TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 156, May 1971

THEORY OF RANDOM EVOLUTIONS WITH APPLICATIONS TO PARTIAL DIFFERENTIAL EQUATIONS

BY
RICHARD GRIEGO AND REUBEN HERSH(1)

Abstract. The selection from a finite number of strongly continuous semigroups by means of a finite-state Markov chain leads to the new notion of a random evolution. Random evolutions are used to obtain probabilistic solutions to abstract systems of differential equations. Applications include one-dimensional first order hyperbolic systems. An important special case leads to consideration of abstract telegraph equations and a generalization of a result of Kac on the classical n-dimensional telegraph equation is obtained and put in a more natural setting. In this connection a singular perturbation theorem for an abstract telegraph equation is proved by means of a novel application of the classical central limit theorem and a representation of the solution for the limiting equation is found in terms of a transformation formula involving the Gaussian distribution.

In a little-known section of an out-of-print book [7], Mark Kac has considered a particle which moves on a line at speed v, and reverses direction according to a Poisson process with intensity a. After showing that such a motion is governed by a pair of partial differential equations, Kac comments, "The amazing thing is that these two equations can be combined into a hyperbolic equation"—namely, the telegraph equation,

$$\frac{1}{v}\frac{\partial^2 F}{\partial t^2} = v\frac{\partial^2 F}{\partial x^2} - \frac{2a}{v}\frac{\partial F}{\partial t}.$$

Kac refers to an earlier discussion of this model by Sidney Goldstein [3], and attributes its origin to G. I. Taylor. A few later papers [8], [13], follow up certain aspects of Kac's analysis.

The present research developed out of an attempt to understand this phenomenon of a stochastic process governed by a hyperbolic (rather than parabolic) equation.

Imagine a machine capable of various modes of operation. Let Λ be the set of possible positions of the control switch. To each $\lambda \in \Lambda$, there corresponds a mode

Received by the editors April 1, 1970.

AMS 1970 *subject classifications*. Primary 35B25, 35C15, 35F10, 35L15, 35L45, 47D05, 60J10; Secondary 35L05, 60F05.

Key words and phrases. Singular perturbations, integral representation of solutions, linear systems, hyperbolic systems, telegraph equation, wave equation, semigroups of operators, groups of operators, probabilistic solution of differential equations, Markov chains, central limit theorem.

(1) Research supported in part by NSF Grants GP-8856 and GP-5967.

of evolution of the machine. Suppose that each instantaneous state of the machine is represented by an element f in a Banach space $\boldsymbol{B}$, and each mode of evolution is given by a semigroup $T_\lambda(t)$ of linear operators on $\boldsymbol{B}$, with infinitesimal generator A_λ. Now suppose some random mechanism flips the switch from one position to another. In other words, we are given a stochastic process $\lambda(t)$ with values in Λ. What can we say about the outcome of the resulting random evolution?

In the present paper, we consider only the simple case when Λ is a finite set, and $\lambda(t)$ is a continuous-time Markov chain. By choosing A_i in interesting ways, we obtain a variety of new and old representation theorems and asymptotic formulas for Cauchy's problem.

For example, if A generates a *group* we are able to represent solutions of $u_{tt}=A^2u-2au_t$ as averages of solutions of $w_{tt}=A^2w$ with the parameter t randomized in a certain manner (Theorem 4).

This representation was found by Kac for $A=d/dx$. He then observed that his verification (by a direct and somewhat lengthy computation) held good also for the n-dimensional wave and telegraph equations, $n>1$. However, the probabilistic interpretation in n dimensions was not clear.

We hope that our formulation, in addition to providing a natural generalization of Kac's work, justifies his view that "one shouldn't have to compute anything at all. Such a statement ought to be provable by pure thought."

The last section of the present paper exploits our generalized Kac formula to treat a class of singular perturbations of Cauchy problems.

We use the classical central limit theorem to show that solutions of

$$\varepsilon u_{tt} = A^2u-2au_t$$

converge as $\varepsilon \to 0$ to solutions of

$$0 = A^2v-2av_t,$$

with $v(0)=u(0)$. The connection of singular perturbations of initial-value problems with the central limit theorem was pointed out by Birkhoff and Lynch [1]. In [8], Pinsky proves a central limit theorem by using a singular perturbation theorem for a concrete system of one-dimensional constant-coefficient hyperbolic equations.

A similar theorem has been proved by Smoller [12] and others, using the spectral representation of A. In these approaches it is required that iA be a selfadjoint operator on Hilbert space; we require only that A generate a uniformly bounded group on a Banach space.

This result has recently been improved and generalized by Schoene [11] by non-probabilistic methods, in an investigation inspired by the present work.

Of independent interest is formula (5.6), which expresses explicitly the semigroup generated by A^2 in terms of the group generated by A. (If $A=d/dx$, this is the familiar solution of the heat equation by a Gaussian kernel.) Formula (5.6) is extended to general higher-order abstract Cauchy problems in [5]. In a forthcoming

publication [14] formulas of this type are used for a nonprobabilistic solution of a broad class of singular perturbation problems.

The results of the present paper were announced in [4], which also includes an application to parabolic systems not discussed in the present paper.

1. **Notations.** For $i=1,\dots,n$, $\{T_i(t), t\geqq 0\}$ is a strongly continuous semigroup of bounded linear operators on a fixed Banach space $\boldsymbol{B}$. A_i is the infinitesimal generator of T_i.

$v=\{v(t), t\geqq 0\}$ is a Markov chain with state space $\{1,\dots,n\}$, stationary transition probabilities $p_{ij}(t)$, and infinitesimal matrix $Q=\langle q_{ij}\rangle=\langle p'_{ij}(0)\rangle$. P_i is the probability measure defined on sample paths $\omega(t)$ for v under the condition $\omega(0)=i$. E_i denotes integration with respect to P_i.

