VISUAL DIFFERENTIAL GEOMETRY *and* FORMS

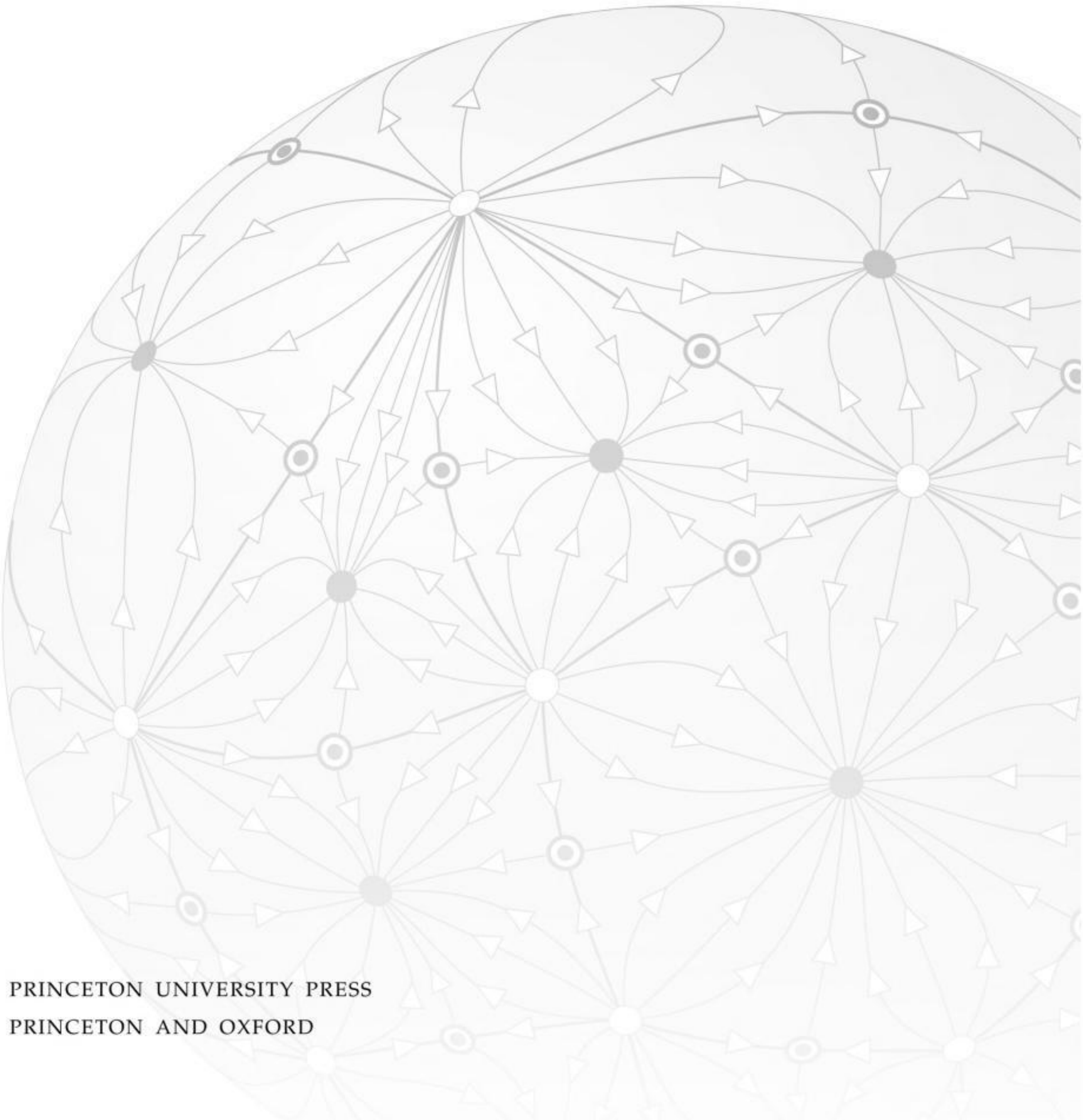
A mathematical drama in five acts

TRISTAN NEEDHAM

VISUAL DIFFERENTIAL GEOMETRY *and* FORMS

A mathematical drama in five acts

TRISTAN NEEDHAM



PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD

Published by Princeton University Press
41 William Street, Princeton, New Jersey 08540
6 Oxford Street, Woodstock, Oxfordshire OX20 1TR

press.princeton.edu

ISBN 9780691203690
ISBN (pbk.) 9780691203706
ISBN (ebook) 9780691219899

Library of Congress Control Number: 2021934723

British Library Cataloging-in-Publication Data is available

Editorial: Susannah Shoemaker, Kristen Hop
Production Editorial: Terri O'Prey
Text Design: Wanda España
Jacket/Cover Design: Wanda España
Production: Brigid Ackerman
Publicity: Matthew Taylor, Amy Stewart
Copyeditor: Gregory W. Zelchenko
Cover image: Stiefel vector field (figure [19.12])

This book has been composed in Palatino (text) and Euler (mathematics)

Printed on acid-free paper. ∞

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Contents

ACT III

Curvature

ACT IV
Parallel Transport

ACT V
Forms

Prologue

The Faustian Offer

> Algebra is the offer made by the devil to the mathematician. The devil says: *"I will give you this powerful machine, and it will answer any question you like. All you need to do is give me your soul: give up geometry and you will have this marvellous machine."* ... the danger to our soul is there, because when you pass over into algebraic calculation, essentially you stop thinking: you stop thinking geometrically, you stop thinking about the meaning.
>
> Sir Michael Atiyah[1]

"*Differential Geometry*" contains the word "*Geometry*."

A tautology? Well, the undergraduate who first opens up the assigned textbook on the subject may care to disagree! In place of geometry, our hapless student is instead confronted with a profusion of *formulas*, and their proofs consist of lengthy and opaque *computations*. Adding insult to injury, these computations are frequently *ugly*, involving a "debauch of indices"[2]—a phrase coined by Élie Cartan (one of the principal heroes of our drama) in 1928. If the student is honest and brave, the professor may be forced to confront an embarrassingly blunt question: "Where has the *geometry* gone?!"

Now, truth be told, most modern texts *do* in fact contain many *pictures*, usually of computer-generated curves and surfaces. But, with few exceptions, these pictures are of specific, concrete examples, which merely *illustrate* theorems whose proofs rest entirely upon symbolic manipulation. In and of themselves, these pictures *explain nothing!*

The present book has *two* distinct and equally ambitious objectives, the first of which is the subject of the first four Acts—to put the "Geometry" back into introductory "Differential Geometry." The 235 hand-drawn diagrams contained in the pages that follow are qualitatively and fundamentally of a different character than mere computer-generated examples. They are the conceptual fruits of many years of intermittent but intense effort—they are the visual embodiment of *intuitive geometric explanations of stunning geometric facts.*

The words I wrote in the Preface to VCA[3] apply equally well now: "A significant proportion of the geometric observations and arguments contained in this book are, to the best of my knowledge, new. I have not drawn attention to this in the text itself as this would have served no useful purpose: students don't need to know, and experts will know without being told. However, in cases where an idea is clearly unusual but I am aware of it having been published by someone else, I have tried to give credit where credit is due." In addition, I have attributed *exercises* that appear to be original, but that are not of my making.

On a personal note (but with a serious mathematical point to follow), the roots of the present endeavour can be traced back decades, to my youth. The story amounts to a tale of two books.

---

[1] *Mathematics in the 20th Century* (Shenitzer & Stillwell, 2002, p. 6)

[2] The full quotation begins to reveal Cartan's heroic stature: "The utility of the absolute differential calculus of Ricci and Levi-Civita must be tempered by an avoidance of excessively formal calculations, where the debauch of indices disguises an often very simple geometric reality. It is this reality that I have sought to reveal." (From the preface to Cartan 1928.)

[3] Given the frequency with which I shall have occasion to refer back to my first book, *Visual Complex Analysis* (Needham 1997), I shall adopt the compact conceit of referring to it simply as VCA.

The *first book* ignited my profound fascination with Differential Geometry and with Einstein's General Theory of Relativity. Perhaps the experience was so intense because it was my *first* love; I was 19 years old. One day, at the end of my first year of studying physics at Merton College, Oxford, I stumbled upon a colossal black book in the bowels of Blackwell's bookshop. Though I did not know it then, the 1,217-page tome was euphemistically referred to by relativity theorists as "The Bible." Perhaps it is appropriate, then, that this remarkable work altered the entire course of my life. Had I not read *Gravitation* (Misner, Thorne, and Wheeler 1973), I would never have had the opportunity[4] to study under (and become lifelong friends with) Roger Penrose, who in turn fundamentally transformed my understanding of mathematics and of physics.

In the summer of 1982, having been intrigued by the mathematical glimpses contained in Westfall's (1980) excellent biography of Newton, I made an intense study of Newton's (1687) masterpiece, *Philosophæ Naturalis Principia Mathematica*, usually referred to simply as the *Principia*. This was the *second book* that fundamentally altered my life. While V. I. Arnol'd[5] and S. Chandrasekhar (1995) sought to lay bare the remarkable nature of Newton's *results* in the *Principia*, the present book instead arose out of a fascination with Newton's *methods*.

As we have discussed elsewhere,[6] Newtonian scholars have painstakingly dismantled the pernicious myth[7] that the results in the 1687 *Principia* were first derived by Newton using his original 1665 version of the calculus, and only later recast into the geometrical form that we find in the finished work.

Instead, it is now understood that by the mid-1670s, having studied Apollonius, Pappus, and Huygens, in particular, the mature Newton became disenchanted with the form in which he had originally discovered the calculus in his youth—which is different again from the Leibnizian form we all learn in college today—and had instead embraced purely geometrical methods.

Thus it came to pass that by the 1680s Newton's algebraic infatuation with power series gave way to a new form of calculus—what he called the "synthetic method of fluxions"[8]—in which the geometry of the Ancients was transmogrified and reanimated by its application to shrinking geometrical figures in their moment of vanishing. *This* is the potent but nonalgorithmic form of calculus that we find in full flower in his great *Principia* of 1687.

Just as I did in VCA, I now wish to take full advantage of Newton's approach throughout this book. Let me therefore immediately spell it out, and in significantly greater detail than I did in VCA, in the vain hope that this second book may inspire more mathematicians and physicists to adopt Newton's intuitive (yet rigorous[9]) methods than did my first.

If two quantities A and B depend on a small quantity ϵ, and their ratio approaches unity as ϵ approaches zero, then we shall avoid the more cumbersome language of limits by following Newton's precedent in the *Principia*, saying simply that, "A is ultimately equal to B." Also, as we did in earlier works (Needham 1993, 2014), we shall employ the symbol $\asymp$ to denote this concept of ultimate equality.[10] In short,

$$\text{“A is ultimately equal to B”} \iff A \asymp B \iff \lim_{\epsilon \to 0} \frac{A}{B} = 1.$$

---

[4]Years later I was privileged to meet with Wheeler several times, and to correspond with him, so I was finally able to thank him directly for the impact that his *Gravitation* had had upon my life.

[5]See Arnol'd and Vasil'ev (1991); Arnol'd (1990).

[6]See Needham (1993), the Preface to VCA, and Needham (2014).

[7]Sadly, this myth originated with Newton himself, in the heat of his bitter priority battle with Leibniz over the discovery of the calculus. See Arnol'd (1990), Bloye and Huggett (2011), de Gandt (1995), Guicciardini (1999), Newton (1687, p. 123), and Westfall (1980).

[8]See Guicciardini (2009, Ch. 9).

[9]Fine print to follow!

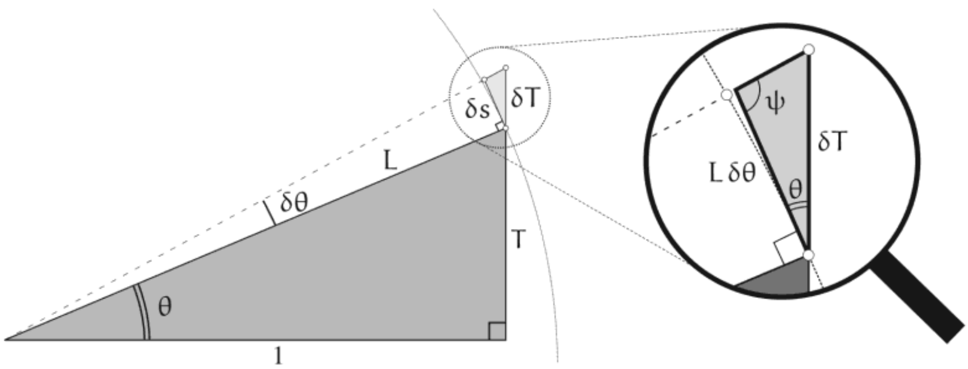
[10]This notation was subsequently adopted by the Nobel physicist, Subrahmanyan Chandrasekhar (see Chandrasekhar 1995, p. 44).

