

Closed Forms: What They Are and Why We Care

Jonathan M. Borwein and Richard E. Crandall

Closed Forms: What They Are

Mathematics abounds in terms that are in frequent use yet are rarely made precise. Two such are *rigorous proof* and *closed form* (absent the technical use within differential algebra). If a rigorous proof is “that which ‘convinces’ the appropriate audience,” then a closed form is “that which looks ‘fundamental’ to the requisite consumer.” In both cases, this is a community-varying and epoch-dependent notion. What was a compelling proof in 1810 may well not be now; what is a fine closed form in 2010 may have been anathema a century ago. In this article we are intentionally informal as befits a topic that intrinsically has no one “right” answer.

Let us begin by sampling the Web for various approaches to informal definitions of “closed form”.

Definitions

First Approach to a Definition of Closed Form. The first comes from MathWorld [56] and so may well be the first and last definition a student or other seeker-after-easy-truth finds.

An equation is said to be a closed-form solution if it solves a given problem in terms of functions and mathematical operations from a given generally accepted set. For example, an infinite sum would generally

not be considered closed-form. However, the choice of what to call closed-form and what not is rather arbitrary since a new “closed-form” function could simply be defined in terms of the infinite sum.

—Eric Weisstein

There is not much to disagree with in this, but it is far from rigorous.

Second Approach. The next attempt follows a September 16, 1997, question to the long-operating “Dr. Math” site¹ and is a good model of what interested students are likely to be told.

Subject: Closed form solutions

Dear Dr. Math: What is the exact mathematical definition of a closed form solution? Is a solution in “closed form” simply if an expression relating all of the variables can be derived for a problem solution, as opposed to some higher-level problems where there is either no solution, or the problem can only be solved incrementally or numerically?

Sincerely,

The answer followed on September 22:

This is a very good question! This matter has been debated by mathematicians for some time, but without a good resolution.

Some formulas are agreed by all to be “in closed form”. Those are the ones which contain only a finite number of symbols, and include only the operators $+$, $-$, $*$, $/$,

Jonathan M. Borwein is professor of mathematics at the University of Newcastle. His email address is jonathan.borwein@newcastle.edu.au.

Richard E. Crandall is professor of Mathematics at Reed College. His email address is crandall@reed.edu.

DOI: <http://dx.doi.org/10.1090/noti936>

¹Available at <http://mathforum.org/dr/math/>.

and a small list of commonly occurring functions such as n th roots, exponentials, logarithms, trigonometric functions, inverse trigonometric functions, greatest integer functions, factorials, and the like.

More controversial would be formulas that include infinite summations or products, or more exotic functions, such as the Riemann zeta function, functions expressed as integrals of other functions that cannot be performed symbolically, functions that are solutions of differential equations (such as Bessel functions or hypergeometric functions), or some functions defined recursively. Some functions whose values are impossible to compute at some specific points would probably be agreed not to be in closed form (example: $f(x) = 0$ if x is an algebraic number, but $f(x) = 1$ if x is transcendental. For most numbers, we do not know if they are transcendental or not). I hope this is what you wanted.

No more formal but representative of many dictionary definitions is:

Third Approach. A coauthor of the current article is at least in part responsible for the following brief definition from a recent mathematics dictionary [17]:

closed form *n.* an expression for a given function or quantity, especially an integral, in terms of known and well understood quantities, such as the evaluation of

$$\int_{-\infty}^{\infty} \exp(-x^2) dx$$

as $\sqrt{\pi}$.

—Collins Dictionary

And of course one cares more for a closed form when the object under study is important, such as when it engages the normal distribution as above.

With that selection recorded, let us turn to some more formal proposals.

Fourth Approach. Various notions of elementary numbers have been proposed.

Definition [31]. A subfield F of \mathbb{C} is *closed under* \exp and \log if (1) $\exp(x) \in F$ for all $x \in F$ and (2) $\log(x) \in F$ for all nonzero $x \in F$, where \log is the branch of the natural logarithm function such that $-\pi < \text{Im}(\log x) \leq \pi$ for all x . The field \mathbb{E} of EL numbers is the intersection of all subfields of \mathbb{C} that are closed under \exp and \log .

—Tim Chow

Tim Chow explains nicely why he eschews capturing all algebraic numbers in his definition, why he wishes only elementary quantities to have closed forms, whence he prefers \mathbb{E} to Ritt's 1948 definition of *elementary numbers* as the smallest algebraically closed subfield \mathbb{L} of \mathbb{C} that is closed under \exp and \log . His reasons include that:

Intuitively, “closed-form” implies “explicit”, and most algebraic functions have no simple explicit expression.

Assuming *Shanuel's conjecture* that *given n complex numbers z_1, \dots, z_n which are linearly independent over the rational numbers, the extension field*

$$\mathbb{Q}(z_1, \dots, z_n, e^{z_1}, \dots, e^{z_n})$$

has transcendence degree of at least n over the rationals, then the algebraic members of \mathbb{E} are exactly those solvable in radicals [31]. We may thence think of Chow's class as the smallest plausible class of closed forms. Only a mad version of Markov would want to further circumscribe the class.

Special Functions

In an increasingly computational world, an explicit/implicit dichotomy is occasionally useful, but not very frequently. Often we will prefer computationally the numerical implicit value of an algebraic number to its explicit tower of radicals, and it seems increasingly perverse to distinguish the root of $2x^5 - 10x + 5$ from that of $2x^4 - 10x + 5$ or to prefer $\arctan(\pi/7)$ to $\arctan(1)$. We illustrate these issues further in Examples 7, 9, and 14.

We would prefer to view all values of classical *special functions* of mathematical physics [54] at algebraic arguments as being closed forms. Again there is no generally accepted class of special functions, but most practitioners would agree that the solutions to the classical second-order algebraic differential equations (linear or, say, Painlevé) are included. But so are various *hypertranscendental functions*, such as Γ , B , and ζ , which do not arise in that way.²

Hence, we do not wish to accept any definition of special function which relies on the underlying functions satisfying natural differential equations. The class must be extensible; new special functions are always being discovered.

A recent *American Mathematical Monthly* review³ of [47] says:

²Of course, a value of a hypertranscendental function at algebraic argument may be very well behaved; see Example 4.

³Available at <http://www.maa.org/maa%20reviews/4221.html>.

There's no rigorous definition of special functions, but the following definition is in line with the general consensus: functions that are commonly used in applications have many nice properties and are not typically available on a calculator. Obviously this is a sloppy definition, and yet it works fairly well in practice. Most people would agree, for example, that the Gamma function is included in the list of special functions and that trigonometric functions are not.

Once again, there is much to agree with and much to quibble about in this reprise. That said, most serious books on the topic are little more specific. For us, special functions are nonelementary functions about which a significant literature has developed because of their importance in either mathematical theory or in practice. We certainly prefer that this literature include the existence of excellent algorithms for their computation. This is all consonant with—if somewhat more ecumenical than—Temme's description in the preface of his lovely book [54, Preface, p. xi]:

[W]e call a function “special” when the function, just like the logarithm, the exponential and trigonometric functions (the elementary transcendental functions), belongs to the toolbox of the applied mathematician, the physicist or engineer.

Even more economically, Andrews, Askey, and Roy start the preface to their important book *Special Functions* [1] by writing:

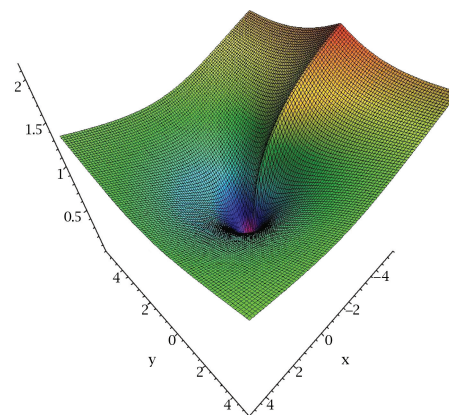
Paul Turan once remarked that special functions would be more appropriately labeled “useful functions”.

With little more ado, they then start to work on the Gamma and Beta functions; indeed, the term “special function” is not in their index. Near the end of their preface, they also write:

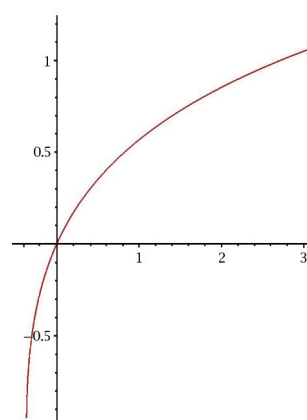
[W]e suggest that the day of formulas may be experiencing a new dawn.

Example 1 (Lambert's W). The *Lambert W* function, $W(x)$, is defined by appropriate solution of $y \cdot e^y = x$ [21, pp. 277–279]. This function has been implemented in computer algebra systems (CAS) and has many uses despite being unknown to most scientists and only relatively recently named [41]. It is now embedded as a primitive in Maple and Mathematica with the same status as any other well-studied special or elementary function. (See, for example, the tome [26].) The CAS know its power series and much more. For instance, in Maple entering:

```
> fsolve(exp(x)*x=1); identify(%);
returns
0.5671432904, LambertW(1)
```



(a) modulus of W



(b) W on real line

Figure 1. The Lambert W function.

We consider this to be a splendid closed form even though, again assuming Shanuel's conjecture, $W(1) \notin \mathbb{E}$ [31]. Additionally, it has only recently rigorously been proven that W is not an elementary function in Liouville's precise sense [26]. We also note that successful simplification in a modern CAS [29] requires a great deal of knowledge of special functions. \square

Further Approaches

Fifth Approach. PlanetMath's offering, as of February 15, 2010,⁴ is certainly in the elementary number corner.

expressible in closed form (Definition) An expression is expressible in a closed form if it can be converted (simplified) into an expression containing only elementary functions, combined by a finite amount of rational operations and compositions.

—PlanetMath

⁴Available at <http://planetmath.org/encyclopedia/ClosedForm4.html>.

