The Arithmetic-Geometric Mean and Fast Computation of Elementary Functions

J. M. Borwein; P. B. Borwein

THE ARITHMETIC-GEOMETRIC MEAN AND FAST COMPUTATION OF ELEMENTARY FUNCTIONS*

J. M. BORWEIN† AND P. B. BORWEIN†

Abstract. We produce a self contained account of the relationship between the Gaussian arithmetic-geometric mean iteration and the fast computation of elementary functions. A particularly pleasant algorithm for π is one of the by-products.

Introduction. It is possible to calculate 2^n decimal places of π using only n iterations of a (fairly) simple three-term recursion. This remarkable fact seems to have first been explicitly noted by Salamin in 1976 [16]. Recently the Japanese workers Y. Tamura and Y. Kanada have used Salamin's algorithm to calculate π to 2^{23} decimal places in 6.8 hours. Subsequently 2^{24} places were obtained ([18] and private communication). Even more remarkable is the fact that all the elementary functions can be calculated with similar dispatch. This was proved (and implemented) by Brent in 1976 [5]. These extraordinarily rapid algorithms rely on a body of material from the theory of elliptic functions, all of which was known to Gauss. It is an interesting synthesis of classical mathematics with contemporary computational concerns that has provided us with these methods. Brent's analysis requires a number of results on elliptic functions that are no longer particularly familiar to most mathematicians. Newman in 1981 stripped this analysis to its bare essentials and derived related, though somewhat less computationally satisfactory, methods for computing π and log. This concise and attractive treatment may be found in [15].

Our intention is to provide a mathematically intermediate perspective and some bits of the history. We shall derive implementable (essentially) quadratic methods for computing π and all the elementary functions. The treatment is entirely self-contained and uses only a minimum of elliptic function theory.

1. 3.141592653589793238462643383279502884197. The calculation of π to great accuracy has had a mathematical import that goes far beyond the dictates of utility. It requires a mere 39 digits of π in order to compute the circumference of a circle of radius 2×10^{25} meters (an upper bound on the distance travelled by a particle moving at the speed of light for 20 billion years, and as such an upper bound on the radius of the universe) with an error of less than 10^{-12} meters (a lower bound for the radius of a hydrogen atom).

Such a calculation was in principle possible for Archimedes, who was the first person to develop methods capable of generating arbitrarily many digits of π. He considered circumscribed and inscribed regular n-gons in a circle of radius 1. Using $n = 96$ he obtained

$$3.1405 \cdots = \frac{6336}{2017.25} < \pi < \frac{14688}{4673.5} = 3.1428.$$

If $1/A_n$ denotes the area of an inscribed regular 2^n-gon and $1/B_n$ denotes the area of a circumscribed regular 2^n-gon about a circle of radius 1 then

(1.1) $$A_{n+1} = \sqrt{A_n B_n}, \qquad B_{n+1} = \frac{A_{n+1} + B_n}{2}.$$

---

†Department of Mathematics, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H8.

This two-term iteration, starting with $A_2 := \frac{1}{2}$ and $B_2 := \frac{1}{4}$, can obviously be used to calculate π. (See Edwards [9, p. 34].) A_{15}^{-1}, for example, is 3.14159266 which is correct through the first seven digits. In the early sixteen hundreds Ludolph von Ceulen actually computed π to 35 places by Archimedes' method [2].

Observe that $A_n := 2^{-n} \operatorname{cosec}(\theta/2^n)$ and $B_n := 2^{-n-1} \cotan(\theta/2^{n+1})$ satisfy the above recursion. So do $A_n := 2^{-n} \operatorname{cosech}(\theta/2^n)$ and $B_n := 2^{-n-1} \cotanh(\theta/2^{n+1})$. Since in both cases the common limit is $1/\theta$, the iteration can be used to calculate the standard inverse trigonometric and inverse hyperbolic functions. (This is often called Borchardt's algorithm [6], [19].)

If we observe that

$$A_{n+1} - B_{n+1} = \frac{1}{2(\sqrt{A_n}/\sqrt{B_n} + 1)}(A_n - B_n)$$

we see that the error is decreased by a factor of approximately four with each iteration. This is linear convergence. To compute n decimal digits of π, or for that matter arcsin, arcsinh or log, requires $O(n)$ iterations.

We can, of course, compute π from arctan or arcsin using the Taylor expansion of these functions. John Machin (1680–1752) observed that

$$\pi = 16 \arctan\left(\frac{1}{5}\right) - 4 \arctan\left(\frac{1}{239}\right)$$

and used this to compute π to 100 places. William Shanks in 1873 used the same formula for his celebrated 707 digit calculation. A similar formula was employed by Leonhard Euler (1707–1783):

$$\pi = 20 \arctan\left(\frac{1}{7}\right) + 8 \arctan\left(\frac{3}{79}\right).$$

This, with the expansion

$$\arctan(x) = \frac{y}{x}\left(1 + \frac{2}{3}y + \frac{2.4}{3.5}y^2 + \cdots\right)$$

where $y = x^2/(1 + x^2)$, was used by Euler to compute π to 20 decimal places in an hour. (See Beckman [2] or Wrench [21] for a comprehensive discussion of these matters.) In 1844 Johann Dase (1824–1861) computed π correctly to 200 places using the formula

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

Dase, an "idiot savant" and a calculating prodigy, performed this "stupendous task" in "just under two months." (The quotes are from Beckman, pp. 105 and 107.)

A similar identity:

$$\pi = 24 \arctan\left(\frac{1}{8}\right) + 8 \arctan\left(\frac{1}{57}\right) + 4 \arctan\left(\frac{1}{239}\right)$$

was employed, in 1962, to compute 100,000 decimals of π. A more reliable "idiot savant", the IBM 7090, performed this calculation in a mere 8 hrs. 43 mins. [17].

There are, of course, many series, products and continued fractions for π. However, all the usual ones, even cleverly evaluated, require $O(\sqrt{n})$ operations $(+, \times, \div, \sqrt{\ })$ to arrive at n digits of π. Most of them, in fact, employ $O(n)$ operations for n digits, which is

essentially linear convergence. Here we consider only full precision operations. For a time complexity analysis and a discussion of time efficient algorithms based on binary splitting see [4].

