

Mathematical Representations of Development Theories¹

BURTON SINGER
SEYMOUR SPILERMAN

ABSTRACT

In this chapter we explore the consequences of particular stage linkage structures for the evolution of a population. We first argue the importance of constructing "dynamic" models of developmental theories and show through a series of examples the implications of various stage connections for population movements. In discussing dynamic models, one thrust of our comments is to identify the sorts of process features about which assumptions must be made in order to convert a static theory about stage connections (the sort of specification commonly presented in life-span psychology) into a dynamic model. A second focus of our discussion concerns inverse problems: How to utilize a model formulation so that the stage linkage structure may be recovered from survey data of the kind collected by developmental psychologists.

I. Introduction

Although time, usually in the guise of age, is a crucial variable in developmental psychology, formal models of developmental phenomena rarely have the character of dynamic representations, in the sense of mimicking the evolution of an empirical process through time. The analytic procedures employed most extensively by life-span psychologists are factor analysis, regression, analysis of variance, scaling, clustering, and variants of these methods (see, for instance, Nesselroade & Reese, 1973). These are powerful

¹ The work reported here was supported by National Science Foundation Grants SOC76-17706 at Columbia University and SOC76-07698 at University of Wisconsin (Madison). Assistance was also provided by the Institute for Research on Poverty at the University of Wisconsin.

techniques for identifying variables that are central to the course of development in a particular substantive area (e.g., intellectual maturation, acquisition of moral values). Also, when applied to panel data, the procedures can yield insights into how the salience of key variables shifts over the life cycle, or over a portion thereof (e.g., stages in infancy, youth, adulthood).

These analytic methods do not, however, lead to *dynamic* formulations of developmental theories, which can be useful in testing predictions from a theory about the evolution of an empirical process, or in comparing the implications of competing explanations. By a dynamic formulation we mean a representation that incorporates into the mathematics the main assumptions about a developmental phenomenon and is specified in such a way that the relevant variables, and their postulated interrelations, are functions of time or subjects's age. In this sense, like the empirical process, it too constitutes an evolving system. As a simple illustration of such a model, consider the following statements of alternative evolutionary mechanisms:

1. *The growth of a process at each instant is proportional to its potential for future growth.*
2. *The growth of a process at each instant is proportional to the product of its current size and its potential for further growth.*

These statements might be proposed as competing explanations of the manner by which information is diffused in a population of size N . In Formulation (1), it does not matter how many persons $y(t)$ know the information of concern at instant t ; only those yet to hear, numbering $N - y(t)$, are salient to the diffusion rate. If the information were propagated by a mass-media source, such as radio or television, rather than by interpersonal communication, this model might apply. Formulation (2), in comparison, is consistent with a process in which those already aware of the information "infect" the uninitiated through contact and conversation. Assuming that the informed and the uninformed mix randomly, the variable governing the

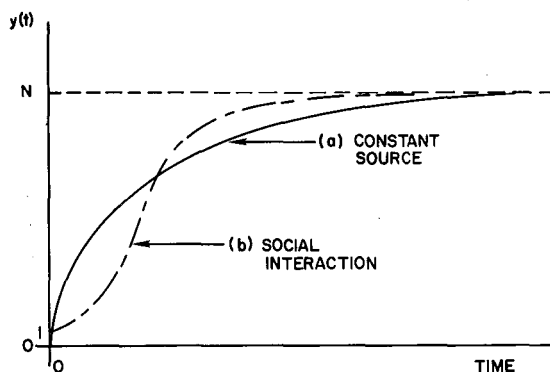


FIGURE 6.1. Illustrative growth curves for diffusion via social interaction and diffusion from a constant source. N = population size; $y(t)$ = number aware of information at time t .

evolution of the process would be $y(t)[N - y(t)]$, which measures the rate at which individuals from the two groups come into contact.

The evolutionary mechanisms, (1) and (2), can be represented by the differential Eq. (1) and (2), respectively,

$$dy(t)/d(t) = k_1[N - y(t)], \quad y(0) = 0 \quad (1)$$

$$dy(t)/d(t) = k_2y(t)[N - y(t)], \quad y(0) = 1 \quad (2)$$

where k_1 and k_2 are constants that adjust for the time unit (e.g., day, year) used in the measurements.² Equations (1) and (2) have for solutions Eq. (3) and Eq. (4),

$$y(t) = N[1 - \exp(-k_1t)] \quad (3)$$

$$y(t) = [N \exp(Nk_2t)]/[N - 1 + \exp(Nk_2t)] \quad (4)$$

which predict the different evolutionary paths displayed in Figure 6.1.

These formulations are *dynamic* in that time appears explicitly as a variable; they are *process models* in that the predicted value of $y(t)$ evolves according to the assumptions of a particular theory. If a researcher has data on the time course of an empirical process, he can test whether Eq. (3), Eq. (4), or a specification of an equivalent sort best approximates his observations. By this exercise it is often possible to select among competing explanations of the mechanism underlying a developmental process. Indeed, these very models have been applied by Coleman, Katz, and Menzel (1957) to data on drug adoptions by physicians (also see Coleman, 1964, pp. 43-45). They concluded that the drug acquisition pattern by socially integrated M.D.'s is best represented by a logistic curve (implying Mechanism 2), whereas isolated M.D.'s adopt according to the constant source model (Mechanism 1), as they are influenced principally by drug advertisements in trade journals. To our knowledge, although developmental psychologists emphasize ontogenetic processes and employ the imagery of an evolutionary system, few attempts have been made to translate their theories into formal models such as the preceding.

In this chapter, we describe the formulation of dynamic models where the objective is to test developmental theories against data or ascertain the consequences of particular assumptions about the structure of a process. To delimit our task, we focus on the sort of mathematics that is appropriate for studying qualitative change. As a result, the tools we introduce are pertinent to theories that postulate stage sequences, a variety of explanation with considerable precedent in developmental psychology (Ausubel & Sullivan, 1970; Kohlberg, 1968; Piaget, 1960). To the degree possible we have written this chapter with a view toward substantive issues and have con-

² The initial condition, $y(0) = 1$, in Eq. (2) is necessary because diffusion through communication cannot begin until at least one person is knowledgeable.

centrated on the translation of theoretical specifications into mathematical formalism; the reader is usually referred elsewhere for mathematical details and estimation procedures. The organization of the chapter is as follows: In the next section we introduce a class of models that is suitable for studying evolutionary processes that incorporate the notion of stage. In section III we describe how particular stage theories can be cast in the framework of the general model. In the fourth section we relax several requirements of the basic model so that it can more realistically represent developmental phenomena.

