The approximation of π

D.C. van Leijenhorst

Abstract

This is the condensed version of a lecture given in the Alcom Seminar of April 6, 1993. An overview was given of old and new algorithms for the approximation of π, and their relation to Computer algebra and Complexity theory. Special attention was paid to the Borwein algorithms involving elliptic modular function theory. The talk was based on the paper [17] by the Borweins and partly on their book [14].

1 History before 1950

Anecdotal material can be found in many textbooks, e.g. [6]. In this summary, we mention just a few facts, relevant to the rest of these notes.

I. In antiquity, π was known to approximately 2 decimal places. In Hellenistic times, A. of Syracuse (-287–212) developed his polygon algorithm and found $3\frac{10}{71} < \pi < 3\frac{1}{7}$. By elementary trigonometry, the areas $a_n^{-1}(b_n^{-1})$ of the inscribed (circumscribed) polygon defined by the center of a unit circle and an arc of angle Θ divided in 2^{n-1} equal pieces, satisfies the linearity convergent recursion $a_{n+1} = \sqrt{a_n b_n}$, $b_{n+1} = (a_{n+1} + b_n)/2$ (and similarly for total chord length). Archimedes used a 96-gon; Ludolph Von Ceulen a $\approx 2^{60}$-gon 1800 years later. He found 34 correct places of π. Remarkably, if one replaces $b_{n+1} = (a_{n+1} + b_n)/2$ by $b_{n+1} = (a_n + b_n)/2$, one obtains Gauss' quadratically convergent AGM (to be discussed later).

II. After the advent of calculus many more formulas for π became available (Vietà, Wallis, Gregory, Machin). Formulas like Machin's

$$\pi = 4 \arctan \frac{1}{5} - \arctan \frac{1}{239}$$

together with Gregory's expansion:

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

formed the basis of most of the high precision calculating up to 1973!

In [18] and [19], the remarkable fact is explained that in partial sums of Gregory's very slowly convergent series for $\frac{\pi}{4} = \arctan 1$ only certain single digits are incorrect (underlined):

$$\begin{aligned} \frac{\pi}{2} &\approx 2\left(1 - \frac{1}{3} + \frac{1}{5} - \cdots - \frac{1}{999999}\right) \\ &= 1,5707\underline{8}6326794489\underline{7}61923132\underline{1}1916397\underline{5}\underline{205}209\ldots \end{aligned}$$

$$\text{difference}: \quad 1 \qquad -1 \qquad 5 \qquad -61$$

Using a computer algebra system and Sloane's wellknown book [35], it was discovered that $1, -1, 5, -61, \ldots$ are the *Euler numbers* $E_{2k}; k = 0, 1, 2, \ldots$ In fact it can be proved that

$$\pi = 4 \sum_{k=1}^{N} (-1)^{k-1}/(2k-1) + (-1)^N \sum_{k=0}^{M} \frac{2E_{2k}}{(2N)^{2k+1}} + R(M, N);$$

$$|R(M,N)| \leq 2|\frac{E_{2M}}{(2N)^{2M+1}}|$$

where $\sec z = \sum_{n=0}^{\infty}(-1)^n E_{2n} z^{2n}/n!$.

By this result, Gregory's expansion becomes useful; the authors calculated 5263 digits of π in about 90 min. This would be a fearsome task if one took the expansion without error term (one would have to sum about 10^{5262} terms!) Incidentally, the use of error terms to improve series like Gregory's is well-known to numerical analysis [36].

The last calculations by hand were made in the 40's by Ferguson (808 places). The most blatantly erranous value of pi is 4.0, found by the algorithm of democratic decision in 1897 (Iowa House of Repr., Bill 246, cf. [27]).

2 Some philosophical considerations

Though π occurs everywhere in geometry and analysis, little is known about it. Most results are number theoretic: π is irrational (Lambert 1771), transcendent over $\mathbb{Q}$ (Lindemann following Hermite, 1882; for a one-page proof see Baker's book [5]); not a Liouville number: $|\pi - p/q| > q^{-14.65}$ for large q (Chudnowski & Chudnowski); e^π is transcendent over $\mathbb{Q}$ (Gelfond, 1929) as is $\pi + \ln 2 + \sqrt{2}.ln3$ (A. Baker, plm. 1968) but it is not known if $\pi + e$ or $\ln \pi$ are irrational.

Bailey [4] proved in 1988 that an algebraic equation over $\mathbb{Z}$ of degree at most 8 for these latter and some other constants has average coefficients at least 1.000.000.000. He used an implementation of the celebrated Ferguson/Forcade multidimensional continued fraction algorithm ([25], [26]) on a Cray-2 with very high precision FFT-based arithmetic.

Other results are statistical. Though π has Kolmogorow complexity $O(\log n)$, the digits seem uniformly distributed and in any further statistical sense very irregular; it is surmised that π is "normal" in the sense of E. Borel (1909); see Wagon [38]. Kanada [28] gives many statistical results (occurrence of digits, blocks, gaps; "poker hand tests" etc.) and curiosa on the first 10^9 decimals. E.g., from digit nr. 754619564 there occurs a block 27182818; from digit nr. 904961770 follows ...31415926...; at point 564665206 π reads ...9999999999,...