The algorithm employed in [17] requires about 1,000,000 operations to compute 1,000,000 digits of π. We shall present an algorithm that reduces this to about 200 operations. The algorithm, like Salamin's and Newman's requires some very elementary elliptic function theory. The circle of ideas surrounding the algorithm for π also provides algorithms for all the elementary functions.

2. Extraordinarily rapid algorithms for algebraic functions. We need the following two measures of speed of convergence of a sequence (a_n) with limit L. If there is a constant C_1 so that

$$|a_{n+1} - L| \leq C_1 |a_n - L|^2$$

for all n, then we say that (a_n) converges to L *quadratically,* or with *second order.* If there is a constant $C_2 > 1$ so that, for all n,

$$|a_n - L| \leq C_2^{-2^n}$$

then we say that (a_n) converges to L *exponentially.* These two notions are closely related; quadratic convergence implies exponential convergence and both types of convergence guarantee that a_n and L will "agree" through the first $O(2^n)$ digits (provided we adopt the convention that .9999...9 and 1.000...0 agree through the required number of digits).

Newton's method is perhaps the best known second order iterative method. Newton's method computes a zero of $f(x) - y$ by

$$(2.1) \qquad\qquad x_{n+1} := x_n - \frac{f(x_n) - y}{f'(x_n)}$$

and hence, can be used to compute f^{-1} quadratically from f, at least locally. For our purposes, finding suitable starting values poses little difficulty. Division can be performed by inverting $(1/x) - y$. The following iteration computes $1/y$:

$$(2.2) \qquad\qquad x_{n+1} := 2x_n - x_n^2 y.$$

Square root extraction ($\sqrt{y}$) is performed by

$$(2.3) \qquad\qquad x_{n+1} := \frac{1}{2}\left(x_n + \frac{y}{x_n}\right).$$

This ancient iteration can be traced back at least as far as the Babylonians. From (2.2) and (2.3) we can deduce that division and square root extraction are of the same order of complexity as multiplication (see [5]). Let $M(n)$ be the "amount of work" required to multiply two n digit numbers together and let $D(n)$ and $S(n)$ be, respectively, the "amount of work" required to invert an n digit number and compute its square root, to n digit accuracy. Then

$$D(n) = O(M(n))$$

and

$$S(n) = O(M(n)).$$

We are not bothering to specify precisely what we mean by work. We could for example count the number of single digit multiplications. The basic point is as follows. It requires

$O(\log n)$ iterations of Newton's method (2.2) to compute $1/y$. However, at the ith iteration, one need only work with accuracy $O(2^i)$. In this sense, Newton's method is self-correcting. Thus,

$$D(n) = O\left(\sum_{i=1}^{\log n} M(2^i)\right) = O(M(n))$$

provided $M(2^i) \geq 2M(2^{i-1})$. The constants concealed beneath the order symbol are not even particularly large. Finally, using a fast multiplication, see [12], it is possible to multiply two n digits numbers in $O(n \log (n) \log \log (n))$ single digit operations.

What we have indicated is that, for the purposes of asymptotics, it is reasonable to consider multiplication, division and root extraction as equally complicated and to consider each of these as only marginally more complicated than addition. Thus, when we refer to operations we shall be allowing addition, multiplication, division and root extraction.

Algebraic functions, that is roots of polynomials whose coefficients are rational functions, can be approximated (calculated) exponentially using Newton's method. By this we mean that the iterations converge exponentially and that each iterate is itself suitably calculable. (See [13].)

The difficult trick is to find a method to exponentially approximate just one elementary transcendental function. It will then transpire that the other elementary functions can also be exponentially calculated from it by composition, inversion and so on.

For this Newton's method cannot suffice since, if f is algebraic in (2.1) then the limit is also algebraic.

The only familiar iterative procedure that converges quadratically to a transcendental function is the arithmetic-geometric mean iteration of Gauss and Legendre for computing complete elliptic integrals. This is where we now turn. We must emphasize that it is difficult to exaggerate Gauss' mastery of this material and most of the next section is to be found in one form or another in [10].

3. The real AGM iteration. Let two positive numbers a and b with $a > b$ be given. Let $a_0 := a, b_0 := b$ and define

$$(3.1) \qquad a_{n+1} := \frac{1}{2}(a_n + b_n), \qquad b_{n+1} := \sqrt{a_n b_n}$$

for n in $\mathbb{N}$.

One observes, as a consequence of the arithmetic-geometric mean inequality, that $a_n \geq a_{n+1} \geq b_{n+1} \geq b_n$ for all n. It follows easily that (a_n) and (b_n) converge to a common limit L which we sometimes denote by $AG(a, b)$. Let us now set

$$(3.2) \qquad c_n := \sqrt{a_n^2 - b_n^2} \quad \text{for } n \in \mathbb{N}.$$

It is apparent that

$$(3.3) \qquad c_{n+1} = \frac{1}{2}(a_n - b_n) = \frac{c_n^2}{4a_{n+1}} \leq \frac{c_n^2}{4L},$$

which shows that (c_n) converges quadratically to zero. We also observe that

$$(3.4) \qquad a_n = a_{n+1} + c_{n+1} \quad \text{and} \quad b_n = a_{n+1} - c_{n+1}$$

which allows us to define a_n, b_n and c_n for negative n. These negative terms can also be generated by the *conjugate scale* in which one starts with $a_0' := a_0$ and $b_0' := c_0$ and defines

(a'_n) and (b'_n) by (3.1). A simple induction shows that for any integer n

(3.5) $$a'_n = 2^{-n}a_{-n}, \quad b'_n = 2^{-n}c_{-n}, \quad c'_n = 2^{-n}b_{-n}.$$

Thus, backward iteration can be avoided simply by altering the starting values. For future use we define the quadratic *conjugate* $k' := \sqrt{1 - k^2}$ for any k between 0 and 1.

