The Annals of Probability
1973, Vol. 1, No. 1, 1–18

PAUL LÉVY, 1886–1971

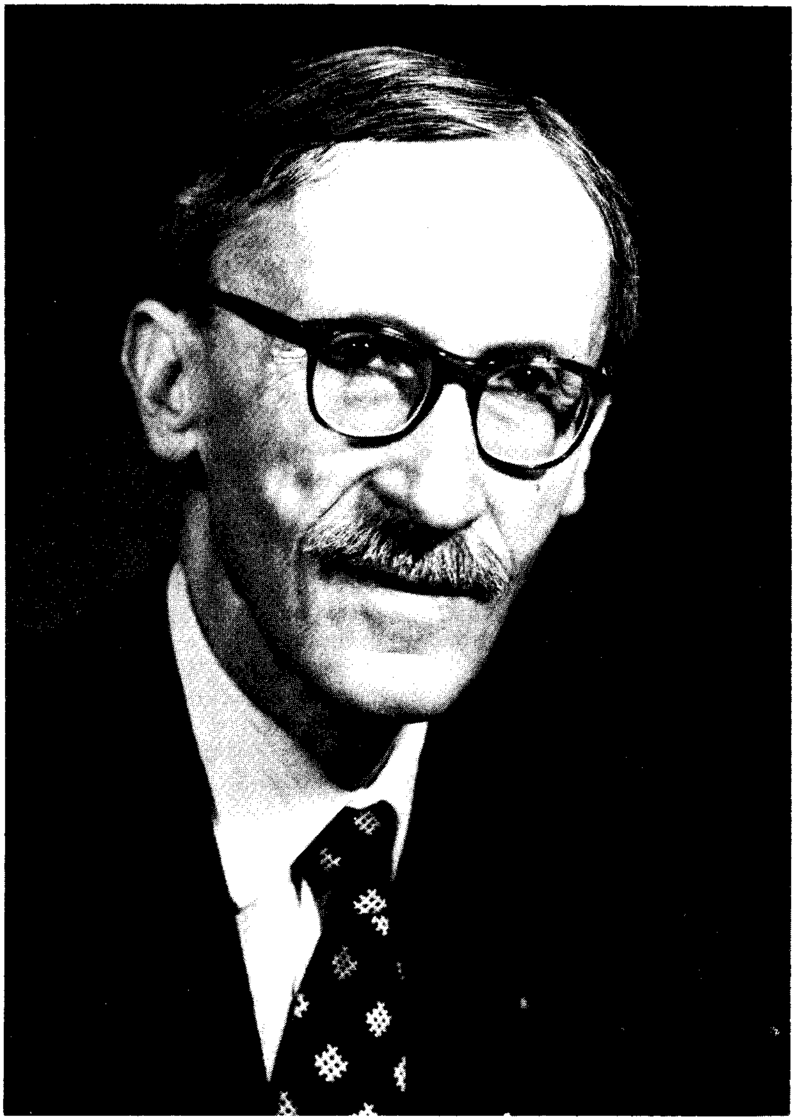
In 1919, at the age of 33, Paul Lévy was asked to give three lectures on Calculus of Probabilities at the Ecole Polytechnique. This began the heroic period in Probability throughout which Paul Lévy was at the center of the stage: During the following twenty years, "Calculus of Probabilities"—which consisted mainly of a collection of small computational problems—became "Probability theory," a full-fledged branch of mathematics, partaking of and contributing to the torrent of twentieth century mathematics, with concepts, problems and results constantly born from its own intuitive background.