2.1 Relevance of these calculations?

10^9 digits comprise about 1 Gb of computer memory, about 750 HD floppies or two CD Rom disks. What sense make these calculations? Some possible answers...:

- Mathematical curiosity about patterns and regularity in the digit sequence;
- The programs are powerful benchmarks and hardware tests for new supercomputers (e.g., by Gosper, 1985)
- Nice applications of techniques from complexity theory;
- Construction of fast algorithms for π but also for various functions [20]
- Intriguing complexity questions (bounds for the costs of calculating the first n digits of certain irrationals, etc.)
- Approximation algorithms e.g. those involving elliptic modular functions are of course mathematically extremely interesting;

- Philosophically (Intuitionism? Is Kanada's list of digits a mathematical result or just the outcome of a physical experiment; and does that matter? etc.)
- The alpinistic argument: π is there!

2.2 Connections with Computer Algebra (CA)

- A system like MAPLE contains a high-precision routine for π - so our subject must belong to CA...
- High precision arithmetic with tricks from complexity theory is necessary for CA as well as for the calculations we discuss here.
- The development of algorithms for π is connected to many deep, partly algebraic questions; e.g. studying nontrivial modular equations really requires a CA system [8] As the Borweins say: if only Ramanujan would have had access to MacSyma...
- Interesting further research connected both to CA and π is evidently possible and desirable.

3 π on the computer

For details we again refer to[14]. Important landmarks for this talk are:

- 1949: von Neumann c.s. 2037 digits on the ENIAC.
- 1973: Guilloud/Boucher; 10^6 decimals on a CDC by Machin-like formulas.
- 1976: Brent/Salamin find their quadratic iteration.
- 1983: Kanada c.s. compute $> 1.6.10^7$ digits by the Brent/Salamin method plus FFT multiple precision arithmetic on a HITAC computer.
- 1985: Gosper finds $1.7.10^7$ digits by a Ramanujan type series.
- 1986: Bailey; $2.9.10^7$ digits by the quartic Borwein algorithm.
- 1986-1989: fierce competition between Kanada and others resulting in about 10^9 digits for π.
- 1992: Chudnovski and Chudnovski on their home-made supercomputer are reported to have calculated 2.10^9 digits ... Ramanujan type series [32].

A graphic representation of Kanada [28] shows an almost linear connection between time (1940-1990) and the log of the number of calculated digits of π.

4 An Example: the Brent/Salamin algorithm

Let $a_n \to \alpha$ $(n \to \infty)$ in $\mathbb{R}$. We shall call this a p^{th} *order process* iff there exists a sequence $\{t_n\}_{n\geq 0}$ in $\mathbb{R}$ such that $t_n \to 0$ $(n \to \infty)$

$$\forall_n |a_n - \alpha| \leq t_n$$

$$\exists_{C\in\mathbb{R}} \exists_{N\in\mathbb{N}} \forall_{n>N} \quad t_{n+1} \leq C t_n^p.$$

If $p = 1$, the process is *linear*; if $p = 2$, the process is *quadratic*. Sometimes we shall loosely speak of "the p^{th} order process α". Obviously, $|a_n - \alpha| = \mathcal{O}(c^{-p^n})$ for some $c > 1$.

Easily, if $a_n \to \alpha$, $b_n \to \beta$ both quadratically, then also

$$a_n b_n \to \alpha\beta,\ a_n{+}b_n \to \alpha{+}\beta,\ \frac{a_n}{b_n} \to \frac{\alpha}{\beta}\ (\beta \neq 0),\ ca_n \to c\alpha$$

are all quadratic. Hence it is possible to "do algebra" with the quadratic processes.

4.1 Gauss AGM

(cf. [14]): $0 < b_0 \leq a_0$; $a_{n+1} = a_n{+}b_n$; $b_{n+1} = \sqrt{a_n b_n}$ is an example of a quadratic process. If $c_{n+1} = \dfrac{a_n - b_n}{2}$; $c_0 = \sqrt{a_0^2 - b_0^2}$ then $b_n \leq b_{n+1} \leq a_{n+1} \leq a_n$ and

$$0 \leq c_{n+2} = \frac{1}{2}\frac{c_{n+1}^2}{(\sqrt{a_n} + \sqrt{b_n})^2}.$$

The common limit of a_n and b_n is called $M(a_0, b_0)$.

4.2 Elliptic integrals

Notations: $0 < k < 1$ (sometimes $k = 0, 1$).

$$K(k) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2\theta}},\ E(k) = \int_0^{\pi/2} \sqrt{1 - k^2 \sin^2\theta}\, d\theta,$$

$$k' = \sqrt{1 - k^2},\ E'(k) = E(k'),\ K'(k) = K(k').$$

4.3 General idea of the Brent/Salamin algorithm

a) We shall see that the AGM provides a quadratic process converging to $K'(k)$. We shall take $k = \frac{1}{2}\sqrt{2}$ so that $k' = k$, $K'(k) = K(k)$.

b) The *relation of Legendre* yields a relation between $E(k), K(k)$ and π

c) *King's formula* gives a relation between E, K, and some quadratic process.

From the relations a), b) and c) eliminate E and K. Thus, π is expressed as an algebraic expression in quadratic processes which, by our cryptic remark above, is itself a quadratic process! This is the Brent/Salamin formula (below).

