# EXPLORATORY EXPERIMENTATION AND COMPUTATION

DAVID H. BAILEY AND JONATHAN M. BORWEIN

Abstract. We believe the mathematical research community is facing a great challenge to re-evaluate the role of proof in light of recent developments. On one hand, the growing power of current computer systems, of modern mathematical computing packages, and of the growing capacity to data-mine on the Internet, has provided marvelous resources to the research mathematician. On the other hand, the enormous complexity of many modern capstone results such as the Poincaré conjecture, Fermat's last theorem, and the classification of finite simple groups has raised questions as to how we can better ensure the integrity of modern mathematics. Yet as the need and prospects for inductive mathematics blossom, the requirement to ensure the role of proof is properly founded remains undiminished.

## 1. Exploratory Experimentation

> The authors' thesis—once controversial, but now a commonplace—is that computers can be a useful, even essential, aid to mathematical research.—Jeff Shallit

Jeff Shallit wrote this in his recent review MR2427663 of [10]. As we hope to make clear, Shallit was entirely right in that many, if not most, research mathematicians now use the computer in a variety of ways to draw pictures, inspect numerical data, manipulate expressions symbolically, and run simulations. However, it seems to us that there has not yet been substantial and intellectually rigorous progress in the way mathematics is presented in research papers, textbooks and

*Date*: July 24, 2012.

Bailey: Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. Email: dhbailey@lbl.gov. This work was supported by the Director, Office of Computational and Technology Research, Division of Mathematical, Information, and Computational Sciences of the U.S. Department of Energy, under contract number DE-AC02-05CH11231. Borwein: Centre for Computer Assisted Research Mathematics and its Applications (CARMA), University of Newcastle, Callaghan, NSW 2308, Australia. Email: jonathan.borwein@newcastle.edu.au.

classroom instruction, or in how the mathematical discovery process is organized.

1.1. **Mathematicians are humans.** We share with George Pólya (1887-1985) the view [25, vol. 2, p. 128] that while learned,

> intuition comes to us much earlier and with much less outside influence than formal arguments.

Pólya went on to reaffirm, nonetheless, that proof should certainly be taught in school.

We turn to observations, many of which have been fleshed out in coauthored books such *Mathematics by Experiment* [10], and *Experimental Mathematics in Action* [3], where we have noted the changing nature of mathematical knowledge and in consequence ask questions such as "How do we teach what and why to students?", "How do we come to believe and trust pieces of mathematics?", and "Why do we wish to prove things?" An answer to the last question is "That depends." Sometimes we wish insight and sometimes, especially with subsidiary results, we are more than happy with a certificate. The computer has significant capacities to assist with both.

Smail [27, p. 113] writes:

> the large human brain evolved over the past 1.7 million years to allow individuals to negotiate the growing complexities posed by human social living.

As a result, humans find various modes of argument more palatable than others, and are more prone to make certain kinds of errors than others. Likewise, the well-known evolutionary psychologist Steve Pinker observes that language [24, p. 83] is founded on

> ...the ethereal notions of space, time, causation, possession, and goals that appear to make up a language of thought.

This remains so within mathematics. The computer offers scaffolding both to enhance mathematical reasoning, as with the recent computation of the Lie group $E_8$, see `http://www.aimath.org/E8/computerdetails.html`, and to restrain mathematical error.

1.2. **Experimental mathodology.** Justice Potter Stewart's famous 1964 comment, *"I know it when I see it"* is the quote with which *The Computer as Crucible* [13] starts. A bit less informally, by *experimental mathematics* we intend [10]:

(a) Gaining insight and *intuition*;
(b) *Visualizing* math principles;
(c) *Discovering* new relationships;
(d) *Testing* and especially *falsifying* conjectures;

(e) *Exploring* a possible result to see if it *merits* formal proof;
(f) *Suggesting* approaches for formal proof;
(g) *Computing* replacing lengthy hand derivations;
(h) *Confirming* analytically derived results.

Of these items (a) through (e) play a central role, and (f) also plays a significant role for us, but connotes computer-assisted or computer-directed proof and thus is quite distinct from *formal proof* as the topic of a special issue of these *Notices* in December 2008; see, e.g., [20].

1.2.1. *Digital integrity, I.* For us (g) has become ubiquitous, and we have found (h) to be particularly effective in ensuring the integrity of published mathematics. For example, we frequently check and correct identities in mathematical manuscripts by computing particular values on the LHS and RHS to high precision and comparing results—and then if necessary use software to repair defects.

As a first example, in a current study of "character sums" we wished to use the following result derived in [14]:

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{(-1)^{m+n-1}}{(2m-1)(m+n-1)^3} \tag{1.1}$$
$$\stackrel{?}{=} 4\,\mathrm{Li}_4\left(\frac{1}{2}\right)-\frac{51}{2880}\pi^4-\frac{1}{6}\pi^2\log^2(2)+\frac{1}{6}\log^4(2)+\frac{7}{2}\log(2)\zeta(3).$$

Here $\mathrm{Li}_4(1/2)$ is a polylogarithmic value. However, a subsequent computation to check results disclosed that whereas the LHS evaluates to $-0.872929289\ldots$, the RHS evaluates to $2.509330815\ldots$. Puzzled, we computed the sum, as well as each of the terms on the RHS (sans their coefficients), to 500-digit precision, then applied the "PSLQ" algorithm, which searches for integer relations among a set of constants [16]. PSLQ quickly found the following:

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{(-1)^{m+n-1}}{(2m-1)(m+n-1)^3} \tag{1.2}$$
$$= 4\,\mathrm{Li}_4\left(\frac{1}{2}\right)-\frac{151}{2880}\pi^4-\frac{1}{6}\pi^2\log^2(2)+\frac{1}{6}\log^4(2)+\frac{7}{2}\log(2)\zeta(3).$$

In other words, in the process of transcribing (1.1) into the original manuscript, "151" had become "51." It is quite possible that this error would have gone undetected and uncorrected had we not been able to computationally check and correct such results. This may not always matter, but it can be crucial.

