LAURENT SCHWARTZ

Quelques réflexions et souvenirs sur Paul Lévy

Astérisque, tome 157-158 (1988), p. 13-28
<http://www.numdam.org/item?id=AST_1988__157-158__13_0>

Société Mathématique de France
Astérisque 157-158 (1988)

Quelques réflexions et souvenirs sur Paul Lévy

par Laurent SCHWARTZ

Un grand-père de Paul Lévy était polytechnicien. Son père aussi, Lucien Lévy ; il a passé l'agrégation de mathématiques et fut examinateur des élèves à Polytechnique, poste équivalent à celui de professeur. C'était un bon mathématicien. Il a écrit un mémoire de géométrie, couronné par l'Académie Royale de Belgique, et plusieurs livres didactiques, sur l'arithmétique, sur les fonctions elliptiques et sur l'analyse infinitésimale. Il a certainement contribué dans une certaine mesure à la formation de Paul Lévy, mais celui-ci s'est avant tout formé tout seul. Paul Lévy a été lui-même professeur à l'Ecole Polytechnique de 1920 à 1958. Sa fille aînée, Marie-Hélène, est devenue ma femme, elle est mathématicienne. Sa deuxième fille, Denise, agrégée d'allemand, a épousé Robert Piron, polytechnicien, et son fils, Jean-Claude, a été polytechnicien. Deux de ses petits-fils sont polytechniciens. L'Ecole Polytechnique a donc joué un grand rôle dans la famille Paul Lévy[1] et, par ailleurs, Lucien Lévy et Paul Lévy ont, à eux deux ensemble, examiné et enseigné à l'Ecole Polytechnique pendant plus de soixante ans. J'ai moi-même étudié son cours de l'Ecole Polytechnique quand j'étais jeune : il était d'une remarquable précision, d'un très grand intérêt et faisait un grand appel à l'intuition. Il s'est beaucoup investi dans son enseignement, et a laissé un très grand souvenir auprès de nombreuses générations d'élèves de l'Ecole Polytechnique.

J'ai fait sa connaissance en 1934 lorsque je me suis fiancé avec sa fille. Je suis alors allé très souvent chez lui dans les années 1934, 1935 et 1936 pour deux raisons : pour voir sa fille, et faire avec lui des probabilités et de l'analyse. Il m'est toujours apparu comme un homme certes sûr de lui, mais extrêmement modeste. Par exemple, la première fois qu'il a été invité au Symposium international de probabilités de Berkeley, après la guerre, il a dit à Jerzey Neymann : "Vous savez, je n'ai pas grand chose d'original à vous exposer ici, j'ai peu de travaux nouveaux". Neymann lui répondit : "Mais vous savez, tout le symposium

[1] Néanmoins, bien évidemment, comme tout scientifique sérieux, il n'attachait aucune importance dans son jugement sur un mathématicien au fait que celui-ci soit ou non passé par telle ou telle grande Ecole ou Université !

va être essentiellement consacré à vos travaux". Il en a été absolument stupéfait. C'était un homme très tranquille, manquant totalement d'agressivité, et même de pugnacité pour se défendre dans le monde mathématique. Il avait énormément de temps libre pour travailler. L'enseignement à l'Ecole Polytechnique n'existait qu'une année sur deux et comprenait quarante conférences, il n'y avait pratiquement pas de contact avec les élèves, et il n'y avait pas de petites classes ou de travaux dirigés sauf dans les dernières années. Le nombre de réunions de conseils était très restreint. Il avait donc pratiquement tout son temps pour faire de la recherche. Il travaillait sans agitation, sentant toujours qu'il avait toute sa vie devant lui pour chercher et trouver. Il avait naturellement toutes les qualités d'énergie du bon chercheur, mais n'avait pas besoin de se surmener et s'arrêtait quand il en avait assez. Il est hors de doute que de telles conditions de travail ont beaucoup favorisé sa recherche.

Bien que connaissant parfaitement le sujet sur lequel il travaillait, il oubliait facilement les autres domaines, et devait très souvent chercher et trouver de nouveau des résultats qu'il avait connus auparavant. Par exemple, une fois, je lui ai demandé s'il connaissait une démonstration simple du théorème de dérivation de Lebesgue (la primitive d'une fonction intégrable admet celle-ci presque partout comme dérivée). Il est allé réfléchir environ une demi-heure, puis est revenu et m'a donné une démonstration fort élégante. Puis nous avons parlé d'autre chose. Je n'ai pris aucune note sur ces démonstrations car, à ce moment-là, j'avais une très grande mémoire et gardais généralement les choses présentes à l'esprit sans rien noter. Quelque six mois plus tard, il est venu me demander si je connaissais une démonstration simple du même théorème ; je lui ai ressorti la sienne. Il m'a répondu : "c'est très ingénieux, jamais je n'aurais trouvé cela moi-même", et a été très étonné quand je lui ai dit que c'est lui-même qui l'avait trouvée et me l'avait donnée. Je dois ajouter ironiquement que moi-même aujourd'hui je ne sais plus quelle était cette démonstration !

