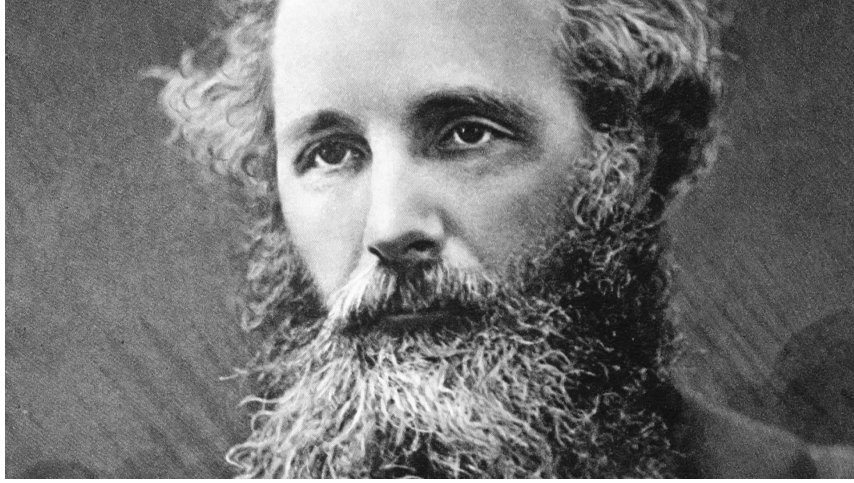


Figure 2.15: James Clerk Maxwell (1831-1879). Credit: <https://readingforlearning.org/2017/08/17/james-clerk-maxwell-and-the-second-great-unification-in-physics/>

and

$$\nabla \times \mathbf{H} = \frac{\partial}{\partial t} \mathbf{D} = \varepsilon \frac{\partial}{\partial t} \mathbf{E}. \quad (2.9)$$

If we apply the curl operator to both sides of Eq. (2.8) and take advantage of a vector identity to resolve the resulting double curl, we find that

$$\begin{aligned} \nabla \times \nabla \times \mathbf{E} &= \nabla(\nabla \cdot \mathbf{E}) - \nabla \cdot (\nabla \mathbf{E}) = \nabla \times \left(-\mu \frac{\partial}{\partial t} \mathbf{H} \right) \\ &= -\mu \frac{\partial}{\partial t} (\nabla \times \mathbf{H}), \end{aligned} \quad (2.10)$$

while taking the partial derivative of Eq. (2.9) with respect to time results in

$$\frac{\partial}{\partial t} (\nabla \times \mathbf{H}) = \varepsilon \frac{\partial^2}{\partial t^2} \mathbf{E}. \quad (2.11)$$

Combining Eqs. 2.10 and 2.11 we arrive at

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla \cdot (\nabla \mathbf{E}) = -\mu\varepsilon \frac{\partial^2}{\partial t^2} \mathbf{E} \quad (2.12)$$

which can be re-written as

$$\nabla^2 \mathbf{E} - \mu\varepsilon \frac{\partial^2}{\partial t^2} \mathbf{E} - \nabla(\nabla \cdot \mathbf{E}) = 0. \quad (2.13)$$

Since $\nabla \cdot \mathbf{D} = \nabla \cdot \varepsilon \mathbf{E} = 0$ when $\rho = 0$ (see Eqs. 2.2 and 2.7), we obtain the well-known expression for the homogeneous wave equation²⁶,

$$\nabla^2 \mathbf{E} - \varepsilon \mu \frac{\partial^2}{\partial t^2} \mathbf{E} = \nabla^2 \mathbf{E} - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \mathbf{E} = 0, \quad (2.14)$$

which shows that the electric field \mathbf{E} propagates with wavelike motion and velocity $v = (\varepsilon \mu)^{-1/2}$. A similar method can be used to derive

$$\nabla^2 \mathbf{H} - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \mathbf{H} = 0. \quad (2.15)$$

In *vacuo*, $\varepsilon \rightarrow \varepsilon_0$ and $\mu \rightarrow \mu_0$, the permittivity and permeability of free space, which implies that the speed of light, c , must be the same as the quantity $(\varepsilon_0 \mu_0)^{-1/2}$. This insight allowed Maxwell to relate the units of magnetism and electrostatics which, in turn, revealed the fundamental nature of electromagnetic force. Electromagnetic waves are discussed in detail in many excellent texts; the present writer's favorites are, in no particular order, [53–59].

2.4.3 A Necessary Word on Notation (Act I)

Even a cursory study of electrical engineering and physics texts (including the usual suspects Jackson [53], Dudley [60], Balanis [54, 61], and Harrington [55]) shows that the notation used for electromagnetic quantities and vector operators is inconsistent, sometimes confusingly so. After much anguished contemplation – to quote L. C. Evans, “notation is a nightmare!” [62] – we decided to use the bold symbols \mathbf{E} , \mathbf{D} , \mathbf{H} and \mathbf{B} to stand for *general, time-varying* electromagnetic fields, well aware that [54, 55, 61] and especially [60] reserve the same notation exclusively for systems that are *time-harmonic*. Breaking with engineering tradition, we chose i – not j – to denote $\sqrt{-1}$ and, since $\mu = \mu_0$ throughout this work, excluded the magnetization \mathbf{M} from Maxwell's equations. With these decisions we followed the standard recommended

²⁶Eqs. 2.14 and 2.15 are often referred to as “electromagnetic wave equations” in textbooks and the published literature.

by the International Union of Pure and Applied Physics (IUPAP) [63] and adopted, to varying degrees of success, by [53, 56–58] as well as all our technical reports. Perhaps more importantly, this convention ensures that differentiation with respect to time is *always explicit*, rendering the results generally applicable irrespective of the time-dependence – harmonic or not – of the fields.

With a nod to North American audiences (and in contrast to European etiquette), ∇v (or $\nabla \mathbf{v}$), $\nabla \cdot \mathbf{v}$ and $\nabla \times \mathbf{v}$ stand, in turn, for the gradient, divergence, and curl operators acting on a scalar or vector field in Euclidean space. The Laplace operator (or Laplacian) is then defined as

$$\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad \text{in } \mathbb{R}^3.$$

To avoid unnecessary confusion, we refrain from using the delta-notation customary to denote the Laplacian in theoretical texts on partial differential equations (*i.e.*, $\Delta v \equiv \nabla^2 v$) and adopt the above symbolism exclusively, which is more widespread in theoretical and mathematical physics.

The D'Alembert operator (or D'Alembertian), in essence the Laplace operator of Minkowski spacetime, is represented by a box (\square) – (*not* \square^2 as is sometimes done in shaky analogy with ∇^2). In standard coordinates $(x, y, z; t)$ it has the form

$$\begin{aligned} \square &= \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \\ &= \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2. \end{aligned}$$

With an eye on the mathematical discussions in Chapters 4 and 6, the dimensionality of a problem in spacetime is indicated by $\mathbb{R}^n \times (t_0, t_1)$, where n denotes the number of space variables and the right-hand bracket the range of time considered. This is nowadays standard notation across all mathematical disciplines.

2.4.4 Ampère’s Missing Term

The fundamental principle underlying the practical faster-than-light emitters presented in Chapter 3 and modeled in Chapters 7 and 8 is the adoption of polarization currents as sources of electromagnetic radiation. It is therefore essential to summarize the derivation of Maxwell’s second curl equation (Eq. (2.5)) and the fields concealed within it.

Maxwell laid out much of our current understanding of electromagnetic radiation in three magisterial papers. In the third, *On Physical Lines of Force* [64], he considers the electrical charges within insulators. Ever ahead of his time²⁷ he states, “In a dielectric under induction” (*i.e.*, subjected to what we would now call an electric field) “we may conceive that the electricity in each molecule is so displaced that one side is rendered positively and the other negatively electrical, but that the electricity remains entirely connected with the molecule, and does not pass from one molecule to another. The effect of this action is to produce a general displacement of the electricity in a certain direction.”

Ref. [65] gives probably the first accurate description of what we now call *polarization* \mathbf{P} , which is due to the slight movement of bound positive and negative charges in opposite directions under the effect of an electric field. This results in a *dipole moment*, defined as the charge displaced multiplied by the distance that it moves; in consequence, any polarization \mathbf{P} is the dipole moment per unit volume. In the same paper Maxwell introduced the concept of *displacement*, a vector proportional to the electric field *that includes the effect of bound charges*.

The mathematics of the day (well before modern vector analysis was developed)

²⁷There are those who believe that, given more time, Maxwell would have discovered the fundamental principles of Special Relativity. In *A Dynamical Theory of the Electromagnetic Field* (1865) he wrote, most notably, about the propagation speed of electromagnetic waves, “This velocity is so nearly that of light that it seems we have strong reason to conclude that light itself (including radiant heat and other radiations if any) is an electromagnetic disturbance in the form of waves” [65].

render Maxwell's treatment of the displacement somewhat lengthy²⁸ but, in present-day notation, it can be expressed as follows. Originally, the differential form of Ampère's equation was

$$\nabla \times \mathbf{H} = \mathbf{J}' \quad (2.16)$$

where, as before, \mathbf{H} is the magnetic field and \mathbf{J}' a current density, albeit not necessarily equal to \mathbf{J} . Taking the divergence of Eq. (2.16) and keeping in mind that $\nabla \cdot \nabla \times \mathbf{v} = 0 \forall \mathbf{v}$, we obtain

$$\nabla \cdot \nabla \times \mathbf{H} \equiv 0 = \nabla \cdot \mathbf{J}'. \quad (2.17)$$

Maxwell realized that if \mathbf{J}' in Ampère's equation were to represent *only* \mathbf{J} , the current density of free charges, then the law of conservation of charge,

$$\nabla \cdot \mathbf{J} + \frac{\partial}{\partial t} \rho = 0, \quad (2.18)$$

would be violated.

As touched on briefly in Section 2.4.2, Gauss and Faraday had already manipulated (very elegantly, as usual) Coulomb's law of electrostatics. It may be written

$$\nabla \cdot \mathbf{D} = \rho$$

in modern-day notation and is already listed as Eq. (2.3) in Section 2.4.4. Here, \mathbf{D} is the above-mentioned displacement, a vector proportional to the electric field that includes the influence of bound charges via the relative permittivity ε_r such that, in a linear, isotropic dielectric,

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} = \varepsilon_0 \varepsilon_r \mathbf{E}. \quad (2.19)$$

Carrying out an operation equivalent to substituting Eq. (2.3) into Eq. (2.18), Maxwell obtained

$$\nabla \cdot \left(\mathbf{J} + \frac{\partial}{\partial t} \mathbf{D} \right) = 0. \quad (2.20)$$

²⁸In fact, it was Heaviside who cast Maxwell's derivations in a more contemporary form.

Hence, as he now had a quantity that takes into account conservation of charge, and whose divergence is zero, as required by Eq. (2.17), he established that

$$\mathbf{J}' \equiv \mathbf{J} + \frac{\partial}{\partial t} \mathbf{D}, \quad (2.21)$$

where $\partial \mathbf{D} / \partial t$ has the units of current density and is therefore, rather unsurprisingly, known as *displacement current density*.

Maxwell's implementation of essential charge conservation resulted in the Ampère-Maxwell equation as given in Eq. (2.5) or, using Eq. (2.19),

$$\nabla \times \mathbf{H} = \mathbf{J} + \varepsilon_0 \frac{\partial}{\partial t} \mathbf{E} + \frac{\partial}{\partial t} \mathbf{P}. \quad (2.22)$$

As shown in depth in Section 2.4.2, $\nabla \times \mathbf{H}$ and $\partial \mathbf{E} / \partial t$ are associated with the propagation of electromagnetic waves whilst the generation of radiation is encompassed by the source terms \mathbf{J} and $\partial \mathbf{P} / \partial t$. An oscillating or accelerating current density of free charges is the basis of conventional radio transmitters, synchrotron light sources, and many other devices [53–61]. However, the charged particles that make up \mathbf{J} have finite rest mass and are usually electrons, which are almost ideally point-like, doubly ruling out their use as superluminal sources (see Section 2.4.1). Bolotovskii and Ginzburg's superb insight was to realize that polarization currents are by their nature extended (since they depend on the relative displacement of oppositely charged particles) and possess zero rest mass. Ergo, they can be employed as faster-than-light sources. It bears repeating that superluminal polarization currents do not violate special relativity as they cannot be used to transmit a signal superluminally; the emitted radiation (as any other) travels *at* the speed of light.

Many natural or man-made phenomena serve as illustrative examples for the operating principle of polarization currents: Tsunami waves that plow through the sea at the speed of a jet plane, strategically placed beacon fires that are lit, one after another, to signal the advancement of an enemy force²⁹, a row of percussionists that

²⁹In mountainous countries such as Switzerland, beacon fires were used as alert systems

can produce a drum roll that exceeds the speed of sound. In *Mexican waves in an excitable medium*, Farkas, Helbing and Vicsek describe “La Ola” (Fig. 2.16), the wave that surges through the rows of (excitable) sports fans in a soccer stadium and rose to fame during the 1986 World Cup in Mexico [66]. Spectators in one section leap to their feet, arms up, and sit down again as the next section rises to repeat the motion. An observer will see a wave of standing people that travels rapidly through the crowd, a rotating pattern whose speed is only limited by the accuracy of timing. It is not necessary that the wave’s constituent elements move very fast (or far away) from their seats; in the same way, the displaced, charged particles that create a superluminal polarization current never broach light speed.



Figure 2.16: (left) “La Ola” as it surges through a stadium [66]. (right) At-the-time UK Prime Minister Theresa May’s wave during a friendly soccer game between France and England “sent twitter trolls into meltdown” according to the Irish Mirror. Credit: <http://Best Image/Vantagenews.com>

As will be described in brief in Section 2.6 and, more expansively, in Chapter 3, extended sources of electromagnetic radiation whose distribution patterns move faster than light *in vacuo* have now been realized in the laboratory on several occasions. In experiments carried out in Russia, the United Kingdom, and the USA, polarization well into the 19th century and the principle is shown famously in *Lord of the Rings: The Return of the King*. On the behest of Gandalf, Pippin lights the first of seven beacons to request troops from Rohan to join Gondor in the looming final battle against Mordor. Once the first fire is lit the propagation of the signal can not be stopped anymore.

currents traveling at up to 95 times the speed of light (and, hence, “hyperluminally”) have been shown to emit radiation [1, 38, 42, 44, 49–51, 67–69].

The Russian examples follow Bolotovskii and Ginzburg’s original proposal very closely [49]: An electromagnetic wave front produced by a high-power laser pulse is incident obliquely onto a planar sheet of material that functions as a photocathode. The intersection between wave front and photocathode propagates faster than light, liberating electrons, via photoemission, that are drawn to a positive-voltage metal grid above, leaving behind a positive space charge. Hence, a wave front of polarization, produced by separation of the photoelectrons from the space charge, travels along faster than c , emitting radiation. Although this is an excellent plaything for research into electromagnetic pulse (EMP) phenomena, it is not a convenient apparatus for communication purposes. The antennas studied in the present dissertation therefore employ polarization currents induced by completely different principles.

Moreover, superluminal polarization currents are quite possibly responsible for the extreme properties of the electromagnetic radiation received from astronomical objects such as pulsars and gamma-ray bursts. As suggested by Bolotovskii and Ginzburg in 1972 [38], superluminal emission appears to be an important and ubiquitous process in the observable universe that may demand significant amendments to many “standard models”.

2.5 Ad Astra

And in that moment, I swear we were infinite.

— Stephen Chbosky, *The Perks of Being a Wallflower*

In the mid-1980’s Houshang Ardavan, at the time a promising young mathematician at the University of Teheran, posited that faster-than-light polarization currents

might be responsible for the radiation received from pulsars, rapidly rotating, highly magnetized neutron stars [70–78]. The inspiration for Ardavan’s theory, although never credited directly, came almost certainly from the authors of [38], who posited in the early 1970’s, “In the pulsar models, a perturbation traveling with a velocity $v > c$ can be produced in the plasma also by magnetodipole radiation or by particle streams emitted from the pulsar”. It is perhaps surprising that the *superluminal model of pulsar emission* never received the same consideration as the standard light-house model, especially given Vitalii Ginzburg’s apparent support³⁰. The reason for including pulsars in this historical account is that the model promoted by Ardavan is closely related to the discussions in Chapter 6. Indeed, the present author contributed to numerical calculations confirming that virtually all the enigmatic features of pulsar radiation can be explained using this elegant theory with few input parameters and no external assumptions [45, 79–81] (By Occam’s razor, this alone should have made it the dominant paradigm.) Although physical and heuristic – rather than mathematical – in form, these results seem to confirm what Bolotovskii and Ginzburg suspected all along: Superluminal emission of electromagnetic radiation is a natural phenomenon that almost certainly occurs throughout the universe and may, therefore, be of importance to the astronomical and astrophysical communities irrespective of their seemingly insurmountable skepticism. The remainder of this section therefore gives a brief description of pulsar physics so that the assumptions behind the model described in Chapter 6 may be understood.

³⁰The present author believes (and supports her opinion at length in Chapter 5) that Ardavan’s ‘great blunder’ was to shackle the superluminal model of pulsar emission to the concept of *nonspherical decay*, *i.e.*, electromagnetic radiation whose intensity falls off more slowly with distance than predicted by the inverse square law. The latter can be unmasked with (comparative) ease as wishful thinking and mars an otherwise beautiful theory.

2.5.1 Pulsars 101

Soon after Jocelyn Bell, quite serendipitously, discovered the first pulsars in the Summer of 1967 [82], Pacini [83,84] and Gold [85,86] suggested independently that rapidly rotating, highly magnetized neutron stars must be responsible for the observed radio pulses since “no other theoretically known astronomical object would possess such short and accurate periodicities” [85]. This insight provided an entirely unexpected verification of a proposal made more than 30 years earlier by Walter Baade and Fritz Zwicky [87], who suggested the existence of the neutron star, which would be the end point of stellar evolution. In 1934 the two astronomers wrote [87], “with all reserve we advance the view that a supernova represents the transition of an ordinary star into a neutron star, consisting mainly of neutrons. Such a star may possess a very small radius and an extremely high density”. However, since such a neutron star would likely be small, cold, inert and emit precious little light, their conjecture seemed, at the time, beyond the possibility of actual astronomical verification. In a

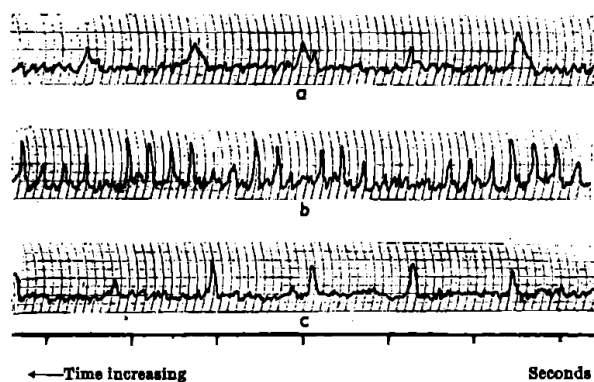


Figure 2.17: Discovery observations of pulsating radio sources recorded by Jocelyn Bell in July 1967 [82].

seminal letter to *Nature*, Gold also pointed out that pulsar radiation per unit emitting volume is bound to be exceptionally high since “the size of the region emitting any one pulse can, after all, not be much larger than the distance light travels in the few milliseconds that represent the lengths of the individual pulses”. He added that

rotational energy is most likely lost through magnetic dipole radiation such that the pulsar “spins down” appreciably with age [85]. Subsequent measurements confirmed that the period of the Crab pulsar lengthens uniformly by 36.48 ± 0.04 ns per day, *i.e.*, by over $1 \mu\text{s}$ per month [88]. This rate of change is not only consistent with the known age of the Crab nebula, thus confirming the association of the pulsar with the supernova observed in AD 1054 (Fig. 2.18), but also provides sufficient energy for the excitation of synchrotron radiation that continuously emanates from the nebula itself [86]. Soon thereafter, Goldreich and Julian [89] concluded that, in spite of their intense surface gravity, rotating magnetic neutron stars “cannot be surrounded by a vacuum” but must possess a “dense magnetosphere”. The particle density within this ‘plasma atmosphere’ can be approximated by

$$n_{\text{GJ}} \equiv n_- - n_+ = 7 \times 10^{-2} B_z P^{-1} \text{ (particles) cm}^{-3}, \quad (2.23)$$

where n_{GJ} is known as the Goldreich-Julian density, B_z stands for the axial component of the magnetic field in Gauss, and P is the period in seconds [89].

In January 1969, a small team at the Steward Observatory in Arizona observed optical signals emanating from the Crab pulsar, confirming the speculation that rotating neutron stars might emit pulses of visible light in addition to radio waves [90]. Two rocket flights, launched shortly thereafter by a team at the Naval Research Laboratory in Washington [91] and the Massachusetts Institute of Technology [92] showed, to everybody’s surprise, that the Crab’s frequency spectrum extends into the X- and gamma-ray regions³¹. Thus, the by then famous neutron star’s frequency spectrum became the broadest of any single radiation source.

53 years after their initial discovery, the catalog of known pulsars, observed by numerous ground-based and orbiting telescopes, has grown to more than 1500 identified sources [93, 94]. Yet, many of the intrinsic characteristics of spinning neutron

³¹X-radiation is absorbed by the Earth’s atmosphere; hence, instruments to detect X-rays must necessarily be taken to high altitude by balloons, sounding rockets, and satellites. Since there was no X-ray telescope orbiting the Earth in 1969, the only viable option was to engage in rocket flights.

stars have remained mysterious and seemingly impervious to human analysis. Most noticeably, neither the region of emission within the pulsar magnetosphere nor the mechanism that produces it have been unambiguously identified, a quandary which caused noted pulsar astronomer Jean Eilek from NRAO to exclaim in frustration, “We know why they pulse. But why do they shine?”³².

2.5.2 The Physics of Neutron Stars

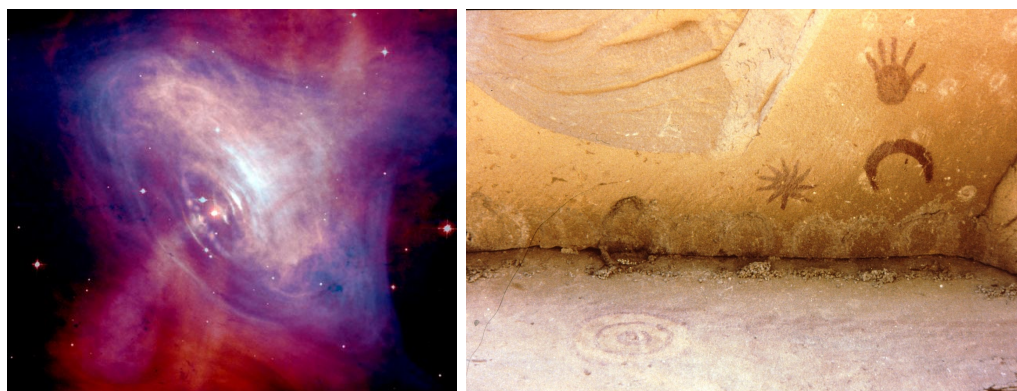


Figure 2.18: (left) The Crab pulsar, a city-sized, magnetized neutron star spinning 30 times per second, lies at the center of this composite image of the Crab Nebula, which combines optical data (red) from the Hubble Space Telescope and X-ray images (blue) from the Chandra Observatory. (right) Petrographs in Chaco Canyon likely depict Supernova 1054 (now the Crab nebula) near the top right corner and Halley’s comet faintly on the floor. Credit: <http://rockartblog.blogspot.com>

A pulsar is born when a massive star exhausts its supply of fuel and, without the opposing force of fusion to balance gravity, collapses in a powerful and violent supernova explosion³³. What is left behind is a ball of tightly-packed neutrons. Though

³²Jean Eilek, 2006, personal communication.

³³The supernova explosion that created the Crab nebula along with its central pulsar in AD 1054 was visible on Earth at high noon and remained discernible during daylight for 23 days. The appearance of a very bright “guest star” in the constellation Taurus was recorded by Chinese and Japanese astronomers around July 4, 1054 and noticed by Anasazi Indians as evidenced by petroglyphs in Navajo Canyon and White Mesa (both AZ) as well as in the Chaco Canyon National Park (NM) (Fig. 2.18).

the latter is only about 20 kilometers in diameter, it contains most of the progenitor's core – about 1.5 to 2 solar masses – making neutron stars the most compact objects in the observable universe apart from black holes³⁴. Due to conservation of angular momentum these neutron balls rotate rapidly, with spin periods ranging from 1.4 milliseconds to about eight seconds [93, 94].

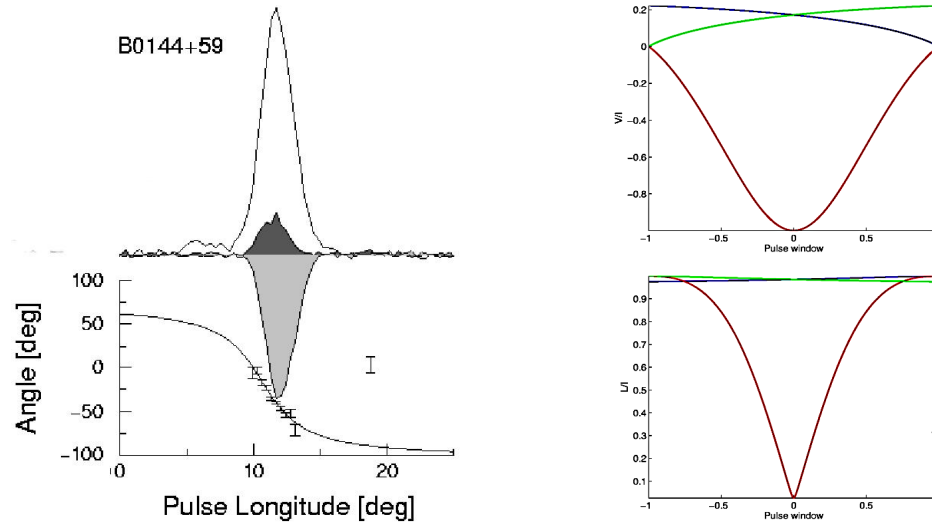


Figure 2.19: (left) Light curve, linear (dark gray) and circular (light gray) polarization, and polarization position angles for PSR B0144+59 [95]. (right) Linear (top) and circular (bottom) polarization for the same pulsar modeled according to Ardavan's superluminal model of pulsar emission.

Pulsars are exceptionally luminous. Their apparent brightness temperature, given by

$$T_b = c^2 S_\nu (2k_B \nu^2 \Omega)^{-1}, \quad (2.24)$$

where S_ν is the measured flux³⁵ at frequency ν , Ω the opening angle of the source

³⁴Indian astrophysicist Subrahmanyan Chandrasekhar showed that the theoretical upper limit to the mass of a white dwarf star – the *Chandrasekhar limit* – is approximately 1.4 solar masses. Above this limit, degenerate electron pressure is insufficient to prevent gravity from collapsing the star further to become a neutron star. If the *Oppenheimer-Volkoff limit*, generally believed to be about 2 to 3 solar masses, is exceeded as well, the collapse will result in a black hole [94].

³⁵Flux (or radiant flux), S , is the total amount of energy that crosses a unit area per unit

as seen from the Earth, and k_B denotes Boltzmann's constant, is generally between 10^{25} and 10^{30} K, but can be as high as 10^{39} K. These values are, of course, stupendously unrealistic unless some form of focusing mechanism or “coherent emission” is assumed. The central neutron star is surrounded by the magnetosphere through which a magnetic field of extraordinary strength (at the order of 10^7 to 10^{11} Tesla) rotates.

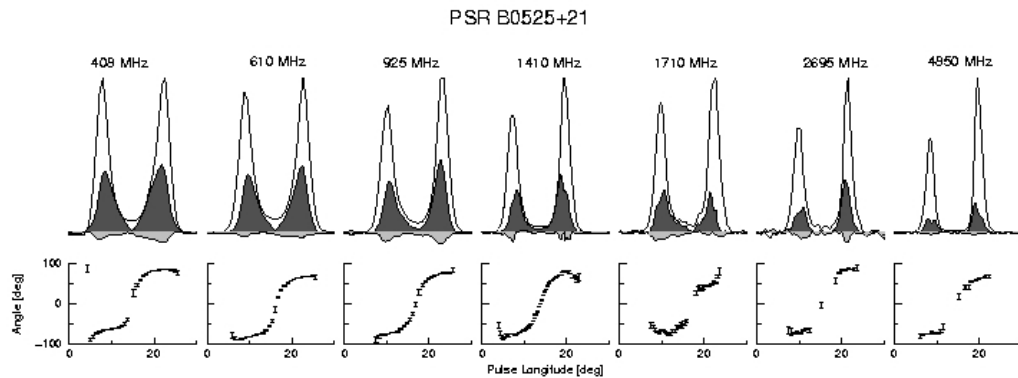


Figure 2.20: Light curves, linear (dark gray) and circular (light gray) polarization, and polarization position angles at different frequencies for PSR B0525+21 [95].

The electromagnetic waves emitted by pulsars are broadband with frequencies extending over an astounding 60 octaves from radio waves to high-energy gamma rays. In general, the observed light curves consist either of a single, sharp pulse (Fig. 2.19) or two closely spaced peaks (Fig. 2.20) and the integrated pulse profiles – constructed from some hundreds or thousands of individual pulses – remain surprisingly consistent across all frequencies. Moreover, pulsar emission is highly polarized – linearly as well as circularly – often with a change of sense through the pulse. Linear polarization is usually dominant; however, very high degrees of circular polarization are occasionally observed (mean values are typically 20 – 30 percent, but can be as high as 100 percent in individual pulses). The position angle of the

time. Flux is measured in Joules per square meter per second ($\text{Joules}/m^2/\text{s}$), or Watts per square meter (Watts/m^2).

radiation ³⁶ follows an S-shaped curve as a function of pulse longitude, often broken by one or more “jumps” (see Fig. 2.19 and the bottom panels of Fig. 2.20). Given these widely accepted findings, radio astronomer D. B. Melrose [96] posited in 1996 that any model seeking to explain the mechanics of pulsar emission must account for the characteristic light curves, the enigmatic polarization properties, and the broad frequency spectrum. However, he conceded that “we cannot expect a theory of pulsar emission to be quantitative in the sense that the theory of synchrotron emission is quantitative” and that “our theories are likely to be useful only in describing the most general features of pulsar emission – the ‘climate’ as opposed to the ‘weather’” [96].

2.5.3 Beyond the Light Cylinder

In the more than five decades since the discovery of spinning neutron stars a plethora of diverse models has been employed to explain the mechanism by which they radiate and to identify emitting regions. Excellent summaries of pulsar data and their interpretation are available [93,94]; yet even a brief perusal of these books will reveal that a comprehensive understanding of the observed radiation remains elusive. In fact, the authors of [94] state, “Despite many attempts to assemble and analyse the very detailed observational data, it is not yet possible to give a detailed explanation of pulsar radiation”.

Although much has been written about neutron stars in the scientific literature and text books, the possibility that the emitting regions might lay *beyond* the light cylinder has found little support, possibly due to the fact that many pulsar astronomers are “in awe of Einstein but poorly versed in electromag [sic]”³⁷. Recent Fermi observations [97,98] along with three-dimensional simulations [99–101], how-

³⁶For those unfamiliar with the concepts of Stokes parameters and position angles, definitions will be given in Chapter 6.

³⁷Jean Eilek, personal communication during the banquet concluding the 2010 New Mexico Symposium in Socorro.

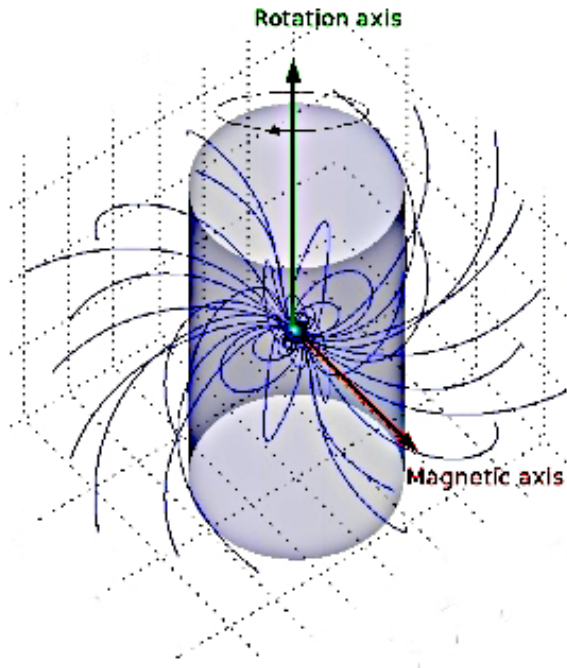


Figure 2.21: Depiction of the magnetic flux lines created by a rotating dipole such as a pulsar. It is essential to understand the following two points. (i) Electromagnetic forces are not instantaneous but are conveyed by virtual photons at the speed of light. The flux lines are a representation of the forces thus provided; their direction mimics that of the magnetic field, and their areal density indicates how strong the field is. Owing to the finite velocity of the virtual photons and the rotation of the source, the flux lines curve backwards. Nevertheless, the pattern of flux lines must rotate at the same angular velocity as the source (this is in fact a requirement of Maxwell's equations [80]). It is therefore unavoidable that the flux lines eventually cross the (velocity of) light cylinder shown in gray. (The latter is the virtual cylindrical surface at which the co-rotating magnetic field reaches the speed of light.) Outside the light cylinder, the tangential velocity of the flux lines therefore exceeds c . This is yet another example of a phenomenon that recurs throughout this document; the carefully timed arrival of a stimulus (in this case the virtual photons) produces a detectable phenomenon (here a magnetic field) that travels superluminally. Maxwell's equations also require that such lines of force must be closed (see the discussion of Eq. (2.4) in Section 2.4.2) although, for clarity, here the closure occurs outside the frame of the figure. Credit: P. Volegov, LANL

ever, seem to support such an argument. In agreement with Maxwell's equations, hydrodynamic calculations show that the pulsar's electromagnetic field distribution must rotate at the same angular velocity as the neutron star itself. Fig. 2.21 shows how the dipole field lines curve backward due to the rotation as they approach (and

pass) the light cylinder. As the forces are conveyed by virtual photons³⁸ traveling at a finite speed of c , the lines of force behave similarly to the jet of water emanating from a spinning garden sprinkler.

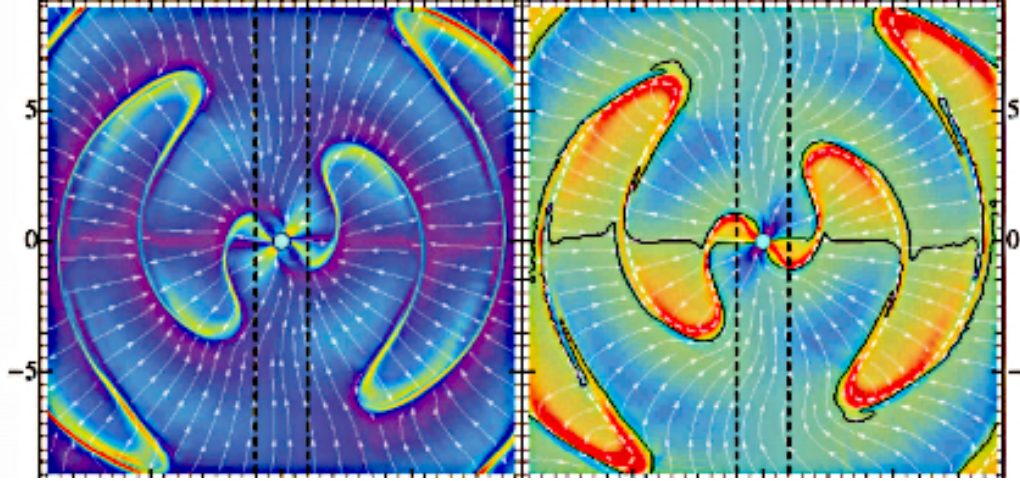


Figure 2.22: Numerical simulations [99, 100] by Constantinos Kalapotharakos, Ioannis Contopoulos and Demos Kazanas. (left) Meridional electric current density multiplied by r^2 (color plot) and direction (arrows). (right) Electric charge density multiplied by r^2 (color plot) and meridional magnetic field direction (arrows).

However, the figure does not include the effects of the pulsar's atmosphere, almost universally believed to be a plasma of electrons and protons along with a handful of positrons in the more energetic regions. The charged particles in the plasma will try to oppose any changes in the electromagnetic fields, an effect known as *screening* [53]. Hydrodynamic calculations demonstrate that this has two effects. (i) The field lines are further distorted and (ii) the remaining large disturbances tend to be in very concentrated regions of strong magnetic field (so-called *sheets*) that

³⁸Virtual photons are the exchange particles that convey electromagnetic interactions. They are transient quantum fluctuations with some of the characteristics of ordinary photons (*e.g.*, they travel at the speed of light c), but whose lifetime Δt is limited by the uncertainty principle $\Delta E \Delta t \gtrsim \hbar/2$. Here, ΔE is the uncertainty (or fluctuation) in energy and \hbar is Planck's constant divided by 2π . It follows that more energetic virtual photons live for shorter times, and, because they travel with speed c , cover a shorter distance during their existence. Hence, electromagnetic interactions get weaker as the distance grows. It is possible to derive the inverse-square law for electrostatic interactions thus [102].

are small in radial extent. Two examples of the results of hydrodynamic models are shown in Figs. 2.22 and 2.23. The latter (and hydrodynamic calculations in

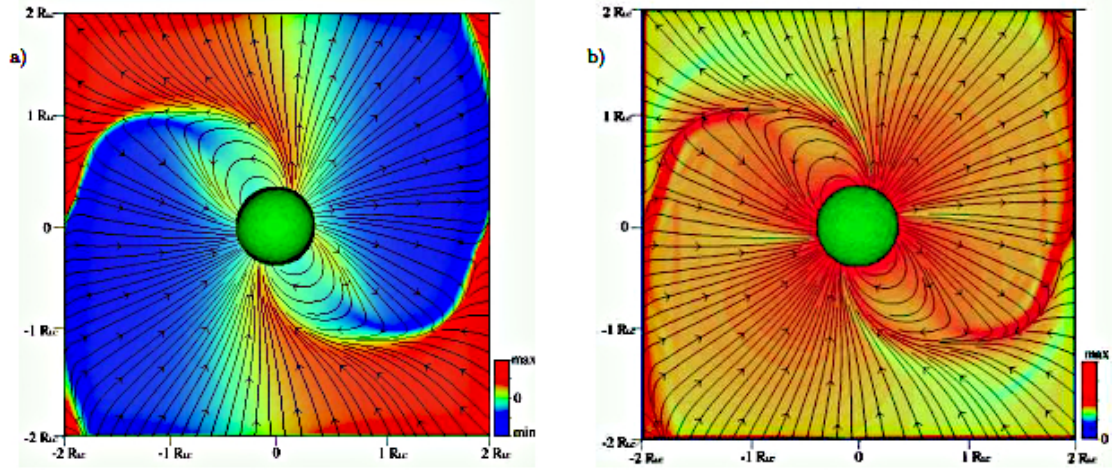


Figure 2.23: Numerical simulations by Anatoly Spitkovsky [101] showing the oblique pulsar magnetosphere with magnetic inclination $\alpha = 60^\circ$ in the corotating frame. (left) Magnetic field lines in the meridional plane; the magnetic field perpendicular to the plane is expressed in color. (right) Here, the color represents the absolute value of the total current $|\nabla \times \mathbf{B}|$.

general) reveal several phenomena that are germane to the current discussion. Firstly, electromagnetic disturbances (polarization currents³⁹) exist outside the light cylinder, and these disturbances rotate at the same angular velocity as the neutron star's magnetic field. As mentioned above, this rigid rotation is a requirement of Maxwell's

³⁹Although already treated in Section 2.4.4, the mechanism by which superluminal polarization currents radiate bears repeating: Polarization \mathbf{P} , defined as the dipole moment per unit volume, results from displacement of positive and negative charges in opposite directions; in this case this is due to the rapidly moving magnetic field. A *polarization current* forms when such a region moves or changes with time t , hence the polarization current density is $\partial \mathbf{P} / \partial t$ and has the same dimensions as a conventional current density of electrons, \mathbf{J} . If such a source distribution oscillates or accelerates, it will emit electromagnetic radiation, just as a current of charged particles does. However, unlike the latter, which possesses rest mass and is therefore limited to speeds which remain below that of light, polarization currents may travel arbitrarily fast because the displacement of their constituent negative and positive particles is negligible; albeit the *pattern* travels faster than c , the *particle* velocities are subluminal. Such a source distribution does not violate special relativity as *it cannot be used to transmit a signal superluminally*; the emitted radiation – as any other – travels at (or below) the speed of light.)

equations, and ensures that any disturbances must travel superluminally at radii outside the light cylinder. Secondly, the most intense disturbances are compact, in that they occupy a small fraction of the pulsar’s atmosphere.

Thus far, the astrophysics community has tended to avoid calculations of the emission of these superluminal disturbances. By contrast, Ardavan’s superluminal model of pulsar emission makes no attempt to explain the origin of the superluminal polarization currents, but instead calculates their emission. The model’s primary assumptions are (i) that the source of radiation is a point-like (*i.e.*, much smaller than any of the other length scales of the pulsar) region of polarization current outside the light cylinder; (ii) this emitting region moves in a circular orbit about the pulsar with the same angular velocity as the rotating neutron-star magnetic field; and (iii) its instantaneous speed must therefore be superluminal.

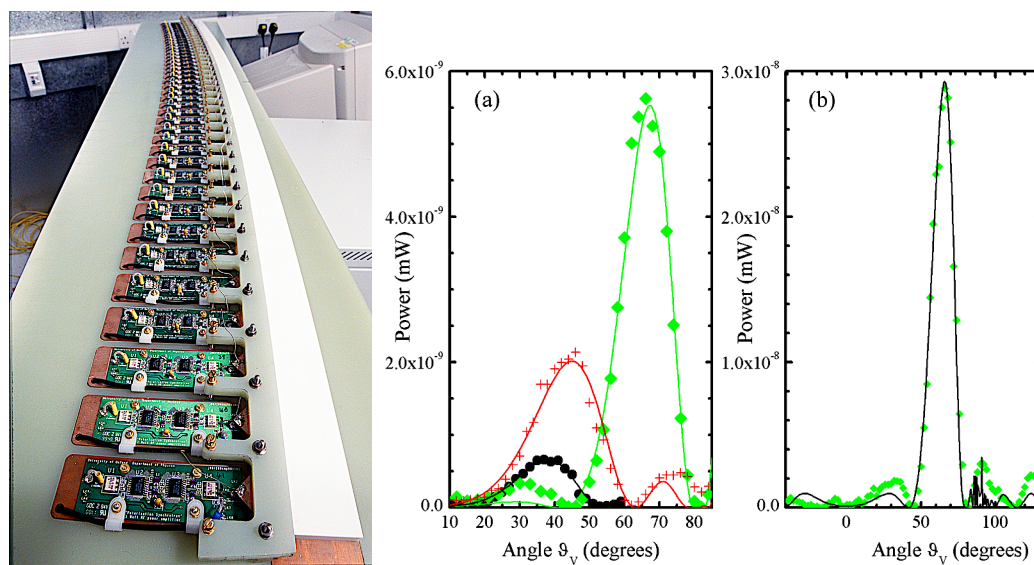


Figure 2.24: (left) Practical superluminal light source. A curved strip of dielectric material is placed between a continuous ground plate and an array of metal electrodes, each of which is connected to an individual amplifier. By turning the amplifiers on and off in sequence one can generate a polarized region that moves along the dielectric at arbitrarily high speeds. (right) Since the source is traveling faster than the speed of light *in vacuo*, it emits “vacuum Čerenkov radiation” at an angle that depends on the speed of the source alone: The higher the source speed, the narrower the cone.

Despite the empirical nature of Ardavan’s model, a comparison of the two previous paragraphs shows that the source assumed in the model matches the attributes of the disturbances predicted by the hydrodynamic simulations quite closely. Chapter 6 describes numerical calculations of the emission of a point source in superluminal rotation and concludes with a comparison of the predictions with actual pulsar data; the agreement is found to be rather convincing.

2.6 “... a waste of tax-payers’ money”

In 2002, John Singleton of the Physics Department of Oxford University, his post-doctoral student, Arzhang Ardavan, and Houshang Ardavan of the University of Cambridge, built the first man-made superluminal light source (Fig. 2.24 (left)). The device was based on the older Ardavan’s theoretical work in pulsar astronomy and the acoustic phenomena elicited by rotating helicopter blades; its constitutes, in effect, the first ground-based simulation of a rotating neutron star. The 2-meter-long device was mounted on a scissor lift and tested on a runway at Turweston Aerodrome in Northampton after air traffic had shut down for the night. The construction of the “polarization synchrotron,” as the machine was called, soon led to numerous articles in the scientific [103–119] and popular press, many of which were neither particularly flattering nor scientifically accurate. “Money Spinner or Loopy Idea?” Edwin Cartlidge asked, somewhat ominously, in *Science* [120], while Martin Durrani of *Physics World* [121] stated the obvious by noting that the “revolutionary device polarizes opinions.” Anthony Hewish, who shared the 1974 Nobel Prize in physics for the discovery of pulsars, declared the device “a waste of tax-payers’ money [122],” claiming cantankerously that the physics “is nonsense” and “simply wrong... The radiation from such a device must be conventional.”

The vitriolic press notwithstanding, subsequent experimental and theoretical work conducted at Los Alamos National Laboratory and in the UK demonstrated un-

ambiguously that polarization currents can indeed be animated to travel faster than the speed of light *in vacuo*, thereby opening a promising and largely unexplored field in electrodynamics (Fig. 2.24 right). In parallel, a team under A. V. Bessarab used the ISKRA-5 laser at Sarov to demonstrate emission by superluminal polarization currents, verifying the fundamental physics involved [49].

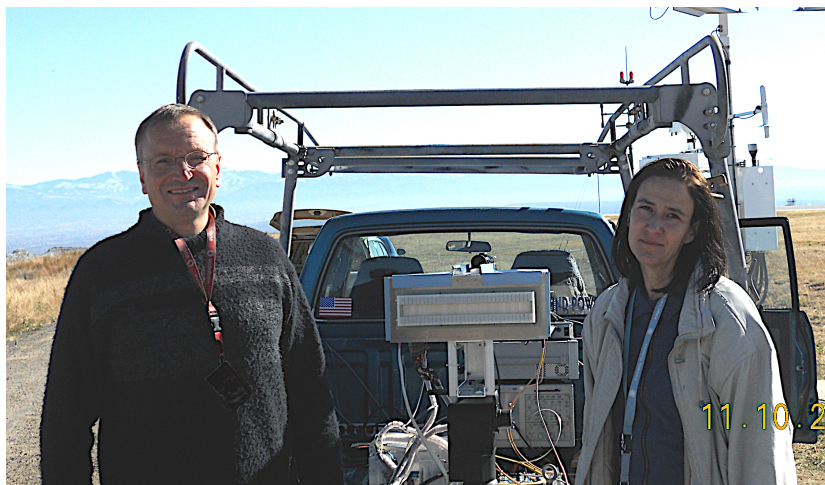


Figure 2.25: John Singleton and Andrea Schmidt at Los Alamos Airport in November 2011, after the first data transmission by a superluminal antenna (center). Herbert Howells’ *Magnificat* as well as excerpts from Louis Vierende’s *Sanctus* and *Agnus Dei* were broadcast across roughly 500 meters of airstrip.

Since 2007, several second-generation practical superluminal sources have been designed, constructed and tested in Los Alamos [67, 68]; an example is shown in Fig. 2.25. Whilst the competing Russian group demonstrated the feasibility of superluminal emission using polarization shock waves in a plasma generated by a photolytically pumped iodine gas laser, the operating principle of the machines in Los Alamos, as well as their British predecessor, is based on electrostatic control and animation of the polarization current; a technique far more amenable for useful and controllable devices. The emission from two of the Los Alamos superluminal antennas is modeled later in this dissertation; therefore the following chapter describes these devices in considerably more detail.

Chapter 3

Meet the Lightslingers: Superluminal Sources in the Laboratory

My congratulations for the definite success.

— Nobel laureate V. L Ginzburg upon learning of the completion of the proof-of-concept superluminal source built at Oxford University.

3.1 Second Prologue

Theoretical work on faster-than-light sources of radiation conducted in Russia and the United Kingdom culminated in the construction of a prototype antenna (Chapter 2 and Fig. 2.24). Whilst the original apparatus was the first of its kind and useful as a proof-of-concept, it was large, unwieldy and based on analogue technology [43, 44]. Subsequently, several technology demonstrators for superluminal emission have been (and are being) built at Los Alamos National Laboratory. The properties of those relevant in the present context are summarized in Table 3.1 while photographs of three of these machines are shown in Fig. 3.1. Four of the anten-

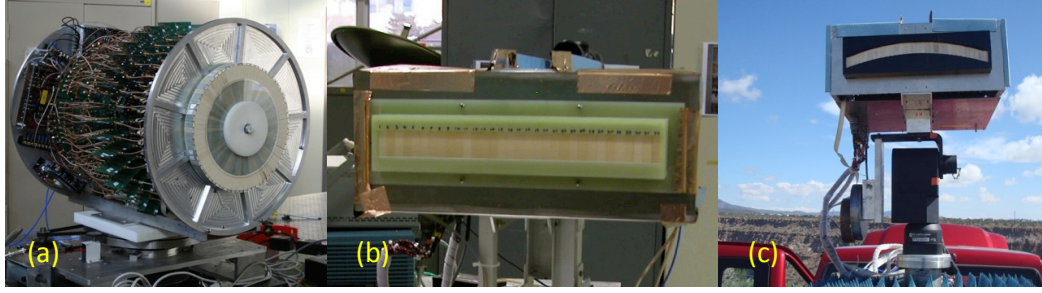


Figure 3.1: Photographs of Technology Demonstrators 1–3. (a) TD 1 (circular), mounted on its rotation table, but with its cylindrical copper outer casing removed to show the circuit boards. (b) TD 2 (linear) in the course of assembly, before its outer casing was fixed. (c) TD 3 (arced) fixed to the pan/tilt mount and mounted on a truck for outdoor tests. In each case, the pale grey alumina that hosts the polarization current is visible on the front of the antenna. The joints visible in the alumina of all three machines are a result of the antennas being built up from modular antenna elements.

nas are *active*, *i.e.*, they contain integral electronic components that synthesize the signals to be transmitted. Computer control allows signal modification on the fly.

Although the author of this dissertation was not actively involved in the design and construction of the Los Alamos faster-than-light sources, she participated in all of their experimental tests and played a large part in visualizing and analyzing the emitted radiation. Moreover, a major concern of this dissertation is to derive a theoretical framework for vacuum Čerenkov radiation resulting from superluminal polarization currents; to demonstrate the validity of this approach, its predictions must be compared with experiments in a quantitative manner (Chapter 7). Bearing in mind the skepticism about superluminal antennas reported in Chapter 2, we provide here sufficient detail of the antennas and measurement techniques to give confidence in the data that are to be simulated. Owing to the fact that patent protection issues have prevented details of the antennas from being published thus far, an appendix with further design details¹ is provided for completeness.

¹The appendix is a summary of as-yet unpublished work carried out by John Singleton, Zhi-fu Wang, Frank Krawczyk, Bill Romero and Quinn Marksteiner.

Table 3.1: The superluminal antennas built at Los Alamos; TD stands for Technology Demonstrator. The name *Lightslinger* was coined by poet, artist, publicist and friend Jabez van Cleef, and used by Los Alamos Technology Transfer for publicity purposes.

Machine	Form	Elements	Purpose
TD 0	40° arc 125 mm radius	8	Test-bed for TD1 components
TD 1 <i>Lightslinger</i>	Full circle 125 mm radius Active control	72	Astrophysics experiments Long-range communications Phase-front studies (radar)
TD 2	Linear accelerator Active control	32	Directed energy. Communications
TD 3	Arched antenna Active control	32	Directed energy Communications
TD4	Linear antenna Passive	32	Communications

The designs of the technology demonstrators were motivated by several considerations. Firstly, they were to be smaller than the proof-of-concept machine (Fig. 2.24), illustrating the scalability of the technology. Secondly, they had to be robust and versatile enough to demonstrate proposed communication and radar applications outdoors. Finally, they would serve as ground-based astrophysics experiments, simulating pulsars and gamma-ray bursts (Chapter 2). In the case of the active antennas, digital signal synthesis was to be used to animate the polarization current.

Though outdoor durability has been emphasized, the compactness of the new antennas has also allowed them to be run in the well-characterized environments of radio-frequency anechoic chambers at Kirtland Air Force Base, Sandia National Laboratories and, latterly, in Building 125 at the National High Magnetic Field Laboratory's Los Alamos campus. By contrast, the large size and relatively long wavelength (60 cm to 1 m) of the original apparatus meant that it could only be

used outside, resulting in the complication of ground reflections [43, 44].

Most of the calculations described in Chapter 7 refer to circular and linear superluminal antennas. Therefore this chapter focuses on descriptions of TD 1 and TD 2. Sections 3.2 to 3.6.4 are highly edited extracts from two reports prepared by the current author and colleagues for the telecommunications company CommScope [68, 69].

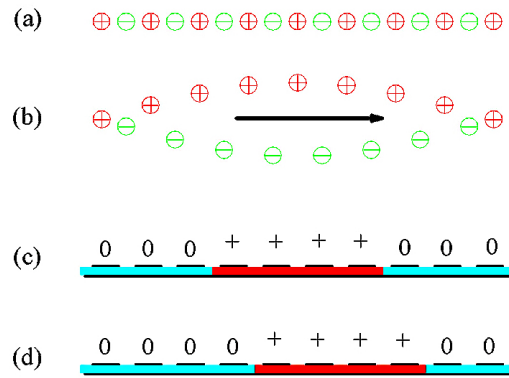


Figure 3.2: Experimental animation of a superluminal polarization current. (a) A simplified dielectric solid containing negative (\ominus) and positive (\oplus) ions. In (b), a spatially-varying electric field has been applied, causing the positive and negative ions to move in opposite directions; a finite polarization \mathbf{P} has therefore been induced. If the field is made to move along the direction of the arrow, the polarized region moves with it. (c) Schematic side view of a practical superluminal source, showing metal electrodes above a strip of dielectric (shaded region) and a ground plate below it. “0” indicates that there is no voltage on that particular upper electrode; the symbol $+$ indicates that a positive voltage is applied. The voltage on the electrodes produces a finite polarization of the dielectric (red shading). (d) By switching the voltages on the electrodes on and off, the polarized region can be made to move along the dielectric.

3.2 Fundamental Principles

The two principles that underpin all electrostatically controlled superluminal antennas are (i) controlled animation of the polarization current and (ii) modular con-

struction. Considering first the former, it is useful to restate Eq. (2.22):

$$\nabla \times \mathbf{H} = \mathbf{J} + \varepsilon_0 \frac{\partial}{\partial t} \mathbf{E} + \frac{\partial}{\partial t} \mathbf{P}.$$

As noted in Chapter 2, the generation of electromagnetic radiation is encompassed by the source terms \mathbf{J} (the current density of free charges) and $\partial \mathbf{P} / \partial t$ (the *polarization current density*); for superluminal emission we of course focus on the latter. The basic principles underlying the production and control of the polarization current are shown in Fig. 3.2. A dielectric solid such as alumina contains positive and negative ions (Fig. 3.2(a)); under the action of an applied electric field, these are displaced in opposite directions, resulting in a finite polarization \mathbf{P} (Fig. 3.2(b)). In the antennas, the electric field is provided by placing a series of electrodes on top of a solid dielectric mounted on a ground plate, or a series of opposing electrodes above and below (Fig. 3.2(c)). The application of voltages to the electrodes creates a polarized region underneath which can then be moved by switching the voltages on the electrodes on and off (Fig. 3.2(d)). The motion imparts a time dependence, leading naturally to a polarization-current density, $\partial \mathbf{P} / \partial t$.

Given the sizes of practical devices used to provide data for analysis (electrode spacing $\sim 1 - 3$ cm), a back-of-the-envelope calculation shows that a wide range of superluminal speeds can be achieved using timings in the 10s – 100s of picosecond range. Far more subtle manipulation of the polarization current is possible by controlling the magnitudes and timings of the voltages applied to the electrodes. In the antennas to be described, the electrode voltages are usually chosen to give a modulated, moving polarization wave of the generic form:

$$\mathbf{P} \propto \cos(\mathbf{k} \cdot \mathbf{r} - \omega t) \cos(\eta t). \quad (3.1)$$

Here, the term $(\mathbf{k} \cdot \mathbf{r} - \omega t)$, and in particular the wave vector \mathbf{k} , describes the motion (speed, direction, acceleration) of the polarization current wave as a function of position \mathbf{r} within the antenna whilst the cosine containing ηt provides modulation.

Turning to the idea of modular design, the individual amplifiers driving the electrodes of the proof-of-concept superluminal source are clearly visible in Fig. 2.24, illustrating how identical units can be used to control the polarization current. The Los Alamos antennas build on this approach by using common electronic modules for all the active machines. The dielectric region and electrodes of *all* of the antennas, including the passive example TD 5, are constructed from large numbers of identical elements as well, leading to streamlining of design and assembly work. These modular dielectric elements are visible in the examples shown in Fig. 3.1.

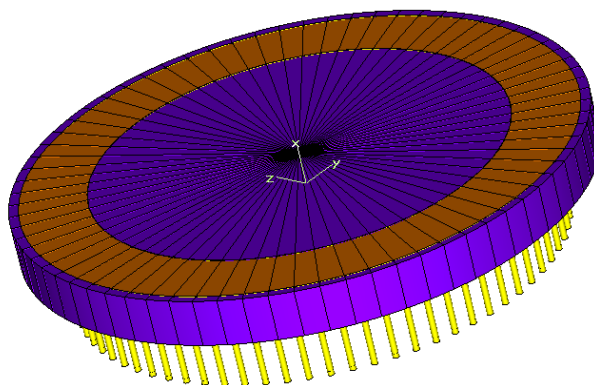


Figure 3.3: Full model of the antenna TD 1. It consists of 72 wedge-shaped elements, each spanning 5° . The structure is driven by 72 individual $50\ \Omega$ coaxial feeds (yellow) from the bottom. The polarization current inhabits the alumina ring (brown) on top of the antenna. Credit: Frank Krawczyk

3.3 Circular Accelerator TD 1

Chapter 2 relates how interesting focusing phenomena occur when a source both accelerates and moves faster than the waves that it emits. In response to this idea, antenna TD 1 (*The Lightslinger*) constrains the polarization current's motion to a circular path, resulting in centripetal acceleration (Fig. 3.1(a)). All of the electronic control modules and dielectric elements used in subsequent antennas were modeled

after TD 1 and so this antenna will be described first, with some design details being relegated to Appendix B.

TD 1 comprises a full circle of dielectric (*i.e.*, spanning 360°), made up of 72 discrete elements; a photo is presented in Fig. 3.1(a) and a schematic shown in Fig. 3.3. Components were tested using the 8-element prototype TD 0, which spans a 40° arc of a 0.125 m radius circle, representing 1/9 of TD 1.

A single mode of a polarization distribution that is both rotating and oscillating in time is described by

$$\mathbf{P}(r, \varphi, z; t) = \mathbf{s}_{r,\varphi,z}(r, z) \cos(m\varphi - \omega t) \cos(\eta t). \quad (3.2)$$

Here, \mathbf{s} describes the direction of polarization, the first cosine term causes circular motion and the second one provides modulation of the entire waveform. Any periodic, rotating and oscillating source can be built up from a superposition of terms like Eq. (3.2) with different \mathbf{s} , m and ω . However, practical machines are not continuous circles, but built up from elements subtending a few degrees, each energized by a separate amplifier as shown in Fig. 3.1, Fig. 3.3, and Table 3.1). In TD 1, each element² or “wedge” subtends an angle $\Delta\varphi = 2\pi/72$ radians = 5° . Hence, a discretized version of Eq. (3.2) is used to give the voltage applied across the electrode pair of the j^{th} element [67, 68]

$$V_j = V_0 \cos(\omega(j\Delta t - t)) \cos(\eta t). \quad (3.3)$$

As above, ω is an angular frequency providing the motion of the polarization current, and η a master modulation angular frequency. Fig. 3.4 illustrates how this approach can simulate a smoothly rotating source. Simulations of the antenna elements carried out using Microwave StudioTM show that fringing effects round out the sharp edges of

²It is worth reminding ourselves that the elements are not in themselves the emitters (these antennas are *not* phased arrays); instead they apply the voltages that induce the polarization current in the dielectric. *The volume polarization current that inhabits the whole of the dielectric emits the radiation.*

the steps in the applied electric field, producing a very smooth waveform, like that shown in grey in Fig. 3.4. Note that there is a subtle difference between a full circle machine such as TD 1 and an antenna that is only the arc of a circle such as TD 3 (Fig. 3.1). For the former, the rotating pattern should be continuous (*i.e.*, contain no discontinuities) both spatially and in the time domain; this restricts m to integer values. The latter have no such restriction, and m can take any value.

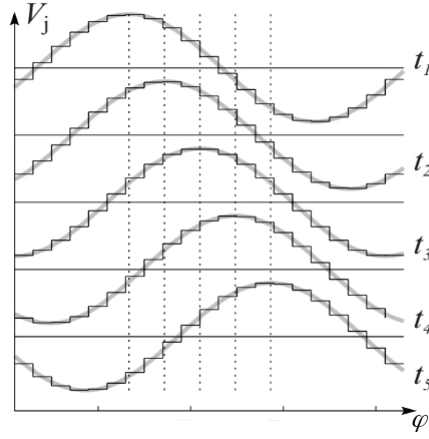


Figure 3.4: The voltage V_j (Eq. (3.6)) on each electrode pair versus the φ coordinates of the electrodes at five equally-spaced consecutive times ($t_1 < t_2 < t_3 < t_4 < t_5$). For clarity, the curve has been plotted for $\omega \gg \eta$. The vertical dotted lines designate the consecutive positions at which V_j is maximum. The sinusoidal curves represent the fundamental Fourier component of the discretized voltage distribution at various times. The constant phase difference between adjacent elements results in propagation of the sinusoidal voltage distribution at a constant speed.

The mean speed of the source (*i.e.*, the instantaneous velocity of the rotating polarization pattern) is $v_{\text{rot}} = \rho_0 \omega$, where ρ_0 is the mean radius of the antenna; this can also be expressed as

$$v_{\text{rot}} = \rho_0 \frac{\Delta\varphi}{\Delta t}. \quad (3.4)$$

The electronics used to generate these signals in TD 1 (discussed in Appendix B.3) set the *phase* of the signal, and not the time delay. The relation between the time

delay and the phase is given by

$$\Delta t = \frac{\Delta\phi}{360f}, \quad (3.5)$$

where $\Delta\phi$ is the phase difference between adjacent elements, in degrees, and $f = \omega/2\pi$ is the carrier frequency.

3.4 Linear Accelerator TD 2

The linear superluminal antenna TD 2 is shown in Fig. 3.1(b). It consists of 32 elements arranged in a straight line; these elements are rectangular versions of the wedges used in TD 1, preserving the same overall dielectric volume and input impedance. As in TD 1, for ease of maintenance, each element contains its own block of alumina dielectric ($\epsilon_r \approx 10$). The center-to-center spacing of the elements is $a = 10.87$ mm.

As before, in order to produce a travelling polarization distribution, the j^{th} ($j = 1, 2, 3, \dots$) element is supplied with a voltage

$$V_j = V_0 \cos[\omega(j\Delta t - t)] \cos \eta t. \quad (3.6)$$

The first cosine term in Eq. (3.6) results in propagation of the voltage, and hence the polarization, along the antenna; the situation is similar to that shown in Fig. 3.4 except that in this case we have *linear*, rather than orbital, motion. The velocity v with which the polarization current distribution propagates is set by adjusting Δt to give $v = a/\Delta t$.

Further design details of TD 2 are given in Appendix C.

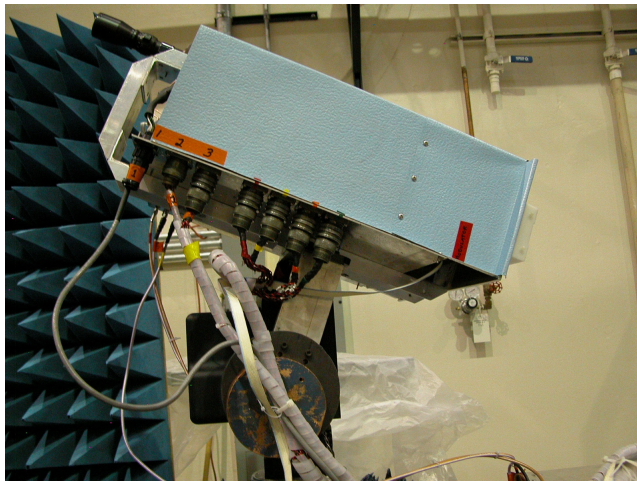


Figure 3.5: Side view of the TD 2 showing the telescope mount that rotates and tilts the antenna. Azimuthal (pan) rotation is about a vertical axis that goes through the center of mass of the antenna when it is horizontal (tilt angle $T = 0$). Polar (tilt) rotation occurs about a pivot point in the center of the black box, which contains the tilt stepper motor and its gearbox. The antenna is shown tilted to $T \approx -20^\circ$. Visible at the rear of the antenna is the telescopic sight used for aligning the antenna along a particular direction. Note that both rotation axes (pan and tilt) are offset by 10s of cm from the geometrical center of the dielectric on the front of the antenna.

3.5 Anechoic Chamber and Test Measurements

3.5.1 Angular Scans

Most of the experiments involving the antennas TD 1 and TD 2 involve mapping the angular distribution of the emitted radiation. For this, TD 2 is mounted on a telescope mount (Directed Perception PTU-D300) that allows polar and azimuthal angles to be adjusted using computer-controlled stepper motors (see Fig. 3.5). Usually, the pan (azimuthal) angle is scanned from -90° to $+90^\circ$ in 0.18° steps. Additional polar angle studies involve setting the tilt angle at fixed values, separated by 5 or 10° . At each tilt angle, the pan angle is again scanned. Note that the telescope mount does not permit polar angles greater than $+30^\circ$.

The radiation from antenna TD 1 is circularly symmetric, and so its angular distribution is measured using a turntable driven by a DC motor. The angular position is monitored using a 10-turn helical potentiometer mounted in the drive gear train. Angular accuracies of about $\pm 0.2^\circ$ are achieved.

3.5.2 Detecting the Radiation

Experiments with antennas TD 1 and TD 2 were carried out in three anechoic chambers: (i) the $20 \times 20 \times 20 \text{ m}^3$ chamber at the Air Force Research Laboratory (AFRL), Kirtland Air Force Base, Albuquerque; (ii) an $8 \times 8 \times 8 \text{ m}^3$ section of the chamber at the FARM Range, Sandia National Laboratories, Albuquerque; and (iii) a $4 \times 4 \times 8 \text{ m}^3$ chamber in Building TSL 125, TA-35, Los Alamos National Laboratory (LANL). All three chambers have walls and ceilings made from metal sheet that acts as a Faraday cage, shielding the experiment from outside RF interference. Inside the chambers, the floor and metal walls and roof are completely covered with RF absorbing tiles. In the case of the AFRL and LANL chambers, these tiles are ECCOSORB VHP-NRL, from Emerson & Cuming Microwave Products. The tiles are made up of a foam base covered in 30 cm long pyramids that are coated in a lossy dielectric, reducing indirect radiation in the chambers by about 50 – 60 dB. In other words, the reflections make up a fraction $\sim 10^{-5}$ of the detected signal. The FARM range chamber is more sophisticated, in that the pyramids vary in size according to a quasi-random pattern; this has the effect of reducing reflected radiation by another 20 dB or so. Hence, experiments in such chambers mimic infinite-free-space measurements rather closely. Fig. 3.6 shows views of the insides of the AFRL and LANL chambers; the cones of absorber can be seen.

Most of the experiments in the anechoic chambers employed one or two calibrated Schwarzbeck Mess-Elektronik UHA 9125 D dipole (hereafter “Schwarzbeck dipole”) antennas as receivers, mounted on a Panasonic RF dielectric tripod on top of a stage

that could be moved along fixed rails. The dipole(s) was/were positioned level with the center of the superluminal antenna under study, and the rails arranged such that the tripod moved along a line that is orthogonal to the array when the pan and tilt angles of the turntable or telescope mount were set to zero. The movement of the dipole(s) to different distances from the superluminal antenna under study used the calibration of the rail system as a primary distance gauge. The distance was also verified using a surveyor's tape measure.

Some measurements employed a calibrated Q-Par QSPDCP2-18SRA dual-circularly polarized antenna in place of the Schwarzbeck dipoles. This antenna is virtually omnidirectional and very broadband (2–18 GHz); it is visible (grey radome) at the rear of the right-hand photograph in Fig. 3.6. By contrast, the Schwarzbeck dipoles must be tuned by fitting stubs and adjusting the balun for different frequency ranges. Consistent data were obtained with the different detection schemes.

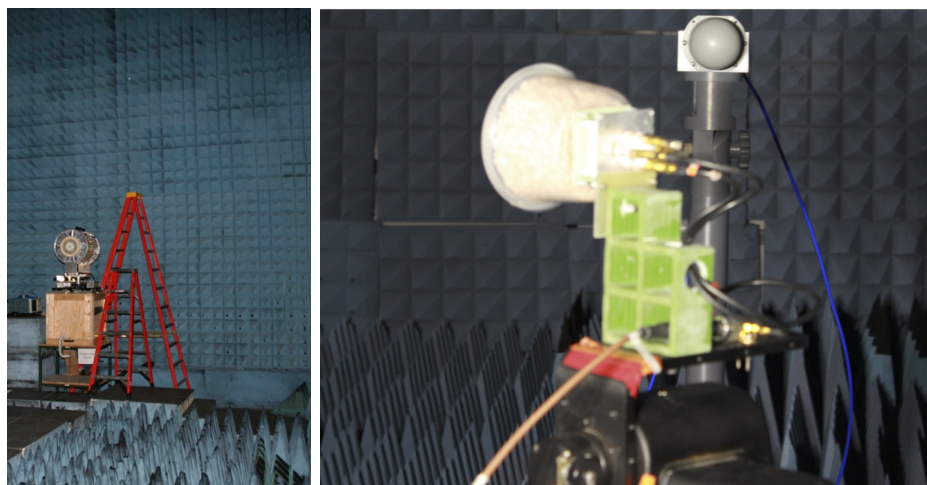


Figure 3.6: Photographs of two of the anechoic chambers used in this dissertation. (left) The $20 \times 20 \times 20 \text{ m}^3$ chamber at Kirtland Air Force Base: TD 1 is in the process of being assembled for tests in the left of the picture. (right) Measurement of a small antenna (foreground) in progress in the $4 \times 4 \times 8 \text{ m}^3$ chamber in TSL 125, TA-35, Los Alamos National Laboratory.

Two calibrated spectrum analyzers (Hewlett-Packard HP8594E and HP8595E)

were used so that either two emission frequencies or two polarizations could be monitored simultaneously. Each spectrum analyzer was set to scan across a 100 kHz span around its chosen detection frequency. At each value of the pan/tilt angle, the peak power detected within this span was measured and transmitted to the control computer. Using peak detection, rather than detection at a fixed frequency, compensates for any slight drift in the emitted frequencies that might occur as ambient conditions change; this is a more serious concern outside, where temperature variations of up to $\pm 10^\circ\text{C}$ were experienced.

In most experiments, the spectrum analyzers were mounted on the same mobile stage as the tripod bearing the detection antennas³. This minimized cable lengths; improperly protected long cables can contribute a parasitic signal. The equipment was shielded with the same ECCOSORB VHP-NRL tiles as used on the chamber walls to prevent interference by stray signals. Communication between the spectrum analyzers and the data-logging computer was maintained using a GPIB interface via a fiber-optic cable running under the absorber tiles on the floor.

3.5.3 Into the Wild

Some experiments were carried out at the FARM Outdoor range at Sandia National Laboratories, which comprises a triangular-cross-section concrete track about 1200 feet long (Fig. 3.7). The triangular cross-section directs some of the ground-reflected radiation away from the detector antenna. The equipment used for these tests was very similar to that employed in the anechoic chambers; in place of the rail arrangement, a specially configured detector truck ran along the apex of the triangular track, carrying the detection antenna and spectrum analyzers. Convenient

³The FARM Anechoic Chamber was the exception; the receiver dipole antennas were mounted on a stepper-motor controlled base on rails which was too small to accommodate the spectrum analyzers. Hence relatively long cables were used to connect the dipoles to the analyzers.

(that is, if you have a 300 foot extension cord) mains outlets are provided at various stages. As before, data were fed back to the control computer using a fiber optic link. Parts of the fiber-optic cable were eaten by pack rats, gophers or raccoons one night, necessitating the rolling up of the fiber on subsequent evenings. Receiver to transmitted distances were measured using survey markers and checked using surveyors' tapes.



Figure 3.7: The FARM Outdoor Range at Sandia National Laboratories. The range has a triangular cross-section concrete surface, designed in an attempt to direct ground-reflected radiation away from the receiver antenna. Antenna TD 2 is in the foreground on its telescope mount; the detector truck, carrying two Schwarzbeck dipoles and two spectrum analyzers, is in the distance. A third antenna is off to one side, used to characterize radiation reflected from the concrete of the range.

In all of the other outdoor tests, the superluminal antenna was placed close to the Los Alamos Airport Weather Station (Fig. 3.8), which provided a convenient source of mains electricity plus an environment relatively free of clutter and with good sight lines to many locations. These experiments employed two receiver antennas, each equipped with a spectrum analyzer set to record the same frequency. One receiver was a Schwarzbeck dipole, placed 20 m from the superluminal antenna under test, connected to a single spectrum analyzer interfaced to the control computer by fiber-optic cable. During angular scans, the control computer recorded the azimuthal angle, the peak power at the chosen frequency and the time. The second antenna

was remote from the superluminal antenna. Hence, the spectrum analyzer at this remote location was interfaced to a second computer, with its clock synchronized to that of the control computer; data recorded were time and peak power. As the clocks of the two computers were synchronized, it was possible to convert time to azimuthal angle for files recorded at the remote computer. In this way, a mapping of power versus azimuthal angle was obtained both 20 m away from the superluminal antenna under test, and at the remote site. Detection equipment there was powered by a 900 W inverter driven by the support truck battery (and the gasoline motor connected to it).



Figure 3.8: View of the transmission site at Los Alamos Airport (the small white boxes just visible to the left of the far end of the runway are the support vehicles parked by the Weather Station) from the receiver site on Pajarito Ski Hill, 10.5 km away.

The remote detector was positioned at four separate locations: (i) the turning circle at the end of Big Rock Loop, Los Alamos, 950 m from the Airport Weather Station; (ii) behind 1015 Los Pueblos, Los Alamos, 2,250 m from the Airport Weather Station; (iii) a clearing close to the base of the Townsight Lift of the Pajarito Ski Hill, 10,500 m from the Airport Weather Station; and (iv) the “10,000 foot” parking lot on

Sandia Peak, 76 km from the Airport Weather Station, and well shielded from the numerous TV, backhaul and cellular transmitters at Sandia Crest by the intervening mountain. Distances were measured using an accurate topographical map (1:37,500 scale) and checked using USGS maps and data from Google Earth. For the two closest stations, it was possible to measure distances to about ± 20 m.

For initial experiments at the second location, the remote antenna was a Schwarzbeck dipole with a mast-head amplifier (MiniCircuits ZX60-33LN-S+) giving an additional 15 dB of gain. An Andrew D4G-1-GR 4 ft dish with bipolar 112875-S4N23 feed, giving 27.3 dB of gain, was used at all other times. The Andrew dish antennas are mounted on stands with pan/tilt adjustment screws. They were first aligned using a high-power telescopic sight and then fine-tuned using the detected signal from a transmitter dipole alongside the superluminal antenna at the Airport.

For ease of reference, the antennas are listed in Table 3.2.

Table 3.2: Detector antenna gains and cable losses for the various experimental locations. The correction in the final column is applied to the raw data so that they can be compared with model calculations in a quantitative manner. It was necessary to use the FARM anechoic chamber’s own long cables for measurements there, resulting in the significant loss listed.

Location	Detector	Antenna gain (dB)	Cable loss (dB)	Correction (dB)
TA-35 Anechoic	Q-Par circular	≈ 0	1.4	+1.4
TA-35 Anechoic	Schwarzbeck dipole	2.1	1.4	-0.7
AFRL Anechoic	Schwarzbeck dipole	2.1	2.0	-0.1
FARM Anechoic	Schwarzbeck dipole	2.1	6.3	+4.2
FARM Outdoor	Schwarzbeck dipole	2.1	1.4	-0.7
Outdoor (2,250 m)	Schwarzbeck dipole + mast-head amp	17.1	1.4	-15.7
Outdoor (other)	Andrew dish + bipolar feed	27.3	2.0	-25.3

3.5.4 Power and Angle Corrections

Adjusting the Power

In order to compare experimental data with the theoretical models described in subsequent chapters, power measurements were corrected for losses incurred in the cables and splitters, and for additional gain due to the dish antenna or masthead amplifier. Corrections are listed in Table 3.2.

The Antenna Coordinate System

The coordinates that describe the superluminal antenna's position with respect to the detection/observation point must be expressed in a modified *spherical polar* system to match the model calculations in Chapter 7. Complications arise because the rotation axis of the turntable used with TD 1 and the pan/tilt axes of the telescope mount for TD 2 do not pass through the geometrical center of the antenna dielectric, but are offset by distances of 10–60 cm, depending on the experiment (Fig. 3.5, Fig. 3.7, Fig. B.11). Therefore, especially at small distances (*e.g.*, in the anechoic chambers), the following corrections must be made. Firstly, there must be a transformation between the pan angle P and tilt angle T (Fig. 3.9) of the turntable or telescope mount to the azimuthal angle ϕ and the polar angle θ of the model coordinate system. Secondly, the fact that the distance between the detector and the geometrical center of the antenna's dielectric varies during rotation must be taken into account.

In the final (model) coordinate system, the separation of the geometrical center of the dielectric from the detection antenna is described by the vector $\mathbf{R} = (R, \phi, \theta)$. To match antenna engineers' coordinate definitions, $\theta = 0$ is when the superluminal antenna rotates in the horizontal (x, y) plane ($T = 0$). Positive θ indicates that \mathbf{R} points above the $\theta = 0$ plane. Finally, when comparing the polarization of the experimental data with models, it is useful to define Cartesian unit vectors $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$,

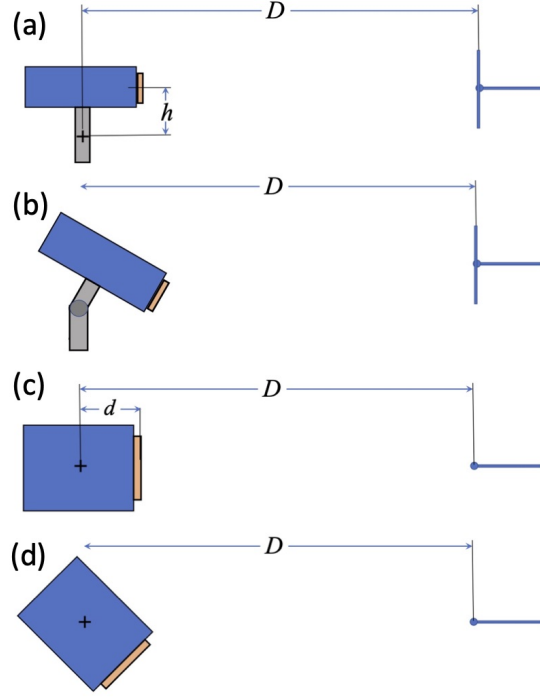


Figure 3.9: Schematic of TD 2 (left) and a dipole receiver (right) positioned for an experiment; the distances d , h and D , and the effects of pan angle P and tilt angle T are shown. The blue box symbolizes the case containing the control electronics; the orange region is the dielectric antenna on the front (*c.f.*, Fig. 3.5). (a) Side view for $T = P = 0$; the black cross represents the tilt axis. (b) Side view for $T = +30^\circ$, $P = 0$. (c) Plan view for $T = P = 0$; the black cross is the pan axis. (d) Plan view for $T = 0$, $P = +45^\circ$.

corresponding to the (x, y, z) directions, for the model coordinate system:

$$\mathbf{R} \cdot \hat{\mathbf{e}}_1 = R \cos \phi \cos \theta, \quad \mathbf{R} \cdot \hat{\mathbf{e}}_2 = R \sin \phi \cos \theta, \quad \mathbf{R} \cdot \hat{\mathbf{e}}_3 = R \sin \theta. \quad (3.7)$$

More usefully for transformation purposes,

$$\tan \phi = \frac{\mathbf{R} \cdot \hat{\mathbf{e}}_2}{\mathbf{R} \cdot \hat{\mathbf{e}}_1} \quad \text{and} \quad \sin \theta = \frac{\mathbf{R} \cdot \hat{\mathbf{e}}_3}{R}. \quad (3.8)$$

A slightly eccentric variation on a standard textbook transformation⁴ relates the unit vectors of the Cartesian coordinate systems of the model $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$ and the

⁴Not only do we have antenna engineers' polar angles $(-\frac{\pi}{2} \leq \theta \leq +\frac{\pi}{2})$, there is also a peculiarity about the implementation of the experimental coordinates in the telescope mount's software; positive values of T tilt the nose of the superluminal antenna *downwards*. Therefore, a positive T will give a positive θ .

experiment $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$

$$\hat{\mathbf{e}}_1 = \hat{\mathbf{i}} \cos P \cos T + \hat{\mathbf{j}} \sin P \cos T - \hat{\mathbf{k}} \sin T, \quad (3.9)$$

$$\hat{\mathbf{e}}_2 = -\hat{\mathbf{i}} \sin P + \hat{\mathbf{j}} \cos P, \quad (3.10)$$

$$\hat{\mathbf{e}}_3 = \hat{\mathbf{i}} \cos P \sin T + \hat{\mathbf{j}} \sin P \sin T + \hat{\mathbf{k}} \cos T. \quad (3.11)$$

We now turn to the separation of the geometrical center of the superluminal antenna's dielectric from the detection antenna. The relevant experimental parameters are shown in Fig. 3.9; they are D , the horizontal distance of the azimuthal (pan) rotation axis from the detector and the horizontal and vertical offsets of the dielectric from the polar (tilt) rotation axis, which are defined as follows: With the antenna horizontal ($T = 0$), the geometrical center of the dielectric is a vertical distance h above the tilt axis (Fig. 3.9(a)) and a horizontal distance d from it (Fig. 3.9(c)). In experiments, the distance between the center of the superluminal antenna's dielectric and the detector is measured when $P = T = 0$; in terms of the above parameters, this is $D - d$ (Fig. 3.9(c)).