With a current Research Assistant, Alex Kaiser at Berkeley, we have started to design software to refine and automate this process and

to run it before submission of any equation-rich paper. This semi-automated integrity checking becomes pressing when verifiable output from a symbolic manipulation might be the length of a Salinger novel. For instance, recently while studying expected radii of points in a hypercube [12], it was necessary to show the existence of a "closed form" for

$$J(t) \quad := \quad \int_{[0,1]^2} \frac{\log(t + x^2 + y^2)}{(1 + x^2)(1 + y^2)} \, \mathrm{d}x \, \mathrm{d}y. \tag{1.3}$$

The computer verification of [12, Thm. 5.1] quickly returned a 100000-character "answer" that could be numerically validated very rapidly to hundreds of places. A highly interactive process stunningly reduced a basic instance of this expression to the concise formula

$$J(2) = \frac{\pi^2}{8} \log 2 - \frac{7}{48}\zeta(3) + \frac{11}{24} \pi \, \mathrm{Cl}_2 \left(\frac{\pi}{6}\right) - \frac{29}{24} \pi \, \mathrm{Cl}_2 \left(\frac{5\pi}{6}\right), \tag{1.4}$$

where $\mathrm{Cl}_2$ is the *Clausen function* $\mathrm{Cl}_2(\theta) := \sum_{n \geq 1} \sin(n\theta)/n^2$ ($\mathrm{Cl}_2$ is the simplest non-elementary Fourier series). Automating such reductions will require a sophisticated simplification scheme with a very large and extensible knowledge base.

1.3. **Discovering a truth.** Giaquinto's [18, p. 50] attractive encapsulation

> In short, discovering a truth is coming to believe it in an independent, reliable, and rational way.

has the satisfactory consequence that a student can legitimately discover things already "known" to the teacher. Nor is it necessary to demand that each dissertation be absolutely original—only that it be independently discovered. For instance, a differential equation thesis is no less meritorious if the main results are subsequently found to have been accepted, unbeknown to the student, in a control theory journal a month earlier—provided they were independently discovered. Near-simultaneous independent discovery has occurred frequently in science, and such instances are likely to occur more and more frequently as the earth's "new nervous system" (Hillary Clinton's term in a recent policy address) continues to pervade research.

Despite the conventional identification of mathematics with deductive reasoning, Kurt Gödel (1906-1978) in his 1951 Gibbs Lecture said:

> If mathematics describes an objective world just like physics, there is no reason why inductive methods should not be applied in mathematics just the same as in physics.

He held this view until the end of his life despite—or perhaps because of—the epochal deductive achievement of his incompleteness results.

Also, we emphasize that many great mathematicians from Archimedes and Galileo—who reputedly said *"All truths are easy to understand once they are discovered; the point is to discover them."*—to Gauss, Poincaré, and Carleson have emphasized how much it helps to "know" the answer beforehand. Two millennia ago, Archimedes wrote, in the Introduction to his long-lost and recently reconstituted *Method* manuscript,

> For it is easier to supply the proof when we have previously acquired, by the method, some knowledge of the questions than it is to find it without any previous knowledge.

Archimedes' *Method* can be thought of as an uber-precursor to today's interactive geometry software, with the caveat that, for example, *Cinderella* actually does provide proof certificates for much of Euclidean geometry.

As 2006 Abel Prize winner Lennart Carleson describes in his 1966 ICM speech on his positive resolution of Luzin's 1913 conjecture (that the Fourier series of square-summable functions converge pointwise a.e. to the function), after many years of seeking a counterexample, he finally decided none could exist. He expressed the importance of this confidence as follows:

> The most important aspect in solving a mathematical problem is the conviction of what is the true result. Then it took 2 or 3 years using the techniques that had been developed during the past 20 years or so.

1.4. **Digital Assistance.** By *digital assistance*, we mean the use of:

(a) *Integrated mathematical software* such as *Maple* and *Mathematica*, or indeed MATLAB and their open source variants.
(b) *Specialized packages* such as CPLEX, PARI, SnapPea, Cinderella and MAGMA.
(c) *General-purpose programming languages* such as C, C++, and Fortran-2000.

(d) *Internet-based applications* such as: Sloane's Encyclopedia of Integer Sequences, the Inverse Symbolic Calculator,[1] Fractal Explorer, Jeff Weeks' Topological Games, or Euclid in Java.[2]
(e) *Internet databases and facilities* including Google, MathSciNet, arXiv, Wikipedia, MathWorld, MacTutor, Amazon, Amazon Kindle, and many more that are not always so viewed.

All entail data-mining in various forms. The capacity to consult the Oxford dictionary and Wikipedia instantly within Kindle dramatically changes the nature of the reading process. Franklin [17] argues that Steinle's "exploratory experimentation" facilitated by "widening technology" and "wide instrumentation," as routinely done in fields such as pharmacology, astrophysics, medicine, and biotechnology, is leading to a reassessment of what legitimates experiment; in that a "local model" is not now a prerequisite. Thus, a pharmaceutical company can rapidly examine and discard tens of thousands of potentially active agents, and then focus resources on the ones that survive, rather than needing to determine in advance which are likely to work well. Similarly, aeronautical engineers can, by means of computer simulations, discard thousands of potential designs, and submit only the best prospects to full-fledged development and testing.

Hendrik Sørenson [28] concisely asserts that experimental mathematics —as defined above—is following similar tracks with software such as *Mathematica*, *Maple* and MATLAB playing the role of wide instrumentation.

> These aspects of exploratory experimentation and wide instrumentation originate from the philosophy of (natural) science and have not been much developed in the context of experimental mathematics. However, I claim that e.g. the importance of wide instrumentation for an exploratory approach to experiments that includes concept formation also pertain to mathematics.

In consequence, boundaries between mathematics and the natural sciences and between inductive and deductive reasoning are blurred and becoming more so. (See also [2].) This convergence also promises some

[1]Most of the functionality of the ISC, which is now housed at `http://carma-lx1.newcastle.edu.au:8087`, is now built into the "identify" function of *Maple* starting with version 9.5. For example, the *Maple* command `identify(4.45033263602792)` returns $\sqrt{3}+e$, meaning that the decimal value given is simply approximated by $\sqrt{3}+e$.

[2]A cross-section of Internet-based mathematical resources is available at `http://ddrive.cs.dal.ca/~isc/portal/` and `http://www.experimentalmath.info`.

relief from the frustration many mathematicians experience when attempting to describe their proposed methodology on grant applications to the satisfaction of traditional hard scientists. We leave unanswered the philosophically-vexing if mathematically-minor question as to whether genuine mathematical experiments (as discussed in [10]) truly exist, even if one embraces a fully idealist notion of mathematical existence. It surely seems to us that they do.

