

Sources and Studies in the History of Mathematics  
and Physical Sciences

Marc Barbut  
Bernard Locker  
Laurent Mazliak

# Paul Lévy and Maurice Fréchet

50 Years of Correspondence in 107  
Letters

 Springer

Paul Lévy and Maurice Fréchet

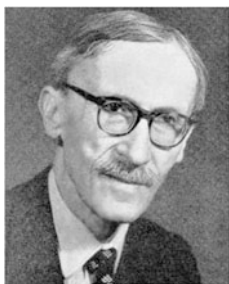
# Sources and Studies in the History of Mathematics and Physical Sciences

For further volumes:  
[www.springer.com/series/4142](http://www.springer.com/series/4142)

Marc Barbut • Bernard Locker • Laurent Mazliak

# Paul Lévy and Maurice Fréchet

50 Years of Correspondence in 107 Letters



Translation of the introductory sections and most of the notes by Peter Kleban, Professor Emeritus LASST & Department of Physics and Astronomy University of Maine

Marc Barbut (*deceased*)  
Professor Barbut wrote this while at EHESS  
Paris, France

Bernard Locker  
UFR de Mathématiques  
Université René Descartes  
Paris, France

Laurent Mazliak  
Laboratoire de Probabilités et Modèles  
Aléatoires  
Université Pierre et Marie Curie  
Paris, France

ISSN 2196-8810

Sources and Studies in the History of Mathematics and Physical Sciences

ISBN 978-1-4471-5618-5

DOI 10.1007/978-1-4471-5619-2

Springer London Heidelberg New York Dordrecht

ISSN 2196-8829 (electronic)

Sources and Studies in the History of Mathematics and Physical Sciences

ISBN 978-1-4471-5619-2 (eBook)

Library of Congress Control Number: 2013955594

Mathematics Subject Classification: 01A60, 60-03, 60J65, 31-03

Based on a translation from the French language edition:

'Paul Lévy - Maurice Fréchet: 50 ans de correspondance mathématique'

edited by Marc Barbut, Bernard Locker and Laurent Mazliak

Copyright © Hermann, Paris, 2004

All Rights Reserved

© Springer-Verlag London 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

*Ten years after the publication of the first edition in French, and though one of us is now among us only in spirit, it is still three generations who are glad to dedicate this book to our friend, colleague and master*

**Bernard Bru**

*for his seventieth birthday, whose energy and immense culture had always been for us a priceless support and whose work has contributed so essentially to the dissemination of the history of mathematics.*

*Marc Barbut †, Bernard Locker and  
Laurent Mazliak*

**In memoriam**

*Since the French version of this book was published, our colleague and friend Marc Barbut passed away in December 2011. His enthusiasm, his endless culture, his passion for the work of the two protagonists (he had met both of them in his youth), were essential ingredients for the completion of this work. We are glad to seize this opportunity to recall his good memory and also to offer a book he appreciated a lot to a larger audience.*

*Bernard Locker and Laurent Mazliak  
Paris, September 2013*

## Preface by *Kai Lai Chung*

Paul Lévy's books were not easy of access in Kunming, China, during the war years at the Southwest Associated Universities, owing to Japanese bombing of the city.<sup>1</sup> I arrived at Princeton in December 1945 but it was later that I saw Captain Gil Hunt poring over *l'Addition*.<sup>2</sup> He told me that von Neumann said: reading Lévy, one soon realizes that the author thinks mathematics in a different way from *nous autres*.

I first saw Lévy at Berkeley, summer of 1950, when he was giving a course which I attended. One day I went to his office and asked him about a result for Markov chains. At once he indicated a counter-example. It must be then or soon after that I mentioned the conjecture regarding the positivity of a transition function  $p_{ij}(t)$ : either it vanishes for all  $t$  or it never does. In his epoch-making paper<sup>3</sup> he began with this occasion. He calls a state  $i$  "stable" if  $-p'_{ii}(0) < \infty$ , "instantaneous" otherwise. A basic property of a stable state he states, is that the time-set in which a stable state occurs is (almost surely) a sequence of disjoint intervals of which the cardinality is finite up to any finite time. This must be obvious to him but missed by Monsieur Doob who published two important papers on denumerable Markov chains. The proof actually needs a little "measure-theoretic" fiddling which I presented at Feller's seminar later. The proof in Lévy's paper of the truth of the conjecture above is valid if all states are stable, as confirmed in the doctorate thesis of R.V. Chacon (at Syracuse University). It was the recently deceased Donald G. Austin who gave a beautiful complete proof of the conjecture by so-called "path-analysis."<sup>4</sup> It was like Lévy that he announced his oversight in a Comptes Rendus note soon after he recognized it. He wrote me to say that he liked Austin's proof.

Turning the clock back to 1947-9, when I was at Cornell, the discovery of Lévy's long paper,<sup>5</sup> not available during the war, caused a sensation with Feller and Kac.

---

<sup>1</sup>See Sem. Stoch. Process. (1988).

<sup>2</sup>The book Lévy (1937).

<sup>3</sup>Ann. Sci. Ec. Norm. Super. 68:327-381 (1951).

<sup>4</sup>See Chung: *Markov chains*, Springer (1960 & 1967).

<sup>5</sup>Compos. Math. 7:283-339 (1939).

Among the numerous new results, the “arcsin” law together with its complement “uniform law” for Brownian Movement was an unheralded event. Feller and I wrote a note transferring both results to the classical setting of coin-tossing, heads or tails with fifty-fifty chance. The resulting statistical interpretation is so contrary to common-sense “intuition” that had Chevalier de Méré been at the gaming table he would certainly have rushed to consult Fermat or Pascal. Indeed the news was even reported in a popular science bulletin, of which Feller kidded me as being the informer—I was not. Lévy’s article contains such a wealth of ideas and methods which remain to be explored to this day.

Probably the best known of Lévy’s multifarious achievements is the family of “infinitely divisible laws” with numerous ramifications and “applications” including the “normal” and “stable” laws (another kind of stability, both named by the discoverer). The general formula for their Fourier transform (ch.f.) is sometimes cited as “Khinchine-Lévy,” in silly English alphabetical order.<sup>6</sup> In a letter Lévy said K. completed it in the unsymmetric Cauchy case (only). In the book by Gnedenko and Kolmogorov, it was properly named. What is not well known is that in Lévy’s hands, it is just the analytic portion of his profound study of the paths of “Lévy process(es),” formerly called “additive,” “differential” etc. I must refer the really interested reader to my Review of Doob’s book *Stochastic Processes*<sup>7</sup> for further comments. Let me add that if Lévy’s other renowned publications are not even mentioned here, it is because of my lack of personal knowledge of them, as well as the lack of SPACE and TIME.

The celebrated Theorem on the convolution of two Normal laws was conjectured by Lévy who actually assumed its truth in his early work on the infinitely divisible family. It was Harald Cramér who proved Lévy’s daring conjecture, using properties of entire function in complex analysis. His Cambridge tract, lent to me by Prof. Pao Lu Hsu in tattered condition, was a precious guide. At the Bicentennial convocation of Princeton University in 1946, Cramér was a visiting professor who gave two successive courses on modern probability. I was assigned to be his assistant, and he approved my dissertation (1947). Prof. Pao Lu Hsu, then in USA, was also an invited delegate, but he did not come (see my Preface to his *Collected Papers*, Springer-Verlag, 1982 with a photo of Hsu, Cramér and Bartlett in Chapel Hill). It was on the bicentennial occasion that I first met Feller, Doob, Abraham Wald among many other mathematicians known to me by name before.

On May 19, 1964 Lévy wrote me a letter part of which announcing his election to *l’Académie* (in much detail), the other part about Mrs Doebelin’s passing. As response I asked him soon after to present a Note to the Comptes Rendus (my final one there). It is about a “pregnant integral equation” obtained by considering a last exit

---

<sup>6</sup>By the way, it would be interesting to show the following result without using Fourier transform: for any stable non-normal law with distribution function  $L$ , we have  $1 - L(x) \sim Cx^{-\alpha}$ , for a constant  $0 < \alpha < 2$  as  $x \rightarrow +\infty$ . The problem has made some progress. The most interesting result is due to Letta and Pratelli who obtained  $P(X > x) \leq cx^{-\alpha}$  for any symmetric stable law.

<sup>7</sup>Bull. Am. Math. Soc. (March 1954). Numerous other references to Lévy are mentioned, including his precursor Martingale.



time from a boundary for Markov chain, which I had spoken of in Budapest, Prague and elsewhere. An almost everywhere solution is easy but I was unable to show it everywhere. He rewrote it after telephoning Jacques Neveu for “information” and split it in two Notes which appeared in May, 1965. The open problem was first solved by Lennart Carleson with his fine analysis, of which I wrote a reproduction in a Chinese journal published in Dalian. Lévy used to write to me while summering in Menton, the border from Ventimiglia, where my wife and I sojourned once.

During our stay in Strasbourg in 1967, where our five-year old son went to French school and learned to speak very well, the Lévy's kindly invited us to dinner in their top-floor apartment on Ave. Théophile Gautier, just above the residence of the novelist Mauriac, we found out later. The guests included the Chevalleys and the Schwartzes (I first met Laurent at Prof. Chevalley's office in Princeton University). Monsieur Lévy offered me a little cup of Port-Salut, warning me of its *parfum*. Those were marvellous days.

After the Lévy Memorial Colloque at Palaiseau in 1987, we were *chez eux* at the Schwartzes, and Laurent told some stories of his *beau-père's* peculiar memory of what he had or had not demonstrated the day before. Mme Schwartz was most kind in giving me a photo of her father and one of Emile Borel, both reproduced in my *Elementary Probability* book with six other luminaries. Lévy had passed away in 1971. That happened to be the first passage time for my return to my homeland and to see my mother in Hangchow, Marco Polo's Qinsai. “TIME MUST HAVE A STOP.”

In 1976 I was able to dedicate a paper to Lévy's memory.<sup>8</sup> I was giving a course on Brownian Motion (sic) with a few faculty auditors (incl. S.T. Yau now in Harvard). A basic theorem in Lévy's *étude profonde*, proved by him using a “vertical window,” baffled me as too hard to do in class, but I was happy when I found a “horizontal window” to replace it, by use of a last-exit time (Theorem 1 of the mentioned paper). The lingo was due to my friend an auditor from Operation Research occurred.

The Brownian excursions being the groundwork for Lévy's fantastic invention of “local time,” I wrote with A.A. Balkema another paper entitled “Lévy's WAY to his local time” in our *Seminar on Stochastic Processes*, 1990 (Birkhäuser). The capitalized WAY is taken from Laotze and Chuangtze, pronounced like “TAO.”

My acquaintance with Maurice Fréchet is brief: I first saw him at a group walk in a Japanese garden in Tokyo, in 1960. He looked very handsome and was most cordial. Unfortunately somehow on the crowd we did not have a chance to talk. However our first contact occurred at the time when Paris fell to the Nazis. Suddenly I received a whole bunch of reprints sent by him (I wish I had kept the large envelope, but his reprints are here.) This must be his kind response to a Note on an inequality by E.J. Gumbel I sent to Borel for possible presentation to the *Académie*. Fréchet also sent a postcard when my professor Pao Lu Hsu and I sent a Note on recurrence (1946). Later I got a galley copy of Fréchet's first volume on probabilities of a finite number of events in the *Actualités* series. I had published a number of

---

<sup>8</sup>Ark. Mat. 14 (1976).

papers on the same subject in the Annals of Mathematical Statistics (1941–1945). Owing to the lack of communication during the war, Fréchet's two volumes and my papers did not seem to *rencontre*. Gumbel invited me to his home in Bronx, New-York City, soon after my arrival in USA near the end of November (1945). We met again in Tokyo, and we had a last supper together at the inn when all other guests had left. Indeed, he was the first mathematician/statistician I met in America. He was most kind to visit me in Princeton where I was a graduate student. Much later he shared an office with a doctorate student of mine at Columbia University. This proves that there is TIME recurrence in this mundane world. I almost forgot to say above that Fréchet's *Généralités* in the Borel collection was one of my first introduction to probability after Borel's treatise. Prof. Hsu had a copy of Fréchet's second volume on Markoff chains which I read. Only much later when I taught a course on general topology I learned Fréchet's pioneer work on that, and other kinds of mathematics (incl. economics, I learned recently from Profs Bru and Barbut). This world is connected in SPACE too.

At home, Stanford, California  
September 1st, 2003

Kai Lai Chung

# Introduction

Despite what the term ‘correspondence’ suggests, the letters included here are only those from Lévy to Fréchet, deposited at the archives of the Academy of Sciences as part of the substantial collection bequeathed to the Academy by the family of Fréchet. There are two reasons why the second half of this exchange does not appear. Fréchet was a lifelong archivist who conserved every document in his possession; the archives of the Academy hold many of Fréchet’s drafts and articles, scientific and otherwise, as well as his extensive correspondence. Also, Lévy made no effort to conserve his letters and documents. There is, in addition, another reason for the asymmetry, for which Lévy bears no responsibility. In 1942, while he was in the South Zone (of wartime France), Lévy’s apartment in Paris was searched and vandalized by the Gestapo and its French auxiliary. Lévy had followed the Ecole Polytechnique to Lyon, where he was removed from his position in December 1940 under the discriminatory *Statut des Juifs* (*Statutes on Jews*). He was reinstated in February 1941 on account of his exceptional service to the State. It was only in 1943, after some rather sordid negotiations between different branches of the French State, that the Ecole Polytechnique finally dealt with this annoying matter and got rid of Lévy by sending him back to the Ecole des Mines. In November 1943 the correspondence is suddenly interrupted. Lévy was in hiding, only to reappear at the Liberation. What nevertheless exists is a rich and passionate set of letters that evoke French scientific life at an important moment in the creation of the modern theory of probability. The protagonists are two mathematicians central to its development.

Maurice René Fréchet was born in 1878 in Maligny in the department of Yonne, north of Burgundy. Settled in Paris when he was still a child, Fréchet’s luck was to meet Jacques Hadamard as his mathematics teacher, who very soon recognized the exceptional qualities of his pupil and Fréchet always benefited from the support of his master. After graduation from the Ecole Normale Supérieure, he began his career as a professional mathematician in 1906, defending a brilliant thesis at the Sorbonne. The thesis introduced the concept of a metric space and developed other ideas that are at the heart of modern topology, especially function spaces and compactness. The most important result of Fréchet’s thesis, which was discovered simultaneously

by Riesz, is the theorem of the representation of continuous linear functionals on a Hilbert space.

Fréchet's first academic position was from 1910 to 1914 as Professor of Rational Mechanics at the University of Poitiers. After World War I, Fréchet became professor at the University of Strasbourg, which had recently been liberated from German occupation. It was at this time that his correspondence with Lévy began, and he undertook a change of mathematical direction, becoming increasingly preoccupied with applied topics. In particular he pursued recent developments in statistics, in which he became the most eminent French participant—see Armatte (2002). Fréchet played a leading role in 20th century mathematics and was an ardent advocate of the role of mathematics in society. At the instigation of his other master Emile Borel, he wrote many widely published popular articles, as well as technical papers, summarizing major advances in mathematics, especially on the theory of probability. His impressive breadth was in part responsible for his unjustly poor reputation among the younger generation of French mathematicians, especially the members of Bourbaki. Fréchet died in Paris in 1973.

Eight years younger than Fréchet, Paul Pierre Lévy was born in 1886 in Paris to a Jewish family which already contained several mathematicians. After exceptionally successful secondary studies, and despite having decided early on to dedicate himself to mathematics, Lévy took the somewhat unusual step of attending the Ecole Polytechnique rather than the Ecole Normale Supérieure, probably as a sign of fidelity to his father who had also been at the Polytechnique. He later entered the *Corps des Mines* and became professor at the Ecole des Mines, which left him relatively little time for research. Nevertheless, benefiting from the benevolent attention of Hadamard (who also became Fréchet's mentor), he undertook a thesis in functional analysis, which he defended in 1912. During the First World War, as part of his military service, he was assigned practical problems of defense, especially those arising from the then new technology, aviation. After the war, Hadamard helped resurrect Lévy's career by assigning him to organize the papers left behind by Gateaux, who had been killed at the front (see Sect. 6). In 1919 Lévy became *chargé de cours* at the Ecole Polytechnique. There he discovered the theory of probability, the area that he would subsequently revolutionize. His approach to mathematics was characterized by outstanding intuition, especially geometric, no doubt as a result of his education as a *polytechnician*. For close to fifteen years he was the leading French probabilist. His influence on the field reflects the orientation of Poincaré or Hadamard who worked often at the interface of mathematics and physics. This is especially apparent in Lévy's greatest achievement, the study of the mathematical properties of Brownian motion, which occupies a major part of his correspondence with Fréchet. He also introduced many fundamental techniques, especially those for dealing with sums of random variables, both independent and dependent. Lévy died in Paris in 1971.

The two mathematicians had rather different career trajectories. Fréchet began his scientific activities by studying very theoretical topics in the topology of function spaces. He gradually moved towards applied mathematics, because of his teaching duties in Strasbourg, becoming an emblematic defender of the application of mathematics, in particular of probability and statistics to various domains. By contrast

Lévy, educated at the Ecole Polytechnique and the Ecole des Mines and professor in a training school for engineers, developed in both his teaching and creative work a style of reasoning that relied heavily on physical arguments rather than the rigorous approach exclusively employed by some mathematicians. This naturally makes certain of his arguments a little difficult to follow, and the letters show that Fréchet did not hesitate to ask for explanations.

The kinship between Lévy and Fréchet evolved under the guidance of their common mentor Hadamard, colored by their desire to be seen as his best, or even favorite, disciple. Emile Borel, the most influential French mathematician of the time, also played an important role in the development of their research careers. Both his name and that of Hadamard appear throughout the correspondence. After more than a few vicissitudes, Fréchet and Lévy assumed, at an interval of eight years (but at the same respectable age of 78) the chairs previously held by their respective advisors at Paris Academy of Sciences. Fréchet was elected to Borel's chair in 1957 and Lévy to that of Hadamard in 1964.

This volume is divided in two parts. First, several sections paint the mathematical landscape in which the correspondence between Fréchet and Lévy took place. The aim is to present the framework within which the mathematical (and especially probabilistic) community evolved during this time, while also introducing certain technical topics that repeatedly appear in the letters. These topics are numbered sections and referred to as Sect. 1, Sect. 2 . . . in the text. The second part of this volume presents the 107 letters, which have been edited and annotated. The volume concludes with a general bibliography that contains the principal references in the two parts. Numerous additional references are included in footnotes.

# Acknowledgements

Ten years after the first publication in French of the present book, we would like to thank again all those who have contributed to our work by their advices and their answers to the numerous questions we addressed to them during the preparation. The revision of the transcription of the letters has been carefully realized by Paul Mazliak and Thierry Cabanal-Duvillard accepted to read carefully the commentaries. Needless to say, we benefited from Bernard Bru's endless culture in order to improve the text. As usually, the remaining errors are entirely ours.

Professor Kai Lai Chung, who had been so close to Paul Lévy during the 1950s, had accepted to write a small preface for the first French edition. He passed away in 2009, aged 92, and it is not without a great emotion that one finds in his text some echoes of his inimitable style and a testimony of some historical moments evoked in the book. Shortly after the publication of the book, he expressed his deep satisfaction that the voice of Paul Lévy could be heard again through the letters we had published. We believe he would be glad today to see these documents accessible to a larger audience.

We thank Paul Lévy's daughters, Mrs Marie-Hélène Schwartz, who deceased recently in 2013 and Mrs Denise Piron, as well as Mrs Florence Lederer, Maurice Fréchet's grand-daughter, for having encouraged us to complete our work, as well as to have provided some beautiful pictures in order to illustrate the book.

Naturally, neither the publication of the first French edition nor the present English version could have been realized without the agreement of Paris Academy of Science through the voice of the Perpetual Secretaries, and without the benevolent welcoming of Mrs Florence Greffe, head of the archival service of this institution. We are glad to thank them here warmly for their support and confidence. Let us also express our gratitude to the publishers Hermann in Paris, and Springer in London, and especially to their competent and enthusiastic staff for their help and constant support.

Last but not least, it is our pleasure to express here how much we are indebted to Peter Kleban, from the University of Maine, for having provided an enormous and so generous work of translation from the French text. We deeply believe that Paul Lévy's letters to Maurice Fréchet deserved to be more widely known by anyone

who is interested in the history of mathematics (and more generally of science) in the 20th century. Peter's work contributed a good deal to an aim for which all the readers will undoubtedly be grateful.

Paris, September 2013

The authors

# Contents

<b>Introduction to the Correspondence</b> . . . . .	1
1 Borel and Hadamard . . . . .	1
2 Probability Theory in France at the Start of the 20th Century . . .	3
3 The Russian and Soviet Schools of Probability Before 1940 . . .	6
4 Several Common Topics of Interest to Lévy and Fréchet . . . . .	13
5 Lévy's Stable Probability Distributions . . . . .	22
6 Lévy and Potential Theory: From Analysis to Probability . . . . .	27
7 Lévy and Brownian Motion . . . . .	37
8 Lévy and Stochastic Integrals . . . . .	42
<b>107 Letters from Paul Lévy to Maurice Fréchet</b> . . . . .	55
<b>References</b> . . . . .	207
<b>Index of Persons Quoted in the Letters</b> . . . . .	211



# Introduction to the Correspondence

## 1 Borel and Hadamard

The two great mathematicians, Jacques Hadamard (1865–1963) and Emile Borel (1871–1956), who were central in French mathematical life for many years, offered mathematical guidance and inspiration to both Maurice Fréchet and Paul Lévy and are frequently referred to in the correspondence. The books of Maz'ja and Shaposhnikova (1998) and Guiraldencq (1999), the article 'Borel' in the *Encyclopedia Universalis* written by Maurice Fréchet as well as the presentation of the life and work of Emile Borel in volume I of Borel (1972) contain extensive descriptions. Moreover, the exchanges between Lévy and Fréchet began just after the end of the First World War in which Borel and Hadamard had been deeply involved—consult Mazliak and Tazzioli (2009).

**Jacques Hadamard** was born in Versailles in 1865. Heir to the great tradition of 19th century analysts, he threw himself into the investigation of analytic functions, focusing on theorems of analytic continuation and the systematic study of the radius of convergence. In particular he obtained what is now known as the Hadamard three-circle theorem, which affirms that the logarithm of the maximum modulus of an analytic function on a disk centered at  $O$  with radius  $r$  is a convex function. From his work, especially that on the Riemann  $\zeta$  function, he drew important conclusions in analytic number theory. In 1896 he announced the remarkable prime number theorem (proven independently by de La Vallée-Poussin) showing that the number of primes less than a positive real number  $x$  grows asymptotically as  $\frac{x}{\ln x}$ .

Under the influence of his illustrious predecessor Poincaré, Hadamard was eager to keep contact with physics, and was a promoter of the systematic study of partial differential equations. He introduced numerous methods for their solution, in particular a general solution for the case of hyperbolic equations.

Hadamard was deeply involved in the scientific life of his time. He began a very active seminar at the Collège de France, which before the Second World War was the obligatory forum for any new results in mathematical analysis in France. In fact, Fréchet and Lévy indicate in their correspondence that this seminar was their natural meeting place. Hadamard was also politically engaged, like Borel as noted below,

but in a different spirit. A committed pacifist largely in sympathy with the goals of the Third International, Hadamard was an ardent partisan of rapprochement with the USSR. He was the guest of honor at the “Week of French science in the USSR” in 1934 (one of the first visits of a French scientist to the Soviet Union). A favorite of the regime, he returned to the Soviet Union several times.

The dramatic developments beginning in 1940 found him an old man of 75, crowned in glory but nonetheless under the threat of racist persecution because of his Jewish origin, despite the loss of two sons for France during the Great War. He felt constrained to accept an invitation to emigrate quickly to the USA in 1942. Returning to France at the Liberation, he lived another twenty years, keeping abreast of the mathematical life of the time. When Hadamard died in 1963 at the age of 98, Lévy was elected to his chair at the Academy of Science.

**Emile Borel** was born at Saint-Affrique in Aveyron in 1871. Throughout his entire life, he kept in close contact with the city of his birth. After very successful secondary-school studies, he entered the Ecole Normale Supérieure, where he immediately chose to pursue mathematics. He began by working on divergent series, for which he introduced what became known as the Borel sum. This quickly led him to a close examination of the idea of measures of sets (Hawkins 1970). In a series of remarkable works over a dozen years, he introduced the basic concepts of measure theory, which allowed Lebesgue to construct his integral and thereby revolutionize analysis. In related work, from 1905 Borel dedicated himself to the theory of probability. At the start of the 20th century he was its most important innovator, opening the way for its axiomatic formulation in terms of measure theory, which was completed in 1933 by Kolmogorov in his *Grundbegriffe der Wahrscheinlichkeitsrechnung* (*Basic Concepts of Probability Theory*) (КОЛМОГОРОВ 1933).

In addition to his research activities, Borel was very attentive to the role of the mathematician in society. He left numerous writings, both philosophical and popular, seeking to explain his ideas. He was constantly involved in political and social life, especially by way of his wife’s connections. She was the daughter of the mathematician Paul Appell, a literary figure and talented author who wrote under the pseudonym Camille Marbo (who won the Fémina prize in 1913). The First World War drew Borel to a task that soon became, for him, a calling. The Pyrrhic victory of 1918 decimated French youth, in particular students and alumni from the Ecole Normale Supérieure (Borel’s nephew among them) of which Borel felt himself the spiritual father as much as director. He resigned from this position, entered national politics and was elected deputy for Aveyron in 1924 in the so-called *Cartel des Gauches*, a union of moderate leftist parties. His friend Paul Painlevé, President of the Conseil in 1925, named him minister of the Navy. From that time, his life was divided between Paris and Saint-Affrique. He was elected Mayor of the latter several times. In 1921 he became a member of the Academy of Science, and was its President in 1934. He also drew on his political influence in the service of his discipline (Catellier and Mazliak 2012). The opening in 1928 of the Institut Henri Poincaré in Paris, which became one of the most important centers for mathematical research in France, was mainly a result of his efforts. Borel fought constantly, with considerable success, to ensure that numerous mathematicians and physicists visited the IHP, especially representatives of the Soviet school of probability, for whom the difficulty

of travel was ever greater. In fact it is mainly due to him that the French school remained in contact with the most active foreign practitioners of the discipline. During the Occupation, Borel was imprisoned several times because of his political views (Mazliak and Shafer 2011). He was thus a model of the engaged scientist. When he died in 1956 at age 85, he was succeeded at the Academy of Science by Maurice Fréchet, whose ideas were very close to his. There is further reference below to the central role played by Borel in the genesis of modern probability theory (Sect. 2).

## 2 Probability Theory in France at the Start of the 20th Century

Thanks to the work of Pierre-Simon Laplace and Denis-Siméon Poisson, probability theory was an important subject in French mathematics at the start of the 19th century. Thereafter it fell somewhat out of favor and only returned to a central position a century later. By that time new techniques, especially set theory, enabled the examination of random phenomena from a fresh perspective. This renewed interest was prompted by two exceptional mathematicians, **Henri Poincaré** (1854–1912) and especially **Emile Borel** (1871–1956).

Poincaré was educated at the Ecole Polytechnique in the second half of the 20th century. Marked by important discoveries in Brownian motion, radioactivity, and gas dynamics, this period was especially suited to his progressive realization of the links between the theory of probability and physics. The first paper in which Poincaré used the theory of probability was his famous *Sur le problème des trois corps et les équations de la mécanique*,<sup>1</sup> which appeared in 1890. Here Poincaré announced his recurrence theorem, showing that the initial conditions for which a dynamical system is not recurrent (in the sense that over time it never returns to a position close to the initial one) are a set with probability zero, probability being defined in terms of the set volume. This paved the way for the use of measure theory in probabilistic reasoning, later initiated by Borel and Lebesgue. Questions of the distribution of initial conditions of dynamical systems also play a central role in Poincaré's entry into the debate on the question of irreversibility in the article *Sur la théorie cinétique des gaz*.<sup>2</sup>

In 1896, the contents of Poincaré's famous course at the Faculté des Sciences de Paris appeared for the first time in print. A second edition was to follow just prior to the author's death in 1912. The course was construed one of mathematical physics, complementing the numerous volumes that Poincaré had written on various physical theories. (His title at the Faculté des Sciences de Paris was Professor of Calculus of Probability and Mathematical Physics.) This series of lessons treated fundamental results in the theory of probability in depth, especially questions pertinent to independence, the law of large numbers, and the error distribution. This

---

<sup>1</sup>(1890) *Acta Math.* 13:1–270.

<sup>2</sup>(1894) *Rev. Gén. Sci. Pures Appl.* 5:513–521. For more on these matters, see Von Plato (1994, Chap. 3) and Mazliak (2014).

work came to replace Joseph Bertrand's classic treatise on probability published by Gauthier-Villars in 1888, and a standard reference work in France.

Emile Borel's personality has been discussed above (see Sect. 1). Borel's route to probability theory was rather different from that of Poincaré and is studied in detail in the paper of Durand and Mazliak (2011). Borel appears to have become interested in probability only when he realized that measure theory, which he had introduced during the final years of the 19th century and which served as basis for the creation of a new integral by Lebesgue, was very useful in studying certain problems connected to randomness. The article of Poincaré on the three-body problem mentioned above is a good example. In that case measure theory provides a structure to express the rarity of initial conditions for non-recurrent trajectories.

Borel's ideas on probability first appeared in 1905 in his *Remarques sur certaines questions de probabilités*.<sup>3</sup> The goal of this paper, established at the outset, is to utilize the then-new measure theory and Lebesgue integral to justify the intuitive result of an example that Poincaré had given in his course of 1896, that choosing a rational point 'at random' from the interval  $[0, 1]$  has probability zero.

In 1909, number theory furnished Borel with material for a notable article entitled *Les probabilités dénombrables et leurs applications arithmétiques*.<sup>4</sup> This remarkable work took probability theory into a new era. Borel announced a first version of the strong law of large numbers, stipulating the almost sure convergence of the arithmetic mean of a series of independent random variables, thereby illustrating the interest of the notion of convergence in measure. Borel's most striking result illuminates the structure of the reals: with probability 1, the frequency of 1s in the binary representation of a number between 0 and 1 is equal to  $\frac{1}{2}$ . This appears to contradict the fact that one can choose "arbitrarily" the sequence of 0s and 1s that comprises the representation of a real.

Despite several spectacular papers, Borel's career in probability never attained the esteem that one might have expected. He was convinced that working on formalizing the theory was only of interest to lesser mathematicians, and hence never pushed as he might to axiomatize probability theory in the framework of measure theory. The program was instead completed, as is well known, by the Russian school (see Sect. 3). In the years following his 1905 article on countable probabilities, Borel mainly published popular articles on mathematics or papers on the philosophy of mathematics that explained his understanding of probabilistic models in applications. His celebrated book *Le Hasard (Chance)* was published in 1914, succeeded by numerous later editions.

The trauma of the Great War without doubt informed a certain disillusionment with mathematics on Borel's part. However it also seems that Borel did not have a very positive opinion of probability theory during these years. In her memoirs, the novelist Camille Marbo, Borel's wife, reported that Borel told her that he had enough of very abstract mathematics and so had *pantouflait* (taken the comfortable

---

<sup>3</sup>(1905) Bull. Soc. Math. Fr. 33:123–128.

<sup>4</sup>(1909) Rend. Circ. Mat. Palermo 27:247–270.

route), working on probability and statistics. Bru et al. (2009) notes a surprising consequence of this thinking: some of Borel's most profound results in the theory of probability are to be found buried in his popular articles or reviews. (Bru et al. (2009) is especially interested in the pamphlet from the collection *Que-Sais-Je* which includes some interesting remarks on the theory of martingales.)

In this overview of the condition of probability theory in France before the appearance of Lévy and his generation, mention should be made of a third figure, **Louis Bachelier** (1870–1946). Bachelier was for a very long time ostracized by French mathematicians, and his prophetic work was only recognized during his lifetime by foreigners, especially members of the Russian school. In 1900 at the Sorbonne Bachelier defended a very original thesis entitled *Théorie de la spéculation* (*The Theory of Speculation*). This work envisaged for the first time a random phenomena (in this case the fluctuations of the stock market) as a mathematical process in continuous time. It in fact corresponded to a model of Brownian motion by Norbert Wiener that was introduced only some twenty years later. Bachelier also laid the groundwork for the study of Markov processes in continuous time. In 1931, in his great paper<sup>5</sup> (see Sect. 3), Kolmogorov referenced Bachelier's notions on Markovian transitions as one origin of his own views.

It was Bachelier's misfortune to be too original—see Bru (2001) and Courtault and Kabanov (2002). This quality was in fact praised by Poincaré in his report on Bachelier's thesis which was laudatory, despite what has often been claimed. In addition, Bachelier 'came from nowhere,' being neither *normalien* nor *polytechnicien*. As a result he was treated with an indifference bordering on disdain. In applied mathematics today the theory of Brownian motion is widely used, and the role of Louis Bachelier as its forefather in the mathematics of finance is universally acknowledged.

Bachelier's most severe and least sympathetic critic was without doubt Lévy (Bru 1999). In fact, before he assumed his first position as professor in 1928, Lévy had read almost nothing by Bachelier, relying on the partial and incomplete reading of a single page of a 1913 publication by Bachelier, which he assumed to be mathematically incorrect. In this work, Bachelier had repeated, very briefly, certain parts of his thesis (which Lévy had never read), using several unfortunate shortcuts, in particular concerning the question of the speed of a Brownian particle. Lévy said that the "erreur de Bachelier (Bachelier's error)," which Bachelier in fact did not make in his thesis, stopped his reading. This also resulted in a distressing incident. In 1926 Bachelier, at the age of 56, was a candidate for the post left vacant by René Louis Baire in Dijon. Maurice Gevrey, who held the chair of Mechanics, asked Lévy for his opinion of the incriminating passage. Lévy responded that there was an error and the post went to Georges Cerf. Bachelier, who for 25 years had held only temporary positions, very poorly paid (or not paid at all), and had received a few grants, blamed Lévy for this new setback in his unfortunate career. He even circulated an accusatory letter to the academic community. This greatly affected Lévy, who be-

---

<sup>5</sup>Kolmogoroff, Andrei (1931) Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.* 104:415–458.

lieved that even without his contribution, Bachelier would never have obtained the Dijon post over a candidate as young and promising as Cerf.

It was only after reading Kolmogorov's article of 1931 that Lévy began to take seriously the content of Bachelier's work. Lévy's letters of September 1943 show how he was ready, when he began the first draft of certain chapters of Lévy (1948), to recognize the accomplishments and priorities of Bachelier on important questions concerning not only Brownian motion but diffusion more generally. Lévy and Bachelier were reconciled at the start of the Second World War. Beginning in 1948, two years after Bachelier's death, Lévy spoke of Brownian motion as 'la fonction déjà étudiée par Wiener et Bachelier (the function already studied by Wiener and Bachelier)'. In his autobiography (Lévy 1970), Lévy revisited his relationship with Bachelier, recognizing the unfairness of his initial judgment.

### 3 The Russian and Soviet Schools of Probability Before 1940

By the time Lévy had become deeply interested in research on the theory of probability, its center of gravity had already moved east. As of the 1920s, Soviet mathematicians used the results obtained by their immediate predecessors (especially Borel) and steered their efforts toward the investigation of probabilistic structures. In the two ensuing decades, this strategy resulted in fundamental advances by the Soviet school of probability. Because of these successes and many allusions to this school featuring in the correspondence between Lévy and Fréchet, the present section is included. The seemingly arbitrary cut-off date of 1940 is introduced only because after the Second World War, the letters between Lévy and Fréchet contain far fewer references to the Soviet school.

#### 3.1 Before the 1917 Revolution

Without question, the founder of the Russian school of probability was Pafnuty Lvovich Chebyshev (1821–1894), a professor in Saint Petersburg (where he was recruited at the urging of Buniakovsky, after failing to obtain a position in Moscow). Because of Chebyshev, Russian research in mathematics, especially in probability, was centered in Saint Petersburg for fifty years. From the beginning of his mathematical career Chebyshev was interested in probability, in particular improving the fundamental limit theorems (the law of large numbers and the central limit theorem) and specifying the assumptions under which they are applicable. The preamble to his thesis, defended in Moscow in 1844, gives an idea of his goal:

Making a belief in all these conclusions via a simple and precise analysis accessible to the majority of students, is a large step to an elementary exposition of the theory of probability.

In all his research in probability, Chebyshev sought to justify the methods used in its applications. Юшкевич (1968) explains how Chebyshev's interest in ballistics

and in mechanics in the 1860s was at the origin of his discovery of the celebrated inequality that bears his name, and the method of moments. Kolmogorov, in a report from 1947 (КОЛМОГОРОВ 1947)—which should be read with a certain filter, in as much as it was written at the height of Soviet pan-Russian nationalism, for which Chebyshev became an emblem—wrote:

[T]he principal is that Chebychev searched continually to obtain precise estimates of the deviation with the [theoretical] limiting laws, that hold for a large but still *finite*<sup>6</sup> number of trials [. . .]. In addition, Chebychev was the first to realize (and make full use of) the significance of a “random variable” and the “mathematical expectation” (average value) of a random variable. The sense of these concepts was already known and appeared as an extension of the expressions “event” and “probability,” but random variables and their mathematical expectation take on a more precise meaning via a practical and supple algorithm.

In the 1920s von Mises’ theory of collectives enabled the idea of an algorithmic construction of randomness. According to Kolmogorov this work had its origins in the Petersburg school of probability.

Chebyshev’s students Andrej Andrejevitch Markov (1856–1922) and Aleksandr Mikhailovitch Lyapunov (1857–1918) took up their professor’s torch. Chebyshev recruited Markov to the Imperial Academy of Science in 1886. Seneta (2001) justly observes that Markov, who lived until 1922, had a courageously resistant attitude to the most repressive measures of the tsarist regime. This, naturally, engendered a rather benevolent attitude toward him—or by association the school of Petrograd<sup>7</sup>—on the part of the Bolsheviks. Markov and Lyapunov continued to study the limit theorems, weakening the hypotheses on identical distributions in the central limit theorem. The first systematic development of the theory of characteristic functions, originally begun by Lagrange and Laplace, is from Lyapunov. Markov in turn, searching for a counterexample to the necessity of independence in the law of large numbers, formulated in 1907 a first model of the chains which now bear his name.<sup>8</sup> There is no doubt that on the eve of the events of 1917, Petrograd was the center of applied mathematics in Russia.

In Moscow, the other capital of the empire, the school of mathematics took on increasing importance and began to rival Saint Petersburg at the start of the 20th century. The notable members of the Moscow group were Dmitri Fedorovitch Egorov (1869–1931) and his student Nikolai Nikolaievitch Luzin (1883–1950). The latter originated a very powerful formalist movement, studying the theory of abstract function spaces along the lines of the work of the French school of Borel, Baire, Lebesgue and Fréchet. It is to this period that the seminal works applying measure

---

<sup>6</sup>Our emphasis.

<sup>7</sup>Petrograd is the name taken by Saint Petersburg from the beginning of the First World War until 1924 when it became Leningrad.

