Improving the Convergence of Newton's Series Approximation for e

Harlan J. Brothers

Harlan Brothers (hjb@thecountryschool.org) is Director of Technology at The Country School in Madison, Connecticut, where he teaches programming, fractal geometry, robotics, and jazz band. He is currently involved in integrating the study of fractals into standard mathematics curricula through his work with Michael Frame and Benoit Mandelbrot at Yale University. He is also an inventor, holding five U.S. patents. On weekends, he can often be found playing guitar in the New Haven area with the 12-piece Latin jazz band, *Sonido Unidad*.

Introduction

It was Isaac Newton who in 1669 [**7**, p. 235] published what is now known as the Maclaurin series expansion for e. Originally derived from the binomial expansion of the limit definition of e,

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n,$$

it is sometimes referred to as the Direct Method:

$$e = \sum_{k=0}^{\infty}\frac{1}{k!} = \frac{1}{0!}+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\frac{1}{5!}+\frac{1}{6!}+\frac{1}{7!}+\cdots. \tag{1}$$

Over the intervening centuries, it has maintained its place as a supremely simple, elegant, and efficient method for calculating e [**1**, p. 313]. Perhaps for this reason, perennial efforts to calculate e to greater precision have focused primarily on computational techniques employing the Binary Splitting method [**3**] and Fast Fourier Transform (FFT) multiplication [**10**]. The core expression as an infinite series has remained unchallenged as the fastest way to e. However, using techniques familiar to any first-year calculus student, it is indeed possible to derive series that converge more rapidly.

Assessing relative rates of convergence

Methods for comparing the rate at which two series converge to the same limit can be broadly divided into two categories of measurement. One is an algebraic assessment based on the decimal place accuracy (d.p.a.) that is gained from a given number of terms; the greater the d.p.a., the faster the series converges. The other can be generally described as the computational time required to achieve a given degree of accuracy as measured in the context of a particular hardware/software configuration. In this case, a smaller number of processor operations indicates that the algorithm is inherently more efficient; it requires less time to compute, and therefore, in a fundamental sense, is a faster method for achieving a given computation.

Most of the new series that follow are faster than (1) from the algebraic standpoint. To the extent tested, they also appear to be faster in the computational sense **[8]**. However, since the rigorous assessment of computational efficiency (also known as "run-time cost") requires familiarity with theoretical computer science, this paper will rely for the most part on the comparison of decimal place accuracy based on an equivalent number of terms.

To understand how d.p.a. is derived, consider the kth partial sum, S_k, of the series in (1). The exact error is then equal to $e - S_k$ and can be approximated by the first missing term, $1/(k+1)!$. The accuracy, on the other hand, is measured by the position of the last correct decimal digit before the error appears. Since we know that the error can be represented by the $k+1$st term, we also know that the last correct digit is supplied by the kth term, $1/k!$. If we therefore set $1/k! = C \cdot 10^{-d}$, for a constant, C, where $1 < C < 10$, and take the $\log_{10}$ of both sides, we can then solve for d to determine that for the series in (1) the d.p.a. $= \lfloor \log_{10}(k!) \rfloor$, where "$\lfloor\ \rfloor$" denotes the floor function. Decimal place accuracy for the new series that follow can be derived in a similar fashion.

Series compression

In an effort to streamline the already spare series in (1), this article examines ways to "compress" the existing terms. We begin by combining and simplifying pairwise so that, for example,

$$\frac{1}{n!} + \frac{1}{(n+1)!}$$

becomes

$$\frac{n+2}{(n+1)!}.$$

Because this single term now represents two terms in (1) we replace n with $2k$ and take the sum over k to obtain

$$e = \sum_{k=0}^{\infty} \frac{2k+2}{(2k+1)!} = \frac{2}{1!} + \frac{4}{3!} + \frac{6}{5!} + \frac{8}{7!} + \frac{10}{9!} + \frac{12}{11!} + \frac{14}{13!} + \frac{16}{15!} + \cdots. \tag{2}$$

Here the evaluation of 20 terms yields e accurate to 47 decimal places (d.p.a. $= \lfloor \log_{10}((2k+1)!/(2k+2)) \rfloor$). By comparison, the Direct Method offers 18 d.p.a. for an equal number of terms.

