JOURNAL OF REGIONAL SCIENCE, VOL. 37, NO. 4, 1997, pp. 653–682

GRAVITY-TYPE INTERACTIVE MARKOV MODELS—PART I: A PROGRAMMING FORMULATION OF STEADY STATES*

Tony E. Smith*
Department of Systems, University of Pennsylvania, Philadelphia, PA 19104, U.S.A.

Shang-Hsing Hsieh
Department of Transportation, National Chiao Tung University, Hsinshu, Taiwan 30050.

ABSTRACT. An important class of interactive Markov migration models is characterized by *gravity-type* transition kernels, in which migration flows in each time period are postulated to vary inversely with some symmetric measure of migration costs and directly with some population-dependent measure of attractiveness. This two-part study analyzes the uniqueness and stability properties of steady states for such processes. In this first part, it is shown that a flow version of the steady-state problem can be given a programming formulation which permits global analysis of steady-state behavior. Within this programming framework, it is shown that when attractiveness is diminished by increased population congestion, the steady states for such processes are unique. The second part of the study will employ these results to analyze the stability properties of such steady states.

1. INTRODUCTION

The class of interactive Markov models first introduced by Matras (1967) and Conlisk (1976) has been widely studied and applied in the social sciences (as for example in Conlisk, 1982, 1990; Bartholomew, 1982, 1985; DePalma and Lefevre, 1983; and Kulkarni and Kumar, 1989). In modeling collective population behavior, De Palma and Lefevre (1989) have shown that interactive Markov models are a direct consequence of population-dependent choice behavior by individuals. With respect to migration behavior in particular, such population dependencies can often be characterized in terms of those agglomeration effects (both positive and negative) which determine the relative attractiveness of population centers to migrators. The most common models of this type are gravity models, in which migration choices are positively influenced by the attraction effects of population centers and negatively influenced by the distance-deterrence effects of moving to these centers (as for

*The authors are indebted to the referees for numerous helpful comments.

Received March 1996; revised November 1996; accepted November 1996.

Blackwell Publishers, 350 Main Street, Malden, MA 02148, USA and 108 Cowley Road, Oxford, OX4 1JF, UK.

example in Poot, 1984; Kanaroglou et al., 1986a, 1986b; Nijkamp and Poot, 1987; Boots and Kanaroglou, 1988).

Although these models have been widely applied, surprisingly little is known about their structural properties. In particular, while steady states for such models are known to exist under very general conditions, there are very few analytical results on either the uniqueness or stability properties of such steady states (Conlisk, 1992). From a theoretical viewpoint, such properties are of special importance for the present class of dynamical models. In particular, such models were originally proposed by Conlisk (1976) as low-dimensional *deterministic approximations* to very high-dimensional Markov chains.[1] Moreover, while for large populations this approximation can be shown to be very good for any initial segment of the deterministic process (Brumelle and Gerchak, 1980 and in Lehoczky, 1980), the *asymptotic* behavior of the deterministic and stochastic versions can in principle be quite different. In particular, such asymptotic behavior is only guaranteed to be the same when the deterministic process exhibits a *unique globally stable* steady state (see the Corollary to Theorem 3.3 in Brumelle and Gerchak, 1980). Thus, if one wishes to employ the steady-state properties of interactive Markov chains to draw inferences about asymptotic population behavior, then it is of prime importance to establish conditions under which such steady states are unique and globally stable.

Hence the central objective of the present two-part paper is to establish conditions for uniqueness and stability of steady states for a rather general class of *gravity-type interactive Markov chains,* characterized by continuously differentiable attraction functions and symmetric deterrence functions. Our approach focuses on the *spatial-flow chains* implicit in such models, and shows that the steady-state problem for these flows can be given an explicit programming formulation. More precisely, it is shown in Part I that the steady states for each such process are equivalent to the Karush-Kuhn-Tucker (*KKT*) points of an appropriately defined minimization problem. Hence uniqueness of steady states is equivalent to uniqueness of minima. In particular, it is shown that for 'pure congestion' processes in which increased population densities always decrease the attractiveness of population centers, the objective function for the associated minimization problem is strictly convex, so that steady states are always unique. More generally, it is shown that for mature systems in which the total population is large relative to the number of regions, there exists at most one 'fully congested' steady state.

In Part II (Smith and Hsieh, 1997) the stability properties of gravity-type interactive Markov chains are analyzed in terms of the continuous (differential equation) versions of such models. More general models have been studied by

[1]However, in some cases it is possible to regard these deterministic models as viable models of behavior in their own right. For further discussion of this point see the concluding remarks to Part II.

DePalma and Lefevre (1983), Haag and Weidlich (1988a, 1988b), and Dendrinos and Sonis (1990), among others. Our main purpose is to show that for the present class of gravity-type models, the objective function in Part I is a global Lyapunov function for appropriately defined continuous versions of the corresponding flow processes. This Lyapunov property in turn yields a number of powerful global stability results. In particular, it is shown that each such process must always converge to its set of steady states. Hence, for the continuous version of pure-congestion processes, the unique steady state is globally asymptotically stable. More generally, the set of locally stable steady states for any gravity-type process is shown to correspond precisely to the set of isolated local minima for its associated Lyapunov function. Finally, it is shown that for the case of unique steady states, global convergence is inherited by all gravity-type interactive Markov chains with sufficiently small adjustments, that is, which are 'sufficiently close' to their continuous-time versions.

To establish the results for Part I, we begin in the next section with a formal definition of gravity-type interactive Markov chains together with several alternative characterizations of their steady states. This is followed in Section 3 by a formulation and analysis of the spatial flows implicit in each gravity-type interactive Markov chain. In particular, the steady states for these spatial-flow chains are given an explicit characterization and are shown to exhibit a one-to-one correspondence with steady states for the underlying interactive Markov chains. In Section 4 the steady-state problem for spatial-flow chains is shown to have a simple programming formulation. In particular, this formulation is employed to show that steady states for the 'pure-congestion' case are always unique. Finally, in Section 5 we lay the groundwork for the stability analysis of interactive Markov chains in Part II. Here it is shown that stability can fail to hold even when steady states are unique. However, if time periods are made short enough to ensure that only a small fraction of the population considers migrating in any given period, then such processes exhibit much greater stability. This motivates an analysis of the limiting version of interactive Markov chains obtained by letting the period length go to zero. The resulting class of *interactive Markov processes* and associated *spatial flow processes* are formalized in this final section, and are analyzed in greater detail in Part II.

2. GRAVITY-TYPE INTERACTIVE MARKOV CHAINS

For any fixed population of N behaving units (*individuals*) distributed over a finite set of spatial locations (*regions*), $i \in I = \{1, \ldots, q\}$, let p_i^t denote the *population fraction* in region, $i \in I$, in time period t.[2] Migration behavior in such a system is said to be governed by a *Markov chain* with *transition matrix,* $\mathbf{M} =$

[2]Alternatively, p_i^t can be interpreted as the probability that a randomly sampled individual occupies region i in period t. The present nonprobabilistic interpretation of p_j^t as a population fraction is designed to be consistent with the notion of 'interactive Markov chains' defined below. See footnote 10 below for further discussion.

$(M_{ij}: i, j \in I)$, iff

$$(1) \qquad p_j^{t+1} = \sum_i p_i^t M_{ij}, \qquad j \in I, \qquad t \in \mathbb{Z}_+$$ [3]

From a behavioral viewpoint, such models are severely limited by the constancy of the transition matrix, **M**, which implies that individual migration decisions are not influenced by the current distribution of the population.[4] Hence, if one postulates that migration behavior in period t does depend on the current *population distribution*, $\mathbf{p}^t = (p_i^t: i \in I)$, then the model in (1) can be generalized by allowing **M** to depend on $\mathbf{p}^t$ as follows:

$$(2) \qquad p_j^{t+1} = \sum_i p_i^t M_{ij}(\mathbf{p}^t), \qquad j \in I, \qquad t \in \mathbb{Z}_+$$

If the probability simplex,

$$\mathbb{P}_q = \left\{ \mathbf{p} = (p_i: i \in I) \in \mathbb{R}_+^q: \sum_i p_i = 1 \right\}$$

is now taken to represent the set of *population distributions* on $I = \{1, \ldots, q\}$ then (following Conlisk, 1976) these models may be formalized as follows:

DEFINITION 1: (*i*) *Each continuously differentiable matrix-valued function,* $\mathbf{M}: \mathbb{P}_q \to \mathbb{R}_+^{q \times q}$, *satisfying* $\Sigma_j M_{ij}(\mathbf{p}) = 1$ *for all* $i \in I$ *and* $\mathbf{p} \in \mathbb{P}_q$ *is designated as a transition function on I.*
(*ii*) *The model in (2) is then designated as an interactive Markov chain*[5] *on I with transition function,* **M**.[6]

In particular, $M_{ij}(\mathbf{p}^t)$ represents the fraction of individuals at i who migrate to region j in period t given current population distribution $\mathbf{p}^t$. For later purposes, we note that if each $\mathbf{p} \in \mathbb{P}_q$ is treated as a row vector and if the associated transition matrix for $\mathbf{p}$ is denoted by $\mathbf{M}(\mathbf{p}) = [M_{ij}(\mathbf{p}): i, j \in I]$, then (2) can be

[3]The following standard notation is employed throughout. Let $\mathbb{Z}_+(\mathbb{R}_+)$ denote the *nonnegative* integers (reals) and let $\mathbb{Z}_{++}(\mathbb{R}_{++})$ denote the *positive* integers (reals).

[4]For a detailed discussion of this point, together with references to additional relevant literature, see Plane (1993).

[5]As mentioned in the introduction, the notion of an 'interactive Markov chain' originated as an approximation to a Markov chain with states corresponding to all possible spatial allocations of a finite population. The practical significance of this approximation is to reduce the dimensionality of the state space down to the number of discrete locations (regions) considered. For our present purposes, these approximation issues can be avoided by focusing exclusively on the deterministic model in which each p_i^t is interpreted simply as a continuous fraction of the population N. It should be noted, however, that these deterministic models do indeed represent instances of Markov chains with continuous state spaces and 'degenerate' transition kernels, as treated for example in Brumelle and Gerchak (1980) [and in more detail by Doob (1953, section V.5)].

[6]Note also that model (2) can be viewed as an instance of the 'universal' discrete-time model of relative dynamics studied by Dendrinos and Sonis (1990, p. 21).

written in matrix form as

$$\mathbf{p}^{t+1} = \mathbf{p}^t\mathbf{M}(\mathbf{p}^t), \qquad t \in \mathbb{Z}_+ \tag{3}$$ [7]

As mentioned in the introduction, it is reasonable to assume that these fractions are positively influenced by the attraction effects of region j and are negatively influenced by the distance-deterrence effects of moving from i to j. In particular, the attractiveness of region j for potential migrants will generally depend on the current population level in j (or factors such as population density that are influenced by this population level). But since the total population N is taken to be fixed, it follows that the current population level in j is given by Np_j^t (and similarly that population density is given by Np_j^t/A_j where A_j is the area of region j). Hence such population factors are determined solely by population fractions, p_j^t, and it follows that *attractiveness* can be written as a function of p_j^t, say $a_j(p_j^t)$. In this context, it may then be postulated that $M_{ij}(\mathbf{p}^t)$ increases with $a_j(p_j^t)$. Similarly, if distance-deterrence effects are representable by some measure of *migration costs*, c_{ij}, between i and j, and if *accessibility*, $f(c_{ij})$, of i to j is a decreasing function of c_{ij}, then it may also be postulated that $M_{ij}(\mathbf{p}^t)$ increases with $f(c_{ij})$. The simplest model incorporating these effects is the classical *gravity model* in which $M_{ij}(\mathbf{p}^t)$ is postulated to be proportional to $a_j(p_j^t)f(c_{ij})$, and hence which (in view of the normalization condition $\Sigma_j M_{ij}(\mathbf{p}) = 1$) takes the explicit form

$$M_{ij}(\mathbf{p}^t) = \frac{a_j(p_j^t)f(c_{ij})}{\sum_k a_k(p_k^t)f(c_{ik})} \tag{4}$$

Example 1: A 'Logit' Model of Migration

This type of interactive Markov model can be illustrated within the choice-theoretic framework of DePalma and Lefevre (1983) as follows. Consider a migration model in which the *housing cost*, h_j, in j depends on the current population, p_j^t, in j, so that the *total moving cost* for a migrant from i to j is given by $h_j(p_j^t) + c_{ij}$. Here it is postulated that the *net utility* of region j for a randomly sampled individual in region i is of the form, $U_{ij}(p_j^t) = B_{ij} - \theta[h_j(p_j^t) + c_{ij}]$, where the 'perceived benefits' of j are assumed to vary among individuals at i, and in particular, are assumed to be independently Gumbel-distributed random variates, B_{ij}, with common 'dispersion parameter,' θ^{-1}.[8] In this context, it is well known that the probability, $M_{ij}(\mathbf{p}^t)$, of a randomly sampled (utility maximizing) migrant from i will choose region j must have the

[7] Alternatively, one may employ *column*-vector notation by employing the transpose of the transition matrix and writing, $\mathbf{p}^{t+1} = \mathbf{M}(\mathbf{p}^t)^T\mathbf{p}^t$ or by redefining $\mathbf{M}_{ij}(\mathbf{p}^t)$ to be the fraction of individuals at j who move to i.