In 1919, Paul Lévy was already a renowned mathematician, with over twenty publications between 1905 and 1914, a period interrupted until 1918 by the war, which he spent in the French artillery. Born in Paris in 1886, son and grandson of mathematicians, he received the usual—in France—scholastic honors for the exceptionally gifted: Prix du Concours Général in greek and also in mathematics, Prix d'Excellence at Lycée Saint Louis in mathematics and also in physics and chemistry, first at the Concours d'entrée of Ecole Normale Supérieure and second at the Concours d'entrée of Ecole Polytechnique; he entered the latter, published in 1905 his first paper—on semiconvergent series—and finished in first place. He spent a year doing his military service and three years at Ecole des Mines. During these three years, 1907 to 1910, he followed courses at the Sorbonne by Darboux and by Picard and at the College de France by Humbert and by Hadamard. In 1910, influenced by Hadamard, he began research on Functional Analysis of Volterra and on Green functions, his main mathematical preoccupation between 1910 and 1914, and from 1919 to 1922. He obtained his Docteur ès sciences degree in 1912 and his thesis became the core of his 1922 book *Leçons d'Analyse fonctionnelle*, which in turn formed the core of his book *Problèmes concrets d'Analyse fonctionnelle* published in 1951.

In 1913 he was appointed Professor at Ecole Nationale des Mines and from 1920 to 1959 he was Professor of Analysis at Ecole Polytechnique. In 1964 he was finally elected to the Académie des Sciences. Beginning in 1905 and continuing almost to the time of his death on December 15, 1971, he published 10 books and over 270 papers, of which over 150 are in Probability theory. Here we shall limit ourselves to some of the highlights of his probabilistic thought.

Paul Lévy was a painter of the probabilistic world. Like the very great painting geniuses, his palette was his own and his paintings transmuted forever our vision of reality. Only a few of his paintings will be described here—some of those which are imprinted indelibly on the vision of every probabilist. His three main, somewhat overlapping, periods were: the limit laws period, the great period of additive processes and of martingales painted in pathtimes colors, and the Brownian pathfinder period.

The three lectures of 1919 were requested to be on "notions of Calculus of Probabilities and the role of Gaussian law in the theory of errors." The books consulted by Paul Lévy were those by Bertrand, by Borel and, especially, by Poincaré; the results of the Russian school—by Tchebichev, Liapounov and Markov, were not even mentioned therein. A glance at Poincaré's book shows how its critical reading



By action of the Council of the Institute of Mathematical Statistics, this first issue of the Annals of Probability is dedicated to the memory of

PAUL LÉVY

by Paul Lévy—a mature and thorough mathematician who, contrary to Poincaré, considered that probability deserved rigorous mathematical handling—led him to the rediscovery of the Russian school results, the general concept of probability law, the method of characteristic functions, and impelled him into his fifteen year quest for "justification" of the Gaussian law. Paul Lévy's ideas and results obtained in 1919 and further developed from 1922 to 1925 were collated in his first book, *Calcul des Probabilités* in 1925.

Only discrete and absolutely continuous laws of real random variables X appear in Poincaré. At once Paul Lévy realizes, independently of von Mises (1919), that they are but particular cases of the general concept of probability laws as defined by distribution functions on the real line R, equivalently, by distributions, i.e. probability measures on R; also he introduces the concept of types of laws.

Poincaré considers real characteristic functions Ee^{uX}, which do not always exist for all $u \in R$. Paul Lévy replaces them by present day characteristic functions Ee^{iuX}, which exist for all $u \in R$, thus creating the apparatus of Fourier–Stieltjes transforms which play nowadays such a central role in harmonic analysis. Unknown to him, some characteristic functions were considered by Cauchy and were used by Liapounov in his proof of "Liapounov's theorem." But Paul Lévy explores them in depth, to the point that, since then, only improvements of detail have been obtained. He establishes by now classical inequalities and creates the fundamental apparatus: correspondence between distributions and characteristic functions transforms convolutions into multiplications and conversely, and is biunivoque and bicontinuous. More precisely, Paul Lévy finds the explicit inversion formula and the so called "continuity theorem" for characteristic functions, except that he assumes convergence on R of sequences of characteristic functions to be uniform in some neighborhood of the origin; in 1933 Bochner showed that it suffices to require continuity at the origin of the limit function.

Thus sprung, fully fledged, from the forehead of Paul Lévy the familiar technique for the present day theory of limit laws within the far reaching extension by Kolmogorov of Poisson's limiting procedure: limit laws are to be those of sequences of sums $\sum_k X_{nk}$, $k = 1, 2, \cdots, k_n \to \infty$ as $n \to \infty$, of independent and uniformly asymptotically negligible summands, i.e., for every $\epsilon > 0$, $\max_k P[|X_{nh}| > \epsilon] \to 0$ as $n \to \infty$.