Il avait des difficultés considérables pour lire les travaux des autres. C'est souvent très difficile pour de nombreux mathématiciens, mais je l'ai rarement vu à ce point-là. Comme il avait par contre une intuition et une créativité énormes, il lui est arrivé très souvent de trouver des résultats antérieurement trouvés par d'autres, mais dont il ne connaissait pas l'existence. C'était toujours un sujet de grand regret pour lui. Il a écrit dans sa Notice, lors de son élection à l'Académie des Sciences en 1964 : "Je ne crois pas qu'il eût été en mon pouvoir de faire mieux que je n'ai fait en consacrant plus de temps à la lecture. Quelques tentatives que j'ai faites pour "forcer mon talent" n'ont pas été heureuses". Il disposait fort heureusement d'informateurs, à savoir les mathématiciens de son entourage avec qui il discutait très souvent et qui lui permettaient de se tenir au courant. Il n'y eut pas en France de probabilistes de même niveau que lui, mais les conversations avec ses collègues, même sur des sujets de probabilités, étaient d'un grand profit pour eux et pour lui ; il a en outre correspondu avec des probabilistes du monde entier, par exemple avec Cramer, Lindeberg et surtout avec des mathématiciens soviétiques, Khintchine et Kolmogorov. L'arrivée d'une lettre de Khintchine ou de Kolmogorov était toujours pour lui un grand sujet de réjouissance. Cette correspondance a dû être interrompue peu avant la guerre, il a reçu en 1937 une lettre de Khintchine assez extraordinaire et a eu connaissance d'un article de lui parlant de la nécessité de lutter contre les fausses idéologies en probabilités, c'était l'époque de la terreur stalinienne, et il a compris qu'il était meilleur de ne plus écrire aux mathématiciens soviétiques. Après la guerre, il a naturellement été aussi en correspondance très active avec les mathématiciens américains et japonais. Il a, en particulier, été un grand nombre de fois aux Etats-Unis, et notamment à Berkeley. Il n'a pratiquement pas eu d'élèves en France : cela tient en grande partie à ce qu'il était professeur à l'Ecole Polytechnique et non à la Sorbonne, et que les polytechniciens, pendant des décennies, ne se sont jamais dirigés vers la recherche, il n'a donc pas pu trouver d'élèves de recherche parmi les jeunes qui suivaient ses cours. Les étudiants d'université et les normaliens lui échappaient presque complètement. Il faut ajouter -et c'est là un peu ce que j'appelais tout à l'heure son manque de pugnacité- qu'il n'a jamais fait beaucoup d'efforts pour aller les trouver. Il n'a pas non plus essayé de tenir, où que ce soit, un séminaire pour attirer des jeunes. Il a accepté l'isolement tel qu'il se présentait. Il a toutefois eu deux élèves, Jean Bass et Daniel Dugué, devenus probabilistes ; un autre, qui lui est resté très fidèle, Michel Loève, de nationalité égyptienne, n'a pas pu se faire naturaliser en France, il est parti aux Etats-Unis juste avant la guerre et a pu y répandre la "bonne parole" de Paul Lévy. Surtout, il a eu un autre élève, Doeblin, extrêmement doué, mais qui malheureusement est mort à la guerre en 1940. J'ai un peu connu Doeblin en 1936, il parlait de se diriger vers les intégrales stochastiques et les équations différentielles stochastiques, ce qui m'a beaucoup impressionné. Mais il n'a pas laissé de travaux sur ce sujet, à ma connaissance. Presque tous les probabilistes français se sont formés en dehors de lui, et c'est très dommage. Ceci explique en partie qu'il n'ait pas été estimé en France, de son vivant, à sa juste valeur. Mais cela tient aussi à ses collègues.

Ses collègues probabilistes parisiens étaient essentiellement Emile Borel (dont il avait suivi les cours à la Sorbonne), Maurice Fréchet et Georges Darmois, qui s'intéressaient d'ailleurs aussi à beaucoup d'autres branches des mathématiques.

Or ils se sont toujours un peu méfiés des idées de Paul Lévy. Après la guerre, Fréchet et Borel, déjà âgés ne se sont plus occupés de probabilités, et Georges Darmois a été professeur de probabilités à la Sorbonne. Le monde universitaire a exercé un ostracisme certain vis-à-vis de Paul Lévy. J'ai moi-même essayé, au cours des années cinquante, d'organiser un cycle de conférences de probabilités par Paul Lévy à la Sorbonne, et je n'ai pas pu y parvenir. Plus curieuse est l'attitude de Jacques Hadamard à son sujet. Hadamard dominait, entre les deux guerres, avec Henri Lebesgue et Elie Cartan, les mathématiques françaises, et le séminaire Hadamard, hebdomadaire, était un remarquable lieu de rencontre de mathématiciens du monde entier. Hadamard a beaucoup estimé les travaux d'analyse de Paul Lévy, mais s'est peu intéressé à ses travaux de probabilités. C'est d'autant plus curieux qu'Hadamard s'est énormément intéressé à la physique, notamment par tout ce qu'il a fait dans la théorie des équations aux dérivées partielles. Mais il considérait que le devoir des mathématiciens était de s'intéresser à la physique et non de devenir physiciens, ils devaient rester les apôtres de la rigueur. Et il a dit une fois qu'il n'avait pas pardonné à Paul Lévy, en se consacrant presque exclusivement aux probabilités, d'avoir quitté les mathématiques pour la physique. Il faut bien dire que les probabilités de Paul Lévy avaient une certaine ressemblance plus avec la physique qu'avec les mathématiques de son temps ; l'attitude de Hadamard reste quand même inexplicable. Paul Lévy en a toujours été très affecté, Hadamard avait été son maître, et il avait une immense admiration (justifiée !) pour lui. Voici ce qu'écrit Paul Lévy, visiblement à l'adresse de Hadamard, dans la Préface de son livre de 1937 sur l'addition des variables aléatoires, 2e édition (1951), page XII : "Je voudrais maintenant dire quelques mots des préventions de certains analystes, et non des moindres, vis-à-vis du calcul des probabilités, ou du moins vis-à-vis des probabilistes, qui n'auraient pas le sens de la rigueur, qui n'auraient pas su à ce point de vue profiter des progrès de l'analyse depuis Cauchy, qui même seraient tentés d'oublier qu'une fonction continue n'est pas toujours dérivable... Ces critiques ont pu être en partie justifiées au siècle dernier. Pour les répéter maintenant, il faut tout ignorer du développement récent du calcul des probabilités..." Et il est bien vrai qu'il y avait ignorance ! Paul Lévy désirait beaucoup entrer à l'Académie et il n'y est entré que très tardivement, comme successeur de Jacques Hadamard, en 1964, à l'âge de 78 ans, 30 ans après sa première candidature.