The limit of (a_n) can be expressed in terms of a *complete elliptic integral of the first kind,*

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

In fact

(3.7) $$I(a, b) = \frac{1}{2} \int_{-\infty}^{\infty} \frac{dt}{\sqrt{(a^2 + t^2)(b^2 + t^2)}}$$

as the substitution $t := a \tan \theta$ shows. Now the substitution of $u := \frac{1}{2}(t - (ab/t))$ and some careful but straightforward work [15] show that

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

It follows that $I(a_n, b_n)$ is independent of n and that, on interchanging limit and integral,

$$I(a_0, b_0) = \lim_{n \to \infty} I(a_n, b_n) = I(L, L).$$

Since the last integral is a simple arctan (or directly from (3.6)) we see that

(3.9) $$I(a_0, b_0) = \frac{\pi}{2} AG(a_0, b_0).$$

Gauss, of course, had to derive rather than merely verify this remarkable formula. We note in passing that $AG(\cdot, \cdot)$ is positively homogeneous.

We are now ready to establish the underlying limit formula.

PROPOSITION 1.

(3.10) $$\lim_{k \to 0^+} \left[\log\left(\frac{4}{k}\right) - I(1, k)\right] = 0.$$

Proof. Let

$$A(k) := \int_0^{\pi/2} \frac{k' \sin \theta \, d\theta}{\sqrt{k^2 + (k')^2 \cos^2 \theta}}$$

and

$$B(k) := \int_0^{\pi/2} \sqrt{\frac{1 - k' \sin \theta}{1 + k' \sin \theta}} \, d\theta.$$

Since $1 - (k' \sin \theta)^2 = \cos^2 \theta + (k \sin \theta)^2 = (k' \cos \theta)^2 + k^2$, we can check that

$$I(1, k) = A(k) + B(k).$$

Moreover, the substitution $u := k' \cos \theta$ allows one to evaluate

(3.11) $$A(k) := \int_0^{k'} \frac{du}{\sqrt{u^2 + k^2}} = \log\left(\frac{1 + k'}{k}\right).$$

356 J. M. BORWEIN AND P. B. BORWEIN

Finally, a uniformity argument justifies

$$(3.12) \qquad \lim_{k \to 0^+} B(k) = B(0) = \int_0^{\pi/2} \frac{\cos \theta \, d\theta}{1 + \sin \theta} = \log 2,$$

and (3.11) and (3.12) combine to show (3.10). $\qquad \square$

It is possible to give various asymptotics in (3.10), by estimating the convergence rate in (3.12).

PROPOSITION 2. *For $k \in (0, 1]$*

$$(3.13) \qquad \left| \log\left(\frac{4}{k}\right) - I(1, k) \right| \le 4k^2 I(1, k) \le 4k^2(8 + |\log k|).$$

Proof. Let

$$\Delta(k) := \log\left(\frac{4}{k}\right) - I(1, k).$$

As in Proposition 1, for $k \in (0, 1]$,

(3.14)

$$|\Delta(k)| \le \left| \log\left(\frac{2}{k}\right) - \log\left(\frac{1 + k'}{k}\right) \right| + \left| \int_0^{\pi/2} \left[\sqrt{\frac{1 - k' \sin \theta}{1 + k' \sin \theta}} - \sqrt{\frac{1 - \sin \theta}{1 + \sin \theta}} \right] d\theta \right|.$$

We observe that, since $1 - k' = 1 - \sqrt{1 - k^2} < k^2$,

$$(3.15) \qquad \left| \log\left(\frac{2}{k}\right) - \log\left(\frac{1 + k'}{k}\right) \right| = \left| \log\left(\frac{1 + k'}{2}\right) \right| \le 1 - k' \le k^2.$$

Also, by the mean value theorem, for each θ there is a $\gamma \in [0, k]$, so that

$$0 \le \left[\sqrt{\frac{1 - k' \sin \theta}{1 + k' \sin \theta}} - \sqrt{\frac{1 - \sin \theta}{1 + \sin \theta}} \right]$$

$$\le \left[\sqrt{\frac{1 - (1 - k^2) \sin \theta}{1 + (1 - k^2) \sin \theta}} - \sqrt{\frac{1 - \sin \theta}{1 + \sin \theta}} \right]$$

$$= \left[\frac{\sqrt{1 + (1 - \gamma^2) \sin \theta}}{\sqrt{1 - (1 - \gamma^2) \sin \theta}} \cdot \frac{2\gamma \sin \theta}{(1 + (1 - \gamma^2) \sin \theta)^2} \right] k$$

$$\le \frac{2\gamma k}{\sqrt{1 - (1 - \gamma^2) \sin \theta}} \le \frac{2k^2}{\sqrt{1 - (1 - k^2) \sin \theta}}.$$

This yields

$$\left| \int_0^{\pi/2} \left[\sqrt{\frac{1 - k' \sin \theta}{1 + k' \sin \theta}} - \sqrt{\frac{1 - \sin \theta}{1 + \sin \theta}} \right] d\theta \right| \le 2k^2 \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k' \sin \theta}} \le 2\sqrt{2}\, k^2 I(1, k)$$

which combines with (3.14) and (3.15) to show that

$$|\Delta(k)| \le (1 + 2\sqrt{2})k^2 I(1, k) \le 4k^2 I(1, k).$$

We finish by observing that

$$kI(1, k) \le \frac{\pi}{2}$$

allows us to deduce that

$$I(1, k) \leqq 2\pi k + \log\left(\frac{4}{k}\right). \qquad \square$$

Similar considerations allow one to deduce that

(3.16) $$|\Delta(k) - \Delta(h)| \leqq 2\pi |k - h|$$

for $0 < k, h < 1/\sqrt{2}$.

The next proposition gives all the information necessary for computing the elementary functions from the AGM.

