Experimental Mathematics In Action

David H. Bailey, Jonathan M. Borwein,
Neil J. Calkin, Roland Girgensohn,
D. Russell Luke and Victor H. Moll

Table of Contents

Preface

These are a set of notes to be distributed at a short course, sponsored by the Mathematical Association of America (MAA) at the annual meeting in San Antonio, January 2006. The final version will be published by A. K. Peters in August 2006.

Two of the authors have already established a web site containing an updated collection of links to many of the URLs mentioned in the two volumes, plus errata, software, tools, and other web useful information on experimental mathematics. This can be found at:

http://www.experimentalmath.info

The authors would like to thank

• The Mathematical Association of America (MAA) for the opportunity to develop this course,

• A. K. Peters for enthusiastically agreeing to publish the corresponding book,

• all of our colleagues who have shared their experimental mathematics experiences with us.

David H. Bailey dhbailey@lbl.gov
Jonathan M. Borwein jborwein@mail.cs.dal.ca
Neil Calkin calkin@ces.clemson.edu
Roland Girgensohn girgen@cecm.sfu.ca
Russell Luke rluke@math.udel.edu
Victor H. Moll vhm@math.tulane.edu

December 2005

Chapter 1

A Philosophical Introduction

Christopher Koch, [150], accurately captures a great scientific distaste for philosophizing:

> *Whether we scientists are inspired, bored, or infuriated by philosophy, all our theorizing and experimentation depends on particular philosophical background assumptions. This hidden influence is an acute embarrassment to many researchers, and it is therefore not often acknowledged.*

That acknowledged, I am of the opinion that mathematical philosophy matters more now than it has in nearly a century. The power of modern computers matched with that of modern mathematical software and the sophistication of current mathematics is changing the way we do mathematics.

In my view it is now both necessary and possible to admit quasi-empirical inductive methods fully into mathematical argument. In doing so carefully we will enrich mathematics and yet preserve the mathematical literature's deserved reputation for reliability—even as the methods and criteria change.

1 Mathematical Knowledge as I View It

Somewhat unusually, I can exactly place the day at registration that I became a mathematician and I recall the reason why. I was about to deposit my punch cards in the 'honours history bin'. I remember thinking

> *If I do study history, in ten years I shall have forgotten how to use the calculus properly. If I take mathematics, I shall still be able to read competently about the War of 1812 or the Papal schism.* (Jonathan Borwein, 1968)

The inescapable reality of objective mathematical knowledge is still with me. Nonetheless, my view then of the edifice I was entering is not that close to my view of the one I inhabit nearly forty years later.

I also know when I became a computer-assisted fallibilist. Reading Imre Lakatos' *Proofs and Refutations*, [157], a few years later while a very new faculty member, I was suddenly absolved from the grave sin of error, as I began to understand that missteps, mistakes and errors are the grist of all creative work. The book, his doctorate posthumously published in 1976, is a student conversation about the Euler characteristic. The students are of various philosophical stripes and the discourse benefits from his early work on Hegel with the Stalinist Lukács in Hungary and from later study with Karl Popper at the London School of Economics. I had been prepared for this dispensation by the opportunity to learn a variety of subjects from Michael Dummett. Dummett was at that time completing his study rehabilitating Frege's status, [101].

A decade later the appearance of the first 'portable' computers happily coincided with my desire to decode Srinivasa Ramanujan's (1887–1920) cryptic assertions about theta functions and elliptic integrals, [47]. I realized that by coding his formulae and my own in the *APL* programming language[1], I was able to rapidly confirm and refute identities and conjectures and to travel much more rapidly and fearlessly down potential blind alleys. I had become a computer-assisted fallibilist; at first somewhat falteringly but twenty years have certainly honed my abilities.

Today, while I appreciate fine proofs and aim to produce them when possible, I no longer view proof as the royal road to secure mathematical knowledge.

2 Introduction

I first discuss my views, and those of others, on the nature of mathematics, and then illustrate these views in a variety of mathematical contexts. A considerably more detailed treatment of many of these topics is to be found in Dave Bailey and my book *Mathematics by Experiment: Plausible Reasoning in the 21st Century*—especially in Chapters One, Two and Seven.

Kurt Gödel may well have overturned the mathematical apple cart entirely deductively, but nonetheless he could hold quite different ideas about legitimate forms of mathematical reasoning, [122]:

> *If mathematics describes an objective world just like physics, there is no reason why inductive methods should not be applied*

[1] Known as a 'write only' very high level language, APL was a fine tool; albeit with a steep learning curve whose code is almost impossible to read later.

> *in mathematics just the same as in physics.* (Kurt Gödel[2], 1951)

While we mathematicians have often separated ourselves from the sciences, they have tended to be more ecumenical. For example, a recent review of *Models. The Third Dimension of Science*, [69], chose a mathematical plaster model of a Clebsch diagonal surface as its only illustration. Similarly, authors seeking examples of the aesthetic in science often choose iconic mathematics formulae such as $E = mc^2$.

Let me begin by fixing a few concepts before starting work in earnest. Above all, I hope to persuade you of the great power of mathematical experimentation—it is also fun—and that the traditional accounting of mathematical learning and research is largely an ahistorical caricature. I recall three terms.

mathematics, n. *a group of related subjects, including algebra, geometry, trigonometry and calculus, concerned with the study of number, quantity, shape, and space, and their inter-relationships, applications, generalizations and abstractions.*

This definition–taken from my Collins Dictionary [41]—makes no immediate mention of proof, nor of the means of reasoning to be allowed. The Webster's Dictionary [1] contrasts:

induction, n. *any form of reasoning in which the conclusion, though supported by the premises, does not follow from them necessarily.*; and

deduction, n. *a process of reasoning in which a conclusion follows necessarily from the premises presented, so that the conclusion cannot be false if the premises are true.*

b. a conclusion reached by this process.

Like Gödel, I suggest that both should be entertained in mathematics. This is certainly compatible with the general view of mathematicians that in some sense "mathematical stuff is out there" to be discovered. In this paper, I shall talk broadly about experimental and heuristic mathematics, giving accessible, primarily visual and symbolic, examples.

3 Philosophy of Experimental Mathematics

> The computer has in turn changed the very nature of mathematical experience, *suggesting for the first time that mathematics, like physics, may yet become an empirical discipline, a place where things are discovered because they are seen.* (David Berlinski, [27])

[2]Taken from a previously unpublished work, [122].

The shift from *typographic* to *digital culture* is vexing for mathematicians. For example, there is still no truly satisfactory way of displaying mathematics on the web–and certainly not of asking mathematical questions. Also, we respect *authority*, [123], but value *authorship* deeply—however much the two values are in conflict, [62]. For example, the more I recast someone else's ideas in my own words, the more I enhance my authorship while undermining the original authority of the notions. Medieval scribes had the opposite concern and so took care to attribute their ideas to such as Aristotle or Plato.

And we care more about the *reliability* of our literature than does any other science, Indeed I would argue that we have reified this notion and often pay lip-service not real attention to our older literature. How often does one see original sources sprinkled like holy water in papers that make no real use of them?

The traditional central role of proof in mathematics is arguably and perhaps appropriately under siege. Via examples, I intend to pose and answer various questions. I shall conclude with a variety of quotations from our progenitors and even contemporaries:

My Questions. What constitutes secure mathematical knowledge? When is computation convincing? Are humans less fallible? What tools are available? What methodologies? What of the 'law of the small numbers'? Who cares for certainty? What is the role of proof? How is mathematics actually done? How should it be? I mean these questions both about the apprehension (discovery) and the establishment (proving) of mathematics. This is presumably more controversial in the formal proof phase.

My Answers. To misquote D'Arcy Thompson (1860–1948) 'form follows function', [218]: rigour (proof) follows reason (discovery); indeed, excessive focus on rigour has driven us away from our wellsprings. Many good ideas are wrong. Not all truths are provable, and not all provable truths are worth proving . . . . Gödel's incompleteness results certainly showed us the first two of these assertions while the third is the bane of editors who are frequently presented with correct but unexceptional and unmotivated generalizations of results in the literature. Moreover, near certainty is often as good as it gets—intellectual context (community) matters. Recent complex human proofs are often very long, extraordinarily subtle and fraught with error—consider, Fermat's last theorem, the Poincaré conjecture, the classification of finite simple groups, presumably any proof of the Riemann hypothesis, [2]. So while we mathematicians publicly talk of certainty we really settle for security.

In all these settings, modern computational tools dramatically change the nature and scale of available evidence. Given an interesting identity buried in a long and complicated paper on an unfamiliar subject, which would give you more confidence in its correctness: staring at the proof, or

confirming computationally that it is correct to 10,000 decimal places?

Here is such a formula, [13]:

$$\frac{24}{7\sqrt{7}}\int_{\pi/3}^{\pi/2}\log\left|\frac{\tan t+\sqrt{7}}{\tan t-\sqrt{7}}\right|dt \quad \stackrel{?}{=} \quad L_{-7}(2) \;= \tag{1.1}$$

$$\sum_{n=0}^{\infty}\left[\frac{1}{(7n+1)^2}+\frac{1}{(7n+2)^2}-\frac{1}{(7n+3)^2}+\frac{1}{(7n+4)^2}-\frac{1}{(7n+5)^2}-\frac{1}{(7n+6)^2}\right].$$

This identity links a volume (the integral) to an arithmetic quantity (the sum). Ir arose out of some studies in quantum field theory, in analysis of the volumes of ideal tetrahedra in hyperbolic space. The question mark is used because, while no hint of a path to a formal proof is yet known, it has been verified numerically to 20,000 digit precision–using 45 minutes on 1024 processors at Virginia Tech.

A more inductive approach can have significant benefits. For example, as there is still some doubt about the proof of the classification of finite simple groups it is important to ask whether the result is true but the proof flawed, or rather if there is still perhaps an 'ogre' sporadic group even larger than the 'monster'? What heuristic, probabilistic or computational tools can increase our confidence that the ogre does or does not exist? Likewise, there are experts who still believe the *Riemann hypothesis*[3] (RH) may be false and that the billions of zeroes found so far are much too small to be representative.[4] In any event, our understanding of the complexity of various crypto-systems relies on (RH) and we should like secure knowledge that any counter-example is enormous.

Peter Medawar (1915–87)—a Nobel prize winning oncologist and a great expositor of science—writing in *Advice to a Young Scientist*, [177], identifies four forms of scientific experiment:

1. The Kantian experiment: generating "the classical non-Euclidean geometries (hyperbolic, elliptic) by replacing Euclid's axiom of parallels (or something equivalent to it) with alternative forms." All mathematicians perform such experiments while the majority of computer explorations are of the following Baconian form.

2. The Baconian experiment is a contrived as opposed to a natural happening, it "is the consequence of 'trying things out' or even of merely messing about." Baconian experiments are the explorations of a happy if disorganized beachcomber and carry little predictive power.

[3]All non-trivial zeroes—not negative even integers—of the zeta function lie on the line with real part 1/2.

[4]See [187] and various of Andrew Odlyzko's unpublished but widely circulated works.

3. Aristotelian demonstrations: "apply electrodes to a frog's sciatic nerve, and lo, the leg kicks; always precede the presentation of the dog's dinner with the ringing of a bell, and lo, the bell alone will soon make the dog dribble." Arguably our 'Corollaries' and 'Examples' are Aristotelian, they reinforce but do not predict. Medawar then says the most important form of experiment is:

4. The Galilean: is "a critical experiment – one that discriminates between possibilities and, in doing so, either gives us confidence in the view we are taking or makes us think it in need of correction." The Galilean the only form of experiment which stands to make Experimental Mathematics a serious enterprise. Performing careful, replicable Galilean experiments requires work and care.

Reuben Hersh's arguments for a humanist philosophy of mathematics, [137, 138], as paraphrased below, become even more convincing in our highly computational setting.

1. Mathematics is human. *It is part of and fits into human culture. It does not match Frege's concept of an abstract, timeless, tenseless, objective reality.*[5]

2. Mathematical knowledge is fallible. *As in science, mathematics can advance by making mistakes and then correcting or even re-correcting them. The "fallibilism" of mathematics is brilliantly argued in Lakatos'* Proofs and Refutations.

3. There are different versions of proof or rigor. *Standards of rigor can vary depending on time, place, and other things. The use of computers in formal proofs, exemplified by the computer-assisted proof of the four color theorem in 1977*[6]*, is just one example of an emerging nontraditional standard of rigor.*

4. Empirical evidence, numerical experimentation and probabilistic proof all can help us decide what to believe in mathematics. *Aristotelian logic isn't necessarily always the best way of deciding.*

5. Mathematical objects are a special variety of a social-cultural-historical object. *Contrary to the assertions of certain post-modern detractors, mathematics cannot be dismissed as merely a new form of literature or religion. Nevertheless, many mathematical objects can be seen as shared ideas, like Moby Dick in literature, or the Immaculate Conception in religion.*

I entirely subscribe to points 2., 3., 4., and with certain caveats about

[5] That Frege's view of mathematics is wrong, for Hersh as for me, does not diminish its historical importance.

[6] Especially, since a new implementation by Seymour, Robertson and Thomas in 1997 which has produced a simpler, clearer and less troubling implementation.

objective knowledge[7] to points 1. and 5. In any event mathematics is and will remain a uniquely human undertaking.

This form of humanism sits fairly comfortably with current versions of **social-constructivism**:

> *The social constructivist thesis is that mathematics is a social construction, a cultural product, fallible like any other branch of knowledge.* (Paul Ernest, [104])

I personally qualify this with "Yes, but much less fallible than most." Associated most notably with the writings of Paul Ernest—an English Mathematician and Professor in the Philosophy of Mathematics Education who carefully traces the intellectual pedigree for his thesis, a pedigree that encompasses the writings of Wittgenstein, Lakatos, Davis, and Hersh among others—social constructivism seeks to define mathematical knowledge and epistemology through the social structure and interactions of the mathematical community and society as a whole.

This interaction often takes place over very long periods. Many of the ideas our students—and some colleagues—take for granted took a great deal of time to gel. The Greeks suspected the impossibility of the three *classical construction problems* [8] and the irrationality of the golden mean was well known to the Pythagoreans.

While concerns about potential and completed infinities are very old, until the advent of the calculus with Newton and Leibnitz and the need to handle fluxions or infinitesimals, the level of need for rigour remained modest. Certainly Euclid is in its geometric domain generally a model of rigour, while also Archimedes' numerical analysis was not equalled until the 19th century.

The need for rigour arrives in full force in the time of Cauchy and Fourier. The treacherous countably infinite processes of analysis and the limitations of formal manipulation came to the fore. It is difficult with a modern sensibility to understand how Cauchy's proof of the continuity of pointwise limits could coexist for half a century with Fourier's clear counter-examples originating in his theory of heat.

By the end of the 19th century Frege's (1848-1925) attempt to base mathematics in a linguistically based *logicism* had foundered on Russell and other's discoveries of the paradoxes of naive set theory. Within thirty

[7] While it is not Hersh's intention, a superficial reading of point 5. hints at a cultural relativism to which I certainly do not subscribe.

[8] Trisection, circle squaring and cube doubling were taken by the educated to be impossible in antiquity. Already in 414 BCE, in his play *The Birds*, Aristophanes uses 'circle-squarers' as a term for those who attempt the impossible. Similarly, the French Academy stopped accepting claimed proofs a full two centuries before the 19th century achieved proofs of their impossibility.

five years Gödel—and then Turing's more algorithmic treatment[9]—had similarly damaged both Russell and Whitehead's and Hilbert's programs.

Throughout the twentieth century, bolstered by the armour of abstraction, the great ship Mathematics has sailed on largely unperturbed. During the last decade of the 19th and first few decades of the 20th century the following main streams of philosophy emerged explicitly within mathematics to replace logicism, but primarily as the domain of philosophers and logicians.

- *Platonism.* Everyman's idealist philosophy—stuff exists and we must find it. Despite being the oldest mathematical philosophy, Platonism—still predominant among working mathematicians—was only christened in 1936.
- *Formalism.* Associated mostly with Hilbert—it asserts that mathematics is invented and is best viewed as formal symbolic games without intrinsic meaning.
- *Intuitionism.* Invented by Brouwer and championed by Heyting, intuitionism asks for inarguable monadic components that can be fully analyzed and has many variants; this has interesting overlaps with recent work in cognitive psychology such as Lakoff and Nunez' work, [158], on 'embodied cognition'[10].
- *Constructivism.* Originating with Markoff and especially Kronecker (1823–1891), and refined by Bishop it finds fault with significant parts of classical mathematics. Its 'I'm from Missouri, tell me how big it is' sensibility is not to be confused with Paul Ernest's 'social constructivism', [104].

The last two philosophies deny the principle of the *excluded middle*, "A or not A", and resonate with computer science—as does some of formalism. It is hard after all to run a deterministic program which does not know which disjunctive gate to follow. By contrast the battle between a Platonic idealism (a 'deductive absolutism') and various forms of 'fallibilism'(a quasi-empirical 'relativism') plays out across all four, but fallibilism perhaps lives most easily within a restrained version of intuitionism which looks for 'intuitive arguments' and is willing to accept that 'a proof is what

[9]The modern treatment of incompleteness leans heavily on Turing's analysis of the *Halting problem* for so-called Turing machines.

[10]"*The mathematical facts that are worthy of study are those that, by their analogy with other facts are susceptible of leading us to knowledge of a mathematical law, in the same way that physical facts lead us to a physical law.*" reflects the cognate views of Henri Poincaré (1854–1912), [194, p. 23] on the role of the *subliminal*. He also wrote *"It is by logic we prove, it is by intuition that we invent."*, [193].

convinces'. As Lakatos shows, an argument convincing a hundred years ago may well now viewed as inadequate. And one today trusted may be challenged in the next century.

As we illustrate in the next section or two, it is only perhaps in the last twenty five years, with the emergence of powerful mathematical platforms, that any approach other than a largely undigested Platonism and a reliance on proof and abstraction has had the tools[11] to give it traction with working mathematicians.

In this light, Hales' proof of Kepler's conjecture that *the densest way to stack spheres is in a pyramid* resolves the oldest problem in discrete geometry. It also supplies the most interesting recent example of intensively computer-assisted proof, and after five years with the review process was published in the *Annals of Mathematics*—with an "only 99% checked" disclaimer.

This process has triggered very varied reactions [152] and has provoked Thomas Hales to attempt a formal computational proof, [2]. Famous earlier examples of fundamentally computer-assisted proof include the *Four color theorem* and proof of the *Non-existence of a projective plane of order 10*. The three raise and answer quite distinct questions about computer assisted proof—both real and specious.

To make the case as to how far mathematical computation has come we trace the changes over the past half century. The 1949 computation of π to 2,037 places suggested by von Neumann, took 70 hours. A billion digits may now be computed in much less time on a laptop. Strikingly, it would have taken roughly 100,000 ENIACs to store the Smithsonian's picture—as is possible thanks to *40 years of Moore's law* in action . . . .

Moore's Law is now taken to be the assertion that *semiconductor technology approximately doubles in capacity and performance roughly every 18 to 24 months*. This is an astounding record of sustained exponential progress without peer in history of technology. Additionally, mathematical tools are now being implemented on parallel platforms, providing *much* greater power to the research mathematician. Amassing huge amounts of processing power will not alone solve many mathematical problems. There are very few mathematical 'Grand-challenge problems' where, as in the physical sciences, a few more orders of computational power will resolve a problem, [59]. There is, however, much more value in *very rapid 'Aha's'* as can be obtained through "micro-parallelism"; that is, where we benefit by being able to compute many simultaneous answers on a neurologically rapid scale and so can hold many parts of a problem in our mind at one time.

To sum up, in light of the discussion and terms above, I now describe

[11] That is, to broadly implement Hersh's central points (2.-4.).

myself a a social-constructivist, and as a computer-assisted fallabilist with constructivist leanings. I believe that more-and-more the interesting parts of mathematics will be less-and-less susceptible to classical deductive analysis and that Hersh's 'non-traditional standard of rigor' must come to the fore.

4 Our Experimental Mathodology

Despite Picasso's complaint that "computers are useless, they only give answers," the main goal of computation in pure mathematics is arguably to yield *insight*. This demands speed or, equivalently, substantial *micro-parallelism* to provide answers on a cognitively relevant scale; so that we may ask and answer more questions while they remain in our consciousness. This is relevant for rapid verification; for validation; for *proofs* and *especially for refutations* which includes what Lakatos calls "monster barring", [157]. Most of this goes on in the daily small-scale accretive level of mathematical discovery but insight is gained even in cases like the proof of the Four color theorem or the Non-existence of a plane of order ten. Such insight is not found in the case-enumeration of the proof, but rather in the algorithmic reasons for believing that one has at hand a tractable unavoidable set of configurations or another effective algorithmic strategy.

In this setting it is enough to equate *parallelism* with access to requisite *more* space and speed of computation. Also, we should be willing to consider all computations as 'exact' which provide truly reliable answers.[12] This now usually requires a careful *hybrid* of symbolic and numeric methods, such as achieved by *Maple*'s liaison with the *umerical Algorithms Group* (NAG) Library[13], see [109, 43]. There are now excellent tools for such purposes throughout analysis, algebra, geometry and topology, see [44, 45, 109, 59, 60].

Along the way questions required by—or just made natural by—computing start to force out older questions and possibilities in the way beautifully described a century ago by Dewey regarding evolution, [98].

> *Old ideas give way slowly; for they are more than abstract logical forms and categories. They are habits, predispositions, deeply engrained attitudes of aversion and preference. Moreover, the conviction persists—though history shows it to be a hallucination—that all the questions that the human mind has asked are questions that can be answered in terms of the alternatives that the questions themselves present. But in fact*

[12] If careful interval analysis can certify that a number known to be integer is larger that 2.5 and less than 3.5, this constitutes a exact computational proof that it is '3'.

[13] See `http://www.nag.co.uk/`.

> *intellectual progress usually occurs through sheer abandonment of questions together with both of the alternatives they assume; an abandonment that results from their decreasing vitality and a change of urgent interest. We do not solve them: we get over them. Old questions are solved by disappearing, evaporating, while new questions corresponding to the changed attitude of endeavor and preference take their place. Doubtless the greatest dissolvent in contemporary thought of old questions, the greatest precipitant of new methods, new intentions, new problems, is the one effected by the scientific revolution that found its climax in the "Origin of Species."* (John Dewey)

Additionally, what is "easy" changes: high performance computing and networking are blurring, merging disciplines and collaborators. This is democratizing mathematics but further challenging authentication—consider how easy it is to find information on *Wikipedia*[14] and how hard it is to validate it.

Moving towards a well articulated Experimental *Mathodology*—both in theory and practice—will take much effort. The need is premised on the assertions that intuition is acquired—we can and must better mesh computation and mathematics, and that visualization is of growing importance—in many settings even three is a lot of dimensions.

"Monster-barring" (Lakatos's term, [157], for refining hypotheses to rule out nasty counter-examples[15]) and "caging" (my own term for imposing needed restrictions in a conjecture) are often easy to enhance computationally, as for example with randomized checks of equations, linear algebra, and primality or graphic checks of equalities, inequalities, areas, etc.

4.1 Eight Roles for Computation

I next recapitulate eight roles for computation that Bailey and I discuss in our two recent books [44, 45]:

#1. Gaining insight and intuition or just knowledge. Working algorithmically with mathematical objects almost inevitably adds insight to the processes one is studying. At some point even just the careful aggregation of data leads to better understanding.

#2. Discovering new facts, patterns and relationships. The number of *additive partitions* of a positive integer n, $p(n)$, is *generated*

[14] *Wikipedia* is an open source project at `http://en.wikipedia.org/wiki/Main_Page`; "wiki-wiki" is Hawaian for "quickly".

[15] Is, for example, a polyhedron always convex? Is a curve intended to be simple? Is a topology assumed Hausdorff, a group commutative?

by

$$1 + \sum_{n\geq 1} p(n)q^n = \frac{1}{\prod_{n=1}^{\infty}(1-q^n)}. \tag{1.2}$$

Thus, $p(5) = 7$ since

$$\begin{aligned} 5 &= 4+1 = 3+2 = 3+1+1 \\ &= 2+2+1 = 2+1+1+1 = 1+1+1+1+1. \end{aligned} \tag{1.3}$$

Developing (1.2) is a fine introduction to enumeration via *generating functions*. Additive partitions are harder to handle than multiplicative factorizations, but they are very interesting, [45, Chapter 4]. Ramanujan used Major MacMahon's table of $p(n)$ to intuit remarkable deep congruences such as

$$p(\mathbf{5n+4}) \equiv 0 \mod \mathbf{5},\ p(\mathbf{7n+5}) \equiv 0 \mod \mathbf{7},$$

and

$$p(\mathbf{11n+6}) \equiv 0 \mod \mathbf{11},$$

from relatively limited data like

$$\begin{aligned} P(q) &= 1 + q + 2\,q^2 + 3\,q^3 + \underline{5}\,q^{\mathbf{4}} + \overline{7}\,q^{\mathbf{5}} + 11\,q^6 + 15\,q^7 \\ &+ 22\,q^8 + \underline{30}\,q^{\mathbf{9}} + 42\,q^{10} + 56\,q^{11} + \overline{77}\,q^{\mathbf{12}} + 101\,q^{13} + \underline{135}\,q^{\mathbf{14}} \\ &+ 176\,q^{15} + 231\,q^{16} + 297\,q^{17} + 385\,q^{18} + \overline{\underline{490}}\,q^{\mathbf{19}} \\ &+ 627\,q^{20}b + 792\,q^{21} + 1002\,q^{22} + \cdots + p(200)q^{200} + \cdots \end{aligned} \tag{1.4}$$

Cases $5n+4$ and $7n+5$ are flagged in (1.4). Of course, it is markedly easier to (heuristically) confirm than find these fine examples of *Mathematics: the science of patterns.*[16] The study of such congruences—much assisted by symbolic computation—is very active today.

#3. Graphing to expose mathematical facts, structures or principles. Consider Nick Trefethen's fourth challenge problem as described in [109, 43]. It requires one to find ten good digits of:

4. What is the global minimum of the function

$$e^{\sin(50x)} + \sin(60e^y) + \sin(70\sin x) + \tag{1.5}$$
$$\sin(\sin(80y)) - \sin(10(x+y)) + (x^2+y^2)/4? \tag{1.6}$$

[16] The title of Keith Devlin's 1996 book, [97].

As a foretaste of future graphic tools, one can solve this problem graphically and interactively using current *adaptive 3-D plotting* routines which can catch all the bumps. This does admittedly rely on trusting a good deal of software.

#4. Rigourously testing and especially falsifying conjectures. I hew to the Popperian scientific view that we primarily falsify; but that as we perform more and more testing experiments without such falsification we draw closer to firm belief in the truth of a conjecture such as: *the polynomial* $P(n) = n^2 - n + p$ *has prime values for all* $n = 0, 1, \ldots, p-2$*, exactly for* Euler's lucky prime numbers*, that is,* p*= 2, 3, 5, 11, 17, and 41.*[17]

#5. Exploring a possible result to see if it *merits* formal proof. A conventional deductive approach to a hard multi-step problem really requires establishing all the subordinate lemmas and propositions needed along the way. Especially if they are highly technical and unintuitive. Now some may be independently interesting or useful, but many are only worth proving if the entire expedition pans out. Computational experimental mathematics provides tools to survey the landscape with little risk of error: only if the view from the summit is worthwhile, does one lay out the route carefully. I discuss this further at the end of the next Section.

#6. Suggesting approaches for formal proof. The proof of the *cubic theta function identity* discussed on [45, pp. 210] shows how an fully intelligible human proof can be obtained entirely by careful symbolic computation.

#7. Computing replacing lengthy hand derivations. Who would wish to verify the following prime factorization by hand?

$$\begin{aligned} & 64226075786769428387925497752087347463 07 \\ = \quad & (2140992015395526641)(1963506722254397)(1527791). \end{aligned}$$

Surely, what we value is understanding the underlying algorithm, not the human work?

#8. Confirming analytically derived results. This is a wonderful and frequently accessible way of confirming results. Even if the result itself is not computationally checkable, there is often an accessible corollary. An assertion about bounded operators on Hilbert space may have a useful consequence for three-by-three matrices. It is also an excellent way to error correct, or to check calculus examples before giving a class.

[17] See [228] for the answer.

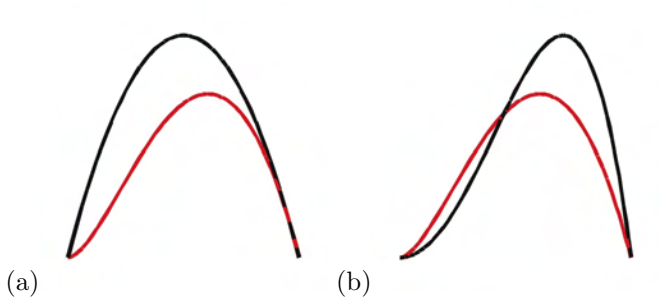


Figure 1.1. #1: Graphical comparison of $y-y^2$ and y^2-y^4 to $-y^2\ln(y)$ (red)

5 Finding Things versus Proving Things

I now illuminate these eight roles with eight mathematical examples. At the end of each I note some of the roles illustrated.

1. **Pictorially comparison** of $y - y^2$ and $y^2 - y^4$ to $-y^2\ln(y)$ when y lies in the unit interval, is a much more rapid way to divine which function is larger than by using traditional analytic methods.

   The figure shows that it is clear in the later case the functions cross, and so it is futile to try to prove one majorizes the other. In the first case, evidence is provided to motivate attempting a proof and often the picture serves to guide such a proof—by showing monotonicity or convexity or some other salient property. □

This certainly illustrates #3 and #4, and perhaps #5.

2. **A proof and a disproof.** Any modern computer algebra can tell one that

$$0 < \int_0^1 \frac{(1-x)^4 x^4}{1+x^2}\,dx = \frac{22}{7} - \pi, \tag{1.7}$$

   since the integral may be interpreted as the area under a positive curve. We are however no wiser as to why! If however we ask the same system to compute the indefinite integral, we are likely to be told that

$$\int_0^t \cdot = \frac{1}{7}\,t^7 - \frac{2}{3}\,t^6 + t^5 - \frac{4}{3}\,t^3 + 4\,t - 4\,\arctan(t)\,.$$

   Then (1.7) is now rigourously established by differentiation and an appeal to the Fundamental theorem of calculus. □

This illustrates points #1 and #6. It also falsifies the bad conjecture that $\pi = 22/7$ and so illustrates #4 again. Finally, the computer's proof is easier (#7) and very nice, though probably it is not the one we would have developed by ourselves. The fact that 22/7 is a continued fraction approximation to π has lead to many hunts for generalizations of (1.7), see [45, Chapter 1]. None so far are entirely successful.

3. **A computer discovery and a 'proof' of the series for** $\arcsin^2(x)$. We compute a few coefficients and observe that there is a regular power of 4 in the numerator, and integers in the denominator; or equivalently we look at $\arcsin(x/2)^2$. The generating function package 'gfun' in *Maple*, then predicts a recursion, r, for the denominators and solves it, as R.

```
>with(gfun):
>s:=[seq(1/coeff(series(arcsin(x/2)^2,x,25),x,2*n),n=1..6)]:
>R:=unapply(rsolve(op(1, listtorec(s,r(m))),r(m)),m);
>[seq(R(m),m=0..8)];
```

yields, $s := [4, 48, 360, 2240, 12600, 66528]$,

$$R := m \mapsto 8\,\frac{4^m\,\Gamma(3/2+m)(m+1)}{\pi^{1/2}\Gamma(1+m)},$$

where Γ is the Gamma function, and then returns the sequence of values

$$[4, 48, 360, 2240, 12600, 66528, 336336, 1647360, 7876440].$$

We may now use Sloane's *Online Encyclopedia of Integer Sequences*[18] to reveal that the coefficients are $R(n) = 2n^2\binom{2n}{n}$. More precisely, sequence A002544 identifies

$$R(n+1)/4 = \binom{2n+1}{n}(n+1)^2.$$

```
> [seq(2*n^2*binomial(2*n,n),n=1..8)];
```

confirms this with

$$[4, 48, 360, 2240, 12600, 66528, 336336, 1647360].$$

Next we write

[18] At `www.research.att.com/~njas/sequences/index.html`

```
> S:=Sum((2*x)^(2*n)/(2*n^2*binomial(2*n,n)),
  n=1..infinity):S=values(S);
```

which returns

$$\frac{1}{2}\sum_{n=1}^{\infty}\frac{(2\,x)^{2n}}{n^2\,\binom{2n}{n}} = \arcsin^2(x).$$

That is, we have discovered—and proven if we trust or verify *Maple*'s summation algorithm—the desired Maclaurin series.

As prefigured by Ramanujan, it transpires that there is a beautiful closed form for $\arcsin^{2m}(x)$ for all $m = 1, 2, \ldots$. In [55] there is a discussion of the use of *integer relation methods*, [44, Chapter 6], to find this closed form and associated proofs are presented. □

Here we see an admixture of all of the roles save #3, but above all #2 and #5.

4. **Discovery without proof.** Donald Knuth[19] asked for a closed form evaluation of:

$$\sum_{k=1}^{\infty}\left\{\frac{k^k}{k!\,e^k}-\frac{1}{\sqrt{2\,\pi\,k}}\right\}-0.084069508727655\ldots. \tag{1.8}$$

Since about 2000 CE it has been easy to compute 20—or 200—digits of this sum, and the 'smart lookup' facility in the *Inverse Symbolic Calculator*(ISC). The ISC at www.cecm.sfu.ca/projects/ISC/ISCmain.html uses a variety of search algorithms and heuristics to predict what a number might actually be. Similar ideas are now implemented as 'identify' in *Maple* and 'Recognize' in *Mathematica*, and are described in [43, 44, 60, 22]. In this case it *rapidly* returns

$$0.084069508727655 \approx \frac{2}{3}+\frac{\zeta\,(1/2)}{\sqrt{2\,\pi}}.$$

We thus have a prediction which *Maple* 9.5 on a 2004 laptop *confirms* to 100 places in under 6 seconds and to 500 in 40 seconds. Arguably we are done. After all we were asked to *evaluate* the series and we now know a closed-form answer. Notice also that the 'divergent' $\zeta(1/2)$ term is formally to be expected! □

[19] Posed as an MAA Problem [149].

We have discovered and tested the result and in so doing gained insight and knowledge while illustrating #1, #2 and #4. Moreover, as described in [45, pp. 15], one can also be lead by the computer to a very satisfactory computer-assisted but very human proof, thus illustrating #6. Indeed, the first hint is that the computer algebra system returned the value in (1.8) very quickly even though the series is very slowly convergent. This suggests the program is doing something intelligent—and it is!

5. A striking conjecture with no known proof strategy is: for $N = 1, 2, 3 \cdots$

$$8^N \zeta\left(\{\overline{2}, 1\}_N\right) \stackrel{?}{=} \zeta\left(\{2, 1\}_N\right). \tag{1.9}$$

Explicitly, the first two cases are

$$8 \sum_{n>m>0} \frac{(-1)^n}{n^2 m} = \sum_{n>0} \frac{1}{n^3} \text{ and } 64 \sum_{n>m>o>p>0} \frac{(-1)^n}{n^2 m\, o^2 p} = \sum_{n>m>0} \frac{1}{n^3 m^3}.$$

The notation should now be clear—we use the 'overbar' to denote an alternation. Such alternating sums are called *multi-zeta values* (MZV) and positive ones are called *Euler sums* after Euler who first studied them seriously. They arise naturally in a variety of modern fields from combinatorics to mathematical physics and knot theory.

There is abundant evidence amassed since 'identity' (1.9) was found in 1996. For example, very recently Petr Lisonek checked the first 85 cases to 1000 places in about 41 HP hours with only the *predicted round-off error*. And the case N=163 was checked in about ten hours. These objects are very hard to compute naively and require substantial computation as a precursor to their analysis.

Formula (1.9) is the *only* identification of its type of an Euler sum with a distinct MZV and we have no idea why it is true. Any similar MZV proof has been both highly non-trivial and illuminating. To illustrate how far we are from proof: can just the case $n = 2$ be proven *symbolically* as has been the case for $n = 1$? □

This identity was discovered by the British quantum field theorist David Broadhurst and me during a large hunt for such objects in the mid-nineties. In this process we discovered and proved many lovely results (see [44, Chapter 2] and [45, Chapter 4]), thereby illustrating #1,#2, #4, #5 and #7. In the case of 'identity' (1.9) we have failed with #6, but we have ruled out many sterile approaches. It is one of many examples where we can now have (near) certainty without proof. Another was shown in (1.1).

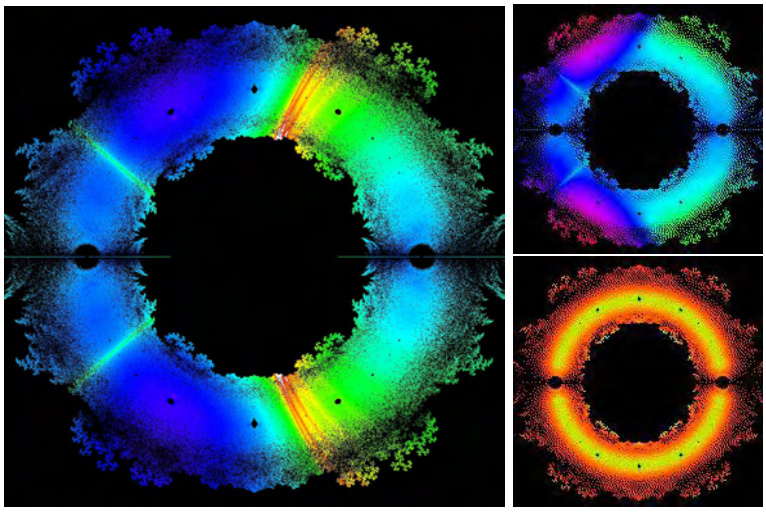


Figure 1.2. #6: "The price of metaphor is eternal vigilance." (Arturo Rosenblueth & Norbert Wiener, [168])

6. **What you draw *is* what you see.** *Roots of polynomials with coefficients 1 or -1 up to degree 18.*

   As the quote suggests, pictures are highly metaphorical. The colouration is determined by a normalized sensitivity of the coefficients of the polynomials to slight variations around the values of the zeros with red indicating low sensitivity and violet indicating high sensitivity. It is hard to see how the structure revealed in the pictures above[20] would be seen other than through graphically data-mining. Note the different shapes—now proven—of the holes around the various roots of unity.

   The striations are unexplained but all re-computations expose them! And the fractal structure is provably there. Nonetheless different ways of measuring the stability of the calculations reveal somewhat different features. This is very much analogous to a chemist discovering an unexplained but robust spectral line. □

This certainly illustrates #2 and #7, but also #1 and #3.

[20] We plot all complex zeroes of polynomials with only -1 and 1 as coefficients up to a given degree. As the degree increases some of the holes fill in—at different rates.

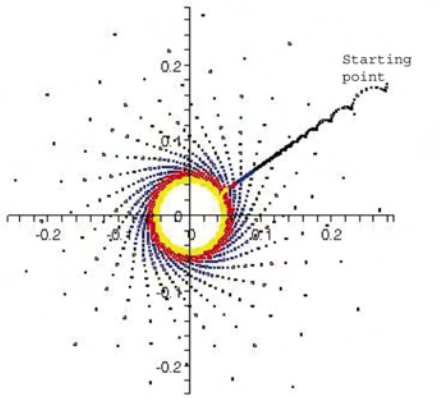


Figure 1.3. #7: "Visual convergence in the complex plane"

7. Visual Dynamics. In recent continued fraction work, Crandall and I needed to study the *dynamical system* $t_0 := t_1 := 1$:

$$t_n \hookleftarrow \frac{1}{n}\, t_{n-1} + \omega_{n-1} \left(1 - \frac{1}{n}\right) t_{n-2},$$

where $\omega_n = a^2, b^2$ for n even, odd respectively, are two unit vectors. Think of this as a **black box** which we wish to examine scientifically. Numerically, all one *sees* is $t_n \to 0$ slowly. Pictorially, we *learn* significantly more[21]. If the iterates are plotted with colour changing after every few hundred iterates, it is clear that they spiral roman-candle like in to the origin:

Scaling by $\sqrt{n}$, and coloring even and odd iterates, *fine structure* appears. We now observe, predict and validate that the outcomes depend on whether or not one or both of a and b are roots of unity (that is, rational multiples of π). Input a p-th root of unity and out comes p spirals, input a non-root of unity and we see a circle. □

This forceably illustrates #2 but also #1, #3, #4. It took my coauthors and me, over a year and 100 pages to convert this intuition into a rigorous formal proof, [13]. Indeed, the results are technical and delicate enough that I have more faith in the facts than in the finished argument. In this sentiment, I am not entirely alone.

[21]... "Then felt I like a watcher of the skies, when a new planet swims into his ken." (*Chapman's Homer*)

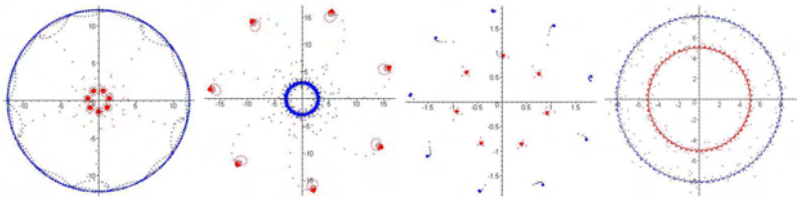


Figure 1.4. #7: The attractors for various $|a| = |b| = 1$

Carl Friedrich Gauss, who drew (carefully) and computed a great deal, once noted, *I* ***have*** *the result, but I do not yet know how to get it.*[22] An excited young Gauss writes: "A new field of analysis has appeared to us, self-evidently, in the study of functions etc." (October 1798). It had and the consequent proofs pried open the doors of much modern elliptic function and number theory.

My penultimate and more comprehensive example is more sophisticated and I beg the less-expert analyst's indulgence. Please consider its structure and not the details.

8. A full run. Consider the *unsolved* **Problem 10738** from the 1999 *American Mathematical Monthly*, [45]:

Problem: For $t > 0$ let

$$m_n(t) \sum_{k=0}^{\infty} k^n \exp(-t)\, \frac{t^k}{k!}$$

be the *nth* moment of a *Poisson distribution* with parameter t. Let $c_n(t)m_n(t)/n!$. Show

a) $\{m_n(t)\}_{n=0}^{\infty}$ is log-convex[23] for all $t > 0$.
b) $\{c_n(t)\}_{n=0}^{\infty}$ is not log-concave for $t < 1$.
c*) $\{c_n(t)\}_{n=0}^{\infty}$ is log-concave for $t \geq 1$.

Solution. (a) Neglecting the factor of $\exp(-t)$ as we may, this reduces to

$$\sum_{k,j\geq 0} \frac{(jk)^{n+1} t^{k+j}}{k!j!} \leq \sum_{k,j\geq 0} \frac{(jk)^n t^{k+j}}{k!\, j!}\, k^2 = \sum_{k,j\geq 0} \frac{(jk)^n t^{k+j}}{k!j!}\, \frac{k^2+j^2}{2},$$

[22] Likewise, the quote has so far escaped exact isolation!

[23] A sequence $\{a_n\}$ is *log-convex* if $a_{n+1}a_{n-1} \geq a_n^2$, for $n \geq 1$ and log-concave when the sign is reversed.

and this now follows from $2jk \leq k^2 + j^2$.

(b) As

$$m_{n+1}(t) = t \sum_{k=0}^{\infty} (k+1)^n \exp(-t) \frac{t^k}{k!},$$

on applying the binomial theorem to $(k+1)^n$, we see that $m_n(t)$ satisfies the recurrence

$$m_{n+1}(t) = t \sum_{k=0}^{n} \binom{n}{k} m_k(t), \qquad m_0(t) = 1.$$

In particular for $t = 1$, we computationally obtain as many terms of the sequence

$$1, 1, 2, 5, 15, 52, 203, 877, 4140 \ldots$$

as we wish. These are the *Bell numbers* as was discovered again by consulting *Sloane's Encyclopedia* which can also tell us that, for $t = 2$, we have the *generalized Bell numbers*, and gives the exponential generating functions.[24] Inter alia, an explicit computation shows that

$$t \frac{1+t}{2} c_0(t)\, c_2(t) \leq c_1(t)^2 = t^2$$

exactly if $t \geq 1$, which completes (b).

Also, preparatory to the next part, a simple calculation shows that

$$\sum_{n \geq 0} c_n u^n = \exp\left(t(e^u - 1)\right). \tag{1.10}$$

(c*)[25] We appeal to a recent theorem, [45], due to E. Rodney Canfield which proves the lovely and quite difficult result below. A self-contained proof would be very fine.

Theorem 5.1. *If a sequence* $1, b_1, b_2, \cdots$ *is non-negative and log-concave then so is the sequence* $1, c_1, c_2, \cdots$ *determined by the generating function equation*

$$\sum_{n \geq 0} c_n u^n = \exp\left(\sum_{j \geq 1} b_j \frac{u^j}{j} \right).$$

[24] Bell numbers were known earlier to Ramanujan—an example of *Stigler's Law of Eponymy*, [45, p. 60].

[25] The '*' indicates this was the unsolved component.

Using equation (1.10) above, we apply this to the sequence $\mathbf{b_j} = \mathbf{t}/(\mathbf{j}-\mathbf{1})!$ which is log-concave exactly for $t \geq 1$. □

A search in 2001 on *MathSciNet* for "Bell numbers" since 1995 turned up 18 items. Canfield's paper showed up as number 10. Later, *Google* found it immediately!

Quite unusually, the given solution to (c) was the only one received by the *Monthly*. The reason might well be that it relied on the following sequence of steps:

A **(Question Posed)** ⇒ Computer Algebra System ⇒ Interface ⇒
Search Engine ⇒ Digital Library ⇒ Hard New Paper ⇒ **(Answer)**

Without going into detail, we have visited most of the points elaborated in Section 4.1. Now if only we could already automate this process!

Jacques Hadamard, describes the role of proof as well as anyone—and most persuasively given that his 1896 proof of the Prime number theorem is an inarguable apex of rigorous analysis.

> *The object of mathematical rigor is to sanction and legitimize the conquests of intuition, and there was never any other object for it.*[26] (Jacques Hadamard)

Of the eight uses of computers instanced above, let me reiterate the central importance of heuristic methods for determining what is true and whether it merits proof. I tentatively offer the following surprising example which is very very likely to be true, offers no suggestion of a proof and indeed may have no reasonable proof.

9. **Conjecture.** *Consider*

$$x_n \quad = \quad \left\{ 16x_{n-1} + \frac{120n^2 - 89n + 16}{512n^4 - 1024n^3 + 712n^2 - 206n + 21} \right\}. \tag{1.11}$$

The sequence $\beta_n = (\lfloor 16x_n \rfloor)$, where (x_n) is the sequence of iterates defined in equation (1.11), precisely generates the hexadecimal expansion of $\pi - 3$.

(Here $\{\cdot\}$ denotes the fractional part and $(\lfloor\cdot\rfloor)$ denotes the integer part.) In fact, we know from [44, Chapter 4] that the first million iterates are correct and in consequence:

$$\sum_{n=1}^{\infty} \|x_n - \{16^n \pi\}\| \quad \leq \quad 1.46 \times 10^{-8} \ldots. \tag{1.12}$$

[26] Hadamard quoted in [195]. See also [194].

where $\|a\| = \min(a, 1-a)$. By the first Borel-Cantelli lemma this shows that the hexadecimal expansion of π only finitely differs from (β_n). Heuristically, the probability of any error is very low. □

6 Conclusions

To summarize, I do argue that reimposing the primacy of mathematical knowledge over proof is appropriate. So I return to the matter of what it takes to persuade an individual to adopt new methods and drop time honoured ones. Aptly, we may start by consulting Kuhn on the matter of paradigm shift:

> *The issue of paradigm choice can never be unequivocally settled by logic and experiment alone. ⋯ in these matters neither proof nor error is at issue. The transfer of allegiance from paradigm to paradigm is a conversion experience that cannot be forced.*[27] (Thomas Kuhn)

As we have seen, the pragmatist philosopher John Dewey eloquently agrees, while Max Planck, [192], has also famously remarked on the difficulty of such paradigm shifts. This is Kuhn's version[28]:

> *And Max Planck, surveying his own career in his Scientific Autobiography, sadly remarked that "a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it."* (Albert Einstein, [156, 192])

This transition is certainly already apparent. It is certainly rarer to find a mathematician under thirty who is unfamiliar with at least one of *Maple*, *Mathematica* or *MatLab*, than it is to one over sixty five who is really fluent. As such fluency becomes ubiquitous, I expect a re-balancing of our community's valuing of deductive proof over inductive knowledge.

In his '23' "*Mathematische Probleme*" lecture to the Paris International Congress in 1900, Hilbert writes[29]

[27] In [206], *Who Got Einstein's Office?* The answer is Arne Beurling.

[28] Kuhn is quoting Einstein quoting Planck. There are various renderings of this second-hand German quotation.

[29] See the late Ben Yandell's fine account of the Hilbert Problems and their solvers, [233]. The written lecture is considerably longer and further ranging that the one delivered in person.

> *Moreover a mathematical problem should be difficult in order to entice us, yet not completely inaccessible, lest it mock our efforts. It should be to us a guidepost on the mazy path to hidden truths, and ultimately a reminder of our pleasure in the successful solution.* (David Hilbert)

Note the primacy given by a most exacting researcher to discovery and to truth over proof and rigor. More controversially and most of a century later, Greg Chaitin invites us to be bolder and act more like physicists.

> *I believe that elementary number theory and the rest of mathematics should be pursued more in the spirit of experimental science, and that you should be willing to adopt new principles...* And the Riemann Hypothesis isn't self-evident either, but it's very useful. *A physicist would say that there is ample experimental evidence for the Riemann Hypothesis and would go ahead and take it as a working assumption. ⋯ We may want to introduce it formally into our mathematical system* (Greg Chaitin, [44, p. 254])

Ten years later:

> *[Chaitin's] "Opinion" article proposes that the Riemann hypothesis (RH) be adopted as a new axiom for mathematics. Normally one could only countenance such a suggestion if one were assured that the RH was undecidable. However, a proof of undecidability is a logical impossibility in this case, since if RH is false it is provably false. Thus, the author contends, one may either wait for a proof, or disproof, of RH—both of which could be impossible—or one may take the bull by the horns and accept the RH as an axiom. He prefers this latter course as the more positive one.*(Roger Heath Brown[30])

Much as I admire the challenge of Greg Chaitin's statements, I am not yet convinced that it is helpful to add axioms as opposed to proving conditional results that start "Assuming the continuum hypothesis" or emphasize that "without assuming the Riemann hypothesis we are able to show." Most important is that we lay our cards on the table. We should explicitly and honestly indicate when we believe our tools to be heuristic, we should carefully indicate why we have confidence in our computations—and where our uncertainty lies— and the like.

On that note, Hardy is supposed to have commented (somewhat dismissively) that Landau, a great German number theorist, would never be the

[30]Roger Heath-Brown's *Mathematical Review* of [72], 2004.

first to prove the Riemann Hypothesis, but that if someone else did so then Landau would have the best possible proof shortly after. I certainly hope that a more experimental methodology will better value independent replication and honour the first transparent proof[31] of Fermat's last theorem as much as Andrew Wiles' monumental proof. Hardy also commented that he did his best work past forty. Inductive, accretive, tool-assisted mathematics certainly allows brilliance to be supplemented by experience and—as in my case—stands to further undermine the notion that one necessarily does ones best mathematics young.

6.1 Last Words

To reprise, I hope to have made convincing arguments that the traditional deductive accounting of Mathematics is a largely ahistorical caricature—Euclid's millennial sway not withstanding.[32] Above all, mathematics is primarily about *secure knowledge* not proof, and that while the aesthetic is central, we must put much more emphasis on notions of supporting evidence and attend more closely to the reliability of witnesses.

Proofs are often out of reach—but understanding, even certainty, is not. Clearly, computer packages can make concepts more accessible. A short list includes linear relation algorithms, Galois theory, Groebner bases, etc. While progress is made "*one funeral at a time*"[33], in Thomas Wolfe's words "*you can't go home again*" and as the co-inventor of the Fast Fourier transform properly observed

> *Far better an approximate answer to the right question, which is often vague, than the exact answer to the wrong question, which can always be made precise.* (J. W. Tuckey, 1962)

[31] If such should such exist and as you prefer be discovered or invented.

[32] Most of the cited quotations are stored at `jborwein/quotations.html`

[33] This grim version of Planck's comment is sometimes attributed to Niels Bohr but this seems specious. It is also spuriously attributed on the web to Michael Milken, and I imagine many others

Figure 1.5. Coxeter's favourite 4D polyhedron

Chapter 2

Algorithms for Experimental Mathematics, Part One

Many different computational methods have been used in experimental mathematics. Just a few of the more widely-used methods are the following:

1. Symbolic computation for algebraic and calculus manipulations.
2. Integer-relation methods, especially the "PSLQ" algorithm.
3. High-precision integer and floating-point arithmetic.
4. High-precision evaluation of integrals and infinite series summations.
5. The Wilf-Zeilberger algorithm for proving summation identities.
6. Iterative approximations to continuous functions.
7. Identification of functions based on graph characteristics.
8. Graphics and visualization methods targeted to mathematical objects.

In this chapter and in Chapter 2 we will present an overview of some of these methods, and give examples of how they have been used in some real-world experimental math research. We will focus on items 2, 3, 4 and 5, mainly because the essential ideas can be explained more easily than, say, the mechanics behind symbolic computation or advanced scientific visualization.

1 High-Precision Arithmetic

We have already seen numerous examples of high-precision numerical calculations. Indeed, such computations frequently arise in experimental mathematics. We shall focus here on high-precision floating-point computation. High-precision integer computation is also required in some aspects of mathematical computation, particularly in prime number computations and symbolic manipulations, but as we shall see, many of the algorithms described below are equally applicable to both types of arithmetic. An excellent presentation of high-precision integer arithmetic is given in [90].

By "arbitrary precision" we mean a software facility that permits one to adjust the level of numeric precision over a wide range, typically extending to the equivalent of thousands or possibly even millions of decimal digits. An extended dynamic range is almost always included as well, since such computations often require a larger range than the $10^{\pm 308}$ range available with the IEEE double format.

For these levels of precision, the best approach is as follows. Define an arbitrary precision datum to be an $(n+4)$-long string of words. The sign of the first word is the sign $\pm$ of the datum, and the absolute value of the first word is n, the number of mantissa words used. The second word contains an exponent p. Words three through $n+2$ are the n mantissa words m_i, each of which has an integer value between 0 and $2^b - 1$, or in other words b bits of the mantissa. Finally, words $n+3$ and $n+4$ are reserved as "scratch" words for various arithmetic routines. One can optionally designate an additional word, placed at the start of the data structure, to specify the amount of memory available for this datum, so as to avoid memory overwrite errors during execution. The value A represented by this datum is

$$A \quad = \quad \pm(2^{pb}m_1 + 2^{(p-1)b}m_2 + 2^{(p-2)b}m_3 + \cdots + 2^{(p-n+1)b}m_n),$$

where it is assumed that $m_1 \neq 0$ and $m_n \neq 0$ for nonzero A. Zero is represented by a string consisting of a sign word and an exponent word, both of which are zero.

There are several variations possible with this general design. One approach is to utilize 64-bit IEEE floating-point words, with $b48$ mantissa bits per word. Addition operations can easily be performed by adding the two vectors of mantissas (suitably shifted to adjust for differences in exponent), and then releasing carries beginning from the last mantissa word back to the first. Multiplication and division can be performed by straightforward adaptations of the long multiplication and long division schemes taught in grammar school, performed modulo 2^{48} instead of modulo 10. The multiplication of two individual 48-bit entities can be performed by simple algorithms, or, on some systems, by using the "fused multiply-add" hardware instruction. Up to $2^5 = 32$ such products can be accumulated

before needing to release carries, since $5 + 48 = 53$, and integers as large as 2^{53} can be accommodated exactly in a 64-bit IEEE word. This approach was taken in the software package described in [20], and available at the URL

`http://www.experimentalmath.info`

This software includes C++ and Fortran-90 translation modules, so that these functions can be invoked from ordinary programs with only minor modifications to the source code.

Another approach is to utilize arrays of integer data, with integer arithmetic operations, since all values in the data structure above are whole numbers. One disadvantage of this approach is it is hard to write programs that are both fast and easily portable to different systems. Nonetheless, some integer-based implementations have been very successful, notably the GNU package, available at the URL

`http://www.gnu.org/software/gmp/gmp.html`

Either way, fast Fourier transforms (FFTs) can be used to accelerate arithmetic for higher levels of precision (approximately 1000 digits or more). The reason for the savings is that an FFT calculation scales only as $n \log_2 n$ in computation al cost, compared with the n^2 for conventional methods (where n is the precision in digits or words).

2 Integer Relation Detection

Integer relation detection methods are employed very often in experimental math applications to recognize a mathematical constant whose numerical value can be computed to at least moderately high precision, and also to discover relations between a set of computed numerical values.

For a given real vector $(x_1, x_2, \cdots, x_n)$, an integer relation algorithm is a computational scheme that either finds the n integers $(a_1, a_2, \cdots, a_n)$, not all zero, such that $a_1 x_1 + a_2 x_2 + \cdots a_n x_n = 0$ or else establishes that there is no such integer vector within a ball of some radius about the origin, where the metric is the Euclidean norm $(a_1^2 + a_2^2 + \cdots + a_n^2)^{1/2}$.

At the present time, the best known integer relation algorithm is the PSLQ algorithm [113] of Helaman Ferguson. Another widely used integer relation detection scheme involves the Lenstra-Lenstra-Lovasz (LLL) lattice reduction algorithm. The PSLQ algorithm, together with related lattice reduction schemes such as LLL, was recently named one of ten "algorithms of the century" by the publication *Computing in Science and Engineering* [11]. In addition to possessing good numerical stability, PSLQ is guaranteed to find a relation in a polynomially bounded number of iterations. The

name "PSLQ" derives from its usage of a partial sum of squares vector and a LQ (lower-diagonal-orthogonal) matrix factorization.

A simplified formulation of the standard PSLQ algorithm, mathematically equivalent to the original formulation, is given in [19][44, pg 230–234]. These references also describe another algorithm, called "multi-pair" PSLQ, which is well-suited for parallel processing, and which runs faster even on a one-CPU system than the standard PSLQ. Two-level and three-level variants of both standard PSLQ and multi-pair PSLQ, which economize on run time by performing most iterations using ordinary 64-bit IEEE arithmetic, are described in [19].

High-precision arithmetic must be used for almost all applications of integer relation detection methods, using PSLQ or any other algorithm. This stems from the fact that if one wishes to recover a relation of length n, with coefficients of maximum size d digits, then the input vector x must be specified to at least nd digits, and one must employ floating-point arithmetic accurate to at least nd digits, or else the true solution will be lost in a sea of numerical artifacts. PSLQ typically recovers relations when the input data is specified to at least 10 or 15 percent greater precision than this minimum value, and when a working precision of at least this level is used to implement PSLQ.

PSLQ operates by constructing a series of matrices A_n, such that the entries of the vector $y_n = A_n^{-1}x$ steadily decrease in size. At any given iteration, the largest and smallest entries of y_n usually normally differ by no more than two or three orders of magnitude. When a relation is detected by the algorithm, the smallest entry of the y_n vector abruptly decreases to roughly the "epsilon" of the working precision (i.e., 10^{-p}, where p is the precision level in digits), and the desired relation is given by the corresponding column of A_n^{-1}. See Figure 2.1, which shows this behavior for a typical PSLQ computation.

The detection threshold in the termination test for PSLQ is typically set to be a few orders of magnitude greater than the epsilon value, in order to allow for reliable relation detection in the presence of some numerical roundoff error. The ratio between the smallest and the largest entry of the vector $A^{-1}x$ when a relation is detected can be taken as a "confidence level" that the relation is a true relation and not an artifact of insufficient numeric precision. Very small ratios at detection, such as 10^{-100}, almost certainly denote a true relation (although, of course, such results are experimental only, and do not constitute rigorous proof).

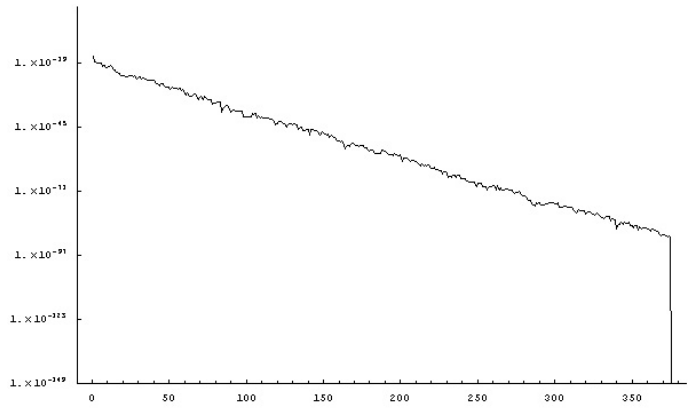


Figure 2.1. Smallest entry of $y_n = A_n^{-1}x$ in a typical PSLQ run, as a function of n.

3 Illustrations and Examples

3.1 The BBP Formula for Pi

Perhaps the best-known application of PSLQ is the 1995 discovery, by means of a PSLQ computation, of the "BBP" formula for π:

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k}\left[\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6}\right]. \qquad (2.1)$$

This formula permits one to calculate directly binary or hexadecimal digits beginning at the nth digit, without the need to calculate any of the first $n-1$ digits, using a simple algorithm and standard machine arithmetic [18].

The genesis of this discovery was the realization, by Peter Borwein and Simon Plouffe, that individual binary digits of $\log 2$ could be calculated, by applying this well-known classical formula:

$$\log 2 = \sum_{k=0}^{\infty} \frac{1}{k2^k}. \qquad (2.2)$$

Suppose we wish to compute a few binary digits beginning at position $d+1$ for some integer $d > 0$. This is equivalent to calculating $\{2^d \log 2\}$, where

$\{\cdot\}$ denotes fractional part. Thus we can write

$$\begin{aligned}\{2^d \log 2\} &= \left\{ \left\{ \sum_{k=0}^{d} \frac{2^{d-k}}{k} \right\} + \sum_{k=d+1}^{\infty} \frac{2^{d-k}}{k} \right\} \\ &= \left\{ \left\{ \sum_{k=0}^{d} \frac{2^{d-k} \bmod k}{k} \right\} + \sum_{k=d+1}^{\infty} \frac{2^{d-k}}{k} \right\}. \qquad (2.3)\end{aligned}$$

We are justified in inserting "mod k" in the numerator of the first summation, because we are only interested in the fractional part of the quotient when divided by k.

Now the key observation is this: The numerator of the first sum in (2.3), namely $2^{d-k} \bmod k$, can be calculated very rapidly by means of the binary algorithm for exponentiation, performed modulo k.

The binary algorithm for exponentiation is merely the formal name for the observation that exponentiation can be economically performed by means of a factorization based on the binary expansion of the exponent. For example, we can write $3^{17} = ((((3^2)^2)^2)^2) \cdot 3 = 129140163$, thus producing the result in only 5 multiplications, instead of the usual 16. If we are only interested in the result modulo 10, then we can calculate $((((3^2 \bmod 10)^2 \bmod 10)^2 \bmod 10)^2 \bmod 10) \cdot 3 \bmod 10)3$, and we never have to store or operate on integers larger than 81. Indeed, this particular calculation can be done in one's head.

Since that we can very rapidly evaluate each term of the first summation in (2.3), and since the second summation can be truncated after just a few terms, it is clear one can quickly calculate, say, the first 40 digits of the binary expansion of $\log 2$, beginning with some position $d + 1$, where $d < 10^7$, using only standard IEEE 64-bit floating-point arithmetic. If one uses 128-bit floating-point arithmetic, or "double-double" arithmetic, then one can calculate more digits beginning at the desired position $d + 1$, and this calculation is reliable for $d \leq 10^{15}$.

After this discovery by Peter Borwein and Simon Plouffe, they immediately began to investigate whether individual digits of π could be computed in this manner. It is clear that we can employ this technique on any constant that can be written using a formula of the form

$$\alpha = \sum_{n=1}^{\infty} \frac{p(n)}{b^n q(n)},$$

where $b > 1$ is an integer and p and q are polynomials with integer coefficients, with q having no zeroes at positive integer arguments. However, at this time (1995), there were no known formulas for π of this form. So they began to investigate (with the help of Bailey's PSLQ computer program)

Position	Hex Digits Beginning at This Position
10^6	26C65E52CB4593
10^7	17AF5863EFED8D
10^8	ECB840E21926EC
10^9	85895585A0428B
10^{10}	921C73C6838FB2
10^{11}	9C381872D27596
1.25×10^{12}	07E45733CC790B
2.5×10^{14}	E6216B069CB6C1

Table 2.1. Computed hexadecimal digits of π.

whether π was a linear combination of other constants that are of this form. These computer runs were redone numerous times over the course of two or three months, as new constants of the requisite form were found in the literature. Eventually, the BBP formula for π was discovered.

Table 2.1 gives some results of calculations that have been done in this manner. The last-listed result, which is tantamount to a computation of the one-quadrillionth binary digit of π, was performed on over 1,700 computers worldwide, using software written by Colin Percival.

The BBP formula for π has even found a practical application—it is now used in the g95 Fortran compiler as part of the procedure for evaluating certain transcendental functions.

3.2 Bifurcation Points in Chaos Theory

One application of integer relation detection methods is to find the minimal polynomial of an algebraic constant. Note that if α satisfies a polynomial $a_0 + a_1 t + \cdots + a_n = 0$, then we can discover this polynomial simply by computing, to high precision, the values of $1, \alpha, \alpha^2, \cdot, \alpha^n$, and then applying PSLQ or some other integer relation scheme to the resulting $(n+1)$-long vector.

The chaotic iteration $x_{n+1} = r x_n (1 - x_n)$ has been studied since the early days of chaos theory in the 1950s. It is often called the "logistic iteration," since it mimics the behavior of an ecological population that, if its growth one year outstrips its food supply, often falls back in numbers for the following year, thus continuing to vary in a highly irregular fashion. When r is less than one iterates of the logistic iteration converge to zero. For r in the range $1 < r < B_1 = 3$ iterates converge to some nonzero limit. If $B_1 < r < B_2 = 1 + \sqrt{6} = 3.449489\ldots$, the limiting behavior bifurcates—every other iterate converges to a distinct limit point. For r with $B_2 < r < B_3$ iterates hop between a set of four distinct limit

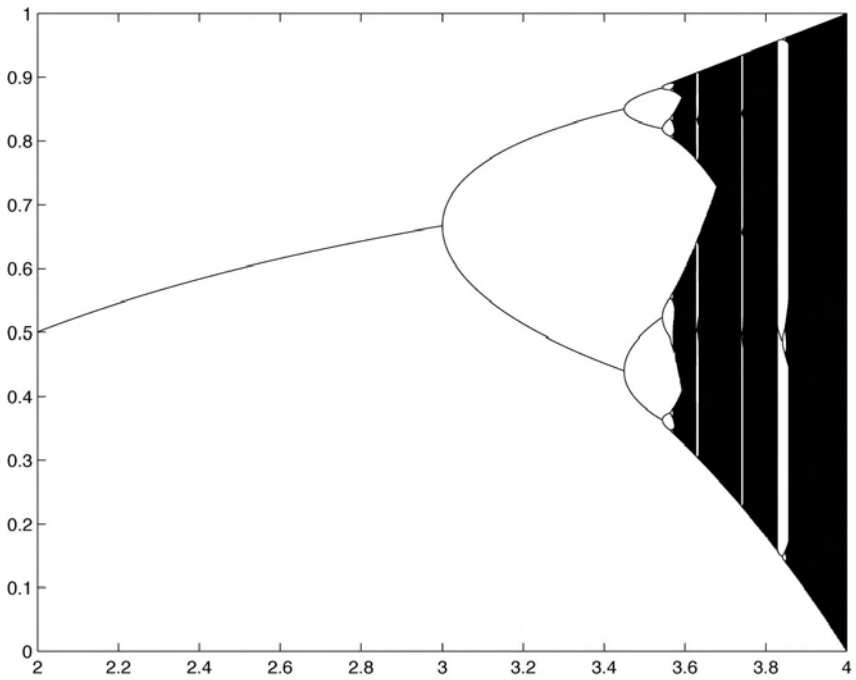


Figure 2.2. Bifurcation in the logistic iteration.

points; when $B_3 < r < B_4$, they select between a set of eight distinct limit points; this pattern repeats until $r > B_\infty = 3.569945672\ldots$, when the iteration is completely chaotic (see Figure 2.2). The limiting ratio $\lim_n (B_n - B_{n-1})/(B_{n+1} - B_n) = 4.669201\ldots$ is known as *Feigenbaum's delta constant*.

It is fairly easy to see that all of these B constants are algebraic numbers, but the bounds one obtains on the degree are often rather large, and thus not very useful. Thus one may consider using PSLQ or some other integer relation algorithm to discover their minimal polynomials.

A highly accurate numerical value of B_3, for instance, can be obtained using a relatively straightforward search scheme. Let $f_8(r, x)$ be the eight-times iterated evaluation of $rx(1-x)$, and let $g_8(r, x) = f_8(r, x) - x$. Imagine a three-dimensional graph, where r ranges from left to right and x ranges from bottom to top (as in Figure 2.2), and where $g_8(r, x)$ is plotted in the vertical (out-of-plane) dimension. Given some initial r slightly less than B_3, we compute a "comb" of function values at n evenly spaced x values (with spacing h_x) near the limit of the iteration $x_{n+1} = f_8(r, x_n)$. In our implementation, we use $n = 12$, and we start with $r = 3.544, x =$

$0.364, h_r = 10^{-4}$, and $h_x = 5 \times 10^{-4}$. With this construction, the comb has $n/2$ negative function values, followed by $n/2$ positive function values. We then increment r by h_r and re-evaluate the "comb," continuing in this fashion until two sign changes are observed among the n function values of the "comb." This means that a bifurcation occurred just prior to the current value of r, so we restore r to its previous value (by subtracting h_r), reduce h_r, say by a factor of four, and also reduce the h_x roughly by a factor of 2.5. We continue in this fashion, moving the value of r and its associated "comb" back and forth near the bifurcation point with progressively smaller intervals h_r. The center of the comb in the x direction must be adjusted periodically to ensure that $n/2$ negative function values are followed by $n/2$ positive function values, and the spacing parameter h_x must be adjusted as well to ensure that two sign changes are disclosed when this occurs. We quit when the smallest of the n function values is within two or three orders of magnitude of the "epsilon" of the arithmetic (e.g., for 2000-digit working precision, "epsilon" is 10^{-2000}). The final value of r is then the desired value B_3, accurate to within a tolerance given by the final value of r_h. With 2000-digit working precision, our implementation of this scheme finds B_3 to 1330-digit accuracy in about five minutes on a 2004-era computer. The first hundred digits are as follows:

$$\begin{aligned} B_3 \;=\;& 3.54409035955192285361596598660480454058309984544457367545 \\ & 78125303058429428588630122562585664248917999626\ldots \end{aligned}$$

With even a moderately accurate value of r in hand (at least two hundred digits or so), one can use a PSLQ program to check to see whether r is an algebraic constant. When $n \geq 12$, the relation

$$\begin{aligned} 0 \;=\;& r^{12} - 12r^{11} + 48r^{10} - 40r^9 - 193r^8 + 392r^7 + 44r^6 + 8r^5 - 977r^4 \\ & -604r^3 + 2108r^2 + 4913 \end{aligned} \tag{2.4}$$

can be recovered.

The significantly more challenging problem of computing and analyzing the constant $B_4 = 3.564407266095\cdots$ is discussed in [19]. In this study, conjectural reasoning suggested that B_4 might satisfy a 240-degree polynomial, and, in addition, that $\alpha = -B_4(B_4 - 2)$ might satisfy a 120-degree polynomial. The constant α was then computed to over 10,000-digit accuracy, and an advanced three-level multi-pair PSLQ program was employed, running on a parallel computer system, to find an integer relation for the vector $(1, \alpha, \alpha^2, \cdots, \alpha^{120})$. A numerically significant solution was obtained, with integer coefficients descending monotonically from 257^{30}, which is a 73-digit integer, to the final value, which is one (a striking result that is exceedingly unlikely to be a numerical artifact). This experimentally discovered polynomial was recently confirmed in a large symbolic computation [154].

Additional information on the Logistic Map is available at

`http://mathworld.wolfram.com/LogisticMap.html`

.

3.3 Sculpture

The PSLQ algorithm, which was discovered in 1993 by Helaman Ferguson. This is certainly a signal accomplishment—for example, the PSLQ algorithm (with associated lattice reduction algorithms) was recently named one of ten "algorithms of the century" by *Computing in Science and Engineering* [11]. Nonetheless Ferguson is even more well-known for his numerous mathematics-inspired sculptures, which grace numerous research institutes in the United States. Photos and highly readable explanations of these sculptures can be seen in a lovely book written by his wife, Claire [112]. Together, the Fergusons recently won the 2002 Communications Award, bestowed by the Joint Policy Board of Mathematics.

Ferguson notes that the PSLQ algorithm can be thought of as a n-dimension extension of the Euclidean algorithm, and is, like the Euclidean scheme, fundamentally a "subtractive" algorithm. As Ferguson explains, "It is also true that my sculptural form of expression is subtractive: I get my mathematical forms by direct carving of stone." [200]

There is a remarkable, as well as entirely unanticipated, connection between Ferguson's PSLQ algorithm and one of Ferguson's sculptures. It is known that the volumes of complements of certain knot figures (which volumes in R^3 are infinite) are finite in hyperbolic space, and sometimes are given by certain explicit formulas. This is not true of all knots. Many of these hyperbolic complements of knots correspond to certain discrete quotient subgroups of matrix groups.

One of Ferguson's well-known sculptures is the "Figure-Eight Complement II" (see Figure 2.3). It has been known for some time that the hyperbolic volume V of the figure-eight knot complement is given by the formula

$$V = 2\sqrt{3}\sum_{n=1}^{\infty}\frac{1}{n\binom{2n}{n}}\sum_{k=n}^{2n-1}\frac{1}{k} \tag{2.5}$$

$$= 2.02988321281930725004240510854 9\ldots \tag{2.6}$$

In 1998, British physicist David Broadhurst conjectured that $V/\sqrt{3}$ is a rational linear combination of

$$C_j = \sum_{n=0}^{\infty}\frac{(-1)^n}{27^n(6n+j)^2}. \tag{2.7}$$

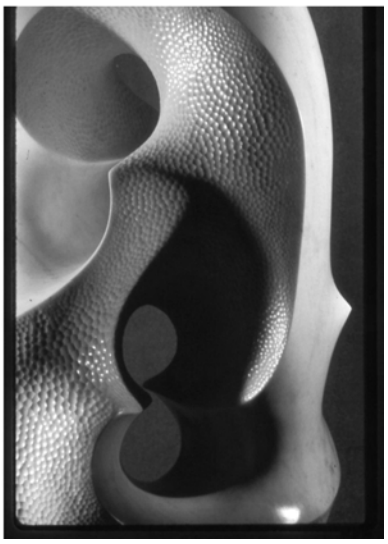


Figure 2.3. Ferguson's "Figure-Eight Knot Complement" sculpture

Indeed, it is, as Broadhurst [66] found using a PSLQ program:

$$V = \frac{\sqrt{3}}{9}\sum_{n=0}^{\infty}\frac{(-1)^n}{27^n}\left(\frac{18}{(6n+1)^2}-\frac{18}{(6n+2)^2}-\frac{24}{(6n+3)^2} -\frac{6}{(6n+4)^2}+\frac{2}{(6n+5)^2}\right). \tag{2.8}$$

You can verify this yourself, using for example the Mathematician's Toolkit, available at

`http://www.experimentalmath.info`

Just type the following lines of code:

```
v = 2 * sqrt[3] * sum[1/(n * binomial[2*n,n]) * sum[1/k, \
  {k, n, 2*n-1}], {n, 1, infinity}]
pslq[v/sqrt[3], table[sum[(-1)^n/(27^n*(6*n+j)^2), \
  {n, 0, infinity}], {j, 1, 6}]]
```

When this is done you will recover the solution vector $(9, -18, 18, 24, 6, -2, 0)$. A proof that formula (2.8) holds, together with a number of other identities for V, is given at the end of this chapter in [44, pg xx]. This proof, by the way, is a classic example of experimental methodology, in that it relies on "knowing" ahead of time that the formula holds.

Ferguson comments that the discovery of this BBP-type expression for V is a "major advance toward understanding the figure-eight knot complement volume." Accordingly, he has carved Broadhurst's formula on the figure-eight knot complement sculptures commissioned by the Clay Mathematics Institute, both the Inner Mongolian black granite piece and the smaller bronzes (the Clay Math Award pieces). As he explains, "Finally the subtractive sculpture and the subtractive algorithm have come together."

3.4 Euler Sums

In April 1993, Enrico Au-Yeung, an undergraduate at the University of Waterloo, brought to the attention of one of us (Borwein) the curious result

$$\sum_{k=1}^{\infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{k}\right)^2 k^{-2} \quad = \quad 4.59987\ldots \tag{2.9}$$

$$\approx \quad \frac{17}{4}\zeta(4) \;=\; \frac{17\pi^4}{360}. \tag{2.10}$$

The function $\zeta(s)$ in (2.10) is the classical *Riemann zeta-function*:

$$\zeta(s) \quad = \quad \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Au-Yeung had computed the sum in (2.10) to 500,000 terms, giving an accuracy of five or six decimal digits. Suspecting that his discovery was merely a modest numerical coincidence, Borwein sought to compute the sum to a higher level of precision. Using Fourier analysis and Parseval's equation, he obtained

$$\frac{1}{2\pi}\int_0^{\pi} (\pi - t)^2 \log^2(2\sin\frac{t}{2})\, dt \quad = \quad \sum_{n=1}^{\infty} \frac{(\sum_{k=1}^{n} \frac{1}{k})^2}{(n+1)^2}. \tag{2.11}$$

The series on the right of (2.11) permits one to evaluate (2.10), while the integral on the left can be computed using the numerical quadrature facility of *Mathematica* or *Maple*. When he did this, Borwein was surprised to find that the conjectured identity holds to more than thirty digits.

The summation in (2.11) is a special case of the *multivariate zeta-function*, defined as

$$\zeta(s_1, s_2, \cdots, s_k) \quad = \quad \sum_{n_1 > n_2 > \cdots > n_k > 0} \prod_{j=1}^{k} n_j^{-|s_j|} \sigma_j^{-n_j},$$

where the $s_1, s_2, \ldots, s_k$ are nonzero integers and $\sigma_j = \text{signum}(s_j)$. A fast method for computing such sums, based on Hölder convolution is discussed

in [52], and is discussed further in Chapter 3. For the time being, it suffices to note that the scheme is implemented in EZFace+, an online tool available at the URL

`http://www.cecm.sfu.ca/projects/ezface+`

We will illustrate its application to one specific case, namely the analytic identification of the sum

$$S_{2,3} \quad = \quad \sum_{k=1}^{\infty} \left(1 - \frac{1}{2} + \cdots + (-1)^{k+1}\frac{1}{k}\right)^2 (k+1)^{-3}. \qquad (2.12)$$

Expanding the squared term in (2.12), we have

$$\sum_{\substack{0<i,j<k \\ k>0}} \frac{(-1)^{i+j+1}}{ijk^3} = -2\,\zeta(3,-1,-1) + \zeta(3,2). \qquad (2.13)$$

Evaluating this in EZFace+ we quickly obtain

$$C = 0.15616693338117691588103590968798819368577670984030387295752935449707503744029579145520565370935814757 8\ldots.$$

Based on our experience with other multivariate zeta constants, we conjectured that this constant satisfies a rational linear relation involving the following constants: $\pi^5, \pi^4\log(2), \pi^3\log^2(2), \pi^2\log^3(2), \pi\log^4(2), \log^5(2)$, $\pi^2\zeta(3), \pi\log(2)\zeta(3), \log^2(2)\zeta(3), \zeta(5), \mathrm{Li}_5(1/2)$. Note that each of these constants can be seen to be have "degree" five. The result is quickly found to be

$$\begin{aligned} C \quad = \quad & 4\,\mathrm{Li}_5\left(\frac{1}{2}\right) - \frac{1}{30}\log^5(2) - \frac{17}{32}\zeta(5) - \frac{11}{720}\pi^4\log(2) + \frac{7}{4}\zeta(3)\log^2(2) \\ & + \frac{1}{18}\pi^2\log^3(2) - \frac{1}{8}\pi^2\zeta(3). \end{aligned}$$

This result has been proved in various ways, both analytic and algebraic. Indeed, all evaluations of sums of the form $\zeta(\pm a_1, \pm a_2, \cdots, \pm a_m)$ with *weight* $w = \sum_k a_m, (k < 8)$, as in (2.13) have been established.

High precision calculations of many of these sums, together with considerable investigations involving heavy use of *Maple's* symbolic manipulation facilities, eventually yielded numerous new, rigorously established results

[42]. A few examples include:

$$\sum_{k=1}^{\infty}\left(1+\frac{1}{2}+\cdots+\frac{1}{k}\right)^2(k+1)^{-4} = \frac{37}{22680}\pi^6-\zeta^2(3),$$

$$\sum_{k=1}^{\infty}\left(1+\frac{1}{2}+\cdots+\frac{1}{k}\right)^3(k+1)^{-6} =$$

$$\zeta^3(3)+\frac{197}{24}\zeta(9)+\frac{1}{2}\pi^2\zeta(7)-\frac{11}{120}\pi^4\zeta(5)-\frac{37}{7560}\pi^6\zeta(3),$$

$$\sum_{k=1}^{\infty}\left(1-\frac{1}{2}+\cdots+(-1)^{k+1}\frac{1}{k}\right)^2(k+1)^{-3} =$$

$$4\,\mathrm{Li}_5\left(\frac{1}{2}\right)-\frac{1}{30}\log^5(2)-\frac{17}{32}\zeta(5)-\frac{11}{720}\pi^4\log(2)+\frac{7}{4}\zeta(3)\log^2(2)$$

$$+\frac{1}{18}\pi^2\log^3(2)-\frac{1}{8}\pi^2\zeta(3), \tag{2.14}$$

where $\mathrm{Li}_n(x)=\sum_{k>0}x^k/k^n$ denotes the polylogarithm function.

3.5 Quantum Field Theory

In another recent development, David Broadhurst (who discovered the identity (2.8) for Ferguson's Clay Math Award sculpture) has found, using similar methods, that there is an intimate connection between Euler sums and constants resulting from evaluation of Feynman diagrams in quantum field theory [67, 68]. In particular, the renormalization procedure (which removes infinities from the perturbation expansion) involves multiple zeta values. He has shown [66], using PSLQ computations, that in each of ten cases with unit or zero mass, the finite part of the scalar 3-loop tetrahedral vacuum Feynman diagram reduces to four-letter "words" that represent iterated integrals in an alphabet of seven "letters" comprising the single 1-form $\Omega=dx/x$ and the six 1-forms $\omega_k=dx/(\lambda^{-k}-x)$, where $\lambda=(1+\sqrt{-3})/2$ is the primitive sixth root of unity, and k runs from 0 to 5. A four-letter word here is a four-dimensional iterated integral, such as

$$U=\zeta(\Omega^2\omega_3\omega_0)=$$

$$\int_0^1\frac{dx_1}{x_1}\int_0^{x_1}\frac{dx_2}{x_2}\int_0^{x_2}\frac{dx_3}{(-1-x_3)}\int_0^{x_3}\frac{dx_4}{(1-x_4)}\sum_{j>k>0}\frac{(-1)^{j+k}}{j^3k}.$$

There are 7^4 such four-letter words. Only two of these are primitive terms occurring in the 3-loop Feynman diagrams: U, above, and

$$V = \mathrm{Re}[\zeta(\Omega^2\omega_3\omega_1)]=\sum_{j>k>0}\frac{(-1)^j\cos(2\pi k/3)}{j^3k}.$$

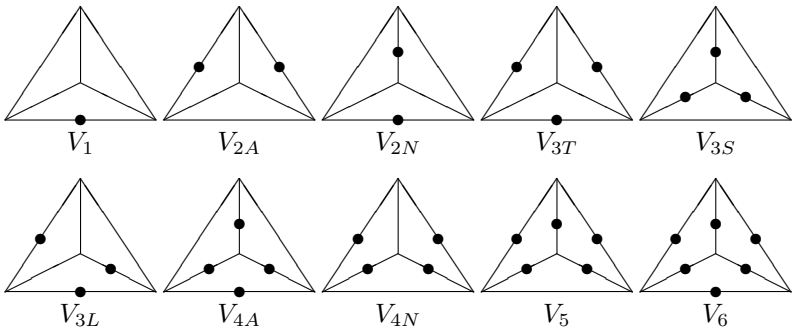


Figure 2.4. The ten tetrahedral configurations.

The remaining terms in the diagrams reduce to products of constants found in Feynman diagrams with fewer loops. These ten cases are shown in Figure 2.4. In these diagrams, dots indicate particles with nonzero rest mass. The formulas that have been found, using PSLQ, for the corresponding constants are given in Table 2.2. In the table the constant $C = \sum_{k>0} \sin(\pi k/3)/k^2$.

