DEPARTMENT OF COMPUTER SCIENCES
↓ THE TECHNOLOGICAL INSTITUTE
↓ NORTHWESTERN UNIVERSITY
↓ EVANSTON, ILLINOIS 60201

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

COMPUTATIONAL COMPLEXITY IN MULTIDIMENSIONAL
NEUTRON TRANSPORT THEORY CALCULATIONS

Progress Report

for period September 1, 1975 - August 31, 1976

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

Erwin H. Bareiss

Northwestern University
Evanston, Illinois

May 1976

Prepared For

THE U.S. ENERGY RESEARCH AND DEVELOPMENT ADMINISTRATION
UNDER CONTRACT NO. E(11-1)-2280

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product or process disclosed or represents that its use would not infringe privately owned rights.

COMPUTATIONAL COMPLEXITY IN MULTIDIMENSIONAL
NEUTRON TRANSPORT THEORY CALCULATIONS

Progress Report
for the period September 1, 1976 - August 31, 1976

Project Abstract

The objectives of our research remain the same as outlined in the original proposal. They are in short as follows. Develop mathematically and computationally founded criteria for the design of highly efficient and reliable multidimensional neutron transport codes to solve a variety of neutron migration and radiation problems. Analyze existing and new methods for performance.

To achieve our goal we have subdivided the problem into three major steps:

- Step 1: Automated error analysis over computational cells.
- Step 2: Global error and stability analysis.
- Step 3: Systems analysis and assessment of computational efficiency.

Each of these steps consists of three parts.

- a.) Rigorous analysis of the analytic behavior of the mathematical solution for any sub-problem.
- b.) Implementation of computer codes. Generation of representative benchmark problems for which the exact solutions are known.
- c.) Evaluation of test results and recommendations.

The exhibits on the next two pages summarize the master plan of the research and the general philosophy for implementation.

QUANTITATIVE AND COMPUTERIZED ERROR ANALYSISA. CELL
CALCULATIONS

QUANTITATIVE ERROR ANALYSIS

B. GLOBAL
ANALYSIS

PROBABILISTIC ERROR ANALYSIS

C. COMPUTATIONAL
COMPLEXITY

COST ANALYSIS

PRINCIPLE: SYSTEMATIC EVALUATION OF KNOWN
AND NEW NUMERICAL METHODS.
"STORE" THE EVALUATION PROCEDURE,
SPECIAL FUNCTIONS, ALGORITHMS,
WITH INSTRUCTIONS ON TAPE.
EASY ACCESS TO ENGINEERS AND
SCIENTISTS.
PROVIDE TABULAR AND VISUAL
DISPLAYS.

IMPLEMENTATION: OPEN ENDED (EXPANDABLE)
COLLECTION OF SUBROUTINES.

ATTITUDE: GIVE USER FREEDOM TO SET HIS
OWN STANDARDS AND TO CHOOSE
PERFORMANCE CRITERIA FOR HIS
PROBLEM.
LET THE COMPUTER DO THE "DIRTY"
WORK.

LINEAR OPERATORS: USE DIMENSIONLESS ANALYSIS
TO REDUCE THE NUMBER OF CELL
CALCULATIONS.

Scope of Investigations

Our aim is to develop a theory of error estimation for coarse network calculations in neutron migration problems and implement the results simultaneously in a computer code for use by engineers and scientists that need not be specialists in numerical analysis. At the same time, as new analytical insight is gained, new numerical methods are developed and tested. We look at the problems from different points of view and then try to make a best judgment for implementation in a large scale computer code.

Significant Results Obtained by May 1976

- a) Implementation of the integer-preserving Gaussian elimination method (two-step method) in a CDC 6400 computer code.

This code inverts matrices, calculates its determinants, or solves systems of linear equations up to order 24 exactly. It is very efficient, and as a general library routine available on request. It provides exact answers for ill-conditioned matrices when ordinary Gaussian elimination or Householder factorization with iterative improvement does not converge. The User's Guide (see Publications, #21; ERDA Report C00-2280-24) is enclosed with this report. Such a code, to name an example, is very helpful in finite element methods where an ill-conditioned algorithm can be made well-conditioned by static reduction.

- b) Modes analysis for one-dimensional transport solutions. A new set of generalized modes was derived that permits representation of transport solutions to arbitrary accuracy. The theoretical results are implemented in the "Mode Analyzer" which yields backward errors up to 10^{-12} .

(if so desired). The method is described in ERDA Report C00-2280-25 (see Publication #22) of which a preliminary copy is enclosed.