It follows [exercise] from the theorems on limits that ultimate equality is an equivalence relation, and that it also inherits additional properties of ordinary equality, e.g., $X \asymp Y \;\&\; P \asymp Q \Rightarrow X \cdot P \asymp Y \cdot Q$, and $A \asymp B \cdot C \Leftrightarrow (A/B) \asymp C$.

Before we begin to apply this idea in earnest, we also note that the jurisdiction of ultimate equality can be extended naturally to things other than numbers, enabling one to say, for example, that two triangles are "ultimately similar," meaning that their angles are ultimately equal.

Having grasped Newton's method, I immediately tried my own hand at using it to simplify my teaching of introductory calculus, only later realizing how I might apply it to Complex Analysis (in VCA), and now to Differential Geometry. Though I might choose any number of simple, illustrative examples (*see* Needham 1993 for more), I will reuse the specific one I gave in the preface to VCA, and for one simple reason: *this* time I will use the "$\asymp$"–notation to *present* the argument rigorously, whereas in VCA I did not. Indeed, this example may be viewed as a recipe for transforming most of VCA's "explanations" into "proofs,"[11] merely by sprinkling on the requisite $\asymp$ s.

Let us show that if $T = \tan\theta$, then $\frac{dT}{d\theta} = 1 + T^2$. See figure below. If we increase θ by a small (ultimately vanishing) amount $\delta\theta$, then T will increase by the length of the vertical hypotenuse δT of the small triangle, in which the other two sides of this triangle have been constructed to lie in the directions $(\theta + \delta\theta)$ and $(\theta + \frac{\pi}{2})$, as illustrated. To obtain the result, we first observe that in the limit that $\delta\theta$ vanishes, the small triangle with hypotenuse δT is ultimately similar to the large triangle with hypotenuse L, because $\psi \asymp \frac{\pi}{2}$. Next, as we see in the magnifying glass, the side δs adjacent to θ in the small triangle is ultimately equal to the illustrated arc of the circle with radius L, so $\delta s \asymp L\,\delta\theta$. Thus,



$$\frac{dT}{L\,d\theta} \asymp \frac{\delta T}{L\,\delta\theta} \asymp \frac{\delta T}{\delta s} \asymp \frac{L}{1} \quad \Longrightarrow \quad \frac{dT}{d\theta} = L^2 = 1 + T^2.$$

So far as I know, Newton never wrote down this specific example, but compare the illuminating directness of his *style*[12] of geometrical reasoning with the unilluminating computations we teach our students today, more than three centuries later! As Newton himself put it,[13] the geometric method is to be preferred by virtue of the "clarity and brevity of the reasoning involved and because of the simplicity of the conclusions and the illustrations required." Indeed, Newton went even further, resolving that *only* the synthetic method was "worthy of public utterance."

[11] I was already using the $\asymp$ notation (both privately and in print) at the time of writing VCA, and, in hindsight, it was a mistake that I did not employ it in that work; this led some to suppose that the arguments presented in VCA were less rigorous than they actually were (and remain).

[12] The best ambassador for Newton's approach will be you yourself. We therefore suggest that you *immediately* try your *own* hand at Newtonian reasoning, by doing Exercises 1, 2, 3, and 4, on page 24.

[13] See Guicciardini (2009, p. 231)

Newton himself did not employ *any* symbol to represent his concept of "ultimate equality." Instead, his devotion to the geometrical *method* of the Ancients spilled over into emulating their *mode* of expression, causing him to write out the words "ultimately have the ratio of equality," every single time the concept was invoked in a proof. As Newton (1687, p. 124) explained, the *Principia* is "written in words at length in the manner of the Ancients." Even when Newton claimed that two ratios were ultimately equal, he insisted on expressing *each ratio* in words. As a result, I myself was quite unable to follow Newton's reasoning without first transcribing and summarizing each of his paragraphs into "modern" form (which was in fact already quite common in 1687). Indeed, back in 1982, this was the catalyst for my private introduction and use of the symbol, $\asymp$.

It is my view that Newton's choice *not* to introduce a symbol for "ultimate equality" was a tragically consequential error for the development of mathematics. As Leibniz's symbolic calculus swept the world, Newton's more penetrating geometrical method fell by the wayside. In the intervening centuries only a handful of people ever sought to repair this damage and revive Newton's approach, the most notable and distinguished recent champion having been V. I. Arnol'd[14] (1937–2010).

Had Newton shed the trappings of this ancient mode of exposition and instead employed some symbol (*any* symbol!) in place of the words "ultimately equal," his dense, paragraph-length proofs in the *Principia* might have been reduced to a few succinct lines, and his mode of thought might still be widely employed today. Both VCA and this book are attempts to demonstrate, very concretely, the continuing relevance and vitality of Newton's geometrical approach, in areas of mathematics whose discovery lay a century in the future at the time of his death in 1727.

Allow me to insert some fine print concerning my use of the words "rigour" and "proof." Yes, my explicit use of Newtonian ultimate equalities in this work represents a quantum jump in rigour, as compared to my exposition in VCA, but there will be some mathematicians who will object (with justification!) that even this increase in rigour is insufficient, and that *none* of the "proofs" in this work are worthy of that title, including the one just given: I did not actually prove that the side of the triangle is ultimately equal to the arc of the circle.

I can offer no *logical* defence, but will merely repeat the words I wrote in the Preface of VCA, more than two decades ago: "My book will no doubt be flawed in many ways of which I am not yet aware, but there is one 'sin' that I have intentionally committed, and for which I shall not repent: many of the arguments are not rigorous, at least as they stand. This is a serious crime if one believes that our mathematical theories are merely elaborate mental constructs, precariously hoisted aloft. Then rigour becomes the nerve-racking balancing act that prevents the entire structure from crashing down around us. But suppose one believes, as I do, that our mathematical theories are attempting to capture aspects of a robust Platonic world that is not of our making. I would then contend that an initial lack of rigour is a small price to pay if it allows the reader to see into this world more directly and pleasurably than would otherwise be possible." So, to preemptively address my critics, let me therefore concede, from the outset, that when I claim that an assertion is "proved," it may be read as, *"proved beyond a reasonable doubt"!*[15]

Separate and apart from the issue of rigour is the sad fact that in rethinking so much classical mathematics I have almost certainly made mistakes: The blame for all such errors is mine, and mine alone. But please do not blame my geometrical tools for such poor craftsmanship—I am *equally* capable of making mistakes when performing symbolic computations! Corrections will be received with gratitude at VDGF.correction@gmail.com.

The book can be fully understood without giving a second thought to the complete arc of the unfolding drama, told as it is in five Acts. That said, I think that plot matters, and that the book's unorthodox structure and title are fitting, for the following reasons. First, I have sought

[14]See, for example, Arnol'd (1990).

[15]Upon reading these words, a strongly supportive member of the Editorial Board of Princeton University Press suggested to my editor that in place of "Q.E.D.," I conclude each of my proofs with the letters, "P.B.R.D."!

to present the ideas as dramatically as I myself see them, not only in terms of their historical development,[16] but also (more importantly) in terms of the cascading, interconnected flow of the ideas themselves, and their startling implications for the rest of mathematics and for physics. Second, more by instinct than design, the role of each of the five Acts does indeed follow (more or less) the classical structure of a Shakespearean drama; in particular, the anticipated "Climax" is indeed Act III: "Curvature." It was in fact years after I had begun work on the book that one day it suddenly became clear to me that what I had been composing all along had been *a mathematical drama in five acts.* That very day I "corrected" the title of the work, and correspondingly changed its five former "Parts" into "Acts":

- Act I: The Nature of Space
- Act II: The Metric
- Act III: Curvature
- Act IV: Parallel Transport
- Act V: Forms

The first four Acts fulfill the promise of a self-contained, *geometrical* introduction to Differential Geometry. Act IV is the true mathematical powerhouse that finally makes it possible to provide *geometric proofs* of many of the assertions made in the first three Acts.

Several aspects of the *subject matter* are as unorthodox (in a first course) as the geometrical methods by which they are treated. Here we shall describe only the three most important examples.

First, the climax *within* the climax of Act III is the *Global Gauss–Bonnet Theorem*—a remarkable link between local geometry and global topology. While the inclusion of this topic is standard, our treatment of it is not. Indeed, we celebrate its centrality and fundamental importance with an extravagant display of mathematical fireworks: we devote *five* chapters to it, offering up *four* quite distinct proofs, each one shedding new light on the result, and on the nature of Differential Geometry itself.

Second, the transition (usually in graduate school) from 2-dimensional surfaces to n-dimensional spaces (called "manifolds") is often confusing and intimidating for students. Chapter 29—the second longest chapter of the book—seeks to bridge this gap by focusing (initially) on the curvature of **3**-dimensional manifolds, which can be *visualized*; yet we frame the discussion so as to apply to *any* number of dimensions. We use this approach to provide an intuitive, geometrical, yet technically complete, introduction to the famous *Riemann tensor*, which measures the curvature of an n-dimensional manifold.

Third, having committed to a full treatment of the Riemann tensor, we felt it would have been *immoral* to have hidden from the reader its single greatest triumph in the arena of the natural world. We therefore conclude Act IV with a prolonged, *geometrical* introduction to Einstein's glorious *General Theory of Relativity*, which explains gravity as the curvature impressed upon 4-dimensional spacetime by matter and energy. This is the third longest chapter of the book. Not only does it treat (in complete geometrical detail) the famous *Gravitational Field Equation* (which Einstein discovered in 1915) but it also explains some of the most recent and exciting discoveries regarding its implications for black holes, gravitational waves, and cosmology!

Now let us turn to Act V, which is quite different in character from the four Acts that precede it, for it seeks to accomplish a *second* objective of the work, one that is quite distinct from the first, but no less ambitious.

Even the most rabid geometrical zealot must concede that Atiyah's diabolical machine (described in the opening quotation) is a *necessary* evil; but if we *must* calculate, let us at least

[16] As I did in VCA, I *strongly* recommend Stillwell's (2010) masterpiece, *Mathematics and Its History*, as a companion to this book, for it provides deeply insightful and detailed analysis of many historical developments that we can only touch on here.

do so gracefully! Fortunately, starting in 1900, Élie Cartan developed a powerful and elegant new method of *computation*, initially to investigate Lie Groups, but later to provide a new approach to Differential Geometry.

Cartan's discovery is called the "Exterior Calculus," and the objects it studies and differentiates and integrates are called "Differential Forms," here abbreviated simply to *Forms*. We shall ultimately follow Cartan's lead, illustrating his method's power and elegance in the final chapter of Act V—the longest chapter of the book—*reproving symbolically results that were proven geometrically in the first four Acts.* But Forms will carry us *beyond* what was possible in the first four Acts: in particular, they will provide a beautifully efficient method of calculating the Riemann tensor of an n-manifold, via its *curvature 2-forms*.

First, however, we shall fully develop Cartan's ideas in their own right, providing a self-contained introduction to Forms that is *completely independent* of the first four Acts. Lest there be any confusion, we repeat, *the first six chapters*—out of seven—*of Act V make no reference whatsoever to Differential Geometry!* We have done this because Forms find fruitful applications across diverse areas of mathematics, physics, and other disciplines. *Our aim is to make Forms accessible to the widest possible range of readers, even if their primary interest is* not *Differential Geometry.*

To that end, we have sought to treat Forms much more intuitively and *geometrically* than is customary. That said, the reader should be under no illusions: the principal purpose of Act V is to construct, at the *undergraduate* level, the "Devil's machine"—a remarkably powerful method of *computation*.

The immense power of these Forms is reminiscent of the complex numbers: a tiny drop goes in, and an ocean pours out—Cartan's Forms explain vastly more than was asked of them by their discoverer, a sure sign that he had hit upon *Platonic* Forms!

To give just one example, Forms unify and clarify *all* of Vector Calculus, in a way that would be a *revelation* to undergraduates, if only they were permitted to see it. Indeed, Green's Theorem, Gauss's Theorem, and Stokes's Theorem are merely different manifestations of a *single* theorem about Forms that is simpler than any of these special cases! Despite the indisputable importance of Differential Forms across mathematics and physics, most *undergraduates* will leave college without ever having seen them, and I have long considered this a scandal. Only a precious handful[17] of undergraduate textbooks (on either Vector Calculus *or* Differential Geometry) even mention their existence, and they are instead relegated to graduate school.

This lamentable state of affairs is now well into its second century, and I see no signs of an impending sea change. In response, Act V seeks not to curse the dark, but rather to light a candle,[18] striving to convince the reader that Cartan's Forms (and their underlying "tensors") are as *simple* as they are beautiful, and that they (and the name Cartan!) deserve to become a standard part of the *undergraduate* curriculum. *This* is the brazenly ambitious goal of Act V. After drowning the reader in *pure* Geometry for the first four Acts, we hope that the computational aspect of this final Act may serve as a suitably cathartic dénouement!

Before we close, let us simply list some housekeeping details:

- First, I have made no attempt write this book as a classroom textbook. While I hope that some brave souls may nevertheless choose to use it for that purpose—as some previously did with VCA—my primary goal has been to communicate a majestic and powerful subject to the reader as honestly and as lucidly as I am able, regardless of whether that reader is a tender neophyte, or a hardened expert.