PROPOSITION 3. *The AGM satisfies the following identity (for all initial values):*

(3.17) $$\lim_{n \to \infty} 2^{-n} \frac{a'_n}{a_n} \log\left(\frac{4a_n}{c_n}\right) = \frac{\pi}{2}.$$

Proof. One verifies that

$$\frac{\pi}{2} = \lim_{n \to \infty} a'_n I(a'_0, b'_0) \qquad \text{(by (3.9))}$$

$$= \lim_{n \to \infty} a'_n I(a'_{-n}, b'_{-n}) \qquad \text{(by (3.8))}$$

$$= \lim_{n \to \infty} a'_n I(2^n a_n, 2^n c_n) \qquad \text{(by 3.5)).}$$

Now the homogeneity properties of $I(\cdot, \cdot)$ show that

$$I(2^n a_n, 2^n c_n) = \frac{2^{-n}}{a_n} I\left(1, \frac{c_n}{a_n}\right).$$

Thus

$$\frac{\pi}{2} = \lim_{n \to \infty} 2^{-n} \frac{a'_n}{a_n} I\left(1, \frac{c_n}{a_n}\right),$$

and the result follows from Proposition 1. $\square$

From now on we fix $a_0 := a'_0 := 1$ and consider the iteration as a function of $b_0 := k$ and $c_0 := k'$. Let P_n and Q_n be defined by

(3.18) $$P_n(k) := \left(\frac{4a_n}{c_n}\right)^{2^{1-n}}, \qquad Q_n(k) := \frac{a_n}{a'_n},$$

and let $P(k) := \lim_{n \to \infty} P_n(k)$, $Q(k) := \lim_{n \to \infty} Q_n(k)$. Similarly let $a := a(k) := \lim_{n \to \infty} a_n$ and $a' := a'(k) := \lim_{n \to \infty} a'_n$.

THEOREM 1. *For $0 < k < 1$ one has·*

(a) $P(k) = \exp(\pi Q(k))$,

(3.19) (b) $0 \leq P_n(k) - P(k) \leq \dfrac{16}{1 - k^2}\left(\dfrac{a_n - a}{a}\right)$,

(c) $|Q_n(k) - Q(k)| \leq \dfrac{a'|a - a_n| + a|a' - a'_n|}{(a')^2}$.

Proof. (a) is an immediate rephrasing of Proposition 3, while (c) is straightforward.

To see (b) we observe that

$$(3.20) \qquad P_{n+1} = P_n \cdot \left(\frac{a_{n+1}}{a_n}\right)^{2^{1-n}}$$

because $4a_{n+1} c_{n+1} = c_n^2$, as in (3.3). Since $a_{n+1} \le a_n$ we see that

$$O \le P_n - P_{n+1} \le \left[1 - \left(\frac{a_{n+1}}{a_n}\right)^{2^{1-n}}\right] P_n \le \left(1 - \frac{a_{n+1}}{a_n}\right) P_0,$$

$$(3.21)$$

$$P_n - P_{n+1} \le \left(\frac{a_n - a_{n+1}}{a}\right) P_0$$

since a_n decreases to a. The result now follows from (3.21) on summation.   □

Thus, the theorem shows that both P and Q can be computed exponentially since (a_n) can be so calculated. In the following sections we will use this theorem to give implementable exponential algorithms for π and then for all the elementary functions.

We conclude this section by rephrasing (3.19a). By using (3.20) repeatedly we derive that

$$(3.22) \qquad P = \frac{16}{1 - k^2} \prod_{n=0}^{\infty} \left(\frac{a_{n+1}}{a_n}\right)^{2^{1-n}}.$$

Let us note that

$$\frac{a_{n+1}}{a_n} = \frac{a_n + b_n}{2a_n} = \frac{1}{2}\left(1 + \frac{b_n}{a_n}\right),$$

and $x_n := b_n/a_n$ satisfies the one-term recursion used by Legendre [14]

$$(3.23) \qquad x_{n+1} := \frac{2\sqrt{x_n}}{x_n + 1} \qquad x_0 := k.$$

Thus, also

$$(3.24) \qquad P_{n+1}(k) = \frac{16}{1 - k^2} \prod_{j=0}^{n+1} \left(\frac{1 + x_j}{2}\right)^{2^{1-j}} = \left(\frac{1 + x_n}{1 - x_n}\right)^{2^{-n}}.$$

When $k := 2^{-1/2}$, $k = k'$ and one can explicitly deduce that $P(2^{-1/2}) = e^{\pi}$. When $k = 2^{-1/2}$ (3.22) is also given in [16].

4. Some interrelationships. A centerpiece of this exposition is the formula (3.17) of Proposition 3.

$$(4.1) \qquad \lim_{n \to \infty} \frac{1}{2^n} \log\left(\frac{4a_n}{c_n}\right) = \frac{\pi}{2} \lim_{n \to \infty} \frac{a_n}{a_n'},$$

coupled with the observation that both sides converge exponentially. To approximate $\log x$ exponentially, for example, we first find a starting value for which

$$\left(\frac{4a_n}{c_n}\right)^{1/2^n} \to x.$$

This we can do to any required accuracy quadratically by Newton's method. Then we compute the right limit, also quadratically, by the AGM iteration. We can compute exp analogously and since, as we will show, (4.1) holds for complex initial values we can also get the trigonometric functions.

There are details, of course, some of which we will discuss later. An obvious detail is that we require π to desired accuracy. The next section will provide an exponentially converging algorithm for π also based only on (4.1). The principle for it is very simple. If we differentiate both sides of (4.1) we lose the logarithm but keep the π!

Formula (3.10), of Proposition 1, is of some interest. It appears in King [11, pp. 13, 38] often without the "4" in the log term. For our purposes the "4" is crucial since without it (4.1) will only converge linearly (like $(\log 4)/2^n$). King's 1924 monograph contains a wealth of material on the various iterative methods related to computing elliptic integrals. He comments [11, p. 14]:

"The limit [(4.1) without the "4"] does not appear to be generally known, although an equivalent formula is given by Legendre (*Fonctions éliptiques,* t. I, pp. 94–101)."

King adds that while Gauss did not explicitly state (4.1) he derived a closely related series expansion and that none of this "appears to have been noticed by Jacobi or by subsequent writers on elliptic functions." This series [10, p. 377] gives (4.1) almost directly.

