LOCKHEED MARTIN ENERGY RESEARCH LIBRARIES
3 4456 0515566 7

CENTRAL RESEARCH LIBRARY
DOCUMENT COLLECTION

4

ORNL-4385
UC-32 – Mathematics and Computers

MATHEMATICS DIVISION

ANNUAL PROGRESS REPORT

FOR PERIOD ENDING DECEMBER 31, 1968

OAK RIDGE NATIONAL LABORATORY
CENTRAL RESEARCH LIBRARY
CIRCULATION SECTION
4500N ROOM 175
LIBRARY LOAN COPY
DO NOT TRANSFER TO ANOTHER PERSON.
if you wish someone else to see this report, send in name with report and the library will arrange a loan.
ORNL-118 (6-97)

OAK RIDGE NATIONAL LABORATORY
operated by
UNION CARBIDE CORPORATION
for the
U.S. ATOMIC ENERGY COMMISSION

ORNL-4385

Contract No. W-7405-eng-26

MATHEMATICS DIVISION
ANNUAL PROGRESS REPORT
for Period Ending December 31, 1968

A. S. Householder, Director
H. P. Carter, Assistant Director
D. A. Gardiner, Assistant Director

Compiled and Edited by
Manuel Feliciano, Jr.
C. W. Nestor, Jr.

MARCH 1969

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee
operated by
UNION CARBIDE CORPORATION
for the
U.S. ATOMIC ENERGY COMMISSION

LOCKHEED MARTIN ENERGY RESEARCH LIBRARIES

3 4456 0515566 7

Reports previously issued in this series are as follows:

ORNL-2283	Period Ending February 28, 1957
ORNL-2652	Period Ending August 31, 1958
ORNL-2915	Period Ending December 31, 1959
ORNL-3082	Period Ending December 31, 1960
ORNL-3264	Period Ending January 31, 1962
ORNL-3423	Period Ending December 31, 1962
ORNL-3567	Period Ending December 31, 1963
ORNL-3766	Period Ending December 31, 1964
ORNL-3919	Period Ending December 31, 1965
ORNL-4083	Period Ending December 31, 1966
ORNL-4236	Period Ending December 31, 1967

Contents

MATHEMATICAL RESEARCH

Questions Relating to the Zeros of a Polynomial 1

Numerical Solution of Integral Equations 2

Stress Analysis of Graphite Under Neutron Irradiation 3

Oscillatory Behavior of a Linear Third-Order Differential Equation and Its Adjoint 4

Construction of Fields with Given Group 5

Numerical Solutions of Ordinary Initial Value Problems 5

Roots and Proper Vectors of a Real Symmetric Matrix 5

Arithmetical Functions Related to the Fibonacci Numbers 6

Amicable Numbers 6

COMPUTER APPLICATIONS

Special Projects 7

- General Purpose System 7
- Spectrum Analysis System 8
- Family of Heat Transfer Programs 9
- Graphics Program Development 11
- Information Retrieval 13
- Programming for the Library KWIC Index 13
- FORTRAN Source Program Documentation 14
- COMMGEN 15
- Multiple Factor Analysis 15
- Statistical Calculations 16
- Two Programming Systems 16
- Mathematical Programming System, MPS/360 16

Lecture Series 17

- FORTRAN Lecture Series 17

Programming Assistance 17

Chemical Technology 18

- Noble-Gas Separation Project 18

Optimization of Sizes and Locations of Reprocessing Plants in an Expanding Economy 19
Simplified Reactor Systems Linear Program for ORNL 21
High-Temperature Gas-Cooled Reactor Fuel Reprocessing 22
Spent Fuel Shipping Casks 22
Preliminary Analysis of Light Absorption Spectra 22

Civil Defense 22

General Engineering and Construction 22

Capsule Design Program 22

Health Physics 23

Dispersion of Finite Neutron Beams Within the Cylindrical Model of Standard Man 23
Estimation of Plutonium Body Burden 23
Dose from Gamma-Ray Sources 23
Water Pollution 24

Instrumentation and Controls 24

Radioactive Isotope Sand Tracer Program 24
FFTF Thermocouple Drift Tests 25
Data Compression and Reformatting Program 25
TRU Product Tank Temperature Anomalies 25
Inventory and Maintenance Records 26
Material Usage Report 26

Metals and Ceramics 26

Alloys 26
Collimation Correction in Small-Angle X-Ray Scattering 26

MSRE 27

An Improved MSRE Data Processing System 27
Calculation of Fission Product Concentrations in the MSRE 27

Neutron Physics 28

O5R – Albedo Version 28
Analysis of Neutron Transport Problems 28
O5R/ACTIFK 28
SNAP Collimator Design 29
Time-Dependent Neutron and Secondary Gamma-Ray Transport in Air over Ground 29
Analysis of TSF Slab Experiments 29
Air Transport 29
OGRE-4D 30
Program DUCAL 30
Intranuclear Cascades 30

Nuclear Data Project 32

Nuclear Systematics 32
Atomic Mass Adjustment 32
Text Capitalization Algorithms 32

ORELA 34

ORIC 35
Data Transmission 35
Physics 35
Relativistic Hartree Calculations 35
Fission Fragment Data Analysis 38
Programming Services 38
Reactor 38
Steam-Graphite Reaction 38
Pressure Vessel Stress Analysis 39
Reactor Chemistry 39
Calculation of the Efficiency of Removal of Particles by an Air-Cooled Condenser 39
Solid State 40
ESR Powder Patterns 40
Subroutines 40
Higher Quality Printing for CDC 1604-A Printed Output 40
FORTRAN Output Without FORMAT Statements 40
Thermonuclear 41
Investigation of Compound Superconductors 41

SYSTEMS PROGRAMMING

IBM/360 Operating Systems 42
Operational Improvements 43
The Large Core Storage Nucleus 44
Operator Training 44
System Studies 44
Compiler Development 45
ORNL FORTRAN 45
General Purpose Compiler 46
Cathode Ray Display 46
Tape Utility 47

COMPUTER OPERATIONS

The IBM Computer 48
The CDC Computers 48
Work Load and Performance 50
Plotters 53
Operating Staff 53

MATHEMATICAL STATISTICS

INTRODUCTION 54

STATISTICAL RESEARCH 55

Combinatorial Aspects of 2^{k-p} Designs 55

Augmenting Response Surface Designs 57

Selecting the Population with the Smallest Variance 60

Selecting the Population with the Largest Mean 62

Prime Numbers Treated as a Markov Chain 63

Nonparametric Tests of Randomness 65

Reliability in a Three-Unit Redundant System 66

The Inverse of a Matrix Occurring in First-Order Moving-Average Models 68

An Inversion Method for Band Matrices 70

Uniqueness of a Minimum Norm Solution 73

Estimation of Parameters in Mixtures of Distributions 75

Application of the Logistic Model to Binomial Response Data 75

Kinetics of Antibody Production 77

Estimation of the Mean and Variance of a Censored Weibull Measurement Distribution 79

Numbers Generated by the Function $\exp(1 - e^x)$ 80

A Generalized Multidimensional Age-Dependent Branching Process 83

STATISTICAL APPLICATIONS 83

Analytical Chemistry 83

An Experimental Optimization Program for the Chemical Assay of tRNA 83

Biology 85

Pathology and Physiology 85

Mammalian Recovery 90

Cytogenetics 93

Radiation Immunology 93

Biochemical Regulation 94

Chemical Technology 95

Estimation of the Parameters in an Anion Exchange Model 95

Electronuclear 96

Electron Microscopy 96

Health Physics 96

Estimation from Pooled Samples 96

Half-Lives of ^{60}Co and ^{137}Cs in Leaves 97

ORIC Lake Study Analysis 97

Honeybee Irradiation Study 98

Numerical Taxonomy of Three Populations of Aquatic Isopods . . . 98
Resistance of Mosquito Fish to Irradiation . . . 98

Metals and Ceramics . . . 99
Relationship Between Fluorine Content and Hardness . . . 99
Analysis of Sintered Aluminum Product Data . . . 99
Defining Regions of Experimentation Under Loss of Rotatability . . . 100
Proposed Experiments in the Heavy Section Steel Technology Program . . . 101

Reactor Standards and Safety . . . 102

STATISTICAL PROGRAMMING . . . 102

The Time-Sharing Computer Terminal . . . 102

Biology Programming . . . 102
Radiation Mutagenesis . . . 102
Pathology Data Control: Autopsy Code . . . 103
Filter Cap Experiment Data . . . 103
Low-Level Radiation Experiment . . . 104
Foreign Spleen Experiment . . . 104

Statistical Methods . . . 104
Random Distributions . . . 104
Data Smoothing . . . 105
Plotting Package . . . 105

Data Processing . . . 105
Pitfall Data . . . 105
Conversion of RBC and Spleen Counts . . . 106
Specific Locus Mutation Study . . . 106

STATISTICAL SEMINARS AND LECTURES . . . 106

Seminars . . . 106

LECTURE SERIES IN STATISTICS AND PROBABILITY . . . 107

PROFESSIONAL ACTIVITIES . . . 108

MATHEMATICS DIVISION PUBLICATIONS . . . 109

MATHEMATICS DIVISION LECTURES AND PAPERS . . . 112

Mathematical Research

S. J. Chang A. S. Householder
J. M. Dolan E. J. Lee
G. K. Haeuslein W. L. Morris
F. W. Stallmann

QUESTIONS RELATING TO THE ZEROS OF A POLYNOMIAL

There are many methods in the literature for solving an algebraic equation, but each presents certain difficulties more or less peculiar to itself. Newton's method is fairly simple and rather rapidly convergent to any given root provided a suitable initial approximation is available. Bairstow's method can replace it for a pair of conjugate complex roots of a real equation, but again a suitable initial approximation is required. Laguerre's method converges more rapidly, and always converges if the roots of the equation are all real. But the arithmetic is somewhat more complicated, when there are complex roots an initial approximation is again required, and there seems to be no way to avoid complex arithmetic. The quotient difference algorithm requires no initial approximations and in principle produces all real roots directly as well as factors containing conjugate complex pairs, when the equation is real. But it can be unstable, and generally it must be followed by Newton, Laguerre, or some other iteration to refine the approximations it gives. Similar remarks apply to the Graeffe method. Lehmer has developed a method that is quite slow, but usually succeeds in isolating the roots sufficiently for the application of Newton or Laguerre. A former ORAU Fellow, G. W. Stewart III,[1] has analyzed this method as to stability and introduced certain improvements. Perhaps the most promising method to be found in the literature is due to J. F. Traub; it produces first a sequence of polynomials approaching $f(z)/(z-r)$, where $f(z)=0$ is the equation to be solved and r is one of its roots. At a certain stage, Newton's method is applied to the quotient of $f(z)$ divided by this polynomial, and a suitable modification takes care of the possibility that r is complex (when f is real), though with certain programming complications. This method is based upon one due to F. L. Bauer and Klaus Samelson, which, in turn, takes its departure from one due to Sebastiao e Silva.

A suggestion due to Stewart has lead to a quite general method which, though it has not yet been tested, seems to have real advantages over all others now known. It rests in a sense on all three of these methods, but goes a step farther. A general step of the method proceeds by two stages: Assuming f to be monic, two polynomials, $p_\nu(z)$ and $q_\nu(z)$, both monic and of degrees m and $n-m$, say, are at hand, and

[1] G. W. Stewart III, *Some Topics in Numerical Analysis,* ORNL-4303 (September 1968) (thesis).

$p_{\nu+1}$ and $q_{\nu+1}$, also monic and of the same degrees, are formed by the algorithms

$$p_{\nu+1}(z)\, q_\nu(z) + h_\nu(z)\, p_\nu(z) = f(z)\ ,$$

$$q_{\nu+1}(z)\, p_\nu(z) + k_\nu(z)\, q_\nu(z) = f(z)\ ,$$

where $h_\nu(z)$ is of degree $n - m - 1$ and $k_\nu(z)$ is of degree $m - 1$ (neither necessarily monic). Convergence can be accelerated, but at some expense in terms of computations per step, if $p_\nu(z)$ in the second step is replaced by $p_{\nu+1}(z)$. Coefficients of the polynomials $p_{\nu+1}$ and h_ν satisfy a system of linear equations whose determinant is the determinantal form of the resultant of p_ν and q_ν; and the same is true of $q_{\nu+1}$ and k_ν. Under suitable circumstances, if $f(z)$ has m zeros whose zeros are all strictly less in modulus than the other $n - m$, then p_ν and q_ν converge to polynomials $p(z)$ and $q(z)$ whose zeros are, respectively, the zeros of smallest and those of largest modulus. In case convergence fails, $p_\nu(z)$ and $q_\nu(z)$ can be replaced in a certain way by polynomials of degrees $m + 1$ and $n - m - 1$, respectively, and normally one would take $m = 1$ to start. The method will be tested and described in detail in an appropriate journal.

The problem of counting the zeros of a given polynomial that are located in a given half plane or a given circular disk is of importance in connection with studies of stability of dynamic systems. It arises also in connection with certain methods, notably Lehmer's, for solving an algebraic equation. There are three standard methods for doing this; all require forming a certain pair of polynomials from the given one. Two of these, one due to Routh and one due to Schur and Cohn, use modified Euclidean algorithms, and the second is the one Lehmer applies. The other is due to Hurwitz and uses resultants and subresultants in bigradient form.[2] It appears, however, that an old theorem due to Hermite, published in 1856, provides a simpler, more general, and more stable computation. This is based on the use of the Bezoutiant form of the resultant, which is a real symmetric matrix of order n, by contrast with the bigradient which is of order $2n - 1$, and the computation requires nothing more than a triangular factorization of the matrix with pivoting along the diagonal. This possibility of pivoting can considerably augment the computational stability, at least in certain cases. Curiously, the Hermite theorem seems to have been largely forgotten. However, it can be shown to be equivalent to that of Hurwitz, and as a by-product a much simpler proof for it has been developed, and one that also provides a much simpler proof of the Hurwitz theorem than any in the literature.

NUMERICAL SOLUTION OF INTEGRAL EQUATIONS

The linear integral equation of the first kind

$$F(x) = \int_a^b K(x, t)\, f(t)\, dt \tag{1}$$

leads to an ill-conditioned linear system if the integral is replaced by a finite sum according to some standard quadrature formula. If $f(t)$ is determined from such a system, small errors in $F(x)$ introduce large oscillations in $f(t)$, which makes this method unsuitable for practical applications. The following procedure has been tried successfully to overcome these difficulties.

We approximate

$$f(t) \approx \sum_{i=1}^{n} a_i(t)\ [F(\xi_i) + \epsilon_i]\ , \tag{2}$$

[2] A. S. Householder, "Bigradients and the Problem of Routh and Hurwitz," *SIAM Rev.* 10, 56–66 (1968).

the ξ_i being the points where $F(x)$ is measured and ϵ_i the measuring errors. Then

$$E_t(f, \epsilon) = f(t) - \sum_{i=1}^{n} a_i(t)\,[F(\xi_i) + \epsilon_i]$$

$$= f(t) - \sum_{i=1}^{n} a_i(t) \left[\epsilon_i + \int_a^b K(\xi_i, \tau) f(\tau)\, d\tau\right] \tag{3}$$

is a linear operator in f and ϵ. The coefficients $a_i(t)$ are then determined by minimizing a suitable norm of E_t.

One possibility for defining a norm arises from the following representation of E_t. Set

$$K^{[1]}(x,t) = -\int_a^t K(x, \tau)\, d\tau,$$

$$U(t, \tau) = 0\,, \quad a \leqslant \tau < t\,,$$

$$U(t, \tau) = -1\,, \quad t \leqslant \tau \leqslant b\,.$$

Then

$$E_t(f, \epsilon) = \int_a^b \left[U(t, \tau) - \sum_{i=1}^{n} a_i(t)\, K^{[1]}(\xi_i, \tau)\right] f(\tau)\, d\tau + \sum_{i=1}^{n} a_i(t)\, \epsilon_i\,, \tag{4}$$

if the $a_i(t)$ are chosen in such a way that

$$\sum_{i=1}^{n} a_i(t)\, K^{[1]}(\xi_i, b) = 1\,. \tag{5}$$

This leads to the following problem: Minimize

$$\int_a^b \left[U(t, \tau) - \sum_{i=1}^{n} a_i(t)\, K^{[1]}(\xi_i, \tau)\right]^2 d\tau + \gamma^2 \sum_{i=1}^{n} a_i^2(t)\,, \tag{6}$$

subject to condition (5). In (6) the first term represents the error, which is introduced in (2), if the exact values $F(\xi_i)$ are given. The second term takes the experimental errors into account under the assumption that they are equally and independently distributed. The parameter γ has to be chosen to balance between these two types of error. Its value is not critical and can be determined either from experience or from some rough a priori estimate of $f(t)$ and the distribution of ϵ_i.

This method is applicable to the inversion of any linear operator, and can be useful in a wide variety of situations in which instabilities and experimental errors are prominent features. We are now using it in a problem of x-ray scattering.

STRESS ANALYSIS OF GRAPHITE UNDER NEUTRON IRRADIATION

For the stress analysis of the graphite moderator under irradiation, the theory of linear viscoelasticity has been applied successfully.[3] The analysis includes the effects of neutron-induced creep and dimensional change for an arbitrary two-dimensional cross section. In fact, it is proved that a viscoelastic solution can be obtained readily by multiplying the corresponding elastic solution by an appropriate function of dose. If ϕ

[3] Will be published as ORNL-TM-2407.

denotes the stress function and the shrinkage function is given by

$$\psi(D, T) = A_1(T)\, D + A_2(T)\, D^2 \ ,$$

then

$$\phi(x, y, D) = \phi^a(x, y) + \phi^b(x, y)\, \frac{G(D)}{G(0)} + \phi^c(x, y)\, F_1(D) + \phi^d(x, y)\, F_2(D) \ ,$$

where ϕ^a, ϕ^b, ϕ^c, and ϕ^d are elastic solutions, with Young's modulus $G(0)$, due to boundary traction, thermal expansion, and shrinkages $A_1(T)$ and $A_2(T)$, respectively, and

$$F_1(D) = \frac{1}{G(0)} \int_0^D G(D - D')\, dD'$$

and

$$F_2(D) = \frac{1}{G(0)} \int_0^D G(D - D') \cdot 2D' \cdot dD' \ .$$

$G(D)$ is understood as the relaxation function for the material with the neutron dose D. Thus the numerical calculation reduces to that for an elastic problem.

Based on the linear viscoelastic model, the toughness for the radiated graphite was shown to depend inversely on its radiated Young's modulus and on the stress concentration factor but to be independent of the history of the loading process which is in general the difference between the elastic analysis and the viscoelastic analysis.[4]

OSCILLATORY BEHAVIOR OF A LINEAR THIRD-ORDER DIFFERENTIAL EQUATION AND ITS ADJOINT

Third-order differential equations are of current interest in applied mathematics because they arise in problems in reactor technology, aerodynamics, and electricity. Of particular interest is the oscillatory behavior of solutions of the linear differential equations

$$y''' + p_2 y'' + p_1 y' + p_0 y = 0 \tag{1}$$

on $[a, \infty)$ and its formal adjoint

$$\left((z' - p_2 z)' + p_1 z \right)' - p_0 z = 0 \tag{1*}$$

also on $[a, \infty)$. Here the coefficients are continuous. Let S and S^* be the solution spaces of (1) and (1*) respectively.

It can be shown that the oscillatory behavior of S and S^* are related under certain circumstances. These relationships lead to some separation theorems among solutions in certain subspaces of S and S^*.

An example was found in which (1) has no nontrivial oscillatory solutions but is not disconjugate on any half-line.

---

[4]Submitted for publication.

CONSTRUCTION OF FIELDS WITH GIVEN GROUP

The question as to which algebraic equations

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n = 0 \tag{1}$$

with coefficients in a field k can be solved by radicals is answered by Galois theory, which associates to each algebraic equation (1) the group of k-automorphisms of its splitting field. Given a finite group G and a field k one might wonder if there are algebraic equations (1) over k the group of which is isomorphic to G. This is true, for example, for all finite symmetric groups and alternating groups.[5] The general problem, however, is still open. For some n the group G can be represented as a permutation group operating on n variables $x_1, x_2, \ldots, x_n$. If the field of all rational functions $g(x_1, x_2, \ldots, x_n) \in k(x_1, x_2, \ldots, x_n)$ which are invariant under G is purely transcendental over k, the existence of an equation (1) with group G has been established.[6] This was the reason for computing the fields of invariants for special groups: the quaternionic group and the dihedral groups D_n. The method applied was basically due to Fischer.[7] In each case the field of invariants turned out to be purely transcendental over k. The computations are rather complicated, and it is not clear how they can be generalized to arbitrary solvable groups. The field k was supposed to contain certain roots of unity. This latter assumption can be dropped for the dihedral groups D_3 and D_4, and presently it is checked if it is redundant for all D_n. This is important if one wants to prove the existence of equations with rational coefficients the group of which is D_n.

NUMERICAL SOLUTIONS OF ORDINARY INITIAL VALUE PROBLEMS

A new algorithm for the numerical solution of ordinary initial value problems was developed and is now available in the ORNL Subroutine Library as ORNL Subroutine D02013.[8] This algorithm represents an attempt to avoid numerical instability by automatically adjusting the so-called step in the solution according to a criterion for which there is a rigorous mathematical basis in the case of linear homogeneous systems.

ROOTS AND PROPER VECTORS OF A REAL SYMMETRIC MATRIX

A new algorithm for the calculation of all roots and all proper vectors of a real symmetric matrix was developed and is now available in the ORNL Subroutine Library as ORNL Subroutine F02013.[9] The method is based upon minimization of associated Weinstein discs in two-dimensional subspaces. The novel feature of the algorithm is an efficient criterion for choosing the various 2-spaces. In comparison with other popular methods, this algorithm is very slow, but it produces roots to greater accuracy and proper vectors that are superior in being more nearly mutually orthogonal.

---

[5] D. Hilbert, "Über die Irreduzibilität ganzer rationaler Functionen mit ganzzahligen Coefficienten," *Crelles J.* **110**, 104–29 (1892).

[6] E. Noether, "Gleichungen mit vorgeschriebener Gruppe," *Math. Ann.* 78, 221–29 (1917–18).

[7] E. Fischer, *Die Isomorphie der Invariantenkörper der endlichen Abelschen Gruppen linearer Transformationen,* Nachrichten Göttinger Akademie der Wissenschaften, 1915, pp. 77–80.

[8] D. W. Altom and W. L. Morris, *An Algorithm for Initial Value Problems,* ORNL-4282 (August 1968).

[9] W. L. Morris, *Minimal Weinstein Discs in a Subspace,* ORNL-4151 (November 1967) (thesis).

ARITHMETICAL FUNCTIONS RELATED TO THE FIBONACCI NUMBERS

Tables of six Fibonacci functions were calculated for all arguments up to 15,600.[10] The tables are believed to be quite reliable since all data processing was done in core memory of the IBM 360/75. Discovery of two of these functions was reported in the previous annual report.

AMICABLE NUMBERS

Early this year an article was published describing search methods for amicable pairs. The article includes a list of 264 new amicable pairs.

Heuristic support is found for a model of the distribution of amicable numbers inferred from observation. The model is in striking agreement with the known distribution. Results are submitted for publication as a conjecture, since a rigorous theory of the model will require deep analytic number theory going beyond any results yet obtained in that field.

In the submitted article a theorem is proved giving the possible forms of even amicable pairs whose sums are not divisible by 9. The article describes a computer search for amicable numbers whose forms are given by a corollary to the theorem. Six new even pairs were found which have sums $\equiv 2 \bmod 9$. Somewhat surprisingly all these numbers have digital roots of 1.

Work on a number of other topics involving aspects of amicable numbers and related subjects has been essentially completed, and papers are being prepared on the following subjects:

1. Probabilistic Finitude of Certain Types of Amicable Pairs,
2. Ratios of Amicable Numbers,
3. A New Search Method for Some Types of Amicable Numbers,
4. Some Probabilistic Aspects of Catalan's Empirical Theorem.

---

[10] Beth H. Hannon and W. L. Morris, *Tables of Arithmetical Functions Related to the Fibonacci Numbers,* ORNL-4261 (June 1968).

Computer Applications

SPECIAL PROJECTS

D. E. Arnurius
T. D. Calton
Nancy E. Dean
R. C. Durfee
Marie H. Eckart
Manuel Feliciano, Jr.
N. B. Gove
G. K. E. Haeuslein
A. F. Joseph
W. L. McMullen, Jr.
C. W. Nestor, Jr.
Shirley B. Opstrup
Marilyn M. Perardi

General Purpose System

We have added new procedures for interpolation, plotting, FORTRAN subroutine linkage, and program control to the ORDEAL system.[1] The program control procedures allow conditional branching and looping within the users' ORDEAL program.

Approximately 2000 procedures have been executed by a small group of users during the period May to October 1968 with production results. The READ and FIT procedures accounted for nearly half of the procedures executed during that period.

We are currently studying and incorporating other procedures for the ORDEAL system. We are extending the FIT procedure for various nonlinear models and will soon make available new procedures for statistical analysis, integration, and smoothing and root solving.

We are finishing a users' manual and should release it early in 1969. A list of the current available procedures as of October 1968 follows, with brief descriptions of each.

Procedure	Brief Description
READ	Procedure READ is used to input data to the ORDEAL program. Modifiers describe how and where to read data.
PRINT DISPLAY WRITE	These three procedures are used to print data from ORDEAL program.
HEADING	Procedure HEADING is used to produce a title or heading at the top of the next page of printed output.

[1] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 12.

Procedure	Brief Description
COLUMNS	Procedure COLUMNS is used to label data columns when using the WRITE procedures.
SUMMATION	Procedure SUMMATION is used to add all the elements of a given data set.
MAXIMUM MINIMUM	These two procedures find the maximum or minimum element within a given data set.
ABMAX ABMIN	These two procedures find the absolute maximum or minimum element within a given data set.
JOIN	Procedure JOIN is used to concatenate two data sets into one data set.
DIMENSION	Procedure DIMENSION determines the size of an existing data set.
DEFINE	Procedure DEFINE defines a data set with certain size and value requirements.
FIT	Procedure FIT fits a model described by modifiers to a given pair of data sets in the least-square sense.
CALL	Procedure CALL allows control to pass to a user-supplied FORTRAN subroutine.
SUBSET	Procedure SUBSET is used to edit data sets.
LOOP	Procedure LOOP is used to repeat a section of the ORDEAL program a number of times depending on conditional modifiers.
JUMP-TO	Procedure JUMP-TO allows the ORDEAL user to branch to a certain section within the ORDEAL program.
LABEL	LABEL identifies a certain section of the ORDEAL program for use by the LOOP or JUMP-TO procedures.
INTERPOLATE	Procedure INTERPOLATE is used to interpolate given data points for intermediate values. The degree and type of interpolation are specified by modifiers.
PLOT	Procedure PLOT is used to plot data set pairs on the Cathode Ray Tube (CRT) or mechanical plotters. Scaling and labeling requirements are specified by modifiers of the procedure.

Spectrum Analysis System

We developed a system for processing, analyzing, and plotting spectral data. The program, GAMSPEC3, is composed of various independent modules which contain subroutines to perform certain requested tasks.

The repertoire of services available at present is limited but includes data smoothing, background subtraction, peak hunting, peak area analysis, library comparison, and plotting (pen-and-ink and cathode ray tube). The system will be expanded to include other services with several methods available for each service.

The program is user-oriented in that the input of the spectral data may be either magnetic tape or cards with variable formats and variable record lengths for both the identification numbers and data. Also, the request cards have a free format; that is, the information may be placed anywhere on the card with the only restriction that options are separated by commas.

For example, if the following cards were used with the system:

```
SPECTRUM 3, 2048.
TAPE ID FØRMAT=(I8), DATA FØRMAT=(10F6.0).
20203 + 20204 – 2. * 20201
REGIØN 1, 2048.
PLØT
LIST SPECTRUM
LIST ØF CØNTINUØUS SUMS
```

REGIØN 100, 2000.
SMØØTH DATA
FIND PEAKS
FIT PEAKS
PLØT, CHANNELS/INCH=50, CYCLES=3, BEGIN=200, END=350.

then three spectra of 2048 channels each would be read in from a tape with the formats specified. Spectrum number 20203 would be added to 20204 and two (2) times 20201 would be subtracted (this might be the background). The REGIØN card indicates that the whole spectrum will be analyzed. A semilog plot will be generated with four cycles, 40 channels per inch, and permanent rescaling. (These are the default options.) A list of this combination of spectra will also be generated, along with a list of continuous sums. The next REGIØN card sets up new boundaries for the spectrum, and the data between channels 100 and 2000 will be smoothed and searched for peaks, and once the peaks are found, they will be analyzed. Finally, another plot will be generated with 50 channels per inch and three cycles. This plot will begin in channel 200 and end in channel 350. This may be used to look at one particular peak.

Figure 1 is a photograph of a spectrum that was plotted using the cathode ray tube plotter.

Family of Heat Transfer Programs

We are developing a family of programs for the IBM 360/75 designed for use in heat transfer problems. It has become clear that no single program can be general enough to serve the needs of the many heat

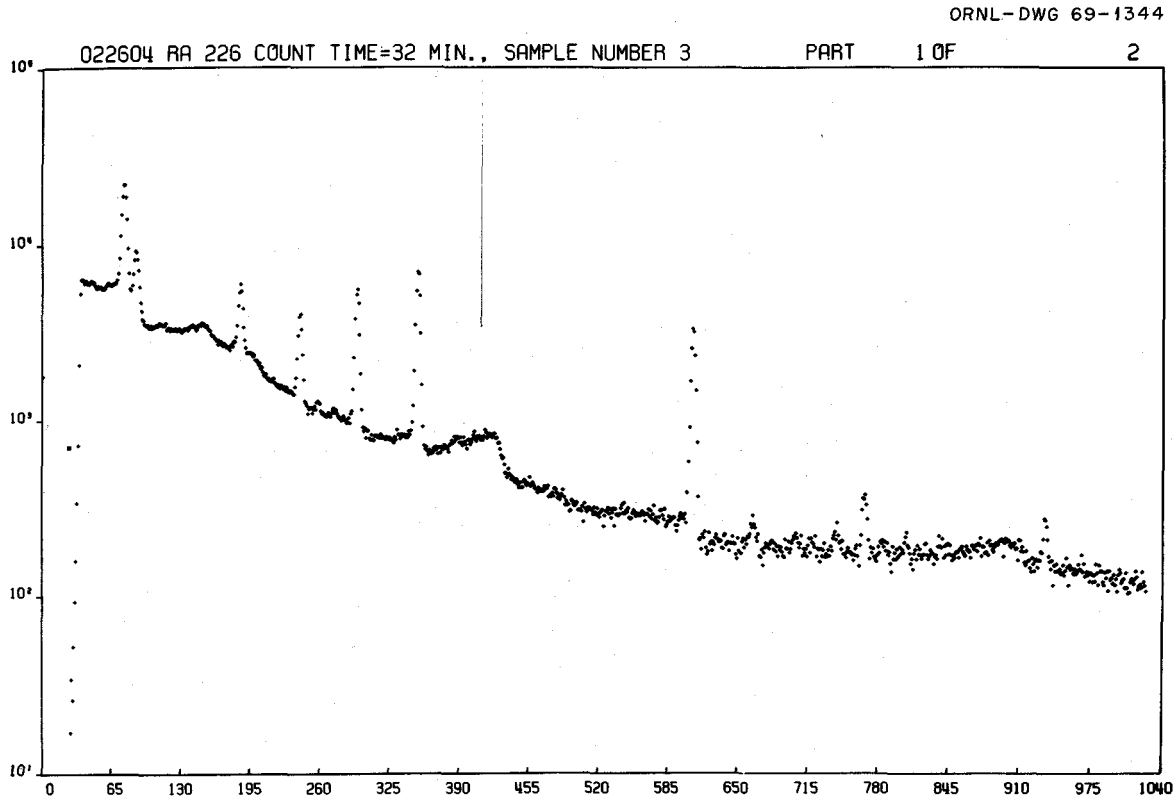


Fig. 1. Spectrum Plotted Using the Cathode Ray Tube Plotter.

transfer analysts in the Laboratory, but it is also clear that many heat transfer problems have features in common. The approach that we have taken is to design a set of programs that can be used in a sequence that best fits a particular need.

The major goal in the development of this family or system of programs is simplicity from the user's point of view. The user must be able to specify his problem with a minimum amount of input, to select computational procedures appropriate to his problem, and to extract the results he really needs. In each of these areas there are important programming considerations, but the chief concern of the program designer is data management. There are interesting similarities between the heat transfer problem and other large-scale problems, and, for the most part, these similarities have to do with the need for handling large amounts of data efficiently.

Our work during the year has been concentrated in the development of programs for the solution of steady-state and transient conduction problems in the commonly encountered two-dimensional geometries. In the steady-state problem we have studied several devices for accelerating convergence and will extend this study to the case in which the thermal properties are temperature dependent. In the transient problem, we include an approximate treatment of melting and freezing using a correct latent heat of fusion.

We have also developed a number of routines for printed and graphical output. In particular, two routines for producing temperature traverses and temperature maps make use of the Calcomp CRT microfilm plotter. Sample pictures are shown in Figs. 2 and 3.

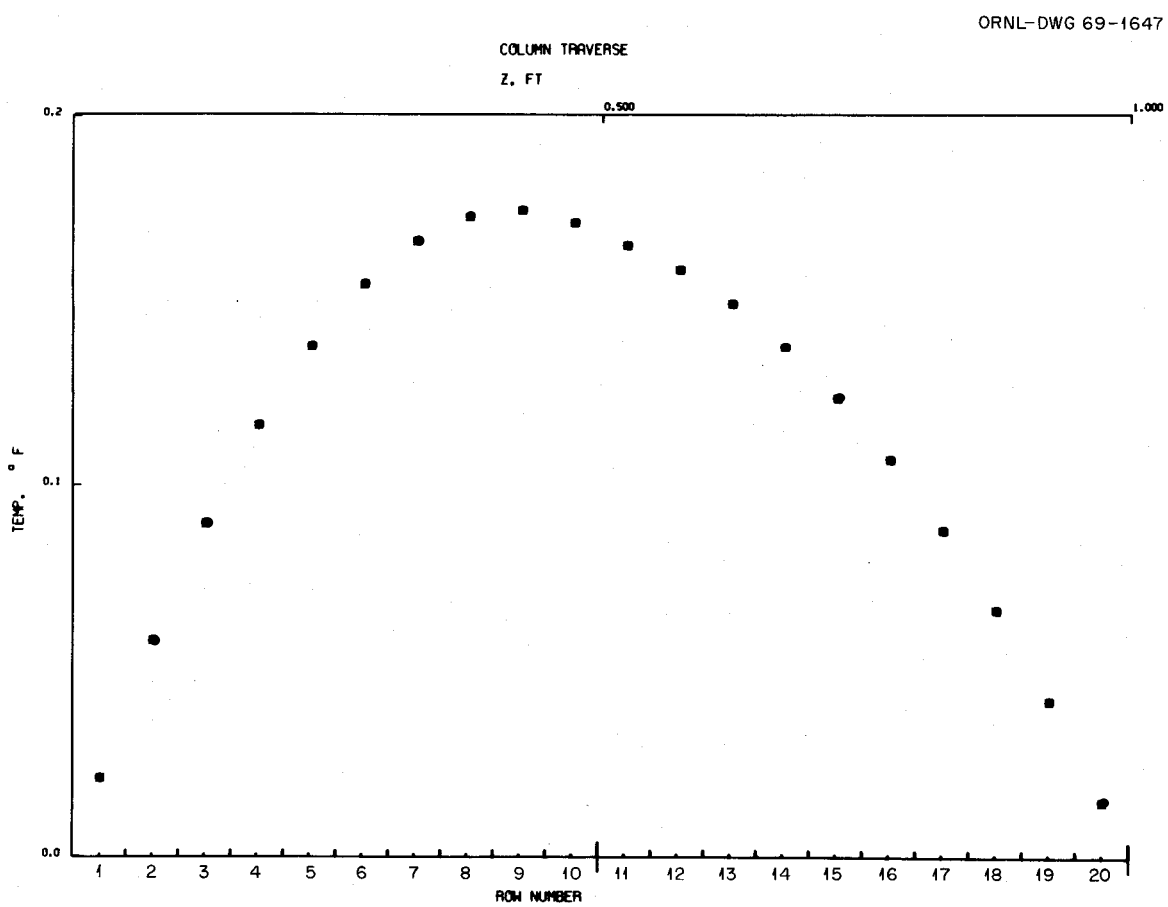


Fig. 2. Temperature Traverse for Sample Problem.

ORNL-DWG 69-1648

HEAT TRANSFER PROBLEM

SAMPLE STEADY STATE CASE

REGIONS

ROWS

A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A

A	0.02	-	0.08
B	0.08	-	0.12
C	0.12	-	0.17
D	0.17	-	0.20

Fig. 3. Temperature Map for Sample Problem.

Graphics Program Development

Early this year three new electrical plotting devices were installed in the computer area. These new devices are faster and better than the previous plotters. Some software development was necessary to make these new plotters compatible with old programs. This is finished, and the new plotters are in heavy use.

We have written a new plotting subroutine package. The emphasis is on usefulness, but some attention has been given to efficient plotting techniques. Its versatility and modular design obviate reprogramming parts of the package; modifications can be made easily in the user's program.

The program can be applied to the cathode ray tube plotter or to either of the pen-and-ink plotters. It is on a disk library, so that the user does not need card decks for the plotting package. Only the parts needed are loaded. Data to be plotted may be single or double precision. The user can specify a scale factor in the x or y coordinates or have the program select scale factors so that all the data fit on the graph. The data to be plotted may be given in user's units or in graph inches. The various parts of the package communicate with each other through a common parameter area, so that scaling information, for example, need not be repeated.

A curve generator has been included to plot any specified function. There are facilities to help make graphs acceptable for publication. For example, when grid lines are drawn an option has been provided to leave a blank space in anticipation of a legend or other insert. A legend subroutine has been provided. Finally, it is not a closed package; new features can and will be added.

Figure 4 shows one of the types of graphs that can be obtained. This graph required several calls to subroutines in the package. We also wrote several subroutines which combine calls to part of the package,

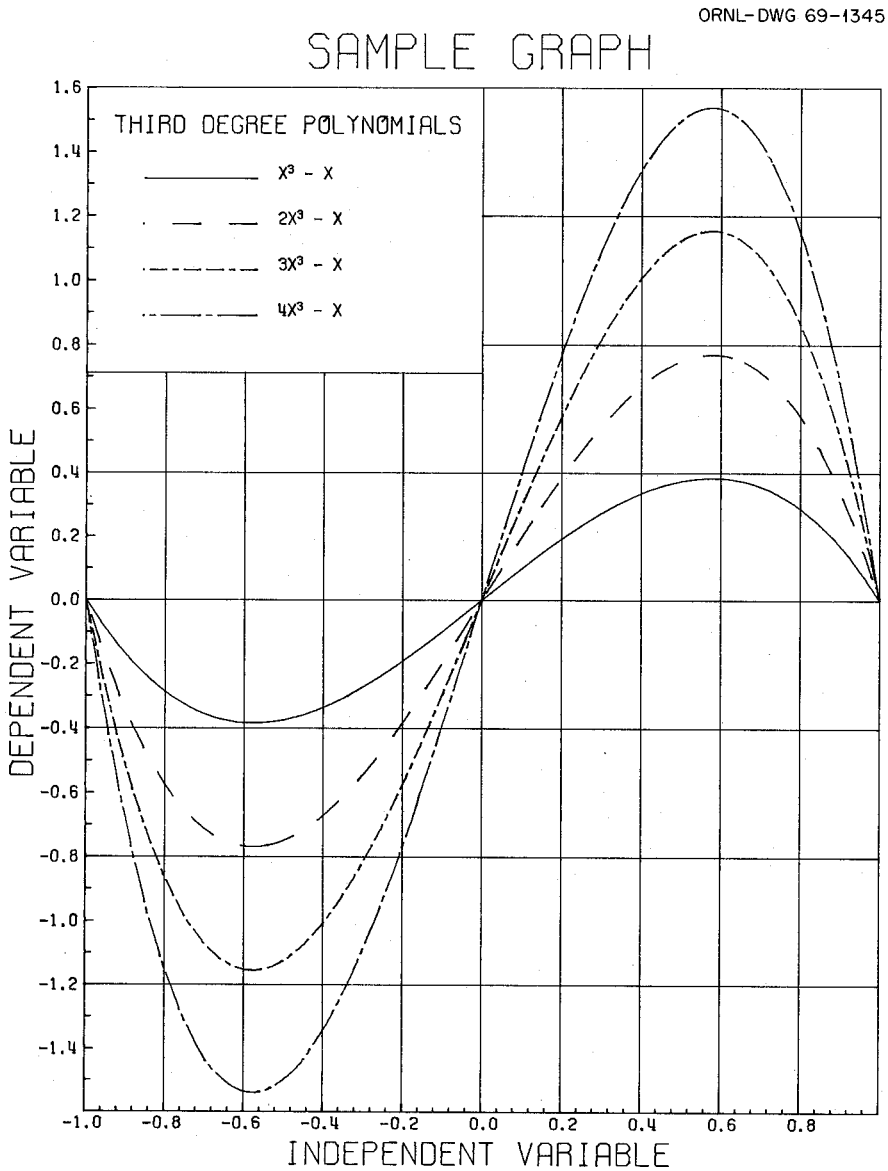


Fig. 4. Sample Plot with Several Calls to the Graphics Package.

so that plotting of routine kinds of graphs becomes very easy. The graph shown in Fig. 5 requires only one call to the package.

