J. Soc. Indust. Appl. Math.
Vol. 6, No. 4, December, 1958
Printed in U.S.A.

REACTOR CRITICALITY AND NONNEGATIVE MATRICES*

GARRETT BIRKHOFF AND RICHARD S. VARGA

1. Multiplicative processes. In this paper, we apply the Perron-Frobenius theory of nonnegative matrices to provide a rigorous mathematical basis for the physical concept of "criticality," and for typical computational interpretations of this concept.[1] Though the applications given are almost entirely new,[2] few new mathematical results are given, and so many proofs will be omitted.

The proper mathematical setting for many concepts of chain reaction theory is provided by the concept of a (finite, stationary) *multiplicative process*, which we now define.

Definition. A real square matrix $P = \| p_{ij} \|$ will be called *nonnegative* if and only if

$$p_{ij} \geqq 0 \tag{1a}$$

for all i, j. Similarly, a real square matrix $Q = \| q_{ij} \|$ will be called *essentially nonnegative* if and only if

$$q_{ij} \geqq 0 \qquad (i \neq j). \tag{1b}$$

A *discrete* multiplicative process with a finite number m of "states" is a system of difference equations of the form

$$N_i(r + 1) = \sum_{j=1}^{m} p_{ij} N_j(r) \qquad (i = 1, \cdots, m), \tag{2}$$

where P is nonnegative. A *continuous* multiplicative process is any system of ordinary differential equations of the form

$$dN_i/dt = \sum_{j=1}^{m} q_{ij} N_j(t) \qquad (i = 1, \cdots, m), \tag{3}$$

where Q is essentially nonnegative.

Such multiplicative processes generalize the usual concept of a stochastic or *Markoff process* [5, §4]. By definition [2, pp. 24, 206], these are multipli-

* Received by the editors January 9, 1958 and in revised form May 23, 1958.

[1] The authors wish to thank the referee for suggesting many improved formulations.

[2] S. Ulam [1] seems to be the first to have recognized the connection between discrete multiplicative (alias "branching") processes and reactor theory. Some related ideas were also ruggested in Report UMM-144 to Westinghouse Electric Corp. by W. F. Bauer, J. F. Carr, R. Dames, G. Graves, and V. Larrowe, (Oct. 1953), and by M. K. Butler and J. M. Cook [13].

cative processes for which (1a) and (1b) are supplemented by the conditions

(4a) $$\sum_{i=1}^{m} p_{ij} = 1 \qquad (j = 1, \cdots, m),$$

or

(4b) $$\sum_{i=1}^{m} q_{ij} = 0 \qquad (j = 1, \cdots, m),$$

respectively. Multiplicative processes are more general than Markoff processes in that they permit birth (fission) and death (escape or absorption), as well as pure diffusion and slowing down.

2. Relation to chain reactions. A natural connection between continuous multiplicative processes and nuclear chain reactions is suggested by the following model for *multigroup diffusion.*

The expected neutron distribution in a nuclear reactor can be subdivided into m "cells" (or neutron "states") in various ways. Commonly, the subdivision is into so-called space and lethargy groups, corresponding to different neutron positions and velocities. In transport theory, there are direction groups as well. If $N_i(t)$ denotes the number of neutrons expected in cell i at time t, and q_{ij} denotes the average rate of production of neutrons in cell i from a single neutron in cell j, then (3) holds [4, §13.2]. The negative terms on the diagonal correspond to loss of neutrons from a cell by leakage, absorption, and slowing down.

Caution. The preceding model assumes a *time-independent* neutron environment. As usual[3] in criticality calculations, one neglects changes in the matrix Q due to depletion, poisoning, expansion by fission heating, etc.

A natural interpretation of *discrete* multiplicative processes is furnished, similarly, by iterative *computation* schemes for calculating critical flux distributions on high-speed computing machines (§10). In nuclear reactor theory [4, 8], it is usual to take the neutron *flux* $\varphi_i = v_i N_i$ as the dependent variable, instead of the neutron density N_i. However, this substitution simply replaces a process P or Q by a diagonally equivalent one $D^{-1}PD$ or $D^{-1}QD$, where D is a positive diagonal matrix. Thus, this change does not alter any of the properties discussed below.

In the simplest case of thermal reactors, if epithermal fissions are neglected and all fission neutrons are assumed fast, we get multigroup equations of the special form

(5) $$A_k\varphi_k = B_k\varphi_{k-1} \qquad (k = 2, \cdots, n),$$

(5′) $$A_1\varphi_1 = \nu B_1\varphi_n \qquad (\nu > 0).$$

[3] See [4], Chs. VII and XII; also [8].

In (5) and (5′), $\varphi_n = \varphi_n(x)$ represents the thermal (slow) neutron flux distribution, φ_1 the fast neutron flux, and $\varphi_2, \cdots, \varphi_{n-1}$ intermediate neutron flux distributions, arranged in order of decreasing velocity (increasing lethargy). The flux φ_k arises from the "slowing down" and scattering of neutrons in φ_{k-1}, which may thus be regarded as a (distributed) source. The matrices A_k, B_k involved will be defined in §9.

It will be shown in §9 that the A_i in (5) and (5′) are nonsingular, and that the matrices A_k^{-1}, B_k are nonnegative. Hence, the composite matrix

$$\begin{pmatrix} 0 & \cdots & 0 & \nu A_1^{-1}B_1 \\ A_2^{-1}B_2 & & 0 & 0 \\ \vdots & & \vdots & \\ 0 & \cdots & A_n^{-1}B_n & 0 \end{pmatrix} \tag{6}$$

defines, for each $\nu > 0$, a (discrete) multiplicative process from (5) and (5′).

Of great physical interest is the spatial *thermal flux distribution* which, for the simple model (5) and (5′), satisfies (in the critical case)

$$\varphi_n = \nu(A_n^{-1}B_n \cdots A_1^{-1}B_1)\varphi_n = T\varphi_n, \tag{7}$$

where

$$T = \nu(A_n^{-1}B_n \cdots A_1^{-1}B_1). \tag{7′}$$

This evidently defines a second multiplicative process from (5) and (5′).

3. General properties. Nonnegative and essentially nonnegative matrices have a few general properties. As these are either known[4] or obvious, we shall state them without proof.

LEMMA 1. *Any nonnegative matrix P has a nonnegative eigenvector, with nonnegative eigenvalue.*

LEMMA 2. *The matrix Q is essentially nonnegative if and only if e^{Qt} is nonnegative for all $t \geqq 0$.*

Similarly, the matrix Q defines a continuous Markoff process (4b), if and only if $P = e^{Qt}$ defines a discrete Markoff process (4a) for any $t > 0$.

LEMMA 3. *For any square matrix Q, with eigenvalues μ_k, the matrices Q, e^{Qt} and $sI + Q$ have the same eigenvectors for any $t \neq 0$ or s. Moreover, the eigenvalues of e^{Qt} are the $\lambda_k = e^{\mu_k t}$, and those of $(sI + Q)^n$ are the $(s + \mu_k)^n$ for any integer n.*

This result is obvious when Q is put into its Jordan canonical form. Combining with Lemmas 1 and 2, we get the

COROLLARY. *Any essentially nonnegative matrix has a nonnegative eigenvector.*

[4] See [9, 11]. The irreducible case was originally treated in [3].

However, for most applications, the matrices P, Q of (1a) and (1b) must satisfy further conditions, of which the following are typical.

Definition. A matrix P will be called *positive* if and only if

$$p_{ij} > 0 \tag{8}$$

for all i, j. A matrix Q will be called *irreducible*[5] if, for any i, j, there exists a finite sequence of indices $k(0) = i, \cdots, k(r) = j$ such that $q_{k(h-1),k(h)} \neq 0$ for $h = 1, \cdots, r$.

In the first interpretation of §2, irreducibility is evidently equivalent to the possibility of ultimate transition from any cell to any other; thus, it implies the geometrical connectedness of the reactor.

LEMMA 4. *$\exp(tQ)$ is positive for all $t > 0$ if and only if Q is essentially nonnegative and irreducible.*

Hence, we define Q to be *essentially positive* if and only if it is essentially nonnegative and irreducible. Again,

LEMMA 5. *If an $n \times n$ matrix Q is essentially positive, then $(sI + Q)^m$ is positive, for all sufficiently large s and m.*

THEOREM 1. *Any essentially positive matrix Q has a unique*[6] *strictly positive eigenvector φ_1, with real simple*[7] *eigenvalue $\mu_1 = M$. Moreover, $\mu_1 > \mathrm{Re}\{\mu_j\}$ for any other eigenvalue μ_j of Q.*

Proof. Since e^{Qt} is positive (Lemma 4), this follows from Lemma 3, and the main result of Perron and Frobenius about positive matrices,[8] which asserts

$$|e^{\mu_1 t}| = e^{\mu_1 t} > |e^{\mu_j t}| = e^{\mathrm{Re}\{\mu_j t\}},$$

if $t > 0$ and $j > 1$.

