1971

THE HISTORY AND NATURE OF PI

Karen A. Rand

THE HISTORY AND NATURE OF π

A Thesis

Submitted to the Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirement for credit in
Arts and Sciences 199^H

in

The Department of Mathematics

by
Karen A. Rand
May, 1971

ABSTRACT

A brief history of π is given, and one proof that π is transcendental is presented.

That the ratio of the length of a circle to its diameter is a constant value for any circle has been understood since ancient times. The value of this ratio is denoted by the Greek letter π. In its long history, attempts have been made to find the value of π and to construct a square having an area equal to that of a given circle.

For a circle of diameter one, the circumference will be of length π; and, since the area of such a circle is $\frac{\pi}{4}$, the desired square is one of side $\frac{\sqrt{\pi}}{2}$. A segment of length π gives rise to the construction of a segment of length $\sqrt{\pi}$[1], and bisecting this segment would give the length of the side of the desired square. Thus, early attempts to construct a segment of the same length as the circumference of a circle are equivalent to attempts to square the circle.

There is a value for π of 256/81 recorded as early as 1700 BC, and Biblical references to the value 3. Attempts by Hippias of Elis, Antiphon, Bryson, and Hippocrates of Chion to determine the value are known. These early attempts were made by means of Euclidean constructions. The methods of Antiphon and Bryson are still to be found in elementary geometry textbooks. They consist of inscribing regular polygons with successively greater numbers of sides in a circle and circumscribing similar polygons about the circle, thus yielding approximations for the area of the circle.

Notable among the Greeks who studied the problem was

Archimedes, who, by inscribing and circumscribing polygons of 96 sides, arrived at 3 1/7 as an upper bound and 3 10/71 as a lower bound for π.

Later Greeks to consider the problem included Hipparchus and Ptolemy, both of whom compiled tables of chords of a circle, the latter expressing π in sexagesimal measure as $3^{\circ}8'3''$ or 3 8/60 3/3600 or about 3.14166.

In India, the value 62832/20000 = 3.1416 may have been discovered as early as 500 AD by Aryabhatta. Bhaskara, born in 1114 AD, offered 3927/1250 as an exact value for π. It has been conjectured that the value of Brahmagupta, $\sqrt{10}$, was derived from the approximation

$$\sqrt{a^2+x} = a+x/(2a+x)$$

by setting a=3 and x=1.

Chinese approximations to π date from the twelfth century BC when Chou-Kong set π equal to 3. Chang Heng (78-139 AD) believed that the ratio of the circumference squared to the perimeter squared of the circumscribed square is 5/8, or that π is $\sqrt{10}$. Wang Fau gave the value 142/45, or about 3.1555. The method of approximating π by inscribing regular polygons was used by Liu Hui to derive the value 157/50 or 3.14. In the fifth century AD, Tsu Ch'ung-chih

arrived at the limits 31.415926 and 31.415927 for 10π and took 22/7 and 355/113 to be approximate values.

In the Middle Ages, Leonardo Pisano arrived at the limits $1440/458\frac{1}{5}$, approximately 3.1427, and $1440/458\frac{4}{9}$, or about 3.1410, as limits for π. These compare favorably with the Archimedean limits of 3.1428 and 3.1408. Pisamo took π to be 1440/458 .

At about the beginning of the seventeenth century, Adriaen Anthonisz rediscovered the value 355/113, possibly by taking the arithmetic means of the numerators and denominators of 333/106 and 377/120, lower and upper limits for π which he arrived at by the Archimedean method.

