Paul Lévy-Maurice Fréchet

50 ans de correspondance

Laurent MAZLIAK
Laboratoire de Probabilités et Modèles Aléatoires
Universités Paris VI et VII
4, Place Jussieu
75252 Paris Cédex 05 France

March 2, 2004

1 Introduction

Le présent exposé rend compte d'un travail d'édition de l'importante correspondance échangée entre deux mathématiciens français majeurs du XXème siècle, Maurice Fréchet et Paul Lévy. Engagé en juin 2002 en commun avec Marc Barbut (EHESS) et Bernard Locker (Paris V), ce travail donne lieu à la publication d'un livre aux éditions Hermann. Celui-ci contient l'ensemble de la correspondance commentée et annotée, ainsi qu'une introduction détaillée, représentant un quart de l'ouvrage, qui tente de synthétiser le cadre scientifique dans lequel s'est développée la pensée des deux savants. L'exceptionnelle durée de la correspondance (de 1918 à 1965), et la longueur importante de nombreuses lettres, ont rendu la mise au point des textes laborieuse. Nous espérons être arrivés à un résultat satisfaisant, qui puisse donner une idée juste de la valeur de la série de documents que nous avons eue en mains.
Pour dire encore deux mots sur la genèse du projet, il faut tout de suite souligner que Bernard Locker en a été l'initiateur. Il avait soutenu en décembre 2001 une thèse à l'Université Paris V, réalisée sous la direction de Bernard Bru, consacrée aux travaux de Paul Lévy pendant la Deuxième Guerre Mondiale, et ses recherches se fondaient de façon importante sur les lettres écrites à Maurice Fréchet dans la période. Il était alors naturel de vouloir prolonger ce travail par l'édition de la correspondance complète (dont B.Locker avait réalisé une première transcription dans sa thèse) et notre petite équipe a été mise en place. Je profite de l'occasion pour remercier Bernard de sa confiance et de m'avoir permis de participer à cette longue et passionnante expérience.
L'ensemble des lettres dont nous allons parler est contenu dans le gigantesque fonds d'archives que la famille de Maurice Fréchet a légué à l'Académie des Sciences de Paris. Nous avons retrouvé 107 lettres de Lévy à Fréchet... et aucune dans l'autre sens. Il est donc nécessaire d'expliquer, pour commencer, cette dissymétrie. On peut d'abord insister sur le fait que Fréchet s'est montré tout au long de sa vie un archiviste au zèle irréprochable: il conservait des montagnes de documents, et en particulier toutes les lettres qu'il recevait, ce qui représente un nombre considérable en quelques soixante-dix années de vie professionnelle.

Ainsi le fonds Fréchet de l'Académie contient-il des échanges avec des dizaines de figures plus ou moins marquantes du siècle, richesse qui d'ailleurs est loin d'avoir été exploitée complètement. Paul Lévy n'a pas eu, loin s'en faut, le même soin à la conservation, et il fait lui-même remarquer quelquefois dans les lettres qu'il a égaré tel ou tel papier important! Mais un autre point éclaire aussi la situation: en 1940, Lévy a dû quitter Paris pour suivre l'Ecole Polytechnique (où il était professeur) qui s'était installée en Zone Sud à Lyon, puis, dans la phase la plus amère de l'Occupation, rentrer dans la clandestinité en raison de ses origines juives. En 1942, son appartement parisien fut saccagé et tout document antérieur à 1940 semble ainsi irrémédiablement perdu. Il a donc fallu se résigner à n'entendre Fréchet qu'à mi-voix. Certaines parties de la correspondance où les lettres sont très rapprochées permettent une tentative de reconstitution approximative; pour d'autres il faut nous contenter des réponses de Lévy, sans avoir les questions posées. Notons aussi que ce que nous avons ne constitue manifestement pas l'ensemble des lettres que Lévy envoya à Fréchet: nous avons la première (ils y font connaissance), pas la dernière (pourquoi la correspondance se serait-elle arrêtée en 1965?), et dans cet intervalle quelques trous, dont le plus important s'étend de septembre 1931 à janvier 1936. Mais il est remarquable que les années noires de la guerre pour Lévy n'aient pas interrompu les envois; bien au contraire, nous devons à cette période, et notamment à l'année 1943 où Lévy se consacre entièrement à l'étude du mouvement brownien, un flot de considérations livrées à Fréchet (plus ou moins seul interlocuteur de Lévy alors) qui formeront une grande partie du livre de 1948 *Mouvement brownien et processus stochastiques.* Notons aussi que la correspondance se poursuit même dans des périodes où les deux hommes se côtoyaient pour ainsi dire quotidiennement. Ils avaient visiblement trouvé là une forme de communication qui leur convenait à tous les deux, permettant d'échanger des idées après les avoir un minimum mises en forme.
De ce fait, et malgré les lacunes, ce qui nous reste est un document passionnant, qui fait côtoyer en direct (c'est émouvant!) la pensée très créative d'un grand scientifique qui se donne entièrement à son travail. Le ton employé dans les lettres restera toujours celui, professionnel et courtois, qui sied entre deux personnes ayant fait leurs études supérieures dans le cadre choisi des écoles de la IIIème république. A l'exception d'une période très particulière, le discours est pratiquement exclusivement mathématique mais on y sent le témoignage d'un temps, quelque peu révolu, où faire des mathématiques était conçu comme un humanisme et où règne encore pour une part l'atmosphère teintée de scientisme qui prévalait dans les années où nacquirent les deux protagonistes.