For a sample path $\omega\in\Omega$ of v, $\tau_j(\omega)$ is the time of the jth jump, and $N(t,\omega)$ is the number of jumps up to time t. $\gamma_j(t,\omega)$ is the occupation time in the jth state up to time t.

DEFINITION 1. A "random evolution" $\{M(t,\omega), t\geqq 0\}$ is defined by the product

$$M(t) = T_{v(0)}(\tau_1)T_{v(\tau_1)}(\tau_2-\tau_1)\cdots T_{v(\tau_{N(t)})}(t-\tau_{N(t)}). \tag{2.1}$$

We will see that the average of a random evolution defines a semigroup, the "expectation semigroup." But first we need two lemmas.

2. **Preliminary lemmas.** The first lemma states, roughly, that the number of jumps performed by a finite-state Markov chain is majorized by a Poisson-like expression.

LEMMA 1. *Let $0\leqq\alpha\leqq -q_{ii}\leqq\beta$ for all i. Then, for $m\geqq 0$, we have $P_i[N(t)=m]\leqq(\beta t)^m e^{-\alpha t}/m!$, for each $t\geqq 0$ and $i=1,\dots,n$.*

Proof. Let $\eta_i=-q_{ii}$. We note $\sum_{j\neq i} q_{ij}=\eta_i$. Also let $p_{ij}^{(m)}(t)=P_i[N(t)=m, v(t)=j]$. We then have $p_{ij}^{(0)}(t)=\delta_{ij}\exp(-\eta_i t)$ and

$$p_{ij}^{(m)}(t) = \sum_{k\neq i}\int_0^t \exp(-\eta_i(t-s))q_{ik}p_{kj}^{(m-1)}(s)\,ds;$$

see [2, p. 228]. Now,

$$\begin{aligned}
p_{ij}^{(m)}(t) &\leqq \sum_{k_1\neq i}\int_0^t \exp(-\alpha(t-t_1))q_{ik_1}p_{k_1j}^{(m-1)}(t_1)\,dt_1\\
&\leqq \sum_{k_1\neq i}\sum_{k_2\neq k_1}\int_0^t\int_0^{t_1} \exp(-\alpha(t-t_1))\exp(-\alpha(t_1-t_2))q_{ik_1}q_{k_1k_2}p_{k_2j}^{(m-2)}(t_2)\,dt_2\\
&\ \ \vdots\\
&\leqq \sum_{k_1\neq i}\cdots\sum_{k_m\neq k_{m-1}}\int_0^t\cdots\int_0^{t_{m-1}} \exp(-\alpha(t-t_m))\\
&\qquad\qquad\cdot q_{ik_1}\cdots q_{k_{m-1}k_m}p_{k_mj}^{(0)}(t_m)\,dt_m\cdots dt_1\\
&\leqq \sum_{k_1\neq i}\cdots\sum_{k_m\neq k_{m-1}}\int_0^t\cdots\int_0^{t_{m-1}} e^{-\alpha t}q_{ik_1}\cdots q_{k_{m-1}k_m}\delta_{k_mj}\,dt_m\cdots dt_1\\
&= \frac{t^m}{m!}e^{-\alpha t}\sum_{k_1\neq i}\cdots\sum_{k_m\neq k_{m-1}} q_{ik_1}\cdots q_{k_{m-1}k_m}\delta_{k_mj}.
\end{aligned}$$

Summing over j, we obtain

$$\begin{aligned} P_i[N(t) = m] &\leqq \frac{t^m}{m!} e^{-\alpha t} \sum_{k_1 \neq i} \cdots \sum_{k_m \neq k_{m-1}} q_{ik_1} \cdots q_{k_{m-1}k_m} \\ &= \frac{t^m}{m!} e^{-\alpha t} \sum_{k_1 \neq i} \cdots \sum_{k_{m-1} \neq k_{m-2}} q_{ik_1} \cdots q_{k_{m-2}k_{m-1}} \eta_{k_{m-1}} \\ &\leqq \frac{\beta t^m}{m!} e^{-\alpha t} \sum_{k_1 \neq i} \cdots \sum_{k_{m-1} \neq k_{m-2}} q_{ik_1} \cdots q_{k_{m-2}k_{m-1}} \\ &\leqq \frac{\beta^m t^m}{m!} e^{-\alpha t}. \qquad \text{Q.E.D.} \end{aligned}$$

In order to state Lemma 2 we need to introduce the concept of conditional expectation for Banach space valued random variables. The following theorem is known; see, for example [10].

THEOREM. *Let $(\Omega, \mathscr{F}, P)$ be a probability space and let $\boldsymbol{B}$ be a Banach space. Let $g: \Omega \to \boldsymbol{B}$ be a Bochner P-integrable function and let $\mathscr{G}$ be a σ-algebra, $\mathscr{G} \subset \mathscr{F}$. Then there exists a function $h: \Omega \to \boldsymbol{B}$ that is Bochner integrable, strongly measurable relative to $\mathscr{G}$, unique a.e. P, and*

$$\int_\Lambda g \, dP = \int_\Lambda h \, dP \quad \textit{for all } \Lambda \in \mathscr{G}.$$

We denote h by $E[g|\mathscr{G}]$ and call it the strong conditional expectation of g relative to $\mathscr{G}$ (and P). Our next lemma extends a standard formula on products of random variables to the case when the first factor is our random operator $M(t)$.

