Recent Computations of Pi

Donal Hurley

1 Early History

We first give a brief sketch of the history of computing π. Details can be found in [1].

Computations of the number π go back to the time of Archimedes (287-212 BC). He inscribed and circumscribed regular polygons on a circle with diameter 1. He began with hexagons and doubled the number of sides to get polygons of 96 sides which yielded the estimate

$$3\frac{10}{71} < \pi < 3\frac{1}{7}$$

By continuing to double the number of sides, one is in principle able to get as many decimal places of π as one desires. However, the convergence is slow, since the error decreases by about a factor of four per iteration. Until the discovery of calculus in the 17th century, efforts at calculating π relied on the method of Archimedes.

With the use of calculus, series were discovered for π. The formula of Leibniz

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$$

has a very slow convergence rate. Various other series and formulae were used in the computation of π, some of the more famous being

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots$$

due to James Gregory (1638–1675), and

$$\frac{\pi}{4} = 4\arctan\left(\frac{1}{5}\right) - \arctan\left(\frac{1}{239}\right)$$

due to John Machin (1680–1752). Machin substituted the Gregory formula for arctan into his formula to get 100 decimal places of π in 1706.

In 1844 Johann Dase (1824–1861) computed π correctly to 200 decimal places using the formula

$$\frac{\pi}{4} = \arctan\left(\frac{1}{2}\right) + \arctan\left(\frac{1}{5}\right) + \arctan\left(\frac{1}{8}\right)$$

and in 1853 William Shanks published 607 places, although the digits after the 527th place were incorrect. This error was not discovered until 92 years later when D.F. Ferguson produced 530 digits in one of the final hand computations. Two years later Ferguson used a desk calculator to get 808 digits.

The advent of digital computers saw a renewal of efforts to calculate even more digits of π. The first such computation was made in 1949 on ENIAC (Electronic Numerical Integrator and Computer) and 2037 digits were produced in 70 hours by John Von Neumann and his colleagues. In 1958 F. Genuys computed 10,000 digits on an IBM 704. In 1961 D. Shanks and J. W. Wrench Jr. calculated 100,000 digits in less than nine hours on an IBM 7090 [7]. The million-digit mark was set by J. Gilloud and M. Bouyer in 1973 in a feat that took under a day of computation on a CDC 7600. All these computations used series for arctan and identities such as Machin's.

Despite the increased speed of the computers, it was realised that there were limits to the number of digits which could be produced. An examination of the rate of convergence of the arctangent series shows that the arctangent method uses $O(n)$ full-precision operations to compute n decimals of π. By an operation we mean one of $+$, $\times$, $\div$, $\surd$. For example, the Shanks and Wrench computation of 100,000 decimals used 105,000 full-precision operations. Thus there are two basic time costs involved in doubling the number of digits; firstly, the number of operations increases by a factor of two, and secondly, the time for each full-precision operation is about twice as long. So doubling the number of digits lengthens computing time by a factor of four.

In 1975 Brent and Salamin [4,6], independently discovered an algorithm that dramatically lowered the time needed to compute large numbers of digits of π. The Brent-Salamin algorothm requires only $O(\log n)$ full-precision operations for n digits of π, and the ideas used go back to the work of Gauss and Legendre in the early part of the 19th century. The formula for the algorithm exploits the speed of convergence of the defining sequences for the arithmetic-geometric mean of two numbers.

2 The Brent-Salamin Algorithm

If a and b are two positive real numbers, with $a > b$, then we have the familiar arithmetic-geometric mean inequality

$$\frac{a+b}{2} \geq \sqrt{ab}$$

Thus, from two positive numbers a and b we get a second pair, $(a+b)/2$ and $\sqrt{ab}$. If we iterate this process we obtain sequences $\{a_n\}$ and $\{b_n\}$ defined by

$$a_0 = a\,, \quad b_0 = b\,, \quad a_{n+1} = \frac{a_n + b_n}{2}\,, \quad b_{n+1} = \sqrt{a_n b_n}$$

The sequence $\{a_n\}$ is strictly decreasing and bounded below, while the sequence $\{b_n\}$ is strictly increasing and bounded above. A simple computation, beginning with $a_{n+1}^2 - b_{n+1}^2$, shows that

$$a_{n+1} - b_{n+1} < \frac{1}{2}(a_n - b_n)$$

and so one concludes that the sequences have a common limit, which is denoted by $\mathrm{AG}(a,b)$. It can also be shown that

$$a_{n+1} - b_{n+1} < \frac{(a_n - b_n)^2}{8\mathrm{AG}(a,b)}$$

so that $a_n - b_n$ approaches 0 quadratically.

