SOME RANDOM WALKS ARISING IN LEARNING MODELS I

SAMUEL KARLIN

Introduction The present paper presents an analysis of certain transition operators arising in some learning models introduced by Bush and Mosteller [2]. They suppose that the organism makes a sequence of responses among a fixed finite set of alternatives and there is a probability p_s^n at moment n that response s will occur. They suppose further that the probabilities $p_s^{(n+1)}$ are determined by the p_s^n, the response s_n made after moment n, and the outcome or event r_n that follows response s_n. We shall examine in detail the one-dimensional models which occur in their theory. These models can be described in simplest form as follows: There exist two alternatives A_1 and A_2, and two possible outcomes r_1 and r_2, for each experiment. There exists a set of Markoff matrices F_{ij} which will apply where choice i was made and outcome r_j occurs. Let p represent the initial probability of choosing alternative A_2, and $1 - p$ the probability of choosing A_1. Depending on the choice and outcome, the vector $(p, 1 - p)$ is transformed by the appropriate F_{ij} into a new probability vector which represents the new probabilities of preference of A_2 and A_1, respectively, by the organism. The psychologist is interested in knowing the limiting form of the probability choice vector $(p, 1 - p)$.

The mathematical description of the simplest process of this type can be formulated as follows: A particle on the unit interval executes a random walk subject to two impulses. If it is located at the point x, then $x \longrightarrow F_1 x = \sigma x$ with probability $1 - \phi(x)$, and $x \longrightarrow F_2 x = 1 - \alpha + \alpha x$ with probability $\phi(x)$. The actual limiting behavior of x depends on the nature of $\phi(x)$. The transition operator representing the change of the distribution describing the position of the particle is given by

$$(TF)(x) = \int_0^{x/\sigma} [1 - \Phi(t)] dF + \int_0^{(x-1+\alpha)/\alpha} \Phi(t) dF.$$

We introduce an additional operator, acting on continuous functions, and given by

---

$$U\pi(t) = (1 - \phi(t))\,\pi(\sigma t) + \phi(t)\,\pi(1 - \alpha + \alpha t).$$

It turns out that T is conjugate to U; hence knowing the behavior of U one obtains much information about T. This interplay shall be exploited considerably. The operator T is not weakly completely continuous, nor does it possess any kind of compactness property; thus none of the classical ergodic theorems apply to this type [3]. The limiting behavior of $T^n F$ depends very sensitively on the assumptions made about the operators F_i and the probabilities $\phi(x)$.

Section 1 treats the case where $\phi(x) = x$. This causes the boundaries 0 and 1 to be absorbing states, and thus the limiting distribution concentrates only at these points. However, the concentration depends on the initial distribution. By examining the corresponding U in detail, we have been able to obtain much additional knowledge. For example, we have shown that if π is m times continuously differentiable then $(U^n\pi)^{(r)}$ converges uniformly for each $0 \leq r \leq m - 1$. It is worth emphasizing that the knowledge of the convergence of the distributions does not imply the uniform convergence of $U^n\pi$ for any continuous function π. Additional arguments are needed for this conclusion. In this connection, we finally remark that R. Bellman, T. Harris, and H. N. Shapiro [1] have analyzed only this case independently. They did not point out the connection between the operators T and U. The methods they used to establish the convergence of $T^n F$ are probabilistic. Our paper in §1 overlaps with theirs in some of the theorems, notably 6, 8, 9, 12, and 15; our results subsume theirs, and their proofs are entirely different from ours. Section 2 considers the case where $\phi(x)$ is monotone increasing and

$$|\phi(x) - \phi(y)| \leq u < 1.$$

This leads to the ergodic phenomonon, or steady-state situation, where the limiting distributions are independent of the starting distributions.

In §3, we examine the situation $\phi(x) = 1 - x$. This corresponds to completely reflecting boundaries, and of course the ergodic phenomenon holds. Other interesting properties of the operators are also developed. We consider in §4 the case where $\phi(x)$ is linear and monotonic decreasing. Section 5 introduces a further possibility where we allow the particle to stand still with certain probability. This type has been statistically examined by M. M. Flood [5]. In §6 we investigate the general ergodic type where $\phi(x)$ is not necessarily linear. The arguments here combine both abstract analysis and probabilistic reasoning involving recurrent event theory. Furthermore, it is worth emphasizing, the proofs given in §6 apply without any modifications to the case where we allow any finite number of impulses acting on the particle. In a future paper we

shall present the extension of this model to the circumstance where changes in time occur continuously and the possible motion of the particle has a continuous or infinite discrete range of values.

The last section studies some of the properties of the limiting distribution in the ergodic types. It is shown in all circumstances that the limiting distribution is either singular or absolutely continuous, and the actual form depends on the value of $\alpha + \sigma$.

Most of the analysis carries over to higher dimensional models where more alternatives are allowed. In a subsequent paper we shall present this theory with other generalizations. We finally note that this paper represents a combination of abstract analysis and probability; it is hoped that the methods used will be useful for future investigations of this type.

It has been brought to my attention by the referee that the material of [6], [7], [8], and [9] relate closely to the content of this paper. This techniques seem to be different.

1. A particle undergoes a random walk on the unit interval subject to the following law: If the particle is at x, then after unit time $x \longrightarrow \alpha + (1 - \alpha)x$ with probability x, and $x \longrightarrow \sigma x$ with probability $1 - x$, where $0 < \alpha, \sigma < 1$. If $F(x)$ represents the cumulative distribution describing the location of x at the beginning of the time interval, with the understanding that $F(x) \equiv 1$ for $x \geq 1$ and $F(x) = 0$ for $x \leq 0$, then the new distribution locating the position of the particle at the end of the time interval is given by

$$(1) \qquad G(x) = TF = \int_0^{x/\sigma} (1 - t) \, dF(t) + \int_0^{(x-\alpha)/(1-\alpha)} t \, dF(t).$$

Indeed, the probability $dG(x)$ that after unit time the particle is located at x can materialize in two ways; namely, the particle was at x/σ and moved with probability $1 - x/\sigma$ to x, or it jumped with probability $(x - \alpha)/(1 - \alpha)$ from $(x - \alpha)/(1 - \alpha)$ to x during the unit time interval. This yields

$$dG(x) = \left(1 - \frac{x}{\sigma}\right) dF\left(\frac{x}{\sigma}\right) + \frac{x - \alpha}{1 - \alpha} \, dF\left(\frac{x - \alpha}{1 - \alpha}\right),$$

which easily implies the conclusion of equation (1).

Equation (1) represents the transition law for the particular Markoff process on hand.

The transformation T is easily seen to furnish a linear bounded mapping of the space of functions of bounded variation (V) on the unit interval into

itself. Furthermore, T takes distributions into distributions and is of norm 1. This section investigates the behavior of T^n for large n with the aim of determining limiting properties of T^n.

We consider the following additional mapping U applied to the space of of continuous functions defined on the unit interval $C[0, 1]$:

$$(2) \qquad (U\pi)(t) = (1 - t)\,\pi(\sigma t) + t\pi(\alpha + (1 - \alpha)t).$$

The operator U has a probabilistic interpretation which we shall speak about later; but its direct relevance to T is given in Theorem 1. The inner-product notation

$$(\pi, F) = \int_0^1 \pi(t)\,dF(t)$$

will be extensively used.

THEOREM 1. *The conjugate map U^* to U is T.*

Proof. It is necessary to verify that $(U\pi, F) = (\pi, TF)$ for any continuous function $\pi(t)$ and any distribution $F(t)$ with $F(t) \equiv 1$ for $t \geq 1$ and $F(t) = 0$ for $t \leq 0-$. Indeed,

$$(U\pi, F) = \int (1 - t)\pi(\sigma t)\,dF(t) + \int t\pi(\alpha + (1 - \alpha)t)\,dF(t).$$

By a change of variable, we get

$$(U\pi, F) = \int \left(1 - \frac{t}{\sigma}\right) \pi(t)\,dF\left(\frac{t}{\sigma}\right) + \int \pi(t)\,\frac{t - \alpha}{1 - \alpha}\,dF\left(\frac{t - \alpha}{1 - \alpha}\right)$$

$$= \int \pi(t)\,dG(t) \text{ where } G(t) = TF.$$

The value of Theorem 1 is that, by studying the iterates of U^n, we deduce corresponding results about the conjugate operators T^n. We proceed now to study this operator U. To be complete, we should denote the operator by $U_{\sigma,a}$, but where no ambiguity arises we shall drop the subscripts. Let W denote the isometry

$$W\pi(t) = \pi(1 - t).$$

Clearly $W^{-1} = W$. We now observe the identity

$$(3) \qquad\qquad U_{1-a,\ 1-\sigma} = W U_{\sigma,\ a} W.$$

The mapping $(\sigma, \alpha) \longrightarrow (1 - \alpha, 1 - \sigma)$ of the parameter space into itself has the effect of mapping the triangle of the unit square bounded above by $1 - \alpha - \sigma = 0$ into the other triangle located in the unit square. This isomorphism property (3) enables us to restrict our attention to the case where $1 - \alpha - \sigma \geq 0$. Corresponding theorems valid for the other circumstances, where $1 - \alpha - \sigma < 0$, are deduced easily by virtue of (3) and will be summarized at the end of this section. From now on in §2, unless explicitly stated otherwise, we shall assume that $1 - \alpha - \sigma \geq 0$.