## 2. Pi, Partitions and Primes

The present authors cannot now imagine doing mathematics without a computer nearby. For example, characteristic and minimal polynomials, which were entirely abstract for us as students, now are members of a rapidly growing box of concrete symbolic tools. One's eyes may glaze over trying to determine structure in an infinite family of matrices including

$$M_4 = \begin{bmatrix} 2 & -21 & 63 & -105 \\ 1 & -12 & 36 & -55 \\ 1 & -8 & 20 & -25 \\ 1 & -5 & 9 & -8 \end{bmatrix} \quad M_6 = \begin{bmatrix} 2 & -33 & 165 & -495 & 990 & -1386 \\ 1 & -20 & 100 & -285 & 540 & -714 \\ 1 & -16 & 72 & -177 & 288 & -336 \\ 1 & -13 & 53 & -112 & 148 & -140 \\ 1 & -10 & 36 & -66 & 70 & -49 \\ 1 & -7 & 20 & -30 & 25 & -12 \end{bmatrix}$$

but a command-line instruction in a computer algebra system will reveal that both $M_4^3 - 3M_4 - 2I = 0$ and $M_6^3 - 3M_6 - 2I = 0$. Likewise, more and more matrix manipulations are profitably, even necessarily, viewed graphically. As is now well known in numerical linear algebra, graphical tools are essential when trying to discern qualitative information such as the block structure of very large matrices. See, for instance, Figure 1.

Equally accessible are many matrix decompositions, the use of Groebner bases, Risch's decision algorithm (to decide when an elementary function has an elementary indefinite integral), graph and group catalogues, and others. Many algorithmic components of a *computer algebra system* are today extraordinarily effective compared with two decades ago, when they were more like toys. This is equally true of extreme-precision calculation—a prerequisite for much of our own work [8, 11, 9]. As we will illustrate, during the three decades that we have seriously tried to integrate computational experiments into research, we have experienced at least 12 Moore's law doublings of computer power and memory capacity [10, 13], which when combined with the utilization of highly parallel clusters (with thousands of processing cores) and

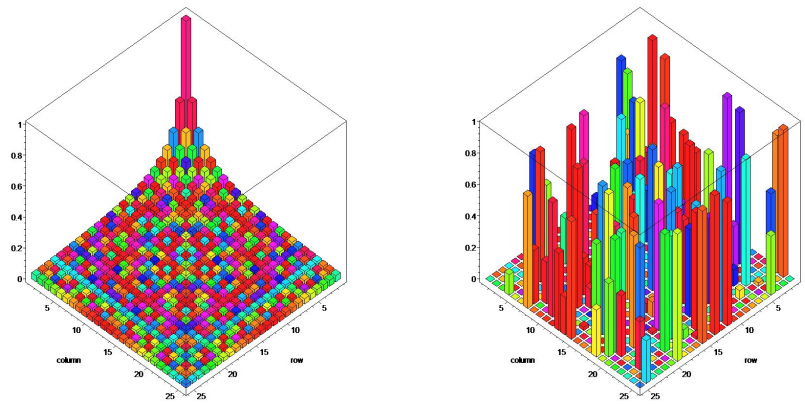


FIGURE 1. Plots of a $25 \times 25$ Hilbert matrix (L) and a matrix with 50% sparsity and random $[0, 1]$ entries (R).

fiber-optic networking, has resulted in six to seven orders of magnitude speedup for many operations.

2.1. **The partition function.** Consider the number of additive partitions, $p(n)$, of a natural number, where we ignore order and zeroes. For instance, $5 = 4+1 = 3+2 = 3+1+1 = 2+2+1 = 2+1+1+1 = 1+1+1+1+1$, so $p(5) = 7$. The ordinary generating function (2.1) discovered by Euler is

$$\sum_{n=0}^{\infty} p(n) q^n \quad = \quad \prod_{k=1}^{\infty} \left(1 - q^k\right)^{-1}. \tag{2.1}$$

(This can be proven by using the geometric formula for $1/(1-q^k)$ to expand each term and observing how powers of $q^n$ occur.)

The famous computation by MacMahon of $p(200) = 3972999029388$ at the beginning of the 20th century, done symbolically and entirely naively from (2.1) on a reasonable laptop, took 20 minutes in 1991 but only 0.17 seconds today, while the many times more demanding computation

$$p(2000) = 4720819175619413888601432406799959512200344166$$

took just two minutes in 2009. Moreover, in December 2008, Crandall was able to calculate $p(10^9)$ in three seconds on his laptop, using the Hardy-Ramanujan-Rademacher 'finite' series for $p(n)$ along with FFT methods. Using these techniques, Crandall was also able to calculate the probable primes $p(1000046356)$ and $p(1000007396)$, each of which has roughly 35000 decimal digits.

Such results make one wonder when easy access to computation discourages innovation: Would Hardy and Ramanujan have still discovered their marvelous formula for $p(n)$ if they had powerful computers at hand?

2.2. **Quartic algorithm for $\pi$.** Likewise, the record for computation of $\pi$ has gone from 29.37 *million* decimal digits in 1986, to 5 *trillion* digits in 2010. Since the algorithm below was used as part of each computation, it is interesting to compare the performance in each case: Set $a_0 := 6 - 4\sqrt{2}$ and $y_0 := \sqrt{2} - 1$, then iterate

$$\begin{aligned} y_{k+1} &= \frac{1-(1-y_k^4)^{1/4}}{1+(1-y_k^4)^{1/4}}, \\ a_{k+1} &= a_k(1+y_{k+1})^4 - 2^{2k+3}y_{k+1}(1+y_{k+1}+y_{k+1}^2). \end{aligned} \tag{2.2}$$

Then $a_k$ converges *quartically* to $1/\pi$—each iteration approximately quadruples the number of correct digits. Twenty-one full-precision iterations of (2.2), which was discovered on a 16K Radio Shack portable in 1983, produce an algebraic number that coincides with $\pi$ to well more than six trillion places. This scheme and the 1976 Salamin–Brent scheme [10, Ch. 3] have been employed frequently over the past quarter century. Here is a highly abbreviated chronology (based on `http://en.wikipedia.org/wiki/Chronology_of_computation_of_pi`).

