Pacific Journal of Mathematics

APPROXIMATION OF SEMI-GROUPS OF OPERATORS

HALE TROTTER

Vol. 8, No. 4 June 1958

APPROXIMATION OF SEMI-GROUPS OF OPERATORS

H. F. TROTTER

1. Introduction. The usual methods for numerically computing the solution of a partial differential equation consist in replacing the differential operators by difference operators which approximate them, and taking the solution of the resulting difference equation as an approximation to the solution of the original equation. The question of *convergence* then arises; that is, when will a sequence of difference equations have the property that their solutions converge to the solution of a given differential equation? We treat this question in an operator-theoretic fashion, and our discussion has much in common with that of Lax and Richtmyer [17], as is pointed out in more detail below. The reader is referred to the bibliography of [17] for a list of the principal papers dealing with this question of convergence.

Our discussion will be limited to the initial value problem (Cauchy problem) for linear equations in the form

$$\frac{\partial}{\partial t}u(t, x) = \Omega u(t, x)\,;\, u(0, x) = f(x) \tag{1.1}$$

in which Ω is linear and constant in time. Here t is a non-negative real variable, and x is a point in some space S. Equation (1.1) is formally of parabolic type, but, as is shown in [14, Chap. XX], the initial value problem for hyperbolic equations can also be put into this form. Lateral conditions (i.e., boundary conditions such as are needed for the heat equation on a finite interval) are considered to be incorporated into the definition of Ω as restrictions on its domain [8, 17].

The process of setting up a sequence of finite difference approximations to (1.1) may be described in the following general terms. For each n, take a positive number h_n and a set $S_n \subset S$ whose points form a suitable grid. The solution to the nth approximating equation is defined inductively, for t an integral multiple of h_n and x a point of S_n by the following system of equations:[1]

[1] Sometimes only the space variable is made discrete, so that u_n is defined by a finite set of simultaneous *differential* equations (cf. [**13**, p. 233]). Theorem 5.2 can be applied to this situation just as Theorem 5.3 can be applied to the case in which the time variable is made discrete and the u_n are defined by (1.2).

This paper was originally accepted by the Transactions of the American Mathematical Society, Received July 3, 1957, by Trans. Amer. Math. Soc. Research done at Princeton University and supported in part by the Office of Ordnance Research U.S. Army.

$$u_n((k+1)h_n, x) = T_n u(kh_n, x) \qquad k = 0, 1, 2, \cdots \tag{1.2}$$

$$u_n(0, x) = f_n(x)$$

where T_n is a linear operator and f_n is a function defined on S_n which suitably approximates f. For example, f_n may be simply the restriction of f to S_n. (Other possible ways of defining f_n are discussed in § 2.) It will be convenient to extend the definition of u_n to all t by setting

$$u_n(t, x) = u_n(kh_n, x) \text{ for } kh_n \leqq t < (k+1)h_n \,. \tag{1.3}$$

In some numerical methods, the relation between the values of u_n at steps k and $k+1$ is given by a set of simultaneous linear equations, so that T_n is defined implicitly rather than by explicit formulae; the particular way in which T_n may be defined will be irrelevant to our discussion.

If the operators Ω_n defined by

$$\Omega_n = h_n^{-1}(T_n - I) \tag{1.4}$$

converge to Ω in some suitable sence, then (1.2) converges formally to (1.1) as $n \to \infty$, and it is plausible that under certain conditions the functions u_n will converge to the solution of (1.1).

It was observed by von Neumann [20] that a system like (1.2) may be *unstable* in the sense that small errors in the initial data may lead to errors in $u_n(t)$ which become unbounded as $n \to \infty$. The definitions of stability given in the literature vary slightly in detail. We adopt essentially the same definition as that used in [17]. We suppose that the space of functions on S_n is normed as a Banach space. (Examples of how such a norm may be defined are given in § 2.) Then the norm of T_n as an operator on this space is defined. We shall say that the system (1.2) is *stable* if

$$\| T_n^k \| \leqq M e^{Kkh_n} \tag{1.5}$$

for some constants K and M independent of n and k. The simpler condition $\| T_n \| \leqq 1$ is satisfied in many applications, and clearly implies (1.5).[2]

Although it is possible to find an example of an unstable system whose solutions converge to the correct result if the approximating functions f_n are appropriately chosen [18], Lax and Richtmyer [17] have shown that in general an unstable system cannot converge. On the other hand, they have shown that stability, together with some reasonable

[2] It should be pointed out that we are concerned only with the behaviour of the *exact* solutions of (1.2). In actual computation the effect of round-off errors must be considered, and the situation becomes more complicated.

assumptions on the limiting behaviour of the Ω_n, is sufficient to imply convergence. Our main result is very similar; however, our hypotheses differ in two respects from those of [17].

In the first place, we do not require that the limit function u and the approximating functions u_n all belong to the same Banach space. Frequently, the T_n arise most naturally as operators on functions which are defined only at a grid of points in the space S. In most practical applications the T_n can be modified so as to become operators on functions defined on the whole space, and this is assumed in [17]. Such modification, however, is usually unnatural, and in the case of random walks and diffusion processes which we discuss in § 6 it is not necessarily possible. Consequently it seems worth while to eliminate this assumption. We introduce the notion of an *approximating sequence of Banach spaces* and define associated concepts of convergence of vectors and operators. Section 2 is devoted to setting up the definitions and giving examples; it also includes some lemmas on the convergence of operators. Section 3 contains some remarks on the adjoint spaces of an approximating sequence which have application in § 6. Kantorovich [16] uses a similar sort of approximation of one Banach space by another, but requires the approximating space to be isomorphic to a subspace of the approximated space. This requirement is unnecessarily restrictive for our purposes.

To explain the second difference, we must describe more precisely what we mean by a solution of (1.1). We give an abstract formulation in terms of a semi-group of operators [14, chap. 20; 15]. Let X be some suitable Banach space of functions on S, and let Ω be a linear operator on X. Suppose that Ω is densely defined and has the property that for every f in its domain there exists a unique function $u(t, x)$ satisying

$$\text{(1.6)} \qquad \text{(i)} \quad u(t,x) \text{ and } \frac{\partial}{\partial t} u(t,x) \text{ are in } X \text{ for all } t \geqq 0$$

$$\text{(ii)} \quad \frac{\partial}{\partial t} u(t,x) = \Omega\, u(t,x) \text{ for } t \geqq 0$$

$$\text{(iii)} \quad \| u(t,x) - f(x) \| \to 0 \text{ as } t \to 0$$

$$\text{(iv)} \quad \| u(t,x) \| \leqq M \| f(x) \| \text{ for } 0 \leqq t \leqq 1$$

Where M is a constant independent of f. Then for each t, setting

$$\text{(1.7)} \qquad [T(t)f](x) = u(t,x)$$

defines an operator on the domain of Ω. Since Ω is densely defined, $T(t)$ (which is bounded by $1 + M^{t+1}$) can be extended to all of X by continuity. The operators $T(t)$ then form a semi-group with Ω as

infinitesimal generator, which is (hypothesis (iii)) strongly continuous at the origin. The conditions (1.6) express the requirement that the initial value problem (1.1) be "well-posed" [cf. 17].

In [17] it is assumed that an operator Ω is given which leads to a well-posed problem, and the operators Ω_n are required to satisfy a consistency condition which may be translated into our terminology as follows:

There exists a dense set of functions f such that $\| (\Omega_n - \Omega)T(t)f \|$ tends to zero uniformly for t restricted to a bounded interval.

Our condition is that $\lim_{n\to\infty} \Omega_n$ be densely defined, and that for some positive λ, $\lim_{n\to\infty} (\lambda - \Omega_n)$ have a dense range. (The precise meaning we attach to "limit of a sequence of operators" is given in § 2.) It is part of our conclusion that the operator to which the Ω_n converge gives rise to a well-posed problem. Although our condition appears to be quite different from that of Lax and Richtmyer, we have found no example in which we could show that one condition was satisfied and the other was not. Our condition seems to be easier to verify in the case of the applications made in § 6.

Our proof of convergence is based on the relation between a semigroup and its resolvent which is made explicit in the Hille-Yosida theorem [14, 22]. In § 4 we develop the relevant facts in a form convenient for our discussion. All the results in this section are well-known; the proof of Theorem 4.1, however, is new.

Section 5 is concerned with the convergence problem proper, and the theorems of this section represent the principal results of the paper.

In § 6 we consider in some detail the convergence of random walks to one-dimensional diffusion processes, and discuss several examples.

This paper is based on a thesis submitted to the Department of Mathematics of Princeton University in June 1956. I wish to thank Professor W. Feller for his guidance in the writing of the thesis, and the National Research Council of Canada for fellowship support during the academic year 1955–56.

2. Approximating sequences of Banach spaces. Let X be a Banach space. A sequence of Banach spaces $\{X_n\}$ together with a sequence of linear maps $P_n: X \to X_n$ is called a *sequence*[3] *of Banach spaces approximating* X if

[3] We could equally well define a system of Banach spaces approximating X using an arbitrary directed set as index set. The proofs of all the main theorems of this paper require only trivial changes of language to adapt them to the more general situation. Unless X is assumed to be separable, it is necessary to use nets in discussing some of the properties of the adjoint approximating sequence (§ 3).

$$\|P_n\| \leqq 1 \tag{2.1}$$

and

$$\lim_{n\to\infty} \|P_n f\| = \|f\| \text{ for every } f \in X . \tag{2.2}$$

Condition (2.2) obviously implies that in some sense the maps P_n become isomorphisms "in the limit". The following lemma gives a precise expression to this idea.

LEMMA 2.1. *Let X' be any finite-dimensional subspace of X, and let Q_n be the restriction of P_n to X'. Then Q_n is one-to-one if n is sufficiently large, and* $\lim_{n\to\infty} \|Q_n^{-1}\| = 1$.

Proof. Take any $\varepsilon > 0$, and let $\{f_i\}$ be a finite set of vectors of unit length in X' such that the ε-neighbourhoods of the f_i cover the unit sphere in X'. Now take N sufficiently large that for all $n \geqq N$, $\max_i(\|f_i\| - \|P_n f_i\|) < \varepsilon$. Then for any g on the unit sphere in X', $\|P_n g\| > 1 - 2\varepsilon$ for all $n \geqq N$. Hence Q_n is one-to-one and

$$1 \leqq \|Q_n\|^{-1} \leqq \|Q_n^{-1}\| \leqq (1 - 2\varepsilon)^{-1} .$$

We now define convergence for sequences of vectors and operators. We shall use the following terminology and notation. By "operator on X" we shall mean "linear transformation defined on a linear subset of X and taking values in X". If A is an operator on X, the linear subset of X on which A is defined is the *domain* of A, written $\mathbf{D}(A)$. The *range* of A, denoted by $\mathbf{R}(A)$, is the linear subset consisting of all $g \in X$ for which there exists an $f \in \mathbf{D}(A)$ with $g = Af$. If A and B are two operators such than $\mathbf{D}(A) \subset \mathbf{D}(B)$ and $Af = Bf$ for all $f \in \mathbf{B}(A)$ then we call B an *extension* of A and write $A \subset B$. The identity operator on any space will be denoted by I.