The next two theorems, which we state for completeness, are immediate from (2).

THEOREM 2. *The operator U preserves the values at 0 and 1.*

THEOREM 3. *The operator U is positive; that is, it transforms positive continuous functions into positive continuous functions.*

In particular, if $\pi_1(t) \geq \pi_2(t)$, for all t, then $U\pi_1 \geq U\pi_2$.

THEOREM 4. *If $\pi, \pi', \cdots, \pi^{(n)} \geq 0$, then $U\pi, (U\pi)', \cdots, (U\pi)^{(n)} \geq 0$.*

Proof. A simple calculation yields

$$(4) \qquad (U\pi)^{(n)} = (1 - t)\sigma^n \pi^{(n)}(\sigma t) + t(1 - \alpha)^n \pi^{(n)}(\alpha + (1 - \alpha)t)$$

$$+ n(1 - \alpha)^{n-1} \pi^{(n-1)}(\alpha + (1 - \alpha)t) - n\sigma^{n-1} \pi^{(n-1)}(\sigma t).$$

Since

$$\sigma t < t < \alpha + (1 - \alpha)t,$$

we conclude since $\pi^{(n-1)}(t)$ is monotonic increasing that

$$\pi^{(n-1)}(\alpha + (1 - \alpha)t) \geq \pi^{(n-1)}(\sigma t) \geq 0.$$

The assumption that $1 - \alpha \geq \sigma$ implies that $(1 - \alpha)^{n-1} \geq \sigma^{n-1}$. As $\pi^{(n)}(t) \geq 0$, it follows that $(U\pi)^{(n)} \geq 0$. The same conclusion and argument apply to $(U\pi)^{(i)}$ for $0 \leq i \leq n - 1$.

In particular, U transforms positive monotonic convex functions into functions of the same kind. Although in the proof of Theorem 4 we assumed the existence of derivatives, the argument can be carried through routinely at the expense of elegance, by use of the general definitions of convexity and monotonicity.

THEOREM 5. *If* $c \geq \pi^{(i)}(t) \geq 0$ *for* $0 \leq i \leq n$, *then* $(U^r\pi)^{(i)}(1) \leq K_i$ *for* $0 \leq i \leq n$ *and hence* $(U^r\pi)^{(i)}(t) \leq K_i$.

Proof. The proof is by induction. By Theorem 2, the theorem is trivially true for $i = 0$. Suppose we have established the result for the ith derivative with $0 \leq i \leq n - 1$. Equation (4) yields

$$(5) \qquad (U\pi)^{(n)}(1) - \pi^{(n)}(1) = c_1(\alpha)\pi^{(n-1)}(1) - c_2(\sigma)\pi^{(n-1)}(\sigma)$$

$$+ [(1-\alpha)^n - 1]\,\pi^{(n)}(1),$$

where $c_1(\alpha)$ and $c_2(\sigma)$ are constants depending only on α and σ respectively, and on n. If

$$\pi^{(n)}(1) > M(\alpha, \sigma, c),$$

where M is a constant sufficiently large, then (5) yields

$$(U\pi)^{(n)}(1) < \pi^{(n)}(1).$$

Since $c_1(\alpha)$ and $c_2(\sigma)$ do not depend on k, and by the induction hypotheses

$$|(U^k\pi)^{n-1}(x)| \leq M$$

uniformly in k and x, we find in general that when $(U^k\pi)^{(n)}(1)$ becomes larger than $M(\alpha, \sigma, c)$, then

$$(U^{k+1}\pi)^{(n)}(1) < (U^k\pi)^{(n)}(1).$$

Consequently, the iterates $(U^k\pi)^{(n)}(1)$ for $k \geq k_0$ are bounded by

$$M(\alpha, \sigma, c) + c_1(\alpha)M + c_2(\sigma)M.$$

This trivially implies the conclusion of Theorem 5.

The proof of the next theorem is due originally to R. Bellman. We present

it for completeness.

THEOREM 6. *There exists at most one continuous solution of $U\pi = \pi$ for which $\pi(0) = 0$ and $\pi(1) = 1$.*

Proof. (By contradiction.) Let π_1 and π_2 denote two solutions with the prescribed boundary conditions. Put $\pi_0 = \pi_1 - \pi_2$; then $\pi_0(0) = \pi_0(1) = 0$. Let t_0 be a point where π_0 achieves its maximum. Since

$$\pi(t_0) = (1 - t_0)\pi(\sigma t_0) + t_0\pi(\alpha + (1 - \alpha)t_0),$$

we deduce that σt_0 is also a maximum point. Iterating, we find by continuity that $\pi(0) \equiv 0$ is the maximum value of $\pi(t)$. A similar argument shows that $0 = \min \pi(t)$, which implies that $\pi_1 = \pi_2$.

THEOREM 7. *For any function $\pi(t) = t^r$ with $\infty > r \geq 1$, $U^n(t^r)$ converges uniformly as $n \longrightarrow \infty$.*

Proof. Clearly $t \geq t^r > p(t)$, where

$$p(t) = \begin{cases} 0 & \text{for } 0 \leq t \leq t_0; \\[2mm] \dfrac{t - t_0}{1 - t_0} & \text{for } t_0 \leq t \leq 1; \end{cases}$$

and t_0 is close to 1 with r fixed. Since Ut is convex by Theorem 4, and the values at 0 and 1 are fixed, we find that $t \geq Ut$. Hence

$$U^n t \geq U^{n+1} t \geq 0,$$

and $\lim U^n t = \theta(t)$ for every t. Since $\theta(t)$ is convex, and by Theorem 5 the derivatives of $U^n t$ at 1 are uniformly bounded, we conclude that $\theta(t)$ is continuous. By Dini's theorem the convergence of $U^n t$ to $\theta(t)$ is uniform. Obviously, $U\theta = \theta$. On the other hand, if t_0 is close to 1 then $(Up)'(1) < p'(1)$ (see the proof of Theorem 5). Since Theorem 4 guarantees the convexity of Up, and the slope at 0 is 0, it follows that $Up \leq p$, and hence $U^n p \leq U^{n+1} p$; therefore $\lim U^n p = \phi(t)$. Again, $\phi(t)$ is a continuous fixed point, and therefore by Theorem 6 we infer that $\phi(t) = \theta(t)$. On account of $U^n t \geq U^n t^r \geq U^n p$, we deduce that $\lim U^n t^r = \phi(t)$ with the convergence being uniform.

We denote this unique fixed point of U by $\phi_{\sigma, a}(t)$, or by $\phi(t)$ whenever no ambiguity arises.

THEOREM 8. *The iterates U^n converge strongly (that is, $U^n \pi$ converges uniformly for any continuous function π).*

Proof. The constant functions are fixed points of U^n. Consequently by Theorem 7, $U^n q$ converges uniformly for any function $q(t)$ in the linear space L spanned by the functions $(1, t^r)$. The set L is dense in the space of continuous functions. Moreover, as $\|U^n\| = 1$, by a well-known theorem of Banach, $U^n q$ converges strongly when applied to any continuous function $q(t)$.

The actual limit is easily seen to be given by

$$(6) \qquad \lim_{n \to \infty} U^n q(t) = q(1)\, \phi_{\sigma, a}(t) + q(0)[1 - \phi_{\sigma, a}(t)].$$

This is an immediate consequence of the fact that the fixed points of U consist of the two-dimensional space spanned by the function 1 and $\phi_{\sigma, a}$. Equation (6) shows that two functions q_1 and q_2 which agree at 0 and 1 have the same limit. This enables us to show:

THEOREM 9. *If $q(t)$ is any bounded function continuous at 0 and 1, then $U^n q$ converges strongly.*

Proof. Let $q(t)$, in addition to being continuous at 0 and 1, possess finite derivatives at 0 and 1. Then clearly there exist two continuous functions $h_1(t)$ and $h_2(t)$ with

$$h_1(t) \geq q(t) \geq h_2(t),$$

where $h_1(0) = h_2(0)$ and $h_1(1) = h_2(1)$. We conclude the result from this using the argument of Theorem 7 and equation (6). If now $q(t)$ is only continuous at 0 and 1, then we can find for any ϵ a $q_\epsilon(t)$ satisfying the properties assumed about $q(t)$ in the first part of the proof with $|q(t) - q_\epsilon(t)| \leq \epsilon$. As $\|U^n\| = 1$, the conclusion of the theorem now follows by a standard argument.