Comment. In this summary, we shall not give the proofs of $a), b), c)$ which can be pieced together using [14]. Some indications:

Ad a) By some clever substitutions $(1/a)K'(b/a)$ appears to be invariant under the *Landen transform*

$$(a, b) \mapsto \left(\frac{a + b}{2}, \sqrt{ab}\right).$$

Taking $b = k, a = 1$ and passing to the limit $M(a, b)$ one then obtains Gauss' result

$$K'(k) = \frac{\pi/2}{M(1, k)}.$$

Ad b) With $E = E(k)$ etc., Legendre's relation([14],[39]) reads

$$EK'{+}E'K{-}KK'{-}\frac{\pi}{2} = 0$$

($2EK-K^2 = \pi/2$ if $k = \sqrt{2}/2$). The proof is by expressing E and K as hypergeometric series (expanding $(1-k^2\sin^2\theta)^{\pm\frac{1}{2}}$ as a series in $\sin\theta$ and integrating term-by-term) and then using the hypergeometric differential equations. E.g.,

$$K(k) = \frac{\pi}{2}F(\frac{1}{2},\frac{1}{2};1;k^2) = \sum_{n=0}^{\infty}\frac{(\frac{1}{2})_n(\frac{1}{2})_n}{(1)_n n!}k^{2n}$$

where $(a)_n \equiv a(a+1)\dots(a+n-1)$.

Ad c) King's formula reads

$$E(k) = (1-\sum_{n=0}^{\infty}2^{n-1}c_n^2)K(k)$$

where the c_n belong to the AGM with $a = 1, b = k$. It is found by investigating the behaviour of E (instead of K) under the Landen transform. $\sum_{n=0}^{\infty} 2^{n-1}c_n^2$ converges quadratically (easy from $c_{n+1} \leq cst \cdot c_n^2$). After elimination of the elliptic integrals, we obtain

$$\begin{aligned}\pi &= \frac{2M(1,\frac{1}{2}\sqrt{2})^2}{1-\sum_{n=0}^{\infty}2^n c_n^2}; \quad \text{i.e.}\\ \pi_n &\rightarrow \pi \text{ quadratically, where}\\ \pi_n &= \frac{2a_{n+1}^2}{1-\sum_{k=0}^{n}2^k c_k^2}.\end{aligned}$$

(Brent/Salamin, 1976; used by Kanada. Cf. [14], the error in π_n is $\mathcal{O}(2^n e^{-\pi 2^{n+1}})$.)

5 Some remarks on complexity theory.

How fast are these algorithms? For example, in each step of the AGM, square roots are taken; is this a slow process?

The standard model for computer arithmetic is the *bit operation model* [1]. In this model, the (sequential) time complexity of a calculation on n-bit numbers is measured as the number of logical gates in a combinational dyadic circuit performing that calculation.

It is easy to show that addition of n-bit numbers takes time $6n-4 = \mathcal{O}(n)$, and multiplication $\mathcal{O}(n^2)$ if we perform both in the classical way.
Let $M(n)$ be the optimal complexity of multiplication of n-bit numbers. The upper bound $M(n) = \mathcal{O}(n\log n\log\log n)$ is known (Schönhage/Strassen, using the Fast Fourier tranform). $M(n)$ is most important, since one can prove that "elementary" calculations (e.g. division) on $\mathcal{O}(n)$ bit numbers in precision $\mathcal{O}(n)$ bits, all cost $\mathcal{O}\big(M(n)\big)$.

The same holds for the calculation of algebraic functions like $\sqrt{x}$ (avoiding singularities). The reason is, that these calculations all can be performed by Newtons tangent method. This quadratic process is "self-correcting" in the sense that half precision calculations yield nearly full precision results. This feature is not shared by our 2^{nd} order processes to approximate π. These all run in time $\mathcal{O}\big(M(n)\log n\big)$.

Elementary functions like e^x and $\log x$ also have complexity $\mathcal{O}\big(M(n)\log n\big)$. Interestingly, the best algorithm for the log are based on the AGM. This is because

$K'(k)$ diverges logarithmically near $k = 0$.
In fact:

$$\left|\frac{2}{\pi}\log x - \frac{1}{M(1,10^{-n})} + \frac{1}{M(1,x10^{-n})}\right| < n\,10^{-2n+2} \quad (n > 3, \frac{1}{2} \leq x \leq 1)$$

The multiple precision arithmetic of the Schönhagen/Strassen algorithm is not easy to implement. In this famous calculations, Kanada uses an ordinary complex (floating point) FFT from the standard software of his supercomputer. He mentions the advantages

- no binary to decimal conversion necessary
- greater speed w.r.t. integer operations.

A drawback is the absolute necessity of a (theoretical or experimental) error analysis. On the contrary, Bailey used modular FFT arithmetic mod 3 large prime numbers.

Many highly interesting questions remain; for example the existence of non-algebraic functions of bit complexity $\mathcal{O}\big(M(n)\big)$, or self-correcting algorithms for π [15].

6 Cotangent expansions and Machin-type formulas

For centuries, Machin's method was the most important algorithm for the calculation of π. It is based upon the identity

$$\frac{\pi}{4} = 4\,\mathrm{arccotg}\,5 - \mathrm{arccotg}\,239 \qquad \text{Machin}$$

(where $\mathrm{arccotg}\,x = \frac{1}{x} - \frac{1}{3x^3} + \frac{1}{5x^5} - \cdots$).
Many generalizations have been found and used; e.g.

$$\frac{\pi}{4} = 5\ \mathrm{arccotg}\,7 + 2\ \mathrm{arccotg}\,26 - 2\ \mathrm{arccotg}\,2057 \qquad \text{Stormer 1896}$$

$$\frac{\pi}{4} = 4\ \mathrm{arccotg}\,5 - 2\ \mathrm{arccotg}\,478 + 2\ \mathrm{arccotg}\,131836323 + \mathrm{arccotg}\,318281039 \qquad \text{Shibata 1982}$$