[8] An alternative quadratic form of such utilities is given in Weidlich (1988, p. 335).

following 'logit' form:[9]

$$M_{ij}(\mathbf{p}^t) = \frac{\exp\{-\theta[h_j(p_j^t) + c_{ij}]\}}{\sum_k \exp\{-\theta[h_k(p_k^t) + c_{ik}]\}}, \qquad i, j \in I$$

This is immediately seen to be of the form (4) with $a_j(p_j^t) = \exp[-\theta h_j(p_j^t)]$ and $f(c_j) = \exp[-\theta c_{ij}]$.[10] A variety of similar examples of (4) can also be developed within this choice-theoretic framework. ■

Note in the above example that if increased population levels reflect increased competition for housing, then h_j may well increase with p_j^t, so that attraction, a_j, decreases with p_j^t. More generally, those cases in which higher population densities always decrease the attractiveness of regions will be designated as 'pure-congestion' cases, and are of special interest in the analysis to follow. Note also that migration costs in this example are taken to be *constant*. More generally, although distribution of population can significantly influence the relative attractiveness of regions, it is assumed in gravity models that this distribution has little effect on transport costs.[11] Hence, it will usually be convenient to suppress c_{ij} and write simply, $f_{ij} = f(c_{ij})$. Our final assumption is that migration costs between i and j are the same as between j and i, so that $f_{ij} = f_{ji}$ for all $i, j \in I$. This *symmetry* assumption is not unreasonable at the interregional scale (and indeed, is standard in essentially all empirical applications of gravity models).[12] With these observations, we now formalize the relevant class of transition functions for our purposes as follows:

DEFINITION 2. (*i*) *For any positive continuously differentiable*[13] attraction functions, a_j: $\mathbb{R}_+ \rightarrow \mathbb{R}_{++}$, $j \in I$, *and positive* accessibility weights, f_{ij}, *with* $f_{ij} = f_{ji}$ *for all* $i, j \in I$, *the positive transition function,* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$: $\mathbb{P}_q \rightarrow \mathbb{R}_{++}^{q \times q}$, *defined for all*

[9]This logit form corresponds to the model in Section 2 of De Palma and Lefevre (1983) with $\mu = \theta^{-1}$. An alternative class of dynamical models involving logit forms can be found in Sonis (1992).

[10]Note as in footnote 2 that the identification of choice probabilities with population fractions, $M_{ij}(\mathbf{p}^t)$, can also be justified by the approximation arguments in Brumelle and Gerchak (1980) and Lehoczky (1980) [as observed by De Palma and Lefevre (1983)].

[11]In particular, while local population congestion can have profound effects on within-city traffic flows, such effects are often less significant at the interregional level. Hence at this scale of analysis, transport costs are usually treated as constant. Possible relaxations of this constancy assumption are discussed in the concluding section of Part II.

[12]It should be noted that apparent asymmetries between migration costs often relate to properties of the origin and destination regions, rather than to flows between them. For example if it is more expensive to locate in one region than another, then such differences can be reflected in their respective attraction factors (see also Haag and Weidlich, 1988a, pp. 16–17). Possible relaxations of this symmetry assumption are discussed in the concluding section of Part II.

[13]It should be noted that continuous differentiability of attraction functions is only required in Section 5, where the differential-equation approximation of interactive Markov chains is developed. All other results are easily seen to hold for *continuous* attraction functions. However, we choose for convenience to maintain the same definitions and assumptions throughout.

$\mathbf{p} \in \mathbb{P}_q$ *by*

(5) $$M_{ij}^{\mathbf{a},\mathbf{f}}(\mathbf{p}) = \frac{a_j(p_j)f_{ij}}{\sum_k a_k(p_k)f_{ik}}, \qquad i,j \in I$$

is designated as a gravity-type transition function *on I. The vector of attraction functions,* $\mathbf{a} = (a_j\colon j \in I)$, *is designated as the* attraction profile *for* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$ *and the matrix of accessibility weights,* $\mathbf{f} = (f_{ij}\colon ij \in I \times I)$, *is designated as the* accessibility matrix *for* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$.[14]
(ii) If in addition each attraction function a_j *is nonincreasing then* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$ *is said to exhibit* pure congestion effects.[15]

Hence, for any given gravity-type transition function in (5), the adjustment process in (2) takes the explicit form

(6) $$p_j^{t+1} = \sum_i p_i^t \frac{a_j(p_j^t)f_{ij}}{\sum_k a_k(p_k^t)f_{ik}}, \qquad j \in I, \qquad t \in \mathbb{Z}_+$$

and is designated as a *gravity-type interactive Markov chain.*

Our interest in such dynamical models focuses on uniqueness and stability of their steady states. A *steady state* for a transition function, **M**, is a distribution, $\mathbf{p} \in \mathbb{P}_q$, which remains invariant in (3), that is, which satisfies the fixed-point condition, $\mathbf{p} = \mathbf{pM}(\mathbf{p})$. The class of steady states for **M** is denoted by

(7) $$S(\mathbf{M}) = \{\mathbf{p} \in \mathbb{P}_q\colon \mathbf{p} = \mathbf{pM}(\mathbf{p})\}$$

For gravity-type transition functions in particular, the fixed-point condition, $\mathbf{p} = \mathbf{pM}(\mathbf{p})$, is seen from (5) to take the explicit form:

(8) $$p_j = \sum_i p_i \frac{a_j(p_j)f_{ij}}{\sum_k a_k(p_k)f_{ik}}, \qquad j \in I$$

It is convenient to record the following useful properties of steady states for later use. If the set of *positive population distributions* is denoted by $\mathbb{P}_q^+ = \mathbb{P}_q \cap \mathbb{R}_{++}^q$, then

[14]This class of models is also an instance of the 'gravitational interaction models' outlined in Dendrinos and Sonis (1990, pp. 165–167).

[15]It should be emphasized that this use of the term 'congestion' is meant only to be suggestive. Many other types of population-related effects could produce such a monotonic relation. In Example 1, for instance, this relation is more accurately described in terms of the housing-market response to the increased demand generated by higher population levels. Similar illustrations could be developed in terms of job-market responses, for example.

PROPOSITION 1: *For each transition function,* $\mathbf{M}$: $\mathbb{P}_q \rightarrow \mathbb{R}_+^{q\times q}$, *there exists at least one steady state, that is,* $S(\mathbf{M}) \neq 0$. *In addition, the steady states for each gravity-type transition kernel,* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, *are positive, that is,* $S(\mathbf{M}^{\mathbf{a},\mathbf{f}}) \subseteq \mathbb{P}_q^+$.

Proof. Since every continuously differentiable function is necessarily continuous, it follows that the mapping, $\mathbf{T}$: $\mathbb{P}_q \rightarrow \mathbb{P}_q$, defined by $\mathbf{T}(\mathbf{p}) = \mathbf{pM}(\mathbf{p})$, $\mathbf{p} \in \mathbb{P}_q$, must also be continuous. But since $\mathbb{P}_q$ is convex and compact, it then follows from the Brouwer Fixed-Point Theorem that there must exist at least one fixed point, $\mathbf{p} = \mathbf{T}(\mathbf{p})$, which is by definition an element of $S(\mathbf{M})$. Moreover, for each gravity-type transition function, $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, the positivity of $\mathbf{a}$ and $\mathbf{f}$ implies that $\mathbf{M}^{\mathbf{a},\mathbf{f}}(\mathbf{p})$ is positive. But since at least one component of $\mathbf{p}$ is positive, and all are nonnegative, it then follows at once from (8) that all components of $\mathbf{p}$ must be positive. ■

The following example shows, however, that such steady states need not be unique:

Example 2: Nonuniqueness of Steady States

As the extreme opposite of pure congestion, one may consider a 'pure agglomeration' case in which higher population densities are always more attractive.[16] In particular, let $q = 3$ and suppose that each attraction function is of the form, $a_j(p_j) = \exp[6p_j]$, $j = 1, 2, 3$. Suppose also that the spatial configuration is symmetric with migration costs, $c_{ii} = 0$ and $c_{ij} = 2$ for all distinct $i, j = 1, 2, 3$, and with accessibilities of the form, $f_{ij} = f(c_{ij}) = \exp[-c_{ij}]$. Then the steady states for this case can be depicted as in Figure 1, where the triangle represents the probability simplex, $\mathbb{P}_3$, with vertices denoting the 'complete concentration' states, (1, 0, 0), (0, 1, 0), (0, 0, 1), in which all of the population resides in a single region. The notion of 'pure agglomeration' suggests that each vertex should be a steady state. Although this is essentially true, it follows from Proposition 1 above that all steady states must be *strictly positive.* Hence the actual steady state shown at vertex (1, 0, 0), for example, is approximately (0.99932, 0.00034, 0.00034) (with the others being corresponding permutations of these values). In addition, symmetry implies that midpoint of the triangle, (⅓, ⅓, ⅓), must also be steady state. This steady state is clearly unstable, however, since slight increases in the population of any region will increase the attractiveness of that region relative to others. (Stability properties will be discussed in detail below). Finally, there are three other 'pairwise' steady states in which all population is essentially shared equally by two regions. In the case of regions 1 and 2, for example, the corresponding pairwise steady state is approximately (0.49625, 0.049625, 0.00750). Hence this pure agglomeration example exhibits *seven* distinct steady states. ■[17]

[16]As with 'congestion,' the term 'agglomeration' is meant only to be suggestive.

[17]For further discussion of this example, see the concluding remarks to Part II.

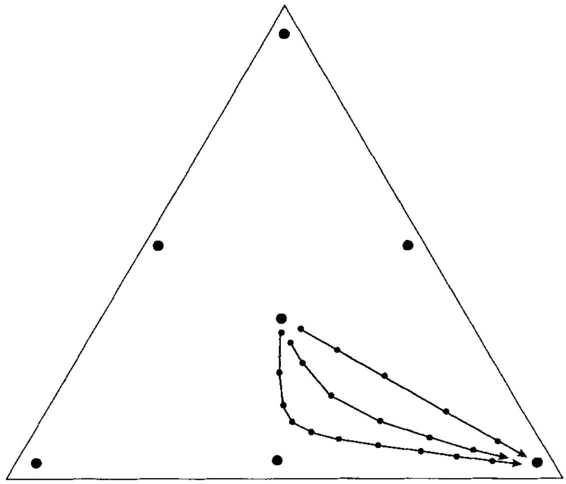

FIGURE 1: Example of Multiple Steady States.