This reflects both much of what is best and what is worst about “the mathematical wisdom of crowds”. For the reasons adduced above, we wish to distinguish—but admit both—those closed forms that give analytic insight from those which are sufficient and prerequisite for effective computation. Our own current preferred class [7] is described next.

Sixth Approach. We wish to establish a set \mathbb{X} of *generalized hypergeometric* evaluations; see [7] for an initial, rudimentary definition which we shall refine presently. First, we want any convergent sum

$$(1) \quad x = \sum_{n \geq 0} c_n z^n$$

to be an element of our set \mathbb{X} , where z is algebraic, c_0 is rational, and for $n > 0$,

$$c_n = \frac{A(n)}{B(n)} c_{n-1}$$

for integer polynomials A, B with $\deg A \leq \deg B$. Under these conditions the expansion for x converges absolutely on the open disk $|z| < 1$. However, we also allow x to be any finite analytic-continuation value of such a series. Moreover, when z lies on a branch cut, we presume both branch limits to be elements of \mathbb{X} . (See ensuing examples for some clarification.) It is important to note that our set \mathbb{X} is closed under rational multiplication due to freedom of choice for c_0 .

Example 2 (First members of \mathbb{X}). The *generalized hypergeometric function* evaluation

$${}_{p+1}F_p \left(\begin{matrix} a_1, \dots, a_{p+1} \\ b_1, \dots, b_p \end{matrix} \middle| z \right)$$

for rational a_i, b_j with all b_j positive has branch cut $z \in (0, \infty)$, and the evaluation is an element of \mathbb{X} for complex z not on the cut (and the evaluation on each side of said cut is also in \mathbb{X}).

The *trilogarithm* $\text{Li}_3(z) := \sum_{n \geq 1} z^n / n^3$ offers a canonical instance. Formally,

$$\frac{1}{z} \text{Li}_3(z) = {}_4F_3 \left(\begin{matrix} 1, 1, 1, 1 \\ 2, 2, 2 \end{matrix} \middle| z \right),$$

and for $z = 1/2$ the hypergeometric series converges absolutely, with

$$\text{Li}_3 \left(\frac{1}{2} \right) = \frac{7}{8} \zeta(3) + \frac{1}{6} \log^3 2 - \frac{\pi^2}{12} \log 2.$$

Continuation values at $z = 2$ on the branch cut can be inferred as

$$\lim_{\epsilon \rightarrow 0^+} \text{Li}_3(2 \pm i\epsilon) = \frac{7}{16} \zeta(3) + \frac{\pi^2}{8} \log 2 \pm i \frac{\pi}{4} \log 2,$$

so both complex numbers on the right here are elements of \mathbb{X} . The *quadralogarithmic* value $\text{Li}_4 \left(\frac{1}{2} \right)$ is thought not to be similarly decomposable but likewise belongs to \mathbb{X} . \square

Now we are prepared to posit

Definition [7]. The *ring of hyperclosure* \mathbb{H} is the smallest subring of \mathbb{C} containing the set \mathbb{X} . Elements of \mathbb{H} are deemed *hyperclosed*.

In other words, the ring \mathbb{H} is generated by all general hypergeometric evaluations under the $\cdot, +$ operators, all symbolized by

$$\mathbb{H} = \langle \mathbb{X} \rangle_{\cdot, +}.$$

\mathbb{H} will contain a great many interesting closed forms from modern research. Note that \mathbb{H} contains all closed forms in the sense of Wilf and Zeilberger [48, Ch. 8], wherein only finite linear combinations of hypergeometric evaluations are allowed.

So what numbers are in the ring \mathbb{H} ? First off, *almost no* complex numbers belong to this ring! This is easily seen by noting that the set of general hypergeometric evaluations is countable, so the generated ring must also be countable. Still, a great many fundamental numbers are provably hyperclosed. Examples follow, in which we let ω denote an arbitrary algebraic number and n any positive integer:

$$\omega, \log \omega, e^\omega, \pi;$$

the dilogarithmic combination

$$\text{Li}_2 \left(\frac{1}{\sqrt{5}} \right) + (\log 2)(\log 3);$$

the elliptic integral $K(\omega)$;

the zeta function values $\zeta(n)$;

special functions such as

the Bessel evaluations $J_n(\omega)$.

Incidentally, it occurs in some modern experimental developments that the real or imaginary part of a hypergeometric evaluation is under scrutiny. Generally, \Re, \Im operations preserve hyperclosure simply because the series (or continuations) at z and z^* can be linearly combined in the ring \mathbb{H} . Referring to Example 2, we see that for algebraic z , the number $\Re(\text{Li}_3(z))$ is hyperclosed, and even on the cut, $\Re(\text{Li}_3(2)) = \frac{7}{16} \zeta(3) + \frac{\pi^2}{8} \log 2$ is hyperclosed. In general, $\Re \left({}_{p+1}F_p \left(\begin{matrix} \dots \\ \dots \end{matrix} \middle| z \right) \right)$ is hyperclosed.

We are not claiming that hyperclosure is any kind of final definition for “closed forms”, but we do believe that any defining paradigm for closed forms must include this ring of hyperclosure \mathbb{H} . One way to reach further is to define a *ring of superclosure* as the closure

$$\mathbb{S} := \langle \mathbb{H}^{\mathbb{H}} \rangle_{\cdot, +}.$$

This ring contains numbers such as

$$e^\pi + \pi^e, \frac{1}{\zeta(3)\zeta(5)},$$

and of course a vast collection of numbers that may not belong to \mathbb{H} itself. If we say that an element of

\mathbb{S} is *superclosed*, we still preserve the countability of all superclosed numbers. Again, any good definition of “closed form” should incorporate whatever is in the ring \mathbb{S} .

Seventh Approach. In a more algebraic topological setting, it might make sense to define closed forms to be those arising as *periods*, that is, as integrals of rational functions (with integer parameters) in n variables over domains defined by algebraic equations. These ideas originate in the theory of elliptic and abelian integrals and are deeply studied [42]. Periods form an algebra and certainly capture many constants. They are especially well suited to the study of L-series, multizeta values, polylogarithms, and the like but again will not capture all that we wish. For example, e is conjectured not to be a period, as is Euler’s constant γ (see the section “Profound Curiosities”). Moreover, while many periods have nice series, it is not clear that all do.

As this takes us well outside our domain of expertise, we content ourselves with two examples originating in the study of Mahler measures. We refer to a fundamental paper by Deninger [39] and a very recent paper of Rogers [50] for details.

Example 3 (Periods and Mahler measures [39]). The logarithmic *Mahler measure* of a polynomial P in n -variables can be defined as

$$\mu(P) := \int_0^1 \int_0^1 \cdots \int_0^1 \log |P(e^{2\pi i \theta_1}, \dots, e^{2\pi i \theta_n})| \times d\theta_1 \cdots d\theta_n.$$

Then $\mu(P)$ turns out to be an example of a period, and its exponential, $M(P) := \exp(\mu(P))$, is a mean of the values of P on the unit n -torus. When $n = 1$ and P has integer coefficients, $M(P)$ is always an algebraic integer. An excellent online synopsis can be found in Dave Boyd’s article, <http://eom.springer.de/m/m120070.htm>. Indeed, Boyd has been one of the driving forces in the field. A brief introduction to the univariate case is also given in [22, pp. 358–359].

There is a remarkable series of recent results—many more discovered experimentally than proven—expressing various multidimensional $\mu(P)$ as arithmetic quantities. Boyd observes that there appears to be a tight connection to K -theory. An early result due to Smyth (see [51], also [53]) is that $\mu(1 + x + y) = L'_3(-1)$. Here L_3 is the Dirichlet L-series modulo three. A partner result of Smyth’s is that $\mu(1 + x + y + z) = 7\zeta(3)/\pi^2$, a number that is certainly hyperclosed since both $\zeta(3)$ and $1/\pi$ are. A conjecture of Deninger [39], confirmed to over fifty places, is that

$$(2) \quad \mu(1 + x + y + 1/x + 1/y) \stackrel{?}{=} \frac{15}{\pi^2} L_E(2)$$

is an L-series value over an elliptic curve E with conductor 15. Rogers [50] recasts (2) as

$$(3) \quad F(3, 5) \stackrel{?}{=} \frac{15}{\pi^2} \sum_{n=0}^{\infty} \binom{2n}{n}^2 \frac{(1/16)^{2n+1}}{2n+1},$$

where

$$F(b, c) := (1+b)(1+c) \times \sum_{n,m,j,k} \frac{(-1)^{n+m+j+k}}{((6n+1)^2 + b(6m+1)^2 + c(6j+1)^2 + bc(6k+1)^2)^2}$$

is a four-dimensional lattice sum.

While (3) remains a conjecture,⁵ Rogers is able to evaluate many values of $F(b, c)$ in terms of Meijer- G or hypergeometric functions. We shall consider the most famous crystal sum, the Madelung constant, in Example 15. \square

It is striking how beautiful combinatorial games can be when played under the rubric of hyper- or superclosure.

Example 4 (Superclosure of Γ at rational arguments). Let us begin with the *Beta function*

$$B(r, s) := \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)},$$

with $\Gamma(s)$ defined, if one wishes, as $\Gamma(s) := \int_0^\infty t^{s-1} e^{-t} dt$. It turns out that for any rationals r, s the Beta function is hyperclosed. This is immediate from the hypergeometric identities

$$\begin{aligned} \frac{1}{B(r, s)} &= \frac{rs}{r+s} {}_2F_1 \left(\begin{matrix} -r, -s \\ 1 \end{matrix} \middle| 1 \right), \\ B(r, s) &= \frac{\pi \sin \pi(r+s)}{\sin \pi r \sin \pi s} \frac{(1-r)_M (1-s)_M}{M! (1-r-s)_M} \\ &\quad \times {}_2F_1 \left(\begin{matrix} r, s \\ M+1 \end{matrix} \middle| 1 \right), \end{aligned}$$

where M is any integer chosen such that the hypergeometric series converges, say $M = \lceil 1+r+s \rceil$. (Each of these Beta relations is a variant of the celebrated Gauss evaluation of ${}_2F_1$ at 1 [1], [54] and is also the reason B is a period.)