A desperate attempt to get large numbers behind the arccotg:

$$\begin{aligned}\frac{\pi}{4} = {} & 2805\ \mathrm{arccotg}\,5257 - 398\ \mathrm{arccotg}\,9466 \\ & +1950\ \mathrm{arccotg}\,12943 + 1850\ \mathrm{arccotg}\,34208 \\ & +2021\ \mathrm{arccotg}\,44179 + 2097\ \mathrm{arccotg}\,85353 \\ & +1484\ \mathrm{arccotg}\,114669 + 1389\ \mathrm{arccotg}\,330182 \\ & +808\ \mathrm{arccotg}\,485298 \qquad \text{Gauss, cf. [29]}\end{aligned}$$

The study of these identities leads to various interesting problems in Diophantine analysis and approximation theory. Much of the relevant research was done in the 30s and 40s by Lehmer and Todd [37]. Verification of these identities is easy, though laborious. Indeed, for small $\varepsilon_j > 0$,

$$\arg\prod_{j=1}^{n}(1+i\varepsilon_j) = \sum_j \mathrm{arctg}\ \varepsilon_j.$$

Take $\varepsilon_j = 1/m_j$, $m_j \in \mathbb{Z}$. Then

$$\sum_j \operatorname{arccotg} m_j = \operatorname{arccotg}\left(\frac{\sum_j (-1)^j S_{n-2j}(\underline{m})}{\sum_j (-1)^j S_{n-2j-1}(\underline{m})}\right).$$

Here, $S_i(\underline{m})$ is the i^{th} elementary symmetric function in $\underline{m} = (m_1, m_2, \ldots)$.

Lehmer [30] studied identities of the form

$$\frac{k\pi}{4} = \sum_{i=1}^{n} a_i \operatorname{arccotg} m_i$$

under the complexity measure $\sum_{i=1}^{n} 1/\log m_i$. Some of his results are:

- *Transformations* like

  $$\operatorname{arccotg} x = 2 \operatorname{arccotg} 2x - \operatorname{arccotg}(4x^3 + 3x)$$

  are of little use in order to reduce complexity, viz. only for $x < 6.6760135\ldots$

- *Formulas* like

  $$\frac{\pi}{4} = 8 \operatorname{arccotg} 10 - \operatorname{arccotg} 100 - 2 \operatorname{arccotg} 1000 + \cdots$$

  are simpler, but better dan Gauss' (in 10-adic arithmetic!)

In his paper [29] *Lehmer* considered the repeated iteration of a function $f(x,y) \mapsto f(x_0, f(x_1, f(x_2, \ldots)))$. In this way

$$\begin{array}{rcl} x+y & \longrightarrow & \sum x_i \\ xy & \longrightarrow & \Pi x_i \\ x+y^{-1} & \longrightarrow & \text{regular continued fraction} \\ x+yc^{-1} & \longrightarrow & \sum x_i c^{-i} \text{ power series, or decimal number.} \end{array}$$

His new example

$$f(x,y) = \operatorname{cotg}(\operatorname{arccotg} x - \operatorname{arccotg} y) = \frac{xy+1}{y-x}$$

leads to

$$\operatorname{cotg}(\operatorname{arccotg} x_0 - \operatorname{arccotg} x_1 + \operatorname{arccotg} x_2 - \cdots).$$

A *regular expansion* satisfies

$$\begin{array}{ll} \forall_i & x_i \in \mathbb{N} \cup \{0\} \\ \forall_i & x_{i+1} \geq x_i^2 + x_i + 1 \quad (> \text{ at the last term of a finite expansion}) \end{array}$$

Then it can be proved that

- Any $x \in \mathbb{R}, x > 0$ possesses a unique regular expansion; finite if $x \in \mathbb{Q}$

  $$\xi_0 = x,\ x_i = \lfloor \xi_i \rfloor,\ \xi_{i+1} = \frac{\xi_i x_i + 1}{\xi_i - x_i}.$$

  E.g.

  $$\frac{65}{37} = \operatorname{cotg}(\operatorname{arccotg} 1 - \operatorname{arccotg} 3 + \operatorname{arccotg} 18 - \operatorname{arccotg} 603).$$

- The ith term is $< x_0^{-2^i}$! (second order convergence, better than continued fractions).

- The unique number $\xi \in \mathbb{R}$ with $x_{i+1} = x_i^2 + x_i + 1$ is not of degree 2 or 3 over $\mathbb{Q}$.

Finally, Lehmer asked for "nice" regular expansions for expressions involving π, e, $\sqrt{2}$, ... (like the continued fraction $(3-e)/(e-1) = 1/6+1/10+1/14+\cdots+1/(4i+2)+\cdots$). He mentions

$$2+\sqrt{2} = \operatorname{cotg}(\operatorname{arccotg} 3 - \operatorname{arccotg} 17 + \operatorname{arccotg} 99 - \cdots)$$

which obeys $x_i = 6x_{i-1} - x_{i-2}$ but is not regular.

John Todd [37] considered the equation $\operatorname{arctg} n = \sum_r f_r \operatorname{arctg} n_r$ over $\mathbb{Z}$. If $\forall_r n_r < n : n$ is called "reducible". He proves the

Theorem *n is reducible $\Longleftrightarrow P(n)$, the largest prime factor of $1+n^2$, is $\leq 2n-1$.*

The proof is elementary and uses the Gaussian integers. Indeed, if $\forall_r f_r = \pm 1$ then one has

$$\frac{\Pi_r(1 \pm i n_r)}{1 + in} \in \mathbb{Q}$$

Todd also developes a reduction algorithm and gives lists of reductions.

In another article with Chowla [22], Todd proves that the numbers n with $P(n) < 2\sqrt{n}$ have density $> 1 - \log 2 \sim 0.3069$ and conjectures that the reducible n (with $P(n^2+1) < 2n$) also have density about 0.3.