As for the "justification" of the Gaussian law in the theory of errors, Poincaré, and following him Paul Lévy, starts from the intuitive idea that the Gaussian law is approximately the law of sums of very large numbers of approximately equally very small independent summands. Poincaré's mathematical translation of this idea was encumbered by many restrictions and treated purely formally—with no consideration for rigor. Yet, gradually Poincaré's idea leads Paul Lévy to the final solution. First, using characteristic functions he rediscovers Liapounov's and Lindeberg's results. But it is only in 1934 that he finds the answer: the limit of the laws $\mathfrak{L}(\sum_k X_{nk})$ is Gaussian if and only if the laws $\mathfrak{L}(\max_k X_{nk})$ converge to the law degenerate at zero. Independently, Feller obtains the necessary and sufficient conditions for convergence to the Gaussian law by exploiting to its full extent Paul Lévy's apparatus of characteristic functions. Paul Lévy on the other hand makes little use of them. For gradually he moved to a more direct way of translating his intuition. He introduced the "Paul Lévy distance" of probability laws, later extended by Prohorov

(1953) to the distance of laws in separable metric spaces, then his dispersion function with basic dispersion inequalities. He reached his quest's goal by decomposing the summands into two parts, uniformly small ones and uniformly bounded away from zero ones, and by means of dispersion inequalities proved that the two sets could be thought of as asymptotically independent. Using these Paul Lévy methods some twenty years later, Kolmogorov obtained beautiful approximation results, since then sharpened by many probabilists.

At one of the 1919 lectures, a listener told Paul Lévy that the Gaussian law was the only stable one. The same day Paul Lévy disproved this statement by rediscovering the symmetric stable laws already known to Cauchy and soon thereafter described in terms of their characteristic functions all stable laws including the nonsymmetric ones as well as the quasi-stable laws. This led him to the introduction and investigation of the nowadays familiar concepts of domains and partial domains of attraction, pursued later by Khintchine (1936).

Paul Lévy's study of stable laws led him to a completely novel approach to the problem of limit laws which for two centuries was the central and, in fact, the only theoretical problem of Calculus of Probabilities. From Bernoulli (1713) and de Moivre (1732) to Liapounov (1900) and Lindeberg (1920), it consisted of the search of conditions for the limit laws of sequences of normed sums of independent random variables with finite second moments to be Gaussian or degenerate. Paul Lévy in his 1925 book transforms the old central limit problem into the search all possible limit laws for sequences of suitably normed sums of independent (and identically distributed) random variables, not necessarily having finite second moments, and then for necessary and sufficient conditions for convergence to any given stable law Poisson's limiting procedure and his limit law stood isolated and ignored until 1934. Then, in one of the most important probability papers ever published, "Sur les intégrales dont les éléments sont des variables aléatoires indépendantes" (1934) *Ann. Scuola Norm. Sup. Pisa*, Paul Lévy, by looking at sample paths of additive processes, discovers and describes explicitly in terms of their characteristic functions infinitely divisible laws—whose building blocks are Poisson type laws. De Finetti already had the idea of such laws and in 1932 Kolmogorov described those with finite seconds moments. Paul Lévy's representation formula for general infinitely divisible laws led soon to the definitive formulation and solution of the present day central limit problem: Paul Lévy indicated and Khintchine (1936) proved that the family of all limit laws in the Kolmogorov formulation is exactly that of all infinitely divisible ones; the same year Paul Lévy characterized the subfamily which consists of all limit laws in the particular case of normed sequences. And in 1938 necessary and sufficient conditions for convergence to any given infinitely divisible law were obtained, independently by Gnedenko, using to its full extent the Paul Lévy method of characteristic functions, and by Doeblin, using to its full extent the Paul Lévy method for the Gaussian limit law.