Vieta (1540-1603) found an infinite sequence of operations by which π could be expressed. Given a polygon of n sides and a polygon of 2n sides, both inscribed in the same circle, he showed the ratio of the area of the first polygon to the area of the second to be the same as the ratio of the supplementary chord of the side of the first to the diameter of the circle. This yielded the expression:

$$\frac{\pi}{2} = \frac{1}{\sqrt{\frac{1}{2}}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}}\cdots} .$$

He also used the Archimedean method with polygons of up to $2^{16} \cdot 6$ sides to arrive at the inequality:

$$3.1415926535 < \pi < 3.1415926537$$

Adrianus Romanus (1561-1615) computed π to fifteen places using polygons of $15 \cdot 2^{24}$ sides, and Ludolph von Ceulen (1539-1610) computed π to 35 places, "a fact that was thought to be so noteworthy as to lead to π being called the Ludolphian number, a name still used in Germany."[2]

Snellius (1580-1626) arrived at narrower limits without increasing the number of sides of the polygons by use of two theorems equivalent to the inequality:

$$\frac{1}{3}(2\sin\theta + \tan\theta) < \theta < 3/(2\csc\theta + \cot\theta).$$

Using hexagons he arrived at narrower limits than Archimedes had with polygons of 96 sides, and he computed π to 34 places with a polygon of 2^{30} sides, compared with Ludolph's fourteen places.

Christian Huyghens (1629-1665) published sixteen theorems proved by geometrical processes but did not find a method of constructing a segment of length π. By use of polygons of 60 sides he calculated π to nine decimal places.

Another approach to the problem by geometrical processes was the attempt by Descartes to construct the diameter of a circle given a segment to be taken as the length of its circumference. While he was able to construct a sequence of segments the lengths of which approached that of the diameter, he was able to solve the problem for all practical purposes without solving the ideal problem.

With the development of analysis emphasis shifted from geometrical determinations and straight edge and compass constructions to an interest in the analytical expression of π in infinite series or products. John Wallis (1616-1703) showed that

$$\frac{\pi}{2} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdot \frac{8}{7} \cdot \frac{8}{9} \cdot \cdots$$

and proved correct the continued fraction

$$\frac{4}{\pi} = 1 + \cfrac{1}{2 + \cfrac{9}{2 + \cfrac{25}{2 + \cfrac{49}{2 + \cdots}}}}$$

which had been communicated to him by Lord Brouncker.

The series

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots \qquad (-1 \leq x \leq 1),$$

discovered by James Gregory in 1670 and by Leibniz in 1673, yields a series for $\pi/4$, but "it converges so slowly as not to be convenient in practice."[3] By taking x to be $\sqrt{1/3}$ rather than 1, a more rapidly convergent series for $\pi/6$ is obtained, which Abraham Sharp used to compute π to 72 decimal places. By use of the series

$$\sin^{-1} x = x + \frac{1}{2} \cdot \frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4} \cdot \frac{x^5}{5} + \cdots$$

Newton set $x = \frac{1}{2}$ and computed π to fourteen places.

Euler, in 1737, formulated the addition formula

$$\tan^{-1}(1/p) = \tan^{-1}(1/(p+q)) + \tan^{-1}(q/(p^2+pq+1))$$

and derived formulae for $\tan^{-1}(x/y)$ in special cases which led to the series

$$\frac{\pi}{4} = \tan^{-1}(\tfrac{1}{2}) + \tan^{-1}(1/(2\cdot 4)) + \tan^{-1}(1/(2\cdot 9)) + \ldots .$$

In 1706 Machin arrived at the formula

$$\frac{\pi}{4} = 4\tan^{-1}(1/5) - \tan^{-1}(1/239)$$

$$= 4(1/5 - (1/3\cdot 5^3) + (1/5\cdot 5^5) - (1/7\cdot 5^7) + \ldots)$$

$$-(1/239 - (1/3\cdot 239^3) + (1/5\cdot 239^5) - (1/7\cdot 239^7) + \ldots)$$

and π was computed to 100 decimal places. In 1719 deLagny computed π to 127 places. Vega computed π to 140 places using a formula due to Euler,

$$\frac{\pi}{4} = 5\tan^{-1}(1/7) + 2\tan^{-1}(3/79)$$

$$= 2\tan^{-1}(1/3) + \tan^{-1}(1/7) ,$$

and corrected the 113th place of deLagny's value. In 1847 Clausen used Machin's formula and the expression

$$\frac{\pi}{4} = 2\tan^{-1}(1/3) + \tan^{-1}(1/7)$$

to compute π to 248 places, after Rutherford's computation of π

to 208 places (of which 152 were correct) and the earlier calculation of π to 200 places by Zacharias Dase using

$$\frac{\pi}{4} = \tan^{-1}(\tfrac{1}{2}) + \tan^{-1}(1/5) + \tan^{-1}(1/8).$$

The degree of accuracy increased to 440 places (Rutherford), 530 and 607 (Shanks), and in 1873-4 to 707 places (Shanks, using Machin's formula).

