

ECONOMETRICA

VOLUME 29

April 1961

NUMBER 2

AGGREGATION OF VARIABLES IN DYNAMIC SYSTEMS¹

BY HERBERT A. SIMON AND ALBERT ANDO

Previous studies have examined dynamic systems that are decomposable into independent subsystems. This article treats of systems that are *nearly* decomposable—systems with matrices whose elements, except within certain submatrices along the main diagonal, approach zero in the limit. Such a system can be represented as a superposition of (1) a set of independent subsystems (one for each submatrix on the diagonal) and (2) an aggregate system having one variable for each subsystem. This superposition separates short-run from long-run dynamics and justifies the ignoring of "weak" linkages in partial equilibrium studies.

I. INTRODUCTION

IN MANY problems of economic theory we need to use aggregates. The general Walrasian system and its more modern dynamic extensions are relatively barren of results for macroeconomics and economic policy. Hence, in our desire to deal with such questions we use highly aggregated systems by sheer necessity, often without having much more than the same necessity as our justification. Perhaps the most important result to date for justifying aggregation under certain circumstances is the Lange–Hicks² condition, about which we shall say more later.

Concern with actual numerical coefficients in the Leontief input–output model renewed interest in aggregation. It was hoped at first, perhaps, that modern computers would handle matrices of about any desired size, and hence would obviate the need for aggregation. By now it is clear that our ambitions have outstripped the computers, for the time required to invert a matrix by known methods increases with the cube of the number of rows and

¹ An earlier version of this paper was the result of the research undertaken in the Graduate School of Industrial Administration, Carnegie Institute of Technology, for the project *Planning and Control of Industrial Operations*, under contract with the Office of Naval Research and was circulated as O. N. R. Research Memorandum, No. 31, December, 1956. Reproduction of this manuscript in whole or in part is permitted for any purpose of the United States Government. Contract N OMR-76001, Project NR 047001. It was also read at the meetings of the Econometric Society, December, 1956. We are indebted to Drs. Allen Newell, Franklin Fisher, and S. Chakravarty for several enlightening discussions on the subject of this paper, and to John Muth for valuable suggestion on the proof given in Section IV.

² Lange, O., *Price Flexibility and Employment*, pp. 103 ff., 1952.

columns. The high cost of inverting large matrices has led to a number of experiments³ to determine whether aggregation can be used to obtain approximate inverses more economically.

Hence, aggregation is a topic of considerable importance regardless of whether we are interested in the general mathematical treatment of large systems—in which case aggregation is essential for conceptual clarity and for effective manipulation of the systems—or interested in numerical computation—in which case aggregation is often necessary to make computation feasible with available means.

Let us consider an example. Suppose that government planners are interested in the effects of a subsidy to a basic industry, say the steel industry, on the total effective demand in the economy. Strictly speaking, we must deal with individual producers and consumers, and trace through all interactions among the economic agents in the economy. This being an obviously impossible task, we would use such aggregated variables as the total output of the steel industry, aggregate consumption and aggregate investment. The reasoning behind such a procedure may be summarized as follows: (1) we can somehow classify all the variables in the economy into a small number of groups; (2) we can study the interactions within the groups as though the interaction among groups did not exist; (3) we can define indices representing groups and study the interaction among these indices without regard to the interactions within each group.

When thus explicitly written out, this reasoning appears rather bold. Yet, we implicitly use it almost every day in our analysis of the economy. We should at least make some attempt to specify the conditions that would justify the aggregation.

The conditions for exact aggregation are very severe. Whether these conditions are strictly satisfied in any practical situation is, of course, not important. We would be perfectly satisfied with aggregative models that were only approximate; we have no illusions that *any* model we might employ is more than an approximate description of reality. In exploring the aggregation problem we seek rules and criteria—exact or heuristic—that indicate what variables to aggregate and that show the circumstances under which aggregation will yield satisfactory approximations.

In the more general sense, justifications for approximation must be related to the decisions that depend on the approximating—if the decisions based on the approximate model are not much “worse” than the decisions based on the more elaborate model according to some criteria, then we may be justified in using the approximate, simpler model. This consideration is strengthened if, while the improvement of the final decision is very slight, the cost of

³ Balderston, J. B., and T. M. Whitin, “Aggregation in the Input-Output Model,” in *Economic Activity Analysis*, ed. by O. Morgenstern, 1954.

working with a larger model is very much greater than that of working with an approximate, simpler model. Furthermore, the relation between the aggregation problem here discussed and the identification problem when the structural equations are misspecified—a relation pointed out recently by Fisher⁴—can be better understood when both of these problems are viewed as parts of a larger decision problem. We shall come back to this point in the concluding section of this paper, but we must first make the statement of our problem more precise.

We shall restrict our attention to aggregation defined within the following algebraic model: We have two sets of (not necessarily all distinct) variables, $x = (x_1, \dots, x_i, \dots, x_n)$, and $y = (y_1, \dots, y_j, \dots, y_m)$, and a system of equations giving y as a function of x ; $y = \phi(x)$. We wish to know under what circumstances there exist two sets of functions, $X_I(x)$, $I = 1, \dots, N$, $N < n$, and $Y_J(y)$, $J = 1, \dots, M$, $M < m$, such that a set of relations between X and Y : $Y = \xi(X)$, can be derived from the given system of equations, ϕ , relating x and y . X and Y are vectors: $X = (X_1, \dots, X_N)$, and $Y = (Y_1, \dots, Y_M)$. Sometimes additional conditions are imposed, e.g., that the functions X and Y are given; or that there exist also relations, $y = y(Y)$, such that if, for a given set of values of x , \bar{x} , $\bar{y} = \phi(\bar{x})$, $\bar{X} = X(\bar{x})$, $\bar{Y} = \xi(\bar{X})$, and $\bar{y} = y(\bar{Y})$, then $|\bar{y}_j - \bar{y}_j| < \epsilon$ for all j , where ϵ is a given positive number.

In the special case where the relations are all linear and the y_i 's are the values of the x_i 's in the following period, the function ϕ can be written as

$$(1.1) \quad x(t+1) = x(t)P$$

where $x(t)$ is an n -dimensional row vector and P is an $n \times n$ matrix of constants. The question raised above will then be specialized to that of conditions under which there exist N functions,

$$(1.2) \quad X_I(t) = X_I[x(t)] \quad (I = 1, \dots, N < n)$$

and a new set of relations

$$(1.3) \quad X(t+1) = X(t)Q$$

where $X(t)$ is an N -dimensional row vector, and Q is an $N \times N$ matrix of constants. The elements of Q are functions of the elements of P . Further, we seek to define an additional set of n functions

$$(1.4) \quad x_i(t) = f_i[X(t)] \quad (i = 1, \dots, n)$$

such that the time path of the x_i 's defined by (1.3) and (1.4) can be considered an acceptable approximation of the time path of the x_i 's defined by (1.1) according to some predetermined criteria.

The Lange-Hicks condition is a criterion of this kind. It states: if two or

⁴ Fisher, F. M., "On the Cost of Approximate Specification in Simultaneous Equation Estimation," *Econometrica*, Vol. 29, April, 1961.

more variables always move together they may be aggregated into a single variable, which will be an appropriately weighted average of the original variables. This is a useful criterion, since it tells us that we may aggregate classes of commodities that are perfect substitutes, or that are approximately so.

At another level, the Lange-Hicks condition is unsatisfactory, for it requires that we know in advance which variables move together. In reality, it may be part of our problem to discover this. We may be confronted with a dynamic system of many variables, and may have to infer from the equations of the system which variables will move together—or will move nearly enough together to warrant aggregation. This is the problem we shall set ourselves here: *to determine conditions that, if satisfied by a (linear) dynamic system, will permit approximate aggregation of variables.* Note that we shall be interested in sufficient, rather than necessary, conditions. Hence, we may also view our task as that of discovering one or more classes of dynamic systems whose variables may be aggregated.

In Section II, we shall present one such class of dynamic systems, which we call "nearly decomposable" systems, and suggest that in such systems, the aggregation of variables described by equations (1.1) through (1.4) can be performed. In Section III, we shall give a physical illustration of such a system. In Sections IV and V, the mathematical analysis underlying the propositions summarized in Section II will be given in some detail. Section VI will present a numerical illustration of the behavior through time of "nearly decomposable" systems, and, as a special case, an application of the aggregation procedure discussed in this paper to the problem of the inversion of matrices. Finally, in Section VII, we shall discuss some further implications of our results and their relations to the recent works of others.

II. PROPERTIES OF NEARLY DECOMPOSABLE MATRICES

The dynamic characteristics of the system (1.1) depend on the properties of the matrix of its coefficients, $P = ||P_{ij}||$. More specifically, we will be interested in the patterns of zeros and near-zeros in the matrix P . Since we are concerned with closed systems, we assume P to be a square matrix.

Let us consider a matrix P^* , that can be arranged in the following form after an appropriate permutation of rows and columns:

$$(2.1) \quad P^* = \begin{array}{c|c|c|c} & P_{11}^* & & \\ & & \ddots & \\ & & & P_N^* \end{array}$$

where the P_I^* 's are square submatrices and the remaining elements, not displayed, are all zero. Then the matrix is said to be completely decomposable.

