Emerging Tools for Experimental Mathematics

Jonathan M. Borwein and Robert M. Corless

1 Introduction and "Warm-up"

> If I can give an abstract proof of something, I'm reasonably happy. But if I can get a concrete, computational proof and actually produce numbers I'm much happier. I'm rather an addict of doing things on the computer, because that gives you an explicit criterion of what's going on. I have a visual way of thinking, and I'm happy if I can see a picture of what I'm working with. —John Milnor, [25, p. 78]

Using mostly elementary examples, we discuss the use of some recent and emerging tools for experimental mathematics. The tools discussed include so-called "inverse symbolic computation", using lattice reduction algorithms such as "LLL" and "PSLQ", and Sloane and Plouffe's integer sequence lookup program. We concentrate on 'computer-assisted discovery' of mathematical results, but a little 'computer-assisted proof' creeps in as well. This paper is *not* a tutorial on how lattice basis reduction algorithms such as LLL or PSLQ actually work; rather, we discuss some of the ways these tools can be used to generate conjectures, and for that, a detailed understanding of the underlying algorithms is not necessary. We do hope, however, to convey some appreciation of their power.

We begin with some warm-up examples, using the *Inverse Symbolic Calculator* (ISC); http:// www.cecm.sfu.ca/ MRG/ INTERFACES.html. The basic idea is simple: given the first few decimal digits of some real number, we want the ISC to guess a formula for what it "really" is.

For example, if we input 3.14626436994198, and click on "simple lookup" (the default) and "Run", the ISC tells us that

```
...
3146264369941972 = (0405) 1/abs(-sr(3)+sr(2))
Your value of 314626436994198 would be here.
3146264469611207 = (0192) (5^(1/2)+4)/(exp(1/2)+1/3)
...
```

This has correctly identified our input as $1/(\sqrt{3}-\sqrt{2}) = \sqrt{3}+\sqrt{2}$, by table lookup. Using the "integer relation" option would get us, instead, the error message that we need at least 16 digits, and then when we change the final 8 to 72, the following answer appears:

```
K = 3.146264369941972 gave the following results:
                                                      2    4
        K satisfies the following polynomial, 1 - 10 X  + X
```

together with some negative results about combinations of other constants.

Now consider a second "warm-up". If we input the number K, computed from the infinite product

$$K = \prod_{n \geq 2} \frac{n^2 - 1}{n^2 + 1} = .2720290549821331\ldots,$$

then the "simple lookup" fails to tell us anything; the "integer relations" option tells us that it is not a simple combination of a few specific constants; but the "smart lookup" tells us that $K/2 = \pi/(\exp(-\pi) - \exp(\pi))$. This is actually wrong—it's got the wrong sign, possibly because signs are ignored for the ISC—but the digits are correctly identified. K is indeed equal to $\pi/\sinh(\pi)$.

As a final warm-up, consider the following two infinite products:

$$\prod_{k \geq 1} \frac{(1 + \frac{1}{k} + \frac{1}{k^2})^2}{(1 + \frac{2}{k} + \frac{3}{k^2})} = 1.84893618285824448\ldots \tag{1}$$

$$\prod_{k \geq 1} \frac{(1 + \frac{1}{k} + \frac{1}{k^2} + \frac{1}{k^3})^2}{(1 + \frac{2}{k} + \frac{3}{k^2} + \frac{4}{k^3})} = 1.797439588835227\ldots \tag{2}$$

Simple lookup, smart lookup, and integer relations as embodied in the ISC all fail to tell us anything about these numbers. In fact, (1) is

$$\prod_{k \geq 1} \frac{(1 + \frac{1}{k} + \frac{1}{k^2})^2}{(1 + \frac{2}{k} + \frac{3}{k^2})} = \frac{3\sqrt{2}}{\pi} \, \frac{\cosh^2(\pi \frac{\sqrt{3}}{2})}{\sinh(\pi \sqrt{2})},$$

but this is a strange enough formula that we aren't surprised that the ISC can't identify it. We do not know any closed form expression for (2), however.

The "generalized expansions" option guesses that there is a simple generating function for the "egyptian fraction" of the reciprocal of (1), namely

$$\frac{x(69x - 2)}{47x - 1},$$

but *this is incorrect*, and it is easy to disprove this conjecture by computing the series expansion

$$\frac{x(69x - 2)}{47x - 1} = 2\,x + \sum_{k \geq 2} \left(25 \cdot 47^{k-2}\right) x^k \tag{3}$$

and evaluating the rational number that is the "egyptian fraction" defined by the coefficients of the series (3):

$$\frac{1}{2} + \sum_{k \geq 2} \frac{1}{25 \cdot 47^{k-2}} = \frac{311}{575} = .540869565\ldots\,, \tag{4}$$

whereas the reciprocal of (1) starts off as .5408515498..., which differs from 311/575 after the fourth decimal place. Similarly, the ISC's generalized expansions return an incorrect egyptian fraction for the reciprocal of (2).

An 'egyptian fraction' is just an ordinary rational written as a sum of reciprocals of natural numbers without repeated entries in the sum. The egyptian fraction (4) corresponding to the generating function (3) is a geometric series.

So, we have seen examples where the ISC tells us something useful, tells us something incorrect, and tells us nothing.

2 A connection between the Lambert W function and Stirling's formula for $n!$

We now look at a more interesting example, using the online version of the Encyclopedia of Integer Sequences [27] (http:// www.research.att.com/ ~njas/ sequences/).

The Lambert W function satisfies

$$W(x)e^{W(x)} = x\,. \tag{5}$$

See [14] for a survey of properties and applications of W, together with some of its history; [16] explores various series for W including the one we discuss in this section. We give a short introduction to this function in Appendix A.

There is a branch point of W at $x = -1/e$, where $W(x) = -1$. See Figure 1, which can be produced in MAPLE by the command

```
> plot( [ t*exp(t), t, t=-5..1 ], -1..3, -4..1 );
```

The two real-valued branches of W are denoted $W_0(x)$ and $W_{-1}(x)$; we also refer to $W_0(x)$ as the *principal branch*. We wish to know more about the function near the branch point at $x = -1/e$. After various experiments, we decide to compute the series of

$$W_0\left(-e^{-1-z^2/2}\right)$$

in MAPLE. We get, very quickly, that $W_0\left(-\exp(-1-z^2/2)\right) =$

$$\begin{aligned} &-1 + z - \frac{1}{3}z^2 + \frac{1}{36}z^3 + \frac{1}{270}z^4 + \frac{1}{4320}z^5 - \frac{1}{17010}z^6 \\ &- \frac{139}{5443200}z^7 - \frac{1}{204120}z^8 - \frac{571}{2351462400}z^9 + O\left(z^{10}\right)\,. \end{aligned} \tag{6}$$

As our first real example of using a new tool, we look up the sequence of denominators 1, 3, 36, 270, 4320, ..., in [27]. [Sometimes denominators have

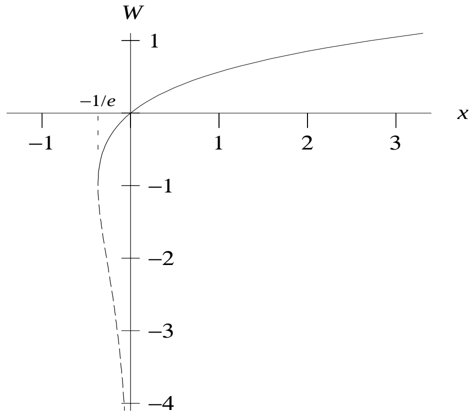

Figure 1: The real branches of the function $W(x)$ that satisfies $W \exp W = x$.