- c) New method for solving the 1-T transport equation. With the asymptotic and the generalized modes described in (b), it is possible to solve the discrete boundary value problem for the transport equation with region-wise constant coefficients exactly. Since this is in effect an analytic solution, no additional interior meshpoints within a region are needed. The solution satisfies exactly the transport equation and thus gives the correct behavior at the boundary and interfaces. The method has been implemented in BEAPAC 1-T to create user-oriented exact benchmark problems. If implemented in a production code, the order of the global matrix could be reduced by a factor of two to eight or more.

We have also obtained interesting preliminary results in factorization iterative techniques (#21, ERDA Report C00-2280-12, copy enclosed), in boundary integral methods for solving the neutron diffusion equation, and in statistical error analysis. We shall summarize these results below.

Compliance with Contract Requirements

To the best of the principal investigator's knowledge, all contract requirements have been fulfilled.

Time Effort

The principal investigator has devoted a minimum of twenty percent of his time on work for the contract since September 1, 1975, and expects to devote 100% of his time on the project during the summer months from June through August.

Review of Work During September 1, 1975 to August 31, 1976

Elementary Remarks:

In the first section we gave a master plan for our research. The question is often asked how it is possible to make meaningful statements on coarse network calculations since no analytic solutions to the transport equation exist. We shall try to explain our thinking very briefly by analogy with systems of ordinary differential equations. Let

$$(1) \quad y' = f(x)y + g(x) \quad y(0) = y_0$$

be a single differential equation. Then the solution is given by

$$(2) \quad y = \left(\exp \int_0^x f dx \right) \left\{ y_0 + \int_0^x \left(\exp - \int_0^{x'} f dx'' \right) g(x') dx' \right\}$$

If (1) is regionwise constant, then in region 1, say,

$$(3) \quad y' = ay + bx + c \quad (y(0) = y_0)$$

and (2) reduces to

$$(4) \quad y = -(c + b/a + bx)/a + z_0 e^{ax}$$

$$z_0 = +(c + b/a)/a + y_0$$

Similarly, when we have a system of linear ordinary differential equations

$$(5) \quad y' = Ay + bx + c \quad y(0) = y_0$$

where A is an $n \times n$ matrix, and, y , b and c are $n \times 1$,

$$\begin{aligned}
 (6) \quad y &= -A^{-1}(c + A^{-1}b + bx) + Mz_0 \\
 z_0 &= M^{-1}(A^{-1}c + A^{-2}b + y_0) \\
 M &= P \exp(Jx)P^{-1} = PTP^{-1}
 \end{aligned}$$

where $J = P^{-1}AP$ is the Jordan canonical form of A . The matrix J is the direct sum of its Jordan blocks $J_k(\lambda_i)$ of order $p+1$,

$$(7) \quad J_k(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix}$$

while $T = \exp(Jx)$ is the direct sum of the blocks

$$(8) \quad T_k = \exp J_k(\lambda_i)x = \begin{bmatrix} 1 & x/1! & \dots & x^p/p! \\ & 1 & \ddots & \vdots \\ & & \ddots & \vdots \\ & & & 1 & x/1! \\ & & & & 1 \end{bmatrix} e^{\lambda_i x}$$

which correspond to the Jordan blocks $J_k(\lambda_i)$. The end effect is the representation of the solution y of (6) by a linear combination of base functions, given by the particular solution and the n homogeneous solutions. Define $w(x)$ by

$$(9) \quad y(x) = Pw(x) \Rightarrow w(0) = P^{-1}y(0)$$

and partition $w(x)$ in correspondence with the Jordan blocks $J_k(\lambda_i)$,

$$(10) \quad w^T(x) = [w_1(x), \dots, w_m(x)]^T$$

Then (6) becomes

$$\begin{aligned}
 w(x) &= w_P(x) + w_H(x) \\
 (11) \quad w_P(x) &= -J^{-1}P^{-1}(c + bx) - J^2P^{-1}b \\
 w_H(x) &= Tw_H(0) \quad \Rightarrow w_{H,k}(x) = T_k w_{H,k}(0) \\
 w_H(0) &= T^{-1}(w(0) - w_P(0)) \Rightarrow w_{H,k}(0) = w_{P,k}(0)
 \end{aligned}$$

which reveals the mathematical structure of $y(x)$.