A sequence $\{f_n\}$, where $f_n \in X_n$, *converges* to $f \in X$ if $\lim \|f_n - P_n f\| = 0$. It is easy to see that (2.2) implies that a sequence cannot converge to more than one $f \in X$. A sequence is *convergent* if there exists an $f \in X$ to which it converges; we call f the *limit* of $\{f_n\}$ and write $f = \lim_{n\to\infty} f_n$.

The *limit of a sequence of operators* $\{A_n\}$, where A_n is an operator on X_n, is the operator on X whose domain consists of those $f \in X$ for which $\{A_n P_n f\}$ converges and whose value for such an f is $\lim_{n\to\infty} A_n P_n f$.

EXAMPLES.

(1) Let X be an arbitrary Banach space, and for every n let $X_n = X$ and $P_n = I$. Then the convergence of vectors is ordinary

convergence, and convergence of operators is the usual strong convergence. Lemmas 2.2, 2.3, and 2.4 below then become results on the strong convergence of operators.

(2) Let X be the space of bounded continuous functions on some topological space S, with the uniform norm. For each n, let S_n be a subset of S and let X_n be some Banach space of functions on S_n (with the uniform norm) which contains all the restrictions of elements of X. For $f \in X$, define $P_n f$ to be the restriction of f to S_n. (Note that there is no requirement that the projection P_n map X *onto* X_n.) Condition (2.1) is obviously satisfied, and if the sets S_n become dense in S in the sense that every open set $U \subset S$ contains points of S_n for n sufficiently large, then (2.2) is also satisfied. In this case $\{f_n\}$ converges to f if and only if

$$\lim_{n\to\infty} \sup_{x\in S_n} |f_n(x) - f(x)| = 0 .$$

(3) Let S be a region of Euclidean space with Lebesgue measure, and let X be $L^p(S)$. For each n, let S be partitioned into measureable sets $S_{n,i}$, each with finite measure $m_{n,i} > 0$. Let X_n be the subspace of X consisting of functions constant on the $S_{n,i}$. For $f \in X$, define $P_n f$ to have the value $m_{n,i}^{-1} \int_{S_{n,i}} f(x)\,dx$. Condition (2.1) is always satisfied. If each partition is a refinement of the preceding one, and if the partitions become sufficiently fine in the sense that the Borel field generated by the collection of all the $S_{n,i}$ contains all Borel-measureable subsets of S, then (2.2) will be satisfied. It is clear that a similar procedure can be followed to get an approximating sequence to the L^p-space over any measure space. Essentially this type of approximation (with $p = 2$) has been used by Douglas [4].

The following lemmas will be needed for later use. Lemma 2.2 is a generalization of the Banach-Steinhaus lemma [1, p. 79], and Lemma 2.3 similarly generalizes an obvious fact about the strong convergence of operators. Lemma 2.4 is a more special result which is used in the proof of Theorem 5.2.

LEMMA 2.2. *For each n, let A_n be an operator on X_n, with domain all of X_n. If there exists a constant M such that $\| A_n \| \leqq M$ for all n, and if $A = \lim_{n\to\infty} A_n$ is densely defined, then A is defined on all of X and $\| A \| \leqq M$.*

Proof. By hypothesis, for any $f \in X$ it is possible find a sequence $\{f^j\}$ converging to f, with every f^j in $\mathbf{D}(A)$. This, by definition, means that for each f^j there exists a $g^j \in X$ such that

$\| A_n P_n f^j - P_n g^j \| \to 0$ as $n \to \infty$. For any i, j,

$$\| g^i - g^j \| = \lim_{n\to\infty} \| P_n(g^i - g^j)\|$$

$$= \lim_{n\to\infty} \| A_n P_n (f^i - f^j) \|$$

$$\leqq M \| f^i - f^j \| .$$

Since the f^i form a convergent sequence, it follows that the g^i form a Cauchy sequence and consequently have a limit g. Now

$$\| A_n P_n f - P_n g \| \leqq \| A_n P_n (f - f^i) \| + \| A_n P_n f^i - P_n g^i \| + \| P_n(g^i - g) \| .$$

The first and last terms on the right may be made arbitrarily small by taking i sufficiently large, independent of n, and the middle term goes to zero as $n \to \infty$. Hence Af is defined and equal to g. It is obvious that $\| A \| \leqq M$.

LEMMA 2.3. *Let $\{A_n\}$ be a sequence of operators satisfying the hypotheses of Lemma* 2.2, *and converging to A. For each n, let B_n be an operator on X_n and let $B = \lim_{n\to\infty} B_n$. Then $AB \subset \lim_{n\to\infty} A_n B_n$.*

Proof. Suppose $f \in \mathbf{D}(AB)$. Then

$$P_n ABf - A_n B_n P_n f = (P_n ABf - A_n P_n Bf) + A_n(P_n B - B_n P_n)f.$$

The first term on the right tends to zero because $A = \lim_{n\to\infty} A_n$. The second term on the right is dominated in norm by $M \| P_n Bf - B_n P_n f \|$ and this tends to zero because $f \in \mathbf{D}(AB) \subset \mathbf{D}(B)$.

COROLLARY. *Let $\{A_n\}$ be a sequence of operators satisfying the hypotheses of Lemma* 2.2 *and converging to A. Then for any positive integer k, $\lim_{n\to\infty} A_n^k = A^k$.*

LEMMA 2.4. *For each n, let A_n be an operator on X_n with an inverse B_n defined on all of X_n. Suppose that $\| B_n \| \leqq M$ for all n, and that both the domain and range of $A = \lim_{n\to\infty} A_n$ are dense in X. Then $B = \lim_{n\to\infty} B_n$ is defined on all of X, and has a dense range. B has an inverse if and only if A has a closed extension, and then B^{-1} is the closure of A.*

Proof. Consider an arbitrary $g \in \mathbf{R}(A)$, $g = Af$.
Then

$$\| B_n P_n g - P_n f \| = \| B_n(P_n g - A_n P_n f) \| \leqq M \| P_n g - A_n P_n f \| .$$

Since $g = \lim_{n\to\infty} A_nP_nf$, the term on the right goes to zero, and consequently $B_nP_ng \to f$. Since $\mathbf{R}(A)$ is dense, it follows from Lemma 2.2 that $B = \lim_{n\to\infty} B_n$ is defined on all of X. Since $Bg = f$ if $g = Af$, $BA \subset I$. Hence $\mathbf{R}(B) \supset \mathbf{D}(A)$ and is dense in X.

If B^{-1} exists it is a closed operator because B is closed; since $BA \subset I$, it is an extension of A. Conversely, suppose A has a closed extension A'. Then, since B is bounded, $A'B$ is a closed extension of AB. But for $g = Af \in \mathbf{R}(A)$, we have $ABg = AB(Af) = A(BA)f = Af = g$, so that AB coincides with I on a dense set. Hence $A'B = I$; therefore B has an inverse, and A' is an extension of it. This is true for *any* closed extension of A, and consequently B^{-1} is the closure of A.

3. Adjoint spaces. If $\{X_n\}$ is a sequence approximating X, with associated projection operators P_n, then the adjoint spaces X_n^* are connected with X_n^* by the adjoint operators $P_n^* : X_n^* \to X^*$. Condition (2.1) of course implies that $\|P_n^*\| \leqq 1$. Condition (2.2) clearly should imply that in some sense the images of the P_n^* become dense in X^* in the limit. This is not generally true in terms of the norm topology on X^*—it is easy to give an example in which the union of the images of the P_n^* is nowhere dense in X^* with respect to this topology. It is more appropriate to consider the weak topology [1] (often called the weak* topology) on X^*. For any $f^* \in X^*, f_1, \cdots, f_m \in X$, and $\varepsilon > 0$ we write $V_{f^*}(f_1, \cdots, f_m; \varepsilon)$ for the set $\{g^* \in X^* : |f^*(f_i) - g^*(f_i)| < \varepsilon, i = 1, 2, \cdots, m\}$. The collection of the "cubical neighbourhoods" forms a base for the neighbourhoods of f^* in the weak topology.

LEMMA 3.1. *Let f^* be any element of X^* and $V = V_{f^*}(f_1, \cdots, f_m; \varepsilon)$ a cubical neighbourhood of it. Then for all sufficiently large n, there exists an $f_n^* \in X_n^*$, with $\|f_n^*\| \leqq 2\|f^*\|$ and $f_n^*P_n \in V$.*[4]

Proof. Let X' be the subspace of X spanned by $f_1, \cdots, f_m$, and as in Lemma 2.1, let Q_n be the restriction of P_n to X'. Let $f^{*\prime}$ be the restriction of f^* to X'. Let n be sufficiently large that Q_n^{-1} exists, and define $f_n^* = f^{*\prime}Q_n^{-1}$. It follows directly from the definitions that $f_n^*P_nf_i = f^*(f_i)$ for any one of the f_i, so that $f_n^*P_n \in V$. Since $\|Q_n^{-1}\| \to 1$, the condition on the norm of f_n^* is satisfied for sufficiently large n.

If $\{f_\alpha^*\}$ is a net with $f_\alpha^* \in X_{n_\alpha}^*$ for each α, we say that $\{f_\alpha^*\}$ *converges weakly to* f^* if $n_\alpha \to \infty$ and $\{f_\alpha^*P_{n\alpha}\}$ converges weakly to f^*. The net $\{f_\alpha^*\}$ is said to be *bounded* if $\|f_\alpha^*\|$ is bounded uniformly with

[4] We write adjoint operators on the right when applied to a vector (dropping the "star" since the position of the operator indicates that it is the adjoint).

respect to α.

From Lemma 3.1 we obtain at once the proposition: for every $f^* \in X^*$ there exists a bounded net $\{f_\alpha^*\}$ with $f_\alpha^* \in X_{n_\alpha}^*$ which converges weakly to it. To construct such a net, let the index set consist of the cubical neighbourhoods of f^*, ordered by inclusion. For each such neighbourhood α, we can pick an $f_\alpha^* \in X_{n_\alpha}^*$, $n_\alpha > \varepsilon^{-1}$ (where ε is the positive number used to define the cubical neighbourhood α), with $\|f_\alpha^*\| \leqq 2\|f^*\|$ and $f_\alpha^* P_{n_\alpha} \in \alpha$. This net converges to f^*. If X is *separable* there exists a uniformly bounded *sequence* $\{f_n^*\}$ with $f_n^* \in X_n^*$ which converges weakly to f^*. To show this, let $f_1, f_2, \cdots$ be a sequence which is dense in X, and let X_m' be the subspace of X spanned by the first m vectors of the sequence. Using the construction of Lemma 3.1 we can define N_m inductively so that $N_{m+1} > N_m$, and for $N_m \leqq n < N_{m+1}$ there exists an $f_n^* \in X_n^*$ with $\|f_n^*\| \leqq 2\|f^*\|$ and $f_n^* P_n f_i = f^*(f_i)$ for $i < m$. Then $\|f_n^* P_n\|$ is uniformly bounded and $\lim_{n\to\infty} f_n^* P_n f_i = f^*(f_i)$ for all the f_i. Hence by a theorem of Banach [1, Theorem 2 of Chapter 7], $\{f_n^* P_n\}$ converges weakly to f^*.