Il faut aussi mentionner l'attitude un peu contradictoire de l'équipe de Bourbaki vis-à-vis des probabilités. Bourbaki s'est intéressé exclusivement aux mesures de Radon, sur des espaces topologiques localement compacts, délaissant complètement les mesures de Borel sur des ensembles munis de tribus. Or les deux théories de la mesure existent, et on ne peut pas négliger l'une pour l'autre. On doit considérer comme une erreur cette attitude dans laquelle j'ai eu aussi ma part de responsabilité ! J'en ai été, en ce qui me concerne, le bénéficiaire et la victime ; le bénéficiaire dans la mesure où, sans les mesures de Radon, je n'aurais pas pu trouver les distributions (et alors le succès des distributions a fatalement engagé tout Bourbaki à ne plus considérer que les mesures de Radon), et la victime dans la mesure où, ayant quitté les probabilités en 1937 à cause de mon service militaire et ensuite la guerre, ayant ensuite été entraîné par les distributions après la guerre, je me suis désintéressé des probabilités pendant une très longue période, pour n'y revenir que vers la fin des années soixante. La notoriété de Bourbaki ne s'est pas fait sentir avant la deuxième guerre mondiale, elle n'a donc en rien contribué à l'isolement de Paul Lévy à cette époque. Par contre, Bourbaki disposait d'un très fort pouvoir dans les mathématiques en France pendant les années d'après-guerre, et a en partie contribué à empêcher les probabilités de se développer à fond en France. Mais contrairement aux contemporains de Paul Lévy, tous ses membres avaient pour lui un respect énorme. Avec les avis souvent très tranchants dont ils étaient coutumiers, ils considéraient que dans la génération des mathématiciens français d'âge moyen entre les deux guerres, les deux plus grands étaient sans conteste Arnaud Denjoy et Paul Lévy ; je crois que ce jugement reste exact.

Les probabilités ont eu un certain essor en France juste après la guerre, en dehors du solitaire Paul Lévy : il y eut Robert Fortet, Edith Mourier, Jean Bass, Jean Ville, Daniel Dugué, André Blanc-Lapierre, qui ne se vexeront pas si je dis qu'ils n'étaient pas du niveau de Paul Lévy. C'est plus aux Etats-Unis et au Japon que les probabilités ont pris un très grand essor après la guerre, notamment avec Doob puis Feller aux Etats-Unis et Ito au Japon, qui ont largement contribué à porter dans ces pays les probabilités à leur très haut niveau actuel. La France ne rayonnait pas de la même lumière. C'est par des voyages aux Etats-Unis ou par le passage, une année, de Michel Loève à Paris, qu'ont été formés Jacques Neveu et Paul-André Meyer, qui ont ensuite formé la majorité des probabilistes français modernes. Paul Lévy n'y était pas parvenu. Aujourd'hui la situation est bien différente, les probabilités constituent une des très grandes branches des mathématiques en France, et la France est un des plus grands pôles des probabilités dans le monde. Il n'était pas initialement dans mon but de faire l'histoire du développement des probabilités dans notre propre pays, mais c'était un fait singulier à signaler que Paul Lévy n'est arrivé à assurer sa succession en France que par l'intermédiaire des Etats-Unis[1].

---

[1] On trouvera une histoire du développement des probabilités dans le livre de Jean Dieudonné : "Abrégé d'histoire des Mathématiques. 1700-1900", Hermann, ed. 1978, 2 vol. ; le chapitre sur les probabilités a été rédigé par Michel Loève.

Je voudrais maintenant aborder un peu le développement de la vie scientifique de Paul Lévy. Né le 15 septembre 1886, il a, au cours de ses années de lycée, montré tout de suite qu'il était très doué. Il a eu au Concours général un premier prix de thème grec et un prix de mathématiques. Il s'est présenté aux grands concours et a été reçu deuxième à l'Ecole Polytechnique et premier à l'Ecole Normale. Il a choisi d'entrer à l'Ecole Polytechnique. Les raisons de ce choix ne sont pas très claires ; il indique dans sa biographie que la date limite du choix de l'Ecole Polytechnique était antérieure à celle du choix de l'Ecole Normale, et que, un peu pris de panique, il a choisi l'Ecole Polytechnique parce que, s'il ne la choisissait pas, il n'aurait plus la possibilité de le faire ensuite ; c'est une raison qui n'est guère convaincante ! Je pense plutôt qu'il a été influencé par son père, examinateur des élèves à l'Ecole Polytechnique. Il en est sorti dans le Corps des Mines, et s'est très vite dirigé vers la recherche mathématique. Il a d'abord été passionné par l'analyse fonctionnelle à laquelle il a consacré ses premiers travaux. Il a soutenu sa thèse sur ce sujet en 1911, avec le jury suivant : Jacques Hadamard, Emile Picard, Henri Poincaré ! C'était sur les équations aux dérivées fonctionnelles partielles. Cela correspondait aux idées à la mode à la suite de travaux de Volterra et de Gâteaux. Ses principales idées ont été exprimées dans son premier livre d'analyse fonctionnelle en 1922 (publié chez Gauthier-Villars, comme ultérieurement tous ses livres scientifiques), livre réédité sous le nom de "Problèmes concrets d'analyse fonctionnelle" en 1951. Chargé du cours Peccot sur ce sujet en 1919, il a été nommé professeur à l'Ecole Polytechnique en 1920. Son livre a eu un assez grand succès sur le moment, mais est un peu ensuite tombé dans l'oubli. Les problèmes traités, liés par exemple à une étude fine de la manière dont une solution du problème de Dirichlet ou de Neumann dépendait du contour (dans la direction des fonctions de lignes et de surfaces de Volterra, Gâteaux, Hadamard), ont peu intéressé les mathématiciens de l'époque suivante. L'analyse fonctionnelle a pris une direction tout à fait différente avec les travaux commencés par Hilbert et Frédéric Riesz, avec les idées topologiques de Maurice Fréchet, poursuivis surtout par Banach et l'Ecole polonaise, Leray-Schauder, ..., pour un avenir très brillant jusqu'à nos jours, où les mathématiciens français ont joué et jouent encore un rôle de tout premier plan ; mais pas dans la direction indiquée par Hadamard et Lévy. Comme le dit Jean Dieudonné, Hadamard et Lévy connaissaient très bien l'analyse linéaire, mais mal l'algèbre linéaire, point de départ inévitable d'une vraie analyse fonctionnelle.