nontrivial common factors with numerators. Cancellation of these common factors will make any "guessing" procedure more difficult.] We find the sequence immediately, and the Encyclopedia gives a reference to the delightful paper [23], which does not mention W or refer to any papers on W, or indeed even use it explicitly. Thus, [23] would not easily be found by a normal citation search. We find out in [23] that equation (6) gives coefficients needed in Stirling's formula for $n!$, which begins

$$n! \sim \sqrt{2\pi n} n^n e^{-n} \left(1 + \frac{1}{12\,n} + \frac{1}{288\,n^2} - \frac{139}{51840\,n^3} - \frac{571}{2488320\,n^4} + O\left(\frac{1}{n^5}\right) \right) .$$

The connection we discover (without doing any work ourselves) is that if

$$W_0\left(-e^{-1-z^2/2}\right) = \sum_{k\geq 0} (-1)^{k-1} a_k\, z^k ,$$

then

$$n! \sim \sqrt{2\pi n}\, n^n e^{-n} \sum_{k\geq 0} \frac{1\cdot 3\cdot 5\cdots(2k+1)}{n^k} a_{2k+1} ,$$

and moreover there is a lovely (and useful!) recurrence relation for the a_k's, namely $a_0 = 1$, $a_1 = 1$, and

$$a_n = \frac{1}{(n+1)a_1}\left(a_{n-1} - \sum_{k=2}^{n-1} k a_k a_{n+1-k} \right) .$$

3 Riemann Surfaces

Tools such as MATLAB and MAPLE permit easy and accurate visualization of Riemann surfaces for elementary functions [15, 28]. Our qualitative understanding of even extremely basic mathematical building blocks can thus be affected by mathematical software tools. See [11] for more discussion of visualization in general; here we concentrate on a simple technique for visualization of Riemann surfaces, namely to make 3-d plots of $\Re f(z)$ or $\Im f(z)$.

It is necessary to *prove* something about this technique—namely, that it really gives us a good picture of the Riemann surface and not just a 3-d plot of the imaginary part (or the real part) of the function involved. This is pursued in more detail in [15], but the key point is that given $w = u+iv = f(z) = f(x+iy)$, then we get an accurate Riemann surface by plotting, say, (x, y, v) if and only if the missing piece of information (here, u) is completely determined once x, y, and v are given. This is simple, if not quite obvious: once we have a smooth three-dimensional surface, each point of which can be associated with a unique value (i.e. ordered pair) of the map $z \mapsto w = f(z)$, then we have a representation of the Riemann surface of f.

This exact association is not automatic. For example, if $w = \ln(z)$ and we plot (x, y, u), then we do *not* get a picture of the Riemann surface for logarithm, because the branch of $v = \Im(w) = \arg(z)$ is not determined from $u = \ln(x^2 + y^2)/2$, x, and y. If we plot (x, y, v), of course, we *do* recover the classical picture of the Riemann surface for $\ln(z)$, because given x, y, and v we can easily find u.

The following short piece of MAPLE code shows how to graph the Riemann surface for the Lambert W function. We urge you to try the following computation, because the dynamic coloured picture you get is much more easily understood than the static black-and-white image in Figure 2. We also urge you to try your hand at your own functions; many others are graphed in [15] and [28].

```
>  w := u + I*v;
```

$$w := u + I\,v$$

```
>  z := w*exp(w);
```

$$z := (u + I\,v)\,e^{(u+I\,v)}$$

```
>  evalc(z);
```

$$u\,e^u \cos(v) - v\,e^u \sin(v) + I\,(v\,e^u \cos(v) + u\,e^u \sin(v))$$

```
>  x := evalc(Re(z));
```

$$x := u\,e^u \cos(v) - v\,e^u \sin(v)$$

```
>  y := evalc(Im(z));
```

$$y := v\,e^u \cos(v) + u\,e^u \sin(v)$$

```
> plot3d([x,y,v], u=-6..1,
        v=-5..5, axes=FRAME,
        orientation=[-110,73],
        labels=["x","y","v "],
        style=PATCHNOGRID,
```

```
      colour=u,
      view=[-1..1,-1..1,-5..5],
      grid=[50,50]);
```

See Figure 2 for a static representation of the results of this plot command.

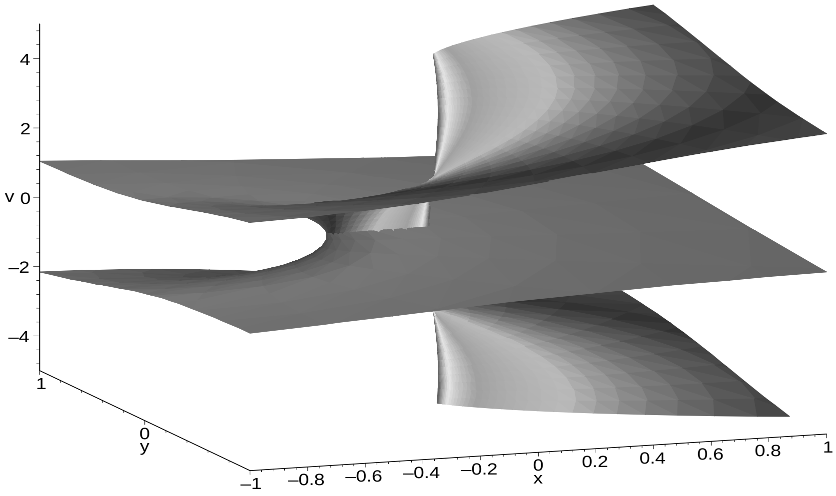

Figure 2: The Riemann surface for the Lambert W function.

3.1 1–1 correspondence proof

Given x, y, and v, we have to solve for u. Of course, one takes the existence of (u, v) for a given (x, y) for granted here; for the Lambert W function, a proof can be found in [14]. We have

$$(u + iv)e^{u+iv} = x + iy \ ,$$

which gives

$$ue^u + ive^u = (x + iy)e^{-iv} = (x + iy)(\cos v - i \sin v) \ ;$$

therefore,

$$ue^u + ive^u = (x \cos v + y \sin v) + i(y \cos v - x \sin v) \ .$$

If $v \neq 0$, and moreover $y \cos v - x \sin v \neq 0$, then dividing the real part by the imaginary part gives u in terms of x, y, and v:

$$u = \frac{v(x \cos v + y \sin v)}{y \cos v - x \sin v} .$$

This solution is unique. Investigation of the exceptional conditions $v = 0$ or $y \cos v - x \sin v = 0$ leads to $u \exp u = x$, which has two solutions if and only if $-1/e \leq x < 0$, in the case $v = 0$, and to the singular condition $u = -\infty$ and $x = y = 0$.

This is precisely what we observe in the graph: two sheets intersect only if $-1/e \leq x < 0$ (note that the colours are different and hence the corresponding sheets on the Riemann surface do not "really" intersect), and all sheets have a singularity at the origin, except the central one, which contains $v = 0$. This is as good a representation of the Riemann surface for the Lambert W function as can be produced in three dimensions.

However, Figure 2 is nowhere near as intelligible as the live MAPLE plot. On a PC, the use of OpenGL by MAPLE allows the plot to be rotated by direct mouse control. This helps to give a good sense of what the surface is really like, in three dimensions.

4 Dynamical systems, numerical analysis, and formal power series

In this section we give a brief overview of a surprising connection between numerical analysis of dynamical systems and formal power series. We begin with a simple question: what, exactly, does the fixed time step *forward Euler* numerical method do to the solution of the simple initial value problem

$$\frac{dy}{dt} = y^2 \tag{7}$$

with $y(0) = y_0$? The numerical procedure is just

$$y_{n+1} = y_n + h y'_n \tag{8}$$

for integer $n \geq 0$, where $y'_n = y_n^2$ and $h > 0$ is the chosen time step.