LEMMA 2. *If $g: \Omega \to \boldsymbol{B}$ is Bochner P_i-integrable for a fixed $i = 1, \ldots, n$, then for each $t \geqq 0$ the function $\omega \to M(t, \omega) g(\omega)$ is Bochner P_i-integrable and*

$$E_i[M(t)g|\mathscr{F}_t](\omega) = M(t, \omega) E_i[g|\mathscr{F}_t](\omega), \tag{2.2}$$

for almost all ω with respect to P_i, where $\mathscr{F}_t$ is the σ-algebra generated by the random variables $v(u)$, $0 \leqq u \leqq t$, that is, $\mathscr{F}_t$ is the past up to time t for the Markov chain.

Proof. By Theorem 3.7.4 of [6], $M(t, \cdot) g(\cdot)$ is P_i-Bochner integrable if and only if $M(t, \cdot) g(\cdot)$ is strongly measurable and $E_i[\|M(t)g\|] < \infty$.

In order to establish strong measurability we must show that there exists a sequence f_n of countably $\boldsymbol{B}$-valued measurable functions of ω so that $\lim_{n \to \infty} \|f_n - M(t)g\| = 0$, a.e. P_i. Such a sequence g_n converging to g exists by the strong measurability of g. In addition, by approximating the random variables involved in the definition of $M(t)$ by discrete random variables we obtain a sequence of countably valued operators $M_n(t)$ so that $f_n = M_n(t) g_n$ is countably $\boldsymbol{B}$-valued and measurable. The strong continuity of our given semigroups then yields the required strong convergence of f_n to $M(t)g$.

In proving $E_i[\|M(t)g\|]$ is finite we have Hille's estimate for each factor $T_i(t)$, namely $\|T_i(t)\| \leqq M_i \exp(k_i t)$. But since the number of jumps $N(t)$ is arbitrarily large, $M(t)$ may be a product of arbitrarily many factors $T_v(\tau_i)$. If the constants M_i are larger than one, then the random operator $M(t)$ may have arbitrarily large norm for fixed t. By Lemma 1 we find, nevertheless, that the norm has finite expected value. Let $M = \max_i M_i$ and $K = \max_i K_i$. Then by Lemma 1,

$$\begin{aligned} E_i[\|M(t)g\|] &= \sum_{m=0}^{\infty} E_i[\|M(t)g\| \mid N(t) = m]P_i[N(t) = m] \\ &\leqq e^{Kt} \sum_{m=0}^{\infty} M^m E_i[\|g\|]P_i[N(t) = m] \\ &\leqq e^{Kt} E_i[\|g\|] \sum_{m=0}^{\infty} M^m \frac{(\beta t)^m}{m!} e^{-\alpha t} \\ &= e^{(K-\alpha+M\beta)t} E_i[\|g\|] < \infty. \end{aligned}$$

Now, in order to prove (2.2), the idea, roughly, is that $M(t)$ depends only on the past $\mathscr{F}_t$ and should then factor out of the conditional expectation. By the above quoted theorem we note first that both conditional expectations in (2.2) make sense since both g and $M(t)g$ are Bochner P_i-integrable. Employing the above approximation to $M(t)$ we see that the operator $M_n(t)$ can be written $M_n(t) = \sum_k L_k I_{\Lambda_k}$ for some $\Lambda_k \in \mathscr{F}_t$, where L_k is a bounded linear operator and I_Λ is the indicator function of Λ. The rest of the proof repeats the standard argument for the case of real random variables, but we include it for completeness. We have for each $\Lambda \in \mathscr{F}_t$,

$$\begin{aligned} \int_\Lambda M_n(t)E_i[g \mid \mathscr{F}_t]\, dP_i &= \sum_k L_k \int_{\Lambda\cap\Lambda_k} E_i[g \mid \mathscr{F}_t]\, dP_i \\ &= \sum_k L_k \int_{\Lambda\cap\Lambda_k} g\, dP_i \qquad (\text{since } \Lambda \cap \Lambda_k \in \mathscr{F}_t) \\ &= \int_\Lambda \Big(\sum_k L_k I_{\Lambda_k}\Big) g\, dP_i = \int_\Lambda M_n(t) g\, dP_i. \end{aligned}$$

Hence $E_i[M_n(t)g \mid \mathscr{F}_t] = M_n(t)E_i[g \mid \mathscr{F}_t]$, a.e. P_i, for each n. By the continuity property of the strong conditional expectation (for example, Theorem 2.2 in [10]) as $n \to \infty$ we obtain (2.2). Q.E.D.

3. **The expectation semigroup.** Let $\tilde{\boldsymbol{B}}$ be the n-fold Cartesian product of $\boldsymbol{B}$ with itself. A generic element of $\tilde{\boldsymbol{B}}$ is denoted by $\tilde{f} = \langle f_i \rangle$ where $f_i \in \boldsymbol{B}$, $i = 1, \ldots, n$. We equip $\tilde{\boldsymbol{B}}$ with any appropriate norm so that $\|\tilde{f}\| \to 0$ as $\|f_i\| \to 0$ for each i.

DEFINITION 2. For $t \geqq 0$ define the (matrix) operator $\tilde{T}(t)$ on $\tilde{\boldsymbol{B}}$ specified componentwise by

$$(\tilde{T}(t)\tilde{f})_i = E_i[M(t)f_{v(t)}]. \tag{3.1}$$

As in Lemma 2, we see that $M(t)f_{v(t)}$ is Bochner integrable. Below we will show that $\tilde{T}(t)$ defines a semigroup on $\tilde{\boldsymbol{B}}$. We call $\tilde{T}(t)$ the "expectation semigroup" associated with the random evolution $M(t)$.