[17] See *Further Reading*, at the end of this book.

[18] Ours is certainly not the first such candle to be lit. Indeed, just as our work was nearing completion, Fortney (2018) published an entire book devoted to this same goal. However, Fortney's work does not include any discussion of Differential Geometry, and, at 461 pages, Fortney's book is considerably longer than the 100-page introduction to Forms contained in Act V of this book.

- My selection of topics may seem eclectic at times: for example, why is no attention paid to the fascinating and important topic of minimal surfaces? Frequently, as in this case, it is for one (or both) of the following two reasons: (1) our focus is on intrinsic geometry, *not* extrinsic[19] geometry; (2) an excellent literature already exists on the subject; in such cases, I have tried to provide useful pointers in the *Further Reading* section at the end of the book.
- *Equations* are numbered with (ROUND) brackets, while *figures* are numbered with [SQUARE] brackets.
- Bold italics are used to highlight the ***definition of a new term***.
- For ease of reference when flipping through the book, noteworthy results are $\boxed{\texttt{framed}}$, while doubly remarkable facts are $\boxed{\boxed{\texttt{double} - \texttt{framed}}}$. In the entire work, only a handful of results are *so* fundamental that they are *triple*-framed; we hope the reader will enjoy finding them, like Easter eggs.
- I have tried to make you, the reader, into an active participant in developing the ideas. For example, as an argument progresses, I have frequently and deliberately placed a pair of logical stepping stones sufficiently far apart that you may need to pause and stretch slightly to pass from one to the next. Such places are marked "[exercise]"; they often require nothing more than a simple calculation or a moment of reflection.
- Last, we encourage the reader to take full advantage of the *Index*; its creation was a painful labour of love!

We bring this Prologue to a close with a broader philosophical objective of the work, one that transcends the specific mathematics we shall seek to explain.

One of the rights [*sic*] of passage from mathematical adolescence to adulthood is the ability to distinguish *true miracles* from *false miracles.* Mathematics itself is replete with the former, but examples of the latter also abound: "I can't believe all those ugly terms cancelled and left me such a beautifully simple answer!"; or, "I can't believe that this complicated expression has such a simple meaning!"

Rather than congratulating oneself in such a circumstance, one should instead hang one's head in shame. For if all those ugly terms cancelled, *they should never have been there in the first place!* And if that complicated expression has a wonderfully simple meaning, *it should never have been that complicated in the first place!*

In my own case, I am embarrassed to confess that mathematical puberty lasted well into my 20s, and I only *started* to grow up once I became a graduate student, thanks to the marvellous twin influences of Penrose and of my close friend George Burnett-Stuart, a fellow advisee of Penrose.

The Platonic Forms of mathematical reality are always perfectly beautiful and they are always perfectly simple; transient impressions to the contrary are manifestations of our own imperfection. My hope is that this book may help nudge the reader towards humility in the face of this perfection, just as my two friends first nudged me down this path, so many years ago amidst the surreal, Escher-like spires of Oxford.

T. N.

Mill Valley, California
Newtonmas, 2019

---

[19]The meanings of "intrinsic" and "extrinsic" are explained in Section 1.4.

Acknowledgements

Roger Penrose transformed my understanding of mathematics and of physics. From the very first paper of his I ever read, when I was 20 years old, the *perfection*, beauty, and almost musical counterpoint of his ideas elicited in me a profound aesthetic exhilaration that I can only liken to the experience of listening to the opening of Bach's Cantata 101 or Beethoven's *Grosse Fuge*.

From the time that I was his student, Roger's ability to unravel the deepest of mysteries through *geometry* left an indelible mark—it instilled in me a lifelong, unshakable *faith* that a geometric explanation must always exist. (My study of Newton's *Principia* later served to *deepen* my belief in the universality of the geometrical approach.) Without that faith, this work could not exist, for it sometimes took me many *years* of groping before I discovered the geometric explanation of a particular mathematical phenomenon.

To be able to count myself amongst Roger's friends has been a great joy and a high honour for 40 years, and my dedication of this imperfect work to Roger can scarcely repay the intellectual debt that I owe to him, but it is the best that I can do.

In order to properly introduce the next person to whom I owe thanks, I am forced to reveal a somewhat shabby detail about myself: when I first came to America from England in 1989, I smoked two packs of cigarettes per day! In 1995 I was finally able to quit: it was the hardest thing I had ever done, and I would likely have failed, had it not been for the invention of the nicotine patch.

Perhaps five years later, in response to the 1997 publication of my *Visual Complex Analysis* (VCA), I received a "fan letter" from a *medical* researcher at Duke University; he planned to visit the Bay Area and asked if we could meet. With some trepidation, I agreed. My visitor turned out to be Professor Jed Rose, *inventor* of my (saviour) nicotine patch! Jed had started out in mathematics and physics, and had never lost his love of those disciplines, but shrewdly calculated that he could have a greater impact if he directed his energies to medical research, instead. I'm so glad that he did!

Once I began work on VDGF (*Visual Differential Geometry and Forms*, this book) in 2011, Jed became my most enthusiastic supporter, demonstrating great generosity in using funds from his medical inventions to buy out some of my teaching, thereby greatly assisting my research for VDGF. Every time Jed visited me and my family in California during the nine years of work on the book, my spirits were lifted by Jed's relentlessly upbeat personality and his belief in the importance of what I was trying to accomplish. And, as the manuscript slowly evolved, Jed offered a remarkably large number of detailed and helpful suggestions and corrections; the finished work is significantly better as a result of his helpful observations. Thus, as you see, Jed helped me in three linearly independent directions, and I cannot thank him enough. And as if all this were not enough, what started out as a purely intellectual relationship, subsequently blossomed into a very warm and close friendship between our two families.

The next key person I would like to thank is Professor Thomas Banchoff, the distinguished geometer of Brown University. During the writing of this book, I managed to arrange for Tom to come to USF as a visiting scholar in two separate years, for one semester each. Tom was extremely generous to me during both of those semesters, offering to read my evolving manuscript and giving me extremely valuable feedback. Each week he would read the latest installment of the manuscript, hot off the presses, and then he and I would meet in his office each Friday afternoon, and go over his detailed corrections and suggestions, written in red pen in the margins, line by

line. Although this partnership sadly ended when the book was only half done, I have adopted essentially all of his helpful suggestions and corrections, and I am immensely grateful to him for sharing his deep geometrical wisdom and expertise with me.

I wish to express my sincere gratitude to Dr. Wei Liu[20] (a physicist specializing in optics research, whom I hope to eventually meet, some fine day). In 2019 he wrote to me to express his appreciation of VCA, and he enclosed a research paper of his[21] that cited my treatment of the Poincaré–Hopf Theorem. This paper totally opened my eyes to how physicists continue to make wonderful use of a beautiful result of Hopf that seems to have completely evaporated from all *mathematical* textbooks. The result is the subject of Section 19.6.4, and it says this: The Poincaré–Hopf Theorem not only applies to vector fields, but also to Hopf's *line fields*,[22] which greatly generalize vector fields and which can have singular points with *fractional* indices. Witness the examples shown in [19.14] on page 212. At my request, Dr. Liu then further assisted me by pointing out many other applications that physicists have made of Hopf's ideas. I have in turn shared his kind guidance with you, dear reader, in the *Further Reading* section at the end of the book.

In addition to the principal players above, I have received all manner of advice, support, and suggestions from colleagues and friends, near and far.

My beloved brother Guy is an anchor whose love and faith in me is too often taken for granted, but it should not be!

Stanley Nel and Paul Zeitz, my friends of 30 years, have always believed in me more than I have believed in myself, and their encouragement has meant a great deal to me over the many years it has taken to create this book, first struggling to discover the needed geometrical insights, then writing and *drawing* the book.

Douglas Hofstadter—whose *Gödel, Escher, Bach* transfixed me (and millions more) as an undergraduate—has honoured me with his support for more than 20 years. First, he has repeatedly and forcefully promoted VCA, both in print and in interviews. Second, he read and provided very valuable feedback on Needham (2014), which was ultimately incorporated into this book.

Dr. Ed Catmull—co-founder with Steve Jobs of Pixar, and later president of both Pixar and Walt Disney Animation Studios—wrote me a very flattering email about VCA back in 1999. At first I was convinced that one of my the University of San Francisco (USF) maths-pals was playing a practical joke on me, but the email was real. Ed invited me to visit the Pixar campus (still in Point Richmond back in those days), gave me a guided tour of the studios, took me out to lunch, and offered me a job! (I will leave it to the reader to assign a numerical value to my stupidity in turning that offer down.) Though my contact with Ed has been sporadic, he has been a faithful bolsterer of VCA (praising it in interviews he has done), he also wrote me a letter of recommendation for a grant, and he has been very supportive of my effort to create VDGF. I deeply appreciate Ed's encouragement over the years, and I thank him for the extremely kind words on the back of this book.

Professor Frank Morgan (whom I know by reputation only) was originally approached by Princeton University Press to provide an *anonymous* review of the manuscript of VDGF. But when he submitted his review to my editor, he *also* sent it to me directly, under his own name. I am very grateful that he chose to do this, as it now allows me to thank him publicly for his concrete suggestions and corrections. Furthermore, I was especially grateful for the tremendous boost his report gave to my *morale* at the time. Finally, I offer him my sincerest thanks for being willing to share his remarkably generous assessment of VDGF on the back of this book.

---

[20] College for Advanced Interdisciplinary Studies, National University of Defense Technology, Changsha, Hunan, P. R. China.

[21] Chen et al. (2020b).

[22] This is the modern terminology; Hopf (1956) originally called them *fields of line elements*.

I am likewise also grateful for the constructive criticism, suggestions, and corrections I have received from all of the truly anonymous reviewers—I have tried to incorporate all of their improvements, and I'm sorry I cannot thank each of them by name.

I thank The M. C. Escher Company for permission to reproduce two modifications of *Circle Limit I*: [5.11] and [5.12], the latter being an explicit mathematical transformation, carried out by John Stillwell, and used with his kind permission. Note that M. C. Escher's *Circle Limit I* is © 2020, The M. C. Escher Company, The Netherlands. All rights reserved. www.mcescher.com.

Finally, I am very grateful to Professor Henry Segerman for supplying me with the image of his *Topology Joke* [18.8], page 191, and for granting me permission to reproduce it here.

This is my second book, and it is also my last book. I therefore wish to not only thank all those who directly helped me create VDGF, listed above, but also those who influenced and supported me much earlier in my life. Some of these people were so seamlessly integrated into the fabric of my existence that they became invisible, and, shamefully, I failed to thank them properly in VCA; now is my last chance to put things right.

First amongst these is Anthony Levy, my oldest friend, from our undergraduate days together at Merton College, Oxford. Inexplicably, Anthony (or Tony, as I knew him then) believed in me long before there was any evidence to support such a belief, and that continued belief in me has buoyed me up repeatedly during periods of mathematical self-doubt over the decades. And, beyond the world of pure intellect, Anthony's sage advice and love have helped me navigate some of the most fraught episodes of my life.

Also from those undergraduate Merton days, I will always be grateful to my two physics tutors, Dr. Michael Baker (1930–2017) and Dr. Michael G. Bowler, who not only taught me a great deal of physics themselves, but who also went out of their way to arrange for me to have more advanced, individual tuition on General Relativity and on spinors, from two remarkable Fellows of Merton, Dr. Brian D. Bramson and Dr. Richard S. Ward. In particular, Dr. Bramson's enthusiasm for science was utterly contagious, and it was he who gave me my very first exposure to the (revelatory!) work of Penrose, and who pushed me to apply to undertake a DPhil in Penrose's "Relativity Group."

Moving forward to my graduate student days studying under Penrose, I wish to thank, once again, my friend George Burnett-Stuart, a fellow advisee of Penrose. George and I shared a small house together on Great Clarendon St. for several years as we carried out our doctoral work, and in the course of our endless discussions of music, physics, and mathematics, George helped me to refine both my conception of the nature of mathematics, and of what constitutes an acceptable explanation within that subject. For better or for worse, George bears great responsibility for the mathematician that I am today.