- 1986: One of the present authors used (2.2) to compute 29.4 million digits of $\pi$. This required 28 hours on one CPU of the new Cray-2 at NASA Ames Research Center. Confirmation using the Salamin-Brent scheme took another 40 hours. This computation uncovered hardware and software errors on the Cray-2.
- Jan. 2009: Takahashi used (2.2) to compute 1.649 trillion digits (nearly 60,000 times the 1986 computation), requiring 73.5 hours on 1024 cores (and 6.348 Tbyte memory) of a Appro Xtreme-X3 system. Confirmation via the Salamin-Brent scheme took 64.2 hours and 6.732 Tbyte of main memory.
- Apr. 2009: Takahashi computed 2.576 trillion digits.
- Dec. 2009: Bellard computed nearly 2.7 trillion decimal digits (first in binary), using the Chudnovsky series given below in (2.10). This took 131 days on a single four-core workstation with lots of disk storage.This took 131 days, but he only used a single four-core workstation with lots of disk storage and even more human intelligence!
- Aug. 2010: Kondo and Yee computed 5 trillion decimal digits using the same formula (2.10) due to the Chudnovskys. This

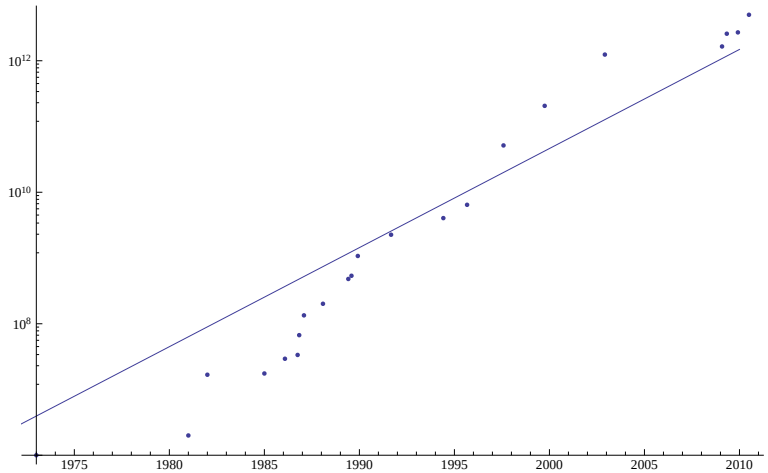


FIGURE 2. Plot of $\pi$ calculations, in digits (dots), compared with the long-term slope of Moore's Law (line).

was first done in binary, then converted to decimal. The binary digits were confirmed by computing 32 hexadecimal digits of $\pi$ ending with position 4,152,410,118,610, using BBP-type formulas for $\pi$ due to Bellard and Plouffe. Additional details are given at `http://www.numberworld.org/misc_runs/pi-5t/announce_en.html`. See also [6]. These digits appear to be 'very normal.'

Daniel Shanks, who in 1961 computed $\pi$ to over 100,000 digits, once told Phil Davis that a billion-digit computation would be "forever impossible." But both Kanada and the Chudnovskys achieved that in 1989. Similarly, the intuitionists Brouwer and Heyting asserted the "impossibility" of ever knowing whether the sequence 0123456789 appears in the decimal expansion of $\pi$, yet it was found in 1997 by Kanada, beginning at position 17387594880. As late as 1989, Roger Penrose ventured, in the first edition of his book *The Emperor's New Mind*, that we likely will never know if a string of ten consecutive sevens occurs in the decimal expansion of $\pi$. This string was found in 1997 by Kanada, beginning at position 22869046249.

Figure 2— shows the progress of $\pi$ calculations since 1970, superimposed with a line that charts the long-term trend of Moore's Law. It is worth noting that whereas progress in computing $\pi$ exceeded Moore's Law in the 1990s, it has lagged a bit in the past decade.

2.2.1. *Digital integrity, II.* There are many possible sources of errors in these and other large-scale computations:

- The underlying formulas used might conceivably be in error.
- Computer programs implementing these algorithms, which employ sophisticated algorithms such as fast Fourier transforms to accelerate multiplication, are prone to human programming errors.
- These computations usually are performed on highly parallel computer systems, which require error-prone programming constructs to control parallel processing.
- Hardware errors may occur—this was a factor in the 1986 computation of $\pi$, as noted above.

So why would anyone believe the results of such calculations? The answer is that such calculations are always double-checked with an independent calculation done using some other algorithm, sometimes in more than one way. For instance, Kanada's 2002 computation of $\pi$ to 1.3 trillion decimal digits involved first computing slightly over one trillion hexadecimal (base-16) digits. He found that the 20 hex digits of $\pi$ beginning at position $10^{12}+1$ are `B4466E8D21 5388C4E014`.

Kanada then calculated these hex digits using the "BBP" algorithm [7]. The BBP algorithm for $\pi$ is based on the formula

$$\pi = \sum_{i=0}^{\infty} \frac{1}{16^i}\left(\frac{4}{8i+1} - \frac{2}{8i+4} - \frac{1}{8i+5} - \frac{1}{8i+6}\right), \tag{2.3}$$

which was discovered using the "PSLQ" integer relation algorithm [16]. Integer relation methods find or exclude potential rational relations between vectors of real numbers. At the start of this millennium, they were named one of the top ten algorithms of the twentieth century by *Computing in Science and Engineering*. The best-known is Helaman Ferguson's PSLQ algorithm [10, 3].

Eventually PSLQ produced the formula

$$\pi = 4\,{}_2\mathrm{F}_1\left(\left.\begin{matrix}1,\frac{1}{4}\\ \frac{5}{4}\end{matrix}\,\right|-\frac{1}{4}\right) + 2\tan^{-1}\left(\frac{1}{2}\right) - \log 5, \tag{2.4}$$

where ${}_2\mathrm{F}_1\left(\left.\begin{matrix}1,\frac{1}{4}\\ \frac{5}{4}\end{matrix}\,\right|-\frac{1}{4}\right) = 0.955933837\ldots$ is a Gaussian hypergeometric function.

From (2.4), the series (2.3) almost immediately follows. The BBP algorithm, which is based on (2.3), permits one to calculate binary or hexadecimal digits of $\pi$ beginning at an arbitrary starting point, without needing to calculate any of the preceding digits, by means of a simple scheme that does not require very high precision arithmetic.

The result of the BBP calculation was `B4466E8D21 5388C4E014`. Needless to say, in spite of the many potential sources of error in both computations, the final results dramatically agree, thus confirming (in a convincing but heuristic sense) that both results are almost certainly correct. Although one cannot rigorously assign a "probability" to this event, note that the chances that two random strings of 20 hex digits perfectly agree is one in $16^{20} \approx 1.2089 \times 10^{24}$.