Despite the amount of time and effort spent on these successive approximations to π, the expressions arrived at for π shed no light on the question of squaring the circle. One of the more important steps which did have bearing on that question was the discovery of the relationship between π and e, the base of the natural logarithm. It was by use of Maclaurin's series for e^x, cos x, and sin x that the formula due to Euler,

$$e^{ix} = \cos x + i \sin x \quad ,$$

was derived. Setting $x = \pi$ in this last equation yields

$$e^{i\pi} = -1,$$

a relationship which was "indispensable later on in making out the true nature of the number π."[4]

J. H. Lambert in 1766 published two theorems, the first that if x is rational and not zero then e^x is not rational; and the second that if x is rational and not zero then tan x is not rational. He drew these conclusions by use of Euler's continued

fraction

$$\frac{e-1}{2} = \cfrac{1}{1+\cfrac{1}{6+\cfrac{1}{10+\cfrac{1}{14+1 \cdots}}}}$$

to derive the formulae

$$\frac{e^x-1}{e^x+1} = \cfrac{1}{\frac{2}{x}+\cfrac{1}{\frac{6}{x}+\cfrac{1}{\frac{10}{x}+\cfrac{1}{\frac{14}{x}+\cdots}}}}$$

and

$$\tan x = \cfrac{1}{\frac{1}{x}-\cfrac{1}{\frac{3}{x}-\cfrac{1}{\frac{5}{x}-\cfrac{1}{\frac{7}{x}-\cfrac{1}{\frac{9}{x}-\cdots}}}}}\,.$$

From these theorems he concluded that $\pi/4$ is irrational. Lambert's proof was not rigorous, but Legendre did prove rigorously that π and π^2 are irrational. Hermite also proved that π and π^2 are irrational, his proofs "containing the germ of the later proof of the transcendency of e and π."[5]

In 1840 it was proved by Liouville that neither e nor e^2 was a root of a quadratic equation with rational coefficients. In the same year he proved the existence of numbers which are not algebraic, that is, not roots of polynomials with rational

coefficients. He used two different methods of proof, one of which is presented in reference (3). Cantor's proof of the existence of such numbers, known as transcendental, appeared in 1874. Since there are countably many rational numbers there are countably many roots of polynomials with rational coefficients, that is, countably many algebraic numbers. However, there are uncountably many real numbers, thus transcendental numbers must constitute the excess.

The problem of the possibility of squaring the circle as it was originally posed was a question of the possibility of a Euclidean construction using straight edge and compass. Such constructions determine points by three methods: the intersection of two lines; the intersection of a circle with a line; and the intersection of two circles. As Cartesian coordinates came into use it became possible to express these figures by equations in two variables and to determine their points of intersection by solving simultaneous equations. Thus any possible Euclidean construction (which consists of a finite number of applications of the three methods of determination) must also be expressible by a finite number of rational operations and square root operations of the coordinates of the points from which the determination is to be made. Additionally, any point determinable by such methods will have coordinates which will be roots of polynomials of degree a power of two, with coefficients which are rational functions of the coordinates of the given points. Thus only certain classes of algebraic numbers are determinable by

classical constructions and no transcendental number will be determinable by such methods.

In 1873 it was demonstrated by Hermite that the number e is transcendental, that is, that, there is no equation of the form

$$c_0 + c_1 e + c_2 e^2 + \cdots + c_n e^n = 0$$

where the c_i's are integers. In 1882 Lindemann extended Hermite's proof to a more general theorem, that no equation of the form

$$0 = c_0 + c_1 (e^{k_1} + e^{k_2} + \cdots + e^{k_n}) + c_2 (e^{m_1} + e^{m_2} + \cdots + e^{m_{n'}}) + \cdots$$

can hold where the coefficients are integral and the exponents algebraic. Thus the equation $e^{i\pi} + 1 = 0$ implies that $i\pi$ is transcendental and this implies that π is transcendental. The proofs of Hermite and Lindemann use methods of complex integration. The latter proof was simplified by Weierstrass in 1885, and other proofs that π is transcendental were published by Stieltjes (1890), Hilbert, Hurwitz, Gordan, and Mertens (1896), and Vahlen (1900). With the proof that π is not an algebraic number the impossibility of squaring the circle was demonstrated.