THEOREM 1′. *Any irreducible, nonnegative matrix P has a unique strictly positive eigenvector φ_1, with positive simple eigenvalue $L = \lambda_1$. Moreover, $\lambda_1 \geqq |\lambda_j|$ for any other eigenvalue λ_j of P, and any nonnegative eigenvector is a scalar multiple of φ_1.*

For this and other results due to Frobenius, see [3].

Definition. The *spectral norm* of a matrix A is the maximum of the absolute values of its eigenvalues λ_j. The *spectral prenorm* of A is the maximum of the real parts of these λ_j.

The following result follows directly from the previous definition and Theorems 1 and 1′.

---

[5] This is also called "indecomposable." See [2, 6, 9, 11] and H. Geiringer, Reissner Anniversary Volume, Ann Arbor, 1949, pp. 365–93.

[6] Up to scalar factors.

[7] An eigenvalue μ of a matrix Q is "simple" if $(\lambda - \mu)$ is not a repeated factor of the determinant $| \lambda I - Q |$.

[8] For a new, geometric proof, see G. Birkhoff, "Extensions of Jentzsch's Theorem," Trans. Amer. Math. Soc., 85 (1957), pp. 219–27.

COROLLARY. *The spectral norm of any nonnegative irreducible matrix P is the L of Theorem 1′; the spectral prenorm of any essentially positive matrix Q is the M of Theorem 1. The spectral norm of e^{Qt} is e^{Mt}, if $t > 0$.*

4. Importance vector. Obviously, (off-diagonal) nonnegativity and irreducibility are unaffected when a matrix A is replaced by its transpose A'. Hence the transpose of any irreducible multiplicative process is also an irreducible multiplicative process. Hence, in Theorems 1 and 1′, P' and Q' also admit positive (column) eigenvectors with the properties described there. Moreover, since $|P - \lambda I| = |P' - \lambda I|$ and $|e^{Qt} - \lambda I| = |e^{Q't} - \lambda I|$, clearly P and P' have the same spectral norm $L = \mu_1$, while Q and Q' have the same spectral prenorm $M = \mu_1 = \max[\mathrm{Re}\{\mu_k\}]$. Hence, if $\boldsymbol{F}'$ is the positive (column, right-) eigenvector of P', $P'\boldsymbol{F}' = L\boldsymbol{F}'$—and similarly for Q.

Taking transposes, we get $\boldsymbol{F}P = L\boldsymbol{F}$, and similarly Q has a positive *left*-eigenvector satisfying $\boldsymbol{F}Q = M\boldsymbol{F}$. We can summarize these facts as follows.

Definition. For any irreducible process (2) or (3), the positive *left*-eigenvector of P (or Q) is called the *importance vector* of the process.

THEOREM 2. *The importance vector $\boldsymbol{F}$ of any irreducible multiplicative process (2) or (3) satisfies respectively $\boldsymbol{F}P = L\boldsymbol{F}$ or $\boldsymbol{F}Q = M\boldsymbol{F}$.*

For any discrete process (2), we have

$$\boldsymbol{F}\cdot\boldsymbol{N}(r+1) = \boldsymbol{F}\cdot P\boldsymbol{N}(r) = L^{r+1}\boldsymbol{F}\cdot\boldsymbol{N}(0). \tag{9}$$

It is a corollary that the real hyperplane $\boldsymbol{F}\cdot\boldsymbol{N} = 0$ is *invariant* under the process. Moreover, since $\boldsymbol{F}\cdot\boldsymbol{\varphi} > 0$ for any positive $\boldsymbol{F}$ and $\boldsymbol{\varphi}$, this hyperplane H is *complementary*[9] to the line containing $\boldsymbol{\varphi}_1$, the positive eigenvector of P.

Similarly, for any continuous process (3),

$$\boldsymbol{F}\cdot\boldsymbol{N}(t) = e^{Mt}\boldsymbol{F}\cdot\boldsymbol{N}(0); \tag{10}$$

the hyperplane $\boldsymbol{F}\cdot\boldsymbol{N} = 0$ is again complementary to $\boldsymbol{\varphi}_1$, and invariant under the process.

If $\boldsymbol{F} = (F_1, \cdots, F_m)$ is the importance vector of an irreducible multiplicative process with matrix P, we may let D be the diagonal matrix with positive diagonal entries $(F_1, \cdots, F_m)$. For this D, one easily shows[10] that $S = L^{-1}DPD^{-1}$ satisfies (4a)—i.e., is a *stochastic* matrix. Therefore,

THEOREM 3. *If P is an irreducible nonnegative matrix, with spectral norm L, then $P = L(D^{-1}SD)$, where D is a positive diagonal matrix, and S satisfies* (4a).

---

[9] In the sense of Birkhoff and MacLane *Survey of modern algebra*, rev. ed., p. 185.

[10] This observation is due to Kolmogoroff; see N. Dmitriev and E. Dynkin, Dokl. Akad. Nauk SSSR, 10 (1946), pp. 167–184.

Similarly, one can prove

THEOREM 3'. *If Q is essentially positive with spectral prenorm M, then $Q = (D^{-1}TD) + MI$, where D is a positive diagonal matrix and T satisfies* (4b).

This result makes many facts about irreducible multiplicative processes follow from the corresponding results about (irreducible) Markoff processes.

Minimax property. Now let Q be essentially positive, and let E^+ be the set of all vectors with strictly positive components. The following result about the "spectral prenorm" M of Q can then be proved as in[11] [12], §13:

$$(11)\quad M(Q) = \sup_{X\in E^+}\{\inf_{Y\in E^+} XQY/XY\} = \inf_{Y\in E^+}\{\sup_{X\in E^+} XQY/XY\},$$

the minimax being assumed when X is the importance (row) vector and Y is the dominant (column) eigenvector of Q.

A curious related theorem holds in game theory: if $XY = XIY$ is replaced by XJY, where J is the (positive) matrix all of whose entries are unity, then the eigenvectors in question define "optimal strategies" for the game with "payoff matrix" Q.

An analogous characterization of the spectral norm L of a nonnegative irreducible matrix P is known [9, p. 648], and

$$(12)\qquad L(P) = \sup_{x\in E^+}\left\{\inf_i \frac{(Px)_i}{x_i}\right\} = \inf_{x\in E^+}\left\{\sup_i \frac{(Px)_i}{x_i}\right\},$$

the minimax being assumed for x the dominant eigenvector of P.

5. Semi-irreducible matrices. Reactor calculations based on multigroup equations like (5) and (5') commonly define multiplicative processes like (6) or (7) which are *semi*-irreducible in the following sense.

Definition. A nonzero matrix P is *semi-irreducible* when, for each j, either all $p_{ij} = 0$, or, for all i, there exists a finite sequence, $k(0) = i, \cdots, k(r) = j$, such that $p_{k(h-1),k(h)} \neq 0$ for $h = 1, 2, \cdots, r$.

The indices of the first kind generate the *null space*[12] of P, while those of the second generate its range. After a permutation Λ of indices, transforming P into $\Lambda P \Lambda^{-1}$, the matrix P assumes the form of Fig. 1, where P_2 is an irreducible square submatrix. One can show that Theorem 1' also applies to semi-irreducible matrices.

THEOREM 1''. *Any semi-irreducible, nonnegative matrix P has a positive eigenvector φ_1 with positive simple eigenvalue $L = \lambda_1$. Moreover, $\lambda_1 \geqq |\lambda_i|$*

[11] In [12], p. 28, line 6, the inequality sign should be reversed.

[12] The null space corresponds physically to that part of the thermal flux in the reflector. For a fuller discussion of this and many other points, see [12]. Continuous multiplicative processes which are semi-irreducible are irreducible for reactors.

for any other eigenvalue λ_j, *and any nonnegative eigenvector with nonzero eigenvalue is a scalar multiple of* φ_1.

