A CLASS OF MARKOV PROCESSES ASSOCIATED WITH NONLINEAR PARABOLIC EQUATIONS

By H. P. McKean, Jr.*

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND THE ROCKEFELLER UNIVERSITY

Communicated by Mark Kac, August 31, 1966

Introduction.—The familiar connection between the Brownian motion and the differential operator $f \to f''/2$, based upon the fact that the Brownian transition function $(2\pi t)^{-1/2} \exp[-(b - a)^2/2t]$ is also the elementary solution of $\partial u/\partial t = (1/2)\partial^2 u/\partial a^2$, is the simplest and most fruitful instance of the connection between Markov processes with constant transition mechanism and linear parabolic equations. The purpose of this paper is to explain a similar connection between a wider class of Markov processes and certain *nonlinear* parabolic equations. Boltzmann's equation from statistical mechanics is a special case, as is Burgers' equation: $\partial u/\partial t = (1/2)\partial^2 u/\partial a^2 - u\partial u/\partial a$.

1. *Markov Processes with Constant Transition Mechanism.*—Given a nice topological space Q, let $x:[0,\infty) \to Q$ be a *sample path* with coordinates $x(t) \equiv x(t+) \in Q (t \geq 0)$, let Z be the associated *universal field* measuring $x(t)$ for each $t \geq 0$, and to each $a \in Q$ attach probabilities $P_a(B)$, defined for $B \in Z$, such that

(*a*) for each $B \in Z$, $P_a(B)$ is a (topological) Borel function of a,

(*b*) for each $a \in Q$, $P_a[x(0) \in db]$ is the unit mass at $b = a$.

$P_a(B)$ should be thought of as *the chance of the event B for paths starting at a.* Define $P_f(B) = \int_Q P.(B)df$ for any nonnegative mass distribution f on Q of total mass $+1$.

$M = [Q,Z,P]$ is *a Markov process with constant transition mechanism* if the separate motions $M_f = [Q,Z,P_f]$ are knit together according to the rule

$$P_f[x_t^+ \in B | x(s) : s \leq t] = P_a(B)$$

for each choice of f, $t \geq 0$, and $B \in Z$. x_t^+ stands for the *shifted path* $x(\cdot + t)$, $a \equiv x(t)$, and the identity is permitted to fail for a P_f-negligible class of paths.

Given $t \geq 0$, define a map of mass distributions by the rule

$$e(tD^*): \quad f \to P_f[x(t) \in db].$$

Because

$$P_a[x(t) \in B] = \int_Q P_a[x(s) \in db]P_b[x(t - s) \in B] \qquad (t \geq s),$$

$e(tD^*)$ is *multiplicative* $[e(tD^*) = e(sD^*)e((t - s)D^*)]$, as the notation suggests, and can be regarded as the (formal) exponential $\Sigma(tD^*)^n/n!$ of the (additive) *forward infinitesimal operator* $D^* = \lim_{t \downarrow 0} t^{-1}[e(tD^*) - 1]$. $u = e(tD^*)f$ is the (formal) solution of the parabolic equation $\partial u/\partial t = D^*u$ with initial data f.

E. B. Dynkin[2] gives a very good account of this whole development.

2. *A Wider Class of Markov Processes.*—Consider the setup of §1, modified as follows: (1) to each mass distribution f on Q of total mass $+1$ attach probabilities $P_f(B)$, defined for $B \in Z$, such that

(*a*) for each $B \in Z$, $P_f(B)$ is a Borel function of f,

(*b*) for each f, $P_f[x(0) \in db] = f$,

but do not suppose that $P_f(B) = \int_Q P_a(B)df$ *as in* §1 [$P_a(B)$ stands for $P_f(B)$ in case f is the unit mass at $b = a$]; (2) let each of the motions $M_f = [Q,Z,P_f]$ be

Markovian in the sense that for any $t \geq s \geq 0$ and $B \in Z$, $P_f[x_t^+ \in B|x(s): \ s \leq t]$ depends upon $a = x(t)$ alone, *but do not suppose that it coincides with* $P_a(B)$ *as in* §1; (3) instead of that, let

$$P_f[x_t^+ \in B|x(s):s \leq t] = P_u[B|x(0) = a]$$

with $a = x(t)$ and $u = P_f[x(t) \in db]$.