Il faut ajouter que la rédaction de ce livre d'analyse fonctionnelle laisse à de nombreux endroits beaucoup à désirer. Le profond sens intuitif de Paul Lévy le prédisposait sûrement plutôt à développer les probabilités comme il les a faites, qu'à développer l'analyse fonctionnelle qui allait suivre une voie très rigoureusement axiomatisée. Toujours est-il que lui-même constata un jour qu'il existe un certain passage de son livre d'analyse fonctionnelle qui "semblait avoir été rédigé par des singes dactylographes". Il a fini, dit-il, par trouver ce qu'il avait voulu dire et l'a écrit à Frédéric Riesz qui en a été très intéressé. Malheureusement il a perdu cette lettre et, quand il l'a recherchée plus tard, pour la deuxième édition de son livre en 1951, il ne s'en est plus souvenu et Frédéric Riesz non plus ! Mais maintes idées de ce livre eurent un grand succès en URSS où il a été traduit en 1967, et ses idées sur les problèmes isopérimétriques, très profondes, ont été reprises par des géomètres (Gromov a fait un exposé dessus au présent Colloque), et aussi son étude de l'inégalité isopérimétrique sur la sphère euclidienne a été utilisée avec fruit par V. Milman pour démontrer le théorème de Dvoretzky sur les sections sphériques des corps convexes, un des plus beaux théorèmes de la géométrie des espaces de Banach.

Il m'a plusieurs fois raconté dans sa jeunesse comment il s'était dirigé de l'analyse fonctionnelle vers les probabilités. Tout d'abord, lorsqu'il était élève à l'Ecole Polytechnique, Poincaré a donné deux conférences sur la théorie des erreurs. C'était en 1905. Poincaré n'a pas été satisfait de sa propre rédaction de ses deux conférences, et a demandé à Paul Lévy, premier de la promotion, de les rédiger. Paul Lévy avait été absent lors des conférences et a dû se renseigner auprès de ses camarades : aucun n'avait rien compris. Néanmoins, par des conversations avec ses camarades et avec Poincaré, il a fini par faire une rédaction que celui-ci a trouvée satisfaisante. Puis il a tout oublié ; en fait certainement pas tout, puisque c'est au cours de ses conférences que Poincaré a développé un peu le rôle de la loi de Gauss dans la théorie des erreurs en exprimant qu'il était dû à l'accumulation d'un grand nombre d'erreurs indépendantes et toutes très petites (remarque déjà faite par Laplace). En 1919, un an avant sa nomination comme professeur à l'Ecole Polytechnique, le directeur des études, Carvallo, lui a demandé de faire trois conférences sur les probabilités et de les rédiger. Il accepta avec enthousiasme. Il regarda la littérature qui existait sur le sujet, trouva qu'elle était assez importante, mais qu'elle ne contenait pas grand-chose. Il écrivit à sa femme : "Les probabilités sont au fond à peu près inexistantes ; j'ai deux mois pour les rédiger ; ce n'est pas assez pour avoir le temps de lire cette littérature, mais c'est largement assez pour avoir le temps de tout retrouver". Et c'est en effet à peu près ce qu'il fit.

Tout d'abord, pour les variables aléatoires réelles, il constata qu'on supposait toujours qu'elles ne pouvaient prendre qu'un nombre fini de valeurs ou que leur répartition était une loi ayant une densité par rapport à la mesure de Lebesgue. Il s'aperçut évidemment tout de suite qu'il fallait prendre des mesures pouvant être singulières par rapport à la mesure de Lebesgue, ayant donc une fonction de répartition F croissante sur R avec $F(-\infty)=0$, $F(+\infty)=1$. Conformément au théorème de Lebesgue, elle était la somme de trois fonctions croissantes : une fonction de sauts, une primitive d'une fonction intégrable et une fonction singulière par rapport à la mesure de Lebesgue. Il fut sûr, d'un seul coup, qu'il fallait faire les probabilités dans ce cadre. C'est pour nous, aujourd'hui, tout à fait classique (mis à part le fait que nous considérons la mesure plus que la fonction de répartition, sauf pour étudier les processus à variation finie). C'était connu à l'époque pour les mesures, non imaginé pour les probabilités. Il fut navré de ne pas pouvoir en parler aux élèves de l'Ecole Polytechnique car l'intégrale de Stieltjes n'était pas au programme, mais il travailla les choses pour lui pendant ces deux mois. Il vit que Poincaré s'était servi d'une fonction caractéristique qui était la transformée de Laplace et comprit aussitôt qu'il fallait utiliser l'intégrale de Fourier, dont il avait fait la connaissance en 1912 (à 26 ans), en même temps que la convolution. En réalité Poincaré retardait lui-même car, déjà, Laplace et Cauchy avaient montré qu'il fallait utiliser l'intégrale de Fourier.

C'est pendant ces années, allant de 1920 à 1922, qu'il trouva rapidement un certain nombre des grands théorèmes des probabilités. Tout d'abord, les deux théorèmes fondamentaux sur la fonction caractéristique : l'inversion de la fonction caractéristique permettant de retrouver la fonction de répartition, et l'équivalence de la convergence d'une suite de fonctions de répartition et de la convergence de leurs fonctions caractéristiques. C'étaient là des théorèmes tout neufs à l'époque.