For each n let B_n be an operator on X_n^*. The *weak limit* of the sequence $\{B_n\}$ is defined to be an operator B on X^*, with domain consisting of all $f^* \in X^*$ such that for every bounded net $\{f_\alpha^*\}$ with $f_\alpha^* \in X_{n_\alpha}^*$ which converges weakly to f^*, the net $\{f_\alpha^* B_{n_\alpha} P_{n_\alpha}\}$ converges weakly to a unique limit. For $f^* \in \mathbf{D}(B)$, f^*B is defined to be this limit.

LEMMA 3.2. *For each n let A_n be a bounded operator on X_n with domain all of X_n, and suppose that $A = \lim A_n$ is defined on all of X. Then the adjoint operators A_n^* converge weakly to A^*.*

Proof. Take any $f^* \in X^*$, and let $\{f_\alpha^*\}$ be any bounded net converging to it weakly. We must show that for every $f \in X$. $\{f_\alpha^* A_{n_\alpha} P_{n_\alpha} f\}$ converges to f^*Af. We have

$$f_\alpha^* A_{n_\alpha} P_{n_\alpha} f - f^*Af = f_\alpha^*(A_{n_\alpha} P_{n_\alpha} - P_{n_\alpha} A)f + (f_\alpha^* P_{n_\alpha} - f^*)Af .$$

The first term on the right goes to zero because $n_\alpha \to \infty$, $A = \lim_{n\to\infty} A_n$, and the f_α^* are uniformly bounded. The second term on the right goes to zero because $\{f_\alpha^*\}$ converges weakly to f^*.

4. Semi-groups and resolvents. Throughout this section we shall be dealing with a fixed Banach space X. Convergence of operators is to be interpreted as strong convergence.

A (*one-parameter*) *semi-group of operators* is a family $\{T(t)\}$ of

bounded operators on X, t ranging over the non-negative real numbers, which satisfies the relation

$$T(t+s) = T(t)T(s) \quad t, s \geqq 0 , \tag{4.1}$$

with $T(0)$ defined to be I. We shall consider only semi-groups for which

$$\lim_{t\to 0} T(t) = I \tag{4.2}$$

and

$$\| T(t) \| \leqq M \qquad \text{all } t \tag{4.3}$$

for some constant M. Semi-groups satisfying (4.2) and (4.3) will be called *proper*.[5]

The operator

$$\Omega = \lim_{t\to 0} t^{-1}(T(t) - I) \tag{4.4}$$

is the *infinitesimal generator* of the semi-group and is always closed and densely defined. One has $g = \Omega f$ if and only if

$$T(t)f = f + \int_0^t T(s)g \, ds . \tag{4.5}$$

For every $\lambda > 0$ there is a bounded operator

$$J(\lambda) = (\lambda - \Omega)^{-1} = \int_0^\infty e^{-\lambda t} T(t) \, dt . \tag{4.6}$$

The family $\{J(\lambda)\}$ is called *the resolvent family of* $T(t)$. The operators $J(\lambda)$ satisfy the relations

$$J(\lambda) - J(\mu) = (\mu - \lambda)J(\lambda)J(\mu) \tag{4.7}$$

$$\| \lambda^m J^m(\lambda) \| \leqq M \qquad \lambda > 0, m = 1, 2, \cdots \tag{4.8}$$

$$\lim_{\lambda\to\infty} \lambda J(\lambda) = I . \tag{4.9}$$

A family of operators satisfying (4.7) will be called a *resolvent family*, and one which satisfies (4.8) and (4.9) as well, a *proper* resolvent family.

If $\{J(\lambda)\}$ is a resolvent family it is clear that any f annihilated by one of the $J(\lambda)$ is annihilated by all of them, and also that $\mathbf{R}(J(\lambda))$ is independent of λ. These remarks, together with (4.9), show that *the operators of a proper resolvent family are one-to-one transformations with*

[5] Condition (4.3) is not a serious restriction. If a semi-group satisfies (4.2) then $\| T(t) \|$ must be bounded for t near zero. It is then well-known [**14**] that $\| T(t) \| \leqq Me^{Kt}$ for some K, so that the closely related semi-group $T'(t) = e^{-Kt}T(t)$ will be proper. We use this trick in proving Theorems 5.2 and 5.3.

dense range. Hence for each λ, $\{J(\lambda)\}^{-1}$ is a densely defined operator. It is easy to show from (4.7) that

$$\Omega = \lambda - \{J(\lambda)\}^{-1} \tag{4.10}$$

is independent of λ; we call it *the infinitesimal generator associated with the resolvent family* $\{J(\lambda)\}$.

A proper semi-group is uniquely determined by its resolvent family. Suppose $\{T(t)\}$ and $\{T'(t)\}$ are two proper semi-groups with the same resolvent family $\{J(\lambda)\}$. For any $f \in X$, $f^* \in X^*$ the function $f^*(J(\lambda)f)$ will be the Laplace transform of both $f^*(T(t)f)$ and $f(T'(t)f)$. The classical uniqueness theorem for the Laplace transform [21, p. 63] then implies the identity of the last two functions (since both are bounded and continuous), and since f and f^* are arbitrary it follows that $T(t) = T'(t)$. This fact shows that $\{J(\lambda)\}$ is the resolvent family of the semi-group $T(t)$ if and only if the operator Ω defined by (4.10) is the same as that defined by (4.4).

The question still remains as to whether every proper resolvent family is the resolvent family of some proper semi-group. The Hille-Yosida theorem [14, 22] provides an affirmative answer. Our next theorem gives an expression for the semi-group in terms of the resolvent. (cf. [14, p. 234]; the proof given there depends on the Post-Widder inversion formula for the Laplace transform.)

THEOREM 4.1. *Let $\{J(\lambda)\}$ be a proper resolvent family. Then the operators*

$$T(t) = \lim_{\lambda \to \infty} \{\lambda J(\lambda)\}^{[\lambda t]} \qquad t \geqq 0$$

are defined on all of X and form a proper semi-group which has $\{J(\lambda)\}$ as its resolvent family.

We first prove several lemmas.

LEMMA 4.1 *Let A and B be two operators which commute and have the property that $\| A^i \|, \| B^i \| \leqq M$ for all positive integers i.*

Then, for any f,

$$\| (A^n - B^n)f \| \leqq nM^2 \| (A - B)f \| .$$

Proof. Since A and B commute

$$(A^n - B^n)f = \sum_{i=0}^{n-1} A^{n-i-1}B^i(A - B)f .$$

The right hand side contains n terms, each with norm less than or equal to $M^2 \| (A - B)f \|$.

LEMMA 4.2. *For any* $f \in \mathbf{D}(\Omega)$ (where Ω is defined by (4.10))

$$\| (\lambda J(\lambda) - I)f \| \leqq \lambda^{-1} M \| \Omega f \| .$$

Proof. This follows from (4.8) since for $f \in \mathbf{D}(\Omega)$,

$$(\lambda J(\lambda) - I)f = J(\lambda)\, \Omega f . \tag{4.11}$$

LEMMA 4.3. *For any* $f \in \mathbf{D}(\Omega^2)$

$$\| \{\lambda J(\lambda) - (2\lambda J(2\lambda))^2\} f \| \leqq (2\lambda)^{-2} M^2 \| \Omega^2 f \| .$$

Proof. Note that

$$J(\lambda) - J(2\lambda) = \lambda J(\lambda) J(2\lambda) \tag{4.12}$$

is a special case of (4.7). Now

$$\begin{aligned}
\lambda J(\lambda)f - (2\lambda J(2\lambda))^2 f &= (\lambda J(\lambda) - I)f - (2\lambda J(2\lambda) + I)(2\lambda J(2\lambda) - I)f \\
&= J(\lambda)\Omega f - (2\lambda J(2\lambda) + I) J(2\lambda)\Omega f && \text{by (4.11)} \\
&= \{J(\lambda) - J(2\lambda) - 2\lambda J^2(2\lambda)\}\Omega f \\
&= J(2\lambda)\{\lambda J(\lambda) - 2\lambda J(2\lambda)\}\Omega f && \text{by (4.12)} \\
&= J(2\lambda)\{\lambda J(\lambda) - I - 2\lambda J(2\lambda) + I\}\Omega f \\
&= J(2\lambda)\{J(\lambda) - J(2\lambda)\}\Omega^2 f && \text{by (4.11)} \\
&= \lambda J(\lambda) J^2(2\lambda)\Omega^2 f && \text{by (4.12)} .
\end{aligned}$$

The conclusion follows from this identity and (4.8).

Proof of Theorem 4.1. Let r be an arbitrary positive number which will be assumed fixed throughout the following discussion. Write r_n as an abbreviation for $2^n r$. Define

$$T(n, t) = \{r_n J(r_n)\}^{[tr_n]} . \tag{4.13}$$

From the definition we have

$$\begin{aligned}
T(n+1, t) &= \{2r_n J(2r_n)\}^{[2tr_n]} \\
&= \{2r_n J(2r_n)\}^{2[tr_n]+\varepsilon}
\end{aligned}$$

where $\varepsilon = [2tr_n] - 2[tr_n] = 0$ or 1. Thus

$$\begin{aligned}
T(n+1, t)f - T(n, t)f = {} & \{2r_n J(2r_n)\}^{2[tr_n]}(\{2r_n J(2r_n)\}^{\varepsilon} - I)f \\
& + (\{2r_n J(2r_n)\}^{2[tr_n]} - \{r_n J(r_n)\}^{[tr_n]})f .
\end{aligned}$$

Estimating the first term on the right by Lemma 4.2 and the second term by Lemmas 4.1 and 4.3, we obtain, for $f \in \mathbf{D}(\Omega^2)$,

$$(4.14)\quad \|\{T(n+m,t)-T(n,t)\}f\| \leqq \sum_{k=0}^{m-1} \|\{T(n+k+1,t)-T(n+k,t)\}f\|$$
$$\leqq \sum_{k=0}^{\infty} r^{-1}2^{-n-k}(M\|\Omega f\| + tM^4\|\Omega^2 f\|)$$
$$\leqq 2r_n^{-1}(M\|\Omega f\| + tM^4\|\Omega^2 f\|)\ .$$

Since the right-hand side goes to zero as $n \to \infty$, the vectors $T(n,t)f$ form a Cauchy sequence and therefore converge. It should be remarked that the convergence is uniform over every bounded t-interval. $\mathbf{D}(\Omega^2)$ is dense in X since it includes $\mathbf{R}(J^2(\lambda))$, and the latter is dense because of (4.9). From (4.8), $\|T(n,t)\| \leqq M$ for all n. Hence, by the Banach-Steinhaus theorem

$$(4.15)\qquad T(t) = \lim_{n\to\infty} T(n,t)$$

is everywhere defined and satisfies (4.3). Taking $n=0$ in (4.14) and letting $m \to \infty$, we obtain the estimate

$$(4.16)\qquad \|\{rJ(r)\}^{[tr]}f - T(t)f\| \leqq 2r^{-1}(M\|\Omega f\| + tM^4\|\Omega^2 f\|$$

for any $f \in \mathbf{D}(\Omega^2)$.