THEOREM 10. *If $|\pi^{(i)}(t)| \leq c_i$ for $0 \leq i \leq m$, then $|U^n \pi^{(i)}(t)| \leq c_i$ for $0 \leq i \leq m$.*

Proof. The proof is by induction. For $r = 0$, the result is trivial since U preserves positivity, and the constant functions are fixed points of U. Suppose we have established the result for $r = m - 1$. We note that

$$U\pi^{(m)} = (1-t)\sigma^m \pi^{(m)}(\sigma t) + t(1-\alpha)^{(m)} \pi^{(m)}(\alpha + (1-\alpha)t)$$

$$+ m(1-\alpha)^{m-1} \pi^{(m-1)}(\alpha + (1-\alpha)t) - m\sigma^{m-1} \pi^{(m-1)}(\sigma t).$$

This easily yields that

$$\max_t \, | \, U\pi^{(m)}(t) \, | \leq \lambda \, \max \, | \, \pi^{(m)}(t) \, | \, + \, C \, \max \, | \, \pi^{(m-1)}(t) \, | \, ,$$

where

$$\lambda \, = \, \max_t \, [\, (1 - t)\sigma^m + t(1 - \alpha)^m] \, < \, 1 \, .$$

Therefore,

$$\max_t \, | \, (U^k\pi)^{(m)}(t) \, | \leq \lambda \, \max \, | \, (U^{k-1}\pi)^{(m)}(t) \, | + C \, \max_t \, | \, (U^{k-1}\pi)^{m-1}(t) \, |$$

$$\leq \lambda \, \max \, | \, (U^{(k-1)}\pi)^{(m)}(t) \, | \, + \, K$$

by our induction hypothesis. Iterating this last inequality gives that

$$\max_t \, | \, (U^k\pi)^{(m)}(t) \, | \leq \sum_{i=0}^{k-1} \lambda^i K \, + \, \lambda^k \, \max_t \, | \, \pi^{(m)}(t) \, | \leq M \, .$$

This establishes the theorem.

THEOREM 11. *If $q(t)$ belongs to C^n (n continuous derivatives), then*

$$\lim_{m \to \infty} \, [\, U^m q(t)]^{(r)}$$

converges uniformly for $0 \leq r \leq n - 1$.

Proof. We prove the theorem only for $r = 1$, for the other cases are similar. On account of Theorem 10, the uniform boundedness of $(U^m q)^{(2)}$ implies the equi-continuity of $U^m q^{(1)}$. Thus we can select a subsequence converging uniformly since $U^m q^{(1)}$ are also uniformly bounded. Let

$$\Psi(t) \, = \, \lim_{i \to \infty} \, U^{m_i} q^{(1)} \, .$$

Since $\lim U^{m_i} q$ converges uniformly to a unique limit $\theta(t)$, we obtain that $\theta'(t) = \Psi(t)$. As $\theta'(t)$ is independent of the subsequence chosen, the conclusion of the theorem easily follows.

THEOREM 12. *The fixed point $\phi_{\sigma,\alpha}$ is analytic for $0 \leq t \leq 1$ with $\phi_{\sigma,\alpha}^{(r)} \geq 0$.*

Proof. Let $p(t)$ denote a function infinitely differentiable with $p^{(r)}(t) \geq 0$ and $p(0) = 0$, $p(1) = 1$. By virtue of Theorem 11 and Theorem 4 we deduce that

$$\lim_{n \to \infty} (U^n p)^{(r)} = \phi_{\sigma,a}^{(r)} \geq 0.$$

Therefore $\phi_{\sigma,a}$ is absolutely monotonic and hence, by a well-known theorem, is analytic.

At this point it seems desirable to summarize the analogous results of Theorems 2 through Theorem 12 for the case where $\alpha + \sigma \leq 1$. We enumerate the corresponding theorems.

THEOREM 4'. *If* $(-1)^{i-1} \pi^{(i)}(t) \geq 0$ *for* $i = 0, 1, 2, \cdots, n$, *and* $\pi(t) \geq 0$, *then* $(-1)^{i-1} (U\pi)^{(i)}(t) \geq 0$.

In particular, positive increasing concave functions are transformed into functions of the same kind.

THEOREM 5'. *If* $C \geq \pi(t) \geq 0$ *and* $C \geq (-1)^{i-1} \pi^{(i)}(t) \geq 0$ *for* $1 \leq i \leq n$, *then* $0 \leq (-1)^{i-1} (U^r \pi)^{(i)}(0) \leq K_i$, *and hence* $|U^r \pi^{(i)}(t)| \leq K_i$ *for* $1 \leq i \leq n$.

Theorem 6 remains unchanged and is valid independent of the conditions on α and σ, provided only they lie in the open unit interval.

Theorem 7 holds with a modification of the proof where $p(t)$ is replaced by the concave function

$$p(t) = \left\{ \begin{array}{ll} 1 & \text{for. } 1 \geq t \geq t_0 \\ \dfrac{1}{t_0} t & \text{for } 0 \leq t \leq t_0 \end{array} \right\},$$

and the functions t^r are replaced by $1 - (1 - t)^r$. These also constitute, with the constant function, a family of functions whose linear span is dense in $C[0, 1]$. This enables us to infer the validity of Theorem 8. Theorems 9, 10, and 11, with suitable changes in their statements which we leave for the reader, are established by simple appropriate modifications similar to that indicated above for Theorem 7. The unique solution $\phi_{\sigma,a}$ for this situation, where $\alpha + \sigma \leq 1$, is completely monotonic and hence analytic. In the remainder of this section the theorems are established without any specification as to the value of $\alpha + \sigma$.

THEOREM 13. *The functions*

$$\phi_m(t) = \sum_{n=m}^{\infty} U^n(t(1-t))$$

converge geometrically to 0.

Proof. It is immediate from (6) that

$$U^n(t(1-t)) = \Psi_n(t)$$

tends uniformly to zero. Since the derivative at 0 and 1 of $t(1-t)$ is 1 and -1, we conclude by Theorem 11 that for n sufficiently large there exists an $n_0(\lambda)$ such that

$$U^{n_0}(t(1-t)) \le \lambda t(1-t)$$

with $\lambda < 1$. Let kn_0 denote the last integer k for which $kn_0 \le m$. We obtain

$$0 \le \phi_m(t) \le \phi_{kn_0}(t) \le \frac{\lambda^k}{1-\lambda} \sum_{i=0}^{n_0-1} U^i(t(1-t)) \le C\lambda^k \le C\rho^{(n_0+1)k} < C\rho^m,$$

where

$$\rho = \lambda^{1/(n_0+1)} < 1.$$

THEOREM 14. If $q(t)$ is continuous, $|q'(1)| < \infty$ and $|q'(0)| < \infty$, then $\lim U^n[q(t)]$ converges geometrically.

Proof. We first establish the result for special functions t^r with $1 \le r \le \infty$. A simple calculation shows that

$$-Ct(1-t) \le U(t^r) - t^r \le Ct(1-t).$$

For $n < m$, we obtain upon continued application of U and summation that

$$-C\sum_{i=m}^{n} U^i(t(1-t)) \le U^n(t^r) - U^m(t^r) \le C\sum_{i=m}^{n} U^i(t(1-t)).$$

The conclusion now follows from Theorem 13. The general function $q(t)$, satisfying the hypothesis of Theorem 14, can be bounded from above and below by two polynomials $P_1(t)$ and $P_2(t)$ which agree at 0 and 1. The result now follows

directly from this fact and the first part of this proof.

We observe easily the identity

$$Ut - t = (\alpha + \sigma - 1)\, t(1 - t).$$

Applying successively U and adding, we obtain

$$(7) \qquad \phi_{\sigma,\alpha} = \lim_{n \to \infty} U^m t = t + (\alpha + \sigma - 1) \sum_{n=1}^{\infty} U_{\sigma,\alpha}^n t(1 - t).$$

This is useful for purposes of calculation.

Some remarks describing the dependence of $\phi_{\sigma,\alpha}$ on σ and α are in order. We consider the following identity:

$$(8) \qquad U_{\sigma,\alpha}^n - U_{\sigma',\alpha'}^n = \sum_{i=0}^{n-1} U_{\sigma,\alpha}^i (U_{\sigma,\alpha} - U_{\sigma',\alpha'})\, U_{\sigma',\alpha'}^{n-i-1}.$$

If $f(t)$ is any function with bounded derivatives, then we obtain by the mean-value theorem that

$$|(U_{\sigma,\alpha} - U_{\sigma',\alpha'})f| \le |(1 - t)[f(\sigma t) - f(\sigma' t)] + t[f(\alpha + (1-\alpha)t) - f(\alpha' + (1-\alpha')t)]|$$

$$\le C(|\sigma - \sigma'| + |\alpha - \alpha'|)\, t(1 - t).$$

Applying equation 8 to $f(t) = \phi_{\sigma',\alpha'}$, and remembering that inequalities are preserved by Theorem 2, we obtain

$$|U_{\sigma,\alpha}^n \phi_{\sigma',\alpha'} - \phi_{\sigma',\alpha'}| \le C(|\sigma - \sigma'| + |\alpha - \alpha'|) \sum_{i=0}^{n-1} U^i(t(1-t)).$$

Allowing n to go to ∞, we have easily that

$$|\phi_{\sigma,\alpha} - \phi_{\sigma',\alpha'}| \le K(|\sigma - \sigma'| + |\alpha - \alpha'|),$$

where $K(\eta)$ is finite, provided that $0 < \eta < \alpha, \alpha'\sigma, \sigma' < 1 - \eta < 1$.

