THEORETICAL POPULATION BIOLOGY 7, 84–93 (1975)

Convergence of the Age Structure: Applications of the Projective Metric*

MARTIN GOLUBITSKY

Department of Mathematics, Queens College, Flushing, New York 11367

EMMETT B. KEELER

The Rand Corporation, Santa Monica, California 90406

AND

MICHAEL ROTHSCHILD

Department of Economics, Princeton, New Jersey 08540

Received March 5, 1973

This paper states necessary and sufficient conditions for the convergence of the age structure (in a discrete time, one-sex model of population growth); it also contains a new and simple proof of the weak ergodic theorem of stable population theory. The main tool used to attain these results is Hilbert's notion of the projective metric. This metric provides a way of defining the distance between positive vectors in $\mathbb{R}^n$ which has two important features: First, the distance between any two positive vectors depends only on the rays on which the vectors lie; and, second, positive matrices act as contractions in this metric.

1. THE PROJECTIVE METRIC

Let x and y be vectors in $\mathbb{R}^n$ with $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$. We shall adopt the following conventions for vector inequalities.

(i) $x \geqslant y$, iff $x_i \geqslant y_i$ for all i;
(ii) $x > y$, iff $x \geqslant y$ and $x \neq y$;
(iii) $x \gg y$, iff $x_i > y_i$ for all i.

The vector x is positive if $x > 0$ (where 0 is the vector all of whose components are 0). The positive orthant is the set in $\mathbb{R}^n$ which consists of all the positive

* Research partially supported by the National Science Foundation and the Faculty Research Award Program of CUNY.

vectors. A vector x is strictly positive if $x \gg 0$. The same terminology applies to matrices. In this paper the set of all strictly positive vectors in $\mathbb{R}^n$ will be denoted Ω.

We define a distance between two vectors in the positive orthant of $\mathbb{R}^n$ and then show that this distance depends only on which rays the given vectors lie. (A ray in $\mathbf{R}^n$ is a half line starting at the origin.) Let x and y be positive vectors in $\mathbb{R}^n$. Define $p(x, y)$ as follows.

(1) Suppose there exist scalars a and b such that $x \leqslant ay$ and $ay \leqslant bx$; then define

$$p(x, y) = \min \ln(b),$$

where this minimum is taken over all pairs (a, b) satisfying the above inequalities. Figure 1 shows how the choice of a and b can be made.

(2) If no such scalars a and b exist, then define $p(x, y) = \infty$.

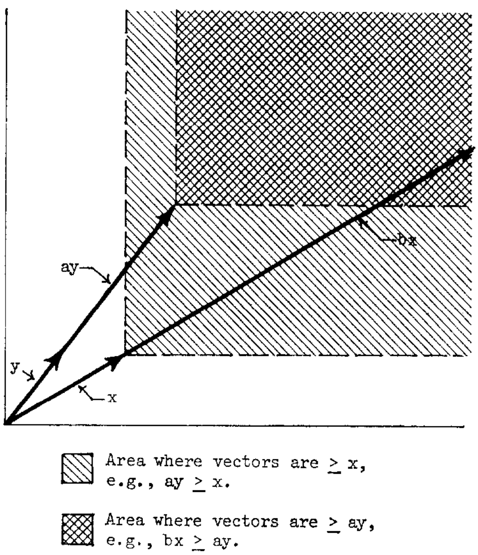

FIGURE 1

A way of computing $p(x, y)$ when x and y are strictly positive, is as follows. Let $r = \max_{1 \leqslant i \leqslant n} (x_i/y_i)$ and $s = \min_{1 \leqslant i \leqslant n} (x_i/y_i)$. Then $p(x, y) = \ln(r/s)$. Note this method will not work when $x = (1, 0) = y$.

As an example of (1), let $x = (1, 1, 2)$ and $y = (3, 2, 1)$; then $p(x, y) = \ln 6$. As an example of (2) let $x = (1, 0)$ and $y = (1, 1)$.

DEFINITION. p is called the projective metric and the number $p(x, y)$ is the projective distance from x to y.

The following lemma, which states some basic facts about p, justifies this terminology.