Armed with these various measured distances, for arbitrary P, T , in the fixed (experimental) coordinate system, the geometrical center of the dielectric is at a position \mathbf{r} , where

$$\begin{aligned} \mathbf{r} = & \hat{\mathbf{i}}(h \cos P \sin T + d \cos P \cos T) \\ & + \hat{\mathbf{j}}(h \sin P \sin T + d \sin P \cos T) + \hat{\mathbf{k}}(h \cos T - d \sin T). \end{aligned} \quad (3.12)$$

The vector position of the detection antenna is $D\hat{\mathbf{i}} + h\hat{\mathbf{k}}$, and so the vector separation of the detection antenna from the center of the dielectric is $\mathbf{R} = D\hat{\mathbf{i}} + h\hat{\mathbf{k}} - \mathbf{r}$. Inserting the expression for \mathbf{r} from Eq. (3.13),

$$\begin{aligned} \mathbf{R} = & \hat{\mathbf{i}}(D - h \cos P \sin T - d \cos P \cos T) \\ & - \hat{\mathbf{j}}(h \sin P \sin T + d \sin P \cos T) + \hat{\mathbf{k}}(h - h \cos T + d \sin T). \end{aligned} \quad (3.13)$$

The magnitude of \mathbf{R} is $(\mathbf{R} \cdot \mathbf{R})^{1/2}$, *i.e.*,

$$\begin{aligned} R = & [(D - h \cos P \sin T - d \cos P \cos T)^2 \\ & + (h \sin P \sin T + d \sin P \cos T)^2 \\ & + (h - h \cos T + d \sin T)^2]^{1/2}. \end{aligned} \quad (3.14)$$

The angles θ and ϕ can now be found by substituting Eqs. (3.10) through (3.14) into Eq. (3.8). This yields (after some necessary tedium)

$$\tan \phi = \frac{-D \sin P}{D \cos P \cos T - h \sin T - d} \quad (3.15)$$

and

$$\sin \theta = \frac{D \cos P \sin T - h(1 - \cos T)}{R}. \quad (3.16)$$

Using these transformations, Eq. (3.14) and the known values of d and h , the experimental coordinates $(D - d)$, P and T can be turned into the “model” coordinates R , ϕ and θ . Note that, owing to the definitions of pan and tilt in the software provided with the telescope mount, positive T gives positive θ , whereas negative P gives positive ϕ . Moreover, because of its cylindrical symmetry, TD 1 is used on a turntable (Fig. B.11), not a telescope mount, so that $T = 0$ for all experiments.

3.5.5 Emitted Frequencies and Driving Voltages

Though the original intent was to run TD 1 and TD 2 mostly with emission frequencies 2.2 and 2.6 GHz, almost all of the experiments described in this dissertation use $\omega/2\pi = 2.5$ GHz and $\eta/2\pi = 100$ MHz so that the emitted frequencies are $f_- = (\omega - \eta)/2\pi = 2.4$ GHz and $f_+ = (\omega + \eta)/2\pi = 2.6$ GHz. The reason for this choice is that a small handful of vector modulators (Appendix B) performed comparatively poorly ($\sim 5\text{--}10^\circ$ phase errors) at 2.2 GHz and slow polarization-current speeds. This problem is briefly explored in Appendix C. However, some outdoor experiments

did indeed employ $\omega/2\pi = 2.4$ GHz and $\eta/2\pi = 200$ MHz; in such cases, the emitted frequencies are $f_- = (\omega - \eta)/2\pi = 2.2$ GHz and $f_+ = (\omega + \eta)/2\pi = 2.6$ GHz.

The control programs of TD 1 and TD 2 also allow the voltage $V_0/2$ to be varied. For most of the experiments, the value of $V_0/2$ was kept at 0.45 V to stay well within the linear regions of the vector modulator chips (Appendix B). The effective input impedance of each element is $50\ \Omega$. Therefore (for example), the total power thus delivered to the TD 2 at each frequency is $32 \times 0.45^2 / (2 \times 50) \approx 65$ mW or 18 dBm.

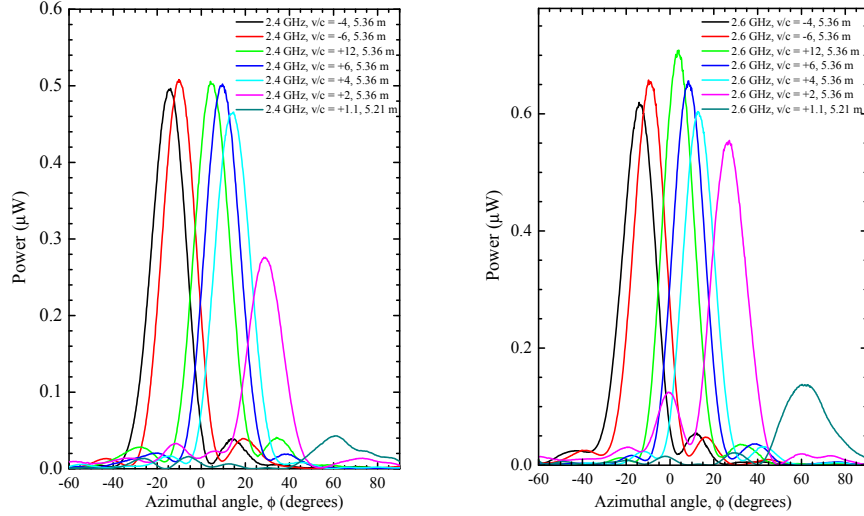


Figure 3.10: Detected power versus azimuthal angle for speeds of $v/c = -4, -6, 12, 6, 4, 2, 1.1$, and frequencies $f_- = 2.4$ GHz (left) and $f_+ = 2.6$ GHz (right), measured at a distance of 5.36 m from the superluminal antenna (5.21 m in the case of $v/c = 1.1$) in the FARM anechoic chamber. Frequencies and speeds are given in the inset keys. In all cases, TD 2 was being run with $\omega/2\pi = 2.5$ GHz and $\eta/2\pi = 100$ MHz.

3.6 Representative Results: Linear Antenna TD 2

3.6.1 Source Speed and Azimuthal Distribution of Radiation

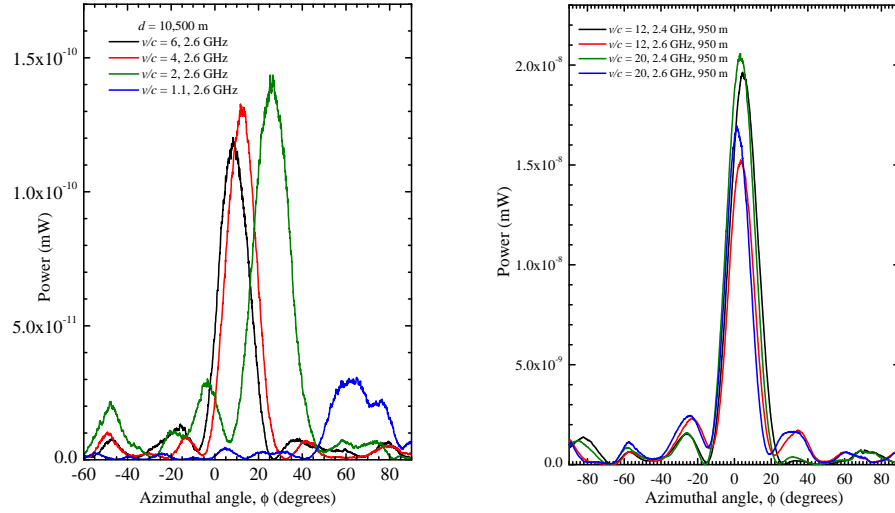


Figure 3.11: Detected power versus azimuthal angle measured outdoors at distances of 10,500 m (left) and 950 m (right); frequencies and speeds are given in the inset keys. In all cases, the superluminal antenna was being run with $\omega/2\pi = 2.5$ GHz and $\eta/2\pi = 100$ MHz.

Fig. 3.10 shows detected power (in μW) versus azimuthal angle ϕ for the linear superluminal antenna TD 2 running at a series of constant speeds ranging from $-6c$ to $+12c$. The data were recorded in the FARM anechoic chamber at a distance of 5.36 m from the antenna (5.21 m in the case of the $v/c = 1.1$ data). When plotted in these linear power units, the azimuthal dependence is clearly dominated by a single, large peak, the angle of which varies with source speed; *i.e.*, the direction of emission can be controlled without mechanically rotating the antenna. Fig. 3.11 displays analogous data recorded during the outdoor experiments, at distances of 10,500 m and 950 m. The plots show that the angular distribution of the emitted radiation, and its steerability, are preserved out to very large distances.

Detailed modeling of the radiation emitted by TD 2 can be found in Chapter 7; however, it is helpful to discuss some features qualitatively. Chapter 2 describes how a polarization current traveling at a constant, superluminal speed v should emit *vacuum Čerenkov radiation*; it is this phenomenon that gives the dependence of the peak power angle on speed seen in Fig. 3.10 and Fig. 3.11. Čerenkov radiation is predicted to propagate out from the antenna in wave fronts that form cones with aperture $\sin^{-1}(c/v)$. The experimental coordinates are such that this should result in emitted power that peaks at an angle

$$\phi = \sin^{-1} \left(\frac{c}{v} \right). \quad (3.17)$$

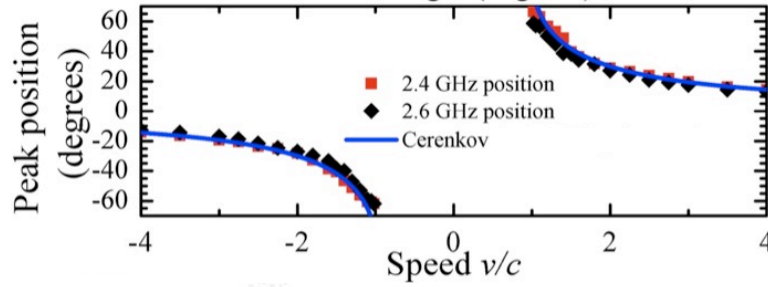


Figure 3.12: Azimuthal angle ϕ at which peak power is detected versus source speed v/c ; data (points) are a compilation of several experiments carried out using TD 2 in the FARM and LANL anechoic chambers. Red (black) points represent peak positions for 2.4 GHz (2.6 GHz). For clarity (*i.e.*, so that details are visible) only data for speeds $-4 \leq v/c \leq 4$ are plotted. The curve labeled “Čerenkov” is Eq. (3.17).

However, Eq. (B.6) shows that the signal applied to the elements of the array may be considered as two separate waves with frequencies $f_+ = (\omega + \eta)/2\pi$ and $f_- = (\omega - \eta)/2\pi$, with respective speeds $v_+/c = \frac{\omega + \eta}{\omega}(v/c)$ and $v_-/c = \frac{\omega - \eta}{\omega}(v/c)$. Hence, the vacuum Čerenkov emission will occur at an angle

$$\phi_- = \sin^{-1} \left(\frac{c}{v} \frac{\omega}{\omega - \eta} \right) \quad (3.18)$$

for the lower (f_-) frequency, and at an angle

$$\phi_+ = \sin^{-1} \left(\frac{c}{v} \frac{\omega}{\omega + \eta} \right) \quad (3.19)$$

for the higher (f_+) frequency. Therefore, the peaks for frequencies 2.4 GHz and 2.6 GHz should lie on either side of the angles predicted by Eq. (3.17); for positive v/c , $\phi_- > \phi_+$, and for negative v/c , $\phi_- < \phi_+$. Fig. 3.12 shows that all of these trends are observed in the data and that the azimuthal angle at which peak power occurs varies as expected with the speed of the source; of course, $\eta/\omega = 1/25$ for these experiments, so that the difference between ϕ_- and ϕ_+ is rather small.

3.6.2 A Qualitative Understanding of the Line Shape

Chapter 7 shows that the emission is entirely vertically polarized for ϕ scans in the horizontal plane ($\theta = 0$). Under these circumstances, at constant speed the antenna's emission will be similar to the diffraction pattern of a wide slit with the transmitted radiation's phase varying linearly with distance across the slit width. This situation is treated in numerous optics text books (*e.g.*, [12]).

We consider the Superluminal Antenna array lying along the y axis, with the 32 elements occupying the range $-\frac{y_0}{2} \leq z \leq \frac{y_0}{2}$. The observation point is in the xy -plane (*i.e.*, $\theta = 0$), at a radial distance R and azimuthal angle ϕ . Taking the formula from [12] and rearranging in terms of the antenna parameters, the amplitude \mathcal{A} is given by

$$\mathcal{A} \propto \int_{-\frac{y_0}{2}}^{\frac{y_0}{2}} \frac{e^{i(\omega_{\pm}t - ky)} e^{iK[(R^2 + 2Ry \sin \phi + y^2)^{1/2} - R]}}{(R^2 + 2Ry \sin \phi + y^2)^{1/2}} dy. \quad (3.20)$$

Here, $\omega_{\pm} = 2\pi f_{\pm}$, $k = \omega_{\pm}/v_{\pm}$ (see Eq. (3.18) and Eq. (3.19)) and $K = \omega_{\pm}/c$; the subscript \pm indicates that f_+ and v_+ should be chosen to model the upper-frequency emission and f_- and v_- that at the lower frequency. Eq. (3.20) may be evaluated numerically; however, for distances $R \gg y_0$, variations in the denominator over the integration range may be neglected. Under such conditions, Eq. (3.20) can be calculated analytically, giving

$$\text{power}(R, \phi) \propto \frac{\sin^2[(K \sin \phi - k)\frac{y_0}{2}]}{(K \sin \phi - k)^2}, \quad (3.21)$$

which will be recognized as a skewed Airy function. The maximum power will occur when $(K \sin \phi - k) = 0$, which yields

$$\sin \phi_{\pm} = \frac{c}{v_{\pm}}, \quad (3.22)$$

in agreement with Eqs. (3.18) and (3.19). In practice, numerical evaluations of Eq. (3.20) are indistinguishable from Eq. (3.21) for $R > 3$ m at the operating frequencies of TD 2. Fig. 3.13 also shows that Eq. (3.21) gives a reasonably satisfactory fit to experimental data, at least for the central maximum.

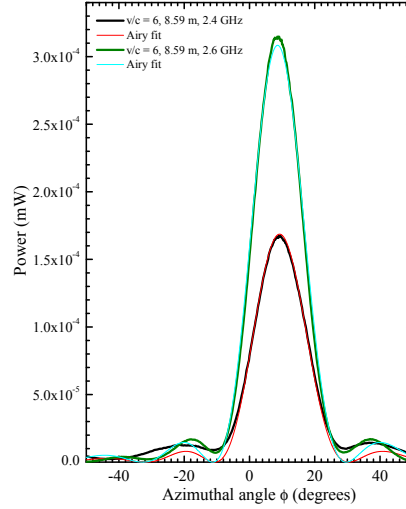


Figure 3.13: The results of fitting Eq. (3.21) to experimental data recorded in the FARM anechoic chamber ($R = 8.59$ m, $v/c = 6$).

3.6.3 Power Versus Speed

Both Fig. 3.10 and Fig. 3.11 show that the efficiency of the emission falls as the source speed is reduced towards the speed of light; typically, the peak power when the source was run at $v/c = 1.1$ was 5 – 8 dB lower than that seen with higher source speeds. This phenomenon will be discussed in more depth in Appendix C,

but it is worth mentioning here that it is due to a process analogous to *total internal reflection*.

Total internal reflection [12] occurs when light is propagating in a dielectric towards an interface with a medium (such as air or vacuum) with a lower dielectric constant. As the angle of incidence rises, the light will eventually undergo total internal reflection. For the interface between a non-magnetic material with relative permittivity ϵ_r and air or vacuum, the critical angle at which this effect starts is $\sin^{-1}(\epsilon_r^{-\frac{1}{2}})$. For alumina ($\epsilon_r \approx 10$), this corresponds to $\phi \approx 18^\circ$.

Whilst this picture is usually applied to situations where the wavelength of light is much less than the size of the dielectric components involved, even at longer wavelengths the effect inhibits the emission of radiation from the superluminal antennas at larger angles; it is, in effect, an angle-dependent impedance mismatch. Empirically, the emitted power seems to decrease like $\cos^2 \phi$, suggesting that the emission comes from a finite thickness “skin” at the surface of the dielectric (interface charge), a signature of poor impedance matching [54].

In Section 3.5.5, the larger phase and amplitude errors obtained in TD 1 and TD 2 when setting slow speeds (corresponding to large emission angles) were mentioned. In part this is due to the confined oscillatory fields not being able to escape; they are therefore reflected back into the drive electronics, making an unambiguous measurement of the phase and amplitude much harder.

3.6.4 Polar Power Dependence

Fig. 3.14 shows the results of the polar angle studies carried out in the LANL anechoic chamber. In these measurements, the tilt angle T (Fig. 3.5) was set at fixed values, separated by 10° , at each value of T , the pan angle P was again scanned from $P = -90^\circ$ to $P = +90^\circ$ in 0.18° steps. As before (*c.f.*, Fig. 3.10), the emission peaks

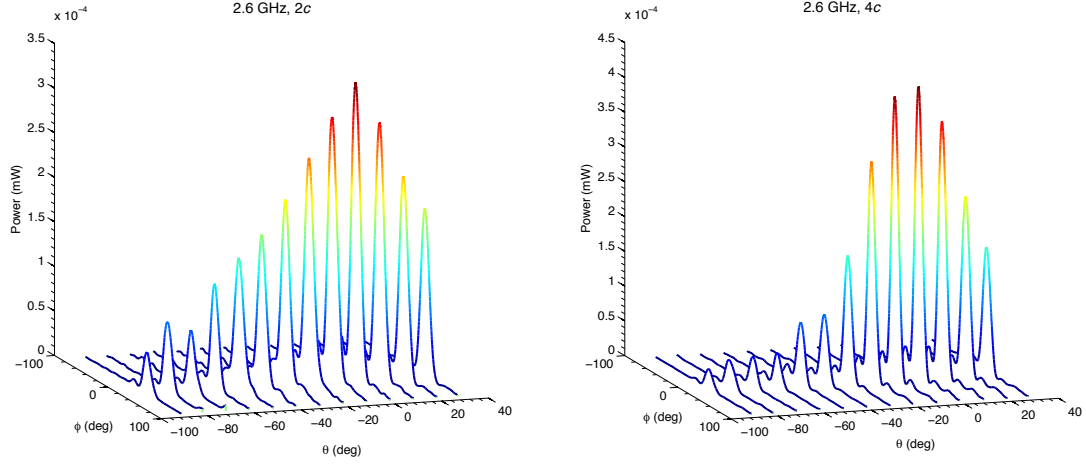


Figure 3.14: Detected power versus azimuthal and polar angles, measured at a distance of 3 m in the LANL anechoic chamber; left $v/c = 2$, 2.6 GHz, right $v/c = 4$, 2.6 GHz. The peak power gradually decreases with increasing polar angle.

at an azimuthal angle close to that expected for Vacuum Čerenkov radiation.

To parameterize these quite sparse data, the peak powers were reflected about zero degrees and fitted to a Lorentzian:

$$\frac{P}{P_{\max}} = \frac{P_0}{P_{\max}} + \left(\frac{2\mathcal{A}}{\pi} \right) \frac{\Gamma}{(4[\theta - \theta_0]^2 + \Gamma^2)}. \quad (3.23)$$

Here, the amplitude \mathcal{A} , the background (angle-independent) power P_0 , the center angle θ_0 (≈ 0) and the width Γ are fit parameters.⁵ Typical fits are shown in Fig. 3.15, where the peak powers have been normalized to the maximum value P_{\max} ; the fitted background was usually less than 10 % of the peak power. The Lorentzian widths Γ are plotted as a function of source speed v/c in the right-hand panel of Fig. 3.15; the polar width is slightly larger around $v/c \approx 2$, later traced to a poor phase setting at this speed.

⁵The fit parameter Γ gives the full width at half-maximum power. This is very similar to the full width at the -3 dB points, since $\log_{10}(2) = 0.3010$ to four significant figures.

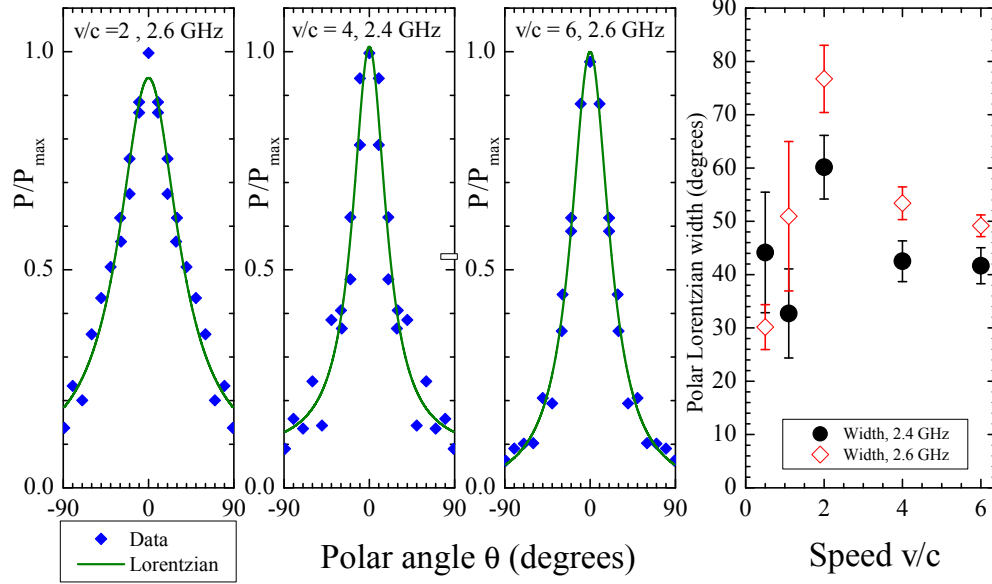


Figure 3.15: The three left-hand panels show fits of a Lorentzian (Eq. (3.23)) to the peak power as a function of polar angle θ ; for clarity, the peak power has been normalized to the maximum value. Speeds and frequencies are labelled in each plot. The right-hand figure shows the fitted Lorentzian linewidths Γ as a function of speed v/c ; subluminal ($v/c = 0.5$) data are also included.

3.7 Range tests of TD 1

Chapter 7 provides theoretical fits of the emission from TD 1, and experimental lineshapes as a function of m (m parameterizes the speed) are presented in some detail there. Here we instead display a few examples of data from the extensive range tests carried out on this circular antenna. The reasons for these range tests were twofold.

(i) Chapter 5 describes the controversy surrounding the prediction of a component of the radiation emitted by a circular superluminal source the intensity of which would decay more slowly with distance than the inverse-square law out to very large

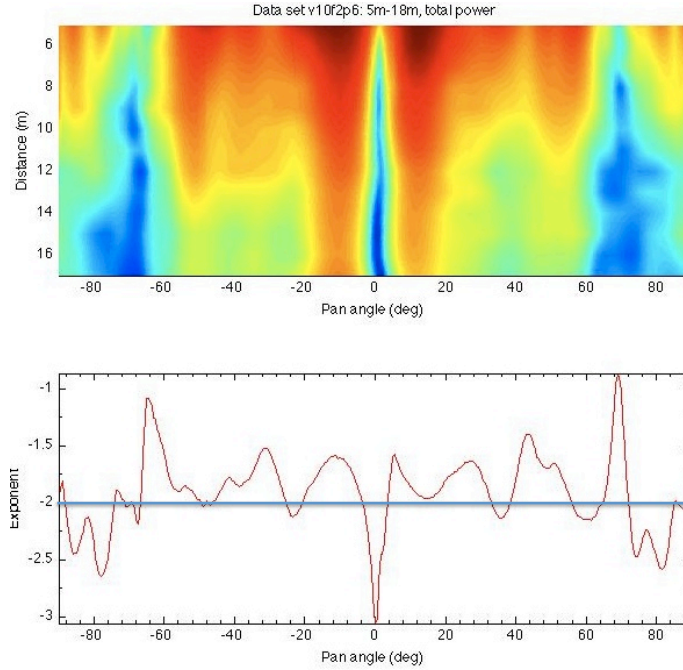


Figure 3.16: Upper: contour plot of total (horizontally plus vertically polarized) power (dBm) versus polar angle θ and transmitter to receiver distance r . Lower: corresponding decay exponent n obtained by fitting Power $P = P_0 r^n$ to the upper data along constant-angle paths. The experiment was carried out using TD 1 in the Kirtland Air Force Base anechoic chamber; the speed setting was $m = 2$, corresponding to a phase difference of 10° between adjacent antenna elements and a speed of $v/c \approx 3.3$; the detected frequency was $f_+ = 2.6$ GHz.

distances. The range tests searched for possible traces of this component but failed to find them because it was later shown to be non-existent (see Chapter 5)!

(ii) Despite the wide angular distribution of the radiation emitted by TD 1 (see Chapter 7), leading to significant ground reflections of the broadcast radiation and therefore potential destructive interference, the fall-off of the detected power with distance was found to be comparable to “state of the art” telecommunications antennas, promoting some interest in possible commercial applications.

The most comprehensive tests were carried out in the large anechoic chamber at Kirtland Air Force Base, allowing path lengths of up to 18 m in an essentially reflection-free environment. Polar scans of the received power (both horizontally and

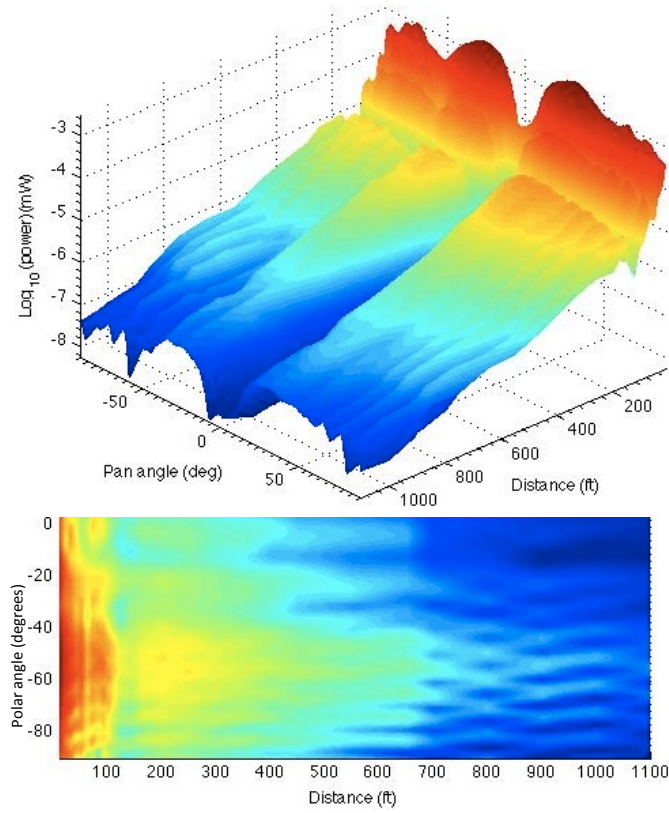


Figure 3.17: Two different views of a contour plot of total (horizontally plus vertically polarized) power (mW - logarithmic scale) versus polar angle θ and transmitter to receiver distance r . The experiment was carried out using TD 1 at the FARM Outdoor range ; the speed setting was $m = 3$, corresponding to a phase difference of 15° between adjacent antenna elements and a speed of $v/c \approx 2.2$; the detected frequency was $f_+ = 2.6$ GHz.

vertically polarized) were collected at distance intervals of 1 m for several values of m . An example of the data is given in Fig. 3.16, showing the format derived by the present author used to compare and evaluate all of the range tests. The upper part shows a contour plot of detected power (in dBm) versus polar angle θ and transmitter to receiver distance r . The lower part displays the corresponding decay exponent n obtained by fitting power $P = P_0 r^n$ to the upper data along constant-angle paths. In similar measurements, conventional antennas (both horns and dipoles were used) showed a uniform decay exponent of $n = -2$ (the expected inverse-square law [54]); by contrast, TD 1 gives an exponent that varies with polar angle to either side of

$n = -2$. As Chapter 7, will show, the reasons for this are now understood; they are not due to the “exotic physics” debunked in Chapter 5.

An example of analogous data collected at the FARM Outdoor Range is shown in Fig. 3.17. An interesting feature is the reproducible speckle pattern seen at larger distances; Fig. 3.18 shows data for just two distances to illustrate the effect better. This phenomenon is analogous to laser speckle [12] and is caused by scattering from various metal objects placed in the desert around the range for airborne RADAR experiments. A later experiment with better focusing control (Section 6.6.2) and analysis discussed in Section 7.8.3 combine to show that at certain pan angles, portions of the source approach these objects at the speed of light and with zero acceleration; therefore all of the temporal variations within this volume of the source appear coherent at the scatterer. The scattered radiation therefore produces very strong interference, resulting in the sharp, reproducible features.

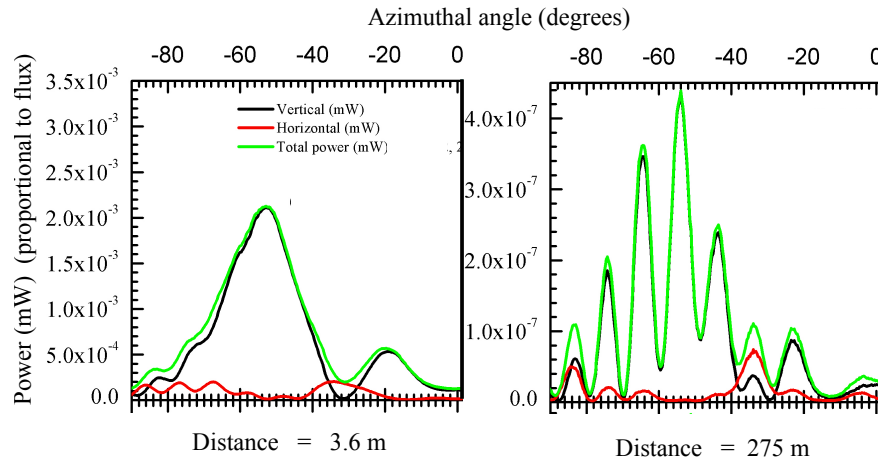


Figure 3.18: Cuts through the data of Fig. 3.17 at two different source-to-detector distances. At smaller distances (left), the radiation pattern is unremarkable; at large distances (right), scattering from objects either side of the range produces reproducible sharp features, the angular position of which depend on the speed settings of TD 1.

Chapter 4

Sources That Move Faster Than Light and How to Treat Them

Empirical evidence can never establish mathematical existence – nor can the mathematician’s demand for existence be dismissed by the physicist as useless rigor. Only a mathematical existence proof can ensure that the mathematical description of a physical phenomenon is meaningful.

— Richard Courant [123]

4.1 Third Prologue

4.1.1 The Consequences of a Vexing Problem

The antennas described in Chapters 2 and 3 represent the first practical attempts to obtain controllable emission of electromagnetic radiation from an accelerated superluminal volume source contained within a dielectric material. Compared to established devices such as phased arrays, there are several distinctive features, not least that the

origin of the radiation (a polarization current, rather than one of free electrons) is not confined to discrete elements such as dipoles. Instead, it is distributed throughout the dielectric and may move faster than the characteristic wave speed. It is legitimate – in fact, necessary – to establish whether the techniques used to simulate the radiation from conventional antennas apply likewise to these unprecedented devices or if it is requisite to resort to first principles and calculate the problem *ab initio*.

Faced with the challenge of building antennas to a three-year time-limit without a proven method to model the emitted radiation, the Los Alamos collaboration turned for help to an enthusiastic proponent of superluminal light sources, the Cambridge University mathematical physicist, astronomer and correspondent of Vitalii Ginzburg, Dr Houshang Ardavan. Aided in practical details and subsidiary calculations by members of the Los Alamos team, Ardavan quickly started to produce theoretical predictions considered relevant to the antennas in production. However, no sooner were the results published in the literature than comments started to appear, followed, at short notice, by responses from the original authors and yet further statements from their challengers. Subsequent investigation revealed that most of Ardavan’s previous publications had similarly engendered a cascade of arguments, counterarguments, and arguments to counter the counterarguments. From the present author’s standpoint, the former physicist Angela Merkel’s remark, “Das hat mir einen grossen Shitstorm eingebracht!” is very apposite.

In all of this controversy, who was correct? In their comments, Ardavan’s detractors tended to assume that readers would have a life-long education in electromagnetic theory, along with a detailed knowledge of the results derived and validated by luminaries such as Maxwell, Kirchhoff, Heaviside and Hadamard. By contrast, many texts in electrical engineering take for granted the validity of and premises behind these techniques, offering little or no explanation. Hence, to unravel the conflict between Ardavan and his critics, one must indeed return to first principles and revisit the mathematical arguments behind wave propagation.



Figure 4.1: *The Great Wave Off Kanagawa*, a woodblock print by the Japanese ukiyo-e artist Hokusai, was published sometime between 1829 and 1833 in the late Edo period. The rogue wave threatening three boats off the coast of the town of Kanagawa can be modeled by Burgers' equation $\partial\psi/\partial t + \psi\partial\psi/\partial x = \epsilon\partial^2\psi/\partial x^2$ which is highly singular once the wave breaks and a 'shock' forms. Credit: Wikipedia ©public domain

It turns out that the issue at stake can be stated quite simply: *The Cauchy problem for the three-dimensional wave equation is improperly posed (in the sense of Hadamard) if the velocity of the source exceeds the characteristic wave speed.* To make this statement fully transparent, the present chapter must deal with two questions in some detail. First, what is a Cauchy problem, and how may it be ill- or well-posed in the sense of Hadamard? Second, what are the ramifications of an improperly set system and, in particular, which mathematical methods will lead to a physically correct answer when we come to calculate the radiated fields of superluminal sources in later chapters?

Finding answers to these propositions demands considerable perseverance and an eye for detail. We will find that a well-posed problem depends on *existence* and *uniqueness* as well as *stability*, where the latter denotes continuous dependence on certain conditions prescribed on a hypersurface in the problem domain. Once a source moves faster than the waves it generates, stability is lost along with any hope for a smooth, “classical” problem; the solution – if it exists at all – will be *weak* or *generalized* from a mathematical physics standpoint. This cross-over from the sub-

to the superluminal regime corresponds to the transition from a *space-like* to a *time-like* problem in mathematics, imposing severe restrictions on the direction in which a system may be solved if causality is to be preserved.

It will soon become clear that the imposition of proper initial and boundary conditions plays an essential part in the search for an acceptable solution to a differential equation. The calculated emission of the antennas throughout this work is assumed to be into *free space* which demands an advanced understanding of infinity – mathematically as well as physically. If the concept of the boundless, propelled to prominence by Georg Cantor, is applied correctly, it can be shown that one or more terms in the Green’s function solution to the three-dimensional wave equation must be *identically* zero. (Kirchhoff’s incomplete grasp of infinity, although very much in accordance with nineteenth century practices, led to his “integral theorem”, an over-specified Green’s function approach [53] which is now – more than 100 years ‘post-Hadamard’ – known to be inapplicable to free-space radiation calculations¹.)

Having examined all of the prerequisites for a valid solution describing the emission from a source that travels faster than the characteristic wave speed, it is possible to undertake a critique of Ardavan’s work. We will show that most of the suspicions expressed by his detractors are, in fact, well founded. (Nevertheless, many of his predictions about astronomical objects such as pulsars, though based on shaky reasoning, are likely to be correct.) The detailed examination of electromagnetic theory also underpins and validates the calculations made in Chapters 6 through 8.

The rather ambitious agenda of this chapter sets out with some “points of principle” that define and explain the basic methods of mathematical physics needed to

¹However, Jackson [53] states that “Kirchhoff’s theory, despite its mathematical inconsistency and its physical deficiencies, works remarkably well in the *optical domain* and has been the basis of most of the work on diffraction” (emphasis added). Ardavan (see, for instance, [117]) likes to quote this section of the seminal text, ignoring both Jackson’s caveat on the subsequent page and the fact that the present problem has very little to do with diffraction since the latter involves radiation that passes through an opening in an opaque screen.

appreciate the arguments to follow. This includes the distinction between well- and ill-posed descriptions of physical systems, the necessity – and beauty! – of properly prescribed boundary conditions, over- and underdetermination as well as the existence, uniqueness and stability of solutions pertaining to the problem at hand. The contributions of Cauchy, Dirichlet, Kovalevskaya and Hadamard to these concepts are introduced since they offer glimpses into the *zeitgeist* of late nineteenth- and early twentieth-century mathematics. As some of the calculations in Chapter 7 and all of those in Chapter 8 are numerical, the possibility of chaotic behavior and its impact on computational methods is discussed. Finally, we will touch upon the idea of classical versus generalized (or weak) solutions, supporting to the use of Green’s function methods in subsequent parts of this dissertation.

Section 4.2 talks about the Cauchy problem, first introducing the classification of second-order partial differential equations (such as the wave equation) into the types proposed by Hadamard. This is decisive in establishing the type of initial and boundary conditions that determine a unique, stable solution to a problem of mathematical physics. The wave equation is shown to be *hyperbolic*, *i.e.*, it is a member of a class of equations that describe disturbances that propagate away from their source. This leads to the idea of *characteristics*, traveling surfaces across which solutions to partial differential equations can have discontinuous derivatives; Mach and Čerenkov cones – already encountered in Chapter 2 – are practical examples. The notion of disturbances that advance through spacetime permits definitions of the *range of influence* and *domain of dependence* along with a discussion of Huygens’ principle which predicts the (rather spectacular) dependence of wave propagation on the dimensionality of the problem. The section culminates in a discussion of the most startling property of supersonic or superluminal emission: The discontinuous change from a properly- to an ill-posed problem, and a space-like to a time-like solution, that occurs as the source’s speed overtakes that of the emitted waves.

The theoretical explorations of this chapter find their conclusion in Section 4.3

and the recognition that a Green’s function approach is an excellent method to study the electromagnetic fields arising from the type of antenna studied here. Moreover, having dissected the mathematics underlying faster-than-light sources step by step and *ab initio*, we set about Chapter 5 fully understanding the blunders and fallacies committed by others working in the field and disseminated recklessly in the published literature.

4.1.2 Some Points of Principle

In theory, the many explicit forms of the scalar wave equation which, in general form, can be stated as²

$$\square \psi(\mathbf{r}, t) \equiv \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \psi(\mathbf{r}, t) = 4\pi q(\mathbf{r}, t) \quad \text{in } \mathbb{R}^3 \times (0, \infty) \quad (4.1)$$

have countless different solutions – an embarrassing wealth of choices. Infinities and zeros may be placed with impunity and results can be found for any reasonably continuous distribution of arguments on any arbitrarily shaped surface. To quote [124], “a marvellous book by one of the very few mathematical giants who can talk intelligibly to physicists prepared to listen” [125],

A differential equation – whether ordinary or partial – admits, in principle ... an infinity of solutions. In elementary cases, where it is possible to integrate the equation completely, the solution obtained contains constants or arbitrary functions. By varying these arbitrary elements, one can find all possible solutions, save certain exceptional ones: One says, then, that one has found the “general integral”.³

²As before, the box (\square) represents the D’Alembert or wave operator, the Laplacian of spacetime. q is a time-varying current or charge density, giving not only the distribution of the radiation sources, but also their time dependence at each point in space. c denotes the wave speed, the rate at which the electromagnetic disturbances propagate, and the scaling constant 4π has been divided out for ease of calculation.

³The translation of this (and any other) quote from [124] is due to the present author,

However, as Hadamard points out in his lectures [126, 127], finding the “general integral” – a pursuit that he dismisses as “the older and classic view” – is of limited use at best, not only because it is, save in the simplest of cases, difficult (or impossible) to obtain, but

above all, because the question does not by any means consist merely in its determination. The question, as set by most applications, does not consist in finding *any* solution of the differential equation, but in choosing, amongst all those possible solutions, *a particular one defined by properly given accessory conditions* [127] [emphasis added].

The point at issue is, therefore, whether one can find – from an infinity of choices! – the *particular* solution corresponding to the *particular* problem at hand [62, 128–135]. The latter usually differs from its fellow candidates in the *boundary* or *hyperboundary conditions* imposed on the differential equation (we will distinguish between the two later): complementary expressions that specify the value of the unknown function ψ and, possibly, some of its derivatives on the surface that encloses the volume of interest in space or spacetime.

It is difficult to overstate the significance of posing proper constraints when solving problems of mathematical physics, that is those that model natural or man-made physical systems. In [124], Hadamard proposes a trio of questions – the last of which he duly credits to Hilbert and Courant [128] – to identify boundary value problems or integral equations that are said to be *well-posed* or *properly posed*.

To know whether we may consider a problem ... well-posed, we must study it in regard to the following questions:

1. Is the problem conceivable?
2. Is it determined?

who feels very strongly that *La théorie des équations aux dérivées partielles*, more than any other work by Hadamard, is deserving of an English edition.

But, in addition, we demand that the solution be continuous in relation to the data, which is to say that very small variations in the conditions may only change the values of the unknown function very little. ... Hence, we add the following question to those posed above:

3. Is the solution of the problem functionally continuous, at least to suitable order k , with respect to the data?

In brief, we would prefer (i) that a solution to the problem does, indeed, exist, (ii) that this solution is unique, and (iii) that it is stable⁴. In practice, this statement is made precise by additionally stipulating the function spaces in which the data and solution are to reside in conjunction with the measure and notion of continuous dependence [136]. Lavrentiev [137] proposes the following rigorous definition, which we adopt here in slightly revised form:

Let X, Y be complete metric spaces and $A\psi$ a function with domain of definition X and range of values⁵ Y . Now consider

$$A\psi = q, \tag{4.2}$$

to which most models of mathematical physics can be reduced if the aim is to investigate solutions ψ when function A acts on X and the right-hand side of the equation – commonly some source or sink – is given. The problem of solving Eq. (4.2) is properly posed if there exists a unique function Bq , defined and continuous over all of Y , that is inverse to $A\psi$.

⁴When we discuss the concept of *continuous dependence on the data*, we consider not only values of the differential operator along with those at the system's boundary, but coefficients, parameters, and the general geometry of the problem as well. Hence, in a certain sense, we might think of many – if not most – models of physical systems (*e.g.*, forced vibration, bifurcation and entire classes of singular perturbation problems) as possibly being improperly posed. As Payne remarks somewhat facetiously in [136], “Perhaps even the concept of improper posedness might be said to be ill-posed.”

⁵For linear problems, X and Y are Banach (*i.e.*, complete normed vector) spaces and A denotes a linear operator. A comprehensible discussion of linear spaces in the context of electromagnetic theory can be found in [60], for a more rigorous treatment we refer the interested reader to introductory texts in functional analysis such as [138] or [139].

The first postulate, *existence*⁶ of the function Bq , implies that the system is not to be *overdetermined* or, in other words, that no incompatible or superfluous conditions should be imposed. Although physicists and engineers rarely encumber themselves with mathematical existence (to the great chagrin of Richard Courant, as evidenced by the quote that opens this chapter), the effort of proving that a mathematical model does, in fact, have a solution in the chosen function space may hardly be dismissed as “useless rigor”. Empirical evidence, no matter how compelling, can, on its own, not ensure “that the mathematical description of a physical phenomenon is meaningful” [123].

From *uniqueness* follows that the conditions must not *underdetermine* (or, grammatically more pleasingly, “insufficiently specify”) the system such that Bq , irrespective of the mathematical means by which it was obtained, is certain to be the *sole* solution. For many physically relevant problems, however, uniqueness is not to be expected. The inverse diffusion (or heat) equation may serve as an example since it is, in general, impossible to deduce the initial from the final state of the system uniquely. In fact, inverse problems in mathematical physics, that is to say models that attempt to calculate from a partial or full knowledge of the solution the causal factors that produced it, are notoriously difficult to treat since they frequently (but by no means always) fail the requirements of uniqueness, continuous dependence on the boundary data, or both. In practice, solutions to inverse systems are often numerically “negotiated” (for instance by regularization schemes or parameter estimation) rather than analytically derived.

The last stipulation, *stability* or *continuity* of the inverse function Bq , is especially important for physical applications since we favor – as does nature – systems

⁶Unless noted otherwise, we shall use the term “uniqueness” in the local, rather than global, sense. This is in agreement with some [132, 140], but at odds with other authors; incidentally, for a few of the problems treated in [136], local existence would follow from the Cauchy-Kovalevskaya theorem (and its extensions), rendering them properly rather than improperly posed as the title of the book suggests.

where arbitrarily small variations in the conditions do not, for any physically relevant arguments, lead to large (or infinite) changes in the solution. As pointed out by Lavrentiev [137], the reason for ostracizing unstable boundary value problems for the better part of a century seems, at least in hindsight, obvious: Assuming that Eq. (4.2) represents an actual physical process (man-made or otherwise), the right-hand side of the equation is a mathematical approximation to instrument assessments (*i.e.*, measurements), derived by intuition or educated guesswork. If, for any given \mathbf{r} and t , the function $Bq(\mathbf{r}, t)$ turns out to be discontinuous, the solution ψ – should it exist at all – can not be uniquely recovered from q at that point, implying that imperceptible measurement errors in the neighborhood of (\mathbf{r}, t) could change the experiment’s outcome appreciably. At a time that predates pop-culture gems such as *The Butterfly Effect* (starring a very young Ashton Kutcher), these models were found to be irrelevant or “peculiar”, rendering them, along with any others whose solution exhibit a similar dependence on the equation’s right-hand side, unsuitable as “translation of some mechanical or physical question” [127]. We remark parenthetically that stability is essential in numerical calculations since unavoidable round-off and truncation errors preclude the exact representation of initial and boundary data in a computer’s memory. If the problem itself is unstable, small errors in the conditions may be propagated by the numerical scheme in such a way that the calculation becomes meaningless.

Although the work of Cauchy, Kovalevskaya, Darboux and Goursat bears witness to the contributions from nineteenth-century mathematicians to the study of well-posed problems in differential equations, the subject gained prominence only after the publication of Hadamard’s rigorously formulated definition in 1902. For many decades, its hold on mathematical physics was such that models failing to meet any one of his three postulates were, somewhat slanderously, labeled “incorrectly set” and, in essence, ignored by the mathematical community. Discussing the problem of determining a solution ψ to the differential equation $\partial^2\psi/\partial x\partial y = 0$ from given val-



Figure 4.2: The Russian mathematician Sofia Kovalevskaya (1850-1891) was the first woman in history to obtain a doctorate (in the modern sense) in mathematics. Appointed Ordinary Professor at Stockholm University and Corresponding Member of the Russian Academy of Sciences, her contributions to analysis, differential equations, and mechanics make her “the greatest known woman scientist before the twentieth century” in the words of historian of science Ann Hibner Koblitz [141]. Credit: <https://www.mpg.de/female-pioneers-of-science/sofia-kovalevskaya> ©public domain

ues on a *closed* curve (which was, even in 1941, known to be improperly posed), Fritz John [142], who went on to author one of the most widely read texts in partial differential equations, writes almost apologetically, “In the case of a hyperbolic equation the Dirichlet problem⁷ certainly is not a “natural” problem of mathematical physics, as its solution may neither exist, nor be uniquely determined, nor depend continuously on the data”. He hurriedly defends his frivolous pursuit by adding that “it is [nevertheless] possible to obtain fairly general positive results in this connection.”

The most celebrated example of its kind, the Cauchy problem for $\nabla^2\psi = 0$, dates back to Hadamard himself [127] and remains included in many contemporary textbooks [62, 130–132, 135]. In two dimensions, the Laplace equation can be given

⁷Cauchy, Dirichlet, Neumann and Robin problems (to list the most widely known) ask for solutions of a differential equation (or an entire class of differential equations) that satisfy boundary or initial conditions of the same name. Detailed descriptions of the latter will follow in later sections of this chapter.

as

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\psi(x, y) = 0 \quad \text{in } \mathbb{R}^2 \times (0, \infty); \quad (4.3)$$

its solutions are harmonic functions whose collective study is known as potential theory, widely exploited in fluid and thermodynamics, electromagnetism, and astronomy. In general, applying Cauchy initial conditions to Eq. (4.3) results in an overdetermined system similar to “an algebraic expression implying more *conditions* than unknowns” [127]. Hence, unless a strong compatibility relation holds among the Cauchy data, existence will not be global. (*Local* existence, however, follows from the Cauchy-Kovalevskaya theorem, which will be discussed briefly in Section 4.2.2.) Moreover, even in the unlikely event that said condition is satisfied and the choice of arguments seems to favor a classical solution⁸, the latter will not depend continuously on the data. This can be shown by returning to Hadamard’s pièce de résistance and imposing the initial conditions

$$\psi(x, 0) = 0 \quad \text{and} \quad \frac{\partial}{\partial y} \psi(x, 0) = \frac{1}{k} \sin kx \quad \text{where} \quad k > 0, \quad (4.4)$$

which arise prominently in hypersonic blunt-body problems [143]. Separation of variables suggests the easily verifiable explicit solution

$$\psi(x, y) = \frac{1}{k^2} \sinh ky \sin kx. \quad (4.5)$$

When k approaches infinity, the Cauchy data (4.4) tend uniformly to zero. The solution $k^{-2} \sinh ky \sin kx$, on the other hand, grows rapidly for any determinate value $y \neq 0$. Physically, Eq. (4.5) represents an oscillation (or “fluting”⁹) of amplitude

⁸To be considered “classical”, we require that the solution to a partial differential equation of order m be at least m times continuously differentiable. Then, if nothing else, all derivatives which appear in the problem statement will exist and be continuous. This is in agreement with [62], who asserts that “a solution with this much smoothness ... is certainly the most obvious notion of solution [of a partial differential equation]”.

⁹Whether coined by Hadamard himself or his translators, the term “fluting”, which denotes decorative grooves intended to make a column appear more perfectly circular in architecture, is used – quite uniquely – as a synonym for “oscillation” in [127].

$k^{-2} \sinh ky$ which grows without bound as y increases and x is held constant; thus, even if k is fixed, the system is unstable in that, for any constant x , $\psi \rightarrow \infty$ as $y \rightarrow \infty$ if $\sin kx \neq 0$.

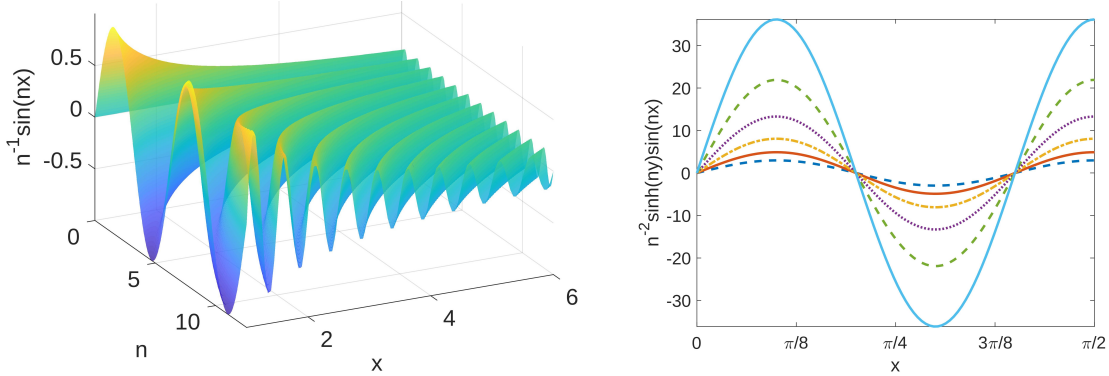


Figure 4.3: The Cauchy problem for the Laplace equation. (left) The initial condition $k^{-1} \sin kx$ tends uniformly to zero when $k \rightarrow \infty$. (right) Even if k is held constant, the solution $k^{-2} \sinh ky \sin kx$ oscillates between limits that increase indefinitely as $y \rightarrow \infty$.

It is obvious that the Cauchy problem for the Laplace equation is improperly posed since small changes in the conditions might give rise to an unstable model or one for which no solution exists at all. Perhaps more intriguingly, the celebrated example shows that any attempt to “smooth” the data by means of polynomial approximations or perturbation methods may lead to solutions that vary wildly from that of the original problem at points off the Cauchy surface. To cite Hadamard once more [127],

Some [geometers] indeed argued that you may always consider any functions as analytic, as, in the contrary case, they could be approximated with any required precision by analytic ones. But, in my opinion, this objection would not apply, the question not being whether such an approximation would alter the data very little, but whether it would alter the solution very little. It is easy to see that ... the two are not at all equivalent.

Hadamard’s dogma to excise ill-posed problems from mathematical physics persisted until well after its architect’s death in 1969. Yet, as Poston and Stewart remark

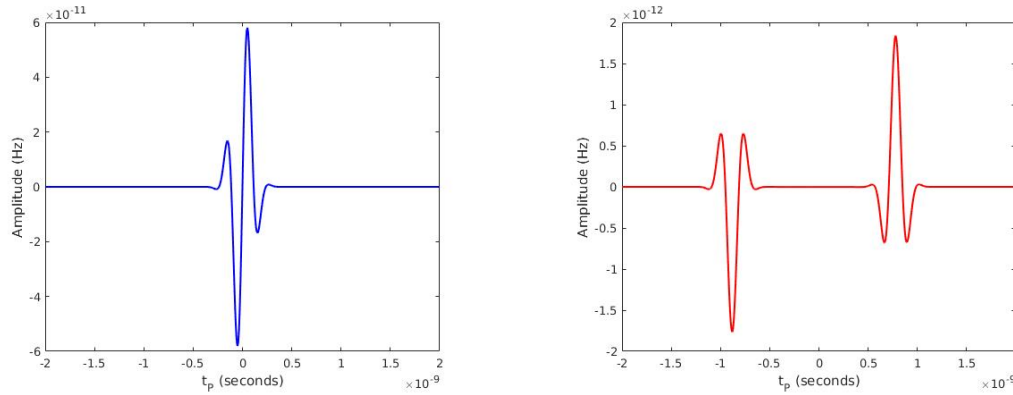


Figure 4.4: Images of a smooth wave-packet-like signal received from a faster-than-light source at two different spacetime positions. (left) A single, coherent image of the source is observed. (right) The image of the source is not only spread over two disjoint volumes, but also distorted (or “scrambled”).

so poetically [29], many things jump. “Water suddenly boils. Ice melts. Earths and moons quake. Buildings fall. The back of a camel is stable, we are told, under a load of N straws, but breaks suddenly under a load of $N + 1$. Stock markets collapse.” Most importantly, attitudes change. As the twentieth century drew to a close, it became increasingly apparent that entire classes of physically relevant problems are, in some sense, unstable and do not lend themselves to classical treatments, yet can be studied with gain. (The monograph by Payne [136] includes a substantial bibliography of work done prior to 1975 while the introductory chapter of [144] attempts to survey some of the more recent literature.) In a twist of supreme irony, systems that are sensitive to undetectably small changes in the conditions and, therefore, appear “chaotic” fueled the imagination of general audiences in unprecedented ways. Works of popular science such as *Chaos* [145] and *Complexity* [146] ascended to the top of the New York Times Bestseller List although (or, perhaps, precisely because) their subject was – at least in the late 80’s – “still so new and so wide-ranging that nobody [knew] quite how to define it, or even where its boundaries [lay]” [146] – a statement that induces nightmares in any self-respecting mathematician.

Eschewing the wildly sensationalist claims that rang in the new millennium and caused, quite unintentionally, a minor conflagration among chaos scientists who had mixed reactions to their new high profile [147], we simply remark with [130] that dynamical systems which are highly sensitive to initial conditions “need careful interpretation, and their mathematical solution may be subtle”. Looking ahead to the remainder of this chapter we state, so far without proof, that the Cauchy problem for the wave equation in three dimensions is properly posed and has an explicit solution *in the classical sense* (given by Kirchhoff’s formula) if the functions that appear either in the equation itself or in the Cauchy data are sufficiently smooth¹⁰ and ψ is – in accordance with Hadamard’s third postulate – impervious to small perturbations of the initial conditions. While this is, in general, true for radiation sources whose speed remains below that of light, it is to be expected that, in the superluminal regime, Eq. (4.1) describes the formation and propagation of caustics or “shocks”, which is to say curves of discontinuity. In consequence, an observer might, depending on his or her position in spacetime, receive multiple “images” of q even if the latter is perfectly smooth in (\mathbf{r}, t) (Fig. 4.4). Of particular concern is the region where the source “breaks the light barrier” and its speed equals precisely that of the radiated waves, as evidenced by the history of supersonic flight in Chapter 2.

The relation between emission and observation time for a charge (or charges) in superluminal motion will be taken into careful consideration throughout the remainder of this text. For the time being, we simply state that the structure of the Cauchy

¹⁰As before, “sufficiently smooth” implies that all derivatives exist and are continuous [62]. Nonetheless, Hans Lewy [148] shows – admittedly with “considerable surprise” – that smoothness does, in general, not guarantee existence or uniqueness. The notable example, named after its author, states that, on $\mathbb{R} \times \mathbb{C}$, there exists a smooth complex-valued function $q(x, t)$ such that the differential equation

$$\frac{\partial}{\partial x} \psi - ix \frac{\partial}{\partial t} \psi = q(x, t)$$

admits no solution on any open set. Note that if q is *analytic*, however, the Cauchy-Kovalevskaya theorem implies the existence of a unique solution.

problem for Eq. (4.1) in the superluminal regime suggests that we forgo the quest for smooth, classical answers. However, the wave equation is well-posed even for faster-than-light sources if we allow for properly defined *generalized* or *weak solutions* that may not be continuously differentiable but, nonetheless, satisfy a partial differential equation (as well as its initial and boundary conditions) in some precisely defined sense.

From a number of eligible methods we choose a Green's function formulation, which uses the impulse response of the system to establish a link between state ψ and source q . Solutions assume the general form

$$\psi(y) = \int G(y, x) \cdot q(x) \, dx \quad (4.6)$$

where G denotes the Green's function, emphasizing the necessity of dividing the physical process to be modeled into cause and effect relations. This mirrors, more so than any other analytical approach, proceedings in experimental physics as the same is – more or less obviously – true for the preparation and execution of any applied experiment [149], not least our own.

4.2 The Beauty of Properly Posed Conditions

4.2.1 On Characteristics, Singularities and the Classification of Second Order Partial Differential Equations

In the realm of *ordinary* differential equations, the question of existence, uniqueness, and stability of a solution has a clear and satisfactory answer with the Picard-Lindelöf theorem¹¹ which establishes that the problem is properly posed assuming merely

¹¹Depending on the author, the Picard-Lindelöf theorem is also known as the *Cauchy-Picard*, *Cauchy-Lipschitz*, and *Cauchy-Lipschitz-Picard theorem*, often adorned with qualifiers such as ‘existence’, ‘uniqueness’, or both. Some texts call it simply *Picard's* or *uniqueness and existence theorem*.

Lipschitz continuity¹². However, the qualitative behavior of systems involving *partial* differential equations is far more subtle and varied. As indicated earlier, the latter are meaningful within the framework of mathematical physics only if they arise as initial or boundary value problems. Specifically, we seek an unknown function ψ that satisfies, throughout $\mathbb{R}^n \times (0, \infty)$, an “indefinite” equation such as Eq. (4.3) along with some “definite” accessory conditions (Eq. (4.4)) [127]. The system of equations is well-posed if the constraints admit precisely one stable solution. According to Hadamard [124, 126, 127], the simplest such problem is Cauchy’s since “it represents, for partial differential equations, the exact analogue of the well-known fundamental one in ordinary differential equations” [127]. These are precisely the initial conditions for which we seek to solve the three-dimensional wave equation, which makes it necessary to examine, in some detail, the problem named after the prolific 19th-century French mathematical analyst.

A Brief Survey of the Cauchy Problem

Consider a second-order partial differential equation for the function ψ in the independent variables y and $\mathbf{x} = x_1, \dots, x_n$. Assuming that this equation can be solved explicitly for $\partial^2\psi/\partial y^2$, it may be stated in the form

¹²A function $\psi(x)$ is *Lipschitz continuous* in its domain of definition X if

$$|\psi(x_1) - \psi(x_2)| \leq L\|x_1 - x_2\|_2 \quad \forall x_1, x_2 \in X$$

holds for some *Lipschitz constant* $L > 0$. ($\|\cdot\|_2$ denotes the Euclidean norm, *i.e.*, the p -norm for $p = 2$.) Similarly, the gradient vector $\nabla\psi(x)$ of a differentiable function $\psi(x)$ is Lipschitz continuous in X if

$$\|\nabla\psi(x_1) - \nabla\psi(x_2)\|_2 \leq L\|x_1 - x_2\|_2 \quad \forall x_1, x_2 \in X$$

where $L > 0$. In other words, for every pair of points x_1 and x_2 on the graph of ψ the absolute value of the slope of the line connecting them is no greater than some real number. L , the modulus of uniform continuity, is the smallest such bound and limits how fast $\psi(x)$ or $\nabla\psi(x)$ may change.

$$\frac{\partial^2}{\partial y^2} \psi = q \left(y, x_1, x_2, \dots, x_n, \psi, \frac{\partial}{\partial y} \psi, \frac{\partial}{\partial x_1} \psi, \frac{\partial}{\partial x_2} \psi, \dots, \frac{\partial}{\partial x_n} \psi, \right. \\ \left. \frac{\partial^2}{\partial x_1 \partial y} \psi, \frac{\partial^2}{\partial x_2 \partial y} \psi, \dots, \frac{\partial^2}{\partial x_n \partial y} \psi, \frac{\partial^2}{\partial x_1^2} \psi, \frac{\partial^2}{\partial x_2^2} \psi, \dots, \frac{\partial^2}{\partial x_n^2} \psi \right). \quad (4.7)$$

The Cauchy or initial value problem¹³ consists of finding a unique and stable solution ψ to Eq. (4.7) which satisfies the initial conditions

$$\psi|_{y=y_0} = f(x_1, \dots, x_n), \quad \frac{\partial}{\partial y} \psi|_{y=y_0} = g(x_1, \dots, x_n) \quad (4.8)$$

on some hypersurface¹⁴ $S_0 \subset \mathbb{R}^n$. Depending on the dimensions n of \mathbb{R} , S_0 may be a line, curve, or surface.

Hence, we recognize Hadamard's *exemplum non gratum* in Section 4.1.2 as the Cauchy problem for the Laplace equation in two variables. It is elementary to verify that the differential equation as well as the initial conditions are real analytic¹⁵ and have, therefore, derivatives of all orders. However, as demonstrated in Fig. 4.3 (right), the solution loses the property of continuous dependence on the Cauchy data.

¹³Although some authors distinguish between the two, we consider the terms *Cauchy* and *initial-value problem* to be synonymous. Treating the two-dimensional case [132] states, “It is not essential that [the initial values] be given along the line $y = y_0$; they may very well be prescribed along some curve L_0 in the xy -plane. In such a context, the problem is called the *Cauchy problem* instead of the initial value problem, although the two names are actually synonymous.”

¹⁴We define a hypersurface (very loosely) as a manifold of dimension $n - 1$ which is embedded in an n -dimensional ambient environment. A manifold, in turn, is taken to be a topological space that locally ‘resembles’ Euclidean space near each point. Hence, a curve that resides in the xy -plane is a one-dimensional manifold, a surface in free space is one of dimension two. All hypersurfaces can be defined by a single implicit equation of the form $\phi(x_1, x_2, \dots, x_n) = 0$; for instance, $x_1^2 + x_2^2 + \dots + x_n^2 - 1 = 0$ denotes a hypersphere.

¹⁵In the simplest of terms, a function ψ is *real analytic* on an open set X if it is locally given by a convergent power series, *i.e.*,

$$\psi(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + \dots$$

where the coefficients $a_0, a_1, \dots, a_n \in \mathbb{R}$ and the series converges to $\psi(x)$ in the neighborhood of x_0 . Notice that analyticity is a much stronger condition than smoothness, *i.e.*, the requirement that ψ be n times – or even infinitely – differentiable.

In an effort to understand why seemingly similar systems display vastly different behaviors and to bring much needed formalism to a disorderly branch of applied mathematics, Hadamard proposed, around 1911 [127], “the classification of linear partial differential equations of the second order into different types”. Inspired by the description of conic sections in analytic geometry, the technique is based on reducing the general problem 4.7 in two variables

$$A \frac{\partial^2}{\partial x^2} \psi + 2B \frac{\partial^2}{\partial x \partial y} \psi + C \frac{\partial^2}{\partial y^2} \psi = q\left(x, y, \psi, \frac{\partial}{\partial x} \psi, \frac{\partial}{\partial y} \psi\right) \quad (4.9)$$

where A, B, C are real valued functions of x and y , to *standard* or *canonical* form. We note, in passing, that Eq. (4.9) is linear if $q\left(x, y, \psi, \partial\psi/\partial x, \partial\psi/\partial y\right)$ can be written $D \partial\psi/\partial x + E \partial\psi/\partial y + F\psi + G$, where D, E, F and G are arbitrary functions of x and y *alone*, and semilinear if not. Since the equation is of quadratic form, its discriminant is given by

$$B^2 - AC. \quad (4.10)$$

With the dual aim of identifying invariant properties on the left-hand side of Eq. (4.9) and, if possible, writing the equation in simpler form, we introduce the one-to-one transformation

$$\xi = \xi(x, y), \quad \eta = \eta(x, y) \quad (4.11)$$

along with the condition that the Jacobian

$$J = \begin{vmatrix} \frac{\partial}{\partial x} \xi & \frac{\partial}{\partial y} \xi \\ \frac{\partial}{\partial x} \eta & \frac{\partial}{\partial y} \eta \end{vmatrix} \quad (4.12)$$

is bounded and nonzero. Assuming that ξ and η are twice differentiable, we use the chain rule to find that

$$\frac{\partial}{\partial x} \psi = \frac{\partial}{\partial \xi} \psi \frac{\partial}{\partial x} \xi + \frac{\partial}{\partial \eta} \psi \frac{\partial}{\partial x} \eta \quad \text{and} \quad \frac{\partial}{\partial y} \psi = \frac{\partial}{\partial \xi} \psi \frac{\partial}{\partial y} \xi + \frac{\partial}{\partial \eta} \psi \frac{\partial}{\partial y} \eta. \quad (4.13)$$

Then,

$$\frac{\partial^2}{\partial x^2} \psi = \frac{\partial}{\partial x} \left[\frac{\partial}{\partial \xi} \psi \frac{\partial}{\partial x} \xi + \frac{\partial}{\partial \eta} \psi \frac{\partial}{\partial x} \eta \right] \quad (4.14)$$

$$\begin{aligned}
 &= \frac{\partial}{\partial x} \left(\frac{\partial}{\partial \xi} \psi \right) \frac{\partial}{\partial x} \xi + \frac{\partial}{\partial \xi} \psi \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} \xi \right) \\
 &\quad + \frac{\partial}{\partial x} \left(\frac{\partial}{\partial \eta} \psi \right) \frac{\partial}{\partial x} \eta + \frac{\partial}{\partial \eta} \psi \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} \eta \right) \\
 &= \frac{\partial}{\partial x} \xi \left[\frac{\partial}{\partial x} \xi \frac{\partial}{\partial \xi} \left(\frac{\partial}{\partial \xi} \psi \right) + \frac{\partial}{\partial x} \eta \frac{\partial}{\partial \eta} \left(\frac{\partial}{\partial \xi} \psi \right) \right] \\
 &\quad + \frac{\partial}{\partial x} \eta \left[\frac{\partial}{\partial x} \eta \frac{\partial}{\partial \eta} \left(\frac{\partial}{\partial \eta} \psi \right) + \frac{\partial}{\partial x} \xi \frac{\partial}{\partial \xi} \left(\frac{\partial}{\partial \eta} \psi \right) \right] \\
 &\quad + \frac{\partial}{\partial \xi} \psi \frac{\partial^2}{\partial x^2} \xi + \frac{\partial}{\partial \eta} \psi \frac{\partial^2}{\partial x^2} \eta \\
 &= \frac{\partial^2}{\partial \xi^2} \psi \left(\frac{\partial}{\partial x} \xi \right)^2 + 2 \frac{\partial^2}{\partial \xi \partial \eta} \psi \frac{\partial}{\partial x} \xi \frac{\partial}{\partial x} \eta + \frac{\partial^2}{\partial \eta^2} \psi \left(\frac{\partial}{\partial x} \eta \right)^2 \\
 &\quad + \frac{\partial}{\partial \xi} \psi \frac{\partial^2}{\partial x^2} \xi + \frac{\partial}{\partial \eta} \psi \frac{\partial^2}{\partial x^2} \eta
 \end{aligned}$$

and, similarly,

$$\begin{aligned}
 \frac{\partial^2}{\partial y^2} \psi &= \frac{\partial^2}{\partial \xi^2} \psi \left(\frac{\partial}{\partial y} \xi \right)^2 + 2 \frac{\partial^2}{\partial \xi \partial \eta} \psi \frac{\partial}{\partial y} \xi \frac{\partial}{\partial y} \eta + \frac{\partial^2}{\partial \eta^2} \psi \left(\frac{\partial}{\partial y} \eta \right)^2 \\
 &\quad + \frac{\partial}{\partial \xi} \psi \frac{\partial^2}{\partial y^2} \xi + \frac{\partial}{\partial \eta} \psi \frac{\partial^2}{\partial y^2} \eta
 \end{aligned} \tag{4.15}$$

$$\begin{aligned}
 \frac{\partial^2}{\partial x \partial y} \psi &= \frac{\partial^2}{\partial \xi^2} \psi \frac{\partial}{\partial x} \xi \frac{\partial}{\partial y} \xi + \frac{\partial^2}{\partial \xi \partial \eta} \psi \left(\frac{\partial}{\partial x} \xi \frac{\partial}{\partial y} \eta + \frac{\partial}{\partial x} \eta \frac{\partial}{\partial y} \xi \right) \\
 &\quad + \frac{\partial^2}{\partial \eta^2} \psi \frac{\partial}{\partial x} \eta \frac{\partial}{\partial y} \eta + \frac{\partial}{\partial \xi} \psi \frac{\partial^2}{\partial y^2} \xi + \frac{\partial}{\partial \eta} \psi \frac{\partial^2}{\partial y^2} \eta.
 \end{aligned} \tag{4.16}$$

It is now easy to see that, in terms of ξ and η ,

$$\alpha \frac{\partial^2}{\partial \xi^2} \psi + 2\beta \frac{\partial^2}{\partial \xi \partial \eta} \psi + \gamma \frac{\partial^2}{\partial \eta^2} \psi = Q \left(\xi, \eta, \psi, \frac{\partial}{\partial \xi} \psi, \frac{\partial}{\partial \eta} \psi \right), \tag{4.17}$$

for some function Q , where α , β and γ are defined as

$$\begin{aligned}
 \alpha &= A \left(\frac{\partial}{\partial x} \xi \right)^2 + 2B \frac{\partial}{\partial x} \xi \frac{\partial}{\partial y} \xi + C \left(\frac{\partial}{\partial y} \xi \right)^2, \\
 \beta &= A \frac{\partial}{\partial x} \xi \frac{\partial}{\partial x} \eta + B \left(\frac{\partial}{\partial x} \eta \frac{\partial}{\partial y} \eta + \frac{\partial}{\partial y} \xi \frac{\partial}{\partial x} \eta \right) + C \frac{\partial}{\partial y} \xi \frac{\partial}{\partial y} \eta, \\
 \gamma &= A \left(\frac{\partial}{\partial x} \eta \right)^2 + 2B \frac{\partial}{\partial x} \eta \frac{\partial}{\partial y} \eta + C \left(\frac{\partial}{\partial y} \eta \right)^2.
 \end{aligned} \tag{4.18}$$

The discriminant of Eq. (4.17) is found to be

$$\beta^2 - \alpha\gamma = J^2(B^2 - AC), \tag{4.19}$$

where J denotes the Jacobian defined in Eq. (4.12), indicating that Eq. (4.10) is invariant under the proposed coordinate transformation as long as J does not vanish [150].

The *type* of a second-order partial differential equation is determined by the sign of Eq. (4.10) and is decisive in establishing the kind of hyperboundary conditions that serve, in a natural way, to determine a solution uniquely [151]. Eq. (4.9) is said to be *elliptic* if, for all x and y in the domain of definition, $B^2 - AC < 0$, *parabolic* if $B^2 - AC = 0$, and *hyperbolic* otherwise. Some equations are of mixed type; for instance, the Tricomi equation

$$\frac{\partial^2}{\partial x^2} + x \frac{\partial^2}{\partial y^2} = 0, \quad (4.20)$$

which plays a central role in the mathematical analysis of transonic flow, is elliptic in the half plane $x > 0$, parabolic along the ‘sonic line’ $x = 0$ and hyperbolic if $x < 0$.

Unlike the Tricomi equation, the wave equation (4.1) is hyperbolic over its entire domain of definition since $A = -1$, $B = 0$ and $C = 1/c^2$ (where c , as before, is the propagation speed for the medium) are constant for all \mathbf{r} and t . Hyperbolic models are, in general, well understood and the initial value problem is, for the most part, properly posed. Not surprisingly, the solutions are ‘wavelike’, which is to say that disturbances in the Cauchy data travel, as shown in Fig. 4.5, with undulating motion and at finite speed c away from their source of origin. In consequence, perturbations in the conditions are not felt everywhere at once; they propagate, in well ordered fashion, as manifolds of $n - 1$ spatial dimensions termed *characteristics*.

The Importance of Being Characteristic

Mathematically, a characteristic represents a boundary across which solutions to partial differential equations with smooth parameters can have discontinuous derivatives [150, 152] or, in the words of [127], “where two solutions of the equation can



Figure 4.5: A single drop of water falling into a pond. The image shows that disturbances in the initial or boundary data travel outward along characteristic curves (in this case concentric circles). Since sound travels faster than the falling drop does, perturbations start to form before it reaches the surface, just as an ambulance can be heard before it passes. Moreover, the rather remarkable photograph demonstrates that, in two dimensions, the solution to the wave equation is non-zero *almost everywhere* in the region $0 \leq \mathbf{r} < ct$. This is in stark contrast to the one- or three-dimensional case, where the influence of the perturbation can only be felt in the vicinity of the characteristics $\mathbf{r} = \pm ct$; the region in between is appropriately called the ‘zone of silence’. Credit: <http://water.engineering.utoronto.ca>

touch”. This is, of course, equivalent to saying that α , γ , or both in Eq. (4.17) must be zero for a manifold to be characteristic; if not, it is considered *noncharacteristic*. In this sense, the Mach (or Čerenkov) cone generated by a point source that travels faster than the wave speed (shown in Fig. 4.6) is simply a wavefront that expands as a characteristic hypersurface and across which the solution ψ may experience jump discontinuities.

To find the characteristics for the general second-order equation (4.9) in the unknowns x and y , we suppose that none of A , B , and C is zero and choose ξ and η such that the coefficients α and γ in Eq. (4.17) vanish. From Eq. (4.18) we have

$$\begin{aligned}\alpha &= A \left(\frac{\partial}{\partial x} \xi \right)^2 + 2B \frac{\partial}{\partial x} \xi \frac{\partial}{\partial y} \xi + C \left(\frac{\partial}{\partial y} \xi \right)^2 = 0, \\ \gamma &= A \left(\frac{\partial}{\partial x} \eta \right)^2 + 2B \frac{\partial}{\partial x} \eta \frac{\partial}{\partial y} \eta + C \left(\frac{\partial}{\partial y} \eta \right)^2 = 0.\end{aligned}\tag{4.21}$$

Since these equations are of the same general form, we may write them as

$$A \left(\frac{\partial}{\partial x} \zeta \right)^2 + 2B \frac{\partial}{\partial x} \zeta \frac{\partial}{\partial y} \zeta + C \left(\frac{\partial}{\partial y} \zeta \right)^2 = 0, \quad (4.22)$$

where ζ stands for either of the functions ξ or η . Dividing by $(\partial\zeta/\partial y)^2$ yields

$$A \frac{\left(\frac{\partial}{\partial x} \zeta \right)^2}{\left(\frac{\partial}{\partial y} \zeta \right)^2} + 2B \frac{\left(\frac{\partial}{\partial x} \zeta \right)}{\left(\frac{\partial}{\partial y} \zeta \right)} + C = 0 \quad (4.23)$$

such that the curve $\zeta = \text{constant}$ is defined by

$$d\zeta = \frac{\partial}{\partial x} \zeta dx + \frac{\partial}{\partial y} \zeta dy = 0. \quad (4.24)$$

Then, along ζ ,

$$\frac{dy}{dx} = - \frac{\left(\frac{\partial}{\partial x} \zeta \right)}{\left(\frac{\partial}{\partial y} \zeta \right)} \quad (4.25)$$

and Eq. (4.23) assumes the quadratic form

$$A \left(\frac{dy}{dx} \right)^2 - 2B \left(\frac{dy}{dx} \right) + C = 0 \quad (4.26)$$

with roots given by

$$\frac{dy}{dx} = \frac{B \pm \sqrt{B^2 - AC}}{A}. \quad (4.27)$$

Characteristic equations (4.27) constitute a pair of ordinary differential equations for families of curves in the xy -plane along which ξ and η are constant. Integration with respect to both independent variables yields the characteristics: As mentioned above, they are lines in one, curves in two and surfaces in three space dimensions.

Elliptic equations have no real-valued characteristics since $B^2 < AC$, which results in a negative discriminant. For the Laplace equation (4.3), commonly considered the text book case, $B^2 - AC = -1$ such that the system is elliptic in the entire plane. Since disturbances in the initial data can only propagate as characteristic manifolds, solutions to elliptic equations cannot have discontinuous derivatives.

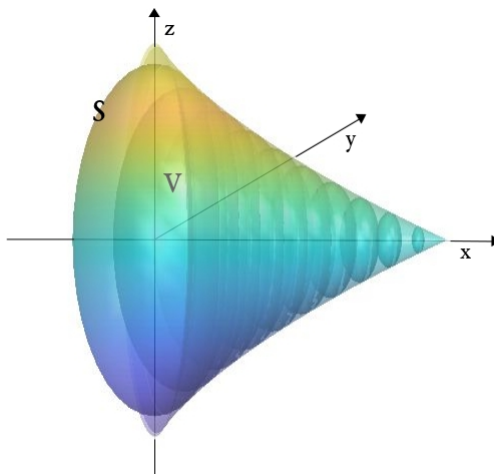


Figure 4.6: The conical hypersurface generated by a point source that both exceeds the characteristic wave speed c and accelerates. Depending on the context, we may think of such manifolds as wavefronts or shock waves.

This indicates that they are well suited to describe equilibrium states where any sharp switches have been smoothed out. As evidenced above and in Section 4.1.2, solutions to the Cauchy problem are, in general, unpredictable for elliptic equations which leaves them ill-posed in the sense of Hadamard.

If $B^2 = AC$, as is the case for parabolic equations, the two characteristics coalesce since the discriminant is zero. The diffusion equation $\partial\psi/\partial t - A\partial^2\psi/\partial x^2 = 0$ serves as an example since $B = AC = 0$ which renders it parabolic across the xt -plane. In one spatial dimension, its (sole) characteristic line is given by $dt^2 = 0$ which is to say that information propagates at infinite speed along $t = \text{constant}$. Although parabolic equations are used to describe a wide variety of time-dependent phenomena in applied science, they tend to be the least foreseeable amongst the three types. Tihonov [153] showed in 1935 that the initial value problem for the heat equation has a unique solution in the class of functions which grow no faster than $\exp\{a|x|^2\}$, where a is a constant. Although this result was extended a few times in the following decades, uniqueness – not to mention continuity – can, in general, not be assumed for parabolic equations. Consequently, the Cauchy problem may – or may not – be

well behaved.

Equations of hyperbolic type, for which $B^2 > AC$, possess two real distinct families of characteristics. Consider, for instance, the homogeneous wave equation in one spatial dimension,

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \psi(x, t) = \frac{\partial^2}{\partial x^2} \psi(x, t) \quad \text{in } \mathbb{R} \times (0, \infty). \quad (4.28)$$

According to Eq. (4.26),

$$-\left(\frac{dt}{dx}\right)^2 + \frac{1}{c^2} = 0 \quad (4.29)$$

or

$$dx^2 - c^2 dt^2 = 0, \quad (4.30)$$

which reduces to

$$dx \pm c dt = 0. \quad (4.31)$$

The integrals are the straight lines

$$x + ct = \xi, \quad x - ct = \eta, \quad (4.32)$$

where ξ and η are constant.