It turns out to be useful to rescale y and t so that $v = hy$ and $\tau = ht$, giving

$$\frac{dv}{d\tau} = v^2 , \tag{9}$$

and (8) becomes

$$v_{n+1} = v_n + v_n^2 . \tag{10}$$

We may then rephrase our question to ask instead what the relationship between v_n and $v(\tau)$ is.

The point of view taken in [13] is that of *backward error analysis*. That is, instead of asking for the difference between $v(n)$ and v_n, we ask instead if there is another differential equation, say

$$\frac{dw}{d\tau} = B(w)w^2 \ , \tag{11}$$

whose solution *interpolates* v_n. That is, we impose the conditions $w(0) = v_0$ and $w(\tau+1) = w(\tau) + w(\tau)^2$ (*cf.* (10)), and see if we can find such a function $B(w)$.

We do this not so we can improve the behaviour of Euler's method for this problem, but rather so that we may understand what Euler's method has done to the problem; for by understanding the function $B(w)$ we will learn something about Euler's method, by comparing (9) to (11).

It turns out that we can use *the method of modified equations* [19] to find as many terms of the Taylor series for $B(w)$ as we desire. When we compute the modified equation for (7) to (say) fifth order, we get

$$\frac{dw}{dt} = \left(1 - w + \frac{3}{2!}w^2 - \frac{16}{3!}w^3 + \frac{124}{4!}w^4 - \frac{1256}{5!}w^5\right) w^2 \ . \tag{12}$$

Now we see the sequence 1, −1, 3, −16, 124, −1256 appearing. This is sequence M3024 in [27], which points us directly to the very beautiful and useful paper [21].

We find in that paper that if

$$B(w) = \sum_{n\geq 0} c_n w^n \ , \tag{13}$$

then

$$c_n = \frac{1}{n-1}\sum_{i=1}^{n-1}\binom{n-i+1}{i+1} c_{n-i} \ ,$$

and this, combined with the functional equation

$$B(w) = \frac{(1+w)^2}{1+2w} B(w+w^2)$$

(which can be iterated to give us two converging infinite products for B), allows us to write an efficient program to evaluate $B(w)$. We can show that $B(w)$ has a pole at $w = -1/2$; by mapping backwards, solving $w + w^2 = -1/2$, we find two more (complex) poles. Iterating this process finds an infinite number of complex poles, approaching the Julia set for the map $v \to v + v^2$ arbitrarily closely; see Figure 3.

The Julia set itself approaches the origin arbitrarily closely. That is, there are poles arbitrarily close to the point of expansion of the series given for B. Thus the series (13) diverges—but, nonetheless, it can be used to evaluate $B(w)$ for w close enough to zero, using MAPLE's built-in sequence acceleration techniques.

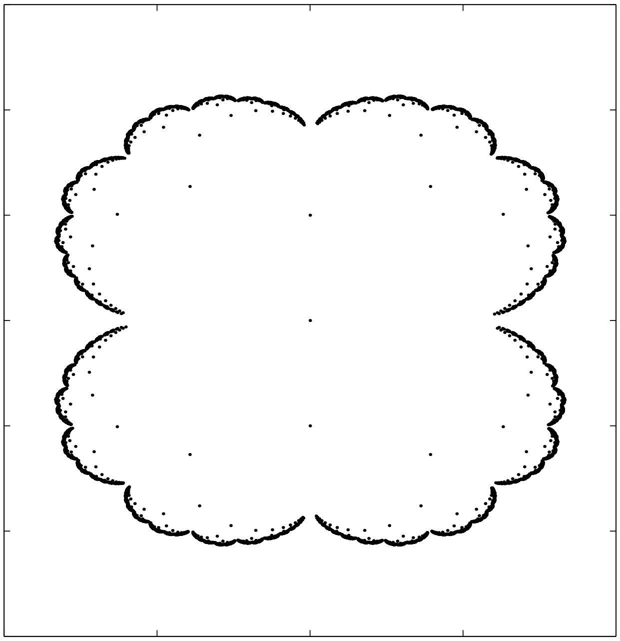

Figure 3: The first 16000 poles of $B(v)$, approaching the Julia set of $v \to v + v^2$.

This is precisely where the convergent infinite products are slow, and hence the series is useful. See [13] for details.

But more to the point, in [21], G. Labelle completely solves the problem of interpolating discrete dynamical systems with continuous dynamical systems, in the domain of formal power series. The mathematical language, however, is quite different from that used in the numerical analysis world. As an example, in [21] the 'modified equation' is termed an 'infinitesimal generator' for the discrete dynamical system. Therefore, simple subject searches might not find [21]. Indeed, a combinatorics journal seems an unlikely place to find the solution of a problem in the numerical analysis of dynamical systems, but the Encyclopedia of Integer Sequences provides a way to search the 'knowledge database' that is keyed on the *examples*, or the *concrete results*, of papers—not the jargon. This, if you like, is a *new kind* of search tool.

5 An integer-relation example

The following is taken from [10]. As a didactic example, suppose that we are interested in finding the value of the definite integral

$$V = \int_0^\infty \frac{\sqrt{x}\ln^5 x}{(1-x)^5}\,dx\ , \tag{14}$$

and that we suspect that V could be expressed as a polynomial in π^2, of low degree, with short rational coefficients.

Such a conjecture might arise naturally from consideration of

$$\begin{aligned}
\int_0^\infty \frac{\sqrt{x}\ln^2 x}{(1-x)^2}dx &= 2\pi^2 \\
\int_0^\infty \frac{\sqrt{x}\ln^3 x}{(1-x)^3}dx &= \frac{1}{4}\pi^2\left(\pi^2-12\right) \\
\int_0^\infty \frac{\sqrt{x}\ln^4 x}{(1-x)^4}dx &= -\frac{1}{3}\pi^2\left(\pi^2-12\right)\ ,
\end{aligned}$$

for example, and we may suppose that these values are known already, for the sake of argument. One can use the "mellin" routine of the "inttrans" package in MAPLE to evaluate all these (and V) symbolically—so this example is really just expository.

To be explicit, we conjecture that

$$V = r_1 + r_2\pi^2 + r_3\pi^6 + r_4\pi^6\ ,$$

where all the r_i are short rational numbers. Instead of trying to derive the coefficients of the this polynomial analytically, we can use numerical approximation and a lattice basis reduction algorithm, the LLL algorithm given in [22], to identify the coefficients heuristically. In an ideal world, we would then know what we had to prove, and, knowing that, would find the proof easier.

We give a short overview of using the LLL algorithm to find integer relations. Suppose that we have a finite set B of n-dimensional linearly independent vectors with rational entries. We call the set

$$L = \left\{ \sum_{v\in B} r_v v \,\middle|\, r_v \in \mathbb{Z} \right\}$$

"the lattice spanned by B". We say that the lattice has dimension n, and that B is a basis for the lattice. There may be many other bases for the lattice, and we often want to find particular bases with nice properties. For many applications, and in particular for finding integer relations, what we would really like to have is "the basis with the shortest Euclidean length". Unfortunately, the problem of determining whether one has *the shortest* basis may be NP-complete [22]. But finding *a short* basis is often just as helpful, and the LLL algorithm [22] can, in polynomial time, find "relatively" short vectors; guaranteed, in fact, to be of length at most $2^{n-1}\ell$, where ℓ the shortest possible. In practice the LLL algorithm often returns vectors much better than this bound.

To proceed in MAPLE, we choose a large constant C and form the following matrix, and use the lattice reduction subroutine.