One proof of the transcendency of π is that of Gordon, presented in reference (3). It begins with the assumption that π is algebraic, from which it follows that $i\pi$ is algebraic. From an equation with integral coefficients of which $i\pi$ is a root, an equation of the form

$$A + e^{B_1} + e^{B_2} + e^{B_3} + \cdots + e^{B_n} = 0$$

is constructed in which symmetric functions of B_i's are integral and A is a positive integer. A function $F(x)$ is constructed so that the non-zero sum

$$F^{(p-1)}(0) + F^{(p)}(0) + F^{(p+1)}(0) + \cdots + F^{(np+p-1)}(0)$$

multiplied by each side of the equation gives the sum of a fraction and an integer, not zero, equal to zero. Such an equation is impossible, the assumption that π is algebraic is contradicted, and it is shown that π is transcendental.

Since some use is made of symmetric functions it is useful to note that a symmetric function in n independent variables is one which "is unaltered by the interchange of any two of the variables."[6] The sum of the cubes of each of the variables, for example, is symmetric in those variables. There are n elementary symmetric functions of $x_1, x_2, \ldots, x_n$. The k th elementary symmetric function of the n variables is the sum of products of k distinct x_i. As an example, given three variables x_1, x_2, and x_3 the elementary symmetric functions are

$$\sum_1 = x_1 + x_2 + x_3 \;,$$

$$\sum_2 = x_1x_2 + x_1x_3 + x_2x_3 \;,$$

and $$\sum_3 = x_1x_2x_3 \;.$$

Also, given any symmetric function of $x_1, x_2, \ldots, x_n$ it can be expressed as sums or products of the n elementary symmetric functions of the n variables. The expression $x_1^2 + x_2^2 + x_3^2 + x_1x_2x_3$ is symmetric in the three variables; it is equal to

$$(\sum_1 x_i)^2 - 2\sum_2 x_i + \sum_3 x_i \;.$$

The general theorem is proved in reference (2).

The proof that π is transcendental begins with the assumption that π is algebraic, that is, that there is a polynomial in x with integral coefficients of which π is a root. From this polynomial it is possible to arrive at an equation with integral coefficients of which $i\pi$ is a root.

Suppose $P(x)$ is the polynomial of which π is a root. Then letting $Q(x) \equiv P(-ix)$ it is apparent that $Q(i\pi) = 0$. However, $Q(x)$ does not necessarily have integral coefficients. Suppose

$$Q(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \;.$$

Then $\overline{Q}(x)$ is defined to be the polynomial arrived at by conjugating the coefficients of $Q(x)$. Then $Q(x)\overline{Q}(x)$ is given by

$$(a_n\overline{a_n})\, x^{2n} + (a_n\overline{a_{n-1}} + \overline{a_n}a_{n-1})\, x^{2n-1} + (a_{n-1}\overline{a_{n-1}} + a_n\overline{a_{n-2}} + \overline{a_n}a_{n-2})x^{2n-2}$$

$$+x^{2n-3}(a_n\overline{a_{n-3}} + \overline{a_n}a_{n-3} + a_{n-1}\overline{a_{n-2}} + \overline{a_{n-1}}a_{n-2}) + \cdots + a_0\overline{a_0} \; .$$

But $a_k\overline{a_k} = |a_k|^2$ and $a_r\overline{a_s} + \overline{a_r}a_s = 2\,\mathrm{Re}(a_r\overline{a_s})$ so the product $Q(x)\overline{Q}(x)$ has real coefficients which, since the coefficients of $Q(x)$ are integers times powers of i, are integers. Thus if π is algebraic, $i\pi$ is algebraic.