Now take the homogeneous wave equation in free space, *i.e.*,

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2\right) \psi(\mathbf{r}, t) = 0 \quad \text{in } \mathbb{R}^3 \times (0, \infty). \quad (4.33)$$

By virtue of spherical symmetry, the Laplace operator assumes the form

$$\begin{aligned} \nabla^2 \bar{\psi}(r, t) &= \frac{\partial^2}{\partial r^2} \bar{\psi} + \frac{2}{r} \frac{\partial}{\partial r} \bar{\psi} \\ &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \bar{\psi} \right) \\ &= \frac{1}{r} \frac{\partial^2}{\partial r^2} (r \bar{\psi}), \end{aligned} \quad (4.34)$$

such that Eq. (4.33) may be rewritten as

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \bar{\psi} - \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) \bar{\psi} = 0. \quad (4.35)$$

In both, Eq. (4.34) and Eq. (4.35), the overline denotes a spherically symmetrical solution $\bar{\psi}(r, t)$ and the polar coordinate $r = \sqrt{x^2 + y^2 + z^2}$ stands for the sphere of interest's radius¹⁶. The last of identities (4.34) can be used to show that the change of variables $\Psi = r\bar{\psi}$ leads to the one-dimensional wave equation (4.28). Although the suggested substitution is not particularly difficult, it is nevertheless shrouded in considerable confusion¹⁷. A dissection of Eq. (4.34) yields

$$\begin{aligned} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \bar{\psi} \right) &= \frac{1}{r^2} \left(r^2 \frac{\partial^2}{\partial r^2} \bar{\psi} + \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} \bar{\psi} \right) \\ &= \frac{\partial^2}{\partial r^2} \bar{\psi} + \frac{2r}{r^2} \frac{\partial}{\partial r} \bar{\psi} \\ &= \frac{\partial^2}{\partial r^2} \bar{\psi} + \frac{2}{r} \frac{\partial}{\partial r} \bar{\psi} \end{aligned} \quad (4.36)$$

and

$$\begin{aligned} \frac{1}{r} \frac{\partial^2}{\partial r^2} (r\bar{\psi}) &= \frac{1}{r} \frac{\partial}{\partial r} \left\{ \frac{\partial}{\partial r} (r\bar{\psi}) \right\} \\ &= \frac{1}{r} \frac{\partial}{\partial r} \left\{ r \frac{\partial}{\partial r} \bar{\psi} + \bar{\psi} \frac{\partial}{\partial r} r \right\} \\ &= \frac{1}{r} \left\{ r \frac{\partial^2}{\partial r^2} \bar{\psi} + \frac{\partial}{\partial r} r \frac{\partial}{\partial r} \bar{\psi} + \bar{\psi} \frac{\partial^2}{\partial r^2} r + \frac{\partial}{\partial r} r \frac{\partial}{\partial r} \bar{\psi} \right\} \\ &= \frac{1}{r} \left\{ r \frac{\partial^2}{\partial r^2} \bar{\psi} + 2 \frac{\partial}{\partial r} r \frac{\partial}{\partial r} \bar{\psi} \right\} \\ &= \frac{\partial^2}{\partial r^2} \bar{\psi} + \frac{2}{r} \frac{\partial}{\partial r} \bar{\psi}, \end{aligned} \quad (4.37)$$

¹⁶The distinction between r and \mathbf{r} is crucial. As defined above and elsewhere, r stands for the *radial component* in a *spherical* coordinate system. \mathbf{r} , on the other hand, denotes a point (x, y, z) in *Cartesian space*.

¹⁷Ockendon *et al.* [150], for instance, suggest writing $\bar{\psi} = r\Psi$ (in our notation), which completely defeats the purpose of the operation. To the credit of the authors, however, the preface to the revised edition states “that on many occasions our zeal in writing the first edition overstretched our accuracy”. This seems to be one such occasion. [154] – ordinarily meticulous to a fault – states merely, “It is clear that [Eq. (4.35)] is the one-dimensional wave equation for $r\psi$ ” (again, our numbering).

which establishes the equalities. Similarly,

$$\frac{\partial^2}{\partial t^2} (r\bar{\psi}) = r \frac{\partial^2}{\partial t^2} \bar{\psi} \quad (4.38)$$

since all derivatives of r with respect to time vanish. It follows that

$$\frac{1}{c^2} \frac{1}{r} \frac{\partial^2}{\partial t^2} (r\bar{\psi}) = \frac{1}{r} \frac{\partial^2}{\partial r^2} (r\bar{\psi}) \quad (4.39)$$

and, after multiplying both sides by r and performing the change of variables $\Psi = r\bar{\psi}$ we arrive at the one-dimensional wave equation (4.28). Transformations (4.34)–(4.39) indicate that the (spherical) characteristic hypersurface for Eq. (4.35) is given by

$$r + ct = \xi, \quad r - ct = \eta, \quad (4.40)$$

where the polar coordinate r is taken to be the location of the disturbance's origin, “which reflects the merciful spreading out” [150] of sound or light waves in space. We note – merely parenthetically at this time – that the singularity at $r = 0$ raises the prospect of focusing in the time domain which is absent in problems of only one spatial dimension [150].

Surprising Effects of Dimensionality

It is not particularly surprising that a simple substitution transforms the spherically symmetric wave equation (4.33) into expression (4.28). That the *qualitative* nature of linear wave propagation depends on the *dimensionality* of the problem, however, is less inferential even though the effect is readily observable in nature. (See, for example, Fig. 4.5 and the associated caption.) By analogy with the one-dimensional case, a perturbation that appears at $r = t = 0$ in free space will develop with time into one that is localized near the spherical ‘shell’ $x_P^2 + y_P^2 + z_P^2 = c^2 t_P^2$. (We will remark on the delicate subject of notation in Section 4.3. For the time being it suffices to recognize that all quantities pertinent to a *field* or *observation point* are denoted with the subscript P.) Hence, a bystander who wishes to sample the signal

must be stationed on the characteristic hypersurface S enclosing the spherical volume V whose radius is given by c times t_P (Fig. 4.6). The implications of this insight are quite striking: A disturbance that propagates through three-dimensional space is felt at field point \mathbf{r}_P *only* at time $t_P = |\mathbf{r}_P|/c$. As will become evident in the sections to follow and, in particular, Chapter 6 which is dedicated to point sources, this is equivalent to stating that the fundamental solution of Eq. (4.1) has support only on the characteristic conoid [155] and explains, *inter alia*, why signals emitted by a point charge are – necessarily – ‘sharp’.

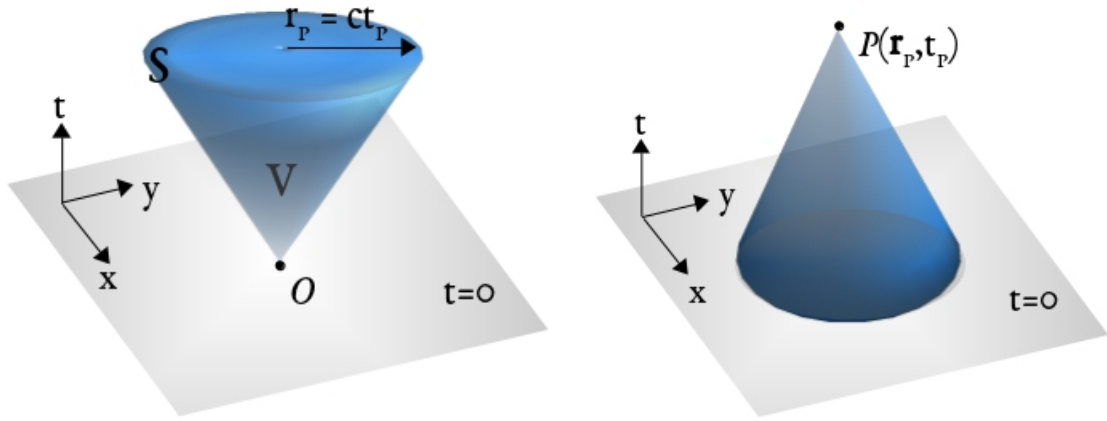


Figure 4.7: (left) The *range of influence* of a disturbance created at point $O(0,0)$ in the xy -plane. In a space of even dimensions the perturbation is felt *everywhere* within the cone of radius $r_P \equiv |\mathbf{r}_P| = ct_P$. (right) The *domain of dependence* of point $P(\mathbf{r}_P, t_P)$ reveals that the solution of $\psi(\mathbf{r}_P, t_P)$ depends solely on the initial data imposed on the shaded area.

The evolution of a two-dimensional axisymmetric (*i.e.*, cylindrical) wave with inception at the origin O of the xy -plane is quite different. Although the perturbation has a sharp leading front or “surface of constant phase” just as in one and three dimensions, it is felt *everywhere* within the cone $x_P^2 + y_P^2 \leq c^2 t_P^2$, termed the *range of influence* and shown in Fig. 4.7 (left). If a pebble is thrown into a pond as in Fig. 4.5, new waves emanate from the center of disturbance even when the stone has vanished. On the other hand, if a field point $P(\mathbf{r}_P, t_P)$ is situated beyond the range of influence, it can not be affected by the initial data imposed on O . This indicates

that the solution $\psi(\mathbf{r}_P, t_P)$ in the right-hand panel of the same figure depends *solely* on the initial data prescribed on the gray-shaded area which can be obtained by tracing the characteristic curves back from P to the xy -plane. The resulting cone is known as the *domain of dependence*.

The – rather spectacular – dependence of the wave equation on the number of independent variables is, at the very least, mentioned in most advanced texts on partial differential equations [62, 132–135] and treated in depth by those dedicated to the Cauchy problem [124, 126, 127, 150, 152]. Discovered by Christiaan Huygens in 1678 and later formalized by Kirchhoff, the principle that bears the former’s name was not fully understood until Hadamard (who else!) stated around the turn of the twentieth century that “Huygens’ (minor) premise holds for no phenomenon governed by a linear partial differential equation of the second order with an odd number of independent variables”¹⁸. To quote Section 2.4.1 of [62] in our notation,

[W]e observe that if n is odd and $n \geq 3$, the data g and h at a given point $\mathbf{r} \in \mathbb{R}^n$ affect the solution ψ only on the *boundary* $\{(\mathbf{r}_P, t_P) \mid t_P > 0, |\mathbf{r}_P - \mathbf{r}| = ct_P\}$ of the cone $C = \{(\mathbf{r}_P, t_P) \mid t_P > 0, |\mathbf{r}_P - \mathbf{r}| < ct_P\}$. On the other hand, if n is even, the data g and h affect ψ within all of C . In other words, a “*disturbance*” *originating at \mathbf{r} propagates along a sharp wavefront in odd dimensions, but in even dimensions it continues to have effects even after the leading edge of the wavefront passes*. This is *Huygens’ principle*. (Italics by the original author.)

The significance of the theorem is that waves *not* governed by Huygens’ principle (*i.e.*, those for which n is even) undergo diffusion [124, 126, 127], a physical effect which appears quite naturally in the form of a “residual integral” [123, 127, 156] in mathematical solutions to problems of even spatial dimensions. An associated conjecture posited by Hadamard remains “in spite of its age” an area of active research [157, 158].

¹⁸An odd number of independent variables implies, of course, an even number of spatial dimensions.

In light of the above discussion, it should not come as a shock that characteristics of all spatial dimensions are closely associated with the propagation of singularities. Along a *noncharacteristic* manifold the Cauchy data uniquely determine the second derivatives of a solution $\psi(\mathbf{r}_P, t_P)$. *Characteristic* ones, on the other hand, allow for jump discontinuities across S . Hence, Eqs. (4.40) divide n -dimensional space into two regions I and II, where the former has been reached by the disturbance, the latter not. Each contains a distinct solution, ψ_I and ψ_{II} , respectively, to the differential equation; together, they define a function $\psi = \psi_I \cup \psi_{II}$ which is singular on S . Mathematically (not to mention physically) this poses a dilemma since ψ can not be considered a solution, generalized or otherwise, of Eqs. (4.7)–(4.8) unless that equation *still holds on S* in some precisely defined sense. This requires certain transition conditions, the simplest of which arises when ψ_I and ψ_{II} as well as their first derivatives coincide on S [142].

In more general terms, we recognize that if an equation of the form (4.7) is to carry a signal it must have *real characteristics*, *i.e.*, be either hyperbolic or parabolic, and only the former will generally lead to a finite propagation speed [159]. Finally, in view of the above discussion, it is apposite to remark that cutting a slice through a three-dimensional problem is *in no way* equivalent to solving the two-dimensional problem.

4.2.2 Sofia Kovalevskaya's Solitary Quest

Systematic investigations into initial and boundary value problems of mathematical physics date back as far as the 17th century with contributions from Huygens, d'Alembert, Riemann, and Kirchhoff. However, it was Sofia Kovalevskaya [160] who, in her doctoral thesis, delivered the first existence theorem for solutions to the Cauchy problem. The immensely talented Russian mathematician used the method of ma-

majorants¹⁹ and a normalization technique due to Carl Gustav Jacobi to show, quite generally, the local existence of unique solutions to a second-order system of partial differential equations with initial conditions on a noncharacteristic surface. Formally, the theorem states that Eq. (4.7) subject to the Cauchy data (4.8) has a unique solution in some neighborhood of the point $(y_0, x_{1_0}, x_{2_0}, \dots, x_{n_0})$ provided that all coefficients which appear in q as well as the functions f and g are analytic with respect to their arguments and the manifold $y_0 = S_0$ on which the initial conditions are prescribed is nowhere characteristic.

As argued in the previous section, characteristics are manifolds across which discontinuities in the second derivatives of the dependent variables propagate or, equivalently, where two solutions of the differential equation “touch” [127]. This contact results invariably in singular solutions, which makes clear why the hypersurface S_0 that bears the initial data can – by definition – not be characteristic. The theorem’s restriction to analytic functions, on the other hand, is due to the fact that the continuation of the latter is determined: If $f(x)$ is analytic in (a, b) , knowledge of the values in any sub-interval (a', b') – regardless how small – allows its calculation for *every* point in (a, b) . For functions that are not analytic, continuation has generally no meaning. A proof of the Cauchy-Kovalevskaya theorem exceeds the scope of the

¹⁹The method of majorants was introduced by Cauchy in 1842 and features prominently in Poincaré’s 1879 dissertation. Generally speaking, it is used to show that a power series in the independent variable $z \in \mathbb{C}$ (derived by the method of undetermined coefficients) that satisfies the differential equation does have a *definite domain of convergence*. Consider

$$f(z) = \sum_{k=0}^{\infty} c_k z^k$$

to be a power series with a positive radius of convergence and

$$g(z) = \sum_{k=0}^{\infty} C_k z^k$$

one with nonnegative coefficients and radius of convergence R . $g(z)$ is a majorant of $f(z)$ if $|c_k| \leq C_k \forall k$.

present work; instead, we refer the interested reader to [140], 14ff., [127], 12ff. and, of course, [160].

In retrospect, Kovalevskaya's theorem seems to have little practical importance since, as evidenced by the Cauchy problem for the Laplace equation (4.3)-(4.4), it does not distinguish between properly and improperly posed problems. More precisely, it covers situations where the solution does not continuously depend on the conditions even if its power series converges to $\psi(\mathbf{x}, y)$ *everywhere*. In addition, due to the exclusion of all but analytic functions, "Kowalewskii's [sic] theorem only gives [answers] for which the concept of the domains of dependence and [range of] influence, so important for hyperbolic equations, is meaningless" [161]. Historically, however, it represents the first existence theorem for a *general category* of partial differential equations. At the very least, it establishes that, within the class of analytic answers to (equally analytic) problems, the conditions required for a particular solution correspond in number to the order of the system and involve one less independent variable than the equation [151]. In the present context, Kovalevskaya's theorem assures us that the wave equation (4.1) has a unique solution *irrespective of the speed of the source* as long as q , together with the initial data, is analytic.

4.2.3 Time- and Space-like Solutions

As we already know, it is not sufficient to take the initial surface S_0 distinct from a characteristic in order to achieve a well-posed problem. Examining – one more time – Hadamard's pièce de résistance (4.3)–(4.4) and, in particular, the result $\psi(x, y) = k^{-2} \sinh kx \sin ky$, it is apparent that the system loses the property of continuous dependence on the Cauchy data *because it admits exponential solutions*.

To support this claim formally, we consider again the homogeneous form of the wave equation in free space given by Eq. (4.33)²⁰. Furthermore, we define a hyper-

²⁰The discussion to follow is much indebted to [151] and [154].

surface S such that

$$S = \sum_{j=1}^n s_j x_j - s_0 t, \quad (4.41)$$

where s_0, \dots, s_n are assumed to be real and n denotes, as before, the number of spatial dimensions. If exponentially growing solutions of the form

$$\psi(\mathbf{r}, t) = e^{S+iM}, \quad M = \sum_{j=1}^n \mu_j x_j - \mu_0 t, \quad (4.42)$$

$\mu_0, \dots, \mu_n \in \mathbb{R}$, can be found to satisfy the differential equation along with the initial conditions assigned on $S = 0$, the Cauchy problem must be *improperly set* since it violates Hadamard's postulate of stability.

Substitution of the proposed solution into Eq. (4.33) yields

$$\begin{aligned} & \nabla^2 e^{S+iM} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} e^{S+iM} \\ &= \nabla^2 \exp\left(\sum_{j=1}^n s_j x_j - s_0 t + i\left(\sum_{j=1}^n \mu_j x_j - \mu_0 t\right)\right) - \\ & \quad \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \exp\left(\sum_{j=1}^n s_j x_j - s_0 t + i\left(\sum_{j=1}^n \mu_j x_j - \mu_0 t\right)\right) \\ &= 0 \end{aligned} \quad (4.43)$$

while integrating twice with respect to time and space results in

$$\begin{aligned} & \sum_{j=1}^n (s_j + i\mu_j)^2 e^{S+iM} - \frac{1}{c^2} (-s_0 - i\mu_0)^2 e^{S+iM} \\ &= \sum_{j=1}^n (s_j^2 + 2is_j\mu_j - \mu_j^2) e^{S+iM} - \frac{1}{c^2} (s_0^2 + 2is_0\mu_0 - \mu_0^2) e^{S+iM}. \end{aligned} \quad (4.44)$$

To satisfy Eq. (4.33) we gather coefficients and set both the real and imaginary parts equal to zero. That is,

$$\sum_{j=1}^n s_j \mu_j - \frac{1}{c^2} s_0 \mu_0 = 0 \quad (4.45)$$

and

$$\sum_{j=1}^n (s_j^2 - \mu_j^2) - \frac{1}{c^2} (s_0^2 - \mu_0^2) = 0. \quad (4.46)$$

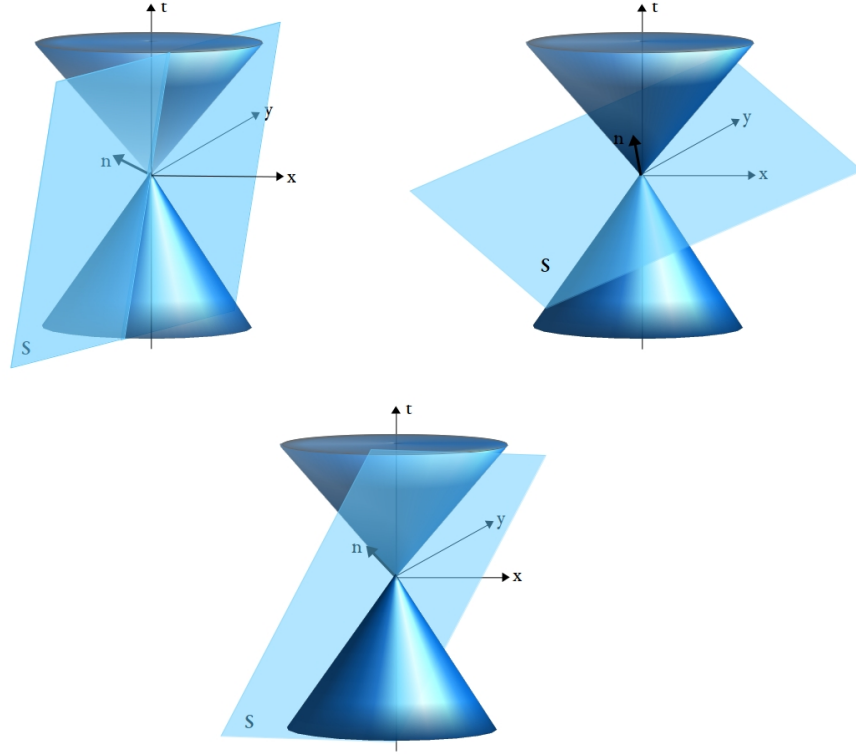


Figure 4.8: Initial surfaces S plotted as a function of x, y, t ; the cones represent the outward spread of the characteristics $\pm(x^2 + y^2)^{1/2} = ct$. (left) If the initial surface S cuts the double cone centered at (\mathbf{r}, t) it is called *timelike*. This is to say that the source exceeds the characteristic speed c as is the case for an object in supersonic or superluminal motion. (right) It is impossible to construct exponentially growing solutions to Eq. (4.33) if the initial plane lies entirely outside the spacetime cone. Such surfaces are termed *spacelike*. (lower) If equality holds in Eq. (4.53), any tangent to S also touches the characteristic cone. In this case, the object moves precisely at c .

If $s_0 = 0$, it is always feasible to find nontrivial values of s_j to generate exponential solutions. Hence, we consider the case $s_0 \neq 0$, solve Eq. (4.45) for μ_0 and substitute the result into Eq. (4.46) such that

$$\mu_0 = \frac{c^2}{s_0} \sum_{j=1}^n s_j \mu_j \quad (4.47)$$

and

$$\sum_{j=1}^n (s_j^2 - \mu_j^2) - \frac{s_0^2}{c^2} + \frac{c^2}{s_0^2} \left(\sum_{j=1}^n s_j \mu_j \right)^2 = 0. \quad (4.48)$$

We may now apply the Cauchy-Schwarz inequality

$$\left(\sum_{j=1}^n s_j \mu_j \right)^2 \leq \left(\sum_{j=1}^n s_j^2 \right) \left(\sum_{j=1}^n \mu_j^2 \right) \quad (4.49)$$

or

$$0 \leq \left(\sum_{j=1}^n s_j^2 \right) \left(\sum_{j=1}^n \mu_j^2 \right) - \left(\sum_{j=1}^n s_j \mu_j \right)^2 \quad (4.50)$$

to find that

$$\begin{aligned} 0 &\leq \sum_{j=1}^n (s_j^2 - \mu_j^2) - \frac{s_0^2}{c^2} + \frac{c^2}{s_0^2} \left(\sum_{j=1}^n s_j^2 \right) \left(\sum_{j=1}^n \mu_j^2 \right) \\ &= \left(\sum_{j=1}^n s_j^2 - \frac{s_0^2}{c^2} \right) \left(1 + \frac{c^2}{s_0^2} \sum_{j=1}^n \mu_j^2 \right). \end{aligned} \quad (4.51)$$

For the right-hand side to be nonnegative, the first factor containing a difference of squares must be greater than or equal to zero, *i.e.*,

$$\sum_{j=1}^n s_j^2 \geq \frac{s_0^2}{c^2}. \quad (4.52)$$

(As mentioned in passing, $s_0 = 0$ is a special case of this result.) In other words, it is possible to create Cauchy problems with solutions that grow exponentially when inequality (4.52) is satisfied. Notice that the condition may be rewritten as

$$\sum_{j=1}^n \left(\frac{\partial}{\partial x_j} S \right)^2 \geq \frac{1}{c^2} \left(\frac{\partial}{\partial t} S \right)^2 \quad (4.53)$$

since

$$\sum_{j=1}^n \left(\frac{\partial}{\partial x_j} S \right) = \sum_{j=1}^n s_j \quad \text{and} \quad \frac{\partial}{\partial t} S = s_0. \quad (4.54)$$

In this form, s_0, \dots, s_n may be interpreted as the components of a vector $\hat{\mathbf{n}}$ normal to the hypersurface $S(\mathbf{r}, t)$ defined in Eq. (4.41). When equality holds, any tangent to

S touches, by necessity, also the characteristic cone centered at (\mathbf{r}, t) . This is easily seen by comparing Eq. (4.53) to Eq. (4.22) for $A = -1$, $B = 0$ and $C = 1/c^2$, *i.e.*, $\left(\frac{\partial}{\partial x} \zeta\right)^2 - \left(\frac{\partial}{\partial y} \zeta\right)^2 = 0$, along with a look at the lower panel of Fig. 4.8. In the case of strict inequality, the initial plane – termed *timelike* by convention – cuts the cone such that its normal lies outside of the latter (Fig. 4.8 (left)). It is for these surfaces, then, that the Cauchy problem is ill-posed.

Consider, on the other hand, an initial surface for which

$$\sum_{j=1}^n \left(\frac{\partial}{\partial x_j} S\right)^2 < \frac{1}{c^2} \left(\frac{\partial}{\partial t} S\right)^2 \quad (4.55)$$

as shown in Fig. 4.8 (right). The spacetime surface $t = 0$, for instance, satisfies this condition. In such cases, it is impossible to construct exponentially growing solutions to the wave equation and the Cauchy problem is well conditioned. It seems natural to call surfaces on which Eq. (4.55) applies *spacelike* in analogy with the vernacular used above.

In this context, it is instructive to examine the homogeneous wave equation in two space variables given by

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \psi - \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right) \psi = 0 \quad \text{in } \mathbb{R}^2 \times (0, \infty). \quad (4.56)$$

According to Eq. (4.52), any spacelike element of S has to fulfill the requirement

$$\left(\frac{\partial}{\partial x_1} S\right)^2 + \left(\frac{\partial}{\partial x_2} S\right)^2 < \frac{1}{c^2} \left(\frac{\partial}{\partial t} S\right)^2. \quad (4.57)$$

This is to say that its normal has a t -component exceeding that of the characteristic cone for the problem, which is given by

$$\left(\frac{\partial}{\partial x_1} S\right)^2 + \left(\frac{\partial}{\partial x_2} S\right)^2 = \frac{1}{c^2} \left(\frac{\partial}{\partial t} S\right)^2. \quad (4.58)$$

In other words, Eq. (4.56) is well-posed only if the Cauchy data specify that the vector normal to S in time is greater than that in space (Fig. 4.9). Otherwise, it is improperly set as in Fig. 4.10.

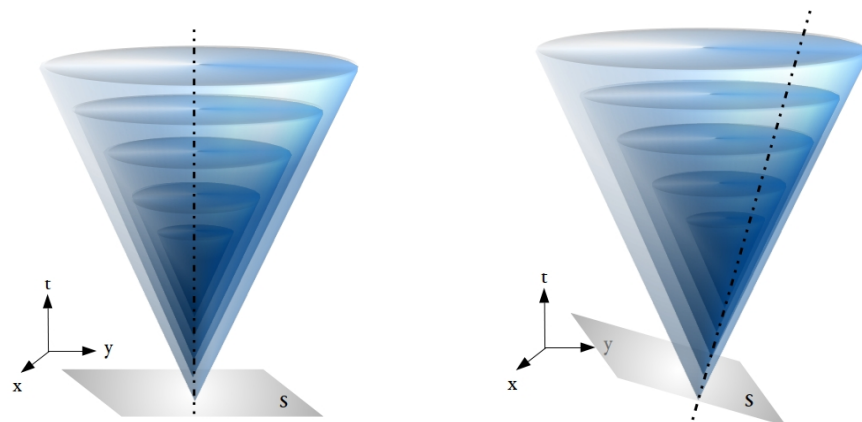


Figure 4.9: The influence of a point charge at $x = vt$, $y = \text{constant}$ in $\mathbb{R}^2 \times (0, \infty)$ (Part I). (left) $v = 0$: The source is stationary. (right) $v < c$: The velocity of the source remains below the characteristic wave speed. Notice that the initial cone contains all subsequent ones and, therefore, constitutes the range of influence for the *entire* problem. The gray shaded area denotes the initial hypersurface S and the broken line the vector $\hat{\mathbf{n}}$ normal to it. In this and the subsequent figure, the ring at the end of each cone represents the spread of waves from the source as time progresses.

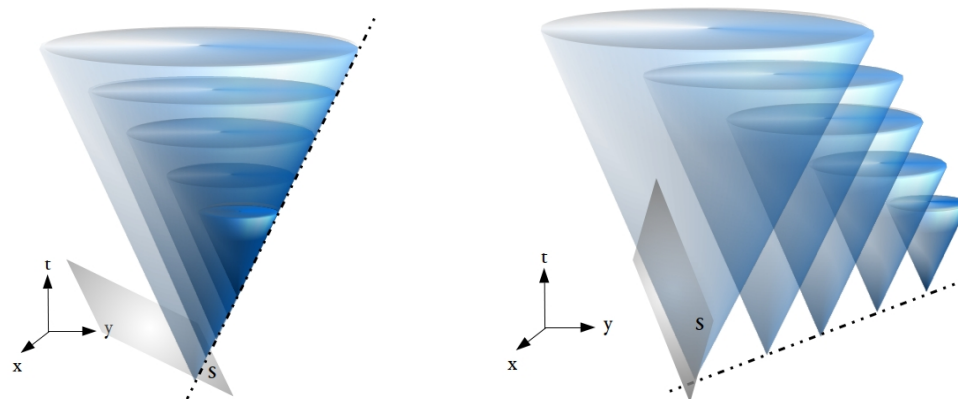


Figure 4.10: The influence of a point charge at $x = vt$, $y = \text{constant}$ in $\mathbb{R}^2 \times (0, \infty)$ (Part II). (left) $v = c$: The source travels precisely at the characteristic wave speed and the emitted wavelets meet tangentially at a single point. (right) $v > c$: The source has broken the sound (or light) barrier and travels at supersonic (or superluminal) speed. As a consequence, the range of influence for the problem has become a function of emission time t . Again, the gray shaded area denotes the initial hypersurface S and the broken line the vector $\hat{\mathbf{n}}$ normal to it.

This explains – finally! – why the Cauchy problem for the wave equation is properly posed only for sources whose tempo remains below the characteristic speed *in medio*. Let us suppose that we seek a solution for Eq. (4.1) where q is a moving point-like element such as an airplane in flight. This problem is adequately modeled by the homogeneous equation in three spatial dimensions (Eq. (4.33)); however, now we impose nonzero Cauchy data on some curve or surface $S \subset \mathbb{R}^3 \times (0, \infty)$. If the source's speed remains below that of the excited waves, *i.e.*, $v < c$, the spherical hypersurface created at some initial point \mathbf{r}_0 at time t_0 contains all subsequent ones and, therefore, constitutes the range of influence for the *entire* problem; this case is shown in a two-dimensional representation in Fig. 4.9 and as a series of three-dimensional Huygens' wavelets in Fig. 4.11. S is then spacelike and the problem is well-posed.

Once $v > c$ (Figs. 4.10 (right) and 4.12 (right)), the Huygens wavelets intersect one another rather than remaining nested in the fashion of a Russian doll. Consequently, the range of influence for the problem becomes a function of excitation time t ; S is timelike and the problem is ill-posed.

The transitional case, $v = c$ (Fig. 4.10 (left)), is highly singular since all of the emitted spheres meet tangentially at the same point (Fig. 4.12 (left)). We will have cause to return to this case in Chapter 6.



Figure 4.11: Huygens wavelets for the problem in three spatial dimensions (Part I). (left) The source is stationary. (right) The speed of the charge remains below c . Both cases are well-posed in the sense of Hadamard.

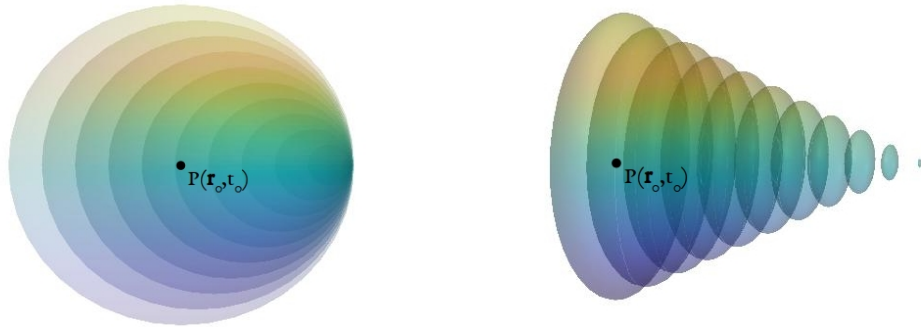


Figure 4.12: Huygens wavelets for the problem in three spatial dimensions (Part II). (left) The charge travels at the characteristic speed. (right) The source has broken the sound (or light) barrier and moves faster than the waves *in medio*. Both cases are ill-posed in the sense of Hadamard.

4.3 Mr. Green's Solution to Cauchy's Problem

The aim of Sections 4.1.2 and 4.2 was to show that the Cauchy (or initial-value) problem for the wave equation has a unique solution for any given source distribution q provided that the functions which appear in either the equation itself or the Cauchy data are sufficiently smooth. As is the case for many other phenomena in science, technology and nature, however, the application of Eq. (4.1) in a physical domain requires not only knowledge of the initial state of the system, but that of its *boundaries* as well. Models of mathematical physics that require both initial and boundary conditions such as, for instance, the description of heat transfer through a rod or electromagnetic waves traveling along a transmission line are known as *initial boundary value problems*, sometimes abbreviated IBVP.

4.3.1 A Necessary Word on Notation (Act II)

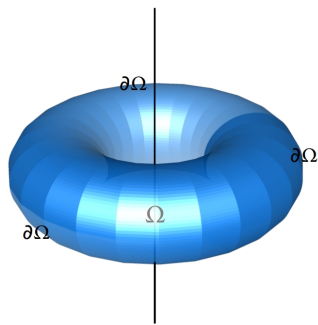


Figure 4.13: A toroidal hypervolume Ω enclosed by the hypersurface $\partial\Omega$. Hyperboundary conditions give the value of function ψ and possibly some of its space and time derivatives on $\partial\Omega$.

The boundary $\partial\Omega$ of a specified domain Ω in three-dimensional Euclidean space is a sheet – called *hypersurface* by convention [131, 125] – that confines Ω in space and time (Fig. 4.13). Intuitively, $\partial\Omega$ may be viewed as the region which can be

approached both from within and without Ω ; mathematically more rigorously, it is defined as the closed, piecewise smooth set of points in the closure, but not belonging to the interior, of Ω (e.g., $\bar{\Omega} = \Omega \cup \partial\Omega$, where \cup denotes the union of the two sets and the overline stands for closure). The bounding surface may be only piecewise connected and part (or all) of it recede to infinity; a familiar example, which constitutes the customary geometry in diffraction [53], involves two concentric spheres – the outer of “infinite” radius (we will comment on the quotation marks, originally introduced by [53], in Chapter 5) – between which lies Ω as shown in Fig. 4.14. The constituent vector elements of $\partial\Omega$ point along the outward normal $\hat{\mathbf{n}}$ – away from the center on the larger sphere and toward it on the smaller one.

The notational difficulty that arises in the derivation of electromagnetic fields far away from the charges and currents where they originated lies in distinguishing retarded quantities from those that are received (or one might say “measured” or “detected”) at a distant observer’s location. The contributions to the radiated fields will originate from the source and, depending on the conditions imposed, perhaps also from the boundary. Breaking with tradition²¹, we reserve the subscript P to denote coordinates and times where we seek the solution ψ – the *field* points – whereas \mathbf{r} and t – the *source* and *boundary* points – define *loci* where data are prescribed. In a nutshell, light emitted at time t and place \mathbf{r} reaches the field point \mathbf{r}_P at t_P . The same notational convention is used for derivatives, where ∇ , $\partial/\partial t$, and $\partial^2/\partial t^2$ are retarded quantities and, in general, $\nabla \neq \nabla_P$, $\partial/\partial t \neq \partial/\partial t_P$, and $\partial^2/\partial t^2 \neq \partial^2/\partial t_P^2$.

Although both source and boundary points contribute to the radiation field $\psi(\mathbf{r}_P, t_P)$ and are, therefore, uniformly labeled (\mathbf{r}, t) , we wish on occasion to distinguish between the two. If this is the case, locations on the boundary will be termed “ \mathbf{r} on $\partial\Omega$ ” or, simply, $\mathbf{r}_{\partial\Omega}$. (Time t , on the other hand, is defined to be the

²¹Most texts mark *retarded* quantities with sub- or superscripts, e.g., (\mathbf{r}', t') [53, 125, 149, 163] or (\mathbf{r}_0, t_0) [129, 164]; some use different variables altogether (e.g., (ξ, τ) in [130, 162, 164]).

same for $Q \in q$ and $Q \in \partial\Omega$ as will be shown in detail in Section 4.3.2 and associated Fig. 4.19.)

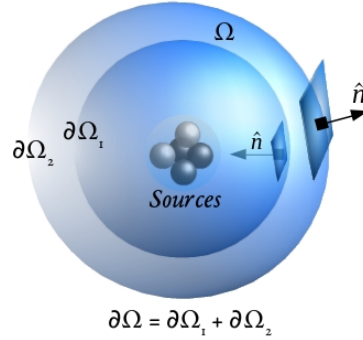


Figure 4.14: A hypersurface that consists of two concentric spheres enclosing the volume Ω . The radiation sources are contained within the smaller sphere.

The task of finding physically relevant solutions to the general²² IBVP (Eq. (4.1)) calls for Dirichlet or Neumann *boundary conditions* which prescribe either the value of ψ or that of its normal gradient $\nabla\psi$ – but not both! – for all $t_{P_0} \leq t \leq t_{P_1}$ everywhere on $\partial\Omega$ (*i.e.*, the vertical edges of Ω in Fig. 4.15). This reflects the reality that the system’s physical domain is restricted by a boundary whose properties, whatever their nature, must be taken into account by the model. As stressed throughout Sections 4.1.2 and 4.2, Cauchy *initial conditions* are to be imposed along the figure’s lower horizontal, which is to say that both displacement ψ and velocity $\partial\psi/\partial t$ must be specified at some initial time t_{P_0} . Since the wave equation is invariant under time-reversal (a proof is to follow shortly) we could equally well prescribe the terminal conditions $\psi|_{t=t_{P_1}}$ and $\partial\psi/\partial t|_{t=t_{P_1}}$ instead. Although this choice is physically of lesser interest since “the future is rarely known” [125] (and the past almost always is) it will

²²We take unbounded space to be a special, limiting case of a bounded region whose surface recedes to infinity as discussed, for example, in [125] or [129].

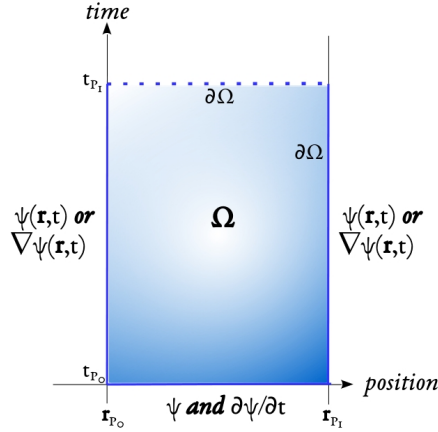


Figure 4.15: The problem of solving the wave equation is well-posed if either Dirichlet or Neumann boundary conditions (but not both) are prescribed along the vertical edges of Ω . Along the lower horizontal, Cauchy initial conditions are wanted. *No restrictions may be placed along the top horizontal*, as this would result in an overdetermined system.

be explored in brief in the chapters to come. No restrictions may be placed along the top horizontal of Fig. 4.15: Any electromagnetic disturbance travels radially outward at precisely the wave speed, c , once it has been emitted at time t and place \mathbf{r} ; this determines uniquely that its arrival time at field point \mathbf{r}_P must be $t_P = t + |\mathbf{r}_P - \mathbf{r}|/c$. (The surprising consequences of this seemingly simple relation will be covered in Chapter 6.) As shown by Hadamard [127], John [142] and many others, attempting to calculate Eq. (4.1) from given values on a *closed*²³ hypersurface $\partial\Omega$ (or, in other words, to choose Dirichlet or Neumann instead of Cauchy initial conditions) will result in either an overdetermined and, in John’s parlance, “unnatural” system or “ineffective boundaries” [125] – restrictions that have no effect on the model and are, therefore, spurious.

²³A hypervolume Ω that is fully bounded in both, space and time, through boundary and initial conditions (or “hyperboundary conditions” for short) is called closed, otherwise it is considered to be open. Depending on the hyperboundary conditions for the particular problem, $\partial\Omega$ can either be open or closed (but, obviously, not both). This differs noticeably from conventions in set theory and topology (used above to characterize $\partial\Omega$) where a closed set is defined as one that contains all of its limit points (which, of course, is always true for $\partial\Omega$).

4.3.2 The Fundamental Properties of the Green's Function

To find the fundamental causal solution for Eq. (4.1) including the satisfaction of all hyperboundary conditions outlined in Fig. 4.15, we consider the wave equation for ψ and the Green's function G or, to be precise, its reciprocal. Proposed in 1828 by George Green in his seminal *Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism* [165], the method that bears his name is as simple as it is elegant: To obtain the field caused by a volume-distributed source q , we calculate the effect of each elementary (or point-like) portion of the source and, by virtue of the superposition principle, sum the solutions. If $G(\mathbf{r}_P|\mathbf{r})$ is the impulse generated by a unit point charge located at \mathbf{r} , then the field caused by the extended source $q(\mathbf{r})$ (as sampled by an observer at \mathbf{r}_P) is the integral of $q(\mathbf{r})G(\mathbf{r}_P|\mathbf{r})$ taken over the space occupied by q [129].



Figure 4.16: George Green, whose formal education lasted, remarkably, only 14 months, worked most of his life in the windmill that his father built in 1807 in Nottingham, UK. Credit: <https://www.greensmill.org.uk/>

In mathematical terms, G – considered to be a function of *field points* [125] and written $G(\mathbf{r}_P, t_P|\mathbf{r}, t)$ if the problem has time dependence – solves Eq. (4.1) whilst satisfying the homogeneous variant of the boundary and initial conditions imposed

on ψ . It is defined as

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t_P^2} - \nabla_P^2\right) G(\mathbf{r}_P, t_P | \mathbf{r}, t) = 4\pi \delta(\mathbf{r}_P - \mathbf{r}) \delta(t_P - t) \quad (4.59)$$

where

$$G(\mathbf{r}_P, t_P | \mathbf{r}, t) = 0 \quad \text{or} \quad \nabla G(\mathbf{r}_P, t_P | \mathbf{r}, t) = 0 \quad \forall \mathbf{r} \in \partial\Omega \quad (4.60a)$$

$$G(\mathbf{r}_P, t_P | \mathbf{r}, t) = 0 \quad t_P < t. \quad (4.60b)$$

$\delta(x)$ is the Dirac delta function that models the density of an idealized point charge as a function equal to zero if $x \neq 0$ and whose integral over the entire real line is unity. Note that the first expression in Eq. (4.60a) represents Dirichlet boundary conditions, the second one Neumann. Eq. (4.60b) arises since it seems reasonable to impose causality as an initial condition, which is to say that G and its time derivative should be zero for $t_P < t$. If an impulse occurs at t , no effects can be present anywhere at an earlier time.

Just as a supersonic airplane leaves a cone of intersecting pressure waves in its wake (consult Section 2.3 and the associated figures for a brief review), the Huygens wavelets emitted by a faster-than-light source will superpose. Take, for instance, the (purely hypothetical) point charge $q(\mathbf{r}, t)$ in constant superluminal rotation which is the basis for most calculations in Chapter 6. Since q ‘outruns’ the waves it generates (which, as all others, must move *at c*) contributions from multiple (\mathbf{r}, t) pairs on its trajectory may arrive simultaneously at an observer or receiver. The inverse, however, is *not* true; q will emit once – and only once! – at any given point in spacetime. Simply put, several wavefronts may intersect at $P(\mathbf{r}_P, t_P)$, but only one can originate from $Q(\mathbf{r}, t)$. Whether the latter contributes to the radiation measured at any given observation point is determined by the sifting property of the Green’s function.

This raises a number of concerns, not least whether the field $\psi(\mathbf{r}_P, t_P)$ radiated by a volume-distributed superluminal source is unique or, more immediately, if the

basic properties of $G(\mathbf{r}_P, t_P | \mathbf{r}, t)$ still hold. Since G is the response to an elementary impulse and therefore a function of \mathbf{r}_P and t_P (which is precisely where multiple wave fronts may meet) this question is not trivial.

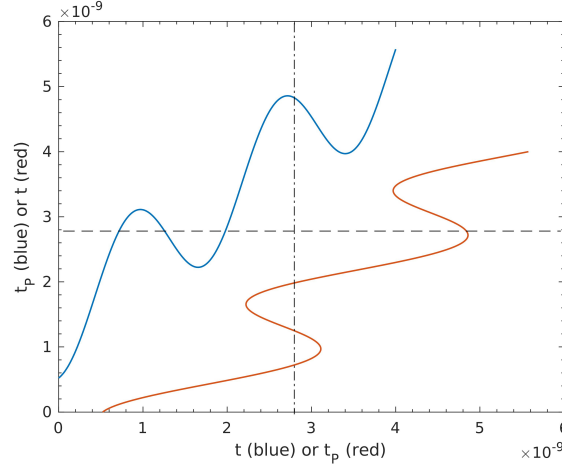


Figure 4.17: Function $h(t) = t + R/c = t_P$ (blue) and its inverse $h^{-1}(t_P) = t_P - R/c = t$ (red). (The function $h(t)$ describes the relationship between reception and retarded time for a point-like source in superluminal rotation; see Chapter 6.) Multiple instances of source time $t_1, t_2, t_3, \dots, t_m$ may arrive concurrently at field point \mathbf{r}_P and time t_P . However, each individual element t_k in t contributes *precisely once* to the range t_P .

As Fig. 4.17 illustrates, the formula that links emission and observation time for the point charge q , *i.e.*,

$$h(t) = t + \frac{R}{c} = t_P \quad (4.61)$$

(blue graph) is a function. (As before, $R = |\mathbf{R}| \equiv |\mathbf{r}_P - \mathbf{r}|$.) Yet, since h is not injective it is not invertible (red graph) and any global uniqueness proof involving $h^{-1}(t_P)$ is doomed to fail. Even if a function is not one-to-one, however, it is often possible to define a *partial inverse* by restricting its domain. If we divide h into m branches, one between each pair of local extrema, none contains more than a single root (Fig. 4.18) and h^{-1} is well defined. It is in this sense that the proofs verifying the fundamental properties of the Green's function should be understood.

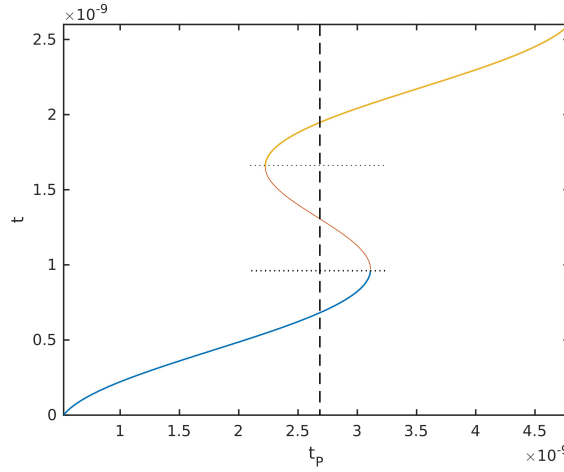


Figure 4.18: The three branches of function $h(t) = t - R/c = t_P$ depicted in blue, red and yellow, respectively. Notice that every branch is one-to-one and can, therefore, be inverted.

We remark in passing and without formal verification that the collection of space-time pairs $(\mathbf{r}_1, t_1), (\mathbf{r}_2, t_2), \dots (\mathbf{r}_m, t_m)$ that map concurrently onto (\mathbf{r}_P, t_P) (*i.e.*, the solution set to $h^{-1}(t_P) = t_P - R/c = t$) is unique owing to causality. Moreover, if an extended source distribution is taken to be a bundle of densely packed rigidly rotating point charges the uniqueness of $\psi(\mathbf{r}_P, t_P)$ can easily be shown by induction.

The basic properties of the Green's function divide roughly into those that affect location $(\mathbf{r}, \mathbf{r}_P)$ and those that concern time (t, t_P) . We will set out with the latter.

Invariance Under Time-Translation

Provided that $\partial\Omega$ is stationary (that is, Ω does not change with time) the Green's function is *invariant under time-translation* and G depends on t_P and t only through the combination $t_P - t$. To appreciate the proof to follow it is crucial to recognize that time t and, in consequence, the time span $t_P - t$, is defined to be *the same* for a point on (or within) the source q and one on the boundary $\partial\Omega$. (See, for instance, the discussion of “Boundaries in the Finite Region” on p. 849 of [129].) As shown in

Fig. 4.19 and discussed in Section 4.3.1, the same is, of course, not true for *position*.

Proof: We subtract t from both, t_P and t such that, without loss of generality, $t = 0$ and $t_P - t = t_P \equiv \tau$. This is equivalent to performing the change of variables $t_P \rightarrow \tau$ in definition (4.59) which transforms $\partial^2/\partial t_P^2$ into $\partial^2/\partial \tau^2$ – the only adjustment of the D'Alembert operator. The equation then becomes

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial \tau^2} - \nabla_P^2\right) G(\mathbf{r}_P, \mathbf{r}; \tau) = 4\pi \delta(\mathbf{r}_P - \mathbf{r}) \delta(\tau), \quad (4.62)$$

where boundary condition (4.60a) remains unaffected since, by assumption, $\partial\Omega$ is fixed and the distance between boundary and observer does not change. Initial condition (4.60b) becomes $G(\mathbf{r}_P, \mathbf{r}; \tau) = 0 \ \forall \ \tau < 0$ once t is subtracted from both sides of the inequality.

Since neither Eq. (4.62) nor the associated conditions involve variables other than \mathbf{r}_P and \mathbf{r} in addition to the time span τ , the solution G must likewise be a function of these unknowns alone. In contrast, as long as the boundary is at a *finite* distance from the observer, G depends on \mathbf{r} and \mathbf{r}_P individually since, as argued above and in Section 4.3.1, $|\mathbf{r}_P - \mathbf{r}|$ is not necessarily equal to $|\mathbf{r}_P - \mathbf{r}_{\partial\Omega}|$. (Reconsider Fig. 4.19 as a visual aid.) ■

The same conclusion can be reached by re-examining the relationship between source and observation time, introduced as Eq. (4.61) above. A simple rearrangement of terms yields $t_P - t = R/c$ or $\tau = R/c$ which reaffirms that t_P and t occur only in tandem in Eq. (4.59)²⁴. Consequentially, G is invariant under the transformation $G(\mathbf{r}_P, t_P | \mathbf{r}, t) \rightarrow G(\mathbf{r}_P, t_P + a | \mathbf{r}, t + a)$ which translates all time variables by an arbitrary amount a . This implies that τ remains unaffected by the exchanges $t_P \rightarrow -t, t \rightarrow -t_P$ and we introduce *the reciprocity relation*

$$G(\mathbf{r}_P, t_P | \mathbf{r}, t) = G(\mathbf{r}_P, -t | \mathbf{r}, -t_P) \quad (4.63)$$

²⁴It is true that t and t_P may occur individually in the formula that calculates $R(t)$. (See, e.g., Eq. (6.3) in Chapter 6.) However, in the present context we are interested only in the radiation emitted by the point-like source element; how it arrived at position (\mathbf{r}, t) is of no import.

as a corollary.

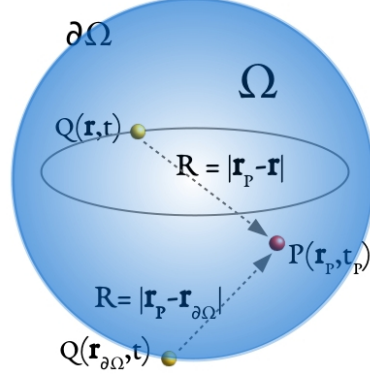


Figure 4.19: Geometry of the Green's function problem in finite space, where $\partial\Omega$ denotes the surface over which the boundary conditions are imposed and Ω the enclosed volume. Whereas time t is the same for source point $Q(\mathbf{r}, t)$ and boundary point $Q(\mathbf{r}_{\partial\Omega}, t)$, generally $|\mathbf{r}_P - \mathbf{r}| \neq |\mathbf{r}_P - \mathbf{r}_{\partial\Omega}|$.

Invariance Under Time-Reversal

In addition, the Green's function that solves Eq. (4.59) is 'reversible' or, in mathematically more proper terms, *invariant under time-reversal*. This is best seen by examining how the homogeneous form of Eq. (4.62) evolves under time-independent boundary conditions.

Proof: Let $G(\mathbf{r}_P, \mathbf{r}; \tau)$ satisfy

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial \tau^2} - \nabla_P^2 \right) G(\mathbf{r}_P, \mathbf{r}; \tau) = 0, \quad (4.64)$$

and define a different, but related, function

$$G_T(\mathbf{r}_P, \mathbf{r}; \tau) = G(\mathbf{r}_P, \mathbf{r}; -\tau), \quad (4.65)$$

where the subscript T indicates time reversal. Exchanging variables such that $\tau \rightarrow -\tau$, $\partial/\partial\tau \rightarrow \partial/\partial(-\tau)$, and $\partial^2/\partial\tau^2 \rightarrow \partial^2/\partial(-\tau)^2$ leads to

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial (-\tau)^2} - \nabla_P^2 \right) G(\mathbf{r}_P, \mathbf{r}; -\tau) = \left(\frac{1}{c^2} \frac{\partial^2}{\partial \tau^2} - \nabla_P^2 \right) G_T(\mathbf{r}_P, \mathbf{r}, \tau) = 0, \quad (4.66)$$

which shows that $G(\mathbf{r}_P, \mathbf{r}; \tau)$ and $G_T(\mathbf{r}_P, \mathbf{r}; \tau)$ *both* satisfy Eq. (4.64). ■

Invariance under time-reversal implies that

$$G(\mathbf{r}_P, \mathbf{r}; \tau) = -G(\mathbf{r}_P, \mathbf{r}; -\tau) \quad (4.67)$$

or, in terms of t_P and t

$$G(\mathbf{r}_P, t_P | \mathbf{r}, t) = -G(\mathbf{r}_P, t | \mathbf{r}, t_P) \quad (4.68)$$

since, by reciprocity relation (4.63), $-t_P \rightarrow t$ and $-t \rightarrow t_P$ ²⁵.

Positional Symmetry

Empirical reasoning suggests that G is *symmetric with respect to location* (i.e., \mathbf{r} , $\mathbf{r}_{\partial\Omega}$, and \mathbf{r}_P). The proof confirming this, however, is lengthy and mirrors the methods used in Eqs. 4.73–4.86 which have yet to be introduced. Hence we posit without formal verification that

$$G(\mathbf{r}_P, t_P | \mathbf{r}, t) = G(\mathbf{r}, t_P | \mathbf{r}_P, t). \quad (4.69)$$

The skeptical reader is encouraged to study the step-by-step *ansatz* offered by Barton on page 239 of [125].

4.3.3 Apropos the Wave Equation I: General Case

The elusive reciprocal (or adjoint) equation for G can now be found by performing the transformations (4.69) and (4.63) in sequence. That is,

²⁵Notice that invariance under time-reversal is neither true nor self-evident for all partial differential equations with time-dependence. While prediction is always possible, *retrodiction* does, for instance, not hold for the diffusion equation where it is usually impossible to deduce the initial from the final state of the system uniquely.

$$\begin{aligned}
\left(\frac{1}{c^2} \frac{\partial^2}{\partial t_P^2} - \nabla_P^2\right) G(\mathbf{r}_P, t_P | \mathbf{r}, t) &= \left(\frac{1}{c^2} \frac{\partial^2}{\partial t_P^2} - \nabla_P^2\right) G(\mathbf{r}, t_P | \mathbf{r}_P, t) \\
&= \left(\frac{1}{c^2} \frac{\partial^2}{\partial t_P^2} - \nabla_P^2\right) G(\mathbf{r}, -t | \mathbf{r}_P, -t_P) \\
&= 4\pi \delta(\mathbf{r}_P - \mathbf{r}) \delta(t_P - t).
\end{aligned} \tag{4.70}$$

To convert the second equality to standard form, we change variables such that $\mathbf{r}_P \rightarrow \mathbf{r}$ and $t_P \rightarrow -t$ (which implies $\partial^2/\partial t_P^2 \rightarrow \partial^2/\partial(-t)^2 = \partial^2/\partial t^2$ and $\nabla_P^2 \rightarrow \nabla^2$). This yields

$$\begin{aligned}
\left(\frac{1}{c^2} \frac{\partial^2}{\partial t_P^2} - \nabla_P^2\right) G(\mathbf{r}, -t | \mathbf{r}_P, -t_P) &= \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2\right) G(\mathbf{r}_P, t_P | \mathbf{r}, t) \\
&= 4\pi \delta(\mathbf{r} - \mathbf{r}_P) \delta(-t + t_P) \\
&= 4\pi \delta(\mathbf{r}_P - \mathbf{r}) \delta(t_P - t)
\end{aligned} \tag{4.71}$$

or

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2\right) G(\mathbf{r}_P, t_P | \mathbf{r}, t) = 4\pi \delta(\mathbf{r}_P - \mathbf{r}) \delta(t_P - t), \tag{4.72}$$

where the conditions are the same as those derived for subscripted quantities in Eqs. 4.60a–4.60b. The last step in Eq. (4.71) makes use of the fact that $\delta(x)$ is an even function of its argument (*i.e.*, $f(x) = f(-x)$) under the weak definition, which suffices here. (See [125], p. 11, for a discussion.) Although Eqs. (4.72) and (4.59) look deceptively similar, the d'Alembert and gradient operators as well as the arguments of the Green's function are now *retarded quantities*, as desired. (The importance of this point will become clear in Chapter 5.)

Multiplying Eq. (4.1) by G , Eq. (4.72) by ψ , and subtracting the latter from the former results in

$$\begin{aligned}
& \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \psi(\mathbf{r}, t) G(\mathbf{r}_P, t_P | \mathbf{r}, t) \\
& - \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) G(\mathbf{r}_P, t_P | \mathbf{r}, t) \psi(\mathbf{r}, t) \\
& = 4\pi G(\mathbf{r}_P, t_P | \mathbf{r}, t) q(\mathbf{r}, t) - 4\pi \psi(\mathbf{r}, t) \delta(\mathbf{r}_P - \mathbf{r}) \delta(t_P - t)
\end{aligned} \tag{4.73}$$

or, using the product rule,

$$\begin{aligned}
& \frac{1}{c^2} \frac{\partial}{\partial t} \left[\frac{\partial \psi}{\partial t} G - \frac{\partial G}{\partial t} \psi \right] - \left(G \nabla^2 \psi - \psi [\nabla^2 G] \right) \\
& = 4\pi G q - 4\pi \psi \delta(\mathbf{r}_P - \mathbf{r}) \delta(t_P - t).
\end{aligned} \tag{4.74}$$

Unless noted otherwise, $G \equiv G(\mathbf{r}_P, t_P | \mathbf{r}, t)$, $\psi \equiv \psi(\mathbf{r}, t)$ and $q \equiv q(\mathbf{r}, t)$ throughout. We may now integrate over the volume of interest, Ω , and with respect to t from t_{P_0} to $t_P +$ ²⁶. This yields

$$\begin{aligned}
& \int_{t_{P_0}}^{t_P +} dt \int_{\Omega} d^3 \mathbf{r} \left\{ \frac{1}{c^2} \frac{\partial}{\partial t} \left[\frac{\partial \psi}{\partial t} G - \frac{\partial G}{\partial t} \psi \right] - \left(G \nabla^2 \psi - \psi [\nabla^2 G] \right) \right\} \\
& = 4\pi \int_{t_{P_0}}^{t_P +} dt \int_{\Omega} d^3 \mathbf{r} \left[G q - \psi \delta(\mathbf{r}_P - \mathbf{r}) \delta(t_P - t) \right] \\
& = \left(4\pi \int_{t_{P_0}}^{t_P +} dt \int_{\Omega} d^3 \mathbf{r} G q \right) - 4\pi \psi(\mathbf{r}_P, t_P),
\end{aligned} \tag{4.75}$$

where the simplification of the very last term on the right-hand side results from the definition of the integral of the delta function. The first term on the left-hand side can be integrated formally with respect to t such that

$$\frac{1}{c^2} \int_{t_{P_0}}^{t_P +} dt \int_{\Omega} d^3 \mathbf{r} \frac{\partial}{\partial t} \left[\frac{\partial \psi}{\partial t} G - \frac{\partial G}{\partial t} \psi \right] = \frac{1}{c^2} \int_{\Omega} d^3 \mathbf{r} \left[\frac{\partial \psi}{\partial t} G - \frac{\partial G}{\partial t} \psi \right]_{t=t_{P_0}}^{t=t_P +}. \tag{4.76}$$

Note that, since $t = t_P +$ implies $t > t_P$, the upper-limit contribution vanishes by the initial condition (4.60b) imposed on G , which leaves

$$-\frac{1}{c^2} \int_{\Omega} d^3 \mathbf{r} \left[\frac{\partial \psi(\mathbf{r}, t_{P_0})}{\partial t} G(\mathbf{r}_P, t_P | \mathbf{r}, t_{P_0}) - \frac{\partial G(\mathbf{r}_P, t_P | \mathbf{r}, t_{P_0})}{\partial t} \psi(\mathbf{r}, t_{P_0}) \right]. \tag{4.77}$$

²⁶To avoid ending the integral precisely at the peak of the delta function, $t_P + \equiv t_P + \epsilon$ with the limit $\epsilon \rightarrow 0+$ to be taken at the end of the calculation.

The crucial – and most elegant – move in the derivation of a solution to Eq. (4.1), however, involves the second volume integral in Eq. (4.75), *i.e.*,

$$\int_{\Omega} d^3\mathbf{r} G \nabla^2 \psi - \psi [\nabla^2 G], \quad (4.78)$$

which we turn into a surface integral over the boundary $\partial\Omega$ by utilizing Green's theorem. Also called “Green's second identity”, the latter is a consequence of Gauss's theorem and the formula $\nabla \cdot (a \nabla b) = a \nabla^2 b + \nabla a \cdot \nabla b$. It can be given as

$$\int_{\Omega} d^3\mathbf{r} \{a \nabla^2 b - b \nabla^2 a\} = \int_{\partial\Omega} d^2\mathbf{r} \hat{\mathbf{n}} \cdot \{a \nabla b - [\nabla a] b\}, \quad (4.79)$$

where $\hat{\mathbf{n}}$ indicates the outward normal to $\partial\Omega$, and its application reduces the second term in Eq. (4.75) to

$$- \int_{t_{P_0}}^{t_P^+} dt \int_{\partial\Omega} d^2\mathbf{r} \hat{\mathbf{n}} \cdot \left[G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t) \nabla \psi(\mathbf{r}_{\partial\Omega}, t) \right] \quad (4.80)$$

$$- [\nabla G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t)] \psi(\mathbf{r}_{\partial\Omega}, t) \Big]. \quad (4.81)$$

Combining the findings from Eq. (4.73) through Eq. (4.80) results in

$$\begin{aligned} & - \frac{1}{c^2} \int_{\Omega} d^3\mathbf{r} \left[\frac{\partial \psi(\mathbf{r}, t_{P_0})}{\partial t} G(\mathbf{r}_P, t_P | \mathbf{r}, t_{P_0}) - \frac{\partial G(\mathbf{r}_P, t_P | \mathbf{r}, t_{P_0})}{\partial t} \psi(\mathbf{r}, t_{P_0}) \right] \\ & - \int_{t_{P_0}}^{t_P^+} dt \int_{\partial\Omega} d^2\mathbf{r} \hat{\mathbf{n}} \cdot \left[G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t) \nabla \psi(\mathbf{r}_{\partial\Omega}, t) \right. \\ & \quad \left. - [\nabla G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t)] \psi(\mathbf{r}_{\partial\Omega}, t) \right] \\ & = \left(4\pi \int_{t_{P_0}}^{t_P^+} dt \int_{\Omega} d^3\mathbf{r} G(\mathbf{r}_P, t_P | \mathbf{r}, t) q(\mathbf{r}, t) \right) - 4\pi \psi(\mathbf{r}_P, t_P) \end{aligned} \quad (4.82)$$

and, after some gathering and rearranging of terms, we arrive at the full causal fundamental (or Green's function) solution for the scalar wave equation including the satisfaction of all boundary and initial conditions on $\psi(\mathbf{r}, t)$ and $\psi(\mathbf{r}_{\partial\Omega}, t)$,

$$\psi(\mathbf{r}_P, t_P) = u(\mathbf{r}_P, t_P) + v(\mathbf{r}_P, t_P) + w(\mathbf{r}_P, t_P) \quad (4.83)$$

where

$$u(\mathbf{r}_P, t_P) = \int_{t_{P_0}}^{t_P^+} dt \int_{\Omega} d^3\mathbf{r} G(\mathbf{r}_P, t_P | \mathbf{r}, t) q(\mathbf{r}, t), \quad (4.84)$$

$$v(\mathbf{r}_P, t_P) = \frac{1}{4\pi} \int_{t_{P_0}}^{t_P^+} dt \int_{\partial\Omega} d^2\mathbf{r} \, \hat{\mathbf{n}} \cdot \left[G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t) \nabla \psi(\mathbf{r}_{\partial\Omega}, t) \right. \\ \left. - [\nabla G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t)] \psi(\mathbf{r}_{\partial\Omega}, t) \right], \quad (4.85)$$

and

$$w(\mathbf{r}_P, t_P) = -\frac{1}{4\pi c^2} \int_{\Omega} d^3\mathbf{r} \left[\frac{\partial G(\mathbf{r}_P, t_P | \mathbf{r}, t_{P_0})}{\partial t} \psi(\mathbf{r}, t_{P_0}) \right. \\ \left. - \frac{\partial \psi(\mathbf{r}, t_{P_0})}{\partial t} G(\mathbf{r}_P, t_P | \mathbf{r}, t_{P_0}) \right]. \quad (4.86)$$

The first integral, u , is the source term which captures radiation emitted by the time-varying source q as it travels through space and time. Therefore, it calculates the inhomogeneous equation $\square u = q(\mathbf{r}, t)$ with homogeneous hyperboundary conditions. The last, w , accounts for the propagation of initial conditions and solves the homogeneous problem $\square w = 0$ with homogeneous boundary and inhomogeneous initial conditions. Under the ‘null initial conditions’ $\psi|_{t=t_{P_0}} = \partial\psi/\partial t|_{t=t_{P_0}} = 0$, which may be assumed here without loss of generality, it is identically zero such that Eq. (4.83) and Eqs. (4.84)–(4.86) reduce to

$$\psi(\mathbf{r}_P, t_P) = u(\mathbf{r}_P, t_P) + v(\mathbf{r}_P, t_P) \quad (4.87)$$

$$= \int_{t_{P_0}}^{t_P^+} dt \int_{\Omega} d^3\mathbf{r} \, G(\mathbf{r}_P, t_P | \mathbf{r}, t) q(\mathbf{r}, t) \\ + \frac{1}{4\pi} \int_{t_{P_0}}^{t_P^+} dt \int_{\partial\Omega} d^2\mathbf{r} \, \hat{\mathbf{n}} \cdot \left[G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t) \nabla \psi(\mathbf{r}_{\partial\Omega}, t) \right. \\ \left. - [\nabla G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t)] \psi(\mathbf{r}_{\partial\Omega}, t) \right]. \quad (4.88)$$

The second of the remaining integrals, the boundary term v , represents contributions to $\psi(\mathbf{r}_P, t_P)$ due to the conditions imposed at the spacial boundary $\partial\Omega$; it resolves the homogeneous equation $\square v = 0$ while satisfying inhomogeneous boundary and homogeneous initial conditions. Depending on the restrictions placed on $\psi(\mathbf{r}_{\partial\Omega}, t)$ (*i.e.*, Dirichlet, Neumann, or mixed) *either G or its gradient is zero for all \mathbf{r} on the*

bounding surface, eliminating one of the two terms²⁷ in v . In consequence, Eq. (4.88) is an explicit construct of the data $q(\mathbf{r}, t)$, $\psi(\mathbf{r}_{\partial\Omega}, t)$ or $\nabla\psi(\mathbf{r}_{\partial\Omega}, t)$, $\psi(\mathbf{r}_P, t_{P_0})$, and $\partial\psi(\mathbf{r}_P, t_{P_0})/\partial t$, not, as claimed in [168], [169] and elsewhere, “merely a mathematical identity” that cannot “be used to calculate the field arising from a given source distribution” – at least if boundary and initial conditions that lead to a well-posed problem are chosen. In fact, this constitutes the intrinsic beauty of Cauchy’s problem: *Boundary values that are not prescribed are not needed.*

4.3.4 Apropos the Wave Equation II: Free Space

In electrical engineering, the general Green’s function solution to the scalar wave equation along with Cauchy data and Dirichlet or Neumann boundary conditions is used to calculate the spread of electromagnetic radiation throughout systems of *finite size* such as resonant cavities, conducting cylinders, and waveguides (see *e.g.*, Ch. 15 of [54]). We, however, aim to describe the fields emitted by a collection of rotating unit point sources into an *infinite domain* which obeys causality and is everywhere quiescent before some initial time $t = 0$. To find the solution of Eq. (4.1) when the bounding surface $\partial\Omega$ recedes to infinity, we proceed as above but use

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t_P^2} - \nabla_P^2\right) G_{\circ}(\mathbf{r}_P, t_P | \mathbf{r}, t) = 4\pi \delta(\mathbf{r}_P - \mathbf{r}) \delta(t_P - t) \quad (4.89)$$

instead of Eq. (4.59), where the subscript \circ indicates quantities in free space and any space integral now extends over all \mathbf{r}_P .

As will be shown in the section to follow, the explicit form of the Green’s function for Eq. (4.89) consists of a delta function scaled by the distance R . $\delta(x)$ is neither a ‘proper’ function nor meaningful in any context other than its multiplication by a

²⁷Recall that, at any one point \mathbf{r} on $\partial\Omega$, either $\psi(\mathbf{r}_{\partial\Omega}, t)$ or $\nabla\psi(\mathbf{r}_{\partial\Omega}, t)$ is given for all $t_{P_0} \leq t \leq t_{P+}$. If ψ is known then, from definition (4.60a) and the discussion following Eq. (4.72), G is zero there and the value of $\nabla\psi$ does not contribute since, in Eq. (4.88), it is multiplied by G . A similar argument holds if $\nabla\psi$ is given.

sufficiently well-behaved test function and subsequent integration over some finitely wide range of x . It is nevertheless possible to predict its behavior when x tends to infinity. Lighthill²⁸, for instance, states in [166] that the sequences equivalent to $e^{-nx^2}(n/\pi)^{1/2}$, $n = 1, 2, 3, \dots$, define a generalized function $\delta(x)$ such that

$$\int_{-\infty}^{\infty} \delta(x) F(x) dx = F(0). \quad (4.90)$$

(A proof can be found on p. 17 of [166].) Due to the negative exponent, this series tends to zero irrespective of the value of n as $x \rightarrow \pm\infty$. As a result conditions 4.60a and 4.60b become quite naturally

$$G_o(\mathbf{r}_P, t_P | \mathbf{r}, t) \rightarrow 0 \text{ and } \nabla G_o(\mathbf{r}_P, t_P | \mathbf{r}, t) \rightarrow 0 \quad \forall \mathbf{r} \in \partial\Omega \mid \partial\Omega \rightarrow \infty \quad (4.91a)$$

$$G_o(\mathbf{r}_P, t_P | \mathbf{r}, t) = 0 \quad t_P < t. \quad (4.91b)$$

Unlike G , G_o depends on \mathbf{r}_P and \mathbf{r} *only* through the combination $R = |\mathbf{R}| \equiv |\mathbf{r}_P - \mathbf{r}|$. This implies that in unbounded space the distance covered by the wave front alone matters; its direction is of no consequence for the calculation of Eq. (4.89).

Proof: The reasoning follows in essence that used to show invariance under time-translation. A change of independent variables from (\mathbf{r}_P, t_P) to (R, τ) , which entails the substitutions $\partial^2/\partial t_P^2 \rightarrow \partial^2/\partial \tau^2$ and $\nabla_P \rightarrow \nabla_R$, transforms the D'Alembert operator for the problem such that

$$\square = \left(\frac{1}{c^2} \frac{\partial^2}{\partial \tau^2} - \nabla_R^2 \right). \quad (4.92)$$

As a result, Eq. (4.89) becomes

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial \tau^2} - \nabla_R^2 \right) G_o(R, \tau) = 4\pi \delta(R) \delta(\tau) \quad (4.93)$$

with hyperboundary conditions

$$G_o(R, \tau) \rightarrow 0 \quad \text{as } R \rightarrow \infty \quad (4.94a)$$

²⁸For an excellent survey of Sir James Lighthill's life and work we direct the interested reader to [167].

$$G_{\circ}(R, \tau) = 0 \quad \text{for } \tau < 0. \quad (4.94b)$$

It is not difficult to see that Eq. (4.93) and the defining relations (4.94a)–(4.94b) can be stated entirely in terms of R , τ , ∇_R and $\partial/\partial\tau$. Hence, the solution G_{\circ} can likewise be expressed in terms of the new variables alone. ■

Notice that this argument succeeds solely because $|\mathbf{r}_P - \mathbf{r}_{\partial\Omega}| \rightarrow \infty$ which results in zero boundary conditions. In contrast, as long as the bounding surface is at a finite distance from the observer, G depends on \mathbf{r} and \mathbf{r}_P *individually* as shown in detail above.

As with the general case, G_{\circ} is invariant under time-reversal and, making use of the spherical symmetry of the problem just as in Eqs. (4.33)–(4.35), we may express the reciprocal of Eq. (4.89) as

$$\begin{aligned} \frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial}{\partial R} G_{\circ}(R, \tau) \right) - \frac{1}{c^2} \left(\frac{\partial^2}{\partial \tau^2} G_{\circ}(R, \tau) \right) \\ = \frac{\partial^2}{\partial R^2} (R G_{\circ}(R, \tau)) - \frac{1}{c^2} \frac{\partial^2}{\partial \tau^2} (R G_{\circ}(R, \tau)) = -\frac{2\delta(R)}{R} \delta(\tau). \end{aligned} \quad (4.95)$$

The transformation of the first delta function on the right-hand side arises from the identity [125]

$$\delta(R) = \frac{1}{4\pi} \frac{\delta(R)}{R^2} \quad \text{in } \mathbb{R}^3 \times (0, \infty).$$

Proceeding exactly as in Eqs. (4.73)–(4.86), we find that the source term u_{\circ} is given by

$$u_{\circ}(\mathbf{r}_P, t_P) = \int_{t_{P_0}}^{t_P^+} dt \int_{\Omega} d^3\mathbf{r} G_{\circ}(R, \tau) q(\mathbf{r}, t), \quad (4.96)$$

and the integral w_{\circ} , which accounts for the propagation of the Cauchy data, by

$$w_{\circ}(\mathbf{r}_P, t_P) = -\frac{1}{4\pi c^2} \int_{\Omega} d^3\mathbf{r} \left[\frac{\partial G_{\circ}(R, t_{P_0})}{\partial t} \psi(\mathbf{r}, t_{P_0}) - \frac{\partial \psi(\mathbf{r}, t_{P_0})}{\partial t} G_{\circ}(R, t_{P_0}) \right]. \quad (4.97)$$

Taken separately, the two terms in Eq. (4.97) solve the homogeneous problem $\square w_{\circ} = 0$ with homogeneous boundary and inhomogeneous initial conditions. As in the general case, they may be ignored without loss of generality.

The boundary term v_o , however, prominently featured and shrouded in mystery in [168] and [169], warrants – precisely for this reason – a mathematically more rigorous treatment. In analogy with expressions (4.96) and (4.97), it can formally be stated as

$$v_o(\mathbf{r}_P, t_P) = \frac{1}{4\pi} \int_{t_{P_0}}^{t_P^+} dt \int_{\partial\Omega} d^2\mathbf{r} \, \hat{\mathbf{n}} \cdot \left[G_o(R_{\partial\Omega}, \tau) \nabla \psi(\mathbf{r}_{\partial\Omega}, t) - [\nabla G_o(R_{\partial\Omega}, \tau)] \psi(\mathbf{r}_{\partial\Omega}, t) \right], \quad (4.98)$$

where the subscript $\partial\Omega$ denotes – as before – values on the bounding surface.

It is apparent that the integration of G over volume in the first and surface in the second term of Eq. (4.88) play very similar parts. What is less evident, however, is how the *same* Green's function that solves Eq. (4.59) under homogeneous boundary conditions can be utilized to determine the field caused by inhomogeneous boundary conditions on the surface. The answer lies in replacing $\partial\Omega$ by a (virtual) *surface distribution of charge* located just inside the bounding sheet. (See [129], pp. 795 ff. for an extensive explanation.) This maneuver is hardly surprising if one is familiar with image charges in electrostatics or ground planes in antenna theory: Boundary conditions on a surface can be substituted for source distributions located infinitely close to the same surface.

This principle – described as “dualism between sources and boundary conditions” in [129] – is worthy of emphasis:

On the one hand, we have an extended source q governed by certain initial conditions. To find its contribution to the radiation field at $P(\mathbf{r}_P, t_P)$ we integrate $G(\mathbf{r}_P, t_P | \mathbf{r}, t) \psi(\mathbf{r}, t)$, which represents one point-like volume element occupied by the source, over Ω .

On the other, we have a bounding surface with certain constraints which can be represented as a source sheet. To find its contribution to the radiation field at $P(\mathbf{r}_P, t_P)$ we integrate either $G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t) \psi(\mathbf{r}_{\partial\Omega}, t)$ or $[\nabla G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t)] \psi(\mathbf{r}_{\partial\Omega}, t)$,

which represents one very small element of the surface²⁹, over $\partial\Omega$.

In both calculations it is very evident that R represents the distance from the location of emission in question, be the latter on the boundary or within the source, to an observer. Hence, if $\partial\Omega \rightarrow \infty$, the distance between any given boundary point $\mathbf{r}_{\partial\Omega}$ and an observer stationed at \mathbf{r}_P becomes, by definition, infinite and $G(\mathbf{r}_P, t_P | \mathbf{r}_{\partial\Omega}, t)$ vanishes due to condition 4.91a. Hence, irrespective of the value of $\psi(\mathbf{r}_{\partial\Omega}, t)$, *the surface integral is identically zero* and the solution to Eq. (4.1) in free space assuming null initial conditions reduces to

$$\psi(\mathbf{r}_P, t_P) = \int_{t_{P_0}}^{t_P^+} dt \int_{\Omega} d^3\mathbf{r} G_o(R, \tau) q(\mathbf{r}, t) \quad (4.99)$$

irrespective of the source speed.

4.3.5 At Last: The Explicit Form of the Green's Function

Various methods can be applied to find the explicit form of the Green's function for the wave equation in free space. Here we proceed to remove the time dependence of G_o in Eq. (4.93) by introducing a Fourier transform with respect to frequency, *i.e.*,

$$G_o(\mathbf{r}_P, t_P | \mathbf{r}, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega \tilde{G}_o(\mathbf{r}_P, \omega | \mathbf{r}, t) e^{-i\omega t_P} \quad (4.100)$$

along with the corresponding inverse transform

$$\tilde{G}_o(\mathbf{r}_P, \omega | \mathbf{r}, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt_P G_o(\mathbf{r}_P, t_P | \mathbf{r}, t) e^{i\omega t_P}. \quad (4.101)$$

²⁹To distinguish between the two cases, [129] on p. 791 introduces the superscript s for elements situated on the boundary, *i.e.*, $G(\mathbf{r} | \mathbf{r}_0^s)$, where \mathbf{r}_0 denotes a retarded quantity.