THEOREM 1. *$\{\tilde{T}(t), t \geqq 0\}$ is a strongly continuous semigroup of bounded linear operators on $\tilde{B}$.*

Proof. By the estimate $E_i[\|M(t)g\|] \leqq e^{(K-\alpha+M\beta)t}E_i[\|g\|]$ established in Lemma 2, we see that $\tilde{T}(t)$ is a bounded linear operator. Also, since $v(\cdot, \omega)$ is continuous at a fixed t for almost all sample paths ω, it is easy to see that $\tilde{T}(t)$ is strongly continuous in t.

We need only check the semigroup property. It suffices to show that for each i, $(\tilde{T}(t+s)\tilde{f})_i = (\tilde{T}(t)\tilde{T}(s)\tilde{f})_i$.

Let $\theta_t\omega$ be the shifted path defined by the requirement that $v(u, \theta_t\omega) = v(u+t, \omega)$ for every $u \geqq 0$. Define $g \circ \theta_t$ by $(g \circ \theta_t)(\omega) = g(\theta_t\omega)$. Then the Markov property of $v(t)$ is expressed by the formula $E_i[g \circ \theta_t | \mathscr{F}_t](\omega) = E_{v(t,\omega)}[g]$ for almost all $\omega[P_i]$. We can omit ω and write simply $E_i[g \circ \theta_t | \mathscr{F}_t] = E_{v(t)}[g]$. It is easy to check that $M(t)$ satisfies the multiplicative formula $M(t+s, \omega) = M(t, \omega)M(s, \theta_t\omega)$ or

$$M(t+s) = M(t)M(s) \circ \theta_t. \tag{3.2}$$

For fixed i we have

$$\begin{aligned}
(\tilde{T}(t+s)\tilde{f})_i &= E_i[M(t+s)f_{v(t+s)}] \\
&= E_i[E_i[M(t+s)f_{v(t+s)} | \mathscr{F}_t]] \\
&= E_i[E_i[M(t)M(s) \circ \theta_t f_{v(s)} \circ \theta_t | \mathscr{F}_t]] && \text{(by (3.2))} \\
&= E_i[M(t)E_i[M(s) \circ \theta_t f_{v(s)} \circ \theta_t | \mathscr{F}_t]] && \text{(by Lemma 2)} \\
&= E_i[M(t)E_{v(t)}[M(s)f_{v(s)}]] && \text{(by Markov property of } v) \\
&= E_i[M(t)(\tilde{T}(s)\tilde{f})_{v(t)}] \\
&= (\tilde{T}(t)\tilde{T}(s)\tilde{f})_i. \qquad \text{Q.E.D.}
\end{aligned}$$

REMARK. M. Pinsky has noted that if $M(t, \omega)$ is any operator-valued multiplicative functional of a Markov chain $v(t)$, i.e., $M(t, \omega)$ satisfies (3.2), and if $M(t, \omega)$ is strongly continuous in $t \geqq 0$ for each ω, then $M(t)$ admits the representation (2.1) as a random evolution for some semigroups $T_i(t)$. In fact, the "backward" order of the factors in our definition of $M(t, \omega)$ was chosen precisely in order to obtain the multiplicative property (3.2), which in turn implies the semigroup property for $\tilde{T}(t)$.

THEOREM 2. *The Cauchy problem for an unknown vector $\tilde{u}(t)$, $t > 0$,*

$$\frac{\partial \tilde{u}_i}{\partial t} = A_i\tilde{u}_i + \sum_{j=1}^{n} q_{ij}\tilde{u}_j, \qquad \tilde{u}(0) = \tilde{f}, \tag{3.3}$$

is solved by $\tilde{u}(t) = \tilde{T}(t)\tilde{f}$.

Proof. Let D_i be the domain of the infinitesimal generator A_i. We first prove that the infinitesimal generator of $\tilde{T}(t)$ is $\tilde{A} = \operatorname{diag}(A_1, \ldots, A_n) + Q$ with domain $D_1 \times \cdots \times D_n$. We need to show

$$(\tilde{A}\tilde{f})_i = A_i f_i + \sum_{j=1}^{n} q_{ij} f_j.$$

If τ is the first jump of v then

$$\begin{aligned}
(\tilde{T}(t)\tilde{f})_i &= E_i[M(t)f_{v(t)};\ \tau > t]+E_i[M(t)f_{v(t)};\ \tau \leqq t] \\
&= T_i(t)f_iP_i(\tau > t)+\int_0^t E_i[M(t)f_{v(t)}|\tau = r]P_i(\tau \in dr) \\
&= T_i(t)f_iP_i(\tau > t)+\int_0^t E_i[T_i(r)E_{v(r)}[M(t-r)f_{v(t-r)}|\tau = r]]P_i(\tau \in dr) \\
&= T_i(t)f_iP_i(\tau > t)+\int_0^t T_i(r)\sum_{j\neq i} E_j[M(t-r)f_{v(t-r)}]P_i(v(\tau) = j)P_i(\tau \in dr) \\
&= T_i(t)f_iP_i(\tau > t)+\int_0^t T_i(r)\sum_{j\neq i} (\tilde{T}(t-r)\tilde{f})_j\left(\frac{q_{ij}}{-q_{ii}}\right)(-q_{ii}p_{ii}(r))\,dr \\
&= T_i(t)f_iP_i(\tau > t)+\int_0^t T_i(r)\sum_{j\neq i} (\tilde{T}(t-r)\tilde{f})_jq_{ij}p_{ii}(r)\,dr.
\end{aligned}$$

Letting s-lim denote limit in the norm of $\boldsymbol{B}$ we thus have, for $f_i \in D_i$,

$$\begin{aligned}
(\tilde{A}\tilde{f})_i &= s\text{-}\lim_{t\searrow 0}\frac{1}{t}\,[(\tilde{T}(t)\tilde{f})_i-f_i] \\
&= s\text{-}\lim_{t\searrow 0} P_i(\tau > t)\,\frac{T_i(t)f_i-f_i}{t}+s\text{-}\lim_{t\searrow 0}\left(\frac{P_i(\tau > t)-1}{t}\right)f_i \\
&\quad +s\text{-}\lim_{t\searrow 0}\frac{1}{t}\int_0^t T_i(r)\sum_{j\neq i}(\tilde{T}(t-r)\tilde{f})_jq_{ij}p_{ii}(r)\,dr \\
&= A_if_i+q_{ii}f_i+\sum_{j\neq i} q_{ij}f_j \\
&= A_if_i+\sum_{j=1}^{n} q_{ij}f_j.
\end{aligned}$$

By standard semigroup theory we obtain the result of the theorem.