Assume now (5) is to be solved by a numerical method of higher order for a coarse network, and that the algorithm represents the solution of a system of linear finite differences of order q . This system is in turn equivalent to a system of qn first order finite differences of the form

$$(12) \quad y_{n+1} = B_n y_n + g_n$$

where y_n is the approximate solution of $y(x_n)$, and B_n is often a sparse matrix of order qn with constant entries. The algorithm is stable if the spectral radius

$$(13) \quad \rho(B) \leq 1.$$

(If B_n is not constant, we have shown only local stability.) If B_n is constant, we can determine also the general solution to (12) with results quite similar to those of (11), and then derive a formula for an error bound $\|y_{n+1} - y(x_n)\|$. In practical calculations (13) is often replaced by the von Neumann stability condition $\rho(B) \leq 1 + O(h)$. The entire procedure can be programmed for the computer.

A quantitative approach is as follows. We use P , $w_P(x)$ and the w_{Hk} ($k = 1, \dots, m$) of (11) to create test problems by appropriate choices

of $w_{H,k}(0)$. These problems are solved on the computer by (12) and the true error in the numerical solution is observed for the coarse network. (At the same time $\|B_n\|$ or $\rho(B)$ can be estimated.) Because of the linearity of the problem, and the fact that the solutions and eigenvalues are continuously dependent on the input parameters, appropriate interpolation is possible to save computer time. For a given particular problem, using appropriate computational care, the solution (6) can be calculated to any desired precision. By making residual checks in (5), the accuracy of (6) can be confirmed locally.

For linear singular differential equations and for partial differential, integral and integro-differential equations, similar considerations can be applied and implemented, but the arguments become lengthy.

On the following page, the stationary neutron transport equation and one form of a general solution is exhibited. If we make in (3) the substitution

$$y = \psi(x, \mu), \quad a = -\frac{\sigma_T}{\mu}, \quad b = 0, \quad c = \frac{\sigma_s}{2\mu} \varphi(x) + \frac{S}{\mu}$$

we have the transport equation in slab-geometry. If we assign μ a set of discrete values $\{\mu_\ell\}$, and let $\varphi = \sum_{(\ell)} f_\ell \psi(x, \mu_\ell)$, then (3) becomes a system of linear equations of the form (5) and our remarks made above apply to this sub-problem on stability and error evaluation.

A final important point will be made. We wrote equation (1) with variable coefficients and gave the corresponding solution (2). In engineering neutron transport calculations, however, the coefficients σ_T (and σ_s) are given, in practically all cases, as regionwise constants. While we do not have the analytic solution for space-dependent coefficients, we do have the general solution for the transport equation with constant

STATIONARY NEUTRON TRANSPORT EQUATION

$$\underbrace{v \cdot \nabla \psi + v \Sigma_T \psi}_{\mathcal{L} \psi} = \underbrace{\int v' \Sigma_s(v' \rightarrow v) \psi(r, v') dv'}_{\mathcal{C} \psi} + Q$$

LOSS OPERATOR

COLLISION OPERATOR

INTERIOR
SOURCE $\mathcal{D}$ $\partial \mathcal{D}$

$$\boxed{\begin{aligned} \mathcal{L} \psi &= \mathcal{C} \psi + Q \\ \psi(r, v) &= f(r, v) \end{aligned}}$$

 $\in X$ $r \in \partial \mathcal{D}$
 $n \cdot v < 0$

$$X = L_1(\mathcal{D} \times \Omega v)$$

 Ω : UNIT SPHERE

GENERAL SOLUTION

$$\boxed{\psi = \psi_0 + \psi_1 + \psi_m}$$

 $\mathcal{D}$
 $\partial \mathcal{D}$

$$\left. \begin{aligned} \mathcal{L} \psi_0 &= Q \\ \psi_0 &= f \end{aligned} \right\}$$

 $\mathcal{D}$
 $\partial \mathcal{D}$

$$\left. \begin{aligned} \mathcal{L} \psi_1 &= \mathcal{C} \psi_0 \\ \psi_1 &= 0 \end{aligned} \right\}$$

$$\psi_0 + \psi_1 = \mathcal{N}_1 f + \mathcal{N}_2 Q$$

 $\mathcal{N}_i$: NONCOMPACT
BOUNDED OPERATORS $\mathcal{D}$
 $\partial \mathcal{D}$

$$\left. \begin{aligned} (\mathcal{L} - \mathcal{C}) \psi_m &= \mathcal{C} \psi_1 \\ \psi_m &= 0 \end{aligned} \right\}$$

$$\psi_2 = \mathcal{K}_1 f + \mathcal{K}_2 Q$$

 $\mathcal{K}_i$: COMPACT OPERATORS

coefficients. Thus we have the necessary tools to attack the problems of practical importance with mathematical rigor. Note, that equation (2) can be used to obtain a sharp bound on the error in the solution if a coefficient is a mean value of a variable quantity.