This raises the following question: What is more securely established, the assertion that the hex digits of $\pi$ in positions $10^{12}+1$ through $10^{12}+20$ are `B4466E8D21 5388C4E014`, or the final result of some very difficult work of mathematics that required hundreds or thousands of pages, that relied on many results quoted from other sources, and that (as is frequently the case) only a relative handful of mathematicians besides the author can or have carefully read in detail?

In the most recent computation using the BBP formula, Tse-Wo Zse of Yahoo! Cloud Computing calculated 256 binary digits of $\pi$ starting at the *two quadrillionth* bit [30]. He then checked his result using the following variant of the BBP formula due to Bellard:

$$\begin{aligned}\pi &= \frac{1}{64}\sum_{k=0}^{\infty}\frac{(-1)^k}{1024^k}\left(\frac{256}{10k+1}+\frac{1}{10k+1}-\frac{64}{10k+3}-\frac{4}{10k+5}\right.\\ &\qquad\left.-\frac{4}{10k+7}-\frac{32}{4k+1}-\frac{1}{4k+3}\right) \qquad (2.5)\end{aligned}$$

In this case, both computations verified that the 24 hex digits beginning immediately after the 500 trillionth hex digit (i.e., after the two quadrillionth binary bit) are: `E6C1294A ED40403F 56D2D764`. More recent related computations are also described in [6].

2.3. **Euler's totient function $\phi$.** As another measure of what changes over time and what doesn't, consider two conjectures regarding $\phi(n)$, which counts the number of positive numbers less than and relatively prime to $n$:

2.3.1. **Giuga's conjecture (1950).** *An integer $n > 1$, is a prime if and only if $\mathcal{G}_n := \sum_{k=1}^{n-1} k^{n-1} \equiv n-1 \mod n$.*

Counterexamples are necessarily *Carmichael numbers*—rare birds only proven infinite in 1994—and much more. In [11, pp. 227] we exploited the fact that if a number $n = p_1 \cdots p_m$ with $m > 1$ prime factors $p_i$ is a counterexample to Giuga's conjecture (that is, satisfies

$s_n \equiv n-1 \bmod n$), then for $i \neq j$ we have $p_i \neq p_j$,

$$\sum_{i=1}^{m} \frac{1}{p_i} \quad > \quad 1,$$

and the $p_i$ form a *normal sequence*: $p_i \not\equiv 1 \mod p_j$ for $i \neq j$. Thus, the presence of '3' excludes $7, 13, 19, 31, 37, \ldots$, and of '5' excludes $11, 31, 41, \ldots$.

This theorem yielded enough structure, using some predictive experimentally discovered heuristics, to build an efficient algorithm to show—over several months in 1995—that any counterexample had at least 3459 prime factors and so exceeded $10^{13886}$, extended a few years later to $10^{14164}$ in a five-day desktop computation. The heuristic is self-validating every time that the programme runs successfully. But this method necessarily fails after 8135 primes; someday we hope to exhaust its use.

While writing this piece, one of us was able to obtain almost as good a bound of 3050 primes in under 110 minutes on a laptop computer, and a bound of 3486 primes and 14,000 digits in less than 14 hours; this was extended to 3,678 primes and 17,168 digits in 93 CPU-hours on a Macintosh Pro, using *Maple* rather than C++, which is often orders-of-magnitude faster but requires much more arduous coding.

An equally hard related conjecture for which much less progress can be recorded is:

2.3.2. **Lehmer's conjecture (1932).** $\phi(n) \big| (n-1)$ *if and only if* $n$ *is prime.* He called this *"as hard as the existence of odd perfect numbers."*

Again, prime factors of counterexamples form a normal sequence, but now there is little extra structure. In a 1997 Simon Fraser M.Sc. thesis, Erick Wong verified the conjecture for 14 primes, using normality and a mix of PARI, C++ and *Maple* to press the bounds of the 'curse of exponentiality.' This very clever computation subsumed the entire scattered literature in one computation but could only extend the prior bound from 13 primes to 14.

For Lehmer's related 1932 question: *when does* $\phi(n) \mid (n+1)$*?*, Wong showed there are eight solutions with no more than seven factors (six-factor solutions are due to Lehmer). Let

$$\mathcal{L}_m := \prod_{k=0}^{m-1} F_k$$

with $F_n := 2^{2^n} + 1$ denoting the *Fermat primes.* The solutions are

$$2, \mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_5,$$

and the rogue pair 4919055 and 6992962672132095, but analyzing just eight factors seems out of sight. Thus, in 70 years the computer only allowed the exclusion bound to grow by one prime.

In 1932 Lehmer couldn't factor 6992962672132097. If it had been prime, a ninth solution would exist: since $\phi(n)|(n+1)$ with $n+2$ prime implies that $N := n(n+2)$ satisfies $\phi(N)|(N+1)$. We say *couldn't* because the number is divisible by 73; which Lehmer—a father of much factorization literature–could certainly have discovered had he anticipated a small factor. Today discovering that

$$6992962672132097 = 73 \cdot 95794009207289$$

is nearly instantaneous, while fully resolving Lehmer's original question remains as hard as ever.

2.4. **Inverse computation and Apéry-like series.** Three intriguing formulae for the Riemann zeta function are

$$\begin{aligned} (a)\ \zeta(2) &= 3\sum_{k=1}^{\infty}\frac{1}{k^2\binom{2k}{k}}, \quad (b)\ \zeta(3) = \frac{5}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k^3\binom{2k}{k}}, \qquad (2.6)\\ (c)\ \zeta(4) &= \frac{36}{17}\sum_{k=1}^{\infty}\frac{1}{k^4\binom{2k}{k}}. \end{aligned}$$

Binomial identity (2.6)(a) has been known for two centuries, while (b)—exploited by Apéry in his 1978 proof of the irrationality of $\zeta(3)$—was discovered as early as 1890 by Markov, and (c) was noted by Comtet [3].

Using integer relation algorithms, bootstrapping, and the "Pade" function (*Mathematica* and *Maple* both produce rational approximations well), in 1996 David Bradley and one of us [3, 11] found the following unanticipated generating function for $\zeta(4n+3)$:

$$\sum_{k=0}^{\infty}\zeta(4k+3)\,x^{4k} = \frac{5}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k^3\binom{2k}{k}(1-x^4/k^4)}\prod_{m=1}^{k-1}\left(\frac{1+4x^4/m^4}{1-x^4/m^4}\right). \quad (2.7)$$

Note that this formula permits one to read off an infinity of formulas for $\zeta(4n+3)$, $n>0$, beginning with (2.6)(b), by comparing coefficients of $x^{4k}$ on the LHS and the RHS.