II. The Concept of Development Stages and a Mathematical Formulation of Stage Progressions

Stage sequences have been postulated for a variety of developmental processes—the evolution of moral behavior (Kohlberg, 1973), cognition (Piaget, 1954), personality (Loevinger, 1966a), and motor skills (Shirley, 1933), to cite but a few topics. There also exist diverse formulations of stage models in the literature of life-span psychology. These differ with respect to the presumed sources of the stages and with regard to the rules governing movement between them. In regard to stage origins, some authors have emphasized maturational considerations, in which individuals are viewed as programmed genetically for particular behaviors or abilities to emerge (Gesell, 1954). The specification of psychosexual stages, keyed to biological activation of the sex glands, provides an illustration (Kohlberg, 1973, p. 181). Others view stages as arising from interactions with the social environment. Kohlberg (1968, pp. 1016–1024), for example, contends that experience with the cultural and physical world is necessary for cognitive stages to take the shapes they do. Still other researchers have adopted the position that stages are a useful research construct around which to discuss development, without insisting that they have an empirical existence (Kaplan, 1966; Reese, 1970).

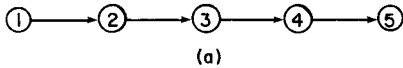
We shall not discuss further the very important issues concerning the etiology of stages, but will focus instead on the mathematical representation of theories about stage connections and on the consequences of various linkage structures for the evolution of individuals among the stages. Formulations of stage connections in a developmental process differ according to whether the progression is viewed as unilineal or multilineal, whether stages in the sequence can be skipped, and whether regression to an earlier level is possible. A second set of considerations pertinent to the structure of developmental theories concerns the age specificity of a stage and the related matter of the variability of duration in a stage. For discussions of

these topics in the context of particular substantive processes, the reader is referred to Emmerich (1968) and Kessen (1962).

To develop the mathematical apparatus for ascertaining the implications of particular stage connections, we discuss both the simplest prototype of a stage theory (for concreteness) and the general mathematical formulation.³ Consider, then, a developmental progression consisting of n stages, in which the linkage is unilineal and there is no possibility of stage skipping or regression. An example of such a structure, with $n = 5$, is presented in panel A of Figure 6.2; henceforth this model is referred to as Example 1. It will be convenient to have available also a matrix representation of the stage linkages. For an arbitrary n -stage structure, we define a matrix M ,

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{bmatrix} \quad (5)$$

whose elements are $m_{ij} = \{\text{probability of transferring from stage } i \text{ to stage } j \text{ when a transition occurs}\}$, where $0 \leq m_{ij} \leq 1$, and $\sum_{j=1}^n m_{ij} = 1$. These restrictions on the elements of M ensure that each row of the matrix constitutes a probability distribution. We require, in addition, that $m_{ii} = 0$ for each stage i that is not an absorbing state of the process; that is, from which individuals can exit. This means we exclude the possibility of within-stage transitions, a type of move that is undefined in most developmental theories. Also, we set $m_{ii} = 1$ for each stage that is an absorbing state of the process.



$$M_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (b)$$

FIGURE 6.2. (a) Representation of a simple unilineal stage structure. (b) Each row of M_1 is a vector of destination probabilities. Thus, if an individual were in Stage 1 before a transition, the Row 1 entries would pertain, and they indicate movement to Stage 2 with probability equal to 1. The main diagonal entries are set equal to 0 (with the exception of Row 5) to indicate that a "move" is not defined apart from a stage transition; that is, there is no notion of movement within a stage. The main diagonal entry of Row 5 is set equal to 1 because this stage is an absorbing state ($m_{5j} = 0$ for $j \neq 5$), and the definition of M_1 (see text) requires $\sum_j m_{5j} = 1$.

³ For a more technical presentation of continuous-time Markov processes see Feller (1968, chapter 17) and Singer and Spilerman (1974). For discussions on the superimposition of theoretical structures on stochastic models see Coleman (1964, chapters 5, 6).

This is done for mathematical convenience and, as we shall see, carries no substantive implications. In the particular case of the unilineal progression (Figure 6.2a), we have the further requirements on \mathbf{M} : $m_{i,i+1} = 1$, and $m_{ij} = 0$ otherwise (except that $m_{55} = 1$). This matrix, \mathbf{M}_1 , is reported in Figure 6.2b.

To this point, though matrix \mathbf{M} conveys important structural information about the process, the description of the stage progression is a static representation. To elaborate the model we must indicate how stage-transition events occur. At a general level of description we assume that the time τ_k spent by an individual in stage i follows some probability distribution,

$$\text{Prob}_i(\tau_k < t | \tau_1, \dots, \tau_{k-1}) \quad (6)$$

where $\tau_1, \dots, \tau_{k-1}$ report the sojourn times in earlier stages. Our imagery, therefore, is the following. An individual originates in stage i at the beginning of the process, $t_0 = 0$. He remains there for an interval τ_1 , specified by a distribution function $\text{Prob}_i(\tau_1 < t)$, and then transfers to stage j with probability m_{ij} . He remains in this stage for a period τ_2 , specified by a conditional probability distribution $\text{Prob}_j(\tau_2 < t | \tau_1)$, then transfers to stage k with probability m_{jk} ; and so forth.⁴ The process continues until some absorbing state is reached, at which point the evolution is terminated. The time path for the unilineal progression associated with the stage linkages of matrix \mathbf{M}_1 is presented in Figure 6.3.

Several further assumptions are necessary to complete the specification of the model. One matter concerns the relevance of an individual's past

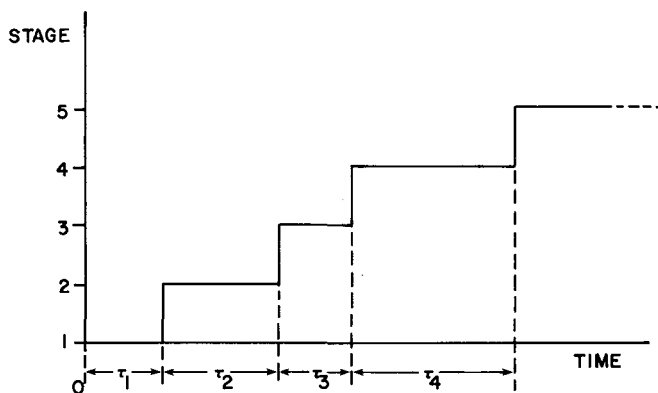


FIGURE 6.3. A sample path description corresponding to the unilineal stage structure of Figure 6.2. It is assumed that there are five stages, which must be traversed sequentially. τ_i is the value of a random variable and denotes the sojourn time for an individual in stage i . Stage 5 is an absorbing state of the process.