We did not seize upon the Beta function arbitrarily, for, remarkably, the hyperclosure of $B(r, s)^{\pm 1}$ leads to compelling results on the Gamma function itself. Indeed, consider for example this product of four Beta-function evaluations:

$$\begin{aligned} &\frac{\Gamma(1/5)\Gamma(1/5)}{\Gamma(2/5)} \cdot \frac{\Gamma(2/5)\Gamma(1/5)}{\Gamma(3/5)} \\ &\quad \cdot \frac{\Gamma(3/5)\Gamma(1/5)}{\Gamma(4/5)} \cdot \frac{\Gamma(4/5)\Gamma(1/5)}{\Gamma(5/5)}. \end{aligned}$$

We know this product is hyperclosed. But upon inspection we see that the product is just $\Gamma^5(1/5)$.

⁵Equation (2) is now proven, see M. Rogers and W. Zudilin, “On the Mahler measure of $1 + X + 1/X + Y + 1/Y$ ”, preprint (2011), <http://arxiv.org/abs/1102.1153>.

Along such lines one can prove that for any positive rational a/b (in lowest terms), we have hyperclosure of powers of the Gamma function in the form

$$\Gamma^{\pm b}(a/b) \in \mathbb{H}.$$

Perforce, we therefore have a superclosure result for any $\Gamma(\text{rational})$ and its reciprocal:

$$\Gamma^{\pm 1}(a/b) \in \mathbb{S}.$$

Again, like calculations show that $\Gamma^b(a/b)$ is a period [42]. One fundamental consequence is thus $\Gamma^{-2}(\frac{1}{2}) = \frac{1}{\pi}$ is hyperclosed; thus every integer power of π is hyperclosed.

Incidentally, deeper combinatorial analysis shows that in spite of our $\Gamma^5(\frac{1}{5})$ Beta-chain above, it really takes only *logarithmically many* (i.e., $O(\log b)$) hypergeometric evaluations to write Gamma-powers. For example,

$$\Gamma^{-7}\left(\frac{1}{7}\right) = \frac{1}{2^3 7^6} {}_2F_1\left(\begin{matrix} -\frac{1}{7}, -\frac{1}{7} \\ 1 \end{matrix} \middle| 1\right)^4 \\ \times {}_2F_1\left(\begin{matrix} -\frac{2}{7}, -\frac{2}{7} \\ 1 \end{matrix} \middle| 1\right)^2 {}_2F_1\left(\begin{matrix} -\frac{4}{7}, -\frac{4}{7} \\ 1 \end{matrix} \middle| 1\right).$$

We note also that for $\Gamma(n/24)$ with n integer, elliptic integral algorithms are known that converge as fast as those for π [27], [22]. \square

The above remarks on superclosure of $\Gamma(a/b)$ lead to the property of superclosure for special functions such as $J_\nu(\omega)$ for algebraic ω and rational ν and for many of the mighty Meijer- G functions, as the latter can frequently be written by Slater's theorem [15] as superpositions of hypergeometric evaluations with composite-gamma products as coefficients. (See Example 8 for instances of Meijer- G in current research.)

There is an interesting alternative way to envision hyperclosure or at least something very close to our above definition. This is an idea of J. Carette [28] to the effect that solutions at algebraic end-points and algebraic initial points for *holonomic* ODEs—i.e., differential-equation systems with integer-polynomial coefficients—could be considered closed. One might say *diffeoclosed*. An example of a diffeoclosed number is $J_1(1)$, i.e., from the Bessel differential equation for $J_1(z)$ with $z \in [0, 1]$; it suffices without loss of generality to consider topologically clean trajectories of the variable over $[0, 1]$. There is a formal ring of diffeoclosure, which ring is very similar to our \mathbb{H} ; however, there is the caution that trajectory solutions can sometimes have nontrivial topology, so precise ring definitions would need to be effected carefully.

It is natural to ask, “what is the complexity of hypergeometric evaluations?” Certainly for the converging forms with variable z on the open unit

disk, convergence is geometric, requiring $O(D^{1+\epsilon})$ operations to achieve D good digits. However, in very many cases this can be genuinely enhanced to $O(D^{1/2+\epsilon})$ [22].

Closed Forms: Why We Care

In many optimization problems, *simple, approximate* solutions are more useful than complex exact solutions.

—Steve Wright

As Steve Wright observed in a recent lecture on *sparse optimization*, it may well be that a complicated analytic solution is practically intractable, but a simplifying assumption leads to a very practical closed form approximation (e.g., in compressed sensing). In addition to appealing to Occam's razor, Wright instances that:

- (a) the data quality may not justify exactness,
- (b) the simple solution may be more robust,
- (c) it may be easier to *explain/actuate/implement/store*,
- (d) and it may conform better to prior knowledge.

As mathematical discovery more and more involves extensive computation, the premium on having a closed form increases. The insight provided by discovering a closed form ideally comes at the top of the list, but efficiency of computation will run a good second.

Example 5 (The amplitude of a pendulum). Wikipedia,⁶ after giving the classical small angle (simple harmonic) approximation

$$p \approx 2\pi\sqrt{\frac{L}{g}}$$

for the period p of a pendulum of length L and amplitude α , develops the exact solution in a form equivalent to

$$p = 4\sqrt{\frac{L}{g}} K\left(\sin \frac{\alpha}{2}\right)$$

and then says:

This integral cannot be evaluated in terms of elementary functions. It can be rewritten in the form of the elliptic function of the first kind (also see Jacobi's elliptic functions), which gives little advantage since that form is also insoluble.

True, an elliptic-integral solution is not elementary, yet the notion of insolubility is misleading for two reasons: First, it is known that for some special angles α the pendulum period can be given a closed form. As discussed in [33], one exact solution is, for

⁶Available at [http://en.wikipedia.org/wiki/Pendulum_\(mathematics\)](http://en.wikipedia.org/wiki/Pendulum_(mathematics)).

$\alpha = \pi/2$ (so pendulum is released from horizontal-rod position),

$$p = \left(2\pi\sqrt{\frac{L}{g}}\right) \frac{\sqrt{\pi}}{\Gamma^2(3/4)}.$$

It is readily measurable in even a rudimentary laboratory that the excess factor here, $\sqrt{\pi}\Gamma^{-2}(3/4) \approx 1.18034$, looks just right; i.e., a horizontal-release pendulum takes 18 percent longer to fall. Moreover, there is an *exact* dynamical solution for the time-dependent angle $\alpha(t)$, namely, for a pendulum with $\alpha(\pm\infty) = \pm\pi$ and $\alpha(0) = 0$; i.e., the bob crosses angle zero (hanging straight down) at time zero, but in the limits of time $\rightarrow \pm\infty$ the bob ends up straight vertical. We have period $p = \infty$, yet the *exact* angle $\alpha(t)$ for given t can be written in terms of elementary functions!

The second misleading aspect is this: K is, for any α , remarkably tractable in a computational sense. Indeed K admits a quadratic transformation

$$(4) \quad K(k) = (1 + k_1) K(k_1), \quad k_1 := \frac{1 - \sqrt{1 - k^2}}{1 + \sqrt{1 - k^2}},$$

as was known already to Landen, Legendre, and Gauss.

In fact all elementary functions to very high precision are well computed via K [22]. So the comment was roughly accurate in the world of slide rules or pocket calculators; it is misleading today if one has access to any computer package. Nevertheless, both deserve to be called closed forms: one exact and the other an elegant approximate closed form (excellent in its domain of applicability, much as with Newtonian mechanics) that is equivalent to

$$K\left(\sin \frac{\alpha}{2}\right) \approx \frac{\pi}{2}$$

for small initial amplitude α . To compute $K(\pi/6) = 1.699075885\dots$ to five places requires using (4) only twice and then estimating the resultant integral by $\pi/2$. A third step gives the ten-digit precision shown. \square

It is now the case that much mathematical computation is *hybrid*: mixing numeric and symbolic computation. Indeed, which is which may not be clear to the user if, say, numeric techniques have been used to return a symbolic answer or if a symbolic closed form has been used to make possible a numerical integration. Moving from classical to modern physics, both understanding and effectiveness frequently demand hybrid computation.

Example 6 (Scattering amplitudes [2]). An international team of physicists, in preparation for the Large Hadron Collider (LHC), is computing scattering amplitudes involving quarks, gluons, and gauge vector bosons in order to predict what results could be expected on the LHC. By default,

these computations are performed using conventional double precision (64-bit IEEE) arithmetic. Then if a particular phase space point is deemed numerically unstable, it is recomputed with double-double precision. These researchers expect that further optimization of the procedure for identifying unstable points may be required to arrive at an optimal compromise between numerical accuracy and speed of the code. Thus they plan to incorporate arbitrary precision arithmetic into these calculations. Their objective is to design a procedure where instead of using fixed double or quadruple precision for unstable points, the number of digits in the higher precision calculation is dynamically set according to the instability of the point. Any subroutine which uses a closed form symbolic solution (exact or approximate) is likely to prove much more robust and efficient. \square

Detailed Examples

We start with three examples originating in [16].

In the January 2002 issue of *SIAM News*, Nick Trefethen presented ten diverse problems used in teaching *modern* graduate numerical analysis students at Oxford University, the answer to each being a certain real number. Readers were challenged to compute ten digits of each answer, with a \$100 prize to the best entrant. Trefethen wrote,

If anyone gets 50 digits in total, I will be impressed.

To his surprise, a total of ninety-four teams, representing twenty-five different nations, submitted results. Twenty of these teams received a full one hundred points (ten correct digits for each problem). The problems and solutions are dissected most entertainingly in [16]. One of the current authors wrote the following in a review [19] of [16]:

Success in solving these problems required a broad knowledge of mathematics and numerical analysis, together with significant computational effort, to obtain solutions and ensure correctness of the results. As described in [16] the strengths and limitations of *Maple*, *Mathematica*, *MATLAB* (The *3Ms*), and other software tools such as *PARI* or *GAP*, were strikingly revealed in these ventures. Almost all of the solvers relied in large part on one or more of these three packages, and while most solvers attempted to confirm their results, there was no explicit requirement for proofs to be provided.