For any n,

$$T(n,s+t) - T(n,s)T(n,t) = T(n,s)T(n,t)(\{r_nJ(r_n)\}^{\epsilon} - I)\ ,$$

where $\varepsilon = [r_n(s+t)] - [r_ns] - [r_nt] = 0$ or 1.
Consequently (Lemma 4.2)

$$(4.17)\qquad \|\{T(n,s+t) - T(n,s)T(n,t)\}f\| \leqq r_n^{-1}M\|\Omega f\|$$

for any $f \in \mathbf{D}(\Omega)$. Taking the limit as $n \to \infty$ in (4.17) shows that the operators defined by (4.15) satisfy (4.1) on a dense subset of X, and hence on all of X, by continuity.

From Lemmas 4.1 and 4.2 we get

$$(4.18)\qquad \|\{T(n,t) - I\}f\| \leqq r_n^{-1}[r_nt]M^3\|\Omega f\| \leqq tM^3\|\Omega f\|\ .$$

This inequality must hold also in the limit and shows that $\lim_{t=0} T(t)f = f$ for all f in the dense set $\mathbf{D}(\Omega)$. By the Banach-Steinhaus theorem it follows that the operators defined by (4.15) satisfy (4.2).

This completes the proof that the operators $T(t)$ which we have constructed form a proper semi-group. The construction, however, depended on the choice of a number r, and we still have to show that the result is independent of this choice. We shall show that the semi-group has the given family $\{J(\lambda)\}$ as its resolvent family, and since the resolvent family of a semi-group determines it uniquely, it will follow that the result of our construction is independent of r.

We show that $\{T(t)\}$ has the original $\{J(\lambda)\}$ as resolvent family by demonstrating that the operator Ω defined by (4.10) is the infinitesimal generator of the semi-group as defined by (4.4). Suppose $f \in \mathbf{D}(\Omega)$ and $g = \Omega f$. For convenience, let t be such that $r_n t$ is an integer for sufficiently large n. Such values of t are dense in the line. Then

$$\begin{aligned}
\int_0^t T(n, s)g\, ds &= \int_0^t \{r_n J(r_n)\}^{[r_n s]}\, \Omega f\, ds \\
&= r_n^{-1} \sum_{k=1}^{k=tr_n} \{r_n J(r_n)\}^{k-1} \Omega f \\
&= \sum_{k=1}^{k=tr_n} \{r_n J(r_n)\}^{k-1} J(r_n)\Omega f \\
&\quad + r_n^{-1}\{\Omega f - T(n, t)\Omega f\} \\
&= \sum_{k=1}^{k=tr_n} \{r_n J(r_n)\}^{k-1}\{r_n J(r_n) - I\} f \\
&\quad + r_n^{-1}\{\Omega f - T(n, t)\Omega f\} \\
&= T(n, t)f - f + r_n^{-1}\{\Omega f - T(n, t)\Omega f\}\ .
\end{aligned}$$

Letting $n \to \infty$ we obtain formula (4.5). (Passage to the limit under the integral sign is justified since $T(n, t)f$ converges uniformly on every bounded t-interval.) The number t was any one of a dense set, and by continuity, (4.5) must in fact hold for all t. This shows that $g = \Omega f$ with Ω defined by (4.4), and the proof of Theorem 4.1 is complete.

5. Convergence of Semi-groups. Throughout this section $\{X_n\}$ will be a sequence of Banach spaces approximating X, with associated projections P_n. We shall use the notational convention that vectors with subscript n are elements of X_n, and operators with subscript n are operators on X_n; vectors and operators without subscript will be associated with X.

A sequence of semi-groups $\{T_n(t)\}$ or resolvent families $\{J^n(\lambda)\}$ will be said to be *uniformly proper* if each member of the sequence is proper, and the constant M in conditions (4.3) or (4.8) may be taken independent of n.

THEOREM 5.1. *Let $\{T_n(t)\}$ be a uniformly proper sequence of semi-groups, and $\{J_n(\lambda)\}$ the sequence of associated resolvent families. Then if the operators $J(\lambda) = \lim_{n-\infty} J_n(\lambda)$ form a proper resolvent family, the sequence $\{T_n(t)\}$ converges to $T(t)$, the proper semi-group having $\{J(\lambda)\}$ as resolvent family.*

Proof. Since the $T_n(t)$ are uniformly bounded, it will be sufficient, by Lemma 2.2, to prove that $T_n(t)f$ converges to $T(t)f$ for a set of f which is dense in X. It has already been remarked (in the paragraph following (4.9)) that any operator in a proper resolvent family has a

dense range. By similar considerations it is easy to show that the square of any operator in a proper resolvent family has a dense range. Hence we need only consider f such that $f = J^2(\mu)g$ for some g and μ. Define $g_n = P_n g$ and $f_n = J_n^2(\mu)g_n$. Letting Ω_n be the infinitesimal generator of $T_n(t)$ we have

$$\begin{aligned} (5.1) \qquad \| \Omega_n f_n \| &= \| (\mu J_n(\mu) - I) J_n(\mu) g_n \| \\ &\leqq \mu^{-1}(M+1) \| g_n \| \leqq \mu^{-1}(M+1) \| g \| \leqq K \end{aligned}$$

and

$$\begin{aligned} (5.2) \qquad \| \Omega_n^2 f_n \| &= \| (\mu J_n(\mu) - I)^2 g_n \| \\ &\leqq (M+1)^2 \| g_n \| \leqq (M+1)^2 \| g \| \leqq K \end{aligned}$$

for some sufficiently large constant K; we also have $\| \Omega f \|$, $\| \Omega^2 f \| \leqq K$. Now

$$\begin{aligned} P_n T(t) f - T_n(t) P_n f = {} & P_n (T(t) - \{rJ(r)\}^{[rt]}) f \\ & + P_n \{rJ(r)\}^{[rt]} f - \{rJ_n(r)\}^{[rt]} P_n f \\ & + \{rJ_n(r)\}^{[rt]} (P_n f - f_n) \\ & + (\{rJ_n(r)\}^{[rt]} - T_n(t)) f_n \\ & + T_n(t)(f_n - P_n f). \end{aligned}$$

Applying (4.16), (5.1), (5.2) and the uniform boundedness of the $T_n(t)$, this yields

$$\begin{aligned} \| P_n T(t) f - Tn(t) P_n f \| \leqq {} & 4r^{-1} K(M + tM^4) + 2M \| P_n f - f_n \| \\ & + \| P_n \{rJ(r)\}^{[rt]} f - \{rJ_n(r)\}^{[rt]} P_n f \| . \end{aligned}$$

For any fixed r, the last two terms go to zero as $n \to \infty$, because $f_n \to f$ and $\{rJ(r)\}^k = \lim_{n \to \infty} \{rJ_n(r)\}^k$ for any fixed k. Thus

$$\limsup_{n \to \infty} \| P_n T(t) f - T_n(t) P_n f \| \leqq 4r^{-1} K(M + tM^4) .$$

Since r may be taken arbitrarily large it follows that $T(t)f = \lim_{n \to \infty} T_n(t) P_n f$.

LEMMA 5.1. *Let $\{J_n(\lambda)\}$ be a uniformly proper sequence of resolvent families, such that for some positive μ, $\lim_{n \to \infty} J_n(\mu)$ is densely defined and has a dense range. Then for every λ, $J(\lambda) = \lim_{n \to \infty} J_n(\lambda)$ is defined on all of X, and $\{J(\lambda)\}$ is a proper resolvent family.*

Proof. That $J(\mu)$ is everywhere defined follows immediately from Lemma 2.2. To show that $\lim_{n \to \infty} J_n(\lambda)$ is everywhere defined we make

use of the relation [14, p. 119]

$$J_n(\mu - \nu) = \sum_{k=1}^{\infty} \nu^{k-1} J_n^k(\mu) \tag{5.3}$$

where, provided that $|\nu| < \mu$, the series converges in the uniform operator norm, uniformly in n. This follows from the formula

$$J_n(\mu - \nu) = \nu^m J_n^m(\mu) J_n(\mu\,\nu) + \sum_{k=1}^{m} \nu^{k-1} J_n^k(\mu)$$

which is easily derived from (4.7) by induction on m, condition (4.8), and the assumption that the $J_n(\lambda)$ are uniformly proper. For each k, $\lim_{n\to\infty} J_n^k(\mu)$ is everywhere defined, by the corollary to Lemma 2.3, and the uniform convergence of (5.3) implies that $\lim_{n\to\infty} J_n(\lambda)$ is everywhere defined for $2\mu > \lambda > 0$. Repetition of the argument, replacing μ by, say, $3\mu/2$, shows the convergence of $\{J_n(\lambda)\}$ for $3\mu > \lambda > 0$. By further repetitions of the argument it can be shown that $J(\lambda) = \lim_{n\to\infty} J_n(\lambda)$ is everywhere defined for all $\lambda > 0$. Relation (4.7) holds for each n, and by Lemma 2.3 it continues to hold in the limit. Condition (4.8) is clearly satisfied by every $J(\lambda)$. To complete the proof it is only necessary to demonstrate (4.9). Since any $f \in \mathbf{R}(J(\mu))$ is in $\mathbf{D}(\Omega)$, where Ω is defined by (4.10), it follows from Lemma 4.2 that (4.9) holds on $\mathbf{R}(J(\mu))$, which is dense by hypothesis. Since the operators $\lambda J(\lambda)$ are uniformly bounded, the conclusion follows by the Banach-Steinhaus theorem.

THEOREM 5.2. *Let $\{T_n(t)\}$ be a sequence of semi-groups satisfying* (4.2) *and the stability condition*

$$\| T_n(t) = \| \leqq M e^{Kt}$$

where M and K are independent of n and t. Let Ω_n be the infinitesimal generator of $T_n(t)$ and define $\Omega = \lim_{n\to\infty} \Omega_n$.