The proof, being straightforward, will be omitted (cf. [12], p. 14). A similar result holds for Q.

$$\begin{pmatrix} 0 & P_1 \\ 0 & P_2 \end{pmatrix}$$

FIG. 1

From Fig. 1, it is evident that the transpose P' of a semi-irreducible nonnegative matrix P is not usually semi-irreducible. Nevertheless, P' has a nonnegative eigenvector with simple eigenvalue L—this follows from Lemma 1, and the fact that the characteristic polynomial of P is the same as that of P'. In the case that P is semi-irreducible, the unique nonnegative eigenvector F of P' corresponding to the simple eigenvalue L will still be called the importance vector of P, and the invariance of the real hyperplane $F \cdot N(r) = 0$ still holds. Moreover, if φ_1 is any positive eigenvector of P, then $F \cdot \varphi_1 > 0$ as before.

6. Criticality. We can now define the concept of criticality, for multiplicative processes defined by irreducible or semi-irreducible matrices.

Definition. A process (3) will be called *subcritical*, *critical*, or *supercritical* according as $M < 0$, $M = 0$, or $M > 0$ in Theorem 1. A process (2) will be called subcritical, critical, or supercritical according as $L < 1$, $L = 1$, or $L > 1$ in Theorems 1′ and 1″. The positive eigenvector of any irreducible or semi-irreducible multiplicative process (2) or (3), normalized to make $\Sigma N_k = 1$, will be called the *dominant distribution.*[13]

Evidently, a semi-irreducible multiplicative process (2) or (3) is critical, if and only if the positive eigenvector φ_1 in its range is invariant, so that $N(r) = \varphi_1$ or $N(t) = \varphi_1$ satisfies (2) or (3), respectively. Markoff processes are always critical, since their transposes admit the invariant left-eigenvector $(1, \cdots, 1)$.

COROLLARY 1.[14] *For any irreducible process* (3),

(*i*) *If supercritical,* $Q\varphi > 0$ *for some* $\varphi > 0$ *and* $Q\varphi \leqq 0$ *for no* $\varphi > 0$,

(*ii*) *If critical,* $Q\varphi = 0$ *for some* $\varphi > 0$, *but* $Q\varphi > 0$ *for no* $\varphi > 0$ *and* $Q\varphi < 0$ *for no* $\varphi > 0$,

(*iii*) *If subcritical,* $Q\varphi < 0$ *for some* $\varphi > 0$, *but* $Q\varphi \geqq 0$ *for no* $\varphi > 0$.

Proof. In all cases, the first statement follows from Theorem 1, letting φ be the positive eigenvector φ_1 of Q. The second statement follows by

[13] In the case of the multiplicative process (6), this (mathematical) critical distribution is proportional to the (physical) critical *flux* distribution. In the case (7), $B_1\varphi_n$ is proportional to the critical fission distribution.

[14] See also Lemma [11], p. 601.

writing $\varphi = c_1\varphi_1 + \psi$, where $F \cdot \psi = 0$ so that $F \cdot \varphi = c_1 F \cdot \varphi_1$; cf. (10). Hence

$$F \cdot Q\varphi = Mc_1(F \cdot \varphi_1) + F \cdot Q = M(F \cdot \varphi).$$

Hence $F \cdot Q\varphi > F \cdot \varphi$, $F \cdot Q\varphi = F \cdot \varphi$, or $F \cdot Q\varphi < F \cdot \varphi$ for all $\phi > 0$, according to whether Q defines a supercritical, critical, or subcritical process (3). Again, since F has all positive entries, $F \cdot \varphi > 0$ if $\varphi > 0$. Hence $F \cdot (Q\varphi - \varphi) \geqq 0$ if $Q\varphi \geqq \varphi$, etc. From this, the second statement of (i)–(iii) above follows.

Using the minimax property of (11), we also have the sharp converse

COROLLARY 2. *An irreducible process* (3):

(i) is supercritical if $Q\varphi > 0$ *for some* $\varphi > 0$,

(ii) is critical if $Q\varphi = 0$ *for some* $\varphi > 0$,

(iii) is subcritical if $Q\varphi < 0$ *for some* $\varphi > 0$.

COROLLARY 1′. *For any semi-irreducible process* (2),

(i) if supercritical, $P\varphi > \varphi$ *for some* $\varphi > 0$, *and* $P\varphi \leqq \varphi$ *for no* $\varphi > 0$,

(ii) if critical, $P\varphi = \varphi$ *for some* $\varphi > 0$, *and* $P\varphi \gtrless \varphi$ *for no* $\varphi > 0$,

(iii) if subcritical, $P\varphi < \varphi$ *for some* $\varphi > 0$, *and* $P\varphi \geqq \varphi$ *for no* $\varphi > 0$.

Since the second conclusion in each of cases *(i)*–*(iii)* is incompatible with the other two cases, we have also the sharp converses

COROLLARY 2′. *A semi-irreducible process* (2):

(i) is supercritical if $P\varphi > \varphi$ *for some* $\varphi > 0$,

(ii) is critical if $P\varphi = \varphi$ *for some* $\varphi > 0$,

(iii) is subcritical if $P\varphi < \varphi$ *for some* $\varphi > 0$.

THEOREM 4. *For any irreducible process* (3),

$$N(t) = Ke^{Mt}\varphi_1 + o(e^{\mu t}), \tag{13}$$

where φ_1 *and* M *are as in Theorem* 1. *In* (13), μ *is any number strictly between* M *and* $\sup_{j>1}\mathrm{Re}\{\mu_j\}$, *and* $K = (F \cdot N(0))/(F \cdot \varphi_1)$, F *being the importance vector and* φ_1 *the positive eigenvector of* Q.

Proof. We can write $N(0) = K\varphi_1 + G$, where

$$F \cdot G = F \cdot N(0) - KF \cdot \varphi_1 = 0.$$

Then, for all $t > 0$,

$$N(t) - Ke^{Mt}\varphi_1 = N(t) - Ke^{Qt}\varphi_1 = e^{Qt}G.$$

But G is in the linear subspace $\mathfrak{M}$, which is invariant under Q and hence under e^{Qt}. On $\mathfrak{M}$, all eigenvalues of e^{Qt} are less than $e^{\mu t}$ in modulus, from which $e^{Qt}G = o(e^{\mu t})$, as $t \to +\infty$, follows. The existence of such μ follows from Theorem 1.

The constant $1/M$ in Theorem 4 may be called the "e-folding time" or "period" ([4], p. 293) of the process (3).

COROLLARY. *For any irreducible process* (3), *if* $N(0) > 0$, *then*

$$\lim_{t\to+\infty} N(t) = \begin{cases} (+\infty, \cdots, +\infty) & \text{if (3) is supercritical} \\ K\varphi_1 & \text{if (3) is critical} \\ (0, \cdots, 0) & \text{if (3) is subcritical} \end{cases} \tag{14}$$

The mathematical *expectation* $\| N(t) \| = \Sigma N_j(t)$ of the neutron flux $N(t)$ should not be confused with the statistical distribution of expected neutron fluxes. The distinction is strikingly illustrated by a theorem of Sevastyanov [5, Theorem 5], which asserts that, in critical processes as defined above, the neutron flux will die out with probability one! This paradoxical fact corresponds to the cumulative possibility that, in infinite time, *all* the neutrons of any one generation will be absorbed. In practice, this possibility can be ignored in any reasonable time interval.

7. Cyclic and primitive matrices. Since $|\lambda_j| = \lambda_1$ is possible in Theorems 1′ and 1″, the theory of *discrete* irreducible processes is less simple. Though it could be derived from Theorem 3 and the literature on stochastic matrices (e.g., [2, 6]), it is more instructive to go back to basic principles of algebra. These involve the concept of a cyclic matrix.

Let P be any $m \times m$ matrix. By a *cycle* of length l, is meant a sequence of l nonzero entries

$$p_{ij_1}, p_{j_1j_2}, \cdots, p_{j_{l-1}i}.$$

Following Frobenius [3], we make the

Definition. The *index* $k(P)$ of an irreducible matrix P is the greatest common divisor (g.c.d.) of the lengths of its cycles. If $k(P) = 1$, P is *primitive*; if $k(P) > 1$, it is *cyclic.*

Frobenius has shown [3, p. 463] that an irreducible nonnegative matrix P is primitive if and only if some positive power of P (and all higher powers) is positive.

Evidently, the matrix of (6) is cyclic: the length of any cycle of P is a multiple of the number n of lethargy groups considered. Hence $k(P)$ is also a multiple of n. In general [6, p. 162], an $m \times m$ matrix P is cyclic of index n if and only if there exists a permutation Λ of indices such that $\Lambda P \Lambda^{-1}$ has the form of Fig. 2, with square blocks on the main diagonal.

