SOME NEW CONNECTIONS BETWEEN PROBABILITY AND CLASSICAL ANALYSIS

By WILLIAM FELLER†

Recent research has revealed the intimate relationship between potential theory and Markov processes, and has supplied new examples of the fertility of a probabilistic approach to problems of classical analysis. Choquet's work on capacities, the Beurling–Deny theory of general potentials, Doob's probabilistic approach to the Dirichlet problem, and Hunt's basic results concerning potentials and Markov processes are closely related, despite the diversity of formal appearances and methods. I had hoped in this address to discuss the interconnections between these theories, but the task proved too overwhelming for the limited time and my own limitations. I am therefore compelled unashamedly to restrict this talk to some related aspects of my own work. I propose to describe the two boundaries and topologies induced by the annihilators of certain operators; to discuss their justification, their use for an abstract theory of so-called boundary value problems, and their connection with an invariant theory of operators of local character (which generalize the ill-defined notion of differential operators).

I shall not endeavor to develop a theory or even to state results in a precise form. Rather I shall try to explain the background and the purpose of the theory by means of a few simple examples, preferably using classical harmonic functions. Although everything will be interpreted probabilistically, the main emphasis is purely analytic.

In the sequel D will always denote a topological space and $\mathbf{C}$ the familiar Banach space of continuous functions in D with

$$\|f\| = \sup_{p \in D} |f(p)|.$$

Unless otherwise stated all operators will act on continuous functions.

1. Boundaries induced by positive operators

1.1. Harmonic functions. It is convenient to start with a probabilistic interpretation of harmonic functions by means of an *ad hoc* constructed random walk. In § 2 we shall approach the Laplacian more directly and more naturally.

† Research in part supported by the Office of Ordnance Research, U.S. Army.

Let D be the open unit disc in the plane and for each point $p \in D$ let $D_p \subset D$ be the greatest open disc centered at p and contained in D. We shall study the operator T from $\mathbf{C}$ to $\mathbf{C}$ for which $Tf(p)$ *equals the arithmetic mean* of f over D_p. Thus

$$Tf = \int_D K(\,.\,, q)\, f(q)\, dq, \tag{1}$$

where $K(p, q)$ equals $|D_p|^{-1}$ or 0 according as $q \in D_p$ or $q \bar{\in} D_p$. Similarly, T^n is induced by a kernel $K^{(n)}$.

This T determines a discrete *random walk* with arbitrary initial position $p \in D$ in which $K^{(n)}(p, .)$ is the density of the random position $Q_n \in D$ after n steps. We obtain a well-defined measure in the space of all sequences $\{Q_n\}$ ($Q_0 = p$, $Q_n \in D$) and it can be shown that almost all sequences are convergent to a point of the boundary circle. For our purposes a less refined purely analytical statement will suffice. For any set $A \subset D$ the probability that $Q_n \in A$ equals $T^n\chi(p)$, where χ is the characteristic function of A. It is easily verified that $T^n f \to 0$ for each $f \in \mathbf{C}$ vanishing at the boundary. Therefore $T^n\chi \to 0$ for each compact A, and this is equivalent to the statement that Q_n approaches the boundary in probability.

This result can be rendered more precise as follows. Let $\Gamma \subset B$ be a set of the boundary circle B, and u_Γ the harmonic function determined, in the classical sense, by the boundary values 1 on Γ and 0 on $B - \Gamma$. Then $u_\Gamma(p)$ *is the probability that* Q_n *approaches* Γ as $n \to \infty$.

Harmonic functions appear in this context because they are eigenfunctions satisfying $\phi = T\phi$. The set $\mathfrak{P}$ of all solutions of this equation such that $0 \leqslant \phi \leqslant 1$ is a convex set and the harmonic functions u_Γ coincide with its *extremals*. The sets

$$\Gamma_\epsilon = \{q \in D \mid u_\Gamma(q) > 1 - \epsilon\} \tag{2}$$

are a system of deleted neighborhoods of Γ. If χ is the characteristic function of Γ_ϵ then $T^n\chi \to u_\Gamma$, and hence

$$u_\Gamma(p) = \lim_{n \to \infty} \operatorname{prob}\{Q_n \in \Gamma_\epsilon\}. \tag{3}$$

From this it follows easily that $u_\Gamma(p)$ is the probability of an actual asymptotic approach to Γ, but for our purposes the weaker statement (3) fully suffices. [An alternative proof is given in the next section.]

The point to be emphasized is that this set-up is not of an analytic nature but can be carried out abstractly for a large class of operators in an arbitrary topological space D. No boundary need be defined *a priori*, and it is natural to *define* a boundary in such a way that the sets Γ_ϵ become neighborhoods of the corresponding boundary sets. Again, D may possess a boundary which does not admit of the interpretation of Γ_ϵ as neighborhoods, and it may be necessary to introduce a new boundary appropriate to the study of the transformation T. The simplest example is obtained by mapping the unit disc D conformally onto a domain $\tilde{D}$ whose boundary $\tilde{B}$ is of a complicated structure with prime ends, etc. Under this conformal map T and its random walk are carried over to $\tilde{D}$, the new eigenfunctions are again harmonic, but obviously the convergence properties of the random walk remain true only if we replace the 'natural' boundary of $\tilde{D}$ by the boundary and topology induced by the conformal map. This, of course, is a probabilistic version of the now familiar observation due to Martin[17] that the study of harmonic functions in complicated domains requires the introduction of an appropriate boundary.

We pass to a more interesting example of a different kind.

1.2. Relativization and isomorphisms. For an arbitrary (not necessarily bounded) $\psi > 0$ harmonic in the disc D we define a new transformation by

$$T_\psi f = \psi^{-1} T(f\psi). \tag{4}$$

Its kernel is given by

$$K_\psi(p,q) = \psi^{-1}(p)\, K(p,q)\, \psi(q). \tag{5}$$

A function $v > 0$ satisfies $v = T_\psi v$ if and only if $\phi = v\psi$ satisfies $\phi = T\phi$ (is harmonic). We have thus a 1-1 correspondence between the positive eigenfunctions of T and T_ψ with 1 and ψ corresponding to ψ^{-1} and 1, respectively.