If instead one chooses to combine consecutive terms in which n is decreasing

$$\frac{1}{n!} + \frac{1}{(n-1)!},$$

these simplify to

$$\frac{n+1}{n!}.$$

Setting $n = 2k$, the resulting expression becomes

$$e = \sum_{k=0}^{\infty} \frac{2k+1}{(2k)!} = \frac{1}{0!} + \frac{3}{2!} + \frac{5}{4!} + \frac{7}{6!} + \frac{9}{8!} + \frac{11}{10!} + \frac{13}{12!} + \frac{15}{14!} + \cdots. \tag{3}$$

Here the evaluation of 20 terms yields 46 correct digits, indicating that it is slightly less accurate than (2). The interested student may wish to derive the formula that quantifies this difference in accuracy.

Not only are the expressions in (2) and (3) aesthetically pleasing; because they trade a simple increment in the numerator to gain greater algebraic efficiency, from an optimized computational standpoint they are, to accuracies of around 200,000 decimal digits, demonstrated to be over 50% faster than the Direct Method [**8**]. It is important to note that comparisons such as this, which are based on custom programming, can differ widely from those obtained from computer algebra packages whose routines are not necessarily optimized for a given set of operations. For example, using the series in (2) for the computation of 10,000 digits of e, the *Timing* function in *Mathematica* indicates a 96% increase in speed over the Direct Method. While at first glance this appears impressive, it is not as accurate an indication of computational efficiency as the smaller but more realistic 53% increase obtained when the programming for the calculation of both series is optimized [**8**].

Deriving new series

Using a combination of compression schemes, an entire family of new and beautiful series expressions can be derived for e. For example, adding (2) and (3) yields

$$2e = \sum_{k=0}^{\infty} \frac{k+1}{k!} = \frac{1}{0!} + \frac{2}{1!} + \frac{3}{2!} + \frac{4}{3!} + \frac{5}{4!} + \frac{6}{5!} + \frac{7}{6!} + \frac{8}{7!} + \cdots, \tag{4}$$

while dividing (2) by 2 results in

$$\frac{e}{2} = \sum_{k=0}^{\infty} \frac{k+1}{(2k+1)!} = \frac{1}{1!} + \frac{2}{3!} + \frac{3}{5!} + \frac{4}{7!} + \frac{5}{9!} + \frac{6}{11!} + \frac{7}{13!} + \frac{8}{15!} + \cdots. \tag{5}$$

Adding (1) to (5) gives us the interesting form

$$\frac{3e}{2} = \sum_{k=0}^{\infty} \frac{(k+3)^{k[\mathrm{mod}\,2]}}{2^{k[\mathrm{mod}\,2]}k!} = \frac{1}{0!} + \frac{2}{1!} + \frac{1}{2!} + \frac{3}{3!} + \frac{1}{4!} + \frac{4}{5!} + \frac{1}{6!} + \frac{5}{7!} + \cdots, \tag{6}$$

while repeated addition of (1) to (4) can be generalized as

$$x \cdot e = \sum_{k=0}^{\infty} \frac{x-1+k}{k!} = \frac{x-1}{0!} + \frac{x}{1!} + \frac{x+1}{2!} + \frac{x+2}{3!} + \frac{x+3}{4!} + \frac{x+4}{5!} + \cdots,$$
$$x \in \mathrm{C}. \tag{7}$$

The same compression technique can be also be applied to the familiar series for $1/e$ [**9**],

$$\frac{1}{e} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = \frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \cdots. \tag{8}$$

This approximation is easily derived from the power series for e^x [**5**, pp. 774–775] by setting $x = -1$ (see equation (12)). Like (1), it offers 18 d.p.a. for 20 terms. Using compression, one can convert this alternating series to a more rapidly converging monotone series,

$$\frac{1}{e} = \sum_{k=0}^{\infty} \frac{1-2k}{(2k)!} = \frac{1}{0!} - \frac{1}{2!} - \frac{3}{4!} - \frac{5}{6!} - \frac{7}{8!} - \frac{9}{10!} - \frac{11}{12!} - \frac{13}{14!} - \cdots, \tag{9}$$

that results in 46 accurate digits for an equal number of terms.