For our later purposes, it is convenient to develop an equivalent form of the steady-state conditions, (8), for gravity-type transition functions:

THEOREM 1: Equivalent Steady-State Conditions. *For any gravity-type transition function,* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, *a distribution,* $\mathbf{p} \in \mathbb{P}_q$, *is an element of* $S(\mathbf{M}^{\mathbf{a},\mathbf{f}})$ *iff there exists a positive scalar,* μ, *such that*

(9) $$p_j = \mu a_j(p_j) \sum_k a_k(p_k) f_{kj}, \qquad j \in I$$

This scalar is uniquely given by $\mu = [\Sigma_{ij} a_i(p_i) a_j(p_j) f_{ij}]^{-1}$

Proof. To see that (9) implies (8), one need only substitute (9) into the right-hand side of (8) and observe from the symmetry of **f** that

$$\begin{aligned}\sum_i p_i \frac{a_j(p_j) f_{ij}}{\sum_k a_k(p_k) f_{ik}} &= \sum_i \left[\mu a_i(p_i) \sum_k a_k(p_k) f_{ki}\right] \frac{a_j(p_j) f_{ij}}{\sum_k a_k(p_k) f_{ik}} \\ &= \mu a_j(p_j) \sum_i \left[\sum_k a_k(p_k) f_{ik}\right] \frac{a_i(p_i) f_{ij}}{\sum_k a_k(p_k) f_{ik}} \\ &= \mu a_j(p_j) \sum_i a_i(p_i) f_{ij} \\ &= p_j\end{aligned}$$

To establish the converse, observe that if **p** satisfies (8) then, letting $a_j = a_j(p_j)$ for notational simplicity, it follows from (8) that for all $j \in$ I,

$$
\begin{aligned}
\frac{p_j}{a_j} = \sum_i p_i \frac{f_{ij}}{\sum_k a_k f_{ik}} &\Rightarrow \frac{p_j}{a_j \sum_k a_k f_{jk}} \\
(10) \qquad &= \sum_i \frac{p_i}{\sum_k a_k f_{ik}} \frac{f_{ij}}{\sum_k a_k f_{jk}} = \sum_i \left(\frac{p_i}{a_i \sum_k a_k f_{ik}} \right) \frac{a_i f_{ji}}{\sum_k a_k f_{jk}} \\
&\Rightarrow \theta_j = \sum_i \theta_i \frac{a_i f_{ji}}{\sum_k a_k f_{jk}}
\end{aligned}
$$

where $\theta_j = p_j/[a_j \Sigma_k a_k f_{jk}]$, $j \in I$. But since all attraction functions and accessibility weights are positive, it then follows from the identity $1 = \Sigma_i (a_i f_{ji}/\Sigma_k a_k f_{jk}]$ that each θ_j is a positive convex combination of $(\theta_1, \ldots, \theta_q)$. This is only possible if all θ_j's are equal. For if $\min_i \theta_i < \max_i \theta_i = \theta_\alpha$, say, then since all positive convex combinations must be strictly less than the maximum value, θ_α, this would imply from (10) that

$$
\max_i \theta_i > \sum_j \left(\frac{a_j f_{\alpha j}}{\sum_k a_k f_{\alpha k}} \right) \theta_j = \theta_\alpha
$$

which contradicts the definition of θ_α. Hence $\min_i \theta_i = \max_i \theta_i$, and all θ_i's are equal. In particular, if we now set $\mu = \theta_1$ then it follows that

$$
\frac{p_j}{a_j \sum_k a_k f_{jk}} = \theta_j = \theta_1 = \mu
$$

for all $j \in I$, and hence that (9) holds for this choice of μ. Finally, the identity, $\mu = [\Sigma_{ij} a_i (p_i) a_j(p_j) f_{ij}]^{-1}$, follows at once from (9) together with the normalization condition, $\Sigma_j p_j = 1$. ■

3. EXTENSION TO SPATIAL FLOWS

To motivate our approach to the analysis of these steady states, we begin by observing that our present notion of state transitions for spatial populations necessarily involves *spatial flows*. Moreover, these flows involve costs which are fundamental determinants of behavior. For gravity-type transition functions in particular, it turns out that the structure of this behavior is most easily revealed by modeling spatial flows explicitly. Hence we now shift our attention from 'population stocks,' which constitute the basic state variables of interactive Markov chains, to 'population flows' which define the changes in these stocks.[18]

In particular, we now consider a space of *flow states*, $ij \in I \times I$, and develop the class of 'spatial flow chains' on $I \times I$, generated by such transition functions. The population distributions, $\mathbf{p}^t = (p_i^t: i \in I)$, for interaction Markov chains are replaced by their corresponding *flow distributions*, $\mathbf{P}_t = [P_t(ij): ij \in I \times I]$, where $P_t(ij)$ denotes the fraction of population in flow state ij in period t, that is, the fraction of population flowing from i to j in period t. If the probability simplex

$$(11) \qquad \mathbb{P}_{q\times q} = \left\{\mathbf{P} = [P(ij): ij \in I \times I] \in \mathbb{R}_+^{q\times q}: \sum_{ij} P(ij) = 1\right\}$$

is taken to represent the set of *flow distributions* on $I \times I$, then for each $\mathbf{P} \in \mathbb{P}_{q\times q}$, we denote the *row marginals* of $\mathbf{P}$ by $P(i\cdot) = \Sigma_j P(ij)$, $i \in I$, and *column marginals* of $\mathbf{P}$ by $P(\cdot j) = \Sigma_i P(ij)$, $j \in I$. The corresponding *row marginal distribution* and *column marginal distributions* for $\mathbf{P}$ are denoted, respectively, by $\mathbf{P}(I\cdot) = [P(i\cdot): i \in I] \in \mathbb{P}_q$ and $\mathbf{P}(\cdot I) = [P(\cdot j): j \in I] \in \mathbb{P}_q$. If $\mathbf{P}_t$ is the relevant flow distribution in period t, then we now take $P_t(i\cdot)$ to be the fraction of population in i at the *beginning* of period t, and take $P_t(\cdot i)$ to be the fraction of population in i at the *end* of period t, so that by definition,

$$(12) \qquad P_t(\cdot i) = P_{t+1}(i\cdot), \qquad i \in I, \qquad t \in \mathbb{Z}_+$$

that is, the population fraction in i at the end of period t is the same as the fraction in i at the beginning of period $t + 1$. With these conventions, it follows from an inspection of (6) that for any given gravity-type transition function, $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, the fraction of population in flow state ij in period t must be given by

$$(13) \qquad \begin{aligned} P_t(ij) &= P_t(i\cdot)M_{ij}^{\mathbf{a},\mathbf{f}}[\mathbf{P}_{t-1}(\cdot I)] = P_t(i\cdot)M_{ij}^{\mathbf{a},\mathbf{f}}[\mathbf{P}_t(I\cdot)] \\ &= P_t(i\cdot)\frac{a_j[P_t(j\cdot)]f_{ij}}{\sum_k a_k[P_t(k\cdot)]f_{ik}}, \qquad ij \in I \times I \end{aligned}$$

Hence (13) defines the sequence of spatial flows generated by $\mathbf{M}^{\mathbf{a},\mathbf{f}}$. This may be expressed in terms of 'flow adjustments' from period to period by evaluating (13) at $t + 1$ and applying (12) to the right-hand side to obtain

$$(14) \qquad P_{t+1}(ij) = P_t(\cdot i)\frac{a_j[P_t(\cdot j)]f_{ij}}{\sum_k a_k[P_t(\cdot k)]f_{ik}}, \qquad ij \in I \times I, \qquad t \in \mathbb{Z}_+$$

[18]Although flow formulations are quite standard in network models (such as in traffic networks for example), they appear to be relatively new to interactive Markov models. A notable exception is the (implicit) flow analysis employed by Brumelle and Gerchak (1989, p. 76) to characterize interactive Markov models as limiting forms of Markov models.

In this form, we now designate the sequence of flow distributions given by (14) for any initial flow distribution, $\mathbf{P}_0 \in \mathbb{P}_{q\times q}$, as the *spatial-flow chain* generated by transition function $\mathbf{M}^{\mathbf{a},\mathbf{f}}$. To see the relation of this spatial-flow chain to the original interactive Markov chain, observe by summing both sides of (14) with respect to i that

$$(15) \qquad P_{t+1}(\cdot j) = \sum_i P_t(\cdot i) \frac{a_j[P_t(\cdot j)]f_{ij}}{\sum_k a_k[P_t(\cdot k)]f_{ik}}, \qquad j \in I, \qquad t \in \mathbb{Z}_+$$

and hence that the *column marginal distributions,* $\mathbf{P}_t(\cdot I) \in \mathbb{P}_q$ define precisely the interactive Markov chain in (6).[19] Thus, if we now define a *steady-state flow distribution* for this spatial-flow chain to be any $\mathbf{P} \in \mathbb{P}_{q\times q}$ satisfying

$$(16) \qquad P(ij) = P(\cdot i) \frac{a_j[P(\cdot j)]f_{ij}}{\sum_k a_k[P(\cdot k)]f_{ik}}, \qquad ij \in I \times I$$

then is also follows at once by summing (16) with respect to i that the column marginal distribution, $\mathbf{P}(\cdot I) \in \mathbb{P}_q$, must satisfy

$$(17) \qquad P(\cdot j) = \sum_i P(\cdot i) \frac{a_j[P(\cdot j)]f_{ij}}{\sum_k a_k[P(\cdot k)]f_{ik}}, \qquad j \in I$$

and hence must be a steady state for $\mathbf{M}^{\mathbf{a},\mathbf{f}}$. Our objective is to study the properties of these steady states in terms of their corresponding steady-state flow distributions.

To do so, we first note by summing (16) with respect to j that every steady-state flow distribution satisfies the natural *flow-balance condition*

$$(18) \qquad P(i\cdot) = P(\cdot i), \qquad i \in I$$

that is, the flows going into i must be the same as the flows going out of i. Using this condition, we may now give the following sharper characterization of these steady-state flow distributions:

THEOREM 2: Steady-State Flows. *For any gravity-type transition function,* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, *a flow distribution,* $\mathbf{P} \in \mathbb{P}_{q\times q}$, *is a steady-state flow distribution for* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$ *iff there exists some scalar* $\mu > 0$ *such that*

$$(19) \qquad P(ij) = \mu a_i[P(\cdot i)]a_j[P(\cdot j)]f_{ij}, \qquad ij \in I \times I$$

Proof. (i) To see that every $\mathbf{P}$ of the form (19) is a steady-state flow distribution for $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, observe first from (19), together with the symmetry of $\mathbf{f}$,

[19]The use of column marginals here is simply a matter of convenience. In particular, if we replace t everywhere by $t + 1$ in (15) and apply (12) once again, we obtain the same adjustment relation for the row marginals.

that $P(ij) = P(ji)$ for all $ij \in I \times I$. Hence, replacing $P(ij)$ by $P(ji)$ in (19) and summing with respect to j we obtain,

$$P(\cdot i) = \mu a_i[P(\cdot i)] \sum_k a_k[P(\cdot k)]f_{ik}, \qquad i \in I \tag{20}$$

Finally, dividing (19) by (20), and multiplying by $P(\cdot i)$ we see that (16) must hold.

(ii) To establish the converse, observe that if $\mathbf{P}$ satisfies (16) then the column marginal distribution, $\mathbf{P}(\cdot I) \in \mathbb{P}_q$, satisfies (17), which implies from Theorem 1 that (20) must hold for some $\mu > 0$. Hence it follows from (20) and (16) together with the positivity of $\mathbf{a}$ and $\mathbf{f}$ that each steady-state flow *distribution is strictly positive.* Thus by letting

$$\theta_i = \frac{P(\cdot i)}{\sum_k a_k[P(\cdot k)]f_{ik}}, \qquad i \in I$$

we see from (16) together with the symmetry of $\mathbf{f}$ that

$$\frac{P(ij)}{P(ji)} = \frac{\theta_i a_j[P(\cdot j)]}{\theta_j a_i[P(\cdot i)]} = \frac{\phi_i}{\phi_j} \tag{21}$$

where $\phi_i = \theta_i/a_i\,[P(\cdot i)]$. Hence by employing the flow-balance condition (18), we may conclude from (21) that

$$\begin{aligned} P(ij)\phi_j = P(ji)\phi_i &\Rightarrow \sum_j P(ij)\phi_j = P(\cdot i)\phi_i = P(i\cdot)\phi_i \\ &\Rightarrow \phi_i = \sum_j \left[\frac{P(ij)}{P(i\cdot)}\right]\phi_j \end{aligned} \tag{22}$$

which together with the positivity of $\mathbf{P}$ and the identity $P(i\cdot) = \Sigma_j P(ij)$ implies that each ϕ_i is a positive convex combination of $(\phi_1, \ldots, \phi_q)$. But this implies, as in the proof of Theorem 1 above, that all ϕ_i's are equal, so that by definition

$$\frac{\theta_i}{a_i[P(\cdot i)]} = \phi_i = \phi_1 = \frac{\theta_1}{a_1[P(\cdot 1)]} \Rightarrow \theta_i = \frac{\theta_1}{a_1[P(\cdot 1)]}a_i[P(\cdot i)]$$

Finally, by letting $\mu = \theta_1/a_1[P(\cdot 1)]$ we may conclude that

$$P(ij) = \theta_i a_j[P(\cdot j)]f_{ij} = \mu a_i[P(\cdot i)]a_j[P(\cdot j)]f_{ij}$$

and hence that (19) holds. ■

As one basic consequence of Theorem 2, observe from the form of (19) that *every steady-state flow distribution,* $\mathbf{P}$, *for a gravity-type transition matrix,* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$ *is symmetric,* that is, that $\mathbf{P} = \mathbf{P}^T$ (where $\mathbf{P}^T$ denotes the transpose of matrix