<sup>8</sup>Seneta (2001) affirms that this occurred in reaction to an article by Nekrassov, explaining that a proof of the existence of free will emerges from the everyday observation of averaging phenomenon in society (for example, the average number of births remains almost the same from one year to the next) because independence of actions is necessary to apply the law of large numbers. Markov, extremely anti-religious, was eager to disprove this necessity.



theory to the theory of functions belong. This very formal research was not without connections to profound questions in philosophy and religion.<sup>9</sup>

Luzin's mathematical education was influenced by his voyage to Göttingen, home of the formalist school headed by Hilbert. Returning to Moscow, Luzin defended his thesis in 1914. Three years later, with the help of Egorov, he founded a research group on fundamental mathematics (affectionately nicknamed "Luzitania"<sup>10</sup>), in which his students participated—including Kolmogorov, after 1922.

There are two other important figures to note in this briefly sketched mathematical landscape. One, Serguei Natanevitch Bernstein (1880–1968), played a major role in the development of modern probability. In 1898 he began a thesis at the Sorbonne on elliptic equations, which he defended in 1904. This put him at the center of activity from which emerged the modern understanding of probability. Bernstein was also an essential point of contact between the French and Russian mathematical communities. In 1911 he began to work on probability in Kharkov, where he had taught since 1907. In 1917 he made an attempt at axiomatizing its foundations (Бернштейн 1917).

Vucinich (1999) notes Bernstein's unyielding attitude after 1917 to the ideological demands of the new revolutionary regime. Bernstein's international reputation, established well before the Soviet era, permitted him to weather the blackest years of the period without great difficulty. Retaliation against him principally took the form of partial concealment, i.e. keeping him obscured behind Khinchin and Kolmogorov (see below).

Finally, as noted by Юшкевич (1968), a very significant figure who should not be overlooked is Evgeni Evgenievitch Slutsky (1880–1948). Slutsky finished his studies in 1911 in Kiev, but not without difficulty, given various political problems. In the years preceding the 1917 revolution he published several articles on mathematical methods for statistics, including a monograph (Слущкий 1912), a very modern exposition. In the 1920s he launched himself into pioneering research in stochastic calculus. These works were recognized as of primary importance by Kolmogorov and Khinchin. Slutsky's research remained little known outside the USSR although in the 1930s the great traveler Fréchet, among the most passionate to maintain contact with the USSR, drew the attention of French mathematicians to several of Slutsky's articles.<sup>11</sup>

---

<sup>9</sup>Luzin and Egorov had close ties to orthodox theological circles, notably to the extraordinary Pavel Alexandrovitch Florensky (1882–1937), Luzin's close friend, who managed both studies in mathematics at the University of Moscow and religion at the Academy of Theology. In 1912, at the Academy, he defended a thesis searching for, à la V. Soloviev, a spiritual interpretation of science. Luzin also spent some time studying theology. The correspondence between Luzin and Florensky is the object of an interesting paper (Демидов et al. 1989). On Florensky, see the beautiful book Betti (2010), and also Žust (2002) for biographical aspects.

<sup>10</sup>Wordplay on the name Luzin, probably an allusion to the Lusitania, sunk in 1916, which was a pretext for the USA to enter the First World War.

<sup>11</sup>Detailed information about Slutsky is available in Locker (2001), as well as in the 1948 eulogy to him written by Kolmogorov (Колмогоров А.Н. (1948) Е.Е. Слущкий. Успехи Математической Наук III(4):142–151).



### 3.2 *Under the New Regime*

Revolutionary torment, as well as the new, post-1917 ideology, caused Russian mathematicians of the old school to progressively modify their discipline and their research strategies. This reorientation favored the emergence of a new generation naturally more in synch with the ideology of the existing regime. As noted in Vucinich (1999), the 1920s were marked by an attempt to erect a structure able to encompass all mathematics. It is easy to imagine that the very abstract ideas of the Moscow school of Egorov and Luzin were not favorably viewed by the new regime, whereas the new ideologues looked with satisfaction on the growing influence of research groups closer to inspirational Marxist ideas that demanded (Graham 1993) the unification of theory and practice.<sup>12</sup> It is interesting to read in this regard Luzin's own description of the situation of Soviet mathematicians at the end of the 1920s, when he had to defend himself with the "energy of despair" against the raging storm aimed at him in July 1936.<sup>13</sup> In the notes of the boisterous meetings of the Commission of the Academy of Science which miraculously have been preserved, Luzin cries out:

Can I ignore the opinion of comrade Kolman, who occupies a high position and accuses me of theoretical abstraction and damage arising from my work? No, certainly not. My goal was not to damage my country. And if those who have a clear view of these matters find that a theoretical activity is damaging, well I preferred to do no more damage. But out of me came theorems. What could I do? Retreat into myself? I decided to publish here all the applied research that could help my country. And the parts of my work that have been called damaging. . . , to publish them abroad. Because I had to take account of the situation. Judge for yourselves if continuing to stubbornly publish theoretical work—that not only has no present application, but for which no future application is anticipated—wouldn't have presented an obstacle to our scientific publishing enterprise.<sup>14</sup>

Remarkably, it was three mathematicians (Колмогоров 1925; Хинчин 1926 and Бернштейн 1927) whose scientific work was principally in probability who formulated a politically correct and rather anti-idealistic view of the fundamentals and elements of mathematics. Their interpretation rose to the level of dogma during the entire Soviet period, with a practical view of probability theory serving as a prime example. Since the founders of the Russian school had developed the theory of probability in a rather practical context, they were easily able to avoid any accusation of idealism. The affirmation that the discipline was at the interface with the "real" world drew on the authority of Chebyshev, Markov and Lyapunov. As a result of this specifically applied character the theory of probability enjoyed an excep-

---

<sup>12</sup>See, for instance, the illuminating resume of the XVIIIth Congress of the Party published as an introduction to the 1939 volume of the *Известия Академии Наук СССР*.

<sup>13</sup>A violent campaign orchestrated by various political bodies accused Luzin of submission to the West, on account of his numerous publications in western journals. See Dugac (2003) and Демидов and Левшин (1999).

<sup>14</sup>See Демидов and Левшин (1999).

tional level of approval during the Soviet period<sup>15</sup> (in Collective 1948, 51 pages are devoted to this theme). Vavilov's 1946 report (Вавилов 1946) considered which fundamental scientific problems the Academy of Science should devote itself to in the next five-year plan. Probability is the first mathematical discipline mentioned, specifically because of its applied aspects. It was grouped with natural sciences and social sciences such as economics.

As in any period of fundamental change, this one was obviously marked by contradictory tendencies. It would be foolhardy to posit an overly happy outlook for the mathematics of randomness. Vucinich (1999) notes that randomness by its nature did not fit a dialectical materialistic interpretation very well, and therefore remained rather suspect in the eyes of those who from the 1920s on took charge of formulating the contours of the new ideology. Vucinich cites, for instance, a contribution to the symposium dedicated to the relation between dialectical materialism and mathematics, organized in 1931 under the auspices of the Communist Academy. This contribution regrets the interest taken by the Leningrad school in the fluctuations of random variables, which are found closer to the anarchic behavior of a capitalist economy than to the planned Soviet economy.

Contacts between Paris and Moscow gradually resumed in the aftermath of the Great War. Leading researchers like Borel and Fréchet sought to maintain relations with their Russian colleagues. The archives of the Institut Henri Poincaré point to the great efforts made by the Institute in this regard, particularly to the tenacity of Borel, who had to deal with an expanding Soviet bureaucracy. Between 1925 and 1939, Bernstein was frequently invited, but could only come to France once, in 1929. Again and again, from 1930 to 1935, Bernstein postponed his arrival, month to month, explaining that his government would not authorize his trip.<sup>16</sup> In 1935 Fréchet mentioned<sup>17</sup> that he had felt the necessity of renewing contacts with Russia as of his installation in Strasbourg in 1919, when relations between the two countries were at a very low point. He managed as best he could to establish an exchange of publications, and mentions in his Kiev speech the personal relationship that he was able to build with Nikolai Mitrofanovitch Krylov. In 1934, Hadamard was one of the first to visit the USSR, followed soon afterwards by Fréchet in 1935, as part of a long journey across eastern Europe.

### 3.3 *Several Important Steps*

The period from 1925 to 1940 witnessed effervescent mathematical creativity in the USSR. The foundations of the modern theory of probability emerged in the papers

---

<sup>15</sup>Which, most ironically, is also the reason for its lack of esteem by the French “structuralist” school (see Maashal 1999).

<sup>16</sup>The archives of the Institut Henri Poincaré also contain an enormous dossier on the physicist Frenkel, who postponed his visit from 1934 to 1937.

<sup>17</sup>Draft of a speech in Kiev, archives of the Paris Academy of Science.

of the three leading figures (Kolmogorov, Khinchin and Bernstein as well as Slutsky, as mentioned) and their impressive list of students.<sup>18</sup>

In 1925, in their only joint paper, Khinchin and Kolmogorov obtained the classical criterion for the convergence of a series of independent random variables with mean equal to zero: if  $\sum_{n \geq 1} E(X_n^2) < +\infty$ ,  $\sum_n X_n$  converges a.s. To prove this, they made use of the celebrated inequality for martingales discovered by Kolmogorov

$$P\left(\max_{1 \leq k \leq n} |S_k| \geq \varepsilon\right) \leq \frac{E(S_n^2)}{\varepsilon^2}$$

where  $S_n = \sum_{k=1}^n X_k$ . During the years that followed, a cascade of results on the law of large numbers flowed from this finding, notably the derivation of a general version of the strong law in 1930 by Kolmogorov: if the  $X_n$  have zero mean and if  $\sum_{n=1}^{\infty} \frac{E(X_n^2)}{n^2} < +\infty$  then  $\sigma_n = \frac{S_n}{n} \rightarrow 0$ , a.s.

In 1924 Khinchin had obtained a first version of the law of iterated logarithms (for Bernoulli variables) that Kolmogorov was able to generalize in 1929 by showing that under reasonable assumptions  $\lim_{n \rightarrow \infty} \frac{S_n}{\sqrt{2B_n \ln \ln B_n}} = 1$  where  $B_n = \sum_{k=1}^n E(X_k^2)$ . Here again, beyond the result itself, it is notable that the proof rests on a technique (large deviations) that itself found considerable further use.

At the beginning of the 1930s, the Soviet school of probability turned its attention to the study of stochastic processes in continuous time. Between 1930 and 1931, Kolmogorov was abroad, mainly at Göttingen in Germany, then in France, where he met Lévy and Fréchet. He stayed for a month in the latter's house in Sanary-sur-Mer in the Var, discussing Markov processes; then he spent several days in Paris where Lévy invited him to his home (Shiryayev 1989). In 1931 Kolmogorov's remarkable memoir<sup>19</sup> on Markov processes appeared. Kolmogorov gave a systematic treatment of the functional equations satisfied by the semi-group of the transitions. This pioneering opus (directly descended from Bachelier's work, as acknowledged by the Soviet school—see Sect. 2) opened the way for the use of analytic techniques for the definition of stochastic processes. Later in the 1930s, Bernstein, in particular, launched a systematic analysis of diffusion processes. Such studies underwent further very rapid development in the 1940s (due especially to Wolfgang Doeblin, Kiyoshi Itô, William Feller, Robert Fortet, John Doob and, naturally, Lévy—see Sects. 7 and 8 below, and also Mazliak 2007, 2009).

In 1933 Kolmogorov published his celebrated *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Kolmogorov 1933), in which he systematically developed the theoretical foundations of probability theory in the framework of measure theory. As a result his voyage to the West, Kolmogorov undoubtedly became convinced that Borel's idea of putting measure theory at the heart of the mathematics of probability needed adjustment in order to accommodate the theory of processes in con-

<sup>18</sup>On Kolmogorov, the reader is invited to consult (Chaumont et al. 2007).

<sup>19</sup>Kolmogoroff, Andrei (1931) Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. Math. Ann. 104:415–458.

tinuous time. In addition, the impact on Kolmogorov of the school of Göttingen, where the spirit of Hilbert and his *Grundlagen der Geometrie* (*Foundations of Geometry*) reigned, inspired the Russian to write a small booklet proposing the rigorous foundations of two essential notions: conditional probabilities and the existence of probability distributions on infinite products of measurable spaces (the famous existence theorem of Kolmogorov under the assumption of compatibility of finite dimensional distributions—see Dellacherie and Meyer (1975)). In his autobiography Lévy (1970) claims (such a bitter remark should be considered with care since Lévy was often obsessed by questions of priority) that he could have written the monograph himself, had he better appreciated its importance. Kolmogorov's approach, albeit purely axiomatic, purports to somehow be constructive (a concession, in more or less good grace, to the mathematical orientation of the Soviets). Fréchet, a believer of the objective character of probability (see Fortet (1948)), was no doubt intrigued by the delicate balance that Kolmogorov struck between a frequentist theory based on finite sets representing the results obtained by repeating a random experiment a finite number of times—the only possible interpretation in connection with real events—and the aspects requiring infinity that are a pure mathematical abstraction.<sup>20</sup> In fact Fréchet wrote to Kolmogorov on this subject and the latter confirmed to him that he conceived the formal definition of probability in terms of “very large but finite sets,” since infinite sets are only a mathematical idealization. Kolmogorov confirmed that for him “formal axiomatics should be accompanied by an analysis of its significance for reality.”<sup>21</sup>

In the 1930s, Khinchin and Kolmogorov regularly notified Lévy and Fréchet of their work.<sup>22</sup> In his book, Lévy (1970) mentions the significant correspondence that he had with Khinchin, notably on the arithmetic of probability distributions. These epistolary discussions culminated at the end of the 1930s with the Lévy-Khinchin formula describing the characteristic function of an indefinitely divisible distribution. It was found by Lévy and then by Khinchin in 1938, using a different method (see Sect. 5). Relations between Lévy and Khinchin deteriorated as a result of an article by Khinchin published at the end of the Stalinist era in a collection dedicated to philosophical problems of modern physics. In it, Khinchin castigated Western ide-

---

<sup>20</sup>See Chap. 7 of Von Plato (1994) on this question.

<sup>21</sup>Letter, August 3, 1939 (archives of the Paris Academy of Science). The very late date of this letter, from a moment when contacts between Soviet and western scientists had almost disappeared, is another sign of the benevolent attitude of the regime toward Kolmogorov.

<sup>22</sup>For example, during the winter of 1933–1934, Kolmogorov wrote to Fréchet (January 29, 1934, archives of the Paris Academy of Science) that he had obtained the criterion for a process in continuous time in the form  $E(|f(x + \Delta) - f(x)|^K) = O(|\Delta|^{1+\varepsilon})$ , which he applied to Brownian motion. Oddly, the famous criterion that bears his name was never published by Kolmogorov himself, but appears for the first time (properly attributed to him) in an article by Slutsky in 1937. Incidentally, the letter of Kolmogorov just mentioned is written in an uncharacteristically anxious style. He wrote for instance: “It requires considerable energy to surmount formal difficulties” suggesting that he was asking Borel for some official help so that he could obtain the permission to travel). Kolmogorov apparently encountered unexpected bureaucratic difficulties regarding his trip to France, a sign of the restrictions that the Soviet authorities were imposing on international travel.

alism. Lévy published a virulently critical response in *Bulletin des Sciences Mathématiques*. In his autobiography (Lévy 1970, p. 109), Lévy alluded to the incident and regretted his response, because it did not take account of the violent Stalinist campaign directed against so-called bourgeois science at the end of the 1940s, and thus may have put Khinchin at risk. Kolmogorov resumed contact with Fréchet at the start of de-Stalinization, and in 1958 was invited for a semester to the University of Paris (see Letter 78).

## 4 Several Common Topics of Interest to Lévy and Fréchet

In the correspondence included here, Lévy is mainly concerned with his own research. However among the letters are many exchanges on topics common to Lévy and Fréchet. Both of them were analysts and probabilists. Fréchet defended his thesis in 1906, five years before Lévy, then immediately joined the school of functional calculus led by his professor and mentor Hadamard. Starting in 1910, he enriched that branch of mathematics with results on abstract sets in *Les ensembles abstraits et le calcul fonctionnel*.<sup>23</sup> In the same year *Leçons sur le calcul des variations (Lessons on the Calculus of Variations)* by Hadamard, assembled and edited by Fréchet, was published. Fréchet had, for example, already clarified the idea of compactness, but his treatment was limited to sequential compactness, a limitation that later incurred the wrath of Bourbaki. Fréchet thus opened the way for modern analysis and certain aspects of topology. In the latter, he contributed at the highest level and was considered the equal of Hausdorff by the Polish school, as is mentioned in several editions of the celebrated treatise on topology by Kuratowski.

In 1918 when the Lévy-Fréchet correspondence began, Lévy, like Fréchet, had mostly worked along lines set out by Hadamard. He attacked some interesting theoretical problems in functional calculus and the calculus of variations, and he applied the obtained results to classical questions of potential theory (see Sect. 6). Hadamard had asked Lévy to organize the papers of Gateaux, who died in the war.<sup>24</sup> Reading them led Lévy to a number of questions, and he felt himself ready to write an epoch making book about functional calculus. However it took until 1922 before Lévy published *Leçons d'analyse fonctionnelle* (Lévy 1922). In 1918, Fréchet already had the ideas of neighborhood, metric space, and in 1925, after having created the corresponding mathematical vocabulary and defined the appropriate structures, he gave the definition of vector spaces, affine spaces and affine topological spaces. Thus, in a few years time, Fréchet gained a considerable lead over Lévy in these emerging areas of analysis. Hadamard approved this approach in which attention focused, at first, on specific problems involving the nature of variables, and only afterwards on the relation between functionals and variables, and finally on hypotheses concerning

---

<sup>23</sup>(1910) *Rend. Circ. Mat. Palermo* 30:1–26.

<sup>24</sup>On Gateaux, see the complete study provided in Mazliak (2011).

sets and transformations between spaces. The French school of analysis, not without considerable grouching, would also follow this path in the aftermath of Fréchet's works.

The correspondence addresses topics in analysis and probability which interested both mathematicians, despite their often very different approaches. These topics appear in the letters in different ways, recurring or sporadically. Since we only have Lévy's letters, Fréchet's contributions seem a bit veiled, despite his considerable mathematical stature. Fully describing the extensive and encyclopedic nature of Fréchet's opus would have required much more than a commentary on his correspondence with Lévy.

To finish this survey of the relevant topics, we present a few details to help introduce the letters.

#### ***4.1 Fréchet's Integral and Lévy's Averages***

In 1898, Borel gave an explicit definition of the measure of a set and measurable sets in his *Leçons sur la théorie des fonctions*. Lebesgue published *Leçons sur la théorie de l'intégration* in 1904. Hawkins' book *Lebesgue's Theory of Integration* (Hawkins 1970) and Alain Michel's book *Naissance de la théorie moderne de l'intégration*<sup>25</sup> trace the origin and history of modern integration theory. They also describe Fréchet's integral, and give an overall viewpoint on these developments. *Éléments d'histoire des mathématiques* by Bourbaki also includes the Fréchet integral in its coverage of the history of integration. In the works of Hadamard, Riesz, Fisher, Fréchet and Radon the concept of the integral, due to its linear and continuity properties, gained increasing connections with functional analysis. In 1907 Riesz found the representation of linear functionals on the space of continuous functions (on a given interval) by using Stieltjès integrals. This improved the 1903 result of Hadamard that defined these functionals as limits of integrals. Later, Riesz and Fréchet independently established the theorem representing linear functionals on  $L^2$  by Lebesgue-Stieltjès integrals. In 1915, in several publications, Fréchet made a decisive step by incorporating certain aspects of Radon's work in his theory of abstract spaces. Fréchet exploited partitions on a function space  $\mathcal{F}$  in order to define, using superior and inferior Darboux sums, the integral of a functional which he denoted  $\int_{\mathcal{F}} G(e) df(e)$ . This is a kind of Stieltjès integral for which the function  $f$  is an additive function of sets and thus a primitive version of an abstract measure. At the end of the First World War, the American works of the British mathematician Daniell appeared, similar to certain of Fréchet's ideas. Daniell constructed an integral on a set of arbitrary elements that can be points in a space of denumerably infinite dimension but found it impossible to give an example in such a space.<sup>26</sup>

---

<sup>25</sup>(1992) Vrin, Paris.

<sup>26</sup>On Daniell, see Aldrich (2007).

At almost the same time, Lévy came up with a notion of the average on a unit sphere<sup>27</sup> in  $L^2$ . This idea proved to be an important step on the way to integration in infinite dimension. These Lévy averages were first introduced, following Gateaux's suggestions and methodology, in a note.<sup>28</sup> In it, Lévy introduced a functional  $\Phi$  defined on the space of continuous functions on  $[0, 1]$ . He considered the subspace of functions that he calls simple of order  $n$ , meaning constant in each of the  $n$  equal intervals of a partition of  $[0, 1]$ . Restricted to this subspace,  $\Phi$  appears as a function  $\varphi$  of the  $n$  values  $z_1, \dots, z_n$  on these intervals. Its average  $\mu_n$  is given by an integral relation

$$\mu_n \cdot \int_{-\infty}^{+\infty} a_{n-1}(\lambda) d\lambda = \int_{-\infty}^{+\infty} \lambda a_{n-1}(\lambda) d\lambda$$

where  $a_{n-1}(\lambda)$  is the  $n - 1$  dimensional volume of the intersection of the unit sphere and the hyperplane  $z_n = \lambda$ . Lévy's idea, which had already occurred to Gateaux, is to let  $n$  go to  $\infty$ . In a situation when  $\mu_n$  goes to a value  $\mu$ , the expression  $\Phi$  has the value  $\mu$  almost everywhere. This results because the measure of the domain in which  $\varphi$  is not between  $\mu - \varepsilon$  and  $\mu + \varepsilon$  goes to 0 as  $n$  goes to  $\infty$ . In the case of averages over spheres, Lévy also notes the necessity of making use of the norm  $L^2$ .

Daniell's integral and Gateaux-Lévy's averages were the source of Wiener's 1923 construction of Brownian motion (see Sect. 7). Wiener's *Differential Space* is the first explicit example of the application of Daniell's integral to a (denumerably) infinite dimensional space. However, in reviewing theories of integration in 1930, Kolmogorov emphasized the importance of Fréchet's integral, which he retained as a source for his *Grundbegriffe* (КОЛМОГОРОВ 1933).

## 4.2 Differentials

In extending of the ideas of derivative and differential to infinite dimensional spaces, Hadamard's school of functional calculus focused on certain questions from the calculus of variations (see Sect. 6). This resulted in a fuzzy cohabitation of various definitions corresponding to different situations. Fréchet's general work on analysis helped greatly to clarify these matters.

Before the First World War, Fréchet introduced the differential of a functional  $U$  as the continuous linear functional  $\delta U$  satisfying  $U(x + \delta x) - U(x) - \delta U(\delta x) = o(r)$  where  $r$  is the distance between the functions  $x$  and  $x + \delta x$ . Simultaneously, Gateaux introduced his own "variation" (aka differential). For a functional  $U$ , this is the limit  $\delta U(\delta x)$  of the ratio  $\frac{U(x+\lambda\delta x)-U(x)}{\lambda}$  when  $\lambda$  tends to 0. Gateaux's definition does not require any metric structure on the space of functions  $x$  and is not always linear in  $\delta x$ . Lévy, inspired by Gateaux, made considerable use of the Gateaux

<sup>27</sup>On this important topic, see the many details contained in the paper of Mazliak (2011).

<sup>28</sup>Lévy, Paul (1919) Sur la notion de moyenne dans le domaine fonctionnel. CRAS (August 25, 1919).



differential in his own work. Starting in 1922 (see Lévy 1922), Lévy compared Gateaux's and Fréchet's differentials and exposed their respective use in the framework of functional calculus. It is clear that in a normed vector space differentiability in Fréchet's sense implies differentiability in Gateaux's sense. Thus the ensuing debate between Lévy and Fréchet focused mainly on the role of linear and metric structures. As a result, these constructs—today part of the foundation of analysis—imposed themselves on Lévy as the natural framework for the calculus of variations, despite his constant concern about economizing structures. In order to reconcile linearity and simplicity, Lévy proposed that  $U$  be differentiable if and only if it admits a variation in Gateaux's sense that is also linear. In a 1937 memoir, Fréchet criticized this idea of differentiability in Lévy's sense because it did not necessarily satisfy the composition formula for partial derivatives  $\frac{d}{dz}U(\phi_1(z), \dots, \phi_n(z))_{z=z_0} = \sum_{i=1}^n \partial_i U(\phi_1(z_0), \dots, \phi_n(z_0))(\frac{d\phi_i}{dz})_{z_0}$  where  $\phi_1, \dots, \phi_n$  are functions of the real variable  $z$ .

In the same memoir Fréchet defined his own differential, those of Gateaux and Lévy, and gave a fourth possibility that follows from work by Volterra and Hadamard. The latter had occasionally been used by Lévy to calculate the variation of a functional  $U$  on a family of curves parameterized by a real number  $\lambda$  (see Sect. 6). If such a  $U$  is defined on set of continuous functions  $t \mapsto x(t, \lambda)$  parametrized by  $\lambda$ ,  $U$  becomes a function of the parameter  $\lambda$ . If one requires that the variation of  $x$  at the point  $t$  be a first-order approximation to the difference  $x(t, \lambda + \delta\lambda) - x(t, \lambda)$ , it is natural to make the definition  $\delta U = \frac{\partial U}{\partial \lambda}[x(t, \lambda)]\delta\lambda$  with  $\delta x = \frac{\partial x}{\partial \lambda}\delta\lambda$ . If  $U$  is in fact defined on a functional space  $E$ , since the preceding formula makes no supposition about the topology of  $E$  or the variation of  $U$  for general  $\delta x \in E$ , one must make this formula compatible with the different types of variation of  $U$  on  $E$ , so that it has the required form when restricted to the parameterized family. Therefore Fréchet, requiring that  $\delta U$  be linear in  $E$ , proposed that the formula for the composition of partial derivatives be part of the definition of  $\delta U$  on  $E$ , as had been suggested by Hadamard. In the variant known as the Hadamard-Fréchet differential, one demands that  $\frac{\partial x}{\partial \lambda}$  belong to the functional space on which  $U$  is defined. In 1937, Fréchet described this differential as “the generalized differential of Monsieur Hadamard” and contended that the extension of Hadamard's definition to functionals had up to then neither been studied nor even proposed. Ten years after Lévy's last discussion on this topic in the second edition of his book (Lévy 1937), Manuel Balanzat published a note presented by Fréchet to the Academy of Science<sup>29</sup> that extended the Hadamard-Fréchet differential to topological vector spaces. If  $X$  and  $Y$  are two such spaces, then  $f : X \rightarrow Y$  is differentiable at  $x_0$  if there exists a continuous linear function  $D$  (the differential of  $f$  at  $x_0$ ) with the following property: for any real function  $g$  satisfying  $g(\lambda_0) = x_0$  that is differentiable at  $\lambda_0$ ,  $G = f \circ g$  is differentiable at  $\lambda_0$  and  $G'(\lambda_0) = D[g'(\lambda_0)]$ . Thus the differential made no use of the metric structure. In that note, Balanzat's principal source is Fréchet's memoir of

<sup>29</sup>Balanzat, Manuel (1960) La différentielle d'Hadamard-Fréchet dans les espaces vectoriels topologiques. CRAS (November 28, 1960).



1937, augmented by his work in topology. The note also contains a short discussion of the different ways to axiomatize differential calculus in topological spaces and in certain Fréchet spaces that are not topological in the Bourbaki sense.

### 4.3 Probability

After the First World War Lévy and Fréchet both became interested in the theory of probability at almost the same time, for pedagogical reasons. They soon realized that there was much to be done (see Sect. 2). Lévy discusses this moment in his autobiography (Lévy 1970), explaining that essentially all required methods and concepts had either to be recovered or created.

Fréchet never overlooked applied research and applications, and his work was quickly noticed by Borel. As opposed to almost all probabilists of that period, Fréchet did not come to probability theory via theoretical physics, set theory, or actuarial studies, but via applications in statistics, economics, and social science.<sup>30</sup> In this regard, he followed a German tradition represented in Strasbourg by G.F. Knapp (1842–1926), professor of Statistics and Economics from 1874 to 1918. This focus on applications marked Fréchet’s entire scientific career. His course in the theory of probability at the IECS (*Institut d’Enseignement Commercial de Strasbourg–Strasbourg Institute for Business Studies*), written with Halbwichs (whose contributions are readily identifiable) entitled *Le calcul des probabilités à la portée de tous* was published in 1924. The two authors presented the theory of probability via applications to insurance and statistics, without using high-level mathematical analysis. They wrote in the introduction “We have not attempted to follow the path of our forerunners either in detail or in outline.” One can easily imagine that Borel, involved in politics (he was deputy in 1924 and even briefly became minister in 1925), appreciated this link between social science, statistics, and the theory of probability.<sup>31</sup> Some of the themes that Fréchet would follow during the next forty years are presented in this book. All his life, Fréchet retained a taste for teaching mathematics applied to social science to students with little scientific background. He returned to this subject in innumerable publications that are collected in his 1955 book *Les mathématiques et le concret*.<sup>32</sup>

Lévy’s approach to mathematics was very different. As a result, beginning with his first notes in CRAS (presented by Hadamard) on characteristic functions and the central limit theorem, Lévy found himself in the center of a dispute between the school of the theory of functions and that of functional calculus, represented by Borel and Hadamard, respectively. Laurent Schwartz notes that Hadamard believed Lévy had abandoned mathematics for physics by devoting himself almost

---

<sup>30</sup>For details on this important moment in Fréchet’s life, see Havlova et al. (2005).

<sup>31</sup>For Borel’s interest in probability, see Durand and Mazliak (2011).

<sup>32</sup>Presses Universitaires de France, Paris (1955).

exclusively to probability, and never pardoned him. Lévy's methods also put him in conflict with Borel (see Lévy 1925). Borel found Lévy's use of rigorous mathematical techniques of little interest, and his analytical reasoning unnecessarily complicated.<sup>33</sup> And yet Lévy's book (Lévy 1925) has a whole chapter on the theory of gases, which must have pleased Borel. In any case, Lévy was very proud of the appreciation that this chapter garnered from physicists. Levi-Civita wrote to Lévy that this was "obviously" the chapter that most interested him. In 1959, Cramer mentioned that Lévy's 1925 book was the first modern treatise on probability

among those books that mark an advance and show the direction in which the science will develop in the future. A central place is given in the book to the general theory of probability distributions, which was novel and audacious, something that the younger generation of today can barely comprehend. . . .<sup>34</sup>

Although physical considerations were in later years sometimes presented by Lévy as a guide to his thinking or to illustrate a property by analogy (see Sect. 7), Lévy did in fact not go on to study connections between probability and physics. Hadamard's reproach is probably for other reasons; perhaps Hadamard hoped the audacity remarked on by Cramer would be applied by Lévy in more noble areas than probability. But it was more likely the supposed absence of rigor in the language used by Lévy to describe the theory of probability that occasioned the reproach. Lévy had to explain this at length in his own preface to Lévy (1937) (which followed Borel's preface). One sees yet again how Fréchet benefited from the support of his two mentors, while Lévy, not succeeding in capturing the interest of Borel, lost a bit of Hadamard's as well.

Among the first topics in probability discussed in the letters, the law of errors took on considerable importance. The two mathematicians worked on and compared their approaches to this topic from the start of the 1920s, and found themselves at the center of a theoretical debate that interested and impassioned an entire international community. The revival of interest in the question of how to justify practical methods of taking into account errors of measurement had various origins, from the experiments done in physics to ideas about its role in the artillery fire of the war that had just ended. In any case, the scientific dispute between Fréchet and Lévy was cordial, but very animated. It figures extensively in an amazing series of letters at the end of 1925, and reemerges during their unfortunate attempt at election to the Academy of Science in 1951.

Lévy embraced the questions of Gauss' law and the central limit theorem from his first lectures in probability theory at the Ecole Polytechnique in 1920. His lectures included various aspects of the convergence to a Gaussian distribution of which he gave several demonstrations and examples, tailored to his audience.<sup>35</sup> In 1921, at a conference of a philosophy meeting, he addressed this question outside the closed

---

<sup>33</sup>See Bru (1999).

<sup>34</sup>Cramer, Harald (1958) Eloge de M. Paul Lévy. In: *Le calcul des probabilités et ses applications* (July 15–20, 1958). Editions du CNRS, Paris. pp. 13–15.

<sup>35</sup>See Barbut and Mazliak (2008).

circle of lecture halls for the first time. In 1924 and 1925, several publications of Lévy and Fréchet both addressed this matter and cited each other. One may follow most of the debate about the law of errors between Lévy and Fréchet in the letters. One of them was published with Lévy's approval (after proofreading the manuscript) in a 1925 article by Fréchet in the Moscow Mathematical Society (see Letter 14) and a passage from a letter by Fréchet was copied by Lévy at the end of a 1929 article (see Letter 24).

A dialogue arose between Fréchet and the Soviet astronomer M.F. Subbotin of Novocherkassk, who was interested in a justification of the corrections made to measurements in the observation of stars. In 1923 he proposed two axioms that any possible error distribution must satisfy.<sup>36</sup> These axioms were inspired by an attempt by Schiaparelli dating from the end of the 19th century, which Subbotin found in the book of Czuber.<sup>37</sup> They were formulated as follows.

Axiom I states that the probability (in modern terms, the probability density) of an error depends only on its size via a certain *error function*  $\varphi(\varepsilon)$  associated to the particular measurement. In a series of measurements of a quantity, the probability that the error lies between  $\varepsilon$  and  $\varepsilon + d\varepsilon$  is given by  $\varphi(\varepsilon) d\varepsilon$ . Subbotin also included a technical requirement that “in general” (sic)  $\varphi$  have a continuous first derivative. This caused Fréchet to object.

Axiom II is a homogeneity principle, which dates to Schiaparelli. It requires that the most probable value of the quantity  $x$ , of which one has observed measurements  $l_1, l_2, \dots, l_n$ , that is the value of  $y$  which maximizes the product  $\varphi(y - l_1) \cdots \varphi(y - l_n)$ , be independent of the unit of measure. It was on the interpretation of this axiom that the debate took place.

Lévy's interpretation was, naturally, the most probabilistic. For him, this axiom signifies that if  $y = F(l_1, \dots, l_n)$  is the value at which  $\varphi(y - l_1) \cdots \varphi(y - l_n)$  attains its maximum, then for all  $\alpha > 0$ ,  $\alpha F(l_1, \dots, l_n)$  is the value at which  $\varphi(y - \alpha l_1) \cdots \varphi(y - \alpha l_n)$  attains its maximum. There is, as Lévy remarks, a connection with his theory of stable probability distributions (see Sect. 5).

Subbotin's first article had the same interpretation of axiom II as Lévy. However, his study made an implicit use of the fact that  $\varphi$  has a continuous derivative everywhere in order to have a simple maximum condition, which led in turn to an error distribution of the form  $\varphi_m(\varepsilon) = \frac{mh}{2\Gamma(m-1)} \cdot e^{-h^m|\varepsilon|^m}$  where  $m > 0$  and the constant  $h$  is chosen to get a probability density. An article by Fréchet<sup>38</sup> then observed the use of this stronger hypothesis, but made an error (with a different interpretation of the axiom) in proposing a counterexample. Surprisingly, this seems to have embroiled Subbotin. In a new note,<sup>39</sup> he accepted Fréchet's conclusions and showed that the general form of a function satisfying the two axioms may be written as a convergent product  $A \prod_{r \geq 1} [\varphi_{m_r}(u)]^{k_r}$ , which can have a non-continuous derivative

<sup>36</sup>(1923) Математический Сборник 31:296–301.

<sup>37</sup>Czuber, Emanuel (1891) Theorie des Beobachtungsfehler. Teubner, Leipzig.

<sup>38</sup>(1924) Математический Сборник 32:5–8.

<sup>39</sup>(1925) CRAS 180:1716–1719.

at 0. Finally, in another note<sup>40</sup> Fréchet showed that the general form of admissible functions of errors obtained by Subbotin can be recovered directly as the solution of a problem of functional optimization.

On receiving Fréchet's article (see Letter 9), Lévy made his disagreement known, and their discussion of Axiom II began. Fréchet's understanding of it is at once tortuous and simple: he required that if the measurements  $l_1, \dots, l_n$  lead to selecting  $x$  as the true value, the measurements  $\alpha l_1, \dots, \alpha l_n$ , taken with the error distribution  $\varphi_1$  deduced from  $\varphi$  by multiplication by  $\alpha$  (and therefore with probability density  $\frac{1}{\alpha}\varphi(\frac{1}{\alpha}x)$ ) lead to selecting  $\alpha x$ . Lévy pointed out to him that this fact holds for any distribution, so that Axiom II becomes vacuous. Subsequently, Fréchet accepted this argument and published an explanation (see Letter 14) in which he only partially gave way, but he did acknowledge publicly that his interpretation of axiom II could not lead to any useful conclusion.

In 1927, the discussion moved to another point. From Fréchet's two articles<sup>41</sup> arose the possibility of replacing the hypothesis of the additivity of errors by consideration of the maximal error. Lévy then wrote a final article on the topic.<sup>42</sup> He criticized Fréchet's hypothesis, justifying once more the for him indisputably additive character of errors and Gauss' law. By now, it had been generally accepted that according to the specific "concrete" situation, one could follow either Lévy or Fréchet's interpretations. This long debate had some unfortunate consequences for the two men. It gave rise to unpleasant remarks about Lévy's supposed errors and the propensity of Fréchet to philosophize rather than do mathematics, remarks that reappeared during their candidacy for the Academy of Science.

By contrast to Lévy, axiomatizing the foundations of the theory of probability and random processes was of immediate interest to Fréchet. He soon realized that the semi-formalized probabilistic notions had much in common with the concepts of general mathematical analysis. Laplace had observed that Bernoulli's theorem arose from a double approximation, the frequency being very likely to be very close to the probability. This very particular convergence corresponds completely to the convergence in measure in the theory of functions. This immediately suggested that random quantities are naturally governed by several modes of convergence, just as functions of a real variable. These modes are not equivalent, which implies that there are several possible versions of the law of large numbers. Clearly, Fréchet was struck by this. He made it the subject of his first course on probabilities at the Institut Henri Poincaré in 1928–1929, and in 1930 published a beautiful article that undoubtedly influenced Kolmogorov, who visited him during the summer of 1930 (see Sect. 3). This point gives an interesting hint about the personalities and thoughts of Fréchet and Kolmogorov. During the second half of the academic year 1928–1929, Fréchet

---

<sup>40</sup>Fréchet, Maurice (1925) Sur la loi des erreurs d'observation. CRAS 181:204–205.

<sup>41</sup>Fréchet, Maurice (1928) Sur la loi de probabilités de l'écart maximum. Ann. Soc. Pol. Math. 6:93–122 and Fréchet, Maurice (1928) Sur l'hypothèse de l'additivité des erreurs partielles. Bull. Soc. Math. Fr. 52:203–216.