4. **The commutative case, and applications to hyperbolic equations.** If the given semigroups commute, we obtain a much simpler formula for $\tilde{T}$. Collecting terms and using the semigroup property for each T_i, we find

$$(\tilde{T}(t)\tilde{f})_i = E_i\left[\prod_{j=1}^{n} T_j(\gamma_j(t))f_{v(t)}\right]. \tag{4.1}$$

Now the jump times τ_j and the number of jumps $N(t)$ are no longer relevant. M depends only on the occupation times $\gamma_i(t)$.

We will repeatedly use the result of the following theorem.

THEOREM 3. *Suppose that the infinitesimal generators A_i are scalar multiples of a single operator, $A_i = c_iA$. Let A generate $T(t)$; we note that negative c_i's are allowed only if $T(t)$ is a group. Then*

$$u_i(t) = E_i\left[T\left(\sum_{j=1}^{n} c_j\gamma_j(t)\right)f_{v(t)}\right] \tag{4.2}$$

solves the system (3.3).

Proof. Now $T_j(t)=T(c_jt)$. By (4.1) we have

$$\begin{aligned} u_i(t) &= (\tilde{T}(t)\tilde{f})_i \\ &= E_i\left[\prod_{j=1}^{n} T(c_j\gamma_j(t))f_{v(t)}\right] \\ &= E_i\left[T\left(\sum_{j=1}^{n} c_j\gamma_j(t)\right)f_{v(t)}\right]. \end{aligned}$$

The random time $\sum c_j\gamma_j(t)$, over which we average, is just what one should expect: the sum of the occupation time in each state multiplied by the speed of evolution in that state.

As an application of Theorem 3, if $A=d/dx$, and $\boldsymbol{B}$ is $C_0(-\infty,\infty)$, then (3.3) is a one-dimensional first-order hyperbolic system in canonical (diagonal) form, and $T(t)f(x)=f(x+t)$. By (4.2) we obtain an elegant, and apparently new, solution formula for such a system:

$$\tilde{u}_i = E_i[f_{v(t)}(x+\textstyle\sum c_j\gamma_j(t))]. \tag{4.3}$$

(The fact that a one-dimensional hyperbolic system governs the evolution of a particle moving on a line with random velocity was noticed by Birkhoff and Lynch [1] and by Pinsky [8].)

In order to study second-order equations, we suppose from now on that $n=2$, $c_1=1$, $c_2=-1$ in Theorem 3. In other words, A generates a group $T(t)$, $-\infty<t<\infty$; $T_1(t)=T(t)$ gives the forward evolution, and $T_2(t)=T(-t)$ the backward evolution. The Markov chain now gives a random reversal of the direction of evolution. Furthermore, we require that the jump process $N(t)$ be a Poisson process. Thus, assume $N(t)$, $t\geqq 0$, is a given standard Poisson process with intensity $a>0$ and $N(0)=0$. The infinitesimal matrix is given by $Q=\left(\begin{smallmatrix}-a & a\\ a & -a\end{smallmatrix}\right)$ and the associated Markov chain $v(t)$ is probabilistically equivalent to the process with sample paths $(-1)^{N(t)}$ for starting point $v(0)=1$ and $(-1)^{N(t)+1}$ for $v(0)=2$.

The system (3.3) then becomes

$$\begin{aligned} du_1/dt &= Au_1-au_1+au_2, & u_1(0) &= f_1, \\ du_2/dt &= -Au_2+au_1-au_2, & u_2(0) &= f_2. \end{aligned} \tag{4.4}$$

Theorem 3 gives us a solution of (4.4) as $u_i(t)=E_i[T(\gamma_1(t)-\gamma_2(t))f_{v(t)}]$.

Letting $f_1=f_2=f\in\mathscr{D}(A^2)$ in (4.4) we easily verify that $v_1=\frac{1}{2}(u_1+u_2)$ solves the second-order Cauchy problem

$$v_1'' = A^2v_1-2av_1', \qquad v_1(0)=f, \qquad v_1'(0)=0; \tag{4.5}$$

where prime denotes d/dt. Our probabilistic solution to (4.5) reduces to

$$v_1(t) = E_1[T(\gamma_1(t)-\gamma_2(t))f/2]+E_2[T(\gamma_1(t)-\gamma_2(t))f/2]. \tag{4.6}$$

Similarly, if $f_1=f_2=g\in\mathscr{D}(A^2)$ in (4.4) then $v_2=\frac{1}{2}(u_1-u_2)$ solves

$$v_2'' = A^2v_2-2av_2', \qquad v_2(0)=0, \qquad v_2'(0)=Ag; \tag{4.7}$$

and the stochastic solution is

$$v_2(t) = E_1[T(\gamma_1(t)-\gamma_2(t))g/2]-E_2[T(\gamma_1(t)-\gamma_2(t))g/2]. \tag{4.8}$$

By linearity, $v=v_1+v_2$ solves

$$v'' = A^2v-2av', \qquad v(0)=f, \qquad v'(0) = Ag. \tag{4.9}$$

With a little more effort we can use our stochastic solution of (4.9) to show that solutions of (4.9) can be expressed in terms of solutions of $w''=A^2w$. More precisely, we have the following generalization of a result of Kac.