P).[20] This symmetry property (which depends critically on the symmetry of **f**) has important implications for steady-state flow behavior. In particular it implies that steady-state flows in such processes are *reversible* in the sense that the 'inverse' process obtained by reversing the time order (analogous to running a movie film backwards) is indistinguishable from the original process, as discussed for the case of Markov chains by Howard (1971, section 9.3). From a practical viewpoint, this symmetry property is also seen to provide a natural test of gravity-type interactive Markov models. For if the system is governed by such a model and is close to a steady state, in the sense that $P_t(ij) \approx P_{t+1}(ij)$ for all ij, then it must true that $P_t(ij) \approx P_t(ji)$. Hence if steady-state behavior is observed to violate this condition then such behavior cannot be consistent with any gravity-type interactive Markov chain.[21]

Next observe from the normalization condition, $\Sigma_{ij}P(ij) = 1$, that the multiplier in (19) must have the explicit form, $\mu = (\Sigma_{ij}\, a_i\, [P(\cdot i)]a_j\, [P(\cdot j)]f_{ij})^{-1}$, and hence depends on both **a** and **f**. Given this characterization, if we now designate the class of *steady-state flow distributions* for $\mathbf{M}^{\mathbf{a},\mathbf{f}}$ by

(23) $$SFD[\mathbf{a}, \mathbf{f}] = \{\mathbf{P} \in \mathbb{P}_{q\times q}: \mathbf{P} \text{ satisfies (19) for some } \mu > 0\}$$

then it is can be shown [Smith and Hseih (1996)] that there is a simple one-to-one correspondence between $S(\mathbf{M}^{\mathbf{a},\mathbf{f}})$ and $SFD[\mathbf{a}, \mathbf{f}]$, and in particular that

THEOREM 3: Bijective Correspondence. *For each gravity-type transition function,* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, *the mapping,* $\mathbf{P}^{\mathbf{a},\mathbf{f}}: S(\mathbf{M}^{\mathbf{a},\mathbf{f}}) \to \mathbb{P}_{q\times q}$, *defined for all* $\mathbf{p} \in S(\mathbf{M}^{\mathbf{a},\mathbf{f}})$ *by*

(24) $$\mathbf{P}_{\mathbf{p}}^{\mathbf{a},\mathbf{f}}(ij) = \mu(\mathbf{p})a_i(p_i)a_j(p_j)f_{ij}, \qquad ij \in I \times I$$

with $\mu(\mathbf{p}) = [\Sigma_{ij}\, a_i(p_i)a_j(p_j)f_{ij}]^{-1}$ *is a bijection from* $S(\mathbf{M}^{\mathbf{a},\mathbf{f}})$ *to* $SFD[\mathbf{a}, \mathbf{f}]$

By employing this bijective correspondence, we may analyze properties of steady-state distributions, $\mathbf{p} \in S(\mathbf{M}^{\mathbf{a},\mathbf{f}})$, entirely in terms of their corresponding steady-state flow distributions, $\mathbf{P}_{\mathbf{p}}^{\mathbf{a},\mathbf{f}} \in SFD[\mathbf{a}, \mathbf{f}]$.

4. A PROGRAMMING FORMULATION OF STEADY-STATE FLOWS

Given these general properties of steady-state flow distributions for $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, we turn now to a more explicit analysis of their structural properties. Our main result is to show that these steady-state flows can be characterized in terms of an appropriately defined programming problem. To construct this program-

[20]Note also from the identity, $P(\cdot i)\mathbf{M}_{ij}^{\mathbf{a},\mathbf{f}}[\mathbf{P}(\cdot I)] = P(ij)$, in (16) that this symmetry property is equivalent to the following 'detailed balance' condition for steady-state population distributions, $P(\cdot i)\mathbf{M}_{ij}^{\mathbf{a},\mathbf{f}}[P(\cdot I)] = P(\cdot j)\mathbf{M}_{ji}^{\mathbf{a},\mathbf{f}}[\mathbf{P}(\cdot I)]$, $i, j \in I$, [see also Weidlich (1988].

[21]Note however that such a test implicitly requires symmetry of deterrence weights. Moreover, even if migration costs are symmetric, there may exist other significant deterrence factors which are asymmetric, and hence which are consistent with steady-state asymmetries. Such behavior is discussed further in the concluding remarks to Part II.

ming problem, we first introduce an extension of the classical (Kullback-Leibler) *divergence function,* D: $\mathbb{R}_{+}^{q\times q} \times \mathbb{R}_{++}^{q\times q} \rightarrow \mathbb{R}$, defined for all nonnegative matrices, $\mathbf{V} = [V(ij): ij \in I \times I] \in \mathbb{R}_{+}^{q\times q}$, and positive matrices, $\mathbf{W} = [W(ij): ij \in I \times I] \in \mathbb{R}_{++}^{q\times q}$, by

$$D(\mathbf{V}, \mathbf{W}) = \sum_{ij} V(ij) \log [V(ij)/W(ij)] \tag{25}$$

where by convention, $0 \log (0) = 0$. Observe in particular that for any fixed matrix, $\mathbf{W} \in \mathbb{R}_{++}^{q\times q}$, the restricted function, $D(\cdot, \mathbf{W})$, is continuously differentiable on the open subset, $\mathbb{R}_{++}^{q\times q} \subset \mathbb{R}_{+}^{q\times q}$. Next, for any attraction profile, $\mathbf{a}$, and accessibility matrix, $\mathbf{f}$, we may employ this divergence function to construct an objective function, Z: $\mathbb{R}_{+}^{q\times q} \rightarrow \mathbb{R}$, defined for all $\mathbf{F} \in \mathbb{R}_{+}^{q\times q}$ by[22]

$$Z(\mathbf{V}) = D(\mathbf{V}, \mathbf{f}) - \sum_{i} \int_0^{V(i\cdot)} \log [a_i(x)]\, dx - \sum_{j} \int_0^{V(\cdot j)} \log [a_j(x)]\, dx \tag{26}$$

With these definitions, we now consider the *programming problem,* $\mathscr{P}[\mathbf{a}, \mathbf{f}]$:

$$\text{minimize: } Z(\mathbf{P}) \qquad \text{subject to: } \mathbf{P} \in \mathbb{P}_{q\times q}$$

where by definition, $\mathbb{P}_{q\times q} \subseteq \mathbb{R}_{+}^{q\times q}$. To analyze this programming problem, we begin by observing that *all local minima for* $\mathscr{P}[\mathbf{a}, \mathbf{f}]$ *are positive,* that is, are elements of the set of *positive distributions,* $\mathbb{P}_{q\times q}^{+} = \mathbb{P}_{q\times q} \cap \mathbb{R}_{++}^{q\times q}$. To see this, consider any distribution, $\mathbf{P} \in \mathbb{P}_{q\times q}$, with $P(ij) = 0$ for some $ij \in I \times I$, and let $I_0(\mathbf{P}) = \{ij: P(ij) = 0\} \neq \emptyset$. Then since our assumptions on $\mathbf{a}$ and $\mathbf{f}$ imply that Z is continuously differentiable on $\mathbb{R}_{++}^{q\times q}$, and since for any choice of positive distribution, $\mathbf{Q} \in \mathbb{P}_{q\times q}^{+}$, we must have $\alpha\mathbf{Q} + (1 - \alpha)\mathbf{P} \in \mathbb{P}_{q\times q}^{+}$ for all $\alpha \in (0, 1]$, it follows that the function, $Z(\alpha) = Z[\alpha\mathbf{Q} + (1 - \alpha)\mathbf{P}]$, is differentiable on $(0, 1]$. In particular, this derivative is seen to be of the form

$$\frac{d}{d\alpha} Z(\alpha) = \frac{d}{d\alpha} D[\alpha\mathbf{Q} + (1 - \alpha)\mathbf{P}, \mathbf{f}] - B(\alpha) \tag{27}$$

where the second term, $B(\alpha)$, is a bounded function on $[0, 1]$, and where the first term is given by

$$\begin{aligned} &\frac{d}{d\alpha} \sum_{ij} \{P(ij) + \alpha[Q(ij) - P(ij)]\} \log (\{P(ij) + \alpha[Q(ij) - P(ij)]\}/f_{ij}) \\ &= \sum_{ij} Q(ij) \log \{P(ij) + \alpha[Q(ij) - P(ij)]\} + R(\alpha) \\ &= \sum_{ij \in I_0(P)} Q(ij) \log [\alpha Q(ij)] + T(\alpha) + R(\alpha) \end{aligned} \tag{28}$$

[22]The formal structure of this objective function is closely related to the 'cumulative cost' functions employed in certain programming formations of stochastic network equilibria [as for example in Smith (1988)]. This relation is discussed further in the concluding section of Part II.

for appropriately defined bounded functions $T(\alpha)$ and $R(\alpha)$ on [0, 1]. Hence it follows from the positivity of **Q** and boundedness of B, T, and R that the derivative in (27) approaches minus infinity as α approaches zero. But since this implies that Z *decreases* in the feasible direction of movement from **P** toward **Q**, it follows that **P** cannot be a local minimum for $\mathscr{P}[\mathbf{a}, \mathbf{f}]$. Hence all analysis of local minima for $\mathscr{P}[\mathbf{a}, \mathbf{f}]$ can be restricted to $\mathbb{P}^{+}_{q\times q}$.

Observe next that all local minima of $\mathscr{P}[\mathbf{a}, \mathbf{f}]$ must satisfy the Karush-Kuhn-Tucker (*KKT*) conditions. Hence, if each distribution, $\mathbf{P} \in \mathbb{P}^{+}_{q\times q}$, satisfying the *KKT*-conditions for $\mathscr{P}[\mathbf{a}, \mathbf{f}]$ is designated as a *KKT-point,* and if the set of such points is denoted by $KKT[\mathbf{a}, \mathbf{f}]$, then our main result is to show that:

THEOREM 4: Programming Equivalence. *For any gravity-type transition function,* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, *the steady-state flow distributions for* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$ *are equivalent to the KKT-points for* $\mathscr{P}[\mathbf{a}, \mathbf{f}]$, *that is*

$$SFD[\mathbf{a}, \mathbf{f}] = KKT[\mathbf{a}, \mathbf{f}] \tag{29}$$

Proof. By Theorem 2 it suffices to show that the *KKT*-points for $\mathscr{P}[\mathbf{a}, \mathbf{f}]$ are precisely the distributions, $\mathbf{P} \in \mathbb{P}^{+}_{q\times q}$, satisfying (19). To do so, let the Lagrangian function for $\mathscr{P}[\mathbf{a}, \mathbf{f}]$ be given by

$$\mathscr{L}(\mathbf{P}, \gamma) = Z(\mathbf{P}) + \gamma\left[1 - \sum_{ij} P(ij)\right] \tag{30}$$

so that the partial derivative, $\nabla_{ij}\mathscr{L}$, of $\mathscr{L}$ with respect to $P(ij)$ takes the form

$$\nabla_{ij}\mathscr{L}(\mathbf{P}, \gamma) = 1 + \log P(ij) - \log f_{ij} - \log a_i[P(i\cdot)] - \log a_j[P(\cdot j)] - \gamma \tag{31}$$

Now consider any *KKT*-point, $\mathbf{P} \in KKT[\mathbf{a}, \mathbf{f}] \subseteq \mathbb{P}^{+}_{q\times q}$, and observe that by definition there is some scalar, $\gamma \in \mathbb{R}$, such that the pair, $(\mathbf{P}, \gamma)$, satisfies the *KKT*-conditions for $\mathscr{P}[\mathbf{a}, \mathbf{f}]$. In particular, the positivity of **P** together with the *KKT*-condition, $\nabla_{ij}\mathscr{L}(\mathbf{P}, \gamma) \cdot P(ij) = 0$, implies that $\nabla_{ij}\mathscr{L}(\mathbf{P}, \gamma) = 0$ for all $ij \in I \times I$. But by setting (31) equal to zero and solving for $P(ij)$ we obtain

$$\log P(ij) = (\gamma - 1) + \log a_i[P(i\cdot)] + \log a_j[P(\cdot j)] + \log f_{ij} \tag{32}$$