Example 7 (Trefethen problem #2 [16], [19]).

A photon moving at speed 1 in the x - y plane starts at $t = 0$ at $(x, y) = (1/2, 1/10)$ heading due east. Around every integer lattice point (i, j) in the plane, a circular mirror of radius $1/3$ has been erected. How far from the origin is the photon at $t = 10$?

Using *interval arithmetic* with starting intervals of size smaller than 10^{-5000} , one can actually find the position of the particle at time 2000, not just at time ten. This makes a fine exercise in very high-precision interval computation, but in the absence of any closed form, one is driven to such numerical gymnastics to deal with error propagation. \square

Example 8 (Trefethen's problem #9 [16], [19]).

The integral $I(a) = \int_0^2 [2 + \sin(10\alpha)]x^\alpha \times \sin(\alpha/(2-x)) dx$ depends on the parameter α . What is the value $\alpha \in [0, 5]$ at which $I(\alpha)$ achieves its maximum?

The maximum parameter is expressible in terms of a *Meijer-G function*, which is a special function with a solid history. While knowledge of this function was not common among the contestants, Mathematica and Maple both will figure this out [15], and then the help files or a Web search will quickly inform the scientist.

This is another measure of the changing environment. It is usually a good idea—and not at all immoral—to data-mine. These Meijer-G functions, first introduced in 1936, also occur in quantum field theory and many other places [8]. For example, the moments of an n -step random walk in the plane are given for $s > 0$ by

$$(5) \quad W_n(s) := \int_{[0,1]^n} \left| \sum_{k=1}^n e^{2\pi x_k i} \right|^s dx.$$

It transpires [24], [36] that for all complex s

$$(6) \quad W_3(s) = \frac{\Gamma(1+s/2)}{\Gamma(1/2)\Gamma(-s/2)} G_{3,3}^{2,1} \left(\begin{matrix} 1, 1, 1 \\ \frac{1}{2}, -\frac{s}{2}, -\frac{s}{2} \end{matrix} \middle| \frac{1}{4} \right).$$

Moreover, for s not an odd integer, we have

$$W_3(s) = \frac{1}{2^{2s+1}} \tan\left(\frac{\pi s}{2}\right) \left(\frac{s}{2}\right)^2 {}_3F_2 \left(\begin{matrix} \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \\ \frac{s+3}{2}, \frac{s+3}{2} \end{matrix} \middle| \frac{1}{4} \right) + \left(\frac{s}{2}\right) {}_3F_2 \left(\begin{matrix} -\frac{s}{2}, -\frac{s}{2}, -\frac{s}{2} \\ 1, -\frac{s-1}{2} \end{matrix} \middle| \frac{1}{4} \right).$$

We have not given the somewhat technical definition of MeijerG, but Maple, Mathematica, Google searches, Wikipedia, the DLMF, or many other tools will.

There are two corresponding formulae for W_4 . We thus know from our “Sixth Approach” section, in regard to superclosure of Γ -evaluations, that both $W_3(q)$, $W_4(q)$ are superclosed for rational argument q for q not an odd integer. We illustrate by showing graphs of W_3, W_4 on the real line in

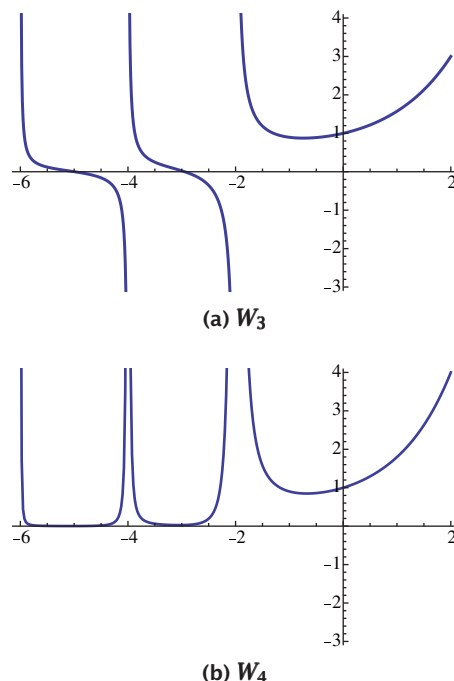


Figure 2. Moments of n -step walks in the plane. W_3, W_4 analytically continued to the real line.

Figure 2 and in the complex plane in Figure 3. The latter highlights the utility of the Meijer-G representations. Note the poles and removable singularities.

The Meijer-G functions are now described in the newly completed *Digital Library of Mathematical Functions*⁷ and as such are now full, indeed central, members of the family of special functions. \square

Example 9 (Trefethen's problem #10 [16], [19]).

A particle at the center of a 10×1 rectangle undergoes Brownian motion (i.e., 2-D random walk with infinitesimal step lengths) till it hits the boundary. What is the probability that it hits at one of the ends rather than at one of the sides?

Hitting the Ends. Bornemann [16] starts his remarkable solution by exploring *Monte-Carlo methods*, which are shown to be impracticable. He then reformulates the problem *deterministically* as the value at the center of a 10×1 rectangle of an appropriate harmonic measure [57] of the ends, arising from a five-point discretization of Laplace's equation with Dirichlet boundary conditions. This is then solved by a well-chosen *sparse Cholesky*

⁷A massive revision of Abramowitz and Stegun, with the now redundant tables removed, is available at <http://dlmf.nist.gov>. The hard copy version is also now out [45]. It is not entirely a substitute for the original version, as coverage has changed.

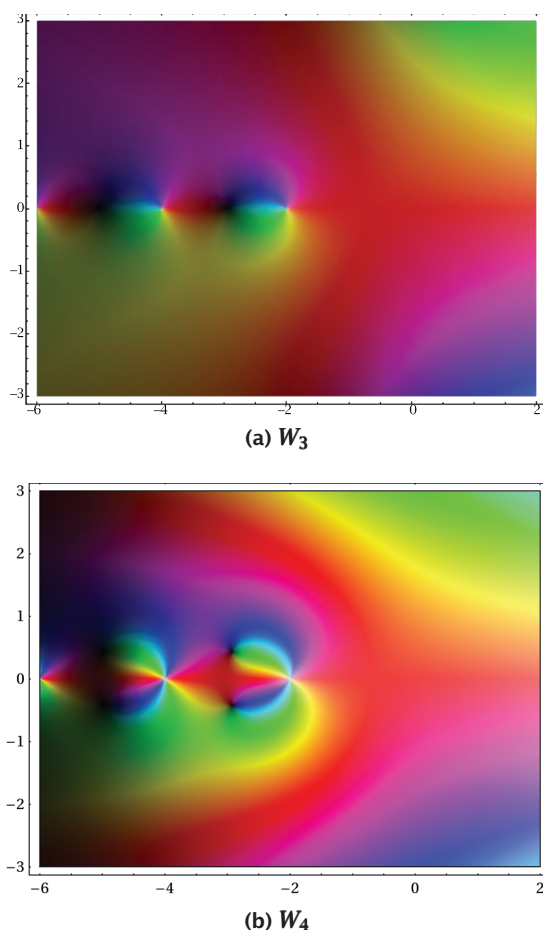


Figure 3. W_3 via (6) and W_4 in the complex plane.

solver. At this point a reliable numerical value of $3.837587979 \cdot 10^{-7}$ is obtained, and the posed problem is solved *numerically* to the requisite ten places.

This is the warm-up. We may proceed to develop two analytic solutions, the first using *separation of variables* on the underlying PDE on a general $2a \times 2b$ rectangle. We learn that

$$(7) \quad p(a, b) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \operatorname{sech} \left(\frac{\pi(2n+1)}{2} \rho \right),$$

where $\rho := a/b$. A second method using *conformal mappings* yields

$$(8) \quad \operatorname{arccot} \rho = p(a, b) \frac{\pi}{2} + \arg K \left(e^{ip(a, b)\pi} \right),$$

where K is again the *complete elliptic integral* of the first kind. It will not be apparent to a reader unfamiliar with inversion of elliptic integrals that (7) and (8) encode the same solution—though they must, as the solution is unique in $(0, 1)$ —and each can now be used to solve for $\rho = 10$ to arbitrary precision. Bornemann ultimately shows that the

answer is

$$(9) \quad p = \frac{2}{\pi} \arcsin(k_{100}),$$

where

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

No one (except harmonic analysts perhaps) anticipated a closed form, let alone one like this.

Where Does This Come From? In fact, [22, (3.2.29)] shows that

$$(10) \quad \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \operatorname{sech} \left(\frac{\pi(2n+1)}{2} \rho \right) = \frac{1}{2} \arcsin k$$

exactly when k_{ρ^2} is parameterized by *theta functions* in terms of the so-called *nome*, $q = \exp(-\pi\rho)$, as Jacobi discovered. We have

$$(11) \quad k_{\rho^2} = \frac{\theta_2^2(q)}{\theta_3^2(q)} = \frac{\sum_{n=-\infty}^{\infty} q^{(n+1/2)^2}}{\sum_{n=-\infty}^{\infty} q^{n^2}}, \quad q := e^{-\pi\rho}.$$

Comparing (10) and (7) we see that the solution is $k_{100} = 6.02806910155971082882540712292 \dots \cdot 10^{-7}$, as asserted in (9).

The explicit form now follows from classical nineteenth-century theory, as discussed, say, in [16], [22]. In fact, k_{210} is the singular value sent by Ramanujan to Hardy in his famous letter of introduction [21], [22]. If Trefethen had asked for a $\sqrt{210} \times 1$ box or, even better, a $\sqrt{15} \times \sqrt{14}$ one, this would have shown up in the answer, since in general

$$(12) \quad p(a, b) = \frac{2}{\pi} \arcsin(k_{a^2/b^2}).$$

Alternatively, armed only with the knowledge that the singular values of rational parameters are always algebraic, we may finish entirely computationally as described in [19]. \square

We finish this section with two attractive applied examples from optics and astrophysics respectively.