Operators of this form will be called *similar* to T. Clearly similarity is transitive, symmetric and reflexive. A closely related transformation has been used by Brelot[1] for harmonic functions. We shall see that the notion of similarity is exceedingly useful, and has its counterpart in similar semigroups and differential equations. Here we use it to illustrate the notion of the boundary induced by T_ψ and to derive a new proof for our interpretation of the harmonic function u_Γ.

Denote by $\mathfrak{P}$ and $\mathfrak{P}_\psi$, respectively, the sets of positive eigenfunctions of T and T_ψ bounded by 1. For simplicity of exposition consider the case $\psi = u_\Gamma$ where ψ is an extremal of $\mathfrak{P}$. The mapping $u\psi \leftrightarrow \phi$ establishes a 1-1 correspondence between $\mathfrak{P}_\psi$ and the subset of $\mathfrak{P}$ of elements such that $\phi \leqslant u_\Gamma$; this correspondence preserves extremals. Now the structure of the set $\mathfrak{P}$ of harmonic functions is best described in terms of the 'natural' boundary of D. Since relations between T, $\mathfrak{P}$, and $D \cup B$ with the natural topology will carry over to T_ψ, $\mathfrak{P}_\psi$ and $D \cup \Gamma$, the set Γ plays for T_ψ the role that B plays for T and may be considered the *boundary induced by* T_ψ.

The same conclusion may be reached probabilistically. From $\psi^{-1} = T_\psi \psi^{-1}$ it follows that $T_\psi^n f \to 0$ for each $f \in \mathbf{C}$ vanishing along Γ. Hence if χ is the characteristic function of a neighborhood U of Γ we have $T_\psi^n \chi \to 1$: thus in the random walk associated with T_ψ the paths converge in probability to the boundary set Γ.

This remark leads to a new proof of the interpretation of u_Γ given in the preceding section. Using (5) it is seen that the relation $T_\psi^n \chi \to 1$ may be rewritten as

$$\int_U K^{(n)}(p,q)\,\psi(q)\,dq \to \psi(p). \tag{6}$$

Now the neighborhood U of Γ may be chosen so small that in it $\psi = u_\Gamma$ is arbitrarily close to 1 and we conclude that in the T-random walk $\mathbf{P}\{Q_n \in U\} \to u_\Gamma(p)$ for each neighborhood of Γ. This is a slightly weakened version of the interpretation of $u_\Gamma(p)$ as *probability of an asymptotic approach to* Γ.

Given this interpretation of u_Γ we see that in the T-random walk K_ψ represents the *conditional* transition probability density given the event that the paths converge to Γ. Probabilistically, then, *the T_ψ-random walk is obtained from the T-random walk by conditioning or relativization:* in the T-walk we pay attention only to paths converging to Γ. More precisely, let $\mathfrak{S}$ be the set of all sequences $\{Q_n\}$, $(Q_0 = p, Q_n \in D)$, with the measure $\mathbf{P}$ induced by T, and let $\mathfrak{S}_\Gamma$ be the subset of sequences converging to Γ. Then the T_ψ-walk assigns zero probability to $\mathfrak{S} - \mathfrak{S}_\Gamma$ and probability $\mathbf{P}\{\mathfrak{A}\} \div \mathbf{P}\{\mathfrak{S}\} = \mathbf{P}\{\mathfrak{A}\}\,\psi^{-1}(p)$ to the subsets $\mathfrak{A} \subset \mathfrak{S}$.

1.3. Abstract construction; restricted and total boundaries. We pass to the extreme case where D is the set of integers $1, 2, \ldots$, with the discrete topology. This has the advantage that no preconceived intuitive notion of boundary obscures the view. The boundaries may be of a complicated topological structure and the present case will clearly reveal the features and problems of the most general set-up.

$\mathbf{C}$ is now the space of bounded functions and we write $\mathbf{f} \in \mathbf{C}$ as a column matrix with elements $f(i)$. We consider an operator T defined by a matrix Π with elements $\Pi(i,j)$ so that in matrix notation $T\mathbf{f} = \Pi\mathbf{f}$. The matrix Π is supposed to be substochastic, that is, its elements are $\geqslant 0$, its row sums $\leqslant 1$. We denote by $\mathfrak{P}$ the set of all eigenvectors $\boldsymbol{\phi}$ such that $\boldsymbol{\phi} = \Pi\boldsymbol{\phi}$

and $\mathbf{0} \leqslant \boldsymbol{\phi} \leqslant \mathbf{1}$, and by $\mathfrak{P}^\infty$ the set of all (possibly unbounded) solutions $\boldsymbol{\phi} > \mathbf{0}$ of $\boldsymbol{\phi} = \Pi\boldsymbol{\phi}$. To avoid trivialities we shall assume: (i) each $\boldsymbol{\phi}$ is strictly positive, (ii) $\mathfrak{P}$ contains at least two independent vectors. [Condition (i) eliminates the nuisance of partitioned matrices requiring words rather than thoughts; (ii) eliminates empty and single-point boundaries.]

Again Π may be interpreted as the matrix of transition probabilities in a *random walk* (Markov chain), the row defects $1 - \Sigma\Pi(\,.\,, j)$ accounting for the possibility of a termination of the process. For an arbitrary initial position $i \in D$ we have a probability measure on the set $\mathfrak{S}$ of all terminating or infinite sequences of integers $\{Q_n\}$. The subset $\mathfrak{S}^{(n)}$ of sequences of length $\geqslant n$ has probability $\Sigma_j \Pi^n(i,j)$, and hence the probability of the set $\mathfrak{S}^{(\infty)}$ of infinite sequences is given by the ith element of $\bar{\boldsymbol{\phi}} = \lim \Pi^n \mathbf{1}$. Note that $\bar{\boldsymbol{\phi}}$ is the *maximal* element of $\mathfrak{P}$.