Finally, exponentiating both sides and setting $\mu = \exp[\gamma - 1] > 0$, we see that

$$P(ij) = \mu a_i[P(i\cdot)]a_j[P(\cdot j)]f_{ij}, \qquad ij \in I \times I$$

and hence that $\mathbf{P} \in SFD[\mathbf{a}, \mathbf{f}]$. Thus $KKT[\mathbf{a}, \mathbf{f}] \subseteq SFD[\mathbf{a}, \mathbf{f}]$. To establish the converse, choose any $\mathbf{P} \in SFD[\mathbf{a}, \mathbf{f}]$ and for μ in (19) set $\gamma = 1 + \log(\mu)$. Then it follows at once that the pair $(\mathbf{P}, \gamma)$ satisfies (38), which together with the positivity of **P** is easily seen to imply that $(\mathbf{P}, \gamma)$ satisfies all *KKT*-conditions for $\mathscr{P}[\mathbf{a}, \mathbf{f}]$. Hence **P** is a *KKT*-point for $\mathscr{P}[\mathbf{a}, \mathbf{f}]$, and it follows that $SFD[\mathbf{a}, \mathbf{f}] \subseteq KKT[\mathbf{a}, \mathbf{f}]$. ■

As an important consequence of this result, we obtain the following sufficient condition for uniqueness of steady-state flow distributions. In particular, recalling Definition (ii), we now show that:

THEOREM 5: Uniqueness of Steady-State Flows. *If* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$ *is a gravity-type transition function with pure congestion effects, then there exists a unique steady-state flow distribution for* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, *that is,* $|SFD[\mathbf{a}, \mathbf{f}]| = 1$

Proof. By Theorem 4 it suffices to show that there is a unique *KKT*-point for $\mathscr{P}[\mathbf{a}, \mathbf{f}]$. But since the constraint set $\mathbb{P}_{q\times q}$ is convex and compact, it is enough to show that the objective function in (26) is strictly convex on $\mathbb{P}_{q\times q}$. Moreover, since the divergence function, $D(\cdot, \mathbf{f})$, is well known to be strictly convex (Kullback, 1968), it remains only to show that the last two terms define convex functions on $\mathbb{P}_{q\times q}$. Finally, since each term is a sum of functions of the form, $\Sigma_i \phi_i[L_i(\mathbf{P})]$, where the L_i's are linear, it suffices to establish convexity of the ϕ_i's. To do so, observe that each ϕ_i is of the form, $\phi_i(z) = \int_0^z g_i(x)\,dx$, for a continuous function, g_i. Hence, each ϕ_i will be convex whenever its derivative, $g_i(z)$, is nondecreasing. But since each g_i is of the form, $g_i(x) = -\log[a_i(x)]$, and since each attraction function, a_i, is hypothesized to be nonincreasing, we may conclude that g_i is nondecreasing, and thus that each ϕ_i is convex. ■

As a direct consequence of this result, we obtain the following uniqueness condition for steady states of gravity-type transition functions:

THEOREM 6: Uniqueness of Steady States. *If* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$ *is a gravity-type transition function with pure congestion effects, then there exists a unique steady state for* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, *that is,* $|S(\mathbf{M}^{\mathbf{a},\mathbf{f}})| = 1$

Proof. By Proposition it is enough to show that there cannot be more than one steady state. Hence suppose that there exist distinct steady states, $\mathbf{p}^1$ and $\mathbf{p}^2$, for some gravity-type transition function, $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, with attraction profile, $\mathbf{a}$, and accessibility matrix, $\mathbf{f}$. Then defining the flow distributions, $\mathbf{P}^1$ and $\mathbf{P}^2$, by

$$P^\alpha(ij) = \mu_\alpha a_i(p_i^\alpha) a_j(p_j^\alpha) f_{ij}, \qquad ij \in I \times I, \qquad \alpha = 1, 2$$

it follows at once from Theorem 3 that $\mathbf{P}^1$ and $\mathbf{P}^2$ must be distinct elements of $SFD[\mathbf{a}, \mathbf{f}]$. But since this together with Theorem 4 is seen to contradict Theorem 5, we may conclude that no such $\mathbf{p}^1$ and $\mathbf{p}^2$ can exist. ■

Hence in cases where higher population densities always induce higher levels of undesirable congestion, there must be a unique steady state for the system. Even when congestion effects hold only for large population levels, this uniqueness property implies that there can be at most one 'fully congested' steady state. In particular, if the system population level, N, is sufficiently large to ensure that each region j is congested for all population levels exceeding some critical fraction, $\bar{p}_j$ of N, then it may be postulated that attraction profile,

a, exhibits *eventual congestion effects* in the sense that each a_j is nonincreasing for all $p_j \geq \bar{p}_j$. Hence if the vector of critical levels is denoted by $\bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_q)$, and if each distribution, $\mathbf{p} \in \mathbb{P}_q$, with $\mathbf{p} \geq \bar{\mathbf{p}}$ is said to be a *fully congested* distribution, then we have the following additional consequence of Theorem 5:

THEOREM 7: Congested Steady States. *If* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$ *is a gravity-type transition function with eventual congestion effects, then there is at most one fully congested steady state for* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$.

Proof. If the derivative of a_j is denoted by a_j', then since a_j is nonincreasing on the closed half line $[\bar{p}_j, \infty)$, it follows from the continuous differentiability of a_j that $a_j'(\bar{p}_j) \leq 0$. Hence defining the function, $\bar{a}_j$: $\mathbb{R}_+ \rightarrow \mathbb{R}_{++}$, by $\bar{a}_j(p_j) = a_j(\bar{p}_j) - a_j'(\bar{p}_j)(\bar{p}_j - p_j)$ for $p_j \in [0, \bar{p}_j)$ and $\bar{a}_j(\bar{p}_j) = a_j(p_j)$ for $p_j \in [\bar{p}_j, \infty)$, it may readily be seen that $\bar{a}_j$ is continuously differentiable and nonincreasing on $\mathbb{R}_+$. Hence by Theorem (26) there exists a unique steady state for the gravity-type transition function, $\mathbf{M}^{\bar{a},\mathbf{f}}$, with attraction profile, $\bar{\mathbf{a}} = (\bar{a}_1, \ldots, \bar{a}_q)$. But since $\mathbf{M}^{\bar{\mathbf{a}},\mathbf{f}}(\mathbf{p}) = \mathbf{M}^{\mathbf{a},\mathbf{f}}(\mathbf{p})$ for all $\mathbf{p} \geq \bar{\mathbf{p}}$, it follows that every steady state, $\mathbf{p}$ for $\mathbf{M}^{\bar{\mathbf{a}},\mathbf{f}}$with $\mathbf{p} \geq \bar{\mathbf{p}}$ is automatically a steady state for $\mathbf{M}^{\mathbf{a},\mathbf{f}}$. Hence there can be at most one such steady state. ■[23]

Note that this does *not* imply the existence of fully congested steady states. (In particular, there exist no fully congested distributions whatsoever unless $\Sigma_j \bar{p}_j \leq 1$.) Conditions for existence of such steady states will be developed in a subsequent paper.

5. STABILITY OF INTERACTIVE MARKOV CHAINS

Given these uniqueness properties of steady states, we turn now to the important question of convergence to steady states. The task of this final section is to motivate the central concepts to be employed in the stability analysis of Part II. To do so, it is appropriate to begin with an example which shows that even when steady states are unique, the sequence of population states defined by (3) need not converge at all.

Example 3: Failure of Convergence

For $q = 3$ consider a pure congestion case with attraction functions of the form, $a_j(p_j) = \alpha_j \exp(-\beta_j p_j)$, with $\alpha_1 = 20$, $\alpha_2 = 15$, $\alpha_3 = 5$, $\beta_1 = 14$, $\beta_2 = 15$, $\beta_3 = 6$. If accessibilities are again given by $f_{ij} = \exp(-c_{ij})$ with migration costs $c_{ii} = 0$ and $c_{ij} = 5$ for all distinct i, $j = 1, 2, 3$, then the unique steady state for this case is given approximately by the point, $\mathbf{s} = (0.299, 0.265, 0.436)$, shown in Figure 2. Here the initial condition, $\mathbf{p}^0 = (0.90, 0.01, 0.09)$, in (3) generates a sequence $(\mathbf{p}^0, \mathbf{p}^1, \mathbf{p}^2, \ldots)$ which converges to a two-point limit cycle defined by the points

[23]As mentioned in footnote 13 above, this proof can easily be modified to cover all *continuous* attraction functions.

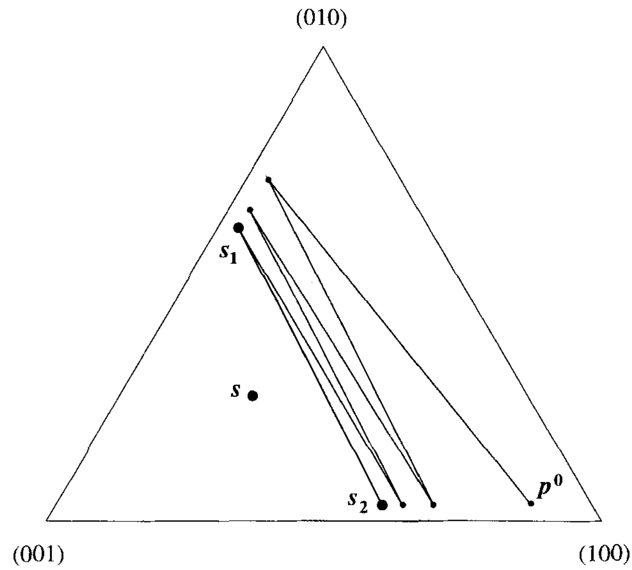

FIGURE 2: Failure of Convergence.

$\mathbf{s}_1 = (0.007, 0.691, 0.302)$ and $\mathbf{s}_2 = (0.705, 0.003, 0.292)$. Hence the sequence never approaches the steady state, **s**. Moreover, since it is clear from Figure 2 that **s** is not on the line joining $\mathbf{s}_1$ and $\mathbf{s}_2$ it also follows that no 'averaging' of this sequence will ever approach **s**. ■

Interactive Markov Processes

An examination of this example shows that failure of convergence results from an 'over adjustment' by migrators in each period.[24] In particular, movement in and out of region 2 (shown by vertical displacement in the Figure) is seen to overshoot the mark in each period. One can gain insight into this type of behavior by focusing on the relevant notion of a 'period' in such models. From an information-theoretic viewpoint, it is implicitly assumed that migration decisions in any given period are based on knowledge about the population distribution in the *previous* period. Hence if each period were to represent a decade, then such migration decisions would presumably be based on information as much as ten years old. But if information flows were indeed this slow, then it is not hard to see how such migration decisions could lead to oscillatory behavior (as illustrated by Example 3, where migration into the least populated region persists long after it has become the most populated one). However, if

[24]Illustrations are given in Dendrinos and Sonis (1990), where such 'over-adjustments' in fact lead to chaotic behavior.

information updating is relatively fast in comparison to the average frequency of migration decisions, then one can expect to observe a much 'smoother' adjustment process. Hence, although the behavioral effects of information lags are not without interest, it is implicitly assumed in most models that periods are sufficiently short to allow such effects to be discounted.

To give meaning to the notion of 'sufficiently short' periods, it is essential to introduce *time,* explicitly. Hence we now reformulate the basic adjustment process in (2) as follows. If t denote a point in real time, and if Δ denotes a positive time interval, then each time period can be represented by the time points t and $t + \Delta$. Here one could in principle simply replace $t + 1$ in (2) by $t + \Delta$. But this brings us to a second implicit assumption in (2), namely that *all* individuals make migration decisions in each period. Although the diagonal components, $M_{ii}(\cdot)$, in the transition function can in principle be taken to include those individuals who do not consider migrating during a given period, the introduction of period length, Δ, makes it clear that $M_{ii}(\cdot)$ must then also depend on Δ. To avoid this complication, it is convenient to interpret the transition function as pertaining only to *migration decisions,* so that $M_{ii}(\cdot)$ represents the fraction of migration decisions that involve only local moves within a given region (or decisions not to migrate at all). This convention allows an explicit separation to be made between migration decision makers and all others in the population. In particular, one may postulate that the fraction, $\alpha(\Delta)$, of individuals making migration decisions within any time interval Δ is an explicit nondecreasing function, $\alpha\colon \mathbb{P}_{++} \rightarrow (0, 1]$, which we now designate as the *participation function* for the population.[25] With this convention, the appropriate adjustment process now takes the form

$$(33) \quad p_j^{t+\Delta} = [1 - \alpha(\Delta)]p_j^t + \alpha(\Delta) \sum_i p_i^t M_{ij}(\mathbf{p^t}), \qquad j \in I, \qquad t/\Delta \in \mathbb{Z}_+$$

where for each region j the first term on the right-hand side represents the fraction of individuals in region j who make no migration decisions in time interval $(t, t + \Delta)$ and hence remain in j, and where the second term represents the fractions of decision makers in all regions who choose to migrate to j (including decision makers in j who choose to stay in j). If (33) is written in vector form as

$$(34) \qquad \mathbf{p}^{t+\Delta} = [1 - \alpha(\Delta)]\mathbf{p}^t + \alpha(\Delta)\mathbf{p}^t\mathbf{M}(\mathbf{p}^t), \qquad t/\Delta \in \mathbb{Z}_+$$

[25]From a probabilistic viewpoint, the population fraction $\alpha(\Delta)$ can also be interpreted as the probability that any randomly sampled individual in the population decides to reconsider (review) his current locational choice in time interval Δ. The present formulation represents a simplified version of the model in De Palma and Lefevre (1983) where the fraction of migration decision makers is allowed to depend on the current population distribution, $\mathbf{p}^t$, as well. A similar model of this type is developed in Boots and Kanaroglou (1988).