Let us denote the number of rows and columns in the I th submatrix in P^* by n_I . Then

$$n = \sum_{I=1}^N n_I.$$

We shall also adopt the following notation for the vector $\{x_i^*(t)\}$ on which P^* operates:

$$x^*(t) = \{x_i^*(t)\} = \{[x_{i_1}^*(t)], \dots, [x_{i_I}^*(t)], \dots, [x_{i_N}^*(t)]\}$$

where $x_{i_I}^*(t)$ is the row vector of a subset of components of $\{x_i^*(t)\}$, so that if

$$x_{i_I}^*(t) = x_i(t)$$

then

$$i = \sum_{J=1}^{I-1} n_J + i_I.$$

It is clear that if the system (1.1) is specified to

$$(2.2) \quad x^*(t) = x^*(0) P^{*t}$$

then the subset $[x_{i_I}^*(t)]$ of $x(t)$ at any stage t depends only on $[x_{i_I}^*(0)]$ and P_I^* and is independent of $[x_{i_J}^*(0)]$ and P_J^* , $I \neq J$.

Let us next consider a slightly altered matrix P defined by

$$(2.3) \quad P = P^* + \varepsilon C$$

where ε is a very small real number, and C is an arbitrary matrix of the same dimension as P^* . In this introductory discussion, the phrase "a very small real number" is intentionally left unprecise; it will be defined later. We shall refer to matrices such as P as *nearly decomposable matrices*. Using P thus defined, we restate equation (1.1) below:

$$(2.4) \quad x(t+1) = x(t)P.$$

It is the dynamic behavior of the system given by (2.4) in which we are interested.

Let the roots of the submatrix P_I^* of P^* be designated as: $\lambda_{1_I}^*, \lambda_{2_I}^*, \dots, \lambda_{n_I}^*$. We assume that these roots are distinct, and the subscripts are so arranged that $\lambda_{1_I}^* > \lambda_{2_I}^* > \dots > \lambda_{n_I}^*$. In addition, if $\lambda_{1_1}^*, \lambda_{1_2}^*, \dots, \lambda_{1_N}^*$ are all distinct, then without loss of generality, we can arrange rows and columns so that $\lambda_{1_1}^* > \lambda_{1_2}^* > \dots > \lambda_{1_N}^*$. When all these conditions are satisfied, i.e., when all roots of P^* are distinct, we can analyze the dynamic behavior of (2.4) with relative ease. In most cases, we feel that this assumption is not too restrictive as a description of reality. There is, however, an important exception, namely, the case where P^* is a stochastic matrix in which case all $\lambda_{1_I}^*$ are

unity. It turns out, fortunately, that our analysis of the case with all distinct roots can readily be extended to the stochastic case. In the remainder of this section, we shall summarize the results of more detailed mathematical analysis presented in Sections IV and V.

Let us define

$$(2.5) \quad \min_{\substack{i,j \\ i \neq j}} |\lambda_i^* - \lambda_j^*| = \delta^*.$$

If P^* is stochastic, the N largest roots of P^* take the identical values of unity. In this case, (2.5) should be interpreted to mean the selection of the minimum of the differences among all roots whose values are not unity and their differences from unity.

Since the roots of a matrix are continuous functions of its elements,⁵ we can define, for any positive real number δ , however small, a small enough ϵ so that, for every root of P^* , λ_i^* , there exists a root of P , λ_i , such that

$$(2.6) \quad |\lambda_i - \lambda_i^*| < \delta.$$

We can choose δ sufficiently smaller than δ^* so that there is no confusion as to which root of P corresponds to any particular root of P^* . The only exceptions to this proposition are the N largest roots of the stochastic P^* . If P^* is stochastic, we have

$$(2.7a) \quad \lambda_{1_I}^* = 1 \quad (I = 1, 2, \dots, N),$$

$$(2.7b) \quad |1 - \lambda_{i_I}^*| > \delta^* \quad (i_I = 2, \dots, n_I; I = 1, \dots, N).$$

Hence, for P , we must have

$$(2.8a) \quad |1 - \lambda_{1_I}| < \delta \quad (I = 1, 2, \dots, N),$$

$$(2.8b) \quad |1 - \lambda_{i_I}| > \delta^* - \delta \quad (i_I = 2, \dots, n_I; I = 1, \dots, N).$$

Because of the correspondence of the characteristic roots of P and P^* described above, we expect $x(t)$, when its time path is defined by (2.4), to exhibit the following dynamic behavior:

(1) In the short run, the behavior of $x_{i_I}(t)$ will be dominated by roots belonging to the I th subset, so that the time path of $x_{i_I}(t)$ will be very close to the time path of $x_{i_I}^*(t)$, and almost independent of $x_{j_J}(t)$, $J \neq I$, and P_J , $J \neq I$. Here, P_J is defined to be the submatrix of P corresponding to P_J^* of P^* . If we are interested in the behavior of the system at this stage, we can treat the system as though it were completely decomposable.

(2) Unlike P^* , P is not completely decomposable, so that the weak links among the subsystems will eventually make their influence felt. But the time required for these influences to appear is long enough so that when

⁵ See, for instance, E. C. Titchmarsh, *The Theory of Functions*, 2nd ed., 1939.

they do become visible, within each subsystem the largest root, λ_{1_I} , will have dominated all other roots, $\lambda_{2_I}, \dots, \lambda_{n_I}$. Thus, at this stage, the variables within each subset, $[x_{I_I}(t)]$, will move proportionately, and the behavior of the whole system will be dominated by N roots, $\lambda_{1_1}, \dots, \lambda_{1_N}$. Notice that, since the variables in each subsystem move roughly proportionately, the Lange-Hicks condition for aggregation is approximately satisfied.

(3) At the end, however, the behavior of $x(t)$ will be dominated by the largest root of P , as in any linear dynamic system.

It is quite clear that all these statements refer to the limiting properties of the system (2.4), and can be made more precise and meaningful only with more careful mathematical analyses of the system, which are presented in Sections IV and V.

III. A PHYSICAL ILLUSTRATION

Before we proceed with a more complete statement of the mathematics that underlies our analysis, it may be useful to provide an example of a physical system that can be approximately decomposed in the manner just described. We shall see that the principle of aggregation we are employing is essentially that which justifies the replacement of microvariables by macrovariables in classical thermodynamics.

Consider a building whose outside walls provide perfect thermal insulation from the environment. The building is divided into a large number of rooms, the walls between them being good, but not perfect, insulators. Each room is divided into a number of offices by partitions. The partitions are poor insulators. A thermometer hangs in each of the offices. Suppose that at time t_0 the various offices within the building are in a state of thermal disequilibrium—there is a wide variation in temperature from office to office and from room to room. When we take new temperature readings at time t_1 , several hours after t_0 , what will we find? At t_1 , there will be very little variation in temperature among the offices within each single room, but there may still be large temperature variations *among* rooms. When we take readings again at time t_2 , several days after t_1 , we find an almost uniform temperature throughout the building; the temperature differences among rooms have virtually disappeared.

The well-known equations for the diffusion of heat allow us to represent this situation by a system of differential equations—or approximately by a system of difference equations. Let $F_{t_I}(t)$ be the temperature of the i th office which is in the I th room, at time t . Let $F(t)$ be the vector consisting of these temperatures as components, $F(t) = [F_{1_1}, F_{2_1}, \dots, F_{t_1}, \dots, F_{1_N}, \dots, F_{n_N}]$. Then

$$(3.1) \quad F(t+1) = F(t)R$$

where R is a matrix whose element, r_{ij} , represents the rate of heat transfer between office i and office j per degree difference in temperature.

A temperature equilibrium *within* each room will be reached rather rapidly, while a temperature equilibrium *among* rooms will be reached only slowly, if the r_{ij} are generally large when i and j are offices in the same room, and are close to zero when i and j are offices in different rooms—that is to say, if the matrix R is nearly decomposable. When this is the case, and as long as we are not interested in the rapid fluctuations in temperature among offices in the same room, we can learn all we want to know about the dynamics of this system by placing a single thermometer in each room—it is unnecessary to place a thermometer in each office.

IV. MATHEMATICAL ANALYSIS

In this section, we shall make precise the meaning of propositions stated at the end of Section II, and provide their proofs for the case where the roots of P^* are all distinct.

Every matrix with distinct roots is similar to a diagonal matrix whose nonzero elements are those roots. Hence, there exist nonsingular matrices Z and Z^* such that

$$(4.1) \quad PZ = Z\Lambda,$$

$$(4.2) \quad P^*Z^* = Z^*\Lambda^*,$$

where Λ and Λ^* are diagonal matrices whose nonzero elements are roots of P and P^* , respectively. Since Z and Z^* are defined only up to a scalar factor and permutations, we select Z^* so that the λ_i^* appear in order of magnitude for each I .