Substituting Eq. (4.101) into Eq. (4.1) we find that the latter now satisfies the inhomogeneous Helmholtz equation³⁰

$$\begin{aligned} \left(\nabla^2 + \frac{\omega^2}{c^2} \right) \tilde{G}_o &= -4\pi \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\infty} dt_P \delta(\mathbf{r}_P - \mathbf{r}) \delta(t_P - t) e^{i\omega t_P} \\ &= -\frac{4\pi}{\sqrt{2\pi}} \delta(\mathbf{r}_P - \mathbf{r}) e^{i\omega t}. \end{aligned} \quad (4.102)$$

Hence, if we define a function g_o such that

$$g_o(\mathbf{r}_P | \mathbf{r}) = \sqrt{2\pi} e^{-i\omega t_P} \quad (4.103)$$

we arrive at

$$(\nabla^2 + k^2) g_o(\mathbf{r}_P | \mathbf{r}) = -4\pi \delta(\mathbf{r}_P - \mathbf{r}), \quad (4.104)$$

where $k = \omega/c$ denotes the wave number. As we saw in the previous section, in the absence of a bounding surface, the Green's function depends only on $R = |\mathbf{r}_P - \mathbf{r}|$. Making use of the spherical symmetry of the problem we utilize, once more, Eqs. (4.33)–(4.35) and write

$$\frac{1}{R} \frac{d^2}{dR^2} R g_o + \frac{\omega^2}{c^2} g_o = -4\pi \delta(\mathbf{R}). \quad (4.105)$$

For $\mathbf{R} \neq 0$, the right hand side of Eq. (4.105) is identically zero and $R g_o$ satisfies the homogeneous equation

$$\frac{1}{R} \frac{d^2}{dR^2} R g_o + \frac{\omega^2}{c^2} g_o = 0 \quad (4.106)$$

which has the solution

$$\begin{aligned} R g_o &= A e^{ikR} + B e^{-ikR} \\ g_o &= \frac{1}{R} (A e^{ikR} + B e^{-ikR}). \end{aligned} \quad (4.107)$$

³⁰The last step is due to the identity

$$\int_{-\infty}^{-\infty} dx \delta(x - x_0) e^{i\omega x} = e^{i\omega x_0}.$$

Near the origin, where the delta function contributes, the second term on the left-hand side of Eq. (4.104) is negligible compared to the first and the equation becomes

$$\nabla^2 g_o = -4\pi\delta(\mathbf{R}), \quad (4.108)$$

which is satisfied by

$$g_o = \frac{1}{R}. \quad (4.109)$$

It is easy to see that this is consistent with Eq. (4.107) provided that $A + B = 1$. In consequence, the general solution for the Green's function is

$$\tilde{G}_o(\mathbf{r}_P, \omega|\mathbf{r}, t) = \frac{1}{\sqrt{2\pi R}} [Ae^{ikR} + (1 - A)e^{-ikR}] e^{i\omega t}. \quad (4.110)$$

Using the inverse transform given in Eq. (4.101), we finally write the time-dependent Green's function as

$$G(\mathbf{r}_P, t_P|\mathbf{r}, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi R}} (Ae^{ikR} + Be^{-ikR}) e^{i\omega t} e^{-i\omega t_P} d\omega \quad (4.111)$$

$$= \frac{1}{\sqrt{2\pi R}} \int_{-\infty}^{\infty} Ae^{i\omega(R/c + t - t_P)} + Be^{i\omega(-R/c + t - t_P)} d\omega \quad (4.112)$$

$$= \frac{A}{R} \delta[t - (t_P - R/c)] + \frac{B}{R} \delta[t - (t_P + R/c)]. \quad (4.113)$$

The first term in Eq. (4.111) is called the *retarded* Green's function because it exhibits the behavior associated with wave disturbances: At position \mathbf{r}_P and a particular reception time t_P , the only contribution to the received signal possible from a point at position \mathbf{r} was emitted at a time $t = t_P - R/c$, where c is the speed at which the electromagnetic disturbances propagate. The second term, the *advanced* Green's function, must be rejected on the ground that it does not obey the causality condition imposed earlier, namely, that no response may be predicted to an event occurring in the future. Hence, $B \equiv 0$, which implies $A = 1$ and

$$G(\mathbf{r}_P, t_P|\mathbf{r}, t) = \frac{\delta[(R/c) - (t_P - t)]}{R}; \quad R, t_P - t > 0. \quad (4.114)$$

Chapter 5

The Curious Case of the Nonspherically Decaying Radiation Field

“Excellent!” I cried. “Elementary,” said he.

— Watson and Holmes in *The Crooked Man*

5.1 The Case

The previous chapter is dedicated to the calculation of the complete and correct solution to the free-space scalar wave equation that satisfies all conditions *irrespective of the source speed* (Eq. (4.99)). In the course of this derivation, it has been shown unambiguously that the surface integral which accounts for contributions from the boundary is identically zero. Despite this, Houshang Ardavan has for decades claimed a primary part for the boundary term in his work, even accusing canonical texts such as [53] of “neglecting” it. Unsurprisingly, this has led to a prolonged and vitriolic

dispute in the scientific literature. To the casual eye, the disagreement seems purely academic: Whether or not the boundary term is non-zero matters little, as Ardavan uses the solution to the scalar wave equation for the *potential* rather than the field as the basis for further calculations, spiriting away the surface integral using an unnecessary gauge transformation. The true nature of the argument is, on the contrary, that he posits the existence of a non-zero boundary term to lend credence to his long-held belief that a smooth source of electromagnetic radiation can, if accelerated beyond the speed of light, generate radiation that decays in power more slowly than the square of the distance *out to infinity*.

Ardavan's sensationalist claims along with the controversy surrounding them, threatened to bring all work on faster-than-light sources of radiation into disrepute. After enduring several years of this unsatisfactory situation, and stimulated by a further Ardavan publication in 2019 [170], the current author and her mentor decided to publish a rebuttal that is intended to exorcise this longstanding mistake for good. The paper is reproduced in its entirety in the following pages. Note that for reasons of completeness, some sections of Chapter 4 are, of necessity, repeated.

5.2 The Evidence

Under consideration for publication in J. Plasma Phys.

1

Flaws in the theory of electromagnetic radiation “whose decay violates the inverse-square law”: mathematical and physical considerations

A. Schmidt[†] and J. Singleton[‡]

National High Magnetic Field Laboratory, MS-E536, Los Alamos National Laboratory,
P.O.Box 1663, Los Alamos, NM 87545, USA

(Received xx; revised xx; accepted xx)

In a recent paper in the present journal, H. Ardavan repeats his longstanding claim that the inhomogeneous wave equation in three dimensions as applied to electromagnetic radiation can yield solutions that fall off more slowly with distance than the inverse-square law *out to infinity* (Ardavan 2019). According to Ardavan, certain parts of the fields radiated by an extended source distribution that rotates faster than light *in vacuo* possess this surprising quality. To satisfy the conservation laws throughout the system, he theorizes that the contribution stemming from the boundary term (*i.e.*, the spatial surface enclosing the volume of interest) does not vanish with distance, but persists and eventually outgrows the source term. Here we show that this argument suffers from several flaws: (i) Kirchhoff’s formula as used by Ardavan is not a proper free-space solution, (ii) the absence of explicit boundary conditions results in an overdetermined problem which is ill-posed in the sense of Hadamard, (iii) the collection of field points (\mathbf{x}_P, t_P) is stationed *on* the boundary which leads to an indeterminate system as both recede to infinity and, most alarmingly, (iv) the value of the Green’s function for the problem is taken to be *the same* in all three integrals that appear in the solution to the wave equation. Once the correct free-space formula and boundary conditions are used, *the boundary term is shown to be identically zero*. This means that the fields can be calculated directly, and renders invalid Ardavan’s claim that the “rôle of the potential” is, in some way, “more fundamental” than that of the fields. In what follows, we will justify all of these contentions with the requisite mathematical rigour. Finally, the absence of the boundary term negates Ardavan’s justification for the presence of electromagnetic radiation “whose decay violates the inverse-square law” out to large distances.

1. Introduction

Since the mid 1980’s, Houshang Ardavan (Ardavan 1984, 1989, 1991, 1994, 1998, 1999; Ardavan & Ffowcs Williams 1995), later joined by several co-authors (Ardavan *et al.* 2003, 2004*d*, 2007, 2008*a*, 2009*c*), has proposed “non-spherically decaying solutions” to the inhomogeneous wave equation, *i.e.*, superpositions of electromagnetic signals whose intensity (power per unit area in the direction of propagation) falls off more slowly with distance than predicted by the inverse-square law *in unbounded space*[†]. [More

[†] Email address for correspondence: aschmidt@lanl.gov

[‡] Email address for correspondence: jsingle@lanl.gov

[†] For the sake of full disclosure, the present authors note that they are, as longstanding members of Los Alamos National Laboratory’s arm of the ‘superluminal research team’, contributors to many of the cited papers. However, in light of the findings to follow, they

recently, these claims have been repeated in a string of hastily filed patents (Ardavan & Ardavan 2017*a,b,c*, 2018) and, more publicly, the paper addressed here (Ardavan 2019).] According to Ardavan, radiation that has such an uncommonly long range arises on condition that an extended source distribution – usually a polarization current $\partial \mathbf{P}/\partial t$ – travels on a circular path faster than light *in vacuo*, leaving in its wake a caustic whose cusp possesses this unusual property.

We disagree emphatically with the statements made by Ardavan. Whilst simulations and experiments (Schmidt-Zweifel 2020) have shown[†] that radiation emitted by practical faster-than-light sources[‡] does possess a non-spherically decaying part, this effect is restricted by the qualitative features of the polarization current (*e.g.*, speed, acceleration, frequency content, volume distribution) and is observed – irrespective of Ardavan’s persistent claims to the contrary (Ardavan & Ardavan 2017*a,b,c*, 2018; Ardavan *et al.* 2003, 2004*d*, 2007, 2008*a*, 2009*c*) – *only over a limited range of distance*. This is true for all of the laboratory-based sources that have been built, including the example that is the basis of the antenna referred to in Ardavan (2019) (see, *e.g.*, Ardavan’s Figures 1 and 2 and associated discussion)[¶].

The *IEEE Standard Definition of Terms for Antennas* defines the radiating near-field or Fresnel region as “that region of the field of an antenna between the reactive near-field region and the far-field region wherein radiation fields predominate and wherein the angular field distribution is dependent upon the distance from the antenna” (IEEE Standards Association 2013). Therefore, the presence of radiation that decays more slowly than the inverse-square law is merely a characteristic of the Fresnel region, even in the case of a superluminal antenna. The superluminal polarization current antennas built by us (Schmidt-Zweifel 2020) and referred to in Ardavan (2019) differ from traditional aeriels, arrays *etc.* in that they are true volume sources. It is therefore possible that the usual formula for the Fresnel region (Balanis 2005) may not hold. For example, the circular superluminal antenna built at Los Alamos (Schmidt-Zweifel 2020) has decay rates that differ from the inverse square law out to distances $\sim 20 - 30$ m (but not beyond), whereas the usual formulae suggest a Fresnel region that is an order of magnitude smaller. By contrast, given the size of practical antennas (Schmidt-Zweifel 2020), the distance over which the nonspherical decay is claimed to occur in Ardavan (2019) becomes as large as 1,000,000 km to 100,000,000 km (page 86).

We are not alone in objecting to Ardavan’s claim. Other authors, such as University of Bristol’s J. H. Hannay (Hannay 1996, 2000, 2001, 2006, 2008, 2009), Nobel laureate A. Hewish (Hewish 1996, 2000), and K. T. McDonald of Princeton University (McDonald 2004*a,b*) have long argued that a superposition of electromagnetic waves whose flux decays more slowly than the square of the distance out to infinity is incompatible with conventional solutions of Maxwell’s equations. (Replies to their comments can be

wish to distance themselves from some of these publications, in particular (Ardavan *et al.* 2006, 2008*a,b*, 2009*b,c*)

[†] Descriptions of dielectric antennas that generate polarization currents exceeding the speed of light – the only ones of their kind! – can be found in Ardavan *et al.* (2004*b,c*); Singleton *et al.* (2004); Schmidt-Zweifel (2013, 2020). A number of technical reports are available to the interested reader upon request.

[‡] Sometimes called *Schott radiation* in the literature (Costa & Kahn 1985).

[¶] The superluminal antenna described in Section 2 of Ardavan (2019) is a hybrid of an antenna designed, built and tested at Los Alamos National Laboratory (LANL) by a collaboration of scientists and engineers led by John Singleton (Schmidt-Zweifel 2020) and a proposed antenna that was never built because it was calculated to be too inefficient. Although Ardavan acted as an external consultant in the early stages of the project, he had no part in the construction and testing of the LANL antenna or the analysis and modeling of the resulting data.

found in Ardavan (2000); Ardavan *et al.* (2004a, 2006, 2008b,c, 2009a,b).) Perhaps more pertinent to Ardavan (2019), no lasting evidence – simulated or observational – has been found to indicate that pulsar radiation outdoes the inverse-square law (Costa & Kahn 1985; Desai 2016).

2. The Point of Contention

Many of Ardavan’s problems revolve around the equation governing the magnetic flux density

$$\nabla^2 \mathbf{B} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{B} = -\frac{4\pi}{c} \nabla \times \mathbf{j} \quad (2.1)$$

where the source is a volume-distributed current density \mathbf{j} (which may be a polarization current density) and c denotes the speed at which the electromagnetic disturbances propagate. In Ardavan (2019) and elsewhere, the free-space Green’s function solution to Eq. (2.1) is stated as[†]

$$\begin{aligned} B_i(\mathbf{x}_P, t_P) = & \frac{1}{c} \int_0^{t_P} dt \int_{\mathcal{D}} d^3\mathbf{x} \left[[\nabla \times \mathbf{j}]_i(\mathbf{x}, t) G(\mathbf{x}_P, t_P | \mathbf{x}, t) \right]_{\{\mathbf{x} \in \mathcal{D}\}} \\ & + \frac{1}{4\pi} \int_0^{t_P} dt \int_{\partial\mathcal{D}} d^2\mathbf{x} \hat{n} \cdot \left[G(\mathbf{x}_P, t_P | \mathbf{x}, t) \nabla B_i(\mathbf{x}, t) \right. \\ & \quad \left. - [\nabla G(\mathbf{x}_P, t_P | \mathbf{x}, t)] B_i(\mathbf{x}, t) \right]_{\{\mathbf{x} \in \partial\mathcal{D}\}} \\ & - \frac{1}{4\pi c^2} \int_{\mathcal{D}} d^3\mathbf{x} \left[\frac{\partial G(\mathbf{x}_P, t_P | \mathbf{x}, t)}{\partial t} B_i(\mathbf{x}, t) \right. \\ & \quad \left. - \frac{\partial B_i(\mathbf{x}, t)}{\partial t} G(\mathbf{x}_P, t_P | \mathbf{x}, t) \right]_{t=0}. \end{aligned} \quad (2.2)$$

and the corresponding Green’s function itself as

$$G(\mathbf{x}, t; \mathbf{x}_P, t_P) = \frac{\delta(t_P - t - R/c)}{R}, \quad (2.3)$$

where, as in Ardavan (2019), $i = 1, 2, 3$ denotes the vector components of \mathbf{B} and $R = |\mathbf{R}| \equiv |\mathbf{x}_P - \mathbf{x}|$. (The subscripts at the end of each integral were added by the present authors to mitigate the arguments to follow.) Proposed in 1828 by George Green in his seminal *Essay on the Application of Mathematical Analysis to the Theory of Electricity and Magnetism* (Green 1850), the method that bears his name is as simple as it is elegant: To obtain the field caused by a volume-distributed source q we calculate the effect of each elementary (or point-like) portion of the source and, by virtue of the superposition principle, sum the solutions. If $G(\mathbf{x}_P | \mathbf{x})$ is the impulse generated by a unit point charge located at source point \mathbf{x} , then the field at \mathbf{x}_P due to $q(\mathbf{x})$ is the integral of $q(\mathbf{x})G(\mathbf{x}_P | \mathbf{x})$ taken over the space occupied by the source (Morse & Feshbach 1953). In consequence, G too has to satisfy the inhomogeneous wave equation, *i.e.*,

$$\nabla^2 G - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} G = 4\pi \delta(\mathbf{x}_P - \mathbf{x}) \delta(t_P - t). \quad (2.4)$$

[†] Adopting the notation used in all of Ardavan’s works, we reserve the subscript P to denote coordinates and times where we seek the solution $B_i(\mathbf{x}_P, t_P)$ – the *field* points – whereas \mathbf{x} and t – the *source* points – define *loci* where data are prescribed. In a nutshell, light must be emitted at time t and place \mathbf{x} to reach field point \mathbf{x}_P at t_P . The same is true for derivatives, where ∇ , $\partial/\partial t$, and $\partial^2/\partial t^2$ are retarded quantities and, in general, $\nabla \neq \nabla_P$, $\partial/\partial t \neq \partial/\partial t_P$, and $\partial^2/\partial t^2 \neq \partial^2/\partial t_P^2$.

The resulting solution (Eq. (2.2)) can be split into three individual integrals: The first is the *source term* which captures radiation, emitted by the time-varying current density $\nabla \times \mathbf{j}$ as it travels through space and time. The last accounts for the propagation of initial conditions. Under the restriction $B_i|_{t=0} = \partial B_i / \partial t|_{t=0} = 0$, which may be assumed here without loss of generality, it is identically zero. The remaining surface integral, commonly called the *boundary term*, represents contributions to $B_i(\mathbf{x}_P, t_P)$ due to the conditions imposed at the spatial boundary $\partial\mathcal{D}$.

We note in passing that the Green's function given in Eq. (2.3) disagrees with its cited origin (Morse & Feshbach (1953), p. 893). The discrepancy between the right-hand sides can be resolved by remembering that $\delta(x)$ is an even function of its argument x . The left-hand side, however, differs in notation from traditional texts in that Ardavan has switched the order of field (\mathbf{x}_P, t_P) and source (\mathbf{x}, t) points. If the reversal is intended to denote the reciprocal it is incorrect; as the authors of Morse & Feshbach (1953) state explicitly on p. 835, “the generalization of the reciprocity relation ... is *not* $G(\mathbf{x}_P, t_P; \mathbf{x}, t) = G(\mathbf{x}, t; \mathbf{x}_P, t_P)$ ”. This is arguably a notational issue of little import as long as it is understood that in Eqs. 2.2 – 2.4 G is a function of *retarded quantities*. Nevertheless, we prefer to use the more conventional form

$$G(\mathbf{x}_P, t_P | \mathbf{x}, t) = \frac{\delta[(R/c) - (t_P - t)]}{R}; \quad R, t_P - t > 0. \quad (2.5)$$

To attain nonspherical radiation decay whilst satisfying the requirement that energy be conserved throughout the system, Ardavan emphasizes the boundary term [*i.e.*, the second integral in Eq. (2.2)] supposedly neglected in text books and standard literature. For example, at the very outset he claims

“The a priori assumption that the retarded field like the retarded potential automatically satisfies the boundary conditions at infinity is moreover unfounded as we shall see in this paper” [page 2 of Ardavan (2019)].

Moreover, in Section 3 (page 19) of Ardavan (2019), he states

“In the case of a source whose distribution pattern rotates superluminally, where the radiation field decays non-spherically (more slowly than R^{-1}) with distance, the boundary term in the retarded solution (3.8) for the field is in fact larger than the source term of this solution, in the limit where the closed surface $\partial\mathcal{D}$ tends to infinity (Ardavan et al. 2008a). Given that the distribution of the radiation field of an accelerated superluminal source in the far zone is not known a priori, to be prescribed as a boundary condition, it follows that the only way one can calculate the free-space radiation field of such sources is via the retarded solution for the potential.”

In what follows we show with the requisite mathematical rigour that the claims made in Sections 1 – 3 of Ardavan (2019) are without merit. In particular, correct treatment shows that the boundary term is *identically zero in problems of this kind*. Whereas this does not necessarily preclude the existence of electromagnetic radiation that disobeys the inverse-square law, the absence of an energy source that fuels Ardavan's mechanism undermines the theory that he describes. Moreover, the absence of a boundary term in problems of this nature renders invalid the statements (Ardavan (2019), page 3)

“There is a fundamental difference between the classical expression for the retarded potential and the corresponding retarded solution of the wave equation that governs the electromagnetic field”

and (Ardavan (2019), page 19)

“the only way one can calculate the free-space radiation field of such sources is via the retarded solution for the potential.”

In other words, there is no difference between the retarded solutions derived from the

two wave equations that govern the electromagnetic potential and the electromagnetic field. To make this distinction (as Ardavan attempts to do) is incorrect.

(Of course it is perfectly legitimate to calculate \mathbf{B} indirectly from the potential

$$A_i(\mathbf{x}_P, t_P) = \frac{1}{c} \int_0^{t_P} dt \int_{\mathcal{D}} d^3\mathbf{x} \left[j_i(\mathbf{x}, t) \frac{\delta[(R/c) - (t_P - t)]}{R} \right]_{\{\mathbf{x} \in \mathcal{D}\}} \quad (2.6)$$

– listed as Eq. (3.7) in Ardavan (2019) – as is routinely done for convenience in antenna theory (Balanis 2005). In fact, most simulations in Schmidt-Zweifel (2020) are based on the potentials and are in excellent agreement with data gathered from the superluminal antennas designed, built and tested at Los Alamos. This point is, however, subsidiary to the larger issues addressed here.)

3. The Importance of Properly Posed Conditions

Although Ardavan ascribes Eq. (2.2) to Morse & Feshbach (1953) it was, in fact, first presented by Gustav Robert Kirchhoff (1824–1887) in *Zur Theorie der Lichtstrahlen* and constitutes the mathematical expression of Huygens’ principle. Though chiefly known nowadays for his contributions to circuit theory, Kirchhoff pioneered the use of Green’s function solutions in electromagnetism. However, despite its empirical application to diffraction theory (as Ardavan points out (Ardavan *et al.* 2008*c,b*)), Eq. (2.2) has two basic shortcomings: (i) Kirchhoff considers “at infinity” to mean “remote from the region of interest” (Jackson 1999) (yet at finite distance) which implies that Eq. (2.2) is not a proper free-space solution and (ii) the problem is ill-posed in the sense of Hadamard (Hadamard 1964) since it is overdetermined. (Fig. 1 gives a diagrammatic representation.)

Of course, Kirchhoff’s notion of infinity is in accord with the beliefs of his day, which is why the term is given in quotation marks throughout the discussion on p. 478 of Jackson (1999). In fact, it was not until 1883 that Cantor removed the ambiguity. As a direct consequence it is understood throughout mathematical physics (Morse & Feshbach 1953; Evans 1998; MyInt-U & Debnath 2006; Debnath 2005; Zauderer 1989; Hilbert & Courant 1931; Hassani 1999; Dudley 1994; Stakgold & Holst 2011; Barton 1991; Duffy 2015), electromagnetism (Jackson 1999; Balanis 2012; Stratton 1941; Griffiths 2015) and antenna theory (Balanis 2005) that free space is more than a spherical region whose bounding surface recedes farther and farther. We demand that, at spatial infinity, *only outward-bound disturbances* may result from sources or initial values that are effectively localized within a finite volume of space (Barton 1991). Since the three dimensional wave equation has solutions that vanish only as $1/R$ when $R \rightarrow \infty$, the boundary conditions imposed on the problem must exclude not only waves arriving from infinitely distant sources but also those produced by radiators or reflectors suitably positioned at infinity. If not, standing waves (or normal modes) characteristic of finite enclosures might persist for all eternity and potentially furnish solutions that are mathematically quite legitimate but incompatible with the laws of physics (Barton 1991).

Even if we take Eq. (2.2) to be the solution to the general (rather than free-space) problem, it is improperly posed in the sense of Hadamard. Whereas the initial conditions are Cauchy, *i.e.*, both B_i and $\partial B_i / \partial t$ at some initial time are given, Kirchhoff does not explicitly prescribe any boundary conditions for the Green’s function on $\partial\mathcal{D}$ (and neither does Ardavan). However, considering that both $G(\mathbf{x}_P, t_P | \mathbf{x}, t)$ and the gradient $\nabla G(\mathbf{x}_P, t_P | \mathbf{x}, t)$ appear in the surface integral of Eq. (2.2) we have to assume that they are Dirichlet *and* Neumann. [Ardavan himself admits as much when he incorrectly describes the bounding surface as “closed”; *e.g.*, see Ardavan (2019), page 19.] The

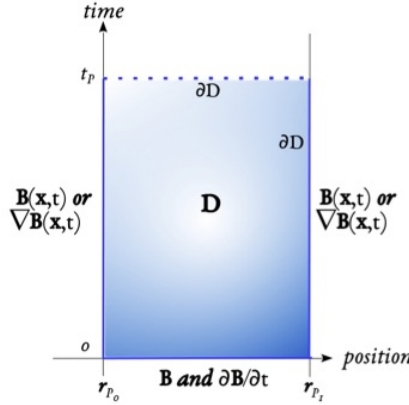


FIGURE 1. The problem of solving the wave equation is well-posed if either Dirichlet or Neumann boundary conditions (but not both) are prescribed along the vertical edges of Ω . Along the lower horizontal, Cauchy initial conditions are wanted. *No restrictions may be placed along the top horizontal*, as this would result in an overdetermined system.

Cauchy problem for hyperbolic equations is in general well-posed under Dirichlet or Neumann boundary conditions prescribed on an *open* surface in space-time (the proof is due to Sofia Kovalevskaya); imposing both, however, results in constraints that are either incompatible or ineffective and a general existence proof will fail. (*Nota bene* that this is true for sources of *all* velocities, not only those that travel faster than light.) The reason why $\partial\mathcal{D}$ must be open is immediately interpretable in terms of causality: At position \mathbf{x}_P and a particular reception time t_P , the only contribution to the received signal possible from a point situated at \mathbf{x} was emitted at $t = t_P - R/c$.

If initial and boundary conditions that lead to a well-posed problem are chosen, *i.e.*,

$$G(\mathbf{x}_P, t_P | \mathbf{x}, t) = 0 \quad \text{or} \quad \nabla G(\mathbf{x}_P, t_P | \mathbf{x}, t) = 0 \quad \forall \mathbf{x} \in \partial\mathcal{D} \quad (3.1a)$$

$$G(\mathbf{x}_P, t_P | \mathbf{x}, t) = 0 \quad t_P < t, \quad (3.1b)$$

either G or its gradient is identically zero for all \mathbf{x} on $\partial\mathcal{D}$, eliminating one of the two terms in the surface integral[†]. (This critique of Kirchhoff's formula is neither new nor original; similar arguments are made by Jackson (1999), pp. 478 and Barton (1991), p. 284.) In consequence, Eq. (2.2) is an explicit construct of the data $(\nabla \times \mathbf{j})_i$, B_i or ∇B_i , $B_i|_{t=0}$, and $\partial B_i/\partial t|_{t=0}$, and *not*, as claimed in Ardavan *et al.* (2009c) and other works by the same author, “merely a mathematical identity”.

4. Boundaries in Free Space

In electrical engineering, the Green's function solution to Eq. (2.1) along with Cauchy data and Dirichlet or Neumann boundary conditions is used to calculate the spread of radiated fields throughout systems of *finite size* such as resonant cavities, conducting cylinders, and waveguides. As might reasonably be expected, advanced texts in electromagnetism discuss the boundary term at length in this context (see e.g. Ch. 15

[†] Eq. (3.1b) arises since it seems reasonable to impose causality as an initial condition, which is to say that G and its time derivative must be zero for $t_P < t$; if an impulse occurs at t , no effects of this impulse should be present anywhere at an earlier time.

of Balanis (2012)). Ardavan (2019), however, aims to describe the field emitted by a collection of rotating unit point sources into an *infinite domain* which obeys causality and is everywhere quiescent before some initial time $t = 0$. It is under these conditions, he claims, that the boundary term outgrows the source term to become the dominant contribution toward the fields. In what is to follow, we will not only show that this assertion is unfounded since it is based on a very elementary misapprehension, but that the boundary term in the free-space solution to Eq. (2.1) is *identically zero*. (Incidentally, this explains why it is “neglected” in standard texts such as Jackson (1999)). Hannay has made an argument based on Maxwell’s equations before (Hannay 2008, 2009); here we provide a detailed *mathematical* explanation to support his findings.

It is apparent that the integration of G over volume in the first term and surface in the second term of Eq. (2.2) play very similar parts. What is less evident, however, is how the *same* Green’s function that solves Eq. (2.4) under homogeneous boundary conditions can be utilized to determine the field caused by inhomogeneous boundary conditions on the surface. The answer lies in replacing $\partial\mathcal{D}$ by a (virtual) *surface distribution of charge* located just inside the bounding sheet. (See Morse & Feshbach (1953), pp. 795 for an extensive explanation.) This maneuver is hardly surprising if one is familiar with image charges in electrostatics or ground planes in antenna theory: Boundary conditions on a surface can be substituted for source distributions located infinitely close to the same surface.

This principle – described as “dualism between sources and boundary conditions” in Morse & Feshbach (1953) – is worthy of emphasis.

On the one hand, we have an extended source – for instance the time-varying current density \mathbf{j} – governed by certain initial conditions. To find its contribution to the radiation field at $P(\mathbf{x}_P, t_P)$ we integrate $G(\mathbf{x}_P, t_P|\mathbf{x}, t)\nabla \times \mathbf{j}(\mathbf{x}, t)$, which represents one point-like volume element occupied by the source, over \mathcal{D} .

On the other, we have a bounding surface with certain constraints which can be represented as a source sheet. To find its contribution to the radiation field at $P(\mathbf{x}_P, t_P)$ we integrate either $G(\mathbf{x}_P, t_P|\mathbf{x}, t)\nabla \mathbf{B}(\mathbf{x}, t)$ or $[\nabla G(\mathbf{x}_P, t_P|\mathbf{x}, t)]\mathbf{B}(\mathbf{x}, t)$, which represents one very small element of the surface, over $\partial\mathcal{D}^\dagger$.

In both calculations it is very evident that R represents the distance from the location of emission in question, be the latter on the boundary or within the source, to an observer. Hence, if $\partial\mathcal{D} \rightarrow \infty$, the distance between any given boundary point \mathbf{x} and an observer stationed at \mathbf{x}_P becomes, by definition, infinite and $G(\mathbf{x}_P, t_P|\mathbf{x}, t)$ vanishes due to the requirement that all disturbances must be outward-bound. This suggests the following conditions for the free-space Green’s function:

$$G(\mathbf{x}_P, t_P|\mathbf{x}, t) \rightarrow 0 \text{ and } \nabla G(\mathbf{x}_P, t_P|\mathbf{x}, t) \rightarrow 0 \quad \forall \mathbf{x} \in \partial\Omega \mid \partial\Omega \rightarrow \infty \quad (4.1a)$$

$$G(\mathbf{x}_P, t_P|\mathbf{x}, t) = 0 \quad t_P < t. \quad (4.1b)$$

Hence, irrespective of the value of $\{B_i(\mathbf{x}, t): \mathbf{x}, t \in \partial\mathcal{D}\}$, the *surface integral is identically zero* and the solution to Eq. (2.1) in free space assuming null initial conditions reduces to

$$B_i(\mathbf{x}_P, t_P) = \frac{1}{c} \int_0^{t_P} dt \int_{\mathcal{D}} d^3\mathbf{x} \left[[\nabla \times \mathbf{j}]_i(\mathbf{x}, t) G(\mathbf{x}_P, t_P|\mathbf{x}, t) \right]_{\{\mathbf{x} \in \mathcal{D}\}} \quad (4.2)$$

irrespective of the source speed.

It is for this reason – and not “assumptions and approximations” [page 2 of Ardavan

[†] To distinguish between the two cases, Morse & Feshbach (1953) on p. 791 introduces the superscript s for elements situated on the boundary, *i.e.*, $G(\mathbf{r}|\mathbf{r}_0^s)$, where \mathbf{r}_0 denotes a retarded quantity.

(2019)] – that advanced texts in electromagnetism such as Jackson (1999)) render the free-space solution to the scalar wave equation in three dimensions

$$\mathbf{B}(\mathbf{x}_P, t_P) = \frac{1}{c} \int d^3\mathbf{x} \frac{[\nabla \times \mathbf{j}]_{\text{ret}}}{|\mathbf{x}_P - \mathbf{x}|}. \quad (4.3)$$

The subscript “ret” signifies the evaluation of points within the integral *only* at the retarded times specified above; in other words, the δ -function in Eq. (2.5) picks out the relevant emission times.

Owing to nonstandard notation (and perhaps a certain lack of meticulousness) it is difficult to understand Ardavan’s conception of the bounding surface. Mathematical forensics performed on Ardavan *et al.* (2007, 2008a, 2009c) and Ardavan *et al.* (2008b,c, 2009b) reveal that R is taken to be both, “the distance ... of the boundary from the source” – which tends to infinity as required for a free-space problem – and “the magnitude of the separation $\mathbf{R} \equiv \mathbf{x}_P - \mathbf{x}$ ”. R_P is the radial component of the spherical polar coordinates that define observation point $P(R_P, \theta_P, \varphi_P)$, later set to approach infinity as well ($R_P \equiv |\mathbf{x}_P| \rightarrow \infty$). The only plausible conclusion that can be drawn from this arrangement is that P coincides with the boundary and, what is more, that both are stationed at infinity. Hence, not heeding the clear direction on p. 791 of Morse & Feshbach (1953) that the field point may *not* lie† on $\partial\mathcal{D}$, Ardavan seems to place it precisely there. Perhaps even more alarmingly, he then proceeds to compare the R_P (or, alternatively, R) dependence of the volume and the surface integral in Eq. (2.2) [a move considered “nonsense” by Hannay (2008)] obviously assuming that the value of the Green’s function (Eq. (2.5)) is the same in all three terms of the equation. As we have shown in detail above, this is usually not true in the general case and *never* in free space.

5. Apropos the Forward and the Inverse Problem

Just as supersonic airplanes leave a cone of intersecting pressure waves in their wake, the Huygens wavelets emitted by a faster-than-light source will superpose. Take, for instance, the (purely hypothetical) point charge $q(\mathbf{x}, t)$ in constant superluminal rotation which is the basis for most of Ardavan’s calculations. Since q outruns the waves it generates (which, as all others, must move *at c*) contributions from multiple (\mathbf{x}, t) pairs on its trajectory may arrive simultaneously at an observer or receiver. The inverse, however, is *not* true; q will emit once – and only once! – at any given point in spacetime. This is a crucial principle worthy of repetition: *Several wavefronts may intersect at $P(\mathbf{x}_P, t_P)$, but only one can originate from $P(\mathbf{x}, t)$.* Whether the latter contributes to the radiation measured at $P(\mathbf{x}_P, t_P)$ is determined by the sifting property of the Green’s function (see Eq. (2.5)). It follows that statements such as “This ... represents the *common emission time* of the *three* wave fronts that are mutually tangential at the cusp curve of the envelope” (emphasis by the present authors), published in Ardavan (1998) are inconsistent with causality‡.

The fact that more than one wavefront can arrive concurrently at a field point raises a number of questions (none of which are addressed in Ardavan (2019)), not least whether the field $\mathbf{B}(\mathbf{x}_P, t_P)$ radiated by a volume-distributed superluminal source is unique or if $G(\mathbf{x}_P, t_P | \mathbf{x}, t)$ – originally a function of (\mathbf{x}_P, t_P) – remains invariant under time translation

† The reason is, evidently, that for an observer situated on the boundary, $|\mathbf{x}_P - \mathbf{x}|_{\{\mathbf{x} \in \partial\mathcal{D}\}}$ is zero in general and indeterminate in free space.

‡ This is especially troublesome as already Einstein stated that it is the requirement of causality (e.g. the grandfather paradox) rather than relativistic considerations that precludes massive or information-carrying bodies from moving superluminally.

and reversal to furnish a solution for Eq. (2.2) in terms of retarded quantities. (A thorough mathematical analysis is forthcoming in Schmidt-Zweifel (2020).) More importantly, it suggests that the forward problem, which is in the *source's* reference frame, can be solved with comparative ease whereas the inverse problem, which is in the *observer's* reference frame leads to the intractable edifice of formulae and regulation techniques presented in Ardavan (2019). This predicament is, of course, not restricted to the present problem; consider, for instance, the inverse diffusion (or heat) equation where it is, in general, impossible to deduce the initial from the final state of the system uniquely (Barton 1991). In fact, inverse problems in mathematical physics, that is to say models that attempt to calculate from a partial or full knowledge of the solution the causal factors that produced it, are notoriously difficult to treat since they frequently (but by no means always) fail the requirements of uniqueness, continuous dependence on the boundary data, or both. In practice, solutions to inverse systems are often numerically “negotiated” (for instance by regularization schemes or parameter estimation) rather than analytically derived.

Ever since Ardavan introduced the “bifurcation surface of an observer” in 1998 (Ardavan 1998), he has attempted to solve the inverse problem; unnecessarily so, it seems. As pointed out above, all quantities that appear in Eqs. 2.5 and 4.2 (or, for that matter, Eq. (2.2)) *are retarded*. Inverse systems can often not be avoided, but in this case, the forward problem as solved and modeled in Schmidt-Zweifel (2020) provides a perfectly good solution.

6. Conclusion

The author of Ardavan (2019) claims, as in previous publications, that an extended source distribution in uniform superluminal rotation – in this case a polarization current $\partial \mathbf{P} / \partial t$ – leaves in its wake a caustic whose cusp diminishes more slowly than expected *out to infinity*. The energy needed to sustain this unusually slow radiation decay whilst satisfying the conservation laws throughout the system is supposedly furnished by the surface integral (*i.e.*, the boundary term) in the solution to the scalar wave equation (Ardavan *et al.* 2006, 2008*a,b,c*, 2009*a,b,c*). The absence – or “neglect” – of this integral in standard texts such as Jackson (1999) and Ardavan’s subsequent insistence that the formula used for “conventional” problems (*i.e.*, Eq. (4.2)) does not apply to a source whose distribution pattern moves faster than light has led to a vitriolic (and very public) dispute in the scientific literature (Hannay 1996, 2000, 2001, 2006, 2008, 2009; Hewish 1996, 2000; McDonald 2004*a,b*; Ardavan 2000; Ardavan *et al.* 2004*a*, 2006, 2008*b,c*, 2009*a,b*).

In hopes of ending this controversy for good we have demonstrated that Ardavan’s argumentation – besides being somewhat circular (see Appendix below) – suffers from several flaws, the most significant of which is that the value of the Green’s function for the problem is taken to be the same in all three integrals that appear in the solution to the wave equation. Once one accepts that the geometry of the problem dictates that the distance from an observer to the boundary is infinite in free space, it is easy to see that the surface integral is *identically zero*. This does not *per se* exclude the existence of radiation that decays more slowly than predicted by the inverse square law. However, without a source to fuel this unusual property of the radiated fields Ardavan will find it difficult indeed to uphold his theory without violating the basic laws of physics.

Acknowledgments

This work was supported by Los Alamos National Laboratory LDRD projects 20200285ER and 20180352ER, and carried out at the National High Magnetic Field

Laboratory, which is funded by NSF Cooperative Agreements DMR-1157490 and 1164477, the State of Florida and U.S. DoE. We thank Edl Schamiloglu, Bruce Carlsten and Frank Krawczyk for helpful discussions.

REFERENCES

- ARDAVAN, A., ARDAVAN, H. & SINGLETON, J. 2004a Synchrotron-Čerenkov Radiation – Response. *Science* **303**, 310–311.
- ARDAVAN, A., ARDAVAN, H. & SINGLETON, J. 2009a Comment on “Radial Dependence of Radiation from a Bounded Source”. *arXiv:physics/0610085v1 [physics.optics]*.
- ARDAVAN, A., HAYES, W., SINGLETON, J., ARDAVAN, H., FOPMA, J. & HALLIDAY, D. 2004b Corrected Article: “Experimental observation of nonspherically-decaying radiation from a rotating superluminal source” [J. Appl. Phys. 96, 4614, (2004)]. *J. Appl. Phys.* **96** (12), 7760–7777.
- ARDAVAN, A., HAYES, W., SINGLETON, J., ARDAVAN, H., FOPMA, J. & HALLIDAY, D. 2004c Experimental observation of nonspherically-decaying radiation from a rotating superluminal source. *J. Appl. Phys.* **96** (8), 4614–4631.
- ARDAVAN, H. 1984 A speed-of-light barrier in classical electrodynamics. *Phys. Rev. D* **29** (2), 207–215.
- ARDAVAN, H. 1989 The speed-of-light catastrophe. *Proc. R. Soc. London A* **424**, 113–141.
- ARDAVAN, H. 1991 The near-field singularity predicted by the spiral Green’s function in acoustics and electrodynamics. *Proc. R. Soc. London A* **433**, 451–459.
- ARDAVAN, H. 1994 The mechanism of radiation in pulsars. *Mon. Not. R. Astr. Soc.* **268**, 361–392.
- ARDAVAN, H. 1998 Generation of focused, nonspherically decaying pulses of electromagnetic radiation. *Phys. Rev. E* **58** (5), 6659–6684.
- ARDAVAN, H. 1999 Method of handling the divergences in the radiation theory of sources that move faster than their waves. *J. Math. Phys.* **40** (9), 4331–4336.
- ARDAVAN, H. 2000 Reply to Comments on “Generation of focused, nonspherically decaying pulses of electromagnetic radiation”. *Phys. Rev. E* **62** (2), 3010–3013.
- ARDAVAN, H. 2019 The electromagnetic radiation whose decay violates the inverse-square law: detailed mathematical treatment of an experimentally realized example. *J. Plasma Phys.* **85** (3), 905850304.
- ARDAVAN, H. & ARDAVAN, A. 2017a Aparatus for generating focused electromagnetic radiation. US Patent 2017/0323697 A1.
- ARDAVAN, H. & ARDAVAN, A. 2017b Equatorially and near-equatorially radiating arc-shaped polarization current antennas and related methods. US Patent 2017/0373404 A1.
- ARDAVAN, H. & ARDAVAN, A. 2017c Polarization current antennas that generate superluminal polarization current waves having acceleration and related methods of exciting such antennas. US Patent 2017/0271778 A1.
- ARDAVAN, H. & ARDAVAN, A. 2018 Equatorially and near-equatorially radiating arc-shaped polarization current antennas and related methods. US Patent 2018/0062273 A1.
- ARDAVAN, H., ARDAVAN, A. & SINGLETON, J. 2003 Frequency spectrum of focused broadband pulses of electromagnetic radiation generated by polarization currents with superluminally rotating distribution patterns. *J. Opt. Soc. Am. A* **20** (11), 2137–2155.
- ARDAVAN, H., ARDAVAN, A. & SINGLETON, J. 2004d Spectral and polarization characteristics of the nonspherically decaying radiation generated by polarization currents with superluminally rotating distribution patterns. *J. Opt. Soc. Am. A* **21** (5), 858–872.
- ARDAVAN, H., ARDAVAN, A., SINGLETON, J., FASEL, J. & SCHMIDT, A. 2006 Spectral and polarization characteristics of the nonspherically decaying radiation generated by polarization currents with superluminally rotating distribution patterns: reply to comment. *J. Opt. Soc. Am. A* **23** (6), 1535–1539.
- ARDAVAN, H., ARDAVAN, A., SINGLETON, J., FASEL, J. & SCHMIDT, A. 2007 Morphology of the nonspherically decaying radiation field generated by a rotating superluminal source. *J. Opt. Soc. Am. A* **24** (8), 2443–2456.
- ARDAVAN, H., ARDAVAN, A., SINGLETON, J., FASEL, J. & SCHMIDT, A. 2008a Fundamental

- role of the retarded potential in the electrodynamics of superluminal sources. *J. Opt. Soc. Am. A* **25** (3), 543–557.
- ARDAVAN, H., ARDAVAN, A., SINGLETON, J., FASEL, J. & SCHMIDT, A. 2008*b* Morphology of the nonspherically decaying radiation generated by a rotating superluminal source: reply to comment. *J. Opt. Soc. Am. A* **25** (9), 2167–2169.
- ARDAVAN, H., ARDAVAN, A., SINGLETON, J., FASEL, J. & SCHMIDT, A. 2008*c* Response to “Comment on ‘Methods of handling the divergences in the radiation theory of sources that move faster than their own waves’”. *arXiv:0805.0304v1 [math-ph]*.
- ARDAVAN, H., ARDAVAN, A., SINGLETON, J., FASEL, J. & SCHMIDT, A. 2009*b* Fundamental role of the retarded potential in the electrodynamics of superluminal sources: reply to comment. *J. Opt. Soc. Am. A* **26** (10), 2109–2113.
- ARDAVAN, H., ARDAVAN, A., SINGLETON, J., FASEL, J. & SCHMIDT, A. 2009*c* Inadequacies in the conventional treatment of the radiation field of moving sources. *J. Math. Phys.* **50** (10), 10.1063/1.3215978.
- ARDAVAN, H. & FFOWCS WILLIAMS, J. E. 1995 Violation of the inverse square law by the emissions of supersonically or superluminally moving volume sources. *arXiv:astro-ph/9506023v1*.
- BALANIS, C. A. 2005 *Antenna Theory: Analysis and Design*, 3rd edn. Hoboken, NJ: John Wiley & Sons, Inc.
- BALANIS, C. A. 2012 *Advanced Engineering Electromagnetics*, 2nd edn. Hoboken, NJ: John Wiley & Sons, Inc.
- BARTON, G. 1991 *Elements of Green’s Functions and Propagation*, 2nd edn. New York: Oxford University Press.
- COSTA, A. & KAHN, F. 1985 Pulsar electrodynamics – The back reaction of the motion of charged particles. *Mon. Not. R. Astr. Soc.* **215**, 701–711.
- DEBNATH, L. 2005 *Nonlinear Partial Differential Equations for Scientists and Engineers*. Cambridge, MA: Birkhäuser Boston, Inc.
- DEBNATH, L. 2008 *Sir James Lighthill and Modern Fluid Dynamics*. London, UK: Imperial College Press.
- DESAI, S. 2016 Do pulsar radio fluxes violate the inverse-square law? *Astrophys. Space Sci.* **361**, 138–148.
- DUDLEY, D. G. 1994 *Mathematical Foundations for Electromagnetic Theory*. New York: John Wiley & Sons, Inc. – IEEE Press.
- DUFFY, D. G. 2015 *Green’s Functions with Applications*, 2nd edn. Boca Raton: CRC Press.
- EVANS, L. C. 1998 *Partial Differential Equations*. Providence, RI: American Mathematical Society.
- GILMORE, R. 1993 *Catastrophe Theory for Scientists and Engineers*. New York, NY: Dover Publications, Inc.
- GREEN, G. 1850 An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism. Originally published as book in Nottingham, 1828. Reprinted in three parts in *Journal für die reine und angewandte Mathematik*.
- GRIFFITHS, D. J. 2015 *Introduction to Electrodynamics*, pearson new international edn. Uttar Pardesh, India: Pearson India Education Services.
- HADAMARD, J. 1964 *La théorie des équations aux dérivées partielles*. Pekin: Editions Scientifiques, (French).
- HANNAY, J. H. 1996 Bounds on fields from fast rotating sources, and others. *Proc. R. Soc. London A.* **452**, 2351–2354.
- HANNAY, J. H. 2000 Comment II on “Generation of focused, nonspherically decaying pulses of electromagnetic radiation”. *Phys. Rev. E* **62** (2), 3008–3009.
- HANNAY, J. H. 2001 Comment on “Method of handling the divergences in the radiation theory of sources that move faster than their waves”. *J. Math. Phys.* **42**, 3973–3974.
- HANNAY, J. H. 2006 Spectral and polarization characteristics of the nonspherically decaying radiation generated by polarization currents with superluminally rotating distribution patterns: comment. *J. Opt. Soc. Am. A* **23** (6), 1530–1534.
- HANNAY, J. H. 2008 Morphology of the nonspherically decaying radiation generated by a rotating superluminal source: comment. *J. Opt. Soc. Am. A* **25** (9), 2165–2166.

- HANNAY, J. H. 2009 Fundamental role of the retarded potential in the electrodynamics of superluminal sources: comment. *J. Opt. Soc. Am. A* **26** (10), 2107–2108.
- HASSANI, S. 1999 *Mathematical Physics: A Modern Introduction to Its Foundations*. New York: Springer-Verlag.
- HEWISH, A. 1996 Problems with the superluminal pulsar model. *Mon. Not. R. Astr. Soc.* **280**, 27–30.
- HEWISH, A. 2000 Comment I on “Generation of focused, nonspherically decaying pulses of electromagnetic radiation”. *Phys. Rev. E* **62** (2), 3007.
- HILBERT, D. & COURANT, R. 1931 *Methoden der mathematischen Physik*, 2nd edn. Berlin: Springer-Verlag, (German).
- IEEE STANDARDS ASSOCIATION 2013 *IEEE Standard for Definitions of Terms for Antennas*, IEEE std 145-2013 edn. (Revision of IEEE Std 145-1993).
- JACKSON, J. D. 1999 *Classical Electrodynamics*, 3rd edn. Hoboken, NJ: John Wiley & Sons, Inc.
- LIGHTHILL, M. J. 1958 *An introduction to Fourier analysis and generalized functions*. Cambridge, UK: Cambridge University Press
- MCDONALD, K. T. 2004a Radial Dependence of Radiation from a Bounded Source. www.hep.princeton.edu/~mcdonald/examples/bounded.pdf.
- MCDONALD, K. T. 2004b Synchrotron-Čerenkov Radiation. *Science* **303**, 310.
- MORSE, P. M. & FESHBACH, H. 1953 *Methods of Theoretical Physics*, , vol. I & II. New York: McGraw-Hill.
- MYINT-U, T. & DEBNATH, L. 2006 *Linear Partial Differential Equations*, 4th edn. Cambridge, MA: Birkhäuser Boston, Inc.
- SCHMIDT-ZWEIFEL, A. C. & SINGLETON, J. 2020 Preprint.
- SCHMIDT-ZWEIFEL, A. C. 2013 Terrestrial and Extraterrestrial Radiation Sources that Move Faster than Light. Master’s thesis, University of New Mexico, digitalrepository.unm.edu/math_etds/45/.
- SCHMIDT-ZWEIFEL, A. C. 2020 Theoretical and Experimental Studies of the Emission of Electromagnetic Radiation by Superluminal Polarization Currents. PhD thesis, University of New Mexico, (Anticipated Dec 2020).
- SINGLETON, J., ARDAVAN, A., ARDAVAN, H., FOPMA, J., HALLIDAY, D. & HAYES, W. 2004 Experimental demonstration of emission from a superluminal polarization current – A new class of solid-state source for MHz – THz and beyond. In *Conference Digest of the 2004 Joint 29th International Conference on Infrared and Millimeter Waves and 12th International Conference on Terahertz Electronics*. IEEE.
- STAKGOLD, I. & HOLST, M. 2011 *Green’s Functions and Boundary Value Problems*. Hoboken, NJ: John Wiley & Sons, Inc.
- STRATTON, J. A. 1941 *Electromagnetic Theory*, later printing edn. New York: McGraw-Hill.
- ZAUDERER, E. 1989 *Partial Differential Equations of Applied Mathematics*, 2nd edn. Hoboken, NJ: John Wiley & Sons, Inc.

7. Appendix: Zero-wavelength problems

We have seen in Sections 2 – 4 of this paper that the boundary term for problems of this kind is identically zero. This is a serious problem for Ardavan (2019), as the boundary term is used as primary motivation for the derivation carried out in the rest of the paper. To clarify this point, the very lengthy discourse that forms Sections 1-3 of Ardavan (2019) may be summarised as follows.

- (i) Without any experimental or theoretical evidence, Ardavan posits the existence of sources whose flux decays more slowly than the inverse square law with distance (*e.g.*, pages 2, 18).
- (ii) Combined with Green’s function (which is taken to be the same as that for a conventional source in free space), he proposes that this results in a boundary term

in the wave equation for the field that grows relative to the conventional integral with increasing distance (pages 18-19).

(iii) The growing boundary term is taken as theoretical evidence for the existence of a non-spherically-decaying component which must then be treated using the potential (pages 2, 3, 19).

Stripped of its padding, the above reasoning is fallacious both (a) because it is circular† and (b) because the effect’s primary manifestation is stated to occur in a term that is identically zero. It is therefore worth asking how Ardavan’s calculations of the non-spherical decay produce effects that appear to be finite. The answer was already mentioned briefly by Hannay (2001); we give a summary of the issue below, with more extended treatments following elsewhere (Schmidt-Zweifel 2020; Schmidt & Singleton 2020).

The emission from a superluminal polarization-current antenna may be calculated using either Equation 2.6 or Equation 4.3. If, like Ardavan, we choose the former and evaluate the potential, the integrand is seen to be the convolution of two functions: (a) \mathbf{j} which describes the density, direction, motion and frequency content of the polarization current within the antenna and (b) the Green’s function. One of the rôles of the latter is to ensure that emission times from various points in the polarization current are correctly mapped onto the desired observation time. In his calculations, Ardavan in effect separates the two terms, focusing on the relationship between retarded time and reception time‡. Whenever one reads of “bifurcation surfaces”, “cusps” or “singularities” or sees statements similar to “giving rise to a higher-order singularity” [*e.g.*, pages 3, 6, 8, 9, 11, 13, 21, 23, 27, 28, 44 – 50, 70, 71, 77, 80, 85 of Ardavan (2019)] then this procedure has occurred or is implied. Most of the mathematical techniques used at great length in Ardavan (2019) [*e.g.*, Hadamard’s method for extracting the finite part of a divergent integral, asymptotic approximations, Taylor expansions] are required to deal with the “singularities” thus generated. However, this effort is completely irrelevant in the calculation of the emission of practical antennas, because the singularities, cusps and infinities *only occur at zero wavelength*. In real life, an antenna must of course work at finite wavelength, λ .

As long as $\lambda \neq 0$, infinities are avoided, and the critical points that would give singularities at $\lambda = 0$ become merely “bright fringes” (Gilmore 1993). For finite λ , an easy but lengthy calculation shows that the broadcast flux associated with the critical points is well behaved and scales as $\lambda^{-4/9}$ (Schmidt-Zweifel 2020; Schmidt & Singleton 2020). (A similar point is made for related optical effects in texts on Catastrophe Theory [see *e.g.*, page 247 and Chapter 13 of Gilmore (1993)].)

We have shown elsewhere (Schmidt-Zweifel 2020) that had Ardavan approached the problem in the usual manner, with the effects of finite wavelength implicit from the very beginning, the solution would have been easier (no singularities), more compact (no need for the Hadamard procedure, Taylor expansions, the method of steepest descent or asymptotic solutions) and would have generated the expected result - a flux that decays with distance as the inverse square law in the far field. Moreover, such a model reproduces experimental data from real antennas quantitatively (Schmidt-Zweifel 2020).

† That is, the distance dependence of flux from non-spherically decaying sources is accepted as a premise; placing this dependence in an equation and speculating about the result then “shows” that such sources exist and need to be treated in a special way.

‡ This is often cloaked by the use of angular coordinates.

5.3 The Verdict

The main focus of the article enclosed above is to show that Ardavan’s use of the boundary term as a justification and motivation for his non-spherical decay proposals is without merit; as stated earlier, the surface integral is identically zero for problems of this kind. However, as the appendix to the paper describes, he still expends many, many pages of labyrinthine mathematics trying to derive this effect. His approach hinges on attempts to treat apparently divergent integrals.

The singularities in question result from the characteristic form of the relationship between emission and observation time for a source in superluminal rotation. In the chapter to follow, the latter is examined in great detail; we show that singularities only occur at infinitesimal points in spacetime, and that they do not present any kind of problem once finite wavelengths are introduced.

Another contentious issue in Ardavan’s latest publication [170] (briefly mentioned in one of our paper’s footnotes) is the so-called “experimental antenna” depicted. Despite his confident statements, this apparatus does not actually exist; it is an amalgam of LANL’s circular antenna TD 1 and a machine that was never built because it would have been extraordinarily inefficient. The reasons for this flaw are covered in Appendix C.2.

Chapter 6

Point Source Models: The Little Engine That Could

We demand rigidly defined areas of doubt and uncertainty!

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

6.1 Fourth Prologue

Thus far, we have – in quite some detail! – explored the mathematical foundations required for calculating the electromagnetic emission from a polarization current in superluminal motion and derived the causal fundamental solution for the problem. As was to be expected, the Green's function has revealed itself as the system's response to an electromagnetic impulse caused by an infinitesimal current element. Therefore, before embarking on simulations of the practical antennas in Chapter 7, we must explore a concept closely related to the cause-effect relationship expressed by the Green's function at the core of the solution to the wave equation: the emission of a *point-like* accelerated faster-than-light source. Most of this chapter will deal with

centripetal acceleration where the source is in faster-than-light travel on a circular trajectory. The reason for considering this geometry is both its resemblance to one of our practical antennas (TD 1, introduced in Chapter 3) and its possible relevance to emission by pulsars.

Initially the geometry of the problem is used to derive the relationship that maps retarded times onto the signal measured by an observer. As described earlier in conjunction with supersonic flight, for an accelerated superluminal source and a certain set of possible field points, disturbances from *more than one* retarded time contribute to the instantaneous, detected intensity. A very small subset of these locations receives signals emitted over *an extended time span* in a single burst of light, a concept known as “temporal focusing”. The similarities of this idea to some of the notions of catastrophe theory [29,30] as applied to optics [171,172], are described, examining focusing in the time domain for possible pathological behavior as the source-to-observer distance becomes very large¹ and deriving the exponent that describes the divergence of the focused intensity as the wavelength tends to zero.

As described in Section 2.5, pulsars are spinning neutron stars that possess very large, off-axis magnetic fields; their periods of rotation range from 1.5 ms to 8.5 s [93, 94]. A quick back-of-the-envelope calculation shows that at surprisingly small distances (85 km for the 1.5 ms pulsar; 40,000 km for the 8.5 s one) from the rotation axis, the pulsar’s magnetic field will plow through its plasma atmosphere faster than the speed of light. Recent magnetohydrodynamic simulations [99, 101] suggest that the magnetic field stirs up compact, rotating, superluminal disturbances (which is to say polarization currents) that act as sources of radiation. These regions of intense current are very small (in some cases room-sized) compared to the radius at which they orbit the neutron star (10s to 100,000s of km), and tiny compared to the source to observer (*i.e.*, pulsar to Earth) distance (kparsecs) [93,94]. The pulsar is thus a “practical antenna” (in that it emits a very characteristic signal received on

¹Plot spoiler: There is no pathological behavior!

Earth [94]) which approximates to our rotating, point-like source model. We therefore calculate the degree of temporal focusing numerically, using parameters relevant to real pulsars.

Having modeled the (quite dramatic) effects of temporal focusing, it is useful to get a more quantitative estimate of the detected radiation intensity. To do this, we investigate the relationship between the Green's function and the Liénard-Wiechert scalar potential, making modifications that, for the first time, generalize the latter to include superluminal sources. The result is used, along with a version of Kepler's problem, to predict the light curve of the point charge and derive its Stokes parameters. Both are found to correspond closely to astronomical observations.

Finally, as the staging post between pulsars and ground-based antennas, the point-like polarization current element is given a finite spatial extent and some frequency content. The resulting effects are most easily understood considering linear acceleration, and so a proof-of-concept experiment is described and then implemented utilizing the passive antenna TD 5. The results give valuable insights as to why pulsar radiation behaves as it does.

6.2 On the Anatomy of a Charge in Superluminal Rotation

6.2.1 Geometrical Considerations and the Function $h(t)$

Consider a localized charge q – such as a polarization-current element of infinitesimal volume – that rotates in the xy -plane at radius r with angular velocity ω . In terms of the cylindrical coordinates r , φ and z the path $\mathbf{r}(t)$ of q is given by

$$r = \text{const}, \quad \varphi = \varphi_0 + \omega t, \quad z = 0, \quad (6.1)$$

where the coordinate φ_0 denotes the initial azimuthal position of φ and is, without loss of generality, assumed to be zero throughout the remainder of this chapter. The wave fronts that are emitted by this point source in an empty and unbounded space can then be described by

$$|\mathbf{r}_P - \mathbf{r}| = c(t_P - t); \quad (6.2)$$

as before, the constant c denotes the wave speed and the spacetime of observation points is defined as $(\mathbf{r}_P, t_P) = (r_P, \varphi_P, z_P; t_P)$. In light of the forensics performed on

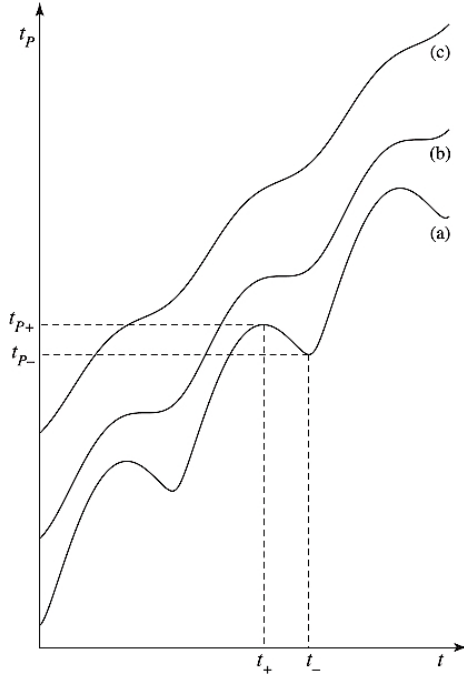


Figure 6.1: Relationship between observation and emission time for a field point P located (a) inside, (b) on the cusp of, and (c) outside the envelope of wave fronts. Here, solutions of $t_P = h(t)$ are intersections of the horizontal lines t_P with the graphs of h . (Curves (a) and (c) were originally published on p. 85 of [37], (b) in [173].) Credit: Houshang Ardavan

characteristic surfaces in Section 4.2.1 it is hardly surprising that Eq. (6.2) describes spheres of radius $c(t_P - t)$ – centered at their respective “launch points” $(r, \omega t, 0)$ – which expand at the characteristic speed *in medio*. In the words of Schott [37] (yet our notation),

If (r_P, φ_P, z_P) be regarded as a variable point, then the equation represents a sphere, whose center is at (r, φ, z) and whose radius is $c(t_P - t)$. The sphere is the position at time t_P of that wave which was emitted by the moving charge when it occupied the position (r, φ, z) , at the time t .

If t_P be regarded as a variable parameter, the equation determines a family of concentric spheres, namely the successive positions of the particular wave emitted at the time t .

If t be regarded as a variable parameter, it determines a family of spheres, whose centers lie on the path of the charge, namely the positions at any time t_P of all the waves emitted up to that time. ...

If both t_P and t be regarded as variable parameters, the characteristic equation represents a doubly infinite system of spheres, namely all the positions of all the waves emitted by the charge.

Inserting Eq. (6.1) into Eq. (6.2) and utilizing the Pythagorean theorem we find that the distance R which separates the source from an observer is given by

$$R(t) = [z_P^2 + r_P^2 + r^2 - 2rr_P \cos(\varphi_P - \omega t)]^{\frac{1}{2}} . \quad (6.3)$$

In consequence, the relationship between t and t_P must satisfy

$$\begin{aligned} t_P &= t + \frac{R(t)}{c} \\ &= t + \frac{1}{c} [z_P^2 + r_P^2 + r^2 - 2rr_P \cos(\varphi_P - \omega t)]^{\frac{1}{2}} \\ &= h(t) . \end{aligned} \quad (6.4)$$

As illustrated by Fig. 6.1, Eq. (6.4) denotes a family of functions whose members assume one of three characteristic forms given by a combination of source speed and observer position. Owing to the cosine term, h may, but need not, be monotonically increasing: If the function is oscillatory as in curve (a), the one-to-one correspondence between source and field time vanishes since multiple contributions from the domain

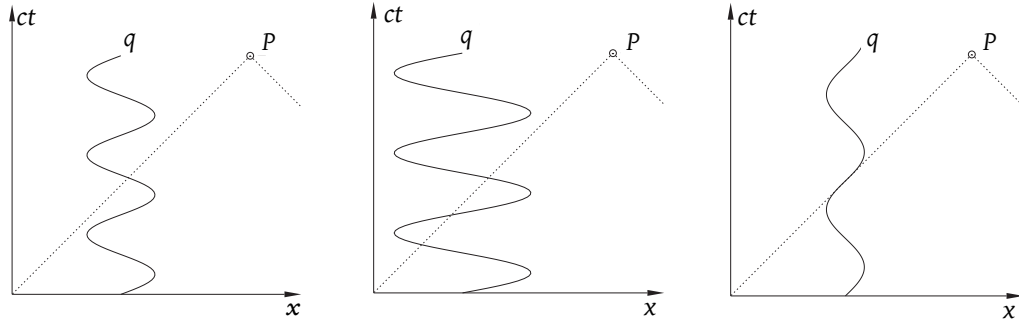


Figure 6.2: Spacetime (*i.e.*, time ct versus a projection of the motion onto the Cartesian x axis) diagrams showing the intersection between the trajectory of a charge q in superluminal rotation and the past light cone of the field point P when P lies outside (left), inside (center), and on the cusp of (right) the envelope of wave fronts. Credit: Arzhang Ardavan

of $h(t)$ may arrive in a single instant of reception². Raising the source speed leads to higher oscillation frequencies which, in turn, increases the number of wavelets³ that aggregate at an observer's location (Fig. 6.3 (left)). Of particular interest is graph (b), which represents the transition between (a) and (c), where only one image of the source is received. Here, a local maximum t_{P+} and minimum t_{P-} coalesce to form a point of inflection. As the separation between the two extrema approaches zero⁴ an extended interval of source time Δt is compressed into a much shorter period of reception δt_P . Section 6.3.1 attempts to quantify (or one might say ‘measure’) this light burst which constitutes a natural manifestation of *temporal focusing* or *focusing in the time domain*. A spacetime diagram depicting the intersection between the

²In Chapter 2 this has been identified as a hallmark of sources that exceed the characteristic wave speed. In contrast, for charges that remain at or below c there is *always* a one-to-one correspondence.

³Schott showed that the number of contributions (which is always odd) increases by two when

$$\tan \sqrt{\left(\frac{\omega}{c}\right)^2 - 1} = \sqrt{\left(\frac{\omega}{c}\right)^2 - 1}; \quad (6.5)$$

that is, for the values $(\omega/c)^2 = 1, 1 + (1.43\pi)^2, 1 + (2.45\pi)^2, \dots$

⁴Since the inflection point itself is a singularity (or “critical point”), the separation between the two extrema can – in practice – never be *exactly* zero.

flight path of a charge q in constant superluminal rotation and the past light cone of observation point P is shown in Fig. 6.2.

For the sake of mathematical (if not physical) clarity we introduce here the ‘natural’ length scale c/ω [173], which allows us to express all distances in *units of light cylinders*. In astronomy and astrophysics, the quantity $r = c/\omega$ is traditionally referred to as the *velocity of light cylinder* – or *light cylinder* in short – since it delineates the (imagined) cylindrical surface where a co-rotating extension of a spinning celestial body would have a speed equal to that of light (*i.e.*, $v/c = 1$ where $v = r\omega$ is the instantaneous linear velocity) [93, 94]. A circumflex ($\hat{}$) will denote a scaled length such that $\hat{r} = r\omega/c$, $\hat{z} = z\omega/c$, $\hat{r}_P = r_P\omega/c$ and $\hat{z}_P = z_P\omega/c$. All distances are now ‘dimensionless’⁵ and the quantity formerly labeled v is recast as $v/c = \hat{v} = \hat{r}$. This becomes evident when we convert Eq. (6.4) to ‘hatted’ quantities, namely

$$\begin{aligned} t_P &= t + \frac{\hat{R}(t)}{\omega} \\ &= t + \frac{1}{\omega} \left[\hat{z}_P^2 + \hat{r}_P^2 + \hat{r}^2 - 2\hat{r}\hat{r}_P \cos(\varphi_P - \omega t) \right]^{\frac{1}{2}} \\ &= \hat{h}(t). \end{aligned} \tag{6.6}$$

Despite their obvious benefits, units of light cylinders will be used judiciously (and sparingly) throughout the remainder of this text since they tend to complicate the interpretation of quantitative measurements.

In the parlance of catastrophe theory, “a beautiful mathematical [field] with many controversial applications” [171], $h(t)$ represents a family of smooth functions whose members may have m *critical points*⁶ ($m = 0, 1, 2, \dots$) which are all *isolated*. As stated informally in Section 4.1.2, a function f defined on an open subset $S \subset \mathbb{R}^n$ is

⁵In reality, they tally light cylinders and are, therefore, measures of velocity rather than distance. In this sense, the “unit of light cylinder” is closely related to the (equally dimensionless) Mach and Čerenkov numbers.

⁶In mathematical disciplines other than catastrophe theory, critical points are conventionally called *singular points* or, simply, *singularities*. Here, we shall use all three terms on occasion and interchangeably.

smooth if it has derivatives of arbitrary order. Eq. (6.4) and Eq. (6.6) are, therefore, smooth functions of their argument, t .

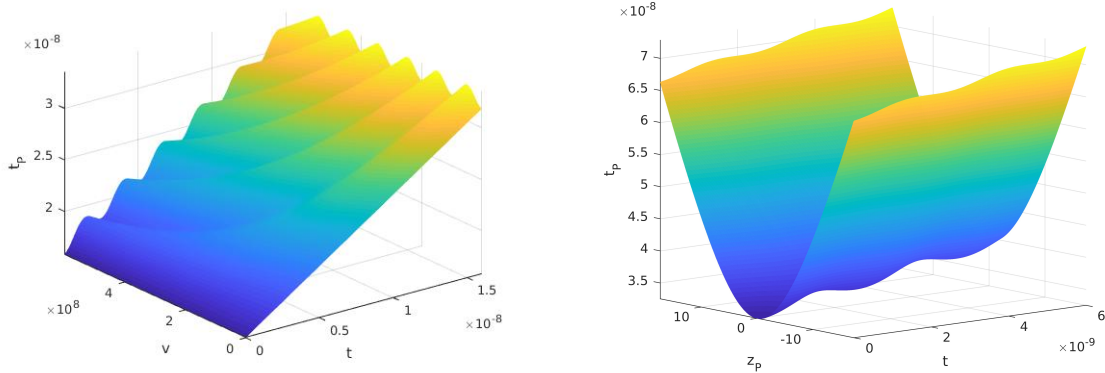


Figure 6.3: Topological representation of the interplay between emission time, t , source speed, v , observer position \mathbf{r}_P and instance of reception, t_P . (left) Higher charge velocities increase the frequency of oscillations in $h(t)$. (right) A field point's position in space-time determines how many contributions from the domain of $h(t)$ arrive simultaneously at $P(\mathbf{r}_P, t_P)$. Along the line $z_P = 0$ the function appears as in curve (a) of Fig. 6.1 and three images of the source are received. Towards the yellow edges of the sheet an extended period of source time is compressed into a single “light burst”.

A full geometric picture of the $h(t)$ set – visualized in Fig. 6.3 for independent variables v (left) and z_P (right) – emerges once the nature of its critical points has been determined as these are the *loci* where a function may change its behavior. Mathematically more precisely, a point $P \in S$ of $f : S \rightarrow \mathbb{R}^n$ is called critical if all partial derivatives $\partial f(P)/\partial x_1, \partial f(P)/\partial x_2, \dots, \partial f(P)/\partial x_n$ of f at P vanish. In other words, singularities appear in places where the gradient of f is zero⁷ and are traditionally classified as *maxima*, *minima* and *saddle points* (or *points of inflection* if $n = 1$). Moreover, P is termed *isolated* if there exists a neighborhood Ω of P in S such that no other point in Ω is critical. It is not difficult to see that, for every rotation of the source (which is moderately superluminal) curve (a) of Fig. 6.1 has

⁷In other branches of mathematics the notion of singularity may include points in the domain of f where the function is not differentiable (which implies that f is not smooth). These cases, however, are of no interest to us or catastrophe theory in general.

two isolated extrema which occur when the first time derivative of h vanishes, *i.e.*,

$$\frac{d}{dt}\widehat{h}(t) = 1 - \frac{\widehat{r}\widehat{r}_P \sin(\varphi_P - \omega t)}{\widehat{R}} = 0. \quad (6.7)$$

From (6.7) it follows immediately that

$$\widehat{R} = \widehat{r}\widehat{r}_P \sin(\varphi_P - \omega t) \quad (6.8)$$

whereas making use of the trigonometric identity $\cos^2 x = 1 - \sin^2 x$ leads to

$$\begin{aligned} \widehat{R}^2 &= \widehat{z}_P^2 + \widehat{r}_P^2 + \widehat{r}^2 - 2\widehat{r}\widehat{r}_P \cos(\varphi_P - \omega t) = \widehat{r}^2\widehat{r}_P^2 \sin^2(\varphi_P - \omega t) \\ &= \widehat{r}^2\widehat{r}_P^2 - \widehat{r}^2\widehat{r}_P^2 \cos^2(\varphi_P - \omega t) \end{aligned} \quad (6.9)$$

or

$$\widehat{r}^2\widehat{r}_P^2 \cos^2(\varphi_P - \omega t) - 2\widehat{r}\widehat{r}_P \cos(\varphi_P - \omega t) - \widehat{r}^2\widehat{r}_P^2 + \widehat{z}_P^2 + \widehat{r}_P^2 + \widehat{r}^2 = 0. \quad (6.10)$$

Recognizing that Eq. (6.10) is quadratic provides a direct solution for $\cos(\varphi_P - \omega t)$, *i.e.*,

$$\begin{aligned} \cos(\varphi_P - \omega t) &= \frac{1}{\widehat{r}\widehat{r}_P} \pm \frac{[\widehat{r}^2\widehat{r}_P^2 - \widehat{r}^2\widehat{r}_P^2(-\widehat{r}^2\widehat{r}_P^2 + \widehat{z}_P^2 + \widehat{r}_P^2 + \widehat{r}^2)]^{\frac{1}{2}}}{\widehat{r}^2\widehat{r}_P^2} \\ &= \frac{1}{\widehat{r}\widehat{r}_P} \left(1 \pm [1 + \widehat{r}^2\widehat{r}_P^2 - \widehat{z}_P^2 - \widehat{r}_P^2 - \widehat{r}^2]^{\frac{1}{2}} \right) \\ &= \frac{1 \pm \Delta^{\frac{1}{2}}}{\widehat{r}\widehat{r}_P} \end{aligned} \quad (6.11)$$

where

$$\Delta \equiv (\widehat{r}^2 - 1)(\widehat{r}_P^2 - 1) - \widehat{z}_P^2. \quad (6.12)$$

It is elementary to see that Eq. (6.11) implies

$$\varphi_P - \omega t = \arccos \left[\frac{1 \pm \Delta^{\frac{1}{2}}}{\widehat{r}\widehat{r}_P} \right] \quad (6.13)$$

such that the two solutions⁸ of Eq. (6.7) are

$$t_{\pm} = \frac{1}{\omega} \left[2\pi + \varphi_P - \arccos \left(\frac{1 \mp \Delta^{\frac{1}{2}}}{\widehat{r}\widehat{r}_P} \right) \right]. \quad (6.14)$$

⁸Strictly speaking, Eq. (6.7) has a doubly infinite set of solutions $t = t_{\pm} + m\pi$, where $m \in \mathbb{Z}$. However, we aim to restrict the present discussion to a single rotation of the charge.

The subscripts $+$ and $-$ denote emission times which result in a local maximum t_{P+} and minimum t_{P-} , respectively. The behavior of t_{\pm} as a function of \hat{z}_P and the composition $\hat{h} \circ t_{\pm} = \hat{h}(t_{\pm}(\hat{z}_P))$ are graphed in Fig. 6.4, showing that the separation between t_+ (blue) and t_- (red) decreases until $t_+ = t_-$ where the two functions coalesce. The corresponding point of inflection $t_{P+} = t_{P-}$ (depicted in curve (b) of

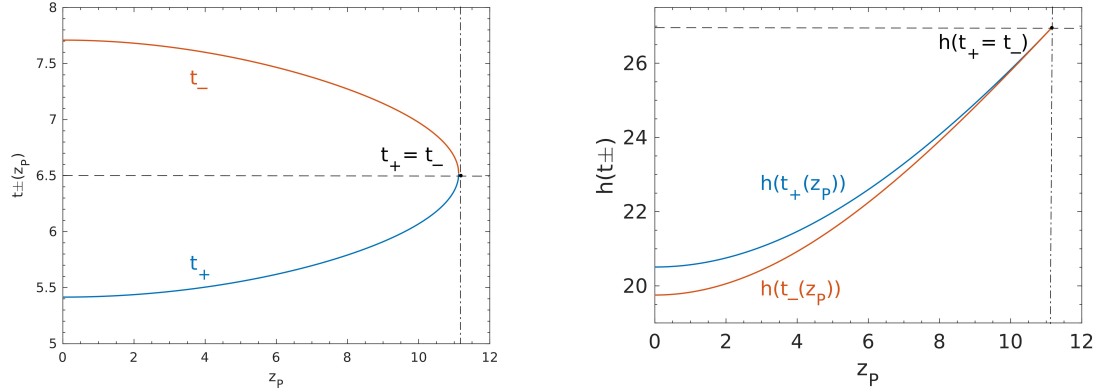


Figure 6.4: (left) t_+ and t_- as functions of \hat{z}_P , where the subscripts $+$ and $-$ denote emission times which result in a local maximum t_{P+} and minimum t_{P-} , respectively. The separation between the two curves decreases and vanishes for $\hat{z}_P \approx 11.12$ where $\hat{h}(t)$ has a point of inflection. (right) The composition $\hat{h}(t_{\pm}(\hat{z}_P)) = \left[2\pi - \arccos((1 \mp \Delta^{1/2})/\hat{r}\hat{r}_P) + [\hat{z}_P^2 + \hat{r}_P^2 + \hat{r}^2 - 2(1 \mp \Delta^{1/2})]^{1/2} \right] / \omega$ displays behavior similar to that of a fold catastrophe [29, 30, 171, 172].

Fig. 6.1) is *degenerate* since both the first and second time derivative of \hat{h} vanish rendering the Hessian trivially zero. To find $\hat{h}(t_+ = t_-) = \hat{h}(t_{\sim})$, we may proceed as above, which is to say

$$\frac{d^2}{dt^2} \hat{h}(t) = \omega \left[-\frac{\hat{r}\hat{r}_P \cos(\varphi_P - \omega t)}{\hat{R}} + \frac{\hat{r}^2 \hat{r}_P^2 \sin^2(\varphi_P - \omega t)}{\hat{R}^3} \right] = 0. \quad (6.15)$$

However, it is more expedient to set t_+ and t_- equal to each other, solve the resulting expression for \hat{z}_P and use the latter to calculate $\hat{h}(t_{\sim})$. Mathematically more explicitly,

$$\hat{z}_{P\sim} = (\hat{r}^2 - 1)^{\frac{1}{2}} (\hat{r}_P^2 - 1)^{\frac{1}{2}} \quad (6.16)$$

which entails

$$\begin{aligned}
 \Delta_{\sim} &= (\hat{r}^2 - 1)(\hat{r}_P^2 - 1) - \hat{z}_{P\sim}^2 \\
 &= (\hat{r}^2 - 1)(\hat{r}_P^2 - 1) - (\hat{r}^2 - 1)(\hat{r}_P^2 - 1) \\
 &= 0
 \end{aligned} \tag{6.17}$$

and, consequently,

$$\begin{aligned}
 \hat{h}(t_{\sim}) &= \frac{1}{\omega} \left[2\pi - \arccos\left(\frac{1}{\hat{r}\hat{r}_P}\right) + [\hat{r}^2\hat{r}_P^2 - \hat{r}^2 - \hat{r}_P^2 + 1 + \hat{r}^2 + \hat{r}_P^2 - 2]^{\frac{1}{2}} \right] \\
 &= \frac{1}{\omega} \left[2\pi - \arccos\left(\frac{1}{\hat{r}\hat{r}_P}\right) + [\hat{r}^2\hat{r}_P^2 - 1]^{\frac{1}{2}} \right].
 \end{aligned} \tag{6.18}$$

We now have derived all of the equations to permit three-dimensional spatial mapping of the regions corresponding to curves (a), (b), and (c) in Fig. 6.1, and more importantly, the associated critical points. This will be carried out in the following section; subsequently, the detailed behavior of $h(t)$ close to the point of inflection in curve (b) will be examined and compared to catastrophe theory as applied to optics.

6.2.2 Graphing $h(t)$ in Three Dimensions

For the sake of clarity, it may be beneficial to depict the spatial distribution of the various members of the $h(t)$ family shown in Fig. 6.1 first (Fig. 6.5 and Fig. 6.6); the details of how these curves and surfaces are calculated then follows. However, there is – once again! – the vexed question of nomenclature. We shall refer to the surfaces that separate the different behaviors of $h(t)$ as *sheets* (T_{\pm} in Fig. 6.5 and the red and blue structures in Fig. 6.6). The *locus* corresponding to $\hat{h}(t_{\sim})$ will be named “the cusp”. As we shall see below, it has some of the traits of, *but is not exactly the same as*, Thom’s canonical *cusp catastrophe* [171, 172]. Nevertheless, this point, at which two sheets meet, is “cusp-like” from a geometrical descriptive standpoint, and the term was used by Schott [37] (predating catastrophe theory by many decades) to describe similar structures in the wavelets emitted by a faster-than-light source.

Previous papers by others (*e.g.*, [75]) have also referred to this feature as a “cusp” and so to avoid confusion, we will retain the same name.

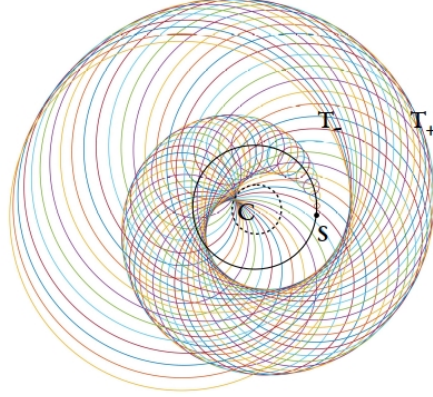


Figure 6.5: Cusp (C) and envelope (T_{\pm}) of the wave fronts emitted by a point charge (S) in constant superluminal rotation. The bold curves show the cross section of the envelope with the rotational plane of the source. The larger of the two black circles marks the orbit (at $r = 2.0 c/\omega$), the smaller one the velocity of light cylinder ($r_P = c/\omega$).

Fig. 6.5 shows a plan view of emitted Huygens’ wavelets and the sheets in the plane of rotation of the source. The emitted wave fronts form a tube-like structure analogous to the Čerenkov envelope described in Chapter 2; however, owing to the centripetal acceleration of the source, the sheets are curved and fold inward to meet on the light cylinder at the cusp. Fig. 6.6 is a three-dimensional representation of part of the two-sheeted envelope.

The expression for T_{\pm} , the temporal relation that determines the position of the two sheets of the envelope, can be found by inserting $t = t_{\pm}$ into Eq. (6.14),

$$T_{\pm} \equiv \hat{h}(t_{\pm}(\hat{z}_P)) = \frac{1}{\omega} \left[2\pi - \arccos\left(\frac{1 \mp \Delta^{1/2}}{\hat{r}\hat{r}_P}\right) + [\hat{z}_P^2 + \hat{r}_P^2 + \hat{r}^2 - 2(1 \mp \Delta^{1/2})]^{1/2} \right] \quad (6.19)$$

in which all quantities are defined as above. The cusp is located where the two sheets meet tangentially (see Fig. 6.5, the left pane of Fig. 6.7 and Eq. (6.16)); it represents the inflection point of curve (b) in Fig. 6.1, and in a subsequent section,

will be shown to correspond to radiation emitted by the source at a point for which $dR/dt = -c$ and $d^2R/dt^2 = 0$. As shown in Fig. 6.5, it kisses the light cylinder in the source's plane of rotation; above and below (Fig. 6.7, right), it spirals outward, defining a *locus* reminiscent of an old-fashioned bed spring, approaching a cone of half angle $\arcsin(c/r\omega)$ as the vertical distance becomes very large. The path of the cusp rotates rigidly with the source, and an observer sitting on the double-funnel-shaped surface that this sweeps out measures the effects of the passing cusp with a frequency $\omega/2\pi$. Mathematically, T_\sim , which indicates the location of the cusp, is defined by a combination of Eq. (6.16) and Eq. (6.18).