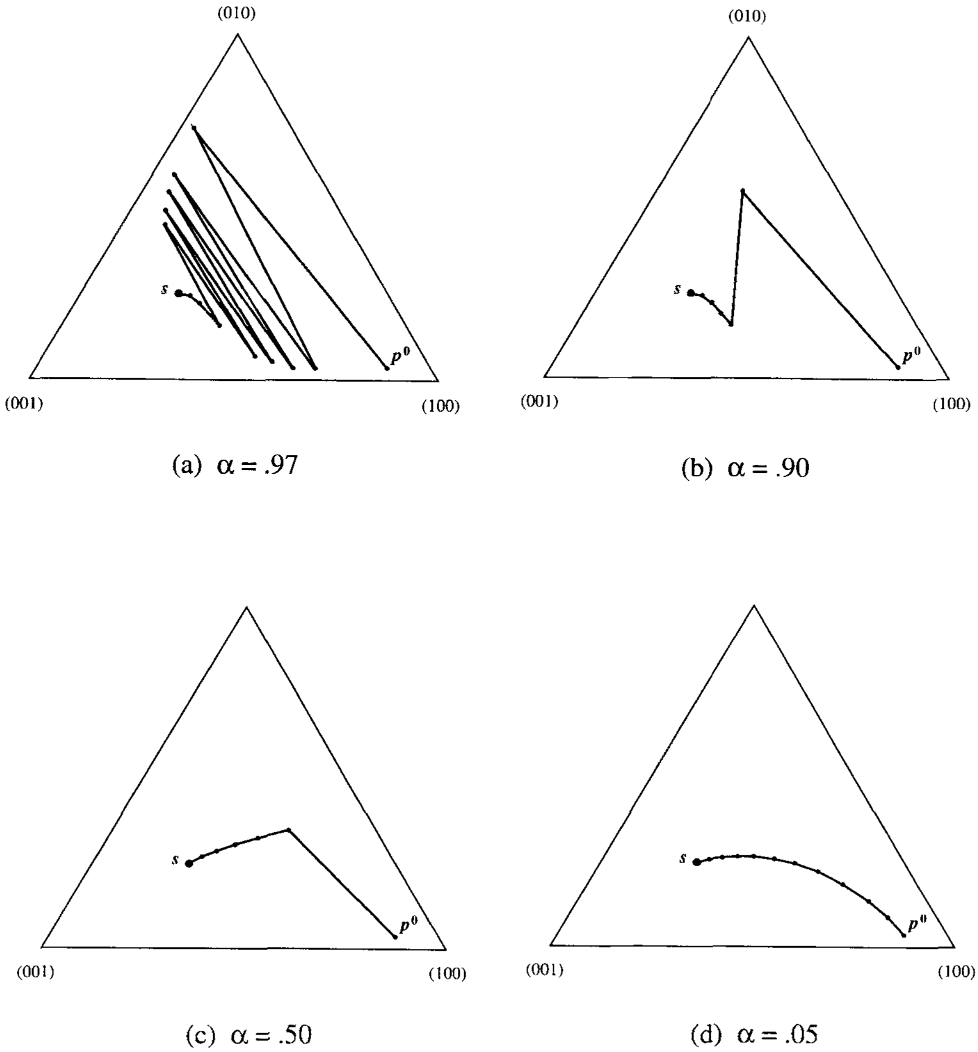

FIGURE 3. Time-Dependent Smoothing.

then (2) is seen to be the special case of (34) in which $\Delta = 1$ and $\alpha(1) = 1$. However, (34) is most usefully regarded as simply a *reparameterization* of (2). For if the $q \times q$ identity matrix is denoted by $\mathbf{I}$, then expression (33) is seen to be an instance of (2) with $\mathbf{M}(\cdot)$ replaced by a new transition function, $[1 - \alpha(\Delta)]\mathbf{I} + \alpha(\Delta)\mathbf{M}(\cdot)$, in which the diagonal elements are singled out for special treatment. In terms of this parameterization, we shall henceforth characterize each *interactive Markov chain* as a triple $(\mathbf{M}, \alpha, \Delta)$ with *transition function,* $\mathbf{M}(\cdot)$, *participation function,* $\alpha(\cdot)$, and *adjustment period,* Δ. Note in particular, that for all $\mathbf{p} \in \mathbb{P}_q$, $\Delta > 0$, and participation functions, $\alpha(\cdot)$

$$\mathbf{p} = \mathbf{pM(p)} \Leftrightarrow \alpha(\Delta)\mathbf{p} = \alpha(\Delta)\mathbf{pM(p)}$$

$$\Leftrightarrow [1 - \alpha(\Delta)]\mathbf{p} + \alpha(\Delta)\mathbf{p} = [1 - \alpha(\Delta)]\mathbf{p} + \alpha(\Delta)\mathbf{pM(p)}$$

$$\Leftrightarrow \mathbf{p} = \mathbf{p}\{[1 - \alpha(\Delta)]I + \alpha(\Delta)\mathbf{M(p)}\}$$

so that the steady states for $(\mathbf{M}, \alpha, \Delta)$ are seen to depend only on $\mathbf{M}$, and not on the choice of α or Δ. Hence for any given transition function, $\mathbf{M}$, the parameterizations, $(\mathbf{M}, \alpha, \Delta)$, yield a family of interactive Markov chains that differ only with respect to their *dynamical* behavior, and not their steady-state behavior. These parameterizations will form the basis for all of our subsequent dynamical analyses.[26]

For any given interactive Markov chain, $(\mathbf{M}, \alpha, \Delta)$, observe that as the adjustment period Δ becomes small the fraction $\alpha(\Delta)$ of individuals making migration decisions should also become small, so it is natural to assume that $\lim_{\Delta \to 0} \alpha(\Delta) = 0$. Under this condition, the key feature of such processes is that adjustment behavior becomes 'smoother' as $\Delta \to 0$. This is illustrated by the modifications of Example 3 shown in Figure 3. Here we see in Figure 3a that when the value of $\alpha = \alpha(\Delta)$ decreases from $\alpha = 1$ to $\alpha = 0.97$ the oscillatory behavior in Figure 2 eventually diminishes and the process converges to $\mathbf{s}$. Hence even when the fraction of nondecision makers in the population is as low as 3 percent, the adjustment process is damped enough (in this example) to ensure eventual convergence to a steady state. Moreover, Figures 3b, 3c, and 3d show that the oscillatory behavior of the process disappears altogether as α decreases and, in particular, that the process is quite smooth when the fraction of migration decision makers is no greater than 5 percent ($\alpha \leq .05$).

Given these observations, our main objective is to study the dynamical behavior of interactive Markov chains $(\mathbf{M}, \alpha, \Delta)$ as Δ becomes small. To do so, observe that by rearranging terms and dividing by Δ, (33) is equivalent to

$$(35) \qquad \frac{p_j^{t+\Delta} - p_j^t}{\Delta} = \frac{\alpha(\Delta)}{\Delta}\left[\sum_i p_i^t M_{ij}(\mathbf{p}^t) - p_j^t\right], \qquad j \in I, \qquad t/\Delta \in \mathbb{Z}_+$$

where the left-hand side is seen to approximate the time derivative, $\dot{p}_j^t$, of p_j as $\Delta \to 0$. In particular, if *participation rate,* $\alpha(\Delta)/\Delta$, (i.e., participation per unit of time) has a well defined limit, say $\lim_{\Delta \to 0} (\alpha(\Delta)/\Delta] = \lambda > 0$, then (35) is approximated for small Δ by the (autonomous) differential equation system[27]

$$(36) \qquad \dot{p}_j(t) = \lambda\left\{\sum_i p_i(t) M_{ij}[\mathbf{p}(t)] - p_j(t)\right\}, \qquad j \in I, \qquad t \in \mathbb{R}_+$$

[26]We return to the explicit analysis of these parameterizations in Section 4 of Part II.

[27]It should be noted at this point that our assumption of *continuously differentiable* attraction functions is needed only to ensure this uniqueness property. Hence the results of this section (as well as most of those in Part II) can be extended to all attraction functions which are at least locally Lipschitzian.

where λ denotes the *limiting participation rate* for the given population. If this differential equation system is written in matrix form as

$$\text{(37)} \qquad \dot{\mathbf{p}}(t) = \lambda\{\mathbf{p}(t)\mathbf{M}[\mathbf{p}(t)] - \mathbf{p}(t)\}, \qquad t \in \mathbb{R}_+$$

then we begin by focusing on the convergence behavior of these limiting processes. Our ultimate objective is to show that certain of these convergence properties are inherited by those associated interactive Markov chains, $(\mathbf{M}, \alpha, \Delta)$, with Δ sufficiently small. In this way, one can give concrete meaning to the above notion of 'sufficiently short' periods.

To analyze these limiting processes, we begin by recalling from the Picard-Lindelöf Theorem (for example Theorem 3.1 in Hale, 1980) that the continuous differentiability of $\mathbf{M}$ implies the existence of a unique (continuously differentiable) solution to (37) for each initial condition, $\mathbf{p}^0 \in \mathbb{P}_q$.[28] Of particular importance for our present purposes is that each solution starting in $\mathbb{P}_q$ must stay in $\mathbb{P}_q$. More formally, if a set $\Omega \subseteq \mathbb{R}^q$ is designated as an *invariant* set for the differential equation system (37) iff for every solution, $\mathbf{p}: \mathbb{R}_+ \rightarrow \mathbb{R}^q$, with $\mathbf{p}(0) \in \Omega$ it is true that $\mathbf{p}(t) \in \Omega$ for all $t \in \mathbb{R}_+$, then we have

PROPOSITION 2: Invariant Sets: *The set of distributions $\mathbb{P}_q$ is always an invariant set for* (32). *In addition, if* $\mathbf{M}$ *is positive then* $\mathbb{P}_q^+$ *also an invariant set.*

Proof. (*i*) If for each $\Delta > 0$ with $\Delta\lambda < 1$ we consider the following difference equation defined for all $n \in \mathbb{Z}_+$ by

$$\mathbf{p}_\Delta^{n+1} = (1 - \Delta\lambda)\mathbf{p}_\Delta^n + \Delta\lambda\mathbf{p}_\Delta^n\mathbf{M}(\mathbf{p}_\Delta^n)$$

and construct a standard 'Euler approximation' $(\mathbf{p}_\Delta^t)$ of (37) defined for all $t \in \mathbb{R}_+$ by the linear interpolation,

$$\mathbf{p}_\Delta^t = \frac{n\Delta - t}{\Delta}\mathbf{p}_\Delta^n + \frac{n\Delta + \Delta - t}{\Delta}\mathbf{p}_\Delta^{n+1}, \qquad t \in [n\Delta, n\Delta + \Delta]$$

then the convexity of $\mathbb{P}_q$ implies that for each choice of initial condition, $\mathbf{p}_\Delta^0 \equiv \mathbf{p}^0 \in \mathbb{P}_q$, this family of functions $(\mathbf{p}_\Delta: \Delta\lambda < 1)$ lies entirely in $\mathbb{P}_q$, and in particular is uniformly bounded. Moreover the continuous differentiability of $\mathbf{M}$ also implies that this family is equicontinuous on $\mathbb{R}_+$. Hence it follows from the compactness of $\mathbb{P}_q$ together with an application of the Arzelà-Ascoli Theorem (Royden, 1968, Theorem 33, p. 179) that $(\mathbf{p}_\Delta)$ converges (uniformly on each bounded time interval $[0, T]$) to a continuous bounded function, $\mathbf{p}: \mathbb{R}_+ \rightarrow \mathbb{P}_q$, which can be shown (by an application of the Lebesgue Dominated Convergence Theorem

[28]Note also that this differential equation system is an instance of the more general family of 'master equations' studied by Haag and Weidlich (1988b) and others. In particular, if we substitute the identity, $\Sigma_{i \neq j} M_{ji}\,[\mathbf{p}(t)] = 1 - M_{jj}[\mathbf{p}(t)]$, into the right-hand side of (36), then this equation is seen to have the equivalent form, $\dot{p}_j(t) = \lambda[\Sigma_{i \neq j}\, p_i(t)M_{ij}[\mathbf{p}(t)] - \Sigma_{i \neq j}\, p_j(t)M_{ji}[\mathbf{p}(t)]]$, which is an instance of the 'pure migratory' equation in Haag and Weidlich (1988b, p. 30).

