Partial differential equations

By L. Fox

This expository paper discusses the present state of our ability to solve partial differential equations. It considers the success achieved in the production of new techniques, machine-oriented techniques, error analysis, mathematical theorems and the solution of practical problems, and contrasts this with corresponding work in the field of linear algebra. Outstanding problems include a determination of the error of finite-difference approximations, the automatic machine production of finite-difference formulae in complicated regions, the smoothing of physical data, and the classification of equations for computing-machine library routines.

I want to make some observations on what seems to me to be the general state of the art and science of solving partial differential equations, and especially what is lacking in that field. I am concerned particularly with what is sometimes called "numerical" solution and sometimes "approximate" solution, and which implies an approach other than the evaluation, at selected values of the independent variables, of the exact solution expressed in closed mathematical form, such as a single function, an infinite or doubly-infinite series, and so on. The numerical approach is of course almost always necessary, because the mathematical solution can rarely be obtained with known techniques.

All research in numerical analysis can be put into one of four or five categories, and we can approach our subject conveniently from these points of view. My categories are

(i) New methods.
(ii) Machine-oriented techniques.
(iii) Error analysis.
(iv) Theorems and mathematical analysis.
(v) Application to practical problems.

Linear algebra

In linear algebra, for example, I would class the original Givens method for finding the latent roots of matrices as a brand-new method, and it has been followed by other work on similarity and unitary transformations, to reduce the matrix to one of simpler form. Some of these methods are more effective than the Givens technique, but they owe their existence to Givens' idea. Under (ii) we have recent analysis of the details of various techniques, to make them as automatic as possible and to avoid, for example, the loss of efficiency in the use of a two-level store in computing machines. An analysis of Givens' method, for example, by Rollett and Wilkinson in a recent paper in this *Journal* showed that it was possible, by reorganizing the calculation to avoid too frequent transfers to and from the backing store, to save some 85% of Mercury time in the treatment of matrices of large orders.

Under (iii) we have the work of Von Neumann and Goldstine, Turing, Wilkinson and others in the assessment of errors in a variety of processes in linear algebra, together with an investigation of stability, for example in avoiding divisions by small numbers with consequent large rounding errors. In (iv) we have various results, such as investigations into bounds for eigenvalues, the relations between various known direct and iterative methods, the fact that a symmetric positive-definite matrix A can be decomposed into LL' with perfect safety without interchanges, and so on.

Finally, the various pieces of knowledge accumulated over the years have given us a fair mastery over any problem likely to arise in the field of linear algebra, from which, incidentally, I am excluding linear programming in which there are several unsolved problems of technique and error analysis.

I mention linear algebra because (*a*) it has received considerable attention since the automatic calculating machine arrived and (*b*) because it has considerable bearing on the subject of partial differential equations. Indeed, when we examine our various categories, we find very little advance in our subject in its own right. In elliptic equations in particular the major advance has been in the extra efficiency of the solution of the corresponding algebraic problem which arises through the use of finite differences.

Partial differential equations

In fact the use of finite differences is still our main technique, at least for equations of parabolic and elliptic types, and I know of no brand-new method of general application. Apart from finite-difference methods we can think of suggestions, none of them particularly new, such as a linear combination of trial solutions for linear differential equations, with weights adjusted to satisfy certain conditions, perhaps the boundary conditions at various points, or even the differential equation at various points. These are the so-called methods of *collocation*, and may be extended with principles of variation or least-squares to give methods like those of Rayleigh, Ritz and Galerkin. They are valuable, I repeat, only in certain restricted classes of problem, and in any case it is not easy to estimate the error of an approximation with a finite number of terms.

Certain methods of a brand-new type have been proposed, for example the kernel-function method of Bergman and Schiffer, the suggestion to replace the differential equation by an integral equation, and the "hyper-circle" method of Synge, but there is very little numerical evidence of successful application, and indeed they receive hardly a mention in the most recent book on partial differential equations by Forsythe and Wasow. Certainly I know nothing about these processes and would welcome any knowledgeable comment in the discussion.