For any such nonnegative irreducible cyclic matrix P of index n, with positive eigenvector $\Phi = (\varphi_1, \cdots, \varphi_n)$, evidently $P^n\Phi = (T_1\varphi_1, \cdots, T_n\varphi_n)$ where $T_k = P_kP_{k-1} \cdots P_1P_n \cdots P_{k+1}$. Hence $L(P_n \cdots P_1) = L(P^n) = |L(P)|^n$, and in particular P is critical if and only if $P_n \cdots P_1$ is critical. Thus, in §2, the multiplicative process represented by (7) and (7′) is critical if and only if the multiplicative process defined by the matrix P of (6) is critical.

$$\begin{pmatrix} 0 & \cdots & 0 & P_1 \\ P_2 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & P_n & 0 \end{pmatrix}$$

FIG. 2

By considering the nonzero entries in $|\lambda I - P|$, one proves similarly without great difficulty [6, p. 166]:

LEMMA 6. *For any $n \times n$ cyclic matrix P of index k,*

$$|\lambda I - P| = \lambda^{\nu}(\lambda^k - \lambda_1^k) \cdots (\lambda^k - \lambda_r^k), \tag{15}$$

where $\nu + kr = n$.

From this result, we deduce the important

COROLLARY. *In Theorem 1″, in order that $\lambda_1 = L > |\lambda_j|$ for $j = 2, \cdots, m$, it is necessary and sufficient that P_2 in Fig. 1 be primitive.*

Combining the first statement with the general theory of matrices whose dominant eigenvalues are simple, we see, much as in the proof of Theorem 4, that (13) applies also to discrete semi-irreducible processes (2). In fact,

THEOREM 4′. *For any semi-irreducible process (2), the trichotomy (14) is valid. If and only if P_2 is primitive (i.e., if and only if $k(P_2) = 1$), we have further*

$$N(r) = KL^r\varphi_1 + o(\rho^r), \qquad K = (F \cdot N(0))/(F \cdot \varphi_1). \tag{16}$$

where φ_1 is the positive eigenvector of P_2, F its importance vector, $L = \lambda_1$ the eigenvalue of φ_1 and ρ is any number strictly between L and $\sup_{j>1}|\lambda_j|$.

The proof of this theorem parallels that of Theorem 4. By analogy with the corollary to Theorem 4, we also have

COROLLARY. *For any semi-irreducible process (2), if $N(0)$ is positive, then*

$$\lim_{r \to +\infty} \Sigma N_j(r) = \begin{cases} +\infty & \text{if (2) is supercritical.} \\ \text{finite} & \text{if (2) is critical.} \\ 0 & \text{if (2) is subcritical.} \end{cases}$$

If P_2 is primitive, then

$$\lim_{r \to \infty} N(r) = \begin{cases} (+\infty, \cdots, +\infty) & \text{if (2) is supercritical.} \\ K\varphi_1 & \text{if (2) is critical.} \\ (0, \cdots, 0) & \text{if (2) is subcritical.} \end{cases}$$

8. Stieltjes matrices. The results of §§3–6 can be used to prove certain facts about the matrices A_i of (5) and (5′). Let $A = -Q$, where Q is essentially nonnegative. We consider the three mutually exclusive possibilities of Theorem 4, Corollary 2:

(a) If Q is *supercritical* (i.e., if $M > 0$), then $A = -Q$ has a nonnegative eigenvector $\varphi_1 \neq 0$ with negative eigenvalue $-M$. It follows that, if A^{-1} exists, $A^{-1} = -Q^{-1}$ has a nonnegative eigenvector with negative eigenvalue $-M^{-1}$. Hence A^{-1} cannot exist and be nonnegative.

(b) If Q is *critical* (i.e., if $M = 0$), then Q and A are singular: their null space includes a nonnegative eigenvector.

(c) If Q is *subcritical* (i.e., if $M < 0$), then the eigenvalues of A are in the half-plane $\mathrm{Re}\{\lambda\} \geqq -M > 0$. Hence A is nonsingular, and A^{-1} has eigenvalues in the *circle* $| \lambda - 1/(2M) | \leqq 1/(2M)$, omitting 0. Therefore all eigenvalues of A^{-1} have a *positive real part.* Furthermore, by Theorem 4, the following integrals are *convergent*, and define a *nonnegative* matrix A^{-1}:

$$\int_0^\infty e^{Qt}\, dt = \int_{-\infty}^0 e^{-Qt}\, dt = -Q^{-1} = A^{-1}.$$

This proves the following result.

THEOREM 5. *Let $A = \| a_{ij} \|$ have nonpositive off-diagonal entries (i.e., let $a_{ij} \leqq 0$ if $i \neq j$). Then A^{-1} exists and is nonnegative if and only if all eigenvalues of A have positive real parts.*

COROLLARY 1. *If A is symmetric and if $a_{ij} \leqq 0$ whenever $i \neq j$, then A^{-1} exists and is nonegative if and only if A is positive definite.*

(For, A is positive definite if and only if its (real) eigenvalues are all positive.)

Appealing to Lemma 4, we get

COROLLARY 2. *In Theorem 5 and Corollary 1 above, A^{-1} is positive if and only if A is irreducible.*

These results sharpen an old Lemma of Stieltjes[15] and some recent improvements on it. Because of this, we shall make the following

Definition. A *Stieltjes matrix* is a symmetric, irreducible matrix with nonpositive off-diagonal entries.

Thus, a matrix A is a Stieltjes matrix if and only if $-A = Q$ is essentially positive and symmetric. We have shown that a Stieltjes matrix A has a positive inverse A^{-1} if and only if $Q = -A$ is subcritical.

9. Multigroup approximations. The usual multigroup approximation to the time-dependent neutron distribution in a heterogeneous reactor involves several functions of position: the diffusion length D_i of the ith lethargy group, its slowing-down cross-section Σ_i, its total cross-section $\Sigma_i' > \Sigma_i$, the fission cross-section Σ_n, and the fission yield ν. The multigroup equations for the flux $\phi_i(x, t)$ of neutrons in the ith lethargy group, with average velocity v_i, are then [4, p. 291]:

[15] T. J. Stieltjes, Acta Math. 9 (1886), pp. 385–400. These results can also be obtained from [11], or from unpublished work of Ky Fan.

$$(17)\quad \frac{\partial \varphi_i}{\partial t} = v_i\left\{\sum_k \frac{\partial}{\partial x_k}\left(D_i \frac{\partial \varphi_i}{\partial x_k}\right) - \Sigma_i'\varphi_i + \Sigma_{i-1}\varphi_{i-1}\right\} \quad (i = 2, \cdots, n),$$

$$(17')\quad \frac{\partial \varphi_1}{\partial t} = v_1\left\{\sum_k \frac{\partial}{\partial x_k}\left(D_1 \frac{\partial \varphi_1}{\partial x_k}\right) - \Sigma_1'\varphi_1 + \nu\Sigma_n\varphi_n\right\}.$$

If the subdivision into lethargy groups, postulated in (17) and (17′), is accompanied by a spatial subdivision into cells (say, into squares or cubes of constant side h), one can approximately represent (17) and (17′) in the form $d\boldsymbol{\varphi}/dt = Q\boldsymbol{\varphi}$ of (3). In fact, one can do this in various ways, depending on the *difference* operator used to approximate the differential operator $\nabla \cdot D_i \nabla = \Sigma\partial(D_i\partial/\partial x_k)/\partial x_k$. Each such approximation leads, in a natural way, to a continuous multiplicative process (3).

In one space dimension, the "best possible" simple approximation is the three-point formula

$$(18)\quad \begin{aligned} \frac{d}{dx}\left[D(x)\frac{d\varphi}{dx}\right] \sim \frac{1}{h^2}\Big\{ & D\left(x + \frac{h}{2}\right)\varphi(x+h) + D\left(x - \frac{h}{2}\right)\varphi(x-h) \\ & - \left[D\left(x+\frac{h}{2}\right) + D\left(x - \frac{h}{2}\right)\right]\varphi(x)\Big\}. \end{aligned}$$

The extension of (18) to $(2n + 1)$-point approximations for (17) and (17′) on rectangular meshes in $n = 2$ and $n = 3$ rectangular space coordinates is well-known and obvious.