4 Definite Integrals and Infinite Series Summations

One particularly useful application of integer relation computations is to evaluate definite integrals and sums of infinite series by means of numerical calculations. We use one of various methods to obtain a numerical value for the integral or summation, then try to identify this value by means of

V_1	=	$6\zeta(3) + 3\zeta(4)$
V_{2A}	=	$6\zeta(3) - 5\zeta(4)$
V_{2N}	=	$6\zeta(3) - \frac{13}{2}\zeta(4) - 8U$
V_{3T}	=	$6\zeta(3) - 9\zeta(4)$
V_{3S}	=	$6\zeta(3) - \frac{11}{2}\zeta(4) - 4C^2$
V_{3L}	=	$6\zeta(3) - \frac{15}{4}\zeta(4) - 6C^2$
V_{4A}	=	$6\zeta(3) - \frac{77}{12}\zeta(4) - 6C^2$
V_{4N}	=	$6\zeta(3) - 14\zeta(4) - 16U$
V_5	=	$6\zeta(3) - \frac{469}{27}\zeta(4) + \frac{8}{3}C^2 - 16V$
V_6	=	$6\zeta(3) - 13\zeta(4) - 8U - 4C^2$

Table 2.2. Formulas found by PSLQ for the ten tetrahedral diagrams.

integer relation methods.

In many cases, one can apply online tools (which employ integer relation techniques combined with large-scale table-lookup schemes) to "blindly" identify numerical values. One of the most popular and effective tools is the Inverse Symbolic Computation (ISC) tool, available at

`http://www.cecm.sfu.ca/projects/ISC`

In other cases, these tools are unable to identify the constant, and custom-written programs must be used, which typically take advantage of knowledge that certain sums are likely to involve a certain class of constants.

As one example, we were inspired by a recent problem in the *American Mathematical Monthly* [5]. By using one of the quadrature routines to be described in the next section, together with the ISC tool and a PSLQ integer relation detection program, we found that if $C(a)$ is defined by

$$C(a) = \int_0^1 \frac{\arctan(\sqrt{x^2+a^2})\,dx}{\sqrt{x^2+a^2}(x^2+1)}, \tag{2.15}$$

then

$$\begin{aligned} C(0) &= \pi\log 2/8 + G/2 \\ C(1) &= \pi/4 - \pi\sqrt{2}/2 + 3\sqrt{2}\arctan(\sqrt{2})/2 \\ C(\sqrt{2}) &= 5\pi^2/96, \end{aligned} \tag{2.16}$$

where $G = \sum_{k\ge 0}(-1)^k/(2k+1)^2$ is Catalan's constant. The third of these results is the result from the *Monthly*. These particular results then led to the following general result, among others:

$$\int_0^\infty \frac{\arctan(\sqrt{x^2+a^2})\,dx}{\sqrt{x^2+a^2}(x^2+1)} = \tag{2.17}$$
$$\frac{\pi}{2\sqrt{a^2-1}}\left[2\arctan(\sqrt{a^2-1}) - \arctan(\sqrt{a^4-1})\right].$$

We will discuss techniques for computing definite integrals and sums of series to high precision in Chapter 3. For the time being, we simply note that both *Mathematica* and *Maple* have incorporated some reasonably good numerical facilities for this purpose, and it is often sufficient to rely on these packages when numerical values are needed.

5 Computation of Multivariate Zeta Constants

One class of mathematical constants that has been of particular interest to experimental mathematicians in the past few years are multivariate zeta constants. Research in this arena has been facilitated by the discovery of methods that permit the computation of these constants to high precision. While Euler-Maclaurin-based schemes can be used (and in fact were used) in these studies, they are limited to 2-order sums. We present here an algorithm that permits even high-order sums to be evaluated to hundreds or thousands of digit accuracy. We will limit our discussion here to multivariate zeta constants of the form

$$\zeta(s_1, s_2, \cdots, s_n) = \sum_{n_1 > n_2 > \cdots > n_k} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_k^{s_k}}, \tag{2.18}$$

for positive integers s_k and n_k, although in general the technique we describe here has somewhat broader applicability.

This scheme is as follows [52]. For $1 \le j \le m$, define the numeric strings

$$a_j = \{s_j + 2, \{1\}_{r_j}, s_{j+1}, \{1\}_{r_{j+1}}, \cdots, s_m + 2, \{1\}_{r_m}\} \tag{2.19}$$

$$b_j = \{r_j + 2, \{1\}_{s_j}, r_{j-1}, \{1\}_{s_{j-1}}, \cdots, r_1 + 2, \{1\}_{s_1}\}, \tag{2.20}$$

where by the notation $\{1\}_n$, we mean n repetitions of 1. For convenience, we will define a_{m+1} and b_0 to be the empty string.

Define

$$\delta(s_1, s_2, \cdots, s_k) = \prod_{j=1}^{k} \left[\sum_{\nu_j 1}^{\infty} 2^{-\nu_j} \left(\sum_{i=j}^{k} \nu_i \right)^{-s_j} \right]. \tag{2.21}$$

Then we have

$$\begin{aligned} \zeta(a_m) = \sum_{j=1}^{m} \Bigg[& \sum_{t=0}^{s_j+1} \delta(s_j + 2 - t, \{1\}^{r_j}, a_{j+1}) \delta(\{1\}^t, b_{j-1}) \\ & + \sum_{u=1}^{r_j} \delta(\{1\}^u, a_{j+1}) \delta(r_j + 2 - v, 1^{s_j}, b_{j-1}) \Bigg] + \delta(b_m) \end{aligned} \tag{2.22}$$

See the discussion in Chapter 3 of [45] for further details. An online tool that implements this procedure is available at

`http://www.cecm.sfu.ca/projects/ezface+`

This procedure has also been implemented as part of the Experimental Mathematician's Toolkit, available at

`http://www.experimentalmath.info`

6 Ramanujan-Type Elliptic Series

Truly new types of infinite series formulas, based on elliptic integral approximations, were discovered by Srinivasa Ramanujan (1887–1920) around 1910, but were not well known (nor fully proven) until quite recently when his writings were widely published. They are based on elliptic functions and are described at length in [46]. One of these is the remarkable formula

$$\frac{1}{\pi} \quad = \quad \frac{2\sqrt{2}}{9801} \sum_{k=0}^{\infty} \frac{(4k)!\,(1103 + 26390k)}{(k!)^4 396^{4k}}. \tag{2.23}$$

Each term of this series produces an additional *eight* correct digits in the result. When Gosper used this formula to compute 17 million digits of π in 1985, and it agreed to many millions of places with the prior estimates, *this concluded the first proof* of (2.23), as described in [46]. Actually, Gosper first computed the simple continued fraction for π, hoping to discover some new things in its expansion, but found none. At about the same time, David and Gregory Chudnovsky found the following rational variation of Ramanujan's formula. It exists because $\sqrt{-163}$ corresponds to an imaginary quadratic field with class number one:

$$\frac{1}{\pi} \quad = \quad 12 \sum_{k=0}^{\infty} \frac{(-1)^k\,(6k)!\,(13591409 + 545140134k)}{(3k)!\,(k!)^3\,640320^{3k+3/2}}. \tag{2.24}$$

Each term of this series produces an additional 14 correct digits. The Chudnovskys implemented this formula using a clever scheme that enabled them to use the results of an initial level of precision to extend the calculation to even higher precision. They used this in several large calculations of π, culminating with a *record computation* (in 1994) to over four billion decimal digits. Their remarkable story was compellingly told by Richard Preston in a prizewinning *New Yorker* article "The Mountains of Pi" [199].

While these Ramanujan and Chudnovsky series are in practice considerably more efficient than classical formulas, they share the property that the number of terms needed increases linearly with the number of digits desired: *if you wish to compute twice as many digits of π, you must evaluate twice as many terms of the series.* Relatedly, the Ramanujan-type series

$$\frac{1}{\pi} \quad = \quad \sum_{n=0}^{\infty} \left(\frac{\binom{2n}{n}}{16^n} \right)^3 \frac{42\,n + 5}{16} \tag{2.25}$$

allows one to compute the billionth binary digit of $1/\pi$, or the like, *without computing the first half* of the series, and is a foretaste of our discussion of Borwein-Bailey-Plouffe (or BBP) formulas.

In some recent papers, Guillera has exhibited several new Ramanujan-style series formulas for reciprocal powers of π, including the following [128, 129, 130]:

$$\frac{128}{\pi^2} = \sum_{n=0}^{\infty}(-1)^n r(n)^5(13+180n+820n^2)\left(\frac{1}{32}\right)^{2n} \quad (2.26)$$

$$\frac{32}{\pi^2} = \sum_{n=0}^{\infty}(-1)^n r(n)^5(1+8n+20n^2)\left(\frac{1}{2}\right)^{2n} \quad (2.27)$$

$$\frac{32}{\pi^3} = \sum_{n=0}^{\infty} r(n)^7(1+14n+76n^2+168n^3)\left(\frac{1}{32}\right)^{2n}. \quad (2.28)$$

where we define the function $r(n)$ as follows:

$$r(n) = \frac{(1/2)_n}{n!} = \frac{1/2\cdot 3/2\cdot\cdots\cdot(2n-1)/2}{n!} = \frac{\Gamma(n+1/2)}{\sqrt{\pi}\,\Gamma(n+1)}.$$

Guillera proved (2.26) and (2.27) using Wilf-Zeilberger's method described in Chapter 3. He ascribes series (2.28) to Gourevich who also found it using integer relation methods. Guillera also provides other series for $1/\pi^2$ based on other Gamma-function values as in (2.23) and (2.24) but for our experiments we restrict ourselves to $r(n)$.

6.1 Experiments with Ramanujan-Type Series

We have attempted to do a more thorough experimental search for identities of this general type. In particular, we searched for formulas of either of the two forms

$$\frac{c}{\pi^m} = \sum_{n=0}^{\infty} r(n)^{2m+1}(p_0+p_1 n+\cdots+p_m n^m)\alpha^{2n} \quad (2.29)$$

$$\frac{c}{\pi^m} = \sum_{n=0}^{\infty}(-1)^n r(n)^{2m+1}(p_0+p_1 n+\cdots+p_m n^m)\alpha^{2n}. \quad (2.30)$$

Here c is some integer linear combination of the constants (d_i, $1 \le i \le 34$):

$$1,\ 2^{1/2},\ 2^{1/3},\ 2^{1/4},\ 2^{1/6},\ 4^{1/3},\ 8^{1/4},\ 32^{1/6},\ 3^{1/2},\ 3^{1/3},\ 3^{1/4},\ 3^{1/6},\ 9^{1/3},$$
$$27^{1/4},\ 243^{1/6},\ 5^{1/2},\ 5^{1/4},\ 125^{1/4},\ 7^{1/2},\ 13^{1/2},\ 6^{1/2},\ 6^{1/3},\ 6^{1/4},\ 6^{1/6},$$
$$7, 36^{1/3},\ 216^{1/4},\ 7776^{1/6},\ 12^{1/4},\ 108^{1/4},\ 10^{1/2},\ 10^{1/4},\ 15^{1/2}.$$

The polynomial coefficients (p_k, $1 \le k \le m$) in (2.29) and (2.30) are each some integer linear combination of the constants (q_i, $1 \le i \le 11$):

$$1,\ 2^{1/2},\ 3^{1/2},\ 5^{1/2},\ 6^{1/2},\ 7^{1/2},\ 10^{1/2},\ 13^{1/2},\ 14^{1/2},\ 15^{1/2},\ 30^{1/2}.$$

Note that the linear combination chosen for a given p_k may be different from that chosen for any of the others. The constant α in (2.29) and (2.30) is chosen from

$$1/2,\ 1/4,\ 1/8,\ 1/16,\ 1/32,\ 1/64,\ 1/128,\ 1/256,\ \sqrt{5}-2,\ (2-\sqrt{3})^2,$$
$$5\sqrt{13}-18,\ (\sqrt{5}-1)^4/128,\ (\sqrt{5}-2)^4,\ (2^{1/3}-1)^4/2,\ 1/(2\sqrt{2}),$$
$$(\sqrt{2}-1)^2,\ (\sqrt{5}-2)^2,\ (\sqrt{3}-\sqrt{2})^4.$$

This list of α constants was taken from a table on page 172 of [46].

These searches were done using a two-level PSLQ integer relation finding program, with 1000-digit precision. Each selection of m and α constituted one separate integer relation search. In particular, for a fixed m and α in (2.29), we calculated the $[34 + 11(m + 1)]$-long set of real numbers

$$d_1,\ d_2,\ \cdots,\ d_{34},$$
$$q_0 \sum_{n=0}^{\infty} r(n)^{2m+1}\alpha^{2n},\ q_1 \sum_{n=0}^{\infty} r(n)^{2m+1}\alpha^{2n},\ \cdots,\ q_{11} \sum_{n=0}^{\infty} r(n)^{2m+1}\alpha^{2n},$$
$$q_1 \sum_{n=0}^{\infty} r(n)^{2m+1}n\alpha^{2n},\ q_2 \sum_{n=0}^{\infty} r(n)^{2m+1}n\alpha^{2n},\ \cdots q_{11} \sum_{n=0}^{\infty} r(n)^{2m+1}n\alpha^{2n},\ \cdots$$
$$q_1 \sum_{n=0}^{\infty} r(n)^{2m+1}n^m\alpha^{2n},\ q_2 \sum_{n=0}^{\infty} r(n)^{2m+1}n^m\alpha^{2n},\ \cdots q_{11} \sum_{n=0}^{\infty} r(n)^{2m+1}n^m\alpha^{2n}$$

and then applied a two-level PSLQ program, implemented using the ARPREC multiple-precision software, to this vector.

After finding a relation with our program, we carefully checked to ensure that it was not reducible to another in the list by an algebraic manipulation. Also, in numerous cases, multiple relations existed. In such cases, we eliminated these one by one, typically by replacing one of the constants in the relation by an unrelated transcendental and re-running the program, until no additional relations were found.

The result of this effort is the following list of relations. As it turns out, each of these are given either implicitly or explicitly in [46] or [129]. But just as important here is the apparent non-existence of additional relations. In particular, if a relation is not shown below for a given α and/or sign choice, that means (as a consequence of our calculations) that there is no such relation with integer coefficients whose Euclidean norm is less than 10^{10}.

For degree $m = 1$, with non-alternating signs:

$$\begin{aligned}
\frac{4}{\pi} &= \sum_{n=0}^{\infty} r(n)^3(1+6n)\left(\frac{1}{2}\right)^{2n} \\
\frac{16}{\pi} &= \sum_{n=0}^{\infty} r(n)^3(5+42n)\left(\frac{1}{8}\right)^{2n} \\
\frac{12^{1/4}}{\pi} &= \sum_{n=0}^{\infty} r(n)^3(-15+9\sqrt{3}-36n+24\sqrt{3}n)\left(2-\sqrt{3}\right)^{4n} \\
\frac{32}{\pi} &= \sum_{n=0}^{\infty} r(n)^3(-1+5\sqrt{5}+30n+42\sqrt{5}n)\left(\frac{(\sqrt{5}-1)^4}{128}\right)^{2n} \\
\frac{5^{1/4}}{\pi} &= \sum_{n=0}^{\infty} r(n)^3(-525+235\sqrt{5}-1200n+540\sqrt{5}n)\left(\sqrt{5}-2\right)^{8n}.
\end{aligned}$$

For degree $m = 1$, with alternating signs:

$$\begin{aligned}
\frac{2\sqrt{2}}{\pi} &= \sum_{n=0}^{\infty} (-1)^n r(n)^3(1+6n)\left(\frac{1}{2\sqrt{2}}\right)^{2n} \\
\frac{2}{\pi} &= \sum_{n=0}^{\infty} (-1)^n r(n)^3(-5+4\sqrt{2}-12n+12\sqrt{2}n)\left(\sqrt{2}-1\right)^{4n} \\
\frac{2}{\pi} &= \sum_{n=0}^{\infty} (-1)^n r(n)^3(23-10\sqrt{5}+60n-24\sqrt{5}n)\left(\sqrt{5}-2\right)^{4n} \\
\frac{2}{\pi} &= \sum_{n=0}^{\infty} (-1)^n r(n)^3(177-72\sqrt{6}+420n-168\sqrt{6}n)\left(\sqrt{3}-\sqrt{2}\right)^{8n}.
\end{aligned}$$

For degree $m = 2$,

$$\begin{aligned}
\frac{8}{\pi^2} &= \sum_{n=0}^{\infty} (-1)^n r(n)^5(1+8n+20n^2)\left(\frac{1}{2}\right)^{2n} \\
\frac{128}{\pi^2} &= \sum_{n=0}^{\infty} (-1)^n r(n)^5(13+180n+820n^2)\left(\frac{1}{32}\right)^{2n}.
\end{aligned}$$

For degree $m = 3$,

$$\frac{32}{\pi^3} = \sum_{n=0}^{\infty} r(n)^7(1+14n+76n^2+168n^3)\left(\frac{1}{8}\right)^{2n}.$$

For degree $m = 4, 5$ we have been unable to find any similar series, with exclusion bounds roughly 10^{10} as before, thereby (so far) dashing our hope to find an infinite family of rational sums extending (2.25), (2.26), 2.27), (2.28). More study, however, will be needed to understand this phenomenon.

6.2 Working with the Series Analytically

While (2.26), 2.27), (2.28) have no "explanation," there are tantalizing echoes of the elliptic theory described in [46] that explains the series for $1/\pi$ as we now partially reprise. We first define the *theta functions* θ_3, θ_4 and θ_2

$$\theta_3(q) := \sum_{n=-\infty}^{\infty} q^{n^2}, \quad \theta_4(q) := \sum_{n=-\infty}^{\infty} (-1)^n q^{n^2}, \quad \theta_2(q) := \sum_{n=-\infty}^{\infty} q^{(n+1/2)^2},$$

for $|q| < 1$. We next identify the *invariant*

$$k_N = \frac{\theta_2^2}{\theta_3^2}\left(e^{-\pi\sqrt{N}}\right).$$

We denote the *complementary modulus* $k' := \sqrt{1-k^2}$ in terms of which it transpires that Jacobi's identity $\theta_3^4 = \theta_4^4 + \theta_2^4$ (see [46]) implies

$$k_N^{'} = \frac{\theta_2^2}{\theta_3^2}\left(e^{-\pi\sqrt{N}}\right).$$

For reasons detailed in [46] and [45] we know that for each natural number N, k_N is algebraic.

For example, $k_1 = 1/\sqrt{2} = k_1'$ while k_{210} is the *singular value* sent to Hardy in Ramanujan's famous 1913 letters of introduction—ignored by two other famous English mathematicians.

$$\begin{aligned} k_{210} := & \left(\sqrt{2}-1\right)^2\left(\sqrt{3}-2\right)\left(\sqrt{7}-6\right)^2\left(8-3\sqrt{7}\right) \\ & \times\left(\sqrt{10}-3\right)^2\left(\sqrt{15}-\sqrt{14}\right)\left(4-\sqrt{15}\right)^2\left(6-\sqrt{35},\right). \end{aligned}$$

Remarkably,

$$k_{100} := \left(\left(3-2\sqrt{2}\right)\left(2+\sqrt{5}\right)\left(-3+\sqrt{10}\right)\left(-\sqrt{2}+\sqrt[4]{5}\right)^2\right)^2$$

arose in Bornemann's solution to Trefethen's 10th problem, [109]: the probability that a Brownian motion starting at the center of a 10×1 box hits the

ends first is $2/\pi \arcsin(k_{100})$. Ramanujan also noticed that the invariants G_N and g_N defined next are often simpler

$$G_N^{-12} := 2k_N k'_N \text{ and } g_N^{-12} := 2k_N/k'^2_N.$$

Note that each of these two latter invariants provides a quadratic formula for k_N. We also need *Ramanujan's invariant of the second kind*

$$\alpha_N := \frac{1/\pi - q\theta_4'(q)/\theta_4(q)}{\theta_3^4(q)} \qquad q := e^{-\pi\sqrt{N}} \tag{2.31}$$

which is also algebraic for integer N, [46]. In the form we have given them all the coefficients are very simple to compute numerically. Hence integer relation methods are easy to apply.

Example 6.1. *(Determining Invariants) The following Maple code produces 20 digits of each of our invariants:*

```
Digits:=16:que:=N->exp(-Pi*sqrt(N)):
kk:=q->(JacobiTheta2(0,q)/JacobiTheta3(0,q))^2:
kc:=q->(JacobiTheta4(0,q)/JacobiTheta3(0,q))^2: k:=kk@que:
l:=kc@que: G:=1/(2*k*l): g:=2*k/l^2:
alpha:=r->(subs(q=exp(-Pi*sqrt(r)),
  (1/Pi-sqrt(r)*4*(q*diff(JacobiTheta4(0,q),q)/JacobiTheta4(0,q)))/
  JacobiTheta3(0,q)^4)):
  a0:=N->(alpha(N)-sqrt(N)*k(N)^2): a1:=N->sqrt(N)*(1-2*k(N)^2):
  b0:=N->alpha(N)/(1-k(N)^2): b1:=N->sqrt(N)*(1+k(N)^2)/(1-k(N)^2):
```

We first explore use of Maple's identify function. Entering

```
> for n to 6 do n,identify(evalf[20](k((n)))) od;
```

returns

$$1/2\sqrt{2},\ -1+\sqrt{2},\ 1/4\sqrt{6}-1/4\sqrt{2},\ 3-2\sqrt{2},$$

$$0.11887694580260010118,\ 0.085164233174742587643.$$

where we have used only the simplest parameter-free version of the "identify" function. Correspondingly

```
> for n to 8 do latex(identify(evalf[20](G((2*n-1))))) od;
```

returns for the first 8 odd values of G_N^{-12}:

$$1,\ 2,\ 2+\sqrt{5},\ 87+4\sqrt{3},\ 4/3\sqrt[3]{199+3\sqrt{33}}+\frac{136}{3}\frac{1}{\sqrt[3]{199+3\sqrt{33}}}+\frac{22}{3},$$

$$18+5\sqrt{13},\ 28+12\sqrt{5}$$

and

```
> for n to 8 do latex(identify(evalf[20](g((2*n))))) od;
```

returns for the first 5 even values of g_N^{-12}:

$$1,\ 1/4\,\sqrt{2},\ 3-2\,\sqrt{2},\ 1/4\,\sqrt{-14+10\,\sqrt{2}},\ 9-4\,\sqrt{5}$$

but fails on the next three.

$$0.034675177060507381314, 0.022419012334044683484, 0.014940167059400883091$$

This can be remedied in many ways. For example,

```
>_EnvExplicit:=true:
   (PolynomialTools[MinimalPolynomial](g(14)^(1/3),4));
   solve(%)[2]; evalf(%/g(14)^(1/3));
```

yields $1-2X-5X^2-2X^3+X^4$ as the polynomial potentially satisfied by g_{14}^{-4}; and then extracts the correct radical

$$1/2+\sqrt{2}-1/2\,\sqrt{5+4\,\sqrt{2}}$$

which is confirmed to 15 places. One may check that $(g_{14}^6+g_{14}^{-6})/2=\sqrt{2}+1$ is an even simpler invariant. Similarly,

```
>_EnvExplicit:=true:(PolynomialTools[MinimalPolynomial]
  (G(25)^(1/12),4));
```

illustrates that G_{25} solves $x^2-x-1=0$ and so is the golden mean, and also shows that the appropriate power of G_N, g_N varies with N. Armed with these tools a fine challenge is to obtain all values of G_N, g_N or k_N up to, say, $N=50$. □

We may now record two families of series of which Ramanujan discovered many cases:

Theorem 6.2. (**Ramanujan-type Series**, [46, p. 182])
(a) For $N>1$

$$\frac{1}{\pi}=\sum_{n=0}^{\infty} r_n^3\,\left\{(\alpha_N-\sqrt{n}\,k_N^2)+n\,\sqrt{N}(k'^2_N-k_N^2)\right\}\,\left(G_N^{-12}\right)^{2n}. \qquad (2.32)$$

(b) For $N\geq 1$

$$\frac{1}{\pi}=\sum_{n=0}^{\infty}(-1)^n\,r_n^3\,\left\{\alpha_N\,k'^{-2}_N+n\,\sqrt{N}\,\frac{1+k_N^2}{1-k_N^2}\right\}\,\left(g_N^{-12}\right)^{2n} \qquad (2.33)$$

Example 6.3. *(Identifying our series) We shall now try to determine which cases of Theorem 6.2 we have recovered.*

Crude code to determine the coefficients is:

```
A:=proc() local N; N:= args[1];
     if nargs >1 then Digits :=args[2] fi;
  identify(evalf(G(N))),
  identify(evalf(a0(N)))+'n'*identify(evalf(a1(N))) end:
B:=proc() local N;
     N:=args[1];if nargs>1 then Digits:=args[2];fi;
  identify(evalf(g(N))),
  identify(evalf(b0(N)))+'n'*identify(evalf(b1(N))):end:
```

for the non-alternating and alternating cases respectively. For example $B(1)$ returns $\sqrt{2}, 3n + 1/4$ which means that

$$\frac{2}{\pi} = \sum_{n=0}^{\infty} (-1)^n \, r(n)^3 (1 + 4n).$$

We leave it to the reader to see that we had recovered the cases $N = 3, 5, 7, 15, 25$ and $N = 4, 6, 10, 18$ of Theorem 6.2 (b). □

Chapter 3

Algorithms for Experimental Mathematics, Part Two

1 The Wilf-Zeilberger Algorithm

One fascinating non-numerical algorithm is the Wilf-Zeilberger (WZ) algorithm, which employs "creative telescoping" to show that a sum (with either finitely or infinitely many terms) is zero. Below is an example of a WZ proof of $(1+1)^n = 2^n$. This proof is from Doron Zeilberger's original *Maple* program, which in turn is inspired by the proof in [235].

Let $F(n,k) = \left(\begin{array}{c} n \\ k \end{array} \right) 2^{-n}$. We wish to show that $L(n) = \sum_k F(n,k) = 1$ for every n. To this end, we construct, using the WZ algorithm, the function

$$G(n,k) \quad = \quad \frac{-1}{2^{(n+1)}} \left(\begin{array}{c} n \\ k-1 \end{array} \right) \left(= \frac{-k}{2(n-k+1)} F(n,k) \right), \qquad (3.1)$$

and observe that

$$F(n+1,k) - F(n,k) \quad = \quad G(n,k+1) - G(n,k). \qquad (3.2)$$

By applying the obvious telescoping property of these functions, we can write

$$\begin{aligned} \sum_k F(n+1,k) - \sum_k F(n,k) \quad &= \quad \sum_k (G(n,k+1) - G(n,k)) \\ &= \quad 0, \end{aligned}$$

which establishes that $L(n+1) - L(n) = 0$. The fact that $L(0) = 1$ follows from the fact that $F(0,0) = 1$ and 0 otherwise.

As an example, we will briefly present here the proof of the identities

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{\binom{4n}{2n}\binom{2n}{n}^4}{2^{16n}}\left(120n^2+34n+3\right) &= \frac{32}{n^2}\\ \sum_{n=0}^{\infty} \frac{(-1)^n\binom{2n}{n}^5}{2^{20n}}\left(820n^2+180n+13\right) &= \frac{128}{\pi^2},\end{aligned}$$

which, as we mentioned in the previous chapter, can be discovered by a PSLQ-based search strategy.

Guillera started by defining

$$G(n,k) = \frac{(-1)^k}{2^{16n}2^{4k}}\left(120n^2+84nk+34n+10k+3\right)\frac{\binom{2n}{n}^4\binom{2k}{k}^3\binom{4n-2k}{2n-k}}{\binom{2n}{k}\binom{n+k}{n}^2} \tag{3.3}$$

He then used the software package EKHAD, which implements the WZ method, obtaining the companion formula

$$F(n,k) = \frac{(-1)^k 512}{2^{16n}2^{4k}}\frac{n^3}{4n-2k-1}\frac{\binom{2n}{n}^4\binom{2k}{k}^3\binom{4n-2k}{2n-k}}{\binom{2n}{k}\binom{n+k}{n}^2} \tag{3.4}$$

When we define

$$H(n,k) = F(n+1,n+k)+G(n,n+k),$$

Zeilberger's theorem gives the identity

$$\sum_{n=0}^{\infty} G(n,0) = \sum_{n=0}^{\infty} H(n,0),$$

which when written out is

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{\binom{2n}{n}^4\binom{4n}{2n}}{2^{16n}}\left(120n^2+34n+3\right) &= \sum_{n=0}^{\infty}\frac{(-1)^n}{2^{20n+7}}\frac{(n+1)^3}{2n+3}\frac{\binom{2n+2}{n+1}^4\binom{2n}{n}^3\binom{2n+4}{n+2}}{\binom{2n+2}{n}\binom{2n+1}{n+1}^2}\\ &\quad+\sum_{n=0}^{\infty}\frac{(-1)^n}{2^{20n}}\left(204n^2+44n+3\right)\binom{2n}{n}^5\\ &= \frac{1}{4}\sum_{n=0}^{\infty}\frac{(-1)^n\binom{2n}{n}^5}{2^{20n}}\left(820n^2+180n+13\right)\end{aligned}$$

after considerable algebra.

Guillera then observes that since $\sum_{n\geq 0} G(n,k) = \sum_{n\geq 0} G(n,k+1)$, then by Carlson's theorem [23] it follows that $\sum_{n\geq 0} G(n,k) = A$ for some A, independent of k, even if k is not an integer. We then note that $0 < G(n,t) \leq 8^{-n}$, so one can interchange limit and sum to conclude that

$$\lim_{t\to 1/2} \sum_{n=1}^{\infty} \mathrm{Re}\,[G(n,t)] \;=\; 0.$$

Thus,

$$A \;=\; \lim_{t\to 1/2} \mathrm{Re}\,[G(0,t)] \;=\; \frac{32}{\pi^2},$$

and we have

$$\sum_{n=0}^{\infty} G(n,k) \;=\; \sum_{n=0}^{\infty} H(n,k) \;=\; \frac{32}{\pi^2}.$$

Guillera's two results follow immediately.

Obviously, this proof does not provide much insight, since the difficult part of the result is buried in the construction of (3.3). In other words, the W-Z method provides "proofs," but these "proofs" tend to be relatively unenlightening. Nonetheless, the very general nature of this scheme is of interest. It possibly presages a future in which a wide class of such identities can be "proved" automatically in a computer algebra system. Details and additional applications of this algorithm are given in [235].

2 Prime Number Computations

It is well known that there is a connection between prime numbers and the Riemann zeta function [44]. Prime numbers crop up in numerous other arenas of mathematical research, and often even in commercial applications, with the rise of RSA-based encryption methods on the Internet. Inasmuch as this research topic is certain to be of great interest for the foreseeable future, we mention here some of the techniques for counting, generating, and testing prime numbers.

The prime counting function $\pi(x)$

$$\pi(x) \;=\; \#\{\text{primes} \leq x\}$$

is of central interest in this research. Table 1 gives $\pi(x)$ for power-of-ten arguments up to 10^{22}. This data was obtained from Eric Weisstein's "World of Mathematics" web site `http://mathworld.wolfram.com`.

It is clear from even a cursory glance at Table 3.1 that the researchers who have produced these counts are not literally testing every integer up to 10^{22} for primality—that would require much more computation than the combined power of all computers worldwide, even using the best known methods to test individual primes. Indeed, some very sophisticated techniques have been employed, which unfortunately are too technical to be presented in detail here. We refer interested readers to the discussion of this topic in the new book *Prime Numbers: A Computational Perspective* by Richard Crandall and Carl Pomerance [90, 140-150]. Readers who wish to informally explore the behavior of $\pi(x)$ may use a sieving algorithm, which is a variant of a scheme originally presented by Eratosthenes of Cyrene about 200 BCE [90, pg. 114].

If one does not require certainty, but only high probability that a number is prime, some very efficient probabilistic primality tests have been discovered in the past few decades. In fact, these schemes are now rou-

x	$\pi(x)$	$\int_2^x dt/\log t$	Difference
10^1	4	5	1
10^2	25	29	4
10^3	168	177	9
10^4	1229	1245	16
10^5	9592	9629	37
10^6	78498	78627	129
10^7	6 64579	6 64917	338
10^8	57 61455	57 62208	753
10^9	508 47534	508 49234	1700
10^{10}	4550 52511	4550 55614	3103
10^{11}	41180 54813	41180 66400	11587
10^{12}	3 76079 12018	3 76079 50280	38262
10^{13}	34 60655 36839	34 60656 45809	1 08970
10^{14}	320 49417 50802	320 49420 65691	3 14889
10^{15}	2984 45704 22669	2984 45714 75287	10 52618
10^{16}	27923 83410 33925	27923 83442 48556	32 14631
10^{17}	2 62355 71576 54233	2 62355 71656 10821	79 56588
10^{18}	24 73995 42877 40860	24 73995 43096 90414	219 49554
10^{19}	234 05766 72763 44607	234 05766 73762 22381	998 77774
10^{20}	2220 81960 25609 18840	2220 81960 27836 63483	2227 44643
10^{21}	21127 26948 60187 31928	21127 26948 66161 26181	5973 94253
10^{22}	2 01467 28668 93159 06290	2 01467 28669 12482 61497	19323 55207

Table 3.1. The prime-counting function $\pi(x)$ and Gauss' approximation.

tinely used to generate primes for RSA encryption in Internet commerce. When you type in your Visa or Mastercard number in a secure web site to purchase a book or computer accessory, somewhere in the process it is quite likely that two large prime numbers have been generated, which were certified as prime using one of these schemes.

The most widely used probabilistic primality test is the following, which was originally suggested by Artjuhov in 1966, although it was not appreciated until it was rediscovered and popularized by Selfridge in the 1970s [90].

Algorithm 2.1. Strong probable prime test.
Given an integer $n = 1 + 2^s t$, for integers s and t (and t odd), select an integer a by means of a pseudorandom number generator in the range $1 < a < n - 1$.

1. Compute $b := a^t \bmod n$ using the binary algorithm for exponentiation (see Algorithm 3.2 in Chapter 3 of [44]). If $b = 1$ or $b = n - 1$ then exit (n is a strong probable prime base a).

2. For $j = 1$ to $s - 1$ do: Compute $b : b^2 \bmod n$; if $(b = n - 1)$ then exit (n is a strong probable prime base a).

3. Exit: n is composite.

This test can be repeated several times with different pseudo-randomly chosen a. In 1980 Monier and Rabin independently showed that an integer n that passes the test as a strong probable prime is prime with probability at least 3/4, so that m tests increase this probability to $1-1/4^m$ [178, 203]. In fact, for large test integers n, the probability is even closer to unity. Damgard, Landrock and Pomerance showed in 1993 that if n has k bits, then this probability is greater than $1 - k^2 4^{2-\sqrt{k}}$, and for certain k is even higher [91]. For instance, if n has 500 bits, then this probability is greater than $1 - 1/4^{28m}$. Thus a 500-bit integer that passes this test even once is prime with prohibitively safe odds—the chance of a false declaration of primality is less than one part in Avogadro's number (6×10^{23}). If it passes the test for four pseudo-randomly chosen integers a, then the chance of false declaration of primality is less than one part in a googol (10^{100}). Such probabilities are many orders of magnitude more remote than the chance that an undetected hardware or software error has occurred in the computation.

A number of more advanced probabilistic primality testing algorithms are now known. The current state-of-the-art is that such tests can determine the primality of integers with hundreds to thousands of digits. Additional details of these schemes are available in [90].

For these reasons, probabilistic primality tests are considered entirely

satisfactory for practical use, even for applications such as large interbank financial transactions, which have extremely high security requirements. Nonetheless, mathematicians have long sought tests that remove this last iota of uncertainty, yielding a mathematically rigorous certificate of primality. Indeed, the question of whether there exists a "polynomial time" primality test has long stood as an important unsolved question in pure mathematics.

Thus it was with considerable elation that such an algorithm was recently discovered, by Manindra Agrawal, Neeraj Kayal, and Nitin Saxena (initials AKS) of the Indian Institute of Technology in Kanpur, India [4]. Their discovery sparked worldwide interest, including a prominent report in the *New York Times* [207]. Since the initial report in August 2002, several improvements have been made. Readers are referred to a variant of the original algorithm due to Lenstra [165], as implemented by Richard Crandall and Jason Papadopoulos [89][45, pg xx].

3 Roots of Polynomials

In Chapter 1, we showed how a relatively simple scheme involving Newton iterations can be used to compute high-precision square roots and even to perform high-precision division. This Newton iteration scheme is, in fact, quite general and can be used to solve many kinds of equations, both algebraic and transcendental. One particularly useful application, frequently encountered by experimental mathematicians, is to find roots of polynomials. This is done by using a careful implementation of the well-known version of Newton's iteration

$$x_{k+1} \quad = \quad x_k - \frac{p(x)}{p'(x)}, \tag{3.5}$$

where $p'(x)$ denotes the derivative of $p(x)$. As before, this scheme is most efficient if it employs a level of numeric precision that starts with ordinary double precision (16-digit) or double-double precision (32-digit) arithmetic until convergence is achieved at this level, then approximately doubles with each iteration until the final level of precision is attained. One additional iteration at the final or penultimate precision level may be needed to insure full accuracy.

Note that Newton's iteration can be performed, as written in (3.5), with either real or complex arithmetic, so that complex roots of polynomials (with real or complex coefficients) can be found almost as easily as real roots. Evaluation of the polynomials $p(x)$ and $p'(x)$ is most efficiently performed using Horner's rule: For example, the polynomial $p(x) = p_0 + p_1x + p_2x^2 + p_3x^3 + p_4x^4 + p_5x^5$ is evaluated as $p(x) = p_0 + x(p_1 + x(p_2 + x(p_3 + x(p_4 + xp_5))))$.

There are two issues that arise here that do not arise with the Newton iteration schemes for division and square root. The first is the selection of the starting value—if it is not close to the desired root, then successive iterations may jump far away. If you have no idea where the roots are (or how accurate the starting value must be), then a typical strategy is to try numerous starting values, covering a wide range of likely values, and then make an inventory of the approximate roots that are found. If you are searching for complex roots, note that it is often necessary to use a two-dimensional array of starting values. These "exploratory" iterations can be done quite rapidly, since typically only a modest numeric precision is required—in almost all cases, just ordinary double precision (16 digits) or double-double precision (32 digits) arithmetic will do. Once the roots have been located in this fashion, then the full-fledged Newton scheme can be used to produce their precise high-precision values.

The second issue is how to handle repeated roots. The difficulty here is that, in such cases, convergence to the root is very slow, and instabilities may throw the search far from the root. In these instances, note that we can write $p(x) = q^2(x)r(x)$, where r has no repeated roots (if all roots are repeated, then $r(x) = 1$). Now note that $p'(x) = 2q(x)r(x) + q^2(x)r'(x) = q(x)[2r(x) + q(x)r'(x)]$. This means that if $p(x)$ has repeated roots, then these roots are also roots of $p'(x)$, and, conversely, if $p(x)$ and $p'(x)$ have a common factor, then the roots of this common factor are repeated roots of $p(x)$. This greatest common divisor polynomial $q(x)$ can be found by performing the Euclidean algorithm (in the ring of polynomials) on $p(x)$ and $p'(x)$. The Newton iteration scheme can then be applied to find the roots of both $q(x)$ and $r(x)$. It is possible, of course, that $q(x)$ also has repeated roots, but recursive application of this scheme quickly yields all individual roots.

In the previous paragraph, we mentioned the possible need to perform the Euclidean algorithm on two polynomials, which involves polynomial multiplication and division. For modest-degree polynomials, a simple implementation of the schemes learned in high school algebra suffices—just represent the polynomials as strings of high-precision numbers. For high-degree polynomials, polynomial multiplication can be accelerated by utilizing fast Fourier transforms and a convolution scheme that is almost identical to the scheme, mentioned in Chapter 1, to perform high-precision multiplication. High-degree polynomial division can be accelerated by a Newton iteration scheme, similar to that mentioned above for high-precision division. See [90] for additional details on high-speed polynomial arithmetic.

We noted above that if the starting value is not quite close to the desired root, then successive Newton iterations may jump far from the root, and eventually converge to a different root than the one desired. In general, suppose we are given a degree-n polynomial $p(x)$ with m distinct complex

roots r_k (some may be repeated roots). Define the function $Q_p(z)$ as the limit achieved by successive Newton iterations that start at the complex number z; if no limit is achieved, then set $Q_p(z)\infty$. Then the m sets $\{z : Q_p(z) = r_k\}$ for $k = 1, 2 \cdots, m$ constitute a partition of the complex plane, except for a filamentary set of measure zero that separates the m sets. In fact, each of these m sets is itself an infinite collection of disconnected components.

The collection of these Newton-Julia sets and their boundaries form pictures of striking beauty, and are actually quite useful in gaining insight on both the root structure of the original polynomial and the behavior of Newton iteration solutions. Some of the most interesting graphics of this type are color-coded plots of the function $N_p(z)$, which is the number of iterations required for convergence (to some accuracy ϵ) of Newton's iteration for $p(x)$, beginning at z (if the Newton iteration does not converge at z, then set $N_p(z) = \infty$). A plot for the cubic polynomial $p(x) = x^3 - 1$ is shown in Figure 3.1.

4 Numerical Quadrature

Experimental mathematicians very frequently find it necessary to calculate definite integrals to high precision. Recall the examples given in Chapter 2, wherein we were able to experimentally identify certain definite integrals as analytic expressions, based only on their high-precision numerical value.

As one example, recently one of the authors (Borwein), together with Greg Fee of Simon Fraser University in Canada, were inspired by a recent problem in the *American Mathematical Monthly* [5]. They found by using a numerical integration scheme described in Chapter 2, together with a PSLQ integer relation detection program, that if $C(a)$ is defined by

$$C(a) = \int_0^1 \frac{\arctan(\sqrt{x^2+a^2})\,dx}{\sqrt{x^2+a^2}(x^2+1)},$$

then

$$\begin{aligned} C(0) &= \pi \log 2/8 + G/2 \\ C(1) &= \pi/4 - \pi\sqrt{2}/2 + 3\sqrt{2}\arctan(\sqrt{2})/2 \\ C(\sqrt{2}) &= 5\pi^2/96, \end{aligned}$$

where $G = \sum_{k\geq 0}(-1)^k/(2k+1)^2$ is Catalan's constant (the third of these results is the result from the *Monthly*). These experimental results then led to the following general result, rigorously established, among others:

$$\int_0^\infty \frac{\arctan(\sqrt{x^2+a^2})\,dx}{\sqrt{x^2+a^2}(x^2+1)} = \frac{\pi}{2\sqrt{a^2-1}}\left[2\arctan(\sqrt{a^2-1}) - \arctan(\sqrt{a^4-1})\right].$$

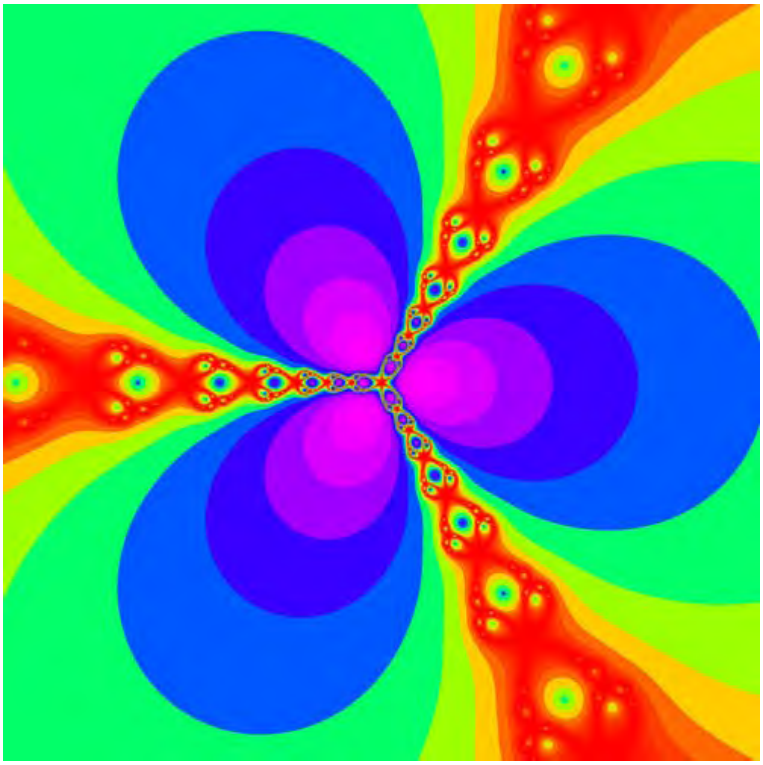


Figure 3.1. Newton-Julia set for $p(x) = x^3 - 1$.

As a second example, two of the present authors (Bailey and Borwein) empirically determined that

$$\begin{aligned}\frac{2}{\sqrt{3}}\int_0^1 \frac{\log^6(x)\arctan[x\sqrt{3}/(x-2)]}{x+1}\,dx \quad &= \quad \frac{1}{81648}\,[-229635L_3(8) \\ &+29852550L_3(7)\log 3 - 1632960L_3(6)\pi^2 + 27760320L_3(5)\zeta(3) \\ &-275184L_3(4)\pi^4 + 36288000L_3(3)\zeta(5) - 30008L_3(2)\pi^6 \\ &-57030120L_3(1)\zeta(7)]\,,\end{aligned}$$

where $L_3(s) = \sum_{n=1}^{\infty}[1/(3n-2)^s - 1/(3n-1)^s]$. General results have been conjectured but not yet rigorously established.

A third result is the following, which was found by one of the present

authors (Borwein) and British physicist David Broadhurst [53]:

$$\frac{24}{7\sqrt{7}} \int_{\pi/3}^{\pi/2} \log \left| \frac{\tan t + \sqrt{7}}{\tan t - \sqrt{7}} \right| dt \quad \stackrel{?}{=} \quad L_{-7}(2) \; = \tag{3.6}$$

$$\sum_{n=0}^{\infty} \left[\frac{1}{(7n+1)^2} + \frac{1}{(7n+2)^2} - \frac{1}{(7n+3)^2} + \frac{1}{(7n+4)^2} - \frac{1}{(7n+5)^2} - \frac{1}{(7n+6)^2} \right].$$

This integral arose out of some studies in quantum field theory, in analysis of the volume of ideal tetrahedra in hyperbolic space. It is the simplest of 998 empirically determined cases where the volume of a hyperbolic knot complement is expressible in terms of an L-series and an apparently unexpected integral or sum [53]. The question mark is used here because although this identity has been numerically verified to 20,000-digit precision (see Chapter 2), as of this date no proof is yet known.

PSLQ computations were also able to recover relations among integrals of this type. Let I_n be the definite integral of (3.6), except with limits $n\pi/24$ and $(n+1)\pi/24$. Then the relations

$$\begin{aligned} -2I_2 - 2I_3 - 2I_4 - 2I_5 + I_8 + I_9 - I_{10} - I_{11} \quad &\stackrel{?}{=} \quad 0, \\ I_2 + 3I_3 + 3I_4 + 3I_5 + 2I_6 + 2I_7 - 3I_8 - I_9 \quad &\stackrel{?}{=} \quad 0 \end{aligned} \tag{3.7}$$

have been numerically discovered, although as before no hint of a proof is known.

In some cases, *Maple* or *Mathematica* is able to evaluate a definite integral analytically, but the resulting expressions are complicated and not very enlightening. For example, although the integrals

$$\begin{aligned} I_1 &= \int_0^1 \frac{t^2 \log(t)\, dt}{(t^2-1)(t^4+1)} \\ I_2 &= \int_0^{\pi/4} \frac{t^2\, dt}{\sin^2(t)} \\ I_3 &= \int_0^{\pi} \frac{x \sin x\, dx}{1+\cos^2 x} \end{aligned}$$

are successfully evaluated by *Maple* and *Mathematica*, the results are somewhat lengthy expressions involving advanced functions and complex entities. We suspect that there are considerably simpler closed-form versions of these integrals. Indeed, we can obtain the following, based solely on the high-precision numerical values of these integrals, combined with integer relation computations:

$$\begin{aligned} I_1 &= \pi^2(2-\sqrt{2})/32 \\ I_2 &= -\pi^2/16 + \pi \log(2)/4 + G \\ I_3 &= \pi^2/4. \end{aligned}$$

The commercial packages *Maple* and *Mathematica* both include rather good high-precision numerical quadrature facilities. However, these packages do have some limitations, and in many cases much faster performance can be achieved with custom-written programs. And in general it is beneficial to have some understanding of quadrature techniques, even if you rely on software packages to perform the actual computation.

We describe here three state-of-the-art, highly efficient techniques for numerical quadrature. Some additional refinements, not mentioned here, are described in [21]. You can try programming these schemes yourself, or you can refer to the C++ and Fortran-90 programs available at `http://www.experimentalmath.info`.

4.1 Tanh-Sinh Quadrature

The scheme we will discuss here is known as "tanh-sinh" quadrature. While it is not as efficient as Gaussian quadrature for continuous, bounded, well-behaved functions on finite intervals, it often produces highly accurate results even for functions with (integrable) singularities or vertical derivatives at one or both endpoints of the interval. In contrast, Gaussian quadrature typically performs very poorly in such instances. Also, the cost of computing abscissas and weights in tanh-sinh quadrature only scales as $p^2 \log p$ in tanh-sinh quadrature, where p is the number of correct digits in the result (for many problems), whereas the corresponding cost for Gaussian quadrature scales as $p^3 \log p$. For this reason, Gaussian quadrature is not practical beyond about 1,000 digits.

The tanh-sinh function quadrature scheme is based on the Euler-Maclaurin summation formula, which can be stated as follows [10, pg. 280]. Let $m \geq 0$ and $n \geq 1$ be integers, and define $h = (b-a)/n$ and $x_j = a + jh$ for $0 \leq j \leq n$. Further, assume that the function $f(x)$ is at least $(2m+2)$-times continuously differentiable on $[a, b]$. Then

$$\begin{aligned} \int_a^b f(x)\,dx &= h\sum_{j=0}^{n} f(x_j) - \frac{h}{2}\left(f(a)+f(b)\right) \\ &\quad -\sum_{i=1}^{m} \frac{h^{2i}B_{2i}}{(2i)!}\left(f^{(2i-1)}(b) - f^{(2i-1)}(a)\right) + E(h,m), \end{aligned}$$

where B_{2i} denote the Bernoulli numbers, and

$$E(h,m) = \frac{h^{2m+2}(b-a)B_{2m+2}f^{(2m+2)}(\xi)}{(2m+2)!}, \tag{3.8}$$

for some $\xi \in (a, b)$.

In the circumstance where the function $f(x)$ and all of its derivatives are zero at the endpoints a and b, the second and third terms of the Euler-Maclaurin formula are zero. Thus the error in a simple step-function approximation to the integral, with interval h, is simply $E(h, m)$. But since E is then less than a constant times $h^{2m+2}/(2m+2)!$, for any m, we conclude that the error goes to zero more rapidly than any power of h. In the case of a function defined on $(-\infty, \infty)$, the Euler-Maclaurin summation formula still applies to the resulting doubly infinite sum approximation, provided as before that the function and all of its derivatives tend to zero for large positive and negative arguments.

This principle is utilized in the error function and tanh-sinh quadrature scheme (see the next subsection) by transforming the integral of $f(x)$ on a finite interval, which we will take to be $(-1, 1)$ for convenience, to an integral on $(-\infty, \infty)$ using the change of variable $x = g(t)$. Here $g(x)$ is some monotonic function with the property that $g(x) \to 1$ as $x \to \infty$, and $g(x) \to -1$ as $x \to -\infty$, and also with the property that $g'(x)$ and all higher derivatives rapidly approach zero for large arguments. In this case we can write, for $h > 0$,

$$\int_{-1}^{1} f(x)\,dx \;=\; \int_{-\infty}^{\infty} f(g(t))g'(t)\,dt \;\approx\; h\sum_{-\infty}^{\infty} w_j f(x_j), \qquad (3.9)$$

where $x_j = g(hj)$ and $w_j = g'(hj)$. If the convergence of $g'(t)$ and its derivatives to zero is sufficiently rapid for large $|t|$, then even in cases where $f(x)$ has a vertical derivative or an integrable singularity at one or both endpoints, the resulting integrand $f(g(t))g'(t)$ will be a smooth bell-shaped function for which the Euler-Maclaurin summation formula applies, as described above. In such cases we have that the error in the above approximation decreases faster than any power of h. The summation above is typically carried out to limits $(-N, N)$, beyond which the terms of the summand are less than the "epsilon" of the multiprecision arithmetic being used.

The tanh-sinh scheme employs the transformation $x = \tanh(\pi/2 \cdot \sinh t)$, where $\sinh t = (e^t - e^{-t})/2$, $\cosh t(e^t + e^{-t})/2$, and $\tanh t = \sinh t/\cosh t$ (alternately, one can omit the $\pi/2$ in this definition). This transformation converts an integral on $(-1, 1)$ to an integral on the entire real line, which can then be approximated by means of a simple step-function summation. In this case, by differentiating the transformation, we obtain the abscissas x_k and the weights w_k as

$$\begin{aligned} x_j &= \tanh[\pi/2 \cdot \sinh(jh)] \\ w_j &= \frac{\pi/2 \cdot \cosh(jh)}{\cosh^2[\pi/2 \cdot \sinh(jh)]}. \end{aligned} \qquad (3.10)$$

Note that these functions involved here are compound exponential, so, for example, the weights w_j converge very rapidly to zero. As a result, the tanh-sinh quadrature scheme is often very effective in dealing with singularities or infinite derivatives at endpoints.

Algorithm 2.3. tanh-sinh quadrature.

Initialize: Set $h := 2^{-m}$.
For $k := 0$ to $20 \cdot 2^m$ do:
Set $t := kh$, $x_k := \tanh(\pi/2 \cdot \sinh t)$ and $w_k := \pi/2 \cdot \cosh t / \cosh^2(\pi/2 \cdot \sinh t)$;
If $|x_k - 1| < \epsilon$ then exit do; enddo.
Set $n_t = k$ (the value of k at exit).

Perform quadrature for a function $f(x)$ on $(-1, 1)$:
Set $S := 0$ and $h := 1$.
For $k := 1$ to m (or until successive values of S are identical to within ϵ) do:
$h := h/2$.
For $i := 0$ to n_t step 2^{m-k} do:
If $(\mathrm{mod}(i, 2^{m-k+1}) \neq 0$ or $k = 1)$ then
If $i = 0$ then $S := S + w_0 f(0)$ else $S := S + w_i(f(-x_i) + f(x_i))$ endif.
endif; enddo; endo.
Result $= hS$.

There are several similar quadrature schemes that can be defined, also based on the Euler-Maclaurin summation formula. For example, using $g(t)\mathrm{erf}t$ gives rise to "error function" or "erf" quadrature; using $g(t) = \sinh(\sinh t)$ gives rise to "sinh-sinh" quadrature, which can be used for functions defined on the entire real line. Additional refinements are described in [21]. An implementation of this scheme on a highly parallel computer system is described in [14].

4.2 Practical Considerations for Quadrature

The tanh-sinh scheme have assumed a function of one variable defined and continuous on the interval $(-1, 1)$. Integrals on other finite intervals (a, b) can be found by applying a linear change of variable:

$$\int_a^b f(t)\,dt \quad = \quad \frac{b-a}{2} \int_{-1}^{1} f\left(\frac{b+a}{2} + \frac{b-a}{2}x\right) dx. \qquad (3.11)$$

Note also that integrable functions on an infinite interval can, in a similar manner, be reduced to an integral on a finite interval, for example:

$$\int_0^\infty f(t)\,dt \quad = \quad \int_0^1 [f(x) + f(1/x)/x^2]\,dx. \tag{3.12}$$

Integrals of functions with singularities (such as "corners" or step discontinuities) within the integration interval (i.e., not at the endpoints) should be broken into separate integrals.

The above algorithm statements each suggest increasing the level of the quadrature (the value of k) until two successive levels give the same value of S, to within some tolerance ϵ. While this is certainly a reliable termination test, it is often possible to stop the calculation earlier, with significant savings in runtime, by means of making reasonable projections of the current error level. In this regard, the authors have found the following scheme to be fairly reliable: Let S_1, S_2, and S_3 be the value of S at the current level, the previous level, and two levels back, respectively. Then set $D_1 := \log_{10}|S_1 - S_2|$, $D_2 := \log_{10}|S_1 - S_3|$, and $D_3 := \log_{10}\epsilon - 1$. Now we can estimate the error E at level $k > 2$ as 10^{D_4}, where $D_4 = \min(0, \max(D_1^2/D_2, 2D_1, D_3))$. These estimation calculations may be performed using ordinary double precision arithmetic.

In the case of tanh-sinh quadrature (or other schemes based on the Euler-Maclaurin summation formula as above), the following formula provides a more accurate estimate of the error term. Let $F(t)$ be the desired integrand function, and then define $f(t) = F(g(t))g'(t)$, where $g(t) = \tanh(\sinh t)$ (or one could use any of the other g functions mentioned above). Then consider

$$E_2(h,m) \quad = \quad h(-1)^{m-1}\left(\frac{h}{2\pi}\right)^{2m} \sum_{j=a/h}^{b/h} D^{2m} f(jh). \tag{3.13}$$

In many cases, this estimate of the error term is exceedingly accurate once h is moderately small, for any integer $m \geq 1$. In most cases, it is sufficient to use $E_2(h,1)$, although higher-order estimates can be used with the bound

$$|E(h,m) - E_2(h,m)| \quad \leq$$
$$2\left[\zeta(2m) + (-1)^m\zeta(2m+2)\right]\left(\frac{h}{2\pi}\right)^{2m} h\sqrt{\int_a^b |D^{2m} f(t)|^2\,dt},$$

to yield rigorous "certificates" of quadrature results. See [12] for details.

Some refinements to these schemes, and to the error estimation procedure above, are described in [21, 12]. Gaussian quadrature and tanh-sinh have been implemented in C++ and Fortran-90 programs, available at `http://www.experimentalmath.info`.

4.3 Higher-Dimensional Quadrature

The tanh-sinh quadrature scheme can be easily generalized to perform two-dimensional (2-D) and three-dimensional (3-D) quadrature. Runtimes are typically many times higher than with one-dimensional (1-D) integrals. However, if one is content with, say, 32-digit or 64-digit results (by using double-double or quad-double arithmetic, respectively), then many two-variable functions can be integrated in reasonable run time (say, a few minutes). One advantage that these schemes have is that they are very well suited to parallel processing. Thus even several-hundred digit values can be obtained for 2-D and 3-D integrals if one can utilize a highly parallel computer, such as a "Beowulf" cluster. One can even envision harnessing many computers on a geographically distributed grid for such a task, although the authors are not aware of any such attempts yet.

One sample computation of this sort, performed by one of the present authors, produced the following evaluation:

$$\int_{-1}^{1}\int_{-1}^{1} \frac{dx\,dy}{\sqrt{1+x^2+y^2}} \quad = \quad 4\log(2+\sqrt{3}) - \frac{2\pi}{3}. \tag{3.14}$$

5 Infinite Series Summation

We have already seen numerous examples in previous chapters of mathematical constants defined by infinite series. In experimental mathematics work, it is usually necessary to evaluate such constants to say several hundred digit accuracy. The commercial software packages *Maple* and *Mathematica* include quite good facilities for the numerical evaluation of series. However, as with numerical quadrature, these packages do have limitations, and in some cases better results can be obtained using custom-written computer code. In addition, even if one relies exclusively on these commercial packages, it is useful to have some idea of the sorts of operations that are being performed by such software.

Happily, in many cases of interest to the experimental mathematician, infinite series converge sufficiently rapidly that they can be numerically evaluated to high precision by simply evaluating the series directly as written, stopping the summation when the individual terms are smaller than the "epsilon" of the multiprecision arithmetic system being used. All of the BBP-type formulas, for instance, are of this category. But other types of infinite series formulas present considerable difficulties for high-precision evaluation. Two simple examples are Gregory's series for $\pi/4$ and a similar

series for Catalan's constant:

$$\begin{aligned} \pi &= 1 - 1/3 + 1/5 - 1/7 + \cdots \\ G &= 1 - 1/3^2 + 1/5^2 - 1/7^2 + \cdots . \end{aligned}$$

We describe here one technique that is useful in many such circumstances. In fact, we have already been introduced to it in an earlier section of this chapter: It is the Euler-Maclaurin summation formula. The Euler-Maclaurin formula can be written in somewhat different form than before, as follows [10, page 282]. Let $m \geq 0$ be an integer, and assume that the function $f(x)$ is at least $(2m+2)$-times continuously differentiable on $[a, \infty)$, and that $f(x)$ and all of its derivatives approach zero for large x. Then

$$\sum_{j=a}^{\infty} f(j) = \int_a^{\infty} f(x)\,dx + \frac{1}{2}f(a) - \sum_{i=1}^{m} \frac{B_{2i}}{(2i)!} f^{(2i-1)}(a) + E, \quad (3.15)$$

where B_{2i} denote the Bernoulli numbers, and

$$E = \frac{B_{2m+2} f^{(2m+2)}(\xi)}{(2m+2)!}, \quad (3.16)$$

for some $\xi \in (a, \infty)$.

This formula is not effective as written. The strategy is instead to evaluate a series manually for several hundred or several thousand terms, then to use the Euler-Maclaurin formula to evaluate the tail. Before giving an example, we need to describe how to calculate the Bernoulli numbers B_{2k}, which are required here. The simplest way to compute them is to recall that [3, page 807]

$$\zeta(2k) = \frac{(2\pi)^{2k} |B_{2k}|}{2(2k)!}, \quad (3.17)$$

which can be rewritten as

$$\frac{B_{2k}}{(2k)!} = \frac{2(-1)^{k+1} \zeta(2k)}{(2\pi)^{2k}}. \quad (3.18)$$

The Riemann zeta function at real arguments s can, in turn, be computed using the formula [65]

$$\zeta(s) = \frac{-1}{2^n (1 - 2^{1-s})} \sum_{j=0}^{2n-1} \frac{e_j}{(j+1)^s} + E_n(s), \quad (3.19)$$

where

$$e_j \quad = \quad (-1)^j \left(\sum_{k=0}^{j-n} \frac{n!}{k!(n-k)!} - 2^n \right) \tag{3.20}$$

(the summation is zero when its index range is null), and $|E_n(s)| < 1/(8^n|1-2^{1-s}|)$. This scheme is encapsulated in the following algorithm.

Algorithm 2.4. Zeta function evaluation.

Initialize: Set $n = P/3$, where P is the precision level in bits, and set $t_1 := -2^n,\ t_2 := 0,\ S := 0$, and $I = 1$.

For $j := 0$ to $2n-1$ do: If $j < n$ then $t_2 := 0$ elseif $j = n$ then $t_2 := 1$ else $t_2 := t_2 \cdot (2n-j+1)/(j-n)$ endif.
Set $t_1 := t_1 + t_2,\ S := S + I \cdot t_1/(j+1)^s$ and $I := -I$; enddo.
Return $\zeta(s) := -S/[2^n \cdot (1-2^{1-s})]$.

A more advanced method to compute the zeta function in the particular case of interest here, where we need the zeta function evaluated at all even integer arguments up to some level m, is described in [15].

We illustrate the above by calculating Catalan's constant using the Euler-Maclaurin formula. We can write

$$\begin{aligned} G &= (1-1/3^2) + (1/5^2 - 1/7^2) + (1/9^2 - 1/11^2) + \cdots \\ &= 8\sum_{k=0}^{\infty} \frac{2k+1}{(4k+1)^2(4k+3)^2} \\ &= 8\sum_{k=0}^{n} \frac{2k+1}{(4k+1)^2(4k+3)^2} + 8\sum_{k=n+1}^{\infty} \frac{2k+1}{(4k+1)^2(4k+3)^2} \\ &= 8\sum_{k=0}^{n} \frac{2k+1}{(4k+1)^2(4k+3)^2} + 8\int_{n+1}^{\infty} f(x)\,dx + 4f(n+1) \\ &\quad -8\sum_{i=1}^{m} \frac{B_{2i}}{(2i)!} f^{(2i-1)}(n+1) + 8E, \end{aligned}$$

where $f(x) = (2x+1)/[(4x+1)^2(4x+3)^2]$ and $|E| < 3/(2\pi)^{2m+2}$. Using $m = 20$ and $n = 1000$ in this formula, we obtain a value of G correct to 114 decimal digits. We presented the above scheme for Catalan's constant because it is illustrative of the Euler-Maclaurin method. However serious computation of Catalan's constant can be done more efficiently using the recently discovered BBP-type formula (given in Table 3.5 of [44]), Ramanujan's formula (given in Item 7 of Chapter 6 in [44]), or Bradley's formula (also given in Item 7 of Chapter 6 in [44]).

One less-than-ideal feature of the Euler-Maclaurin approach is that high-order derivatives are required. In many cases of interest, successive derivatives satisfy a fairly simple recursion and can thus be easily computed with an ordinary handwritten computer program. In other cases, these derivatives are sufficiently complicated that such calculations are more conveniently performed in a symbolic computing environment such as *Mathematica* or *Maple*. In a few applications of this approach, a combination of symbolic computation and custom-written numerical computation is required to produce results in reasonable runtime[16].

6 Apery-Like Summations

Here we present a detailed case study in identifying sums of a certain class of infinite series, by means of a multi-step approach that is broadly illustrative of the experimental methodology in mathematics. This origins of this work lay in the existence of infinite series formulas involving central binomial coefficients in the denominators for the constants $\zeta(2)$, $\zeta(3)$, and $\zeta(4)$. These formulas, as well the role of the formula for $\zeta(3)$ in Apéry's proof of its irrationality, have prompted considerable effort during the past few decades to extend these results to larger integer arguments. The formulas in question are

$$\zeta(2) = 3\sum_{k=1}^{\infty} \frac{1}{k^2 \binom{2k}{k}}, \tag{3.21}$$

$$\zeta(3) = \frac{5}{2}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3 \binom{2k}{k}}, \tag{3.22}$$

$$\zeta(4) = \frac{36}{17}\sum_{k=1}^{\infty} \frac{1}{k^4 \binom{2k}{k}}. \tag{3.23}$$

Identity (3.21) has been known since the 19th century, while (3.22) was variously discovered in the last century and (3.23) was noted by Comtet [87, 54, 224, p. 89].

These results led many to conjecture that the constant $\mathcal{Q}_5$ defined by the ratio

$$\mathcal{Q}_5 := \zeta(5) \Big/ \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^5 \binom{2k}{k}}$$

is rational, or at least algebraic. However, 10,000-digit PSLQ computations have established that if $\mathcal{Q}_5$ is a zero of a polynomial of degree at most 25

with integer coefficients, then the Euclidean norm of the vector of coefficients exceeds 1.24×10^{383}. Similar computations for $\zeta(5)$ have yielded a bound of 1.98×10^{380}. These computations lend credence to the belief that $\mathcal{Q}_5$ and $\zeta(5)$ are transcendental.

Given the negative result from PSLQ computations for $\mathcal{Q}_5$, the authors of [50] systematically investigated the possibility of a multi-term identity of this general form for $\zeta(2n+1)$. The following were recovered early in experimental searches using computer-based integer relation tools [50, 49]:

$$\zeta(5) = 2\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^5\binom{2k}{k}} - \frac{5}{2}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3\binom{2k}{k}} \sum_{j=1}^{k-1} \frac{1}{j^2}, \tag{3.24}$$

$$\zeta(7) = \frac{5}{2}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^7\binom{2k}{k}} + \frac{25}{2}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3\binom{2k}{k}} \sum_{j=1}^{k-1} \frac{1}{j^4} \tag{3.25}$$

$$\begin{aligned}\zeta(9) &= \frac{9}{4}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^9\binom{2k}{k}} - \frac{5}{4}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^7\binom{2k}{k}} \sum_{j=1}^{k-1} \frac{1}{j^2} + 5\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^5\binom{2k}{k}} \sum_{j=1}^{k-1} \frac{1}{j^4} \\ &+ \frac{45}{4}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3\binom{2k}{k}} \sum_{j=1}^{k-1} \frac{1}{j^6} - \frac{25}{4}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3\binom{2k}{k}} \sum_{j=1}^{k-1} \frac{1}{j^4} \sum_{j=1}^{k-1} \frac{1}{j^2},\end{aligned} \tag{3.26}$$

$$\begin{aligned}\zeta(11) &= \frac{5}{2}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^{11}\binom{2k}{k}} + \frac{25}{2}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^7\binom{2k}{k}} \sum_{j=1}^{k-1} \frac{1}{j^4} \\ &- \frac{75}{4}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3\binom{2k}{k}} \sum_{j=1}^{k-1} \frac{1}{j^8} + \frac{125}{4}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3\binom{2k}{k}} \sum_{j=1}^{k-1} \frac{1}{j^4} \sum_{i=1}^{k-1} \frac{1}{i^4}.\end{aligned} \tag{3.27}$$

The general formula

$$\sum_{k=1}^{\infty} \frac{1}{k(k^2-x^2)} \frac{1}{2}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3\binom{2k}{k}} \frac{5k^2-x^2}{k^2-x^2} \prod_{m=1}^{k-1} \left(1 - \frac{x^2}{m^2}\right) \tag{3.28}$$

was obtained by Koecher [151] following techniques of Knopp and Schur.

Using bootstrapping and an application of the "Pade" function (which in both *Mathematica* and *Maple* produces Padé approximations to a rational function satisfied by a truncated power series) the following remarkable and unanticipated results were produced [50]:

$$\sum_{k=1}^{\infty} \frac{1}{k^3(1-x^4/k^4)} \frac{5}{2}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3\binom{2k}{k}(1-x^4/k^4)} \prod_{m=1}^{k-1} \left(\frac{1+4x^4/m^4}{1-x^4/m^4}\right). \tag{3.29}$$

Following an analogous—but more deliberate—experimental-based procedure, as detailed below, we provide a similar general formula for $\zeta(2n+2)$ that is pleasingly parallel to (3.29). It is:

Theorem 6.1. *Let x be a complex number not equal to a non-zero integer. Then*

$$\sum_{k=1}^{\infty} \frac{1}{k^2 - x^2} = 3 \sum_{k=1}^{\infty} \frac{1}{k^2 \binom{2k}{k} (1 - x^2/k^2)} \prod_{m=1}^{k-1} \left(\frac{1 - 4\,x^2/m^2}{1 - x^2/m^2} \right). \quad (3.30)$$

Note that the left hand side of (3.30) is trivially equal to

$$\sum_{n=0}^{\infty} \zeta(2n+2) x^{2n} = \frac{1 - \pi x \cot(\pi x)}{2x^2}. \quad (3.31)$$

Thus, (3.30) generates an Apéry-like formulae for $\zeta(2n)$ for every positive integer n.

We describe this process of discovery in some detail here, as the general technique appears to be quite fruitful and may well yield results in other settings.

We first conjectured that $\zeta(2n+2)$ is a rational combination of terms of the form

$$\sigma(2r; [2a_1, \cdots, 2a_N]) \quad := \quad \sum_{k=1}^{\infty} \frac{1}{k^{2r} \binom{2k}{k}} \prod_{i=1}^{N} \sum_{n_i=1}^{k-1} \frac{1}{n_i^{2a_i}}, \quad (3.32)$$

where $r + \sum_{i=1}^{N} a_i = n + 1$, and the a_i are listed in nonincreasing order (note that the right-hand-side value is independent of the order of the a_i). This dramatically reduces the size of the search space, while in addition the sums (3.32) are relatively easy to compute.

One can then write

$$\sum_{n=0}^{\infty} \zeta(2n+2)\, x^{2n} \quad \stackrel{?}{=} \quad \sum_{n=0}^{\infty} \sum_{r=1}^{n+1} \sum_{\pi \in \Pi(n+1-r)} \alpha(\pi)\, \sigma(2r; 2\pi)\, x^{2n}, \quad (3.33)$$

where $\Pi(m)$ denotes the set of all additive partitions of m if $m > 0$, $\Pi(0)$ is the singleton set whose sole element is the null partition $[\,]$, and the coefficients $\alpha(\pi)$ are complex numbers. In principle $\alpha(\pi)$ in (3.33) could depend not only on the partition π but also on n. However, since the first few coefficients appeared to be independent of n, we found it convenient to assume that the generating function could be expressed in the form given above.

For positive integer k and partition $\pi = (a_1, a_2, \ldots, a_N)$ of the positive integer m, let

$$\widehat{\sigma}_k(\pi) \quad := \quad \prod_{i=1}^{N} \sum_{n_i=1}^{k-1} \frac{1}{n_i^{2a_i}}.$$

Then

$$\sigma(2r; 2\pi) \quad = \quad \sum_{k=1}^{\infty} \frac{\widehat{\sigma}_k(\pi)}{k^{2r}\binom{2k}{k}},$$

and from (3.33), we deduce that

$$\begin{aligned}
\sum_{n=0}^{\infty} \zeta(2n+2)\, x^{2n} &= \sum_{n=0}^{\infty} \sum_{r=1}^{n+1} \sum_{\pi \in \Pi(n+1-r)} \alpha(\pi)\sigma(2r; 2\pi) x^{2n} \\
&= \sum_{k=1}^{\infty} \frac{1}{\binom{2k}{k}} \sum_{r=1}^{\infty} \frac{x^{2r-2}}{k^{2r}} \sum_{n=r-1}^{\infty} \sum_{\pi \in \Pi(n+1-r)} \alpha(\pi)\, \widehat{\sigma}_k(\pi) x^{2(n+1-r)} \\
&= \sum_{k=1}^{\infty} \frac{1}{\binom{2k}{k}(k^2 - x^2)} \sum_{m=0}^{\infty} x^{2m} \sum_{\pi \in \Pi(m)} \alpha(\pi)\, \widehat{\sigma}_k(\pi) \\
&= \sum_{k=1}^{\infty} \frac{1}{\binom{2k}{k}(k^2 - x^2)} P_k(x)
\end{aligned}$$

where

$$P_k(x) \quad := \quad \sum_{m=0}^{\infty} x^{2m} \sum_{\pi \in \Pi(m)} \alpha(\pi)\, \widehat{\sigma}_k(\pi), \tag{3.34}$$

whose closed form is yet to be determined. Our strategy, as in the case of (3.29) [49], was to compute $P_k(x)$ explicitly for a few small values of k in a hope that these would suggest a closed form for general k.

Some examples we produced are shown below. At each step we "bootstrapped" by assuming that the first few coefficients of the current result are the coefficients of the previous result. Then we found the remaining coefficients (which are in each case unique) by means of PSLQ computations. Note below that in the sigma notation, the first few coefficients of each expression are simply the previous step's terms, where the first argument of σ (corresponding to r) has been increased by two. These initial terms (with coefficients in bold) are then followed by terms with the other partitions as arguments, with all terms ordered lexicographically by partition

(shorter partitions are listed before longer partitions, and, within a partition of a given length, larger entries are listed before smaller entries in the first position where they differ; the integers in brackets are nonincreasing):

$$\zeta(2) = 3\sum_{k=1}^{\infty}\frac{1}{\binom{2k}{k}k^2} = 3\sigma(2,[0]), \tag{3.35}$$

$$\zeta(4) = \mathbf{3}\sum_{k=1}^{\infty}\frac{1}{\binom{2k}{k}k^4} - 9\sum_{k=1}^{\infty}\frac{\sum_{j=1}^{k-1}j^{-2}}{\binom{2k}{k}k^2} = \mathbf{3}\sigma(4,[0]) - 9\sigma(2,[2]) \tag{3.36}$$

$$\begin{aligned}\zeta(6) &= \mathbf{3}\sum_{k=1}^{\infty}\frac{1}{\binom{2k}{k}k^6} - \mathbf{9}\sum_{k=1}^{\infty}\frac{\sum_{j=1}^{k-1}j^{-2}}{\binom{2k}{k}k^4} - \frac{45}{2}\sum_{k=1}^{\infty}\frac{\sum_{j=1}^{k-1}j^{-4}}{\binom{2k}{k}k^2} \\ &\quad + \frac{27}{2}\sum_{k=1}^{\infty}\sum_{j=1}^{k-1}\frac{\sum_{i=1}^{k-1}i^{-2}}{j^2\binom{2k}{k}k^2}, \end{aligned}\tag{3.37}$$

$$= \mathbf{3}\sigma(6,[]) - \mathbf{9}\sigma(4,[2]) - \frac{45}{2}\sigma(2,[4]) + \frac{27}{2}\sigma(2,[2,2]) \tag{3.38}$$

$$\begin{aligned}\zeta(8) &= \mathbf{3}\sigma(8,[]) - \mathbf{9}\sigma(6,[2]) - \frac{\mathbf{45}}{\mathbf{2}}\sigma(4,[4]) + \frac{\mathbf{27}}{\mathbf{2}}\sigma(4,[2,2]) - 63\sigma(2,[6]) \\ &\quad + \frac{135}{2}\sigma(2,[4,2]) - \frac{27}{2}\sigma(2,[2,2,2]) \end{aligned}\tag{3.39}$$

$$\begin{aligned}\zeta(10) &= \mathbf{3}\sigma(10,[]) - \mathbf{9}\sigma(8,[2]) - \frac{\mathbf{45}}{\mathbf{2}}\sigma(6,[4]) + \frac{\mathbf{27}}{\mathbf{2}}\sigma(6,[2,2]) - \mathbf{63}\sigma(4,[6]) \\ &\quad + \frac{\mathbf{135}}{\mathbf{2}}\sigma(4,[4,2]) - \frac{\mathbf{27}}{\mathbf{2}}\sigma(4,[2,2,2]) - \frac{765}{4}\sigma(2,[8]) + 189\sigma(2,[6,2]) \\ &\quad + \frac{675}{8}\sigma(2,[4,4]) - \frac{405}{4}\sigma(2,[4,2,2]) + \frac{81}{8}\sigma(2,[2,2,2,2]). \end{aligned}\tag{3.40}$$

Next from the above results, one can immediately read that $\alpha([\,]) = 3$, $\alpha([1]) = -9$, $\alpha([2]) = -45/2$, $\alpha([1,1]) = 27/2$, and so forth. Table 1 presents the values of α that we obtained in this manner.

Using these values, we then calculated series approximations to the

Partition	Alpha	Partition	Alpha	Partition	Alpha
[empty]	3/1	1	-9/1	2	-45/2
1,1	27/2	3	-63/1	2,1	135/2
1,1,1	-27/2	4	-765/4	3,1	189/1
2,2	675/8	2,1,1	-405/4	1,1,1,1	81/8
5	-3069/5	4,1	2295/4	3,2	945/2
3,1,1	-567/2	2,2,1	-2025/8	2,1,1,1	405/4
1,1,1,1,1	-243/40	6	-4095/2	5,1	9207/5
4,2	11475/8	4,1,1	-6885/8	3,3	1323/2
3,2,1	-2835/2	3,1,1,1	567/2	2,2,2	-3375/16
2,2,1,1	6075/16	2,1,1,1,1	-1215/16	1,1,1,1,1,1	243/80
7	-49149/7	6,1	49140/8	5,2	36828/8
5,1,1	-27621/10	4,3	32130/8	4,2,1	-34425/8
4,1,1,1	6885/8	3,3,1	-15876/8	3,2,2	-14175/8
3,2,1,1	17010/8	3,1,1,1,1	-1701/8	2,2,2,1	10125/16
2,2,1,1,1	-6075/16	2,1,1,1,1,1	729/16	1,1,1,1,1,1,1	-729/560
8	-1376235/56	7,1	1179576/56	6,2	859950/56
6,1,1	-515970/56	5,3	902286/70	5,2,1	-773388/56
5,1,1,1	193347/70	4,4	390150/64	4,3,1	-674730/56
4,2,2	-344250/64	4,2,1,1	413100/64	4,1,1,1,1	-41310/64
3,3,2	-277830/56	3,3,1,1	166698/56	3,2,2,1	297675/56
3,2,1,1,1	-119070/56	3,1,1,1,1,1	10206/80	2,2,2,2	50625/128
2,2,2,1,1	-60750/64	2,2,1,1,1,1	18225/64	2,1,1,1,1,1,1	-1458/64
1,1,1,1,1,1,1,1	2187/4480				

Table 3.2. Alpha coefficients found by PSLQ computations

functions $P_k(x)$, by using formula (3.34). We obtained:

$$\begin{aligned}
P_3(x) &\approx 3-\frac{45}{4}x^2-\frac{45}{16}x^4-\frac{45}{64}x^6-\frac{45}{256}x^8-\frac{45}{1024}x^{10}-\frac{45}{4096}x^{12}-\frac{45}{16384}x^{14}\\
&\quad -\frac{45}{65536}x^{16}\\
P_4(x) &\approx 3-\frac{49}{4}x^2+\frac{119}{144}x^4+\frac{3311}{5184}x^4+\frac{38759}{186624}x^6+\frac{384671}{6718464}x^8\\
&\quad +\frac{3605399}{241864704}x^{10}+\frac{33022031}{8707129344}x^{12}+\frac{299492039}{313456656384}x^{14}\\
P_5(x) &\approx 3-\frac{205}{16}x^2+\frac{7115}{2304}x^4+\frac{207395}{331776}x^6+\frac{4160315}{47775744}x^8+\frac{74142995}{6879707136}x^{10}\\
&\quad +\frac{1254489515}{990677827584}x^{12}+\frac{20685646595}{142657607172096}x^{14}+\frac{336494674715}{20542695432781824}x^{16}
\end{aligned}$$

$$\begin{aligned}
P_6(x) &\approx 3 - \frac{5269}{400}x^2 + \frac{6640139}{1440000}x^4 + \frac{1635326891}{5184000000}x^6 - \frac{5944880821}{18662400000000}x^8 \\
&\quad - \frac{212874252291349}{67184640000000000}x^{10} - \frac{141436384956907381}{241864704000000000000}x^{12} \\
&\quad - \frac{70524260274859115989}{87071293440000000000000000}x^{14} - \frac{31533457168819214655541}{313456656384000000000000000000}x^{16} \\
P_7(x) &\approx 3 - \frac{5369}{400}x^2 + \frac{8210839}{1440000}x^4 - \frac{199644809}{5184000000}x^6 - \frac{680040118121}{18662400000000}x^8 \\
&\quad - \frac{278500311775049}{67184640000000000}x^{10} - \frac{84136715217872681}{241864704000000000000}x^{12} \\
&\quad - \frac{22363377813883431689}{87071293440000000000000000}x^{14} - \frac{5560090840263911428841}{313456656384000000000000000000}x^{16}.
\end{aligned}$$

With these approximations in hand, we were then in a position to attempt to determine closed-form expressions for $P_k(x)$. This can be done by using either "Pade" function in either *Mathematica* or *Maple*. We obtained the following:

$$\begin{aligned}
P_1(x) &\stackrel{?}{=} 3 \\
P_2(x) &\stackrel{?}{=} \frac{3(4x^2-1)}{(x^2-1)} \\
P_3(x) &\stackrel{?}{=} \frac{12(4x^2-1)}{(x^2-4)} \\
P_4(x) &\stackrel{?}{=} \frac{12(4x^2-1)(4x^2-9)}{(x^2-4)(x^2-9)} \\
P_5(x) &\stackrel{?}{=} \frac{48(4x^2-1)(4x^2-9)}{(x^2-9)(x^2-16)} \\
P_6(x) &\stackrel{?}{=} \frac{48(4x^2-1)(4x^2-9)(4x^2-25)}{(x^2-9)(x^2-16)(x^2-25)} \\
P_7(x) &\stackrel{?}{=} \frac{192(4x^2-1)(4x^2-9)(4x^2-25)}{(x^2-16)(x^2-25)(x^2-36)}
\end{aligned}$$

These results immediately suggest that the general form of a generating function identity is:

$$\sum_{n=0}^{\infty} \zeta(2n+2)x^{2n} \stackrel{?}{=} 3\sum_{k=1}^{\infty} \frac{1}{\binom{2k}{k}(k^2-x^2)} \prod_{m=1}^{k-1} \frac{4x^2-m^2}{x^2-m^2}, \qquad (3.41)$$

which is equivalent to (3.30).

We next confirmed this result in several ways:

1. We symbolically computed the power series coefficients of the LHS and the RHS of (3.41), and have verified that they agree up to the term with x^{100}.

2. We verified that $\mathcal{Z}(1/6)$, where $\mathcal{Z}(x)$ is the RHS of (3.41), agrees with $18-3\sqrt{3}\pi$, computed using (3.31), to over 2,500 digit precision; likewise for $\mathcal{Z}(1/2)=2,\ \mathcal{Z}(1/3)=9/2-3\pi/(2\sqrt{3}),\ \mathcal{Z}(1/4)=8-2\pi$ and $\mathcal{Z}(1/\sqrt{2})=1-\pi/\sqrt{2}\cdot\cot(\pi/\sqrt{2})$.

3. We then affirmed that the formula (3.41) gives the same numerical value as (3.31) for the 100 pseudorandom values $\{m\pi\}$, for $1\le m\le 100$, where $\{\cdot\}$ denotes fractional part.

Thus, we were certain that (3.30) was correct and it remained only to find a proof of Theorem 1.

Chapter 4

Exploration and Discovery in Inverse Scattering

The mathematical study of physical processes is conventionally split between modeling and, once the modeling is sufficiently mature, manipulation and data analysis. The modeling stage of exploration has a natural experimental component that conforms to the usual picture conjured by the word "experiment", that is, a laboratory with machines that funnel physical processes through a narrowly defined, well controlled environment for the purpose of isolating and observing phenomena. Here mathematics serves as a language to describe physical processes and to focus rational exploration.

In this chapter we focus on the data analysis side of experimental mathematics, and, as such, we take the models for granted. Indeed, we make no distinction between the models and the physical processes they describe. We are interested instead in the inverse problem of determining the model from an observation, or as we will refer to it, the solution to the forward model .