For hyperbolic equations in two dimensions, likewise, we have nothing particularly new. In addition to finite-difference methods, of course, we can here integrate along characteristics, and this in my view is usually the better approach, since the characteristics separate regions of discontinuity which would seriously affect the finite-difference formulae. The characteristic method, moreover, has recently been extended to problems in three dimensions by Butler and others, and indeed this is a bright light in my gloomy account of lack of progress.

Machine-oriented techniques

In category (ii), of course, there have been substantial, elegant and exciting advances in methods, particularly for the elliptic equation, of solving the algebraic problem, called by Forsythe and Wasow the "discretized" problem. The algebraic problem here involves matrices of large order but of very particular forms, with large blocks of zero elements. If the problem is small enough we can solve it by a direct method, and for various purposes and reasons this is very attractive. In general, however, the small size of our machine store is a prohibitive factor, and iteration is necessary. In place of the paper-and-pencil relaxation method of Southwell, with its emphasis on refined tricks and dodges which are difficult to program, and which in any case do not use the modern computer to best advantage, we have developed some extremely fast and effective methods of automatic iteration. The convergence of the simple automatic Liebmann process has been improved by the "successive point over-relaxation" methods of Young and Frankel, and we have progressed through line iteration and Chebyshev acceleration to the systematic alternating-direction methods of Douglas and others, applicable both to elliptic and parabolic equations and to problems in more than two independent variables.

At the same time a number of beautiful theorems have been discovered, part of my category (iv), which indicate the circumstances under which these various iterative devices will have optimum success. We have all heard, for example, of the phrases "property (A)" and "consistent ordering" in connection with Young's theory of successive over-relaxation, and we know that the simple five-point formula for the Laplace operator satisfies the requirements, whereas the more accurate nine-point formula is satisfactory in this respect for line iteration but not for point iteration. Other results of this kind are coming to light every day.

Error analysis

When we come to category (iii), however, we find a rather mixed state of affairs. We know a lot about the stability of step-by-step processes of explicit and implicit kinds, through the pioneering Fourier-type analysis of Von Neumann and its extensions by Lax and Richtmyer and others, and of the similar use of matrix algebra in this connection. We even have some theorems, in the associated category (iv), connecting stability with convergence. Generally speaking, if stability is assured (that is if the asymptotic behaviour is right), and if the finite-difference equations are compatible with the differential equation (that is if we can choose our mesh lengths, or certain parameters thereof, in such a way that the local truncation error vanishes with the mesh size), then we also have convergence—that is the finite-difference solution tends correctly in the limit to the true solution at all points.

But we are very considerably in the dark about the difference between the true solution and that of the algebraic problem at any fixed finite intervals. I sometimes think that the excitement of the discoveries about the algebraic problem, particularly for elliptic equations and particularly in America, has caused its exponents to lose complete sight of the fact that our real problem is differential, and not algebraic. Certainly comparatively little attention has been paid to this point during the last decade of great algebraic discovery. I suppose the problems are harder, and those mathematicians, still rather small in number, who have become seriously interested in numerical analysis have found that the algebraic field produces greater immediate reward.

The problems are certainly formidable. What, for example, is the nature of the function g in the expression

$$f(x, y) - f(x, y, h) = g_1(h), \tag{1}$$

or in the corresponding

$$\lambda - \lambda(h) = g_2(h) \tag{2}$$

in the eigenvalue problem, where $f(x, y)$ and λ are the true solutions of the differential systems, and $f(x, y, h)$ and $\lambda(h)$ are approximations obtained from truncated finite-difference formulae with interval h? The basis of Richardson's "deferred approach to the limit" rests on such knowledge, and particularly on the assumption that the right-hand side has a convergent Taylor's series or at least a semi-convergent series about $h = 0$. But this is true only in particularly simple cases. It can be proved, for example, that for the solution of