THEOREM 4. *Let A generate a strongly continuous group of bounded linear operators on a Banach space. If $w(t)$ is the unique solution of the abstract "wave equation"*

$$w_{tt} = A^2w, \qquad w(0)=f, \qquad w_t(0) = Ag, \tag{4.10}$$

where $f, g \in \mathscr{D}(A^2)$, then

$$u(t) = E\left[w\left(\int_0^t (-1)^{N(s)}\,ds\right)\right], \tag{4.11}$$

where $N(t)$ is a Poisson process with intensity $a>0$, solves the abstract "telegraph equation"

$$u_{tt} = A^2u-2au_t, \qquad u(0)=f, \qquad u_t(0) = Ag. \tag{4.12}$$

Proof. First, we easily verify that the difference of occupation times can be written

$$\gamma_1(t)-\gamma_2(t) = \int_0^t (-1)^{N(s)}\,ds \quad \text{for } v(0) = 1$$

and

$$\gamma_1(t)-\gamma_2(t) = -\int_0^t (-1)^{N(s)}\,ds \quad \text{for } v(0) = 2.$$

Now, $\frac{1}{2}[(T(t)+T(-t))f+(T(t)-T(-t))g]$ solves (4.10) and is by the assumed uniqueness equal to w. Letting $\tau(t)=\int_0^t (-1)^{N(s)}\,ds$ we have

$$\begin{aligned} u(t) &= E[w(\tau(t))] \\ &= E[(T(\tau(t))+T(-\tau(t)))f/2+(T(\tau(t))-T(-\tau(t)))g/2] \\ &= E_1[T(\gamma_1(t)-\gamma_2(t))f/2]+E_2[T(\gamma_1(t)-\gamma_2(t))f/2] \\ &\quad +E_1[T(\gamma_1(t)-\gamma_2(t))g/2]-E_2[T(\gamma_1(t)-\gamma_2(t))g/2] \\ &= v_1(t)+v_2(t) \end{aligned}$$

which we saw by (4.6), (4.8) and (4.9) solves (4.12).

Kac derived this formula in the special case $A=d/dx$ by a formal passage to the limit from a discrete random walk. He then observed that his a posteriori rigorous

verification of the formula was valid also for the n-dimensional wave and telegraph equations. This seemed rather mysterious, since there appeared to be no random process associated with the higher dimensional problems, in the way in which random alternation between right and left translation was associated with the one-dimensional problem.

Our abstract set-up makes clear what is going on. We now have $A^2=\Delta=\sum(\partial^2/\partial x_i^2)$, and the stochastic solution formula (4.6) comes from a random alternation between evolution forward or backward according to the generator $(\Delta)^{1/2}$. In one dimension, $(\Delta)^{1/2}=d/dx$ and the group is translation.

Given an equation $u_{tt}=Cu-2au_t$, our justification of representing u by formula (4.11) depends on whether C has a square root A which generates a group. In Theorem 6 below we show that such a square root exists if Cauchy's problem is well posed for $u_{tt}=Cu$. Applying this theorem to $C=\Delta$ on the appropriate L^2-space we obtain the justification of Kac's n-dimensional result.

5. **A singular perturbation problem.** Observe that the system

$$\begin{aligned} \frac{du_1}{dt} &= \frac{1}{\varepsilon^{1/2}}Au_1-\frac{a}{\varepsilon}u_1+\frac{a}{\varepsilon}u_2, \\ \frac{du_2}{dt} &= \frac{-1}{\varepsilon^{1/2}}Au_2+\frac{a}{\varepsilon}u_1-\frac{a}{\varepsilon}u_2 \end{aligned} \tag{5.1}$$

is equivalent to the second-order equation

$$\varepsilon u_{tt} = A^2u-2au_t. \tag{5.2}$$

We want to use (5.1) to study the asymptotic behavior of the solution of (5.2) as $\varepsilon\to 0^+$.

Let $N(t)$ be the Poisson process associated with the system (5.1) when $\varepsilon=1$. Employing a change of scale we have that $N_\varepsilon(t)=N(t/\varepsilon)$ is a Poisson process with intensity a/ε. Letting

$$\tau_\varepsilon(t) = \int_0^t (-1)^{N_\varepsilon(s)}\,ds, \tag{5.3}$$

we have $\tau_\varepsilon(t)=\varepsilon\tau(t/\varepsilon)$, where $\tau(t)$ is obtained from (5.3) for $\varepsilon=1$.

We thus see that (5.1) is obtained from (4.4) by replacing $T(t)$ by $T(t/\varepsilon^{1/2})$ and $N(t)$ by $N_\varepsilon(t)$. By the discussion in Theorem 4 we have that

$$\begin{aligned} u^\varepsilon(t) = E[&(T(\varepsilon^{1/2}\tau(t/\varepsilon))+T(-\varepsilon^{1/2}\tau(t/\varepsilon)))f/2 \\ &+(T(\varepsilon^{1/2}\tau(t/\varepsilon))-T(-\varepsilon^{1/2}\tau(t/\varepsilon)))g/2] \end{aligned} \tag{5.4}$$

solves (5.2) with initial conditions $u^\varepsilon(0)=f$, $u_t^\varepsilon(0)=Ag$.

We will need the following analogue of the Helly-Bray theorem. The proof is a straightforward adaptation to our situation of the proof of the classical theorem, so we omit it.