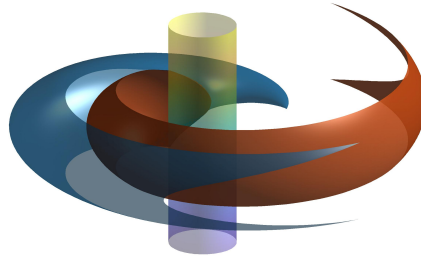


Figure 6.6: Development of the two-sheeted envelope in three dimensions. Here, the velocity of the source is $r\omega = 2.5c$.

6.2.3 Relationship of $h(t)$ to Catastrophe Theory and Optics

As noted in the caption to Fig. 6.4, the transition from curve (a) to curve (c) via curve (b) in Fig. 6.1 is very reminiscent of the fold catastrophe, important in catastrophe theory as applied, for example, to simple mechanical components responding to deformation [29, 30, 171, 172]. Take, for instance, a beam under stress; for a small

force, the beam will bend in a controlled and reversible manner, but eventually, a straw will break the camel’s back.

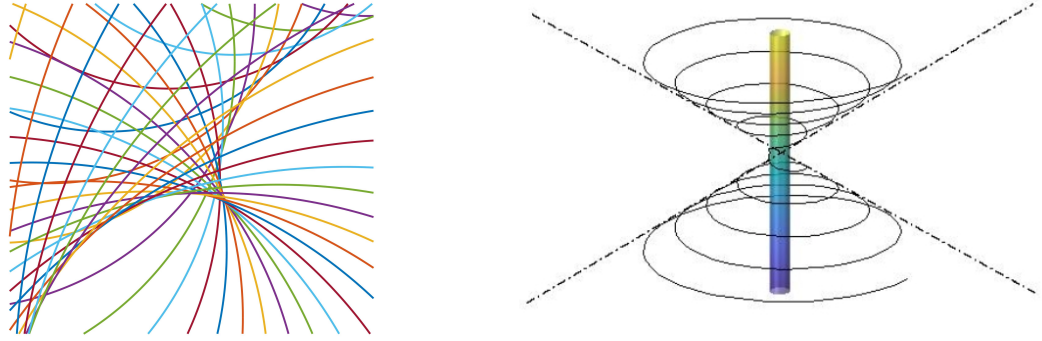


Figure 6.7: The cusp “catastrophe”. (left) In two dimensions, the region around the cusp – the single point where the two sheets of the envelope meet – is almost identical to the caustic that forms when parallel rays are reflected from a curved surface. (See, for example, Figure 13.3 in [30].) (right) If t , \hat{r}_P , $\hat{\varphi}_P$ and \hat{z}_P are all variable parameters, the cusp is a curve that touches the light cylinder before spiraling out above and below the plane of rotation, approaching the cone of polar angle $\arcsin c/r\omega$ in the far field. This expanding spiral shape may be derived analytically from Eq. (6.18) and numerically from Eq. (6.48).

Using such a *mechanical* analogy, we might call the various instances of $h(t)$ potential or harmonic functions and the horizontal axis a spatial displacement or force. In these terms, the minimum of curve (a) represents a stable equilibrium, whilst the other critical point – the maximum – is unstable. As long as the minimum persists, the mechanical system is considered balanced since it can rest in the stable equilibrium configuration. However, a control parameter may be introduced that alters the form of the potential function. As both change, the stable and unstable extrema approach and ultimately annihilate each other in a point of inflection (*c.f.*, curve (b)). The catastrophe occurs exactly at this “tipping” or “bifurcation point” [171,172] and thereafter, for further changes in the control parameter in the same direction, the mechanical system always fails (*i.e.*, has no stability)⁹.

⁹This nomenclature is why some previous treatments of rotating superluminal sources

In Fig. 6.1, the transitions between the various curves are driven by observer position. Curve (a) persists whilst $\hat{z}_P < (\hat{r}^2 - 1)^{1/2}(\hat{r}_P^2 - 1)^{1/2}$; when \hat{z}_P is precisely $(\hat{r}^2 - 1)^{1/2}(\hat{r}_P^2 - 1)^{1/2}$, the maximum and minimum combine to form the inflection point of curve (b). Having pointed out the resemblance of Fig. 6.1 to a classic problem in catastrophe theory, it is natural to ask whether the behavior of $h(t)$ follows one of Thom's canonical forms of catastrophe [171, 172]. Perhaps more importantly, the implications for measurements and calculations of a catastrophe that occurs in time, rather than energy, should be considered.

Although the behavior of $h(t)$ looks suspiciously like that of a fold catastrophe, it depends on more than one control parameter, unlike Thom's canonical fold [171, 172]. This should already be evident in the conditions given in the previous paragraph, which depend on both radial and vertical distance between the source and observer. Fig. 6.8 gives an illustrative summary of the spatial regions associated with curves (a), (b) and (c). With the observer positioned in the plane of rotation of the source (Fig. 6.8 (left)), the monotonic increase of $h(t)$ with no critical points will be observed within the light cylinder. Outside the light cylinder, $h(t)$ will exhibit oscillatory behavior [two critical points]; the transition between the two occurs exactly *on* the light cylinder, with the pointed nose of the inner envelope corresponding to the doubly degenerate critical point.

Fig. 6.8 (right) shows the boundary – composed entirely of singularities – that separates members of the $h(t)$ family without critical points from those with two as z_P varies. As mentioned in the discussion of Fig. 6.7, at large z_P , this sheet tends towards a cone with half angle $\sin^{-1}(c/v)$; as $z_P \rightarrow 0$, the boundary settles on the light cylinder (*c.f.*, Fig. 6.7, right-hand side), such that the overall shape is that of two conjoined funnels. The doubly degenerate critical point will be observed on the surface, hitting the observer with the period of rotation of the source, $2\pi/\omega$. Starting inside the funnel, an observer can therefore traverse from a function with no critical

have invoked so-called “bifurcation surfaces” [75].

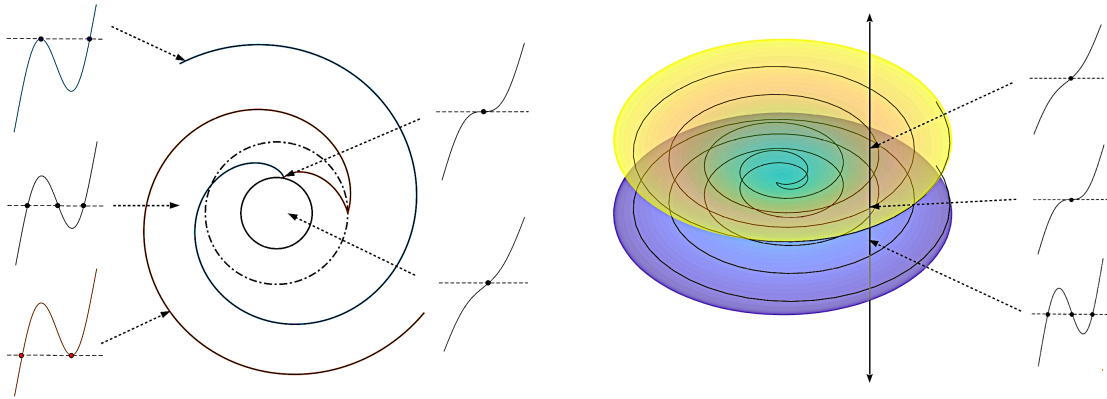


Figure 6.8: Illustrations of the regions in which $h(t)$ has no, one or two critical points. (left) the plane of rotation of the source. The light cylinder is indicated by a solid black circle, the path of the source by a black circle and the sheets by blue and red curves (*c.f.* Fig. 6.5). The form of $h(t)$ for each area is shown by the curves plotted around the figure, with arrows linking each curve to the place in which it would be observed. (right) The “double-funnel” structure (yellow and purple surfaces) sketched out by the spiraling path of the cusp (fine black curves). Again, the form of $h(t)$ for each region is shown by the curves plotted around the figure, with arrows linking each curve to the place in which it would be observed. The black vertical line illustrates the second observer path described in the text.

points to one with two (*i.e.*, move from inside to outside) by following a radial path (increasing r_P) or, if his/her initial $r_P > c/\omega$, by staying at constant r_P and changing z_P . As long as r_P is finite, continuing along the latter path will eventually cause the observer to re-enter the funnel. Of course, any combination of these two tactics will allow for a traversal of the doubly-degenerate critical point.

There are therefore two control parameters (r_P and z_P), so that the catastrophe has some of the attributes of a canonical cusp [171, 172]. However, the path of the doubly degenerate critical point maps out a three-dimensional object (the double funnel), reminiscent of the behavior of a swallowtail catastrophe [171, 172].

It will now be clear that our catastrophe does not follow any of Thom’s canonical forms, but is something of an intermediate case lying between the examples mentioned above. The interesting question from an electrical engineering or optics standpoint is the degree of the singularity [30], or one might say, seriousness, of the

catastrophe. This can be expressed as the rate at which the intensity I diverges as the wavelength λ tends to zero. Gilmore [30] (p. 330 ff.) shows that the neighbor-

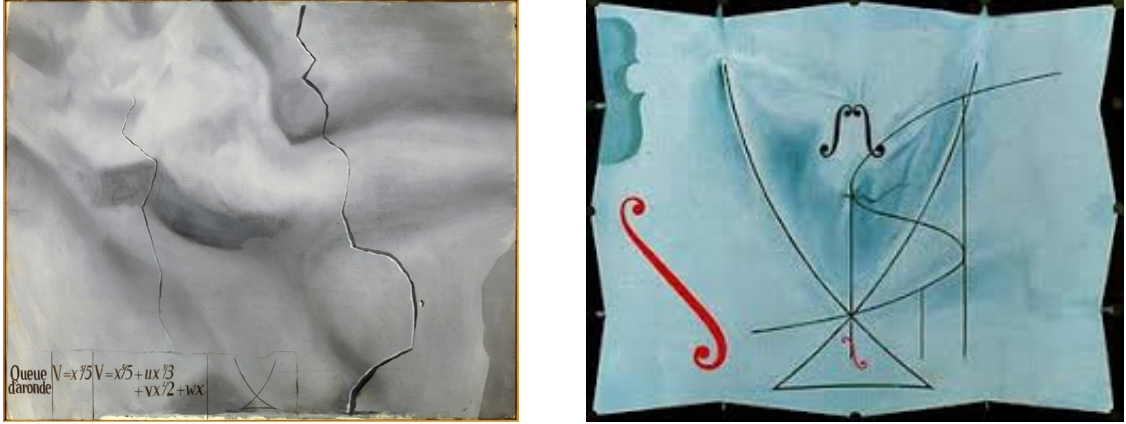


Figure 6.9: Two paintings by Salvador Dalí inspired by Thom's canonical catastrophes. (left) *El rapte topològic d'Europa*. Homenatge a René Thom (Topological Abduction of Europe. Homage to René Thom) depicts the definition of the swallowtail ($V = x^5/5 + ux^3/3 + vx^2/2 + wx$) in the lower left corner. (right) *La queue d'aronde* — *Série des catastrophes* (The swallowtail — Series of catastrophes) was Salvador Dalí's last painting. Completed in May 1983, it shows the fold, cusp and swallowtail catastrophes. Credit: <https://www.salvador-dali.org>

hood of a degenerate critical point gives an approximate intensity $I \propto \lambda^{-\sigma}$ in the limit $\lambda \rightarrow 0$ and lists values of the exponent σ for some of the simple canonical catastrophes. For example, the fold yields $\sigma = 1/6$, the cusp $\sigma = 1/3$ and the swallow tail $\sigma = 4/9$; as the degree of the catastrophe increases, σ asymptotically approaches $1/2$. As long as $\lambda \neq 0$, infinite intensity is avoided, and our doubly degenerate critical point will yield what Gilmore [30] terms “a bright fringe”. A semi-empirical estimate of $\sigma = 4/9$ for our case will be given below. Such behavior is not unusual, especially in the context of wave theory or optics, in that families of functions often exhibit focusing [12, 30]. This is not a property of individual paths or trajectories described by the functions, but occurs when neighboring trajectories make contact or converge, as happens in curves (a) and (b) of Fig. 6.1. In general, the regions in which there is such a focus are referred to as “caustics”. In the limit of zero wavelength, the density of trajectories will be infinite on a caustic, corresponding to infinite electromagnetic

fields or intensities [30].

Moving onto more realistic situations, the analogous zero-wavelength approach of geometrical optics gives reasonable predictions for beam trajectories in the limit of short wavelength [12]. In other words, for the latter, the positions of *foci*, bright fringes and other observables measured in an optical system will be very close to those predicted by the caustics of the zero-wavelength calculations or geometrical constructions [12]. Though no longer infinite, the measured fields can be intense [30], as anyone who has witnessed the explosive dielectric breakdown of air caused by the focusing of a Class IV laser beam will testify. Therefore, unlike Huygens wavelets or rays - useful constructs but with no physical reality – one could say that the effects of a caustic are real and measurable, as when we burn a piece of paper using the Sun’s light focused through a lens.

Nevertheless, it is precisely on caustics that predictions made using the trajectories of zero-wavelength calculations will fail. As long as the length of the waves is not exactly zero, solutions based on the wave equation produce finite intensities, never infinite ones. At best, we can say that for various classes of caustic, careful trajectory-based models can be derived that give asymptotic approximations to wave fields in the short-wavelength limit. All of this should be borne in mind in the following sections, where we examine the details of the caustic associated with the bifurcation point shown in curve (b) of Fig. 6.1.

6.2.4 In the Far-field

The foregoing discussion has introduced the idea of caustics or *foci*, taken originally from geometrical optics. This represents a more formal description of a phenomenon already mentioned qualitatively in Chapter 2; there we stated that one of the primary motivations for studying accelerated superluminal polarization currents is the possibility of combining signals from several different emission times such that they

arrive at an observation point within a very short time interval. In effect, Huygens wavelets are focused in the time domain and the term “temporal focusing” was coined to describe this effect.

The function $h(t)$ plotted in Fig. 6.1 gives a gauge for the degree of temporal focusing (or, one might say, of the presumed *practical* importance of the event). If we consider radiation emitted by the source over a short interval Δt , then it will be detected by an observer over a time interval δt_P , where

$$\delta t_P \approx \left(\frac{d}{dt} t_P \right) \Delta t, \quad (6.20)$$

and we may introduce a ‘temporal focusing factor’ given by

$$\frac{\Delta t}{\delta t_P} \approx \frac{d}{dt} t = \left(\frac{d}{dt} h(t) \right)^{-1}. \quad (6.21)$$

This quantity will become very large when $dh(t)/dt \rightarrow 0$, which forces us to consider the detailed behavior of curve (b) in Fig. 6.1 at, or close to, the bifurcation point (or point of inflection). We know that for $dh(t)/dt = 0$, Eq. (6.21) is infinite. However, this occurs at a *single point of infinitesimal length*, and thus is likely to be of minimal influence once the temporal ‘fuzziness’ caused by considering finite wavelengths is introduced (we will return to this claim in the sections to follow). Of more immediate importance is the region of t *around* the bifurcation point since $(dh(t)/dt)^{-1}$ will likely grow quickly over an extended range of t when the separation between source and observer increases. As will be described later, a quantitative investigation is relevant to astronomy; as we have seen in Chapter 2, the model of a charge in superluminal rotation is a possible description of pulsars, which are at very large distances (kparsecs) from Earth. Accordingly, it is apposite to ask whether some ‘ultra-far-field’ effect contributes to the brightness of pulsar emission as observed by astronomers. To answer this and other questions, we use metrical, rather than dimensionless, variables.

The inflection point of $h(t)$ is defined by

$$\frac{d\hat{h}}{dt} = \frac{d^2\hat{h}}{dt^2} = 0, \quad (6.22)$$

from which it follows that

$$\frac{dh}{dt} = \frac{d^2h}{dt^2} = 0. \quad (6.23)$$

Using Eq. (6.4), which relates $h(t)$ to $R(t)$, the distance between source and receiver, we obtain

$$\frac{dh}{dt} = 1 + \frac{1}{c} \frac{dR(t)}{dt} = 0 \quad (6.24)$$

and

$$\frac{d^2h}{dt^2} = \frac{d^2R(t)}{dt^2} = 0. \quad (6.25)$$

These conditions have an important geometrical consequence. Eq. (6.24) shows that at the tipping point,

$$\frac{dR}{dt} = -c, \quad (6.26)$$

which is to say that the source was approaching the observer at the speed of light and without acceleration when the radiation was emitted. As before (see discussion of Eq. (6.14)), we label the time at which the first and second derivatives of h with respect to t are zero as t_{\sim} , and the corresponding retarded distance as R_{\sim} .

Squaring Eq. (6.3), we obtain

$$R^2(t) = z_P^2 + r_P^2 + r^2 - 2rr_P \cos(\varphi_P - \omega t) \quad (6.27)$$

whereas subsequent differentiation yields

$$2R \frac{dR}{dt} = -2\omega r r_P \sin(\varphi_P - \omega t). \quad (6.28)$$

After dividing through by 2 and differentiating (yet) again, we have

$$\left(\frac{dR}{dt} \right)^2 + R \frac{d^2R}{dt^2} = \omega^2 r r_P \cos(\varphi_P - \omega t). \quad (6.29)$$

Exactly at the bifurcation point, we may substitute the results of Eqs. (6.25) and (6.26) such that the two differentials of R yield

$$\sin(\varphi_P - \omega t_\sim) = \frac{cR_\sim}{\omega r r_P} \quad \text{and} \quad \cos(\varphi_P - \omega t_\sim) = \frac{c^2}{\omega^2 r r_P}. \quad (6.30)$$

We now use these identities to examine how

$$\frac{dh}{dt} = 1 + \frac{1}{c} \frac{dR(t)}{dt} = 1 - \frac{\omega r r_P}{cR} \sin(\varphi_P - \omega t) \quad (6.31)$$

varies close to¹⁰ $t = t_\sim$. Consider a variation in time around the point of inflection such that $t = t_\sim + \delta_t$. Substituting into Eq. (6.31), we obtain

$$\frac{dh}{dt} = 1 - \frac{\omega r r_P}{cR} \sin(\varphi_P - \omega(t_\sim + \delta_t)) \quad (6.32)$$

$$= 1 - \frac{\omega r r_P}{cR} [\sin(\varphi_P - \omega t_\sim) \cos(\omega \delta_t) - \cos(\varphi_P - \omega t_\sim) \sin(\omega \delta_t)] . \quad (6.33)$$

while using the relationships given in Eq. (6.30) leads to

$$\frac{dh}{dt} = 1 - \frac{R_\sim}{R} \cos(\omega \delta_t) + \frac{c}{\omega R} \sin(\omega \delta_t) . \quad (6.34)$$

We now consider $\delta_t \ll 2\pi/\omega$ in order to use small angle expressions for sine and cosine, and also large source-to-observer distances such that $R \gg r$. Under such constraints, to leading-order terms in small quantities,

$$\frac{dh}{dt} \approx \frac{(\omega \delta_t)^2}{2} + \frac{c \delta_t}{R} . \quad (6.35)$$

A precursory glance suggests that the rightmost term dominates and $dh/dt \propto 1/R$, leading to an ever-flatter curve (and hence greater focusing in the time domain) as R gets larger. However, this is only true when

$$\frac{\omega^2 \delta_t^2}{2} < \frac{c \delta_t}{R} , \quad (6.36)$$

that is, whilst

$$\omega \delta_t < \frac{2c}{R\omega} = \frac{2}{\widehat{R}} , \quad (6.37)$$

¹⁰In the final term, we have substituted for dR/dt from Eq. (6.28).

i.e., less than 2 divided by the source to observer distance in dimensionless units of light cylinders¹¹. In other words, although there *is* a time interval very close to the bifurcation point over which $(dh/dt)^{-1} \propto R$, the fraction of the orbit that this interval represents is proportional to $1/R$, becoming ever smaller as R increases, thus cancelling out any possible enhancement of the signal due to this effect at large distances. This finding will be confirmed by numerical calculations of the temporal focusing factor later in this chapter.

6.2.5 Wavelength Scaling Behavior of the $h(t)$ “Catastrophe”

Rather than looking at point-to-point variations in the gradient of $h(t)$, another way of estimating the degree of temporal focusing is to consider the average enhancement over a small interval of time close to the point of inflection. This is most easily accomplished by expanding $h(t)$ as a Taylor series about $t = t_{\sim}$. As the first and second derivatives of $h(t)$ are identically zero at $t = t_{\sim}$, we obtain

$$h(t) = h(t_{\sim}) + \frac{1}{3!}(t - t_{\sim})^3 \left(\frac{d^3 h}{dt^3} \right)_{(t=t_{\sim})} + \dots, \quad (6.38)$$

i.e., the leading-order term in $(t - t_{\sim})$ is the cube. For small deviations of t away from the bifurcation, this term will dominate. Rearranging gives

$$\delta t_P = h(t) - h(t_{\sim}) \frac{1}{3!}(t_{\sim} + \delta t)^3 \left(\frac{d^3 h}{dt^3} \right)_{(t=t_{\sim})}. \quad (6.39)$$

We now differentiate Eq. (6.29) with respect to time, remembering that $d^2 R/dt^2 = 0$ at $t = t_{\sim}$, and using Eq. (6.24) and Eq. (6.30) such that

$$\left(\frac{d^3 h}{dt^3} \right)_{(t=t_{\sim})} = \frac{1}{c} \left(\frac{d^3 R}{dt^3} \right)_{(t=t_{\sim})} = \omega^2. \quad (6.40)$$

Writing $\delta t = t - t_{\sim}$ and rearranging, the “temporal-focusing factor” becomes

$$\frac{\delta t}{\delta t_P} = \frac{6}{\omega^2(\delta t)^2}. \quad (6.41)$$

¹¹In the final section of this chapter, this number will be shown to be very small indeed, even for very low-frequency pulsars.

Given the discussion of the previous section, it is reassuring that this result is independent of both the frequency of the orbit and the distance of the source from the observer; it depends *only on the dimensionless fraction $\omega\delta t$ of the orbit over which the emission occurs*. The smaller this fraction is, the greater the degree of temporal focusing. Alternatively, the expression may be rearranged as follows;

$$\frac{\delta t}{\delta t_P} = \frac{6^{\frac{1}{3}}}{(\omega\delta t_P)^{\frac{2}{3}}}, \quad (6.42)$$

showing that *a shorter observation time will yield a greater enhancement*. Again, because any phenomenon measured by the observer must repeat with the angular frequency of rotation of the source ω , the result is independent of ω as well as the distance of the rotating source from the observer; it only depends on the dimensionless fraction $\omega\delta t_P$ of the orbit over which the observation is made.

Thus far, the results in this and the previous section have been completely independent of the size and rotation frequency of the source. However, let us for a moment consider the point-like rotating source oscillating at a frequency ν . For this frequency to be recognisable as a component of the emitted electromagnetic radiation, the oscillation must persist for an interval of observer time corresponding to at least a small number of periods, that is for a time interval $\delta t_P \sim n/\nu$, where n is the small number. Substituting into Eq. (6.42) yields

$$\frac{\delta t}{\delta t_P} \sim \left(\frac{c}{\omega\lambda} \right)^{\frac{2}{3}}, \quad (6.43)$$

where $\lambda = c/\nu$ is the wavelength of the emitted radiation. Electromagnetic fields superpose linearly, and so Eq. (6.43) represents the way in which the electric or magnetic field is enhanced close to the point of inflection as $\lambda \rightarrow 0$. To permit a comparison with the canonical catastrophes, we require the scaling behavior of the *intensity*. Squaring Eq. (6.43) results in an intensity scaling of $1/\lambda^{\frac{4}{9}}$, the behavior of a canonical swallow's tail [29, 30, 171, 172]. As has been mentioned before, a true crisis (infinity) only occurs at $\lambda = 0$, that is, *for zero wavelength*. Just as in geometrical optics, the caustic or focus only becomes a singularity in that limit; once

the wavelength is finite, we have emphatic, but not excessive, effects [30]. We must now examine where, relative to the source, the latter will occur.

6.3 All About Temporal Focusing

6.3.1 General Idea

The object of this section is to evaluate the degree of temporal focusing produced by a small, point-like source in superluminal rotation numerically. In order to make quantitative predictions, an orbit radius and angular frequency relevant to an astronomical object - a pulsar - are chosen (Section 2.5). The reason for this selection is that – though controversial - the superluminal model of pulsar emission has been remarkably successful in predicting the observed radiation spectra of pulsars over many orders of magnitude of frequency using only a small number of adjustable parameters [79–81].

Taking up the story from Section 2.5, relativistic hydrodynamic calculations show that the active emitting region is likely to be very small, many orders of magnitude smaller than the orbit radius or circumference. In addition, the observation point (Earth) is very far away from the source. Therefore a far-field model that invokes a point-like source in superluminal rotation has a good chance of capturing some of the essential features of pulsar observations.

6.3.2 Definition of a Practical “Temporal Focusing Factor”

Fig. 6.10 replots two of the curves from Fig. 6.1; to determine the length of source time corresponding to a given reception period, we move a “window” along $h(t)$. The window’s height is chosen to be one nanosecond, a period of reception time that allows

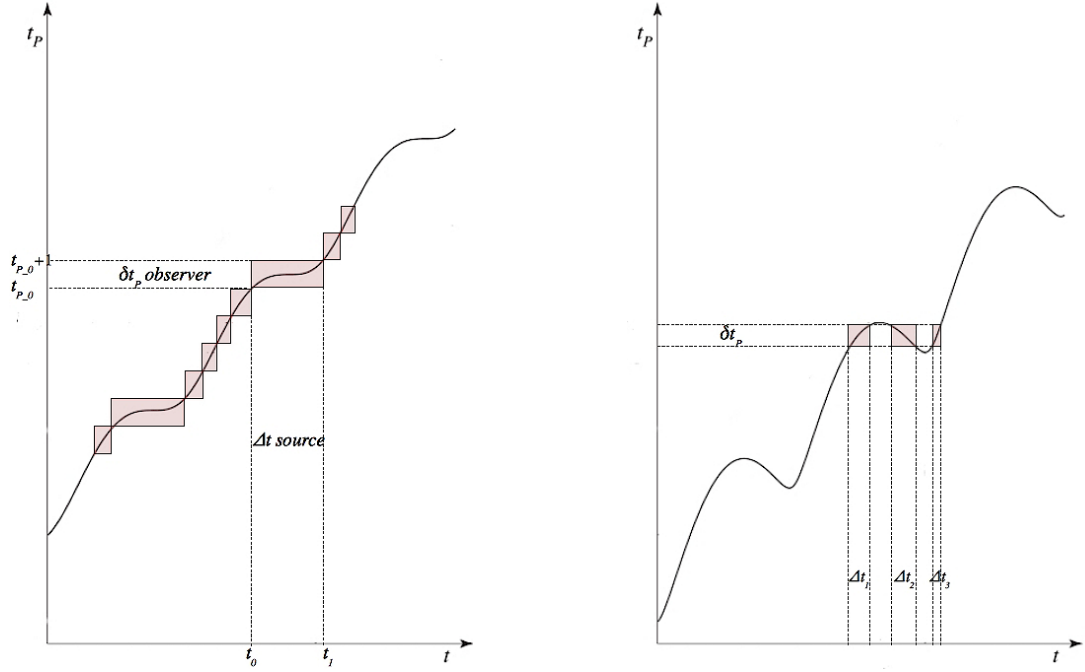


Figure 6.10: Left: The curve $t_P = h(t)$ (see Eq. (6.4)) for an observation point on, or very close to, the cusp, *e.g.*, $\theta_P = \theta_{P_C}$. The reception time window δt_P (observer) slides along the curve, and the corresponding limits in the source's time frame, t_0 and t_1 , are found. The degree of temporal focusing is $\Delta t / \delta t_P$, where $\Delta t = t_1 - t_0$. Right: The curve $t_P = h(t)$ for the case $\theta_{P_C} < \theta_P < 180^\circ - \theta_{P_C}$. Here, there are three windows of Δt corresponding to δt_P . To assess the degree of temporal focusing, $\Delta t / \delta t_P$ is again used, but with $\Delta t = \Delta t_1 + \Delta t_2 + \Delta t_3$.

realistic comparisons with the radiation received from pulsars, since it is around the minimum time interval resolvable in much current astronomical instrumentation. The width of the window is determined by the intersections of the horizontal lines t_{P_0} and $t_{P_0} + 1 \text{ ns}$ with h .

Calculating the degree of temporal focusing of $h(t)$ is mathematically reminiscent of taking the Lebesgue measure, which is usually introduced as part of the modern theory of integration [179]. In the Lebesgue scheme, instead of partitioning the domain of a function f , the range of f is subdivided, *i.e.*, $\min f \geq y_0 < y_1 < \dots < y_n \geq \max f$, to form the sum $\sum_{i=1}^n y_{i-1} \cdot \text{measure}(\{x | f(x) \in [y_{i-1}, y_i]\})$. Here, $\text{measure}(\{x | f(x) \in [y_{i-1}, y_i]\})$ is the sum of the lengths of those subintervals of $[a, b]$

on which $y_{i-1} \leq f(x) \leq y_i$ or, in the present case, the differences Δ_i in Fig. 6.10. In other words, it is determined how much of the domain is mapped by the function to some value between two end points in the range. While the Lebesgue measure is traditionally intimately connected with the notion of integrability, we use it here to analyze the behavior of the function $h(t)$ as it transitions through its characteristic forms and to derive quantitative – *e.g.*, measurable – conclusions.

Since we wish particularly to examine the radiation sampled by an observer situated on (or very close to) the cusp, the first step is to choose such an observation position (r_P, φ_P, z_P) approximately. The value of z_P is then refined until the desired accuracy is achieved.

The process begins with a choice of r_P and the use of the far-field approximation for θ_P , the polar angle of the cusp location,

$$\sin \theta_P = \frac{c}{r\omega}. \quad (6.44)$$

The relation between θ_P , r_P and z_P then gives a first estimate for z_P , namely

$$\tan \theta_P = \frac{r_P}{z_P}. \quad (6.45)$$

To refine the location more precisely, we use Eq. (6.27) and Eq. (6.28) in combination with Eq. (6.30), *i.e.*,

$$\cos(\varphi_P - \omega t) = \frac{c^2}{rr_P\omega^2} \quad \text{and} \quad (6.46)$$

$$\sin(\varphi_P - \omega t) = (1 - \cos^2(\varphi_P - \omega t))^{\frac{1}{2}} = \left(1 - \frac{c^4}{r^2 r_P^2 \omega^4}\right)^{1/2}, \quad (6.47)$$

to eliminate the time-dependent terms, yielding

$$\frac{r^2 r_P^2 \omega^2 \left(1 - \frac{c^4}{r^2 r_P^2 \omega^4}\right)}{r_P^2 + r^2 - 2 \frac{c^2}{rr_P \omega^2} rr_P + z_P^2} = c^2. \quad (6.48)$$

It is Eq. (6.48) that defines the expanding, spiral-shaped *locus* of the cusp curve shown in Fig. 6.7.

Rearranging terms and substituting the linear velocity v for $r\omega$, we find that

$$\frac{c^2}{v^2} \left(1 + \frac{v^2 - c^2}{\omega^2(r_P^2 + z_P^2)} \right) = \frac{r_P^2}{r_P^2 + z_P^2} = \sin^2 \theta_P. \quad (6.49)$$

Eq. (6.49) gives the means to refine z_P until the cusp is located to the desired precision. This is done by iterating through

$$z_P = r_P / \tan \theta_P \quad (6.50)$$

and

$$\theta_{P_{new}} = \arcsin \left[\frac{c}{v} \left(1 + \frac{v^2 - c^2}{\omega^2(r_P^2 + z_P^2)} \right)^{1/2} \right] \quad (6.51)$$

until convergence for θ_P is reached. This procedure, albeit simple, tends to converge to a precision of 10^{-15} in less than 20 iterations.

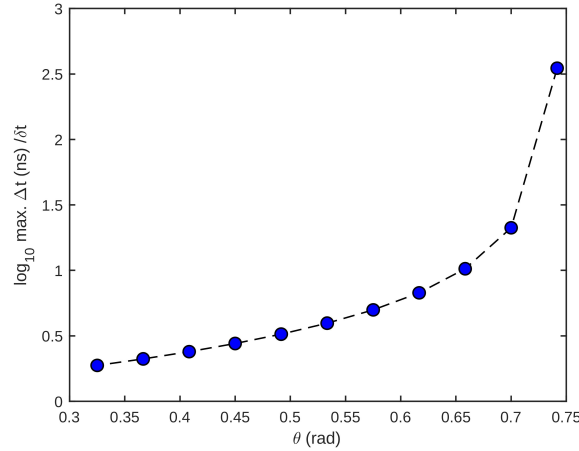


Figure 6.11: Logarithm of the degree of temporal focusing as a function of polar angle θ . Note the rapid increase as θ approaches the cusp angle.

Once the values of r_P and z_P have been determined, the function $t_P = h(t)$ (Eq. (6.4)) can be generated for an arbitrarily small neighborhood around the cusp location. As shown in Fig. 6.10 (left), the window of observation time, $\delta t_P = 1\text{ns}$, can be moved along this curve and the corresponding upper and lower limits in the

source-time domain, t_0 and t_1 , determined. In the discussion that follows below, we will plot the ratio $\Delta t/\delta t_P$, where $\Delta t = t_1 - t_0$, as a measure of the degree of temporal focusing. The root-finding procedure used to find the retarded times t and their values will be outlined in Section 6.5.

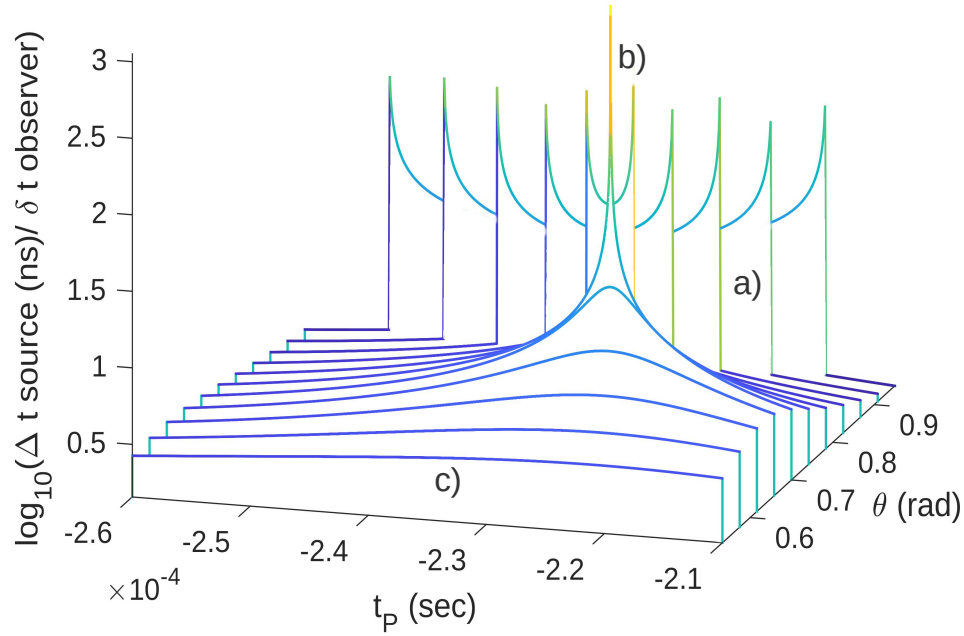


Figure 6.12: Waterfall plot for the degree of temporal focusing (a) inside, (b) on the cusp of, and (c) outside the envelope of the wave fronts (see Figures 6.1 and 6.8). $\Delta t/\delta t_P$ exhibits a pronounced spike at locations for which $\theta_P = \theta_{PC}$. For $\theta_P < \theta_{PC}$ or $\theta_P > 180^\circ - \theta_{PC}$, a single, broader peak is observed, but for $\theta_{PC} < \theta_P < 180^\circ - \theta_{PC}$, *e.g.*, region (b), the three intervals of source time that arrive within one interval of observation time result in a double-horned structure.

Having located the cusp, z_P can be raised or lowered to assess the degree of temporal focusing on either side. Fig. 6.10 (right) shows $t_P = h(t)$ for a location $\theta_{PC} < \theta_P < 180^\circ - \theta_{PC}$, where θ_{PC} denotes the exact angle of observation of the cusp. In this case, there are values for t_P at which either – or both – edges of the window δt_P intersect the curve $t_P = h(t)$ on more than one occasion. In such cases we calculate $\Delta t = \sum_i \Delta t_i$ as in Fig. 6.10 (right), where the Δt_i denote the extent of source time

for each contribution. Fig. 6.11 shows the results of a calculation of $\Delta t/\delta t_P$ as a function of polar angle. Values of $\omega = 4.5 \times 10^3/2\pi$ Hz and $r = 10^5$ m were chosen to reflect a pulsar similar in size and characteristics to the Crab pulsar [93,94]. Note the large degree of focusing – the vertical scale is logarithmic – in the time domain as the cusp angle is approached.

The degree of temporal focusing for several θ_P is shown in Fig. 6.12. $\Delta t/\delta t_P$ exhibits a pronounced spike at locations for which $\theta_P = \theta_{PC}$, *i.e.*, those that are subject to cusp radiation. For $\theta_P < \theta_{PC}$ or $\theta_P > 180^\circ - \theta_{PC}$, a single hump is observed, but for $\theta_{PC} < \theta_P < 180^\circ - \theta_{PC}$, the three intervals of source time that arrive within one interval of observation time result in a double-horned structure. As will be seen later, these features resemble closely those observed in pulsar light curves.

We have quantitatively analyzed (“measured”) the relation between emission and observation time of a radiation source in constant superluminal rotation and found that intriguing phenomena occur on the cusp and within the envelope of the emitted wave fronts. To investigate these unusual features further, we now proceed to calculate the more sophisticated Liénard-Wiechert potentials and fields for these locations, thereby introducing amplitude in addition to phase information.

6.4 In Search of the Point-Source Response

The expressions for the potentials and fields emitted by a single charged particle traveling along an arbitrary path – in essence a generalized solution to Maxwell’s equations – were formulated independently by Alfred-Marie Liénard (1869-1958) [174] and Emil Wiechert (1861-1928) [175,176] around the end of the 19th century. (Equivalent solutions were given by Sommerfeld (1904) [36] and Schott (1912) [37]; for detailed historical and mathematical accounts see, among many others, [37], [58], and [177].)

Since these *formulae* were proposed before the advent of Einstein’s special theory of relativity, they describe electromagnetic phenomena that propagate in “absolute space” or a “luminiferous ether”¹².

As shown by Maxwell (Section 2.4) and emphasized throughout the present work, electromagnetic waves, once emitted by a charged particle, will propagate at the speed of light *irrespective of how fast the particle is moving*, just as sound waves propagate at a speed independent of the source’s velocity. It is universally acknowledged that Maxwell’s equations (Eq. (2.2) – Eq. (2.5)) are invariant under Lorentz transformations and the results derived by Liénard and Wiechert, in consequence, relativistic [53]. Here we will rederive their results, which are nowadays encountered in the analysis of synchrotron radiation, and generalize them such that the speed of the source is arbitrary rather than restricted to the subluminal regime¹³.

Given as

$$\Phi(\mathbf{r}_P, t_P) = \left[\frac{q}{(1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}})R} \right]_{\text{ret}}, \quad \mathbf{A}(\mathbf{r}_P, t_P) = \left[\frac{q\boldsymbol{\beta}}{(1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}})R} \right]_{\text{ret}}, \quad (6.52)$$

in most standard texts (see, for example, [53]) the Liénard-Wiechert potentials are built directly from Maxwell’s equations and yield the complete, relativistically correct, time-varying electromagnetic field for a point charge q in (subluminal) motion. Here, as before, $R = |\mathbf{r}_P - \mathbf{r}|$ denotes the spatial separation between source and observer, $\hat{\mathbf{n}} = \mathbf{R}/R$ designates the radiation direction, and $\boldsymbol{\beta} = \mathbf{v}/c$ is the velocity vector of the source scaled by the speed of light, *i.e.*, the ‘Čerenkov number’. The subscript “ret” indicates that the quantity within the brackets is to be evaluated at the retarded time.

¹²In fact, Wiechert opens his essay on elementary laws in electrodynamics with the words, “Die neuere, sich auf Maxwell stützende Elektrodynamik ist durch Unterscheidung zwischen Aether und Materie im Innern der sinnlich wahrnehmbaren Körper in so weitem Maasse zu den Ansichten der älteren Schule zurückgekehrt, dass der einstige Gegensatz nicht mehr besteht.”

¹³Perepelitsa [178] reaches a similar conclusion for the special case of a tachyon moving on a rectilinear trajectory.

Chapter 7 will explain in detail how expressions for the electric and magnetic fields may, in general, be calculated directly from the potentials by using the relations $\mathbf{E} = -\nabla\Phi - \partial\mathbf{A}/\partial t$ and $\mathbf{H} = \nabla \times \mathbf{A}/\mu$ (listed as Eqs. (7.8) and (7.9) there). The derivation, which is non-trivial and requires a number of steps, can be found in Section 7.2 or, alternatively, in most standard textbooks. Here, we simply state the result and give the expressions for the fields [53] as

$$\mathbf{E}(\mathbf{r}_P, t_P) = q \left[\frac{\hat{\mathbf{n}} - \boldsymbol{\beta}}{\gamma^2(1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}})^3 R^2} \right]_{\text{ret}} + \frac{q}{c} \left[\frac{\hat{\mathbf{n}} \times \{(\hat{\mathbf{n}} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}\}}{(1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}})^3 R} \right]_{\text{ret}} \quad (6.53)$$

and

$$\mathbf{H}(\mathbf{r}_P, t_P) = \frac{1}{\mu} \hat{\mathbf{n}} \times \mathbf{E} \quad (6.54)$$

where

$$\gamma = \sqrt{\frac{1}{1 - \boldsymbol{\beta}^2}},$$

denotes the Lorentz factor, a quantity encountered ubiquitously in electrodynamics. Jackson ([53], p. 657) points out that Eqs. (6.53) and (6.54) divide themselves naturally into “velocity fields”, which are independent of acceleration, and “acceleration fields”, which depend linearly on $\dot{\boldsymbol{\beta}}$. The former are essentially static, falling off as R^{-2} , whereas the latter typically radiate, diminishing as R^{-1} .

Whilst the formulation of the Liénard-Wiechert potentials as given in Eq. (6.52) is sufficient to describe charges that remain subluminal, they must be generalized if the source is to travel arbitrarily fast and multiple retarded times may occur. We propose

$$\Phi(\mathbf{r}_P, t_P) = q \sum_{t_{\text{ret}}} \left[\frac{1}{|1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}}| R} \right], \quad \mathbf{A}(\mathbf{r}_P, t_P) = q \sum_{t_{\text{ret}}} \left[\frac{\boldsymbol{\beta}}{|1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}}| R} \right], \quad (6.55)$$

where we added summations and absolute-value brackets in the factor $|1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}}|$. Both are typically omitted from textbook derivations since, for the subluminal regime, retarded times are unique and $1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}}$ must be positive. As expected, the potentials

diverge when $\beta \cdot \hat{\mathbf{n}} = 1$ and the instantaneous distance from source to observer changes as c at the retarded time. Furthermore, it is easy to see that Eqs. (6.52) reduce to Eqs. (6.55) for a point source whose velocity remains below that of light.

Both the Liénard-Wiechert scalar and vector potential are fundamental causal solutions to the three-dimensional wave equation (4.1) for a unit point charge or, in the present context, a polarization-current element of infinitesimal extent. In the proof to follow we use $\Phi(\mathbf{r}, t)$ (but could equally well have chosen $\mathbf{A}(\mathbf{r}, t)$) which must satisfy

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \Phi(\mathbf{r}, t) = 4\pi q(\mathbf{r}, t) \quad \text{in } \mathbb{R}^3 \times (0, \infty). \quad (6.56)$$

Likewise, the Green's function G (or, rather, its reciprocal, see Section 4.3.3 for a refresher), introduced as Eq. (4.114) in Section 4.3.5, solves Eq. (4.72), restated here for convenience as

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) G(\mathbf{r}_P, t_P | \mathbf{r}, t) = 4\pi q(\mathbf{r}, t).$$

Hence, to prove that Φ is the Green's function for Eq. (4.59) it is sufficient to show that $\Phi(\mathbf{r}, t) = G(\mathbf{r}_P, t_P | \mathbf{r}, t)$, where $q = 1$ (since the charge is of unit strength) and Φ as well as G represent the contribution from a single spacetime pair (\mathbf{r}, t) .

Proof: We take the argument of the delta function in Eq. (4.114), rewrite it slightly and find its derivative with respect to time, *i.e.*,

$$\begin{aligned} g(t) &= t - t_P + \frac{R(t)}{c} \\ &= t - t_P + \frac{|\mathbf{r}_P - \mathbf{r}(t)|}{c} = 0 \end{aligned} \quad (6.57)$$

and

$$\begin{aligned} \frac{d}{dt} g(t) &= \frac{d}{dt} \left(t - t_P + \frac{|\mathbf{r}_P - \mathbf{r}(t)|}{c} \right) \\ &= 1 + \left| \frac{d}{dt} \frac{-\mathbf{r}(t)}{c} \right| \\ &= 1 - \frac{\dot{\mathbf{r}}}{c} \end{aligned} \quad (6.58)$$

$$\begin{aligned}
&= 1 - \frac{\dot{\mathbf{r}}}{c} \left(\frac{|\mathbf{r}_P - \mathbf{r}|}{R} \right) \\
&= 1 - \frac{\mathbf{v}}{c} \cdot \hat{\mathbf{n}} \\
&= 1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}}.
\end{aligned}$$

Now we may use the third generalized scaling property,

$$\int_{-\infty}^{+\infty} dt \delta(g(t)) f(t) = \sum_i \frac{f(t_i)}{|g'(t_i)|}, \quad (6.59)$$

where $f(t) = 1/R$ and $g(t) = t - t_P + R(t)/c$. In consequence,

$$\begin{aligned}
\int_{-\infty}^{+\infty} dt G(\mathbf{r}_P, t_P | \mathbf{r}, t) &= \int_{-\infty}^{+\infty} dt \delta(g(t)) f(t) \\
&= \int_{-\infty}^{+\infty} dt \delta(t - t_P + R/c) \frac{1}{R} \\
&= \sum_{t_{\text{ret}}} \frac{1}{R|1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}}|} = \sum_{t_{\text{ret}}} \Phi.
\end{aligned} \quad (6.60)$$

where the summation extends over all roots of $g(x)$, which are assumed to be simple.

■

Eqs. (6.58)–(6.60) confirm the results that we derived using the three dimensional wave equation (Chapter 4) and basic Huygens techniques: The retarded potential $\Phi(\mathbf{r}_P, t_P) = G(\mathbf{r}_P, t_P | \mathbf{r}, t)$ radiated by a point source in faster-than-light motion is discontinuous on the envelope and cusp of the wave fronts. If the hypersurface ∂X is approached from the outside, the sum (Eq. (6.60)) has only a single term and yields a finite value for G , but if it is approached from the inside, two retarded times coalesce at an extremum of h and the equation yields a divergent value.

Having verified that the potentials (Eq. (6.55)) are indeed fundamental causal solutions of Maxwell's equations, the Liénard-Wiechert fields of a point charge q traveling arbitrarily fast on a given trajectory $\mathbf{r}(t)$ as observed at spacetime coordinates (\mathbf{r}_P, t_P) can be written as

$$\mathbf{E}(\mathbf{r}_P, t_P) = q \sum_{t_{\text{ret}}} \left[\frac{\hat{\mathbf{n}} - \boldsymbol{\beta}}{\gamma^2 |1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}}|^3 R^2} \right] + \left[\frac{\hat{\mathbf{n}} \times \{(\hat{\mathbf{n}} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}\}}{|1 - \boldsymbol{\beta} \cdot \hat{\mathbf{n}}|^3 c R} \right] \quad (6.61)$$

and

$$\mathbf{H}(\mathbf{r}_P, t_P) = \frac{1}{\mu} \hat{\mathbf{n}} \times \mathbf{E}. \quad (6.62)$$

Notice that, for a charge that moves faster than c , the Lorentz factor γ is complex, not real as in the subluminal regime. As a result, the intensity of the radiation due to a superluminal source oscillates as a function of its frequency. This unique feature of faster-than-light sources explains, as we shall later see, the oscillations observed in the frequency spectrum of the Crab pulsar.

6.5 The Relation between Emission and Observer Time: An Old Problem Revisited

The evaluation of the radiation fields given in Eq. (6.61) and Eq. (6.62) poses no difficulties once the number of retarded times t and their values have been determined. Hence, we must find the root(s) of function $h(t)$ (Eq. (6.4)), *i.e.*, solutions to the relation

$$t - t_P + c^{-1}[z_P^2 + r_P^2 + r^2 - 2r_P r \cos(\varphi_P - \omega t)]^{1/2} = 0, \quad (6.63)$$

which is transcendental.

For mathematical ease, if, again, at the cost of physical transparency, we introduce the Lagrangian variables $\hat{\varphi} \equiv \varphi - \omega t$ and $\hat{\varphi}_P \equiv \varphi_P - \omega t_P$, capturing the rigid rotation of the patterns of the source and the radiation field [80]. Using distances scaled by units of light cylinders as introduced in Section 6.2, Eq. (6.3) becomes

$$\phi = h(\varphi) \equiv \varphi + \left[\hat{R}_P^2 + \hat{r}^2 - 2\hat{r}\hat{R}_P \cos(\varphi - \varphi_P) \right]^{1/2}, \quad (6.64)$$

where $\phi \equiv \hat{\varphi} - \hat{\varphi}_P$ and $\hat{R}_P^2 = \hat{z}_P^2 + \hat{r}_P^2$. A numerical difficulty with (6.64) is that a large (*e.g.*, astronomical) \hat{R}_P swamps the oscillations in the cosine term. We remedy

this by subtracting $\widehat{R}_0 \equiv (\widehat{R}_P^2 + \widehat{r}^2)^{1/2}$ from both sides, resulting in

$$\begin{aligned}\Delta\phi = \Delta h(\varphi) &\equiv \varphi + \widehat{R}(\varphi) - \widehat{R}_0 \\ &= \varphi - \frac{2\widehat{r}\widehat{r}_P \sin \varphi}{(\widehat{R}_0^2 - 2\widehat{r}\widehat{r}_P \sin \varphi)^{1/2} + \widehat{R}_0},\end{aligned}\tag{6.65}$$

where we have also chosen $\varphi_P = -3\pi/2$ for convenience. In the limit that $\widehat{R}_P \rightarrow \infty$, this becomes

$$\Delta\phi = \varphi - \widehat{r} \sin \theta_P \sin \varphi,\tag{6.66}$$

which is of the form of Kepler's equation¹⁴ for elliptical orbits,

$$M = E - e \sin E.\tag{6.67}$$

Simply stated, Kepler's equation relates the eccentric anomaly E , the mean anomaly M , and the eccentricity e in an elliptic orbit. It is, up to this day, of fundamental importance in celestial mechanics, but cannot be inverted directly to determine E – the angular parameter that defines the position of the body. Hence, despite its apparent simplicity, papers devoted to Kepler's Problem have appeared in virtually every decade from 1650 to the present, endowing it with an undeniable luster and allure for the modern practitioner, although a number of satisfactory solutions are long known. Peter Colwell's text [180] provides an excellent survey of the problem and its treatment over the centuries, beginning with early attempts by Kepler himself,¹⁵ Newton, Euler, Gauss and Cauchy. Infinite series solutions – the most

¹⁴What we call Kepler's equation and Kepler's method of solving it appeared, in fact, in the ninth century writings of Habash-al-Hasib in connection with problems of parallax.

¹⁵Kepler's own solution can be summarized as follows: Given e and M , guess an approximate solution E_0 for E and calculate

$$M_0 = E_0 - e \sin E_0.$$

Let $E_1 = E_0 + (M - M_0)$ and calculate

$$M_1 = E_1 - e \sin E_1.$$

Then $E_2 = E_1 + (M - M_1)$ should yield a better approximation of E .

elegant – use Lagrange’s Theorem, Bessel functions and Levi-Civita as well as Lie series while modern treatments are due to Burniston and Siewert [181], Ioakimidis and Papadakis [182], and Delves and Lyness [183]. In our case, the “eccentricity” e is greater than one when multiple solutions for retarded times are possible, which is to say that the orbit is “nearly parabolic.”¹⁶

The correspondence of Eq. (6.66) with Kepler’s equation brings to the fore the serious numerical difficulty in solving either of the transcendental equations (6.65) or (6.66): That e near one and M near zero results in severe subtractive cancellation [180]. In orbital mechanics, this special case of determining the eccentric anomaly near the pericenter of a near-parabolic orbit is usually merely a side issue, but for our superluminal sources, the analogous situation is a central concern: As described in Section 6.3.1, we are interested precisely in determining retarded times near the cusp of the envelope, which are the strongest contributors to the observed field. For typical pulsar observations, $\hat{r} \sin \theta_{\text{p}}$ is very near indeed to unity. Close to the cusp, the increment above one is of the order of \hat{R}_{p}^{-2} , with the scaled distances \hat{R}_{p} to pulsars being of order 10^9 to 10^{15} . The oscillatory region of g is correspondingly tiny: in radians, $\varphi_- - \varphi_+$ is of order \hat{R}_{p}^{-1} , and $\phi_+ - \phi_-$, \hat{R}_{p}^{-3} . *Hence, one of the central problems of the present thesis is one in numerical analysis: To find one or more solutions to Kepler’s equation near the pericenter of a near-parabolic orbit, where e is near one and M near zero.*

6.5.1 An Iterative Solution to Kepler’s Equation

Whilst various numerical methods of approximation have been suggested to find the (usually sole) root of Kepler’s equation, most contemporary solutions are based on an iterative algorithm developed by Newton and later refined by Raphson.¹⁷

¹⁶An elliptic or a hyperbolic orbit with eccentricity close to 1 is said to be “nearly parabolic.” Convention holds that this is the case if the boundary $\delta = |e - 1| \leq 0.01$.

¹⁷For an exhaustive list of references, see [180].

Given a real-valued continuous function f , we wish to construct a sequence (x_i) , using iteration, which converges to a solution of $f(x) = 0$. Newton's method for the solution of this problem is formally defined as

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}, \quad i = 0, 1, 2, \dots, \quad (6.68)$$

with prescribed starting value x_0 , where we implicitly assume that $f'(x_k) \neq 0$ for all $k \geq 0$.

Hence, Newton's method is a simple iteration of the function

$$g(x) = x - f(x)/f'(x), \quad (6.69)$$

that, unlike other iterative approximation algorithms, requires the evaluation of both the function $f(x)$ and its derivative $f'(x)$ at arbitrary points x . Geometrically, it consists of extending the tangent line at a current point x_i until it crosses zero, then setting the next guess x_{i+1} to the abscissa of that zero-crossing.

The power of Newton's method lies in its rate of convergence, which is quadratic, where quadratic convergence of a sequence is defined as follows:

Suppose that $\xi = \lim_{i \rightarrow \infty} x_i$. We say that the sequence (x_i) converges to ξ *with at least order* $q > 1$, if there exists a sequence (ϵ_i) of positive real numbers converging to 0, $\mu > 0$, such that

$$|x_i - \xi| \leq \epsilon_i, \quad i = 0, 1, 2, \dots, \quad \text{and} \quad \lim_{i \rightarrow \infty} \frac{\epsilon_{i+1}}{\epsilon_i^q} = \mu. \quad (6.70)$$

If (6.70) holds with $\epsilon_i = |x_i - \xi|$ for $i = 0, 1, 2, \dots$, then the sequence (x_i) is said to converge to ξ *with order* q . In particular, if $q = 2$, the sequence (x_i) is said to converge to ξ *quadratically*.

Taylor's theorem states that any function $f(x)$ which has a continuous second derivative can be represented by an expansion about a point that is close to a root of $f(x)$. Suppose this root is ξ . Then the expansion about x_i is of the form

$$f(\xi) = f(x_i) + f'(x_i)(\xi - x_i) + R, \quad (6.71)$$

where the Lagrange form of the remainder R is

$$R = \frac{1}{2!} f''(\chi_i)(\xi - x_i)^2, \quad (6.72)$$

with $x_i \leq \chi_i \leq \xi$.

Since ξ is a root, (6.71) assumes the form

$$f(\xi) = 0 = f(x_i) + f'(x_i)(\xi - x_i) + \frac{1}{2} f''(\chi_i)(\xi - x_i)^2. \quad (6.73)$$

Rearranging and dividing Eq. (6.73) by $f'(x_i)$ gives

$$\frac{f(x_i)}{f'(x_i)} + (\xi - x_i) = -\frac{f''(\chi_i)}{2f'(x_i)}(\xi - x_i)^2. \quad (6.74)$$

Recalling that x_{i+1} is defined by (6.68), we find that

$$\underbrace{\xi - x_{i+1}}_{\epsilon_{i+1}} = -\frac{f''(\chi_i)}{2f'(x_i)} \underbrace{(\xi - x_i)^2}_{\epsilon_i^2}, \quad (6.75)$$

that is,

$$\epsilon_{i+1} = -\frac{f''(\chi_i)}{2f'(x_i)} \epsilon_i^2. \quad (6.76)$$

Taking absolute values on both sides yields

$$|\epsilon_{i+1}| = \frac{|f''(\chi_i)|}{2|f'(x_i)|} \epsilon_i^2, \quad (6.77)$$

which shows that the rate of convergence is quadratic. As in (6.68) above, we require that $f'(x) \neq 0 \quad \forall x \in I$, where I is the interval $[\xi - r, \xi + r]$ for some $r \geq |(\xi - x_0)|$, and that $f''(x)$ be finite $\forall x \in I$. Furthermore, the prescribed starting value x_0 needs to be *sufficiently close* to the root.

While Newton's method has very attractive local convergence properties, its global behavior can be complicated and erratic [184–187]. Many texts in numerical analysis include detailed studies of pathological cases, where the algorithm yields grossly inaccurate, meaningless corrections, enters a nonconvergent cycle or attempts

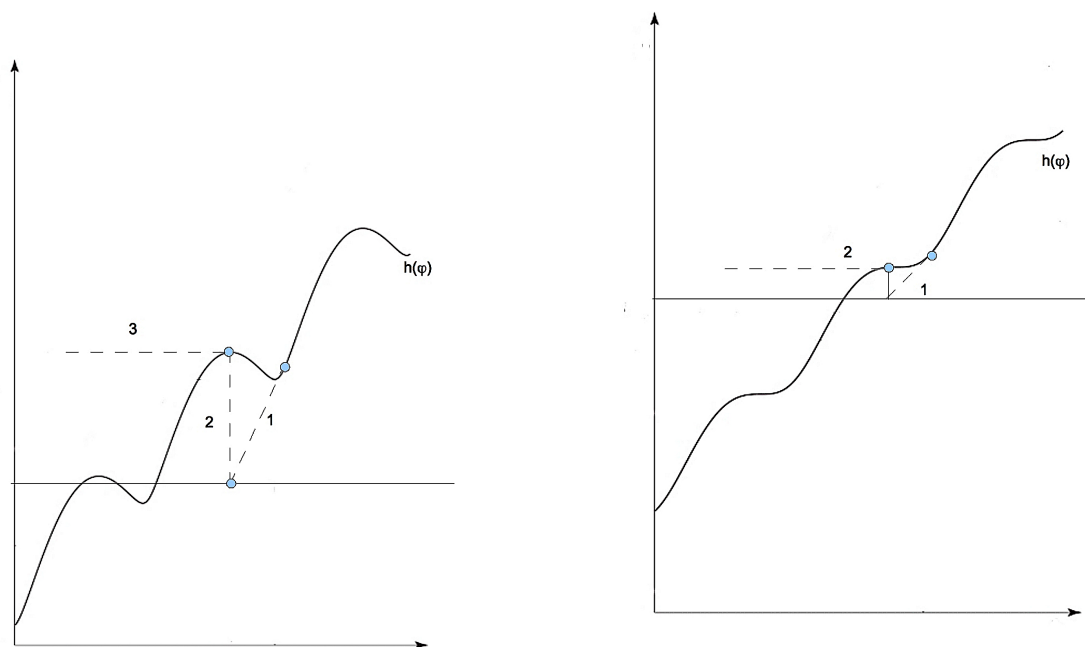


Figure 6.13: Pathological cases in which Newton's method as applied to Eq. (6.66) will fail to converge. Carefully chosen bracketing bounds and the insertion of a bisection step whenever Newton-Raphson would take the solution out of bounds prevent calamities such as these.

to locate a root at infinity [184–187]. While a formal analysis of the Newton-Raphson formula as applied to Kepler's equation for nearly parabolic orbits is beyond the scope of this work, we include two examples (Fig. 6.13) which illustrate how the procedure will fail to locate the roots of Eq. (6.66) if the starting values φ_0 are ill-chosen.

Since the global convergence properties of Newton's method are rather poor, it is necessary to design routines that i) guard against the method making inappropriate corrections while ii) ensuring that rapid convergence be maintained. The former is achieved fairly easily by utilizing a combination of bisection and the Newton-Raphson formula as suggested in [188]. The hybrid algorithm takes a bisection step whenever the Newton-Raphson procedure would take the solution out of bounds or is not reducing the bracket size – the extent of the region around the function's crossing of the abscissa – swiftly enough. The latter, namely to bracket the roots such that few

bisection steps need to be taken and quadratic convergence remains ensured, requires more diligence and ingenuity.

The first step is to assess the characteristic combination of emission and observation time for each individual point of the radiated field, *i.e.*, to determine whether the observer is located inside, on the cusp of, or outside the envelope. To this end we evaluate the difference $\Delta \equiv (\hat{r}_P^2 - 1)(\hat{r}^2 - 1) - \hat{z}_P^2$. If this is nonpositive, then h is nonoscillatory and we have a single zero of $h(t) - t_P$ somewhere on the closed interval bounded by $t_P - c^{-1}(r^2 + r_P^2 + z_P^2 \pm 2rr_P)^{1/2}$. We may employ Newton's method as outlined and modified above or any other numerical procedure that locates a single root on a continuous interval. If, on the other hand, Δ is positive, there may be multiple retarded times, each bracketed by a sequence of adjacent local extrema with alternating signs (or zero). We locate the greatest minimum $h(t_{min_0})$ not exceeding t_P and the least minimum $h(t_{min_k})$ greater than t_P . Then the extrema of h between $h(t_{min_0})$ (inclusive) and $h(t_{min_k})$ (exclusive) delimit a sequence of monotonic intervals, each of which contains a single zero of $h(t) - t_P$, and we apply our numerical root-finder to each of these.

Whilst the approach described above is sufficiently accurate for terrestrial distances, difficulties arise if we move to astronomical length scales, especially if we wish to determine the eccentric anomaly near the pericenter of the nearly parabolic orbit, as outlined above. (The closest known pulsar to Earth is PSR J0108-1431 in the direction of the constellation Cetus, at a distance of about 85 parsecs or 280 light years from the sun [94]).

Our solution to finding retarded times in this critical region is to precisely locate the turning points (φ_+, ϕ_+) and (φ_-, ϕ_-) of h using multiple-precision arithmetic and approximate $\Delta\Phi$ in this region by a cubic interpolating the turning points and matching the zero derivatives at those points. We find that this cubic is typically a good approximation to $\Delta\Phi$ over a range several orders of magnitude larger than $\varphi_- - \varphi_+$, and it can be directly inverted to yield three real roots for $\phi_- < \phi < \phi_+$

or one elsewhere. Beyond this range, $\Delta\Phi$ is well enough behaved that, starting from a linear first approximation to the root, Newton's modified iteration method will converge quickly. There may be other regions of the parameter space, however, where a more involved interpolation technique will be needed to produce a root or a first approximation. However, it appears that approximating the Kepler function by a cubic Hermite spline with knots at multiples of π or $\pi/2$ might be a viable approach to solving Kepler's problem in the near-parabolic case.

6.6 From Point Source to Wavepacket

Sections 6.6.1 to 6.6.4 follow the treatment used in the recently submitted paper [189] by the present author and colleagues. It provides a crucial bridge between the point source and the model of pulsars that concludes this chapter.

6.6.1 Concept

After studying a point-like accelerated superluminal source of radiation, a logical development is to extend the source primarily in one dimension. Now, therefore, we muse upon what might be termed a “wavepacket” of polarization current that has finite extent in both space and time and which contains a spread of frequencies. As in previous sections, the consequences of temporal focusing will be shown to be important.

We consider a wavepacket that moves on a linear trajectory, with acceleration arranged so that temporal focusing occurs at a particular point.¹⁸ Besides representing the simplest form of temporal focusing, this configuration has the advantage that it

¹⁸Although we describe the problem using the two-dimensional (x, y) plane, it will be quickly seen that the focusing occurs in three dimensions on a *ring* of points disposed symmetrically around the y -axis.

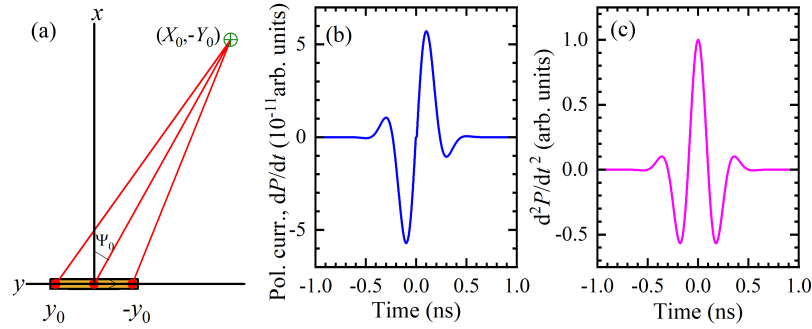


Figure 6.14: (a) Experimental concept. An element of polarization current (red) moves along the dielectric antenna (dark yellow shading) such that the component of its velocity in the direction of the target [green cross at $(X_0, -Y_0)$] is always c , the speed of light in the surrounding medium, here assumed to be vacuum. The center of the antenna is at $(0, 0)$. (b) Notional time dependence of the polarization current dP/dt sent along the antenna. (c) Derivative of the curve shown in (b) with respect to time t . The “arbitrary units” in (c) are equivalent to those in (b), the large scaling factor occurring because of the rapid time dependences of the waveforms.

(or something close to it) can be realized experimentally using the linear antennas described in Chapter 3. Fig. 6.14(a) shows a plan view of such an antenna’s dielectric of length $2y_0$ with its center at $(0, 0, 0)$ lying along the Cartesian y -axis.¹⁹ As in the experimental antennas, the dielectric has rectangular cross-section; its depth $2x_0$ (extent in the x direction) and height $2z_0$ (extent in the z direction) are symmetrical about the y axis; both are much less than $2y_0$. The target point for temporal focusing is situated in the (x, y) plane at a radial distance R_0 ; in this case we follow telecommunications engineers and use the angle Ψ_0 “off boresight” (*i.e.*, away from the normal to the antenna’s center) to describe the target’s azimuthal position. As everything of interest lies in the (x, y) ($z = 0$) plane (but see the footnote on the previous page), for conceptual simplicity we also drop the Cartesian z coordinate for the time being. Thus, the target point is at $(X_0, -Y_0)$, where

$$X_0 = R_0 \cos \Psi_0 \quad \text{and} \quad |Y_0| = R_0 \sin \Psi_0. \quad (6.78)$$

¹⁹This coordinate system is chosen to match those used for modeling antennas elsewhere in this dissertation.

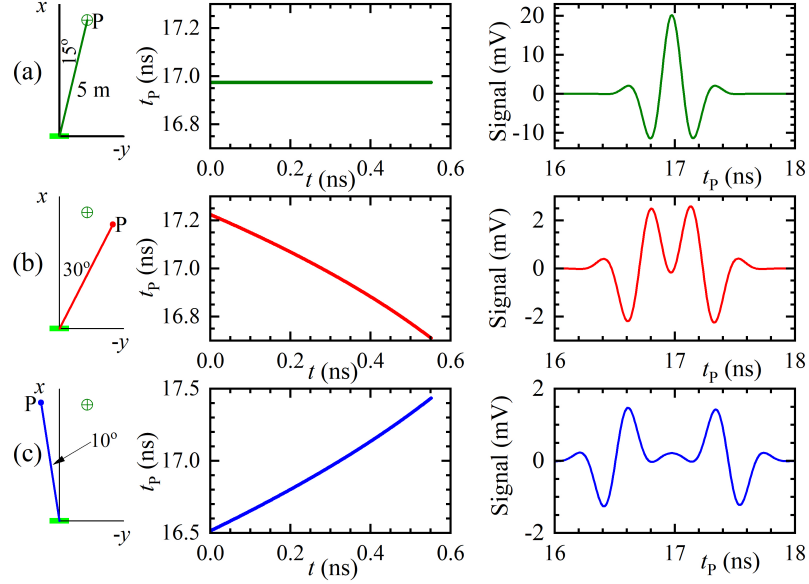


Figure 6.15: Each row of the figure shows the effect of moving the detector to various positions P that are 5 m away from the center of the antenna (light green). In each row, the left panel is a schematic of the detector position, the center shows the relationship between arrival time t_p of radiation emitted at time t by a point accelerating along the dielectric antenna, and the right is the signal detected due to the polarization current shown in Fig. 6.14(b) being accelerated along the antenna. In all cases, the target (green cross) is 5 m from the antenna center at an angle of 15° to the x axis. Row (a): detector at the target position. Row (b): detector placed on a line making an angle of 30° with the x axis. Row (c): detector placed on a line making an angle of -10° with the x axis.

We now consider a point in a polarization current that is moving through the antenna's dielectric along the y -axis; The instantaneous distance r between the point at $(0, y)$ and the target at (X_0, Y_0) is given by

$$r^2 = X_0^2 + (Y_0 + y)^2. \quad (6.79)$$

Suppose that the point moves in such a way that the component of its velocity towards the target point is always c , the speed of light *in vacuo*,²⁰ that is, $(dr/dt) =$

²⁰The surrounding medium is assumed to be a vacuum.

–c. Differentiating Eq. (6.79) with respect to time t , inserting the above value for (dr/dt) and rearranging, we obtain the point's velocity along the y -axis:

$$\frac{dy}{dt} = -c \frac{[X_0^2 + (Y_0 + y)^2]^{\frac{1}{2}}}{Y_0 + y}. \quad (6.80)$$

Integrating Eq. (6.80), and assuming that the point commences its journey along the antenna at $y = y_0$ and time $t = 0$, we obtain a relationship between the point's position y and time t :

$$t = \frac{1}{c} \left[(X_0^2 + (Y_0 + y_0)^2)^{\frac{1}{2}} - (X_0^2 + (Y_0 + y)^2)^{\frac{1}{2}} \right]. \quad (6.81)$$

We now consider a detector placed at a general point P with coordinates $(X, -Y)$ in the (x, y) plane. The radiation emitted by the point as it travels along the antenna will reach P at a time t_P given by

$$\begin{aligned} t_P &= t + \frac{1}{c} (X^2 + (Y + y)^2)^{\frac{1}{2}} \\ &= \frac{1}{c} \left[(X_0^2 + (Y_0 + y_0)^2)^{\frac{1}{2}} - (X_0^2 + (Y_0 + y)^2)^{\frac{1}{2}} + (X^2 + (Y + y)^2)^{\frac{1}{2}} \right]. \end{aligned} \quad (6.82)$$

It should be obvious that if, and only if, $X = X_0$ and $Y = Y_0$, then $t_P = \text{a constant}$. For all other choices of detector position, t_P is a function of y and therefore of t .

This situation is illustrated in the first two columns of Fig. 6.15. The intended target $(X_0, -Y_0)$ is at a distance of 5 m from the antenna center and a target angle of 15° (Fig. 6.15, row (a), left column); if the detector P is placed *exactly* at this position, then $t_P = \text{constant}$ (row (a), center column). The constant in question is of course the transit time of light from $y = y_0$, the place at which the point source enters the antenna at $t = 0$, to the focus; subsequently the accelerated motion of the point source along the antenna exactly compensates for the changing point-to-focus distance. For this configuration, one might say (transit time of source along antenna from y_0 to y) + (transit time of light from y to focus) = constant. If, on the other hand, the detector position P is not at $(X_0, -Y_0)$ [Fig. 6.15, rows (b) and (c)], then t_P becomes a function of t .

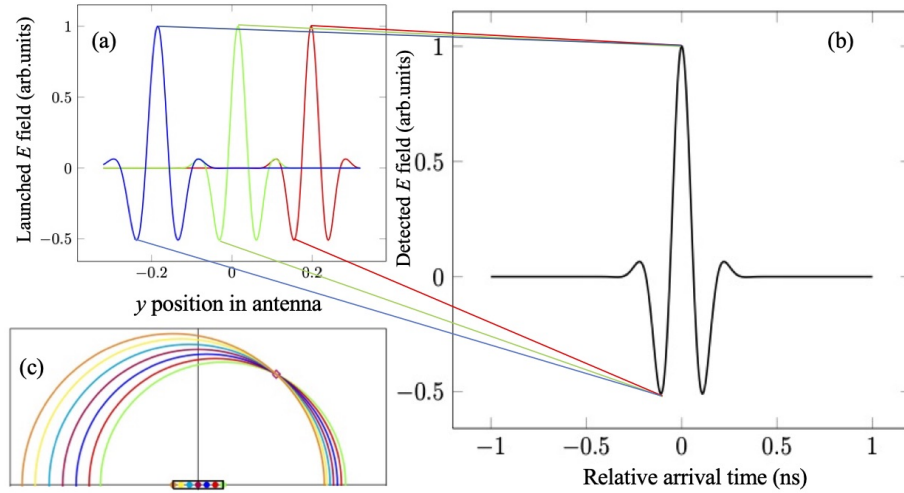


Figure 6.16: (a) Launched E -field [$\propto (d^2P/dt^2)$], corresponding to a polarization current similar to that in Fig. 6.14(b), at three different times (denoted by red, green and blue curves) during its transit along the antenna; note that the waveform “stretches out” due to the acceleration. (b) The corresponding detected E -field at the target point; colored lines linking (a) and (b) show schematically the principle that radiation from a particular point on the traveling waveform arrives at the same time at the target. This is because the acceleration compensates exactly for the different distances between source point and detection locations. Hence, features in the launched E -field are reinforced in the correct time sequence in the detected signal. (c) The same principle is illustrated using Huygens wavelets; colored dots in the antenna (dielectric outlined by black lines) represent the positions of a particular point on waveform (a) at different times during its transit of the antenna; semicircles of the same color show corresponding emitted Huygens wavelets arriving at the focus point (orange diamond) simultaneously.

Thus far, the situation is not very different from some of the point source examples discussed in previous sections. However, we are now interested in a polarization current that is both extended in the y direction, and travels along the antenna with the characteristic motion described above, *i.e.*, all points within it have a velocity component of c towards the target. As mentioned above, the extent of the antenna in the x and z directions is small compared to its length. In the experimental demonstration below, the x and z extents of the antenna are also small compared to the wavelength of the emitted radiation. Therefore we ignore the slight variations in distance caused by the non-zero x depth and z height of the antenna and represent the

motion of the volume-distributed polarization current by a vector function depending only on y and t

$$\frac{d\mathbf{P}}{dt} = \mathbf{f}(y, t) = \mathbf{f}[t - p(y)]. \quad (6.83)$$

Constant phase points within the polarization current are represented by $t - p(y) = \text{constant}$; differentiating this with respect to t results in

$$1 = \frac{dp}{dt} = \frac{dp}{dy} \frac{dy}{dt} \quad (6.84)$$

Substituting from Eq. (6.80), rearranging and integrating, we obtain

$$p(y) = -\frac{[X_0^2 + (Y_0 + y)^2]^{\frac{1}{2}}}{c}. \quad (6.85)$$

Together, Eq. (6.83) and Eq. (6.85) describe the motion of an extended polarization current, all of the points within which approach the target at a speed of c . This means that *all* of the radiation emitted from such a point during its passage through the antenna will arrive at the same time at the target.

To envisage what this implies, we consider the effect of sending the polarization-current wavepacket shown in Fig. 6.14(b) along the antenna using the acceleration scheme described above (*i.e.*, target angle of 15° at a distance of 5 m from the antenna center). The resulting signals for the detector positions given in the first column of Fig. 6.15 are shown in the third column of the same figure; Section 7.7, will describe how such calculations are carried out.

At the target angle and distance [(Fig. 6.15, row (a))], the detected signal reproduces the shape of the *time derivative* of the polarization-current wavepacket [Fig. 6.14(c)] *exactly*. Away from the target position [Fig. 6.15, rows (b) and (c)], the detected signal is much smaller and has altered frequency content and shape.

Two features of Fig. 6.15 must be considered. Firstly, the perfect reproduction of some aspect of the transmitted signal at the focus point may be understood as follows. The acceleration scheme ensures that all of the emission from a particular point on

the wavepacket as it traverses the antenna will arrive at the focus *simultaneously* (see center plot of row (a) of Fig. 6.15). Indeed, emission from any point on the wavepacket will behave similarly, the arrival time being the time at which that point entered the antenna at $y = y_0$ plus the time taken for light to travel from $(0, y_0)$ to the target. This ensures that all of the emitted radiation arrives at the focus in the correct time sequence, thereby exactly reproducing the time dependence of the emission of the whole packet.

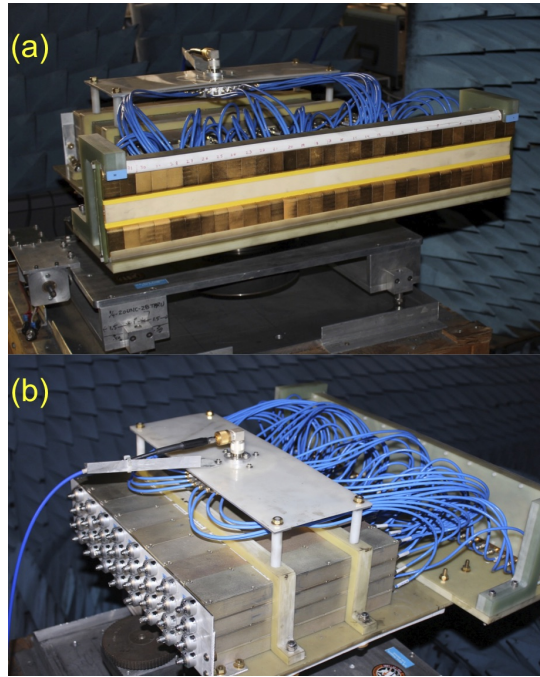


Figure 6.17: (a) Front view of passive antenna TD 4 used in the demonstration experiment mounted on the turntable in the TA-35 anechoic chamber. It has 32 alumina elements spanning a total length of 0.64 m. (b) Rear view of the antenna showing the 32-way splitter feeding 32 independent ATM P1214 mechanical phase shifters. The dials for adjusting the phase are visible on the lower left of the picture.

Secondly, there is the derivative shape of the signal received at the focus. As we will see in Section 7.7, under the focus conditions, the magnetic vector potential \mathbf{A} at the focus position and time t_P becomes proportional to the time dependence of the polarization current at the retarded time t ; hence the launched electric field, which

is proportional to the time differential of \mathbf{A} is proportional to the time derivative of the polarization current. Therefore it is the *launched E -field* that is reproduced at the target position. Fig. 6.16 illustrates both of these ideas schematically.

Now that the concept has been outlined, the following sections will describe a closely related experimental demonstration and the method for calculating the emitted radiation.

6.6.2 Experimental Demonstration

Passive antenna TD 4 (Fig. 6.17) was used to provide a simple demonstration. The experiment was carried out in the TA-35 anechoic chamber with Schwarzbeck dipoles as detectors (Section 3.5.2). The only variation on standard procedure (Chapter 3) was that the signal from the dipoles could be sent to a Mini-Circuits TVA-82-213A broadband amplifier that allowed the time-dependent voltage to be viewed and/or digitized using a Tektronix TDS7404 digital oscilloscope.

The description in the previous section has been framed in terms of a single traveling wavepacket. However, detecting such a single pulse, especially if it contains a wide spread of frequencies, presents technical difficulties for a laboratory with relatively limited resources. Instead, we choose to transmit and detect what is in effect a train of wavepackets. This is in the form of a continuous broadband signal with a distinctive shape, based on a mixture of harmonics of 0.90 GHz and synthesized by mixing outputs from phase-locked TTI TGR6000 and Agilent N9318 function generators.

The synthesized signal is conveyed to a Mini-Circuits TVA-82-213A amplifier, the output of which drives a 32-way splitter feeding 32 independent ATM P1214 mechanical phase shifters [Fig. 6.17(b)]. The latter are used to set the time-delays of the signals sent to each antenna element, reproducing the above acceleration scheme.

For the purposes of keeping the “information focus” well within the anechoic chamber, $X_0 = 3.03$ m and $Y_0 = 0.64$ m are chosen, yielding target distance $R_0 = 3.09$ m and target azimuthal angle $\Phi_0 = 11.9^\circ$.