$$\nabla^2 f + \lambda f = 0 \tag{3}$$

in the interior of a rectangle, with f vanishing on the boundaries, the use of the five-point formula gives such a series in even powers of h, and h^2-extrapolation is possible. For the nine-point formula we have h^4- extrapolation. If there is an internal angle greater

than π, however, as in the shape

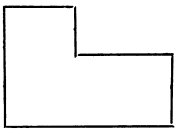

Fig. 1

of the famous L-shaped membrane (Fig. 1) this hypothesis is certainly not true, and as far as I know the expression for the error has not been found either in $f(x, y, h)$ or in $\lambda(h)$.

A re-entrant corner is not the only danger. For example we have similar difficulties in the neighbourhood of the points $x = 0, t = 0; x = 1, t = 0$ in the parabolic equation

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial f}{\partial t}, \tag{4}$$

with initial conditions on $t = 0$ and boundary conditions on the lines $x = 0$ and $x = 1$.

If

$$\lim_{x \to 0} f(x, 0) \neq \lim_{t \to 0} f(0, t) \tag{5}$$

we are in serious trouble, and if the limiting higher derivatives do not satisfy the differential equation and its derived forms at the corner we have similar troubles in decreasing degree.

With discontinuous boundary conditions even on a straight boundary, and particularly with non-linear equations, the solution may behave in a way which prohibits accurate representation in particular regions by finite-difference equations, and therefore nullifies the deferred approach, whose success demands a local truncation error of the same form at every point.

It is interesting to note, in passing, that discontinuities in the specified conditions for hyperbolic equations are often much less serious. The discontinuities may be propagated only along characteristics, and the solution may have quite different forms in regions separated by certain characteristics. Each within its own region may be perfectly well behaved, to such an extent that the integration along characteristics with a simple trapezoidal rule can be supported quite correctly with h^2-extrapolation. There may, of course, be other difficulties with the characteristic method, such as the presence of interfaces, the development of shocks, and so on, about which considerable work is being done, particularly at Aldermaston, and of which some has been published in the book of the 1961 Summer School at Oxford (Fox, 1962).

Returning to the parabolic and elliptic case, and particularly to the latter, how are we going to estimate the accuracy of our computed result? I assume, without apology, that this is desirable, though its practice is not widespread in these days of automatic computation! There are about four possible methods. First, we can sometimes prove, in the absence of rounding error (and this I ignore completely since in stable methods its effect is negligible) that the maximum difference between the true solution and the finite-difference solution with a particular formula is less than some known multiple of the largest value M of some derivatives. This possibility, and in particular a knowledge of M deduced from the data alone, rather than from numerical differentiation of the approximate solution, occurs so rarely that it can be discounted.

Second, we might try to use the deferred approach to the limit. This also is unsound unless we can determine in advance the nature of the expression for the error. Without this knowledge the method is dangerous, and I know of several cases, for example in the determination of the eigenvalues for the L-shaped membrane, in which extrapolations have given consistent but wrong results.

Third, we can decrease the interval successively until the solution stops changing. This is sometimes the only feasible and sound method, for example in cases of many singularities or discontinuities, such as at interfaces in nuclear reactor problems. But it is extremely laborious and time-consuming, and even dangerous. Again I know of cases in which, for the eigenvalue problem, successive intervals have given the same result to a certain precision, whereas the true solution differs quite substantially therefrom.

Fourth, we can apply a method which is very successful in solving the linear algebraic simultaneous equations

$$Ax = b. \tag{6}$$

From an alleged solution x_0 we compute the residual vector

$$r = b - Ax_0, \tag{7}$$

and find a correction x_1 to x_0 from the equation

$$Ax_1 = r. \tag{8}$$

In a linear differential equation we are trying to solve something like

$$k(D)f = g, \tag{9}$$

where $k(D)$ is a differential operator and g a function of the independent variables. Our first approximation satisfies

$$k(D)f_0 = g + \epsilon, \tag{10}$$

and if we can find ϵ we can correct our approximation with the addition of an f_1 which satisfies

$$k(D)f_1 = \epsilon \tag{11}$$

and the homogeneous versions of the various boundary conditions. For non-linear equations we can obtain linear equations for first-order corrections.