LEMMA 1.1. *Let x, y, and z be positive vectors in $\mathbb{R}^n$. Then,*

(i) $p(x, y) = p(rx, sy)$ *where r and s are positive scalars. Thus p only depends on the rays generated by x and y in the positive orthant.*
(ii) $p(x, y) \geqslant 0$.
(iii) $p(x, y) = 0$ *iff* $x = ay$ *for some positive scalar a.*
(iv) $p(x, y) = p(y, x)$.
(v) $p(x, y) \leqslant p(x, z) + p(z, y)$ (*the triangle inequality*).

Proof. (i) Suppose $x \leqslant ay \leqslant bx$, then $rx \leqslant r(ay) = (ra/s)(sy)$. Let $a' = ra/s$. Now $a'(sy) = r(ay) \leqslant r(bx) = b(rx)$. So $rx \leqslant a'(sy)$ and $a'(sy) \leqslant b(rx)$. Thus a "b" which works to compute $p(x, y)$ also works to compute $p(rx, sy)$. The process is reversible, so $p(rx, sy) = p(x, y)$.

(ii) If $x \leqslant ay \leqslant bx$, then $x \leqslant bx$, which implies $b \geqslant 1$. So $\ln b \geqslant 0$.

(iii) Since $x \leqslant x$, $p(x, x) = \ln(1) = 0$. By (i), $p(x, ax) = 0$. Conversely suppose that $p(x, y) = 0$. Then we can take $b = 1$. Thus $x \leqslant ay \leqslant x$ and $x = ay$.

(iv) If $x \leqslant ay \leqslant bx$, then $y \leqslant (b/a)\, x \leqslant (b/a)(ay) = by$. So a "$b$" which works to compute $p(x, y)$ also works to compute $p(y, x)$ and $p(x, y) = p(y, x)$.

(v) If either $p(x, z)$ or $p(z, y) = \infty$, the inequality holds trivially. So let $p(x, z) = \ln b$ and $p(z, y) = \ln b'$. Then $x \leqslant az \leqslant bx$ and $z \leqslant a'y \leqslant b'z$, for constants a and a'. It follows that

$$x \leqslant az \leqslant a(a'y) \leqslant a(b'z) = b'(az) \leqslant b'(bx),$$

or

$$x \leqslant (aa')y \leqslant (bb')x.$$

Hence, $p(x, y) \leqslant \ln(bb') = \ln b + \ln b' = p(x, z) + p(z, y)$.

Note that the projective distance between any two vectors in Ω, that is, between any two strictly positive vectors, is finite.

LEMMA 1.2. *Fix y in Ω. Then the real-valued function f defined on Ω by $f(x) = p(x, y)$ is continuous.*

Proof. As long as x and y are in Ω, there exist constants a and b such that $x \leqslant ay \leqslant bx$. It should be clear that the choice of a and b can be made continuously as x varies. (Look at Fig. 1.)

PROPOSITION 1.3. *Let x and y belong to the positive orthant of $\mathbb{R}^n$ and let S be a nonnegative $n \times n$ matrix. Then $p(Sx, Sy) \leqslant p(x, y)$. If $S \gg 0$ (all the* *entries of S are positive), then S is a strict contraction relative to p;* i.e., *$p(Sx, Sy) \leqslant K_S p(x, y)$ for all positive vectors x and y where K_S is a constant <1. The contraction constant K_S varies continuously with the entries of S.*

Proof. The complete proof, due to Birkhoff (1957), is too complicated to present here. The first part however, is easy. If $p(x, y) = \infty$, then $p(Sx, Sy) \leqslant p(x, y)$. If $p(x, y) < \infty$, then for some a and b,

$$x \leqslant ay \leqslant bx.$$

Since S is a positive matrix,

$$Sx \leqslant aSy \leqslant bSx$$

so that if $p(x, y) = \ln b$, $p(Sx, Sy) \leqslant \ln b$.[1]