Proposition 1 may be found in Bowman [3]. Of course, almost all the basic work is to be found in the works of Abel, Gauss and Legendre [1], [10] and [14]. (See also [7].) As was noted by both Brent and Salamin, Proposition 2 can be used to estimate log given π. We know from (3.13) that, for $0 < k \leq 10^{-3}$,

$$\left| \log\left(\frac{4}{k}\right) - I(1, k) \right| < 10k^2 |\log k|.$$

By subtraction, for $0 < x < 1$, and $n \geq 3$,

$$(4.2) \qquad |\log(x) - [I(1, 10^{-n}) - I(1, 10^{-n}x)]| < n\, 10^{-2(n-1)}$$

and we can compute log exponentially from the AGM approximations of the elliptic integrals in the above formula. This is in the spirit of Newman's presentation [15]. Formula (4.2) works rather well numerically but has the minor computational drawback that it requires computing the AGM for small initial values. This leads to some linear steps (roughly $\log(n)$) before quadratic convergence takes over.

We can use (3.16) or (4.2) to show directly that π is exponentially computable. With $k := 10^{-n}$ and $h := 10^{-2n} + 10^{-n}$ (3.16) yields with (3.9) that, for $n \geq 1$,

$$\left| \log(10^{-n} + 1) - \frac{\pi}{2}\left[\frac{1}{AG(1, 10^{-n})} - \frac{1}{AG(1, 10^{-n} + 10^{-2n})} \right] \right| \leq 10^{1-2n}.$$

Since $|\log(x + 1)/x - 1| \leq x/2$ for $0 < x < 1$, we derive that

$$(4.3) \qquad \left| \frac{2}{\pi} - \left[\frac{10^n}{AG(1, 10^{-n})} - \frac{10^n}{AG(1, 10^{-n} + 10^{-2n})} \right] \right| \leq 10^{1-n}.$$

Newman [15] gives (4.3) with a rougher order estimate and without proof. This analytically beautiful formula has the serious computational drawback that obtaining n digit accuracy for π demands that certain of the operations be done to twice that precision.

Both Brent's and Salamin's approaches require *Legendre's relation:* for $0 < k < 1$

$$(4.4) \qquad I(1, k)J(1, k') + I(1, k')J(1, k) - I(1, k)I(1, k') = \frac{\pi}{2}$$

where $J(a, b)$ is the *complete elliptic integral of the second kind* defined by

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

The elliptic integrals of the first and second kind are related by

$$(4.5) \qquad J(a_0, b_0) = \left(a_0^2 - \frac{1}{2} \sum_{n=0}^{\infty} 2^n c_n^2 \right) I(a_0, b_0)$$

where, as before, $c_n^2 = a_n^2 - b_n^2$ and a_n and b_n are computed from the AGM iteration.

Legendre's proof of (4.4) can be found in [3] and [8]. His elegant elementary argument is to differentiate (4.4) and show the derivative to be constant. He then evaluates the constant, essentially by Proposition 1. Strangely enough, Legendre had some difficulty in evaluating the constant since he had problems in showing that $k^2 \log (k)$ tends to zero with k [8, p. 150].

Relation (4.5) uses properties of the ascending Landen transformation and is derived by King in [11].

From (4.4) and (4.5), noting that if k equals $2^{-1/2}$ then so does k', it is immediate that

$$(4.6) \qquad \pi = \frac{[2AG(1, 2^{-1/2})]^2}{1 - \sum_{n=1}^{\infty} 2^{n+1} c_n^2} \, .$$

This concise and surprising exponentially converging formula for π is used by both Salamin and Brent. As Salamin points out, by 1819 Gauss was in possession of the AGM iteration for computing elliptic integrals of the first kind and also formula (4.5) for computing elliptic integrals of second kind. Legendre had derived his relation (4.4) by 1811, and as Watson puts it [20, p. 14] "in the hands of Legendre, the transformation [(3.23)] became a most powerful method for computing elliptic integrals." (See also [10], [14] and the footnotes of [11].) King [11, p. 39] derives (4.6) which he attributes, in an equivalent form, to Gauss. It is perhaps surprising that (4.6) was not suggested as a practical means of calculating π to great accuracy until recently.

It is worth emphasizing the extraordinary similarity between (1.1) which leads to linearly convergent algorithms for all the elementary functions, and (3.1) which leads to exponentially convergent algorithms.

Brent's algorithms for the elementary functions require a discussion of incomplete elliptic integrals and the Landen transform, matters we will not pursue except to mention that some of the contributions of Landen and Fagnano are entertainingly laid out in an article by G.N. Watson entitled "The Marquis [Fagnano] and the Land Agent [Landen]" [20]. We note that Proposition 1 is also central to Brent's development though he derives it somewhat tangentially. He also derives Theorem 1(a) in different variables via the Landen transform.

5. An algorithm for π. We now present the details of our exponentially converging algorithm for calculating the digits of π. Twenty iterations will provide over two million digits. Each iteration requires about ten operations. The algorithm is very stable with all the operations being performed on numbers between $\frac{1}{2}$ and 7. The eighth iteration, for example, gives π correctly to 694 digits.

THEOREM 2. *Consider the three-term iteration with initial values*

$$\alpha_0 := \sqrt{2}, \quad \beta_0 := 0, \quad \pi_0 := 2 + \sqrt{2}$$

given by

$$(i) \quad \alpha_{n+1} := \frac{1}{2} (\alpha_n^{1/2} + \alpha_n^{-1/2}),$$

(ii) $\beta_{n+1} := \alpha_n^{1/2}\left(\dfrac{\beta_n + 1}{\beta_n + \alpha_n}\right)$,

(iii) $\pi_{n+1} := \pi_n\beta_{n+1}\left(\dfrac{1 + \alpha_{n+1}}{1 + \beta_{n+1}}\right)$.