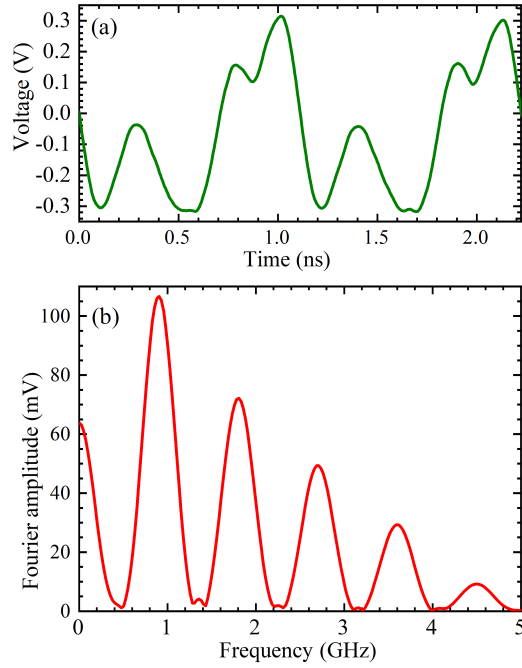


Figure 6.18: (a) The voltage measured by placing the dipole receiver 3 mm in front of the 16th antenna element as a function of time; in effect, this is the desired transmitted signal. (b) Fourier transform of the waveform shown in (a). Note the easily recognised “triangular” pattern of harmonics of 0.9 GHz.

The time-dependence of the broadcast waveform is recorded by placing the receiver dipole 3 mm in front of the 16th element of the antenna and observing the signal on the oscilloscope. As long as the shortest emitted wavelength is much larger than distance from the dielectric to the detector, the calculations of Section 7.7 can be used to show that the E -field thus detected by the dipole is, to a good approximation, $\propto \partial^2 \mathbf{P} / \partial t^2$, where \mathbf{P} is the polarization passing the point in the dielectric closest to the detector antenna. Hence, an analogue of Fig. 6.14(c) for the experimental wavetrain is captured. Moreover, because the equipment used is exactly the same, any frequency-dependent artefacts are identical in the measurements of

the broadcast and received signals, making a comparison analogous to that between Fig. 6.14(b) and the third column of Fig. 6.15 simpler and more direct.

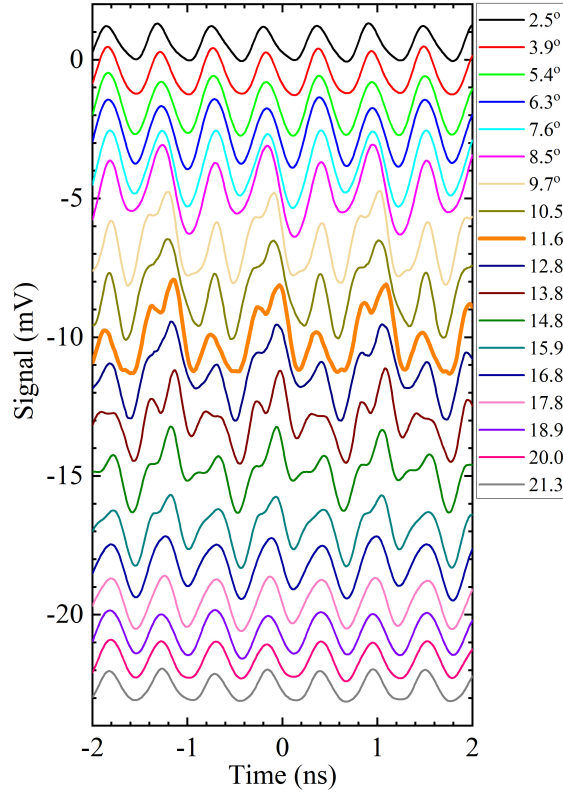


Figure 6.19: Time dependence of the signal received by the detector dipole for an antenna-to-detector distance of 3.0 m and for the azimuthal angles show in the key. The shape of the transmitted waveform (Fig. 6.18(b)) is only reproduced close to the “target” angle of 11.6° (orange trace).

The waveform used for the experiments is selected by adjusting the outputs and phasing of the two signal generators and is shown in Fig. 6.18(a). It is chosen because (i) it has a very distinctive time-dependent shape (*e.g.*, the double peak followed by two differing minima, one relatively broad) and (ii) an easily recognized “triangular” Fourier spectrum [Fig. 6.18(b)]. These traits aid in the rapid location of ranges of distance and azimuthal angle over which the broadcast signal was reproduced accurately. In all that follows, the azimuthal angles are corrected for the fact that the antenna rotation axis is slightly offset from the center of the antenna (see Chapter 3).

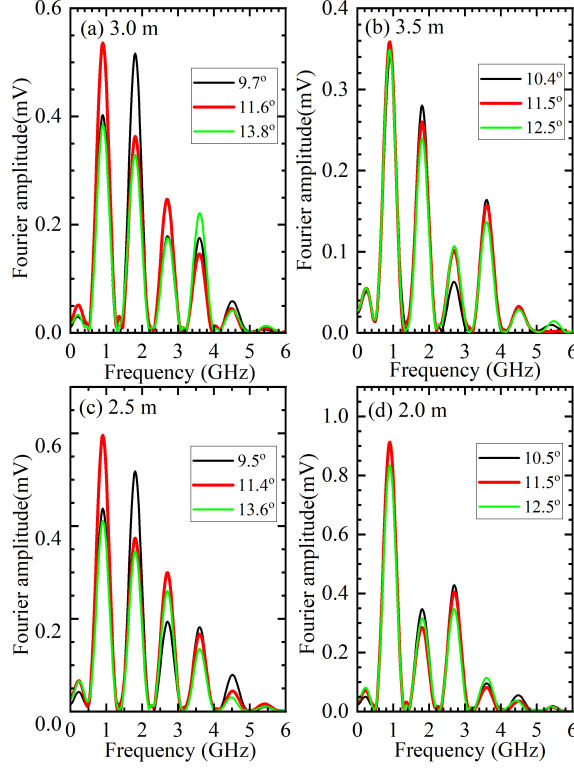


Figure 6.20: Fourier transforms of the detector signal for azimuthal angles (shown in key) on either side of (green and black) and close to or at the target angle of 11.6° (red) and at different antenna to detector distances: (a) 3.0 m, (b) 3.5 m, (c) 2.5 m and (d) 2.0 m.

6.6.3 Experimental Results

Preliminary surveys were carried out by sweeping the transmitter azimuthal angle at several closely spaced distances around the expected focus distance whilst carefully observing the received signal on the oscilloscope or spectrum analyzer. It was found that slight phase-setting errors resulted in actual target coordinates $R_0 \approx 3.00$ m and $\Phi_0 \approx 11.6^\circ$ (*c.f.* planned values of 3.09 m and 11.9°).

Once this focus position was established, the transmitter-to-receiver distance was

fixed at 3.0 m and the oscilloscope trace of the received signal recorded for several fixed antenna azimuthal angles spaced by $\approx 1^\circ$. The results of this procedure are shown in Fig. 6.19; Comparing this with Fig. 6.18(a), it is clear that the broadcast signal (double peak followed by a narrower and then a wider minimum) is only reproduced faithfully at an azimuthal angle of 11.6° (orange, thicker curve). The time-dependent signals for angles 12.8° and 10.5° show distinct differences from the broadcast waveform; one only has to move a few more degrees away from the target angle for any resemblance to the broadcast signal to be lost.

This picture is confirmed by Fourier transforms of the oscilloscope data, shown in Fig. 6.20(a). At an angle of 11.6° (red trace), the expected “triangular” Fourier spectrum [*c.f.* Fig. 6.18(b)] is produced. On moving $\approx \pm 2^\circ$ away from this, the relative amplitudes of the various harmonics of 0.9 GHz change quite dramatically, showing that the frequency content present in the broadcast signal is being scrambled.

Measurements were then repeated at fixed transmitter-to-receiver distances either side of the target distance of $R_0 = 3.0$ m [Fig. 6.20(b)-(d)]. Even azimuthal angles close to the target value (red traces) failed to yield the broadcast “triangular” Fourier spectrum [compare with Fig. 6.20(a) and Fig. 6.18(b)], showing that the frequency content of the original broadcast signal is only reproduced when the distance *and* the azimuthal angle are close to the target values. Fourier transforms taken over wider angular ranges are given in the contour plots of Fig. 6.21, showing that the “triangular” Fourier spectrum is not recovered as one moves farther from the target angle. Fig. 6.22 shows the effect on the time dependence of the received signal caused by keeping the azimuthal angle close to the target value of 11.6° and varying the transmitter-to-receiver distance. Comparing Fig. 6.22 with Fig. 6.18(a), it is clear that the broadcast signal’s time dependence (double peak followed by a narrower and then a wider minimum) is only reproduced faithfully at distances close to the target value of 3.0 m (orange, thicker curve).

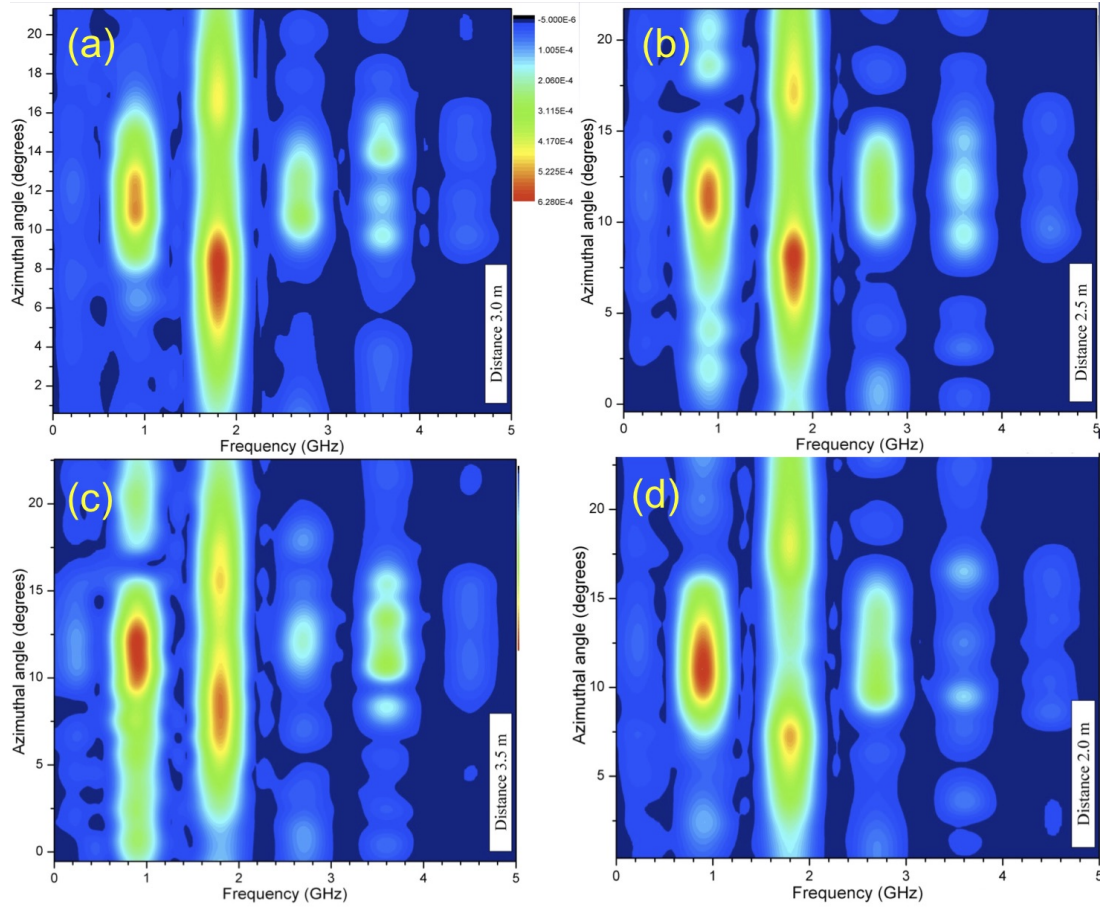


Figure 6.21: Fourier transforms of the detector signal plotted as contour plots versus frequency and azimuthal angle for different antenna to detector distances: (a) 3.0 m, (b) 2.5 m, (c) 3.5 m and (d) 2.0 m.

6.6.4 Implications for Pulsars

It is now clear that a plot of reception time t_P versus emission time t for a point-like source in superluminal motion exhibits ‘plateaux’ at, and very close to, a special polar angle determined by the source’s tangential speed (Fig. 6.1). We have already seen that these plateaux are not, in fact, flat. However, there is a reasonable region of t over which $dt_P/dt \ll 1$, so that a situation similar to that in Fig. 6.15(a) may arise.

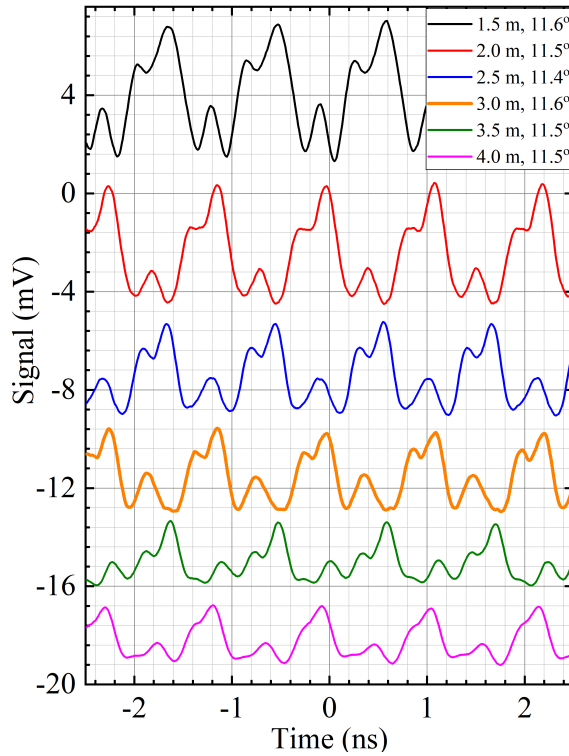


Figure 6.22: Time dependence of the signal received by the detector dipole for azimuthal angle close to the target of 11.6° and for several different antenna-to-detector distances shown in the key. The shape of the transmitted waveform (Fig. 6.18(b)) is only reproduced close to the “target” distance of 3.0 m (orange trace).

Pulsars can potentially emit electromagnetic radiation via many mechanisms [94], including thermal emission along with other processes in their hot plasma atmospheres, and dipole radiation from the rotating magnetic field of the neutron-star core. Why then, given all of these choices, might the pulsed radiation detected on Earth be dominated by the emission from a small volume of superluminal polarization current? The similarity of the plateaux in Fig. 6.1 to the flat line in Fig. 6.15(a) provides an important clue. At and very close to the focus polar angle and over a short window of t_P , the frequency content of all of the emission processes occurring within the rotating polarization-current element will reproduce exactly, and result in a detected signal with greatly enhanced amplitude; the result is similar to coherent emission [12], but via a completely different mechanism. At all other observation

angles and times, radiation from the emission processes will superpose incoherently [compare with Fig. 6.15(b) and (c)], leading to a much diminished amplitude, and scrambled frequency content. The sharp focusing in the time domain at the focus polar angle is likely to allow the radiation produced by the superluminal mechanisms to dominate the pulses. These preliminary qualitative discussions will be verified quantitatively by calculations in Chapter 7. *(Note that this explanation of the brightness of pulsar pulses does not depend on the spurious proposal of non-spherical decay advocated by Ardavan (see Chapter 5).)*

6.7 Results from this Chapter Applied to Pulsars

6.7.1 Fields Predicted by the Liénard-Wiechert Expressions

The previous section has given a plausible reason as to why the pulsed component of pulsar radiation received on Earth will be dominated by the emission from superluminal polarization currents outside the light cylinder. Chapter 2 described how these polarization currents occupy a volume with lengthscales that are small compared to those of the orbit radius, and very small compared to the source-to-observer (on Earth) distance. A legitimate starting point is to use the Liénard-Wiechert expression for the fields, as generalized to include point-like sources in superluminal motion. Therefore, having evaluated the relevant retarded times using the method described in Section 6.5.1, Eq. (6.61) is evaluated numerically and used along with Eq. (6.62) to yield the Liénard-Wiechert fields²¹ at spacetime points (\mathbf{r}_P, t_P) . Typical results for the intensity $I \propto |\mathbf{E}|^2$ are shown in Fig. 6.23 for the plane of the source's orbit (a), and for the limiting cone of the cusp (b), $\theta_P = \sin^{-1} c/v$ (here θ_P is the polar angle of the observer – see Fig. 6.8). Most of the radiated energy occurs on or very close to the cone. As this region of high electromagnetic intensity spirals out

²¹The results described below were published in part as [79].

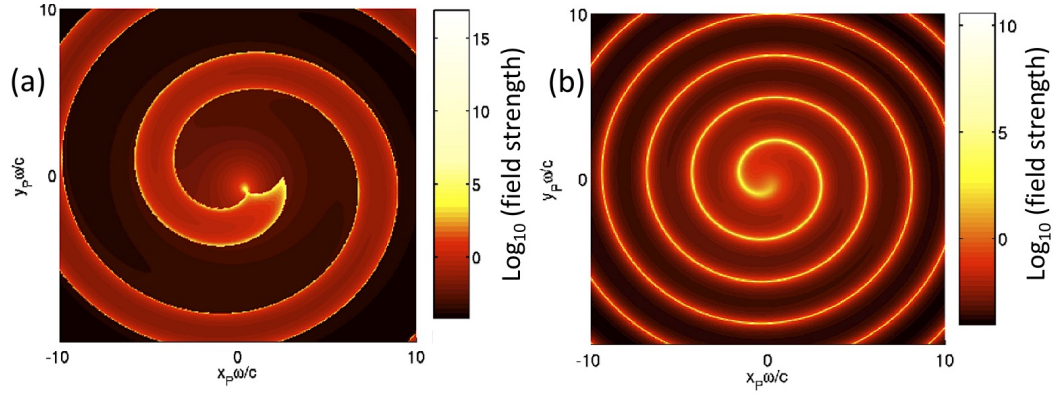


Figure 6.23: (left) Field strength in the plane of the source's orbit, for $\hat{r} = 2.5$. Note the high radiation intensity along the inner edges of the envelope and near the cusp. (right) Relative radiation intensity on the limiting cone of the cusp.

and into the far-field, it sweeps past an observer once for every rotation of the source, delivering a sharp and powerful pulse of electromagnetic radiation. This is in exact correspondence to the ‘spikes’ observed in the temporal focusing plots of Fig. 6.12. For the astronomical scales on which both calculations are carried out, such a pulse has a width of ≤ 1 ns.

Encouragingly, around 50 per cent of the pulsars in the Parkes²² data base [191] exhibit a single pulse; virtually all of the rest display double pulses (see *e.g.*, the data shown in Fig. 2.20), a phenomenon that we will explain in a later section (6.7.2). At first sight, there seems to be a discrepancy in the pulse duration. Typical data [93, 94] (*e.g.*, Fig. 2.17 and Fig. 2.19) suggest that the pulse width is at the order of milliseconds, much longer than the few nanoseconds predicted. However, this apparent width is almost certainly due to instrumental time resolution: In most telescopes, effective sensitivity is the most important consideration, and averaging of very many pulses is used to maximize the signal²³. In other words, time resolution is sacrificed to sensitivity.

²²The Parkes data base is a kind of online telephone directory for all pulsars that have been observed; it contains pulse shape, frequency, flux and distance data for each one.

²³J. Middleditch and T. Hankins, numerous personal communications.

Nevertheless, pulses from the Crab Pulsar are sufficiently intense to permit high time resolution recordings of individual pulses. Once this is accomplished, structure on much shorter timescales reveals itself. Instrumental issues still restrict the apparent widths to microseconds, but members of the observing team (T. Hankins, personal communication) believe that the true duration may be on a nanosecond timescale, in accord with the model predictions plotted in Fig. 6.23 and Fig. 6.12.

From pulse to pulse, the timing and height of these sharp features will probably vary, eventually building up the broader, time-averaged pulses [93,94] (Fig. 2.17 and Fig. 2.19). The latter are very predictable and regular, reflecting the stable rotation of the neutron star. The pulse-to-pulse variation, on the other hand, likely reflects the turbulence and atmospheric variations in the emitting plasma or “weather” [93,94].

6.7.2 Stokes Parameters

The electric field radiated from a compact, superluminal source in a circular orbit is predicted by Eq. (6.61) to have an orientation, or polarization, that changes rapidly with time and location. To visualize this, we employ Stokes parameters, which provide a description of the polarization state of electromagnetic radiation and are widely used in the analysis of astronomical data (see, *e.g.*, pages 25–37 of [190]). Named after Sir George Stokes (1819–1903), these quantities are defined as

$$\begin{aligned} I &= |\mathbf{E}_{\parallel}|^2 + |\mathbf{E}_{\perp}|^2 & \frac{U}{I} &= \frac{2|\mathbf{E}_{\parallel}/\mathbf{E}_{\perp}| \cos \xi}{|\mathbf{E}_{\parallel}/\mathbf{E}_{\perp}|^2 + 1} \\ \frac{Q}{I} &= \frac{|\mathbf{E}_{\parallel}/\mathbf{E}_{\perp}|^2 - 1}{|\mathbf{E}_{\parallel}/\mathbf{E}_{\perp}|^2 + 1} & \frac{V}{I} &= \frac{2|\mathbf{E}_{\parallel}/\mathbf{E}_{\perp}| \sin \xi}{|\mathbf{E}_{\parallel}/\mathbf{E}_{\perp}|^2 + 1}, \end{aligned} \quad (6.86)$$

where the subscripts \parallel and \perp denote directions parallel and perpendicular to an axis chosen to be in the plane transverse to the direction of light propagation given by the Poynting vector $\mathbf{S} = \mathbf{E} \times \mathbf{H}$. The parameter ξ is the phase difference, at the frequency of interest, between \mathbf{E}_{\parallel} and \mathbf{E}_{\perp} . In visualizing and discussing the results

of Eq. (6.61), we use the following.

$$\begin{aligned}
 \text{Degree of circular polarization:} & \quad \frac{V}{I}. \\
 \text{Degree of linear polarization:} & \quad \frac{L}{I} = \frac{\sqrt{Q^2 + U^2}}{I}. \\
 \text{Polarization position angle:} & \quad \frac{1}{2} \tan^{-1} \frac{U}{Q}.
 \end{aligned} \tag{6.87}$$

These represent the standard procedure for depicting pulsar data [93].

Since the field of the rotating source itself rotates rigidly (Chapter 2), an observer at spherical coordinates $(r_P, \varphi_P, \theta_P)$ samples, during each rotation period, the field on the arc that lies at the intersection of the cone $\theta_P = \text{constant}$ and the sphere $r_P = \text{constant}$. For the limiting cone of the cusp, this arc lies partly inside and partly outside the envelope. Owing to the discontinuous change in the strength of the field across the envelope, the intensity of the radiation received along this curve has a pulsed distribution consisting of two sharply peaked components; the closer the opening angle of the cone to $\sin^{-1} c/v$, the smaller the width of this two-component pulse, until, on the cusp itself, the two peaks combine to form a single flash of light.

Moving off the cusp further into the volume enclosed by the two sheets of the envelope ($\theta_{P_C} < \theta_P < 180^\circ - \theta_{P_C}$), Fig. 6.8 would lead one to expect three retarded contributions. This is indeed shown in Fig. 6.24 (a), which depicts the time dependence (in units of pulse window²⁴) of the intensity from each of these retarded time contributions. The total yields a double-peaked structure connected by a saddle, which is very reminiscent of the temporal focusing calculations for a similar location (Fig. 6.12). Thus, the “light curve” of the pulses can be seen to reflect the degree of temporal focusing directly. If one imagines a little rounding of the sharp peaks in Fig. 6.24 (a) due to the limited time-resolution of the observing equipment (see above), the predicted light-curve becomes very similar to the observational data plotted in Fig. 2.20.

²⁴The pulse emission from pulsars is usually confined to a time span of a few per cent of the pulsar period – the “pulse window”.

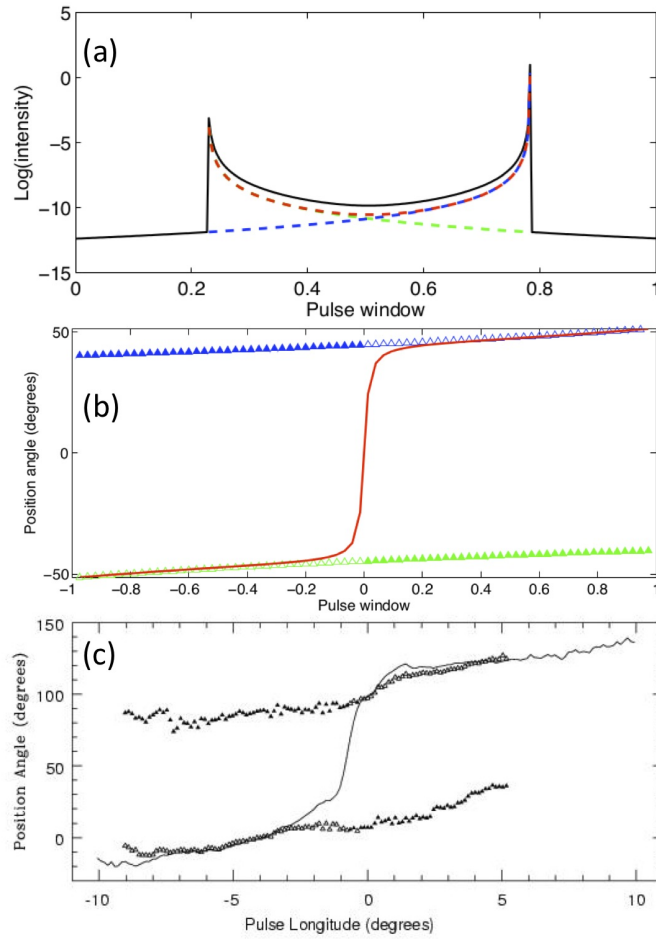


Figure 6.24: (a) The relative strengths of the three radiation modes (green, red, blue) and the total field (black) as experienced by an observer inside the envelope ($\theta_{P_C} < \theta_P < 180^\circ - \theta_{P_C}$) near the cusp on a sphere of large radius. Note that the contribution from the third retarded time (green) is much stronger than that from the first (blue) near the beginning of the pulse, with these rôles reversed near the end. Furthermore, two of the contributions are stronger than the third everywhere except in the middle of the pulse. (b) Calculated polarization position angle of the contributions from the three retarded times (green, red, and blue triangles) are shown relative to one another for the same source and observation arc as above. The position angles of the dominant contributions are depicted using open triangles, those of the weakest contributions with filled triangles. The position angle of the second contribution (red) closely follows the average position angle, bridging the first and third contributions. (c) Observational data [192]: position angle histogram of PSR 2016+28 at 1404MHz. Note the presence of the “mystery third component” bridging the middle of the pulse. The three retarded times (b) reproduce all of the features of the observational data (c), including the 90° swing.

Fig. 6.24 (b) shows the calculated polarization position angles (Eq. (6.87)) of the contributions from the three retarded times across the pulse window, where open triangles denote the dominant and closed ones the weakest contributions, respectively. Two of the radiation modes remain approximately parallel to each other for the duration of the pulse, displaying very little change in their position angles. The third contribution, however, closely follows the average position angle, bridging the remaining two in an S-shaped curve. The numerical results are strikingly similar to the polarization information extracted from 1404 MHz observations of PSR B2016+28 by McKinnon [192] [Fig. 6.24 (c)]. McKinnon notes some intriguing polarization properties while all but conceding that standard models offer no explanation for the observed phenomena:

Polarization observations of the radio emission from PSR B2016+28 at 1404 MHz reveal properties that are consistent with two, very different, interpretations of the pulsar’s viewing geometry. The pulsar’s average polarization properties show a rapid change in position angle (PA) near the pulse center, suggesting that the observer’s sightline nearly intersects the star’s magnetic pole. But single pulse, polarization observations of the pulsar show nearly orthogonal modes of polarization following relatively flat and parallel PA trajectories across the pulse, suggesting that the sightline is far from the pole. Additionally, PA histograms reveal a “modal connecting bridge,” of unknown origin, joining the modal PA trajectories over much of the pulse and following the rapid PA change shown in the average data.

Note that the swing in position angle and the near orthogonality of the two polarization modes are entirely consistent with a superluminal source, which does not require a particular viewing angle with respect to the star’s magnetic pole. Furthermore, the “modal connecting bridge of unknown origin” is likely to be the signal of the second of three contributing times characteristic of a superluminal source.

The polarization position angles for the entire field, scaled by their intensity, are

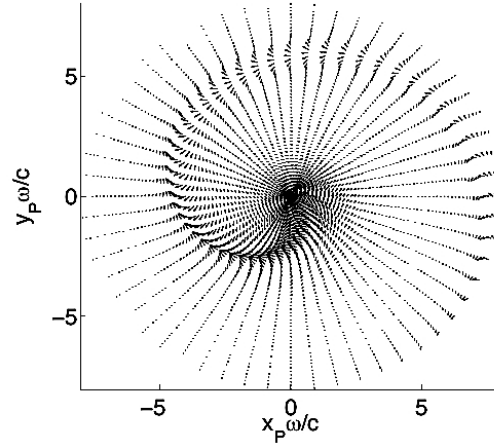


Figure 6.25: Polarization position angles and intensities for the same source as depicted in Fig. 6.23 (right). The arrows represent the direction and magnitude of the electric field for a given field point.

shown in Fig. 6.25. The aesthetically pleasing, filigreed structure seems organic in its appearance, showing that the relative directions and sizes vary subtly over the duration of a pulse.

6.7.3 Intrinsic Characteristics of the Radiation

The previous two sections have shown that both the temporal focusing factor and the Liénard-Wiechert fields of a localized source in superluminal rotation show a consistent set of intrinsic characteristics; before embarking on a calculation of the expected frequency dependence of the emitted radiation (Section 6.7.4), it is useful to summarize them.

1. The radiation is sharply focused along a narrow, rigidly rotating spiral-shaped locus that tracks the cusp of the envelope of the emitted wave fronts.
2. For moderately superluminal sources, it consists of either one or three concurrent polarization modes (depending on the position of the observer relative

to the cusp) that arise from contributions to the field from differing retarded times.

3. At each edge of the pulse, two of the modes are roughly equal in strength, dominating over the third. Near the middle of the pulse, the three modes are of comparable intensity (Fig. 6.24).
4. The position angle of one of the modes, as well as that of the total field, swings across the beam by as much as 180° , while the position angles of the other two modes remain approximately orthogonal (Fig. 6.25).
5. One of the three modes is highly circularly polarized and differs in its sense of polarization from the other two.
6. Two of the modes are highly linearly polarized across the entire pulse.

The similarity of the calculated radiation to observational data from pulsars is quite marked, supporting the plausibility of this model.

6.7.4 Frequency Content of the Radiation

As the frequency dependence of the emitted radiation will in essence be a Fourier transform of its time dependence [12], the sharpness of the peaked structures plotted in Fig. 6.12 and Fig. 6.24 are expected to lead to very broadband emission; as has already been mentioned, pulsar radiation has been observed to cover nearly 30 orders of magnitude in frequency [93, 94]. Another consequence of the sharp peaks will be pulse timings that remain roughly the same at all frequencies. This is in agreement with many observations; see, for instance, Fig. 2.20.

The emission of a small superluminal volume element is in many ways analogous to synchrotron radiation; in the latter, a highly relativistic compact bunch of electrons emits radiation as it undergoes centripetal acceleration [177]. The frequency

spectrum of the radiation is, in essence, the Fourier transform of the electric field of the bunch as it passes the observer. Given that the bunch is both fast and compact, the time-dependent electric field is very sharply peaked: Hence the corresponding synchrotron emission is broadband and rather featureless [177].

In the superluminal model of pulsar emission, the rôle of the electron's electric field is fulfilled by the sharply-peaked electric field of the radiation [see Fig. 6.23 (single pluse) and Fig. 6.24 (double pulse)]. However, since the source itself is superluminal, the Lorentz factor $1/\sqrt{1-v^2/c^2}$ is imaginary. This has the effect of rendering exponential terms in the synchrotron radiation *formulae* oscillatory; in contrast to conventional synchrotron radiation with its smooth spectrum, the point-like superluminal source has an emission spectrum that oscillates as a function of frequency. The radiation spectrum is given by the following set of equations (for a complete mathematical treatment, see [80, 81]) in which $n = 2\pi f/\omega$, and Ai and Ai' are the Airy function and its derivative, respectively:

$$\begin{aligned} \frac{dP_n}{d\Omega_P} \propto & S_1(n)^2 \text{Ai}^2 \left[-\left(\frac{2}{n}\right)^{1/3} \frac{\Omega}{\omega} \right] \\ & + S_2(n)^2 \left(\frac{2}{n}\right)^{2/3} \text{Ai}'^2 \left[-\left(\frac{2}{n}\right)^{1/3} \frac{\Omega}{\omega} \right] \\ & + 2S_3(n)^2 \left(\frac{2}{n}\right)^{1/3} \text{Ai} \left[-\left(\frac{2}{n}\right)^{1/3} \frac{\Omega}{\omega} \right] \\ & \times \text{Ai}' \left[-\left(\frac{2}{n}\right)^{1/3} \frac{\Omega}{\omega} \right], \end{aligned} \quad (6.88)$$

in which

$$S_1(n) = n^{2/3} |K_r| K_{\varphi_0} \left(|\bar{s}_r|^2 + |\bar{s}_\varphi \cos \theta_P - \bar{s}_z \sin \theta_P|^2 \right)^{1/2},$$

$$S_2(n) = n^{2/3} |K_r| K_{\varphi_0} \left(|\bar{s}_\varphi|^2 + |\bar{s}_r|^2 \cos^2 \theta_P \right)^{1/2},$$

and

$$S_3(n) = n^{2/3} |K_r| K_{\varphi_0} \times \left\{ \Im \left[\bar{s}_r^* \cos \theta_P (\bar{s}_\varphi \cos \theta_P - \bar{s}_z \sin \theta_P) - \bar{s}_r \bar{s}_\varphi^* \right] \right\}^{1/2}.$$

Here, $\bar{s}_{r,\varphi,z}$ are the Fourier components of the source densities $s_{r,\varphi,z}|_{\hat{r}=\csc \theta_P}$ with respect to z , \Im and the superscript star denote the imaginary part and the conjugate of a complex variable, respectively, and $(R_P, \theta_P, \varphi_P)$ are the spherical polar coordinates of the observation point P . The function K_{φ_0} is defined by

$$K_{\varphi_0} = (-1)^{n+m} \sin \left(\frac{\pi \Omega}{\omega} \right) \left(\frac{\mu_+}{n - \mu_+} + \frac{\mu_-}{n - \mu_-} \right), \quad (6.89)$$

where

$$\mu_+ = \left\lfloor \frac{\Omega + \omega}{\omega} \right\rfloor \quad \text{and} \quad \mu_- = \left\lfloor \frac{\Omega - \omega}{\omega} \right\rfloor. \quad (6.90)$$

Though Eq. (6.88) appears complicated, the following points should be noted:

1. The function is generic; it should apply to all rotating faster-than-light sources. Therefore the frequency-dependent emission of *all* pulsars should have a qualitatively similar form.
2. Whilst details of the spectrum can be adjusted by varying the relative sizes of S_1, S_2 , and S_3 , the overall appearance of the spectrum is determined by just two adjustable parameters, ω (the angular velocity of the source) and Ω , a frequency at which the pulsar's atmosphere responds resonantly to electromagnetic disturbances. This is thought to correspond to the plasma frequency of the free electrons, $\omega_P = \left(\frac{Ne^2}{\varepsilon_0 m_e} \right)^{1/2}$, where N is the number density of electrons, e the electron charge and m_e denotes the electron mass.
3. If there are subsequent, higher frequency, resonances in the plasma atmosphere, the emission will be enhanced at those frequencies. The only notable resonance expected for a pulsar's atmosphere would be the electron cyclotron frequency, $\omega_c = \frac{eB}{m_e}$, where B is the magnetic field in the emitting region of the atmosphere.

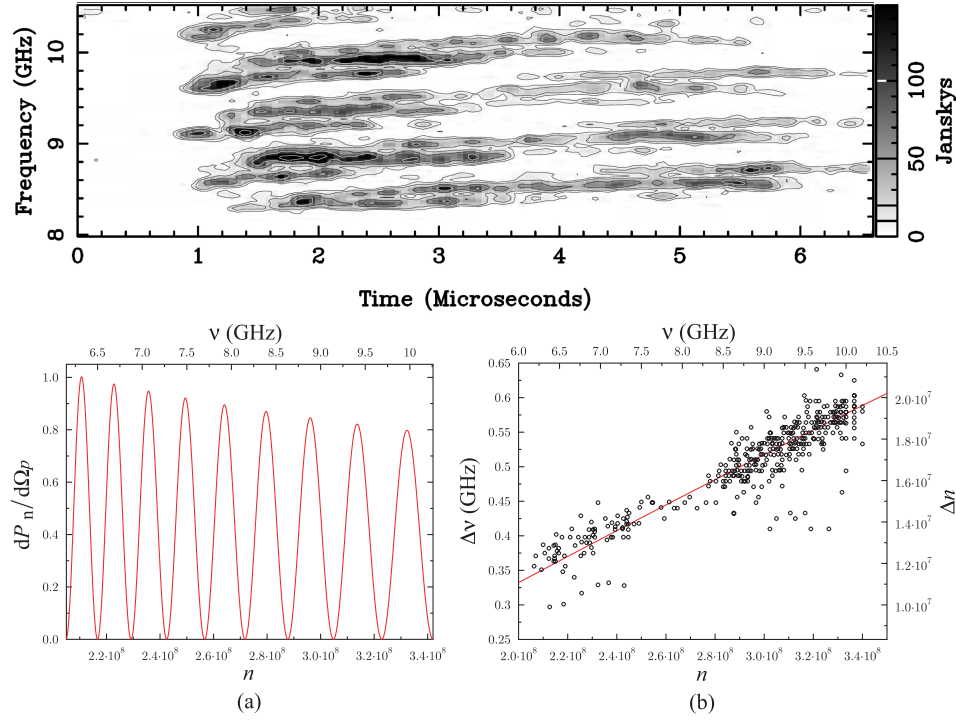


Figure 6.26: (top) Superluminal emission is, analogous to that of synchrotrons, very broad-band, but the Lorentz factor γ is imaginary. Hence, the intensity oscillates as a function of frequency. These oscillations are observed in Crab interpulses. (bottom) Predictions (left) fit observations (right) very well. The predicted oscillations of the spectrum of the emission for $\omega/(2\pi) = 30.3$ Hz and $\Omega/\omega \simeq 1.9 \times 10^4$, shown in (a), have the same spacing as those of the emission bands in the observed spectrum of the Crab pulsar.

Typical fits are shown in Fig. 6.26 and Fig. 6.27. From such models, the electron density and magnetic field at the emitting region can be deduced: Thus far, this has been done for 12 pulsars [80, 81]. The values deduced are in accord with the expectations for neutron stars and their atmospheres. Note particularly that the emission predicted by Eq. (6.88) continues to fit Crab pulsar data to very high photon energies [Fig. 6.27(a)], describing X-rays measured by the FERMI satellite and even pulsed gamma-rays at energies above 100GeV, detected by the VERITAS array of atmospheric Čerenkov telescopes [193]. The authors of [193] note that “the detection cannot be explained on the basis of present (*i.e.*, non-superluminal) pulsar models,” and conclude that the photon spectrum of pulsed emission between 100MeV

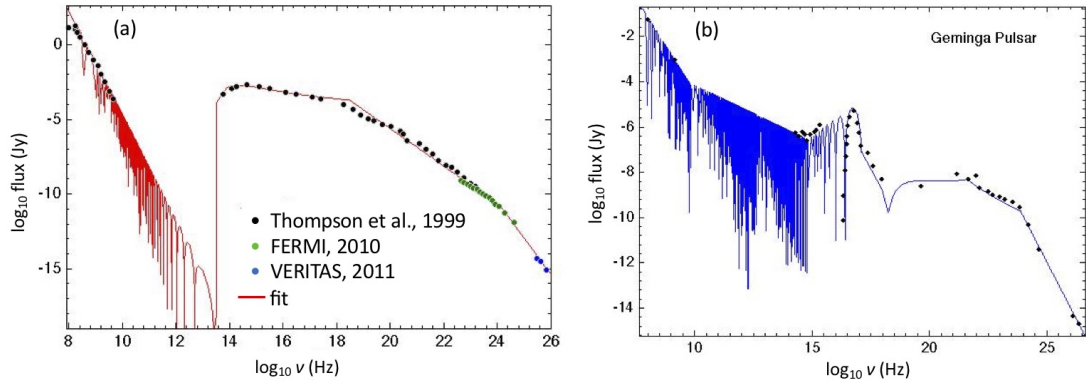


Figure 6.27: (a) The points show observational data (where available) on the spectrum of the Crab pulsar. The curves show the spectral distribution $\log(dP_n/d\Omega_P)$, predicted by Eq. (6.88), versus $\log n$ and $\log \nu$ for $\nu = n\omega/(2\pi) \simeq 30.3n$ Hz and $\Omega/\omega \simeq 1.9 \times 10^4$. In the model, the recovery of intensity at the ultraviolet peak ($\sim 10^{15}$ Hz) is caused by resonant enhancement due to the azimuthal modulation frequency $m\omega/(2\pi) \simeq 3 \times 10^{13}$ Hz. Given the size of the pulsar’s magnetic field, this is likely to be due to cyclotron resonance of electrons. The steepening of the gradient of the spectrum by -1 at 2.4×10^{18} Hz reflects a transition through the Rayleigh distance. (b) A similar fit for the Geminga pulsar. The overall characteristics of pulsar emission are determined by the *superluminal nature* of the source: Only the *details* depend on the pulsar “atmosphere”, *i.e.*, its relative permittivity. One further parameter, namely the electron cyclotron frequency, is used to fit the peak at higher ν .

and 400GeV is best described by a broken power law Fig. 6.27(a). As Fig. 6.27(a) shows, the superluminal model yields a comparable goodness of fit for the high-energy data observed by FERMI and VERITAS, while, at the same time, accounting for the *entire* frequency spectrum of the Crab pulsar, covering more than 18 orders of magnitude.

6.7.5 Summary

It will be clear from the previous sections that the electromagnetic emission of a small superluminal source in a circular orbit is extremely broadband. Moreover, the pulses themselves appear to be coherent. The latter behaviour can be understood in terms of the part of the source that approaches the observer at the speed of light

and no acceleration. Consequently, all of the phase information from this part of the source will collapse onto a single arrival time, making the source appear coherent.

Once one accepts that the emission from superluminal polarization currents dominates pulsar observations from Earth, the superluminal model of pulsar emission explains all the following data: (i) the enigmatic polarization properties (*e.g.*, swing in position angle); (ii) the apparent radiation temperature and pulse shape; (iii) the oscillatory intensity seen in the Crab pulsar at around 10 GHz; and (iv) the broadband radiation spectrum, covering 16-18 orders of magnitude of frequency. The latter point (*i.e.*, a single model fitting the entire broadband spectrum) shows the great advantage of the the present model; though we can show that, *e.g.*, the “slot gap” or “outer gap” models [93,94] give comparable quality fits to parts of the data, *they are only applicable to a small frequency window*, and so a plethora of other (unrelated) phenomena must be invoked to cover the entire spectral range.

Chapter 7

Green's Functions at Work

The beauty of mathematics only shows itself to her more patient followers.

— Maryam Mirzakhani

7.1 Fifth Prologue

The time has come to put the various principles derived or discussed to gainful employment. To demonstrate the validity of the approach, expressions for the time-dependent magnetic vector potential produced by the antennas described in Chapter 3 are derived. In two cases - circular and linear superluminal antennas - the expressions are evaluated numerically and compared with data measured during the experiments of Chapter 3. Before embarking on modelling the antennas, there is a short but necessary introduction to the potentials that will be employed.

7.2 The Secret Lives of Potentials

In the analysis of radiation and scattering problems it is common practice to calculate the time-dependent electric and magnetic field intensity \mathbf{E} and \mathbf{H} due to a charge (or collection of charges) in motion [61]. Generated either directly from the differential equations governing the fields or via the analogous expressions for the potentials, these solutions are applied widely to determine, among other things, the radiation received from apertures and current elements or scattered by structures of various materials and shapes [54, 61].

The *homogeneous* wave equations (Eq. (2.14) and Eq. (2.15)) describe the propagation of electric and magnetic fields at a distance from their source of origin. Close to the varying charges and currents that generate them, however, ρ and \mathbf{J} are nonzero and must be taken into account when solving problems of mathematical physics. Brief calculations (or, alternatively, the consultation of standard texts such as [128–131, 194] in addition to those cited above) show that the *inhomogeneous* electromagnetic wave equations can be stated as

$$\nabla^2 \mathbf{E} - \varepsilon \mu \frac{\partial^2}{\partial t^2} \mathbf{E} = \frac{1}{\varepsilon} \nabla \rho + \mu \frac{\partial}{\partial t} \mathbf{J} \quad (7.1)$$

and

$$\nabla^2 \mathbf{H} - \varepsilon \mu \frac{\partial^2}{\partial t^2} \mathbf{H} = -\nabla \times \mathbf{J}. \quad (7.2)$$

As in Eq. (2.14) and Eq. (2.15), the left hand sides of the two expressions describe the motion of the emitted waves (*i.e.*, the D'Alembert operator¹ or “D'Alembertian” act-

¹Without going into unnecessary detail, we take the D'Alembert operator – represented by a box (\square) – to be the Laplace operator of Minkowski spacetime. In standard coordinates ($t; x, y, z$) it has the form

$$\begin{aligned} \square &= \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \\ &= \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2. \end{aligned}$$

ing on the fields) while the nonzero right hand sides now represent the sources. The latter make evident that electromagnetic waves are generated by gradients in charge density, ρ , and time-varying or circulating current densities, \mathbf{J} . While it is possible to obtain the fields for a specific boundary-value problem directly by integration over the explicit expression for \mathbf{J} , the source terms are often inconveniently complicated. Hence, it is more common (and, generally, much simpler) to calculate the electromagnetic potentials first and derive the fields later by differentiation [54, 61].

In the simplest of terms, a potential is taken to be a function whose derivative yields a field [53]. Hence, fields are associated with *forces*, potentials with *energy*. \mathbf{A} and Φ , the magnetic vector and electric scalar potential, are used most widely, whereas the Hertz vector potentials Π_e and Π_m , from which both Φ and \mathbf{A} can be found by taking derivatives, are rarely mentioned more than *obiter* in standard textbooks (and, if so, as end-of-chapter exercise for the diligent student [54, 61]) although many classical radiation and propagation problems can be solved elegantly and concisely by using them [195].

Since the magnetic flux density is always solenoidal (see Gauss' law of magnetism (Eq. (2.3)) we may substitute $\nabla \times \mathbf{A}$ for \mathbf{B} where \mathbf{A} , as before, is the magnetic vector potential – an arbitrary vector-valued function. Hence, the divergence of the curl of \mathbf{A} is zero ($\nabla \cdot (\nabla \times \mathbf{A}) = 0$) and the magnetic field intensity can now be defined as

$$\mathbf{H} = \frac{1}{\mu} \mathbf{B} = \frac{1}{\mu} (\nabla \times \mathbf{A}). \quad (7.3)$$

This allows us to state Faraday's law (Eq. (2.4)) as

$$\nabla \times \mathbf{E} = -\mu \frac{\partial}{\partial t} \mathbf{H} = -\frac{\partial}{\partial t} (\nabla \times \mathbf{A}) \quad (7.4)$$

and, in consequence,

$$\nabla \times \mathbf{E} + \frac{\partial}{\partial t} (\nabla \times \mathbf{A}) = \nabla \times \mathbf{E} + \nabla \times \frac{\partial}{\partial t} \mathbf{A} = 0. \quad (7.5)$$

Use of the distributive property (*i.e.*, $\nabla \times (\mathbf{A} + \mathbf{B}) = \nabla \times \mathbf{A} + \nabla \times \mathbf{B}$) results in

$$\nabla \times \left(\mathbf{E} + \frac{\partial}{\partial t} \mathbf{A} \right) = 0 \quad (7.6)$$

while applying the vector identity $\nabla \times (-\nabla\Phi) = 0$ leads to

$$-\nabla\Phi = \mathbf{E} + \frac{\partial}{\partial t} \mathbf{A} \quad (7.7)$$

where $\nabla\Phi$ represents the gradient of some position-dependent scalar function, namely the electric scalar potential. We notice that both the electric and magnetic field intensity can be expressed in terms of the two potentials \mathbf{A} and Φ :

$$\mathbf{E} = -\nabla\Phi - \frac{\partial}{\partial t} \mathbf{A} \quad (7.8)$$

and

$$\mathbf{H} = \frac{1}{\mu}(\nabla \times \mathbf{A}). \quad (7.9)$$

The definition of \mathbf{E} and \mathbf{H} with regard to the potentials (Eq. (7.8) and Eq. (7.9)) satisfies the two homogeneous Maxwell's equations (*i.e.*, $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{E} + \partial/\partial t \mathbf{B} = 0$) identically. To recast Eq. (2.2) and Eq. (2.5) we return to Eq. (7.3), take the curl on both sides of the equation, and make use of Ampère's law:

$$\begin{aligned} \mu(\nabla \times \mathbf{H}) &= \mu\mathbf{J} + \varepsilon\mu\frac{\partial}{\partial t} \mathbf{E} = \nabla \times \nabla \times \mathbf{A} \\ &= \nabla(\nabla \cdot \mathbf{A}) - \nabla \cdot \nabla \mathbf{A} \\ &= \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} \end{aligned} \quad (7.10)$$

where $\mathbf{E} = -\nabla\Phi - \partial/\partial t \mathbf{A}$, as defined above. Hence,

$$\begin{aligned} \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} &= \mu\mathbf{J} + \varepsilon\mu\frac{\partial}{\partial t} \left(-\nabla\Phi - \frac{\partial}{\partial t} \mathbf{A} \right) \\ &= \mu\mathbf{J} - \varepsilon\mu\nabla\left(\frac{\partial}{\partial t} \Phi\right) - \varepsilon\mu\frac{\partial^2}{\partial t^2} \mathbf{A}, \end{aligned} \quad (7.11)$$

and, after re-arranging terms, we arrive at the wave equation

$$\nabla^2 \mathbf{A} - \varepsilon\mu\frac{\partial^2}{\partial t^2} \mathbf{A} = -\mu\mathbf{J} + \nabla\left(\nabla \cdot \mathbf{A} + \varepsilon\mu\frac{\partial}{\partial t} \Phi\right). \quad (7.12)$$

Thus, the quartet of equations (2.2) – (2.5) has been whittled down to two – if coupled – expressions. Elementary electromagnetism (and basic calculus), however,

suggest that there must be some freedom in the choice of Φ and \mathbf{A} : While \mathbf{E} and \mathbf{H} are unique and impervious to change, they are both defined in terms of *derivatives* of the potentials, which is to say that there exists an infinite family of possible potentials that will all lead to the same fields [53].

Since \mathbf{H} is defined through Eq. (7.3) in terms of \mathbf{A} , the vector potential is arbitrary to the extent that the gradient of some scalar function Λ can be added. In consequence, \mathbf{H} is left unchanged by the transformation

$$\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla \Lambda, \quad (7.13)$$

a technique commonly called *gauge transformation* of the vector potential. A similar expression can be used to transform the scalar potential while leaving \mathbf{E} invariant, namely

$$\Phi \rightarrow \Phi' = \Phi - \frac{\partial}{\partial t} \Lambda. \quad (7.14)$$

It should be evident that, if we choose a set of potentials that satisfy the Lorenz condition²

$$\nabla \cdot \mathbf{A} - \frac{1}{c^2} \frac{\partial}{\partial t} \Phi = 0, \quad (7.15)$$

we are left with an inhomogeneous wave equation for the vector potential that can be written in the familiar form

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{A} - \nabla^2 \mathbf{A} = \mu \mathbf{J}. \quad (7.16)$$

Likewise,

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \Phi - \nabla^2 \Phi = \frac{\rho}{\epsilon}. \quad (7.17)$$

It is worthwhile to note that Eq. (7.16) and Eq. (7.17), along with condition Eq. (7.15), form a set of equations equivalent to Maxwell's quartet Eq. (2.2)-Eq. (2.5), as observed by Ludvig Lorenz and his colleagues [53].

²Named after Ludvig Valentin Lorenz (1829–1891), a Danish mathematician and physicist. Not to be confused with the *Dutch* physicist Hendrik Lorentz.

Expressions for the *retarded* potentials, which we will discuss in detail in the sections to come, describe the electromagnetic potentials due to a time-varying current or charge distribution *in the past*. On that account, they reflect the causal behavior associated with a wave disturbance: The signal takes a finite time, corresponding to its propagation speed, c , to travel the distance from \mathbf{r} to \mathbf{r}_P , where an effect is produced or measured at time t_P . The same principle underlies the derivation of special relativity, such that all rigorously-defined descriptions of electromagnetism – including Maxwell's equations – are necessarily relativistic [53, 196]. Of course the Green's function for the problem (see Chapter 4) automatically select the correct time in the past, ensuring that causality is not violated. We now put these ideas into practice.

7.3 Magnetic Vector Potentials: Linear and Quasi-Linear Superluminal Antennas

7.3.1 Constant Motion

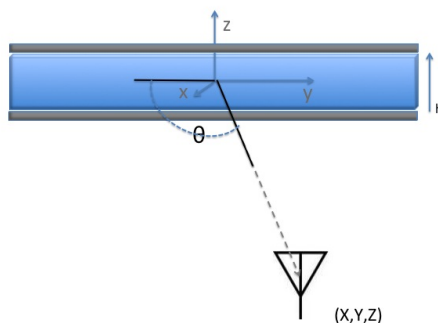


Figure 7.1: Simple model of a rectilinear superluminal antenna centered at the origin of a Cartesian coordinate system. The receiver or detector is shown at position (X, Y, Z) .

Fig. 7.1 shows a simple model of a linear superluminal antenna such as TD 2 (see Fig. 3.1). It consists of a strip of dielectric, the material to be polarized by the applied electric field, shaded blue. To match the alumina used in the practical antennas, we assume a relative permittivity $\varepsilon_r = 10$ and permeability $\mu = 4\pi \times 10^{-7}$ H/m throughout this chapter. The center of the antenna lies on the origin of a Cartesian coordinate system (x, y, z) with unit vectors $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$. As described in Chapter 3, metal electrodes (shaded gray in Fig. 7.1) are placed on the top and bottom of the dielectric solid, forming what is, in effect, a series of capacitors (Fig. 3.2).

Chapter 3 also described how simulations of superluminal emitters performed with off-the-shelf electromagnetic software packages show that fringing fields of adjacent electrode pairs lead to a voltage phase that varies slightly under the electrode; *i.e.*, the phase is more smoothly varying with y than the discrete arrangement of electrodes would suggest. Therefore, we simulate this voltage phase variation as a continuous phenomenon, and the electric potential can be expressed as

$$V(y, t) = \frac{V_0}{2} e^{i(\omega t - ky)} \quad \text{for } z = +\frac{h}{2}, \quad (\text{upper electrodes}) \quad (7.18)$$

$$\text{and} \quad V(y, t) = -\frac{V_0}{2} e^{i(\omega t - ky)} \quad \text{for } z = -\frac{h}{2}, \quad (\text{lower electrodes}) \quad (7.19)$$

where h is the height of the strip, ω the frequency and k , the propagation constant (or wave vector), is defined as $v = \frac{\omega}{k}$, where v is the phase velocity of the polarization current, also referred to as the source speed. The voltage in the dielectric as a function of time, t , and Cartesian coordinates y and z can hence be given as

$$V(y, z, t) = \frac{V_0}{2(h/2)} z e^{i(\omega t - ky)} = \frac{V_0}{h} z e^{i(\omega t - ky)}. \quad (7.20)$$

The electric field may now be defined as

$$\mathbf{E}(y, z, t) = -\nabla V \quad (7.21)$$

$$= -\frac{V_0}{h} \left(-ikz e^{i(\omega t - ky)} \hat{\mathbf{e}}_2 + e^{i(\omega t - ky)} \hat{\mathbf{e}}_3 \right) \quad (7.22)$$

$$= -\frac{V_0}{h} e^{i(\omega t - ky)} \left(-ikz \hat{\mathbf{e}}_2 + \hat{\mathbf{e}}_3 \right). \quad (7.23)$$

For any dielectric material, an applied electric field \mathbf{E} creates internal electric dipole moments, resulting in a polarization \mathbf{P} ; in a linear medium,

$$\mathbf{P} = \varepsilon_0 \chi \mathbf{E} = \varepsilon_0 (\epsilon_r - 1), \quad (7.24)$$

where χ is the electric susceptibility and $\varepsilon_0 = 8.85418782 \times 10^{-12}$ F/m is the permittivity of free space. Then,

$$\mathbf{P} = \varepsilon_0 (\epsilon_r - 1) \mathbf{E} = -\varepsilon_0 (\epsilon_r - 1) \frac{V_0}{h} e^{i(\omega t - ky)} \left(-ikz \hat{\mathbf{e}}_2 + \hat{\mathbf{e}}_3 \right) \quad (7.25)$$

and the polarization current density is

$$\frac{\partial \mathbf{P}}{\partial t} = -\varepsilon_0 (\epsilon_r - 1) \frac{V_0}{h} \omega e^{i(\omega t - ky)} \left(kz \hat{\mathbf{e}}_2 + i \hat{\mathbf{e}}_3 \right) = \mathbf{J}. \quad (7.26)$$

As has been shown already in the discussion of Eq. (7.16) and Eq. (7.17) (repeated here for emphasis), in the presence of sources, the inhomogeneous wave equations that govern the scalar and vector potentials Φ and \mathbf{A} are

$$\begin{aligned} \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \Phi - \nabla^2 \Phi &= \frac{1}{\varepsilon_0} \rho \\ \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{A} - \nabla^2 \mathbf{A} &= \mu_0 \mathbf{J} \end{aligned}$$

where ρ - the free charge density - and \mathbf{J} can be thought of as the sources that generate the potentials Φ and \mathbf{A} from which, in turn, the fields \mathbf{E} and \mathbf{H} may be calculated. If these source densities are known, then the causal (*e.g.*, Green's function) solutions of the above wave equations are given by

$$\Phi(\mathbf{r}_P, t_P) = \frac{1}{4\pi\varepsilon_0} \int_V \frac{\rho(\mathbf{r}, t)}{|\mathbf{r}_P - \mathbf{r}|} d^3\mathbf{r} \quad (7.27)$$

$$\mathbf{A}(\mathbf{r}_P, t_P) = \frac{\mu_0}{4\pi} \int_V \frac{\mathbf{J}(\mathbf{r}, t)}{|\mathbf{r}_P - \mathbf{r}|} d^3\mathbf{r}, \quad (7.28)$$

which we recognize immediately as a special form of Eq. (4.99). These expressions describe the scalar or vector potential for electromagnetic fields of a time-varying

current or charge distribution, where \mathbf{J} and ρ are the retarded source densities. $\mathbf{r}_P = (X, Y, Z)$ is the observer's position and $\mathbf{r} = (x, y, z)$ denotes a point within the source volume. The retarded or source time – denoted t as before – is related to the time t_P at the observation point (*e.g.*, the receiving antenna) by

$$t = t_P - R/c, \quad (7.29)$$

where

$$R^2 = (X - x)^2 + (Y - y)^2 + (Z - z)^2 \quad (7.30)$$

and c denotes the speed of electromagnetic waves in the medium (assumed to be uniform) between the source and the observer ³. Substituting Eq. (7.26) into the expression for the retarded vector potential (Eq. (7.28)) gives

$$\mathbf{A}(\mathbf{r}_P, t_P) = \frac{\mu_0}{4\pi} J_0 \int_V \frac{e^{i(\omega t - ky)} (kz\hat{\mathbf{e}}_2 + i\hat{\mathbf{e}}_3)}{|\mathbf{r}_P - \mathbf{r}|} d^3\mathbf{r} \quad (7.31)$$

where $J_0 = -\omega\varepsilon_0(\varepsilon_r - 1)\frac{V_0}{h}$ is the current density. Using relation Eq. (7.29), Eq. (7.31) can be re-written as

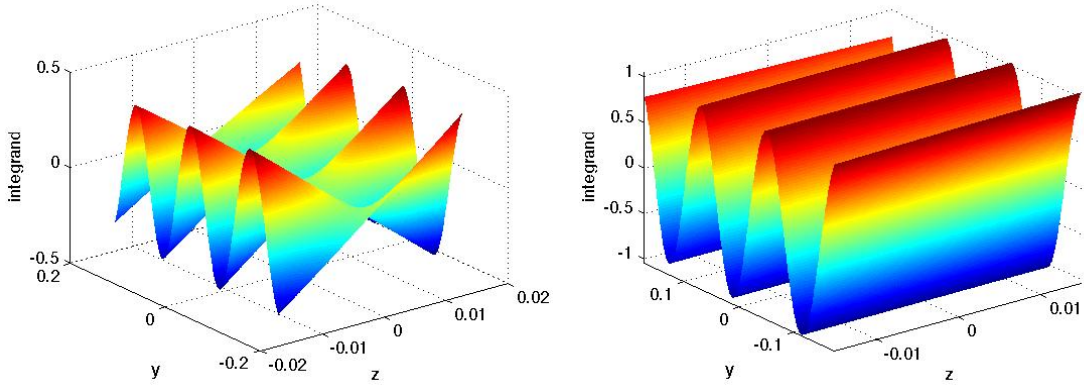
$$\mathbf{A}(\mathbf{r}_P, t_P) = \frac{\mu_0}{4\pi} J_0 e^{i\omega t_P} \int_V \frac{e^{-i(\omega R/c + ky)} (kz\hat{\mathbf{e}}_2 + i\hat{\mathbf{e}}_3)}{|\mathbf{r}_P - \mathbf{r}|} d^3\mathbf{r} \quad (7.32)$$

As Fig. 7.2 shows, the two integrands are smooth and the integration can therefore be performed using a routine method such as recursive adaptive Lobatto quadrature, which is more efficient than adaptive Simpson quadrature for smooth integrands at high accuracies (*e.g.*, low error tolerances).

7.3.2 Implementing Acceleration I

We now wish to add non-uniform acceleration along the y-axis to the model proposed in the previous section. Hence, instead of the propagation constant (wavevector) k ,

³As stated above, the δ -function within the Green's function of the problem automatically selects the correct retarded time.

Figure 7.2: The two integrands of Eq. (7.31) as functions of y and z .

we choose the function $p(y)$ and its derivative to represent the source's speed and acceleration, respectively.

The electric potential at the two electrode arrays is now given by

$$V(y, z; t) = \frac{V_0}{2(h/2)} z e^{i(\omega t - p)} = \frac{V_0}{h} z e^{i(\omega t - p)} \quad (7.33)$$

and the resulting electric field is of the form

$$\mathbf{E}(y, z; t) = -\nabla V \quad (7.34)$$

$$= -\frac{V_0}{h} \left[\frac{\partial}{\partial y} z e^{i(\omega t - p)} \hat{\mathbf{e}}_2 + \frac{\partial}{\partial z} z e^{i(\omega t - p)} \hat{\mathbf{e}}_3 \right] \quad (7.35)$$

$$= -\frac{V_0}{h} \left[-iz \frac{\partial}{\partial y} p e^{i(\omega t - p)} \hat{\mathbf{e}}_2 + e^{i(\omega t - p)} \hat{\mathbf{e}}_3 \right] \quad (7.36)$$

$$= \frac{V_0}{h} e^{i(\omega t - p)} \left[-iz \frac{\partial}{\partial y} p \hat{\mathbf{e}}_2 + \hat{\mathbf{e}}_3 \right]. \quad (7.37)$$

The polarization current may then be calculated as

$$\mathbf{P} = \varepsilon_0(\varepsilon_r - 1)\mathbf{E} \quad (7.38)$$

$$= -\varepsilon_0(\varepsilon_r - 1) \frac{V_0}{h} e^{i(\omega t - p)} \left[-iz \frac{\partial}{\partial y} p \hat{\mathbf{e}}_2 + \hat{\mathbf{e}}_3 \right] \quad (7.39)$$

and its density as

$$\frac{\partial \mathbf{P}}{\partial t} = -\varepsilon_0(\varepsilon_r - 1) \frac{V_0}{h} \omega e^{i(\omega t - p)} \left[z \frac{\partial}{\partial y} p \hat{\mathbf{e}}_2 + i \hat{\mathbf{e}}_3 \right] \quad (7.40)$$

$$= J_0 e^{i(\omega t - p)} \left[z \frac{\partial}{\partial y} p \hat{\mathbf{e}}_2 + i \hat{\mathbf{e}}_3 \right] \quad (7.41)$$

$$= \mathbf{J}, \quad (7.42)$$

where $J_0 = -\omega \varepsilon_0 (\varepsilon_r - 1) \frac{V_0}{h}$, as before, is the current density. Hence, the expression for the vector potential \mathbf{A} becomes

$$\mathbf{A}(\mathbf{r}_P, t_P) = \frac{\mu_0}{4\pi} J_0 \int_V e^{i(\omega t - p)} \left[z \frac{\partial p}{\partial y} \hat{\mathbf{e}}_2 + i \hat{\mathbf{e}}_3 \right] \quad (7.43)$$

As an example, one of the $p(y)$ chosen for outdoor tests was

$$p(y) = \frac{\omega}{\omega_0} \sin^{-1} \frac{\omega_0 y}{u} \quad (7.44)$$

with the derivative

$$\frac{\partial p}{\partial y} = \frac{\partial}{\partial y} \left(\frac{\omega}{\omega_0} \sin^{-1} \frac{\omega_0 y}{u} \right) = \frac{\omega}{u} \frac{1}{\sqrt{1 - \left(\frac{\omega_0 y}{u} \right)^2}}. \quad (7.45)$$

A related acceleration scheme will be studied in depth in Section 7.7.

7.3.3 Implementing Acceleration II

In some antenna tests, it was more convenient to implement a position-dependent wavevector for the polarization current. Consequently, the voltage in the dielectric is very similar in form to that of the constant-speed case:

$$V = \frac{V_0 z}{h} e^{i(\omega t - k(y)y)}. \quad (7.46)$$

However, instead of a constant k , we have a position-dependent wavevector $k(y)$.

The mathematical operations are similar, but the y dependence of $k(y)$ produces an extra term in the polarization current density:

$$\frac{\partial \mathbf{P}}{\partial t} = -\varepsilon_0 (\varepsilon_r - 1) \omega \frac{V_0}{h} e^{i(\omega t - k(y)y)} \left[\hat{\mathbf{e}}_2 z \left(k(y) + y \frac{dk(y)}{dy} \right) + i \hat{\mathbf{e}}_3 \right]. \quad (7.47)$$

7.3.4 Arced Antenna

The prototype arced superluminal antenna (TD 3) is shown in the right-hand side of Fig. 3.1. A position \mathbf{r} within the dielectric is defined by Cartesian coordinates (x, y, z) . The upper and lower z limits of the dielectric are z_+ and z_- , both dependent on y and defined by

$$z_+ = (r_0^2 - y^2)^{\frac{1}{2}} - r_0 + \frac{h}{2} \quad \text{and} \quad z_- = (r_0^2 - y^2)^{\frac{1}{2}} - r_0 - \frac{h}{2}. \quad (7.48)$$

Here y is the horizontal distance from the central plane of the dielectric, h is its height and r_0 is its radius of curvature. The z -coordinate of the center of the dielectric is given by

$$z = (r_0^2 - y^2)^{\frac{1}{2}} - r_0. \quad (7.49)$$

Just as in (linear) TD 2, the electrodes above and below the dielectric apply equal and opposite voltages V_+ and V_- respectively, approximated by continuous functions

$$V_+ = \frac{V_0}{2} e^{i(\omega t - ky)} \quad \text{and} \quad V_- = -\frac{V_0}{2} e^{i(\omega t - ky)}, \quad (7.50)$$

where V_0 , ω and k are constants. Hence, the voltage at any point in the dielectric is

$$\begin{aligned} V(y, z, t) &= \frac{V_0}{h} \left(z - \left[(r_0^2 - y^2)^{\frac{1}{2}} - r_0 \right] \right) e^{i(\omega t - ky)} \\ &= \frac{V_0}{h} \left[z e^{i(\omega t - ky)} - s(y) e^{i(\omega t - ky)} \right], \end{aligned} \quad (7.51)$$

where $s(y) = (r_0^2 - y^2)^{\frac{1}{2}} - r_0$. The electric field \mathbf{E} is given by

$$\mathbf{E} = -\nabla V = -\frac{V_0}{h} e^{i(\omega t - ky)} \left[\hat{\mathbf{e}}_2 \left(-ik(z - s(y)) - \frac{ds}{dy} \right) + \hat{\mathbf{e}}_3 \right], \quad (7.52)$$

where

$$\frac{ds}{dy} = -\frac{y}{(r_0^2 - y^2)^{\frac{1}{2}}}. \quad (7.53)$$

As before $\mathbf{P} \equiv \varepsilon_0(\varepsilon_r - 1)\mathbf{E}$, so the polarization current is

$$\frac{\partial \mathbf{P}}{\partial t} = -\varepsilon_0(\varepsilon_r - 1)\omega \frac{V_0}{h} e^{i(\omega t - ky)} \left[\hat{\mathbf{e}}_2 \left(k(z - s(y)) - i \frac{ds}{dy} \right) + i \hat{\mathbf{e}}_3 \right]. \quad (7.54)$$

7.3.5 Other Non-Circular Cases

To calculate the radiation fields, we need the magnetic vector potential, $\mathbf{A}(X, Y, Z, t_P)$, at the observation point (X, Y, Z) and observation time t_P . This is obtained from Eq. (7.28), repeated here for ease of reference:

$$\mathbf{A}(X, Y, Z, t_P) = \frac{\mu_0}{4\pi} \int \int \int \frac{\mathbf{J}(x, y, z, t)}{|\mathbf{r}_P - \mathbf{r}|} dx dy dz.$$

Here, $\mu_0 (\equiv 4\pi \times 10^{-7}$ in SI units) is the permeability of free space; $\mathbf{r}_P = (X, Y, Z)$ is the observer position, t is the retarded time and $\mathbf{r} = (x, y, z)$ is a point within the dielectric. For a dielectric antenna, the current density \mathbf{J} is equivalent to the polarization current density

$$\mathbf{J} \equiv \frac{\partial \mathbf{P}}{\partial t}. \quad (7.55)$$

Inserting the various expressions derived above for $\frac{\partial \mathbf{P}}{\partial t}$, it transpires that the integrals that give \mathbf{A} at a remote point (X, Y, Z) and time t_P have the same general form:

$$\mathbf{A} = \frac{\mu_0}{4\pi} P_0 \omega e^{i\omega t_P} \int_V e^{-i\phi} \left[\frac{(a - ib)\hat{\mathbf{e}}_2 + i\hat{\mathbf{e}}_3}{R} \right] dx dy dz, \quad (7.56)$$

where

$$P_0 = -\frac{\varepsilon_0(\varepsilon_r - 1)V_0}{h}, \quad (7.57)$$

$$R = [(X - x)^2 + (Y - y)^2 + (Z - z)^2]^{\frac{1}{2}} \quad (7.58)$$

and

$$\phi = \frac{\omega R}{c} + p. \quad (7.59)$$

Expressions for p , a and b for the various cases discussed thus far are given in Table 7.1. The term ϕ results from a substitution for the retarded time

$$t = t_P - \frac{R}{c} \quad (7.60)$$

Condition	p	a	b
Constant speed v	$ky \equiv \frac{\omega y}{v}$	kz	0
LSLA with accel.	$k(y)y \equiv \frac{\omega y}{v(y)}$	$(k(y) + \frac{dk(y)}{dy})z$	0
“Arcsine” accel.	$\frac{\omega}{\omega_0} \arcsin(\frac{\omega_0 y}{u})$	$\frac{\omega z}{u} (1 - \frac{\omega_0^2 y^2}{u^2})^{-\frac{1}{2}}$	0
Curved arc	$ky \equiv \frac{\omega y}{v}$	$z - s(y) =$ $z - \left[(r_0^2 - y^2)^{\frac{1}{2}} - r_0 \right]$ NB z limits of integration are a function of y ; See Eq. (7.48).	$\frac{ds}{dy} = -\frac{y}{(r_0^2 - y^2)^{1/2}}$

Table 7.1: Parameters for the magnetic vector potential integral, Eq. (7.56). Coordinates x , y and z denote positions within the dielectric.

and the varying phase of the polarization current density due to its motion.

Note that the y component (*i.e.*, parallel to $\hat{\mathbf{e}}_3$) is entirely due to the motion of the polarization current. In addition, there are no frequencies beside the animation frequency ω present in the radiation received at the observation point (X, Y, Z) .

7.4 Magnetic Vector Potentials: Circular Superluminal Antenna

7.4.1 Simulation of TD 1

To treat the circular superluminal antenna TD 1 (left frame of Fig. 3.1), it is useful to define positions \mathbf{r} within the dielectric using the cylindrical polar coordinates (ρ, ψ, x) with their origin at the geometrical center of the antenna. The coordinate x is defined as before;

$$z = \rho \cos \psi \quad \text{and} \quad y = \rho \sin \psi. \quad (7.61)$$

Owing to the need to analyze the emitted radiation using detectors sensitive to linear polarization, it is useful to retain the Cartesian unit vectors used before; the unit vectors of the CSLA dielectric are

$$\hat{\rho} = \hat{\mathbf{e}}_2 \sin \psi + \hat{\mathbf{e}}_3 \cos \psi, \quad (7.62)$$

$$\hat{\psi} = \hat{\mathbf{e}}_2 \cos \psi - \hat{\mathbf{e}}_3 \sin \psi \quad (7.63)$$

and

$$\hat{\mathbf{x}} = \hat{\mathbf{e}}_1. \quad (7.64)$$

The gradient operator in this coordinate system is

$$\nabla = \hat{\rho} \frac{\partial}{\partial \rho} + \hat{\psi} \frac{1}{\rho} \frac{\partial}{\partial \psi} + \hat{\mathbf{x}} \frac{\partial}{\partial x}. \quad (7.65)$$

Finally, the voltage in the dielectric is

$$V = \frac{V_0}{\Delta \rho} (\rho - \rho_0) e^{i(\omega t - m\psi)}, \quad (7.66)$$

where ρ_0 is the mean radius of the dielectric, $\Delta \rho$ is its radial extent and $\frac{\omega}{m}$ is the angular velocity of the polarization current, m being a non-zero integer. The polarization \mathbf{P} is

$$\mathbf{P} \equiv \varepsilon_0(\varepsilon_r - 1)\mathbf{E} = -\varepsilon_0(\varepsilon_r - 1)\nabla V = -\frac{V_0}{\Delta \rho} \varepsilon_0(\varepsilon_r - 1) \left\{ \hat{\rho} - im \left[\frac{\rho - \rho_0}{\rho} \right] \hat{\psi} \right\} e^{i(\omega t - m\psi)}. \quad (7.67)$$

Differentiating with respect to (retarded) time t and rearranging in terms of the Cartesian unit vectors, we obtain

$$\begin{aligned} \frac{\partial \mathbf{P}}{\partial t} = & -\frac{V_0}{\Delta \rho} \varepsilon_0(\varepsilon_r - 1) \omega \times \\ & \left\{ \hat{\mathbf{e}}_2 \left(i \sin \psi + m \left[\frac{\rho - \rho_0}{\rho} \right] \cos \psi \right) + \hat{\mathbf{e}}_3 \left(i \cos \psi - m \left[\frac{\rho - \rho_0}{\rho} \right] \sin \psi \right) \right\}. \end{aligned} \quad (7.68)$$

Finally, we need to evaluate the following integral to give $\mathbf{A}(X, Y, Z, t_P)$:

$$\mathbf{A} = \frac{\mu_0}{4\pi} P_0 \omega e^{i\omega t_P} \times \int_V e^{-i(\frac{\omega R}{c} + m\psi)} \left[\frac{\hat{\mathbf{e}}_2(i \sin \psi + \xi \cos \psi) + \hat{\mathbf{e}}_3(i \cos \psi - \xi \sin \psi)}{R} \right] \rho \, d\rho d\psi dx, \quad (7.69)$$

with

$$P_0 = -\varepsilon_0(\varepsilon_r - 1) \frac{V_0}{\Delta\rho}, \quad (7.70)$$

$$\xi = m \left[\frac{\rho - \rho_0}{\rho} \right] \quad (7.71)$$

and

$$R = [(X - x)^2 + (Y - \rho \sin \psi)^2 + (Z - \rho \cos \psi)^2]^{\frac{1}{2}}. \quad (7.72)$$

The volume of integration is

$$\rho_0 - \frac{\Delta\rho}{2} \leq \rho \leq \rho_0 + \frac{\Delta\rho}{2}, \quad 0 \leq \psi \leq 2\pi, \quad \text{and} \quad -\frac{\Delta x}{2} \leq x \leq \frac{\Delta x}{2}, \quad (7.73)$$

where Δx is the thickness of the dielectric.

7.4.2 Circular Antenna with Unidirectional Polarization

We treat a conjectural circular antenna where the polarization current is always parallel to the unit vector $\hat{\mathbf{e}}_3$ (*i.e.*, parallel to the z direction). This has been proposed as suitable for radar applications, but not yet constructed. The calculation was also used to diagnose possible phase errors in TD 1.

Using the same cylindrical polar coordinate system as in the previous section, we evaluate the following integral to give $\mathbf{A}(X, Y, Z, t_P)$:

$$\mathbf{A} = \frac{\mu_0}{4\pi} P_0 \omega e^{i\omega t_P} \int \int \int e^{-i(\frac{\omega R}{c} + m\psi)} \left\{ \frac{\rho}{R} \right\} \hat{\mathbf{e}}_3 \, d\rho d\psi dx \quad (7.74)$$

with

$$P_0 = -\varepsilon_0(\varepsilon_r - 1) \frac{V_0}{\Delta\rho}, \quad (7.75)$$

and

$$R = [(X - x)^2 + (Y - \rho \sin \psi)^2 + (Z - \rho \cos \psi)^2]^{\frac{1}{2}}. \quad (7.76)$$

The volume of integration is

$$\rho_0 - \frac{\Delta\rho}{2} \leq \rho \leq \rho_0 + \frac{\Delta\rho}{2}, \quad 0 \leq \psi \leq 2\pi, \quad \text{and} \quad -\frac{\Delta x}{2} \leq x \leq \frac{\Delta x}{2}, \quad (7.77)$$

where Δx is the dielectric thickness. All other symbols have their usual meanings.

7.5 Turning Potentials into Fields

We have derived integrals that yield the magnetic vector potentials for antennas TD 1 (circular), TD 2 and TD 5 (linear) and TD 3 (arc) (and indeed all present and as-yet unbuilt antennas of similar geometries). To compare theory with experiment, we now must calculate the *measurable fields* \mathbf{E} and \mathbf{H} at the observation point.

7.5.1 Much Needed Vector Algebra

Since the magnetic flux is always solenoidal (*e.g.*, $\nabla \cdot \mathbf{B} = 0$), the magnetic flux density \mathbf{B} can now be calculated from \mathbf{A} as

$$\mathbf{B}(X, Y, Z, t_P) = \nabla \times \mathbf{A} = \left(\hat{\mathbf{e}}_1 \frac{\partial}{\partial X} + \hat{\mathbf{e}}_2 \frac{\partial}{\partial Y} + \hat{\mathbf{e}}_3 \frac{\partial}{\partial Z} \right) \times \mathbf{A}(X, Y, Z, t_P) \quad (7.78)$$

and the magnetic field intensity is

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} = \frac{1}{\mu_0} \nabla \times \mathbf{A}. \quad (7.79)$$

The electric field is found from Maxwell's fourth equation

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t_P} \quad (7.80)$$

where \mathbf{J} and \mathbf{D} are the free current density and the electric displacement, respectively. In an insulating, uniform, isotropic medium,

$$\mathbf{J} = 0 \quad \text{and} \quad \mathbf{D} = \varepsilon_0 \mathbf{E}. \quad (7.81)$$

Hence,

$$\nabla \times \mathbf{H} = \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t_P}. \quad (7.82)$$

Since \mathbf{E} varies with t_P as $\mathbf{E} \propto e^{i\omega t_P}$, it must be the case that

$$\varepsilon_0 \frac{\partial \mathbf{E}}{\partial t_P} = i\omega \varepsilon_0 \mathbf{E} \quad (7.83)$$

and

$$\mathbf{E} = \frac{-i(\nabla \times \mathbf{H})}{\omega \varepsilon_0}. \quad (7.84)$$

Alternatively, we may use the electric scalar potential to find an expression for the electric field. Substituting Eq. (7.79) into Maxwell's curl equation yields

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t_P} = -\frac{\partial(\nabla \times \mathbf{A})}{\partial t_P} \quad (7.85)$$

or, in time-harmonic form,

$$\nabla \times \mathbf{E} = -i\omega \nabla \times \mathbf{A}. \quad (7.86)$$

Eq. (7.86) can be re-written as

$$\nabla \times \mathbf{E} + i\omega(\nabla \times \mathbf{A}) = 0 \quad (7.87)$$

$$\nabla \times [\mathbf{E} + i\omega \mathbf{A}] = 0 \quad (7.88)$$

$$\nabla \times (-\nabla \Phi) = 0, \quad (7.89)$$

where the scalar function Φ represents the electric scalar potential. It follows that

$$\mathbf{E} + i\omega\mathbf{A} = -\nabla\Phi \quad (7.90)$$

and

$$\mathbf{E} = -\nabla\Phi - i\omega\mathbf{A}. \quad (7.91)$$

Since $\rho = 0$ in, and along the surface of, the dielectric $\Phi = 0$ and \mathbf{A} satisfies the homogeneous wave equation. Hence, the fields are given by Eq. (7.78) and

$$\mathbf{E} \simeq -\frac{\partial\mathbf{A}}{\partial t_P} = -i\omega\mathbf{A}. \quad (7.92)$$

One of the validation methods used to check the numerical simulations of the antennas was to compare the results of Eq. (7.84) and Eq. (7.92); if all was well (and it almost invariably was), both yielded the same answer to within numerical precision.