<sup>42</sup>Lévy, Paul (1929) Sur quelques travaux relatifs à la théorie des erreurs. Bull. Soc. Math. Fr. 53:11–32.

lectured in Paris on the theory of random variables in a manner very close to his course at the IECS in Strasbourg. He defined

a random variable as a number  $X(E)$  of which the value is fixed by the result  $E$  of a trial belonging to a certain category  $C$  of trials. . . . One sees that  $X$  is a functional with a numerical value  $X(E)$ , with variable  $E$  which belongs to a domain  $C$ . In the following, whenever we consider several random variables  $X(E), Y(E), \dots$ , at once, we will assume that they are all defined over the same category  $C$  of trials, so that the symbol  $C$  can be omitted.<sup>43</sup>

Fréchet noted that by making the probability play the role of a linear measure, the types of convergence of numerical functions of real variables, namely convergence almost everywhere, convergence in measure, and convergence in the average of order  $r$ , could be extended to sequences of random variables. Therefore one may extend the usual properties of numerical functions of a real variable to random variables, notably the usual connections between the various modes of convergence, the possible metrizable, the questions of compactness (in Fréchet's sense), etc. It was, incidentally, Fréchet who introduced various designations such as *convergence in probability* by analogy with the *convergence in measure* used by analysts. Even though  $C$  was clearly an abstract set, it would be an exaggeration to think that Fréchet presented his students at the IECS and the IHP with an explicit set-theoretic axiomatics of the theory of probability. In fact, he followed the 1843 treatise *Exposition of the Theory of Randomness* by Cournot very closely, interpreting it in his own way and using a more general analysis to develop it. However, for Fréchet as for Borel, probability defined on a category of trials is a physical (or psychological) reality, while the linear measure of Borel is a concept belonging to the realm of mathematics. Fréchet did not seem to find it useful to clearly affirm in his books or courses that both could be related to the general definition of an additive function on abstract sets. For him, the theory of probability and the theory of linear measure do not belong to the same world, but this does not prevent them from being linked by the same mathematical structures and finding from this connection a basis for the richness of their possible applications. Thus he put himself in a purely French tradition, reaching from Poincaré to Borel, that declined to see real mathematics in probability theory.<sup>44</sup> Kolmogorov belonged to another generation and a different culture. The frontiers between analysis and probability must have appeared strangely Byzantine to him. Nonetheless, he always considered Fréchet a mentor.

With a few reservations, Lévy would little by little adopt Fréchet's understanding of probability spaces, i.e. via categories of trials. In Lévy (1937) he adopted the language employed by his colleague almost totally, and referred to it explicitly, using Borel's *fields*, the ancestor of modern  $\sigma$ -algebras with a then somewhat fuzzy definition; he later improved this definition and called the category of trials a *closed field*. In the same publication, he also returned to his previous work, dating from 1924, addressing the question of the definition of a probability distribution on an

---

<sup>43</sup>See Fréchet and Halbwachs (1924, Chap. V).

<sup>44</sup>See Mazliak (2014).

abstract set. His notable idea was to distribute the positive mass over the elements of a successively refined partition. This allowed him to construct a probability measure on a set that has the cardinality of the continuum, thus implicitly using the axiom of choice. He connected this approach to the integrals of Fréchet and Daniell, but never succeeded in relating this theory of partitions to the Fréchet integral. Lévy always defended a subjective interpretation of probability, and never gave a complete and general model of probability in the now-standard Doob form  $(\Omega, \mathcal{A}, P)$ , the only exception being the special case of  $[0, 1]$  equipped with Borel sets and the Lebesgue measure. Quite the contrary, Lévy's practice permits one to leave  $\Omega$  and the  $\sigma$ -algebra out and only speak of events for which there is no doubt of having a well determined probability because of their definition or because their probability is fixed via asymptotic conditions dependent on a chain of simpler events of well defined probability (for example, the zero-one law). This did not stop Lévy from being very precise and complete on questions of measurability (see Sect. 7). Lévy especially objected to the space  $\Omega$  and its technicalities for stochastic processes.<sup>45</sup> However this did not hinder him from considering questions as profound as the separability of processes or their modifications. For Lévy, a stochastic process  $X$  is a function that depends on chance at every instant, and the miracle of the choice of an element  $\omega$  determining the trajectory  $X(\cdot, \omega)$  does not correspond to any concrete situation, and certainly not to the philosophy of randomness that he wished to defend.

"In principle," wrote Lévy, "a stochastic process is a phenomenon for which chance intervenes at every instant  $t$  of its evolution. For Doob, a stochastic process is simply a random function  $X(t)$  of a variable  $t$  which one can imagine to be time. Whatever the set  $E$  of possible realizations of  $X(t)$ , a symbolic value  $\omega$  may be associated to each element of  $E$ . This formulation permits Doob to consider  $X(t)$  as a well-defined function  $x_t(\omega)$  of  $t$  and a random variable  $\omega$ , the choice of which encodes all successive influences of chance. It is a notation that is often convenient, even though it gives, as if born in one instant, that which for me is essentially a perpetual future. It should be remarked that if the set  $E$  has only the cardinality of the continuum, one can represent  $\omega$  as a scalar variable, or more precisely as a variable chosen between 0 and 1 with uniform probability. . . ."<sup>46</sup>

## 5 Lévy's Stable Probability Distributions

One of Paul Lévy's principal contributions to the theory of probability and its applications is his theory of stable probability distributions. This theory appears often, implicitly or explicitly, in his correspondence with Maurice Fréchet (especially in Letters 9 and 21).

---

<sup>45</sup>In the complementary notes to the second edition of Lévy (1937) Lévy mentions that Fréchet concurred.

<sup>46</sup>Lévy (1937), note II in the second edition of 1954.

Lévy designated a family of probability distributions on  $\mathbb{R}$  (or the corresponding random variables) as a *distribution type*<sup>47</sup> if that family is invariant under translation and changes of scale; in other words, if an affine transformation  $X \mapsto aX + b$  ( $a > 0$ ) of the random variable  $X$  gives a variable governed by the same family of laws (see Letter 9).

Classic examples of distribution types are normal (Gaussian) distributions  $\mathcal{N}(m, \sigma)$ , exponential distributions  $P(X > x) = e^{-(x-x_0)/h}$ ,  $x \geq x_0$  and the Pareto distributions  $P(X > x) = (\frac{x_0+c}{x+c})^\alpha$ ,  $\alpha > 0$ ,  $x \geq x_0 > -c$ .

A distribution type is called *stable* if for any two independent random variables  $X_1$  and  $X_2$  belonging to the distribution type and  $a_1$  and  $a_2$  two positive constants, there exists a positive real  $a(a_1, a_2)$  such that the variable  $X$  defined by

$$aX = a_1X_1 + a_2X_2 \quad (1)$$

belongs to the same distribution type. More generally, if random variables  $X_1, X_2, \dots, X_n$  belong to a stable distribution type, and  $a_1, a_2, \dots, a_n$  are given positive reals, one can find  $a(a_1, \dots, a_n)$  such that  $X$  defined by

$$a(a_1, \dots, a_n)X = \sum_{k=1}^n a_k X_k \quad (2)$$

with  $a_i > 0 \forall i$  and  $a(a_1, \dots, a_n) > 0$ , belongs to the same distribution type.

Stability may of course be defined by operations other than the addition of independent random variables: for example stability with respect to the maximum (as for the Fréchet, Gumbel and Weibull distributions) or for positive variables by the product, as is the case for the log-normal distribution. Here we always assume stability with respect to addition which is the only case considered by Lévy.

The importance of stable distributions for the theory of probability and its applications derives mainly from the following: if a normalized sum of  $n$  independent random variables with the same distribution<sup>48</sup> has a limiting distribution as  $n$  goes to infinity, that limit is obviously stable under addition.

The best-known stable distribution is the normal distribution, and until at least 1919 it was the only one known, with the exception of the Cauchy distribution. This is the probability density  $\frac{1}{\pi} \frac{a}{a^2 + (x-x_0)^2}$  ( $a > 0$ ) named by Lévy. Poisson, and later Cauchy, demonstrated that it is stable.

Lévy began his research on stable distributions in 1919, but his definitive characterization of them did not appear until a joint note with A.Y. Khinchin soberly entitled *Sur les lois stables*.<sup>49</sup> Up to a translation, for such a distribution the distribution function  $F$  is specified by its characteristic function  $\varphi(z) = \int_{-\infty}^{\infty} e^{izt} dF(t)$ ,

---

<sup>47</sup>French: *type de loi*.

<sup>48</sup>Or an average, or more generally a positive linear combination.

<sup>49</sup>CRAS (February 3, 1936).

namely

$$\varphi(z) = \exp\left[-\left(a + i \frac{z}{|z|} b\right) |z|^\alpha\right] \quad (3)$$

with  $0 < \alpha \leq 2$ ,  $a > 0$  and

$$\left| b \cos\left(\frac{\pi}{2}\alpha\right) \right| \leq a \sin\left(\frac{\pi}{2}\alpha\right). \quad (4)$$

In (3), the parameters  $a$  and  $b$  take into account the units and possible asymmetry of the random variable, respectively, so that a stable distribution type is defined for each possible value of  $\alpha$ . Here,  $L_\alpha$  denotes the stable distribution type with parameter  $\alpha$ .

The note just referenced demonstrates inequality (4). There Lévy treats the case  $0 < \alpha < 1$  and Khinchin the case  $1 < \alpha < 2$ . For  $\alpha = 2$ , equation (3) defines the normal distribution, for  $\alpha = 1$  and  $b = 0$ , the Cauchy distribution. For all other values of  $\alpha$ , although the distribution  $L_\alpha$  is well defined by (3) and by an inverse Fourier transform, an explicit expression for the distribution function is not known. However approximate expressions are available, a point to be considered later.

At the Ecole Polytechnique in 1919, Lévy lectured several times on the theory of probabilities, as a substitute for Humbert. He covered the central limit theorem, mentioned the name *stable limit* of a sequence of distributions, and employed the Fourier transform.<sup>50</sup> In his autobiography (Lévy 1970, p. 77), he wrote

I would probably never have dreamt of this problem without a remark by Captain Lhoste,<sup>51</sup> who heard my lectures in 1919. . . . He told me (using different language, of course) that Gauss' law is the only stable distribution. I did not dare reply. But the same evening. . . .

What he discovered that same evening was that all distributions with the characteristic function

$$\varphi(z) = \exp(-a|z|^\alpha), \quad a > 0, 0 < \alpha \leq 2 \quad (5)$$

are stable, and when the distribution is symmetric, only these. He did not publish this result until 1922.<sup>52</sup> By the academic year 1920–1921 Lévy held the chair of Analysis at the Ecole Polytechnique. The part of his course dedicated to the theory of probability was by then much more substantial than his lectures of 1919. In the course, he showed that the only stable distributions with finite variance are normal distributions. The proof is based on the functional equation (7) below, verified by the characteristic functions of the distributions. In the 1922 note, Lévy gave the definition (1) of stable distributions; he demonstrated again that the Cauchy distribution is

<sup>50</sup>See pp. 495–513 of Humbert's lecture notes of his course in analysis at the Ecole Polytechnique (first division, 1919); the note 319; and also Barbut and Mazliak (2008).

<sup>51</sup>One can read about Lhoste in Hadjadji Seddik-Ameur, Nacira (2003) Les tests de normalité de Lhoste. Math. Sci. Hum. 162:19–43.

<sup>52</sup>Lévy, Paul (1922) Sur le rôle de la loi de Gauss dans la théorie des erreurs. CRAS 174:855–857.



stable, and gave the explicit relation between the coefficients of the two expressions (1) and (2) above

$$a^\alpha = \sum_k a_k^\alpha. \quad (6)$$

In the case  $\alpha = 2$ , i.e. Gauss' law, (6) becomes the familiar expression for the variance of a sum of independent random variables. Finally, in the same note Lévy introduced the expression *domaine d'attraction* (*domain of attraction*) of a stable distribution, and stated that this domain includes all distributions for which the expectation  $E(|X|^p)$  is finite when  $p < \alpha$  and infinite when  $p > \alpha$ . In 1923 three notes to the CRAS of Paris appeared; one on April 23, which proved the existence of a stable distribution for every  $\alpha \in (0, 2]$ ; one on May 7, which gave the expression (3) above (but not the inequality (4)) and proved the existence of asymmetric stable distributions; and finally one on May 14 which gave an example of a distribution  $L_\alpha$  obtained as a limiting distribution.

In 1923 Lévy also wrote the important article *Théorie des erreurs. La loi de Gauss et les lois exceptionnelles*.<sup>53</sup> The *exceptional distributions* are, of course, the cases  $L_\alpha$  with  $\alpha < 2$ . Lévy drew attention to the fact that taking the average of an increasing number of random *errors* with stable distributions having  $\alpha \leq 1$ , far from increasing the precision of a measurement, on the contrary makes the imprecision of the measurement tend to infinity, because for such distributions not only is the variance infinite (as in fact holds for all  $\alpha < 2$ ) but also the expected value. This article contains Lévy's proof, celebrated for its elegance and conciseness, of the formula (3) or (5) for the characteristic function of a stable distribution, starting from the functional equation

$$\varphi(az) = \varphi(a_1z)\varphi(a_2z). \quad (7)$$

Because of the definition (1), the characteristic function of a stable distribution must satisfy this condition.

Maurice Fréchet used Lévy's proof in his article *Fréchet, Maurice. 1927. Sur la loi de probabilité de l'écart maximum*.<sup>54</sup> Fréchet applied Lévy's technique not to the characteristic function but to the distribution function itself to establish the form of the distributions stable with respect to the maximum of  $n$  independent random variables. This appears to have been the origin of an argument between Fréchet and Lévy (see Letter 21 and especially Lévy's article *Sur quelques travaux relatifs à la théorie des erreurs*).<sup>55</sup>

Aside from the book of Lévy (1925) and the article of 1929, Lévy did not return to stable distributions until 1936 in his *Note commune* co-authored with Khinchin, in which he completed the one-dimensional case, and another note<sup>56</sup> which extends

<sup>53</sup>(1924) Bull. Soc. Math. Fr. 52:49–85.

<sup>54</sup>Ann. Soc. Pol. Math. 5:93–116.

<sup>55</sup>(1929) Bull. Soc. Math. Fr. 53:1–21.

<sup>56</sup>Intégrales à éléments aléatoires indépendants et lois stables à  $n$  variables. CRAS (March 17, 1936).

the idea to vector random variables in  $n$  dimensions. Then Lévy (1937) appeared. In his autobiography (Lévy 1970, p. 78), Lévy wrote: ‘I believe that this time I can say that in my book of 1937, the theory took its definitive form’. In fact this was only a stopping point. The definitive theory was developed later in the work of Doebelin, Lévy, Feller and the Soviet school. Doebelin (1938) and Gnedenko (1939), especially, discovered the characterization of the domains of attraction of the stable distributions.

In his work (Lévy 1937, § 56 and 57, pp. 198–204) Lévy identifies two properties that relate stable distributions to those discovered empirically by Pareto for the statistical distributions of income and inheritances.<sup>57</sup> The simplest Paretian form for a random variable  $X$  is

$$P(X > x) = \left(\frac{x_0}{x}\right)^\alpha, \quad x \geq x_0 > 0, \alpha > 0. \quad (8)$$

The first property is that the probability distribution function  $L_\alpha$  ( $0 < \alpha < 2$ ) is asymptotically Paretian in the sense that for  $x$  large<sup>58</sup>

$$P(X > x) \equiv \frac{K}{x^\alpha}. \quad (9)$$

In other words, for large values of the variable  $X$ , Pareto’s Law with exponent  $\alpha$  is a good approximation to the stable distribution  $L_\alpha$ .

One can also obtain the distribution  $L_\alpha$  as follows:

- (1) Choose  $n$  independent random variables  $X_1, X_2, \dots, X_n$  from a *Paretian urn*, i.e. a distribution of form (8) with exponent  $\alpha$  and,
- (2) if  $1 < \alpha < 2$ , define

$$Y_n = \frac{1}{n^{1/\alpha}} \sum_{k=1}^n (X_k - m), \quad (10)$$

where  $m = x_0 \frac{\alpha}{\alpha-1}$  is the expectation of the Paretian distribution (8)

- (3) or, if  $0 < \alpha \leq 1$ , define

$$Z_n = \frac{1}{n^{1/\alpha}} \sum_1^n X_k. \quad (11)$$

(The expectation of the Paretian distribution is infinite for this range of  $\alpha$ .)

As  $n$  goes to infinity, the probability distribution of  $Y_n$  for the larger  $\alpha$  values and that of  $Z_n$  for the smaller approach the stable distribution  $L_\alpha$ . This second property is the analog of the central limit theorem for non-Gaussian stable probability distributions. In order to assure convergence, it is necessary to replace the divisor  $\sqrt{n}$

<sup>57</sup>Pareto, Vilfredo. *Cours d’Économie Politique* Professé a l’Université de Lausanne. Volume I (1896), Volume II (1897). F. Rouge, Lausanne.

<sup>58</sup>See the note 6 to the Preface by Kai Lai Chung.

of the sum of independent random variables by  $n^{1/\alpha}$ , which diverges more rapidly. When  $\alpha = 1$  the limit is the Cauchy distribution, and here the arithmetic mean appears; for  $\alpha = \frac{1}{2}$ , the divisor is  $n^2$ .

Since these two properties are quite general, the argument is a satisfying explanation of why distributions that are (at least approximately) asymptotically Paretian appear so frequently: for example, the distribution of the frequency of words in a sufficiently long text, the populations of cities, the length of rivers, etc. Fréchet (see for example Fréchet 1955), and later Benoît Mandelbrot and other authors, give multiple further examples in the physical sciences (physical geography, the physics of “sandpiles”), in the social sciences (demographics and economics), and even in the somewhat controversial subject of mathematical finance. It is appropriate and satisfying that these distributions are known as *Pareto-Lévy probability distributions*, which honors both their discoverer, Pareto, and their principal theoretician, Lévy.

## 6 Lévy and Potential Theory: From Analysis to Probability

Before he focused on probability, Paul Lévy was an analyst, a part of his career that is somewhat overlooked today. Lévy’s expertise in potential theory was considerable, and resulted in studies in mathematical analysis as well as in probability theory by way of Brownian motion. Several of Lévy’s letters to Fréchet mention these matters, which bestow a certain continuity on Lévy’s work.

There is extensive documentation of the intense and complex history of classical potential theory before 1945, which marked the earliest appearance of axiomatic and probabilistic theories.<sup>59</sup> The solution of the Dirichlet problem, the goal of generations of mathematical physicists and mathematicians, was the central problem of the classical theory. Up to the Second World War, the focus of attention alternated between the conditions on the domain, the boundary itself, and the boundary conditions. Research drew on all the central developments of functional analysis and contributed to the reanimation of entire branches of mathematics. The history of the probabilistic theory of potential has also been extensively investigated.<sup>60</sup>

A brief overview of this immense subject will open the way to a deeper understanding of the spirit of Lévy’s work and certain of his letters to Fréchet. The work of George Green, made accessible by the young Sir William Thomson (the future Lord Kelvin) in 1845, is founded on three integral formulas which Green applied to the theory of electricity. These formulas also apply to gravity and heat, evidencing a formal mathematical analogy between these different physical phenomena.

---

<sup>59</sup>For details, see Dieudonné (1978), Bouleau (1986), Locker (2001), Taton (1961), Brelot (1964), Choquet (1990), Chung (1995) and the very complete article by De la Pradelle (1988).

<sup>60</sup>See the historical notices in the work of Dellacherie and Meyer (1975), the text by Bouleau (1986) and numerous remarks and notes by Doob (1953, 1980).

The three formulas in question are:

$$\iiint_V v \Delta u \, dv + \iiint_V \nabla u \nabla v \, dv = \iint_S v \frac{\partial u}{\partial v} \, ds \quad (\text{GI})$$

$$\iiint_V (v \Delta u - u \Delta v) \, dv = \iint_S \left( v \frac{\partial u}{\partial v} - u \frac{\partial v}{\partial v} \right) \, ds \quad (\text{GII})$$

$$u(P) = -\frac{1}{4\pi} \iiint_V \frac{1}{r} \Delta u \, dv + \frac{1}{4\pi} \iint_S \frac{\partial u}{\partial v} \frac{1}{r} \, ds - \frac{1}{4\pi} \iint_S u \frac{\partial}{\partial v} \frac{1}{r} \, ds \quad (\text{GIII})$$

In these formulas the normal derivative  $\frac{\partial}{\partial v}$  is oriented inward, the volume integrals are over  $V$  and the surface integrals are over the surface  $S$  which bounds  $V$  with normal derivatives evaluated on  $S$ . The distance from the point  $P$  inside  $V$  to the point  $M$  being integrated over is  $r$ . On the right hand side of the third identity (GIII), the first term is the potential caused by a volume density  $\frac{\Delta u}{4\pi}$ , the second term is caused by a surface density  $\frac{1}{4\pi} \frac{\partial u}{\partial v}$  and the third term is the potential due to a double layer potential<sup>61</sup> of moment  $\frac{1}{4\pi} u$ , which Lévy put to use in a very clever way in 1910.

The Green formulas are a succinct presentation of the relations between functions and potentials. They reveal the central feature of potential theory, namely that a sufficiently regular function on a sufficiently regular domain is the sum of three potentials. If the function is in addition harmonic, it is the sum of the potential of a single layer and the potential of a double layer.<sup>62</sup> Equation (GIII) gives a harmonic function  $u$  at every point in the volume  $V$  in terms of its normal derivative on the surface; the Neumann problem<sup>63</sup> in fact consists of determining a harmonic function in a domain in terms of the values of its normal derivative on the boundary. By manipulating (GII) one is led to introduce the function  $H(P, M)$  such that, for fixed  $P$  in  $V$ , the Green function  $G(P, M) = \frac{1}{PM} - H(P, M)$ <sup>64</sup> is a harmonic function of  $M$  in  $V$  which vanishes on the boundary. One arrives at the equation

$$4\pi u(P) = \iint_S u(M) \frac{\partial G(P, M)}{\partial v_M} \, ds(M) \quad (*)$$

<sup>61</sup>See footnote 62 below.

<sup>62</sup>The double layer potential consists of a surface distribution of dipole moments all oriented toward the same side of a double-sided surface. The double layer potential may also be viewed as the result of two surface charge distributions when the surfaces are infinitesimally close. The usage of the word “distribution” here follows that of physics of the period, that is as a density of electric charges. During the first quarter of the 20th century this was a continuous function. Only later, in the work of Schwartz, did this concept as well as that of a *unit-doublet* find definitions not requiring a limit by means of distribution theory.

<sup>63</sup>This refers to the Berlin academic Carl Neumann (1832–1925), who was editor of *Mathematische Annalen*.

<sup>64</sup>In two dimensions the potential is logarithmic and the singularity which must be removed is  $\ln PM = \ln r$ , rather than  $1/PM = 1/r$ .

which determines  $u$  by its values on the boundary.<sup>65</sup> This same equation was employed by Lévy during the Second World War to discover new relations between Brownian motion and the harmonic measure.

The “harmonic Dirichlet problem”<sup>66</sup> seeks to determine a harmonic function  $u$  in a domain by its values on the boundary. The Green function solves this problem if the domain and its boundary satisfy certain appropriate conditions. Green used ideas from electrostatics to derive equation (\*), which also gives the temperature  $u(P)$  at a point  $P$  in an isotropic homogeneous body when the boundary is held at a temperature  $u(M)$ . This visualization played an important role in potential theory during Lévy’s time and was very significant for Lévy’s intuition about diffusion and Brownian motion.

The general Dirichlet problem is to determine a function  $u$  that satisfies  $\Delta u = \rho$  in a domain  $\Omega$  for a given function  $\rho$  by specifying its values on the boundary. Under sufficient conditions on the domain, the solution is obtained by adding to the solution of the corresponding harmonic Dirichlet problem the unique function  $v$  which vanishes on the boundary and satisfies  $\Delta v = \rho$  in the interior of the open domain. In three dimensions, one has  $v(P) = \frac{1}{4\pi} \iiint_{\Omega} \rho(M)G(M, P) dM$ . The existence of these solutions for domains of which the shape (i.e. the geometry and topology) is that of a real conductor is evident from the results of physical experiments and common sense. The work of mathematical physicists gave this observation a rigorous basis.

Dirichlet’s principle shows that the solution of the Dirichlet problem is characterized as the minimum of an energy integral. This guarantees the existence of a solution and serves as a method of calculating it. However, the extent of its applicability came into question with the development of a mathematical potential theory encompassing domains of more general shape. This occurred, notably, in Weierstrass’ criticisms of the work of Riemann<sup>67</sup> who had used the Dirichlet principle in his theorem on holomorphic transformations and conformal representations. As for the method of Green, not only does the issue of the conditions necessary for the existence of the Green function in a general domain arise, but finding an explicit Green function generally entails intractable computations.<sup>68</sup>

---

<sup>65</sup>Independently of Green, Gauss’ mean-value theorem for harmonic functions also shows that the potential is determined by its values on the boundary. This result allows one to derive the Poisson representation formula.

<sup>66</sup>This designation is used deliberately, since it was introduced by Lévy and employed in 1926 by Bouligand in the volume of the *Mémorial des Sciences Mathématiques* he devoted to these questions.

<sup>67</sup>Especially for the question of the theorem on the conformal representation in the complex plane which is related to the questions considered here. It is well known that in two dimensions, potential theory has a close relation to the theory of holomorphic functions.

<sup>68</sup>The Green function is itself the solution of a Dirichlet problem. Except for domains with a boundary consisting of curves or surfaces of a simple geometry, for which the Green function is known from the work of Poisson, Liouville, Cauchy and their successors, very few explicit examples are available. However numerical methods may be used to determine Green functions to a high degree of accuracy.

One of Lévy's first articles as a professional mathematician<sup>69</sup> concerns a question in potential theory. It announces his thesis defense of 'Integro-differential equations defining functions of line' (Lévy 1911) of the following year, when he also published another paper.<sup>70</sup> Lévy's thesis committee was composed of Hadamard, Picard and Poincaré, three eminent scientists who had worked on potential theory. Picard was interested in ways of solving the differential and integro-differential equations of mathematical physics and published extensively on potential theory between 1900 and 1904. He gained distinction by introducing the famous 'Picard Principle' of successive approximations in his lectures at the Sorbonne. Poincaré introduced his 'balayage method' in 1887 (see Bouleau 1986). In his authoritative article of 1887, he put forward a solution for 'a conductor of arbitrary shape' after having criticized the solutions of Riemann, Schwarz and Harnack.<sup>71</sup> Despite the efforts of Hilbert, the systematic utilization of Dirichlet's principle fell into question.<sup>72</sup>

In 1906 Hadamard furnished new tools for potential theory utilizing his functional calculus. He discovered variational equations for the Green function of a domain whose boundary undergoes continuous deformations. In such circumstances, the Green function ceases to be a symmetric function of two points and becomes a function of the domain and of its boundary (and hence a function of line in Volterra's sense). This framework admits a continuous deformation of the domain and the corresponding Green function may be characterized.

As Hadamard's student Lévy was seemingly destined to pursue the French approach to potential theory, which for a time favored the treatment of Dirichlet's problem by functional analysis, before returning to Poincaré's method. In the work of Hadamard and then Lévy, Green function techniques were brought to center stage, together with the equations, including functional derivatives, that they imply and the generalizations that they entail. In fact, the longest part of Lévy's thesis is devoted to these generalizations.

Lévy began his five chapter opus with a very dense and relatively long introduction, constituting a third of the total 130 pages. In this introduction, Lévy presents what he calls "a summary of the works of Hadamard and Volterra on the functions of line," an expression taken from the works of Volterra and his Italian school, later rebaptized functionals by Hadamard. The two sole examples given by Lévy are:

---

<sup>69</sup>Lévy, Paul (1909) Sur les valeurs de la fonction de Green dans le voisinage du contour. *Bull. Soc. Math. Fr.* 34:186–190.

<sup>70</sup>Lévy, Paul (1911) Une généralisation de la méthode de Fredholm pour la résolution du problème de Dirichlet. *J. Ec. Polytech.* 15:197–210. Fredholm, the Swedish creator of the theory of integral equations that bears his name, developed this method in 1900.

<sup>71</sup>He called Riemann's method (which uses Dirichlet's principle) 'simple and elegant but not rigorous,' and the other methods are quite insufficient in Poincaré's opinion for solving Dirichlet's problem.

<sup>72</sup>Hilbert declared at the congress of 1900: 'Every problem in the calculus of variations has a solution provided that one makes the appropriate hypotheses on the boundary conditions and, if necessary sufficiently extends the concept of solution'. However in 1901, in order to give a proper proof of Riemann's theorem, he had to substitute the calculus of variations for very sophisticated theorems on equicontinuous families of harmonic functions.

“the coefficient of mutual induction of two boundary curves” and “the Green function which is a function of two points of a plane, the form of which depends on a closed contour  $C$ . This is thus a function of two points and of the contour  $C$ . It is of a type of analogous functions that play an important role in mathematical physics.”

By far the most original part of Lévy’s thesis was Chap. I, which covers higher-order functional derivatives. With these new derivatives (especially those of order two), Lévy became a creative mathematician and set out the tools that would later permit him to generalize the Laplacian and the Dirichlet equation to functional calculus.

In his second chapter, Lévy studied equations of the form

$$\delta\Phi_A^B = \int_C f(\Phi_A^B, \Phi_M^A, \Phi_M^B; A, B, M) dS(M).$$

Chapter III was dedicated to the integro-differential equation satisfied by the Green functions. It examines the consequences of the Hadamard variational equation for the Green function

$$\delta G(A, B) = \frac{1}{(n-2)S_n} \oint_{\Sigma} \frac{\partial G(A, M)}{\partial \nu_M} \cdot \frac{\partial G(M, B)}{\partial \nu_M} (\nu, \delta M) d\sigma(M)$$

in dimension  $n > 2$ .

Chapter IV is concerned with the potential of nested charged slabs, which it treats by successive approximation, an application of Picard’s method.

Finally, in Chap. V, which is more an elegant exercise than a fundamental contribution, Lévy addresses the solution of a functional version of the equation  $\Delta \Delta u = 0$ .

Aside from the appealing theoretical questions he solved in his thesis (Hadamard told him<sup>73</sup> “since you have taken up this subject, I leave it to you”), Lévy remained true to the French tradition of mathematical physics. He used the most recent tools of functional analysis, which seemed to him to be the most appropriate, and did not neglect the physical interpretation of his results. Such interpretations served him as a guide to various topics, including questions of the relation between potential theory and probability in Brownian motion (cf. Sect. 7). Between 1912 and 1919 Lévy published several articles touching on potential theory, gave a presentation of this research in 1922 (Lévy 1922), and in 1935 published a note on tensorial forms of Green functional equations.<sup>74</sup>

Lévy’s first letters to Fréchet are about functional analysis and included a list of the types of functions that interested him. The function spaces that are today canonical in analysis can be read between the lines. The main topics are various features of Green functions and the Laplace and Dirichlet equations. Above all, Lévy rails against Volterra’s formulation of the Laplace equation in terms of “functions of line,” as may be read in the scathing Letter 5. It is important to emphasize the extent

<sup>73</sup>Lévy (1970, p. 42).

<sup>74</sup>Lévy, Paul (1935) Sur une forme tensorielle des équations fonctionnelles des fonctions de Green et de Neumann. CRAS 200:1723–1725.

to which Lévy presented himself as an analyst who followed Hadamard and differentiated himself from Volterra, because the theoretical links between Volterra's "functions of line" and the functionals of Hadamard are in fact very close. Laurent Schwartz wrote in 1988 that the title of Lévy's thesis "corresponded to current ideas that followed from the work of Volterra."<sup>75</sup> Once he had completed his thesis, Lévy possessed the mathematical tools to modify and extend certain facets of Volterra's calculations relevant to potential theory.<sup>76</sup> Not only in his thesis, but in his entire opus of analytical research (which often consisted of reworking certain problems so that they fell under the conditions set out in his thesis), Lévy kept to generalizations of functional derivatives, especially the Laplacian. He thus inserted himself into the framework built by Hadamard, becoming his mentor's theoretical and technical successor. It was only after the First World War that Hadamard prompted Lévy to investigate Fréchet's techniques and asked him to assemble the posthumous edition of Gateaux's work.<sup>77</sup> In his 1922 book (Lévy 1922) and his letters to Fréchet, Lévy used Volterra's notation for functional variations and derivatives, but clearly departed from Volterra's conclusions. Volterra used the equation  $\delta U = \int_0^1 \phi(t) \delta x(t) dt$  for the variation of a functional, and Lévy specified that in cases where that equation applies "we say that  $U$  is differentiable in the sense of M. Volterra. . . M. Volterra represents it by  $\phi(t) = U'[x(t)]$ ." Lévy also gave cases to which Volterra's formula does not apply.<sup>78</sup> He took care to distinguish in the first-order terms between the variation or differential in the sense of Fréchet and the variation or differential in the sense of Gateaux. He defined the corresponding second variation and paid particular attention to second variations of the form  $\delta^2 U = \int_0^1 f(t) \delta x(t) dt + \int_0^1 \int_0^1 g(t, t_1) \delta x(t) \delta x_1(t_1) dt dt_1$  where the kernels  $f$  and  $g$  generalize the square ( $\frac{\partial^2}{\partial x^2}$ ) and rectangle ( $\frac{\partial^2}{\partial x \partial y}$ ) second partial derivatives. Lévy thus highlighted the absence of a term in Volterra's formula for the second variation and gave, as a generalization of the Laplace equation, the equation  $\int_0^1 f(t) dt = 0$  in place of Volterra's expression  $\int_0^1 g(t, t) dt = 0$ , which Lévy considered to be meaningless (Lévy 1922, p. 86). In 1970 (Lévy 1970), he revisited this question, which had always seemed to him to be one of his principal contributions to the calculus of variations and potential theory, maintaining that as of 1901 Hadamard "had drawn attention to the importance of the second term" which "plays such a large role in Gateaux's work and mine."

---

<sup>75</sup>Astérisque (1988) n° 157–158.

<sup>76</sup>See the article Fichera, Gaetano (1994) Vito Volterra and the birth of functional analysis. In Pier (1994), pp. 171–183.

<sup>77</sup>For this fascinating story, see the detailed study of Mazliak (2011).

<sup>78</sup>For the classic functionals of mechanics of the form  $\int_0^1 f(t, x, \dot{x}) dt$ , the variation cannot be put in this form, except in certain special cases. For linear functionals, which were characterized in 1903 by Hadamard and transformed in 1907 by Riesz into a representation as Stieltjès integrals  $\int_0^1 x(t) dF(t)$ , the variation can again not be put in Volterra's form, because of the possibility of points of discontinuity of  $F$ . Also, Volterra's form is not possible when  $U(x) = x(\tau)$  (i.e. when  $F$  is the Dirac measure at the point  $\tau$ ) for which  $\delta U = (\delta x)(\tau)$ .



By 1924, analysis and potential theory had made significant progress. Lebesgue became aware of the very advanced results of the new American school of potential theory. The Americans had maintained tight relations with the German school. Oliver Dimon Kellog (whose book *Foundations of Potential Theory* is still in print by Dover and has become a classic on American campuses) wrote his thesis in German, guided by Hilbert, and then went to the University of Missouri, where together with William Osgood (president of the AMS from 1905 to 1906) he helped develop the discipline in the USA. Griffith Conrad Evans (president of the AMS from 1924 to 1926) was a student of Osgood and an eminent specialist in potential theory. By 1924, he had numerous publications in this area. At about the same time the great German specialist Arthur Korn, who met Poincaré in 1891 and won the Valliant prize of the Paris Academy of Science in 1907, had established good relations on the international scene, especially with the Americans. The French school, by contrast seemed quite diffident and inward looking. Thus Lebesgue regarded the American work as so important that he followed up on Wiener's note of March 24, 1924 in the CRAS<sup>79</sup> with an 'Observation' of his own and with the presentation of a note by Georges Bouligand, of which Lebesgue is probably the source. Lebesgue wrote:

Last January, an article by G. Bouligand gave me the occasion to present several results obtained in the course of research on insolvable examples of the Dirichlet problem, research that began in 1913.<sup>80</sup> Wiener spoke at length about my article, but what he did not say, and I would like to mention, is that during recent years the Dirichlet problem has been, unbeknownst to me, the object of fruitful research by various American researchers, especially Wiener. Their theorems often go beyond mine, and they have so well addressed the various questions that concerned me that had I been aware of their results I would have undoubtedly found it pointless to follow up on my results of 1913. . . . The work of the American scientists was not known to Bouligand either, however one should add that the important note by Bouligand in 1919<sup>81</sup> seems to have been overlooked by the Americans. Now that these researchers are aware of each other, one may hope that they will become yet more fruitful.

Lévy turned most of his attention to probability in 1920, but did continue for some time to work on analysis. In addition Wiener, who originated the mathematical theory of Brownian motion in 1923, participated assertively in potential theory and the French, American and German schools converged on this problem. Wiener responded to the articles of Oskar Perron from the USA, discovering one of the final steps that resolved the Dirichlet problem in the case of a continuous boundary.<sup>82</sup> He also took important steps toward the solution of the problem with a discontinu-

---

<sup>79</sup>Wiener, Norbert (1924) Une condition nécessaire et suffisante de possibilité pour le problème de Dirichlet. CRAS 178:1050–1053.

<sup>80</sup>Lebesgue discovered a famous counter-example of insolvability of the Dirichlet problem (called Lebesgue's spine). This demonstrates the close relation between the existence of a solution and the topology of the boundary. He was preceded in this type of example by the Polish mathematician Zaremba.

<sup>81</sup>Lebesgue is referring to the two notes in CRAS by Bouligand on November 3 and 17, 1919.

<sup>82</sup>Wiener, Norbert (1925) Note on a paper of O. Perron. J. Math. Phys. 4:21–32.

ous boundary. Perron and Wiener are at the origin of the PWB (Perron, Wiener and Brelot) methods, which play an important role in potential theory.

In 1926 Georges Bouligand took up the task of writing a survey of potential theory for *Mémorial des Sciences Mathématiques* (Bouligand 1926). He cited Lévy extensively—for his thesis work and his sophisticated introduction of the double layer potential—as well as the other members of the French school. The German contributions are prominently referenced, and several pages are devoted to the Americans, through an analysis of the results of Wiener and Phillips. Lévy's work, however, soon became outdated and the Green function approach of Hadamard did not achieve its desired goals. Wiener's theory of Brownian motion and his remarkable results in classical potential theory (including the characterization of sets with vanishing electrostatic capacity, important for probability) included all Wiener needed to begin probabilistic potential theory. It is a bit surprising he did not do so. Instead the Japanese Shizuo Kakutani began this work in 1944 (Bouleau 1986), and the theorem which carries his name is recognized today as the starting point of the probabilistic potential theory developed after the Second World War. It should be recognized, however, that Kakutani's theorem had already been proven by Lévy in 1943, as can be witnessed in the riveting letters of that year (Letters 44–54).

The concept of harmonic measure, initiated in 1935 by the Nevanlinna (1936) played an important role in the finalization of classical potential theory and generated some beautiful results on holomorphic functions. In the 1930s the French school—and before long the entire international community concerned with potential theory—was dominated by the imposing figure of Marcel Brelot. Just before the Second World War, by introducing the idea of resolute boundary conditions, Brelot expanded the Dirichlet problem to functions integrable with respect to the harmonic measure and gave the problem, in the words of Choquet, a 'definitive form.' At the end of the war, using integrability over the ramified harmonic measure,<sup>83</sup> Brelot gave the necessary and sufficient conditions for solvability. This emphasizes the importance of the result published independently by Lévy in 1943 and Kakutani at the end of 1944. The connections between Brownian motion and potential theory were already known, and then were extended to random walks, leading to meetings between probabilists and potential theory specialists (Courant, Friedrichs and Lewy in 1928, Petrovski, Khinchin and Kolmogorov during the 1930s, etc.). This marked an important step; the tool which allowed the solution of the central problem of classical potential theory also ushered in the theory of probability.