Now can we find the function ϵ? I assert that if we cannot find ϵ with reasonable accuracy at every point, then we have not solved our problem with any degree of confidence. Indeed the determination of ϵ will involve derivatives of all orders up to that of the differential equation, and at least some of these will be the important quantities in practical applications. The standard method of finding ϵ uses formulae for numerical differentiation, and this is usually safe in regions where the differences converge, so that the function can be approximated by interlacing polynomials, and with safe-

guards about possible masking of periodic behaviour. Near singularities, such as the re-entrant corner in the L-membrane, the differences will behave in a way which indicates the applicability or not of finite-difference formulae, and we may have to reduce the interval in a region whose boundaries the behaviour of the function will specify for us. Alternatively, we might use the methods of Motz and Woods, extended recently by Walsh, in which we abandon finite differences altogether in the offending region, and use formulae deduced from a linear combination of special functions which satisfy the differential equation and boundary conditions in the region of the corner. This is an extremely useful method, and has been applied with success by Walsh to the L-shaped eigenvalue problem, with a relatively large interval and a small fraction of the computing time involved in the use of finite differences at the necessary small interval. It is the only method I know, incidentally, which gives the eigenfunction accurately in the neighbourhood of the corner.

In the equation

$$\nabla^2 f + \lambda f = 0, \qquad (12)$$

in two Cartesian co-ordinates, and with zero values of f on the lines meeting at the re-entrant corner, the appropriate functions are the Bessel functions $J_{\frac{1}{2}n}(r\sqrt{\lambda}) \sin(n\theta)$, where the corner is the origin of polar co-ordinates (r, θ). Unfortunately it is not always easy, in more difficult cases, to find functions which satisfy all our requirements, and this is another field of research in which success would be extremely valuable. It belongs to my category (iv).

The difference correction

The method of the difference correction, which I advocated several years ago for both ordinary and partial differential equations (Fox, 1947), is of course a particular technique of the kind I have been talking about, involving the computation of the perturbation term ε. This idea does not seem to have penetrated deeply into the literature of automatic computation, and indeed is not mentioned at all by Forsythe and Wasow. I am more puzzled than hurt by this neglect. Certainly we have to do some differencing, involving extra programming, extra space, and some difficulties in automatic inspection of differences, but machines are getting larger and programming easier (or so everybody tells me), and if we are concerned with accuracy, as we certainly should be, I should have thought that something like this was essential.

In passing I would like to comment on a remark about this method made by Dr. Wilkes in his review of the second edition of *Modern Computing Methods*. He advocated, for various reasons, the direct inclusion of higher differences in the adopted form of the Lagrangian formulae for the determination of the first approximation, which would avoid the differencing and indeed give in one application a very correct result. After some examination I think I would accept this proposal for the solution of integral equations, and for some parabolic equations, but for ordinary differential equations and some partial differential equations this approach has some inconveniences. For any step-by-step process, particularly in ordinary differential equations, it is dangerous to use a recurrence formula of greater order than that arising naturally from the differential equation: the result is almost always "strong instability," in the phraseology of Dahlquist. This is not necessarily true for boundary-value problems, either ordinary or partial, and even for parabolic equations of not too complicated form we have found it possible to incorporate the difference correction or at least the major part of it into a Lagrangian form, with interesting effects on stability criteria. But with two-point problems in ordinary differential equations, and with elliptic partial differential equations, the resulting Lagrangian formulae give rise to matrices of less compact form, so that not only is the arithmetic of both direct and iterative methods of solution more complicated and lengthy, but we may also destroy the valuable property (A) which may be needed for the production of optimum accelerating devices. In any event the fundamental point remains valid, in all cases, that we do not know what order of Lagrangian formula to take until we know the solution, and for any approximate solution the determination of our error function ϵ is still necessary.