7.5.2 Numerical Methods for the Derivatives

Simple first derivatives may be calculated with the central-difference formula, which can be derived by the Taylor series expansion. If the function $f(x)$ can be evaluated at values that lie to the left and right of x , then the best two-point *formulae* will involve abscissas that are chosen symmetrically on both sides of x .

Assume that $f \in [a, b]$, where a and b are in the space of complex numbers \mathbf{C}^3 . Moreover, let $x - h, x + h \in [a, b]$. Then

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}.$$

Furthermore, there exists $c = c(x) \in [a, b]$ such that

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + e_{\text{trunc}}(f, h),$$

where

$$e_{\text{trunc}}(f, h) = -\frac{h^2 f'''(c)}{6} = \mathbf{O}(h^2).$$

e_{trunc} is the truncation error, which, in this case, is of order h^2 . The step size h is taken to be < 1 . Since it is not desirable to choose h very small, we may wish to approximate $f'(x)$ with a formula that has a truncation error of higher order, *e.g.*, $\mathbf{O}(h^4)$. Hence, assume that $f \in \mathbf{C}^5[a, b]$ and that $x - 2h, x - h, x + h, x + 2h \in [a, b]$. Then,

$$f'(x) \approx \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h}, \quad \text{and}$$

$$f'(x) \approx \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h} + e_{\text{trunc}}(f, h),$$

where

$$e_{\text{trunc}}(f, h) = \frac{h^4 f^{(5)}(c)}{30} = \mathbf{O}(h^4)$$

The expression to calculate the curl of the potential then becomes

$$\begin{aligned} \nabla \times \mathbf{A} = & \left[\frac{-\mathbf{A}_z(X, Y+2h, Z) + 8\mathbf{A}_z(X, Y+h, Z) - 8\mathbf{A}_z(X, Y-h, Z) + \mathbf{A}_z(X, Y-2h, Z)}{12h} \right. \\ & \left. - \frac{-\mathbf{A}_y(X, Y, Z+2h) + 8\mathbf{A}_y(X, Y, Z+h) - 8\mathbf{A}_y(X, Y, Z-h) + \mathbf{A}_y(X, Y, Z-2h)}{12h} \right] \hat{\mathbf{e}}_1 \\ & + \left[\frac{-\mathbf{A}_x(X, Y, Z+2h) + 8\mathbf{A}_x(X, Y, Z+h) - 8\mathbf{A}_x(X, Y, Z-h) + \mathbf{A}_x(X, Y, Z-2h)}{12h} \right. \\ & \left. - \frac{-\mathbf{A}_z(X+2h, Y, Z) + 8\mathbf{A}_z(X+h, Y, Z) - 8\mathbf{A}_z(X-h, Y, Z) + \mathbf{A}_z(X-2h, Y, Z)}{12h} \right] \hat{\mathbf{e}}_2 \\ & + \left[\frac{-\mathbf{A}_y(X+2h, Y, Z) + 8\mathbf{A}_y(X+h, Y, Z) - 8\mathbf{A}_y(X-h, Y, Z) + \mathbf{A}_y(X-2h, Y, Z)}{12h} \right. \\ & \left. - \frac{-\mathbf{A}_x(X, Y+2h, Z) + 8\mathbf{A}_x(X, Y+h, Z) - 8\mathbf{A}_x(X, Y-h, Z) + \mathbf{A}_x(X, Y-2h, Z)}{12h} \right] \hat{\mathbf{e}}_3. \end{aligned} \tag{7.93}$$

To obtain the electric field, the curl of \mathbf{H} is then calculated (see Eq. (7.84)) using an identical routine; consequently, in the programs, \mathbf{A} is evaluated at 19 closely spaced points. The spacing is made very small with respect to the wavelength of the radiation, but large enough to ensure a quiet E-field after the second curl operation (and suitable integration at practical distances ~ 10 km).

7.6 Results from Calculations

7.6.1 Input Parameters

The integrals for the magnetic vector potential described in the previous section as well as the curl operations necessary for the fields were evaluated using a full numerical integration routine coded in MATLAB™. The resulting numerical results were then compared with experimental data from the linear (TD 2) and circular (TD 1) superluminal antennas run under a variety of conditions.

Input parameters for the calculations were derived from the experiments (Chapter 3); the practical prototype antennas emit at two frequencies, 2.4 GHz and 2.6 GHz; for some of the numerical simulations, an average value $f = 2.5$ GHz was used to save processing time, so that $\omega = 2\pi \times 2.5 \times 10^9$ radians per second. Dimensions corresponding to the real antennas are used; for example, TD 2, shown in the central panel of Fig. 3.1 has a dielectric approximately 0.348 m long in the y direction, 0.035 m high in the z direction and 0.004 m thick in the x direction; the material used is alumina with $\epsilon_r \approx 10$. For the purpose of evaluating phases, the effective thickness in the x direction is $0.004\sqrt{\epsilon_r} \approx 0.0126$ m; this is employed in the exponential term of the integrand. The corresponding dimensions of the TD 1 (Fig. 3.1 , left panel) are $\rho_0 = 0.127$ m, $\Delta\rho = 0.035$ m, and $\Delta x = 0.004$ m.

For all antenna configurations, the Matlab code generates the three-dimensional vector fields (\mathbf{E}, \mathbf{H}) of the radiation and their individual phases, allowing various polarization measurements (orthogonal linear, left and right circular) to be modelled. The phase fronts may also be extracted from these numerical data. To enable power measurements to be simulated, the Poynting vector (radiated power per unit area, and its direction) [54]

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}^\dagger, \quad (7.94)$$

where † denotes the complex conjugate, is also computed. The calculations permit

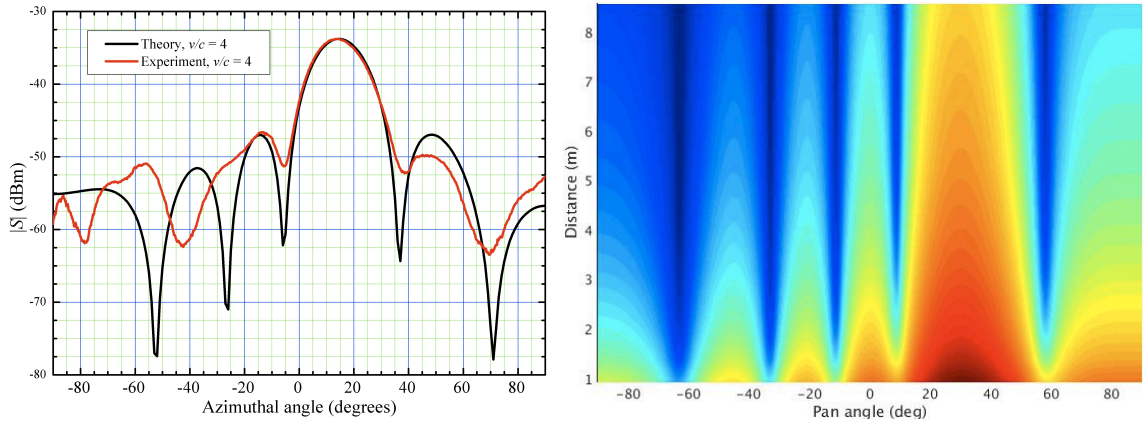


Figure 7.3: Left: modulus of Poynting vector $|S|$ versus azimuthal (pan) angle ϕ (black) predicted by the model of TD 2 (Fig. 3.1, center panel). The frequency is 2.4 GHz, $v/c = 4$ and the antenna to detector distance is 5 m. Experimental data for the same antenna, measured at the same distance (5 m) in the Sandia FARM anechoic chamber are shown in red. After correction for the gain of the detector antenna, there is a very good quantitative match between experiment and theory. Right: modulus of Poynting vector $|S|$ (colour contours) versus azimuthal (pan) angle ϕ and distance predicted by the model of the linear superluminal antenna. The frequency used was 2.5 GHz and the source speed was $v/c = 2$. The simulation is in good agreement with experimental data measured in the FARM and TA-35 anechoic chambers.

a full three-dimensional mapping (distance/range R , azimuthal (pan) angle ϕ and polar (tilt) angle θ of the radiation) to be carried out. This, of course, generates an enormous amount of information to be compared in detail with experimental measurements; for the present purposes, a relatively short summary is sufficient.

7.6.2 Linear Antenna: Superluminal Source Speed $v/c > 2$

Most of the numerical calculations with constant source speed aim to simulate TD 2 (Fig. 3.1, center panel), with which many experimental tests have been done. Fig. 7.3 shows a comparison of the numerical model and experimental data measured in the Sandia FARM anechoic chamber. The experimental conditions [voltage applied to the electrodes, polarization-current speed ($v/c = 4$), dielectric dimensions) were used as model input parameters. After correction for the gain of the receiver antenna and

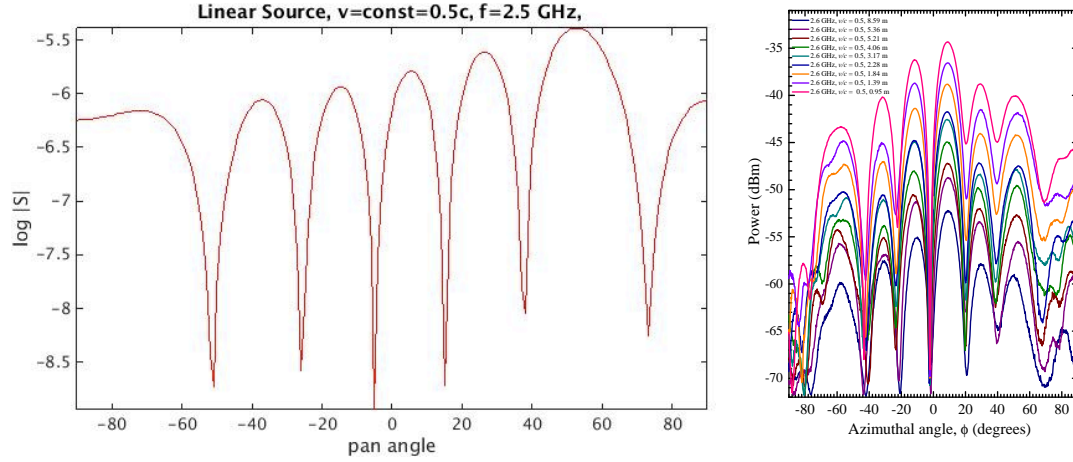


Figure 7.4: Left: modulus of Poynting vector versus azimuthal (pan) angle ϕ predicted by the model of linear superluminal antenna TD 2; the subluminal source speed is $v/c = 0.5$. The frequency used is 2.5 GHz and the antenna to detector distance 5 m. Right: experimental data from TD 2 shown as power versus azimuthal angle ϕ . Data are given for several distances in the range 0.95 – 8.59 m, and for the subluminal source speed $v/c = 0.5$; data are shown for $f_+ = 2.6$ GHz. The inset keys explain the colour-coding of the data, which were measured in the FARM anechoic chamber.

cabling used in the experiment, there is a very good quantitative match between data and theory, especially close to the main lobe. At larger angles, the match is less good. This is understandable, because the subsidiary minima are very dependent on the precise phases of the signals applied to each antenna element, which were subject to errors of a few degrees in the experiment. Similar quantitative agreement between model and anechoic chamber data was obtained for all speeds above $v/c = 2$. As is the case with the experimental data, the simulated curves are very similar to skewed Airy functions (Section 3.6.2).

Appendix C.2 discusses the issues encountered when running TD 2 at superluminal speeds $1 \leq v/c \leq 2$, and in particular the lower-than-expected output power.

7.6.3 Linear Antenna: Subluminal Source Speeds

Results from the simulation of the linear antenna TD 2 for a source speed of $v/c = 0.5$ are given in the left-hand section of Fig. 7.4 for a frequency of 2.5 GHz and a source-to-detector distance of 5 m. Note that the peak power per unit area (Poynting vector magnitude) is about 15 – 20 dB less than for comparable superluminal numerical data for $v/c > 2$ (*e.g.*, Fig. 7.3). These simulations are in good agreement with experimental results from TD 2 run at $v/c = 0.5$ in the FARM Range anechoic chamber; typical data are shown in the right-hand side of Fig. 7.4. In both simulations and experiments, the angular dependence of the emitted power is highly oscillatory; moreover, the maximum intensities are some 15 – 20 dB lower than those observed when $v/c \geq 2$.

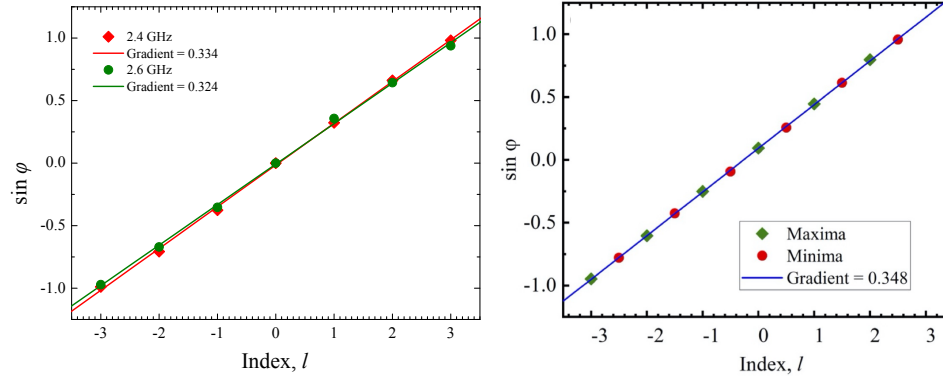


Figure 7.5: Left: a plot of $\sin \phi$ versus index l , where ϕ denotes the experimental angle at which a minimum in power occurs, for a distance of 8.59 m (see Fig. 7.4, right). Experimental data are points, and the lines are straight-line fits, giving gradients of 0.33 and 0.32 for 2.4 GHz and 2.6 GHz points respectively. Right: $\sin \phi$ versus index l for the simulation shown in the left side of Fig. 7.4; the frequency is 2.5 GHz (intermediate between the experimental values) and positions of both minima (blue) and maxima (red) are plotted. The gradient of the fitted line is 0.34, very close to the experimental values found in the left-hand plot.

The power oscillations in Fig. 7.4 are similar to those from a two-slit diffraction experiment in which the light from one of the slits is $(2j + 1)\pi$ out of phase with that from the other, where j is an integer. In the far field, such a two-slit experiment

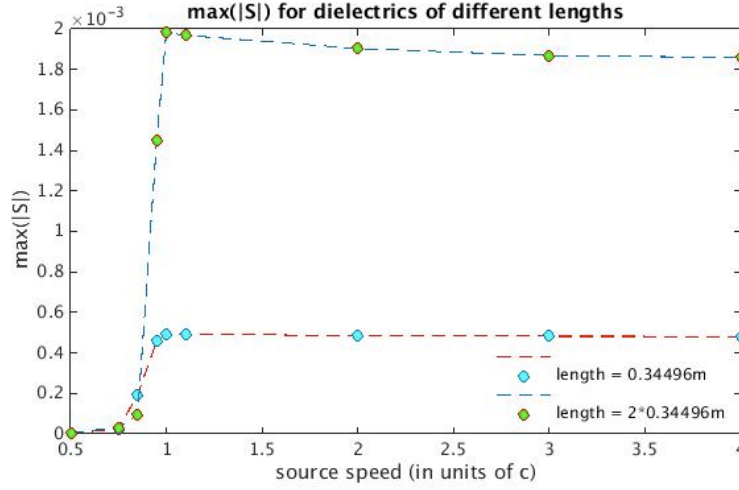


Figure 7.6: Peak Poynting vector magnitude $|S|$ versus source speed v/c simulated for two linear superluminal antennas. Blue points are for a length of 0.345 m (similar to TD 2) and green for 0.690 m (similar to TD 5). Note that the cutoff in power as v/c falls below 1 is sharper for the longer antenna.

would give minima that occur when [12]

$$b \sin \phi = l\lambda, \quad (7.95)$$

where b is the spacing of the slits, l is an integer and λ is the wavelength. Hence, a plot of $\sin \phi$ versus l should be a straight line, with gradient λ/b . Such plots are shown in Fig. 7.5 for minima positions derived from 8.59 m distance data such as those in the right-hand side of Fig. 7.4 and from the simulations shown in the left-hand side of Fig. 7.4. Taken with the relevant wavelengths $\lambda = c/f$, the gradients of the experimental data for 2.4 GHz and 2.6 GHz yield $b = 370$ mm and $b = 360$ mm, and the simulation $b = 347$ mm, values that are very close to the 348 mm overall length of the 32 elements.

This result may be understood qualitatively as follows. Ideally, *no* vacuum Čerenkov radiation should be emitted for $|v/c| < 1$, as the emission angle $\sin^{-1} c/v$ becomes imaginary [46–48]. In this context, “ideal” implies an infinitely-long source of identical elements, in which the radiation from all elements superposes to produce

no net emission. However, the linear superluminal antenna is of finite length, so that the radiation measured in the $v/c = 0.5$ experiments is likely to come mostly from the *ends* of the antenna; the elements at the ends have adjacent elements on only one side. Hence, one might expect that about half of their emitted power would be canceled out, so that the two end elements behave like a double-slit experiment emitting a total power $\sim (1/2) \times 2 \times (1/32) = 1/32$ of the total power of the array. Converting into dB, $10 \log_{10}(1/32) = -15$ dB, explaining why the peak subluminal emission is 15 – 20 dB lower than the peak vacuum Čerenkov power produced at speeds $v/c > 2$, where all 32 elements contribute. The separation b deduced above is very similar to the separation of elements 1 and 32 in the Superluminal Antenna. Moreover, for $v/c = 0.5$, the phase difference of elements $j = 1$ and $j = 32$ is very close to 11π , explaining why the “interference pattern” has *minima* quite close to the values of ϕ given by Eq. (7.95).

The longer the superluminal antenna becomes, the better it will approximate an infinitely long Čerenkov path; we would expect the power cutoff to be sharper as v falls below c . The simulations reproduce this behaviour, as is shown in Fig. 7.6.

7.6.4 Circular Superluminal Antenna

Fig. 7.7 shows a numerical simulation (black) of the circular superluminal antenna TD 1 compared with data (red) measured in the FARM anechoic chamber. The simulation and data are for $m = 1$ ($v/c \approx 6.6$) (see Eq. (7.66)) and the model input parameters ($\rho_0 = 127.0$ mm, $\Delta\rho = 35.0$ mm, $\Delta x = 4.0$ mm, $V_0 = 0.45$ V) correspond to the experimental conditions. There is a good match between model and data; the distribution of power between vertical and horizontal polarizations is very similar, and the relative heights of the principal maximum and the side lobes is correctly reproduced. There is a slight difference between model and data for the side-lobe position and the depth of the adjacent minima; we return to this below.

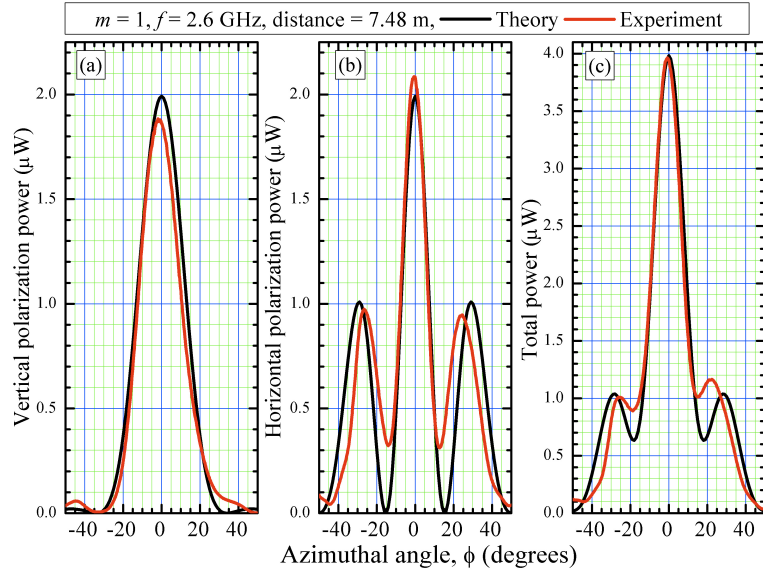


Figure 7.7: Model (black) of circular antenna TD 1 ($m = 1$, $f = 2.6$ GHz) compared with experimental data (red) measured in the FARM Anechoic Chamber: (a) vertical polarization power; (b) horizontal polarization power; (c) total power. Powers are in μW and other parameters are in the key.

Fig. 7.8 shows the predicted power for TD 1 [vertical (a) and horizontal (b) polarizations, and total (c)] as a function of azimuthal (pan) angle for various speed indices (see Eq. (7.66)) ranging from $m = 1$ (speed $\approx 6.6c$) to $m = 6$ (speed $\approx 1.1c$). The calculation reproduces the experimental spreading of the antenna patterns with increasing m , and predicts the observed relative proportions of horizontally- and vertically polarized radiation correctly. It also reproduces the observed fall off in peak power with increasing m .

7.6.5 Using the Model to Explore Antenna Design

Having seen that the numerical model reproduces experimental data for TD 1 and TD 2 quantitatively, we can use it as a design and diagnostic tool to establish how varying key parameters (frequency, dielectric dimensions) affects antenna performance. Since TD 1 is much less like a conventional antenna than TD 2, we con-

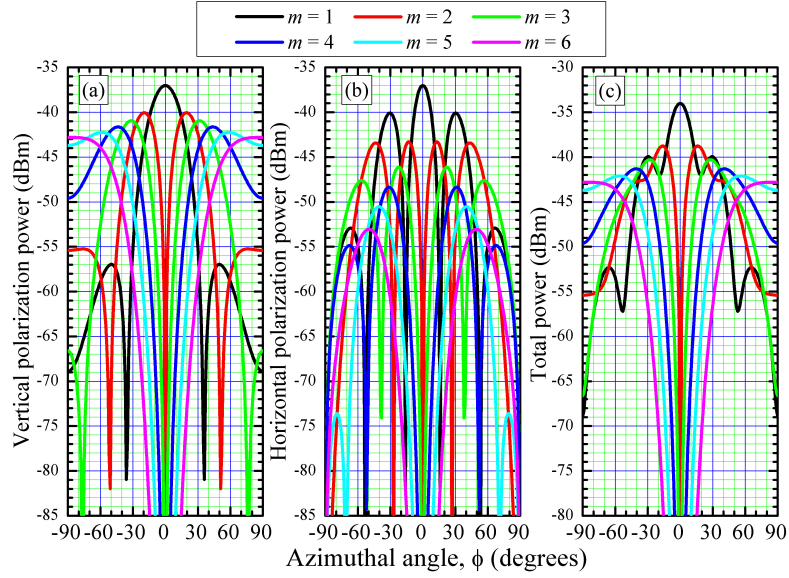


Figure 7.8: Predicted power (dBm) from circular antenna TD 1 at a range of 100 m for $f = 2.5$ GHz and different m values (shown in key): (a) vertical polarization power; (b) horizontal polarization power; (c) total power.

centrate here on the former. (Sections 7.6.5 to 7.6.7 are highly edited extracts from a report prepared by the current author and colleagues for the telecommunications company CommScope [69]. The results therein form the basis of a future paper.)

Fig. 7.9 shows the effect on the power output of TD 1 of varying the frequency $f = \frac{\omega}{2\pi}$; the dimensions of the experimental machine ($\rho_0 = 127.0$ mm, $\Delta\rho = 35.0$ mm, $\Delta x = 4.0$ mm) were used in the simulation, and the oscillatory voltage applied to the elements is kept constant at $V_0 = 0.45$ V. The index $m = 1$ was chosen for clarity because the antenna pattern is simplest; analogous effects are seen for $2 \leq m \leq 6$. The antenna pattern narrows as f increases; the effect is particularly marked close to 90° . As is observed in experimental data, the peak power P_{peak} grows with increasing frequency; running several simulations for $2.0 \leq f \leq 3.0$ GHz, shows that $P_{\text{peak}} \propto f^4$. This can be understood as follows. Eq. (7.68) reveals that the act of differentiating \mathbf{P} to obtain $\partial\mathbf{P}/\partial t$ multiplies the integral that gives \mathbf{A} Eq. (7.69)) by ω . Fields (\mathbf{E}, \mathbf{H}) are obtained by one further differentiation of \mathbf{A} , which introduces another power of

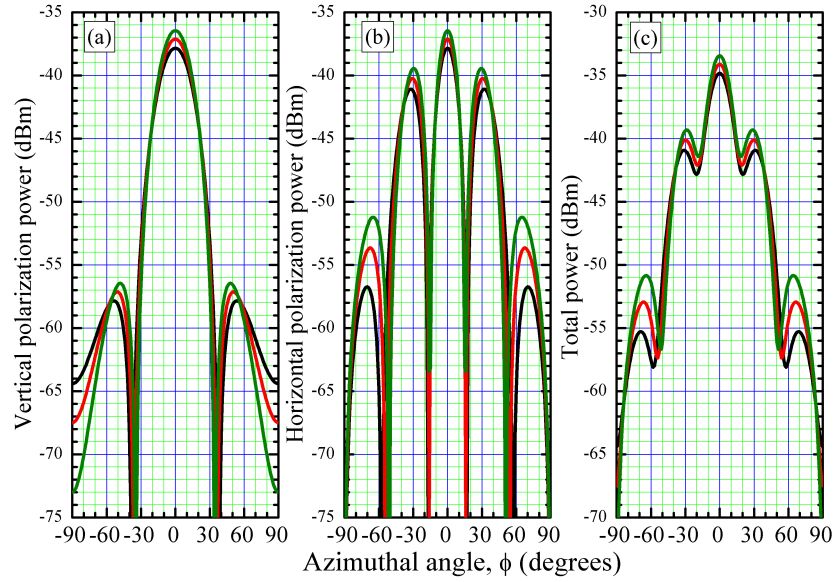


Figure 7.9: Model calculations of the power received at a distance of 100 m from circular antenna TD 1 ($m = 1$, $\Delta\rho = 35.0$ mm, $\rho_0 = 127.0$ mm, $\Delta x = 4.0$ mm, $V_0 = 0.45$ V) for frequencies $f = 2.4$ GHz (black) 2.5 GHz (red) and 2.6 GHz (green): (a) vertical polarization power; (b) horizontal polarization power; (c) total power.

ω ; this is most easily seen if one remembers that $\mathbf{E} \approx -\partial\mathbf{A}/\partial t$. Power is proportional to the square of the fields, resulting in the observed ω^4 (*i.e.*, f^4) dependence of the peak power.

Fig. 7.10 shows the effect of varying the mean radius ρ_0 on a circular antenna similar to TD 1. The antenna pattern for $m = 2$ is shown; qualitatively similar behaviour was observed for other values of m . All other dimensions were kept the same as in the experimental antenna and the frequency used was $f = 2.5$ GHz. Increasing ρ_0 sharpens/narrows the features of the antenna pattern. The peak power increases as $P_{\text{peak}} \propto \rho_0^2$. (On going from $\rho_0 = 100$ mm to $\rho_0 = 140$ mm, the power goes up by a factor $1.4^2 \approx 2 \approx 3$ dB.) This can be again understood by considering Eqs. 7.68 and 7.69; the electric field amplitude remains the same (because V_0 and $\Delta\rho$ are unchanged), but volume occupied by the polarization current increases in proportion to ρ_0 . As mentioned above, the power is proportional to the square of

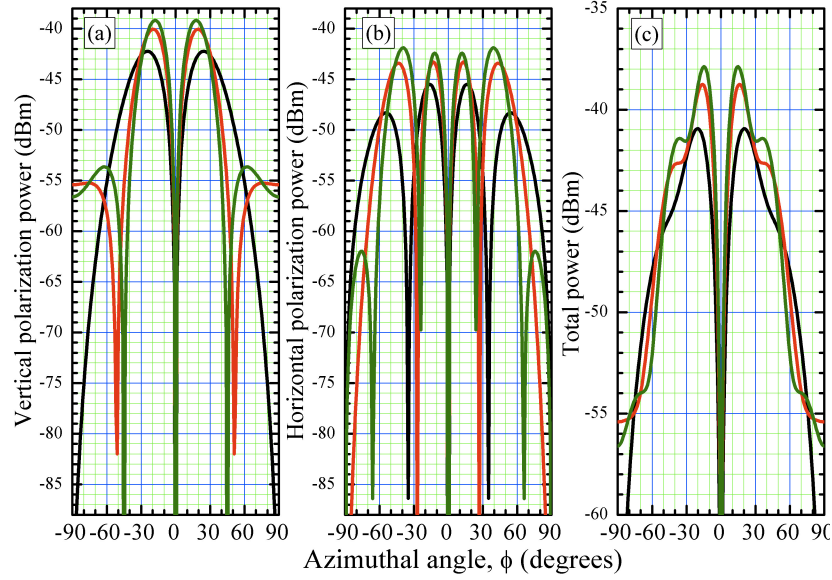


Figure 7.10: The effect of varying the mean dielectric radius ρ_0 in a circular antenna similar to TD 1 for $m = 2$ and $f = 2.5$ GHz; $\rho_0 = 100.0$ mm (black), 127.0 mm (red) 140.0 mm (green). All other parameters are the same as in Fig. 7.9.

the fields, leading to the observed $P_{\text{peak}} \propto \rho_0^2$.

Fig. 7.11 shows the effect of varying the radial width, $\Delta\rho$, on the antenna pattern. Increasing $\Delta\rho$ sharpens/narrows the antenna pattern; a careful choice of $\Delta\rho$ can cut the power radiated at azimuthal angles $|\phi| \geq 60^\circ$ dramatically (see red curves for $m = 2$, corresponding to $\Delta\rho = 60.0$ mm.) In this case, the peak power is roughly independent of $\Delta\rho$; again we consult Eqs. 7.68 and 7.69, where it can be seen that the electric field amplitude is proportional to $1/\Delta\rho$. The decrease in electric field amplitude as $\Delta\rho$ grows is exactly compensated by the increase in the volume occupied by the polarization current, which is proportional to $\Delta\rho$.

Finally, the calculations examined the effect of varying the dielectric depth Δx . For relatively small values of Δx compared to the wavelength $\lambda = 120$ mm of the radiation, the pattern does not change significantly; the power merely grows in proportion to Δx^2 . This is due to the volume occupied by the polarization current being proportional to Δx . However, once the effective (optical) depth of the dielectric,

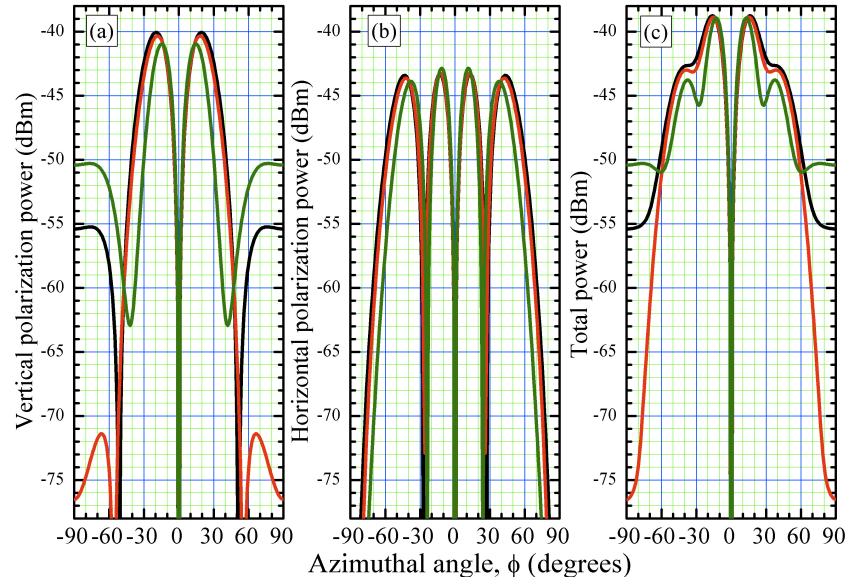


Figure 7.11: Model calculations of the power received at a distance of 100 m from a circular antenna similar to TD 1 ($m = 2$, $f = 2.5$ GHz, $\rho_0 = 127.0$ mm, $\Delta x = 4.0$ mm, $V_0 = 0.45$ V) for dielectric radial widths $\Delta\rho = 35.0$ mm (black) 60.0 mm (red) and 120.0 mm (green): (a) vertical polarization power; (b) horizontal polarization power; (c) total power.

$\varepsilon_r^{\frac{1}{2}}\Delta x$, becomes a significant fraction of λ , the antenna pattern changes significantly (Fig. 7.14), and the peak power no longer follows the simple proportionality. This is due to interference between radiation emitted by the back and front of the dielectric. Consequently, there is a relative increase in the importance of side lobes at higher azimuthal angles; in the case illustrated, this gives a higher proportion of horizontal polarization. A more detailed description of these effects is in Appendix C.2.

7.6.6 The Model as a Diagnostic Tool

The numerical model can also be used to diagnose antenna performance issues. Fig. 7.12 shows data from TD 1 ($m = 2$, $f = 2.6$ GHz, antenna to detector distance 92 m) obtained on the FARM outdoor range. At the time, there was an airborne RADAR landscape imaging experiment that included a high-power aircraft-

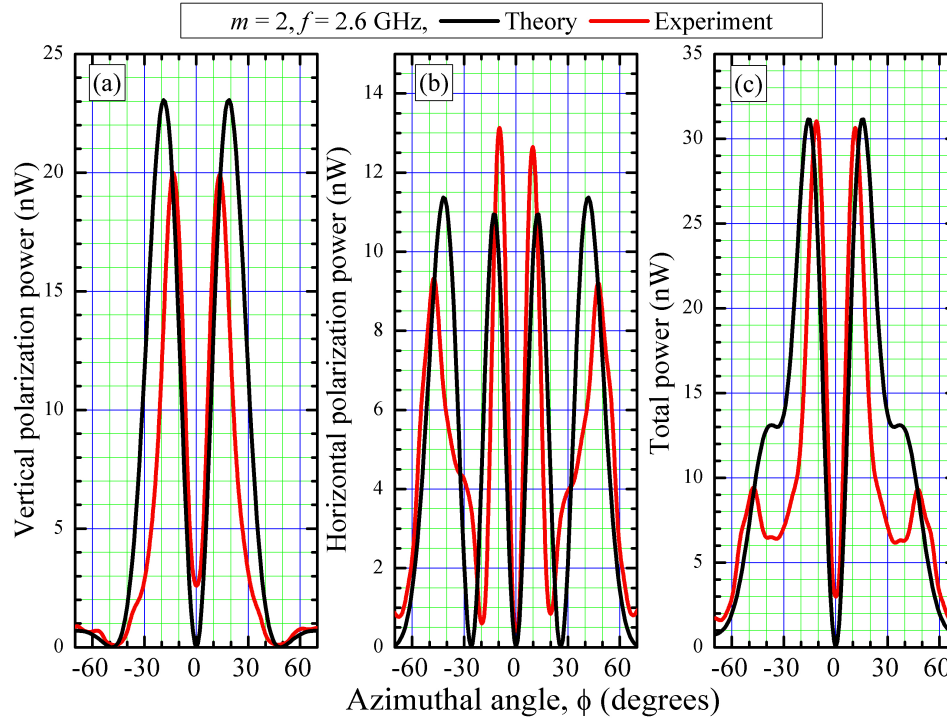


Figure 7.12: Model calculations of the power received at a distance of 92 m from TD 1 [$m = 2, f = 2.6 \text{ GHz}$] (black) compared to experimental data (red) recorded at the FARM outdoor range in the presence of electromagnetic noise; (a) vertical polarization power; (b) horizontal polarization power; (c) total power.

to-ground telemetry link functioning nearby, and the consequent background signals interfered with setting the phase of the voltages sent to the antenna elements. This has affected the positions of the experimental maxima compared to the model, plus the relative sizes of the horizontal and vertical polarizations. However, once the angular-velocity variations caused by the phase settings (functionally equivalent to a $\pm 10\%$ variation in m) were included in the model, the experimental antenna pattern was reproduced (Fig. 7.13). This is not just a useful diagnostic result; variations in the polarization-current angular velocity around a circular antenna could be *designed* to shift and sharpen features in the emitted radiation in a desirable manner.

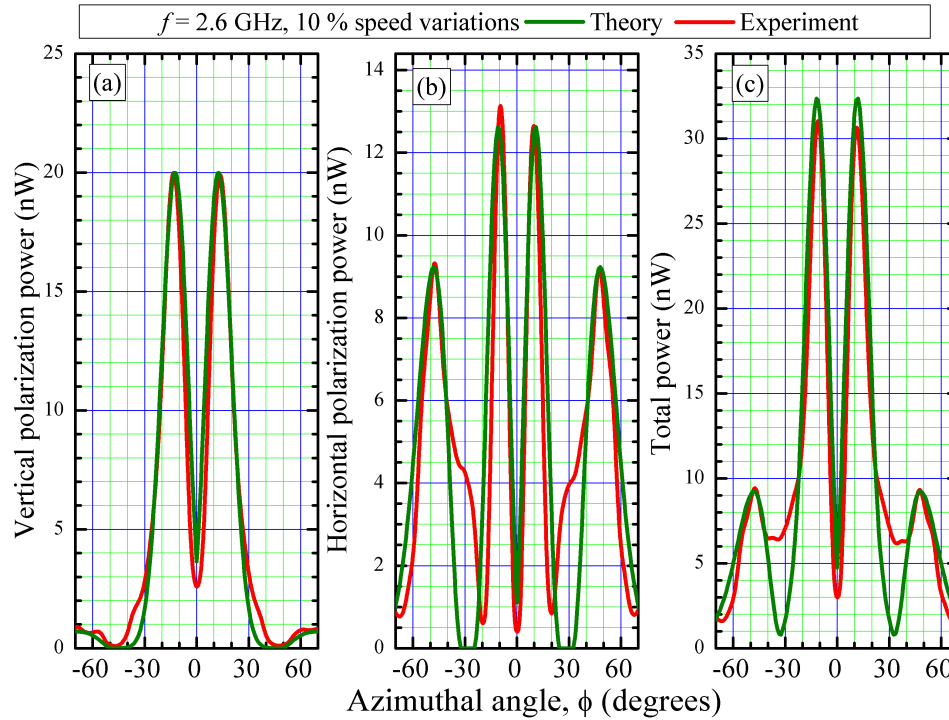


Figure 7.13: The numerical model incorporating a $\pm 10\%$ variation in source angular velocity caused by the noisy phase settings (green) compared with the same experimental data as in Fig. 7.12 (red). (a) vertical polarization power; (b) horizontal polarization power; (c) total power.

7.6.7 Non-Spherical Decay Exponents

Numerical data sets such as those shown in Figs. 7.7 to 7.13 (and their continuations out to distances of up to 100 km) show a decay exponent very close to -2 (the inverse square law) for all angles. However, the experimental data have yielded exponents as high as -1.3 , especially at distances $\sim 3 - 20$ m (see *e.g.*, Fig. 3.16).

The reason for the discrepancy seems to stem from the “perfection” and continuity of the numerical sources; in the initial implementation of the numerical model, the voltages animating the polarization current are perfectly sinusoidal (Eq. (7.66)), with no discontinuities or noise. By contrast, the acoustical sources that generate sonic booms fluctuate on short timescales and are much more localized (*i.e.*, point-like).

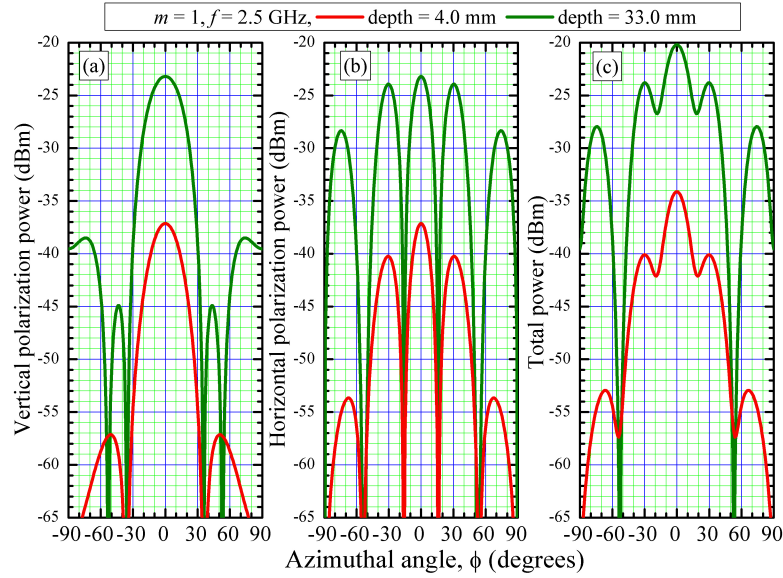


Figure 7.14: Model calculations of the power received at a distance of 100 m from an antenna similar to TD 1 ($m = 1$, $f = 2.5$ GHz, $\rho_0 = 127.0$ mm, $\Delta\rho = 35.0$ mm, $V_0 = 0.45$ V) for dielectric depths $\Delta x = 4.0$ mm (red) and 33.0 mm (green): (a) vertical polarization power; (b) horizontal polarization power; (c) total power.

Consequently, pulses, noise and/or fluctuations in a circular superluminal source may give rise to signals with decay exponents $p > -2$. The distance range over which this non-spherical decay occurs depends critically on the lengthscale and timescale of the disturbances in the source.

We have combined this idea with the numerical model; introducing temporal fluctuations in the phase and amplitude of the source results in numerical data such as those presented in Fig. 7.15. The simulation involves fluctuations with a timescale ~ 5 ps, typical of the “fur” on the experimental voltages applied to the elements of TD 1, observed using a high speed oscilloscope. Over source-to-detector distances $\sim 3 - 20$ m, these fluctuations give rise to decay exponents that are typically greater than -2 (Fig. 7.15) at azimuthal angles $|\phi| < 70^\circ$. Such values are comparable to the exponents observed in experiments (Fig. 3.16) over similar distance ranges.

In TD 1, these imperfections, noise and fluctuations give rise to bursts of polar-

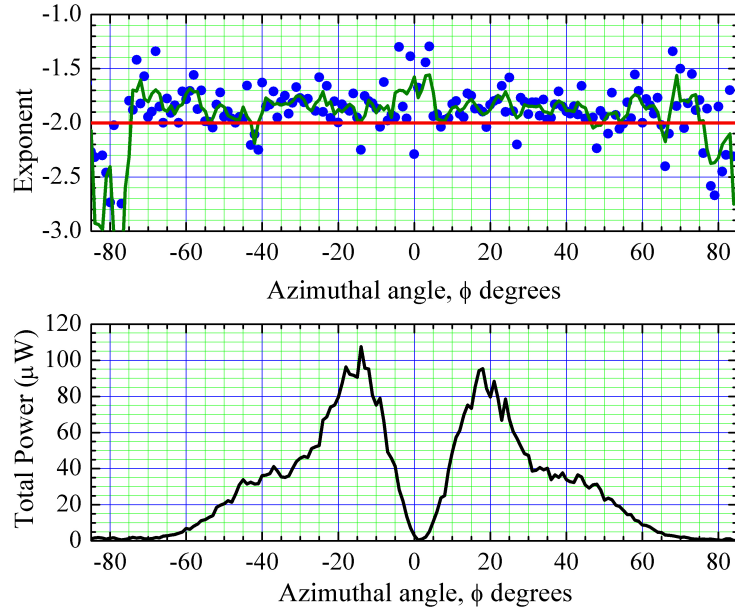


Figure 7.15: Numerical model of TD 1 in which voltage and phase fluctuations of typical timescale ~ 5 ps are introduced. The parameters are similar to those of the experiment in Fig. 3.16 ($m = 2$, $f = 2.5$ GHz, $\rho_0 = 127.0$ mm, $\Delta\rho = 35.0$ mm, $\Delta x = 4.0$ mm). The lower panel shows a snapshot of the noisy detected power; the upper panel is the corresponding distance exponent observed from 4 m to 10 m. Departures from -2 (red line) indicate deviations from the inverse-square law; the effect encourages upward deviations over distances that are related to the timescale of the fluctuations, indicating an improvement over the inverse-square law. The green line is a coarse average of the values evaluated at discrete angles.

ization current that propagate around the dielectric, in effect functioning as compact sources. In analogous sonic-boom sources, the emitting region is not polarized in a well-defined way, so that emissions from source points at different space-time locations (retarded times) can combine to give temporal focusing. By contrast, in an ideal circular superluminal antenna (Eq. (7.69)), the plane of polarization of the source rotates as it orbits the dielectric, with orthogonal directions in quadrature; this will affect the ability of radiation from different retarded times to reinforce to produce temporal focusing. In TD 1, phasing errors give rise to a component of the polarization current that possesses a unidirectional polarization for at least part of its orbit around the antenna, allowing enhanced temporal focusing.

7.7 Focusing Effects: Linear Antenna

We now turn to simulations of the emission from the wavepackets of polarization current described in Section 6.6. In this case, the emitted fields no longer have a harmonic time dependence; instead, solutions covering an extended reception time t_P must be derived. The Cartesian coordinate system first used in Section 7.3 is employed to describe the problem.

7.7.1 Formalism

Consider a voltage applied symmetrically across the dielectric in the vertical (z) direction, given by $V_0 e^{i\omega[t-p(y)]} e^{-\alpha^2[t-p(y)]^2}$. Note that this consists of a Gaussian convoluted with a travelling wave⁴; both have the (y, t) dependence required for the motion described in Section 6.6. As before, let the dielectric extend from $z = -\frac{h}{2}$ to $z = +\frac{h}{2}$ in the vertical direction; assuming that the dielectric is uniform, the potential at a general position y, z will be

$$V(y, z, t) = V_0 \frac{z}{h} e^{i\omega[t-p(y)]} e^{-\alpha^2[t-p(y)]^2}. \quad (7.96)$$

After substituting from Eq. (6.85), applying the operation

$$\mathbf{P} = \varepsilon_0(\varepsilon_r - 1)\mathbf{E} = \varepsilon_0(\varepsilon_r - 1)(-\nabla V), \quad (7.97)$$

and differentiating with respect to time, we obtain a polarization-current density

$$\mathbf{J}(y, z, t) = \frac{\partial \mathbf{P}}{\partial t} = -\varepsilon_0(\varepsilon_r - 1)(j_y \hat{\mathbf{e}}_2 + j_z \hat{\mathbf{e}}_3) e^{i\omega[t-p(y)]} e^{-\alpha^2[t-p(y)]^2}, \quad \text{with} \quad (7.98)$$

$$j_y = z \left[\frac{i\omega(Y_0 + y) - 2\alpha^2 t(Y_0 + y)}{c(X_0^2 + (Y_0 + y)^2)^{\frac{1}{2}}} - \frac{2\alpha^2(Y_0 + y)}{c^2} \right] \times$$

⁴For algebraic convenience, this expression is complex. In comparisons with (real) experimental data, we must choose a phase, for example by taking the real or imaginary component of the final computed signal (or even some other angle in the complex plane).

$$\left[i\omega - \frac{2\alpha^2(X_0^2 + (Y_0 + y)^2)^{\frac{1}{2}}}{c} - 2\alpha^2 t \right] - \frac{2z\alpha^2(Y_0 + y)}{c(X_0^2 + (Y_0 + y)^2)^{\frac{1}{2}}} \quad \text{and}$$

$$j_z = \left(i\omega - \frac{2\alpha^2(X_0^2 + (Y_0 + y)^2)^{\frac{1}{2}}}{c} - 2\alpha^2 t \right).$$

The definition of $p(y)$ is given in Eq. (6.85) and X_0 and Y_0 are specified by Eq. (6.78).

As before, in calculating the emitted radiation, we consider only the contribution of the polarization currents in the dielectric; there are no conduction currents in such an insulator, and we neglect the free charges that exist only at the interface between the dielectric and the electrodes. \mathbf{A} is obtained using Eq. (7.28), repeated for clarity:

$$\mathbf{A}(\mathbf{r}_P, t_P) = \frac{\mu_0}{4\pi} \iiint \frac{\mathbf{J}(\mathbf{r}, t)}{|\mathbf{r}_P - \mathbf{r}|} d^3\mathbf{r},$$

where \mathbf{r} is a coordinate within the dielectric. The integration is carried out over the volume of the dielectric; t is the retarded time, given by $t = t_P - \frac{|\mathbf{r}_P - \mathbf{r}|}{c}$, where c denotes the speed of electromagnetic waves in the medium (assumed to be uniform) between the source and the observer. The corresponding radiation fields are derived from differentiating \mathbf{A} with respect to the relevant observer's coordinates (X, Y, Z, t_P) :

$$\mathbf{E}(X, Y, Z, t_P) = \frac{\partial \mathbf{A}(X, Y, Z, t_P)}{\partial t_P}, \quad (7.99)$$

and

$$\mathbf{B}(X, Y, Z, t_P) \equiv \nabla \times \mathbf{A}$$

$$= \left(\hat{\mathbf{e}}_1 \frac{\partial}{\partial X} + \hat{\mathbf{e}}_2 \frac{\partial}{\partial Y} + \hat{\mathbf{e}}_3 \frac{\partial}{\partial Z} \right) \times \mathbf{A}(X, Y, Z, t_P). \quad (7.100)$$

Eqs. (7.28), (7.99) and (7.100) are calculated numerically (see Section 7.5.2).

7.7.2 Results and Potential Applications

Simulations using the above procedure are plotted in Fig. 6.15; a further example is given here to illustrate possible future lines of research. These numerical results employ a polarization current burst with the time dependence shown in Fig. 7.16 which is accelerated using the same parameters as in Section 6.6.2. Similarly, the dielectric dimensions are based on the antenna TD 4 (Fig. 6.17) used in those experiments.

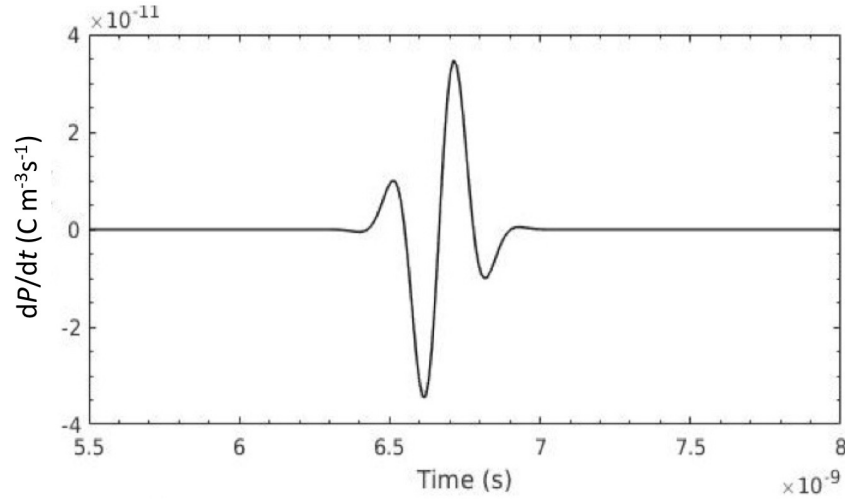


Figure 7.16: Time dependence of the polarization current used for the simulations plotted in this section. This is based on a carrier frequency of $\omega/2\pi = 3.6$ GHz and an adjustable Gaussian envelope with $\alpha = 4 \times 10^9 \text{ s}^{-1}$ (see Eq. (7.98)).

The simulated power at the “focus” distance of 3.09 m is shown as a power-contour plot versus pan (azimuthal) angle and frequency in Fig. 7.17 (left); the right-hand plot is an expansion around the focus angle visualized using slices through the contour plot at 1° intervals. The maximum power (yellow) occurs at a pan angle of $\approx 11.9^\circ$, the desired focus direction (Section 6.6.2) and at 3.6 GHz, the carrier frequency (Fig. 7.16). Although the acceleration scheme is straightforward, the signal is “scrambled” rapidly with increasing angular distance from the focus at 11.9° since the relationship between emission and reception time becomes neither linear nor 1:1 (Fig. 6.15). Under such conditions, the received signal loses its original frequency

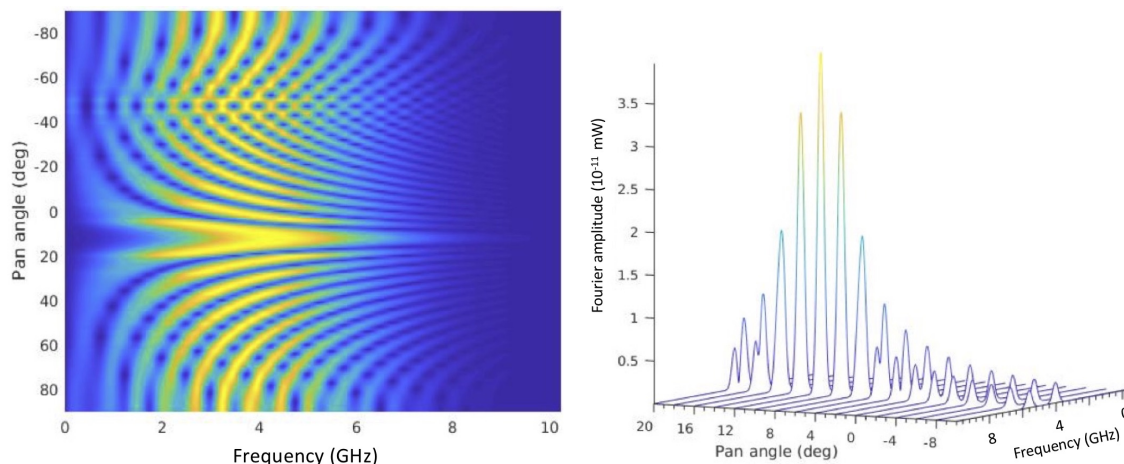


Figure 7.17: Left: simulated detected power at an antenna-to-receiver distance of 3.09 m plotted as power contours (dark blue: weak; yellow: strong) versus frequency and pan (azimuthal) angle. As intended with the acceleration scheme used (Section 6.6.2), the signal has maximum strength at a pan angle close to 11.9° , and at a frequency of 3.6 GHz (the carrier frequency). Right: expansion of the angular region close to the focus using slices through the contour plot at 1° intervals.

content rapidly and becomes “illegible”.

This technique represents a contrast to conventional radio transmission methods. In many instances of the latter, signals are broadcast with little or no directivity, selectivity of reception being achieved through the use of one or more narrow frequency bands. In place of this, the technique discussed here uses a spread of frequencies to transmit information to a particular location; the signal is weaker and has a scrambled time dependence elsewhere. A possible application may be in proposed 5G neighbourhood networks, where a single active antenna will sequentially spray bursts of information into a selection of target buildings around it; ensuring that neighbours cannot easily understand what you are transmitting and receiving will be an important component.

The results reported here represent merely a proof-of-concept; much remains to be explored using simulations and experiment, but a few preliminary remarks may be made. Note that the depth of focus (*i.e.*, the range of distance over which the signal

is comprehensible) depends strongly on the frequency spread of the broadcast signal. The waveform used in the experiment, which encompasses frequencies from 0.9 to 4.5 GHz, results in a received signal that distorts relatively quickly as the detector moves out beyond the target distance of $R_0 = 3.0$ m at the target angle (Fig. 6.22). By contrast, a relatively narrow-band broadcast signal [*e.g.*, Fig. 7.17] will be recognizable at the target angle over a wider range of detector distances.

7.8 Focusing Effects: Circular Superluminal Sources

7.8.1 Formalism

We now consider signals broadcast by a wavepacket-type source traveling on a circular orbit at superluminal speeds. The wavepacket's extension will primarily be in its instantaneous direction of travel, *i.e.*, around the circular orbit path.

In the case of the linear antenna, the important condition for receiving unscrambled signals is that the t_P versus t relationship should be flat [*i.e.*, $t_P = \text{a constant}$, as shown in Fig. 6.15, row (a), central plot] at the location of the detector. For circular superluminal sources, Fig. 6.1 illustrates that at certain special detector positions, part of the t_P versus t relationship displays a plateau. However, the extended discussion and analysis of Fig. 6.1 showed that this plateau is not truly flat, zero gradient only being achieved at a single point in the center. Nevertheless, as long as the source contains only frequencies that are smaller than $\sim 1/(\Delta t_P)$, where Δt_P is the observer-time width of the plateau, “information focusing” analogous to that shown for the linear antenna in Fig. 6.15(a) may also be possible with wavepackets of polarization current in circular superluminal motion. In the circular case, we expect the focusing to occur at all antenna-to-detector distances, but only at a precise polar angle that is slightly distance dependent until the far field is reached (Section 6.2.2).

We again employ the dimensions and properties of TD 1 (Fig. 3.1(a)) for the model ⁵; Positions \mathbf{r} within the antenna are described using the cylindrical polar coordinate system of Section 7.4.1. The potential within the circular dielectric due to voltages applied to the electrodes is⁶

$$V = \frac{V_0}{\Delta\rho} (\rho - \rho_0) e^{i(\omega t - m\psi)} e^{-\alpha^2 \left(t - \frac{m\psi}{\omega}\right)^2}. \quad (7.101)$$

Just as for the linear antenna (Eq. (7.96)), the parameter α controls the time-extent of the wavepacket; however, in the present circular case, the speed (*i.e.*, the angular velocity) and hence the focus angle is controlled by the integer m . Standard identities are again used to obtain the polarization current density \mathbf{J} :

$$\mathbf{E} = -\nabla V, \quad \mathbf{P} = \varepsilon_0(\varepsilon_r - 1)\mathbf{E} \quad \text{and} \quad \mathbf{J} = \frac{\partial \mathbf{P}}{\partial t}.$$

After some algebra, we have

$$\mathbf{J} = -P_0 f(t, \psi) \left[\mathbf{e}_+ \Gamma(t, \psi) + \mathbf{e}_- \left(\frac{\rho - \rho_0}{\rho} \right) \left(\frac{2m\alpha^2}{\omega} + \Gamma(t, \psi) \aleph(t, \psi) \right) \right] \quad (7.102)$$

where

$$P_0 = \varepsilon_0(\varepsilon_r - 1) \frac{V_0}{\Delta\rho}, \quad f(t, \psi) = e^{i(\omega t - m\psi) - \alpha^2 \left(t - \frac{m\psi}{\omega}\right)^2},$$

$$\mathbf{e}_+ = \hat{\mathbf{e}}_2 \cos \psi + \hat{\mathbf{e}}_3 \sin \psi, \quad \mathbf{e}_- = \hat{\mathbf{e}}_2 \cos \psi - \hat{\mathbf{e}}_3 \sin \psi,$$

$$\Gamma(t, \psi) = \frac{2m\alpha^2\psi}{\omega} + i\omega - 2\alpha^2 t \quad \text{and} \quad \aleph(t, \psi) = \frac{2m\alpha^2 t}{\omega} - im - \frac{2m^2\alpha^2\psi}{\omega^2}.$$

We can then evaluate the magnetic vector potential $\mathbf{A}(X, Y, Z, t_P)$ at the observation position (X, Y, Z) and observation time t_P using

$$A(X, Y, Z, t_P) = \frac{\mu_0}{4\pi} \iiint \frac{\mathbf{J}(\rho, \psi, x, t) \rho \, d\rho d\psi dx}{R}, \quad (7.103)$$

⁵*i.e.*, the dielectric extends from a radius $\rho = \rho_0 - \frac{\Delta\rho}{2}$ to $\rho = \rho_0 + \frac{\Delta\rho}{2}$ and from $x = -\frac{\Delta x}{2}$ to $x = +\frac{\Delta x}{2}$, with $\rho_0 = 0.1254$ m, $\Delta\rho = 0.035$ m and $\Delta x = 0.004$ m.

⁶For algebraic convenience, this expression is complex. In comparisons with (real) experimental data, we must choose a phase, for example by taking the real or imaginary component of the final computed signal (but not both!).

where $R^2 = (X - x)^2 + (Y - \rho \sin(\psi))^2 + (Z - \rho \cos(\psi))^2$ and the retarded time $t = t_P - \frac{R}{c}$. The electric and magnetic fields at (X, Y, Z, t_P) are found from

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t_P} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A}. \quad (7.104)$$

Steps (7.103) and (7.104) are evaluated numerically (Section 7.5.2).

7.8.2 Results

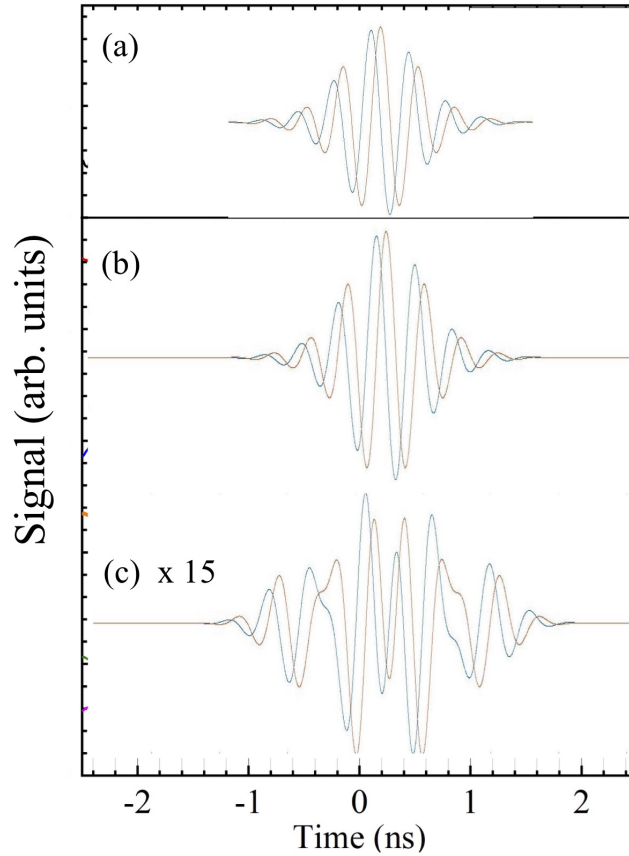


Figure 7.18: Calculated signal focusing for circular antenna TD 1; blue traces represent the real part of the complex exponential and brown the imaginary. (a) Broadcast E -field [*c.f.* Fig. 6.14(c) for the linear antenna]. (b) Detected signal ($\propto E$) at the focus polar angle $\theta = 27.3^\circ$. (c) Detected signal at $\theta = 40.0^\circ$. Note that signal (c) is much smaller in amplitude than (b); for clarity, it is amplified by a factor 15 compared to (b). In (b) and (c), the antenna to detector distance is 100 m.

Fig. 7.18 shows an example of the calculations for the emission of a polarization-

current wavepacket in antenna TD 1. The speed was set to give a focus at a far-field polar angle of $\theta = 27.3^\circ$ ($m = 3$ corresponding to $v/c \approx 2.2$); close to this angle, the detected signal [Fig. 7.18(b)] reproduces the shape and time dependence of the broadcast E -field [Fig. 7.18(a)]. As the detector polar angle moves away from the focus [*e.g.*, Fig. 7.18(c)], the detected signal falls dramatically in amplitude (by a factor ≈ 15 , *i.e.*, a reduction in power of a factor ~ 200) and its frequency content alters so that it differs markedly in shape from the broadcast signal.

7.8.3 Relevance to Pulsars and Speckle Patterns

Antenna TD 1 modelled in Fig. 7.18 is in some ways analogous to a pulsar, but of course works on much smaller lengthscales. Pulsars can potentially emit electromagnetic radiation via many mechanisms, including thermal emission and other processes in their hot, plasma atmospheres, and dipole radiation from the rotating magnetic field of the neutron-star core [93,94]; why then, is the pulsed radiation detected on Earth dominated by the small volume of superluminal polarization current? Fig. 7.18 provides an important clue. At the focus polar angle, the frequency content of the processes occurring within the rotating polarization-current element will reproduce *exactly*, and reinforce quasi-coherently to give a detected signal with greatly enhanced amplitude [*c.f.* Fig. 7.18(b)]. Mechanisms involved in generating this frequency content could include plasma resonance and electron cyclotron resonance, for example, both known to occur in pulsar atmospheres [93,94]. At all other observation angles, radiation from these (and other) processes will superpose incoherently, leading to a greatly reduced amplitude, and scrambled frequency content [*c.f.*, Fig. 7.18(c)]. It is therefore the sharp temporal focusing at the focus polar angle that allows the radiation produced by the superluminal (outside the light cylinder) mechanisms to dominate, and *not* the exotic nonspherical decay erroneously proposed by Ardavan (see Chapter 5).

Moreover, the “speckle patterns” generated by TD 1 and described in Section 3.7 can now be understood. At certain polar angles, portions of the rotating polarization-current source approach the metallic objects distributed around the FARM Range at the speed of light and with zero acceleration; therefore all of the temporal variations within these volumes of the source appear coherent at the scatterer. The scattered radiation therefore produces very strong interference patterns, resulting in the sharp, reproducible features. The effect is analogous to the speckle patterns generated by lasers (truly coherent sources of radiation [12]), but here, the effective coherence is generated by a very different mechanism.

Chapter 8

Epilogue

It's only forever, not long at all.

— Jareth the Goblin King

8.1 Current Work

The process of designing modern communication, sensing and radar systems poses ever-mounting challenges as the complexity of the platforms along with the performance metrics grows steadily [61]. Incremental parametric analysis and antenna pattern synthesis, followed by antenna reconfiguration, are traditionally used to assess the effects of modifying the design variables of an initial guess, which may either be based on prior knowledge or on a promising new contender [197]. In general, the use of a large set of design variables results in many potential solutions that, under constraints such as efficiency, operation frequency, electrical size and bandwidth, rarely yield a winning candidate that more than approximates the desired radiation pattern [61]. Moreover, if the design specifications call for an *in-situ*, high-fidelity analysis of the platform in which the antenna system is embedded (*e.g.*, the ceramic

armor of a tank, the deck of an aircraft carrier), traditional antenna optimization methods fail due to the level of geometrical detail, the complexity of the scattering problem, and the sheer size of the computational domain [198].

In previous chapters, we have considered simple variations in antenna parameters using only the most basic antenna geometries (circular, linear and curved); yet we were able to produce significant changes in the emitted radiation patterns. As mentioned sporadically throughout this work, extended faster-than-light sources are distinct from emitters that employ surface currents of free electrons on localized elements such as dipoles to produce radiation. They are true volume sources whose radiation characteristics are entirely dependent on the speed and acceleration of the moving polarization along with the size, shape, and dielectric constant of the solid that contains it. Changing any one of these parameters will – subtly or radically – alter the antenna’s “light curve”. The speed and acceleration of the polarization current squints the lobes in any desired direction; the dielectric shape accounts for the beam width, height of the side lobes and the position of nulls; and the dielectric constant, ϵ_r , determines the optimal operation frequency. In consequence, a superluminal antenna can be designed to emit a desired radiation pattern by changing (*i.e., optimizing*) the topology and material of the dielectric. Since dielectrics and metamaterials are easy to cut, shape or cast, novel topologies may be created using multivariate material and topological optimization methods, approaches which have, so far, hardly been pursued in antenna design [13], primarily due to the challenges associated with the fabrication of inhomogeneous materials, limited access to analysis tools and, frankly, in traditional antennas, the absence of a topology to be optimized. The optimized product can be cut or printed in three dimensions.

The design of algorithms to perform the optimization is very much an area of ongoing and vibrant research whose detailed description exceeds the scope of this (or any) dissertation. Hence, a brief outline of the chosen technique – a multivariate topological optimization algorithm on tetrahedral elements – will have to suffice.

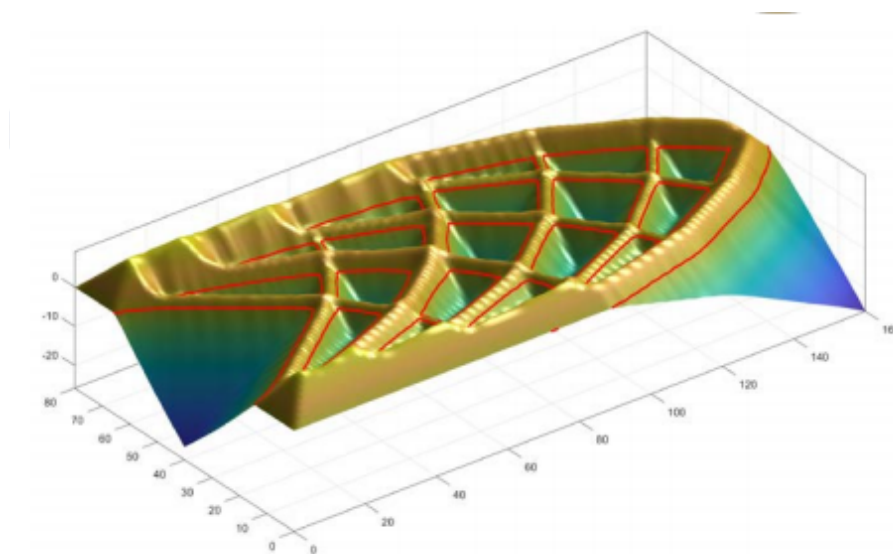


Figure 8.1: Structural optimization of an airplane wing. Credit: Alicia Kim

Though still in its infancy, the sequential linear programming level-set (SLIPLES) method has been used with success in aeronautical engineering by our friends at UCSD [200].

The level-set method (LSM) was developed as a mathematical tool for tracking the motion of interfaces in two or three dimensions [201, 202]. The crucial idea is to represent them using a discretized implicit function, which then evolves under a velocity field. This approach naturally allows for complicated phenomena to occur, such as interface merging and splitting, and the generation of voids within the material, which is crucial for our designs. In two dimensions, the level set method amounts to representing the shape boundary Θ using an auxiliary function ϕ , called the level-set function. Θ is represented as the zero level contour of ϕ , *i.e.*, $\Theta = \{(x, y) | \phi(x, y) = 0\}$ and manipulated “implicitly” by ϕ . When the curve moves in the normal direction with a speed v , then ϕ satisfies the level-set equation $\partial\phi/\partial t = v|\Delta\phi|$, where $|\cdot|$ is the Euclidian norm and t is time. This is a Hamilton-Jacobi partial differential equation that can be solved by using approved techniques of computational electromagnetics.

Adding sequential linear programming steps to the level-set algorithm allows for the handling of multiple constraints and simultaneously optimizes non-level set design variables. This requires a) the discretization of the boundary integrals to estimate function changes and b) the formulation of an optimization sub-problem. Replacing the discretized form of the velocity function V in the above equation by a boundary movement function z (*i.e.*, $z = \Delta t \cdot V$) removes the optimization problem from the (fictitious) time domain: The changes in the objective and constraints are now explicit functions of z , and the optimization sub-problem can force the boundary function to meet the constraints whilst improving the objective [203,204]. The challenge to adapting the SLIPLES algorithm to optimize superluminal antennas rather than airplane wings lies in calculating the discretized objective function after each deformation of the three-dimensional boundary, which poses a conceivable obstacle to rapid simulations. To avoid the potentially high computational cost of the procedure it is imperative to employ a powerful computational numerical technique. Amongst a plethora of potential candidates, we chose a discontinuous Galerkin time-domain method (DGTd) performed on a tetrahedral mesh.

The methods to generate an *in-situ* analysis of a superluminal antenna-system are still in their infancy. What is clear, however, is that complex domains can be decomposed into large fixed parts that have no parametric dependency and portions which contain the antenna system and are, therefore, electrically small. This naturally divides the calculations to be performed into “online” and affordable, upfront “offline” components and results in the same computational cost as if the antenna(s) were radiating into free space. Notice that the DGTd method can – with minor modifications – be employed to facilitate such a modular design composition. Once the program performs to satisfaction, additive manufacture may be used to produce a prototype, perhaps incorporating metamaterials [200], to design specification.

8.2 Back to the Future...

... and apologies for what must be the most overused reference in all of pop-culture history.

— The Present Author

We are nearing the end of this dissertation and, with it, the culmination of more than 15 years of analysis, research and experimentation. Over time, we succeeded in calculating and characterizing the electromagnetic radiation stemming from both, superluminal point sources and extended faster-than-light polarization currents. More recently, we returned to first principles and generalized the scalar three-dimensional wave equation to include sources that travel faster than the emitted Huygens wavelets. Unwilling to ignore the consequences of our calculations (and with a heavy heart), we bid farewell to the notion of “nonspherical decay”, a theory that, if proven correct, would have made us all (very) rich and famous indeed.

On a brighter note, we were able to show unequivocally that antennas driven by superluminal polarization currents are effective emitters of tightly focused packets of electromagnetic radiation that have a number of attractive qualities. They are true volume sources that inhabit an electrical insulator (rather than periodic structures of electrical conductors acting as point- or line-like sources) which makes them more efficient and resilient than conventional phased arrays. Since the constituent modular components of the Lightslingers are tasked with moving a polarization current through the dielectric rather than emitting a signal, the failure of individual elements does not impact the quality or content of the message. Moreover, such an antenna can potentially be shaped into almost any form, uniquely optimized for a particular application.

However, owing to the complexity of a discontinuous volume of integration, any calculation of a superluminal polarization current of arbitrary shape will involve

techniques of computational electromagnetics. Commercial software suites, although attractively packaged and readily available, will not suit our needs since they tend to be too slow for the *in-situ* analysis of a potential superluminal antenna-system. (In addition, Davidson [205] warns somewhat ominously, “[M]ore and more frequently, codes are being applied by users unfamiliar with the basic formulations underlying the codes, and not infrequently to problems for which the codes were not designed.”)

The nodal discontinuous Galerkin method mentioned in the previous section has been chosen for its local mass conservation property, great flexibility in handling nonmatching grids, elements of various types and shapes and, especially, potential for successful parallelization. Once a prototype MATLAB™ program has been shown to reproduce analytically derived results faithfully, much faster¹ C and C++ codes based on the same principles will be employed and integrated into Alicia Kim’s sequential linear programming level-set algorithm.

This is by far not the end. It is just the beginning. With microwave antennas counting in the tens of millions worldwide, even a small advantage in efficiency or versatility will result in big gains (literally and figuratively). Given sufficient funding and fruitful collaborations with universities and commercial entities, optimized superluminal polarization-current antennas might well provide the latter.

¹In numerical computing, ease of usage and raw speed seem to be eternally at odds. C++ has long been known to outdo MATLAB™ by a factor of at least 10, but is more difficult to master and certainly less forgiving.

Appendices

Appendix A

Equations (6.4) *et seq.* in Metrical Units

$$\begin{aligned}
 t_{\text{P}} &= t + \frac{R(t)}{c} \\
 &= t + \frac{1}{c} \left[z_{\text{P}}^2 + r_{\text{P}}^2 + r^2 - 2rr_{\text{P}} \cos(\varphi_{\text{P}} - \omega t) \right]^{1/2} \\
 &= h(t).
 \end{aligned} \tag{A.1}$$

$$\frac{dh}{dt} = 1 - \frac{\omega}{c} \frac{rr_{\text{P}}}{R} \sin(\varphi_{\text{P}} - \omega t) = 0. \tag{A.2}$$

$$t_{\pm} = \frac{1}{\omega} \left[\varphi_{\text{P}} - \arccos \left(\frac{c^2}{\omega^2} \frac{(1 \mp \Delta^{1/2})}{rr_{\text{P}}} \right) \right] \tag{A.3}$$

where

$$\Delta \equiv \frac{\omega^2}{c^2} \left[\left(\frac{\omega}{c} r^2 - \frac{c}{\omega} \right) \left(\frac{\omega}{c} r_{\text{P}}^2 - \frac{c}{\omega} \right) - z_{\text{P}}^2 \right] \tag{A.4}$$

$$\frac{d^2 h}{dt^2} = \frac{\omega^2}{c} \left(\frac{r_{\text{P}} r \cos(\varphi_{\text{P}} - \omega t)}{R} \right) - 2 \frac{\omega^2}{c} \left(\frac{r_{\text{P}} r \sin(\varphi_{\text{P}} - \omega t)}{R} \right)^2 = 0. \tag{A.5}$$

Appendix B

Circular Antenna TD 1: Design Details

A summary is given here of the design processes involved in building circular superluminal antenna TD 1. This work was carried out primarily by Zhi-Fu Wang, Quinn Marksteiner, Frank Krawczyk, Bill Romero and John Singleton, with input from Jim Potter and Larry Earley. The details are summarized from various reports that are no longer generally available, and are included here for completeness. This information should enable the reader to judge that TD 1 is likely to be a reliable source of radiation due to superluminal, volume-distributed polarization currents and therefore worthy of modeling.

B.1 Dielectric Element Design for TD 1

The design shown in Fig. 3.3 requires approximately radial electric fields to excite the polarization currents within the dielectric around the periphery of the circular antenna. The circle is broken up into 72 narrow wedges, each subtending 5° . Each of these elements is driven separately from a coaxial line connected to the control

electronics. The 72 individual drives allow voltages corresponding to Eq. 3.3 to be imposed on each element. The carrier frequency $f = \omega/2\pi$ for the radiation pattern was chosen to be between 2.4 and 2.5 GHz. Modulation frequencies $\eta/2\pi$ of up to 200 MHz are imposed; the transmission must be broadband with a total range of 20%, to avoid reflections back into the driving electronics. Several challenges had therefore to be overcome.

1. The drive signals are transported to the antenna in $50\ \Omega$ coaxial lines. Thus, each of the 72 wedges must provide a transformation from a coaxial TEM mode to the radial electric field applied across the dielectric slabs. While this type of drive is common in accelerator structures or other resonant circuits [61], where TEM modes feed into eigenmodes of a totally different pattern, a more complex, field-shaping geometry is required for what is in effect an open antenna.
2. The dielectric that hosts the polarization current represents a large impedance mismatch with air; the effective impedance is $\sim 10s\ \Omega$ in the dielectric, compared to $377\ \Omega$ in free space. The design is a compromise between selecting a material of a sufficiently large permittivity (to obtain a strong polarization as a result of the radial field) and a reduction of the reflection at the mismatch. Mismatch conditions also need to be considered at the transition from the $50\ \Omega$ coaxial cables to the propagation region.
3. To generate radial fields, distinct anode and ground electrodes need to be provided on either side of the region where the polarization-current pattern propagates. This requires a break in the conduction path between opposite sides of the dielectric, conflicting with the need to shield the fields from radiating out from anywhere else *but* the dielectric. The break for the RF wall currents, while maintaining very low leakage of RF fields was a major design driver.
4. There is a need to impose very accurate timing, amplitude and frequency variations onto the radiation pattern around the circumference of the dielectric

circle. This requires a design of the individual wedge elements that provides very good reproducibility in assembly.

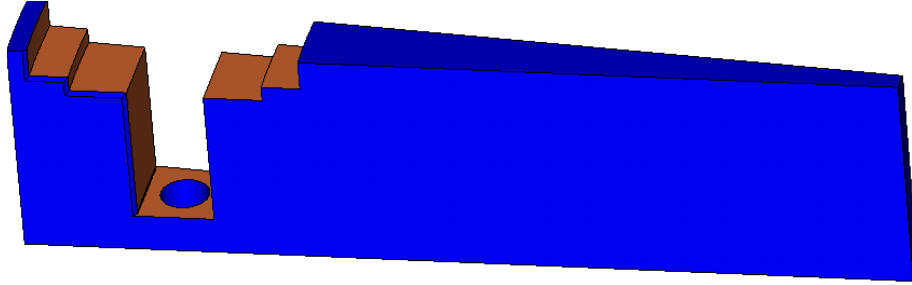


Figure B.1: Wedge-shaped individual antenna element: the body is made up of G10 (blue). The cut-out is the propagation region for the dipole (*i.e.*, radial) mode. The sides of the cut-out are plated (copper color) for shielding purposes. Credit: Frank Krawczyk

Fig. B.1 shows the body of the wedge-shaped element made from G10, a machinable, fiberglass-epoxy composite (in photographs of the antennas, *e.g.*, Fig. 3.1, this material looks grey-green). The cut-out is the propagation region. The coaxial mode couples to this region from the bottom, then it is transformed into a radial pattern and transported to the top, where it excites the polarization current; this, in turn, radiates into free space. The cutout is located so that the radius at its center is one wavelength at 2.4 GHz (125 mm).

The critical element for the field-shaping transformation from the TEM mode to the radial mode is a connection of the coaxial inner conductor to one side of the rectangular propagation region (electrode) and the connection of the outer conductor to the other side of this region (ground). This establishes the radial field in the dielectric at the top of the slot. The connection to the ground is straightforward due to the good accessibility of the outer conductor. The connection of the inner conductor requires careful shaping to establish a smooth change in impedance, maximizing transmission of the signal. Fig. B.2 shows the shaped center conductor providing a current path from the coaxial inner conductor to the side electrode.