2 Biographie sommaire des deux protagonistes

Né en 1878 à Maligny (Yonne), Maurice René Fréchet, après sa sortie de l'Ecole Normale Supérieure, commença sa carrière de mathématicien professionnel en soutenant en 1906, sous la direction de Jacques Hadamard, une brillante thèse à la Sorbonne créant le concept d'espace métrique et développant les notions qui fondèrent la topologie moderne, tout spécialement autour des espaces de fonctions et de la compacité. Son résultat le plus célèbre, découvert simultanément par Riesz, est le théorème de représentation des fonctionnelles linéaires continues sur les espaces de Hilbert. Sa première chaire universitaire (de mécanique

rationnelle) fut obtenue à l'Université de Poitiers, où il professa jusqu'en 1914. Nommé après la Ière Guerre Mondiale - au moment donc où commence la correspondance avec Lévy - professeur à l'Université de Strasbourg afin de la refranciser après l'occupation allemande, Fréchet va effectuer un virage dans sa carrière mathématique en s'intéressant de plus en plus aux aspects appliqués de la discipline et tout particulièrement aux développements récents de la statistique dont il va devenir pendant trente ans le plus éminent représentant français. Il joue au cours du XXème siècle un rôle académique de premier plan, et se fait un ardent propagandiste des mathématiques dans la société de son temps en écrivant, à l'instar de son maître Borel, de multiples articles de large diffusion, en plus de traités techniques faisant le point sur de grandes avancées mathématiques, notamment celles qui touchaient au calcul des probabilités. Ce brillant éclectisme sera d'ailleurs partiellement la cause du manque de considération dans lequel le tiendra, très injustement, la jeune génération française et particulièrement les membres de Bourbaki. Il meurt en 1973 à Paris.
De huit ans le cadet de Fréchet, Paul Pierre Lévy nait en 1886 à Paris dans une famille d'origine juive qui comprenait déjà plusieurs mathématiciens. Après des études exceptionnellement brillantes, il choisit assez curieusement d'entrer à l'Ecole Polytechnique plutôt qu'à l'Ecole Normale Supérieure alors qu'il avait très tôt décidé de se consacrer à la recherche mathématique. Il entra ensuite dans le corps des Mines, et devint professeur à l'Ecole des Mines, ce qui lui laissa relativement peu de temps pour la recherche, comme il le raconte dans son autobiographie. Néanmoins, sous la direction bienveillante d'Hadamard (comme ce fut le cas pour Fréchet), il entreprit une thèse en Analyse Fonctionnelle qu'il soutint en 1912. La Ière Guerre Mondiale va l'amener à s'occuper, dans le cadre de son engagement militaire, de problèmes pratiques de défense (notamment de la toute nouvelle situation de l'aviation). Au sortir de la guerre, Hadamard le charge de la récupération des papiers de Gateaux, tombé au Front. Par ailleurs, en 1919, il est nommé chargé de cours à l'Ecole Polytechnique, et à cette occasion, il découvre la discipline qu'il allait être amené à révolutionner, le calcul des probabilités. Sa manière de pratiquer les mathématiques, marquée d'une intuition stupéfiante, notamment géométrique, doit sans doute beaucoup à sa formation de polytechnicien. L'orientation qu'il donnera à la recherche en probabilités - il sera pendant une quinzaine d'années l'unique grand probabiliste français - se situera ainsi dans l'héritage du style de Poincaré, à l'interface entre les mathématiques et la physique. Ceci se ressentira spécialement dans son plus grand achèvement, la théorie fine du mouvement brownien, qui occupe une part majeure de la correspondance avec Fréchet. On lui doit aussi d'avoir introduit de multiples techniques fondamentales, en particulier pour l'étude des sommes de variables aléatoires (indépendantes et dépendantes). Il meurt à Paris en 1971.
La mise en perspective des carrières des deux mathématiciens est ainsi étonnante. Maurice Fréchet commence son activité scientifique par des sujets extrêmement théoriques sur la topologie des espaces de fonctions et va progressivement se rapprocher des mathématiques appliquées, de par ses responsabilités dans la communauté universitaire, au point de devenir leur défenseur emblématique. Paul Lévy, de par sa formation X-Mines et sa situation de professeur en Ecole d'Ingénieurs, va cultiver dans son enseignement et son travail créatif un style personnel où il aimera à raisonner souvent avec des explications teintées de considérations de physique au détriment de la rigueur obsessionnelle de certains

mathématiciens. Naturellement, ce fait rend parfois ses raisonnements quelques peu difficiles à suivre, et on constate à plusieurs reprises que Fréchet ne s'est pas privé de demander des explications.
Lévy et Fréchet eurent des rapports placés dès le début sous la tutelle spirituelle de leur maître Hadamard. C'est donc dans une atmosphère de complicité où transparaît en demi-teinte le désir de chacun d'être le meilleur élève, voire le préféré du Maître, que le dialogue entre les deux hommes va se développer. Emile Borel, le mathématicien français le plus influent de son époque, aura lui aussi un rôle de premier plan dans le déroulement de leur vie de chercheurs. Avec Hadamard, il est présent tout au long de leur correspondance. Après bien des vicissitudes, le destin les fera équitablement succéder, à huit ans d'intervalle (donc au même âge respectable de 78 ans), aux fauteuils de leurs deux grands maîtres à l'Académie, Borel (Fréchet est élu en 1957) et Hadamard (Lévy est élu en 1964).