We are free to assume that $i\pi$ is a root of the equation

$$C(x-a_1)(x-a_2)(x-a_3)\cdots(x-a_s) = 0 \qquad (C \neq 0)$$

where the numbers

$$C,\; C\sum_1 a_i\;,\; C\sum_2 a_i\;,\ldots,\; Ca_1a_2\cdots a_s$$

are integers, the notation $\sum_k a_i$ referring to the k th elementary symmetric function of $a_1, a_2, \ldots, a_s$. From the factorization we see that one of $a_1, a_2, \ldots, a_s$ is $i\pi$ and since $e^{i\pi}+1 = 0$, the equation

$$(1+e^{a_1})(1+e^{a_2})\cdots(1+e^{a_s}) = 0$$

must hold. Multiplying this out yields an equation of the form

$$1+e^{a_1}+e^{a_2}+\cdots+e^{a_s}+e^{a_1+a_2}+e^{a_1+a_3}+\cdots+e^{a_1+a_s}$$

$$+\cdots+e^{a_1+a_2+\cdots+a_s} = 0.$$

This can be reduced to the form

$$A+e^{B_1}+e^{B_2}+\cdots+e^{B_n} = 0$$

where none of $B_1, \ldots, B_n$ are zero and where A is an integer

greater than or equal to one.

The elementary symmetric functions of Ca_1, $Ca_2, \ldots, Ca_s$ are integers since they may be expressed as powers of C times elementary symmetric functions of $a_1, a_2, \ldots, a_s$, and these products are integers. To prove that the elementary symmetric functions of $CB_1, CB_2, \ldots, CB_n$ we show that the sums of products taken p together of a+b letters $x_1, x_2, \ldots, x_a, y_1, y_2, \ldots, y_b$ may be expressed in terms of elementary symmetric functions of the separate sets $x_1, x_2, \ldots, x_a$ and $y_1, y_2, \ldots, y_b$. A simple example may serve to clarify the proof. If we are given two sets x_1, x_2 and y_1, y_2 and the elementary symmetric function

$$\sum_2 (x,y) = x_1x_2 + x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2 + y_1y_2 ,$$

we can express this as

$$x_1x_2 + (x_1+x_2)(y_1+y_2) + y_1y_2 = \sum_2 x_i + (\sum_1 x_i)(\sum_1 y_k) \sum_2 y_k .$$

The proof is a generalization of this example. By $\sum_p (x,y)$ is denoted the p^{th} elementary symmetrical function of the x's and y's; by $\sum_r x_i$, the r th elementary symmetric function of the x's; and by $\sum_s y$, the s th elementary symmetric function of the y's, where $r \leq a$, $s \leq b$ and $p \leq a+b$.

Case one occurs when $p < a$; it is seen that $\sum_p (x,y)$ must contain in each separate product of which it is composed at most p factors from either the set of x's or the set of y's (at most b factors from the set of y's if $b \leq p$). Thus, $\sum_p (x,y)$ is either

$$\sum_p x + \sum_{p-1} x \sum_1 y + \ldots + \sum_{p-b} x \sum_b y \quad ,$$

or

$$\sum_{p} x + \sum_{p-1} x \sum_{1} y + \dots + \sum_{b} y .$$

Case two, for which $p \geq a$, yields

$$\sum_{p} (x,y) = \sum_{a} x \sum_{p-a} y + \sum_{a-1} x \sum_{p-a+1} y \quad \dots ,$$

the series concluding with $\sum_{p-b} x \sum_{b} y$ for $b < p$, or with $\sum_{p} y$ for $b \geq p$.

These give all the possible products of p of the letters $x_1, \dots, x_a, y_1, \dots, y_b$; thus, the symmetric functions of the letters in the union of the two sets may be expressed in terms of the elementary symmetric functions of the letters in the separate sets.

The demonstration may be extended by an induction argument to elementary symmetric functions of letters in any number of separate sets.