From the argument leading to (2.6), it is clear that the λ_i 's are functions of ε . Let us take a particular value of δ , say δ_0 , and choose a value of ε , ε_0 , which satisfies (2.6). Then, we can define

$$(4.3) \quad \delta_i(\varepsilon_0) = \lambda_i(\varepsilon_0) - \lambda_i^*, \quad |\delta_i(\varepsilon_0)| < \delta_0;$$

$$(4.4) \quad v_i = \frac{\delta_i(\varepsilon_0)}{\delta_0}, \quad |v_i| < 1;$$

and a diagonal matrix whose elements are v_i 's, i.e.,

$$(4.5) \quad V = \begin{vmatrix} v_1 \\ & \ddots \\ & & v_i \\ & & & \ddots \\ & & & & v_n \end{vmatrix}.$$

We can then write

$$(4.6) \quad A = A^* + \delta_0 V.$$

Substitution of (2.3) and (4.6) with these choices of ε and δ into (4.1) yields

$$(4.7) \quad (P^* + \varepsilon_0 C)Z = Z(A^* + \delta_0 V)$$

and

$$(4.8) \quad P^*Z - ZA^* = \delta_0 ZV - \varepsilon_0 CZ.$$

As $\delta_0 \rightarrow 0$, and hence $\varepsilon_0 \rightarrow 0$, the Z 's remaining bounded, the right hand side of (4.8) approaches a matrix whose elements are all zero. Comparison of the left hand side of (4.8) with (4.2) indicates that, if z_{ij} and z_{ij}^* are elements of Z and Z^* , we must have a relation

$$(4.9) \quad \lim_{\delta_0 \rightarrow 0} z_{ij} = cz_{ij}^*$$

for all i and j , where c is some constant. Since c is arbitrary, we let $c = 1$. Thus, we can state:

THEOREM 4.1: *There is a way of selecting the matrices Z such that, for an arbitrary positive real number ζ , there exists ε_0 such that, for $\varepsilon < \varepsilon_0$,*

$$(4.10) \quad \max_{i,j} |z_{ij} - z_{ij}^*| < \zeta.$$

Let us take a value of ε , ε_1 , that satisfies the condition (4.10). Corresponding to this ε_1 , we can define $\lambda_i(\varepsilon_1)$, $i = 1, \dots, n$. These values can then be inserted in (4.1) to yield a specific Z . We then define, using this Z and the Z^* given by (4.2), a new set of values

$$(4.11) \quad \zeta_{ij} = z_{ij}^* - z_{ij}.$$

Because of the way in which the ζ_{ij} 's are constructed, we are assured that $|\zeta_{ij}| < \zeta$ for all i and j . Let us further define

$$(4.12) \quad u_{ij} = \frac{\zeta_{ij}}{\zeta}.$$

We note that $|u_{ij}| < 1$. We can then write

$$(4.13) \quad Z = Z^* + \zeta U$$

where U is the matrix whose elements are the u_{ij} 's. We know that

$$(4.14) \quad x(t) = x(0)P^t; \quad x^*(t) = x^*(0)P^{*\dagger}.$$

Consider next vectors $y(t)$ and $y^*(t)$ defined by

$$(4.15a) \quad y(t) = x(t)Z^{-1}$$

$$(4.15b) \quad y^*(t) = x^*Z^{*-1}.$$

Substituting (4.15) into (4.14), we obtain:

$$(4.16a) \quad y(t) = x(0)(Z^{-1}Z)P^tZ^{-1} = y(0)(ZPZ^{-1})^t = y(0)\Lambda^t$$

$$(4.16b) \quad y^*(t) = x^*(0)(Z^{*-1}Z^*)P^{*t}Z^{*-1} = y^*(0)(Z^*P^*Z^{*-1})^t = y^*(0)\Lambda^{*t}.$$

The inverse transformations of (4.15), when the results of (4.16) are substituted into them, yield:

$$(4.17a) \quad x(t) = y(0)\Lambda^t Z;$$

$$(4.17b) \quad x^*(t) = y^*(0)\Lambda^{*t} Z^*.$$

Let us now look at elements of (4.17) more closely. They have the form

$$(4.18) \quad x_{IJ}(t) = \sum_{i=1}^n z_{ij} \lambda_i^t y_i(0).$$

It is obvious, from the structure of the decomposable matrix ρ^* , that

$$z_{I^J}^* = 0 \quad \text{for } I \neq J.$$

This, together with (4.13), implies that

$$(4.19) \quad z_{I^J} = \zeta u_{I^J} \quad \text{for } I \neq J.$$

Hence, we can divide the right hand side of (4.18) into five sets of terms according to the scheme:

$$(4.20) \quad x_{IJ}(t) = \zeta u_{1I^J} \lambda_{1I}^t y_{1I}(0) + z_{1J^J} \lambda_{1J}^t y_{1J}(0) \\ + \zeta \sum_{\substack{I=2 \\ I \neq J}}^N u_{1I^J} \lambda_{1I}^t y_{1I}(0) + \sum_{i_J=2}^{n_J} z_{i_J^J} \lambda_{i_J}^t y_{i_J}(0) + \zeta \sum_{I=1}^N \sum_{\substack{i_I=2 \\ I \neq J}}^{n_I} u_{i_I^J} \lambda_{i_I}^t y_{i_I}(0).$$

In the special case where $J = 1$, the first term is absent. In order to discuss the meaning and implications of (4.20), let us give names to each term on the right hand side of (4.20) as follows:

$$(4.21) \quad x_{J^J}(t) = \zeta S_j^{(1)} + S_j^{(2)} + \zeta S_j^{(3)} + S_j^{(4)} + \zeta S_j^{(5)}.$$

For almost all choices of $y(0)$, each of the five terms on the right side of (4.21) will be nonzero at $t = 0$. We limit ourselves to this case, which is the general one. The following propositions are now self evident, and we state them as a theorem:

THEOREM 4.2. PART (1): Since $\lambda_{1J} > \lambda_{ij}, j = 2, \dots, n$, for any real positive number η_0 there exists an integer T_0 such that, for $t > T_0$,

$$(4.22) \quad \frac{|S_j^{(4)}|}{|S_j^{(2)}|} < \eta_0.$$

For a given $T_1 > T_0$ and arbitrary positive real number η_1 , there exists a number ζ_1 , and hence a number ε_1 by Theorem 4.1, such that, for $t < T_1$ and $\varepsilon < \varepsilon_1$,

$$(4.23) \quad \frac{|\zeta(S_j^{(1)} + S_j^{(3)} + S_j^{(5)})|}{|S_j^{(2)} + S_j^{(4)}|} < \eta_1.$$

Theorem 4.2, Part 1, states that, for a sufficiently small ε , the system characterized by a nearly decomposable matrix P behaves in a manner similar to the behavior of the completely decomposable system P^* for small t . This is even clearer when we express $x^*(t)$ as

$$(4.24) \quad x_{ij}^*(t) = z_{1J}^* \lambda_{1J}^{*t} y_{1J}^*(0) + \sum_{i=2}^n z_{ij}^* \lambda_{ij}^{*t} y_{ij}^*(0)$$

and compare (4.24) with (4.20), remembering that

$$z_{ij} \rightarrow z_{ij}^*, \quad \lambda_i \rightarrow \lambda_i^* \quad \text{as } \varepsilon \rightarrow 0.$$

THEOREM 4.2 PART (2): Given ε_1 satisfying condition (4.23), for an arbitrary positive real number η_2 , there exists a number $T_2 > T_1$ such that, for $t > T_2$,

$$(4.25) \quad \frac{|S_j^{(4)} + \zeta S_j^{(5)}|}{|\zeta S_j^{(1)} + S_j^{(2)} + \zeta S_j^{(3)}|} < \eta_2$$

and, for any positive real number η_3 , there exists a number $T_3 > T_2$ such that, for $t > T_3$,

$$(4.26) \quad \frac{|S_j^{(2)} + \zeta S_j^{(3)} + S_j^{(4)} + \zeta S_j^{(5)}|}{\zeta |S_j^{(1)}|} < \eta_3.$$

These two inequalities are direct consequences of the fact that we have arranged the index of roots in such a way that $\lambda_{1I} > \lambda_{iI}$ for $i = 1, \dots, n$, $\lambda_{1I} > \lambda_{1J}$ for $I < J$.

We may summarize the above propositions by saying that the dynamic behavior of the system represented by the nearly decomposable matrix P can be analyzed in the following four stages: (1) $t < T_0$, where S_2 and S_4 dominate the rest of the terms in the summation (4.20); (2) $T_0 < t < T_1$, where S_2 dominates all the rest; (3) $t > T_2 > T_1$ where S_1, S_2 , and S_3 together

dominate S^4 and S^5 ; and (4) $t > T_3 > T_2$, where S_1 finally dominates all the rest.

If $|\lambda_{1I}| \leq 1$ so that the system as a whole is stable, we may use the terminology: stage (1) is the short-run dynamics; stage (2), the short-run equilibrium; stage (3), the long-run dynamics; stage (4), the long-run equilibrium. Note that the terms S_1 , S_2 , and S_3 involve only the aggregate variables $y_{1I}(0)$, $I = 1, \dots, N$.

Thus we conclude: In the short-run, or for stages (1) and (2), we may treat our system as though it consists of N independent subsystems; in the long-run, i.e., for stages (3) and (4), we may look at our system as a set of relations among N aggregative variables, the y_{1I} 's, ignoring the relations within each of the subsystems.