According to the customary nomenclature, $M = [Q,Z,P]$ is *a Markov process with nonconstant transition mechanism,* but of a very special class. Roughly speaking, the transition mechanism depends upon t *via the distribution* $u = P_f[x(t) \in db]$, *only.* The map $f \rightarrow u$ is still multiplicative, but now the infinitesimal operator $D^*f = \lim_{t \downarrow 0} t^{-1}[u - f]$ can be *nonadditive,* as will be illustrated below in actual examples. u is still a solution of $\partial u/\partial t = D^*u$, but the exponential formula $u = e(tD^*)f$ is not correct any more; a replacement will be indicated in §5.

3. *A Simple Example: 2-State Maxwellian Gas.*—Consider the space Q: ± 1, take independent e_1, e_2, etc. $= \pm 1$ with common distribution f, and let $n = n(t)$ be the population of an independent standard branching process with $n(0) = 1$ and distribution $P_f[n(t) = j] = e^{-t}(1 - e^{-t})^{j-1}$ $(j \geq 1)$, subject to the rule that, conditional on $n(s):s \leq t_1$, $n(t_1 + t_2)$ should be identical in law to the sum of $n(t_1)$ independent copies of $n(t_2)$. Define $P_f(B)$ to be the probabilities induced on the space of sample paths x: $[0, \infty) \rightarrow Q$ by the recipe $x(t) = e_1e_2 \ldots e_n$ $(t \geq 0)$.

$M = [Q,Z,P]$ satisfies all the conditions of §2, and after a little computation, it develops that

$$\frac{\partial u}{\partial t}(e_1) = D^*u(e_1) = \begin{matrix} u(-1)^2 + u(+1)^2 - u(+1) & e_1 = +1 \\ 2u(-1)u(+1) - u(-1) & e_1 = -1 \end{matrix}$$
$$= \int\int [u(e_1^*)u(e_2^*) - u(e_1)u(e_2)]de_2do.$$

$u(e)$ is short for $u(t,e)$, $*$ stands for the *collision:*

$$\begin{matrix} e_1 \\ e_2 \end{matrix} \rightarrow \begin{matrix} e_1^* \\ e_2^* \end{matrix} = \textit{either} \begin{matrix} e_1e_2 \\ e_2 \end{matrix} \textit{ or } \begin{matrix} e_1 \\ e_1e_2 \end{matrix},$$

$\int de_2$ means sum over $e_2 = \pm 1$, and $\int do$ means sum over the two possible outcomes of the collision $e \rightarrow e^*$. $\partial u/\partial t = D^*u$ is expressed in this form to exhibit its similarity to Boltzmann's equation, which also falls into the class described in §2; see McKean[5] and also the end of §5 for additional information on this subject.

4. *Diffusions.*—A simpler example is provided by the motion on $Q = R^1$ corresponding to

$$\frac{\partial u}{\partial t} = D^*u = \frac{1}{2}\frac{\partial^2 u}{\partial a^2} - \frac{\partial}{\partial a} c_1u,$$

with

$$c_1 = c_1[u] = E_f[x(t)] - a = \int_{R^1} (b - a)udb.$$

[u now stands for the density function $\frac{\partial}{\partial b} P_f[x(t) < b]$. f is also regarded as a density function; this will simplify the notation slightly.] A natural conjecture is that this motion should be a standard Brownian motion subject to a *drift = the displacement from the average position,* meaning by this a solution of

$$x(t) = x(0) + y(t) + \int_0^t [\text{ave}(x_1) - x]ds$$

with $x(0)$ distributed according to f, y an independent standard Brownian motion, and x_1 an independent copy of x, the average sign standing for the expectation over x_1. Because $E_f[x(t)] = E_f[x(0)] \equiv m$, the problem simplifies to $x = x(0) + y + \int_0^t [m - x]$, and it is easily verified that the solution

$$x(t) = e^{-t}x(0) + (1 - e^{-t})m + \int_0^t e^{-(t-s)}dy$$

satisfies all the conditions of §2. It is also easily verified that

$$u = \frac{\partial}{\partial b} P_f[x(t) < b] = \int_{-\infty}^{+\infty} [(1 - e^{-2t})]^{-1/2} \exp\left[- \frac{(b - e^{-t}a - (1 - e^{-t})m)^2}{1 - e^{-2t}} \right] f da.$$

Now consider a more elaborate problem:

$$\frac{\partial u}{\partial t} = D^* u = \frac{1}{2} \frac{\partial^2}{\partial a^2} c_2^2 u - \frac{\partial}{\partial a} c_1 u$$

with coefficients c_1 and c_2 expressible as integrals

$$c = c[u] = \int_{R^n} c(a,b_1,\ldots,b_n) u \otimes \ldots \otimes u \, db_1 \ldots db_n.$$

Guided by the above example, it is natural to conjecture that the underlying motion is the solution of

$$x(t) = x(0) + \int_0^t \text{ave}[c_2(x,x_1,\ldots,x_n)]dy + \int_0^t \text{ave}[c_1(x,x_1,\ldots,x_n)]ds$$

with $x(0)$ distributed according to f, and y an independent standard Brownian motion; x_1, etc., being independent copies of x; and the average sign standing for the expectation over these copies. This recipe is easily checked if the integrands $c(a, b_1$, etc.) belong to $C^1(R^{n+1})$.