We proceed to introduce a *restricted* or $\mathfrak{P}$*-boundary* B, and a *total* or $\mathfrak{P}^\infty$*-boundary* $B^\infty \supset B$. We begin with the extremely simple special case where $\boldsymbol{\phi} = \Pi\boldsymbol{\phi}$ has only *finitely* many independent solutions.

(*a*) *The* $\mathfrak{P}$*-boundary*. Let $\mathfrak{P}$ be spanned by N non-zero vectors $\boldsymbol{\phi}^{(1)}, \ldots, \boldsymbol{\phi}^{(N)}$. These can be chosen as extremal elements of $\mathfrak{P}$, which amounts to saying that $\|\boldsymbol{\phi}^{(k)}\| = 1$ and $\bar{\boldsymbol{\phi}} = \boldsymbol{\phi}^{(1)} + \ldots + \boldsymbol{\phi}^{(N)}$. For fixed k and $\epsilon > 0$ put $\Gamma_\epsilon^{(k)} = \{i \mid \phi^{(k)}(i) > 1-\epsilon\}$. As $\epsilon \to 0$ we get a nest of non-empty sets with empty intersection; from $\bar{\boldsymbol{\phi}} \leqslant \mathbf{1}$ we conclude that for fixed $\epsilon > \frac{1}{2}$ the sets $\Gamma_\epsilon^{(j)}$ and $\Gamma_\epsilon^{(k)}$ are non-overlapping ($j \neq k$).

The *restricted boundary* B consists of N points $\beta^{(1)}, \ldots, \beta^{(N)}$ such that $\Gamma_\epsilon^{(k)}$ is a deleted neighborhood of $\beta^{(k)}$. We can extend the definition of each $\boldsymbol{\phi} \in \mathfrak{P}$ to $D \cup B$ by putting $\phi^{(j)}(\beta^{(k)}) = 1$ or 0 according as $j = k$ or $j \neq k$. Then each $\boldsymbol{\phi}$ is continuous in $D \cup B$, and the 'Dirichlet problem' is soluble: to prescribed boundary values there corresponds exactly one $\boldsymbol{\phi} \in \mathfrak{P}$.

Finally with an obvious notation, $\Pi^n(i, \Gamma_\epsilon^{(k)}) \to \phi^{(k)}(i)$, as $n \to \infty$ for each fixed ϵ. From this one deduces that in the random walk starting at i the sample sequences $\{Q_n\}$ converge with probability $\phi^{(k)}(i)$ to $\beta^{(k)}$; with probability $1 - \bar{\phi}(i)$ they terminate; and the probability of no convergence is 0.

(*b*) *The* $\mathfrak{P}^\infty$*-boundary* can be introduced directly, but it is more convenient to use the similarity transformation introduced in § 1.2. A matrix Π_ψ is *similar* to Π if either $\Pi_\psi = \Pi$ or

$$\Pi_\psi(i,j) = \psi^{-1}(i)\,\Pi(i,j)\,\psi(j), \tag{7}$$

where $\boldsymbol{\psi} \in \mathfrak{P}^\infty$. We recall that similarity is transitive, reflexive, and symmetric and that the mapping $\boldsymbol{\phi} \leftrightarrow \mathbf{v}$ where $v(i) = \phi(i)\,\psi^{-1}(i)$ establishes a 1-1 correspondence between $\mathfrak{P}^\infty$ and $\mathfrak{P}^\infty_\psi$.

To see the relation between B and the corresponding restricted boundary B_ψ induced by Π_ψ consider the typical case $\boldsymbol{\psi} = \boldsymbol{\phi}^{(1)} + \boldsymbol{\phi}^{(2)}$. To $\boldsymbol{\phi}^{(3)}, \ldots, \boldsymbol{\phi}^{(N)}$ and to each unbounded $\boldsymbol{\phi} \in \mathfrak{P}^\infty$ there correspond unbounded vectors in $\mathfrak{P}^\infty_\psi$ and the boundary B_ψ reduces to two points whose neighborhoods coincide with the neighborhoods of $\beta^{(1)}$ and $\beta^{(2)}$. For the interpretation of this Π_ψ in terms of *conditional probabilities* see § 1.2.

In general, if B_ψ is the restricted boundary induced by Π_ψ we shall identify points of B and B_ψ with coinciding systems of neighborhoods. With this identification we define the *total boundary* B^∞ *induced by* Π *as the union of the boundaries* B_ψ as $\boldsymbol{\psi}$ ranges over $\mathfrak{P}^\infty$. All *similar matrices induce the same total boundary*, and if $\mathfrak{P}^\infty$ is spanned by M independent vectors, then B^∞ contains exactly M points.

Note that $D \cup B^\infty$ need not be compact. (No compactification seems natural or desirable for our purposes.)

(c) *The maximal ideal boundaries*. When $\mathfrak{P}$ is not spanned by denumerably many elements the extremal elements of $\mathfrak{P}$ do not correspond to points of the prospective boundary, but rather to sets of positive capacity. No satisfactory definition of points and neighborhoods is known. Now both $\mathfrak{P}$ and $\mathfrak{P}^\infty$ have a lattice structure similar to that of harmonic functions, and the correspondence between $\mathfrak{P}^\infty$ and $\mathfrak{P}^\infty_\psi$ is a lattice isomorphism. This makes it possible to define points of B and B^∞ by maximal ideals in $\mathfrak{P}$ and $\mathfrak{P}^\infty$, respectively (see[5]). Unfortunately these boundaries are absurdly large. For example, sample sequences converge to sets rather than to points; each $\boldsymbol{\phi} \in \mathfrak{P}$ has continuous boundary values which is at variance with the desirable model of harmonic function in a disc D with the natural topology. That the lattices $\mathfrak{P}$ and $\mathfrak{P}^\infty$ are isomorphic to lattices of continuous functions on some Hausdorff spaces is, of course, well known (see, for example, Kadison[15]). To us the main point is that this Hausdorff space appears as a *boundary* of D and is useful as such.