Finally, it is perfectly clear how our argument above can be extended to cover the special case where P^* is stochastic. If P^* is stochastic, the largest roots of all P_I^* 's are unity, and this violates the condition for the existence of the matrix Z^* . However, each submatrix, P_I^* , will have a similar matrix, Λ_I^* . Then, we can define Z_I^* such that

$$(4.27) \quad P_I^* Z_I^* = Z_I^* \Lambda_I^*.$$

Let us now construct, for each Z_I^* and Λ_I^* , an $n \times n$ matrix appropriately bordered with zeros. We shall designate these bordered matrices by the same symbols as are employed for the corresponding $n_I \times n_I$ matrices. We then define

$$Z^* = \sum_{I=1}^N Z_I^* ; \quad \Lambda^* = \sum_{I=1}^N \Lambda_I^*.$$

Although the bordered Z_I^* 's are singular, Z^* defined above is not singular, and we can define Z^{*-1} to be the inverse of Z^* . When Z^* , Λ^* , Z^{*-1} thus defined are inserted into equation (4.2) and the following equations, the argument can proceed in the same manner as above.

V. A SPECIAL CASE: STOCHASTIC MATRICES

The proof of the basic theorem presented in the preceding section has the advantage that it is straightforward and fairly general. In studying this problem, however, we have found it rather difficult to appreciate the implications of the mathematical results adequately, and we feel that another, somewhat more specialized, but substantially equivalent way of stating our theorems is very helpful. At least, it enables us to follow the dynamic process of the system in more detail. In the present section, we shall state these alternative theorems, leaving their proofs to the Appendix. Having given the general proof in the preceding section, the sole purpose of stating

the alternative theorems is to provide for the reader some aid which we ourselves have found useful in getting further insight into the nature of the problem. This being the case, we shall restrict ourselves to the simplest case of the stochastic matrix, and sacrifice mathematical rigor whenever it interferes with the simplicity of our presentation.

Let us restate our systems:

$$(5.1a) \quad x(t+1) = x(t)P,$$

$$(5.1b) \quad x^*(t+1) = x^*(t)P^*.$$

These relations are identical to equations (2.4) and (2.2) except that now x , x^* , and P , P^* are restricted to probability vectors and stochastic matrices, respectively. We assume that the matrix C in (2.3) has the necessary property to keep both P and P^* stochastic. Thus, we may think of the system as having n possible states, the subscript running over these states; $x_i(t)$ is the unconditional probability that the system is in the i th state in period t ; P_{ij} is the conditional probability that the system is in the j th state in period $(t+1)$, given that the system is in state i in period t . We note that the relations (2.8) among the roots of P hold for sufficiently small ε .

We wish to express $x(t)$ and $x^*(t)$ in terms of the roots of the respective matrices. In Section IV we have done so by means of the theorem asserting the existence of a similar diagonal matrix. We shall proceed in a somewhat different manner in this section, as follows: We rewrite (5.1a) and (5.1b) as

$$(5.2a) \quad x(t) = x(0)P^t,$$

$$(5.2b) \quad x^*(t) = x^*(0)P^{*\ell}.$$

Now, for any nonsingular $n \times n$ matrix A whose roots, k_1, \dots, k_n , are distinct, there exists a unique set of n matrices, $\alpha^{(1)}, \dots, \alpha^{(n)}$ with the following characteristics⁶:

- (i) $\alpha^{(\varrho)} \cdot \alpha^{(\sigma)} = \alpha^{(\varrho)}$, $\varrho = 1, \dots, n$. (idempotency),
- (ii) $\alpha^{(\varrho)} \cdot \alpha^{(\sigma)} = 0$; $\varrho \neq \sigma$; $\varrho, \sigma = 1, \dots, n$ (orthogonality),
- (iii) $\sum_{\varrho=1}^n \alpha^{(\varrho)} = I$,
- (iv) $\sum_{\varrho=1}^n k_\varrho \alpha^{(\varrho)} = A$,

where 0 in (ii) is an $n \times n$ matrix whose elements are all zero, and I in (iii) is $n \times n$ identity matrix. It is easy to see that, from these properties, it follows that

$$(v) \quad A^t = \sum_{\varrho=1}^n k_\varrho^t \alpha^{(\varrho)}.$$

⁶ See, for instance, J. H. M. Wedderburn, *Lectures on Matrices*, 1934, pp. 25ff.

Using this representation, we can express P^t thus:

$$(5.3) \quad P^t = \sum_{\varrho=1}^n \lambda_{\varrho}^t \pi^{(\varrho)}$$

where the $\pi^{(\varrho)}$'s are matrices associated with P that satisfy the conditions (i) to (iv) above. Remembering the classification of roots described by (2.8), we divide the terms in the right hand side of (5.3) into three parts:

$$(5.4) \quad P^t = \pi^{(1_I)} + \sum_{I=2}^N \lambda_{1_I}^t \pi^{(1_I)} + \sum_{I=1}^N \sum_{\varrho=2}^{n_I} \lambda_{\varrho_I}^t \pi^{(\varrho_I)}.$$

We cannot expand P^* directly into idempotent matrices as above because the N largest roots of P^* are all unity. However, any non-decomposable submatrix P_I^* of P^* can be so expanded, and we may write

$$(5.5) \quad P_I^{*\ell} = \pi^{*(1_I)} + \sum_{\varrho_I=2}^{n_I} \lambda_{\varrho_I}^{*\ell} \pi^{*(\varrho_I)} \quad \text{for all } I.$$

As in the argument we used at the end of Section IV, let us construct $n \times n$ matrices by bordering those in (5.5) with the appropriate number of rows and columns of zeros, and designate these by the same symbols as those used for the $n_I \times n_I$ matrices in (5.5). Then

$$(5.6) \quad P^* = \sum_{I=1}^N P_I^* = \sum_{I=1}^N \pi^{*(1_I)} + \sum_{I=1}^N \sum_{\varrho_I=2}^{n_I} \lambda_{\varrho_I}^{*t} \pi^{*(\varrho_I)}$$

and for the t th power of P^*

$$(5.7) \quad P^{*\ell} = \sum_{I=1}^N \pi^{*(1_I)} + \sum_{I=1}^N \sum_{\varrho_I=2}^{n_I} \lambda_{\varrho_I}^{*\ell} \pi^{*(\varrho_I)}.$$

When we compare equations (5.4) and (5.7) with equations (4.20) and (4.24), we see that they are analogous expressions, but they also have some differences. Equation (4.20) was very convenient for obtaining information on the variations in the relative influences of various λ 's as the size of ε varied. Equation (5.4) gives clearer indications of the characteristics of the time path of x , as we shall show presently.

We first note that, since the $\lambda_{1_I}^*$'s are unity and the λ_{1_I} 's are very close to unity for all I , the first summation term on the right hand side of equation (5.7) and the first and second summation terms of equation (5.4) remain almost unchanged for a relatively small t . This means that, for x to behave very much like x^* for small t as indicated by Theorem 4.2, Part 1, $\pi^{*(\varrho_I)}$ must approach $\pi^{*(\varrho_I)}$ for $\varrho = 2, \dots, n$ and $I = 1, \dots, N$, as ε goes to zero. Further-

more, if x is to exhibit the behavior described by Part 2 of Theorem 4.2 when t becomes so large that the λ_I^t 's are no longer nearly unity, the elements of $\pi^{(1)_I}$, $I = 1, \dots, N$, must be functions of j , J , and I , but independent of i .

Before we proceed to state our basic propositions as theorems, we need a few additional definitions, which we list below:

$P_{ij}^{(t)}$: elements of the matrix P^t , i.e., the t th power of the matrix P ;

$P_{ij}^{*(t)}$: elements of the matrix P^{*t} ;

\bar{x}^* : equilibrium value of x^* ;

$$\bar{x}_{i|I}^* = \frac{\bar{x}_I^*}{\sum_{i_I=1}^{n_I} \bar{x}_{i_I}^*}, \quad I = 1, \dots, N.$$

Note that the vector $[\bar{x}_{i|I}^*]$, $i = 1, \dots, n_I$, is the characteristic vector of P_I^* associated with the root of unity.

$$(5.8) \quad \pi_{IJ}^{(1)_l} = \sum_{i_I=1}^{n_I} \sum_{j_J=1}^{n_J} \bar{x}_{i_I|I}^* \pi_{i_I J}^{(1)_l}$$

for $l = 1, \dots, N$, $I = 1, \dots, N$, and $J = 1, \dots, N$.

$$(5.9) \quad P_{IJ} = \sum_{l=1}^N \lambda_{1_l} \pi_{IJ}^{(1)_l}.$$

The subscript of λ and superscript of π are written 1_l instead of the usual 1_I to avoid confusion.

In terms of these definitions, the following three theorems are proved in the Appendix.

THEOREM 5.1: *For an arbitrary positive real number ξ_2 there exists a number ε_2 such that for $\varepsilon < \varepsilon_2$,*

$$\max_{i,j} |\pi_{ij}^{(\varrho_l)} - \pi_{ij}^{*(\varrho_l)}| < \xi_2$$

for $\varrho_l = 2, \dots, n_l$, and $l = 1, \dots, N$.

THEOREM 5.2: *For an arbitrary positive real number ω there exists a number ε_ω such that for $\varepsilon < \varepsilon_\omega$,*

$$\max_{i,j} |\pi_{ij}^{(1)_l} - \bar{x}_{j|J}^* \pi_{IJ}^{(1)_l}| < \omega$$

for $l = 1, \dots, N$, $I = 1, \dots, N$, $J = 1, \dots, N$.