It is worthwhile to discuss the nature of $\phi_{\sigma,\alpha}$ for (σ,α) lying on the boundary of the unit square. First, we observe by direct verification that when $\alpha + \sigma = 1$, then $\phi_{\sigma,\alpha}(x) = x$. Next let $\alpha = 0$ and $\sigma < 1$; then

$$U\phi = (1-x)\phi(\sigma x) + x\phi(x).$$

Therefore, if ϕ is a fixed point with $\phi(0) = 0$ and $\phi(1) = 1$, then for $x \neq 1$ we have that $\phi(x) = \phi(\sigma x)$, and hence $\phi(x) \equiv \phi(0) = 0$ $(0 \leq x < 1)$ provided that ϕ is continuous at 0. Similarly, when $\sigma = 1$ and $\alpha < 1$ then the only fixed point ϕ continuous at 1 and satisfying $\phi(0) = 0$, $\phi(1) = 1$, is $\phi(x) \equiv 1$ for $0 < x \leq 1$. On the other two boundaries of the unit square the solutions are easily calculated and turn out as follows: If $0 < \sigma < 1$ is arbitrary and $\alpha \leq 1$, then

$$\phi_{\sigma,1} = 1 - \prod_{r=0}^{\infty}(1 - \sigma^r x),$$

while when $\sigma = 0$, $0 < \alpha < 1$, then

$$\phi_{\sigma,a} = \prod_{r=0}^{\infty} L^r x,$$

where $L^0 = I$ and the operation L applied to x gives $\alpha + (1-\alpha)x$. Finally for $\alpha = 0$, $\sigma = 1$ the operator U reduces to the identity mapping. We now investigate the dependence of $\phi_{\sigma,a}$ on σ and α as we allow σ and α to tend to the boundary. We limit our attention for definiteness to studying the case where $(\sigma, \alpha) \rightarrow (\sigma_0, 0)$ with $\sigma_0 < 1$, and we show that $\phi_{\sigma,a}$ converges pointwise to 0 for $0 \leq x < 1$, and $\phi_{\sigma,a}(1) \equiv 1$ otherwise. Moreover, the convergence is uniform in any interval $0 \leq x \leq 1 - \delta < 1$. Let $(\sigma_n, \alpha_n) \longrightarrow (\sigma_0, 0)$; then without loss of generality we may assume that $1 - \sigma_n - \alpha_n > 0$. Therefore the ϕ_{σ_n,a_n} are convex, monotonic increasing and positive, with $\phi_{\sigma_n,a_n}(0) = 0$. Also, for any interior interval $0 \leq x \leq 1 - \sigma < 1$, the first derivatives ϕ'_{σ_n,a_n} are uniformly bounded. Since this implies the ϕ_{σ_n,a_n} are equi-continuous over the subinterval, and as $0 \leq \phi_{\sigma_n,a_n} \leq 1$, we can select a subsequence which may be denoted as ϕ_{σ_r,a_r} converging to $\Psi(t)$ uniformly, for any interval of the form $0 \leq x \leq 1 - \delta < 1$. As

$$\phi_{\sigma_r,a_r}(1) = 1,$$

we get $\Psi(1) = 1$ and similarly $\Psi(0) = 0$. The uniform convergence of ϕ_{σ_r,a_r} guarantees the continuity of Ψ at zero.

Put

$$U_r = U_{\sigma_r,a_r}, \quad U_0 = U_{\sigma_0,0} \quad \text{and} \quad \phi_r = \phi_{\sigma_r,a_r}.$$

We consider the following identity:

$$\Psi - U_0\Psi = (\Psi - \phi_r) + (\phi_r - U_r\Psi) + (U_r\Psi - U_0\Psi) = I_1 + I_2 + I_3 .$$

We take a fixed $x < 1$; then trivially $|I_1| = |\Psi - \phi_r| \leq \epsilon$ when r is sufficiently large. Also

$$|I_2| = |\phi_r - U_r\Psi| = |U_r\phi_r - U_r\Psi| = |(1-x)[\phi_r(\sigma_r x) - \Psi(\sigma_r x)]$$

$$+ x[\phi_r(\alpha_r + (1-\alpha_r)x) - \Psi(\alpha_r + (1-\alpha_r)x)]| .$$

But for $x = x_0 < 1$ fixed, we observe that $\alpha_r + (1-\alpha_r)x_0$ varies in an interval $\leq 1 - \delta$ as $\alpha_r \longrightarrow 0$, and the same applies to $\sigma_r x$. The uniform convergence of $\phi_r \longrightarrow \Psi$ inside $0 \leq x \leq 1 - \delta$ yields $|I_2| \leq \epsilon$. By construction, $|I_3| \leq \epsilon$ for r large. Thus we infer the equality $\Psi = U_0\Psi$ for $0 \leq x < 1$, and by direct verification for $x = 1$. However, the fixed point to the equation $U_0\Psi = \Psi$ with $\Psi(0) = 0$, $\Psi(1) = 1$ and Ψ continuous at 0 is $\Psi(x) = 1$ for $0 \leq x < 1$ and $\Psi(1) = 1$. Thus the limit function Ψ is the same for every subsequence of ϕ_{σ_n, a_n}, and hence we deduce that $\phi_{\sigma_n, \dot{a}_n}$ converges pointwise. We furthermore note that Ψ is independent of $\sigma_0 < 1$. A similar analysis applies to the case where $(\sigma, \alpha) \longrightarrow (1, \alpha)$ $(\alpha > 0)$. The continuity properties of the solution for the other two boundaries yield to simpler analysis. Summarizing, we have established the following theorem:

THEOREM 15. *The fixed points $\phi_{\sigma,a}$ satisfy the following continuity properties: If $0 < \eta < \alpha$, $\alpha' \leq 1$ and $0 \leq \sigma, \sigma' \leq 1 - \eta$, then*

$$|\phi_{\sigma,a} - \phi_{\sigma',a'}| \leq K(\eta)[|\sigma - \sigma'| + |\alpha - \alpha'|] .$$

If $(\sigma, \alpha) \longrightarrow (\sigma_0, 0)$ with $\sigma_0 < 1$, then $\phi_{\sigma,a}(x) \longrightarrow 0$ pointwise for $0 \leq x < 1$ and $\phi_{\sigma,a}(1) \equiv 1$. If $(\sigma, \alpha) \longrightarrow (1, \alpha_0)$ with $\alpha_0 > 0$, then $\phi_{\sigma,a}(x) \longrightarrow 1$ pointwise for $0 < x \leq 1$.

Finally, a word concerning convergence of $U^n\pi$ for π continuous when the parameter values lie on the boundary. When $\alpha = 0$, $\sigma < 1$, then $U^n\pi$ converges pointwise. The same conclusion holds when $\alpha > 0$ and $\sigma = 1$. On the other two boundaries the convergence is uniform for $U^n\pi$. We omit the proofs.

We now return to the study of the operator T.

THEOREM 16. *For any distribution the iterates T^nF converge in the sense*

of distributions to the distribution

$$G(x) = I_1(x) \int \phi_{\sigma,a} \, dF + I_0(x) \int (1 - \phi_{\sigma,a}) \, dF,$$

where $I_0(x)$ and $I_1(x)$ are the distributions concentrating fully at 0 and 1 respectively.

Proof. From the convergence of $U^n \pi$ for any continuous function π and Theorem 1 follows the weak*convergence of $T^n F$. This is equivalent to the convergence of $T^n F$ in the sense of distributions. The actual form of

$$\lim_{n \to \infty} T^n F = G$$

as given in the theorem follows directly from (6).

By choosing the distribution $F = I_{x_0}$, we obtain from Theorem 6 that $\phi_{\sigma,a}(x_0)$ represents the probability with which the limiting distribution concentrates at 1, or in other words - as can be easily shown - the probability with which the particle beginning at x_0 will converge to 1. This furnishes a probability interpretation to the fixed point of the operator U which is different from a constant.

In connection with Theorem 8, we remark that $U^n \pi$ cannot converge for an arbitrary Lebesgue measureable bounded function. In fact, if we assume that $U^n \pi$ converges for every bounded measureable function $\pi(t)$, then $T^n F$ would converge weakly if F were absolutely continuous. Since the space of all integrable functions $L[0, 1]$ is weakly complete, and T maps distributions into distribution, we could find a fixed point $TF = F$ with F absolutely continuous and total variation 1. However, in view of (16) the only fixed distributions which exist concentrate only at 0 and 1, and hence cannot be absolutely continuous.

Finally, we present a slight application of Theorem 14. We show that the expected position of the particle converges geometrically for any starting distribution, although the iterated distributions converge slowly to the limiting distribution. The expected position of the particle is given by

$$\int_0^1 x \, dF(x) = (x, F),$$

where F is the cumulative distribution describing the position. The expected

position at the nth step is given by

$$(x, T^n F) = (U^n x, F).$$

On account of Theorem 14, $U^n x$ converges geometrically, which establishes the assertion. The same conclusion applies to all the moments. This observation is very useful for computational and estimation purposes.

Finally, we note that the spectrum of the operator T cannot consist of the isolated point 1. Otherwise, by standard techniques one can show that $U^n \pi$ converges for any measurable bounded function π.