The CB_i's are separated into sets according to the number of Ca_j's that form them; if $CB_i = C(a_{i1} + a_{i2} + \dots + a_{ir}) \not\equiv 0$ and $CB_j = C(a_{j1} + a_{j2} + \dots + a_{jr}) \not\equiv 0$, then CB_i and CB_j are members of the same set. Symmetric functions of the CB_i's in one set are symmetric functions of the Ca_j's themselves and hence may be expressed in terms of the elementary symmetric functions of the Ca_j's. Since these are integers, it follows that symmetric functions of the CB_i's in separate sets are integral.

It is now essential to consider the function

$$F(x) = \frac{x^{p-1}}{(p-1)!} \quad C^{np+p-1}\,((x-B_1)(x-B_2)\cdots(x-B_n))^p$$

for p prime and greater than all of A, n, C, $|C^n B_1 B_2 \cdots B_n|$. By appropriate manipulation of this function we will construct an integer which when multiplied by $A + e^{B_1} + e^{B_2} + \cdots + e^{B_n}$ will give the sum of a non-zero integer and a fraction. It is necessary to consider several different ways of expressing $F(x)$ in the course of these manipulations. By separating C^{np+p-1} into $C^{p-1} \cdot C^{np}$ we see

$$F(x) = \frac{(Cx)^{p-1}}{(p-1)!}\,(C^n\,(x-B_1)(x-B_2)\,\cdots\,(x-B_n))^p,$$

and multiplying out the last portion of the expression gives

$$F(x) = \frac{(Cx)^{p-1}}{(p-1)!}\quad((Cx)^n - q_1(Cx)^{n-1} + q_2(Cx)^{n-2} - \ldots + (-1)^n q_n)^p$$

where $q_j = \sum_j CB_i$, thus q_j is an integer. $F(x)$ may be expressed as a polynomial in x of degree $np+p-1$ where the coefficients of all powers of x less than $p-1$ are zero. Further, the coefficients which correspond to powers of x between $p-1$ and $np+p-1$ are integers divided by $(p-1)!$, that is,

$$F(x) = c_{p-1}x^{p-1} + c_p x^p + \cdots + c_{np+p-1}\,x^{np+p-1}$$

where c_k is an elementary symmetric function of the q_j times a power of C times 1 or -1, depending on k, all divided by the quantity $(p-1)!$.

Since $F(x)$ is such a polynomial, we see that the derivatives

$$F'(0),\ F''(0),\ F'''(0),\ \ldots\ ,\ F^{(p-2)}(0)$$

are all zero and the higher derivatives evaluated at zero are

$$F^{(p-1)}(0) = (p-1)!\ c_{p-1} = C^{p-1}(-1)^{np}q_n{}^p,$$

$$F^{(p)}(0) = p!\ c_p,$$

$$F^{(p+1)}(0) = (p+1)!\ c_{p+1},$$

$$\vdots$$

$$F^{(np+p-1)}(0) = (np+p-1)!\ c_{np+p-1}.$$

But since c_k is an integer times $1/(p-1)!$ these derivatives at zero are integers. $F^{(p-1)}(0)$ is not divisible by p since q_n is the product $C^nB_1B_2\cdots B_n$, which is not zero, and since we assumed that p was a prime greater than the absolute value of this product. However, for the derivatives of order greater than p-1, $k!\ c_k = (k!/(p-1)!)\cdot w$ (w an integer)

$$= k\,(k-1)(k-2)\cdots p\frac{(p-1)!}{(p-1)!}\cdot w$$

which is an integer multiple of p.

$F(x)$ may be written as $(x-B_i)^p\cdot G_i(x)$ where $G_i(x)$ is defined for each i to be $F(x)\cdot(x-B_i)^{-p}$. Any derivative of this product may be expressed as the finite sum of derivatives of $G_i(x)$ and $(x-B_i)^p$ as follows:

$$F^{(d)}(x) = (G_i(x)\cdot(x-B_i)^p)^{(d)}$$

$$= \sum_{u=0}^{d} \binom{d}{u} (G_i(x))^{(d-u)} ((x-B_i)^p)^{(u)}$$

where $\binom{d}{u}$ is defined as $d!/(u!\,(d-u)!)$.