A decade later, following a quite analogous but much more deliberate experimental procedure, as detailed in [3], we were able to discover a similar general formula for $\zeta(2n+2)$ that is pleasingly parallel to (2.7):

$$\sum_{k=0}^{\infty}\zeta(2k+2)\,x^{2k} = 3\sum_{k=1}^{\infty}\frac{1}{k^2\binom{2k}{k}(1-x^2/k^2)}\prod_{m=1}^{k-1}\left(\frac{1-4\,x^2/m^2}{1-x^2/m^2}\right). \quad (2.8)$$

As with (2.7), one can now read off an infinity of formulas, beginning with (2.6)(a). In 1996, the authors could reduce (2.7) to a finite form that they could not prove, but Almqvist and Granville did a year later. A decade later, the Wilf-Zeilberger algorithm [29, 23]—for which the inventors were awarded the Steele Prize—directly (as implemented in *Maple*) certified (2.8) [10, 3]. In other words, (2.8) was both discovered and proven by computer.

We found a comparable generating function for $\zeta(2n+4)$, giving (2.6) (c) when $x=0$, but one for $\zeta(4n+1)$ still eludes us.

2.5. **Reciprocal series for $\pi$.** Truly novel series for $1/\pi$, based on elliptic integrals, were discovered by Ramanujan around 1910 [3, 10, 31]. One is:

$$\frac{1}{\pi} = \frac{2\sqrt{2}}{9801}\sum_{k=0}^{\infty}\frac{(4k)!\,(1103+26390k)}{(k!)^4 396^{4k}}. \tag{2.9}$$

Each term of (2.9) adds eight correct digits. Gosper used (2.9) for the computation of a then-record 17 million digits of $\pi$ in 1985—thereby completing the first proof of (2.9) [10, Ch. 3]. Shortly thereafter, David and Gregory Chudnovsky found the following variant, which lies in the quadratic number field $Q(\sqrt{-163})$ rather than $Q(\sqrt{58})$:

$$\frac{1}{\pi} = 12\sum_{k=0}^{\infty}\frac{(-1)^k\,(6k)!\,(13591409+545140134k)}{(3k)!\,(k!)^3\,640320^{3k+3/2}}. \tag{2.10}$$

Each term of (2.10) adds 14 correct digits. The brothers used this formula several times, culminating in a 1994 calculation of $\pi$ to over four billion decimal digits. Their remarkable story was told in a prizewinning *New Yorker* article [26]. Remarkably, as we already noted earlier, (2.10) was used again in 2010 for the current record computation of $\pi$.

2.5.1. **Wilf-Zeilberger at work.** A few years ago Jésus Guillera found various Ramanujan-like identities for $\pi$, using integer relation methods. The three most basic—and entirely rational—identities are:

$$\frac{4}{\pi^2} = \sum_{n=0}^{\infty}(-1)^n r(n)^5(13+180n+820n^2)\left(\frac{1}{32}\right)^{2n+1} \tag{2.11}$$

$$\frac{2}{\pi^2} = \sum_{n=0}^{\infty}(-1)^n r(n)^5(1+8n+20n^2)\left(\frac{1}{2}\right)^{2n+1} \tag{2.12}$$

$$\frac{4}{\pi^3} \stackrel{?}{=} \sum_{n=0}^{\infty} r(n)^7(1+14n+76n^2+168n^3)\left(\frac{1}{8}\right)^{2n+1}, \tag{2.13}$$

where $r(n) := (1/2\cdot 3/2\cdot\cdots\cdot(2n-1)/2)/n!$ .

Guillera proved (2.11) and (2.12) in tandem, by very ingeniously using the Wilf-Zeilberger algorithm [29, 23] for formally proving hypergeometric-like identities [10, 3, 19, 31]. No other proof is known, and there seem to be no like formulae for $1/\pi^N$ with $N \geq 4$. The third, (2.13), is almost certainly true. Guillera ascribes (2.13) to Gourevich, who used integer relation methods to find it.

We were able to "discover" (2.13) using 30-digit arithmetic, and we checked it to 500 digits in 10 seconds, to 1200 digits in 6.25 minutes, and to 1500 digits in 25 minutes, all with naive command-line instructions in *Maple*. But it has no proof, nor does anyone have an inkling of how to prove it; especially, as experiment suggests, since it has no 'mate' in analogy to (2.11) and (2.12) [3]. Our intuition is that if a proof exists, it is more a verification than an explication and so we stopped looking. We are happy just to "know" that the beautiful identity is true (although it would be more remarkable were it eventually to fail). It may be true for no good reason—it might just have no proof and be a very concrete Gödel-like statement.

In 2008 Guillera [19] produced another lovely pair of third-millennium identities—discovered with integer relation methods and proved with creative telescoping—this time for $\pi^2$ rather than its reciprocal. They are

$$\sum_{n=0}^{\infty} \frac{1}{2^{2n}} \frac{\left(x+\frac{1}{2}\right)_n^3}{(x+1)_n^3} \left(6(n+x)+1\right) = 8x \sum_{n=0}^{\infty} \frac{\left(\frac{1}{2}\right)_n^2}{(x+1)_n^2}, \tag{2.14}$$

and

$$\sum_{n=0}^{\infty} \frac{1}{2^{6n}} \frac{\left(x+\frac{1}{2}\right)_n^3}{(x+1)_n^3} \left(42(n+x)+5\right) = 32x \sum_{n=0}^{\infty} \frac{\left(x+\frac{1}{2}\right)_n^2}{(2x+1)_n^2}. \tag{2.15}$$

Here $(a)_n = a(a+1)\cdot\cdot(a+n-1)$ is the *rising factorial*. Substituting $x = 1/2$ in (2.14) and (2.15), he obtained respectively the formulae

$$\sum_{n=0}^{\infty} \frac{1}{2^{2n}} \frac{(1)_n^3}{\left(\frac{3}{2}\right)_n^3} (3n+2) = \frac{\pi^2}{4}, \qquad \sum_{n=0}^{\infty} \frac{1}{2^{6n}} \frac{(1)_n^3}{\left(\frac{3}{2}\right)_n^3} (21n+13) = 4\,\frac{\pi^2}{3}.$$

## 3. Formal Verification of Proof

In 1611, Kepler described the stacking of equal-sized spheres into the familiar arrangement we see for oranges in the grocery store. He asserted that this packing is the tightest possible. This assertion is now known as the Kepler conjecture, and has persisted for centuries without rigorous proof. Hilbert implicitly included the irregular case of the Kepler conjecture in problem 18 of his famous list of unsolved problems

in 1900: *whether there exist non-regular space-filling polyhedra?* the regular case having been disposed of by Gauss in 1831.