Our maximal ideal boundaries serve well for an orientation and as a guide, but the introduction of a less clumsy and more appropriate boundary is an open, and promising, problem.

[*Postscript*. A partial solution has now been obtained by J. L. Doob[19] who uses Martin's original construction. The Martin boundary is sufficiently small for sample sequences to converge toward points. In the finite case, however, this boundary may be bigger than ours; its neighborhoods are larger and this could lead to difficulties in connection with boundary value problems and non-minimal semigroups.]

2. Semigroups and differential equations

2.1. Orientation. We turn to the more interesting study of a family of positive operators $\{T(t)\}$, $t \geqslant 0$, from $\mathbf{C}$ to $\mathbf{C}$ with $\|T(t)\| \leqslant 1$, and with the semigroup property $T(t+s) = T(t)\,T(s)$. The probabilistic counterpart to a fixed $T(t)$ is a (possibly terminating) random walk with jumps taking place at times $t, 2t, \ldots$. To the whole semigroup there corresponds a random motion (Markov process) in D with continuous time: the sample space $\mathfrak{S}$ is the space of functions Q defined in an interval $0 \leqslant t < \tau \leqslant \infty$ such that $Q(t) \in D$, and $Q(0) = p \in D$ is a given point. The semigroup induces a $\mathbf{P}$-measure in $\mathfrak{S}$ such that for each Borel set $A \subset D$ we have $\mathbf{P}\{Q(t) \in A\} = T(t)\,\chi(p)$ where χ is the characteristic function of A. Needless to say with the $\mathbf{P}$-measure almost all paths are reasonably regular. In particular, they are continuous when the semigroup is generated by a differential operator ('diffusion processes').

If a path Q is defined only within a finite interval $0 \leqslant t < \tau$ (that is, if the process terminates at time τ) then, except on a null set of paths, as $t \uparrow \tau$ either $Q(t) \to q \in D$ or $Q(t)$ has no point of accumulation in D. In the first case we say that the process terminates at q, in the second that it *terminates 'at the boundary'*. However, it remains to justify this expression.

For this purpose we shall have to introduce a boundary B induced by the *generator* Ω of the semigroup. It will be seen that all operators $T(t)$, $t > 0$, *induce the same boundary* $\pi \subset B$, and that the process *either terminates at* $B - \pi$ *or approaches* π *asymptotically* without reaching it.

We say that the semigroup is *generated by* Ω if for a dense set of 'smooth' $f \in \mathbf{C}$

$$\lim_{t \to 0} t^{-1}\{T(t)\,f - f\} = \Omega f \in \mathbf{C} \tag{8}$$

in the sense of pointwise convergence. According to this definition, introduced in[11], Ω may generate many semigroups; the infinitesimal generator in the sense of Hille[13] is a contraction $\Omega|\Sigma$ of Ω obtained by imposing *lateral conditions* (see § 4.3).

Putting $T(t)\,f = u(t, .)$ the function u will satisfy the functional equation $u_t = \Omega u$ with the 'initial condition' $u(0, .) = f$. (This is literally true for smooth f, and in an operational sense for all f.) In classical terms we are concerned with 'solving' this equation. We consider first a family of 'similar' generators and then the relation between generators with the same annihilators.

2.2. The Laplacian. Isomorphisms. We return to harmonic functions and take as example the familiar *heat equation* $u_t = \Delta u$ in the unit disc D. Among the positive semigroups generated by Δ there exists a *minimal semigroup*. In classical terms $u(t, .) = T(t)f$ is the solution of $u_t = \Delta u$ with initial condition $u(0, .) = f$ and *boundary condition* that u vanishes at the circle B. Associated with it is the *Wiener process* in D terminating at B. Physically the process represents a homogeneous diffusion in D with 'absorbing boundaries', or heat conduction with zero temperature at the boundary.

For this minimal semigroup $\|T(t)\| < 1$ and therefore the boundary π induced by $T(t)$ is empty, but we are concerned with the boundary induced by harmonic functions, that is, the *annihilators* of the generator Δ. In this connection, of course, we adhere to the natural topology of the closed disc. As in § 1.1 let u_Γ be the harmonic function in D determined by the boundary values 1 on the set $\Gamma \subset B$ and 0 on $B-\Gamma$. For the process starting at the point $p \in D$ we have the following analogue to the results of § 1.1.

$u_\Gamma(p)$ is the probability that the process terminates at the boundary set Γ. The probability that this happens before time t is $v(t, p)$ where v is the solution of $v_t = \Delta v$ with zero initial values and boundary values 1 on Γ and 0 on $B-\Gamma$.

This assertion becomes plausible on observing that for bounded harmonic ψ obviously

$$T(t)\,\psi = \psi - v(t, .), \tag{9}$$

where $v_t = \Delta v$ and v has zero initial values and boundary values coinciding (almost everywhere) with those of ψ. Now $T(t)\,1(p) = \mathbf{P}\{Q(t) \in D\}$ is the probability that the process does not terminate before it. Together with (9) this implies the assertion for the particular case $u_\Gamma = 1$, or $\Gamma = B$, at least if we accept as fact that the process does not terminate in the interior of D. We shall only indicate how the general assertion may be reduced to this particular case by a generalization of the method of *similarity transformations* or isomorphisms, introduced in § 1.2.

For positive harmonic ψ we define a new semigroup of positive operators from $\mathbf{C}$ to $\mathbf{C}$ by

$$T_\psi(t)\,f = \psi^{-1}\,T(t)\,(f\psi). \tag{10}$$

If $0 \leqslant f \leqslant 1$ then $0 \leqslant T_\psi(t)f \leqslant \psi^{-1} \leqslant T(t)\,\psi \leqslant 1$ and hence $\|T_\psi(t)\| \leqslant 1$. A glance at (8) shows that the semigroup $\{T_\psi\}$ is generated by the operator Δ_ψ defined by

$$\Delta_\psi f = \psi^{-1}\Delta(f\psi). \tag{11}$$

For simplicity of exposition we now restrict the consideration to $\psi = u_\Gamma$. First note that ψ^{-1} is unbounded near $B-\Gamma$ and close to 1 near Γ, and that $T_\psi(t)^{-1} \leqslant \psi^{-1}$. From this it is easy to conclude that *the random* *process corresponding to* $\{T_\psi\}$ *terminates at* Γ. Our interpretation of u_Γ is obtained from this by adapting the argument following (6). As at the end of § 1.2 we remark that the kernel of the T_ψ-semigroup represents the conditional transition probability densities of the T-process given that the path terminates at Γ. In the sense explained above the T_ψ-*process is therefore simply the restriction of the* T-*process to the paths terminating at* Γ.