The derivation of the following decreasing series for e is left to the reader.

$$e = \sum_{k=0}^{\infty} \frac{3-4k^2}{(2k+1)!} = \frac{3}{1!} - \frac{1}{3!} - \frac{13}{5!} - \frac{33}{7!} - \frac{61}{9!} - \frac{97}{11!} - \frac{141}{13!} - \frac{193}{15!} - \cdots. \tag{10}$$

Here the evaluation of 20 terms produces 47 correct digits.

For a more extensive list of similarly derived expressions involving e, see **[2]**.

Combining an arbitrary number of terms

There is, of course, no reason to stop at pairwise compression. For example, in a similar manner to (3), combining

$$\frac{1}{n!} + \frac{1}{(n-1)!} + \frac{1}{(n-2)!}$$

and substituting $3k$ for n gives

$$e = \sum_{k=0}^{\infty} \frac{(3k)^2+1}{(3k)!} = \frac{1}{0!} + \frac{3^2+1}{3!} + \frac{6^2+1}{6!} + \frac{9^2+1}{9!} + \frac{12^2+1}{12!} + \frac{15^2+1}{15!} + \cdots, \tag{11}$$

resulting in 78 correct digits after 20 terms.

When combining more than two terms, the reference term (that containing just $n!$ in the denominator) can be the largest, the smallest, or any term in-between. Regardless of its position, as one combines more terms, the polynomial in the numerator of the summand increases in order to degree $T-1$, where T represents the total number of consecutive terms that are combined into a single term.

A simple recursion formula will generate the numerator of the summand for any value of T where, as in (2), the terms being combined contain values of n that are sequentially increasing:

$$\text{numerator}_T = (n+T-1)\cdot(\text{numerator}_{T-1}) + 1$$
$$\text{numerator}_1 = 1.$$

Below is *Mathematica* code that takes an arbitrary number of terms, shows their compressed simplified form, the summand (with k substituted for n), and the approximate numerical error relative to e.

```
Clear[k]; TC = 1; t = 20;
numerat[1] = 1;
numerat[TC_] := numerat[TC - 1]*(n + TC - 1) + 1;
denominat = (n + TC - 1)!;
summand = Together[Expand[numerat[TC]/denominat]]
expressn = summand /. n -> (TC*k)
k = t; acc = Floor[N[-Log[10, expressn]]] + 15;
Print["Error ≈ ", N[E - N[Sum[expressn,{k,0,t - 1}],acc],acc ]]
```

Here, TC represents the number of terms in (1) to be combined into a single term of the new series while t sets the number of these new terms to be evaluated.

Compressing e^x

It was Leonhard Euler who, in his *Introductio in analysin infinitorum*, 1748, picked up where Newton left off [**6**, pp. 155–157]. Using the limit definition of e, he derived the power series for e^x:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \frac{1}{0!} + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \frac{x^7}{7!} + \cdots, \quad x \in \mathrm{R}. \tag{12}$$

As in the case of (1), this series can be pairwise compressed to obtain

$$e^x = \sum_{k=0}^{\infty} \frac{x^{(2k-1)}(x+2k)}{(2k)!} = \frac{1}{0!} + \frac{x(x+2)}{2!} + \frac{x^3(x+4)}{4!} + \frac{x^5(x+6)}{6!} + \frac{x^7(x+8)}{8!} + \cdots, \quad \{x \in \mathrm{R} \mid x \neq 0\} \tag{13}$$

and

$$e^x = \sum_{k=0}^{\infty} \frac{x^{(2k)}(x+2k+1)}{(2k+1)!} = \frac{x+1}{1!} + \frac{x^3+3x^2}{3!} + \frac{x^5+5x^4}{5!} + \frac{x^7+7x^6}{7!} + \frac{x^9+9x^8}{9!} + \cdots, \quad x \in \mathrm{R}. \tag{14}$$

It is interesting to note that, while the form of the numerators in (13) and (14) are interchangeable, when written in the form of (13) each term contains exactly one operator for addition, multiplication, division, exponentiation, and the factorial function while in the form of (14) each numerator contains both x^k and its derivative, kx^{k-1}.