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & C\cdot 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & C\cdot\pi^2 \\ 0 & 0 & 1 & 0 & 0 & 0 & C\cdot\pi^4 \\ 0 & 0 & 0 & 1 & 0 & 0 & C\cdot\pi^6 \\ 0 & 0 & 0 & 0 & 1 & 0 & C\cdot\pi^8 \\ 0 & 0 & 0 & 0 & 0 & 1 & C\cdot V \end{bmatrix}$$

```
>  readlib( lattice );
```

$$\mathbf{proc}()\ \ldots\ \mathbf{end}$$

We work to 30 digits for this example. In general, one has to experiment to find how many digits one needs.

```
>  Digits := 30;
```

$$\mathrm{Digits} := 30$$

We compute an approximation for the value V that we wish to identify, and approximations of the quantities that we wish to relate to V.

```
>  V := Int( sqrt(x)*ln(x)^5/(1-x)^5, x=0..infinity);
```

$$V := \int_0^\infty \frac{\sqrt{x}\,\ln(x)^5}{(1-x)^5}\,dx$$

```
>  lastcol := [ seq( evalf( Pi^(2*i)), i=0..4 ), evalf(V) ];
```

lastcol := [1., 9.86960440108935861883449099988,
97.4090910340024372364403326888, 961.389193575304437030219443653,
9488.53101607057400712857550392, −16.6994737192290704961872434007]

We now choose a large constant C. We use the size of C to penalize vectors that do not combine to zero.

```
>  C := 10^15;
```

$$C := 1000000000000000$$

We construct the rows of the matrix that we need, as follows.

```
>  for i to 6 do
>    row.i := [ seq(0, j=1..7) ]:
>    row.i[i] := 1:
>    row.i[7] := C* lastcol[i]:
>  od:
>  B := [ seq( row.i, i=1..6 ) ];
```

$$
\begin{aligned}
B := [[&1,\ 0,\ 0,\ 0,\ 0,\ 0,\ .1000000000000000\,10^{16}],\\
&[0,\ 1,\ 0,\ 0,\ 0,\ 0,\ .986960440108935861883449099988\,10^{16}],\\
&[0,\ 0,\ 1,\ 0,\ 0,\ 0,\ .974090910340024372364403326888\,10^{17}],\\
&[0,\ 0,\ 0,\ 1,\ 0,\ 0,\ .961389193575304437030219443653\,10^{18}],\\
&[0,\ 0,\ 0,\ 0,\ 1,\ 0,\ .948853101607057400712857550392\,10^{19}],\\
&[0,\ 0,\ 0,\ 0,\ 0,\ 1,\ -.166994737192290704961872434007\,10^{17}]]
\end{aligned}
$$

Now we call the "lattice" routine to compute a short basis for the set generated by these rows.

```
>  lattice( B );
```

$$
\begin{aligned}
[[&0,\ 120,\ 140,\ -15,\ 0,\ 24,\ .622\,10^{-11}],\\
&[-16743,\ 51,\ 156,\ 10,\ -1,\ -55,\ 6738.90916826007994],\\
&[35146,\ -443,\ -57,\ -16,\ -1,\ 21,\ 19729.34720281002100],\\
&[6349,\ -2221,\ 94,\ 2,\ 0,\ -269,\ -7554.67120587589348],\\
&[-2452,\ -99,\ 8,\ -3,\ 2,\ 805,\ 5948.36266979182662],\\
&[32181,\ 345,\ 9,\ -11,\ -1,\ 982,\ -19383.09100001444674]]
\end{aligned}
\tag{15}
$$

All of these new basis vectors are of the form

$$\left[r_1, r_2, r_3, r_4, r_5, r_6, C \sum_{i=1}^{6} r_i a_i \right] ,$$

where the r_i are integers. This is because each new vector is an integer linear combination of rows of the initial matrix. Because the initial matrix was an augmented identity matrix, the coefficients of the requisite integer combination show up in the result. Because we chose C to be so large, looking for a short vector in this space really biases the search towards places where the integer linear combination of the final column is zero, if there are any. Hence we suspect, from the first row of (15), that

$$120\pi^2 + 140\pi^4 - 15\pi^6 + 24V = 0$$

or

$$V = \frac{5}{24}\pi^2 \left(3\pi^4 - 28\pi^2 - 24\right) .$$

Issuing the following Maple command lends credence to our suspicion.

```
>  evalf( V - 5/24*Pi^2*(3*Pi^4-28*Pi^2-24), 100);
```

$$-.8\,10^{-97}$$

There is a simpler Web-based implementation, which uses the "EZface" to emulate a more comprehensive GNU MP implementation of this method. Go to http:// www.cecm.sfu.ca/ MRG/ INTERFACES.html, click on "EZface", and type in the following:

```
lindep([1.,
9.869604401089358618834490999988,
97.40909103400243723644033268888,
961.3891935753044370302194436 53,
9488.531016070574007128575503 92,
-16.699473719229070496187243400 7])
```

Then, select 30 digits of precision, and click “evaluate”. Very quickly, the vector

```
0, -120., -140., 15., 0, -24.
```

is returned—voilá, our integer relation.

Issuing the command “lindep” calls a subroutine that looks for short integer linear dependencies among the given vector of numbers. Again its results are to be considered as “possible relations”, to be proved later.

The LLL approach is very powerful and, if used with imagination, offers rich possibilities for discovery.

One danger, indeed a ‘mathematical sin’, is indicated by the following quotes:

> “Number theory is full of things that everyone *knows* are true, but no one is able to prove.” —Peter Borwein
>
> “I have absolutely no interest in proving things that I know are true.” —Henry Abarbanel, at a dynamical systems workshop at Penn State in 1994.

We discuss ‘redemption’ of sins such as this in the final section.

Numerical instability in the LLL algorithm may cause difficulty, as well. Here we have simply worked to “enough” digits to mitigate its effects—that is, we are trying to buy more accuracy by paying for more precision. This is often expensive, and PSLQ, discussed in Section 6.1, is better.

6 How solvable is ‘solvable’?

This example is also taken from [10]. The following problem arises when thinking about modular (theta) functions; see [6]. If we define

$$\begin{aligned} a(q) &:= \sum_{m,n\in\mathbb{Z}} q^{m^2+mn+n^2} \\ b(q) &:= \sum_{m,n\in\mathbb{Z}} \omega^{n-m} q^{m^2+mn+n^2} \\ c(q) &:= \sum_{m,n\in\mathbb{Z}} q^{(n+1/3)^2+(n+1/3)(m+1/3)+(m+1/3)^2} \end{aligned}$$

where $\omega = \exp(2\pi i/3)$, then we have

$$a^3 = b^3 + c^3$$

and a lovely parameterization of the ${}_2F_1$ hypergeometric function [4]:

$$F\left(\left.\begin{matrix}\frac{1}{3},\frac{2}{3}\\ 1\end{matrix}\right|\frac{c^3}{a^3}\right)=a\,.$$

Choosing $q=\exp(-2\pi\sqrt{N/3})$ for $N\in\mathbb{Q}$, it can be shown that $s_N := c/a$ is an algebraic number expressible by radicals; see [6]. If N is a positive integer, then s_N is called the Nth *cubic singular value*. What can we discover computationally about s_N? For example, can we determine radical formulae for the higher order cubic singular values?

RMC—removed reference to CL. Is is necessary? Should we put it back?

