Z. Wahrscheinlichkeitstheorie verw. Gebiete
33, 237–251 (1976)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

A Class of Nonlinear Partial Differential Equations and the Associated Markov Processes

Geoffrey C. Berresford

Department of Mathematics, Long Island University, C.W. Post Center, Greenvale, N.Y. 11548, USA

Table of Contents

1. Introduction

The connection between Markov processes and linear parabolic partial differential equations is well known. To take the simplest example, the transition density for Brownian motion

$$p(t, x, y) = \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2t}$$

is also the Green function (elementary solution) of the heat equation

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2}.$$

A second example comes from the Feynman-Kac formula:

$$p(t, x, y) = E_x\{e^{-\int_0^t k(X_\tau)\,d\tau}\,\delta(X_t - y)\} \tag{1}$$

is the elementary solution of the heat equation with a "cooling" term

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2} - k(x)\,p.$$

Here E_x stands for the expectation of Brownian motion X_t beginning at time zero at position x.

Each of the equations above is linear. Since most "real" problems are nonlinear, one would like to extend this relationship to nonlinear parabolic equations. For example, the evolution equations

$$\frac{\partial p}{\partial t}=\frac{1}{2}\frac{\partial^2 p}{\partial x^2}-p(1-p)$$

and

$$\frac{\partial p}{\partial t}=\frac{1}{2}\frac{\partial^2 p}{\partial x^2}-p(1-p)^2$$

arise in the study of competing genotypes. They are of the form

$$\frac{\partial p}{\partial t}=\frac{1}{2}\frac{\partial^2 p}{\partial x^2}-p\,g(p) \tag{2}$$

where $g(p)$ is defined for $0\leqq p\leqq 1$, non-negative, and once (or more) differentiable. Instead of (2) with $x\in\mathbb{R}$ we let x take only integer values, making things technically simpler. Thus (2) is replaced by

$$\frac{\partial p}{\partial t}=\frac{1}{2}\Delta p-p\,g(p) \tag{3}$$

with initial data

$$p(0,x)=p_0(x) \tag{4}$$

where $x\in L=\{\dots,-2,-1,0,1,2,\dots\}$, $t\geqq 0$, and here and henceforth Δ stands for the second symmetric difference in x:

$$\Delta p(t,x)=p(t,x-1)-2p(t,x)+p(t,x+1).$$

The question arises: *What kind of a Markov process has transition density satisfying* (3) *and* (4) *and what is its interpretation?* We propose to make such a process out of simple random walks. Instead of a single particle, as in the first two examples, we take many, interacting according to a rule determined by g. The solution of (3)–(4) is then obtained as the probability density of a "typical" particle.

The Feynman-Kac formula (1) can be interpreted as the expectation of a function of the Brownian position X_t (in (1) we chose a δ function), with the added feature that the Brownian particle is "killed" at an exponentially distributed time T,

$$P\{T>t\,|\,X_\tau,\ \tau\leqq t\}=e^{-\int_0^t k(X_\tau)\,d\tau}.$$

Notice that the killing time depends on the path. In the present Markov process, the killing time of the first particle, say, depends on the paths of all the particles via the concentration of other particles at the location of the first. This feature introduces a nonlinearity into the forward equation. In this sense what we do is an extension of the Feynman-Kac formula.

With only minor modification one can change $g(p)$ to $g(p,x)$. One need only replace $g\left(\frac{\#_i(t)}{n}\right)$ by $g\left(\frac{\#_i(t)}{n},x_i(t)\right)$ where $x_i(t)$ is the location of the i-th particle at time t.

One can pass formally to a diffusion as follows. Let the times between jumps be exponentially distributed,[1] with mean $1/\gamma$, let the lattice width m go to zero and $\gamma\to\infty$, keeping $m^2\gamma=1$. Then the lattice forward equation (3) becomes the diffusion forward equation (2).

2. The N-particle Process

We place a "particle" at a point of the lattice

$$L=\{\dots,-2,-1,0,1,2,\dots\}$$

and associate with it an exponential holding time T, at which moment it steps a unit distance to the left or right, the two possibilities being equally likely. The distribution of T is

$$P\{T>t\}=e^{-\gamma t},\qquad t\geqq 0 \tag{0}$$

for some fixed $\gamma>0$. Henceforth we take $\gamma=1$. Each time a jump occurs, we assign the particle a new exponential holding time having the same distribution, independent of all past events. The renewal property of exponential holding times

$$P\{T>t+s\,|\,T>t\}=e^{-s}=P\{T>s\}$$

means that the random walk is Markovian. We call this a standard random walk. We now place n of these particles on L and let them move independently *except* that each particle is "killed" at a time determined by the others as follows. Let $\#_i(t)$ be the number of *other* particles which at time t are at the same place as the i-th particle. Then the killing time k_i for the i-th particle is defined as the first root $k=k_i$ of

$$\int_0^k g\left(\frac{\#_i(t)}{n}\right)dt=T_i'. \tag{1}$$

Here, $T_1',\dots,T_n'$ are new independent exponential holding times with law

$$P\{T_i'>t\}=e^{-t},$$

one to each particle; they are independent of each other and of everything else. At the moment k_i the i-th particle jumps to a state ∞ where it remains forever after. Since the event

$$k_i>t$$

is equivalent to

$$T_i'>\int_0^t g\left(\frac{\#_i(s)}{n}\right)ds$$

[1] See Section 2, Eq. (0).