Suppose that

(i) *Ω is densely defined*

(ii) *for some $\lambda > K$, $\mathbf{R}(\lambda - \Omega)$ is dense in X.*

Then the closure of Ω is the infinitesimal generator of a semi-group $T(t)$ which satisfies (4.2), *and $T(t) = \lim_{n\to\infty} T_n(t)$.*

Proof. Define $T_n'(t) = e^{-Kt} T_n(t)$ and $\Omega_n' = \Omega_n - K$. Then Ω_n' is the infinitesimal generator of $T_n'(t)$ and the semi-groups $T_n'(t)$ from a uniformly proper sequence. Also $\Omega' = \lim_{n\to\infty} \Omega_n'$ is densely defined and $\mathbf{R}(\lambda - K - \Omega')$ is dense in X. The sequence $J_n'(\lambda - K) = (\lambda - K - \Omega_n')^{-1}$ is uniformly

bounded, From Lemma 2.4 it follows that $J'(\lambda - K)$ is everywhere defined and has a dense range. By Lemma 5.1, the operators $J'(\lambda) = \lim_{n\to\infty} J'_n(\lambda)$ form a proper resolvent family, which by Theorem 5.1 is the resolvent family of a semi-group $T'(t)$ such that $T'(t) = \lim_{n\to\infty} T'_n(t)$. Since $J'(\lambda)$ has an inverse, this inverse (by Lemma 2.4) is the closure of $\lambda - K - \Omega'$, and it follows that the closure of Ω' is the infinitesimal generator of $T'(t)$. The results stated for $T(t)$ follow immediately from what we have proved about $T'(t)$.

LEMMA 5.2. *Let h be a given positive number, and T an operator such that $\| T^k \| \leqq M$ for all n. Then $\Omega = h^{-1}(T - I)$ is the infinitesimal generator of a semi-group $S(t)$ such that $\| S(t) \| \leqq M$.*

Proof. Define

$$S(t) = \sum_{k=0}^{\infty} (k\,!)^{-1}(t\Omega)^k = e^{-th^{-1}} \sum_{k=0}^{\infty} (k\,!)^{-1}(th^{-1}T)^k .$$

That Ω is the infinitesimal generator of $S(t)$ can be verified by term-by-term differentiation of the first expression given. From the second expression we obtain

$$\| S(t) \| \leqq e^{-th^{-1}} \sum_{k=0}^{\infty} (k\,!)^{-1}(th^{-1})^k M \leqq M.$$

LEMMA 5.3. *Let h, T, Ω and $S(t)$ be as in Lemma 5.2. For a fixed t, let $k = [th^{-1}]$. Then for any f,*

$$\| S(t)f - T^k f \| \leqq hM^2(\tfrac{1}{2}t \| \Omega^2 f \| + \| \Omega f \|) .$$

Proof. Iteration of (4.5) gives

$$\begin{aligned} S(h)f &= f + \int_0^h S(t)\Omega f \, dt \\ &= f + h\Omega f + \int_0^h \int_0^s S(t)\Omega^2 f \, dtds \\ &= Tf + \int_0^h \int_0^s S(t)\Omega^2 f \, dtds . \end{aligned}$$

Hence $\| S(h)f - Tf \| \leqq \frac{1}{2}h^2 \| \Omega^2 f \|$. By Lemma 4.1

$$\| S(kh)f - T^k f \| \leqq \tfrac{1}{2}M^2 kh^2 \| \Omega^2 f \| \leqq \tfrac{1}{2}M^2 ht \| \Omega^2 f \|$$

when $k = [th^{-1}]$. Also

$$\| S(t)f - S(kh)f \| = \| S(kh)\{S(t - kh) - I\}f \| \leqq hM^2 \| \Omega f \| .$$

THEOREM 5.3. *Let $\{h_n\}$ be a sequence of positive numbers converging to zero, and $\{T_n\}$ a sequence of operators satisfying the stability condition*

$$|| T_n^k || \leqq Me^{Kkh_n}$$

where M *and* K *are constants independent of* n *and* k. *Let* $\Omega_n = h_n^{-1}(T_n - I)$ *and define* $\Omega = \lim_{n\to\infty} \Omega_n$. *Suppose that*

(i) Ω *is densely defined*

(ii) *for some* $\lambda > K$, $R(\lambda - \Omega)$ *is dense in* X.

Then the closure of Ω *is the infinitesimal generator of a semi-group* $T(t)$ *and*

$$T(t) = \lim_{n\to\infty} T_n[th_n^-] .$$

Proof. Define $T_n' = e^{-Kh_n}T_n$, so that for any k, $|| T_n'^k || \leqq M$. Then

$$\Omega_n' = h_n^{-1}(T_n' - I) = e^{-Kh_n}\Omega_n - h_n^{-1}(1 - e^{-Kh_n})$$

and

$$\Omega' = \lim_{n\to\infty} \Omega_n' = \Omega - K .$$

Hence Ω' is densely defined and $\mathbf{R}(\lambda - K - \Omega')$ is dense. Let $S_n(t)$ be the semi-group generated by Ω_n'. Lemma 5.2 shows that the hypotheses of Theorem 5.2 are satisfied and hence there is a proper semi-group $T'(t) = \lim_{n\to\infty} S_n(t)$, with infinitesimal generator the closure of Ω'. For a fixed t, define $k_n = [th_n^{-1}]$. We shall show that $\lim_{n\to\infty} T_n'^{h_n} = T'(t)$. As in the proof of Theorem 5.1 it suffices to show that $T'(t)f = \lim_{n\to\infty} T_n'^{k_n}P_nf$ for f of the form $f = J^2(\lambda)g$ for some g and λ. Define $f_n = J_n^2(\lambda)P_ng$; then as in the proof of Theorem 5.1 there exists a constant C such that $|| \Omega_n^2 f_n ||$, $|| \Omega_n f_n || \leqq C$ for all n. Then

$$\begin{aligned} || S_n(t) - T_n'^{k_n})P_nf || &\leqq S_n(t)(P_nf - f_n) || \\ &\quad + || (S_n(t) - T_n'^{k_n})f_n || \\ &\quad + || T_n'^{k_n}(f_n - P_nf) || \\ &\leqq 2M || f_n - P_nf || + h_nM^2C(t + 1) . \end{aligned}$$

Since $f_n \to f$ and $h_n \to 0$ this shows that

$$\lim_{n\to\infty} T_n'^{k_n}P_nf = \lim_{n\to\infty} S_n(t)P_nf = T'(t) .$$

Hence

$$\lim T_n^{k_n} = \lim e^{Kh_nk_n}T_n'^{k_n} = e^{Kt}T'(t) ,$$

which is a semi-group $T(t)$ with the closure of $\Omega' + K = \Omega$ as infinitesimal generator.

6. Random walks and diffusion processes. Throughout this section, S will denote either the compactified real line $[-\infty, \infty]$, or some closed

subinterval of it. Let S' be any Borel subset of S, and suppose that for every $x \in S'$ a positive Borel measure μ_x is given such that

(i) $\mu_x(S) \leqq 1$

(ii) $\mu_x(S - S') = 0$

(iii) $\mu_x(A)$ is a Borel function of x for every Borel set A.

We consider a particle which executes a *random walk* in the following way. Let h be a positive constant. If the particle is at a point $x \in S'$ at time kh (k a non-negative integer) then it remains at x during the interval $[kh, (k+1)h)$ and at time $(k+1)h$ "jumps" so that the probability that it goes to any Borel set A is $\mu_x(A)$. (If $\mu_x(S) < 1$ then the particle disappears with probability $1 - \mu_x(S)$.) The number h is the *basic time-interval* of the random walk, and the measures μ_x are the *one-step transition probability distributions*. The set S' is the *support* of the random walk. If a probability distribution ν on S' for the initial position of the particle is fixed, then the random walk gives rise to what is essentially a discrete parameter Markov process with stationary transition probabilities. For rigorous definition and further details [3, p. 190 ff.] may be consulted.

For any bounded Borel function f on S', the random walk determines a new function Tf defined by setting

$$(Tf)(x) = \int_{S'} f(y)\, d\mu_x(y) \qquad x \in S'.$$

Condition (iii) implies that Tf is again a Borel function. The *one-step transition operator* T is obviously linear and positivity preserving. The space of bounded Borel functions on S' is a Banach space under the norm $\|f\| = \sup_{x \in S'} |f(x)|$, and $\|T\| \leqq 1$ relative to this norm. The adjoint transformation may be considered as acting on the Borel measures on S', and for any such measure ν, νT is given by

$$(\nu T)(A) = \int T\chi_A \, d\nu$$

where A is an arbitrary Borel set and χ_A is its characteristic function. If ν is a probability measure giving the distribution of the initial position of the particle then νT^k is the distribution for its position after k jumps [cf. 3, p. 191]. Thus a random walk is completely characterized by its support, its basic time-interval, and its one-step transition operator, and we may speak of "the random walk $\{S', h, T\}$".

A *diffusion process* (with stationary transition probabilities) on S is characterized by giving for each $x \in S$ and $t > 0$ a measure $\mu_{x,t}$ such that $\mu_{x,t}(S) \leqq 1$, which is to be interpreted as the probability distribution for the position at time $s + t$ of a particle which is at x at time s. To

begin with, it is necessary to have $\mu_{x,t}(A)$ a Borel function of x for every t and every Borel set A [3, p. 255]. Then a family of operators $T(t)$ on the space of bounded Borel functions can be defined by setting

$$T(t)f(x) = \int f(y)\, d\mu_{x,t}(y) .$$

As in the case of a random walk, we may consider the adjoint transformation; for ν a Borel measure on S

$$(\nu T(t))(A) = \int T(t)\chi_A \, d\nu$$

for every Borel set A. If ν gives the distribution for the initial position of the particle then $\nu T(t)$ gives the distribution for its position at time t. The further conditions which the measures $\mu_{x,t}$ must satisfy are most easily expressed in terms of the operators $T(t)$, which we call the *transition operators* of the process. In the first place, we impose the condition that $\mu_{x,t}$ be a continuous function of x with respect to the weak topology on the measures; this is equivalent to requiring that $T(t)f$ be continuous when f is continuous (cf. [11]). Secondly, we would like to require that for any continuous f, $T(t)f$ converge uniformly to f as $t \to 0$, but certain processes (those with absorbing barriers) do not satisfy this condition. One or both end-points of S may be absorbing barriers; intuitively speaking, an end-point is an absorbing barrier if a diffusing particle disappears immediately when it reaches that point. In such a case, the range of $T(t)$ contains only functions which vanish at the absorbing barrier(s). Throughout the rest of this section, let X be the Banach space of *continuous functions on S which vanish at those end-points which are absorbing barriers* for the diffusion process under consideration. We shall require that $T(t)f$ converge uniformly to f as $t \to 0$ for all $f \in X$. For the process to have the Markov property, the operators $T(t)$ must have the semi-group property (4.1). It is obvious that $\| T(t) \| \leqq 1$. The conditions imposed so far may be summarized as:

The operators $T(t)$ form a proper semi-group of operators on the space X.

Finally, in order to restrict attention to diffusion processes rather than more general types of Markov process, we suppose that

The infinitesimal generator of the semi-group $T(t)$ is a restriction of an operator of the form

$$\Omega' = \frac{d}{dm}\frac{d}{dx} \tag{6.1}$$

where m is a strictly increasing function of x.