Already in his paper on the Gaussian limit law, Paul Lévy considered the influence of terms which are not very small and became convinced that a Gaussian law could be decomposed into a finite number of Gaussian laws only. He was so convinced of it that part of the paper is devoted to various consequences of this conjecture. It was beautifully proved by Cramér (1936) using characteristic functions and the fact that an entire one of order two must be Gaussian. It is interesting to note that Paul Lévy remained quite unhappy with this purely technical proof and hoped for one which would correspond to his intuition.

Following the Lévy–Cramér result, Khintchine studied decompositions of infinitely divisible laws, and Raikov proved that the Poisson law could be decomposed only into Poisson ones. Gradually, Paul Lévy and Khintchine, pursuing this line of thought, created a new branch of Probability theory—the arithmetic of probability laws. And recently Linnik has obtained new and deep results in this difficult subject.

In 1925 a participant in Hadamard's seminar noticed that in one of the proofs in Paul Lévy's book on Functional Analysis, completeness of L_2-spaces was assumed, and tried without much success to prove this property using convergence in measure. Paul Lévy intervened, and with the help of Banach who was also present, gave a proof. It was published and, naturally, it was found to be a proof of the Riesz-Fisher theorem (1907). Yet, this seminar led Paul Lévy to examine and compare various types of convergence based upon the concept of probability measure. Meanwhile he decided "to consider other problems than those relative to the characteristic function or the Gaussian law." First he decided to do for continued fractions of numbers chosen at random between 0 and 1, briefly considered in Borel's 1909 paper, what Borel did therein for their decimal expansions. He rediscovered and proved (1929) a conjecture of Gauss pròved independently by Kuzmin (1928), and sharpened the conjectured convergence of laws to almost sure convergence. Paul Lévy returned over and over again to the continued fractions subject pursued also by Khintchine, Doeblin and Ryll-Nardzewski. But it is in 1930 that he begins his second extraordinary productive and fruitful period. He relies more and more on direct mathematical formulation of his intuitive insights, and introduces and investigates the concentration function and its inverse function—the dispersion function, the symmetrization method and use of medians with corresponding inequalities. He shows that for independent random variables X_n, with probability one, either the series $\sum (X_n - a_n)$ converges for some series of constants a_n or is essentially divergent (i.e. diverges whatever be these constants), proves that for $\sum X_n$ almost sure convergence and convergence in probability are equivalent, and in 1939 states also equivalence with convergence in law. He rediscovers and sharpens Khintchine's law of the iterated logarithm and proves that for sequences of successive sums $S_n = X_1 + \cdots + X_n$ of essentially divergent series, $P[S_n > s_n$ infinitely often$]$ is either 0 or 1 so that there are only two classes of sequences (s_n) for (S_n). This gives its full meaning to the search for criteria for this dichotomy pursued further by Cantelli, Kolmogorov and Feller.

In 1934–35, with about 20 publications, Paul Lévy's research reached the height of its originality and the power and fruitfulness of ideas and results. While finding at last the "justification" of the Gaussian law which started his probabilistic research in 1919, he deeply transforms the whole of probabilistic thinking. General additive processes are introduced and in a few weeks so thoroughly analyzed that only improvements of detail have since been achieved; as a consequence general infinitely divisible laws are explicitly described. Martingales are created and analyzed. Random times dominate the search and are utilized throughout. Random analysis with its basic concepts and problems is born. These ideas and results, collated in the Paul Lévy 1937 monograph "Théorie de l'addition des variables aléatores," marks the beginning of a new era of Probability theory. In 1967, the thirtieth anniversary of its publication was celebrated all over the probability world; more than 150 probabilists sent their congratulations to Paul Lévy.