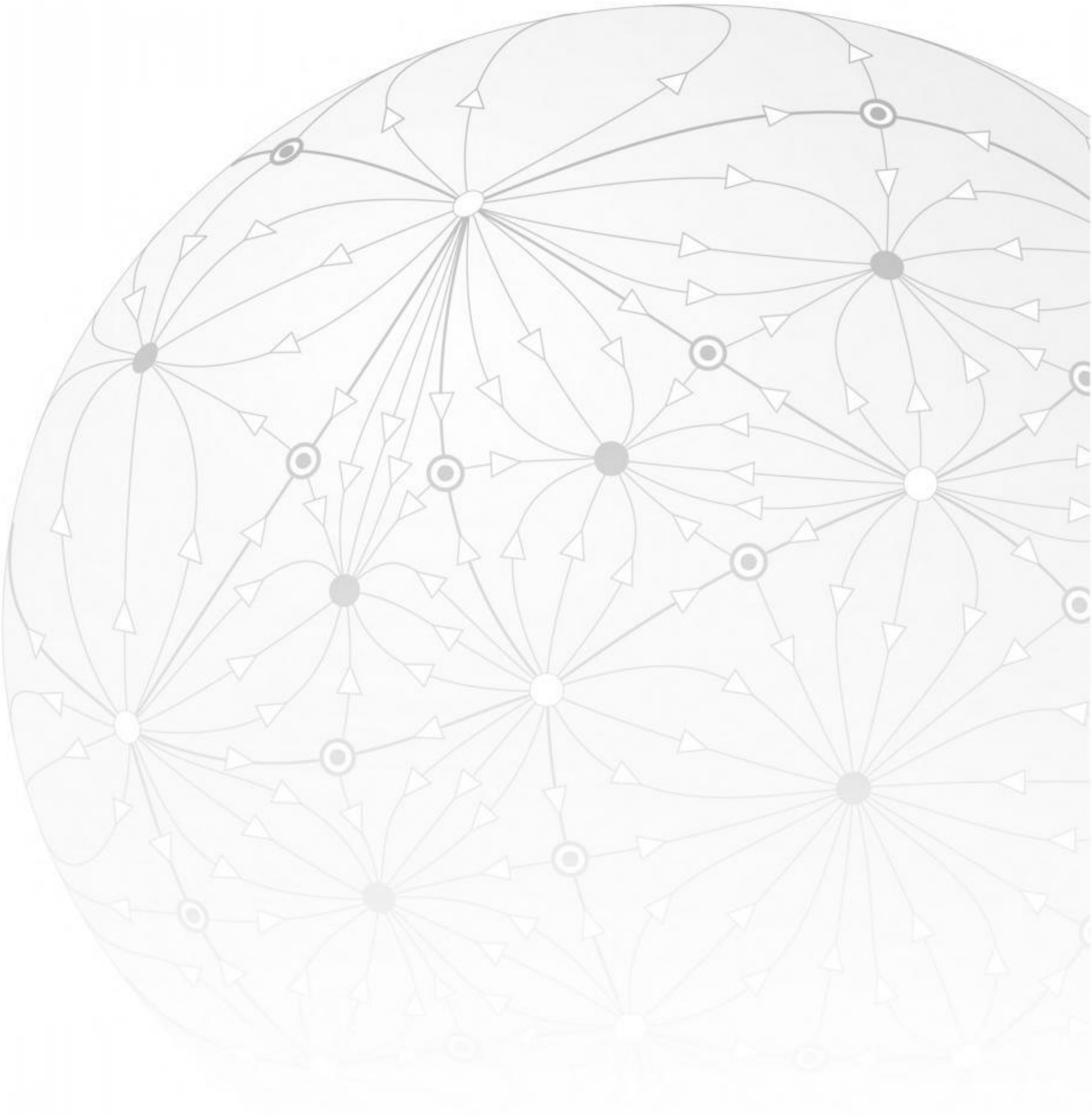
Moving forward again, to my life at USF, I am grateful to John Stillwell—who just retired—and his wonderful wife, Elaine, for 20 happy and productive years as colleagues and friends. While I have picked his brains many times during the course of writing this book, my greatest debt is to his many *writings*. Indeed, our relationship began with a "fan letter" I wrote to him in reaction to the first edition of his magnificent magnum opus, *Mathematics and Its History*. A few years later, while serving as Associate Dean for Sciences, I was able to lure John away from Australia, creating a professorship for him at USF. Both VCA and VDGF owe a great deal to John's holistic grasp of the entire expanse of mathematics, and his ability to use this perspective to give back *meaning* to mathematical ideas, all of which he has so generously shared with the world through his many wonderful books.

In 1996, I concluded the Acknowledgements of VCA with these words: "Lastly, I thank my dearest wife Mary. During the writing of this book she allowed me to pretend that science was the most important thing in life; now that the book is over, she is my daily proof that there is something even more important." Today, more than two decades later, my love for Mary has only grown more profound, but I now have two *more* daily proofs than I had before!

In 1999, Mary and I were blessed with twins: Faith and Hope have been our dazzling beacons of pride and joy ever since. I'm sorry that VDGF has hung like a dark cloud over the life of my family for almost half of my daughters's lives, and that it has robbed us of so much time together. Yet it is the love of these three souls that has given my life meaning and purpose, and has sustained me throughout the long struggle that created this work.

ACT I

The Nature of Space



Chapter 1

Euclidean and Non-Euclidean Geometry

1.1 Euclidean and Hyperbolic Geometry

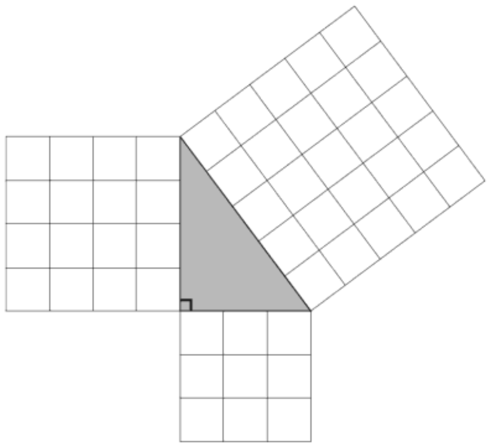
Differential Geometry is the application of calculus to the geometry of space that is *curved*. But to understand space that is curved we shall first try to understand space that is *flat*.

We inhabit a natural world pervaded by curved objects, and if a child asks us the meaning of the word "flat," we are most likely to answer in terms of the *absence* of curvature: a smooth surface *without* any bumps or hollows. Nevertheless, the very earliest mathematicians seem to have been drawn to the singular simplicity and uniformity of the flat plane, and they were rewarded with the discovery of startlingly beautiful facts about geometric figures constructed within it. With the benefit of enormous hindsight, some of these facts can be seen to *characterize* the plane's flatness.

One of the earliest and most profound such facts to be discovered was Pythagoras's Theorem. Surely the ancients must have been awed, as any sensitive person must remain today, that a seemingly unalloyed fact about *numbers*,

$$3^2 + 4^2 = 5^2,$$

in fact has *geometrical* meaning, as seen in [1.1].[1]



[1.1] *Pythagoras's Theorem: the geometry of* $3^2 + 4^2 = 5^2$.

While Pythagoras himself lived in Greece around 500 BCE, the theorem bearing his name was discovered much earlier, in various places around the world. The earliest known example of such knowledge is recorded in the Babylonian clay tablet (catalogued as "Plimpton 322") shown in [1.2], which was unearthed in what is now Iraq, and which dates from about 1800 BCE.

The tablet lists ***Pythagorean triples***:[2] integers (a, b, h) such that h is the hypotenuse of a right triangle with sides a and b, and therefore $a^2 + b^2 = h^2$. Some of these ancient examples are impressively large, and it seems clear that they did not stumble upon them, but rather possessed a mathematical process for generating solutions. For example, the fourth row of the tablet records the fact that $13500^2 + 12709^2 = 18541^2$.

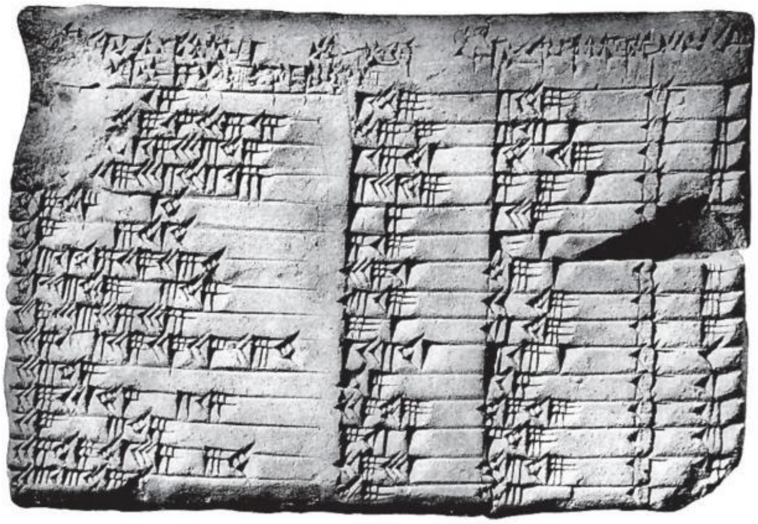
The deeper knowledge that underlay these ancient results is not known,[3] but to find the first evidence of the "modern," logical, deductive approach to mathematics we must jump 1200 years into the future of the clay tablet. Scholars believe that it was Thales of Miletus (around 600 BCE)

---

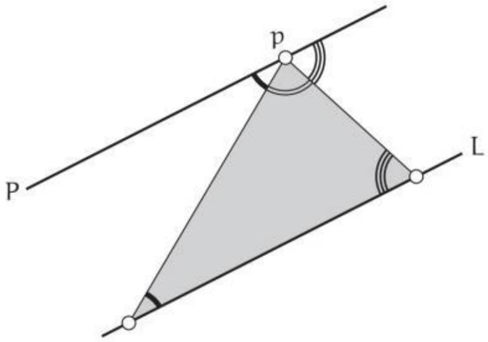
[1] We repeat what was said in the Prologue: *equations* are labelled with parentheses (*round* brackets)—(...), while *figures* are labelled with *square* brackets—[...].

[2] In fact the tablet only records *two* members (a, h) of the triples (a, b, h).

[3] In the seventeenth century, Fermat and Newton reconstructed and generalized a *geometrical* method of generating the general solution, due to Diophantus. See Exercise 5.



[1.2] *Plimpton 322: A clay tablet of Pythagorean triples from 1800 BCE.*



[1.3] *Euclid's Parallel Axiom:* P *is the* unique *parallel to* L *through* p*, and the angle sum of a triangle is* π.

who first pioneered the idea of deducing new results from previously established ones, the logical chain beginning at a handful of clearly articulated assumptions, or ***axioms***.

Leaping forward again, 300 years beyond Thales, we find one of the most perfect exemplars of this new approach in Euclid's *Elements*, dating from 300 BCE. This work sought to bring order, clarity, and rigour to geometry by deducing everything from just five simple axioms, the fifth and last of which dealt with parallel lines.

Defining two lines to be ***parallel*** if they do not meet, Euclid's Fifth Axiom[4] is illustrated in [1.3]:

> ***Parallel Axiom.*** *Through any point* p *not on the line* L *there exists precisely one line* P *that is parallel to* L.

But the character of this axiom was more complex and less immediate than that of the first four, and mathematicians began a long struggle to dispense with it as an assumption, instead seeking to show that it must be a logical *consequence* of the first four axioms.

This tension went unresolved for the next *2000 years*. As the centuries passed, many attempts were made to prove the Parallel Axiom, and the number and intensity of these efforts reached a crescendo in the 1700s, but all met with failure.

Yet along the way useful *equivalents* of the axiom emerged. For example: *There exist similar triangles of different sizes* (Wallis in 1663; see Stillwell (2010)). But the very first equivalent was already present in Euclid, and it is the one still taught to every school child: *the angles in a triangle add up to two right angles.* See [1.3].

The explanation of these failures only emerged around 1830. Completing a journey that had begun 4000 years earlier, Nikolai Lobachevsky and János Bolyai independently announced the

[4]Euclid did not state his axiom in this form, but it is logically equivalent.

discovery of an entirely new form of geometry (now called ***Hyperbolic Geometry***) taking place in a new kind of plane (now called the ***hyperbolic plane***). In this Geometry the first four Euclidean axioms hold, but the parallel axiom does *not*. Instead, the following is true:

> ***Hyperbolic Axiom.*** *There are at least two parallel lines through* p *that do not meet* L. (1.1)

These pioneers explored the logical consequences of this axiom, and by purely abstract reasoning were led to a host of fascinating results within a rich new geometry that was bizarrely different from that of Euclid.

Many others before them, perhaps most notably Saccheri (in 1733; see Stillwell 2010) and Lambert (in 1766; see Stillwell 2010), had discovered some of these consequences of (1.1), but their aim in exploring these consequences had been to find a *contradiction*, which they believed would finally prove that Euclidean Geometry to be the One True Geometry.

[1.4] *Johann Heinrich Lambert (1728–1777).*

Certainly Saccheri believed he had found a clear contradiction when he published "Euclid Freed of Every Flaw." But Lambert is a much more perplexing case, and he is perhaps an unsung hero in this story. His results penetrated so deeply into this new geometry that it seems impossible that he did not at times believe in the reality of what he was doing. Regardless of his motivation and beliefs[5], Lambert (shown in [1.4]) was certainly the first to discover a remarkable fact[6] about the angle sum of a triangle under axiom (1.1), and his result will be central to much that follows in Act II.

Nevertheless, Lobachevsky and Bolyai richly deserve their fame for having been the first to recognize and fully embrace the idea that they had discovered an entirely new, consistent, non-Euclidean Geometry. But what this new geometry really *meant*, and what it might be useful for, even they could not say.[7]

Remarkably and surprisingly, it was the *Differential Geometry of curved surfaces* that ultimately resolved these questions. As we shall explain, in 1868 the Italian mathematician Eugenio Beltrami finally succeeded in giving Hyperbolic Geometry a concrete interpretation, setting it upon a firm and intuitive foundation from which it has since grown and flourished. Sadly, neither Lobachevsky nor Bolyai lived to see this: they died in 1856 and 1860, respectively.