Example 10 (Mirages [46]). In [46] the authors, using geometric methods, develop an exact but implicit formula for the path followed by a light ray propagating over the earth with radial variations in the refractive index. By suitably simplifying, they are able to provide an explicit integral closed form. They then expand it asymptotically. This is done with the knowledge that the approximation is good to six or seven places, more than enough to use it on optically realistic scales. Moreover, in the case of quadratic or linear refractive indices, these steps may be done *analytically*.

In other words, as advanced by Wright, a tractable and elegant approximate closed form is



(a) A superior mirage



(b) An inferior mirage (Photo © Ctein.)

Figure 4. Two impressive mirages.

obtained to replace a problematic exact solution. From these forms interesting qualitative consequences follow. With a quadratic index, images are uniformly magnified in the vertical direction; only with higher-order indices can nonuniform vertical distortion occur. This sort of knowledge allows one, for example, to correct distortions of photographic images as in Figure 4 with confidence and efficiency. \square

Example 11 (Structure of stars). The celebrated *Lane-Emden equation*, presumed to describe the pressure χ at radius r within a star, can be put in the form

$$(13) \quad r^{n-1} \frac{d^2 \chi}{dr^2} = -\chi^n,$$

with boundary conditions $\chi(0) = 0$, $\chi'(0) = 1$, and positive real constant n , all of this giving rise to a unique trajectory $\chi_n(r)$ on $r \in [0, \infty)$. (Some authors invoke the substitution $\chi(r) := r\theta(r)$ to get an equivalent ODE for temperature θ ; see [30].) The beautiful thing is, where this pressure trajectory crosses zero for positive radius r is supposed to be the *star radius*; call that zero z_n .

Amazingly, the Lane-Emden equation has known exact solutions for $n = 0, 1, 5$, the pressure

trajectories for which indices n being respectively

$$(14) \quad \chi_0(r) = -\frac{1}{6}r^3 + r,$$

$$(15) \quad \chi_1(r) = \sin r,$$

$$(16) \quad \chi_5(r) = \frac{r}{\sqrt{1 + r^2/3}}.$$

The respective star radii are thus closed forms $z_0 = \sqrt{6}$ and $z_1 = \pi$, while for (16), with index $n = 5$ we have infinite star radius (no positive zero for the pressure χ_5).

In the spirit of our previous optics example, the Lane-Emden equation is a simplification of a complicated underlying theory—in this astrophysics case, hydrodynamics—and one is rewarded by some closed-form star radii. But what about, say, index $n = 2$? We do not know a closed-form function for the χ trajectory in any convenient sense. What the present authors *have* calculated (in 2005) is the $n = 2$ star radius, as a high-precision number

$$z_2 = 4.352874595946124676973570061526142628112365363213008835302151 \dots$$

If only we could gain a closed form for this special radius, we might be able to guess the nature of the whole trajectory! \square

Recent Examples Relating to Our Own Work

Example 12 (Ising integrals [5], [8]). We recently studied the following classes of integrals [5]. The D_n integrals arise in the Ising model of mathematical physics (showing ferromagnetic temperature-driven phase shifts; see Figure 5 and [32]), and the C_n have tight connections to quantum field theory [8]:

$$C_n = \frac{4}{n!} \int_0^\infty \cdots \int_0^\infty \frac{1}{\left(\sum_{j=1}^n (u_j + 1/u_j)\right)^2} \frac{du_1}{u_1} \cdots \frac{du_n}{u_n},$$

$$D_n = \frac{4}{n!} \int_0^\infty \cdots \int_0^\infty \frac{\prod_{i < j} \left(\frac{u_i - u_j}{u_i + u_j}\right)^2}{\left(\sum_{j=1}^n (u_j + 1/u_j)\right)^2} \frac{du_1}{u_1} \cdots \frac{du_n}{u_n},$$

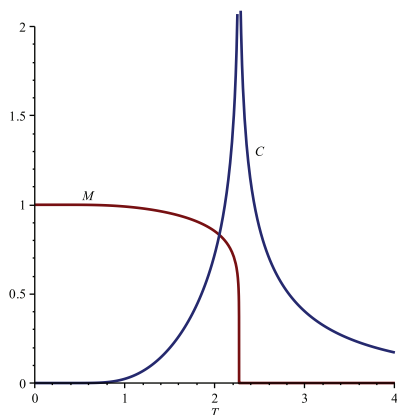
$$E_n = 2 \int_0^1 \cdots \int_0^1 \left(\prod_{1 \leq j < k \leq n} \frac{u_k - u_j}{u_k + u_j} \right)^2 dt_2 dt_3 \cdots dt_n,$$

where (in the last line) $u_k = \prod_{i=1}^k t_i$.

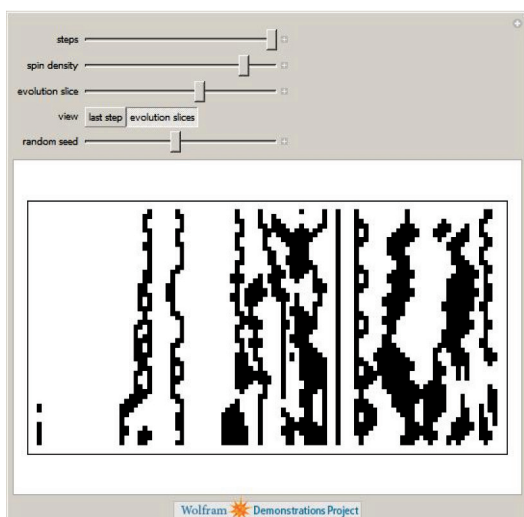
Needless to say, evaluating these multidimensional integrals to high precision presents a daunting computational challenge. Fortunately, in the first case, the C_n integrals can be written as one-dimensional integrals:

$$C_n = \frac{2^n}{n!} \int_0^\infty p K_0^n(p) dp,$$

where K_0 is the *modified Bessel function*. After computing C_n to 1000-digit accuracy for various n , we were able to identify the first few instances of C_n



(a) Critical temperature



(b) Wolfram Player Demonstration

Figure 5. The 2-dimensional Ising Model of Ferromagnetism (a) plotting magnetization C (blue, with peak) and specific heat M (red, decaying) per site against absolute temperature T (image provided by Jacques Perk) [44, pp. 91–93, 245].

in terms of well-known constants, e.g.,

$$C_3 = L_{-3}(2) := \sum_{n \geq 0} \left(\frac{1}{(3n+1)^2} - \frac{1}{(3n+2)^2} \right),$$

$$C_4 = \frac{7}{12} \zeta(3),$$

where ζ denotes the Riemann zeta function. When we computed C_n for fairly large n , for instance,

$$C_{1024} = 0.63047350337438679612204019271087890435458707871273234 \dots,$$

we found that these values rather quickly approached a limit. By using the new edition of the Inverse Symbolic Calculator,⁸ we identified this

⁸Available at <http://isc2.carma.newcastle.edu.au/>.

numerical value as

$$\lim_{n \rightarrow \infty} C_n = 2e^{-2\gamma},$$

where γ is the *Euler constant*; see the section “Profound Curiosities”. We later were able to prove this fact—this is merely the first term of an asymptotic expansion—and thus showed that the C_n integrals are fundamental in this context [5].

The integrals D_n and E_n are much more difficult to evaluate, since they are not reducible to one-dimensional integrals (as far as we can tell); but with certain symmetry transformations and symbolic integration we were able to symbolically reduce the dimension in each case by one or two.

In the case of D_5 and E_5 , the resulting 3-D integrands are extremely complicated (see Figure 6), but we were nonetheless able to numerically evaluate these to at least 240-digit precision on a highly parallel computer system. This would have been impossible without the symbolic reduction. We give the integral in extenso to show the difference between a humanly accessible answer and one a computer finds useful.

In this way, we produced the following evaluations, all of which, except the last, we subsequently were able to prove:

$$D_2 = 1/3,$$

$$D_3 = 8 + 4\pi^2/3 - 27L_{-3}(2),$$

$$D_4 = 4\pi^2/9 - 1/6 - 7\zeta(3)/2,$$

$$E_2 = 6 - 8 \log 2,$$

$$E_3 = 10 - 2\pi^2 - 8 \log 2 + 32 \log^2 2,$$

$$E_4 = 22 - 82\zeta(3) - 24 \log 2 + 176 \log^2 2,$$

$$- 256(\log^3 2)/3 + 16\pi^2 \log 2 - 22\pi^2/3,$$

and

(17)

$$E_5 \stackrel{?}{=} 42 - 1984 \text{Li}_4(1/2) + 189\pi^4/10 - 74\zeta(3)$$

$$- 1272\zeta(3) \log 2 + 40\pi^2 \log^2 2$$

$$- 62\pi^2/3 + 40(\pi^2 \log 2)/3 + 88 \log^4 2$$

$$+ 464 \log^2 2 - 40 \log 2,$$

where Li denotes the polylogarithm function.