LEMMA 3. *For each real t let $T(t)$ be a bounded linear operator on the Banach space $\boldsymbol{B}$. Assume that $T(t)$ is a strongly continuous function of t, and that, for some constant M, it satisfies $\|T(t)\| \leqq M$ for all t. Let Z_n be a sequence of random variables converging in distribution to a random variable Z as $n \to \infty$. Then*

$$E[T(Z_n)f] \to E[T(Z)f] \quad \textit{strongly}, \quad \textit{for each } f \in \boldsymbol{B},$$

or in terms of the respective distribution functions,

$$\int_{-\infty}^{\infty} T(s)f\,dF_n(s) \to \int_{-\infty}^{\infty} T(s)f\,dF(s) \quad \textit{strongly}.$$

We are now ready for the singular perturbation theorem.

THEOREM 5. *Assume A generates a uniformly bounded strongly continuous group $T(t)$ of bounded linear operators on a Banach space. If u^ε is the unique solution, for $t>0$, of*

$$\varepsilon u^\varepsilon_{tt} = A^2u^\varepsilon - 2au^\varepsilon_t, \qquad u^\varepsilon(0) = f, \qquad u^\varepsilon_t(0) = Ag, \tag{5.5}$$

where $f, g \in \mathscr{D}(A^2)$, then for all $t \geqq 0$, $u^\varepsilon(t)$ converges strongly as $\varepsilon \to 0^+$ to

$$u^0(t) = (2\pi t/a)^{-1/2} \int_{-\infty}^{\infty} T(s)f \exp(-as^2/2t)\,ds, \tag{5.6}$$

and furthermore $u^0(t)$ is a solution, for $t>0$, of

$$u^0_t = A^2u^0/2a, \qquad u^0(0) = f. \tag{5.7}$$

Proof. By the above remarks we have that $u^\varepsilon(t)$ is given by (5.4) and we wish to use this representation to calculate the limit as $\varepsilon \to 0^+$.

We first note that as $t \to \infty$ the distribution of $(a/t)^{1/2}\tau(t)$ converges to $N(0, 1)$, the standard normal distribution. For, if τ_k is the time of the kth jump for the given Markov chain associated with $N(t)$ and $\tau_0 = 0$, the random variables $X_k = \tau_k - \tau_{k-1}$, $k \geqq 1$, are independent and have a common exponential distribution with parameter $a>0$, so that $E[X_k] = 1/a$ and $\operatorname{Var}[X_k] = 1/a^2$ for all k. For each n, if t_n is such that $X_1 + \cdots + X_n \leqq t_n \leqq X_1 + \cdots + X_{n+1}$ then $\tau(t_n) = \int_0^{t_n} (-1)^{N(s)}\,ds$ lies between

$$X_1 - X_2 + \cdots \pm X_n \quad \text{and} \quad X_1 - X_2 + \cdots \pm X_n \mp X_{n+1}.$$

By the strong law of large numbers, $t_n \sim n/a$ with probability one. Each difference $X_k - X_{k+1}$ has mean zero and variance $2/a^2$. Applying the central limit theorem to the sequence $X_1 - X_2$, $X_3 - X_4, \ldots$, we see that the distribution of

$$(a/(2n)^{1/2})(X_1 - X_2 + \cdots + X_{2n-1} - X_{2n})$$

converges to $N(0, 1)$. Then as $t \to \infty$ and $t_{2n} \sim 2n/a$, the distribution of $(a/t)^{1/2}\tau(t)$ converges to $N(0, 1)$.

This in turn implies that the distribution of $\varepsilon^{1/2}\tau(t/\varepsilon)$ converges as $\varepsilon \to 0$, for fixed t, to $N(0, t/a)$, the normal distribution with mean zero and variance t/a. By

Lemma 3 we have $E[T(\varepsilon^{1/2}\tau(t/\varepsilon))f] \to E[T(Z)f]$ strongly as $\varepsilon \to 0$ for each f, where Z is a $N(0, t/a)$-random variable. By the expression (4.3) for $u^\varepsilon(t)$ we have that

$$u^\varepsilon \to E[T(Z)f/2]+E(T(-Z)f/2]+E[T(Z)g/2]-E[T(-Z)g/2] \tag{5.8}$$

or, by the symmetry of Z,

$$u^\varepsilon \to E[T(Z)f], \quad \text{strongly as } \varepsilon \to 0, \tag{5.9}$$

that is,

$$u^0(t) \equiv \lim_{\varepsilon\to 0} u^\varepsilon(t) = (2\pi t/a)^{-1/2} \int_{-\infty}^{\infty} T(s)f \exp(-as^2/2t)\, ds. \tag{5.10}$$

To complete the proof we need to show that $u^0(t)$ solves (5.7). Since $\|T(t)\| \leqq M$ for some M, the integral in (5.10) converges uniformly, even after differentiation under the integral sign arbitrarily many times. Applying $A^2/2a$ under the integral sign is the same as replacing $T(s)f$ by $T''(s)f/2a$ (which exists, by the hypothesis on f). Integration by parts twice then gives the same result as differentiation under the integral sign with respect to t. The initial condition is satisfied because the Gaussian kernel in the integral acts like a δ-function as $t \to 0$. The proof is complete.

Formula (5.6) has previously been found (without reference to its probabilistic significance) by Romanoff [9]. It says that the solution of (5.7) is given by an average of random solutions of $w_{tt}=A^2w$, but now averaged with respect to a normally distributed time, instead of with respect to the Poisson distribution as before (when $\varepsilon>0$).

Convergence holds in Theorem 5 even if $T(t)$ is unbounded and the constant a is pure imaginary (see [11]). Of course, one must give up probabilistic methods to prove such a result. For imaginary a there is a physical interpretation in terms of relativistic quantum mechanics, instead of the previous physical interpretation in terms of a random evolution.