2. In this second model the random walk is described as follows: If the particle is at x, then $x \longrightarrow \alpha + (1 - \alpha)x$ with probability $\phi(x)$ and $x \longrightarrow \sigma x$ with probability $1 - \phi(x)$, where

$$|\phi(x) - \phi(y)| \le \mu < 1.$$

The analogous transition operator to (1) becomes

$$(9) \quad G(x) = TF = \int_0^{x/\sigma} (1 - \phi(t))(dF(t)) + \int_0^{(x-a)/(1-a)} \phi(t) \, dF(t),$$

with the same understanding concerning F applying as before. Let

$$(10) \qquad U\pi = [1 - \phi(t)] \pi(\sigma t) + \phi(t) \pi(\alpha + (1 - \alpha)t).$$

In this section, we take $0 < \alpha, \sigma < 1$; the case where boundary values for α and σ are considered is easy to handle but not of great interest. The spaces on which they operate are the same as in §1. Again, in a similar manner to Theorem 1, we obtain:

THEOREM 17. *The operator T is conjugate to the operator U.*

We now further assume that $\phi(t)$ is monotonic increasing. This model includes the important case where $\phi(t) = \lambda + \mu t$, where $\lambda + \mu \le 1$; and whenever $\lambda + \mu = 1$ then $\lambda > 0$.

THEOREM 18. *The operator U preserves positivity and positive monotonic increasing functions.*

Proof. Direct verification.

Since the hypothesis on $\phi(t)$ implies either $\phi(1) < 1$ or $\phi(0) > 0$, we analyze the case where $\phi(1) < 1$. The other circumstance can be treated in an analogous manner. Furthermore, we now assume that if $\phi(0) = 0$, then $\phi'(0)$ exists and is finite.

THEOREM 19. *If* $\pi(t)$ *is monotonic increasing bounded and positive, then* $U^n \pi$ *converges uniformly to a constant.*

The proof can be carried out easily using the techniques employed above.

The hypothesis on $\phi(t)$ easily yields the fact that the only continuous fixed points of $U\pi = \pi$ are constant functions. The proof is similar to the proof used in Theorem 6. This fact directly connects with the result of Theorem 21 below. First, we complete the proof of convergence of $U^n \pi$ for any continuous function $\pi(t)$.

THEOREM 20. *The operators* $U^n \pi$ *converge uniformly for any continuous function.*

Proof. Since $\| U^n \| = 1$, and the space of all monotonic positive continuous functions spans a dense subset of the set of all continuous functions, the theorem follows by a well-known theorem of Banach.

THEOREM 21. *For any distribution* F, *the distribution* $T^n F$ *converge as distributions to a unique distribution* G *for which* $TG = G$ *which is independent of* F.

Proof. The weak*convergence of $T^n F$ follows directly from Theorem 20 and Theorem 16. To complete the proof we must establish that if $\lim T^n F = G$ and $\lim T^n H = K$, then $G = K$. Indeed, let Ψ denote any continuous function. We have that

$$(11) \quad (\Psi, G-K) = \lim_{n \to \infty} (\Psi, T^n(F-H)) = \lim_{n \to \infty} (U^n\Psi, F-H) = a\left(\int dF - \int dH\right) \equiv 0,$$

as F and H are distributions. Hence

$$\int \Psi(t)\, dF(t) = \int \Psi(t)\, dK(t)$$

for any continuous function Ψ, and therefore $G = K$.

It seems extremely difficult to determine the complete nature of this unique

fixed distribution. We shall say more about it in a future section. We denote it by $F_{\sigma, a}$.

THEOREM 22. *The distributions $F_{\sigma, a}$ is a continuous function of σ, α; that is, if $(\sigma_n, \alpha_n) \longrightarrow (\sigma, \alpha)$ with $0 < \sigma$, $\alpha < 1$, then $F_{\sigma_n, a_n} \longrightarrow F_{\sigma, a}$ at every point of continuity of $F_{\sigma, a}$.*

Proof. Let $(\sigma_n, \alpha_n) \longrightarrow (\sigma, \alpha)$; by Helly's theorem we can choose a subsequence $F_r = F_{\sigma_{n_r} a_{n_r}}$ converging to the distribution F at every continuity point. Write T_r for $T_{\sigma_{n_r} a_{n_r}}$ and T for $T_{\sigma, a}$. Let $\pi(t)$ denote any fixed continuous function. We consider the quantity

$$(\pi, F - TF) = (\pi, F - F_r) + (\pi, F_r) - (\pi, TF_r) + (\pi, TF_r - TF).$$

Since $F_r \longrightarrow F$ as distributions, we find for r sufficiently large that $|(\pi, F - F_r)| < \epsilon$. Now we note that

$$|(\pi, F_r) - (\pi, TF_r)| = |(\pi, T_r F_r) - (\pi, TF_r)| = |(U_r \pi - U\pi, F_r)|.$$

Since $U = U_{\sigma_{n_r} a_{n_r}}$ converges strongly to $U = U_{\sigma, a}$, as is trivial to verify, it follows that U_r converges uniformly to $U\pi$. Whence, as F_r are distributions, we infer that

$$|(U_r \pi - U\pi, F_r)| \leq \max_t |U_r \pi - U\pi| < \epsilon$$

when r is chosen large enough. Evidently, with r large we get as before that

$$|(\pi, T(F_r - F))| = |(U\pi, F_r - F)| \leq \epsilon.$$

Therefore we obtain for r large that $|(\pi, F - TF)| \leq 3\epsilon$, and hence $(\pi, F) = (\pi, TF)$. Since π is any continuous function, we infer $F = TF$ and therefore $F = F_{\sigma, a}$ by Theorem 21. Consequently, as any limit distribution of F_{σ_n, a_n} must be $F_{\sigma, a}$ the conclusion of Theorem 22 is now immediate.

3. The model considered in this section is with $\phi(x) = 1 - x$. In this case ϕ is monotonic decreasing. The operator U becomes

$$(12) \qquad\qquad U\pi(t) = t\pi(\sigma t) + (1 - t)\pi(1 - \alpha + \alpha t).$$

Note that we have replaced α by $1 - \alpha$. This is only for convenience in Theorem 28, and does not restrict any generality. In this model the closer the particle moves to the ends 0 and 1 the greater probability there is of moving back

into the interior. The situation described here is of completely reflecting boundaries. Again it is easy to show that the only continuous fixed points $U\pi = \pi$ are the constant function. Therefore, we shall find as in §2 that the distributions describing the position of the particle converge to a limit distribution independent of the initial distribution. We first proceed to analyze convergence properties of $U^n\pi$. In this case it is no longer true that U preserves the class of positive monotonic functions. Only positivity is conserved by the mapping U. However, a new quality as described in Theorem 23 serves here well.

Throughout this section in order to avoid trivial changes of proof and different results at times, we suppose that $0 < \alpha, \sigma < 1$.

THEOREM 23. *If $\pi(t)$ has a continuous derivative, then*

$$\max_t |(U\pi)'(t)| \leq \max_t |\pi'(t)|,$$

with equality holding if and only if $\pi(t)$ is linear.

Proof. By direct computation, we obtain

$$U\pi'(t) = t\sigma\pi'(\sigma t) + (1-t)\alpha\pi'(1-\alpha+\alpha t) + \pi(\sigma t) - \pi(1-\alpha+\alpha t).$$

Hence, with the aid of the mean-value theorem we get

$$(13) \quad \max_t |U\pi'(t)| \leq \max_t |t\sigma\pi'(\sigma t) + (1-t)\alpha\pi'(1-\alpha+\alpha t)|$$

$$+ (\sigma t - (1-\alpha) - \alpha t) \left| \frac{\pi(\sigma t) - \pi(1-\alpha+\alpha t)}{\sigma t - (1-\alpha) - \alpha t} \right|$$

$$\leq \max_t [t\sigma + (1-t)\alpha + 1 - \alpha - (\sigma-\alpha)t] \max_t |\pi'(t)| = \max_t |\pi'(t)|.$$

If equality holds, then let t_0 denote a point where

$$\max_t |\pi'(t)| = |\pi'(t_0)|.$$

It follows easily from (13) that

$$(14) \quad \max |\pi'(t)| = |\pi'(\sigma t_0)| = |\pi'(1-\alpha+\alpha t_0)| = \left| \frac{\pi(\sigma t_0) - \pi(1-\alpha+\alpha t_0)}{\sigma t_0 - (1-\alpha) - \alpha t_0} \right|.$$

This yields that $\pi(t)$ is linear for $\sigma t_0 \le t \le 1 - \alpha + \alpha t_0$, or otherwise somewhere between σt_0 and $1 - \alpha + \alpha t_0$ the slope has greater magnitude than the slope of the chord subtended by $\pi(t)$ at these points. Equation (14) yields also that σt_0 and $(1 - \alpha + \alpha t_0)$ are maximum points of $\pi'(t)$. Repeating this argument successively then shows that equality in (13) requires $\pi(t)$ to be linear.

THEOREM 24. *If $\pi(t)$ belongs to C^m ($\pi(t)$ possesses m continuous derivatives), then $\max_t |(U^n \pi)^{(r)}(t)|$ is uniformly bounded in n for each r ($0 \le r \le m$).*

Proof. The proof is similar to that of Theorem 10.

THEOREM 25. *If $\pi(t)$ possesses two continuous derivatives, and $\sigma \neq \alpha$, then $U^n \pi$ converges uniformly to a constant.*

Remark. The reason why the two cases $\sigma = \alpha$ and $\sigma \neq \alpha$ are distinguished, and necessarily so, will be explained later.