This non-Euclidean Geometry had in fact already manifested itself in various branches of mathematics throughout history, but always in disguise. Henri Poincaré (beginning around 1882) was the first not only to strip it of its camouflage, but also to recognize and exploit its power

---

[5]I thank Roger Penrose for making me see that Lambert deserves greater credit than he is usually granted. Penrose did so by means of the following analogy: "Should we not give credit to Einstein for the cosmological constant, even if he introduced it for the wrong reasons? And should we blame him for later retracting it, calling it the "greatest blunder of my life"? Or what about General Relativity itself, which Einstein seemed to become less and less convinced was the right theory (needing to be replaced by some kind of non-singular unified field theory) as time went on?" [Private communication.]

[6]If you cannot wait, it's (1.8).

[7]Lobachevsky did in fact put this geometry to use to evaluate previously unknown integrals, but (at least in hindsight) this particular application must be viewed as relatively minor.

in such diverse areas as Complex Analysis, Differential Equations, Number Theory, and Topology. Its continued vitality and centrality in the mathematics of the 20th and twenty-first centuries is demonstrated by Thurston's work on 3-manifolds, Wiles's proof of ***Fermat's Last Theorem***, and Perelman's proof of the ***Poincaré Conjecture*** (as a special case of Thurston's ***Geometrization Conjecture***), to name but three examples.

In Act II we shall describe Beltrami's breakthrough, as well as the nature of Hyperbolic Geometry, but for now we wish to explore a different, simpler kind of non-Euclidean Geometry, one that was already known to the Ancients.

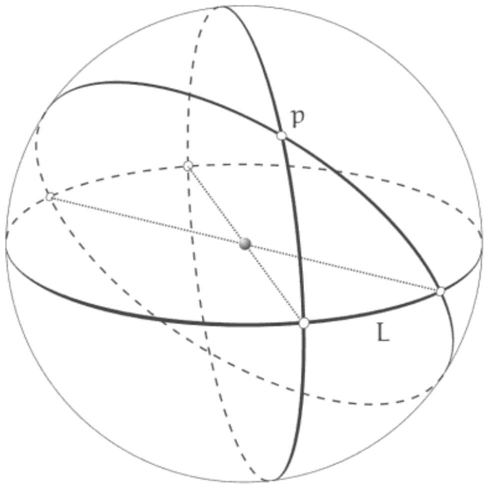
1.2 Spherical Geometry

To construct a non-Euclidean Geometry we must deny the existence of a unique parallel. The Hyperbolic Axiom assumes two or more parallels, but there is one other logical possibility—*no* parallels:

> ***Spherical Axiom.*** *There are* no *lines through* p *that are parallel to* L*:* every *line meets* L. (1.2)

Thus there are actually *two* non-Euclidean[8] geometries: spherical and hyperbolic.

As the name suggests, Spherical Geometry can be realized on the surface of a sphere—denoted S^2 in the case of the *unit* sphere—which we may picture as the surface of the Earth. On this sphere, what should be the analogue of a "straight line" connecting two points on the surface? Answer: the shortest route between them! But if you wish to sail or fly from London to New York, for example, what *is* the shortest route?



[1.5] *The great circles of* S^2 *intersect in pairs of antipodal points.*

The answer, already known to the ancient mariners, is that the shortest route is an arc of a ***great circle***, such as the equator, obtained by cutting the sphere with a plane passing through its centre. In [1.5] we have chosen L to be the equator, and it is clear that (1.2) is satisfied: every line through p meets L in a pair of ***antipodal*** (i.e., diametrically opposite) points.

In the plane, the shortest route is also the *straightest* route, and in fact the same is true on the sphere: in a precise sense to be discussed later, the great circle trajectory bends neither to the right nor to the left as it traverses the spherical surface.

There are other ways of constructing the great circles on the Earth that do not require thinking about planes passing through the completely inaccessible centre of the Earth. For example, on a globe you may map out your great circle journey by holding down one end of a piece of string on London and pulling the string tightly over the surface so that the other end is on New York. The taut string has

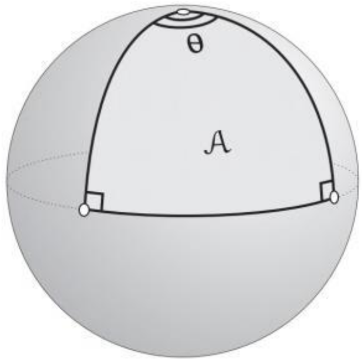
[8] Nevertheless, the reader should be aware that in modern usage "non-Euclidean Geometry" is usually synonymous with "Hyperbolic Geometry."

automatically found the shortest, straightest route—the shorter[9] of the two arcs into which the great circle through the two cities is divided by those cities.

With the analogue of straight lines now found, we can "do geometry" within this spherical surface. For example, given three points on the surface of the Earth, we can connect them together with arcs of great circles to obtain a "triangle." Figure [1.6] illustrates this in the case where one vertex is located at the north pole, and the other two are on the equator.

But if this non-Euclidean Spherical Geometry was already used by ancient mariners to navigate the oceans, and by astronomers to map the spherical night sky, what then was so shocking and new about the non-Euclidean geometry of Lobachevsky and Bolyai?

The answer is that this Spherical Geometry was merely considered to be inherited from the *Euclidean* Geometry of the 3-dimensional space in which the sphere resides. No thought was given in those times to the sphere's internal 2-dimensional geometry as representing an alternative to Euclid's plane. Not only did it violate Euclid's fifth axiom, it also violated a much more basic one (Euclid's first axiom) that we can always draw a unique straight line connecting two points, for this fails when the points are antipodal.



[1.6] *A particularly simple "triangle" on the sphere.*

On the other hand, the Hyperbolic Geometry of Lobachevsky and Bolyai was a much more serious affront to Euclidean Geometry, containing familiar lines of infinite length, yet flaunting multiple parallels, ludicrous angle sums, and many other seemingly nonsensical results. Yet the 21-year-old Bolyai was confident and exuberant in his discoveries, writing to his father, *"From nothing I have created another entirely new world."*

We end with a tale of tragedy. Bolyai's father was a friend of Gauss, and sent him what János had achieved. By this time Gauss had himself made some important discoveries in this area, but had kept them secret. In any case, János had seen further than Gauss. A kind word in public from Gauss, the most famous mathematician in the world, would have assured the young mathematician a bright future. But Nature and nurture sometimes conspire to pour extraordinary mathematical gifts into a vessel marred by very ordinary human flaws, and Gauss's reaction to Bolyai's marvellous discoveries was mean-spirited and self-serving in the extreme.

First, Gauss kept Bolyai in suspense for six months, then he replied as follows:

> Now something about the work of your son. You will probably be shocked for a moment when I begin by saying that *I cannot praise it*, but I cannot do anything else, since to praise it would be to praise myself. The whole content of the paper, the path that your son has taken, and the results to which he has been led, agree almost everywhere with my own meditations, which have occupied me in part for 30–35 years.

Gauss did however "thank" Bolyai's son for having "saved him the trouble"[10] of having to write down theorems he had known for decades.

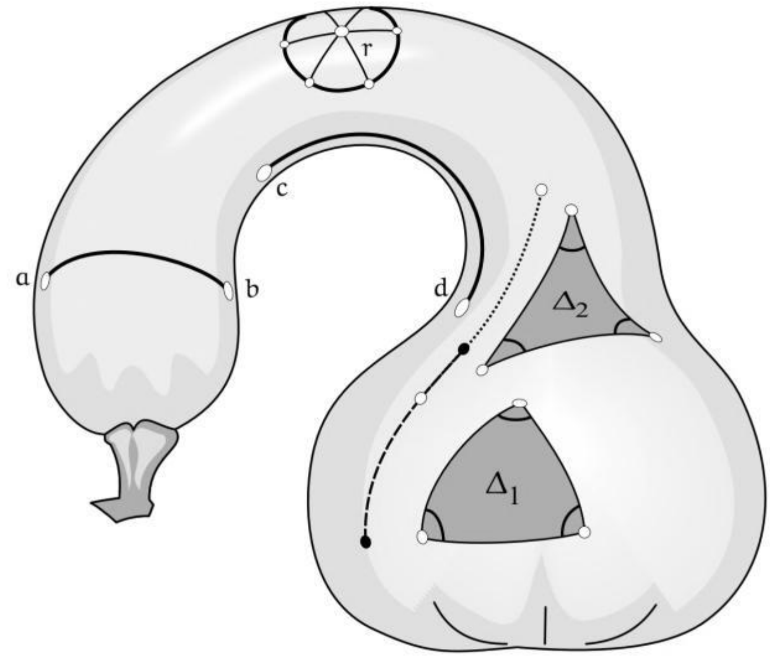
János Bolyai never recovered from the surgical blow delivered by Gauss, and he abandoned mathematics for the rest of his life.[11]

---

[9]If the two points are antipodal, such as the north and south poles, then the two arcs are the *same* length. Furthermore, the great circle itself is no longer unique: *every* meridian is a great circle connecting the poles.

[10]Gauss had previously denigrated Abel's discovery of elliptic functions in precisely the same manner; see Stillwell (2010, p. 236).

[11]If this depresses you, turn your thoughts to the uplifting counterweight of Leonhard Euler. An intellectual volcano erupting with wildly original thoughts (some of which we shall meet later) he was also a kind and generous spirit. We cite one, parallel



[1.9] *The* **intrinsic geometry** *of the surface of a crookneck squash:* **geodesics** *are the equivalents of straight lines, and triangles formed out of them may possess an angular excess of either sign, depending on how the surface bends:* $\mathcal{E}(\Delta_1) > 0$ *and* $\mathcal{E}(\Delta_2) < 0$.

are connected by *multiple* geodesics, and this *nonuniqueness* occurs on more general surfaces, too. What *is* true is that any two points that are *sufficiently close together* can be joined by a unique geodesic segment that is the shortest route between them.

Just as a line segment in the plane can be extended indefinitely in either direction by laying down overlapping segments, so too can a geodesic segment be extended on a curved surface, and this extension is unique. For example, in [1.9] we have extended the dashed geodesic segment connecting the black dots, by laying down the overlapping dotted segment between the white points.

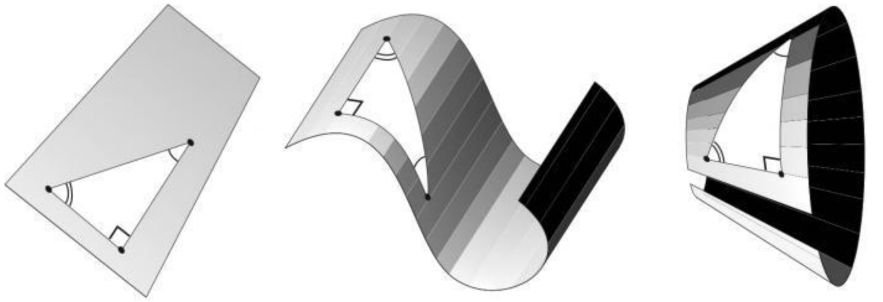
Because of the subtleties associated with the length-minimizing characterization of geodesics, before long we will provide an alternative, purely *local* characterization of geodesics, based on their straightness.

With these caveats in place, it is now clear how we should define distance within a surface such as [1.9]: the distance between two sufficiently close points a and b is the length of the geodesic segment connecting them.

Figure [1.9] shows how we may then define, for example, a "circle of radius r and centre c" as the locus of points at distance r from c. To construct this ***geodesic circle*** we may take a piece of string of length r, hold one end fixed at c, then (keeping the string taut) drag the other end round on the surface. But just as the angles in a triangle no longer sum to π, so now the circumference of a circle no longer is equal to $2\pi r$. In fact you should be able to convince yourself that for the illustrated circle the circumference is *less* than $2\pi r$.

Given three points on the surface, we may join them with geodesics to form a ***geodesic triangle***; [1.9] shows two such triangles, Δ_1 and Δ_2:

- Looking at the angles in Δ_1, it seems clear that they sum to *more* than π, so $\mathcal{E}(\Delta_1) > 0$, like a triangle in Spherical Geometry.



[1.10] *Bending a piece of paper changes the extrinsic geometry, but not the intrinsic geometry.*

- On the other hand, it is equally clear that the angles of Δ_2 sum to *less* than π: $\mathcal{E}(\Delta_2) < 0$, and (as we shall explain) this opposite behaviour is in fact exhibited by triangles in *Hyperbolic Geometry*. Note also that if we construct a circle in this saddle-shaped part of the surface, the circumference is now *greater* than $2\pi r$.