Lévy was only able to publish his result after the Liberation of France, whereas Kakutani's work did not become known until the Japanese capitulation in 1945. The result is that under suitable assumptions, for a domain  $D$  with boundary  $F$  and a subset  $F'$  of  $F$ , the Dirichlet problem  $\Delta U = 0$  in  $D$  with boundary condition  $U = 1$  on  $F'$  and  $0$  on  $F \setminus F'$  has a unique solution  $P(A, F')$  for each point  $A \in D$  which is the harmonic measure of  $F'$  with respect to  $F$ . For a domain with a Green

---

<sup>83</sup>CRAS (November 26, 1945).

function  $G$ , one also has  $P(A, F') = \int_{F'} \frac{\partial G(A, M)}{\partial v} d\sigma(M)$ .<sup>84</sup> Today Kakutani's theorem may be stated as follows: the conditional expectation value  $g(x) = E_A(f(B_\tau))$ , where  $\tau$  is the earliest time at which a Brownian variable  $B$  starting at  $x$  reaches the boundary  $F$  on a region of the plane is a harmonic function of  $x$  which solves the Dirichlet problem  $\Delta u = 0$  in  $D$  with boundary condition  $f$ . This formulation is completely equivalent to the theorem stated by Lévy and then by Kakutani. Their explicit statements differ very little:

Lévy (1948): "The probability that a Brownian motion beginning at a point  $A$  inside a domain  $D$  attains the boundary  $F$  for the first time on  $F'$  is the harmonic measure of  $F'$  with respect to  $F$  viewed from the point  $A$ ."

Kakutani (1944b): "Let  $D$  be a domain in the Gaussian plane  $R_2$ , and let  $E$  be a closed set on the boundary  $\text{Bd}(D)$  of  $D$ . Then, under certain assumptions on  $D$  and  $E$ , the probability that the Brownian motion starting from a point  $\xi \in D$  will enter into  $E$  without entering into  $\text{Bd}(D) - E$  before it, is equal to the harmonic measure in the sense of Nevanlinna of  $E$  with respect to the domain  $D$  and the point."

Lévy's result holds in dimension  $n$ , so it might seem more general than Kakutani's, which is limited to dimension two. However a careful analysis (see Locker 2001) of Kakutani (1944a) shows that Kakutani already knew at the end of 1944 that his theorem held for arbitrary dimension. Although Kakutani's proof is written with complex variables, which shows his preoccupation with two dimensions and includes several crucial steps of a probabilistic nature that were not yet proven in 1944, it nevertheless immediately generalizes to arbitrary dimension. This analysis also shows that the assumptions of Lévy and Kakutani on the type of domains and their boundaries are essentially identical. The assumptions are explicit for Kakutani and implicit for Lévy. However, for Lévy they are made clear by the theorems from analysis that he used which require that the domain  $D$  must be simply connected and its boundary composed of Jordan curves.

There is a technical difference between Lévy's and Kakutani's results. Lévy's approach to the harmonic measure is via Green functions (as one would expect, given his extensive study of Green functions after the First World War and the way in which he preferred to be guided by analogy with the theory of heat conduction), whereas Kakutani completely avoided their use. There is also a difference in the consequences for Lévy and Kakutani. The theorem did not lead Lévy beyond considerations of Brownian motion and its trajectories, although he was able (Lévy 1948)

---

<sup>84</sup>This version of the theorem is sufficient to understand the Lévy-Fréchet correspondence. Modern definitions of the harmonic measure are more subtle and their connection with the Dirichlet problem more general. This more recent work was originated by Wiener and then extended by Brelot. The harmonic measure appears as a Radon measure  $H$ . If  $f$  is super-harmonic in a given open set and  $F$  is the set of super-harmonic functions  $g$  such that  $\underline{\lim} g(x) \geq f(y)$  (for  $x$  an interior point and  $y$  on the boundary), one may consider  $\overline{H}(f) = \inf F$  and  $\underline{H}(f) = -\overline{H}(-f)$ . When  $\underline{H}(f) = \overline{H}(f) = H(f)$  the boundary condition  $f$  is resolute, meaning that  $H(f)$  is a generalized solution of the Dirichlet problem. Wiener had already shown that  $f \mapsto H(f)(x)$  is a Radon measure  $H_x$ . This Radon measure is nothing but the functional representation of the measure. In 1945 Marcel Brelot considered the compactified projective space of a real vector space of arbitrary finite dimension, and redefined the ramified harmonic measure by an extension of the previous method to the space, making use of Daniell's integral.

to connect it with some very significant work of the Soviet school from the 1930s. This link allowed an expansion of the theorem's probabilistic significance by connecting it to the law of the iterated logarithm, via the work of Bernstein, Khinchin, Kolmogorov and Petrovski. Kolmogorov was able to deduce this law from Petrovski (1933, 1935) by considering parabolic equations, but he never published a proof. It was Lévy (1948) who, having heard of it directly from Kolmogorov, made this result known, some years after Ville (1939), Doeblin (1940) and Fortet (1943) had alluded to it. By contrast, in his December 1944 article, Kakutani presented his theorem as the basis for five additional theorems, among them one that established that the sets of capacity zero already characterized by Wiener are composed of precisely those points which Brownian motion almost surely does not reach. These five theorems became part of a program to represent harmonic functions with Brownian motion. Kakutani thus originated a strategy that ultimately led to the fusion of important areas of probability theory and potential theory.

It is nonetheless true that Lévy found the theorem before Kakutani and that Lévy, who gave three different proofs of it, differs from Kakutani in at least the two dimensional case, which makes use of the conformal invariance of planar Brownian motion (see Sects. 7 and 8). Today, this proof is taken as the starting point for the application of the theory of conformal martingales to holomorphic functions. Its extraordinary content, only briefly alluded to in the letters, is as follows: For a circle in the plane with center  $A$ , the harmonic measure of an arc  $F'$  of the circle viewed from  $A$  reduces to the measure of the angle at the center defined by the arc, divided by  $2\pi$ . Due to the homogeneity of Brownian motion starting at  $A$ , this is also the probability that the motion leaves the disk for the first time via the arc. A conformal transformation maps the general case into this simple situation.

To conclude this brief overview of an enormous subject, potential theory, it is necessary to touch on Lévy's research on Markov processes, begun in 1950. This is mentioned in the last letters of the correspondence and also by Kai Lai Chung in the Preface to this volume. The rediscovery of Doeblin by the Americans in 1950 and the work of Doob, Chung and also Hunt led to a renewal of interest in the theory of Markov processes, which in turn ushered in a merging of whole areas of potential theory and probability (see Bouleau 1986). One might add that Lévy, whose theory used "fictitious states" (the set of which constitutes the "Chung boundary set"), divided his classification into seven types of instantaneous states (the states  $E$  for which the sojourn time is zero (see Lévy 1970)). In this way he generalized the idea of a Markov process by replacing the time with surfaces in Euclidean space of dimension  $n$ , showing that the corresponding generalization of Brownian motion is no longer Markovian in the usual sense. Thus, beginning in 1955, Lévy was plunged into the difficult task of separating Markovian from strictly Markovian processes. His last academic publication (at age 84!) is entitled *Un nouveau point de vue pour l'étude des processus Markoviens*.<sup>85</sup> In it, using the example of Brownian motion on a curve in the form of a figure eight, i.e. a curve with a double point, Lévy attempted

---

<sup>85</sup>CRAS (May 25, 1970).

to relate Markov processes to strictly Markov processes (Lévy 1970) by introducing Riemann surfaces above the trajectories. C. Dellacherie and P.-A. Meyer (1975)<sup>86</sup> put Lévy among the ‘ancestors of the study of trajectories.’ They add that Lévy had considered ‘irregular Markov chains [and] that his work remained misunderstood for a long time.’

## 7 Lévy and Brownian Motion

In 1827 the Scottish botanist Robert Brown looked through his microscope at small particles found in pollen grains suspended in water and noticed their unusual, haphazard motion. For the next sixty years explanations of this phenomenon were primarily animist in nature before a physical reason emerged: In 1887 Father Ignace Carbonnelle attributed Brownian motion to thermal agitation; a year later Georges Gouy performed experiments varying several parameters and concluded that “this motion is the subtle result of thermal agitation.”<sup>87</sup>

At the end of the 19th century and dawn of the twentieth, certain features of Brownian motion were deemed relevant to basic question of physics, namely: are matter and energy continuous or discontinuous and is the validity of the Carnot principle an absolute law or a statement of high probability? Although the second half of the 19th century—thanks to Maxwell and Boltzmann—had seen probabilities penetrate deep into physics, notably in statistical mechanics and the dynamics of gases, nevertheless the meaning and significance of probability for physics had been obscured by philosophical issues. Now it became an important tool for the study of certain physical phenomena for which a particle description seemed necessary.

The earliest mathematical model relevant to Brownian motion appeared in Louis Bachelier’s thesis,<sup>88</sup> defended in 1900 (Bachelier 1900). The thesis includes a model of the fluctuations  $X(t)$  of the stock market as independent increments  $dX(t) = X(t + \delta t) - X(t)$  following a Gaussian law with standard deviation  $\sqrt{\delta t}$ . Such stochastic processes are now called Wiener processes, Wiener-Lévy processes, or linear Brownian motion. The equation  $\frac{\partial f}{\partial t} = \frac{\partial^2 f}{\partial x^2}$  satisfied by the density  $f(x, t)$  of the process is also the heat equation for a homogeneous substance in steady state. Bachelier included it under the name “radiation of probability,” but it was a dozen years before he attempted to connect his work with the topic of this section, Brownian motion. In the meantime physicists, notably Albert Einstein, had found an explanation for the phenomenon and begun a mathematical description of it. Beginning in 1902 Einstein developed a new molecular kinetic theory of heat. Thus Brownian motion was incorporated in the new physics of the day and merged with the theory of probability.<sup>89</sup>

<sup>86</sup>Chapters XII–XVI, p. 340, revised edition of 1987.

<sup>87</sup>An analysis of Gouy’s work and its importance for physics appeared in 1941 (Picard 1941).

<sup>88</sup>See the discussion in Sect. 2.

<sup>89</sup>At this time Einstein also revolutionized mechanics with his theory of special relativity and reintroduced the particle picture of light, i.e. photons and the photoelectric effect.

In a 1905 *Annalen der Physik* article (Einstein 1905) Einstein described the mathematical properties of Brownian motion implied by his new molecular kinetic theory. The physical existence of Brownian motion validated his mathematical theory. In 1906, again in *Annalen der Physik*, Einstein showed how one could determine molecular dimensions, and in a third contribution gave a consistent theory of Brownian motion, including gravitational effects. From the diffusion equation  $\frac{\partial f}{\partial t} = D \frac{\partial^2 f}{\partial x^2}$ , where  $f$  is the number of particles per unit volume at the point with coordinate  $x$  and  $D$  the diffusion constant, Einstein showed  $f(x, t)$  to be proportional to  $\frac{1}{\sqrt{4D\pi}} \frac{e^{-x^2/(4Dt)}}{\sqrt{t}}$ . Up to a multiplicative factor, this is just a Gaussian with standard deviation  $\sqrt{2Dt}$ . Experiments, especially those of Jean Perrin (who learned of Einstein's results from Langevin), could then measure molecular masses and Avogadro's number by using Einstein's theory. Einstein was in fact not the only one to deal with Brownian motion. In the same year (1906) and the same journal (*Annalen der Physik*), Marian von Smoluchowski began a series of very mathematical articles on physical Brownian motion (see Von Plato 1994 for details).

Now it would seem that the Brownian increments, being proportional to  $\sqrt{\delta t}$ , rule out any consideration of particle speed, and in fact in his book *Les atomes* (Perrin 1912) Jean Perrin declared that

the changes to the trajectory are so numerous and so rapid that they are impossible to follow. The apparent average speed varies wildly in magnitude and direction, without tending to a limit as the observation time decreases. This is a case in which it is natural to think of those functions without derivatives imagined by mathematicians.

In the original edition of *Les atomes*, Perrin further remarked that the trajectories are nonetheless continuous “because we cannot imagine that the particle goes from one position to another without cutting any plane which separates the two positions.” Other mathematicians and physicists constructed models of physical Brownian motion, or similar motions. Paul Langevin (1908), Adrian Fokker (1914) and Max Planck (1917) constructed models with continuous trajectories for which one can find a distribution of speeds. The process formalized by Leonard Ornstein and George Eugene Uhlenbeck (1930) of the Lorentz school, known today as the Ornstein-Uhlenbeck process, is sometimes considered as an alternative to Brownian motion. It can be derived from Brownian motion  $X$  by the analytic formula  $U(t) = e^{-t/2} X(e^t)$ .<sup>90</sup>

Wiener as of 1922 and as of 1934 Lévy each gave a mathematically consistent theory of Brownian motion as a continuous process with continuous but non-differentiable trajectories. Wiener described general relations between his constructions and statistical mechanics. Lévy made frequent use of the analogy between Brownian motion and the theory of heat as a guide or aide in his research (see, e.g., Letter 45). Aside from the forays of Bachelier twenty years earlier, this work signaled the emergence of Brownian motion as a subject for mathematical research. Wiener, with his Hilbertian approach, was more analyst than probabilist. Lévy, by

---

<sup>90</sup>For more on these questions as well as a general history of Brownian motion, see Kahane (1998) and Pier (1994).

contrast, made Brownian motion a central theme in the development of the theory of stochastic processes in continuous time. In Wiener's *Differential Space* (Wiener 1923), his seminal article of 1923, the continuity of the trajectories is assumed *a priori*. Lévy began, in 1934, a reformulation of Brownian motion which led him to an effective procedure for constructing trajectories that are almost surely continuous.

Wiener had already produced three articles on Brownian motion before writing *Differential Space*. In the references to all these articles, one finds Lévy, with his *Leçons d'analyse fonctionnelle (Lessons in Functional Analysis)* (Lévy 1922), in company with Gateaux, Daniell and Einstein.<sup>91</sup>

In Wiener's *Differential Space*, the basic constituents of the theory are the measures on the space of continuous functions associated to the increments of the function  $f$  which represents a trajectory of the Brownian motion. The differences (or increments)  $f(t) - f(t')$  between two positions of the particle are thus the bricks of Wiener's mathematical edifice. Wiener constructs a measure on the space of continuous trajectories that assigns a probability to all "reasonable" events constructed with the Brownian motion. He does this in two different ways. The first is inspired by the work of Lévy and Gateaux, and the second by Daniell's integral, which was the most general and abstract extension of the idea of an integral then available. Wiener's measure  $\mu$  can be represented by its definition on certain functions  $F$  that only depend on a finite number of increments of functions

$$\begin{aligned} \mu(F) &= \pi^{-n/2} \prod_1^n (t_k - t_{k-1})^{-1/2} \\ &\times \int_{\mathbb{R}^n} du \exp \left[ - \sum_1^n \frac{u_k^2}{t_k - t_{k-1}} \right] F(u_1, \dots, u_1 + \dots + u_n). \end{aligned}$$

Daniell's theorem then allows an extension to all  $\mu$ -integrable functionals and the Lévy and Gateaux averages (see Mazliak 2011 and Sects. 6 and 8) give an equivalent extension. Lévy mentions in his autobiography (Lévy 1970) how close he came in his *Lessons in Functional Analysis* (Lévy 1922) to Wiener's measure, but at that time (1922) Lévy's methods were familiar only to a very few specialists, and Daniell's integral appeared so sophisticated that Wiener felt obliged to project his new measure on the better understood Lebesgue measure. Wiener was able to take account of the observable properties of Brownian motion and its expected mathematical behavior: independence of the increments of position  $X(t) - X(s)$ , which follow a Gaussian law with standard deviation  $\sqrt{t - s}$ , and non-differentiable trajectories. Wiener also identified other more subtle mathematical properties such as the almost sure Hölder continuity with exponent  $\frac{1}{2} - \varepsilon$  of the trajectories in a given interval.

By 1923 Wiener had constructed the first mathematical model of Brownian motion. In a certain sense this was the first model of a stochastic process in continuous time, ten years before Kolmogorov's *Grundbegriffe* (Sect. 3). A new mathematical

---

<sup>91</sup>For details on this transition from Gateaux and Lévy's works to Wiener's, see Mazliak (2011).



dimension was given to Brownian motion in the 1930s by Wiener's work with Paley and Zygmund. This introduced a second mathematical formulation making use of the Paley-Wiener series

$$X(t) + iY(t) = \frac{\xi' + i\eta'}{\sqrt{2\pi}} + \sum_1^{\infty} \frac{(\xi_n + i\eta_n)(\cos nt - 1)}{\sqrt{n\pi}} + \sum_1^{\infty} \frac{\xi'_n + i\eta'_n}{\sqrt{n\pi}} \sin nt$$

which represents the motion as a trigonometric series with random independent Gaussian coefficients. This series had been introduced by Wiener in a paper in 1924, with the collaboration of Lévy. Wiener had visited Lévy at Pougues-les-Eaux during the summer of 1922. Lévy describes this pivotal encounter in Lévy (1970), relating how Wiener became familiar with the theory of countable probabilities on which Lévy, following Borel, focused. As a result, Wiener was able to offer a new presentation of *Differential Space* (Wiener 1923) in the *Bulletin de la Société mathématique de France*.<sup>92</sup> On the first page of that article, written in French, Wiener “thanks M. Paul Lévy for his help in editing the French version of this article.” This is followed by the representation of the motion as a trigonometric series and a proof of its equivalence to the representation in *Differential Space* (Wiener 1923). Thus, Lévy was involved alongside Wiener in the mathematical history of Brownian motion, at the time of Wiener's first publications. In 1934 Lévy returned to work on Brownian motion again and began to combine it with his theory of additive processes. Then the motion appears naturally via the decomposition of such processes into three parts, with Wiener's function being the continuous part. In 1934 Lévy published two notes and a supplement on these matters.<sup>93</sup> He further refined these articles in 1937 (Lévy 1937).

Lévy began his reformulation of Brownian motion with the idea that it could be a paradigm for random functions. For Lévy, a random function was indeed a function, but one that is created by means of an intervention of randomness at each instant. To account for this development through the existence of a stochastic process, it is necessary to specify an effective procedure for construction of the process. This is quite different from Wiener's approach, which *a priori* takes the space of trajectories to be a space of continuous functions and constructs a measure that incorporates the required properties, even if he uses the theory of countable probabilities for that aim. As culminating point of the reformulation that he began in 1934, Lévy proposed the “constructive definition of Brownian motion,” which consisted of constructive procedures using successively finer approximations. At each step the approximation employs a finite number of random points, for which the joint laws satisfy the properties of independent increments. In the limit, the required properties of Wiener's function remain satisfied and continuity, assumed *a priori* by Wiener, is assured. In

<sup>92</sup>Wiener, Norbert (1924) Un problème de probabilités dénombrables. Bull. Soc. Math. Fr. 52:569–578.

<sup>93</sup>Lévy, Paul (1934) Les généralisations de l'espace différentiel de N. Wiener. CRAS (February 26); Lévy, Paul (1934) Sur les espaces *V* et *W*. CRAS (March 26); and Lévy, Paul (1934) Complément à l'étude des espaces *V* et *W*. CRAS (May 7).



articles in 1940 and 1948 (Lévy 1940 and Lévy 1948, respectively), the interpolation formula is the only constructive method that Lévy retains. This interpolative method is related to the first works by Bernstein on the Kolmogorov problem in the 1930s, i.e. the construction of all Markov processes with continuous trajectories (see Sect. 8).

Lévy’s beautiful constructive definition of Brownian motion reveals many properties in arbitrary dimension that he would discover later. Lévy worked with subdivisions associated with the dyadic numbers in  $[0, 1]$ . If  $X(t)$  is the function of linear Brownian motion, and  $t_0 < t_1 < t_2$  are three values of  $t$  in the interval under consideration, given  $X(t_0)$  there are two equivalent stochastic methods to determine  $X(t_1)$  and  $X(t_2)$ . Either one determines the two successive increments  $X' = X(t_1) - X(t_0)$  and  $X'' = X(t_2) - X(t_1)$  independently, or first finds the total increment, and then interpolates, that is determines  $X(t_1)$  according to the conditional law for that variable when  $X(t_0)$  and  $X(t_2)$  are known. This conditional law is a Gaussian of mean  $m_1 = \frac{(t_2-t_1)X(t_0)+(t_1-t_0)X(t_2)}{t_2-t_0}$  and variance  $\sigma_1^2 = \frac{(t_2-t_1)(t_1-t_0)}{t_2-t_0}$ . Thus one can construct a conditional linear interpolation by writing

$$X(t_1) = [X(t_1) - E(X(t_1)/X(t_0), X(t_2))] + E(X(t_1)/X(t_0), X(t_2))$$

and exploiting the independence of the two terms in the sum, the first having variance  $\sigma_1^2$  and the second having the mean  $m_1$  given above.

It is clear that if the two successive time intervals defined by the three values of  $t$  above have the same length, then the difference between  $X(t_1)$  and its conditional expectation given  $X(t_0)$  and  $X(t_2)$  is a Gaussian variable of standard deviation  $\sqrt{2}$  times smaller than the conditional expectation given  $X(t_0)$  only (i.e. a simple extrapolation). This suggests good properties of convergence: given  $X(t_0)$ , one may determine  $X(t_1)$ , then, by successive interpolations,  $X(\frac{1}{2})$ ,  $X(\frac{1}{4})$  and  $X(\frac{3}{4})$ ... and continue dyadically. Let  $(X_n(t))_{n \geq 1}$  be the sequence of continuous functions obtained by this series of interpolations, and  $\delta_n$  the maximum of  $|X_n(t) - X_{n+1}(t)|$ . Due to the Gaussian character of this difference, one finds  $P(\delta_n > (\frac{1}{\sqrt{2}})^{n+2}y) \leq \frac{2^{n+1}}{y\sqrt{2\pi}}e^{-y^2/2}$ . Letting  $y = c\sqrt{2n \ln 2}$  with  $c > 1$ , one sees that this probability is a term in a convergent series. Lévy concluded that

according to Cantelli’s lemma (or Borel’s, when the  $\delta_n$  are independent),<sup>94</sup> there exists almost surely a finite number  $N$  such that, for  $n > N$ , one has  $\delta_n < c \frac{1}{(\sqrt{2})^{n+1}} \sqrt{n \text{Log } 2}$ , which establishes the almost sure uniform convergence to an evidently continuous limit  $X(t)$  on  $[0, 1]$ . It is easy to check that  $X(t)$  obtained this way verifies all the conditions of the theoretical definition....

Lévy’s proof is an exemplary moment in the history of the theory of probability. It is the most direct and most probabilistic way to both the construction of Brownian motion and the continuity of the trajectories, i.e. to standard Brownian motion. It also embodies Lévy’s understanding of randomness and it combines Lévy’s entire

---

<sup>94</sup>The correspondence makes clear that when using the Borel-Cantelli lemma, Lévy always emphasized the advantage of Borel’s result over Cantelli’s in the independent case.

and unvarying notion of stochastic processes: the influence of randomness at every instant, and its incorporation in a concrete, constructive algorithm. This proof is extensively discussed in the treatise of Blanc-Lapierre and Fortet (1953) intended for engineers and physicists, where it sets the stage for introducing the most useful properties of Brownian motion. It is also the origin of practical algorithms. Lévy's construction is no doubt the only one possible in a work combining mathematical theory with the applications, calculations, and models of engineers and physicists. Itô, who read this proof in Lévy (1948), wrote in 1965 in the first pages of Itô (1965) that he only retained two formulations of standard Brownian motion, Wiener's and Lévy's, and that each had its advantages. In 1940 Lévy not only put his formulation at the start of Lévy (1940), but drew many consequences for Brownian motion from it. He certainly experienced a sense of victory, but later in life it was mixed with remorse: in 1970 he wrote of his regret "to have left to Wiener the discovery of the function  $X(t)$ ," adding "I think, incidentally, that my manner of defining  $X(t)$  for Brownian motion is clearer than Wiener's."

Following his insightful work on linear Brownian motion and Brownian motion in a plane (which, because of its singular nature, requires special treatment), Lévy continued his research, investigating Brownian motion with several parameters and Brownian motion with values in a Hilbert space. He also considered Brownian motion in various geometries (on a Riemann sphere, using the work of Francis Perrin and even to some extent in hyperbolic space). Loève (1973) notes the surprise with which, after the war, Lévy discovered the deterministic nature of Brownian motion in a Hilbert space with a parameter; Lévy mentions this to Fréchet in the Letter 58.

At the Second Berkeley Symposium on Probability and Statistics in 1950, shortly before his 65th birthday, Lévy announced to applause his discovery of the Lévy area law (Sect. 8). This victory came after fifteen years of effort by the then incontestable master of Brownian motion. From 1934 to 1950 practically all the advances in the subject were due to Lévy: characterizations, detailed studies of trajectories, and analysis of the geometrical properties of the trajectories. In addition Lévy made innumerable interesting remarks, later amplified. It is clear why Brownian motion imbued his correspondence with Fréchet during these years.

## 8 Lévy and Stochastic Integrals

The stochastic integral is a major theme in the correspondence of Lévy and Fréchet, especially during the Second World War. The following brief history of stochastic calculus from 1928 until the end of the 1950s is limited to those developments explicitly or implicitly referred to by Lévy in his letters.

Between 1934 and 1950 Lévy invented, investigated and used various types of stochastic integrals. In 1950 he showed that many processes could be represented in the form of an integral with respect to Brownian motion, and then studied these representations. In his 1964 Notice sur les travaux scientifiques de Paul Lévy in support of his election to the Academy of Science, Lévy attached the greatest importance to a certain type of integral that in 1939 had allowed him to define the

area determined by an arc of Brownian motion and its chord (Sect. 7). It was this kind of integral that he wished had been denoted *stochastic*, and in Lévy (1970) he regrets that this did not come to pass. The usage established after the Second World War is that an integral  $\int f dg$  is called stochastic if at least one of the two functions  $f$  or  $g$  is random, and if it satisfies certain regularity properties. The construction, study and interpretation of these integrals is closely tied to the need for a stochastic calculus suitable for applications (especially in statistics and physics). However purely mathematical considerations arising from both the extension of functional analysis to random functions and the needs of probability theory itself played an essential role in the creation of various versions of stochastic calculus. For Lévy, as can be seen in his exchanges with Fréchet, an integral  $\int f dg$  should only be called stochastic if it can be defined on the individual trajectories of Brownian motion as the stochastic limit of Stieltjès sums estimated at random times, a point of view that Lévy expressed in his correspondence. Thus, Lévy's stochastic integrals appear to the modern reader as doubly stochastic, and it should come as no surprise that the integral  $\int f dg$  which defines the Lévy area has a meaning on the trajectories. There is no paradox; the Lévy integral is not the Itô integral!

We begin our account in 1928, when Slutsky constructed a whole class of stochastic processes for the first time, those that today are called second order processes. He clarified the possible types of convergence of probabilities and elucidated different notions of continuity, differentiability, and stochastic integrability. In fact the terms stochastic derivative and stochastic integral have their origin in Slutsky's work.<sup>95</sup> For second order processes, it is natural that convergence in quadratic mean plays a privileged role. Slutsky's stochastic integral is thus of the form  $\int x(t) dt$  for  $x(t)$  continuous in quadratic mean. Integrals of this type, defined by convergence in quadratic mean of Riemann or Riemann-Stieltjès sums, have been called Riemann-Stieltjès stochastic integrals since the 1950s (Blanc-Lapierre and Fortet 1953), with  $\int x(t) dt$  defined as the limit in quadratic mean of Riemann sums. In 1937 Slutsky gave more general sufficient conditions on the trajectories that allow a definition of the integral.

Wiener is given credit for a first approach to stochastic integration with respect to Brownian trajectories—even though he never used the term—for square integrable non-random functions. Doob and Itô traced the origin of this integral to *Differential Space*, Wiener's seminal article of 1923 laying out his original theory of Brownian motion (Sect. 7). However its explicit introduction came in 1933 and 1934 in Wiener's *Notes on Random Functions*<sup>96</sup> and Chap. IX of his book *Fourier Transforms in the Complex Domain*,<sup>97</sup> and represents the major break between ordinary

---

<sup>95</sup>In his important note Slutsky, Eugène (1928) Sur les fonctionnelles éventuelles continues et dérivables dans le sens stochastique. CRAS 187:878–880, Slutsky defined and constructed “regular random functions,” i.e. the processes called second order today. He showed that a quadratic mean existed as well as an integral which he wrote as  $(S) \int_a^b x_t dt$  and called the *stochastic integral*.

<sup>96</sup>Wiener, Norbert (1933). Math. Z. 37:647–668.

<sup>97</sup>Wiener, Norbert (1934) Colloq. Publ.-Am. Math. Soc. XIX. Most of this text was written with Raymond Paley who died in a skiing accident in 1933.

and stochastic integration. If neither  $f$  nor  $g$  are of bounded variation—and in 1923 Wiener had shown that Brownian trajectories are not of bounded variation—the integral  $\int f dg$  has no possible interpretation in Lebesgue-Stieltjès theory.<sup>98</sup> Wiener dealt with this problem, giving in the process two successive but equivalent definitions of stochastic integration. The 1933 version begins with the class of functions  $f$  of bounded variation and defines its integral with respect to  $g$  via integration by parts, then extends the definition to the class of continuous functions. The 1934 version, in the style of Wienerian folklore, uses the Paley-Wiener series (Sect. 7). The formal primitives of that trigonometric series have good convergence properties and permit one, via integration by parts, to obtain integrals of trigonometric polynomials. An extension to  $L^2$  using Hilbert space techniques then follows. These techniques are common in mathematical analysis today, but at the time were rarely used in probability theory, which was not yet completely embedded in the framework of measure theory. Wiener had therefore to make use of the theory of Rademacher functions.

Wiener showed that the limiting object, defined either way, exists on almost all trajectories. In both of these versions,  $\int f dg$  (with  $f \in L^2$  and  $g$  Brownian motion) is nothing but a random function according to Wiener's theory, satisfying the requirements of Einstein and Smoluchowsky (Sect. 7). Its construction is thus strongly influenced by statistics and physics and can be understood as an isometry between the space  $L^2$  and the Gaussian space of Brownian increments.

As of the years 1942 to 1944 Itô originated stochastic calculus with respect to Brownian motion. This work became the basis of theories of integration with respect to martingales. Itô's 1944 article *Stochastic Integral* remains the most well-known of the period. Its only references are to Doob, Lévy and Wiener. In a subsequent letter to Lévy, Itô explained that his goal had been to give an integral sense to the differential expression  $dX(t) = A(t, X(t))dt + B(t, X(t))dg(t)$  (where  $g$  is a Brownian motion). Itô and his colleague Kakutani, following Doob, put themselves squarely in Kolmogorov's camp, contrary to Lévy (Sect. 3). Itô's formula

$$\int_0^t F'(g(\tau, \omega)) dg(\tau, \omega) = F(g(t, \omega)) - F(g(0, \omega)) - \frac{1}{2} \int_0^t F''(g(\tau, \omega)) d\tau$$

figures explicitly in *Stochastic Integral*; Itô considered it as expressing the essential difference between ordinary and stochastic integration.

The construction of Itô's stochastic integral can be described as follows. Riemann sums define the integral of simple processes defined on subdivisions of  $[0, 1]$  that before the time  $t$  ( $0 < t < 1$ ) are independent of the Brownian increments after  $t$ . Closure of these simple processes under the norm  $L^2([0, 1] \times \Omega)$  gives the integral of processes  $L^2([0, 1] \times \Omega)$ , independent of the time  $t$  of the Brownian increments after  $t$ . By weakening the topology, Itô was finally able to extend the integral to the class of processes that are sufficiently close in probability to the preceding class. Itô's integral clearly makes no sense on individual trajectories; Lévy subsequently

---

<sup>98</sup>See Sect. I.8, p. 43 ff, *Naive Stochastic Integration is Impossible* in Protter (1990).

called it a global procedure by contrast to his integral, which may be defined on trajectories.

With the resumption of relations between Japanese scholars and the international community after the war, Doob made note of the extraordinary versatility of Itô's integral. Doob's remark that  $g^2(t) - t$  is a martingale together with the introduction of classical conditional expectation values in the theory of martingales opened the way for the modern theory of stochastic integration. Lévy was at the origin of the theory of martingales<sup>99</sup> and recognized their appearance in Brownian motion. It is therefore puzzling that he did not even try to use them to attack stochastic integration theory (see Dellacherie 1980), and sometimes found himself among those resisting this approach.

The first contact between Lévy and Itô was an exchange of letters in April 1954. Decades later, after receiving a prize from the Kyoto Academy in 1998, Itô described the difficulties that confronted him and the extent of his isolation during the long years when no one showed interest in his research. He greatly admired Lévy, who motivated his work and inspired his espousal of probability at the start of the war. When an article by Lévy arrived in Japan<sup>100</sup> Itô was in awe. However he was completely ignorant of Lévy's integrals, the essential account having been published in an American journal in 1940.<sup>101</sup> He only found out about them with the appearance of Lévy (1948) which gave a rather abbreviated account of Lévy's work on stochastic integrals during the war and was never completed, Lévy being absorbed by questions about Markov processes (Sect. 4). In 1934, as Lévy began his reformulation of Brownian motion (Sect. 7), the first stochastic integrals appeared, which in 1934 Lévy named 'integrals whose elements are independent random variables.'<sup>102</sup> Lévy's book (Lévy 1937) brought notice of his integrals to a much broader circle than his habitual audience. It contained equations of the form  $X(T) = \int_0^T \xi_t \sqrt{dt}$  and  $Y(T) = \int_0^T \xi_t \sqrt{d\sigma^2(t)}$ , where the positive function  $\sigma^2$  is non-decreasing in  $t$  and the  $\xi_t$  are independent Gaussians. These were extraordinary, due to the presence of  $\sqrt{dt}$ <sup>103</sup> and became somewhat notorious, even being attributed to some eccentricity of Lévy's (see, especially, the remarks of Schwartz<sup>104</sup>). In fact, they are

<sup>99</sup>On this topic see Mazliak (2009).

<sup>100</sup>According to certain accounts, including Itô's notices of the American Mathematical Society ((1998) 45:1455), Lévy's papers arrived via Romania.

<sup>101</sup>Lévy's American article, in which his stochastic integral is defined, is included in the bibliography of an article by Kakutani (Kakutani, Shinzo (1944) On Brownian motion in  $n$ -space. Proc. Imp. Acad. (Tokyo) (November 20)). Lévy's results apparently arrived in Japan from the USA before Pearl Harbor.

<sup>102</sup>Lévy, Paul (1934) Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. Ann. Sc. Norm. Super. Pisa 3:337–366.

<sup>103</sup>The notation  $\sqrt{dt}$  had in fact been introduced by Lévy in various publications and appeared in Lévy (1937) where it appears in a note to the chapter entitled *Les intégrales à éléments aléatoires indépendants*.

<sup>104</sup>Number 157–158 of *Astérisque* (1988).

already implicitly present in the elements of a Wiener space  $W$ . Decomposing additive processes Lévy showed that the continuous component is Wiener Brownian motion. Eliminating components of Poisson type and the various jumps, a Brownian component may be recovered by a procedure closely related to Lévy's 1934 study of random series with independent terms or with terms forming a Markov chain. Sums of weighted increments  $\Delta X = \xi_t \sqrt{\Delta t}$  on nested subdivisions with step  $\sqrt{\Delta t}$ , when properly reordered into the terms of a series, almost surely approach a limit  $X(T)$  independent of the choice of the subdivisions. A generalization leads immediately to  $Y(T)$ , a Gaussian process with independent increments of variance  $\sigma^2$ , which Lévy called a generalized Brownian motion. The time change  $\tau = \sigma^2(t)$  then connects  $Y(T)$  to a standard Wiener-Lévy process. Lévy had followed Bernstein's work on diffusion during the 1930s. In particular he knew very well the connections between the equation  $dZ(t) = A(t, Z(t))dt + B(t, Z(t))\xi_t \sqrt{dt}$  satisfied by a process  $Z$  and the diffusion equation satisfied by its probability density  $\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 (B^2 u)}{\partial x^2} - \frac{\partial (Au)}{\partial u}$ . In Lévy (1948), enriched by his deliberations during the war, he again took up his work on diffusion and stochastic differential equations. His principal sources were two memoirs by Bernstein<sup>105</sup> and certain of Kolmogorov's results.

Lévy's approach to the method of Cauchy-Bernstein for solving stochastic differential equations required that the convergence of the Cauchy approximations not only occur in distribution, as for Bernstein, but also with respect to spatial topologies important for probability calculus, especially almost sure convergence. The connections between stochastic differential equations, diffusion equations, and stochastic integrals are not so easy to detect in Lévy (1948) because these topics arose in various places, as they were needed. Skorokhod's 1961 book<sup>106</sup> recognizes Lévy as the first to elucidate the methods of Bernstein for differential equations and to be able to study the trajectories of the solutions. In this context an integral of a non-random function  $B(t)$  with respect to a Brownian motion  $X(T) = \int_0^T B(t)\xi_t \sqrt{dt}$  may be recast as a solution of the stochastic partial differential equation  $dX(t) = B(t)\xi_t \sqrt{dt}$ . Further, the equation  $\frac{\partial u}{\partial t} = \frac{1}{2} B^2(t) \frac{\partial^2 u}{\partial x^2}$  shows that this  $X$  is an additive process that becomes Brownian motion<sup>107</sup> on redefining the time via  $\tau(t) = \int_0^t B^2(s) ds$ .

At the end of the 1930s Lévy had little use for Wiener's integral, which remained poorly understood by probabilists and considered by many analysts as useful only for constructing counter-examples in the theory of trigonometric series. As December 1938 drew to a close Lévy immersed himself in a study of the geometry of the planar Brownian curve. Between 1939 and 1941 he produced six articles on what he called in Lévy (1940) "stochastic integrals of an entirely new type." Lévy's view

<sup>105</sup>Bernstein, Serge (1934) *Principes de la théorie des équations différentielles stochastiques*. Труды Физического Математического Института Стеклова 5:95–124 and Bernstein, Serge (1938) *Equations différentielles stochastiques*. Hermann, Paris.

<sup>106</sup>Скоруход Анатолий В. (1961) *Исследования по теории случайных процессов*. Киевский Университет, Киев. English translation (1965): *Studies in the theory of random processes*. Addison-Wesley.