THEOREM 5.3: *The right hand side of equation (5.9) is the idempotent expansion of the matrix P_{IJ} .*

The implications of the above theorems are quite clear. Since the λ_{1_l} are almost unity for $l = 1, \dots, N$ as indicated in (2.8), for a relatively small t , say $t < T_2$, $\lambda_{1_l}^t$, $l = 1, \dots, N$, will be very close to unity. Hence the first two terms on the right hand side of equation (5.4) will not change very much, while the first term on the right hand side of equation (5.7) will not change at all. Hence, for $t < T_2$, the time behavior of x and x^* are completely determined by the last terms on the right hand side of equation (5.4) and (5.7), respectively. But, Theorem 5.1 asserts that the π 's appearing in the last term of equation (5.4) can be made as close as we please to the corresponding π^* 's in equation (5.7), by taking ϵ sufficiently small. We may recall that $\lambda_i \rightarrow \lambda_i^*$ as $\epsilon \rightarrow 0$. Hence, for $t < T_2$ the time path of x must be very close to the time path of x^* . Note that T_2 can be made as large as we please by taking ϵ sufficiently small.

Since the $\lambda_{1_l}^*$ are independent of ϵ , and for $i_l = 2, \dots, n_l$ less than unity as is indicated in (2.7), for any positive real ξ_1 we can define T_1^* such that for $t > T_1^*$ the absolute value of the last summation term on the right hand side of (5.7) is less than ξ_1 . T_1^* is independent of ϵ . On the other hand, (2.8) and Theorem 5.1 insure not only that for any positive real ξ_1 there exists T_1 such that for $t > T_1$ the absolute value of the last summation term on the right hand side of (5.4) is less than ξ_1 , but that $T_1 \rightarrow T_1^*$ for the same ξ_1 as $\epsilon \rightarrow 0$. That is, for any positive real number ξ_1 , there exist T_1^* and T_1 such that

$$(5.10) \quad \left| \sum_{l=1}^N \sum_{e_l=2}^{n_l} \lambda_{1_l}^{*t} \pi^{*(e_l)} \right| < \xi_1 \quad \text{for } t > T_1^*,$$

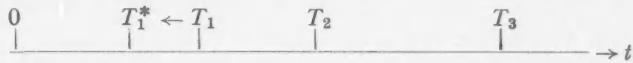
$$(5.11) \quad \left| \sum_{l=1}^N \sum_{e_l=2}^{n_l} \lambda_l^t \pi(e_l) \right| < \xi_1 \quad \text{for } t > T_1 \text{ and} \\ T_1 \rightarrow T_1^* \text{ as } \epsilon \rightarrow 0.$$

Since T_2 can be made as large as we please by taking ϵ sufficiently small while T_1^* is independent of ϵ , let us take ϵ such that T_2 is very much larger than T_1 . Provided that ϵ is not identically zero so that none of λ_{1_l} except λ_1 is identically unity, the second summation term on the right hand side of equation (5.4) will eventually become negligible as t becomes indefinitely large. Let us define T_3 , corresponding to an arbitrary positive real number ξ_3 , such that for $t > T_3$

$$(5.12) \quad \max_{i,j} \left| \sum_{l=2}^N \lambda_{1_l}^t \pi_{ij}^{(1_l)} \right| < \xi_3.$$

T_3 also increases without limit as $\epsilon \rightarrow 0$.

We show the relations among the various T 's schematically below:



For $T_2 < t < T_3$, corresponding to the period which we called the long run dynamics in Section IV, the last summation term on the right hand side of equation (5.4) has very little influence on the time path of x . The path is determined by the first and second summation terms on the right hand side of equation (5.4). But, Theorem 5.2 asserts that the elements of $\pi^{(1)}_l$, $l = 1, \dots, N$, are functions only of j , I , J , and independent of i . That is, for any I, J , and l , $[\pi_{I|J}^{(1)}, \dots, \pi_{I|J}^{(1)}, \dots, \pi_{I|J}^{(1)}]$, are proportional to the characteristic vector of P_J^* associated with the root of unity, and are the same for $i_I = 1, \dots, n_I$. Hence, $[x_{1|I}, \dots, x_{I|I}, \dots, x_{n|I}]$ will move, for the period $T_2 < t < T_3$, keeping a roughly constant proportionality relationship among individual elements for a given I . This permits us, following the Hicks-Lange condition, to replace the n -dimensional vector by an N -dimensional vector, and the $n \times n$ matrix P by an $N \times N$ matrix $\|P_{IJ}\|$.

The usefulness of Theorem 5.3 will become apparent in the next section when we shall discuss the application of our theorems to the problems of inverting nearly decomposable matrices. Here we merely note that, since (1) $\pi_{i|I}^*$ has $\tilde{x}_{1|I}^*$ as its rows where its elements are not zero, (2) $\pi_{i|I}^*$ are orthogonal to $\pi_{i'|I}^*$ for $i = 2, \dots, n$, (3) $\pi_{i|I} \rightarrow \pi_{i|I}^*$ for $i_I = 2, \dots, n_I$, we can express P_{IJ} as

$$(5.13) \quad P_{IJ} = \sum_{i_I=1}^{n_I} \sum_{j_J=1}^{n_J} \tilde{x}_{i|I}^* P_{i|I} j_J.$$

VI. A NUMERICAL ILLUSTRATION

The above results can be understood more readily with the aid of a simple numerical example. Let us consider a nearly decomposable stochastic matrix

$$(6.1) \quad P = \begin{vmatrix} .9700 & .0295 & .0005 & 0 \\ .0200 & .9800 & 0 & 0 \\ 0 & 0 & .9600 & .0400 \\ .0002 & .0002 & .0396 & .9600 \end{vmatrix}$$

and the corresponding completely decomposable matrix

$$(6.2) \quad P = \begin{vmatrix} .9700 & .0300 & 0 & 0 \\ .0200 & .9800 & 0 & 0 \\ 0 & 0 & .9600 & .0400 \\ 0 & 0 & .0400 & .9600 \end{vmatrix} = \begin{vmatrix} P_1^* & 0 \\ 0 & P_2^* \end{vmatrix}.$$

We can compute $\tilde{x}_j^*|_J$:

$$(6.3) \quad [\tilde{x}_j^*]_1 = [.4 .6],$$

$$(6.4) \quad [\tilde{x}_j^*]_2 = [.5 .5];$$

and P_{IJ} :

$$(6.5) \quad \|P_{IJ}\| = \begin{vmatrix} .9998 & .0002 \\ .0002 & .9998 \end{vmatrix}.$$

The roots of P and their selected powers are

$$\lambda_1 = 1;$$

$$\lambda_2 = .9996; \quad \lambda_2^{128} = .9511; \quad \lambda_2^{(128)^2} = .001724;$$

$$\lambda_3 = .952004; \quad \lambda_3^{128} = 1.845 \times 10^{-3}; \quad \lambda_3^{(128)^2} = 1.091 \times 10^{-350};$$

$$\lambda_4 = .9202; \quad \lambda_4^{128} = 2.381 \times 10^{-5}; \quad \lambda_4^{(128)^2} = 1.622 \times 10^{-592}.$$

The idempotent expansion of P is given by

$$(6.6) \quad P^t = \pi^{(1)} + \lambda_2^t \pi^{(2)} + \lambda_3^t \pi^{(3)} + \lambda_4^t \pi^{(4)},$$

i.e.,

$$(6.7) \quad P^{(t)} = (1.)^t \begin{vmatrix} 0.20050125 & 0.29824561 & 0.25062657 & 0.25062657 \\ 0.20050125 & 0.29824561 & 0.25062657 & 0.25062657 \\ 0.20050125 & 0.29824561 & 0.25062657 & 0.25062657 \\ 0.20050125 & 0.29824561 & 0.25062657 & 0.25062657 \end{vmatrix}$$

$$+ (.9996)^t \begin{vmatrix} .198685261 & .296522699 & -.248356576 & -.250865228 \\ .202740062 & .302574184 & -.253425078 & -.255984927 \\ -.199496221 & -.297732997 & .249370276 & .251889168 \\ -.197501259 & -.294755666 & .246876573 & .249370276 \end{vmatrix}$$

$$+ (.952004)^t \begin{vmatrix} .6004874 & -.5952777 & .0019774 & -.0807124 \\ -.4030171 & .3995206 & -.0013271 & .0541701 \\ -.0008141 & .0008070 & -.0000027 & .0001094 \\ .0000403 & -.0000399 & .0000001 & -.0000054 \end{vmatrix}$$

$$+ (.92019)^t \begin{vmatrix} .0002636 & -.0007692 & -.0062681 & .0062939 \\ -.0000878 & .0002397 & .0020960 & -.0021003 \\ -.0209766 & .0577601 & .5007744 & -.5017791 \\ .0208776 & -.0569881 & .4983870 & .4993957 \end{vmatrix}.$$

From (6.7), we also readily compute P^{128} (the matrix for the "middle-run" dynamic system) and $P^{(128)^2}$, the matrix for the long run:

$$(6.8) \quad P^{128} = \begin{vmatrix} .390089 & .579037 & .016631 & .014244 \\ .392503 & .586246 & .011831 & .009419 \\ .009465 & .013138 & .487509 & .489888 \\ .011385 & .015999 & .485107 & .487509 \end{vmatrix},$$

$$(6.9) \quad P^{(128)^2} = \begin{vmatrix} .200776 & .298656 & .250286 & .250282 \\ .200782 & .298664 & .250279 & .250275 \\ .200222 & .297829 & .250973 & .250976 \\ .200225 & .297833 & .250970 & .250973 \end{vmatrix}.$$

Note that, if we neglect roots smaller than 0.002, we have

$$(6.10) \quad P^{128} = \pi^{(1)} + \lambda_2^{128} \pi^{(2)},$$

$$(6.11) \quad P^{(128)^2} = \pi^{(1)}.$$

The reader's attention is called to the behavior of elements of P^t from P^{128} to $P^{(128)^2}$ —elements of P_1^t and P_2^t maintain the same proportion over the columns and independently of rows within each submatrix while moving toward the full equilibrium.