BEAPAC 1-T:

BEAPAC 1-T is a computer code for the IBM 360/195 that performs an error and stability evaluation for different types of one-dimensional transport equations and their numerical methods of solution for one and two cell problems (multicell problems can be handled too). A set of widely used algorithms is incorporated in matrix representation. The analysis is performed in the spirit of the previous section, but the mathematical properties of the solutions are more complicated. How the system works, and what algorithms are used, is described in the ERDA Report C00-2280-25, "An Automated Approach to Quantitative Error Analysis." A preliminary copy of this report is enclosed for the reader's information. The set of base functions has been changed from the previous code. The exhibit on the following page summarizes these base functions as used in the mode analyzer. It should be noticed that they are functionals of the generalized eigenmodes $\varphi_\nu(\mu)e^{-x/\nu}$ whose test functions are a set of first order splines and Holder-continuous functions. In the form given, it is possible to approximate any given flux shape by expansion in these homogeneous base functions (and a particular solution in the inhomogeneous case) pointwise to any desired accuracy. With the previous set of base functions we obtained heavy fluctuations near the boundary (similar to Gibbs' phenomenon in Fourier series expansions) although the solution did converge in the L_1 -norm.

THE MODE ANALYZER

GENERAL SOLUTION:

$$\psi(x, \mu) = \psi_0^+ + \psi_0^- + \psi_T$$

$$\psi_0^\pm = a^\pm \frac{\nu_0}{\nu_0 \mp \mu} e^{\mp x/\nu_0}; \quad \psi_T = \int_{-1}^{+1} A(\nu) \phi_\nu(\mu) e^{-x/\nu} d\nu$$

$$\phi_\nu(\mu) = \frac{\nu}{\nu - \mu} + \delta(\nu - \mu) \left\{ 1 - \frac{c\nu}{2} \ln \frac{1+\nu}{1-\nu} \right\}$$

$$\text{LET } A(\nu) = \sum_{(k)} A_k(\nu)$$

BASE FUNCTIONS:

$$\psi_0^+; \psi_0^-; \psi_k = \int_{\nu_{k-1}}^{\nu_k} A_k(\nu) \phi_\nu(\mu) d\nu$$

ALL BASE FUNCTIONS SATISFY EXACTLY

$$(\mathcal{L} - \mathcal{C}) \psi_k = 0$$

Work to be done by August 31, 1976:

Implement BEAPAC T. This will be an error analysis program patterned after BEAPAC 1-T. Provisions are made for x-geometry, x,y-geometry, R, θ -geometry and x,y,z-geometry. Graphic display will be two-dimensional profiles, three-dimensional plots, and on-line printer plots for fast and inexpensive output. The code will be prepared to have interface capability with TWOTRAN, TRIPLET, FENT, DIF30, THREETRAN. The corresponding matrix analysis will be postponed for the next period.

Complete Report C00-2280-25. This report is intended to be a joined publication with Argonne National Laboratory.

Complete User's Guide for BEAPAC 1-T.

Transfer the new code BEAPAC 1-T and BEAPAC T for the IBM 360/195 to Northwestern University's CDC 6400.

Construct analytic solutions for the two- and three-dimensional corner problem by a combination of the methods, such as described in C00-2280-1, and later reports.

Interface and Corner Problem in Diffusion Theory

Since the neutron diffusion equation is a widely used approximation to neutron transport theory, we decided to include neutron diffusion as an option in the error analysis program. In particular, it appeared that the corner problem was not well understood, as we pointed out in report C00-2280-17. Since we are interested in the quantitative as well as the analytical behavior of the solution of the diffusion equation we created a simple benchmark problem as shown on the exhibit of the following page. The traditional approach in series expansions of the eigenfunction of the Laplace equation and applied to the Poisson equation is a very simple

DIFFUSION APPROXIMATION: BENCHMARK PROBLEM

$$-\nabla D(\theta) \nabla \phi(r, \theta) = S(\theta)$$

$$\left. \frac{\partial \phi}{\partial r} \right|_{r=1} = 0$$

DIMENSIONAL ANALYSIS: REDUCE NUMBER OF PARAMETER TO 3

$$D_1 \quad \Sigma_1 \quad \Sigma_2$$

$$(D_2 = 1, S_2 = 0, S_1 = 1)$$

CENTER PROFILE

INTERFACE PROFILE

ANGULAR PROFILE

FLUX PROFILES

matter. However, when we applied the technique to the neutron diffusion equation equation, we run into considerable computational difficulties. The exhibits of the next two pages show the general representation of the solution to our model problem, and the result of a residual check in reconstructing the source term from a rather elaborate series expansion solution. We extended our research to multiregion problems. The results turned out to be as expected. It is not sufficient to restrict oneself to the lowest eigenvalues, and few term series expansions are very inaccurate.