Making use of this product rule, we see that

$$F'(B_i),\ F''(B_i),\ \ldots\ ,\ F^{(p-1)}(B_i) \qquad\qquad (1 \leq i \leq n)$$

must vanish.

We next consider the derivatives

$$F^{(p)}(B_i) = \binom{p}{0} p!\ G_i\ (B_i),$$

$$F^{(p+1)}(B_i) = \binom{p+1}{1}\ p!\ G_i'(B_i),$$

$$\vdots$$

$$F^{(np+p-1)}(B_i) = \binom{np+p-1}{np-1}\ p!\ G_i^{(np-1)}(B_i)\ .$$

The series for these derivatives collapse in each case to a single term because except for the pth derivative of $(x-B_i)^p$, the derivatives of this expression are zero at B_i. The pth derivative is $p!$.

The sums

$$\sum_{i=1}^{n} F^{(p)}(B_i),\ \sum_{i=1}^{n} F^{(p+1)}(B_i),\ \ldots\ ,\ \sum_{i=1}^{n} F^{(np+p-1)}(B_i)$$

take the form

$$\sum_{i=1}^{n} \binom{p}{0} p!\ G_i\ (B_i) = \binom{p}{0} p! \sum_{i=1}^{n} G_i(B_i),$$

$$\sum_{i=1}^{n} \binom{p+1}{1} p!\, G_i'(B_i) = \binom{p+1}{1} p! \sum_{i=1}^{n} G_i'(B_i),$$

$$\vdots$$

$$\sum_{i=1}^{n} \binom{np+p-1}{np-1} p!\, G_i^{(np-1)}(B_i) = \binom{np+p-1}{np-1} p! \sum_{i=1}^{n} G_i^{(np-1)}(B_i) .$$

The value $1/(p-1)!$ is a factor in each of the derivatives of each of the $G_i(x)$ so it may be factored out of each sum. What remains in each sum is a symmetric function in the CB_i's, that is, an integer. We are left, then, with p integers $h_1, h_2, \ldots, h_p$ multiplied respectively by

$$\binom{p}{0} \cancel{\frac{p!}{(p-1)!}}, \binom{p+1}{1} \frac{p!}{(p-1)!}, \ldots, \binom{np+p-1}{np-1} \frac{p!}{(p-1)!} .$$

Thus each of the sums is an integer multiple of p.

Next, we make the definition

$$Kp = (p-1)!\, c_{p-1} + p!\, c_p + \ldots + (np+p-1)!\, c_{np+p-1},$$

where c_k is the coefficient of x^k in $F(x)$. Since

$$Kp = F^{(p-1)}(0) + F^{(p)}(0) + \ldots + F^{(np+p-1)}(0),$$

we note that Kp is an integer not divisible by p.

We shall now consider what happens when the equation

$$A + e^{B_1} + e^{B_2} + \ldots + e^{B_n} = 0$$

is multiplied by Kp. Since p was chosen greater than A, $Kp \cdot A$

is an integer not divisible by p.

For each of the e^{Bm} $(1 \leq m \leq n)$,

$$Kp\ e^{Bm} = \left(\sum_{r = p-1}^{np+p-1} r!\ c_r \right) \left(\sum_{t=0}^{\infty} \frac{(B_m)^t}{t!} \right)$$

which when the series are multiplied yields an infinite series of the form

$$\sum_{r= p-1}^{np+p-1} c_r \left\{ B_m^{\ r} + rB_m^{\ r-1} + r(r-1)\ B_m^{\ r-2} + \ldots + r! + \frac{B_m^{\ r+1}}{r+1} + \frac{B_m^{\ r+2}}{(r+1)(r+2)} + \ldots \right\}.$$

However, this series may also be expressed as

$$F(Bm) + F'(Bm) + F''(Bm) + \ldots + F^{(np+p-1)}(Bm)$$

$$+ \sum_{r= p-1}^{np+p-1} c_r\ B_m^{\ r} \left\{ \frac{Bm}{r+1}\quad \frac{Bm^2}{(r+1)(r+2)} \quad \ldots \right\}.$$