Practical problems

Finally, I must make two comments about my category (v)—practical problems. In most cases, particularly for elliptic problems and perhaps especially for the free boundary case, in which some extra condition is imposed which will help to fix the position of some part of the boundary, the boundaries are not often of conveniently simple shape. Indeed the automatic machine production of appropriate finite-difference or other formulae, for use near a curved boundary, is a formidable undertaking, and I have seen no research of any great value in this field, even in two dimensions. The adjusting of the boundary to a more simple shape must, I think, have a dangerous influence on the result, except possibly in the eigenvalue problem, for which there are certain exclusion and inclusion theorems due to Polya and Szego.

Second, there are some problems, for example in tidal flow in rivers and in numerical weather prediction, in which the given set of data comes from physical measurements, and may be rough or at least not smooth. It is not a known mathematical function, and yet we have to differentiate it, perhaps more than once, before our main numerical work can start. Little is known, I suspect, though the Meteorological Office is investigating the problem, of the effect on the solution of roughness in the data, or on what kinds of initial smoothing processes will produce the best results.

Conclusion

One very last word. The SPADE project, started some years ago on the West Coast of America, which

seemed to be intended, without too much exaggeration, to allow us to write, in our program, "solve partial differential equations of type T, with the following specifications," has died a lingering death. I think it was stillborn, since the topic has too many parameters for a single program. But we should at least try to find into what categories our various problems will fit, and to make the number of categories reasonably small, so that not every fresh problem needs a separate complete program.

This is a problem very suitable for university research, and I have a private regret that our Computing Laboratories have such poor facilities, both in machines and particularly in men, and that these facilities are occupied so much with teaching and the interminable preoccupation with programming languages, that we seem to make little progress in mathematics. The solution of partial differential equations is essentially a problem in mathematics, and neither ALGOL nor FORTRAN will do it for us.

Note added in proof: I am grateful for a written comment by Dr. J. C. P. Miller drawing attention to recent work by Lieberstein and Fichera. This work is concerned with solving partial differential equations in terms of a series involving standard functions, such as polynomials in the several variables concerned, with coefficients to be determined to give a suitable approximation to the solution desired. A great deal of emphasis is laid on the determination of error bounds that can be calculated—at least for the overall integral—e.g. of the square of the error. These papers, however, are too new to me for authoritative evaluation.

Bibliography

BERGMAN, S., and SCHIFFER, M. (1953). *Kernel Functions and Elliptic Partial Differential Equations in Mathematical Physics*, New York: Academic Press.

BUTLER, D. S. (1960). "The Numerical Solution of Hyperbolic Systems of Partial Differential Equations in three Independent Variables," *Proc. Roy. Soc.*, A, Vol. 255, p. 232.

DAHLQUIST, G. (1956). "Convergence and Stability in the Numerical Integration of Ordinary Differential Equations," *Math. Scand.*, Vol. 4, p. 33.

DOUGLAS, J. (1955). "On the Numerical Integration of $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = \partial u/\partial t$ by Implicit Methods," *J. Soc. Indust. Appl. Math.*, Vol. 3, p. 92.

FICHERA, G. (1959). "Boundary Problems in Differential Equations," *Proceedings* of a Symposium conducted by the Math. Research Center at the University of Wisconsin, Madison, April 1959. Madison, 1960, p. 97.

FICHERA, G. (1960). "Sul concetto di problema 'ben posto' per una equazione differenziale," *Conferenza tenuta nell'Istituto Matematico dell'Università di Roma il* 29.1.1960.

FICHERA, G. (—). "Formule di maggiorazione connesse ad una classe di transformazioni lineari," *Annali di Matematica Pura ed Applicata*, Vol. 36, Série IV, p. 274.