It has been shown by Feller [10, 11] that all one-dimensional diffusion

processes satisfying certain regularity conditions have associated semi-groups whose infinitesimal generators can be put into the required form by a suitable choice for the coordinate function x.

Suppose that for each n a random walk $\{S_n, h_n, T_n\}$ is given, and that a diffusion process with transition semi-group $T(t)$ is also given. We shall say that the sequence of random walks *converges* to the diffusion process if

(i) For every probability measure ν on S, there exists a sequence of probability measures ν_n converging weakly to it, with the support of ν_n contained in S_n.

(ii) For every such sequence ν_n, and every $t > 0$, the probability distribution for the position at time t of a particle starting with initial distribution ν_n and executing the nth random walk converges weakly to the probability distribution at time t for the position of a particle starting with initial distribution ν and following the diffusion process. The weak convergence referred to above is *weak convergence relative to the space* X.

Now let X_n be the Banach space of bounded Borel functions on S_n, and define $P_n : X \to X_n$ by taking $P_n f$ to be the restriction of f to S_n for any $f \in X$. Suppose that the S_n become dense in S, as in Example 2, § 2; then the X_n form a sequence of Banach spaces approximating X. The adjoint space to X consists of the Borel measures on S (with the exception of measures giving non-zero mass to end-points which are absorbing barriers). The Borel measures with support included in S_n may be considered as elements of X_n^*, and for any such measure ν_n, $\nu_n P_n = \nu_n$. It is clear that condition (i) will be satisfied. Condition (ii) may be restated as follows. For any sequence of probability measures ν_n converging weakly to ν, $\nu_n T_n^{[th_n^{-1}]}$ converges weakly to $\nu T(t)$. Thus as a direct consequence of the definition of weak limit in § 3, (ii) will be satisfied if $T^*(t)$ is the weak limit of $T_n^{*[th_n^{-1}]}$. Finally, appealing to Lemma 3.2, we obtain the following *sufficient* condition for convergence.

> In order that a sequence of random walks converge to a diffusion process as described above, it is sufficient that the S_n become dense in S and that $T(t) = \lim_{n\to\infty} T_n^{[th_n^{-1}]}$.

Our subsequent discussion will be directed to giving conditions under which this criterion holds.

The simplest example is the convergence of the symmetric random walk to Brownian motion [6, Chap. 14]. Let S be the real line. The standard Brownian motion process is obtained by taking

$$\mu_{x,t}(A) = \int_A (2\pi t)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}t^{-1}(x-y)^2\right\}\, dy.$$

The nth symmetric random walk has the integral multiples of n^{-1} for support, and basic time-interval $h_n = n^{-2}$. For $x \in S_n$,

$$\mu_x = \frac{1}{2}\delta(x + n^{-1}) + \frac{1}{2}\delta(x - n^{-1}),$$

where $\delta(x)$ denotes the measure giving unit weight to the point x.

The convergence in this case follows immediately from a special case of the central limit theorem (namely, the normal approximation to the binomial distribution). From more general versions of the central limit theorem it may be shown that many different random walks converge to Brownian motion. The precise form of the one-step transition probability distributions is largely irrelevant; essentially all that matters is the behaviour of the mean displacements and mean square displacements. Our goal is to establish similar results for the more general diffusion processes described above.

Before giving our next set of definitions, we must fix some notational conventions. Assuming that some definite diffusion process is under discussion, we normalize the function m occurring in (6.1) so that it is continuous on the right in the interior of S and continuous (possibly with values $\pm\infty$) at the end-points. The expression $\int_a^b f\, dm$ denotes the integral over the half-open interval $(a, b]$ if $a < b$, and the negative of the integral over $(b, a]$ if $a > b$. The derivatives of functions in $\mathbf{D}(\Omega)$ may have simple discontinuities at the discontinuities of m [cf. 10], and we use $f'(x)$ to denote always the right-hand derivative of f at x.

The necessary and sufficient conditions for the central limit theorem [5] involve *truncated* means and variances of the distributions concerned, rather than the actual means and variances (which need not exist). Let C be a covering of S by intervals, and for every $x \in S$ define C_x to be the union of those elements of C which contain x. Then for a random walk with transition probabilities μ_x we introduce the following functions, defined for all x in the support of the random walk.
The *residual probability* at x is

$$P^C(x) = \mu_x(S - C_x). \tag{6.2}$$

The *truncated mass defect* at x is

$$q^C(x) = 1 - \mu_x(C_x). \tag{6.3}$$

The *truncated mean displacement* at x is

$$k^c(x) = \int_{c_x} (y - x)\, d\mu_x(y) \,. \tag{6.4}$$

The (*generalized*) *truncated mean square displacement* at x is

$$s^c(x) = \int_{c_x} v_x(y)\, d\mu_x(y) \,. \tag{6.5}$$

where

$$v_x(y) = \int_x^y \{m(u) - m(x)\}\, du$$

is the solution of $\Omega v = 1$ which satisfies the conditions $v(x) = v_x'(x) = 0$.[6] For functions in $\mathbf{D}(\Omega)$ we obtain the " Taylor expansion "

$$\begin{aligned} f(y) &= f(x) + \int_x^y f'(u)\, du \\ &= f(x) + (y - x)f'(x) + \int_x^y \int_x^u \Omega f(w)\, dm(w)\, du \\ &= f(x) + (y - x)f'(x) + v_x(y)(\Omega f(x) + E(x, y)) \end{aligned} \tag{6.6}$$

where $E(x, y)$ is an error term whose absolute value does not exceed the oscillation of Ωf on the interval (x, y). Putting this estimate together with the definitions (6.2) – (6.5), we obtain

$$\begin{aligned} Tf(x) &= \int_{c_x} f(y)\, d\mu_x(y) + \int_{S - c_x} f(y)\, d\mu_x(y) \\ &= f(x) + s^c(x)\Omega f(x) - q^c(x)f(x) \\ &\quad + k^c(x)f'(x) + s^c(x)E(x) \\ &\quad + O(p^c(x)\, \|f\|) \end{aligned} \tag{6.7}$$

where $|E(x)|$ does not exceed the oscillation of Ωf on the interval C_x.

A family of coverings will be said to *contain arbitrarily fine members* if it contains a refinement of every finite covering of S by open intervals.

Suppose a diffusion process with transition semi-group $T(t)$ and a sequence of random walks $\{S_n, h_n, T_n\}$ are given and that the space X and the approximating sequence X_n, P_n are defined as described above. Define $\Omega_n = h_n^{-1}(T_n - I)$.

LEMMA 6.1. *Suppose that there exists a family of coverings containing arbitrarily fine members such that for every covering C in the family*

[6] In the case of Brownian motion, $m(x) = 2x$, $v_x(y) = (x - y)^2$ and $s^c(x)$ is the usual truncated mean square displacement. The remark that the definitions given above furnish the appropriate generalization is due to W. Feller.

$p_n^C(x) = o(h_n)$ and $s_n^C(x) = O(h_n)$, uniformly for $x \in S_n$ as $n \to \infty$.[7] *Then for any $f \in \mathbf{D}(\Omega)$,*

$$\lim_{n\to\infty} \sup_{x\in S_n} |\, \Omega_n P_n f(x) - \Omega f(x)\,|$$
$$\leqq \lim_{n\to\infty} \sup_{x\in S_n} h_n^{-1}\,|\, - q_n^C(x) f(x) + k_n^C(x) f'(x) + (s_n^C(x) - h_n)\Omega f(x)\,|\,.$$

Proof. The functions in X are continuous on a compact set; hence Ωf in particular is uniformly continuous. For any $\varepsilon > 0$, there is a covering C in the postulated family which is sufficiently fine that the oscillation of Ωf on any C_x is less then ε. Then for any $x \in S_n$, (6.7) gives

$$(6.8)\quad \Omega_n P_n f(x) - \Omega f(x) = h_n^{-1}\{- q_n^C(x) f(x) + k_n^C(x) f'(x) + (s_n^C(x) - h_n)\,\Omega f(x) + O(\varepsilon s_n^C(x) + p_n^C(x)\,\|f\|)\}\,.$$

From the assumptions, the last term on the right is $O(\varepsilon)$, uniformly in n, and since ε is arbitrary, the conclusion follows.

In order to apply Theorem 5.3 to show that $\lim_{n\to\infty} T_n^{[th_n^{-1}]} = T(t)$ we must show that $\lim_{n\to\infty} \Omega_n P_n f = \Omega f$ for sufficiently many functions f. According to the definition, this means that we must show that

$$\lim_{n\to\infty} \sup_{x\in S_n} |\, \Omega_n P_n f(x) - \Omega f(x)\,| = 0\,.$$

We shall in all cases assume the following

Condition A. There exists a family of coverings such that the hypotheses of Lemma 6.1 are satisfied, and such that

(i) $q_n^C(x) = o(h_n)$

(ii) $k_n^C(x) = o(h_n)$

(iii) $s_n^C(x) = h_n + o(h_n)$

as $n \to \infty$, uniformly for x bounded away from the end-points of S.

We remark that since f and Ωf are bounded for any $f \in \mathbf{D}(\Omega)$ while f' is bounded except perhaps near the end-points of S, these conditions imply that

$$(6.9)\qquad -q_n^C(x) f(x) + k_n^C(x) f'(x) + (s_n^C(x) - h_n)\Omega f = o(h_n)$$

for any $f \in \mathbf{D}(\Omega)$, uniformly for x bounded away from the end-points of S.

[7] The first assumption is essentially the Lindeberg condition [**9**], strengthened by the requirement of uniformity, and implies that $\lim_{n\to\infty} \Omega_n$ is of local character.

The conditions to be imposed on q, k, and s near the boundaries are more complicated and will depend on the boundary conditions used to define Ω as a contraction of Ω'. The classification of types of boundaries and possible boundary conditions has been given in [8], and justification for the assertions made in what follows regarding the behaviour of the solutions of the homogeneous equation $(\lambda - \Omega')u = 0$ is to be found there. In [8] a different canonical form for the operator Ω' is used, so that the statements require some translation to fit our situation; [19] contains a summary of what we need in terms of the present notation. For simplicity, we shall consider only boundary conditions under which Ω is of local character at the end-points of S as well as in the interior. This restriction rules out any interaction between the boundaries of the sort described in [8] and they can be considered separately. We shall discuss in detail only the left-hand boundary. The modifications necessary to deal with the right-hand boundary will be obvious.

Since $T(t)$ is a proper semi-group, the resolvent $(\lambda - \Omega)^{-1}$ is defined on all of X. (From here on we take λ to have some fixed positive value.) Let Y be the linear subset of X consisting of all functions which are constant in some neighbourhood of each end-point. Suppose $g \in Y$ has the value c in a neighbourhood of an end-point, and let $f = (\lambda - \Omega)^{-1}g$. Then $(\lambda - \Omega)(f - c\lambda^{-1}) = g - c = 0$ on this neighbourhood. Hence in this neighbourhood, f has the form

$$f(x) = bu(x) + \text{constant} \tag{6.10}$$

where b is some constant and u is the solution (unique up to a constant multiple) of $(\lambda - \Omega')u = 0$ which satisfies the boundary condition for Ω at the end-point in question.