Inverse problems are not new to applied mathematics. The inverse scattering techniques we detail below, however, are still in their infancy and point to a shifting approach to inverse problems in general which we would like to take a moment to contrast with more conventional methodologies. Statistical regression and estimation, for instance, have long been fundamental to empirical studies which seek to find the most likely model to fit some observed data. Here the model for the data is given *a priori* with undetermined parameters that are selected by an optimization procedure which yields a best fit to the observation according to some criteria, often depending on an assumed noise or stochastic model. Quantitative confidence intervals can be obtained that provide bounds on the likelihood that a given model assumption is true through a procedure known as hypothesis testing. This process of quantitatively deducing a likely explanation for

an observation fits very naturally in the mold of experimental mathematics, and, indeed, might well be considered the template for experimental mathematics as it is applied elsewhere. What distinguishes conventional inverse problems from the case study presented here is the amount of a priori information we shall allow ourselves and the nature of the information we shall glean from our mathematical experiments. We assume relatively little at the outset about the model for the data and, through a series of qualitative feasibility tests, progressively refine our assumptions. The feasibility tests are boolean in nature, very much in the spirit of a yes/no query to an oracle that is constructed from the governing equations. As certain possibilities are ruled out, we can then begin the process of obtaining quantitative estimates for specific parameters. At the final stages of the process our techniques will merge with conventional parameter estimation techniques, but it is the process of getting to that stage of refinement that is of recent vintage and a topic of current research.

Another reason for placing the spotlight on inverse scattering is to highlight the role of computational experimentation in the development of mathematical methodologies. As mentioned above, the techniques we discuss are in their infancy. What is meant by this is that the settings in which they can be mathematically proven to work are far fewer than those in which they have been computationally demonstrated to work. At the same time, uniqueness proofs for the solutions to boundary value problems in partial differential equations have been the motivation for many computational techniques. This interplay between computation and theory is central to experimental mathematics.

The following is a case study of recent trends in *qualitative* exploration of the mathematical structure of models for the scattering of waves and their corresponding inverse operators. This approach has its roots in functional analysis methods in partial differential equations and has merged with engineering and numerical analysis. Applications range from acoustic scattering for geophysical exploration to electromagnetic scattering for medical imaging. As the mathematical models for acoustic and electromagnetic scattering share identical formulations in many instances, it follows that the methods we describe below apply to both electromagnetic and acoustic scattering . For ease, however, our case study is for acoustic scattering.

1 The Physical Experiment

The experiment consists of sending a wave through a compact region $\mathbb{D} \subset \mathbb{R}^2$ of isotropic material with a small amplitude, monochromatic, time-

harmonic acoustic plane wave denoted

$$u^i(x;\widehat{\eta},\kappa) \equiv e^{\mathrm{i}\kappa\widehat{\eta}\cdot x}, \quad x \in \mathbb{R}^2 \tag{4.1}$$

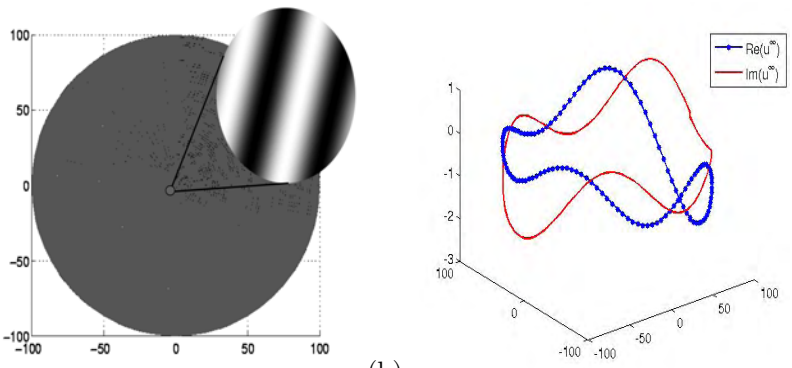
and parameterized by the incident direction $\widehat{\eta}$ in the set $\mathbb{S}$ of unit vectors in $\mathbb{R}^2$ and wavenumber $\kappa > 0$ – and recording the resulting *far field pattern* for the scattered field, denoted by $u^\infty(\cdot,\widehat{\eta},\kappa) : \partial\mathbb{D} \to \mathbb{C}$ at points $\widehat{x}$ uniformly distributed around $\partial\mathbb{D}$, the boundary of $\mathbb{D}$. The incident field at points $x \in \mathbb{D}$, a circle of radius 100 centered a the origin, and resulting far field pattern are shown in Figure 4.1. The restriction to two-dimensional space is a matter of computational convenience – the theory applies equally to $\mathbb{R}^3$. The length scales are determined by κ which is inversely proportional to the wavelength, $\omega = 2\pi/\kappa$ *physical units*. The time-dependence of the wave has already been factored out since the time-harmonicity only contributes an $e^{-\mathrm{i}\omega t}$ factor to the waves. The isotropy assumption on the region $\mathbb{D}$ means that there is no preferential direction of scattering. This is not the same as assuming that there are no obstructions to the wave since, otherwise, the far field pattern would be zero. By "far field" we mean that the radius of $\mathbb{D}$ is large relative to the wavelength, and that whatever is causing the scattering lies well inside the interior of $\mathbb{D}$. Since our experiments are at a single fixed frequency, $\kappa = 2$, we shall drop the explicit dependence of the fields on κ. The representation of the fields as mappings onto $\mathbb{C}$ is also, to some degree, a matter of convenience. For acoustic experiments it is possible to measure the *phase* of the fields, that is, the real and imaginary parts. This is not always the case for electromagnetic experiments [171] though we avoid these complications.

The experiment is repeated at N incident directions $\widehat{\eta}_n$ equally distributed on the interval $[-\pi,\pi]$. For each incident direction $\widehat{\eta}_n$, we collect N far field measurements at points $\widehat{x}_n$. The resulting arrays of data are $N \times N$ complex-valued matrices shown in Figure 4.2.

Our goal is to determine as much as possible about the scatterer(s) that produced this data. We will be satisfied with being able to locate and determine the size and shape of the scatterers. If we can determine some of the physical properties of the material making up the scatterers, we will be at the cutting edge of inverse scattering research. We begin with a review of acoustic scattering. This theory is classical, hence our treatment is terse. Readers interested in thorough derivations of the tools we explore are referred to [80] from which we borrow our formalism.

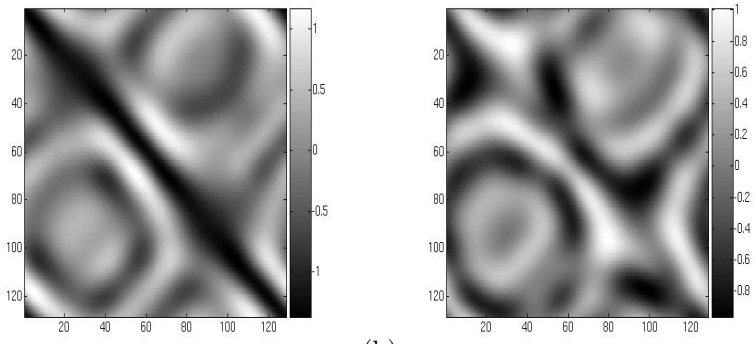
2 The Model

Our a priori knowledge is limited to the incident waves u^i and the substrate medium – what we will henceforth call the *background medium* – in



(a) (b)

Figure 4.1. **(a) Real part of incident field with direction $\widehat{\eta} = -\pi/12$, and wavenumber $\kappa = 2$ on the region $\mathbb{D}$. The inset is a close-up view of the field. (b) Far field data, real and imaginary parts, corresponding to scattering due to the passing of the incident plane wave shown in (a) through the region $\mathbb{D}$.**



(a) (b)

Figure 4.2. Far field data, real (a) and imaginary (b) parts, from a series of acoustic scattering experiments differing in the direction of the incident field. Each experiment is at the same incident wavenumber $\kappa = 2$.

$\mathbb{D}$, namely that they are small-amplitude, monochromatic plane waves traveling in an isotropic medium. The waves travel in space, and as such are represented as vector-valued functions of their position. Since the medium is isotropic, however, the vector components of the wave fields are not coupled, thus we can treat each of the spatial components of the wave as independent scalar waves obeying the same governing equations. As mentioned above, if there were no obstruction to u^i in $\mathbb{D}$ the far field pattern would be zero. As this is apparently not the case, there must be some *scat-*

tered field generated by u^i which we denote by u^s. Together the scattered field and the incident field comprise the *total field*, denoted u, where

$$u(x;\widehat{\eta}) = e^{\mathrm{i}\kappa\widehat{\eta}\cdot x} + u^s(x,\widehat{\eta}). \tag{4.2}$$

For slowly varying scattering media , the governing equations for acoustic scattering are given by the Helmholtz equation

$$\left(\triangle + n(x)\kappa^2\right) u(x) = 0, \quad x \in \mathbb{R}^2, \tag{4.3}$$

where $\triangle$ denotes the Laplacian, and $n : \mathbb{R}^2 \to \mathbb{C}$ is the *index of refraction of the medium* . In terms of the speed of propagation through the medium, this is given by

$$n(x) := \frac{c_0^2}{c^2(x)} + i\sigma(x), \tag{4.4}$$

where $c_0 > 0$ denotes the sound speed of the background medium, $c : \mathbb{R}^2 \to \mathbb{R}_+ \setminus \{0\}$ is the sound speed inside the scatterer, and $\sigma : \mathbb{R}^2 \to \mathbb{R}_+$ is a function that models the influence of absorption. The "slowly-varying" assumption on the medium means that derivatives of order 1 and higher are negligible. Note that at a point x in the background medium, $\frac{c_0^2}{c^2(x)} = 1$. We assume that the background medium is nonabsorbing, that is $\sigma(x) = 0$, hence $n(x) = 1$ in the background medium.

If the domain $\mathbb{D}$ contains in its interior one or more sound reflectors ($c^2(x) \ll 1$) or a highly absorbing materials ($\sigma \gg 1$) then we restrict the domain in Eq.(4.3) from $\mathbb{R}^2$ to the open exterior domain $\Omega^o = \mathbb{R}^2 \setminus \Omega$ where the compact domain Ω contains the support of the scatterers. The scatterers are then modeled as *obstacles* which behave as either *sound-soft*, or *sound-hard* obstacles, or some mixture of these. This is modeled with Dirichlet, Neumann or Robin boundary conditions on $\partial\Omega$:

$$u = f \quad \text{or} \quad \frac{\partial u}{\partial \nu} = f, \quad \text{or} \quad \frac{\partial u}{\partial \nu} + \lambda u \;=\; f, \quad \text{on } \partial\Omega, \tag{4.5}$$

where, f is continuous on $\partial\Omega$, ν is the unit outward normal, and λ is an *impedance function* .

In either case, slowly-varying media or obstacle scattering, the scattered field u^s is a decaying field satisfying what is known as the *Sommerfeld radiation condition*

$$r^{\frac{1}{2}}\Big(\frac{\partial}{\partial r} - i\kappa\Big)u^s(x) \to 0, \;\; r = |x| \to \infty, \tag{4.6}$$

uniformly in all directions.

We prefer the formulation as slowly-varying media to obstacle scattering since the latter can be viewed as an ideal limiting case of the former. If,

in the course of our investigation it appears that there is some obstacle scattering present we can always refocus our investigation to take this into account.

For inhomogeneous media, the far field pattern – what is actually measured – is modeled by the leading term in the asymptotic expansion for u^s as $|x| \to \infty$, namely

$$u^s(x,\widehat{\eta}) = \beta \frac{e^{\mathrm{i}\kappa|x|}}{|x|^{1/2}} u^\infty(\widehat{x},\widehat{\eta}) + o\left(\frac{1}{|x|^{1/2}}\right), \qquad \widehat{x} = \frac{x}{|x|} \tag{4.7}$$

where

$$\beta = \left(\frac{e^{\mathrm{i}\frac{\pi}{2}}}{8\pi\kappa}\right)^{1/2} \tag{4.8}$$

and

$$u^\infty(\widehat{x},\widehat{\eta}) \equiv -\kappa^2 \int_{\mathbb{R}^2} e^{-\mathrm{i}\kappa y\cdot\widehat{x}} m(y) u(y,\widehat{\eta})\ dy. \tag{4.9}$$

In an abuse of notation we have replaced points $\widehat{x} \in \partial\mathbb{D}$ with directions on the unit sphere $\widehat{x} \in \mathbb{S}$, which is tantamount to letting $\partial\mathbb{D} \to \infty$ in all directions, consistent with the asymptotic expression for u^∞.

Since we will be trying to reconstruct volumetric information from boundary measurements, it shouldn't come as a surprise that a central tool in our analysis is Green's theorem

$$\int_{\mathbb{D}} (v\triangle u + \nabla v\cdot\nabla u)\ dx = \int_{\partial\mathbb{D}} v\frac{\partial u}{\partial \nu}\ ds \tag{4.10}$$

for $u \in C^2(\mathbb{D})$ and $v \in \mathbb{C}^1(\mathbb{D})$, $\partial\mathbb{D}$ sufficiently smooth and the unit outward normal ν. For $v \in \mathbb{C}^2(\mathbb{D})$ this is sometimes given as

$$\int_{\mathbb{D}} (v\triangle u - u\triangle v)\ dx = \int_{\partial\mathbb{D}} v\frac{\partial u}{\partial \nu} - u\frac{\partial v}{\partial \nu}\ ds \tag{4.11}$$

We will also make use of Green's formula

$$\begin{aligned} u(x) \quad = \quad & \int_{\partial\mathbb{D}} \left(\Phi(x,y)\frac{\partial u}{\partial\nu}(y) - u(y)\frac{\partial\Phi(x,y)}{\partial\nu(y)}\right)\ ds(y) \\ & - \int_{\mathbb{D}} \left(\triangle u(y) + \kappa^2 u(y)\right)\Phi(x,y)\ dy, \quad x \in \mathbb{D}, \end{aligned} \tag{4.12}$$

where $\Phi(\cdot,y) : \mathbb{R}^2 \setminus \{y\} \to \mathbb{C}$ is the free space ($n(x) \equiv 1$) radiating fundamental solution to Eq.(4.3):

$$\Phi(x,y) := \frac{i}{4} H_0^{(1)}(\kappa|x-y|), \qquad x \neq y \tag{4.13}$$

with $H_0^{(1)}$ denoting the zero-th order Hankel function of the first kind.

For smooth boundaries $\partial\Omega$, Eq.(4.7), Eq.(4.12), and the asymptotic expansion of Φ with respect to $|x|$ yield

$$u^\infty(\widehat{x},\widehat{\eta}) = \beta \int_{\partial\Omega} \left(u^s(y,\widehat{\eta}) \frac{\partial e^{-\mathrm{i}\kappa\widehat{x}\cdot y}}{\partial\nu(y)} - \frac{\partial u^s(y,\widehat{\eta})}{\partial\nu} e^{-\mathrm{i}\kappa\widehat{x}\cdot y} \right), \quad \widehat{x},\widehat{\eta}\in\mathbb{S} \tag{4.14}$$

Using Eq.(4.3) for the scattered field u^s, Eq.(4.6) and Eq.(4.10), it can be shown that the far field enjoys a symmetry

$$u^\infty(\widehat{x},\widehat{\eta}) = u^\infty(-\widehat{\eta},-\widehat{x}). \tag{4.15}$$

This relation is the well known *reciprocity* relation for the far field.

A very useful tool in our analysis and algorithms are fields that can be represented as the superposition of plane waves. Define the *Herglotz Wave Operator* $\mathcal{H}: L^2(-\Lambda) \to H^1(\mathbb{R}^m)$ by

$$(\mathcal{H}g)(x) \equiv \int_\Lambda e^{i\kappa x\cdot(-\widehat{y})} g(-\widehat{y}) ds(\widehat{y}), \quad x\in\mathbb{R}^2 \tag{4.16}$$

where Λ is an open subset of $\mathbb{S}$, $g\in L^2(-\mathbb{S})$, and H^1 denotes the Sobolev space of order 1 defined by

$$\|h\|_{H^1(\mathbb{R}^m)} \equiv \left(\int_{\mathbb{R}^m} |v(x)|^2 + |\nabla v(x)|^2 dx \right)^{1/2}.$$

Theorem 2.1. *Let $u^s:\mathbb{R}^2\to\mathbb{C}$ and $u^\infty:\mathbb{S}\to\mathbb{C}$ denote the scattered and far fields, respectively, due to excitation from an incident plane wave u^i at a fixed wavenumber $\kappa>0$ with direction $-\widehat{\eta}$, $u^i(x,-\widehat{\eta}) \equiv e^{i\kappa x\cdot(-\widehat{\eta})}$, $x\in\mathbb{R}^2$, $\widehat{\eta}\in\mathbb{S}$. The solution to the scattering problem corresponding to the incident field $v_g^i \equiv (\mathcal{H}g)(x)$ for $\mathcal{H}$ given by Eq.(4.16) is $v_g = v_g^i + v_g^s$, where*

$$v_g^s(x) \;=\; \int_{\mathbb{S}} u^s(x,-\widehat{\eta}) g(-\widehat{\eta}) ds(\widehat{\eta}), \qquad x\in\mathbb{R}^2\setminus\overline{\Omega}. \tag{4.17}$$

The incident field $v_g^i \equiv (\mathcal{H}g)(x)$ given by Eq.(4.16) is called a Herglotz wave function*. The corresponding far field pattern is given by*

$$v_g^\infty(\widehat{x}) \;=\; \int_\Lambda u^\infty(\widehat{x},-\widehat{\eta}) g(-\widehat{\eta}) ds(\widehat{\eta}), \qquad \widehat{x}\in\mathbb{S}. \tag{4.18}$$

Proof. The proof follows from the linearity and boundedness of the particular scattering problem. Linearity implies that scattering from an incident field that is the superposition of incident plane waves can be represented as the superposition of scattering from independent scattered plane waves.

By boundedness of the scattering operator from $C(\Omega)$ into $C_{loc}(\mathbb{R}^2 \setminus \overline{\Omega})$, the limit for the integration can be performed and we obtain the stated results. □

Remark 2.1. *The signs in the expressions for v_g^i, v_g^s, and v_g^∞ above have been chosen so that the backprojection mapping between the far field measurements and the scattered field, which we derive later, has a natural interpretation in terms of a physical aperture in the far field. Note that the function g is defined on $-\Lambda$, the mirror image of the interval Λ: $\widehat{\eta} \in \Lambda \in \mathbb{S} - \widehat{\eta} \in -\Lambda$. Using the far field reciprocity relation Eq.(4.15) we see that the far field is defined on Λ with any incident wave direction $-\widehat{x}$.*

Remark 2.2. *By Eq.(4.15) we have*

$$v_g^\infty(\widehat{x}) \;=\; \int_\Lambda u^\infty(\widehat{\eta}, -\widehat{x}) g(-\widehat{\eta}) ds(\widehat{\eta}), \qquad \widehat{x} \in \mathbb{S}. \tag{4.19}$$

□

It is worthwhile taking a second look at the Herglotz wave operator Eq.(4.16). This operator, restricted to the boundary of some region $\Omega_t \subset \mathbb{R}^m$ with piecewise C^2 boundary and connected exterior, is injective with dense range [81]. In particular, we can construct the density $g(\cdot; z)$ such that $v^i_{g(\cdot,z)}(x) \approx \Phi(x,z)$ and $\frac{\partial v^i_{g(\cdot,z)}}{\partial n}(x) \approx \frac{\partial \Phi(x,z)}{\partial n}$ arbitrarily closely for $x \in \partial\Omega$ and $z \in \Omega^o$. By Eq.(4.12) and Eq.(4.18), for $z \in \Omega^o$ and any generic scattered field u^s satisfying Eq.(4.3) on Ω^o and the radiation condition Eq.(4.6), we have

$$\begin{aligned}
u^s(z) &= \int_{\partial\Omega} \left\{ \Phi(y,z) \frac{\partial u^s}{\partial \nu}(y) - \frac{\partial \Phi(y,z)}{\partial \nu(y)} u^s(y) \right\} ds(y), \\
&\approx \int_{\partial\Omega} \left\{ v^i_{g(\cdot;z)}(y) \frac{\partial u^s}{\partial \nu}(y) - \frac{\partial v^i_{g(\cdot;z)}(y)}{\partial \nu(y)} u^s(y) \right\} ds(y), \\
&= \int_\Lambda \int_{\partial\Omega} \left\{ e^{i\kappa y \cdot (-\widehat{x})} \frac{\partial u^s}{\partial \nu}(y) - \frac{\partial e^{i\kappa y \cdot (-\widehat{x})}}{\partial \nu(y)} u^s(y) \right\} ds(y) g(-\widehat{y}; z) ds(\widehat{y}) \\
&= \int_\Lambda u^\infty(\widehat{y}) g(-\widehat{y}; z) ds(\widehat{y})
\end{aligned} \tag{4.20}$$

We can rewrite the above relation in a more suggestive manner for the specific case of scattering from an incident plane wave. Define the *far field operator* $\mathcal{A} : L^2(\mathbb{S}) \to L^2(\mathbb{S})$ corresponding to the far field data u^∞ by

$$\mathcal{A}f(\widehat{x}) \equiv \int_{\mathbb{S}} u^\infty(\widehat{x}, -\widehat{\eta}) f(-\widehat{\eta}) \; ds(\widehat{\eta}). \tag{4.21}$$

When the integral is over an open subset $\Lambda \subset \mathbb{S}$ we denote the far field operator restricted to the limited aperture Λ by $\mathcal{A}_\Lambda$. Then, using the reciprocity relation Eq.(4.15), we can rewrite Eq.(4.20) as

$$u^s(z, -\widehat{x}) \approx (\mathcal{A}g(\cdot; z))(\widehat{x}). \tag{4.22}$$

If the scattering problem has a unique solution, then knowledge of the scattered field u^s everywhere is all we need in order to determine the location, shape and properties of the scatterer. According to Eq.(4.22), we can reconstruct u^s by acting on some well-chosen densities g with the far field operator. In other words, the scattered field is in the *range* of the far field operator. Before exploring the range of $\mathcal{A}$, however, we should convince ourselves that the scattering problem does indeed have a unique solution.

2.1 Weak Formulation

The scattering model is conventionally reformulated as an integral equation, which admits both existence and uniqueness theorems, as well as numerical algorithms. To do this we make use of the *volume potential*

$$(\mathcal{V}\varphi)(x) \equiv \int_{\mathbb{R}^m} \Phi(x, y) m(y) \varphi(y)\, dy, \quad x \in \mathbb{R}^2, \tag{4.23}$$

where m is the *scattering potential*,

$$m \equiv 1 - n. \tag{4.24}$$

Note that from the assumptions above $m(x) = 0$ outside the scattering inhomogeneity, and, from the assumption that the inhomogeneity lies in the interior of $\mathbb{D}$, m has compact support.

If u satisfies Eq.(4.2), Eq.(4.3) and Eq.(4.6), and is twice continuously differentiable on $\mathbb{R}^2$, i.e. $u \in C^2(\mathbb{R}^2)$, then by Eq.(4.12) applied to u we have

$$\begin{aligned} u(x) = \int_{\partial r\mathbb{B}} \left(\frac{\partial u}{\partial \nu} \Phi(x, y) - u(y) \frac{\partial \Phi(\cdot, y)}{\partial \nu(y)}(x) \right) ds(y) \\ -\kappa^2 \int_{r\mathbb{B}} \Phi(x, y) m(y) u(y)\, dy \end{aligned} \tag{4.25}$$

where $r\mathbb{B}$ is the ball of radius r with $\operatorname{supp}(m) \in \operatorname{int}(r\mathbb{B})$ and $x \in \operatorname{int}(r\mathbb{B})$, and ν is the unit outward normal to $r\mathbb{B}$. Here we have used the fact that u satisfies Eq.(4.3) and $m = 1 - n$. Applying Eq.(4.12) to the incident field u^i yields

$$u^i(x) = \int_{\partial r\mathbb{B}} \left(\frac{\partial u^i}{\partial \nu} \Phi(x, y) - u^i(y) \frac{\partial \Phi(\cdot, y)}{\partial \nu(y)}(x) \right) ds(y) \tag{4.26}$$

Moreover, Eq.(4.10) applied to u^s together with the boundary condition Eq.(4.6) gives

$$0 = \int_{\partial r\mathbb{B}} \left(\frac{\partial u^s}{\partial \nu} \Phi(x,y) - u^s(y) \frac{\partial \Phi(\cdot,y)}{\partial \nu(y)}(x) \right) \, ds(y). \tag{4.27}$$

Altogether, Eq.(4.25)-(4.27) and Eq.(4.2) yield the well known *Lippmann-Schwinger* integral equation

$$u = u^i - \kappa^2 \mathcal{V} u. \tag{4.28}$$

From Eq.(4.2) it follows immediately from Eq.(4.28) that

$$u^s(x,\widehat{\eta}) = -\kappa^2 (\mathcal{V}u)(x,\widehat{\eta}) \tag{4.29}$$

With a little more work [80, Theorem 8.1 and 8.2] it can be shown that if $u \in C(\mathbb{R}^2)$ solves Eq.(4.28), then u satisfies Eq.(4.2)-(4.6) for $n \in C^1(\mathbb{R}^2)$. We summarize this discussion with the following restatement of [80, Theorem 8.3].

Theorem 2.2 (Helmholtz-Lippmann-Schwinger equivalence)**.** *If $u \in C^2(\mathbb{R}^2)$ satisfies Eq.(4.2)-(4.6) for $n \in C^1(\mathbb{R}^2)$, then u solves Eq.(4.28). Conversely, if $u \in C^2(\mathbb{R}^2)$ solves Eq.(4.28), then u solves Eq.(4.2)-(4.6).*

The previous theorem assumed that the solutions to Eq.(4.2)-(4.6) and equivalently Eq.(4.28) exist, which begs the question about existence and uniqueness.

Theorem 2.3. *For each $\kappa > 0$ there exists a unique u satisfying Eq.(4.2), Eq.(4.3) and Eq.(4.6) with u^i given by Eq.(4.1). Moreover, u depends continuously on u^i with respect to the max-norm.*

Proof sketch. The main ideas are sketched here. Interested readers are referred to the works of Leis [164], Reed and Simon [205] and Colton and Kress [80]. The integral operator $\mathcal{V}$ has a weakly singular kernel Φ, hence is a compact operator on $C(\mathbb{D})$ where $\Omega \in \text{int}\,(\mathbb{D})$ for $\mathbb{D}$ bounded. Standard results in integral equations [155, 208] ensure that, if the homogeneous Lippmann-Schwinger equation

$$u + \kappa^2 \mathcal{V} u = 0,$$

has only the trivial solution, then the inhomogeneous equation has a unique solution and the inverse operator $(I+\kappa^2\mathcal{V})^{-1}$ exists and is bounded in $C(\mathbb{D})$. Since the inverse operator is bounded, the solution u depends continuously on u^i with respect to the max-norm on $C(\mathbb{D})$. From the previous discussion, the homogeneous Lippmann-Schwinger equation is equivalent to

$$\left(\triangle + n(x)\kappa^2\right) u(x) = 0, \quad x \in \mathbb{R}^2, \tag{4.30}$$

and

$$r^{\frac{1}{2}}\Big(\frac{\partial}{\partial r} - i\kappa\Big)u(x) \to 0, \;\; r = |x| \to \infty, \tag{4.31}$$

Green's Theorem Eq.(4.10) and the unique continuation principle [164] are employed to show that the only solution to Eq.(4.30)-(4.31) is the trivial solution. □

The natural question to ask next is what about the index of refraction n is encoded in the far field data u^∞. This is the central question in inverse scattering and one that often leads to algorithms. In the next section we shall motivate some representative inverse scattering algorithms by the uniqueness results at their foundation.

3 The Mathematical Experiment: Qualitative Inverse Scattering

A long standing problem for the mathematical community has been to determine when the scattering potential is uniquely determined by the far field. The answer to this question is subtle and varied, depending on the setting – whether $\mathbb{R}^2$ or $\mathbb{R}^3$, isotropic or anisotropic, multi-frequency or single frequency, Dirichlet, Neumann, or inhomogeneous media. In three or more dimensions for isotropic media, key results affirming the recovery of $n(x)$ from $u^\infty(\widehat{x};\widehat{\eta})$, where u^∞ is known for all $\widehat{x},\widehat{\eta} \in \mathbb{S}$, can be found in Nachman [182], Novikov [185], and Ramm [204]. In two dimensions, the case we consider here, positive results are more difficult. A quick heuristic as to why this might be the case is to look at the dimensions of the data versus those of the unknown index of refraction n. In $\mathbb{R}^3$ the dimension of $n : \mathbb{R}^3 \to \mathbb{C}$ is 3 while the data $u^\infty : \mathbb{S} \times \mathbb{S} \to \mathbb{C}$ is on two two-dimensional spheres, $\widehat{x} \in \mathbb{S}$ and $\widehat{\eta} \in \mathbb{S}$, so that the problem appears to be overdetermined. By contrast, in $\mathbb{R}^2$ the unknown $n : \mathbb{R}^2 \to \mathbb{C}$ and the data $u^\infty : \mathbb{S} \times \mathbb{S} \to$ are both two-dimensional since, in this case, the spheres $\mathbb{S}$ are one-dimensional. The same *kind* of data in $\mathbb{R}^2$ does not stretch as far is it does in $\mathbb{R}^3$.

Two dimensional results for *small* potentials are given by Novikov [186], for *most* potentials in an open dense subset of the Sobolev space $W^{1,\infty}$ by Sun and Uhlmann [216], and later by the same authors for *discontinuous* potentials in $L^\infty(\mathbb{D})$ [217], for special kinds of potentials by Nachman [183] (though not necessarily for fixed $\kappa > 0$), and for a broad class of potentials at *almost all* fixed $\kappa \in \mathbb{R}_+$ by Eskin [105]. In some cases [182, 183] the proofs are *constructive* and yield algorithms for inversion. More generic results give practitioners courage to try inversion, but provide little specific guidance as to exactly how much information is really required. Most of the

results, with some exceptions that we will explore below, require a continuum of far field measurements from a continuum of incident field directions. Negative results would indeed be useful, though very few are known that indicate when the information is *not* sufficient for unique recovery.

3.1 Where Is the Scatterer? How Big is It?

The first thing we might naturally want to know about the scatterer is where it is and, approximately, how big it is. The first method we study appeared in [172] and belongs to a class of algorithms that share similar features. With these methods one constructs a test domain $\Omega_t(z)$, parameterized by the reference point $z \in \mathbb{R}^2$, and an *indicator function* μ that is a function of the test domain $\Omega_t(z)$ and the far field pattern u^∞ measured on some open subset $\Lambda \subset \mathbb{S}$ for a fixed incident direction $-\widehat{x}$. These far field measurements correspond to a segment of a column of the matrix shown in Figure 4.2. The indicator function μ is "small" for all $z \in \mathbb{R}^2$ such that $\Omega \subset \text{int}\,(\Omega_t(z))$ and infinite for those z where $\Omega_t(z) \cap \Omega = \emptyset$. The size of the test domain Ω_t and the set of points $\left\{ z \in \mathbb{R}^2 \mid \mu(\Omega_t(z), u^\infty(\Lambda, -\widehat{x})) < \infty \right\}$ will determine the approximate size and location of the scatterer. A crucial advantage of this method and others like it [142, 198] is that they only require *one* incident wave (single frequency and single incident direction) on a limited aperture, $\Lambda \subset \mathbb{S}$.

Definition 3.1 (scattering test response)**.** *Given the far field pattern $u^\infty(\widehat{\eta}, -\widehat{x})$ for $\widehat{\eta} \in \Lambda \subset \mathbb{S}$ due to an incident plane wave $u^i(\cdot, -\widehat{x})$ with fixed direction $-\widehat{x}$, let $v_g^i \equiv (\mathcal{H}g)(x)$ denote the incident field defined by Eq.(4.16) and $v_g^\infty(\widehat{x})$ denote the corresponding far field pattern given by Eq.(4.19). We define the* scattering test response *for the test domain Ω_t by*

$$\mu(\Omega_t, u^\infty(\Lambda, -\widehat{x})) := \sup \left\{ \left| v_g^\infty(\widehat{x}) \right| \;\middle|\; g \in L^2(-\Lambda) \text{ with } \|v_g^i\|_{\Omega_t} = 1 \right\}. \tag{4.32}$$

Theorem 3.2 (behavior of the scattering test response)**.** *Let the bounded domain Ω_t have a C^2 boundary. If $\Omega \subset \Omega_t$, then there is a constant $c \in \mathbb{R}$ such that*

$$\mu(\Omega_t, u^\infty(\Lambda, -\widehat{x})) \leq c.$$

If, on the other hand, $cl\,(\Omega) \cap cl\,(\Omega_t) = \emptyset$, and $\mathbb{R}^2 \setminus (cl\,(\Omega) \cup cl\,(\Omega_t))$ is connected, then we have

$$\mu(\Omega_t, u^\infty(\Lambda, -\widehat{x})) = \infty.$$

Proof. This was proved in [172, Theorem 3.2]. We reproduce the proof here. When $\Omega \subset \Omega_t$ the scattering map $u^i \mapsto u^\infty$ is bounded for any generic

incident field u^i that is an entire solution to the Helmholtz equation. Hence there exists a constant c such that for all u^i satisfying

$$\left\|u^i\right\|_{C(\Omega_t)} = 1,$$

we have

$$\|u^\infty\|_{C(\mathbb{S})} \leq c.$$

This completes the proof of the first statement.

To prove the second statement, we consider two disjoint domains, Ω_t' and Ω', satisfying $\Omega_t \subset \Omega_t'$, $\Omega \subset \Omega'$, and $\overline{\Omega_t'} \cap \overline{\Omega'} = \emptyset$. We further require that the interior homogeneous Dirichlet problems for Ω_t' and Ω' have only the trivial solution. Then the Herglotz wave operator $H : L^2(-\Lambda) \to L^2(\partial(\Omega_t' \cup \Omega'))$, defined by

$$(\mathcal{H}g)(x) := v_g^i(x)\Big|_{\partial(\Omega_t' \cup \Omega')},$$

has dense range. This can be shown using the techniques of [81]. Choose $y \notin \overline{\Omega_t'} \cup \overline{\Omega'}$ such that the far field pattern $w^\infty(\widehat{x}, y)$ for scattering of an incident point source $\Phi(\cdot, y)$ by Ω is not zero. This is always possible since, by the *mixed reciprocity relation* [197, Theorem 2.1.4], we have

$$w^\infty(\widehat{x}, y) = \gamma u^s(y, -\widehat{x})$$

and $u^s(\cdot, -\widehat{x})$ cannot vanish on an open subset of $\mathbb{R}^2$. Next, construct $v_g^i(x)$ satisfying

$$\left\|v_g^i\right\|_{C(\Omega_t')} = 1, \quad \text{and} \quad \left\|v_g^i - \beta\Phi(\cdot, y)\right\|_{C(\Omega')} \leq 1. \tag{4.33}$$

Then since $\Omega_t \subset \Omega_t'$, we have

$$\mu(\Omega_t, u^\infty(\Lambda, -\widehat{x})) \geq \left|v_g^\infty(\widehat{x})\right|. \tag{4.34}$$

By definition $\Omega \subset \Omega'$; thus

$$\left|v_g^\infty(\widehat{x}) - \beta w^\infty(\widehat{x}, y)\right| \leq c,$$

for some constant c, which, by the triangle inequality, yields

$$\left|v_g^\infty(\widehat{x})\right| \geq \beta \left|w^\infty(\widehat{x}, y)\right| - c. \tag{4.35}$$

Together, Eq.(4.33)-(4.35) yield

$$\mu(\Omega_t, u^\infty(\Lambda, -\widehat{x})) \geq \beta \left|w^\infty(\widehat{x}, y)\right| - c$$

for all $\beta \in \mathbb{R}$. This completes the proof. □

Definition 3.3 (corona of Ω)**.** *Let $\Omega_t(z) \subset \mathbb{R}^2$ be a domain parameterized by $z \in \mathbb{R}^2$. Define the* corona *of the scatterer Ω, relative to the scattering test response μ given in Eq.(4.32) by*

$$\mathbb{M}_\mu \;:=\; \bigcup_{\substack{z\in\mathbb{R}^2 \\ s.t.\quad \mu(\Omega_t(z),u^\infty(\Lambda,-\widehat{x}))<\infty}} \Omega_t(z). \tag{4.36}$$

Corollary 3.1 (approximate size and location of scatterers)**.** *Let $\Omega_t(z) \subset \mathbb{R}^2$, with $\mathbb{R}^2 \setminus cl(\Omega_t(z))$ connected, be a bounded domain large enough that there is some $z \in \mathbb{R}^2$ for which $\Omega \subset \Omega_t(z)$. Then we have*

$$\mathbb{M}_\mu \subset \bigcup_{z\in\mathbb{R}^2} \{\Omega_t(z) \mid cl(\Omega_t(z)) \cap cl(\Omega) \neq \emptyset\} \tag{4.37}$$

and the scatterer Ω is a subset of its corona, $\mathbb{M}_\mu$.

Proof. For points z with $\mu(\Omega_t^0(z), u^\infty(\Lambda, -\widehat{x})) < \infty$, by Theorem 3.2 we have $\text{cl}\left(\Omega_t^0(z)\right) \cap \overline{\Omega} \neq \emptyset$, from which we immediately obtain the relation (4.37). For $\Omega \subset \Omega_t^0(z)$ we have $\mu(\Omega_t^0(z), u^\infty(\Lambda, -\widehat{x})) < \infty$ and thus the support of the scatterer is a subset of its corona. □

Before we develop the numerical algorithm based on these facts we would like to point out some challenges in the theory above. The observant reader will notice that we didn't mention what happens to μ when $\Omega \cap \Omega_t \neq \emptyset$ but $\Omega \cap (\mathbb{R}^2 \setminus \Omega_t) \neq \emptyset$. If $\mu(\Omega_t^0(z), u^\infty(\Lambda, -\widehat{x})) = \infty$ when $\Omega \cap \Omega_t \neq \emptyset$ but $\Omega \cap (\mathbb{R}^2 \setminus \Omega_t) \neq \emptyset$, then the set inequality Eq.(4.37) can be tightened considerably. Unfortunately, because the fields can be analytically continued, the answer to this question is not as sharp as we might hope. Also note that the constant c in Theorem 3.2 is not resolved. All we know is that c is smaller than infinity, but that isn't saying much. Naturally, we would like to know more about this constant. Finally, note that the optimization problem embedded in Eq.(4.32) is an infinite dimensional problem which we must solve at *each* point $z \in \mathbb{D}$. Even with all these challenges, we can get a surprising amount of information from a naive implementation of this test, as we demonstrate below.

The algorithm we shall develop is for full aperture data, that is $\Lambda = \mathbb{S}$. The same basic algorithm can be adapted for the limited aperture settings. We address first the computation of the scattering test response μ. By Eq.(4.32), we must solve, at each point z, an optimization problem over the densities $g \in L^2(\mathbb{S})$. We construct densities such that the incident field v_g^i approximates the fundamental solution to Eq.(4.3), $\Phi(x, y)$ on $x \in \partial\Omega_t(z)$ for $y \in \mathbb{R}^2 \setminus \Omega_t(z)$, that is, at each point $z \in \mathbb{D}$ we solve the following ill-posed equation for g:

$$v_g^i(x) = \Phi(x, y) \text{ for } x \in \partial\Omega_t(z),\ y \in \mathbb{R}^2 \setminus \Omega_t(z) \tag{4.38}$$

where v_g^i is given by Eq.(4.16). Denote the Herglotz wave operator Eq.(4.16) corresponding to v_g^i by $\mathcal{H}_z$. Since Eq.(4.38) is ill-posed, we solve the Tikhonov-regularized least-squares problem

$$\underset{g\in L^2(\mathbb{S})}{\text{minimize}} \ \|\mathcal{H}_z g - \Phi(\cdot, y)\|^2 + \alpha\|g\|^2 \tag{4.39}$$

whose solution is

$$g(\cdot; y, z, \alpha) := (\alpha I + \mathcal{H}_z^*\mathcal{H}_z)^{-1}\mathcal{H}_z^*\Phi(\cdot, y). \tag{4.40}$$

This yields

$$v_g^i(\cdot; y, z, \alpha) \ \approx \ \Phi(\cdot, y) \quad \text{on } \partial\Omega_t(z).$$

We thus exchange the infinite dimensional optimization problem in Eq.(4.32) for the parameterized, finite dimensional optimization problem

$$\mu(\Omega_t(z), u^\infty(\mathbb{S}, -\widehat{x})) = \sup_{y\in\mathbb{R}^2} \left|v_g^\infty(\widehat{x}; y, z, \alpha_0)\right| \tag{4.41}$$

where

$$g_*(\widehat{\eta}; y, z, \alpha_0) = \left((\alpha_0 I + \mathcal{H}_z^*\mathcal{H}_z)^{-1}\mathcal{H}_z^*\Phi\right)(\widehat{\eta}, y) \tag{4.42}$$

and

$$g = \frac{g_*}{\|v_{g_*}^i\|}.$$

Next, we introduce a specific approximation domain that allows further efficiencies due to symmetry of the fundamental solution Φ. Let Ω_r^0 be a circle of radius r centered on the origin. We construct the test domain parameterized by points $z \in \mathbb{R}^2$ by linear translations of the domain Ω_r^0: $\Omega_t(z) \equiv \Omega_r^0 + z$. In [170, Proposition 2] it is shown that

$$g_*(\widehat{\eta}; y, z, \alpha_0) = e^{-\mathrm{i}\kappa z\cdot\widehat{\eta}} g_*(\widehat{\eta}; y, 0, \alpha_0), \tag{4.43}$$

where g_* is the respective solution to Eq.(4.39). Similarly, rotations of the point y around the origin translate to shifts in the density with respect to $\mathbb{S}$:

$$g_*(\widehat{\eta}; R_{\widehat{w}} y, z, \alpha_0) = e^{-\mathrm{i}\kappa z\cdot\widehat{\eta}} g_*(R_{-\widehat{w}}\widehat{\eta}; y, 0, \alpha_0), \tag{4.44}$$

where $R_{\widehat{w}}$ is the rotation in the direction $\widehat{w}$ in the plane. Thus, for fixed $r > 0$ with $y = r\widehat{y}$, where $\widehat{y} \equiv y/|y|$, we have the explicit, closed-form expressions for the densities g_*:

$$g_*(\widehat{\eta}; y, z, \alpha_0) = e^{-\mathrm{i}\kappa z\cdot\widehat{\eta}} g_*(R_{-\widehat{y}}\widehat{\eta}; (r, 0), \alpha_0), \tag{4.45}$$

where

$$g_*(\widehat{\eta}; (r, 0), 0, \alpha_0) = \left((\alpha_0 I + \mathcal{H}_0^*\mathcal{H}_0)^{-1}\mathcal{H}_0^*\Phi\right)(\widehat{\eta}; (r, 0)). \tag{4.46}$$

We need only compute Eq.(4.46) along the line $r > 0$ and solve Eq.(4.41) for these explicitly calculated densities.

At the time of this writing, no satisfactory algorithm exists for the solution of Eq.(4.41). Instead, we propose the following technique in the form of a conjecture.

Conjecture 3.2. *Let $\Omega_t(0)$ be a circle of radius r centered at the origin where r is large enough that $\Omega \subset \Omega_t(0) + z$ for some $z \in \mathbb{R}^2$. For each $\widehat{y} \in \mathbb{S}$, let $g(\widehat{\eta}, r\widehat{y}, 0)$ solve*

$$(\mathcal{H}g(\widehat{\eta}))\,(\partial\Omega_t(0), r\widehat{y}) = \Phi(\partial\Omega_t(0), r\widehat{y}). \tag{4.47}$$

Define the partial scattering test response *, $\delta : \mathbb{R}^2 \to \mathbb{R}_+$, by*

$$\delta(z) \equiv \int_{\mathbb{S}} |(\mathcal{A}_\Lambda g_z)(\widehat{y})|\ ds(\widehat{y}) \tag{4.48}$$

where $g_z \equiv e^{\mathrm{i}\kappa\widehat{\eta}\cdot z} g(\widehat{\eta}, r\widehat{y}, 0)$ and $\mathcal{A}_\Lambda$ is the far field operator given by Eq.(4.21) restricted to $\Lambda \subset \mathbb{S}$. Then, for any $\widehat{x} \in \mathbb{S}$ there exist constants $0 < M' < M$ such that

$$\delta(z) \begin{cases} > M & \forall\ z \in \mathbb{R}^2 \quad \textit{where} \quad \Omega \cap \Omega_t(0) + z = \emptyset \\ < M' & \forall\ z \in \mathbb{R}^2 \quad \textit{where} \quad \Omega \subset \textit{int}\,(\Omega_t(0) + z). \end{cases} \tag{4.49}$$

Using the efficient formulation Eq.(4.45), we need only solve one ill-posed integral equation, Eq.(4.46) and calculate the integral in Eq.(4.49) at each sample point z. The corona of the scatterer is identified by a jump in $\delta(z)$.

We can test this conjecture numerically. If Conjecture 3.2 is true, then the scatterer lies in the region $\mathbb{G} + \Omega_t(0)$ where $\Omega_t(0)$ is a circle centered on the origin and $\mathbb{G}$ is the set of points $z \in \mathbb{D}$ where Eq.(4.49) is satisfied. To generate Figure 4.3 we used a test domain radius of $r = 5\sqrt{2}$. In order to cover the domain $\mathbb{D} = [-100, 100] \times [-100, 100]$ this required us to sample on a 20×20 grid for a pixel size of 10×10. The dark region $\mathbb{G}$ shown in the figure are the points z on this grid where Eq.(4.49) is satisfied with $M = 9$, indicating that there is something of interest in the region $\mathbb{G} + \Omega_t(0)$.

There are two principal advantages of this method: first, only a single incident direction is needed, and, second, one can cast as large a net as desired, depending on the radius of the test domain Ω_t, in order to determine the approximate location and size of the scatterer without the need to sample at many points $z \in \mathbb{D}$. We will see in section 3.3 whether the conjectured region does indeed contain something of interest.

3.2 Is the Scatterer Absorbing?

Recall from section 2 that the scattered field we seek lies in the range of the far field operator $\mathcal{A}$ defined by Eq.(4.21). The first question that comes to

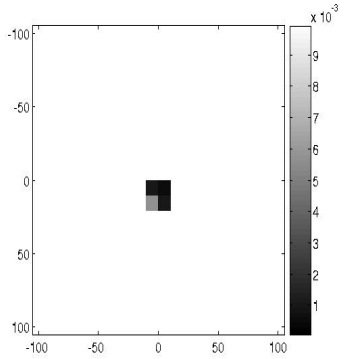


Figure 4.3. Plotted is the value of the integral in Eq.(4.49) at the reference points $z \in \mathbb{D}$ where $\mathbb{D}$ is shown in Figure 4.1(a). The test domain $\Omega_t(0)$ is a circle of radius $r = 5\sqrt{2}$ and with a cutoff value in Eq.(4.49) of $M = 9$. If conjecture 3.2 is true, then the scatterer lies in the region $\mathbb{G} + \Omega_t(0)$ where $\mathbb{G}$ is the set of points in the dark region above.

mind, then, regarding the range of $\mathcal{A}$ is, what are it's spectral properties, that is, how do its eigenvalues behave. We obtain a partial answer with the next theorem.

Theorem 3.4. *The far field operator has at most a countable number of discrete eigenvalues with zero as the only cluster point.*

Proof. This follows from the fact that $\mathcal{A}$ is a compact operator [205, 208]. □

This is disappointing. Even if zero is not and eigenvalue, the decay of the eigenvalues of $\mathcal{A}$ will cause numerical instabilities and sensitivity to noise in the determination of the range of $\mathcal{A}$. This is symptomatic of *ill-posed problems* like the one we are faced with here. While the spectral properties of the far field operator point to some of the difficulties in recovering scatterers from far field measurements, the next theorem due to Colton and Kress [79, 80] shows that the spectrum of $\mathcal{A}$ easily provides qualitative information about the nature of the scatterer, whether it is absorbing or or not, from the location of the eigenvalues of $\mathcal{A}$.

Theorem 3.5. *Let the scattering inhomogeneity have index of refraction defined in Eq.(4.4) mapping $\mathbb{R}^2$ to the upper half of the complex plane. The scattering inhomogeneity is nonabsorbing, that is, $\Im n(x) = 0$ for all x, if and only if the eigenvalues of $\mathcal{A}$ lie on the circle centered at*

$\frac{1}{2\kappa}\left(\Im(\beta^{-1}), \Re(\beta^{-1})\right)$ *and passing through the origin. Otherwise, the eigenvalues of* $\mathcal{A}$ *lie on the interior of this disk.*

Proof. Let $v_g = v_g^i + v_g^s$ satisfy Eq.(4.3) with v_g^s satisfying Eq.(4.6) and v_g^i defined by Eq.(4.16) where $\Lambda = \mathbb{S}$ and where $g \in L^2(\mathbb{S})$ is an eigenfunction of $\mathcal{A}$, that is, $\mathcal{A}g = \lambda g$. By Eq.(4.10) we have

$$\int_\Omega (\overline{v_g}\Delta v_g - v_g \Delta \overline{v_g})\, dx = \int_{\partial\Omega} \overline{v_g}\frac{\partial v_g}{\partial \nu} - v_g \frac{\partial \overline{v_g}}{\partial \nu}\, ds. \tag{4.50}$$

But v_g satisfies Eq.(4.3) so the left hand side of Eq.(4.50) satisfies

$$\int_\Omega (\overline{v_g}\Delta v_g - v_g \Delta \overline{v_g})\, dx = -2\kappa^2 \mathrm{i} \int_\Omega \Im(n(x))|v_g(x)|^2\, dx. \tag{4.51}$$

According to Eq.(4.2), the right hand side of Eq.(4.50) can be expanded to

$$\begin{aligned}\int_{\partial\Omega} \overline{v_g}\frac{\partial v_g}{\partial \nu} - v_g \frac{\partial \overline{v_g}}{\partial \nu}\, ds &= \int_{\partial\Omega} \overline{v_g^s}\frac{\partial v_g^s}{\partial \nu} - v_g^s \frac{\partial \overline{v_g^s}}{\partial \nu}\, ds \\ &\quad -2\mathrm{i}\Im \int_{\partial\Omega} \overline{v_g^s}\frac{\partial v_g^i}{\partial \nu} - v_g^i \frac{\partial \overline{v_g^s}}{\partial \nu}\, ds + \int_{\partial\Omega} \overline{v_g^i}\frac{\partial v_g^i}{\partial \nu} - v_g^i \frac{\partial \overline{v_g^i}}{\partial \nu}\, ds \\ &= \int_{\partial\Omega} \overline{v_g^s}\frac{\partial v_g^s}{\partial \nu} - v_g^s \frac{\partial \overline{v_g^s}}{\partial \nu}\, ds - 2\mathrm{i}\Im \int_{\partial\Omega} \overline{v_g^s}\frac{\partial v_g^i}{\partial \nu} - v_g^i \frac{\partial \overline{v_g^s}}{\partial \nu}\, ds.\end{aligned} \tag{4.52}$$

Now, Eq.(4.6) gives

$$v_g^s \frac{\partial \overline{v_g^s}}{\partial \nu} = \frac{-\mathrm{i}\kappa}{|x|} v_g^\infty \overline{v_g^\infty} + O\left(\frac{1}{|x|^2}\right), \qquad |x| \to \infty.$$

So by Eq.(4.11),

$$\int_{\partial\Omega} \overline{v_g^s}\frac{\partial v_g^s}{\partial \nu} - v_g^s \frac{\partial \overline{v_g^s}}{\partial \nu}\, ds = 2\mathrm{i}\kappa \int_{\mathbb{S}} |v_g^\infty|^2\, ds. \tag{4.53}$$

Expanding the expression for v_g^i, together with a change in the order of integration and Eq.(4.14) yields

$$\begin{aligned}\int_{\partial\Omega} \overline{v_g^s}\frac{\partial v_g^i}{\partial \nu} - v_g^i \frac{\partial \overline{v_g^s}}{\partial \nu}\, ds &= \int_{\mathbb{S}} g(\widehat{y}) \int_{\partial\Omega} \overline{v_g^s}\frac{\partial e^{\mathrm{i}\kappa \widehat{y}\cdot x}}{\partial \nu} - e^{\mathrm{i}\kappa \widehat{y}\cdot x}\frac{\partial \overline{v_g^s}}{\partial \nu}\, ds(x)\, ds(\widehat{y}) \\ &= \overline{\beta^{-1}} \int_{\mathbb{S}} g(\widehat{y}) \overline{v_g^\infty}(\widehat{y})\, ds(\widehat{y}).\end{aligned} \tag{4.54}$$

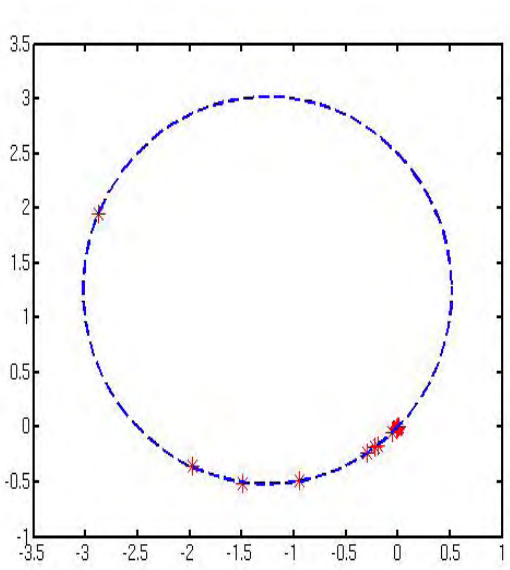


Figure 4.4. The eigenvalues (asterisks) of the far field matrix shown in Figure 4.2 are shown to line up on the circle passing through the origin with center $1/2\kappa(\Im\beta, \Re\beta)$ for $\kappa = 2$ and β given by Eq.(4.8). This implies that the inhomogeneity is nonabsorbing .

In summary, Eq.(4.50)-(4.54) give

$$\begin{aligned} -2\kappa^2 \mathrm{i} \int_\Omega \Im(n(x))|v_g(x)|^2 \, dx &= 2\mathrm{i}\kappa \int_{\mathbb{S}} |v_g^\infty|^2 \, ds + \overline{\beta^{-1}} \int_{\mathbb{S}} g(\widehat{y})\overline{v_g^\infty}(\widehat{y}) \, ds(\widehat{y}) \\ &\quad -\beta^{-1} \int_{\mathbb{S}} \overline{g}(\widehat{y}) v_g^\infty(\widehat{y}) \, ds(\widehat{y}) \\ &= 2\mathrm{i}\kappa \int_{\mathbb{S}} |v_g^\infty|^2 \, ds - 2\mathrm{i}\Im\left(\beta^{-1} \left\langle v_g^\infty, \ g \right\rangle\right). \end{aligned}$$

By Eq.(4.18) we have that $\mathcal{A}g = v_g^\infty$, and since g is an eigenfunction of $\mathcal{A}$, we have

$$-\frac{\kappa}{||g||^2} \int_\Omega \Im(n(x))|v_g(x)|^2 \, dx = |\lambda|^2 - \Im\left(\frac{\lambda}{\kappa\beta}\right). \tag{4.55}$$

This completes the proof. □

As figure Figure 4.4 shows, the eigenvalues of the sampled far field operator corresponding to our data set lie on the circle passing through the

origin with center

$$1/2\kappa(\Im\beta, \Re\beta) = \sqrt{\pi}\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$$

for $\kappa = 2$ and β given by Eq.(4.8). From the previous theorem we conclude that the scatterer is nonabsorbing. [1]

In some applications, this may be all that is needed to distinguish whether the scatterer is of interest. For instance, the presence of leukemia causes the bone marrow to exhibit an elevated electric permittivity and diminished conductivity as compared to a healthy individual. This, difference can, in theory, be detected by a shift in the location of the eigenvalues of the far field operator corresponding to electromagnetic measurements analogous to the data shown in Figure 4.2 [84]. Whether this is a viable diagnostic tool depends on a number of factors: the sensitivity of the eigenvalues to electrical parameter shifts; whether the shift, or *signal* can be distinguished from random variations, or *noise*; and whether the amount of data needed to achieve a sufficient signal is practical - i.e. the required electromagnetic radiation should not do more damage to the patient than the disease!

3.3 What Is the Shape of the Scatterer?

We now use the information from the previous method to refine our search of the domain $\mathbb{D}$ with the goal of determining the approximate shape of the scatterer. The method we shall explore is the *linear sampling* method developed by Colton, Kirsch and others [78, 86, 9, 179, 77]. As with the method in the previous section, this method and others like it (see [197, 147, 82, 83]) use the blow-up of some function to indicate whether a sampled point $z \in \mathbb{R}^2$ is on the interior of the unknown scatterer. The linear sampling method is a type of *feasibility* test for the solution of what is known as the *interior transmission problem*. This partial differential equation is recast as a linear integral equation parameterized by z where the kernel of integral operator is the far field data u^∞. It is shown that the interior transmission problem almost always has at most one solution, and if $z \in \text{int}\,(\Omega)$, then a solution does indeed exist. Moreover as $z \to \partial\Omega$ from the interior of Ω then the solution to the related linear integral equation becomes unbounded, indicating the infeasibility of the interior transmission problem. We use the blow-up of this putative solution to image the boundary of the scatterer. This is described in more detail next.

[1] The reader is encouraged to reproduce this observation with the data set provided. See notes.

To begin, we define the interior transmission problem:

$$\triangle w(x) + \kappa^2 n(x) w(x) = 0, \quad \triangle v(x) + \kappa^2 v(x) = 0 \text{ for } x \in \text{int}\,(\Omega) \tag{4.56}$$

$$w - v = f(\cdot, z), \quad \frac{\partial w}{\partial \nu} - \frac{\partial v}{\partial \nu} = \frac{\partial f}{\partial \nu} \quad \text{on } \partial\Omega. \tag{4.57}$$

From our first experiment, shown in Figure 4.4 we have determined that the medium that generated the data shown in Figure 4.2 is nonabsorbing, that is, $\Im(n(x)) = 0$ for all x. If it had been even slightly absorbing ($\Im(n(x)) > 0$ as would most likely be the case for any physical data) then we could have been certain that there are no nontrivial solutions to the homogeneous problem Eq.(4.56)-(4.57) with $f = 0$, hence the inhomogeneous problem will have a unique solution when a solution exists. But even in our case, it was shown [85] that the set of values of κ for which the solution to Eq.(4.56)-(4.57) with $f = 0$ has a nontrivial solution – called *transmission eigenvalues* – is a discrete set. We can therefore be almost sure that, for our data, κ is *not* a transmission eigenvalue.

We observe further the following fact.

Theorem 3.6. *Let*

$$f(y) = h_p^{(1)}(\kappa|y|) Y_p(\widehat{y}), \tag{4.58}$$

a spherical wave function of order p, and let β be given by Eq.(4.8). The integral equation

$$\int_{\mathbb{S}} u^\infty(\widehat{x}; \widehat{y}) g(-\widehat{x})\, ds(\widehat{x}) = \frac{\mathrm{i}^{p-1}}{\beta\kappa} Y_p(\widehat{y}), \qquad \widehat{y} \in \mathbb{S} \tag{4.59}$$

has a solution $g \in L^2(\mathbb{S})$ if and only if there exists $w \in C^2(int\,(\Omega)) \cap C^1(\Omega)$ and a function v given by

$$v(x) = \int_{\mathbb{S}} e^{\mathrm{i}\kappa x \cdot (-\widehat{y})} g(-\widehat{y})\, ds(\widehat{y}) \tag{4.60}$$

such that the pair (w, v) is a solution to Eq.(4.56)-(4.57)

Proof. Let $w^i \equiv v$ and $w^s \equiv f$. Then w solves the scattering problem

$$\left(\triangle + \kappa^2 n(x)\right) w(x) = 0 \text{ for } x \in \mathbb{R}^2 \tag{4.61}$$

$$w = f + v \tag{4.62}$$

$$r^{\frac{1}{2}} \Big(\frac{\partial}{\partial r} - i\kappa \Big) f(x) \to 0, \;\; r = |x| \to \infty. \tag{4.63}$$

The solution to this problem is also the solution to the interior transmission problem Eq.(4.56)-(4.57) (note that v given by Eq.(4.60) is an entire solution to the Helmholtz equation). Now by Theorem 2.1 and the observation that

$$f(y) = \frac{e^{\mathrm{i}\kappa|y|}}{|y|^{1/2}} \frac{\mathrm{i}^{p-1}}{\kappa} Y_p(\widehat{y}) + o\left(\frac{1}{|y|^{1/2}}\right), \qquad \widehat{y} = \frac{y}{|y|},$$

we have that $w^\infty(\widehat{y}) = \frac{\mathrm{i}^{p-1}}{\beta\kappa} Y_p(\widehat{y})$. Hence the solvability of Eq.(4.56)-(4.57) is equivalent to the solvability of Eq.(4.59). □

Finally, it can be shown [80, Theorem 10.25] that there exists a unique weak solution to Eq.(4.56)-(4.57) with $f(x;z) \equiv \Phi(x,z)$ for every $z \in \text{int}\,(\Omega)$ with Φ given by Eq.(4.13), that is the pair (w,v) satisfies

$$(I + \kappa^2 \mathcal{V})w = v \qquad \text{on} \;\; \text{int}\,(\Omega) \tag{4.64}$$

and

$$-\kappa^2 \mathcal{V} w = \Phi(\cdot, z) \qquad \text{on} \;\; \partial\mathbb{B} \tag{4.65}$$

where $\mathcal{V}$ is the volume integral defined by Eq.(4.23) and $\mathbb{B} \subset \mathbb{R}^2$ is a ball with $\text{int}\,(\Omega) \subset \mathbb{B}$.

To recap the logic thus far, we observe that Eq.(4.59) has a solution if and only if there is a corresponding solution to Eq.(4.56)-(4.57) with f given by Eq.(4.58); moreover, Eq.(4.56)-(4.57) with $f = \Phi(x,z)$ is solvable for every $z \in \text{int}\,(\Omega)$. The natural thing to ask is, for $z \in \mathbb{R}^2 \setminus \Omega$, or, just as $z \to \partial\Omega$ from $\text{int}\,(\Omega)$, what happens to solutions to

$$\int_{\mathbb{S}} u^\infty(\widehat{x};\widehat{y}) g(-\widehat{x})\, ds(\widehat{x}) = \Phi^\infty(\widehat{y}, z), \qquad \widehat{y} \in \mathbb{S} \tag{4.66}$$

where $\Phi^\infty(\widehat{y}, z)$ is the far field pattern of the fundamental solution $\Phi(y,z)$?

Theorem 3.7 (linear sampling)**.** *For every $\epsilon > 0$ and $z \in \Omega$ there exists a $g(\cdot;z) \in L^2(\mathbb{S})$ satisfying*

$$\|\mathcal{A}g - \Phi^\infty(\cdot, z)\|_{L^2(\mathbb{S})} \leq \epsilon \tag{4.67}$$

such that

$$\lim_{z \to \partial\Omega} \|g\|_{L^2(\mathbb{S})} = \infty \quad \textit{and} \quad \lim_{z \to \partial\Omega} \|v_g\|_{L^2(\mathbb{S})} = \infty \tag{4.68}$$

where v_g is given by Eq.(4.16).

Proof. Let the pair $(w(\cdot;z), v(\cdot;z))$ be the weak solution to Eq.(4.56)-(4.57) with $f = \Phi(\cdot, z)$, that is, w and v satisfy Eq.(4.64)-(4.65). Denote the linear space of Herglotz wave functions by

$$\mathbb{H} \equiv \left\{ h \;\middle|\; h(x) = \int_{\mathbb{S}} e^{\mathrm{i}\kappa x \cdot \widehat{\eta}} g(\widehat{\eta})\, ds(\widehat{\eta}) \text{ for } g \in L^2(\mathbb{S}) \right\}.$$

For this space, $v \in \text{cl}\,(\mathbb{H})$, thus for every $\widetilde{\epsilon} > 0$ and $z \in \text{int}\,(\Omega)$ there is a $g(\cdot;z) \in L^2(\mathbb{S})$ such that

$$\|v(\cdot, z) - v_g(\cdot)\|_{L^2(\Omega)} \leq \widetilde{\epsilon}, \tag{4.69}$$

where

$$v_g(x) = \int_{\mathbb{S}} e^{\mathrm{i}\kappa x\cdot\widehat{\eta}} g(\widehat{\eta}; z)\ ds(\widehat{\eta}).$$

Next, note that the inverse operator $(I+\kappa^2\mathcal{V})^{-1}$ is defined and continuous thus, for some $c>0$ and

$$w_g \equiv (I+\kappa^2\mathcal{V})^{-1} v_g,$$

we have

$$\|w(\cdot,z) - w_g(\cdot)\|_{L^2(\Omega)} \leq c\widetilde{\epsilon}. \tag{4.70}$$

By the continuity of $\mathcal{V}$ Eq.(4.70) yields the following bound on the seminorm

$$\left\|\kappa^2\mathcal{V}w_g - \Phi(\cdot,z)\right\|_{C(\partial\mathbb{B})} \leq c'\widetilde{\epsilon} \tag{4.71}$$

for some $c'>0$. Denote the scattered field due to the incident field $u^i = e^{\mathrm{i}\kappa^2(-\widehat{\eta})}$ by $u^s(x,-\widehat{\eta})$. From Theorem 2.1 we have $w_g = v_g - \kappa^2\mathcal{V}w_g$ and

$$-\kappa^2\mathcal{V}w_g = \int_{\mathbb{S}} u^s(x,-\widehat{\eta}) g(-\widehat{\eta})\ ds(\widehat{\eta}),$$

hence

$$w_g^\infty(\widehat{x}) \equiv \int_{\mathbb{S}} u^\infty(\widehat{x},-\widehat{\eta}) g(-\widehat{\eta})\ ds(\widehat{\eta}) = (\mathcal{A}g)(\widehat{x}) \quad \widehat{x}\in\mathbb{S}. \tag{4.72}$$

Note that the solution to the scattering problem on the exterior region $\mathbb{R}^2\setminus\mathbb{B}$ depends continuously on the boundary conditions on $\partial\mathbb{B}$, thus

$$\|\mathcal{A}g - \Phi^\infty(\cdot,z)\|_{L^2(\mathbb{S})} \leq c''\widetilde{\epsilon} \tag{4.73}$$

for some $c''>0$. Letting $\epsilon = c''\widetilde{\epsilon}$ completes the first part of the proof.

The proof of Eq.(4.68) is technical. We sketch the basic idea here and refer the reader to [77] for details. The idea is to use the continuous embedding of $C(\partial\Omega)$ in the Sobolev space $H^{3/2}(\partial\Omega)$ to achieve the bound

$$\|\Phi(\cdot,z)\|_{C(\partial\Omega)} \leq c\|\kappa^2\mathcal{V}w(\cdot,z)\|_{H^3/2(\partial\Omega)}$$

By the trace theorem and the fact that $(I+\kappa^2\mathcal{V})^{-1}$ and $\mathcal{V}$ are bounded, one obtains the bound

$$\|\kappa^2\mathcal{V}w(\cdot,z)\|_{H^3/2(\partial\Omega)} \leq c\|v(\cdot,z)\|_{L^2(\Omega)}.$$

Now, combining Eq.(4.69) with the previous inequalities yields

$$\|\Phi(\cdot,z)\|_{C(\partial\Omega)} \leq c\left(\|v_g\|_{L^2(\Omega)} + \widetilde{\epsilon}\right)$$

whence the right limit of Eq.(4.68). Since v_g with g bounded in $L^2(\mathbb{S})$ is also bounded in $L^2(\Omega)$, the left inequality of Eq.(4.68) follows immediately. □

The algorithm suggested by Theorem 3.7 follows almost immediately and shares many features of the scattering test response of section 3.1. First note that the equation

$$\mathcal{A}g = \Phi^\infty(\cdot, z) \tag{4.74}$$

is ill-posed, albeit linear, with respect to g. As in section 3.1 we regularize the problem by solving the regularized least squares problem

$$\underset{g\in L^2(\mathbb{S})}{\text{minimize}} \ \|\mathcal{A}g - \Phi^\infty(\cdot, z)\|^2 + \alpha\|g\|^2. \tag{4.75}$$

The solution to this problem is

$$g(\cdot; z, \alpha) \equiv (\alpha I + \mathcal{A}^*\mathcal{A})^{-1}\mathcal{A}^*\Phi^\infty(\cdot, z). \tag{4.76}$$

Since we already know from the scattering test response approximately where and how big the scatterer is, we needn't calculate Eq.(4.76) at all points $z \in \mathbb{D}$, but rather just on the corona $\mathbb{M}_\mu$, or, if we are confident of Conjecture 3.2, then on the corona $\mathbb{M}_\delta$ calculated using δ in Eq.(4.48). We identify the boundary of the scatterer by those points z_j on a grid where the norm of the density $g(\cdot; z_j, \alpha)$ becomes large relative to the norm of the density at neighboring points.

In Figure 4.5 we show the value of $\|g(\cdot; z_j, \alpha)\|_{L^2(\mathbb{S})}$ at points $z_j \in \mathbb{G} \subset \mathbb{M}_\delta$ with $\alpha = 10^{-8}$ fixed and a cutoff value of 2. The resulting image indicates that the scatterer consists of two distinct scatterers of different size.

For our implementations we do not take as much care with the choice of the regularization parameter α as we could. A more precise implementation would calculate an optimal α at each sample point z_j (see [86]). Obviously, since Eq.(4.74) is regularized, $\|g(\cdot; z, \alpha)\|_{L^2(\mathbb{S})}$ will be bounded, so a reasonable cutoff will have to be chosen which will affect the estimate for the shape and extent of the boundary of the scatterer. This is clearly a weakness of the technique, but it does not appear in practice to be significant. The contours of Figure 4.5 indicate that the shape estimate is robust with respect to this cutoff.

What is more problematic about linear sampling, however, is that it only says that *there exists* a g that is unbounded in norm, it doesn't tell us how to calculate that g or even how common these densities are. Theorem 3.7 only provides *sufficient* conditions for the blow-up of the density

$g(\cdot; z, \alpha)$. We cannot exclude the possibility of a phantom scatterer consisting of points $z \in \mathbb{R}^2 \setminus \Omega$ satisfying

$$\|\mathcal{A}g - \Phi^\infty(\cdot, z)\|_{L^2(\mathbb{S})} \leq \epsilon \tag{4.77}$$

such that $\|g\|_{L^2(\mathbb{S})}$ is "small". We observe, however, that $\|g\|_{L^2(\mathbb{S})}$ remains relatively large for all $z \in \mathbb{R}^2 \setminus \Omega$.

Kirsch [147, 148] amended the linear sampling method to close this gap. In particular, he showed that

$$(\mathcal{A}^*\mathcal{A})^{\frac{1}{4}} g = \Phi^\infty(\cdot, z) \tag{4.78}$$

has a solution if *and only if* $z \in \text{int}\,(\Omega)$. Imaging using $(\mathcal{A}^*\mathcal{A})^{\frac{1}{4}}$ has become known as the *factorization method* . Arens [9] later showed that the density computed by Eq.(4.76) corresponds to a Tikhonov-regularized solution to Eq.(4.78), thus, in this setting, the linear sampling is indistinguishable from the more rigorous method of Kirsch. This is indeed a remarkable coincidence since it is not at all obvious why one would use Tikhonov regularization (or others like it, including More-Penrose inversion) other than the fact that it is easy to implement. Had a different regularization, such as maximum entropy, been the method of choice, linear sampling might not have worked so well, and Kirsch's correct method never discovered.

Remark 3.3 (Open Problem)**.** *Does there exist a regularization of Eq.(4.74) such that the linear sampling method is guaranteed to fail.*

Kirsch's factorization is limited to nonabsorbing scatterers since it relies on the fact that $\mathcal{A}$ is normal, a property that is lost as soon as $\Im(n(x)) \neq 0$. Linear sampling, on the other hand, still works in these and many more exotic settings, so the search for a complete theory behind this successful technique continues.

3.4 Refining the Information

To maintain some sense of mystery about the scatterer we are trying to tease out of the data in Figure 4.2, we have purposely withheld the true answer to each of the above queries about the nature of the scatterer. Unfortunately, inverse scattering is not like an Arthur Conan Doyle novel – there is never a tipping point in our investigation in which, with all the clues in place, we can simply deduce the solution. We still need to peek at the answer to see if we are on the right track. To close this case study, we reveal the true answer in order to see how we did, and we discuss current trends for refining the information. The true scatterer was a series of 6 circles, two with radius 0.5 centered at $(0, 0)$ and $(3, -3)$, and four with radius 1 centered at $(\pm 1, \pm 1)$. The circles centered at $(0, 0)$, $(-1, -1)$, $(1, 1)$, and

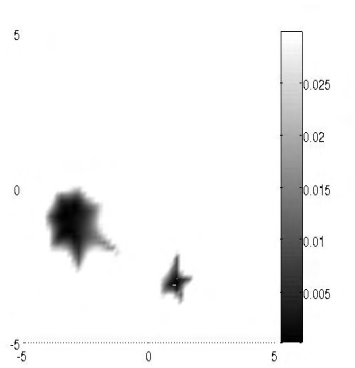


Figure 4.5. Scatterer estimated using the linear sampling method. Shown is $\|g(\cdot; z_j, \alpha)\|_{L^2(\mathbb{S})}$ for $g(\cdot; z_j, \alpha)$ given by Eq.(4.76) with $\alpha = 10^{-8}$ for all grid points z_j on the domain $[-6, 6] \times [-6, 6]$ sampled at a rate of 40 points in each direction. The cutoff is 2.

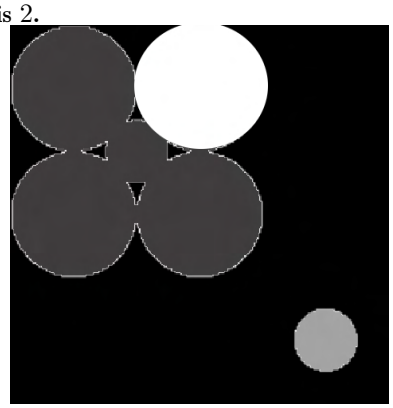


Figure 4.6. The true scatterer consisting of 6 circles of different sizes and indices of refraction indicated by the color.

$(1, -1)$ all had an index of refraction of $n(x) = \frac{401}{400}$ while the circles centered at $(3, -3)$ and $(-1, 1)$ had $n(x) = 201$ and $n(x) = 401$ respectively.

We correctly located the scatterer and determined its approximate size using the partial scattering test response in section 3.1. We also correctly determined that the scatterer is nonabsorbing. Our implementation of linear sampling did a good job of estimating the boundary of the scatterers with the largest index of refraction, but missed the weak scatterers. To distinguish between weak and strong scatterers one must, to some degree, estimate the (relative) index of refraction of the scatterers. This, however,

comes at the cost of more data.

4 Current Research

The methods we have reviewed are not exhaustive. There are numerous variations and alternative strategies. Current research is tending toward meta-algorithms that use different techniques in concert in order to progressively tease more information out of the data. At the same time, these and other methods have made it possible to do imaging in increasingly diverse settings. Questions such as what constitutes and image, and what features allow one to discriminate one object from another are fundamental to the science of imaging. In many cases specific knowledge about a very special case can allow for easy discrimination, however this likely does not generalize. The methods we illustrated here are applications of general principles that can be applied in a wide variety of settings. Research on extending these principles to other physical applications is ongoing. What will not change about these methods is the constant interplay between computational experimentation and analysis.

Chapter 5

Exploring Strange Functions on the Computer

Every computer algebra system has built-in routines for representation, computation, manipulation and display of elementary and less-elementary functions of analysis, number theory or probability. Typically, these are smooth functions since often they are solutions of differential equations or are otherwise defined in a similarly analytical way. It is therefore no surprise that many of these techniques do not extend to non-smooth ("strange") functions without such convenient differential equations, integral representations or polynomial approximations. In this chapter we want to show that even such strange functions can be explored on the computer and useful theorems about them can be discovered. Sometimes it is a question of finding the right technique for their analysis, sometimes it is the even more basic matter of visualizing them, i.e., of getting the computer to draw them. We will demonstrate these two points on two classes of examples: certain nowhere differentiable functions, and certain probability distribution functions.

1 Nowhere Differentiable Functions

The first example of a continuous, nowhere differentiable function which became widely known is due to Karl Weierstraß. Weierstraß introduced his example in a talk held at the Akademie der Wissenschaften in Berlin, in 1872 ([227]; the function was first published by Paul du Bois-Reymond in 1875, [100]). Weierstraß proved that the continuous function

$$C_{a,b}(x) \;:=\; \sum_{n=0}^{\infty} a^n \cos\left(b^n \cdot 2\pi x\right),$$

Figure 5.1. Approximation to the Weierstraß function $C_{0.9,7}$.

where $0 < a < 1$ and $b \in 2\mathbb{N} + 1$, nowhere has a finite or infinite derivative if $a \cdot b > 1 + \frac{3}{2}\pi$. Figure 5.1 shows an approximation to $C_{0.9,7}$ (the series evaluated to 60 terms at the points $i/(4 \cdot 7^n)$ for $n = 4$).

Until the publication of Weierstraß's example, 19th century mathematicians had been of divided opinions whether such a continuous, nowhere differentiable (cnd) function could exist. Although that question was now settled, analysts continued to be fascinated by the Weierstraß example and by cnd functions in general. Other, sometimes simpler, cnd functions were found, general constructions were given, and of course many aspects of the Weierstraß example were investigated. One such aspect is the precise range of the parameters for which $C_{a,b}$ is nowhere differentiable. It is clear that $C_{a,b}$ is continuously differentiable when $|a|\, b < 1$. But what happens for $1 \le |a|\, b \le 1 + \frac{3}{2}\pi$? Or for even intergers or reals b? Despite much effort, no real progress on this question was made in the years after Weierstraß. In fact, it took more than forty years until finally in 1916, G.H. Hardy ([131]) proved the strongest possible result: Both $C_{a,b}$ and the corresponding sine series

$$S_{a,b}(x) \;:=\; \sum_{n=0}^{\infty} a^n \sin\left(b^n \cdot 2\pi x\right)$$

have no finite derivative anywhere whenever b is a real greater than 1, and $ab \ge 1$. (Hardy also proved that for small values of $ab \ge 1$, the functions can have infinite derivatives.) This settled the most important questions.

However, Hardy's methods are not easy. They use results which lie a good deal "deeper" than the simple question: Is this function, given by a uniformly convergent series, differentiable somewhere?

Therefore, and because of the fascination mathematicians have often felt with such pathological but also beautiful objects, research into the Weierstraß functions and into cnd functions in general has continued and continues until today. Several approaches to the Weierstraß functions have been proposed, putting them, for example, into the context of lacunary Fourier series ([116, 145]), almost periodic functions ([136, 135]), functional equations ([120, 121]), or treating them in their own right (leading to short proofs of non-differentiability in [33, 25]). Here we will report on the functional equations approach since it leads to a simple proof of non-differentiability which can be applied to larger classes of functions, and since some key discoveries can be made on the computer. This approach works for integer values of b (a case which is usually investigated separately from the case of arbitrary real b since it employs different, simpler methods), and for didactic reasons we will restrict the discussion to the case $b = 2$ (the inclusion of higher integer values of b would afford more formalism without corresponding gain) and to the function $S_{a,2} = \sum_{n=0}^{\infty} a^n \sin(2^n \cdot 2\pi x)$ as shown in Figure 5.2 for $a = \frac{1}{2}$ and $a = \frac{3}{4}$. Finally, because of periodicity, it makes sense to restrict our attention to functions defined on the interval $[0, 1]$.

1.1 Functional Equations

We do not need computers to find promising functional equations for the Weierstraß functions. In fact, it has been known for a long time and is easy to verify that $S_{a,2}$ satisfies

$$S_{a,2}\left(\frac{x}{2}\right) \;=\; a\, S_{a,2}(x) + \sin(\pi x)$$

for all $x \in [0, 1]$. In words: $S_{a,2}\left(\frac{x}{2}\right)$ (i.e., a condensed version of $S_{a,2}$ on $\left[0, \frac{1}{2}\right]$) consists of a rescaled version of itself plus a sine wave.

Unfortunately, this functional equation alone is not sufficient to characterize the Weierstraß function on the interval $[0, 1]$. There are many continuous functions, among them differentiable a.e. functions, which satisfy the same functional equation on $[0, 1]$. Thus we cannot expect to infer interesting theorems about the Weierstraß function from this functional equation alone. We therefore attempt to add a second functional equation such that the two equations together hopefully are characteristic of the Weierstraß function. Many choices for this second functional equation are possible, of course, but it will turn out that the most natural idea is to

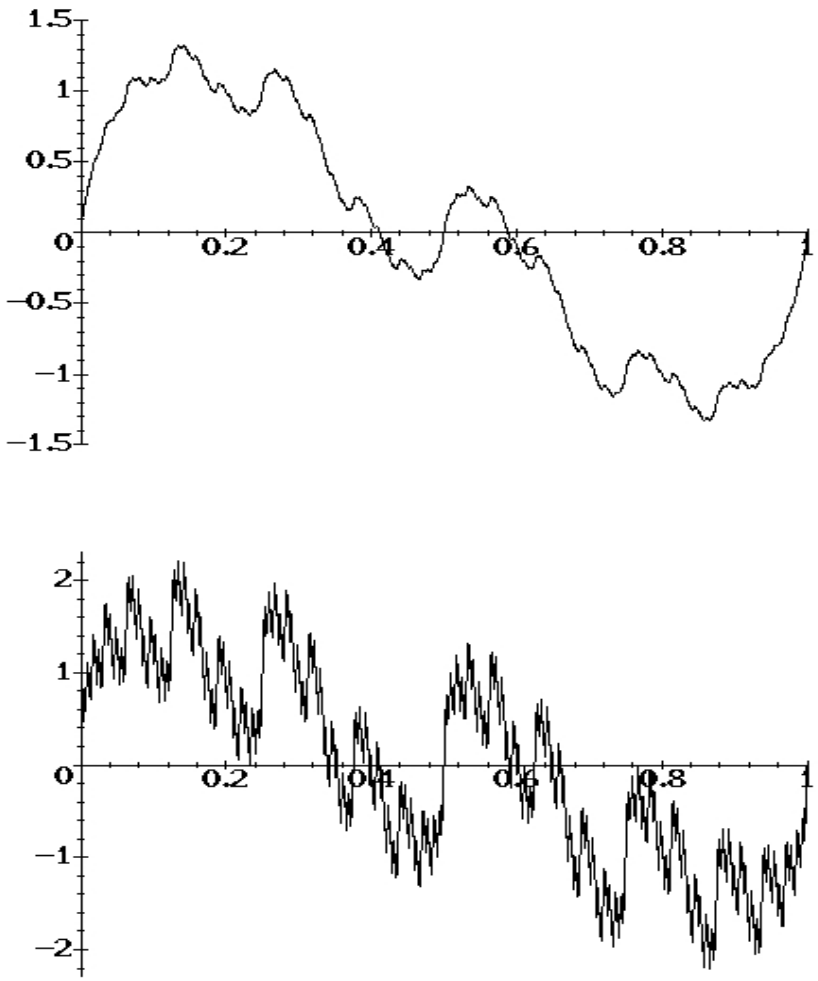


Figure 5.2. The Weierstraß functions $S_{1/2,2}$ and $S_{3/4,2}$.

replace the term $\frac{x}{2}$ in the functional equation by $\frac{x+1}{2}$, i.e., to condense the function not on the first half of the unit interval, but on the second half. This leads to

$$S_{a,2}\left(\frac{x+1}{2}\right) \;=\; a\,S_{a,2}(x) - \sin(\pi x)$$

for all $x \in [0,1]$, also easily verified.

Are these two functional equations together characteristic for the Weierstraß function? To answer this question, we consider functional equations of this type in general. For given constants $-1 < a_0, a_1 < 1$ and perturbation functions $g_0, g_1 : [0,1] \to \mathbb{R}$ consider the system (F) consisting of the two functional equations

$$f\left(\frac{x}{2}\right) \;=\; a_0 f(x) + g_0(x) \tag{F_0}$$

$$f\left(\frac{x+1}{2}\right) \;=\; a_1 f(x) + g_1(x) \tag{F_1}$$

for unknown $f : [0,1] \to \mathbb{R}$. (Special cases of such functional equations have been investigated by G. de Rham in the 1950's, see [95, 96].)

Generalizing in this way makes sense because it turns out that many other "strange" function besides the Weierstraß sine series $S_{a,2}$ can be found which satisfy such a system. It is for example immediate that the Weierstraß cosine series $C_{a,2}$ solves such a system if g_0, g_1 are chosen as $\pm\cos(\pi x)$. Another example is the so-called Takagi function, after the Japanese mathematician Teiji Takagi, who introduced this function in 1903 ([221]). It is defined, for $|a| < 1$, as the series

$$T_a(x) \;:=\; \sum_{n=0}^{\infty} a^n \, d\,(2^n x)\,,$$

where d is the 1-periodic saw-tooth function $d(x) := \operatorname{dist}(x, \mathbb{Z})$. This function (see Figure 5.3 for $a = \frac{1}{2}$ and $a = \frac{3}{4}$) is sometimes considered as the simplest cnd function possible; therefore it (or a variant) is often given as an example in introductory texts. It is again easy to check that this function satisfies the system (F) on $[0,1]$ with $a_0 = a_1 = a$, $g_0(x) = \frac{x}{2}$ and $g_1(x) = \frac{1-x}{2}$. More examples of cnd solutions of (F) with simple (differentiable) perturbation functions are given in [120, 121].