In 1994, Thomas Hales, now at the University of Pittsburgh, proposed a five-step program that would result in a proof: (a) treat maps that only have triangular faces; (b) show that the face-centered cubic and hexagonal-close packings are local maxima in the strong sense that they have a higher score than any Delaunay star with the same graph; (c) treat maps that contain only triangular and quadrilateral faces (except the pentagonal prism); (d) treat maps that contain something other than a triangular or quadrilateral face; and (e) treat pentagonal prisms.

In 1998, Hales announced that the program was now complete, with Samuel Ferguson (son of mathematician-sculptor Helaman Ferguson) completing the crucial fifth step. This project involved extensive computation, using an interval arithmetic package, a graph generator, and *Mathematica*. The computer files containing the source code and computational results occupy more than three Gbytes of disk space. Additional details, including papers, are available at `http://www.math.pitt.edu/~thales/kepler98`. For a mixture of reasons—some more defensible than others—the *Annals of Mathematics* initially decided to publish Hales' paper with a cautionary note, but this disclaimer was deleted before final publication.

Hales [20] has now embarked on a multi-year program to certify the proof by means of computer-based formal methods, a project he has named the "Flyspeck" project. As these techniques become better understood, we can envision a large number of mathematical results eventually being confirmed by computer, as instanced by other articles in the same issue of the *Notices* as Hales' article.

## 4. Limits of Computation

A remarkable example is the following:

$$\int_0^\infty \cos(2x) \prod_{n=1}^\infty \cos(x/n)\,\mathrm{d}x \quad = \tag{4.1}$$

$$0.392699081698724154807830422909937860524645434187231595926\ldots$$

The computation of this integral to high precision can be performed using a scheme described in [5]. When we first did this computation, we thought that the result was $\pi/8$, but upon careful checking with the numerical value

$$0.392699081698724154807830422909937860524646174921888227621\ldots,$$

it is clear that the two values disagree beginning with the 43rd digit!

Richard Crandall [15, §7.3] later explained this mystery. Via a physically motivated analysis of *running out of fuel* random walks, he showed that $\pi/8$ is given by the following very rapidly convergent series expansion, of which formula (4.1) above is merely the first term:

$$\frac{\pi}{8} = \sum_{m=0}^{\infty} \int_0^{\infty} \cos[2(2m+1)x] \prod_{n=1}^{\infty} \cos(x/n)\, \mathrm{d}x. \tag{4.2}$$

Two terms of the series above suffice for 500-digit agreement.

As a final sobering example, we offer the following "sophomore's dream" identity

$$\begin{aligned} \sigma_{29} &:= \sum_{n=-\infty}^{\infty} \operatorname{sinc}(n)\operatorname{sinc}(n/3)\operatorname{sinc}(n/5)\cdots\operatorname{sinc}(n/23)\operatorname{sinc}(n/29) \\ &= \int_{-\infty}^{\infty} \operatorname{sinc}(x)\operatorname{sinc}(x/3)\operatorname{sinc}(x/5)\cdots\operatorname{sinc}(x/23)\operatorname{sinc}(x/29)\, dx, \end{aligned} \tag{4.3}$$

where the denominators range over the odd primes, which was first discovered empirically. More generally, consider

$$\begin{aligned} \sigma_p &:= \sum_{n=-\infty}^{\infty} \operatorname{sinc}(n)\operatorname{sinc}(n/3)\operatorname{sinc}(n/5)\operatorname{sinc}(n/7)\cdots\operatorname{sinc}(n/p) \\ &\stackrel{?}{=} \int_{-\infty}^{\infty} \operatorname{sinc}(x)\operatorname{sinc}(x/3)\operatorname{sinc}(x/5)\operatorname{sinc}(x/7)\cdots\operatorname{sinc}(x/p)\, dx. \end{aligned} \tag{4.4}$$

Provably, the following is true: The "sum equals integral" identity, for $\sigma_p$ remains valid at least for $p$ among the first $10^{176}$ primes; but stops holding after some larger prime, and thereafter the "sum less the integral" is strictly positive, but *they always differ by much less than one part in a googolplex* $= 10^{100}$. An even stronger estimate is possible assuming the Generalized Riemann Hypothesis (see [15, §7] and [8]).

## 5. Concluding Remarks

The central issues of how to view experimentally discovered results have been discussed before. In 1993, Arthur Jaffe and Frank Quinn warned of the proliferation of not-fully-rigorous mathematical results and proposed a framework for a "healthy and positive" role for "speculative" mathematics [21]. Numerous well-known mathematicians responded [1]. Morris Hirsch, for instance, countered that even Gauss published incomplete proofs, and the 15,000 combined pages of the

proof of the classification of finite groups raises questions as to when we should certify a result. He suggested that we attach a label to each proof – e.g., "computer-aided," "mass collaboration," "constructive," etc. Saunders Mac Lane quipped that "we are not saved by faith alone, but by faith and works," meaning that we need both intuitive work and precision.

At the same time, computational tools now offer remarkable facilities to confirm analytically established results, as in the tools in development to check identities in equation-rich manuscripts, and in Hales' project to establish the Kepler conjecture by formal methods.

The flood of information and tools in our information-soaked world is unlikely to abate. We have to learn and teach judgment when it comes to using what is possible digitally. This means mastering the sorts of techniques we have illustrated and having some idea why a software system does what it does. It requires knowing when a computation is or can—in principle or practice—be made into a rigorous proof and when it is only compelling evidence, or is entirely misleading. For instance, even the best commercial linear programming packages of the sort used by Hales will not certify any solution though the codes are almost assuredly correct. It requires rearranging hierarchies of what we view as hard and as easy.

It also requires developing a curriculum that carefully teaches experimental computer-assisted mathematics. Some efforts along this line are already underway by individuals including Marc Chamberland at Grinnell (`http://www.math.grin.edu/~chamberl/courses/MAT444/syllabus.html`), Victor Moll at Tulane, Jan de Gier in Melbourne, and Ole Warnaar at University of Queensland.