⁴ In the present example $i, j, k = 1, 2, 3$, respectively.

movement history to the course of his subsequent evolution among the stages. We assume

1. *Knowledge of current stage conveys all information that is relevant to forecasting future movements.*

Stated technically, if $m_{ij,ab\dots f} = \{\text{probability of moving from stage } i \text{ to stage } j \text{ at the occurrence of a transition, given prior sojourns in stages } a, b, \dots, f\}$, then

$$m_{ij,ab\dots f} = m_{ij}.$$

(This assumption is superfluous in the current example of a unilineal progression since there is only one possible path, but it is relevant to the evolution of a population in less restrictive models.) We indicate in the next section that this specification has been employed in descriptions of stage linkages in developmental psychology.

For an initial baseline class of models, we further assume

2. *The sojourn time in stage i is exponentially distributed; that is,*

$$\begin{aligned} \text{Prob}_i(\tau_k < t | \tau_1, \dots, \tau_{k-1}) &= \text{Prob}_i(\tau_k < t) \\ &= F_i(t) = 1 - e^{-\lambda_i t}. \end{aligned} \quad (7)$$

Use of the exponential distribution amounts to specifying that the probability of departing from stage i during the infinitesimal interval $(t, t + dt)$, conditional on being in stage i at time t , equals

$$\frac{f_i(t) dt}{1 - F_i(t)} = \frac{\lambda_i \exp(-\lambda_i t) dt}{1 - [1 - \exp(-\lambda_i t)]} = \lambda_i dt$$

where $f_i(t)$ is the density function corresponding to $F_i(t)$. This result, in turn, indicates that the probability of leaving stage i is independent of duration in the stage, and is tantamount to specifying an absence of aging. So new entrants have the same likelihood of departing as individuals who have been in the stage for some period of time. The parameter λ_i , incidentally, has an interpretation as the rate of movement out of stage i ; consequently, $1/\lambda_i$ equals the expected duration in stage i .

3. *Finally, we require that if the data pertain to the movements of a population, rather than to the transitions of a single individual, the population is homogeneous with respect to the structure of the evolutionary process.*

This does not mean that all persons have the same duration τ_i in stage i , but that τ_{ic} , the time spent in stage i by individual c , follows the single exponential distribution $F_i(t) = 1 - \exp(-\lambda_i t)$. Stated less formally, duration in a stage is a random variable with the underlying distribution of holding times the same for all individuals. Similarly, where alternative

destinations are available to persons in stage i , homogeneity means that all have the same list of probabilities for making the various transitions, not that they move identically.

It is worth dwelling on the conceptual status of the preceding assumptions. The question of the structure of \mathbf{M} is a familiar topic to developmental psychologists, since stage theories are commonly specified at this level. Assumptions (1)–(3) can be viewed as “side conditions,” aspects of the process to which researchers have generally not been sensitive, though see Kessen (1960) and Emmerich (1968) for provocative comments on precisely these matters. What is made evident by formulating a dynamic model is that development theorists must address these auxiliary questions if complete models are to be specified. The particular assumptions we have made constitute a gross simplification of reality; this is especially true of Specification (2), which postulates an absence of duration effects, and Specification (3), which postulates population homogeneity. These assumptions do, however, provide a convenient starting point from which to consider more realistic formulations, which are developed in the next sections.

We now wish to convey the implications of Assumptions (1)–(3) for the movements of individuals among the stages. We denote by $p_{ij}(t)$ the probability that an individual in stage i at time 0 moves to stage j by time t . (This probability differs from m_{ij} in that the latter refers to movement proclivities at the occurrence of a transition, not over widely spaced time intervals.) With this specification in hand, the evolution of a population among the stages is described by the system of integral equations,

$$p_{ij}(t) = \delta_{ij} \exp(-\lambda_i t) + \sum_k \int_0^t \lambda_i \exp(-\lambda_i u) m_{ik} p_{kj}(t-u) du \quad (8)$$

$$0 \leq i, j \leq n$$

where $\delta_{ij} = 1$ if $i = j$, and 0 otherwise. This expression, known as the backward equations for a continuous-time Markov process (Feller, 1971, p. 484), is amenable to the following interpretation:

1. When $i \neq j$, $p_{ij}(t)$ consists of the sum of products of three factors: the probability of a first departure from stage i at time u , the probability of a stage i to stage k transition at that instant, and the probability of transferring to stage j by some combination of moves in the interval $t - u$. The summation is over all intermediate stages k and over all time divisions u in the interval $(0, t)$.

2. When $i = j$, in addition to the above term, there is the possibility of not transferring out of stage i during $(0, t)$. This probability is given by the first term.

If we represent by $\mathbf{P}(t)$ the matrix of elements $p_{ij}(t)$.

$$\mathbf{P}(t) = \begin{bmatrix} p_{11}(t) & \cdots & p_{1n}(t) \\ \vdots & & \vdots \\ p_{n1}(t) & \cdots & p_{nn}(t) \end{bmatrix}$$

$0 \leq p_{ij}(t) \leq 1$, $\sum_j p_{ij}(t) = 1$, then the integral equations (Eq. 8) have the convenient solution,

$$\mathbf{P}(t) = e^{\Lambda[\mathbf{M}-\mathbf{I}]t}, \quad \mathbf{P}(0) = \mathbf{I}. \quad (9)$$

In this representation Λ is a diagonal matrix,

$$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

whose entries are the reciprocals of the expected duration times in each stage, \mathbf{I} is the identity matrix, and \mathbf{M} is the array specified in Eq. (5), which describes the pattern of movement between the stages. Furthermore, by the expression e^{Λ} , Λ an arbitrary square matrix, we mean the power series in Λ ,

$$e^{\Lambda} = \sum_{n=0}^{\infty} \frac{\Lambda^n}{n!} \quad (10)$$

which can be evaluated by standard numerical methods (see, e.g., Gantmacher, 1960).

It is useful to recapitulate what is accomplished by this mathematical formulation. The matrix $\mathbf{P}(t)$ relates the distribution of a population among stages at time t to its distribution at time 0, in the sense that a typical entry, $p_{ij}(t)$, conveys the probability of moving from stage i to stage j during the interval $(0, t)$. The model is dynamic in that $\mathbf{P}(t)$ is a function of time; with the passage of time, $\mathbf{P}(t)$ describes the evolution of the population among the stages. Equation (9) shows how the matrix $\mathbf{P}(t)$ is built up from the arrays \mathbf{M} and Λ . However, while this equation is useful as a calculating formula, the logic of the process is conveyed more adequately by the integral equation (8).