The following observations help the efficiency of the computations. It is known that

$$\begin{aligned}a(q) &= \theta_3(q)\theta_3(q^3)+\theta_2(q)\theta_2(q^3)\\ b(q) &= (3a(q^3)-a(q))/2\\ c(q) &= (a(q^{1/3})-a(q))/2\,,\end{aligned}$$

where

$$\theta_2(q)=\sum_{n\in\mathbb{Z}} q^{(n+1/2)^2}\qquad\text{and}\qquad \theta_3(q)=\sum_{n\in\mathbb{Z}} q^{n^2}$$

are the classical theta functions. The lacunarity of these series allows for very rapid computation.

6.1 A useful transformation

A further transformation, which makes the as yet unknown minimal polynomial simpler, is useful. After examining the patterns in the first few cases s_1, s_2, s_3, …, and using the analogous classical quadratic singular values (where one sees the forms $4k_N^2(1-k_N^2)$ or $(1-k_N^2)/2k_N$ depending on the parity of N), the authors of [10] thought to look at

$$G_N := \left(\frac{1}{2}-s_N\right)^2 \qquad\qquad N\not\equiv 0 \bmod 3$$

or

$$g_N := \frac{3s_N}{1-s_N^3} \qquad\qquad N\equiv 0 \bmod 3\,;$$

the minimal polynomial for G_N or g_N then has lower degree than the minimum polynomial for s_N. This makes the polynomial easier to find by the PSLQ algorithm.

The PSLQ algorithm (see [17]) and the LLL algorithm can both be used to find integer relations (and hence minimal polynomials for an algebraic number α, by looking for an integer relation among $1, \alpha, \ldots, \alpha^m$). However, PSLQ can also produce *negative* results. If PSLQ fails to find an integer relation, then one can usually say that *there is no such relation with coefficients less than a* *computable bound*, effectively proving that there is no "simple" relation of the guessed form.

The authors of [10] used these ideas to 'decode' the numerical values of s_N into radical form, up to $N = 100$, and some values beyond, such as $N = 110$ and 154. They used a variety of strategies to verify the results; some ingenuity was necessary in order to extract the radicals. For $N < 53$, they computed P_N, the minimal polynomial for G_N or g_N; they then tried factoring P_N over different quadratic number fields until they got a factor of degree 4 or less, which they solved in radicals. This approach failed at $N = 53$, where they first had to use a special MAPLE program for finding a radical for any solvable quintic. (See ftp:// calfor.lip6.fr:/ pub/ softwares/ Maple/ quinticV2.gz.) The radical returned for $N = 53$ has over 7500 symbols in it. Kevin Hare at the CECM refined it to an equivalent but simpler radical with 'only' 860 symbols. MAPLE was able to verify symbolically that this simpler radical solved P_{53}. In general, determining that a symbolic equation is indeed zero is, in certain classes of expressions, computationally undecidable [26].

Indeed, the point of this whole exercise was to determine how good both MAPLE's symbolic tools and PSLQ's numerical ones were on "grand challenge" examples. Experience with exercises such as this have led to improvements in both tools.

If symbolic verification is not possible, reassurance that the results are correct can often be obtained by using Klein's *absolute invariant* [3, p. 115]

$$J_2(x) = \frac{4}{27}\frac{(1-x^2(1-x^2))^3}{x^4(1-x^2)^2}\,,$$

and its cubic counterpart

$$J_3(x) = \frac{1}{64}\frac{(9-8x^3)^3}{x^9(1-x^3)}\,.$$

If our computed s_N is correct, then it is related to the (known) classical singular value k_{3N} by

$$J_2(k_{3N}) = J_3(s_N)\,. \tag{16}$$

The identity (16) can be derived from Proposition 5.8 in [3, p. 185, (5.5.26)]. It can be checked symbolically in MAPLE for the radicals arising in the cases $N \leq 10$. For larger N, some human intervention is required. For $N = 70$, the verification requires use of k_{210}, the computation of which Hardy called "one of the most striking of Ramanujan's results" [20, p. 229]. We note that purely numerical computation, together with analytic reasoning about such computation (some of which is automatable) can be used to verify the results. Standard irrational number theoretic techniques allow one to show that either $J_2(k_{210}) = J_3(S)$ or $|J_2(k_{210}) - J_3(S)| > 10^{-6400}$, where S is our heuristically guessed radical formula for s_{70}. Given this knowledge, a few minutes of CPU time establishes that $|J_2(k_{210}) - J_3(S)| < 10^{-6400}$, and thus $J_2(k_{210}) = J_3(S)$.

7 Final Vignettes

Integer relation algorithms have already helped to discover many new results. We list a few of these, again taken from [10]. The number of such results continues to climb. We have to tell the algorithms what kind of relationship to look for, but, given that, the algorithms allow previously impossible jumps.

7.1 Zeta value series

The formula for $\zeta(3)$ used by Apéry to prove that $\zeta(3)$ is irrational, namely

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

has no analog for $\zeta(2n+1)$ with $n \geq 2$. It can be shown using PSLQ (or more simply the Euclidean algorithm since there are only two unknown integers) that if a formula like

$$\zeta(5) = \frac{p}{q} \sum_{k \geq 1} \frac{(-1)^{k+1}}{k^5 \binom{2k}{k}}$$

exists, then the integers p and q are larger than 10^{300}.

There is a similar but more complicated formula for $\zeta(5)$, due to Koecher, that *does* suggest generalization, however. Borwein and Bradley used an LLL algorithm to determine the new coefficients [7]. They found that

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

and they discovered similar formulas for $\zeta(4n+3)$ for $2 \leq n \leq 10$ that involve linear combinations of sums of the form

$$H_{m,n} = \sum_{k \geq 1} \frac{(-1)^{k+1}}{k^{4m+3} \binom{2k}{k}} \sum_{j=1}^{k-1} \frac{1}{j^{4(n-m)}}$$

and multiple dimensional analogues. They conjectured the following generating function:

$$\begin{aligned} \sum_{n \geq 0} \zeta(4n+3) z^{4n} &= \sum_{k \geq 1} \frac{1}{k^3 (1 - z^4/k^4)} \\ &= \frac{5}{2} \sum_{k \geq 1} \frac{(-1)^{k+1}}{k^3 \binom{2k}{k}} \frac{1}{(1 - z^4/k^4)} \prod_{m=1}^{k-1} \frac{1 + 4z^4/m^4}{1 - z^4/m^4}, \end{aligned} \tag{17}$$

where the final infinite sum is quite unexpected. However, from the first ten cases it was apparent that the series had the form

$$\frac{5}{2} \sum_{k \geq 1} \frac{(-1)^{k+1}}{k^3 \binom{2k}{k}} \frac{1}{(1 - z^4/k^4)} P_k(z)$$

for as yet undetermined P_k; and there were abundant data to compute

$$P_k(z) = \prod_{m=1}^{k-1} \frac{1+4z^4/m^4}{1-z^4/m^4} \ .$$

They reduced the conjectured formula to an equivalent finite sum

$$\frac{5}{2}\sum_{k=1}^{n} \binom{2k}{k} \frac{k^2}{4n^4+k^4} \prod_{j=1}^{k-1} \frac{n^4-j^4}{4n^4+j^4} = \frac{1}{n^2} \tag{18}$$

($1 \le n < \infty$) that was subsequently proved by Almkvist and Granville [1]. Series expansion of the finite products in (17) gives a rapidly converging series for any $\zeta(4n+3)$, which someone may be able to use to settle the question of the irrationality of the ζ values.