Now assume that f solves (F). Exploring, we find that if we put $x = 0$ into equation (F_0), then we get that necessarily $f(0) = \dfrac{g_0(0)}{1-a_0}$. Similarly, from ($\mathrm{F}_1$) with $x = 1$ we get that $f(1) = \dfrac{g_1(1)}{1-a_1}$. This can be continued. Putting $x = 1$ into Equation (F_0), we get

$$f\left(\frac{1}{2}\right) \;=\; a_0 f(1) + g_0(1) \;=\; a_0\,\frac{g_1(1)}{1-a_1} + g_0(1), \tag{5.1}$$

and putting $x = 0$ into Equation (F_1), we get

$$f\left(\frac{1}{2}\right) \;=\; a_1 f(0) + g_1(0) \;=\; a_1\,\frac{g_0(0)}{1-a_0} + g_1(0). \tag{5.2}$$

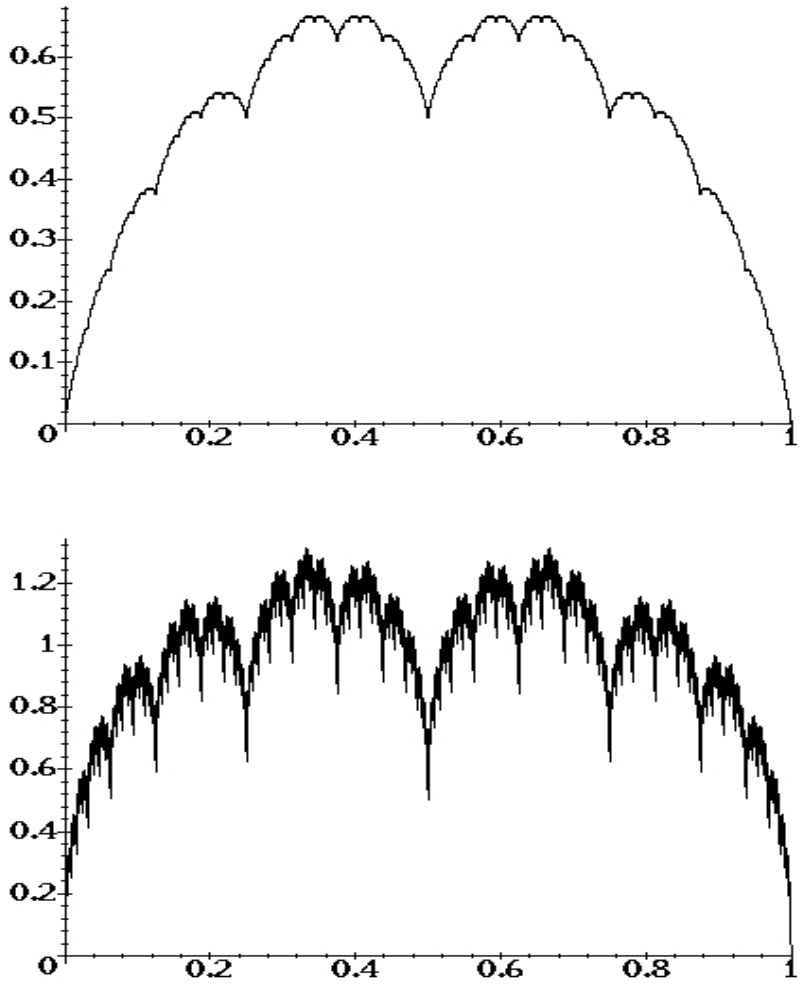


Figure 5.3. The Takagi functions $T_{1/2}$ and $T_{3/4}$.

Thus, if f solves (F), then the expressions in (5.1) and (5.2) must be equal. Therefore the condition

$$a_0 \frac{g_1(1)}{1-a_1} + g_0(1) \;=\; a_1 \frac{g_0(0)}{1-a_0} + g_1(0) \qquad (*)$$

must be satisfied if the system (F) has a solution.

We can now continue to compute values of the solution f at other points x out of already computed values. Using $f\left(\frac{1}{2}\right)$, we get $f\left(\frac{1}{4}\right) = a_0 f\left(\frac{1}{2}\right) + g_0\left(\frac{1}{2}\right)$ and $f\left(\frac{3}{4}\right) = a_1 f\left(\frac{1}{2}\right) + g_1\left(\frac{1}{2}\right)$. From this, we can then compute $f\left(\frac{1}{8}\right), f\left(\frac{3}{8}\right), f\left(\frac{5}{8}\right), f\left(\frac{7}{8}\right)$, then $f\left(\frac{2i+1}{16}\right)$, and so on. In this way, we see that the functional equations fix the values of f at all the dyadic rationals $f\left(\frac{i}{2^n}\right)$. Since these are dense in $[0,1]$, we see that if there is a continuous solution of (F), then it must be unique. By closer inspection of this argument we are now led to the following theorem which says that the condition $(*)$ above is not only necessary for existence of a solution, but under natural conditions also sufficient.

Theorem 1.1. *If the perturbation functions g_0, g_1 are continuous and condition $(*)$ is satisfied, then the system (F) has a unique continuous solution.*

Having realized that this theorem holds, it is now in fact not difficult to prove the theorem directly, e.g. with the use of Banach's fixed point theorem. Just define the operator Tf by

$$(Tf)(x) \;=\; \begin{cases} a_0 f(2x) + g_0(2x) & \text{for } x \in [0, \frac{1}{2}], \\ a_1 f(2x-1) + g_0(2x-1) & \text{for } x \in [\frac{1}{2}, 1], \end{cases}$$

and check that it satisfies the assumptions of Banach's theorem with a subset of $(C[0,1], \|\cdot\|_\infty)$ as the Banach space.

It is interesting to visualize the recursive procedure to compute values of the solution at the dyadic rationals, described above, on the computer: In the first step, we compute $f(0)$ and $f(1)$ and connect these points by a straight line in the x-y-plane. In the second step, we compute from these the value $f\left(\frac{1}{2}\right)$, and again connect the three known points on the graph of f by straight lines. We repeat this with $f\left(\frac{1}{4}\right)$ and $f\left(\frac{3}{4}\right)$ in the next step, then $f\left(\frac{1}{8}\right), f\left(\frac{3}{8}\right), f\left(\frac{5}{8}\right), f\left(\frac{7}{8}\right)$, and so on. This leads to a sequence of piecewise affine approximations $f^{(n)}$ of f, where $f^{(n)}$ equals f on the dyadic rationals $\frac{i}{2^n}$, see Figure 5.4 (left).

So far, this may look straightforward, but it is already a big step towards a new understanding of the Weierstraß function. The next step now is an insight: the insight to look at the data in a different way which will in fact turn out to be essential for the development of what follows. Thus a general rule about experimental mathematics is confirmed once more: Being able to compute and draw quickly is not sufficient; necessary is an active experimenter's mind, interacting with the results and figures and adding ideas to the mix.

In this case, the idea is to look not only at how the approximations approach the solution f, but to ask oneself what is added in each step. Thus on the right-hand side of Figure 5.4, we see the differences $f_n :=$

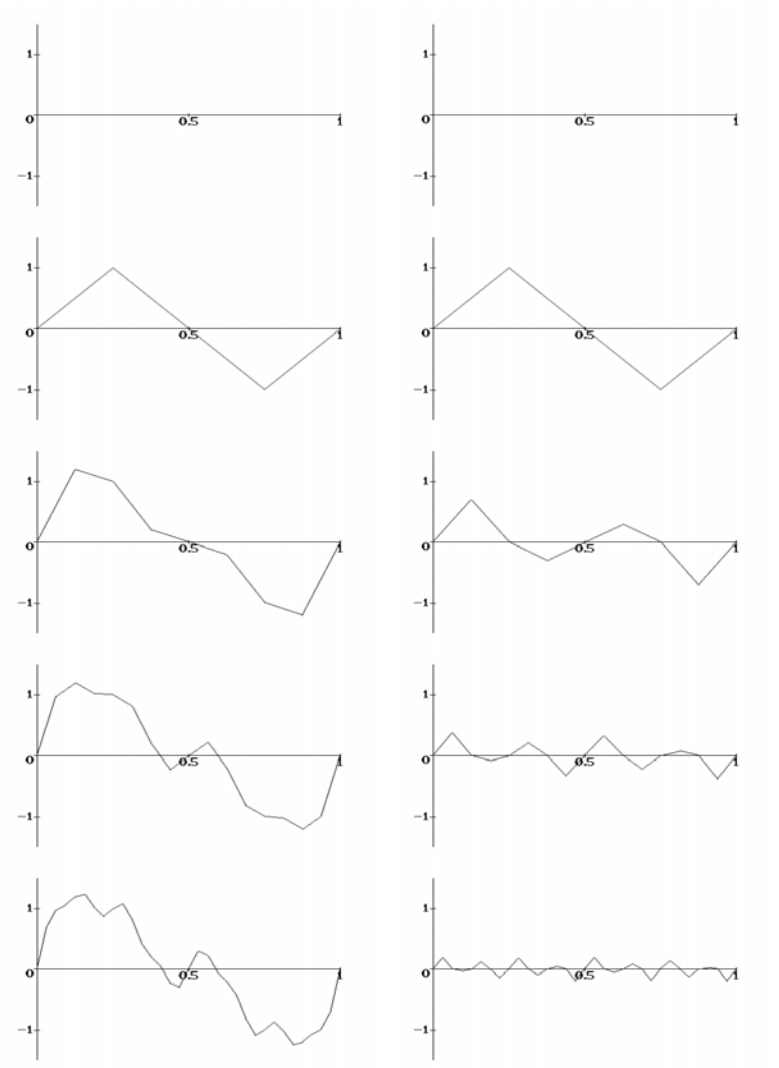


Figure 5.4. Piecewise affine approximations to the Weierstraß function $S_{1/2,2}$.

$f^{(n)} - f^{(n-1)}$. As expected, these are sawtooth functions (remember that $f^{(n)}\left(\frac{i}{2^n}\right) = f^{(n+1)}\left(\frac{i}{2^n}\right) = f\left(\frac{i}{2^n}\right)$). What is not expected in such irregular functions as the Weierstraß functions is the visual regularity of the f_n, apparent in Figure 5.4: each of these differences, especially in the later figures, seems just to consist of two scaled-down copies of its predecessor.

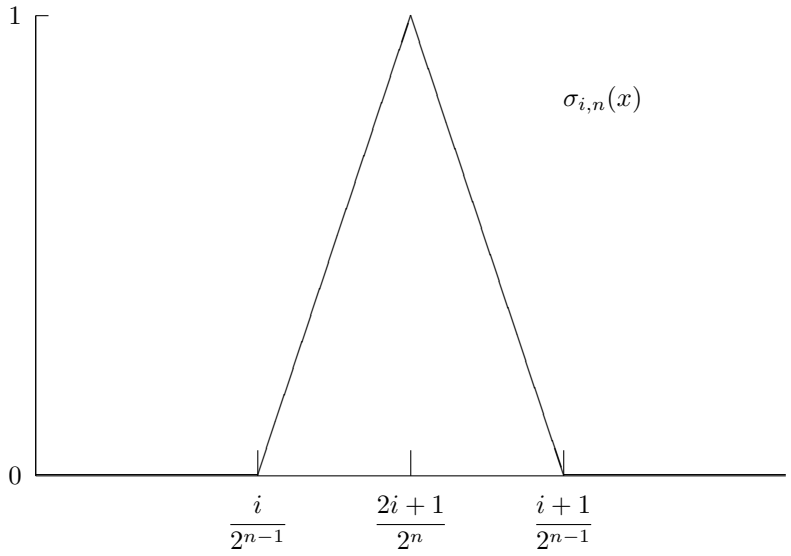


Figure 5.5. An element $\sigma_{i,n}$ of the Schauder basis.

Having discovered such a pattern where none was expected, its origins and implications should now be pursued.

1.2 Schauder Bases

The mathematical tool that we need to explore these patterns is the classical Schauder basis of the space $C[0,1]$. It was introduced by J. Schauder in 1927 ([210]). Define functions $\sigma_{i,n} : [0,1] \to \mathbb{R}$ by $\sigma_{0,0}(x) = 1 - x$, $\sigma_{1,0}(x) = x$, and, for $n \in \mathbb{N}$ and $i = 0, \ldots, 2^{n-1} - 1$, $\sigma_{i,n}$ as the piecewise linear function connecting the points $(0,0)$, $\left(\frac{i}{2^{n-1}}, 0\right)$, $\left(\frac{2i+1}{2^n}, 1\right)$, $\left(\frac{i+1}{2^{n-1}}, 0\right)$, $(1,0)$; see Figure 5.5.

Then it is known (and not difficult to prove directly) that every $f \in C[0,1]$ has a unique, uniformly convergent expansion of the form

$$f(x) \;=\; \gamma_{0,0}(f)\,\sigma_{0,0}(x) + \gamma_{1,0}(f)\,\sigma_{1,0}(x) + \sum_{n=1}^{\infty} \sum_{i=0}^{2^{n-1}-1} \gamma_{i,n}(f)\,\sigma_{i,n}(x),$$

where the coefficients $\gamma_{i,n}(f)$ are given by

$$\gamma_{0,0}(f) \;=\; f(0), \quad \gamma_{1,0}(f) = f(1), \quad \text{and}$$

$$\gamma_{i,n}(f) \;=\; f\left(\frac{2i+1}{2^n}\right) - \frac{1}{2}\,f\left(\frac{i}{2^{n-1}}\right) - \frac{1}{2}\,f\left(\frac{i+1}{2^{n-1}}\right).$$

For given $f \in C[0,1]$, let

$$f_n(x) \;:=\; \sum_{i=0}^{2^{n-1}-1} \gamma_{i,n}(f)\,\sigma_{i,n}(x)$$

and

$$f^{(n)}(x) \;:=\; \gamma_{0,0}(f)\,\sigma_{0,0}(x) + \gamma_{1,0}(f)\,\sigma_{1,0}(x) + \sum_{k=1}^{n} f_k(x)$$

and

$$r_n(x) \;:=\; f(x) - f^{(n)}(x) \;=\; \sum_{k=n+1}^{\infty} f_k(x).$$

Note that $r_n(x) = 0$ for all dyadic rationals of the form $x = \frac{i}{2^n}$. Remember also that Figure 5.4 shows the $f^{(n)}$'s and f_n's for the Weierstraß. Thus the pattern discovered in that figure visually translates into a (recursive) pattern of the $\gamma_{i,n}(f)$'s which we will explore in the next subsection.

For our purposes it is important that differentiability properties of a function can be established by looking at its Schauder coefficients. Statements of this type have apparently been noticed for the first time by G. Faber in 1910 (who used the Schauder basis prior to Schauder but in a different mathematical guise; see [110, 111]). Faber proved the following theorem.

Theorem 1.2. *Assume that $f \in C[0,1]$ has a finite derivative at some point x_0. Then*

$$\lim_{n\to\infty} 2^n \cdot \min\left\{|\gamma_{i,n}(f)| : i = 0, \ldots, 2^{n-1}-1\right\} \;=\; 0.$$

Proof. Note that if $f'(x_0)$ exists, then

$$f'(x_0) \;=\; \lim_{n\to\infty} \frac{f(v_n) - f(u_n)}{v_n - u_n}$$

for all sequences $(u_n), (v_n) \subseteq [0,1]$ with $u_n \le x_0 \le v_n$ and $u_n < v_n$ and $v_n - u_n \to 0$.

Now for given $x_0 \in [0,1]$ and $n \in \mathbb{N}$ choose the dyadic rationals u_n, v_n as shown in Figure 5.6 (possibly $x_0 = u_n$ or $x_0 = v_n$ if x_0 is itself a dyadic rational). If $f'(x_0) \in \mathbb{R}$ exists, then

$$\begin{aligned} f'(x_0) &= \lim_{n\to\infty} \frac{f(v_n) - f(u_n)}{v_n - u_n} \\ &= \lim_{n\to\infty} \left[\sum_{k=1}^{n} \frac{f_k(v_n) - f_k(u_n)}{v_n - u_n} + \frac{r_n(v_n) - r_n(u_n)}{v_n - u_n} \right] \\ &= \lim_{n\to\infty} \left[\sum_{k=1}^{n} \pm\delta_{i_k,k}(f) + 0 \right], \end{aligned}$$

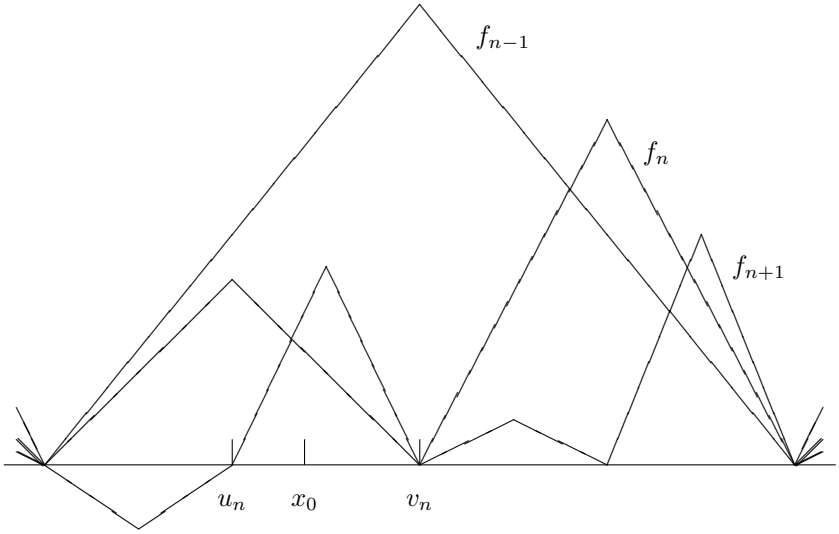


Figure 5.6. Definition of the points u_n and v_n.

for some sequence i_k, where $\delta_{i,k}(f)$ is the slope of the Schauder triangle $\gamma_{i,k}(f)\,\sigma_{i,k}(x)$ in the expansion of f; thus $\delta_{i,k}(f) = 2^k\gamma_{i,k}(f)$.

This implies that the series $\sum_{k=1}^{\infty} \pm 2^k \gamma_{i_k,k}(f)$ is convergent, hence the summand $\pm 2^k\gamma_{i_k,k}(f)$ necessarily tends to 0, and therefore also

$$\lim_{k\to\infty} 2^k \min_i |\gamma_{i,k}(f)| = 0.$$

□

Note that non-differentiability of the Takagi function T_a, whose Schauder coefficients satisfy $\gamma_{i,n}(T_a) = a^n$, directly follows from this theorem.

1.3 Non-differentiability

By the results of the previous subsection, we can infer differentiability properties of an $f \in C[0,1]$ from its Schauder coefficients. But how can we compute the Schauder coefficients of, say, the Weierstraß functions? In subsection 1.1 we had already discovered a pattern in these coefficients. With the notation from subsection 1.2 we can now make this discovery more precise and then use it.

The pattern in fact translates into a recursion formula for the $\gamma_{i,n}(f)$. Since the functional equation was instrumental in discovering the pattern,

it is now no surprise that this recursion formula is based on the functional equation as well. Now formulation and proof of the next theorem are straightforward.

Theorem 1.3. *Assume* $(*)$ *and let* f *be the continuous solution of the system (F) with continuous* g_0, g_1*. Then*
(i) $\gamma_{0,0}(f) = f(0) = \frac{g_0(0)}{1-a_0}$ *and* $\gamma_{1,0}(f) = f(1) = \frac{g_1(1)}{1-a_1}$,

$$\begin{aligned} \text{(ii) } \gamma_{0,1}(f) &= \left(a_1 - \tfrac{1}{2}\right) f(0) - \tfrac{1}{2} f(1) + g_1(0) \\ &= \left(a_0 - \tfrac{1}{2}\right) f(1) - \tfrac{1}{2} f(0) + g_0(1), \end{aligned}$$

$$\begin{aligned} \text{(iii) } \gamma_{i,n+1}(f) &= a_0 \gamma_{i,n}(f) + \gamma_{i,n}(g_0) && \text{for } i = 0, \ldots, 2^{n-1} - 1, \\ \gamma_{i,n+1}(f) &= a_1 \gamma_{i-2^{n-1},n}(f) + \gamma_{i-2^{n-1},n}(g_1) && \text{for } i = 2^{n-1}, \ldots, 2^n - 1. \end{aligned}$$

If we let $\delta_{i,n}(f) := 2^n \gamma_{i,n}(f)$, then the recursion step (iii) of the theorem becomes

$$\begin{aligned} \delta_{i,n+1}(f) &= 2a_0 \delta_{i,n}(f) + 2\delta_{i,n}(g_0) && \text{for } i = 0, \ldots, 2^{n-1} - 1, \\ \delta_{i,n+1}(f) &= 2a_1 \delta_{i-2^{n-1},n}(f) + 2\delta_{i-2^{n-1},n}(g_1) && \text{for } i = 2^{n-1}, \ldots, 2^n - 1. \end{aligned}$$

Also, let $\underline{\delta}_n(f) := \min_i |\delta_{i,n}(f)|$. If it can be proved that $\underline{\delta}_n(f) \not\to 0$, then f must be non-differentiable by Theorem 1.2.

It is now relatively easy to check this condition for the Weierstraß function $f = S_{a,2}$ via the recursion. In Table 5.1 some of its Schauder coefficients and their minimum are listed. The numbers strongly suggest that the minimum indeed does not tend to 0, thus numerically already confirming the function's non-differentiability. Our task now is to prove this strictly.

The Weierstraß function $f = S_{a,2}$ satisfies the system (F) with $a_0 = a_1 = a$ and $g_0(x) = -g_1(x) = \sin(\pi x)$, and thus

$$\delta_{i,n}(g_0) \;=\; -\delta_{i,n}(g_1) \;=\; 2^n \sin\left(\pi \frac{2i+1}{2^n}\right) \cdot \left(1 - \cos\frac{\pi}{2^n}\right).$$

Using the recursion step and estimating, we get the following recursive estimate for the minima $\underline{\delta}_n(f)$:

$$\underline{\delta}_{n+1}(f) \;\geq\; 2\,|a| \cdot \underline{\delta}_n(f) - 2^{n+1} \cos\frac{\pi}{2^n} \left(1 - \cos\frac{\pi}{2^n}\right).$$

Iterating this and then doing a final estimate, we get

$$\underline{\delta}_{n+k}(f) \;\geq\; (2\,|a|)^k \left[\underline{\delta}_n(f) - 2^n \left(1 - \cos\frac{\pi}{2^{n-1}}\right)\right].$$

This says that if we can find an index n such that $\underline{\delta}_n(f) > 2^n \left(1 - \cos\frac{\pi}{2^{n-1}}\right)$, then it follows that $\underline{\delta}_k(f) \to \infty$ for $k \to \infty$. Such an index can easily be

n	$\delta_{0,n}$	$\delta_{1,n}$	$\delta_{2,n}$	$\delta_{3,n}$	$\delta_{4,n}$	$\underline{\delta}_n$
1	0.000000					0.000000
2	4.000000	−4.000000				4.000000
3	5.656854	−2.343146	2.343146	−5.656854		2.343146
4	6.122935	−1.217927	3.468364	−5.190774	5.190774	1.217927
5	6.242890	−0.876323	3.979611	−4.587717	5.793830	0.876323
6	6.273097	−0.786864	4.124884	−4.392211	6.032054	0.786864
7	6.280662	−0.764241	4.162348	−4.340269	6.097976	0.764241
8	6.282555	−0.758569	4.171786	−4.327088	6.114869	0.732594
9	6.283028	−0.757150	4.174150	−4.323780	6.119118	0.704348
10	6.283146	−0.756795	4.174741	−4.322952	6.120182	0.696645
11	6.283175	−0.756707	4.174889	−4.322745	6.120448	0.694677
12	6.283183	−0.756684	4.174926	−4.322693	6.120515	0.693818

Table 5.1. Some Schauder coefficients for $S_{1/2,2}$.

found by computing the δ_n's for $S_{a,2}$ and comparing. We get the following results:

$$\begin{array}{lcll} \underline{\delta}_3\left(S_{a,2}\right) & = & 8\left|a\right| - 4\sqrt{2} + 4 & \\ & > & 8 - 4\sqrt{2} \;=\; 2^3\left(1 - \cos\frac{\pi}{2^2}\right), & \text{for } \left|a\right| > \frac{1}{2}; \\ \underline{\delta}_5\left(S_{a,2}\right) & > & 2^5\left(1 - \cos\frac{\pi}{2^4}\right), & \text{for } a = \frac{1}{2}; \\ \underline{\delta}_4\left(S_{a,2}\right) & > & 2^4\left(1 - \cos\frac{\pi}{2^3}\right), & \text{for } a = -\frac{1}{2}. \end{array}$$

Thus the non-differentiability of the Weierstraß sine series $S_{a,2}$ is completely proved. □

Note that this is more than a proof: it is a method. It can be used to examine any non-differentiable function f that satisfies a system of type (F) with simple perturbation functions (simple enough so that the recursion for the Schauder coefficients of f is manageable). Another, rather easy example is given by the Takagi function. Its perturbation functions g_0, g_1 are linear, so that their Schauder coefficients vanish. If we did not know that the function was already defined via its Schauder expansion, then the recursion would almost instantly show that the condition of Theorem 1.2 is fulfilled. The Weierstraß cosine series $C_{a,2}$ can also be analyzed with this method. Matters are slightly more complicated, however, since in this case it turns out that $\underline{\delta}_n\left(C_{a,2}\right) \to 0$ (see Table 5.2). Thus for this example the method has to be refined a bit, making matters slightly more complicated (but not inherently more difficult). Full details are given in [120, 121]. With such modifications, the Weierstraß cosine series and many other non-differentiable functions become part of this method, their treatment now being possible in a unified (and quite algebraic, i.e., computational) way.

n	$\delta_{0,n}$	$\delta_{1,n}$	$\delta_{2,n}$	$\delta_{3,n}$	$\delta_{4,n}$	$\underline{\delta}_n$
1	−4.000000					4.000000
2	−4.000000	−4.000000				4.000000
3	−2.343146	−5.656854	−5.656854	−2.343146		2.343146
4	−1.217927	−5.190774	−6.122935	−3.468364	−3.468364	1.217927
5	−0.614871	−4.679527	−5.781331	−3.348409	−3.588319	0.614871
6	−0.308177	−4.384620	−5.509543	−3.110184	−3.392814	0.308177
7	−0.154182	−4.232107	−5.359982	−2.965015	−3.253435	0.154182
8	−0.077102	−4.155213	−5.283459	−2.889048	−3.178206	0.077102
9	−0.038553	−4.116687	−5.244979	−2.850638	−3.139888	0.038553
10	−0.019277	−4.097413	−5.225712	−2.831379	−3.120641	0.019277
11	−0.009638	−4.087776	−5.216074	−2.821743	−3.111007	0.009638
12	−0.004819	−4.082956	−5.211255	−2.816924	−3.106188	0.004819

Table 5.2. Some Schauder coefficients for $C_{1/2,2}$.

2 Bernoulli Convolutions

Consider the discrete probability density on the real line with measure $1/2$ at each of the two points ± 1. The corresponding measure is the so-called Bernoulli measure, denoted $b(X)$. For every $0 < q < 1$, the infinite convolution of measures

$$\mu_q(X) \quad = \quad b(X) * b(X/q) * b(X/q^2) * \cdots \tag{5.3}$$

exists as a weak limit of the finite convolutions. The most basic theorem about these infinite Bernoulli convolutions is due to Jessen and Wintner ([144]). They proved that μ_q is always continuous, and that it is either absolutely continuous or purely singular. This statement follows from a more general theorem on infinite convolutions of purely discontinuous measures (Theorem 35 in [144]); however, it is not difficult to prove the statement directly with the use of Kolmogoroff's 0-1-law. The problem is to decide for which values of the parameter q the measure is singular, and for which q it is absolutely continuous.

This question can be recast in a more real-analytic way by defining the distribution function F_q of μ_q as

$$F_q(t) \quad = \quad \mu_q(-\infty, t], \tag{5.4}$$

and to ask for which q this continuous, increasing function F_q is singular, and for which it is absolutely continuous. Note that F_q satisfies $F_q(t) = 0$ for $t < -1/(1-q)$ and $F_q(t) = 1$ for $t > 1/(1-q)$.

Approaching this question experimentally, the first step would be to draw these functions or their derivatives for some well-chosen values of q. However, there is no easy way of computing (let alone graphing) the distribution function or density of a probability measure. One possibility that has been proposed is to generate random numbers which follow this distribution (or an approximation thereof) and count how often these numbers fall into each interval of a partition of $\mathbb{R}$. This method, however, is relatively slow and imprecise.

Thus we once more need an idea. After the success we had with functional equations in the treatment of non-differentiable functions, this idea is to use functional equations for the present problem as well.

In fact, it turns out that functional equations can be used to directly define the distribution function F_q: It can be proved that F_q is the only bounded solution of the functional equation

$$F(t) \quad = \quad \frac{1}{2}F\left(\frac{t-1}{q}\right) + \frac{1}{2}F\left(\frac{t+1}{q}\right) \tag{5.5}$$

with the above restrictions. Moreover, if F_q is absolutely continuous and thus has a density $f_q \in L^1(\mathbb{R})$, then f_q satisfies the functional equation

$$2q\, f(t) \;=\; f\left(\frac{t-1}{q}\right) + f\left(\frac{t+1}{q}\right) \tag{5.6}$$

almost everywhere. This functional equation is a special case of a much more general class of equations, namely two-scale difference equations. Those are functional equations of the type

$$f(t) \;=\; \sum_{n=0}^{N} c_n\, f(\alpha t - \beta_n) \qquad (t \in \mathbb{R}), \tag{5.7}$$

with $c_n \in \mathbb{C}$, $\beta_n \in \mathbb{R}$ and $\alpha > 1$. They were first discussed by Ingrid Daubechies and Jeffrey C. Lagarias, who proved existence and uniqueness theorems and derived some properties of L^1-solutions [92, 93]. One of their theorems, which we state here in part for the general equation (5.7) and in part for the specific case (5.6), is the following:

Theorem 2.1. *(a) If $\alpha^{-1}(c_0 + \cdots + c_N) = 1$, then the vector space of $L^1(\mathbb{R})$-solutions of (5.7) is at most one-dimensional.*

(b) If, for given $q \in (0,1)$, equation (5.6) has a nontrivial L^1-solution f_q, then its Fourier transform satisfies $\widehat{f_q}(0) \neq 0$, and is given by

$$\widehat{f_q}(x) \;=\; \widehat{f_q}(0) \prod_{n=0}^{\infty} \cos(q^n\, x). \tag{5.8}$$

In particular, for normalization we can assume $\widehat{f_q}(0) = 1$.

(c) On the other hand, if the right-hand side of (5.8) is the Fourier transform of an L^1-function f_q, then f_q is a solution of (5.6).

(d) Any nontrivial L^1-solution of (5.7) is finitely supported. In the case of (5.6), the support of f_q is contained in $[-1/(1-q), 1/(1-q)]$.

This implies in particular that the question of whether the infinite Bernoulli convolution (5.3) is absolutely continuous is equivalent to the question of whether (5.6) has a nontrivial L^1-solution. Now what is known about these questions?

It is relatively easy to see that in the case $0 < q < 1/2$, the solution of (5.5) is singular; it is in fact a Cantor function, meaning that it is constant on a dense set of intervals. This was first proved by R. Kershner and A. Wintner [146].

It is also easy to see that in the case $q = 1/2$, there is an L^1-solution of (5.6), namely $f_{1/2}(t) = \frac{1}{4}\,\chi_{[-2,2]}(t)$. Moreover, this function can be used to construct a solution for every $q = 2^{-1/p}$ where p is an integer, namely

$$f_{2^{-1/p}}(t) = 2^{(p-1)/2} \cdot \Big[f_{1/2}(t) * f_{1/2}(2^{1/p}t) * \cdots * f_{1/2}(2^{(p-1)/p}t) \Big]. \quad (5.9)$$

This was first noted by Wintner via the Fourier transform [230],

$$\widehat{f}(x) \;=\; \int_{-\infty}^{\infty} f(t)\, e^{-ixt}\, dt.$$

Explicitly, we have

$$\begin{aligned} \widehat{f_{2^{-1/p}}}(x) \;&=\; \prod_{n=0}^{\infty} \cos(2^{-n/p}x) \;=\; \prod_{m=0}^{\infty} \prod_{k=0}^{p-1} \cos(2^{-(m+k/p)}x) \\ &=\; \widehat{f_{1/2}}(x) \cdot \widehat{f_{1/2}}(2^{-1/p}x) \cdots \widehat{f_{1/2}}(2^{-(p-1)/p}x), \end{aligned}$$

which is equivalent to (5.9) by the convolution theorem.

Note that the regularity of these solutions $f_{2^{-1/p}}$ increases when p and thus $q = 2^{-1/p}$ increases: $f_{2^{-1/p}} \in C^{p-2}(\mathbb{R})$. From the results given so far, one might therefore surmise that (5.6) would have a nontrivial L^1-solution for every $q \geq 1/2$ with increasing regularity when q increases. This supposition, however, would be wrong, and it came as a surprise when Erdős proved in 1939 [102] that there are some values of $1/2 < q < 1$ for which (5.6) does *not* have an L^1-solution, namely, the inverses of Pisot numbers. A *Pisot number* is defined to be an algebraic integer greater than 1 all of whose algebraic conjugates lie inside the unit disk. The best known example of a Pisot number is the golden mean $\varphi = (\sqrt{5}+1)/2$. The characteristic property of Pisot numbers is that their powers quickly approach integers: If a is a Pisot number, then there exists a θ, $0 < \theta < 1$, such that

$$\operatorname{dist}(a^n, \mathbb{Z}) \;\leq\; \theta^n \quad \text{for all } n \in \mathbb{N}. \quad (5.10)$$

Erdős used this property to prove that if $q = 1/a$ for a Pisot number a, then $\limsup_{x\to\infty} \left|\widehat{f}_q(x)\right| > 0$. Thus in these cases, f_q cannot be in $L^1(\mathbb{R})$ since that would contradict the Riemann-Lebesgue lemma. Erdős's proof uses the Fourier transform $\widehat{f}_q$: Consider, for $N \in \mathbb{N}$,

$$\left|\widehat{f}_q(q^{-N}\pi)\right| \;=\; \prod_{n=1}^{\infty} |\cos(q^n\pi)| \cdot \prod_{n=0}^{N-1} \left|\cos(q^{-n}\pi)\right| \;=:\; C \cdot p_N,$$

where $C > 0$. Moreover, choose $\theta \neq 1/2$ according to (5.10) and note that

$$\begin{aligned} p_N &= \prod_{\substack{n=0 \\ \theta^n \leq 1/2}}^{N-1} \left|\cos(q^{-n}\pi)\right| \cdot \prod_{\substack{n=0 \\ \theta^n > 1/2}}^{N-1} \left|\cos(q^{-n}\pi)\right| \\ &\geq \prod_{\substack{n=0 \\ \theta^n \leq 1/2}}^{N-1} \cos(\theta^n\pi) \cdot \prod_{\substack{n=0 \\ \theta^n > 1/2}}^{N-1} \left|\cos(q^{-n}\pi)\right| \\ &\geq \prod_{\substack{n=0 \\ \theta^n \leq 1/2}}^{\infty} \cos(\theta^n\pi) \cdot \prod_{\substack{n=0 \\ \theta^n > 1/2}}^{\infty} \left|\cos(q^{-n}\pi)\right| \;=\; C' \;>\; 0, \end{aligned}$$

independently of N. □

In 1944, Raphaël Salem [209] showed that the reciprocals of Pisot numbers are the only values of q where $\widehat{f_q}(x)$ does not tend to 0 for $x \to \infty$. In fact, no other $q > 1/2$ are known at all where F_q is singular. Moreover, the set of explicitly given q with absolutely continuous F_q is also not very big: The largest such set known to date was found by Adriano Garsia in 1962 [117]. It contains reciprocals of certain algebraic numbers (such as roots of the polynomials $x^{n+p} - x^n - 2$ for $\max\{p, n\} \geq 2$) besides the roots of 1/2. More recently, in [94], "bad behaviour", such as unboundedness or non-membership to L^2, of the density, if it exists, is established for certain classes of algebraic numbers.

The most recent significant progress then was made in 1995 by Boris Solomyak [213], who developed new methods in geometric measure theory which he then used to prove that F_q is in fact absolutely continuous for almost every $q \in [1/2, 1)$. (See also [189] for a simplified proof and [188] for a survey and some newer results.)

Thus the set of $q \in [1/2, 1)$ with absolutely continuous F_q has measure 1/2, but does not equal the whole interval. Which numbers are in this set, which are not? Apart from the results cited above, nothing specific is known; in particular, membership to this set has been decided only for some algebraic numbers.

Our goal here is to generate drawings of the functions F_q, or rather, since more structure is visible, of their densities f_q. This is helpful in obtaining some better visual feeling for their properties. Interestingly, drawing good approximations to these functions is not easy or immediate. It turns out that use of the functional equations appears to be a good way to generate pictures.

In fact, define a map B_q, mapping the set D of L^1-functions f with

support in $[-1/(1-q), 1/(1-q)]$ and with $\widehat{f}(0) = 1$ into itself, by

$$(B_q f)(t) \;=\; \frac{1}{2q}\left(f\left(\frac{t-1}{q}\right) + f\left(\frac{t+1}{q}\right)\right) \quad \text{for } t \in \mathbb{R}.$$

One can picture the action of the operator B_q on some $f \in \mathcal{D}$ as putting two rescaled copies of f into the two corners of $[-1/(1-q), 1/(1-q)]$ and adding them.

Now note that the fixed points of B_q are the solutions of (5.6) and that B_q is nonexpansive. Therefore, one may have hope that by iterating the operator, it may be possible to approximate the solution. Explicitly, we are interested in the iteration

$$\begin{aligned} f^{(0)} \in \mathcal{D} \;&:=\; \left\{ f \in L^1(\mathbb{R}) :\; \operatorname{supp} f \subseteq \left[-\tfrac{1}{1-q}, \tfrac{1}{1-q}\right],\; \widehat{f}(0) = 1 \right\}, \\ f^{(n)} \;&:=\; B_q\, f^{(n-1)} \text{ for } n \in \mathbb{N}. \end{aligned} \tag{5.11}$$

Since B_q is continuous, it is clear that if a sequence of iterates $B_q^n f^{(0)}$ converges in $L^1(\mathbb{R})$ for some initial function $f^{(0)} \in \mathcal{D}$, then the limit will be a fixed point of B_q. No general proof of convergence is known, even under the assumption that a fixed point exists. Under this assumption, it is, however, possible to prove a weaker result, namely convergence in the mean: *If a solution $f_q \in L^1(\mathbb{R})$ with $\widehat{f_q}(0) = 1$ of (5.6) exists, then for every initial function $f^{(0)} \in \mathcal{D}$, we have*

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{k=0}^{n-1} B_q^k f^{(0)} - f_q \right\|_1 \;=\; 0.$$

This theorem follows from properties of Markov operators [161] and from a result by Mauldin and Simon [175], showing that if an L^1-density f_q exists then it must be positive a.e. on its support.

Thus to summarize: *If the iteration (5.11) converges in $L_1(\mathbb{R})$ for some $f^{(0)} \in \mathcal{D}$, then the limit is a solution of (5.6). If, on the other hand, (5.6) has a solution $f_q \in \mathcal{D}$, then the iteration (5.11) converges in the mean to f_q for every $f^{(0)} \in \mathcal{D}$.*

The iterate $f^{(n)} := B_q^n f^{(0)}$ can of course be expressed explicitly: Let $S_n := \{\pm 1 \pm q \pm q^2 \pm \cdots \pm q^{n-1}\}$. Then

$$(B_q^n f^{(0)})(t) \;=\; \frac{1}{(2q)^n} \sum_{s \in S_n} f^{(0)}\left(\frac{t+s}{q^n}\right).$$

The most natural choice for $f^{(0)}$ in this context seems to be $f^{(0)} := \frac{1-q}{2}\,\chi_{[-1/(1-q),1/(1-q)]}$. Then we get: *If the limit*

$$\lim_{n\to\infty} \frac{1-q}{2\cdot(2q)^n} \sum_{s\in S_n} \chi_{[-\frac{q^n}{(1-q)}-s,\, \frac{q^n}{(1-q)}-s]} \tag{5.12}$$

exists in $L^1(\mathbb{R})$, then it is the solution in $\mathcal{D}$ of (5.6).

Thus we now have to different methods to approximate and graph solutions to (5.6): Either by iterating B_q or by directly plotting the function in (5.12) for some n. Interestingly, a mixture of both methods seems to work best with respect to computing time, dramatically better, in fact, than either of the two methods alone. The trick is to first compute an approximation by (5.12) and then to use this as a starting function for the iteration.

In this way, Figures 5.7 and 5.8 have been computed. They show the indicated iterates, where the nth iterate is defined as $B_q^n f^{(0)}$ with, as above, $f^{(0)} := \frac{1-q}{2}\,\chi_{[-1/(1-q),1/(1-q)]}$. Each function is evaluated at $2^{13}+1$ points, equidistantly spanning the interval $[-1/(1-q), 1/(1-q)]$. In the left-hand columns, the evaluated points $(t, f(t))$ are connected by lines (presupposing continuity); in the right-hand columns, they are just marked by dots. The pictures show how regularity of the solutions (if they exist) generally increases with q, but with exceptions.

For $q = 1/2$, $q = \sqrt{1/2}$ and $q = 3/4$, convergence of the iteration seems relatively clear from the pictures. The picture for $q = 1/2$ is only there to provide an anchor; in this case our $f^{(0)}$ is already the fixed point. The picture for $q = \sqrt{1/2}$ shows the solution given by (5.9). Note, however, that this solution was computed by iterating B_q with a step function a starting point. Thus the picture clearly shows convergence of this iteration, even without taking the mean. For $q = 3/4$, the iteration seems to behave in exactly the same manner, with a continuous limit function.

For $q = 0.6$, $q = (\sqrt{5}-1)/2$ and $q = 2/3$, the situation seems more complicated. For the golden mean, the iteration cannot converge to any meaningful function, and the figure shows this (but compare Figure 5.1 which shows an approximation to a *continuous* function!). For $q = 0.6$, the figure looks only marginally better, so that the situation is unclear: Does the figure show approximation to an L^1-function? Or even to a continuous function? Finally for $q = 2/3$, the iterate seems to be much closer to a definite, maybe even continuous function which would then be a solution of (5.6).

Thus investigation of figures such as these leads to experimentally derived conjectures, questions and ideas for further exploration. We close this chapter with some of those.

Convergence of the iteration. Can convergence (pointwise or uniform or in L^1) of the iteration (5.11) with arbitrary or specific $f^{(0)}$ be proved in the case $q = 2^{-1/p}$? (The picture indicates this for $p = 2$; for $p = 1$, it is in general not true—for example, for $f^{(0)} = \frac{1}{2}\left(1 - \frac{|x|}{2}\right)$ every point in $[-2, 2] \times [0, \frac{1}{2}]$ is then a limit point of some sequence $(x_n, f^{(n)}(x_n))$.) Can convergence be proved for other specific values of q (such as $q = 3/4$,

as the picture indicates)? Can convergence of the iteration be proved under the sole assumption that a fixed point exists?

Boundedness of the iteration. Can a value of q be identified where the sequence of iterates remains bounded (for arbitrary or specific $f^{(0)}$)? Are there values of q with unbounded iteration? What about the golden mean or any other Pisot number, specifically?

Continuity of solutions. The figures definitely indicate existence of a continuous solution for $q = 3/4$, almost certainly for $q = 2/3$ and maybe for $q = 0.6$. Comparison with Figures 5.2 and 5.3 seems to show a comparable small-scale behaviour of the functions shown there for $a = 3/4$ with the iterate shown here for $q = 2/3$. Can some correspondence be established between equation (5.6) and functional equations of the type (F) as investigated there?

Existence of a limit function for $q \to 1$. Is it possible that the solutions f_q (so far they exist), *suitably rescaled,* converge to some well-defined function for $q \to 1$, either pointwise or in some other sense?

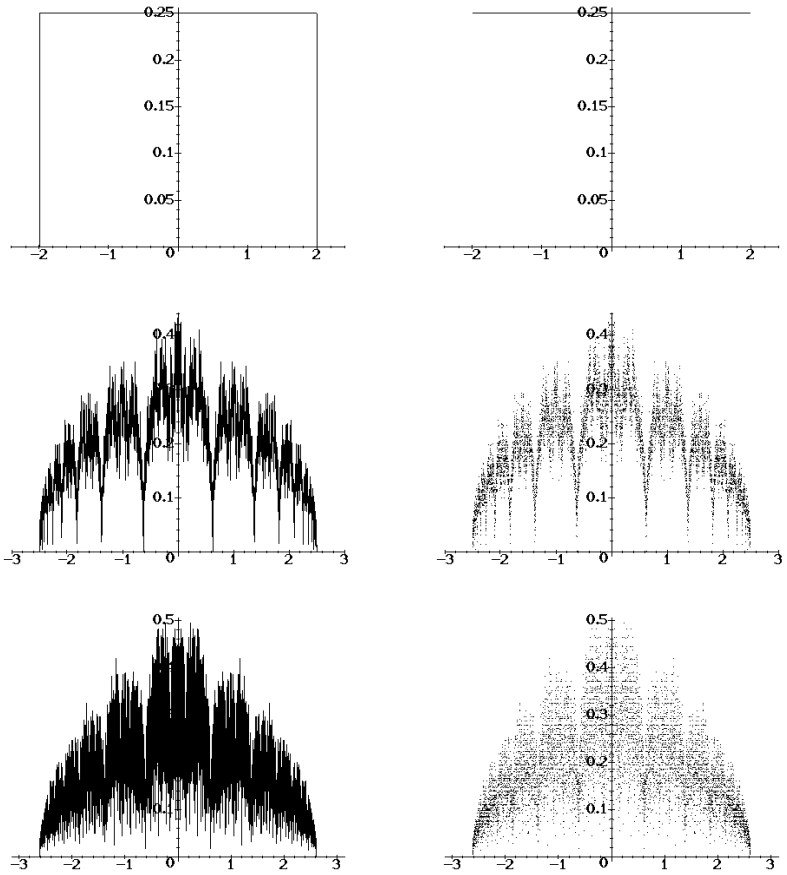


Figure 5.7. 26th iterate for $q = 1/2$ (top), $q = 0.6$ (middle) and $q = (\sqrt{5}-1)/2$ (bottom).

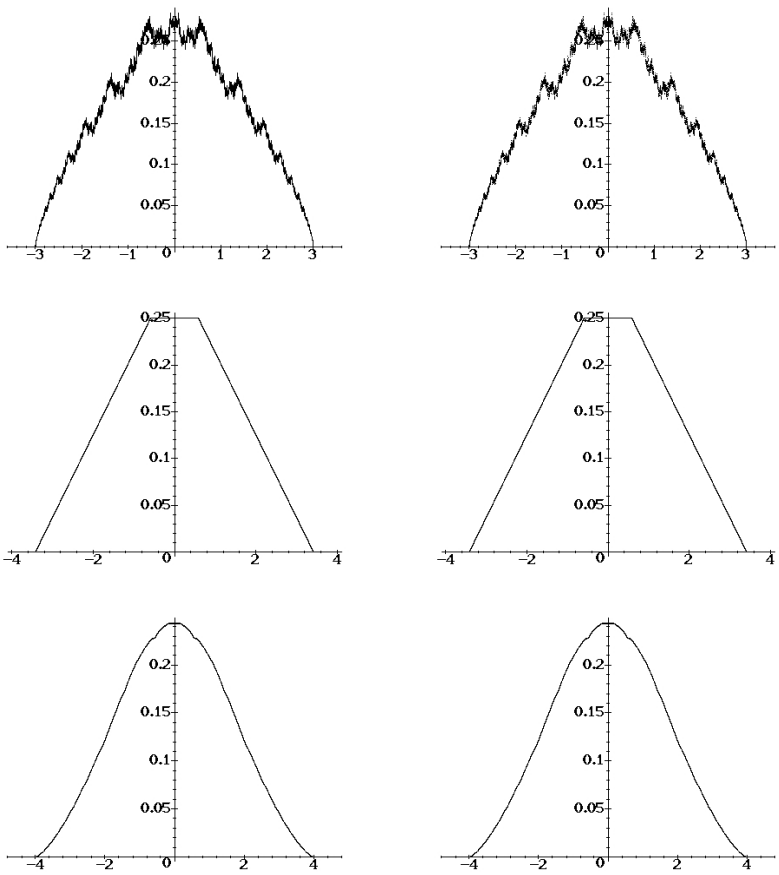


Figure 5.8. 26th iterate for $q = 2/3$ (top), $q = 1/\sqrt{2}$ (middle) and $q = 3/4$ (bottom).

Chapter 6

Random Vectors and Factoring Integers: A Case Study

Mathematics is often presented as a fait accompli, a finished result, with little attention being paid to the process of discovery or to the mistakes made, wrong turns taken, etc. However, especially where experimental mathematics is concerned, we can learn from a great deal from experience, both our own and that of others. New directions can be suggested by previous mistakes.

In this chapter we will consider as a case study the experiences of one of us supervising undergraduate research, paying particular attention to the role of experiment in introducing students to research at a much earlier stage than is traditional.

For several years, Clemson University has hosted a National Science Foundation supported Research Experiences for Undergraduates program in number theory and combinatorics, supervised jointly by Neil Calkin and Kevin James, with the assistance of two graduate students, Tim Flowers and Shannon Purvis. Each summer nine students are brought to Clemson University for eight weeks to learn what research in mathematics is about. We will focus on the research of two teams, Kim Bowman and Zach Cochran in 2004, and Katie Field, Kristina Little and Ryan Witko in 2005. At the time of their respective research experiences, all of the undergraduates had completed their junior year, with the exception of Katie Field, who had completed her sophomore year.

Probably the biggest difficulty in having (extremely bright) undergraduates do research is that they lack the breadth and depth of education we traditionally expect of students in graduate studies: this is especially true of students who are not from elite institutions such as Berkeley and Princeton. As a consequence we spend a lot of time in the first few weeks of the REU giving an intensive introduction to the mathematics required to study various problems. In order for the students to be able to get started

on research immediately, we try to choose topics having a significant experimental and computational component, and try to get the participants involved in computing in the first few days of the program.

The students listed above chose to work on projects related to the Quadratic Sieve algorithm for integer factorization.

Throughout this section, n will denote a (large) integer which we wish to factor, and $\mathcal{B} = \{p_1, p_2, \ldots p_k\}$ will denote a finite set of primes, typically the first k primes p for which n is a square (mod p) In practice, k is usually chosen so that $\log k = \sqrt{\log n}$: this choice arises from optimizing the running time of the sieve.

1 Integer Factorization

One of the oldest and most fundamental theorems in mathematics is the fundamental theorem of arithmetic.

> *Every positive integer has a unique factorization into a product of primes.*

An existence theorem, it led to one of the earliest problems in algorithmic mathematics (together with finding solutions to Diophantine equations): given that an integer n *can* be factored, how do we actually *find* integers a, b so that $n = ab$?

Currently the best general purpose factorization algorithms are the General Number Field Sieve (the GNFS) and the Quadratic Sieve. Both of these algorithms work by finding distinct square roots of a residue mod n. The questions we discuss here apply to both, but since QS is much simpler, we will restrict our attention to it. Our discussion of the Quadratic Sieve will follow closely that in Pomerance's "A Tale of Two Sieves"[196].

1.1 Fermat's Method

The idea of factoring an integer n by finding square roots mod n can be traced back to Fermat . He observed that if we can find integers a and b so that

$$n = a^2 - b^2$$

then we immediately have the factorization

$$n = (a - b)(a + b).$$

It is clear that provided that n is odd, the reverse is also true: a factorization $n = (2u+1)(2v+1)$ leads to $n = (u+v+1)^2 - (u-v)^2$, so n can also be written as the difference of two squares.

1.2 Kraitchik's Improvement

The first major advance in this direction was by Maurice Kraitchik , who observed that it wasn't necessary that n be a *difference* of two squares: it was enough that

$$a^2 \equiv b^2 \pmod{n}$$

and

$$a \not\equiv \pm b \pmod{n}.$$

Indeed, if $a+b$ and $a-b$ are non-zero $\pmod{n}$, then $n|(a+b)(a-b)$ implies $n|\gcd(a+b,n)\gcd(a-b,n)$, so both $\gcd(a+b,n)$ and $\gcd(a-b,n)$ must be non-trivial factors of n.

The question now is how to find distinct a, b so that $a^2 \equiv b^2 \pmod{n}$. Kraitchik's method was to consider the values taken by $f(x) = x^2 - n$ for values of x slightly larger than $\sqrt{n}$. Since these are all known to be congruent to squares mod n, if we can find a set of these values whose product is a square, then we might be able to use these to factor n.

It is easy to construct an example: to factor the integer $n = 2449$, we observe that $\lceil\sqrt{n}\rceil = 50$: we consider the following squares $\pmod{n}$:

$$\begin{aligned}
50^2 - n &= 3 \times 17 \\
51^2 - n &= 2^3 \times 19 \\
52^2 - n &= 3 \times 5 \times 17 \\
53^2 - n &= 2^3 \times 3^2 \times 5 \\
54^2 - n &= 467 \\
55^2 - n &= 2^6 \times 3^2 \\
56^2 - n &= 3 \times 229 \\
57^2 - n &= 2^5 \times 5^2.
\end{aligned}$$

Rewriting these equations as congruences modulo n, we obtain

$$\begin{aligned}
50^2 &\equiv 3 \times 17 && \pmod{n} \\
51^2 &\equiv 2^3 \times 19 && \pmod{n} \\
52^2 &\equiv 3 \times 5 \times 17 && \pmod{n} \\
53^2 &\equiv 2^3 \times 3^2 \times 5 && \pmod{n} \\
54^2 &\equiv 467 && \pmod{n} \\
55^2 &\equiv 2^6 \times 3^2 && \pmod{n} \\
56^2 &\equiv 3 \times 229 && \pmod{n} \\
57^2 &\equiv 2^5 \times 5^2 && \pmod{n}.
\end{aligned}$$

Hence we have

$$(50 \times 52 \times 53 \times 57)^2 \equiv (2^4 \times 3^2 \times 5^2 \times 17)^2 \pmod{n}$$

that is,

$$2424^2 \equiv 657^2 \pmod{2449}.$$

We compute $\gcd(2424 + 657, 2449) = 79$ and $\gcd(2424 - 657, 2449) = 31$, and we note that both 31 and 79 are non-trivial factors of n. Indeed, $n = 31 \times 79$. (We should also point out that we could have used $55^2 \equiv (2^3 \times 3)^2 \pmod{2449}$ via Fermat's method).

Exercise 1.1. *Use Kraitchik's idea to factorize 2041.*

Exercise 1.2. *Construct an example in which we don't get an easy factorization via Fermat's method.*

1.3 Brillhart and Morrison

Brillhart and Morrison [181] found an easy method to systematize the search for a subset whose product is a square: they observed that by considering the vectors of prime exponents in the prime factorizations of $f(x) - n$, a subset had a product which is a square if and only if the corresponding vectors summed to the zero vector (mod 2). In the example above, the fact that $(3 \times 17)(3 \times 5 \times 17)(2^3 \times 3^2 \times 5)(2^5 \times 5^2)$ is a square corresponds to the fact that the vectors of prime exponents sum to a vector all of whose entries are even.

$$\begin{pmatrix}0\\1\\0\\0\\0\\0\\1\end{pmatrix} + \begin{pmatrix}0\\1\\1\\0\\0\\0\\1\end{pmatrix} + \begin{pmatrix}3\\2\\1\\0\\0\\0\\0\end{pmatrix} + \begin{pmatrix}5\\0\\2\\0\\0\\0\\0\end{pmatrix} = \begin{pmatrix}8\\4\\4\\0\\0\\0\\2\end{pmatrix}$$

Hence searching for products which are squares is equivalent to finding linear dependencies among the prime exponent vectors considered over the binary field $\mathbb{F}_2$.

A useful and important technical observation here is that if a prime p divides $x^2 - n$, then since $x^2 \equiv n \pmod{p}$, n must necessarily be a quadratic residue mod p. Hence any prime for which n is *not* a quadratic residue will always have exponent zero, and might as well be ignored in all our computations. In the example above, this explains why the primes 7, 11 and 13 never appear as factors.

Brillhart and Morrison suggested setting a bound B, and restricting attention to those x for which $x^2 - n$ factored completely over the set $\mathcal{B}$ of primes (for which n is a quadratic residue) less than B. Call a value $x^2 - n$ which factors completely over $\mathcal{B}$ $\mathcal{B}$-smooth.

1.4 Pomerance and Sieving

A naive implementation of the methods as described above would proceed by attempting to factor $x^2 - n$ for each x, checking whether it factors over $\mathcal{B}$ and creating the exponent vector or discarding it as appropriate. Pomerance made the critical observation that this step could be improved immensely by sieving. Rather than factoring one $f(x) = x^2 - n$ at a time, he suggested considering the primes p in $\mathcal{B}$ one at a time, and using the observation that if $p|x^2 - n$ then $p|(x+p)^2 - n$ to extract factors of p from many $f(x)$ values at once. More precisely, find a value x_0 so that $p|f(x_0)$, and then sieve out $x_0, x_0 + p, x_0 + 2p, \ldots$.

The advantage here is that for each prime in $\mathcal{B}$, the work of determining whether $p|x^2 - n$ is shared over all values of x under consideration. The disadvantage is that the method inherently needs the range of x being considered to be fixed in advance. This leads to three questions:

1. How many x do we need to consider to obtain l $\mathcal{B}$-smooth numbers?

2. How large does l have to be in order for a set of l $\mathcal{B}$-smooth numbers to have a product which is a square?

If $|\mathcal{B}| = k$, then clearly if we have $k+1$ exponent vectors, each having k coordinates, then the set must be linearly dependent. However, it is possible that the first linear dependency could occur far earlier. Since the Quadratic Sieve is a sieve, in order to obtain the full efficiency of the algorithm the set of x for which we will attempt to factor $x^2 - n$ has to be determined in advance: adding just a few more x values after the fact is expensive.

2 Random Models

We can regard the Quadratic Sieve as taking a range $1, 2, \ldots, R$, and for each i in the range setting $x = \lfloor \sqrt{n} \rfloor + 1$ and returning either FAIL (if $x^2 - n$ doesn't factor over the factor base $\mathcal{B}$) or a vector $\underline{v}_i$, the entries of which are the parities of the exponents of the primes in the factor base in the factorization of $x^2 - n$. Although the vectors $\underline{v}_i$, and the indices for which we get FAIL are deterministic, they are hard to predict, and (in a strong sense) look random.

Although most questions in number theory are completely deterministic, it is often possible to obtain insight (and sometimes even theorems) by considering random models which behave similarly to the integers.

Erdős and Kac [103] introduced the idea of probabilistic methods in number theory in 1945. Building on work of Hardy and Ramanujan [134], they showed that the number of prime factors of a randomly chosen large

integer, suitably normalized, approximately follows a normal distribution. This led to a host of similar theorems.

As we noted above, the Quadratic Sieve can be viewed as producing a sequence of vectors whose entries are distributed in a quasi-random fashion. If we choose an appropriate probabilistic model, we may be able to predict the behaviour of the Quadratic Sieve based on the behaviour of the random model.

Of course, this would not prove that the sieve behaves in a particular fashion: however, it would suggest possible parameter choices which could be tested empirically.

We will discuss two different types of models, with various different possibilities for parameters. In each model, vectors will be chosen independently with replacement from the same distribution.

The first model will be vectors chosen uniformly from the set of vectors of weight w (and the generalization of this to vectors with a given distribution of weights).

In the second model, vectors will have entries which are chosen independently, but with different probabilities.

3 The Main Questions

The main questions now are the following:

- What is the right random model for vectors arising from sieves?
- Given a model for randomly generating vectors in $\mathbb{F}_2^k$, how many vectors do we need to select before we get a linear dependency?

3.1 The Constant Weight Model

Let $\omega(m)$ be the number of distinct prime factors of m: Hardy and Ramanujan [134] proved that if $g(m) \to \infty$, then the number l_n of integers m in $\{1, 2, 3, \ldots, n\}$ for which

$$|\omega(m) - \log\log m| > g(m)(\log\log m)^{1/2}$$

satisfies $l_n = o(n)$.

Less precisely, they showed that almost all numbers m have about $\log\log m$ distinct prime factors, and that the probability of differing from this by much more than $(\log\log m)^{1/2}$ tends to 0.

Erdős and Kac [103], in their seminal paper introducing probabilistic number theory, showed that if m is chosen uniformly from $\{1, 2, 3, \ldots, n\}$, then the distribution of

$$\frac{\omega(m) - \log\log m}{(\log\log m)^{1/2}}$$

converges to the normal distribution as n tends to infinity.

This suggests the following approach to developing a random model for the vectors arising from the sieve: the vectors v_i which are produced by the Quadratic Sieve correspond to numbers in the range $(1, 2, 3, \ldots n-1)$, so they ought to have about $\log\log n$ distinct prime factors. Furthermore, Canfield, Erdős and Pomerance [71] showed that given a bound B the proportion of B-smooth numbers less than n is about e^{-u}, where $u = \log n / \log B$. Hence for our first model for each i in the range we will generate FAIL with probability $1 - e^{-u}$, and with probability e^{-u} we will pick a vector of weight $w = \log\log n$ uniformly from $\mathbb{F}_2^k$. (Recall that in the implementations of the Quadratic Sieve, k is typically chosen so that $\log k \simeq \sqrt{\log n}$: hence $\log\log n \simeq 2\log\log k$.)

3.2 The Independent Entries Model

Fix a prime p: since a random integer chosen uniformly from a large range is divisible by p^e with probability about $1/p^e$, the probability that it is divisible by an odd power of p is

$$\frac{1}{p} - \frac{1}{p^2} + \frac{1}{p^3} - \frac{1}{p^4} + \cdots = \frac{1}{p+1}.$$

Hence we might choose the following model: take as the factor base a set $\mathcal{B} = \{p_1, p_2, \ldots, p_k\}$: each component of the vector $\underline{v}_i$ is chosen uniformly and independently, and the probability that the j component of $\underline{v}_i$ is 1 is

$$\Pr(\underline{v}_i[j] = 1) = \frac{1}{p_j + 1}$$

As we shall see later, we will want to modify this slightly so that

$$\Pr(\underline{v}_i[j] = 1) = \alpha_j$$

where α_j will be chosen to take into account some subtle complications.

4 Bounds

It is a simple fact from elementary linear algebra that if we have $l \geq k+1$ then any set of l vectors in a vector space of dimension k is linearly dependent. This fact has typically been used in implementations of the Quadratic Sieve to determine a range to be sieved over: choose an l somewhat larger than k, and use Canfield, Erdős and Pomerance's estimates on the proportion of smooth numbers to determine the size of range required to produce l $\mathcal{B}$-smooth numbers. If we could determine a bound L, significantly less than k, so that with high probability any L exponent vectors are linearly

dependent, then this would improve the running time of the sieve in the following ways:

1. The sieving stage of the algorithm would need to return fewer $\mathcal{B}$-smooth numbers: this would decrease the running time
2. The linear algebra stage of the algorithm would involve finding a linear dependency in a matrix with fewer rows, also decreasing the running time
3. It is possible that stopping the algorithm earlier will allow us to use pre-processing on the vectors which will significantly decrease the size of the linear algebra problem.
4. The change would allow us to optimize the parameters k and $\mathcal{B}$ based on the new running times: this might also improve the running time.

Of these, the sieving stage is likely to be least significant: this stage of the algorithm is easily parallelized, and indeed, the most recent examples of factorization records have been set by parallel implementations.

The matrix stage has more potential: this stage is now becoming the bottleneck for factoring.

The reoptimizing would be necessary to take full advantage of speedups especially in the linear algebra phase.

Upper bound for the constant weight model Consider the constant weight $w = w(k)$ model: recall that we pick l vectors uniformly from the set of vectors of weight w in $\mathbb{F}_2^k$. The probability that a given vector has a 0 in the jth coordinate is $(1-w/k)$. Since the vectors are chosen independently, the probability that all of the vectors have a zero in the jth coordinate is

$$\left(1 - \frac{w}{k}\right)^l.$$

Hence if we pick k vectors this way, the expected number of coordinates that are never 1 is about ke^{-w}. Hence our vectors are contained in a subspace of dimension less than $k(1 - e^{-w})$, and the first linear dependency will probably occur for some $l < k(1-e^{-w})$. Since we only get information about the expected number of coordinates we need to use a sharp concentration result to show that the probability that we are not close to the expected value is small: this is technical, but standard and straightforward.

Upper bound for the $\alpha_j = 1/(p_j + 1)$ model As is usually the case, similar methods work for the α_j model, but the analysis is rather more delicate and depends on the behaviour of the α_j's. We recall that in this model we obtain vectors $\underline{v}_i$, which form the rows of a matrix $A = (a_{ij})$, so that

$$\Pr(a_{ij} = 1) = \alpha_j$$

independently. Hence the probability that column j is empty (i.e. has no 1's) is $(1-\alpha_j)^l$, and in most of the cases we'll be considering, either this will be very small, or $(1-\alpha_j)^l \simeq e^{-l\alpha_j}$. Thus the expected number of non-empty columns $E(C_l)$ is given by

$$E(C_l) = \sum_{j=1}^{k} 1-(1-\alpha_j)^l \simeq \sum_{j=1}^{k} 1-e^{-l\alpha_j}$$

In the most simplistic model of this type, we take $\alpha_j = 1/(p_j+1)$: we make the reasonable assumption that our factor base contains about half of the primes up to B and that $p_j \sim 2j\log j$. Now it is easy to see that

$$E(C_l) \simeq \sum_{j=1}^{k} \left(1-e^{-l/(p_j+1)}\right) \tag{6.1}$$

If $E(C_l) < l$, then (after again applying an appropriate sharp concentration result) we can conclude that with high probability the vectors $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_l\}$ are linearly dependent.

The sum in equation 6.1 can be estimated by splitting it into appropriate ranges. A suitable choice of ranges is
(I) $1 \le j \le l/(\log l)^2$
(II) $l/(\log l)^2 < j \le l$
(III) $l < j \le k$.
Since the summands are all at most 1, and we are concerned with the size of the sum relative to l, the contribution of range (I) is negligible. We approximate the second range by an integral, replacing p_j+1 by $2j\log j$: it is then easy to see that the contribution of this range is $O\left(l\frac{\log\log l}{\log l}\right)$. Finally, in region (III), $l/(p_j+1)$ is small, so the summand is approximately $\frac{l}{p_j+1}$. By Merten's theorem (and assuming that about half of the primes are in the sum), the contribution of this region is about $l(\log\log k - \log\log l)/2$. Hence if $\log\log k - \log\log l < 2$ we have $E(C_l) < l$. Writing $l = k^\delta$, this gives $-\log\delta < 2$, that is, $\delta > e^{-2}$. Hence if we have many more than k^{1/e^2} vectors in this model we expect to see linear dependence.

At this point the students investigating the model got very excited, since this would lead to an immense speedup in the Quadratic Sieve: however after implementing the Quadratic Sieve, they discovered that the predictions were far too optimistic: this led to a re-examination of the assumptions of the model.

Lower Bounds Lower bounds for these models are both more difficult, and less suggestive of suitable choices of parameters for the sieve. However, we can obtain some results in both choices of model.

Calkin [70] and Balakin, Kolchin and Khokhlov [24, 153] showed independently that in the case of vectors of fixed constant weight w the following is true:

Theorem 4.1. *There is a constant $\beta = \beta(w)$ so that if $b > \beta$ if $l = l(k) < kb$ then as $k \to \infty$, the probability that a set of l vectors chosen uniformly from the vectors of weight w in $\mathbb{F}_2^k$ is linearly independent tends to 1.*

Essentially this theorem says that we need about βk vectors to obtain a linearly dependent set: careful analysis shows that

$$\beta(w) \simeq 1 - \frac{e^{-w}}{\log 2}$$

and that if $w = w(k) \to \infty$ as $k \to \infty$ (and $w < k/2$) then $\beta(w) \to 1$.

How would we discover and prove such a theorem? We want to know how likely it is that a set of vectors is dependent: can we determine the probability that a given set of vectors sums to the zero vector? This suggests trying Markov chain approach: start at the origin and add random vectors of weight w using a transition matrix $T = T_w$ to compute the probability that we end up back at the origin.

Since the transition matrix will have size $2^k \times 2^k$ it is natural to try to use the symmetry of the problem to reduce the complexity. Note that we can do this by grouping together points in $\mathbb{F}_2^k$ by weight: then the probability T_{ij} of moving from a vector of weight i to a vector of weight j by adding a random vector of weight w is precisely

$$T_{ij} = \frac{\binom{i}{\frac{w+i-j}{2}}\binom{k-i}{\frac{w-i+j}{2}}}{\binom{k}{w}}$$

(if we have i 1's in a vector, and we flip x of them to 0's and flip $w - x$ of the other 0's to 1's, then we end up with $j = i - x + (w - x)$ 1's as a result: hence $x = (w + i - j)/2$). Using this transition matrix, we can obtain an expression for the probability of a random walk on the hypercube returning to the origin after a given number of steps.

Experimentation in Maple now reveals the beautiful fact that the transition matrix T is diagonalizable, and that its eigenvalues and eigenvectors have rational entries, which can be expressed in terms of sums of binomial coefficients. [1] Hence we can write down an expression for $E(2^s)$ which is amenable to analysis, leading to the theorem above.

[1] It also turns out the eigenvectors for T_w are independent of w: hence for two distinct weights w and w', the matrices T_w and $T_{w'}$ commute: this corresponds to the observation that if we start with a vector of weight w and add a vector of weight w' the effect is the same as starting with a vector of weight w' and adding a vector of weight w. This is intimately related to the fact that there is an *association scheme*, the *Hamming scheme* here.

A key observation now is that this is exactly the probability that a set of vectors sums to zero: if we sum this over all subsets of an l-set of vectors, we obtain the expected number of linear combinations which sum to zero, that is, the expectation $E(2^s)$ of the size of the left null space of the matrix A whose rows are the vectors $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_l$ (that is, s is the co-rank of A). If $E(2^s)$ is close to 1, then the set $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_l\}$ is likely to be linearly independent. This method works by considering the left null-space of A. One can perform a similar calculation, giving exactly the same results, using the right null-space of A.

Since we now have upper and lower bounds for the constant weight vector models, it makes sense to ask whether there is a sharp threshold for linear dependence: that is, does there exist a threshold value b_t so that
if $b < b_t$, then $\lim_{n\to\infty} Pr(\text{dependent}) = 0$
if $b > b_t$, then $\lim_{n\to\infty} Pr(\text{dependent}) = 1$.

This in fact was where Bowman and Cochran started their investigations, generating random binary vectors with exactly three 1's and computing when the set of vectors generated first becomes dependent.

Exercise 4.1. *Fix the dimension k of the vector space, and find a way of uniformly generating binary vectors of weight 3. For various values of k generate random vectors until you obtain a linearly dependent set. Compare the number of vectors required to the upper and lower bounds given above. Do you think that there is a sharp threshold?*

Exercise 4.2. *Investigate the eigenstructure of other structured matrices: for example, consider the Toeplitz matrices*

$$T_n = \begin{pmatrix} 1 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 1 & 0 & \ddots & \\ 0 & 1 & 1 & 1 & \ddots & \\ 0 & 0 & 1 & 1 & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ 0 & & & & & 1 \end{pmatrix}$$

Compute the eigenvalues for various values of n, plotting the values in increasing order. Guess, with the help of the inverse symbolic calculator if necessary, a formula for the jth largest eigenvalue of T_n. Compute the corresponding eigenvectors, and guess the formula for their entries. For more information on this and related problems, see Doust, Hirshhorn and Ho [99]

Similar methods, both for the left and the right null space can be applied to the α_j model: in this case the bounds obtained are harder to analyze, and of course depend on the behaviour of the parameters α_j.

5 Which Model Is Best?

The constant weight model is easily seen to be inappropriate: the model assumes that the probability that a large prime divides $x^2 - n$ is the same as the probability that a small prime does. Hence the depressing fact that the lower and upper bounds for the number of vectors required to obtain an independent set with high probability are both asymptotic to k, the dimension of the space, is irrelevant to the Quadratic Sieve.

The second model, where

$$\Pr(\underline{v}_i[j] = 1) = \frac{1}{p_j + 1}$$

requires more thought. In this model, the upper bounds on probable linear dependencies occur *far* earlier than k. Indeed, upon proving these bounds the students immediately set to coding up the Quadratic Sieve to see if the behaviour occurred in reality. Unfortunately, the tests that they ran suggested that the model was far too optimistic. Indeed, although the model seemed as if it ought to be much better than the constant weight model, it gave results which were much further off!

This led to much head scratching, trying to figure out why a seemingly reasonable model could give results that were so far out of touch with reality.

The key point here is the following: first, the reason that the model is incorrect is that it fails to take into account the fact that we are conditioning on numbers being $\mathcal{B}$-smooth. We will give a simple example in a moment to show how drastically this can change probabilities. The reason that even seemingly small changes can have a large impact on the sieve is that we will end up estimating sums which behave like

$$\sum_{i<x} \frac{1}{i} \simeq \log x$$

instead of

$$\sum_{i<x} \frac{1}{i \log i} \simeq \log \log x.$$

This will mean that a term which can be controlled in the inappropriate model will never be small in the corrected model.

Why can smoothness change things drastically? Consider the following question: what is the probability that a random integer is divisible by 2? If the integer is chosen uniformly from a large interval, then the answer is very close to $1/2$. However, if we choose uniformly from $\mathcal{B}$-smooth numbers less than x, then we see the following.

1. If $\mathcal{B} = \{2\}$, then there are $\lfloor \log_2 x \rfloor + 1$ $\mathcal{B}$-smooth numbers up to x, and all but 1 is divisible by 2. Hence the probability that a random $\mathcal{B}$-smooth number is even is

$$1 - \frac{1}{\lfloor \log_2 x \rfloor + 1}$$

2. If $\mathcal{B} = \{2, 3\}$, then the number of $\mathcal{B}$-smooth numbers up to x is equal to the number of non-negative integer solutions to the inequality

$$i \log 2 + j \log 3 \leq \log x$$

For large x, this in turn is about equal to the area of the corresponding triangle, namely

$$\frac{1}{2} \frac{\log x}{\log 2} \frac{\log x}{\log 3} = \frac{1}{2} \log_2 x \log_3 x$$

Now the number of these which are not divisible by 2 is exactly the number powers of 3 less than x, namely $\lfloor \log_3 x \rfloor + 1$: hence the probability that a random $\mathcal{B}$-smooth number less than x is even is about

$$1 - \frac{\lfloor \log_3 x \rfloor + 1}{\frac{1}{2} \log_2 x \log_3 x} \simeq 1 - \frac{2}{\log_2 x}.$$

5.1 Refining the Model

We wish now to refine the model to compensate for conditioning on $\mathcal{B}$-smoothness. A first attempt at this will proceed by modifying α_j, the probability of divisibility by the jth prime p_j in $\mathcal{B}$. Naively (by which we mean that this is what we actually did and thought correct for several months!) we let B be the largest element of $\mathcal{B}$, fix a prime $p < B$, and consider the number of B-smooth numbers up to n, and the number of B-smooth numbers divisible by p: the second quantity is exactly equal to the number of B-smooth numbers less than n/p. Using Canfield, Erdős and Pomerance's estimates for the number of B-smooth numbers and estimating some asymptotics, this leads to the following approximation:

$$\Pr(p \text{ divides } m \mid m \text{ is } B\text{-smooth and } \leq n) \simeq \frac{1}{p^{1-\delta}}$$

in which $\delta = \frac{\log u}{\log B}$ and $u = \frac{\log n}{\log B}$. Using the parameters usually chosen for the Quadratic Sieve, this gives $\log B \simeq \log k \simeq \sqrt{\log n}$, so $u \simeq \sqrt{\log n}$ and

$$\delta \simeq \frac{\frac{1}{2} \log \log n}{\log B} \simeq \frac{\log \log k}{\log k}.$$

When p is small, p^δ is very close to 1. However, when, say p is not $o(k)$, which is the case for most primes in $\mathcal{B}$, then

$$p^\delta \simeq \exp\left(\frac{\log p \log\log k}{\log k}\right) \simeq \log k \simeq \log p.$$

Is this model now appropriate? It is rather hard to compare empirical data from the Quadratic Sieve at this stage, because our estimates have been a little careless with constants: however, if our model was inappropriate, and in fact some aspect which we have ignored should be considered, and would introduce a factor, say, of $\log\log\log n$, would we be able to tell? If we run the Quadratic Sieve on toy problems, say with n about 10^{15}, then $\log\log\log n$ is much less than 2: even $\log\log n \simeq 3.5$ in this range!

However, having realized that there is a problem with conditioning on smoothness, let us now try to take it into account at every stage.

Let $\mathcal{P}$ denote the set of all primes for which n is a quadratic residue (mod p). It is known that this set has relative density about $1/2$ in the set of all primes: however, we will mainly be concerned with the small primes in the set (those less than n, say) and much less is known about the number of small primes in the set. In the examples we have studied, $\mathcal{P}$ appears to contain about half the primes in reasonably sized intervals.

Now $\mathcal{B}$ will be the set of primes $\{p_1, p_2, \ldots, p_k\}$ less than B which are contained in the set $\mathcal{P}$. We will write $\mathbb{Z}_\mathcal{P}$ for the set of $\mathcal{P}$-smooth numbers, and $\mathbb{Z}_\mathcal{B}$ for the set of $\mathcal{B}$-smooth numbers.

Now, the probability that the Quadratic Sieve returns a vector rather than FAIL should be about the probability that an integer m chosen uniformly in $\{1, 2, \ldots, n\}$ is $\mathcal{B}$-smooth given that it is $\mathcal{P}$-smooth. This leads to the following natural questions: given a set $\mathcal{P}$ of primes (of relative density $1/2$ in the set of all primes) and a finite subset $\mathcal{B}$ of $\mathcal{P}$

1. What is the number $Z_\mathcal{P}(y)$ of $\mathcal{P}$-smooth integers up to y?
2. What is the number $Z_\mathcal{B}(y)$ of $\mathcal{B}$-smooth integers up to y?

Then the probability that $x^2 - n$ returned by the sieve is $\mathcal{B}$ smooth should be about $Z_\mathcal{B}(n)/Z_\mathcal{P}(n)$. Furthermore, the probability that $x^2 - n$ is divisible by p, given that it is $\mathcal{B}$-smooth ought to be about $Z_\mathcal{B}(pn)/Z_\mathcal{B}(n)$.

Landau [159] and Wirsing [231] have addressed various versions of these questions: in particular, if y is allowed to be arbitrarily large, then Wirsing gave precise asymptotics for the behaviour of the function $Z_\mathcal{P}(y)$. When $|\mathcal{B}|$ is sufficiently small, $Z_\mathcal{B}(y)$ is also easy to estimate. However, in the regions relevant to the sieve, results are harder to obtain. We are interested in the behaviour of $\mathbb{Z}_\mathcal{P}(y)$ when y is somewhat smaller than n, and $\mathcal{P}$ is the set of primes for which n is a quadratic residue (mod p). Since $\mathcal{P}$ is defined

in terms of n, it seems unlikely that asymptotic results will be possible for values this small. However, there is some hope that the behaviour is similar for small values, and since we are primarily interested in heuristics, we will make this assumption.

Theorem 5.1 (Wirsing)**.** *Let $\mathcal{P}$ be a set of primes for which there exist constants δ and K so that*

$$\sum_{p<y, p\in\mathcal{P}} 1 \sim \delta \frac{y}{\log y}$$

and

$$\prod_{p<y, p\ in\mathcal{P}} \left(1 - \frac{1}{p-1}\right) \sim K(\log y)^{\delta}.$$

Then

$$\mathbb{Z}_{\mathcal{P}}(y) \sim \frac{\delta K e^{-\gamma\delta}}{\Gamma(\delta+1)} \frac{y}{(\log y)^{1-\delta}}$$

(in which γ is the Euler-Mascheroni constant, and Γ is the Gamma function).

Exercise 5.1. *When $\mathcal{P}$ is the set of primes congruent to 1 (mod 4), what are the values of the constants δ and K?*

Exercise 5.2. *Pick a moderately large value n (the product of two large primes would be appropriate). Using Maple or Mathematica, estimate the constants δ and K appearing in Wirsing's theorem.*

Exercise 5.3. *For those who know about quadratic reciprocity: suppose that n is the product $n = q_1 q_2$ of two primes. Find a condition* on *p for a prime p to be such that n is a quadratic residue* $\pmod p$*. Assuming Dirichlet's theorem (that if a and m are relatively prime then the number of primes congruent to a* $\pmod m$ *is asymptotic to $1/\phi(m)$) deduce that the relative density of $\mathcal{P}$ in the set of all primes should be $1/2$. Compare the value of δ from the previous exercise to 1/2.*

Exercise 5.4. *For small y, compare the value of $\mathbb{Z}_{\mathcal{P}}(y)$ to the value predicted (for* large *y) by Wirsing's theorem. Does our assumption that $\mathbb{Z}_{\mathcal{P}}(y)$ behaves similarly for small and for large y seem justified?*

Exercise 5.5. *Devise some statistical tests to determine whether the heuristic that*

$$\alpha_j = Pr(p_j | x^2 - n) \simeq c \frac{\log(p_j)}{p_j}$$

is a reasonable one. For example, the number $x^2 - n$ lies between $\sqrt{n}$ and n and so the sums

$$\frac{1}{2}\log n < \sum_{p_j | x^2 - n} \log p_j < \log n$$

and

$$\sum_{j=1}^{k} \alpha_j \log p_j$$

should be about the same. (This ignores primes dividing $x^2 - n$ to the second or higher power, but their effect is small). Note that this heuristic illustrates again the fact that the size of $\mathcal{B}$ has an impact on the probability that a $\mathcal{B}$-smooth number is divisible by p.

Following the heuristics suggested by the previous section, it seems reasonable to assume that the probability α_j that $p_j | x^2 - n$, given that $x^2 - n$ is $\mathcal{B}$-smooth should be about $c \log p_j / p_j$ for some constant c.

Since we also have the heuristic that about half the primes are in our factor base, by the prime number theorem we expect that $p_j \simeq 2j \log j$, and so $\alpha_j \simeq c/j$.

We now let A be an $l \times k$ binary array, entries chosen independently, in which probability that an entry in column j is 1 is $\alpha_j = c/j$. We now consider the number of empty columns, the number of columns with exactly one 1 (which we will refer to as a *solon* for solo 1) and the number of columns containing exactly two 1's (which we will refer to as *colons*). Let $X_r(A)$ be the number of columns containing exactly r 1's. Earlier we considered the number of non-empty columns, $k - X_0$. We now refine our analysis to include solons: we will take colons into account later.

If a column is a solon, then the row containing the corresponding 1 cannot appear in a non-trivial row-dependency. Thus we can replace A by the array with the solon and its corresponding row removed. Since two solons can have the same row containing their solo 1, we can remove more columns than rows. There are $X_1(A)$ solons, and for each solon every row is equally likely to contain the solo 1. Hence the probability that a given row is not deleted is

$$\left(1 - \frac{1}{l}\right)^{X_1} \simeq e^{-X_1/l}$$

and so the expected number of rows left after deleting X_1 solons is about

$$l e^{-X_1/l}$$

Now the rows of array A are linearly dependent if

$$e^{-X_1/l} l > k - X_0 - X_1.$$

Computing expected values here,

$$E(X_0) \simeq \sum_{j=1}^{k} e^{-l\alpha_j} \simeq \int_1^k e^{-cl/x}\ \mathrm{d}x$$

$$\simeq \int_0^k e^{-cl/x}\ \mathrm{d}x = ke^{-cl/k} - clEi(1, cl/k)$$

in which $Ei(a, b) = \int_1^\infty e^{-bx}x^{-a}\ \mathrm{d}x$ is the exponential integral (as defined by Maple). Furthermore,

$$E(X_1) = \simeq \sum_{j=1}^{k} l\alpha_j e^{-l\alpha_j} \simeq \int_1^k \frac{cl}{x} e^{-cl/x}\ \mathrm{d}x$$

$$\simeq \int_0^k \frac{cl}{x} e^{-cl/x}\ \mathrm{d}x = clEi(1, cl/k)$$

So we expect that the number of rows and columns remaining will be equal after removing empty columns and solons if

$$l \exp(-E(X_1)/l) = k - E(X_0) - E(X_1)$$

is satisfied: that is,

$$l \exp(-cEi(1, cl/k)) = k(1 - \exp(-cl/k))$$

Solving this equation numerically with Maple, we obtain the following values:

c	l/k threshold
0.2	0.0008980931
0.4	0.1400018494
0.6	0.4119714260
0.8	0.6120819888
1.0	0.7394711744
1.2	0.8207464220
1.4	0.8739519006
1.6	0.9097291738
1.8	0.9343624832
2.0	0.9516659561

For the value $c = 0.72$ (which arises in the experimental evidence discussed below) we obtain a threshold of .5423300259.

Reducing the size of the final linear algebra problem We include colons in our analysis: their impact is not to reduce the number of rows

needed to obtain the first dependency (there is a small impact, but it is negligible, so we won't discuss it), but rather to reduce the size of the final linear algebra question under consideration.

Suppose that a column is a colon, that is, it contains exactly two 1's, and that these 1's appear in rows r_1 and r_2. Then in any linear combination of rows summing to zero, either both the rows r_1 and r_2 appear, or neither appear. Hence the rows of A are linearly dependent if and only if the rows of A' are, where A' is the array obtained by replacing r_1 by $r_1 + r_2$, replacing r_2 by $\underline{0}$, and deleting the colon-column.

A good combinatorial model for this is to take a graph G having as vertex set the set of rows of A, and as edge set the set of pairs $\{r_1, r_2\}$ corresponding to colons. Then replacing rows by their sum corresponds to contracting edges in the graph, replacing connected components by a single vertex. Components containing a vertex corresponding to a 1 appearing in a solon get deleted completely.