To illustrate this model in the setting of a simple unilineal progression (matrix \mathbf{M}_1 of Figure 6.2), we must specify average waiting times in Stages 1, 2, 3, and 4. We assume these to be .5, 1, 2, and 5 years, respectively.

Consequently, we have for matrix Λ ,

$$\Lambda = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & .5 & 0 & 0 \\ 0 & 0 & 0 & .2 & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 \end{bmatrix} \quad (11)$$

where the choice of λ_5 is arbitrary. Since Stage 5 is an absorbing state, the notion of waiting time to a departure has no meaning. (Mathematically, $[\mathbf{M} - \mathbf{I}]_{55} = [m_{55} - 1] = [1 - 1] = 0$, so λ_5 bears no influence on the calculations.) Now, from \mathbf{M}_1 , Λ , and \mathbf{I} , we have

$$\Lambda(\mathbf{M}_1 - \mathbf{I}) = \begin{bmatrix} -2 & 2 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -.5 & .5 & 0 \\ 0 & 0 & 0 & -.2 & .2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (12)$$

For the illustrative times $t = 1, 2$, and 4 years, we obtain, from Eq. (9) for $\mathbf{P}(t)$,

$$\mathbf{P}(1) = \begin{bmatrix} .1353 & .4651 & .3263 & .0691 & .0041 \\ .0000 & .3679 & .4773 & .1438 & .0110 \\ .0000 & .0000 & .6065 & .3537 & .0398 \\ .0000 & .0000 & .0000 & .8187 & .1813 \\ .0000 & .0000 & .0000 & .0000 & 1.0000 \end{bmatrix}, \quad (13)$$

$$\mathbf{P}(2) = \begin{bmatrix} .0183 & .2340 & .4641 & .2482 & .0354 \\ .0000 & .1353 & .4651 & .3394 & .0602 \\ .0000 & .0000 & .3679 & .5041 & .1281 \\ .0000 & .0000 & .0000 & .6703 & .3297 \\ .0000 & .0000 & .0000 & .0000 & 1.0000 \end{bmatrix}, \quad (14)$$

and

$$\mathbf{P}(4) = \begin{bmatrix} .0003 & .0360 & .2881 & .4843 & .1913 \\ .0000 & .0183 & .2340 & .5079 & .2398 \\ .0000 & .0000 & .1353 & .5233 & .3413 \\ .0000 & .0000 & .0000 & .4493 & .5507 \\ .0000 & .0000 & .0000 & .0000 & 1.0000 \end{bmatrix}. \quad (15)$$

These values of $\mathbf{P}(t)$ describe the evolution of individuals among the stages, subject to the assumptions about the process structure detailed above. The entries $p_{ij}(t)$ refer to proportions⁵ of the population who have

⁵ If the observations are on a single individual the interpretation of $p_{ij}(t)$ is in terms of the probability of a stage i to stage j move between times 0 and t .

moved between particular stages in the relevant time interval. For example, according to the entries in the top row of $P(1)$, if observations are taken 1 year apart, we would expect 13% of the population in Stage 1 at time 0 to still be there, 46% to have moved to Stage 2, and 33% to have reached Stage 3. By comparison, over a 4-year interval, less than 1% would remain in Stage 1, 48% would have reached Stage 4, and 19% would be in the terminal stage of the process.

The results from the three calculations reveal that, even though the progression is unilinear with all individuals characterized by the same parameters, if observations were taken on the population at two time points, $t = 0$ and $t = t_1$, the array⁶ $\hat{P}(t_1)$ might be interpreted as evidence for a more complex theory, such as one permitting stage skipping or population heterogeneity in the rate or pattern of movement. Furthermore, the correspondence between the matrix constructed from the population locations at two time points, $\hat{P}(t_1)$, and the rule governing stage transitions, M_1 , decreases with time. Thus, different researchers observing the same population at two time points but with different spacing intervals might draw contrary conclusions about the stage-linkage structure even though the single mechanism, M_1 of Figure 6.2 governs its evolution. Only with a formal model of the process could one hope to uncover its underlying structure.

III. Models of More Elaborate Stage Theories

The matrix M contains structural information about stage linkages. Since theories of development are commonly posed at the level of specifying this array, flexibility in incorporating a variety of particular formulations would appear to be an important feature of a general framework for describing evolutionary behavior. In this section we focus on the issue of translating stage theories into M -matrices, and illustrate the evolution of $P(t)$, the transition matrix for a population based on its locations at times 0 and t , under alternative specifications of M . As we have noted, auxiliary information about the process, concerning the distribution of waiting-time intervals and the form of population heterogeneity, is required for a full description of a dynamic model. In the next section we therefore elaborate upon these "side conditions" and outline ways in which our initial assumptions can be relaxed.

No technical difficulties arise in reformulating the continuous-time Markov model to accommodate more elaborate theories of stage linkages than

⁶ A caret over a matrix or over an element in a matrix will mean that it should be viewed as estimated directly from data rather than calculated from a mathematical model.

the structure in Figure 6.2. We illustrate the procedure with a few examples.⁷

A. Unilineal Progressions

1. A Unilineal Progression that Permits Stage Skipping

The formulation of such a structure is diagrammed in Figure 6.4a its translation into an M -matrix is reported in Figure 6.4b. The principal new feature is that, supplementing the deterministic sequence of Figure 6.2, it is now possible to move directly from Stage 2 to Stage 4 and from Stage 3 to Stage 5, when transition out of the relevant origin location takes place. We must also specify the probabilities of following the alternate paths. In the present example, lacking information as to the relative magnitudes of the various probabilities, we assume all destinations to be equally likely; that is, we prescribe $m_{23} = m_{24} = .5$, and $m_{34} = m_{35} = .5$. In practice, estimates of the transition probabilities would be assigned on the basis of theory or from observation on the empirical process.