7 The Borwein Algorithms

The p-th order Borwein algorithms are remarkable since their form is very simple, though their derivation isn't at all so! In this summary we shall try to explain the underlying ideas, leaving out most of the calculations.

The main ingredients of their construction are:

1. The theta-function solution of the AGM and, thereby, the expression of elliptic integrals in terms of theta-functions.

2. The possibility to express various functions $f(q)$ (related to theta functions) in their values $f(q^p)$ by means of an algebraic relation over $\mathbb{Q}$ between them.

3. The fact that algebraic functions can be computed quickly.

4. The existence of a function $\alpha(q)$ approximating π whose values in q and q^p are also algebraically related; leading to a fast recursive approximation of π by numbers $\alpha(q_0^i)$.

We shall now discuss these points briefly. Most of the material is taken from [14].

7.1 Theta and modular functions

The *theta functions*, defined for $q \in \mathbb{C}, |q| < 1$,

$$\theta_2 = \theta_2(q) = \sum_{n=-\infty}^{\infty} q^{(n+\frac{1}{2})^2} = 2q^{1/4}\prod_{n=1}^{\infty}(1-q^{2n})(1+q^{2n})^2$$
$$\theta_3 = \theta_3(q) = \sum_{n=-\infty}^{\infty} q^{n^2} = \prod_{n=2}^{\infty}(1-q^{2n})(1+q^{2n-1})^2$$
$$\theta_4 = \theta_4(q) = \sum_{n=-\infty}^{\infty} (-1)^n q^{n^2} = \prod_{n=1}^{\infty}(1-q^{2n})(1-q^{2n-1})^2$$

are specializations of the Jacobian theta functions $\theta_i(z,q)$ which are important in the analysis of doubly periodic complex functions, such as the Weierstrass function $\wp(z)$.

In this theory many beautiful relations between the θ_i have been derived. The most important to our purpose are

$$\theta_3^4 = \theta_4^4 + \theta_2^4 \quad (7.1)$$
$$\theta_3^2 + \theta_4^2 = 2\theta_3^2(q^2) \quad (7.2)$$
$$\theta_3\theta_4 = \theta_4(q^2) \quad (7.3)$$

By elementary monotonicity arguments it is not difficult to show that for $0 < k < 1$ $\exists!_q \quad k = k(q) = \theta_2^2(q)/\theta_3^2(q)$. Then, in our former notation, $k' = \theta_4^2(q)/\theta_3^2(q)$ (by (7.1)). Also, ...(7.2) and ...(7.3) are the AGM with $a_n = \theta_3^2(q^{2^n}), b_n = \theta_4^2(q^{2^n})$! Hence

$$M(1,k') = \theta_3(q)^{-2},\ K(k) = \frac{\pi}{2}\theta_3^2(q)$$

for $k = k(q)$ (Gauss).

A second important property, easy to derive by Poisson summation, is

$$\sqrt{s}\theta_3(e^{-\pi s}) = \theta_3(e^{-\pi/s})$$

and similarly

$$k(e^{-\pi s}) = k'(e^{-\pi/s}).$$

Hence, for $k = k(q), q = e^{-\pi s}$,

$$\frac{M(1,k')}{M(1,k)} = \frac{K'(k)}{K(k)} = s.$$

So

$$q = e^{-\pi K'(k)/K(k)}.$$

7.2 Modular functions

The Borwein algorithms are based on algebraic relations *(modular equations)* over $\mathbb{Q}$ (or $\mathbb{Z}$) between functions θ_i in the points q and q^m, for some $m \in \mathbb{N}$.
Many such relations are constructed in a standard way in the theory of complex multiplication ([33],[10]). For example, let $q = e^{\pi i t}$ and

$$\lambda(t) = k^2(q) = \theta_2(q)^4/\theta_3(q)^4 = 16q\prod_{n=1}^{\infty}\left\{\frac{1+q^{2n}}{1+q^{2n+1}}\right\}^8.$$

Then $\lambda(t{+}2) = \lambda(t), \lambda(-1/t) = 1{-}\lambda(t)$ (by Poisson) hence λ is invariant under the group

$$G(\lambda) = \left\{t \to \frac{at+b}{ct+d} \mid a,b,c,d \in \mathbb{Z}, ad{-}bc = 1, a,d \text{ odd }; b,c \text{ even}\right\}$$

which leaves $H = \{t \mid \text{Im } t > 0\}$ invariant.
If

$$C = \{t \to \frac{t+2i}{p} \mid 0 \le i \le p{-}1\} \cup \{t \to pt\}$$

then ([14], p. 121)

$$W_p = \prod_{B \in \mathcal{B}} \Big(x{-}\lambda\big(B(t)\big)\Big)$$

is in $\mathbb{Z}[x,\lambda]$ and is invariant under $G(\lambda)$.

Analogously, modular equations in Klein's Absolute Invariant

$$J = \frac{4}{27}(1{-}\lambda{+}\lambda^2)^3/\lambda^2(1{-}\lambda)^2$$

can be constructed.

The coefficients of modular equations (even of low degree) are often quite large. Manipulating them demands advanced computer algebra systems like MACSYMA, as does the verification of Ramanujans more exotic relations like

$$\frac{\theta_3^4(q^3)}{\theta_3^4(q^9)} = \left\{\frac{\theta_3(q)}{\theta_3(q^9)} - 1\right\}^3$$

(Berndt,[8]).