It is of some interest to see whether or not our results are useful for inverting the nearly decomposable matrix, P . We know that

$$(6.12) \quad P^{-1} = \sum_{l=1}^N \sum_{\theta_l=1}^{n_l} \lambda_l^{-1} \pi^{(\theta_l)}.$$

This relation is of little use for computational purposes, since the π 's are very difficult to compute. But, in the case of nearly decomposable matrices, this leads to a potentially useful approximation. We note that, from Theorem 5.1,

$$(6.13) \quad P^{-1} \simeq \sum_{l=1}^N \lambda_l^{-1} \pi^{(1_l)} + P^{*-1} - \sum_{I=1}^N \pi^{*(1_I)}.$$

Since P^* is completely decomposable, it is much simpler to invert. The $\pi^{*(1_l)}$ are, of course, matrices with identical rows, and their rows are the characteristic vectors of the respective submatrices P_I , $I = 1, \dots, N$, associated with roots of unity. Hence, our problem reduces to that of computing the first term on the right hand side of (6.13). By Theorem 5.2, we can write

$$(6.14) \quad \sum_{l=1}^N \lambda_l^{-1} \pi^{(1_l)} = \bar{x}_{J|J}^* \sum_{l=1}^N \lambda_l^{-1} \pi^{(1_l)}.$$

But, by Theorem 5.3, we have

$$(6.15) \quad \sum_{l=1}^N \lambda_l^{-1} \pi^{(1_l)} = P_{IJ}^{(-1)}.$$

Hence,

$$(6.16) \quad \sum_{l=1}^N \lambda_{1_l}^{-1} \pi_{ij}^{(1_l)} = \bar{x}_{j|J}^* P_{IJ}^{(-1)}.$$

To summarize: in order to obtain P^{-1} , (i) find the inverses of P_I^* ; (ii) compute $\bar{x}_{j|J}^*$ and form $\pi^{*(1_l)}$ directly from them; (iii) form the aggregate matrix $\|P_{IJ}\|$ by (5.13); (iv) invert $\|P_{IJ}\|$; (v) substitute these results into (6.13) to obtain P^{-1} .

Going back to the numerical example of P and P^* given by (6.1) and (6.2), we readily find

$$(6.17a) \quad P_1^{*-1} = \begin{vmatrix} 1.031578947 & -0.31578947 \\ -0.021032617 & 1.021052620 \end{vmatrix},$$

$$(6.17b) \quad P_2^{*-1} = \begin{vmatrix} 1.043478260 & -0.43478260 \\ -0.43478260 & 1.043478260 \end{vmatrix}.$$

$[\bar{x}_{j|J}^*]$ and $\|P_{IJ}\|$ have already been given by (6.3), (6.4), and (6.5). We also have

$$(6.18) \quad \|P_{IJ}\|^{-1} = \begin{vmatrix} 1.00020008 & -0.00020008 \\ -0.00020008 & 1.00020008 \end{vmatrix}.$$

Substituting these results into (6.16), we have

$$(6.19) \quad \sum_{l=1}^2 \lambda_{1_l}^{-1} \pi_{ij}^{(1_l)} = \begin{vmatrix} .400080032 & .600120048 & -.000100040 & -.000100040 \\ .400080032 & .600120048 & -.000100040 & -.000100040 \\ -.000800320 & -.001200480 & .500100040 & .500100040 \\ -.000800320 & -.001200480 & .500100040 & .500100040 \end{vmatrix}.$$

Further substitutions into (6.13) yield

$$(6.20) \quad P^{-1} \simeq \begin{vmatrix} 1.031658979 & -0.34588990 & -.000100040 & -.000100040 \\ -0.020972585 & 1.021172668 & -.000100040 & -.000100040 \\ -.000800320 & -.001200480 & 1.043578300 & -.043378220 \\ -.000800320 & -.001200480 & -.043378220 & 1.043578300 \end{vmatrix}.$$

This result may be compared with the inverse obtained by the perturbation method, which is probably the most efficient way of obtaining the approximate inverse of a matrix of the type treated here.⁷

Let C be defined by

$$(6.21) \quad P = P^* + \varepsilon C$$

where ε is a small positive number. Let \hat{C} be defined by

$$(6.22) \quad P^{-1} = P^{*-1} + \varepsilon \hat{C}$$

⁷ See, for instance, Courant, R. and D. Hilbert, *Methods of Mathematical Physics*, pp. 42-43 and 343-350.

where ε is the same number as in equation (6.21). Then, we have

$$(6.23) \quad P \cdot P^{-1} = (P^* + \varepsilon C)(P^{*-1} + \varepsilon \hat{C}).$$

Ignoring the terms in higher powers of ε , we get:

$$(6.24) \quad \hat{C} = -P^{*-1}CP^{*-1}.$$

If, in the numerical example of (6.1) and (6.2), we take ε to be 0.0001, we have

$$(6.25) \quad C = \begin{vmatrix} 0 & -5 & 5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 2 & -4 & 0 \end{vmatrix}$$

$$(6.26) \quad \varepsilon \hat{C} = \begin{vmatrix} -.000011 & .000527 & -.000538 & +.000022 \\ .000000 & -.000011 & +.000011 & .000000 \\ .000009 & .000009 & -.000018 & .000001 \\ .000211 & .000206 & .000436 & .000018 \end{vmatrix}.$$

$$(6.27) \quad P^{-1} \approx \begin{vmatrix} 1.031568 & -.031052 & -.000538 & .000022 \\ -.021053 & 1.021042 & .000011 & .000000 \\ .000009 & .000009 & 1.043460 & -.043477 \\ -.000211 & -.000206 & -.043042 & 1.043460 \end{vmatrix}.$$

The direct multiplication of P^{-1} given by (6.27) with P given by (6.1) shows that the approximate inverse obtained by the perturbation method is accurate to the 6th decimal place. The estimate of the inverse given by (6.20), obtained by the aggregation procedure, has an error of about 0.0001 in each element in the product. This relatively large error is caused by the restriction, implicit in the aggregation procedure, that the matrix elements *within* each subset are replaced by the eigenvectors corresponding to the middle-run equilibrium.

On the other hand, the perturbation method assumes implicitly that there is a zero probability that the system will move more than once from a state belonging to one subset to a state belonging to another subset. Hence, we can conclude that, while the perturbation method gives us a good approximation to the short-run behavior of the system, and hence to P^{-1} , its approximation to the middle-run behavior (and hence to $(P^{-1})^t$ for relatively large t) will be much less good. The aggregation method permits us to study this middle-run behavior, as we have seen earlier in this section.

For those readers who may be interested in the comparison, we give P^{-1} obtained by equating $t = -1$ in equation (6.7):

$$(6.28) \quad P^{-1} = \begin{vmatrix} 1.0306708 & -.0302846 & -.0026491 & -.0783382 \\ -.0206318 & 1.0209660 & -.0019932 & .0492225 \\ .0440999 & -.1027626 & 1.0505158 & -.0510149 \\ -.0408082 & .1067156 & -.0495267 & 1.0505238 \end{vmatrix}$$

VII. SOME CONCLUDING COMMENTS

In the preceding sections, we have analyzed the structure of dynamic systems represented by nearly-decomposable matrices. We have seen that such systems may be viewed as composite systems, constructed by the superposition of: (1) terms representing interactions of the variables within each subsystem; and (2) terms representing interactions among the subsystems. We concluded that, over a relatively short period, the first group of terms dominates the behavior of the system, and hence each subsystem can be studied (approximately) independently of other subsystems. Over a relatively long period of time, on the other hand, the second group of terms dominates the behavior of the system, and the whole system moves, keeping the state of equilibrium within each subsystem—i.e., the variables within each subsystem move roughly proportionately. Hence, the variables within each subsystem can be aggregated into indexes representing the subsystem.