3 Quelques illustrations tirées de la correspondance

3.1 Le style de Lévy

Quelques mots tout d'abord concernant le discours mathématique de Lévy. Tout ceux qui ont eu à s'affronter à ses livres savent qu'il n'est pas aisé à suivre (c'est le moins qu'on puisse dire). Dans ses recherches sur le calcul des probabilités, Lévy ne se départit jamais d'une certaine vision physique des choses et (il le dit lui-même) préfère souvent faire "sentir" la façon dont les choses se passent que formaliser les démonstrations. De ce fait, il ne s'est jamais approprié les axiomatisations diverses de l' "Analyse Moderne", et en particulier celles de Kolmogorov pour le modèle probabiliste et de Doob pour les processus. Pour Lévy par exemple, un processus stochastique se construit instant par instant sous l'impulsion du hasard et ceci explique aussi pourquoi il sera toujours obsédé par l'obtention de procédé effectifs de construction. Les lettres, qui sont souvent une tentative première pour exposer ses intuitions, sont de ce fait régulièrement délicates à suivre. Le dialogue avec Fréchet ne manque pas de piquant également pour ces raisons. Fréchet, qui est par bien des côtés un structuraliste avant la lettre (la lecture de sa thèse par exemple montre une façon de rédiger étonnemment en avance sur son temps), est visiblement assez souvent décontenancé par des affirmations de Lévy, ce qui oblige ce-dernier à redoubler d'efforts pour mettre au point ses idées.

3.2 1918-1921: Analyse fonctionnelle

On a un peu oublié aujourd'hui que Lévy, jusqu'à l'âge de quarante ans fut un analyste et un spécialiste de la théorie du potentiel, avant de devenir probabiliste. A la fin du XIXème siècle, l'école italienne autour de Volterra s'était préoccupée de définir un calcul des variations pour les "fonctions de lignes", autrement dit les fonctions de fonctions. Dans l'étude du problème de Dirichlet pour des contours se déformant, Hadamard s'était saisi de ce cadre pour obtenir des équations satisfaites par les fonctions de Green correspondantes. C'est lors d'un cours d'Hadamard au Collège de France en 1909, comme le raconte lui-même

Lévy, que lui vint l'idée d'étudier pour elles-mêmes les équations en question et cela devint le sujet de sa thèse qu'il soutint quelques années plus tard. Mais il ne s'arrêta pas là, et continua à généraliser ses études en dimension trois. Le fait qu'Hadamard lui demande de s'occuper des papiers laissés impubliés par Gateaux, tombé au Front, lui donna un tremplin à la sortie de la guerre pour poursuivre ses extensions aux espaces de dimension infinie, et en particulier pour chercher à élaborer le problème de Dirichlet en dimension infinie (généralisation du laplacien). Chargé du Cours Peccot en 1919, il y fait quelques leçons d'analyse fonctionnelle, qui composeront son premier livre paru en 1922. C'est à l' occasion de la préparation de ces cours qu'Hadamard lui conseilla aussi de prendre contact avec Fréchet qui s'était occupé de calcul différentiel dans les espaces de dimension infinie. Gateaux, dans ses notes, s'était en particulier intéressé à la validité d'une propriété de moyenne pour des fonctionnelles définies sur un espace de dimension infinie (en particulier pour la sphère unité de l'espace des fonctions continues de [0,1] dans $\mathbb{R}$). Il avait défini une notion de moyenne pour une telle fonctionnelle Φ en passant à la limite sur la moyenne (finie-dimensionnelle) de $\Phi[x_n]$ où x_n est la n-ième section de la fonction x, c'est-à-dire la fonction prenant sur chacun des intervalles $[\frac{i}{n}, \frac{i+1}{n}]$ la valeur constante égale à la moyenne de x sur cet intervalle. En fait, Gateaux n'avait réussi à justifier ce passage à la limite que dans le cas où l'approximation de $\Phi[x]$ par $\Phi[x_n]$ est uniforme, et avait émis quelques doutes quant à la portée de cette hypothèse. Lévy va confirmer que cette hypothèse est trop restrictive; il va montrer qu'il est possible de considérer des fonctionnelles beaucoup plus générales pour lesquelles une convergence uniforme peut être également obtenue.

extrait de la lettre du 16 février 1919

...je représente d'abord une fonction x par sa n ième section x_n au sens de Gateaux, c'est à dire par une fonction x_n constante dans chacun des intervalles $(0, 1/n), (1/n, 2/n)$,...,$(n-1/n, n/n)$ et égale dans chaque intervalle à la moyenne de x. La fonctionnelle $\Phi\ |[x]|$ est approchée, mais non uniformément, par $\Phi\ |[x_n]|$. Gateaux considère à plusieurs reprises les fonctionnelles pour lesquelles cette convergence est uniforme dans tout domaine fini. Elles sont très particulières. En raisonnant d'une manière approchée je puis dire: pour que $\mid \Phi - \Phi_n \mid < \varepsilon$ pour n assez grand, il faut que la contribution due à la variation de x dans chacun des intervalles d'amplitude $\dfrac{1}{n}$ à cette intégrale puisse être limitée par $\dfrac{\varepsilon}{n}$ (ε indépendant de la fonction x). Or si $x = x_n + a$ sur la moitié de l'intervalle et $x_n - a$ sur l'autre, cette contribution est à peu près égale à

$$(4)\frac{1}{2n}\int_{x_n}^{x_n+a} \Phi'_x\ |[\xi(s), t]|\ d\lambda - \frac{1}{2n}\int_{x_n-a}^{x_n} \text{expression identique}$$

t valeur particulière de l'intervalle considéré
$\xi(s)$ fonction $= x_n(s)$ en dehors de l'intervalle considéré
$x_n \pm a\lambda$ sur les deux moitiés de cet intervalle.
Ceci est le produit de $1/n$ par une quantité qui n'est très petite quels que soient x et a que si Φ'_x est indépendant de ξ, c'est à dire si le terme $A_0\delta x$ de la formule (2) appliqué à Φ'_x manque. Donc les fonctionnelles de Gateaux sont les fonctionnelles pour lesquelles

$$\Phi''_{x^2} = 0,$$

linéaires par rapport à chacun des x, mais beaucoup trop particulières à mon point de vue.