we have the law that, conditional upon the paths [2] up to time t,

$$P\{k_i > t\} = P\left\{T_i' > \int_0^t g\left(\frac{\#_i(s)}{n}\right) ds\right\} = e^{-\int_0^t g\left(\frac{\#_i(s)}{n}\right) ds}.$$

Moreover, by the renewal property,

$$P\{k_i \in (t, t+dt] \mid k_i > t\} = g\left(\frac{\#_i(t)}{n}\right) dt.$$

with the same conditioning. Thus the killing Markovian: if a particle has not been killed by time t, it behaves as if it begins anew at time t, any past time spent in the company of other particles haaing no effect on its future. Stated more colloquially, a particle being killed "at a given moment" depends only on the number of other particles that are visiting it at that moment.

Recall that the particles are indistinguishable, since the initial distribution and interaction rules are the same for each particle. It is then appropriate to look not at the paths of the individual particles but at the (time dependent) empirical distribution [3]

$$\mathcal{X} = \{\mathcal{X}^t(j)\}$$

where

$$\mathcal{X}^t(j) = \frac{\text{number of particles at place } j \text{ at time } t}{n}.$$

Clearly $\mathcal{X}$ is a Markov process. We look at $\mathcal{X}$ because, although the n-particle process is random, the empirical distribution process $\mathcal{X}_n = \mathcal{X}$ for n particles converges (in probability) as $n \uparrow \infty$ to a deterministic process $\mathcal{X}_\infty$. This may be regarded as the "law of large numbers" for the process.

3. The Law of Large Numbers

Consider the space $\mathbf{X}_n$ of (possibly) defective empirical probability distributions of n particles on L, of total mass $0, 1/n, 2/n, \ldots, 1-1/n$, or 1. The defect will account for particles sent to ∞. Let $\mathbf{X}$ be the space of all (possibly) defective probability distributions on L. Notice that $\mathbf{X}_n$ becomes dense in $\mathbf{X}$ (in the usual weak topology). Also note that $\mathbf{X}$ is compact. Let $C(\mathbf{X})$ be the Banach space of continuous functions on $\mathbf{X}$ with the usual supremum norm. Let ψ be a C^1 function[4] on $\mathbf{X}$.

2 The paths are now on $L \cup \{\infty\}$.

3 Note that we will use script typeface to denote processes (e.g. $\mathcal{X}$, $\mathcal{Y}$) while saving X, Y etc. for fixed functions.

4 Let $X \in \mathbf{X}$ and let Y be a signed measure on L with total variation $\leqq 1$, such that for small ε, $X + \varepsilon Y \in \mathbf{X}$. Then ψ is a C^1 function on $\mathbf{X}$ if for all such X and Y

$$\psi(X+\varepsilon Y) = \psi(X) + \varepsilon \frac{\partial \psi}{\partial X} \cdot Y + o(\varepsilon) \quad \text{as } \varepsilon \to 0$$

where the linear map $\partial\psi/\partial X$ of Y into $\mathbb{R}^1$ is continuous and bounded as a function of X and $\varepsilon^{-1} o(\varepsilon) \to 0$ uniformly in X and Y.

The empirical distribution process $\mathscr{X}_n$ described in Section 2 moves in $\mathbb{X}_n$. Its generator at $X \in \mathbb{X}_n$ is easily computed:

$$\begin{aligned}
G_n\psi(X) &= \sum_{l\in L} nX(l)\left[\frac{1}{2}\psi\left(X-\frac{\delta(l)}{n}+\frac{\delta(l-1)}{n}\right)-\psi(X)+\frac{1}{2}\psi\left(X-\frac{\delta(l)}{n}+\frac{\delta(l+1)}{n}\right)\right]\\
&\quad+\sum_l nX(l)\,g\left(X(l)-\frac{1}{n}\right)\left[\psi\left(X-\frac{\delta(l)}{n}\right)-\psi(X)\right]\\
&=\sum nX(l)\frac{1}{n}\left[\frac{1}{2}\frac{\partial\psi}{\partial X}(-\delta(l)+\delta(l-1))+\frac{1}{2}\frac{\partial\psi}{\partial X}(-\delta(l)+\delta(l+1))\right]\\
&\quad+\sum nX(l)\,g\left(X(l)-\frac{1}{n}\right)\frac{1}{n}\frac{\partial\psi}{\partial X}(-\delta(l))+o(1)\\
&=\frac{\partial\psi}{\partial X}\left(\frac{1}{2}\sum X(l)\,(\delta(l+1)-2\delta(l)+\delta(l-1))\right)\\
&\quad+\frac{\partial\psi}{\partial X}\left(-\sum X(l)\,g\left(X(l)-\frac{1}{n}\right)\delta(l)\right)+o(1)\\
&=\frac{\partial\psi}{\partial X}\left(\frac{1}{2}\Delta X-Xg\left(X-\frac{1}{n}\right)\right)+o(1),
\end{aligned}$$