For physical reasons, the v_i, D_i, Σ_i and ν in (17) and (17′) are all *nonnegative*. Hence, if A_i denotes the difference approximation (18) to $-\Sigma\partial/\partial x_k(D_i\partial/\partial x_k) + \Sigma_i$, and B_i denotes the matrix corresponding to multiplication by Σ_{i-1}, then $-A_i$ is an essentially nonnegative matrix and B_i is a nonnegative diagonal matrix. Hence (17) and (17′) lead to systems of vector differential equations of the form

$$(19)\quad d\boldsymbol{\varphi}_i/dt = v_i\{-A_i\boldsymbol{\varphi}_i + B_i\boldsymbol{\varphi}_{i-1}\} \quad (i = 2, \cdots, n),$$

$$(19')\quad d\boldsymbol{\varphi}_1/dt = v_1\{-A_1\boldsymbol{\varphi}_1 + \nu B_1\boldsymbol{\varphi}_n\},$$

where the $Q_i = -A_i$ are essentially nonnegative and the B_i are nonnegative. Hence equations (19) and (19′) define a continuous multiplicative process like (3), of the form $d\boldsymbol{\Phi}/dt = Q\boldsymbol{\Phi}$, where

$$(19^*)\quad Q = \begin{pmatrix} -v_1A_1 & 0 & \cdots & 0 & \nu B_1 \\ v_2B_2 & -v_2A_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & v_nB_n & -v_nA_n \end{pmatrix}.$$

If the reactor is geometrically *connected*, and if it contains some fissionable material, then the matrix Q of (19*) is then essentially positive, and the submatrices A_i are Stieltjes matrices. Hence, the results of §§3–6 are applicable.

Furthermore, the A_i are diagonally dominant,[16] and hence positive definite. Therefore, by Theorem 5, we have the further

COROLLARY 3. *In the usual difference-equation approximations for the multigroup diffusion equations* (17) *and* (17′), *the matrices* A_i^{-1} *are positive and the* $A_i^{-1}B_i$ *are nonnegative.*

In fact, for $i > 1$, the $A_i^{-1}B_i$ are positive.

10. Critical flux calculations. In critical flux calculations, the time-derivatives in (17) and (17′) are set equal to zero, giving rise to discrete multiplicative processes (2) of the special form (5) and (5′). We shall now apply the theory of §§3–6 to this case.

As shown in §7, the matrix P of (6), derived from (5) and (5′), is cyclic of index $k(P) = n > 1$. It follows from Theorem 4′ that simple iteration of the process $N(r + 1) = P \cdot N(r)$ defined by (6) will not in general converge to the critical flux distribution. However, this apparent disadvantage can be turned into a positive advantage by the simple device of replacing (6) by the multiplicative process defined by (7), which is nonnegative and *semi-irreducible*, but with a *pimitive* square submatrix P_2 (cf. Fig. 1).

To see that the two processes define the same (critical) thermal flux distribution, we proceed as follows. The matrices $A_i^{-1}B_i$ are nonnegative by Corollary 3 above, and the product T of (7′) is also nonnegative and semi-irreducible. Hence, by Theorems 1 and 1″, each process has a *unique* positive eigenvector. If φ_n is the thermal flux distribution vector component of the positive eigenvector $\Phi_1 = (\varphi_1, \cdots, \varphi_n)$ of P in (6), then

$$\varphi_n = A_n^{-1}B_n\varphi_{n-1} = \cdots = \nu(A_n^{-1}B_n \cdots A_1^{-1}B_1)\varphi_n = T\varphi_n .$$

Hence φ_n is the unique positive eigenvector of T.

We shall now turn our attention to finding the positive eigenvector of (7) and (7′). In digital calculations based on (7) and (7′), one calculates each $\varphi_k = A_k^{-1}B_k\varphi_{k-1}$ by an iterative numerical procedure; the *inner iterations* involved can be interpreted as solving a source problem, which will be discussed in §11. Assuming that this problem can be solved, the vector $T\varphi_n$ of the process represented by (7) and (7′) can be calculated with arbitrary accuracy. The iteration of this process is referred to as an *outer iteration*.

Theorem 4′ shows that such an outer iteration will necessarily converge to the dominant (critical) flux distribution. Moreover, this simple iteration

[16] A. Ostrowski, Comment. Math. Helv. 10 (1937), pp. 69–96. For further details and generalizations to various boundary conditions, see [7, p. 14]. The case of non-isotropic diffusion is, however, quite different.

process is intuitively attractive, because by §2 it can be interpreted as an ideal diffusion process with absorption and multiplication: it is analogous to what actually happens physically.

Though simple iteration, based on (7), is a multiplicative process which is sure to converge, other processes may converge more rapidly.[17] Though nonnegative matrices define intuitively attractive multiplicative processes, matrices whose entries have mixed signs may be more computationally efficient.

11. Source problems. With reactors containing sources of constant strength, (3) is replaced by

$$dN_i/dt = \sum_{j=1}^{n} q_{ij} N_j(t) + S_i, \qquad S_i \geqq 0. \tag{20}$$

Unless zero is an eigenvalue of Q, there is a unique *equilibrium* distribution N_0 such that $QN_0 = -S$; the general solution of (20) is then N_0 plus the general solution of (3). In the *subcritical* case where $M(Q) < 0$, Theorem 5 shows that $-Q^{-1}$ is *positive*. Hence, comparing with Theorem 4, we see

THEOREM 6. *For any irreducible process* (20), *with all* $N_i(0) \geqq 0$, *all* $S_i \geqq 0$ *and some* $S_i > 0$, *we have*

$$\begin{cases} \lim_{t\to\infty} \Sigma N_j(t) = +\infty & \textit{if (3) is critical or supercritical.} \\ \lim_{t\to\infty} N(t) = -Q^{-1}S & \textit{if (3) is subcritical.} \end{cases} \tag{21}$$

Source problems also arise in critical flux calculations. Clearly, we can write any single equation (5) or (5′) in the simplified form

$$A\varphi = -\nabla\cdot(D\nabla\varphi) + \Sigma'\varphi = S(x), \tag{22}$$

where $\nabla\cdot D\nabla$ may be approximated by (18), for example. Hence the "inner" iterations of §10 can be regarded as solutions of a *one*-velocity-group source problem.

We now discuss various specific processes for solving source problems[18] (22); no real generality is lost by considering the case

$$\begin{aligned} -h^2\delta^2\varphi = 4\varphi(x, y) - \varphi(x + h, y) - \varphi(x, y + h) - \varphi(x - h, y) \\ - \varphi(x, y - h) = S(x, y), \end{aligned} \tag{23}$$

of the plane difference analog to the Laplace equation. We shall show that the theory of §§3–8 sheds fresh light[19] on various iterative processes for solving (23).

---

[17] For example, Chebyshev polynomials in the matrix T of (7′) have been used successfully to accelerate convergence; see Bettis Atomic Power Division of Westinghouse Electric Corporation, Report WAPD-TM-70.

[18] For Dirichlet type boundary conditions.

[19] I.e., it gives results apparently not in the literature. For the latter, see [7] or [10].

The primary questions concern the convergence to zero, as $r \to \infty$, of the error function $E_r(x, y)$ after r iterations, under different iterative processes. This is equivalent to having spectral norm less than one (being "subcritical").

Modified Schmidt Processes transform the error function according to

$$E_{r+1} = E_r + kh^2\delta^2 E_r . \tag{24}$$

This is *nonnegative* (i.e., a "multiplicative process") if and only if $0 \leqq k \leqq \frac{1}{4}$, which is the usual Courant-Friedrichs-Lewy "stability" condition, as given in [14]. As the mesh length h tends to zero, this is asymptotically the same as the exact condition for being stable (having spectral norm less than one). In a connected domain, (24) is irreducible and nonnegative. If $k < \frac{1}{4}$, it is primitive; in the case $k = \frac{1}{4}$ of the ordinary Schmidt process, it is however *cyclic* of index two.

The Gauss-Seidel or *Liebmann process*, which affects the error vector by

$$4E_{r+1}(x, y) = E_r(x + h, y) + E_r(x, y + h) + E_{r+1}(x - h, y) + E_{r+1}(x, y - h) \tag{25}$$

is also a discrete multiplicative process; it can easily be shown to be nonnegative, semi-irreducible with a primitive square submatrix, and stable.

The class of implicit processes defined by

$$E_{r+1} = E_r + kh^2[\rho\delta^2 E_r + (1 - \rho)\delta^2 E_{r+1}] \tag{26}$$

is also of interest. Laasonen's choice $\rho = 0$ defines (26) as a multiplicative process, which is also stable (alias convergent or subcritical). The choice $\rho = \frac{1}{2}$ of Crank-Nicolson[20] has the highest order of accuracy, but is not positive.