Returning to the analytical relationship between the T- and the T_ψ-semigroup note that $\Delta_\psi \phi = 0$ if and only if $\phi\psi$ is harmonic. The positive (possible unbounded) annihilators of Δ and Δ_ψ stand in a 1-1 correspondence (which is a lattice isomorphism). The bounded annihilators of Δ_ψ correspond to the harmonic functions dominated by $\psi = u_\Gamma$, and thus the set Γ is the appropriate boundary for $\{T_\psi\}$ just as the circle B is for $\{T\}$. In short we find the same situation as in § 1.3.

The circle B, *induced by the harmonic functions, represents the 'total' boundary for the family of all similar semigroups* $\{T_\psi(t)\}$ [*or all differential equations* $u_t = \Delta_\psi u$]. *The boundary induced by the bounded annihilators of* Δ_ψ *corresponds to the subset* $\Gamma \subset B$.

That Γ is the appropriate boundary for the T_ψ-semigroup reflects the fact that the common *range of the transformations* $T_\psi(t)$ *is characterized by the side condition that* $T_\psi(t)f$ *vanishes along* Γ just as $T(t)f$ vanishes along B.

We have here the simplest example of a 'boundary condition' and see that it relates to our boundary rather than the 'natural' one. In the terminology of classical differential equations, Δ_ψ is a differential operator with coefficients singular along $B - \Gamma$, and boundary conditions can be imposed only along Γ. How vague and unsatisfactory such descriptions can be is known from the simple Sturm–Liouville theory in one dimension.

2.3. The active boundary. Singular operators. A new phenomenon may be described in connection with the operator $\Omega = \omega\Delta$ in the disc D where $\omega > 0$ is continuous in D but may tend to zero or infinity near the circle B. This operator has the same annihilators as the Laplacian Δ and it is interesting to compare the semigroups generated by $\omega\Delta$ with those generated by Δ. In classical terms we are concerned with the integration of the parabolic differential equation $u_t = \omega\Delta u$ which may be singular owing to the behavior of ω near B. We describe here the main features of the theory carried out in[6] for denumerable spaces (the Kolmogoroff differential equations) by methods of much wider applicability.

There exists a uniquely determined *minimal* positive semigroup $\{T(t)\}$ generated by $\omega\Delta$ and for it $\|T(t)\| \leqslant 1$. However, for an appropriate choice of ω it may happen that $T(t)\,1 = 1$: the minimal semigroup is in this case *unique*; by contrast to the heat equation no boundary conditions need or can be imposed in this case, and the induced random process (diffusion) does not terminate at a finite time.

With an arbitrary $\omega > 0$ let $\{T(t)\}$ be the minimal semigroup generated by $\omega\Delta$ in the unit disc D. Then the following is true.

(*a*) *The passive boundary* π. There exists a set $\pi \subset B$ of the unit circle (determined up to a null set) such that for each $t > 0$ the eigenfunctions ϕ of $\phi = T(t)\,\phi$ such that $0 \leqslant \phi \leqslant 1$ coincide with the positive harmonic functions (the annihilators of $\omega\Delta$) dominated by u_π, the harmonic function with boundary values 1 on π and 0 on $B-\pi$. In this sense *the boundary induced by each* $T(t)$ *coincides with* π. In the case $\omega = 1$ (the heat equation) π is empty. For any set $\Gamma \subset \pi$ the value $u_\Gamma(p)$ is the probability in the $T(t)$-process starting at the point $p \in D$ that Γ is *asymptotically approached as* $t \to \infty$; the probability of reaching Γ at a finite time is zero.

(*b*) *The active boundary* $A = B-\pi$. For each set $\Gamma \subset A$ the value $u_\Gamma(p)$ *is the probability that the process will terminate at* Γ. The probability that this occurs before time t equals $v(t,p)$ where v is a solution of $v_t = \omega\Delta v$ with zero initial values and boundary values 1 on Γ and 0 on $A-\Gamma$.

No boundary conditions can be imposed along the passive boundary. This summary explains the relations between minimal *semigroups generated by operators with the same annihilators.*

In analytic terms we may characterize the active and passive boundaries as follows. For $\lambda \geqslant 0$ the bounded positive solutions of

$$\lambda\phi - \omega\Delta\phi = 0 \tag{12}$$

form a convex set $\mathfrak{P}_\lambda$ endowed by a lattice structure similar to that of harmonic functions. Now there exists a 1-1 correspondence (lattice isomorphism) between the $\mathfrak{P}_\lambda$ for $\lambda > 0$ on one hand, and the harmonic functions dominated by u_A on the other hand. Thus the *active boundary is induced by each* $\mathfrak{P}_\lambda$ *for* $\lambda > 0$. An alternative interpretation may be given in terms of the resolvent $(\lambda-\omega\Delta)^{-1}$ and a discrete random walk associated with it.

3. The adjoint boundary

3.1. Duality. So far we have restricted consideration to operators acting on functions. Actually the study of an operator T from **C** to **C** cannot be separated from the study of the adjoint operator T^* which takes *measures* into measures. In probability T^* is the primary notion, although we are compelled to take T as the basic tool of the theory of Markov processes. The reasons are discussed in[7]; see also Dynkin[3].