Burgers' equation

$$\frac{\partial u}{\partial t} = D^* u = \frac{1}{2} \frac{\partial^2 u}{\partial a^2} - u \frac{\partial u}{\partial a}$$

corresponds to the case $c_2 = 1$, $c_1 = u(a)/2$. J. Cole[1] expressed the solution as

$$\exp\left(\int_{-\infty}^{a} u\right) = \int_{-\infty}^{+\infty} \frac{e^{-(b-a)^2/2t}}{\sqrt{2\pi t}} \exp\left(\int_{-\infty}^{b} f\right) db.$$

The underlying motion is the solution of

$$x(t) = x(0) + y(t) + \frac{1}{2} \int_0^t u[s,x(s)]ds = x(0) + y(t) + \text{ave(t)},$$

t being the *local time*

$$\text{t} = \lim_{\epsilon \downarrow 0} (4\epsilon)^{-1} \text{ measure } (s \leq t: \ |x_1(s) - x(s)| < \epsilon)$$

that x spends in coincidence with an independent copy x_1 before time t.

5. *Tagged Molecule in a Bath of Like Molecules.*—A process M of the class described in §2 can be pictured as the motion of a tagged molecule in a bath of infinitely many like molecules; the first instance of such a model is due to Kac.[4]

Consider $n < \infty$ like molecules performing a joint Markovian motion with constant transition mechanism on the n-fold product Q^n of Q and let the corresponding (additive) *backward* infinitesimal operator be

$$D_n = \sum_{i \leq n} D(i) + \binom{n}{1}^{-1} \sum_{i<j\leq n} D(ij) + \binom{n}{2}^{-1} \sum_{i<j<k\leq n} D(ijk) + \ldots + D(12\ldots n).$$

[*A backward* operator D is the dual of a *forward* infinitesimal operator D^* as described in §1.] $D(1)$ is an (additive) 1-molecule backward infinitesimal operator acting on the first coordinate of Q^n. $D(i)$ is a copy of $D(1)$ acting on the ith coordinate ($i \leq n$). $D(12) = D(21)$ is a 2-molecule operator acting on the pair of coordinates 12. $D(ij)$ is a copy of $D(12)$ acting on the pair ij ($i < j \leq n$), etc. $D(1)$ governs the motion of molecule number 1 in isolation, $D(12)$ governs the interaction (double collisions) between molecules 1 and 2, $D(123)$ the (triple) collisions between molecules 1, 2, 3, etc. D_n^* commutes with permutations of the coordinates so that if the initial n-molecule distribution f^n is symmetrical, then so is $v = e(tD_n^*)f^n$ at any later time.

Given a 1-molecule distribution function f, let f^n be the outer product $f\otimes \ldots \otimes f$ so that the molecules are independent at time $t = 0$ with common distribution f. A formal computation indicates that as $n \uparrow \infty$, $v = e(tD_n^*)f^n$ tends to the infinite outer product $u^\infty = u\otimes u\otimes$, etc., of the 1-molecule distribution function $u = \lim_{n\uparrow\infty} P_{f^n}[x_1(t) \in db]$. Kac,[4] describing the first known instance of this phenomenon, called it *the propagation of chaos.* u is the (formal) solution of $\partial u/\partial t = D^*u$, D^* being the dual of the (additive) map $D = D(1) + D(12) + D(123) +$ etc. applied to the infinite outer product $f^\infty = f\otimes f\otimes$, etc., i.e.,

$$\int_Q D^*f\varphi = \sum_{n=1}^{\infty} \int_{Q^n} f\otimes \ldots \otimes f D(12\ldots n)\varphi.$$