Explication (sommaire) de texte: On ne sait pas trop ce que Fréchet a pu comprendre de ce que Lévy lui a écrit, car ce dernier s'est passablement emmêlé dans ses notations. Nous avons heureusement une rédaction finale de ce passage dans les Leçons d'Analyse Fonctionnelle, qui permet de reconstituer à peu près ce qu'il avait en tête.
Lévy considère des fonctionnelles Φ qui admettent une variation première du type de Volterra, c'est à dire

$$\Phi(x+\delta x)-\Phi(x)\approx\int_0^1\Phi'_x\ |[x;t]|\ \delta x(t)dt$$

et une variation seconde

$$\Phi'_x\ |[x+\delta x;t]|\ -\Phi'_x\ |[x;t]|\approx\Phi''_{x^2}\ |[x;t]|\ \delta x(t)+\int_0^1\Phi''_{xx_1}\ |[x;t]|\ (s)\delta x(s)ds.$$

Il suppose en outre que toutes ces fonctions sont continues.
Il considère alors x_n une fonction constante sur les intervalles $[\frac{i-1}{n},\frac{i}{n}]$. On se demande quelle erreur on commet en remplaçant x_n par la fonction qui vaut $x_n\pm a$ sur $[\frac{i-1}{n},\frac{i}{n}]$; si elle est de l'ordre de $\frac{\varepsilon}{n}$, on peut le faire successivement en rajoutant à chaque fois un intervalle de longueur $\frac{1}{n}$, et obtenir que $\Phi(x)$ et $\Phi(x_n)$ sont proches de moins de ε, où x est la fonction où on a changé x_n sur chaque intervalle.
Fixons un $1\leq i\leq n$. On note ξ la fonction $\xi=x_n\pm a$ sur $[\frac{i-1}{n},\frac{i}{n}]$ et x_n ailleurs, et $\xi_\lambda=x_n\pm\lambda a$ sur $[\frac{i-1}{n},\frac{i}{n}]$ et x_n ailleurs.
Soit $\varphi(\lambda)=\Phi(\xi_\lambda)$.

$$\Phi(\xi)-\Phi(x_n)=\varphi(1)-\varphi(0)=\int_0^1\varphi'(\lambda)d\lambda=\int_0^1 d\lambda\int_{\frac{i-1}{n}}^{\frac{i}{n}}\Phi'_x\ |[\xi_\lambda;t]|\ \frac{d\xi_\lambda}{d\lambda}dt$$

$$=\int_{\frac{i-1}{n}}^{\frac{i}{n}}dt\int_0^1 d\lambda\frac{d\xi_\lambda}{d\lambda}\Phi'_x\ |[\xi_\lambda;t]|$$

qu'on peut remplacer, en choisissant deux points t_i et t'_i dans chacun des demi-intervalles par

$$\approx\frac{a}{2n}[\int_0^1 d\lambda\Phi'_x\ |[\xi_\lambda;t_i]|-\int_0^1 d\lambda\Phi'_x\ |[\xi_\lambda;t'_i]|]$$

$$\approx\frac{a}{2n}\left[\sum d\lambda\left\{(\Phi'_x\ |[\xi_{\lambda+d\lambda};t_i]|\ -\Phi'_x\ |[\xi_\lambda;t_i]|)-(\Phi'_x\ |[\xi_{\lambda+d\lambda};t'_i]|\ -\Phi'_x\ |[\xi_\lambda;t'_i]|)\right\}\right]$$

$$\approx\frac{a}{n}\left[\sum d\lambda\left(\Phi''_{x^2}|[\xi_\lambda;t_i]\frac{d\xi_\lambda}{d\lambda}(t_i)+\int_{\frac{i-1}{n}}^{\frac{i}{n}}\Phi''_{xx_1}[\xi_\lambda;t_i](s)\frac{d\xi_\lambda}{d\lambda}(s)ds\right)-\sum d\lambda\left(\Phi''_{x^2}[\xi_\lambda;t'_i]\frac{d\xi_\lambda}{d\lambda}(t_i)+\int_{\frac{i-1}{n}}^{\frac{i}{n}}\Phi''_{xx_1}|[\xi_\lambda;t'_i](s)\frac{d\xi_\lambda}{d\lambda}(s)ds\right)\right]$$

Pour a petit, ξ_λ est proche de x_n, donc, en vertu de la continuité, $\Phi''_{x^2}\ |[\xi_\lambda;t'_i]|\approx\Phi''_{x^2}\ |[x_n;t'_i]|$, et donc le terme estimé est encore de l'ordre de

$$\approx\frac{2a^2}{n}\Phi''_{x^2}\ |[x_n;t_i]|+\text{terme de l'ordre de }\frac{1}{n^2}.$$

Une somme de n quantités de ce type ne peut donc être rendue arbitrairement petite que si $\Phi''_{x^2}\ |[x_n;t]|=0$.