Overrelaxation. The Young-Frankel method of Successive Overrelaxation is applicable to systems $A\boldsymbol{x} = \boldsymbol{k}$, where $A = \| a_{ij} \|$, an $n \times n$ matrix has all positive diagonal entries $a_{ii} > 0$, and satisfies Young's Property (A) [10, p. 93]:

(A) There exist two nonempty subsets S, T with $S \cap T = 0$, $S \cup T = \{1, 2, \cdots, n\}$, such that $a_{ij} \neq 0$ implies $i = j$ or, if $i \neq j$, then $i \in S$, $j \in T$ or $j \in S$, $i \in T$.

Let D be the positive diagonal matrix with diagonal entries $d_{ii} = 1/a_{ii}$, and let $B = -DA + I$. The reduced (iteration) matrix B is cyclic of index

[20] Proc. Cambridge Philos. Soc. 43 (1947), 5–67; for Laasonen, see Acta Math. 81 (1949), 309–317. The "alternating direction" implicit method of Peaceman-Rachford, described in J. Soc. Indust. Appl. Math. 3 (1953), pp. 28–41, is not usually order-preserving.

2 if, and only if, A satisfies property (A). Hence, if we apply (15) with $k = 2$, we have a simple proof of Young's result that the nonzero eigenvalues of B occur in $\pm$ pairs. Also, the proof is equally valid for generalizations to block property (A).[21]

Though the concepts of cyclic and nonnegative matrices are very useful in deriving the properties of Successive Overrelaxation with optimum ω, the process itself is *not* nonnegative on the error vector. Neither is "second-order Richardson", as defined in [15].

On the other hand, the Perron-Frobenius theory can be used for estimating the optimum *successive overrelaxation factor* ω_b [10, p. 95] for a class of cases including (23). Using the minimax property (12), nontrivial upper and lower bounds can be found for ω_b [7, p. 21].

Successive Overrelaxation can also be generalized to the following approximation to the Laplace equation for a triangular mesh in the complex z-plane,

$$(27) \quad u_{r+1}(z) - \tfrac{1}{3}\{u_r(z - h) + u_r(z - h\omega) + u_r(z - h\omega^2)\} = 0,$$

where ω is a primitive cube root of unity. This is cyclic of index three; one of us will discuss this generalization elsewhere.

12. Parameter ν. Criticality studies involve a matrix $P = A + \nu B$ depending linearly on a parameter ν, in many applications besides (6), (7) and (7′). For instance, ν might correspond to enrichment by U^{235}, to control rod effectiveness, or to reactor size. In such cases, one looks for the value of ν making the matrix $A + \nu B$ define a critical process. The following result is therefore of interest.

LEMMA 7. *Let A and B be nonnegative, $A + B$ semi-irreducible, and let $A + B$ contain a cycle with a nonzero entry from B. Then $\sigma(\nu) = L(A + \nu B)$ increases monotonely with ν, for $\nu > 0$.*

Proof. We shall assume for simplicity that $A + B$ is irreducible; the extension to the semi-irreducible case is easy. Let F_1 be the importance vector of $A + \nu B$, and φ_2 the eigenvector of $A + \nu' B$. If L is the spectral norm of $A + \nu B$, and L' that of $A + \nu' B$, where $\nu' > \nu$, then

$$F_1 L' \varphi_2 = F_1(A + \nu' B)\varphi_2 = F_1(A + \nu B)\varphi_2 + \Delta\nu F_1 B \varphi_2$$
$$= LF \cdot \varphi_2 + \Delta\nu F_1 B \varphi_2 > LF_1 \cdot \varphi_2 ,$$

since $\Delta\nu = \nu' - \nu > 0$, and all entries of F_1 and φ_2 are positive. Hence $L' > L$, as claimed. In the semi-irreducible case, a similar argument goes

[21] For such generalizations, see Arms, Gates and Zondek, J. Soc. Indust. Appl. Math. 4 (1956), 220–229.

through because all cycles lie in the range of $A + \nu B$, on which $A + \nu B$ is irreducible.

Further, since $L(P)$ is a simple root of $| P - \lambda I | = 0$, by the Implicit Function Theorem we know

LEMMA 8. *In Lemma 7, $\sigma(\nu)$ is analytic.*

In most applications (e.g., in (6)), $\sigma(0) = 0$. Moreover, we can prove

LEMMA 9.[22] *In Lemma 7, $\nu(+\infty) = +\infty$.*

Proof. If k is the length of the cycle of $A + B$ containing a nonzero entry from B, then $(A + \nu B)$ has a nonzero diagonal entry with some positive pth power of ν as a factor. Hence

$$(28) \qquad [L(A + \nu B)]^k = L([A + \nu B]^k) \geqq K\nu^p,$$

from which the conclusion follows by Lemma 7.

Combining Lemmas 7–9, we obtain

THEOREM 7. *Under the hypotheses of Lemma 7, $A + \nu B$ is critical for just one value ν_{cr}; it is subcritical if $\nu < \nu_{cr}$, and supercritical if $\nu > \nu_{cr}$.*

A similar result holds for continuous multiplicative processes. Moreover if P is *symmetric*, then the "norm" $N(P)$ of P satisfies

$$(29) \qquad N(P) = \sup_{X \neq O} \{XPP'X'/XX'\}^{1/2} = \sup | \lambda_k | = L(P).$$

But, for any matrices, $N(A + B) \leqq N(A) + N(B)$; therefore

THEOREM 8. *For symmetric matrices, the spectral norm is a convex function; hence $L(P)$ is a convex function of P, for symmetric, essentially positive matrices.*

The preceding argument applies also to *normal* matrices,[23] or matrices P such that $PP' = P'P$. There is however a catch: the sum of two normal matrices is not in general normal. The case of symmetric matrices seems, therefore, to be the most important case for which $L(A + B) \leqq L(A) + L(B)$.

13. Transition matrix. The concept of a matrix of transition probabilities is familiar from the theory of radioactive disintegration and other stochastic processes. We now extend it to any continuous multiplicative process (3) with negative diagonal entries q_{jj}. These entries are necessarily negative if (3) is subcritical.

Clearly, $\exp(q_{jj}t)$ is the probability that a neutron will remain in cell j through time t uninterruptedly under (3), so that $1/(-q_{jj}) = \tau_j$ is the mean sojourn time there. Hence $q_{ij} \exp(q_{jj}t)\, dt$ will be the expectation that a

[22] In Lemma 9, if B is semi-irreducible, then we have the more precise formula $\sigma(\nu) \sim \nu \cdot L(B)$ as $\nu \to \infty$, where $L(B) > 0$ is the largest eigenvalue of B.

[23] Another generalization is due to Peter Lax, Report NYO-7974, AEC Computing Facility, Institute of Mathematical Sciences, New York University, 1957.

neutron will appear in cell i by direct transiton from cell j in the time interval from t to $t + dt$. Integrating over $0 \leqq t \leqq +\infty$, we see that the *transition expectation* from cell j to cell i is

$$(30) \qquad r_{ij} = q_{ij}/-q_{jj} = \int_0^\infty q_{ij} \exp (q_{jj} t)\, dt \geqq 0, \qquad i \neq j.$$

Also, the idea of transition leads us to define

$$(30') \qquad r_{ii} = 0 \qquad (i = 1, \cdots, m).$$

Hence we make the

Definition. If all $q_{jj} < 0$, the *transition matrix* associated with the multiplicative process (3) is the matrix R defined by (30) and (30′).

THEOREM 9. *If all $q_{jj} < 0$, then the discrete process $N(r + 1) = RN(r)$ defined by the transition matrix* (30) *and* (30′) *is subcritical, critical, or supercritical, according as* (3) *is subcritical, critical, or supercritical. In the critical case, Q and R have the same importance vector.*

Proof. As in §4, Q and its transpose Q' have the same spectrum and hence criticality. If (3) has the importance vector F, then for $j = 1, \cdots, m$ we will have

$$(31) \qquad -q_{jj} \sum_k r_{kj} F_k = \sum_{k \neq j} q_{kj} F_k = MF_j - q_{jj} F_j .$$

Dividing through by $-q_{jj} > 0$, we see that $R'F > F$, $R'F = F$ or $R'F < F$ according as Q is supercritical, critical, or subcritical. The desired conclusions then follow from Theorem 4 and its corollaries.