A remarkable connection with classical algebra is the fact that the general quintic equation can be solved in terms of θ-functions, using a modular equation with $m = 5$ and Galois group A_5 (Hermite, Kronecker).

Note that W_p is zero for $x = \lambda(q^p)$ so, indeed, one has found an algebraic relation over $\mathbb{Q}$ between $k^2(q) = \theta_2^4(q)/\theta_3^4(q)$ and $k^2(q^p)$. In fact, let $k = k(q),\ l = k(q^{1/p})$ then $\varphi(k,l) \equiv W_p(k^2,l^2) = 0$.

The *Multiplier* is defined as

$$M_p = \frac{K(k)}{K(l)} = \frac{\theta_3(q)^2}{\theta_3(q^{1/p})^2}.$$

Using W_p, it is not difficult to show that $M_p \in \mathbb{Z}(k,l)$.

Example. In [17] the Borweins and Bailey indicate the following p^{th} order algorithm for $K(k)/\pi$. Just take

$$\theta_3^2(q) = \frac{\theta_3^2(q)}{\theta_3^2(q^p)} \cdot \frac{\theta_3^2(q^p)}{\theta_3^2(q^{p^2})} \cdots = \frac{1}{M_p(k_0,k_1)} \cdot \frac{1}{M_p(k_1,k_2)} \cdots$$

To understand this, note that

- W_p is known; so k_{i+1} follows from k_i by solving an algebraic equation (this can be done fast, as we have seen).
- M_p is known.
- $$\theta_3^2(q){-}\prod_{j \le i} \frac{1}{M_p(k_{j-1},k_j)} = \theta_3^2(q)\left(1{-}\frac{1}{\theta_3^2(q^{p^i})}\right) = \mathcal{O}(q^{p^i}).$$

Unfortunately, this is not (yet) an algorithm for the approximation of π. However, we shall see that the same ideas will apply.

7.3 The approximation function. Ramanujan's Singular Values

Let $q = e^{-\pi K'/K} = e^{i\pi t} = e^{-\pi\sqrt{r}},\ t \in H, r > 0$. The final component of the p^{th} order algorithms for π is an approximation function (which will be written as a function of r instead of q):

$$\alpha(r) = \frac{E'(k)}{K(k)} - \frac{\pi}{4K^2(k)}.$$

It is called the "2^{th} singular value"; according to the Borweins it is already implicit in Ramanujan's work. α can be expressed in θ-functions:

$$\alpha(r) = \frac{\frac{1}{\pi} - 4\sqrt{r}\frac{\sum_{-\infty}^{\infty}(-1)^{n^2} n q^{n^2}}{\sum_{-\infty}^{\infty}(-1)^{n^2} q^{n^2}}}{(\sum_{-\infty}^{\infty} q^{n^2})^4}.$$

Hence,

$$\alpha(r) - \frac{1}{\pi} = \mathcal{O}(\sqrt{r}e^{-\pi\sqrt{r}}).$$

Some properties of α:

$$\alpha(r^{-1}) = \frac{1}{\sqrt{r}} - \alpha(r)\sqrt{r} \quad \text{(by Poisson)}$$

and

$$\frac{\pi}{4} = K(k)\left(\sqrt{r}E(k) - \left(\sqrt{r} - \alpha(r)\right)K(k)\right)$$

(a form of Legendre's identity).

The behaviour of α under the transformation $q \to q^p$ (or $r \to p^2 r$) is expressed by the formula:

$$\alpha(p^2 r) = \frac{\alpha(r)}{M_p^2} - \sqrt{r}\epsilon_p(k_0, k_1)$$

with

$$k_i = k(q^{p^i}),\ M_p = M_p(k_0, k_1),\ \epsilon_p(k_0, k_1) = k_0^2 M^2 - pk_1^2 + \frac{pk_1'^2 k_1}{M_p}\frac{\partial M_p}{\partial k_0}.$$

Again, we shall not give the full derivation but note that the formula can be found by differentiation of $K(k_1) = M_p K(k_0)$ w.r.t. k_0 and using the various relations between E and K, e.g.:

- $\dot{K}$ can be expressed in E and K;
- the above form of Legendre's identity.

The approximation function α has the remarkable property that for $r \in \mathbb{Q}, \alpha(r)$ is algebraic over $\mathbb{Q}$. E.g., choose $r = p^{-1}$; then $k_0 = k_1'$ so

$$\begin{cases} W_p(k'^2, k_1^2) = 0 \\ k_1'^2 = 1 - k_1^2 \end{cases}$$

Hence k_1 and k_1' are algebraic. Also, M_p was rational in k_0, k_1; so from the formula for $\alpha(p^2 r)$ it follows that $\alpha(p)$ is algebraic!
In this way $\alpha(p)$ can be calculated. Many values of α are known; e.g.

$$\alpha(12) = 264 + 154\sqrt{3} - 188\sqrt{2} - 108\sqrt{6}.$$

The above results lead to an iteration, the "General Borwein Algorithm":

- Let

$$\begin{array}{rcll} \alpha_0 & = & \alpha(r); & k_0 = k(q) \\ \alpha_n & = & \alpha(p^{2n}r); & k_n = k(q^{p^n}) \end{array}$$

- For $n > 0$ compute k_{n+1} from k_n by solving

$$W_p(k_n^2, k_{n+1}^2) = 0$$

quickly, by Newton's method.