Using matrix M_2 , together with the A array of Eq. (11), whose entries describe the rate of movement by individuals out of each stage, we obtain for $P(1)$ and $P(4)$, from Eq. (9):

$$P(1) = \begin{bmatrix} .1353 & .4651 & .1632 & .2012 & .0352 \\ .0000 & .3679 & .2387 & .3177 & .0757 \\ .0000 & .0000 & .6065 & .1768 & .2166 \\ .0000 & .0000 & .0000 & .8187 & .1813 \\ .0000 & .0000 & .0000 & .0000 & 1.0000 \end{bmatrix}, \quad (16)$$

$$P(4) = \begin{bmatrix} .0003 & .0360 & .1440 & .4104 & .4093 \\ .0000 & .0183 & .1170 & .3964 & .4683 \\ .0000 & .0000 & .1353 & .2617 & .6030 \\ .0000 & .0000 & .0000 & .4493 & .5507 \\ .0000 & .0000 & .0000 & .0000 & 1.0000 \end{bmatrix}. \quad (17)$$

These $P(t)$ arrays are the transition matrices a researcher should expect to observe if the stage locations of individuals are surveyed 1 year or 4 years apart, assuming that the population evolves according to the linkage specification M_2 together with the auxiliary conditions outlined in the preceding section. The entries are different from those obtained with the simple unilineal progression [Eq. (13) and (15)], yet the same pattern of zeros and nonzeros is present. Without a formal model of the evolution of the process, a researcher would be unable to predict the different implications of these structures.

⁷ We begin here with Example 2; Example 1 refers to the structure in Figure 6.2.

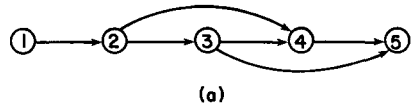


FIGURE 6.4. (a and b) **Representation of a unilineal progression in which stage skipping is permitted. All destination stages corresponding to an origin location are assumed to occur with equal probability. See Figure 6.2 legend for details on interpretation of M_2 .**

$$M_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & .5 & .5 & 0 \\ 0 & 0 & 0 & .5 & .5 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(b)

2. A Unilineal Progression with Stage Skipping and the Possibility of Regression

We now superimpose on the linkage structure the possibility of reverting to an earlier stage. This arrangement is diagrammed in Figure 6.5a, in which we have provided for the possibility of backward flows from Stage 2 to Stage 1, from Stage 3 to Stage 2, and from Stage 5 to Stage 4. The M-matrix corresponding to this model is reported in Figure 6.5b. Again, where multiple destinations correspond to an origin stage, we have arbitrarily assigned equal values to the m_{ij} 's. There is one additional alteration in M_3 , in comparison with the M-matrices of earlier examples. Because there now exists a possibility of regressing from the terminal stage to an earlier level, $m_{55} \neq 1$. To maintain our conceptual imagery, in which within-stage transitions are undefined, we set $m_{54} = 1$ and $m_{55} = 0$. Note that the former value does not imply a high rate of departure from Stage 5, since the rate of movement is controlled by λ_5 . It only means that all transitions from Stage 5 are directed to Stage 4.

To obtain $P(t)$ we use M_3 and Λ in conjunction with Eq. (9). Here the element λ_5 in Eq. (11) is no longer arbitrary, as movement out of Stage 5 is a possibility. We shall assume that such reversions are rare, and hence

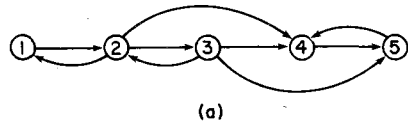


FIGURE 6.5. (a and b) **Representation of a unilineal progression in which stage skipping and regression to an earlier level are permitted. All destination stages corresponding to an origin location are assumed to occur with equal probability. See Figure 6.2 legend for additional details on interpretation.**

$$M_3 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ .33 & 0 & .34 & .33 & 0 \\ 0 & .33 & 0 & .34 & .33 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

(b)

specify the average waiting time to a transition from Stage 5 to be 8 years; that is, $\lambda_5 = .125$. With these assumptions, we obtain for our illustrative calculations at $t = 1, 4$:

$$P(1) = \begin{bmatrix} .2043 & .5240 & .1153 & .1374 & .0190 \\ .0871 & .4758 & .1742 & .2217 & .0411 \\ .0094 & .0858 & .6215 & .1461 & .1371 \\ .0000 & .0000 & .0000 & .8292 & .1708 \\ .0000 & .0000 & .0000 & .1067 & .8933 \end{bmatrix}, \quad (18)$$

$$P(4) = \begin{bmatrix} .0330 & .1560 & .1652 & .4025 & .2433 \\ .0259 & .1246 & .1500 & .4174 & .2820 \\ .0135 & .0739 & .1846 & .3542 & .3738 \\ .0000 & .0000 & .0000 & .5523 & .4477 \\ .0000 & .0000 & .0000 & .2798 & .7202 \end{bmatrix}. \quad (19)$$

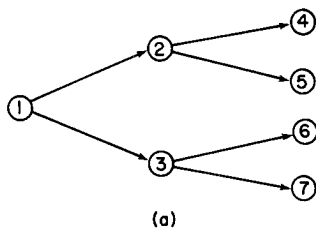
If we compare the $P(1)$ matrices and the $P(4)$ matrices from the three examples [i.e., Eqs. (13), (16), (18), and (15), (17), (19)], we can acquire a fair idea of the implications of different stage interconnections for the evolution of a population among the statuses. We also emphasize the fact that if a population is surveyed at two time points, especially widely spaced time points, the structure of the stage linkages (matrix M) that generated the observations may not be obvious from inspecting the empirically determined transition array, $\hat{P}(t_1)$. We will return to the issue of identifying the correct structure and recovering matrix M when the observations on a process are widely spaced; we conclude this discussion on translating theoretical specifications of stage linkages into M -matrices with two examples of multilineal sequences that have been described in the developmental psychology literature.

3. A Divergent Multiple Progression

This stage linkage structure (see Van den Daele, 1969, Figures 2, 4) has the diagrammatic representation of Figure 6.6a; its corresponding M -matrix is presented in Figure 6.6b. Because Stages 4-7 are specified to be terminal states of the process, the corresponding rows of M_4 have 1's in the main diagonal. Van den Daele provides no discussion of waiting time distributions to departure from the various stages; hence the model remains incomplete as an evolutionary process.