The concept of a geodesic belongs to the so-called ***intrinsic geometry*** of the surface—a fundamentally new view of geometry, introduced by Gauss (1827). It means the geometry that is knowable to tiny, ant-like, intelligent (but 2-dimensional!) creatures living *within* the surface. As we have discussed, these creatures can, for example, define a geodesic "straight line" connecting two nearby points as the shortest route within their world (the surface) connecting the two points. From there they can go on to define triangles, and so on. Defined in this way, it is clear that the intrinsic geometry is unaltered when the surface is bent (as a piece of paper can be) into quite different shapes in space, as long as distances *within* the surface are not stretched or distorted in any way. To the ant-like creatures within the surface, such changes are utterly undetectable.

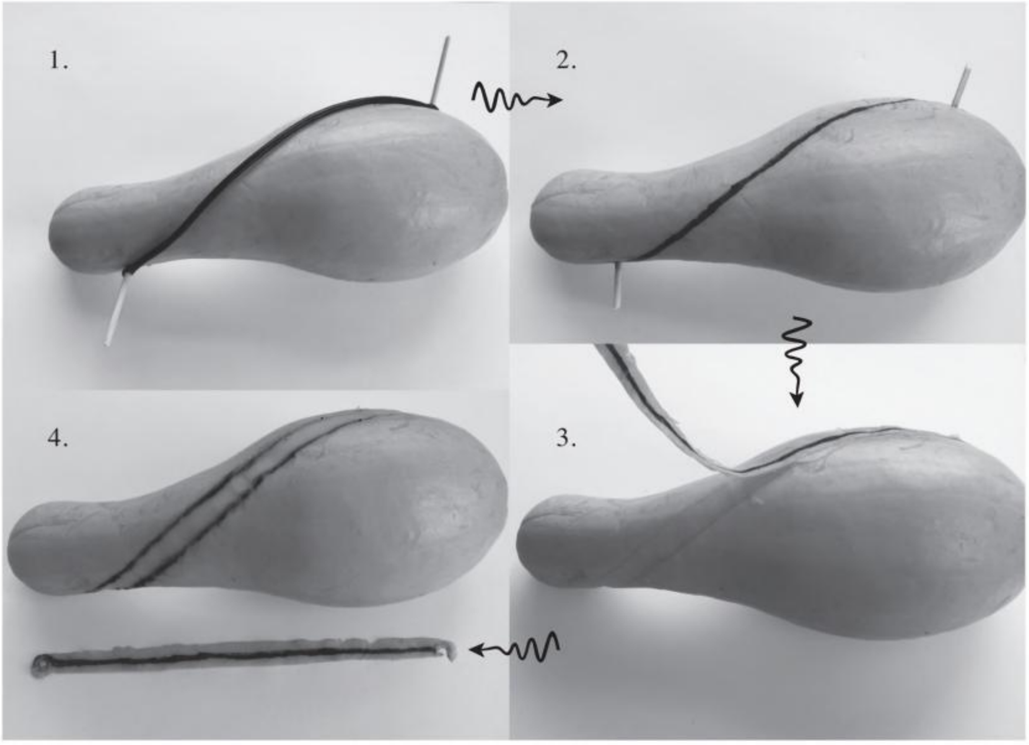
Under such a bending, the so-called ***extrinsic geometry*** (how the surface sits in space) most certainly does change. See [1.10]. On the left is a flat piece of paper on which we have drawn a triangle Δ with angles $(\pi/2)$, $(\pi/6)$, and $(\pi/3)$. Of course $\mathcal{E}(\Delta) = 0$. Clearly we can bend such a flat piece of paper into either of the two (extrinsically) curved surfaces on the right.[14] However, *intrinsically* these surfaces have undergone no change at all—they are both as flat as a pancake! The illustrated triangles on these surfaces (into which Δ is carried by our stretch-free bending of the paper) are identical to the ones that intelligent ants would construct using geodesics, and in both cases $\mathcal{E} = 0$: geometry on these surfaces is Euclidean.

Even if we take a patch of a surface that is intrinsically *curved*, so that a triangle within it has $\mathcal{E} \neq 0$, it too can generally be bent somewhat without stretching or tearing it, thereby altering its extrinsic geometry while leaving its intrinsic geometry unaltered. For example, cut a ping pong ball in half and gently squeeze the rim of one of the hemispheres, distorting that circular rim into an oval (but not an oval lying in a single plane).

1.5 Constructing Geodesics via Their Straightness

We have already alluded to the fact that geodesics on a surface have at least *two* characteristics in common with lines in the plane: (1) they provide the *shortest* route between two points that are not too far apart *and* (2) they provide the "straightest" route between these points. In this section we seek to clarify what we mean by "straightness," leading to a very simple and *practical* method of constructing geodesics on a physical surface.

[14]But note that we must first trim the edges of the rectangle to bend it into the shape on the far right.



[1.11] *On the curved surface of a fruit or vegetable, peel a narrow strip surrounding a geodesic, then lay it flat on the table. You will obtain a* straight line *in the plane!*

Most texts on Differential Geometry pay scant attention to such practical matters, and it is perhaps for this reason that the construction we shall describe is surprisingly little known in the literature.[15] In sharp contrast, in this book we *urge* you to explore the ideas by all means possible: theoretical contemplation, drawing, computer experiments, and (especially!) physical experiments with actual surfaces. Your local fruit and vegetable shop can supply your laboratory with many interesting shapes, such as the yellow summer squash shown in [1.11].

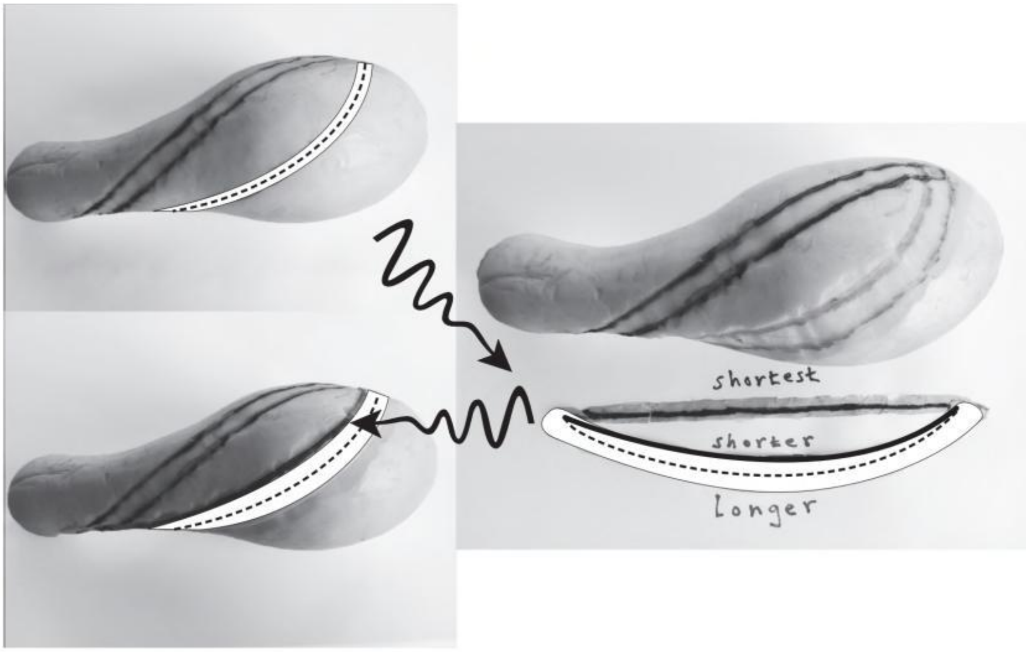
We can now use this vegetable to reveal the hidden straightness of geodesics via an experiment that we hope you will repeat for yourself:

1. On a fruit or vegetable, construct a geodesic by stretching a string over its curved surface.
2. Use a pen to trace the path of the string, then remove the string.
3. Make shallow incisions on either side of (and close to) the inked path, then use a vegetable peeler or small knife to remove the narrow strip of peel between the two cuts.
4. Lay the strip of peel flat on the table, and witness the marvellous fact that the geodesic within the peeled strip has become a *straight line* in the plane!

But why?!

To understand this, first let us be clear that although the strip is free to bend in the direction perpendicular to the surface (i.e., perpendicular to itself), it is *rigid* if we try to bend it sideways, tangent to the surface. Now let us employ proof by contradiction, and imagine what would happen if such a peeled geodesic did *not* yield a straight line when laid flat on the table. It is both a

[15]One of the rare exceptions is Henderson (1998), which we strongly recommend to you; for more details, see the *Further Reading* section at the end of this book.



[1.12] *Suppose that the illustrated dotted path is a geodesic such that a narrow (white) strip surrounding it does* not *become a straight line when laid flat in the plane. But in that case we can shrink the dotted path in the plane (towards the shortest, straight-line route in the plane) thereby producing the solid path. But if we then reattach the strip to the surface, this solid path is still shorter than the original dotted path, which was supposed to be the shortest path within the surface—a contradiction!*

drawback and an advantage of conducting such physical experiments that they will simply not *permit* us to construct something that is impossible, as is required in our desired mathematical proof by contradiction. Nevertheless, let us *suppose* that there exists a geodesic path, such as the dotted one shown on top left of [1.12], that when peeled and laid flat on the table (on the right) does *not* become a straight line.

The shortest route between the ends of this dotted (nonstraight) plane curve is the straight line connecting them. (As illustrated, this is the path of the *true* geodesic we already found using the string—but pretend you don't know that for now!) Thus we may shorten the dotted curve by deforming it slightly towards this straight, shortest route, yielding the solid path along the edge of the peeled strip. Therefore, after reattaching the strip to the surface (bottom left) the solid curve provides a shorter route over the surface than the dotted one, which we had supposed to be the shortest: a contradiction! Thus we have proved our previous assertion:

> *If a narrow strip surrounding a segment* G *of a geodesic is cut out of a surface and laid flat in the plane, then* G *becomes a segment of a* straight line. (1.6)

We are now very close to the promised simple and practical construction of geodesics. Look again at step 3 of [1.11], where we peeled off the strip of surface. But imagine now that we are *reattaching* the strip to the surface, instead. Ignore the history of how we got to this point: what are we actually doing right now in this reattachment process? We have picked up a narrow straight strip (of three-dimensional peel—but mathematically idealized as a two-dimensional strip) and we have unrolled it back onto the surface into the shallow channel from which we cut it. But here is the crucial observation: this shallow channel need not exist—the *surface* decides where the strip must go as we unroll it!

Thus, as a kind of time-reversed converse of (1.6), we obtain a remarkably simple and practical method[16] of constructing geodesics on a physical surface:

> *To construct a geodesic on a surface, emanating from a point* p *in direction* **v**, *stick one end of a length of narrow sticky tape down at* p *and unroll it onto the surface, starting in the direction* **v**. (1.7)

(Note, however, that this does *not* provide a construction of the geodesic connecting p to a specified *target* point q.)

If this construction seems too simple to be true, please try it on any curved surface you have to hand. You can check that the sticky tape[17] is indeed tracing out a geodesic by stretching a string over the surface between two points on the tape: the string will follow the same path as the tape. But note that, as a promised bonus, this new tape construction works on *any* part of a surface, even where the surface is concave towards you, so that the stretched-string construction breaks down.

Of course all of this is a concrete manifestation of a mathematical idealization. A totally flat narrow strip of tape of nonzero width *cannot*[18] be made to fit perfectly on a genuinely curved surface, but its centre line *can* be made to rest on the surface, while the rest of the tape is tangent to the surface.

1.6 The Nature of Space

Let us return to the history of the discovery of non-Euclidean Geometry, and take our first look at how these two new geometries differ from Euclid's.

As we have said, Euclidean Geometry, is characterized by the vanishing of $\mathcal{E}(\Delta)$. Note that, unlike the original formulation of the parallel axiom, *this statement can be checked against experiment*: construct a triangle, measure its angles, and see if they add up to π. Gauss may have been the first person to ever conceive of the possibility that physical space might not be Euclidean, and he even attempted the above experiment, using three mountain tops as the vertices of his triangle, and using light rays for its edges.

Within the accuracy permitted by his equipment, he found $\mathcal{E} = 0$. Quite correctly, Gauss did not conclude that physical space is definitely Euclidean in structure, but rather that if it is *not* Euclidean then its deviation from Euclidean Geometry is extremely small. But he did go so far as to say (see Rosenfeld 1988, p. 215) that he wished that this non-Euclidean Geometry might apply to the real world. In Act IV we shall see that this was a prophetic statement.

---

[16] This important fact is surprisingly hard to find in the literature. After we (re)discovered it, more than 30 years ago, we began searching, and the earliest mention of the underlying idea we could find at that time was in Aleksandrov (1969, p. 99), albeit in a less practical form: he imagined pressing a flexible metal ruler down onto the surface. Later, the basic idea also appeared in Koenderink (1990), Casey (1996), and Henderson (1998). *However*, we have since learned that the essential idea (though not in our current, *practical* form) goes all the way back to Levi-Civita, more than a century ago! See the footnote on page 236.

[17] We recommend using masking tape (aka painter's tape) because it comes in bright colours, and once a strip has been created, it can be detached and reattached repeatedly, with ease. A simple way to create narrow strips (from the usually wide roll of tape) is to stick a length of tape down onto a kitchen cutting board, then use a sharp knife to cut down its length, creating strips as narrow as you please.