This enables us to reduce the size of the array A considerably: if we have no cycles in the graph, then the contraction of each edge reduces the number of components by 1. Furthermore, when we contract edges or delete components in the graph, we can create new edges (for example, if a column in A contains four 1's, in rows r_1, r_2, r_3 and r_4, and if we have a colon with 1's in rows r_1 and r_2, then when we replace r_1 with $r_1 + r_2$ and r_2 with $\underline{0}$, the four 1's will become a colon).

Exercise 5.6. *Suppose that a column of A contains exactly three 1's and that A has X_1 solons. After deleting the solons, what is the probability that the column now contains two 1's?*

Exercise 5.7. *Suppose A has X_1 solons, X_2 colons, and X_3 columns with exactly three 1's. If we remove the colons and then the solons, the columns with exactly three 1's can remain unchanged, can become colons, solons or empty. What is the expected number of each that will be produced.*

Exercise 5.8. *Suppose A has X_1 solons, X_2 colons, and X_r columns with exactly r 1's. If we remove the colons and then the solons, the columns with exactly r 1's can become columns with s 1's, $0 \le s \le r$. What is the expected number of each that will be produced?*

Exercise 5.9. *Consider a stochastic model for the deletion of colons and solons along the following lines: consider the random variable $X(t) = (X_0(t), X_1(t), X_2(t), dots)$, where $X_i(t)$ is the number of columns containing exactly i 1's after t rounds of deletions of solons and colons. More precisely, $X(t-1)$ is the number before the tth round of deletions: in round t, remove all colons (by replacing colon-rows with their sum), and then remove all solons. Now update $X(t)$. Develop heuristics (by making reasonable assumptions where necessary) for the dynamics of $X(t)$.*

6 Experimental Evidence

The students involved in this project ran several different experiments, and we have followed up on these. To begin with, Tim Flowers implemented the quadratic sieve to produce real data: working with n around 10^{15} and a factor base bound $B = \exp(\sqrt{\log(n)\log\log(n)})$, iterated deletions reduced arrays of size around 500×2000 to arrays of size around 65×60.

In addition, the arrays produced suggest that (for n in this range) the probabilities α_j satisfy $\alpha_j \simeq 0.72/j$.

Exercise 6.1. *Explain the seemingly paradoxical fact that increasing the number of columns in A* decreases *the size of the final non-zero array after iterated deletions. For example, the values in the table are typical. Each row is the result of iterated deletions of solons and colons for a random array A with $\alpha_j = 0.72/j$. The final size is the number of non-zero rows and columns.*

Initial size	*Final size*
10000×10000	4700×2300
10000×15000	2250×1300
10000×20000	800×480
10000×25000	300×180
10000×30000	150×100
10000×35000	100×50
10000×40000	47×22
10000×45000	38×20

We fixed $c = 0.72$ and constructed random arrays of size $l \times k$ for various values of l, k, and computed various statistics:

- the proportion of successful trials: trials for which iterated removal of solons and colons left an array with more non-zero rows than columns (implying the rows of the initial array are linearly dependent)
- the proportion of trials for which iterated removal of solons and colons left a non-zero array with at most as many rows as columns (so that we can't conclude that the rows of the initial array are linearly dependent)
- the maximum and minimum number of non-zero rows and columns remaining after iterated removal of solons and colons
- the average number of rows and columns remaining after iterated removal of solons and colons.

We observed in almost all trials we either ended up with a zero matrix (so the original rows were independent) or with more rows than columns (so the original rows were dependent). This is consistent with the following conjecture:

Conjecture 6.1. *Generate vectors in GF_2^k independently with $\alpha_j = 0.72/j$. Then almost surely (in the sense that the probability approaches 1 as $k \to \infty$) the first linear dependency gives a matrix which reduces to a $t + 1 \times t$ matrix after iterated removal of solons and colons.*

Essentially, this conjecture says that there are two conditions, one of which implies the other, and that usually, the two coincide exactly. This is a common situation in the theory of random graphs, in which for example, if the edges of the complete graph on n vertices are placed in random order, the edge which makes the graph connected is almost surely the edge which makes the minimum degree equal to 1.

In studying threshold functions for the occurrence of combinatorial structures, for example in random graphs (see for example, Bollobas [34]) it is frequently the case that the threshold behaves (suitably scaled) like the function $e^{e^{-x}}$ Figure 6.1 shows the proportion of successful trials with $k = 20000$, for l from 1000, 1100, 1200, ... 3000, with 100 trials each. (Figure 6.2) shows the function $e^{e^{-x}}$, suitably scaled and shifted to match this: the overlay of the two pictures in Figure 6.3 shows that the match is close but not perfect. However, it is close enough to suggest that there is a similar threshold behaviour going on.

We can also compare the threshold functions for finding dependence this way for various values of k: figure 6.4 shows the proportion of successes for $k = 20000, 30000$ and 40000: the horizontal axis is scaled to give l/k. This figure is consistent with the hypothesis that there is a sharp threshold β around 0.2.

7 Conclusions

After considering various probabilistic possible models for the Quadratic Sieve, it appears that we have found one which is a reasonably good match. Since almost all of our arguments were heuristic and our evidence almost purely experimental, there are two obvious routes to follow now. First, implement the Quadratic Sieve to take account of the ideas discussed here, and estimate the improved running time. Second, put the heuristics on a sounder footing, proving some of the natural assumptions that we have made. We are currently working on both of these tasks.

One of the advantages of removing solons and colons is that it can be done in parallel: most linear algebra algorithms are rather hard to paral-

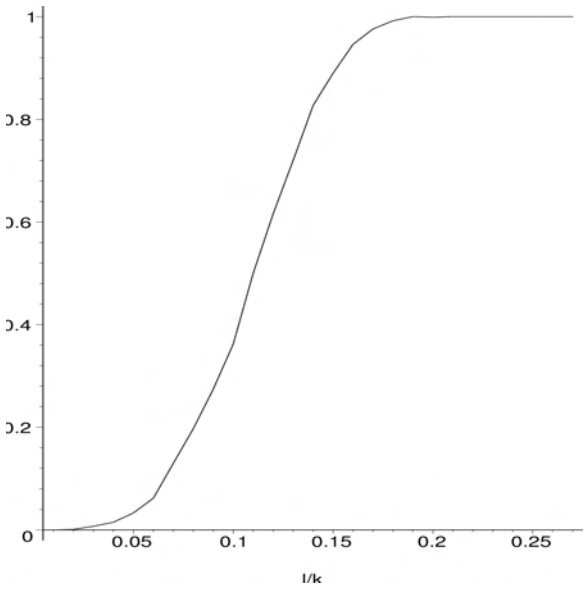


Figure 6.1. Probability l rows of length 20000 are dependent.

lelize, so if we can preprocess the matrix in this fashion, then it may make the remaining linear algebra tasks much faster to perform. (In addition, if the iterated deletions can be parallelized effectively, it is *possible* that this might make the Quadratic Sieve faster in implementations than algorithms such as the Number Field Sieve in which the linear algebra phase is hard to parallelize).

One might ask why these techniques have not been used already: it appears to be the case that they only become useful if l is sufficiently *small* compared to k, and since most implementations aimed to get l *large* in order to ensure dependency, the usefulness of iterated deletion of solons and colons has not been noticed.

While it is clear that the ideas here will speed up the Quadratic Sieve, we don't expect that this will have a huge impact: currently the Quadratic Sieve is only the second fastest general purpose factoring algorithm, and the improvements suggested here are unlikely to lift it into first place. The fastest algorithm, the General Number Field Sieve, is similar in nature, in that it generates a sequence of vectors over GF_2, and a linear dependency is used to generate a factorization. However, heuristics suggest that the vectors generated are somewhat denser than for the Quadratic Sieve: perhaps having $\alpha_j = c \log j/j$. If this is the case, then the expected number of solons and colons is so small that the techniques discussed here will not

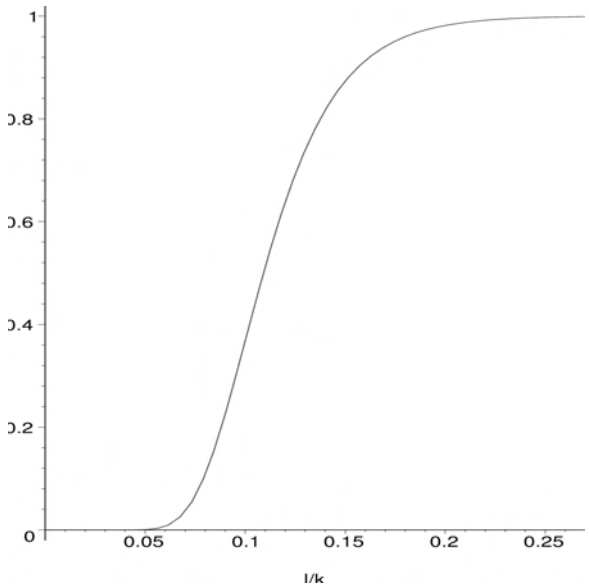


Figure 6.2. Scaled, shifted plot of $e^{-e^{-x}}$.

have any effect.

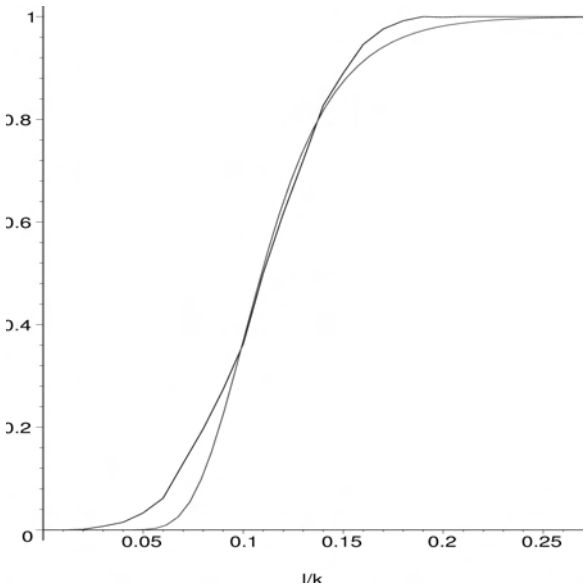


Figure 6.3. Overlay of probability, $e^{e^{-x}}$.

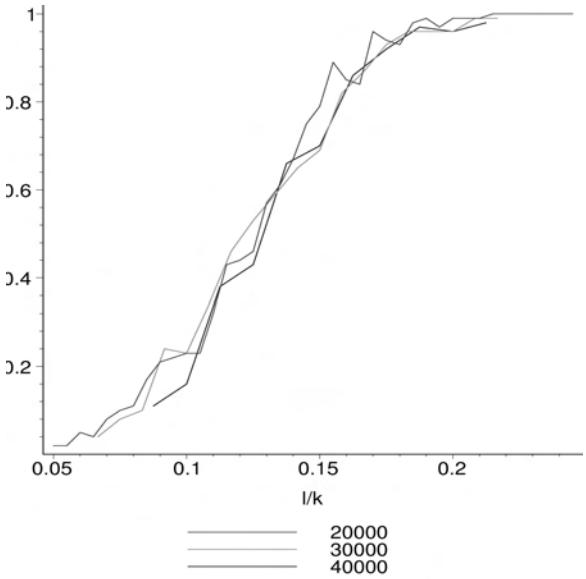


Figure 6.4. Probabilities of dependencies for various l, $k = 20000, 30000, 40000$.

Chapter 7

A Selection of Integrals from a Popular Table

1 Introduction

The evaluation of definite integrals is one of the most intriguing topics of elementary mathematics. Every student of Calculus is exposed to a variety of techniques that *sometimes* work, but they always leave the feeling of just being a collection of tricks.

The goal of this chapter is to introduce the reader to the methods of *Experimental Mathematics* in the context of definite integrals. We hope to convey that, in spite of wonderful collections such as *Tables of Integrals, Series and Products* by I.S. Gradshteyn and I. M. Rhyzik [124] and sophisticated symbolic languages such as Mathematica, there is a lot of interesting things to do. *The use of Mathematica is essential in our approach to this question.*

One of the features of definite integration is that similar integrands produce problems of different levels of complexity. The reader is already aware of this phenomenon: the integral

$$\int_0^\infty e^{-x}\,dx = 1, \tag{7.1}$$

is elementary, the normal integral

$$\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}, \tag{7.2}$$

can be obtained by elementary methods, but the next examples in this family

$$\int_0^\infty e^{-x^3}\,dx = \Gamma\left(\frac{4}{3}\right) \tag{7.3}$$

requires the introduction of the *gamma function*

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \tag{7.4}$$

to obtain the answer. This is a remarkable function that satisfies

$$\Gamma(n+1) = n! \tag{7.5}$$

so it provides an analytic extension of factorials. Under some mild conditions this is in fact unique. The evaluation (7.2) corresponds to the special value

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \tag{7.6}$$

At this point it is natural to ask whether $\Gamma(4/3)$ in (7.3) can be expressed in terms of more elementary functions. This is a difficult question that will not be addressed here, but see [75] for an introduction.

The fact that definite integrals are given as specific values of special functions is familiar to students. This is central to the question of what constitutes an *acceptable answer* to a required evaluation. For instance, every student will evaluate

$$\int_0^1 \frac{dx}{\sqrt{2-x^2}} = \frac{\pi}{4} \tag{7.7}$$

and consider it a reasonable good answer, but the similar problem

$$\int_0^1 \frac{dx}{\sqrt{9-x^2}} = \sin^{-1}\left(\frac{1}{3}\right) \tag{7.8}$$

will force the issue of *simplifying the answer*. Even at this level, this is related to deep interesting questions: for which integers n is the number $\sin(\pi/n)$ expressable in radicals? It is surprising to the beginner that the answer lies in the realm of Abstract Algebra. See [219] for a historical discussion of this topic.

A remarkable feature of integration is that *complicated real numbers* appear as definite integrals, in which the integrand is relatively simple. One does not need to complicate the integrand to produce a difficult problem. For example, in entry 4.241.11 of [124] we find

$$\int_0^1 \frac{\ln x \, dx}{\sqrt{x(1-x^2)}} = -\frac{\sqrt{2\pi}}{8}\left[\Gamma\left(\frac{1}{4}\right)\right]^2. \tag{7.9}$$

The gamma function defined in (7.4) makes a new appearance. The concept of *complicated* needs to be formalized: but it should be clear that (7.9)

is more complicated than (7.8).

The reader will find in [39] many other interesting features involved in the evaluation of definite integrals.

2 The Project and its Experimental Nature

The central question of our research can be described in simple terms: given a function

$$f = f(x; p_1, p_2, \cdots, p_n) \tag{7.10}$$

that depends on a set of *parameters*: $p_1, \cdots, p_n$, we want to express the definite integral

$$I = I(f; p_1, \cdots, p_n; a, b) = \int_a^b f(x; p_1, p_2, \cdots, p_n)\, dx, \tag{7.11}$$

in terms of the parameter set $\{p_1, \cdots, p_n, a, b\}$.

As such the problem is too general and simple to solve. Define g to be a *primitive* of f and use the fundamental theorem of calculus to evaluate I. The problem becomes more interesting if we fix the class of possible primitives. For instance: given a *rational function* f, find the value of I in the rational class.

Our long term plan is to develop a *complete theory of definite integration*. As a special part of this project, we would like to provide proofs of the many formulas appearing in the classical table of integralsi, such as [124]. The material presented there is enormous. This collection has as ancestors the table compiled by Bierens de Haan [32] and the beautifully typed tables [126] and [127]. There are some new tables that present interesting evaluations, for instance A. Apelblat [8]

In these notes we present some of the methods that we use in our work. Two aspects that appear througout this presentation are:

(1) **Integral representations**. Many special functions admit representations as integrals. For example,

$$\tan^{-1} x = \int_0^x \frac{dt}{1+t^2}. \tag{7.12}$$

We read these representations from right to left, that is, we see the special function as giving us evaluations of the definite integral.

(2) **Interesting numbers**. The values of integrals produce real numbers that *appear in other contexts*. For example,

$$\int_0^\infty \frac{t^{x-1}\,dt}{e^t+1} = (1-2^{1-x})\Gamma(x)\zeta(x) \tag{7.13}$$

where

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x} \tag{7.14}$$

is the famous **Riemann zeta function**. This function appeared in the study of distribution of prime numbers and its zeros are worth some money[1] See

`http://www.claymath.org/millennium/Riemann_Hypothesis`

specially the article by Peter Sarnak on that website. Therefore its special values have some intrinsic interest. We think of *special values* of *interesting functions* as the *primitive blocks* that form real numbers. For instance,

$$\int_0^{\pi/2} \Big(\frac{x}{\sin x}\Big)^4\,dx = -\frac{1}{12}\pi^3 + 2\pi\ln 2 + \frac{1}{3}\pi^3\ln 2 - \frac{3}{2}\pi\zeta(3) \tag{7.15}$$

is an expression formed by the blocks $\ln 2, \pi$ and $\zeta(3)$. The problem of simplication of these combinations is rather difficult: if the expression $\zeta(4)/\pi^4$ appears, then it has been known since Euler that it reduces to $1/90$. On the other hand, we do not know what to do about with $\zeta(3)/\pi^3$. The case of odd values of ζ seems to be much more difficult. W. Zudilin informs me that *A general feeling is that* π, $\zeta(3)$, $\zeta(5)$, $\zeta(7), \cdots$ *are algebraically independent, but this seems to be problem forever.*

3 Families and Individuals

Among the formulas presented in [124] some of them can be treated as membres of a larger family. For example, in 4.224.6 we find

$$\int_0^{\pi/2} \ln\cos x\,dx = -\frac{\pi}{2}\ln 2 \tag{7.16}$$

that is related to 4.224.8:

$$\int_0^{\pi/2} (\ln\cos x)^2\,dx = \frac{\pi}{2}\left[(\ln 2)^2 + \frac{\pi^2}{12}\right]. \tag{7.17}$$

[1] We believe that are much easier ways to make $1,000,000.

This leads to the consideration of the general case

$$LC_n := \int_0^{\pi/2} (\ln \cos x)^n \, dx. \tag{7.18}$$

A direct symbolic computation yields

$$\begin{aligned} LC_1 &= -\tfrac{1}{2}ac \\ LC_2 &= \tfrac{1}{2}a^2c + \tfrac{1}{24}c^3 \\ LC_3 &= -\left(\tfrac{1}{2}a^3c + \tfrac{3}{4}b^3c + \tfrac{1}{8}ac^3\right) \\ LC_4 &= \tfrac{1}{2}a^4c + 3ab^3c + \tfrac{1}{4}a^2c^3 + + \tfrac{19}{480}c^5, \end{aligned} \tag{7.19}$$

where we have employed the notation $a := \ln 2$, $b := \zeta(3)^{1/3}$, $c := \pi$. The values of $\zeta(2n)$ have been well-known since the eighteenth century. The Bernoulli numbers defined by the series

$$\frac{x}{e^x - 1} = 1 - \frac{1}{2}x + \sum_{n=1}^{\infty} B_{2n} \frac{x^{2n}}{(2n)!}, \tag{7.20}$$

are rational numbers with alternating signs. It turns out that

$$\zeta(2n) = (-1)^{n-1} \frac{2^{2n-1}}{(2n)!} B_{2n} \, \pi^{2n} \tag{7.21}$$

so that $\zeta(2n)$ is a rational multiple of π^{2n}. The situation for the odd values of ζ is much more difficult. The reader can find in [224] an informal report on the reaction to the proof of irrationality of $\zeta(3)$ given by R. Apery. The constant $\zeta(3)$ is become known as *Apery's constant*. A different proof of this result was presented by F. Beukers in [30]. His proof is based on the representation

$$\int_0^1 \int_0^1 \frac{P_n(x)\, P_n(y)}{1 - xy} \, \ln xy \, dx \, dy = 2(a_n - b_n \zeta(3)). \tag{7.22}$$

Here b_n are integers, a_n are rationals numbers such that $2\mathrm{LCM}[1, \cdots, n]a_n$ is an integer, and

$$P_n(z) = \frac{1}{n!} \frac{d^n}{dz^n} \left[z^n (1 - z)^n \right] \tag{7.23}$$

is the classical Legendre polynomial. This is one more instance of integrals at the center of interesting Mathematics.

From the data in (7.19) we see that $(-1)^n LC_n$ is a polynomial in the variables a, b, c with positive rational coefficients. Morever, this is a homogeneous polynomial of degree $n + 1$. Using a symbolic language we can test this conjecture. Indeed, the next term in the family is

$$LC_5 = -\left(\frac{1}{2}a^5c + \frac{15}{2}a^2b^3c + \frac{5}{12}a^3c^3 + \frac{5}{8}b^3c^3 + \frac{19}{96}ac^5 + \frac{45}{4}cd^5 \right), \tag{7.24}$$

where the new variable d is $\zeta(5)^{1/5}$. The conjecture has now been verified for $n = 5$.

The next steps in the analysis of this example are:
• Prove the conjecture. The reader will find in [31] a recurrence for the integrals LC_n. This should help.
• Discover a reason for its existence.
• Find a closed form for the coefficients in LC_n.

There are some other evaluations that seem to be *individuals*. The formula 4.229.7

$$\int_{\pi/4}^{\pi/2} \ln\ln\tan x\,dx = \frac{\pi}{2}\ln\left\{\frac{\Gamma(3/4)}{\Gamma(1/4)}\sqrt{2\pi}\right\} \tag{7.25}$$

that is the subject of the delightful paper by Ilan Vardi [225], is apparently in this category. The proof of (7.25) is based on Dirichlet L-series. This is evidence that to evaluate integrals one needs to learn Analytic Number Theory.

The flexibility provided by changes of variables permits us to represent Vardi's integrals in many forms. Some of these can be found in [124]. The new variable $u = \ln\tan x$ gives 4.371.1:

$$\int_0^\infty \frac{\ln u}{\cosh u}\,du = \pi\ln\left\{\frac{\Gamma(3/4)}{\Gamma(1/4)}\sqrt{2\pi}\right\}. \tag{7.26}$$

One of the advantages of a table in paper form is that it allows for *browsing*. A neighbor of the previous integral is 4.371.3

$$\int_0^\infty \frac{\ln x\,dx}{\cosh^2 x} = \ln\pi - 2\ln 2 - \gamma. \tag{7.27}$$

Here γ is Euler's constant defined by

$$\gamma = \lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{k} - \ln n. \tag{7.28}$$

The question of whether γ is rational is still open. We are sure that definite integrals will appear in its solution. J. Sondow has many interesting examples of integrals for γ, see

`http://home.earthlink.net/~jsondow.`

The two examples mentioned above are part of the family

$$LT_n := \int_0^\infty \frac{\ln x\,dx}{\cosh^n x}. \tag{7.29}$$

We encourage the reader to use a symbolic language to explore this new series of integrals.

Browsing also allows to find interesting examples in [124]. For instance, on page 575 we find as 4.375.1

$$\int_0^\infty \ln\cosh\frac{x}{2}\,\frac{dx}{\cosh x} = G + \frac{\pi}{4}\ln 2 \tag{7.30}$$

and a symbolic evaluation reveals that there is a sign error: the correct value of the integral is $G - \frac{\pi}{4}\ln 2$. Here G, known as *Catalan's constant*, is defined by

$$G = \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n-1)^2}. \tag{7.31}$$

The value of this integral appears written correctly as formula BI(259)(11) in [32]. It appears there in the equivalent form

$$\int_0^\infty \frac{\ln(e^{x/2}+e^{-x/2})}{e^x+e^{-x}}\,dx = \frac{\pi}{8}\ln 2 + \frac{1}{2}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^2}. \tag{7.32}$$

Symbolic extensions of this example are not very succesful: my current version of Mathematica is unable to evaluate

$$\int_0^\infty \ln\cosh\frac{x}{3}\,\frac{dx}{\cosh x}. \tag{7.33}$$

One of the intrinsic problems with symbolic languages is that the result in the evaluation depends how the integrand is input. For instance, Mathematica gives the value of

$$\int_0^\infty \ln\cosh x\,\frac{dx}{\cosh 3x} \tag{7.34}$$

as a complicated expression involving the *polylogarithm function* defined by

$$\mathrm{PolyLog}[k,x] := \sum_{n=1}^{\infty}\frac{x^n}{n^k}. \tag{7.35}$$

A simpler expression seems unlikely.

4 An Experimental Derivation of Wallis' Formula

The evaluation

$$J_{2,m} := \int_0^\infty \frac{dx}{(x^2+1)^{m+1}} = \frac{\pi}{2^{2m+1}}\binom{2m}{m} \tag{7.36}$$

is presented in some Calculus books. We present here an experimental derivation of it.

The first step is to produce convincing evidence that the right hand side of (7.36) is correct. This can be achieved by a symbolic evaluation of the integral. The Mathematica command

```
Integrate[1/(x^2+1)^(m+1), {x,0,Infinity}]
```

requests the value of $J_{2,m}$. The response involves the condition $\mathrm{Re}\, m > -\frac{1}{2}$ that guarantees convergence of the integral. This can be added to the request via the command

```
Assumptions -> Re[m] > -1/2.
```

The answer provided by Mathematica is given in terms of the gamma function as

$$J_{2,m} = \frac{\sqrt{\pi}}{2}\frac{\Gamma(\frac{1}{2}+m)}{\Gamma(1+m)}. \tag{7.37}$$

The expression for $J_{2,m}$ given in (7.36) now follows directly from the duplication formula for the gamma function

$$\Gamma(2x) = 2^{2x-1}\Gamma(x+\tfrac{1}{2})\Gamma(x)/\Gamma(\tfrac{1}{2}) \tag{7.38}$$

and the value $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$.

We now present an alternative form of obtaining (7.36) as an example of our experimental technique. Using a symbolic language (like Mathematica) we obtain the first few values of $J_{2,m}$ as

$$\left\{\frac{\pi}{4}, \frac{3\pi}{16}, \frac{5\pi}{32}, \frac{35\pi}{256}, \frac{63\pi}{512}, \frac{231\pi}{2048}\right\}. \tag{7.39}$$

It is reasonable to conjecture $J_{2,m}$ is a rational multiple of π. We now begin the exploration of this rational number by defining the function

$$f_2(m) := \mathrm{Denominator}(J_{2,m}/\pi). \tag{7.40}$$

Use the command

```
FactorInteger[Denominator[J[2,m]/Pi]]
```

to conclude that the denominator of $f_2(m)$ is a power of 2. Making a list of the first few exponents we (experimentally) conclude that

$$f_3(m) := 2^{2m} f_1(m)/\pi \tag{7.41}$$

is an integer. The first few values of this function are

$$\{1,\ 3,\ 10,\ 35,\ 126,\ 462,\ 1716,\ 6435,\ 24310,\ 92378\}. \tag{7.42}$$

Now comes the difficult part: we want to obtain a closed form expression for $f_3(m)$ directly from this list. For this we employ the *On-Line Encyclopedia of Integer Sequences* at the site

`http://www.research.att.com/~njas/sequences/.`

Entering the first four values we find that

$$f_3(m) = \binom{2m+1}{m+1} \tag{7.43}$$

is reasonable. This can be checked by computing more data. The expression for f_3 leads to the proposed form of $J_{2,m}$.

The exploration of Sloane's list is also a wonderful learning experience. The reader should use it and learn about the sequence

$$\{1,\ 3,\ 10,\ 35,\ 126,\ 462,\ 1717\,\} \tag{7.44}$$

that we got when we made a mistake and typed 1717 instead of 1716.

We now prove that

$$J_{2,m} = \int_0^\infty \frac{dx}{(x^2+1)^{m+1}} = \frac{\pi}{2^{2m+1}}\binom{2m}{m}, \tag{7.45}$$

where m is a nonnegative integer. The change of variables $x = \tan\theta$ converts $J_{2,m}$ to its trigonometric form

$$J_{2,m} = \int_0^{\pi/2} \cos^{2m}\theta\, d\theta = \frac{\pi}{2^{2m+1}}\binom{2m}{m}, \tag{7.46}$$

which is known as Wallis's formula. The proof of (7.46) is elementary and sometimes found in calculus books (see e.g. [160], page 492). It consists of first writing $\cos^2\theta = 1 - \sin^2\theta$ and using integration by parts to obtain the recursion

$$J_{2,m} = \frac{2m-1}{2m} J_{2,m-1}, \tag{7.47}$$

and then verifying that the right side of (7.46) satisfies the same recursion and that both sides yield $\pi/2$ for $m = 0$.

We now present a new proof of Wallis's formula. We have

$$J_{2,m} = \int_0^{\pi/2} \cos^{2m}\theta\, d\theta = \int_0^{\pi/2} \left(\frac{1+\cos 2\theta}{2}\right)^m d\theta.$$

Now introduce $\psi = 2\theta$ and expand and simplify the result by observing that the odd powers of cosine integrate to zero. The inductive proof of (7.46) requires

$$J_{2,m} = 2^{-m} \sum_{i=0}^{[m/2]} \binom{m}{2i} J_{2,i}. \tag{7.48}$$

Note that $J_{2,m}$ is uniquely determined by (7.48) along with the initial value $J_{2,0} = \pi/2$. Thus (7.46) now follows from the identity

$$f(m) := \sum_{i=0}^{[m/2]} 2^{-2i} \binom{m}{2i} \binom{2i}{i} = 2^{-m} \binom{2m}{m} \tag{7.49}$$

since (7.49) can be written as

$$A_m = 2^{-m} \sum_{i=0}^{[m/2]} \binom{m}{2i} A_i,$$

where

$$A_i = \frac{\pi}{2^{2i+1}} \binom{2i}{i}.$$

It remains to verify the identity (7.49). This can be done *mechanically* using the theory developed by Wilf and Zeilberger, which is explained in [184, 190]; the sum in (7.49) is the example used in [190] (page 113) to illustrate their method. The command

```
ct(binomial(m,2i)  binomial(2i,i) 2^{-2i}, 1, i, m,N)
```

produces

$$f(m+1) = \frac{2m+1}{m+1}\, f(m), \tag{7.50}$$

and one checks that $2^{-m}\binom{2m}{m}$ satisfies the same recursion. Note that (7.47) and (7.50) are equivalent since

$$J_{2,m} = \frac{\pi}{2^{m+1}} f(m).$$

This proof is more complicated than using (7.47), but the method behind it applies to other rational integrals. This was the first step in a series of results on rational Landen transformations. The reader will find in [38] the details for even rational functions and [174] for recent developments.

5 A Hyperbolic Example

Section 3.511 of the table of integrals [124] contains the evaluation of several definite integrals whose integrands are quotients of hyperbolic functions. For instance, 3.511.2 reads

$$\int_0^\infty \frac{\sinh(ax)}{\sinh(bx)}\, dx = \frac{\pi}{2b} \tan \frac{a\pi}{2b} \tag{7.51}$$

while the next formula 3.511.3 is

$$\int_0^\infty \frac{\sinh(ax)}{\cosh(bx)}\,dx = \frac{\pi}{2b}\sec\frac{a\pi}{2b} - \frac{1}{b}\beta\left(\frac{a+b}{2b}\right). \tag{7.52}$$

The reader will find in [124], formula 8.371.1, the function β given by the integral representation

$$\beta(x) = \int_0^1 \frac{t^{x-1}}{1+t}\,dt. \tag{7.53}$$

Introduce the notation

$$g_1(a,b) = \int_0^\infty \frac{\sinh(ax)}{\cosh(bx)}\,dx \tag{7.54}$$

and we now discuss its evaluation. Naturally the answer given in the table is in terms of the β-function, so its integral representation suggests a procedure how to start. *Checking the value of an integral is much easier than finding it.* Simply let $t = e^{-2bx}$ to obtain

$$g_1(a,b) = \beta\left(\frac{1}{2} - \frac{a}{2b}\right) - \beta\left(\frac{1}{2} + \frac{a}{2b}\right). \tag{7.55}$$

We now use properties of β to reduce this answer to the one given in [124]. This function is related to the classical *digamma* function

$$\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)} \tag{7.56}$$

via

$$\beta(x) = \frac{1}{2}\left(\psi\left(\frac{x+1}{2}\right) - \psi\left(\frac{x}{2}\right)\right). \tag{7.57}$$

The functional equation for the gamma function

$$\Gamma(1+x) = x\Gamma(x) \tag{7.58}$$

yields the identity

$$\psi(1+x) = \psi(x) + \frac{1}{x}. \tag{7.59}$$

In turn this produces

$$\beta(x+1) = -\beta(x) + \frac{1}{x}. \tag{7.60}$$

Similarly

$$\psi\left(\frac{1}{2}+x\right) = \psi\left(\frac{1}{2}-x\right) + \pi\tan\pi x \tag{7.61}$$

produces

$$\beta(1-x) = \beta(x) + \pi \operatorname{cosec} \pi x. \tag{7.62}$$

Using these identities in (7.55) yields (7.52).

But suppose that the reader is not familiar with these functions. Is there anything that one can learn from the integrals by using a symbolic language?

A blind evaluation[2] produces an answer in terms of the function

$$h(x) := \text{HarmonicNumber}[x] := \sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+x} \right) \tag{7.63}$$

as

$$\begin{aligned} g_1(a,b) &= \frac{\text{HarmonicNumber}\left[\frac{1}{4}(-3+\frac{a}{b})\right]}{4b} - \frac{\text{HarmonicNumber}\left[\frac{1}{4}(-1+\frac{a}{b})\right]}{4b} \\ &+ \frac{\text{HarmonicNumber}\left[-\frac{a+b}{4b}\right]}{4b} - \frac{\text{HarmonicNumber}\left[-\frac{a+3b}{4b}\right]}{4b}. \end{aligned}$$

The answer is written exactly as it appears in the Mathematica code. To simplify it we apply the command

`FunctionExpand`

that produces

$$\begin{aligned} g_1(a,b) &= -\frac{\text{PolyGamma}\left[0, \frac{1}{4} - \frac{a}{4b}\right]}{4b} + \frac{\text{PolyGamma}\left[0, \frac{3}{4} - \frac{a}{4b}\right]}{4b} \\ &+ \frac{\text{PolyGamma}\left[0, \frac{1}{4} + \frac{a}{4b}\right]}{4b} - \frac{\text{PolyGamma}\left[0, \frac{3}{4} + \frac{a}{4b}\right]}{4b} \end{aligned}$$

The polygamma function appearing here is simply $\psi(x)$ and it is easy to see that this expression for g_1 is equivalent to (7.55). We also see from here that $4bg_1(a,b)$ is a function of the single parameter $c = \frac{a}{b}$. This is elementary. The change of variables $t = bx$ yields

$$4bg_1(a,b) = 4 \int_0^{\infty} \frac{\sinh(ct)}{\cosh t}\, dt. \tag{7.64}$$

Thus we write

$$g_2(c) = 4bg_1(a,b). \tag{7.65}$$

[2]This is one in which we simply input the question and ask Mathematica for an answer.

and the previous evaluation is written as

$$\begin{aligned} g_2(c) &= -\text{PolyGamma}\left[0, \tfrac{1}{4} - \tfrac{c}{4}\right] + \text{PolyGamma}\left[0, \tfrac{3}{4} - \tfrac{c}{4}\right] \\ &\quad +\text{PolyGamma}\left[0, \tfrac{1}{4} + \tfrac{c}{4}\right] - \text{PolyGamma}\left[0, \tfrac{3}{4} + \tfrac{c}{4}\right]. \end{aligned}$$

In order to obtain some information about the function g_2 we can use Mathematica to create a list of values. The command

```
T_{1} :=  Table[g_{2}[n], {n,1, 10}]
```

produces

$$T_1 = \left\{\infty, -8, \infty, \tfrac{16}{3}, \infty, -\tfrac{104}{15}, \infty, \tfrac{608}{105}, \infty, -\tfrac{2104}{315}\right\}. \tag{7.66}$$

This indicates the presence of singularities for the function g_2 at the odd integers and a clear pattern for the signs at the even ones. We then compute the table of values of $(-1)^n g_2(2n)$ for $1 \leq n \leq 9$:

$$T_2 = \left\{8, \tfrac{16}{3}, \tfrac{104}{15}, \tfrac{608}{105}, \tfrac{2104}{315}, \tfrac{20624}{3465}, \tfrac{295832}{45045}, \tfrac{271808}{45045}, \tfrac{4981096}{765765}\right\}.$$

From here we experimentally conclude that g_2 has a singularity at the odd integers and that $r_n = (-1)^n g_2(2n)$ is a positive rational number. We leave the singularity question aside and explore the properties of the sequence r_n.

The beginning of the sequence of denominators agrees with that of the *odd semi-factorials*:

$$(2n-1)!! := (2n-1)\cdot(2n-3)\cdot(2n-5)\cdots 5\cdot 3\cdot 1 \tag{7.67}$$

that begins as

$$T_3 = \{1, 3, 15, 105, 945, 10395, 135135, 2027025\}. \tag{7.68}$$

Thus it is natural to consider the function

$$g_3(n) = (-1)^n (2n-1)!! g_2(2n). \tag{7.69}$$

The hope is that $g_3(n)$ is an integer valued function. The first few values are given by

$$T_4 = \{8, 16, 104, 608, 6312, 61872, 887496, 12231360\}. \tag{7.70}$$

Observe that they are all even integers. To examine their divisibility in more detail, we introduce the concept of p-adic valuation.

Let $n \in \mathbb{N}$ and p be a prime. The p-adic valuation of n, denoted by $\nu_p(n)$, is defined as the exact power of p that divides n. Now define

$$g_4(n) = \nu_2(g_3(n)), \tag{7.71}$$

and its first few values are

$$T_5 = \{0,\ 1,\ 0,\ 2,\ 0,\ 1,\ 0,\ 3,\ 0,\ 1,\ 0,\ 2,\ 0,\ 1,\ 0,\ 4\}. \tag{7.72}$$

From this data we conjecture that

$$g_4(n) = 3 + \nu_2(n), \tag{7.73}$$

so that the (conjectured) integer

$$g_5(n) := \frac{g_4(n)}{2^{3+\nu_2(n)}} \tag{7.74}$$

is odd.

We now continue this process and examine the number $\nu_3(g_5(n))$. It seems that $g_5(n)$ is divisible by 3 for $n \geq 5$, so we consider

$$g_6(n) = \nu_3(g_5(n)). \tag{7.75}$$

Extensive symbolic calculations show that for $i \geq 2$, we have

$$g_6(3i-1) = g_6(3i) = g_6(3i+1) \tag{7.76}$$

We pause our experiment here. The reader is invited to continue this exploration and begin proving some of these results.

The method outlined here is referred by us as the *peeling strategy*: given an expression for which we desire a closed-form, we use a symbolic language to peel away recognizable parts. The process ends successfully if one is able to reduce the original expression to a known one.

6 A Formula Hidden in the List

It is impossible for a table of integrals to be complete. In the process of studying integrals of rational functions and after completing our study of Wallis' formula we consider the integral

$$N_{0,4}(a;m) := \int_0^\infty \frac{dx}{(x^4+2ax^2+1)^{m+1}}. \tag{7.77}$$

We were surprised not to find it in [124]. The table does contain formula 3.252.11

$$\int_0^\infty (1+2\beta x+x^2)^{\mu-\frac{1}{2}} x^{-\nu-1}\,dx = \tag{7.78}$$

$$2^{-\mu}(\beta^2-1)^{\mu/2}\Gamma(1-\mu)B(\nu-2\mu+1,-\nu)P^{\mu}_{\nu-\mu}(\beta)$$

where B is the *beta function* and P^{μ}_{ν} is the *associated Legendre function*. We will not discuss here the technical details required to prove this. Using an appropriate representation of these functions, we showed that

$$P_m(a) := \frac{1}{\pi}2^{m+3/2}(a+1)^{m+1/2}N_{0,4}(a;m) \tag{7.79}$$

is a polynomial in a, given by

$$P_m(a) = 2^{-2m}\sum_{k=0}^{m}2^k\binom{2m-2k}{m-k}\binom{m+k}{m}(a+1)^k. \tag{7.80}$$

The reader will find the details in [36]. It turns out that this was not the original approach we followed to establish (7.79). At the time we were completely unaware of the *hypergeometric world* involved in (7.78). Instead we used

$$\int_0^{\infty}\frac{dx}{bx^4+2ax^2+1} = \frac{\pi}{2\sqrt{2}}\,\frac{1}{\sqrt{a+\sqrt{b}}} \tag{7.81}$$

to produce in [37] the expansion

$$\sqrt{a+\sqrt{1+c}} = \sqrt{a+1} + \frac{1}{\pi\sqrt{2}}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}N_{0,4}(a;k-1)c^k. \tag{7.82}$$

The expression for $P_m(a)$ was a corollary of Ramanujan Master Theorem; see [28] for details on this and many other results of Ramanujan. It was a long detour, forced by ignorance of a subject. It was full of surprises.

The coefficients of the polynomial $P_m(a)$ are given by

$$d_l(m) = 2^{-2m}\sum_{k=l}^{m}2^k\binom{2m-2k}{m-k}\binom{m+k}{m}\binom{k}{l}, \tag{7.83}$$

have many interesting properties. Many of which we discovered by *playing around* with them in a computer.

• The sequence $\{d_l(m) : 0 \le l \le m\}$ is *unimodal*. That is, there is an index l^* such that $d_i(m) \le d_{i+1}(m)$ if $0 \le i \le l^*-1$ and the inequality is reversed if $l^* \le i \le m$. In our study of $d_l(m)$, we discovered a general unimodality criteria. The details are presented in [7] and [35]. A property stronger than unimodality is that of *log-concavity*: a sequence of numbers $\{a_l\}$ is called log-concave if it satisfies $a_l^2 \ge a_{l-1}a_{l+1}$. We have conjectured that the coefficients $d_l(m)$ are logconcave, but so far we have not been able to prove this.

• The representation (7.83) provides an efficient way to compute $d_l(m)$ for m fixed if l is large relative to m. Trying to produce an alternative form, that would give an efficient way to compute them when l is small, we established in [40] the existence of two families of polynomials $\alpha_l(m)$ and $\beta_l(m)$ such that

$$d_l(m) = \frac{1}{l!\, m!\, 2^{m+l}} \left(\alpha_l(m) \prod_{k=1}^{m} (4k-1) - \beta_l(m) \prod_{k=1}^{m} (4k+1) \right). \quad (7.84)$$

The degrees if α_l and β_l are l and $l-1$ respectively. For instance, the linear coefficient of $P_m(a)$ is given by

$$d_1(m) = \frac{1}{m!\, 2^{m+1}} \left((2m+1) \prod_{k=1}^{m} (4k-1) - \prod_{k=1}^{m} (4k+1) \right), \quad (7.85)$$

and the quadratic one is

$$d_2(m) = \frac{1}{m!\, 2^{m+2}} \left((2m^2+2m+1) \prod_{k=1}^{m} (4k-1) - (2m+1) \prod_{k=1}^{m} (4k+1) \right).$$

On an day without new ideas[3], we started computing zeros of the polynomials α_l and β_l. We were very surprised to see that *all of them* were on a the vertical line $\operatorname{Re} m = -\frac{1}{2}$. In the summer of 2002, while working at SIMU (Summer Institute in Mathematics for Undergraduates), the author had a good idea about how to solve this problem: ask John Little. The result is true. The polynomial $A_l(s) := \alpha_l(\frac{s-1}{2})$ satisfies the recurrence

$$A_{l+1}(s) = 2sA_l(s) - (s^2 - (2l-1)^2)A_{l-1}(s) \quad (7.86)$$

and the location of the zeros can be deduced from here. The details will appear in [169].

This example illustrates our working philosophy: there are many classes of definite integrals that have very interesting Mathematics hidden in them. The use of a Symbolic Language often helps you find the correct questions.

7 Some Experiments on Valuations

We now report on some experiments concerning the p-adic valuation of the coefficients $d_l(m)$ defined in the previous section. The expression

$$d_l(m) = 2^{-2m} \sum_{k=l}^{m} 2^k \binom{2m-2k}{m-k} \binom{m+k}{m} \binom{k}{l} \quad (7.87)$$

[3] Like most of them.

shows that $\nu_p(d_l(m)) \geq 0$ for $p \neq 2$.

We now describe results of symbolic calculations of the p-adic valuation of

$$f(m,l) = \alpha_l(m) \prod_{k=1}^{m} (4k-1) - \beta_l(m) \prod_{k=1}^{m} (4k+1). \tag{7.88}$$

The coefficient $d_l(m)$ and $f(m,l)$ are related via $f(m,l) = l!m!2^{m+l}d_l(m)$. These two functions are computationally equivalent because the p-adic valuations of factorials are easily computed via

$$\nu_p(m!) = \sum_{k=1}^{\infty} \lfloor \frac{m}{p^k} \rfloor. \tag{7.89}$$

Naturally the sum is finite and we can end it at $k = \lfloor \log_p m \rfloor$. An alternative is to use a famous result of Legendre [162, 125]

$$\nu_p(m!) = \frac{m - s_p(m)}{p-1} \tag{7.90}$$

The command

```
g[m_,l_,p_]:= IntegerExponent[ f[m,l],p ]
```

gives directly the p-adic valuation of the coefficient $f(m,l)$. For example, for $m = 30, l = 10$ and $p = 7$ we have $g(30,10,3) = 18$. Indeed,

$$f(30,10) = 2^{30} \cdot 3^{18} \cdot 5^{10} \cdot 7^6 \cdot 11^4 \cdot 13^3 \cdot 17^2 \cdot 19^2 \cdot 23^2 \cdot N \tag{7.91}$$

where N is the product of five primes.

We have evaluated the function $g(m,l,p)$ for large classes of integers m, l and primes p.

The remainder of this section describes our findings. The situation is quite different according to whether $p = 2$ or odd. In the first case, $g(m,l,2)$ has a well defined structure. In the latter, $g(m,l,p)$ has a random flavor. In this case we have been only able to predict its asymptotic behavior.

The case $p = 2$.

• The experiments show that, for fixed l, the function $g(m,l,2)$ has blocks of length $2^{\nu_2(l)+1}$. For example, the values of $g(m,1,2)$ for $1 \leq m \leq 20$ are

$$\{2,2,3,3,2,2,4,4,2,2,3,3,2,2,5,5,2,2,3,3\}.$$

Similarly, $g(m, 4, 2)$ has blocks of length 8: it begins with eight 11, followed by eight 12, continued by eight 11, and so on.

• The graphs of the function $g(m, l, 2)$, where we reduced the repeating blocks to a single value are shown in the next figures.

The main experimental result is that the graph of $g(m, l, p)$ has an *initial segment* from which the rest is determined by adding a *central piece* followed by a *folding rule*. For example, in the case $l = 1$, the first few values of the reduced table for $g(m, 1, 2)$ are

$$\{2,\ 3,\ 2,\ 4,\ 2,\ 4,\ 2,\ 3,\ 2,\ 5, \ldots\}$$

The ingredients are:

initial segment: $\{2,\ 3,\ 2\}$,

central piece: the value at the center of the initial segment, namely 3.

rules of formation: start with the initial segment and add 1 to the central piece and reflect.

This produces the sequence

$$\{2,3,2\} \to \{2,3,2,4\} \to \{2,3,2,4,2,3,2\} \to \{2,3,2,4,2,3,2,5\} \to$$

$$\to \{2,3,2,4,2,3,2,5,2,3,2,4,2,3,2\}.$$

The details are shown in Figure 7.1.

The difficulty with this procedure is that, at the moment, we have no form of determining the initial segment nor the central piece. Figure 7.2 shows the beginning of the function $g(m, 9, 2)$. From here one could be tempted to predict that this graph extends as in the case $l = 1$. This is not correct as it can be seen in Figure 7.3. The new pattern described seems to be the correct one as shown in Figure 7.4.

The initial pattern could be quite elaborate. Figure 7.5 illustrates the case $l = 53$.

In [40] we have given details of the only analytic result: the 2-adic valuation of $d_1(m)$ is given by

$$\nu_2(d_1(m)) = 1 - m + \nu_2\left(\binom{m+1}{2}\right) - \nu_2(m!). \tag{7.92}$$

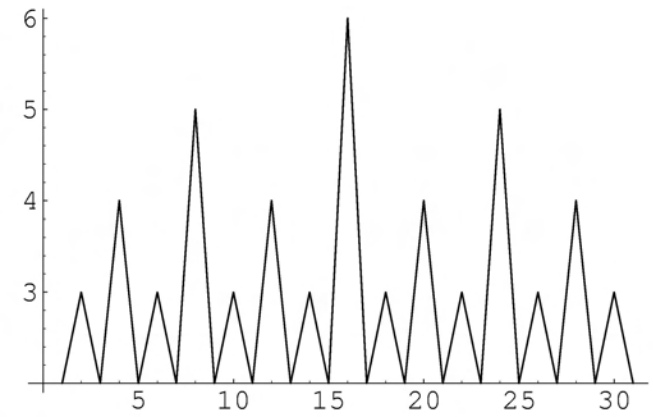


Figure 7.1. The 2-adic valuation of $d_1(m)$

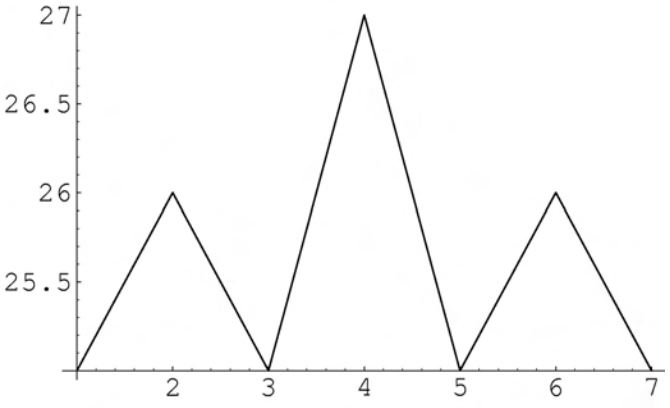


Figure 7.2. The beginning of $g(m, 9, 2)$.

The case of odd primes.

The p-adic valuation of $d_l(m)$ behaves quite differently for p an odd prime. Figure 7.6 shows the values for $l = 1$ and $p = 3$. There is a clear linear growth, of slope $\frac{1}{2}$ and Figure 7.7 shows the deviation from this linear function. We suspect that the error is bounded.

This behavior persists for other values of l and primes p. The p-adic valuation of $d_l(m)$ has linear growth with slope $1/(p-1)$. Figures 7.8, 7.10, 7.9, 7.11 show four representative cases.

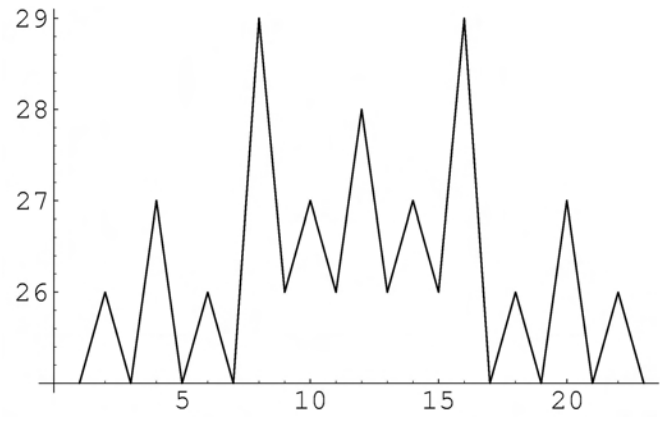


Figure 7.3. The continuation of $g(m, 9, 2)$.

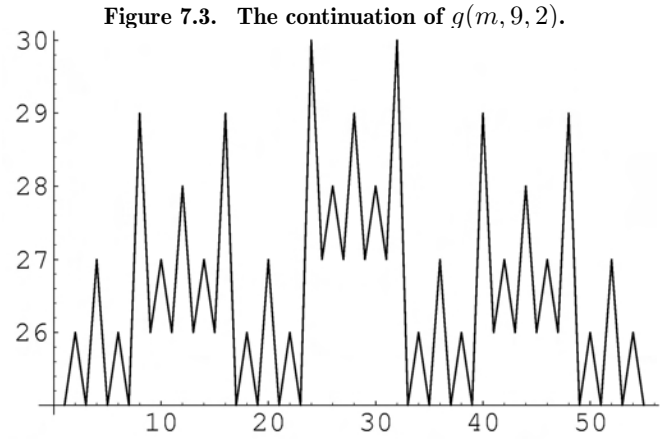


Figure 7.4. The pattern of $g(m, 9, 2)$ persists.

Conjecture. Let p be an odd prime. Then

$$\nu_p(d_l(m)) \sim \frac{m}{p-1}. \tag{7.93}$$

8 An Error in the Latest Edition

A project of the size of [124] is bound to have some errors. In the last edition of the table, we found formula 3.248.5 as a new addition to this

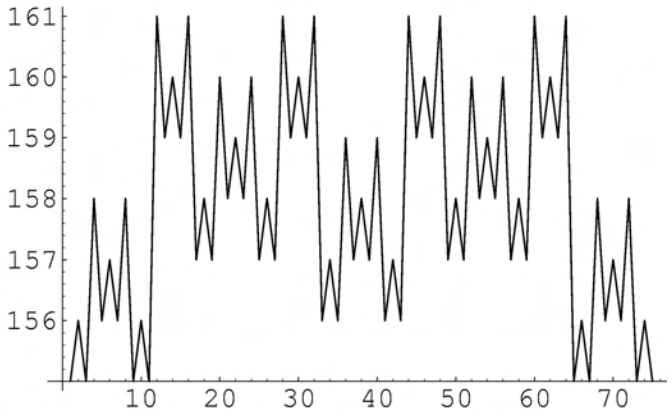


Figure 7.5. The initial pattern for $l = 53$.

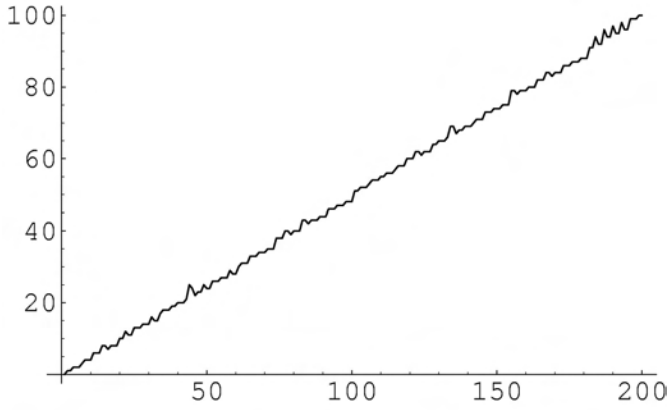


Figure 7.6. The 3-adic valuation of $d_1(m)$.

great collection. It has a beautiful structure given by nested radicals. Let

$$\varphi(x) = 1 + \frac{4x^2}{3(1+x^2)^2}, \tag{7.94}$$

then the evaluation states that

$$\int_0^\infty \frac{dx}{(1+x^2)^{3/2}\,\left[\varphi(x)+\varphi(x)^{1/2}\right]^{1/2}} = \frac{\pi}{2\sqrt{6}}. \tag{7.95}$$

After several failed attemps at proving it, the author decided to check it numerically. *It is incorrect*[4]. Normally this is disappointing, but in this

[4]We should have checked earlier.

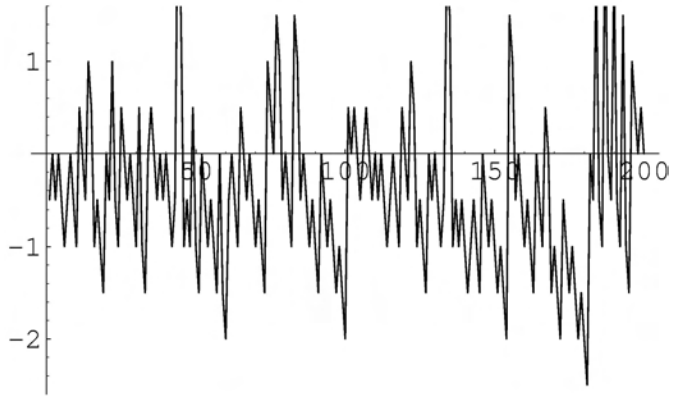


Figure 7.7. The deviation from linear growth.

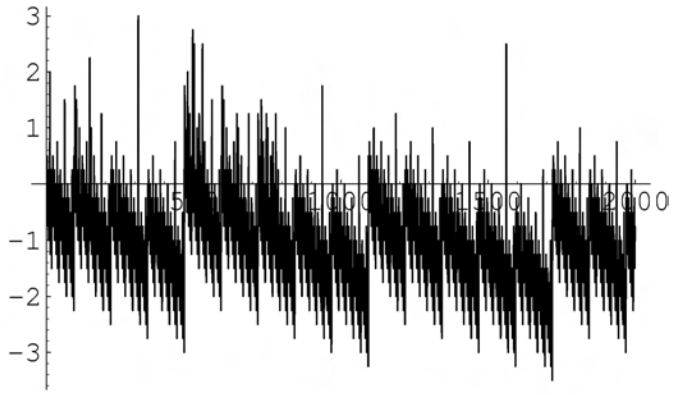


Figure 7.8. The error for $l = 1$ and $p = 5$.

case the structure of the evaluation leads to two interesting questions:

• *The direct problem.* Find the correct value of the integral. As usual there is no method that is guaranteed to succeed and a good idea is required.

• *The inverse problem.* Given that $\pi/2\sqrt{6}$ is the correct answer, find a modification of the integrand that produces that answer. This modification is expected to be close to the integrand given in the table, perhaps a typo: the exponent $\frac{3}{2}$ in the term $1 + x^2$ perhaps is $\frac{2}{3}$; or the 4 in the expression for $\varphi(x)$ is a 14. There is no systematic approach to solving this problem, but it is a lot of fun to experiment with.

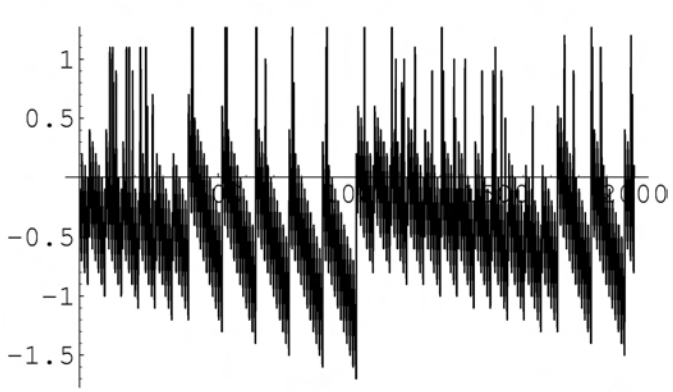


Figure 7.9. The error for $l = 1$ and $p = 11$.

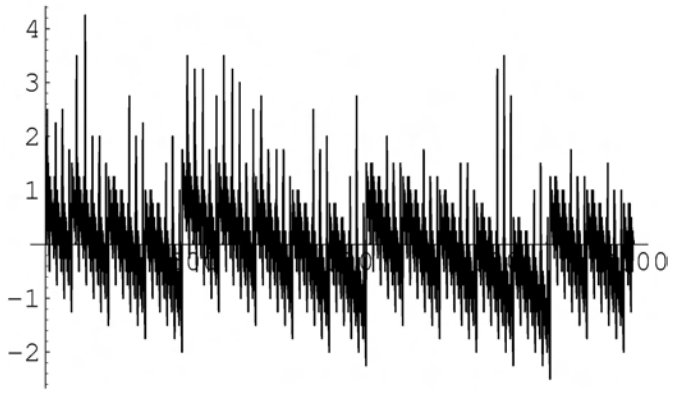


Figure 7.10. The error for $l = 5$ and $p = 5$.

We have not explore this example in detail but it seems to have many beautiful alternative representations. For instance, let

$$a[x, p] := \sqrt{x^4 + 2px^2 + 1}. \tag{7.96}$$

Then the integral is $I(\frac{5}{3}, 1)$ where

$$I(p, q) = \int_0^\infty \frac{dx}{a[x, p]^{1/2}\, a[x, q]^{1/2} (a[x, p] + a[x, q])^{1/2}}. \tag{7.97}$$

It is interesting to observe that if

$$b[x, p] := \sqrt{x^2 + 2px + 1} \tag{7.98}$$

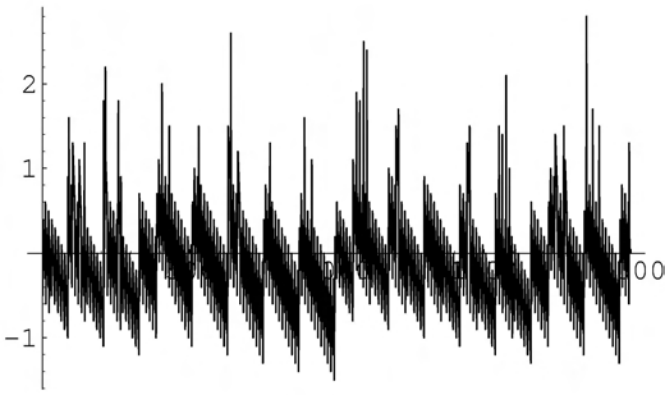


Figure 7.11. The error for $l = 5$ and $p = 11$.

then the integral is $J(\frac{5}{3}, 1)$ where

$$J(p,q) = \frac{1}{2}\int_0^\infty \frac{dx}{b[x,p]^{1/2}\, b[x,q]^{1/2}(b[x,p]+b[x,q])^{1/2}}. \tag{7.99}$$

9 Some Examples Involving the Hurwitz Zeta Function

There are many evaluations in [124] where the Hurwitz zeta function

$$\zeta(z,q) = \sum_{k=0}^{\infty} \frac{1}{(k+q)^z} \tag{7.100}$$

appears as part of the value of the integral. For example, 3.524.1 states that

$$\int_0^\infty x^{\mu-1}\frac{\sinh\beta x}{\sinh\gamma x}\,dx = \frac{\Gamma(\mu)}{(2\gamma)^\mu}\left\{\zeta\left[\mu,\frac{1}{2}\left(1-\frac{\beta}{\gamma}\right)\right]-\zeta\left[\mu,\frac{1}{2}\left(1+\frac{\beta}{\gamma}\right)\right]\right\}$$

valid for $\operatorname{Re}\gamma > |\operatorname{Re}\beta|, \quad \operatorname{Re}\mu > -1$. The identity (7.101) can be written in the *typographically* simpler form

$$\int_0^\infty x^{\mu-1}\frac{\sinh bx}{\sinh x}\,dx = \frac{\Gamma(\mu)}{2^\mu}\left\{\zeta\left[\mu,\frac{1}{2}(1-b)\right]-\zeta\left[\mu,\frac{1}{2}(1+b)\right]\right\},$$

that illustrates the fact that 3.524.1 has only two independent parameters.

The table [124] contains no examples in which $\zeta(z,q)$ appears in the integrand. A search of the 3-volume compendium [202] produces

$$\int \zeta(z,q)dq = \frac{1}{1-z}\zeta(z-1,q), \tag{7.101}$$

which is an elementary consequence of

$$\frac{\partial}{\partial q}\zeta(z-1,q) = (1-z)\zeta(z,q), \tag{7.102}$$

and in Section 2.3.1 we find the moments

$$\int_0^\infty q^{\alpha-1}\zeta(z,a+bq)dq = b^{-\alpha}B(\alpha,z-\alpha)\zeta(z-\alpha,a). \tag{7.103}$$

Motivated mostly by the lack of explicit evaluations, we initiated in [106, 107] a systematic study of these type of integrals. Our results are based mostly on the expressions for the Fourier coefficients of $\zeta(z,q)$ given in Section 2.3.1 of [202] as

$$\int_0^1 \sin(2\pi q)\zeta(z,q)dq = \frac{(2\pi)^z}{4\Gamma(z)}\ \csc\left(\frac{z\pi}{2}\right). \tag{7.104}$$

Some interesting evaluations are obtained from Lerch's identity

$$\frac{d}{dz}\zeta(z,q)|_{z=0} = \ln\Gamma(q) - \ln\sqrt{2\pi}. \tag{7.105}$$

For instance, the reader will find in [106]) the elementary value

$$\int_0^1 \ln\Gamma(q)dq = \ln\sqrt{2\pi}, \tag{7.106}$$

and the surprising one

$$\int_0^1 (\ln\Gamma(q))^2\,dq = \frac{\gamma^2}{12} + \frac{\pi^2}{48} + \frac{1}{3}\gamma\ln\sqrt{2\pi} + \frac{4}{3}\left(\ln\sqrt{2\pi}\right)^2 - (\gamma + 2\ln\sqrt{2\pi})\frac{\zeta'(2)}{\pi^2} + \frac{\zeta''(2)}{2\pi^2}. \tag{7.107}$$

An elementary proof of (7.106), due to T. Amdeberham, uses only the elementary properties of the gamma function. The first step is to partition the interval $[0,1]$ to obtain

$$\int_0^1 \ln\Gamma(x)\,dx = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{\infty}\ln\Gamma(k/n). \tag{7.108}$$

The Riemann sum can be written as

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}\ln\Gamma(k/n) &= \frac{1}{n}\ln\left(\prod_{k=1}^{n}\Gamma\left(\tfrac{k}{n}\right)\right)\\ &= \frac{1}{n}\ln\left(\prod_{k=1}^{n/2}\Gamma\left(\tfrac{k}{n}\right)\Gamma\left(1-\tfrac{k}{n}\right)\right)\end{aligned}$$

and using the reflection formula

$$\Gamma(z)\Gamma(1-z)=\frac{\pi}{\sin\pi z} \tag{7.109}$$

we obtain

$$\frac{1}{n}\sum_{k=1}^{n}\ln\Gamma(k/n)=\ln\sqrt{\pi}-\ln\left(\prod_{k=1}^{n/2}\sin(\pi k/n)\right)^{1/n}. \tag{7.110}$$

The identity

$$\prod_{k=1}^{n-1}\sin\left(\frac{\pi k}{n}\right)=\frac{n}{2^{n-1}} \tag{7.111}$$

written as

$$\left(\prod_{k=1}^{n/2}\sin\left(\frac{\pi k}{n}\right)\right)^{1/n}=\frac{(2n)^{1/2n}}{\sqrt{2}} \tag{7.112}$$

yields the evaluation.

We now observe that, for $n=1$ and $n=2$, the integral

$$LG_n:=\int_0^1(\ln\Gamma(q))^n\,dq \tag{7.113}$$

is a homogeneous polynomial of degree n in the variables γ, $\pi, \ln\sqrt{2\pi}$, $\zeta(2)$, $\zeta'(2)$ and $\zeta''(2)$. This is similar to (7.19). In this problem the weights are assigned experimentally as follows:

- The weight of a rational number is 0,
- The constants π, γ have weight 1 and so does $\ln\sqrt{2\pi}$, that is,

$$w(\pi)=w(\gamma)=w(\ln\sqrt{2\pi}=1, \tag{7.114}$$

- The weight is extended as $w(ab)=w(a)+w(b)$.
- The value $\zeta(j)$ has weight j, so that $w(\zeta(2))=2$ is consistent with $\zeta(2)=\pi^2/6$.

• Differentiation increases the weight by one, so that $\zeta''(2)$ has weight 4.

We have been unable to evaluate the next example

$$LG_3 := \int_0^1 (\ln \Gamma(q))^3 \, dq, \tag{7.115}$$

but in our (failed) attempts we have established in [108] a connection between LG_n and the Tornheim-Zagier sums

$$T(a,b;c) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{n^a m^b (n+m)^c}. \tag{7.116}$$

These sums will reappear in Chapter 8. An extensive bibliography on these *multiple zeta series* has been compiled by M. Hoffman and it can be found in

`http://www.usna.edu/Users/math/meh/biblio.html`

In our current work, we have been able to express the sums $T(a,b;c)$ in terms of integrals of triple products of the Bernoulli polynomials and the function

$$A_k(q) = k\zeta'(1-k,q). \tag{7.117}$$

The classical identity

$$B_k(q) = -k\zeta(1-k,q) \tag{7.118}$$

shows the similarity between A_k and B_k.

This is just one more example of the beautiful Mathematics that one is able to find by experimenting with integrals.

Chapter 8

Experimental Mathematics: a Computational Conclusion

We have now seen experimental computationally-assisted mathematics in action in areas pure and applied, discrete and continuous. In each case perhaps Goethe said it best:

> *"Mathematicians are a kind of Frenchmen: whatever you say to them they translate into their own language, and right away it is something entirely different."*(Johann Wolfgang von Goethe[1])

Goethe was right and we do it to wonderful effect. Returning however to views bout the nature of mathematics Greg Chaitin takes things much further than Gödel did.

> *Over the past few decades, Gregory Chaitin, a mathematician at IBM's T.J. Watson Research Center in Yorktown Heights, N.Y., has been uncovering the distressing reality that much of higher math may be riddled with unprovable truths–that it's really a collection of random facts that are true for no particular reason. And rather than deducing those facts from simple principles,* "I'm making the suggestion that mathematics is done more like physics in that you come about things experimentally,*" he says.* "This will still be controversial when I'm dead. It's a major change in how you do mathematics." (Time Magazine, Sept 4, 2005)

Chaitin was featured in a brief article, *the Omega Man*, which is the name of his new book on computational complexity.

[1] Maximen und Reflexionen, no. 1279

This 'Ω' refers to Chaitin's seminal halting probability constant[2]

$$\Omega := \sum_{\pi} 2^{-\#(\pi)}$$

where π ranges over halting Turing machines and $\#(\pi)$ is its length . While intrinsically non-computable and *algorithmically irreducible*, the first 64 bits are provably known

$\Omega := 0.0000001000000100001000001000011101110011001001111000100100111 00\ldots$

"Most of mathematics is true for no particular reason," Chaitin says. "Maths is true by accident."

1 A Little More History

In his '23' "*Mathematische Probleme*"[3] lecture to the Paris International Congress in 1900, David Hilbert wrote

> *""Moreover a mathematical problem should be difficult in order to entice us, yet not completely inaccessible, lest it mock our efforts. It should be to us a guidepost on the mazy path to hidden truths, and ultimately a reminder of our pleasure in the successful solution."*

Here, I'll also add some of the missing references about the **History of Computer Experiments** in Math [Shallit] and in Science from Lehmer to Simons.

> Ask Dr. Edward Witten of the Institute for Advanced Study in Princeton, New Jersey what he does all day, and it's difficult to get a straight answer. "*There isn't a clear task*," Witten told CNN. *"If you are a researcher you are trying to figure out what the question is as well as what the answer is.*
>
> *"You want to find the question that is sufficiently easy that you might be able to answer it, and sufficiently hard that the answer is interesting. You spend a lot of time thinking and you spend a lot of time floundering around."*(Ed Witten[4])

[2] `http://www.umcs.maine.edu/ chaitin/`

[3] See the late Ben Yandell's fine account of the Hilbert Problems and their solvers in *The Honors Class*, AK Peters, 2002.

[4] Witten was interviewed on CNN June 27, 2005

2 Putting Lessons in Action

In "Proof and beauty", the *Economist* of March 31, 2005 wrote

> *" Just what does it mean to prove something? Although the Annals will publish Dr Hales's paper, Peter Annals, an editor of the Annals, whose own work does not involve the use of computers, says that the paper will be accompanied by an unusual disclaimer, stating that the computer programs accompanying the paper have not undergone peer review. There is a simple reason for that, Dr Sarnak says-it is impossible to find peers who are willing to review the computer code. However, there is a flip-side to the disclaimer as well-Dr Sarnak says that the editors of the Annals expect to receive, and publish, more papers of this type-for things, he believes, will change over the next 20-50 years. Dr Sarnak points out that maths may become "a bit like experimental physics" where certain results are taken on trust, and independent duplication of experiments replaces examination of a colleague's paper.*
>
> *"Why should the non-mathematician care about things of this nature? The foremost reason is that mathematics is beautiful, even if it is, sadly, more inaccessible than other forms of art. The second is that it is useful, and that its utility depends in part on its certainty, and that that certainty cannot come without a notion of proof. Dr Gonthier, for instance, and his sponsors at Microsoft, hope that the techniques he and his colleagues have developed to formally prove mathematical theorems can be used to "prove" that a computer program is free of bugs-and that would certainly be a useful proposition in today's software society if it does, indeed, turn out to be true."*

3 Visual Computing

In a similar light *Visual Computing* is in its infancy and we can only imagine its future, but already there are many interesting harbingers both theoretical and applied ...

3.1 The Perko Pair

In many knot compendia produced since the 19th century, the first two knots in Figure 8.1 below have been listed as distinct ten crossing knots. They were however shown to be the same by Ken Perko in 1974. This is

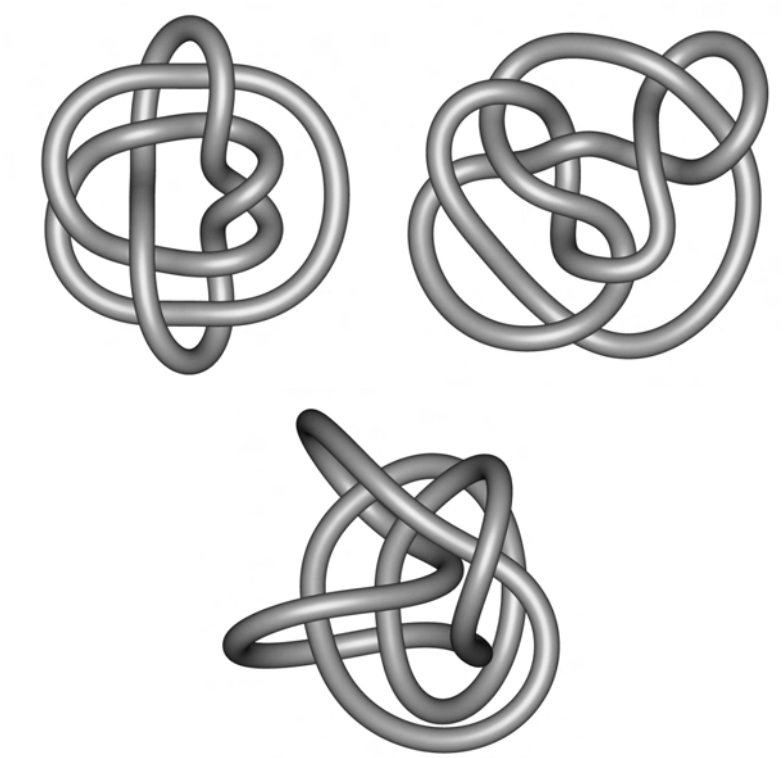


Figure 8.1. The Perko Pair of Knots

best illustrated dynamically in a program like *Knotplot* which will diffeomorphically adjust both into the third knot.

3.2 Fractal Cards and Chaos Games

Deterministic Constructions Not all impressive discoveries require a computer. Elaine Simmt and Brent Davis describe lovely constructions made by repeated regular paper folding and cutting—but no removal of paper—that result in beautiful fractal, self-similar, "pop-up" cards[5].

Nonetheless, in Figures 8.2 and 8.3 we choose to show various iterates of a pop-up Sierpinski triangle built in software, on turning those paper cutting and folding rules into an algorithm given in [44, pp 94–95]. This should be enough to let one start folding.

[5] "Fractal Cards: A Space for Exploration in Geometry and Discrete Maths,",*Math Teacher*, **91** (1998), 102–108.

Figure 8.2. The 1st and 2nd iterates of a Sierpinski card

Figure 8.3. The 3rd and 7th iterates of a Sierpinski card

Note the similarity to the Pascal triangles given in Figure 8.4. This is clarified when we note the pattern modulo two:

1
1 1
1 **2** 1

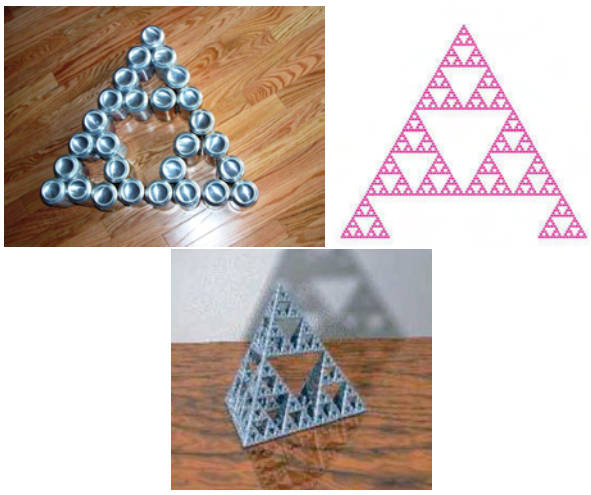


Figure 8.4. Pascal's Triangle modulo two above a Sierpinski construction

Figure 8.5. Self-similarity in Granada and Chartres

1 3 3 1
1 **4 6 4** 1
1 5 **10 10** 5 1
16 15 **20** 15 **6** 1
1 7 21 35 35 21 7 1

...

with the even numbers emboldened.

One may consider more complex modular patterns as discussed in Andrew Granville's online paper on binomial coefficients at `www.cecm.sfu.ca/organics/papers/granville/support/pascalform.html`. Likewise, we draw the reader's attention to the recursive structures elucidated in our chapter on strange functions. And always art can be an additional source of mathematical inspiration and stimulation as in the rose window from Chartres and the view of the Alhambra shown in Figure 8.5.

Random Constructions These prior constructions are all deterministic, and so relate closely to the ideas about cellular automata discussed at length in Stephen Wolfram's *A New Kind of Science*. But, as we shall see random constructions lead naturally to similar self-replicative structures.

In [44, §2.4] we described how to experimentally establish that we do indeed obtain a Sierpinski triangle from Pascal's triangle and noted also that simple random constructions led to the Sierpinski triangle or *gasket*: construct an "arbitrary" triangle, with vertices $(x_1, y_1), (x_2, y_2)$ and (x_3, y_3). Specify a point (x, y) within the triangle. Then iterate indefinitely the following construction: First select a random integer r in the set $(1, 2, 3)$, and

then construct a new point (x', y') as follows:

$$(x', y') = \left(\frac{1}{2}(x + x_r), \frac{1}{2}(y + y_r)\right). \tag{8.1}$$

The corresponding graph will yield a Sierpinski gasket asymptotically, as we see below. Even more intriguing is the following link with genetic processes.

4 A Preliminary Example: Visualizing DNA Strands

Suppose that we have a strand of DNA, that is, a string whose elements may be any of four symbols (bases): A (adenosine), C (cytosine), G (guanine), or T (thymine).

$$\text{ATGGACTTCCAG} \tag{8.2}$$

Each DNA sequence can be identified with a unique sub-square in the unit square $[0, 1] \times [0, 1]$ as follows. Identify each base, A, C, G, and T, with a vertex of the square (see Figure 8.6). Let x_n denote the sub-square corresponding to the subsequence of length n, and let ω_n be the n^{th} base. With each iteration, we will subdivide the square into four smaller squares. The initial region x_0 is the entire unit square. Given x_{n-1}, calculate the next region x_n as follows.

1. Find the n^{th} base ω_n in the sequence.

2. Subdivide the square x_{n-1} into four smaller squares. Set x_n to be the sub-square of x_{n-1} closest to the vertex ω_n. That is, if $\omega_n = \text{A}$, then x_n is the upper left sub-square of x_{n-1}; if $\omega_n = \text{G}$, then x_n is the lower right sub-square of x_{n-1}.

3. Plot x_n as follows. Replace x_{n-1} with the square x_{n-1} subdivided into four smaller squares, and shade in the sub-square x_n.

The first few sub-squares x_n, corresponding to the subsequences of length n of the sequence in (8.2), are shown in Figure 8.6. Note that different sequences yield distinct sub-squares in $[0, 1] \times [0, 1]$, with a one-to-one correspondence between dyadic sub-squares of the unit square and finite sequences of DNA.

Here we iterate a simple algorithm that calculates a new x_n from the previous x_{n-1}, plus an additional piece of information (ω_n), and displays the x_n graphically. As we describe below, this is the main idea behind the *chaos game*.

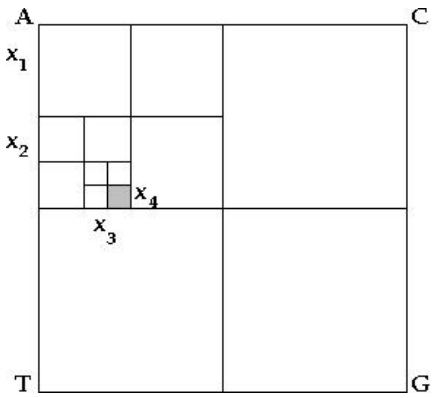


Figure 8.6. Visualizing DNA sequences of length n

5 What Is a Chaos Game?

A *chaos game* is a probabilistic algorithm for visualizing the *attractor* of an *iterated function system*. Chaos games can be used to generate fractals, since in many cases the attractor of an iterated function system is a fractal.

What does this mean? Suppose that we are working in a metric space X, say $X = \mathbf{R}^2$, and we have a finite collection $\{w_i\}_{i=1}^N$ of contractive maps (possibly composed with translations). The pair $(X, \{w_i\}_{i=1}^N)$ is called an *iterated function system* (IFS). A basic proposition in the study of dynamical systems states that there exists a unique compact set $A \subset X$ such that

$$A = \bigcup_{i=1}^{N} w_i(A),$$

and for every nonempty compact set $S \subset X$, the Hausdorff distance

$$h(\bigcup_{i=1}^{N} w_i^n(S), A) \to 0$$

as $n \to \infty$ (that is, $\cup_{i=1}^N w_i^n(S)$ converges to A or to a subset of A). The set A is called the *attractor* of the iterated function system $(X, \{w_i\}_{i=1}^N)$. If $X = \mathbf{R}$ or $X = \mathbf{R}^2$, we can choose some compact set S and graph the sets $\cup_{i=1}^N w_i^n(S)$ for successively larger and larger n to visualize successively better approximations to the attractor A. The chaos game algorithm gives a different method for visualizing the attractor.

To set up a chaos game, we first must associate a probability p_i with each of the contractive mappings w_i in our iterated function system. That is, $0 \le p_i \le 1$ for each i and $\sum_{i=1}^N p_i = 1$. In each iteration of the chaos

game algorithm, we will select one of the maps w_i randomly, with p_i being the probability of selecting map w_i. The triple $(X, \{w_i\}_{i=1}^N, \{p_i\}_{i=1}^N)$ is called an *iterated function system with probabilities* (IFSP). We initialize the chaos game by selecting a point $x_0 \in X$ to be a fixed point of one of the maps, say w_1. We begin by plotting x_0. We then iterate the following steps:

1. Select a map w_{i_n} at random according to the probabilities p_i.
2. Set $x_n = w_{i_n}(x_{n-1})$.
3. Plot x_n.

When using this algorithm to visualize an attractor of an IFSP, we will stop after a pre-set maximum number of iterations. For example, a pixel on a computer screen is a fixed width, ℓ. Since the maps w_i are contractive, eventually the distance between successive points x_k and x_{k+1} will be less than ℓ, and we will not be able to plot any new points on our image of the attractor. We can calculate or estimate the number of iterations M such that, no matter what trajectory we have taken (that is, no matter which maps w_i were chosen in each iteration), the distance between x_M and x_{M+1} will be less than ℓ, and stop after M iterations.

Since the point x_0 is initialized to be in the attractor, all subsequent points x_n will also be in the attractor. Running the chaos game algorithm for a small number of iterations will give only the barest outline of the shape of the attractor. Eventually we fill in more detail, however. We can also start over after a finite number of iterations, re-initialize the chaos game, and plot additional points x'_n along a (probably) different trajectory, to add more detail to our plot of points in the attractor.

More information about chaos games and iterated function systems can be found in the papers by Mendivil and Silver [**?**] and by Ashlock and Golden [**?**], discussed further below.

5.1 Examples of Chaos Games

Sierpinski Triangle The Sierpinski Triangle can be generated from an iterated function system with three maps:

$$w_1 : x \mapsto \frac{x + v_1}{2}$$
$$w_2 : x \mapsto \frac{x + v_2}{2}$$
$$w_3 : x \mapsto \frac{x + v_3}{2}$$

where x is a point in $\mathbf{R}^2$, v_1 is the point $(0, 1)$, v_2 is the point $(-1, 0)$, and v_3 is the point $(0, \sqrt{3})$ (these are three vertices of an equilateral triangle in

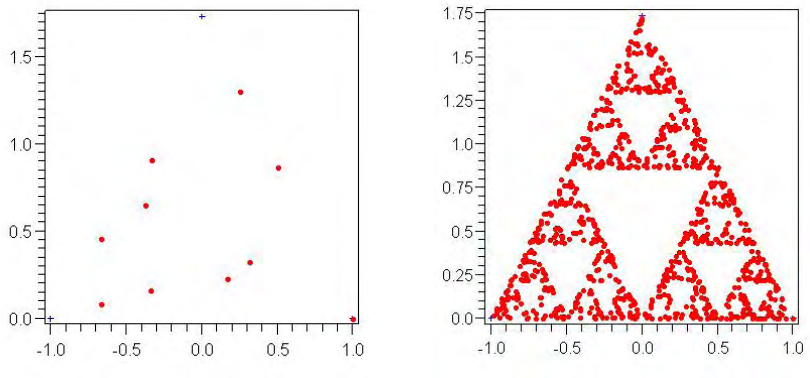


Figure 8.7. Chaos game for the Sierpinski Triangle after 10 and 1000 iterations

$\mathbf{R}^2$). Each map w_i maps the current point x to a point halfway in between x and the vertex v_i. We will choose the maps w_i according to a uniform probability distribution, with $p_i = \frac{1}{3}$ for each i, to make an IFSP out of this iterated function system.

For each i, the point v_i is a fixed point of the map w_i (eg. $w_1(v_1) = \frac{v_1+v_1}{2} = v_1$), so we may initialize the chaos game by setting $x_0 := v_1$ and plotting x_0. Several iterations of the chaos game are shown in Figure 8.7.