where $o(1)\to 0$ independently of X as $n\uparrow\infty$. Here $\delta(l)$ denotes the unit mass at the place $l\in L$. Note that all sums are finite. In the last line note that $Xg(X-1/n)$ stands for the function with value $X(l)\,g(X(l)-1/n)$ at the place l. We now let $n\to\infty$ and require $X_n\in\mathbb{X}_n$ to converge to $X\in\mathbb{X}$, obtaining

$$G_n\psi(X_n)\to G_\infty\psi(X)=\frac{\partial\psi}{\partial X}\left(\frac{1}{2}\Delta X-Xg(X)\right) \tag{1}$$

where the convergence is uniform in X. The limiting generator G_∞ is a first order operator and so regulates a deterministic process $\mathscr{X}_\infty$: the associated semigroup $\exp(tG_\infty)$ acts on $C(\mathbb{X})$ by translation along solutions of

$$\dot{X}=\tfrac{1}{2}\Delta X-Xg(X). \tag{2}$$

If g is non-negative and once continuously differentiable it is easy to prove existence and uniqueness of solutions to (2) in the space $\mathbb{X}$. Clearly the semigroup $\exp(tG_\infty)$ is strongly continuous and contractive.

The purpose of this section is to prove

The Law of Large Numbers for the $\mathscr{X}_n$ Process:

$\mathscr{X}_n\to\mathscr{X}_\infty$ in probability.

This will be done by proving the strong convergence of the semigroups $\exp(tG_n)\to\exp(tG_\infty)$. The main tool is the Trotter-Kato theorem:

Let $\mathbb{X}$ be a topological space and $C=C(\mathbb{X})$ the space of continuous functions on $\mathbb{X}$ with the supremum norm. Let $\mathbb{X}_n$ be a subset of $\mathbb{X}$ and C_n the restriction of C to $\mathbb{X}_n$. Let

$$\|\psi\|_n=\sup_{X\in\mathbb{X}_n}|\psi(X)|$$

and assume that $\mathbb{X}_n$ becomes dense in $\mathbb{X}$ so that $\|\psi\|_n \to \|\psi\|$ for $\psi \in C$. For $n=1,2,\ldots$, let $\exp(tG_n)$ be a strongly continuous semigroup of contraction operators on C_n with infinitesimal generator G_n and similarly for $\exp(tG_\infty)$ on C.

The Trotter-Kato Theorem. *If there is a subspace $C' \subset C$ such that*

(a) *C' is dense in the domain of G_∞,*

(b) *the closure of the restriction of G_∞ to C' is identical to G_∞, and*

(c) *for $\psi \in C'$, $\|G_n\psi - G_\infty\psi\|_n \to 0$,*

then $\exp(tG_n)$ *converges to* $\exp(tG_\infty)$, *in the sense that*

$$\|e^{tG_n}\psi - e^{tG_\infty}\psi\|_n \to 0$$

for every ψ in C.

The following smoothness condition on the limiting semigroup is sufficient:

If ψ is of class C^1 then so is $\exp(tG_\infty)\psi$, *and*

$$\left|\frac{\partial}{\partial X} e^{tG_\infty}\psi(X)\right| \leqq A e^{Bt} \tag{3}$$

with constants A and B independent of X and t.

We will now prove the sufficiency. Assume for simplicity that $B<1$. By the Hille-Yosida theorem,

$$R_1 = (1-G_\infty)^{-1}$$

is bounded and invertible on $C(\mathbb{X})$, and so maps C^1, which is dense[5] in $C(\mathbb{X})$, onto a set C' which is dense in the domain of G_∞.

We now show that this set fulfills the conditions for C' in the Taotter-Kato theorem. We have just verified condition (a). Choose a ψ in the domain of G_∞. Then approximate $(1-G_\infty)\psi$ by $\phi_n \in C^1$ and define

$$(1-G_\infty)^{-1}\phi_n = \psi_n \in C'.$$

By the continuity of R_1,

$$\psi_n = (1-G_\infty)^{-1}\phi_n \to (1-G_\infty)^{-1}(1-G_\infty)\psi = \psi.$$

Now this convergence combined with

$$\phi_n = (1-G_\infty)\psi_n \to (1-G_\infty)\psi$$

gives

$$G_\infty\psi_n \to G_\infty\psi.$$

So the closure of the restriction of G_∞ to C' is an extension of G_∞, but since G_∞ is closed, they must be equal. With condition (b) verified, only (c) remains. Recall

[5] The denseness is proved as follows. Any bounded function f on L defines a linear function on X via

$$\textstyle\sum_{l\in L} X(l)\, f(l).$$

Polynomials are then defined as finite linear combinations of finite products of such objects. The Stone-Weierstrass theorem implies that polynomials (which are a subset of C^1) are dense in $C(\mathbb{X})$ and so C^1 is dense.

that

$$\psi(X) \equiv (1-G_\infty)^{-1}\,\phi(X) = \int_0^\infty e^{-t}(e^{tG_\infty}\phi)(X)\,dt.$$

We want to show that the generators G_n converge on the function space C' in the sense that, for $\phi \in C^1$

$$G_n\psi(X) - G_\infty\psi(X) \tag{4}$$

is uniformly small on $\mathbb{X}_n$. This, however, comes from the smoothness condition (3). First

$$\psi(X+\varepsilon Y) = \int_0^\infty e^{-t}(e^{tG_\infty}\phi)(X)\,dt + \int_0^\infty e^{-t}\int_0^\varepsilon \frac{\partial}{\partial(X+\delta Y)}(e^{tG_\infty}\phi)\cdot Y\,d\delta\,dt.$$