4. A Convergent Multiple Progression

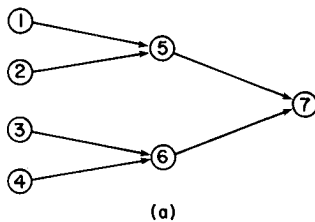
This stage sequence (see Van den Daele, 1969, Figure 2) is depicted in Figure 6.7a, and its associated M -matrix is reported in Figure 6.7b. In this instance, the structure consists of a collection of deterministic unilineal progressions, the specific sequence for an individual being contingent upon his entry stage. Note also that the assumption of irrelevance of past history,



$$M_4 = \begin{bmatrix} 0 & .5 & .5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .5 & .5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .5 & .5 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

FIGURE 6.6. (a and b) Representation of a divergent multiple progression. [Source: Van den Daele (1969, Figures 2,3).]

which is posited in this formulation, is one of the side conditions we have required (Assumption 1 in the preceding section). In particular, this specification appears in the fact that knowledge of the path by which one has reached Stage 5 (or Stage 6) is of no value in forecasting, or understanding, an individual's subsequent movements. Van den Daele (1969) discusses several additional models of stage linkages, such as "partially convergent, divergent progression," and "partially divergent, convergent progression." Since the procedure used in converting flow structures into M-matrices should be evident at this point, discussions of these specifications are not presented.



$$M_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

FIGURE 6.7. (a and b) Representation of a convergent multiple progression. [Source: Van den Daele (1969, Figure 2).]

To recapitulate, we have shown that, subject to several side conditions, it is possible to construct formulations of a range of developmental phenomena that mimic the evolutionary character of the observed process. With such a model one can forecast the movements of a population among the stages. By carrying out the requisite calculations for different specifications of the stage linkages and comparing the predictions, it is possible to ascertain the ways in which rather complex theories produce divergent implications and to design testing schemes that maximize the possibility of rejecting one or another formulation as a description of the empirical process. Of equal importance, it is often possible to work backward, starting with observations on the stage locations of a population at a few widely spaced time points, and derive the structure of the stage linkages compatible with the data.

B. An Inverse Problem

Until this point we have assumed that observations have been made on an empirical process in a way such that M and Λ can be estimated directly from the data, or that theories are available that specify the values of their entries. We then sought to derive the evolution of the process subject to the presumed structure. In developmental psychology, it is not uncommon for a researcher to have many observations on a few individuals (e.g., Piaget, 1954). Such a data-collection scheme approximates *sample path information*, a complete history of movements and waiting times of the sort illustrated in Figure 6.3. Detailed observation of a few subjects is a research strategy not without its costs, however. One learns little about the frequency of rare events (e.g., regression to an earlier stage, stage skipping, rare development paths) and acquires only the most rudimentary knowledge about the variation of duration times in a stage. It is therefore not surprising that investigators who rely on this approach tend to be oriented to uncovering universal rules (e.g., Piaget, 1960) rather than to elucidating individual differences and ascertaining the variety of developmental patterns.

Partly because of the limitations of small data sets, it is becoming increasingly common to employ survey methods in which a large population, sometimes thousands of individuals, is observed (or interrogated) at a very few time points (e.g., Baltes & Nesselrode, 1972). The spacing intervals in such panel studies are usually wide—often 1 or more years elapse between interviews—so it is not unusual for some subjects to have made multiple moves while others have made one or zero shifts between stages. The transition matrices that can be constructed directly from such observations are $P(t)$ arrays, rather than M arrays, and the stage linkages may not be readily discernible. Indeed, determination of the movement structure that underlies the evolution of the population can be a difficult task.

One approach to ascertaining the stage linkages from survey data involves

consideration of the *inverse problem* to the mathematical formulation of the evolutionary model [Eq. (9)]. Stated formally, we have available the matrix $\hat{P}(t_1)$, constructed from observations on the stage locations of individuals at time 0 and t_1 . The typical entry in this matrix is $\hat{p}_{ij}(t_1) = n_{ij}(t_1)/n_i$, where $n_i = \{\text{number of individuals in stage } i \text{ at time } 0\}$ and $n_{ij}(t_1) = \{\text{number of persons who started in stage } i \text{ at time } 0 \text{ and are in stage } j \text{ at time } t_1\}$. We wish to inquire whether it is possible to recover a unique M -matrix for the process and, where the answer is affirmative, we wish to estimate this matrix.

The first step in solving the inverse problem is to take the logarithm of both sides of Eq. (9).

$$Q = \Lambda[M - I] = t_1^{-1} \ln \hat{P}(t_1). \quad (20)$$

Just what we mean by the logarithm of matrix $\hat{P}(t_1)$, the conditions under which a solution to Eq. (20) will exist, and the circumstances under which the solution will be unique are complex issues that are discussed at length in Singer and Spilerman (1976). Assuming we can obtain a valid and unique Q -matrix from these calculations, a second task, separating M from Λ , still remains. In many instances, though, this matter is of little concern, since the pattern of zeros and nonzeros in Q and $M - I$ will be identical, and development theories are often posed at the level of identifying permissible transitions. Moreover, because zeros are typically present in many main diagonal cells of M in models of developmental structures, a complete or near complete separation between M and Λ can frequently be effected.

We conclude this section with an example of the calculations associated with the inverse problem. Suppose observations taken on a population at times 0 and t_1 have produced the transition matrix.

$$\hat{P}(t_1) = \begin{bmatrix} .0224 & .2633 & .2402 & .1262 & .3479 \\ .0064 & .1758 & .2460 & .1735 & .3983 \\ .0216 & .0288 & .3758 & .5060 & .0678 \\ .0365 & .0744 & .0288 & .6794 & .1809 \\ .0005 & .0960 & .0460 & .0178 & .8397 \end{bmatrix} \quad (21)$$

Such data would appear to be consistent with a variety of evolutionary mechanisms. From inspection of $\hat{P}(t_1)$ we do know that regression to some earlier stage must be possible; otherwise all entries below the main diagonal would be zero. Little else about the structure of M , however, can be inferred from inspection of $\hat{P}(t_1)$. Indeed, because of the sizable nonzero elements in most cells of the matrix, a researcher might conclude that direct transitions are possible between most pairs of stages.