FORSYTHE, G. E., and WASOW, W. R. (1960). *Finite-difference Methods for Partial Differential Equations*, New York and London: John Wiley.

FOX, L. (1947). "Some Improvements in the use of Relaxation Methods for the Solution of Ordinary and Partial Differential Equations," *Proc. Roy. Soc.*, A, Vol. 190, p. 39.

FOX, L. (Ed.) (1962). *The Numerical Solution of Ordinary and Partial Differential Equations*, Oxford: Pergamon.

FRANKEL, S. P. (1950). "Convergence Rates of Iterative Treatments of Partial Differential Equations," *Math. Tab., Wash.*, Vol. 4, p. 65.

GIVENS, W. (1954). "Numerical Computation of the Characteristic Values of a Real Symmetric Matrix," Oak Ridge National Laboratory, ORNL–1574.

LAX, P. D., and RICHTMYER, R. D. (1956). "Survey of the Stability of Linear Finite-difference Equations," *Comm. Pure Appl. Math.*, Vol. 7, p. 159.

LIEBERSTEIN, H. M. (1960). "A Continuous Method in Numerical Analysis Applied to Examples from a new Class of Boundary-value Problems," Math. Research Center Technical Summary Report /175, University of Wisconsin.

MOTZ, H. (1946). "The Treatment of Singularities of Partial Differential Equations by Relaxation Methods," *Quart. Appl. Math.*, Vol. 4, p. 371.

POLYA, G., and SZEGO, G. (1951). *Isoperimetric Inequalities in Mathematical Physics*, Princeton University Press.

RICHTMYER, R. D. (1957). *Difference Methods for Initial-value Problems*, New York: Interscience.

ROLLETT, J. S., and WILKINSON, J. H. (1961). "An Efficient Scheme for the Co-diagonalization of a Symmetric Matrix by Givens' Method in a Computer with a Two-level Store," *The Computer Journal*, Vol. 4, p. 177.

SOUTHWELL, R. V. (1946). *Relaxation Methods in Theoretical Physics*, Oxford University Press.

SYNGE, J. L. (1957). *The Hypercircle in Mathematical Physics*, Cambridge University Press.

TURING, A. M. (1948). "Rounding-off Errors in Matrix Processes," *Quart. J. Mech.*, Vol. 1, p. 287.

VON NEUMANN, J., and GOLDSTINE, H. H. (1947). "Numerical Inverting of Matrices of High Order," *Bull. Amer. Math. Soc.*, Vol. 53, p. 1021.

WALSH, J. E. (1960). D.Phil. Thesis, Oxford.

WILKES, M. V. Review of *Modern Computing Methods* (Second Edition, 1961), *The Computer Bulletin*, Vol. 5, p. 75, September 1961.

WILKINSON, J. H. (1960). "Rounding Errors in Algebraic Processes," *Information Processing*, pp. 44–53. Paris: UNESCO; Munich: Oldenbourg; London: Butterworths.
WOODS, L. C. (1953). "The Relaxation Treatment of Singular Points in Poisson's Equation," *Quart. J. Mech.*, Vol. 6, p. 163.
YOUNG, D. (1954). "Iterative Methods for Solving Partial Differential Equations of Elliptic Type," *Trans. Amer. Math. Soc.* Vol. 76, p. 92.

Note on a method of forming a sorting key for a partly ordered list, and an application

By D. M. Collison

In commercial applications, calculating the keys for items to be sorted is usually trivial, since they are rarely time-dependent. In on-line applications the keys may well vary with time, and if a complicated decision is involved in calculating them, any method of simplifying the decision is useful. One way may be to assume the answer, and check the assumption in a simpler manner after the calculation. For a key with two different forms, A and B, the conventional structure of the sorting routine is shown in Fig. 1, and the suggested one in Fig. 2. The new routine behaves like a flip-flop, the decisions switching it from one state to the other; and it can be extended for keys with more than two different forms, although the result is not as efficient. The main advantage of such a structure is that it will work better if the list is already partly ordered, as the cross-overs are less efficient than the conventional structure.