Thus, the system of variables in the case just described can be represented as a two-level hierarchy, with the aggregative variables at the higher level. Now, there is no reason why we need to restrict ourselves to a two-level hierarchy. For, in such a hierarchy, each of the subsystem variables at the lower level might be an aggregate of variables at a still lower level of aggregation. The matrix of a three-level hierarchy, for example, might look something like this:

$$\begin{vmatrix} P_1 & Q_1 & S_1 & S_2 \\ Q_2 & P_2 & S_3 & S_4 \\ \hline R_1 & R_2 & P_3 & Q_3 \\ R_3 & R_4 & Q_4 & P_4 \end{vmatrix}$$

In this matrix, the elements of the submatrices designated as Q 's are of the first order of smallness, and the elements of the R 's and S 's of the second order of smallness. At the first level of aggregation, there will be four aggregative variables corresponding to the four submatrices along the diagonal, respectively. At the second level of aggregation, there will be two aggregative variables, corresponding to the blocks indicated by broken lines.

It is of some interest to consider the implications of our analysis concerning the computation involved in inverting a matrix in this context. To invert a

matrix like the one above, we would first invert the matrices P_1 , then the two aggregative matrices

$$\left\| \begin{array}{cc} P_1 & Q_1 \\ Q_2 & P_2 \end{array} \right\| \text{ and } \left\| \begin{array}{cc} P_3 & Q_3 \\ Q_4 & P_4 \end{array} \right\|,$$

and finally the second level aggregative matrix.

In ordinary methods of matrix inversion, the number of multiplications increases as the cube of the size of the matrix. On the other hand, under the method of inversion suggested above, if the size of the largest matrix to be inverted at any level of aggregation remains constant as n increases, then the number of matrices to be inverted will increase proportionately with n , their size will not increase, and the total number of multiplications will increase only slightly more than proportionately with n .

It may be objected that decomposable matrices are rare objects, mathematically speaking, and nearly decomposable matrices almost as rare. For if the elements of a matrix are selected in any ordinary way by a random process, the probability that the matrix will be decomposable is zero. There is every reason to believe, however, not only that near-decomposability is a very common characteristic of dynamic systems that exist in the real world, but also that many economists and other social scientists conduct their research as though this were the case. As we have pointed out in the introductory remarks, every time economists construct indices representing groups of variables and construct a theory in terms of such indices, they are implicitly assuming that the economic system they study is constructed in the hierarchical form described above.

It is interesting to note that many recent discussions on the problems of aggregation can be looked at as studies of various aspects of a system like the one analyzed here.

Interest in the aggregation of the static input-output matrices largely stems from the desire to facilitate the inversion of the matrices involved.⁸ As we have shown in Section VI, the method suggested by our analysis is not so efficient as the perturbation method, in the sense that the product of the original matrix with its inverse obtained by our method deviates further from

⁸ See, for instance, J. B. Balderston and T. M. Whitin, *op. cit.*; O. Morgenstern and Whitin, "Aggregation and Errors in Input-Output Models," Logistics Research Project, George Washington University, cited in F.T. Moore, "A Survey of Current Inter-Industry Models," in *Input-Output Analysis, an Appraisal*, Studies in Income and Wealth, vol. 18, p. 228 (National Bureau of Economic Research). J. C. H. Fei, "A Fundamental Theorem for the Aggregation Problem of Input-Output Analysis," *Econometrica*, 1956, pp. 400-412; S. B. Noble, "Structure and Classification in Resource Flow Models," The George Washington University Logistic Research Project, Serial T-100/59, May 19, 1959; and A. Ghosh, "Input-Output Analysis with Substantially Independent Groups of Industries, *Econometrica*, 1960, pp. 88-96.

the identity matrix than the product of the original matrix with the inverse obtained by the perturbation method. Our analysis provides, however, a better ground for interpreting the approximate result than the mechanical application of the perturbation method. Furthermore, the goodness of the approximation must be judged in terms of the decisions that will be based on the approximation rather than the closeness of the product to the identity matrix, and in this sense, it is by no means clear that the perturbation method is superior to ours.

Some of the recent work of Theil can also be related to our analysis.⁹ One of his aggregation problems may be looked at in the following way: Suppose that we have a set of observations over time on the x 's, and estimate from these observations the matrix of coefficients P . Suppose further that we classify the x 's into a few groups, construct indices representing these groups, consider a new linear system relating these indices, and estimate the new aggregate coefficients from the same set of observations. Theil investigated, among other things, the relations between the micro and macro coefficients then estimated. Theil's results show that the conditions under which there exists a relatively simple relation between micro and macro variables are very severe. Our result suggests that, if the underlying structure generating the x 's is nearly decomposable, then, as the x 's are aggregated, the unit of time over which observations are made should be changed accordingly. However, a more complete analysis of the relation between our results and those of Theil must be deferred to a future study.

In a paper presented at the December, 1959, meetings of the Econometric Society, F. Fisher explored yet another problem that is closely related to our analysis.¹⁰ The conditions that permit the identification of a structural relationship ordinarily state that a certain set of structural coefficients are identically zero, and the estimation can be carried out on the assumption that there exist no relations among the variables under consideration other than those explicitly stated in the system. Suppose, however, that these conditions are only approximately satisfied. Fisher has shown that, as the approximation of these conditions becomes better and better, the estimates of the structural parameters, obtained by the usual estimation method such as the limited information method or the generalized least squares method, are asymptotically consistent.

In the framework of our analysis, Fisher's problem is analogous to that of specifying the conditions under which one of the subsystems can be treated in isolation from all other parts of the system for the purposes of estimation.

The comparison of our result with that of Fisher immediately raises two

⁹ H. Theil, *Linear Aggregation of Economic Relations*, 1954.

¹⁰ F. M. Fisher, *op. cit.*

important questions. The first has already been suggested by Fisher; since his results apply to what he calls block recursive systems, nearly triangular matrices must possess some properties analogous to those we have shown in nearly decomposable matrices. The investigation of these analogous properties would open a way to generalize Goodwin's justification of partial dynamics to the case where the coupling is only "nearly-unilateral," as well as to the case where the whole system is "nearly-block-recursive."¹¹ The second is that, according to our analysis, even a very weak link will eventually make its influence felt given a long enough period of time. Thus, in interpreting Fisher's result, we must be careful to choose the appropriate period over which the observations are made.

Finally, we note that there are a number of other discussions in the economic literature upon which the notion of nearly-decomposable systems appears to throw some light. We have already pointed out that the argument here may be regarded as a statement of the circumstances under which the Lange-Hicks condition will be satisfied. It can easily be seen from our analysis that if the micro-system is dynamically stable, this will also be true of the aggregated system, since the characteristic roots of the aggregative matrix are also roots of the original matrix. This stability theorem has been proved earlier by Tamotsu Yokoyama.¹² Yokoyama assumes that the Lange-Hicks condition is satisfied, and derives the stability theorem from this assumption.

Samuelson points out that aggregation of commodities can seek its justification in either of two kinds of principles that, at first blush, appear rather antithetical to each other.¹³ On the one hand (the Lange-Hicks condition), we can aggregate the parts of a subsystem when these are much more closely linked with each other than they are with the rest of the system. On the other hand, we can aggregate a set of variables if each of them is linked with the remainder of the system in just the same way as are the others. Our analysis of near-decomposability shows that the former condition is really a special case of the latter. For if x_i and x_j are variables belonging to different subsets of a nearly decomposable system, then p_{ij} is very small, but $p_{ij}^{(t)}$, for sufficiently large t , is almost independent of i . That is to say, the linkage between i and j is negligible in the short run, and satisfies the second condition in the middle run for which it is not negligible.

Carnegie Institute of Technology and Massachusetts Institute of Technology

¹¹ R. M. Goodwin, "Dynamic Coupling with Especial Reference to Markets Having Production Lags," *Econometrica*, 1947, pp. 181-204.

¹² Tamotsu Yokoyama, "A Theory of Composite Commodity," *Osaka Economic Papers*, May, 1952.

¹³ P. A. Samuelson, *Foundations of Economic Analysis*, 1948, pp. 144-46.

APPENDIX

PROOF OF THEOREM 5.1:

We note first that the elements of the t th power of a matrix are continuous functions of the elements of the original matrix. Hence, for any positive real number ξ_2' and an arbitrary integer T_2 there exists a number ϵ_2' such that for $t < T_2$ and $\epsilon < \epsilon_2'$

$$(A.1) \quad \max_{i,j} |P_{ij}^{[t]} - P_{ij}^{*[t]}| < \xi_2'$$

$P_{ij}^{[t]}$ and $P_{ij}^{*[t]}$ are elements of t th power of matrices P and P^* .