3.3 1925-1928: Loi des erreurs

La découverte des probabilités par Lévy s'est faite, comme on l'a dit, à travers quelques heures d'enseignement qu'il donne à l'Ecole Polytechnique à partir de 1919. Lors de ces leçons, il doit en particulier s'occuper de l'important sujet de la loi des erreurs. Il donne dès 1920 plusieurs justifications du théorème central limit, notamment par l'utilisation des transformées de Fourier (fonctions caractéristiques). Il commence aussi à étudier les lois stables (par rapport à l'addition) et montre que les seules lois stables de variance finie sont les lois normales.

L'obtention d'axiomes satisfaisants pour justifier la loi des erreurs a été l'obsession de nombreux mathématiciens depuis Laplace et Gauss. On peut en particulier citer Bienaymé et Tchebitcheff. Le regain d'intérêt pour une justification pratique du mode de prise en compte des erreurs de mesure avait dans les années 1920 des causes diverses, allant des expériences mises en place pour la vérification des nouvelles théories physiques à des arrières pensées liées au rôle qu'elles avaient eu dans les tirs d'artillerie de la guerre qui s'étaient achevée peu avant. L'astronome soviétique M.F.Subbotin, de Novocherkassk, intéressé à justifier les choix de corrections dans les mesures d'observation des étoiles, avait en 1923 (Mat.Sb.,31,296-301) proposé deux axiomes auxquelle devaient satisfaire les lois d'erreurs possibles.

L'axiome I dit que la probabilité d'une erreur (c'est à dire, en termes modernes, la densité de la loi) dépend seulement de la grandeur de cette erreur et on la note $\varphi(\varepsilon)$. Il faut comprendre qu'on se place en référence à une série de mesures d'une quantité physique: la probabilité que l'erreur soit comprise entre ε et $\varepsilon + d\varepsilon$ est donnée par $\varphi(\varepsilon)d\varepsilon$ où φ est une certaine *fonction d'erreur* associée à la mesure qu'on effectue. Subbotin ajoute aussi une exigence technique, qui soulèvera la critique de Fréchet: il demande que φ soit à dérivée première continue *en général* (sic).

Le deuxième axiome est un principe d'homogénéité, qui était déjà chez Schiaparelli; il demande que la valeur la plus probable de la quantité x dont on a observé des mesures $l_1, l_2, \ldots, l_n$, c'est à dire le point y qui réalise le maximum de $\varphi(y - l_1) \ldots \varphi(y - l_n)$, ne dépende pas de l'unité de mesure envisagée. C'est sur l'interprétation de cet axiome qu'un différent va naître entre les deux mathématiciens.

L'interprétation de Lévy est la plus probabiliste. Pour lui, cet axiome signifie que si on note $F(l_1, \ldots, l_n)$ le point y qui réalise le maximum de $\varphi(y - l_1) \ldots \varphi(y - l_n)$, pour tout $\alpha > 0$, $\alpha.F(l_1, \ldots, l_n)$ réalise le maximum de $\varphi(y - \alpha l_1) \ldots \varphi(y - \alpha l_n)$.

Subbotin a, dans son premier article, la même interprétation de l'axiome II que Lévy. Cependant, il fait dans son étude une maladresse en supposant implicitement que φ est à dérivée continue partout (pour avoir une condition de maximum élémentaire), et arrive ainsi à déduire comme forme nécessaire de la loi d'erreur $\varphi_m(\varepsilon) = \frac{mh}{2\Gamma(m^{-1})} e^{-h^m|\varepsilon|^m}$ où $m > 0$ et h est fixée pour qu'il s'agisse d'une densité de probabilités. En 1924 (Мат.Сб., 32, 5-8), Fréchet relève avec justesse ce renforcement de l'hypothèse, mais se trompe (en interpréntant différemment l'axiome II) en proposant un contre-exemple. De manière étonnante, il semble avoir alors embrouillé Subbotin, qui dans une nouvelle Note (CRAS Paris (180,1925, 1716-1719)) accepte les conclusions de Fréchet et montre qu'en fait la forme générale d'une fonction satisfaisant ses deux axiomes s'écrit sous la forme d'un

produit convergent $A.\prod_{r\geq 1}[\varphi_{m_r}(u)]^{k_r}$ qui peut être à dérivée non continue en 0.

A la réception de l'article de Fréchet, Lévy signale son désaccord et la discussion s'engage sur l'axiome II. La façon dont Fréchet le comprend est en fait assez alambiquée, et en même temps plus évidente: il demande que si les mesures $l_1, \ldots, l_n$ ont conduit à prendre x, les mesures $\alpha l_1, \ldots, \alpha l_n$ amenées avec la loi d'erreur φ_1 déduite de φ par multiplication par α (donc de densité $\frac{1}{\alpha}\varphi(\frac{1}{\alpha}x)$) conduisent à prendre αx. Lévy lui signale évidemment que ce résultat est valable quelle que soit la loi considérée, et donc qu'interprété ainsi, l'axiome II est sans intérêt. Fréchet finira par se ranger à cet argument et publiera une mise au point où il ne rend d'ailleurs que partiellement les armes, mais reconnaît publiquement que son interprétation de l'axiome II n'amène aucune conclusion utile.