7.2 Independent computation of digits of π

The following formula, discovered using the PSLQ algorithm, allows rapid computation of hexadecimal digits of π independently of previous digits [2]:

$$\pi = \sum_{k\ge 0} \left(\frac{1}{16}\right)^k \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6}\right) \ . \tag{19}$$

Bailey, Borwein, and Plouffe knew that a fast algorithm would result from a formula of this form, and deliberately used a computer search to find it; some have called this approach *mathematical reverse engineering*. Once known, the formula can be proved very concisely by a human [2]. Interestingly, the following MAPLE session shows that it can now be proved almost automatically, too.

```
> p := Sum( (1/16)^k*(4/(8*k+1)-2/(8*k+4)-1/(8*k+5)-1/(8*k+6)),
k=0..infinity );
```

$$p := \sum_{k=0}^{\infty} (\frac{1}{16})^k \, (4\,\frac{1}{8\,k+1} - 2\,\frac{1}{8\,k+4} - \frac{1}{8\,k+5} - \frac{1}{8\,k+6})$$

The following shows a temporary increase in complexity. This phenomenon is called "intermediate expression swell".

```
> value( p );
```

$$\frac{47}{15}\,\mathrm{hypergeom}([1, \frac{1}{2}, \frac{3}{4}, \frac{5}{8}, \frac{1}{8}], [\frac{3}{2}, \frac{13}{8}, \frac{9}{8}, \frac{7}{4}], \frac{1}{16})$$
$$+ \frac{47}{8192} \frac{\mathrm{hypergeom}([2, \frac{3}{2}, \frac{13}{8}, \frac{9}{8}, \frac{7}{4}], [\frac{5}{2}, \frac{21}{8}, \frac{17}{8}, \frac{11}{4}], \frac{1}{16})}{\frac{123669}{40960} - \frac{819}{40960}\sqrt{241}}$$
$$+ \frac{1504}{36855} \frac{(\frac{391}{8192} - \frac{1}{8192}\sqrt{241})\,\mathrm{hypergeom}([2, \frac{3}{2}, \frac{13}{8}, \frac{9}{8}, \frac{7}{4}], [\frac{5}{2}, \frac{21}{8}, \frac{17}{8}, \frac{11}{4}], \frac{1}{16})}{(\frac{151}{240} + \frac{1}{240}\sqrt{241})\,(\frac{151}{240} - \frac{1}{240}\sqrt{241})}$$

$$+\frac{47}{1920}\,\frac{(\frac{391}{8192}-\frac{1}{8192}\sqrt{241})\,\mathrm{hypergeom}([3,\,\frac{5}{2},\,\frac{21}{8},\,\frac{17}{8},\,\frac{11}{4}],\,[\frac{7}{2},\,\frac{29}{8},\,\frac{25}{8},\,\frac{15}{4}],\,\frac{1}{16})}{(\frac{151}{240}+\frac{1}{240}\sqrt{241})\,(\frac{151}{240}-\frac{1}{240}\sqrt{241})\,(\frac{511819}{8192}-\frac{1309}{8192}\sqrt{241})}$$

Looking at those conjugate radicals in the denominators suggests expansion—this step is natural but not automatic.

```
> normal( %, expanded );
```

$$\begin{aligned}&\frac{47}{15}\,\mathrm{hypergeom}([1,\,\frac{1}{2},\,\frac{5}{8},\,\frac{3}{4},\,\frac{1}{8}],\,[\frac{3}{2},\,\frac{9}{8},\,\frac{13}{8},\,\frac{7}{4}],\,\frac{1}{16})\\&\quad+\frac{271}{39312}\,\mathrm{hypergeom}([2,\,\frac{3}{2},\,\frac{9}{8},\,\frac{13}{8},\,\frac{7}{4}],\,[\frac{5}{2},\,\frac{17}{8},\,\frac{21}{8},\,\frac{11}{4}],\,\frac{1}{16})\\&\quad+\frac{1}{20944}\,\mathrm{hypergeom}([3,\,\frac{5}{2},\,\frac{17}{8},\,\frac{21}{8},\,\frac{11}{4}],\,[\frac{7}{2},\,\frac{25}{8},\,\frac{29}{8},\,\frac{15}{4}],\,\frac{1}{16})\qquad(20)\end{aligned}$$

As an aside, equation (20) is an interesting identity itself. In the notation of [18], it implies (once the proof is completed) that

$$\begin{aligned}\pi &= \frac{47}{15}F\left(\begin{matrix}1,\frac{1}{2},\frac{5}{8},\frac{3}{4},\frac{1}{8}\\ \frac{3}{2},\frac{9}{8},\frac{13}{8},\frac{7}{4}\end{matrix}\middle|\frac{1}{16}\right)+\frac{271}{39312}F\left(\begin{matrix}2,\frac{3}{2},\frac{9}{8},\frac{13}{8},\frac{7}{4}\\ \frac{5}{2},\frac{17}{8},\frac{21}{8},\frac{11}{4}\end{matrix}\middle|\frac{1}{16}\right)\\&\quad+\frac{1}{20944}F\left(\begin{matrix}3,\frac{5}{2},\frac{17}{8},\frac{21}{8},\frac{11}{4}\\ \frac{7}{2},\frac{25}{8},\frac{29}{8},\frac{15}{4}\end{matrix}\middle|\frac{1}{16}\right).\end{aligned}$$

The next step the mechanical proof of the Bailey-Borwein-Plouffe π formula simplifies the hypergeometric functions:

```
> convert( %, StandardFunctions );
```

$$\begin{aligned}&-\frac{1}{2}\sqrt{2}(\ln(1-\frac{1}{2}\sqrt{2})-\ln(1+\frac{1}{2}\sqrt{2})+\frac{1}{2}\sqrt{2}\ln(\frac{1}{2})-\sqrt{2}\arctan(1)-2\arctan(\frac{1}{2}\sqrt{2})\\&\quad-\frac{1}{2}\sqrt{2}\ln(\frac{5}{2})-\sqrt{2}\arctan(\frac{1}{3}))+\ln(\frac{3}{4})-\ln(\frac{5}{4})+\frac{1}{2}\sqrt{2}(\ln(1-\frac{1}{2}\sqrt{2})-\ln(1+\frac{1}{2}\sqrt{2})\\&\quad-\frac{1}{2}\sqrt{2}\ln(\frac{1}{2})+\sqrt{2}\arctan(1)-2\arctan(\frac{1}{2}\sqrt{2})+\frac{1}{2}\sqrt{2}\ln(\frac{5}{2})+\sqrt{2}\arctan(\frac{1}{3}))+\ln(\frac{1}{2})\\&\quad-\ln(\frac{3}{2})+2\arctan(\frac{1}{2})\end{aligned}$$

The next step is not necessary, but it slows down the computation so we can see that many of the terms in the above formula simply cancel.

```
> expand( % );
```

$$\frac{1}{2}\pi+2\arctan(\frac{1}{3})+2\arctan(\frac{1}{2})$$

Now, finally, our answer is plain:

```
> simplify( % );
```

$$\pi$$

A somewhat more efficient version of (19) was discovered by Fabrice Bellard. This has led Colin Percival, an undergraduate student at Simon Fraser, to design an ingenious parallel internet computation of staggeringly high order hexadecimal digits of π. Details may be found at http:// www.cecm.sfu.ca/ projects/ pihex/: the five trillionth bit of π is '0'.

7.3 Fast series for the Catalan constant

Consider the *Catalan constant*, which can be defined by

$$G = \sum_{k \geq 0} \frac{(-1)^k}{(2k+1)^2} \tag{21}$$

or alternatively by

$$G = -\int_0^{\pi/4} \log \tan \theta \, d\theta = -\int_0^1 \frac{\log u}{1+u^2} \, du \, .$$

This is perhaps the simplest constant whose irrationality is still unsettled.