We seek a bound for the latter part of this series and find

$$c_r Bm^r \left\{ \frac{Bm}{r+1} + \frac{Bm^2}{(r+1)\ (r+2)} \ldots \right\} \leq \left| c_r\ Bm^r \right| \left\{ \frac{|Bm|}{r+1} \quad \frac{|Bm|^2}{(r+1)(r+2)} \quad \ldots \right\}$$

$$\leq \left| c_r\ Bm^{\ r} \right| \sum_{t=0}^{\infty} \frac{|Bm|^t}{t!}.$$

This bound is dependent on m. However, for any m,

$$\sum_{r= p-1}^{np+p-1} \left| c_r\ Bm^r \right| e^{|Bm|} = e^{|Bm|} \sum_{r=p-1}^{np+p-1} \left| c_r \quad Bm^r \right|$$

$$\leq e^{|Bm|} \frac{|Bm|^{p-1}}{(p-1)!}\ C^{np+p-1} \left\{ (|Bm| + |B_1|)(|Bm| + |B_2|) \right.$$

$$+ \cdots (|Bm| + |Bn|)\}^p \quad .$$

But this is bounded by

$$e^{\bar{B}} \frac{\bar{B}^{p-1}}{(p-1)!} C^{np+p-1} \left\{(\bar{B} + |B_1|)(\bar{B} + |B_2|) \cdots (\bar{B} + |Bn|)\right\}^p$$

where $\bar{B}$ is the maximum of $|B_1|, |B_2|, \ldots, |Bn|$. This is of the form $P \cdot Q^p/(p-1)!$ where P and Q are independent of p and m. We have shown

$$Kp(A + e^{B_1} + e^{B_2} + e^{B_3} + \ldots + e^{B_n})$$
$$= KpA + \sum_{m=1}^{n} (F^{(p-1)}(Bm) + F^{(p)}(Bm) + \ldots + F^{(np+p-1)}(Bm))$$
$$+ \quad L$$

where $L \leq nP \cdot Q^p/(p-1)!$ and the preceding portion of the sum is a non-zero integer. Since p was restricted only in that it had to be prime and greater than A, n, C, and $|C^n B_1 B_2 \cdots Bn|$, such an equation may be derived for any p no matter how large. But choosing p sufficiently large so that

$$\left|\frac{nP \cdot Q^p}{(p-1)!}\right| < 1$$

would necessitate that the sum of a non-zero integer and a fraction less in absolute value than one be zero. It is possible to choose such a p since the series $\sum_{n=0}^{\infty} (Q^n/(n-1)!)$ converges for all Q and thus $\lim_{n \to \infty} (Q^n/(n-1)!) = 0$. The subsequence $\{Q^p/(p-1)!\}$ must have the same limit as $p \to \infty$. It is, however, contradictory that the sum of a non-zero integer and a fraction smaller than one should be zero. Thus no equation of the form

$$A + e^{B_1} + e^{B_2} + \dots + e^{B_n} = 0$$

can hold and the assumption that π is algebraic has been disproved.

Footnotes

[1] L. E. Dickson, "Constructions with Ruler and Compasses; Regular Pôlygons," Monographs on Topics of Modern Mathematics, ed. J.W.A. Young (New York:Longmans, Green and Co., 1932), p. 355.

[2] D.E. Smith, "The History and Transcendence of π,"in Young, p. 395.

[3] Smith in Young, p. 397.

[4] E.W. Hobson, "Squaring the Circle,"Squaring the Circle and other Monographs (New York: Chelsea Publishing Co., 1953), p. 42.

[5] Hobson, p. 44.

[6] L.E. Dickson, Elementary Theory of Equations(New York: John Wiley & Sons, Inc., 1914), p. 63.

Bibliography

(1) Cajori, Florian. A History of Mathematics. New York: Macmillan and Co., 1895.

(2) Dickson, L.E. Elementary Theory of Equations. New York: John Wiley & Sons, Inc., 1914.

(3) Hobson, E.W., et al. Squaring the Circle and Other Monographs. New York:Chelsea Publishing Co., 1953.

(4) Young, J.W.A., ed. Monographs on Topics of Modern Mathematics. New York: Longmans, Green and Co., 1932.