Next we looked at the interface and corner problem using the Kontorovich-Lebedev transform. The representation of the solution near the corner for the infinite corner problem appeared very promising, but again we encountered computational difficulties in evaluating the integrals. One chief reason is the occurrence of a heavily oscillating argument.

We returned to the Green's function approach. Originally we hoped that a Green's function for the infinite domain would be sufficient to reduce the problem to a boundary and interface integral equation problem. At that time, the analytic work involved turned out to be prohibitive. We have now a very promising Green's function with very desirable properties that lend themselves to computational treatment. A summary of the work done so far is enclosed to this report, entitled "Interface and Corner Problem in Diffusion Theory," ERDA Report C00-2280-27. It is a preliminary report for the reader's information only.

Work to be done by August 31, 1976:

We shall continue to work on the "Interface and Boundary Value Integral Equation Method" and complete the preliminary report mentioned above. As it stands now, the method is very promising as a theoretical foundation for a very efficient and accurate numerical method to solve the neutron diffusion

ANALYTIC SOLUTION

SERIES EXPANSION (TRADITIONAL APPROACH)

$$-\nabla D(\theta) \nabla \varphi = S(\theta)$$

POISSON EQUATION

$$\varphi = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{ij} \chi_i(\theta) J_{\nu_i}(\mu_{ij} r)$$

$$\chi_i = \begin{cases} \cos \nu_i(\pi - \alpha) \cos \nu_i \theta \\ \cos \nu_i \alpha \cos \nu_i(\pi - \theta) \end{cases}$$

IN REGION 1

IN REGION 2

CHARACTERISTIC EQUATIONS:

$$\text{DET} \begin{bmatrix} \cos \nu_i \alpha & \cos \nu_i(\pi - \alpha) \\ D \sin \nu_i \alpha & -\sin \nu_i(\pi - \alpha) \end{bmatrix} = 0; \quad J'_{\nu_i}(\mu_{ij}) = 0$$

$$-\nabla D(\theta) \nabla \varphi + \sum \varphi = S(\theta)$$

NEUTRON DIFFUSION

$$\text{IF } \nu_k \neq 2j \quad \varphi = \sum_{\ell=0}^{\infty} \sum_{k=0}^{\infty} \sum_{i=0}^{\infty} a_{ik\ell} \chi_i J_{\nu_k}(\mu_{k\ell} r)$$

$$\text{IF } \nu_k = 2j \quad \text{ADD TERMS} \quad r^{2j} (\ell n r)^m$$

$$\text{IF } \nu_k \approx 2j \quad \text{NUMERICALLY UNSTABLE}$$

RECONSTRUCTION OF SOURCE TERM

CENTER PROFILE

--- GIVEN SOURCE
— CALCULATED SOURCE
USING 12 TERMS IN
SERIES EXPANSION

equation. We shall extend our numerical experimentation, but restrict ourselves to smaller benchmark problems.

Matrix Theory

A report on the application of relaxation techniques to factorization iterative procedures was finalized. In this report a precise meaning is given to the notion of iterative procedures of order q and the two-parameters overrelaxation technique of Woznicki is shown to be a factorization iterative method of the second order; some consequences of this fact are developed.

Other results include the refinement of known properties of the extrapolation principle and estimations of the powers of the Oliphant relaxation technique for consistently ordered matrices and of the Chebyshev technique for Stieltjes matrices.

This report was written by R. Beauwens and is enclosed (ERDA Report C00-2280-12).

Another report on factorization iterative methods for neutron migration problems with numerical experimentation is in progress, but seems unlikely to be completed before August 31, 1976.

Statistical Error Analysis

In solving large linear systems, the absolute forward error bounds (and often the backward error bounds as popularized by J. H. Wilkinson) are in almost all cases so conservative that the code users do ignore them. It has been observed that the errors in floating point arithmetic are not samples from a Gaussian distribution. On the other hand, the rounding errors under addition with fixed point arithmetic have a Gaussian probability distribution. If we replace the basic operations in floating point arithmetic by computations with logarithms, addition- and subtraction logarithms,

then all basic operations are reduced to fixed-point additions, subtraction and well controlled probability density transformations. This is seen easily from the simple example

$$c = a + b \Rightarrow \log c = \log a + \log f(\log b - \log a)$$

where $f(x) = 1 + e^x$. Hence there is a possibility to find bounds for the probabilistic floating point errors as transforms of Gaussian distributions. These ideas open a new field of thought for computer arithmetic.

