Experimental Mathematics via Inverse Symbolic Computation

David M. Bradley

Centre for Experimental and Constructive Mathematics
Simon Fraser University

http://www.cecm.sfu.ca/~dbradley/

June 12, 1997

Introduction

Is mathematics created or discovered? Or both?

Traditionally, mathematics has distinguished itself from the empirical sciences, in that the latter must fit their theories to experimental reality.

On the other hand, mathematicians, within certain constraints, are free to choose their own reality.

If mathematics is primarily created, rather than discovered, then experimentation should have little or no influence in mathematical development.

However, if discovery plays a significant role, then we should expect that experimentation should also.

In previous centuries, experimentation in mathematics was mostly limited to trial and error hand-calculuation. With the advent of digital (and now molecular!) computers, a revolution is taking place in how mathematicians go about their work.

In fact, the past 20 years has seen a dramatic reconcretization of mathematics, in which fields such as number theory, classical analysis, and special functions have received new infusions fueled by advances in hardware, software, and algorithms.

A whole new palette of tools is available to the researcher, many of which, with a little effort, can operate like extensions of the mind, multiplying our ability to generate examples, test hypotheses, and build intuition by unprecendented factors.

The current generation of computers can not only verify results, but *predict* them.

The sophistication of computational tools continues to improve, yielding more than just a quantitative increase in mathematical results; rather, a *qualitative* shift is taking place.

Probably the best kind of tool is the one that enables people to build better tools.

As tool is built upon tool, a peculiar meld of mind and machine emerges. The distinction between mind and machine is increasingly blurred. Many of my most impressive results of the past two years could not and would not have been discovered without the assistance of sophisticated inverse symbolic computational software.

"Computers are useless. They can only give you answers."

Pablo Picasso (1881-1973)

Picasso was wrong.

• Answers most definitely are useful.

• In the realm of experimental mathematics, computers generate far more questions than answers.

• Computers are now sophisticated enough to make conjectures of their own.

• Computer-generated conjectures have led to questions that suggest fundamental new avenues of research.

• In many cases, consideration of computational issues has led to significant paradigm shifts.

Fractals and Chaos

• The study of fractals and chaos was simply too difficult to undertake before the advent of high-speed digital computers.

• Simply stated, it was humanly impossible to generate any but the crudest of fractal images.

*****

But...

Just as the study of fractals unveiled fantastically baroque images of stunning beauty, massive computer searches and inverse symbolic techniques are revealing equally enticing analytic objects whose beauty lies in the more conceptual realms of infinite series, continued fractions, and other iterative expansions.

What is Inverse Symbolic Computation?

First, what is symbolic computation?

<u>Partial Answer by Examples</u>:

• Evaluate a sum, an integral, a product,...
• Compute a derivative, a determinant,...
• Verify a formula
• Given a question, find the answer

The inverse problem to

"given a question, find the right answer"

is

Given an answer,

Find the "Right" Question!

Hence, Inverse Symbolic Computation:

- Given a number, ask "What combination of constants and special function values most likely produced it?"

- Given a sequence, ask "What is the generating function?" or "Does the sequence have a name?" "Does it arise in other contexts?"

- Given a finite set of functions and operations, ask "Does there exist a formula which looks something like $\{\cdots\}$ and involves these functions and operations? (And if so, what is it?)"

Inverse Symbolic Computational Tools

Inverse Symbolic Calculator
- Input a number, and find out where it comes from.
- Combines table lookup with "smart lookup" to check whether the number is a simple combination of known constants.
- Author: Simon Plouffe, formely of CECM, now at Wolfram Research
- http://www.cecm.sfu.ca/projects/ISC/

Encyclopedia of Integer Sequences
- Does for sequences what the ISC does for real numbers.
- Expanded and updated version of Sloane's handbook
- Text by Neil Sloane & Simon Plouffe, Academic Press, 1995, ISBN: 0-12-558632-9
- http://www.research.att.com/~njas/sequences/

Maple's GFUN package
-Given a sequence of numbers, the package will try to guess a recurrence, find the ordinary/exponential generating function, solve the recurrence, etc.