Paris- 9 rue Chernoviz
le 29 octobre 1925
Mon cher collègue,
C'est décidément bien difficile de s'entendre par lettre, et c'est d'ailleurs de ma faute, car je ne me suis encore pas suffisamment expliqué dans mes dernières lettres.
(...)
Prenons donc pour éviter toute difficulté sur ce point, une origine bien déterminée, et deux systèmes d'unités différentes. La même loi se trouve avoir deux expressions analytiques différentes; le même résultat se trouve exprimé par deux nombres différents. Mais c'est tout de même la même loi et le même résultat; il faut bien que la même loi donne le même résultat. Je considère que ce n'est rien qu'une tautologie de l'affirmer, et que les calculs que l'on peut faire pour le vérifier prouvent seulement que le calculateur a su ne pas s'embrouiller dans ses notations; pour ma part il me faut un peu d'attention pour suivre un calcul, si simple soit-il, et il n'ajoute rien à une vérité qui s'impose à mon entendement comme un axiome sans lequel aucun raisonnement ne serait possible; c'est l'axiome qui consiste à dire que ce dont je parle signifie quelque chose, et que la vérité est la même dans toutes les langues.
Ce qui suit maintenant est une question psychologique, et l'expérience peut me donner un démenti. Je crois qu'en entendant dire "Le résultat ne doit pas changer par un changement d'unité" 9 personnes sur 10 comprendront qu'on n'a pas voulu énoncer une simple tautologie, mais un principe dont on puisse tirer quelque chose. Il faut donc comprendre qu'il s'agit de deux lois différentes, réductibles l'une à l'autre par un changement d'unité (de sorte qu'elles auront la même expression analytique avec deux unités différentes convenablement choisies, ou deux expressions distinctes avec la même unité ; je comprends qu'avec ces deux lois la valeur probable (ou la plus probable) déduite de n mesures est la même. C'est là une propriété qui n'est pas vraie en général mais qui peut l'être pour certaines lois, et le problème de la détermination de ces lois peut être intéressant.
Sommes-nous enfin d'accord ? Sinon il nous faudra aller à Locarno, ou à la Haye, conférence verbale ou arbitrage. Cinq minutes de conversation valent mieux que plusieurs lettres.
Bien cordialement à vous.
P.Lévy

3.4 1936-1939: fondement des probabilités, arithmétique des lois de probabilités

Comme on l'a déjà dit, Lévy n'a jamais vu véritablement la nécessité d'axiomatiser le calcul des probabilités. Fréchet, par contre, a été sensible au fait de préciser les notions utilisées. Ses propres travaux pour faire sortir les notions de mesure et d'intégrale des cadres topologiques, qui furent repris avec enthousiasme et infiniment développés par l'école polonaise, eurent une influence certaine sur les *Grundbegriffe* de Kolmogorov qui mentionne explicitement en ce sens Fréchet, avec qui il avait passé de nombreuses heures lors de son séjour français de 1930.

Fréchet et Lévy découvrent tardivement, semble-t-il, les *Grundbegriffe*, à la fin 1935, comme en témoigne la lettre suivante. Elle offre un exemple intéressant des difficultés de l'"ancienne" génération des probabilistes, à laquelle appartenaient Lévy et Fréchet, pour l'utilisation de concepts abstraits de théorie de la mesure comme le principe des classes monotones pour la démonstration de la loi 0-1.

Paris le 29/1/36

Mon cher Collègue,

Je vous confirme que dans le théorème de Kolmogoroff, p. 60 de ses Grunbegriffe, la condition d'indépendance des x_n mentionnée au bas de la page est absolument inutile.

La démonstration est très simple, et correcte. Il faut arriver à se dégager de l'impression que c'est un tour de passe-passe. Elle utilise bien la notion essentielle qui est la suivante: la probabilité de la suite illimitée des x_n ne peut être considérée comme bien définie que dans le cas où elle apparaît comme la limite (au sens de la convergence en probabilité) de la probabilité d'une propriété de l'ensemble des n premières variables- qui alors, si elle est réalisée pour n très grand, entraîne avec une probabilité voisine de l'unité , la propriété étudiée. La conséquence cherchée est immédiate.

Ma démonstration dégage mieux ces idées , je crois; mais on les sent implicitement chez K.

- Avez-vous remarqué dans K, p.55, un résultat résolvant un problème très voisin de celui qui nous occupait? Mais il s'agit de convergence en probabilité, et non de convergence presque sûre, de $\frac{1}{n}(x_1 + x_2 + \ldots + x_n)$. Pour la convergence presque sûre, il ne fait que rappeler p.59 le résultat de sa note de 1930.

Bien cordialement à vous .

P.Lévy.

C'est dans un article de 1934 au Bulletin des Sciences Mathématiques que Lévy a énoncé la loi 0-1. Lévy se contente en fait de démontrer le résultat dans le cas où les variables X_n suivent une loi uniforme sur [0,1]. Dans ce cas, le tirage d'une réalisation de la suite $(X_n)_{n\geq 0}$ peut être conçu comme celui d'un point dans un cube de côté 1 à une infinité de dimensions, la loi de probabilité étant donnée par la mesure de Lebesgue. L'argument de Lévy repose alors sur une observation qu'il affirme comme évidente et qui est équivalente de fait au résultat de classe monotone: il fait remarquer (en termes modernes) que pour tout événement A de la tribu $\mathcal{F} = \sigma(X_0, X_1, \ldots, X_n, \ldots)$ (qui s'écrit donc sous la forme $[(X_0, X_1, \ldots, X_n, \ldots) \in B]$, où B est un ensemble mesurable de $\mathbb{R}^{\mathbb{N}}$), et pour tout $\varepsilon > 0$,

on peut trouver $n > 0$ et $D_n \in \mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$ tels que $P(D_n \Delta A) < \varepsilon$. En fait, sa justification est de dire que les ensembles mesurables dans le cube infini-dimensionnels sont obtenus par *les constructions de M.Lebesgue* à partir des "intervalles" du cube, qui sont les ensembles du type $]a_0, b_0[\times]a_1, b_1[\times \ldots \times]a_n, b_n[\times[0,1] \times [0,1] \ldots$, de la même façon que les Boréliens de $\mathbb{R}$ à partir des intervalles ouverts réels.