We wrote subroutine PERSPECT to draw a perspective plot of a set of points as observed from a given view point. The points are projected on a picture plane that is perpendicular to the line of observation (origin to view point) and unit distance from view point. All projections converge at the view point. The picture plane may be rotated around the line of observation. The projection on the picture plane is drawn, with points connected by straight lines, on the Calcomp 835 cathode ray tube plotter. A sample plot is shown in Fig. 6.

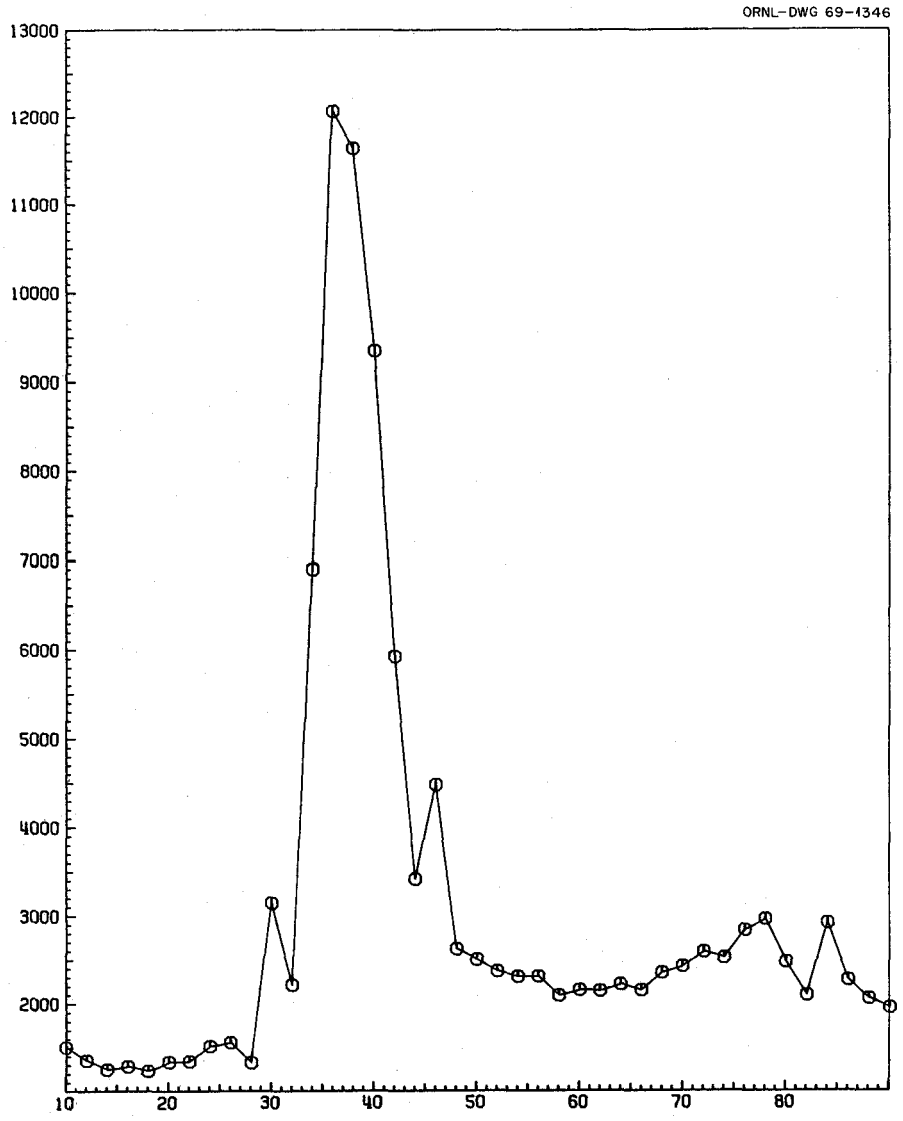


Fig. 5. Sample Plot with One Call to the Graphics Package.

Information Retrieval

We have completed the first version of ORLOOK, an information retrieval program for small files. Any file consisting of cards or card images on tape may be searched. The program will accept a large variety of file formats so that files need not be rekeypunched for use with this program. Retrieval is based on string matching. Retrieval requests involving Boolean "and" and "or" are also possible.

Programming for the Library KWIC Index

The ORNL Central Research Library uses a KWIC index program which was written by Bell Laboratories for the IBM 7090. This program has been rewritten for the IBM 360 system. In the new version, trivial key word entries are not eliminated immediately. At the end, all entries are sorted and run

ORNL-DWG 69-1347

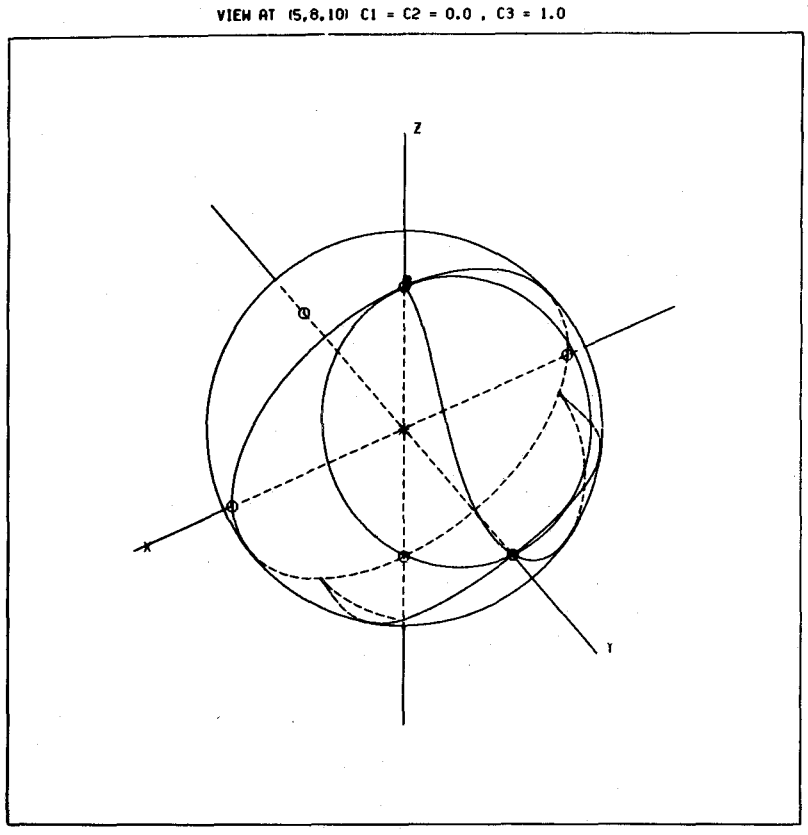


Fig. 6. Sample Plot Prepared with Subroutine PERSPECT.

against the sorted list of trivial key words. This method results in considerable savings of execution time. Future versions should accept control cards to permit change of input formats.

FORTRAN Source Program Documentation

We have written a computer package, FORTDOC, for the IBM 360/75 which will help programmers document their FORTRAN source programs. The package may also prove helpful in debugging programs or converting programs from one FORTRAN compiler to another. FORTDOC takes a FORTRAN source program as input and produces an expanded source listing, a new source deck, and several information tables. The package will accept ORNL FORTRAN, which includes FORTRAN IV, FORTRAN 63, and FORTRAN II as subsets. This package uses some of the subroutines written for the ORNL FORTRAN compiler.[2]

The expanded source listing contains resequenced statement numbers, DO loop brackets, sequence numbers, and statement-type indicators. There are three information tables. The first is a cross-reference table for statement numbers, the second is a cross-reference table for variable names, and the third is an input-output table defining all input-output activity in the program. All of these features are explained in detail in a report to be published.[3]

---

[2] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 30.

[3] R. C. Durfee, *FORTRAN Source Program Documentation,* ORNL-CF-68-12-7 (in preparation). (For internal use only.)

COMMGEN

We wrote a PL/1 program, COMMGEN, which will alter FORTRAN programs so that arrays may be dimensioned at execution time. The COMMON and DIMENSION statements in the main programs are written with variable dimensions such as:

```
COMMON /FAST/ A(IMAX,KMAX), B, C(100,LMAX)

DIMENSION D(IMAX,JMAX), E(KMAX), F(LMAX,2,10), G(FIVE)
```

When the user wishes to compile and execute his program, he supplies his FORTRAN deck preceded by the necessary input cards as input to COMMGEN. The input cards might look like:

```
NVAR = 5        PRINT        PUNCH

IMAX = 20    JMAX = 10    KMAX = 100    FIVE = 5

LMAX = (KMAX/JMAX+2*IMAX)
```

COMMGEN then replaces all the specified variables (IMAX, JMAX, KMAX, LMAX, and FIVE) with their appropriate values, writes a new FORTRAN program onto disk, prints the new program (PRINT was specified), and punches the new program (PUNCH was specified). The next job step then compiles and executes the new FORTRAN program which was written onto disk.

If a variably dimensioned program is used frequently with different dimensions, the user may wish to put it onto disk. COMMGEN will then read the program from disk rather than cards. The first input card would have an extra parameter:

```
NVAR = 5    INPUT = DISK    PRINT    PUNCH
```

Multiple Factor Analysis

Cluster Analysis. – We have written a cluster analysis program to analyze a group of variables to find the best set of clusters which describe these variables. The clusters found generally agree closely with the common factors calculated by another program, FACTOR ANALYSIS I.[4]

The problem may be thought of geometrically with each variable representing a test vector in the common factor space. The program then calculates the cosines between every pair of test vectors. Using these cosines, the program finds those test vectors which are grouped or clustered together. Most of the time there are at least as many clusters as there are common factors.

The input to this program is either correlation matrix or the principal axis factor matrix from FACTOR ANALYSIS I.[4]

Primary Cluster Structure. – We have written a program, PRIMARY, to find the primary factors and the primary structure matrix from a raw score matrix. This program is to be used in conjunction with two

---

[4]R. C. Durfee, *Multiple Factor Analysis,* ORNL-TM-1798 (Mar. 10, 1967).

other programs, CLUSTER ANALYSIS and FACTOR ANALYSIS I.[4] These two programs define the clusters and the number of factors, N, in our common factor space. PRIMARY uses the cluster centroids from the N best clusters to approximate the primary factors. Once we know the primary factors we can find the hyperplanes which bound our common factor space. The primary factors lie along the intersections of these hyperplanes. Now we are able to find the primary structure matrix, because the oblique vectors which define our structure matrix are perpendicular to the hyperplanes.

It is hoped that our primary structure matrix will be very similar to the simple structure matrix which can be calculated by another program, FACTOR ANALYSIS III.[4] This primary structure matrix can be used to interpret the N factors which describe our original raw score matrix.

Statistical Calculations

We have written two programs, CORR and CORRMISS, to calculate frequencies, means, standard deviations, and correlations for a group of variables. During the calculation of the correlation matrix, the first program replaces missing data with their means. The second program is more sophisticated; it maintains a frequency count of the valid observations between every pair of variables. The missing data may then be eliminated so that they do not influence the calculation of the correlations.

Two Programming Systems

We received and made available for general local use two programming systems: the Continuous System Modeling Program (CSMP) and the PL/1 Formac Interpreter. CSMP is an IBM-supported system and the PL/1 Formac Interpreter is an IBM-distributed system. The CSMP system was developed to simulate continuous systems and it provides an application-oriented input language that accepts problems given in the form of an analog block diagram or a system of ordinary differential equations. With the PL/1 Formac Interpreter the PL/1 language is extended to include the Formac language. Formac provides the capability to perform formal manipulations (differentiation, simplification, etc.) on mathematical expressions. Only a little knowledge of PL/1 is required to use the Formac capabilities in the PL/1 Formac Interpreter. User's manuals[5] for these two systems and descriptions of the necessary control cards for local use are available from the ORNL Computer Librarian. Moreover, persons familiar with these programming systems are available for consultation.

It should be noted that several errors in the CSMP system were uncovered early this year. Since then a new version of CSMP has been supplied by IBM. No evaluation of the new version is available yet.

Mathematical Programming System, MPS/360

During 1968 we used the mathematical programming system, MPS,[6] to solve several large linear programming problems on the IBM 360/75 computer. Linear programming is a mathematical technique for determining the optimum solution to a system of linear inequalities.

During 1967 and the first few months of 1968, we used version 1 of MPS, which consisted primarily of a set of linear programming procedures. In February of this year we received version 2, which contained a separable programming algorithm and several other new features. Separable programming is a mathematical

---

[5] *System/360 Continuous System Modeling Program,* Publication No. H20-0367-2; *PL/1 Formac Interpreter,* Publication No. 360D-03.3.004.

[6] *Mathematical Programming System/360 Application Description, IBM Application Program,* H20-0136-3 (1968).

programming technique to solve problems containing nonlinear (but separable) variables within a linear programming framework. Since February, four modifications have been made to MPS to correct known errors and speed up execution time.

At present, MPS is composed of three programs: (1) MPS/360, which contains a control language and control language compiler,[7] a set of linear and separable programming procedures,[8] and a FORTRAN interface (READCOMM);[9] (2) MPS/360 Report Generator,[10] which consists of a report language and three procedures to compile and execute the report program; and (3) MARVEL,[11] a language processor for matrix generation, output analysis, and report writing. MPS will handle a maximum of 4095 constraints (or rows) with our computer configuration. In addition to solving the original linear programming problem, MPS will allow the user to vary cost or requirement data to determine the effect on the optimal solution.

LECTURE SERIES

Nancy B. Alexander Marie H. Eckart

FORTRAN Lecture Series

We conducted three series of lectures on the FORTRAN language, two at the X-10 site and one at the Y-12 site. About 210 persons attended.

We made some changes in the format and content which improved audience participation and interest.

PROGRAMMING ASSISTANCE

Nancy B. Alexander Marie H. Eckart
Nancy E. Dean Carol D. Paulk

We have continued the programming assistance service that was begun in July 1967. This service is available to all users of the Laboratory computing facilities. Usually two programmers are available to answer programming questions. About 8000 questions were asked, 42% of them by phone.

Below is a list of categories of questions asked and the percentage of the total number of questions.

JOB CONTROL LANGUAGE	
IBM 360	14
CDC 1604-A	4
DEBUGGING	
ORNL FORTRAN	12
FORTRAN-IV (Level H)	9
FORTRAN-63	9
MESSAGES and CODES	10
INPUT/OUTPUT ERRORS	5
FORTRAN LANGUAGE	16
SYSTEMS PROBLEMS and ERRORS	3
MISCELLANEOUS	18

[7] *Mathematical Programming System/360 Control Language, IBM Application Program,* H20-0290-2 (1968).

[8] *Mathematical Programming System/360 Linear and Separable Programming User's Manual, IBM Application Program,* H20-0476 (1968).

[9] *Mathematical Programming System/360 Read Communications Format (READCOMM), IBM Application Program,* H20-0372 (1968).

[10] *Mathematical Programming System/360 Application Description, IBM Application Program,* H20-0136-3, pp. 14–15 (1968).

[11] *MARVEL/360 Program Description Manual, IBM Application Program,* H20-0496 (1968).

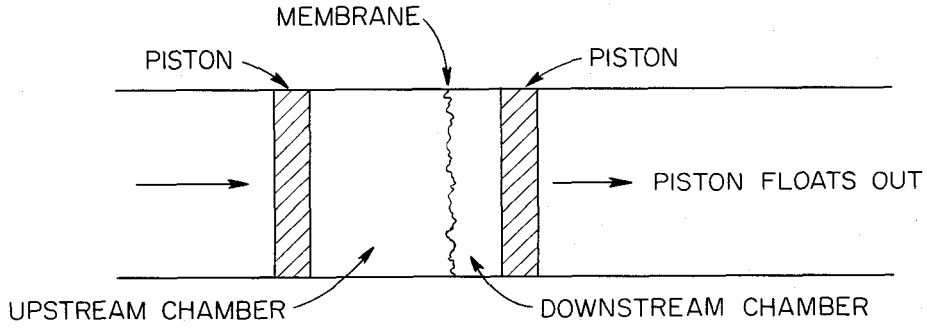
CHEMICAL TECHNOLOGY

D. E. Arnurius
Nancy E. Dean
R. C. Durfee
Sandra H. Merriman
C. W. Nestor, Jr.
Carol D. Paulk
V. A. Singletary

Noble-Gas Separation Project

Flow systems for separating mixtures of gas (oxygen, nitrogen, krypton, and xenon) by the use of permselective membranes are being studied by members of the Chemical Technology Division. In each stage of the flow system a mixture of gas is passed into a chamber separated by a membrane from another chamber. The noble gases permeate the membrane at a faster rate than the carrier gases and thereby concentrate in the downstream chamber of the stage. With the appropriate number and arrangement of the stages in a flow system, a prescribed concentration of a noble gas in the last downstream chamber is theoretically obtainable.

An integral part of the study is the determination of the permeation rates of a given membrane for the several gases subject to varying operating conditions that are controllable. The preliminary part of the study has dealt with an investigation of a single modified stage in which a binary gas mixture (oxygen or nitrogen with krypton or xenon) is considered.



The pressure in the two chambers is constant, with the upstream pressure greater than the downstream pressure. If $\hat{P}(V)$ represents the partial pressure of the noble gas in the upstream chamber for a given volume V, then the function $\hat{P}$ satisfies (approximately) the differential equation

$$P(V) + \frac{1-a}{V} - \frac{b}{V^0 - V}\hat{P}(V) = -\frac{bV^0\,\hat{P}(V^0)}{V(V^0 - V)}\,, \qquad V \in [0, V^0]\,, \tag{1}$$

where V^0 is the initial volume in the upstream chamber (initial volume of downstream chamber is zero), $a > b + 1 > 1$, and a and b are functions of the operating environment and physical properties of the membrane. The solution to the differential equation (1) is at this time the best analytic approximation to the system. Laboratory measurements of $\hat{P}$ have been made, and they agree favorably with the values given by the solution of (1). Because of this agreement the solution of (1) is used to approximate $\hat{P}$ on $[0, V^0]$.

It is easily shown that the solution of (1) is given by

$$\hat{P}(V) = \frac{b\,\hat{P}(V^0)\,V^0}{V^{1-a}(V^0 - V)^b} \int_V^{V^0} \frac{(V^0 - t)^{b-1}}{t^a}\,dt\,, \qquad V \in [0, V^0]\,. \tag{2}$$

In order to evaluate (2), we have found that it is more convenient to regard the partial pressure of the noble gas as a function of the cut fraction

$$F \equiv \frac{V^0 - V}{V^0}\,, \qquad V \in [0, V^0]\,, \tag{3}$$

which represents the fraction of the volume that has passed through the membrane. With this it follows from (2) with $\hat{P}(V^0) = 1$ and from (3) that

$$P(F) = \hat{P}(V) = \frac{b}{F^b} \int_{1-F}^{1} [U - (1 - F)]^{b-1}\, U^{a-b-1}\, dU\,, \qquad F \in [0, 1]\,. \tag{4}$$

Repeated integration by parts of the integral in (4) yields an infinite series representation for P:

$$P(F) = \sum_{i=0}^{\infty} A_i F^i\,, \qquad F \in [0, 1]\,, \tag{5}$$

where

$$A_0 = 1\,, \qquad A_i = \prod_{k=1}^{i} \left(1 - \frac{a}{b + k}\right).$$

Moreover, it can be shown that $A_i \to 0$ and $iA_i \to 0$ as $i \to \infty$.

We have written a computer program to approximate P by using partial sums of the series in (5) with a varying number of terms. The number of terms in the partial sum is small and is chosen so as to obtain a prescribed number of significant digits in the approximation. It should be noted that rounding errors in the computer calculation are ignored, but since only three or four significant digits are necessary and since the important region for F is $[0.25, 0.5]$, their influence will be small. As F increases in $[0, 1]$ the number of terms in the approximation increases. This number is reported next to the approximation. A sample plot of $P(F)$ vs F is also shown in Fig. 7.

Optimization of Sizes and Locations of Reprocessing Plants in an Expanding Economy

A project was started last year[12] to determine the optimum size, time, and location of nuclear fuel reprocessing plants in the United States during the period 1970–2039. A linear programming model was developed to minimize the total cost, and the mathematical programming system, MPS[13], for the IBM 360/75 was used to solve the problem. The United States was divided into eight geographical regions which were the same as those used by the USAEC and the Federal Power Commission.[14] Seventy time periods were used, each period corresponding to a year. In the model were considered, among other items, shipping

---

[12] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236.

[13] *Mathematical Programming System/360 Linear and Separable Programming User's Manual,* IBM Application Program, H20-0476, 1968.

[14] *Federal Power Commission National Power Survey – 1964,* Part II, 1964.

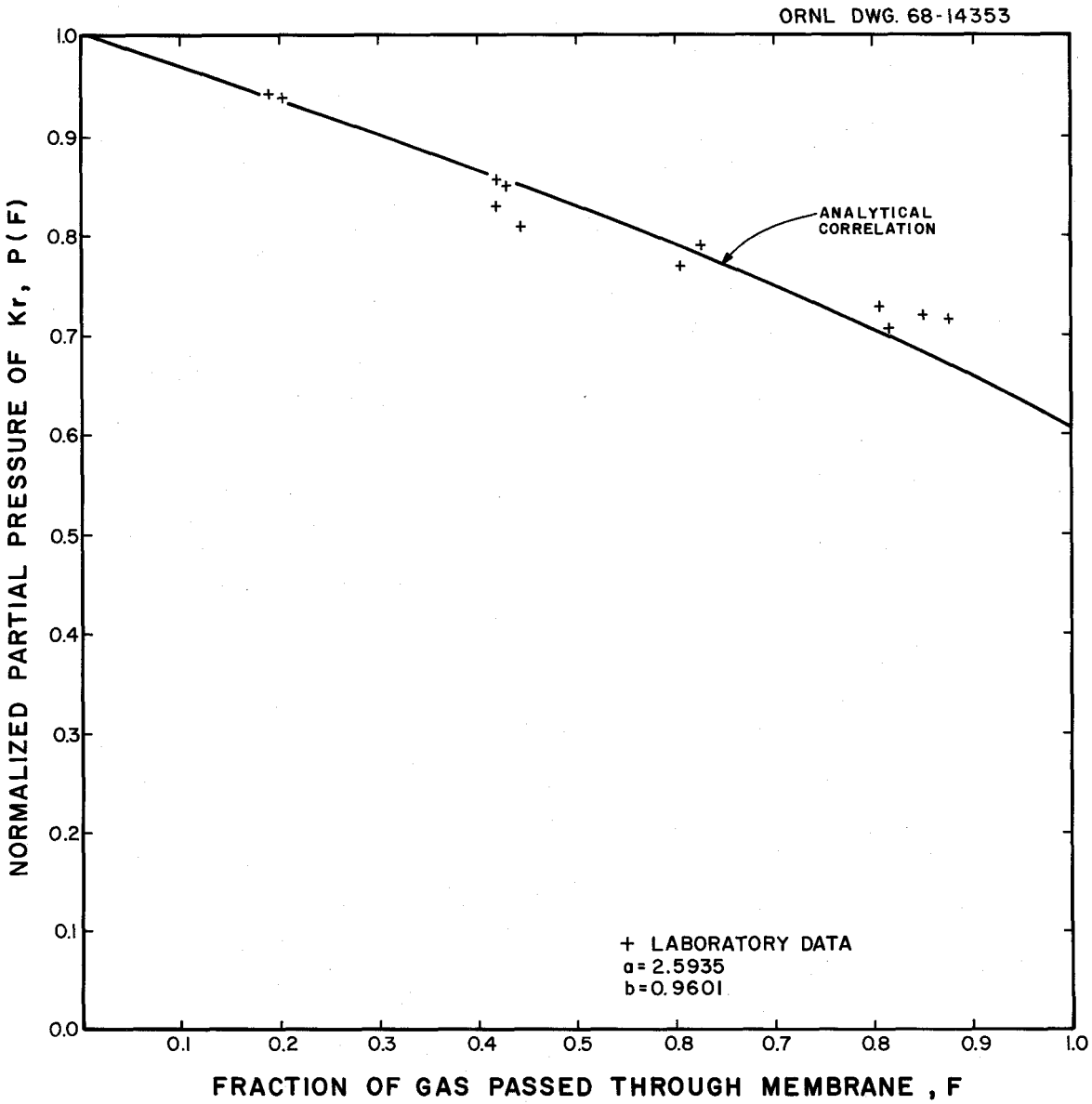


Fig. 7. Comparison of Analytical Correlation with Laboratory Data for Kr-O_2 System.

costs between reactors and reprocessing sites, stockpiling of spent fuel at reactors or reprocessing sites, and present-worth levelizing of costs.

Several nonlinearities had to be considered which MPS could not handle. Therefore the nonlinear cost coefficients were linearized, and MPS was used to solve an approximation to the model. Using the results of this solution, better cost coefficients were calculated, and another MPS run was made. Rather than calculate coefficients by hand we decided to combine MPS and the FORTRAN program NEWCOST. NEWCOST calculates coefficients from the results of an MPS run and returns to MPS for another run. The more stable the cost coefficients become, the closer we are to the nonlinear solution of our model.

Initially we did not set upper limits on the size of the reprocessing plants. The optimum was reached with a few very large plants. The total cost obtained from the first solution was $541 million with nine

plants to be built throughout the 70 years. We ran the problem with upper limits on the plant sizes. This resulted in a total cost of $589 million with 45 plants to be built.

When linear programming is used to obtain a solution to a problem containing nonlinearities, the solution may be a local optimum rather than the global optimum. The objective function and constraints describe a response surface containing hills and valleys. A linear programming solution is found by starting at some point on the response surface and continuously moving "uphill" until a peak is reached. The linear program is unable to find some higher peak which may be nearby.

In February we received a new version of MPS which contained a separable programming algorithm to handle some nonlinearities. In the algorithm the nonlinear functions are approximated by a series of straight line segments. Separable programming does not necessarily give us the global optimum but it usually produces better results than does the linear programming algorithm. The total cost obtained from separable programming was $580 million with 46 plants to be built. The total computer time to reach the solution was significantly less with the new version of MPS than with the iterative process described previously.

The results of this study were used to form the following conclusions with respect to long-term economy in the reprocessing of spent nuclear fuel:

1. Spent fuels generated prior to about 1972 should be placed in inventory.
2. The first reprocessing plant should be constructed about 1972 in the Northeast.
3. The plant should have a capacity of some 900 metric tons per year, which is about 50% oversized with respect to 1972 demand.
4. A second and larger plant having a capacity of up to 1500 metric tons should be constructed in the East Central or Southeast United States in 1976 or 1977.
5. The policy should be one of deferring new plant construction until stockpiling costs of spent fuels become equal to capital savings available from increasing plant sizes. The effect of shipping costs should also be considered.
6. The savings resulting from building very large plants are very significant.

The accuracy of the forecasted items in our model should be fairly good during the early years of the planning period.

Simplified Reactor Systems Linear Program for ORNL

A project has been started to study reactor economics in order that a least-cost pattern may be found for the advent of various reactors and operating modes over the next 70 years. A linear programming model is being developed similar to the one developed at Pacific Northwest Laboratories.[15] The mathematical programming system[16] for the IBM 360/75 will be used to solve the problem.

The linear programming model is designed to select the optimal schedule from among the various reactors and modes which will have become technically feasible during the course of the planning period. The various activities in the model include such things as the installation and operation of reactors of various types and operating modes, the mining of U_3O_8, and the stockpiling of fissile and fertile materials.

---

[15] *The U.S. Electrical Power Economy – a Linear Programming Model,* W. I. Neef, Economic Systems Analysis Unit, Mathematics Department, Pacific Northwest Laboratories, Richland, Wash. (May 1967).

[16] *Mathematical Programming System/360 Linear and Separable Programming User's Manual,* IBM Application Program, H20-0476, December 1967.

The model will allow for a 70-year planning horizon with time periods of two years each. The reactors will be operated at either a high capacity level or a low capacity level. The objective function to be minimized will be the cost of all activities in the solution in billions of dollars.

Initially we will have about 35,000 activities (or columns) and 1000 constraints (or rows) in the matrix.

High-Temperature Gas-Cooled Reactor Fuel Reprocessing

As part of the endeavor to determine the best head end reprocessing flowsheet for advanced HTGR fuels, small reprocessing experiments are being conducted to test the feasibility of a given reprocessing procedure for a given fuel composition. We wrote a computer program to calculate the average composition of Dragon fuel compacts starting with the basic fuel particle composition and coating parameters. Work continues by us and the experimenter on a second computer program that will calculate the recoveries and losses at various stages in the reprocessing system. These programs will greatly facilitate the evaluation of the reprocessing system by the experimenter.

Spent Fuel Shipping Casks

We modified a program which describes the thermal behavior of spent fuel shipping casks during a fire so that the ambient temperature can be dropped in order to simulate an environmental cooldown. We also expanded the code to handle casks made from materials other than lead.

Preliminary Analysis of Light Absorption Spectra

We completed a program[17] for the preliminary analysis of light absorption spectra.

CIVIL DEFENSE

D. E. Arnurius
E. S. Cutler
W. L. McMullen, Jr.

We have continued the work reported last year[18] in support of the ORNL Civil Defense Project study of national security attitudes of the American public. We have edited the results of some 120 opinion surveys so that they can be used for computer analysis. We wrote a PL/1 program which can be used to edit any survey question in the computerized survey data base.

GENERAL ENGINEERING AND CONSTRUCTION

N. B. Gove
Ruth B. Hofstra

Capsule Design Program

We wrote a FORTRAN program TRAP-1 to calculate the temperature distribution in a concentric array of materials of different types exposed to a flux of neutrons and gamma rays. Heat is generated in the

---

[17] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 15; *Math. Div. Ann. Progr. Rept. Dec. 31, 1966,* ORNL-4083, p. 29.

[18] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 11.

materials by absorption of gamma rays and, in some cases, by fission. Heat is removed by water flowing through one of the regions in the concentric array. In accordance with experiments performed by the Design Department, some regions contain gas mixtures. An iterative procedure is used to account for thermal expansion.

HEALTH PHYSICS

D. E. Arnurius
A. M. Craig, Jr.
Sandra H. Merriman
C. L. Thompson
G. G. Warner

Dispersion of Finite Neutron Beams Within the Cylindrical Model of Standard Man

We rewrote the Monte Carlo neutron transport code for estimating dose in cylindrical phantoms[19] to allow the calculation of dose from the dispersion of neutrons from finite beams incident at the center of the side of the cylinder as a function of the distance from the beam.

We made corresponding changes to the gamma-ray transport code which analyzes the gamma rays generated by the neutron problem.

Estimation of Plutonium Body Burden

We designed a computer code to use the entire record of an employee's urinary excretion of plutonium to produce an estimate of his intakes as well as of the plutonium present in his body at any time during or after exposure.[20] We used ORDEAL[21] to fit the logarithms of the excretion data by the least-squares method because we believe that the data fit a power function of the form

$$F(a, \alpha, t) = at^{-\alpha} .$$

Studies have been made and are continuing to enable one to determine, in the light of daily fluctuation in the data, whether a given high fluctuation is the result of a patient's excretion variability or is actually the result of a new intake.

Dose from Gamma-Ray Sources

The computer code used to estimate internal dose in a man phantom from gamma-ray exposure situations has been reported previously.[22] In previous results reported from the use of this phantom, the phantom was taken as homogeneous. We modified the code in such a way as to take into account differences in density and composition in human internality.[23] We then used the code to provide estimates of dose from sources of photons uniformly distributed in the masses of the lungs, the skeleton, four regions

---

[19] *Math. Div. Ann. Progr. Rept. Dec. 31, 1966,* ORNL-4083, p. 24.

[20] *Health Phys. Div. Ann. Progr. Rept. July 31, 1967,* ORNL-4168, p. 273.

[21] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, pp. 12–13; see "General Purpose System," this report.

[22] H. L. Fisher, Jr., and W. S. Snyder, *Health Phys. Div. Ann. Progr. Rept. July 31, 1966,* ORNL-4007, p. 221.

[23] G. G. Warner and A. M. Craig, Jr., *ALGAM, a Computer Program for Estimating Internal Dose from Gamma-Ray Sources in a Man Phantom,* ORNL-TM-2250 (1968).

of the gastrointestinal tract, the bladder, the liver, the spleen, and the kidneys. We used results from these runs to investigate the validity of the reciprocity theorem.[24]

We also used the code to provide dose estimates in the man phantom subjected to variously located external point sources of ^{60}Co and ^{137}Cs according to specifications which simulated three arrangements of the whole-body irradiation facilities at ORAU Medical Division.

Because of the smallness of the ovaries and their remoteness from the sources of photons located in the upper trunk, we have had difficulty obtaining statistically reliable estimates of dose due to the lack of sufficient number of actual interactions in them. We amended the code to provide another kind of statistical estimate of dose in the right ovary based on the hypothetical situation in which every photon has at least one interaction there. We obtained this estimate by

$$\sum_{1}^{n} \frac{\varphi \times W \times \sigma_T \times e^{-\sigma_T R} \times \Delta E \times C}{R^2} ,$$

where n is the number of interactions, φ is the probability of Compton scattering in the angle necessary to achieve the hypothetical collision of the photon with the right ovary, W is the photon weight, σ_T is the total macroscopic cross section for the scattering medium, ΔE is the energy to be "absorbed," C is a constant appropriate to convert energy to dose, and R is the distance from the location of the preceding interaction to the center of the right ovary. For simplicity's sake at this time, we assumed that there is a single medium between the right ovary and the preceding interaction site.

Water Pollution

A part of the Laboratory's research in the area of water pollution is being conducted on the ORIC pond. The pond is supplied from two input sources: a spring located in its northeast corner and precipitation runoff from 9.5 acres of grassed watershed which drains into the pond. Periodically, flow data and chemical data are collected at these two input sources and the one output source.

We wrote a computer program which determines and plots material balances for nutrients entering and leaving the pond. These data are to be applied to a mathematical model from which the concentration of nutrients can be estimated for some future time.

INSTRUMENTATION AND CONTROLS

A. M. Craig, Jr.
Judy C. Johns

Radioactive Isotope Sand Tracer Program

Sand movements in a littoral system are being studied by making radiation intensity measurements in an area before and after injection of sand tagged with ^{133}Xe. The measurements are made on the beach, in the surf, and offshore to a water depth of 50 ft with a mobile amphibious detection system.[25] We have

[24] W. S. Snyder and G. G. Warner, "Dose to Body Organs from a Source of Photons in One of Them: Validity of the Reciprocity Theorem," Proceedings of the First European Congress on Radioprotection, Menton, France, Oct. 9–11, 1968 (to be published).

[25] E. E. Fowler, *Isotopes Radiation Technol.* 5, 171 (1968).

written a program for the IBM 360/75 to process the data collected in these experiments. This program calculates and plots the vehicle track and detector track, the normal radiation background, and the radiation intensity after injection of the tagged sand, corrected for background and decay.

FFTF Thermocouple Drift Tests

We wrote a program for the IBM 360/75 to prepare graphical records, using the CRT plotter, of the behavior of thermocouples tested in the Fast Flux Test Facilities thermocouple experiments. A sample of the output is shown in Fig. 8, where the deviation of the temperature indicated by a test thermocouple from the temperature indicated by a reference thermocouple is plotted as a function of time.

Information collected in this project will be used to prepare purchase specifications for thermocouples to be used in the Fast Flux Test Facility.

Data Compression and Reformatting Program

We wrote a program for the IBM 360/75 to compress large amounts of data stored on magnetic tape by a high-speed data-logging device into a form acceptable to an existing program for the CDC 1604-A.

TRU Product Tank Temperature Anomalies

Long-term and short-term temperature anomalies have been observed in several of the Transuranium Facility product tanks. Since these temperatures were logged on magnetic tape during the periods of interest, we wrote a program for the IBM 360/75 to retrieve the data from the tape and to prepare detailed graphs of temperatures vs time.

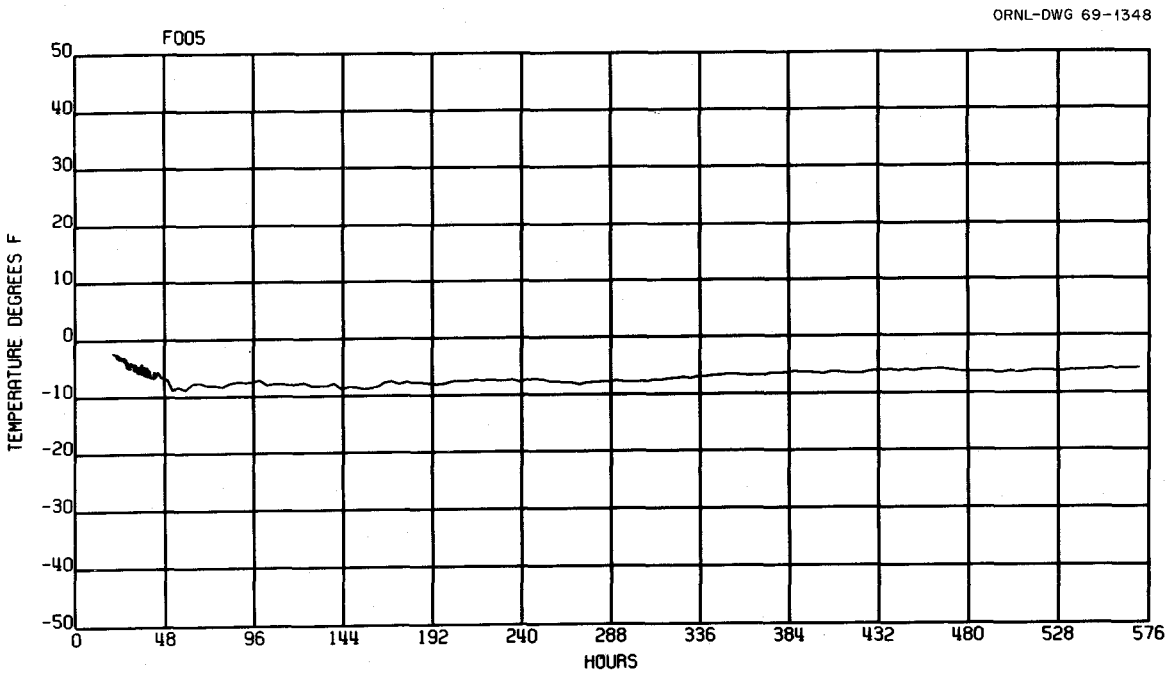


Fig. 8. CRT Plot of Temperature vs Time.

Inventory and Maintenance Records

The monitoring systems development group of the Instrumentation and Controls Division services several hundred health physics instruments. To aid in the control of inventory and maintenance costs of these instruments, we wrote a program for the IBM 360/75 to prepare periodic reports on maintenance costs and to flag those instruments requiring frequent service or showing abnormally high maintenance costs.

Material Usage Report

We wrote a program for the IBM 360/75 to prepare monthly reports of the materials obtained from stores by each individual in the Instrumentation and Controls Division, with the totals for each group and the total for the Division.

METALS AND CERAMICS

D. E. Arnurius
N. B. Gove
Ruth B. Hofstra
Sandra H. Merriman

Alloys

We prepared several programs for data analysis and file management of alloy test data. Types of information include chemical analysis, temperature, flux and stress history, and creep-rupture test results. The reduced area and true strain are computed. A variety of sorting and listing options are available.

Collimation Correction in Small-Angle X-Ray Scattering

The x-ray intensity $I(\theta)$ scattered at small angles θ is used to determine the size and shape of very small particles in biological, chemical, and metallurgical applications. In one small-angle experiment being conducted at present, intensity requirements force the use of a beam collimation system which seriously distorts the desired intensity $I(\theta)$ into the observed intensity $J(\theta)$. It has been shown[26] that these two intensities are related by the following integral equations:

$$J(\theta) = \int_{-\infty}^{\infty} W_w(u)\, F(\theta + u)\, du \ ,$$

$$F(\theta) = \int_{-\infty}^{\infty} W_l(u)\, I(\sqrt{\theta^2 + u^2})\, du \ ,$$

where $W_l(u)$ and $W_w(u)$ are the slit-length and the slit-width weighting functions determined by the geometry of the collimation system.

We wrote a computer program to find the corrected intensities using the method described in "Numerical Solution of Integral Equations," this report, for solving integral equations.[27]

---

[26] R. W. Hendricks and Paul W. Schmidt, "Calculation of Weighting Functions for Collimation Corrections in Small-Angle X-Ray Scattering," *Acta Phys. Austriaca,* XXVI/2-3 (1967).

[27] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 4.

MSRE

D. E. Arnurius
E. J. Lee
W. L. McMullen, Jr.
C. W. Nestor, Jr.

An Improved MSRE Data Processing System

We have redesigned the MSRE data processing system for the CDC 1604-A and have implemented the newly designed system on the IBM 360/75.[28] The new system takes advantage of several hardware and software features of the IBM 360/75 system, employs some better programming techniques, and, most important, eliminates most known deficiencies in the system for the CDC 1604-A.

The new system is made up of two main stages. The first stage consists of a comprehensive editing program. The editor deletes physical errors and corrects logical errors in the raw data. Moreover, the processing of raw data is not terminated whenever a physical anomaly occurs in a data record as is the case in the old system. The edited data may be recorded in two formats. Data recorded in one of the formats may be read by the MSRE's monitoring computer and by the CDC 1604-A data processing system. Data recorded in the other format may be read easily using standard FORTRAN procedures.

The second stage performs the desired analyses of data generated in the first stage. Basically, this stage consists of a supervisor program, data management subprograms, and a collection of analysis routines written as modules. The supervisor program accepts requests for particular data management tasks and analyses of data generated in the first stage. It then fetches the data requested for analysis, makes these data available in the format (the format is specified at execution time) required by the analysis routine, and performs the analysis. The modular construction of the analysis routines and the specification of data format at execution time permit the insertion or modification of an analysis module with minimal programming effort.

Calculation of Fission Product Concentrations in the MSRE

During the past year we completed fission product concentration calculations[29] for the ^{235}U phase of the MSRE experiment, and we are currently preparing data for similar calculations for the ^{233}U phase of operations.