Integer Relations Algorithms
- Lattice Basis Reduction (Ferguson & Forcade)
- "LLL" (Lenstra, Lenstra & Lovasz)
- "PSLQ", (Dave Bailey, NASA Ames & Ferguson)
- Given a vector of real numbers, output a vector of integers such that the dot product is zero to within working precision.

Inverse Symbolic Computation

- What is 1.1981402347355922074...?

If $a_0 = 1$, $b_0 = \sqrt{2}$, and

$$a_{n+1} = \frac{a_n + b_n}{2}, \quad b_{n+1} = \sqrt{a_n b_n},$$

then

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \frac{\pi/2}{\int_0^1 \frac{dt}{\sqrt{1-t^4}}}$$

$$= 1.1981402347355922074...$$

The convergence is quadratic - 10 iterations already yields nearly 1400 digits of accuracy.

In 1799, Gauss observed this purely numerically, and wrote that this result

"will surely open a whole new field of analysis."

The Inverse Symbolic Calculator (ISC)

To what extent can we automate Gauss's incredible insight?

Obviously, Gauss was familiar with the initial digits of the decimal expansions of π, the complete elliptic integrals, and probably simple rational multiples of these as well.

So, it would make sense to compile a table of "famous" constants, and for a first crack, try a simple table-lookup.

Even better, one could try simple rational multiples of tabulated constants, and various linear combinations and products of these.

Simon Plouffe's "Inverse Symbolic Calculator" is a practical instantiation of this idea. Essentially, it is a calculator with a big screen and only one button, which answers the question "What is this number made of?"

The number $r := 1.6180339887498948482\ldots$ happens to be the unique positive real number satisfying

$$\int_0^\infty \frac{dx}{(1+x^r)^r} = 1.$$

So, what is r? Is it just the number that happens to work or is it special in some other contexts as well?

Ask the ISC, and we find that, most likely,

$$r = \frac{1+\sqrt{5}}{2},$$

which is indeed the case.

The so-called "golden ratio" can turn up in the most unexpected places.

Detecting Integer Relations

• Given a vector of real numbers (or their decimal approximations), output a vector of integers such that the dot product is zero to within working precision.

• Let $x = (x_1, x_2, \ldots, x_n)$ be a vector of real numbers. Then x is said to <u>possess</u> <u>an</u> <u>integer</u> <u>relation</u> if there exist integers a_j not all zero such that

$$\sum_{j=1}^{n} a_j x_j = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0.$$

Problem: Find the integers a_j if they exist. If they do not exist, then truncating the x_j to working precision, obtain lower bounds on the size of the a_j.

• Euclid's algorithm gives a solution for $n = 2$.

• Euler, Jacobi, Poincare, Minkowski, Perron, Brun, Bernstein and others tried to find a general algorithm for $n > 2$.

• The first general algorithm, *Lattice Basis Reduction,* was discovered in 1977 by Ferguson and Forcade, however it suffered from numerical stability.

• Improvements were given by Lenstra, Lentra and Lovasz (LLL) and Dave Bailey of NASA Ames (PSLQ).

Applications of Integer Relation Detection

• Suppose α can be computed to high precision. Form the vector $x = (1, \alpha, \alpha^2, \ldots, \alpha^n)$ and apply an integer relation detecting algorithm. If a relation is found, the integers a_j satisfy

$$\sum_{j=0}^{n} a_j \alpha^j = a_0 + a_1\alpha + \cdots a_n\alpha^n = 0,$$

i.e. α is algebraic of degree $\leq n$. If no relation is found, bounds are obtained within which no such annihilating polynomial can exist.

• One can also check whether α satisfies an identity of the form

$$\alpha^p = 2^a 3^b 5^c \pi^d \left[\zeta(3)\right]^m \left[\Gamma(1/4)\right]^k \cdots,$$

say, by taking logarithms.