La convergence en loi devait être précisée car elle ne l'était pas réellement, et il saisit parfaitement qu'une suite de mesures converge étroitement (comme nous le disons aujourd'hui) vers une loi limite, si, et seulement si, pour tout ensemble borélien, sa probabilité pour les mesures converge vers sa probabilité pour la mesure limite lorsque la frontière de cet ensemble est négligeable pour la mesure limite. (Ici les F_n, fonctions de répartition, convergent vers F, si elles convergent simplement en tout point de continuité de F. Mais sur la droite, la topologie étroite des mesures est métrisable, et il introduisit une distance de deux fonctions de répartition). D'autre part, la convergence des fonctions caractéristiques n'est pas immédiate puisqu'elle devait être uniforme sur tout intervalle borné, et, d'autre part, il fallait utiliser un artifice de compacité pour passer de la convergence des fonctions caractéristiques à la convergence en loi des variables aléatoires ou convergence étroite des mesures. C'est tellement vrai que tous ces théorèmes n'étaient pas évidents, qu'il m'en parlait encore avec fierté en 1934, époque où pourtant ils faisaient partie normalement du certificat de calcul des probabilités à la Faculté des Sciences de Paris dans les cours de Georges Darmois et d'Emile Borel. C'est encore plus vrai si l'on songe que Borel, qui, par ailleurs, était celui des mathématiciens français qui estimait le plus et comprenait le mieux Paul Lévy (n'oublions pas qu'il a le premier trouvé et démontré la loi forte des grands nombres, pour des variables aléatoires bornées), jugeait assez sévèrement ces théorèmes car, disait-il, les résultats qu'on peut tenter d'en déduire ne sont pas suffisants pour justifier l'emploi de "méthodes aussi transcendantes". C'est à la même époque qu'il trouva ce qu'on appelle les lois stables, employées aujourd'hui couramment par les probabilistes et les spécialistes de la géométrie des espaces de Banach. Une petite partie de ces résultats avait déjà été trouvée par Cauchy et Polya. De nouvelles réflexions apparurent à cette époque sur la convergence des variables aléatoires. Elles résultent d'une séance du séminaire Hadamard en 1924. Un des auditeurs, Noaillon, lui fit observer qu'il y avait une lacune dans son livre d'analyse fonctionnelle : il avait admis (page 122) que, si une suite d'éléments de L^2 converge mutuellement vers zéro, c'est-à-dire est une suite de Cauchy, alors elle converge dans L^2. Autrement dit, il avait admis implicitement que l'espace L^2 était complet. Tous les auditeurs du séminaire Hadamard furent alors stupéfaits et se rendirent compte qu'ils ignoraient absolument si c'était vrai ! Tout le monde se décida à le chercher. A une séance ultérieure, Noaillon donna une démonstration qu'il avait trouvée, qui dura trois quarts d'heure, qui était très compliquée et à laquelle personne ne comprit rien. (Mais rendons-lui justice : cette démonstration était très probablement exacte, quoique non terminée). Tout d'un coup, Lévy eut un éclair et se rendit compte qu'il avait simplement démontré que la suite de Cauchy convergeait nécessairement en probabilité vers une limite. Il restait à démontrer que, si une suite est de Cauchy dans L^2 et si elle converge en probabilité vers une limite, elle converge dans L^2 vers cette limite. C'est une méthode un peu contournée car la convergence en probabilités est plus compliquée que la convergence dans L^2. En réalité, on démontrait-là qu'on peut extraire une suite partielle, presque sûrement convergente, et on savait que, si une suite est telle que de toute sous-suite on peut extraire une nouvelle sous-suite convergeant presque sûrement, elle converge en probabilité. Aujourd'hui, on montre plutôt que, si l'on a une suite de Cauchy dans L^2, on peut en extraire une suite partielle qui converge à la fois presque sûrement et dans L^2, et on sait que toute suite de Cauchy dont une suite partielle converge, est elle-même convergente. On ne passe pas par la convergence en probabilité. De toute façon, il restait un petit problème à résoudre : montrer qu'une suite de Cauchy dans L^2 qui converge en probabilité converge aussi dans L^2. C'est Stefan Banach, présent à la séance, qui résolut cette dernière petite difficulté. Tous furent d'accord, compte tenu de la contribution de Paul Lévy, pour que ce soit lui qui publie la note indiquant le résultat, en citant ceux qui y avaient contribué. A ce séminaire avaient assisté Jacques Hadamard, Paul Lévy, Emile Borel et Stefan Banach ! C'était en 1924. Banach avait à cette époque trente-deux ans. Il allait dans les années qui suivent fonder les principaux théorèmes qui l'ont rendu célèbre, au sujet des espaces vectoriels normés complets, appelés universellement aujourd'hui espaces de Banach. Et tous ces mathématiciens ignoraient que l'espace L^2 était complet. Comment pouvait-on faire de l'analyse fonctionnelle, ou tout simplement de l'analyse, comment pouvait-on se promener dans la rue en ignorant que l'espace L^2 était complet, en 1924 ! Le plus comique est qu'Emile Picard, qui était l'un des plus cultivés de tous, téléphona quelques jours après à Paul Lévy pour lui indiquer que Frédéric Riesz et M. Fischer avaient démontré ce même théorème pour tous les espaces L^p en 1907 ! Mais il en ignorait la démonstration. Ce résultat était resté inconnu de 1907 à 1924, aussi bien en France qu'en Pologne (sauf en partie de Picard et sûrement de Fréchet), alors qu'il était couramment manipulé par les écoles de Hilbert et F. Riesz. (Il y a sans doute des liens entre cette coupure, France + Pologne, Allemagne + Hongrie, et celle de la guerre de 1914-18). Et plus spécialement Paul Lévy l'avait cru évident pour L^2 dans son livre d'analyse fonctionnelle. A titre d'amusement, pour la petite histoire, il n'est pas mauvais de dire qu'à la même époque, aux environs de 1907, Frédéric Riesz avait montré que le dual de L^2 était L^2, en assimilant l'espace L^2 à l'espace ℓ^2 grâce aux bases orthonormées, et que, enthousiasmé par ce résultat, il avait essayé, pendant trois semaines, de montrer que, pour p quelconque, le dual de L^p était aussi L^p ! Il n'avait pas tardé à voir que c'était L^q, où q est l'exposant conjugué de p, mais il avait suffisamment cherché à démontrer que le dual de L^p était L^p pour que toute la communauté mathématique de Budapest en soit informée. L'article de Paul Lévy parut dans le Bulletin des Sciences Mathématiques, tome XLIX, 1925, sous le titre : "Sur le théorème de MM. M. Fischer et Fr. Riesz sur la convergence en moyenne".