In applications, one meets equations of the form

$$u_{tt} = Cu-2au_t \quad \text{or} \quad \varepsilon u_{t\bar{t}} = Cu-2au_t. \tag{5.11}$$

To apply Theorem 4 or 5, it is necessary to know that a group $T(t)$ exists whose generator A satisfies $A^2=C$. This is true under the reasonable assumption that (5.11) is well posed for $a=0$.

THEOREM 6. *If C is a closed linear operator such that Cauchy's problem is well posed in $\boldsymbol{B}$ for $w_{tt}=Cw$, then there exists a strongly continuous group of operators $T(t)$ on $\boldsymbol{B}$, whose generator A satisfies $A^2=C$.*

Proof. Let $w(t)$ satisfy $w_{tt}=Cw$, $w(0)=f$, $w'(0)=0$. If we rewrite $w_{tt}=Cw$ as a first-order system, we find that $\left(\begin{smallmatrix} 0 & I \\ C & 0 \end{smallmatrix}\right)$ generates a group, and

$$\begin{pmatrix} w \\ w' \end{pmatrix} = \exp\left(t\begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix}\right)\begin{pmatrix} f \\ 0 \end{pmatrix}.$$

By a basic lemma of Hille, there exist constants M, k, such that

$$\left\| \exp t \begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix} \right\| \leqq M \exp (k|t|),$$

so that $\|w\| \leqq M \exp (k|t|)\|f\|$. Let

$$v(t) = (4\pi t)^{-1/2} \int_{-\infty}^{\infty} w(s) \exp (-s^2/4t)\, ds.$$

Evidently the integral converges strongly and uniformly, and can be differentiated under the integral sign. The same arguments as in the last paragraph of the proof of Theorem 5 now show that $v' = Cv$, $v(0) = f$. This means that C generates a strongly continuous semigroup. The well-known theory of fractional powers of closed operators developed by Bochner, Phillips and Balakrishnan (see [15] for a readable exposition) assures us that $-C$ has a square root which generates a holomorphic semigroup. If we let $iA = (-C)^{1/2}$, then clearly $A^2 = C$, and A is closed and densely defined with nonempty resolvent. From this it follows by Theorem 23.9.5 of Hille-Phillips [6] that A generates a strongly continuous group $T(t)$; indeed,

$$w(t) = \tfrac{1}{2}(T(t) + T(-t))f. \tag{5.12}$$

We are indebted to Jerry Goldstein for suggesting the use of the Bochner-Phillips-Balakrishnan theory in this argument. Notice that $T(t)$ by no means is unique. The "odd part" of $T(t)$ cancels out in (5.6) and (5.12). Theorems 4 and 5 do not depend on which group-generating square root of A^2 one uses, but only on the fact that one exists.

Bibliography

1. Garrett Birkhoff and Robert E. Lynch, *Numerical solution of the telegraph and related equations*, Proc. Sympos. Numerical Solution of Partial Differential Equations (Univ. of Maryland, 1965), Academic Press, New York, 1966, pp. 289–315. MR **34** #3800.

2. K. L. Chung, *Markov chains with stationary transition probabilities*, Die Grundlehren der math. Wissenschaften, Band 104, Springer-Verlag, Berlin, 1960. MR **22** #7176.

3. S. Goldstein, *On diffusion by discontinuous movements, and on the telegraph equation*, Quart. J. Mech. Appl. Math. **4** (1951), 129–156. MR **13**, 960.

4. R. J. Griego and R. Hersh, *Random evolutions, Markov chains, and systems of partial differential equations*, Proc. Nat. Acad. Sci. U.S.A. **62** (1969), 305–308.

5. R. Hersh, *Explicit solution of a class of higher-order abstract Cauchy problems*, J. Differential Equations **8** (1970), 570-579.

6. E. Hille and R. S. Phillips, *Functional analysis and semi-groups*, rev. ed., Amer. Math. Soc. Colloq. Publ., vol. 31, Amer. Math. Soc., Providence, R. I., 1957. MR **19**, 664.

7. M. Kac, *Some stochastic problems in physics and mathematics*, Magnolia Petroleum Co., Lectures in Pure and Applied Science, no. 2, 1956.

8. M. Pinsky, *Differential equations with a small parameter and the central limit theorem for functions defined on a finite Markov chain*, Z. Wahrscheinlichkeitstheorie Verw. Gebiete **9** (1968), 101–111. MR **37** #3651.

9. N. P. Romanoff, *On one-parameter groups of linear transformations*. I, Ann. of Math. (2) **48** (1947), 216–233. MR **8**, 520.

10. Frank S. Scalora, *Abstract martingale convergence theorems*, Pacific J. Math. (1961), 347–374. MR **23** #A684.

11. A. Schoene, *Semi-groups and a class of singular perturbation problems*, Indiana U. Math. J. **20** (1970), 247-263.

12. J. A. Smoller, *Singular perturbations of Cauchy's problem*, Comm. Pure Appl. Math. **18** (1965), 665–677. MR **32** #2709.

13. S. Kaplan, *Differential equations in which the Poisson process plays a role*, Bull. Amer. Math. Soc. **70** (1964), 264–268. MR **28** #1409.

14. L. Bobisud and R. Hersh, *Perturbation and approximation theory for higher-order abstract Cauchy problems*, Rocky Mt. J. Math. (to appear).

15. K. Yosida, *Functional analysis*, Springer-Verlag, New York, 1968.

University of New Mexico,
Albuquerque, New Mexico 87106
Courant Institute of Mathematical Sciences, New York University,
New York, New York 10012