The transition matrix R also arises naturally in connection with *source* problems (20). We look for the solution of (20) which is time-independent, so that

$$(32) \qquad -QN = S.$$

By definition, N satisfies (32) if and only if $-q_{jj} N_j = \sum_{k \neq j} q_{jk} N_k + S_j$, for $j = 1, 2, \cdots, n$, or

$$(32') \qquad N_j = \sum_k r_{jk} N_k + \tau_j S_j ,$$

or

$$(32'') \qquad N = RN + DS$$

where D is a diagonal matrix with diagonal entries $\tau_j = 1/(-q_{jj})$.

The *Gauss iteration process* [10, p. 100] for solving (32) consists in iterating (32′), with only unimproved values of N_k on the right side. Hence, the discrete process (2) defined by the transition matrix R of (3) is identical, if all $q_{jj} < 0$, with the process defined on the error function by the Gauss iteration process for solving $AN = S$, with $A = -Q$. Theorem 9 therefore implies the following

COROLLARY. *In the time-independent source problem* (32), *Gauss iteration is convergent if and only if the process* (3) *is subcritical.*

14. Thermal up-scattering. As an example to show how the preceding mathematical techniques can be applied to solve new problems, we consider the case of so-called thermal up-scattering.

In many group diffusion problems, "thermal" neutrons may themselves be subdivided into two or more lethargy groups. Due to the possibility of "up-scattering", neutrons in a given thermal group can then arise from lethargy groups having *lower* velocity. This defines a continuous multiplicative process through equations which may be written as

$$\frac{1}{v_i}\frac{\partial \varphi_i}{\partial t} = -A_i\varphi_i + \sum_{k \neq i} B_{ik}\varphi_k \qquad (1 \leqq i \leqq N), \tag{33}$$

where φ_i denotes the flux in the ith lethargy group, and the nonnegative matrices B_{ik} have entries corresponding to the possibility of scattering between lethargy groups in the various spatial cells. The theory of the continuous multiplicative process (3) defined by (33) is the same as before, since the matrix Q involved is still essentially positive.

The computational problem, however, requires a somewhat different analysis. For computational purposes, *all* the thermal lethargy groups are lumped together into a single "thermal group", containing all lethargy groups into which up-scattering is possible. The "inner iterations" used to solve the source problem for ordinary lethargy groups must be supplemented by "thermal iterations".

The source problem for the thermal group in the discrete process defined by (33) may be written in the matrix form

$$A_i\varphi_i = s_i + \sum_{k \neq i} B_{ik}\varphi_k \qquad (n \leqq i, k \leqq N). \tag{33'}$$

Here, s_i is the flux coming to the ith group from fission sources and scattering from epithermal groups; the B_{ik} refer to scattering between thermal groups. The matrices A_i are nonsingular as before, so that (33′) is equivalent to

$$\varphi_i = A_i^{-1}s_i + \sum_{k \neq i} A_i^{-1}B_{ik}\varphi_k \qquad (n \leqq i, k \leqq N). \tag{33''}$$

If $\Phi = (\varphi_n, \cdots, \varphi_N)$ represents the total thermal neutron flux, then the system (33″) may be written

$$\Phi = C\Phi + S, \tag{34}$$

where the matrix C is defined by

$$C = \begin{bmatrix} 0 & A_n^{-1}B_{n,n+1} & \cdots & A_n^{-1}B_{nN} \\ A_{n+1}^{-1} B_{n+1,n} & 0 & \cdots & A_{n+1}^{-1} B_{n+1,N} \\ \vdots & \vdots & & \vdots \\ A_N^{-1} B_{Nn} & A_N^{-1}B_{N,n+1} & \cdots & 0 \end{bmatrix}. \tag{35}$$

The A_i^{-1} are positive, so that C is nonnegative and in general irreducible.

Equation (34) can be solved by Gauss-Seidel iteration, which reduces to the Liebmann process (25) in the case of the Laplace equation. The matrix C is formed only implicitly, since, in general, the matrices A_i^{-1} themselves are not found explicitly, but approximated by means of inner iterations. Thus, the application of the Gauss-Seidel iteration method to the matrix C is equivalent to the method of "successive *block* displacements",[24] the block submatrices being precisely the A_i's.

Theorem 10. *In any critical process* (33), *Gauss and Gauss-Seidel iteration of* (34) *are convergent.*

Proof. In the *critical* case, the thermal groups by themselves define a *subcritical* process, by the monotonicity theorems of §12. Hence, as shown in Theorem 9 Gauss iteration is convergent. But it is known[25] that Gauss-Seidel iteration is then also convergent—and at *least* as rapidly convergent as Gauss iteration (which reduces to the Schmidt process (24) in the case of the Laplace equation). A similar conclusion holds, by the same argument, for any component of any critical process. Moreover the convergence of outer iterations follows as in §10.

Remark. If one allows up-scattering and down-scattering between *adjacent* lethargy groups only, then the matrix C of (34) is *cyclic* of index two. It follows that, in this case, one can again use Successive Overrelaxation to speed up the convergence of thermal iterations.

15. Complex eigenvalues. In many reactor problems like those considered above, the matrices involved have only real eigenvalues. Thus, this is the case with source problems (22), since the A_i are then symmetric. It has been conjectured implicitly that this is true for the matrix Q in typical reactor problems; this would be convenient, because it would permit an expansion of the time-dependent flux in a series of exponentials [4, p. 357]. It has also been conjectured that T has only real eigenvalues in actual reactors.[26] This has an important computational application,

---

[24] L. J. Arms, L. D. Gates, B. Zondek, loc. cit.

[25] P. Stein and R. L. Rosenberg, J. London Math. Soc. 23 (1948), 111–118.

[26] See Bettis Atomic Power Division of Westinghouse Electric Corporation, Report WAPD-TM-70, where a discussion is given as to how to apply Chebysheff-polynomials in the matrix T.

because it enables one to speed up convergence by using Chebysheff polynomials in T, in the outer iteration cycle. Some empirical support for this conjecture is lent by the fact that the use of such polynomials does, in fact, speed up convergence.[27]

We shall now discuss the reality of the eigenvalues of the matrices P, Q and T arising from *bare homogeneous reactor* problems.

In such reactors, the vector space of flux-components $(\varphi_1, \cdots, \varphi_n)$ is the direct sum of *subspaces* S_j, in each of which the spatial distribution is proportional to an eigenfunction $u_j(x)$ of the scalar Helmholtz equation $\nabla^2 u + \beta^2 u = 0$ for some $\beta = \beta_j$, or of its discrete analog. Moreover, each such subspace is *invariant* under the (linear) operators involved in equations (17) and (17′). Hence, the matrix P of (6) acts on each jth invariant subspace like an $n \times n$ matrix of the same form, but with $A_k^{-1}B_k$ replaced by $\Sigma_k/(\Sigma_k' + \beta_j^2 D_k)$.

It follows that the matrix T of (7) and (7′) acts on each jth invariant subspace like the *scalar* matrix $\nu \prod_k [\Sigma_k/(\Sigma_k' + \beta_j^2 D_k)] I_j$, where I_j is the $n \times n$ identity matrix. Consequently, T has *all real eigenvalues.*

On the other hand, the matrix of (6) will always have complex eigenvalues, if $n > 2$. For, P being similar to a *bidiagonal* matrix, having nonzero entries only on the principal diagonal and one adjacent diagonal, its characteristic polynomial is easily calculated as

$$(36) \qquad | \lambda I - P | = \prod_j \{\lambda^n - \nu \prod_k [\Sigma_k/(\Sigma_k' + \beta_j^2 D_k)]\}.$$

This evidently has all real eigenvalues if and only if $n = 1$ or 2 (cf. [4], p. 242).

The matrix Q arising in the nominal[28] reactor kinetics problem defined by (17) and (17′), for the same (ideal) reactor, is also bidiagonal. Moreover,

$$(37) \qquad | I - Q | = \prod_j \left\{ \prod_{i=1}^n [\mu + v_i(\beta_j^2 D_i + \Sigma_i')] - \nu \prod_{i=1}^n v_i \Sigma_i \right\}.$$

For Q to have complex eigenvalues, it is necessary and sufficient that the number of intersections of each polynomial curve

$$(38) \qquad y = f_j(\mu) = \prod_i [\mu + v_i(\beta_j^2 D_i + \Sigma_i']$$

with the horizontal line $y = \nu \prod v_i \Sigma_i$ be n.

[27] Some *a priori* support may also be found in the idea that inhomogeneities lead typically to cycles of length two (core fission → slowing down in moderator → core fission) or less; 2×2 nonnegative matrices have real eigenvalues.

[28] Nominal, becaused delayed neutrons are ignored, and because the time scales involved in (17) and (17′) are entirely fictitious in few-group calculations.