Pour revenir à 1924, c'est de cette séance du séminaire Hadamard que datent de nombreuses réflexions de Paul Lévy sur les divers types de convergence des variables aléatoires : convergence en loi, convergence en probabilité, convergence en moyenne d'ordre p, convergence presque sûre. Ce sont aussi devenues des choses classiques et qui, quand j'étais jeune, étaient enseignées dans les cours de calcul des probabilités à la Faculté des Sciences de Paris. Paul Lévy jugea qu'il était temps d'écrire ce qu'il savait sur les probabilités dans un livre, qui est donc son deuxième livre, "Calcul des probabilités" (1925). Dans les années suivantes, il démontra encore de nombreux théorèmes, qui, souvent, avaient été partiellement, ou même presque complètement faits par d'autres. Il démontra la loi du logarithme itéré en 1929 ; mais Emile Picard lui dit que Khintchine l'avait trouvée quelques années auparavant ; il s'aperçut avec confusion que le résultat de Khintchine était de 1924. Il démontra, en 1930, le théorème des trois séries pour la convergence d'une série de variables aléatoires indépendantes ; or Kolmogorov l'avait démontré quelques semaines avant lui. Il en fut d'autant plus confus que, trois semaines avant, il avait reçu une visite de Kolmogorov avec lequel il avait longuement parlé et qui lui avait apporté un article contenant ce résultat mais qu'il n'avait pas lu. En 1935, il trouva le théorème des probabilités zéro ou 1 mais Kolmogorov l'avait trouvé en 1934. Enfin, dans ces mêmes années trente, il a complètement élucidé le rôle de la loi de Gauss dans l'addition des variables aléatoires indépendantes. On pourrait considérer comme une catastrophe qu'il ait trouvé autant de résultats déjà juste connus un peu auparavant. Il n'en est rien. Il a toujours perfectionné ces résultats de manière notable. Il a su les manier de façon parfaite et en tirer un grand nombre de conséquences nouvelles. Il en a poussé les applications aussi loin que possible. Par ailleurs, cela montre aussi son grand esprit inventif, car il a incontestablement trouvé ces résultats en ignorant absolument qu'ils existaient et en y travaillant entièrement seul.

Il est difficile de dire exactement qui a trouvé le théorème central limite. Voici l'énoncé :

Soit $(S_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles ; on dit qu'elle est asymptotiquement gaussienne pour n infini s'il existe une suite $(a_n)_{n\in\mathbb{N}}$ de nombres réels et une suite $(\tau_n)_{n\in\mathbb{N}}$ de nombres >0 telles que la loi de $(S_n-a_n)/\tau_n$ tende vers la loi de Gauss normale. Supposons que $S_n = X_{n,1} + X_{n,2} + \ldots + X_{n,p_n}$, où les $X_{n,k}$, $k=1,2,\ldots,p_n$, sont de variables indépendantes. Si alors les $X_{n,k}$ ont une moyenne nulle, et des moments d'ordre 2 finis, $\sigma_{n,k}$, $\sum_{k=1}^{p_n} \sigma^2_{n,k} = \tau^2_n$, et si, pour tout $\epsilon>0$, $\sum_{k=1}^{p_n} \int_{|x|>\epsilon\tau_n}^{+\infty} x^2\, \mu_{n,k}(dx)$, où $\mu_{n,k}$ est la loi de $X_{n,k}$, tend vers 0 pour $n \to +\infty$, alors $(S_n)_{n\in\mathbb{N}}$ est asymptotiquement gaussienne, et la loi de S_n/τ_n tend vers la loi normale de Gauss.

Sous une forme un peu moins générale, ce sont, indépendamment et à peu près simultanément, Lindeberg et Paul Lévy qui ont trouvé ce résultat. La démonstration de Paul Lévy est immédiate par la fonction caractéristique ; celle de Lindeberg utilise la convolution sans Fourier, donc est, en un sens, plus "élémentaire" et un peu plus longue. Paul Lévy de toute façon admirait beaucoup la méthode de Lindeberg et correspondit longuement avec lui. C'est peut-être cette situation qui a provoqué le jugement négatif de Borel sur les méthodes de Paul Lévy. Mais aujourd'hui, avec le triomphe moderne de Fourier, qui se glorifierait d'utiliser la convolution et pas Fourier ? Mais c'est sûrement Paul Lévy qui a trouvé le théorème le plus général sur la condition pour qu'une somme de variables aléatoires indépendantes soit asymptotiquement gaussienne. On ne doit pas en effet supposer l'existence de moments, car des valeurs des $X_{n,k}$, très grandes mais très peu probables, suppriment les moments mais n'influent pas sur le résultat. L'énoncé général est le suivant :

Soient S_n , $X_{n,k}$, comme ci-dessus, mais sans hypothèse de moments. Soit $a_{n,k}$ une valeur médiane de $X_{n,k}$, a_n une valeur médiane de S_n. Soit $\alpha \in]0,1[$, et soit L_n le plus petit nombre tel que $P\{S_n \in [a_n-L_n, a_n+L_n]\} \geq \alpha$ (si par exemple $\alpha=1/2$, L_n est une valeur médiane de $|S_n-a_n|$). Alors on dit que les $|X_{n,k}-a_{n,k}|$ sont, pour n grand, individuellement négligeables devant la somme S_n-a_n si, quel que soit $\epsilon>0$, $\sup\limits_{0\leq k\leq p_n} P\{|X_{n,k}-a_{n,k}| > \epsilon L_n\}$ tend vers 0 pour n tendant vers $+\infty$. On dit que les $|X_{n,k}-a_{n,k}|$ sont uniformément négligeables devant S_n-a_n si :
$P\{\sup\limits_{0\leq k\leq p_n} |X_{n,k}-a_{n,k}| > \epsilon L_n\}$ ou bien $\sum\limits_{0\leq k\leq p_n} P\{|X_{n,k}-a_{n,k}| > \epsilon L_n\}$ tend vers 0 pour n infini (le plus grand des $|X_{n,k}-a_{n,k}|$, $0\leq k\leq p_n$, est négligeable devant S_n-a_n, pas seulement chacun individuellement). Alors une condition suffisante pour que S_n devienne gaussienne quand $n \to +\infty$, est que les $|X_{n,k}-a_{n,k}|$ deviennent globalement négligeables, et c'est nécessaire s'ils deviennent individuellement négligeables. Le résultat est indépendant du choix de α, $0<\alpha<1$.

Ce théorème lui donna beaucoup de mal. Il était nécessaire de montrer, pour l'affiner encore plus, que la somme de deux variables indépendantes ne peut être gaussienne que si chacune l'est ; je me rappelle qu'il a cherché longtemps ce théorème (et moi aussi à ce moment) mais c'est finalement Cramer qui l'a trouvé ; Paul Lévy ne s'en est jamais consolé. On appelle aujourd'hui ce résultat le théorème de Lévy-Cramer, car suggéré par Lévy en 1934 et trouvé par Cramer en 1936. La démonstration est d'ailleurs simple, elle fait appel à la théorie des fonctions entières.