Proof. In view of Theorem 23 and Theorem 24, the first and second derivatives of $U^n \pi$ are uniformly bounded. Thus $U^n \pi$ and $(U^n \pi)'$ constitute equicontinuous families of functions. We can thus select a subsequence n_i such that $U^{n_i} \pi$ converges uniformly to $\phi(t)$, and $(U^{n_i} \pi)'$ converges uniformly to $\phi'(t)$. It follows trivially that $U^{n_i + 1} \pi$ tends uniformly to $U\phi$ and

$$U^{n_i + 2} \pi \longrightarrow U^2 \phi.$$

Moreover, by virtue of Theorem 23,

$$(15) \qquad \max_t |(U^{n_i} \pi)'| \ge \max_t |(U^{n_i + 1} \pi)'| \ge \max_t |(U^{n_{i+1}} \pi)'|.$$

Hence

$$\lim_{i \to \infty} \max_t |(U^{n_i} \pi)'| = \lim_{i \to \infty} \max_t |(U^{n_i + 1} \pi)'| = \lim_{i \to \infty} \max_t |(U^{n_i + 2} \pi)'|.$$

Therefore, by the uniform convergence of the derivatives, we secure

$$\max_t |\phi'(t)| = \max_t |(U\phi)'(t)| = \max_t |(U^2 \phi)'(t)|.$$

Invoking Theorem 23 yields that $\phi(t)$ and $U\phi(t)$ are linear. However, if $\alpha \neq \sigma$ and $\phi(t)$ contains a term with t, then $U\phi$ is quadratic. This impossibility

forces $\phi(t)$ to be identically a constant. Let i be chosen sufficiently large so that

$$| U^{n_i} \pi - c | \leq \epsilon .$$

Then

$$| U^{n_i+1} \pi - c | \leq t | U^{n_i} \pi(\sigma t) - c | + (1 - t) | U^{n_i} \pi(1 - \alpha + \alpha t) - c | < \epsilon .$$

Repeating this argument shows that

$$| U^{n_i+p} \pi - c | \leq \epsilon$$

for any p. This establishes that $U^n \pi$ converges uniformly to c.

THEOREM 26. *If $\pi(t)$ is continuous and $\sigma \neq \alpha$, then $U^n \pi$ converges uniformly.*

Proof. The space of all functions with two continuous derivatives spans linearly a dense subset of the space of all continuous functions. Since $||U^n|| = 1$, we obtain the result using Theorem 25 and a well-known theorem of Banach.

In the next two theorems we establish the uniform convergence of $U^n \pi$ for the case where $1 > \sigma = \alpha > 0$. We note in this case the interesting fact that U applied to a polynomial does not increase its degree. Particularly,

$$Ux^n = [\alpha^n - n\alpha^{n-1}(1 - \alpha)]x^n + P_{n-1}(x),$$

where $P_{n-1}(x)$ denotes a polynomial of degree $n - 1$.

THEOREM 27. *If $P(t)$ is any polynomial, then $U^k P$ converges uniformly to a constant and the convergence is geometric.*

Proof. The proof is by induction on the degree of the polynomial. Clearly if P is a constant $= c$ then $U^k P \equiv c$. Suppose we have shown for any polynomial P_{n-1} of degree $\leq n - 1$ that the iterates $U^k P_{n-1}$ converge uniformly. To complete the proof, it is enough to verify that $U^k x^n$ converges uniformly. Let

$$\lambda = \alpha^n - n\alpha^{n-1}(1 - \alpha);$$

then $| \lambda | < 1$ since $1 > \alpha > 0$. We obtain

$$Ux^n = \lambda x^n + P_{n-1}(x).$$

Repeating, we get, for $k \geq 1$,

$$U^k x^n = \lambda^k x^n + \sum_{r=0}^{k-1} \lambda^r U^{k-r-1} P_{n-1} .$$

This last sum is of the form

$$c_k = \sum_{r=0}^{k} a_r b_{k-r},$$

with $\sum |a_r| < \infty$, and $\lim_{k=\infty} b_k(x)$ exists. It is a well-known theorem that $\lim c_k(x)$ exists uniformly whenever

$$b_k(x) = U^{k-1} P_{n-1}$$

converges uniformly. Thus, $U^k x^n$ converges uniformly to a fixed point which must be a constant function. Finally we note that in this case where $\sigma = \alpha$ (the rate of learning, so to speak, is the same regardless of the outcome of the experiment), then $U^n P$ for any polynomial converges geometrically. The proof can be carried through easily by induction.

This yields the fact that the expected position converges geometrically to a limiting expected position with similar statements applying to higher moments.

THEOREM 28. *If $\pi(t)$ is continuous and $\sigma = \alpha > 0$, then $U^n \pi$ converges uniformly.*

Proof. Similar to Theorem 26, since the set of all polynomials is dense.

We now note the important example that when $\alpha = \sigma = 0$ it is no longer true that $U^n \pi$ converges. It is easily verified that in this case $U^{2n} \pi$ and $U^{2n+1} \pi$ converge separately but that a periodic phenomenon occurs otherwise. The argument of Theorem 27 breaks down in this case as the quantity λ is -1. We only mention that other difficult convergence behavior occurs when α, σ traverse the boundary of the unit square for this model. In particular, when $\alpha = 1$ and $\sigma < 1$ it is not hard to show that $U^n_{\sigma, \alpha} \pi$ does not necessarily converge for every continuous function π, and even for the circumstance where π is a polynomial. The case where $\sigma = \alpha = 1$ produces for U the identity operator for which the convergence of U^n is trivial. For $\alpha < 1$ and $\sigma = 1$ we can conclude again a lack

of convergence. However, when $\alpha = 0$ and $1 > \sigma > 0$, or $\sigma = 0$ and $1 > \alpha > 0$, then $U_{\sigma,a}^n \pi$ converges for every continuous function π.

We return now to the hypothesis $0 < \alpha$, $\sigma < 1$.

THEOREM 29. *If $\pi(t)$ belongs to C^m, then $(U^k \pi)^{(r)}(t)$ converges uniformly for $0 \leq r \leq m$.*

Proof. This follows easily from Theorems 24, 26, and 28. Let

Let

$$TF = \int_0^{x/\sigma} t \, dF(t) + \int_0^{(x+a-1)/a} (1-t) \, dF(t).$$

This represents the transition law for the distribution describing the position of of the particle for this model. By arguments analogous to those employed in the preceding sections, we can establish the following theorems, using the conjugate relationship between T and U.

THEOREM 30. *For any distribution F the distributions $T^n F$ converge as distributions to a unique distribution $F_{\sigma,a}$ for which $TF_{\sigma,a} = F_{\sigma,a}$, which is independent of F.*

THEOREM 31. *The distributions $F_{\sigma,a}$ constitute a continuous family of distributions in the sense of Theorem 22.*

Again it seems very difficult to determine any more explicit information about $F_{\sigma,a}$.

4. **The model examined here** is such that $1 - \phi(x) = \lambda x + \mu$, with $\lambda + \mu \leq 1$ and at least $1 > \lambda$ or $0 < \mu$. The operator U has the form

$$(16) \qquad U\pi = (\lambda x + \mu)\, \pi(\sigma x) + (1 - \lambda x - \mu)\, \pi(1 - \alpha + \alpha x).$$

Of course, as before, $0 < \alpha$, $\sigma < 1$. Convergence questions for $U^n \pi$ turn out to be very elementary in this case in view of the following theorem which is easily proven.

THEOREM 32. *If $\pi(x)$ has a bounded derivative, then*

$$\max_x |(U\pi)'(x)| \leq a \max_x |\pi'(x)|$$

with $a < 1$.

An immediate consequence of Theorem 32 is that $(U^k \pi)'$ converges geometrically to 0. Let T denote the transition operator of distributions for this model. In the standard way, we obtain:

THEOREM 33. *For any distribution F the distributions $T^n F$ converge to the distribution $F_{\sigma,\alpha}$ which is a continuous function of (σ, α), and $TF_{\sigma,\alpha} = F_{\sigma,\alpha}$. Moreover, $F_{\sigma,\alpha}$ is independent of F.*

5. **This section is devoted to some variations** of the preceding models. A new feature added first is that we allow in addition to the two impulses of motions towards the two fixed points 0 and 1 by the transformations

$$F_1 x = \sigma x \quad \text{and} \quad F_2 x = 1 - \alpha + \alpha x$$

the possibility of a third motion where the particle stands still with certain probability. These models are particularly important in learning problems, and much statistical investigation on this type has been done by M. M. Flood [5]. They are referred to as the pure models. The mathematical description of the first model of this type is as follows: A particle x on the unit interval is subject to three random impulses: (1) $x \longrightarrow \sigma x$ with probability $\pi_1(1-x)$; (2) $x \longrightarrow 1 - \alpha + \alpha x$ with probability $\pi_2 x$; and (3) $x \longrightarrow x$ with probability $(1-\pi_1)(1-x) + (1-\pi_2)x$, where $0 \leq \pi_1, \pi_2 \leq 1$. This is similar to model I where absorption takes place at the boundaries 0 and 1. The operator analogous to (2) becomes

$$(17) \quad U\pi = \pi_1(1-x)\pi(\sigma x) + [(1-\pi_1)(1-x) + (1-\pi_2)x]\,\pi(x)$$

$$+ \pi_2 x\, \pi(1 - \alpha + \alpha x).$$

Again, let T denote the transition operator which maps the distribution locating the particle into the corresponding distribution at the end of the experiment. Theorem 1 is valid for this setup, and T is consequently conjugate to U. It is easy to verify that U fulfills the conditions of Theorems 2 and 3 and also preserves the property of monotone increasing functions. Furthermore, we obtain:

THEOREM 34. *If π, π' and $\pi'' \geq 0$, then $(U\pi)'' \geq 0$ if and only if*

$$(1 - \sigma)\pi_1 + \pi_2(\alpha - 1) \geq 0,$$

and otherwise $U\pi$ preserves with π and $\pi' \geq 0$ the property of concavity.