The last member can be written

$$\begin{aligned}
&\int_0^\varepsilon d\delta \int_0^\infty e^{-t}\frac{\partial}{\partial(X+\delta Y)}(e^{tG_\infty}\phi)\cdot Y\,dt\\
&\quad= \int_0^\varepsilon d\delta \int_0^\infty e^{-t}\left[\frac{\partial}{\partial X}(e^{tG_\infty}\phi)\cdot Y + o(1)\right]dt\\
&\quad= \varepsilon\left[\int_0^\infty e^{-t}\frac{\partial}{\partial X}(e^{tG_\infty}\phi)\,dt\right]\cdot Y + o(\varepsilon),
\end{aligned}$$

as the reader may easily check, and

$$\int_0^\infty e^{-t}\frac{\partial}{\partial X}e^{tG_\infty}\phi\,dt$$

is bounded and continuous with respect to X. Thus ϕ is a C^1 function of X and this proves (4).

We now show that the smoothness condition (3) is satisfied by our semigroup $\exp(tG_\infty)$. We will prove the following

Lemma. *The solution $X = X^t$ of*

$$\dot{X} = \tfrac{1}{2}\Delta X - Xg(X) \tag{5}$$

is of class C^1 in its initial data X^0 and $\left|\dfrac{\partial X^t}{\partial X^0}\right| \leqq Ae^{Bt}$ with constants A, B independent of X^0 and t.

For $\phi \in C^1$, we can then write

$$\frac{\partial}{\partial X^0}(e^{tG_\infty}\phi) = \frac{\partial \phi}{\partial X^t}\cdot\frac{\partial X^t}{\partial X^0}$$

and the smoothness condition will be satisfied.

Proof of Lemma. Let $X_1^0 \in \mathbb{X}$ and let Y^0 be a signed measure on L with total mass $\leqq 1$, and such that for small ε, X_1^0 and $X_1^0 + \varepsilon Y^0$ are both in $\mathbb{X}$. Let X_1 and X_2 be the solutions to (5) with initial data (respectively) X_1^0 and $X_1^0 + \varepsilon Y^0$. The difference $X_3 = X_2 - X_1$ satisfies the equation

$$\begin{aligned}\dot{X}_3 &= \tfrac{1}{2}\Delta X_3 - [X_2 g(X_2) - X_1 g(X_1)] \\ &= \tfrac{1}{2}\Delta X_3 - [(X_2 - X_1) g(X_2) + X_1 (g(X_2) - g(X_1))] \\ &= \tfrac{1}{2}\Delta X_3 - (X_2 - X_1)\left[g(X_2) + X_1 \frac{g(X_2) - g(X_1)}{X_2 - X_1}\right] \\ &= \tfrac{1}{2}\Delta X_3 - X_3 k(t, l).\end{aligned} \tag{6}$$

where $k(t,l)$, defined by the square bracket in (6), is bounded since $g \in C^1[0,1]$ and the values of X are all between 0 and 1. Let E_l be the expectation for the random walk $W(t)$ beginning at place l at time $t=0$. Since X_3 has initial data εY^0, the Feynman-Kac formula gives

$$X_3^t(l) = \varepsilon E_l\{e^{-\int_0^t k(t-s,\, W(s))\,ds}\, Y^0(W(t))\}. \tag{7}$$

Therefore, as $\varepsilon \to 0$, X_2 tends to X_1 uniformly in l and uniformly in the initial data X_1^0 (since $|k|$ remains under a bound B depending only on g). We therefore have $g(X_2) \to g(X_1)$ and so, as $\varepsilon \to 0$, $k(t,x)$ converges to

$$\tilde{k}(t,x) = g(X_1^t) + X_1^t g'(X_1^t)$$

uniformly in the initial data X_1^0. Thus (7) gives

$$\begin{aligned}\lim_{\varepsilon \to 0} \frac{X_2^t(l) - X_1^t(l)}{\varepsilon} &= \frac{\partial X_1^t}{\partial X_1^0} \cdot Y^0(l) \\ &= E_l\{e^{-\int_0^t \tilde{k}(t-s,\, W(s))\,ds}\, Y^0(W(t))\}\end{aligned} \tag{8}$$

uniformly in the initial data. Clearly $\partial X_1^t / \partial X_1^0$ is continuous in X^0 and

$$\left|\frac{\partial X_1^t}{\partial X_1^0}\right| \leqq A e^{Bt}$$

with the constant A independent of X_1^0 and t. This completes the proof of the lemma.

We now conclude from the Trotter-Kato theorem that $\exp(tG_n)\psi = E\psi(\mathscr{X}_n^t)$ converges (uniformly in the initial data) to $\exp(tG_\infty)\psi = \psi(\mathscr{X}_\infty^t)$. The above is simply the statement that $\mathscr{X}_n$ converges in probability to $\mathscr{X}_\infty$, where $\mathscr{X}_\infty$ evolves by

$$\dot{\mathscr{X}}_\infty = \tfrac{1}{2}\Delta \mathscr{X}_\infty - \mathscr{X}_\infty g(\mathscr{X}_\infty).$$

This concludes the proof of the law of large numbers for $\mathscr{X}_n$.