In the case of D_2 , D_3 , and D_4 , these are confirmations of known results. We tried but failed to recognize D_5 in terms of similar constants (the 500-digit numerical value is accessible⁹ if anyone wishes to try to find a closed form or, in the manner of the hard sciences, to confirm our data values). The conjectured identity shown here for E_5 was confirmed to 240-digit accuracy, which is 180 digits beyond the level that could reasonably be ascribed to numerical round-off error. Thus we are quite

⁹Available at <http://crd.lbl.gov/~dhbailey/dhbpapers/ising-data.pdf>.

$$\begin{aligned}
E_5 = & \int_0^1 \int_0^1 \int_0^1 [2(1-x)^2(1-y)^2(1-xy)^2(1-z)^2(1-yz)^2(1-xyz)^2 \\
& (-[4(x+1)(xy+1)\log(2)(y^5z^3x^7 - y^4z^2(4(y+1)z+3)x^6 - y^3z((y^2+1)z^2 \\
& + 4(y+1)z+5)x^5 + y^2(4y(y+1)z^3 + 3(y^2+1)z^2 + 4(y+1)z-1)x^4 \\
& + y(z(z^2+4z+5)y^2 + 4(z^2+1)y+5z+4)x^3 + ((-3z^2-4z+1)y^2 \\
& - 4zy+1)x^2 - (y(5z+4)+4)x-1)] / [(x-1)^3(xy-1)^3(xyz-1)^3] \\
& + [3(y-1)^2y^4(z-1)^2z^2(yz-1)^2x^6 + 2y^3z(3(z-1)^2z^3y^5 + z^2(5z^3+3z^2 \\
& + 3z+5)y^4 + (z-1)^2z(5z^2+16z+5)y^3 + (3z^5+3z^4-22z^3-22z^2 \\
& + 3z+3)y^2 + 3(-2z^4+z^3+2z^2+z-2)y+3z^3+5z^2+5z+3)x^5 \\
& + y^2(7(z-1)^2z^4y^6 - 2z^3(z^3+15z^2+15z+1)y^5 + 2z^2(-21z^4+6z^3 \\
& + 14z^2+6z-21)y^4 - 2z(z^5-6z^4-27z^3-27z^2-6z+1)y^3 + (7z^6 \\
& - 30z^5+28z^4+54z^3+28z^2-30z+7)y^2 - 2(7z^5+15z^4-6z^3-6z^2+15z \\
& + 7)y+7z^4-2z^3-4z^2-2z+7)x^4 - 2y(z^3(z^3-9z^2-9z+1)y^6 \\
& + z^2(7z^4-14z^3-18z^2-14z+7)y^5 + z(7z^5+14z^4+3z^3+3z^2+14z \\
& + 7)y^4 + (z^6-14z^5+3z^4+84z^3+3z^2-14z+1)y^3 - 3(3z^5+6z^4-z^3 \\
& - z^2+6z+3)y^2 - (9z^4+14z^3-14z^2+14z+9)y+z^3+7z^2+7z+1)x^3 \\
& + (z^2(11z^4+6z^3-66z^2+6z+11)y^6 + 2z(5z^5+13z^4-2z^3-2z^2+13z \\
& + 5)y^5 + (11z^6+26z^5+44z^4-66z^3+44z^2+26z+11)y^4 + (6z^5-4z^4 \\
& - 66z^3-66z^2-4z+6)y^3 - 2(33z^4+2z^3-22z^2+2z+33)y^2 + (6z^3 \\
& + 26z+6)y+11z^2+10z+11)x^2 - 2(z^2(5z^3+3z^2+3z+5)y^5 \\
& + z(22z^4+5z^3-22z^2+5z+22)y^4 + (5z^5+5z^4-26z^3-26z^2+5z+5)y^3 \\
& + (3z^4-22z^3-26z^2-22z+3)y^2 + (3z^3+5z^2+5z+3)y+5z^2+22z+5)x \\
& + 15z^2+2z+2y(z-1)^2(z+1)+2y^3(z-1)^2z(z+1)+y^4z^2(15z^2+2z+15) \\
& + y^2(15z^4-2z^3-90z^2-2z+15)+15] / [(x-1)^2(y-1)^2(xy-1)^2(z-1)^2 \\
& (yz-1)^2(xyz-1)^2] - [4(x+1)(y+1)(yz+1)(-z^2y^4+4z(z+1)y^3 \\
& + (z^2+1)y^2-4(z+1)y+4x(y^2-1)(y^2z^2-1)+x^2(z^2y^4-4z(z+1)y^3 \\
& - (z^2+1)y^2+4(z+1)y+1)-1)\log(x+1)] / [(x-1)^3x(y-1)^3(yz-1)^3] \\
& - [4(y+1)(xy+1)(z+1)(x^2(z^2-4z-1)y^4+4x(x+1)(z^2-1)y^3 \\
& - (x^2+1)(z^2-4z-1)y^2-4(x+1)(z^2-1)y+z^2-4z-1)\log(xy+1)] / \\
& [x(y-1)^3y(xy-1)^3(z-1)^3] - [4(z+1)(yz+1)(x^3y^5z^7+x^2y^4(4x(y+1) \\
& + 5)z^6-xy^3((y^2+1)x^2-4(y+1)x-3)z^5-y^2(4y(y+1)x^3+5(y^2+1)x^2 \\
& + 4(y+1)x+1)z^4+y(y^2x^3-4y(y+1)x^2-3(y^2+1)x-4(y+1))z^3 \\
& + (5x^2y^2+y^2+4x(y+1)y+1)z^2+((3x+4)y+4)z-1)\log(xyz+1)] / \\
& [xy(z-1)^3z(yz-1)^3(xyz-1)^3]] \\
& / [(x+1)^2(y+1)^2(xy+1)^2(z+1)^2(yz+1)^2(xyz+1)^2] dx dy dz.
\end{aligned}$$

Figure 6. The reduced multidimensional integral for E_5 , which integral has led via extreme-precision numerical quadrature and PSLQ to the conjectured closed form given in (17).

confident in this result, even though we do not have a formal proof [5].

Note that every one of the D, E forms above, including the conjectured last one, is hyperclosed in the sense of our “Sixth Approach” section. \square

Example 13 (Weakly coupling oscillators [49], [6]). In an important analysis of coupled *Winfree oscillators*, Quinn, Rand, and Strogatz [49] developed a certain N -oscillator scenario whose bifurcation phase offset ϕ is implicitly defined, with a conjectured asymptotic behavior: $\sin \phi \sim 1 - c_1/N$,

with experimental estimate $c_1 = 0.605443657\dots$. In [6] we were able to derive the exact theoretical value of this “QRS constant” c_1 as the unique zero of the Hurwitz zeta $\zeta(1/2, z/2)$ on $z \in (0, 2)$. In so doing, we were able to prove the conjectured behavior. Moreover, we were able to sketch the higher-order asymptotic behavior, something that would have been impossible without discovery of an analytic formula.

Does this deserve to be called a closed form? In our opinion, resoundingly yes unless all inverse functions such as that in Bornemann’s (12) are to be eschewed. Such constants are especially interesting in light of even more recent work by Steve Strogatz and his collaborators on *chimera*, coupled systems which can self-organize in parts of their domain and remain disorganized elsewhere; see Figure 7 taken from [43]. In this case, observed numerical limits still need to be put in closed form. \square

It is a frequent experience of ours that, as in Example 13, the need for high-accuracy computation drives the development of effective analytic expressions (closed forms?), which in turn typically shed substantial light on the subject being studied.

Example 14 (Box integrals [3], [7], [23]). There has been recent research on the calculation of the expected distance of points inside a hypercube to the hypercube. Such expectations are also called “box integrals” [23]. So, for example, the expectation $\langle |\vec{r}| \rangle$ for random $\vec{r} \in [0, 1]^3$ has the closed form

$$\frac{1}{4}\sqrt{3} - \frac{1}{24}\pi + \frac{1}{2}\log(2 + \sqrt{3}).$$

Incidentally, box integrals are not just a mathematician’s curiosity; the integrals have been used recently to assess the randomness of brain synapses positioned within a parallelepiped [38]. Indeed, we had cognate results for

$$\Delta_d(s) := \int_{[0,1]^d} \int_{[0,1]^d} \|x - y\|_2^s dx dy,$$

which gives the moments of the distance between two points in the hypercube.

In a lovely recent paper [52] Stephan Steinerberger has shown that in the limit, as the dimension goes to infinity,

$$\begin{aligned}
(18) \quad & \lim_{d \rightarrow \infty} \left(\frac{1}{d} \right)^{s/p} \int_{[0,1]^d} \int_{[0,1]^d} \|x - y\|_p^s dx dy \\
& = \left(\frac{2}{(p+1)(p+2)} \right)^{s/p}
\end{aligned}$$

for any $s, p > 0$. In particular, with $p = 2$, this gives a first-order answer to our earlier published request for the asymptotic behavior of $\Delta_d(s)$.

A quite recent result is that all box integrals $\langle |\vec{r}|^n \rangle$ for integer n and dimensions 1, 2, 3, 4, 5 are *hyperclosed*, in the sense of our “Sixth Approach”

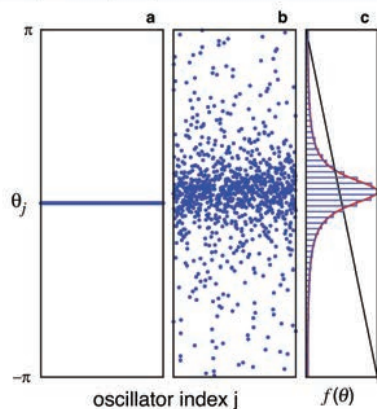


FIG. 1 (color online). Snapshot of a chimera state, obtained by numerical integration of (1) with $\beta = 0.1$, $A = 0.2$, and $N_1 = N_2 = 1024$. (a) Synchronized population. (b) Desynchronized population. (c) Density of desynchronized phases predicted by Eqs. (6) and (12) (smooth curve) agrees with observed histogram.

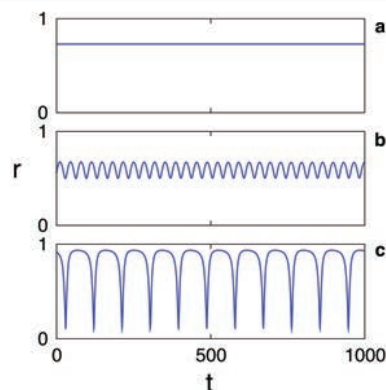


FIG. 2 (color online). Order parameter r versus time. In all three panels, $N_1 = N_2 = 128$ and $\beta = 0.1$. (a) $A = 0.2$: stable chimera; (b) $A = 0.28$: breathing chimera; (c) $A = 0.35$: long-period breather. Numerical integration began from an initial condition close to the chimera state, and plots shown begin after allowing a transient time of 2000 units.