Some Integer Relation Results

$$\int_0^{\pi/3} \log(\tan\theta)\, d\theta = \int_0^{\pi/6} \log(\tan\theta)\, d\theta,$$

$$2\int_0^{\pi/4} \log(\tan\theta)\, d\theta = 3\int_0^{\pi/12} \log(\tan\theta)\, d\theta,$$

$$\begin{aligned} 2\int_0^{\pi/4} & \log(\tan\theta)\, d\theta \\ &= 5\int_0^{3\pi/20} \log(\tan\theta)\, d\theta \\ &\quad -5\int_0^{\pi/20} \log(\tan\theta)\, d\theta, \end{aligned}$$

and many, many more...

Series Acceleration Formulae for Catalan's Constant

Catalan's constant is

$$G := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2}.$$

Let $L(n)$ denote the nth Lucas number (M0155 in Sloan and Plouffe's Encyclopaedia). The Lucas numbers satisfy the recursion

$$L(n) = L(n-1) + L(n-2), \qquad n > 2,$$

with initial conditions $L(1) = 1$, $L(2) = 3$.

Theorem 1

$$G = \frac{\pi}{8}\log\left(\frac{10+\sqrt{50-22\sqrt{5}}}{10-\sqrt{50-22\sqrt{5}}}\right) + \frac{5}{8}\sum_{n=0}^{\infty} \frac{L(2n+1)}{(2n+1)^2\binom{2n}{n}}.$$

Proof. (Sketch) First, define

$$T(r) := \int_0^{\pi r} \log(\tan\theta)\, d\theta, \qquad 0 \leq r \leq \tfrac{1}{2}.$$

Lemma 1 *The Fourier Series expansion*

$$T(r) = -\sum_{n=0}^{\infty} \frac{\sin((2n+1)2\pi r)}{(2n+1)^2}, \qquad 0 \leq r \leq \tfrac{1}{2},$$

holds.

Proof. Let $z = e^{-2ix}$, $0 \leq x \leq \frac{1}{2}\pi$. By multisecting the power series for $\log(1+z)$, we have

$$\begin{aligned} 2\sum_{n=0}^{\infty} \frac{\cos(4n+2)x}{2n+1} &= 2\Re \sum_{n=0}^{\infty} \frac{e^{-(4n+2)ix}}{2n+1} \\ &= \Re \log \frac{1+e^{-2ix}}{1-e^{-2ix}} \\ &= -\log(\tan x). \end{aligned}$$

Lemma 1 now follows on integrating and setting $x = \pi r$.

Corollary 1

$$\begin{aligned} G &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2} \\ &= -T(\tfrac{1}{4}) \\ &= -\int_0^{\pi/4} \log(\tan\theta)\, d\theta. \end{aligned}$$

Proof. Put $r = \frac{1}{4}$ in the Fourier Series expansion of Lemma 1.

We seek series acceleration formulae for G. The idea is to employ integer relations between log(tan) integrals. Ideally, the relations should involve integrals with reduced integration ranges. When re-expanded into series, the reduced integration range is manifested as a continuous analog of bunching several terms together, yielding a more rapidly convergent series.

Continuing with the proof of Theorem 1, we write

$$\begin{aligned} -\frac{2}{5}G &= \frac{2}{5}\int_0^{\pi/4} \log(\tan\theta)\, d\theta \\ &= \int_0^{3\pi/20} \log(\tan\theta)\, d\theta - \int_0^{\pi/20} \log(\tan\theta)\, d\theta \\ &= \int_{\pi/20}^{3\pi/20} \log(\tan\theta)\, d\theta \\ &= \theta \log(\tan\theta)\Big|_{\pi/20}^{3\pi/20} - \int_{\pi/20}^{3\pi/20} \frac{\theta \sec^2\theta}{\tan\theta}\, d\theta \\ &= \frac{\pi}{20} \log\left(\frac{\tan^3(\frac{3\pi}{20})}{\tan(\frac{\pi}{20})}\right) - \int_{\pi/10}^{3\pi/10} \frac{\frac{1}{2}\theta\, \frac{1}{2} d\theta}{\sin\frac{1}{2}\theta \cos\frac{1}{2}\theta}. \end{aligned}$$

