$$\mathbf{E}_{n+1} = A^n\mathbf{E}_0 + A^{n-1}\mathbf{U}_0 + A^{n-2}\mathbf{U}_1 + \cdots + \mathbf{U}_{n-1}. \tag{18}$$

From (18) it follows at once that the criterion for the stability of (17) is identical with that for (9), namely that Δx and Δy must be chosen so that $\alpha + \beta \leqq \frac{1}{2}$.

It may be briefly mentioned that the above analysis may be extended to the more general case of the boundary conditions $pT + q(\partial T/\partial n) = F(t)$ where p and q take on prescribed values along the boundary. It may also be mentioned that the above analysis may be extended to problems with cylindrical and spherical symmetry.

Yeshiva University
New York, New York: and
AVCO Corporation
Wilmington, Massachusetts

High Precision Calculation of Arcsin x, Arccos x, and Arctan x

By I. E. Perlin and J. R. Garrett

1. Introduction. In this paper a polynomial approximation for Arctan x in the interval $0 \leqq x \leqq \tan \pi/24$, accurate to twenty decimal places for fixed point routines, and having an error of at most 2 in the twentieth significant figure for floating point routines is developed. By means of this polynomial Arctan x can be calculated for all real values of x. Arcsin x and Arccos x can be calculated by means of the identities:

$$\text{Arctan} \frac{x}{\sqrt{1 - x^2}} = \text{Arcsin } x = \frac{\pi}{2} - \text{Arccos } x.$$

2. Polynomial Approximation for Arctan x. A polynomial approximation for the Arctangent is obtained from the following Fourier series expansion, given by Kogbetliantz [1], [2] and Luke [3].

$$\text{Arctan } (x \tan 2\theta) = 2 \sum_{i=0}^{\infty} \frac{(-1)^i(\tan \theta)^{2i+1}}{2i + 1} T_{2i+1}(x), \tag{2.1}$$

where $T_i(x)$ are the Chebyshev polynomials, i.e., $T_i(\cos y) = \cos (iy)$. The expansion (2.1) is absolutely and uniformly convergent for $|x| \leqq 1$ and $0 \leqq \theta < \pi/4$.

An approximating polynomial is obtained by truncating (2.1) after n terms. Thus,

$$P(x \tan 2\theta) = 2 \sum_{i=0}^{n-1} \frac{(-1)^i(\tan \theta)^{2i+1}}{2i + 1} T_{2i+1}(x). \tag{2.2}$$

The truncation error is

$$|\epsilon_T| \leqq \tan 2\theta\, (\tan \theta)^{2n} |x|. \tag{2.3}$$

Received December 11, 1959; in revised form, February 16, 1960. The work reported in this paper was sponsored by the Air Force Missile Development Center, Alamogordo, New Mexico.

When $x \tan 2\theta$ is replaced by M and $T_{2i+1}(x)$ are expressed in terms of x, (2.2) becomes

$$P(M) = \sum_{r=0}^{n-1} \frac{(-1)^r B_r M^{2r+1}}{2r+1} \tag{2.4}$$

where

$$B_r = (1 - \tan^2 \theta)^{2r+1} \sum_{k=0}^{n-r-1} \binom{2r+k}{k} (\tan \theta)^{2k}.$$

With a choice of $n = 9$ and $\tan \theta = \tan \pi/48$, the following polynomial approximation for Arctan x is obtained.

$$P(x) = a_1 x + a_3 x^3 + \cdots + a_{17} x^{17} \tag{2.5}$$

where

$a_1 =$	1.0					
$a_3 =$	−0.33333	33333	33333	33160	7
$a_5 =$	0.19999	99999	99998	24444	8
$a_7 =$	−0.14285	71428	56331	30652	9
$a_9 =$	0.11111	11109	07793	96739	3
$a_{11} =$	−0.09090	90609	63367	76370	73
$a_{13} =$	0.07692	04073	24915	40813	20
$a_{15} =$	−0.06652	48229	41310	82779	05
$a_{17} =$	0.05467	21009	39593	88069	41

The truncation error is $|\epsilon_T| < 6 \cdot 10^{-22} |x|$.