Because the speed of convergence of a series is determined by the rate at which its denominators increase relative to its numerators, the convergence of these power series improves dramatically for smaller x. It is therefore possible to achieve extremely rapid convergence by employing a technique referred to as *powering* [**4**]. Powering consists of using small values for x and then exponentiating by $1/x$. If x takes the form of 2^{-n}, then it is only necessary to square the result n consecutive times in order to approximate e.

For example, using $x = \frac{1}{2}$ in (14) gives

$$\sqrt{e} = \sum_{k=0}^{\infty} \frac{4k+3}{2^{(2k+1)}(2k+1)!} = \frac{3}{2\cdot 1!} + \frac{7}{2^3\cdot 3!} + \frac{11}{2^5\cdot 5!} + \frac{15}{2^7\cdot 7!} + \frac{19}{2^9\cdot 9!} + \frac{23}{2^{11}\cdot 11!} + \cdots, \tag{15}$$

which for 20 terms is accurate to 59 decimal places. Squaring the result gives e, also accurate to 59 decimal places. For $x = \frac{1}{16}$, the same number of terms produces 96 correct decimal places. Raising this result to the 16th power (squaring it 4 consecutive times) yields e accurate to 94 decimal places. This represents twice the d.p.a. of (2) and over five times the d.p.a. of (1) . The slightly lower accuracy after exponentiation results from the fact that the original error is being multiplied by a factor of $1/x$.

Mathematica code to generate the summand for an arbitrary number of terms in a compressed power series for e can be found at [**2**].

Students should also feel encouraged to explore other well-known series, such as those associated with π, to examine which of them might lend themselves to effective compression. Examples, along with newly derived series, are available at [**2**].

Conclusion

Initial findings indicate that when evaluated on a term by term basis, from both the algebraic and computational standpoints, these new series are substantially faster than (1) for computing the digits of e. When combined with the Binary Splitting method, they are still faster but, to the extent tested, only by about 1.5% [**8**]. It is as yet unexplored as to whether series derived from (12) might be successfully used, perhaps in conjunction with the Binary Splitting method, to achieve substantially faster results than are currently possible. But most importantly, the greatest value of these expressions may lie simply in the process of obtaining them; the methods are exploratory, fun, and well within the grasp of first-year calculus students. Indeed, how often do undergraduates have the opportunity to make personal discoveries that build directly on the work of giants like Newton?

It is hoped that the inherent symmetry and numerical beauty of these newly derived expressions might provide inspiration to students, educators, and all who are drawn by the allure of numbers.

References

1. G. Arfken, *Mathematical Methods for Physicists*, 3rd ed., Academic Press, 1985.
2. H. Brothers, http://www.brotherstechnology.com/math/cmj-supp.html.
3. X. Gourdon and P. Sebah, http://www.brotherstechnology.com/math/binarysplit.html.
4. ———, http://www.brotherstechnology.com/math/gourdon-sebah.html.
5. J. Gullberg, *Mathematics: From the Birth of Numbers*, W. W. Norton & Company, 1997.
6. E. Maor, *e: The Story of a Number*, Princeton University Press, 1994.
7. I. Newton, *The Mathematical Papers of Isaac Newton, Vol. II 1667–1670* (D. T. Whitehead, ed.), Cambridge University Press, 1968.
8. S. Wedeniwski (IBM Deutschland), personal communication, January 14, 2001.
9. Eric Weisstein, http://mathworld.wolfram.com/e.html.
10. ———, http://mathworld.wolfram.com/FastFourierTransform.html.

Acknowledgments. The author would like to thank Sebastian Wedeniwski of IBM Deutschland for his assessment regarding the computational efficiency of these series. Thanks also to Professor John Knox, University of Georgia, for his keen editorial eye and to Professor Robert Speiser, Brigham Young University, for his early support and insightful comments. The author would also like to acknowledge the work of the innominate referee whose detailed comments contributed to a clearer presentation.