If we are willing to assume that matrix $\hat{P}(t_1)$ was generated by a continuous-time Markov process—that is, via the evolution of the structure $P(t) = e^{\Lambda(M-I)t}$, for some matrices Λ and M , that satisfy the definitional restrictions enumerated in connection with Eqs. (5) and (9)—we can solve

for $\Lambda(\mathbf{M} - \mathbf{I})t_1$ using Eq. (20). This yields the array,

$$\Lambda(\mathbf{M} - \mathbf{I})t_1 = \begin{bmatrix} -4 & 4 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 \\ .20 & 0 & 0 & -.40 & .20 \\ 0 & .25 & 0 & 0 & -.25 \end{bmatrix}. \quad (22)$$

In this instance Λt_1 and \mathbf{M} can be separated by employing the following argument. From our earlier examples we know that a main diagonal element m_{ii} of \mathbf{M} will equal zero if any off-diagonal entry in the same row, m_{ij} , is different from zero. According to Eq. (22), each row of matrix \mathbf{M} must have at least one nonzero off-diagonal element; therefore, $m_{ij} = 0$ for all values of i . With this information we can obtain Λt_1 uniquely,

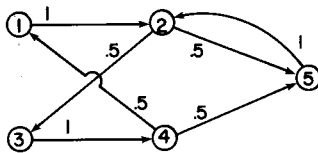
$$\Lambda t_1 = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & .4 & 0 \\ 0 & 0 & 0 & 0 & .25 \end{bmatrix} \quad (23)$$

and solving for \mathbf{M} provides the structure \mathbf{M}_4 reported in Figure 6.8a. The schematic representation of the stage linkages implied by \mathbf{M}_4 is shown in Figure 6.8b, in which probabilities of the various moves have been appended to the paths.

The point to be emphasized is that it is not apparent from inspecting matrix $\hat{\mathbf{P}}(t_1)$ in Eq. (21) that the underlying stage linkages are those reported in Figure 6.8, nor would any static analytic procedure be likely to lead a researcher to the correct conclusion. What is necessary is to construct a model of the evolution of the process and solve the implied inverse problem for the parameters that correspond to the particular data set. (In the present example, we have assumed that the underlying model is a continuous-time

$$\mathbf{M}_4 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & .5 & 0 & .5 \\ 0 & 0 & 0 & 1 & 0 \\ .5 & 0 & 0 & 0 & .5 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

(a)



(b)

FIGURE 6.8. (a and b) Stage sequence structure implied by $\hat{\mathbf{P}}(t_1)$ in Eq.(21). The process is assumed to evolve according to a continuous-time Markov formulation. Entries indicate the probability of a stage i to stage j move when transition takes place. Probabilities of the various transitions are attached to the appropriate paths.

Markov process (i.e., Specifications 1-3 of the preceding section) and have solved for the matrices Λt_1 and M that are compatible with the observed array $\hat{P}(t_1)$, in that they would have given rise to this array if the postulated evolutionary process were approximately correct.

IV. Alternative Specifications of the Side Conditions

In this section we discuss relaxing two of the more burdensome specifications of the model, in the sense that they are likely to be inappropriate as characterizations of developmental processes. We first consider the requirement that the duration intervals in a stage must follow an exponential distribution (Assumption 2 of the second section). Following these comments we turn to the requirement that the population be homogeneous with respect to the process parameters Λ and M (Assumption 3).

A. More General Waiting Times than Exponential

The exponential distribution is frequently employed in the literature of reliability theory to describe duration intervals in a system state (stage in the current application). It has the advantages of being mathematically tractable and approximating reality in situations where the probability of a state change is uninfluenced by aging or time in the state. For example, if the process states are "alive" and "not alive," then over the middle age ranges of many animal species, the age-specific mortality rate is relatively constant and the duration intervals (in the "alive" state) are reasonably well captured by the exponential distribution. Similarly, when mortality results from exogenous events—accidents—the distribution of ages at failure can often be approximated by the exponential.

In a great many situations in social research, however, we know that proneness to changing state is a function of duration. In particular, this has been suggested with respect to residence location (McGinnis, 1968) and employment affiliation (Ginsberg, 1971). In these applications it has been argued that the duration-specific departure rate decreases with time, giving rise to the phenomenon of *cumulative inertia*—the longer an individual remains in a state the less likely he is to leave in the immediate future. The substantive explanations for a declining departure rate involve the growing investment an individual has made, with duration, in friendships (in the first instance) and in seniority in his place of work (in the second). There is no mathematical reason, however, to assume a declining departure rate in choosing $F_i(t)$; and in other substantive contexts a different specification may be more appropriate. For a superb review of stochastic models incorporating the notion of duration dependence, see Hoem (1972).

A convenient way to generalize the Markov model to accommodate a variety of duration-time distributions is to begin with the integral equation representation for transition probabilities. Equation (8) is a special case of the formulation,

$$p_{ij}(t) = \delta_{ij}[1 - F_i(t)] + \sum_k \int_0^t f_i(u) m_{ik} p_{kj}(t-u) du \quad 0 \leq i, j \leq n \quad (24)$$

in which the terms are identical with those of the earlier equation except that $f_i(u)$ replaces the exponential density $\lambda_i \exp(-\lambda_i u)$, and $F_i(t)$ [the distribution function corresponding to $f_i(t)$], replaces $[1 - \exp(-\lambda_i t)]$. A theoretically appropriate choice may now be made for $F_i(t)$.

As an illustration, one candidate for $F_i(t)$, in the case of a declining departure rate, is the two-parameter family of functions

$$F_i(t) = 1 - \exp(-\lambda_i t^{\gamma_i}) \quad \lambda_i > 0, \quad 0 < \gamma_i < 1. \quad (25)$$

Here the probability of departing from state i during the infinitesimal interval $(t, t + dt)$, conditional on the process being in state i at time t , equals

$$\frac{f_i(t) dt}{1 - F_i(t)} = \frac{(\lambda_i \gamma_i t^{\gamma_i-1}) \exp(-\lambda_i t^{\gamma_i})}{\exp(-\lambda_i t^{\gamma_i})} dt = \lambda_i \gamma_i t^{\gamma_i-1} dt.$$

Because of the restriction on γ_i in Eq. (25), t^{γ_i-1} is a decreasing function of time, and the declining failure-rate aspect of the distribution is evident.

The general formulation in Eq. (24), built up from duration-time distributions and transition probabilities between states, generates a class of models known as semi-Markov processes. These generally do not have simple representations for the matrices $P(t)$, analogous to Eq. (9), and the solution of the system of Eq. (24) requires numerical integration methods.

B. Population Heterogeneity

To this point we have assumed that the matrices Λ and M of Eq. (9) are identical for all individuals. This does not mean that all persons move identically since the process is probabilistic; it does imply, though, that individual-level characteristics are unrelated to the structural parameters of the process. In other words, homogeneity means that considerations of genetic makeup, intelligence, sensory stimulation, and other factors by which individuals differ from one another do not portend distinct evolutionary paths in the developmental process under consideration.