[18] This is a consequence of a fundamental theorem we shall meet later, called the *Theorema Egregium*.

Chapter 2

Gaussian Curvature

2.1 Introduction

The proportionality constant,

$$\mathcal{K} = +\frac{1}{R^2},$$

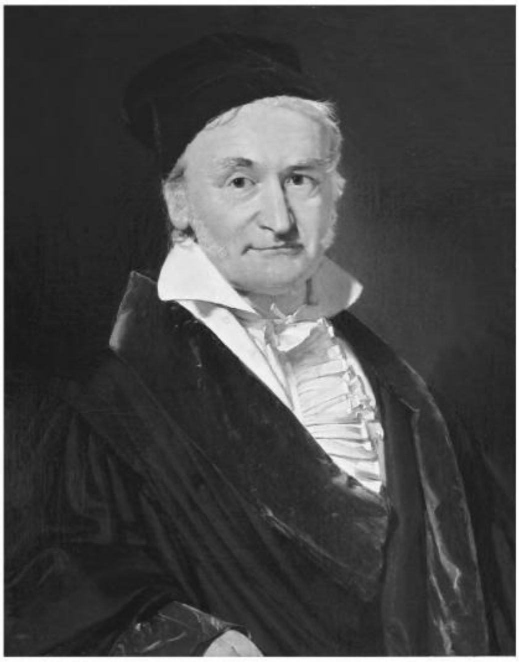
that enters into Spherical Geometry via Harriot's result (1.3), is called the ***Gaussian curvature***[1] of the sphere. The smaller the radius R, the more tightly curved is the surface of the sphere, and the greater the value of the Gaussian curvature $\mathcal{K}$.

Likewise, in Hyperbolic Geometry the negative constant

$$\mathcal{K} = -\frac{1}{R^2}$$

occurring in (1.8) is *again* called the Gaussian curvature, for reasons we shall explain shortly.

This intrinsic[2] concept $\mathcal{K}$ was announced by Gauss (after a decade of private investigation) in his revolutionary "General Investigations of Curved Surfaces,"[3] published in 1827.



[2.1] *Carl Friedrich Gauss (1777–1855).*

As we now explain, Gauss introduced this concept to measure the *curvature at each point* of a general, irregular surface such as that depicted in [1.9]. This one idea of curvature will dominate all that is to come. According to Harriot's and Lambert's results (1.8),

$$\mathcal{K} = \frac{\mathcal{E}(\Delta)}{\mathcal{A}(\Delta)} = \textit{angular excess per unit area.}$$

In both Spherical and Hyperbolic Geometry this interpretation holds for a triangle Δ of any size and any location. But on a more general surface such as in [1.9] this definition makes no sense, for even the *sign* of $\mathcal{E}$ varies between triangles, such as Δ_1 and Δ_2, that reside in different parts of the surface.

---

[1] Also called the ***Gauss curvature***, ***intrinsic curvature***, ***total curvature***, or just plain ***curvature***.

[2] As we shall discuss later, Olinde Rodrigues had arrived at and published the same concept as early as 1815, but from an *extrinsic* point of view. Gauss was not aware that he had been anticipated in this way.

[3] Gauss (1827).

[2.2] *The Gaussian curvature $\mathcal{K}(p)$ at a point is the angular excess per unit area as a geodesic triangle shrinks to that point. In this example, $\mathcal{K}(p) > 0$ and $\mathcal{K}(q) < 0$.*

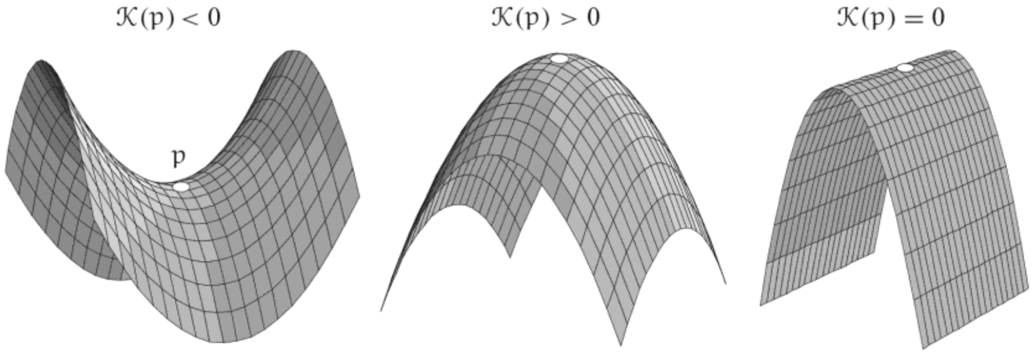
To define the Gaussian curvature *at* a point p on such a surface we now imagine a small geodesic triangle Δ_p containing p, and then allow the triangle to shrink down towards p.

Using the construction of geodesics discovered in the previous section, [2.2] depicts such a sequence of shrinking triangles, converging towards a point on the surface of an inflatable swimming pool ring, the mathematical name for which is a ***torus***. We now define the Gaussian curvature $\mathcal{K}(p)$ at p to be the limit as this triangle shrinks down towards p:

$$\mathcal{K}(p) = \lim_{\Delta_p \to p} \frac{\mathcal{E}(\Delta_p)}{\mathcal{A}(\Delta_p)} = \textit{angular excess per unit area at } p. \tag{2.1}$$

At this stage it is *not* meant to be obvious to you that this limit exists, independently of the shape of the triangle and the precise manner in which it shrinks; this will be proved later. As our drama unfolds we shall discover several other ways[4] of interpreting the Gaussian curvature and of calculating its value in concrete cases.

[4]For a mathematical concept to be truly fundamental it must lie at the intersection of different branches of mathematics. Thus it is to be expected that each of these branches will provide a seemingly distinct yet equally natural way of looking at one and the same concept.



[2.3] *The Gaussian curvature $\mathcal{K}$ is the local angular excess per unit area: its sign is negative if the surface looks like a saddle, positive if it's like a hill, and it vanishes if it's like a curled piece of paper.*

The definition in (2.1) extends beyond triangles. If we replace Δ_p with a small n-gon then (see Ex. 10),

$$\boxed{\mathcal{E}(n\text{-gon}) \equiv [\text{angle sum}] - (n-2)\pi,} \tag{2.2}$$

and the interpretation of curvature in (2.1) as *angular excess per unit area* applies without change.

Inspection of the inflatable pool ring in [2.2] should make it clear that $\mathcal{K}(p) > 0$ at every point p on the outer half, where the immediate neighbourhood of p resembles a hill, whereas $\mathcal{K}(q) < 0$ at every point q on the inner half, where the immediate neighbourhood of q resembles a saddle. Figure [2.3] summarizes this.

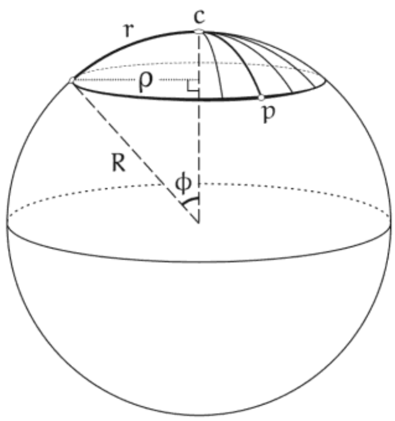
2.2 The Circumference and Area of a Circle

But why is $\mathcal{K}(p)$ so important? Yes, clearly it controls small triangles to some extent, but there is so much more to geometry than just triangles! The answer is that while we may have chosen to define $\mathcal{K}(p)$ (for the moment) in terms of small triangles, we will gradually discover that the curvature has an iron grip over *every* aspect of geometry within the surface. Let us give just two examples for now.

In [1.9] we indicated how a "circle of radius r" centred at c could be defined by taking the end p of a geodesic segment cp of fixed length r and swinging it fully around c. Let us calculate the circumference $C(r)$ of such a circle constructed on the sphere of radius R.

Referring to [2.4], we see that

$$\rho = R\sin\phi \quad \text{and} \quad \phi = \frac{r}{R} \quad \Longrightarrow \quad C(r) = 2\pi R\sin(r/R). \tag{2.3}$$



[2.4] *A circle of radius r on a sphere of radius R has circumference $C(r)$, given by $C(r) = 2\pi R\sin(r/R)$.*

Just as the curvature governs the departure of the angle sum of a triangle from the Euclidean prediction of π, so too does it govern the departure of $C(r)$ from the Euclidean prediction of $2\pi r$. To see this, recall the power series for sine:

$$\sin\phi = \phi - \frac{1}{3!}\phi^3 + \frac{1}{5!}\phi^5 + \cdots.$$

Thus, as ϕ vanishes,

$$\phi - \sin\phi \asymp \frac{1}{6}\phi^3.$$

(We remind the reader that here, $\asymp$ denotes Newton's concept of ***ultimate equality***, as introduced in the Prologue.) It follows from (2.3) that as r shrinks to zero,

$$2\pi r - C(r) = 2\pi R[(r/R) - \sin(r/R)] \asymp \frac{\pi r^3}{3R^2}.$$

In other words, the inhabitants of S^2 can now determine the curvature of their world by examining the circumference of a small circle, just as easily as they previously could by examining the angles of a small triangle:

$$\boxed{\mathcal{K} \asymp \frac{3}{\pi}\left[\frac{2\pi r - C(r)}{r^3}\right].} \tag{2.4}$$

Remarkably, as we will be able to show much later, in Act IV, this formula continues to correctly measure the Gaussian curvature on a *general* surface! (Note that the power of r in the denominator could have been anticipated: we know that $\mathcal{K}$ has dimensions of $1/(\text{length})^2$, and circumference is a length, so we require $(\text{length})^3$ in the denominator.)

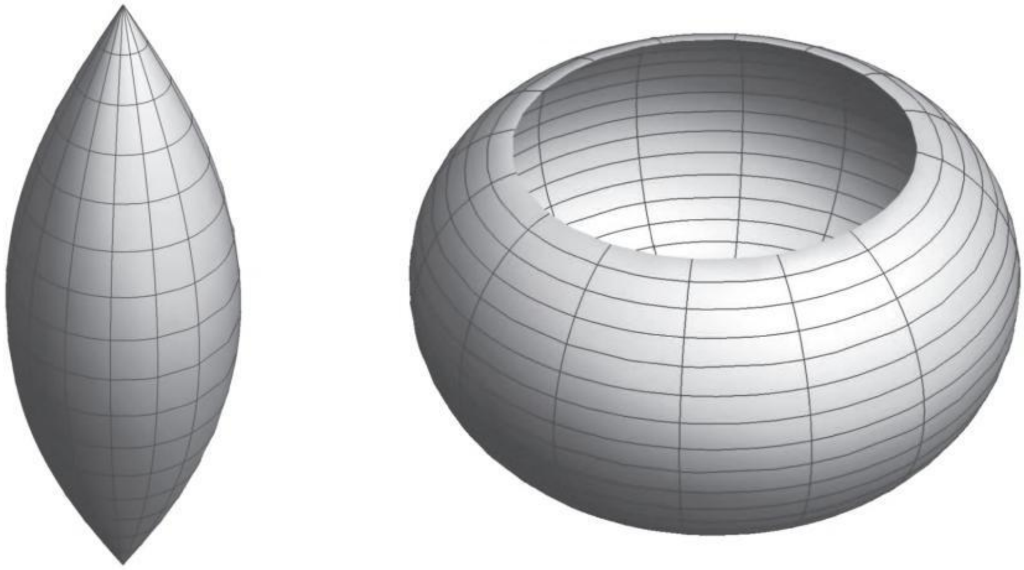
Continuing with this example, let us instead examine the *area* $\mathcal{A}(r)$ of the polar cap bounded by this circle. Again it is the curvature that governs how the area departs from the Euclidean prediction of πr^2. With the assistance of the formula for the polar cap (see Ex. 10, p. 85) it is not hard to verify [exercise] that, in fact,

$$\boxed{\mathcal{K} \asymp \frac{12}{\pi}\left[\frac{\pi r^2 - \mathcal{A}(r)}{r^4}\right].} \tag{2.5}$$

And *again* this formula turns out to be universal! (Again, the same reasoning as above explains the fourth power in the denominator.)

While we are not yet in a position to prove the universality of (2.4) and (2.5), we can at least see that they do indeed yield the correct *sign* at each point of a variably curved surface, such as that shown in [1.9]. For if the immediate vicinity of a point on such a surface is positively curved, then it is hill-shaped near that point (as it everywhere in the region of [1.9] containing Δ_1). Thus both the circumference and area of a small circle centred there will indeed be squeezed by the curvature and be *less* than they would have been in a flat Euclidean plane. Thus, both the preceding formulas yield $\mathcal{K} > 0$, as they should.