Paul Lévy writes that "at the beginning of 1934 I suddenly noticed that any stable law leads, as does the Gaussian, to a random function that we can obtain, like that of Wiener, by an interpolation method. I then decided to define the general form of a function $X(t)$ with independent increments, in other words of an additive process .... It suffices to consider the case with $X(0) = 0$." He shows that one can delete from this process its degenerate discontinuities by including them within a "centering" numerical function so as to leave an additive process whose only almost sure discontinuities are nondegenerate jumps at fixed times, independent of the choice of the centering function. In turn, this set of fixed discontinuities, which is countable, can be deleted. At this point Paul Lévy introduces unconditionally convergent series of independent random variables and shows that a series $\sum X_n$ which is not essentially divergent can be made unconditionally convergent by selecting suitably the constants a_n to be subtracted from the X_n. Thus he was left with an additive process $X_c(t)$ with almost surely no fixed discontinuities. It may have moving discontinuities—at random times; more precisely sample paths $X(\cdot, \omega)$ may have jumps at a countable time set varying with ω. The set of all those moving jumps is that of, in general, a "compensated sum of infinitesimal increments" of Poisson type additive processes. There remains finally a sample continuous additive process. And Paul Lévy proves it to be Brownian, thus obtaining the converse of the celebrated Wiener's result. The complete Paul Lévy analysis was formalized by Itô as follows:

$$X_c(t) = \eta(t) + \int_{R-\{0\}} \{x\nu_t(dx) - \frac{x}{1+x^2} L_t(dx)\}$$

where $\eta(\cdot)$ is a Brownian process and the $\nu_{st}[x, y) = \nu_t[x, y) - \nu_s[x, y)$, the numbers of jumps in the time interval $[s, t)$ of heights in $[x, y)$ $xy > 0$, are Poisson random variables with parameters $L_{st}[x, y) = L_t[x, y) - L_s[x, y)$ independent for disjoint time intervals $[s, t)$ and independent for disjoint height intervals $[x, y)$. Only Paul Lévy's admirable probabilistic intuition and sophisticated mathematical technique could in one sweep completely overcome the many delicate problems arising in and leading to the foregoing decomposition and analysis of sample paths of additive processes. The 1934 Paul Lévy representation formula for characteristic functions of general infinitely divisible laws, i.e. laws of the $X(t)$, followed at once from this analysis; soon thereafter, Khintchine gave a direct and purely analytic proof of this formula.

Beginning in 1934, Paul Lévy creates and investigates the martingale concept (which owes its name to Ville). Like the Markov dependence concept, the martingale one was born from, and its investigation was guided by, the results obtained in the case of independent summands. More precisely, they both were born from an attempt to preserve the law of large numbers.

Paul Lévy noticed that the Bienaimé equality and the Tchebichev method of proving the law of large numbers remain valid when, instead of centering the summands at their expectations (in the independence case), they are centered at their conditional expectations given the predecessors. Thus he is led to work with summands X_n such that, for $n = 1, 2, \cdots$, almost surely, $E(X_{n+1} \mid X_1, \cdots, X_n) = 0$, or equivalently $E(S_{n+1} \mid S_1, \cdots, S_n) = S_n(S_n = X_1 + \cdots + X_n)$, i.e. with martingales $(S_1, S_2, \cdots)$. He observes that Kolmogorov's inequality remains valid for martingales. Exactly as in the case of independence, various almost sure convergence and stability theorems follow.

But, as usual, Paul Lévy goes much farther. The first martingale convergence theorem is the celebrated Paul Lévy 0–1 law. It is perhaps one of the most beautiful results of probability theory; its proof is quite sophisticated while it sounds intuitively obvious when stated in gambling terms: If A is an event defined on the sequence $(X_1, X_2, \cdots)$ of arbitrary random variables then $P(A \mid X_1, X_2, \cdots, X_n)$ converges almost surely to 1 on A and to 0 on A^c. In intuitive terms it becomes: A gambler plays a series of related or unrelated games of any kind with numerical outcomes. If, say, he will be ruined eventually then, except for a miracle, his chances of ruin evaluated in terms of already known outcomes approach 100% and the chances of not being ruined approach 0%.

Furthermore, Paul Lévy uses systematically random times. He measures "time" τ (random) by the sum of the variances conditioned upon the predecessors, (i.e. every game is given a duration $\sigma^2(X_{n+1} \mid X_1, \cdots, X_n)$), introduces the corresponding sums $S(\tau)$ and under various conditions obtains a large number of their properties such as Gaussian convergence for $S(\tau)/\tau^{\frac{1}{2}}$ and the law of the iterated logarithm in terms of τ.

