On the Numerical Treatment of Heat Conduction Problems with Mixed Boundary Conditions

By Arnold N. Lowan

Abstract. The two-dimensional problem of heat conduction in a rectangle where the temperature is prescribed over a portion of the boundary while the temperature gradient is prescribed over the remainder of the boundary, may be treated numerically by replacing the differential equation of heat conduction and the equations expressing the given initial and boundary conditions by their difference analogs.

Received September 15, 1959.
and solving the resulting system. It is shown that if the scheme is to be stable the
intervals $\Delta x$ and $\Delta y$ must be chosen so that $k\Delta t/(\Delta x)^2 + k\Delta t/(\Delta y)^2 \leq \frac{1}{2}$.

Consider the two-dimensional problem of heat conduction in a rectangle, Figure 1,
when the temperature is prescribed over the thin line portion of the boundary,
while the temperature gradient is prescribed over the heavy line portion of the
boundary. This is a typical problem with “mixed” boundary conditions and should
not be confused with the considerably simpler problem when the temperature is
prescribed over certain complete sides of the rectangle, while the temperature
gradient is prescribed over the remaining sides. As far as the writer is aware no
analytical solution of the mixed boundary value problem above formulated (or
of the analogous problem for the cylinder) is to be found in the literature. We must
therefore (if interested in numerical answers) resort to the alternative of substituting
for the differential equation of heat conduction and for the equations expressing
the initial and boundary conditions their appropriate difference analogs, and solving
the resulting system.

The mathematical formulation of the problem is as follows:

\[
\begin{align*}
(1) \quad \frac{\partial T}{\partial t} &= k \left( \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) \quad 0 \leq x \leq a, 0 \leq y \leq b, t > 0 \\
(2) \quad T(x, y, 0) &= f(x, y) \\
(3) \quad T(0, y, t) &= 0 \quad 0 \leq y \leq b \\
(4) \quad T(x, 0, t) &= 0 \quad 0 \leq x \leq c_1 \\
(5) \quad \left[ \frac{\partial}{\partial y} T(x, y, t) \right]_{y=0} &= 0 \quad c_1 \leq x \leq a \\
(6) \quad \left[ \frac{\partial}{\partial x} T(x, y, t) \right]_{x=a} &= 0 \quad 0 \leq y \leq b \\
(7) \quad \left[ \frac{\partial}{\partial y} T(x, y, t) \right]_{y=b} &= 0 \quad c_2 \leq x \leq a \\
(8) \quad T(x, b, t) &= 0 \quad 0 \leq x \leq c_2 
\end{align*}
\]

where for the sake of simplicity we have at first assumed that the prescribed
temperature and temperature gradient are $= 0$. The difference analogs of the above
equations are:

\[
\begin{align*}
(1^*) \quad T_{h,k,n+1} &= \beta T_{h,k-1,n} + \alpha T_{h-1,k,n} + (1 - 2\alpha - 2\beta) T_{h,k,n} \\
&\quad + \alpha T_{h+1,k,n} + \beta T_{h,k+1,n} \quad h = 1, 2, 3, \cdots M, k = 1, 2, 3 \cdots N \\
(2^*) \quad T_{h,k,0} &= f(h\Delta x, k\Delta y) \\
(3^*) \quad T_{0,k,n} &= 0 \quad 1 \leq k \leq N \\
(4^*) \quad T_{h,0,n} &= 0 \quad 1 < h < c_1/\Delta x \\
(5^*) \quad T_{h,1,n} &= T_{h,0,n} \quad c_1/\Delta x \leq h \leq a/\Delta x 
\end{align*}
\]
\( T_{M+1,k,n} = T_{M,k,n} \quad 1 \leq k \leq N \)

\( T_{h,N+1,n} = T_{h,N,n} \quad \frac{c_t}{\Delta x} \leq h \leq \frac{a}{\Delta x} \)

\( T_{h,N+1,n} = 0 \quad 1 \leq h \leq \frac{c_t}{\Delta x} \)

where

\[
T_{h,k,n} = T(h\Delta x, k\Delta y, n\Delta t); \quad \alpha = \frac{k\Delta t}{(\Delta x)^2}; \quad \beta = \frac{k\Delta t}{(\Delta y)^2}
\]

\( \Delta x = a/(M + 1) \) and \( \Delta y = b/(N + 1) \).

