USING INTEGER RELATIONS ALGORITHMS FOR FINDING RELATIONSHIPS AMONG FUNCTIONS

MARC CHAMBERLAND

Nearly three decades ago the first integer relations algorithm was developed. Given a set of numbers $\{x_1, \ldots, x_m\}$, an integer relations algorithm seeks integers $\{a_1, \ldots, a_m\}$ such that $a_1x_1+\cdots+a_mx_m = 0$. One of the most popular and efficient of these is the PSLQ algorithm, listed as one of the top ten algorithms of the 20^{th} century[2]. This algorithm either finds the integers or obtains lower bounds on the sizes of coefficients for which such a relation will hold. PSLQ has been implemented in both Maple and Mathematica. Typically a high degree of numerical precision is needed for PSLQ to run effectively. If the precison is not sufficiently high, "large" coefficients are produced suggesting a relation has not been found.

PSLQ has been used to find relationships between various constants. Its first well-known success was in finding the formula

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

discovered by Bailey, Borwein, and Plouffe[4]. This opened up the whole area of what are now called (after the preceeding authors) BBP series. Examples include the identity

$$\frac{\pi^2}{6} = \sum_{k=0}^{\infty}\left\{\frac{G^2}{(5k+1)^2} - \frac{G}{(5k+2)^2} - \frac{G^2}{(5k+3)^2} + \frac{G^5}{(5k+4)^2} + \frac{2\,G^5}{(5k+5)^2}\right\} g^{5k}$$

where $g = (\sqrt{5}-1)/2$ and $G = (\sqrt{5}+1)/2 = g^{-1}$, or the conjecture that

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

equals

$$\frac{24}{7\sqrt{7}}\int_{\pi/3}^{\pi/2} \log\left|\frac{\tan t + \sqrt{7}}{\tan t - \sqrt{7}}\right| dt.$$

This last "identity" has been verified to over 20,000 decimal digits [7], but a traditional proof is still lacking. Indeed, PSLQ has subsequently been used to find a variety of other relationships involving infinite series, integrals and special functions. Many interesting examples can be found in the recent article [5] and the book [3]. Despite such diverse successes, mathematicians have been slow to utilize this tool.

Date: December 4, 2007.

1991 *Mathematics Subject Classification.* Primary 11.

Key words and phrases. Integer relations algorithms, PSLQ, Ramanujan, Fermat's Last Theorem, Eisenstein series, Fibonacci numbers .

This paper piggybacks on the PSLQ algorithm to experimentally find integer linear relationships among **functions**. The idea is straightforward. Suppose it is suspected that a function $f(x_1, \ldots, x_m)$ can be expressed as a rational linear combination of the functions $g_1(x_1, \ldots, x_m), \ldots, g_n(x_1, \ldots, x_m)$. Evaluate all these functions at some random point in the intersection of their domains and apply the PSLQ algorithm. As a check, run PSLQ again at another point to produce a second set of coefficients. If the second response is a scaled version of the first response, we conjecture that a relationship has been found.

Some examples of this approach are buried in the literature. In [11], the authors use this approach to show that if

$$J(x) := \int_0^x \frac{(\log(1-t))^2}{2t} dt$$

where $0 \leq x \leq 1$, then

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

The current note aims to demonstrate this PSLQ approach for functions in a variety of contexts.

It should be noted that in some cases, one can solve this problem using linear algebra. If the common domain admits at least n points where exact function evaluations can be made and the n equations are linearly independent, then the sought-after coefficients can be determined exactly. Even if this is the case, however, it may be easier to automate the process using the PSLQ approach.

The rest of this paper is a collection of applications.

2. Powers of Sine and Cosine

This is the simplest example because the existence of such formulas is well-known. The function $\sin((2n-1)x)$ can be written as a linear combination of $\sin(x), \sin^3(x), \ldots, \sin^{2n-1}(x)$. Using the PSLQ approach, one finds, for example, that

$$\begin{aligned} \sin(19x) \quad = \quad & 19\sin(x) - 1140\sin(x)^3 + 20064\sin(x)^5 - 160512\sin(x)^7 \\ & +695552\sin(x)^9 - 1770496\sin(x)^{11} + 2723840\sin(x)^{13} \\ & -2490368\sin(x)^{15} + 1245184\sin(x)^{17} - 262144\sin(x)^{19} \end{aligned}$$

This is obtained with the following Maple commands:

```
> with(IntegerRelations):
> Digits := 100;
> x := 0.8234789345738979234583945;
> PSLQ( [sin(19*x), seq(sin(x)^(2*k-1), k=1..10)] );
```

While using linear algebra directly is not easily possible because the sin function has too few simple evaluations, a variety of other means are possible to find the coefficients $\{a_1, \ldots, a_{10}\}$ such that

$$\sin(19x) = a_1 \sin(x) + a_2 \sin(3x) + \cdots + a_{10} \sin(19x)$$

- Expand using the multiple-angle formula for $\sin(nx)$ and $\cos(nx)$ recursively.
- Use Fourier series for each of the powers of $\sin(x)$ then solve a linear system.