<sup>107</sup>Lévy (1948, pp. 71–72).

of Brownian motion was built around scale invariance and stochastic self-similarity. Starting at the beginning of 1939 Lévy used these fractal properties and stochastic measures to establish that the area measure of the curve is almost surely zero. He believed thus that the area defined by the arc of the curve and its cord should exist in some sense, similar to the area defined by those extraordinary curves without tangents that had fascinated him since his youth. He constructed entire classes of fractals, whether stochastic or not, that served as models for Brownian motion and reproduced some of its properties. He calculated the characteristic functions of the corresponding Lévy areas. Some of his examples have entered fractal folklore (Lévy flights, Lévy dust, Lévy dragon, etc.)—see Mandelbrot (1975). The bends and loops imposed by randomness on the trajectories of Brownian motion down to the smallest distances imply that defining the area by the usual Green-Riemann formula  $A(t) = \frac{1}{2} \int_0^t \{X dY - Y dX\} = \int_0^t X dY$  where  $X$  and  $Y$  are the coordinates of the two dimensional Brownian motion makes no sense in the usual theory of integration. To circumvent this problem, Lévy introduced another independent randomness in the time scale with a probability  $P'$ . The corresponding Stieltjès sums  $S_n = \sum \frac{f(t_{i+1})+f(t_i)}{2} [g(t_{i+1}) - g(t_i)]$  are calculated on nested random subdivisions of the integration interval  $[a, b]$  with a random step tending  $P'$ -almost surely to zero. Between 1939 and 1940, in several stages and with great finesse, Lévy proved the following assertions:

- (a) Given a series  $\{t_n\}$  that is dense in the integration interval  $[a, b]$  and with  $t_0 = a$ , the series of Stieltjès sums  $\{S_n\}$  of the form given above converges almost surely in quadratic mean to a limit  $S$  independent of the series  $\{t_n\}$ .
- (b) Random independent choices of the points  $t_n$ , each following a uniform distribution on  $[a, b]$  will lead almost surely to a dense sequence in  $[a, b]$ .
- (c) In the computations, one may exchange the probability  $P'$  used for the time scale and the probability  $P$  for the trajectories by application of Fubini's theorem. This theorem had always been emphasized by Lévy as a fundamental tool for the study of random processes.

The formal integral that gives the Lévy area is thus defined on  $P$ -almost every trajectory with a  $P'$ -almost sure limit. In Lévy's articles from 1938 to 1941 stochastic integrals of an entirely new type are defined by a series of areas of polygons with vertices chosen randomly on non-random curves. It is clear that this construction corresponds to the constructions leading to limiting objects which are among the curves without tangents of which each part is similar to the whole, i.e. that possess self-similar properties. This early vision of fractals may be found in a remarkable article of 1941<sup>108</sup> in which Lévy compares his integrals with the features that one can observe with more and more powerful microscopes.

During the Second World War and up to 1954, Lévy worked at extending his method of stochastic integration. He defined stochastic integration for non-random functions by applying the 0–1 law to the superior and inferior limits of random

---

<sup>108</sup>Lévy, Paul (1941) Intégrales stochastiques. Bull. Soc. Math. Fr.-Sud-Est 67–74.



Stieltjès sums regarded as asymptotic events. In Lévy's parlance (which he even used in his notes to the CRAS), a function is *stoch-int* if the two limits exist and are equal with probability one. If  $f$  is Lebesgue integrable, it is also *stoch-int* in Lévy's sense. In addition Lévy compared the range of application of his stochastic integrals of the form  $\int f dg$ , with  $f$  and  $g$  continuous and non-random, with that of Young's integral and of various non-stochastic extensions of Stieltjès integrals. Lévy also differentiated various modes of convergence of the Stieltjès sums defining his integrals, with the terminology *stoch-int a.s.*, *stoch-int Pr* (for convergence in probability), etc. The Lévy area thus appears on almost every Brownian trajectory as a Stieltjès integral that is stochastic in Lévy's sense. Encouraged by this result, Lévy expected that integrals  $\int f dg$ , where both  $f$  and  $g$  are random processes, would be stochastically well-defined in his sense on almost every trajectory.

Lévy's work on Gaussian processes, which he presented in 1950 in Berkeley, used an integral (that he was careful not to call stochastic) of the form  $\int_0^t F(t, u) f_u \sqrt{du}$ , later rewritten as  $\int F(t, u) dY(u)$  where  $Y(u)$  encodes Brownian motion. This integral is defined as the limit of a Riemann sum on partitions of the integration interval, the sum being a Laplacian variable well-defined by its covariance matrix. The limit is thus Gaussian and its covariance is obtained from the kernel  $F$ . In 1950 Lévy also gave an interpretation of the Lévy area using a generalization of Wiener's integral. Later developments in stochastic calculus showed, however, that it was Itô who had uncovered the Brownian properties which best allowed an extension of stochastic integration by using martingales.

In 1954, at a time when he would have no further occasion to work on stochastic integration, Lévy gave his personal view of its history and summarized his own theory, which he never ceased to defend:

Our stochastic integrals have nothing in common with what various authors, since Slutsky, refer to under that name. Ours is a definition that applies to individual functions. We use it, in particular, for each instance of a random function. Such a definition, to our knowledge, cannot be found before 1940 except in Paley-Wiener. . . . However, these authors start with the representation of a complex function  $g(t) = Z(t)$  by the Fourier-Wiener formula. For the general theory, it is interesting to start instead with the intrinsic definitions of  $f(t)$  and  $g(t)$  and Riemann-Stieltjès sums. Since 1944, a theory that uses Riemann sums has been developed by K. Itô,<sup>109</sup> who replaces our sum  $\sum \frac{f(t_{i+1})+f(t_i)}{2} [g(t_{i+1}) - g(t_i)]$  by  $\sum f(t_i) [g(t_{i+1}) - g(t_i)]$  and does not attempt to find a definition applicable to individual functions.

Thus, although Lévy's name is nowadays not immediately associated with stochastic integration, there is abundant evidence of his deep and protracted involvement in the development of that field.

---

<sup>109</sup>In his seminal article (Itô, Kyoshi (1944) Stochastic integrals. Proc. Imp. Acad. (Tokyo) 20), Itô integrates certain simple processes with respect to Brownian motion. These Riemann sums were the object of an exchange of letters between Itô and Lévy in April 1954. Itô's construction, especially useful in the space  $L^2$ , was later employed for the development of a theory of stochastic integration with respect to martingales. See Dellacherie (1980).



Paris - 29 décembre 1918  
 9 rue Chernoviz


Monsieur et cher collègue,

Ayant eu votre adresse par M. Hadamard, je me permets de vous demander s'il vous serait possible de me faire parvenir un exemplaire de votre mémoire sur une extension de la méthode de Jacobi Hamilton. Il me serait utile pour la préparation du cours Peccot, que j'espère faire cette année.

Les *Annali di Matematica* sont peu répandues en France, et étant mobilisé, il ne m'est pas facile d'aller travailler à la bibliothèque de la Sorbonne, qui ne prête pas les périodiques. Pous m'excuserez de m'adresser à vous dans ces conditions.

Prenez, je vous prie, à mes sentiments dévoués

Paul Lévy



**Fig. 1** First letter from Lévy to Fréchet (29 December 1918). (Photo: © Académie des Sciences, Paris)

Paris - 38 Av. Th. Gautier  
19-10-65

Mon cher collègue et confrère,

La Commission du tricentenaire s'est réunie hier après le Comité secret (jusqu'à 18 h. 15 !). Il a été décidé que chaque Section devrait prendre la responsabilité du travail fait en son nom.

Montel, Garnier et moi avons tenu une réunion préliminaire. Denjoy est dans le Midi. Je prévins Julia, qui voudra peut-être nous réunir le mercredi 27 après la séance (rennis au mercredi à cause de la séance de l'Institut du lundi). Je crois bien faire de vous prévenir de cette éventualité.

J'ai dû exposer longuement à Julia les critiques auxquelles mon travail a donné lieu. Il est trop long, et je n'aurais pas dû être les vivants ! Excusez-moi de ne pas vous donner plus de détails. Il est trop tard pour tout recommencer.

Bien cordialement

P. Lévy




Fig. 2 Letter from Lévy to Fréchet (19 October 1965). (Photo: © Académie des Sciences, Paris)

**Fig. 3** Maurice Fréchet in 1910. (Photo: © Private collection F. Lederer)



**Fig. 4** Suzanne and Maurice Fréchet and their children Hélène and Henri in 1915. (Photo: © Private collection F. Lederer)



**Fig. 5** Paul Lévy in 1913.  
(Photo: © Private collection  
D. Piron)



**Fig. 6** Paul Lévy in 1933.  
(Photo: © Private collection  
D. Piron)

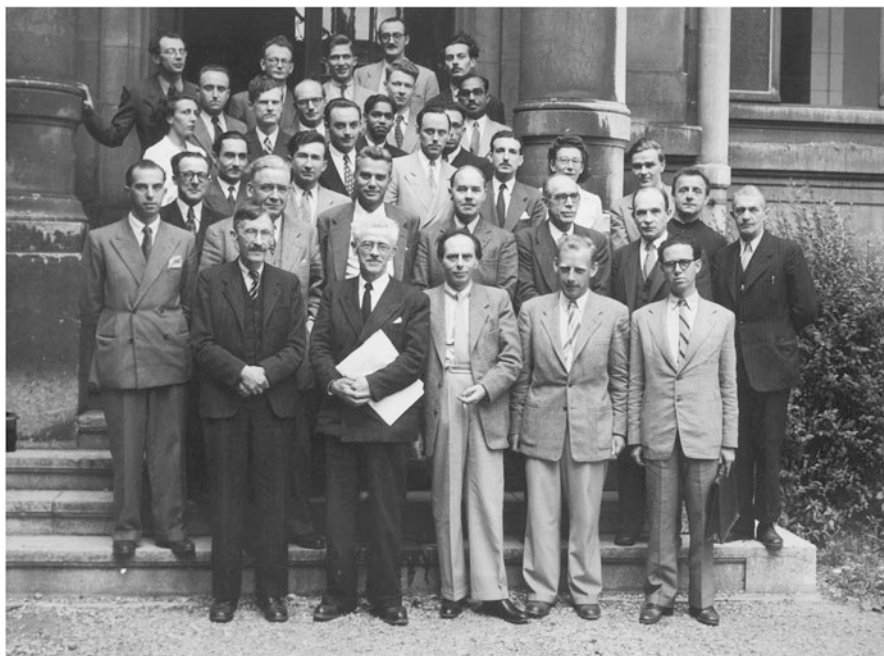


**Fig. 7** Suzanne and Paul Lévy in 1952. (Photo: © Private collection D. Piron)



**Fig. 8** Maurice Fréchet in 1942. (Photo: © Private collection F. Lederer)





**Fig. 9** Colloque International sur le Calcul des Probabilités, Lyon 1948. First row: Paul Lévy and Maurice Fréchet. On the picture one can find among others J. Doob, R. Fortet, D. Van Dantzig, E. Mourier, J. Kampé de Fériet, A. Blanc-Lapierre. . . . (Photo: © Private collection F. Lederer)

# 107 Letters from Paul Lévy to Maurice Fréchet

## Lettre 1

Paris- 29 décembre 1918

9 rue Chernoviz

Monsieur et cher collègue,

Ayant eu votre adresse par Monsieur Hadamard,<sup>1</sup> je me permets de vous demander s'il vous serait possible de me faire parvenir un exemplaire de votre mémoire sur une extension de la méthode de Jacobi-Hamilton.<sup>2</sup> Il me serait utile pour la préparation du cours Peccot,<sup>3</sup> que j'espère faire cette année. Les *Annali di Matematica* sont peu répandues en France, et étant mobilisé, il ne m'est pas facile d'aller travailler à la Bibliothèque de la Sorbonne, qui ne prête pas les périodiques. Vous m'excuserez de m'adresser à vous dans ces conditions.

Croyez je vous prie, à mes sentiments dévoués.

Paul Lévy

## Lettre 2

Paris- 6 janvier 1919

9 rue Chernoviz

---

<sup>1</sup>From the outset, the correspondence between Lévy and Fréchet took place under the intellectual sanction of their mentor Hadamard. In 1918 Hadamard was 54, and one of the most influential French mathematicians. Throughout his long life (he died in 1963 at age 98) Lévy and Fréchet were at all times alert to his approval or encouragement. Lévy suffered the consequences by losing favor from the moment he “wasted the heritage” by taking up the theory of probability (see Sect. 4).

<sup>2</sup>Fréchet, Maurice (1905) Sur une extension de la méthode de Jacobi-Hamilton. *Ann. Math. Pura Appl.* XI:187–199. In this article Fréchet studied the calculus of variations in  $n$  dimensions using the functional techniques of Volterra.

<sup>3</sup>The Cours Peccot, financed by the Peccot Foundation, began in 1900 with a three year series of lectures by Borel, who was followed by Lebesgue. The Cours was designed to give mathematicians under 30 a platform for the presentation of their results, and a small remuneration. It is the only course at the Collège de France remaining from that period. Lévy's book *Leçons d'Analyse Fonctionnelle (Lessons in Functional Analysis)*, published by Gauthier-Villars in 1922 (see the letters of 1919) resulted from his lectures.

Mon cher collègue,

Je ne suis pas du tout surpris que vous n'ayez pas vos mémoires sous la main. Je vous avais demandé l'un d'eux que pour le cas où vous pourriez me l'envoyer sans dérangement. Mais ne vous inquiétez pas ; je pourrai m'arranger pour le trouver. Je vais vous envoyer ma thèse et 3 mémoires publiés par le cercle de Palerme. Je puis pour vous éviter une perte de temps vous dire que je considère comme sans intérêt les § 13 à 16 de ma thèse,<sup>4</sup> que je regrette d'avoir écrits. Par contre je crois que ce qui peut vous intéresser le plus, si vous ne désirez pas approfondir l'application aux fonctions de Green, c'est mon dernier mémoire de Palerme (et aussi pour les éléments les § 9, 11 et 12 de ma thèse)<sup>5</sup>

Je vous signale aussi un mémoire paru dans le Bull.Soc.Math (1918) sur l'application à l'étude de la distribution de l'électricité ;<sup>6</sup> je n'en ai pas de tirage à part.

Je viens aussi de rédiger une note sur la généralisation d'un théorème de M. Hadamard sur les transformations ponctuelles. Mais avant de la présenter, je voudrais être d'accord avec vous sur un point sur lequel il me semble qu'actuellement je ne le suis pas tout à fait.<sup>7</sup>

Soit  $v = (A, B)$  le "voisinage" de deux éléments d'une classe  $(V)$ .<sup>8</sup>

Par définition si

$$v(A, B) < \varepsilon, \quad v'(B, C) < \varepsilon$$

---

<sup>4</sup>Lévy is referring to his work on differentials of order greater than 2, which he did not use in the rest of his thesis. Perhaps he regretted this part of his thesis because he saw it as a pointless digression.

<sup>5</sup>At the request of the Circolo di Palermo, Lévy had written a long article Lévy, Paul (1914) Sur l'intégration des équations aux dérivées fonctionnelles partielles. Rend. Circ. Mat. Palermo 37:1–56 giving the principal results of his thesis.

<sup>6</sup>Lévy, Paul (1918) Sur la variation de la distribution de l'électricité sur un conducteur dont la surface se déforme. Bull. Soc. Math. Fr. 46:35–69. Here, Lévy applied the theory of Green functions which he had developed after writing his thesis.

<sup>7</sup>This note was never submitted, since Lévy realized that he had read Fréchet too quickly. (See the next letter.)

<sup>8</sup>Fréchet had introduced central ideas of general topology in his 1906 thesis, most of which was published in a long article Fréchet, Maurice (1906) Sur quelques points du calcul fonctionnel. Rend. Circ. Mat. Palermo 32:1–74. In particular, he introduced the idea of *écarts* (gaps), which soon after he denoted "distance" (see the next letter), in the classes  $(E)$ , and an intermediate idea, called class  $(V)$ , as follows. A set is a class  $(V)$  if for every pair of points  $A$  and  $B$  one may associate a positive number  $(A, B)$  called the *voisinage* (*neighborhood*) of  $A$  and  $B$ , such that  $A \neq B \iff (A, B) \neq 0$  and there exists a function  $f(\varepsilon)$  which goes to 0 with  $\varepsilon$  such that  $(A, B) \leq \varepsilon$  and  $(B, C) \leq \varepsilon$  imply  $(A, C) \leq f(\varepsilon)$ . This idea, clearly connected to the concept of precompactness, was later discarded, especially after the publication of an article by E.W. Chittenden (*On the equivalence of ecart and voisinage*. Trans. Am. Math. Soc. 18:161–166 (1917)) which demonstrated the existence of a distance yielding a topology equivalent to the class  $(V)$ . This idea of neighborhood disappeared from Fréchet (1928), where Fréchet defined the classes  $(V)$  completely differently, based on neighborhood. As is seen here, Lévy did not fully understand Fréchet's requirement that the same  $\varepsilon$  applied to the upper bounds on  $(AB)$  and  $(BC)$ , contrary to the case for a distance.



on a  $v''(A, C) < f(\varepsilon)$ .

Je dis qu'on peut trouver une fonction  $\varphi(v)$  assez lentement croissante pour que  $e = \varphi(v)$  vérifie la condition que vous imposez à l'écart. Donc contrairement à ce qui semble résulter de votre thèse, les notions de classe (V) et de classe (E) ne sont pas distinctes.

Pour le démontrer je considère la suite infinie dans les deux sens

$$\dots \varepsilon_{-2}, \varepsilon_{-1}, \varepsilon_0, \varepsilon_1, \varepsilon_2, \dots$$

$\varepsilon_0$  étant arbitraire et  $\varepsilon_{i+1} = f(\varepsilon_i)$ . Les  $\varepsilon$  d'indices négatifs très grands sont infiniment petits.

Je définis  $\varphi(v)$  pour  $\varepsilon_0 < v < \varepsilon_1$ , de manière que  $e$  varie de 1 à 2 quand  $v$  varie de  $\varepsilon_0$  à  $\varepsilon_1$ , et dans les autres intervalles de manière que

$$v' = f(v) \text{ entra\^i ne } e' = 2e$$

$e$  varie de 0 à  $\infty$  et vérifie toutes les conditions que vous imposez à l'écart dans une classe (E).

Sommes nous bien d'accord ? L'identité entre vos classes (V) et vos classes (E) a-t-elle déjà été signalée par d'autres ? Je serais heureux de le savoir avant de présenter à l'Académie ma note où je parle de l'écart dans une classe (V)- je voudrais être sûr de ne pas faire de confusion.

- Au sujet des travaux de Gateaux j'ai précisément appris hier que Monsieur Hadamard les avait mis en sûreté à l'Ecole Normale pendant la guerre et vient de les retirer. Rien n'est donc encore publié<sup>9</sup>

Bien à vous P.Lévy

### Lettre 3

Paris- 14 janvier 1919

Mon cher collègue,

Merci de vos deux lettres, et de vos mémoires que j'ai commencé à lire avec intérêt.

Votre objection est parfaitement juste. C'est moi qui ai eu une distraction singulière en ne m'apercevant pas qu'il y avait dans la définition de l'écart de votre thèse autre chose que le fait de faire  $f(\varepsilon) = 2\varepsilon$  dans la définition du voisinage. Il m'a paru évident que, quelque soit la définition initiale du voisinage, on pouvait

---

<sup>9</sup>René Gateaux died in October 1914, at the start of World War I. His highly original work enhancing Volterra's extension of differential calculus to functionals attracted considerable attention. His name is still attached to a definition of the differential that is more general than Fréchet's, since it does not require a metric structure on the space of functions—see Sect. 4. Hadamard, who was collecting Gateaux's unpublished work, asked Lévy to write several articles on it for publication in the Bulletin de la Société Mathématique de France (the first is in volume 47, p. 47 (1919)). At the time Lévy was working on a formulation of the Dirichlet problem for functionals, and was therefore a natural choice for this task. On April 14, 1919, Lévy presented a first note to the Comptes-Rendus, extending Gateaux's results on the Laplace equation. For a complete study of the case of Gateaux, see Mazliak (2011).

la modifier de manière que  $f(\varepsilon) = 2\varepsilon$ , et je ne me suis pas aperçu que j'énonçais beaucoup plus.

La question me paraît, comme à vous, beaucoup plus complexe que je ne l'avais cru d'abord. Il me semble qu'on doit pouvoir aisément prouver par des exemples que  $(A, B)$  étant le voisinage, on n'est pas sûr de pouvoir trouver une fonction  $\varphi$  telle que  $\varphi((A, B))$  soit un écart. Mon principe de démonstration ne s'applique donc pas. Je pense quand même, comme vous, que l'énoncé reste exact, mais je n'ose plus l'affirmer !

J'essayerai d'y repenser, mais je ne veux pas me laisser entraîner trop loin, car j'ai des préoccupations beaucoup moins générales, ou moins abstraites, que les vôtres, et ai surtout en vue des applications pour lesquelles ces questions ne se posent pas. Je ne considère que des ensembles de fonctions, et c'est pour éviter de préciser les restrictions de continuité que je leur impose ou la définition de l'écart, que je puis avoir intérêt à employer votre langage abstrait.

J'emploierai très volontiers votre nouvelle expression de distance,<sup>10</sup> d'autant plus que je trouve commode de représenter les éléments d'un ensemble par des points d'un espace idéal. Mais je me demande si pour quelques questions il n'y aurait pas intérêt à préciser encore cette définition, en supposant par exemple que,  $A$  et  $B$  étant donnés, l'égalité

$$(AB) = (AC) + (BC)$$

est réalisée et n'est réalisée que pour des éléments dépendant d'un seul paramètre formant entre  $A$  et  $B$  une ligne continue qu'on pourrait appeler droite. Pour les ensembles de fonctions, je particulariserais plus encore en supposant que la distance entre  $u$  et  $u + \lambda v$ ,  $\lambda$  étant une constante,  $u$  et  $v$  des fonctions, est indépendante de  $u$  et proportionnelle à  $\lambda$ .<sup>11</sup>

Pour préciser plus encore, la définition qui me paraît la plus intéressante est la suivante :

$$\left[ (A, B)^2 \right] = \int_0^1 (v - u)^2 ds, \quad \left. \begin{array}{l} A \text{ représentant } u(s) \\ B \text{ ----- } v(s) \end{array} \right\} 0 \leq s \leq 1$$

étant entendu non seulement que je mettrai la définition de limite en accord avec celle-ci, mais encore que je ne considérerai pas comme distinctes deux fonctions dont la distance serait nulle.

Ainsi deux fonctions intégrables ne différant que pour les points d'un ensemble de mesure nulle ne seront pas distinctes.

Mais je pense qu'en cherchant ces précisions je suis d'accord avec vous, car évidemment vous n'avez donné vos définitions générales que pour laisser à chacun la possibilité de les particulariser selon ses besoins.

<sup>10</sup>See note 8.

<sup>11</sup>It is interesting to note, some eighty years later, how much hesitation surrounded the notion of non-vectorial metric structures. The introduction of the  $L^2$  norm mentioned in the next sentence of this letter illustrates the tumult of those years concerning functional spaces tied to the then-new measure theory, which extended Lebesgue's ideas.

Vos considérations sur les intégrales, et le rôle des ensembles de fonctions également continues m'ont beaucoup intéressé.<sup>12</sup> Là encore, je cherche à préciser, sans le faire de manière trop artificielle, mais au point d'arriver à écrire par exemple des formules généralisant celle de Green dans la théorie de l'équation de Laplace au cas d'équations aux dérivées fonctionnelles.

Je pense par exemple limiter les oscillations ou irrégularités des fonctions en limitant une intégrale  $I$  telle que  $\int u'^2 ds$ , ou du moins en considérant comme "moins probables" les fonctions pour lesquelles  $I$  serait trop grand. Je ne puis préciser sans allonger beaucoup ces explications, mais je voulais vous indiquer dans quelle voie j'espère trouver une généralisation précise et maniable de l'intégrale définie. D'après ce que m'a dit M.Hadamard, peut-être y a-t-il des indications sur ce sujet dans les papiers de Gateaux.<sup>13</sup>

Bien cordialement à vous.

P.Lévy

PS- Je m'excuse d'avoir tardé jusqu'ici à vous envoyer mes mémoires. Mais maintenant que vous me le demandez j'attends encore

- Si vous passez par Paris, je serai très heureux de faire votre connaissance. Je vous indique à cet effet mon N° de téléphone, Auteuil 19-57 ; et que je suis en général l'après midi à l'ex-direction des Inventions, 23<sup>bis</sup> rue de l'Université, 6<sup>e</sup> étage et téléph Fleurus 05-71 (ou 72, 73, 74)<sup>14</sup>

#### Lettre 4

Paris rue Chernoviz

12 février 1919

Mon cher collègue,

Je profite d'une bronchite qui me retient en chambre pour répondre de suite à votre lettre du 7.

Je pense absolument comme vous que nous avons tout intérêt à maintenir le contact entre nous. Moi du moins j'y ai un intérêt évident. Vous avez établi en vue de

---

<sup>12</sup>Lévy is probably referring to Fréchet's article Fréchet, Maurice (1915) Sur l'intégrale d'une fonctionnelle étendue à un espace abstrait. Bull. Soc. Math. Fr. 43:248–265 which proposes an extension of Lebesgue's integral, augmenting Radon's work. The role of equi-continuous sets of functions mentioned is probably an allusion to one of the extensions of the Arzela-Ascoli theorem, which Fréchet proposed in his thesis and several articles after it.

<sup>13</sup>Lévy only found partial fulfillment of his expectations in Gateaux's work and was led to develop his own notion of an integral via the mean (see Sect. 4 and Mazliak 2011).

<sup>14</sup>On November 20, 1917, after the fall of the Painlevé government, the post of *Sous-Secrétaire d'Etat des Inventions* (Assistant Minister for Inventions) was dissolved. By ministerial decree it was reconstituted as the *Direction des Inventions, des Etudes et Expériences Techniques* (Directorate of Invention, Study and Technical Expertise), headed by Jules-Louis Breton. This agency came under the jurisdiction of the *Ministère de la Guerre* (Ministry of War) on November 26, 1918, and then under the *Ministère de l'Instruction Publique* (Ministry of Public Education) on April 14, 1919. Finally, on December 29, 1922, it was transformed into the *Office National de Recherche Scientifique Industrielle et des Inventions* (National Bureau of Industrial Scientific Research and Inventions), which merged in 1938 with the newly created CNRSA (see Roussel 1989 for more details).

certaines applications votre théorie des ensembles abstraits ; je suis un de ceux qui applique cette théorie, m'intéressant surtout aux ensembles de fonctions et je ne puis pas ne pas être intéressé par ce que vous avez fait.

La première fois que j'ai eu à utiliser la notion de distance, c'était pour préciser l'énoncé d'un théorème sur les fonctionnelles implicites. C'était relativement facile et vous pourrez voir dans les Comptes Rendus de janvier et dans les prochains bulletins de la Société Mathématique la solution à laquelle j'ai abouti.<sup>15</sup>

Depuis j'ai travaillé,<sup>16</sup> comme je crois vous l'avoir déjà écrit, à généraliser les notions d'intégrale multiple, et les théories relatives à l'équation de Laplace, les formules de Green. La définition de la distance qui permet de généraliser les théories du potentiel est, naturellement, toujours la même ;  $r^2 = \int (f - \phi)^2 dx$ , ou si on développe  $f$  et  $\phi$  en séries de fonctions orthogonales  $\sum (a_n - b_n)^2$ . Je suis d'accord avec Gateaux dans les papiers duquel j'ai trouvé la théorie du potentiel dans le cas de la sphère.

Pourtant dans cette même théorie, je ne puis, ou ne pouvais que difficilement, me passer d'une autre définition de la distance. Cela tient à ce que la notion d'intégrale dans l'espace fonctionnel ne peut guère être conçue que comme limite de l'intégrale à  $n$  dimensions, et il est important de légitimer le passage à la limite, ce que Gateaux ne me semble pas avoir fait, ce qui l'a obligé à limiter beaucoup ses théories. Je crois que je viens d'y réussir - peut-être arriverais-je à simplifier ce que j'ai fait. Voici où pour le moment j'en suis.

Pour cela il faut une définition de la distance jouissant de la propriété suivante :

- 1° Dans tout domaine fini (par rapport à la notion usuelle de distance), en excluant une portion du domaine dont la mesure soit une fraction très petite \* de celle du domaine total, on peut trouver une suite de fonctions en nombre fini telle que la distance à l'une d'elle d'une fonction quelconque de la partie restante du domaine soit inférieure à n'importe quel nombre donné d'avance
- 2° Les fonctionnelles usuelles sont uniformément continues, par rapport à cette définition de la distance, dans une fraction aussi grande qu'on veut de tout domaine fini.

Vous voyez sans peine qu'avec cette double propriété, la notion d'intégrale se légitime sans peine.

J'arrive au but avec la définition suivante :

Soit  $u(s)$  et  $v(t)$  deux fonctions, définies dans le même intervalle 0 à 1. J'établis entre  $s$  et  $t$  une correspondance sans souci de la continuité, mais uniforme et con-

---

<sup>15</sup>On January 20, 1919, Lévy presented a note and sent the corresponding article Lévy, Paul (1920) Sur les fonctions de lignes implicites. Bull. Soc. Math. Fr. 48:13–27. In them he introduced the notion of distance in  $L^2$ .

<sup>16</sup>For a better understanding of the rest of this letter, see the next letter, where Lévy's arguments are more precise.

servant la mesure de tout sous-ensemble ; et je considère la quantité

$$\sqrt{\int [(u - v)^2 + (s - t)^2] ds} \quad (ds \text{ ou } dt \text{ indifféremment})$$

son minimum lorsqu'on fait varier la correspondance considérée est la distance.\*\*

Cette définition, (ou une autre équivalente au point de vue où je me place), a-t-elle déjà été considérée ?

J'ai d'ailleurs trouvé dans Gateaux des préoccupations voisines. Il étudie à plusieurs reprises les fonctionnelles telles que si l'on divise  $(0, 1)$  en intervalles  $< h$  assez petit, et si  $u$  et  $v$  ont même valeur moyenne dans chaque intervalle, on ait  $|\Phi_u - \Phi_v| < \varepsilon$ . Avec une telle continuité la définition d'intégrale n'est pas difficile à légitimer, mais elle est beaucoup trop restrictive bien qu'elle comprenne les fonctionnelles de la forme

$$\Phi_p = \int_0^1 \cdots \int_0^1 f(x_1, \dots, x_p) u(x_1) \cdots u(x_p) dx_1 \cdots dx_p$$

que Volterra et vous-même avez souvent considérées. Je considère ces fonctionnelles comme la limite pour  $n$  infini de fonctions de  $n$  variables linéaires par rapport à chacune des variables.

La différentielle seconde, par exemple, d'une fonctionnelle  $\Phi$ , est en principe pour moi de la forme

$$\int_0^1 f_2(x) \delta u^2 dx + \int_0^1 \int_0^1 F_2(x_1, x_2) \delta u_1 \delta u_2 dx_1 dx_2$$

L'équation de Laplace se généralise par

$$\int_0^1 f_2(x) dx = 0 \tag{1}$$

et non par

$$\int_0^1 F_2(x, x) dx = 0, \tag{2}$$

équation considérée par Volterra,<sup>17</sup> qui admet implicitement  $f_2(x) = 0$ .

---

<sup>17</sup>After the publication of his book *Equations intégrales et intégral-différentielles (Integral and Integral-differential Equations)* (Gauthier-Villars (1913)), Volterra published a long article Volterra, Vito (1914) *Equazioni integro-differenziali ed equazioni alle derivate funzionali*. Rend. Accad. Lincei 23:551–577. In it he attempted to solve certain types of functional equations by using representation theorems such as those found by Gateaux and published by the latter at the Accademia dei Lincei (December 21, 1913). Lévy quickly realized, and never stopped repeating, that Volterra had made an error by neglecting the not absolutely continuous term in the second derivative. See Sect. 4, as well as Lévy's contribution *Fonctions de lignes et équations aux dérivées fonctionnelles (Functions of Line and Equations with Functional Derivatives)* to the XIIth international congress of the history of science, held in Moscow in 1971.

Les problèmes relatifs à cette dernière équation sont en principe des problèmes mal posés, en ce sens que toute fonctionnelle peut être approchée autant qu'on veut par une solution de cette équation- comme elle peut l'être par une solution de l'équation<sup>18</sup>

$$\Phi'[[u(x), \xi]] = c, \quad \xi \text{ valeur particulière.}$$

Ces problèmes ressemblent au problème classique : trouver une fonction continue de  $x$  égale à 1 pour  $x = 0$  et  $x = 1$ , et rendant minima  $\int_0^1 y^2 dx$ . On peut rendre cette intégrale aussi petite qu'on veut, mais non nulle. Je citerai comme analogue l'équation

$$y(a) + \int_0^1 \varphi[x, y(x)] dx = 0$$

qui peut être vérifiée par une fonction approchant autant qu'on veut (au sens usuel du mot distance) de toute fonction donnée.

On n'évite ces difficultés qu'en considérant des équations pour lesquelles  $\Phi'$  ou  $f_2(x)$  n'interviennent qu'intégrées dans un intervalle de mesure linéaire finie, ou  $F_2(x, x)$  intégrée dans une aire de mesure finie. L'équation (2) est donc d'un type un peu bâtard, et Volterra qui a considéré (Lincei, 1er sem.1914) des équations de cette nature ne semble pas s'en être douté.<sup>19</sup> Au contraire, il ne se doute pas que l'hypothèse  $f_2(x) = 0$  qu'il fait implicitement est aussi restrictive que ne l'est l'hypothèse  $\frac{\partial^2 \varphi}{\partial x_1^2} = \frac{\partial^2 \varphi}{\partial x_2^2} = \dots = \frac{\partial^2 \varphi}{\partial x_n^2} = 0$  pour une fonction de  $n$  variables.<sup>20</sup>

Au contraire l'équation (1) se prête à la généralisation du problème de Dirichlet, problème résolu par Gateaux pour la sphère et à peu près pour une surface quelconque, et dont la solution est plus simple que dans le cas de l'espace ordinaire.

Bien entendu, il est intéressant d'étudier les fonctions de la forme  $\Phi_p$  considérée plus haut, comme aussi de chercher par contre la forme la plus générale des fonctionnelles d'ordre  $p$  ce que vous avez fait avec succès, mais pour ma part je cherche à mettre au premier plan un type de fonctionnelles d'une généralité intermédiaire, qui correspondra à ce qu'est dans la théorie des fonctions ordinaires la fonction ayant des dérivées finies et continues jusqu'à l'ordre  $p$ .

J'insiste à ce sujet parce que votre avis m'intéresse particulièrement. Ayant étudié et souvent considéré les fonctionnelles  $\Phi_p$ , je serais heureux de savoir si vous estimez aussi qu'elles doivent être considérées comme limite de fonctions de  $n$  variables vérifiant  $n$  équations aux dérivées partielles du second ordre - et par suite comme assez particulières.

Bien cordialement à vous.

P.Lévy

<sup>18</sup>Lévy, following Volterra's notations, used the brackets  $[[ \ ]]$  to enclose the argument of a functional.

<sup>19</sup>In Lévy (1922) Lévy was more direct, claiming that Volterra's equations were of no interest.

<sup>20</sup>On this point, see the next letter.

Je ne réponds pas directement à certains points de votre lettre car je n'ai qu'à tirer mon profit des intéressants renseignements que vous me donnez, et aucune nouvelle observation à présenter.

Je vous signale comme publication pouvant vous intéresser : G.C. Evans - The Cambridge colloquium 1916 - Part I - Functionals and their applications ; selected topics ; including integral equations. - (publié par Amer.Math.Soc.).

Monsieur Hadamard me l'a prêté. Peut-être vous le prêtera-t-il aussi, bien que votre éloignement rende la chose plus difficile.<sup>21</sup>

- \* j'ai dû renoncer aux exclusions basées sur la condition de Lipschitz ou l'égalité de continuité au sens habituel, qui ne vérifient pas cette condition de porter une fraction très petite du domaine.
- \*\* avec cette définition,  $\Phi$  est uniformément continue si  $\Phi'[[u(x), \xi]]$  existe et est continue en  $\xi$ , et de plus également continue (lorsque  $u(x)$  varie).

## Lettre 5

Paris 16 février 1919

9 rue Chernoviz,

Mon cher collègue,

Il est vrai que ma dernière lettre manquait de précision. Je vais essayer d'en apporter un peu dans celle ci, bien que je craigne qu'il ne me faille pour cela écrire des volumes.<sup>22</sup>

D'abord quand je parle de fonctions ou de fonctionnelles usuelles, j'entends avant tout celles qu'on rencontre en physique mathématique. Fonctions : solutions d'équations aux dérivées partielles. Fonctionnelles : celles du calcul des variations ; intégrales de Dirichlet et intégrales analogues qu'on rencontre en généralisant les théories de Jacobi-Hamilton, aire d'une surface minimum limitée à un contour, fonctions de Green, capacité électrique d'un conducteur isolé, densité électrique en un point de la surface... etc.

Je ne définis pas les conditions de continuité a priori, mais je prends celles qui se rencontrent dans les applications que j'ai en vue. J'arrive ainsi à certaines conceptions qui changeront peut-être si je veux étudier la physique moléculaire et les fonctions monogènes considérées par Borel dont les pôles sont partout denses.<sup>23</sup> Mais voici où j'en suis pour le moment :

Fonctions : Domaine  $D_n$  des fonctions admettant des dérivées finies et continues jusqu'à l'ordre  $n$  ; domaine  $D'_n$  des fonctions analogues aux précédentes mais ayant des singularités isolées. Ces fonctions sont celles qui s'imposent le plus, toutefois l'importance des développements en série de fonctions orthogonales me conduit à généraliser un peu et considérer le domaine  $D_{\mathcal{L}}$  des fonctions intégrables et de carré intégrable au sens de Lebesgue (Je crois que c'est bien intégrable ; je n'ai pas le

---

<sup>21</sup>It is clear how proud Lévy was that Hadamard had lent him this work.

<sup>22</sup>Indeed, the full explanation of the topics in this letter constitutes most of Chap. VI of Lévy (1922).

<sup>23</sup>See Sect. 1.

Lebesgue sous la main et j'ai peur de faire une confusion de mots entre intégrable et sommable. Je veux dire fonction telle que l'ensemble des points où la fonction  $x(s)$  est comprise entre  $x_i$  et  $x_{i+1}$  a toujours une mesure bien déterminée  $\varepsilon_i$  et où  $\sum \varepsilon_i x_i$  a une limite déterminée ainsi que  $\sum \varepsilon_i x_i^2$  si les  $x_{i+1} - x_i$  tendent uniformément vers 0 dans tout intervalle fini  $(\xi_0, \xi_1)$ .

Cet ensemble  $D_{\mathcal{L}}$  présente l'avantage que les fonctionnelles définies dans  $D_0$  et qui ne seraient pas dans  $D_{\mathcal{L}}$  ne se sont guère rencontrées dans les applications que j'ai en vue. J'ai donc intérêt à étendre le domaine de fonctions de  $D_0$  à  $D_{\mathcal{L}}$ , restreindre ainsi le domaine de fonctionnelles à des fonctionnelles susceptibles de représentations plus commodes.

Donc en principe je considère le domaine  $D_{\mathcal{L}}$ . Les définitions de la distance, continuité, seront faites pour ce domaine, et s'il m'arrive de parler de fonctionnelles qui ne sont définies que dans  $D_0$ , ou  $D_1$ , ou  $D_2$ , je le préciserai.