Gauss (1777–1855) studied these limits in his work on elliptic integrals. A complete elliptic integral of the first kind is given by

$$K(a,b) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{a^2\cos^2\theta + b^2\sin^2\theta}}$$

The change of variable $t = a\tan\theta$ yields

$$K(a,b) = \int_0^{\infty} \frac{dt}{\sqrt{(t^2+a^2)(t^2+b^2)}}$$

A further substitution helps to make the connection between this integral and $\mathrm{AG}(a,b)$. If we put $u = (t - \beta^2/t)/2$ in the last integral, we get

$$\begin{aligned} K(a,b) &= \int_0^{\infty} \frac{du}{\sqrt{(u^2+a_1^2)(u^2+b_1^2)}} \\ &= K(a_1,b_1) \end{aligned}$$

Repeating this, we have

$$K(a,b) = K(a_1,b_1) = \ldots = K(a_n,b_n) = \ldots$$

By continuity of the integral K in its arguments, we have

$$K(a,b) = K\left(\mathrm{AG}(a,b), \mathrm{AG}(a,b)\right) = \frac{\pi}{2\mathrm{AG}(a,b)}$$

Thus

$$\mathrm{AG}(a,b) \int_0^{\pi/2} \frac{d\theta}{\sqrt{a^2\cos^2\theta + b^2\sin^2\theta}} = \frac{\pi}{2} \qquad (1)$$

This was used to compute elliptic integrals by Gauss.

A second relation of Gauss relates the arithmetic-geometric mean $\mathrm{AG}(a,b)$ to complete elliptic integrals of the second kind. These integrals are

$$E(a,b) = \int_0^{\pi/2} \sqrt{a^2\cos^2\theta + b^2\sin^2\theta}\,d\theta$$

Since the elliptic integrals satisfy the homogeneity relations

$$K(\lambda a, \lambda b) = \frac{1}{\lambda}K(a,b)\,, \quad E(\lambda a, \lambda b) = \lambda E(a,b)$$

the variables can be normalised to $a = 1$. There is a relation between these due to Legendre (1752–1833). For $0 < x < 1$ and $0 < y < 1$, where $x^2 + y^2 = 1$,

$$K(1,x)E(1,y) + K(1,y)E(1,x) - K(1,x)K(1,y) = \frac{\pi}{2} \qquad (2)$$

For a proof of this, see [2] and [3]. Using the relation (2), Gauss then proved the following:

$$E(a,b) = \left[a^2 - \sum_{n=0}^{\infty} 2^{n-1}\left(a_n^2 - b_n^2\right)\right] K(a,b) \qquad (3)$$

The details of this are in [3].

Following the presentation in [5], we now derive a formula for π. If in (2), $x = y = 1/\sqrt{2}$, then with $K = K(1, 1/\sqrt{2})$ and $E = E(1, 1/\sqrt{2})$, we have

$$2KE - K^2 = \frac{\pi}{2} \qquad (4)$$

In the Gauss relations (1) and (3), if $a = 1$ and $b = 1/\sqrt{2}$, then

$$K = \frac{\pi}{2\mathrm{AG}(1, 1/\sqrt{2})} \quad \text{and} \quad E = (1 - S)K \tag{5}$$

where

$$S = \sum_{n=0}^{\infty} 2^{n-1} (a_n^2 - b_n^2)$$

From (4) and (5)

$$2K^2(1 - S) - K^2 = \frac{\pi}{2}$$

that is,

$$\frac{\pi^2}{4\left(\mathrm{AG}(1, 1/\sqrt{2})\right)^2}(1 - 2S) = \frac{\pi}{2}$$

giving

$$\pi = \frac{2\left(\mathrm{AG}(1, 1/\sqrt{2})\right)^2}{1 - 2S} \tag{6}$$

But

$$\begin{aligned} 1 - 2S &= 1 - \sum_{n=0}^{\infty} 2^n (a_n^2 - b_n^2) \\ &= 1 - (1 - \frac{1}{2}) - \sum_{n=1}^{\infty} 2^n (a_n^2 - b_n^2) \end{aligned}$$

since $a_0 = a = 1$ and $b_0 = b = 1/\sqrt{2}$. Thus

$$1 - 2S = \frac{1}{2} - \sum_{n=1}^{\infty} 2^n (a_n^2 - b_n^2)$$

Substituting this into (6) we get

$$\pi = \frac{4\left(\mathrm{AG}(1, 1/\sqrt{2})\right)^2}{1 - \sum_{n=1}^{\infty} 2^{n+1} (a_n^2 - b_n^2)}$$

This was discovered by Salamin in 1973 [6] and independently by Brent at the same time [4].