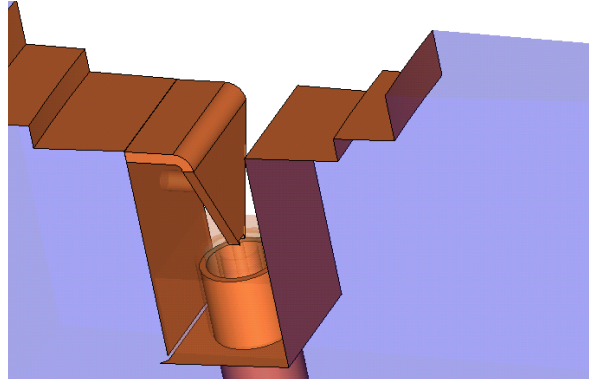


Figure B.2: Field shaping element; the coaxial inner conductor (not shown for clarity) is connected to the left side of the plating by an element that shapes the electric field and provides a slow impedance change from bottom to top. The outermost conductor of the coaxial lines contacts the ground at the penetration from the bottom to the propagation region. Credit: Frank Krawczyk

The need for shielding of the dipole mode toward the bottom and the sides (in the radial direction) requires an elaborate separation between electrode and ground plating. This must as far as possible prevent RF leaking out into the environment, or else the signal from the polarization-current emission will be contaminated. The solution is to use a coaxial choke with low field at the location of the break; this in effect functions as a sleeve balun [61]. This structure was implemented using two concentric coaxial cables. The inner cable carries the signal into the structure, and the outer is shorted at the bottom and has a break in its outer conductor. The distance between short and break is approximately one quarter of a wavelength.

The two cables are flush at the top, where the full wave exits into the structure. In the outer cable a quarter wave builds up with a null close to the break (see Fig. B.4 (left)). Only a small portion of the wave radiates out of the break. The propagation region is completely plated to prevent radiation from leaving anywhere else than at the top of the wedge. This plating is only broken in one thin line azimuthally across the wedge. The use of a break at the coax to minimize leakage and a second break where electrodes needed to be separated results in a low leakage.

To obtain a propagating polarization current radiating into free space, a dielectric medium is used as the top layer of the antenna wedges; this hosts the polarization current. For a large amplitude of the radiated fields one should use a dielectric with a large volume and/or a high relative permittivity, ϵ_r . A very large permittivity will in some circumstances aggravate the impedance mismatch between the dielectric top layer and free space (we return to this in Appendix C.2). Hence, so-called *Alumina 99%*, with $\epsilon_r \approx 10$, was employed. To further address the impedance mismatch, thin slabs of dielectric (4 mm) were chosen, but spanning a relatively large radial extent to increase the volume. Fig. B.3 shows the radial cross-section of the top of the antenna.

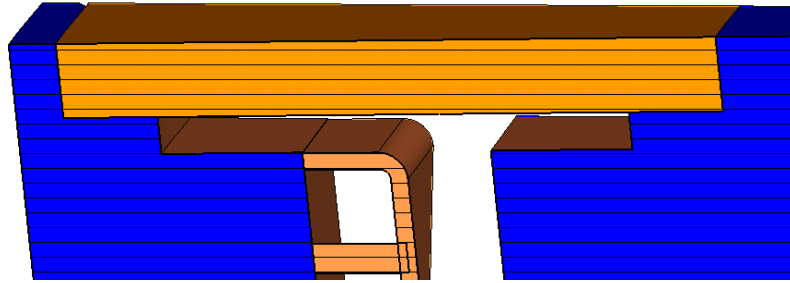


Figure B.3: The top of each antenna element is filled with alumina. The radial field in this region sets up the polarization current. To minimize impedance mismatch, the slab has been made relatively thin and wide (radially). The increase in radial dimension without a large discontinuity in impedance change has been accomplished using two steps, one filled with air and the second containing the alumina. Credit: Frank Krawczyk

To provide a complete record of the design, Fig. B.4 (right) shows other features that help in lowering the RF leakage (a copper block that accepts the concentric coaxial lines), and improving the impedance transition along the propagation path (a shorting pin between the field shaping device and one of the electrodes). In subsequent tests, the copper block was found to play a very limited role in the shielding and impedance matching; it was omitted from the elements used in antennas TD 2 and TD 3.

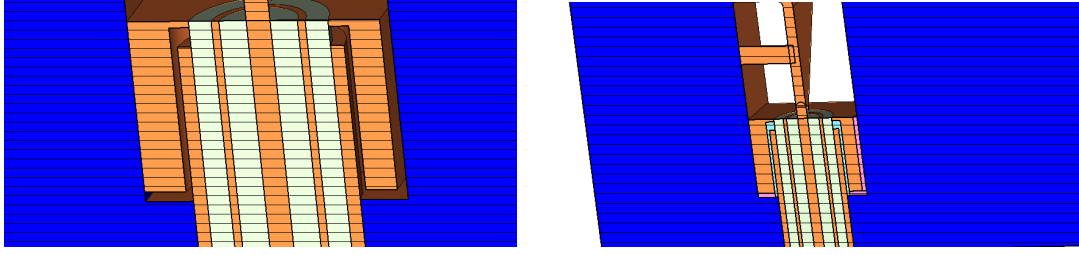


Figure B.4: (left) Shielded break in RF-conduction; this cross-section (solids are indicated by hashing) shows the concentric coaxial lines and a shielding block. A gap between the block and the coaxial lines provides the primary break in electrical contact. The block shields the second break between the electrode and ground plating (lower left of cut-out). The low leakage through the breaks can be attributed to the choice of break locations and the effect of the outer (larger) coaxial line, in which a quarter wave builds up with a minimum at the primary break location. (right) Additional impedance match features: the copper block below the field-shaping element and the pin between the field-shaping element and the electrode are additional features to help the mitigation of impedance mismatches. Credit: Frank Krawczyk

B.2 Performance of Antenna Element Wedges.

Simulations of the design were carried out by Frank Krawczyk using CST Microwave Studio; a model of an individual wedge was studied employing azimuthal periodic boundary conditions; these assumed that the wedge was surrounded by others excited with identical signals, *i.e.*, $\Delta\phi = 0$). Desired characteristics were the good transmission of power into the wedge, characterized by S_{11} at the coaxial feed. Fig. B.5 shows the range where the return loss is better than -10 dB; it includes the required band from 2.16 to 2.64 GHz. A snapshot of the electric field amplitude indicates the field shaping from coaxial mode to a radiating dipole pattern [Fig. B.6].

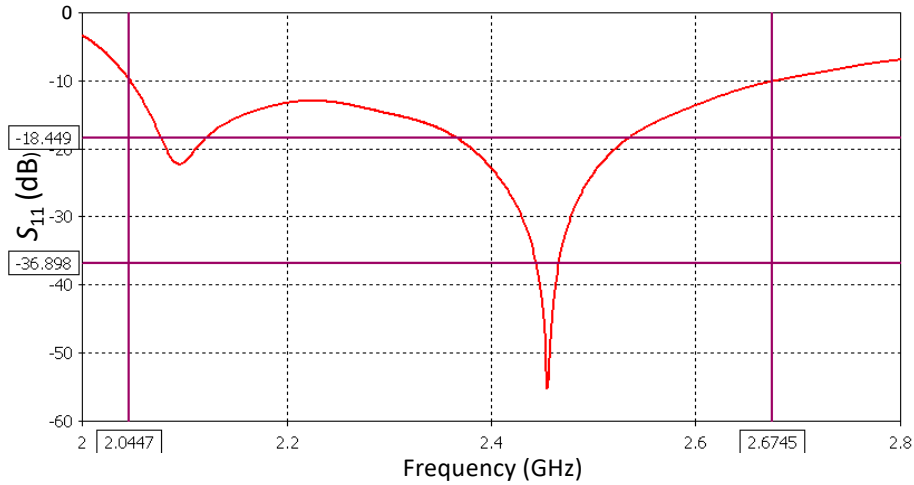


Figure B.5: S -parameter analysis for an individual wedge with periodic boundary conditions: a good measure for the suitability of the wedges is the S_{11} parameter that describes reflections from the antenna element back into the coaxial drive. Assuming a reflection of -10 dB or better as definition for the bandwidth, the wedges have a bandwidth of 630 MHz between 2.04 GHz and 2.67 GHz, which includes the required range of operation from 2.2 GHz to 2.6 GHz or 2.4 ± 0.2 GHz. Credit: Frank Krawczyk

B.3 Control Electronics for TD 1

The phase and amplitude of the voltage sent to each antenna element is controlled by a pair of AD8349 RF vector modulators (VM)¹.

The operation of VMs is described on the manufacturer's website, but will be summarized here; specifications are given in Table B.1. Fig. B.7 shows VM input and output connections; the device accepts signals RFIN I and RFIN Q of angular frequency ω_1 from a 90° splitter (quad splitter). Ideally, the quad splitter will produce a pair of equal-amplitude inputs

$$\text{RFIN I} = \sin(\omega_1 t) \quad \text{and} \quad \text{RFIN Q} = \cos(\omega_1 t). \quad (\text{B.1})$$

The output of the VM is controlled by voltages $I(t)$ and $Q(t)$ applied across the

¹Vector modulators are sometimes referred to as *vector multipliers*.

Table B.1: Manufacturer's specifications for the AD8349 Vector Modulator (VM).

1 degree rms quadrature error @ 1.9 GHz
0.2 dB I/Q amplitude balance @ 1.9 GHz
Broad frequency range: 0.8 GHz to 2.7 GHz
Sideband suppression: 46 dBc @ 0.8 GHz
Sideband suppression: 36 dBc @ 1.9 GHz
Modulation bandwidth: dc to 70 MHz
0 dBm output compression level @ 0.8 GHz
Noise floor: 147 dBm/Hz
Single 2.7 V to 5.5 V supply
Quiescent operating current: 45 mA

IBBP/IBBM and QBBP/QBBM terminals respectively. Each of these is a differential input centered around a common mode-voltage $V_c \approx 0.5$ V:

$$I(t) = \text{IBBP} - \text{IBBM} \quad \text{and} \quad Q(t) = \text{QBBP} - \text{QBBM}. \quad (\text{B.2})$$

Voltages $I(t)$ and $Q(t)$ can either be DC inputs or RF inputs with a frequency of up to 240 MHz. The (theoretical/ideal) output of the VM is given by

$$V_{\text{out}}(t) = I(t) \sin(\omega_1 t) + Q(t) \cos(\omega_1 t) \quad (\text{B.3})$$

In active antennas such as TD 1, we use DC voltages $I(t) = I_1$ and $Q(t) = Q_1$, resulting in an output signal

$$V_{1\text{out}} = I_1 \sin(\omega_1 t) + Q_1 \cos(\omega_1 t), \quad (\text{B.4})$$

which has amplitude $A_1 = \sqrt{I_1^2 + Q_1^2}$ and phase $P_1 = \tan^{-1}(I_1/Q_1)$. The phase is adjusted through 360° by using both positive and negative values of I_1 and Q_1 .

For each antenna-element feed, two VMs are used to control the amplitude and phase of two input signals with angular frequencies ω_1 and ω_2 ; these signals are

then added together using a splitter. If the two VM output amplitudes A_1 and A_2 are chosen to be equal, such that $A_1 = A_2 = A$, then the resulting signal is $V_{\text{out}} = A [\cos(\omega_1 t + P_1) + \cos(\omega_2 t + P_2)]$, or

$$V_{\text{out}} = 2A \cos \left[\frac{1}{2}(\omega_1 + \omega_2)t + \frac{1}{2}(P_1 + P_2) \right] \cos \left[\frac{1}{2}(\omega_1 - \omega_2)t + \frac{1}{2}(P_1 - P_2) \right]. \quad (\text{B.5})$$

Thus, the output waveform is the desired modulated signal, with carrier frequency $\omega = (\omega_1 + \omega_2)/2$, carrier phase $P_c = (P_1 + P_2)/2$, modulation frequency $\eta = (\omega_1 - \omega_2)/2$, and modulation phase $P_m = (P_1 - P_2)/2$. In order to reproduce Eq. (3.3), the modulation phase on each antenna element is held constant; the carrier phase is adjusted to match Eq. (3.5). This procedure is rendered somewhat more transparent if Eq. (3.3) is rearranged in the following form

$$V_j = \frac{V_0}{2} (\cos[(\omega + \eta)t - \omega j \Delta t] + \cos[(\omega - \eta)t - \omega j \Delta t]). \quad (\text{B.6})$$

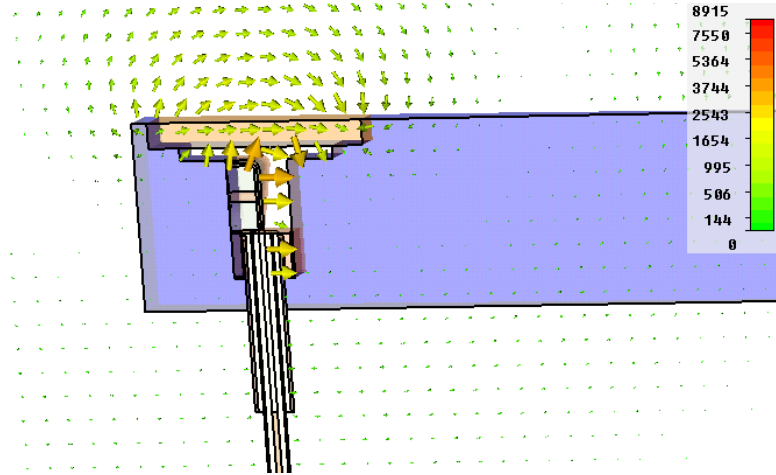


Figure B.6: Snapshot of the dipole electric field radiation pattern: the electric field in the central cut-plane shows the field amplitude and direction mostly in the propagation region. The pattern in the coaxial region is not visible due to overall scaling issues. The radiation pattern with only small field contributions (logarithmic scale) elsewhere exhibits the desired field transformation. Credit: Frank Krawczyk

It will be seen that V_j may be determined by using two VMs to independently

adjust the amplitude and phase of two cosinusoidal voltages at frequencies $f_- = (\omega - \eta)/2\pi$ ($\equiv \omega_1/2\pi$) and $f_+ = (\omega + \eta)/2\pi$ ($\equiv \omega_2/2\pi$).

Fig. B.8 shows the output waveform from two VMs used in this way. The signal is clearly of the correct form (*c.f.* Eq. (B.5)). Fig. B.9 shows the equivalent measured frequency spectrum; as desired, the 2.2 GHz and the 2.6 GHz amplitudes are almost equal, and contributions at other frequencies are suppressed by more than -30 dB. The modulation and carrier phases P_m and P_c are controlled by setting the phases of

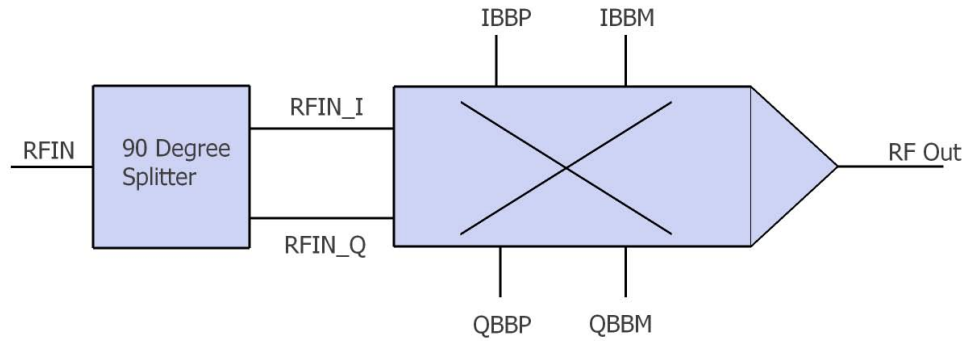


Figure B.7: Schematic of the Vector Modulator (VM) inputs and outputs. VMs are used in TD 1 to generate a single sideband of the RF signal for each antenna element.

each individual sideband, P_1 and P_2 , which are in turn controlled by the four inputs I_1 , Q_1 , I_2 , and Q_2 . The external control of the I and Q inputs is accomplished by using Linear Technologies LTC1990 digital to analog converters (DACs). Each individual DAC has eight outputs, which can control the four DC I and Q inputs for a single output waveform (each I and Q input require two voltages to maintain the 0.5 V common-mode offset). These DACs are controlled by serial communication from an 8051 processor, and can be daisy-chained so that (for example) the 72 elements required for TD 1 are controlled from just two processors.

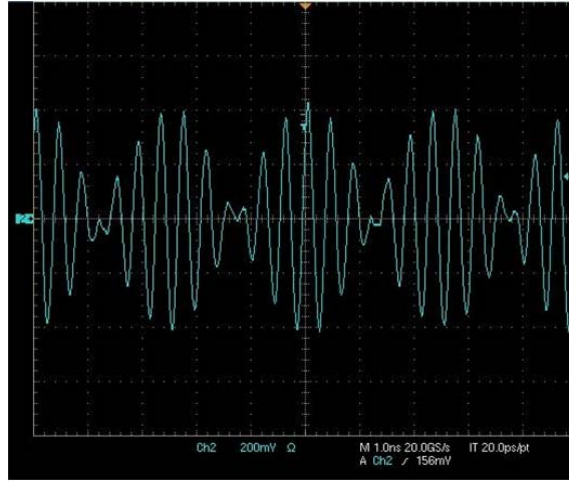


Figure B.8: Oscilloscope signal of amplitude-modulated waveform with a 2.4 GHz carrier frequency and 200 MHz modulation frequency.

B.4 Labview Control Software.

A scalable² Labview amplitude and phase control routine originally developed by Quinn Marksteiner is used to control the circuits that set the parameters used in Eq. (B.3). The routine first uses these equations to set the amplitude and phases, and then iteratively compensates for errors in the components and wedge antenna elements until the desired phase and amplitude has been reached.

This approach requires monitoring of the signal sent to each antenna element using a Tektronix TDS 7404 digital oscilloscope. A coupler attached to each antenna element takes $\frac{1}{10}$ of the RF power into a binary switchstack that is digitally controlled to compare any selected element signal with a reference signal, using two channels of the oscilloscope. The TDS 7404 can sample two channels at 10 GS/s (*i.e.*, 10^{10} samples per second), which is more than adequate for the maximum frequency of 2.6 GHz used. The LabVIEW program measures the phase of each sideband by taking a fast Fourier transform of a large sampling time of data, resulting in a res-

²In this context “scalable” means that the same Labview routines are used in the control programs for the 72-element TD 1 and the 32 element TD 2 and TD 3.

olution of a fraction of a degree. This was shown to be better than the resolution that could be achieved by looking at the signal directly in the time domain. In

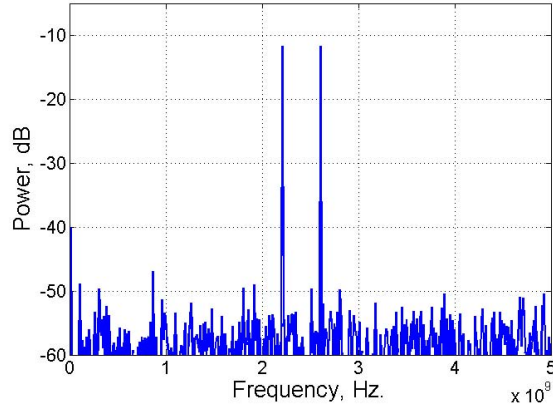


Figure B.9: The frequency spectrum of the signal from two VMs.

the active antennas, the connection between the antenna elements, VMs and phase measuring system is made using identical circuit boards and a minimum of cabling. Despite this, it was found that channel-to-channel phase and amplitude differences were introduced by non-identical behavior of various components, chiefly the dielectric antenna elements and some parts of the switchstack. There are also significant variations in performance of the VM chips. In the test phase of TD 1, the current author measured the phase and amplitude of the electric field at the surface of each dielectric element using a compact dipole antenna, allowing the variations due to the components to be assessed. The measured phases and amplitudes were iteratively improved by reprogramming the voltages fed to the VM circuits, until satisfactory agreement with the desired phase and amplitude at each element was achieved. The resulting VM input voltage values have therefore been shifted to compensate for the differences between desired and measured phase and output voltage for each element due to component errors; these shifts were then stored in a look-up table in the control computer. It was found necessary to create a look-up table for each speed (m) setting used in TD 1.

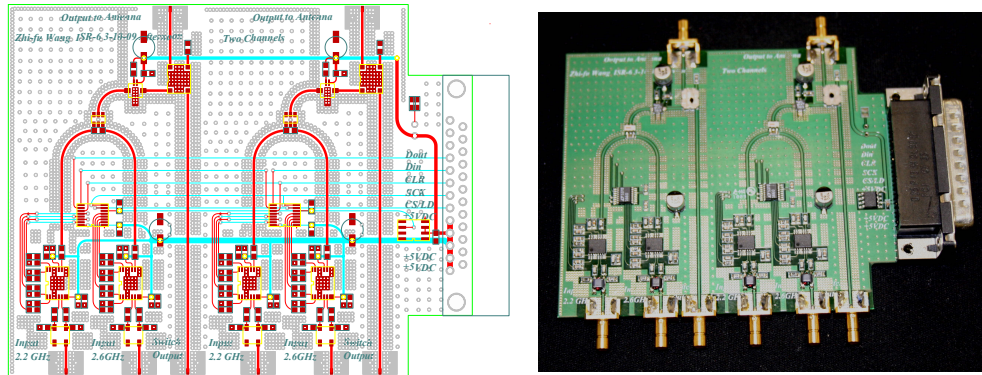


Figure B.10: Left: PCB Artist layout of the four VMs (to control signals to two wedge antenna elements) on their PCB. The lower SMB connectors are (from left to right) 2.2 GHz in, 2.6 GHz in, element monitor signal out, 2.2 GHz in, 2.6 GHz in, element monitor signal out. The upper SMB connectors are the outputs to the two antenna elements. The bus plug on the right brings in power and instructions to the D-to-A converters that set the I and Q values for each VM. Right: a photograph of the finished board.

On setting a speed (m value), initial values of the amplitude and phase are fed to the VM chips. The switchstack then connects each antenna element's coupler to the oscilloscope, so that the phase and amplitude at each frequency, on each element, can be compared with a reference signal, making allowance for the offsets stored in the look-up table for the selected speed. These values are directly compensated for in the Labview code, which iteratively shifts the voltages supplied to the VMs until satisfactory agreement is achieved. At higher speeds [smaller values of m (1, 2, 3)], this compensation method reduces the error in phase to $\leq 2^\circ$, and the amplitude errors to ≤ 0.3 dB. At slower speeds ($m \geq 4$), the errors are worse (5° and 1 dB respectively), for reasons that will be explained in Appendix C.2.

B.5 Circuit Boards and Final Assembly

The initial demonstrations of all of the above control systems was achieved using discrete components and VM demonstrator boards, connected by coaxial cable and

used with the eight-element TD 0. The final 72 channel TD 1 system uses custom printed-circuit boards (PCBs) that were fabricated by Advanced Circuits Corp. All of the RF components used in the custom PCBs were surface-mount versions of the packaged components used in the 8 channel TD 0 prototype. The custom PCBs were designed by Zhi-Fu Wang (LANL) using the commercial software PCB Artist. Each board uses careful layout to maintain equal signal paths to allow easier phase control of the 72 channels. Each PCB has four layers and strict $50\ \Omega$ traces for all signals. The PCBs employ SMB connectors for the coaxial cable interconnects. An example is shown in Fig. B.10.

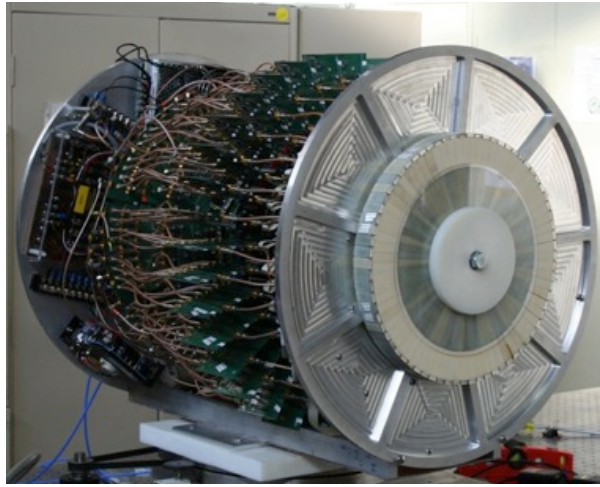


Figure B.11: Enlargement of Fig. 3.1(a), showing the TD 1 antenna without its casing. The cylindrical arrangement of the circuit boards is clearly visible. Those nearest the front of the antenna contain the VMs and the phase monitoring electronics.

The PCBs are designed so that they could be assembled into the “wedding cake” arrangement shown in Fig. B.11. This has the advantage of keeping all cables the same length, but the disadvantage that some of the circuit boards (those at the rear) are also wedge-shaped!

The frequencies $f_- = (\omega - \eta)/2\pi$ and $f_+ = (\omega + \eta)/2\pi$ are provided by a pair of NovaSource G6 signal generators, each of which feeds a splitter/amplifier tree

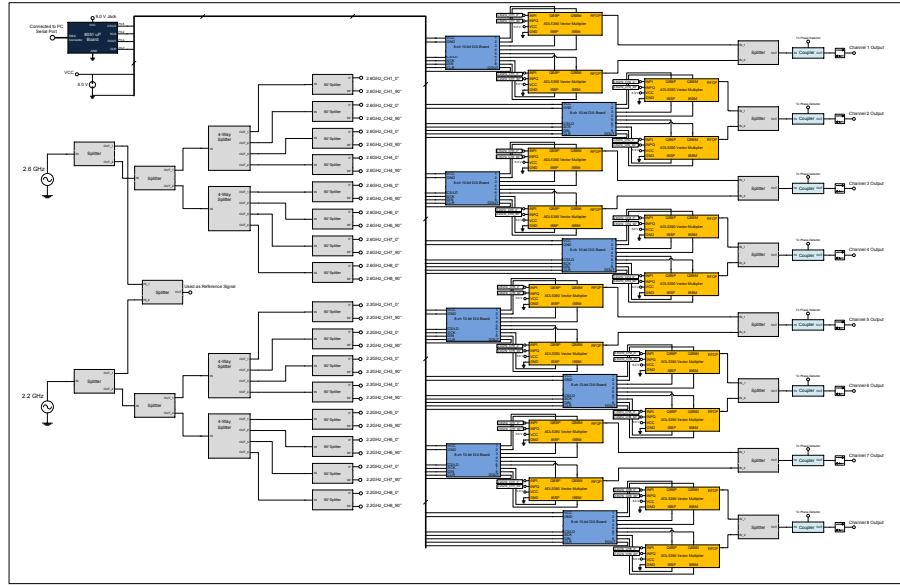


Figure B.12: Block diagram for one of the nine 8-channel sections of TD 1. The left section depicts the splitters (grey). The first row shows the 2.2 GHz and 2.6 GHz signals coming from the signal sources after initial splitting and amplification. These are split, with one half being recombined as a reference channel for phase comparison. Further to the right, the signals are again split, first by two-way splitters, then by four-way splitters and finally by quad (90°) splitters; the latter provide in-phase and quadrature signals to feed the VMs. To one side, the black box represents one of the 8051 microprocessors which instructs the D-to-A converters (there are two in TD 1). The right section shows the VMs (orange), the 8-channel, 10 bit D-to-A converters (blue) that provide the voltages for the phase-setting inputs of the VM, the combiners (grey) and the couplers (pale blue). The latter take $\frac{1}{10}$ of each channel's signal to the binary switchstack for phase comparison with the reference signal.

that drives the ADL5390s. Either or both of the internal signal generators may be replaced by an external source for testing modulation schemes. The manufacturer's specifications for the signal generators is given in Table B.5; frequencies were set from the control computer.

The final, assembled, TD 1 is shown in Fig. B.11; it has the NovaSource signal generators, microprocessor and DC power supplies in the rear (supported by a circular aluminum plate), the distribution boards in the middle and the amplitude and

Table B.2: Manufacturer's specifications for the NovaSource G6 signal generators used to provide the primary frequencies for the TD 1 antenna.

Frequency tunable	2.0 to 3.0 GHz
SSB phase noise	-60dBc/Hz@1kHz removed
Spurious signals	-60dBc
Harmonics	-35dBc

phase control PCBs as close to the antenna elements (supported by another circular aluminum plate) as possible. There is a central mechanical support that connects the two end plates and holds all of the boards in place, with longitudinal PCBs used as bus boards to provide the interconnections. A lower support connects the whole assembly to a turntable that allows the antenna to be rotated. When in use, the sides of TD 1 are covered by a cylindrical copper mesh. Rather than preventing the broadcast of stray signals from TD 1's electronics, this was found necessary to shield the innards from *outside* interference; in particular, the outdoor range at Sandia National Laboratories was host to several competing RADAR experiment that impeded the phase setting procedure if the cover was not in place.

Finally, Fig. B.12 shows the block diagram for eight of the 72 channels of TD 1. The complete antenna comprises 9 identical blocks.

Appendix C

Linear Superluminal Antenna TD 2

C.1 Design Summary

TD 2 was designed and built by Zhu-Fu Wang and Frank Krawczyk under the overall supervision of John Singleton. Labview control software was provided by Quinn Marksteiner. Subsequent modifications to the design were made by Jamie Wigger and John Singleton.

The TD 2 antenna elements are rectangular versions of the TD 1 “wedges” and the control system functions very similarly to that of TD 1. Rearranging Eq. (3.6)

$$V_j = \frac{V_0}{2} (\cos[(\omega + \eta)t - \omega j \Delta t] + \cos[(\omega - \eta)t - \omega j \Delta t]), \quad (\text{C.1})$$

it is seen that V_j may be determined by adjusting the amplitude and phase of cosinusoidal voltages at frequencies $f_- = (\omega - \eta)/2\pi$ and $f_+ = (\omega + \eta)/2\pi$. As before, this is accomplished by driving each element of the antenna using two Analog Devices AD8349 Vector Modulator chips, one to control the amplitude and phase of the signal at frequency f_- and the second to control that at f_+ . Ten percent of the signal delivered to each element is split off and used for phase and amplitude monitoring of the various V_j . A binary switch stack allows the monitor signal of each element in

turn to be connected to a Tektronix TDS 7404 digital phosphor oscilloscope, where it is compared to a reference signal.

The control blocks are thus identical to those in TD 1, but as there are only 32 antenna elements only four blocks, each containing 16 VM chips, are required. Once again, the frequencies f_- and f_+ are provided by a pair of NovaSource G6 signal generators, each of which feeds a 32-way splitter/amplifier tree that drives the VMs. Either or both of the internal signal generators may be replaced by an external source for testing modulation schemes, for example.

Rather than specifying phases as in TD 1, the speed is specified using time intervals. When constant speeds are used, the time interval Δt is the same for each element. However, the polarization current may also be accelerated, by making Δt vary across the array of elements.

Speed setting and general control of TD 2 is done via a LabView program running on a control computer, which communicates with the oscilloscope via a GPIB bus. The same computer also controls the telescope mount that changes the orientation of the antenna, and records data from the two spectrum analyzers used to monitor emitted radiation.

C.2 Retardation and Field Confinement for $v/c < 2$

Chapters 3 and 7 allude to two unplanned features of antennas TD 1 and TD 2:

- (i) For source speeds $1 \leq v/c \leq 2$, TD 1 and TD 2 give less output power than expected; *i.e.*, they are less efficient at these low speeds than they are for $v/c \geq 2$;¹
- (ii) The control software experienced difficulties setting the phases of frequencies f_+ and f_- at speeds $v/c \leq 2$. This issue was studied by Jamie Wigger, Connor Bailey

¹At source speeds $v/c \geq 2$, tests and models show that the superluminal antennas have better efficiency and directivity than an equivalently sized phased array.

and John Singleton using antenna TD 2. Their experimental data were analyzed applying numerical models developed by the present author (Chapter 7). The results are of considerable importance for future antenna design.

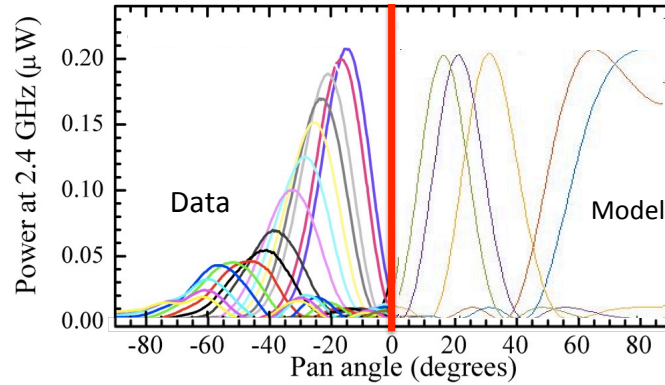


Figure C.1: Comparison of experimental data from linear antenna TD 2 run at several constant speeds $-1.1 > v/c > -4.1$ (measured in the TA-35 anechoic chamber; left of red line) with numerical model evaluated for $v/c = 1, 1.1, 2, 3, 4$ under the same conditions (right of red line). The source-to-detector distance was 5.36 m and $f = 2.4$ GHz. in both cases.

C.2.1 The Problem

The numerical models of Chapter 7 provide excellent quantitative predictions of the output power of linear antennas such as TD 2 for speeds $v/c \geq 2$ and $v/c < 1$. However, at intermediate superluminal speeds, the quantitative agreement is less good. Fig. C.1 shows a comparison of experimental data from TD 2 run at several constant speeds $-1.1 > v/c > -4.1$ (measured in the TA-35 anechoic chamber; left of red line) with numerical model results calculated for $v/c = 1, 1.1, 2, 3, 4$ under the same conditions (right of red line). The source-to-detector distance was 5.36 m and $f = 2.4$ GHz. Though it gets the lineshapes and peak position correct, the original form of the numerical model does not reproduce the experimental fall-off in power as the speed decreases towards c . We now explore the reasons for this disagreement.

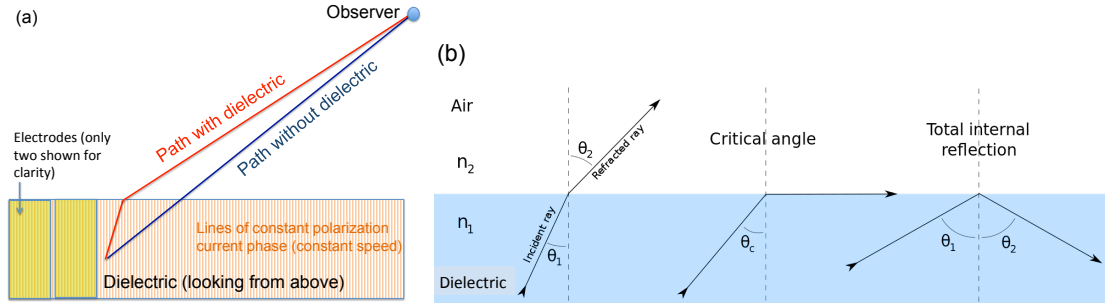


Figure C.2: (a) Plan view of a linear antenna running at a constant speed, showing lines of constant polarization-current phase (orange). Two electrodes are shown (yellow); the rest are omitted for clarity. According to Fermat's principle, electromagnetic disturbances follow the path with shortest transit time from source to observer. In the absence of the dielectric, this would be a straight line (dark blue), but with the dielectric present, the path includes a change in direction at the surface of the dielectric (red). (b) Schematic of electromagnetic waves moving from a dielectric to air; for large enough angles of incidence, there is total internal reflection. Credit: John Singleton

The reason is that the numerical model originally treated the polarization current as if it were naked, without examining the effect of the host dielectric. The mechanisms involved - which we refer to as *retardation* and *confinement* - are illustrated in Fig. C.2. We will discuss each in turn, and then also describe a *control issue* associated with running active antennas at slow speeds.

C.2.2 Retardation

Fig. C.2 (a) shows a schematic view of a linear antenna similar to TD 2 running at a (fairly slow) speed (no acceleration), showing lines of constant polarization-current phase (orange). According to Fermat's principle, electromagnetic disturbances follow the path from source to observer with the shortest transit time. In the absence of the dielectric, this would be a straight line (dark blue), but, with the dielectric present, the path includes a change in direction at the surface, leading to a longer distance travelled (red), or, to put it another way, a longer transit time. We refer to this as *retardation*; it causes additional phase differences between the signals emitted from

different depths in the dielectric, altering the expected antenna pattern. The effect will be reduced at higher speeds, where the central lobe is emitted in a direction that is close to orthogonal to the polarization-current path, but for slower speeds and thick dielectrics it can cause potentially serious errors. This is illustrated by the simulation shown in Fig. C.3 for a linear antenna running at $v/c = 1.1$. The blue curve shows the effect of introducing the phase delays caused by the presence of the dielectric, which in this case is 25 mm thick (all other parameters are the same as for the 348 mm-long TD 2). The black curve is for the equivalent naked polarization current. The desired peak emission is suppressed by about 30 dB. Note that the subsidiary lobes emitted closer to $\phi = 0$ are much less affected.

There is an interesting contrast between the numerical data for the thick linear antenna (Fig. C.3) and those for the thick circular antenna (Fig. 7.14). In the latter case, the spread of source speeds due to the finite radial width of the source ($\Delta\rho = 35.0$ mm), averages out the equivalent of the sharp fringes seen close to $\phi \approx \pm 80^\circ$ in the case of the linear antenna.

C.2.3 Confinement

Fig. C.2(b) shows light propagating in dielectric towards an interface with a medium (such as air or vacuum) with a lower dielectric constant. As the angle of incidence rises, eventually the light will undergo total internal reflection. For the interface between a non-magnetic material with relative permittivity ϵ_r and air or vacuum, the critical angle at which this starts is $\sin^{-1}(\epsilon_r^{-\frac{1}{2}})$. For alumina, this corresponds to $\phi \approx 18^\circ$. Whilst this picture is usually applied to situations where the wavelength of light is much less than the size of the dielectric components involved, even at longer wavelengths the effect inhibits the emission of radiation from the superluminal antennas at larger angles. Empirically, the emitted power seems to decrease like $\cos^2 \phi$, suggesting that the emission comes from a finite thickness “skin” at the surface

of the dielectric (interface charge). This is a sign of an impedance mismatch between alumina and air due to the “total internal reflection” effect once the emission angle approaches 90° .

C.2.4 Active Phase Control at Slow Speeds.

The confinement effects discussed in the previous section result in reduced emission as the speed falls towards c . Consequently, power is reflected back into the antenna elements, increasing S_{11} and enhancing the amount of crosstalk between adjacent antenna elements.

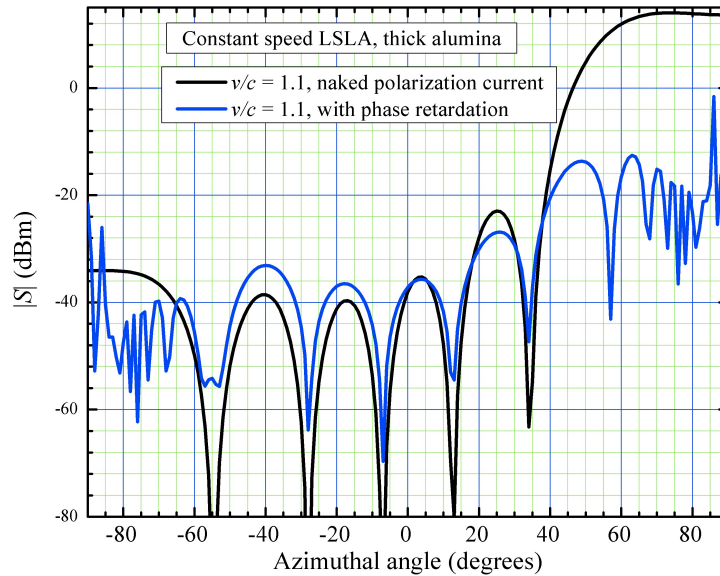


Figure C.3: Simulation of a 25.0 mm thick linear superluminal antenna (LSLA) running at $v/c = 1.1$ and $f = 2.4$ GHz; all other parameters are the same as for TD 2. The black curve is for a “naked” polarization current. The blue curve shows the effect of introducing the phase delays caused by the presence of the dielectric. In this particular case (slow speed, thick dielectric), the desired peak emission is suppressed by about 30 dB.

Such effects exacerbate the difficulties in setting the large phase differences between adjacent elements that are necessary for low speeds. When the phase difference between adjacent antenna elements is $\approx 5^\circ$ corresponding to $v/c \approx 7$, a

small amount of crosstalk is irrelevant; when the phase differences are around 30° ($v/c \approx 1$), crosstalk is much more serious.

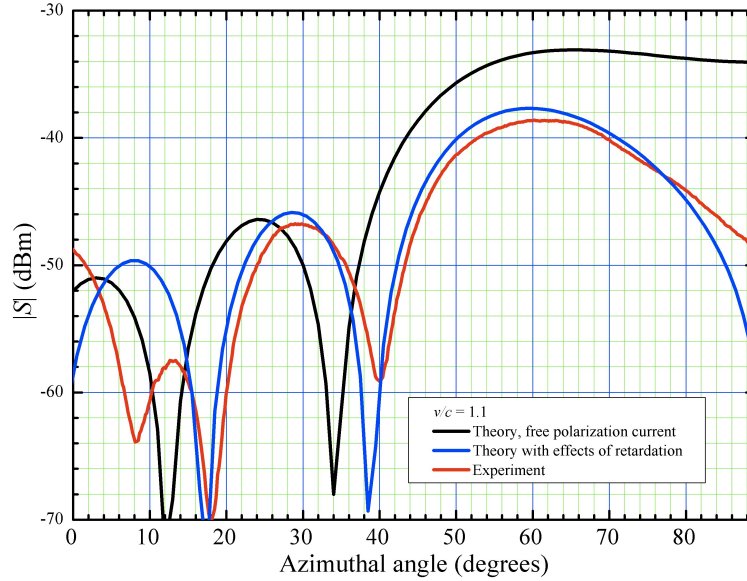


Figure C.4: Linear antenna TD 2 emitted power at 2.4 GHz for $v/c = 1.1$ as a function of azimuthal angle ϕ . The black curve is the prediction of the numerical model for a naked polarization current (no effects of the dielectric included). The blue curve is the prediction of the numerical model including the effects of the dielectric (phase delays and internal reflection); once these corrections are made, the agreement with experimental data from the FARM anechoic chamber (red) is much improved. The remaining disagreement between red and blue curves close to $\phi = 0$ is due to phase-setting errors in TD 2 on the day of the experiment.

Detailed examination of the files recorded by the phase-setting routines in all of our actively controlled experimental antennas reveals a fall off in performance as the the phase difference between the signals sent to adjacent elements increases. The increased S_{11} and the associated crosstalk between the antenna elements leads to signals of inappropriate phase being fed back to the phase monitoring system. In correcting this, the automated phase setting routine attempts to put more power into the vector modulator in question to maintain the desired phase. However, this can lead to imbalances in power between the vector modulators; the automated system then compensates for this by lowering the power to *all* vector modulators

until uniformity is restored. The net result is less power fed to the whole antenna.

Such phenomena illustrates the difficulties associated with any antenna of this scale when run close to light speed; a partial solution is to have many more, closely spaced, narrow electrodes, so that there is much less of a phase difference between adjacent electrodes at low speeds.

C.2.5 Modelling the Effects

All of these effects can be incorporated into the numerical models of Chapter 7; Fig. C.4 shows a typical example for TD 2 running at 2.4 GHz. The black curve is the prediction for a naked polarization current with speed $1.1c$. The blue curve shows the effect of retardation and confinement caused by the presence of the dielectric, leading to a reasonably good agreement with the experimental data (red curve). The residual disagreement close to $\phi = 0$ is primarily due to errors in the experimental phase settings on the day of the measurement (see Section 7.6), which influence subsidiary lobes more severely.

References

- [1] B. M. Bolotovskii and A. V. Serov, *Radiation of superluminal sources in vacuum*, Radiation Physics and Chemistry, 2006, **75**, 813–24.
- [2] Homer, *The Iliad*. With an English Translation by A. T. Murray, Ph.D., Harvard University Press, Cambridge, MA, William Heinemann Ltd., London, 1924.
- [3] L. Casson, *The Ancient Mariners: Seafarers and Sea Fighters of the Mediterranean in Ancient Times*, Macmillan, New York, 1959.
- [4] D. Savitsky, *On the subject of high-speed monohulls*. Presented to the Greek Section of the Society of Naval Architects and Marine Engineers, Athens, Greece, October 2, 2003, oa.upm.es/14340/2/Documentacion/3_Formas/Savitskyreport_conSemidesplazamiento.pdf. Accessed 31 October 2018.
- [5] L. Casson, *Ships and Seamanship in the Ancient World*, Princeton University Press, Princeton, NJ, 1971.
- [6] Homer, *The Odyssey*. Translated by George Herbert Palmer, The Riverside Press, Cambridge, MA, 1892.
- [7] T. von Kármán, *Aerodynamics: Selected Topics in the Light of Their Historical Development*, Dover Publications, Inc., Mineola, NY, 2004, originally published by Cornell University Press, Ithaca, NY, 1954.
- [8] E. W. J. W. Mach, *Über Erscheinungen an fliegenden Projektilen*, Populärwissenschaftliche Vorlesungen, 5. Auflage, Leipzig, 1923.
- [9] W. G. Pohl, *Peter Salcher und Ernst Mach. Schlierenfotografie von Überschall-Projektilen*, Plus Lucis, 2003, **2/2002** and **1/2003**, 22–6.
- [10] W. F. Merzkirch, *Mach's Contribution to the Development of Gas Dynamics*. Ernst Mach: Physicist and Philosopher. Edited by R. S. Cohen and R. J. Seeger, Boston Studies in the Philosophy of Science, D. Reidel Publishing Company, Dordrecht, Holland, 2010, 42–59.

- [11] H. Reichenbach, *Contributions of Ernst Mach to Fluid Dynamics*, Ann. Rev. Fluid Mech., 1983, **15**, 1-29.
- [12] G. Brooker, *Modern Classical Optics*, Oxford Master Series in Physics, Oxford University Press, Oxford, UK, 2003.
- [13] T. Wolfe, *The Right Stuff*, Farar, Straus and Giroux, 1979, reissued by Picador, 2008.
- [14] A. Blackburn, *Mach Match*, Air and Space Magazine, January 01, 1999.
- [15] *Death at Farnborough*, Time Magazine, September 15, 1952.
- [16] B. Mele, M. Ostieri and R. Tognaccini, *Aircraft Lift and Drag Decomposition in Transonic Flows*, Journal of Aircraft, 2017, **54**, 5:1933–44.
- [17] *The Farnborough Tragedy*, Flight Magazine, September 1952.
- [18] T. Gold, *The ‘Double Bang’ of Supersonic Aircraft*, Nature, 1952, **170**, 808.
- [19] C. H. E. Warren, *Noise from Aircraft at Supersonic Speeds*, Nature, 1953, **171**, 214–5.
- [20] P. Rothwell, Nature, 1953, **171**, 215–6.
- [21] G. M. Lilley, R. Westley, A. H. Yates and J. R. Busing, *The Supersonic Bang*, Nature, 1953, **171**, 994–7.
- [22] J. Ackeret, *Akustische Phänomene bei hohen Fluggeschwindigkeiten*, Neue Zürcher Zeitung, September 18, 1952 (German).
- [23] M. Roy, *A propos du gong sonique*, Comptes rendus de l’ Académie des Sciences, Paris, 1952, **235**, 756–9 (French).
- [24] *Sonic Booms and Superbooms*, Nature, 1968, **217**, 399.
- [25] *Concorde: You asked a pilot. Concorde pilot Peter Benn answered your questions*, Transcript of a BBC News interview, October 24, 2003.
- [26] J.-C. L. Wanner, J. Vallee, C. Vivier and C. Thery, *Theoretical and Experimental Studies of the Focus of Sonic Booms*, J. Acoust. Soc. Am., 1972, **52**, 1:13–32 .
- [27] M. Downing, N. Zamot, C. Moss, D. Morin, E. Wolski, S. Chung, K. Plotkin and D. Maglieri, *Controlled focused sonic booms from maneuvering aircraft*, J. Acoust. Soc. Am., 1998, **104**, 1:112–21.
- [28] K. J. Plotkin, *State of the art of sonic boom modeling*, J. Acoust. Soc. Am., 2002, **111**, 1:530–6.

- [29] T. Poston and I. Stewart, *Catastrophe Theory and Its Applications*, Dover Publications, Inc., Mineola, NY, 1996.
This is an unabridged, unaltered republication of the work originally published by Pitman Publishing, Ltd., London, 1978, in the series “Surveys and Reference Works in Mathematics”.
- [30] R. Gilmore, *Catastrophe Theory for Scientists and Engineers*, Dover Publications, Inc., Mineola, NY, 1993.
This Dover edition, first published in 1993, is an unabridged, unaltered republication of the work first published by John Wiley & Sons (“A Wiley-Interscience Publication”), New York, 1981.
- [31] T. Anger and F. Coulouvrat, *Numerical simulation of sonic boom focusing*, AIAA Journal, 2002, **40**, 9:1726–34 .
- [32] R. Marchiano and F. Coulouvrat, *Numerical simulation of shock wave focusing at fold caustics, with application to sonic boom*, J. Acoust. Soc. Am., 2003, **114**, 4:1758–71 .
- [33] Albert Einstein, *Zur Elektrodynamik bewegter Körper*, Annalen der Physik, 1905, **17**, 891–921 (German).
- [34] G. B. Malykin and E. A. Romanets, *Superluminal Motion (Review)*, Optics and Spectroscopy, 2011, **112**, 6:920–34.
- [35] Oliver Heaviside, *Electrical Papers*, First Edition, Macmillan and Co., New York and London, 1894.
- [36] A. Sommerfeld, *Zur Elektronentheorie (3 Teile)*, Nach. Kgl. Ges. Wiss. Göttingen, Math. Naturwiss. Klasse, 1904, 99–130, 363–439, 1905, 201–35 (German).
- [37] G. A. Schott, *Electromagnetic Radiation and the Mechanical Reactions Arising from it*, Cambridge University Press, Cambridge, UK, 1912.
- [38] B. M. Bolotovskii and V. L. Ginzburg, *The Vavilov-Čerenkov Effect and the Doppler Effect in the Motion of Sources with Superluminal Velocity in Vacuum*, Sov. Phys. Usp., 1972, **15**, 2:184–92.
- [39] B. M. Bolotovskii and V. P. Bykov, *Radiation by charges moving faster than light*, Sov. Phys. Usp., 1990, **33**, 6:477–87.
- [40] *Nobel Lectures, Physics 1942-1962*, Elsevier Publishing Company, Amsterdam, 1964.
- [41] Emilio Segrè, *Nuclei and Particles*, Second Edition, W. A. Benjamin, Inc., Reading, MA, 1977.

- [42] A. Ardavan, W. Hayes, J. Singleton, H. Ardavan, J. Fopma and D. Halliday, *Experimental observation of nonspherically-decaying radiation from a rotating superluminal source*, J. Appl. Phys., 2004, **96**, 8:4614–31.
- [43] A. Ardavan, W. Hayes, J. Singleton, H. Ardavan, J. Fopma and D. Halliday, *Corrected Article: “Experimental observation of nonspherically-decaying radiation from a rotating superluminal source” [J. Appl. Phys. 96, 4614, (2004)]*, J. Appl. Phys., 2004, **96**, 12:7760–77.
- [44] J. Singleton, A. Ardavan, H. Ardavan, J. Fopma, D. Halliday and W. Hayes, *Experimental demonstration of emission from a superluminal polarization current – A new class of solid-state source for MHz – THz and beyond*, Conference Digest of the 2004 Joint 29th International Conference on Infrared and Millimeter Waves and 12th International Conference on Terahertz Electronics, IEEE Cat. No. 04EX857, 2004.
- [45] A. C. Schmidt-Zweifel, *Terrestrial and Extraterrestrial Radiation Sources that Move Faster than Light*, Master Thesis, 2013, digitalrepository.unm.edu/math_etds/45/. Accessed 5 July 2018.
- [46] I. Frank and I. Tamm, *Coherent Visible Radiation of Fast Electrons Passing Through Matter*. In I. E. Tamm: Selected Papers. Edited by B. M. Bolotovskii, V. Y. Frenkel and R. Peierls, Springer-Verlag, Berlin · Heidelberg, 1991.
- [47] V. L. Ginzburg, *Certain Theoretical Aspects of Radiation Due to Superluminal Motion in a Medium*, Sov. Phys. Usp., 1960, **2**, 6:874–93.
- [48] V. L. Ginzburg, *Vavilov-Čerenkov effect and anomalous Doppler effect in a medium in which wave phase velocity exceeds velocity of light in vacuum*, Sov. Phys. JETP, 1972, **35** 1:92–3.
- [49] A. V. Bessarab, A. A. Gorbunov, S. P. Martynenko, N. A. Prudkoi, *Faster-than-light EMP source initiated by short X-ray pulse of laser plasma*. IEEE Trans. Plasma Sci., 2004, **32**, 1400–03.
- [50] A. V. Bessarab, S. P. Martynenko, N. A. Prudkoi, A. V. Soldatov, V. A. Terenkhin, *Experimental study of electromagnetic radiation from a faster-than-light vacuum macroscopic source*, Radiation Physics and Chemistry, 2006, **75**, 825–31.
- [51] Yu. N. Lazarev, P. V. Petrov, *A high-gradient accelerator based on a faster-than-light radiation source*, Technical Physics, 2000, **45**, 971–9.
- [52] D. Fleisch, *A Student’s Guide to Maxwell’s Equations*, Seventh Printing, Cambridge University Press, Cambridge, UK, 2009.

A website hosted by the author, and available through www.cambridge.org/9780521877619, contains interactive solutions to every problem in the text.

- [53] J. D. Jackson, *Classical Electrodynamics*, Third Edition, John Wiley & Sons, Inc., New York, 1999.
- [54] C. A. Balanis, *Advanced Engineering Electromagnetics*, Second Edition, John Wiley & Sons, Inc., Hoboken, NY, 2012.
- [55] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York, 1961.
- [56] J. A. Stratton, *Electromagnetic Theory*, Later Printing Edition, McGraw-Hill, New York, 1941.
- [57] B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*. The Oxford Classic Text Edition, Oxford University Press, Oxford, UK, 2013.
- [58] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*. Translated by Morton Hamermesh, Fourth Revised English Edition, Butterworth-Heinemann, Oxford, UK, 1975.
- [59] O. D. Jefimenko, *Electricity and Magnetism: An Introduction to the Theory of Electric and Magnetic Fields*, Second Edition, Electret Scientific, Waynesburg, PA, 1989.
- [60] D. G. Dudley, *Mathematical Foundations for Electromagnetic Theory*. The IEEE Series on Electromagnetic Wave Theory, John Wiley & Sons, Inc. – IEEE Press, New York, 1994.
- [61] C. A. Balanis, *Antenna Theory: Analysis and Design*, Third Edition, John Wiley & Sons, Inc., 2005.
Authorized reprint by Wiley India Pvt., Ltd., for sale on the Indian Sub-continent only. The present author acquired her copy from a private seller on Amazon.com for half the prize of any other edition, new or used.
- [62] L. C. Evans, *Partial Differential Equations*, American Mathematical Society, Providence, RI, 1998.
- [63] International Union of Pure and Applied Physics, *Symbols, Units, Nomenclature and Fundamental Constants in Physics*, 2010 Reprint of the 1987 Revision, Prepared by E. R. Cohen and P. Giacomo, SUNAMCO 87–1.

- [64] J. C. Maxwell, *On Physical Lines of Force (4 parts)*, The London, Edinburgh and Dublin Philosophical Magazine, 1861, **4**, 21:161–75, 281–91, 338–48, 23:12–24, 85–95.
- [65] J. C. Maxwell, *A Dynamical Theory of the Electromagnetic Field*, Phil. Trans. R. Soc. Lond. 1865, **155**, 459–512.
This publication followed a December 8, 1864, presentation by Maxwell to the Royal Society.
- [66] I. Farkas, D. Helbing, T. Vicsek, *Social behaviour: Mexican waves in an excitable medium – The stimulation of this concerted motion among expectant spectators is explained*, Nature, 2002, **419**, 131–2.
- [67] J. Singleton *et al.*, *Eighteen-month report on LDRD 20080085 DR: Construction and use of superluminal emission technology demonstrators with applications in radar, astrophysics, and secure communications*, 2009.
- [68] J. Singleton, A. Schmidt, J. Wigger, *Tests of the Linear Superluminal Antenna done for and in collaboration with CommScope*, Internal Report, 2012.
- [69] A. C. Schmidt-Zweifel, J. Singleton, H. Lu, J. Wigger, C. B. Bailey and F. L. Krawczyk, *Current Experimental and Theoretical Understanding of Practical Antennas Employing Superluminal Polarization Currents: a Report of the CRADA between Los Alamos National Laboratory and CommScope*, Internal Report, 2017.
- [70] H. Ardavan, *Is the light cylinder the site of emission in pulsars?*, Nature, 1981, **289**, 1:44–5.
- [71] H. Ardavan, *Pulsar electrodynamics*, The Astrophysical Journal, 1981, **251**, 674–86.
- [72] H. Ardavan, *Relativistic beaming of the pulsar emission*, Mon. Not. R. Astr. Soc., 1982, **198**, 627–35.
- [73] H. Ardavan, *Spectrum of the coherent radio emission from pulsars*, Mon. Not. R. Astr. Soc., 1982, **199**, 667–82.
- [74] H. Ardavan, *A speed-of-light barrier in classical electrodynamics*, Phys. Rev. D, 1984, **29**, 2:207–15.
- [75] H. Ardavan, *The speed-of-light catastrophe*, Proc. R. Soc. London A., 1989, **424**, 113–41.
- [76] H. Ardavan, *The near-field singularity predicted by the spiral Green’s function in acoustics and electrodynamics*, Proc. R. Soc. London A., 1991, **433**, 451–59.

- [77] H. Ardavan, *The mechanism of radiation in pulsars*, Mon. Not. R. Astr. Soc., 1994, **268**, 361–92.
- [78] H. Ardavan and J. E. Ffowcs Williams, *Violation of the inverse square law by the emissions of supersonically or superluminally moving volume sources*, 1995, arXiv:astro-ph/9506023v1. Accessed 25 June 2018.
- [79] A. Schmidt, H. Ardavan, J. Fasel, J. Singleton, and A. Ardavan, *Occurrence of concurrent ‘orthogonal’ polarization modes in the Liénard-Wiechert field of a rotating superluminal source*. MPE Report 291. Edited by W. Becker and H. H. Huang, Proc. 363rd WE-Heraeus Seminar on Neutron Stars and Pulsars, 2007, 124–127.
- [80] H. Ardavan, A. Ardavan, J. Singleton, and M. Perez, *Mechanism of generation of the emission bands in the dynamic spectrum of the Crab pulsar*, Mon. Not. R. Astr. Soc., 2008, **388**, 873–83.
- [81] H. Ardavan, A. Ardavan, J. Singleton, J. Fasel, W. Junor, J. Middleditch, M. R. Perez, A. Schmidt, P. Sengupta, P. Volegov, *Comparison of multiwavelength observations of 9 broad-band pulsars with the spectrum from an extended current with a superluminally rotating distribution*, 2009, ArXiv:0908.1349v1 [astro-ph].
- [82] A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, R. A. Collins, *Observation of a rapidly pulsating radio source*, Nature, 1968, **217**, 709–13.
- [83] F. Pacini, *Energy emission from a neutron star*, Nature, 1967, **216**, 567–8.
- [84] F. Pacini, *Rotating neutron stars, pulsars and supernova remnants*, Nature, 1968, **219**, 145–6.
- [85] T. Gold, *Rotating neutron stars as the origin of the pulsating radio sources*, Nature, 1968, **218**, 731–2.
- [86] T. Gold, *Rotating neutron stars and the nature of pulsars*, Nature, 1969, **221**, 25–7.
- [87] W. Baade, F. Zwicky, F., *On super-novae*, Proc. Nat. Acad. Sci., 1934, **20**, 254–9.
- [88] D. W. Richards, J. M. Comella, *The period of pulsar NP 0532*, Nature, 1969, **222**, 551–2.
- [89] P. Goldreich, W. H. Julian, *Pulsar electrodynamics*, Ap. J., 1969, **157**, 869–80.
- [90] W. J. Cocke, M. J. Disney, D. J. Taylor, *Discovery of optical signals from pulsar NP 0532*, Nature, 1969, **221**, 525–7.

- [91] G. Fritz, R. C. Henry, J. F. Meekins, T. A. Chubb, H. Friedmann, *X-ray pulsar in the Crab nebula*, Science, 1969, **164**, 709–12.
- [92] H. Bradt, S. Rappaport, W. Mayer, R. E. Nather, B. Warner, M. Mcfarlane, J. Kristian, *X-ray and optical observations of the pulsar NP 0532 in the Crab nebula*, Nature, 1969, **222**, 728–30.
- [93] D. R. Lorimer and M. Kramer, *Handbook of Pulsar Astronomy*, Cambridge University Press, Cambridge, UK, 2005.
- [94] A. G. Lyne, F. Graham-Smith, *Pulsar Astronomy*, Cambridge University Press, Cambridge, UK, 2006.
- [95] A. von Hoensbroech, J. Kijak, and A. Krawczyk, *On the high frequency polarization of pulsar radio emission*, Astron. Astrophys., 1998, **334**, 571–85.
- [96] D. B. Melrose, *Whence the pulses? Pulsars: Problems and Progress*, ASP Conference Series, 1996, **105**, 139–46.
- [97] S. M. Ransom *et al.*, *Three millisecond Pulsars in Fermi LAT Unassociated Bright Sources*, The Astrophysical Journal Letters, 2011, **727**, 1:16–22.
- [98] L. Guillemot *et al.*, *Multiwavelength analysis of four millisecond pulsars*, AIP Conference Proceedings, 2011, **1357**, 241.
- [99] C. Kalapotharakos, I. Contopoulos, and D. Kazanas, *The extended pulsar magnetosphere*, Mon. Not. R. Astr. Soc., 2012, **420**, 2793–8.
- [100] I. Contopoulos and C. Kalapotharakos, *The pulsar synchrotron in 3D: curvature radiation*, Mon. Not. R. Astr. Soc., 2010, **404**, 767–78.
- [101] A. Spitkovsky, *Time-dependent Force-free Pulsar Magnetospheres: Axisymmetric and Oblique Rotators*, The Astrophysical Journal Letters, 2006, **648**, 1:51–54.
- [102] T. Lancaster and S.J. Blundell *Quantum Field Theory for the Gifted Amateur*, Oxford University Press, Oxford, UK, 2014.
- [103] J. H. Hannay, *Bounds on fields from fast rotating sources, and others*, Proc. R. Soc. London A., 1996, **452**, 2351–4.
- [104] J. H. Hannay, *Comment II on “Generation of focused, nonspherically decaying pulses of electromagnetic radiation”*, Phys. Rev. E, 2000, **62**, 2:3008–9.
- [105] J. H. Hannay, *Comment on “Method of handling the divergences in the radiation theory of sources that move faster than their waves”*, J. Math. Phys., 2001, **42**, 3973–4.

- [106] J. H. Hannay, *Spectral and polarization characteristics of the nonspherically decaying radiation generated by polarization currents with superluminally rotating distribution patterns: comment*, J. Opt. Soc. Am. A, 2006, **23**, 6:1530–4.
- [107] J. H. Hannay, *Morphology of the nonspherically decaying radiation generated by a rotating superluminal source: comment*, J. Opt. Soc. Am. A, 2008, **25**, 9:2165–6.
- [108] J. H. Hannay, *Fundamental role of the retarded potential in the electrodynamics of superluminal sources: comment*, J. Opt. Soc. Am. A, 2009, **26**, 10:2107–8.
- [109] A. Hewish, *Problems with the superluminal pulsar model*, Mon. Not. R. Astr. Soc., 1996, **280**, 27–30.
- [110] A. Hewish, *Comment I on “Generation of focused, nonspherically decaying pulses of electromagnetic radiation”*, Phys. Rev. E, 2000, **62**, 2:3007.
- [111] K. T. McDonald, *Radial Dependence of Radiation from a Bounded Source*, Princeton University, 2004, www.hep.princeton.edu/~mcdonald/examples/bounded.pdf. Accessed 14 June 2018.
- [112] K. T. McDonald, *Synchrotron-Čerenkov Radiation*, Science, 2004, **303**, 310.
- [113] H. Ardavan, *Reply to Comments on “Generation of focused, nonspherically decaying pulses of electromagnetic radiation”*, Phys. Rev. E, 2000, **62**, 2:3010–13.
- [114] H. Ardavan, A. Ardavan, J. Singleton, J. Fasel, and A. Schmidt, *Response to “Comment on ‘Methods of handling the divergences in the radiation theory of sources that move faster than their own waves’ ”*, 2008, arXiv:0805.0304v1 [math-ph]. Accessed 14 June 2018.
- [115] H. Ardavan, A. Ardavan, J. Singleton, J. Fasel, and A. Schmidt, *Spectral and polarization characteristics of the nonspherically decaying radiation generated by polarization currents with superluminally rotating distribution patterns: reply to comment*, J. Opt. Soc. Am. A, 2006, **23**, 6:1535–9.
- [116] H. Ardavan, A. Ardavan, J. Singleton, J. Fasel, and A. Schmidt, *Morphology of the nonspherically decaying radiation generated by a rotating superluminal source: reply to comment*, J. Opt. Soc. Am. A, 2008, **25**, 9:2167–9.
- [117] H. Ardavan, A. Ardavan, J. Singleton, J. Fasel, and A. Schmidt, *Fundamental role of the retarded potential in the electrodynamics of superluminal sources: reply to comment*, J. Opt. Soc. Am. A, 2009, **26**, 10:2109–13.

- [118] A. Ardavan, H. Ardavan, and J. Singleton, *Comment on “Radial Dependence of Radiation from a Bounded Source”*, 2009, arXiv:physics/0610085v1 [physics.optics]. Accessed 14 June 2018.
- [119] A. Ardavan, H. Ardavan, and J. Singleton, *Synchrotron-Čerenkov Radiation – Response*, Science, 2004, **303**, 310–1.
- [120] Edwin Cartlidge, *Money spinner or loopy idea?* Science, 2003, **301**, 1463–5.
- [121] M. Durrani, *Revolutionary device polarizes opinions*, Physics World, August 2001.
- [122] *Lawbreakers?* The Economist, August 17, 2000.
- [123] R. Courant, *Dirichlet’s Principle, Conformal Mapping, and Minimal Surfaces*, Interscience Publishers, Inc., New York, 1950.
Reprinted in 2005 by Dover Publications, Inc., Mineola, NY.
- [124] J. Hadamard, *La théorie des équations aux dérivées partielles*, Editions Scientifiques, Pekin, 1964. (French)
- [125] G. Barton, *Elements of Green’s Functions and Propagation*, Second Edition (with corrections), Oxford University Press, New York, 1991.
- [126] J. Hadamard, *Four Lectures on Mathematics. Delivered at Columbia University in 1911*, Columbia University Press, New York, 1915.
Released as Project Gutenberg eBook “for the use of anyone, anywhere at no cost and with almost no restrictions whatsoever” at <http://www.gutenberg.org/ebooks/29788>.
- [127] J. Hadamard, *Lectures on Cauchy’s Problem in Linear Partial Differential Equations*, Yale University Press, New York, 1923. Reprinted in 1952 and 2003 by Dover Publications, Inc., Mineola, NY.
This and the previous bibliographical entry are two of the very few mathematical texts by Hadamard that are available in English. The present writer feels very strongly that, at the very least, [124] would be deserving of a translation.
- [128] D. Hilbert and R. Courant, *Methoden der mathematischen Physik*, 1. und 2. Band, Zweite verbesserte Auflage, Verlag von Julius Springer, Berlin, 1931. (German)
- [129] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*, Volumes I & II, McGraw-Hill, New York, 1953.
- [130] I. Stakgold, *Boundary Value Problems of Mathematical Physics*, Volumes I & II, SIAM, Philadelphia, 2000.

- [131] S. Hassani, *Mathematical Physics: A Modern Introduction to Its Foundations*, Springer-Verlag New York, Inc., New York, 1999.
- [132] T. MyInt-U and L. Debnath, *Linear Partial Differential Equations*, Fourth Edition, Birkhäuser Boston, Inc., Cambridge, MA, 2006.
- [133] L. Debnath, *Nonlinear Partial Differential Equations for Scientists and Engineers*, Birkhäuser Boston, Inc., Cambridge, MA, 2005.
- [134] S. Salsa, *Partial Differential Equations in Action. From Modelling to Theory*, Springer-Verlag Italia, Milano, Italy, 2008.
- [135] E. Zauderer, *Partial Differential Equations of Applied Mathematics*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs, and Tracts Founded by Richard Courant, Second Edition, John Wiley & Sons, Inc., New York, 1989.
- [136] L. E. Payne, *Improperly Posed Problems in Partial Differential Equations*, Regional Conference Series in Applied Mathematics, Society of Industrial and Applied Mathematics, Philadelphia, PA, 1975.
- [137] M. M. Lavrentiev, *Some Improperly Posed Problems of Mathematical Physics*, First English Edition, Springer Tracts in Natural Philosophy, Volume 11, Springer-Verlag Berlin · Heidelberg, 1967.
- [138] J. B. Conway, *A Course in Functional Analysis*, Second Edition, Graduate Texts in Mathematics, Volume 96, Springer Science + Business Media, LLC, New York, 1990.
- [139] M. Schechter, *Principles of Functional Analysis*, Second Edition, Graduate Studies in Mathematics, Volume 36, American Mathematical Society, Providence, RI, 2001.
- [140] I. G. Petrovsky, *Lectures on Partial Differential Equations*. Translated by A. Shenitzer, First English Edition, Interscience Publishers, Inc., New York, 1954.
- [141] A. H. Koblitz, *A Convergence of Lives. Sofia Kovalevskaia: Scientist, Writer, Revolutionary*. First published in hardcover by Birkhäuser Boston, Inc., 1983; second printing (revised), 1988. Reprinted in hardcover by Rutgers University Press, 1993. First published in paperback by Rutgers University Press, New Brunswick, NJ, 1993.
- [142] F. John, *The Dirichlet Problem for a Hyperbolic Equation*, Am. J. Math., 1941, **63**, 1:141–54.

- [143] W. D. Hayes and R. F. Probstein, *Hypersonic Inviscid Flow*, Dover Publications, Inc., Mineola, NY, 2004.
This Dover edition is an unabridged republication of *Hypersonic Flow Theory, Second Edition: Volume I: Inviscid Flows*, published by Academic Press in 1959. (Volume II never materialized.)
- [144] K. A. Ames and B. Straughan, *Non-Standard and Improperly Posed Problems*, Mathematics in Science and Engineering, Volume 194, Academic Press, Inc., San Diego, CA, 1997.
- [145] J. Gleick, *Chaos: Making a New Science*. First published in the United States of America by Viking Penguin, Inc., New York, 1987.
- [146] M. M. Waldrop, *Complexity: The Emerging Science at the Edge of Order and Chaos*, Simon & Schuster, New York, 1992.
- [147] D. Diamond, *James Gleick's Survival Lessons*, Wired Magazine, December 1999. www.wired.com/1999/08/gleick/. Accessed 18 January 2019.
- [148] H. Lewy, *An Example of a Smooth Linear Partial Differential Equation Without Solution*, Ann. Math., 1957, **66**, 1:155–8.
- [149] T. Rother, *Green's Functions in Classical Physics*. Lecture Notes in Physics, Volume 938, Springer International Publishing AG, Cham, Switzerland, 2017.
- [150] J. Ockendon, S. Howison, A. Lacey, and A. Movchan, *Applied Partial Differential Equations*, Second Edition, Oxford University Press, Oxford, UK, 2003.
In the preface to the revised edition, the authors concede that “on many occasions [their] zeal in writing the first edition overstretched [their] accuracy”.
- [151] P. R. Garabedian, *Partial Differential Equations*, John Wiley & Sons, Inc., New York, 1964.
- [152] H. Ockendon and J. Ockendon, *Waves and Compressible Flow*. Texts in Applied Mathematics, Volume 47, Springer-Verlag New York, Inc., New York, 2004.
- [153] A. Tychonoff, *Théorèmes d'unicité pour l'équation de la chaleur*, Mat. Sb., 1935, **42**, 2:199–216, mi.mathnet.ru/msb6410. Accessed 15 April 2019.
- [154] J. L. Davis, *Mathematics of Wave Propagation*, Princeton University Press, Princeton, NJ, 2000.
- [155] *CRC Handbook of Lie Group Analysis of Differential Equations*. Volume 2: Applications in Engineering and Physical Sciences. Edited by N. H. Ibragimov, CRC Press, Boca Raton, 1995.

- [156] J. Hadamard, *The problem of diffusion of waves*, Ann. Math., 1942, **43**, 510–22.
- [157] M. Belger, R. Schimmig and V. Wunsch, *A Survey on Huygens' Principle*, Zeitschrift für Analysis und ihre Anwendungen, 1997, **16**, 1:9–36.
- [158] P. Günther, Huygens' Principle and Hadamard's Conjecture, The Mathematical Intelligencer, 1991, **13**, 2:56–63.
- [159] G. F. Carrier and C. E. Pearson, *Partial Differential Equations. Theory and Technique*, Academic Press, Inc., New York, 1976.
- [160] S. Kovalevskaya, *Zur Theorie der partiellen Differentialgleichungen*, Journal für die reine und angewandte Mathematik, 1875, **80**, 1–32.
The Russian mathematician published her doctoral dissertation as “Sophie von Kowalevsky”, the German spelling of her name.
- [161] P. Prasad, *Partial differential equations*, Wiley Eastern Limited, New Delhi, India, 1984.
Published in the Western Hemisphere by Halsted press, a division of John Wiley & Sons, Inc., New York.
- [162] I. Stakgold and M. Holst, *Green's Functions and Boundary Value Problems*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs, and Tracts Founded by Richard Courant, Third Edition, John Wiley & Sons, Inc., Hoboken, NJ, 2011.
- [163] G. F. Roach, *Green's Functions*, Second Edition, Cambridge University Press, Cambridge, UK, 1982.
- [164] D. G. Duffy, *Green's Functions with Applications*, Second Edition, CRC Press, Boca Raton, 2015.
- [165] G. Green, *An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism*. Originally published as book in Nottingham, 1828. Reprinted in three parts in Journal für die reine und angewandte Mathematik, 1850, **39**, 1:73–89, 1852, **44**, 4:356–74, and 1854, **47**, 3:161–221.
- [166] M. J. Lighthill, *An introduction to Fourier analysis and generalised functions*. Cambridge Monographs on Mechanics, First Edition, Cambridge University Press, Cambridge, UK, 1958.
This slim monograph on generalized functions, penned by “one of the towering figures in twentieth century applied mathematics and fluid mechanics” saw *eight* reprints between 1959 and 1980. Despite Debnath's glowing ode to its author (see Chapter 4 of [167]), the text – dedicated to Paul Dirac, Laurent Schwartz, and George Temple – has since fallen a bit out of fashion. Nevertheless, we

consider it by far the best introduction to the topic for mathematical hobbyists who seek a treatment as elementary and free from complications as possible.

- [167] L. Debnath, *Sir James Lighthill and Modern Fluid Dynamics*, Imperial College Press, London, UK, 2008.
- [168] H. Ardavan, A. Ardavan, J. Singleton, J. Fasel, and A. Schmidt, *Fundamental role of the retarded potential in the electrodynamics of superluminal sources*, J. Opt. Soc. Am. A, 2008, **25**, 3:543–57.
- [169] H. Ardavan, A. Ardavan, J. Singleton, J. Fasel, and A. Schmidt, *Inadequacies in the conventional treatment of the radiation field of moving sources*, J. Math. Phys., 2009, **50**, 10:10.1063/1.3215978.
- [170] H. Ardavan, *The electromagnetic radiation whose decay violates the inverse-square law: detailed mathematical treatment of an experimentally realized example*, J. Plasma Phys., 2019, **85**, 3:905850304.
- [171] D. P. L. Castrigiano and S. A. Hayes, *Catastrophe Theory*, Second Edition, Westview Press, A Member of the Perseus Books Group, Cambridge, MA, 2004. First edition published by Addison-Wesley Publishing Company, 1993.
- [172] R. Thom, *Stabilité Structurelle et Morphogénèse*, Benjamin, New York, 1972.
- [173] H. Ardavan, *Generation of focused, nonspherically decaying pulses of electromagnetic radiation*, Phys. Rev. E, 1998, **58**, 5:6659–84.
- [174] A. M. Liénard, *Champ électrique et magnétique produit par une charge électrique concentrée en un point et animée d'un mouvement quelconque*, G. Carré et C. Naud, Paris, 1898.
- [175] E. Wiechert, *Grundlagen der Elektrodynamik*, Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmal in Göttingen, B. G. Teubner, Leipzig, 1899.
- [176] E. Wiechert, *Elektrodynamische Elementargesetze*, Annalen der Physik, 1901, **309**, 4:667–89.
- [177] P. J. Duke, *Synchrotron Radiation*, Oxford Series on Synchrotron Radiation, Oxford University Press, Oxford, UK, 2000.
- [178] V. F. Perepelitsa, *The Lienard-Wiechert potentials for charged tachyons and several remarks on the tachyon Cherenkov radiation*, 2015, arXiv:1502.06551v2 [hep-th]. Accessed November 29, 2020.
- [179] F. Burk, *Lebesgue Measure and Integration: An Introduction*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs, and Tracts Founded by Richard Courant, John Wiley & Sons, Inc., New York, 1997.

- [180] P. Colwell, *Solving Kepler's Equation over Three Centuries*, Willmann-Bell, Richmond, VA, 1993.
- [181] E. Burniston and C. Siewert, *The use of Riemann problems in solving a class of transcendental equations*, Proceedings of the Cambridge Philosophical Society, 1973, **73**, 111–8.
- [182] N. I. Ioakimidis and K. Papadakis, *A new simple method for the analytical solution of Kepler's equation*, Celestial Mechanics, 1985, **35**, 305–16.
- [183] L. Delves and J. Lyness, *A numerical method for locating the zeros of an analytic function*, Mathematics of Computation, 1967, **21**, 543–77.
- [184] W. Gautschi, *Numerical Analysis: An Introduction*, Birkhäuser, Basel, 1997.
- [185] E. Süli and D. Mayers, *An Introduction to Numerical Analysis*, Cambridge University Press, Cambridge, UK, 2003.
- [186] F. B. Hildebrand, *Introduction to numerical analysis*, Dover Publications, Inc., Mineola, NY, 1956.
- [187] C. M. Bender, S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, New York, 1978.
- [188] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, Cambridge University Press, Cambridge, UK, 1986.
- [189] J. Singleton, A. C. Schmidt, C. B. Bailey, J. Wigger and F. L. Krawczyk, *On the Information Carried by Electromagnetic Radiation Launched from Accelerated Polarization Currents*, Phys. Rev. Applied, submitted, 2020.
- [190] S. Chandrasekhar, *Radiative Transfer*, Dover Publications, Inc., Mineola, NY, 1960.
- [191] <https://heasarc.gsfc.nasa.gov/W3Browse/all/pmpulsar.html>. Accessed October 23, 2020.
- [192] M. M. McKinnon, *The transition between nonorthogonal polarization modes in PSR B2016+28 at 1404 MHz*, ApJ, 2003, **590**, 1026–034.
- [193] The VERITAS Collaboration, *Detection of Pulsed Gamma Rays Above 100 GeV from the Crab Pulsar*, Science, 2011, **334**, 69–72.

- [194] D. J. Griffiths, *Introduction to Electrodynamics*, Pearson New International Edition, Fourth Edition, Pearson India Education Services, Uttar Pardesh, India, 2015.
This edition is manufactured in India and authorized for sale only in India, Bangladesh, Bhutan, Pakistan, Nepal, Sri Lanka and the Maledives. The present writer acquired her copy at a very reasonable price from a private seller in Bangladesh.
- [195] E. A. Essex, *Hertz vector potentials of electromagnetic theory*, AJP, 1977, **45**, 11:1099–101.
- [196] A. P. French, *Special Relativity*, The M.I.T. Introductory Physics Series, Thomas Nelson and Sons, London, 1968.
- [197] M. J. Inman, J. M. Earwood, A. Z. Elsherbeni, and C. E. Smith, *Bayesian Optimization Techniques For Antenna Design*, Progress In Electromagnetics Research, 2004, **49**, 71–86. doi:10.2528/PIER04021302
- [198] B. MacKie-Mason, B. A. Greenwood, and Z. Peng, *Adaptive and parallel surface integral equation solvers for very large-scale electromagnetic modeling and simulation*, Progress in Electromagnetics Research. Invited paper for the Commemorative Collection on the 150-Year Anniversary of Maxwell’s Equations, 2015, **154**, 143–62.
- [199] G. Kiziltas, D. Psychoudakis, J. L. Volakis, and N. Kikuchi, *Topology Design Optimization of Dielectric Substrates for Bandwidth Improvement of a Patch Antenna*, IEEE Trans. Antennas Propag., 2003, **51**, 10:2732–43.
- [200] P. D. Dunning and H. A. Kim, *Introducing the Sequential Linear Programming Level-set Method for Topology Optimization*, Struct. Multidiscip. Optim., 2014, **51**, 3:631–43.
- [201] S. Osher and J. A. Sethian, *Fronts Propagating with Curvature-dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations*, J. Comp. Phys., 1988, **79**, 1:12–49.
- [202] J. A. Sethian and A. Wiegmann, *Structural Boundary Design via Level Set and Immersed Interface Methods*, J. Comp. Phys., 2000, **163**, 2:489–528.
- [203] E. Hassan, E. Wadbro and M. Berggren, *Patch And Ground Plane Design Of Microstrip Antennas By Material Distribution Topology Optimization*, Prog. Electromag. Res. B, 2014, **59**, 89–102.
- [204] E. Hassan, D. Noreland, R. Augustine, E. Wadbro and Martin Berggren, *Topology Optimization of Planar Antennas for Wideband Near-Field Coupling*, IEEE Trans. Antennas Propag., 2015, **63**, 9:4208–13.

- [205] D. B. Davidson, *Computational Electromagnetics for RF and Microwave Engineering*, Second Edition, Cambridge University Press, Cambridge, UK, 2011.