LEMMA 6.2. *Let $\tilde{\Omega}$ be the restriction of Ω to those functions in* $\mathbf{D}(\Omega)$ *which are of the form* (6.10) *in some neighbourhood of each boundary. Then $\tilde{\Omega}$ is densely defined, and* $\mathbf{R}(\lambda - \tilde{\Omega})$ *is dense in X.*

Proof. It is clear that $\tilde{\Omega}$ is densely defined. From the remarks preceding the lemma, $\mathbf{R}(\lambda - \tilde{\Omega}) \supset Y$ which is dense in X.

Let r denote the left-hand end-point of S. We first consider the case where r is a *natural* boundary, so that no additional boundary condition may be imposed. The functions $u(x)$, $u'(x)$, and $\Omega u(x)$ all tend to zero as x approaches r. Hence for $f \in \mathbf{D}(\tilde{\Omega})$, f' and Ωf tend to zero as $x \to r$. Then if

Condition B.1. For every covering of the family postulated in condition A

$$q_n^c(x) = o(h_n)$$

$$k_n^c(x) = O(h_n)$$

$$s_n^c(x) = O(h_n)$$

uniformly in some neighbourhood of r,

is satisfied, it is clear that for every $f \in \mathbf{D}(\tilde{\Omega})$ and every $\varepsilon > 0$, there exists a neighbourhood N of r such that

$$\lim_{n\to\infty} \sup_{x \in N \cap S_n} h_n^{-1} \mid - q_n^c(x)f(x) + k_n^c(x)f'(x) + (s_n^c(x) - h_n)\Omega f(x) \mid \leqq \varepsilon \,. \tag{6.11}$$

In the case of an *exit* boundary, we shall suppose that the absorbing barrier condition $f(r) = 0$ is imposed. Then the functions $u(x)$ and $\Omega u(x)$ vanish at r, while $u'(x)$ is continuous at r and has a finite nonzero value there. Since $f(r) = 0$, any $f \in \mathbf{D}(\tilde{\Omega})$ is a multiple of u in some neighbourhood of r. At an exit boundary, r is finite and we may integrate to obtain

$$f(x) = f(r) + (x - r)f'(r) + o(x - r) \,.$$

Making use of this and the continuity of f and Ωf, we see that if

Condition B.2. For every covering of the family postulated in Condition A

$$\begin{aligned} &(\mathrm{i}) && (x - r)q_n^c(x) = O(h_n) \\ &(\mathrm{ii}) && k_n^c(x) = O(h_n) \\ &(\mathrm{iii}) && s_n^c(x) = O(h_n) \\ &(\mathrm{iv}) && -(x - r)q_n^c(x) + k_n^c(x) = o(h_n) \end{aligned}$$

uniformly in some neighbourhood of r,

is satisfied, then for any $f \in \mathbf{D}(\tilde{\Omega})$ and $\varepsilon > 0$ there exists a neighbourhood of r on which (6.11) holds.

At an *entrance* boundary r is infinite while m is finite. The function u has the property that $u'(r) = 0$. Hence for $f \in \mathbf{D}(\tilde{\Omega})$, $f'(r) = 0$, and integrating with respect to m yields

$$f'(x) = \{m(x) - m(r)\}\Omega f(r) + o(m(x) - m(r)) \,.$$

Using this it is easy to see that

Condition B.3. For every covering of the family postulated in Condition A

$$\begin{aligned} &(\mathrm{i}) && \{m(x) - m(r)\}k_n^c(x) = C(h_n) \\ &(\mathrm{ii}) && s^c(x) = O(h_n) \end{aligned}$$

(iii) $\{m(x) - m(r)\}k_n^c(x) + s_n^c(x) - h_n = o(h_n)$

(iv) $q_n^c(x) = o(h_n)$

uniformly in a neighbourhood of r,

implies (6.11) for some neighbourhood N.

At a *regular* boundary, m and r are both finite, and the boundary condition is of the form

$$a\Omega f(r) + bf'(r) + cf(r) = 0 \tag{6.12}$$

where a, b, and c are constants, not all zero. For $f \in \mathbf{D}(\Omega)$ we have

$$\Omega f(x) = \Omega f(r) + o(1)$$

and by integrating obtain

$$f'(x) = f'(r) + \{m(x) - m(r)\}\Omega f(r) + o(m(x) - m(r))$$

and

$$f(x) = f(r) + (x - r)f'(r) + v_r(x)\Omega f(r) + o(v_r(x)) .$$

Under these circumstances, the following condition is sufficient to give the conclusion (6.11) for any $\varepsilon > 0$ and any $f \in \mathbf{D}(\Omega)$.

Condition B.4. For every covering of the family postulated in condition A

(i) $v_r(x)q_n^c(x) = O(h_n)$

(ii) $\{m(x) - m(r)\}k_n^c(x) = O(h_n)$

(iii) $s_n^c(x) = O(h_n)$

(iv) $c\{k_n^c(x) - (x - r)q_n^c(x)\} + bq_n^c(x) = o(h_n)$

(v) $a\{k_n^c(x) - (x - r)q_n^c(x)\}$
$-b\{s_n^c(x) - h_n + \{m(x) - m(r)\}k_n^c(x) - v_r(x)q_n^c(x)\} = o(h_n)$

(vi) $c\{s_n^c(x) - h_n + \{m(x) - m(r)\}k_n^c(x) - v_r(x)q_n^c(x)\}$
$+ aq_n^c(x) = o(h_n)$,

omitting (vi) if $a = b = 0$.

The last three conditions are of course not independent, and if neither a, b, nor c is zero any two of them imply the third. Condition (vi) can be omitted if $a = b = 0$ because this is the absorbing barrier condition and hence for $f \in \mathbf{D}(\Omega), f(r) = \Omega f(r) = 0$.

Putting all these results together, we obtain as the final result

THEOREM 6.1. *If a diffusion process and a sequence of random walks are such that*

(i) *The supports S_n become dense in S and*

(ii) *Condition* A *and the appropriate conditions from among* B.1, B.2, B.3, B.4, *and their analogues for the right-hand boundary are satisfied*

then the random walks converge to the diffusion process.

Proof. For any $\varepsilon > 0$ and any $f \in \mathbf{D}(\tilde{\Omega})$ the conditions imposed near the boundaries imply the existence of neighbourhoods of the boundaries on which (6.11) holds. Taken with Condition A and Lemma 6.1, this implies that for $f \in \mathbf{D}(\tilde{\Omega})$, $\lim_{n\to\infty} \Omega_n P_n f = \tilde{\Omega} f$. Hence, from Lemma 6.2, $\lim_{n\to\infty} \Omega_n$ is densely defined and $\lambda - \lim_{n\to\infty} \Omega_n$ has a dense range. Since the operators T_n are all bounded by 1, Theorem 5.3 applies to give the desired conclusion.

For our first example we shall take Brownian motion on the half-line $[0, \infty]$, for which there is a regular boundary at the origin and a natural boundary at infinity. Let $\{h_n\}$ be a sequence of positive numbers converging to zero, and define $d_n = h_n^{1/2}$. Consider a sequence of random walks, with the nth walk having a basic time interval h_n and support S_n consisting of the non-negative integral multiples of d_n. Away from the origin take the ordinary symmetric random walk with

$$\mu_x^n = \frac{1}{2}\delta(x + d_n) + \frac{1}{2}\delta(x - d_n)$$

for $x \in S_n$, $x \neq 0$. The behaviour at the origin will, of course depend on the boundary condition to be imposed on the diffusion process. As the family of coverings required in condition A we shall take the family of all finite coverings of $[0, \infty]$ by open intervals. Since $d_n \to 0$ it is easy to see that for any fixed covering, for sufficiently large n, $p_n^\sigma(x) = q_n^\sigma(x) = k_n^\sigma(x) = 0$ and $s_n^\sigma(x) = h_n$ for all $x \in S_n$, $x \neq 0$. Thus Condition A is satisfied, and so is Condition B.1 at the natural boundary. Furthermore, whatever condition of type B.4 is imposed at the regular boundary, it will automatically be satisfied in any neignbourhood of that boundary except at the origin itself. At the origin the Conditions B.4 reduce to

(i) $s_n^\sigma(0) = O(h_n)$

(ii) $ck_n^\sigma(0) + bq_n^\sigma(0) = o(h_n)$

(iii) $ak_n^\sigma(0) - b\{s_n^\sigma(0) - h_n\} = o(h_n)$

(iv) $c\{s_n^\sigma(0) - h_n\} + aq_n^\sigma(0) = o(h_n)$

where (iv) can be omitted in the absorbing barrier case.

In order to obtain convergence to the *absorbing barrier* process, it is sufficient that no particle which reaches the origin returns to the interior of the interval (i.e., once a particle is at the origin, it either stays there or disappears). In this case, $k_n^\sigma(0) = s_n^\sigma(0) = 0$, and since in

the absorbing barrier process $a = b = 0$ and condition (iv) can be omitted, the conditions are satisfied.

It is slightly more complicated to obtain a sequence of random walks converging to a diffusion process with an *elastic barrier* condition $a = 0$, $b = 1$, $c = -\alpha$ (which by taking $\alpha = 0$ specializes to the *reflecting barrier* condition). It can be done by letting a particle reaching the origin be "reflected" to the point d_n with probability $(1 + \alpha d_n)^{-1}$ and disappear with probability $\alpha d_n(1+\alpha d_n)^{-1}$. This gives $q_n^c(0) = \alpha d_n(1+\alpha d_n)^{-1}$, $k_n^c(0) = d_n(2 + \alpha d_n)^{-1}$, and $s_n^c(0) = d_n^2(1 + \alpha d_n)^{-1}$. Remembering that $d_n^2 = h_n$, it easy to verify that the conditions given above are satisfied.

As a further variation of Brownian motion, let us consider the diffusion process on $[-\infty, \infty]$ defined by taking

$$\begin{aligned} m[x] &= 2x - 1 \qquad x < 0 \\ &= 2x + 1 \qquad x \geqq 0 \,. \end{aligned}$$

This is known to give a Brownian motion process modified by the introduction of a "delay" at the origin so that the set of times for which a particle is at the origin is a nowhere-dense set of positive Lebesgue measure. We shall show that this process is the limit of symmetric random walks, modified so that a particle at the origin has probability $1 - d_n$ of staying there at the next jump, and probability $d_n/2$ of jumping to each of its neighbours. (We use h_n and d_n with the same meaning as in the previous example). As before, we have $q_n^c(x) = k_n^c(x) = 0$ for all $x \in S_n$, while $s_n^c[x] = h_n$ for all $x \neq 0$. Noting that $v_0(y) = y^2$ for $y \geqq 0$, and $v_0(y) = y^2 + 2\,|\,y\,|$ for $y < 0$, we obtain

$$\begin{aligned} s_n^c(0) &= d_n(d_n^2 + d_n) \\ &= h_n + d_n^3 = h_n + o(h_n) \,. \end{aligned}$$

The boundaries are both natural and Conditions A and B.1 can obviously be satisfied by taking the required family of coverings to consist of all finite coverings of the line.