- The reduction between $\alpha(p^2r)$ and $\alpha(r)$ yields a recursion

$$\alpha_{n+1} = m_n^2\alpha_n - p^n\sqrt{r}\varepsilon_n$$

where m_n and ε_n are complicated, but known expressions in k_n and k_{n+1}. By the approximation property of α one has:

$$\alpha_n - \pi^{-1} = \mathcal{O}(p^n\sqrt{r}e^{-\pi p^n\sqrt{r}})$$

hence for $rp^{2n} \geq 1$ there is p^{th} *order convergence*!

- Of course, r must be chosen in such a way that the initial value $\alpha_0 = \alpha(r)$ is known.

Examples

$$\begin{aligned} &p = 2, \\ &\alpha_{n+1} = (1 + k_{n+1})^2\alpha_n - 2^{n+1}\sqrt{r}k_{n+1}, \\ &k_{n+1} = \frac{1 - k'_n}{1 + k'_n} = \frac{1 - \sqrt{1 - k_n^2}}{1 + \sqrt{1 - k_n^2}}. \end{aligned}$$

The convergence is quadratic! Taking two steps at once, one obtains the famous *Quartic Algorithm*

$$\begin{aligned} &\alpha_0 = \alpha(r),\ y_0 = \sqrt{k(q)}, \\ &y_{n+1} = \frac{1 - \sqrt[4]{(1 - y_n^4)}}{1 + \sqrt[4]{(1 - y_n^4)}}, \\ &\alpha_{n+1} = (1 + y_{n+1})^4\alpha_n - 4^{n+1}\sqrt{r}\ y_{n+1}(1 + y_{n+1} + y_{n+1}^2), \\ &r = 4 : \alpha(4) = 6 - 4\sqrt{2}, k(q) = \sqrt{2} - 1. \end{aligned}$$

For $p = 5$ one has:

$$\begin{aligned} &\alpha(25r) = s^2\alpha(r) - \sqrt{r}\left\{\frac{s^2 - 5}{2} + \sqrt{s(s^2 - 2s + 5)}\right\}, \\ &s = 1/M_5(k_0, k_1). \end{aligned}$$

a quintic algorithm; etc.!

All these algorithms obviously have bit complexity $\mathcal{O}(M(n)\ \log n)$. However, the "constants within the $\mathcal{O}$-symbol" will differ. In fact Kanada prefers Brent/Salamin above the quartic Borwein algorithm for this reason.

8 Ramanujan's Series

To conclude our presentation of important methods for the approximation of π, we mention the spectacular series

$$\Pi^{-1} = \frac{\sqrt{8}}{9801}\sum_{n=0}^{\infty}\frac{(4n)!}{(n!)^4}\frac{(1103+26390n)}{(396)^{4n}} \qquad \text{(Ramanujan)}$$

This series has been used by Gesper in 1985 (though a strict correctness proof was not yet known). The following series (due to Chudnowski's [23] and Borwein's [16])

$$\pi^{-1} = \frac{6541681608}{(640320)^{3/2}}\sum_{n=0}^{\infty}\left(\frac{13591409}{545140134}+n\right)\frac{(6n)!)}{(3n)!(n!)^3}\frac{(-1)^n}{(640320)^{3n}}$$

is used in the computer algebra system MAPLE (presumably because MAPLE possesses a good standard package for hypergeometric functions). Many more series of this type, e.g.

$$\pi^{-1} = \sum_{n=0}^{\infty}\frac{(\frac{1}{6})_n(\frac{1}{2})_n(\frac{5}{6})_n}{2\sqrt{3}(n!)^3}(-1)^n \cdot \frac{((212175710912\sqrt{61}+1657145277365)+(13773980892672\sqrt{61}+107578229802750)n)}{(451720356265155784716800 0+5783686501836674471040 00\sqrt{61})^{n+\frac{1}{2}}}$$

are given in [16]. Such series are found in the following way:

- As we have seen, K and E are expressible as hypergeometric series (h.g.).
- By Clausen's Identity, the square of a h.g. is a "generalized h.g.". E.g.

  $$K(k)^2 \cdot \frac{4+4k^2}{\pi^2} = F(\frac{1}{4},\frac{3}{4},\frac{1}{2};1,1;\gamma) \overset{D}{=} \sum_{n\geq 0}\frac{(\frac{1}{4})_n(\frac{3}{4})_n(\frac{1}{2})_n}{(1)_n(1)_n n!}\gamma^{2n}$$

  where $\gamma = (4kk^{'2})/(1+k^2)^2$
- Now substitute θ-functions in γ! One has $k = k(e^{-\pi\sqrt{r}})$; take $r = N$.

Now $\alpha(N)$ and $k(e^{-\pi\sqrt{N}})$ are algebraic over $\mathbb{Q}$, and γ can be expressed in terms of these. For example, taking

$$\alpha(58) = \left(\frac{\sqrt{29}+5}{2}\right)^6 (99\sqrt{29}-444)(99\sqrt{2}-70-13\sqrt{29})$$

and noting that $W_{58}(k_{58}^2, 1-k_{58}^2) = 0$ one obtains Ramanujan's series.
More generally, π is approximated as

$$\pi^{-1} = A\sum_{n\geq 0}\frac{(q_1)_n(q_2)_n(q_3)_n}{(n!)^3}(B+Cn)D^n$$

with $q_1, q_2, q_3 \in \mathbb{Q}$ and A, B, C, D algebraic numbers of small degree (quadratic or quartic).

The convergence of all these series is only linear. Yet they are beautiful and very useful in practice.

References

[1] A.V. Aho, J.E. Hopfcroft and J.D. Ullman, "The Design and Analysis of Computer Algorithms", Addison-Wesley, 1974.

[2] G.E. Andrews e.a. (eds.), "Ramanujan Revisited", Acad. Press, 1988.