Ramanujan discovered the following series for G, which converges *much* more quickly than (21) [3, p. 386]:

$$G = \frac{\pi}{8} \log(2+\sqrt{3}) + \frac{3}{8} \sum_{k \geq 0} \frac{1}{(2k+1)^2 \binom{2k}{k}} \, . \tag{22}$$

After many false starts, David Bradley found a new family of series that includes (22). One member of this family is

$$G = \frac{\pi}{8} \log \left(\frac{10+\sqrt{50-22\sqrt{5}}}{10-\sqrt{50-22\sqrt{5}}} \right) + \frac{5}{8} \sum_{k \geq 0} \frac{L_{2k+1}}{(2k+1)^2 \binom{2k}{k}} \, . \tag{23}$$

where the *Lucas numbers* L_n are given by $L_n = L_{n-1} + L_{n-2}$ with $L_0 = 2$, $L_1 = 1$.

The general formula for Bradley's family of series is proved using certain identities among log tangent integrals. For example, (23) is proved using

$$2 \int_0^{\pi/4} \log(\tan \theta) \, d\theta = 5 \int_0^{3\pi/20} \log(\tan \theta) \, d\theta - 5 \int_0^{\pi/20} \log(\tan \theta) \, d\theta \, .$$

This identity was discovered by an LLL integer relation algorithm. It turns out to be quite easy to search for such relations among log tangent integrals, whereas looking for resummations of the original series (by LLL) is quite difficult.

David Broadhurst has, in his pursuit of new insights for theoretical physics, computationally probed more of these constants [12]. Based on an extraordinary blend of intuition, methodical use of PSLQ, and computer-assisted proofs, he was led to remarkable binary identities for polylogarithmic constants such as $\zeta(3)$, $\zeta(5)$, and Catalan's constant. His formula for Catalan's constant is:

$$\begin{aligned} G \quad = \quad & \frac{3}{2} \sum_{i=0}^{\infty} \frac{1}{16^i} \left(\frac{1}{(8\,i+1)^2} - \frac{1}{(8\,i+2)^2} + \frac{1}{2} \frac{1}{(8\,i+3)^2} \right. \\ & \qquad \left. - \frac{1}{4} \frac{1}{(8\,i+5)^2} + \frac{1}{4} \frac{1}{(8\,i+6)^2} - \frac{1}{8} \frac{1}{(8\,i+7)^2} \right) \end{aligned}$$

$$-\frac{1}{4}\sum_{i=0}^{\infty}\frac{1}{16^i}\left(\frac{1}{(8\,i+1)^2}+\frac{1}{2}\,\frac{1}{(8\,i+2)^2}+\frac{1}{8}\,\frac{1}{(8\,i+3)^2}-\frac{1}{64}\,\frac{1}{(8\,i+5)^2}\right.$$
$$\left.-\frac{1}{128}\,\frac{1}{(8\,i+6)^2}-\frac{1}{512}\,\frac{1}{(8\,i+7)^2}\right)\,.$$

Thus, digits of both G and π may be computed in the same fashion, and we might hope that the formula sheds some light on the normality of Catalan's constant. [Recall that a number is 'normal' if its digits occur with equal frequency.]

8 Sin, Redemption and Cautionary Tales

Experimental mathematics cannot supplant rigorous mathematics. Dropping the latter for the former would indeed be a 'sin'. We have seen at least one example of a false computer generated conjecture—namely the egyptian fractions example in Section 1—and we could come up with many more [5]. Experimental mathematics is, however, a good supplement to rigorous mathematics. It can enrich our subject and, when used with discipline, can significantly assist mathematical discovery. We have also seen examples where the computer can assist with the proof.

As a final demonstration, consider the power series

$$J(x)=\sum_{n_1>n_2>0}\frac{x^{n_1}}{n_1^2 n_2}\,.$$

In [8], a functional relation was sought in pursuit of a proof of the identity $J(1)=8J(-1)$. For $0\le x\le 1$,

$$\begin{aligned}J(x) &= \int_0^x \frac{\ln^2(1-t)}{2t}\,dt \;=\; \zeta(3)+\frac{1}{2}\,\ln^2(1-x)\ln(x)\\ &+ \ln(1-x)\mathrm{polylog}(2,1-x)-\mathrm{polylog}(3,1-x)\,.\end{aligned}$$

It can be shown that

$$J(-x)=-J(x)+\frac{1}{4}J(x^2)+J\left(\frac{2x}{x+1}\right)-\frac{1}{8}J\left(\frac{4x}{(x+1)^2}\right)\,. \qquad (24)$$

Jon, eliminating this example, which I still find somewhat weak (especially when the point about proof has already been made), would also eliminate a reference and save some much-needed space.

This relation was found, once the ingredients were determined by inspection, by evaluating (24) (actually, a version of it with undetermined coefficients) at a random point and then using LLL. Another successful strategy is to evaluate each J function at enough specific values of x to enable one to solve linear equations for the unknown coefficients.

If $L(x)$ and $R(x)$ denote the left-hand and the right-hand sides of (24), respectively, then computer manipulations (under the assumption $0<x<1$) show that $dL/dx = dR/dx$: mechanically differentiating both sides and using "simplify" reduces the difference between the two to zero. Observing that $L(0)=R(0)=0$ completes a proof of (24).

8.1 Knowing 'the answer' might limit us

We are all familiar with examples of the value of 'doing things ourselves'. It is now trivial in most computer algebra systems (CAS) to compute very large values of the *partition function* with little or no thought, directly from the generating function

$$P(q) = \frac{1}{\prod_{n=1}^{\infty}(1-q^n)} .$$

The well-known exact finite series for values of the partition function, due to Radamacher [24], and its wonderful infinite, asymptotic precursor due to Ramanujan and Hardy, might well have seemed less worthy of discovery, had CAS been available then. We must be careful to ensure that our use of new tools neither limits us to what they can find for us nor supresses our interest in things easily computed.

This really will require attention: for example, the authors of [9] report in their conclusions that had they been aware of the answers in the Encyclopedia, they might not have bothered to prove what they did—and their results went beyond those in the Encyclopedia!

9 Concluding Remarks

It's not over yet. The merging of text and tools that can be anticipated over the next few years will make an enormous difference—we can expect greater insight while reading mathematical materials, and easier access to yet more powerful tools—but we make no detailed predictions, because the most significant, qualitative, changes to the work environment are by their nature unexpected.

Acknowledgements

This work was supported by NSERC. David Boyd, Mark Giesbrecht, David Jeffrey, and Glen Ord were all helpful, but the contributions of Petr Lisoněk were exceptional. Many good suggestions from the referees and editor, for which we are grateful, were used to improve the paper.

A The Lambert W function in brief

If you have used MAPLE to solve transcendental equations, you may already have encountered the Lambert W function, defined by (5). The history and some of the properties of this remarkable function are described in [14]. This function provides a beautiful new look at much of undergraduate mathematics, in addition to some new puzzles of intrinsic interest.

Here are some of the elementary properties of W.

1. On $0 \le x < \infty$ there is one real-valued branch $W(x) \ge 0$ (see Figure 1). On $-1/e < x < 0$ there are two real-valued branches. We call the branch

that has $W(0) = 0$ the principal branch. On this branch, it is easy to see that $W(e) = W(1 \cdot e^1) = 1$.

2. The derivative of W can be found by implicit differentiation to be

$$\begin{aligned} \frac{d}{dx} W(x) &= \frac{1}{(1 + W(x))e^{W(x)}} \\ &= \frac{W(x)}{(1 + W(x))x} \end{aligned}$$

where the second formula follows on using $\exp W(x) = x/W(x)$, and holds if $x \neq 0$. We may use the first formula to find the value of the derivative at $x = 0$, and we see the singularity is just a removable one.