A. Roundoff errors. The rounding of a number x to the nearest number in a number system using base b and s places behind the radix point may be treated as rounding $b^s x$ to the nearest integer. If numbers are uniformly distributed on the number line, the errors made in rounding to the nearest integer have a rectangular density function, $f(x) = 1$ on the unit interval $-1/2 \leq x < 1/2$ and $f(x) = 0$ elsewhere, since all numbers in the interval are equally likely to occur as a result of the roundoff operation.

This statistical treatment of roundoff errors is unsatisfactory for two chief reasons:

1. Naturally occurring numbers are not uniformly distributed on the number axis. Rather [1] they empirically follow, when expressed in terms of a fixed base b , a logarithmic law of distribution of leading digits

$$p(a) = \log_b (a+1)/a$$

for $a = 1, \dots, b - 1$. This corresponds to a uniform density of the mantissas of their logarithms to this base (or, for that matter, to any other base). Obviously, for numbers which are rounded from a distribution of this kind, roundoff errors will not be uniformly distributed either.

2. The motivation behind machine floating point number systems lies in the importance of the relative difference of numbers $|x - y|/|x|$ over that of the absolute difference $|x - y|$. Since these machines are actually hybrid systems, the important points are obscured; they consist of a series of arithmetic progressions of machine numbers which themselves progress geometrically. It is possible to design number systems for machine use which concentrate on the latter aspect by retaining the exponent part of the floating point representation and discarding completely the fractional part. For useful systems the base can no longer be 2 or 10, obviously, but will be a small number such as 2^{-10} . A number is thus represented on the machine by d^k , where d is the fixed base and k is an integer. The machine binary pattern to represent x may be that which previously was the integer representation of k . In such a constant relative difference system, numbers are preferably rounded in the logarithmic sense, i.e., x will be rounded to d^k if it is true that $d^{k-1/2} \leq x < d^{k+1/2}$. This interval, it may be noted, is not symmetric about d^k in the absolute sense; $\log_d x$, however, falls in the interval

$$k - 1/2 \leq \log_d x < k + 1/2,$$

which is symmetric about $\log_d d^k = k$. Certain considerations justify the assumption that in this interval the logarithms of x are uniformly distributed. This means, however, that x itself will have the density function

$$f(x) = 1/(x \log d)$$

in this interval. In this case we are not really interested in the distribution of leading digits, though the reciprocal property for them returns

as soon as we write the number x in terms of their representation in base b .

Meaningful statistical analysis of propagated errors must take into account the above considerations. In investigations in this area, the reciprocal distribution apparently will play a leading role, particularly since it has also been demonstrated that the distribution persists under certain repeated arithmetic operations [2]. There remains the question of the relationship of the reciprocal distribution to the normal distribution which bears investigation. Some preliminary results not yet in a form suitable for inclusion here are at hand. Some efforts to investigate these questions are planned for the coming summer months.

B. Constant relative difference number systems. We present some preliminary notes dealing with the properties of machine number systems in general. As can be seen from these, the systems alluded to above, also called logarithmic number systems, since in effect the logarithm of x to the base d is stored as its representation, have certain practical advantages:

1. Number representation may be more efficient than in the use of normal floating point number systems;
2. It is possible to introduce naturally dynamic precision, that is, precision under control of programs, into computation without changing word length.

The introduction of dynamic precision into computation permits a closer approximation to a true machine realization of the mathematical concept of real number. A real number, in terms of the Cauchy definition, is a number that can be represented by convergent sequences of rational numbers. The computationally practical aspect is that if a certain rational approximation

(the only ones that can be used) is at some point no longer suitable, the sequence indicates how a better one can be found. For our purposes no real generality is lost if the sequence is considered to consist of terms $d_k^{a_k}$, where $d_k \rightarrow 0$ as k increases without limit, and a_k is an integer. In connection with the constant relative difference systems it is not even essential that d_k be rational.

Change of precision (with a corresponding change of number range) is achieved without changing word length simply by changing the value of d . This requires the conversion of numbers needed to the new base - the algorithm to do this is very simple, though it must still be determined whether this can be done without using an auxiliary representation. In algorithms performing the cardinal operations and the necessary conversions to and from the external (ordinary decimal) representation, d will appear as a global parameter. The careful study of efficient algorithms needed here is indicated.