Then π_n converges exponentially to π and

$$|\pi_n - \pi| \le \frac{1}{10^{2^n}}.$$

Proof. Consider the formula

(5.1)
$$\frac{1}{2^n}\log\left(4\,\frac{a_n}{c_n}\right) - \frac{\pi}{2}\,\frac{a_n}{a_n'}$$

which, as we will see later, converges exponentially at a uniform rate to zero in some (complex) neighbourhood of $1/\sqrt{2}$. (We are considering each of $a_n, b_n, c_n, a_n', b_n', c_n'$ as being functions of a complex initial value k, i.e. $b_0 = k$, $b_0' = \sqrt{1 - k^2}$, $a_0 = a_0' = 1$.)

Differentiating (5.1) with respect to k yields

(5.2)
$$\frac{1}{2^n}\left(\frac{\dot{a}_n}{a_n} - \frac{\dot{c}_n}{c_n}\right) - \frac{\pi}{2}\,\frac{a_n}{a_n'}\left(\frac{\dot{a}_n}{a_n} - \frac{\dot{a}_n'}{a_n'}\right)$$

which also converges uniformly exponentially to zero in some neighbourhood of $1/\sqrt{2}$. (This general principle for exponential convergence of differentiated sequences of analytic functions is a trivial consequence of the Cauchy integral formula.) We can compute $\dot{a}_n$, $\dot{b}_n$ and $\dot{c}_n$ from the recursions

$$\dot{a}_{n+1} := \frac{\dot{a}_n + \dot{b}_n}{2},$$

(5.3)
$$\dot{b}_{n+1} := \frac{1}{2}\left(\dot{a}_n\sqrt{\frac{b_n}{a_n}} + \dot{b}_n\sqrt{\frac{a_n}{b_n}}\right),$$

$$\dot{c}_{n+1} := \frac{1}{2}(\dot{a}_n - \dot{b}_n),$$

where $\dot{a}_0 := 0$, $\dot{b}_0 := 1$, $a_0 := 1$ and $b_0 := k$.

We note that a_n and b_n map $\{z \mid Re(z) > 0\}$ into itself and that $\dot{a}_n$ and $\dot{b}_n$ (for sufficiently large n) do likewise.

It is convenient to set

(5.4)
$$\alpha_n := \frac{a_n}{b_n} \quad \text{and} \quad \beta_n := \frac{\dot{a}_n}{\dot{b}_n}$$

with

$$\alpha_0 := \frac{1}{k} \quad \text{and} \quad \beta_0 := 0.$$

We can derive the following formulae in a completely elementary fashion from the basic relationships for a_n, b_n and c_n and (5.3):

(5.5)
$$\dot{a}_{n+1} - \dot{b}_{n+1} = \frac{1}{2}(\sqrt{a_n} - \sqrt{b_n})\left(\frac{\dot{a}_n}{\sqrt{a_n}} - \frac{\dot{b}_n}{\sqrt{b_n}}\right),$$

$$(5.6) \qquad 1 - \frac{a_{n+1}\,\dot{c}_{n+1}}{\dot{a}_{n+1}\,c_{n+1}} = \frac{2(\alpha_n - \beta_n)}{(\alpha_n - 1)(\beta_n + 1)},$$

$$(5.7) \qquad \alpha_{n+1} = \frac{1}{2}(\alpha_n^{1/2} + \alpha_n^{-1/2}),$$

$$(5.8) \qquad \beta_{n+1} = \alpha_n^{1/2}\left(\frac{\beta_n + 1}{\beta_n + \alpha_n}\right),$$

$$(5.9) \qquad \alpha_{n+1} - 1 = \frac{1}{2\alpha_n^{1/2}}(\alpha_n^{1/2} - 1)^2,$$

$$(5.10) \qquad \alpha_{n+1} - \beta_{n+1} = \frac{\alpha_n^{1/2}}{2}\frac{(1 - \alpha_n)(\beta_n - \alpha_n)}{\alpha_n(\beta_n + \alpha_n)},$$

$$(5.11) \qquad \frac{\alpha_{n+1} - \beta_{n+1}}{\alpha_{n+1} - 1} = \frac{(1 + \alpha_n^{1/2})^2}{(\beta_n + \alpha_n)} \cdot \frac{(\alpha_n - \beta_n)}{(\alpha_n - 1)}.$$

From (5.7) and (5.9) we deduce that $\alpha_n \to 1$ uniformly with second order in compact subsets of the open right half-plane. Likewise, we see from (5.8) and (5.10) that $\beta_n \to 1$ uniformly and exponentially. Finally, we set

$$(5.12) \qquad \gamma_n := \frac{1}{2^n}\left(\frac{\alpha_n - \beta_n}{\alpha_n - 1}\right).$$

We see from (5.11) that

$$(5.13) \qquad \gamma_{n+1} = \frac{(1 + \alpha_n^{1/2})}{2(\beta_n + \alpha_n)}\gamma_n$$

and also from (5.6) that

$$(5.14) \qquad \frac{\gamma_n}{1 + \beta_n} = \frac{1}{2^{n+1}}\left(1 - \frac{a_{n+1}\,\dot{c}_{n+1}}{\dot{a}_{n+1}\,c_{n+1}}\right).$$

Without any knowledge of the convergence of (5.1) one can, from the preceding relationships, easily and directly deduce the exponential convergence of (5.2), in $\{z \mid |z - \frac{1}{2}| \leq c < \frac{1}{2}\}$. We need the information from (5.1) only to see that (5.2) converges to zero.

The algorithm for π comes from multiplying (5.2) by $a_n/\dot{a}_n$ and starting the iteration at $k := 2^{-1/2}$. For this value of k $a_n' = a_n$, $(\dot{a}_n') = -\dot{a}_n$ and

$$\frac{1}{2^{n+1}}\left(1 - \frac{a_{n+1}\,\dot{c}_{n+1}}{\dot{a}_{n+1}\,c_{n+1}}\right) \to \pi$$

which by (5.14) shows that

$$\pi_n := \frac{\gamma_n}{1 + \beta_n} \to \pi.$$

Some manipulation of (5.7), (5.8) and (5.13) now produces (iii). The starting values for α_n, β_n and γ_n are computed from (5.4). Other values of k will also lead to similar, but slightly more complicated, iterations for π.