3.5 1939-1945: Mouvement brownien

Comme on l'a vu, à partir de 1941 Lévy va progressivement entrer dans la clandestinité et ceci, paradoxalement, va être l'occasion pour lui de se concentrer sur des recherches très profondes autour du mouvement brownien, recherches entreprises à la veille de la guerre et qui avaient donné lieu à un gros article publié en 1940 dans l'*American Journal of Mathematics.* Ce célèbre papier sera d'ailleurs pour de nombreuses personnes, et en particulier Itô, un viatique pour leurs études futures.
Dans la première partie de cette période, Lévy est occupé principalement par la définition et l'étude des propriétés de son intégrale stochastique par rapport au mouvement brownien. Je renvoie à la thèse de B.Locker pour un examen détaillé de ces questions. A partir de 1942, il va se replonger dans l'étude des propriétés fines de la trajectoire brownienne, et en particulier de la distribution des zéros sur la droite réelle. La lettre suivante offre un beau témoignage de ces recherches qui aboutiront à la publication du livre de 1948 *Processus stochastiques et mouvement brownien.*
Montbonnot - le 1er juin 1943
Mon cher Collègue,
Je reviens sur le P.S. de ma lettre de dimanche. Le résultat obtenu me parait assez important et, comme je ne peux pas présenter en ce moment de Note à l'Académie, je vous serais obligé de conserver la présente lettre.
Il s'agit de la fonction aléatoire du mouvement brownien, $X(t)$, déjà étudiée dans mon livre sur les variables aléatoires et dans mes deux mémoires de 1939, ainsi que par divers savants étrangers J'appelle E l'ensemble des racines situées dans un intervalle (a, b) fini ou ayant une limite infinie; on suppose essentiellement que, pour la limite finie, ou pour les limites finies, $X(t)$ soit nul; de toute façon a et b sont presque sûrement des points d'accumulation des racines de $X(t)$.
Dans ces conditions
1° La nature stochastique de E est invariante par le changement de t en $t+h$, a et b étant bien entendu changés en $a+h$ et $b+h$
2° La nature stochastique de E est aussi invariante par le changement de t en λt (a et b étant changés en λa et λb)
3° Elle est aussi, si $a = 0$, $b = \infty$, invariante par le changement de t en k^2/t. C'est une conséquence évidente du fait que la corrélation entre $\frac{X(t)}{\sqrt{t}}$ et $\frac{X(t')}{\sqrt{t'}}$ est absolument symétrique (en posant $t = e^u$, $X(t) = \sqrt{t}\varphi(u), \varphi(u)$ dépend d'un processus, du type que Ville étudie actuellement, et qui est invariant par le changement de u en $-u$, Cf le théorème 2 de mon mémoire de l'A.J.M.
Ce dernier résultat entraîne la conséquence suivante : si on suppose connue une racine t_0 de

$X(t)$, donc une racine $t'_0 = k^2/t_0$ de $Y(t') = kX(t)/t$ ($t' = k^2/t$), la nature stochastique de l'ensemble des racines de $X(t)$ entre 0 et t'_0 se déduit par $t' = k^2/t$ de celle de l'ensemble des racines de $X(t)$ entre t_0 et l'infini. Or $Y(t')$ dépend précisément du processus stochastique du mouvement brownien, à cela près qu'on suppose $Y(t'_0) = 0$. (En termes précis, cela signifie qu'on considère les éventualités à probabilité positive : il y a une racine de $X(t)$ entre t_0 et $t_0 + dt_0$; on étudie dans ces conditions les probabilités conditionnelles, et on passe à la limite. J'ai souvent expliqué que lorsqu'on supposait réalisé un événement dont la probabilité a priori était nulle, c'est de cette manière qu'il faut définir la probabilité conditionnelle).
On voit ainsi qu'on peut ramener l'intervalle infini (t_0, ∞) à l'intervalle fini $(0, t'_0)$, l'origine étant d'ailleurs n'importe quel point intérieur à (t_0, ∞).
Finalement, la nature stochastique de E étant invariante par les substitutions qui servent de base au groupe des homographies, on voit que
Théorème: - La nature stochastique de E est invariante par n'importe quelle homographie.
J'ai découvert ce théorème, comme le prouve le P.S. de ma lettre de dimanche, en partant d'une propriété individuelle des intervalles e dont la réunion constitue l'ensemble complémentaire de E. Je m'étais proposé, en supposant connues deux racines a et b de $X(t)$ de chercher la probabilité que deux nombres t_0 et t_1 de l'intervalle (a, b) appartiennent à un même intervalle e, c'est à dire qu'ils ne soient séparés par aucune racine de $X(t)$. J'ai trouvé qu'elle ne dépend que du rapport anharmonique des nombres a, b, t_0 et t_1. Comme la nature stochastique des e détermine celle de E, le théorème ci-dessus en résulte. Il m'a d'ailleurs surpris. Mais je viens d'en trouver l'explication ci-dessus qui le rattache aux propriétés générales du mouvement Brownien.
Quoique nous soyons souvent en désaccord sur la valeur de ce j'appelle des considérations intuitives, je pense que vous trouvez comme moi que cette explication est plus intéressante qu'une vérification par le calcul.
Bien cordialement à vous.
P.L.
P.S Je pense aussi communiquer ce résultat à Ville.
Mon théorème entraîne pour l'arc de courbe du mouvement brownien plan limité à un point double des conséquences presque évidentes mais un peu longues à exposer. Je ne veux pas vous ennuyer davantage avec ces questions.