D is now extended to nice sums of tame functions on Q^∞ so as to be a derivation under the outer product of such functions, and with this extension, it develops that

$$\frac{\partial^n}{\partial t^n} \int_Q u\varphi = \int_{Q^{n+1}} u\otimes \ldots \otimes u\, D^n\varphi \qquad (n \geq 0),$$

which suggests the (formal) exponential sum

$$\int_Q u\varphi = \sum_{n=0}^{\infty} \frac{t^n}{n!} \int_{Q^{n+1}} f\otimes \ldots \otimes f\, D^n\varphi = \int_{Q^\infty} f^\infty e(tD)\varphi$$

and leads us to express the solution of $\partial u/\partial t = D^*u$ as

$$u = e(tD)^*f^\infty.$$

A more elaborate (formal) computation gives the rule

$$\lim_{n\uparrow 0} P_{f^n}[x_1(t_1) \in B_1,\ x_1(t_1 + t_2) \in B_2,\ x_1(t_1 + t_2 + t_3) \in B_3, \text{ etc.}] = \int_{Q^\infty} f^\infty e(t_1D)\varphi_1 e(t_2D)\varphi_2 e(t_3D)\varphi_3, \text{ etc.},$$

φ_1 being the indicator function of $B_1 \subset Q$, etc. Using this, it is easy to convince oneself that the motion of the (tagged) molecule number 1 satisfies all the conditions of §2.

Going backwards, if M is one of the motions of §2 and if its infinitesimal operator

D^* is expressible as the dual of $D = D(1) + D(12) +$ etc. applied to f^∞ as above, then M can be regarded as *the motion of a tagged molecule in an infinite bath of like molecules*, formally at least.

McKean[5] verified this recipe for the 2-state Maxwellian gas of §3, for a caricature of a Maxwellian gas proposed by Kac,[4] for a 3-dimensional Maxwellian gas with cutoff, and for a caricature of Burgers' equation. It is an open problem to substantiate it for a gas of hard balls, though the fact cannot be doubted. D. P. Johnson[3] has carried out the verification for the 2-state case $[\pm 1]$ under the condition that $D^*f(+1)$ is a real analytic function of $f(+1)$.

6. *Increase of Chaos.*—A very seductive conjecture is that *chaos increases* in the infinite gas. This means that if the initial ∞-molecule distribution is symmetrical but *not* chaotic [i.e., not an outer product $f^\infty = f\otimes f\otimes$, etc.], it becomes more nearly so as time passes. A symmetrical initial distribution can be expressed as an average of chaotic distributions f^∞, and granting that chaos propagates as in §5, the ∞-molecule distribution at time $t \geq 0$ will be the *same* average of $u^\infty = u\otimes u\otimes$, etc. $[u = e(tD)^*f^\infty]$. Now the expression

$$H = \lim_{n\uparrow\infty} \frac{1}{n}\int_{Q^n} \text{ave}[u^n] \lg \frac{\text{ave}\ [u^n]}{[\text{ave}\ u]^n} = \text{ave}\left[\int_Q u \lg \frac{u}{\text{ave}\ u}\right]$$

is ≤ 0 and vanishes only in the chaotic case. As such, it can be thought of the degree of nonindependence of the molecules (lack of chaos), and it is natural to believe that H should increase with time. H. Kesten [private communication] disproved this, so the problem is to find a pleasing definition of the degree of chaos that permits this conjecture to hold.

7. *A Conjecture about Parabolic Equations.*—Guided by the model of §5, it is clear that the motion underlying $\partial u/\partial t = \partial^2 u^{2n+1}/\partial a^2$ involves meetings at the *same* place of $n + 1$ molecules. Because three or more standard Brownian particles cannot meet, it is conjectured that solutions with general nonnegative initial data do not exist if $n \geq 2$. It is likewise conjectured that $\partial u/\partial t = \partial^2 u/\partial a^2 - \partial u^{n+1}/\partial a$ does not have good solutions for $n \geq 2$.

8. *Conclusion.*—A wide class of statistical mechanical transport problems can be expressed as $\partial u/\partial t = D^*u$. Because of the connection with the underlying motion via the formula $u = P_f[x(t) \in db]$, a new field of application of probabilistic methods is opened up. The model of a tagged molecule in an infinite bath of like molecules not only permits us to give an explicit formula for the solution $[u = e(tD)^*f^\infty]$, but in simple cases can actually clarify the statistical-mechanical problem.

* The partial support of the National Science Foundation (NSF GP 4364) is gratefully acknowledged.

[1] Cole, J. D., "On a quasi-linear parabolic equation occurring in aerodynamics," *Quart. Appl. Math.*, 9, 225–236 (1951).

[2] Dynkin, E. B., *Markov Processes* (Berlin: J. Springer, 1965).

[3] Johnson, D. P., thesis, Massachusetts Institute of Technology (unpublished).

[4] Kac, M., *Probability and Related Topics in the Physical Sciences* (New York: Interscience, 1958).

[5] McKean, H. P., Jr., "An exponential formula for solving Boltzmann's equation and its probabilistic interpretation," *J. Comb. Th.*, to be published.