To analyse the error one considers

$$\frac{\gamma_{n+1}}{1 + \beta_{n+1}} - \frac{\gamma_n}{1 + \beta_n} = \left[\frac{(1 + \alpha_n^{1/2})^2}{2(\beta_n + \alpha_n)(1 + \beta_{n+1})} - \frac{1}{(1 + \beta_n)}\right]\gamma_n$$

and notes that, from (5.9) and (5.10), for $n \geqq 4$,

$$|\alpha_n - 1| \leq \frac{1}{10^{2^n+2}} \quad \text{and} \quad |\beta_n - 1| \leq \frac{1}{10^{2^n+2}}.$$

(One computes that the above holds for $n = 4$.) Hence,

$$\left| \frac{\gamma_{n+1}}{1 + \beta_{n+1}} - \frac{\gamma_n}{1 + \beta_n} \right| \leqq \left| \frac{1}{10^{2^n+1}} \right| |\gamma_n|$$

and

$$\left| \frac{\gamma_n}{1 + \beta_n} - \pi \right| \leqq \frac{1}{10^{2^n}}. \qquad \square$$

In fact one can show that the error is of order $2^n e^{-\pi 2^{n+1}}$.

If we choose integers in $[\delta, \delta^{-1}]$, $0 < \delta < \frac{1}{2}$ and perform n operations $(+, -, \times, \div, \sqrt{\ })$ then the result is always less than or equal to δ^{2^n}. Thus, if $\gamma > \delta$, it is not possible, using the above operations and integral starting values in $[\delta, \delta^{-1}]$, for every n to compute π with an accuracy of $O(\gamma^{-2^n})$ in n steps. In particular, convergence very much faster than that provided by Theorem 2 is not possible.

The analysis in this section allows one to derive the Gauss-Salamin formula (4.6) without using Legendre's formula or second integrals. This can be done by combining our results with problems 15 and 18 in [11]. Indeed, the results of this section make quantitative sense of problems 16 and 17 in [11]. King also observes that Legendre's formula is actually equivalent to the Gauss–Salamin formula and that each may be derived from the other using only properties of the AGM which we have developed and equation (4.5).

This algorithm, like the algorithms of §4, is not self correcting in the way that Newton's method is. Thus, while a certain amount of time may be saved by observing that some of the calculations need not be performed to full precision it seems intrinsic (though not proven) that $O(\log n)$ full precision operations must be executed to calculate π to n digits. In fact, showing that π is intrinsically more complicated from a time complexity point of view than multiplication would prove that π is transcendental [5].

6. The complex AGM iteration. The AGM iteration

$$a_{n+1} := \frac{1}{2}(a_n + b_n), \qquad b_{n+1} := \sqrt{a_n b_n}$$

is well defined as a complex iteration starting with $a_0 := 1$, $b_0 := z$. Provided that z does not lie on the negative real axis, the iteration will converge (to what then must be an analytic limit). One can see this geometrically. For initial z in the right half-plane the limit is given by (3.9). It is also easy to see geometrically that a_n and b_n are always nonzero.

The iteration for $x_n := b_n/a_n$ given in the form (3.23) as $x_{n+1} := 2\sqrt{x_n}/x_{n+1}$ satisfies

$$(6.1) \qquad (x_{n+1} - 1) = \frac{(1 - \sqrt{x_n})^2}{1 + x_n}.$$

This also converges in the cut plane $\mathbb{C} - (-\infty, 0]$. In fact, the convergence is uniformly exponential on compact subsets (see Fig. 1). With each iteration the angle θ_n between x_n and 1 is at least halved and the real parts converge uniformly to 1.

It is now apparent from (6.1) and (3.24) that

$$(6.2) \qquad P_n(k) := \left(\frac{4a_n}{c_n} \right)^{2^{1-n}} = \left(\frac{1 + x_n}{1 - x_n} \right)^{2^{-n}}$$
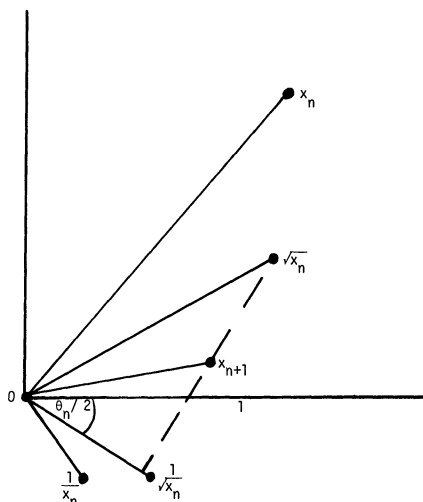
FIG. 1.

and also,

$$Q_n(k) := \frac{a_n}{a_n'}$$

converge exponentially to analytic limits on compact subsets of the complex plane that avoid

$$D := \{z \in \mathbb{C} \mid z \notin (-\infty, 0] \cup [1, \infty)\}.$$

Again we denote the limits by P and Q. By standard analytic reasoning it must be that (3.19a) still holds for k in D.

Thus one can compute the complex exponential—and so also cos and sin—exponentially using (3.19). More precisely, one uses Newton's method to approximately solve $Q(k) = z$ for k and then computes $P_n(k)$. The outcome is e^z. One can still perform the root extractions using Newton's method. Some care must be taken to extract the correct root and to determine an appropriate starting value for the Newton inversion. For example $k := 0.02876158$ yields $Q(k) = 1$ and $P_4(k) = e$ to 8 significant places. If one now uses k as an initial estimate for the Newton inversions one can compute $e^{1+i\theta}$ for $|\theta| \leq \pi/8$. Since, as we have observed, e is also exponentially computable we have produced a sufficient range of values to painlessly compute $\cos\theta + i\sin\theta$ with no recourse to any auxiliary computations (other than π and e, which can be computed once and stored). By contrast Brent's trigonometric algorithm needs to compute a different logarithm each time.