It will be convenient to consider the \( MN \) temperatures \( T_{h,k,n} \) with \( h = 1, 2, 3, \ldots M \) and \( k = 1, 2, 3, \ldots N \) as the components of an \( M \times N \) - dimensional vector to be denoted by \( T_n \). It will also be convenient to replace the two subscripts \( h \) and \( k \) identifying the lattice point \( P_{hk} \) (i.e., the point with coordinate \( x = h\Delta x \) and \( y = k\Delta y \)) by the single subscript \( p \) running from \( p = 1 \) to \( p = MN \) with the understanding that for the \( M \) lattice points corresponding to \( k = 1 \), \( p \) runs from \( 1 \) to \( M \), for the next set of \( M \) lattice points corresponding to \( k = 2 \) \( p \) runs from \( M + 1 \) to \( 2M \), \ldots etc. The system of \( MN \) equation (1*) may then be written in the matrix—vector form

\[
T_{n+1} = AT_n.
\]

As is well known, to prove the stability of (9), it suffices to prove that if \( S \) denotes the largest of the sums of absolute values of the elements of the rows of \( A \) then \( S \leq 1 \).

Let \( \Omega \) denote the set of lattice points closest to the boundary. It is clear that if the difference equation (1*) is applied to lattice points lying inside of \( \Omega \), the resulting equation (which is in fact equation (1*) itself) has the five non-vanishing coefficients \( \beta, \alpha, 1 - 2\alpha - 2\beta, \alpha \) and \( \beta \). If we assume \( 1 - 2\alpha - 2\beta \geq 0 \) or \( \alpha + \beta \leq \frac{1}{2} \) it is clear that the sum of the absolute values of the coefficients is \( = 1 \). We shall show that if (1*) is applied to lattice points belonging to the set \( \Omega \), the resulting equation may be characterized by the fact that the sum of absolute values of the coefficients is smaller than unity. Consider for instance the form taken by (1*) when applied to a point of \( \Omega \) such that \( h\Delta x \leq c_1 \). Since for such a point \( T_{h,0} = 0 \) the resulting equation has the four non-vanishing coefficients \( \alpha, 1 - 2\alpha - 2\beta, \alpha \) and \( \beta \). If again we assume \( 1 - 2\alpha - 2\beta > 0 \), it follows that the sum of the absolute values of the coefficients is \( = 1 - \beta < 1 \). In an entirely similar manner it is shown that if (1*) is applied to lattice points in \( \Omega \) for which \( h\Delta x \geq c_2 \) and the boundary condition (5*) is taken into account, the sum of the absolute values of the coefficients of the resulting equation is \( = 1 - \alpha < 1 \). Similar conclusions may be drawn in the case of all lattice points in \( \Omega \). Thus the quantity \( S \) previously defined is \( = 1 \). The stability of the difference scheme under consideration is thus proven, provided that the intervals \( \Delta x \) and \( \Delta t \) are chosen so that

\[
\alpha + \beta = \frac{k\Delta t}{(\Delta x)^2} + \frac{k\Delta t}{(\Delta y)^2} \leq \frac{1}{2}.
\]

In the above discussion we assumed that the prescribed temperature is \( = 0^\circ C \) on the thin line portion of the boundary. If instead, nonvanishing temperatures are
prescribed on this portion, the criterion of stability is the same as before, since the
error vector satisfies \((1^*)\) and evidently is \(= 0\) on the portion of the boundary in
question.