Judith Grabner has noted that a large impetus for the development of modern rigor in mathematics came with the Napoleonic introduction of regular courses: lectures and textbooks force a precision and a codification that apprenticeship obviates. But it will never be the case that quasi-inductive mathematics supplants proof. We need to find a new equilibrium. That said, we are only beginning to tap new ways to enrich mathematics. As Jacques Hadamard said [25]:

> The object of mathematical rigor is to sanction and legitimize the conquests of intuition, and there was never any other object for it.

Never have we had such a cornucopia of ways to generate intuition. The challenge is to learn how to harness them, how to develop and how to transmit the necessary theory and practice. The Priority Research Centre for Computer Assisted Research Mathematics and its

Applications (CARMA), `http://www.newcastle.edu.au/research/centres/carmacentre.html`, which one of us directs, hopes to play a lead role in this endeavor: an endeavor which in our view encompasses an exciting mix of exploratory experimentation and rigorous proof.

## References


[1] M. Atiyah, et al (1994). "Responses to 'Theoretical Mathematics: Toward a Cultural Synthesis of Mathematics and Theoretical Physics,' by A. Jaffe and F. Quinn," *Bulletin of the American Mathematical Society*, vol. 30, no. 2 (Apr 1994), pp. 178-207.

[2] J. Avigad (2008). "Computers in mathematical inquiry," in *The Philosophy of Mathematical Practice*, P. Mancuso ed., Oxford University Press, pp. 302–316.

[3] D. Bailey, J. Borwein, N. Calkin, R. Girgensohn, R. Luke and V. Moll (2007). *Experimental Mathematics in Action*, A K Peters, Natick, MA, 2007.

[4] D. H. Bailey and J. M. Borwein (2008). "Computer-assisted discovery and proof." *Tapas in Experimental Mathematics*, 21–52, in *Contemporary Mathematics*, vol. 457, American Mathematical Society, Providence, RI, 2008.

[5] D. H. Bailey, J. M. Borwein, V. Kapoor and E. Weisstein. "Ten Problems in Experimental Mathematics," *American Mathematical Monthly*, vol. 113, no. 6 (Jun 2006), pp. 481–409.

[6] D. H. Bailey, J. M. Borwein, A. Mattingly and G. Wightwick, "The Computation of Previously Inaccessible Digits of $\pi^2$ and Catalan's Constant," *Notices of the American Mathematical Society*, to appear.

[7] D. H. Bailey, P. B. Borwein and S. Plouffe, (1997). "On the Rapid Computation of Various Polylogarithmic Constants," *Mathematics of Computation*, vol. 66, no. 218 (Apr 1997), pp. 903-913.

[8] R. Baillie, D. Borwein, and J. Borwein (2008). "Some sinc sums and integrals," *American Math. Monthly*, vol. 115 (2008), no. 10, pp. 888–901.

[9] J. M. Borwein (2005). "The SIAM 100 Digits Challenge," Extended review in the *Mathematical Intelligencer*, vol. 27 (2005), pp. 40–48.

[10] J. M. Borwein and D. H. Bailey (2008). *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, extended second edition, A K Peters, Natick, MA, 2008.

[11] J. M. Borwein, D. H. Bailey and R. Girgensohn (2004). *Experimentation in Mathematics: Computational Roads to Discovery*, A K Peters, Natick, MA, 2004.

[12] J. M. Borwein, O-Yeat Chan and R. E. Crandall, "Higher-dimensional box integrals." Submitted *Experimental Mathematics*, January 2010.

[13] J. M. Borwein and K. Devlin (2008). *The Computer as Crucible*, A K Peters, Natick, MA, 2008.

[14] J. M. Borwein, I. J. Zucker and J. Boersma, "The evaluation of character Euler double sums," *Ramanujan Journal*, **15** (2008), 377–405.

[15] R. E. Crandall (2007). "Theory of ROOF Walks," 2007, available at `http://www.perfscipress.com/papers/ROOF11_psipress.pdf`.

[16] H. R. P. Ferguson, D. H. Bailey and S. Arno (1999). "Analysis of PSLQ, An Integer Relation Finding Algorithm," *Mathematics of Computation*, vol. 68, no. 225 (Jan 1999), pp. 351–369.

[17] L. R. Franklin (2005). "Exploratory Experiments," *Philosophy of Science*, **72**, pp. 888–899.

[18] M. Giaquinto (2007). *Visual Thinking in Mathematics. An Epistemological Study*, Oxford University Press, New York, 2007.

[19] J. Guillera (2008). "Hypergeometric identities for 10 extended Ramanujan-type series," *Ramanujan Journal*, vol. 15 (2008), pp. 219–234.

[20] T. C. Hales, "Formal Proof," *Notices of the AMS*, vol. 55, no. 11 (Dec. 2008), pp. 1370–1380.

[21] A. Jaffe and F. Quinn (1993). "'Theoretical Mathematics': Toward a Cultural synthesis of Mathematics and Theoretical Physics," *Bulletin of the American Mathematical Society*, vol. 29, no. 1 (Jul 1993), pp. 1–13.

[22] M. Livio (2009). *Is God a Mathematician?*, Simon and Schuster, New York, 2009.

[23] M. Petkovsek, H. S. Wilf, D. Zeilberger, $A = B$, A K Peters, Natick, MA, 1996.

[24] S. Pinker (2007). *The Stuff of Thought: Language as a Window into Human Nature*, Allen Lane, New York, 2007.

[25] G. Pólya (1981). *Mathematical discovery: On understanding, learning, and teaching problem solving*, (Combined Edition), New York, John Wiley and Sons, New York, 1981.

[26] R. Preston (1992) "The Mountains of Pi," *New Yorker*, 2 Mar 1992, `http://www.newyorker.com/archive/content/articles/050411fr_archive01`.

[27] D. L. Smail (2008). *On Deep History and the Brain*, Caravan Books, University of California Press, Berkeley, CA, 2008.

[28] H. K. Sørenson (2009). "Exploratory experimentation in experimental mathematics: A glimpse at the PSLQ algorithm," *Philosophy of Mathematics: Sociological Aspects and Mathematical Practice*. In press.

[29] H. S. Wilf and D. Zeilberger, "Rational Functions Certify Combinatorial Identities," *Journal of the American Mathematical Society*, vol. 3 (1990), pp. 147–158.

[30] Tse-Wo Zse, personal communication to the authors, July 2010.

[31] W. Zudilin (2008). "Ramanujan-type formulae for $1/\pi$: A second wind," 19 May 2008, available at `http://arxiv.org/abs/0712.1332`.