(Royden, 1968, Theorem 16, p. 229) to satisfy (37) with initial condition $\mathbf{p}^0$ (see for example in Kushner and Clark, 1978, pp. 20–21). But by the Picard-Lindelöf Theorem this solution is unique, and hence $\mathbb{P}_q$ is an invariant set for (37).
(*ii*) From the continuity of **M** and compactness of $\mathbb{P}_q$ it follows for each $ij \in I \times I$, the component function M_{ij} achieves a minimum in $\mathbb{P}_q$, which together with the positivity of M_{ij} implies that $\min_{\mathbf{p}\in\mathbb{P}_q} M_{ij}(\mathbf{p}) > 0$. Hence letting $\alpha = \min_{ij} \{\min_{\mathbf{p}\in\mathbb{P}_q} M_{ij}(\mathbf{p})\} > 0$, it follows from (36) that

$$(38) \qquad \dot{p}_j(t) \geq \lambda\left\{\sum_i p_i(t)\alpha - p_j(t)\right\} = \lambda\{\alpha - p_j(t)\}, \qquad j \in I, \qquad t \in \mathbb{R}_+$$

Thus, if for any solution of (37) with $\mathbf{p}(0) \in \mathbb{P}_q^+$ we let $\epsilon_j = \min\{p_j(0), \alpha/2\} > 0$ for each $j \in I$, then it follows from (38) that $\dot{p}_j(t) > 0$ whenever $p_j(t) = \epsilon_j$. But since $p_j(0) \geq \epsilon_j$ by definition, we may then conclude that $p_j(t) \geq \epsilon_j > 0$ for all $t \in \mathbb{R}_+$, and hence that $\mathbf{p}(t) \in \mathbb{R}^q_{++}$ for all $t \in \mathbb{R}_+$. Finally since (*i*) implies that $\mathbf{p}(t) \in \mathbb{P}_q$ we must have $\mathbf{p}(t) \in \mathbb{P}_q \cap \mathbb{R}^q_{++} = \mathbb{P}_q^+$ for all t whenever $\mathbf{p}(0) \in \mathbb{P}_q^+$. ■

With these observations, we may now formalize this class of limiting processes as follows:

DEFINITION 3: (*i*) *For any transition function,* **M**: $\mathbb{P}_q \to \mathbb{R}_+^{q\times q}$, *the differential equation system in* (37) *is designated as a* interactive Markov process *on I with intensity parameter,* $\lambda > 0$.[29] *Each such process is denoted by the pair* (**M**, λ).
(*ii*) *Every continuously differentiable function,* **p**: $\mathbb{R}_+ \to \mathbb{P}_q$, *with values,* $\mathbf{p}(t)$, *satisfying* (37) *is said to be an* adjustment path *for process* (**M**, λ) *with* starting point, $\mathbf{p}(0) \in \mathbb{P}_q$. *Such paths are also denoted by* $\mathbf{p}(\cdot)$.
(*iii*) *For each gravity-type transition matrix,* $\mathbf{M}^{\mathbf{a},\mathbf{f}}$, *the interactive Markov process* $(\mathbf{M}^{\mathbf{a},\mathbf{f}}, \lambda)$ *is also said to be of* gravity type.
(*iv*) *In addition, if each* a_j *is nonincreasing then* $(\mathbf{M}^{\mathbf{a},\mathbf{f}}, \lambda)$ *is designated as a gravity-type interactive Markov process with* pure congestion effects.

Next observe that since *steady states,* $\mathbf{p} \in \mathbb{P}_q$, for the differential equation system (37) are defined by the condition, $\dot{\mathbf{p}} = 0$, it follows that the steady states for (**M**, λ) are precisely the steady states, $S(\mathbf{M})$, for the transition function, **M**, in (3) above. Hence we shall use the two concepts interchangeably.

Extension to Spatial-Flow Processes

To analyze convergence to these steady states, we again focus on the spatial flows associated with such processes. As in the discrete case, we begin with a space of *flow states,* $ij \in I \times I$, and develop a class of continuous 'spatial-flow processes' on $I \times I$, with flow paths, $\mathbf{P}_t = [P_t(ij): ij \in I \times I]$, where

[29]From a formal viewpoint this non-probabilistic model is more accurately described as a *deterministic version* of a Markov process, as for example in De Palma and Lefevre (1983, Section 4) [In particular, the present model is seen to be an instance of model (11) in De Palma and Lefevre where the intensity rate, λ, corresponds to their constant 'review rate' *R*.] However, the present terminology serves to emphasize the parallel between these continuous-time models and interactive Markov chains.

$P_t(ij)$ denotes the fraction of population in flow state ij at time t, that is, the fraction of population flowing from i to j at time t. But unlike the analysis of Section 4 that involved only steady-state properties, we are now interested in the full range of *adjustment paths* generated by each process. Hence, while it was sufficient in Section 4 to focus on the single the most natural representation of spatial flows (designated there as the *spatial-flow chain* associated with the given interactive Markov chain), it is now important to emphasize that there are *many* different spatial-flow processes which are consistent with the same interactive Markov process.[30] In particular, it will be shown in Part II that for purposes of stability analysis, it is convenient to study a 'symmetric flow version' of gravity-type interactive Markov processes which differs from the natural continuous analog of spatial-flow chains (designated as the 'canonical flow version' in Example (35) below).

To develop the full range of possibilities here, it is convenient to begin with a very general definition of 'flow processes' and then specialize to those 'spatial-flow processes' of interest. Hence, if $\mathbb{P}_{q\times q}$ in (11) again denotes the set of possible flow distributions on $I \times I$, then we now consider the general class of processes governed by differential equations of the form

$$\dot{P}_t = \sigma[\mathbf{\Phi}(\mathbf{P}_t) - \mathbf{P}_t], \qquad t \in \mathbb{R}_+ \tag{39}$$

for some continuously differentiable function, $\mathbf{\Phi}$: $\mathbb{P}_{q\times q} \rightarrow \mathbb{P}_{q\times q}$, and positive scalar, σ. The properties of system (39) closely parallel those of (37). First of all the steady states for this system are again defined by the condition that $\dot{\mathbf{P}} = 0$, so that a flow distribution, $\mathbf{P} \in \mathbb{P}_{q\times q}$, is a *steady state* for (39) iff $\mathbf{\Phi}(\mathbf{P}) = \mathbf{P}$. If the set of such steady states is denoted by

$$S(\mathbf{\Phi}) = \{\mathbf{P} \in \mathbb{P}_{q\times q}: \mathbf{\Phi}(\mathbf{P}) = \mathbf{P}\} \tag{40}$$

then the argument in Proposition 1 again shows that there is always at least one fixed point, $\mathbf{\Phi}(\mathbf{P}) = \mathbf{P}$, for $\mathbf{\Phi}$ and hence that $S(\mathbf{\Phi}) \neq \emptyset$. Next observe from the continuous differentiability of $\mathbf{\Phi}$ that (again by the Picard-Lindelöf Theorem) there exists a unique continuously differentiable matrix-valued function, $\mathbf{P}$: $\mathbb{R}_+ \rightarrow \mathbb{R}^{q\times q}$, satisfying (39) for any choice of initial conditions, $\mathbf{P}_0 \in \mathbb{P}_{q\times q}$. Moreover, if we now designate a set, $\Omega \subseteq \mathbb{R}^{q\times q}$, as an *invariant set* for (39) iff for every solution, $\mathbf{P}$: $\mathbb{R}_+ \rightarrow \mathbb{R}^{q\times q}$, with $\mathbf{P}_0 \in \Omega$ it is true that $\mathbf{P}_t \in \Omega$ for all $t \in \mathbb{R}_+$, then we have the following parallel to Proposition 2:

PROPOSITION 3: Invariant Flow Sets: *The set of distributions* $\mathbb{P}_{q\times q}$ *is always an invariant set for* (39)

[30]This is easily seen in the discrete case by observing that each interative Markov chain corresponds to the *marginal* distributions of its associated spatial-flow chain, and that in general there are many different joint distributions with the same marginal distributions. An explicit example for the case of interactive Markov chains is given in Part II (footnote 1).

Proof. If for each $\Delta > 0$ with $\Delta\sigma < 1$ we now consider the difference equation defined for all $n \in \mathbb{Z}_+$ by

$$\mathbf{P}_{n+1}^{\Delta} = (1 - \Delta\sigma)\mathbf{P}_n^{\Delta} + \Delta\sigma\mathbf{\Phi}(\mathbf{P}_n^{\Delta})$$

together with the Euler approximations $(\mathbf{P}_t^{\Delta})$ of (39) defined for all $t \in \mathbb{R}_+$ by

$$\mathbf{P}_t^{\Delta} = \frac{n\Delta - t}{\Delta}\mathbf{P}_n^{\Delta} + \frac{n\Delta + \Delta - t}{\Delta}\mathbf{p}_{n+1}^{\Delta}, \qquad t \in [n\Delta, n\Delta + \Delta]$$

then the convexity of $\mathbb{P}_{q\times q}$ again implies that for each choice of initial condition, $\mathbf{P}_0^{\Delta} \equiv \mathbf{P}_0 \in \mathbb{P}_{q\times q}$, the family of functions $(\mathbf{P}^{\Delta}\colon \Delta\sigma < 1)$ lies entirely in $\mathbb{P}_{q\times q}$. Hence the same argument as in Proposition 2 shows that $(\mathbf{P}^{\Delta})$ converges to the unique solution of (37) with initial condition $\mathbf{P}_0$, and thus that $\mathbb{P}_{q\times q}$ is an invariant set for (39). ■

Hence solutions to (39) starting from any initial flow distribution $\mathbf{P}_0 \in \mathbb{P}_{q\times q}$ always yield flow distributions, $\mathbf{P}_t \in \mathbb{P}_{q\times q}$, at each point in time, and thus can potentially describe meaningful 'flow adjustments' in this sense. However, it is important to note that if such a flow interpretation is to make physical sense, then every *steady state* of the system must satisfy the additional 'balancing' constraint that total flow into each region i be exactly the same as total flow out of region i. Hence, if for each distribution, $\mathbf{P} = [P(ij)\colon ij \in I \times I]$, we again denote the *row marginals* by $P(i\cdot) = \Sigma_j P(ij)$, $i \in I$, and *column marginals* by $P(\cdot j) = \Sigma_i P(ij)$, $j \in I$, then (as a parallel to (18) for the discrete case) we require that each steady state, $\mathbf{P} \in S(\mathbf{\Phi})$, satisfy the *flow-balance condition:*

(41) $$P(i\cdot) = P(\cdot i), \qquad i \in I$$

These observations may be summarized as follows:

DEFINITION 4. (*i*) *For any continuously differentiable function,* $\mathbf{\Phi}\colon \mathbb{P}_{q\times q} \to \mathbb{P}_{q\times q}$, *and positive scalar,* σ, *the differential equation system in* (39) *is designated as a* flow process, $(\mathbf{\Phi}, \sigma)$, *on* $I \times I$ *iff every steady state,* $\mathbf{P} \in S(\mathbf{\Phi})$, *satisfies flow-balance condition* (41).
(*ii*) *Each continuously differentiable matrix-valued function,* $\mathbf{P}\colon \mathbb{R}_+ \to \mathbb{P}_{q\times q}$, satisfying (39) *is designated as a* flow-adjustment path *for* $(\mathbf{\Phi}, \sigma)$ *with* starting point, $\mathbf{P}_0 \in \mathbb{P}_{q\times q}$. *Such flow-adjustment paths are also denoted by* $(\mathbf{P}_t)$.