4. Propagation of Chaos

The interaction between the particles in the n-particle process is determined by (2.1). The division by n suggests that the effect of any given particle on any other becomes very slight for large n. This suggests that if we start the particles independently at time zero, the independence will hold for later times in the limit $n \to \infty$. Let

$$p_{m,n}^t = p_{m,n}^t(l_1, \ldots, l_m), \qquad l_i \in L, \qquad m \leqq n,$$

be the marginal distribution at time t of the first m of the n particles. Since the distributions are symmetric it does not matter which particles we choose. The particular being initially distributed independently, each according to some common distribution p^0 means that $p^0_{m,n}$ factors:

$$p^0_{m,n}(l_1, \dots, l_m) = p^0(l_1) \dots p^0(l_m).$$

"Propagation of chaos" means that this same factoring of the marginals holds in the limit $n \uparrow \infty$ at later times $t > 0$. More precisely, for fixed m and $n \uparrow \infty$,

$$p^t_{m,n}(l_1, \dots, l_m) \to p^t(l_1) \dots p^t(l_m)$$

where p^t is the solution of

$$\frac{\partial p}{\partial t} = \frac{1}{2} \Delta p - p g(p) \tag{1}$$

with initial data p^0.

The meaning of p^t is clear from the statement for $m=1$: p^t is the distribution of a single "typical" particle at time t (having averaged over all possible positions of the other particles). Therefore the result is that in the limit $n \uparrow \infty$ the individual particles become independent at any fixed time with distribution regulated by (1).

Proof of Propagation of Chaos. Because of the symmetry of the distribution function we consider not the locations of the individual particles but events like

$$M = \{\text{among the first } m \text{ of the } n \text{ particles, } M(l) \text{ sit at place } l\}. \tag{2}$$

Let X be a fixed empirical distribution of n particles. Then the probability of the event M is

$$\pi(X) = \binom{n}{m}^{-1} \prod_l \binom{nX(l)}{M(l)}.$$

Now let $P_{m,n}\{M|X\}$ be the probability that at time t event M occurs, given that the particles have initial empirical distribution X. Let $P_{m,n}\{M|p^0 \otimes \dots \otimes p^0\}$ be the same except that the particles are originally distributed independently, each according to the common distribution p^0.

Step 1. We must prove the convergence[6]

$$P_{m,n}\{M|X\} \to \binom{m}{M} \prod_l [\mathscr{X}^t_\infty(l)]^{M(l)} \tag{3}$$

[6] In the next line we use the notation

$$\binom{m}{M} = \frac{m!}{\prod_l M(l)!}.$$

M stands for both a sequence $\{M(l)\}$ and an event. The meaning will be clear from the context.

uniformly in X. To do this we take $\mathscr{X}_n^0 = X$ and observe that

$$P_{m,n}\{M \mid X\} = e^{tG_n}\pi(X).$$

This converges uniformly in X to

$$e^{tG_\infty}\binom{m}{M}\prod_l \mathscr{X}_\infty^0(l)^{M(l)} = \binom{m}{M}\prod_l \mathscr{X}_\infty^t(l)^{M(l)},$$

completing step 1.

Step 2. We choose fixed positions $l_1, \ldots, l_m$; these determine occupation numbers $M = \{M(l)\}$:

$M(l)$ = the number of these position which are equal to l.

Let M also stand for the event (2). Then

$$\begin{aligned} p^t_{m,n}(l_1, \ldots, l_m) &= \binom{m}{M}^{-1} \sum_{(l', \ldots, l'_m)\in M} p^t_{m,n}(l'_1, \ldots, l'_m) \\ &= \binom{m}{M}^{-1} P_{m,n}\{M \mid p^0 \otimes \cdots \otimes p^0\} \\ &= \binom{m}{M}^{-1} \sum_{X\in \mathbb{X}_n} \left[\binom{n}{nX}\prod_l p^0(l)^{nX(l)}\right] P_{m,n}\{M \mid X\}. \end{aligned} \tag{4}$$

Now, by the classical weak law of large numbers, as $n\uparrow\infty$, the weights in the square bracket of (4) peak to unit mass at $X = p^0$. This, combined with the convergence (3) and the uniformity of the latter in X, proves that

$$p^t_{m,n}(l_1, \ldots, l_m) \to \prod_l \mathscr{X}_\infty^t(l)^{M(l)} = \mathscr{X}_\infty^t(l_1) \ldots \mathscr{X}_\infty^t(l_m).$$

This proves the propagation of chaos, since we know $p^t = \mathscr{X}_\infty^t$ solves (1).