On trouve donc dans cette lettre un bon exemple des aller-et-retours de l'intuition géométrique de la pensée de Lévy entre des questions assez particulières (l'estimation d'une loi conditionnelle sur une excursion du mouvement brownien) et des propriétés générales de la trajectoire brownienne. Dans une première étape, il a calculé que pour $t_0 < t < u < t_1$, la probabilité que X ne s'annule pas entre t et u sachant que $X(t_0) = X(t_1) = 0$ est donnée par $\frac{2}{\pi} \arccos \sqrt{\frac{(t_1 - t_0)(u - t)}{(u - t_0)(t_1 - t)}}$. Elle est donc invariante par une transformation homographique h du temps et on peut donc mettre en correspondance l'ensemble E des zéros et $h(E)$. Lévy dit s'être étonné d'une propriété aussi générale, qui ne faisait pas intervenir spécifiquement la loi conditionnelle précédente, et a pensé qu'il devait pouvoir la rattacher à des propriétés

générales du mouvement brownien. Or, effectivement, l'invariance par translation, *scaling* et inversion du temps qu'il explicite alors lui permet de remplir son programme. A noter que dans le livre de 1948, les choses seront remises dans l'ordre naturel: l'invariance de E est obtenue dès la page 22 comme conséquence des propriétés mentionnées sur les transformations du temps, alors que le calcul de la probabilité conditionnelle est renvoyé dans la partie sur l'étude approfondie à la page 219.

3.6 1951-1956: Elections à l'Académie des Sciences

En 1951, Elie Cartan vient de mourir et la première élection à la Section de Géométrie depuis 1942 (au fauteuil de Lebesgue) va se tenir. Les candidatures de Fréchet et de Lévy à l'Académie des Sciences vont être l'occasion de tracas divers. En fait, elles constituent le seul sujet extra-mathématique traité avec quelque longueur dans la correspondance. L'histoire est assez compliquée: s'y mèlent la volonté de Julia de promouvoir Garnier, spécialiste de géométrie algébrique plutôt que Lévy ou Fréchet, une tentative stratégique de Lévy de passer par la voie des Académiciens libres, avant de revenir à celle de la section de Géométrie, et un vague malaise lié à l'impossibilité où Lévy s'était trouvé pour se présenter en 1942, à un moment où il avait des chances assez importantes de succès. Fréchet et Lévy vont essayer de se soutenir mutuellement pour contrer la campagne semble-t-il assez virulente faite en faveur de Garnier, mais sans trop de succès. L'élection de 1952 va être tendue, et Garnier sera finalement élu. Lévy est tombé gravement malade et a été opéré en juin 1952 pendant que Fréchet se voir offrir comme lot de consolation une place de correspondant. Comme on sait, les deux élections suivantes, aux fauteuils de Borel puis Hadamard seront finalement pour Fréchet et Lévy.

Paris - 4-7-51

Mon cher collègue,

Il faut qu'à deux jours d'intervalle je revienne sur ma dernière lettre.

J'ai causé longuement avec Julia, et je me suis décidé à lui rappeler sa sortie de 1936, qu'il ne se rappelait pas. Mais il m'a rappelé qu'en 1920, quand j'ai été nommé à l'X, Humbert lui avait dit "présentez-vous ; vous aurez la 2ième ligne et cela vous fera connaître"; qu'il n'avait pas fait campagne, pas songé à la 1ère ligne, et que dans ces conditions il était absurde de croire je lui en voulais.

De plus, il est très entier et combattif; il m'a parlé avec enthousiasme des travaux de Garnier (qui a d'ailleurs incontestablement abordé et résolu des problèmes difficiles). Il m'a été difficile de l'interrompre pour lui parler des miens. Je me suis peut-être trompé en croyant qu'il m'attaque pour faire le jeu de Garnier. Personne ne parlera pour vous ni pour moi avec autant d'éloquence que lui pour Garnier.

Enfin, tant à cause de cela qu'à cause de la confiance manifestée par un autre ami de Garnier, qui m'a très gentiment conseillé de faire demi-tour pour ne pas aller au devant d'un échec très probable, et cela après avoir consulté "un bon nombre d'électeurs", me fait me dire que nous n'arriverons pas à donner l'impression que le jeu est entre vous et moi. Naturellement si vous arrivez à donner cette impression, ce sera tant mieux. Mais j'en doute, et je me rends compte que les 20 amis que je croyais sûrs ne le sont pas tous.

Il ne reste donc qu'un tout petit point de ma lettre précédente (et un autre, qu'il ne faut pas perdre de temps). Il est certain qu'il y a des gens qui disent que vous êtes plus philosophe que mathématicien ; il vaut mieux que vous le sachiez.
Bien cordialement à vous
P.Lévy