Figure 7. Simulated chimera (figures and parameters from [43]).

section. It turns out that five-dimensional box integrals have been especially difficult, depending on knowledge of a hyperclosed form for a single definite integral $J(3)$, where

$$(19) \quad J(t) := \int_{[0,1]^2} \frac{\log(t + x^2 + y^2)}{(1 + x^2)(1 + y^2)} dx dy.$$

A proof of hyperclosure of $J(t)$ for algebraic $t \geq 0$ is established in [23, Thm. 5.1]. Thus $\langle |\vec{r}|^{-2} \rangle$ for $\vec{r} \in [0, 1]^5$ can be written in explicit hyperclosed form involving a 10^5 -character symbolic $J(3)$; the authors of [23] were able to reduce the 5-dimensional box integral down to “only” 10^4 characters. A companion integral $J(2)$ also starts out with about 10^5 characters but reduces stunningly to only a few dozen characters, namely,

$$(20) \quad J(2) = \frac{\pi^2}{8} \log 2 - \frac{7}{48} \zeta(3) + \frac{11}{24} \pi \text{Cl}_2\left(\frac{\pi}{6}\right) - \frac{29}{24} \pi \text{Cl}_2\left(\frac{5\pi}{6}\right),$$

where Cl_2 is the *Clausen function* $\text{Cl}_2(\theta) := \sum_{n \geq 1} \sin(n\theta)/n^2$ (Cl_2 is the simplest nonelementary Fourier series).

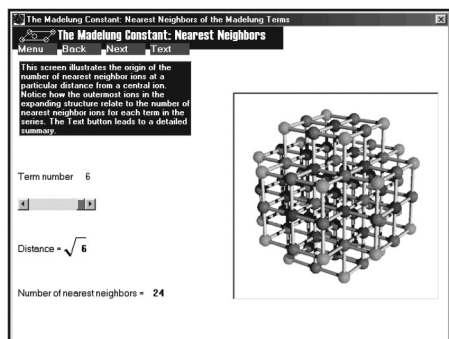
Automating such reductions will require a sophisticated simplification scheme with a very large and extensible knowledge base. With a current research assistant, Alex Kaiser at Berkeley, we have started to design software to refine and automate this process and to run it before submission of any equation-rich paper (see [9]). This semi-automated integrity checking becomes pressing when, as above, verifiable output from a symbolic manipulation can be the length of a Salinger novella. \square

Profound Curiosities

In our treatment of numbers enjoying hyperclosure or superclosure, we admitted that such numbers are countable, and so almost all complex numbers cannot be given a closed form along such lines. What is stultifying is: *How do we identify an explicit number lying outside such countable sets?* The situation is tantamount to the modern bind in regard to *normal numbers*, numbers which to some base have statistically random digit structure in a certain technical sense. The bind is, though almost all numbers are *absolutely normal* (i.e., normal to every base 2, 3, ...), we do not know a single fundamental constant that is provably absolutely normal. (We do know some “artificial” normal numbers; see [14].)

Here is one possible way out of the dilemma. In the theory of computability, the existence of *noncomputable* real numbers, such as an encoded list of halting Turing machines, is well established. The celebrated *Chaitin constant* Ω is a well-known noncomputable. So a “folk” argument goes: *Since every element of the ring of hyperclosure \mathbb{H} can be computed via converging series, it should be that $\Omega \notin \mathbb{H}$.* A good research problem would be to make this heuristic rigorous.

Let us focus on some constants that *might* not be hyperclosed (nor superclosed). One such constant is the celebrated *Euler constant* $\gamma := \lim_{n \rightarrow \infty} \sum_{k=1}^n 1/k - \log n$. We know of no hypergeometric form for γ ; said constant may well lie outside \mathbb{H} (or even \mathbb{S}). There *are* expansions for



A screen from the Madelung Constant section of Lattice Energetics.

(a) NaCl nearest neighbors



(b) CMS Prize Sculpture

Figure 8. Two representations of salt.

the Euler constant, such as

$$\gamma = \log \pi - 4 \log \Gamma\left(\frac{3}{4}\right) + \frac{4}{\pi} \sum_{k \geq 1} (-1)^{k+1} \frac{\log(2k+1)}{2k+1},$$

and even more exotic series (see [13]). But in the spirit of the present treatment, we do not want to call the infinite series closed, because it is not hypergeometric per se. Relatedly, the classical Bessel expansion is

$$K_0(z) = -\left(\ln\left(\frac{z}{2}\right) + \gamma\right) I_0(z) + \sum_{n=1}^{\infty} \frac{\sum_{k=1}^{n-1} \frac{1}{k}}{(n!)^2} \left(\frac{z^2}{4}\right)^n.$$

Now $K_0(z)$ has a (degenerate) Meijer-G representation—so potentially is superclosed for algebraic z —and $I_0(z)$ is accordingly hyperclosed, but the nested harmonic series on the right is again problematic. Again, γ is conjectured not to be a period [42].

Example 15 (Madelung constant [22], [37], [58]). Another fascinating number is the *Madelung constant*, \mathcal{M} , of chemistry and physics [22, Section 9.3]. This is the potential energy at the origin of an oscillating-charge crystal structure (most often

said crystal is NaCl (salt), as illustrated in Figure 8). Image (b) is of a Helaman Ferguson sculpture based on \mathcal{M} which is awarded biannually by the Canadian Mathematical Society as part of the David Borwein Career Award) and is given by the formal (conditionally convergent [18]) sum

$$(21) \quad \mathcal{M} := \sum_{(x,y,z) \neq (0,0,0)} \frac{(-1)^{x+y+z}}{\sqrt{x^2 + y^2 + z^2}} = -1.747564594633\dots$$

and has never been given what a reasonable observer would call a closed form. Nature plays an interesting trick here: There are other crystal structures that *are* tractable, yet somehow this exquisitely symmetrical salt structure remains elusive. In general, even-dimensional crystal sums are more tractable than odd for the same modular function reasons that the number of representations of a number as the sum of an even number of squares is. But this does not make them easy, as illustrated by Example 2.

Here we have another example of a constant having no known closed form yet is rapidly calculable. A classical rapid expansion for the Madelung constant is due to Benson:

$$(22) \quad \mathcal{M} = -12\pi \sum_{m,n \geq 0} \operatorname{sech}^2\left(\frac{\pi}{2} \sqrt{(2m+1)^2 + (2n+1)^2}\right),$$

in which convergence is exponential. Summing for $m, n \leq 3$ produces $-1.747564594\dots$, correct to eight digits. There are a great many other such formulae for \mathcal{M} (see [22], [35]).

Through the analytic methods of Buhler, Crandall, Tyagi, and Zucker since 1999 (see [35], [37], [55], [58]), we now know approximations such as

$$\mathcal{M} \approx -\frac{1}{8} - \frac{\log 2}{4\pi} + \frac{8\pi}{3} + \frac{1}{\sqrt{8}} + \frac{\Gamma(\frac{1}{8})\Gamma(\frac{3}{8})}{\pi^{3/2}\sqrt{2}} + \log \frac{k_4^2}{16k_4'},$$

where $k_4 := ((2^{1/4} - 1)/(2^{1/4} + 1))^2$. Two remarkable things: First, this approximation is good to the same thirteen decimals we give in (21); the missing $O(10^{-14})$ error here is a rapidly, exponentially converging—but alas infinite—sum in this modern approximation theory. Second, this six-term approximation itself is indeed hyperclosed, the only problematic term being the Γ -function part, but we did establish in our “Sixth Approach” section that $B(1/8, 3/8)$ and also $1/\pi$ are hyperclosed, which is enough. Moreover, the work of Borwein and Zucker [27] also settles hyperclosure for that term. \square

Certainly we have nothing like a proof, or even the beginnings of one, that \mathcal{M} (or γ) lies outside \mathbb{H} (or even \mathbb{S}), but we ask on an intuitive basis, Is a constant such as the mighty \mathcal{M} telling us that it is not hyperclosed, in that our toil only seems to bring more “closed-form” terms into play, with no exact resolution in sight?

Concluding Remarks and Open Problems

- We have posited several approaches to the elusive notion of “closed form”, but what are the intersections and interrelations of said approaches? For example, can our “Fourth Approach” be precisely absorbed into the evidently more general “Sixth Approach” (hyperclosure and superclosure)?

- How do we find a *single* number that is provably not in the ring of hyperclosure \mathbb{H} ? (Though no such number is yet known, *almost all* numbers are, as noted, not in said ring!) The same question persists for the ring of superclosure, \mathbb{S} . Furthermore, how precisely can one create a *field* out of $\mathbb{H}^{\mathbb{H}}$ via appropriate operator extension?

- Though \mathbb{H} is a subset of \mathbb{S} , how might one prove that $\mathbb{H} \neq \mathbb{S}$? (Is the inequality even true?) Likewise, is the set of closed forms in the sense of [48, Ch. 8] (only finite linear combinations of hypergeometric evaluations) properly contained in our \mathbb{H} ? And what about a construct such as $\mathbb{H}^{\mathbb{H}^{\mathbb{H}}}$? Should such an entity be anything really new? Lest one remark on the folly of such constructions, we observe that most everyone would say $\pi^{\pi^{\pi}}$ is a closed form!

- Having established the property of hyperclosure for $\Gamma^b(a/b)$, are there any cases where the power b may be brought down? For example, $1/\pi$ is hyperclosed, but what about $1/\sqrt{\pi}$?

- What is a precise connection between the ring of hyperclosure (or superclosure) and the set of periods or of Mahler measures (as in Example 3)?

- There is expounded in reference [23] a theory of “expression entropy”, whereby some fundamental entropy estimate gives the true complexity of an expression. So, for example, an expression having one thousand instances of the polylog token Li_3 might really involve only about 1,000 characters, with that polylogarithm token encoded as a single character, say. (In fact, during the research for [23] it was noted that the entropy of Maple and Mathematica expressions of the *same* entity often had widely varying text-character counts but similar entropy assessments.)

On the other hand, one basic notion of “closed form” is that explicitly infinite sums not be allowed. Can these two concepts be reconciled? Meaning, can we develop a theory of expression entropy by which an explicit, infinite sum is given infinite entropy? This might be difficult, as, for example, a sum $\sum_{n=1}^{\infty} \frac{1}{n^{3/2}}$ takes only a few characters to symbolize, as we just did! If one can succeed, though, in thus resolving the entropy business for expressions, “closed form” might be rephrased as “finite entropy”.