To avoid new notations let D be an open domain of the plane and suppose that T is of the form (1) with an arbitrary *positive* kernel. The adjoint transformation acts on all measures, but it is convenient to consider only absolutely continuous measures and treat T^* as an operator on densities. Let then **L** be the Banach space of integrable functions in D with the usual norm

$$N(u) = \int_D |u|\, dq. \tag{13}$$

Then T^* as an operator from **L** to **L** carries $u \in \mathbf{L}$ into

$$T^*(u) = \int_D u(p)\, K(p, .)\, dp. \tag{14}$$

It is positive and $N(T^*) \leqslant 1$. The transformation (14) remains meaningful for all $u > 0$, although the integral may diverge.

In principle the construction of a boundary induced by T^* should follow the method used for T, but fortunately *an extremely simple trick will save us the trouble of a repeated construction.*

We start from the set $\mathfrak{P}^*$ of all finite eigenfunctions $u > 0$ of $u = T^*u$. They need not be integrable, but for simplicity we shall suppose that each $u \in \mathfrak{P}^*$ is strictly positive and continuous in D.

For an arbitrary $u \in \mathfrak{P}^*$ define a kernel $\tilde{K}$ by

$$\tilde{K}(p, q) = u(q)\, K(q, p)\, u^{-1}(p), \tag{15}$$

where $p \in D$, $q \in D$. Clearly

$$\int_D \tilde{K}(p, q)\, dq = 1 \tag{16}$$

for each p, and thus $\tilde{K}$ represents the kernel of a new transformation $\tilde{T}$ from **C** to **C**.

If $v \in \mathfrak{P}^*$ then $\phi = vu^{-1}$ is a continuous eigenfunction of $\tilde{T}\phi = \phi$, not necessarily bounded. Conversely, to each positive eigenfunction of $\tilde{T}\phi = \phi$ there corresponds the element $v = \phi u \in \mathfrak{P}^*$. This establishes a 1-1 correspondence between $\mathfrak{P}^*$ and the set $\tilde{\mathfrak{P}}^\infty$ of positive eigenfunctions of $\tilde{T}$. (These sets are endowed with lattice structures and the correspondence is a lattice isomorphism, but we omit these details.) In § 1.3 we have seen that $\tilde{T}$ induces a *total boundary* relative to the set $\mathfrak{P}^\infty$ of all positive eigenfunctions of $\tilde{T}\phi = \phi$. If in (15) we replace u by another element of $\mathfrak{P}^*$ then $\tilde{T}$ is replaced by a *similar* operator (as defined in § 1.3) and the total boundary remains unchanged. This justifies the

Definition. The adjoint boundary B^* induced by T is the total boundary induced by the operator $\tilde{T}$ acting from **C** to **C**. It is independent of the choice of $u \in \mathfrak{P}^*$.

Probabilistic interpretation. Consider first the case where $N(u) = 1$ and interpret u as the stationary probability density of the position (at any time) in a Markov chain with transition probability densities K. This process is defined for all integral values of the time parameter from $-\infty$ to ∞. In[16] Kolmogoroff pointed out that in this process $\tilde{K}(p,q)$ represents the conditional probability density of the position q at time n given that at time $n+1$ the position is p. Moreover, the same relationship exists between the higher order transition probability densities of the K-chain and the $\tilde{K}$-chain. In other words, for the K-chain, $\tilde{K}$ represents the transition probability densities in the negative time direction: the $\tilde{K}$-chain is obtained from the K-chain by *reversing the time direction.*

Roughly speaking, then, the boundary induced by T represents the directions towards which the sample sequences can converge, and the adjoint boundary the directions from which the process can originate. This description applies to, and becomes more concrete in connection with, continuous time parameter processes and leads to an interpretation of the boundary conditions for differential equations.

Analytically there is no change in the situation when the integral (13) diverges, and it is therefore annoying that Kolmogoroff's intuitive interpretation of $\tilde{K}$ breaks down. However, as Derman[2] pointed out, it may be salvaged for non-integrable u by considering a whole family of chains.

3.2. Relations between the two topologies. It is natural to ask whether and how the topology of $D \cup B$ induced by T is related to the topology of $D \cup B^*$ induced by T^*. A first answer is that every imaginable situation can arise. For the familiar symmetric operators of analysis the two spaces are identical. Still simpler is the other extreme, where B and B^* are disjoint and have disjoint neighborhoods. Most vexing are the intermediate cases. Among the examples given in[5] there appears the following configuration:

The boundary B consists of m points $\beta^{(1)}, \ldots, \beta^{(m)}$ and B^* consists of n points $\gamma^{(1)}, \ldots, \gamma^{(n)}$. Each deleted neighborhood of B is a deleted neighborhood of B^* and vice versa. However, each neighborhood of $\gamma^{(1)}$ contains a neighborhood of $\beta^{(1)} \cup \beta^{(2)}$ and each neighborhood of $\beta^{(1)}$ or $\beta^{(2)}$ is contained in a neighborhood of $\gamma^{(1)}$. Roughly speaking, the point $\gamma^{(1)}$ is the same as the set $\beta^{(1)} \cup \beta^{(2)}$. Similarly, two points of B may be equivalent to the union of the three points of B^*, etc.

These phenomena lead to many new problems connecting topological and analytical problems and are interesting in connection with boundary conditions; see the end of § 4.3.

4. Background and program

4.1. The problems. Given a topological space D an important problem of probability theory is to find the most general Markov process on D. (Hunt's beautiful results in [14] permit us to reformulate this in terms of *potentials*.) Space does not permit us here to analyze why and how this problem is reduced to that of finding semigroups of operators from **C** to **C**. Anyhow, the following slightly more general problem is of obvious interest in itself.

Find all operators Ω *generating* [*in the sense of* (8)] *positive semigroups* $\{T(t)\}$ *from* **C** *to* **C**. We have omitted the restriction $\|T(t)\| \leqslant 1$, which becomes more and more untenable even for probability theory and excludes semigroups of interest in potential theory, in diffusion (with creation of masses), and heat conduction. [We note in passing that our method of isomorphisms cannot be fully exploited as long as one adheres to the conventional Banach norm. It would be interesting and highly desirable to reformulate the whole theory free from restriction to normed spaces utilizing the *Köthe–Mackey* concept of dual spaces.]