Twin Dragon Let $d_0 = (0,0)$ and $d_1 = (1,0)$ be points in $\mathbf{R}^2$. Let M be the twin dragon matrix

$$M = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Then the maps

$$w_1 : x \mapsto M^{-1}(x + d_0)$$
$$w_2 : x \mapsto M^{-1}(x + d_1)$$

form an iterated function system. We can assign the probabilities $p_1 = p_2 = \frac{1}{2}$ to the maps w_1 and w_2 to make an IFSP. Note that the origin $d_0 = (0,0)$ is a fixed point of the map w_1, thus we can initialize $x_0 := d_0 = (0,0)$ in the chaos game, and plot x_0. Several iterations are shown in Figure 8.8.

Scaling Functions and Wavelets We follow the notation of Mendivil and Silver [**?**] in this section, and present a simpler chaos game algorithm here than is given in [**?**]. While Mendivil and Silver's algorithm can additionally be used in wavelet analysis and synthesis, our simplified algorithm suffices to visualize the scaling function and wavelet. A scaling function $\phi(x)$ for a

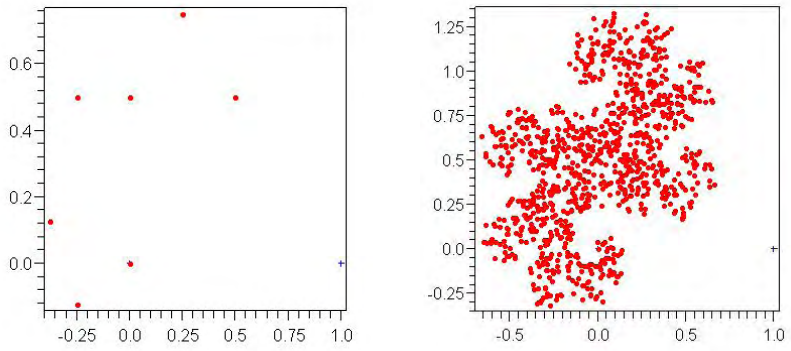


Figure 8.8. Chaos game for the Twin Dragon after 10 **and** 1000 **iterations**

multiresolution analysis satisfies a *scaling equation*

$$\phi(x) = \sum_{n \in \mathbf{Z}} h_n \phi(2x - n),$$

where the coefficients h_n are real or complex numbers. Suppose that the scaling function $\phi(x)$ is compactly supported. Then only finitely many of the coefficients, say $h_0, h_1, \ldots, h_N$, are nonzero, and $\phi(x)$ is supported on the interval $[0, N]$. In this case, a wavelet $\psi(x)$ associated with the multiresolution analysis can be generated by the formula

$$\psi(x) = \sum_{k \in \mathbf{Z}} (-1)^k h_{N-k} \phi(2x - k). \tag{8.3}$$

The wavelet $\psi(x)$ is also supported on the interval $[0, N]$. Both $\phi(x)$ and $\psi(x)$ depend only on the sequence of coefficients $h_0, h_1, \ldots, h_N$. Compactly supported scaling functions and wavelets can thus be generated by starting from a suitable sequence of coefficients. In fact, the first example of a compactly supported wavelet was constructed by Daubechies in just this manner. See [**?**] for more information about this construction, or about wavelets and multiresolution analysis in general.

These scaling functions and wavelets may not have a closed form representation. In practice, this is not a problem, since it is the sequence of coefficients that are used in applications, rather than the scaling function or wavelet themselves. Indeed, the Fourier transform of the sequence of coefficients

$$m(\xi) = \sum_n h_n e^{-2\pi i n x i}$$

is called the *low-pass filter* for the scaling function $\phi(x)$, and is an important tool in studying multiresolution analyses and wavelets.We may wish to visualize the scaling function $\phi(x)$ or wavelet $\psi(x)$, however. This can be done as follows using a chaos game.

We first vectorize the scaling equation, as follows. Define $V_\phi : [0,1] \to \mathbf{R}$ by

$$V_\phi(x) = \begin{bmatrix} \phi(x) \\ \phi(x+1) \\ \vdots \\ \phi(x+N-1) \end{bmatrix}.$$

Define matrices T_0 and T_1 by

$$(T_0)_{i,j} = h_{2i-j-1}$$
$$(T_1)_{i,j} = h_{2i-j}.$$

Let $\tau : [0,1] \to [0,1]$ be the mapping $\tau(x) = 2x \pmod 1$. Then the scaling equation can be written

$$V_\phi(x) = T_\omega V_\phi(\tau x), \tag{8.4}$$

where ω is the first digit of the binary expansion of x. That is, $2x = \omega + \tau x$.

The chaos game in this example will update the values of both x and $V_\phi(x)$ in each iteration. To update x, we will use the mappings $w_0, w_1 : [0,1] \to [0,1]$ defined by

$$w_0(x) = \frac{x}{2}$$
$$w_1(x) = \frac{x}{2} + \frac{1}{2}.$$

To update $V_\phi(x)$, we will use the mappings T_0 and T_1 defined above. To initialize the chaos game, we set $x_0 = 0$, and set $V_{\phi,0}$ to be a fixed point of the map T_0 (that is, $V_{\phi,0}$ is a specific vector

$$V_{\phi,0} = \begin{bmatrix} v_{0,0} \\ v_{1,0} \\ \vdots \\ v_{N,0} \end{bmatrix},$$

where we will set $\phi(x_0) := v_{0,0}$, $\phi(x_0+1) := v_{1,0}$, and so on up to $\phi(x_0 + N - 1) := v_{N,0}$).

For each iteration of the chaos game, we choose $\alpha_n = 0, 1$ uniformly, and update

$$x_n = w_{\alpha_n}(x_{n-1})$$
$$V_{\phi,n} = \begin{bmatrix} v_{0,n} \\ v_{1,n} \\ \vdots \\ v_{N,n} \end{bmatrix} = T_{\alpha_n} V_{\phi,n-1}.$$

We set $V_\phi(x_n)$ to be equal to $V_{\phi,n}$, so that $\phi(x_n) = v_{0,n}$, $\phi(x_n + 1) = v_{1,n}$, and so on to $\phi(x_n + N - 1) = v_{N,n}$. Then we plot the points $(x_n, \phi(x_n))$, $(x_n + 1, \phi(x_n + 1))$, through $(x_n + N - 1, \phi(x_n + N - 1))$.

The wavelet $\psi(x)$ can be visualized similarly. We may vectorize the equation 8.3 by forming the vector

$$V_\psi(x) = \begin{bmatrix} \psi(x) \\ \psi(x+1) \\ \vdots \\ \psi(x+N-1) \end{bmatrix}$$

and defining mappings H_0 and H_1 by

$$(H_0)_{i,j} = (-1)^k h_{N-k},\ k = 2i - j - 1$$
$$(H_1)_{i,j} = (-1)^{k'} h_{N-k'},\ k' = 2i - j.$$

We then use V_ψ, H_0, and H_1 in place of V_ϕ, T_0, and T_1, respectively, in the chaos game described above.

Several iterations of this algorithm for the Daubechies' D_3 wavelet, one of the first examples of a compactly supported wavelet, are shown in Figure 8.9. The Daubechies' D_3 wavelet is supported on the interval $[0, 3]$, and has coefficients $h_0 = \frac{1+\sqrt{3}}{4}$, $h_1 = \frac{3+\sqrt{3}}{4}$, $h_2 = \frac{3-\sqrt{3}}{4}$, and $h_3 = \frac{1-\sqrt{3}}{4}$ [**?**].

5.2 More on Visualizing DNA

In the opening example, we used a deterministic sequence of DNA bases to update a point x_n in the attractor of an iterated function system, rather than randomly choosing which of the four vertices of the square to average the previous point x_{n-1} with. If we instead choose the base ω_n randomly, then we have a true chaos game.

Let $p_{\beta,k}(seq)$ be the probability that base β (β = A, C, G, or T) follows the subsequence *seq* of length k. Thus $p_{\mathrm{A},1}(\mathrm{T})$ is the probability that x_n = A given that x_{n-1} = T, while $p_{\mathrm{A},2}(\mathrm{GG})$ is the probability that x_n = A given that x_{n-1} = G and that x_{n-2} = G. If the n^{th} base is selected

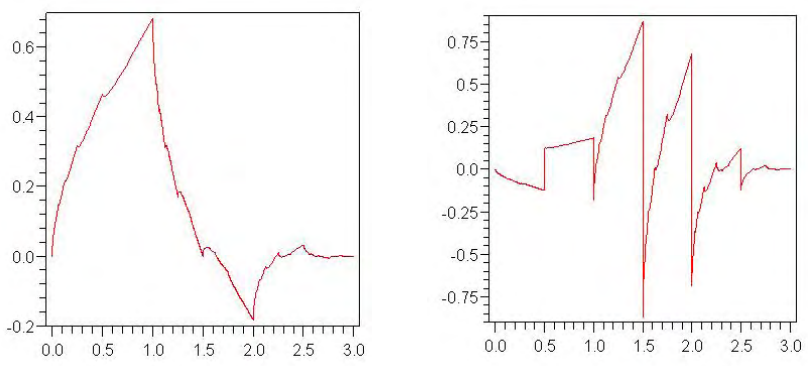


Figure 8.9. Chaos game for the Daubechies' D_3 wavelet (L) and scaling function (R) after 1000 iterations

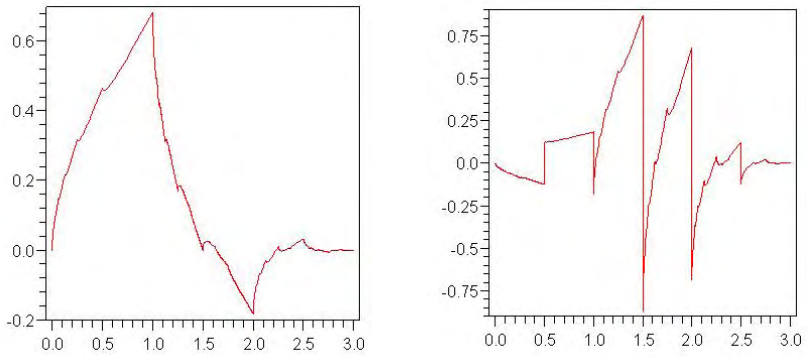


Figure 8.10. Chaos games for two different bacteria

uniformly at random, with no dependence on the preceding sequence, then we do not get a very interesting picture out of the chaos game. However, in most organisms, these probabilities are not uniformly distributed. As well, the probabilities $p_{\beta,k}(seq)$ vary by organism. Ashlock and Golden [**?**] have shown that different organisms yield distinct chaos games (see Figure 8.10).

6 Hilbert's Inequality and Witten's Zeta-function

We next explore a variety of pleasing connections between analysis, number theory and operator theory, while exposing a number of beautiful inequalities originating with Hilbert. We shall first establish the afore-mentioned

inequality [133, 215] and then apply it to various multiple zeta values. In consequence we obtain the norm of Hilbert's matrix.

6.1 Hilbert's (easier) Inequality

A useful preparatory lemma is

Lemma 6.1. *For* $0 < a < 1$ *and* $n = 1, 2, \ldots$

$$\sum_{m=1}^{\infty} \frac{1}{(n+m)(m/n)^a} < \int_0^{\infty} \frac{1}{(1+x)\,x^a}\,dx < \frac{(1/n)^{1-a}}{1-a} + \sum_{m=1}^{\infty} \frac{1}{(n+m)(m/n)^a},$$

and

$$\int_0^{\infty} \frac{1}{(1+x)\,x^a}\,dx = \pi \csc(a\,\pi).$$

Proof. The inequalities comes from using standard rectangular approximations to a monotonic integrand and overestimating the integral from 0 to $1/n$.

The evaluation of the integral is in various tables and is known to *Maple* or *Mathematica*. We offer two other proofs: (i) write

$$\begin{aligned}
\int_0^{\infty} \frac{1}{(1+x)\,x^a}\,dx &= \int_0^1 \frac{x^{-a} + x^{a-1}}{1+x}\,dx \\
&= \sum_{n=0}^{\infty} (-1)^n \left\{ \frac{1}{n+1-a} + \frac{1}{n+a} \right\} \\
&= \sum_{n=1}^{\infty} (-1)^n \left\{ \frac{1}{n+a} - \frac{1}{n-a} \right\} + \frac{1}{a} \\
&= \frac{1}{a} + \sum_{n=1}^{\infty} \frac{(-1)^n\,2a}{a^2 - n^2} = \pi \csc(a\,\pi),
\end{aligned}$$

since the last equality is the classical partial fraction identity for $\pi \csc(a\,\pi)$.

(ii) Alternatively, we begin by letting $1 + x = 1/y$,

$$\begin{aligned}
\int_0^{\infty} \frac{x^{-a}}{1+x}\,dx &= \int_0^1 y^{a-1}\,(1-y)^{-a}\,dy \\
&= B(a, 1-a)\Gamma(a)\,\Gamma(1-a) = \frac{\pi}{\sin(a\,\pi)}.
\end{aligned}$$

□

Combining the arguments in (i) and (ii) above actually derives the identity

$$\Gamma(a)\,\Gamma(1-a) = \frac{\pi}{\sin(a\,\pi)},$$

from the partial fraction for cosec or vice versa—especially if we appeal to the Bohr-Mollerup theorem to establish $B(a, 1-a) = \Gamma(a)\,\Gamma(1-a)$.

Theorem 6.2. (**Hilbert Inequality.**) *For non-negative sequences* (a_n) *and* (b_n)*, not both zero, and for* $1 \le p, q \le \infty$ *with* $1/p + 1/q = 1$ *one has*

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_n\,b_m}{n+m} < \pi\csc\left(\frac{\pi}{p}\right)\|a_n\|_p\,\|b_n\|_q. \tag{8.5}$$

Proof. Fix $\lambda > 0$. We apply Hölder's inequality with 'compensating difficulties' to obtain

$$\begin{aligned}
\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_n\,b_m}{n+m} &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_n}{(n+m)^{1/p}(m/n)^{\lambda/p}}\,\frac{b_m}{(n+m)^{1/q}(n/m)^{\lambda/p}} \qquad (8.6)\\
\le\ &\left(\sum_{n=1}^{\infty}|a_n|^p\sum_{m=1}^{\infty}\frac{1}{(n+m)(m/n)^{\lambda}}\right)^{1/p}\left(\sum_{m=1}^{\infty}|b_m|^q\sum_{n=1}^{\infty}\frac{1}{(n+m)(n/m)^{\lambda q/p}}\right)^{1/q}\\
<\ &\pi\,\left|\csc\left(\pi\,\lambda\right)\right|^{1/p}\,\left|\csc\left((q-1)\pi\,\lambda\right)\right|^{1/q}\,\|a_n\|_p\,\|b_m\|_q,
\end{aligned}$$

so that the left hand side of (8.5) is no greater than $\pi\csc\left(\frac{\pi}{p}\right)\|a_n\|_p\,\|b_n\|_q$ on setting $\lambda = 1/q$ and appealing to symmetry in p, q. □

The integral analogue of (6.2) may likewise be established. There are numerous extensions. One of interest for us later is

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_n\,b_m}{(n+m)^{\tau}} < \left\{\pi\csc\left(\frac{\pi(q-1)}{\tau q}\right)\right\}^{\tau}\|a_n\|_p\,\|b_n\|_q, \tag{8.7}$$

valid for $p, q > 1, \tau > 0, 1/p + 1/q \ge 1$ and $\tau + 1/p + 1/q = 2$. The best constant $C(p, q, \tau) \le \{\pi\csc(\pi(q-1)/(\tau q))\}^{\tau}$ is called the *Hilbert constant* [114, §3.4].

For $p = 2$, (6.2) becomes Hilbert's original inequality:

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_n\,b_m}{n+m} \le \pi\sqrt{\sum_{n=1}^{\infty}|a_n|^2}\sqrt{\sum_{n=1}^{\infty}|b_n|^2}, \tag{8.8}$$

though Hilbert only obtained the constant 2π, [132].

A fine direct Fourier analytic proof starts from the observation that

$$\frac{1}{2\pi\,i}\int_0^{2\,\pi}(\pi - t)\,e^{int}\,dt = \frac{1}{n}$$

for $n = 1, 2 \ldots$, and deduces

$$\sum_{n=1}^{N} \sum_{m=1}^{N} \frac{a_n\, a_m}{n+m} \frac{1}{2\pi\, i} \int_0^{2\,\pi} (\pi - t) \sum_{k=1}^{N} a_k\, e^{ikt} \sum_{k=1}^{N} b_k\, e^{ikt}\, dt. \tag{8.9}$$

We recover (8.8) by applying the integral form of the Cauchy-Schwarz inequality to the integral side of the representation (8.9).

Exercise 6.1. *Likewise*

$$\sum_{n=1}^{N} \sum_{m=1}^{N} \frac{a_n\, a_m}{(n+m)^2} \frac{1}{2\pi} \int_0^{2\pi} \left(\zeta(2) - \frac{\pi t}{2} + \frac{1}{4} \right) \sum_{k=1}^{N} a_k\, e^{ikt} \sum_{k=1}^{N} b_k\, e^{ikt}\, dt,$$

and more generally

$$\sum_{n=1}^{N} \sum_{m=1}^{N} \frac{a_n\, a_m}{(n+m)^\sigma} \frac{1}{2\pi\, i^\sigma} \int_0^{2\pi} \psi_\sigma \left(\frac{t}{2\pi} \right) \sum_{k=1}^{N} a_k\, e^{ikt} \sum_{k=1}^{N} b_k\, e^{ikt}\, dt$$

since

$$\psi_{2n}(x) = \sum_{k=1}^{\infty} \frac{\cos(2k\pi x)}{k^{2n}}, \qquad \psi_{2n+1}(x) = \sum_{k=1}^{\infty} \frac{\sin(2k\pi x)}{k^{2n+1}}$$

where $\psi_\sigma(x)$ are related to the *Bernoulli polynomials* by

$$\psi_\sigma(x) = (-1)^{\lfloor (1+\sigma)/2 \rfloor}\, B_\sigma(x)\, \frac{(2\pi)^\sigma}{2\,\sigma!},$$

for $0 < x < 1$. It follows that

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{a_n\, a_m}{(n+m)^\sigma} \le \|\psi_\sigma\|_{[0,1]}\, \|a\|_2\, \|b\|_2,$$

where for $n > 0$ we compute

$$\|\psi_{2n}\|_{[0,1]} = \psi_{2n}(0) = \zeta(2n) \qquad \text{and} \qquad \|\psi_{2n+1}\|_{[0,1]} = \psi_{2n+1}(1/4) = \beta(2n{+}1).$$

This and much more of the early 20th century history—and philosophy—of the *" 'bright' and amusing"* subject of inequalities charmingly discussed in Hardy's retirement lecture as London Mathematical Society Secretary, [132]. He comments [132, p. 474] that *Harald Bohr is reported to have remarked "Most analysts spend half their time hunting through the literature for inequalities they want to use, but cannot prove."*

This remains true, though more recent inequalities often involve less linear objects such as entropies, divergences, and log-barrier functions [44, 63] such as the *divergence estimate* [61, p. 63]:

$$\sum_{n-1}^{N} p_i \log\left(\frac{p_i}{q_i}\right) \geq \frac{1}{2}\left(\sum_{n=1}^{N} |p_i - q_i|\right)^2,$$

valid for any two strictly positive sequences with $\sum_{i=1}^{N} p_i = \sum_{i=1}^{N} q_i = 1$.

Two other high-spots in Hardy's essay are *Carleman's inequality*

$$\sum_{n=1}^{\infty} (a_1\, a_2 + \ldots + a_n)^{1/n} \leq e \sum_{n=1}^{\infty} a_n,$$

see also [**?**, p. 284], and Hardy's own

$$\sum_{n=1}^{\infty} \left(\frac{a_1 + a_2 + \cdots + a_n}{n}\right)^p \leq \left(\frac{p}{p-1}\right)^p \sum_{n=1}^{\infty} a_n^p, \tag{8.10}$$

for $p > 1$. Hardy comments [132, p. 485] that his *"own theorem was discovered as a by-product of my own attempt to find a really simple and elementary proof of Hilbert's."* For $p = 2$, Hardy reproduces Elliott's proof of (8.10), writing "*it can hardly be possible to find a proof more concise or elegant*". It is as follows. Set $A_n := a_1 + a_2 + \cdots + a_n$ and write

$$\frac{2a_n\, A_n}{n} - \left(\frac{A_n}{n}\right)^2 = \frac{A_n^2}{n} - \frac{A_{n-1}^2}{n-1} + (n-1)\left(\frac{A_n}{n} - \frac{A_{n-1}}{n-1}\right)^2 \geq \frac{A_n^2}{n} - \frac{A_{n-1}^2}{n-1}$$

something easy to check symbolically, and sum to obtain

$$\sum_n \left(\frac{A_n}{n}\right)^2 \leq 2\sum_n \frac{a_n\, A_n}{n} \leq 2\sqrt{\sum_n a_n^2}\sqrt{\sum_n \left(\frac{A_n}{n}\right)^2},$$

which proves (8.10) for $p = 2$; and this easily adapts to the general case.

Finally we record the (harder) *Hilbert inequality* is

$$\left|\sum_{n\neq m\in\mathbf{Z}} \frac{a_n\, b_m}{n-m}\right| < \pi\sqrt{\sum_{n=1}^{\infty} |a_n|^2}\sqrt{\sum_{n=1}^{\infty} |b_n|^2}, \tag{8.11}$$

the best constant π being due to Schur in (1911), [180]. There are many extensions—with applications to prime number theory, [180].

6.2 Witten ζ-functions

Let us recall that initially for $r, s > 1/2$:

$$\mathcal{W}(r,s,t) := \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{1}{n^r\, m^s\, (n+m)^t}$$

is a *Witten ζ-function*, [234, **?**, 88]. We refer to [234] for a description of the uses of more general Witten ζ-functions. Ours are also—more accurately—called *Tornheim double sums*, [**?**]. Correspondingly

$$\zeta(t,s) := \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{1}{m^s\, (n+m)^t} = \sum_{n>m>0} \frac{1}{n^t\, m^s}$$

is an *Euler double sum* of the sort met in Chapters 1 and 2.

There is a simple algebraic relation

$$\mathcal{W}(r,s,t) = \mathcal{W}(r-1,s,t+1) + \mathcal{W}(r,s-1,t+1). \tag{8.12}$$

This is based on writing

$$\frac{m+n}{(m+n)^{t+1}} = \frac{m}{(m+n)^{t+1}} + \frac{n}{(m+n)^{t+1}}.$$

Also

$$\mathcal{W}(r,s,t)\mathcal{W}(s,r,t), \tag{8.13}$$

and

$$\mathcal{W}(r,s,0) = \zeta(r)\,\zeta(s) \quad \text{while} \quad \mathcal{W}(r,0,t) = \zeta(r,t). \tag{8.14}$$

Hence, $\mathcal{W}(s,s,t) = 2\,\mathcal{W}(s,s-1,t+1)$ and so

$$\mathcal{W}(1,1,1) = 2\,\mathcal{W}(1,0,2) = 2\,\zeta(2,1) = 2\,\zeta(3).$$

In particular, (8.15) implies that $\zeta(3) \leq \pi^3/3$, on appealing to (9.9) below. For many proofs of this basic identity $\zeta(2,1) = \zeta(3)$ we refer to [**?**]. We note that the analogue to (8.12), $\zeta(s,t)+\zeta(t,s) = \zeta(s)\,\zeta(t)-\zeta(s+t)$, shows that $\mathcal{W}(s,0,s) = 2\,\zeta(s,s) = \zeta^2(s)-\zeta(2s)$. In particular, $\mathcal{W}(2,0,2) = 2\,\zeta(2,2) = \pi^4/36 - \pi^4/90 = \pi^4/72$.

Exercise 6.2. *Let $a_n := 1/n^r$, $b_n := 1/n^s$. Then the inequality (8.8) becomes*

$$\mathcal{W}(r,s,1) \leq \pi\,\sqrt{\zeta(2r)}\,\sqrt{\zeta(2s)}. \tag{8.15}$$

Similarly, the inequality (8.5) becomes

$$\mathcal{W}(r,s,1) \le \pi \csc\left(\frac{\pi}{p}\right) \sqrt[p]{\zeta(pr)}\ \sqrt[q]{\zeta(qs)}. \tag{8.16}$$

Indeed, (8.7) can be used to estimate $\mathcal{W}(r,s,\tau)$ for somewhat broader $\tau \neq 1$.

More generally, recursive use of (8.12) and (8.13), along with initial conditions (8.14) shows that *all integer $\mathcal{W}(s,r,t)$ values are expressible in terms of double (and single) Euler sums.* As we shall see in (8.21) the representations are necessarily homogeneous polynomials of *weight* $r+s+t$. All double sums of weight less than 8 and all those of odd weight reduce to sums of products of single variable zeta values, [**?**]. The first impediments are because $\zeta(6,2), \zeta(5,3)$ are not reducible.

We also observe that in terms of the *polylogarithm* $Li_s(t) := \sum_{n>0} t^n/n^s$ for real s, the representation (8.9) yields

$$\mathcal{W}(r,s,1)\frac{1}{2\pi\, i}\ \int_{-\pi}^{\pi} \sigma\, \mathrm{Li}_r(-e^{i\sigma})\, \mathrm{Li}_s(-e^{i\sigma})\, d\sigma. \tag{8.17}$$

This representation is not numerically effective. It is better to start with $\Gamma(s)/(m+n)^t = \int_0^1 (-\log\sigma)^{t-1}\, \sigma^{m+n-1}\, d\sigma$ and so to obtain

$$\mathcal{W}(r,s,t)\frac{1}{\Gamma(t)}\ \int_0^1 \mathrm{Li}_r(\sigma)\, \mathrm{Li}_s(\sigma)\, \frac{(-\log\sigma)^{t-1}}{\sigma}\, d\sigma. \tag{8.18}$$

This real variable analogue of (8.17) is much more satisfactory computationally. For example, we recover an analytic proof of

$$2\,\zeta(2,1) = \mathcal{W}(1,1,1) \int_0^1 \frac{\ln^2(1-\sigma)}{\sigma}\, d\sigma = 2\,\zeta(3). \tag{8.19}$$

Moreover, we may now discover analytic as opposed to algebraic relations. Integration by parts yields

$$\mathcal{W}(r,s+1,1) + \mathcal{W}(r+1,s,1) = \mathrm{Li}_{r+1}(1)\, \mathrm{Li}_{s+1}(1) = \zeta(r+1)\,\zeta(s+1). \tag{8.20}$$

So, in particular, $\mathcal{W}(s+1,s,1) = \zeta^2(s+1)/2$.

Symbolically, *Maple* immediately evaluates $\mathcal{W}(2,1,1) = \pi^4/72$, and while it fails directly with $\mathcal{W}(1,1,2)$, we know it must be a multiple of π^4 or equivalently $\zeta(4)$; and numerically obtain $\mathcal{W}(1,1,2)/\zeta(4) = .4999999999999999998\ldots$.

Continuing, for $r+s+t = 5$ the only terms to consider are $\zeta(5), \zeta(2)\zeta(3)$, and *PSLQ* yields the weight five relations:

$$\mathcal{W}(2,2,1) = \int_0^1 \frac{\mathrm{Li}_2\,(x)^2}{x}\, dx = 2\,\zeta\,(3)\,\zeta(2) - 3\,\zeta\,(5)\,,$$

$$\mathcal{W}(2,1,2) = \int_0^1 \frac{\mathrm{Li}_2\,(x)\log(1-x)\log(x)}{x}\,dx = \zeta\,(3)\,\zeta(2) - \frac{3}{2}\,\zeta\,(5)\,,$$

$$\mathcal{W}(1,1,3) = \int_0^1 \frac{\log^2(x)\,\log^2(1-x)}{2\,x}\,dx = -2\,\zeta\,(3)\,\zeta(2) + 4\,\zeta\,(5)\,,$$

$$\mathcal{W}(3,1,1) = \int_0^1 \frac{\mathrm{Li}_3\,(x)\log(1-x)}{x}\,dx - \zeta\,(3)\,\zeta(2) + 3\,\zeta\,(5)\,,$$

as predicted.

Likewise, for $r+s+t=6$ the only terms we need to consider are $\zeta(6), \zeta^2(3)$ since $\zeta(6), \zeta(4)\,\zeta(2)$ and $\zeta^3(2)$ are all rational multiples of π^6. We recover identities like

$$\mathcal{W}(3,2,1) = \int_0^1 \frac{\mathrm{Li}_3\,(x)\;\mathrm{Li}_2\,(x)}{x}\,dx = \frac{1}{2}\,\zeta^2\,(3)\,,$$

consistent with the equation below (8.20).

The general form of the reduction, for integer r, s and t, is due to Tornheim and expresses $\mathcal{W}(r,s,t)$ in terms of $\zeta(a,b)$ with weight $a+b = N := r+s+t$, [**?**]:

$$\mathcal{W}(r,s,t) = \sum_{i=1}^{r\vee s}\left\{\binom{r+s-i-1}{s-1} + \binom{r+s-i-1}{r-1}\right\}\zeta\,(i, N-i)\,. \quad (8.21)$$

Various other general formulas are given in [**?**] for classes of sums such as $\mathcal{W}(2n+1, 2n+1, 2n+1)$ and $\mathcal{W}(2n, 2n, 2n)$.

6.3 The Best Constant

It transpires that the constant π used in Theorem 6.2 is best possible, [133].

Exercise 6.3. *Let us numerically explore the ratio*

$$\mathcal{R}(s) := \frac{\mathcal{W}(s,s,1)}{\pi\,\zeta(2s)}$$

as $s \to 1/2^+$. *Note that* $\mathcal{R}(1) = 12\,\zeta(3)/\pi^3 \sim 0.4652181552\ldots$.

Initially, we may directly sum as follows:

$$\begin{aligned}
\mathcal{W}(s,s,1) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{m^{-s}n^{-s}}{m+n} &= 2\sum_{n=1}^{\infty}\frac{1}{n^{2s}}\sum_{m=1}^{n-1}\frac{1/n}{(m/n)^s(m/n+1)} + \frac{\zeta(2s+1)}{2} \\
&\le 2\,\zeta(2s)\int_0^1 \frac{x^{-s}}{1+x}\,dx + \frac{\zeta(2s+1)}{2} \\
&\le 2\sum_{n=1}^{\infty}\frac{1}{n^{2s}}\sum_{m=1}^{n}\frac{1/n}{(m/n)^s(m/n+1)} + \frac{\zeta(2s+1)}{2} \\
&= 2\sum_{n=1}^{\infty}\frac{1}{n^{2s}}\sum_{m=1}^{n-1}\frac{1/n}{(m/n)^s(m/n+1)} + \frac{3\zeta(2s+1)}{2} \\
&= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{m^{-s}n^{-s}}{m+n} + \zeta(2s+1).
\end{aligned}$$

We deduce that $\mathcal{R}(s) \sim \mathcal{I}(s) := 2/\pi \int_0^1 x^{-s}/(1+x)\,dx$ as $s \to 1/2$.

Further numerical explorations seem in order. For $1/2 < s < 1$, (8.18) is hard to use numerically and led us to look for a more sophisticated attack along the line of the Hurwitz zeta and Bernoulli polynomial integrals used in [**?**], or more probably the expansions in [88]. Viz,

$$\mathcal{W}(r,s,t) = \int_0^1 E(r,x)\,E(s,x)\,\overline{E(t,x)}\,dx \tag{8.22}$$

where $E(s,x) := \sum_{n=1}^{\infty} e^{2\pi i n x}\,n^{-s} = \mathrm{Li}_s\left(e^{2\pi i x}\right)$, using the formulae

$$E(s,x) = \sum_{m=0}^{\infty}\zeta(s-m)\,\frac{(2\pi i\,x)^m}{m!} + \Gamma(1-s)\,(-2\pi i\,x)^{s-1},$$

for $|x| < 1$ and

$$E(s,x) = -\sum_{m=0}^{\infty}\eta(s-m)\,\frac{(2x-1)^m\,(\pi i)^m}{m!},$$

with $\eta(s) := (1-2^{1-s})\,\zeta(s)$, for $0 < x < 1$ as given in [88, (2.6) and (2.9)].

Indeed, carefully expanding (8.22) with a free parameter $\theta \in (0,1)$,

leads to the following efficient formula *when neither r nor s is an integer:*

$$\begin{aligned}\Gamma(t)\mathcal{W}(r,s,t) &= \sum_{m,n\geq 1}\frac{\Gamma(t,(m+n)\theta)}{m^r n^s (m+n)^t}\\ &+ \sum_{u,v\geq 0}(-1)^{u+v}\frac{\zeta(r-u)\zeta(s-v)\theta^{u+v+t}}{u!v!(u+v+t)}\\ &+ \Gamma(1-r)\sum_{v\geq 0}(-1)^v\frac{\zeta(s-v)\theta^{r+v+t-1}}{v!(r+v+t-1)}\\ &+ \Gamma(1-s)\sum_{u\geq 0}(-1)^u\frac{\zeta(r-u)\theta^{s+u+t-1}}{u!(s+u+t-1)}\\ &+ \Gamma(1-r)\Gamma(1-s)\frac{\theta^{r+s+t-2}}{r+s+t-2}. \qquad (8.23)\end{aligned}$$

When one or both of r, s is an integer, a limit formula with a few more terms results. We can now accurately plot $\mathcal{R}$ and $\mathcal{I}$ on $[1/3, 2/3]$, as shown in Figure 8.11, and so are lead to:

Conjecture. $\lim_{s\to 1/2}\mathcal{R}(s) = 1$.

Proof. To establish this, we denote $\sigma_n(s) := \sum_{m=1}^{\infty} n^s m^{-s}/(n+m)$ and appeal to Lemma 6.1 to write

$$\begin{aligned}\mathcal{L}: &= \lim_{s\to 1/2}(2s-1)\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{n^{-s}\,m^{-s}}{n+m} = \lim_{s\to 1/2}(2s-1)\sum_{n=1}^{\infty}\frac{1}{n^{2s}}\,\sigma_n(s)\\ &= \lim_{s\to 1/2}(2s-1)\sum_{n=1}^{\infty}\frac{\{\sigma_n(s)-\pi\csc(\pi\,s)\}}{n^{2s}} + \lim_{s\to 1/2}\pi\,(2s-1)\zeta(2s)\csc(\pi\,s)\\ &= 0+\pi,\end{aligned}$$

since, by another appeal to Lemma 6.1, the parenthetic term is $O(n^{s-1})$ while in the second $\zeta(2s) \sim 1/(2s-1)$ as $s \to 1/2^+$.

In consequence, we see that $\mathcal{L} = \lim_{s\to 1/2}\mathcal{R}(s) = 1$, and—at least to first-order—inequality (8.8) is best possible, see also [143]. □

Likewise, the constant in (6.2) is best possible. Motivated by above argument we consider

$$\mathcal{R}_p(s) := \frac{\mathcal{W}((p-1)s, s, 1)}{\pi\,\zeta(ps)},$$

and observe that with $\sigma_n^p(s) := \sum_{m=1}^{\infty}(n/m)^{-(p-1)s}/(n+m) \to \pi\csc\left(\frac{\pi}{q}\right)$,

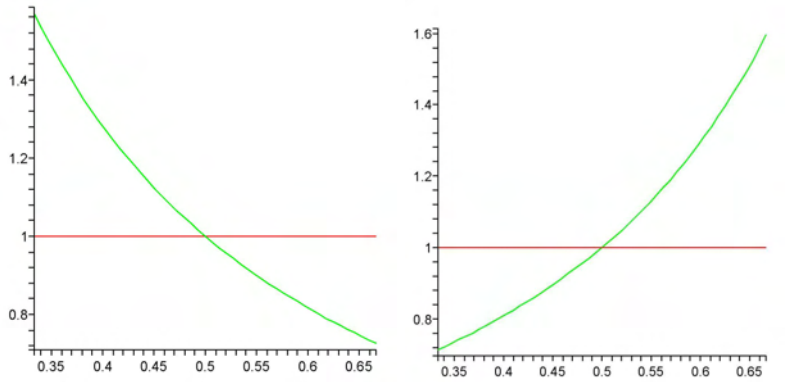


Figure 8.11. $\mathcal{R}$**(left) and and** $\mathcal{I}$ **(right) on** $[1/3, 2/3]$

we have

$$\begin{aligned}
\mathcal{L}_p : &= \lim_{s\to 1/p} (ps-1) \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{n^{-s}\, m^{-(p-1)s}}{n+m} = \lim_{s\to 1/p} (ps-1) \sum_{n=1}^{\infty} \frac{1}{n^{ps}}\, \sigma_n^p(s) \\
&= \lim_{s\to 1/p} (ps-1) \sum_{n=1}^{\infty} \frac{\{\sigma_n^p(s) - \pi \csc(\pi/q))\}}{n^{ps}} + \lim_{s\to 1/p} (2s-1)\zeta(ps)\, \pi \csc\left(\frac{\pi}{q}\right) \\
&= 0 + \pi \csc\left(\frac{\pi}{q}\right).
\end{aligned}$$

Setting $r := (p-1)s, s \to 1/p^+$ we check that $\zeta(ps)^{1/p}\, \zeta(qr)^{1/q} = \zeta(ps)$ and hence the best constant in (8.16) is the one given. To recapitulate in terms of the celebrated infinite *Hilbert matrices*, $\mathcal{H}_0 := \{1/(m+n)\}_{m,n=1}^{\infty}$, and $\mathcal{H}_1 := \{1/(m+n-1)\}_{m,n=1}^{\infty}$, [**?**, pp. 250–252], we have actually proven:

Theorem 6.3. *Let* $1 < p, q < \infty$ *be given with* $1/p + 1/q = 1$. *The Hilbert matrices* $\mathcal{H}_0$ *and* $\mathcal{H}_1$ *determine bounded linear mappings from the sequence space* ℓ^p *to itself such that*

$$\|\mathcal{H}_1\|_{p,p} = \|\mathcal{H}_0\|_{p,p} = \lim_{s\to 1/p} \frac{\mathcal{W}(s,(p-1)s,1)}{\zeta(ps)} = \pi \csc\left(\frac{\pi}{p}\right).$$

Proof. Appealing to the isometry between $(\ell^p)^*$ and ℓ^q, and given the evaluation $\mathcal{L}_p$ above, we directly compute the operator norm of $\mathcal{H}_0$ as

$$\|\mathcal{H}_0\|_{p,p} = \sup_{\|x\|_p=1} \|\mathcal{H}_0 x\|_p = \sup_{\|y\|_q=1} \sup_{\|x\|_p=1} \langle \mathcal{H}_0 x, y\rangle = \pi \csc\left(\frac{\pi}{p}\right).$$

Now clearly $\|\mathcal{H}_0\|_{p,p} \leq \|\mathcal{H}_1\|_{p,p}$. For $n \geq 2$ we have

$$\sum_{m=1}^{\infty} \frac{1}{(n+m-1)(m/n)^a} \leq \sum_{m=1}^{\infty} \frac{1}{(n-1+m)(m/(n-1))^a} \leq \pi \csc(\pi a),$$

an appeal to Lemma 6.1 and Theorem 6.2 shows $\|\mathcal{H}_0\|_{p,p} \geq \|\mathcal{H}_1\|_{p,p}$. □

A delightful operator-theoretic introduction to the Hilbert matrix $\mathcal{H}_0$ is given by Choi in his Chauvenet prize winning article [74] while a recent set of notes by G.J. O. Jameson, see [143], is also well worth accessing.

In the case of (8.7), Finch [114, §4.3] comments that the issue of best constants is unclear in the literature. He remarks that even the case $p = q = 4/3, \tau = 1/2$ appears to be open. It seems improbable that the techniques of this note can be used to resolve the question. Indeed, consider $\mathcal{R}_{1/2}(s, \alpha) : \mathcal{W}(s, s, 1/2)/\zeta(4s/3)^\alpha$, with the critical point in this case being $s = 3/4$.

Numerically, using (8.23) we discover that $\log(\mathcal{W}(s, s, 1/2))/\log(\zeta(4s/3)) \to 0$. Hence, for any $\alpha > 0$, the requisite limit, $\lim_{s\to 3/4} \mathcal{R}_{1/2}(s, \alpha) = 0$, which is certainly not the desired norm. What we are exhibiting is that the subset of sequences $(a_n) = (n^{-s})$ for $s > 0$ is *norming* in ℓ^p for $\tau = 1$ but not apparently for general $\tau > 0$.

One may also study the corresponding behaviour of Hardy's inequality (8.10). For example, setting $a_n := 1/n$ and denoting $H_n := \sum_{k=1}^{n} 1/k$ in (8.10) yields

$$\sum_{n=1}^{\infty} \left(\frac{H_n}{n}\right)^p \leq \left(\frac{p}{p-1}\right)^p \zeta(p).$$

Application of the integral test and the evaluation

$$\int_1^{\infty} \left(\frac{\log x}{x}\right)^p dx = \frac{\Gamma(1+p)}{(p-1)^{p+1}},$$

for $p > 1$ easily shows the constant is again best possible.

7 Computational Challenge Problems

In [44] we gave ten "challenge problems" and indicated that the solutions could be found scattered through [44] and [45]. Subsequently and annotated and much enhanced version of the promised solutions has been published in the American Mathematical Monthly [17] and is reproduced in full as Chapter 9. Our set was triggered by Nick Trefethen's *SIAM 100 Digit Challenge*, wonderfully described in [109] and reviewed in [43].

We conclude this lecture with a visit some to two the high spots of our problem set.

7.1 Finding $\zeta(3,1,3,1)$

Problem 9. *Calculate*

$$\sum_{i>j>k>l>0} \frac{1}{i^3 j k^3 l}.$$

Extra credit: Express this constant as a single-term expression involving a well-known mathematical constant.

History and context. In the notation introduced before, we ask for the value of $\zeta(3,1,3,1)$. The study of such sums in two variables, as we noted, originates with Euler. These investigations were apparently due to a serendipitous mistake. Goldbach wrote to Euler [44, pp. 99–100]:

> When I recently considered further the indicated sums of the last two series in my previous letter, I realized immediately that the same series arose due to a mere writing error, from which indeed the saying goes, "Had one not erred, one would have achieved less. [*Si non errasset, fecerat ille minus*]."

Euler's *reduction formula* is

$$\zeta(s,1) \quad = \quad \frac{s}{2}\,\zeta(s+1) - \frac{1}{2}\sum_{k=1}^{s-2} \zeta(k+1)\zeta(s+1-k),$$

which *reduces* the given double Euler sums to a sum of products of classical ζ-values. Euler also noted the first *reflection formulas*

$$\zeta(a,b) + \zeta(b,a) \quad = \quad \zeta(a)\zeta(b) - \zeta(a+b),$$

certainly valid when $a > 1$ and $b > 1$. This is an easy algebraic consequence of adding the double sums. Another marvelous fact is the *sum formula*

$$\sum_{\Sigma a_i = n, a_i \geq 0} \zeta\left(a_1+2, a_2+1, \cdots, a_r+1\right) \quad = \quad \zeta(n+r+1) \quad (8.24)$$

for nonnegative integers n and r. This, as David Bradley observes, is equivalent to the generating function identity

$$\sum_{n>0} \frac{1}{n^r(n-x)} \quad = \quad \sum_{k_1>k_2>\cdots k_r>0} \prod_{j=1}^{r} \frac{1}{k_j - x}.$$

The first three nontrivial cases of (8.24) are $\zeta(3) = \zeta(2,1), \zeta(4) = \zeta(3,1) + \zeta(2,2)$, and $\zeta(2,1,1) = \zeta(4)$.

Solution. We notice that such a function is a generalization of the zeta-function. As before, we define

$$\zeta(s_1, s_2, \cdots, s_k; x) \sum_{n_1 > n_2 > \cdots > n_k > 0} \frac{x_1^n}{n_1^{s_1} n_2^{s_2} \cdots n_r^{s_r}}, \tag{8.25}$$

for $s_1, s_2, \ldots, s_k$ nonnegative integers. We see that we are asked to compute $\zeta(3, 1, 3, 1; 1)$. Such a sum can be evaluated directly using the EZFace+ interface at
`http://www.cecm.sfu.ca/projects/ezface+`, which employs the Hölder convolution, giving us the numerical value

$$\begin{aligned}&0.00522956956353096010093065228389923158989042078463463552254744897214886954466015007497545432485610401627\ldots.\end{aligned} \tag{8.26}$$

Alternatively, we may proceed using differential equations. It is fairly easy to see [45, sec. 3.7] that

$$\frac{d}{dx}\zeta(n_1, n_2, \cdots, n_r; x) = \frac{1}{x}\zeta(n_1 - 1, n_2, \cdots, n_r; x), \qquad (n_1 > 1), \tag{8.27}$$

$$\frac{d}{dx}\zeta(n_1, n_2, \cdots, n_r; x) = \frac{1}{1-x}\zeta(n_2, \cdots, n_r; x), \qquad (n_1 = 1), \tag{8.28}$$

with initial conditions $\zeta(n_1; 0) = \zeta(n_1, n_2; 0) = \cdots \zeta(n_1, \cdots, n_r; 0) = 0$, and $\zeta(\cdot; x) \equiv 1$. Solving

```
> dsys1 > diff(y3131(x),x) = y2131(x)/x,
> diff(y2131(x),x) = y1131(x)/x,
> diff(y1131(x),x) = 1/(1-x)*y131(x),
> diff(y131(x),x) = 1/(1-x)*y31(x),
> diff(y31(x),x) = y21(x)/x,
> diff(y21(x),x) = y11(x)/x,
> diff(y11(x),x) = y1(x)/(1-x),
> diff(y1(x),x) = 1/(1-x);
> init1 = y3131(0) = 0,y2131(0) = 0, y1131(0) = 0,
>                    y131(0)=0,y31(0)=0,y21(0)=0,y11(0)=0,y1(0)=0;
```

in *Maple*, we obtain 0.005229569563518039612830536519667669502942 (this is valid to thirteen decimal places). Maple's `identify` command is unable to identify portions of *this* number, and the inverse symbolic calculator does not return a result. It should be mentioned that both *Maple* and the ISC identified the constant $\zeta(3, 1)$ (see the remark under the "history and context" heading). From the hint for this question, we know this is a single-term expression. Suspecting a form similar to $\zeta(3, 1)$, we

search for a constants c and d such that $\zeta(3,1,3,1) = c\pi^d$. This leads to $c = 1/81440 = 2/10!$ and $d = 8$.

Further history and context. We start with the simpler value, $\zeta(3,1)$. Notice that

$$-\log(1-x) = x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots,$$

so

$$\begin{aligned} f(x) &= -\log(1-x)/(1-x) = x + (1+\frac{1}{2})x^2 + (1+\frac{1}{2}+\frac{1}{3})x^3 + \cdots \\ &= \sum_{n \geq m > 0} \frac{x^n}{m}. \end{aligned}$$

As noted in the section on double Euler sums,

$$\frac{(-1)^{m+1}}{\Gamma(m)} \int_0^1 x^n \log^{m-1} x \; dx \frac{1}{(n+1)^m},$$

so integrating f using this transform for $m = 3$, we obtain

$$\begin{aligned} \zeta(3,1) &= \frac{(-1)}{2} \int_0^1 f(x) \log^2 x \; dx \\ &= 0.27058080842778454787900092 4\ldots. \end{aligned}$$

The corresponding generating function is

$$\sum_{n \geq 0} \zeta\left(\{3,1\}_n\right) x^{4n} = \frac{\cosh(\pi x) - \cos(\pi x)}{\pi^2 x^2},$$

equivalent to Zagier's conjectured identity

$$\zeta(\{3,1\}_n) = \frac{2\,\pi^{4n}}{(4n+2)}.$$

Here $\{3,1\}_n$ denotes n-fold concatenation of $\{3,1\}$.

The proof of this identity (see [45, p. 160]) derives from a remarkable factorization of the generating function in terms of hypergeometric functions:

$$\sum_{n \geq 0} \zeta(\{3,1\}_n)\, x^{4n} = {}_2\mathrm{F}_1\left(x\frac{(1+i)}{2}, -x\frac{(1+i)}{2}; 1; 1\right) {}_2\mathrm{F}_1\left(x\frac{(1-i)}{2}, -x\frac{(1-i)}{2}; 1; 1\right).$$

Finally, it can be shown in various ways that

$$\zeta(\{3\}_n) = \zeta(\{2,1\}_n)$$

for all n, while a proof of the numerically-confirmed conjecture

$$\zeta(\{2,1\}_n) \stackrel{?}{=} 2^{3n}\,\zeta(\{-2,1\}_n) \tag{8.29}$$

remains elusive. Only the first case of (8.29), namely,

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\sum_{m=1}^{n-1}\frac{1}{m} = 8\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\sum_{m=1}^{n-1}\frac{1}{m} \qquad (=\zeta(3))$$

has a self-contained proof [45]. Indeed, the only other established case is

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\sum_{m=1}^{n-1}\frac{1}{m}\sum_{p=1}^{m-1}\frac{1}{p^2}\sum_{q=1}^{p-1}\frac{1}{q} = 64\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\sum_{m=1}^{n-1}\frac{1}{m}\sum_{p=1}^{m-1}\frac{(-1)^p}{p^2}\sum_{q=1}^{p-1}\frac{1}{q} \quad (=\zeta(3,3)).$$

This is an outcome of a complete set of equations for multivariate zeta functions of depth four.

As we discussed in Chapter 1, there has been abundant evidence amassed to support identity (8.29) since it was found in 1996. This is the *only* identification of its type of an Euler sum with a distinct multivariate zeta-function.

7.2 $\pi/8$ or not?

Problem 8. *Calculate*

$$\pi_2 = \int_0^{\infty}\cos(2x)\prod_{n=1}^{\infty}\cos\left(\frac{x}{n}\right)\,dx.$$

History and context. The challenge of showing that $\pi_2 < \pi/8$ was posed by Bernard Mares, Jr., along with the problem of demonstrating that

$$\pi_1 = \int_0^{\infty}\prod_{n=1}^{\infty}\cos\left(\frac{x}{n}\right)\,dx < \frac{\pi}{4}.$$

This is indeed true, although the error is remarkably small, as we shall see.

Solution. The computation of a high-precision numerical value for this integral is rather challenging, owing in part to the oscillatory behavior of $\prod_{n\geq 1}\cos(x/n)$, see Figure 8.12, but mostly because of the difficulty of computing high-precision evaluations of the integrand. Note that evaluating thousands of terms of the infinite product would produce only a few correct digits. Thus it is necessary to rewrite the integrand in a form more suitable for computation as discussed in Chapter 3.

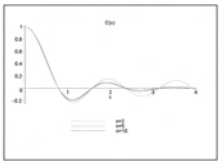


Figure 8.12. Approximations to $\prod_{n\geq 1}\cos(x/n)$.

Let $f(x)$ signify the integrand. We can express $f(x)$ as

$$f(x) \quad = \quad \cos(2x)\left[\prod_{1}^{m}\cos(x/k)\right]\exp(f_m(x)), \tag{8.30}$$

where we choose m greater than x and where

$$f_m(x) \quad = \quad \sum_{k=m+1}^{\infty}\log\cos\left(\frac{x}{k}\right). \tag{8.31}$$

The kth summand can be expanded in a Taylor series [3, p. 75], as follows:

$$\log\cos\left(\frac{x}{k}\right) \quad = \quad \sum_{j=1}^{\infty}\frac{(-1)^j 2^{2j-1}(2^{2j}-1)B_{2j}}{j(2j)!}\left(\frac{x}{k}\right)^{2j},$$

in which B_{2j} are again Bernoulli numbers. Observe that since $k > m > x$ in (8.31), this series converges. We can then write

$$f_m(x) \quad = \quad \sum_{k=m+1}^{\infty}\sum_{j=1}^{\infty}\frac{(-1)^j 2^{2j-1}(2^{2j}-1)B_{2j}}{j(2j)!}\left(\frac{x}{k}\right)^{2j}. \tag{8.32}$$

After applying the classical identity, [3, p. 807], that

$$B_{2j} \quad = \quad \frac{(-1)^{j+1}2(2j)!\zeta(2j)}{(2\pi)^{2j}}$$

and interchanging the sums, we obtain

$$f_m(x) \quad = \quad -\sum_{j=1}^{\infty}\frac{(2^{2j}-1)\zeta(2j)}{j\pi^{2j}}\left[\sum_{k=m+1}^{\infty}\frac{1}{k^{2j}}\right]x^{2j}.$$

Note that the inner sum can also be written in terms of the zeta-function, as follows:

$$f_m(x) \quad = \quad -\sum_{j=1}^{\infty} \frac{(2^{2j}-1)\zeta(2j)}{j\pi^{2j}} \left[\zeta(2j) - \sum_{k=1}^{m} \frac{1}{k^{2j}}\right] x^{2j}.$$

This can now be reduced to a compact form for purposes of computation as

$$f_m(x) \quad = \quad -\sum_{j=1}^{\infty} a_j b_{j,m} x^{2j}, \tag{8.33}$$

where

$$a_j \quad = \quad \frac{(2^{2j}-1)\zeta(2j)}{j\pi^{2j}}, \tag{8.34}$$

$$b_{j,m} \quad = \quad \zeta(2j) - \sum_{k=1}^{m} 1/k^{2j}. \tag{8.35}$$

We remark that $\zeta(2j), a_j$, and $b_{j,m}$ can all be precomputed, say for j up to some specified limit and for a variety of m. In our program, which computes this integral to 120-digit accuracy, we precompute $b_{j,m}$ for $m = 1, 2, 4, 8, 16, ..., 256$ and for j up to 300. During the quadrature computation, the function evaluation program picks m to be the first power of two greater than the argument x, and then applies formulas (8.30) and (8.33). It is not necessary to compute $f(x)$ for x larger than 200, since for these large arguments $|f(x)| < 10^{-120}$ and thus may be presumed to be zero.

The computation of values of the Riemann zeta-function can be done in various ways but, since what we really require is the entire set of values $\{\zeta(2j) : 1 \le j \le n\}$ for some n, by a convolution scheme described in [15]. It is important to note that the computation of both the zeta values and the $b_{j,m}$ must be done with a much higher working precision (in our program, we use 1600-digit precision) than the 120-digit precision required for the quadrature results, since the two terms being subtracted in formula (8.35) are very nearly equal. These values need to be calculated to a *relative* precision of 120 digits.

With this evaluation scheme for $f(x)$ in hand, the integral (8.30) can be computed using, for instance, the tanh-sinh quadrature algorithm, which can be implemented fairly easily on a personal computer or workstation and is also well suited to highly parallel processing [21, 14] and [45, p. 312]. This algorithm approximates an integral $f(x)$ on $[-1, 1]$ by transforming

it to an integral on $(-\infty, \infty)$ via the change of variable $x = g(t)$, where $g(t) = \tanh(\pi/2 \cdot \sinh t)$:

$$\int_{-1}^{1} f(x)\,dx \;=\; \int_{-\infty}^{\infty} f(g(t))g'(t)\,dt \;=\; h \sum_{j=-\infty}^{\infty} w_j f(x_j) + E(h) \tag{8.36}$$

Here $x_j = g(hj)$ and $w_j = g'(hj)$ are abscissas and weights for the tanh-sinh quadrature scheme (which can be precomputed), and $E(h)$ is the error in this approximation.

The function $g'(t) = \pi/2 \cdot \cosh t \cdot \mathrm{sech}^2(\pi/2 \cdot \sinh t)$ and its derivatives tend to zero very rapidly for large $|t|$. Thus, even if the function $f(t)$ has an infinite derivative, a blow-up discontinuity, or oscillatory behavior at an endpoint, the product function $f(g(t))g'(t)$ is in many cases quite well behaved, going rapidly to zero (together with all of its derivatives) for large $|t|$. In such cases, the Euler-Maclaurin summation formula can be invoked to conclude that the error $E(h)$ in the approximation (8.36) decreases very rapidly—faster than any power of h. In many applications, the tanh-sinh algorithm achieves quadratic convergence (i.e., reducing the size h of the interval in half produces twice as many correct digits in the result).

The tanh-sinh quadrature algorithm is designed for a finite integration interval. In this problem, where the interval of integration is $[0, \infty)$, it is necessary to convert the integral to a problem on a finite interval. This can be done with the simple substitution $s = 1/(x + 1)$, which yields an integral from 0 to 1.

In spite of the substantial computation required to construct the zeta- and b-arrays, as well as the abscissas x_j and weights w_j needed for tanh-sinh quadrature, the entire calculation requires only about one minute on a 2004-era computer, using the ARPREC arbitrary precision software package available at `http://crd.lbl.gov/~dhbailey/mpdist`. The first hundred digits of the result are the following:

0.39269908169872415480783042290993786052464543418723159592681228516209324713993854617901651274745536677 7....

A *Mathematica* program capable of producing 100 digits of this constant is available on Michael Trott's website:
`http://www.mathematicaguidebooks.org/downloads/N_2_01_Evaluated.nb`.

Using the Inverse Symbolic Calculator, for instance, one finds that this constant is likely to be $\pi/8$. But a careful comparison with a high-precision value of $\pi/8$, namely,

0.39269908169872415480783042290993786052464617492188822762186807403847705078577612482850435316776463349 7...,

reveals that they are *not* equal—the two values differ by approximately 7.407×10^{-43}. Indeed, these two values are provably distinct. This follows from the fact that

$$\sum_{n=1}^{55} 1/(2n+1) \quad > \quad 2 \; > \; \sum_{n=1}^{54} 1/(2n+1).$$

See [45, chap. 2] for additional details. We do not know a concise closed-form expression for this constant.

Further history and context. Recall the *sinc* function

$$\operatorname{sinc} x \quad = \quad \frac{\sin x}{x},$$

and consider, the seven highly oscillatory integrals:

$$\begin{aligned}
I_1 &= \int_0^\infty \operatorname{sinc} x \, dx \; = \; \frac{\pi}{2}, \\
I_2 &= \int_0^\infty \operatorname{sinc} x \operatorname{sinc}\left(\frac{x}{3}\right) \, dx \; = \; \frac{\pi}{2}, \\
I_3 &= \int_0^\infty \operatorname{sinc} x \operatorname{sinc}\left(\frac{x}{3}\right) \operatorname{sinc}\left(\frac{x}{5}\right) \, dx \; = \; \frac{\pi}{2}, \\
&\ddots \\
I_6 &= \int_0^\infty \operatorname{sinc} x \operatorname{sinc}\left(\frac{x}{3}\right) \cdots \operatorname{sinc}\left(\frac{x}{11}\right) \, dx \; = \; \frac{\pi}{2}, \\
I_7 &= \int_0^\infty \operatorname{sinc} x \operatorname{sinc}\left(\frac{x}{3}\right) \cdots \operatorname{sinc}\left(\frac{x}{13}\right) \, dx \; = \; \frac{\pi}{2}.
\end{aligned}$$

It comes as something of a surprise, therefore, that

$$\begin{aligned}
I_8 &= \int_0^\infty \operatorname{sinc} x \operatorname{sinc}\left(\frac{x}{3}\right) \cdots \operatorname{sinc}\left(\frac{x}{15}\right) \, dx \\
&= \frac{467807924713440738696537864469}{935615849440640907310521750000}\pi \approx 0.499999999992646\pi.
\end{aligned}$$

When this was first discovered by a researcher, using a well-known computer algebra package, both he and the software vendor concluded there was a "bug" in the software. Not so! It is fairly easy to see that the limit of the sequence of such integrals is $2\pi_1$. Our analysis, via Parseval's theorem, links the integral

$$I_N \quad = \quad \int_0^\infty \operatorname{sinc}(a_1 x) \operatorname{sinc}(a_2 x) \cdots \operatorname{sinc}(a_N x) \, dx$$

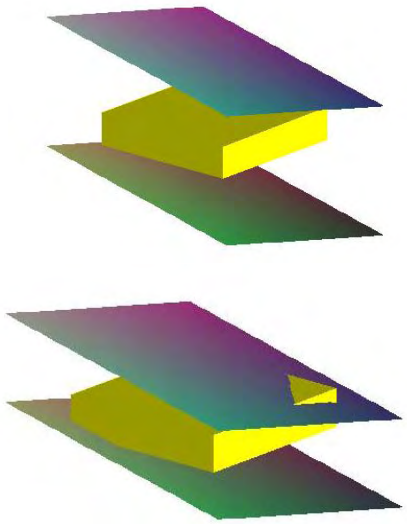


Figure 8.13. The 'Bite' in Three Dimensions

with the volume of the polyhedron P_N described by

$$P_N \quad = \quad \{x : |\sum_{k=2}^{N} a_k x_k| \leq a_1, |x_k| \leq 1, 2 \leq k \leq N\},$$

for $x = (x_2, x_3, \cdots, x_N)$. If we let

$$C_N \quad = \quad \{(x_2, x_3, \cdots, x_N) : -1 \leq x_k \leq 1, 2 \leq k \leq N\},$$

then

$$I_N \quad = \quad \frac{\pi}{2a_1} \frac{\text{Vol}(P_N)}{\text{Vol}(C_N)}.$$

Thus, the value drops precisely when the constraint $\sum_{k=2}^{N} a_k x_k \leq a_1$ becomes *active* and bites the hypercube C_N, as in Figure 8.13. That occurs when $\sum_{k=2}^{N} a_k > a_1$. In the foregoing,

$$\frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{13} < 1,$$

but on addition of the term 1/15, the sum exceeds 1, the volume drops, and $I_N = \pi/2$ no longer holds. A similar analysis applies to π_2. Moreover,

it is fortunate that we began with π_1 or the falsehood of $\pi_2 = 1/8$ would have been much harder to see.

Additional information on this problem is available at
`http://mathworld.wolfram.com/InfiniteCosineProductIntegral.html` and
`http://mathworld.wolfram.com/BorweinIntegrals.html`.

Chapter 9

Exercises

In this chapter we collect a series of exercises.

1 Exercises for Chapter 1

1. Nick Trefethen's fourth digit-challenge problem was given as (1.5) in Subsection 4.1. Find a numerical or graphical method to obtain ten good digits of the solution which occurs near $(-0.15, 0.29, -.028)$ with value -3.32834.

2. Prove that

$$0 < \frac{1}{3164} \int_0^1 \frac{x^8 \, (1-x)^8 \, \left(25 + 816\, x^2\right)}{1+x^2} dx = \frac{355}{113} - \pi$$

and derive the estimate that

$$\frac{355}{113} - \frac{911}{2630555928} < \pi < \frac{355}{113} - \frac{911}{5261111856}$$

3. **Powers of arcsin**. The formula for $\arcsin^2(x)$ discovered in the text is the first example of a family of formulas that can be experimentally discovered as is described in [55].

   (a) Show that

$$\arcsin^4\left(\frac{x}{2}\right) = \frac{3}{2} \sum_{k=1}^{\infty} \left\{ \sum_{m=1}^{k-1} \frac{1}{m^2} \right\} \frac{x^{2\,k}}{\binom{2\,k}{k} k^2}.$$

   More generally, show that for $|x| \le 2$ and $N = 1, 2, \ldots$

$$\frac{\arcsin^{2N}\left(\frac{x}{2}\right)}{(2N)!} = \sum_{k=1}^{\infty} \frac{H_N(k)}{\binom{2\,k}{k} k^2} \, x^{2\,k}, \tag{9.1}$$

where $H_1(k) = 1/4$ and

$$H_{N+1}(k) := \sum_{n_1=1}^{k-1} \frac{1}{(2n_1)^2} \sum_{n_2=1}^{n_1-1} \frac{1}{(2n_2)^2} \cdots \sum_{n_N=1}^{n_{N-1}-1} \frac{1}{(2n_N)^2}.$$

(b) Show that for $|x| \le 2$ and $N = 0, 1, 2, \ldots$

$$\frac{\arcsin^{2N+1}\left(\frac{x}{2}\right)}{(2N+1)!} = \sum_{k=0}^{\infty} \frac{G_N(k)\binom{2\,k}{k}}{2(2k+1)4^{2k}} x^{2k+1}, \tag{9.2}$$

where $G_0(k) = 1$ and

$$G_N(k) := \sum_{n_1=0}^{k-1} \frac{1}{(2n_1+1)^2} \sum_{n_2=0}^{n_1-1} \frac{1}{(2n_2+1)^2} \cdots \sum_{n_N=0}^{n_{N-1}-1} \frac{1}{(2n_N+1)^2}.$$

Proof. The formulae for $\arcsin^k(x)$ with $1 \le k \le 4$ are given on pages 262–63 of [28]. Berndt's proof implicitly gives the desired result since it establishes that for all a

$$e^{a\,\arcsin(x)} = \sum_{n=0}^{\infty} c_n \frac{x^n}{n!} \tag{9.3}$$

where

$$c_{2n+1} = a \prod_{k=1}^{n} \left(a^2 + (2k-1)^2\right), \qquad c_{2n} = \prod_{k=1}^{n} \left(a^2 + (2k-2)^2\right).$$

Now expanding the power of a^n on each side of (9.3) provides the asserted formula.

Another proof can be obtained from the hypergeometric identity

$$\frac{\sin(ax)}{a\sin(x)} = {}_2\mathrm{F}_1\left(\frac{1+a}{2}, \frac{1-a}{2}; \frac{3}{2}; \sin^2(x)\right)$$

given in [46, Exercise 16, p. 189]. □

Maple can prove identities such as (9.3) as the following code shows.

```
> ce:=n->product(a^2+(2*k)^2,k=0..n-1):
> co:=n->a*product(a^2+(2*k+1)^2,k=0..n-1):
> sum(ce(n)*x^(2*n)/(2*n)!,n=0..infinity) assuming x>0;
                        cosh(a arcsin(x))
> simplify(expand(sum(co(n)*x^(2*n+1)/(2*n+1)!,n=0..infinity)))
  assuming x>0;
                        sinh(a arcsin(x))
```

4. **A proof of Knuth's problem.** Why should such an identity hold? One clue was provided by the surprising speed with which *Maple* was able to calculate a high-precision value of this slowly convergent infinite sum. Evidently, the *Maple* software knew something that we did not. Looking under the hood, we found *Maple* was using the Lambert W function, which is the functional inverse of $w(z) = ze^z$.

   Another clue was the appearance of $\zeta(1/2)$ in the experimental identity, together with an obvious allusion to Stirling's formula in the original problem. This led us to conjecture the identity

$$\sum_{k=1}^{\infty}\left(\frac{1}{\sqrt{2\pi k}} - \frac{(1/2)_{k-1}}{(k-1)!\sqrt{2}}\right) \quad = \quad \frac{1}{\sqrt{2\pi}}\zeta\left(\frac{1}{2}\right), \tag{9.4}$$

   where $(x)_n$ denotes the rising factorial or Pochhammer function $x(x+1)\cdots(x+n-1)$, and where the binomial coefficients in (9.4) are the same as those of the function $1/\sqrt{2-2x}$. *Maple* successfully evaluated this summation, as shown on the RHS. We now needed to establish that

$$\sum_{k=1}^{\infty}\left(\frac{k^k}{k!e^k} - \frac{(1/2)_{k-1}}{(k-1)!\sqrt{2}}\right) \quad = \quad -\frac{2}{3}.$$

   Guided by the presence of the Lambert W function

$$W(z) \quad = \quad \sum_{k=1}^{\infty}\frac{(-k)^{k-1}z^k}{k!},$$

   an appeal to Abel's limit theorem suggested the conjectured identity

$$\lim_{z\to 1}\left(\frac{dW(-z/e)}{dz} + \frac{1}{2-2z}\right) \quad = \quad 2/3.$$

   Here again, *Maple* was able to evaluate this summation and establish the identity. □

   Such *instrumental* use of the computer is one of the most exciting features of experimental mathematics. The next three exercises explore the partition function and follow material in [46], see also [45].

5. **Euler's pentagonal number theorem** is

$$Q(q) := \prod_{n\geq 1}(1-q^n) \quad = \quad \sum_{n=-\infty}^{\infty}(-1)^n q^{(3n+1)n/2}. \tag{9.5}$$

One would be less prone to look at Q on the way to the partition function today when computation is very snappy. Determine empirically the series for $Q^3(q)$.

6. **Jacobi's triple product** in general form is

$$\prod_{n\geq 1}(1+xq^{2n-1})(1+x^{-1}q^{2n-1})(1-q^{2n}) \quad = \quad \sum_{n=-\infty}^{\infty} x^n q^{n^2}. \tag{9.6}$$

(a) Deduce from (9.6) the pentagonal number theorem of the previous exercise and that

$$\left(\sum_{n=-\infty}^{\infty}(-1)^n q^{(3n+1)n/2}\right)^3 \quad = \quad \sum_{m=0}^{\infty}(2m+1)(-1)^m\, q^{m(m+1)/2}. \tag{9.7}$$

7. **Modular properties of the partition function.**

(a) Prove that the partition function of 5n+4 is divisible by 5.
Proof Sketch. With Q as in (9.5), we obtain

$$\begin{aligned} qQ^4(q) &= q\,Q(q)Q^3(q) \qquad (9.8)\\ &= \sum_{m\geq 0}\sum_{n=-\infty}^{\infty}(-1)^{n+m}(2m+1)q^{1+(3n+1)n/2+m(m+1)/2}, \end{aligned}$$

from the triple product and pentagonal number theorems above. Now consider when k is a multiple of 5, and discover this can only happen if $2m+1$ is divisible by 5 as is the coefficient of q^{5m+5} in $qQ^4(q)$. Then by the binomial theorem,

$$(1-q)^{-5} \equiv (1-q^5)^{-1} \mod 5.$$

Consequently, the coefficient of the corresponding term in $qQ(q^5)/Q(q)$ is divisible by 5. Finally,

$$q+\sum_{n>1}p(n-1)q^n = qQ^{-1}(q) = \frac{qQ(q^5)}{Q(q)}\prod_{m=1}^{\infty}\sum_{n=0}^{\infty}q^{5mn},$$

as claimed. □

(b) Try to adapt extend this argument to show $p(7n+6)$ is divisible by 7.

The next three exercises are taken from [48] where many other such results may be found. We wish to show:

$$\zeta(2,1) := \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{m=1}^{n-1} \frac{1}{m} = \sum_{n=1}^{\infty} \frac{1}{n^3} = \zeta(3) \tag{9.9}$$

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{m=1}^{n-1} \frac{1}{m} = \sum_{n=1}^{\infty} \frac{1}{n^3} = 8 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \sum_{m=1}^{n-1} \frac{1}{m} := 8\,\zeta(\overline{2},1). \tag{9.10}$$

8. **Two proofs that** $\zeta(2,1) = \zeta(3)$.

   (a) A really quick proof of (9.9) considers

$$\begin{aligned} S &:= \sum_{n,k>0} \frac{1}{nk(n+k)} = \sum_{n,k>0} \frac{1}{n^2} \left(\frac{1}{k} - \frac{1}{n+k} \right) = \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{k=1}^{n} \frac{1}{k} \\ &= \zeta(3) + \zeta(2,1). \end{aligned} \tag{9.11}$$

   On the other hand,

$$\begin{aligned} S &= \sum_{n,k>0} \left(\frac{1}{n} + \frac{1}{k} \right) \frac{1}{(n+k)^2} = \sum_{n,k>0} \frac{1}{n(n+k)^2} + \sum_{n,k>0} \frac{1}{k(n+k)^2} \\ &= 2\zeta(2,1), \end{aligned}$$

   by symmetry.

   (b) Show that

$$\zeta(3) = \int_0^1 (\log x) \log(1-x) \frac{dx}{x}. \tag{9.12}$$

   Hint. Let $\varepsilon > 0$. Expand the integrand to get

$$\sum_{n=1}^{\infty} \frac{1}{(n+\varepsilon)^2} = \int_0^1 \int_0^1 \frac{(xy)^{\varepsilon}}{1-xy}\, dx\, dy.$$

   Differentiate with respect to ε and let $\varepsilon = 0$ to obtain

$$\begin{aligned} \zeta(3) &= -\frac{1}{2} \int_0^1 \int_0^1 \frac{\log(xy)}{1-xy}\, dx\, dy = -\frac{1}{2} \int_0^1 \int_0^1 \frac{\log x + \log y}{1-xy}\, dx\, dy \\ &= -\int_0^1 \int_0^1 \frac{\log x}{1-xy}\, dx \end{aligned}$$

   by symmetry. Now integrate with respect to y to get (9.12).

(c) Hence

$$2\zeta(3) = \int_0^1 \frac{\log^2 x}{1-x}\,dx = \int_0^1 \log^2(1-x)\frac{dx}{x} = \sum_{n,k>0}\int_0^1 \frac{x^{n+k-1}}{nk}\,dx$$
$$= \sum_{n,k>0} \frac{1}{nk(n+k)} = \zeta(2,1) + \zeta(3)$$

on appealing to the first half of the first proof. □

9. **A proof that** $8\,\zeta(2,\overline{1}) = \zeta(2,1)$. Let

$$J(x) := \sum_{n>k>0} \frac{x^n}{n^2 k}, \qquad 0 \le x \le 1.$$

(a) Show that

$$J(-x) = -J(x) + \frac{1}{4}J(x^2) + J\left(\frac{2x}{x+1}\right) - \frac{1}{8}J\left(\frac{4x}{(x+1)^2}\right). \quad (9.13)$$

Hint. Differentiate.

(b) Putting $x = 1$ gives $8\,J(-1) = J(1)$ immediately, which is (9.10).

Further proofs may be found in the Exercises for Chapter 8.

10. **AMM Problem 11103, October 2004.** Show that for positive integer n

$$2^{1-n}\sum_{k=1}^{n} \frac{\binom{n}{2\,k-1}}{2\,k-1} = \sum_{k=1}^{n} \frac{1}{k\binom{n}{k}}.$$

Solution. We show that the ordinary generating function of each side is the same. Indeed, applying the binomial theorem while interchanging sum and integral, allows one to write

$$\begin{aligned} 2\sum_{n=1}^{\infty}\sum_{k=1}^{n} \frac{\binom{n}{2k-1}}{2k-1}\left(\frac{y}{2}\right)^n &= \int_0^1 \frac{\sum_{n=1}^{\infty}\left(\frac{y(1+t)}{2}\right)^n - \left(\frac{y(1-t)}{2}\right)^n}{t}dt \\ &= \int_0^1 \frac{4\,y}{(y+yt-2)\,(y-yt-2)}dt \\ &= \frac{-\ln(1-y)}{1-y/2}. \end{aligned} \quad (9.14)$$

Also, using the beta function to write $1/(k\binom{n}{k})$ as $\int_0^1 t^{k-1}(1-t)^{n-k}\,dt$, produces

$$\begin{aligned}\sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{1}{k\binom{n}{k}}\,y^n &= \int_0^1 \frac{\sum_{n=1}^{\infty}\{t^n-(1-t)^n\}}{2t-1}y^n\,dt \\ &= \int_0^1 \frac{y}{(1+yt-y)(1-yt)}\,dt \\ &= \frac{-\ln(1-y)}{1-y/2}. \end{aligned} \tag{9.15}$$

2 Exercises for Chapter 2

1. Find rational coefficients a_i such that the identity

$$\begin{aligned}\pi &= a_1\arctan\frac{1}{390112}+a_2\arctan\frac{1}{485298} \\ &\quad +a_3\arctan\frac{1}{683982}+a_4\arctan\frac{1}{1984933} \\ &\quad +a_5\arctan\frac{1}{2478328}+a_6\arctan\frac{1}{3449051} \\ &\quad +a_7\arctan\frac{1}{18975991}+a_8\arctan\frac{1}{22709274} \\ &\quad +a_9\arctan\frac{1}{24208144}+a_{10}\arctan\frac{1}{201229582} \\ &\quad +a_{11}\arctan\frac{1}{2189376182}\end{aligned}$$

holds. Also show that an identity with even simpler coefficients exists if $\arctan 1/239$ is included as one of the terms on the RHS.

2. The constant $\zeta(3)$ can be written as the BBP series

$$\zeta(3) = \frac{1}{a_0}\sum_{k=0}^{\infty}\frac{1}{2^{12k}}\sum_{j=1}^{23}\frac{a_j}{24k+j}$$

for certain integers a_i. Find the integers a_i, in lowest terms.

3. The constant $\arctan(4/5)$ can be written as the BBP series

$$\arctan\left(\frac{4}{5}\right) = \frac{1}{a_0}\sum_{k=0}^{\infty}\frac{1}{2^{20k}}\sum_{j=1}^{39}\frac{a_j}{40k+j}$$

for certain integers a_i. Find the integers a_i, in lowest terms.

4. Calculate π, $\sqrt{2}$, e, $\log 2$ and $\log 10$ to 100,000 digits, and then tabulate the ten single-digit frequencies, and the 100 double-digit frequencies. One statistical procedure for testing the hypothesis that the empirical frequencies of n-long strings of digits are random is the χ^2 test. The χ^2 statistic of the k observations $X_1, X_2, \cdots, X_k$ is defined as

$$\chi^2 = \sum_{i=1}^{k} \frac{(X_i - E_i)^2}{E_i}$$

where E_i is the expected value of the random variable X_i. In this case $k = 10$ for single-digit frequencies (100 for double-digit frequencies), $E_i = 10^{-n}d$ for all i (here $d100,000$), and (X_i) are the observed digit counts (or double-digit counts). The mean of the χ^2 statistic is $k-1$, and its standard deviation is $\sqrt{2(k-1)}$.

5. Calculate the BBP sequences associated with $\log 2$, up to say 1000 terms, and then tabulate the frequency of appearance in the sixteen subintervals $[0, 1/16), [1/16, 2/16), \cdots, [15/16, 1]$. Use the chi-square test mentioned in the previous exercise to test the hypothesis that the iterates of the BBP sequence are uniformly distributed in the unit interval. Do the same for the BBP sequence associated with π. Note that high precision arithmetic software must be used in calculations of the iterates.

3 Exercises for Chapter 3

1. The integral

$$\frac{2}{\sqrt{3}} \int_0^1 \frac{\log^8(x) \arctan[x\sqrt{3}/(x-2)]}{x+1}\, dx$$

satisfies a linear relation involving the constants $L_3(10)$, $L_3(9)\log 3$, $L_3(8)\pi^2$, $L_3(7)\zeta(3)$, $L_3(6)\pi^4$, $L_3(5)\zeta(5)$, $L_3(4)\pi^4$, $L_3(3)\zeta(7)$, $L_3(2)\pi^8$ and $L_3(1)\zeta(9)$, where $L_3(s) = \sum_{n=1}^{\infty} [1/(3n-2)^s - 1/(3n-1)^s]$. Find the integer coefficients of this relation.

2. Check if the relation given in the previous problem extends to an analogous relation for the integral with $\log^{10} x$ in the numerator. How about for $\log^{12} x$?

3. Evaluate the following integrals, by numerically computing them and then trying to recognize the answers, either by using the Inverse Symbolic Calculator at `http://www.cecm.sfu.ca/projects/ISC`, or by using a PSLQ facility, such as that built into the Experimental Mathematician's Toolkit, available at `http://www.experimentalmath.info`. All of the answers are simple expressions involving familiar mathematical constants such as π, e, $\sqrt{2}$, $\sqrt{3}$, $\log 2$, $\zeta(3)$, G (Catalan's constant), and γ (Euler's constant). Many of these can be evaluated analytically using symbolic computing software. The intent here is to provide exercises for numerical quadrature and constant recognition facilities.

(a) $\displaystyle\int_0^1 \frac{x^2\,dx}{(1+x^4)\sqrt{1-x^4}}$

(b) $\displaystyle\int_0^\infty xe^{-x}\sqrt{1-e^{-2x}}\,dx$

(c) $\displaystyle\int_0^\infty \frac{x^2\,dx}{\sqrt{e^x-1}}$

(d) $\displaystyle\int_0^{\pi/4} x\tan x\,dx$

(e) $\displaystyle\int_0^{\pi/2} \frac{x^2\,dx}{1-\cos x}$

(f) $\displaystyle\int_0^{\pi/4} (\pi/4 - x\tan x)\tan x\,dx$

(g) $\displaystyle\int_0^{\pi/2} \frac{x^2\,dx}{\sin^2 x}$

(h) $\displaystyle\int_0^{\pi/2} \log^2(\cos x)\,dx$

(i) $\displaystyle\int_0^1 \frac{\log^2 x\,dx}{x^2+x+1}$

(j) $\displaystyle\int_0^1 \frac{\log(1+x^2)\,dx}{x^2}$

(k) $\displaystyle\int_0^\infty \frac{\log(1+x^3)\,dx}{1-x+x^2}$

(l) $\displaystyle\int_0^\infty \frac{\log x\,dx}{\cosh^2 x}$

(m) $$\int_0^1 \frac{\arctan x}{x\sqrt{1-x^2}}$$

(n) $$\int_0^{\pi/2} \sqrt{\tan t}\, dt$$

Answers: (a) $\pi/8$, (b) $\pi(1+2\log 2)/8$, (c) $4\pi(\log^2 2 + \pi^2/12)$, (d) $(\pi\log 2)/8 + G/2$, (e) $-\pi^2/4 + \pi\log 2 + 4G$, (f) $(\log 2)/2 + \pi^2/32 - \pi/4 + (\pi\log 2)/8$, (g) $\pi\log 2$, (h) $\pi/2(\log^2 2 + \pi^2/12)$, (i) $8\pi^3/(81\sqrt{3})$, (j) $\pi/2 - \log 2$, (k) $2(\pi\log 3)/\sqrt{3}$, (l) $\log\pi - 2\log 2 - \gamma$, (m) $[\pi\log(1+\sqrt{2})]/2$, (n) $\pi\sqrt{2}/2$.

4. **Evaluation of infinite series.** Evaluate the following infinite series, by numerically computing them and then trying to recognize the answers, either by using the Inverse Symbolic Calculator at

   `http://www.cecm.sfu.ca/projects/ISC,`

   or else by using a PSLQ facility, such as that built into the Experimental Mathematician's Toolkit, available at

   `http://www.experimentalmath.info`

   All of the answers are simple expressions involving familiar mathematical constants such as π, e, $\sqrt{2}$, $\sqrt{3}$, $\log 2$, $\zeta(3)$, G (Catalan's constant), and γ (Euler's constant).

(a) $$\sum_0^\infty \frac{50n-6}{2^n\binom{3n}{n}}$$

(b) $$\sum_0^\infty \frac{2^{n+1}}{\binom{2n}{n}}$$

(c) $$\sum_0^\infty \frac{12n2^{2n}}{\binom{4n}{2n}}$$

(d) $$\sum_0^\infty \frac{(4n)!(1+8n)}{4^{4n}n!^4}$$

(e) $$\sum_0^\infty \frac{(4n)!(19+280n)}{4^{4n}n!^4 99^{2n+1}}$$

(f) $$\sum_0^\infty \frac{(2n)!(3n)!4^n(4+33n)}{n!^5 108^n 125^n}$$

$$\text{(g)} \quad \sum_0^\infty \frac{(-27)^n(90n+177)}{16^n\binom{3n}{n}}$$

$$\text{(h)} \quad \sum_0^\infty \frac{275n-158}{2^n\binom{3n}{n}}$$

$$\text{(i)} \quad \sum_0^\infty \frac{8^n(520+6240n-430n^2)}{\binom{4n}{n}}$$

$$\text{(j)} \quad \sum_0^\infty \frac{\binom{2n}{n}}{n^2 4^n}$$

$$\text{(k)} \quad \sum_0^\infty \frac{(-1)^n}{n^3 2^n\binom{2n}{n}}$$

$$\text{(l)} \quad \sum_0^\infty \frac{8^n(338-245n)}{3^n\binom{3n}{n}}$$

$$\text{(m)} \quad \sum_1^\infty \frac{(-9)^n\binom{2n}{n}}{6n^2 64^n} - \sum_1^\infty \frac{3^n\binom{2n}{n}}{n^2 16^n}$$

Answers: (a) π, (b) $\pi+4$, (c) $3\pi+8$, (d) $2/(\pi\sqrt{3})$, (e) $2/(\pi\sqrt{11})$, (f) $15\sqrt{3}/(2\pi)$, (g) $120-64\log 2$, (h) $6\log 2-135$, (i) $45\pi-1164$, (j) $\pi^2/6-2\log^2 2$, (k) $\log^3 2/6-\zeta(3)/4$, (l) $162-6\pi\sqrt{3}-18\log 3$, (m) $\pi^2/18+\log^2 2-\log^3 3/6$.

These examples are due to Gregory and David Chudnovsky.

4 Exercises for Chapter 4

1. **Huygens' principle.** Huygens' Principle for the propagation of optical waves through some aperture states that

> *"'light falling on the aperture propagates as if every [surface] element [ds] emitted a spherical wave the amplitude and phase of which are given by that of the incident wave [u^i]"* [214].

An analogous statement holds for scattering from obstacles, which we derive precisely here. Let Ω have C^2 boundary. We illuminate Ω with a plane wave $u^i(x,\widehat{\eta}) = e^{i\kappa\widehat{\eta}\cdot x}$. Suppose that the obstacle is

sound soft, that is, the total field $u(x,\widehat{\eta}) = u^i(x,\widehat{\eta}) + u^s(x,\widehat{\eta}) = 0$ for $x \in \partial\Omega$. Show that

$$u^s(x,\widehat{\eta}) = -\int_{\partial\Omega} \frac{\partial u(y,\widehat{\eta})}{\partial \nu(y)} \Phi(x,y) ds(y), \quad x \in \Omega^o. \tag{9.16}$$

Interpret this in terms of Huygens' principle above.

Guide. Apply Green's formula Eq.(4.12) to the scattered field u^s (with justification for why you can do this), Green's theorem Eq.(4.11) to the incident field u^i, and use the boundary condition. Note that this argument does not make use of the explicit designation of the incident field, thus the statement Eq.(9.16) can be more broadly applied to *any* entire incident field v^i. □

2. **Far field reciprocity.** Suppose that the scatterer Ω generates a total field u satisfying *Dirichlet* boundary conditions: $u = 0$ on $\partial\Omega$. Prove the *reciprocity* relation for the far field given by Eq.(4.15).