2. The Weak Ergodic Theorem

The discrete, one-sex model of population growth may be sketched as follows. Break the population into n equally spaced age groups and let v_0 be the vector whose ith component is the number of people in the ith age group; v_0 is a vector in the positive orthant of $\mathbb{R}^n$. If we let $\| v \| = | v_1 | + \cdots + | v_n |$ where $v = (v_1, \ldots, v_n)$, then $\tilde{v}_0 = v_0 / \| v_0 \|$ gives the vector of percentages of people in each age group; $\tilde{v}_0$ is the age structure vector at time 0. Suppose that each age group has the birth rates $(b_1, \ldots, b_n)$ and survival rates $(s_1, \ldots, s_{n-1})$. The survival rate for the oldest age group s_n is necessarily equal to 0. In the discrete time model, birth rates necessarily include some component of survival; only in continuous models is fertility completely separated from mortality. This is an inconvenient, although not unimportant fact, which we shall henceforth ignore.

Let

$$T = \begin{pmatrix} b_1 & b_2 \cdots b_{n-1} & b_n \\ s_1 & 0 \cdots 0 & 0 \\ 0 & s_2 \cdots 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \cdots s_{n-1} & 0 \end{pmatrix}.$$

Then the population and age structure vectors next period are given by $v_1 = Tv_0$ and $\tilde{v}_1 = v_1 / \| v_1 \|$. Of course, we must measure time so that it takes exactly one period to move from one age category to the next. Thus, if the data in T represent birth and survival rates for 5-year intervals, a single time period is 5

[1] See Birkhoff (1966, Chap. 16), Bushell (1973), and Keeler (1972) for other discussions and applications of the projective metric.

years long. T is called a "population matrix" and is a positive $n \times n$ matrix. The Perron–Frobenius Theorem[2] states that a positive matrix like T has a unique positive eigenvalue λ whose modulus is exceeded by no other eigenvalue and a unique positive eigenvector e with $Te = \lambda e$ and $\| e \| = 1$. Such a matrix T is primitive if T has no other eigenvalue whose modulus is λ. A sufficient condition for a population matrix T to be primitive is that the survival rates s_i, the last birthrate, b_n, and birth rates in the middle age groups are nonzero (i.e., two successive age groups, not including the first, have positive birth rates).[3] With these conditions $T^l \gg 0$ for some integer l.

If the birth and survival rates are constant, then after k time periods the age structure is $\tilde{v}_k = T^k v_0 / \| T^k v_0 \|$.

These conditions are sufficient to ensure the convergence of the age structure v_k to the vector e independent of the initial population. This result is the Strong Ergodic Theorem of stable population theory.

THEOREM 2.1 (The Strong Ergodic Theorem). *Under the conditions stated* $\lim_{k\to\infty} \tilde{v}_k = e$.

Proof. This is a simple consequence of the Perron–Frobenius Theorem. We present a different proof here, one based on Proposition 1.3, which is almost identical to the proof of the weak ergodic theorem which we give below. Since both e and $\tilde{v}_k$ are of unit length, it will suffice to prove that $\lim_{k\to\infty} p(\tilde{v}_k, e) = \lim_{k\to\infty} p(T^k v_0, e) = 0$. Let $S = T^l \gg 0$; then, $T^k = T^{k-l[k/l]} S^{[k/l]} = U_k S^{[k/l]}$, where $[k/l]$ is the greatest integer in k/l and $U_k = T^{k-l[k/l]}$. If k/l is an integer, U_k is the identity matrix; in any event, U_k is positive and, by Proposition 1.3, $p(U_k x, U_k y) \leqslant p(x, y)$ for all positive x and y.

Since e is an eigenvector of T, e is an eigenvector of T^k. Thus,

$$\begin{aligned} p(T^k v_0, e) &= p(T^k v_0, T^k e) = p(U_k S^{[k/l]} v_0, U_k S^{[k/l]} e) \\ &\leqslant p(S^{[k/l]} v_0, S^{[k/l]} e) < K_S^{[k/l]-1} p(S v_0, S e), \end{aligned}$$

where $K_S < 1$ is the contraction constant whose existence is guaranteed by Proposition 1.3. Since $S \gg 0$, Sv_0 and Se are in Ω so that $p(Sv_0, Se)$ is finite. Clearly $\lim_{k\to\infty} p(T^k v_0, e) \to 0$.