Within this general framework, we now consider those flow processes on $I \times I$ which 'generate' interactive Markov processes on I. To do so, observe first that for any flow-adjustment path, $(\mathbf{P}_t)$, both of the corresponding marginal distributions, $\mathbf{P}_t(I\cdot)$ and $\mathbf{P}_t(\cdot I)$, generate paths in $\mathbb{P}_q$ that can in principle represent adjustment paths for interactive Markov processes. Hence, as a parallel to the discrete case, we again focus on the 'column' marginals, $\mathbf{P}_t(\cdot I)$, and consider those flow processes with column marginals corresponding to a given interactive Markov process. To do so, observe first that since by definition the identity, $P_t(\cdot j) = \Sigma_i P_t\,(ij)$, implies the corresponding identity, $\dot{P}_t(\cdot j) = \Sigma_i$

$\dot{P}_t(ij)$, it follows that the column marginals of a given flow process correspond to the interactive Markov process in (36) iff the following *marginal consistency condition* is satisfied:

$$(42) \qquad \dot{P}_t(\cdot j) = \lambda\left\{\sum_i P_t(\cdot i)M_{ij}[\mathbf{P}_t(\cdot I)] - P_t(\cdot j)\right\}, \qquad j \in I, \qquad t \in \mathbb{R}_+$$

However, there are certain 'trivial' flow processes which automatically satisfy this condition, as illustrated by the following example.

Example 4: Product-Flow Processes

For any interactive Markov process $(\mathbf{M}, \lambda)$ on I with adjustment paths $\mathbf{p}(\cdot)$ in $\mathbb{P}_q$, we may construct a unique flow process with 'flow-adjustment' paths given by the simple 'product' condition, $P_t(ij) \equiv p_i(t)p_j(t)$. Observe that by definition the identity, $\mathbf{P}_t(I\cdot) \equiv \mathbf{p}(t) \equiv \mathbf{P}_t(\cdot I)$, implies that each steady state of this process automatically satisfies (41) and hence defines a 'flow process.' Moreover, this identity also implies that the column marginals for each flow-adjustment path agree identically with the adjustment paths for $(\mathbf{M}, \lambda)$, and hence that (42) is automatically satisfied. To construct the differential equation specification of this process observe that we may differentiate the identity, $P_t(ij) \equiv p_i(t)p_j(t)$, and employ (36) to obtain

$$\begin{aligned}\dot{P}_t(ij) = \dot{p}_i(t)p_j(t) + p_i(t)\dot{p}_j(t) &= \lambda\left\{\sum_k p_k(t)M_{ki}[\mathbf{p}(t)] - p_i(t)\right\}p_j(t) \\ &\quad + p_i(t)\lambda\left\{\sum_k p_k(t)M_{kj}[\mathbf{p}(t)] - p_j(t)\right\} \\ &= 2\lambda\left[\sum_k p_k(t)\{\tfrac{1}{2}p_j(t)M_{ki}[\mathbf{p}(t)] + \tfrac{1}{2}p_i(t)M_{kj}[\mathbf{p}(t)]\} - p_i(t)p_j(t)\right]\end{aligned}$$

Hence, by setting $\alpha = 2\lambda$ and $\mathbf{\Phi}_{ij}(\mathbf{P}_t) = \Sigma_k P_t(\cdot k)\{\tfrac{1}{2}P_t(\cdot j)M_{ki}[\mathbf{P}_t(\cdot I)] + \tfrac{1}{2}P_t(\cdot i)M_{kj}[P_t(\cdot I)]\}$ we obtain an explicit *product flow process,* $(\mathbf{\Phi}, \sigma)$, which always generates the interactive Markov process $(\mathbf{M}, \lambda)$. (Note also that this same argument yields a product-flow process which generate *any* given process with paths in $\mathbb{P}_q$.) ■

However, although the steady states, $\mathbf{P}$, of this product flow process always generate the steady states of $(\mathbf{M}, \lambda)$, the individual flow shares in $\mathbf{P}$ do *not* necessarily correspond to those implicit in the transition function, $\mathbf{M}$. In particular, observe that for any steady-state flow distribution, $\mathbf{P}$, the fraction of all flow from i that goes to j is given by $P(ij)/P(i\cdot)$. But by the interpretation of $\mathbf{M}$ it follows that $M_{ij}[\mathbf{P}(\cdot I)]$ must represent the steady-state fraction of population at i who migrates to j (which by (41) is equivalent to $M_{ij}[\mathbf{P}(I\cdot)]$). Hence in order that a given flow process $(\mathbf{\Phi}, \sigma)$ yield meaningful steady-state flows consistent with $(\mathbf{M}, \lambda)$, it is necessary that each steady state for $(\mathbf{\Phi}, \sigma)$

satisfy the following additional *transitional consistency condition:*[31]

$$P(ij)/P(i\cdot) = M_{ij}[\mathbf{P}(I\cdot)], \qquad ij \in I \times I \tag{43}$$

It may be verified by inspection that simple product-flow process in Example 4 generally fails to satisfy this condition. Hence we now focus on those flow processes which satisfy *both* of these consistency conditions:

DEFINITION 5: *For any given interactive Markov process,* $(\mathbf{M}, \lambda)$, *a flow process,* $(\mathbf{\Phi}, \sigma)$, *is designated as a* flow version *of* $(\mathbf{M}, \lambda)$ *iff* $(\mathbf{\Phi}, \sigma)$ *satisfies the marginal consistency condition* (42), *and each steady state,* $\mathbf{P} \in S(\mathbf{\Phi})$, *satisfies the transitional consistency condition* (43).

As a first illustration of these concepts, it is appropriate to develop a continuous analog of the *spatial-flow chains* in (13) and (14) above which shows that there always exists as least one flow version of any interactive Markov process $(\mathbf{M}, \lambda)$:

Example 5: Canonical Flow Version

For any interactive Markov process $(\mathbf{M}, \lambda)$ consider the process $(\mathbf{\Phi}^M, \sigma)$ with $\sigma = \lambda$ and with operator, $\mathbf{\Phi}^M: \mathbb{P}_{q\times q} \rightarrow \mathbb{P}_{q\times q}$, defined for all $\mathbf{P} \in \mathbb{P}_{q\times q}$ by

$$\mathbf{\Phi}^M_{ij}(\mathbf{P}) = P(\cdot i)M_{ij}[\mathbf{P}(\cdot I)], \qquad ij \in I \times I \tag{44}$$

It then follows from (39) and (44) that a distribution, $\mathbf{P} \in \mathbb{P}_{q\times q}$, is a steady state for $(\mathbf{\Phi}^M, \lambda)$, iff

$$P(\cdot i)M_{ij}[\mathbf{P}(\cdot I)] = P(ij), \qquad ij \in I \times I \tag{45}$$

Hence summing (45) with respect to j we see that (41) holds identically, and hence that $(\mathbf{\Phi}^M, \lambda)$ is a flow process. Moreover, by using (41) and dividing both sides of (45) by $P(i\cdot)$, we also see that each steady state for $(\mathbf{\Phi}^M, \lambda)$ satisfies the transitional consistency condition (43). Finally, by substituting (44) into (39) and summing both sides with respect to i, we see that (42) holds, and hence that $(\mathbf{\Phi}^M, \lambda)$ is a flow version of $(\mathbf{M}, \lambda)$, which we now designate as the *canonical flow version* of $(\mathbf{M}, \lambda)$. ■

An important additional example involving 'symmetric' flows will be developed in Part II. There it will be shown that the stability properties of these *symmetric flow versions* of gravity-type interactive Markov processes can be fully analyzed in term of the objective function, Z, in (26). In particular, it will be shown that this objective function yields a global Lyapunov function for such processes on an appropriately defined invariant flow set containing all steady states of the process.

[31]Note also from (41) that this condition is essentially a generalization of condition (16).

REFERENCES

Bartholomew, David J. *Stochastic Models for Social Processes,* 3rd Edition. New York: Wiley.

———. 1985. "Interactive, Threshold and Non-Linear Models for Social Systems," *Journal of Mathematical Sociology,* 11, 43–45.

Boots, Barry N. and Pavlos Kanaroglou. 1988. "Incorporating the Effects of Spatial Structure Indiscrete Choice Models of Migration," *Journal of Regional Science,* 28, 495–510.

Brumelle, Shelby L., and Yigal Gerchak. 1980. "A Stochastic Model Allowing Interaction Among Individuals and Its Behavior for Large Populations," *Journal of Mathematical Sociology,* 7, 73–90.

Conlisk, John. 1976. "Interactive Markov Chains," *Journal of Mathematical Sociology,* 4, 157–185.

———. 1982. "The Law of Supply and Demand as a Law of Markov Chains," *Journal of Economic Theory,* 26, 1–16.

———. 1992. "Stability and Monotonicity for Interactive Markov Chains," *Journal of Mathematical Sociology,* 17, 127–143.

De Palma, Andre and Claude Lefevre. 1983. "Individual Decision-Making in Dynamic Collective Systems," *Journal of Mathematical Sociology,* 9, 103–124.

Dugundji, James. 1966. *Topology.* Boston: Allyn and Bacon.

Haag, Gunther and Wolfgang Weidlich. 1988a. "Concepts of the Dynamic Migration Model," in Haag, Gunther and Wolfgang Weidlich (eds.), *Interregional Migration,* Berlin: Springer-Verlag. 9–20.

———. 1988b. "The Migratory Equations of Motion," in Haag, Gunther and Wolfgang Weidlich (eds.), *Interregional Migration.* Berlin: Springer-Verlag. 32–37.

Hahn, Wolfgang. 1967. *Stability of Motion.* New York: Springer-Verlag.

Hale, Jack K. 1980. *Ordinary Differential Equations.* Malabar, Florida: Krieger Publishing.

Hildenbrand, Werner. 1974. *Core and Equilibria of a Large Economy.* Princeton, N.J.: Princeton University Press.

Howard, Ronald A. 1971. *Dynamic Probablistic Systems. Volume 1: Markov Models.* New York: Wiley.

Kanaroglou, Pavlos, Kao-Lee Liaw and Yorgo Y. Papageorgiou. 1986a. "An Analysis of Migratory Systems: 1. Theory," *Environment and Planning A,* 18, 913–928.

———. 1986b. "An Analysis of Migratory Systems: 2. Operational Framework," *Environment and Planning A,* 18, 1039–1060.

Kullback, Solomon. 1968. *Information Theory and Statistics.* New York: Dover.

Kulkarni, Manjusha S. and Arvind Kumar. 1989. "Interactive Markov Models of 'Progressive' Trends," *Journal of Mathematical Sociology,* 14, 45–65.

Kushner, Harold J. and Dean S. Clark. 1978. *Stochastic Approximation Methods for Constrained and Unconstrained Systems.* New York: Springer-Verlag.

Lehoczky, John P. 1980. "Approximations for Interactive Markov Chains," *Journal of Mathematical Sociology,* 7, 139–157.

Matras, Judah. 1967. "Social Mobility and Social Structure: Some Insights from the Linear Model," *American Sociological Review,* 32, 608–614.

MacLane, Saunders and Garrett Birkhoff. 1967. *Algebra.* New York: MacMillan.

Nijkamp, Peter and Jacques Poot. 1987. "Dynamics of Generalized Spatial Interaction Models," *Regional Science and Urban Economics,* 17, 367–390.

Plane, David A. 1993. "Requiem for the Fixed-Transition-Probability Migrant," *Geographical Analysis,* 25, 211–223.

Poot, Jacques. 1984. Models of New Zealand Migration and Residential Mobility, unpublished Ph.D. dissertation, Victoria University of Wellington, Wellington, New Zealand.

Royden, H. L. 1968. *Real Analysis.* Second Edition. New York: MacMillan.

Smith, Tony E. and Shang-Hsing Hsieh. 1996. "Gravity-Type Interactive Markov Models—Part I: A Programming Formulation of Steady States," Working Paper No. 96-06, Department of Systems, University of Pennsylvania.

———. 1997. "Gravity-Type Interactive Markov Models—Part II: Lyapunov Stability of Steady States," *Journal of Regional Science,* 37,

Sonis, Michael. 1992. "Dynamics of Continuous and Discrete Logit Models," *Socio-Spatial Dynamics,* 3, 35–53.

Weidlich, Wolfgang. 1988. "Solutions of the Master Equation," in Haag, Gunther and Wolfgang Weidlich (eds.), *Interregional Migration,* Berlin: Springer-Verlag, 327–355.