Astérisque

KAI-LAI CHUNG

Reminiscences of some of Paul Levy's ideas in Brownian motion and in Markov Chains

Astérisque, tome 157-158 (1988), p. 29-36
<http://www.numdam.org/item?id=AST_1988__157-158__29_0>

Société Mathématique de France
Astérisque 157-158 (1988)

Reminiscences of some of Paul Lévy's ideas in Brownian Motion and in Markov Chains

by Kai-Lai CHUNG

We begin with a resume. Let $\{P(t),\ t \geq 0\}$ be a semigroup of stochastic matrices with elements $p_{ij}(t), (i,j) \in I \times I$, where I is a countable set, satisfying the condition

$$\lim_{t \downarrow 0} p_{ii}(t) = 1. \tag{1}$$

It is known that $p'_{ij}(0) = q_{ij}$ exists and

$$0 \leq q_i = -q_{ii} \leq +\infty, \quad 0 \leq q_{ij} < \infty,\ i \neq j; \tag{2}$$

$$\sum_{j \neq i} q_{ij} \leq q_i. \tag{3}$$

The state i is called *stable* if $q_i < +\infty$, and *instantaneous* if $q_i = +\infty$ (Lévy's terminology). The matrix $Q = (q_{ij})$ is called *conservative* when equality holds in (3) for all i.

If the convergence in (1) holds uniformly with respect to all i, or equivalently if the set of all q_i is bounded, then we have

$$P(t) = e^{Qt} \tag{4}$$

Let $\{X(t),\ t \geq 0\}$ be a Markov chain with $P(t)$ as its transition matrix, separable and measurable. Then in the special case just mentioned, almost all sample functions are step functions in any finite time interval. The Poisson process is an example, as well as the case when I is a finite set.

Before Lévy, the regularity properties of the sample functions of a general Markov chain have been investigated by Doob by martingale methods (1942, 1945). To describe the *allure* of a typical path, let us start it at a stable state i. The Markov property implies that it will remain at i during a sojourn time ρ_1 with $P(\rho_1 > t) = e^{-q_i t}$. Unless $q_i = 0$, ρ_1 is finite but $X(\rho_1+)$ need not exist if inequality holds in (3), or if there is some instantaneous j; in fact the path may encounter an infinity of states immediately after ρ_1 and so the analysis is halted. To avoid such a quick termination let us assume that all states are stable and the Q-matrix is conservative, also that all $q_i > 0$ to exclude a trivial case. Then at the time ρ_1 the path will jump from i to j with probability q_{ij}/q_i for all $j \neq i$, and we can resume the analysis starting with j. The path will remain in j during another sojourn time ρ_2 with $P(\rho_2 > t) = e^{-q_j t}$, then jump again, and so on. The analysis proceeds by induction until the time

$$\rho_1 + \rho_2 + \ldots + \rho_n + \ldots = \tau.$$

If $\tau = +\infty$ then the entire path has been traced, as in the Poisson case. The discovery of the possibility that τ may be finite with positive or even full probability caused a sensation, and much confusion. Read the Prologue of my Strasbourg Lectures for some historical perspective. The tracing of the path has been stopped in its track, what happens after τ? Confusion arose because a wrong question: "what can we do after τ?" was asked. The proper question is of course "what will the path do after τ?". The path exists, and there is nothing we can do except to find it! Looking back, we now realize that the problem lies in the insufficiency of the initial date given by the Q-matrix, and further structure of the paths must be searched out. This leads to a boundary theory which yields new clues to the paths but cannot deal with the general situation. I do not believe there is any complete solution, and certainly bulldozing the countable state space into something unrecognizable is no solution at all.

Lévy forsook the old way of tracking the path and instead plunged in midstream, as it were. If i is stable it is intuitively clear that the set $\{t : X(t) = i\}$ is a collection of disjoint (maximal) intervals. He made the crucial observation that the number of these i-intervals is finite up to any finite time (almost surely). This requires a proof; a short one is given in my Strasbourg Lectures. Once this is established, it follows that the i-intervals may be ordered in sequence and that their lengths form independent and identically distributed random variables. Moreover, these lengths are also independent of "everything outside the intervals". Hence if all states are stable, this global picture gives a bird's-eye view of the paths, with an abundance of mutual independence among various portions thereof. He applied this idea to the following theorem, one of the finest in the theory. (In his 1951 paper he attributed its origin to this result, which I discussed with him at the Berkeley Symposium a year before. Thus some conferences yield fruits.)