3. The function $y = W(\exp z)$ satisfies

$$y + \log y = z \; .$$

This function appears, for example, in convex optimization. Consider the *convex conjugate*, $f^*(s) = \sup_r rs - f(r)$, of the function $f(r) = r \ln(r/(1-r)) - r$. Calculation shows that $f^*(s)$ is just $W(\exp s)$.

4. $W(x)$ has a Taylor series about $x = 0$ with rational coefficients. Similarly, $W(\exp z)$ has a Taylor series with rational coefficients about $z = 1$. MAPLE computes the first few terms to be

$$\begin{aligned} W(e^z) \;\; := \;\; & 1 + \frac{1}{2}(z-1) + \frac{1}{16}(z-1)^2 - \frac{1}{192}(z-1)^3 - \frac{1}{3072}(z-1)^4 \\ & + \frac{13}{61440}(z-1)^5 - \frac{47}{1474560}(z-1)^6 - \frac{73}{41287680}(z-1)^7 \\ & + \frac{2447}{1321205760}(z-1)^8 - \frac{16811}{47563407360}(z-1)^9 \\ & - \frac{15551}{1902536294400}(z-1)^{10} + O((z-1)^{11}) \end{aligned}$$

Here is an exact formula for the coefficients of the nth derivative of $W(\exp z)$, in terms of second-order Eulerian numbers $\left\langle\!\left\langle {n \atop k} \right\rangle\!\right\rangle$ [18]. This formula comes from the following expression for the nth derivative of $W(\exp z)$, which is stated in [14]. Once the answer is known, the proof is an easy induction, which we leave for the reader.

The derivatives of $W(\exp z)$ are

$$\frac{d^n}{dz^n} W(e^z) = \frac{q_n(W(e^z))}{(1 + W(e^z))^{2n-1}} \; , \tag{25}$$

where $q_n(w)$ is a polynomial of degree n satisfying the recurrence relation

$$q_{n+1}(w) = -(2n-1)wq_n(w) + (w + w^2)q_n'(w) \; , \qquad n > 1 \tag{26}$$

and having the explicit expression

$$q_n(w) = \sum_{k=0}^{n-1} \left\langle\!\!\left\langle {n-1 \atop k} \right\rangle\!\!\right\rangle (-1)^k w^{k+1} \; . \tag{27}$$

If $n = 1$ we have $q_1(w) = w$, and it is convenient to put $q_0(w) = w/(1+w)$; this isn't a polynomial, but it makes things work out right. This means that our series for $W(\exp z)$ about $z = 1$ is just

$$W(e^z) = \sum_{n \geq 0} \frac{q_n(1)}{n! 2^{2n-1}} (z-1)^n \; . \tag{28}$$

References

[1] G. Almkvist and A. Granville. Borwein and Bradley's Apéry-like formulae for $\zeta(4n+3)$. *Manuscript*, 1998.

[2] David Bailey, Peter Borwein, and Simon Plouffe. On the rapid computation of various polylogarithmic constants. *Math. Comp.*, 66:903–913, 1997.

[3] Jonathan M. Borwein and Peter B. Borwein. *Pi and the AGM*. John Wiley & Sons, New York, 1987.

[4] Jonathan M. Borwein and Peter B. Borwein. A cubic counterpart of Jacobi's identity and the AGM. *Trans. Amer. Math. Soc.*, 323:691–701, 1991.

[5] Jonathan M. Borwein and Peter B. Borwein. Strange series evaluations and high precision fraud. *American Mathematical Monthly*, 99:622–640, 1992.

[6] Jonathan M. Borwein, Peter B. Borwein, and Frank G. Garvan. Some cubic identities of Ramanujan. *Trans. Amer. Math. Soc.*, 343:35–47, 1994.

[7] Jonathan M. Borwein and David M. Bradley. Empirically determined Apéry-like formulae for zeta(4n+3). *Experimental Mathematics*, 6:181–194, 1997.

[8] Jonathan M. Borwein, David M. Bradley, David J. Broadhurst, and P. Lisoněk. Special values of multidimensional polylogarithms. Technical Report CECM Research Report 98-106, http:// www.cecm.sfu.ca/ preprints, July 1998.

[9] Jonathan M. Borwein and Kwok-Kwong Stephen Choi. On the representations of $xy + xz + yx$. Technical Report 98-119, http:// www.cecm.sfu.ca/ preprints, 1998.

[10] Jonathan M. Borwein and Petr Lisoněk. Applications of integer relation algorithms. *Discrete Mathematics, accepted August*, 1998.

[11] Peter Borwein and Loki Jörgenson. Visible structures in number theory. *http:// www.cecm.sfu.ca/ ~loki/ Papers/ Numbers/*, 1998.

[12] David J. Broadhurst. Polylogarithmic ladders, hypergeometric series and the ten millionth digits of $\zeta(3)$ and $\zeta(5)$. *preprint*, January 1998.

[13] Robert M. Corless. Error backward. In *Proceedings of Chaotic Numerics*, volume 172 of *AMS Contemporary Mathematics*, pages 31–62, 1994.

[14] Robert M. Corless, Gaston H. Gonnet, David E. G. Hare, David J. Jeffrey, and Donald E. Knuth. "On the Lambert W function". *Advances in Computational Mathematics*, 5:329–359, 1996.

[15] Robert M. Corless and David J. Jeffrey. Graphing elementary Riemann surfaces. *ACM Sigsam Bulletin: Communications in Computer Algebra*, 32(1):11–17, 1998.

[16] Robert M. Corless, David J. Jeffrey, and Donald E. Knuth. A sequence of series for the Lambert W function. In W. Küchlin, editor, *Proceedings of ISSAC '97, Maui*, pages 197–204, 1997.

[17] H. R. P. Ferguson, D. H. Bailey, and S. Arno. Analysis of PSLQ, an integer relation finding algorithm. Technical Report NAS Technical Report NAS-96-005, NASA Ames Research Center, Moffett Field, CA, April 1996.

[18] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics*. Addison-Wesley, 1994.

[19] D. F. Griffiths and J. M. Sanz-Serna. On the scope of the method of modified equations. *SIAM J. Sci. Stat. Comput.*, 7:994–1008, 1986.

[20] G. H. Hardy. *Ramanujan*. Chelsea, New York, 1940.

[21] Gilbert Labelle. "Sur l'inversion et l'itération continue des séries formelles". *European Journal of Combinatorics*, 1:113–138, 1980.

[22] A. K. Lenstra, H. W. Lenstra Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:515–534, 1982.

[23] George Marsaglia and John C. Marsaglia. "A new derivation of Stirling's approximation to $n!$". *American Mathematical Monthly*, 97:826–829, 1990.

[24] H. Radamacher. On the partition function $p(n)$. *Proceedings of the London Mathematical Society*, 43:241–254, 1937.

[25] Ed Regis. *Who Got Einstein's Office?* Addison-Wesley, 1986.

[26] Daniel Richardson. Some unsolvable problems involving elementary functions of a real variable. *Journal of Symbolic Logic*, 33:511–520, 1968.

[27] N. J. A. Sloane and Simon Plouffe. *The Encyclopedia of Integer Sequences*. Academic Press, San Diego, 1995.

[28] Michal Trott. Visualization of Riemann surfaces of algebraic functions. *Mathematica in Education and Research*, 6(4):15–36, 1997.

Jonathan M. Borwein,
Centre for Experimental and Constructive Mathematics, Simon Fraser University, Vancouver, B.C., Canada, V5A 1S6, jborwein@cecm.sfu.ca

Robert M. Corless,
Department of Applied Mathematics, University of Western Ontario, London, Ontario, Canada, N6A 5B7, Rob.Corless@uwo.ca