On the other hand, if the surface is saddle-shaped near the point, the opposite happens. Recall that we pointed out in [1.9] that a circle drawn in the saddle-shaped part of the surface (where Δ_2 is located) will have $C(r) > 2\pi r$. To grasp this, stand up and hold one arm out at right angles to your body. If you spin around on your heels, the tip of your hand will trace out a horizontal circle. Now repeat this pirouette, but this time wave your arm up and down as you turn; clearly the tip of hand has travelled *further* than before. But this waving up and down is just what happens when we trace out a circle on a saddle-shaped surface, and therefore both of the preceding formulas yield $\mathcal{K} < 0$, as they should.



[2.5] *Nonspherical surfaces of revolution exist that possess constant positive curvature, but these necessarily have either spikes or edges.*

We have said that curvature has an "iron grip" on geometry, but just how absolute is this control? For example, if we know that a patch of surface has constant positive curvature $\mathcal{K} = (1/R^2)$, must it in fact be a portion of a sphere of radius R? Well, take a ping pong ball and cut it in half—now flex one of the hemispheres slightly. Clearly we have obtained a new *non*spherical patch of surface, but since we have not stretched distances within the surface, geodesics and angles are unchanged, and therefore according to the definition (2.3) the curvature $\mathcal{K}$ has not changed. Thus we certainly can obtain at least patches of surface of constant curvature that are not extrinsically spherical, although they all have identical intrinsic geometry.

Figure [2.5] illustrates that even if we restrict attention just to surfaces of revolution, the sphere is not the only one of constant positive curvature. In fact there is an entire family of such surfaces, with the sphere representing the boundary case between the two illustrated types (see Ex. 22). Though they hardly look like spheres, an intelligent ant living on either of these surfaces would never know that she wasn't living on a sphere. Well, that's almost true: eventually she might discover sharp creases or spikes at which the surface is not smooth, or else she might run into an edge where the surface ends. In 1899 Heinrich Liebmann proved[5] that if a surface of constant positive curvature does not suffer from these defects then it can *only* be a sphere.

Ignoring such superficial extrinsic differences, can two surfaces have *constant* positive curvature $\mathcal{K} = (1/R^2)$ and yet have genuinely different *intrinsic* geometries? More explicitly, if our intelligent ant were suddenly transported from one surface to the other, could she devise an experiment to discover that her world had changed? In 1839 Minding (one of Gauss's few students) discovered the answer: "no!" In other words, Minding found[6] that if two surfaces have *constant* positive curvature $\mathcal{K} = (1/R^2)$ then their intrinsic geometries are locally *identical* to that of the sphere of radius R.

We have discussed the fact that the inner rim of the pool ring in [2.2] has negative curvature, but it is not *constant* negative curvature. Indeed, if C is the circle of contact between the ring and the ground, separating the inner and outer halves, then it's clear that the negative curvature $\mathcal{K}(q)$

[5]The proof will have to wait till Section 38.11.

[6]The proof will have to wait till Act IV (Exercise 7, p. 336).

image not available

(i) Treating (2.4) as an approximate equality (instead of an exact ultimate equality), estimate $\mathcal{K}$. From these *intrinsic* measurements, estimate the *extrinsic* radius R of the fruit. Compare your answer with a direct measurement of R.

(ii) Continuing from (i), suppose that your measurements of r and $C(r)$ are perfect. Use the third term of the (decreasing and alternating) Taylor expansion of $\sin(r/R)$ to show that an upper bound on the percentage error in $\mathcal{K}$ that results from applying (2.4) in the manner of part (i)—i.e., *without* taking the limit implied by the ultimate equality—is given by

$$\left|\frac{\Delta\mathcal{K}}{\mathcal{K}}\right| < 5\left[\frac{r}{R}\right]^2 \%$$

Deduce that even for a circle as large as the one you constructed, the error cannot be larger than approximately 3% !

(iii) Use the result of (ii) to deduce a formula for the upper bound of the percentage error in R.

14. ***Negative* Curvature.** Using the technique described in the footnote on page 14, or otherwise, manufacture narrow strips of sticky tape, ideally coloured masking tape. Then use (1.7) to conduct the following experiments.

(i) By following the instructions that accompany [5.3], page 53, construct your own, personal pseudosphere out of discs of radius R—the more cones, the better; the bigger, the better!

(ii) Starting at a point on the circular base of radius R, launch geodesics in various directions, and try to predict their course *before* you lay the tape down onto the surface. When a strip of tape runs out, continue the geodesic by simply *overlapping* a new strip with the old one, as illustrated in [1.9]. With the sole exception of the meridian geodesic that heads straight up the surface—tracing a ***tractrix*** generator of the surface of revolution—note that every geodesic initially heads up the pseudosphere but then turns around and comes back down the pseudosphere, ultimately returning to the base circle.

(iii) Construct a right-angled geodesic triangle, Δ, measure its angles, and hence estimate its angular excess, $\mathcal{E}(\Delta)$. Estimate (as best you can) its areas, $\mathcal{A}(\Delta)$. Hence estimate the (constant) curvature $\mathcal{K}$ of your pseudosphere, using

$$\mathcal{K} = \frac{\mathcal{E}(\Delta)}{\mathcal{A}(\Delta)}.$$

(iv) The larger the triangle, the larger (i.e., more negative) the value of $\mathcal{E}(\Delta)$, making its measurement easier and more accurate. But the tradeoff is that it becomes harder to accurately estimate the area $\mathcal{A}(\Delta)$. To overcome this difficulty, do the following. Make narrow strips of the same kind that you use to generate geodesics, but create them all with the *same* (accurately measured) width, W, perhaps 1/4 inch. Now fill your Δ with these strips, cutting them off when they hit an edge. Remove the strips and lay them end-to-end on a flat surface, and measure the total length, L. Then $\mathcal{A}(\Delta) \asymp LW$.

(v) Repeat (iii) with several more triangles, but no longer restrict them to be right-angled, because (iv) now makes it easy to measure $\mathcal{A}(\Delta)$ for any shape of triangle. Verify that (within experimental error) all triangles yield the *same* value of $\mathcal{K}$.

(vi) Assuming that

$$\mathcal{K} = -\frac{1}{R^2},$$

estimate R, and compare this to the actual radius of the discs you used to construct your pseudosphere.

Further Reading

NOTES: In the following, I merely list title and author; full details can be found in the bibliography. Some of the following works are included simply because they are highly relevant, and I therefore believe they deserve to be brought to your attention (even if they are not quite my cup of tea). Most, however, are included because I consider them to be gems, and I strongly recommend them to you. Many other excellent books sit on my bookshelves, and I seek their counsel often, and yet they are not included here simply in order to cut down this (already long) list to a manageable size—I apologize to the authors of all those excellent works for failing to highlight them here. Finally, I also apologize to the authors of the wonderful works I have yet to discover!

Global Recommendations

First, let me set the stage with six works that I hold to be *invaluable*, the content of each of which spans multiple Acts of this book.

- *The Road to Reality*, by Roger Penrose.

  An extraordinary panorama of almost all of physics, and most of mathematics, by a master of both. Many of the insights can only be found here, and they are brought to life by Penrose's remarkable (and beautiful!) hand drawings.

- *Gravitation*, by Misner, Thorne, and Wheeler.

  Almost 50 years after its original publication in 1973, this classic remains one of the very best introductions to Einstein's geometrical theory of gravity (General Relativity) and to the Differential Geometry upon which it rests. It also contains one of the best, most geometrical introductions to Forms, including the curvature 2-forms that allow one to calculate the Riemann tensor efficiently. The new (2017) edition from Princeton University Press is beautifully done, and contains a new introduction by Charles Misner and Kip Thorne, discussing the exciting developments in the field since the book's original publication.

- *Differential Geometry in the Large*, by Heinz Hopf.

  Hopf was not only one of the towering figures of twentieth-century mathematics, he was also a master of exposition. Here, ideas of Differential Geometry and Topology (many of which are due to Hopf himself) come together in a beautiful way, explained with remarkable clarity and simplicity. Every time I return to this *Meisterwerk*, I feel that some beneficent magician has inserted more wonderful ideas into its pages, for I swear that *this* beautiful idea wasn't on the page the last time I looked!

- *Elementary Differential Geometry* (revised 2nd edition), by Barrett O'Neill.

  First published in 1966, this trail-blazing text pioneered the use of Forms at the undergraduate level. Today, more than a half-century later, O'Neill's work remains, in my view, the single most clear-eyed, elegant, and (ironically) *modern* treatment of the subject available—present company excepted!—at the undergraduate level.

- *Geometrical Methods of Mathematical Physics*, by Bernard Schutz.

  This work—now 40 years old!—is a timeless treasure trove, covering manifolds, tensors, Lie derivatives and Lie groups, Forms, Riemannian Geometry, gauge theories, and a host of other applications to physics. To achieve this, Professor Schutz channels *Star Trek*'s Mr. Spock. His *Vulcan* half enables him to erect a logical structure of crystalline perfection, in which everything is concisely and rigorously proven, and—unlike my (I hope delicious) cheeseburger approach in this book—Kashrut is strictly observed: concepts that depend on the existence of a metric are scrupulously and explicitly separated from those that do not. But, in tandem with this, Schutz is able to harness his *human* half to provide a wealth of intuition that reveals the underlying *geometric* reality.

- *Mathematics and Its History* (3rd edition), by John Stillwell.

  A remarkable panorama of all of mathematics through the lens of history. But make no mistake, this is not primarily a book about history, rather it is fundamentally a work about the interconnectedness and meaning of mathematics itself, all explained in a rather concise style (relative to mine!), with deep insight and lucidity.

Geometry in General

The following works are concerned with geometry in general, but especially with Hyperbolic Geometry. (*Differential* Geometry has its own category.)

- *Geometry and the Imagination*, by David Hilbert and S. Cohn-Vossen.

  A magnificent, deeply insightful survey of geometry, focusing on intuitive understanding, by one of the greatest mathematicians of the 20th century. The diagrams (drawn by K. H. Naumann and H. Bödeker) are *astonishingly* beautiful, to the point of causing me envy!

- *Experiencing Geometry*, by David W. Henderson and Daina Taimina.

  A highly unusual approach, philosophically akin to mine (but using the Moore method), focused on intuitive, *experimental* investigations of geometry. It contains significant discussion of parallel transport and holonomy. The overlap of their approach with mine is made clear by this quotation from the preface: "This book is based on a view of proof as a *convincing communication that answers—Why?*" (Their italics.)

- *Introduction to Geometry* (2nd edition), by H.S.M. Coxeter.

  A wonderfully clear survey by a modern master.

- *Geometry*, by Brannan, Esplen, and Gray.

  An excellent modern survey of geometry, based on Klein's vision of groups of transformations.

- *Euclidean and Non-Euclidean Geometries: Development and History*, by Marvin J. Greenberg.

  A valuable, detailed history of the development of Hyperbolic Geometry, including lengthy quotations from critical, private letters of Gauss, Bolyai, and many others.

- *The Poincaré Half-Plane*, by Saul Stahl.

  The title says it all.

- *Geometry Revealed*, by Marcel Berger.
  A much more advanced survey of geometry, with a focus on conceptual proofs and unsolved problems, by one of the great geometers of the 20th century.

Topology

- *Intuitive Topology*, by V. V. Prasolov.
  Less than 100 pages long, and filled with diagrams, this super-friendly introduction lives up to its title!

- *Euler's Gem*, by David S. Richeson.
  A masterful, mathematically accurate, yet riveting account of Euler's polyhedral formula, its history and the connected mathematical ideas.

- *Surface Topology*, by P. A. Firby and C. F. Gardiner.
  A very gentle, nicely illustrated introduction to the fundamental *geometric* ideas of topology.

- *First Concepts of Topology*, by W. G. Chinn and N. E. Steenrod.
  Another very gentle, nicely illustrated introduction to the fundamental *geometric* ideas of topology.

- *Topology: A Very Short Introduction*, by Richard Earl
  This remarkable little book lives up to its title, covering a huge range of fundamental ideas in just 140 pages, and it does so in a very clear, elementary, informal style. This is my new favourite introduction to the subject.

- *The Shape of Space* (3rd edition), by Jeffrey R. Weeks.
  A wonderfully lucid, engaging, elementary treatment of the topology of two and three dimensional spaces. The last of the four parts of the book deals with the possibility of detecting the topology of the Universe! An appendix contains John Horton Conway's famous ZIP Proof of the Classification Theorem for surfaces, beautifully illustrated by George K. Francis.

- *Three-Dimensional Geometry and Topology*, by William P. Thurston.
  Thurston won the Fields Medal for discovering that 3-manifolds are fundamentally built out of Hyperbolic Geometry. In this book you will hear Thurston's discoveries in his own distinctive voice, and although the difficulty of the *topology* accelerates rapidly, the first 100 pages provide a relatively elementary, highly original introduction to Hyperbolic Geometry that should not be missed.

Hopf's Line Fields and the Poincaré–Hopf Theorem in Physics

In this book I have sought to draw attention to *line fields* and Hopf's beautiful result that the Poincaré–Hopf Theorem applies to them, too, ((19.9), p. 213). These ideas have all but disappeared from modern *mathematics* textbooks, and I strongly believe that it is past time for a revival. That said, *physicists* never lost sight of the value of these ideas, and they have sustained them with wonderful new discoveries.

Index