Given a generator Ω we face the problem of *finding all positive semigroups generated by it and to discover the analytic and probabilistic relations among them.* The first part leads to a strict formulation of the vague and unsatisfactory notion '*boundary conditions*' for differential equations. The problem is to construct all possible lateral conditions; in this form it is analogous to the construction of self-adjoint contractions in Hilbert space, but it leads to new angles.

Finally, it is important to *find the generator of the adjoint semigroup.* (In special cases this amounts to finding the physicist's Fokker–Planck, or continuity, equation.)

This point of view links together operators which classically seemed worlds apart. For example, when D is the set of integers we are led to infinite systems of ordinary differential equations that can be treated precisely as partial differential equations; in fact, we obtain the boundary conditions with an analogue of '*normal derivatives*' at the boundary in a form which is applicable also, say, to harmonic functions in a domain with non-differentiable boundary. Again, when D is the real line, the semigroup for which $u(t, .) = T(t)f$ is harmonic in the half plane $t > 0$ is generated by a Riesz potential, and Elliott[4] has shown that its restriction to finite intervals is closely related to a second-order differential operator.

It is possible now to see, in rough outlines, the way to a rather general solution of the problems stated. Space does not permit us to go into details, and we proceed instead to indicate how the first problem is connected with an intrinsic theory of differential operators (or their analogue) in arbitrary spaces.

4.2. Operators of local character. The concept of a differential operator is defined only in special spaces and depends on a co-ordinate system. We replace it by the more meaningful and more general notion of an operator of local character. The positive semigroups generated by an operator of local character present an obvious analytic interest. (The corresponding Markov processes are the only ones whose path functions are continuous with probability one; see Ray[18].) They are a natural generalization of the second-order elliptic differential operators in Euclidean spaces: on one hand, such operators share the main properties of the Laplacian and generate positive semigroups. On the other hand, the derivation of the diffusion equation based on the local character condition (which I regret having introduced in [12] under the misleading name of Lindeberg condition) shows that no differential operator of higher order shares this property.

For an operator Ω of local character to generate a positive semigroup such that $\|T(t)\| \leqslant 1$ it is necessary, and very likely also sufficient, that it have the following *positive maximum property*: for each f in the (local) domain of Ω such that f attains a positive local maximum at the point p one has $\Omega f(p) \leqslant 0$. Dropping the norm condition $\|T(t)\| \leqslant 1$ leads to the *weak maximum condition* which requires $\Omega f(p) \leqslant 0$ only at points p such that $f(p) = 0$ and $f \leqslant 0$ in a neighborhood of p.

Operators of local character with the maximum property promise to be a fertile analogue in topological spaces of the Laplacian and general elliptic operators. It is a challenging problem (now within reach) to find a canonical expression for the operators of this class. A complete answer exists only in one dimension[9], but unpublished results of McKean point toward a solution in Euclidean spaces.

The notion of *induced boundaries* plays an important role in this connection. To explain it, consider an operator Ω of local character with the maximum property in a one-dimensional interval I. It is easily seen that Ω can have at most two independent annihilators. We may suppose, if necessary, that the domain of Ω has been extended so that Ω possesses the maximal number of annihilators compatible with its definition. We say then that the point $p \in I$ is regular if in a neighborhood N of p there exist two independent annihilators, that is, if the local topology induced by Ω agrees with the given topology. At a point p near which there exists only one annihilator, Ω induces a topology in which each interval has but one boundary point (is half open). This is reflected both in the analytic and the probabilistic properties of the corresponding processes. For example, all paths starting at a deficient point go in one direction (see Dynkin[3]). Similarly, in an arbitrary topological space D it will be desirable to avoid singularities and pathologies by requiring that in the interior of D the *local topology* induced by Ω agrees with the given topology.

An indication of the general character of our operators is obtained from the *canonical form* of an operator Ω of local character with the positive maximum property in an open interval I without singular points. It is given by the following theorem[9]. The interval I can be parametrized by a 'canonical scale' x in such a way that each f in the domain of Ω has one-sided derivatives with respect to x and they are continuous except for jumps. (We denote them indiscriminately by f'.) Moreover, there exist two Borel measures m and γ in I, the m-measure of each interval being positive, such that for each f in the domain of Ω,

$$\Omega f . dm = df' - f d\gamma \tag{17}$$

in the sense that the integrals are equal. Of course, if m and γ are absolutely continuous this canonical form reduces to $\Omega f = af'' - cf$. However, this traditional form requires artificial restrictions on the coefficients, whereas (17) is of an intrinsic nature and its theory is simpler and more flexible. Thus, in the equation of the vibrating string our m represents the mechanical mass and γ an attractive elastic force[10]. The form (17) now covers cases where either the mass or the force are concentrated at individual points, which are usually treated by artificial passages to the limit, whereas the generalized form of Ω makes available once and for all the basic existence and expansion theorem.

4.3. Lateral conditions. A *conditioning set* Σ for Ω is a set such that for $f \in \mathbf{C}$ and each $\lambda > \lambda_0$ there exists exactly one solution $F \in \Sigma$ of

$$\lambda F - \Omega F = f. \tag{18}$$

Using the Hille–Yosida theory it can be shown that each semigroup generated by Ω corresponds to a conditioning set. For example, the minimal semigroups generated by $\Omega = \omega\Delta$ (§ 2.3) correspond to restricting the domain of Ω to functions vanishing at the active boundary. The problem 'to find a semigroup generated by the differential operator Ω restricted by the conditioning set Σ' is the rigorous formulation of the vague and not always soluble problem of 'integrating the differential equation $u_t = \Omega u$ with the boundary condition $u \in \Sigma$'.