Multiplication and division pose no problems in these systems, as the algorithms to perform them are independent of d . Addition and subtraction require integer-valued functions of integers, which are defined as the rounded results of transcendental functions. Since the values of the transcendental functions are to be rounded, it would appear that simpler approximations for them ought to suffice than if a number of places behind the decimal point were required. Conjectures concerning criteria which may be used to terminate approximate methods for computing the transcendental functions at a point where the true rounded result is assured are under investigation.

Work to be done before August 31, 1976.

A report on constant relative computer arithmetic is to be completed.

It is hoped to clarify the direction which should be taken, and to lay the foundation on which a meaningful probabilistic error analysis can be built.

References

- [1] Benford, Frank, "The Law of Anomalous Numbers," Proc. Amer. Phil. Soc., 78 (1938), 551-572.
- [2] Hamming, R. W., "On the Distribution of Numbers," Bell Sys. Tech. Jour., 49, No. 8 (1970), 1609-1625.

Preliminary Bibliographical Material on Statistical Error Analysis

The publications listed here represent early papers on the subject and references to the subject in text books in numerical analysis. Later material is mostly not covered and will be added as research continues. It is expected that some of the material presented in the articles and books will be summarized in a set of notes to be used for reference.

1. Bjerkammar, Theory of Errors and Generalized Matrix Inverses, Elsevier Publishing Co., 1973, chap. 2, 8-320.
2. Feller, William, An Introduction to Probability Theory and Applications, Vol. 1 and 2, Wiley and Sons, 1957.
3. Forsythe, G. E., "Roundoff Errors in Numerical Integration on Automatic Machinery," Bull. Amer. Math. Soc., 56 (1950), 61-62.
4. Goldstine, H. H. and von Neumann, John, "Numerical Inverting of Matrices of High Order, II," Proc. Amer. Math. Soc., 2 (1951), 188-202.
5. Hamming, R. W., "On the Distribution of Numbers," Bell System Tech. Jour., 49, No. 8 (Oct. 1970), 1609-1625.
6. Hamming, R. W., Numerical Methods for Scientists and Engineers, 2nd ed., McGraw-Hill, 1973, 33-39.
7. Hildebrand, F. B., Introduction to Numerical Analysis, McGraw-Hill Publishing Co., 1956, 17-21.
8. Householder, A. S., Principles of Numerical Analysis, McGraw-Hill, 1953, 14-15.
9. Huskey, H. D. and Hartree, D. R., "On the Precision of a Certain Procedure of Numerical Integration," Jour. Res. Nat. Bureau of Stand., 42 (1949), 57-62.
10. Inman, S., "The Probability of a Given Error Being Exceeded in Approximate Computation," Math. Gaz., 34 (1950), 99-113.
11. Kuperman, I. B., Approximate Linear Equations, van Nostrand-Reinhold Co., 1971, chap. 12, 196-219.
12. Rademacher, H. A., "On the Accumulation of Errors in Processes of Integration on High-Speed Computing Machines," Proc. Symp. Large-Scale Digital Calc. Mach., Harvard Univ. Press, 1948, 176-185.
13. Henrici, Peter, Error Propagation for Difference Methods, Wiley & Sons, New York, 1963, chap. 5, 37-47.
14. Parker, Jr., D. S., "The Statistical Theory of Relative Errors in Floating Point Computation," Univ. of Illinois, Rept. No. UIUCDCS-R-76-787, March 1976.

Publications supported under Contract No. E(11-1)-2280

1972 - 1973

1. Bareiss, E. H. and I. K. Abu-Shumays, "Finite Elements in Neutron Transport Theory," Northwestern University, AEC Report C00-2280-1 (June, 1973)
2. Abu-Shumays, I. K. and E. H. Bareiss, "Adjoining Appropriate Singular Elements to Transport Theory Computations," Northwestern University, AEC Report C00-2280-2 (June, 1973)
3. Bareiss, E. H. and I. K. Abu-Shumays, "Computational Complexity in Multidimensional Transport Theory Calculations, Progress Report," Northwestern University, AEC Report C00-2280-3 (May, 1973)