Avec cela je définis la convergence par la convergence en moyenne, et la distance par

$$r = \sqrt{\int_0^1 [u(s) - v(s)]^2 ds}.$$

Pourquoi ?<sup>24</sup> D'abord parce que les fonctionnelles qui m'intéressent pour les applications sont continues avec cette définition, sauf lorsqu'il y a un point spécialement en évidence dans leur définition, cas qui de toute façon mérite d'être traité à part. Elles sont évidemment continues a fortiori si on définit la limite par la convergence uniforme, mais cette définition n'est pas assez restrictive ; il serait hardi de considérer comme continue une fonctionnelle qui pour la fonction  $x(s) = 0$  si  $s < \sigma$ ,  $x(s) = 1$  si  $s \geq \sigma$  ne se réduirait pas à une fonction continue de  $\sigma$ .

D'autre part il ne faut pas perdre de vue que, si on définit les fonctions par diverses représentations analytiques, il ne faut pas oublier qu'une fonction est une succession de valeurs de  $x$  dont la représentation ne peut se faire de manière plus conforme à sa véritable nature qu'en la représentant d'une manière approchée par une suite de traits horizontaux. C'est ce qu'on fait dans la théorie élémentaire de l'intégrale définie. La fonctionnelle apparaît comme limite d'une fonction  $f(x_1, \dots, x_n)$ . Si on définit la distance dans l'espace à  $n$  dimensions, le passage à la limite n'est possible qu'en divisant par  $\sqrt{n}$ , et on obtient la définition ci-dessus.

Avec cette définition, toute la géométrie et l'analyse ordinaire se généralise. Sphère, angle, inversion, orthogonalité, potentiel, fonctions harmoniques, théorèmes des surfaces minima, courbure des lignes et des surfaces, et même avec d'étranges simplifications dans ces théories.

D'ailleurs<sup>25</sup> si je définis  $x(s)$  par une série de fonctions orthogonales

$$x(s) = \sum c_n \varphi_n(s)$$

<sup>24</sup>The entirety of Chap. II of Lévy (1922) is devoted to a discussion of the different kinds of continuity of interest in function spaces. Lévy was an unceasing advocate of the norm  $L^2$ .

<sup>25</sup>Geometrical properties of the space  $L^2$  constitute Chap. VII of Lévy (1922).



la fonctionnelle devient une fonction des  $c_n$  et la géométrie dont je parle devient identique à celle de votre espace  $\Omega$ .<sup>26</sup> Le passage d'un point de vue à l'autre est un simple changement de coordonnées rectangulaires. La réduction d'une quadrique à ses axes conduit suivant les cas à l'un ou l'autre système de coordonnées. La généralisation de la notion d'intégrale aboutit au même résultat en partant de l'un ou l'autre point de vue.

Cela dit je passe aux fonctionnelles. Pour celles que j'ai rencontrées dans les applications j'ai toujours rencontré les circonstances suivantes :

Si aucun point de l'intervalle de variation de  $s$  n'est en évidence dans la définition de  $\Phi$ , j'ai

$$\delta\Phi = \int \Phi'_x|[x(s); t]|\delta x(t) dt. \quad (1)$$

Si un point particulier  $\sigma$  est en évidence, cette formule reste exacte sauf dans le voisinage de ce point. Mais pour le voisinage de ce point il faut modifier la formule.

Deux cas suivant que  $\Phi'_x$  devient infini ou non pour  $t$  voisin de  $\sigma$ . Le cas où cette fonction devient infinie est fréquent dans les applications (fonctions de Green ; densité électrique en un point d'un conducteur). Néanmoins il me paraît devoir être traité après l'autre, de même que les singularités des fonctions analytiques ne doivent être étudiées que quand on a déterminé les aspects de ces fonctions aux points réguliers. Je suppose donc  $\Phi'_x$  continue pour  $t = \sigma$ . Les termes dépendant de  $\sigma$  sont alors de la forme

$$A_0\delta x(\sigma) + A_1\delta x'(\sigma) + \dots$$

ou si je veux que  $\Phi$  soit définie dans  $D_{\mathcal{L}}$  ou au moins dans  $D_0$ , de la forme  $A_0\delta x(s)$ . Donc la formule (1) devient

$$\delta\Phi_{\sigma} = \int_0^1 \Phi'_x|[x(s), t]|\delta x(t) dt + A_0\delta x(\sigma). \quad (2)$$

Cette formule s'applique en particulier à  $\Phi'_x|[x(s), \sigma]$ . Si j'en porte la variation dans  $\delta\Phi$ , j'ai

$$\delta_2\Phi = \int_0^1 \Phi''_{x^2}[\delta x(t)]^2 dt + \int_0^1 \int_0^1 \Phi''_{x_1x_2}(t_1, t_2)\delta x(t_1)\delta x(t_2) dt_1 dt_2 \quad (3)$$

qui généralise

$$\delta_2\phi(x_1, \dots, x_n) = \sum \frac{\partial^2\phi}{\partial x_i^2} dx_i^2 + \sum_{i=1}^n \sum_{j=1, i \neq j}^n \frac{\partial^2\phi}{\partial x_i \partial x_j} dx_i dx_j. \quad (4)$$

---

<sup>26</sup>The space  $\Omega$  is the Hilbert space of square summable real sequences, denoted  $\ell_2$  today. Fréchet studied its geometric properties in Fréchet, Maurice (1908) *Essai de géométrie analytique à une infinité de coordonnées*. *Nouv. Ann. Math.* 8:97–116 and 289–317. Lévy shows the connection between  $L^2$  and  $\ell_2$ , which holds generally in a Hilbert space (see Dixmier 1985).

Supprimer la première intégrale équivaudrait à supprimer les termes carrés dans (4) et ne considérer que des formes linéaires par rapport à chacune des variables. Il y a lieu de remarquer que si je considère le potentiel

$$\Phi = \int_S \frac{\varphi(M)}{r^{n-2}} dS, \quad r^2 = \sum_{i=1}^n (x_i - \xi_i)^2, \quad \xi_i \text{ coordonnées de } M$$

dans l'expression  $\delta_2 \Phi$ , les termes carrés, en nombre  $n$  seulement, représentent pourtant un même ordre de grandeur que les autres et à la limite on arrive bien à une expression de la forme (3).

Ce serait donc une erreur de négliger la première intégrale ; on arriverait ainsi à une notion de la quadrique qui ne comprendrait pas la sphère

$$\int_0^1 x^2(t) dt = 1.$$

-Je reviens au sujet. Je considère donc maintenant les fonctionnelles  $\Phi$  auxquelles les formules (2) et (3) s'appliquent,  $\Phi, \Phi', \Phi'', \Phi'_{x^2}, \Phi''_{x_1 x_2}$ , et de plus  $\frac{d\Phi'_x}{dx}$  étant bornés dans tout domaine fini extrait de  $D_{\mathcal{L}}$ .

Ceci dit 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  $(0, 1/n), (1/n, 2/n), \dots, (n-1/n, n/n)$  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  $|\Phi - \Phi_n| < \varepsilon$  pour  $n$  assez grand, il faut que la contribution due à la variation de  $x$  dans chacun des intervalles d'amplitude  $\frac{1}{n}$  à cette intégrale puisse être limitée par  $\frac{\varepsilon}{n}$  ( $\varepsilon$  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<sup>27</sup>

$$\frac{1}{2n} \int_{x_n}^{x_n+a} \Phi'_x |[\xi(s), t]| d\lambda - \frac{1}{2n} \int_{x_n-a}^{x_n} \text{expression identique} \quad (5)$$

$t$  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  $\Phi'_x$  est indépendant de  $\xi$ , c'est à dire si le terme  $A_0 \delta x$  de la formule (2)

---

<sup>27</sup>Lévy's ideas are somewhat confused here, so it is unclear what Fréchet made of his explanations. Lévy was perhaps thinking of integration by parts. The proof that Gateaux's functionals are harmonic (in the functional sense) was finalized in Lévy (1922, first part, nos. 85–89).

appliqué à  $\Phi'_x$  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.

Pour mes fonctionnelles (p7 lignes 1 à 3),<sup>28</sup> je constate que si la grandeur des différentes valeurs de  $x$  qui contribuent dans l'intervalle considéré à donner la valeur moyenne  $x_n$  est très importante, par contre il est indifférent de savoir en quels points de l'intervalle ces valeurs sont atteintes. Si je divise l'échelle verticale des  $x$  en  $n$  intervalles tels que les valeurs de l'intervalle horizontal considéré pour lesquelles les  $x$  soient compris dans chacun des intervalles verticaux constituent des ensembles de mesure  $\frac{1}{n^2}$ , la détermination exacte de chacun de ces ensembles est sans importance. En vue de bien évaluer l'intégrale

$$\int_{\frac{i}{n}}^{\frac{i+1}{n}} dt \int_0^1 \Phi'_x [x_n(s) + \lambda y_n(s); t] d\lambda, \quad (6)$$

[ $y_n = 0$  en dehors de l'intervalle,  $= x - x_n$  dans l'intervalle], seules sont importantes les valeurs moyennes de  $x$  dans chacun de ces intervalles. Si j'exclus les trop grandes valeurs de  $x$  en disant que dans le domaine fini

$$\int_0^1 x^2 ds < R^2$$

j'exclus les fonctions  $x$  pour lesquelles les valeurs de  $x$  supérieures à  $N$  contribuent pour plus de  $\frac{R^2}{N}$  à la valeur de cette intégrale, j'arrive à cette conclusion que les valeurs moyennes de  $x$  dans chacun de mes petits intervalles me représentent assez bien cette fonction pour que la valeur de  $\Phi$  soit atteinte avec une erreur  $< \frac{\varepsilon}{n}$ , donc si je fais la même chose dans tous les intervalles d'amplitude  $1/n$ , la valeur de  $\Phi$  est atteinte à  $\varepsilon$  près.

Donc j'arrive à représenter  $x$  par  $X_n$  de la manière suivante :

je divise l'intervalle de variation de  $s$  en  $n$  intervalles  $e_i$  ( $i = 1, \dots, n$ ) égaux à  $1/n$ . Dans chacun de ces intervalles  $e_i$ , je définis  $n$  ensembles  $e_{i,j}$  de mesure  $\frac{1}{n^2}$  comprenant l'un les plus petites valeurs de  $x$ , puis celles qui viennent ensuite, et ainsi de suite. Soit  $\xi_{i,j}$  la valeur moyenne de  $x$  dans  $e_{i,j}$ .

Je divise maintenant  $e_i$  en  $n$  intervalles consécutifs égaux  $e'_{ij}$  et je considère la fonction  $X_n = \xi_{i,j}$  dans  $e'_{ij}$ . Je la considère comme la  $n$  ième approximation de  $x$ . Et en excluant du domaine considéré les fonctions exclues p.9<sup>29</sup>  $N$  est très grand,  $\Phi|[X_n]|$  tend uniformément vers  $\Phi|[x]|$ .

<sup>28</sup>Lévy refers here and below to the pagination of his letters. The lines are those from the earlier paragraph that begins with "Je reviens au sujet."

<sup>29</sup>I.e., that satisfy  $\int_0^1 x^2 ds < R^2$ .

Je veux alors mettre en évidence une définition de la distance avec laquelle je puisse dire que  $X_n$  tend uniformément vers  $x$  lorsque  $\Phi|[X_n]$  tend uniformément vers  $\Phi|[x]$ .

Ce serait alors la suivante :

Soient  $x(s)$  et  $y(t)$  définies entre 0 et 1.

Je divise l'intervalle 0, 1 en  $p$  intervalles égaux et j'établis une correspondance entre  $s$  et  $t$  en permutant ces intervalles. J'ai ainsi  $p'$  valeurs de

$$\int_0^1 \sqrt{(x-y)^2 + (s-t)^2} ds \quad (ds \text{ ou } dt)$$

dont la plus petite sera désignée par  $\mu_p$ . Les quantités  $\mu_1, \mu_2, \dots, \mu_{2i}$  forment une suite décroissante, ou du moins non croissante, dont la limite  $\mu$  sera la distance entre  $x$  et  $y$ .

- J'ai écrit vite, et certains détails ne sont pas encore au point. En me relisant je vois que la forme sous laquelle j'ai exclu les valeurs trop grandes de  $x$  ne permet pas les conclusions que j'ai énoncées. Il faut que je choisisse  $\varepsilon$  d'abord,  $N$  ensuite, de sorte que ma phrase de la page 10, lignes 2 à 4<sup>30</sup> n'est pas tout à fait correcte. Mais je suis à peu près sûr de l'ensemble, et j'espère que vous me comprendrez, quand vous en aurez le temps.

Bien cordialement à vous

P.Lévy.

## Lettre 6

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

SECRETARIAT A LA SORBONNE

Place de la Sorbonne

A PARIS

Paris, le 19 février 1921

Cher Monsieur,

Vous connaissez sans doute l'existence du séminaire de Monsieur Hadamard, réunions tenues au Collège de France sous sa direction, et où chacun analyse des mémoires, suivant sa spécialité.

M.Hadamard m'avait demandé de me charger du mémoire de Chittenden (Circ. Palermo 1er sem 1915) sur la general Analysis.<sup>31</sup> Après avoir parcouru ce mémoire, et le livre de Moore<sup>32</sup> dont la lecture préalable, au moins partielle, est indispensable,

<sup>30</sup>The assertion above that  $\Phi|[X_n]$  goes uniformly to  $\Phi|[x]$ .

<sup>31</sup>Lévy is referring to the article with a somewhat obscure title Chittenden, Edward W. (1915) Infinite developments and the composition property  $(K_{12}B_1)_*$  in general analysis. Rend. Circ. Mat. Palermo 39:81–108. E.W. Chittenden was engaged as well as his colleague E.H. Moore in research on General analysis and topology. He explored the foundations of Fréchet's functional calculus in his own *On the Foundations of the Calcul fonctionnel of Fréchet*. Trans. Am. Math. Soc. 19:66–67 (1918). This work was cited numerous times in Fréchet (1928).

<sup>32</sup>Lévy is probably alluding to the notes of E.H. Moore's lectures of 1906 entitled *Introduction to a Form of General Analysis*, given at Yale's New Haven Mathematical Colloquium, and published in 1910. E.H. Moore appears again in Letter 96.

je me suis persuadé que je ne pourrai pas cette année mener cette lecture à bonne fin. Je lis péniblement l'anglais, et je ne suis pas du tout habitué aux notations dont se servent ces auteurs.

Comme il s'agit du domaine de la science qui est le vôtre, je suis persuadé que, si vous ne connaissez pas déjà le mémoire de Chittenden, vous le lirez sans difficulté. Je pense aussi que vous venez quelquefois à Paris. Ne vous serait-il pas possible de venir faire cette analyse au Collège de France, un vendredi ou un samedi, à 5 heures du soir.

Je serais heureux de cette occasion de faire votre connaissance,<sup>33</sup> et je suis persuadé que l'auditoire gagnera au change, car je ne sais pas ce que j'arriverais à dire sur ce sujet que je ne connais pas tandis que vous le connaissez à fond.

Bien cordialement à vous.

Paul Lévy

9 rue Chernovicz 16e

[*In footnote (Fréchet's handwriting): Répondre peu probable*]

## Lettre 7

Paris- 9 rue Chernovicz

le 2 décembre 1924

Mon cher Collègue,

J'ai bien reçu vos tirages à part, récents et anciens, et vous en remercie bien vivement. Vous avez eu raison d'attirer mon attention sur quelques points de vos travaux, qui m'avaient échappé. J'aurais dû vous citer dans mes Leçons d'analyse fonctionnelle plus souvent que je ne l'ai fait.

Le résumé que je prépare pour le Mémorial<sup>34</sup> me donnera l'occasion de réparer en partie ces erreurs ; je dis seulement en partie, car en soixante pages je dois me borner, et ne puis dire tout ce que je voudrais dire.

Toutefois je dois insister sur l'intégration dans l'espace  $\Omega$ . Votre note sur les singularité des espaces à un très grand nombre de dimensions touche à ce sujet. Elle est datée du 29 juillet ; je suppose que c'est du 29 juillet 1914 ;<sup>35</sup> elle serait donc antérieure, sinon aux travaux de Gateaux, du moins à leur publication, et dans ce cas il serait utile de la citer. Mais sous quelle référence ?

J'ai lu avec intérêt votre dernière note du Bulletin des Sciences.<sup>36</sup>

---

<sup>33</sup>It is a little surprising that two years after they first exchanged letters, the mathematicians still had not met.

<sup>34</sup>At the end of 1924, at the request of Henri Villat, Lévy wrote a 56-page booklet on functional analysis for the collection *Mémorial des Sciences Mathématiques*. This was published in 1925, numbered V. Principally, it summarizes the book of Lévy (1922). However, as mentioned in the letters below, it includes some new concepts, especially remarks on Wiener's differential space on p. 51. Fréchet's work is cited frequently.

<sup>35</sup>This is a contribution to the meeting of the AFAS at Le Havre in 1914 (pp. 146–147).

<sup>36</sup>Probably Fréchet, Maurice (1924) Prolongement des fonctionnelles continues sur un ensemble abstrait. Bull. Sci. Math. 47:171–192, in which Fréchet took up several of Lévy's ideas from Lévy (1922).

Croyez, mon cher collègue à mes sentiments les plus cordialement dévoués.  
P.Lévy

### Lettre 8

Paris - 9 rue Chernoviz

le 7 décembre 1924

Mon cher collègue,

Un mot pour vous remercier de votre lettre et vous dire de ne pas vous donner la peine de m'envoyer vos mémoires sur l'intégration d'une fonctionnelle (Soc.Math.1915) et sur les fonctions additives d'ensembles abstraits (fund.Math tIV), que vous avez déjà eu l'obligeance de m'envoyer.

J'étais arrivé de mon côté à ces résultats, après vous, mais avant d'avoir lu vos mémoires. C'est une chose qui m'arrive souvent, car mon travail est d'un meilleur rendement quand je réfléchis par moi-même aux questions qui m'intéressent que quand je lis la plume à la main, et le plus souvent dans un mémoire nouveau je ne fais que glaner quelques idées pour y réfléchir à loisir. Cela fait que bien souvent je ne sais pas moi-même ce qui est mien et ce qui ne l'est pas. Ceci dit pour vous expliquer les omissions de ma bibliographie. Mais il faut qu'elles soient comblées pour le mémorial.

A propos de vos idées sur l'intégration dans un ensemble abstrait, et de celles de Daniell, j'estime qu'elles s'adaptent surtout aux ensembles ayant la puissance du continu. C'est une idée un peu vague, que j'ai depuis longtemps, et que je cherche à préciser pour mon article du Mémorial.<sup>37</sup> Vous le verrez si j'y arrive. C'est d'ailleurs le seul point où je toucherai explicitement aux ensembles abstraits.

Dans le même ordre d'idée, je compte parler de l'espace différentiel de N.Wiener.<sup>38</sup> Je pense que son travail ne vous a pas échappé. Les calculs sont beaucoup trop longs mais l'idée fondamentale est très intéressante.

A propos de calcul des Probabilités, je n'ai rien publié en dehors des Comptes Rendus, de la Société Math., et de la revue de Métaphysique et Morale (je crois du moins que c'est dans cette revue que doivent paraître les conférences faites au congrès de philosophie).<sup>39</sup> Je n'ai d'exemplaires disponibles que de mon article de la Soc. Math. et vais vous en envoyer un.

J'ai aussi consacré quelques Leçons à ce sujet dans mon cours de l'X, mais j'ai très peu d'exemplaires et si cela vous intéresse, je vous serais obligé d'emprunter à M.Villat celui que je lui ai envoyé.

Bien cordialement à vous.

P.Lévy.

---

<sup>37</sup>See note 34.

<sup>38</sup>First appearance of Wiener's *Differential Space* and thus of the modern theory of Brownian motion. Wiener read Lévy (1922) attentively, and found in it a number of techniques that he used in *Differential Space* (see Sect. 7, the note 34) and Mazliak (2011).

<sup>39</sup>Lévy is probably referring to his article included in the last part of his book (Lévy 1925).

## Lettre 9

Paris- 9 rue Chernoviz

le 16 octobre 1925

Mon cher collègue,

Merci beaucoup pour l'envoi de vos tirages à part.<sup>40</sup>

Excusez moi d'ajouter, à propos de la loi des erreurs, qu'il me semble qu'il y a un malentendu sur le sens du principe d'homogénéité (axiome II de M.Soubbotine).

Pour n'importe quelle loi  $L$ , si des mesures  $x_1, x_2, \dots, x_n$  conduisent à prendre  $x$  comme valeur, on peut en déduire que, pour la loi  $L_a$  déduite de la précédente par le changement de  $x$  en  $ax$ , des mesures  $ax_1, \dots, ax_n$  conduisent à prendre  $ax$  comme valeur la plus probable. L'axiome de M.Soubbotine implique que cela soit vrai, non pour la loi  $L_a$  mais pour la loi  $L$ ; (sans cela cela ne vaudrait pas la peine de l'énoncer).

Or dans l'exemple que vous donnez, vous changez les constantes  $h$  et  $k$  qui définissent la loi  $L$  en  $H = ah$  et  $K = ak$ , qui définissent une loi  $L_a$  autre que  $L$ . On ne peut donc pas dire que votre exemple vérifie le principe de M Soubbotine.

Je ne connais pas le travail de cet auteur, mais à première vue il me paraît avoir raison en affirmant qu'un principe d'homogénéité de la nature considérée conduit à des lois dépendant de deux paramètres, un paramètre d'homogénéité et un autre. Il serait plus intéressant d'ailleurs d'énoncer un tel principe en mettant en évidence la valeur probable, et non la plus probable.

J'ai moi-même, en analysant un mémoire de Baer et Bernstein (v par ex mon mémoire du Bulletin du cinquantenaire de la Soc. Math §12 avant-dernier alinéa de ce §),<sup>41</sup> fait observer qu'un principe de cette nature (présenté d'ailleurs par les auteurs sous une forme que je critique), conduit nécessairement à ce que j'appelle une loi stable; mais il reste à justifier l'exposant 2 qui donne la loi de Gauss, et à moins d'écarter pour des raisons un peu métaphysiques l'hypothèse d'une loi admettant une erreur quadratique infinie (ce que fait Bienaymé), on n'y arrive que par la méthode utilisée par M. Lindeberg et moi-même.

Je n'ai encore fait que jeter un coup d'œil sur votre autre mémoire, mais la question m'intéresse et je vous promets de le lire attentivement.

Croyez, mon cher collègue, à mes sentiments les plus cordialement dévoués.

P.Lévy.

---

<sup>40</sup>This letter is the beginning of a veritable avalanche of correspondence on the law of errors. Fréchet sent Lévy the article that he had published in 1924 (Математический Сборник 32:5–8). This was a response to the two axioms for admissible error functions proposed by M.F. Subbotin. Lévy and Fréchet each had different interpretations of Subbotin's second axiom. See Sect. 4.

<sup>41</sup>In his article Lévy, Paul (1924) Théorie des erreurs. La loi de Gauss et les lois exceptionnelles. Bull. Soc. Math. Fr. 52:49–85, Lévy strongly criticized the 1915 attempt in Baer, W.S. and Bernstein, Felix (1915) Ein Axiomensystem des Methode des kleinsten Quadrate. Math. Ann. 76:284–294 to base the law of errors on axioms supposedly independent of the idea of probability. In particular, Lévy showed that the introduction of a "parameter of precision," on which the two German mathematicians based their theory, was not possible in the absence of a probabilistic structure. In his article Lévy also denounced the unconditional veneration of the axiomatic method by "Hilbert's students."

**Lettre 10**

Paris- 9 rue Chernoviz  
 le 21 octobre 1925,  
 Mon cher collègue,

- 1° Au sujet du principe d'homogénéité de M.Soubbotine,<sup>42</sup> je vous ai écrit que l'exactitude du résultat d'après lequel il définissait des lois à deux paramètres seulement me paraissait probable. Je me suis mal exprimé ; cela pourrait faire croire que j'ai des raisons plausibles de le croire exact. Je voulais seulement dire que cela ne me paraît pas improbable, à première vue.
- 2° A première vue également, je suis peu porté à croire à l'existence d'une loi ayant les propriétés que vous indiquez au P.S. Remarquez que,  $\varphi(x_1, x_2, \dots, x_n)$  désignant la valeur probable de  $x$  après  $n$  mesures ayant donné  $x_1, x_2, \dots, x_n$ , on ne dispose que d'une fonction arbitraire de  $n$  variables pour rendre  $\varphi(x_1, \dots, x_n)$  égal à une fonction donnée de ces  $n$  variables, et cela quelque grand que soit  $n$ . Le problème semble donc n'avoir de solution que dans des cas très particuliers.
- 3° Je n'ai aucune objection au sujet des idées de Daniell. Je lis l'anglais avec effort, et connaît Daniell surtout à travers Wiener ;<sup>43</sup> si rien d'important ne m'a échappé, Daniell a donné, non une définition de l'intégrale, mais une extension de la notion d'intégrale d'un domaine restreint à un domaine plus vaste. C'est du Lebesgue. Il m'a semblé que dans un exposé très succinct, il convenait d'indiquer avec précision les théories constructives,<sup>44</sup> et n'indiquer que brièvement les travaux ayant pour objet de définir avec précision le champ d'application de ces théories, ce que j'ai fait p.43.<sup>45</sup>

Je suis le premier à regretter d'avoir dû être si bref. Vous avez pu voir dans mes autres livres la place que j'accorde à Lebesgue lorsque je ne suis pas borné.

Bien cordialement à vous  
 P.Lévy

**Lettre 11**

Paris- 9 rue Chernoviz  
 le 23 octobre 1925,  
 Mon cher collègue,

---

<sup>42</sup>See note 40.

<sup>43</sup>Lévy knew Wiener's work thoroughly. Wiener had proposed a construction of the Brownian measure based on an abstract extension of Daniell's integral. Lévy was exaggerating a bit in suggesting that this was only a small extension of Lebesgue's methods, since Lebesgue never considered generalizations to infinite dimension spaces See Sect. 7 and Mazliak (2011).

<sup>44</sup>This is a good testimony of Lévy's constructivist bent, particularly evident in his work on Brownian motion.

<sup>45</sup>The allusion is to the booklet in the collection *Mémorial sur l'Analyse Fonctionnelle* (see footnote 34). On the page cited, Lévy makes passing and laconic mention of the generalization of the integral to infinite dimension studied by Fréchet, Daniell and Wiener "qui, d'ailleurs, se sont inspirés dans ce but des idées de M.H. Lebesgue (whose work, incidentally, was inspired by the ideas of H. Lebesgue)."



Je ne sais pas si je dois vous dire que nous sommes d'accord, ou que nous ne le sommes pas, car au fond ce n'est qu'une question de convention qui nous sépare. On a le droit d'énoncer tel axiome que l'on veut, pour en étudier les conséquences.

Ce que je dis, je crois vous l'avoir dit nettement dès ma première lettre, c'est que l'interprétation de l'axiome qui justifierait votre raisonnement en ferait une propriété évidente appartenant à toutes les lois et dont on ne peut rien tirer. L'autre donne au contraire une propriété très restrictive.

Je n'ai pas lu l'article de M.S.<sup>46</sup> Le plus probable me paraît être qu'il a énoncé son axiome dans l'idée que je lui attribue, et qu'en lisant votre note il ne s'est pas rendu compte du malentendu. Il y a là une question de fait facile à éclaircir.

Je pose le problème : "Quelles sont les courbes unicursales dont l'équation ne change pas par changement d'unité ?" Je réponds : "Les droites passant par l'origine." Je trouverais un peu spécieux que l'on vienne me dire : toute courbe possède cette propriété, car elle appartient à un faisceau de courbes homothétiques qui la possède. Ce serait mal comprendre mon énoncé. Je vous avoue franchement que c'est un peu l'effet que me produit votre raisonnement.

[*Paragraphe barré par Lévy* : Il est bien entendu qu'en général, quand on parle de loi de probabilités, on considère la grandeur mesurée  $X$  comme fixe et la mesure  $x$  comme variable ; l'autre conception est possible et l'on ne peut passer de l'une à l'autre que par application du principe de la probabilité des causes et l'hypothèse que les différentes valeurs de  $X$  ont la même probabilité a priori. Mais alors l'axiome I, probabilité fonction de  $|X - x|$  seulement, entraîne cette conséquence qu'on a la même loi à l'un et l'autre point de vue. Je ne vois pas bien le rapport entre cette position et celle qui nous occupe.] (Cet alinéa provenait de ce que j'avais mal compris un passage de votre lettre)

- Je ne comprends pas pourquoi vous dites que mon interprétation, celle qui n'est pas une tautologie, ne correspond à rien d'intuitif.<sup>47</sup> Si des mesures  $x_1, x_2, \dots, x_n$  conduisent à prendre  $x = f(x_1, \dots, x_n)$  comme meilleure valeur pour la quantité mesurée, je trouve bien naturel de penser que les mesures  $\lambda x_1, \dots, \lambda x_n$  conduisent à prendre  $\lambda x$ . Et cela n'est pas évident et n'est pas vrai pour n'importe quelle loi. Je pense, ayant ainsi précisé la manière dont j'interprète l'axiome II de S, qu'il ne peut y avoir ambiguïté. D'ailleurs l'énoncé est correct, car si l'on énonce la loi algébriquement, (sans y mettre de paramètre), les deux cas que je considère correspondent bien aux mêmes grandeurs énoncées avec deux unités différentes.

Bien cordialement à vous.

P.Lévy.

## Lettre 12

Paris- le 25 octobre 1925

9 rue Chernoviz

Mon cher collègue,

<sup>46</sup>Subbotin's 1923 article in *Математический Сборник*. See note 40.

<sup>47</sup>See Sect. 4 and the long discussion in Fréchet's article Fréchet, Maurice (1925) Sur la loi des erreurs d'observation. *Математический Сборник* 33:705–710.

Nous sommes bien à peu près d'accord, mais il y a tout de même une petite différence entre nos deux manières de voir.

Pour moi, lorsqu'on me dit : "En changeant A, on ne change pas B", je considère que changer A, c'est changer quelque chose. Autrement la phrase est une tautologie.

Lorsqu'on me parle de changer d'unité, dans le cas qui nous occupe, je suis forcé de comprendre qu'on me parle, non de la même loi et de sa nouvelle expression analytique, car cela ne changerait rien, mais de la nouvelle loi qui après le changement d'unité a la même expression analytique que l'ancienne loi avec l'ancienne unité.

Je pense être ainsi plus clair que dans ma lettre de vendredi,<sup>48</sup> que je n'ai mise à la poste qu'hier.

Bien cordialement à vous.

Paul Lévy

### Lettre 13

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.

Quand on parle d'une loi d'erreur, j'admet qu'on parle de quelque chose de bien déterminé. Si avec la même loi d'erreur, des mesures d'une quantité  $X$  ont donné  $x_1, x_2, \dots, x_n$ , la valeur probable (ou la plus probable) qui en résulte pour  $X$  est quelque chose de bien déterminé- indépendant du système des notations ou des unités employées. Le principe d'homogénéité, dans le sens où l'entend votre note, n'est véritablement qu'une tautologie.

J'ai eu tort de vous écrire que je sous-entendais que la probabilité d'une erreur  $x - X$  était fonction de  $x - X$  ; il suffit que ce soit une fonction déterminée de  $x$  et de  $X$  et il ne peut en être autrement ; (seulement si elle n'est pas fonction de  $x - X$  seulement, il faut bien comprendre que l'on n'a pas le droit d'appliquer indifféremment la même expression analytique à cette loi de  $X$  et à  $X.\text{const}$  ; ce seraient deux lois différentes)

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.

---

<sup>48</sup>Friday October 23, 1925. This refers to the preceding letter.

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" \* [*Note de Lévy en marge* : \* bien entendu, je prends cette phrase rapprochée du texte qui la précède dans votre note ; car si je la voyais isolée je pourrais très bien la comprendre autrement] 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

#### Lettre 14

Paris - 9 rue Chernoviz le 30/11/25<sup>49</sup>

Mon cher collègue,

J'ai bien reçu votre lettre. Vous pouvez naturellement imprimer mes 10 lignes à la suite de votre article,<sup>50</sup> dans quelque recueil que ce soit ; je n'ai aucune préférence. Bien cordialement à vous P.Lévy

<sup>49</sup>This is a post card addressed to 'Monsieur Fréchet 141a, rue principale Oberhausbergen par Strasbourg,' with a blue stamp 'semeuse 30 centimes' postmarked 30/11/25 at the post office on the rue de Longchamp, Paris.

<sup>50</sup>At the end of his article Fréchet, Maurice (1925) Sur la loi des erreurs d'observation. Математический Сборник 33:705–710, Fréchet included the following text under the title 'Lettre adressée à M.Fréchet par M.Paul Lévy après lecture du manuscrit de l'article ci-dessus' (Letter sent to M. Fréchet by M. Paul Lévy after reading the manuscript of the above article):

Paris, 9 rue Chernoviz, le 23 novembre 1925

Mon cher collègue,

Vous avez très clairement exposé mes idées.

Notre discussion se réduit au fond à bien peu de chose, car on peut appeler ce qu'on veut "axiome II" ; le point important est celui que vous indiquez à la fin de votre article : "peut-on considérer l'axiome en question comme une hypothèse fondée sur le sens commun ?".

Dans mon interprétation, l'axiome II est un principe d'homogénéité qui revient à dire qu'il n'y a pas d'unité privilégiée. Bien que, comme vous le rappelez, je préfère chercher la justification de la loi de Gauss dans des considérations tout à fait différentes, je suis porté à attribuer une certaine vraisemblance à un tel principe, et ne suis sans doute pas le seul.

Pour ce qui est de votre interprétation, je crois que nous sommes d'accord : on ne peut pas mettre en doute la valeur de l'axiome ainsi interprété, mais on ne peut rien en tirer.

Bien cordialement à vous, P.Lévy

P.S.- Au cas où vous désiriez publier la présente lettre à la suite de votre article, je n'y ferais aucune objection.

**Lettre 15**

Paris 14/10/28

Mon cher Collègue,

Merci d'abord pour la photo du bain de pieds dans l'Adriatique ; elle a amusé mes enfants et nous a permis d'évoquer de bons souvenirs.

Pour la question qui nous occupe, les choses me paraissent bien claires.

- 1° Le résultat de Hausdorff - Whittaker est exact, et vous l'avez inexactement cité dans votre article.<sup>51</sup> Mais ce résultat exact est banal, et il faut n'avoir pas compris la loi des grands nombres pour le donner comme une exception.<sup>52</sup>

La formule de composition adoptée est en effet

$$\varepsilon = \frac{2}{\pi} \left( \varepsilon_1 + \frac{1}{3}\varepsilon_2 + \frac{1}{5}\varepsilon_3 + \dots \right)$$

or, dans l'addition des erreurs, les poids sont proportionnels aux carrés et comme  $\sum \frac{1}{(2n+1)^2}$  est fini, chaque erreur partielle est une fraction finie de l'erreur totale. Or on ne peut appliquer la loi des grands nombres que quand chaque erreur est infiniment petite par rapport à l'erreur totale.

L'importance de cette condition de divergence non réalisée dans l'ex ci-dessus est indiquée à la page 236 de mon livre.<sup>53</sup> Quant au résultat tel que vous le citez, il serait faux.

- 2° Je pense maintenant à ce que vous me dites pour soutenir, avec quelques atténuations, votre idée que la loi forte des grands nombres implique que dès le début on ait des lois se comportant à l'infini comme celle de Gauss.

J'ai dit en effet, et je reste de cet avis, qu'avec la relation de Fourier

$$\varphi(z) = \int_{-\infty}^{+\infty} e^{izx} f(x) dx,$$

l'allure de  $\varphi(z)$  à l'origine est liée à celle de  $f(x)$  à l'infini. Mais précisons cette relation. La formule

$$\frac{d^p \varphi(z)}{dz^p} = i^p \int_{-\infty}^{+\infty} e^{izx} x^p f(x) dx$$

---

<sup>51</sup>Here, following Fréchet's 1928 article Fréchet, Maurice (1928) Sur l'hypothèse de l'additivité des erreurs partielles. Bull. Sci. Math. 52:203–216, an ongoing discussion of the law of errors began again. In the 1928 article, Fréchet, questioning the principle of additivity, examines the situation when one instead considers the maximum error. In 1929, in an attempt to complete the discussion, Lévy published an article entitled Sur quelques travaux relatifs à la théorie des erreurs in the Bulletin des Sciences Mathématiques. See Sects. 4 and 5.

<sup>52</sup>A not very diplomatic remark, especially since Lévy had read Fréchet's article too quickly; see Letter 17.

<sup>53</sup>This refers to the Lyapunov condition, which shows that  $\sum_{k=1}^n X_k / \sum_{k=1}^n E(X_k^2)$  converges to a Gaussian law if  $X_n$  is a series of independent variables with zero mean such that  $\sum_n E(X_n^2) = +\infty$ .

nous montre que la convergence de  $E_p = \int_{-\infty}^{+\infty} x^p f(x) dx$ , ( $f(x) \geq 0$ ) suffit à entraîner l'existence des dérivées de  $\varphi(z)$  jusqu'à l'ordre  $p$ , les valeurs de ces dérivées à l'origine étant liées à  $E_1, E_2, \dots, E_p$ .

Supposons seulement l'existence de  $E_2$  et par suite de  $E_1$ . Par un simple changement d'unité et d'origine on a  $E_1 = 0, E_2 = 1$  d'où

$$\varphi(z) = 1 + 0z - \frac{1}{2}z^2 + \dots \quad \text{et} \quad \psi(z) = \log \varphi(z) = -\frac{z^2}{2} [1 + \omega(z)],$$

$\omega(0)$  étant nul. Donc le fait que  $\psi(z)$  soit équivalent à  $-\frac{1}{2}z^2$  est bien lié à une propriété de  $f(x)$  à l'infini. Mais cette propriété à l'infini, c'est simplement la convergence de  $\int_{-\infty}^{+\infty} x^2 f(x) dx$ .

Donc  $f(x)$  peut se comporter comme  $\frac{1}{1+x^4}$ , ou  $e^{-|x|}$ , ou  $e^{-x^2}$ , ou bien tendant vers zéro plus rapidement encore, ou même être nul à partir d'une certaine valeur de  $x$  (comme dans la théorie classique du jeu de pile ou face), on peut toujours appliquer la loi des grands nombres. Vous voyez que vous donnez une idée inexacte de cette condition suffisante (et presque nécessaire, la convergence de  $\int_{-\infty}^{+\infty} x^\alpha f(x) dx$  pour tout  $\alpha < 2$  est nécessaire), disant que  $f(x)$  doit se comporter à l'infini comme la fonction  $e^{-x^2}$ .

- 3° Je crois en effet qu'il n'y a rien à corriger dans votre article polonais,<sup>54</sup> mais pour moi il est sans grand rapport avec la théorie des erreurs, en raison de la loi de composition adoptée qui ne me paraît guère avoir de chance de se rencontrer en pratique.<sup>55</sup>
- 4° Je pense que vous ne pouvez qu'être d'accord avec moi sur la matérialité des faits qui résultent des 1° et 2° ci dessus. Sauf avis contraire de vous, je considérerai qu'il en est ainsi, et que je puis faire état de votre accord.

Je vais sans doute profiter tout de même de l'occasion pour envoyer un article sur la question au Bull.d Sc.<sup>56</sup> Je voudrais à la fois plaider encore une fois la cause de la loi de Gauss, si je puis dire ainsi, mais en même temps me défendre de l'erreur absurde de croire qu'elle s'applique partout.