The importance of the martingale concept cannot be overemphasized. In Doob's hands (1940) martingales became a powerful tool within Probability theory as well as within Analysis in general. Martingales, Markov dependence and stationarity are the only three dependence concepts so far isolated which are sufficiently general and sufficiently amenable to investigation yet with a great number of deep properties.

In 1939 begins the third stage of Paul Lévy's probabilistic revolution. Once more he returns to his first love, the Gaussian law, with his intuition deepened and his technique sharpened while living during random times along the paths of martingales and of additive processes. A new branch of Probability theory is born: fine structure of Brownian paths.

Pioneering work in Brownian notion was done at the very beginning of the 20th century by Bachelier, Einstein and Smolukovski; it was purely formal. The first rigorous approach is due to Wiener (1923). He defined Brownian motion constructively and proved that almost all its paths were continuous. Later (1933), jointly with Paley and Zygmund, he discovered that almost all Brownian paths were nowhere differentiable. As always, Paul Lévy went much further. In fact, his concepts and results were so fruitful that they continue to dominate and inspire works in Brownian motion and more generally those in Diffusion processes. It suffices to glance, say, through the Itô and McKean book *Diffusion Processes and their Sample Paths.*

Let $X(t)$ be the Brownian motion function with $X(0) = 0$, omitting the negligible set of noncontinuous paths.

First, Paul Lévy considerably sharpens Wiener's results. He gives a new constructive definition, proves its projective invariance, finds a continuity modulus: $|X(t + h) - X(t)| < c\,[2|h|\log 1/|h|]^{\frac{1}{2}}$ for $|h| \leqq \tau$ (random) when $c > 1$ but not when $c < 1$, deduces its law of the iterated logarithm, shows that the arc length of $(t, X(t))$ is infinite in every nondegenerate time interval and, in fact, proves that if $\{t_1, t_2, \cdots\}$ is a countable set dense in $[0, t)$ then, letting S_n be the sum of squares of increments of $X(t)$ over the n first t's (reordered increasingly), S_n converges to t almost surely, so that almost no sample path is of bounded variation.

Next, Paul Lévy looks at the function defined by $M(t) = \max_{s \leq t} X(s), m(t) = \inf_{s \leq t} X(s)$, shows that the functions $|X(\cdot)|$, $M(\cdot) - X(\cdot)$, $X(\cdot) - m(\cdot)$, have the same probability distribution, and rediscovers Bachelier's formal results about the distributions of $M(t)$, $(M(t), X(t))$ and $(M(t), X(t), m(t))$. Then he finds that $|X(\cdot)|$ and $M(\cdot) - X(\cdot)$ have relative extrema everywhere dense.

He introduces "passage times" in the shape of the inverse function $\tau(\cdot)$ of $M(\cdot)$ and shows that it is a stable process of coefficient $\frac{1}{2}$—sums of positive jumps with Poissonian number $\nu(h)$ of parameter $x(2/\pi h)^{\frac{1}{2}}$ of jumps of height larger than h in each interval of length x. This relates directly to his deepest and most influential results for the set of absolute minima of $|X(\cdot)|$ and $M(\cdot) - X(\cdot)$, which is the (random) set Z of zeros of $X(\cdot)$: Z is closed, uncountable, with no isolated points, of Lebesgue measure 0, and (due later to Taylor) of Hausdorff–Besicovitch dimension $\frac{1}{2}$, and its largest member in $[0, t]$ obeys the celebrated Paul Lévy "Arcsin law". Furthermore, Paul Lévy constructs a random time scale, the local time $\tau(t)$ which describes times spent on Brownian paths at 0; in fact his "mesure du voisinage" is $(2/\pi)^{\frac{1}{2}}\tau(t)$ and is at the root of various "time changes" so important nowadays in traveling along Brownian, Diffusion, and Markov paths.