But,

$$\begin{aligned} -\int_{\pi/10}^{3\pi/10} \frac{\frac{1}{2}\theta\, \frac{1}{2} d\theta}{\sin\frac{1}{2}\theta \cos\frac{1}{2}\theta} &= -\frac{1}{2}\int_{\pi/10}^{3\pi/10} \frac{\theta}{\sin\theta}\, d\theta \\ &= -\int_{\sin\frac{\pi}{10}}^{\sin\frac{3\pi}{10}} \frac{\sin^{-1} x}{\sqrt{1-x^2}} \cdot \frac{dx}{2x} \\ &= -\int_{\sin\frac{\pi}{10}}^{\sin\frac{3\pi}{10}} \sum_{n=1}^{\infty} \frac{(2x)^{2n-2}}{n\binom{2n}{n}}\, dx \\ &= \frac{1}{4}\sum_{n=0}^{\infty} \frac{\phi^{2n+1} - \tau^{2n+1}}{(2n+1)^2\binom{2n}{n}}, \end{aligned}$$

where

$$\phi := 2\sin(\pi/10) = \frac{\sqrt{5}-1}{2},$$

$$\tau := 2\sin(3\pi/10) = \frac{\sqrt{5}+1}{2}.$$

The proof of Theorem 1 is complete once we

a) note that the Lucas numbers satisfy

$$L(n) = \left(\frac{1+\sqrt{5}}{2}\right)^n + \left(\frac{1-\sqrt{5}}{2}\right)^n, \qquad n \geq 0,$$

and

b) verify the non-trivial algebraic denesting relationship

$$\frac{\tan^3(\frac{3\pi}{20})}{\tan(\frac{\pi}{20})} = \frac{10 - \sqrt{50 - 22\sqrt{5}}}{10 + \sqrt{50 - 22\sqrt{5}}}.$$

Success or Failure?

The simpler relationship

$$2\int_0^{\pi/4} \log(\tan\theta)\,d\theta = 3\int_0^{\pi/12} \log(\tan\theta)\,d\theta,$$

yields Ramanujan's result (which he proved by quite different methods):

$$G = \frac{\pi}{8}\log(2+\sqrt{3}) + \frac{3}{8}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2\binom{2n}{n}}.$$

But, recall the adage

"If at first you don't succeed ...

... redefine success."

We have

$$\sum_{n=0}^{\infty}\frac{L(2n+1)}{(2n+1)^2\binom{2n}{n}} = \frac{8}{5}G + \frac{\pi}{5}\log\left(\frac{10-\sqrt{50-22\sqrt{5}}}{10+\sqrt{50-22\sqrt{5}}}\right),$$

where

$$G := \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^2}$$

is Catalan's constant.

Some Much Deeper Results
(Born of LLL/PSLQ)

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

- For all positive integers n,

$$\frac{5}{2}\sum_{k=1}^{n}\binom{2k}{k}\frac{n^2k^2}{4n^4+k^4}\prod_{j=1}^{k-1}\frac{n^4-j^4}{4n^4+j^4} = 1.$$

- For all positive integers n,

$$\frac{1}{\pi}\int_0^{\infty}\frac{dy}{1+y^2}\prod_{j=0}^{n-1}\frac{4y^2-(j/n)^4}{y^2+(j/n)^4} = \binom{2n}{n}.$$

- For all positive integers n,

$$ {}_6F_5\left(\begin{array}{c} n+1,\ n+1,\ 2n\pm in,\ \pm in \\ n+1/2, n, 2n+1, n+1\pm in \end{array}\middle| -\frac{1}{4}\right) = \frac{2}{5}\binom{2n}{n}\prod_{j=1}^{n-1}\frac{n^4-j^4}{4n^4+j^4}.$$

- For all complex numbers z,

$$\sum_{k=1}^{\infty}\frac{1}{k^3(1-z^4/k^4)} = \frac{5}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k^3\binom{2k}{k}}\frac{1}{1-z^4/k^4}\prod_{j=1}^{k-1}\frac{1+4z^4/j^4}{1-z^4/j^4}.$$