3. Procedure for Calculation Arctan x. Subdivide the interval $(0, \infty)$ into seven intervals as follows: $0 \leqq u < \tan \pi/24$, $\tan [(2j-3)\pi/24] \leqq u < \tan [(2j-1)\pi/24]$ for $j = 2, 3, 4, 5, 6$, and $\tan 11\pi/24 \leqq u < \infty$. For $|x|$ on the first interval use (2.5). For $|x|$ on the $(j+1)$st interval, $(j = 1, 2, 3, 4, 5)$, the formula employed is

$$\text{Arctan} |x| = \frac{j\pi}{12} + \text{Arctan}\, t_j, \tag{3.1}$$

where

$$t_j = \frac{|x| - \tan \frac{j\pi}{12}}{1 + |x| \tan \frac{j\pi}{12}}$$

is used to obtain a value t_j such that $|t_j| \leqq \tan \pi/24$. Arctan t_j is calculated by (2.5) and Arctan $|x|$ from (3.1). When $|x|$ is in the seventh interval

$$\text{Arctan}\,|x| = \frac{\pi}{2} - \text{Arctan}\,\frac{1}{|x|}, \tag{3.2}$$

and

$$\frac{1}{|x|} \leqq \tan\frac{\pi}{24}.$$

The constants $\tan j\pi/24$, $(j = 1, 2, \cdots, 11)$ and $\pi/2$ are:

$\tan \pi/24 =$	0.13165 24975 87395 85347 2
$\tan \pi/12 =$	0.26794 91924 31122 70647 3
$\tan \pi/8 =$	0.41421 35623 73095 04880 2
$\tan \pi/6 =$	0.57735 02691 89625 76450 9
$\tan 5\pi/24 =$	0.76732 69879 78960 34292 3
$\tan \pi/4 =$	1.00000 00000 00000 00000 0
$\tan 7\pi/24 =$	1.30322 53728 41205 75586 8
$\tan \pi/3 =$	1.73205 08075 68877 29352 7
$\tan 3\pi/8 =$	2.41421 35623 73095 04880 2
$\tan 5\pi/12 =$	3.73205 08075 68877 29352 7
$\tan 11\pi/24 =$	7.59575 41127 25150 44052 6
$\pi/2 =$	1.57079 63267 94896 61923 1.

4. Error Analysis.

A. General.

Errors arising from calculations by a computer may be classified into three categories according to Householder [4], namely: (1) truncation errors, (2) propagated errors, and 3) round-off errors. For the propagated error, if x and y are approximated by x' and y', respectively, and the errors in each are denoted by $\epsilon(x)$ and $\epsilon(y)$, then:

$$|\epsilon(x \pm y)| \leqq |\epsilon(x)| + |\epsilon(y)|,$$

$$\left|\frac{\epsilon(xy)}{x'y'}\right| \leqq \left|\frac{\epsilon(x)}{x'}\right| + \left|\frac{\epsilon(y)}{y'}\right|, \qquad x'y' \neq 0$$

$$\left|\frac{\epsilon\left(\dfrac{x}{y}\right)}{\dfrac{x'}{y'}}\right| \leqq \frac{\left|\dfrac{\epsilon(x)}{x'}\right| + \left|\dfrac{\epsilon(y)}{y'}\right|}{1 - \left|\dfrac{\epsilon(y)}{y'}\right|}, \qquad x'y' \neq 0.$$

For round-off error it is assumed that rounding is accomplished in the following manner. If λ digits are to be retained and the $(\lambda + 1)$st digit is $\geqq 5$, add one to the preceding digit; otherwise do not change the preceding digit. With this convention the round-off error in fixed point arithmetic is easily determined. For floating point arithmetic use is made of the following result. Let

$$x = (x_1\beta^{-1} + x_2\beta^{-2} + \cdots + x_\lambda\beta^{-\lambda})\beta^\rho, \qquad x_1 \neq 0,$$

and

$$y = (y_1\beta^{-1} + y_2\beta^{-2} + \cdots + y_\lambda\beta^{-\lambda})\beta^\sigma, \qquad y_1 \neq 0,$$

and $x \oplus y$ represent addition, subtraction, multiplication or division. The round-off error in $x \oplus y$, $\epsilon(x \oplus y)$, is given by

$$|\epsilon(x \oplus y)| \leqq |x \oplus y| \left(\frac{\beta}{2}\right) \beta^{-\lambda}.$$

Another result useful in floating point arithmetic is the following: If $\left|\frac{x' - x}{x'}\right| < a\beta^{-\tau}$, $(\tau \geqq 1)$, then x' differs from x by at most a units in the τth significant digit. The preceding results are easily established.

B. Errors in Arctan x.

a) $$0 \leqq x < \tan \frac{\beta}{24}.$$

For fixed point arithmetic it shall be assumed that twenty-one decimals are used. The truncation error $|\epsilon_T| < 8 \cdot 10^{-23}$. The error ϵ_p due to errors in the coefficients in (2.5) is $|\epsilon_p| < 1.2 \cdot 10^{-24}$. The round-off error ϵ_R is $|\epsilon_R| < 6.35 \cdot 10^{-22}$. Hence, the total error is less than $8 \cdot 10^{-22}$, and the calculated value of Arctan x is accurate to at least twenty decimal places.