If we now define

$$\pi_n = \frac{4a_{n+1}^2}{1 - \sum_{n=1}^{\infty} 2^{n+1} (a_n^2 - b_n^2)}$$

then from error analysis, it can be shown that π_n converges to π quadratically [2,6]. This means, roughly, that the number of correct digits doubles from one value of π_n to the next.

The Brent-Salamin algorithm was implemented in Japan in 1983 by Y. Kanada, Y. Tamura, S. Yoshino and Y. Ushiro to compute 16,000,000 digits in less than 30 hours.

In recent years the algorithm has been modified by the brothers Jonathan and Peter Borwein (natives of St. Andrews, Scotland, and both at Dalhousie University, Novia Scotia) to obtain iterative algorithms for computing π. Details of these are in [2] and [3]. These algorithms are now being implemented to compute π. In January 1986, D.H. Bailey of the NASA Ames Research Center produced 29,360,000 decimal places using one of the Borwein algorithms iterated 12 times on a Cray-2 supercomputer. A year later, Y. Kanada and his colleagues carried out one more iteration to obtain 134,217,000 places on a NEC SX-2 supercomputer. Earlier this year Kanada computed 201,326,000 digits on a new supercomputer manufactured by Hitachi, requiring only six hours of computing time.

3 Utility

One may ask what is the point of all of this, since about 40 decimal places is all one requires for any application imaginable. One use is that the calculation of π has become a benchmark in measuring the sophistication and reliability of the computers that carry it out. In addition, pursuit of more accurate ways has led researchers into intriguing and unexpected areas of number theory. Finally, a statistical analysis of the first 10,000,000 by Y. Kanada shows that the digits are distributed in a way that is expected from the conjecture that π is normal. This means that the frequency of appearance of each string s of digits of length m is asymptotically equal to 10^{-m} i.e.,

$$\lim_{n \to \infty} \frac{N(s, n)}{n} = 10^{-m}$$

where $N(s,n)$ is the number of occurrences of s in the first n digits of π. Because of this, the digits of π are sometimes used in algorithms to generate sequences of random numbers.

References

[1] P. Beckman, *A History of* π, The Golem Press (1970).

[2] J.M. Borwein and P.B. Borwein, *The arithmetic-geometric mean and fast computation of elementary functions*, SIAM Review 26 (1984), 351–366.

[3] J.M. Borwein and P.B. Borwein, *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*, J. Wiley and Sons Inc., (1986).

[4] R.P. Brent, *Multiple-precision zero-finding methods and the complexity of elementary function evaluation*, in *Analytic Computational Complexity*, ed. J.F. Traub, Academic Press (1976), 151–176.

[5] H. Flanders, *Computing* π, College Mathematics Journal 18 (1987), 230–235.

[6] E. Salamin, *Computation of* π *using arithmetic-geometric mean*, Math. of Computation 30 (1976), 565–570.

[7] D. Shanks and J.W. Wrench, *Calculation of* π *to 100,000 decimals*, Math. of Computation 16 (1962) 76–99.

Mathematics Department
University College
Cork

Topological Equivalents of the Axiom of Choice

S.D. McCartan

Recall that, within the terms of Von Neumann-Bernays-Gödel set theory, one form of the axiom of choice (abbreviated AC) is stated as follows:

If $\{X_i : i \in I\}$ *is a non-empty disjoint family of non-empty sets, then there exists a set* C *such that* $C \cap X_i$ *is a singleton for each* $i \in I$.

The axiom of choice has become virtually indispensable in mathematics since a large number of important results have been obtained from it in almost all branches of the subject without leading to a contradiction. However, although this axiom is consistent with, yet independent of, the other axioms of set theory, its status has long been a source of controversy and not all mathematicians are willing to accept it. Perhaps the principal appeal of the axiom of choice resides in the extensive list of its logical equivalents which exist in apparently disparate areas of mathematics. A fairly comprehensive dossier of these was compiled by the Rubins [4] in 1963.

Most topologists side with the majority of mathematicians, assume the axiom of choice, and do not hesitate to use it whenever necessary. Indeed some would argue that the following proposition (usually known as Tychonoff's theorem) constitutes the single most important result in general topology:

The product of a family of non-empty compact topological spaces is compact.

The point here is that Tychonoff's theorem is logically equivalent to the axiom of choice (see [3]). In this note some other such topological equivalents are introduced.

Classically a topological space (X,τ) is said to be a T_0*-space* (T_1*-space*) if and only if for every pair of distinct points in X there exists a τ-neighbourhood of one which does not contain the other (exist τ-neighbourhoods of each which do not contain the other). Properties like T_0 and T_1, when possessed by a topological space, essentially express a degree of separation enjoyed by the