It is unrealistic to assume that the birth and survival rates do not change with time. However, if these rates are constant over a single time period, then ergodic analysis is still possible. Let T_k be the matrix of birth and survival rates during the kth time period. Then after k periods, population and age structure vectors are $v_k = T_k \cdot T_{k-1} \cdots T_1 v_0$ and $\tilde{v}_k = v_k / \| v_k \|$. It is no longer true that

[2] See Nikaido (1968), for a proof. We assume that T is indecomposable.

[3] If $b_n = 0$ and b_k is the last nonzero birth rate, it is common to consider only the $k \times k$ matrix composed of the first k rows and columns of T. This essentially determines all facts of demographic interest. See Nikaido (1968) and Parlett (1972).

the $\tilde{v}_k$'s converge, but under rather modest assumptions on the T_k's, it is still true that for large k's, the vectors $\tilde{v}_k$ are independent of v_0. This is the Weak Ergodic Theorem which we now prove.

THEOREM 2.2 (The Weak Ergodic Theorem). *Let $T_1, T_2, \ldots$ be a sequence of primitive population matrices satisfying $M \leqslant T_k \leqslant N$ for all k, where M and N are fixed positive matrices and $M^l \gg 0$. Let $S_k = T_k \cdot T_{k-1} \cdots T_1$. Then if v_0 and w_0 are any positive vectors* $\lim_{k\to\infty} p(S_k v_0, S_k w_0) = 0$.

Remark. Since the projective distance depends only on rays, this theorem states that $p(\tilde{v}_k, \tilde{w}_k) \to 0$. It is possible to use these techniques to prove a slightly stronger result; that

$$p(\tilde{v}_k, \tilde{w}_k) < CK^{[k/l]-1}, \qquad K < 1,$$

where the constants C and K and l depend on M and N and not on v_0, w_0, or the T_k's. Thus the speed of convergence can be bounded independent of v_0 and w_0. See Keeler (1973) for details.

As in the proof of the strong ergodic theorem, the key is to apply Proposition 1.3 to products of the T_k's taken l at a time. The following lemma states that this can be done.

LEMMA 2.3. *There exists a constant $K < 1$ such that if S is any l-fold product of the T_k's, i.e., $S = T_{k_1} \cdots T_{k_l}$, then S contracts projective distance by at least K.*

Proof of Lemma 2.3. The boundedness assumption in the T_k's implies that $M^l \leqslant S \leqslant N^l$. By Proposition 1.3, S contracts distance in the projective metric by a factor K_S. Recall that a set in $\mathbb{R}^{n^2}$ is compact if it is closed and bounded and that any continuous function on a compact set achieves its minimum and maximum. Now the set of matrices S satisfying $M^l \leqslant S \leqslant N^l$ is closed and bounded in $\mathbb{R}^{n^2}$ and therefore compact. Since K_S varies continuously with S (Proposition 1.3 again), there is a $K > 0$ such that $K_S \leqslant K < 1$ for all such S.

Proof of Theorem 2.1. Again, let $[k/l]$ be the greatest integer in k/l. Then,

$$S_k = U_k V_{[k/l]} \cdots V_1,$$

where $V_i = T_{il} T_{il-1} \cdots T_{(i-1)l+1}$ and U_k is the identity matrix if k/l is an integer and otherwise $U_k = T_k \cdots T_{[k/l]+1}$. In either case U_k is a positive (but not necessarily strictly positive) matrix so that by Proposition 1.3, $p(U_k x, U_k y) \leqslant p(x, y)$ for all positive x and y. Then,

$$\begin{aligned} p(S_k v_0, S_k w_0) &= p(U_k V_{[k/l]} \cdots V_1 v_0, U_k V_{[k/l]} \cdots V_1 w_0) \\ &\leqslant p(V_{[k/l]} \cdots V_1 v_0, V_{[k/l]} \cdots V_1 w_0) \\ &\leqslant K^{[k/l]-1} p(V_1 v_0, V_1 w_0), \end{aligned}$$

where $K < 1$ is the contraction constant whose existence is guaranteed by Lemma 2.3. Since $V_1 \gg 0$, $p(V_1 v_0, V_1 w_0)$ is finite and $K^{[k/l]-1} p(V_1 v_0, V_1 w_0)$ converges to 0.[4]