[3] D.H. Bailey, "The Computation of π to 29,360,000 Decimal Digits Using Borweins' Quartically Convergent Algorithm", Math. of Computation Vol. 50 (1988) pp 283–296.

[4] D.H. Bailey, "Numerical Results on the Transcendence of Constants Involving π, e, and Euler's Constant", Math. of Computation 50 (1988) pp 275–281.

[5] A. Baker, "Transcendental Number Theory", Cambridge Univ. Press, 1975.

[6] P. Beckmann, "A History of Pi", Golem Press, 1971.

[7] R. Bellman, "A Brief introduction to Theta Functions"; Holt, Rinehart and Winston, Inc., 1961.

[8] B.C. Berndt, "Ramanujan's Notebooks, Parts I,II,III; Springer, 1985/89/91.

[9] F. van der Blij, "Modulaire vormen en getaltheorie"; Lecture Notes 1972 by D.C. van Leijenhorst.

[10] A Borel e.a., "Seminar on Complex Multiplication", Springer Lect. Notes in MAth. 26 (1966).

[11] J.M. Borwein and P.B. Borwein, "The Arithmetic-Geometric Mean and Fast Computation of Elementary Functions", SIAM Rev. 26 (1984) pp. 351–366.

[12] J.M. Borwein and P.B. Borwein, "An Explicit Cubic Iteration for π"; BIT 26 (1985), pp. 123–126.

[13] J.M. Borwein and P.B. Borwein, "More Quadratically Convergent Algorithms for π"; Math. of Computation 46 (1986) pp. 247–253.

[14] J.M. Borwein and P.B. Borwein, "Pi and the AGM" (Wiley and Sons, 1987)

[15] J.M. Borwein and P.B. Borwein, "On the Complexity of Familiar Functions and Numbers", SIAM Rev. vol. 30 nr. 4 (1988) pp. 589–601.

[16] J.M. Borwein and P.B. Borwein, "More Ramanujan–type Series for $1/\pi$"; in G.E. Andrews e.a. "Ramanujan's Revisited", Acad. Press, 1988; pp. 359–396.

[17] J.M. Borwein, P.B. Borwein and D.H. Bailey, "Ramanujan, Modular Equations, and Approximations to Pi or How to Compute one Billion Digits of Pi", Am. Math. Monthly 96 (1989) pp. 201–219.

[18] J.M. Borwein, P.B. Borwein, and K. Dilcher, "Pi, Euler Numbers, and Asymptotic Expansions", Am. Math. Monthly 96 (1989) pp. 681–687.

[19] J.M. Borwein, P.B. Borwein, and M.A. Limber, "Maple as a High Precision Calculator", MapleTech Issue 8 (fall 1992) pp. 39–44.

[20] R.P. Brent, "Fast Multiple–Precision Evaluation of Elementary Functions", JACM 23 (1976) pp. 242–251.

[21] R.P. Brent and E.M. McMillan, "Some New Algorithms for High–Precision Computation of Euler's Constant", Math. of Computation 34 (1980) pp. 305–312.

[22] S.D. Chowla and J. Todd, "The Density of Reducible Integers"; Can. J. Math. 1 (1949) pp. 297–299.

[23] D.V. Chudnovski and G.V. Chudnovski, "Approximations and Complex Multiplication According to Ramanujan", in G.E. Andrews e.a. (Eds.), "Ramanujan Revisited", Acad. Press, 1988; pp. 375–472.

[24] A. Erdélyi (ed), "Higher Transcendental Functions", Vol. 1,2,3; McGraw–Hill, 1953–5.

[25] H. Ferguson, "A short Proof of the Existence of Vector Euclidean Algorithms", Proc. Am. Math. Soc., vol. 97 (1986) pp. 8–10.

[26] H. Ferguson, "A Noninductive $GL(n, \mathbb{Z})$ Algorithm That Constructs Integral Linear Relations for $n\mathbb{Z}$-Linearly Dependent Real Numbers", J. Algor. 8 (1987) pp. 131–145.

[27] Guinness Book of World Records, Bantam, 1989, p. 182.

[28] Y. Kanada, "How to Calculate π up to 1 billion Decimal Places", Lecture sheets copies, Delft, 1992.

[29] D.H. Lehmer, "A Cotangent Analogue of Continued Fractions", Duke Math. J., vol. 4 (1938) pp. 323–340.

[30] D.H. Lehmer, "On Arccotangent Relations for π", Am. Math. Monthly 45 (1938) pp. 657–664.

[31] D.C. van Leijenhorst, Lecture Notes "T2,CC1,CC2", KU Nijmegen, Dept. of Theor. Computer Science.

[32] Richard Preston, "The Mountains of Pi", The New Yorker, March 2 1992, ("Profiles", pp. 36 ff).

[33] B. Schoeneberg, "Elliptic Modular Functions", Springer, 1974.

[34] J-P. Serre, "Cours d'Arithmétique", Presses Universitaires de France, 1970.

[35] N.J.A. Sloane, "A Handbook of Integer Sequences", Acad. Press, 1973.

[36] A. van der Sluis, "Dictaat Toegepaste Analyse", pg. 1; Utrecht, Univ. Dept. of Math., 1968.

[37] J. Todd, "A Problem on Arc Tangent Relations", Am. Math. Monthly 8 (1949) pp. 517–528.

[38] S. Wagon "Is π Normal?, The Math. Intelligencer 7 (1985) pp. 65–67.

[39] E.T. Whittaker and G.N. Watson, "A Course of Modern Analysis", Cambridge Univ. Press, 1902 (1965).