There is reason to believe, however, that individual differences are present in the course of development in many processes (Kohlberg, 1968, p. 1024; Werner, 1957). We therefore desire a formulation in which the movement pattern is parameterized in terms of variables that differentiate among

persons. To construct a general specification of heterogeneity within the conceptual framework of a Markov process, we assume that, corresponding to Eq. (9), the stage transitions by individual c have the structure

$$P_c(t) = \exp[\Lambda_c(M_c - I)t]. \quad (26)$$

This formula indicates that each person is characterized by a pair of matrices, Λ_c and M_c , and his evolution, in turn, is described by $P_c(t)$. Thus, our formulation begins with a separate Markov process for each individual.

This approach directs a researcher to identify the variables that describe heterogeneity; that is, to ascertain which factors account for individual differences in the matrices M and Λ . Thus, not only does a heterogeneity formulation lead to more realistic models of evolutionary processes, in that allowance is made for individual differences, but it stresses the analytic tasks of specifying the variety of developmental patterns in a population and ascertaining which attributes make an individual more prone to following one set of paths rather than another.

One form of heterogeneity involves the distribution of M -matrices in a population. Focusing on these arrays serves to emphasize individual differences in proneness to making particular moves when a transition takes place. We shall not discuss this form of heterogeneity in the present chapter and direct the interested reader instead to McFarland (1970), Spilerman (1972a), and Singer and Spilerman (1974). A second form of heterogeneity stresses individual differences in the Λ -matrix; that is, in the rates at which departures occur for persons in the various states. We conclude this section with a simple formulation of population heterogeneity in which it is assumed that the individual differences can be expressed in the latter way.

To simplify the discussion, we further require the nonzero entries in the diagonal matrix Λ to be equal for an individual; that is, $\lambda_i = \lambda$ for all i . This means we are specifying identical departure rates from all states. As a result, Eq. (9) reduces to

$$P(t|\lambda) = e^{\lambda t(M-I)} \quad (27)$$

where $P(t|\lambda)$ denotes the transition matrix for an individual having a rate of movement value equal to λ . We shall assume that Eq. (27) describes the evolution of an individual drawn at random from the population.

Heterogeneity is incorporated into the formulation by specifying a density function $g(\lambda)$ that describes the distribution of λ -values in the population. We now define the population-level transition matrix corresponding to times 0 and t to be

$$P(t) = \int_0^{\infty} P(t|\lambda) g(\lambda) d\lambda = \int_0^{\infty} e^{t\lambda(M-I)} g(\lambda) d\lambda. \quad (28)$$

This formula expresses the population-level matrix as a weighted average

of the individual-level arrays, $\mathbf{P}(t|\lambda)$, the weights reflecting the population proportions associated with particular λ -values.

To complete this specification of heterogeneity it is necessary to select a density function $g(\lambda)$ to describe the distribution of λ -values. One useful choice is the gamma family of functions

$$g(\lambda) = \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} \quad \lambda > 0, \quad \alpha > 0, \quad \beta > 0 \quad (29)$$

which is flexible enough to describe a variety of unimodal curves. With this selection of $g(\lambda)$, a convenient representation of the population-level matrix $\mathbf{P}(t)$ is obtained (Spilerman, 1972b, p. 608):

$$\mathbf{P}(t) = \left(\frac{\beta}{\beta + t} \right)^\alpha \left[\mathbf{I} - \frac{t}{\beta + t} \mathbf{M} \right]^{-\alpha}. \quad (30)$$

The transition probabilities [Eq. (30)] do not describe the evolution of a Markov process; however, they do describe the movement of a population in which each individual follows a Markov process, with individual differences being specified by $g(\lambda)$ in Eq. (29).

In analogy with our earlier discussion of the inverse problem for Markov chains, the present formulation can be used with observations taken at widely spaced time points, 0 and t_1 , together with estimates of α and β to yield an estimate of the underlying transition mechanism \mathbf{M} , according to the matrix equation

$$\mathbf{M} = \left(\frac{\beta + t_1}{t_1} \right) \left(\mathbf{I} - \frac{\beta}{\beta + t_1} [\hat{\mathbf{P}}(t_1)]^{-1/\alpha} \right). \quad (31)$$

Thus, from observations of the sort collected in many surveys, even under an assumption of population heterogeneity in the rate of movement, it may be possible to recover the matrix of stage linkages that governs the evolution of the process.

V. Conclusions

In this chapter we have explored the consequences of particular stage-linkage structures for the evolution of a population. One thrust of our comments has been to identify the sorts of process features concerning which assumptions must be made in order to convert a static theory about stage connections into a dynamic model. A second focus in our discussion has been on inverse problems; how to utilize a model formulation so that the stage-linkage structure (matrix \mathbf{M}) may be recovered from survey data of the kind usually collected by developmental psychologists.

We have presented only the most rudimentary sorts of stage structures.

Indeed, even within the Markov framework we have limited our consideration to a subset of these models; namely, those that are time-stationary (i.e., A and M are not functions of time). By this specification we have excluded the possibility of accommodating age-dependent transition laws, a consideration of substantial importance in developmental psychology. (An extension of the models discussed here to incorporate both age dependence and cohort effects is, however, a feasible undertaking but one with an increase in mathematical complexity.) Furthermore, all the models we have discussed entail a low dependence of future movements on the transition history of an individual, given his current stage.⁸ Restrictions of these sorts are likely to be reasonable for some processes, unreasonable for others. Appropriate models of developmental phenomena must, therefore, be constructed from a list of known characteristics about an empirical process.

We also point out that the concept of stage merges with the notion of state as the number and sorts of permissible transitions increase. "Stage" seems conceptually rooted to the idea of progress (i.e., development) and would be an appropriate component of a theory that sees the system's statuses as genetically determined or as facilitating the conditions for succeeding statuses to come into play.⁹ The mathematical framework we have introduced is also compatible with a notion of state, in which there is extensive opportunity to cycle among the statuses. State formulations have been suggested in the psychology literature in relation to anxiety, moods, etc. (e.g., Kessen, 1962, pp. 72-73).

As a final set of considerations in relation to the structure of stage models, we note that all the formulations we have addressed are models of solitary processes. We have proceeded as if intelligence, cognition, motor skills, and personality development unfolded autonomously. In reality there no doubt exist extensive dependencies among some of these processes. Mathematical models of interacting developmental phenomena could be formulated, but clear empirically based specifications of such dependencies are still lacking.

⁸ The time-stationary Markov formulations postulate irrelevance of prior stage affiliations, durations in those stages, and duration in current stage. The last two of these restrictions can be eliminated by introducing nonstationary semi-Markov models as delineated, for example, in Hoem (1972).

⁹ Stages in childhood such as "walking" or "reading" expose an individual to entirely new sets of experiences that may be prerequisites for the onset of more advanced behaviors.