Je pense avoir un peu de temps maintenant, d'autres recherches que j'avais [*La suite verticalement en marge gauche* : entreprises se terminent par un échec complet ; j'ai travaillé pendant plus d'un mois et Cerf plus de 8 jours, sur un système dont l'intégrale générale devait être prévue a priori ; le fait qu'il n'y a pas d'autre intégrale que celles prévisibles a priori rend ce résultat peu intéressant, alors qu'une intégrale nouvelle eût été intéressante. Enfin ce résultat vient d'être publié par Mineur<sup>57</sup> dans la Soc.Maths. Malgré cela il faut que je publie une Note pour

<sup>54</sup>See note 81.

<sup>55</sup>Here Lévy returns to Fréchet's idea of considering the maximum error, published in Fréchet's 1929 article in the Bulletin des Sciences Mathématiques: Sur quelques travaux relatifs à la théorie des erreurs. See Letter 20 and those following.

<sup>56</sup>See note 51.

<sup>57</sup>Mineur, Henri (1928) Sur les Ondes de gravitation. Bull. Soc. Math. Fr. 56:50–73.

mettre les choses au point et rectifier une erreur de ma première Note ; mais après cela je pourrai faire autre chose.

Le nouveau livre de Lebesgue<sup>58</sup> va faire concurrence au vôtre. On voudrait se tenir au courant, et la Science marche trop vite ; on ne peut que se cantonner dans une toute petite spécialité. J'ai été stupéfait des progrès récents de la chimie physique exposés récemment par Langevin. On en arrive à prévoir les propriétés des atomes et leurs groupements en molécules, en partant d'une hypothèse unique sur la forme de l'équation définissant les champs électromagnétiques du proton et de l'électron ; je croyais rêver.<sup>59</sup>

Bien cordialement

P.Lévy.]

## Lettre 16

Paris 17 octobre 1928

Mon cher Collègue,

J'ai mis- plutôt je vais mettre à la poste, en même temps que cette lettre, votre Note pour M. Picard.<sup>60</sup> Je n'ai pas d'objection, et je trouve en effet qu'il vaut mieux que vous rectifiez vous-même vos petites erreurs matérielles. J'ai pris sur moi de corriger un lapsus (vous aviez écrit convergente au lieu de divergente). Il m'a semblé qu'il n'était pas la peine de vous renvoyer la Note pour cela, d'autant plus que vous verrez les épreuves.

J'ai plus travaillé dans Bertrand, Poincaré, Borel que dans les vieux auteurs, et hésite toujours beaucoup sur les questions de priorité<sup>61</sup> C'est Polya qui m'avait signalé la priorité de Cauchy sur les fonctions caractéristiques, et j'ai indiqué quelque part que Poisson devait avoir eu une vague idée de ces questions ; je l'avais vu par hasard dans un mémoire de Bienaymé (Journal de Liouville 1867- sauf erreur le même N° contient la démonstration du Th. de Bernoulli par Tchebichef).<sup>62</sup> Quant à Laplace, je crois que sa fonction génératrice n'a aucun rapport avec la fonction

<sup>58</sup>In 1926 Lebesgue, at the request of Gauthier-Villars, wrote a second edition of his famous *Leçons sur l'intégration et la recherche de fonctions primitives* which appeared in 1928. A presentation of the book by Lebesgue himself was published in the *Bull. Sci. Math.* 53, in 1929. In the same volume, Bouligand reviewed Fréchet's book *Les Espaces abstraits et leur théorie considérée comme introduction à l'analyse générale* which was published concurrently by Gauthier-Villars.

<sup>59</sup>Interesting evidence of the unity of science at that time, and of the attention Lévy paid to mathematical physics. The 1920s saw great progress in the study of chemical bonds, following de Broglie's introduction of wave mechanics in 1924. In 1926, E. Schrödinger's model of the hydrogen atom appeared, and the following year, the memoir of W. Heitler and F. London on the stability of the H<sub>2</sub> molecule opened the way for theoretical chemistry.

<sup>60</sup>This 'note' by Fréchet is in fact the letter (reproduced in note 87) included at the end of an article by Lévy; see Letter 24. The delivery of this letter is a story in itself; see Letter 24.

<sup>61</sup>The question of priorities for the method of characteristic functions reappears periodically in Lévy's writings. See, for example, Letter 30.

<sup>62</sup>The article by Bienaymé which improves the method of least squares was in fact published in 1952 in the *J. Math. Pures Appl.* 17:33–78, then reprinted in the *Journal de Liouville* in 1867 as an introduction to an article by Chebyshev.

caractéristique ; j'avoue l'avoir seulement parcouru, ayant l'impression qu'on avait fait des progrès depuis et pensant que ce qu'il y a d'essentiel dans Laplace devait se trouver dans les ouvrages de didactique modernes- car je crois que les auteurs de ces ouvrages ne sont pas en général comme moi, qui aime réfléchir aux questions qui m'intéressent mais ne lis qu'avec effort ce qu'on fait les autres. Malgré tout cela, je serais surpris que Cauchy ne soit pas le premier à avoir développé d'une manière systématique, les applications de la fonction caractéristique.

- Ayant Abel et Laplace chez moi, je viens d'interrompre ma lettre pour y jeter un coup d'œil
- Abel A l'endroit cité, il y a bien un Mémoire sur les fonctions qu'il appelle génératrices. Il donne quelques propriétés, maintenant banales, relatives à leurs développement en séries ; quelques exemples simples. C'est au fond un mémoire sur la formule de Fourier, il l'applique au cas de plusieurs variables. Mais aucune indication du fait que la relation

$$\int f_1(y)f_2(x-y)dy = f(x)$$

pour une fonction, entraîne pour l'autre

$$\varphi_1(z)\varphi_2(z) = \varphi(z)$$

C'est ce fait - que j'avais déjà appliqué à l'étude des fonctions de Green (Circolo 1912, 2° sem p.192)<sup>63</sup> - qui est fondamental en calcul des probabilités.

- Laplace - La fonction génératrice est tout autre chose. C'est une série de Taylor, et ce qu'elle engendre, c'est la suite des coefficients. En jetant un coup d'œil sur la table des matières, je ne vois pas où il pourrait être question de fonction caractéristique (du moins dans son Calcul des Probabilités t.7 de ses œuvres complètes).

Pour Lebesgue, c'est bien une deuxième édition de ses Leçons sur l'Intégration ; elle a dû paraître ces jours-ci.<sup>64</sup>

- Je croyais votre nomination à Paris absolument sûre, et suis surpris d'apprendre que vous êtes encore dans l'attente. Bauer, lui, s'est installé à Paris avant d'être sûr. En tout cas, je pense, en ce qui vous concerne, que votre venue ne peut plus beaucoup tarder.<sup>65</sup>

Bien cordialement  
P.Lévy.

---

<sup>63</sup>Evidence, yet again, of a connection between Lévy's work on probability and his previous research in functional analysis.

<sup>64</sup>See note 58.

<sup>65</sup>At the end of 1928, at the request of Borel, Fréchet was indeed named Professor—but without chair, which means that he could be employed to teach any mathematical lectures if needed—in the Faculté des Sciences de Paris.

**Lettre 17**

Paris, 9 rue Chernoviz

Le 29 octobre 1928

Mon cher Collège,

Je retrouve votre tirage à part du Bull.d Sc.,<sup>66</sup> que j'avais égaré dans un tas de papiers mis dans une malle quand je suis parti pour l'Italie, et que je n'avais pas sous les yeux en vous écrivant, il y a une quinzaine de jours. Je m'aperçois avec stupéfaction que votre citation d'Hausdorff est parfaitement correcte.

C'est moi qui ai commis l'erreur de ne pas voir de suite que le résultat d'Hausdorff s'appliquait pour la convergence de la série  $\sum \frac{1}{(2n+1)^2}$ , chaque erreur partielle ne pouvant pas être dans ces conditions considérée comme très petite par rapport à l'erreur totale. Cette circonstance m'a au contraire frappé quand vous m'avez envoyé la traduction exacte du passage cité de H ; cela tient sans doute à ce que j'ai sous les yeux explicitement une formule du type

$$x = \frac{x_1}{3} + \frac{x_2}{5} + \dots + \frac{x_n}{2n+1} + \dots,$$

et, ne pouvant croire que cette circonstance ne m'avait pas frappé la première fois, croyant d'ailleurs que vous aviez écrit  $x = \frac{x_1 + \dots + x_n}{n}$ , je vous ai écrit que votre citation était inexacte.

Je m'excuse de cette erreur, je devrais même dire de ces erreurs, puisque j'ai successivement, et à tort, accusé le résultat de H. d'être inexact et vous de l'avoir mal cité.

Je m'empresse de vous prévenir, car il se peut que cela vous fasse changer quelque chose, au moment de la correction des épreuves, à votre Note rectificative. J'ajoute d'ailleurs que, à mon avis, cette rectification, ou plutôt cette addition, reste nécessaire pour empêcher le lecteur de rester sur l'impression que le résultat de H est en contradiction avec les énoncés généralement admis pour la loi des grands nombres.

Quant au fond de votre article, là encore, je vois que je me le rappelais mal, et vous croyais plus hostile à la loi de Gauss que vous ne l'êtes en effet. Nous ne sommes séparés que par des nuances de pensée.<sup>67</sup>

Ainsi, je suis bien d'accord sur ce que ma justification de la loi de Gauss est une explication, et non une démonstration logique ; et il ne peut en être autrement, puisque à la base d'une construction logique il faut des axiomes que l'on explique, mais que l'on ne démontre pas ; je sais déjà que nous sommes d'accord, au point de vue philosophique, sur le rôle des axiomes. Seulement cette explication ne vous satisfait pas ; elle me satisfait pleinement, en ce sens qu'elle me fait comprendre pourquoi la loi de Gauss joue un rôle particulier dans un grand nombre de questions

---

<sup>66</sup>Fréchet, Maurice (1928) Sur l'hypothèse de l'additivité des erreurs partielles. Bull. Soc. Math. Fr. 52:203–216.

<sup>67</sup>The discussion was concluded with Lévy's exposition of their respective points of view in §5 of his 1929 article in the Bulletin des Sciences Mathématiques. See note 83.



et pourquoi on la rencontre si souvent, avec un degré d'approximation plus ou moins grand ; elle m'aurait même permis de la prévoir a priori.

Au point de vue théorique, votre objection à la formule d'addition des infiniment petits est incontestable, et il n'est sans doute pas impossible de trouver d'application possible de votre théorie. Seulement ce que je dis, c'est que dans la plupart des cas où l'on a prétendu appliquer la loi de Gauss, la composition des erreurs se fait bien suivant une formule qui équivaut à l'addition pour les erreurs très petites ; que du moins, si je ne peux le démontrer, cela me paraît bien naturel de l'admettre, à l'exclusion de toute autre formule. Si demain vous rencontrez un cas différent où ce soit votre formule qui paraisse s'appliquer le plus naturellement, je ne ferai aucune difficulté à admettre que l'on est dans un cas échappant à la théorie classique ; (ce qui n'empêche pas qu'en attendant votre formule me paraisse assez artificielle, mais ce qui est artifice aujourd'hui, peut cesser de l'être demain).

Faisons une comparaison. Il s'agit par exemple d'expliquer pourquoi il pleut plus dans les Cévennes qu'en Champagne ;<sup>68</sup> on m'explique que les montagnes accrochent les nuages venus de l'Océan. On m'explique de même qu'il pleut moins au centre du Sahara à cause de la grande chaleur et l'éloignement de mers. Ces explications me satisfont. On peut m'objecter que ce ne sont pas des démonstrations, que je n'ai pas utilisé toutes les données et que l'influence d'une donnée négligée peut tout remettre en question. Je répondrai que j'ai tenu compte des données essentielles, et que j'ai une explication grossièrement mais sûrement exacte de phénomènes que mon ignorance m'empêche de comprendre exactement en détail.

De même, en calcul des probabilités, c'est pour moi une chose essentielle de comprendre la raison de cette convergence vers la loi de Gauss, que l'on observe si souvent, et qui n'a rien de mystérieux. C'est une chose utile d'observer que cette convergence n'a peut-être pas lieu dans tous les cas possibles et imaginables ; mais cette remarque ne doit pas faire perdre de vue l'essentiel, et je reste persuadé que c'est l'essentiel dans la théorie des erreurs que de mettre en évidence ce rôle de la loi de Gauss.

Ce rôle n'étant qu'asymptotique\* je ne m'attends d'ailleurs pas à une précision parfaite.

Je crois que tout cela n'est pas en contradiction avec vos idées. Ce qui finalement me surprend dans votre article, c'est ce que vous dites de l'insuffisance des vérifications expérimentales. J'aurais cru qu'une série de 500 observations, qui doit dessiner une courbe des probabilités totales assez précise, ne pouvait guère permettre d'hésiter entre les deux lois de Laplace. Mais j'avoue n'avoir pas étudié la question de très près, et que admettant qu'on ne peut trouver la loi  $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$  avec une précision absolue, je dois admettre que l'imprécision du résultat permette une certaine hésitation entre deux lois différentes<sup>69</sup>

---

<sup>68</sup>In the 1929 article, these places became Spain and France!

<sup>69</sup>In his 1924 article in *Математический Сборник*, Fréchet mentioned that one of his students had used 500 data from C.S. Pierce's experimental geodesic curves to test the validity of the Gaussian distribution of the errors, and that the test was non conclusive.

Je ne suis toujours pas décidé à vous répondre dans une Note dans le Bulletin, mais comme vous le voyez, il ne peut s'agir que de discuter des nuances de pensée. Si j'ai cru la discussion plus grave, c'est sans doute que je vous avais lu trop vite à un moment où j'étais plongé dans des calculs sur la généralisation du  $ds^2$  de Schwarzschild,<sup>70</sup> et je m'excuse encore de ce que j'ai pu dire d'inexact.

Bien cordialement

P.Lévy.

[*In left margin*: \* doublement, puisqu'il faut à la fois que  $n$  soit très grand et la mesure précise.]

### Lettre 18

Paris - 9 rue Chernoviz

Le 3 novembre 1928

Mon cher Collègue,

Je n'ai pour le moment à vous répondre qu'au sujet du battage des cartes. On passe des  $\alpha_{i,j}^{(n)}$  aux  $\alpha_{i,j}^{(n+1)}$  par la formule

$$\alpha_{i,j}^{(n+1)} = \alpha_{i,1}^{(n)}\alpha_{1,j} + \dots + \alpha_{i,p}^{(n)}\alpha_{p,j},$$

et pour passer de l'inégalité - supposée vraie pour  $n$ - à l'inégalité analogue pour  $n + 1$ , il suffit d'appliquer et de préciser la remarque connue que la moyenne

$$x = x_1\alpha_1 + \dots + x_p\alpha_p$$

(où  $\alpha_1, \dots, \alpha_p \geq 0$  et  $\sum \alpha_n = 1$ ) est comprise entre le plus petit et le plus grand des  $x_i$ . Supposons qu'on ait pour les  $x_i$  les limites extrêmes  $1/p - \varepsilon$  et  $1/p + \varepsilon$  en marge gauche : l'une des deux étant effectivement atteinte, et que le plus petit des  $\alpha_i$  soit  $\geq \mu$ .

$$x - 1/p = \sum \alpha_i \left( x_i - \frac{1}{p} \right) = \sum' \alpha_i \left( x_i - \frac{1}{p} \right) + \sum'' \alpha_i \left( x_i - \frac{1}{p} \right).$$

$\sum'$  contenant les termes positifs,  $\sum''$  les termes négatifs. Si d'ailleurs tous les  $(x_i - \frac{1}{p})$  ne sont pas nuls, le plus grand des  $|x_i - \frac{1}{p}|$  ayant été supposé  $= \varepsilon$ , on a<sup>71</sup>

$$\sum' \left| x_i - \frac{1}{p} \right| = \sum'' \left| x_i - \frac{1}{p} \right| \geq \varepsilon$$

$\sum' \alpha_i (x_i - \frac{1}{p}) < \varepsilon(1 - \mu)A$  (en remplaçant les  $x$  par le plus grand)

<sup>70</sup>The metric introduced by Karl Schwarzschild (1873–1916) which allowed an explicit solution of Einstein's gravitational equations. See note 57.

<sup>71</sup>In fact,  $\sum' = \sum'' = \varepsilon + (\text{sum of terms} \leq \varepsilon) \geq \varepsilon$ . See Lévy's better exposition of the argument in the next letter, which was written for Fréchet, who was not satisfied with the proof given by Lévy here.

$\sum'' \alpha_i (x_i - \frac{1}{p}) |> -\varepsilon(1 - \mu)A$  (en remplaçant les  $x$  par le plus petit)

$$\text{d'où } x - \frac{1}{p} < \varepsilon A(1 - 2\mu)$$

et de même  $x - \frac{1}{p} > -\varepsilon A(1 - 2\mu)$  d'où  $|x - 1/p| \leq A(1 - 2\mu)$  formule de récurrence que l'on applique en remplaçant  $\alpha_k$  par  $\alpha_{k,j}$  et  $x_k$  par  $\alpha_{k,i}^{(n)}$ . On a ainsi une formule de récurrence d'où résulte la formule à démontrer.

Je n'ai pas développé le calcul, qui n'est pas aisé à reconstituer. Je croyais d'ailleurs me rappeler qu'il était dans Poincaré,<sup>72</sup> mais, d'après ce qu'Hadarnard m'a dit depuis, j'ai dû me tromper, je ne l'ai pas recherché n'ayant pas Poincaré chez moi.

Ce calcul a été repris et développé par Hostinsky<sup>73</sup> et Hadarnard Comptes Rendus 1928 1er semestre<sup>74</sup> qui, outre l'extension où les  $\alpha_{k,j}$  varient d'un coup à l'autre, ont étudié le cas où  $p$  augmente indéfiniment. On a alors une formule du type

$$f_{n+1}(t) = \int_0^1 f_n(t)\alpha_n(t) dt, \quad \alpha_n > \mu \text{ et } \int_0^1 \alpha_n dt = 1,$$

et on obtient de même la tendance vers une valeur constante de  $f(t)$ , suivant une loi exponentielle. M. Hadarnard a développé ces circonstances et leurs applications physiques à Bologne ;<sup>75</sup> mais le synchronisme rigoureux de sa communication et de la mienne m'a empêché de l'entendre.

Je ne retrouve pas la référence du Mémoire développé de Hostinsky ; mais ses Notes des Comptes Rendus vous montreront de quoi il s'agit. J'ajoute que cela ne m'a rien appris ; je savais tout cela quand j'ai fait mon livre et n'ai eu que le tort de ne pas le développer. Mais mon but n'était que de faire comprendre sur un cas simple la tendance au nivellement déjà signalée, quoique peut-être d'une manière moins précise, par Poincaré.<sup>76</sup>

Comme travaux récents sur le calcul des probabilités m'ayant intéressé, je vous signale ceux de Harald Cramer (Skandinavisk Aktuarietidskrift, 1928) il précise la tendance vers la loi de Gauss en limitant supérieurement

$$F_n(x) - \int_{-\infty}^x e^{-\frac{\xi^2}{2}} d\xi,$$

<sup>72</sup>Poincaré, Henri (1912) Calcul des Probabilités (Deuxième édition). Gauthier-Villars, Paris. Chapitre XVI, pp. 301–314.

<sup>73</sup>(1928) CRAS 186:59–61. On Hostinsky's works, consult Bru (2003) and Mazliak (2007).

<sup>74</sup>(1927) CRAS 185:5–9.

<sup>75</sup>Hadarnard, Jacques (1928) Sur le battage des cartes et ses relations avec la mécanique statistique. Rendiconti del Congresso Internazionale dei matematici, Bologna, 1928 5:133–139.

<sup>76</sup>Poincaré's method is perhaps less precise and more complicated, but it is certainly more general. See Fréchet's book Fréchet, Maurice (1938) Méthode des fonctions arbitraires. Traité du Calcul des Probabilités et de ses applications; fascicule 3. Gauthier-Villars, Paris.

$F_n$  désignant la fonction des probabilités relative à  $\frac{x_1 + \dots + x_n}{n}$ .

Il est probable que je ne vais rien publier sur notre discussion, je n'aurais à faire que de petites remarques qui n'en valent pas la peine.

Bien cordialement

P.Lévy.

### Lettre 19

Paris- 9 rue Chernoviz

le 9/11/28

Mon cher Collègue,

Il me semble que ma lettre du 3 est bien exacte, sauf la petite correction que je viens faire au crayon au point où vous êtes arrêté par mon  $\varepsilon$ . Dans la 1ère parenthèse explicative, que je me souviens avoir rajoutée après coup, il suffit d'écrire  $x_i - 1/p$  au lieu de  $\alpha_{ii}$ , je l'avais ajouté au dernier moment, me disant que vous ne verriez peut-être pas comment j'avais opéré, et oubliant déjà que je n'avais pas opéré de même pour les deux formules. Pour plus de clarté, je reprends les formules

pour les  $x_i$  un des  $|x_i - \frac{1}{p}| = \varepsilon$  les autres  $\leq \varepsilon$ <sup>77</sup>

$$\sum' \left| x_i - \frac{1}{p} \right| = \sum'' \left| x_i - \frac{1}{p} \right| \geq \varepsilon$$

pour les  $\alpha_i$  un des  $\alpha_i = \mu$ , les autres  $\geq \mu$ <sup>78</sup>

$$\sum' \alpha_i = 1 - \sum'' \alpha_i \leq 1 - \mu, \quad \sum'' \alpha_i \leq 1 - \mu$$

Donc

$$\begin{aligned} \sum' \alpha_i \left( x_i - \frac{1}{p} \right) &\leq \varepsilon \sum' \alpha_i \leq \varepsilon (1 - \mu) \\ \left| \sum'' \alpha_i \left( x_i - \frac{1}{p} \right) \right| &\geq \mu \sum'' \left| x_i - \frac{1}{p} \right| \geq \varepsilon \mu \\ \sum'' \alpha_i \left( x_i - \frac{1}{p} \right) &\leq -\varepsilon \mu \\ \sum \alpha_i \left( x_i - \frac{1}{p} \right) &\leq (1 - 2\mu)\varepsilon \end{aligned}$$

Vous remarquerez que le calcul se simplifie si l'on se contente de  $1 - \mu$  au lieu de  $1 - 2\mu$ .

<sup>77</sup>See note 71.

<sup>78</sup>One has  $\sum'' \alpha_i \geq A'' \mu \geq \mu$  because  $A''$ , the number of terms in  $\sum''$  is  $\geq 1$ . Thus  $1 - \sum'' \leq 1 - \mu$ .

Peut-être ai-je eu tort de ne pas mettre dans ma précédente lettre  $\leq$  (au lieu de  $<$ ) pour la formule de récurrence. L'égalité est en effet possible pour  $p = 2$ ,  $\alpha_1 = \mu$ ,  $\alpha_2 = 1 - \mu$ ,  $x_1 - \frac{1}{2} = -\varepsilon$ ,  $x_2 - \frac{1}{2} = \varepsilon$ . Mais la formule de mon livre reste exacte puisqu'on part d'une valeur de  $|\alpha_{i,j} - \frac{1}{p}|$  qui est  $1/2 - \mu \leq (1 - 2\mu)$  pour  $n = 1$ , sauf pour  $\alpha_1 = \alpha_2 = 1/2$  cas où l'égalité est réalisée dès le début.

2° Il est bien évident que si l'égalité  $\alpha_{i,j}^{(n)} = 1/p$  est réalisée pour  $n$  elle est réalisée pour  $n + 1, n + 2, \dots$ . Je n'avais pas cru utile dans mon livre d'insister sur un point si évident. Autrement, il m'eût été facile de poser

$$m_n = \max \text{ de } |\alpha_{i,j}^{(n)} - 1/p|,$$

et d'écrire explicitement la relation de récurrence  $m_{n+1} \leq m_n \cdot (1 - 2\mu)$ .

3° Je ne suis pas sûr d'avoir reconstitué ma démonstration de 1924-25. Je me dis qu'elle était peut-être encore plus simple, et je ne l'ai même pas mise en note (mon plan me rendant difficile de la mettre dans le texte) ; c'est qu'elle m'avait paru excessivement simple et que j'étais convaincu que n'importe qui la reconstituerait aisément. Mes souvenirs sont précis sur ce point, et je vous avoue que j'ai été un peu surpris l'autre jour du petit effort qu'il m'a fallu pour retrouver le résultat. Mais je suis sûr d'être bien passé par la formule de récurrence  $m_{n+1} \leq m_n \cdot (1 - 2\mu)$ .

4° Je précise que je n'ai ni avant ni après mon livre, rien publié d'autre sur le sujet. J'avais d'ailleurs bien compris que c'était pour vous le point essentiel.

J'étais resté sur l'impression que je n'avais fait que répéter Poincaré, avec quelques modifications de détails, jusqu'à il y a environ 6 mois. A ce moment Hadamard m'a dit : "Je viens seulement de découvrir ce que vous dites du battage des cartes ; vous dites que vous citez Poincaré, ce qui fait que j'ai passé la nuit sur ce paragraphe ; or Poincaré n'a rien fait de si simple"<sup>79</sup>

Je regarderai tout de même un jour de nouveau Poincaré pour savoir à quoi m'en tenir.

Je vous avoue que je n'ai pas l'impression d'avoir fait du nouveau quand j'ai simplifié un détail dans une démonstration d'un autre ; seulement quand je vois d'autres se flatter d'avoir trouvé une méthode nouvelle simple, il m'arrive de répondre "Il y a longtemps que j'enseigne cela sans savoir que c'est nouveau". C'est par Hostinsky que j'ai appris que j'avais fait du nouveau pour la battage des cartes. Autant l'histoire de la Science, dans ses grandes lignes, est importante, autant je me dis qu'on attache en général trop d'importance à la priorité de tel ou tel détail. Mais je puisqu'il en est ainsi et que l'homme est homme, je me laisse aller à réclamer ma priorité quand j'apprends ainsi avoir fait une découverte !

Bien cordialement.

P.Lévy.

---

<sup>79</sup>See note 76.

## Lettre 20

Paris - 9 rue Chernowiz

le 13 novembre 1928<sup>80</sup>

Mon cher Collègue,

Je viens de revoir votre mémoire polonais,<sup>81</sup> et les pages 107 à 109 sur lesquels vous aviez attiré mon attention. Les circonstances qui se présentent dans cette question sont tout à fait analogues à celles qui se présentent dans le problème que j'ai étudié ;<sup>82</sup> dans le cas de convergence de la série  $A_1^\alpha, A_2^\alpha, \dots, A_n^\alpha, \dots$ , il ne saurait être question de convergence vers une loi limite indépendante des lois composantes ; pour cette raison bien simple qu'une des variables, la première par exemple, est du même ordre de grandeur que la résultante des autres, et s'il y avait convergence, par hasard, avec un certain système de lois, elle disparaîtrait si on ajoutait en tête une variable de plus.

Par contre, si  $\sum A_n^\alpha$  diverge, chaque erreur partielle est négligeable devant l'ensemble et c'est dans ce cas que la convergence vers une loi limite, la loi stable de la classe, se manifeste. Il n'y a à ce sujet aucune différence essentielle entre ce problème et le mien.

Je crois que ce que vous dites est correct, mais vous avez l'air de considérer comme mystérieuse une circonstance bien simple. La résultante de  $n$  erreurs est (dans votre cas comme dans le mien) de la forme  $B_n x_n$ ,  $x_n$  obéissant à une loi bien déterminée, et  $B_n$  augmentant indéfiniment. Il est alors bien clair que la probabilité d'une inégalité de la forme  $B_n |x_n| < A$ ,  $A$  étant fini, tend vers zéro. La variable  $B_n x_n$  augmente indéfiniment. Sauf dans des cas infiniment peu probables, la variable  $x_n$  est finie et a une loi limite.

Il me semble alors que votre rédaction serait à revoir d'assez près. La remarque que  $H(Z) = 0$  est banale, et ne mérite pas, il me semble, l'importance que vous lui donnez, d'autant plus que dans un cas comme celui du N° VI, p.112,  $F_n(A_n x_n)$ , fonction relative à la  $n$ ième composante non réduite, tend vers zéro ; la variable résultante étant toujours  $\geq A_n x_n$ , c'est vrai a fortiori pour la loi  $H_n(Z)$  non réduite ;  $H_n(Z) \Rightarrow H(Z) = 0$ . Cela n'empêche pas que la loi réduite a une limite bien déterminée, appartient à la classe  $C_\alpha$ , contrairement à ce que l'énoncé en italique au milieu de la p.104 peut faire penser.

J'ai beaucoup insisté dans mon livre, p.201 et suivantes, sur ces difficultés que l'on rencontre en raisonnant sur les lois non réduites alors qu'on met au contraire en évidence des circonstances très simples en raisonnant sur les lois réduites.

---

<sup>80</sup>The five letters that follow are the continuation of the discussion about the law of errors begun in October (Letters 15, 16 and 17).

<sup>81</sup>Fréchet, Maurice (1927) Sur la loi de probabilité de l'écart maximal. Ann. Soc. Pol. Math. 5:93–116. One of the points under debate was that Fréchet's very elegant method for finding the law of the maximum  $F(x) = e^{-ax^{-\alpha}}$  is, *mutatis mutandis*, the same as that which allowed Lévy, starting in 1919, to find the characteristic function of the stable symmetric laws  $\varphi(t) = e^{-a|t|^\alpha}$ . At the end of the article which he wrote in response (see note 51), Lévy reproduced an extract from a letter which Fréchet sent him on November 22, 1928. See also Sect. 4.

<sup>82</sup>See Sects. 4 and 5.

Si vous réfléchissez un peu à ces circonstances, vous verrez que la solution générale du problème que vous posez en haut de la p.109 ne présente aucune difficulté. C'est un corollaire évident, qu'il est à peine utile d'énoncer, du fait que, la loi réduite tendant vers une limite, le coefficient de réduction augmente indéfiniment.

On peut par contre avoir une loi réduite mixte, comportant une probabilité  $\beta$  pour la valeur  $+\infty$  et une loi de probabilité continue, avec  $\int_0^{+\infty} f(x) dx = 1 - \beta$  pour les valeurs finies. Il faudrait définir d'abord les classes un peu plus générales  $C'_\alpha$ , comprenant les lois comprises entre  $C_{\alpha-\varepsilon}$  et  $C_{\alpha+\varepsilon}$ , et prendre moitié d'erreurs obéissant à  $C'_0$ , et le reste obéissant à un des  $C_\alpha$ . Dans mon problème, comme dans le vôtre, il me semble, on arrive à la loi mixte indiquée (Voyez mon livre p.268, étude du cas où  $\alpha = 0$  ; il suffit de composer les lois considérées à cet endroit et d'autres lois du domaine d'attraction de la loi de Gauss, et de prendre le coefficient de réduction déduit de l'étude de ces dernières, en négligeant les premières. L'ensemble est alors identique à ce qu'on obtient en composant une variable obéissant à la loi de Gauss réduite avec une variable ne pouvant prendre que les valeurs 0 et  $\infty$ . En composant, on trouve une probabilité nulle pour 0, une probabilité positive pour l'infini, et la loi de Gauss au facteur  $1 - \beta$  près pour les valeurs finies).

Bien cordialement à vous

P.Lévy

J'ai écrit un peu vite ; mais même s'il y a des inexactitudes dans les détails, je crois bien que l'esprit de ces remarques est exact.

## Lettre 21

Paris - 9 rue Chernoviz

le 21 novembre 1928

Mon cher Collègue,

Il faut que je précise ma dernière lettre pour bien vous montrer que le problème posé dans votre article polonais est facile à résoudre dans tous les cas.<sup>83</sup> Il n'y a d'ailleurs que trois cas à distinguer.

*Premier cas* : La série

$$B^\alpha = A_1^\alpha + A_2^\alpha + \dots + A_n^\alpha + \dots$$

est convergente. Vous avez élucidé ce cas.  $Z$  obéit à une loi limite bien déterminée, qui n'est en général pas la loi stable de la famille normale considérée.  $B$  étant fini, la distinction entre  $Z$  et la variable réduite  $z = \frac{Z}{B}$  est sans importance.

*Deuxième cas* : la série est divergente ;  $B_n^\alpha$  désignera  $A_1^\alpha + \dots + A_n^\alpha$  ; et  $\frac{M_n}{B_n}$  tend vers zéro ( $M_n$  étant le plus grand des nombres  $A_1, \dots, A_n$ ).

On pose alors  $Z_n = B_n z_n$  ; on est dans le cas où la loi des grands nombres s'applique sans difficulté (Voyez votre formule (39), écrite pour  $n$  fini, et passez à la limite). A la limite,  $z_n$  obéit à la loi non stable, qui donne une probabilité nulle à la

---

<sup>83</sup>The resolution of these points may be found in paragraph 8 of Lévy, Paul (1929) Sur quelques travaux relatifs à la loi des erreurs. Bull. Sci. Math. 53:11-32, as noted by Lévy at the end of this letter and already mentioned in Letters 15 and 17.

valeur 0 ;  $B_n$  augmente indéfiniment. Donc, quelque grand que soit  $\zeta$ , la probabilité de  $Z_n < \zeta$  tend vers zéro.

*Troisième cas* : La série est divergente ; mais  $\frac{M_n}{B_n}$  ne tend pas vers zéro. Alors les  $A_n$  ne sont pas bornés.

Dans la définition de la famille normale, il me suffit d'ailleurs de supposer que la fonction majorante  $\mu(z)$  existe pour  $z > \zeta$  ; comme elle doit s'annuler à l'infini, on peut en augmentant  $\zeta$ , supposer  $\mu(z) < \frac{1}{2}$  (Il est essentiel dans toute cette théorie de ne faire intervenir la fonction majorante que pour  $z$  très grand). Alors, pour toutes les lois réduites de la famille normale, la probabilité de l'intervalle  $(0, \zeta)$  est  $< p < 1$ .

Soit  $X_n = A_n x_n$  la  $n$  ième erreur composante. La probabilité de l'inégalité  $X_n < A_n$  est  $< p$  et comme les  $A_n$  ne sont pas bornés (et a fortiori ne tendent pas vers 0) on a une infinité de fois  $x < A_n \zeta$ , quelque soit  $x$  positif. On a donc une infinité de fois

$$F_n(x) \leq F_n(A_n \zeta) < p < 1.$$

La formule de composition relative aux lois non réduites  $H(x) = F_1(x) \cdots F_n(x) \cdots$  donne alors  $H(x) = 0$ .

Par contre, il est bien évident, s'il ne s'agit pas de familles normales, qu'on peut, tant dans le second cas que dans le troisième, s'arranger pour que  $H(x)$  soit positif. Il suffit, pour une certaine valeur  $\xi$  de  $x$ , de prendre arbitrairement les valeurs des  $F_n(\xi)$  de manière que le produit converge. Prenons par exemple

$$\text{pour } x < \varphi(n), \quad \log F_n(x) = -\frac{1}{n^2 x^2}, \quad (\varphi(n) \text{ indéfiniment croissante})$$

cela n'empêchera pas de déterminer  $F_n(x)$ , pour  $x$  assez grand, de manière que

$$\text{pour } x > \varphi(n), \quad \log F_n(x) = -\left(\frac{A_n}{x}\right)^\alpha.$$

les  $A_n$  ayant telles valeurs que nous voudrions de manière à nous placer dans celui des cas précédents que nous aurons choisi, et de raccorder les deux déterminations précédentes de manière que  $F_n(x)$  soit continu et constamment croissant.

Alors, pour tout  $x$  fixé,  $F(x)$  est positif et croissant.

- Je me décide à rédiger un nouvel article sur ces questions.<sup>84</sup> J'ose espérer que vous trouverez qu'il n'y a pas de bien grandes différences entre nos points de vue.

Bien cordialement

P.Lévy.

## Lettre 22

Paris- 9 rue Chernowiz

22/11/28

Mon cher Collègue,

---

<sup>84</sup>See note 83.



Je suis stupéfait que votre Note ne soit pas parvenue à Monsieur Picard.<sup>85</sup>

Il est vrai que je ne l'ai pas mise à la poste ; mais je l'ai déposée moi-même à l'Institut. Après avoir fermé la lettre que je venais de vous écrire je me suis aperçu que j'allais chez Gauthier-Villars (si je me rappelle bien c'était un mercredi matin et j'avais à corriger les épreuves de ma dernière note, on peut le vérifier par les dates) il m'a semblé aussi simple de passer par l'Institut que de chercher une boîte aux lettres- et plus sûr !

Mes souvenirs sont très précis ; je me vois traversant le Pont des Arts, apercevant une boîte aux lettres à l'entrée de la cour de l'Institut, trouvant que le carnet de timbres que j'ai généralement sur moi était épuisé ; après m'être demandé un instant si cela pouvait se faire, j'ai pris le timbre dont vous avez eu soin de munir votre lettre, qui était mal collé, afin de mettre tout de suite à la boîte la lettre que je vous écrivais (en bonne justice je vous dois 0F,50). Puis j'ai été de l'autre côté de l'horloge, ai trouvé la loge fermée, et ai mis la lettre dans l'ouverture ménagée dans la porte de la loge. Il m'a semblé que cette ouverture ne correspondait pas à une boîte, mais que la lettre tombait par terre, mais étant donnée cette installation le concierge doit être habitué à ramasser les lettres par terre.

Je pense que Monsieur Picard n'a pas déménagé et que j'ai bien eu raison d'aller du côté de la Coupole opposé à la cour de l'Institut. Mais même dans ce cas la lettre n'aurait pas du être perdue.

Le fait du timbre enlevé exclut toute hypothèses que quelqu'un ait pris la lettre pour en conserver le timbre. J'aime d'ailleurs à penser que le concierge de l'Institut est à l'abri d'un tel soupçon.

Donc je ne comprends pas ; je suis aussi sûr qu'il est possible de l'être cette fois. Je n'ai pas conservé la lettre poche restante et n'ai pas le doute qui serait possible si je l'avais mise à la poste avec d'autres.

Bien cordialement à vous.

P.Lévy.

### **Lettre 23**

Paris -9 rue Chernoviz

23/11/28

Mon cher Collègue,

Un petit mot pour vous dire que j'ai remis ce matin chez M.Picard le petit travail<sup>86</sup> qui constituera la réponse à votre Mémoire. La rédaction en était naturellement assez avancée l'autre jour ; mais il m'arrive souvent en me relisant de trouver qu'il faut tout recommencer et je ne pensais pas quand je vous ai écrit que j'aurais si vite fini.

Vous pouvez voir que j'ai tâché de concilier nos deux points de vue plutôt que de vous contredire.

J'ai profité de ce que je repassai chez le concierge de M.Picard pour examiner l'état des lieux ; cet examen a confirmé mes souvenirs. Le concierge m'a dit- ce dont

---

<sup>85</sup>See note 60.

<sup>86</sup>See note 83.

je me doutais- que l'ouverture était faite exprès ; et qu'il était impossible qu'une lettre mise là n'ait pas été remise à M.Picard. Je pense n'avoir rien à me reprocher qui ne m'empêche pas d'être navré que cette lettre que vous m'avez confiée ne soit parvenue à destination.

Bien cordialement

P.Lévy.