5. Examples

If $g(p) = p$, then the i-th particle executes a random walk until the time k when

$$\int_0^k \frac{\#_i(t)}{n}\, dt \tag{1}$$

reaches an exponential holding time T_i'. But (1) says: add up the total time each other particle spends in coincidence with the i-th particle and divide by the number of particles. This can be interpreted as the average time that another "typical" particle on the lattice spends in coincidence with the i-th particle. The i-th particle is killed when this average coincidence time reaches the exponential holding time T_i'. The density for this process is the solution to

$$\frac{\partial}{\partial t} p = \tfrac{1}{2} \Delta p - p^2,$$

$$p(0, \cdot) = p^0(\cdot).$$

In the case $g(p) = 1 - p$ we have a particle being killed not when its average coincidence time becomes too high but when it is not high enough (compared with elapsed time t). This is a sort of communal survival rather than communal killing. Here the probability density solves

$$\frac{\partial p}{\partial t} = \frac{1}{2} \Delta p - p + p^2.$$

6. Fluctuations

This section is purely formal in intent; however it can be made completely rigorous. In Section 3 we showed that the $\mathscr{X}_n$ process converged in probability to a process $\mathscr{X}_\infty$ satisfying the equation

$$\frac{\partial}{\partial t} X = \frac{1}{2} \Delta X - X g(X). \tag{1}$$

One can consider the fluctuations about this limiting "mean" behavior. The question is analogous to passing from the weak law of large numbers, governing the empirical mean of a family of independent observations, to the central limit theorem, governing the fluctuation about the mean. We define the fluctuation process

$$\mathscr{Y}_n = \sqrt{n}\,[\mathscr{X}_n - \mathscr{X}_\infty] \tag{2}$$

where each $\mathscr{Y}_n(i)$ can be any real number, as distinct from the bounded $\mathscr{X}_n$ process, and we are suppressing the time variable t. We will show that

$$\mathscr{X}_n = \mathscr{X}_\infty + \frac{1}{\sqrt{n}} \mathscr{Y}_\infty + \text{lower order terms}$$

where $\mathscr{Y}_\infty$, the formal limit of $\mathscr{Y}_n$, behaves like an Ornstein-Uhlenbeck process.

Let ψ be a C^2 function[7] on $\mathbb{Y}$, where $\mathbb{Y}$ is the space of functions on the integers. Then the generator G_n of $\mathscr{Y}_n$ can be computed at a place $Y = \sqrt{n}\,[X_n - X_\infty]$ as

[7] Let $\mathbb{Z}$ be all functions on the integers which have compact support. We say Ψ is a C^2 function on $\mathbb{Y}$ if, for any $Y \in \mathbb{Y}$, $Z \in \mathbb{Z}$, one has

$$\Psi(Y + \varepsilon Z) = \Psi(Y) + \varepsilon \frac{\partial \Psi}{\partial Y} \cdot Z + \frac{1}{2} \varepsilon^2 \frac{\partial^2 \Psi}{\partial Y^2} \cdot Z \otimes Z + o(\varepsilon^2),$$

where $\partial \Psi / \partial Y$ is a linear map from $\mathbb{Z}$ into the real numbers, $\partial^2 \Psi / \partial Y^2$ is a linear map from $\mathbb{Z} \otimes \mathbb{Z}$ into the reals, and both are continuous in Y.
We will use the notation $Z^{\otimes 2}$ for $Z \otimes Z$.

follows:

$$G_n\psi(Y)=\sum_{l\in L} nX_n(l)\left[\frac{1}{2}\psi\left(Y-\frac{\delta(l)}{\sqrt{n}}+\frac{\delta(l-1)}{\sqrt{n}}\right)+\frac{1}{2}\psi\left(Y-\frac{\delta(l)}{\sqrt{n}}+\frac{\delta(l+1)}{\sqrt{n}}\right)-\psi(Y)\right]$$
$$+\sum_l nX_n(l)\,g\left(X_n(l)-\frac{1}{n}\right)\left[\psi\left(Y-\frac{\delta(l)}{\sqrt{n}}\right)-\psi(Y)\right]$$
$$-\frac{\partial\psi}{\partial Y}(\sqrt{n}(\tfrac{1}{2}\Delta X_\infty-X_\infty g(X_\infty)))$$
$$=\sum n\,X_n(l)\frac{1}{2}\left[\frac{\partial\psi}{\partial Y}\left(-\frac{\delta(l)}{\sqrt{n}}+\frac{\delta(l-1)}{\sqrt{n}}\right)+\frac{1}{2}\frac{\partial^2\psi}{\partial Y^2}\left(-\frac{\delta(l)}{\sqrt{n}}+\frac{\delta(l-1)}{\sqrt{n}}\right)^{\otimes 2}\right.$$
$$\left.+\frac{\partial\psi}{\partial Y}\left(-\frac{\delta(l)}{\sqrt{n}}+\frac{\delta(l+1)}{\sqrt{n}}\right)+\frac{1}{2}\frac{\partial^2\psi}{\partial Y^2}\left(-\frac{\delta(l)}{\sqrt{n}}+\frac{\delta(l+1)}{\sqrt{n}}\right)^{\otimes 2}\right]$$
$$+\sum n\,X_n(l)\,g\left(X_n(l)-\frac{1}{n}\right)\left[\frac{\partial\psi}{\partial Y}\left(-\frac{\delta(l)}{\sqrt{n}}\right)+\frac{1}{2}\frac{\partial^2\psi}{\partial Y^2}\left(-\frac{\delta(l)}{\sqrt{n}}\right)^{\otimes 2}\right]$$
$$-\sqrt{n}\frac{\partial\psi}{\partial Y}(\tfrac{1}{2}\Delta X_\infty-X_\infty\,g(X_\infty))+o(1).$$