Proof. The proof can be carried through by direct computation.

We remark that the remainder of the analogue to Theorem 4 does not carry over under the condition stated in Theorem 34. Moreover, noting that we have here changed α into $1 - \alpha$ as compared to §2, we obtain for $\pi_1 = \pi_2 = 1$ the condition of §1 for preservation of convexity, and so on.

The analogues of Theorems 5, 6, 7, and 8 easily extend to this model by the same methods, and we obtain that $U^n \pi$ converges uniformly to a limit given by

$$(18) \qquad [1 - \phi_{\sigma, a, \pi_1, \pi_2}(x)]\, \pi(0) + \phi_{\sigma, a, \pi_1, \pi_2}(x)\, \pi(1),$$

where $\phi_{\sigma, a, \pi_1, \pi_2}$ is the unique continuous fixed point of $U\phi = \phi$ with $\phi(0) = 0$ and $\phi(1) = 1$. The entire theory of geometric convergence, continuity of ϕ as a function of σ, α, π_1, and π_2, and the form of the limiting distribution of the particle established for the model of §1 remains valid with slight changes in the proofs. The general conclusion is that introducing a probability of standing still has no effect on the convergence of the distributions or its limiting form provided only the essential feature of absorbing boundaries still persists. Finally, in this connection we remark that for special boundary values of the parameters π_1 and π_2 the motion may become a drift to one or other of the end points; for example, $\pi_1 = 0$, $\pi_2 > 0$.

6. **We treat in this section,** the following general nonlinear one-dimensional learning model. The particle moves with probability $\phi(x)$ from x to $1 - \alpha + \alpha x$ and with probability $1 - \phi(x)$ from x to σx. The function is only continuous with the additional important requirement for this case that $\phi(x) \geq \delta > 0$ and $1 - \phi(x) > \delta > 0$ for all x in the unit interval. This excludes the types of models discussed in §§1 and 3, but includes some subcases of the examples investigated in §§2 and 4. However, in those cases we obtained much stronger results about the rate of convergence of derivatives, and so on. The transition operators become

$$(20) \qquad TF = \int_0^{x/\sigma} [1 - \phi(t)]\, dF(t) + \int_0^{(x-1+a)/a} \phi(t)\, dF(t),$$

and T is adjoint to

$$(21) \qquad (U\pi)(t) = (1 - \phi(t))\, \pi(\sigma t) + \phi(t)\, \pi(1 - \alpha + \alpha t).$$

We shall show that $U^n \pi$ converges uniformly for any continuous function $\pi(t)$. The proof of this fact shall be based on the following highly intuitive proposition. Let an experiment be repeated with only two possible outcomes, success

or failure at each trial. Suppose further that the probability of success p_n at the nth trial depends on the outcome of the previous trial, but that these conditional probabilities satisfy $p_n \geq \eta > 0$; that is, regardless of the previous number of failures the conditional probability of success is always at least $\eta > 0$. Then the recurrent event of a success run of length r with r fixed is a certain event; that is, with probability 1 it will occur in finite time. This result can be deduced in a standard way using the theory of recurrent events [4].

We turn back now to the examination of $U^n \pi$. Let

$$F_1 x = \sigma x \quad \text{and} \quad F_2 x = 1 - \alpha + \alpha x$$

and by Fx denote the operation that either F_1 or F_2 is applied. We note the important obvious fact that

$$(22) \qquad\qquad |F^r x - F^r y| \leq \lambda^r |x - y|,$$

with $0 < \lambda < 1$, where F^r denotes r applications of F_1 and F_2 in some order acting on x and y in the same way.

Next, we need the important lemma:

LEMMA. *If* $|\phi^{(m)}(t)| \leq K$ *for* $m = 0, 1, \cdots,$ *and* $|\pi^{(m)}(t)| \leq K_1,$ *then* $|U^n \pi^{(m)}(t)| \leq K_2$ *uniformly in n and t.*

Proof. The proof is similar to that of Theorem 24.

Now let $\pi(t)$ denote a continuously differentiable function. Consider the following identity:

$$(23) \quad U^n \pi(x) - U^n \pi(y) = (1 - \phi(x))(1 - \phi(y))[U^{n-1} \pi(F_1 x) - U^{n-1} \pi(F_1 y)]$$

$$+ \phi(x) \phi(y)[U^{n-1} \pi(F_2 x) - U^{n-2} \pi(F_2 y)]$$

$$+ (1 - \phi(y)) \phi(x)[U^{n-1} \pi(F_2 x) - U^{n-1} \pi(F_1 y)]$$

$$+ \phi(y)(1 - \phi(x))[U^{n-1} \pi(F_1 x) - U^{n-1} \pi(F_2 y)].$$

We continue to apply this identity to the factors $U^{n-1} \pi(\cdot) - U^{n-1} \pi(\cdot)$; and when any term of the form $U^m \pi(F^r w) - U^n \pi(F^r z)$ is achieved, then that factor is allowed to stand without any further reduction. All other terms are reduced to expressions involving as factors $\pi(\cdot) - \pi(\cdot)$. Thus we obtain

$$U^n \pi(x) - U^n \pi(y) = I_1^n + I_2^n,$$

when I_1 consists of terms of the form

$$\sum p_k [U^{mk} \pi(F^r w_k) - U^{mk} \pi(F^r z_k)],$$

and $\sum p_k \leq 1$ while I_2 consists of the remaining terms. We now conceive of the following probability model. Let two particles undergo the random walk described by this model starting from x and y, respectively. We say a success occurs if the same impulse activates both particles, and otherwise failure occurs. The probability of success is given initially by

$$\phi(x) \, \phi(y) + [1 - \phi(x)][1 - \phi(y)] \geq 2\delta^2 > 0,$$

and it is easily seen that each p_k, where p_k is the conditional probability of success occurring on the kth trial, satisfies

$$p_k \geq 2\delta^2 > 0.$$

Consequently, a success run of length r is certain to happen in finite time. In particular as $n \longrightarrow \infty$, $I_2^n \longrightarrow 0$, since I_2^n is bounded by twice the probability of no success run in n trials times K. On the other hand, in view of the lemma and equation (22) we secure that $I_1^n \leq C\lambda^r$. Therefore,

$$\varlimsup_{n \to \infty} \; | U^n \pi(x) - U^n \pi(y)| \leq C\lambda^r,$$

which can be made arbitrarily small as $r \longrightarrow \infty$. Hence, if

$$\lim \, U^n \pi(y) = a$$

exists for a single y, then

$$\lim_{n \to \infty} \; U^n \pi(x) = a$$

for every x. Since a subsequence can be found so that

$$\lim_{i \to \infty} \, U^{n_i} \pi(x) = a$$

for one x and hence for all x, an argument used in the close of the proof of

Theorem 25 shows that

$$\lim_{n \to \infty} U^n \pi(x) = a.$$

The lemma easily implies that the convergence is uniform. Using the fact that $||U^n|| = 1$, we can sum up the conclusions for this nonlinear model as follows:

THEOREM 35. *If $\pi(t)$ is continuous, then $\lim_{n \to \infty} U^n \pi$ exists uniformly converging to a constant limit.*

THEOREM 36. *If $\phi(t)$ belongs to C^m, and $\pi(t)$ is in C^m, then*

$$\lim_{n \to \infty} (U^n \pi)^{(m)}(t) = 0$$

with convergence uniform in t.

THEOREM 37. *For any distributions F, $T^n F$ converges to a distribution $F_{\sigma,a}$ independent of F with $TF_{\sigma,a} = F_{\sigma,a}$ and $F_{\sigma,a}$ continuous with respect to σ, α.*

This last theorem follows on account of the conjugate relationship of T and U.

Finally, we note that the method used in this section can be employed to analyze the random walks with any number of impulses

$$F_i x = (1 - \alpha_i) m_i + \alpha_i x.$$

7. In the present section we investigate the nature of the limiting distribution obtained in the various models. In the case where the boundaries were absorbing states as in §§ 1 and 5, we find that the limiting distribution is discrete and concentrates at the two ends 0 and 1. The weight at 1 depends on the starting distribution F and is given by

$$\int_0^1 \phi_{\sigma,a}(x) \, dF(x),$$

where $\phi_{\sigma,a}$ is the unique continuous fixed point of $U\phi = \phi$ with $\phi(0) = 0$ and $\phi(1) = 1$. Many properties of $\phi_{\sigma,a}$ are developed in those sections. In all the other types the ergodic property was seen to hold and the limiting distribution was independent of the initial distribution. Let us deal with the following

general type. The random walk is given by $x \longrightarrow F_1 x = \sigma x$ with probability $1 - \phi(x)$, and $x \longrightarrow F_2 x = 1 - \alpha + \alpha x$ with probability $\phi(x)$, where $1 - \delta \geq \phi(x) \geq \delta > 0$. The relevant operators are given by equations (20) and (21). Let the limiting distribution be denoted by $F_{\sigma,\alpha}$.