Substituting (5.4) and (5.7) into (A.1), we obtain

$$(A.2) \quad \max_{i,j} \left| \sum_{I=1}^N \pi_{ij}^{*(1_I)} - \pi_{ij}^{(1_I)} - \sum_{I=2}^N \lambda_{1_I}^t \pi_{ij}^{(1_I)} + \sum_{I=1}^N \sum_{\varrho_I=2}^n \lambda^{*t} \pi_{ij}^{*(\varrho_I)} - \sum_{I=1}^N \sum_{\varrho_I=2}^n \lambda_{\varrho_I}^t \pi_{ij}^{(\varrho_I)} \right| < \xi_2'.$$

Introducing for convenience

$$(A.3) \quad D_{ij}^{[t]} = \sum_{I=1}^N \pi_{ij}^{*(1_I)} - \pi_{ij}^{(1_I)} - \sum_{I=2}^N \lambda_{1_I}^t \pi_{ij}^{(1_I)},$$

we rewrite (A.2):

$$(A.4) \quad \max_{i,j} |D_{ij}^{[t]} + \sum_{I=1}^N \sum_{\varrho_I=2}^n (\lambda_{\varrho_I}^{*t} - \lambda_{1_I}^t) \pi_{ij}^{(\varrho_I)} + \sum_{I=1}^N \sum_{\varrho_I=2}^n \lambda_{\varrho_I}^{*t} (\pi_{ij}^{*(\varrho_I)} - \pi_{ij}^{(\varrho_I)})| < \xi_2'.$$

By (2.6), $(\lambda_{\varrho_I}^{*t} - \lambda_{1_I}^t) \rightarrow 0$ as $\epsilon \rightarrow 0$. Hence, for any positive real ξ_2 , we can choose a new ϵ_2'' so that for $\epsilon < \epsilon_2''$,

$$(A.5) \quad \max_{i,j} |D_{ij}^{[t]} + \sum_{I=1}^N \sum_{\varrho_I=2}^n \lambda_{\varrho_I}^{*t} (\pi_{ij}^{*(\varrho_I)} - \pi_{ij}^{(\varrho_I)})| < \xi_2'.$$

But, since λ^* are all distinct for $i_I \geq 2$, and $D_{ij}^{[t]}$ becomes independent of t as $\lambda_{1_I}^t \rightarrow 1$ for all I , this inequality can hold for all $t \leq T_0$ only if the coefficients of λ^* become vanishingly small; hence, for any positive real number ξ_2 , we can choose a value of ϵ, ϵ_2 so that for $\epsilon < \epsilon_2$,

$$(A.6) \quad \max_{i,j} |\pi_{ij}^{*(\varrho_I)} - \pi_{ij}^{(\varrho_I)}| < \xi_2$$

for $\varrho_I = 2, \dots, n_I$ and $I = 1, \dots, N$.

PROOF OF THEOREM 5.2:

As indicated in the text and in the above proof of Theorem 5.1, we know that (1) as $\epsilon \rightarrow 0$, $T_1 \rightarrow T_1^*$, (2) T_2 can be made as large as we please for a given ξ_2 by taking small enough ϵ , and (3) T_1^* is independent of ϵ . Hence, by taking ϵ sufficiently small, we can make T_2 very much larger than T_1 for a given set of ξ_1 and ξ_2 . Then, for $T_1 < t < T_2$, we have

$$(A.7) \quad |P_{ij}^{[t]} - \tilde{x}_{j_J|J}^{* \Delta_{IJ}}| = |(P_{ij}^{[t]} - P_{ij}^{*[t]}) + (P_{ij}^{*[t]} - \tilde{x}_{j_J|J}^{* \Delta_{IJ}})| \leq \xi_2' + \xi_1$$

for $i_I = 1, \dots, n_I$, $j_J = 1, \dots, n_J$, $I, J = 1, \dots, N$, and where Δ_{IJ} is Kronecker's delta.

Let us now consider $P_{ij}^{[t]}$ for a large enough t so that it can be expressed in the following scheme:

$$(A.8) \quad P_{ij}^{[t]} = \sum_k \sum_m P_{ik}^{[t_1]} P_{km}^{[t_2]} P_{mj}^{[t_3]}$$

where $t = t_1 + t_2 + t_3$, and

$$\begin{aligned} T_1 &< t_1 < T_2, \\ T_1 &< t_3 < T_2, \\ T_1 &< t_1 + t_3 < T_2, \\ T_2 &< t_2. \end{aligned}$$

Then, conditions (5.11) and (A.7) assures us that, for any real positive number ω_0 , there exists a value of ϵ , ϵ_{ω_0} , such that for $\epsilon < \epsilon_{\omega_0}$ we have

$$(A.9) \quad \sum_{l=1}^N \lambda_{1_l} \pi_{ij}^{(1_l)} = \sum_k \sum_m [\tilde{x}_{K|K}^* \Delta_{IK}] [\sum_{l=1}^N \lambda_{1_l}^{t_2} \pi_{km}^{(1_l)}] [\tilde{x}_{j|L}^* \Delta_{MJ}] + \omega_{ij}$$

$$\max_{ij} |\omega_{ij}| < \omega_0$$

where the usual 1_I , the subscript of λ and the superscript of π_{ij} , is replaced by 1_l in order to avoid the confusion with the identification of a block of rows and columns of P in this context.

(A.9) becomes, after a rearrangement of terms,

$$(A.10) \quad \sum_{l=1}^N \lambda_{1_l}^{t_2} \pi_{ij}^{(1_l)} = \sum_{l=1}^N \lambda_{1_l}^{t_2} x_{j|J}^* [\sum_{I=1}^{n_I} \tilde{x}_{I|I}^* \sum_{J=1}^{n_J} \pi_{i'j'}^{(1_l)}] + \omega_{ij}$$

$$\max_{ij} |\omega_{ij}| < \omega_0.$$

For this to hold for all permissible values of t_2 , we must have

$$(A.11) \quad \lambda_{1_l}^{(t_1+t_3)\pi_{ij}^{(1_l)}} = x_{j|J}^* \sum_{I=1}^{n_I} \tilde{x}_{I|I}^* \sum_{J=1}^{n_J} \pi_{i'j'}^{(1_l)} + \omega_l,$$

$$\omega_l \rightarrow 0 \text{ as } \omega_0 \rightarrow 0, l = 1, \dots, N.$$

Comparison of (A.11) with the definition (5.8), remembering that $\lambda_{1_l}^t \rightarrow 1$ for $l = 1, \dots, N$ and for $t < T_2$, and that $t_1 + t_3 < T_2$, yields the result summarized in Theorem 5.2.

PROOF OF THEOREM 5.3:

Let us define

$$(A.12) \quad \tilde{P}_{ij} = \sum_{l=1}^N \lambda_{1_l} \pi_{ij}^{(1_l)}.$$

The omission of terms $2_l, \dots, n_l$ will guarantee that we may write

$$(A.13) \quad \tilde{P}_{ij}^{[t]} = \tilde{x}_{j|J}^* \sum_{l=1}^N \lambda_{1_l}^{t_2} \pi_{IJ}^{(1_l)}.$$

Suppose that there exists an $N \times N$ nonsingular matrix Q satisfying the condition

$$(A.14) \quad \tilde{P}_{ij} = \tilde{x}_{j|J}^* Q_{IJ},$$

$$(A.15) \quad \tilde{P}_{ij}^{[t]} = \tilde{x}_{j|J}^* Q_{IJ}^{[t]},$$

for all t .

Let the idempotent expansion of Q be given by

$$(A.16) \quad Q_{IJ}^{(t)} = \sum_{l=1}^N \mu_l^t \phi_{IJ}^{(l)}$$

where μ_l are characteristic roots and $\phi^{(l)}$ are corresponding idempotent matrices. Substituting (A.16) in (A.15) and comparing the result with (A.13), we must have

$$(A.17) \quad \sum_{l=1}^N \mu_l^t \phi_{IJ}^{(l)} = \sum_{l=1}^N \lambda_{1_l}^t \pi_{IJ}^{(1_l)}$$

for all t . As we shall show presently, this implies that

$$(A.18) \quad \mu_l = \lambda_{1_l} \quad (l = 1, \dots, N),$$

$$(A.19) \quad \phi^{(l)} = \pi^{(1_l)} \quad (l = 1, \dots, N).$$

These conditions obviously imply that

$$(A.20) \quad [P_{IJ}] = [Q_{IJ}],$$

proving Theorem 5.3.

To justify equation (A.18), we note that $\lambda_{1_1} > \lambda_{1_2} > \dots > \lambda_{1_N}$.

Let us also assume, without loss of generality, that the subscript of μ is arranged so that $\mu_1 > \mu_2 > \dots > \mu_N$. Dividing both sides of (A.17) by λ_{1_1} and rearranging terms, we have

$$(A.21) \quad \pi_{IJ}^{(1_1)} = \phi_{IJ}^{(1)} \left(\frac{\mu_1}{\lambda_{1_1}} \right)^t + \sum_{l=2}^N \pi_{IJ}^{(1_l)} \left(\frac{\lambda_{1_l}}{\lambda_{1_1}} \right)^t + \sum_{l=2}^N \phi_{IJ}^{(1_l)} \left(\frac{\mu_l}{\lambda_{1_1}} \right)^t.$$

Because of the way the matrix Q , and hence μ are defined, $|\mu_l| \leq 1$ for $l = 1, \dots, N$. Hence, as t becomes large, (A.21) may be written as

$$(A.22) \quad \pi_{IJ}^{(1_1)} = \phi_{IJ}^{(1)} \left(\frac{\mu_1}{\lambda_{1_1}} \right)^t + \mu_0$$

where μ_0 is an arbitrarily small real number, since $\pi_{IJ}^{(1_1)}$ and $\phi_{IJ}^{(1)}$ are not zero. For (A.22) to hold for all t , we must have

$$(A.23) \quad \pi_{IJ}^{(1_1)} = \phi_{IJ}^{(1)},$$

$$(A.24) \quad \lambda_{1_1} = \mu_1.$$

It is clear that this argument can be repeated N times, justifying (A.18) and (A.19).