   Guide. Following the elegant argument of Colton and Kress [80], use Green's Theorem Eq.(4.11), the Helmholtz equation for the incident and scattered waves Eq.(4.3), the radiation condition Eq.(4.6) for the scattered wave, and the boundary integral equation for the far field pattern Eq.(4.14) to show that

$$\beta\left(u^\infty(\widehat{x},\widehat{\eta}) - u^\infty(-\widehat{\eta},-\widehat{x})\right) = \int_{\partial\Omega} \left(u(\cdot,\widehat{\eta}) \frac{\partial u(\cdot,-\widehat{x})}{\partial \nu} - u(\cdot,-\widehat{x}) \frac{\partial u(\cdot,\widehat{\eta})}{\partial \nu} \right),$$

   where β, in the two-dimensional setting, is given by Eq.(4.8).

   Show that this implies Eq.(4.15) when u satisfies the boundary condition $u(\cdot,\widehat{\eta}) = 0$ on $\partial\Omega$ for all $\widehat{\eta} \in \mathbb{S}$. □

3. **Resolvent kernel for the Dirichlet Laplacian.** In the proof of Theorem 3.2 we made use of the *resolvent kernel for the Dirichlet Laplacian*. This is the total field for scattering due to an incident *point source*:

$$w^i(x,z) \equiv \Phi(x,z), \qquad x,z \in \mathbb{R}^2,\ x \neq z \tag{9.17}$$

   where Φ is defined by Eq.(4.13). The total field satisfies the boundary value problem

$$\begin{aligned} (\triangle + \kappa^2) w(x,z) &= -\delta(x-z), \qquad x,z \in \Omega^o; \\ w(x,z) &= 0, \qquad x \in \partial\Omega. \end{aligned} \tag{9.18}$$

Problem Eq.(9.18) is uniquely solvable [80] with $w = w^i + w^s$ for the scattered field w^s satisfying Eq.(4.6).

Show that

$$w^s(x,z) = -\int_{\partial\Omega} \frac{\partial w(y,z)}{\partial \nu(y)} \Phi(x,y)\, ds(y), \quad x,z \in \Omega^o, \tag{9.19}$$

analogous to Eq.(9.16).

4. **Symmetry of the resolvent kernel.** Continuing with the previous problem, show that the incident field w^i is spatially symmetric,

$$w^i(x,z) = w^i(z,x), \tag{9.20}$$

thus w^s and w also have this property.

5. **Asymptotic behavior.** Let w satisfy Eq.(9.18) in Ω^o and $w = w^i + w^s$ with the incident point source $w^i(x,z) = \Phi(x,z)$. Let $u(z,-\widehat{x})$ be the total field from scattering due to the incident plane wave $u^i(z,-\widehat{x}) = e^{-\mathrm{i}\kappa\widehat{x}\cdot z}$ with $u(z,-\widehat{x}) = 0$ on $\partial\Omega$. For β given by Eq.(4.8), show that the following relation holds as $|x| \to \infty$:

$$w(x,z) = \frac{e^{\mathrm{i}\kappa|x|}}{|x|^{1/2}} \left\{ \beta u(z,-\widehat{x}) + O\left(|x|^{-1}\right) \right\}. \tag{9.21}$$

6. **Mixed reciprocity.** Use Eq.(9.21), together with the asymptotic behavior of a radiating solution to the Helmholtz equation given by Eq.(4.7) to show the *mixed reciprocity relation*:

$$w^\infty(\widehat{x},z) \;=\; \beta u^s(z,-\widehat{x}), \qquad \widehat{x} \in \mathbb{S},\; z \in \Omega^o \tag{9.22}$$

for β given by Eq.(4.8).

7. **Potential theory.** In this exercise we use *potential theoretic* techniques to calculate the solution to the *exterior Dirichlet problem*, that is, u^s satisfying Eq.(4.3) on Ω^o with the boundary condition $u^s = f$ on $\partial\Omega$ and the radiation condition Eq.(4.6). To do this, we introduce the *acoustic single-* and *double-layer operators* given respectively as

$$\begin{aligned} (S\varphi)(x) &\equiv 2\int_{\partial\Omega} \varphi(y)\Phi(x,y)\, ds(y), \quad x \in \partial\Omega \\ (K\varphi)(x) &\equiv 2\int_{\partial\Omega} \varphi(y)\frac{\partial\Phi(x,y)}{\partial\nu(y)}\, ds(y), \quad x \in \partial\Omega, \end{aligned} \tag{9.23}$$

where Φ is the two-dimensional fundamental solution given by Eq.(4.13). It can be shown [80] that, if the potential φ satisfies the integral equation

$$(I + K - \mathrm{i}\gamma S)\varphi = f \qquad (\gamma \neq 0), \tag{9.24}$$

then u^s satisfies the exterior Dirichlet problem where u is given explicitly by

$$u(x) = \int_{\partial\Omega} \left(\frac{\partial \Phi(x,y)}{\partial \nu(y)} - \mathrm{i}\gamma \Phi(x,y) \right) \varphi(y)\ ds(y), \quad x \in \mathbb{R}^2 \setminus \partial\Omega. \tag{9.25}$$

Show that the far field pattern due to scattering from a sound-soft obstacle with an incident plane wave of direction $\widehat{\eta}$ is given by

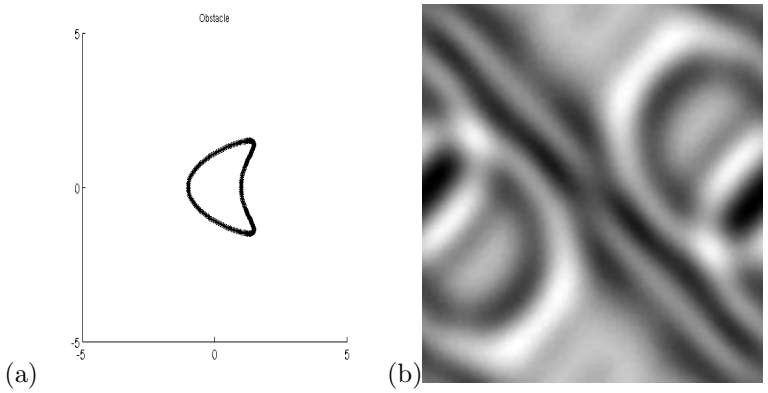
$$u^\infty(\widehat{x}, \widehat{\eta}) = \beta \int_{\partial\Omega} \left(\frac{\partial e^{-\mathrm{i}\kappa \widehat{x}\cdot y}}{\partial \nu(y)} - \mathrm{i} e^{-\mathrm{i}\kappa \widehat{x}\cdot y} \right) \varphi(y)\ ds(y), \quad \widehat{x} \in \mathbb{S}, \tag{9.26}$$

where φ satisfies

$$(I + K - \mathrm{i}S)\varphi = -e^{\mathrm{i}\kappa x\cdot\widehat{\eta}}. \tag{9.27}$$

8. **Compute the far field pattern.** In this exercise, you will write a computer program to generate the far field data for scattering from the *sound soft*, kite-shaped obstacle shown in Figure 9.1 whose boundary is given parametrically by

$$\partial\Omega(\theta) \equiv (\cos\theta + 0.65\cos(2\theta) - 0.65, 1.5\sin\theta) \tag{9.28}$$



(a) (b)

Figure 9.1. (a) Kite-shaped obstacle given by Eq.(9.28) and generated by the MATLAB code in Figure 9.2 (b) Real part of far field data for Exercise 8.

```
function [bdy_mat,dbdy_mat]=...
        Make_phantom(theta_res)

% Parameters:
theta_res=240;
objrot=0;

%-------------------------------------------
% Parametric grid generation:
%-------------------------------------------
h=2*pi/(theta_res-1);
t_vec=[0:h:2*pi];
rot_mat = [cos(objrot) -sin(objrot); sin(objrot) cos(objrot)];
bdy_mat=zeros(2,theta_res);
bdy2_mat=zeros(2,theta_res);
dbdy_mat = bdy_mat;

%-----------------
% kite
%-----------------
% boundary:
bdy_mat(1,:)  =  -(cos(t_vec) + 0.65 * cos(2*t_vec) - 0.65);
bdy_mat(2,:)  =  1.5 * sin(t_vec);

% Normal derivative:
% In order to get the correct unit outward normal to the kite
% shaped obstacle we have to multiply the derivative by -1.  The
% peculiarity is due to the way it is parameterized.
dbdy_mat(1,:) = -(sin(t_vec) + 2*.65*sin(2*t_vec));
dbdy_mat(2,:) = -1.5*cos(t_vec);
```

Figure 9.2. MATLAB/OCTAVE code for constructing kite shaped object shown in Figure 9.1

We give the MATLAB^{TM}code (also runs with OCTAVE) to generate the obstacle and its normal derivative in Figure 9.2.

Use Eq.(9.26)-(9.27) to calculate the far field data at 128 points equally distributed on $[-\pi, \pi]$ (full aperture) for a wavelength $\kappa = 3$ and 128 incident field directions $\widehat{\eta}$ coincident with the far field "measurement" points. Your answer should have real part resembling Figure 9.1(b). As a guide, we show below a a partial MATLAB^{TM}code (also runs in OCTAVE).

```
% PARAMETERS:
kappa=10;  % wavenumber
nff=128; % number of far field measurements
n_inc = 128; % number of incident fields
apang = pi; % aperture angle:  pi=full aperture.

% r*_mat are matrices that keep track of the distances between
% the source points, y=(y1,y2), and the integration points, x=(x1,x2),
% on the boundary of the
% scatterer.  Note that these are parameterized by theta_res points on -pi
% to pi.  dx*_mat are the matrices of normal derivatives to the boundary.
% The following involves a ‘‘regularization" of the
% point source, which has a singularity at x=y.
% A more sophisticated numerical quadrature would
% be appropriate for applications that require high accuracy,
% but for the purposes of this exercise, our rough approach is
% very effective.

% matrix of differences for the kernel
tmp_vec = ones(theta_res,1);
x1_mat=tmp_vec*bdy_mat(1,:);
x2_mat=tmp_vec*bdy_mat(2,:);
dx1_mat=tmp_vec*dbdy_mat(1,:);
dx2_mat=tmp_vec*dbdy_mat(2,:);

min_mat=eps*ones(size(x1_mat));
r1_mat=x1_mat.'-(x1_mat);
r2_mat=x2_mat.'-(x2_mat);
r_mat=max(sqrt(r1_mat.^2 + r2_mat.^2),min_mat);
dr_mat=sqrt(dx1_mat.^2 + dx2_mat.^2);


% Discrete kernel of the integral operators:
S_mat=2*(i/4*besselh(0,1,kappa*r_mat).*dr_mat); % Hankel function
K_mat=2*(1i*kappa/4*(-dx2_mat.*r1_mat+dx1_mat.*r2_mat).*(-besselh(1,1,kappa*r_mat))
% the derivative of the Hankel function changes the order of the function
% in K_mat

A_mat=eye(theta_res)+(2*pi/theta_res)*(K_mat-i*S_mat);

% incident field at the boundary of the scatterer:
% incident field has direction d_mat in Cartesian coordinates
% the incident directions are not perfectly symmetric across the the aperture
```

```
h_inc=2*apang/n_inc;
t_inc_vec=[-apang+h_inc/2:h_inc:apang-h_inc/2];

% incident directions  - Cartesian
d_mat(1,:)=cos(t_inc_vec);
d_mat(2,:)=sin(t_inc_vec);

% normal to the boundary
d1_mat=bdy_mat(1,:)'*(d_mat(1,:));
d2_mat=bdy_mat(2,:)'*(d_mat(2,:));

b_mat=-exp(i*kappa*(d1_mat+d2_mat));

% density
phi_mat=A_mat\b_mat;


% Now the far field.
% the measurements are not perfectly symmetric across the the aperture
hff=2*apang/nff;
tff_vec=[-apang+hff/2:hff:apang-hff/2];

% far field points
ffgrid_mat(1,:)=cos(tff_vec);
ffgrid_mat(2,:)=sin(tff_vec);

% matrix of differences for the far field kernel
tmp_vec = ones(theta_res,1);
ffx1_mat = (tmp_vec * ffgrid_mat(1,:)).';
ffx2_mat = (tmp_vec * ffgrid_mat(2,:)).';
tmp_vec = ones(nff,1);

x1_mat = tmp_vec * bdy_mat(1,:);
x2_mat = tmp_vec * bdy_mat(2,:);
dx1_mat = tmp_vec * dbdy_mat(1,:);
dx2_mat = tmp_vec * dbdy_mat(2,:);

r1_mat=ffx1_mat.*x1_mat;
r2_mat=ffx2_mat.*x2_mat;
r_mat=r1_mat+r2_mat;
dr_mat=sqrt(dx1_mat.^2 + dx2_mat.^2);

ffM_mat=exp(-i*kappa*r_mat).*dr_mat;
```

```
ffL_mat=kappa*(dx2_mat.*ffx1_mat-dx1_mat.*ffx2_mat).*exp(-i*kappa*r_mat);
beta=exp(i*pi/4)/sqrt(8*pi*kappa);
ff_mat=beta*(2*pi/theta_res)*(ffL_mat+ffM_mat)*phi_mat;
```

9. **Is the scatterer absorbing?** Show either numerically or analytically that the singular values of the far field operator $\mathcal{A}$ corresponding to the data in Exercise 8 lie on the circle centered at $\frac{1}{2\kappa}\left(\Im(\beta^{-1}), \Re(\beta^{-1})\right)$ and passing through the origin, hence, the scatterer, as we already know, is nonabsorbing.

10. **Linear sampling.** At each point z_j corresponding to a sample point on a rectangular grid in the computational domain $[-5,5]\times[-5,5]$, compute the density g_{z_j} that satisfies the unconstrained, Tikhonov-regularized optimization problem

$$\underset{g\in\mathbb{C}^n}{\text{minimize}}\ \|\mathcal{A}g-\Phi^{\infty}(z_j,\cdot)\|^2+\alpha\|g\|^2$$

where $\mathcal{A}$ is the discrete far field operator generated from the data computed in Exercise 8 and the regularization parameter $\alpha\approx 1e-7$. Generate a surface plot of the value $\|g_j\|$ versus z_j. If you choose the dynamic range of your surface plot correctly you should see something with a hole resembling the kite shown in Figure 9.1(a). *Hint.* The solution to the optimization problem can be written in closed form via the normal equations.

5 Exercises for Chapter 5

1. **The Takagi function for $0<a<\frac{1}{2}$.** Discuss and prove the differentiability properties of the Takagi function $T_{a,2}$ for $0<a<\frac{1}{2}$.

   Hint: It is relatively easy to conjecture that for $a<\frac{1}{2}$, $T_{a,2}$ does *not* have a derivative at any dyadic rational, but is differentiable everywhere else. This conjecture, however, is only "almost" true: There is precisely one exceptional value of a strictly between 0 and $\frac{1}{2}$ for which the conjecture is wrong. What is this value?

2. **Differentiability of the Weierstraß cosine series.** Prove that the Weierstraß cosine series $C_{a,2}$ for $1/2\le a<1$ is nowhere differentiable.

3. **Singular functions.** Choose $0<a<1$. Discuss and prove the differentiability properties of the unique continuous solution of the

system

$$\begin{aligned} f\left(\frac{x}{2}\right) &= a\, f(x) \\ f\left(\frac{x+1}{2}\right) &= (1-a)\, f(x) + a \end{aligned}$$

on $[0,1]$. *Hint:* Except for $a = \frac{1}{2}$, the function is *singular* (i.e., strictly monotone, with $f'(x) = 0$ a.e.). This example is from [95].

4. **The Cantor function.** Show that the system of functional equations

$$\begin{aligned} f\left(\frac{x}{3}\right) &= \frac{1}{2}\, f(x) \\ f\left(\frac{x+1}{3}\right) &= \frac{1}{2} \\ f\left(\frac{x+2}{3}\right) &= \frac{1}{2}\, f(x) + \frac{1}{2} \end{aligned}$$

on $[0,1]$ has a unique continuous solution, and discuss its differentiability properties.
Hint: The solution is a (the) *Cantor function*. In general, we call an $f : [0,1] \to [0,1]$ a Cantor function if it is surjective, but constant on each of a collection of intervals which together have full measure (such as the complement of the Cantor set). This example is from [211].

5. **One-sided differentiability.** Prove the following sharper version of Theorem 1.2: *Assume that $f \in C[0,1]$ has a finite one-sided derivative at some point in $[0,1]$. Then*

$$\lim_{n\to\infty} 2^n \cdot \min\left\{|\gamma_{i,n}(f)| : i = 0, \ldots, 2^{n-1} - 1\right\} \;=\; 0.$$

6. **The Minkowski function.** Minkowski's ?-function is defined as follows. Let $x \in [0,1)$ with simple continued fraction $[0; a_1, a_2, a_3, \ldots]$. Then

$$?(x) \;:=\; \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{2^{a_1+\cdots+a_k-1}}.$$

This function maps the rational numbers onto the dyadic rationals and the quadratic irrationals onto the rationals.

Find a system of functional equations which characterizes the Minkowski function and then use the functional equations to derive its differentiability properties.

7. **The Riemann function.** The function

$$R(x) := \sum_{n=1}^{\infty} \frac{\sin(n^2 \pi x)}{n^2}$$

was apparently communicated by Riemann to his pupils as an example of a function which does not have a finite derivative at a dense set of points (see [100, 131]). While this statement is true, it was proved only much later that this function has, in fact, finite derivatives at certain rational values (see [118, 119]).

Find a basis suitable for analyzing this function.

8. **A binary recursion.**

(a) Let $x \geq 0$. Set

$$a_0 = x \quad \text{and} \quad a_{n+1} = \frac{2a_n}{1 - a_n^2} \quad (\text{with } a_{n+1} = -\infty \text{ if } a_n = \pm 1).$$

Prove that

$$\sum_{\substack{a_n < 0 \\ n \geq 0}} \frac{1}{2^{n+1}} = \frac{\arctan x}{\pi}.$$

(b) In general (see also [58]), let an interval $I \subseteq \mathbb{R}$ and subsets $D_0, D_1 \subseteq I$ with $D_0 \cup D_1 = I$ and $D_0 \cap D_1 = \emptyset$ be given, as well as functions $r_0 : D_0 \to I$, $r_1 : D_1 \to I$. Then consider the system (S) of the following two functional equations for an unknown function $f : I \to [0, 1]$.

$$2f(x) = f(r_0(x)) \quad \text{if } x \in D_0, \tag{S_0}$$
$$2f(x) - 1 = f(r_1(x)) \quad \text{if } x \in D_1. \tag{S_1}$$

Such a system always leads to an iteration:

$$a_0 = x \quad \text{and} \quad a_{n+1} = \begin{cases} r_0(a_n), & a_n \in D_0, \\ r_1(a_n), & a_n \in D_1. \end{cases}$$

Prove that

$$f(a_0) = f(x) = \sum_{\substack{a_n \in D_1 \\ n \geq 0}} \frac{1}{2^{n+1}}.$$

(c) Find parameters such that the following functions can be written as solutions of the system (S) and thus can be computed via their binary expansion: $\ln(x)/\ln(2)$ on $[1, 2]$, $\arccos(x)/\pi$ on $[-1, 1]$, and $\begin{cases} \arctan(x)/\pi, & x \in [0, \infty), \\ 1 + \arctan(x)/\pi, & x \in (-\infty, 0) \end{cases}$.

(d) Find parameters such that a singular function or a Cantor function is a solution of (S).

(e) Experiment by plotting the solution for arbitrary (or intelligent) choices of the parameters.

9. **The Schilling equation.** The Schilling equation is the functional equation

$$4q\, f(qt) \;=\; f(t+1) + 2f(t) + f(t-1) \quad \text{for } t \in \mathbb{R}$$

with a parameter $q \in (0,1)$. It has its origin in physics, and although it has been studied intensively in recent years, there are still many open questions connected with it. The main question is to find values of q for which the Schilling equation has a nontrivial L^1-solution. Discuss this question!

Hint: If an L^1-function f satisfies (5.6), then a rescaled version of $f * f$ satisfies the Schilling equation.

10. **Unboundedness of the iteration.** Is the sequence of iterates $B_q^n f^{(0)}$ unbounded for $q = (\sqrt{5}-1)/2$ (or any other Pisot or non-Pisot q) and for, say, $f^{(0)} = \frac{1-q}{2}\,\chi_{[-1/(1-q),1/(1-q)]}$?

11. **Plotting f_q.** Find a good algorithm for computing and plotting the iterates $B_q^n f^{(0)}$ or any other approximation to f_q, for arbitrary q.

In fact, let's make this a contest. If you have a good (= fast) algorithm, then send it to `girgencecm.sfu.ca`. If possible, include an implementation of your algorithm in Maple. If your algorithm is faster than mine (on my laptop), then you will receive an honorable mention on our web page or in subsequent editions of this book!

6 Exercises for Chapter 6

Some of the following exercises are relatively straightforward: others are open-ended and could be used as the starting point for an experimental research project.

1. Use Kraitchik's method to factorize 2041.

2. Find an n for which Kraitchik's method works, but we don't get an easy factorization via Fermat's method.

3. Rediscover the diagonalization of the $(k+1) \times (k+1)$ matrix T

$$T_{ij} = \frac{\binom{i}{\frac{w+i-j}{2}}\binom{k-i}{\frac{w-i+j}{2}}}{\binom{k}{w}}$$

by computing explicit examples in Maple or Mathematica: use Sloane's database to help identify sequences.

4. Investigate the eigenstructure of other structured matrices: for example, consider the Toeplitz matrices

$$T_n = \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 1 & 1 & \dots & 0 \\ 0 & 0 & 1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & & & & & 1 \end{pmatrix}$$

Compute the eigenvalues for various values of n, plotting the values in increasing order. Guess, with the help of the inverse symbolic calculator if necessary, a formula for the jth largest eigenvalue of T_n. Compute the corresponding eigenvectors, and guess the formula for their entries. For more information on this and related problems, see Doust, Hirshhorn and Ho [99]

5. Fix the dimension k of the vector space, and find a way of uniformly generating binary vectors of weight 3. For various values of k generate random vectors until you obtain a linearly dependent set. Compare the number of vectors required to the upper and lower bounds given above. Do you think that there is a sharp threshold?

6. When $\mathcal{P}$ is the set of primes congruent to 1 (mod 4), what are the values of the constants δ and K appearing in Wirsing's theorem?

7. Pick a moderately large value n (the product of two large primes would be appropriate). Using Maple or Mathematica, estimate the constants δ and K appearing in Wirsing's theorem.

8. For those who know about quadratic reciprocity: suppose that n is the product $n = q_1 q_2$ of two primes. Find a condition *on* p for a prime p to be such that n is a quadratic residue (mod p). Assuming Dirichlet's theorem (that if a and m are relatively prime then the number of primes congruent to a (mod m) is asymptotic to $1/\phi(m)$) deduce that the relative density of $\mathcal{P}$ in the set of all primes should be 1/2. Compare the value of δ from the previous exercise to 1/2.

9. For small y, compare the value of $\mathbb{Z}_{\mathcal{P}}(y)$ to the value predicted (for *large* y) by Wirsing's theorem. Does our assumption that $\mathbb{Z}_{\mathcal{P}}(y)$ behaves similarly for small and for large y seem justified?

10. Devise some statistical tests to determine whether the heuristic that

$$\alpha_j = \Pr(p_j | x^2 - n) \simeq c\frac{\log(p_j)}{p_j}$$

is a reasonable one. For example, the number $x^2 - n$ lies between $\sqrt{n}$ and n and so the sums

$$\frac{1}{2}\log n < \sum_{p_j | x^2 - n} \log p_j < \log n$$

and

$$\sum_{j=1}^{k} \alpha_j \log p_j$$

should be about the same. (This ignores primes dividing $x^2 - n$ to the second or higher power, but their effect is small). Note that this heuristic illustrates again the fact that the size of $\mathcal{B}$ has an impact on the probability that a $\mathcal{B}$-smooth number is divisible by p.

11. Generate $l \times k$ random matrices with $\alpha_j = c/j$ for various values of c. Iteratively remove colons and solons from the array until all columns are either empty or have at least three 1's in. If there are more non-zero rows than columns, then the rows of the original array are linearly dependent.

    - What is the running time of the algorithm to delete all colons and solons?
    - How large should l be as a function of k so that the rows of the initial array are linearly dependent?
    - How large should l be as a function of k so that the final array has more non-zero rows than columns? Suggested values to try: $l = 0.5k$, $l = 0.6k$, $l = 0.9k$, $l = k + 1$
    - How many non-zero rows and columns does the final array have?

12. Suppose that a column of A contains exactly three 1's and that A has X_1 solons. After deleting the solons, what is the probability that the column now contains two 1's?

13. Suppose A has X_1 solons, X_2 colons, and X_3 columns with exactly three 1's. If we remove the colons and then the solons, the columns with exactly three 1's can remain unchanged, can become colons, solons or empty. What is the expected number of each that will be produced?

14. Suppose A has X_1 solons, X_2 colons, and X_r columns with exactly r 1's. If we remove the colons and then the solons, the columns with exactly r 1's can become columns with s 1's, $0 \le s \le r$. What is the expected number of each that will be produced?

15. Consider a stochastic model for the deletion of colons and solons along the following lines: consider the random variable

$$X(t) = (X_0(t), X_1(t), X_2(t), \dots),$$

where $X_i(t)$ is the number of columns containing exactly i 1's after t rounds of deletions of solons and colons. More precisely, $X(t-1)$ is the number before the tth round of deletions: in round t, remove all colons (by replacing colon-rows with their sum), and then remove all solons. Now update $X(t)$. Develop heuristics (by making reasonable assumptions where necessary) for the dynamics of $X(t)$.

16. Explain the seemingly paradoxical fact that increasing the number of columns in A *decreases* the size of the final non-zero array after iterated deletions. For example, the values in the table are typical. Each row is the result of iterated deletions of solons and colons for a random array A with $\alpha_j = 0.72/j$. The final size is the number of non-zero rows and columns.

Initial size	Final size
10000×10000	4700×2300
10000×15000	2250×1300
10000×20000	800×480
10000×25000	300×180
10000×30000	150×100
10000×35000	100×50
10000×40000	47×22
10000×45000	38×20

7 Exercises for Chapter 7

1. Give a direct proof of the Wallis' formula

$$\int_0^\infty \frac{dx}{(x^2+1)^{m+1}} \quad = \quad \frac{\pi}{2^{2m+1}} \binom{2m}{m}. \tag{9.29}$$

Hint. Convert the integral into a value of the beta function

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, dt. \tag{9.30}$$

The value $\Gamma(1/2) = \sqrt{\pi}$ can be obtained directly from Mathematica or you can reduce it to a well-known evaluation.

Generalize the previous argument to obtain the value

$$\int_0^\infty \frac{dx}{(x^4+1)^{m+1}} = \frac{\pi}{m!\,2^{2m+3/2}} \prod_{k=1}^{\infty}(4k-1). \qquad (9.31)$$

2. In the evaluation of the quartic integral $N_{0,4}(0;m)$ one can prove by completely elementary means that the polynomial P_m in (7.80) is also given by

$$P_m(a) = \sum_{j=0}^{m} \binom{2m+1}{2j}(a+1)^j \qquad (9.32)$$
$$\times \sum_{k=0}^{m-j} \binom{m-j}{k}\binom{2(m-k)}{m-k} 2^{-3(m-k)}(a-1)^{m-k-j}.$$

Prove that these two forms coincide.

Compute the value of $P_m(1)$ using both forms of P_m to produce

$$\sum_{k=0}^{m} 2^{-2k}\binom{2k}{k}\binom{2m-k}{m} = \sum_{k=0}^{m} 2^{-2k}\binom{2k}{k}\binom{2m+1}{2k}. \qquad (9.33)$$

Mathematica converts this identity into

$$-\frac{2^{2m+1}\sqrt{\pi}}{\Gamma(-1/2-2m)\,\Gamma(2m+2)} = \frac{2^{2m+1}\,\Gamma(3/2+2m)}{\sqrt{\pi}\,\Gamma(2m+2)}. \qquad (9.34)$$

Use elementary properties of the gamma function to check this directly.

Use the WZ-method to prove (9.33). Check that both sides satisfy the recursion

$$(2m+3)(2m+2)f(m+1) = (4m+5)(4m+3)f(m). \qquad (9.35)$$

It would be interesting to provide a direct proof of (9.33).

3. The study of 2-adic properties of the coefficients appearing in the quartic integral requires an elementary fact of binomial coefficients:

Fact. The central binomial coefficient

$$C_m := \binom{2m}{m} \tag{9.36}$$

is even. Moreover $\frac{1}{2}C_m$ is odd if and only if m is a power of 2. This exercise outlines a new proof. The exact power of 2 that divides m is denoted by $\nu_2(m)$.

a) Prove that $m \geq \nu_2(m!)$. **Hint**. Use (7.89).

b) Check that $\nu_2((2^n)!) = 2^n - 1$. Now let a be the largest integer such that $m = 2^a + b$, then

$$\nu_2(m!) = \nu_2((2^a)!) + \nu_2(b!). \tag{9.37}$$

c) To conclude, check that

$$\nu_2(C_m) = 2^a + b - \nu_2((2^a + b)!) > 1. \tag{9.38}$$

4. The polynomial $2^{2m}P_m(a)$ has positive integer coefficients. It seems to have many interesting properties when considered modulo a prime p. Use a symbolic language to explore them. You can report your results to vhm@math.tulane.edu

5. The integers

$$b_{l,m} \quad = \quad \sum_{k=l}^{m} 2^k \binom{2m-2k}{m-k}\binom{m+k}{m}\binom{k}{l} \tag{9.39}$$

appear in the polynomial P_m as

$$P_m(a) = 2^{-2m}\sum_{l=0}^{m} b_{l,m}a^l. \tag{9.40}$$

The sequence $\{b_{l,m} : 0 \leq l \leq m\}$ is known to be unimodal. See [35] for a proof. We have conjectured that this sequence is logconcave. Use the WZ-method to check that $b_{l,m}$ satisfies

$$b_{l+1,m} = \frac{2m+1}{l+1}b_{l,m} - \frac{(m+l)(m+1-l)}{l(l+1)}b_{l-1,m}. \tag{9.41}$$

Therefore the sequence is logconcave provided

$$(m+l)(m+1-l)b_{l-1,m}^2 + l(l+1)b_{l,m}^2 - l(2m+1)b_{l-1,m}b_{l,m} \geq 0. \tag{9.42}$$

Prove that the left-hand side attains its minimum at $l = m$ with value $2^{2m}m(m+1)\binom{2m}{m}^2$. *We cannot do this part.*

A generalization. The coefficients $\{b_{l,m}\}$ seem to have a property much stronger than logconcavity. Introduce the operator $\mathfrak{L}$ on the space of sequences, via

$$\mathfrak{L}(a_l) := a_l^2 - a_{l-1}a_{l+1}. \tag{9.43}$$

Therefore $\{a_l\}$ is logconcave if $\mathfrak{L}(a_l)$ is nonnegative. We say that $\{a_l\}$ is r-logconcave if $\mathfrak{L}^{(k)}(a_l) \geq 0$ for $0 \leq k \leq r$. The sequence $\{a_l\}$ is ∞-logconcave if it is r-logconcave for every $r \in \mathbb{N}$.

Conjecture 7.1. *For each $m \in \mathbb{N}$, the sequence $\{b_{l,m} : 0 \leq l \leq m\}$ is ∞-logconcave.*

The binomial coefficients is the canonical sequence on which these issues are tested. The solution of the next conjecture should provide guiding principles on how to approach Conjecture 7.1.

Conjecture 7.2. *For $m \in \mathbb{N}$ fixed, the sequence of binomial coefficients $\{\binom{m}{l}\}$ is ∞-logconcave.*

A direct calculation proves the existence of rational functions $R_r(m,l)$ such that

$$\mathfrak{L}^{(r)}\binom{m}{l} = \binom{m}{l}^{2^r} R_r(m,l). \tag{9.44}$$

Moreover R_r satisfy the recurrence

$$R_{r+1}(m,l) = R_r^2(m,l) - \left[\frac{l(m-l)}{(l+1)(m-l+1)}\right]^{2^r} \times R_r(m,l-1)R_r(m,l+1).$$

Therefore we *only* need to prove that $R_r(m,l) \geq 0$. This could be difficult.

6. The integral $N_{0,4}(a;m)$ appeared in a most intriguing edxpansion. The main result of [37] is that

$$\sqrt{a+\sqrt{1+c}} \;=\; \sqrt{a+1} + \frac{1}{\pi\sqrt{2}}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}N_{0,4}(a;k-1)c^k.$$

The proof employ **Ramanujan's Master Theorem**:

Theorem 7.3. *Suppose F has a Taylor expansion around $c = 0$ of the form*

$$F(c) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \varphi(k) c^k. \tag{9.45}$$

Then, the moments of F, defined by

$$M_n = \int_0^{\infty} c^{m-1} F(c)\, dc, \tag{9.46}$$

can be computed via $M_n = \Gamma(m)\varphi(-m)$.

Our expansion comes from the value

$$\int_0^{\infty} \frac{dx}{bx^4 + 2ax^2 + 1} = \frac{\pi}{2\sqrt{2}} \frac{1}{\sqrt{a + \sqrt{b}}}. \tag{9.47}$$

This explains the appearance of the double square root.

It would be interesting to provide a different type of proof. The reader should first use a symbolic language to find a closed-form expression for the coefficients in

$$\sqrt{b + \sqrt{a + \sqrt{1 + c}}} = \sum_{k=0}^{\infty} \rho_k(a, b) c^k. \tag{9.48}$$

The polynomials

$$P_k^*(a, b) = b^k P_k\left(\frac{a}{b}\right) \tag{9.49}$$

will play a role in this expression. The polynomials $P_k^*(a, b)$ are the *homogenization* of $P_k(a)$.

7. Define the function

$$e_l(m) = \frac{(m - l)!l!m!}{(m + l)!2^m} b_{l,m} \tag{9.50}$$

where $b_{l,m}$ are the integers defined in Exercise 5. Use the recurrence for $b_{l,m}$ to check that

$$Q_{m-j}(m) = e_{m-j}\left(\frac{m}{2}\right), \tag{9.51}$$

for j indepedent of m, is a polynomial in m. Confirm that Q_{m-j} is of degree j. For example $Q_{m-3}(m) = m^3 + 2m + 3$. Compute the value of $Q_{m-j}(-1)$ and check that $m+1$ divides $Q_{m-j}(m)$ for j odd.

8. The function

$$\text{jump}_i(j) \quad = \quad \nu_2(i+j) - \nu_2(i), \quad i,\, j \in \mathbb{N}, \tag{9.52}$$

seems to appear in the description of the 2-adic values of the integers $b_{l,m}$ in Exercise 5. The index i is fixed and the word *seems* is to be taken as 'we have a conjecture, that we are not ready to make public'. Use a symbolic language to generate interesting conjectures about $jump_i$.

9. The integral 4.351 of Gradshteyn and Rhyzik [124] states that

$$\int_0^1 (1-x)e^{-x}\,\ln x\,dx = \frac{1-e}{e}. \tag{9.53}$$

Define the function

$$q_n \quad = \quad \int_0^1 x^n e^{-x}\,\ln x\,dx + n!\,(\gamma + \Gamma(0,1))\,, \tag{9.54}$$

where

$$\Gamma(a,z) = \int_z^\infty t^{a-1}e^{-t}\,dt \tag{9.55}$$

is the *incomplete gamma function* and γ is the Euler's constant

$$\gamma = \lim_{n\to\infty} 1 + \frac{1}{2} + \frac{1}{3} + \cdots \frac{1}{n} - \ln n. \tag{9.56}$$

Explore the sequences of positive integers a_n, b_n defined by

$$q_n = a_n - b_n e^{-1}. \tag{9.57}$$

Mathematica yields the first few values

$$\begin{array}{rr} a_1 = 1 & b_1 = 1 \\ a_2 = 3 & b_2 = 4 \\ a_3 = 11 & b_3 = 17 \\ a_4 = 50 & b_4 = 84 \\ a_5 = 274 & b_4 = 485 \end{array}$$

In particular develop recurrences, explore divisibility properties, examine their growth, and so on.

10. Study the integral

$$I_n \quad = \quad \int_0^1 \frac{(1-x)^n\,dx}{(1+x^2)\,\ln x}. \tag{9.58}$$

Use a symbolic language to check that

$$I_n = \ln a_n + b_n \ln \pi + c_n \ln \Gamma(\tfrac{1}{4}) + d_n \ln \Gamma(\tfrac{3}{4}), \tag{9.59}$$

for $a_n \in \mathbb{Q}$ and integers b_n, c_n, d_n. Observe that $c_n = -d_n$ and that $c_{4n+2} = d_{4n+2} = 0$. Explain this.

11. Changes of variables are a powerful tool for the evaluation of definite integral. Sometimes they produce unexpected results. One of them is illustrated in this exercise.

a) Let R be a rational function with real coefficients. Assume that the integral

$$I \quad = \quad \int_{-\infty}^{\infty} R(x)\,dx \tag{9.60}$$

is convergent. Introduce the change of variables

$$y = \frac{x^2-1}{2x} \tag{9.61}$$

to obtain

$$I \quad = \quad \int_{-\infty}^{\infty} R_1(y)dy \tag{9.62}$$

where

$$R_1(y) = R(y-\sqrt{y^2-1})\left(1-\frac{y}{\sqrt{y^2+1}}\right)+R(y+\sqrt{y^2-1})\left(1+\frac{y}{\sqrt{y^2+1}}\right). \tag{9.63}$$

Check that R_1 is again a rational function. The degree of R_1 is at most the degree of R.

b) Discuss the result of part a) in the special case

$$R(x) = \frac{1}{ax^2+bx+c}. \tag{9.64}$$

The convergence of the integral requires that the *discriminant*

$$D(a,b,c) := b^2 - 4ac \tag{9.65}$$

be negative. Conclude that

$$\int_{-\infty}^{\infty} \frac{dx}{ax^2+bx+c} \quad = \quad \int_{-\infty}^{\infty} \frac{dx}{a_1x^2+b_1x+c_1} \tag{9.66}$$

where

$$\begin{aligned} a_1 &= \frac{2ac}{a+c} \\ b_1 &= -\frac{b(a-c)}{a+c} \\ c_1 &= \frac{(a+c)^2-b^2}{a+c}. \end{aligned} \tag{9.67}$$

This is the rational version of the classical *elliptic Landen transformation*: the elliptic integral

$$G(a,b) \quad = \quad \int_0^{\pi/2} \frac{d\theta}{\sqrt{a^2\cos^2\theta + b^2\sin^2\theta}} \tag{9.68}$$

is invariant under the transformation

$$a_1 = \frac{a+b}{2} \text{ and } b_1 = \sqrt{ab}. \tag{9.69}$$

c) Check that the discriminant D is preserved by (9.67), that is, $D(a_1,b_1,c_1) = D(a,b,c)$.

d) Iterate (9.67) to produce a sequence (a_n, b_n, c_n) such that the quadratic integral is preserved. Prove, or convince yourself that, there exists a number L such that

$$a_n \to L, \quad L, \quad b_n \to 0, \quad c_n \to L. \tag{9.70}$$

Use the invariance of the integral to conclude that

$$I \quad = \quad \frac{\pi}{L}. \tag{9.71}$$

This allows you to evaluate the integral I by computing the iterates of (9.67).

e) Use a symbolic language to develop transformations for the integral

$$U_6 \quad := \quad \int_0^{\infty} \frac{cx^4+dx^2+e}{x^6+ax^4+bx^2+1}. \tag{9.72}$$

The denominator of the integrand has been normalized to be monic and have constant term 1. This was the original integral discussed in [38]. The iteration of this transformation appears in [73].

The classical elliptic Landen transformation is at the center of the arithmetic geometric mean (AGM). Many good things have come from here. We expect the same to happen with these rational versions.

The rational Landen transformations were developed originally in [38] and the details of the example discussed in this exercise appear in [173, 174]. The reader will find in [46] and [176] general information about the AGM.

8 Exercises for Chapter 8

1. **Computing values of $\mathcal{W}(a,b,c)$.**

   (a) For $r+s+t=6$ the only terms we need to consider are $\zeta(6), \zeta^2(3)$ since $\zeta(6), \zeta(4)\,\zeta(2)$ and $\zeta^3(2)$ are all rational multiples of π^6. We recovered

   $$\mathcal{W}(3,2,1) = \int_0^1 \frac{\mathrm{Li}_3(x)\ \mathrm{Li}_2(x)}{x}\,dx = \frac{1}{2}\,\zeta^2(3)\,,$$

   consistent with the equation below (8.20). Find all weight six and weight seven evaluations.

   (b) At weight eight one irreducible must be introduced, say $\zeta(6,2)$.

2. Prove the inequalities given in the preparatory Lemma 6.1 used for Hilbert's inequality of Theorem 6.2.

3. Show, as outlined in the text, that the constant in Hardy's inequality given in Theorem 8.10 is best possible.

4. Show that $\sigma_n := \frac{2}{\pi}\int_0^\infty \mathrm{sinc}^n$ is strictly decreasing. Aliev [6] uses this among other tools to show that for any nonzero a in Z^n, there exist linearly independent vectors (x_k) in Z^n with $\langle a, x_k\rangle = 0$ such that

   $$\|x_1\|\|x_2\|\cdots\|x_n\| < \frac{\|a\|}{\sigma_n}.$$

5. **Ramanujan's AGM continued fraction** is the object

$$\mathcal{R}_\eta(a,b) = \cfrac{a}{\eta + \cfrac{b^2}{\eta + \cfrac{4a^2}{\eta + \cfrac{9b^2}{\eta + \ddots}}}},$$

certainly valid for $a, b, \eta > 0$.

Our discussion of $\mathcal{R}_\eta$ follows [56, 57, 45]. It enjoys attractive algebraic properties such as a striking arithmetic-geometric mean relation

$$\mathcal{R}_\eta\left(\frac{a+b}{2}, \sqrt{ab}\right) = \frac{\mathcal{R}_\eta(a,b) + \mathcal{R}_\eta(b,a)}{2} \tag{9.73}$$

and has elegant links with elliptic-function theory [45]. We may restrict attention to $\eta = 1$, since $\mathcal{R}_\eta(a,b) = \mathcal{R}_1(a/\eta, b/\eta)$. Now $\mathcal{R}_1$ is far from easy to compute naively.

(a) Indeed, inspection of $\mathcal{R}_1$ yields the *reduced continued fraction* form:

$$\begin{aligned}\mathcal{R}_1(a,b) &= \frac{a}{[c_0; c_1, c_2, c_3, \dots]} \\ &:= \cfrac{c}{c_0 + \cfrac{1}{c_1 + \cfrac{1}{c_2 + \cfrac{1}{c_3 + \ddots}}}}\end{aligned}$$

where the q_i are all positive real numbers.

(b) Show that c_n satisfies:

$$c_n = \frac{n!^2}{(n/2)!^4}\, 4^{-n}\, \frac{b^n}{a^n} \quad \sim \quad \frac{2}{\pi\, \mathbf{n}}\, \frac{b^n}{a^n}.$$

for even n, while for odd n

$$c_n = \frac{((n-1)/2!)^4}{n!^2}\, 4^{n-1}\, \frac{a^{n-1}}{b^{n+1}} \quad \sim \quad \frac{\pi}{2\, ab\, \mathbf{n}}\, \frac{a^n}{b^n}.$$

This representation leads immediately to a proof that for any positive real pair a, b the fraction $\mathcal{R}_1(a,b)$ converges. Indeed, by the *Seidel-Stern theorem*, [56, 57], a reduced continued fraction

converges iff $\sum c_i$ diverges. In our case, such divergence is evident for every choice of real $a, b > 0$. Note for $a = b$, divergence of $\sum c_i$ is only *logarithmic* hence and somewhat surprisingly this is the hardest case to compute.

(c) There are, however, beautiful numerical series involving the *complete elliptic integral*

$$K(k) = \int_0^{\pi/2} \frac{1}{\sqrt{1 - k^2 \sin^2 \theta}}\, d\theta.$$

Below we write $\mathrm{K} := \mathrm{K}(k),\ \mathrm{K}' := \mathrm{K}(k')$ with $k' := \sqrt{1 - k^2}$.
For $0 < b < a$ and $k := b/a$ we have

$$\mathcal{R}_1(a, b) = \frac{\pi a\, \mathrm{K}}{2} \sum_{n \in Z} \frac{\operatorname{sech}\left(n\pi \frac{\mathrm{K}'}{\mathrm{K}}\right)}{\mathrm{K}^2 + \pi^2 a^2 n^2}. \tag{9.74}$$

Correspondingly, for $0 < a < b$ and $k := a/b$ we have

$$\mathcal{R}_1(a, b) = 2\pi b\, \mathrm{K} \sum_{n \in Z, \mathrm{odd}} \frac{\operatorname{sech}\left(n\pi \frac{\mathrm{K}'}{2\mathrm{K}}\right)}{4\mathrm{K}^2 + \pi^2 b^2 n^2}. \tag{9.75}$$

Since K is fast computable—and well implemented in *Maple* and *Mathematica*—for a and b not too close these formulae are very effective. Moreover, Poisson transformation, for $0 < b < a$, yields

$$\begin{aligned} \mathcal{R}_1(a, b) &= \mathcal{R}_1\left(\frac{\pi a}{2\mathrm{K}'}, \frac{\pi a}{2\mathrm{K}'}\right) + \frac{\pi}{\cos \frac{\mathrm{K}'}{a}} \frac{1}{e^{2\mathrm{K}/a} - 1} \\ &+ 8\pi a\, \mathrm{K}' \sum_{0 < d \in Z, \mathrm{odd}} \frac{(-1)^{(d-1)/2}}{4\mathrm{K}'^2 - \pi^2 d^2 a^2} \frac{1}{e^{\pi d\, \mathrm{K}/\mathrm{K}'} - 1} \end{aligned}$$

Thus, once we obtain an effective formula for $\mathcal{R}(a) := \mathcal{R}_1(a, a)$ we can compute $\mathcal{R}_1$ for all $a, b > 0$ since Ramanujan's identity (9.73) allows us to compute only with $a > b > 0$.
This is the gist of the next exercise:

6. **Closed forms for $\mathcal{R}$.** By viewing (9.74) as a Riemann sum as $b \to a^-$, prove that for all $a > 0$

$$\mathcal{R}(a) = \int_0^\infty \frac{\operatorname{sech}\left(\frac{\pi}{2}\frac{x}{a}\right)}{1 + x^2}\, dx$$

(a) Derive that

$$\begin{aligned}\mathcal{R}(a) &= 2\,a\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{1+(2\,k-1)\,a}\\ &= \frac{1}{2}\left(\psi\left(\frac{3}{4}+\frac{1}{4a}\right)-\psi\left(\frac{1}{4}+\frac{1}{4a}\right)\right)\\ &= \frac{2a}{1+a}\,{}_2\mathrm{F}_1\left(\frac{1}{2a}+\frac{1}{2},1;\frac{1}{2a}+\frac{3}{2};-1\right)\\ &= 2\int_0^1\frac{t^{1/a}}{1+t^2}\,dt.\end{aligned}$$

(b) Conclude that

$$\mathcal{R}(a)=\int_0^{\infty}e^{-x/a}\,\mathrm{sech}(x)\,dx.$$

(c) Show that

$$\mathcal{R}(a)=\frac{2a}{1+a}-\mathcal{R}\left(\frac{a}{1+2a}\right).$$

(d) Determine the value of $\mathcal{R}(n)$ for $n=1,2,3,\ldots$ or even for n rational. For example, $\mathcal{R}(1)=\log 2$, $\mathcal{R}(3)=\pi/\sqrt{3}-\log 2$ and $\mathcal{R}(3/2)=\pi+\sqrt{3}\,\log(2-\sqrt{3})$. No closed form is known for $\mathcal{R}_{\infty}(a,b)$ when $a\neq b$.

7. The values of $\zeta(2,1),\zeta(2,\overline{1}),\zeta(\overline{2},1)$, and $\zeta(\overline{2},\overline{1})$ all reduce to sums of products of one dimensional (alternating) zeta functions. Thus, the only possible basis elements are $\zeta(3)$ and $\zeta(2)\log(2)$. Hence, recover

$$\zeta(2,1)=\zeta(3)\quad \zeta(\overline{2},1)=\frac{1}{8}\,\zeta(3)$$

$$\zeta(2,\overline{1})=\zeta(3)-\frac{3}{2}\,\zeta(2)\log(2),$$

$$\zeta(\overline{2},\overline{1})=\frac{3}{2}\,\zeta(2)\log(2)-\frac{13}{8}\,\zeta(3).$$

8. **Uses of the Dilogarithm and Trilogarithm.** We describe material taken from [48].

(a) Consider the power series

$$J(x):=\zeta_x(2,1)=\sum_{n>k>0}\frac{x^n}{n^2k},\qquad 0\le x\le 1.$$

and show that

$$J(x) = \int_0^x \frac{dt}{t} \int_0^t \frac{du}{1-u} \int_0^v \frac{dv}{1-v} = \int_0^x \frac{\log^2(1-t)}{2t}\, dt.$$

(b) *Maple* readily evaluates

$$\int_0^x \frac{\log^2(1-t)}{2t}\, dt = \zeta(3) + \frac{1}{2}\log^2(1-x)\log(x) + \log(1-x)\mathrm{Li}_2(1-x) - \mathrm{Li}_3(1-x) \tag{9.76}$$

where $\mathrm{Li}_s(x) := \sum_{n=1}^{\infty} x^n/n^s$ is the classical *polylogarithm.*

(c) Verify (9.76) by differentiating both sides by hand, and checking that (9.76) holds as $x \to 0+$. Thus, deduce

$$J(x) = \zeta(3) + \frac{1}{2}\log^2(1-x)\log(x) + \log(1-x)\mathrm{Li}_2(1-x) - \mathrm{Li}_3(1-x).$$

Let $x \to 1-$ to prove (**??**).

(d) In *Ramanujan's Notebooks*, we also find that

$$J(-z) + J(-1/z) = -\frac{1}{6}\log^3 z - \mathrm{Li}_2(-z)\log z + \mathrm{Li}_3(-z) + \zeta(3) \tag{9.77}$$

and

$$J(1-z) = \frac{1}{2}\log^2 z \log(z-1) - \frac{1}{3}\log^3 z - \mathrm{Li}_2(1/z)\log z - \mathrm{Li}_3(1/z) + \zeta(3). \tag{9.78}$$

(e) Put $z = 1$ in (9.77) and employ the well-known *dilogarithm evaluation*

$$\mathrm{Li}_2(-1) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} = -\frac{\pi^2}{12}$$

to obtain (9.10).

(f) Put $z = 2$ in (9.78) and employ Euler's *dilogarithm evaluation*

$$\mathrm{Li}_2\left(\frac{1}{2}\right) = \sum_{n=1}^{\infty} \frac{1}{n^2\, 2^n} = \frac{\pi^2}{12} - \frac{1}{2}\log^2 2$$

along with Landen's *trilogarithm evaluation* (see Lewin)

$$\mathrm{Li}_3\left(\frac{1}{2}\right) = \sum_{n=1}^{\infty} \frac{1}{n^3\, 2^n} = \frac{7}{8}\zeta(3) - \frac{\pi^2}{12}\log 2 + \frac{1}{6}\log^3 2$$

to obtain (9.10) yet again.

(g) These evaluations follow respectively from the functional equations

$$\mathrm{Li}_2(x) + \mathrm{Li}_2(1-x) \quad = \quad \zeta(2) + \log(x)\log(1-x) \quad (9.79)$$

and

$$\mathrm{Li}_3(x) + \mathrm{Li}_3(1-x) + \mathrm{Li}_3(x/(x-1)) \qquad (9.80)$$

$$= \zeta(3) - \frac{1}{2}\,\log(x)\log^2(1-x) + \zeta(2)\,\log(1-x) + \frac{1}{6}\,\log^3(1-x)$$

and prove both of these by symbolically determining that they have equal derivatives on both sides.

(h) Once the component functions in (9.79), (9.80), or (9.13) are known, the coefficients can be deduced by computing each term to high precision with a common transcendental value of x and employing a linear relations finding algorithm. This is also an excellent way to 'error-correct'

9. **AMM Problem 11115, November 2004.** Evaluate the limit of

$$E_m := \left(\sum_{k=1}^{m} \frac{1}{k}\right)^2 - \sum_{k=1}^{m} \frac{\sum_{j=1}^{\max(k,m-k)} 1/j}{k},$$

as m goes to ∞.

(a) First, show the limit exists.
Hint. Exhibit E_m as a double Riemann sum for the integral following:

$$E_m = \frac{1}{m} \sum_{k=1}^{m} \frac{\frac{1}{m}\sum_{1+\max(k,m-k)}^{m} \frac{1}{j/m}}{k/m}$$

which converges to

$$\int_0^1 \frac{1}{x} \int_{(1-x)\vee x}^{1} \frac{1}{y}\, dy\, dx = -\int_0^{1/2} \frac{\ln(1-x)}{(1-x)} \frac{dx}{x} = 0.8224670336\ldots.$$

We note that the 'identify' function in *Maple* will identify the constant from the numeric value of the integral: $0.8224670336\ldots$. This is much more difficult from the original sum where $10,000$ terms only provides $0.82236685\ldots$.

(b)
$$-\int_0^{1/2} \frac{\ln(1-x)}{(1-x)} \frac{dx}{x} = \frac{\pi^2}{12},$$
as is known to both Maple and Mathematica and can be obtained from the value of dilogarithm at 1/2, or otherwise. Indeed, with the assumption that $0 < t < 1$, *Maple* returns
$$-\int_0^{t} \frac{\ln(1-x)}{(1-x)} \frac{dx}{x} = \frac{1}{2}\ln^2(1-t) + \operatorname{dilog}(1-t),$$
as differentiation yet again confirms. (Note *Maple* uses $\operatorname{dilog}(x) = \operatorname{Li}_2(x)$.) Now use (f) of the previous exercise.

(c) Evaluate
$$\sum_{n=1}^{\infty} \frac{H_n}{n\,2^n} \text{ and } \sum_{n=1}^{\infty} \frac{H_{n-1}}{n\,2^n},$$
with $H_n := 1 + 1/2 + \cdots 1/n$.

10. **Some Ising Integrals.** Define
$$I_n := \int_0^{\infty} \frac{\left(\prod_{n\geq k>j\geq 1} \frac{u_k-u_j}{u_k+u_j}\right)^2}{(\sum_{j=1}^n u_j + 1/u_j)^2} \frac{du_1}{u_1} \cdots \frac{du_n}{u_n}.$$

(a) Use the symmetry of the I_n integrands to reduce to the simplex where $u_k > u_{k+1}$ and then use the change of variables $u_k := \prod_{i=1}^k t_i$ to represent the integral as
$$I_n = 2\int_0^1 \iota_n\, dt_2\, dt_3 \cdots dt_n$$
with
$$\iota_n(t_2, t_3 \ldots, t_n) := n! \left(\prod_{n\geq k>j\geq 1} \frac{u_k/u_j - 1}{u_k/u_j + 1}\right)^2 \frac{1}{(1+\sum_{k=2}^n w_k)(1+\sum_{k=2}^n v_k)}$$
where
$$w_k := \prod_{i=2}^{k} t_i, \quad v_k := \prod_{i=k}^{n} t_i,$$
since the integral in $t_1 = u_1$ is now easy to obtain.

(b) Determine that $\iota_1 = 1$ and $\iota_2 = 2!\,(t_2-1)^2/(t_2+1)^4$ while
$$\iota_3 = 3!\,\frac{(t_2-1)^2(t_2t_3-1)^2(t_3-1)^2}{(t_2+1)^2(t_2t_3+1)^2(t_3+1)^2(t_2+t_2t_3+1)(t_2t_3+t_3+1)}$$

(c) Hence,

$$I_2 = 4 \int_0^1 \frac{(x-1)^2}{(x+1)^4}\, dx = \frac{1}{6}$$

$$I_3 = 24 \int_0^1 \int_0^x \iota_3(x,y)\, dx\, dy,$$

which integral *Maple* can reduce to

$$1 + \zeta(2) - \frac{27}{8} L_{-3}(2),$$

(most easily from the same integral over the positive orthant which is 6 times the integral on the square!) The code:

```
> p:=(x-1)^2*(x-y)^2*(y-1)^2/(x+1)^2/(x+y)^2/(y+1)^2/(1+y+x)/(y+x+x*y):
> Int(Int(p, x = 0 .. infinity),y = 0 .. infinity): evalc(value(%)));
```

returns
$18\,i\,\mathrm{dilog}\left(1/2 - 1/2\,i\sqrt{3}\right)\sqrt{3} - 18\,i\,\mathrm{dilog}\left(1/2 + 1/2\,i\sqrt{3}\right)\sqrt{3} + 24 + 4\,\pi^2$.

The value of I_4 is known to be $-1/6 + 4\pi^2/9 + 7/2\,\zeta(3)$ but the status of higher values is open, [220].

11. **Levy Constants.** [114] Let p_n/q_n be the n-th convergent of a number $\alpha > 0$. Levy showed in 1929 that $\lambda(\alpha) := \lim_{n\to\infty} \frac{\log q_n}{n}$ exists and is a constant, $\overline{\lambda}$, for almost all irrationals. It is also known that $\lambda(\beta)$ exists for all quadratic irrationals β, and is dense in $[\log G, \infty]$ where G is the golden mean [232].

(a) Numerically explore the constant and attempt to identify $\overline{\lambda}$ numerically and then symbolically.

(b) Explore whether either π or e appears to behave 'normally'.

(c) For a quadratic β with purely-periodic part fraction $[b_1, b_2, \ldots b_N]$ show that

$$\lambda(\beta) = \frac{\log \sigma(B_N)}{N}$$

where σ denotes the spectral radius and B_N is the matrix

$$\prod_{k=1}^{N} \begin{bmatrix} b_k & 1 \\ 1 & 0 \end{bmatrix}.$$

Hint. $\overline{\lambda} \approx 1.1865691104156$.

9 Additional Exercises

We collect various additional examples and exercises. We first solve some recent *Monthly Problems*. In each case a few lines of computer algebra code either provides the solution or at least confirms it.

1. **AMM Problem 11148, April 2005**. Show that

$$\int_0^\infty \frac{(x^8 - 4x^6 + 9x^4 - 5x^2 + 1)\,dx}{x^{12} - 10x^{10} + 37x^8 - 42x^6 + 26x^4 - 8x^2 + 1} = \frac{\pi}{2}.$$

Solution.

(a) For convenience, first prove the following lemma.

Lemma 9.1. *Let C be a positively-oriented simple closed curve containing n distinct complex numbers $a_1, a_2, \ldots, a_n$. Let $Q(z) := \prod_{k=1}^n (z - a_k)$. Then*

$$\int_C \frac{z^m}{Q(z)}\,dz = \begin{cases} 0, & m = 0, 1, \ldots, n-2 \\ 2\pi i, & m = n-1 \end{cases}$$

(b) Let I denote the integral. A partial fraction decomposition yields

$$\begin{aligned} I &= \int_0^\infty \left[\frac{(1+x)^2\,dx}{2(x^6 + 4x^5 + 3x^4 + 4x^3 - 2x^2 - 2x + 1)} \right. \\ &\qquad \left. + \frac{(1-x)^2\,dx}{2(x^6 - 4x^5 + 3x^4 - 4x^3 - 2x^2 + 2x + 1)} \right] \qquad (9.81) \\ &= \int_{-\infty}^\infty \frac{(1-x)^2\,dx}{2(x^6 - 4x^5 + 3x^4 - 4x^3 - 2x^2 + 2x + 1)} \\ &= \int_{-\infty}^\infty \frac{y^2\,dy}{2(y^6 + 2y^5 - 2y^4 - 4y^3 + 3y^2 + 4y + 1)} \\ &= \int_{-\infty}^\infty \left[\frac{(-i - u + iu^2)\,du}{4(u^3 + u^2 + iu^2 - 2u + iu - 1)} + \frac{(i - u - iu^2)\,du}{4(u^3 + u^2 - iu^2 - 2u - iu - 1)} \right]. \end{aligned}$$

(c) Check that the roots of $u^3 + u^2 + iu^2 - 2u + iu - 1$ are in the lower half-plane, while those of $u^3 + u^2 - iu^2 - 2u - iu - 1$ (its complex conjugate) are in the upper half-plane. Denote these latter roots as $\{\alpha_k\}$, $k = 1, \ldots, 3$. Apply the lemma to (9.81) over a standard half-circle contour integral to get

$$I = 2\pi i \left(-\frac{i}{4} \right) = \frac{\pi}{2}.$$

All of this can be done entirely symbolically in a computer algebra system such as *Maple*—indeed this is how we discovered the basis for this proof. A generalization is straightforward.

2. Prove that

Theorem 9.2. *Let* $q(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$, $c_k \in \mathbb{C}$, *be a polynomial whose roots all lie in the upper half-plane, and* $p(x) = -i\,x^{n-1} + d_{n-2}x^{n-2} + \cdots + d_1 x + d_0$, $d_k \in \mathbb{C}$. *Then*

$$\int_{-\infty}^{\infty} \mathrm{Re}\left(\frac{p(x)}{q(x)}\right) = \pi.$$

There are other like generalizations.

3. **AMM Problem 11113, November 2004.** Evaluate

$$I_k(a,b) := \int_0^\infty \int_0^\infty \frac{e^{-k\sqrt{x^2+y^2}}\sin(ax)\sin(by)}{\sqrt{x^2+y^2}\,xy}\,dx\,dy$$

in closed form for $a, b, k > 0$. *Solution.* Since $I_k(a,b) = k\,I_1(a/k, b/k)$, as a simple change of variables shows, we study only $I(a,b) := I_1(a,b)$. The desired formula is

$$\begin{aligned} I(a,b) &= \frac{\pi\,b}{4}\ln\left(b^2+1\right) + \frac{\pi\,a}{4}\ln\left(a^2+1\right) \\ &- \frac{\pi\,b}{2}\ln\left(\sqrt{a^2+1+b^2}-a\right) - \frac{\pi\,a}{2}\ln\left(\sqrt{a^2+1+b^2}-b\right) \\ &- \frac{\pi}{2}\arctan\left(\frac{ab}{\sqrt{a^2+1+b^2}}\right). \end{aligned}$$

To prove this let

$$K(a,b) := \frac{d^2 I(a,b)}{da\,db}.$$

Then a change to polar variables produces

$$K(a,b) = \frac{\pi}{2\sqrt{1+a^2+b^2}}.$$

Finally a careful integration produces the formula in (9.82). Having found (9.82) it is easier to confirm it by (symbolic) differentiation on observing that each term on the right vanishes when $a = 0$ or $b = 0$.

4. **Solution to AMM Problem 11152, May 2005** Evaluate

$$\int_0^1 \frac{\log\left(\cos(\pi x/2)\right)}{x(1+x)}\,dx.$$

Solution. The answer is $C := \log^2 2/2 - \log(2)\log(\pi)$. This value may be obtained immediately by placing a *Maple*-induced 50-digit approximation of the integral into the "Smart-Lookup" option of the *Inverse Symbolic Calculator*.

Let I denote the integral. We again use the *dilogarithm* function $\mathrm{Li}_2(x) := -\int_0^x \log(1-z)/z\,dz$, though only its simplest property will be needed. Using the infinite product expansions for $\sin(x)$ and $\cos(x)$, we have

$$\begin{aligned}
I &= \int_0^1 \log(\cos(\pi x/2))/x\,dx - \int_1^2 \log(\sin(\pi x/2))/x\,dx \\
&= -\int_1^2 \log(\pi x/2)/x\,dx \\
&+ \sum_{k=1}^{\infty}\left[\int_0^1 \frac{1}{x}\log\left(1-\frac{x^2}{(2k-1)^2}\right)dx - \int_1^2 \frac{1}{x}\log\left(1-\frac{x^2}{4k^2}\right)dx\right] \\
&= C + \sum_{k=1}^{\infty}\left[\int_0^1 \frac{\log\left(1-\frac{x}{2k-1}\right)}{x}\,dx + \int_0^1 \frac{\log\left(1+\frac{x}{2k-1}\right)}{x}\,dx\right. \\
&\qquad \left. - \int_1^2 \frac{\log\left(1-\frac{x}{2k}\right)}{x}\,dx - \int_1^2 \frac{\log\left(1+\frac{x}{2k}\right)}{x}\,dx\right] \\
&= C + \sum_{k=1}^{\infty}\left[-\mathrm{Li}_2\left(\frac{1}{2k-1}\right) - \mathrm{Li}_2\left(-\frac{1}{2k-1}\right)\right. \\
&\qquad \left. +\mathrm{Li}_2\left(\frac{1}{k}\right) - \mathrm{Li}_2\left(\frac{1}{2k}\right) + \mathrm{Li}_2\left(-\frac{1}{k}\right) - \mathrm{Li}_2\left(-\frac{1}{2k}\right)\right] \\
&= C.
\end{aligned}$$

The last equality, a telescoping collapse, is justified since $\mathrm{Li}_2(x) = x + O(x^2)$ implies the sum is absolutely convergent.

Alternatively, one may rewrite the telescoping argument as a change of variables of the form

$$\int_1^2 \log(\cos(\pi x/2^N))/x\,dx = \int_{2^{-N}}^{2^{1-N}} \log(\cos(\pi t))/t\,dt.$$

5. **AMM Problem 11001, August-September 2004.** Evaluate

$$\int_0^{\infty} a\arctan\left(\frac{b}{\sqrt{a^2+x^2}}\right)\frac{1}{\sqrt{a^2+x^2}}\,dx \tag{9.82}$$

for $a, b > 0$.

Solution. An answer is

$$\int_0^\infty a \arctan\left(\frac{b}{\sqrt{a^2+x^2}}\right)\frac{1}{\sqrt{a^2+x^2}}\,dx = \frac{a\pi}{2}\left\{\ln\left(b+\sqrt{a^2+b^2}\right)-\ln(a)\right\}.$$

Indeed, making the two substitutions $x \to k\,t \to k\tan(s)$, where $k := a/b$, in (9.82) shows it suffices to establish that

$$\int_0^{1/2\,\pi} \sec(s)\,\arctan\left(\frac{k}{\sec(s)}\right)\,ds = \frac{\pi}{2}\,\ln\left(k+\sqrt{k^2+1}\right). \quad (9.83)$$

In turn, differentiation of (9.83) with respect to k, yields

$$\frac{\pi}{\sqrt{1+k^2}} = \frac{\pi}{\sqrt{1+k^2}}.$$

Since both sides of (9.83) are zero when $k = 0$, an appeal to the fundamental theorem of calculus ends the argument.

6. **AMM Problem 11159, May 2005.** For $|a| < \pi/2$, evaluate in closed form

$$I(a) := \int_0^{\pi/2}\int_0^{\pi/2} \frac{\cos s\,ds\,dt}{\cos(a\cos s\cos t)}.$$

Hint. The series

$$\sec x = \sum_{n=0}^{\infty} \frac{(-1)^n E_{2n}}{(2n)!}\,x^{2n},$$

wherein E_{2n} are the even *Euler numbers*, converges for $|x| < \pi/2$.

[The answer is $\pi/(2a)\,\log(\sec a + \tan a)$.]

Solution. (Due to David Bradley) Expand secant as a Maclaurin series in even powers of a:

$$\sec(a\cos\psi\cos\varphi) = \sum_{n=0}^{\infty} \frac{(-1)^n E_{2n}}{(2n)!}\,a^{2n}\cos^{2n}\psi\cos^{2n}\varphi. \quad (9.84)$$

Let $0 < r < 1$. We temporarily strengthen the restriction on a to $|a| \le r\pi/2$. The coefficients E_{2n} in (9.84) are the Euler numbers $1, -1, 5, -61, \ldots$ and satisfy an inequality of the form

$$\frac{|E_{2n}|}{(2n)!} < C\left(\frac{2}{\pi}\right)^{2n+1}, \qquad n \ge 0,$$

for C a constant independent of n (one can take $C = 2$) Thus, the series in (9.84) converges absolutely and uniformly in $[-r\pi/2, r\pi/2]$.

Substitute into the integral and interchange summation and integration (as justified by uniform convergence) to get

$$I(a) = \sum_{n=0}^{\infty} \frac{(-1)^n E_{2n}}{(2n)!}\, a^{2n} \int_0^{\pi/2} \cos^{2n+1}\psi \, d\psi \int_0^{\pi/2} \cos^{2n}\varphi \, d\varphi.$$

In light of the known evaluations [124, 3.621]

$$\int_0^{\pi/2} \cos^{2n+1}\psi \, d\psi = \frac{2 \cdot 4 \cdots (2n)}{1 \cdot 3 \cdots (2n+1)}$$

and

$$\int_0^{\pi/2} \cos^{2n}\varphi \, d\varphi = \frac{1 \cdot 3 \cdots (2n-1)}{2 \cdot 4 \cdots (2n)} \frac{\pi}{2},$$

we have

$$I(a) = \frac{\pi}{2} \sum_{n=0}^{\infty} \frac{(-1)^n E_{2n}}{(2n)!} \frac{a^{2n}}{2n+1}.$$

Clearly $I(0) = \pi/2$. If $0 < |a| \le r\pi/2$, then interchange sum and integral again:

$$I(a) = \frac{\pi}{2a} \sum_{n=0}^{\infty} \frac{(-1)^n E_{2n}}{(2n)!} \int_0^a t^{2n}\, dt = \frac{\pi}{2a} \int_0^a \sec t \, dt = \frac{\pi}{2a} \log(\sec a + \tan a). \tag{9.85}$$

Since $0 < r < 1$ is arbitrary, (9.85) is actually valid for $0 < |a| < \pi/2$. □

What is perplexing here is that we need to know secant has a power series on $(-\pi/2, \pi/2)$ but nothing else!

7. **AMM Problem 11164, May 2005.** Evaluate

$$\sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} \sum_{i=1}^{k} \frac{\sum_{j=1}^{i} 1/j}{i}$$

for $n = 1, 2, \ldots$.

Hint. Sum the first few terms.

8. Evaluate $A_n B_n$ for $n = 1, 2, \ldots$ when

$$A_n := \int_{-\infty}^{\infty} \frac{\sin(\pi x)}{\pi x} \prod_{k=1}^{n} \left(1 - \frac{x^2}{k^2}\right)^{-1} dx$$

$$B_n := \int_{-\infty}^{\infty} \frac{\cos(\pi x)}{\pi} \prod_{k=1}^{n} \left(1 - \frac{x^2}{(k-1/2)^2}\right)^{-1} dx$$

Hint. Evaluate the first few terms.

9. **Szegő curves.** While the exponential has no zeroes, its partial Taylor series at zero do. We wish to explore the zeroes of such partial sums of the exponential: $\{s_n(z)\}$ where $s_n(z) := \sum_{k=0}^{n} z^k/k!$.

   (a) Begin by computing and plotting the zeroes of s_n for various values of n as in Figure 9.3. Remarkably, the shape of the curves appears very stable and regular. Indeed, doing the same for the normalized curves as in Figure 9.4 strongly suggests there is a limit curve.

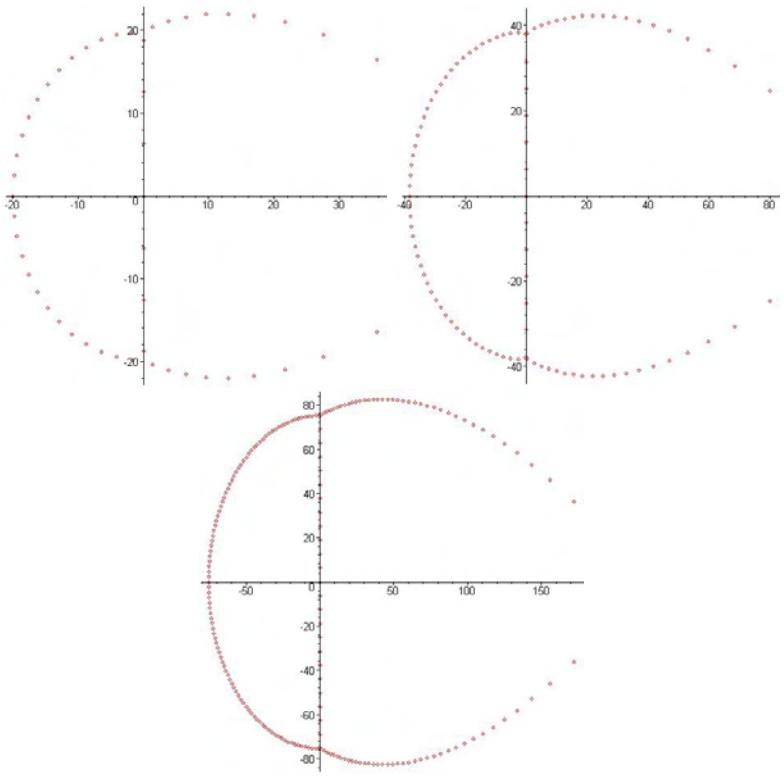


Figure 9.3. The zeroes of s_n for $n = 50, 100, 200$

Maple code to draw such zeros is simple:

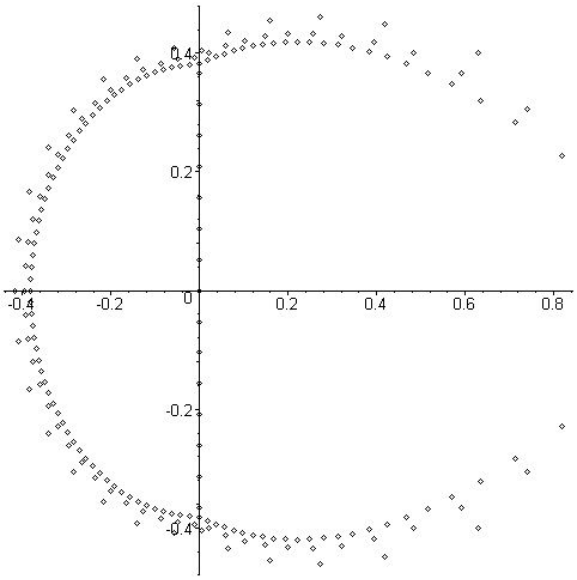


Figure 9.4. The zeroes of the normalized s_n for $n = 30, 60, 120$

```
> draw:=(l,N)->complexplot([fsolve(l(N),x,complex)],
                           style=point,color=black):
 > sn:=N->sum((N*x)^n/n!,n=1..N):draw(sn(50));
```

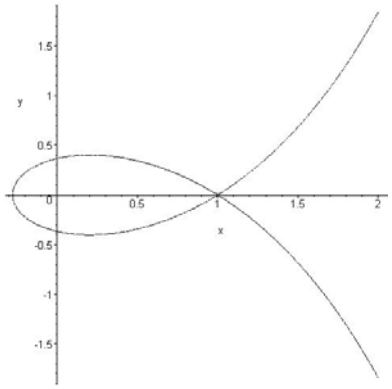


Figure 9.5. The Szegő curve for $\exp(z)$

(b) In 1924, Szegő showed (with no such tools) that the zeros of the *normalized partial sums*, $s_n(nz)$, of e^z tended to the curve, now

called the *Szegő curve* S, shown in Figure 9.5, where

$$S := \left\{ z \in \mathbf{C} : |ze^{1-z}| = 1 \text{ and } |z| \leq 1 \right\}.$$

Via modern weighted potential theory, Pritzger and Varga [201] recover these zero distribution results of Szegő, along with an asymptotic formula for the weighted partial sums $\{e^{-nz}s_n(nz)\}_{n=0}^{\infty}$. They go on to show that $G := \text{Int } S$ is the largest universal domain such that the weighted polynomials $e^{-nz}P_n(z)$ are dense in the set of functions analytic in G. More generally, they show that if $f(z)$ is analytic in G and continuous on $\overline{G}$ with $f(1) = 0$, then there is a sequence of polynomials $\{P_n(z)\}_{n=0}^{\infty}$, with $\deg P_n \leq n$, such that

$$\lim_{n\to\infty} \|e^{-nz}P_n(z) - f(z)\|_{\overline{G}} = 0,$$

where $\|\cdot\|_{\overline{G}}$ denotes the supremum norm on $\overline{G}$. The reader may wish to follow up by consulting Zemyan's recent article [**?**].

10. **Clausen's function and the figure-eight knot complement volume.** As discussed in [44], the volume of the figure-eight knot complement, which we take to be defined by the *log sine integral*

$$V = -2\int_0^{\pi/3} \log\left(2\sin\left(\frac{t}{2}\right)\right)\, dt, \tag{9.86}$$

is also given by ternary BBP formula

$$\frac{3\sqrt{3}}{2}\,V = \sum_{k=0}^{\infty}\left(\frac{-1}{27}\right)^k\left(\frac{9}{(6k+1)^2} - \frac{9}{(6k+2)^2} - \frac{12}{(6k+3)^2} - \frac{3}{(6k+4)^2} + \frac{1}{(6k+5)^2}\right). \tag{9.87}$$

Recall that the *Clausen function* is

$$\mathrm{Cl}_2(\theta) = \sum_{n>0} \frac{\sin(n\theta)}{n^2}, \tag{9.88}$$

which satisfies $\mathrm{Cl}_2(\pi/2) = G$. Requisite details about Clausen's function are to be found in Lewin [167].

(a) Show that integration of (9.86) gives

$$\begin{aligned} V &= i\left\{\mathrm{Li}_2\left(e^{-i\pi/3}\right) - \mathrm{Li}_2\left(e^{i\pi/3}\right)\right\} \\ &= 2\,\mathrm{Im}\mathrm{Li}_2\left(\frac{1+i\sqrt{3}}{2}\right) = 2\,\mathrm{Cl}_2\left(\frac{\pi}{3}\right) = 3\,\mathrm{Cl}_2\left(\frac{2\,\pi}{3}\right). \end{aligned}$$

(b) Show that a hypergeometric equivalent formulation of (9.87) is

$$\begin{aligned}\frac{V}{\sqrt{3}} &\stackrel{?}{=} 2\,F\left(\frac{1}{6},\frac{1}{6},1;\frac{7}{6},\frac{7}{6};\frac{-1}{27}\right)-\frac{1}{2}\left(\frac{1}{3},\frac{1}{3},1;\frac{4}{3},\frac{4}{3};\frac{-1}{27}\right)\\ &- \frac{8}{27}F\left(\frac{1}{2},\frac{1}{2},1;\frac{3}{2}\frac{3}{2};\frac{-1}{27}\right)-\frac{1}{24}F\left(\frac{2}{3},\frac{2}{3},1;\frac{5}{3},\frac{5}{3};\frac{-1}{27}\right)\\ &+ \frac{2}{225}F\left(\frac{5}{6},\frac{5}{6},1;\frac{11}{6},\frac{11}{6};\frac{-1}{27}\right).\end{aligned}$$

With some effort this is expressible in dilogarithms leading to

$$\begin{aligned}V &\stackrel{?}{=} \operatorname{Im}\left\{4\operatorname{Li}_2\left(\frac{i\sqrt{3}}{3}\right)-\frac{8}{3}\operatorname{Li}_2\left(\frac{i\sqrt{3}}{9}\right)\right.\\ &\qquad \left.+\operatorname{Li}_2\left(\frac{1}{2}-\frac{i\sqrt{3}}{6}\right)+8\operatorname{Li}_2\left(-\frac{1}{2}+\frac{i\sqrt{3}}{6}\right)\right\}.\end{aligned}$$

(c) Now, Lewin in Equation (5.5) of [167] gives

$$\operatorname{Im}\operatorname{Li}_2\left(re^{i\theta}\right)=\omega\log(r)+\frac{1}{2}\operatorname{Cl}_2\left(2\omega\right)-\frac{1}{2}\operatorname{Cl}_2\left(2\omega+2\theta\right)+\frac{1}{2}\operatorname{Cl}_2\left(2\theta\right),$$

where $\omega=\arctan(r\sin\theta/(1-r\cos\theta))$. Using this, a proof that (9.87) holds is reduced to showing that, with $\alpha=\arctan(\sqrt{3}/9)$,

$$4\operatorname{Cl}_2\left(\frac{\pi}{3}\right)=2\operatorname{Cl}_2\left(2\,\alpha\right)+\operatorname{Cl}_2\left(\pi+2\,\alpha\right)-3\operatorname{Cl}_2\left(\frac{5}{3}\pi+2\,\alpha\right),$$

which is true by applying the two variable identities for Clausen's function given in Equations (4.61) and (4.63) of [167], with $\theta=\pi/3$.

11. **The origin of formula (1.1).** A much harder unproven identity is

$$\frac{7\sqrt{7}}{4}\,L_{-7}\left(2\right)=3\operatorname{Cl}_2\left(\alpha\right)-3\operatorname{Cl}_2\left(2\,\alpha\right)+\operatorname{Cl}_2\left(3\,\alpha\right) \tag{9.89}$$

with $\alpha=2\arctan\left(\sqrt{7}\right)$.

Show that (9.89) is equivalent to (1.1), indeed it is the form in which the identity was originally found using PSLQ.

12. **An extremal problem.** For $t>0$, $a>1$, let

$$f_a(t):=a^{-t}+a^{-1/t},$$

and set $f_a(0) := 1$ (the limiting value). Show that

$$\sup_{t\geq 1} f_a(t) = \max\left(\frac{2}{a}, 1\right).$$

Hint. Since $f_a(t) = f_a(1/t)$, and $f_a(t)$ is continuous on $[0,\infty)$ and differentiable on $(0,\infty)$, it suffices to show that $f_a(t)$ has a unique critical point (necessarily a minimum) in $(0,1)$. (See Figure 9.6.) For, in that case $f_a(t)$ must assume its maximum value at the larger of $f_a(0) = 1$ and $f_a(1) = 2/a$. Now the condition for a critical point becomes $a^{(t^2-1)/2t} = t$ and taking logs this is

$$(t^2-1)\log(a) = 2t\log(t).$$

Substituting $t = \sqrt{u}$ we see that it suffices to show that there is a unique $u \in (0,1)$ satisfying

$$g_a(u) := (u-1)\log(a) = \sqrt{u}\log(u) =: h(u).$$

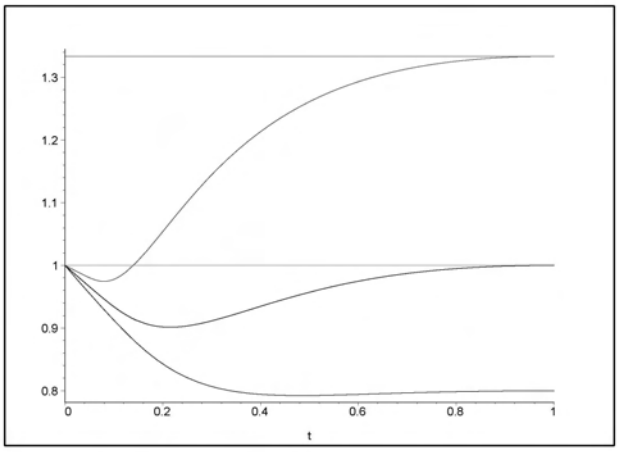


Figure 9.6. The function f_a on $[0,1]$ for $1 < a < 2$, $a = 2$, and $a > 2$.

Since $g_a(u)$ is affine, while $h(u)$ is convex on $[0,1]$, and $-\log(a) = g_a(0) < h(0) = 0 = h(1) = g_a(1)$, it follows that there is a *unique* $u \in (0,1)$ such that $g_a(u) = h(u)$, as illustrated in Figure 9.7. □

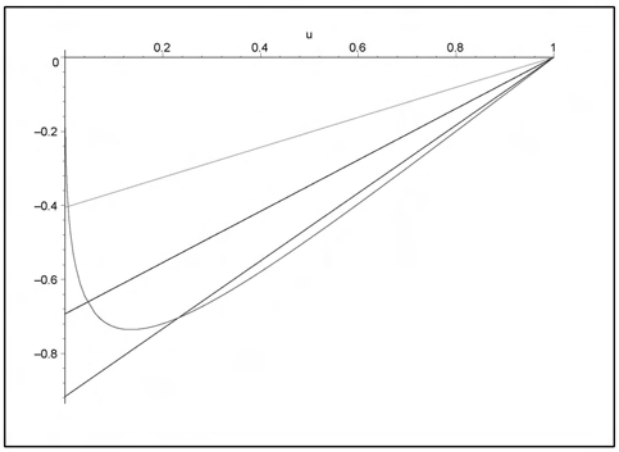


Figure 9.7. The functions h and g_a on $[0,1]$ for $1 < a < 2$, $a = 2$, and $a > 2$.

13. **Centrally symmetric polytopes**, [26, p. 276]. No example is known of a d-dimensional centrally symmetric polytope with fewer than 3^d faces. Can one exist?

14. **Minkowski's convex body theorem**, [26, p. 295]. One of the most beautiful and potent results relating geometry and number theory is Minkowski's theorem which asserts that in $\mathbb{R}^d$ a (compact) convex body A must meet an integer lattice Λ as soon as $\operatorname{vol} A > 2^d \det \Lambda$. (That $\operatorname{vol} A \geq 2^d \det \Lambda$ suffices in the compact case.)

    (a) Establish Siegel's result that

$$2^d = \operatorname{vol} A + \frac{4^d}{\operatorname{vol} A} \sum_{u \in \mathbb{Z}^d} \left| \int_{A/2} \exp\{-2\pi\, i \langle u, x \rangle\} \right|^2$$

    Hint. Again Fourier techniques are suggested. Use Parseval's formula applied to the function

$$\sum_{u \in \mathbb{Z}^d} \chi_{u + A/2}$$

    (b) Deduce Minkowski's theorem.

(c) Show this fails for non-symmetric bodies and that it is bwst possible (in 2 dimensions).