For simplicity consider an operator Ω such that for $\lambda > 0$ the equation $\lambda\xi - \Omega\xi = 0$ admits exactly m independent solutions $\xi \in \mathbf{C}$. Then the *active boundary* A induced by Ω consists of m points $\beta^{(1)}, \ldots, \beta^{(m)}$. For example, if Ω is a second-order differential operator in one dimension, say of the form (17), $m \leqslant 2$. All F in the domain of Ω will be continuous in $D \cup A$. The solution $F^{\min}$ of (18) corresponding to the *minimal* semigroup generated by Ω satisfies the boundary condition $F^{\min}(\beta^{(j)}) = 0$. Every other solution is of the form

$$F = F^{\min} + \sum_{j=1}^{m} \Phi^{(j)}\xi^{(j)}, \tag{19}$$

where $\xi^{(j)}$ is the solution of $\lambda\xi - \Omega\xi = 0$ such that $\xi^{(j)}(\beta^{(k)}) = 0$ or 1 according as $j \neq k$ or $j = k$, and where $\Phi^{(j)}$ is a certain functional of f.

Our problem now consists in determining these arbitrary functionals in their dependence on λ in such a way that *the range* Σ of *the transformation* $f \to F$ *is independent of* λ, and to describe all possible such ranges (conditioning sets). In [6] and [8] this problem is solved by a method of wide applicability. It is interesting that even in the case of differential equations the lateral condition $F \in \Sigma$ need not be of local character; hence the classical notion of boundary condition is too restrictive.

The *adjoint* semigroup is generated by an operator Ω^* acting on measures. It is noteworthy that even for a differential operator Ω the adjoint Ω^* need not be of local character. It appears, however, that for the minimal semigroup Ω^* shares the local character of Ω.

It is customary to consider (in Euclidean spaces) only differential operators Ω which are symmetric or so nearly symmetric that Ω and Ω^* induce the same boundary. New and interesting phenomena occur for truly unsymmetric operators. An abstract formulation of many classical problems is to find semigroups generated by Ω whose adjoint is generated by a given formal adjoint Ω^* of Ω.

Suppose that the minimal semigroup generated by Ω has an adjoint generated by Ω^*. Let the *adjoint boundary* consist of n points $\gamma^{(1)}, \ldots, \gamma^{(n)}$ corresponding to n independent solutions $\eta^{(k)}$ of $\lambda\eta - \Omega^*\eta = 0$. For the semigroup corresponding to (19) to be generated by Ω^* it is necessary and sufficient that each functional $\Phi^{(j)}$ is a linear combination of the $\eta^{(k)}$, so that the general solution of our problem involves mn free parameters to be determined in such a way that the range of the transformation (19) be independent of λ.

The solution of this problem given in [6] shows clearly the abstract generalization of the *normal derivatives* appearing in the mixed boundary value problem for harmonic functions. (One advantage of the abstract approach is to derive the boundary conditions in a form that is always meaningful, whereas the classical normal derivatives impose a regularity condition on the boundary.)

The general solution seems also to indicate that the topological question whether $\beta^{(j)}$ and $\gamma^{(k)}$ have disjoint neighborhoods is related to the behavior of the inner product of $\xi^{(j)}$ with $\eta^{(k)}$ as $\lambda \to \infty$ (see § 3.2). However, the precise way in which the relations between the two topologies induced by Ω are reflected in the analytical properties of Ω is an open and promising problem.

REFERENCES

[1] Brelot, M. Le problème de Dirichlet. Axiomatique et frontière de Martin. *J. Math. pures appl.* (9), 35, 297–335 (1956).

[2] Derman, C. A solution to a set of fundamental equations in Markov chains. *Proc. Amer. Math. Soc.* 5, 332–334 (1954).

[3] Dynkin, E. B. The infinitesimal operators of Markov processes. [In Russian, English summary.] *Teoria Verojatnostei i le Primenenija*, 1, 38–60 (1956).

[4] Elliott, J. The boundary value problems and semi-groups associated with certain integro-differential operators. *Trans. Amer. Math. Soc.* 75, 300–331 (1954).

[5] Feller, W. Boundaries induced by non-negative matrices. *Trans. Amer. Math. Soc.* 83, 19–54 (1956).

[6] Feller, W. On boundaries and lateral conditions for the Kolmogorov differential equations. *Ann. Math.* (2), 65, 527–570 (1957).

[7] Feller, W. The general diffusion operator and positivity preserving semi-groups in one dimension. *Ann. Math.* (2), 60, 417–436 (1954).

[8] Feller, W. Generalized second order differential operators and their lateral conditions. *Ill. J. Math.* 1, 459–504 (1957).

[9] Feller, W. On the intrinsic form for second order differential operators. *Ill. J. Math.* 2, 1–18 (1958).

[10] Feller, W. On the equation of the vibrating string. *J. Math. Mech.* 8, 339–348 (1959).

[11] Feller, W. The parabolic differential equation and the associated semi-groups of transformations. *Ann. Math.* (2), 55, 468–519 (1952).

[12] Feller, W. Zur Theorie der stochastischen Prozesse. *Math. Ann.* 113, 113–160 (1936).

[13] Hille, E. and Phillips, R. S. *Functional Analysis and Semi-groups.* Revised edition. Amer. Math. Soc. Colloquium Publ. no. 31 (1957).

[14] Hunt, G. A. Markoff processes and potentials. I–III. *Ill. J. Math.* 1, 44–93 (1957); 2, 151–213 (1958).

[15] Kadison, R. V. A representation theory for commutative topological algebra. *Mem. Amer. Math. Soc.* no. 7 (1951).

[16] Kolmogoroff, A. Zur Theorie der Markoffschen Ketten. *Math. Ann.* 112, 155–160 (1935).

[17] Martin, R. S. Minimal positive harmonic functions. *Trans. Amer. Math. Soc.* 49, 137–172 (1941).

[18] Ray, D. Stationary Markov processes with continuous paths. *Trans. Amer. Math. Soc.* 82, 452–493 (1956).

[*Added in proof*

[19] Doob, J. L. Discrete potential theory and boundaries. To appear in the *J. Math. Mech.*]