3. Necessary and Sufficient Conditions for Convergence of Age Structure

The strong ergodic theorem states that if the population matrices T_k are constant in time ($T_k = T$ for all k) and primitive then the age structure converges. The weak ergodic theorem allows T_k to vary (with some boundedness assumptions) and concludes that the age structure, in the long run, does not depend on the initial population distribution. In general, it is not true that the age structure must converge. We now present conditions which are both necessary and sufficient for the convergence of the age structure.

We make the boundedness assumption of Section 2 on the sequence T_k, namely, $M \leqslant T_k \leqslant N$ for all k where $M^l \gg 0$ for some l. This implies that the survival rates and the middle and last birth rates of each T_k are bounded away from zero. Thus each T_k is primitive, has a unique positive eigenvalue λ_k which dominates all other eigenvalues of T_k (in modulus), and has a unique positive engenvector e_k with $\| e_k \| = 1$.

THEOREM 3.1. *With the assumptions and notation just given, the age-structure vector converges to a vector* e iff $\lim_{k \to \infty} e_k = e$.

Remark. Recall that in a compact set any sequence has a convergent subsequence. Moreover, a sequence in a compact set converges if every convergent subsequence converges to the same point.

Proof. Given an initial age-structure vector $\tilde{v}_0 = v_0$, define, as usual, $\tilde{v}_k = T_k \tilde{v}_{k-1} / \| T_k \tilde{v}_{k-1} \|$. The problem is not changed if we multiply each T_k by some positive scalar. So we may assume that $\lambda_k = 1$ for all k.

Necessity. Assume $\lim_{k \to \infty} \tilde{v}_k = e$. We must show that $\lim_{k \to \infty} e_k = e$. Since the e_k's are all unit length vectors, they vary within a compact set. Let $e_{k_1}, e_{k_2}, \ldots$, be a subsequence converging to f. By the remark above it is sufficient to show that $e = f$. Since the T_{k_i}'s vary within a compact subset, there is a convergent subsequence. So, by passing to a subsequence if necessary, we may assume that $\lim_{i \to \infty} T_{k_i} = T$, with $M \leqslant T$. The assumptions on M guarantee

[4] A similar proof of the weak ergodic theorem could be fashioned by using Lemma 3.3 of Furstenberg and Kesten (1960) to do the work of Proposition 1.3. For other discussions of the basic results of stable population theory, see, e.g., Bourgeois-Pichat (1968), Coale (1972), Lopez (1961), and Parlett (1970).

that T has a unique positive eigenvector with eigenvalue 1. (Note the modulus of the largest eigenvalue depends continuously on the matrix. Since that eigenvalue is 1 for each T_{k_i} it must be 1 for T.) In fact $Tf = \lim_{i\to\infty} T_{k_i}(e_{k_i}) = \lim_{i\to\infty} e_{k_i} = f$, so f must be that positive eigenvector. On the other hand

$$\lim_{i\to\infty} \tilde{v}_{k_i} = e = \lim_{i\to\infty} T_{k_i}(\tilde{v}_{k_i-1})/\| T_{k_i}(\tilde{v}_{k_i-1})\| = Te/\| Te \|.$$

Since $T^l \gg 0$, $Te \neq 0$ for any positive vector e. So e is also an eigenvector for T with positive entries and unit length. By the uniqueness of such a vector $e = f$.

Sufficiency. We assume that $\lim_{k\to\infty} e_k = e$ and show that $\lim_{k\to\infty} \tilde{v}_k = e$. Recall that l is the integer for which $M^l \gg 0$.

Part I. It is clearly sufficient to show that $\lim_{k\to\infty} \tilde{v}_{a+kl} = e$ for $a = 1, 2,..., l$. Moreover, since the $\tilde{v}_k$'s all have unit length, all we need do is show that any convergent subsequence of the sequence $\tilde{v}_{a+l}, \tilde{v}_{a+2l},...$, converges to e. (Recall the remark above.)