---

[28] *Math. Div. Ann. Progr. Rept. Dec. 31, 1965,* ORNL-3919, pp. 41–42.

[29] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 15.

NEUTRON PHYSICS

A. M. Craig, Jr.
Arline H. Culkowski
Margaret B. Emmett
Manuel Feliciano, Jr.
N. B. Gove
J. R. Stockton
C. L. Thompson
G. G. Warner

Much of the work done for this division involved the use or modification of existing Monte Carlo programs.[30–32]

O5R – Albedo Version

We have completed a FORTRAN IV version of O5R to handle albedo problems. This version allows a medium in the defined system to be treated as a reflecting medium, each neutron interaction with a boundary of this medium being treated as a reflection at the incident point in a manner defined by the user. Parameters of the reflected neutron are determined using input data defining the reflecting properties of the medium.

This version of O5R has been tested for problems allowing only albedo reflections by reproducing the results reported by Maerker and Muckenthaler[33] for concrete ducts using their doubly differential dose albedo data for concrete. We have confirmed the validity of the code for standard O5R neutron transport test problems, and are currently testing the code for problems where both types of neutron interactions are allowed.

Analysis of Neutron Transport Problems

We have completed and tested the ORNL FORTRAN version of the analysis package for O5R problems, ACTIFK, including options for handling albedo problems. We have converted this version to FORTRAN IV (Level H) for the IBM 360/75.

O5R/ACTIFK

We have combined the analysis package for O5R neutron transport problems, ACTIFK, with the FORTRAN IV (Level H) albedo version of O5R. The combined code allows in-core analysis of an O5R neutron problem during the neutron transport. All options available to users of O5R and ACTIFK in sequence are available in the combined code.

We are presently converting the standard supporting subroutines described in the ACTIFK manual, and testing them for as many options as is practical.

---

[30] D. C. Irving *et al.*, *O5R, a General Purpose Monte Carlo Neutron Transport Code,* ORNL-3622 (February 1965).

[31] S. K. Penny *et al.*, *OGRE, a Monte Carlo System for Gamma-Ray Transport Studies,* ORNL-3805 (April 1966).

[32] F. B. K. Kam and K. D. Franz, *ACTIFK, a General Analysis Code for O5R,* ORNL-3856 (September 1966).

[33] R. E. Maerker and F. J. Muckenthaler, *Monte Carlo Calculations Using the Albedo Concept,* ORNL-TM-1557 (Sept. 21, 1966).

SNAP Collimator Design

We have made amendents to the geometry and analysis code for the O5R program used for SNAP core mapping calculations.

Time-Dependent Neutron and Secondary Gamma-Ray Transport in Air over Ground

We made extensive production calculations of the time, energy, angle, and spatial distribution of neutrons in an air over ground geometry using the modified O5R Monte Carlo program previously reported.[34] We completed the analysis of these results by producing Calcomp plots of the detailed results. We found it necessary to modify as well as write additional subprograms for the plotting package INTRIGUE[35] (CDC 1604-A version) in order to allow the flexibility required in displaying the results.

In order to reduce the computation time for tracking and analyzing neutron histories generated by Monte Carlo programs, we combined the above-mentioned version of O5R and the analysis program ACTIFK in ORNL FORTRAN on the IBM 360/75 so that analysis was performed during the tracking of the history. In addition to this, we added time-dependent heating rates in the ground for the air over ground problem. We added gamma generation routines so that the complete time-dependent secondary gamma source from neutrons of all energies could be determined.[36] We made changes in the logic of O5R to increase the efficiency of tracking low-energy neutrons.

We modified the four-dimensional analysis routines which are used in ACTIFK for use with a next flight estimator; thus both the boundary crossing and the next flight estimators may be used in determining a four-dimensional flux. We are using these routines in analyzing time-dependent experiments performed both at ORNL and elsewhere.

Analysis of TSF Slab Experiments

We have written a program for the CDC 1604 to generate the source term for the analysis of slab experiments done at the Tower Shielding Facility (TSF). The program either calculates neutron source parameters directly from the input data or gets them from an O5R collision tape which was generated during a previous calculation. The calculations include biasing the source in different ways for the different thicknesses of the slab. We write the source parameters on a tape in batches, and then we read and analyze this tape after the specified number of source neutrons has been generated.

Air Transport

We have written several codes for the IBM 360/75 to edit the results from the two-dimensional discrete ordinates code DOT[37] for radiation transport, and from the 4-D analysis of neutron and gamma-ray transport problems in the air-ground geometry by O5R-ACTIFK and OGRE-4D.

Basic data from the problems can be presented in several user-directed formats. Data from several runs with nonoverlapping source energy spectra may be combined to estimate results from a continuous source

---

[34] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 17.

[35] D. K. Trubey and M. B. Emmett, *A CDC-1604 Subroutine Package for Making Linear, Logarithmic and Semilogarithmic Graphs Using the Calcomp Plotter,* ORNL-3447 (June 1963).

[36] *Neutron Phys. Div. Ann. Progr. Rept. May 31, 1968,* ORNL-4280, pp. 78–80.

[37] F. R. Mynatt, *A User's Manual for DOT – a Two-Dimensional Discrete Ordinates Code with Anisotropic Scattering,* K-1694 (to be published).

energy spectrum. Data from several independent solutions of the same problem may be combined to estimate the results from a problem of the combined sample size.

OGRE-4D

We have developed a flexibly dimensioned, time-dependent, batch-process OGRE-system Monte Carlo gamma-ray transport code to analyze secondary gamma rays generated during an O5R point isotropic source neutron transport problem in the air-ground geometry.[38]

OGRE-4D combines time dependent random walk procedures from OGRE with the 4-D analysis methods of the 4-D ACTIFK analysis package for O5R problems. Source parameters are read from a magnetic tape generated by O5R, a specific two-media infinite slab geometry replaces the general geometry routines of OGRE, and photons are processed in batches so that statistics may be based on batches of photons. OGRE-4D calculates the time-, angle-, energy-, and space-dependent flux at the air-ground interface, determines the total and time-dependent dose rate at the interface, and, as an option, determines the total and time-dependent ground heating as a function of depth in the ground.

The total and time-dependent dose rate is also determined as a function of the type of gamma, as determined in the O5R problem, so that fractional contributions of several gamma source types can be calculated.

Program DUCAL

We converted DUCAL,[39] which computes energy spectra at both discrete and continuous energy levels for given spin and parity of gamma rays evolved through neutron capture in a given nuclide, from the FORTRAN 63 language on the CDC 1604-A computer to FORTRAN IV (Level H), Opt = O, language on the IBM 360/75 computer.

Intranuclear Cascades

Medium-Energy Cascades. – We have continued work on this project.[40] We computed from new absorption parameters and phase shift data values for differential scattering cross sections for pion-nucleon reactions up to 2500 MeV. We also computed and plotted Coulomb scattering, interference, nuclear scattering, and total scattering. These computations indicate that nuclear scatterings predominate. We used this and other information to update the cross-section tape.

We refined the integration and interpolation subroutines. We added tests to reject cascades which do not preserve nuclear mass and charge. This updated version of MECC is called MECC4.

We converted the isobar model subprograms for use in the IBM 360/75 computer.

High-Energy Cascades. – We have designed a program to simulate the effect of high-energy bombardment of nuclei. The distribution of collision sites within the nucleus is expressed in terms of an absorption coefficient. The absorption coefficient is expressed as an integral of the cross section over the nucleon momentum distribution. Two approaches to the estimation of this integral are planned, an analytic approach for cases where the cross section and the integrand can be expressed as truncated power series, and use of numerical quadrature formulas elsewhere.

---

[38] *Neutron Phys. Div. Ann. Progr. Rept. May 31, 1968,* ORNL-4280, p. 78.

[39] K. J. Yost, "A Method for the Calculation of Neutron Capture Gamma-Ray Spectra," *Nucl. Sci. Eng.* 32, 62 (1968).

[40] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 19.

ORNL-DWG 69-1349

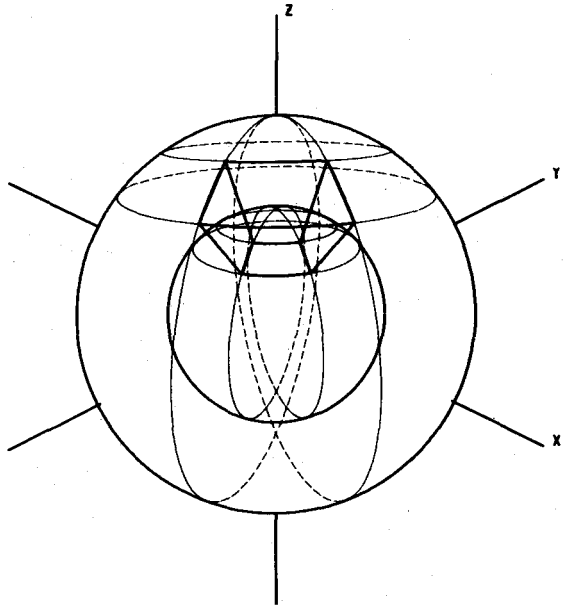


Fig. 9. Nuclear Regions in $r\theta\phi$ Coordinate System.

Nuclear matter is represented by a density function. As the cascade develops, some nucleons become high-energy projectiles. This depletes the nuclear matter, and so a simple density depletion subroutine is planned, for which the nuclear volume is assumed to be divided into regions. After a region has been depleted, nuclear matter may "flow" into the region, and so a crude model for propagation of density depletions is proposed.

The nuclear regions are defined in an (r, θ, φ) system as shown in Fig. 9. The number of shells and regions per shell are controlled by the user. We developed an algorithm for determining quickly which pairs of regions are neighbors.

NUCLEAR DATA PROJECT

Judith K. Brinkley N. B. Gove
Sara E. Webb

Nuclear Systematics

We wrote a program to draw graphs of neutron separation energies vs mass number (Fig. 10) and of proton separation energies vs neutron number.[42,43] The graphs confirm an earlier hypothesis[44] that the separation energy of the *n*th proton increases monotonically (except for odd-even effects) as the neutron number is increased, showing little, if any, discontinuity at magic numbers. An analogous statement can be made for the separation energy of the *n*th neutron as the proton number is increased. Extrapolation by means of the graphs should, therefore, be particularly reliable since it is essentially magic-number independent. We prepared a table of 24 anomalous cases in the systematics.

Atomic Mass Adjustment

In collaboration with A. H. Wapstra of the Instituut voor Kernphysisch Onderzoek, Amsterdam, a new mass adjustment project is planned. This will update the 1964 adjustment,[45] but the increase in input data and the improvements needed have led to a design for a new computer program.

A survey has been made of experimental results from which atomic masses can be calculated. Relevant types of data include mass spectrometer measurements, reaction thresholds, and radioactive decay energies. An experimental result of one of these types can be regarded as a linear relationship among two or more atomic masses. About 4500 such relationships have been found so far. The number of different nuclear species involved is about 1500. Thus some 4500 or more linear algebraic equations in 1500 unknowns are expected.

The collection of data is complete and is being checked. We wrote several programs to assist in the cross-checking and bookkeeping procedures and are writing a program to make a least-squares fit to the system of overdetermined equations.

Text Capitalization Algorithms

The Recent References file on the Nuclear Data Project was prepared for publication using an IBM 1401 equipped with a 120-character printing chain (courtesy of Computing Technology Center, Union Carbide, Oak Ridge). To save input preparation cost, the file was keypunched in the usual all-caps fashion, and an IBM 360 program was written to decide what should be upper case, lower case, and superscript. The first algorithm consists in making the first letter of every word a capital letter. There are many exceptions to this rule as well as some difficulty in defining a word. Cases of capitalization inside a word (O'Hara, d'Arcy, MeV, ORNL, XIV) occur. In titles most words are capitalized, but prepositions and articles are not. Chemical element symbols are recognized (usually), and atomic mass numbers are made superscript. The overall error rate is currently about 1 per 20 lines of text. An edit feature is included to permit correction of these errors. A sample of the output is shown in Fig. 11.

---

[42] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 19.
[43] N. B. Gove and M. Yamada, *Nucl. Data* A4(3), 237–63 (1968).
[44] M. Yamada and Z. Matumoto, *J. Phys. Soc. Japan* **16**, 1497 (1967).
[45] J. H. E. Mattauch, W. Thiele, and A. H. Wapstra, *Nucl. Phys.* **67**, 73 (1965).

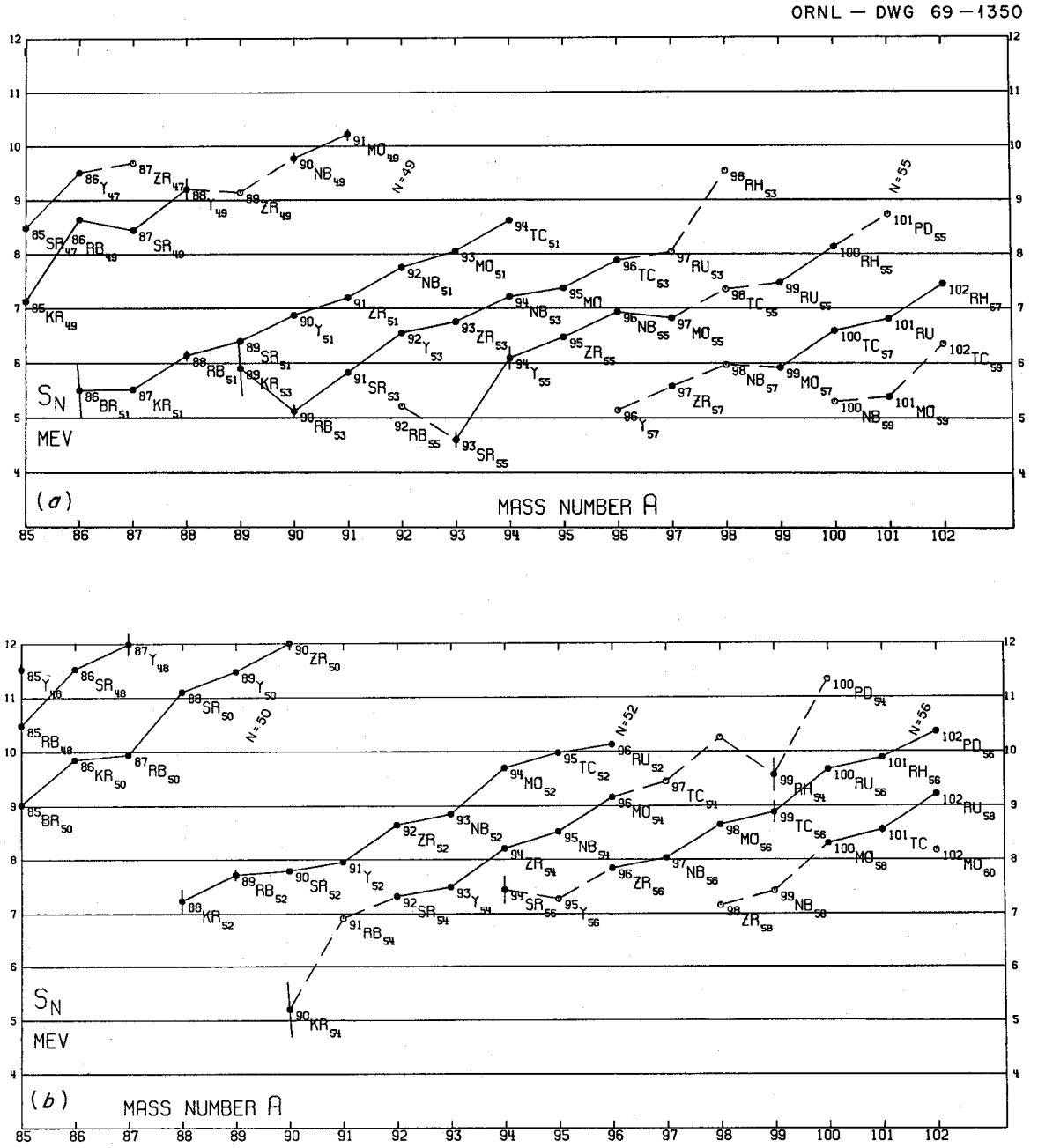


Fig. 10. Neutron Separation Energies. (*a*) For odd-*N* nuclei, A = 85 to 102; (*b*) for even-*N* nuclei, A = 85 to 102.

ORNL-DWG 69-1351

```
68Ya04  T.Yamazaki, G.T.Ewan, S.G.Prussin - Phys.Rev.Letters 20,
          1376(1968)
        Level and Isomer Systematics in Even Sn Isotopes
68Ya05  K.Yagi, Y.Saji, T.Ishimatsu, Y.Ishizaki, M.Matoba,
          Y.Nakajima, C.Y.Huang - J.Phys.Soc.Japan 24, 1167(1968)
        Level Structure of 118Sn from a 119Sn(p,d) Reaction
68Ya06  G.C.Yang, J.B.Sampson - Bull.Am.Phys.Soc. 13, No.3, 435,
          DH8 (1968)
        Electron Spin Resonance of Ru in TiO2
68Ye01  A.Yen, F.Ajzenberg-Selove, B.Rosner - Nucl.Phys. A111,
          100(1968)
        Excited States of 93Mo
68Yo01  H.J.Young, J.Rapaport - Phys.Letters 26B, 143(1968)
        The 58Ni(3He,p)60Cu Reaction
68Yo02  H.J.Young, J.Rapaport - Bull.Am.Phys.Soc. 13, No.1, 105,
          GG10(1968)
        Ni62(He3,p)Cu64 and Ni64(He3,p)Cu65 Reactions
68Yo04  D.H.Youngblood, B.H.Wildenthal, C.M.Class - Phys.Rev. 169,
          859(1968)
```

Fig. 11. Sample Output from Text Capitalization Program.

ORELA

Nancy A. Betz
J. G. Sullivan, Jr.
D. R. Winkler

Because Systems Engineering Laboratories (SEL) could not meet scheduled delivery of the phase I data acquisition computer system,[46] an interim computer system was delivered in June 1968. The phase I system is now scheduled for delivery early in 1969.

We encountered no problems in transferring the Executive Monitor System[47] to the interim computer. We developed disk file, maintenance, and debugging routines, and modified some SEL routines, such as the assembler and loader, to run under the executive monitor system.

Although DAC-1, Data Acquisition Computer One, differs from DAC-2 in interrupt assignments, memory size, and other ways, the executive monitor system is being written to make these differences transparent to user programs in order to facilitate the use of either computer.

Acceptance tests for the linear accelerator are scheduled to be completed in March 1969. We hope that during these tests we can run experiments that will use the data acquisition computers.

The bids on phase II, the Intermediate Analysis and Coordinating Computer, have been received and are being evaluated.

---

[46] *Neutron Phys. Div. Ann. Progr. Rept. May 31, 1968,* ORNL-4280, p. 138.

[47] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 15.

ORIC

T. D. Calton
N. B. Gove

Data Transmission

We have modified the Houston Automatic Spooling Program (HASP)[48] to allow data communication between the IBM 360/75 and the ORIC terminal.[49] The HASP program was written with the option of generating code for Remote Job Entry (RJE) support. The HASP program was generated with the RJE coding and modified to handle the two types of data transmission from the ORIC terminal.

The types (job or data for storage in the IBM 360/75) are indicated by an identification record. If the identification record indicates that a job follows, the identification record is replaced by a job card required by the IBM 360/75 operating system. The job is then placed in the job queue to await execution. If the identification record indicates that data follow, then the data are preceded by a group of control cards necessary to define a job that transfers the data to IBM 360/75 storage assigned to ORIC.

The identification record contains either a job name or data file name. The data file name is used to label the data file created for ORIC. The job name is used to construct the job card.

We made additional modifications to the RJE coding to preface output to the ORIC terminal with an identification card which indicates the type of output (print or punch) and the number of records to be transmitted.

The ORIC terminal has been operational since May 1968 and was released to ORIC users July 1968. During the period July to October 1968, approximately 30 jobs were transmitted to the IBM 360/75 from the ORIC terminal with satisfactory results.

PHYSICS

H. J. Hargis C. W. Nestor, Jr.
T. C. Tucker

Relativistic Hartree Calculations

We have converted the series of programs[50] for the relativistic Hartree calculation of the electronic structures of atoms and ions for use in the IBM 360/75. Aside from a considerable increase in computing speed (roughly a factor of 8), the chief benefit to the users of the programs is that an IBM 2314 disk pack is now used for the storage of both the programs and the wave function library. The user need not submit a large program deck or large amounts of input data, and he also has random access to the data in the library rather than the serial access provided by magnetic tape. To date, operating experience with the 2314 has been entirely satisfactory. A recent list of the library contents is shown in Table 1.

---

[48] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 32.
[49] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 16.
[50] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, pp. 21–23.

Table 1.

THE CONTENTS OF THE RELATIVISTIC WAVE FUNCTION LIBRARY ON THE IBM/360/2314 ON 02-27-69 WERE

```
SER.NO.    94 Z=  7.NITROGEN POINT NUCLEUS FREE                    NBG            FREE  0.0    FM 0.0    FM 1.50EX 01-27-69
SER.NO.    21 Z=  7. NITROGEN 13   NEUTRAL                         NBG            FREE  2.2403FM 0.5666FM 1.50EX 10-08-68

SER.NO.    47 Z= 10. NEON                                                        FREE  2.7016FM 0.5666FM 1.50EX 12-13-68
SER.NO.    97 Z= 10.NEON 21 NEUTRAL                             NBGRWS= 4.00000BOHR  2.7572FM 0.5666FM 1.50EX 01-20-68

SER.NO.   103 Z= 11.SODIUM 22 POINT NUCLEUS FREE                 NBG            FREE  0.0    FM 0.0    FM 1.50EX 01-27-69
SER.NO.    93 Z= 11.SODIUM 22 NEUTRAL                            NBGRWS= 4.00000BOHR  2.8109FM 0.5666FM 1.50EX 01-14-69
SER.NO.    85 Z= 11.SODIUM 22 NEUTRAL                            NBGRWS= 3.00000BOHR  2.8109FM 0.5666FM 1.50EX 01-13-69

SER.NO.   123 Z= 13.ALUMINUM 27 NEUTRAL                          NBGRWS= 4.00000BOHR  3.0553FM 0.5666FM 1.50EX 01-17-69

SER.NO.    82 Z= 14.SILICON 29 NEUTRAL                           NBGRWS= 4.00000BOHR  3.1439FM 0.5666FM 1.50EX 01-19-69

SER.NO.    70 Z= 15. PHOSPHORUS 31, W.S. R=4.0                   NBGRWS= 4.00000BOHR  3.2282FM 0.5666FM 1.50EX 12-20-68

SER.NO.     5 Z= 17.CHLORINE                                                     FREE  3.3857FM 0.5666FM 1.00EX 08-20-68
SER.NO.     8 Z= 17.CL -                                            RWS= 4.00000BOHR  3.3857FM 0.5666FM 1.00EX 08-27-68

SER.NO.     2 Z= 18.ARGON                                                        FREE  3.5655FM 0.5666FM 1.00EX 09-18-68
SER.NO.    87 Z= 18.ARGON 38 NEUTRAL                            NBGRWS= 4.00000BOHR  3.4956FM 0.5666FM 1.50EX 01-13-69
SER.NO.    77 Z= 18.ARGON 38 NEUTRAL                            NBGRWS= 3.00000BOHR  3.4956FM 0.5666FM 1.50EX 01-12-69

SER.NO.   115 Z= 19. POTASSIUM 40 NEUTRAL                        NBGRWS= 4.00000BOHR  3.5655FM 0.5666FM 1.50EX 01-16-69
SER.NO.    78 Z= 19. POTASSIUM 40 NEUTRAL                        NBGRWS= 3.00000BOHR  3.5655FM 0.5666FM 1.50EX 01-12-69

SER.NO.    86 Z= 20. CALCIUM 40                                  NBGRWS= 4.00000BOHR  3.5655FM 0.5666FM 1.50EX 01-13-69
SER.NO.    54 Z= 20. CALCIUM 40                                  NBGRWS= 3.00000BOHR  3.5655FM 0.5666FM 1.40EX 12-17-68

SER.NO.   107 Z= 21.SCANDIUM 45 NEUTRAL                          NBGRWS= 4.00000BOHR  3.7298FM 0.5666FM 1.50EX 01-20-69

SER.NO.   106 Z= 22.TITANIUM 48 NEUTRAL                          NBGRWS= 4.00000BOHR  3.8223FM 0.5666FM 1.50EX 01-20-69

SER.NO.    96 Z= 23.VANADIUM 51 NEUTRAL                          NBGRWS= 4.00000BOHR  3.9108FM 0.5666FM 1.50EX 01-20-68

SER.NO.    99 Z= 24.CHROMIUM 52 NEUTRAL                          NBGRWS= 4.00000BOHR  3.9395FM 0.5666FM 1.50EX 01-20-68

SER.NO.    88 Z= 25.MANGANESE 55 NEUTRAL                         NBGRWS= 4.00000BOHR  4.0233FM 0.5666FM 1.50EX 01-19-69

SER.NO.    23 Z= 26. IRON          (CWN)                                        FREE  0.0    FM 0.0    FM 1.50EX 02-05-69
SER.NO.    33 Z= 26. IRON          (CWN)                                        FREE  0.0    FM 0.0    FM 1.50EX 02-06-69
SER.NO.   124 Z= 26.IRON 57 NEUTRAL                              NBGRWS= 4.00000BOHR  4.0774FM 0.5666FM 1.50EX 01-17-69

SER.NO.   112 Z= 27.COBALT 60 NEUTRAL                            NBGRWS= 4.00000BOHR  4.1560FM 0.5666FM 1.50EX 01-17-69

SER.NO.   120 Z= 28.NICKEL 60 NEUTRAL                            NBGRWS= 4.00000BOHR  4.1560FM 0.5666FM 1.50EX 01-17-69
SER.NO.    75 Z= 28.NICKEL 60 NEUTRAL                            NBGRWS= 3.00000BOHR  4.1560FM 0.5666FM 1.50EX 01-12-69

SER.NO.     6 Z= 29.COPPER                                                       FREE  0.0    FM 0.0    FM 0.0 EX 08-22-68
SER.NO.    62 Z= 29.COPPER  WS ATOM                              NBGRWS= 4.00000BOHR  4.2567FM 0.5666FM 1.40EX 12-19-68
SER.NO.   118 Z= 29.COPPER NEUTRAL                               NBGRWS= 4.00000BOHR  4.2567FM 0.5666FM 1.50EX 01-17-69

SER.NO.   121 Z= 30.ZINC 67 NEUTRAL                              NBGRWS= 4.00000BOHR  4.3294FM 0.5666FM 1.50EX 01-17-69

SER.NO.    34 Z= 31. GALLIUM       (CWN)                                        FREE  0.0    FM 0.0    FM 1.50EX 02-06-69

SER.NO.   122 Z= 32.GERMANIUM 71 NEUTRAL                         NBGRWS= 4.00000BOHR  4.4229FM 0.5666FM 1.50EX 01-17-69
SER.NO.     1 Z= 32.GERMANIUM     (4S)2(4P)2                         RWS= 3.30700BOHR  4.6000FM 0.5700FM 1.50EX 09-04-68

SER.NO.     3 Z= 36.KRYPTON                                                      FREE  0.0    FM 0.0    FM 1.50EX


SER.NO.    61 Z= 36.KRYPTON WS ATOM                              NBGRWS= 4.00000BOHR  4.6623FM 0.5666FM 1.40EX 12-19-68
SER.NO.    80 Z= 36.KRYPTON 82 NEUTRAL                           NBGRWS= 4.00000BOHR  4.6623FM 0.5666FM 1.50EX 01-12-69

SER.NO.    24 Z= 40. ZIRCONIUM    (CWN)                                        FREE  0.0    FM 0.0    FM 1.50EX 02-05-69
SER.NO.    39 Z= 40. ZIRCONIUM    (CWN)                                        FREE  0.0    FM 0.0    FM 1.50EX 02-06-69
SER.NO.    51 Z= 40.ZIRCONIUM 90 NEUTRAL                         NBGRWS= 4.00000BOHR  4.8226FM 0.5666FM 1.50EX 02-06-69
SER.NO.    71 Z= 40.ZIRCONIUM 90 NEUTRAL                         NBGRWS= 4.00000BOHR  4.8226FM 0.5666FM 1.50EX 01-24-69

SER.NO.     7 Z= 46.PALLADIUM                                                    FREE  0.0    FM 0.0    FM 1.50EX
SER.NO.    64 Z= 46.PALLADIUM NEUTRAL W.S.                       NBGRWS= 4.00000BOHR  5.1844FM 0.5666FM 1.40EX 12-19-68
SER.NO.    40 Z= 46.PALLADIUM NEUTRAL W.S.                       NBGRWS= 4.00000BOHR  5.1844FM 0.5666FM 1.50EX 01-09-69

SER.NO.    32 Z= 50. TIN           (CWN)                                        FREE  0.0    FM 0.0    FM 1.50EX 02-05-69
SER.NO.    30 Z= 50.TIN 120 NEUTRAL                              NBGRWS= 4.00000BOHR  5.3483FM 0.5666FM 1.50EX 01-24-69

SER.NO.    10 Z= 54.XENON                                                       FREE  0.0    FM 0.0    FM 1.50EX
SER.NO.    65 Z= 54.XENON NEUTRAL W.S.                           NBGRWS= 4.00000BOHR  5.5628FM 0.5666FM 1.40EX 12-19-68
SER.NO.    76 Z= 54.XENON NEUTRAL W.S.                           NBGRWS= 4.00000BOHR  5.5031FM 0.5666FM 1.50EX 01-12-69

SER.NO.    81 Z= 56.BARIUM 135 NEUTRAL                           NBGRWS= 4.00000BOHR  5.5775FM 0.5666FM 1.50EX 01-25-69

SER.NO.   133 Z= 57.LANTHANUM 139 NEUTRAL                        NBGRWS= 4.00000BOHR  5.6356FM 0.5666FM 1.50EX 01-22-69

SER.NO.    52 Z= 58. CERIUM (4F)(5D)(6S)2                 (CWN)                 FREE  0.0    FM 0.0    FM 1.50EX 12-17-68
SER.NO.    53 Z= 58.CERIUM   (4F)2(6S)2     (CWN)                               FREE  0.0    FM 0.0    FM 1.50EX 12-17-68
SER.NO.    27 Z= 58.CERIUM   (4F)2(6S)2     (CWN)                               FREE  0.5650FM 0.0567FM 1.50EX 12-13-68
SER.NO.    44 Z= 58.CERIUM   (4F)2(6S)2     (CWN)                               FREE  5.6499FM 0.5666FM 1.50EX 12-16-68
SER.NO.    68 Z= 58. CERIUM  (4F)(5D)(6S)2            (CWN)                     FREE  5.6499FM 0.5666FM 1.50EX 12-20-68
SER.NO.    66 Z= 58.CERIUM NEUTRAL                               NBGRWS= 4.00000BOHR  5.6499FM 0.5666FM 1.40EX 01-02-69
SER.NO.   116 Z= 58.CERIUM 140 NEUTRAL                           NBGRWS= 4.00000BOHR  5.6499FM 0.5666FM 1.50EX 01-16-69

SER.NO.   132 Z= 59.PRASEODYMIUN 141 NEUTRAL                     NBGRWS= 4.00000BOHR  5.6642FM 0.5666FM 1.50EX 01-22-69

SER.NO.   130 Z= 62.SAMARIUM 150 NEUTRAL                         NBGRWS= 4.00000BOHR  5.7897FM 0.5666FM 1.50EX 01-21-69
```

Table 1 (continued)

```
SER.NO.   119 Z= 63.EUROPIUM 152 NEUTRAL                                    NBGRWS= 4.00000BOHR  5.8165FM 0.5666FM 1.50EX 01-17-69

SER.NO.   101 Z= 64.GADOLINIUM 156 NEUTRAL                                  NBGRWS= 4.00000BOHR  5.8705FM 0.5666FM 1.40EX 01-15-69

SER.NO.    91 Z= 65.TERBIUM 159 NEUTRAL                                     NBGRWS= 4.00000BOHR  5.9102FM 0.5666FM 1.40EX 01-14-69

SER.NO.    42 Z= 66. DYSPROSIUM            (CWN)                                     FREE  0.0   FM 0.0   FM 1.50EX 12-11-68
SER.NO.    58 Z= 66. DYSPROSIUM (FINITE NUCLEUS)                   (CWN)              FREE  5.9493FM 0.5666FM 1.50EX 12-18-68
SER.NO.    84 Z= 66.DYSPROSIUM 160 NEUTRAL                                  NBGRWS= 4.00000BOHR  5.9232FM 0.5666FM 1.40EX 01-13-69

SER.NO.   104 Z= 68.ERBIUM 166 NEUTRAL                                      NBGRWS= 4.00000BOHR  6.0006FM 0.5666FM 1.50EX 01-28-69

SER.NO.    90 Z= 69.THULIUM NEUTRAL W.S.                                    NBG            FREE  6.0386FM 0.5666FM 1.40EX 12-31-68
SER.NO.    11 Z= 69.THULIUM                                                              FREE  0.0   FM 0.0   FM 1.50EX
SER.NO.    38 Z= 69.THULIUM NEUTRAL                                         NBGRWS= 4.00000BOHR  6.0386FM 0.5666FM 1.30EX 01-09-69
SER.NO.    92 Z= 69.THULIUM NEUTRAL W.S.                                    NBGRWS= 4.00000BOHR  6.0386FM 0.5666FM 1.40EX 12-31-68
SER.NO.   109 Z= 69.THULIUM 169 NEUTRAL                                     NBGRWS= 4.00000BOHR  6.0386FM 0.5666FM 1.50EX 01-16-69

SER.NO.    19 Z= 74. TUNGSTEN                   (CWN)                                 FREE  0.0   FM 0.0   FM 1.50EX 12-04-68
SER.NO.     9 Z= 74. TUNGSTEN             (CWN)                                       FREE  0.6220FM 0.0567FM 1.50EX 12-13-68
SER.NO.    43 Z= 74. TUNGSTEN             (CWN)                                       FREE  6.2219FM 0.5666FM 1.50EX 12-16-68
SER.NO.    79 Z= 74.TUNGSTEN 184 NEUTRAL                                    NBGRWS= 4.00000BOHR  6.2219FM 0.5666FM 1.50EX 01-19-69
SER.NO.   127 Z= 74.TUNGSTEN 184 NEUTRAL                                    NBGRWS= 4.00000BOHR  6.2219FM 0.5666FM 1.50EX 01-17-69

SER.NO.    12 Z= 79.GOLD                                                                FREE  0.0   FM 0.0   FM 1.50EX
SER.NO.    17 Z= 79.GOLD                                                       RWS= 3.00000BOHR  0.0   FM 0.0   FM 1.50EX 11-15-68
SER.NO.    72 Z= 79.GOLD                                                       RGPZ 3.00000BOHR  0.0   FM 0.0   FM 1.50EX 12-20-68


SER.NO.    56 Z= 81. THALLIUM                        (CWN)                             FREE  0.0   FM 0.0   FM 1.50EX 12-17-68
SER.NO.     4 Z= 81. THALLIUM                        (CWN)                             FREE  0.6470FM 0.0567FM 1.50EX 12-13-68
SER.NO.    14 Z= 81.TL +      (6S)2                                                    FREE  6.4620FM 0.5666FM 1.50EX 09-22-68
SER.NO.    15 Z= 81.TL +      (6S)1(6P1/2)1                                            FREE  6.4620FM 0.5666FM 1.50EX 09-23-68
SER.NO.    16 Z= 81.TL +      (6S)1(6P3/2)1                                            FREE  6.4620FM 0.5666FM 1.50EX 09-24-68
SER.NO.    50 Z= 81. THALLIUM                        (CWN)                             FREE  6.4620FM 0.5666FM 1.50EX 12-16-68
SER.NO.   102 Z= 81.THALLIUM 204 NEUTRAL                                    NBGRWS= 4.00000BOHR  6.4510FM 0.5666FM 1.40EX 01-15-69
SER.NO.   125 Z= 81.THALLIUM 204 NEUTRAL                                    NBGRWS= 4.00000BOHR  6.4510FM 0.5666FM 1.50EX 01-17-69

SER.NO.    26 Z= 82.LEAD,FREE ATOM,POINT NUCLEUS,XM=1.5                                 FREE  0.0   FM 0.0   FM 1.50EX 11-15-68
SER.NO.    36 Z= 82.LEAD,FREE ATOM,POINT NUCLEUS,XM=1.5                                 FREE  0.6490FM 0.0567FM 1.50EX 12-16-68
SER.NO.    48 Z= 82.LEAD (CWN)                                                           FREE  6.4949FM 0.5666FM 1.50EX 12-16-68

SER.NO.    29 Z= 88. RADIUM                 (CWN)                                        FREE  0.0   FM 0.0   FM 1.50EX 11-20-68
SER.NO.   138 Z= 88.RADIUM 226 NEUTRAL                                      NBGRWS= 4.00000BOHR  6.6859FM 0.5666FM 1.50EX 01-24-69

SER.NO.    31 Z= 92. URANIUM,FREE ATOM, POINT NUCLEUS, XM = 1.0, R(MAX) = 40            FREE  0.0   FM 0.0   FM 1.00EX 11-04-68
SER.NO.    60 Z= 92. URANIUM                        (CWN)                               FREE  0.7330FM 0.0569FM 1.00EX 02-07-69
SER.NO.   110 Z= 92. URANIUM                        (CWN)                               FREE  3.6650FM 0.2845FM 1.00EX 02-07-69
SER.NO.    69 Z= 92. URANIUM                        (CWN)                               FREE  7.3300FM 0.5690FM 1.00EX 02-07-69
SER.NO.    22 Z= 92. URANIUM,FREE ATOM, POINT NUCLEUS, XM = 1.2                          FREE  0.0   FM 0.0   FM 1.20EX 10-18-68
SER.NO.    13 Z= 92.URANIUM                                                              FREE  0.0   FM 0.0   FM 1.50EX
SER.NO.    20 Z= 92. URANIUM,FREE ATOM, POINT NUCLEUS, XM = 1.5, TIGHT CONV              FREE  0.0   FM 0.0   FM 1.50EX 10-18-68
SER.NO.    45 Z= 92. URANIUM,FREE ATOM, FINITE NUCLEUS  (SMALL)            (            FREE  0.7330FM 0.0569FM 1.50EX 12-12-68
SER.NO.    41 Z= 92. URANIUM,FREE ATOM, FINITE NUCLEUS  ( HALF SIZE)  (CWN)              FREE  3.6650FM 0.2845FM 1.50EX 12-11-68
SER.NO.    37 Z= 92. URANIUM,FREE ATOM, FINITE NUCLEUS          (CWN)                   FREE  7.3300FM 0.5690FM 1.50EX 12-10-68
SER.NO.   134 Z= 92.URANIUM 236 NEUTRAL                                     NBGRWS= 4.00000BOHR  6.7876FM 0.5666FM 1.50EX 01-23-69
SER.NO.    95 Z= 92.URANIUM                                                 NBGRWS= 3.25395BOHR  7.3300FM 0.5690FM 1.40EX 01-14-69
SER.NO.    28 Z= 92. URANIUM                (CWN)                              RWS= 3.25395BOHR  7.3300FM 0.5690FM 1.50EX 12-04-68

SER.NO.    25 Z= 95. AMERICIUM   (CWN)                                                  FREE  0.0   FM 0.0   FM 1.50EX 02-05-69
SER.NO.    73 Z= 95. AMERICIUM 243     (CWN)                                            FREE  0.6860FM 0.0567FM 1.50EX 01-03-69

SER.NO.    83 Z= 96. CURIUM                 (CWN)                                        FREE  0.6900FM 0.0567FM 1.50EX 12-27-68

SER.NO.    59 Z= 97. BERKELIUM                       (CWN)                               FREE  0.0   FM 0.0   FM 1.50EX 12-19-68
SER.NO.    35 Z= 97. BERKELIUM                       (CWN)                               FREE  0.7190FM 0.0567FM 1.50EX 12-16-68
SER.NO.    63 Z= 97. BERKELIUM                       (CWN)                               FREE  6.8961FM 0.5666FM 1.50EX 12-19-68
SER.NO.   135 Z= 97.BERKELIUM 247 NEUTRAL                                   NBGRWS= 4.00000BOHR  6.8961FM 0.5666FM 1.50EX 01-24-69

SER.NO.    89 Z= 99. EINSTEINIUM 254        (CWN)                                        FREE  0.6900FM 0.0567FM 1.50EX 12-31-68

SER.NO.    49 Z=101. MENDELEEVIUM 256       (CWN)                                        FREE  0.6980FM 0.0567FM 1.50EX 01-02-68

SER.NO.    67 Z=102. NOBELIUM 258           (CWN)                                        FREE  0.7010FM 0.0567FM 1.50EX 01-02-69

SER.NO.    74 Z=103. LAWRENCIUM             (CWN)                                        FREE  0.7000FM 0.0567FM 1.50EX 01-10-69

SER.NO.   105 Z=104. FICTICIUM +                (CWN)                                     FREE  0.7000FM 0.0567FM 1.50EX 01-16-69

SER.NO.   111 Z=105. MERETRICIUM +              (CWN)                                     FREE  0.7000FM 0.0567FM 1.50EX 01-16-69

SER.NO.   117 Z=106. MENDACIUM   +              (CWN)                                     FREE  0.7000FM 0.0567FM 1.50EX 01-17-69

SER.NO.   100 Z=107. PREVARICACIUM +  (CWN)                                               FREE  0.7000FM 0.0567FM 1.50EX 01-20-68
SER.NO.   113 Z=107. PREVARICACIUM    (CWN)    (6D5/2) 1                                  FREE  0.7000FM 0.0567FM 1.50EX 01-21-69

SER.NO.   108 Z=110. PHONCLIUM        (CWN)    (6D5/2) 5 (7S1/2) 1                        FREE  0.7000FM 0.0567FM 1.50EX 01-29-69
SER.NO.   131 Z=110. PHONCLIUM        (CWN)    (6D5/2) 4                                  FREE  0.7000FM 0.0567FM 1.50EX 01-22-69
SER.NO.    55 Z=110. PHONCLIUM        (CWN)    (6D5/2) 4                                  FREE  7.3258FM 0.5666FM 1.50EX 02-12-69
SER.NO.   128 Z=110. PHONCLIUM        (CWN)    (6D5/2) 5 (7S1/2) 1                        FREE  7.3258FM 0.5666FM 1.50EX 02-17-69


SER.NO.    18 Z=112. BOSCCLIUM        (CWN)    (6D5/2) 6 (7S1/2) 2                        FREE  0.7000FM 0.0567FM 1.50EX 02-03-69
SER.NO.   114 Z=112. BOSCCLIUM        (CWN)    (6D5/2) 6 (7S1/2) 2                        FREE  7.3430FM 0.5666FM 1.50EX 02-14-69

SER.NO.   129 Z=114. COGITARIUM  (7S)2 (7P)2  (CWN)                                       FREE  7.3601FM 0.5666FM 1.50EX 02-20-69
```

Fission Fragment Data Analysis

In experimental studies of the fission process of heavy nuclei, we are interested in obtaining relationships among certain fission products, for example, the number of neutrons emitted as a function of fragment mass. Knowledge of the details of such distribution is essential in our endeavor to understand the fission process. In the experimental approach used at the Laboratory, measurements are made of the energies of the two fragments and the time of flight of one of the fragments. The data recorded from the observation of a fission event are three quantities which are functionally related to the energies and time. Through the use of auxiliary experiments we are able to transform instrumental output into the quantities of interest. Statistical error considerations make it necessary to include the measurements from several thousand events in an analysis. Analysis in this case is largely concerned with obtaining the distributions of interest through the application of kinematic conditions which the fission particles must meet.