- Expand the right side with Taylor expansions and match the first ten coefficients. Maple commands:

```
> ex := sin(19*x) - sum( a[i]*sin(x)^(2*i-1), i=1..10 );
> ex2 := series( ex, x, 20 );
> solve( {seq(coeff(ex2,x^(2*k-1)),k=1..10)}, {seq(a[k],k=1..10)} );
```

- Differentiate the equation ten times then solve a linear system with $x = 0$. Maple commands:

```
> ex := sin(19*x) - sum(a[k]*sin(x)^(2*k-1), k=1..10 );
> solve( { seq( diff( ex, x$j), j=1..10 ) }, {seq(a[i], i=1..10)} );
```

The point here is that the PSLQ approach finds these coefficients with little overhead. It should be noted that if n is replaced with larger values, more digital precision is needed.

3. Lamé-like Equations and Fermat's Last Theorem

Before general theories to approach Fermat's Last Theorem were developed by Kummer and Sophie Germain, mathematicians considered each exponent one at a time. The case for exponent $n = 7$ (see [17, 18]) was first proved by Lamé (1839). The proof was substantially shortened by Lebesgue (1840), but both used the equation

$$\frac{(x+y+z)^7 - (x^7+y^7+z^7)}{7(x+y)(x+z)(y+z)} = (x^2+y^2+z^2+xy+xz+yz)^2 + xyz(x+y+z).$$

It seems such identities were not found for higher values of the exponent n. Using Fermat's Little Theorem, one may show that the expression

$$\frac{(x+y+z)^p - (x^p+y^p+z^p)}{p(x+y)(x+z)(y+z)}$$

is a polynomial for all odd p. One has

$$\frac{(x+y+z)^3 - (x^3+y^3+z^3)}{3(x+y)(x+z)(y+z)} = 1$$

and

$$\frac{(x+y+z)^5 - (x^5+y^5+z^5)}{5(x+y)(x+z)(y+z)} = x^2+y^2+z^2+xy+xz+yz$$

However, how can one find a compact form for higher cases such as $n = 11$ or $n = 13$? Define the symmetric functions

$$h_1 := x+y+z,\ h_2 := x^2+y^2+z^2+xy+xz+yz,\ h_3 := xyz$$

Now apply the PSLQ approach (to random x and y) to a linear combination of the functions

$$h_1^8, h_1^6h_2, h_1^5h_3, h_1^4h_2^2, h_1^3h_2h_3, h_1^2h_3^2, h_1h_2^2h_3, h_3^2h_2, h_2^4$$

to produce the formula

$$\begin{aligned}\frac{(x+y+z)^{11} - (x^{11}+y^{11}+z^{11})}{11(x+y)(x+z)(y+z)} &= h_1^6h_2 - 2h_1^4h_2^2 - 2h_1^3h_2h_3 + h_1^2h_2^3 + h_1^2h_3^2 \\ &\quad +5h_1h_2^2h_3 + h_2h_3^2 + h_2^4\end{aligned}$$

Similarly, one has for $n = 13$ the formula

$$\begin{aligned}\frac{(x+y+z)^{13}-(x^{13}+y^{13}+z^{13})}{13(x+y)(x+z)(y+z)} &= h_1h_3^3+2h_2^2h_3^2-2h_1^4h_3^2+3h_1h_2^3h_3\\ &\quad -8h_1^3h_2^2h_3+3h_1^5h_2h_3+h_2^5-3h_1^2h_2^4\\ &\quad +h_1^4h_2^3+h_1^6h_2^2+h_1^2h_3h_5+h_1h_2^2h_5\end{aligned}$$

These equations can be easily — albeit tirelingly — proved by algebraic expansions. Some care is needed in defining the set of functions over which we look for a linear combination. If this set of functions is linearly dependent (this happens if we added the term h_1^8 in the last equation), many different combinations can be obtained. While one is tempted to use Maple's built-in partition capabilities, the resulting set of functions is linearly dependent. It is not clear whether the equations for $n = 11$ and $n = 13$ can be used to gain further insight into Fermat's equation.