As a final example, in which it is not so evident *a priori* what the appropriate boundary conditions are, we consider the limiting behaviour of a sequence of random walks encountered in genetic theory. These processes are discussed in [7, p. 232 ff.] and their genetic interpretation is described there. We have made some inessential changes in notation for purposes of convenience.

The processes take place on the interval $0 \leqq y \leqq 1$. (We use y as coordinate to reserve x for the natural scale used to express Ω' in the form (6.1).) The nth random walk has basic time-interval $h_n = n^{-1}$ and support the integer multiples of n^{-1}. The one-step transition probabilities are

$$\mu_y^n = \sum_{k=0}^{n} \binom{n}{k} p_{n,y}^k q_{n,y}^{n-k}(y_k) \tag{6.13}$$

so that

$$T_n f(y) = \sum_{k=0}^{n} \binom{n}{k} p_{n,y}^k q_{n,y}^{n-k}(y_k)$$

where $y_k = kn^{-1}$, $p_{n,y} = 1 - q_{n,y} = y(1 - rn^{-1}) + (1 - y)sn^{-1}$, and r and s are non-negative constants.

It is easy to check that the formal limit of the operators $\Omega_n = n(T_n - I)$ is the differential operator

$$\Omega' = \frac{1}{2} y(1 - y)D^2 + \{s - (r + s)y\}D \tag{6.14}$$

where D denotes differentiation with respect to y.[8] The natural scale x, and the monotone function m needed to express Ω' in the form (6.1) may be taken as [10, formula 4.2]

$$x(y) = \int_{1/2}^{y} z^{-2s}(1 - z)^{-2r}\, dz \tag{6.15}$$

$$m(y) = 2\int_{1/2}^{y} z^{2s-1}(1 - z)^{2r-1}\, dz\ .$$

The nature of the boundary at $y = 0$ depends on the value of s; the nature of the boundary at $y = 1$ depends similarly on the value of r. We shall discuss only the left-hand boundary in detail, since there is obviously a complete symmetry. Checking with the criteria given in [8] or [19], we see that the boundary at $y = 0$ is an exit boundary if $s = 0$, a regular boundary if $0 < 2s < 1$, and an entrance boundary if $2s \geq 1$.

It would presumably be possible to show that the sequence of random walks under consideration satisfies our Condition A and the appropriate conditions at the boundary, but the functions x and m are not elementary, and it would be complicated to obtain satisfactory estimates for the mean displacements, etc. We shall instead make direct use of Lemma 6.2 and Theorem 5.3.

We assert that *if f has a continuous second derivative on the closed interval* $[0, 1]$ *then* $\lim_{n\to\infty} \Omega_n f = \Omega' f$. To prove this, observe first that for the constant function 1, for $g(y) = y$, and for $h(y) = y^2$, simple calculations from the elementary formulas for the mean and variance of a binomial distribution give

[8] Goldberg [12] discusses the solutions of the equation $\partial f/\partial t = \Omega' f$ under various boundary conditions.

$$\Omega_n \mathbf{1} = 0 = \Omega \mathbf{1}$$

$$\Omega_n g(y) = s - (r + s)y = \Omega g(y)$$

$$\Omega_n h(y) = (1 + 2s)y - (1 + 2r + 2s)y^2 + O(n^{-1})$$
$$= \Omega h(y) + O(n^{-1}) .$$

It is also easy to verify (for instance, by estimates obtained from the normal approximation to the binomial distribution) that for any $\varepsilon > 0$,

$$p_n^\varepsilon(y) = \int_{|z-y|>\varepsilon} d\mu_y^n(z) = o(n^{-1})$$

uniformly in y. Then using a second-order Taylor expansion, any f with a continuous second derivative can be approximated over the interval $(y - \varepsilon, y + \varepsilon)$ by a linear combination of $\mathbf{1}$, g, and h so that the error at z is less than $(y - z)^2 E_\varepsilon$ where E_ε is the maximum oscillation of $D^2 f$ on any interval of length 2ε. Estimating f by such a linear combination, we obtain

$$\Omega_n f(y) = \Omega f(y) + O(n^{-1} + E_\varepsilon) + n p_n^\varepsilon(y) .$$

Since E_ε can be made arbitrarily small by suitable choice of ε and for each such choice $n p_n^\varepsilon(y) = o(1)$, our assertion holds as stated.

We shall complete the proof that the random walks (6.13) converge to a diffusion process associated with the operator (6.14) by showing that (if a suitable boundary condition is imposed in the regular boundary case) all functions in $\mathbf{D}(\tilde{\Omega})$, where $\tilde{\Omega}$ is defined as in Lemma 6.2, have continuous second derivatives on the closed interval. Then we shall have $\lim_{n\to\infty} \Omega_n$ densely defined and $\mathbf{R}(\lambda - \lim_{n\to\infty} \Omega_n)$ dense, so that Theorem 5.3 will apply to give the desired conclusion. Of course, every $f \in \mathbf{D}(\Omega')$ has a continuous second derivative on the open interval, but since the coefficient of D^2 in Ω' vanishes at the end-points, the continuity of $\Omega' f$ on $[0, 1]$ does not imply the continuity of $D^2 f$ there.

The homogeneous equation $(\lambda - \Omega')u = 0$ may be put into the form of the standard hypergeometric equation

$y(1 - y)u''(y) + \{c - (a + b + 1)y\}'(y) - abu = 0$ by determing a, b, and c from the equations

$$c = 2s$$
$$a + b + 1 = 2r + 2s$$
$$ab = 2\lambda .$$

The solutions of this equation are given in terms of hypergeometric functions in section 2.3.1 of [2], whose notation we adopt. In the case

of a non-regular boundary there is a unique solution which is bounded at 0. This solution is

$$u = yF(a - c + 1, b - c + 1; 2; y)$$

for the exit boundary case, $c = 2s = 0$, and

$$u = F(a, b; c; y)$$

for the entrance boundary case, $c = 2s \geq 1$. Since the hypergeometric function is analytic at the origin, it has a continuous second derivative there.

In the case of a regular boundary, $0 < c = 2s < 1$, there are two independent solutions which are bounded at 0, namely

$$u_1 = F(a, b; c; y)$$

and

$$u_2 = y^{1-c}F(a - c + 1,\ b - c + 1;\ 2 - c; y)\ .$$

We now impose the reflecting barrier condition, $df/dx \to 0$ as $x \to 0$. From (6.15)

$$\frac{df}{dx} = y^{2s}(1 - y)^{2r}D$$

and it is easily verified that $du_1/dx \to 0$ at the boundary, but that du_2/dx does not. Thus solutions of the homogeneous equation which satisfy the boundary condition imposed are multiples of u_1, and have a continuous second derivative at 0.

References

1. S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932.
2. Bateman Manuscript Project (ed. A. Erdélyi), *Higher Transcendental Functions*, vol. 1, New York, 1953.
3. J.L. Doob, *Stohastic processes*, John Wiley & Sons, New York, 1953.
4. J. Douglas, Jr., *On the relation between stability and convergence in the numerical solution of linear parabolic and hyperbolic differential equations*, J. Soc. Indust. Appl. Math. **4** (1956), 20–37.
5. W. Feller, *Uber den zentralen Grenzwertsatz der Wahrscheinlickeitsrechnung*, Math. Z. **40** (1935), 521–559.
6. ———, *An introduction to probability theory and its applications*, vol. 1, John Wiley & Sons, New York, 1950.
7. ———, *Diffusion processes in genetics*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1951, 227–246.
8. ———, *The parabolic differential equations and associated semi-groups of transformations*, Ann. of Math., **55** (1952), 468–519.
9. ———, *Diffusion processes in one dimension*, Trans. Amer. Math. Soc., **77** (1954) 1–31.

10. ———, *On differential operators and boundary conditions*, Comm. Pure and Appl. Math., **8** (1955), 203–216.
11. ———, *The general diffusion operator and positivity preserving semi-groups in one dimension*, Ann. of Math. **60** (1954), 417–436.
12. S. Goldberg, *On a singular diffusion equation*, Thesis, Cornell, 1950.
13. D.R. Hartree, *Numerical analysis*, Oxford University Press, 1952.
14. E. Hille, *Functional analysis and semi-groups*, Amer. Math. Soc. Colloquium Publications, vol. 31, New York, 1948.
15. ———, *The abstract Cauchy problem and Cauchy's problem for parabolic differential equations*, Journal d'Analyse Mathématique, **3** (1953–1954), 81–196.
16. L.V. Kantorovitch, *Functional and applied mathematics*, Uspehi Matem. Nauk **3** (1948); Bureau of Standards report 1509, 1952.
17. P.D. Lax and R. D. Richtmyer, *Survey of the stability of linear finite difference equations*, Comm. Pure Appl. Math. **9** (1956), 267–293.
18. W. Leutert, *On the convergence of approximate solutions of the heat equation to the exact solution*, Proc. Amer. Math. Soc. **2** (1951), 433–439.
19. H.P. McKean, Jr., *Elementary solutions for certain parabolic partial differential equations*, Trans. Amer. Math. Soc. **82** (1956), 519–548.
20. G. G. O'Brien, M. A. Hyman, and S. Kaplan, *A study of the numerical solution of partial differential equations*, J. Math. Phys. **29** (1951), 223–251.
21. D.V. Widder, *The Laplace transform*, Princeton University Press, 1941.
22. K. Yosida, *On the differentiability and the representation of one-parameter semi-group of operators*, J. Math. Soc. Japan. **1** (1949).

QUEEN'S UNIVERSITY KINGSTON, CANADA

Mathematical papers intended for publication in the *Pacific Journal of Mathematics* should be typewritten (double spaced), and the author should keep a complete copy. Manuscripts may be sent to any one of the four editors. All other communications to the editors should be addressed to the managing editor, E. G. Straus at the University of California, Los Angeles 24, California.

50 reprints per author of each article are furnished free of charge; additional copies may be obtained at cost in multiples of 50.

The *Pacific Journal of Mathematics* is published quarterly, in March, June, September, and December. The price per volume (4 numbers) is $12.00; single issues, $3.50. Back numbers are available. Special price to individual faculty members of supporting institutions and to individual members of the American Mathematical Society: $4.00 per volume; single issues, $1.25.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 2120 Oxford Street, Berkeley 4, California.

Printed at Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.), No. 6, 2-chome, Fujimi-cho, Chiyoda-ku, Tokyo, Japan.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION

The Supporting Institutions listed above contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its content or policies.

Pacific Journal of Mathematics
Vol. 8, No. 4 June, 1958