The energy-energy-time experiments (EET) were initiated before the IBM 360/75 computer was purchased by the Laboratory. Until now data analysis for the EET experiments has been done with the CDC 1604-A computer, in steps, by a number of programs. Some of the arrays of interest contain approximately 66,000 elements, and analysis with the 1604-A programs required the use of a number of magnetic tapes for storage of arrays and output from intermediate programs. Because of the needs of this task and the need to make some changes in the techniques of analysis, we wrote a new code for the 360/75 which utilizes slow core and disk storage. The time-consuming multistage task of the old codes has been assumed by one program which performs the analysis with one job submission and a few minutes of computing time. We will soon add routines for plotting to the output routine, with one for making contour plots of three-dimensional distributions. We have structured the program so that calculations in the center-of-mass system of the fission products can be accommodated with minimal effort.

PROGRAMMING SERVICES

Many times a programmer is called upon to write a very specific program to perform what are termed minor calculations, to make small modifications to an existing program, or to rework an existing program to reduce its execution time. Taken collectively these activities represent a significant fraction of the work done by the programming staff.

REACTOR

D. E. Arnurius
T. D. Calton
N. B. Gove
C. W. Nestor, Jr.
H. M. Nicholson
V. A. Singletary

Steam-Graphite Reaction

Water vapor that is released accidentally into the coolant stream of a high-temperature gas-cooled nuclear reactor is carried with the coolant stream into the fuel element channels, where it diffuses into the graphite matrix of the unclad fuel element. At high temperatures, reaction takes place between water vapor and the graphite, which results in graphite loss from the fuel element matrix.

We have written a computer program to predict the rate of loss of the graphite as a function of time and radial position in the cylindrical fuel elements. The program allows for the variation of the diffusion coefficient of graphite with temperature and the time-integrated local rate of graphite loss. Turbulent and laminar flow conditions can be studied.

The results of this program provide determination of the local and time- or position-integrated graphite loss in the matrix as a function of time and position.

Pressure Vessel Stress Analysis

The designs of hot water heaters, nuclear reactors, boilers, submarines, interstate oil lines, and numerous other pressure vessel structures are based in part on the ASME[51] pressure vessel code. Certain structural specifications in this code are being examined in the hope that new specifications can be established that would permit a more economical design in many pressure vessel structures. One of the first structures to come under investigation was the cylinder-to-cylinder structure. Explicit solutions (involving infinite series and integrals over an infinite ray) for the internal pressure and out-of-plane bending moments for the cylinder-to-cylinder structure have been derived by the General Technology Corporation under support by the Navy. These solutions have been checked by a consultant to the Reactor Division, and certain minor errors were found and corrected.

From detailed notes received from the Reactor Division, we wrote a computer program to perform the necessary calculations for the out-of-plane bending moments. We are now checking and refining this program. At present our attention is centered on the problem of numerically approximating the integrals

$$\int_1^\infty J_{n-m}(tr) + J_{n+m}(tr)\, H_m^{(1)}(tr)\, dr, \qquad n = 1, 3 \text{ and } m = 1, 2, \ldots, M,$$

where J and $H^{(1)}$ represent respectively Bessel and Hankel functions of the first kind and where $t = a(1 + i)$ with $a > 0$. Soon after this problem is solved, the program should become fully operational. The problem solved by the computer program is an approximation to the real problem. If the results of this approximation are reasonable for the real problem, a parameter study will be made with the program.

REACTOR CHEMISTRY

D. E. Arnurius
Sandra H. Merriman

Calculation of the Efficiency of Removal of Particles by an Air-Cooled Condenser

We wrote a computer program to aid in a preliminary study that focused on the calculation of the fraction of the total incoming particles (particles plus air are mixed with steam at about 100°C and are led into the bottom of an air-cooled condenser) which have reached the condenser wall at any height up the condenser. We approximated certain terms in the radial and vertical velocity components of the particles by finite differences.

---

[51] American Society of Mechanical Engineers.

SOLID STATE

E. J. Lee
C. W. Nestor, Jr.

ESR Powder Patterns

We wrote several routines to calculate theoretical powder pattern spectra for rhombic, axial, and cubic symmetries and weighted combinations thereof. These routines will be used in analyzing spin resonance data from the lunar samples expected next year in the Solid State Division.

SUBROUTINES

D. E. Arnurius
R. C. Durfee
C. W. Nestor, Jr.
V. A. Singletary

Higher Quality Printing for CDC 1604-A Printed Output

We have written a FORTRAN/360 program to read a standard CDC 1604-A output tape and print the information contained therein on a printer attached to the IBM 360/75 computer. A translation table for the 15 special characters available on the printers attached to the CDC 160-A but not available on those attached to the IBM 360/75 may be defined by the user at execution time. We hope that the use of this program will be restricted to those CDC 1604-A jobs where the printed output will be part of a report or document.

FORTRAN Output Without FORMAT Statements

We wrote three subroutines, PRINTF, PRINTI, and LIST, which print FORTRAN variables in a fixed format with user-supplied headings. PRINTF and PRINTI print any of the following types of variables:

1. a complete matrix,
2. a single row of a matrix,
3. a single column of a matrix,
4. a single element of a matrix,
5. a vector,
6. a simple variable.

PRINTF is used with floating point variables, while PRINTI is used with fixed point variables. For example:

```
      CALL PRINTF( RMATRIX, NROWS,NCOLS, NRMAX, 'ROWS', 'COLUMNS ',
     1                 ' CORRELATION     MATRIX $' )
```

The LIST subroutine is designed to print any number and type of simple variables. For example:

```
      CALL LIST( 4, ' A,B,I,J=$', A,B,I,J )
```

These subroutines may be used on the IBM 7090 or the IBM 360/75.

THERMONUCLEAR

J. D. McDowell
C. W. Nestor, Jr.
F. W. Stallmann

Investigation of Compound Superconductors

We are currently investigating the relationships between heat flow, temperature, voltage, and current in a compound superconductor.

Systems Programming

Nancy B. Alexander
R. T. Boughner
A. M. Craig, Jr.
Janet B. Estill
C. E. Hammons
R. P. Rannie
J. R. Stockton
J. G. Sullivan
Carolyn P. Walker

IBM/360 OPERATING SYSTEMS

During 1968 the IBM 360/75 computer has been operated using three versions of the System/360 operating system received from IBM. This operating system performs multiprogramming with a fixed number of tasks (MFT). Our operation of the system has been carried out with memory allocated to either two or three tasks. Waiting time in higher priority tasks is used for computation in lower priority tasks to produce overlapping of tasks. The task with the highest priority contains the Houston Automatic Spooling Priority (HASP) system. The HASP system controls the card readers, card punches, and printers and provides for remote job entry and operator communication.

The task with the lowest priority is the problem program. For three months during the summer the system was run with a third, intermediate-priority, teleprocessing task. This was run as an experiment to study interactive communication between the IBM 360/75 computer and two IBM 2740 communication terminals (typewriters).

Several hardware devices were attached to the 360/75 during the year. Testing of the IBM 2321 data cell began late in January, and by March the number of problems encountered in using it made it clear that the data cell would not perform well as a part of the IBM 360/75 configuration. The most serious problem with the data cell was the inability of its storage control unit to transfer executable programs from the data cell to memory. As a consequence, the data cell was returned to IBM in June and was replaced by the IBM 2314 unit.

The IBM 2314 direct access storage facility (DASF) has a storage capacity of over 233,000,000 bytes of data on eight on-line disk drives. Prior to the acquisition of this device the direct access storage on the eight IBM 2311 disk drives was 58,000,000 bytes. The 2314 DASF has increased the available direct access space by a factor of 5 in addition to providing twice the data transfer rate of the 2311 disk drives.

Because of the additional space, it is now possible to provide storage on 2311 disk drives for frequently used programs and small data arrays. In addition, dismountable disk packs can now be used for storage of large data arrays. Large amounts of "scratch" space previously available only on magnetic tape can now be allocated to the 2314 disk drives, eliminating the delays associated with tape operations.

Communication between the IBM 360/75 and the SEL 840A computer at the ORIC facility via the IBM 2701 data adapter unit under control of the HASP program was begun in May.

In September the LCS nucleus was used in normal operations. (See "The Large Core Storage Nucleus" in this report.) This modification produced a significant reduction in the job overhead time. Several routines were added to the LCS nucleus by October, and studies are under way to determine which additional routines should be included.

Generation of an experimental Multiprogramming with a Variable number of Tasks (MVT) system was begun in late October. This MVT system will be used to obtain operating experience on the IBM 360/75 computer in preparation for MVT operation on the IBM 360/91 computer scheduled to be installed early in 1969. Use of the present operating system (MFT) on the IBM 360/75 computer is expected to continue until the MVT system is demonstrably equal in performance to the present system and until the MVT system can provide the remote job entry facility presently available under HASP.

OPERATIONAL IMPROVEMENTS

During the year a number of improvements have been incorporated into the operating system of the IBM 360/75 computer. These improvements are designed to achieve faster job execution or greater facility in running the operating system, or, hopefully, both.

The creation of the large core storage nucleus, described separately, was one endeavor in systems improvement. A quantitative measure of the reduction in job execution time resulting from the use of the LCS nucleus and other operational improvements was obtained by running some or all of a battery of 18 bench-mark jobs following successive "improvements" in the operating system. The execution time for these jobs, measured to the nearest 0.01 sec, provides a measure of the reduction in system overhead time. Timing studies with these bench-mark jobs provide a measure of system operation times and other parameters of the IBM 360/75 system. Use of the LCS nucleus resulted in a saving of about 8 sec per job or 1 hr/day.

Manual loading into the card reader of the control cards used to execute the startup procedure has been eliminated. The control cards are now stored on a disk, and the system automatically issues a command to start processing by reading these control cards from the disk.

Removal of temporary data sets from all system direct access devices, regardless of the number and type of devices in the operating configuration, is now provided by the new GETDSCB program as one step in the system startup procedure. Another step in the system startup procedure now provides automatic cataloging of data sets which are dependent on the systems residence device in use. This feature facilitates the use of any of the three types of devices now available for systems residence and provides a wider range of operating configurations which can be run when one or more of the direct access devices are inoperative.

Adoption of a standard procedure for the allocation of magnetic tapes has enabled premounting of tape volumes, with a decrease in job overhead time for jobs using magnetic tape. The storage of frequently used programs on direct access devices has also increased job throughput. Techniques to accomplish premounting of direct access volumes have been developed to save disk mounting time. Improved methods for space allocation on direct access devices have been adopted, frequently used modules have been included in the system library, and so far as possible the system has been directed toward the utilization of the highest speed direct access devices.

Changes have been made to the object deck handling routine HEXTAP, renamed HEXHAN. The examples printed out now include the data set name, unit, and volume serial number for the particular job being run. The special control cards needed for using a HEXHAN data set are punched with the proper information.

New cataloged procedures have been added to the IBM 360/75 procedure library and old procedures changed according to programmers' requirements. New procedures provide for the execution of assemblers,

compilers, and link editors in various combinations. Special tape conversion procedures, procedures for system startup, a procedure to correct tape drive allocation problems, and a number of procedures to delete temporary data sets from direct access devices have also been added.

Maintenance on the system library routines has continued, and new routines have been added to the system library. Existing routines have been checked, extended, and supplied with error messages. The magnetic tape manipulation routine has been improved through the addition of still further options, and utility programs making use of these routines have been added to the system library. Additional error messages have been added to the existing tape buffering package, and the maximum record length has been extended to 2,000,000 bytes. The package has been modified to facilitate future changes. The Calcomp plotting routines have been updated for use with the new Calcomp pen-and-ink plotters, and plotting routines for the Calcomp cathode ray tube plotter have been added to the library. An improved system of incorporating information concerning the library routines within the library itself has been put into operation.

The Large Core Storage Nucleus

The memory of the IBM 360/75 computer contains 2560K (K = 1024) bytes. There are 512K bytes of main memory (0.75 μsec) and 2048K bytes of large core storage (LCS) memory (8 μsec).

The nucleus of the operating system takes up 54K bytes of the 512K bytes of main memory. Of this amount, 34K bytes are the nucleus proper, and 24K bytes are occupied by frequently used routines and work areas. Incorporation of these routines into the nucleus speeds up operation of the system because they no longer have to be loaded from direct access devices each time they are to be executed.

In order to expand the number of these routines and increase system performance without taking even more main memory away from the problem programmer, a program was written that transfers these routines from main memory to a region of LCS designated as the LCS nucleus. The program then releases the main memory occupied by those routines which have been moved. This liberated main memory is then available for use by the problem programmer.

The creation of the LCS nucleus does not detract from the available LCS, since the LCS nucleus uses a portion of LCS previously reserved for the HASP, teleprocessing programs, and other system functions.

Figure 12 shows memory at three times: without extension of the main memory nucleus, with extension of the nucleus but before creation of the LCS nucleus, and the final configuration after creation of the LCS nucleus.

OPERATOR TRAINING

Training courses for the IBM 360/75 computer operators have been conducted approximately once every week by the systems group. Each course lasts approximately 1½ hr and is held during the weekly scheduled preventive maintenance service on the system. Approximately eight operators attend each course.

The topics covered in these courses included information on changes in hardware configuration, operating systems and procedures, operation of the system with various devices inoperative, handling of abnormal conditions, dealing with special requests by users, and efficient operation of all systems facilities.

SYSTEM STUDIES

A study is under way to determine which of the routines in the supervisor call (SVC) library on the system resident direct access devices are used most frequently. Routines from this library are loaded into

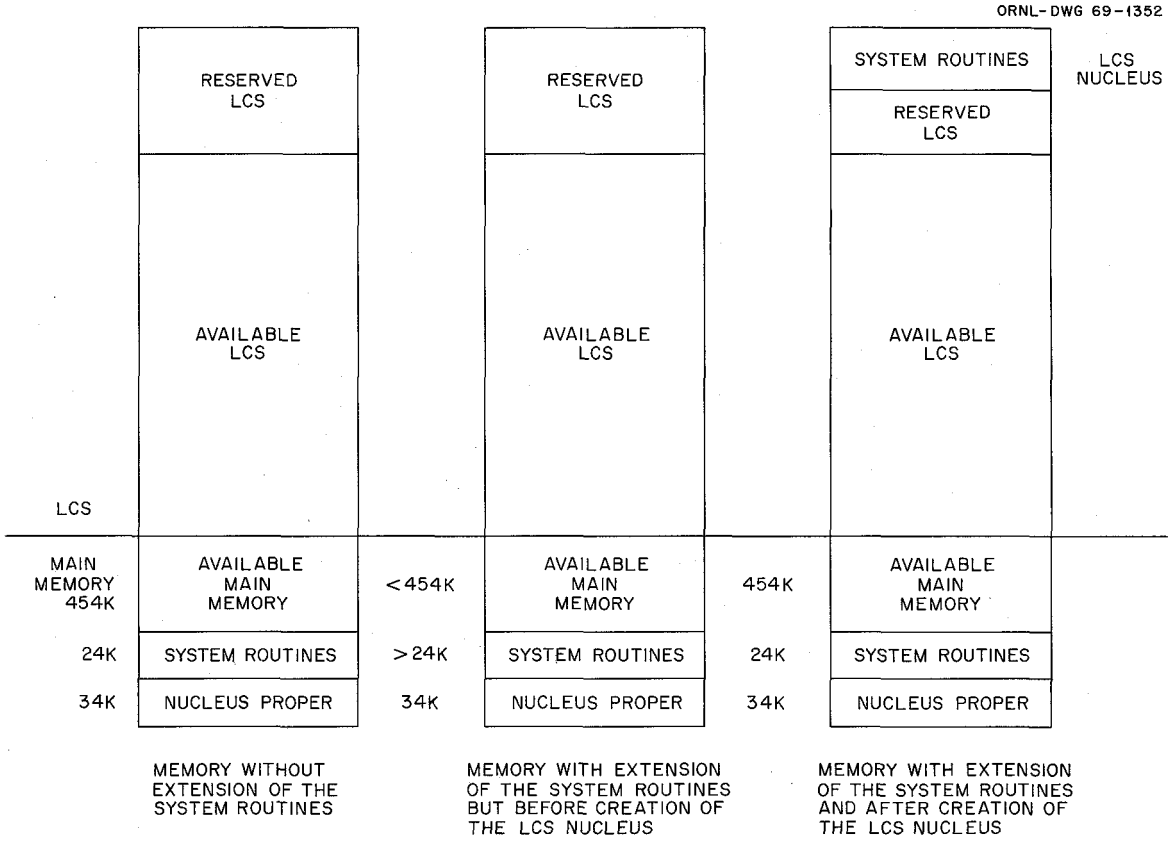


Fig. 12. Memory of the IBM 360/75 Computer – the LCS Nucleus for Extended System Routines.

memory from the devices and then executed. Because the loading of routines from this library may occur often, delays averaging one half of the revolution time of the systems resident device plus the loading time (~9.4 msec for the IBM 2301 drum) may occur for each load. It may be possible to achieve further significant reductions in system overhead time by making the more frequently used of these routines resident within memory.

Programs have been written and changes have been incorporated in the operating system to monitor and tabulate loads from the SVC library in order to determine where the greatest gains can be expected.

COMPILER DEVELOPMENT

ORNL FORTRAN

We have maintained in service the compiler for the ORNL FORTRAN language.[1-3] This compiler is in use on the IBM 360/75 computer. A survey made of users of the IBM 360/75 computer shows that two-thirds of them make use of ORNL FORTRAN.

We will make few, if any, revisions of the compiler. We will, of course, attempt to correct any errors found in it.

---

[1] *Math. Div. Ann. Progr. Rept. Dec. 31, 1965,* ORNL-3919, p. 56.
[2] *Math. Div. Ann. Progr. Rept. Dec. 31, 1966,* ORNL-4083, p. 54.
[3] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 30.

General Purpose Compiler

We have worked on a new compiler[3] designed to compile a number of languages. Originally the compiler was planned in three phases: (1) translate source language program to a process language program, (2) eliminate programming inefficiencies, thus reducing execution time, and (3) produce object code.

We found the use of a process language to be both restrictive and unwieldy. For this reason we replaced the first phase with a conversion of the source language program to a "matrix" representation. A row of the matrix represents information about elements of the program. Examples of elements of a program are local variables, variables in COMMON, and a fixed-point register. A column of the matrix is headed by an indicator which specifies, for example, one of the arithmetic operations, branches, or definitions of a statement label. We can handle matrices up to 8000 by 8000; however, we find that a satisfactory size is 2000 by 6000. Since the matrix resulting from a program is necessarily sparse, we are able to save storage space by storing the entries in a special manner.

At present we have phases 1 and 3 finished and have used them. We will not release this compiler for use until the first version of phase 2 is completed. On this we are working.

By a general purpose compiler we mean one that can be used to compile a number of languages. As such this compiler is restricted to FORTRAN-like languages. Even so restricted, this is a versatile compiler. At present we are working to use the compiler on the BASIC language. We have found that the compiler is working as expected.

CATHODE RAY DISPLAY

An IBM 1800 computer and IBM 2250 cathode ray tube (CRT) display unit were installed at ORNL in June 1968 to be rented for a year. The computer is interfaced with the IBM 360/75 computer. The configuration consists principally of: (1) an 8K central processor, (2) a disk unit, (3) a card reader/punch, (4) a typewriter, (5) a display unit and independent buffer, and (6) the input/output (I/O) channels for the 2250 unit and 360/75 interface.

This particular configuration of the computer is nonstandard, and the software available for this system does not efficiently support this configuration. However, with IBM's cooperation, we developed a more efficient, ad hoc, monitor system from a disk-monitor system for the IBM 1130 computer. The 1130 order codes are a subset of those for the 1800, but the interrupt handling is different. Once the differences were taken care of, the monitor system became quite useful.

The CRT unit has several very interesting and useful features. The display is self-regenerating; thus the only requirements for the 1800 computer in conjunction with the display are to initiate I/O operations to the display's buffer and to handle various interrupts caused by I/O and manual operations. This leaves the 1800 computer essentially free to handle other displays, I/O operations, or computational problems.

The display has a programmed function keyboard (PFK) with which one can indicate to a monitoring program up to 8192 different operations to be performed or user-defined programs to be executed.

Another feature of the display is the light-detecting pen. This feature in conjunction with the PFK can be used to indicate specific areas on the display to the monitoring program. For example, one could indicate specific points in a spectral display to be added or deleted.

The alphanumeric keyboard is a feature useful for data input. For example, FORTRAN program modifications may be typed in, or data point values may be given to plotting programs. These data are displayed as they are typed, and the characters are automatically displayed by the display unit with no programming or interrupt handling required for their generation.

With these features properly utilized, an experimental physicist could interact with the large 360/75 system by having rapid access to preliminary plots of data, study these, make parametric changes, and submit the problem again to the larger system. Without these features, problems of this nature are presently beset by the turnaround time on the local computers, and each modification of parameters would cost one day in time.

A rudimentary interaction system has been developed allowing a limited amount of interaction with the small computer system. A depression of one of the PFK's causes a directory of selectable functions to be displayed, whereupon the user may indicate via the keyboard a particular function desired. For instance, a least-squares program may be initialized, data points typed in via the alphanumeric keyboard, and finally a plot of the polynomial generated may be viewed.

Further work is being done to expand this system so that eventually interaction with the 360/75 system may be achieved. Other data manipulation operations will be added to the PFK directory, and an attempt will be made to inveigle a few physicists or mathematicians to experiment with the system and supply an evaluation of it.

TAPE UTILITY

A tape utility program is available on the IBM 360/75 computer. With it one can copy tapes, list the contents of the tapes in hexadecimal or alphanumerics, and perform other tape manipulations. One of the features of this utility is that one does not need to know the parity, record size, or density of the tape in order to use the code. The program will find the correct parity and density and indicate this on the printer output. The utility program gives the length in bytes for each record printed and indicates the maximum record length in bytes encountered in each file of data printed or scanned. When copying tapes the number of records copied is indicated at the end of each file. One can choose the mode of printout, the number of records to be printed, and the position of the tape before reading. A user's manual for the program is being prepared.

Computer Operations

J. E. Parham C. S. Williams

THE IBM COMPUTER

The 360/75 computer system, described in detail in the last annual report,[1] underwent one significant alteration during the year. It was discovered that the 2321 data cell drive, attached to the 360/75 via a selector subchannel on the 2870 multiplexer channel and operated with IBM-supplied software, would not properly perform the task for which it was primarily intended (i.e., the loading and execution of program modules stored there by various users). No solution to this problem was found, even after extensive effort on the part of IBM; therefore the 2321 data cell drive and 2841 control unit were returned to IBM, and a different device was ordered.

In July 1968, the Laboratory took delivery of an IBM 2314 direct access storage facility as a replacement for the data cell. The 2314 has a storage capacity of 207,014,400 bytes on eight removable disk packs. The high reliability of performance and faster access time of the direct access storage facility more than compensate for the fact that its storage capacity is less than that of the data cell. The new device was attached to the 360/75 immediately, passed its acceptance test without incident, and is now operating routinely.

Due to the saturated condition of the 360/75 system and to an increasing need for faster processor speeds on some classes of problems, a proposal was submitted and approved for the acquisition of a "super computer," the IBM 360/91. This computer will not only relieve the backlog of long-running jobs on the 360/75, but will be able to run large problems not heretofore possible. Delivery of the 360/91 is expected in early 1969, and a report of its performance will be made in the next report.

At the end of the report period, the complete IBM 360/75 system is as shown in Table 2.

THE CDC COMPUTERS

The CDC 1604-A and 160-A systems have been described in an earlier report,[2] and no changes have been made to the equipment. System performance continues to be quite good, even though the equipment has been heavily utilized for almost six years. Such problems as do arise are centered in the magnetic tape units, which are electromechanical devices and heavily utilized. Even then, computer system downtime is minimized, since repair can usually be made off-line without interrupting the computer itself.

In the light of the continued heavy use of the CDC computer systems, permission has been secured from the AEC to retain these computers as long as utilization of them warrants it.

[1] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236.

[2] *Math. Div. Ann. Progr. Rept. Dec. 31, 1965,* ORNL-3919.

Table 2. The IBM 360/75 Computer System

Quantity	Type and Model	Description
1	2075-I	Processing unit (0.2 μsec cycle time)
2	2365-3	Processor storage; each unit contains 262,144 bytes (8 binary bits) of magnetic core storage with 0.75 μsec access time
2	2361-1	Large-capacity storage; each unit contains 1,048,576 bytes of magnetic core storage with 8 μsec access time; the two units are interleaved, however, giving an effective access time of about 4 μsec
1	2860-2	Selector channel (2)
1	2870-1	Multiplexer channel with two selector subchannels
1	2314-1	Direct access storage facility; this unit has a storage capacity of 207,014,400 bytes on eight removable disk packs
1	2301-1	Drum storage; this unit has 4,090,000 bytes of storage with a transfer rate of 1,200,000 bytes/sec
1	2701	Data adapter
1	2403-3	Tape unit and control; contains one tape drive with seven-track recording head; recording densities of 200, 556, or 800 bpi give transfer rates of 22.5, 62.5, or 90 KB respectively
2	2402-3	Tape units; each unit contains two tape drives; one drive has a seven-track head as described above, and the other three drives have nine-track heads and record at 800 bpi, giving a transfer rate of 90 KB
8	2311-1	Disk storage drives; each unit has 7,250,000 bytes of storage on a removable disk pack and has a transfer rate of 156 KB
3	1403-N1	Printers; 1100 lines/min; 132 characters/line
1	2540-1	Card read-punch (1000 cards/min reading; 300 cards/min punching)
1	2501-B2	Card reader (1000 cards/min)
1	2520-B2	Card punch (500 cards/min)
2	2841-1	Disk control units
1	2821-1	Printer and read-punch control unit
2	2821-2	Printer control units
1	2820	Drum control unit

WORK LOAD AND PERFORMANCE

A number of figures follow which indicate the use of the various computer systems and which give some indication of equipment performance. Figure 13 gives the productive hours and downtime recorded for the IBM 360/75 for the past three years; Fig. 14 gives the same information for the CDC 1604-A. Figures 15 and 16 show the use by division of the 360/75 and 1604-A, respectively, while Fig. 17 indicates the monthly work load on both computers in terms of number of jobs per scheduled work day.

As indicated in Fig. 13, the work load on the IBM 360/75 remained at the saturation level for the entire year. Improved maintenance schedules and procedures mentioned in the previous report have continued to be successful; as indicated in Fig. 13, downtime on the system was kept to an average of 20 to 30 hr/month.

Figure 14 indicates a decline in the work load on the CDC 1604-A following the usual summer peak loads. This decline is attributed not only to the usual year-end letup in work but also to the anticipated switching of work from the 1604-A to the faster, larger 360/75 computer. Downtime on the 1604-A stayed at a remarkably low level throughout the year, as shown in Fig. 14.

Utilization of Computing Center services continues to increase, particularly in use of the IBM 360/75. Figure 17 shows that in May 1968 the number of jobs submitted for the 360/75 exceeded for the first time the number of jobs submitted for the 1604-A and has continued to increase since that time.

In 1968, over 700 different users of the computer systems submitted almost 157,000 jobs, requiring in excess of 12,400 hr of computer processing time.

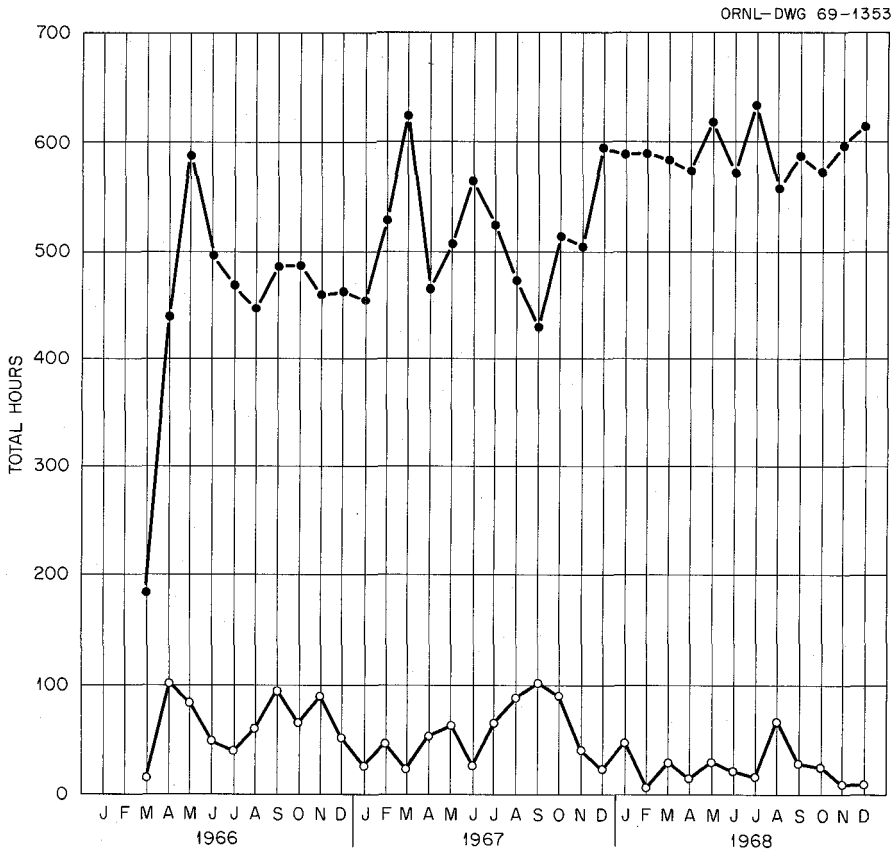


Fig. 13. IBM 360/75 Productive Time and Downtime.

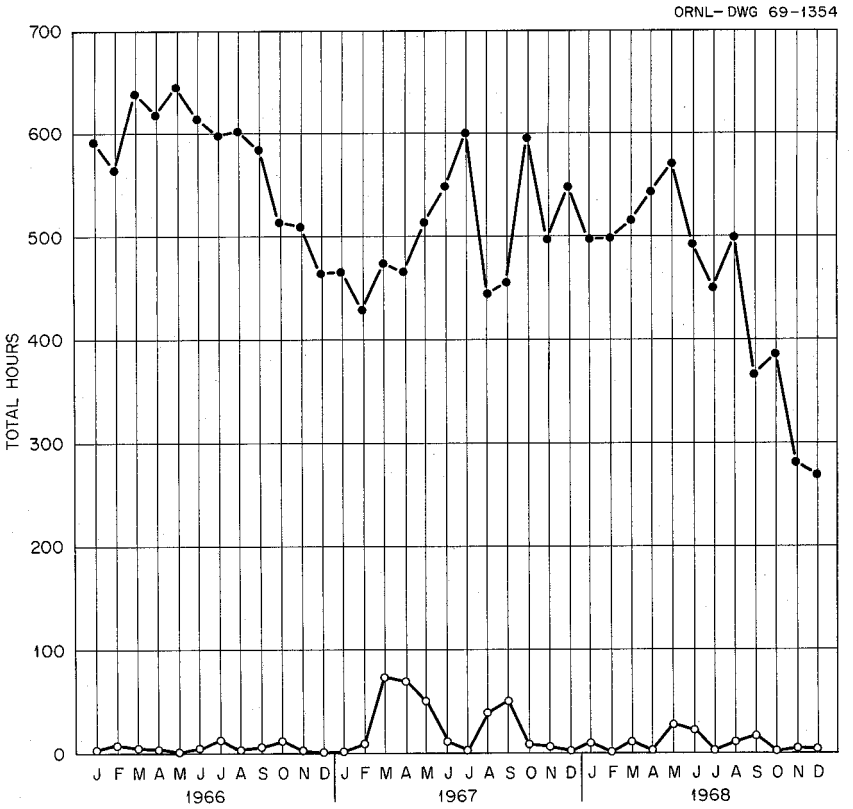


Fig. 14. CDC 1604-A Productive Time and Downtime.

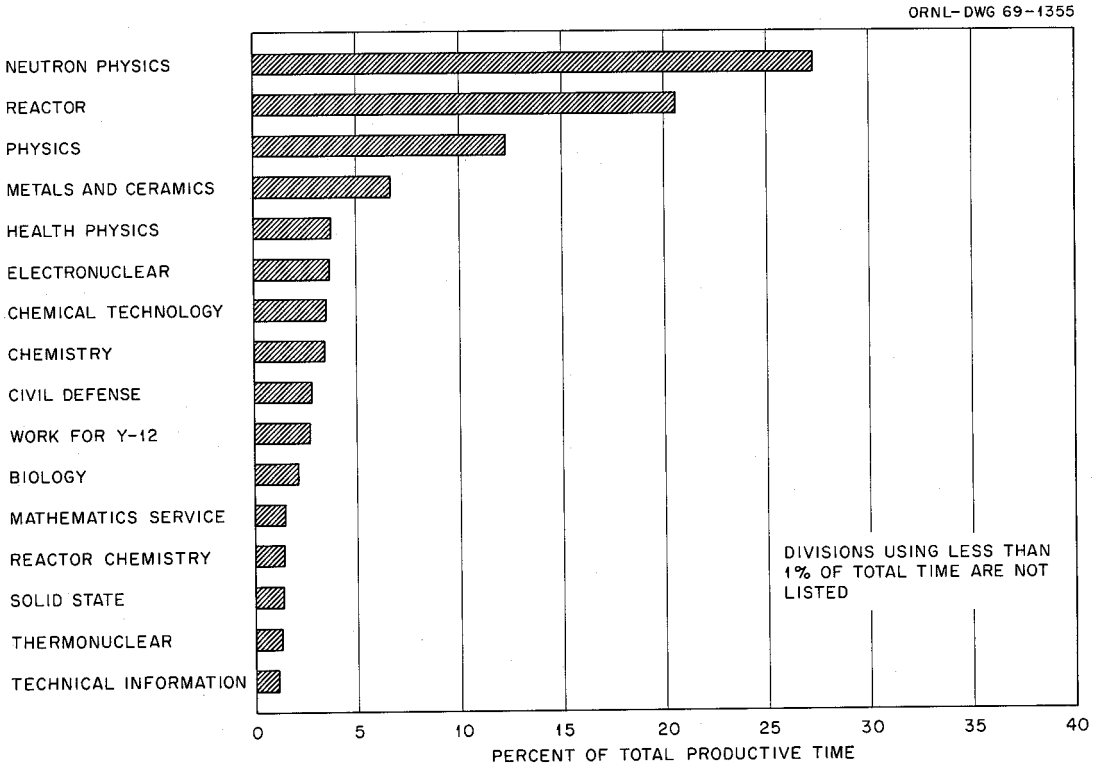


Fig. 15. Divisional Use of the IBM 360/75 Computer.

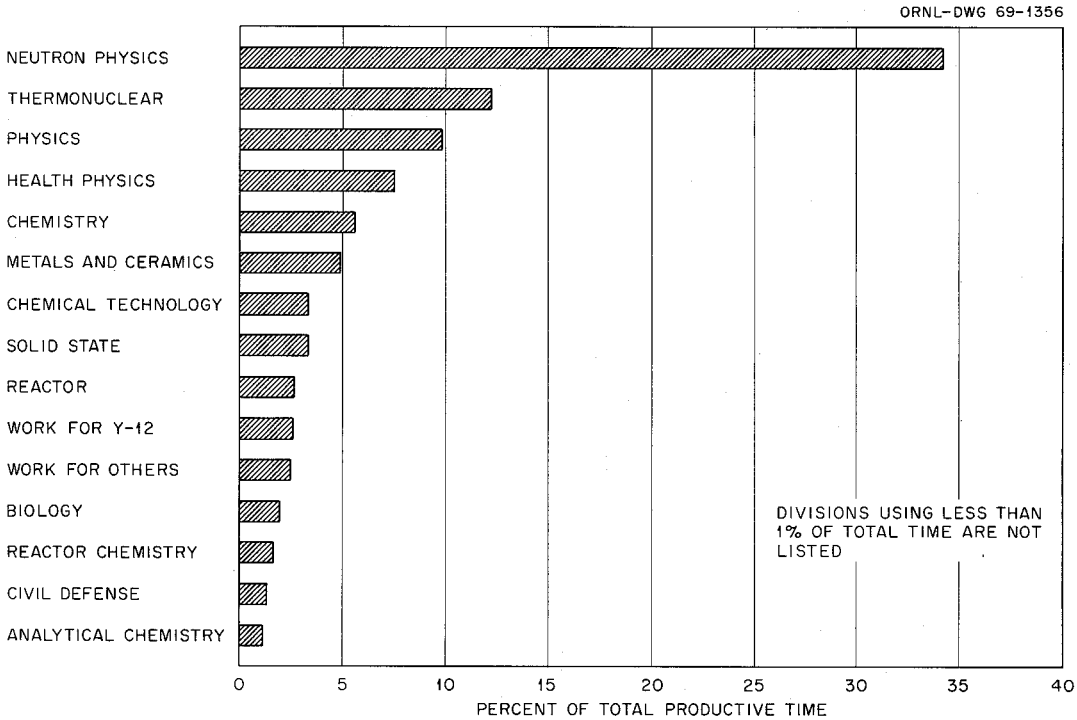


Fig. 16. Divisional Use of CDC 1604-A Computer.

Fig. 17. Work Load at the ORNL Computing Center.

PLOTTERS

In January 1968, the Computing Center took delivery of three Calcomp (California Computer Products) plotting systems: two digital incremental plotter systems for producing pen-and-ink plots and one electronic digital incremental plotter system which produces plots on a cathode ray tube and from there to 35-mm film. The new systems passed acceptance without incident and were put to use immediately, replacing two older pen-and-ink plotter systems, which were released to other divisions in the Laboratory. Installed at the same time were an automatic film developer and a film viewer/printer. These devices are utilized in connection with the cathode ray tube plotter. Plotting equipment at the ORNL Computing Center is now as follows:

Narrow Incremental Plotter

1 Calcomp 765 plotter. This unit operates at 1687 steps/sec when in "ZIP" mode, otherwise 450 steps/sec. Plots can be up to 11 in. wide and up to 120 ft long. Full-step size is 0.005 in.; half-step size is 0.0025 in.

1 Calcomp 770 magnetic tape unit.

Wide Incremental Plotter

1 Calcomp 763 plotter. This unit operates at 1687 steps/sec when in "ZIP" mode, otherwise 450 steps/sec. Plots can be up to $29^{1}/_{12}$ in. wide and up to 120 ft long. Full-step size is 0.005 in.; half-step size is 0.0025 in.

1 Calcomp 770 tape unit.

Cathode Ray Tube Plotter

1 Calcomp 835 electronic digitial plotter. This unit produces images on the face of a cathode ray tube at the rate of 24,000 steps/sec. Image is recorded on 35-mm film.

1 Calcomp 780 magnetic tape unit.

OPERATING STAFF

Functions of the various service groups in the Computer Services Section have been described in previous reports. Some turnover was experienced in the keypunch group, but the personnel level in this group was maintained and the level of output has not suffered. Additions were made to the operator staff, bringing the number of computer operators to 24 and permitting the beginning of a fully scheduled six-day work week. Personnel levels remained the same in the other service groups as they continued the level of service offered to all staff members of the Laboratory.

Mathematical Statistics

J. J. Beauchamp
W. J. Blot[1]
J. A. Carpenter
T. A. DeRouen[2]
J. R. Doner[2]
D. A. Gardiner
D. G. Gosslee
T. L. Hebble
D. G. Hoel
Norma C. Hull
Kay L. Kannan[2]
M. A. Kastenbaum[3]
Mary A. Kau
E. Leach
Claudia S. Lever
W. E. Lever
R. G. Lovell[4]
F. L. Miller, Jr.
T. J. Mitchell
C. J. Mode[5]
W. H. Olson[6]
Diana B. Smith
R. L. Taylor[2]
V. R. R. Uppuluri

Introduction

This part of the annual report summarizes some of the activities of the Statistics Department of the Mathematics Division during calendar year 1968. As in previous reports, it is organized in three main sections, Statistical Research, Statistical Applications, and Statistical Programming, with an additional section listing seminars on statistical topics given under the auspices of the Mathematics Division.