The most stable way to compute P_n is to use the fact that one may update c_n by

(6.3) $$c_{n+1} = \frac{c_n^2}{4a_{n+1}}.$$

One then computes a_n, b_n and c_n to desired accuracy and returns

$$\left(\frac{4a_n}{c_n}\right)^{1/2^n} \quad \text{or} \quad \left(\frac{2(a_n + b_n)}{c_n}\right)^{1/2^n}.$$

This provides a feasible computation of P_n, and so of exp or log.

In an entirely analogous fashion, formula (4.2) for log is valid in the cut complex plane. The given error estimate fails but the convergence is still exponential. Thus (4.2) may also be used to compute all the elementary functions.

7. Concluding remarks and numerical data. We have presented a development of the AGM and its uses for rapidly computing elementary functions which is, we hope, almost entirely self-contained and which produces workable algorithms. The algorithm for π is particularly robust and attractive. We hope that we have given something of the flavour of this beautiful collection of ideas, with its surprising mixture of the classical and the modern. An open question remains. Can one entirely divorce the central discussion from elliptic integral concerns? That is, can one derive exponential iterations for the elementary functions without recourse to some nonelementary transcendental functions? It would be particularly nice to produce a direct iteration for e of the sort we have for π which does not rely either on Newton inversions or on binary splitting.

The algorithm for π has been run in an arbitrary precision integer arithmetic. (The algorithm can be easily scaled to be integral.) The errors were as follows:

Iterate	Digits correct	Iterate	Digits correct
1	3	6	170
2	8	7	345
3	19	8	694
4	41	9	1392
5	83	10	2788

Formula (4.2) was then used to compute 2 log (2) and log (4), using π estimated as above and the same integer package. Up to 500 digits were computed this way. It is worth noting that the error estimate in (4.2) is of the right order.

The iteration implicit in (3.22) was used to compute e^{π} in a double precision Fortran. Beginning with $k := 2^{-1/2}$ produced the following data:

Iterate	$P_n - e^{\pi}$	$a_n/b_n - 1$
1	1.6×10^{-1}	1.5×10^{-2}
2	2.8×10^{-9}	2.8×10^{-5}
3	1.7×10^{-20}	9.7×10^{-11}
4	$< 10^{-40}$	1.2×10^{-21}

Identical results were obtained from (6.3). In this case $y_n := 4a_n/c_n$ was computed by the two term recursion which uses x_n, given by (3.23), and

$$(7.1) \qquad y_0^2 := \frac{16}{1 - k^2}, \qquad y_{n+1} = \left(\frac{1 + x_n}{2}\right)^2 y_n^2.$$

One observes from (7.1) that the calculation of y_n is very stable.

We conclude by observing that the high precision root extraction required in the AGM [18], was actually calculated by inverting $y = 1/x^2$. This leads to the iteration

$$(7.2) \qquad x_{n+1} = \frac{3x_n - x_n^3 y}{2}$$

for computing $y^{-1/2}$. One now multiplies by y to recapture $\sqrt{y}$. This was preferred because it avoided division.

REFERENCES

[1] N. H. ABEL, *Oeuvres complètes,* Grondahl and Son, Christiana, 1881.

[2] P. BECKMAN, *A History of Pi,* Golem Press, Ed. 4, 1977.

[3] F. BOWMAN, *Introduction to Elliptic Functions,* English Univ. Press, London, 1953.

[4] R. P. BRENT, *Multiple-precision zero-finding and the complexity of elementary function evaluation,* in Analytic Computational Complexity, J. F. Traub, ed., Academic Press, New York, 1975, pp. 151–176.

[5] ———, *Fast multiple-precision evaluation of elementary functions,* J. Assoc. Comput. Mach., 23 (1976), pp. 242–251.

[6] B. C. CARLSON, *Algorithms involving arithmetic and geometric means,* Amer. Math. Monthly, 78(1971), pp. 496–505.

[7] A. CAYLEY, *An Elementary Treatise on Elliptic Functions,* Bell and Sons, 1895, republished by Dover, New York, 1961.

[8] A. EAGLE, *The Elliptic Functions as They Should Be,* Galloway and Porter, Cambridge, 1958.

[9] C. H. EDWARDS, Jr. *The Historical Development of the Calculus,* Springer-Verlag, New York, 1979.

[10] K. F. GAUSS, *Werke,* Bd 3, Gottingen, 1866, 361–403.

[11] L. V. KING, *On the direct numerical calculation of elliptic functions and integrals,* Cambridge Univ. Press, Cambridge, 1924.

[12] D. KNUTH, *The Art of Computer Programming. Vol. 2: Seminumerical Algorithms,* Addison-Wesley, Reading, MA, 1969.

[13] H. T. KUNG AND J. F. TRAUB, *All algebraic functions can be computed fast,* J. Assoc. Comput. Mach., 25(1978), 245–260.

[14] A. M. LEGENDRE, *Exercices de calcul integral,* Vol. 1, Paris, 1811.

[15] D. J. NEUMAN, *Rational approximation versus fast computer methods,* in Lectures on Approximation and Value Distribution, Presses de l'Université de Montréal, Montreal, 1982, pp. 149–174.

[16] E. SALAMIN, *Computation of π using arithmetic-geometric mean,* Math. Comput. 135(1976), pp. 565–570.

[17] D. SHANKS AND J. W. WRENCH, Jr. *Calculation of π to* 100,000 *decimals,* Math. Comput., 16 (1962), pp. 76–79.

[18] Y. TAMURA AND Y. KANADA, *Calculation of π to* 4,196,293 *decimals based on Gauss-Legendre algorithm,* preprint.

[19] J. TODD, *Basic Numerical Mathematics,* Vol. 1, Academic Press, New York, 1979.

[20] G. N. WATSON, *The marquis and the land agent,* Math. Gazette, 17 (1933), pp. 5–17.

[21] J. W. WRENCH, Jr. *The evolution of extended decimal approximations to π.* The Mathematics Teacher, 53(1960), pp. 644–650.