Dans les années 1936-37, il était devenu le maître à penser incontesté pour les lois gaussiennes relatives aux sommes d'un grand nombre de variables aléatoires indépendantes. Tous ses résultats font l'objet de son troisième livre, de 1937, "Théorie de l'addition des variables aléatoires". Telles étaient les conséquences très lointaines des deux conférences de Poincaré sur la théorie des erreurs, quand il était élève à l'Ecole Polytechnique !

C'est aussi à cette époque qu'il a trouvé ses premiers résultats sur le mouvement brownien, sur les processus additifs (à accroissements aléatoires indépendants) liés aux lois indéfiniment indivisibles, sur l'arithmétique des lois indéfiniment divisibles. Il a commencé aussi à parler des intégrales stochastiques, et même des équations différentielles stochastiques. J'ai été complètement éloigné de lui après cette date ; j'ai commencé mon service militaire en 1937, la guerre est venue tout de suite après, et nous avons été séparés pendant la guerre et l'occupation de la France par les nazis. Il a vécu dans la famille de sa deuxième fille, Denise Piron (Robert Piron n'était pas juif), dans des conditions qui lui ont apporté, sinon un certain confort, du moins une certaine sécurité et une bonne ambiance familiale qui lui a rendu supportable cette dure période. Quant aux mille difficultés de la vie quotidienne, il les accepta avec la sereine philosophie qui a toujours été la sienne. A ma connaissance, le remplacement de carreaux cassés a été le seul travail manuel qu'il ait jamais fait. A partir d'une certaine date, nous lui avons procuré de toute façon de fausses cartes d'identité, mais il a eu un très grand scrupule à s'en servir ; il avait l'habitude de porter sur lui à la fois la vraie carte et la fausse carte et, une fois, alors qu'il était contrôlé dans un train par un officier allemand, il a montré la vraie carte portant le nom de Lévy ; il ne lui est fort heureusement rien arrivé. Après cette malencontreuse expérience, il a éliminé tout papier d'identité au nom de Lévy.

Il avait près de soixante ans en 1945, mais c'est d'après guerre que datent encore certains de ses meilleurs travaux. Il a alors complètement étudié le mouvement brownien et a publié l'ensemble de ses résultats en 1948 dans son quatrième livre : "Processus stochastiques et mouvement brownien". Ces résultats sont absolument fantastiques. Il a étudié les problèmes d'aire de la courbe brownienne dans $\mathbb{R}^2$ (qui est nulle, il l'a très finement expliqué)(1), de points multiples du mouvement brownien, d'aire entourée par le lacet brownien. Il a été jusqu'à introduire le temps local et les excursions browniennes, comme le montre l'exposé de Chung dans ce même colloque, et a aussi réfléchi aux chaînes de Markov. Il était vraiment au coeur du calcul des probabilités modernes. Il a également abordé les mouvements browniens à plusieurs paramètres de temps et a trouvé, à l'âge

---

(1) La nullité de cette aire équivaut, par Fubini, à dire qu'un ensemble réduit à un point est polaire, c'est-à-dire n'est pas rencontré par la trajectoire brownienne.

de soixante-seize ans, en 1962, un très joli résultat : le mouvement brownien à paramètre de temps hilbertien de dimension infinie est déterministe ! Il mourut en 1971, âgé de quatre-vingt-cinq ans, après avoir publié un dernier livre, qui est plus qu'une autobiographie : "Quelques aspects de la pensée d'un mathématicien", Blanchart, 1970 ; quelques passages de cet article sont empruntés à ce livre.

Je crois qu'il n'est pas exagéré de dire que les grands fondateurs du calcul des probabilités modernes sont Kolmogorov et Lévy. Mais Kolmogorov a introduit complètement la rigueur en probabilité, en introduisant en 1933 l'espace Ω des épreuves et toute l'axiomatique nécessaire. Il est assez extraordinaire que Paul Lévy se soit toujours refusé à utiliser cet espace. Il est vrai que les probabilistes dans leurs formules n'écrivent jamais la variable ω, mais ils savent bien qu'elle existe et l'utilisent quand il le faut. Paul Lévy ne savait pas qu'elle existait, ne voulait pas qu'elle existât et ne connaissait pas la notion de tribu. Pour faire tout ce qu'il a fait en se passant de Ω, il faut évidemment être un virtuose d'acrobatie mathématique. Et souvent certains de ses énoncés étaient surprenants. Certaines de ses notations aussi. Par exemple, quand il voulait écrire une intégrale stochastique brownienne, au lieu d'écrire dW_t, comme tout le monde le fait aujourd'hui, il écrivait $\xi_t\sqrt{dt}$, ξ_t étant une variable aléatoire gaussienne normale. Il savait bien que ce n'était pas vrai mais c'était une simple notation. Toutefois les notations ont souvent des conséquences. Les mathématiciens, y compris moi-même, ont été longtemps rebutés par la "fonction" de Dirac. Je crois que cette notation pour l'intégrale stochastique était peu heureuse ; en tout cas, personnellement, elle m'a tout à fait éloigné de l'étude de ces questions, que je n'ai reprise que bien plus tard, à la fin des années soixante, quand je suis revenu aux probabilités, en étudiant l'intégrale stochastique dans les exposés modernes. L'omission de Ω était lourde de conséquences. Par exemple, Paul Lévy n'avait aucune définition rigoureuse des temps d'arrêt. Toutes ses conceptions sur l'indépendance ou la dépendance relative des variables aléatoires étaient de nature physique : il dit bien que deux évènements A, B sont indépendants si et seulement si $\mathbf{P}\{A \text{ et } B\}=\mathbf{P}(A)\mathbf{P}(B)$, mais il ne définit pour la première fois l'indépendance de deux variables réelles aléatoires X,Y, que dans son livre de 1948 (par $\mathbf{P}\{X<x \text{ et } Y<y\}=\mathbf{P}\{X<x\}\mathbf{P}\{Y<y\}$), sans par ailleurs jamais définir les variables aléatoires autrement que comme résultats d'un tirage au sort ; deux variables aléatoires sont indépendantes si elles résultent de deux choix faits tout à fait indépendamment l'un de l'autre, c'est-à-dire elles sont indépendantes si elles sont indépendantes.