The technical staff of the Statistics Department consists of 11 statisticians, 4 programmers, and 4 EDPM personnel. The statisticians engage in mathematical research on a variety of topics in probability and statistics and offer a consulting service in the methods of statistics, including the design of experiments. The consulting work is performed with researchers throughout the Oak Ridge National Laboratory and to a lesser extent with the personnel of the U.T.-AEC Agricultural Research Laboratory and the Oak Ridge Y-12 Plant, this latter on a formal basis. Much of the research reported here was suggested by the consulting effort.

[1] On leave of absence.
[2] Temporary summer employee.
[3] Director's Division.
[4] Student trainee.
[5] Research participant.
[6] ORAU Fellow.

The Statistics Department conducted a Lecture Series in Statistics and Probability at the Laboratory between June 10 and August 12 of this year. The series consisted of ten lectures, more or less self-contained, and was open to all research personnel in the Oak Ridge area. Average attendance was slightly more than 100 persons. The lecture notes which were distributed at each lecture were edited and compiled for publication as an ORNL report.[7] A questionnaire mailed to 165 of the attendees, of whom about one-third responded, revealed that the series was, in the main, well received and gave encouragement to the Department to offer similar lectures in the future. The titles of the lectures are listed in the section "Lecture Series in Statistics and Probability."

Statistical Research

COMBINATORIAL ASPECTS OF 2^{k-p} DESIGNS

A 2^k factorial design matrix $\mathbf{D}_k$ is a $2^k \times k$ array of plus and minus signs such that every one of the 2^k possible ordered combinations of k signs is represented as a row of $\mathbf{D}_k$. In experimental applications, each column of $\mathbf{D}_k$ represents one of k *variables*, and each row represents one of 2^k experimental *runs*. The (i, j)th element of the design matrix indicates whether the jth variable is to take its high level (+) or its low level (−) in the ith run.

It can be shown that the rows of $\mathbf{D}_k$ form a commutative group R_k, where the product of two rows $(a_1, a_2, \ldots, a_k)$ and $(b_1, b_2, \ldots, b_k)$ is defined to be the row $(a_1b_1, a_2b_2, \ldots, a_kb_k)$. Any row times itself is the unit element: $(+, +, \ldots, +)$.

We note that R_k is isomorphic to another group H_k, whose 2^k elements are "words" representing subsets of the k variables. The word of H_k corresponding to a row of $\mathbf{D}_k$ is the one composed of those variables which exhibit a minus sign in that row. The product of any two words W_1 and W_2 is that word which contains those variables present in W_1 *or* W_2 but not in both. The unit element of H_k is that word which contains no variables.

Under the definition of product introduced in connection with the rows of the design matrix $\mathbf{D}_k$, the columns of $\mathbf{D}_k$ can be multiplied together in all possible ways to generate a $2^k \times 2^k$ array, called the *factorial estimation matrix* $\mathbf{X}_k$. Every column of $\mathbf{X}_k$ (except for the unit element, the column of pluses) can be represented uniquely as the product of some subset of the columns $(1, 2, \ldots, k)$. In fact, it can be shown that the set of columns of $\mathbf{X}_k$ is isomorphic to H_k.

Any subgroup R_{k-p} of R_k is necessarily of order 2^{k-p}, for some $p \leqslant k$, and can be used to partition R_k into 2^p sets, namely, R_{k-p} and all its cosets. Together, these sets comprise a family of 2^{k-p} *fractional factorial designs,* of which R_{k-p} is the *principal member*. Obviously, the choice of R_{k-p}, which partitions R_k into sets of order 2^{k-p}, also partitions the elements of H_k through the isomorphic relationship between R_k and H_k. The subgroup of H_k which corresponds to R_{k-p} is called the intrablock subgroup S_{k-p}.

The partition of R_k effected by R_{k-p} induces another partition of H_k. This can be seen by noting that there are 2^p columns of $\mathbf{X}_k$ whose elements in R_{k-p} consist entirely of plus signs. These 2^p columns form a group. The corresponding subgroup T_p of H_k is called the *defining relationship* of the design and can be used to partition H_k into T_p and its cosets.

---

[7] J. J. Beauchamp (ed.), *Lecture Series in Probability and Statistics,* ORNL-4347, to be published.

Corresponding to every family of 2^{k-p} fractional factorial designs, therefore, are two subgroups of H_k: (1) the intrablock subgroup S_{k-p}, with 2^{k-p} elements, and (2) the defining relationship T_p, with 2^p elements. It can further be shown that the elements of T_p are those words of H_k which have an even number of variables in common with all the words of S_{k-p}. We note, then, that if T_p is the defining relationship corresponding to the intrablock subgroup S_{k-p}, then S_{k-p} is the defining relationship corresponding to the intrablock subgroup T_p.

Associated with each word in H_k, and therefore with each word in T_p and S_{k-p}, is a property called the *length* of the word, which is merely the number of variables contained in that word. One property of the defining relationship T_p which has recently proved of interest both in categorizing designs and in testing for the equivalence of two designs is the word length vector $\mathbf{t}' = (t_0, t_1, \ldots, t_k)$, where t_l is the number of words of length l in T_p.

The purpose of this research is to investigate the connection between the word length vector $\mathbf{t}'$ of the defining relationship T_p and the word length vector $\mathbf{s}'$ of the intrablock subgroup S_{k-p}, and to develop methods for computing these vectors which do not require the construction of the groups themselves.

Theorem: Let $\mathbf{t}'$ be the word length vector of the defining relationship T_p of a 2^{k-p} fractional factorial design, and let $\mathbf{s}'$ be the word length vector of the intrablock subgroup S_{k-p}. Then

$$t_l = \sum_{v=l}^{k} \sum_{j=0}^{v} s_j \, 2^{p-v} \, (-1)^{v+j-l} \binom{k-j}{k-v} \binom{v}{l} \quad . \tag{1}$$

Outline of Proof: Recalling that each column of $\mathbf{X}_k$ corresponds to an element of H_k, we define the $2^k \times 1$ vector function $\eta(z)$ such that

$$\eta_j = z^{l(j)} \, ,$$

where $l(j)$ is the length of the word in H_k which corresponds to the jth column of $\mathbf{X}_k$. Now let $\mathbf{X}_{k-p}$ be the matrix consisting of those rows of $\mathbf{X}_k$ which correspond to R_{k-p}, and let $\mathbf{r}_i$ be the ith row of $\mathbf{X}_{k-p}$, $i = 1, 2, \ldots, 2^{k-p}$. It can be shown that

$$(1/2^{k-p}) \sum_{i=1}^{2^{k-p}} \mathbf{r}_i'$$

is a vector of 0's and 1's, where the 1's appear only in those positions which correspond to the words of T_p. The function

$$f(z) = (1/2^{k-p}) \left(\sum_{i=1}^{2^{k-p}} \mathbf{r}_i' \right) \eta$$

is then a polynomial in z such that the coefficient of z^l is the number of words of length l in T_p. That is,

$$f(z) = \sum_{l=0}^{k} t_l z^l \, .$$

On the other hand, it can be shown that $\mathbf{r}_i' \, \eta$ is equal to $(1-z)^{n_i}(1+z)^{k-n_i}$, where n_i is the length of the word in H_k corresponding to the ith row of $\mathbf{X}_{k-p}$. Hence,

$$\sum_{l=0}^{k} t_l z^l = \sum_{i=1}^{2^{k-p}} (1-z)^{n_i}(1+z)^{k-n_i} = \sum_{l=0}^{k} s_l(1-z)^l(1+z)^{k-l} . \tag{2}$$

Some algebra is necessary to express the right-hand side of (2) explicitly as a polynomial in z, whereupon we equate coefficients of z^l to obtain (1).

Given the runs of a 2^{k-p} design, therefore, (1) may be used to calculate the word length vector $\mathbf{t}'$ of T_p without deriving T_p itself. If the runs are not given but a set of generators of the defining relationship T_p is given, then it is useful to be able to construct, directly from this set, the word length vector $\mathbf{s}'$ of S_{k-p}. An algorithm for doing this has been developed but not will not be discussed here, except to say that it involves the application of the well-known Yates algorithm to a $2^{k-p} \times 1$ vector which is obtained directly from the generators of T_p.

The result (1) has also been extended to provide formulas for the word length vectors associated with the cosets of T_p, which are more commonly called the *alias relationships* of the design.

AUGMENTING RESPONSE SURFACE DESIGNS

In a previous report,[8] a method of choosing additional runs for a response surface design was proposed and discussed. Basically, this method involves first searching the region of interest R in the factor space for the point p_a at which the variance of the fitted (or predicted) response is maximized. This point, which is in effect the "weakest" point in R, is then chosen for the next experimental run. Assuming therefore that p_a will be run, it is possible to recalculate the variance function over R and find a new "weak point" at which the second additional run will be performed. This procedure can be continued until a set of additional runs has been selected for the next phase of the experiment.

This method, which was originally[8] based on intuitive appeal, can also be shown to satisfy a well-known criterion used in experimental design, namely, minimization of the "generalized variance" of the estimated parameters of the model.

If we let $\mathbf{X}_n$ be the $n \times k$ estimation matrix for the estimated parameters of the model, we define the *generalized variance* of these estimates to be

$$G = |\mathbf{X}_n' \mathbf{X}_n|^{-1} .$$

An important property of the generalized variance is that when the observations are normally distributed, the design which minimizes G is the design which minimizes the volume of the confidence ellipsoid for the parameters.

Now suppose that the original design, corresponding to the estimation matrix $\mathbf{X}_n$, is augmented with r runs. The new estimation matrix $\mathbf{X}_{n+r}$ is an $(n+r) \times k$ matrix which can be written in partitioned form as:

$$\mathbf{X}_{n+r} = \begin{bmatrix} \mathbf{X}_n \\ \cdots \\ \mathbf{A}_r \end{bmatrix} .$$

If $\mathbf{T}_n = (\mathbf{X}_n' \mathbf{X}_n)^{-1}$, then it can be shown that

[8] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, pp. 43–46.

$$|\mathbf{T}_{n+r}| = \frac{|\mathbf{T}_n|}{|\mathbf{I}_r + \mathbf{A}_r\mathbf{T}_n\,\mathbf{A}_r'|}, \tag{1}$$

where $\mathbf{I}_r$ is the $r \times r$ identity matrix. In particular, if $r = 1$, that is, if we add just one point to the design, and if we let $\mathbf{a}'$ be the corresponding row by which $\mathbf{X}_n$ is augmented, we have

$$|\mathbf{T}_{n+1}| = \frac{|\mathbf{T}_n|}{1 + \mathbf{a}'\mathbf{T}_n\,\mathbf{a}} = \frac{|\mathbf{T}_n|}{1 + (V_a/\sigma^2)},$$

where V_a is the variance at p_a after the first n runs and σ^2 is the variance of the individual observations. To minimize $|\mathbf{T}_{n+1}|$, therefore, it is necessary to choose p_a so that V_a is maximized. Hence the choice, for the next run, of the point at which the variance of the predicted response is the largest will actually minimize the generalized variance of the coefficients.

In most situations, however, it is desirable to add more than one point to the current experimental design. It should be pointed out that the minimization of the generalized variance for each additional run, as we propose, will not necessarily result in the lowest possible generalized variance at the conclusion of the augmentation. It would be preferable to choose the r additional runs all at once in such a way that $|\mathbf{T}_{n+r}|$ as given by (1) will be minimized. This appears to pose a rather formidable problem, but until a solution is found the "one-at-a-time" addition of runs seems quite satisfactory.

Previously,[8] we considered searching only a particular path in the region of interest for a point of maximum variance. Programs have since been written to search the entire region of interest, and Figs. 18 and 19 illustrate the results of the procedure in some examples.

Figure 18*a* shows the ten initially run experimental points, as well as the contours of constant variance of the estimated response, assuming this response to be quadratic in the independent variables X_1 and X_2.

ORNL-DWG 68-2647

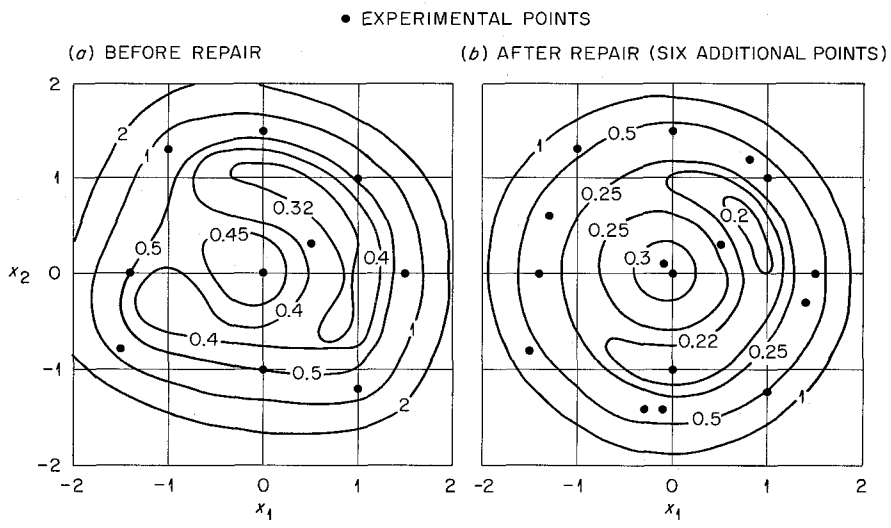


Fig. 18. Variance Contours (Second-Order Model).

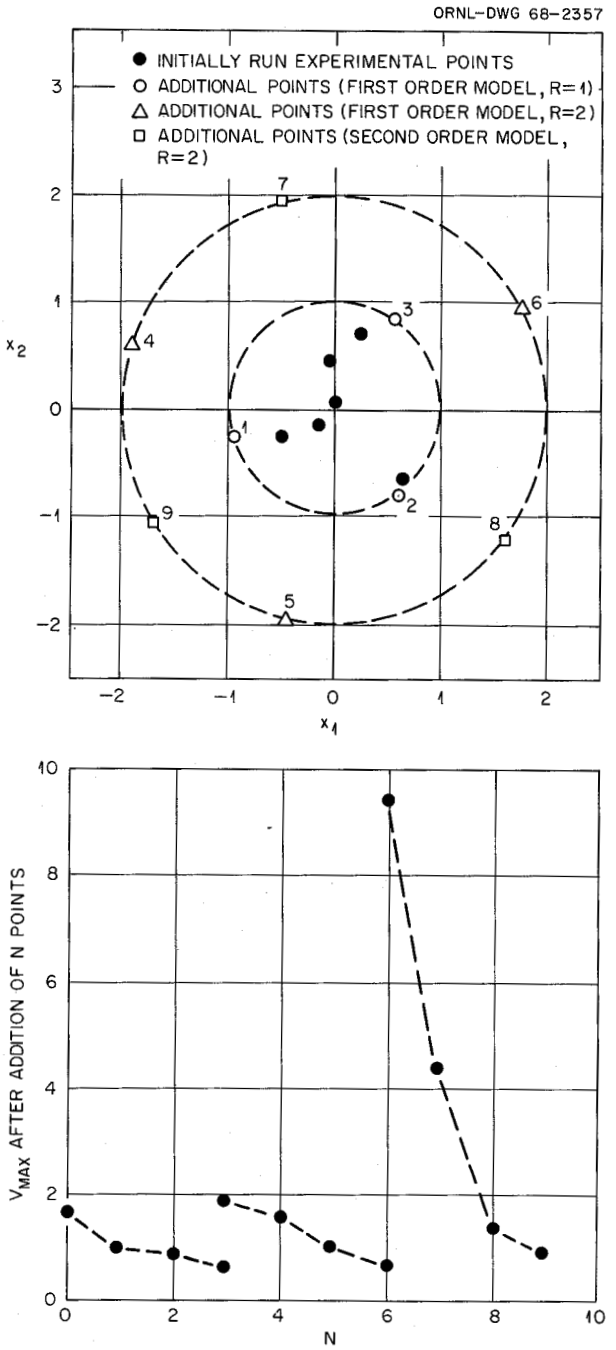


Fig. 19. Sequential Repair and Augmentation of a Response Surface Design.

Had the experiment been run at the initially prescribed conditions (not shown), these contours would have all been circular. When six experimental points are added to this design in accordance with our procedure, the resulting variance contours are as shown in Fig. 18*b*. To a certain extent the "rotatability" of the design (as characterized by circular variance contours) has been recaptured through the choice of further experimental runs.

Figure 19 illustrates how this method may be used not only to repair the original design but to augment the current design when the experimenter chooses to expand his region of interest or change the form of his

model. The discrete jumps in the graph of V_{max} as a function of N correspond to these decisions on the part of the experimenter, since the calculation of V depends not only on the definition of the region of interest but also on the model.

In more than two dimensions, computerized grid search techniques to find the point of maximum variance are extremely time consuming, except in some special cases. Elsewhere in this report is a description of a random search technique, which in practice is clearly superior to the grid search methods used here.

SELECTING THE POPULATION WITH THE SMALLEST VARIANCE

Let $\pi_1, \pi_2, \ldots, \pi_k$ represent k populations and X_{ij} denote the jth observation from population π_i. It is assumed that the observations $\{X_{ij}\}$, $i = 1, \ldots, k$; $j = 1, 2, \ldots$, are independent random variables for all i and j and that the distribution of X_{ij} depends upon an unknown parameter θ_i. Denote the ordered θ_i's by $\theta_{[1]} \leqslant \theta_{[2]} \leqslant \ldots \leqslant \theta_{[k]}$ and define $\pi_{[i]}$ to be the population associated with $\theta_{[i]}$.

We wish to construct a sequential procedure based on $\{X_{ij}\}$ which chooses with a guaranteed probability the "best" population $\pi_{[k]}$. For this guaranteed probability of correct selection we use the indifference zone approach, which is described as follows.

Let $\rho(\theta_{[i]}, \theta_{[j]})$ be a measure of distance between populations $\pi_{[i]}$ and $\pi_{[j]}$ for $i \geqslant j$. Then given the two constants ρ^* and P^* with $\rho^* > 0$, $1 > P^* > 1/k$, we select $\pi_{[k]}$ with probability at least P^* whenever $\rho(\theta_{[k]}, \theta_{[k-1]}) \geqslant \rho^*$. In other words, we require a procedure for which

$$P(\text{correct selection}) \geqslant P^* \text{ if } \rho(\theta_{[k]}, \theta_{[k-1]}) \geqslant \rho^* .$$

A general method has been developed for the construction of sequential procedures which satisfy the above error requirements. To illustrate, consider the problem of selecting the normal population with the smallest variance. Specifically, let X_{ij} be normally distributed with unknown mean μ_i and unknown variance σ_i^2. Next, define $\theta_i = 1/\sigma_i^2$ and

$$\rho(\theta_{[i]}, \theta_{[j]}) = \theta_{[i]}/\theta_{[j]} .$$

Now, if we let the ordered σ_i^2's be denoted by $\sigma_{[1]}^2 \leqslant \sigma_{[2]}^2 \leqslant \ldots \leqslant \sigma_{[k]}^2$, then our error requirement becomes

$$P(\text{correct selection}) \geqslant P^* \quad \text{if} \quad \sigma_{[2]}^2/\sigma_{[1]}^2 \geqslant \rho^* ,$$

where correct selection means the selection of the population with the smallest variance $\sigma_{[1]}^2$. Next we define

$$\overline{X}_{im} = \sum_{j=1}^{m} X_{ij}/m ,$$

$$S_{im}^2 = \sum_{j=1}^{m} (X_{ij} - \overline{X}_{im})^2 ,$$

$$R_{ji}(m) = S^2_{jm}/S^2_{im} ,$$

$$l_{ij}(n) = \frac{n}{2} \log \frac{\tau_0}{\tau_1} + n \log \left[\frac{(1/\tau_0) + R_{ij}(n+1)}{(1/\tau_1) + R_{ij}(n+1)} \right] ,$$

where $\tau_0 = 1/\rho^*$ and $\tau_1 > \tau_0$. It then can be shown that the following sequential procedure satisfies the above error requirements.

Sequential Procedure. Take an initial sample of size 2 from each population and calculate $l_{ij}(1)$ for $i = 1, \ldots, k, j = 1, \ldots, k$. If

$$l_{ij}(1) > a \quad \text{and} \quad l_{ji}(1) \leqslant a ,$$

where

$$a = \log [(k-1)/(1-P^*)] ,$$

we eliminate population π_j from further consideration. If

$$l_{ij}(1) > a \quad \text{and} \quad l_{ji}(1) > a ,$$

then we eliminate π_j if

$$l_{ij}(1) > l_{ji}(1) ;$$

otherwise π_i is eliminated. If only one population remains we terminate the procedure and declare the remaining population to be best. If more than one population remains we proceed to the next stage and take a single observation from each of the remaining populations. The l_{ij}'s are again computed, and the same elimination rule is used except that $l_{ij}(2)$ replaces $l_{ij}(1)$. We proceed in this manner until only one population remains, which we then declare to be best.

This procedure has two important features. The first is that for $\tau_1 < \rho^*$ the procedure is truncated. By this we mean that there is a known maximum to the number of stages which the procedure may require before termination. The second property is that the procedure will eliminate obviously "poor" populations early in the sampling and thus save on the total number of observations taken.

A sequential selection procedure can also be constructed when the measure of distance is

$$\rho(\theta_{[k]}, \theta_{[k-1]}) = \theta_{[k]} - \theta_{[k-1]} = 1/\sigma^2_{[1]} - 1/\sigma^2_{[2]} .$$

It can be proved, however, that for this measure no procedure exists which is truncated.

For further details and references see Hoel.[9]

---

[9] D. G. Hoel, *A Sequential Procedure for Selecting the Population with the Smallest Variance from k Normal Population,* ORNL-4355 (in press).

SELECTING THE POPULATION WITH THE LARGEST MEAN

Let X_{ij} denote the jth observation from population π_i ($i = 1, 2, \ldots, k; j = 1, 2, \ldots$). $\{X_{ij}\}$ is assumed to be a sequence of independent normal random variables with

$$EX_{ij} = \mu_i, \qquad \operatorname{var} X_{ij} = \sigma^2 ,$$

where σ^2 is finite and *known.* Let us denote the ranked μ's by $\mu_{[1]} \leqslant \mu_{[2]} \leqslant \ldots \leqslant \mu_{[k]}$. Paulson[10] has given the following sequential procedure for selecting population $\pi_{[k]}$, the populations associated with $\mu_{[k]}$. He employes the indifference zone approach with

$$P(\text{correct selection}) \geqslant 1 - \alpha, \quad \text{if} \quad \mu_{[k]} - \mu_{[k-1]} \geqslant \Delta ,$$

as the error requirement.

Paulson's Procedure. At stage r we estimate population π_j if

$$\sum_{s=1}^{r} X_{js} < \max_{\nu} \left\{ \sum_{s=1}^{r} X_{\nu s} \right\} - a_\lambda + r\lambda ,$$

where the maximum is taken over the remaining populations. Here λ is a fixed quantity with $0 < \lambda < \Delta$ and

$$a_\lambda = \frac{\sigma^2}{\Delta - \lambda} \log \frac{k-1}{\alpha} .$$

We may modify Paulson's procedure by removing the normality assumption in the following manner. Suppose that X_{ij} has an unknown distribution function F_i with

$$EX_{ij} = \mu_i, \qquad \operatorname{var} X_{ij} \leqslant \sigma^2 ,$$

where σ^2 is finite and *known.* Next modify Paulson's procedure by replacing a_λ with

$$a_\lambda^* = 2 \frac{\sigma^2}{\Delta - \lambda} \left(\frac{k-1}{\alpha} \right) - \frac{\Delta - \lambda}{2} .$$

In order to show that this modified procedure satisfies the error requirement, we suppose that $\mu_{[k]} - \mu_{[k-1]} \geqslant \Delta$. It can then be shown that

$$P(\pi_{[k]} \text{ is eliminated}) \leqslant (k-1) P\left(\sum_{s=1}^{n} Y_s > a_\lambda^* \text{ for some } n \right) ,$$

where

[10] E. Paulson, *Ann. Math. Statist.* **35**, 174–80 (1964).

$$Y_s = -X_{[k]s} + X_{\nu s} + \lambda , \qquad \nu \neq [k] .$$

Letting

$$S_n = \sum_{s=1}^{n} (Y_s - \lambda + \Delta) ,$$

we obtain

$$P(\pi_{[k]} \text{ is eliminated}) \leqslant (k-1) P \{S_n^2 > [a_\lambda^* + n(\Delta - \lambda)]^2 \text{ for some } n\} .$$

Now S_n^2 is a nonnegative submartingale, and thus we have from Chow's[11] generalization of the Hájek-Rényi inequality that

$$P(\pi_{[k]} \text{ is eliminated}) \leqslant (k-1) \sum_{i=1}^{\infty} 2\sigma^2 [a_\lambda^* + i(\Delta - \lambda)]^{-2} \leqslant \alpha .$$

PRIME NUMBERS TREATED AS A MARKOV CHAIN

It is a fascinating observation that there are infinitely many primes that end in 1, 3, 7, and 9, and that each kind is in equal proportion. We may define the state of a prime number by looking at its last digit, so that any prime number can only be in one of the four states 1, 3, 7, or 9. We tried to use the existing statistical methods to prove or disprove that, asymptotically, the prime numbers appear in each of these four states with probability $1/4$. But the fact that we cannot observe a prime number at random hinders the possibility of testing the hypothesis that $(1/4, 1/4, 1/4, 1/4)$ is the probability vector associated with the four states. In other words, if we have the largest known prime number available to us, we can obtain the exact proportions of the primes that end in 1, 3, 7, or 9 up to that number. No statistical test will help to prove or disprove what happens in the limit, since we have not picked a random sample from the entire population of primes.

The fact that we cannot observe a random prime number can be overcome by the following approach. It is well known that there are an infinite number of primes, that there is no known periodicity of primes ending in 1, 3, 7, or 9, and that there is no formula for the nth prime. It is not unreasonable, therefore, to assume that the primes appear in nature in each state at random and that when a prime is observed in state i (= 1, 3, 7, or 9), there is a certain probability π_{ij} with which the next observed prime will be in state j. In other words, we may assume that whenever primes are observed, the transitions may be treated as random realizations from a finite-state, stationary, and ergodic Markov chain.[12]

We might consider the particular realization of primes observed among the sequence of natural numbers, viz., . . ., 11, 13, 17, 19, 23, 29, 31, . . ., as a particular sample function among all possible paths of this ergodic Markov chain associated with primes. This particular sample function is useful in estimating the transition probability matrix $\pi = (\pi_{ij}, i, j = 1, 3, 7, 9)$ by observing the relative proportion of transitions $\hat{\pi}_{ij}$ in the primes up to the largest known prime. Since we have a finite-state Markov chain, these estimates

[11] Y. S. Chow, *Proc. Am. Math. Soc.* **11**, 107–11 (1960).

[12] W. Feller, *An Introduction to Probability Theory and Its Application,* 3d ed., Wiley, New York, 1968.

Table 3. Observed Transitions from Primes in State i to State j ($i, j = 1, 3, 7, 9$) Among the First 788,000 Primes

State	1	3	7	9	Total
1	33,712	61,087	62,994	39,163	196,956
3	45,773	30,993	57,529	62,708	197,003
7	50,536	54,121	31,012	61,434	197,103
9	66,935	50,802	45,568	33,628	196,933
Total	196,956	197,003	197,103	196,933	787,995

Table 4. $\hat{\pi} = (\hat{\pi}_{ij}, i, j = 1, 3, 7, 9)$: Estimate of the Transition Probability Matrix π Based on the First 788,000 Primes

State	1	3	7	9	Total
1	0.17116	0.31016	0.31984	0.19884	1.
3	0.23235	0.15732	0.29202	0.31831	1.
7	0.25639	0.27458	0.15734	0.31169	1.
9	0.33989	0.25797	0.23139	0.17075	1.

Table 5. Eighth Power of $\hat{\pi}$ Given in Table 4

State	1	3	7	9
1	0.24994	0.25001	0.25013	0.24992
3	0.24994	0.25001	0.25013	0.24992
7	0.24994	0.25001	0.25013	0.24992
9	0.24994	0.25001	0.25013	0.24992

$\hat{\pi}_{ij}$ have all the desirable properties. Furthermore, they converge to the elements of the matrix π as all the primes become available.

In Table 3, we give the observed matrix indicating the number of transitions from state i to state j among the first 788,000 primes. For example, 33,712 times we observed a prime in state 1 immediately preceded by a prime in state 1, and so on. In Table 4, we give the same matrix in terms of relative proportions for each state i and denote it by $\hat{\pi} = (\hat{\pi}_{ij}, i, j = 1, 3, 7, 9)$. From theoretical considerations we know that for an ergodic Markov chain, the powers of the transition probability matrix, π^n, converge to a matrix each row vector of which is the stationary distribution

$$\lim_{n\to\infty} \pi_{ij}^n = \alpha_j, \qquad i, j = 1, 3, 7, 9.$$

We would also expect $\alpha_1 = \alpha_3 = \alpha_7 = \alpha_9 = 1/4$. We note that $\hat{\pi}$ is an estimate of π, and from Table 5, we see that even $\hat{\pi}^8$ is reasonably close to the matrix which has all its elements equal to $1/4$. The corresponding results for the first 17,980 primes are available in a technical memorandum.[13]

[13] V. R. Rao Uppuluri, *Prime Numbers Treated as a Markov Chain,* ORNL-TM-2317 (1968).

NONPARAMETRIC TESTS OF RANDOMNESS

Work on nonparametric tests of independence in time series was discussed in a previous report.[14] Three closely related processes, the Ornstein-Uhlenbeck process, the first-order moving-average process, and the first-order autoregressive process, have been studied. For each process the null hypothesis to be tested is that the parameter ρ is equal to zero against the alternative that ρ is greater than zero. In each case the locally most powerful rank order test has been derived. The statistic of the test is a noncircular serial correlation coefficient using order statistics. The asymptotic properties of this statistic and other rank serial correlation coefficients have been studied.

Due to the discovery of other research in this area,[15] the objectives of this work have changed. Now the small-sample properties of the tests based on these serial correlation coefficients are being investigated with the aim of making the tests more easily applicable. Computation of rank order probabilities under the alternative hypothesis is one of the most important aspects of this work. Several approaches are under consideration, but much work remains to be done.

Another model for serial dependence is also being considered. It is based on a model for bivariate distributions proposed by Gumbel.[16]

Gumbel's bivariate distribution of the random variables X and Y is given by

$$F(x, y) = F_1(x)\, F_2(y)\ \{1 + \alpha[1 - F_1(x)]\ [1 - F_2(y)]\},$$

where $-1 \leqslant \alpha \leqslant 1$ and $F_1(x)$ and $F_2(y)$ are the cumulative distribution functions of X and Y respectively. Then the bivariate probability density function is given by

$$f(x, y) = f_1(x)\, f_2(y)\ \{1 + \alpha[2F_1(x) - 1]\ [2F_2(y) - 1]\}.$$

A model for serial dependence between $X_1, \ldots, X_n$ is given by the conditional probability density function

$$f(x_i|x_{i-1}) = f(x_i)\ \{1 + \alpha[2F(x_{i-1}) - 1]\ [2F(x_i) - 1]\},$$

$i = 2, \ldots, n$, where $F(x)$ is the cumulative distribution function of X_i, $i = 1, \ldots, n$, $f(x) = dF(x)/dx$, and $-1 \leqslant \alpha \leqslant 1$. The rank order probabilities do not depend on F but only on α. So without loss of generality, F may be taken as the uniform [0, 1] distribution, in which case the density becomes

$$f(x_i|x_{i-1}) = [1 + \alpha(2x_{i-1} - 1)\,(2x_i - 1)], \qquad 0 \leqslant x_{i-1}, x_i \leqslant 1\,.$$

The integrals are tractable, and the rank order probabilities can be computed exactly. This model is being investigated further.

---

[14] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, pp. 53–54.

[15] R. Aiyer, Ph.D. thesis, University of California, Berkeley, 1968.

[16] E. J. Gumbel, *J. Am. Statist. Assoc.* 55, 698–707 (1960).

RELIABILITY IN A THREE-UNIT REDUNDANT SYSTEM

The system under consideration satisfies the following conditions:

1. The system is composed of three identical units, each with identical constant failure rate λ.
2. The repair rate is the same for each unit and is equal to μ.
3. At any time t the system must be in one of the following states:

   State 3 – All three units are operable and operating.

   State 2 – One unit is in the failed state and the two other units are operable and operating.

   State 1 – Two of the units are in the failed state and the other unit is operable and operating.

   State 0 – All three units have failed.
4. Transitions from states 3, 2, or 1 are possible, but transition from state 0 is not possible (i.e., once the system is in state 0 it stays there).

Calabro[17] suggests the assumption that the time to failure has an exponential distribution with mean given by

$$\frac{11\lambda^2 + 7\lambda\mu + 2\mu^2}{6\lambda^3} .$$

This assumption leads to an approximate measure of reliability of the system given by

$$R(t) = 1 - \exp\left\{ -\frac{6\lambda^3 t}{11\lambda^2 + 7\lambda\mu + 2\mu^2} \right\} .$$

In the following discussion, it is shown that there is no need for the assumption.

From conditions 1 and 2 given above, the probability that a unit fails in the interval Δt is equal to $\lambda\,\Delta t$, and the probability that a unit is repaired in the interval Δt is equal to $\mu\,\Delta t$. Let $P_i(t)$, $i = 0, 1, 2, 3$, equal the probability that a unit is in state i at time t. It can be shown that the following equations are satisfied:

$$\begin{aligned} P_3(t+\Delta t) &= P_3(t)\,(1-\lambda\,\Delta t)^3 + P_2(t)\,(1-\lambda\,\Delta t)^2\mu\,\Delta t + P_1(t)\,(1-\lambda\,\Delta t)\,(\mu\,\Delta t)^2 \\ &= P_3(t)\,(1-3\lambda\,\Delta t) + P_2(t)\mu\,\Delta t + o(\Delta t)\,, \end{aligned} \tag{1}$$

$$\begin{aligned} P_2(t+\Delta t) &= P_3(t)\binom{3}{1}\lambda\,\Delta t(1-\lambda\,\Delta t)^2 + P_2(t)\left[(1-\lambda\,\Delta t)^2(1-\mu\,\Delta t) + \binom{2}{1}\lambda\,\Delta t\,\mu\,\Delta t(1-\lambda\,\Delta t)\right] \\ &\quad + P_1(t)\left[\binom{2}{1}\mu\,\Delta t(1-\mu\,\Delta t)(1-\lambda\,\Delta t) + \lambda\,\Delta t(\mu\,\Delta t)^2\right] \\ &= P_3(t)3\lambda\,\Delta t + P_2(t)\,[1-(\mu+2\lambda)\Delta t] + P_1(t)\,(2\mu\,\Delta t) + o(\Delta t)\,, \end{aligned} \tag{2}$$

[17] S. R. Calabro, *Reliability Principles and Practices*, McGraw-Hill, New York, 1962.

$$P_1(t+\Delta t) = P_3(t)\binom{3}{2}(\lambda\,\Delta t)^2(1-\lambda\,\Delta t) + P_2(t)\left[\binom{2}{1}\lambda\,\Delta t(1-\lambda\,\Delta t)(1-\mu\,\Delta t) + (\lambda\,\Delta t)^2\mu\,\Delta t\right]$$
$$+ P_1(t)\left[(1-\lambda\,\Delta t)(1-\mu\,\Delta t)^2 + \binom{2}{1}(1-\mu\,\Delta t)\mu\,\Delta t\,\lambda\,\Delta t\right]$$
$$= P_2(t)2\lambda\,\Delta t + P_1(t)\,[1-(\lambda+2\mu)\Delta t] + o(\Delta t)\,, \tag{3}$$

$$P_0(t+\Delta t) = P_3(t)\,(\lambda\,\Delta t)^3 + P_2(t)\,(1-\mu\,\Delta t)\,(\lambda\,\Delta t)^2 + P_1(t)\,(1-\mu\,\Delta t)^2\,\lambda\,\Delta t + P_0(t)$$
$$= P_1(t)\lambda\,\Delta t + P_0(t) + o(\Delta t)\,. \tag{4}$$

Using the definition of the derivative of a function, that is,

$$\frac{dP_i(t)}{dt} = P_i'(t) = \lim_{\Delta t\to 0}\frac{P_i(t+\Delta t) - P_i(t)}{\Delta t}\,, \qquad i = 0, 1, 2, 3, \tag{5}$$

Eqs. (1) to (4) now become

$$P_3'(t) = \mu P_2(t) - 3\lambda P_3(t)\,, \tag{6}$$

$$P_2'(t) = 2\mu P_1(t) - (\mu + 2\lambda)P_2(t) + 3\lambda P_3(t)\,, \tag{7}$$

$$P_1'(t) = -(\lambda + 2\mu)P_1(t) + 2\lambda P_2(t)\,, \tag{8}$$

$$P_0'(t) = \lambda P_1(t)\,. \tag{9}$$

Since

$$\sum_{i=0}^{3} P_i(t) = 1$$

for all t, there are only three independent equations in (6) to (9). Without loss of generality Eqs. (6) to (8) are used to solve for $P_1(t), P_2(t), P_3(t)$; $P_0(t)$ is found from $P_0(t) = 1 - P_1(t) - P_2(t) - P_3(t)$. By using Laplace transforms the solutions to Eqs. (6) to (8) are given by

$$P_1(t) = 6\lambda^2\left[\frac{e^{\rho_1 t}}{(\rho_1-\rho_2)(\rho_1-\rho_3)} + \frac{e^{\rho_2 t}}{(\rho_2-\rho_1)(\rho_2-\rho_3)} + \frac{e^{\rho_3 t}}{(\rho_3-\rho_1)(\rho_3-\rho_2)}\right], \tag{10}$$

$$P_2(t) = 3\lambda\left[\frac{(\lambda+2\mu+\rho_1)e^{\rho_1 t}}{(\rho_1-\rho_2)(\rho_1-\rho_3)} + \frac{(\lambda+2\mu+\rho_2)e^{\rho_2 t}}{(\rho_2-\rho_1)(\rho_2-\rho_3)} + \frac{(\lambda+2\mu+\rho_3)e^{\rho_3 t}}{(\rho_3-\rho_1)(\rho_3-\rho_2)}\right], \tag{11}$$

$$P_3(t) = \frac{(\rho_1-s_1)(\rho_1-s_2)}{(\rho_1-\rho_2)(\rho_1-\rho_3)}e^{\rho_1 t} + \frac{(\rho_2-s_1)(\rho_2-s_2)}{(\rho_2-\rho_1)(\rho_2-\rho_3)}e^{\rho_2 t} + \frac{(\rho_3-s_1)(\rho_3-s_2)}{(\rho_3-\rho_1)(\rho_3-\rho_2)}e^{\rho_3 t}\,, \tag{12}$$

where ρ_1, ρ_2, ρ_3 are the characteristic roots, which are assumed to be distinct, of the matrix

$$\begin{pmatrix} -3\lambda & \mu & 0 \\ 3\lambda & -(\mu + 2\lambda) & 2\mu \\ 0 & 2\lambda & -(\lambda + \mu) \end{pmatrix}$$

and $(\rho - s_1)(\rho - s_2) = \rho^2 + 3(\lambda + \mu)\rho + 2\mu^2 + 2\lambda^2 + \lambda\mu$. In addition,

$$P_0(t) = 1 + k_1 e^{\rho_1 t} + k_2 e^{\rho_2 t} + k_3 e^{\rho_3 t}, \tag{13}$$

where

$$k_1 = -\frac{\rho_1^2 + (6\lambda + 3\mu)\rho_1 + 2\mu^2 + 11\lambda^2 + 7\lambda\mu}{(\rho_1 - \rho_2)(\rho_1 - \rho_3)}, \tag{14}$$

$$k_2 = -\frac{\rho_2^2 + (6\lambda + 3\mu)\rho_2 + 2\mu^2 + 11\lambda^2 + 7\lambda\mu}{(\rho_2 - \rho_1)(\rho_2 - \rho_3)}, \tag{15}$$

$$k_3 = -\frac{\rho_3^2 + (6\lambda + 3\mu)\rho_3 + 2\mu^2 + 11\lambda^2 + 7\lambda\mu}{(\rho_3 - \rho_1)(\rho_3 - \rho_2)}. \tag{16}$$

Since $P_0(t)$ represents the probability of the system being in state 0 at time t, that is, the probability of failure at time t, the expected time to failure is given by

$$\int_0^\infty t\, dP_0(t) = \frac{k_1}{\rho_1} + \frac{k_2}{\rho_2} + \frac{k_3}{\rho_3}. \tag{17}$$

Using the expressions for k_1, k_2, k_3 given in (14) to (16), it can be shown that (17) reduces to

$$\int_0^\infty t\, dP_0(t) = \frac{11\lambda^2 + 7\lambda\mu + 2\mu^2}{6\lambda^3}. \tag{18}$$

Since the reliability of this system, which is equal to $1 - P_0(t)$, involves the computation of the eigenvalues ρ_1, ρ_2, ρ_3, it should be pointed out that there are standard techniques available for finding them.