In the application, it was necessary to sort items with Cartesian coordinates (x, y) into angular order. Forming the polar coordinate Θ was slow and inefficient, as the key was only needed for the sorting, and the following key f was used (the case $x = y = 0$ could not occur).

$x \geqslant y \geqslant 0$	$f = y/8x$	$-\frac{1}{2}$
$y \geqslant \lvert x \rvert$	$f = -x/8y$	$-\frac{1}{4}$
$-x \geqslant \lvert y \rvert$	$f = y/8x$	
$-y \geqslant \lvert x \rvert$	$f = -x/8y$	$+\frac{1}{4}$
$x \geqslant -y \geqslant 0$	$f = y/8x$	$+\frac{1}{2}$

f is continuous over the interval $[-\frac{1}{2}, +\frac{1}{2}]$ and any two keys can be compared by one subtraction. The key for $y = 0$, $x > 0$ should be calculated by only one of the first or last forms, and not by both.

The routine was implemented on the Elliott 502. This computer has no modulus facility and the decision whether $|y| > |x|$ takes about six instructions. Form A, therefore, divides y by x and tests for overflow (one instruction), while form B divides x by y. The actual routine to deal with the five different forms is rather more complicated than Fig. 2.

Two points arose in the implementation and are included to guide possible users. The first occurs if the hardware division process does not round properly and the full length of the quotient is used. Adjacent items may be sorted into the wrong order. Luckily the 502 forms an extra digit for the quotient and rounds accordingly; with other computers it may be necessary to round and truncate the quotient by program. Another possibility is that both x/y and y/x may cause overflow. This occurs on the 502 if $x = y$, but it was possible to make provision for it without slowing down the two main loops of the routine.

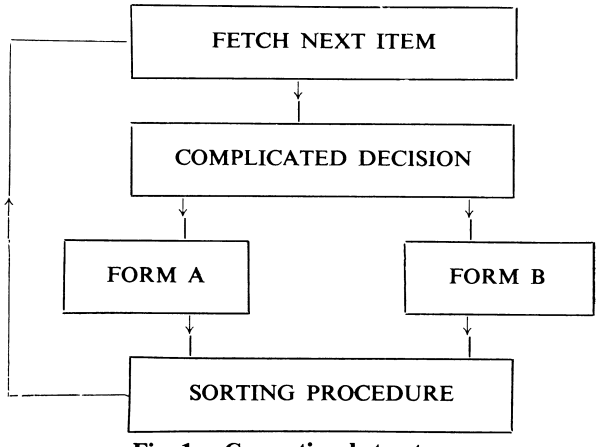

Fig. 1.—Conventional structure

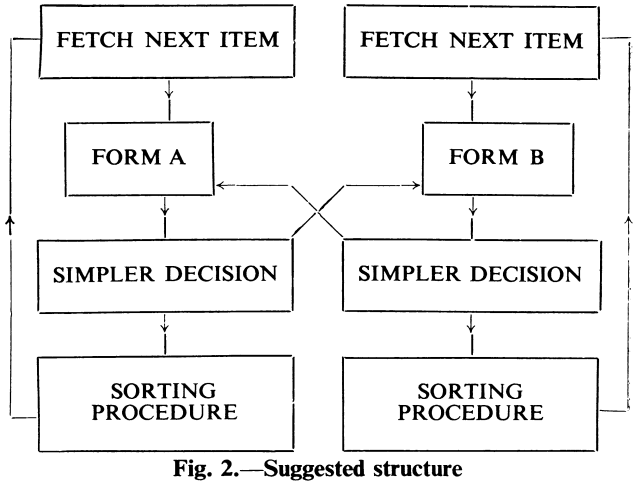

Fig. 2.—Suggested structure

The author would like to thank Elliott Brothers (London) Ltd. for permission to publish this note; and Mr. E. M. Shorter, who implemented the routine on the 502.