15. **Inequalities for sinc integrals.** Suppose that $\{a_n\}$ is a sequence of positive numbers. Let $s_n := \sum_{k=1}^{n} a_k$ and set

$$\tau_n := \int_0^\infty \prod_{k=0}^{n} \operatorname{sinc}(a_k x)\, dx.$$

Show that

(a)
$$0 < \tau_n \le \frac{\pi}{2\, a_0},$$
with equality if $n = 0$, or if $a_0 \ge s_n$ when $n \ge 1$.

(b) If $a_{n+1} \le a_0 < s_n$ with $n \ge 1$, then
$$0 < \tau_{n+1} \le \tau_n < \frac{1}{a_0}\frac{\pi}{2}.$$

(c) If $a_0 < s_{n_0}$ with $n_0 \ge 1$, and $\sum_{k=0}^{\infty} a_k^2 < \infty$, then there is an integer $n_1 \ge n_0$ such that
$$\tau_n \ge \int_0^\infty \prod_{k=0}^{\infty} \operatorname{sinc}(a_k x)\, dx \ge \int_0^\infty \prod_{k=0}^{\infty} \operatorname{sinc}^2(a_k x)\, dx > 0$$
for all $n \ge n_1$.

Observe that applying the result to different permutations of the parameters will in general yield different inequalities.

Proof (a). That $\tau_0 = \pi/(2a_0)$ is a standard result (proven e.g., by contour integration or Fourier analysis with the integral in question being improper). Assume therefore that $n \ge 1$, and define the following convolutions

$$F_0 := \frac{1}{a_0}\sqrt{\frac{\pi}{2}}\chi_{a_0}, \quad F_n := (\sqrt{2\pi})^{1-n} f_1 * f_2 * \cdots * f_n,$$

where $f_n := \frac{1}{a_n}\sqrt{\frac{\pi}{2}}\chi_{a_n}$.

By induction, for $n \ge 1, F_n(x)$ is even, vanishes on $(-\infty, -s_n) \cup (s_n, \infty)$, and is positive on $(-s_n, s_n)$. Moreover, $F_{n+1} = \frac{1}{\sqrt{2\pi}} F_n * f_{n+1}$, so that

$$F_{n+1}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_n(x-t) f_{n+1}(t)\, dt = \frac{1}{2a_{n+1}} \int_{x-a_{n+1}}^{x+a_{n+1}} F_n(u)\, du.$$

Hence $F_{n+1}(x)$ is absolutely continuous on $(-\infty,\infty)$ and, for almost all $x \in (-\infty,\infty)$,

$$\begin{aligned} 2a_{n+1}F'_{n+1}(x) &= F_n(x+a_{n+1}) - F_n(x-a_{n+1}) \\ &= F_n(x+a_{n+1}) - F_n(a_{n+1}-x). \end{aligned}$$

Since $(x+a_{n+1}) \geq \max\{(x-a_{n+1}),(a_{n+1}-x)\} \geq 0$ when $x > 0$, it follows that if $F_n(x)$ is monotone non-increasing on $(0,\infty)$, then $F'_{n+1}(x) \leq 0$ for a.a. $x \in (0,\infty)$, and so $F_{n+1}(x)$ is monotone non-increasing on $(0,\infty)$. This monotonicity property of F_n on $(0,\infty)$ is therefore established by induction for all $n \geq 1$. Also F_n is the cosine transform (FCT) of $\sigma_n(x) := \prod_{k=1}^{n} \mathrm{sinc}(a_k x)$, and σ_n is the FCT of F_n. Thus, all our functions and transforms are even and are in $L_1(0,\infty) \cap L_2(0,\infty)$.

Hence, by Parseval's theorem,

$$\tau_n = \int_0^\infty F_n(x)F_0(x)\,dx = \frac{1}{a_0}\sqrt{\frac{\pi}{2}}\int_0^{\min(s_n,a_0)} F_n(x)\,dx. \qquad (9.90)$$

When $a_0 \geq s_n$, the final term is equal to

$$\frac{1}{a_0}\sqrt{\frac{\pi}{2}}\sqrt{\frac{\pi}{2}}\sigma_n(0) = \frac{1}{a_0}\frac{\pi}{2}$$

since $\sigma_n(x)$ is continuous on $(-\infty,\infty)$; and when $a_0 < s_n$, the term is positive and less than $\pi/(2a_0)$ since $F_n(x)$ is positive and continuous for $0 < x < s_n$.

(b). Note again that $F_{n+1} = \frac{1}{\sqrt{2\pi}} F_n * f_{n+1}$, and hence, for $y > 0$,

$$\begin{aligned} \int_0^y F_{n+1}(x)\,dx &= \frac{1}{\sqrt{2\pi}}\int_0^y dx \int_{-\infty}^{\infty} F_n(x-t)f_{n+1}(t)\,dt \\ &= \frac{1}{2a_{n+1}}\int_0^y dx \int_{-a_{n+1}}^{a_{n+1}} F_n(x-t)\,dt \\ &= \frac{1}{2a_{n+1}}\int_{-a_{n+1}}^{a_{n+1}} dt \int_0^y F_n(x-t)\,dx \\ &= \int_0^y F_n(u)\,du + \frac{1}{2a_{n+1}}(I_1+I_2), \end{aligned}$$

where

$$I_1 := \int_{-a_{n+1}}^{a_{n+1}} dt \int_{-t}^{0} F_n(u)\,du \quad \text{and} \quad I_2 := \int_{-a_{n+1}}^{a_{n+1}} dt \int_{y}^{y-t} F_n(u)\,du.$$

Now $I_1 = 0$ since $t \mapsto \int_{-t}^{0} F_n(u)\,du$ is odd, and for $y \geq a_{n+1}$,

$$\begin{aligned} I_2 &= \int_0^{a_{n+1}} dt \int_y^{y-t} F_n(u)\,du + \int_{-a_{n+1}}^{0} dt \int_y^{y-t} F_n(u)\,du \\ &= -\int_0^{a_{n+1}} dt \int_{y-t}^{y} F_n(u)\,du + \int_0^{a_{n+1}} dt \int_y^{y+t} F_n(u)\,du \\ &= \int_0^{a_{n+1}} dt \int_{y-t}^{y} \big(F_n(u+t) - F_n(u)\big)\,du \leq 0 \end{aligned}$$

since $F_n(u)$ is non-increasing for $u \geq y - t \geq y - a_{n+1} \geq 0$. Hence

$$\int_0^y F_{n+1}(x)\,dx \leq \int_0^y F_n(x)\,dx \text{ when } a_{n+1} \leq y < s_n. \tag{9.91}$$

It follows from (9.90), and (9.91) with $y = a_0$, that $0 < \tau_{n+1} \leq \tau_n$ if $a_{n+1} \leq a_0 < s_n$.

(c). Let $\rho(x) := \lim_{n\to\infty} \sigma_n^2(x) = \prod_{k=1}^{\infty} \operatorname{sinc}^2(a_k x)$ for $x > 0$. The limit exists since $0 \leq \operatorname{sinc}^2(a_k x) < 1$, and there is a set A differing from $(0,\infty)$ by a countable set such that $0 < \operatorname{sinc}^2(a_k x) < 1$ whenever $x \in A$ and $k = 1, 2, \ldots$. Now

$$\operatorname{sinc}(a_k x) = 1 - \delta_k, \text{ where } 0 \leq \frac{\delta_k}{a_k^2} \to \frac{x^2}{3} \text{ as } k \to \infty,$$

so that $\sum_{k=1}^{\infty} \delta_k < \infty$, and hence, by standard theory of infinite products, $\sigma(x) := \lim_{n\to\infty} \sigma_n(x)$ exists and $\sigma^2(x) = \rho(x) > 0$ for $x \in A$. It follows, by part (b), that

$$\tau_n \geq \int_0^{\infty} \sigma_n^2(x)\,dx \geq \int_0^{\infty} \rho(x)\,dx > 0$$

for all $n \geq n_1$, where $n_1 \geq n_0$ is an integer such that $a_{n+1} \leq a_0$ for all $n \geq n_1$. In addition, by dominated convergence,

$$\lim_{n\to\infty} \tau_n = \int_0^{\infty} \sigma(x)\,dx \geq \int_0^{\infty} \rho(x)\,dx,$$

and we are done. □

16. **A discrete dynamical system.** One of the advantages of a symbolic computer package is how accessible it makes initial study of discrete dynamical systems. Recall that a map is *chaotic* if (i) it has a dense set of periodic orbits, (ii) exhibits *sensitive dependence*

on initial conditions and (iii) is *topologically transitive* meaning that orbits mix completely.

It is well known that "period three implies chaos" for one dimensional systems. In particular, Sharkovsky's theorem given in [44, p. 79] implies that a continuous self-map φ of the reals with a period-three point will have periodic points of all orders.

The following example due to Marc Chamberland shows how completely this fails in $\mathbb{R}^2$. Consider the dynamics $z_{n+1} := \varphi(z_n)$ of the map

$$(u, v) = \varphi(y, x^2 - y^2).$$

Show that

(a) all points in the open unit square are attracted to zero;

(b) there is a 3-cycle but no 2-cycle;

(c) there are divergent orbits.

It is instructive to plot the points for which the system appears to converge as in Figure 9.8.

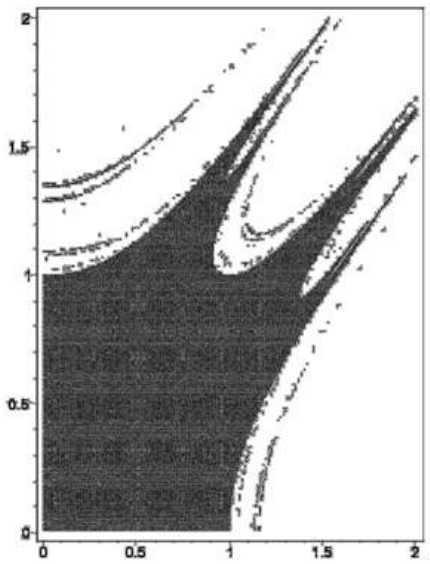


Figure 9.8. The region of convergence for φ

17. **Volume, surface area and ζ.** Recall that the volume and surface area of the ball in d-dimensional space are given by

$$V(d) := \frac{\pi^{d/2}}{\Gamma\left(d/2+1\right)} \qquad S(d) := \frac{2\,\pi^{d/2}}{\Gamma\left(d/2\right)}.$$

(a) Plot $V_\zeta := V/\zeta$ and $S_\zeta := S/\zeta$.

(b) We observe, as Douglas S. Robertson did, that S_ζ appears symmetric about $z = 1/2$. This is illustrated in Figure 9.9.

(c) Show that the symmetry

$$\frac{S(z)}{\zeta(z)} = \frac{S(1-z)}{\zeta(1-z)}$$

is equivalent to Riemann's functional equation for ζ, namely

$$\pi^{-(1-z)/2}\Gamma\left(\frac{1-z}{2}\right)\zeta(1-z) = \pi^{-z/2}\Gamma\left(\frac{z}{2}\right)\zeta(z)$$

(d) Use the duplication formula for Γ to derive a formula originally conjectured by Euler

$$\zeta(1-z) = 2^{1-z}\pi^{-z}\cos\left(\frac{z}{2}\right)\Gamma(z)\zeta(z).$$

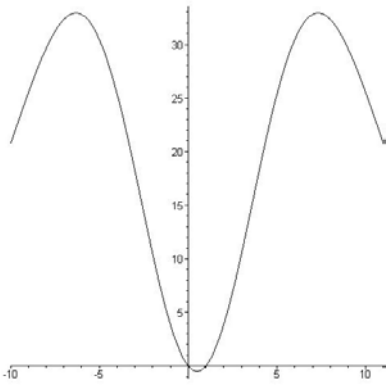


Figure 9.9. The symmetry of S/ζ

18. Evaluate

$$\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\frac{H_j\left(H_{k+1}-1\right)}{kj\left(k+1\right)\left(j+k\right)},$$

where as before $H_k := \sum_{i=1}^{k} 1/i$.

Hint. The answer is a combination of $\zeta(2), \zeta(3), \zeta(5)$, and $\zeta(2)\zeta(3)$.

19. **A multi-dimensional binomial series.** We consider, following Benoist Cloitre, the series

$$B_N := \sum_{n_i, n_2, \ldots, n_N \geq 0} \frac{1}{\binom{2\sum_{i=1}^{N} n_i}{\sum_{i=1}^{N} n_i}}.$$

By computing the first two values exactly

$$B_1 := \frac{4}{3} + \frac{2}{9}\,\frac{\pi}{\sqrt{3}}, \quad B_2 := 2 + \frac{4}{9}\,\frac{\pi}{\sqrt{3}}$$

and the next two via PSLQ

$$B_3 = 3 + \frac{20}{27}\,\frac{\pi}{\sqrt{3}}, \quad B_4 = \frac{40}{9} + \frac{280}{243}\,\frac{\pi}{\sqrt{3}}$$

we discover that the values all appear to be of the form $B_N = a_N + b_N\,\pi/\sqrt{3}$ with well structured rational positive coordinates. Given the binomial coefficients, it is reasonable to look for a two-term recurrence. Again integer relation methods apply and we quickly identify the first few terms of the recursion

$$3(n+1)\,B_{n+2} - (7n+6)\,B_{n+1} + 2(2n+1)\,B_n = 0. \tag{9.92}$$

[It is advisable to limit the number of terms one must sum to obtain sufficient digits of B_n.]

(a) Prove that (9.92) holds and so B_N is of the conjectured form.

(b) Determine a closed form for B_N.

(c) Explore

$$B_N(r,s) := \sum_{n_i, n_2, \ldots, n_N \geq 0}{}' \frac{1}{(\sum_{i=1}^{N} n_i)^r \binom{2\sum_{i=1}^{N} n_i}{\sum_{i=1}^{N} n_i}^s}.$$

for $r, s = 1, 2 \ldots$.

20. **A Gaussian integer zeta function.** Evaluate in closed form

$$\zeta_G(N) := \sum_{\mathbb{Z}_{(i)}}{}' \frac{1}{z^N} = \sum_{m,n}{}' \frac{1}{(m + i\,n)^N}$$

for positive, even integer $N > 1$. Here as always the '′' denotes that summation avoids the pole at 0.

Hint. This is implicitly covered in [45, pp. 167–170] and relies on analysis of the *Weierstrass* $\wp$ *function*, see [3, Ch. 18]. We recall that

$$\wp(x) := {\sum}' \frac{1}{(2in + 2im - x)^2} - \frac{1}{(2in + 2im)^2}.$$

We then differentiate twice and extract the coefficients of $\zeta_G(2n)$. *For N divisible by four*, the sum is actually is a rational multiple of powers of the *invariants*

$$g_2 = K\left(\frac{1}{\sqrt{2}}\right)^4 = \left(\frac{1}{4}\,\beta\left(\frac{1}{4}\right)\right)^4,$$

and $g_3 = 0$. This is the so called *lemniscate case* of $\wp$. Here $\beta(x) := \beta(x, x)$ a central Beta-function.

(a) Show that the general formula for $N \geq 1$ is

$$\zeta_G(4N) = p_N\,\frac{\left\{2\,K\left(1/\sqrt{2}\right)\right\}^{4N}}{(4\,N-1)},$$

where $p_1 = 1/20$ and

$$p_N = \frac{3\,\sum_{m=1}^{N-1} p_m\,p_{N-m}}{(4N+1)(2N-3)},$$

for $N > 1$. The next three values are $p_2 = 1/1200$, $p_3 = 1/156000$ and $p_4 = 1/21216000$. The corresponding values of $q_N := 16^N\,p_N$ are

$$\frac{4}{5}, \frac{16}{75}, \frac{128}{4875}, \frac{256}{82875}.$$

Finally note that q_N satisfies the same recursion. This leads to the simple expression

$$\zeta_G(4N) = \frac{q_N}{4\,N-1}\,K\left(\frac{1}{\sqrt{2}}\right)^{4N},$$

where $q_1 = 16/5$, and for $n > 1$

$$q_N = \frac{3\,\sum_{m=1}^{N-1} q_m\,q_{N-m}}{(4N+1)(2N-3)}.$$

(b) Show that for N congruent to two mod four, the sum is

$$\zeta_G(4N+2) = \sum_{Z(i)}{}' \frac{1}{z^{4N+2}} = 4\,\zeta(4N+2),$$

since

$$2 \sum_{m,n>0}{}' \frac{1}{z^{4N+2}} = \sum_{m,n>0} \frac{(m+in)^{4n+2}}{(m^2+n^2)^{4N+2}} + \sum_{m,n>0} \frac{(n+im)^{4n+2}}{(m^2+n^2)^{4N+2}}$$
$$\sum_{m,n>0} \frac{\mathrm{Re}\left((m+in)^{4n+2} + (n+im)^{4n+2}\right)}{(m^2+n^2)^{4N+2}} = 0$$

as the terms cancel pairwise. If we observe that $\pi^2 = \beta(1/2)^2$ we may more uniformly write

$$\zeta_G(4N) = \frac{q_N}{(4\,N-1)4^{4N}}\,\beta(1/4)^{4N},$$

$$\zeta_G(4N+2) = \frac{b_{4N+2}\,2^{4N+1}}{(4\,N-1)(4N+2)!}\,\beta(1/2)^{4N+2},$$

where b_N is the $N-$th Bernoulli number. By contrast, it is easily seen that

$$\zeta_G(2N+1) = 0.$$

21. **The generalized quantum sum** is defined by

$$\mathcal{Y}(a,b,c) := \sum_{}{}' \frac{1}{(z-a)(z-b)(z-c)}$$

with a,b,c fixed, not on the lattice. Show that $\mathcal{Y}$ has a formal power series development in $\zeta_G(n)$ for even n, just by writing the summand as

$$1/z^3\,(1+a/z+a^2/z^2+...)(1+b/z+b^2/z^2+...)(1+c/z+c^2/z^2+...)$$

This leads for $|a|, |b|, |c| < 1$ to

$$\mathcal{Y}(a,b,c) = \sum_{n=1}^{\infty} \pi_n(a,b,c)\,\zeta_G(2n+2)$$

where

$$\pi_n(a,b,c) := \sum_{k_1+k_2+k_3=n} a^{k_1} b^{k_2} c^{k_3}.$$

In particular, $\pi_n(a,a,a) = \binom{n}{2}\,a^3$. There is a corresponding expansion for more than three linear terms.

22. **A golden example.** Evaluate the following sum:

$$\Phi := \sum_{k=0}^{\infty} \left\{ \frac{G^2}{(5k+1)^2} - \frac{G}{(5k+2)^2} - \frac{G^2}{(5k+3)^2} + \frac{G^5}{(5k+4)^2} + \frac{2\,G^5}{(5k+5)^2} \right\} g^{5k} \tag{9.93}$$

where $g := (\sqrt{5}-1)/2$ is the *golden ration* and $G := (\sqrt{5}+1)/2 = g^{-1}$.

Hint. The answer, $\Phi = \pi^2/50$, was discovered empirically by Benoit Cloitre using integer relation methods. In the irrational base g, the constant Φ has digits independently computable since (9.93) gives an identity of BBP type. (See [44, Chapter 4].)

(a) With computer algebra assistance, work backwards to needing to show

$$\Phi = \operatorname{Re} \operatorname{Li}_2 \left(2\cos(\theta)\, e^{\pi\, i\, \theta} \right)$$

for $\theta := 2\pi/5$, where again $\operatorname{Li}_2(x) := \sum_{n=1}^{\infty} x^n/n^2$ is the *dilogarithm*.

(b) Show that or all real θ

$$\sum_{n=1}^{\infty} \frac{\cos(n\theta)\,(2\,\cos(\theta))^n}{n^2} = \left(\frac{\pi}{2} - \theta \right)^2.$$

This is equation (5.17) in Lewin [166] and is relatively easy to obtain.

(c) Show that $\theta = 2\pi/5$ makes $2\cos(\theta) = (\sqrt{5}-1)/2 = g$. Moreover $2\cos(n\theta)$ takes the values $2, g, -1/g, -1/g, g$ modulo 5. Thus, obtain (9.93).

(d) Any other rational multiple of π has a like form—nicest when $2\cos(n\theta)$ is rational or quadratic. Thus,with $\theta := \pi/3, 2\pi/3$ we obtain, inter alia, an evaluation of $\zeta(2)$. With $\theta := 3\pi/8, 5\pi/12$ obtain nice convergent identities.

23. The following evaluation is given in [141] and used in several interesting applications.

Theorem 9.3. *For $0 < a < b$, if f is defined by the requirement that $f(x)^a - f(x)^b = x^a - x^b$ and f is decreasing, then*

$$-\int_0^1 \frac{\log f(x)}{x}\, dx = \frac{\pi^2}{3ab}, \tag{9.94}$$

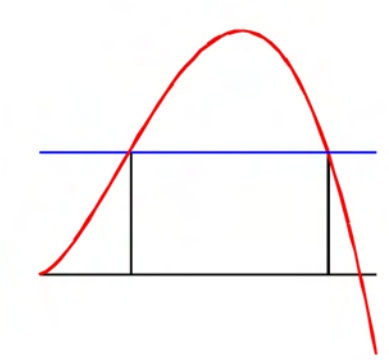


Figure 9.10. The construction of f

The picture in Figure 9.10 shows how f is defined.

The proof is quite elaborate and relies on (i) a reduction to the case $b = 1$, and (ii) application of the following identity after splitting the integral in 9.94 at $1/(a+1)$.

(a) Prove that for $0 < x < 1/(a+1)$

$$G_a(1-x) := \sum_{n=1}^{\infty} \frac{\Gamma((a+1)n)}{\Gamma(an)} \frac{(x(1-x)^a)^n}{(na)\,n!} = -\log(1-x).$$

(b) We request a geometric proof of (9.94). We do not know one.

(c) The function f can be determined explicitly in some cases:

i. $(b = 2a)$ $f(x) = 3(1-x^a)^{1/a}$.
ii. $(b = 3a)$ $f(x) = 3(-x^a/2 + \sqrt{4-3x^{2a}})^{1/a}$.
iii. $(2b = 3a)$ $f(x) = 3(-x^a/2 + +\sqrt{-3x^{2a}+2x^a+1})^{1/a}$.

24. Determine the limit of $\gamma_n := n \int_0^1 t^{n-1}\,((1+t)/2)^n\,dt$, as n goes to infinity.

Hint. It may help to show that the value $\gamma_n = {}_2\mathrm{F}_1\,(1, -n; n+1; 1/2)$.

25. **AMM Problem 11103, October 2004: a second proof.** In the Exercises for Chapter 1 we explored this problem. Show that for positive integer n

$$2^{1-n} \sum_{k=1}^{n} \frac{\binom{n}{2\,k-1}}{2\,k-1} = \sum_{k=1}^{n} \frac{1}{k\binom{n}{k}}.$$

(a) Use integer relation methods to predict the following recursion

$$(n+1)\,u_n - (2n+4)\,u_{n+1} + (3n+4)\,u_{n+2} = 0 \tag{9.95}$$

for the hypergeometric sum on the right-hand side of the requested identity.

(b) Use the Wilf-Zeilberger method to prove this.

(c) Prove that the left-hand series satisfies the same recursion (9.95) and initial conditions and hence coincides with the right,

26. **A difficult limit.** This originates with Mike Hirschorn's treatment in [139] of a solution to AMM problem 10886.

Prove that

$$\gamma_n := \frac{1}{2^n \binom{2n}{n}} \frac{\sum_{k=0}^{n} \frac{\binom{n}{k}}{n+k}}{\sum_{k=0}^{n} (-1)^k \frac{\binom{n}{k}}{n+k}}$$

converges as n goes to ∞ and determine the limit as follows:

(a) Compute enough terms to make plausible that the limit is $2/3$—with error roughly $2/(27n)$.

(b) Verify that also

$$\gamma_n = n\,2^{-n} \int_0^1 t^{n-1}(1+t)^n\,dt.$$

(c) Show that

$$\sum_{n=0}^{\infty} \gamma_n x^n = \int_0^1 \frac{2(1+t)x}{\left(2-(t+t^2)x\right)^2}\,dt \geq \frac{x(3-x)}{(1-x)(4-x)}. \tag{9.96}$$

(d) Hence, show that $\gamma_n \geq 2/3 + 4^{-n}/3$.

(e) Use Wilf-Zeilberger methods to verify a third representation:

$$\gamma_n = \frac{1}{\binom{2n}{n}} \sum_{k\geq 0} 2^{-k} \binom{2n-1-k}{n-1}.$$

(f) Finally, estimate that

$$\gamma_n \leq \frac{1}{2} + \frac{2n}{2n-1} \sum_{k\geq 1} 2^{-(2k+1)} \leq \frac{2}{3} + \frac{6}{2n-1}.$$

27. **The number** 1729 **revisited.** Famously, G.H. Hardy while visiting Ramanujan in hospital remarked that his cab number 1729 was very uninteresting. To which Ramanujan replied that it was very interesting being the smallest number expressible as the sum of two integer cubes in two distinct ways:

$$1729 = 10^3 + 9^3 = 12^3 + 1.$$

(a) Determine the second smallest number with this property, that is, with $r_3(n) = 2$.

(b) As Hirschorn [140] observes 1729 is also special in being a solution to $n = x^3 \pm 1$. Indeed, in the *Lost Notebook* Ramanujan states that the recursion for $u_n = (x_n, y_n, z_n)$

$$u_{n+3} := 82\,u_{n+2} + 82\,u_{n+1} - u_n \tag{9.97}$$

with initial conditions $u_1 := (9, 10, 12)$, $u_2 := (791, 812, 1010)$, and $u_3 := (65601, 67402, 83802)$ solves

$$x_n^3 + y_n^3 - z_n^3 = (-1)^{n+1},$$

for all (positive and negative) integer n. Thus, we have infinitely many near misses to Fermat's equation. For example, $u_{-1} = (-2, 2, 1)$ and

$$x_{12} = 12247547739697622322431,$$

$$y_{12} = 12583657892407702716002,$$

and

$$z_{12} = 15645544827332108296610.$$

(c) There is a 3×3 matrix M, with minimal polynomial $1 - 82\,t - 82\,t^2 + t^3$ and determinant -1, such that $u_n = M\,u_{n-1}$. We have

$$M := \begin{bmatrix} 63 & 68 & -104 \\ -80 & -85 & 131 \\ -64 & -67 & 104 \end{bmatrix} \text{ with } M^{-1} = \begin{bmatrix} 63 & 104 & -68 \\ 64 & 104 & -67 \\ 80 & 131 & -85 \end{bmatrix}$$

(d) Try to prove that if $a^3 + b^3 - c^3 = \pm 1$ then $M(a, b, c) = (a', b', c')$ satisfies $a'^3 + b'^3 - c'^3 = \mp 1$.

(e) Failing that, begin with

$$\left(x^2 + 9\,x - 1\right)^3 + \left(2\,x^2 + 10\right)^3 = \left(x^2 - 7\,x - 9\right)^3 + \left(2\,x^2 + 4\,x + 12\right)^3.$$

Now replace x by v/u, multiply by u^6, and obtain the resulting homogeneous identity

$$\left(10\,u^2+2\,v^2\right)^3+\left(9u^2+7\,uv-v^2\right)^3=\left(12\,u^2+4\,uv+2\,v^2\right)^3 + \left(u^2-9\,uv-v^2\right)^3.$$

(f) Check that the sequence $s_0 := 0, s_1 := 1$ and $s_{n+1} := 9s_n + s_{n-1}$ for all $n > 1$ satisfies $s_n^2 - 9s_n s_{n-1} - s_{n-1}^2 = (-1)^{n-1}$. Hence,

$$(9s_n^2+7s_n s_{n-1}-s_{n-1}^2, 10s_n^2+2s_{n-1}^3, 12s_n^2+4s_n s_{n-1}+2s_{n-1}^2)$$

parameterizes the sequence (u_n).

28. **Integer relations as integer knapsack problems**. An *integer knapsack problem* is an integer programming problem with one constraint, such as

$$\min \sum_{k=1}^{N} \omega_k \quad \text{subject to} \quad \sum_{k=1}^{N} \omega_k \alpha_k = \beta,\ \omega_1 \geq 0, \omega_2 \geq 0, \ldots, \omega_N \geq 0, \tag{9.98}$$

where the nonnegative weights ω_k are required to be integers. If we view β as a quantity we wish to express in terms of real (or vector) quantities $\mathcal{A} := \{\alpha_1, \alpha_2 \ldots, \alpha_N\}$, then 9.98 will solve a positive integer relation problem. Since there are excellent integer programming algorithms, it makes sense to investigate this as a tool in various settings:

(a) *When all signs are known* as when one is checking a known formula but has forgotten the exact constants, e.g., in Machin's formula

$$\arctan(1) = \omega_1 \arctan\left(\frac{1}{5}\right) - \omega_2 \arctan\left(\frac{1}{239}\right),$$

or Euler's formula

$$\arctan(1) = \omega_1 \arctan\left(\frac{1}{2}\right) + \omega_2 \arctan\left(\frac{1}{5}\right) + \omega_3 \arctan\left(\frac{1}{8}\right).$$

(b) *When all signs are anticipated.* This can also be something provable. Indeed, for $\zeta(4N+1)$, very pretty three term representations arise from a couple of PSLQ observations by Simon Plouffe. For $N = 1, 2, 3 \ldots$

$$\left\{2-(-4)^{-N}\right\}\sum_{k=1}^{\infty}\frac{\coth(k\pi)}{k^{4N+1}} \quad - \quad (-4)^{-2\,N}\sum_{k=1}^{\infty}\frac{\tanh(k\pi)}{k^{4N+1}} = Q_N \times \pi^{4N+1},$$

where the quantity Q_N in (9.99) is an explicit rational:

$$Q_N := \sum_{k=0}^{2N+1} \frac{B_{4N+2-2k}B_{2k}}{(4N+2-2k)!(2k)!}\left\{(-1)^{\binom{k}{2}}(-4)^N 2^k + (-4)^k\right\}.$$

On substituting

$$\tanh(x) = 1 - \frac{2}{\exp(2x)+1} \quad \text{and} \quad \coth(x) = 1 + \frac{2}{\exp(2x)-1}$$

one may solve for $\zeta(4N+1)$. We list two examples:

$$\zeta(5) = \frac{1}{294}\pi^5 - \frac{2}{35}\sum_{k=1}^{\infty}\frac{1}{(1+e^{2k\pi})k^5} + \frac{72}{35}\sum_{k=1}^{\infty}\frac{1}{(1-e^{2k\pi})k^5}.$$

and

$$\zeta(9) = \frac{125}{3704778}\pi^9 - \frac{2}{495}\sum_{k=1}^{\infty}\frac{1}{(1+e^{2k\pi})k^9} + \frac{992}{495}\sum_{k=1}^{\infty}\frac{1}{(1-e^{2k\pi})k^9}.$$

This sign pattern sustains for all $\zeta(4N+1)$ and allows one to determine sufficient coefficients to validate or discover the general formula.

(c) The well known series for $\arcsin^2(x)$ generalizes fully, as we saw in the Exercises for Chapter 1. The seed identity is well known:

$$\arcsin^2\left(\frac{x}{2}\right) = \frac{1}{2}\sum_{k=1}^{\infty}\frac{x^{2\,k}}{\binom{2\,k}{k}k^2}. \tag{9.99}$$

The second, slightly rewritten is less well known:

$$\arcsin^4\left(\frac{x}{2}\right) = \frac{3}{2}\sum_{k=1}^{\infty}\left\{\sum_{m=1}^{k-1}\frac{1}{m^2}\right\}\frac{x^{2\,k}}{\binom{2\,k}{k}k^2},$$

and when compared, they suggest the third and fourth—subsequently confirmed numerically—are suggested by the prior if flimsy pattern:

$$\arcsin^6\left(\frac{x}{2}\right) = \frac{45}{4}\sum_{k=1}^{\infty}\left\{\sum_{m=1}^{k-1}\frac{1}{m^2}\sum_{n=1}^{m-1}\frac{1}{n^2}\right\}\frac{x^{2\,k}}{\binom{2\,k}{k}k^2}.$$

$$\arcsin^8\left(\frac{x}{2}\right) = \frac{315}{2}\sum_{k=1}^{\infty}\left\{\sum_{m=1}^{k-1}\frac{1}{m^2}\sum_{n=1}^{m-1}\frac{1}{n^2}\sum_{p=1}^{n-1}\frac{1}{p^2}\right\}\frac{x^{2\,k}}{\binom{2\,k}{k}k^2}.$$

In this case all signs are positive and positive integer relation methods again apply. These relations were found by hunting over larger sets of multidimensional sums.

(d) In particular, for $N = 1, 2, \ldots$

$$\sum_{k=1}^{\infty} \frac{H_N(k)}{\binom{2\,k}{k} k^2} = \frac{\pi^{2N}}{6^{2N}\,(2N)!},$$

with the $H_N(k)$ multi-dimensional harmonic numbers. The first few righthand side values are

$$\frac{1}{72}\,\pi^2, \frac{1}{31104}\,\pi^4, \frac{1}{33592320}\,\pi^6 \text{ and } \frac{1}{67722117120}\,\pi^8.$$

(e) Let

$$\mathcal{Z}(s,t) := \sum_{n>m>0} \frac{(-1)^{n-1}}{n^s}\,\frac{\chi_3(m)}{m^t},$$

where χ_3 is the character modulo 3 (which is ± 1 when $n \equiv \pm 1$ modulo 3 and is zero otherwise). Then

$$\begin{aligned}
\mathcal{Z}(2N+1,1) &= \frac{L_{-3}\,(2\,N+2)}{4^{\,1+N}} - \frac{1+4^{-N}}{2}\,L_{-3}\,(2\,N+1)\log(3) \\
&+ \sum_{k=1}^{N} \frac{1-3^{-2\,N+2\,k}}{2}\,L_{-3}\,(2\,N-2\,k+2)\,\alpha\,(2\,k) \qquad (9.100) \\
&- \sum_{k=1}^{N} \frac{1-9^{-k}}{1-4^{-k}}\,\frac{1+4^{-N+k}}{2}\,L_{-3}\,(2\,N-2\,k+1)\,\alpha\,(2\,k+1) \\
&- 2\,L_{-3}\,(1)\,\alpha\,(2\,N+1)\,,
\end{aligned}$$

where α is the alternating zeta-function, and L_{-3} is the *primitive L-series modulo 3*. One first numerically evaluates such $\mathcal{Z}$ sums as integrals via

$$\mathcal{Z}(s,1) = \int_0^1 \frac{\mathrm{L}_{-3}(s:-x)}{1+x}\,dx,$$

where

$$L_3(s:x) = \sum_{m=1}^{\infty} \frac{\chi_3(m)x^m}{m^s} = xxx,$$

and $\chi_3(n)$ repeats $1, -1, 0$ modulo three. Notice also that the sign pattern will become obvious well before the exact form of the coefficients—especially those in the third line of (9.100). This sort of evaluation, and much more about character Euler sums, may be pursued in [64].

(f) *For small* $\mathcal{A}$, say $N < 6$, one may try all 2^N sign permutations. For example, one may remember that

$$\Gamma(2x) = \pi^{\alpha(x)} 2^{\beta x+\gamma} \Gamma(x)\Gamma(x+1/2),$$

but not remember the rational constants. Taking logarithms and using PSLQ at a few rational values of x will recreate the formula. Likewise, *for medium sized* $\mathcal{A}$, where one knows or anticipates many of the signs on may fix those signs and search over all remaining combinations.

Finally, one may well have *under-determined systems* (with too many relations) and wish to select a relation with a predetermined sign configuration.

29. **Cheating God somehow.** Consider that

$$\sum_{n=1}^{\infty} \frac{\lfloor n \tanh(\pi) \rfloor}{10^n} \stackrel{?}{=} \frac{1}{81}$$

is valid to **268** places, while

$$\sum_{n=1}^{\infty} \frac{\lfloor n \tanh\left(\frac{\pi}{2}\right) \rfloor}{10^n} \stackrel{?}{=} \frac{1}{81}$$

is valid to just **12** places. Both are actually *transcendental numbers*. Correspondingly the *simple continued fractions* for $\tanh(\pi)$ and $\tanh\left(\frac{\pi}{2}\right)$ are respectively

$$[0, 1, \mathbf{267}, 4, 14, 1, 2, 1, 2, 2, 1, 2, 3, 8, 3, 1]$$

and

$$[0, 1, \mathbf{11}, 14, 4, 1, 1, 1, 3, 1, 295, 4, 4, 1, 5, 17, 7].$$

Bill Gosper describes how continued fractions let you "see" what a number is. "[I]t's completely astounding ... it looks like you are cheating God somehow.'

30. **A sincing feeling.** Let

$$I_N := \int_0^{\infty} \prod_{n=0}^{N} \operatorname{sinc}\left(\frac{t}{2n+1}\right) dt. \tag{9.101}$$

(a) Confirm both numerically and symbolically that each $I_N = \pi/2$ for $N < 7$, but $I_7 = .4999999999264685932\,\pi/2$. This occurs because $\sum_{n=1}^{6} 1/(2n+1) < 1 < \sum_{n=1}^{7} 1/(2n+1)$. (See [45].)

(b) Now consider

$$J_N^1 := \int_0^\infty \prod_{n=0}^{N} \operatorname{sinc}\left(\frac{t}{3n+1}\right) dt,$$

and

$$J_N^2 := \int_0^\infty \prod_{n=0}^{N} \operatorname{sinc}\left(\frac{t}{3n+2}\right) dt.$$

Verify that $J_2(N) = \pi$ for $N < 5$ but not for $N = 5$, and relate this to the behaviour of $\sum_{n=1}^{N} 1/(3n+2)$.

(c) Attempt the same tasks for J_N^1.

31. **Coincidences do occur.** The approximations

$$\pi \approx \frac{3}{\sqrt{163}} \log(640320) \quad \text{and} \quad \pi \approx \sqrt{2}\,\frac{9801}{4412}$$

occur for deep number theoretic reasons—the first good to 15 places, the second to eight. By contrast

$$e^\pi - \pi = \mathbf{19.99909997}9189475768\ldots$$

most probably for no good reason. This seemed more bizarre on an eight digit calculator. Likewise, as spotted by Pierre Lanchon recently

$$e = \overline{\mathbf{10.1011011111100001}0}101000101100\ldots$$

while

$$\pi = 11.0010\overline{\mathbf{0100001111110110101}}01000\ldots.$$

have 19 bits agreeing in base two, with one read right to left and the other right to left!

More extended coincidences are almost always contrived such as in Exercise 29 and in Exercise 30 (c), and strong heuristics exist for believing results like the empirical ζ-function evaluation of (9.100). But recall our discussion of the cosine integrals in Chapter 8, and the famous *Skewes number*

$$\int_2^x \frac{dt}{\log t} \geq \pi(x) \quad \text{with first known failure around } 10^{360}$$

and the *Merten Conjecture*

$$\left|\sum_{k=1}^{n} \mu(k)\right| \leq \sqrt{n} \quad \text{with first known failure around } 10^{110}$$

counter-examples. Here μ is the*Möbius function*.

32. **Roman numeration.** Consider the following Roman 'fraction'.

$$\mathrm{II} = \frac{\mathrm{XXII}}{\mathrm{VIII}}$$

A Hungarian school contest asked for a movement of one symbol to make a 'true' identity.

Answer. The proposers thought

$$\overline{\mathrm{II}} = \frac{\mathrm{XXII}}{\mathrm{VII}}.$$

33. For general integer N, determine the inverse of $M_1 := A+B-C$ and $M_2 := A+B+2C$ when A, B, C are $N \times N$ matrices with entries

```
A:=N->Matrix(N,N,(j,k)->(-1)^(k+1)*binomial(2*N-j,2*N-k));
B:=N->Matrix(N,N,(j,k)->(-1)^(k+1)*binomial(2*N-j,k-1));
C:=N->Matrix(N,N,(j,k)->(-1)^(k+1)*binomial(j-1,k-1));
```

Hint. $A^2 = C^2 = I$ and $B = CA$.

34. **Some convexity properties** [191]. Examine the convexity and log-convexity properties of

$$s_k(x) := \sum_{k=0}^{N} \frac{x^k}{k!} \quad \text{and} \quad t_k(x) := e^x - \sum_{k=0}^{N} \frac{x^k}{k!},$$

as a function of $x > 0$, for $k = 1, 2, \ldots$.

35. **Two trigonometric inequalities** [191].

(a) Show that

$$\frac{4}{\pi}\frac{x}{x^2-1} \le \tan\left(\frac{\pi}{2}\right) \le \frac{\pi}{2}\frac{x}{x^2-1},$$

for $0 < x < 1$.

(b) Show that

$$\left(\frac{\sin x}{x}\right)^3 > \cos x,$$

for $0 < x < \pi/2$.

(c) In what sense are these best possible?

36. **Dyson's conjecture** [212] For non-negative integers $a_1, a_2, \cdots, a_n$, determine the constant term of

$$\prod_{1 \le i \ne j \le n} \left(1 - \frac{x_i}{x_j}\right)^{a_j}.$$

Hint. The coefficient is

$$\frac{(a_1 + a_2 + \ldots + a_n)!}{a_1! a_2! \ldots a_n!}.$$

37. **Other chaos games.** Determine how to adjust the random fractal triangle of Chapter 8 to generate the hexagon in Figure 9.11, and like shapes. *Hint*: consult
`http://math.bu.edu/DYSYS/applets/fractalina.html`

Figure 9.11. A random Sierpinski hexagon

38. **Fourteen proofs of a result about tiling a rectangle** [226, 229]. Prove that whenever a rectangle is tiled by rectangles each of which has at least one integer side, then the ambient rectangle has at least one integer side.

39. **"But, My Lord, being by you found out, I wonder nobody else found it out before, when, now known, it is so easy."** Discoveries often have this feature. They may be rapidly transformative as this account by Garry J. Tee (Auckland) makes clear. Immediately after the publication of Napier's Miraculous Canon of Logarithms (Edinburgh 1614), Briggs began teaching logarithms at Gresham College. He convinced the Honorable East India Company that they needed logarithms, to enable their captains to navigate ships to India and return. The Honorable Company paid Edward Wright (Savilian Professor of Geometry at Oxford) to translate Napier's Latin text into English, and in 1615 Henry Briggs undertook the laborious journey from London to Edinburgh, taking the manuscript translation for checking by Napier. Briggs was acquainted with John Marr (or Mair), who was compass-maker and dial-maker to the kings James

6th and Charles 1st. Marr witnessed that meeting, and he reported it to the prominent astrologist William Lilly (1602-1681), who was consulted by the Commonwealth Government about an auspicious date on which to execute King Charles 1st. Auckland City Library has a copy of "Mr Lilly's History of His Life and Times" (published 1715), with Marr's report of that meeting: "I will acquaint you with one memorable story related unto me by Mr John Marr, an excellent mathematician and Geometer whom I conceive you remember. He was servant to King James and Charles. At first, when the Lord Napier of Merchiston made public his logarithms, Mr Briggs, then Reader of the Astronomy lecture at Gresham College in London, was so surprised with admiration of them, that he could have no quietness in himself, until he had seen that noble person, the Lord Merchiston, whose only invention they were. He acquaints Mr Marr herewith, who went into Scotland before Mr Briggs, purposely to be there when these so learned persons should meet. Mr Briggs appoints a certain day, when to meet in Edinburgh, but failing thereof, the Lord Napier was doubtful he would not come. It happened one day as John Marr and the Lord Napier were speaking of Mr Briggs; "Ah, John," says Merchiston, "Mr Briggs will not come now". At the very instant one knocks on his gate. John Marr hasted down, and it proved Mr Briggs to his great contentment. He brings Mr Briggs up into my Lord's chamber, where almost one quarter of an hour was spent, each beholding other almost with admiration, before one word was spoke. At last Mr Briggs began; *"My Lord, I have undertaken this long journey purposely to see your person, and to know by what engine of wit or ingenuity you came first to think of this most excellent help unto Astronomy, namely logarithms. But, My Lord, being by you found out (i.e. discovered), I wonder nobody else found it out before, when, now known, it is so easy*". He was nobly entertained by the Lord Napier, and every summer after that, during the Lord's being alive, this venerable man Mr Briggs went into Scotland purposely to visit him. These two persons were worthy men in their time, and yet the one, namely Lord Merchiston, was a great lover of Astrology; but Briggs the most satyrical man against it that was ever known. But the reason hereof I conceive was that Briggs was a severe Presbyterian, and wholly conversant with persons of that judgement; whereas the Lord Merchiston was a general scholar, and deeply read in all divine and human histories. This is the same Merchiston who made the most serious and learned exposition upon the *Revelation of St. John*, which is the best that ever yet appeared in the world." Briggs and Napier independently invented decimal logarithms, and in 1615 Briggs agreed to compute tables of the new form of logarithms. He

visited Napier again in 1616 and prepared to visit him again in 1617, but Napier died on 1617 April 4.

40. **Which sequence grows faster?** For $b > a > 0$ and $n \geq 0$ which sequence is ultimately larger: $a^{(b^n)}$ or $b^{(a^n)}$ and when?

41. **What is eerie about this limit?** Define the sequence e_n by $e_1 = 0, e_2 = 1$, and $e_n = e_{n-1} + e_{n-2}/(n-2)$. What is the limit of n/e_n as n approaches infinity? What is the rate of convergence?

42. **Jeff Tupper's self-referent fact.** (What is it?) Graph the set of points (x, y) such that

$$\frac{1}{2} < \left\lfloor \mathrm{mod}\left(\left\lfloor \frac{y}{17} \right\rfloor 2^{-17\lfloor x \rfloor - \mathrm{mod}(\lfloor y \rfloor, 17)}, 2\right)\right\rfloor$$

in the region $0 < x < 107$ and $N < y < N + 17$, where N is the following 541-digit integer:

960939379918958884971672962127852754715004339660129306651505
519271702802395266424689642842174350718121267153782770623355
993237280874144307891325963941337723487857735749823926629715
517173716995165232890538221612403238855866184013235585136048
828693337902491454229288667081096184496091705183454067827731
551705405381627380967602565625016981482083418783163849115590
225610003652351370343874461848378737238198224849863465033159
410054974700593138339226497249461751545728366702369745461014
655997933798537483143786841806593422227898388722980000748404 7

The picture in Figure 9.12 shows what transpires.

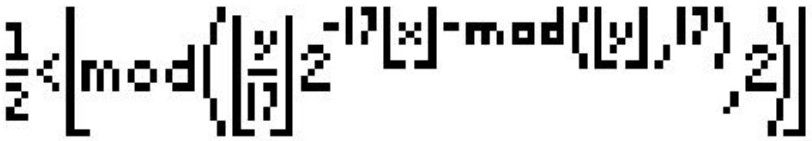


Figure 9.12. The answer is ...

43. For general integer N, determine the inverse of $M_1 := A + B - C$ and $M_2 := A + B + 2C$ when A, B, C are $N \times N$ matrices with entries

```
A:=N->Matrix(N,N,(j,k)->(-1)^(k+1)*binomial(2*N-j,2*N-k));
B:=N->Matrix(N,N,(j,k)->(-1)^(k+1)*binomial(2*N-j,k-1));
C:=N->Matrix(N,N,(j,k)->(-1)^(k+1)*binomial(j-1,k-1));
```

Hint. $A^2 = C^2 = I$ and $B = CA$.

44. **Some convexity properties** [191]. Examine the convexity and log-convexity properties of

$$s_k(x) := \sum_{k=0}^{N} \frac{x^k}{k!} \quad \text{and} \quad t_k(x) := e^x - \sum_{k=0}^{N} \frac{x^k}{k!},$$

as a function of $x > 0$, for $k = 1, 2, \ldots$.

45. **Two trigonometric inequalities** [191].

(a) Show that

$$\frac{4}{\pi} \frac{x}{x^2 - 1} \leq \tan\left(\frac{\pi}{2}\right) \leq \frac{\pi}{2} \frac{x}{x^2 - 1},$$

for $0 < x < 1$.

(b) Show that

$$\left(\frac{\sin x}{x}\right)^3 > \cos x,$$

for $0 < x < \pi/2$.

(c) In what sense are these best possible?

46. **Dyson's conjecture** [212] For non-negative integers $a_1, a_2, \cdots, a_n$, determine the constant term $c_n(a_1, a_2, \ldots, a_n)$ of

$$F_n(a_1, a_2, \ldots, a_n) := \prod_{1 \leq i \neq j \leq n} \left(1 - \frac{x_i}{x_j}\right)^{a_j}.$$

Hint. First show that $c_2(a_1, a_2) = \binom{a_1 + a_2}{a_1}$. Now show inductively that the coefficient is

$$d_n(a_1, a_2, \ldots, a_n) := \frac{(a_1 + a_2 + \ldots + a_n)!}{a_1! a_2! \ldots a_n!}.$$

To prove this, observe that

$$F_n(a_1, a_2, \ldots, a_n) = \sum_{i=1}^{n} F_n(a_1, \ldots, a_{i-1}, a_i - 1, a_{i+1}, \ldots, a_n), \tag{9.102}$$

while

$$F_n(a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n) = F_{n-1}(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n) \tag{9.103}$$

and hence c_n satisfies the same recursion. Finally show that d_n also satisfies (9.102) and (9.103). In [212] this idea is extended neatly to other coefficients of F_n. One empirically determines d_n and then corroborates as above.

47. Show that for $n = 1, 2, 3 \ldots$

$$\frac{2}{\pi} \int_0^\infty \left(\frac{\sin x}{x} \right)^n dx = \frac{n}{2^{n-1}} \sum_{r=0}^{\lceil n/2 \rceil - 1} (-1)^r \frac{(n-2\,r)^{n-1}}{(n-r)!\,r!}.$$

48. **Putnam Problem A6, 1999** Suppose $u_1 = 1, u_2 = 2, u_3 = 24$, and

$$u_n = \frac{6\,u_{n-1}^2 u_{n-3} - 8\,u_{n-1} u_{n-2}^2}{u_{n-2} u_{n-3}}.$$

Show that u_n is always a multiple of n. *Hint.* Determine the recursion for u_n/u_{n-1}.

49. Confirm the following Bernoulli number congruences for prime $p > 3$, [76].

(a)
$$\sum_{k=1}^{p-1} \frac{1}{k} \equiv -\frac{p^2}{3}\, B_{p-3} \bmod p^3$$

(b)
$$\sum_{k=1}^{p-1} \frac{\sum_{j=1}^{k-1} 1/j}{k} \equiv -\frac{p}{3}\, B_{p-3} \bmod p^2$$

(c)
$$\sum_{k=1}^{p-1} \frac{1}{k^2} \equiv \frac{2p}{3}\, B_{p-3} \bmod p^2$$

(d)
$$\sum_{i+j+k=p} \frac{1}{ijk} \equiv -2\, B_{p-3} \bmod p$$

Bibliography

[1] *Random House Webster's Unabridged Dictionary*. Random House, 1999.

[2] Proof and beauty. *The Economist*, March 31 2005.

[3] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, New York, 1970.

[4] M. Agrawal, N. Kayal, and N. Saxena. PRIMES is in P. *Annals of Mathematics*, 160:781–793, 2004.

[5] Z. Ahmed. Definitely an integral. *Amer. Math. Monthly*, 109:670–671, 2002.

[6] I. Aliev. Siegel's lemma and sum-distinct sets. 2005. preprint.

[7] J. Alvarez, M. Amadis, G. Boros, D. Karp, V. Moll, and L. Rosales. An extension of a criterion for unimodality. *Elec. Jour. Comb.*, 8:1–7, 2001.

[8] A. Apelblat. *Tables of Integrals and Series*. Verlag Harry Deutsch, Thun; Frankfurt am Main, 1996.

[9] T. Arens. Why linear sampling works. *Inverse Problems*, 20:163–173, 2004.

[10] K. E. Atkinson. *An Introduction to Numerical Analysis*. John Wiley and Sons, Hoboken, NJ, 1989.

[11] D. H. Bailey. Integer relation detection. *Computing in Science and Engineering*, 2(1):24–28, 2000.

[12] D. H. Bailey and J. M. Borwein. Effective error bounds in Euler-Maclaurin based quadrature schemes. `http://crd.lbl.gov/~dhbailey/dhbpapers/em-error.pdf`, 2005.

[13] D. H. Bailey and J. M. Borwein. Experimental Mathematics: examples, methods and implications. *Notices Amer. Math. Soc.*, 52(5):502–514, 2005.

[14] D. H. Bailey and J. M. Borwein. Highly parallel, high-precision numerical integration. `http://crd.lbl.gov/~dhbailey/dhbpapers/quadparallel.pdf`, 2005.

[15] D. H. Bailey, J. M. Borwein, and R. E. Crandall. On the Khintchine constant. *Mathematics of Computation*, 66:417–431, 1997.

[16] D. H. Bailey, J. M. Borwein, and R. Girgensohn. Experimental evaluation of Euler sums. *Experimental Mathematics*, 3:17–30, 1994.

[17] D. H. Bailey, J. M. Borwein, V. Kapoor, and E. Weisstein. Ten computational challenge problems. *Amer. Math. Monthly*, 2006. In press.

[18] D. H. Bailey, P. B. Borwein, and S. Plouffe. On the rapid computation of various polylogarithmic constants. *Mathematics of Computation*, 66:903–913, 1997.

[19] D. H. Bailey and D. J. Broadhurst. Parallel integer relation detection: techniques and applications. *Mathematics of Computation*, 70:1719–1736, 2000.

[20] D. H. Bailey, Y. Hida, X. S. Li, and B. Thompson. ARPREC: An Arbitrary Precision Computation Package. `http://crd.lbl.gov/~dhbailey/dhbpapers/arprec.pdf`, 2002.

[21] D. H. Bailey and X. S. Li. A comparison of three high-precision quadrature schemes. `http://crd.lbl.gov/~dhbailey/dhbpapers/quadrature.pdf`, 2002.

[22] D.H. Bailey and J.M. Borwein. Experimental Mathematics: recent developments and future outlook. In B. Engquist and W. Schmid, editors, *Mathematics Unlimited—*2001 *and Beyond*, volume 1, pages 51–66. Springer-Verlag, 2000.

[23] W. N. Bailey. *Generalized Hypergeometric Series*. Cambridge University Press, Cambridge, 1935.

[24] G. V. Balakin, V. F. Kolchin, and V. I. Khokhlov. Hypercycles in a random hypergraph. *Diskret. Mat.*, 3(3):102–108, 1991.

[25] A. Baouche and S. Dubuc. La non-dérivabilité de la fonction de Weierstrass. *L'Enseignement Mathématique*, 38:89–94, 1992.

[26] A. Barnikov. *A Course in Convexity*. Amer. Math. Soc., 2002.

[27] D. Berlinski. Ground zero: a review of The pleasures of counting, by T. W. Koerner. *The Sciences*, pages 37–41, July/August 1997.

[28] B. Berndt. *Ramanujan's Notebooks, Part I*. Springer-Verlag, New York, 1985.

[29] P. C. Berry, A. D. Falkoff, and K. E. Iverson. Using the computer to compute: a direct but neglected approach to teaching Mathematics. *IBM New York Scientific Center Technical Report*, (320-2988), May, 1970.

[30] F. Beukers. A note on the irrationality of $\zeta(2)$ and $\zeta(3)$. *Bull. London Math. Soc.*, 11:268–272, 1979.

[31] M. G. Beumer. Some special integrals. *Amer. Math. Monthly*, 68:645–647, 1961.

[32] D. Bierens de Haan. *Nouvelles tables d'integrales definies*. Hafner, New York, 2nd edition, 1957.

[33] G. Boese and W. Lutter. Stetige, nirgends differenzierbare Funktionen und nicht rektifizierbare Kurven. *Mathematische Semesterberichte*, 30:228–254, 1983.

[34] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[35] G. Boros and V. Moll. A criterion for unimodality. *Elec. Jour. Comb.*, 6:1–6, 1999.

[36] G. Boros and V. Moll. An integral hidden in Gradshteyn and Ryzhik. *Jour. Comp. Applied Math.*, 106:361–368, 1999.

[37] G. Boros and V. Moll. The double square root, Jacobi polynomials and Ramanujan's master theorem. *Jour. Comp. Applied Math.*, 130:337–344, 2001.

[38] G. Boros and V. Moll. Landen transformation and the integration of rational functions. *Math. Comp.*, 71:649–668, 2001.

[39] G. Boros and V. Moll. *Irresistible Integrals*. Cambridge University Press, New York, 1st edition, 2004.

[40] G. Boros, V. Moll, and J. Shallit. The 2-adic valuation of the coefficients of a polynomial. *Scientia*, 7:37–50, 2001.

[41] E. J. Borowski and J. M. Borwein. *Dictionary of Mathematics, 2nd revised edition*. Harper Collins, 2002.

[42] D. Borwein, J. M. Borwein, and R. Girgensohn. Explicit evaluation of Euler sums. *Proceedings of the Edinburgh Mathematical Society*, 38:273–294, 1995.

[43] J. M. Borwein. The 100 digit challenge: an extended review. *Math Intelligencer*, 2005.

[44] J. M. Borwein and D. H. Bailey. *Mathematics by Experiment: Plausible Reasoning in the 21st century*. A K Peters Ltd, Natick, MA, 2003.

[45] J. M. Borwein, D. H. Bailey, and R. Girgensohn. *Experimentation in Mathematics: Computational Paths to Discovery*. A K Peters Ltd, Natick, MA, 2003.

[46] J. M. Borwein and P. B. Borwein. *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*. CMS Series of Monographs and Advanced books in Mathematics, John Wiley, Hoboken, NJ, 1987.

[47] J. M. Borwein, P. B. Borwein, and D. A. Bailey. Ramanujan, modular equations and pi or how to compute a billion digits of pi. *Amer. Math. Monthly*, 96:201–219, 1989. Reprinted in *Organic Mathematics Proceedings*, (`http://www.cecm.sfu.ca/organics`), April 12, 1996. Print version: *CMS/AMS Conference Proceedings*, **20** (1997), ISSN: 0731-1036.

[48] J. M. Borwein and D. H. Bradley. 32 Goldbach variations. *Int. J. Num. Theory*, 2005. in press.

[49] J. M. Borwein and D. M. Bradley. Searching symbolically for Apéry-like formulae for values of the Riemann zeta function. *ACM SIGSAM Bulletin of Algebraic and Symbolic Manipulation*, 30(2):2–7, 1996.

[50] J. M. Borwein and D. M. Bradley. Empirically determined Apéry-like formulae for $\zeta(4n+3)$. *Experimental Mathematics*, 6(3):181–194, 1997.

[51] J. M. Borwein and D. M Bradley. On two fundamental identities for Euler sums. Preprint, February 2005.

[52] J. M. Borwein, D. M. Bradley, D. J. Broadhurst, and P. Lisoněk. Special values of multidimensional polylogarithms. *Transactions of the American Mathematical Society*, 353:907–941, 2001.

[53] J. M. Borwein and D. Broadhurst. Determination of rational Dirichlet-zeta invariants of hyperbolic manifolds and Feynman knots and links. `http://arxiv.org/hep-th/9811173`, 1998.

[54] J. M. Borwein, D. J. Broadhurst, and J. Kamnitzer. Central binomial sums, multiple Clausen values, and zeta values. *Experimental Mathematics*, 10:25–41, 2001.

[55] J. M. Borwein and M. Chamberland. Integer powers of arcsin. *Mathematics Magazine*, April 2005. D-drive preprint 288.

[56] J. M. Borwein and R. Crandall. On the Ramanujan AGM fraction. Part II: the complex-parameter case. *Experimental Mathematics*, 13:287–296, 2004.

[57] J. M. Borwein, R. Crandall, and G. Fee. On the Ramanujan AGM fraction. part I: the real-parameter case. *Experimental Mathematics*, 13:275–286, 2004.

[58] J. M. Borwein and R. Girgensohn. Addition theorems and binary expansions. *Canadian Journal of Mathematics*, 47:262–273, 1995.

[59] J.M. Borwein and P.B. Borwein. Challenges for mathematical computing. *Computing in Science & Engineering*, 3:48–53, 2001.

[60] J.M. Borwein and R. Corless. Emerging tools for Experimental Mathematics. *Amer. Math. Monthly*, 106:889–909, 1999.

[61] J.M. Borwein and A.S. Lewis. *Convex Analysis and Nonlinear Optimization*. CMS Books. Springer-Verlag, 2000.

[62] J.M. Borwein and T.S. Stanway. Knowledge and community in mathematics. *The Mathematical Intelligencer*, 2005.

[63] J.M. Borwein and Q.J. Zhu. *Techniques of Variational Analysis*. CMS Books. Springer-Verlag, 2005.

[64] J.M. Borwein, I. J. Zucker, and J. Boersma. The evaluation of character Euler double sums. *The Ramanujan Journal*, 2005. accepted, CoLab Preprint 260.

[65] P. Borwein. An Efficient Algorithm for the Riemann Zeta Function. `http://www.cecm.sfu.ca/preprints/1995pp.html`, 1995.

[66] D. J. Broadhurst. Massive 3-loop feynman diagrams reducible to sc* primitives of algebras of the sixth root of unity. http://lanl.arxiv.org/abs/hep-th/9803091, 1998.

[67] D. J. Broadhurst, J. A. Gracey, and D. Kreimer. Beyond the triangle and uniqueness relations: non-zeta counterterms at large N from positive knots.

[68] D. J. Broadhurst and D. Kreimer. Association of multiple zeta values with positive knots via Feynman diagrams up to 9 loops. *Physics Letters*, B383:403–412, 1997.

[69] J. K. Brown. Solid tools for visualizing science. *Science*, pages 1136–37, November 19 2004.

[70] N. J. Calkin. Dependent sets of constant weight binary vectors. *Combin. Probab. Comput.*, 6(3):263–271, 1997.

[71] E. R. Canfield, Paul Erdős, and Carl Pomerance. On a problem of Oppenheim concerning "factorisatio numerorum". *J. Number Theory*, 17(1):1–28, 1983.

[72] G. J. Chaitin. Thoughts on the Riemann hypothesis. *Math. Intelligencer*, 26(1):4–7, 2004.

[73] M. Chamberland and V. Moll. Dynamics of the degree six Landen transformation. *Discrete and Dynamical Systems*, 2006. Preprint.

[74] M. Choi. Tricks or treats with the Hilbert matrix. *American Mathematical Monthly*, 90:301–312, 1983.

[75] T. Chow. What is a closed-form number? *Amer. Math. Monthly*, 106:440–448, 1999.

[76] Ji Chung-Gang. A simple proof of a curious congruence by Zhao. *Proceedings of the American Mathematical Society*, 133:3469–3472, 2005.

[77] D. Colton, H. Haddar, and M. Piana. The linear sampling method in inverse electromagnetic scattering theory. *Inverse Problems*, 19:S105–S137, 2003.

[78] D. Colton and A. Kirsch. A simple method for solving inverse scattering problems in the resonance region. *Inverse Problems*, 12(4):383–93, 1996.

[79] D. Colton and R. Kress. Eigenvalues of the far field operator for the Helmholtz equation in an absorbing medium. *SIAM J. Appl. Math.*, 55:1742–1735, 1995.

[80] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer-Verlag, New York, 2nd edition, 1998.

[81] D. Colton and R. Kress. On the denseness of Herglotz wave functions and electromagnetic Herglotz pairs in Sobolev spaces. *Math. Meth. Appl. Sci*, 24:1289–1303, 2001.

[82] D. Colton and P. Monk. A novel method for solving the inverse scattering problem for time-harmonic waves in the resonance region. *SIAM J. Appl. Math.*, 45:1039–53, 1985.

[83] D. Colton and P. Monk. A novel method for solving the inverse scattering problem for time-harmonic waves in the resonance region II. *SIAM J. Appl. Math.*, 46:506–23, 1986.

[84] D. Colton and P. Monk. A linear sampling method for the detection of leukimia using microwaves. *SIAM J. Appl. Math*, 58(3):926–941, 1998.

[85] D. Colton and L. Päivärinta. Transmission eigenvalues and a problem of Hans Lewy. *J. Comp. and Appl. Math.*, 117:91–104, 2000.

[86] D. Colton, M. Piana, and R. Potthast. A simple method using Morozov's discrepancy principle for solving inverse scattering problems. *Inverse Problems*, 13(6):1477–1493, 1997.

[87] L. Comtet. *Advanced Combinatorics: The Art of Finite and Infinite Expansion*. D. Reidel Publishing, Dordrecht, Holland, 1974.

[88] R. E. Crandall and J. P. Buhler. On the evaluation of Euler sums. *Experimental Mathematics*, 3:275–284, 1994.

[89] R. E. Crandall and J. S. Papadopoulos. On the implementation of the AKS-class primality tests. 2003.

[90] R. E. Crandall and C. Pomerance. *Prime Numbers: A Computational Perspective*. Springer-Verlag, Heidelberg, 2001.

[91] I. Damgard, P. Landrock, and C. Pomerance. Average case error estimates for the strong probable prime test. *Mathematics of Computation*, 61:177–194, 1993.

[92] I. Daubechies and J. C. Lagarias. Two-scale difference equations. I. Existence and global regularity of solutions. *SIAM Journal of Mathematical Analysis*, 22:1338–1410, 1991.

[93] I. Daubechies and J. C. Lagarias. Two-scale difference equations. II. Local regularity, infinite products of matrices, and fractals. *SIAM Journal of Mathematical Analysis*, 23:1031–1079, 1992.

[94] F. De-Jun and Y. Wang. Bernoulli convolutions associated with certain non-Pisot numbers. *Advances in Mathematics*, 187:173–194, 2004.

[95] G. de Rham. Sur certaines équations fonctionnelles. In *École Polytechnique de l'Université de Lausanne, Ouvrage publié à l'occasion de son Centennaire*, pages 95–97. 1953.

[96] G. de Rham. Sur une courbe plane. *Journal de Mathématiques Pures et Appliquées*, 35:25–42, 1956.

[97] K. Devlin. *Mathematics: the Science of Patterns*. Owl Books, 1996.

[98] J. Dewey. *Influence of Darwin on Philosophy and Other Essays*. Prometheus Books, 1997.

[99] I. Doust, M. D. Hirschhornand, and J. Ho. Trigonometric identities, linear algebra, and computer algebra. *Amer. Math. Monthly*, 112(2):155–164, 2005.

[100] P. du Bois-Reymond. Versuch einer Classification der willkürlichen Functionen reeller Argumente nach ihren Aenderungen in den kleinsten Intervallen. *Journal für die reine und angewandte Mathematik (Crelle's Journal)*, 79:21–37, 1875.

[101] M. Dummett. *Frege: Philosophy of Language*. Harvard University Press, 1973.

[102] P. Erdős. On a Family of Symmetric Bernoulli Convolutions. *Transactions of the American Mathematical Society*, 61:974–976, 1939.

[103] P. Erdös and M. Kac. The Gaussian law of errors in the theory of additive number theoretic functions. *Amer. J. Math.*, 62:738–742, 1940.

[104] P. Ernest. *Social Constructivism As a Philosophy of Mathematics*. State University of New York Press, 1998.

[105] G. Eskin. The inverse scattering problem in two dimensions at fixed frequency. *Comm. in Part. Diff. Eq.*, 26(5&6):1055–1090, 2001.

[106] O. Espinosa and V. Moll. On some definite integrals involving the hurwitz zeta function. part 1. *The Ramanujan Journal*, 6:159–188, 2002.

[107] O. Espinosa and V. Moll. On some definite integrals involving the hurwitz zeta function. part 2. *The Ramanujan Journal*, 6:449–468, 2002.

[108] O. Espinosa and V. Moll. The evaluation of Tornheim double sums. Part 1. *Journal of Number Theory*, 2005. In press.

[109] S. Wagon F. Bornemann, D. Laurie and J. Waldvogel. *The SIAM 100 Digit Challenge: A Study in High-Accuracy Numerical Computing*. SIAM, Philadelphia, 2004.

[110] G. Faber. Über stetige Funktionen I. *Mathematische Annalen*, 66:81–93, 1908.

[111] G. Faber. Über stetige Funktionen II. *Mathematische Annalen*, 69:372–443, 1910.

[112] C. Ferguson. *Helaman Ferguson: Mathematics in Stone and Bronze*. Meridian Creative Group, Erie, PA, 1994.

[113] H. R. P. Ferguson, D. H. Bailey, and S. Arno. Analysis of PSLQ, An Integer Relation Finding Algorithm. *Mathematics of Computation*, 68:351–369, 1999.

[114] S. R. Finch. *Mathematical Constants*. Cambridge University Press, 2003.

[115] G. Forsythe. Pitfalls in computation, or why a math book isn't enough. *Amer. Math. Monthly*, 77:931–956, 1970.

[116] G. Freud. Über trigonometrische Approximation und Fouriersche Reihen. *Mathematische Zeitschrift*, 78:252–262, 1962.

[117] A. Garsia. Arithmetic Properties of Bernoulli Convolutions. *Transactions of the American Mathematical Society*, 102:409–432, 1962.

[118] J. Gerver. The differentiability of the Riemann function at certain rational multiples of π. *American Journal of Mathematics*, 92:33–55, 1970.

[119] J. Gerver. More on the differentiability of the Riemann function. *American Journal of Mathematics*, 93:33–41, 1971.

[120] R. Girgensohn. Functional equations and nowhere differentiable functions. *Aequationes Mathematicae*, 46:243–256, 1993.

[121] R. Girgensohn. Nowhere differentiable solutions of a system of functional equations. *Aequationes Mathematicae*, 47:89–99, 1994.

[122] K. Gödel. Some basic theorems on the foundations. In *Collected Works, Vol. III. Unpublished essays and lectures.* Oxford University Press, New York, 1995.

[123] J. Grabiner. Newton, Maclaurin, and the authority of Mathematics. *Amer. Math. Monthly*, pages 841–852, December 2004.

[124] I. S. Gradshteyn and I.i M. Rhyzik. *Table of Integrals, Series, and Products*. Edited by A. Jeffrey and D. Zwillinger. Academic Press, New York, 6th edition, 2000.

[125] R. Graham, D. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison Wesley, Boston, 2nd edition, 1994.

[126] W. Grobner and N. Hofreiter. *Integraltafel*, volume 1. Springer-Verlag, Vienna, 1st edition, 1949.

[127] W. Grobner and N. Hofreiter. *Integraltafel*, volume 2. Springer-Verlag, Vienna, 1st edition, 1950.

[128] J. Guillera. Some Binomial Series Obtained by the WZ-Method. *Advances in Applied Mathematics*, 29:599–603, 2002.

[129] J. Guillera. About a new kind of Ramanujan-type series. *Experimental Mathematics*, 12(4):507–510, 2003.

[130] J. Guillera. Generators of some Ramanujan formulas. *The Ramanujan Journal*, to appear.

[131] G. H. Hardy. Weierstrass's non-differentiable function. *Transactions of the American Mathematical Society*, 17:301–325, 1916.

[132] G. H. Hardy. Prolegomena to a chapter on inequalities. In *Collected Papers*, volume 2, pages 471–489. Oxford University Press, 1967.

[133] G. H. Hardy, J. E. Littlewood, and G. Pólya. "hilbert's double series theorem" and "On Hilbert's inequality. In *Inequalities*, pages 226–227 and 308–309. Cambridge University Press, 2nd edition, 1988. §9.1 and Appendix III.

[134] G. H. Hardy and S. Ramanujan. The normal number of prime factors of a number n. *Quarterly Journal of Mathematics*, 48:76–92, 1917.

[135] M. Hata. On Weierstrass's non-differentiable function. *Comptes Rendus des Séances de l'Académie des Sciences. Série I. Mathématique*, 307:119–123, 1988.

[136] M. Hata. Singularities of the Weierstrass type functions. *Journal d'Analyse Mathématique*, 51:62–90, 1988.

[137] R. Hersch. Fresh breezes in the philosophy of mathematics. *Amer. Math. Monthly*, pages 589–594, August 1995.

[138] R. Hersh. *What is Mathematics Really?* Oxford University Press, 1999.

[139] M. D. Hirschhorn. A difficult limit. *Australian Math. Gazette*, 31:125–127, 2004.

[140] M. D. Hirschhorn. Ramanujan and Fermat's last theorem. *Australian Math. Gazette*, 31:256–257, 2004.

[141] A. Holroyd, T. Liggett, and D. Romik. Integrals, partitions and cellular automata. *Trans. Amer. Math. Soc.*, 356:3349–3368, 2004.

[142] M. Ikehata. Enclosing a polygonal cavity in a two-dimensional bounded domain from Cauchy data. *Inverse Problems*, 15:1231–41, 1999.

[143] G. J. O. Jameson. Hilbert's inequality on ℓ^2. Lecture Notes.

[144] B. Jessen and A. Wintner. Distribution functions and the Riemann zeta function. *Transactions of the American Mathematical Society*, 38:48–88, 1935.

[145] J. P. Kahane. Lacunary Taylor and Fourier series. *Bulletin of the American Mathematical Society*, 70:199–213, 1964.

[146] R. Kershner and A. Wintner. On symmetric Bernoulli convolutions. *American Journal of Mathematics*, 57:541–548, 1935.

[147] A. Kirsch. Characterization of the shape of a scattering obstacle using the spectral data of the far field operator. *Inverse Problems*, 14:1489–1512, 1998.

[148] A. Kirsch. Factorization of the far field operator for the inhomogeneous medium case and an application to inverse scattering theory. *Inverse Problems*, 15:413–429, 1999.

[149] Donald Knuth. American Mathematical Monthly Problem 10832. *American Mathematical Monthly*, November 2002.

[150] C. Koch. "thinking about the conscious mind," a review of john r. searle's *Mind. A Brief Introduction,* oxford university press, 2004. *Science*, pages 979–980, November 5 2004.

[151] M. Koecher. Letter (German). *Mathematics Intelligencer*, 2(2):62–64, 1979/80.

[152] G. Kolada. In Math, computers don't lie. or do they? *New York Times*, April 6 2004.

[153] V. F. Kolchin and V. I. Khokhlov. On the number of cycles in a random nonequiprobable graph. *Diskret. Mat.*, 2(3):137–145, 1990.

[154] I. Kotsireas and K. Karamanos. Exact computation of the bifurcation point b_4 of the logistic map and the Bailey-Broadhurst conjectures. pages 2417–2423, 2004.

[155] R. Kress. *Linear integral equations*, volume 82 of *Applied Mathematical Sciences*. Springer Verlag, New York, 2 edition, 1999.

[156] T.S. Kuhn. *The Structure of Scientific Revolutions, 3rd ed.* University of Chicago Press, 1996.

[157] I. Lakatos. *Proofs and Refutations*. Cambridge Univ. Press, 1976.

[158] G. Lakoff and R. E. Nunez. *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. Basic Books, 2001.

[159] E. Landau. *Handbuch der Lehre von der Verteilung der Primzahlen. 2 Bände*. Chelsea Publishing Co., New York, 1953. 2d ed, With an appendix by Paul T. Bateman.

[160] R. Larson, B. Edwards, and D. Heyd. *Calculus of a Single Variable*. Houghton Mifflin Company, Boston-New York, sixth edition, 1998.

[161] A. Lasota and M. C. Mackey. *Chaos, Fractals and Noise: Stochastic Aspects of Dynamics*. Springer-Verlag, Heidelberg, 1994.

[162] A. M. Legendre. *Theorie des Nombres*. Firmin Didot Freres, Paris, 1830.

[163] D. H. Lehmer. Computer technology applied to the theory of numbers. In W. J. LeVeque, editor, *Studies in Number Theory*, pages 117–151. Prentice-Hall, New York, 1969.

[164] R. Leis. *Initial Boundary Value Problems in Mathematical Physics*. B.G. Teubner, Stuttgart, 1986.

[165] H. W. Lenstra Jr. Primality testing with cyclotomic rings. *preprint*, 2002.

[166] L. Lewin. *Dilogarithms and Associated Functions*. Macdonald, London, 1958.

[167] L. Lewin. *Polylogariths and Associated Functions*. North Holland, New York, Oxford, 1981.

[168] R. C. Lewontin. In the beginning was the word. *Science (*Human Genome Issue*)*, pages 1263–1264, February 16 2001.

[169] J. Little. On the zeroes of two families of polynomials arising from certain rational integrals. *Rocky Mountain Journal*, 35, 2005.

[170] D. R. Luke. Multifrequency inverse obstacle scattering: the point source method and generalized filtered backprojeciton. *Math. and Computers in Simulation*, 66(4–5):297–314, 2004.

[171] D. R. Luke, J. V. Burke, and R. G. Lyon. Optical wavefront reconstruction: Theory and numerical methods. *SIAM Rev.*, 44:169–224, 2002.

[172] D. R. Luke and R. Potthast. The no response test – a sampling method for inverse scattering problems. *SIAM J. Appl. Math.*, 63(4):1292–1312, 2003.

[173] D. Manna and V. Moll. Rational Landen transformations on $\mathbb{R}$. 2006. Submitted for puiblication.

[174] D. Manna and V. Moll. A simple example of a new class of Landen transformations. *American Mathematical Montly*, 2006. In press.

[175] R. D. Mauldin and K. Simon. The equivalence of some Bernoulli convolutions to Lebesgue measure. *Proceedings of the American Mathematical Society*, 126:2733–2736, 1998.

[176] H. P. McKean and V. Moll. *Elliptic Curves: Function Theory, Geometry, Arithmetic*. Cambridge University Press, New York, NY, 1997.

[177] Peter B. Medawar. *Advice to a Young Scientist*. HarperCollins.

[178] L. Monier. Evaluation and comparison of two efficient probabilistic primality testing algorithms. *Theoretical Computer Science*, 12:97–108, 1980.

[179] P. Monk. *Finite Element Methods for Maxwell's Equations*. Clarendon Press, Oxford, 2003.

[180] H. Montgomery. *Ten lectures on the interface between analytic number theory and Harmonic Analysis*. Amer. Math. Soc., 1994. CBMS Regional Conference Series in Mathematics.

[181] M. A. Morrison and John J. Brillhart. A method of factoring and the factorization of F_7. *Math. Comp.*, 29:183–205, 1975. Collection of articles dedicated to Derrick Henry Lehmer on the occasion of his seventieth birthday.

[182] A. Nachman. Reconstructions from boundary measurements. *Ann. Math.*, 128:531–576, 1988.

[183] A. Nachman. Global uniqueness for a two-dimensional inverse boundary value problem. *The Annals of Mathematics*, 143(1):71–96, 1996.

[184] I. Nemes, M. Petkovsek, H. Wilf, and D. Zeilberger. How to do MONTHLY problems with your computer. *Amer. Math. Monthly*, 104:505–519, 1997.

[185] R. Novikov. Multidimensional inverse spectral problems for the equation $-\triangle\phi + (v(x) - eu(x))\phi$. *Trans. Func. Anal. and its Appl.*, 22:263–272, 1988.

[186] R. Novikov. The inverse scattering problem on a fixed energy for the two-dimensional Schrödinger operator. *J. Funct. Anal.*, 1992:409–463, 1992.

[187] A. Odlyzko. The 10^{22}-nd zero of the Riemann zeta function. Dynamical, spectral, and arithmetic zeta functions. *Contemp. Math.*, 290:139–144, 2001.

[188] Y. Peres, W. Schlag, and B. Solomyak. Sixty years of Bernoulli convolutions. In C. Bandt, S. Graf, and M. Zaehle, editors, *Progress in Probability: Fractals and Stochastics II*, pages 39–65. Birkhäuser, 2000.

[189] Y. Peres and B. Solomyak. Absolute continuity of Bernoulli convolutions, a simple proof. *Mathematical Research Letters*, 3:231–239, 1996.

[190] M. Petkovsek, H. Wilf, and D. Zeilberger. *A=B*. A. K. Peters, Ltd., 1st edition, 1996.

[191] I. Pinelis. On l'Hopital type rules for monotonicty. *Jour. Inequal. Pure Appl. Math.*, 2005. in press.

[192] M. Planck. *Scientific Autobiography and Other Papers, trans. F. Gaynor.* New York, 1949.

[193] H. Poincaré. *Mathematical definitions in education.* 1904.

[194] H. Poincaré. Mathematical invention. In Raymond Ayoub, editor, *Musings of the Masters*, pages 20–30. MAA, 2004.

[195] G. Polya. *Mathematical Discovery: On Understanding, Learning, and Teaching Problem Solving (Combined Edition).* Wiley & Sons, New York, 1981.

[196] C. Pomerance. A tale of two sieves. *Notices Amer. Math. Soc.*, 43(12):1473–1485, 1996.

[197] R. Potthast. *Point Sources and Multipoles in Inverse Scattering Theory.* Chapman & Hall, London, 2001.

[198] R. Potthast, J. Sylvester, and S. Kusiak. A 'range test' for determining scatterers with unknown physical properties. *Inverse Problems*, 19(3):533–547, 2003.

[199] R. Preston. The Mountains of Pi. *New Yorker*, Mar. 2, 1992.

[200] P. Preuss. Mathematics: The Statue in the Stone. *Berkeley Lab Research Review*, 23:12–13, Fall 2000.

[201] I. E. Pritsker and R. S. Varga. The Szego curve, zero distribution and weighted approximation. *Trans. Amer. Math. Soc.*, 349:4085–4105, 1997.

[202] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev. *Integrals and Series.* Gordon and Breach Science, Heidelberg, 1992.

[203] M. Rabin. Probabilistic algorithms for testing primality.

[204] A. G. Ramm. Recovery of the potential from fixed energy scattering data. *Inverse Problems*, 4:877–886, 1988.

[205] M. Reed and B. Simon. *Methods of Modern Mathematical Physics*, volume III. Scattering Theory. Academic Press, New York, 1979.

[206] E. Regis. *Who got Einstein's office?* Addison Wesley, 1988.

[207] S. Robinson. New method said to solve key problem in Math. *New York Times*, Aug. 8 2002.

[208] W. Rudin. *Functional Analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill, San Francisco, 1991.

[209] R. Salem. A remarkable class of algebraic integers, proof of a conjecture by Vijayaraghavan. *Duke Mathematical Journal*, 11:103–108, 1944.

[210] J. Schauder. Zur Theorie stetiger Abbildungen in Funktionalräumen. *Mathematische Zeitschrift*, 26:47–65, 1927.

[211] Wacław Sierpiński. Sur une système d'équations fonctionnelles, définissant une fonction avec un ensemble dense d'intervalles d'invariabilité. *Bulletin International de l'Academie des Sciences de Cracovie A*, pages 577–582, 1911.

[212] A. V. Sills and D. Zeilberger. Disturbing the Dyson conjecture (in a good way). 2005. Preprint.

[213] B. Solomyak. On the random series $\sum \pm\lambda^i$ (an Erdős problem). *Annals of Mathematics*, 142:611–625, 1995.

[214] A. Sommerfeld. *Optics*. Academic Press, New York, 1954.

[215] J. M. Steele. *The Cauchy-Schwarz Master Class*. MAA Problem Books Series. MAA, 2004.

[216] Z. Sun and G. Uhlmann. Generic uniqueness for an inverse boundary value problem. *Duke Math. J.*, 62:131–155, 1991.

[217] Z Sun and G. Uhlmann. Inverse scattering for singular potentials in two dimensions. *Trans. Amer. Math. Soc.*, 338:363–374, 1993.

[218] D'Arcy Thompson. *On Growth and Form*. Dover Publications, 1992.

[219] J. P. Tignol. *Galois's Theory of Algebraic Equations*. World Scientific Co. Pte. Ltd, Singapore, 2nd edition, 2002.

[220] C. Tracy. Integrals for Ising susceptibilty expansion. Personal communication.

[221] T.Takagi. A simple example of the continuous function without derivative. *Proceedings of the Physico-Mathematical Society of Japan (Ser. II)*, 1:176–177, 1903.

[222] T. Tymoczko. The four-color problem and its philosophical significance. *J. Philos.*, 76:57–83, 1979.

[223] T. Tymoczko. Computer, proofs and mathematics: a philosophical investigation of the four-color problem. *Math. Magazine*, 53:131–138, 1980.

[224] A. van der Poorten. A proof that Euler missed: Apéry's proof of the irrationality of $\zeta(3)$. *Mathematical Intelligencer*, 1(4):195–203, 1978/79.

[225] I. Vardi. Integrals, an introduction to analytic number theory. *Amer. Math. Mon.*, 95:308–315, 1988.

[226] S. Wagon. Fourteen proofs of a result about tiling a rectangle. *American Mathematical Monthly*, 94(7):601–617, 1987.

[227] K. Weierstraß. Über continuirliche Functionen eines reellen Arguments, die für keinen Werth des letzteren einen bestimmten Differentialquotienten besitzen. (Gelesen in der Königl. Akademie der Wissenschaften am 18. Juli 1872). In *Mathematische Werke von Karl Weierstraß, vol. 2*, pages 71–74. Mayer & Müller, Berlin, 1895.

[228] E. W. Weisstein. Lucky number of Euler. In *MathWorld–A Wolfram Web Resource*.

[229] P. Winkler. *Mathematical puzzles: a connosseur's collection*. A. K. Peters, 2004.

[230] A. Wintner. On convergent Poisson convolutions. *American Journal of Mathematics*, 57:827–838, 1935.

[231] E. Wirsing. Über die Zahlen, deren Primteiler einer gegebenen Menge angehören. *Arch. Math.*, 7:263–272, 1956.

[232] Jun Wu. On the Lévy constant for quadratic irrationals. *Proc. Amer. Math. Monthly*. e-published Dec. 15, 2005.

[233] B. Yandell. *The Honors Class*. A. K. Peters, 2002.

[234] D. Zagier. Values of zeta function and their applications. In *Proceedings of the First European Congress of Mathematics*, volume 2, pages 497–512, 1994.

[235] D. Zeilberger. Identities in Search of Identity. *Theoretical Computer Science*, 117:23–38, 1993.

[236] H. Zimmer. *Computational problems, methods, and results in Algebraic Number Theory*. Springer-Verlag, New York, 1972.

Index