Theorem. *For any i and j, either $p_{ij}(t) = 0$ for all $t \geq 0$, or $p_{ij}(t) > 0$ for all $t > 0$.*

The case $i = j$ is easy. Now assume all states stable and $i \neq j$. Suppose $p_{ij}(t_0) > 0$, then $P_i\{T_j \leq t_0\} > 0$, where P_i is the probability starting from i, and T_j is the hitting probability of j. Let $I_n(k)$ denote the nth k-interval, and $|I_n(k)|$ its length. Then the global description above implies that

$$T_j = \Sigma |I_n(k)|$$

where the sum is taken over all $I_n(k)$ contained in $[0, T_j)$. By hypothesis $T_j < \infty$ on a set of positive probability. Hence by Egorov's theorem, for any $\epsilon > 0$ there exists finite integers K and N such that

$$T_j = S + R \leq S + \frac{\epsilon}{2},$$

where S is the sum $\sum$ restricted to $k \leq K$ and $n \leq N$, and R is the rest. There are only a finite number of permutations of the $I_n(k)$'s after the restriction, hence there is a subset Λ of the previous set with $P(\Lambda) > 0$ on which

$$T_j = \sum_{v=1}^{m} |I_{n_v}(k_v)| + R, \text{ and } R \leq \frac{\epsilon}{2}, \tag{5}$$

where $\{n_{v_1} k_{v_1} 1 \leq v \leq m\}$, $k_v \leq K, n_v \leq N$, is a fixed sequence not depending on the sample function. The set Λ is defined by a specific ordering of the intervals, hence it is independent of their lengths. It follows that

$$P(\Lambda; \tau \le \epsilon) \ge P\left(\Lambda; R \le \frac{\epsilon}{2}\right) P\left(\sum_{v=1}^{m} |I_{n_v}(k_v)| \le \frac{\epsilon}{2}\right) > 0$$

because each $|I_n(k)|$ is exponentially distributed. Q. E. D.

This proof does not seem to extend to the case when there are instantaneous states. D. G. Austin first proved the general case by a brilliant probabilistic argument using the right separability of the process and Lebesgue's theorem on differentiation of monotone functions. Later D. Ornstein gave another more analytic proof. All three proofs are given in my book on Markov chains. An exposition of Lévy's proof was included in R. V. Chacon's dissertation as a special assignment.

Lévy gave a tantalizing example of a Markov chain with only stable states and no jumps at all, all discontinuities being of the second kind; in particular all $q_{ij} = 0$ for $i \neq j$. Take a strictly increasing function on $[0, \infty)$ with jumps at all the rationals, the size of the jump J_r at r being randomized with $P(J_r > t) = e^{-q_r t}$, and all J_r's are independent. The set of q_r's are chosen as follows:

$$\forall r : \; q_{r+1} = q_r; \; \sum_{0 \le r < 1} (1/q_r) < \infty.$$

This will ensure that almost all the functions are finite and increases to infinity. The right continuous inverse of each such function is a continuous singular monotone function. There is a Markov chain whose sample functions are the collection of these inverses. Thus each path goes through all the positive reals in their natural order, sojourning in each rational but passing through all the irrationals in (Lebesgue) null time. This example is a veritable Columbus's egg stood on its flattened head. It is possible to write down explicit formulas for $X(t)$ as well as $P(t)$, as I did in my book, but it is a tedious and not very enlightening task. Lévy gave a number of such examples of Markov chains by prescribing the sample functions. Often they seem intuitively clear but require painstaking verification *après coup*. This provokes a curious, and I think important question: are there more effective ways of recognizing a Markov process without going through the usual formalities?

A week ago I received the fourth instalment of Dellacherie-Meyer's tomes, which gives an account of Lévy's ideas on local time and excursions of a Brownian motion on the line, together with later developments. Time being short I may therefore confine myself to a few remarks to fill some gaps, perhaps.

(I) Lévy derived a multitude of formulas for the excursions by means of the equivalent process:

$$Y_t = M_t - X_t,$$

where X is the Brownian motion, and $M_t = \max_{0 \leq s \leq t} X_s$. He showed that Y and $|X|$ are equivalent processes, so their zero-sets are also equivalent. Most of his calculations rely on the vertical variation of the space variable, the values of X, M and Y. I found it easier to do the calculations by using the horizontal variation of the time variable t, using only X itself. Having recovered several of his key formulas this way, I looked for something new to do; so at the suggestion of my colleague D. Iglehart, computed the exact distribution of the maximum of the (positive) excursion straddling t, conditioned on the location and duration. This turns out to be expressible by a theta function and its derivative. I could not verify its monotonicity and asked Iglehart to plot it on a computer. Due to faulty transmission by telephone, a slight error in the plotted formula led to a curve decidedly not monotone. After the error was corrected the monotonicity was, of course, confirmed. Later I learned that the distribution had been found by N. H. Kuiper in a statistical test of random points on a circle. He became director of IHES, you know.