4. Ramanujan's 6-8-10 Equation and Beyond

Among Ramanujan's many beautiful equations is the 6-8-10 equation

$$\begin{aligned}&64[(a+b+c)^6+(b+c+d)^6-(c+d+a)^6-(d+a+b)^6+(a-d)^6-(b-c)^6]\\ &\quad\times[(a+b+c)^{10}+(b+c+d)^{10}-(c+d+a)^{10}-(d+a+b)^{10}+(a-d)^{10}-(b-c)^{10}]\\ &\quad=45[(a+b+c)^8+(b+c+d)^8-(c+d+a)^8-(d+a+b)^8+(a-d)^8-(b-c)^8]^2\end{aligned}$$

when $ad = bc$. Berndt and Bhargava[8] cite this as "one of the most fascinating identities we have ever seen." More concisely, let

(4.1)

$$f_m := (1+x+y)^m+(-x-y-xy)^m-(-y-xy-1)^m-(xy+1+x)^m+(-1+xy)^m-(-x+y)^m$$

Ramanujan's equation may be compactly stated as

$$45f_8^2 = 64f_6f_{10}. \tag{4.2}$$

Proofs of equation (4.2) may be found in [9] and [16]. It is also noted (see [8]) that $f_2 = 0$ and $f_4 = 0$. Without computation, this follows directly from the observation that $x = 0, 1, -1, -2, -1/2$ (and by symmetry $y = 0, 1, -1, -2, -1/2$) are zeros of f_{2m}.

Attempts to generalize equation (4.2) had been unfruitful until the recent work of Hirschhorn[14]. He introduced the glaring minus signs in equation (4.2) which are not seen when m is even. By allowing odd powers, he found

$$21f_5^2 = 25f_3f_7$$

One also easily notes that $f_1 = 0$. Looking for other possible equations with the PSLQ approach, the following new discoveries were made:

$$8f_5f_6 = 5f_3f_8$$

$$15f_6f_7 = 7f_3f_{10}$$

$$330f_7^2 = 539f_5f_9 - 245f_3f_{11}$$

$$308f_{10}^2 = 525f_8f_{12} - 300f_6f_{14}$$

$$1763580f_{11}^2 = 2735810f_9f_{13} - 1172490f_7f_{15} + 144837f_5f_{17} + 71995f_2f_{19}$$

$$6395400f_{14}^2 = 10445820f_{12}f_{16} - 5448212f_{10}f_{18} + 1460151f_8f_{20} + 49980f_6f_{22}$$

It is unclear whether more such identities exist or if there is a proof of these formulas beyond mindless expansion.

5. Eisenstein Series

An important tool in the study of partitions and modular forms are Eisenstein series. In particular, consider the functions

$$E_{2k}(q) := 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} \frac{n^{2k-1} q^n}{1-q^n}$$

where B_j is the j^{th} Bernoulli number. It is well-known from the theory of modular forms of level one [15, p.118] that one may always write E_{2k} as a rational linear combination of $E_4^a E_6^b$ where $4a + 6b = 2k$ and a and b are non-negative integers. Some oft-cited examples are

$$E_8 = E_4^2, \; E_{10} = E_6 E_4.$$

Using the PSLQ approach, one quickly finds, for example, that

$$691 E_{12} = 250 E_6^2 + 441 E_4^3$$

$$77683 E_{22} = 20500 E_4 E_6^3 + 57183 E_4^4 E_6$$

Note that since E_{2k} cannot be evaluated exactly except at $q = 0$, a direct linear algebra approach is ineffective. Of course one could *approximate* the series at other values of q, perform the linear algebra, and note that the coefficient values are "close" to integers, but the PSLQ approach is less cumbersome. An analytic approach is to expand the functions E_{2k} as power series about $q = 0$ and compare like powers of q to determine the desired coefficients.

6. Fibonacci Identities

The Fibonacci numbers have proven to be a magnet for both budding and seasoned mathematicians alike. These numbers admit many beautiful formulas, sometimes bringing in their cousins, the Lucas numbers. Two fascinating equations are

$$F_{2n} = -F_{n-1}^2 + F_{n+1}^2,$$

$$F_{3n} = -F_{n-1}^3 + F_n^3 + F_{n+1}^3.$$

These equations suggest a generalization for F_{kn}. Since F_n can be calculated exactly, a direct linear algebra approach seems like a natural route to take. Moreover, the PSLQ approach will encounter problems because the Fibonacci numbers are integers, so there are many integer linear combinations to be found. How can the PSLQ approach be used in this scenario?