In any event, we feel strongly that the value of closed forms increases as the complexity of the objects we manipulate computationally and inspect mathematically grows, and we hope we

have illustrated this. Moreover, we belong to the subset of mathematicians that finds fun in finding unanticipated closed forms.

Acknowledgements

Thanks are due to David Bailey and Richard Brent for many relevant conversations and to Armin Straub for the complex plots of W_3 and W_4 .

References

- [1] G. E. ANDREWS, R. ASKEY, and R. ROY, *Special Functions*, Cambridge University Press, 1999.
- [2] DAVID BAILEY, ROBERTO BARRIO, and JONATHAN BORWEIN, High-precision computation: Mathematical physics and dynamics, *Applied Mathematics and Computation* **218** (2012), 10106–10121, <http://dx.doi.org/10.1016/j.amc.2012.03.087>.
- [3] D. BAILEY, J. BORWEIN, N. CALKIN, R. GIRGENSOHN, R. LUKE, and V. MOLL, *Experimental Mathematics in Action*, A K Peters, Natick, MA, 2007.
- [4] D. H. BAILEY and J. M. BORWEIN, 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, D. BORWEIN, J. M. BORWEIN, and R. E. CRANDALL, Hypergeometric forms for Ising-class integrals, *Experimental Mathematics* **16** (2007), 257–276.
- [6] D. H. BAILEY, J. M. BORWEIN, and R. E. CRANDALL, Resolution of the Quinn-Rand-Strogatz constant of nonlinear physics, *Experimental Mathematics* **18** (2008), 107–116.
- [7] D. H. BAILEY, J. M. BORWEIN, and R. E. CRANDALL, Advances in the theory of box integral, *Mathematics of Computation*, **79** (2010), 1839–1866.
- [8] D. H. BAILEY, J. M. BORWEIN, D. M. BROADHURST, and L. GLASSER, Elliptic integral representation of Bessel moments, *J. Phys. A: Math. Theory* **41** (2008), 5203–5231.
- [9] D. H. BAILEY, J. M. BORWEIN, and A. KAISER, Automated simplification of large symbolic expressions, preprint, August 2012.
- [10] D. H. BAILEY, J. M. BORWEIN, V. KAPOOR, and E. WEISSTEIN, Ten problems in experimental mathematics, *American Mathematical Monthly* **113** (2006), 481–409.
- [11] D. H. BAILEY, P. B. BORWEIN, and S. PLOUFFE, On the rapid computation of various polylogarithmic constants, *Mathematics of Computation* **66** (1997), 903–913.
- [12] R. BAILLIE, D. BORWEIN, and J. BORWEIN, Some sinc sums and integrals, *American Math. Monthly* **115** (2008), 888–901.
- [13] D. BAILEY and R. CRANDALL, On the random character of fundamental constant expansions, *Experimental Mathematics* **10** (2001), 175–190.
- [14] ———, Random generators and normal number, *Experimental Mathematics* **11** (2002), 527–547.
- [15] FOLKMAR BORNEMANN, *How Mathematica and Maple Get Meijer’s G-function into Problem 9*, preprint, 2005. Available at <http://www-m3.ma.tum.de/bornemann/Numerikstreifzug/Chapter9/MeijerG.pdf>.
- [16] F. BORNEMANN, D. LAURIE, S. WAGON, and J. WALDVOGEL, *The SIAM 100-Digit Challenge: A Study*

- in *High-Accuracy Numerical Computing*, SIAM, Philadelphia, 2004.
- [17] E. J. BOROWSKI and J. M. BORWEIN, *Web-Linked Dictionary of Mathematics*, Smithsonian/Collins Edition, 2006.
 - [18] D. BORWEIN, J. M. BORWEIN, and K. F. TAYLOR, Convergence of lattice sums and Madelung's constant, *J. Math. Phys.* **26** (1985), 2999–3009.
 - [19] J. M. BORWEIN, The SIAM 100 Digits Challenge, extended review in the *Mathematical Intelligencer* **27** (2005), 40–48.
 - [20] J. M. BORWEIN and D. H. BAILEY, *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, extended second edition, A K Peters, Natick, MA, 2008.
 - [21] J. M. BORWEIN, D. H. BAILEY, and R. GIRGENSOHN, *Experimentation in Mathematics: Computational Roads to Discovery*, A K Peters, Natick, MA, 2004.
 - [22] J. M. BORWEIN and P. B. BORWEIN, *Pi and the AGM*, John Wiley, 1987.
 - [23] J. BORWEIN, D. NUYENS, A. STRAUB, and JAMES WAN, Some arithmetic properties of short random walk integrals, *Ramanujan Journal* **26** (2011), 109–132.
 - [24] JONATHAN M. BORWEIN, ARMIN STRAUB, and JAMES WAN, Three-step and four-step random walks, *Experimental Mathematics*, in press, January 2013.
 - [25] J. M. BORWEIN and K. DEVLIN, *The Computer as Crucible*, A K Peters, Natick, MA, 2008.
 - [26] M. BRONSTEIN, R. M. CORLESS, J. H. DAVENPORT, and D. J. JEFFREY, Algebraic properties of the Lambert W function from a result of Rosenlicht and of Liouville, *Integral Transforms and Special Functions* **19** (2008), 709–712.
 - [27] J. M. BORWEIN and I. J. ZUCKER, Fast evaluation of the gamma function for small rational fractions using complete elliptic integrals of the first kind, *IMA J. Numerical Analysis* **12** (1991), 519–526.
 - [28] J. CARETTE, private communication, 2010.
 - [29] ———, Understanding expression simplification, *Proceedings of International Conference on Symbolic and Algebraic Computation*, ACM, 2004, pp.72–79.
 - [30] S. CHANDRASEKHAR, *An Introduction to the Study of Stellar Structures*, Dover, New York, 1967.
 - [31] T. Y. CHOW, What is a closed-form number? *American Mathematical Monthly* **106** (1999), 440–448.
 - [32] B. A. CIPRA, An introduction to the Ising model, *Amer. Math. Monthly* **94** (1987), no. 10, 937–959.
 - [33] R. E. CRANDALL, *Topics in Advanced Scientific Computation*, Springer, New York, 1996.
 - [34] ———, *Theory of ROOF Walks*, 2007. Available at <http://people.reed.edu/~crandall/papers/ROOF.pdf>.
 - [35] ———, New representations for the Madelung constant, *Experimental Mathematics* **8** (1999), 367–379.
 - [36] ———, *Analytic representations for circle-jump moments*, PSiPress, 21 November 2009; http://www.perfscipress.com/papers/AnalyticWn_psiPress.pdf
 - [37] R. E. CRANDALL and J. E. BUHLER, Elementary expansions for Madelung constants, *J. Phys. A: Math. Gen.* **20** (1987), 5497–5510.
 - [38] R. CRANDALL, *On the fractal distribution of brain synapses*, preprint, 2010.
 - [39] C. DENINGER, Deligne periods of mixed motives, theory and the entropy of certain actions, *J. Amer. Math. Soc.* **10** (1997), 259–281.
 - [40] H. R. P. FERGUSON, D. H. BAILEY, and S. ARNO, Analysis of PSLQ, an integer relation finding algorithm, *Mathematics of Computation* **68**, no. 225 (1999), 351–369.
 - [41] BRIAN HAYES, Why W? *American Scientist* **93** (2005), 1004–1008.
 - [42] M. KONTSEVICH and D. ZAGIER, Periods. In *Mathematics Unlimited—2001 and Beyond*, Springer-Verlag, 2005, pp. 771–808.
 - [43] E. A. MARTENS, C. R. LAING, and S. H. STROGATZ, Solvable model of spiral wave chimeras, *Physical Review Letters* **104** (2010), 044101.
 - [44] B. MCCOY and TAI TSUN WU, *The Two-Dimensional Ising Model*, Harvard Univ. Press, 1973.
 - [45] F. W. J. OLVER, D. W. LOZIER, R. F. BOISVERT, and C. W. CLARK, *NIST Handbook of Mathematical Functions*, Cambridge University Press, 2010.
 - [46] B. D. NENER, N. FOWKES, and L. BORREDON, Analytical models of optical refraction in the troposphere, *JOSA* **20** (2003), 867–875.
 - [47] K. OLDHAM, J. MYLAND, and J. SPANIER, *An Atlas of Functions: With Equator, the Atlas Function Calculator*, Springer-Verlag, 2009.
 - [48] M. PETKOVSEK, H. WILF, and D. ZEILBERGER, *A = B*, A K Peters, Ltd., 1996. Available at <http://www.math.upenn.edu/~wilf/AeqB.html>.
 - [49] D. QUINN, R. RAND, and S. STROGATZ, Singular unlocking transition in the Winfree model of coupled oscillators, *Phys. Rev. E* **75** (2007), 036218-1-10.
 - [50] M. ROGERS, *Hypergeometric formulas for lattice sums and Mahler measures*, preprint, 2010.
 - [51] C. J. SMYTH, *Explicit formulas for the Mahler measures of families of multivariable polynomials*, preprint, 2003.
 - [52] S. STEINERBERGER, Extremal uniform distribution and random chord lengths, *Acta Mathematica Hungarica* **130** (2011), 321–339.
 - [53] C. J. SMYTH, The Mahler measure of algebraic numbers: A survey, Conference Proceedings, University of Bristol, 3–7 April 2006. LMS Lecture Note Series, vol. 352, Cambridge University Press, Cambridge, 2008, pp. 322–349.
 - [54] N. TEMME, *Special Functions, an Introduction to the Classical Functions of Mathematical Physics*, John Wiley & Sons, New York, 1996.
 - [55] S. TYAGI, New series representation for the Madelung constant, *Progr. Theoret. Phys.* **114** (2005), 517–521.
 - [56] E. W. WEISSTEIN, *Closed-Form Solution*, from MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/Closed-FormSolution.html>.
 - [57] B. L. WALDEN and L. A. WARD, A harmonic measure interpretation of the arithmetic-geometric mean, *American Mathematical Monthly* **114** (2007), 610–622.
 - [58] I. J. ZUCKER, Some infinite series of exponential and hyperbolic functions, *SIAM J. Math. Anal.* **15** (1984), 406–413.