It is easy to check that

$$\sqrt{n}(X_n\,g(X_n)-X_\infty\,g(X_\infty))=Y(g(X_\infty)+X_n\,g'(X_\infty))+o(1),$$

with $Y=\sqrt{n}(X_n-X_\infty)$. If we let ∇X stand for the first symmetric difference

$$(\nabla X)(l)=\tfrac{1}{2}X(l+1)-\tfrac{1}{2}X(l-1)$$

we get

$$G_n\psi(Y)=\sqrt{n}\frac{\partial\psi}{\partial Y}\cdot\frac{1}{2}\Delta X_n$$
$$+\frac{1}{4}\frac{\partial^2\psi}{\partial Y^2}(\tfrac{1}{2}\Delta\sqrt{X_n}-\nabla\sqrt{X_n})^{\otimes 2}+\frac{1}{4}\frac{\partial^2\psi}{\partial Y^2}(\tfrac{1}{2}\Delta\sqrt{X_n}+\nabla\sqrt{X_n})^{\otimes 2}$$
$$+\sqrt{n}\frac{\partial\psi}{\partial Y}(-X_n\,g(X_n))+\frac{1}{2}\frac{\partial^2\psi}{\partial Y^2}(\sqrt{X_n\,g(X_n)})^{\otimes 2}$$
$$-\sqrt{n}\frac{\partial\psi}{\partial Y}(\tfrac{1}{2}\Delta X_\infty-X_\infty\,g(X_\infty))+o(1)$$
$$=\frac{\partial\psi}{\partial Y}(\tfrac{1}{2}\Delta Y-Y(g(X_\infty)+X_n\,g'(X_\infty))$$
$$+\frac{1}{2}\frac{\partial^2\psi}{\partial Y^2}(\tfrac{1}{2}\Delta\sqrt{X_n})^{\otimes 2}+\frac{1}{2}\frac{\partial^2\psi}{\partial Y^2}(\nabla\sqrt{X_n})^{\otimes 2}$$
$$+\frac{1}{2}\frac{\partial^2\psi}{\partial Y^2}(\sqrt{X_n\,g(X_n)})^{\otimes 2}+o(1).$$

Therefore we have $G_n \to G_\infty$, at least formally, where

$$\begin{aligned} G_\infty \psi(Y) = \frac{1}{2} \frac{\partial^2 \psi}{\partial Y^2} [(\tfrac{1}{2}\Delta \sqrt{X_\infty})^{\otimes 2} + (\nabla \sqrt{X_\infty})^{\otimes 2} + (\sqrt{X_\infty g(X_\infty)})^{\otimes 2}] \\ + \frac{\partial \psi}{\partial Y} (\tfrac{1}{2}\Delta Y - Y(g(X_\infty) + X_\infty g'(X_\infty))). \end{aligned} \tag{3}$$

If we let $\mathscr{Y}_\infty$ be the formal limit[8] of $\mathscr{Y}_n$, we may write its generator (3) symbolically as

$$G_\infty \psi = \frac{1}{2} \frac{\partial^2 \psi}{\partial Y^2} \sigma^2 + \frac{\partial \psi}{\partial Y} m \tag{4}$$

where σ^2 is independent of Y, $m = M \cdot Y$ is linear in Y, and both σ^2 and m depend on t only via $\mathscr{X}_\infty$.

A sample (phase) point for $\mathscr{Y}_\infty$ is a function on the integers. This function then evolves in time. We make several observations from (3):

(a) $\sigma^2 \geqq 0$.

By this we mean the matrix σ^2 in $\frac{1}{2} \frac{\partial^2 \psi}{\partial Y^2} \sigma^2 = \frac{1}{2} \sum_{i,j} \frac{\partial^2 \psi}{\partial Y^2}(i,j)\, \sigma^2(i,j)$ is non-negative definite. This is so since σ^2 is the sum of three matrices each of the form $\sigma^2(i,j) = a(i)\, a(j)$ and such matrices are non-negative definite.

(b) *At time zero $\mathscr{Y}_\infty$ is Gaussian distributed (in its spatial variable).* This is because the particles are placed independently at time zero, and $\mathscr{Y}_\infty^0$ is just the limit of

$$\begin{aligned} &\sqrt{n}[X_n(l) - X_\infty(l)] \\ &= \frac{\text{number of particles at place } l \text{ at time zero } - n p^0(l)}{\sqrt{n}}. \end{aligned}$$

As $n \uparrow \infty$ the classical central limit theorem proves the claim.

We can compute the variance of

$$\int \lambda\, dY = \sum_l \lambda(l)\, Y(l)$$

for any $\lambda = \{\lambda(l)\}$, where $Y = \mathscr{Y}_\infty^0$, as follows:

$$E(\int \lambda\, dY)^2 = E(\sum \lambda(l) \sqrt{n}(X_n(l) - X_\infty(l)))^2 \tag{5}$$

---

[8] Note that the generators G_n and G_∞ are not homogeneous in time, but the dependence upon time is only via $\mathscr{X}_\infty$. We can rid ourselves of the time dependence by introducing the extra space variable $\mathscr{X}_\infty$, so we now look at the process of the *pair* $(\mathscr{Y}_n, \mathscr{X}_\infty)$ converging to the pair $(\mathscr{Y}_\infty, \mathscr{X}_\infty)$. The generator (3) will be altered by the addition of a term $\frac{\partial \psi}{\partial X_\infty}(\tfrac{1}{2}\Delta X_\infty - X_\infty g(X_\infty))$. This extra variable makes the generator homogeneous in time and we may then use the Trotter-Kato theorem. The only substantial change is that you must now use not once – but twice differentiable functions for making C'.