1973 - 1974

4. Derstine, K. L., E. H. Bareiss and I. K. Abu-Shumays, "An Automated Error Analysis Code for Transport Theory Calculations," Northwestern University, AEC Report C00-2280-4 (revision)
5. Beauwens, R., "Convergence Analysis of Some Factorization Iterative Methods for M-Matrices," Northwestern University, AEC Report C00-2280-5 (October, 1973)
6. Abu-Shumays, I. K., "Transcendental Functions Generalizing the Exponential Integrals," Northwestern University, AEC Report C00-2280-6 (November, 1973)
7. Abu-Shumays, I. K. and E. H. Bareiss, "Singular Elements in Variational and Finite Element Transport Computations," Trans. Am. Nucl. Soc. 17, p. 236, C00-2280-13 (1973)
8. Abu-Shumays, I. K. and E. H. Bareiss, "Adjoining Appropriate Singular Elements to Transport Theory Computations," J. Math. Anal. Appl. 48 (1974), 200-222, C00-2280-18
9. Beauwens, R., "On the Point and Block Factorization Iterative Methods for Arbitrary Matrices and the Characterization M-Matrices," Northwestern University, AEC Report C00-2280-7 (January, 1974)
10. Stakgold, Ivar, "Global Estimates for Nonlinear Reaction and Diffusion," Northwestern University, AEC Report C00-2280-8 (May, 1974), Chapter in "Ordinary and Partial Differential Equations," Proceedings of the Conference held in Dundee, Scotland, 26 - 29 March 1974, Lecture Notes in Mathematics, Springer Verlag, Berlin-Heidelberg, New York (252-266) 1974.
11. Bareiss, E. H. "Computational Complexity in Multidimensional Transport Theory Calculations, Progress Report," Northwestern University, AEC Report C00-2280-9 (May, 1974)
12. Payne, L. E. and Ivar Stakgold, "Isoperimetric Inequalities for a Critical Reactor," AEC Report C00-2280-10 (available in June, 1974)

1974 - 1975

13. Bareiss, E. H., "Diffusion and Transport Theory Solutions Near Interior Corners and Edges," Abstract, Proceedings International Congress of Mathematicians, Vancouver, B.C., C00-2280-15 (August, 1974)
14. Bareiss, E. H., and S. R. Vickery, "The Interior Corner Problem for the Neutron Diffusion Equation," Northwestern University AEC Report C00-2280-17 (November 1974)
15. Abu-Shumays I. K., E. H. Bareiss, and K. L. Derstine, "A Benchmark Error Analysis Package for One-Dimensional Neutron Transport Calculations, BEAPAC-1T" Northwestern University, AEC-Report C00-2280-19, Trans., ANS, 19 (166-67) October 1974
16. Stakgold, Ivar, and L. E. Payne, "On Hopf Type Maximum Principles for Convex Domains, Northwestern University ERDA Report C00-2280-20 (May 1975) (Submitted to Oxford Quarterly Journal)
17. Bareiss, E. H., and K. L. Derstine, "Quantitative and Computerized Error Analysis," Northwestern University ERDA Report C00-2280-21 (May 1975) to be published in Proceedings of Fourth National Conference on Transport Theory, Blacksburg, VA (March 1975)
18. Bareiss, E. H., "Computational Complexity in Multidimensional Transport Theory Calculations, Progress Report," Northwestern University, ERDA Report C00-2280-22

1975 - 1976

19. Bareiss, E. H., and I. K. Abu-Shumays, "Numerical Solution of the Neutron Transport Equation," accepted for publication in SIAM Jour. Num. Anal. ERDA Report C0-2280-23
- *20. Beauwens, R., "On the Application of Relaxation Techniques to the Factorization Iterative Procedures," Northwestern University, ERDA Report C00-2280-12 (CS/ES 76-1)
- *21. Bareiss, E. H., and R. Kobbe, "Two-Step, User's Guide" Northwestern University, ERDA Report C00-2280-24
- *22. Bareiss, E. H., and K. L. Derstine, "An Automated Approach to Quantitative Error Analysis," Northwestern University, ERDA Report C00-2280-25 (CS/ES 76-2, ANL 76-59)
- *23. Bareiss E. H., "Computational Complexity in Multi-dimensional Transport Theory Calculations, Progress Report," Northwestern University, ERDA Report C00-2280-26 (CS/ES 76-3)

In Preparation

- *24. Bareiss, E. H., et al, "Interface and Corner Problem in Diffusion Equation" C00-2280-27
- 25. Bareiss, E. H., "Ordering of Factorization Iterative Methods for Neutron Migration Problems"
- 26. Bareiss, E. H., and J. L. Lloyd, "An Integral Equation for the Solution of the Two-Dimensional Neutron Transport Equation in a Convex Region with Reflecting Boundaries"
- 27. Bareiss, E. H., et al., "Direct Methods for Sparse Matrices," Northwestern University, ERDA Report C00-2280-11
- 28. Bareiss, E. H. and A. A. Grau, "New Results in Statistical Error Analysis"
- 29. Grau, A. A., "Constant Relative Difference Arithmetic"