The trick is to "roughen" the Fibonacci "function" by extending the domain to non-integer values. This is accomplished via the Binet formula

$$F_n = \frac{1}{\sqrt{5}} \left(\alpha^n + \left(\frac{-1}{\alpha} \right)^n \right)$$

where $\alpha = (1 + \sqrt{5})/2$. Using this as a definition for the Fibonacci numbers, one sees that the domain of definition extends to all rationals a/b in lowest form where b is odd. The identity $F_n - F_{n-1} - F_{n-2} = 0$ extends to this new domain since it only depends on the fact $\alpha^2 - \alpha - 1 = 0$. Since all Fibonacci identities eventually depend on Binet's formula, the sought-after identities must also extend to this new domain. This means we may use these rationals as test points with the PSLQ approach to conjecture Fibonacci identities. This approach yields examples such as

$$6F_{4n} = -F_{n-2}^4 - 3F_{n-1}^4 + 3F_{n+1}^4 + F_{n+2}^4,$$

$$120F_{6n} = -F_{n-3}^6 - 4F_{n-2}^6 + 20F_{n-1}^6, -20F_{n+1}^6 + 4F_{n+2}^6 + F_{n+3}^6$$

$$240F_{7n} = F_{n-3}^7 + 8F_{n-2}^7 + 40F_{n-1}^7 - 60F_n^7 - 40F_{n+1}^7 + 8F_{n+2}^7 + F_{n+3}^7.$$

Proceeding with other values of k suggests that F_{kn} is a rational linear combination of k k^{th} powers of Fibonacci numbers. After analyzing these formulas, one conjectures

$$\left(\prod_{m=1}^{2n} F_m\right) F_{(2n+1)j} = \sum_{k=-n}^{n} r_{n-k} F_{j+k}^{2n+1} \prod_{p=0}^{n-|k|-1} \frac{F_{2n-p}}{F_{1+p}}$$

where $r : \mathbb{Z} \to \mathbb{Z}$ is defined by

$$r_n = \begin{cases} 1, & n \equiv 0, 1 \mod 4 \\ -1, & n \equiv 2, 3 \mod 4 \end{cases}$$

Note that the form suggests some kind of "Fibonacci binomial coefficient".

References

[1] Bailey, D.H. A Compendium of BBP-Type Formulas for Mathematical Constants. Preprint, http://crd.lbl.gov/~dhbailey/dhbpapers/index.html, (2000).

[2] Bailey, D. *Integer Relation Detection. Computing in Science and Engineering*, 2(1):24–28, (2000).

[3] Borwein, J.M. and Bailey, D. *Mathematics by Experiment*. AK Peters, Natick, (2004).

[4] Bailey, D.H., Borwein, P.B. and Plouffe, S. On the Rapid Computation of Various Polylogarithmic Constants. *Mathematics of Computation*, 66(218):903–913, (1997).

[5] Borwein, J.M., Bailey, D., Kapoor, V. , Weisstein, E.W. Ten Problems in Experimental Mathematics. *American Mathematical Monthly*, 113(6):481–509, (2006).

[6] Bailey, D.H., Crandall, R.E. On the random character of fundamental constant expansions. *Experimental Mathematics*, 10(2):175–190, (2001).

[7] Bailey, D.H. Book Review of *The SIAM 100-Digit Challenge: A study in high-accuracy numerical computing*, Bulletin of the American Mathematical Society, 42(2005), 545–548.

[8] Berndt, B., Bhargava, S. Ramanujan — For Lowbrows. *American Mathematical Monthly*, 100(7):644–656, (1993).

[9] Berndt, B., Bhargava, S. A Remarkable Identity found in Ramanujan's Third Notebook. *Glasgow Mathematical Journal*, 34:341–345, (1992).

[10] Borwein, D., Borwein, J.M., Galway,W.F. Finding and Excluding b-ary Machin-Type BBP Formulae. *Canadian Journal of Mathematics*, 56(5), (2004), 897–925.

[11] Borwein, D., Bradley, D., Broadhurst,D., and Lisonek,P. Special values of multiple polylogarithms. *Transactions of the American Mathematical Society*, 353(3), (2001), 907–941.

[12] R. Brent. Computing Aurifeuillian factors. *Computational algebra and number theory (Sydney, 1992)*, Mathematics and its Applications, 325:201–212, Kluwer, Dordrecht, (1995).

[13] J. Brillhart, D.H. Lehmer, J.L. Selfridge, B. Tuckerman, and S.S. Wagstaff,Jr. *Factorizations of* $b^n \pm 1$. American Mathematical Society, Providence, (1983).

[14] Hirschhorn, M. Two or Three Identities of Ramanujan. *American Mathematical Monthly*, 105:52–55, (1998).

[15] Koblitz, N. *Introduction to Elliptic Curves and Modular Forms*. Springer, New York, (1984).

[16] T. Nanjundiah. A Note on an Identity of Ramanujan. *American Mathematical Monthly*, 100(5):485–487, (1993).

[17] Ribenboim, P. *13 Lectures on Fermat's Last Theorem*. Springer, New York, (1979).

[18] Ribenboim, P. *Fermat's Last Theorem for Amateurs*. Springer, New York, (1999).

Department of Mathematics and Statistics, Grinnell College, Grinnell, IA 50112, USA
E-mail address: chamberl@math.grinnell.edu