Il m'a d'ailleurs très souvent expliqué, en citant aussi des travaux philosophiques de Von Mises, que les probabilités avaient une base complètement physique mais qu'on pouvait ensuite en déduire un certain nombre de règles permettant de faire des mathématiques. C'est ce qu'il n'a pas pu faire comprendre à Jacques Hadamard ! En fait, c'est exact, comme pour la théorie du potentiel, mais chez lui la physique se poursuivait très longtemps. L'absence de Ω l'a empêché de comprendre la différence entre la propriété de Markov simple et la propriété de Markov forte. Paul Lévy a refusé l'exemple de Chung relatif au processus de Markov trivial sur une lemniscate de Bernoulli (en forme de 8), en disant que le problème était mal posé et que tout processus de Markov simple était de Markov fort. En effet, disait-il, on complique "artificiellement" le problème, on le pose mal car la lemniscate possède un point double, et il faut dédoubler ce point comme on dédouble des points d'une région du plan complexe en introduisant une surface de Riemann au-dessus du plan. A ce moment-là, ce processus de Markov devient effectivement fort. Mais évidemment, si on prend des exemples beaucoup plus compliqués que celui-là, cette démultiplication par surface de Riemann ne peut plus régler le problème ! Dans son autobiographie de 1965, il répète que l'exemple donné par Chung est un problème mal posé ! Il eut aussi une discussion très vive avec Fréchet avant la guerre : Fréchet, qui aimait beaucoup les espaces abstraits, voulut introduire l'espace de toutes les variables aléatoires réelles (sans Ω !). Paul Lévy a refusé cet espace mais il n'est pas -et pour cause- arrivé à le faire comprendre à Fréchet. Il lui disait simplement : "Si vous prenez l'espace de toutes les variables aléatoires, vous perdez le système de toutes leurs corrélations". A quoi Fréchet lui répondait qu'il suffisait de se donner, en même temps que toutes les variables, l'ensemble de toutes leurs corrélations mutuelles ou encore la loi simultanée de toutes les variables. La discussion pouvait durer longtemps ! Ajoutons que Paul Lévy utilise constamment la distance en probabilité de deux variables aléatoires X,Y, justement introduite par Fréchet : $\operatorname{Inf}_{\epsilon>0} (P\{|X-Y|>\epsilon\} \leq \epsilon)$, ceci sans espace Ω ! Il va plus loin. Il a expliqué que l'expression d'un processus, comme une fonction du temps et de l'épreuve, $X(t,\omega)$, masque toute la réalité du calcul des probabilités. Quand on connaît le processus jusqu'à l'instant t, dit-il, on ne connaît pas encore l'avenir, et c'est justement là qu'est l'esprit même du calcul des probabilités. Alors que, si l'on écrit $X(t,\omega)$, dès que ω est donné, on connaît la trajectoire jusqu'au bout ! "Le choix de la variable ω résume toutes les interventions successives du hasard. C'est une notation souvent commode, bien qu'elle donne comme un tout né en un instant ce qui, pour moi, est essentiellement un perpétuel devenir" (livre de 1937, 2e édition, 1954, page 360). Pourtant il connaissait bien le jeu des espérances conditionnelles. Il avait justement eu une discussion avec un philosophe au sujet de la droite réelle. Ce philosophe lui disait que, si l'on définit la droite réelle par la méthode Dedekind, en appelant nombre irrationnel une coupure dans l'ensemble des nombres rationnels, il n'était pas possible d'obtenir pour cette droite toutes les propriétés intuitives que nous en connaissons. On ne pouvait, par exemple, imaginer le mouvement continu d'un point sur cette droite. Paul Lévy lui avait expliqué (et il m'en avait parlé) que justement on y arrive sans aucune difficulté car on a une définition précise de la continuité avec ϵ,η. Et il disait à ce philosophe : "Du moment qu'il n'y a pas une seule chose relative à la droite, répondant à notre intuition et qui ne puisse pas être traduite par des expressions mathématiques, cela prouve que notre définition mathématique rend compte de tous les phénomènes qui concernent la droite. Nous trouverons peut-être un jour des phénomènes intuitifs dont notre définition ne rend pas compte, à ce moment on devra enrichir la définition, mais on n'en a pas encore trouvé". Petit coup de pied du diable : la droite non standard, avec des points statiques infiniment voisins. Et il est bien évident qu'il en est ainsi de toutes les définitions mathématiques. Ce que l'on donne comme définition de l'espace euclidien tri-dimensionnel, de la notion d'angle ou d'aire ou de volume, comme définition d'une variété et d'un vecteur tangent à cette variété, se fait toujours en des termes qui ne satisfont pas immédiatement notre intuition, mais qui permettent ensuite par leurs propriétés, d'en rendre compte complètement. Paul Lévy le savait bien et savait bien l'expliquer lui-même, mais il ne l'admettait pas dans le cas des probabilités. Il s'en est quelquefois servi dans le cas d'une suite de variables réelles aléatoires, avec $\Omega = [0,1]^{\mathbb{N}}$, mais pas vraiment, et jamais pour les processus "où le hasard intervient à tous les instants".

J'arrive au terme de cette description de l'homme et du mathématicien, basée en grande partie sur la connaissance étroite que j'ai eue de lui pendant un grand nombre d'années. J'espère n'avoir rien travesti de la réalité. Je voudrais surtout dire que Paul Lévy n'est pas un homme du passé, de la naissance duquel on célèbre aujourd'hui (avec un an de retard !) le centenaire. Il a été un des plus grands créateurs des probabilités, pas seulement pour le passé et le présent, mais pour l'avenir. Il suffit d'assister à n'importe quel séminaire ou cours de probabilités aujourd'hui pour voir qu'on y étudie encore des problèmes posés et souvent, en grande partie, résolus par Paul Lévy. Son oeuvre est extraordinairement vivante et ne passera pas à l'état de repos avant de longues années.