THE INVERSE OF A MATRIX OCCURRING IN FIRST-ORDER MOVING-AVERAGE MODELS

Let $\{\epsilon_t\}$ be a sequence of independent identically distributed normal variates with zero mean and unit variance. The model

$$X_t = \epsilon_t + \alpha\epsilon_{t-1}, \qquad t = 1, 2, \ldots; |\alpha| < 1,$$

is referred to as a first-order moving-average process. Given a set of observations $x_1, x_2, \ldots, x_n$, the maximum likelihood estimate of the parameter α depends on the inverse of the variance-covariance matrix of $x_1, x_2 \ldots, x_n$. This covariance matrix, which is denoted by V_n^{-1}, is given by the $n \times n$ tridiagonal matrix

$$V_n^{-1} = \begin{pmatrix} 1+\alpha^2 & \alpha & 0 & 0 & \ldots & 0 \\ \alpha & 1+\alpha^2 & \alpha & 0 & \ldots & 0 \\ 0 & \alpha & 1+\alpha^2 & \alpha & \ldots & 0 \\ . & . & . & . & \ldots & . \\ 0 & . & . & . & 1+\alpha^2 & \alpha \\ 0 & . & . & . & \alpha & 1+\alpha^2 \end{pmatrix} .$$

Durbin[18] in 1959 and Hannan[19] in 1960 gave an approximation to the inverse of the above tridiagonal matrix as

$$V_n = \frac{1}{1-\alpha^2} \begin{pmatrix} 1 & -\alpha & \alpha^2 & \ldots & (-\alpha)^{n-1} \\ -\alpha & 1 & -\alpha & \ldots & (-\alpha)^{n-2} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ (-\alpha)^{n-1} & (-\alpha)^{n-2} & (-\alpha)^{n-3} & \ldots & 1 \end{pmatrix} .$$

We obtained the exact form for V_n invoking a theorem due to Guttman[20] and Ukita[21] which gives a necessary and sufficient condition for a tridiagonal symmetric matrix to have an inverse.

Theorem (Guttman-Ukita): Let $V = (v_{ij})$ be a symmetric, invertible $n \times n$ matrix. A necessary and sufficient condition for V^{-1} to be a tridiagonal matrix is that each element of the original symmetric matrix in the jth row (starting at the main diagonal and proceeding to the right) must have a constant relation with the corresponding element in the first row. In other words,

$$v_{2k} = \lambda_{12} v_{1k}, \quad 2 \leqslant k \leqslant n ,$$
$$v_{3k} = \lambda_{13} v_{1k}, \quad 3 \leqslant k \leqslant n ,$$
$$\vdots$$
$$v_{jk} = \lambda_{1j} v_{1k}, \quad j \leqslant k \leqslant n ,$$

and

$$v_{nk} = \lambda_{1n} v_{1k}, \quad k = n .$$

From $V_n^{-1} = (u_{ij})$, in order to find V_n, it is apparent that it is enough to obtain the first row, $v_{11}, v_{12}, \ldots, v_{1n}$, and $\lambda_{12}, \lambda_{13}, \ldots, \lambda_{1n}$. Then by using the above theorem and the symmetry of V_n, we can

[18] J. Durbin, *Biometrika* 46, 306–16 (1959).

[19] E. J. Hannan, *Time Series Analysis,* Methuen, London, and Wiley, New York, 1960.

[20] L. Guttman, *Psychometrika* 20, 173–95 (1955).

[21] Y. Ukita, *J. Hokkaido College of Art and Literature* 6, 66–75 (1955).

completely determine V_n. The first row of V_n can easily be obtained since V_n^{-1} is a tridiagonal matrix, and the cofactors of u_{1j} can easily be written down. Since it is just as simple to write down the cofactors of the diagonal elements u_{ii}, we can then find $\lambda_{12}, \lambda_{13}, \ldots, \lambda_{1n}$.

Let us denote $1 + \alpha^2$ by β and redefine the tridiagonal matrix V_n^{-1} by

$$u_{ii} = \beta \ ,$$

$$u_{ji} = u_{ij} = 0 \qquad \text{if } |i - j| \geqslant 2 \ ,$$

and

$$u_{ji} = u_{ij} = \alpha \qquad \text{if } |i - j| = 1 \qquad \text{for } i, j = 1, 2, \ldots, n \ .$$

Let δ_n denote the value of the determinant of V_n^{-1}. Expanding δ_n by the elements of its first row, we can obtain the recursive relation

$$\delta_n = \beta\delta_{n-1} - \alpha^2\delta_{n-2} \ , \qquad n = 2, 3, \ldots ,$$

with the definition $\delta_0 = 1$ and $\delta_1 = \beta$.

In fact, one can even solve this for δ_n and obtain the closed form

$$\delta_n = \sum_{k=0}^{[(n+2)/2]} (-1)^k \binom{n-k}{k} \beta^{n-2k} \alpha^{2k} \ ,$$

where $[x]$ denotes the integral part of x.

Now the cofactors of the first row of the tridiagonal matrix V_n^{-1} can easily be obtained as

$$v_{1j} = \frac{(-1)^{j-1} \alpha^{j-1} \delta_{n-j}}{\delta_n} \ , \qquad 1 \leqslant j \leqslant n \ .$$

Therefore

$$\lambda_{1j} = \frac{v_{jj}}{v_{1j}} = \frac{(-1)^{j-1} \delta_{j-1}}{\alpha^{j-1}} \ , \qquad 1 \leqslant j \leqslant n \ .$$

AN INVERSION METHOD FOR BAND MATRICES

In a recent paper, Kounias[22] applied the following fact in finding the inverse of certain patterned matrices. Let A^{-1} be an invertible $r \times r$ matrix and consider the system of linear equations

[22] E. G. Kounias, *J. Math. Anal. Appl.* 21, 695–98.

$$A^{-1}\begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_r \end{pmatrix} = \begin{pmatrix} w \\ w^2 \\ \vdots \\ w^r \end{pmatrix} \tag{1}$$

where $Z_1, Z_2, \ldots, Z_r$ and w are auxiliary variables. It is clear that if we can find $Z_i = f(w, i), i = 1, 2, \ldots, r$, and expand $f(w, i)$ in any region of w where the expansion is possible, then the coefficient of w^j will be the (i, j)th element of A. It turns out that this can be easily done when we have matrices which have associated lower-order difference equations subject to boundary conditions.

Kounias illustrated this in the case of the following $r \times r$ band matrix:

$$A^{-1} = \begin{pmatrix} a & b & 0 & 0 & \ldots & 0 & 0 \\ c & a & b & 0 & \ldots & 0 & 0 \\ 0 & c & a & b & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & & c & a \end{pmatrix} \tag{2}$$

with the associated difference equation

$$cZ_{i-1} + aZ_i + bZ_{i+1} = w^i, \qquad i = 1, 2, \ldots, \tag{3}$$

and with the boundary conditions

$$Z_0 = 0 \quad \text{and} \quad Z_{r+1} = 0 . \tag{4}$$

We solved the difference equation (3) by the following method and obtained Z_i as a polynomial of rth degree in w. This is different in analytical form from that given by Kounias, where Z_i has a power series type representation in w.

In order to solve the difference equation (3) we will first express it in the following vector notation. It is equivalent to

$$\begin{aligned} Z_{i+1} &= -\frac{a}{b} Z_i - \frac{c}{b} Z_{i-1} + \frac{1}{b} w^i , \qquad b \neq 0 , \qquad i = 1, 2, \ldots \\ &= \alpha Z_i + \beta Z_{i-1} + (w^i/b) , \end{aligned}$$

where $\alpha = (a/b)$ and $\beta = -(c/b)$. Therefore

$$\begin{pmatrix} Z_{i+1} \\ Z_i \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ 1 & 0 \end{pmatrix} \begin{pmatrix} Z_i \\ Z_{i-1} \end{pmatrix} + \begin{pmatrix} w^i/b \\ 0 \end{pmatrix}$$

or

$$\nu_i = E\,\nu_{i-1} + \xi_{i-1}\,, \qquad i = 1, 2, \ldots, \tag{5}$$

where

$$\nu_i = \begin{pmatrix} Z_{i+1} \\ Z_i \end{pmatrix}, \qquad E = \begin{pmatrix} \alpha & \beta \\ 1 & 0 \end{pmatrix}, \qquad \text{and} \quad \xi_{i-1} = \begin{pmatrix} w^i/b \\ 0 \end{pmatrix}.$$

The solution of (5) is given by

$$\nu_i = E^i \nu_0 + E^{i-1}\,\xi_0 + E^{i-2}\xi_1 + \ldots + E\xi_{i-2} + \xi_{i-1}\,, \qquad i = 1, 2, \ldots, \tag{6}$$

where

$$\nu_0 = \begin{pmatrix} Z_1 \\ Z_0 \end{pmatrix} = \begin{pmatrix} Z_1 \\ 0 \end{pmatrix} \quad \text{and} \quad \xi_0 = \begin{pmatrix} w/b \\ 0 \end{pmatrix}.$$

If we multiply both sides of (6) by the row vector (0, 1) we obtain

$$Z_i = e_{21}^{(i)}\,Z_1 + e_{21}^{(i-1)}\frac{w}{b} + e_{21}^{(i-2)}\frac{w^2}{b} + \ldots + e_{21}^{(1)}\frac{w^{i-1}}{b}, \qquad i = 2, 3, \ldots, \tag{7}$$

where $e_{21}^{(i)}$ represents the second-row, first-column element of the ith power of the 2×2 matrix E. From (7) and $Z_{r+1} = 0$, we can first solve for Z_1 and then obtain $Z_2, Z_3, \ldots, Z_r$. Thus, to summarize, the inverse of the matrix A^{-1} is given by

$$be^{(r+1)}A =$$

$$\begin{pmatrix} -e^{(r)} & -e^{(r-1)} & \ldots & -e^{(2)} & -e^{(1)} \\ e^{(1)}e^{(r+1)} - e^{(2)}e^{(r)} & -e^{(2)}e^{(r-1)} & \ldots & -e^{(2)}e^{(2)} & -e^{(2)}e^{(1)} \\ e^{(2)}e^{(r+1)} - e^{(3)}e^{(r)} & e^{(1)}e^{(r+1)} - e^{(e)}e^{(r-1)} & \ldots & -e^{(3)}e^{(r-2)} & -e^{(3)}e^{(1)} \\ \vdots & \vdots & & & \\ e^{(r-1)}e^{(r+1)} - e^{(r)}e^{(r-1)} & e^{(r-1)}e^{(r+1)} - e^{(r)}e^{(r-1)} & & & -e^{(r)}e^{(1)} \end{pmatrix} \tag{8}$$

where $e^{(i)} = e_{21}^{(i)}$ is the second-row, first-column element of the ith power of the matrix E, namely,

$$E = \begin{pmatrix} \alpha & \beta \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -(a/b) & -(c/b) \\ 1 & 0 \end{pmatrix}.$$

Since the formula (8) for the inverse of the matrix A^{-1} involves the elements of the higher powers of a 2×2 matrix E, we shall now give a formula for the rth power of an arbitrary 2×2 matrix. Since it is very easy to find the characteristic roots of a 2×2 matrix from the trace and the determinant of the matrix, we can easily prove the following:

Lemma: Let

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$$

be an arbitrary 2×2 matrix with eigenvalues λ_1 and λ_2. Then

$$M^r = \begin{cases} \dfrac{\lambda_1^r - \lambda_2^r}{\lambda_1 - \lambda_2} M - \dfrac{\lambda_1^{r-1} - \lambda_2^{r-1}}{\lambda_1 - \lambda_2} \lambda_1 \lambda_2 I & \text{if } \lambda_1 \neq \lambda_2 , \\ r \lambda^{r-1} M - (r-1)\lambda^r I & \text{if } \lambda_1 = \lambda_2 , \qquad r = 1, 2, \ldots . \end{cases}$$

UNIQUENESS OF A MINIMUM NORM SOLUTION

Let A be an $n \times p$ matrix, X a $p \times 1$ vector, and Y an $n \times 1$ vector. For a consistent equation, $AX = Y$, a general solution can be expressed in the form

$$X = GY + (H - I)Z ,$$

where Z is an arbitrary $p \times 1$ vector, $GA = H$, and G is any generalized inverse of A, that is, $AGA = A$; $(H - I)$ is a linear transformation which maps E^p onto a linear subspace of dimension equal to the rank of $(H - I)$. By adding GY the subspace is translated to the linear manifold $\beta = \{X : X = GY + (H - I)Z \text{ for } Z \in E^p\}$.

The existence and uniqueness of a minimum norm solution in β is trivial when the norm is the Euclidean norm,

$$||X|| = \sqrt{\sum_{i=1}^{p} |X|^2} ,$$

since the problem reduces to finding the distance (in E^p) from the origin to the linear manifold β. However, for an arbitrary norm on E^p does there exist a unique element of smallest norm? A sufficient condition on the norm is given for the uniqueness of a minimum norm solution since it will be shown that existence always holds in a finite-dimensional normed space.

Since a finite-dimensional normed space is complete and corresponding subspaces are closed, and since translation is a bicontinuous operation, β is closed. Let $d = \inf \{t : ||X|| = t \text{ and } X \in \beta\}$. There exists $Z \in C(0, d)$ such that Z is a limit point or an isolated point of $\beta [N(Z, \epsilon) \cap \beta \neq \phi$ for $\epsilon > 0]$, or otherwise for each $Z \in C(0, d)$ there exists $\epsilon_Z > 0$ such that $N(Z, \epsilon_Z) \cap \beta = \phi$. This forms an open covering for the compact set $C(0, d)$ and thus is reducible to a finite subcovering, $\{N(Z', \epsilon_{Z'})\}$. Hence, there exists a $\delta > 0$ such that $N(0, d + \delta) \cap \beta = \phi$, which implies that $\inf \{t : ||X|| = t \text{ and } X \in \beta\} \geqslant d + \delta > d$, and thus we have a contradiction. If W is an isolated point of β, then the proof is trivial. Let W be a limit point of β on $C(0, d)$; then there exists a sequence in β, say X_n, such that $||X_n - W||$ converges to zero. Since the space is complete and β is closed, $W \in \beta$ and W is of minimum norm, that is, $||W|| = d$. Therefore β being closed is sufficient for the existence of a minimum norm solution.

Since a linear subspace is convex, β, being a translate of a subspace, is also convex. We say that a normed space is rotund if every point of $C(0, r)$ is an extreme point, that is, $C(0, r)$ meets no line in three

points. An equivalent definition can be that a space is rotund iff $a \neq b$, and $||a|| = ||b|| \leqslant r$ implies that $||ta + (1-t)b|| < r$ for $0 < t < 1$.

In a finite-dimensional rotund normed space every closed convex subset has a unique element of minimum norm. The set being closed provides the existence. Suppose there exist two distinct elements, x and y, of smallest norm in a convex set S. Then $\frac{1}{2}x + \frac{1}{2}y \in S$, and

$$||\tfrac{1}{2}x + \tfrac{1}{2}y|| \leqslant \tfrac{1}{2}\,||x|| + \tfrac{1}{2}\,||y|| = ||x|| \text{ or } = ||y||\,.$$

However, equality cannot hold since the space is rotund; hence a contradiction, and thus uniqueness is established.

The obvious question is whether a rotund norm is also a necessary condition for the uniqueness of a minimum norm solution. Example 1 gives a nonrotund space where there does not exist a unique minimum norm solution for a given matrix A, and Example 2 gives a nonrotund space where a unique norm solution does exist. Therefore rotundity of the norm is a sufficient condition for the uniqueness of a minimum norm solution but is not a necessary condition.

Example 1. $||X|| = |x_1| + |x_2|$

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, Z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

$$G = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, GY + (H - I)Z = \begin{pmatrix} y_1 + z_2 \\ -z_2 \end{pmatrix}$$

$AX = Y$ is consistent iff $y_2 = 2y_1$. For

$$Y = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

the set of minimum norm solution is

$$\left\{X{:}X = \begin{pmatrix} t \\ 1-t \end{pmatrix} \text{ for } 0 \leqslant t \leqslant 1\right\}.$$

Example 2. $||X|| = |x_1| + |x_2|$

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, Z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

$$G = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, GY + (J - I)Z = \begin{pmatrix} y_1 + 2Z_2 \\ -Z_2 \end{pmatrix}$$

$AX = Y$ is consistent iff $y_2 = 2y_1$. For

$$Y = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

the unique minimum norm solution is

$$X = \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

ESTIMATION OF PARAMETERS IN MIXTURES OF DISTRIBUTIONS

A previous report[23] described the motivation and practical solution of investigations that yielded observations assumed to represent mixtures of Gaussian distributions. Since that time, analytical and Monte Carlo studies of biases and covariances of the estimates were conducted; the results of these were presented in a technical memorandum.[24]

The analytical study considered the mixtures of two normal distributions with equal, known variances and known sample sizes. The case of three or more distributions was deemed to be unfeasible for an analytic attack and was therefore investigated by Monte Carlo methods.

APPLICATION OF THE LOGISTIC MODEL TO BINOMIAL RESPONSE DATA

We consider an experimental situation in which groups of animals are subjected to various doses of radiation, and the percent mortality at the end of a given time period is recorded as a function of dose. When this function is plotted, it very often takes the form of an S-shaped *dose-response* curve, various properties of which (such as the LD50) are of interest to the experimenter.

The three models most commonly used for fitting an empirical dose-response curve are: (1) the integrated normal model (associated with the probit transformation), (2) the angular model (associated with the arc sine transformation), and (3) the logistic model (associated with the logit transformation). We shall consider here only the last of these, the logistic model, which is based on the assumption that P_i, the probability of response (death) at the ith dose x_i, is related to x_i through the *logistic function*

$$P_i = [1 + \exp(-\alpha - \beta x_i)]^{-1} . \tag{1}$$

From (1), we obtain

$$L_i = L(P_i) = \ln\left(\frac{P_i}{1 - P_i}\right) = \alpha + \beta x_i ,$$

where $L(P)$ is known as the *logit transformation.*

Our present interest is centered about unifying and extending the results already in the literature concerning the properties of the logistic function and the logit transformation, estimation of the parameters involved, and the analysis of experiments when the logistic model is assumed.

In working with the logistic model, it is convenient to consider the "logits" $\{l_i\}$, where $l_i = \ln(p_i/q_i)$ and p_i $(= 1 - q_i)$ is the observed proportion of subjects responding in the ith group. It is assumed that the number $(n_i p_i)$ responding in the ith group follows a binomial distribution with expectation $n_i P_i$, where P_i is

[23] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4326, p. 47.

[24] D. G. Gosslee and K. O. Bowman, *Evaluation of Maximum Likelihood Estimates of Parameters in Mixtures of Normal Distributions,* ORNL-TM-2110 (1968).

given by (1). Under these assumptions, it can be shown that l_i is asymptotically normally distributed with mean $L_i = \ln [P_i/(1 - P_i)]$ and variance $[1/n_i P_i(1 - P_i)]$.

Given a set of observed responses, how are the parameters α and β to be estimated? Berkson[25] has proposed and vigorously advocated a modification of the minimum Pearson χ^2 estimation procedure, which he calls *minimum logit* χ^2 *estimation* and which requires α and β to be chosen so that

$$\text{logit } \chi^2 = \sum_{i=1}^{k} n_i p_i q_i (l_i - \alpha - \beta x_i)^2 \tag{2}$$

is minimized. Other authors, notably Cornfield,[26] have proposed iterative methods for obtaining the *maximum likelihood estimates* of α and β. The question of whether it is "better" to use the minimum logit χ^2 or the maximum likelihood estimators has been the subject of much controversy in the literature and is still an open question. The answer seems to depend not only on one's definition of "better" but also on the sample sizes, number of doses, and the unknown true values of α and β. We have begun to look into this question ourselves through an extensive Monte Carlo study.

There is an interesting relationship between the equations used in Cornfield's maximum likelihood procedure and the minimum logit χ^2 estimation procedure. This can be shown by rewriting Cornfield's equations and taking the "initial values" of the $\{P_i\}$ to be the *observed* p_i. The resulting equations are identical to Berkson's equations for the minimum logit χ^2 estimates. In a sense, then, the minimum logit χ^2 estimates can be thought of as the result of the first iteration in the Cornfield maximum likelihood procedure. One particular advantage of this approach is that Cornfield's iterative method, whose convergence properties are generally sensitive to the choice of "starting point," has, in our experience, always converged when required to pass through the minimum logit χ^2 estimates.

This iterative procedure for maximum likelihood estimation can be extended to the general linear logistic model:

$$\ln [P_u/(1 - P_u)] = \sum_{i=1}^{s} f_{ui}\theta_i , \qquad u = 1, 2, \ldots, N , \tag{3}$$

in which there are s parameters to be estimated from the data. The $\{f_{ui}\}$ are known constants which depend on the levels of the set of independent variables under investigation. In matrix notation, the appropriate equations for the iterative solution θ take the form

$$(\mathbf{F}'\mathbf{W}^{(0)}\mathbf{F})\,(\theta - \theta^{(0)}) = \mathbf{F}'\mathbf{d}^{(0)} , \tag{4}$$

where $\mathbf{F} = \{f_{ui}\}$ is an $N \times s$ matrox of constants, $\mathbf{W}^{(0)}$is a diagonal matrix of temporary "weights": $W_{uu}^{(0)} = n_u P_u^{(0)}(1 - P_u^{(0)})$, and $d_u^{(0)} = n_u(p_u - P_u^{(0)})$, $u = 1, 2, \ldots, N$. Given $\theta^{(0)}$, a temporary estimate of θ, we can calculate $\mathbf{d}^{(0)}$ and $\mathbf{W}^{(0)}$ and solve (4) for θ. (We assume that $\mathbf{F}$ is such that $\mathbf{F}'\mathbf{W}^{(0)}\mathbf{F}$ is invertible. If it is not, a reparameterization can be made in (3) which will result in a nonsingular $\mathbf{F}'\mathbf{W}^{(0)}\mathbf{F}$ matrix.) The value of θ thus obtained then becomes the temporary estimate for the next iteration.

---

[25] J. Berkson, *J. Am. Statist. Assoc.* 39, 357–65 (1944).

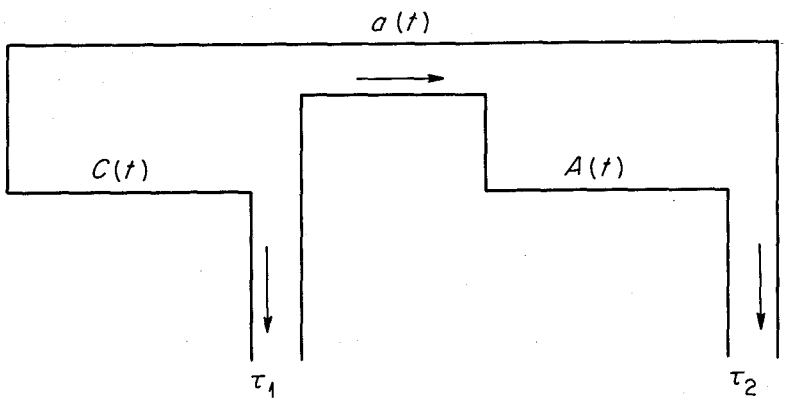
[26] J. Cornfield, *Statistics and Mathematics in Biology,* pp. 327–44, Iowa State College Press, Ames, 1954.

It can be shown that the equations given by Cornfield for $s = 2$ and by Dyke and Patterson[27] when the $\{\theta_i\}$ represent factorial effects, can be derived from (4). The equation (4) can also be rewritten in such a way that when the observed $\{p_u\}$ are substituted for initial values $\{P_u^{(0)}\}$, the minimum logit χ^2 estimates are obtained at the end of the first iteration, as described above for the case $s = 2$.

A further result of this work connects the inverse of the matrix $(\mathbf{F}'\mathbf{W}^{(0)}\mathbf{F})$ at the final iteration with the maximum likelihood estimate of the Fisher information matrix, thus justifying its use as an approximate variance-covariance matrix for the maximum likelihood estimate of θ.

KINETICS OF ANTIBODY PRODUCTION

Consider the following diagram which is used to describe the kinetics of antibody production:



In the above diagram $C(t)$ represents the number of cells available at time t, $A(t)$ represents the amount of antibody present at time t, and $a(t)$ represents the rate at which the cells are producing antibody. The constants τ_1 and τ_2 are associated with the rate at which cells are being depleted and the rate at which antibody is being produced (or the catabolic rate) respectively. The following system of differential equations is assumed to describe this model:

$$\frac{dC(t)}{dt} = -\tau_1 C(t) , \tag{1}$$

$$\frac{dA(t)}{dt} = a(t) C(t) - \tau_2 A(t) , \tag{2}$$

$$\frac{da(t)}{dt} = \tau_1 C(t) . \tag{3}$$

It can be shown for the case when $\tau_1 \neq \tau_2$ that the solutions to Eqs. (1) to (3) are given by

$$C(t) = C(0)\, e^{-\tau_1 t} , \tag{4}$$

[27] G. V. Dyke and H. D. Patterson, *Biometrics* 8, 1–12 (1952).

$$A(t) = \left[\frac{a(0)\,C(0)}{(\tau_2 - \tau_1)} + \frac{C^2(0)}{(\tau_2 - \tau_1)} \right] e^{-\tau_1 t} + \frac{C^2(0)}{(2\tau_1 - \tau_2)}\; e^{-2\tau_1 t} + K e^{-\tau_2 t}\,, \tag{5}$$

where

$$K = A(0) + \frac{a(0)\,C(0)}{(\tau_1 - \tau_2)} + \frac{C^2(0)}{(\tau_1 - \tau_2)} - \frac{C^2(0)}{(2\tau_1 - \tau_2)}$$

and

$$a(t) = a(0) + C(0) - C(0)\, e^{-\tau_1 t}\,. \tag{6}$$

For the case when $\tau_1 = \tau_2 = \tau$, the solutions of (1) to (3) are given by

$$C(t) = C(0)\, e^{-\tau t}, \tag{7}$$

$$A(t) = \left[\, a(0)\, C(0) + C^2(0) \,\right] t e^{-\tau t} + \frac{C^2(0)}{\tau}\, e^{-2\tau t} + K' e^{-\tau t}\,, \tag{8}$$

where

$$K' = A(0) - \frac{C^2(0)}{\tau}$$

and

$$a(t) = a(0) + C(0) - C(0) e^{-\tau t}\,. \tag{9}$$

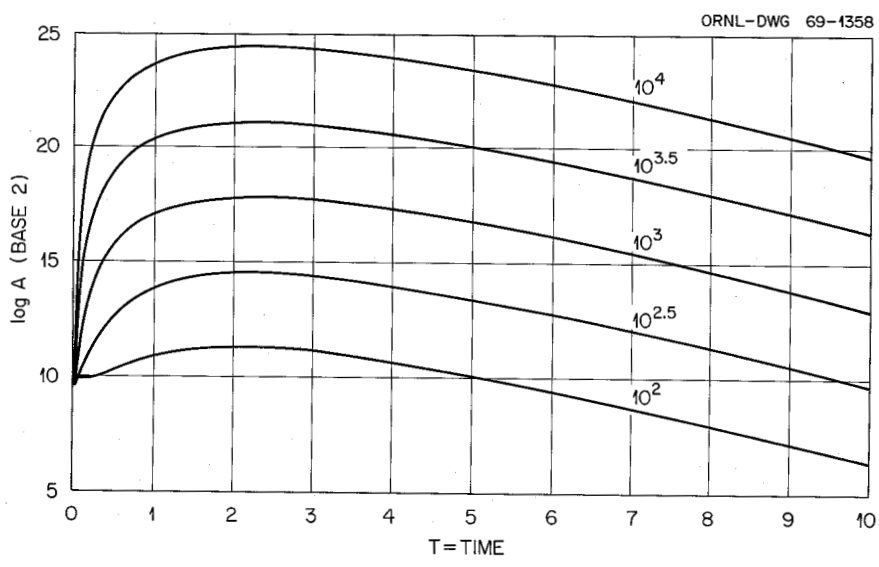


Fig. 20. Antibody Production Model $C(0)$ = 102(0.5)4 $a(0)$ = 0.25, TAU = LN2.**

By choosing various values of $C(0)$, $A(0)$, $a(0)$, τ_1, and τ_2 (or τ), it is possible to plot the solutions (4) to (6) or (7) to (9) in order to see how well the hypothesized model describes the physical situation. Figure 20 shows some representative plots of $\log_2 A(t)$ for different values of $C(0)$ with $\tau = \ln 2$ and $a(0) = 0.25$. Investigation is continuing in this direction.

ESTIMATION OF THE MEAN AND VARIANCE OF A CENSORED WEIBULL MEASUREMENT DISTRIBUTION

Shipments of a certain product arrive with an unknown amount of impurities in them. Before a shipment can be accepted, the mean amount of impurities in the shipment must be estimated, to see if it is less than an acceptable value.

The distribution of impurities in this product is known to be highly skewed. So to estimate the mean and variance of the amount of impurities in the shipment, a highly flexible skewed distribution for the impurities must be assumed. A distribution which fits this requirement is the two-parameter Weibull distribution, which is given by

$$F(m) = 1 - e^{-m^{\gamma}/\theta}, \qquad m \geqslant 0, \theta, \gamma > 0,$$

with mean

$$\mu = \theta^{1/\gamma}\,\Gamma(1 + 1/\gamma)$$

and variance

$$\sigma^2 = \theta^{2/\gamma}\{\Gamma(1 + 2/\gamma) - [\Gamma(1 + 1/\gamma)]^2\}.$$

In the situation considered here the problem is further complicated by the fact that the measuring devices on hand can only measure the amount of impurities which are actually greater than a fixed value M.

If N samples from the shipment are submitted for measurement and $N - R$ values are recorded, while R values can be specified only as less than M ($R < N$), the values of θ and γ can be estimated by the method of maximum likelihood.

The likelihood function to be maximized is given by

$$e^L = [1 - e^{-(M^{\gamma}/\theta)}]^R \prod_{i=R+1}^{N} \left[\left(\frac{\gamma}{\theta}\right) m_i^{\gamma-1}\, e^{(m_i^{\gamma}/\theta)}\right],$$

where m_i, $i = R + 1, \ldots, N$, are the $N - R$ measurements which were obtainable by the measuring device.

Although the first partial derivatives of L with respect to θ and γ are nonlinear, they are such that a solution for θ can be found in terms of γ.

Thus following a first guess γ_0 for $\hat{\gamma}$, a new guess for $\hat{\gamma}$ can be calculated by using

$$\gamma = \gamma_0 - F(\gamma_0) \Big/ \left\{ \frac{\partial F(\gamma)}{\partial \gamma} \bigg|_{\gamma_0} \right\},$$

where

$$F(\gamma) = [1 - e^{-M^\gamma/\theta(\gamma)}] \left[RM^\gamma + \sum_{i=R+1}^{N} m_i^\gamma - (N-R)\,\theta(\gamma) \right] - RM^\gamma$$

and

$$\theta(\gamma) = \frac{\sum_{i=R+1}^{N} m_i^\gamma (\ln m_i - \ln M)}{(N-R)/\gamma + \sum_{i=R+1}^{N} (\ln m_i - \ln M)} .$$

Following the convergence of the iterative procedure, the mean and variance of the distribution of impurities can be estimated by

$$\hat{\mu} = \hat{\theta}^{1/\hat{\gamma}}\, \Gamma(1 + 1/\hat{\gamma})$$

and

$$\hat{\theta}^2 = \hat{\theta}^{2/\hat{\gamma}}\, \{\Gamma(1 + 2/\hat{\gamma}) - [\Gamma(1 + 1/\hat{\gamma})]^2\}$$

respectively.

The solution for the case when the R largest values are censored is given by Cohen.[28]

NUMBERS GENERATED BY THE FUNCTION exp $(1 - e^x)$

A probability distribution associated with Stirling numbers of the second kind was studied before.[29] This led to a sequence of new numbers, referred to as the C numbers, which have properties similar to those of Bell numbers. The first few C_i, $i = 1, 2, \ldots$, are given by $C_0 = 1, C_1 = -1, C_2 = 0, C_3 = 1, C_4 = 1, C_5 = -2, C_6 = -9, C_7 = -9, C_8 = 50$. It turns out that the generating function of $\{C_n, n = 0, 1, 2, \ldots\}$ is given by exp $(1 - e^x)$. The $\{C_n\}$ is defined in an ad-hoc manner in terms of its generating function, and the following propositions are proved.

Definition: The sequence $\{C_n, n = 0, 1, 2, \ldots\}$ is defined by its exponential generating function,

$$\sum_{k=0}^{\infty} C_k \frac{x^k}{k!} = \exp(1 - e^x) .$$

A series representation for C_k is given by:

Proposition 1:

$$C_k = \exp\left[\sum_{r=0}^{\infty} (-1)^r \frac{r^k}{r!} \right], \qquad k = 0, 1, 2, \ldots .$$

[28] A. C. Cohen, *Technometrics* 7, 579–88 (1965).

[29] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, pp. 65–67.

A relationship between C_k and the Stirling numbers of the second kind σ_k^j is given by:
Proposition 2:

$$C_k = \sum_{j=1}^{k} (-1)^j \sigma_k^j .$$

Using proposition 1, we can obtain a recursive relation between the C numbers as:
Proposition 3:

$$C_{k+1} = -\sum_{j=0}^{k} \binom{k}{j} C_j , \qquad k = 0, 1, 2, \ldots , \quad C_0 = 1 .$$

In the next proposition, it is shown that C_n satisfies an nth order difference equation.
Proposition 4:

$$\Delta^n C_1 = \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j} C_{j+1} = -C_n , \qquad n = 1, 2, \ldots .$$

This can be used in computing $C_1, C_2, \ldots, C_n$ for small values of n.

In the next three propositions, some relations between $\{C_n\}$ and the Bell numbers $\{B_n\}$ are derived.
Proposition 5:

$$\sum_{k=0}^{n} \binom{n}{k} B_k C_{n-k} = 0 , \qquad n = 1, 2, \ldots , \text{ with } B_0 = C_0 = 1 .$$

Proposition 6:

$$\sum_{j=0}^{n} \binom{n}{j} C_j B_{n+1-j} = 1 , \qquad n = 0, 1, 2, \ldots .$$

Proposition 7:

$$\sum_{j=0}^{n} \binom{n}{j} B_j C_{n+1-j} = -1 , \qquad n = 0, 1, 2, \ldots .$$

In the next three propositions three determinantal representations for C_n are given.

Proposition 8:

$$\frac{C_n}{n!} = (-1)^n \begin{vmatrix} \frac{B_1}{1!} & B_0 & 0 & 0 & 0 \\ \frac{B_2}{2!} & \frac{B_1}{1!} & B_0 & 0 & 0 \\ \frac{B_3}{3!} & \frac{B_2}{2!} & \frac{B_1}{1!} & B_0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{B_n}{n!} & \frac{B_{n-1}}{(n-1)!} & \frac{B_{n-2}}{(n-2)!} & \cdots & \frac{B_1}{1!} \end{vmatrix}$$

Proposition 9:

$$C_{n+1} = (-1)^n \begin{vmatrix} 1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 1 & 1 & 0 & \cdots & 0 \\ 1 & 3 & 3 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ \binom{n}{0} & \binom{n}{1} & \binom{n}{2} & \cdot & \cdot & \cdots & \binom{n}{n} \end{vmatrix}$$

Proposition 10:

$$C_{n+1} = (-1)^{n+1} \begin{vmatrix} 1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{1!} & 1 & 2 & 0 & 0 & \cdots & 0 \\ \frac{1}{2!} & \frac{1}{1!} & 1 & 3 & 0 & \cdots & 0 \\ \frac{1}{3!} & \frac{1}{2!} & \frac{1}{1!} & 1 & 4 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{n!} & \frac{1}{(n-1)!} & \frac{1}{(n-2)!} & \cdot & \cdot & \cdots & 1 \end{vmatrix}$$

A GENERALIZED MULTIDIMENSIONAL AGE-DEPENDENT BRANCHING PROCESS

In the classical age-dependent branching process of Bellman and Harris,[30] the death and the production of offspring by an individual coincide. For many biological applications, however, it is necessary to consider a model in which an individual may give birth to offspring at random points of time during his life-span. A generalized multidimensional age-dependent branching process is a stochastic model for the growth of a population of particles (individuals) which may produce new particles at random points during their life-span. Each individual may belong to one of a finite number of types. Such processes have applications in both biology and physics.

A stochastic model of a generalized multidimensional age-dependent branching process with applications to the dynamics of cell populations and population genetics is described in an ORNL report.[31]

Statistical Applications

ANALYTICAL CHEMISTRY

An Experimental Optimization Program for the Chemical Assay of tRNA

An experimental optimization program which utilizes statistical experimental design and response surface techniques has been applied to the assay of transfer ribonucleic acids (tRNA). The assay procedure is based on the aminoacylation reaction first described by Hoagland, Zamecnik, and Stephenson.[32] This procedure, which is discussed in some detail by Rubin, Kelmers, and Goldstein,[33] essentially involves the incubation of the tRNA with a ^{14}C-labeled amino acid and other ingredients to form, among other things, a tRNA–amino acid ester linkage. This aminoacylated tRNA is then precipitated by acid, excess amino acid washed away, and the precipitate counted.

The object of the optimization program is to determine which combination of ingredients yields the most complete reaction for a fixed amount of tRNA. The effects of the following 11 variables were considered:

1. pH of the solution,
2. type of monovalent metal ions (potassium or ammonium),
3. concentration of potassium or ammonium ions,
4. type of reducing agent (ME or GSH),
5. concentration of ME or GSH,
6. time of incubation,
7. buffer type (Na cacodylate or Na glycinate),
8. concentration of labeled amino acid,

---

[30] T. E. Harris, *The Theory of Branching Process*, Springer-Verlag, 1963.

[31] C. J. Mode, *A Generalized Multi-Dimensional Age-Dependent Branching Process*, ORNL-4331 (November 1968).

[32] M. B. Hoagland, P. C. Zamecnik, and M. L. Stephenson, *Biochim. Biophys. Acta* 24, 215 (1957).

[33] I. B. Rubin, A. D. Kelmers, and G. Goldstein, *Anal. Biochem.* 20, 533–44 (1967).

9. concentration of ATP,
10. concentration of magnesium ions,
11. enzyme concentration.

The response corresponding to a particular combination of levels of these factors was taken to be the difference between the counts per minute obtained from the aminoacylated tRNA and the counts per minute obtained when a control sample (or blank) with no tRNA was run at the same factor levels.

Experiment 1. – The initial two-level design consisted of three parts:

a) A 2_{III}^{11-7} design (16 runs) chosen to provide unconfounded estimates of the main effects and two-factor interactions of the "major" variables and the main effects of the "minor" variables.

b) Replication of a half-fraction (8 runs) of the design described in *a*.

c) A set of eight "center points" in which the levels of the continuous variables were set midway between their low and high levels. The levels of the three qualitative variables at these eight points formed a replicated 2_{III}^{3-1} fractional factorial.

This experiment contained, therefore, 32 runs. The results indicated the presence of six large main effects, which pointed the way to the location of the next experiment.

Experiment 2. – On the basis of the results of experiment 1, it was decided to fix seven of the variables and vary the remaining four. The seven "fixed" variables were chosen because (1) they were thought now to be unimportant or (2) the results of experiment 1 had suggested an optimum choice, for example, the choice of ME rather than GSH as reducing agent. The effects of the four remaining factors (pH, amino acid concentration, ATP concentration, and enzyme concentration) were then studied by means of a 2^4 factorial design (16 runs) with repeat points at a half-fraction of that design (8 runs) and four center points. An analysis of main effects and interactions indicated the general direction in which to move the center of the design to improve the response.

Experiment 3. – The results of experiment 2 suggested that we were sufficiently close to the "optimum" to be able to benefit from the fitting of a second-order response surface in the same four variables. A central composite design, with replicated axial and center points, requiring 35 runs, was proposed and carried out. After the rejection of two points regarded as "outliers" a satisfactory fit was obtained with a quadratic surface. Unfortunately, we were apparently not as close to the optimum as we had thought, for the maximum fitted response within the experimental region occurred on the boundary of that region. The location of this predicted maximum was determined through the use of the "ridge analysis" method proposed by Draper.[34]

Experiment 4. – At this point the experimenter wished to check the effect of several buffers, in addition to the original two. A 3 × 6 factorial design was set up to examine six buffers at three points of interest in the other variables. One of these points of interest was the predicted maximum obtained from the analysis of experiment 3, and the actual response there was satisfyingly high. There was evidence, however, that the optimal point for one buffer may not be optimal for another. Further designs will attempt to exploit this possibility to improve the current "optimum." In addition, several of the other variables which have been fixed since the first design will be checked again.

[34] N. R. Draper, *Technometrics* 5, 469–79 (1963).

BIOLOGY

Pathology and Physiology

Low-Level Experiment. – The primary objective of this experiment is to determine the dose-response relation for radiation-induced carcinogenesis and life shortening under conditions of low-level irradiation in the mouse. Secondarily, data on the biochemical, physiological, microbiological, and morphological changes associated with late somatic effects of radiation and aging will be correlated with the dose-response relation. Preliminary results were sufficiently definitive by August 1967 to verify that the protocol, procedures, and facilities for the experiment were satisfactory. Irradiation of mice for the main experiment began then. Each week 120 mice are added to the experiment according to the plan shown in Table 6.

Table 6. Revised Radiation Dose Schedule for the Low Level Experiment

Dose (rads)[a]	Total No. of Mice to be Irradiated	No. of Mice Irradiated Each Week[c]
0	4000[b]	32
10	3000	24
25	1000	8
50	1000	8
100	1000	8
150	1000	8
300	4000[b]	32
Total	15,000	120

[a]Whole-body ^{137}Cs gamma rays, ~50 rads/min.

[b]Number includes mice for serial sacrifice.

[c]Mice set up in groups of eight per cage.

Because of fiscal limitations the number of mice to be set up has been reduced from 40,000 of each sex to 15,000 females. This has involved omission of males, omission of groups exposed at low intensities, and an increase in the lowest radiation dose given.