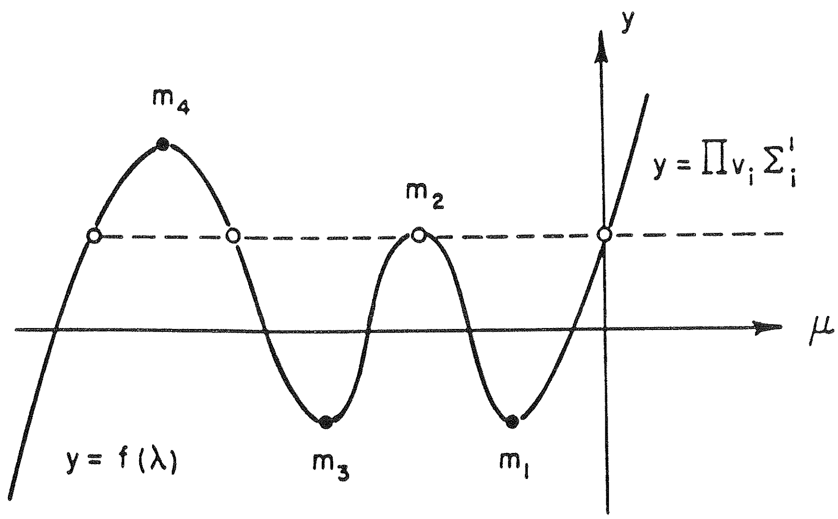

FIG. 3

The variation in the coefficients of (3) is extreme, since the neutron velocities v_i vary by a factor of 4×10^8 or more. This makes the problem very difficult. To get some idea of the facts, we therefore suppose that absorption and fission can be neglected, and that $\Sigma_i' = \Sigma_i$ are velocity-independent. We shall see that, in few-group approximations, Q will then usually have only real eigenvalues. We shall confine our attention to the case $j = 1$ of the *dominant* mode (positive eigenfunction), writing $\beta_1 = \beta$ and $f_1(\mu) = f(\mu)$ for simplicity.

In Fig. 3, we have sketched a sample curve $y = f(\lambda)$. It crosses the μ-axis at $\mu = -v_i\Sigma_i'\mu = \alpha_i$ n times (in Fig. 3, $n = 5$). Between successive α_i, $f(\mu)$ has one maximum or one minimum, and the signs of these alternate. The condition that Q should have (two or more) complex eigenvalues is thus simply that $y = \nu\prod v_i\Sigma_i$ should lie above at least one of these maxima. In Fig. 3, we have sketched a borderline case, in which $y = \nu\prod v_i\Sigma_i$ at the lowest of these maxima. In the critical case, since $|\mu I - Q|$ has a factor μ, clearly $f(0) = \nu\prod v_i\Sigma_i'$, this case is depicted in Fig. 3. The critical case is of course of greatest interest.

With three equally spaced lethargy groups, we can then write $v_1\Sigma_1 = \alpha^{-1}$, $v_2\Sigma_2 = 1$, $v_3\Sigma_3 = \alpha$ for $\alpha > 1$, in suitable units. In the borderline critical case,

$$
\begin{aligned}
f(\mu) &= \nu\prod v_i\Sigma_i' = (\mu - \alpha - 1)(\mu - \alpha) + 1 = \mu(\mu^2 - A\mu + A), \\
A &= \alpha^{-1} + 1 + \alpha.
\end{aligned}
\tag{39}
$$

The quadratic factor is a perfect square $(\mu - m_2)^2$ if and only if $A = 4$, or $\alpha = \frac{1}{2}(3 + \sqrt{5}) = 1.97^{-}$. In this case, the borderline lethargy spacing is about $2n(v_{i+1}/v_i) = 1.4$.

A similar but more complicated calculation can be made in case $n = 4$, with $v_i\Sigma_i = \alpha^{i-5/2}$. A preliminary rough estimate gives the borderline value of α in this case as about *three*. These estimates suggest that *complex eigenvalues are likely to arise if twenty or more equally spaced lethargy groups are used* in multigroup diffusion problems.

A general theorem confirming this view can also be stated.

THEOREM 11. *If $n \geqq 4$, and if the smallest value α of $v_i\Sigma_i'$ exceeds one-fifth of the fourth smallest $v_i\Sigma_i'$, then Q has at least one complex eigenvalue.*

Roughly interpreted, the preceding result suggests that complex eigenvalues may be expected if the change in lethargy $n(E_0/E)$ between successive velocity groups is less than 0.8. We emphasize that it is only *sufficient* to imply the existence of complex eigenvalues, and not necessary.

Proof. By the previous discussion, it suffices to prove that $\prod v_i\Sigma_i'$ exceeds the relative maximum $f(\lambda_2)$ nearest the origin. If the smallest four $v_i\Sigma_i'$ are α_1, α_2, α_3, α_4, then $-\alpha_2 > \lambda_2 > -\alpha_3$. The maximum of the ratio $\prod(\lambda - \alpha_i)/\alpha_1\alpha_2\alpha_3\alpha_4$, subject to the condition

$$\alpha = \alpha_1 \leqq \alpha_2 < \lambda < \alpha_3 \leqq \alpha_4 \leqq 5\alpha,$$

clearly occurs when $\alpha_1 = \alpha_2 = \alpha; \lambda = 3\alpha; \alpha_3 = \alpha_4 = 5\alpha$. Hence it is 16/25, proving that

$$\prod_{i=1}^{4} (\lambda_2 + v_i \Sigma_i') \Big/ \prod_{i=1}^{4} v_i \Sigma_i' < 16/25.$$

For all $v_i\Sigma_i'$ other than the smallest four, however, clearly $\lambda_2 + v_i\Sigma_i' < v_i\Sigma_i$ since $\lambda_2 < 0$. Hence $\prod v_i\Sigma_i' > \prod(\lambda_2 + v_i\Sigma_i')$, completing the proof.

REFERENCES

1. C. J. EVERETT AND S. ULAM, *Multiplicative systems. I.* Proc. Nat. Acad. Sci. U.S.A., 34 (1948), pp. 403–405.
2. M. FRÉCHET, *Méthode des Fonctions Arbitraires . . .*, vol. 2, Gauthier-Villars, Paris, 1936.
3. G. FROBENIUS, *Über Matrizen aus nicht negativen Elementen*, Sitzungsberichte der Akademie der Wissenschaften zu Berlin, (1912), pp. 456–477.
4. S. GLASSTONE AND M. C. EDLUND, *The Elements of Nuclear Reactor Theory*, van Nostrand, New York, 1952.
5. T. E. HARRIS, *Some mathematical models for branching processes*, Second Berkeley Symposium on Mathematical Statistics and Probability, pp. 305–328, Univ. of Calif. Press, 1951.
6. V. ROMANOVSKY, *Recherches sur les chaînes de Markoff*, Acta Math., 66 (1936), pp. 147–251.
7. R. S. VARGA, *Numerical solution of multigroup diffusion problems in x–y geometry*, I. R. E. Trans. Prof. Group Nuclear Sci., NS-4 (1957), pp. 52–62.
8. A. M. WEINBERG, *Current status of nuclear reactor theory*, Amer. J. Phys., 20 (1952), pp. 1–12.
9. H. WIELANDT, *Unzerlegbare, nicht negative Matrizen*, Math. Z., 52 (1950), pp. 642–648.

10. D. YOUNG, *Iterative methods for solving partial difference equations of elliptic type*, Trans. Amer. Math. Soc., 76 (1954), pp. 92–111.

11. G. DEBREU AND I. N. HERSTEIN, *Non-negative square matrices*, Econometrica 21 (1953), pp. 597–607.

12. GARRETT BIRKHOFF AND RICHARD S. VARGA, *Reactor criticality and non-negative matrices*, Report WAPD-166, Bettis Atomic Power Division of Westinghouse Electric Corporation, 1957.

13. M. K. BUTLER AND J. M. COOK, *Univac Programs for the solution of one-dimensional multigroup reactor equations*, Report ANL-5437, Argonne National Laboratory, 1956.

14. G. O'BRIEN, M. HYMAN AND S. KAPLAN, *A study of the numerical solution of partial differential equations*, J. Math. Phys., 29 (1951), pp. 223–251.

15. S. P. FRANKEL, *Convergence rates of iterative treatments of partial differential equations*, Math. Tables Aids Comput., 4 (1950), pp. 65–75.

HARVARD UNIVERSITY,
BETTIS ATOMIC POWER DIVISION,
WESTINGHOUSE ELECTRIC CORPORATION,
PITTSBURGH