Let x_i be the actual location of the i-th particle. Since $nX_n(l)$ is the number of particles at l, (5) becomes

$$E\left(\frac{1}{\sqrt{n}}\sum_l \lambda(l)\, nX_n(l) - \sqrt{n}\int \lambda\, dX_\infty\right)^2$$

$$= E\left(\frac{1}{\sqrt{n}}\sum_1^n \lambda(x_i) - \sqrt{n}\int \lambda\, dp^0\right)^2 = \frac{n}{n}\int \lambda^2\, dp^0 + \frac{n(n-1)}{n}\left(\int \lambda\, dp^0\right)^2$$

$$-2n\left(\int \lambda\, dp^0\right)^2 + n\left(\int \lambda\, dp^0\right)^2 \to \int \lambda^2\, dp^0 - \left(\int \lambda\, dp^0\right)^2.$$

By polarization one easily finds

$$E\{\int \lambda\, dY \int \mu\, dY\} = \int \lambda \mu\, dp^0 - \int \lambda\, dp^0 \int \mu\, dp^0,$$

or, if we let λ, μ be indicator function of the sets A, B in Y we get the familiar correlation

$$E\{Y(A)\, Y(B)\} = p^0(A \cap B) - p^0(A)\, p^0(B),$$

which is related to the "Brownian bridge."

(c) *Once the process is Gaussian at epoch $t=0$ it remains Gaussian at any epoch t.* This is because we may write

$$d\mathcal{Y}_\infty = \sigma\, d\mathcal{E} + m\, dt \tag{6}$$

with $m = M \cdot Y$, or in coordinates,

$$d\mathcal{Y}_\infty(k) = \sum_{l \in L} \sigma_{k,l}\, d\mathcal{E}_l + \sum_l M_{k,l}\, \mathcal{Y}_\infty(l)\, dt,$$

where σ is a symmetric square root of σ^2, and the $\mathcal{E}$'s are independent standard Brownian motions. Since σ and M are independent of Y, we may solve (6) explicitly, given $\mathcal{Y}_\infty^0$, showing that $\mathcal{Y}_\infty^t$ remains Gaussian.

Note that the stochastic equation (6) is the "Boltzman" equation

$$X = \tfrac{1}{2}\Delta X - X g(X)$$

linearized around $\mathcal{X}_\infty$ and driven by a white noise $\delta\, d\mathcal{E}$. The parameters of this noise are actually determined by the requirement that the process has as its equilibrium distribution the one computed in (b) (at least when $\mathcal{X}_\infty$ is stationary). This has been the heuristic for construction "fluctuating B-equations" used by physicists. See [1] and [5].

(d) *In some cases m acts as a "restoring" force.* Take a g such that $g + x g' \geqq 0$ on $0 \leqq x \leqq 1$; for example $g(x) = x$. Then

$$m = M \cdot Y = \left(\tfrac{1}{2}\Delta - (g(X_\infty) + X_\infty\, g'(X_\infty))\right) \cdot Y$$

and both operators Δ and $-(g + X_\infty\, g')$ are negative, acting to "restore" Y to zero.

Properties (a)–(d) permit us to call $\mathcal{Y}_\infty$ an Ornstein-Uhlenbeck process. Although in our case σ^2 and M are not constant, because they depend upon $\mathcal{X}_\infty$, we may speak of $\mathcal{Y}_\infty$ as an Ornstein-Uhlenbeck process "guided" by $\mathcal{X}_\infty$.

The author wishes to thank his thesis advisor, Professor Henry P. McKean, Jr., who provided invaluable aid during all stages of this work.

References

1. Fox, R.F., Uhlenbeck, G.E.: Contributions to the Theory of Non-Equilibrium Thermodynamics I and II. Phys. Fluids **13**, 1893–1902 and 2881–2890 (1970)
2. Goldstein, J.: Semigroups of Operators and Abstract Cauchy Problems. New Orleans: Tulane University Department of Mathematics 1970
3. Grünbaum, F. Alberto: Propagation of chaos for the Boltzman Equation. Arch. Rational Mech. Anal. **42**, 323–345 (1971)
4. Kac, M.: Foundations of Kinetic Theory. In: Proc. 3rd Berkeley Sympos. Math. Statist. Probab., III, 171–197, Univ. of Calif. Press (1956)
5. Kac, M.: Some Probabilistic Aspects of the Boltzman Equation. In: The Boltzman Equation, Theory and Applications. Acta Phys. Austraila Suppl. X, ed. Cohen and Thirring, eds. (1973)
6. Kato, T.: Perturbation Theory for Linear Operators. Berlin, Heidelberg, New York: Springer 1966
7. Trotter, H.: Approximation of Semigroups of Operators. Pacific J. Math. **8**, 887–919 (1958)

Received February 8, 1975; in revised form September 25, 1975