Starting with his 1940 paper on "Plane Brownian Motion," Paul Lévy introduced and investigated Brownian motions whose values or whose parameter lie in n-dimensional Euclidean spaces and in infinitely-dimensional Hilbert space E_ω. From 1948 on he was also concerned with "Laplacian" (or Gaussian) processes—essentially integrals with respect to Brownian motion. The wealth of his results, as first summarized in his 1948 book *Processus stochastiques et mouvement brownien*, then completed in its second edition in 1965, is tremendous. Let us mention one only, simply because it surprised Paul Lévy himself. If a Brownian motion defined in E_ω is known in a sphere, however small its radius, then it is determined in the whole of E_ω. He "explained" to himself this astounding determinism, which recalls that of analytic functions but exists only because of infinite dimensionality of E_ω, as due, contrary to common mathematical sense, not to the regularity but to such an "extreme irregularity that all the information about the process in the whole of E_ω is already contained in any sphere."

Throughout his research Paul Lévy met and had to examine Markov processes. But his ideas did not have the extraordinary impact of the foregoing ones. However, in connection with Brownian motion in several parameters he introduced and investigated Markov processes in several parameters. But the most influential was his 1951 paper on stationary ones with a denumerable state space. There he flushed into the open the totally unexpected and, then, monstrous possibilities of instantaneous and fictitious (or boundary) states which now are part and parcel of Markov process theory. And so are those of one of his very few direct students, Doeblin, who died in 1940 at the age of 25, with 26 publications between 1936 and 1940. Let us only mention that he was the first to proceed to the analysis of Markov paths—under a uniform continuity condition for transition probabilities which leads to step sample paths; Doob (1945), upon removing the uniformity restriction discovered sample discontinuities more complicated than jumps, and then came Lévy's 1951 paper. There is no doubt that Doeblin was to be one of the great ones. In fact, Paul Lévy writes: "One can count on the fingers of one hand those mathematicians who, since Abel and Galois, are dead so young in leaving such an important oeuvre."

The foregoing, among Paul Lévy's ideas and results were selected partly because they are (or ought to be) known to every probability student. Thus he may realize Paul Lévy's dominant role in the evolution of the old Calculus of Probabilities to the by now classifical Probability theory. Yet, even today his writings are a treasure of ideas and results still awaiting further research. Fréchet's comment, "Your results are more or less complex according to one's perspective," says it well, provided "complex" is understood also as "deep."

If one may dare to try to extract the essence of Paul Lévy's probabilistic thinking, it may be said that his leitmotif since the beginning (1919) is the Gaussian law; over and over again he starts from it and inexorably returns to it. But above all he is a traveller along sample paths (this is why for him Markov property is always strong Markov property). Statistical physics has a familiar—the "Maxwell demon" who travels along individual paths of particles subject to the deterministic laws of mechanics; his clock is the same along all paths and he encounters effects of extremely large numbers of deterministic phenomena in extremely short time intervals. In Probability theory we now also have a familiar—the "Lévy demon" who travels along individual sample paths of stochastic processes subject to "successive interventions of hazard"; his (random) clock depends upon the paths and he encounters effects of, most frequently uncountable, chance phenomena in extremely short path time intervals. In fact, Maxwell's demon is but a degenerate form of Lévy's demon.

Paul Lévy had very few direct probability students since he was not a professor at a university and did not teach Probability theory at Ecole des Mines or at Ecole Polytechnique. His influence was primarily through his writings and he was recognized as a great probabilist in Russia and in the United States long before it happened in his own country. In his writings he described leisurely the mathematical world he lived in, in clear and beautiful French.

He was a very modest man while believing fully in the power of rational thought. His 1970 book *Quelques aspects de la pensée d'un mathématicien* consists of two parts: "Autobiographie mathématique" and "L'évolution de mes idées sur la philosophie." They are a beautifully written completely candid portrait of the man and the mathematician. Myself, whenever I pass by the Luxembourg gardens, I still see us there strolling, sitting in the sun on a bench; I still hear him speaking carefully his thoughts. I have known a great man.

MICHEL LOÈVE
UNIVERSITY OF CALIFORNIA, BERKELEY