Since $T_k{}^l \gg 0$ for all k, each e_k is strictly positive. Furthermore $T_k{}^l \geqslant M^l$ for all k implies that all the entries of e_k are bounded away from zero, so all the e_k's and any limit point of the e_k's are in Ω.

Part II. For any $\epsilon > 0$, there is an integer $K > 0$ such that $p(e_k, e) < \epsilon$ for all $k \geqslant K$. (Use Lemma 1.2 and the fact that $e_k \to e$ in Ω.) For such a k, we claim that

$$p(\tilde{v}_{k+l}, e) < Cp(\tilde{v}_k, e) + 2l\epsilon, \tag{$*$}$$

where C is a constant <1 independent of k.

Let $S = T_{k+l} \cdot T_{k+l-1} \cdots T_{k+1}$. Then,

$$p(\tilde{v}_{k+l}, e) = p(S\tilde{v}_k/\| S\tilde{v}_k \|, e) = p(S\tilde{v}_k, e)$$

by Lemma 1.1(i). Now $p(S\tilde{v}_k, e) \leqslant p(S\tilde{v}_k, Se) + p(Se, Se_{k+1}) + p(Se_{k+1}, e)$ by Lemma 1.1(v). By Proposition 1.3, there is a $C < 1$ so that $p(Sx, Sy) \leqslant Cp(x, y)$ for all x, y, and S. So

$$p(\tilde{v}_{k+l}, e) \leqslant C\,p(v_k, e) + \epsilon + p(S\,e_{k+1}, e).$$

So we need only show that $p(S\,e_{k+1}, e) \leqslant (2l - 1)\epsilon$. Let $S' = T_{k+l} \cdots T_{k+2}$ so that $S = S' \cdot T_{k+1}$. Then since $T_{k+1}\,e_{k+1} = e_{k+1}$,

$$\begin{aligned} p(S\,e_{k+1}, e) &= p(S'e_{k+1}, e) \\ &\leqslant p(S'\,e_{k+1}, S'e) + p(S'e, S'e_{k+2}) + p(S'e_{k+2}, e) \\ &\leqslant p(e_{k+1}, e) + p(e, e_{k+2}) + p(S'e_{k+2}, e), \end{aligned}$$

the last inequality following from Proposition 1.3. So,

$$p(S\,e_{k+1}, e) \leqslant 2\epsilon + p(S'e_{k+2}, e).$$

A simple induction shows that

$$\begin{aligned} p(S\,e_{k+1}, e) &\leqslant 2(l-1)\epsilon + p(T_{k+l}e_{k+l}, e) \\ &\leqslant (2l-1)\epsilon. \end{aligned}$$

Part III. From $(*)$ we see that when $k > K$,

$$p(\tilde{v}_{a+(k+1)l}, e) \leqslant C\,p(\tilde{v}_{a+kl}, e) + 2l\epsilon.$$

Repeated applications of this formula show that

$$\begin{aligned} p(\tilde{v}_{a+(k+t)l}, e) &\leqslant C^t p(\tilde{v}_{a+k}, e) + 2l\epsilon(1 + C + \cdots + C^{t-1}) \\ &= C^t p(\tilde{v}_{a+k}, e) + 2l\epsilon(1 - C^t)/(1 - C). \end{aligned}$$

Since $C < 1$ and ϵ may be arbitrarily small, it follows that $\lim_{k\to\infty} p(\tilde{v}_{a+kl}, e) = 0$. Thus if $v_{a+k_1 l}, \tilde{v}_{a+k_2 l}, \ldots$ is a convergent subsequence of the $\tilde{v}_{a+kl}$'s we have that $p(\lim_{i\to\infty} \tilde{v}_{a+k_i l}, e) = 0$. By Lemma 1.1(iii), $\lim_{i\to\infty} \tilde{v}_{a+k_i l}$ is a scalar multiple of e. Since both are positive and of unit length, they must be equal. Q.E.D.