(II) The excursions of Brownian motion bear remarkable resemblance to those of a Markov chain with a single sticky recurrent boundary (see my Strasbourg Lectures). By identifying the explicit formulas for excursions with the general ones in the chain theory, hidden meanings of certain quantities are revealed. This is because in the final analytic expressions factorization and cancellation have taken place without our notice. Here is an example. One of the deeper formulas for Markov chains is the last exit decomposition:

$$p_{ij}(t) = \int_0^t p_{ii}(s)\, g_{ij}(t-s)ds.$$

The analogue for Brownian motion is (for $y > 0$)

$$p(t;\ 0,y) = \int_0^t p(s;\ 0,0)g(t-s,0,y)ds \tag{6}$$

which is identical to the first entrance formula owing to symmetry. We have

$$p(s;\ 0,0) = \frac{1}{\sqrt{2\pi s}}, \quad g(t;\ 0,y) = \frac{y}{\sqrt{2\pi t^3}}e^{-\frac{y^2}{2t}}.$$

Putting these quantities together we obtain for the right member of (6):

$$\begin{aligned}&\int_0^t \frac{1}{2\pi\sqrt{s(t-s)}}\frac{y}{t-s}e^{-y^2/(t-s)}ds \\ &= \int_0^t P(\gamma(t) \in ds)P\left(|X(t)| \in dy \;\middle|\; \gamma(t) = s\right) ds\end{aligned} \tag{7}$$

where $\gamma(t)$ is the last zero of X before t. Both probabilities in the last-written integral are derived by Lévy by his methods. The first is the arcsin law made famous by Feller's propaganda (mentioned by Lévy in his *Notice sur les travaux*). The second apparently was not understood by other authors until my 1976 paper which resulted from my attempt to unravel it. It is Theorem 42.5 in Lévy's book and it plays a key role in his *étude profonde.*[1]

(III) However pretty those excursions may be, Lévy's grand scheme is to string them all together on a new time scale, the local time at zero, and recover the Brownian motion as a Poisson point process run by the local clock. An illustration of this idea is his proof of the following theorem, the *pièce de resistance* of his conception of "*mesure du voisinage*", later known as local time.

Theorem. *We have almost surely*

$$\lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} m\left(\{s \leq t : \ |X(t)| < \epsilon\}\right) = L(t)$$

where m is the Lebesgue measure and $L(t)$ is the local time at t.

Recall that Lévy defined $L(.)$ by an inversion of a strictly increasing purely jumping stable process of index $1/2$. It is a profound analogue of his construction of the singular Markov chain discussed earlier.

Lévy's proof, given at the end of his great paper *Sur certains processus stochastiques homogènes, Compostio math.* **7**, 1939[2], and not reproduced in his book (1948/1965), runs as follows. The total occupation time of $(0, \epsilon)$ by $|X|$ up to time t is the sum of the same occupation time $u_\epsilon(\varphi)$ during all the excursions φ up to local time $L(t)$:

$$\sum_{\varphi} u_\epsilon(\varphi) = \int_{[0, L(t)] \times R_+} u_\epsilon dN \tag{8}$$

Here N is the Poisson point process of the excursions, whose mean measure is the Lévy measure of the inverse local time, namely the stable process with exponent $1/2$, given explicitly by Lévy in his work on "Lévy Processes". The expectation of the right member in (8) can be computed since we can compute the occupation time during an excursion, and there is independance between disjoint excursion intervals. (It has a neat density.) We need also an estimate for the second moment, which is supplied in my paper (dedicated to Lévy). Now under very general conditions on the first two moments of a Poisson sum like that in (8), the value of the random sum is asymptotically equivalent to its expectation in the limit, here as $\epsilon \downarrow 0$. Thus the theorem is proved by a straightforward computation of the expectation (a nice integral), and an adequate bound for the second moment, exactly as in the grand tradition of classical probability.

The same method gives quick proofs of a number of similar but easier results: the downcrossing result and Kingman's Cesàro mean result, etc. (Notes by A. A. Balkema on this topic exist.) Despite later alternative approaches to these matters, Lévy's original way should be preserved, not as a museum piece, but as a monument conjuring up the past and beckoning to the future.

Footnotes

(1) Lévy's grand tradition of deriving a wealth of explicit formulas has been continued in the recent work by Biane and Yor.

(2) It is regrettable that this paper was not cited in the volume by Dellacherie-Meyer mentioned above, but I was pleased to see it publicized in the exhibition at *École polytechnique* during the conference. When I first met Lévy in 1950, I asked him for a reprint of this paper and was told that all his papers were burned by the Nazis. His treatise *Theorie de l'addition des variables aléatoires* was not accessible to me during the war because library collections were stored in underground caves in China to escape from Japanese bombing.