During the past year computer programs for storing,[35] retrieving, and analyzing pathology and hematology data[36] have been developed to complement those used for handling mouse ancestry, production, treatment, and life history data. We now make about 2000 transactions per week on about 8000 live experimental animals, 1000 dead experimental animals, 3000 breeders, and 6000 dead breeders. The programs which handle the life history or anamnestic data use tapes which contain information on over 30,000 mice. This includes data from preliminary experiments.[37] A small number of experimental mice have already died or been sacrificed. Although data on pathologic effects in these animals are highly preliminary, they disclose the expected increase in mortality with dose, at least at dose levels exceeding 50

[35]A. A. Brooks, *DSED, a Data Set Edit Program and the All Mighty Mouse Subprogram,* ORGDP-K-1748.

[36]A. Hume and A. A. Brooks, *SADS, Statistical Analysis of Data Subsets,* ORGDP-K-DP-3310.

[37]*Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 74.

rads. Of considerable interest is the evidence suggesting small effects even at the lowest dose level, 10 rads. Such evidence is highly preliminary and will require substantially larger numbers of animals for verification.

Sample Size Determination for Lens Opacity Experiments. – In the study of the effects of different types of radiation on the induction of lens opacification in mice, the problem of animal mortality arises because the mice are observed for long periods of time. Therefore it is desirable to start the experiment with enough animals so that there will be a sufficient number of animals alive to observe at the end of the experimental period. To arrive at estimates of the necessary sample sizes, this problem was approached as a hypothesis testing problem. That is, at a particular time after a group of mice have received a known dose and type of radiation, the mean percent lens opacity, μ, is estimated from the observations taken on this group of mice. The variance of these observations is denoted by σ^2, and it is desired to test the null hypothesis $\mu = \mu_0$ against the alternative $\mu = \mu_1$, where $\mu_1 \neq \mu_0$. If the probabilities of type I and type II error are fixed at α and β, respectively, then it can be shown that the sample size, n, necessary to satisfy these probabilities should satisfy:

$$n \geqslant \left[\frac{(t_1 + t_2)\sigma}{\mu_1 - \mu_0}\right]^2 ,$$

where

$$P\{-t_1 < t < t_1\} = 1 - \alpha ,$$

$$P\{t > t_2\} = \beta ,$$

and t is the Student's t statistic. Since t_1 and t_2 both depend on n, the determination of n must be done iteratively.

It is now assumed that the sample size determined from the above procedure is the number of animals observed at the end of the experimental period. From the information available on accumulative mortality, it is possible to estimate the probability, p, that an animal receiving a known dose survives the complete experimental period. Hence the expected number of animals needed at the beginning of the experiment for the radiation dose of interest is given by n/p. In many experiments the mice are grouped together into cages containing k animals, so n/p is divided by k and rounded to the next largest integer in order to determine the number of cages required.

Some Approximate Expressions for Variances Related to the Michaelis-Menton Equation. – The Michaelis-Menton equation is a particular example from the general set of equations referred to as rectangular hyperbolas and may be written as

$$v = \frac{V_{\max}\,[S]}{(K_m + [S])} , \tag{1}$$

where v is the initial reaction velocity between an enzyme and different concentrations of substrate, $[S]$ represents the concentration of the substrate, K_m is the Michaelis-Menton constant, and $V_{\max}$ is the maximum reaction velocity when the enzyme exists entirely in the combined form. Consider the following expression:

$$K(i) = \frac{iV_p}{V - V_p} , \tag{2}$$

where V and V_p are different values of the parameter V_{max} in (1) and i is a known constant. Bliss and James[38] and Hanson *et al.*[39] discuss the methods used to obtain the maximum likelihood estimators of the parameters V_{max} and K_m in (1) and also the variances of these estimators. Let $\hat{V}_p$ and $\hat{V}$ be the maximum likelihood estimators of V_p and V, respectively; then $K(i)$ is estimated by $\hat{K}(i)$, which is found by substituting $\hat{V}_p$ and $\hat{V}$ into (2). If it is assumed that $\hat{V}_p$ and $\hat{V}$ are independent random variables, then the variance of $\hat{K}(i)$ may be approximated by

$$\operatorname{Var}\hat{K}(i) \doteq \frac{(i\hat{V})^2}{(\hat{V}-\hat{V}_p)^4}\operatorname{Var}\hat{V}_p + \frac{(i\hat{V}_p)^2}{(\hat{V}-\hat{V}_p)^4}\operatorname{Var}\hat{V}. \tag{3}$$

The quantity $K(i)$ is also related to K and K_p by the expression

$$K(i) = \frac{iK}{K_p - K}, \tag{4}$$

where K and K_p are different values of the parameter K_m in (1). If the maximum likelihood estimators of K and K_p, denoted by $\hat{K}$ and $\hat{K}_p$, respectively, are substituted into (4), $K(i)$ may be estimated by $\hat{\hat{K}}(i)$. The variance of $\hat{\hat{K}}(i)$ may be approximated by

$$\operatorname{Var}\hat{\hat{K}}(i) \doteq \frac{(i\hat{K}_p)^2}{(\hat{K}-\hat{K}_p)^4}\operatorname{Var}\hat{K} + \frac{(i\hat{K})^2}{(\hat{K}-\hat{K}_p)^4}\operatorname{Var}\hat{K}_p$$

when it is assumed that $\hat{K}$ and $\hat{K}_p$ are independent variables.

If there are two different sets of parameters for Eq. (1), then there are two distinct curves when these functions of $[S]$ are plotted. Denote the two functions by

$$v_1 = \frac{V[S]}{K + [S]}$$

and

$$v_2 = \frac{V_p[S]}{K_p + [S]}.$$

The point of intersection of these two functions, S^*, is found by solving $v_1 = v_2$. If the maximum likelihood estimators for the parameters are used in the expression for S^*, it may be estimated by

$$\hat{S}^* = \frac{\hat{V}\hat{K}_p - \hat{V}_p\hat{K}}{\hat{V}_p - \hat{V}}.$$

Under the assumption that the two sets, $(\hat{V}, \hat{K})$ and $(\hat{V}_p, \hat{K}_p)$, of the maximum likelihood estimators are independent, the variance of $\hat{S}^*$ is estimated by:

$$\operatorname{Var}\hat{S}^* \doteq \frac{\hat{V}_p^2(\hat{K}_p-\hat{K})^2}{(\hat{V}_p-\hat{V})^4}\operatorname{Var}\hat{V} + \frac{\hat{V}^2(\hat{K}-\hat{K}_p)^2}{(\hat{V}_p-\hat{V})^4}\operatorname{Var}\hat{V}_p + \frac{\hat{V}_p^2}{(\hat{V}_p-\hat{V})^2}\operatorname{Var}\hat{K}$$
$$+ \frac{\hat{V}^2}{(\hat{V}_p-\hat{V})^2}\operatorname{Var}\hat{K}_p + \frac{2\hat{V}_p^2(\hat{K}-\hat{K}_p)}{(\hat{V}_p-\hat{V})^3}\operatorname{Cov}(\hat{V},\hat{K}) + \frac{2\hat{V}^2(\hat{K}-\hat{K}_p)}{(\hat{V}_p-\hat{V})^3}\operatorname{Cov}(\hat{V}_p,\hat{K}_p).$$

---

[38] C. I. Bliss and A. T. James, *Biometrics* 22, 573–602 (1966).

[39] K. R. Hanson, R. Ling, and E. Havir, *Biochem. Biophys. Res. Commun.* 29, 194–97 (1967).

If there are several independent estimates of $K(i)$, for example, $\hat{K}_1(i)$, $\hat{K}_2(i)$, . . . , $\hat{K}_n(i)$, and if the weighted average

$$\hat{\bar{K}}(i) = \frac{\sum_{j=1}^{n} w_j \hat{K}_j(i)}{\sum_{j=1}^{n} w_j}$$

is used as the estimator of $K(i)$, then

$$\text{Var}\, \hat{\bar{K}}(i) = \frac{\sum_j w_j^2 \text{ Var } \hat{K}_j(i)}{\left(\sum_j w_j\right)^2},$$

or, if $w_j = 1$ for all j,

$$\text{Var}\, \hat{\bar{K}}(i) = \frac{\sum_j \text{Var}\, \hat{K}_j(i)}{n^2}.$$

Confidence Limits on the Difference in Age at Death. – Consider the experimental situation where there are two groups of mice, one group which contains N_1 mice that have been irradiated with a known dose of radiation, and the second group which contains N_2 mice that have not been irradiated (i.e., a control group). Let x_i, $i = 1, 2, \ldots, N_1$, be the age at death of the ith animal to die in the irradiated group, and let y_j, $j = 1, 2, \ldots, N_2$, be the age at death of the jth animal to die in the control group. Therefore $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_{N_1}$ and $y_1 \leqslant y_2 \leqslant \ldots \leqslant y_{N_2}$ for the irradiated and control groups respectively. Although N_1 and N_2 are known, only $N_1/2 < n_1 \leqslant N_1$ values of the x's and $N_2/2 < n_2 \leqslant N_2$ values of the y's are actually observed; that is, these are censored samples.

The observations x_i, $i = 1, 2, \ldots, N_1$, and y_j, $j = 1, 2, \ldots, N_2$, are particular observed values of the random variables X and Y respectively. The probability density functions of X and Y are given by $f_1(X)$ and $f_2(Y)$, respectively, where f_1 and f_2 differ in location only. Since it is desired to estimate the difference in location of f_1 and f_2, one approach would be to estimate the medians of f_1 and f_2, denoted by ν_1 and ν_2, respectively, and then look at their difference. If N_1 is odd then ν_1 is estimated by x_k, where $k = (N_1 + 1)/2$. If N_1 is even then ν_1 is estimated by $(x_k + x_{k+1})/2$, where $k = N_1/2$. Similar results hold for the y's and N_2. The 95% confidence limits on the difference $\nu_2 - \nu_1$ are given by (1) the lower confidence limit, $y_{s'} - x_{r'}$, where

$$s' = \frac{N_2}{2} - \frac{1.96\sqrt{N_2}\sqrt{N_1 + N_2}}{2(\sqrt{N_1} + \sqrt{N_2})},$$

$$r' = \frac{N_1}{2} + \frac{1.96\sqrt{N_1}\sqrt{N_1 + N_2}}{2(\sqrt{N_1} + \sqrt{N_2})},$$

and (2) the upper confidence limit, $y_s - x_r$, where

$$s = \frac{N_2}{2} + \frac{1.96\sqrt{N_2}\sqrt{N_1 + N_2}}{2(\sqrt{N_1} + \sqrt{N_2})},$$

$$r = \frac{N_1}{2} - \frac{1.96\sqrt{N_1}\sqrt{N_1 + N_2}}{2(\sqrt{N_1} + \sqrt{N_2})}.$$

If $f_1(X)$ and $f_2(Y)$ are both normal, then a more efficient approach considers these samples as type I singly censored samples on the right; that is, of the N_1 (or N_2) observations on the random variable X (or Y), full measurement is possible only if $X \leqslant x_0$ (or $Y \leqslant y_0$) and there are n_1 (or n_2) observations on X (or Y) where full measurement has been made. For the experiment under consideration, n_1 (or n_2) of the N_1 (or N_2) animals in the irradiated (or control) group have died up to the moment of observation, and x_0 (or y_0) is the last recorded age at death. Let μ_1 and σ_1^2 be the mean and variance, respectively, of X. Let μ_2 and σ_2^2 be the mean and variance, respectively, of Y. The estimators of μ_1, σ_1^2, μ_2, and σ_2^2 are given by

$$\hat{\mu}_1 = \bar{x} - \hat{\lambda}_1(\bar{x} - x_0),$$

$$\hat{\sigma}_1^2 = s_1^2 + \hat{\lambda}_1(\bar{x} - x_0)^2,$$

$$\hat{\mu}_2 = \bar{y} - \hat{\lambda}_2(\bar{y} - y_0),$$

$$\hat{\sigma}_2^2 = s_2^2 + \hat{\lambda}_2(\bar{y} - y_0)^2,$$

where s_1^2, s_2^2, $\bar{x}$, and $\bar{y}$ are the respective sample variances and means for X and Y. Cohen[40] gives tables to obtain the estimates $\hat{\lambda}_1$ and $\hat{\lambda}_2$. In addition, he also gives expression for Var $\hat{\mu}_1$ and Var $\hat{\mu}_2$. From these expressions and under the assumption of independence of X and Y, an approximate 95% confidence interval on $\mu_2 - \mu_1$ may be found from:

$$\hat{\mu}_2 - \hat{\mu}_1 \pm 1.96\left(\mathrm{Var}\,\hat{\mu}_1 + \mathrm{Var}\,\hat{\mu}_2\right).$$

Heat Resistance of Spores. – In order to test the heat resistance of the spore *Bacillus globigii* in a near vacuum for NASA, it was necessary to find a method to deposit spores uniformly on spore pads. The method developed uses an aerosol sprayed into a polycarbonate seeding chamber (modified rat cage) which has 20 pad positions. Four chambers were constructed. A uniformity trial was conducted to test uniformity of spore deposition over pad position as well as from chamber to chamber. No significant nonuniformity was found. An experiment was then designed to test the effect of two temperatures and five periods of exposure to heat on spore survival. The spore survival rates were almost uniformly low, and further experimentation is in progress in a region of lower temperature and shorter periods of exposure to heat.

---

[40] A. C. Cohen, *Technometrics* 1, 229–31 (1959).

Mammalian Recovery

A Stochastic Model for Cellular Proliferation. – Various models have been proposed to describe the kinetics of cellular proliferation. The basis for these models is the assumption that the generation times are not constant but have some sort of probability distribution, independent of the population size.

A probabilistic model has been proposed in which the generation times vary according to population size. The model explains the observed distribution of generation times by considering each generation time, T, to be the sum of two successive time intervals. The first is a "waiting time," t, and the second a "division time," τ. The waiting time interval is the length of time from cell birth up to the time when the cell prepares to divide. This time is assumed to have a negative exponential distribution and to be dependent on the population size. The division time is composed of two intervals. The first interval is a constant length of time which is the minimum time required for cell division. The second interval is a random time which is assumed to have a gamma distribution.

Thus we have

$$T = t + \tau ,$$

where t is distributed as

$$\lambda f(N)\, e^{-\lambda f(N)\, t}$$

for some nondecreasing function $f(N)$ and where τ is distributed as

$$\frac{\mu^r}{\Gamma(r)} (\tau - \tau_0)^{r-1}\, e^{-\mu(\tau - \tau_0)}$$

for $\tau > \tau_0$.

The results of computer simulations based on this model agree very well with some experimental results in bacterial cultures. In particular, by incorporating the dependence on N a considerable improvement was seen in modeling the bacterial data of Kubitschek.[41] Some time lapse data on hemopoietic stem-cell populations will soon become available for further computer simulation work. Finally, the dependence between sibs and the effects of immigration-emigration will be brought into the model.

Secondary Disease (Homologous). – A previous report[42] described the latest in a series of experimental designs to investigate mortality from secondary disease in irradiated mice which have been treated with homologous bone marrow. The purpose of that design was to investigate these regions in the factor space (where the factors are age of donor, day of injection, cell dose, and sex) in which the mortality was expected to be high. The location of each of these regions is shown in Fig. 21.

The data from this experiment have now been collected and analyzed. Considering our prior expectations, the overall mortality at 90 days was disappointingly low, a result which may be due in part to the change in strain of mouse which the experimenter was unfortunately required to make in order to proceed with the experiment.

Within each region, however, significant effects of one or more of the factors were found. The main effects, in particular, contributed to the overall picture of the response surface which seems to be emerging

---

[41] H. E. Kubitschek, *Exptl. Cell Res.* **26**, 439–50 (1962).

[42] *Math. Div. Ann. Progr. Rept. Dec. 31, 1966,* ORNL-4083, pp. 88–89.

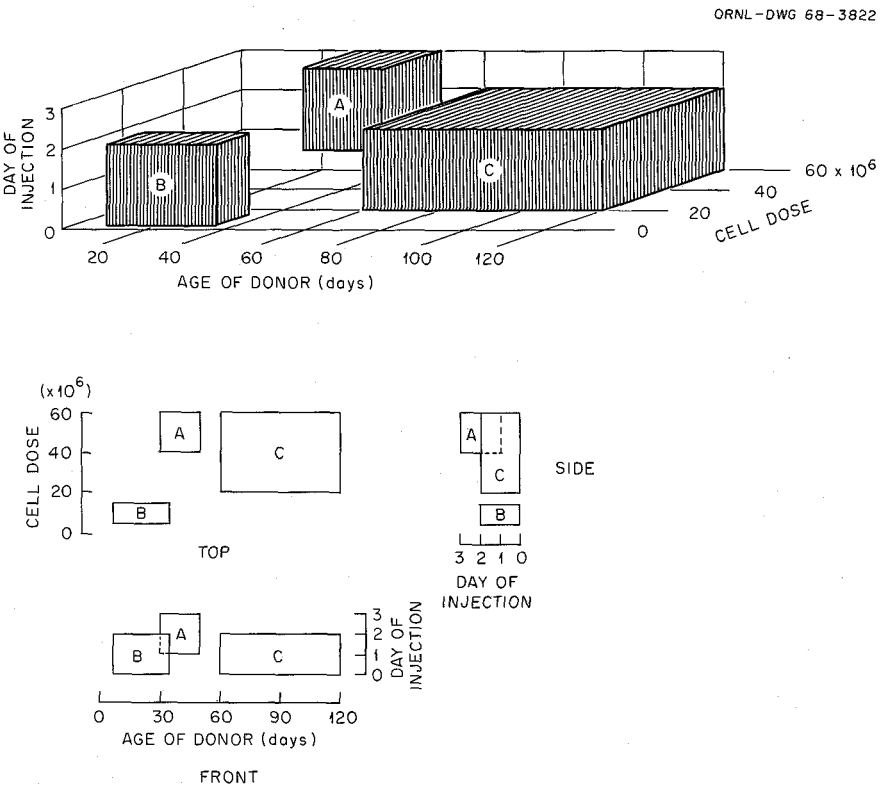


Fig. 21. Design VI (Homologous) – Regions of Experimentation.

from this series of experiments. Interactions among the factors did not play an important role in this particular experiment, except in region *B*, where several two-factor interactions were observed.

In addition to the usual factorial analysis of the effects of the controlled design variables, the design of this experiment permitted an investigation of cage-to-cage variation in mortality as well as unexplained variation in mortality with the time of the experiment. The latter source of variation, which appears to be quite important, was handled by dividing the total experiment into a sequence of "time blocks," or subexperiments. The "balancing" of the levels of the design variables within each time block is shown schematically in the lower part of Fig. 22. As the figure indicates, each time block is basically a little fractional factorial design in itself.

To reduce the adverse effects of cage-to-cage variation in mortality (beyond that ascribable to differences in treatments) the design was also blocked with respect to cages. This was the first design in this series in which animals representing several different treatments were placed in the same cage. The composition of each of the two types of cage is shown in the upper part of Fig. 22, where each dot represents a pair of identically treated mice. Although the variation due to cages (which is still not understood) was not as large as the variation from one time block to the next (which is also not understood), it still lurks as a potential troublemaker which should not be neglected.

Currently, our work in this homologous secondary disease program is centered around an attempt to combine the results of the whole series of experiments, using a single empirical statistical model.

ORNL-DWG 68-3823

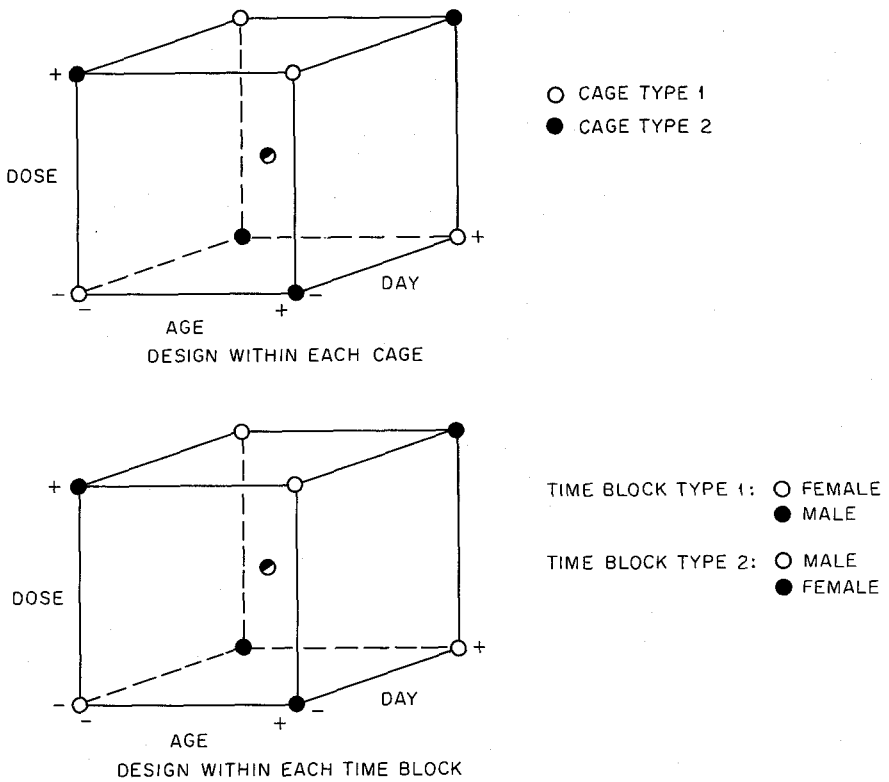


Fig. 22. Design VI (Homologous) – Layout Within Cages and Time Blocks.

Secondary Disease (Heterologous). – Experimental design techniques are also being utilized in the study of secondary disease associated with *heterologous* bone marrow transplants. A previous report[43] described the results of the first experiment in this series, which was a central composite design in the four variables age of donor, day of injection, cell dose, and sex.

A second central composite design has been constructed about the point of lowest mortality as predicted from the results of the first experiment. This second experiment, which is now under way, marks the introduction of a new factor, "environment," representing the effect of maintaining the mice in filter-top cages throughout their lives. Preliminary results indicate that this factor may be quite important.

The design requires 84 "groups," where each group represents one cage of five identically treated mice, there being a total of 420 mice in the experiment. The basic ingredients of the design are:

1. A replicated 2_V^{5-1} fractional factorial design in the five variables (32 groups).
2. A set of 44 groups, 11 at each of the four possible combinations of the levels of the qualitative factors (sex and environment). Within each of these sets of 11 groups are 4 groups corresponding to axial points in each of the variables "day" and "dose," 4 groups corresponding to axial points in the variable "age," and three "center points" in these quantitative variables.

[43] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, pp. 72–73.

3. A replicated 2_{III}^{3-1} design (8 groups) located in another region of the factor space which indicated a possibility of low mortality in the previous experiment.

At the request of the experimenter, the runs of this design are not to be performed in completely randomized order but instead are partitioned into a sequence of blocks. Within each block the order of runs is random, and the blocks themselves are performed in random order. However, the age of donor marrow within each block is constant, a circumstance which might lead to some difficulty in interpreting the effect of age of donor if there is a good deal of variation with respect to time. This risk was accepted by the experimenter in view of the logistic problems he would face if required to do a completely randomized design.

Cytogenetics

The Human Karyotype. – A statistical study of the human karyotype was described in a previous report.[44] A paper containing the results of the study has been submitted for publication.[45] Measurements of human chromosomes from 100 individuals (50 males and 50 females) have been analyzed. Mean long and short chromosome arm lengths and their respective dispersions were estimated for each chromosome. The mean and dispersion of the ratio of short arm to long arm were also estimated for each chromosome.

The estimated mean vectors and dispersion matrices are being used to determine the predictive densities and posterior probabilities for classifying chromosomes. Program CHROME, mentioned in a previous report,[46] computes the predictive density and posterior probability that a triple of measurements identify a particular chromosome. Another section of the program uses these predictive densities to determine the karyotype of the cell. This section is still being revised. It must be flexible enough to find an acceptable karyotype for most "normal" cells but still reject the abnormal cells.

Leukemia Cells. – The purpose of this investigation was to determine if cells from leukemic patients exhibit any depression of metabolic activity in the absence of certain nutrients which are usually synthesized by normal cells. Metabolic activity was determined by incorporation of radioactive precursors of DNA, RNA, and protein into cell fractions. Since several different patients were studied in experiments in which a variety of nutrients were employed, the statistical analyses were of importance in uncovering consistent, significant differences which occurred in most cell samples with two of the nutrients.

The experimental error was estimated from the variation among the nine patients. Tests of homogeneity of variance indicated that the variances could be pooled for all of the nutrients within a DNA, RNA, and protein class. The resulting pooled variances were used to construct confidence limits on the means for each combination of nutrient and DNA, RNA, and protein class. The conclusions were drawn on examination of the table of means and confidence limits.

Radiation Immunology

Interaction of Cell Types in the Generation of Antibody-Forming Cells – a Stochastic Model. – A stochastic model was considered for the number of antibody-forming cells generated as the result of a dose of lymphoid cells which were introduced into a diffusion-chamber culture. The antibody-forming cells are

---

[44] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 75.

[45] M. A. Bender and M. A. Kastenbaum, "Statistical Analysis of the Normal Human Karyotype," submitted to *Am. J. Human Genet.*

[46] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 94.

assumed to arise as the result of an interaction between antigen-reactive cells and progenitor cells. Each antigen-reactive cell was assumed to be able to react with, at most, S progenitor cells. Also the number of progenitor cells in a dose was assumed to be very large relative to the number of antigen-reactive cells. Thus when a progenitor cell interacts with an antigen-reactive cell, the probability P_D (where D represents the dose level) of another progenitor cell reacting with the same or another antigen-reactive cell will not be affected.

The number of antigen-reactive cells present in a dose of size D was assumed to be a Poisson random variable with mean $\lambda_D = AD$. The value of A can be estimated by the use of limiting dilution experiments.

Based on the above assumptions, the probability that I interactions had occurred at a dose of size D was found to be

$$P(I|D) = \sum_k \binom{kS}{I} P_D^I \, (1 - P_D)^{(kS-I)} \, e^{-AD} \, (AD)^k / k! \, ,$$

$$I/S \leqslant k \leqslant \infty \, , \quad 0 \leqslant I \leqslant \infty \, ,$$

where k is the number of antigen-reactive cells present at a given dose.

Thus the mean number of interactions which occur at a dose of size D is given by

$$E(I|D) = SP_D AD \, .$$

If a cell producing the indicated type of reaction goes through an average of x divisions, then the resulting maximum number of antibody-forming cells formed (disregarding death) would be

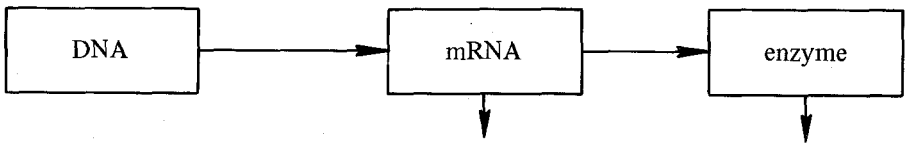
$$E(\text{number of antibody-forming cells}|D) = E(I|D) \, 2^x = SP_D AD 2^x \, .$$

The form of P_D which seemed to best describe the experimental results was

$$P_D = \begin{cases} AD/S, & AD < S \, , \\ 1 \quad , & AD \geqslant S \, . \end{cases}$$

Biochemical Regulation

A Model for the Kinetics of Enzyme Production. – Consider the following compartmental system:



for enzyme production. We denote the quantity of mRNA at time t by N_t and the amount of enzyme by M_t. It is assumed that the input to the mRNA pool is a constant c_1 and the input to the enzyme compartment is a fixed proportion, c_2, of the amount of mRNA. The output of both compartments is assumed to be exponential with parameter λ_1 for the mRNA and λ_2 for the enzyme. With this information we can write the equations

$$\frac{dN_t}{dt} = c_1 - \lambda_1 N_t ,$$

$$\frac{dM_t}{dt} = c_2 N_t - \lambda_2 M_t ,$$

which will yield the solution

$$M_t = e^{-\lambda_2 t} \left\{ M_0 + \frac{(N_0 - c_1/\lambda_1)c_2}{\lambda_1 - \lambda_2} \left[1 - e^{-(\lambda_1 - \lambda_2)t} \right] - \frac{c_1 c_2}{\lambda_1 \lambda_2} \left(1 - e^{\lambda_2 t} \right) \right\} \quad \text{for } \lambda_1 \neq \lambda_2$$

and

$$M_t = \frac{c_1 c_2}{\lambda_1^2} + e^{-\lambda_1 t} \left[(N_0 - c_1/\lambda_1)c_2 t + M_0 - \frac{c_1 c_2}{\lambda_1^2} \right] \quad \text{for } \lambda_1 = \lambda_2 ,$$

where M_0, N_0 are the amounts present at time $t = 0$.

We wish to know the effects on M_t if either c_1 or c_2 is suddenly decreased after the system has been in a steady-state condition. If c_2 is decreased then M_t decreases as e^{-bt}, while if c_1 is decreased then M_t decreases as $e^{-bt}(t + d)$ with an inflection point at $1/\lambda$.

CHEMICAL TECHNOLOGY

Estimation of the Parameters in an Anion Exchange Model

In the anion exchange model for M^{3+}-SO_4 complexes, it can be shown that the concentration of the total amount of metal in the aqueous phase is given by the following expression:

$$\Sigma[\mathrm{M}]_{\mathrm{aq}} = 2[n] + \sqrt{[n]} \left\{ \frac{K_{n0}}{G_{n0}[S]^{1.5}} + \frac{K_{n1}}{G_{n1}[S]^{.5}} + \frac{K_{n2}}{G_{n2}[S]^{-.5}} + \frac{K_{n3}}{G_{n3}[S]^{-1.5}} \right\} . \tag{1}$$

In (1), $[n]$ represents the unknown concentration of the neutral sulfate species, the K's are unknown thermodynamic constants, $[S]$ is the known sulfate ion concentration, and the G's are known constants.

For some fixed values of $[S]$ and the G's, the concentration of the total amount of metal in the aqueous phase is observed by comparing the counts per minute per milliliter of a sample with the counts per minute per milliliter of a standard whose concentration is known. Therefore the observed value of $\Sigma[\mathrm{M}]_{\mathrm{aq}}$ is found by multiplying the counts per minute per milliliter of the sample by a constant. Using the observed total count of the sample, which is assumed to be a Poisson random variable whose mean is proportional to (1), a weighted least-squares procedure was proposed to estimate the parameters in (1). The

weights were taken to be the inverse of the mean of the Poisson random variable. Since these weights were also functions of the unknown parameters, an iterative procedure was developed to determine the weighted least-squares estimates of the parameters. In addition, expressions for the variances of the weighted least-squares estimates were derived. Although the proposed technique has worked well for some examples, there have been situations where unrealistic estimates were found. Work is continuing in an effort to modify the iterative estimation procedure so that constraints may be placed upon the estimates of the parameters.

ELECTRONUCLEAR

Electron Microscopy

The requirements for high-resolution microscopy include a high-voltage supply regulated to a few parts in 10^7 during the limited time of exposure. One of the critical components in achieving stability is a string of several hundred resistors. The temperature coefficient, which is the rate of change in resistance as a function of temperature, must be near zero and within defined limits. Given estimates of the temperature coefficients of each resistor, a string can be made with an estimated temperature coefficient near zero.

The temperature coefficient of a resistor can be estimated with small standard error by measuring resistance at six different temperatures. The standard error of the temperature coefficient of the string is a function of the errors of estimation for the individual resistors. The increase in error caused by the necessity of reducing the number of measurements by omitting one or more temperatures was considered.

The resistance is a cubic function of temperature, and the temperature coefficient is defined as the slope of the curve at the given temperature. The capability of determining the best subset of temperatures from a given set was developed even though a general solution was not found. One subset of five temperatures was found which resulted in a loss of efficiency of only 7%. The second best subset of five and the best subset of four each have a loss of efficiency of 20%.

HEALTH PHYSICS

Estimation from Pooled Samples

In Radiation Ecology continuous sampling programs monitor levels of radioactivity in experimentally contaminated environments. Because of the minute size and low levels of radioactivity in many organisms, pooling of individuals is necessary before an observation is made. It was assumed that the random variables representing the total weight and total radioactivity from a sample are each a sum of the same number of normal independent and identically distributed random variables. Although it was assumed that the above random variables were not independent for a particular sample, from one replicated sample to another independence of these random variables was assumed. From these assumptions and the observed total weight and total radioactivity of replicated pooled samples, unbiased estimates of the mean weight and mean radioactivity per individual were found along with the variances of these estimates.

In addition, a measure of the mean radioactivity per unit weight was defined by the ratio

$$\frac{\text{mean radioactivity per individual}}{\text{mean weight per individual}}.$$

An estimate of this ratio was derived along with an expression for its variance.

A computer program, PITFALL, which is described in the statistical programming section of this report, has been written to compute the statistics mentioned above after the original data have been sorted according to a specified classification procedure.

Half-Lives of ^{60}Co and ^{137}Cs in Leaves

Leaves of alder, sycamore, and oak, some tagged with ^{60}Co, some tagged with ^{137}Cs, and some not tagged, were exposed in small nylon bags in a brook, in a pond, and on a forest floor. Periodic measurements of the loss of weight were made on the untagged leaves and of the amount of ^{60}Co or ^{137}Cs present on the tagged leaves at ten time periods over the course of a year. These measurements provided data to relate leaf species and sites to weight loss rates and half-lives during the first year of exposure. Single bags of untagged leaves and duplicate bags of tagged leaves were exposed.

The ten measurements for each bag separately (45 bags in all) were analyzed according to the regression model

$$\log Y_i = \alpha + \beta t_i \,, \quad i = 1, 2, \ldots, 10 \,,$$

where Y_i is the measurement made at the ith time period, t_i, in order to estimate the coefficient β, which represents the rate of weight loss or rate of loss of ^{60}Co or ^{137}Cs as the case may be. These rates then constituted a derived response which was subjected to an analysis of variance in order to test the effects of the factors (leaf species and sites) on the rates.

Interval estimates of the half-lives were easily obtained since the half-life is inversely proportional to the coefficient β.

ORIC Lake Study Analysis

In order to study the problem of eutrophication in ORIC pond, a set of 14 characteristics were chosen to describe the existing conditions of the pond. A preliminary study in the summer of 1966 had so few sample points that it could not give an accurate picture of the whole pond. It was recommended that more stations be chosen throughout the pond and that a small number of observations be taken at each station so that after analysis of the results the number of stations might be reduced.

Sixteen stations were chosen evenly spaced throughout the pond, and two more were placed at flow points. Four sets of observations on the 14 characteristics were taken on two successive days in August 1967. The results were analyzed separately for each characteristic as a 4 × 18 factorial experiment. A Duncan multiple range test showed a highly significant difference among stations for each variable and indicated which stations differed from the others. As a result of this test, it was possible to put the stations into five groups so that the stations within each group would give fairly consistent results for some of the characteristics. By picking one or two stations from each group, the experimenter can still obtain estimates for the whole pond but reduce his costs considerably.

A study was made of the relative efficiencies of a proposed 1968 study with the number of stations reduced to $1/3$, $1/2$, $2/3$, or $3/4$ of the number in the 1967 study.The results showed no consistent pattern in change of efficiency but indicated a large loss in efficiency for some characteristics and a small gain in efficiency for others.

Honeybee Irradiation Study

Statistical analysis was used to assess the results of a series of experiments testing the response of honeybees to prompt ionizing radiation. Doses of 500, 1000, 2000, and 4000 rads of ^{60}Co gamma radiation were given to four hives of Cordovan honeybees, and six hives were used as controls. The effects of the irradiation were assessed for the flight activity, pollen collection, and mortality of the colonies. The design involved both crossed and nested factors; the days after radiation were crossed with hives and radiation levels, and hives were nested within radiation levels. The number of flights of the bees during 2-min periods was counted several times each. Since the number of counts varied, the total number of flights for each day was weighted by the number of counts taken that day to minimize the effect of the different numbers of counts. The interaction between days and radiation levels was not significant, which indicates that the weighted analysis had the desired effect. Duncan's multiple range test was used to determine which radiation levels gave significantly different results.

Other experiments tested the effects of irradiation on mean lifetime of different strains of bees. The effects of different temperatures and different foods on mean lifetime were also studied. All of these experiments used designs with three or four nested factors.

Numerical Taxonomy of Three Populations of Aquatic Isopods

A minimax multivariate classification procedure, given by Anderson,[47] was constructed to classify an appropriate aquatic isopod into one of three populations, *Lirceus fontinales, Lirceus alabamae,* or *Lirceus lineatus.*

The procedure was based on 30 dimensionless variables which were determined from 27 characteristics measured on 240 aquatic isopods, 192 from *Lirceus fontinales,* 11 from *Lirceus alabamae,* and 37 from *Lirceus lineatus.*

Due to the similarity of the three populations and the small number of observations in the *Lirceus alabamae* and *Lirceus lineatus* populations, the probability that a new aquatic isopod would be misclassified was found to be approximately 0.33.

Resistance of Mosquito Fish to Irradiation

In a study of the effect of a radioactive environment on the resistance of *Gambusia affinis affinis* (mosquito fish) to irradiation, groups of fish from each of two ponds were subjected to various doses of irradiation. One of the ponds contains radioactive waste material, while the other does not. In addition to the effect of radioactive environment, a "temperature effect" was also studied. After irradiation, groups of fish were acclimated to, and then maintained at, two different temperatures: a "normal" temperature of 25° and a "cold" temperature of 20°.

In each group the mortality at 30 days was taken to be the response, and a dose-response curve was estimated for each of the four possible combinations of pond (radioactive or clean) and temperature (20° or 25°).

A good fit to all four sets of data was obtained through the use of a form of the logistic model which expressed the probability of death at dose X to be

$$P(X) = [1 + \exp(-\alpha - \beta X^3)]^{-1},$$

[47] T. W. Anderson, *Introduction to Multivariate Statistical Analysis,* pp. 126–52, Wiley, New York, 1958.

where α and β depend on the pond and on the temperature. Under the assumption that the mortality at each dose is binomially distributed, maximum likelihood estimates of α and β were found for each dose-response curve using a nonlinear estimation program with the minimum logit χ^2 estimates as initial estimates. Since maximum likelihood estimation was used, the goodness of fit was tested by referring to the asymptotic distribution of the logarithm of Wilks' ratio of likelihoods.

Tests of the equality of the two dose-response curves at each temperature and of each pond were made, using a normal approximation to the bivariate distribution of the two-parameter estimates for each curve. These tests indicated that differences in mortality due to environment (ponds) were not significant, as opposed to differences associated with temperature, where the colder temperature clearly increased the probability of survival.

A similar experiment is currently being performed at a "warm" (30°) temperature level.

METALS AND CERAMICS

Relationship Between Fluorine Content and Hardness

The relationship between fluorine content and hardness (Vickers diamond pyramid test) of deposit based on data from the tungsten chemical vapor deposition experimental program was investigated. Festa and Danko[48] suggest that deposit hardness (Knoop test) can be "used as a preliminary quality control check on fluorine content." From laboratory data it can be demonstrated that hardness is not a reasonable means of process control.

Since fluorine content is determined by chemical analysis, it is expected and assumed that the error associated with measuring fluorine content is substantially smaller than that of measuring hardness. Hence the regression of hardness on fluorine content was considered. Using a linear relationship as did Festa and Danko, the 95% confidence intervals about fluorine content were computed at three locations on the tungsten deposit. These intervals enclose the entire expected range of fluorine content, a condition not suitable for process control. Although these intervals were somewhat narrower when the regression of fluorine content on hardness was considered, the conclusions were the same.

Analysis of Sintered Aluminum Product Data

Two analysis of variance procedures were adapted to analyze data from experiments on sintered aluminum product (SAP) billets. The first analysis was used to determine the reproducibility of tensile strength of extruded billets. The second analysis was used to test the effect of annealing on tensile strength using selected billets with varying iron content. The effect of iron on tensile strength could not be tested, since iron content was completely confounded with billet-to-billet variation.

Two sets of data were submitted, and the manner in which these data were collected suggested the mathematical models upon which the analysis of variance procedures were based. The data were then used to illustrate the required calculations. Since these computations are to be applied to data from future experiments, some suggestions for collecting the data were made which would produce more meaningful results.

---

[48] J. V. Festa and J. C. Danko, *Proc. Conf. Chemical Vapor Deposition of Refractory Metals, Alloys, and Compounds*, p. 349 (1967).

Defining Regions of Experimentation Under Loss of Rotatability

A method of defining regions for future experimentation under loss of rotatability was suggested in an earlier report.[49] This method consists in searching the space of the independent process variables for areas of minimum information. A computer program has been written to calculate the information at randomly selected points in the space of k ($\leqslant 10$) variables. It was then applied to results obtained from chemical vapor deposition (CVD) of tungsten experiments.[50] Based on the use of this program five additional runs are proposed.

Rotatability, a feature built into many statistical experiment designs, assures the experimenter that the precision of parameter estimates is constant at a given distance from the center of the design. If the proposed design levels are not attained, rotatability is impaired. This procedure for defining regions of experimentation is an attempt to recapture in part the feature of rotatability.

Information is defined to be the ratio of the residual mean square to the variance of a predicted value. Under this definition, information is a dimensionless quantity which depends only on the inverse $(\mathbf{X}'\mathbf{X})^{-1}$, where $\mathbf{X}$ is the augmented design matrix for a linear model.

A computer program entitled INSEAR2 has been written for the IBM 360/75 to compute the information I at M points selected at random. The region of interest from which points are chosen is a k-dimensional hypersphere with radius R. For specified M, R, $(\mathbf{X}'\mathbf{X})^{-1}$, and k, the program prints out the ten smallest values of I and their corresponding coordinates.

This program was applied to data from the tungsten CVD project in which the effects of four process variables were investigated. Based on the $(\mathbf{X}'\mathbf{X})^{-1}$ of the original 81 experiments, the information was computed for 1000 randomly selected points in four-space. The point of minimum information was then chosen as the next experiment to be performed. A new $(\mathbf{X}'\mathbf{X})^{-1}$ matrix was then formed based on the 82 experimental points. Using the new inverse matrix, INSEAR2 was used again to define a second point of minimum information. This procedure was repeated until five proposed experiments were generated (see Table 7).