For floating point arithmetic using twenty-one significant digits, $|\epsilon_T| < 6 \cdot 10^{-22}x$, $|\epsilon_p| < 10^{-22}x$, and $|\epsilon_R| < 1.03 \cdot 10^{-20}x$. Hence, $|\epsilon(\text{Arctan } x)| < 1.1 \cdot 10^{-20}x$, and

$$\left|\frac{\epsilon(\text{Arctan } x)}{\text{Arctan } x}\right| < 1.2 \cdot 10^{-20}.$$

Thus the calculated value of Arctan x differs from the true value by at most two units in the twentieth significant digit.

b) $$\tan \frac{\pi}{24} \leqq x < \tan \frac{11\pi}{24}.$$

For fixed point arithmetic $|\epsilon(t_j)| < 7.8 \cdot 10^{-22}$. The propagated error in Arctan t_j due to this error in t_j is $|\epsilon(\text{Arctan } t_j)| < 7.8 \cdot 10^{-22}$. The total error in Arctan t_j is then $|\epsilon(\text{Arctan } t_j)| < 1.5 \cdot 10^{-21}$. The error in Arctan x is then $|\epsilon(\text{Arctan } x)| < 2.5 \cdot 10^{-21}$, and the calculated value of Arctan x is accurate to twenty decimal places.

For floating point arithmetic

$$\left|\frac{\epsilon(\text{Arctan } x)}{\text{Arctan } x}\right| < 7.2 \cdot 10^{-20},$$

and the calculated value of Arctan x differs from the true value by at most eight units in the twentieth significant digit.

c) $$\tan \frac{11\pi}{24} \leqq x < \infty.$$

For fixed point arithmetic $|\epsilon(\text{Arctan } x)| < 2.3 \cdot 10^{-21}$, and hence the calculated value is accurate to twenty decimal places.

For floating point arithmetic

$$\left|\frac{\epsilon(\text{Arctan } x)}{\text{Arctan } x}\right| < 1.25 \cdot 10^{-20},$$

and hence the calculated value differs from the true value by at most two units in the twentieth significant digit.

C. Errors in Arcsin x and Arccos x.

The error for floating point arithmetic using twenty-one significant digits will be given. Arcsin x will be calculated by means of

$$\text{Arcsin } x = \text{Arctan } \frac{x}{\sqrt{1 - x^2}}.$$

The quantity $1 - x^2$ is calculated by means of $1 - x^2 = (1 - x)(1 + x)$. Then

$$\left|\frac{\epsilon(\text{Arctan } x)}{\text{Arcsin } x}\right| < 10^{-19},$$

and Arcsin x will be correct to within one unit in the nineteenth significant figure.

The error for Arccos x is similar except that a round-off error due to subtraction is introduced. This error does not affect the conclusion that Arccos x will have been obtained correctly to within one unit in the nineteenth significant figure.

5. Conclusions. From the standpoint of machine application the procedure given is economical and yields precise results. It uses only twenty stored constants; the calculation of Arctan x requires a maximum of only eleven multiplications and one division; the calculation of Arcsin x and Arccos x requires an additional multiplication, division, and square root.

Rich Electronic Computer Center
Engineering Experiment Station
Georgia Institute of Technology
Atlanta, Georgia

1. E. G. Kogbetliantz, "Computation of Arctan N for $-\infty < N < \infty$ using an electronic computer," *IBM Jn. Res. and Dev.*, v. 2, 1958, p. 43–53.
2. E. G. Kogbetliantz, "Computation of Arcsin N for $0 < N < 1$ using an electronic computer," *IBM Jn. Res. and Dev.*, v. 2, 1958, p. 218–222.
3. Y. L. Luke, "On the computation of Log Z and Arctan Z," *MTAC*, v. 11, 1957, p. 16–18.
4. A. S. Householder, *Principles of Numerical Analysis*, McGraw Hill Book Company, New York, 1953.

The Calculation of Toroidal Harmonics

By A. Rotenberg

1. Introduction. It is the purpose of this note to describe the mathematical techniques employed in a code [5] for the IBM 704 to calculate toroidal harmonics (associated Legendre functions of half-integral order). We use recurrence techniques similar to those used by Goldstein and Thaler [1] in calculating Bessel func-

Received October 27, 1959.