We now distinguish two cases: (a) $\sigma \geq 1 - \alpha$ and (b) $\sigma < 1 - \alpha$. Let us examine case (b) first. We note that the union of the image sets $F_1[0, 1] + F_2[0, 1]$ of F_1 and F_2 applied to the unit interval does not overlap with the open subinterval $(\sigma, 1 - \alpha)$. Any two applications of F_1 and F_2 leaves empty the two additional open intervals $(\sigma^2, (1 - \alpha)\sigma)$ and $(\sigma(1 - \alpha), (1 - \alpha)^2)$. Proceeding in this way, we find that the limit of the total set covered by n applications of $F_i (i = 1, 2)$ in any arrangement is a Cantor set C. It is easily seen that $F_{\sigma,\alpha}$ must concentrate its full probability on this set C.

Now let

$$\pi_{t_0}(x) = \left\{ \begin{array}{ll} 1 & \text{if } x = t_0 \\[2mm] 0 & \text{if } x \neq t_0 \end{array} \right\}.$$

We show that $U^n \pi_{t_0}(x)$ converges uniformly to zero. Note that $U\pi_{t_0}(t)$ is zero for every t except at most one value of t; namely, $F_1^{-1} t_0$ or $F_2^{-1} t_0$. Of course, if $\sigma < t_0 < 1 - \alpha$, then neither inverse exists for that t_0; and otherwise only one exists and

$$| U\pi_{t_0} | \leq \max_x [\phi(x), 1 - \phi(x)] \leq 1 - \delta.$$

Similarly, $U^n \pi_{t_0} \leq (1 - \delta)^n$, from which the assertion follows. We now observe that

$$(\pi_{t_0}, F_{\sigma,\alpha}) = (\pi_{t_0}, T^n F_{\sigma,\alpha}) = (U^n \pi_{t_0}, F_{\sigma,\alpha}) \longrightarrow 0.$$

Consequently, the probability of $F_{\sigma,\alpha}$ at t_0 is zero for any t_0 with $0 \leq t_0 \leq 1$. Summing up, we have established:

THEOREM 38. *If $\sigma < 1 - \alpha$, then the limiting distribution $F_{\sigma,\alpha}$ is a singular distribution (probability zero at every point) spread on a Cantor-like set.*

We turn now to examine case (a) where $\sigma \geq 1 - \alpha$. We note first that at least one of the two mappings F_1^{-1} or F_2^{-1} is defined for every x in the unit interval. Let $\pi(t)$ denote any continuous positive function defined on the unit interval so that $\pi(t) \geq \eta > 0$ for some subinterval $t_0 - h \leq t \leq t_0 + h$ $(h > 0)$.

Since at least F_1^{-1} or F_2^{-1} exists at t_0 (say F_1^{-1}), we obtain $F_1^{-1} t_0 = t_1$. We construct t_2 from t_1 in the same way and continue this for n steps, obtaining $t_n = F^{-n} t_0$, where F^{-n} denotes a specific order of application of F_1^{-1} or F_2^{-1} a total of n times. Let F^n denote the reverse order of the operators obtained by passing from t_0 to t_n. We note that

$$| F^n x - F^n y | \le \lambda^n | x - y | \le \lambda^n ,$$

where $\lambda < 1$. Choose n so large that $\lambda^n < h$; then for every x we get that

$$| F^n x - F^n t_n | = | F^n x - t_0 | < h .$$

Consequently, as

$$1 > 1 - \delta \ge \phi(x) \ge \delta > 0 ,$$

$U^n \pi$ is positive for all x since $F^{-n}[t_0 - h, \, t_0 + h]$ covers the entire unit interval and $\pi(t) \ge \eta > 0$ on this initial interval which is spread out by the term in U^n involving F^n. We have thus shown:

THEOREM 39. *If $\sigma > 1 - \alpha$, the operator U is strictly positive; that is, for each positive continuous function $\pi(t)$ there exists an n depending upon π so that $U^n \pi$ is strictly positive.*

Now let $\pi_{t_0}(t)$ be defined as before. Again we establish that $U^n \pi_{t_0}$ converges uniformly to zero. To this end we observe that $U \pi_{t_0}$ has at most two possible values at $F_1^{-1} t_0$ and $F_2^{-1} t_0$ given by $1 - \phi(F_1^{-1} t_0)$ and $\phi(F_2^{-1} t_0)$, respectively, while $U \pi_{t_0} \equiv 0$ elsewhere. Also, $U^2 \pi_{t_0}$ has at most four possible values and the maximum value that could be achieved for $U^2 \pi_{t_0}$ is

$$\max \{ [1 - \phi(F_1^{-1} t_0)][1 - \phi(F_1^{-2} t_0)], \ \phi(F_2^{-1} t_0) \phi(F_2^{-2} t_0),$$

$$[1 - \phi(F_1^{-1} t_0)] \phi(F_2^{-1} F_1^{-1} t_0) + \phi(F_2^{-1} t_0) [1 - \phi(F_1^{-1} F_2^{-1} t_0)] \}.$$

To secure a bound for the maximum of $U^n \pi_{t_0}$, let us consider the same repeated-experiment model set up in the previous section. The conditional probabilities of success p_n at the nth trial satisfy the uniform inequalities $1 > 1 - \eta \ge p_n \ge \eta > 0$, where success in this case is taken to be an application of the impulse F_1 to the particle. It is readily seen by standard inequalities that the probability of securing k ($k \le n$) successes converges uniformly to zero as $n \longrightarrow \infty$. Moreover, it follows directly that $\max_k$ (probability of k successes) is a bound for

$U^n \pi_{t_0}$, and hence $U^n \pi_{t_0} \longrightarrow 0$. We deduce as before that $F_{\sigma,\alpha}$ has probability zero for every t. Thus the cumulative distribution of F is continuous. Let $F = F_1 + F_2$, where F_1 is absolutely continuous and F_2 is singular. Observing that the transition operator transforms absolutely continuous measures into absolutely continuous measures and singular measures into singular measures, we find that $TF_1 = F_1$ and $TF_2 = F_2$. However, as the fixed distribution is unique, we deduce that either F_1 or F_2 vanishes.

THEOREM 40. *If* $\sigma \geq 1 - \alpha$, *then the unique distribution* $F_{\sigma,\alpha}$ *is either absolutely continuous or singular. Furthermore,* $F_{\sigma,\alpha}$ *has positive measure in every open interval.*

Proof. We have demonstrated all the conclusions of the theorem but the last. Let $\pi(t)$ denote a continuous positive function bounded by 1, and zero outside an open interval I, and 1 on a closed subinterval I' of I. By virtue of Theorem 39 there exists an n such that $U^n \pi \geq \delta > 0$ for all t. We note that

$$(\pi, F_{\sigma,\alpha}) = (\pi, T^n F_{\sigma,\alpha}) = (U^n \pi, F_{\sigma,\alpha}) \geq \delta > 0.$$

But

$$\int_I dF_{\sigma,\alpha} \geq (\pi, F_{\sigma,\alpha}) \geq \delta > 0,$$

and the proof of the theorem is complete.

We close with the conjecture that when $\sigma \geq 1 - \alpha$, then $F_{\sigma,\alpha}$ is always absolutely continuous. An example where this is the case is furnished by $\phi(x) \equiv 1/2$, $\sigma = 1/2 = 1 - \alpha$, where $F_{\sigma,\alpha}(x) = x$.

REFERENCES

1. R. Bellman, T. Harris, and H. N. Shapiro, *Studies on functional equations occurring in decision processes*, RM 878, RAND Corporation, July, 1952.

2. R. R. Bush and C. F. Mosteller, *A mathematical model for simple learning*, Psych. Rev. 58 (1951), 313-323.

3. J. L. Doob, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc. 63 (1948), 393-421.

4. W. Feller, *An introduction to probability theory and its applications*, Wiley, New York, 1950.

5. M. M. Flood, *On game learning theory*, RM 853, RAND Corporation, May 30, 1952.

6. O. Onicescu and G. Mihoc, *Sur les chaînes de variables statistiques,* Bull. Sci. Math. 2, 59 (1935), 174 - 192.

7. W. Doeblin and R. Fortet, *Sur les chaînes à liaisons complètes,* Bull. Soc. Math. France 65 (1937), 132 - 148.

8. R. Fortet, (Thèse) *Sur l'iteration des substitutions algébrique linéaires à une infinite de variables,* Revista, No. 424, Año 40, Lima, 1938.

9. Ionescu Tulcea and G. Marinescu, *Sur certaines chaînes à liaisons complètes,* C. R. Acad. Sci. Paris, 227 (1948), 667 - 669.

CALIFORNIA INSTITUTE OF TECHNOLOGY