Table 7. Proposed CVD Experiments

Run No.	Information	Levels of Process Variables			
		Temperature (°C)	Pressure (torr)	Total Gas Flow Rate (cm^3/min)	Percent WF_6 in Inlet Gas
82	0.062	613	6.3	5020	6.17
83	0.201	546	7.2	4490	5.61
84	0.180	597	11.6	5570	6.59
85	0.203	600	10.2	1224	5.62
86	0.277	709	13.2	1136	5.90

To visualize what effect these proposed runs have on the information profile, we can plot lines of constant information in the space of any two variables. These contour lines are shown in Fig. 23 for two different planes.

[49] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 87.
[50] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 82.

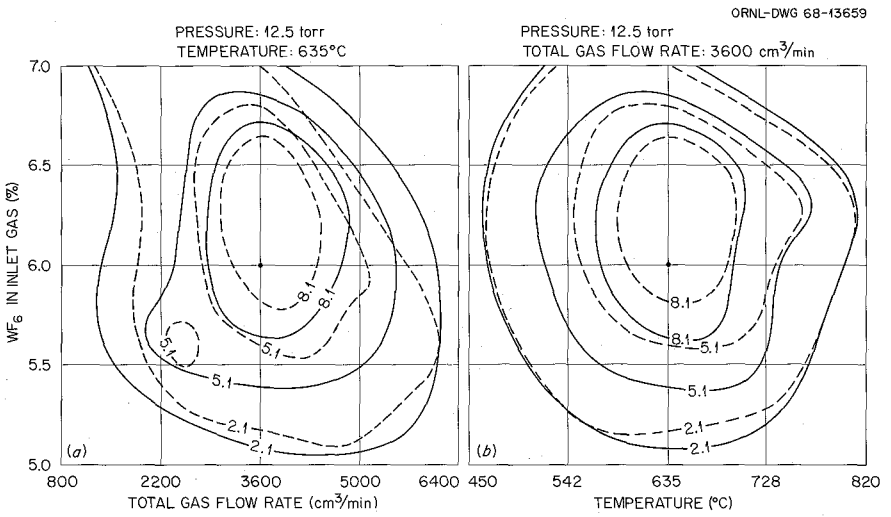


Fig. 23. Chemical Vapor Deposition of Tungsten – Contours of Constant Information.

The dashed lines represent the information contours based on the original 81 experiments and are taken from Figs. 12*c* and 12*d* of the last annual report.[51] By adding the five runs, the information profile takes the shape displayed by solid lines. The closed circle indicates the center of the original proposed design. As more and more runs are added, the contour lines will approach the concentric circles which would be generated by a true rotatable design.

Proposed Experiments in the Heavy Section Steel Technology Program

An evaluation of the proposed simulated service tests in the HSST program and outlined in a preliminary draft of a report by H. T. Corten was made from the viewpoint of statistics. Three projects were of primary interest:

1. Investigation of fracture strength as a function of temperature using tensile specimens 6 in. thick (12 experiments).
2. Investigation of the influence of crack configuration using tensile specimens 6 in. thick (8 experiments).
3. Nozzle-corner simulation (8 experiments).

The proposed experiments provide a sound basis for an experimental program. A recommendation was made that some runs be replicated to provide a measure of experimental error. In one project, an alternate proposal was presented which would permit error estimation without increasing the number of experiments.

[51] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 88.

REACTOR STANDARDS AND SAFETY

Statistical and mathematical approaches to reactor standards and safety problems are being developed with several divisions of the Laboratory. The problems include the estimation of reliability of safety systems, the testing of materials used in containment vessels, and special reactor fuel and reactor control problems. The reliability problem is discussed in this section, and other problems are summarized elsewhere in this report or are in progress. The derivation of the reliability of certain redundant systems is given in the research section.

It is well known that the reliability of components is eventually limited by cost, properties of materials, and quality of production. The reliability of a system of such components can be improved further by the use of multiple components in parallel, in series, or in combined redundancy.

The optimum allocation of components rapidly becomes a complex problem due to the many possible configurations and varying reliabilities and costs of each component. Several programs were written using algorithms appearing in the literature for determining an optimum allocation for systems of components in parallel or series with different reliabilities and costs. Additional complications arise due to the need and desirability of repairing components which have failed, of testing components on line, and of reducing the frequency of false alarms.

The mathematical concepts of probability are used to obtain solutions which are often more complete and more clearly summarize the results than a less sophisticated mathematical solution or a nonmathematical solution. It is especially important to define terms precisely and to check the validity of assumptions. The assumption of independence of individual components is particularly critical and is under investigation.

Many journal articles and reports were studied to help investigators, and a few manuscripts were reviewed.

Statistical Programming

THE TIME-SHARING COMPUTER TERMINAL

Since March 1967, the Statistics Department of the Mathematics Division has been using a remote terminal connected to the multi-access computing (MAC) facility operated by C-E-I-R, Inc. The central facility consists of one General Electric 235 computer and one General Electric 420 computer, situated at Silver Spring, Maryland. The programming languages BASIC, EXTENDED BASIC, FORTRAN, and ALGOL are available for use through the facility. Figure 24 shows the use of the time-sharing facility by month from installation date March 22, 1967, through October 31, 1968.

BIOLOGY PROGRAMMING

Radiation Mutagenesis

A series of data analysis programs to characterize *Neurospora crassa* adenine-3 loci mutations have been completed and will appear in a future TM publication. The programs give the experimenter a tabulation of the data from the genetic analysis of ad-3 mutants. The tabulations are used by the experimenter in evaluating the genetic effects of different mutagens in terms of survival and the spectrum of recessive lethal mutations. They also permit comparison of the dose-effect curves for each of the different classes of

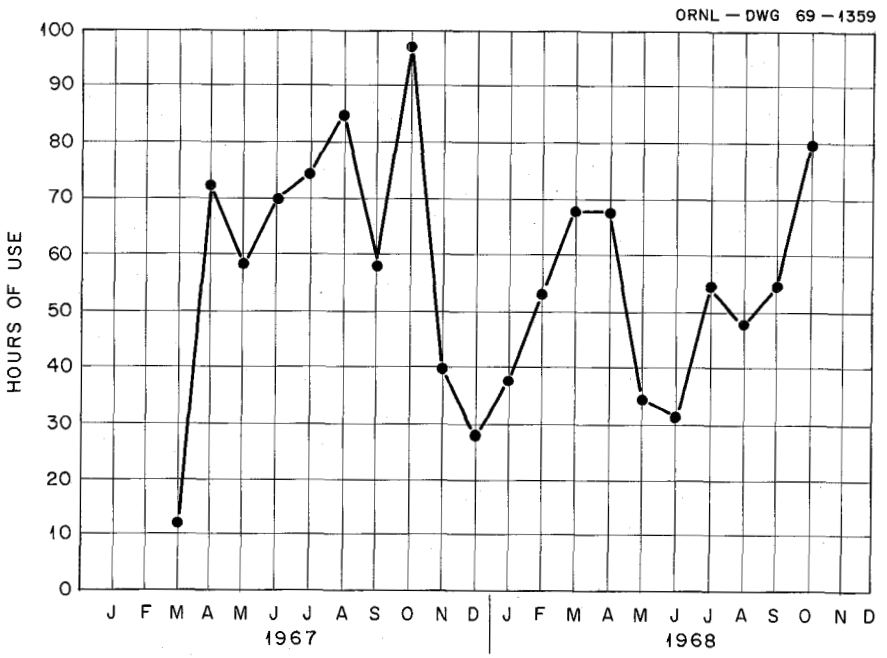


Fig. 24. Use of C-E-I-R On-Site Computing Service.

mutations with regard to slope and distance. Some of the induction agents studied this past year have been uv with and without photoreactivation, gamma irradiation under weightlessness (Biosatellite), beta irradiation under weightlessness (Gemini XI), x irradiation under anoxia, x irradiation at different dose rates, and a wide variety of chemical mutagens.

Pathology Data Control: Autopsy Code

A program SNOPROOF has been written to enable the pathologist to proofread historical and pathological data that have been recorded and punched into cards. Pathologists use a coding scheme known as the SNOP code to record any diagnoses made through an autopsy. This code consists of many different combinations of alphameric characters that indicate the various diagnoses made by the pathologist.

SNOPROOF produces an output that is used for checking against any errors that were made in the recording or the keypunching of the data. The output consists of an orderly listing of the historical data on the animal and a listing of the pathological codes converted to English words and phrases. From this output the pathologist is able to detect and correct any errors previously made.

Filter Cap Experiment Data

The data from a filter cap experiment were collected by attaching hygrometer probes to mouse cages containing five or ten mice. The cages were covered with different types of filter caps to provide a biological physical barrier between the mouse and the environment. Temperature and humidity were measured within the cage for one-week periods to study the effect of filter type and number of mice on these parameters. Other cage environment parameters were measured by other means.

A computer program WOLNY was written which plots the temperature and humidity data with 1 in. equal to 12 hr of test data. The external chamber and cages 1, 2, 3, and 4 may all be plotted on a single graph, or the graph may be restricted to three curves per graph. The simple statistics of means and ranges were computed for temperature and humidity for each cage and each of the four tests.

The filtered mouse cage appears to have a very significant influence on the experimental mouse and possibly the health of people working with mice. The determinations performed provide valuable data on physical properties of filters.

Low-Level Radiation Experiment

The nonlinear model

$$Y = 1 - a^{(1-e^{bx})},$$

where Y is cumulative mortality and x is age at death, was used to fit data obtained from preliminary parts of the Low-Level Experiment. The model is extremely sensitive. Frequently it was necessary to compute a sum-of-squares surface in order to find values of a and b which would lead to convergence.

A program was written to plot as many as ten sets of observed values and calculated curves on the same grid.

The program NONLIN described in STATPAK[52] was used to fit the model. All programs were written for the CDC 1604, and the plotting was done on the Calcomp plotter.

Foreign Spleen Experiment

During the course of studies on the beneficial effect of injection of hemopoietic cells into irradiated mice, it was noted that effects were extremely variable. At times the mice were harmed by the procedure, rather than helped. A study of the interaction of many parameters including age, sex, strain, radiation dose, physiologic and immunologic alterations of donor mice and recipient mice, and importance of timing and dose of injected cells resulted in the accumulation of large amounts of data. It was felt that the best return of information, correlation of experimental variables, and storage for future use would be attained by a computer-based system.

A program was written to isolate selected cages and separate the mice into categories of interest to the experimenter. One study separates the mice according to disposal methods, lists the day of death, and calculates cumulative percent mortality. Graphs of cumulative percent mortality vs day of death are also constructed.

The project is not complete, and it is expected that other examinations of the data will be made.

STATISTICAL METHODS

Random Distributions

A program for studying various distributions of random variables is being written. In it a sample is drawn from uniform, normal, gamma, exponential, binomial, or Poisson distributions. The sample is

[52] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 96.

ordered, some statistical computations are performed, and estimates of expectation of the order statistics are calculated. The ordered sample is plotted against these estimates using arithmetic, logarithmetic, normal, or gamma scales, whichever best suits the distribution.

Thus far the program draws samples and performs the statistical calculations for all the named distributions and calculates expectations of order statistics for the uniform and exponential. An algorithm for determining expectations from the gamma distribution has been written and is being debugged.

Data Smoothing

Program ESTIMATE smooths data, performs desired calculations, and smooths the results. It was written to do the three operations in one run rather than three runs which formerly were required using existing programs.

The smoothing is done by a five-point quadratic moving average, that is, a value $y(t_i)$ is smoothed by fitting a quadratic curve through five successive points $y(t_{i-2})$, $y(t_{i-1})$, $y(t_i)$, $y(t_{i+1})$, $y(t_{i+2})$ by least squares. The smoothed value $\hat{y}(t_i)$ is taken to be the value at t_i.[53]

Program NLLS obtained from the ORNL subroutine library was modified and used as the basis for least-squares fitting. It was altered to accept the data five points at a time and to store the estimated values. For the first two points, $\hat{y}(t_1) = y(t_1)$ and $\hat{y}(t_2) = y(t_2)$. The last two points of the experiment were handled in a similar manner.

For flexibility, calculations are performed in a subroutine which returns the results to the main program for smoothing.

Plotting Package

A general plotting package has been written which will compute and plot confidence limits on a regression line or confidence limits on the estimate of a future observation.

The input constants for the calculations may be easily obtained from the linear regression program BIMD29GR.[54] The plotted graph will contain the computed regression line, the observed values, and the upper and lower confidence lines. The program is called PLOTCONF and was written for the CDC 1604 in FORTRAN 63.

DATA PROCESSING

Pitfall Data

Using the formulas described in the Health Physics section under the title "Estimation from Pooled Samples," a computer program PITFALL was written to sort the original data according to a specified classification procedure and to compute the statistics desired. The "Pitfall" data are stored on magnetic tape in no specific order. They contain three fields of interest: species, time interval, and location of trap (peripheral or central), any one, two, three, or none of which may be specified for sorting and storing on a scratch tape for further processing.

---

[53] Carlo M. Metzler, *Estimations of Transport Rates by Radioisotope Studies of Non-Steady-State Systems,* Ph.D. dissertation, North Carolina State University, 1965.

[54] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 96.

Modifications of the program have been necessary to suit unique types of problems. For example, in one problem certain species were further classified by weight (small, medium, and large). The investigator wished the data run seasonally or by year. Since some species might be classed as small in one month and might be classed as medium in another, an option to produce punched cards containing the data fitting the primary reference was instituted. Thereafter, the data could be pooled in any manner desired and used in a modified version of the program as input.

Conversion of RBC and Spleen Counts

It has become necessary to make some changes in the program GOODMAN2[55] that processes counts of radioactivity in red blood cells and the spleen and also processes hematocrit counts from mice. In rare cases some counts are missing, but they must be treated as such and not as counts of zero. The program was altered to detect the difference between missing counts and counts of zero and to handle them accordingly. The same basic calculations of correction for background, ratios and percentages, and standard analyses of variances are made in addition to the multiple comparisons of means.

Specific Locus Mutation Study

Work has continued on the Specific Locus Mutation Study.[56] At this time experiments 01–06 and experiments 08–17 have been analyzed and the analyses turned over to the investigator. Because experiment 07 (the "G" experiment) data are recorded differently from the other experiments, a new program for processing the data was required. The animals in the experiment are classified by group and by row and by column. The group classification is dependent upon the date of the experiment, the row classification designates the time between irradiation and mating, and the column classification has to do with information from birth until weaning.

Five tables are prepared. Two tables classify the number of mutations in a three-way table (groups by rows by columns) and a two-way table (rows by columns). Three other tables classify the number of abnormalities in a four-way table (observer by groups by rows by columns), a three-way table (groups by rows by columns), and a two-way table (rows by columns).

All data are stored on tapes, and sorting for the various problems is completed quickly by the IBM 360 SORT/MERGE utility program.

Statistical Seminars and Lectures

SEMINARS

Ball, S. J. (Instrumentation and Controls Division, ORNL), "Applications of Spectral Analysis," May 29, 1968.

DeRouen, T. A., "Applications of the Logistic Function to the Analysis of Binomial Response Data," September 13, 1968.

---

[55] *Math. Div. Ann. Progr. Rept. Dec. 31, 1965,* ORNL-3919, p. 102.

[56] *Math. Div. Ann. Progr. Rept. Dec. 31, 1967,* ORNL-4236, p. 97.

Dolan, J. M., "Mathematical Models for Counters," February 21, 1968.

Doner, J. R., "Polya's Theory of Counting," August 7, 1968.

Draper, N. R. (University of Wisconsin), "Examining Residuals in Regression Analysis," October 9, 1968.

Draper, N. R., "Further Rotatable Designs and Residuals in Two Dimensions," October 9, 1968.

Good, I. J. (Virginia Polytechnic Institute), "The Chief Entities," May 6, 1968.

Gosslee, D. G., "Some Remarks on Reliability, Standby Redundancy, and Optimization," April 17, 1968.

Kryter, R. C. (Instrumentation and Controls Division, ORNL), "Applications of Spectral Analysis to Reactor Signals," June 5, 1968.

Kryter, R. C., "Application of the Fast Fourier Transform Algorithm to Spectral Analysis of Time Series," June 26, 1968.

Laska, E. M. (Research Center of Rockland State Hospital, Orangeburg, New York), "Robust Estimation of Location Parameter," February 28, 1968.

Lever, W. E., "Discussion of a Paper by R. R. Prairie on Probit Analysis," January 17, 1968.

Miller, F. L., "Discussion on Exponential Survival with Concomitant Variables," January 31, 1968.

Miller, F. L., "A Heuristic Approach to Time Series Analysis," March 13, 20, 27, April 3, 24, and May 1, 1968.

Miller, F. L., "Estimation of Spectral Densities in Time Series Analysis," May 22, 1968.

Mode, C. J., "Some Remarks on the Foundations of Generalized Age-Dependent Branching Processes," July 3, 1968.

Mode, C. J., "On the Mean and Covariance Functions of Generalized Age-Dependent Branching Processes: Biological Applications and Limit Theorems and Approximations," July 24 and 31, 1968.

Olson, W. H., "Martingales, the Ergodic Theorem and Strong Laws of Large Numbers," May 8 and 15, 1968.

Olson, W. H., "An Asymptotic Density of the Eigenvalues of Some Random Matrices," November 6, 1968.

Rao, J. N. K. (Texas A and M University), "Principles of Statistical Sampling," April 29, 1968.

Taylor, R. L., "Some Aspects of the Generalized Inverse and Uniqueness of a Minimum Norm Solution," August 14, 1968.

Lecture Series in Statistics and Probability

Lecture I: D. A. Gardiner, "Probability, Random Variables, and Distribution Functions," June 10, 1968.

Lecture II: V. R. R. Uppuluri, "Characterization Theorems and Limit Theorems," June 17, 1968.

Lecture III: M. A. Kastenbaum, "Contingency Tables," June 24, 1968.

Lecture IV: M. A. Kastenbaum, "Confidence Regions," July 1, 1968.

Lecture V: W. E. Lever, "Distribution-Free Tests," July 8, 1968.

Lecture VI: T. L. Hebble, "Regression," July 15, 1968.

Lecture VII: J. J. Beauchamp, "Nonlinear Estimation," July 22, 1968.

Lecture VIII: D. G. Gosslee, "Analysis of Variance," July 29, 1968.

Lecture IX: T. J. Mitchell, "Design of Experiments and Response Surface Methodology," August 5, 1968.

Lecture X: V. R. R. Uppuluri, "Elements of Stochastic Processes," August 12, 1968.

Professional Activities

Beauchamp, J. J., Chairman, Session on Analysis of Contingency Tables, Annual Meetings of the American Statistical Association and Biometric Society, ENAR-WNAR, Pittsburgh, Pennsylvania, August 22, 1968.

Gardiner, D. A., Secretary-Treasurer, Section on Physical and Engineering Sciences, American Statistical Association.

Gardiner, D. A., Member, Advisory Board, Section on Physical and Engineering Sciences, American Statistical Association.

Gardiner, D. A., Member, Editorial Board, Society for Industrial and Applied Mathematics.

Mitchell, T. J., Consultant, University of Tennessee Memorial Research Center Hospital, Knoxville, Tennessee.

Uppuluri, V. R. R., Chairman, Session on Dynamic Systems, Meetings of American Statistical Association, Biometrics Society (ENAR), and Institute of Mathematical Statistics, Blacksburg, Virginia, April 8–10, 1968.

Uppuluri, V. R. R., Consultant, System Development Corporation, Santa Monica, California.

Mathematics Division Publications

Altom, D. W., and W. L. Morris, *An Algorithm for Initial Value Problems,* ORNL-4282 (August 1968).

Beauchamp, J. J., and R. G. Cornell, "Simultaneous Estimation by Partial Totals for Compartmental Models," *J. Am. Statist. Assoc.* **63**, 573–83 (1968).

Beauchamp, J. J. (with E. L. Frome), "Maximum Likelihood Estimation of Survival Curve Parameters," *Biometrics* **24**, 595–606 (1968).

Beauchamp, J. J. (ed.), *Lecture Series in Statistics and Probability,* ORNL-4347 (in press).

Boughner, R. T., *Program ENGY,* ORNL-CF-67-12-18 (Jan. 25, 1968).

Boughner, R. T., *The Discretization Error in Finite Difference Solutions to the Linearized Navier-Stokes Equations for Incompressible Fluid Flow at Large Reynolds Numbers,* ORNL-TM-2165 (May 1968) (University of Tennessee M.S. thesis).

Chang, S. J., "A Note on the Fracture Theory of Griffith-Irwin Applied to a Linear Viscoelastic Material," submitted to *Intern. J. Fracture Mech.*

Chang, S. J., "Elastic-Viscoplastic Torsion for a Sokolovsky's Oval Section," submitted to *Acta Mech.*

Dolan, J. M., "On the Relationship Between the Oscillatory Behavior of a Linear Third-Order Equation and Its Adjoint," accepted for publication in *J. Differential Equations.*

Gosslee, D. G., and K. O. Bowman, *Evaluation of Maximum Likelihood Estimates of Parameters in Mixtures of Normal Distributions,* ORNL-TM-2110 (1968).

Gove, N. B., "Beta Decay Energy Systematics," *Proc. Third Intern. Conf. on Atomic Masses* (ed. by R. C. Barber), University of Manitoba Press, 1967.

Gove, N. B. (with C. T. Chen, F. W. Hurley, J. T. Miller), "Recent References, $A = 21$ to $A = 228$," *Nucl. Data* **B2**(2), 1 (1967).

Gove, N. B. (with M. Yamada), "Shell-Independent Systematics, S_n (Isotonic) and S_p (Isotopic)," *Nucl. Data* **A4**(3), 237 (1968).

Haeuslein, G. K., and Shirley B. Opstrup, *ORNL Book Catalog,* ORNL-TM-2185 (March 1968).

Hammons, C. E., *An Extrapolation Algorithm for Special Ordinary Initial Value Problems Based on an Elementary Trigonometric Multistep Formula,* ORNL-TM-2166 (June 1968) (University of Tennessee M.S. thesis).

Hannon, Beth H., and W. L. Morris, *Tables of Arithmetical Functions Related to the Fibonacci Numbers,* ORNL-4261 (June 1968).

Hoel, D. G., "Sequential Testing of Sample Size," *Technometrics* **10**, 331–41 (1968).

Hoel, D. G., "Closed Sequential Tests of an Exponential Parameter," *Biometrika* **55**, 387–91 (1968).

Hoel, D. G., and M. Mazumdar, "An Extension of Paulson's Selection Procedure," *Ann. Math. Statist.* **39** (December 1968).

Hoel, D. G., *A Sequential Procedure for Selecting the Population with the Smallest Variance from k Normal Populations,* ORNL-4355 (October 1968).

Hoel, D. G., and M. Maxumdar, "A Class of Sequential Tests for an Exponential Parameter," submitted to *J. Am. Statist. Assoc.*

Hoel, D. G., "Some Modifications and Applications of Wald's OC Formula," submitted to *Ann. Inst. Statist. Math.*

Hoel, D. G. (with D. P. Gaver, Jr.), "Comparison of Certain Small-Sample Poisson Probability Estimates," submitted to *Technometrics.*

Householder, A. S., "Moments and Characteristic Roots, II," *Numerische Math.* **11**, 126–28 (1968).

Householder, A. S., "Bigradients and the Problem of Routh and Hurwitz," *SIAM Rev.* **10**(1), 56–66 (January 1968).

Householder, A. S., "Norms and the Localization of Roots of Matrices," *Bull. Am. Math. Soc.* **74**(5), 816–30 (September 1968).

Joseph, A. F., *Program GAMSPEC 3,* ORNL-CF-68-11-22 (Nov. 15, 1968).

Lee, E. J., "Amicable Numbers and the Bilinear Diophantine Equation," *Math. Comput.* **22**, 181–87 (January 1968).

Lee, E. J., "On Divisibility by Nine of the Sums of Even Amicable Pairs," submitted to *Math. Comput.*

Lee, E. J., *Calculation of Properties of Steam from the N.E.L. Steam Tables,* ORNL-TM-2045 (June 1968).

Lee, E. J., "A Note on the Distribution of Amicable Pairs," submitted to *Math. Comput.*

Lever, W. E., *Part 1 – the Limiting Distribution of the Likelihood Ratio Statistic* $-2ln\lambda_n$, *Under Local Alternatives; Part 2 – Minimum Average Risk Decision Procedures for the Noncentral Chi-Square Distribution,* ORNL-4225 (March 1968).

Lever, W. E. (with D. L. Groves and T. Makinodan), "Interaction of Cell Types in the Generation of Antibody-Forming Cells," submitted to *Nature* (October 1968).

Lever, W. E., "Confidence Limits for Quantiles of Mortality Distributions," accepted for publication as a Note in *Biometrics* (October 1968).

Mitchell, T. J. (with N. R. Draper), "Construction of the Set of 256-Run Designs of Resolution $\geqslant 5$ and the Set of Even 512-Run Designs of Resolution $\geqslant 6$ With Special Reference to the Unique Saturated Designs," *Ann. Math. Statist.* **39**, 246–55 (1968).

Nelson, Paul, and T. S. Kress, *Theory of an Initial-Boundary Value Problem Occurring in the Study of Fission-Product Deposition,* ORNL-4277 (September 1968).

Nestor, C. W., Jr., *A Comparison of Nonrelativistic Hartree-Fock Calculations for Selected Atoms and Ions,* ORNL-TM-2251 (June 1968).

Nestor, C. W., Jr., *CAL Root-Finding Program,* ORNL-CF-68-6-47 (June 25, 1968).

Nestor, C. W., Jr., *Program for Input Form Preparation,* ORNL-CF-68-11-11 (Nov. 11, 1968).

Nestor, C. W., Jr., and T. C. Tucker (with T. A. Carlson and F. B. Malik), "Calculation of Electron Shake-Off for Elements from Atomic Number Equals 2–92 with the Use of Self-Consistent Field Wave Functions," *Phys. Rev.* **169**, 27 (1968).

Nestor, C. W., Jr., K. T. R. Davies, S. J. Krieger, and M. Baranger, "An Effective Nucleon-Nucleon Potential for Use in Nuclear Hartree-Fock Calculations," *Nucl. Phys.* **A113**, 14 (1968).

Opstrup, Shirley B., and Ann S. Klein, *ORNL Journal Control System,* ORNL-4314 (in press).

Stallmann, F. W., "Numerical Solution of Integral Equations," submitted to *Numerische Math.*

Stewart, G. W., III, "On the Continuity of the Generalized Inverse," *J. Soc. Ind. Appl. Math.* (in press).

Stewart, G. W., III, "On Some Methods for Solving Equations Related to Schroder's Iterations," submitted to *Soc. Ind. Appl. Math. Rev.*

Stewart, G. W., III, *Some Topics in Numerical Analysis,* ORNL-4304 (September 1968) (University of Tennessee Ph.D. thesis).

Stewart, G. W., III, "Accelerating the Orthogonal Iteration for the Eigenvectors of a Hermitian Matrix," submitted to *Numerische Math.*

Stewart, G. W., III, "Error Analysis of the Algorithm for Shifting the Zeros of a Polynomial by Synthetic Division," submitted to *Math. Comput.*

Stewart, G. W., III, "On Lehmer's Method for Finding the Zeros of a Polynomial," submitted to *Math. Comput.*

Sullivan, J. G., *Branching in Fortran,* ORNL-TM-1998 (September 1967).

Sullivan, J. G., *The Effect on Program Execution of the Use of Fortran Multisubscripted Variables,* ORNL-TM-1969 (September 1967).

Sullivan, J. G., *Removal of Constant Terms from DO Loops,* ORNL-TM-2007 (Oct. 25, 1967).

Sullivan, J. G., *Multiple-Precision Arithmetic,* ORNL-CF-68-4-16 (Apr. 10, 1968).

Sullivan, J. G., *Fortran Input/Output for Tapes and Direct Access Devices,* ORNL-TM-2142 (February 1968).

Sullivan, J. G., R. M. Rush, and N. B. Alexander, *Printer Plotting Package,* ORNL-CF-68-6-26 (June 17, 1968).

Tucker, T. C., L. D. Roberts, C. W. Nestor, Jr., T. A. Carlson, and F. B. Malik, "A Relativistic Self-Consistent Field Calculation of the Wave Functions, Eigenvalues, Isotope Shifts and the 6*S* Hyperfine Splitting Coupling Constant as a Function of Pressure for Metallic Gold in the Wigner-Seitz Model," submitted to *Phys. Rev.*

Tucker, T. C., L. D. Roberts, C. W. Nestor, Jr., T. A. Carlson, and F. B. Malik, "Calculation of the Electron Binding Energies and X-Ray Energies for the Superheavy Elements 114, 126, and 140 Using Relativistic Self-Consistent-Field Atomic Wave Functions," *Phys. Rev.* **174**, 118 (1968).

Uppuluri, V. R. R. (with J. S. Olson), "Ecosystem Maintenance and Transformation Models as Markov Processes with Absorbing Barriers," pp. 31–35 in *Some Mathematical Models in Biology,* University of Michigan, 1967.

Uppuluri, V. R. R. (with A. S. Paulson), "A Bivariate Geometric Distribution," (Abstract) *Technometrics* **10**, 425 (1968).

Uppuluri, V. R. R., "Information Theory," by R. B. Ash (book review), *Rev. Intern. Statist. Inst.* **36**, 95 (1966).

Uppuluri, V. R. R., *Prime Numbers Treated as a Markov Chain,* ORNL-TM-2317 (August 1968).

Uppuluri, V. R. R., and J. A. Carpenter, "The Inverse of a Matrix Occurring in First-Order Moving Average Models," accepted for publication in *Sankhya, Ser. A,* **31** (1968).

Uppuluri, V. R. R., and John A. Carpenter, "An Inversion Method for Band Matrices," submitted for publication to *J. Math. Anal. Appl.*

Uppuluri, V. R. R., and John A. Carpenter, "Numbers Generated by the Function exp $(1 - e^x)$," submitted for publication to *Fibonacci Quart.*

Uppuluri, V. R. R., and J. R. Doner, "A Markov Chain Structure for Riff Shuffling," submitted to *SIAM J. Appl. Math.*

Warner, G. G. (with A. M. Craig, Jr.), *ALGAM, a Computer Program for Estimating Internal Dose from Gamma-Ray Sources in a Man Phantom,* ORNL-TM-2250 (June 1968).

Mathematics Division Lectures and Papers

Boughner, R. T. (with R. G. Alsmiller, Jr.), "Solar Neutron Transport in the Earth's Atmosphere," American Nuclear Society Meeting, Toronto, Canada, June 9–13, 1968.

Beauchamp, J. J., "Statistics and Probability," Advanced Health Physics Course, Oak Ridge Associated Universities, Special Training Division, Oak Ridge, Tennessee, June 11, 1968.

Durfee, R. C., "Multiple Factor Analysis," Association for Computing Machinery Mid-Southeast Chapter Meeting, Gatlinburg, Tennessee, November 1, 1968.

Gardiner, D. A., "Statistical Solutions to Some Problems in Nuclear Energy Research," European Meeting 1968, Amsterdam, Netherlands, September 2, 1968.

Gosslee, D. G., "Analysis of Microspectrophotometric Measurements of DNA of Megakaryocytes," Annual Symposium on Biomathematics and Computer Science in the Life Sciences, Houston, Texas, March 1968.

Gosslee, D. G., "Reliability, Redundancy and Golf," Tennessee Section of The American Society for Quality Control, Oak Ridge, Tennessee, May 1968.

Gove, N. B., "Pattern Recognition," Talladega College, Talladega, Alabama, October 18, 1967; Pennsylvania Military College, Chester, Pennsylvania, December 4, 1967; Shippensburg State College, Shippensburg, Pennsylvania, December 5, 1967; East Carolina University, Greenville, North Carolina, May 9, 1968.

Gove, N. B., "Computer Handling of Nuclear Data," First International CODATA Conference on the Generation, Collection, Evaluation, and Dissemination of Numerical Data for Science and Technology, Frankfurt, Germany, July 2, 1968.

Gove, N. B., "Physics and Interdisciplinary Research," Gordon Research Conference on Scientific Information Problems in Research, New London, New Hampshire, July 15, 1968.

Hoel, D. G., and T. J. Mitchell, "A General Model for the 'Stem Cell' Population of Cell Renewal Systems," Annual Southeastern Sectional Meeting of the Society for Experimental Biology and Medicine, Oak Ridge, Tennessee, November 2, 1968.

Hoel, D. G., "Discussion of 'Engineering Analysis of Experimental Data' by Robert P. Benedict," Annual Meeting of the American Society of Mechanical Engineers, New York, December 2, 1968.

Householder, A. S., "The Numerical Solution of an Algebraic or Transcendental Equation," University of Michigan, Ann Arbor, June 1968.

Householder, A. S., "Some Recent Developments in the Numerical Treatment of an Algebraic Equation," Boeing Scientific Research Laboratories, Seattle, Washington, December 1968.

Householder, A. S., "Bezoutiants and the Problem of Routh and Hurwitz," Boeing Scientific Research Laboratories, Seattle, Washington; and The University of Washington, Seattle, December 1968.

Joseph, A. F., "Gamma-Ray Spectrum Analysis," Association for Computing Machinery Mid-Southeast Chapter Meeting, Gatlinburg, Tennessee, November 1, 1968.

Lever, W. E., "The Limiting Distribution of the Likelihood Ratio Statistic $-2\ln\lambda_n$, Under Local Alternatives," Department of Statistics, The Florida State University, Tallahassee, February 14, 1968.

Mitchell, T. J., "A Method of 'Repairing' a Response Surface Design," Ninth Annual Quality Control Clinic, American Society for Quality Control, Knoxville, Tennessee, March 15–16, 1968.

Mitchell, T. J., "Repairing Response Surface Designs," Spring Meeting of the American Statistical Association, Biometrics Society, and Institute of Mathematical Statistics, Blacksburg, Virginia, April 8–10, 1968.

Mitchell, T. J., D. A. Gardiner, M. A. Kastenbaum, and C. C. Congdon, "Factorial Design – Response Surface Study of Mortality from Secondary Disease in Mouse Radiation Chimeras," 52d Annual Meeting of the Federation of American Societies for Experimental Biology, Atlantic City, New Jersey, April 15–20, 1968.

Mitchell, T. J., "Experimental Design," U.S. Naval Research Conference, Oak Ridge, Tennessee, June 12, 1968.

Mitchell, T. J. (with N. R. Draper), "Evolutionary Operation and Response Surface Methodology," Short Course, Engineering and Physical Sciences Extension, University of California at Los Angeles, Los Angeles, California, July 8–13, 1968.

Nelson, Paul (with R. B. Perez, G. M. Watson, R. M. Carroll, and O. Sisman), "A Dynamic Method to Study In-Pile Fission-Gas Release," American Nuclear Society Meeting, Toronto, Canada, June 9–13, 1968.

Uppuluri, V. R. R., "Statistics at a Modern Research Laboratory," Sweet Briar College, Sweet Briar, Virginia, February 12, 1968.

Uppuluri, V. R. R., "Mathematical Models of Growth," ORNL-UT Biomedical Program, Oak Ridge, Tennessee, April 19, 1968.

Uppuluri, V. R. R., "Mathematical Models in Ecology," Ecology Graduate Program, University of Tennessee, Oak Ridge, Tennessee, May 7, 1968.

Uppuluri, V. R. R., "Random Difference Equations," Applied Mathematics Division, Argonne National Laboratory, Argonne, Illinois, October 3, 1968.

Uppuluri, V. R. R., "A Probability Distribution Arising in a Riff-Shuffle," American Association for the Advancement of Science, Annual Meetings, Dallas, Texas, December 26, 1968.

Warner, G. G. (with W. S. Snyder and Mary R. Ford), "A Study of Individual Variation of Excretion of Plutonium by Man and of Its Significance in Estimating the Systemic Burden," 13th Annual Bioassay and Analytical Chemistry Meeting, University of California, Berkeley, October 12–13, 1967.

Warner, G. G. (with W. S. Snyder), "Dose to Body Organs from a Source of Photons in One of Them – Validity of the Reciprocity Theorem," Premier Congres Europeen de Radioprotection, Menton, France, October 9–11, 1968.

ORNL-4385
UC-32 – Mathematics and Computers

INTERNAL DISTRIBUTION

1. Biology Library
2–4. Central Research Library
5–6. ORNL – Y-12 Technical Library
Document Reference Section
7–26. Laboratory Records Department
27. Laboratory Records, ORNL R.C.
28. N. B. Alexander
29. C. L. Allen
30. D. W. Altom
31. D. E. Arnurius
32. J. J. Beauchamp
33. N. A. Betz
34. D. S. Billington
35. W. J. Blot
36. C. J. Borkowski
37. K. O. Bowman
38. G. E. Boyd
39. R. B. Briggs
40. J. K. Brinkley
41. T. D. Calton
42. J. A. Carpenter
43. H. P. Carter
44. S. J. Chang
45. A. M. Craig, Jr.
46. A. H. Culkowski
47. F. L. Culler
48. Eila Cutler
49. N. E. Dean
50. M. J. Dolan
51. R. C. Durfee
52. M. H. Eckart
53. M. B. Emmett
54. J. B. Estill
55. M. Feliciano
56. J. L. Fowler
57. J. H. Frye, Jr.
58. J. L. Gabbard
59–63. D. A. Gardiner
64. D. G. Gosslee
65. N. B. Gove
66. W. R. Grimes
67. G. K. Haeuslein
68. C. E. Hammons
69. H. J. Hargis
70. T. L. Hebble
71. D. G. Hoel
72. R. B. Hofstra
73. A. Hollaender
74–78. A. S. Householder
79. N. C. Hull
80. J. C. Johns
81. W. H. Jordan
82. A. J. Joseph
83–87. M. A. Kastenbaum
88. M. A. Kau
89. M. T. Kelley
90. S. P. King
91. L. A. Kurtz
92. J. A. Lane
93. C. E. Larson
94. E. Leach
95. E. J. Lee
96. C. S. Lever
97. W. E. Lever
98. J. L. Liverman
99. R. S. Livingston
100. H. G. MacPherson
101. F. C. Maienschein
102. J. D. McDowell
103. W. L. McMullen
104. S. H. Merriman
105. F. L. Miller, Jr.
106. T. J. Mitchell
107. K. Z. Morgan
108. Paul Nelson, Jr.
109. C. W. Nestor, Jr.
110. H. M. Nicholson, Jr.
111. W. H. Olson
112. J. E. Parham
113–114. R. B. Parker
115. C. D. Paulk
116. M. M. Perardi
117. C. A. Preskitt
118. R. P. Rannie
119. H. E. Seagren
120. E. D. Shipley
121. V. A. Singletary
122. M. J. Skinner
123. D. B. Smith
124. A. H. Snell

125. W. S. Snyder
126. J. E. Stockton
127. J. G. Sullivan
128. E. H. Taylor
129. C. L. Thompson, III
130. T. C. Tucker
131. V. R. R. Uppuluri
132. C. P. Walker
133. G. G. Warner
134. A. M. Weinberg
135. C. S. Williams
136. D. R. Winkler
137. E. O. Wollan
138. Gale Young

EXTERNAL DISTRIBUTION

139. C. O. Alford, Oak Ridge Associated Universities, Oak Ridge, Tennessee
140. Gilbert W. Beebe, Follow-Up Agency, National Academy of Sciences, Washington, D.C.
141. Joseph M. Cameron, National Bureau of Standards, Washington, D.C.
142. George B. Darling, Director, Atomic Bomb Casualty Commission, Hiroshima, Japan
143. D. J. Dessart, Mathematics Department, University of Tennessee, Knoxville, Tennessee
144. C. W. Edington, Atomic Energy Commission, Washington, D.C.
145. Sidney Fernbach, Computation Division, Lawrence Radiation Laboratory, Livermore, California
146. Wallace Givens, Argonne National Laboratory, Argonne, Illinois
147. W. W. Grigorieff, Oak Ridge Associated Universities, Oak Ridge, Tennessee
148. John J. Harley, Director, Health and Safety Laboratory, Atomic Energy Commission, New York, New York
149. Internal Computation Centre, Casella postale no. 10053, Zona dell'E.U.R., Rome, Italy
150. John L. Jaech, Battelle Northwest, Richland, Washington
151. J. S. Kirby-Smith, U.S. Atomic Energy Commission, Washington, D.C.
152. William L. Morris, Department of Mathematics, University of Houston, Houston, Texas
153. John R. Pasta, Dept. of Computer Science, University of Illinois, Urbana, Illinois
154. Harry Polachek, U.S. Atomic Energy Commission, Washington, D.C.
155. Yoshio Shimamoto, Brookhaven National Laboratory, Upton, Long Island, New York
156. L. R. Shenton, Computer Center, University of Georgia, Athens, Georgia
157. C. V. L. Smith, U.S. Atomic Energy Commission, Washington, D.C.
158. F. W. Stallmann, Mathematics Department, University of Tennessee, Knoxville, Tennessee
159. G. W. Stewart III, Dept. of Computer Sciences, University of Texas, Austin, Texas
160. John R. Totter, U.S. Atomic Energy Commission, Washington, D.C.
161. R. W. Touchberry, U.S. Atomic Energy Commission, Washington, D.C.
162. Jack Vanderryn, U.S. Atomic Energy Commission, Washington, D.C.
163. D. J. Rose, Department of Nuclear Engineering, Room 24-207, Massachusetts Institute of Technology, Cambridge, Massachusetts
164. Documents Room, Computing Center, University of Notre Dame, Notre Dame, Indiana
165. J. A. Swartout, Union Carbide Corporation, New York, N.Y.
166. Laboratory and University Division, AEC, ORO

167–416. Given distribution as shown in TID-4500 under Mathematics and Computers category (25 copies – CFSTI)