We shall now discuss the modifications in the above analysis if the prescribed
temperature gradient does not vanish. Let the above boundary conditions \((5), (6)\)
and \((7)\) be replaced by

\[
\frac{\partial T}{\partial y} = \phi_1(x, t) \quad c_1 \leq x \leq a, y = 0
\]

\[
\frac{\partial T}{\partial x} = \Phi(y, t) \quad x = a, 0 \leq y \leq b
\]

\[
\frac{\partial T}{\partial y} = \phi_2(x, t) \quad c_2 \leq x \leq a, y = b.
\]

The difference analogs of the last three equations are

\[
(10^*) \quad T_{h,0,n} = T_{h,1,n} + \Delta y \cdot \phi_1(h \Delta x, n \Delta t) \quad c_1/\Delta x \leq h \leq a/\Delta x
\]

\[
(11^*) \quad T_{M+1,k,n} = T_{M,k,n} - \Delta x \cdot \Phi(k \Delta y, n \Delta t) \quad 1 \leq k \leq N
\]

\[
(12^*) \quad T_{h,N+1,n} = T_{h,N,n} - \Delta y \cdot \phi_2(h \Delta x, n \Delta t) \quad c_2/\Delta x \leq h \leq a/\Delta x
\]

If the difference equation \((1^*)\) is applied to the lattice points for which the last
three equations hold, we ultimately get

\[
(13) \quad T_{h,1,n+1} = \alpha T_{h-1,1,n} + (1 - 2\alpha - 2\beta) T_{h,1,n} + \alpha T_{h+1,1,n} + \beta T_{h,2,n} + U_{h,1,n} \quad c_1/\Delta x \leq h \leq a/\Delta x
\]

\[
(14) \quad T_{M,k,n+1} = \beta T_{M,k-1,n} + \alpha T_{M-1,k,n} + (1 - 2\alpha - 2\beta) T_{M,k,n} + \beta T_{M,k+1,n} + U_{M,k,n} \quad 1 \leq k \leq N
\]

\[
(15) \quad T_{h,N,n+1} = \beta T_{h,N-1,n} + \alpha T_{h-1,N,n} + (1 - 2\alpha - 2\beta) T_{h,N,n} + T_{h+1,N,n} + U_{h,N,n} \quad 1 \leq k \leq N
\]

where we have put

\[
U_{h,1,n} = \beta \Delta y \cdot \phi_1(h \Delta x, n \Delta t) \quad c_1/\Delta x \leq h \leq a/\Delta x
\]

\[
U_{M,k,n} = -\alpha \Delta x \cdot \Phi(k \Delta y, n \Delta t) \quad 1 \leq k \leq N
\]

\[
U_{h,N,n} = -\beta \Delta y \cdot \phi_2(h \Delta x, n \Delta t) \quad c_2/\Delta x \leq h \leq a/\Delta x.
\]

The system of \(MN\) equations obtained by applying \((1^*)\) to the \(MN\) lattice
points may be written in the form

\[
(17) \quad T_{n+1} = AT_n + U_n
\]

where \(U_n\) is a vector whose non-vanishing components are defined in \((16)\) and whose
remaining components are \(= 0\). Since the error vector \(E_n\) satisfies the difference
equation \((17)\) it is readily seen that
(18) \[ E_{n+1} = A^nE_0 + A^{n-1}U_0 + A^{n-2}U_1 + \cdots + U_{n-1}. \]

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

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 \( 0 \leq x \leq \tan \pi/24 \), 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:

\[ \arctan \frac{x}{\sqrt{1-x^2}} = \arcsin x = \frac{\pi}{2} - \arccos x. \]

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].

\[ \arctan(x \tan 2\theta) = 2 \sum_{i=0}^{n} \frac{(-1)^i(\tan \theta)^{2i+1}}{2i+1} T_{2i+1}(x), \]

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 \( |x| \leq 1 \) and \( 0 \leq \theta < \pi/4 \).

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

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

The truncation error is

\[ |e_r| \leq \tan 2\theta \ (\tan \theta)^{2n} |x|. \]