COROLLARY 3.2. *Suppose the sequence T_k converges to T where $T^l \gg 0$. Then the age-structure vectors always converge to the unique unit length positive eigenvector of T.*

Finally we make some remarks on the demographic interest of these results.

COROLLARY 3.3. *Suppose that the sequence of age-structure vectors $\tilde{v}_k$ converges to $e = (e_1, \ldots, e_n)$ and that T and T' are limit points of the sequence of population matrices T_k. Let $(b_1, \ldots, b_n, s_1, \ldots, s_{n-1})$ and $(b_1', \ldots, b_n', s_1', \ldots, s_{n-1}')$ be the birth and survival rates of T and T', respectively. Then there exist positive constants a and a' such that*

$$a \sum_{i=1}^{n} b_i e_i = a' \sum_{i=1}^{n} b_i' e_i, \tag{1}$$

and

$$a s_j = a' s_j', \qquad \text{for} \quad 1 \leqslant j \leqslant n-1. \tag{2}$$

Proof. Let $a = 1/\| Te \|$ and $a' = 1/\| T'e \|$. In proving the necessity part of Proposition 3.1, we showed that

$$e = Te/\| Te \| = aTe = T'e/\| T'e \| = a'T'e.$$

Writing the equation $aTe = a'T'e$ component-by-component yields the desired conclusion.

The demographic meaning of these results is clear: An age structure will approach a constant only if the crude birth rates ($\sum_{i=1}^{n} b_i e_i$) and each age-specific survival rate approach constants or if, in the limit, these rates vary proportionally and simultaneously. The fact that the age structure converges imposes no other restrictions on the asymptotic behavior of the entries of T_k. The sufficiency part of Theorem 3.1 guarantees that this asymptotic behavior for the crude birth rates and the age-specific survival rates is sufficient to guarantee convergence of the age structure.

Remark. A version of these results was obtained by Bourgeouis-Pichat (1968), who discussed the relationships which must hold among the different demographic schedules when the age structure does not vary. His monograph, which contains many results of the same general form as Corollary 3.3, emphasizes the practical, empirical implications of stable population theory.

Acknowledgments

We are grateful to Jane Menken for guidance on demographic matters, to Nathan Keyfitz for suggesting we combine our earlier papers, Golubitsky and Rothschild (1973) and Keeler (1973), and to Joel Cohen for referring us to the work of Bourgeois-Pichat, and Furstenberg and Kesten.

References

Birkhoff, G. 1957. Extensions of Jentzch's theorem, *Trans. Amer. Math. Soc.* **85**, 219–227.

Birkhoff, G. 1967. "Lattice Theory," 3d ed., American Mathematical Society, Providence, RI.

Bourgeois-Pichat, J. 1968. "The Concept of a Stable Population: Applications to the Study of Countries with Incomplete Demographic Statistics," United Nations, Department of Economic and Social Affairs, Population Studies, No. 39, ST/SOA/Series A/39.

Bushell, P. J. 1973. Hilbert's metric and positive contraction mappings in a Banach space. *Arch. Rational Mech. Anal.* **52**, 330–338.

Coale, A. J. 1972. "The Growth and Structure of Human Populations: A Mathematical Investigation," Princeton University Press, Princeton, NJ.

Furstenberg, H. and Kesten, H. 1960. Products of random matrices, *Ann. Math. Statist.* **31**, 451–469.

Golubitsky, M. and Rothschild, M. 1973. On the convergence of the age structure, Memo No. 152, Econometric Research Program, Princeton University, Princeton, NJ.

Keeler, E. B. 1972. A twisted turnpike, *Internat. Econ. Rev.* **13**, 160–166.

Keeler, E. B. 1973. "A New Method in Ergodic Analysis," unpublished memo.

Lopez, A. 1961. "Some Problems in Stable Population Theory," Office of Population Research, Princeton University, Princeton, New Jersey.

Nikaido, H. 1968. "Convex Structures and Economic Theory," Academic Press, New York.

Parlett, B. 1970. Ergodic properties of populations 1: the one-sex model, *Theor. Pop. Biol.* **1**, 191–207.