## Lettre 24

ECOLE NATIONALE SUPERIEURE DES MINES

LABORATOIRE

60, Boulevard Saint-Michel, VIe

TELEPHONE : [*Correction manuelle de Lévy* : Fleurus 58-17]

Paris, le 24 novembre 1928

Mon cher Collègue,

Votre extrait de lettre fera très bien à la fin de mon article ;<sup>87</sup> je verrai si cela n'entraîne pas pour le dit article des modifications de détail, mais en tout cas aucun changement essentiel.

Je trouve votre livre en arrivant ce matin à l'Ecole des Mines et vous en remercie bien vivement. C'est en effet un sujet qui méritait d'être mieux connu. A l'Ecole des Mines le temps dont je dispose ne me permet guère que de dire : il faut que vous sachiez que cela existe ; vous vous reporterez en cas de besoin aux ouvrages spéciaux.

---

<sup>87</sup>At the end of his article on the theory of errors in the Bulletin des Sciences Mathématiques, 1929, Lévy reproduced an extract from the letter that Fréchet sent him on November 22, 1928:

Je ne verrais que des avantages à voir mentionner dans le Mémoire que vous m'annoncez, mon accord avec vous sur l'utilité d'apporter les précisions suivantes à ma dernière publication dans ce Bulletin (mai 1928), *Sur l'hypothèse de l'additivité des erreurs partielles*. L'exemple de M.Hausdorff que j'avais signalé concerne le cas où la somme des carrés des écarts quadratiques moyens des erreurs composantes converge. Il n'est donc pas en contradiction avec le théorème fondamental de la page 236 de votre livre où vous avez formellement exclu le cas d'une telle convergence. L'exemple m'a paru intéressant parce qu'il donnait un exemple explicite simple de la nécessité de cette exclusion. Il n'était d'ailleurs parvenu à ma connaissance qu'au moment de corriger les épreuves de mon article. D'autre part, il y aurait lieu de préciser le sens du troisième alinéa de mon article, page 203, où j'écrivais : "Tout d'abord, les lois de probabilités admises pour les erreurs ne sont pas aussi générales qu'elles paraissent : elles sont infiniment voisines, pour  $x$  très grand, de l'intégrale de Laplace." La dernière phrase visait la formule (74) de la page 212 de votre livre :

$$\overline{\psi(t)} = -\frac{t^2}{2} [1 + \omega(t)],$$

où  $\omega(t)$  est petit avec  $(t)$ , et elle laissait de côté les autres conditions. Mon observation risquait donc de suggérer à la fois trop et trop peu. Il convient donc de lire : "Tout d'abord les lois de probabilités admises pour les erreurs ne sont donc pas aussi générales qu'elles pourraient le paraître ; elles ont déjà avec la loi de Gauss (comme d'ailleurs aussi avec la première loi de Laplace) cette analogie d'avoir un écart quadratique moyen fini. Et celui-ci est soumis, entre autre, à la condition de convergence mentionnée plus haut."

Bien cordialement  
P.Lévy.

### Lettre 25

Paris, 38 rue Théophile Gautier Auteuil 90.30

le 26 /10/1930

Mon cher Collègue,

Comment ai-je pu ne pas vous envoyer mon mémoire sur le th. de Fischer et Riesz,<sup>88</sup> où je vous cite largement, à propos des fonctionnelles linéaires ?<sup>89</sup>

Malheureusement, je ne puis plus que vous en communiquer un exemplaire que je vous adresse séparément, en vous priant de me le renvoyer quand vous n'en aurez plus besoin ; (il n'y pas urgence car j'en ai un autre ; mais j'aimerais en conserver deux). J'y joins mon dernier mémoire du même bulletin<sup>90</sup> - pour être sûr que vous l'ayez et 3 autres<sup>91</sup> que vous pouvez garder.

Vous n'y trouverez d'ailleurs pas ce que vous semblez croire. Il s'agit d'une démonstration du th. de F. et R. reposant sur la notion de convergence en mesure, et non d'une définition de la moyenne généralisée (sujet que j'ai abordé seulement dans le Bull de la Soc. Math. 1926 ;<sup>92</sup> j'ai seulement précisé depuis dans Annali di Matematica 1928<sup>93</sup> un point que certains lecteurs avaient mal compris)

Bien cordialement

P.Lévy

Vous remarquerez ma nouvelle adresse.

### Lettre 26

Paris, 16/3/31 - 38 rue Théophile Gautier

Mon cher Collègue,

---

<sup>88</sup>Lévy, Paul (1925) Sur le théorème de M.Fischer et F.Riesz sur la convergence en moyenne. Bull. Soc. Math. Fr. 49:1–14. In Lévy's lapidary formulation, the theorem states that for a series of functions in  $L^2([0, 1])$ , pointwise convergence implies convergence in mean.

<sup>89</sup>Lévy is citing a theorem of Fréchet which shows that a continuous linear form  $U$  on  $L^2$  may be written  $U(\varphi) = \int_0^1 \varphi(x)f(x) dx$  where  $f$  is square integrable.

<sup>90</sup>Probably Lévy, Paul (1929) Sur les lois de probabilités dont dépendent les quotients complets et incomplets d'une fraction continue. Bull. Soc. Math. Fr. 57:178–194. This article is mentioned again, for another reason, in the 1947 Letter 61. Another possibility is Lévy, Paul (1930) Sur la croissance des fonctions entières. Bull. Soc. Math. Fr. 57:127–149.

<sup>91</sup>Lévy published 6 articles in the Bulletin in 1929 and 1930. Exactly which ones are the 'three others' is unclear. The article Lévy, Paul (1929) Sur quelques travaux relatifs à la théorie des erreurs. Bull. Soc. Math. Fr. 53:1–21 is a response to Fréchet's two articles Fréchet, Maurice (1927) Sur la loi de l'écart maximum. Ann. Soc. Pol. Math. 5:93–116 and Fréchet, Maurice (1928) Sur l'hypothèse de l'additivité des erreurs partielles. Bull. Soc. Math. Fr. 52:203–216 (see Letters 15–24).

<sup>92</sup>Continuing Borel's work, Lévy published an article on different types of summation: Lévy, Paul (1926) Sur les conditions d'application et de régularité des procédés de sommation des séries divergentes. Bull. Soc. Math. Fr. 54:1–25.

<sup>93</sup>Lévy, Paul (1928) Fonctions à croissance et itération d'ordre fractionnaire. Ann. Mat. 6:269–298.

Je vais passer à la Sorbonne déposer quelques tirages à part,<sup>94</sup> et je joins un mot comme suite à notre conversation téléphonique.

La probabilité de convergence de  $\sum x_n$ , les  $x_n$  étant indépendants, est toujours 0 ou 1 et je donne dans une Note<sup>95</sup> présentée aujourd'hui les conditions nécessaires et suffisantes pour qu'on soit dans l'un ou l'autre cas.

Je viens de m'apercevoir qu'en outre, dans le cas de convergence, bien qu'il y ait en général semi-convergence, la loi de probabilité dont dépend la somme est indépendante de l'ordre des termes. On en déduit aisément que

$$1 \pm \frac{1}{2} \pm \frac{1}{3} \cdots \pm \frac{1}{n} \pm \cdots \quad \text{en tenant compte de } \pm \frac{1}{2} \pm \frac{1}{4} \pm \cdots \sim 4x$$

$$\text{et } 4(x_1 + \frac{x_2}{2} + \cdots + \frac{x_n}{n} \pm \cdots)$$

( $x_n$  étant une variable choisie au hasard entre  $-\frac{1}{2}$  et  $+\frac{1}{2}$ , avec la densité de probabilité unité) dépendent de la même loi de probabilité. Il y a d'autres cas particuliers amusants à étudier.

La série  $\sum \pm \frac{1}{n}$  a déjà été étudiée par Norbert Wiener.<sup>96</sup> Je vois qu'à ce sujet il cite Steinhaus, je vais rechercher dans le mémoire cité.

Il est entendu que si, après avoir lu ma Note, vous trouvez une priorité à me signaler, je vous en serai reconnaissant.

Bien cordialement

P.Lévy.

## Lettre 27

Paris 38 rue Théophile Gautier 18/3/31

Mon cher Collègue,

Je vous remercie bien vivement. Kolmogoroff m'avait laissé 5 ou 6 mémoires que je n'avais pas encore pu étudier ; j'ai pu au reçu de votre lettre, hier soir, me reporter à la source citée, constater qu'il s'agissait bien de mon théorème, qu'il était attribué à Khintchine et Kolmogoroff, (Moscou 1925), et ce matin citer cette référence en addition à ma Note.<sup>97</sup> Il était temps !

J'avais lundi précisément trouvé le Mémoire de Steinhaus dans *Studia Math.* ;<sup>98</sup> il traite un cas très particulier du problème en question, et cela m'avait encore plus convaincu que la solution générale n'était pas connue. Ainsi ai-je été très surpris en recevant votre lettre.

<sup>94</sup>Probably Lévy, Paul (1931) Sur le gain maximum au cours d'une partie de pile ou face. CRAS 192:258–259. At the same time, Lévy published another article on this topic: Lévy, Paul (1931) Nuove formule relative al giuoco di testa e croce. G. Ist. Ital. Attuari 6:3–36. A shortened version appeared in (1931) J. Ec. Polytch. 3–23.

<sup>95</sup>Lévy, Paul (1931) Quelques théorèmes sur les probabilités dénombrables. CRAS 192:658–659, which Lévy presented the same day at the Academy. As the next letter shows, Lévy was unaware of the results of Khinchin and Kolmogorov, despite their having been published five years before.

<sup>96</sup>In 1923, at their first meeting, Wiener showed Lévy the charms of this series. Lévy declared himself "enchanted." See Sect. 7.

<sup>97</sup>His note to CRAS (session of March 16, 1931), cited in Letter 26.

<sup>98</sup>See Letter 26.

Dans ma Note, qui va donc paraître avec l'addition faite ce matin, le seul résultat important, qui me reste et que je vous signale, est le théorème IV.

J'ai écrit à Steinhaus que le résultat annoncé comme probable à la fin de son dernier Mémoire de Studia est faux. Il faut enlever le 2 sous le radical. Dans le cas particulier, l'hypoténuse n'est pas plus grande que les deux composantes, les grandes valeurs des deux composantes n'étant pas réalisées simultanément. Cela est d'ailleurs bien évident, car si vous divisez le cercle en  $p$  parties ( $p$  arbitrairement grand mais fixe) les grandes valeurs de  $\sum a_n e^{i\varphi_n}$  ont autant de chances d'être réalisées avec un argument qu'avec un autre ; on peut donc supposer leur argument compris entre  $-\frac{\pi}{p}$  et  $+\frac{\pi}{p}$ , et, avec une erreur relative très petite, assimiler cette somme à sa partie réelle.

Je réponds maintenant à vos objections sur mon Calcul des Probabilités.

1° Je crois que vous avez été troublé par le mot variable ;  $\lambda(x)$  est en l'espèce quelque chose d'aussi bien déterminé que  $\sin x$ , et quand je définis la valeur probable par  $\sum \alpha_i \lambda\{x_i\}$ ,  $\alpha_i$  est la probabilité de  $x = x_i$  ; elle ne change pas si on l'ajoute à d'autres termes.

Je crois que c'est de Finetti ou en tout cas un italien qui a fait observer que la valeur probable n'est pas exactement la même chose si l'on étudie la valeur probable de  $\lambda$  en  $f^{99}$  de  $x$  ou en partant de la loi de probabilité de  $\lambda$  ; mais la distinction n'intervient que dans des cas exceptionnels. Mon énoncé général, ne précisant pas les hypothèses faites sur  $\lambda(x)$ , est sans doute critiquable ; mais l'application au cas où  $\lambda$  est une fonction à variation bornée ne l'est pas.

Vous demandez ensuite si "la probabilité de  $E$ , pour  $x$  donné, étant  $\alpha$ , il en est de même si l'on ne sait rien sur  $x$ ."

Evidemment oui ; par application des axiomes fondamentaux, et je vois bien ce qui vous arrête, à moins que ce ne soit le souvenir de difficultés rencontrées dans des cas analogues, mais non identiques ; je puis vous en citer un exemple que j'ai rencontré récemment.

Soit  $x_n$  le gain après  $n$  coups de pile ou face,  $y_n$  le plus grand des nombres  $x_1, x_2, \dots, x_n$ . Il s'agissait d'avoir la loi de probabilités de  $y_n$ , connaissant  $x_n$ . Si l'on sait que  $y_n = x^\nu$ ,  $\nu$  étant connu, le problème était facile ; mais était-il correct d'affirmer que  $\mathcal{P}\{y_n > N\}$  est inférieur par exemple à la plus grande des probabilités calculées en faisant successivement  $\nu = 1, 2, \dots, n$ . Evidemment non, parce que pour certaines parties le maximum  $y_n$  est atteint pour plusieurs valeurs de  $\nu$ , et dans le compte des cas possibles ces cas sont comptés plusieurs fois. Toute difficulté disparaît si je précise que  $\nu$  est le plus petit entier pour lequel  $y_n = x^\nu$  ; alors chaque cas est bien compté une fois et une seule.

Je suis persuadé que vous arriverez à cette conclusion qu'il n'y pas de difficulté, moyennant des hypothèses sur  $\lambda(x)$ , que j'aurais dû préciser, mais qui sont bien vérifiées dans l'application que j'avais en vue.

Quant à l'exemple de la sphère, c'est un exemple du type classique où la probabilité n'est pas bien définie. Mais dans la composition des probabilités indépendantes, si  $x$  et  $y$  ont des lois déterminées, il en est de même dans le plan

---

<sup>99</sup>Sic.

dont dépend le point  $x, y$ , et par suite de celle dont dépend  $x + y$  ; la déterminer c'est de l'analyse pure.

2° Il faut être timide avec le transfini, sans doute. Toutefois en reprenant le raisonnement de la page 330, il me semble correct. Il faudrait que je voie le Théorème de Vitali dont vous me parlez. Si vous voulez poursuivre cette discussion (pour ma part je serai content de tirer la chose au clair) voulez-vous m'indiquer la référence.<sup>100</sup>

Pour le 1°, si vous n'êtes pas convaincu, je crois qu'il vaudrait mieux en parler de vive-voix.

- Pour vous prouver que je ne suis pas infallible, et à toutes fins utiles, je vous signale quelques errata

Calcul des Probabilités - p.202. Tout ce passage est à revoir ; j'ai perdu de vue en l'écrivant que  $F(x)$ , fonction monotone croissante de 0 à 1, a une limite croissant de  $x_0 \geq 0$  à  $x_1 \leq 1$  mais qu'on n'a pas le droit d'affirmer que  $x_0 = 0$  et  $x_1 = 1$ .

Calcul des Probabilités- p.161, lignes 1 à 5. C'est faux, cela devient correct en remplaçant l'exposant 2 dans  $I_n$  par l'exposant  $n$  (résultat qui m'a été communiqué par Chokhate<sup>101</sup>)

Sulla legge forte dei grandi numeri, p.7 du tirage à part, remarque 2°. Le résultat énoncé est exact mais le raisonnement défectueux ;<sup>102</sup> même remarque pour le bas de la p.18 du même mémoire.

Bien cordialement

P.Lévy.

## Lettre 28

Mane Fetan. La Trinité sur mer, le 19/9/31

Morbihan

Mon cher Collègue,

Etes-vous à Paris, et si oui pouvez-vous - si cela ne vous dérange pas, me rendre un petit service : compléter les indications bibliographiques de 2 mémoires de Lindeberg, que vous avez sans doute en tirage à part

---

<sup>100</sup>The last section of Lévy (1925) (which includes page 330) is entitled Note sur les lois de probabilités dans les espaces abstraits and is a reproduction of a text published in Revue de Métaphysique et de Morale. Lévy suggested extending the Lebesgue measure to all the subsets of  $[0, 1]$ . He noted that an application of Zermelo's theorem allows the assignment of an arbitrary value to the non-measurable sets but that this is of no practical interest. In the second (1957) edition of Lévy (1937), he mentions on page 370 that the application of Zermelo's theorem was in fact not justified in this case, as had been pointed out to him by Steinhaus. Fréchet no doubt had his reservations about this abstruse remark by Lévy. Perhaps Fréchet had mentioned Vitali's covering theorem to Lévy as an avenue of attacking the question from another angle.

<sup>101</sup>The identity of 'Chokhate' is unclear: the word is badly written.

<sup>102</sup>This remark and the following seem to indicate that the offprints in question at the start of Letter 26 are from Lévy, Paul (1931) Sulla legge forte dei grandi numeri. G. Ist. Ital. Attuari 6:3-23.

Annales Ac. Scientiarum Fennicae - A.16 (1920) p... à ... ?

Math.Zeitschrift - 15 (1922) p ?<sup>103</sup>

et d'envoyer ces indications à H.Steinhaus ul Kadecka 14 Lwów. Il les attend pour le bon à tirer d'un travail de moi qu'il publie dans *Studia*,<sup>104</sup> et semble pressé. Si vous n'êtes pas à Paris, il sera inutile de vous occuper de cette lettre quand elle vous rejoindra, car j'y rentre le 26.

J'en profite pour vous parler d'autre chose. Je crois vous avoir parlé d'un autre petit travail que je publie, qui comprend des observations et rectifications à mon calcul des probabilités. A la demande de M. Dickstein (de Varsovie) qui le publie dans son journal au nom barbare (math et phys),<sup>105</sup> j'ai changé la rédaction primitive que je vous avais, je crois, remise. J'ai supprimé une note sur le battage des cartes, donnant en quelques lignes la démonstration de la formule de mon livre sur laquelle vous m'aviez un jour demandé des explications<sup>106</sup> - et je ne sais ce qui ce jour m'avait empêché de retrouver la démonstration très simple de cette formule dont j'ai mis dans mon livre "un calcul simple montre que...". M.Dickstein m'avait demandé de donner les explications nécessaires pour rendre mon article accessible à qui n'avait pas mon livre sous la main, et il m'a semblé dès lors que cette petite question ne méritait plus la place qu'il faudrait lui attribuer. L'intérêt de la copie dactylographique qui me reste n'étant plus que de constituer la réponse à ces questions posées par vous, je la tiens à votre disposition - à moins que je ne vous en aie déjà donné une, ce qui est possible.

Bien cordialement

P.Lévy.

## Lettre 29

Paris le 29/1/36

Mon cher Collègue,

Je vous confirme que dans le théorème de Kolmogoroff, p. 60 de ses *Grünbegriffe*, la condition d'indépendance des  $x_n$  mentionnée au bas de la page est absolument inutile.<sup>107</sup>

<sup>103</sup>Lindeberg's two articles, Lindeberg, J.W. (1920) Die Exponentialgesetz in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fenn.* 16:1 and Lindeberg, J.W. (1922) Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Z.* 15:211–225, gave a new proof of the central limit theorem with a condition on moments of order 3.

<sup>104</sup>Lévy, Paul (1931) Sur les séries dont les termes sont des variables éventuelles indépendantes. *Studia Mat.* 3:119–155.

<sup>105</sup>This is *Prace Matematyczno-Fizyczne*. The Polish mathematician Samuel Dickstein (1851–1939), University of Warsaw, was the editor. See Mostowski, A. (1949) *La vie et l'œuvre de Samuel Dickstein*. *Pr. Mat.-Fiz.* 47:VII–XII.

<sup>106</sup>See Letters 18 and 19 of 1928.

<sup>107</sup>This is the 'strong' version of the 0–1 law: if an event  $A \in \sigma(x_1, \dots, x_n, \dots)$  ( $f(x_1, x_2, \dots) = 0$  for Kolmogorov) is independent of  $(x_1, x_2, \dots, x_n)$  for all  $n$ ,  $P(A) = 0$  or 1, even without requiring the mutual independence of the  $x_i$ . That said, Lévy made an error, perhaps due to a poor translation of the German. In Kolmogorov's remark cited, the considered independence means that the value of the function  $f$  is not changed when  $(x_1, x_2, \dots, x_n)$  is modified: it is not stochastic independence.

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<sup>108</sup> 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 ?<sup>109</sup> Mais il s'agit de convergence en probabilité, et non de convergence presque sûre, de  $\frac{1}{n}(x_1 + x_2 + \dots + x_n)$ . Pour la convergence presque sûre, il ne fait que rappeler p.59 le résultat de sa note de 1930.<sup>110</sup>

Bien cordialement à vous.

P.Lévy.

**Lettre 30**

Paris, le 21/12/36

Mon cher Collègue,

J'ai écrit à Feldheim, à la suite de notre conversation d'hier, que vous accepteriez de lui remettre la mise en page de mon livre.<sup>111</sup> Si cela ne vous gênait pas de le faire assez tôt pour qu'il puisse transcrire, d'après l'autre exemplaire corrigé par moi et destiné à Doebelin, les corrections que j'ai faites, et le remporter chez lui, je suis sûr que vous lui rendriez service.

The hypothesis of mutual (stochastic) independence of the  $x_i$  introduced by Kolmogorov is thus necessary.

<sup>108</sup>Lévy is referring to his own demonstration of the 0-1 law, given at a meeting of the Société Mathématique de France on May 23, 1934 and published as Lévy, Paul (1935) Propriétés asymptotiques des sommes de variables aléatoires enchaînées. Bull. Soc. Math. Fr. 59:84-96. It is interesting that Lévy does not cite Kolmogorov's memoir in this article (the proofs of which were corrected by Lévy in 1935, as the note added proves). Lévy must have discovered Kolmogorov's work just before writing the present letter, which is evidence that he and Fréchet had only recently discovered the memoir. Abstract measure theory was relatively new in 1936 (Polish school only formulated the contemporary practice of  $\sigma$ -algebras and reasoning by generation at the beginning of the 1920s). This explains Fréchet's apparent uncertainty, and his desire for Lévy's opinion of Kolmogorov's proof.

<sup>109</sup>Kolmogorov's result is a necessary and sufficient condition for a generalized law of large numbers, i.e. convergence to 0 in probability of  $S_n - E(S_n)$ , with  $S_n = \frac{X_1 + \dots + X_n}{n}$ , where the  $X_i$  are independent but do not necessarily have the same law. This condition is expressed in terms of the median.

<sup>110</sup>If  $\sum_n \frac{\text{Var}(X_n)}{n^2} < +\infty$ , then  $\frac{X_1 + \dots + X_n}{n} - E\left(\frac{X_1 + \dots + X_n}{n}\right) \rightarrow 0$ , a.s. See Kolmogorov, Andrei (1930) Sur la loi forte des grands nombres. CRAS 191:910-912.

<sup>111</sup>Lévy, Paul (1937) Théorie de l'Addition des variables aléatoires. Gauthier-Villars, Paris.



J'espère pouvoir donner le volume définitif dans 5 ou 6 semaines.

Je reviens d'autre part sur notre conversation d'hier. Il est certain que l'on découvre parfois des théorèmes importants et faciles qui avaient échappé aux chercheurs antérieurs et que ce n'est pas condamner un théorème que de dire qu'il est facile. Mais lorsqu'il s'agit d'un cas particulier d'un problème général résolu depuis longtemps, sauf dans quelques cas particuliers difficiles qui ont fait l'objet de recherches récentes, je crois très sincèrement que ce serait assez ridicule d'aller chercher un cas particulier du théorème classique pour le monter en épingle. S'il se trouve qu'il prenne de l'importance en vue d'une application, on le mentionnera à ce moment ; mais en faire un théorème important en soi me ferait le même effet que d'attacher de l'importance à une formule telle que  $(3,141592)^2 = \dots$  etc., sous prétexte qu'elle permet le calcul de  $\pi^2$  ; elle ne devient importante que par l'application que l'on en fait.

C'est le cas de la valeur médiane. Le caractère de mon enseignement à l'Ecole Polytechnique m'oblige à me borner au point de vue de Bernoulli.<sup>112</sup> Mais j'explique toujours à mes élèves que la détermination expérimentale d'une suite de  $n$  valeurs indépendantes d'une variable aléatoire  $X$  donne une courbe de fréquence dessinant la loi de probabilité, avec une précision d'autant plus grande que  $n$  est plus grand. La valeur médiane, ou l'écart probable  $\frac{x''-x'}{2}$ ,  $x'$  et  $x''$  étant les valeurs pour lesquelles la fonction de répartition a les valeurs  $\frac{1}{4}$  et  $\frac{3}{4}$ , sont obtenues avec une précision croissante. Je l'explique toujours à mon cours p.281 de mon cours d'analyse de l'X. Mais lorsqu'on prend la moyenne, on introduit une difficulté supplémentaire (qui d'ailleurs n'existe que dans des cas particuliers) parce que des valeurs rares mais très grandes peuvent jouer un grand rôle dans le calcul de la moyenne. Si j'ai souvent, à la Société mathématique, ou dans des mémoires récents, insisté sur ce rôle, c'est parce qu'il me semblait qu'il pouvait n'être pas compris (quoique, sauf erreur, déjà signalé par Bienaymé<sup>113</sup>).

Tout cela s'applique au point de vue de la loi forte des grands nombres. Le rôle de la valeur médiane est élucidé depuis longtemps ; c'est un cas particulier banal des résultats de Borel et Cantelli. Mais si Kolmogoroff et moi nous sommes occupés de la moyenne, c'est parce que c'était il y a peu d'années le problème difficile non encore résolu. Vous avez pu constater que l'impression de Darmon a exactement coïncidé avec la mienne- et si j'avais la chance de pouvoir faire sur le calcul des probabilités un cours moins élémentaire que celui que je fais à l'X (3 leçons en tout sur ce sujet), je vous renverrais à n'importe lequel de mes élèves.

Bien cordialement à vous

P.Lévy.

P.S. En ce qui concerne mes placards, il est bien entendu que mes observations manuscrites sont pour vous personnellement. Si vous vouliez en faire état, je vous serais obligé de me demander si je maintiens mon opinion sur le point dont il s'agit.

<sup>112</sup>In his course at the Ecole Polytechnique, Lévy only considered this elementary case.

<sup>113</sup>For example in his 1852 article in the Journal de Mathématiques Pures et Appliquées, in which he improved the method of least squares.

Ainsi, en reprenant mon Laplace l'autre jour, je suis tombé sur un passage qui ne doit pas être celui que vous m'avez signalé il y a 1 ou 2 ans, mais qui donne raison à ce que vous m'avez dit qu'à propos de la fonction caractéristique il faudrait mentionner Laplace. Je vous donne donc raison sur ce point et vais rectifier mon §14 en conséquence.<sup>114</sup> Il peut y en avoir d'autres qui sont sortis de mon esprit mais sur lesquels je ne maintiendrais pas une opinion quelquefois exprimée un peu vite.

### Lettre 31

Paris - le 23/12/36

Mon cher Collègue,

Je réponds aux points principaux de votre lettre.

1° Au sujet de la note de Ville,<sup>115</sup> son théorème est contenu dans le th. de Glivenko-Cantelli,<sup>116</sup> exposé aux pages 260–261 de votre livre, sur lequel vous attirez mon attention. Vous me demandez une démonstration précise : la voici, en adoptant vos notations.

Soit  $(m', m'')$  l'intervalle médian pour  $X$ , c'est à dire que  $m' \leq m''$ ,  $F(m' + 0) < \frac{1}{2}$ ,  $F(m'' - 0) > \frac{1}{2}$  et quel que soit  $\varepsilon > 0$ ,  $F(m' - \varepsilon) \leq \frac{1}{2}$ ,  $F(m'' + \varepsilon) \geq \frac{1}{2}$ . Alors sauf dans des cas de probabilité nulle,  $F_n(m' - \varepsilon) \rightarrow F(m' - \varepsilon) < \frac{1}{2}$  et  $F_n(m'' + \varepsilon) \rightarrow F(m'' + \varepsilon) > \frac{1}{2}$ ; donc il existe un  $N$  tel que pour  $n > N$  on ait  $F_n(m' - \varepsilon) < \frac{1}{2} < F_n(m'' + \varepsilon)$ . Donc une médiane expérimentale, c'est à dire un nombre  $\mu_n$  tel que  $F_n(\mu_n) = \frac{1}{2}$ , est compris entre  $m' - \varepsilon$  et  $m'' + \varepsilon$ , cela quelque petit que soit  $\varepsilon$ , c.q.f.d.

On peut préciser en distinguant plusieurs cas. Si  $m' = m'' = m$ ,  $\mu_n$  tend (toujours presque sûrement) vers  $m$ . Si de plus  $F(m - 0) < \frac{1}{2} < F(m + 0)$ ,  $F_n(m - 0)$  et  $F_n(m + 0)$  tendent vers  $F(m - 0)$  et  $F(m + 0)$ ,  $\mu_n$  est exactement égal à  $m$ , pour  $n$  assez grand. Si au contraire  $m' < m''$ , en prenant pour fixer les idées  $n$  impair de manière à avoir non un intervalle médian mais une médiane unique,  $\mu_n$  est compris entre  $m' - \varepsilon$  et  $m'' + \varepsilon$ ; on ne peut rien dire de plus. Les fréquences des deux éventualités  $X \leq m'$  et  $X \geq m''$  sont en effet comparables à la fréquence d'un des cas au jeu de pile ou face, qui est presque sûrement une infinité de fois  $> \frac{1}{2}$  et une infinité de fois  $< \frac{1}{2}$ ; cela revient à dire que  $\mu_n$  est tantôt  $< m'$  tantôt  $> m''$ .

Je suis tout à fait d'accord sur l'intérêt d'une synthèse mettant en valeur différents résultats, suivant l'importance qu'ils ont à tel ou tel point de vue; intérêt pédagogique, éventuellement. Mais ce qui pour moi était imprévisible, et me fait regretter de ne pas avoir développé tout cela depuis longtemps, c'est que vous y

<sup>114</sup>The correction was made in Lévy (1937, §14).

<sup>115</sup>CRAS (December 14, 1936). J.-A. Ville's note, presented by Borel, shows the convergence a.s. of the medians of a series of observations of independent random variables with the same law to the theoretical median.

<sup>116</sup>The real power of the Glivenko-Cantelli theorem is that it establishes the *uniform* convergence of the distribution functions of the observations, although here Lévy only makes use of the simple convergence resulting from the law of large numbers. Lévy remarks on this himself in the next letter.

verriez autre chose qu'une mise au point utile du point de vue pédagogique, pour les débutants.

Je vous signale que la référence Glivenko, 2, manque dans votre bibliographie, à moins qu'on ait mis 3 et 4 au lieu de 2 et 3. D'ailleurs, même au point de vue des regrets que je puis avoir, la date n'a pas pour moi grand intérêt, car je ne puis vous dire exactement depuis quand je connaissais le th. en question, peut-être en 1930, date à laquelle mon attention s'est trouvée très spécialement attirée sur la loi forte des grands nombres.

2° Au sujet de Laplace, je viens de lire ou relire vos p. 105 à 107. Cela coïncide bien avec mes souvenirs d'une lettre que vous m'avez écrite. Je ne suis qu'en partie d'accord.

La fonction génératrice de Laplace, pour moi, diffère surtout de la fonction caractéristique par le fait que vous mentionnez p. 106 : Laplace ne considérait que des valeurs entières, le problème d'inversion est alors bien élémentaire, et les formules que vous citez p.107 ne suffisent pas pour justifier votre phrase : "cette propriété précieuse était aussi connue de Laplace". Pour le reste je suis d'accord.

Le passage que j'ai découvert par hasard (si l'on peut appeler hasard le fait que j'ai fait des recherches rapides mais méthodiques après avoir su que Darmais est d'accord avec vous) me paraît plus probant. C'est le Chap.IV du livre II, N° 20.<sup>117</sup> Il n'y a pas d'effort pour atteindre les lois les plus générales, mais je vois l'esquisse de la méthode que j'ai contribué à mettre au point.<sup>118</sup> Cela m'a décidé à envoyer une modification à G.V. (j'espère qu'il sera temps) et je regrette de ne pas m'être décidé comme vous à dire loi de Laplace au lieu de Loi de Gauss.

Je ne serai pas nationaliste ; mais quand dans la table des auteurs cités dans le livre de Tornier,<sup>119</sup> je ne vois pas un savant français postérieur à Poisson (malgré le rôle qu'y jouent les intégrales de Lebesgue et de Jordan) je me dis que nous avons le devoir de défendre nos grands hommes.

Bien cordialement à vous

P.Lévy

Je viens de relire votre lettre pour m'assurer que je n'oublie rien d'important.

Je suis bien d'accord pour le § 301 de mon cours. J'ai simplement voulu dire qu'au point de vue de la loi forte des grands nombres, le th. de Ville est très analogue à ce que j'avais dit du point de vue de Bernoulli, et exprimé le regret que le caractère de mon enseignement à Polytechnique ne m'ait pas permis d'aborder les points de vue de Borel-Cantelli.

Nous sommes d'accord sur le fond. Les différences de point de vue sont inévitables et font l'intérêt des échanges d'idées. J'ai évidemment une psychologie très

<sup>117</sup>Laplace, P.-S. de (1820) *Théorie Analytique des Probabilités*. 3ème Edition. pp. 318–324. Laplace introduced an early version of the characteristic function to determine the law of a sum of random variables and began an approach to a kind of central-limit theorem.

<sup>118</sup>Lévy made the characteristic function a central tool in his study of sums of independent random variables. He appears not to have been ready to abandon its paternity, even to Laplace !

<sup>119</sup>The German mathematician Erhard Tornier published *Wahrscheinlichkeitsrechnung und allgemeine Integrationstheorie (Probability Theory and Generalized Integration Theory)* in 1936 (Springer).

différente de la votre (je pense que le “mathématicien moyen” est entre nous). Je me dis que si vous pouviez voir se dérouler le film de ma pensée, vous accrocheriez au passage des idées que je ne songe pas à retenir, parce que je me dis (et l’expérience prouve que c’est souvent à tort) “il n’est pas possible qu’on n’ait pas déjà eu cette idée”. Or il arrive que - et cela est indépendant de l’intérêt propre des idées ainsi écartées- ce soit en les prenant pour point de départ et en étudiant systématiquement les conséquences - que l’on trouve des extensions auxquelles on n’a pas songé d’abord. La conclusion de ces [*La suite verticalement en marge* remarques est que nous nous complétons assez bien et que, comme vous le dites, les échanges de point de vue entre nous ne sont pas inutiles. Mais j’aurais beau vous voir tous les jours, je n’arriverais ni à vous communiquer toutes mes idées ni à m’assimiler toutes les vôtres. Vous ne pouvez pas vous faire une idée de l’effort qu’il me faut pour suivre la pensée des autres,<sup>120</sup> et j’ai presque toujours intérêt à me contenter d’un coup d’œil sur les résultats suivi d’une reconstitution de la démonstration. Et sauf lorsqu’il s’agit d’une idée nouvelle et simple (comme pour le th. de Cramer) s’il me faut après l’échec de mes efforts de reconstitution reprendre la démonstration la plume à la main, j’ai l’impression que l’essentiel, “l’idée” me manque. Sans doute, “logiquement”, avez vous raison. Mais psychologiquement, à mon point de vue personnel, une phrase même vague qui indique l’idée est préférable à un raisonnement détaillé.<sup>121</sup> Cette “confession” vous explique ce qui parfois a pu vous surprendre chez moi.]

### Lettre 32

Paris - le 24/12/36

Mon cher Collègue,

En réfléchissant à ma lettre d’hier, je précise encore un point. J’avais sous les yeux en vous écrivant la p.261 de votre livre. Mais à cet endroit, ce qui vous a coûté quelque effort, c’est d’établir la convergence uniforme. Ne considérant que deux valeurs  $x' - \varepsilon$  et  $x'' + \varepsilon$  ( $\varepsilon$  arbitrairement petit mais fixe) je n’en ai pas besoin. Il est donc bien entendu que mon raisonnement de 3 ou 4 lignes ne s’appuie que sur le théorème de Borel-Cantelli, ou même de Borel tout seul, relatif au cas de Bernoulli.

D’autre part, à la fin de ma lettre, je ne sais si je me suis bien expliqué. Ce qui domine chez moi, c’est la sensation qu’il me faut un effort moins grand et en tous cas beaucoup moins de temps pour préciser ces théorèmes au moment où j’en ai besoin que pour étudier la littérature et savoir lesquels ont déjà été énoncés. Je parle aussi bien du théorème de Polya<sup>122</sup> (que sans connaître sa priorité j’ai retrouvé après lui) que de différents théorèmes de Glivenko, Cantelli, Steinhaus et autres, que je

<sup>120</sup>Lévy thought of himself as a solitary worker. He mentioned this in his autobiographical book (Lévy 1970), in which he regretted his occasional inability to identify results worthy of publication.

<sup>121</sup>Lévy’s writing style fits this description very well.

<sup>122</sup>Lévy must be alluding to Polya’s article: Polya, George (1935) Zwei Aufgaben aus der Wahrscheinlichkeitsrechnung. Naturforschung Gesellschaft Zürich 80:123–130, in which Polya studied the mean distance separating two neighboring points in a configuration of  $N$  points chosen randomly in a sphere in  $q$  dimensions.

connaissais avant eux et ai omis d'énoncer parce qu'il s'agissait de remarques que je ne considérais que comme des points à préciser pour les débutants qui risqueraient de ne pas les voir tout seuls.

Dans le cas du th. de Jessen Steinhaus,<sup>123</sup> c'est certainement une erreur que j'ai commise, de ne pas souligner une idée nouvelle et féconde que j'avais seulement notée en passant. Dans le cas dont il s'agit, je n'arrive à trouver qu'un très petit nombre d'idées : celle de convergence uniforme, connue depuis longtemps ; puis la loi forte des grands nombres de Borel-Cantelli. Ces deux points connus, tous les théorèmes de Polya Glivenko Cantelli Ville ne me paraissent pas dépasser ce que Darmon propose à ses élèves comme sujet de composition de licence (bien entendu, je ne parle que d'un th. précis de Polya, et non de l'ensemble de son œuvre. C'est pour moi un homme de tout premier ordre).

Quant à "l'intérêt intrinsèque" il ne doit pas être jugé en fonction de vous ou de moi, mais compte tenu de ce que le mathématicien moyen n'est pas comme moi, il ne trouvera pas si on ne lui a pas appris le théorème dont il se trouve avoir besoin. Je vous parle donc uniquement de mes "réactions psychologiques" sans vouloir dire que j'aie raison ou tort. D'ailleurs, lorsqu'il s'agit d'une synthèse comme celle que constitue votre livre, même au point de vue de ces réactions, il ne me viendrait jamais à l'idée de m'étonner de voir une remarque facile, mais qui tient sa place dans l'ensemble, qui y est indispensable. Je sais - surtout maintenant que l'expérience me montre quelle erreur je suis naturellement tenté de commettre, capable d'en faire aussi - et le nouveau travail que j'entreprends vous le montrera. Mais je ne le fais que quand c'est nécessaire pour un exposé d'ensemble.

J'ai été sollicité de collaborer à une revue philosophique.<sup>124</sup> Peut-être développerai-je un jour ces considérations psychologiques.<sup>125</sup>

Bien cordialement à vous.

P.Lévy.

### Lettre 33

le 26/12/36

Mon cher Collègue,

Je vous communique un résultat simple, et pour moi inattendu.<sup>126</sup>

<sup>123</sup>(1936) CRAS 202:1961–1963, in which Steinhaus proposes an explicit construction of a series of independent uniformly distributed random variables chosen by a recursive use of the Peano curve. In 1930 the Danish mathematician Borge Jessen proposed a different construction.

<sup>124</sup>It is unclear to which journal Lévy was referring. One possibility is the *Revue de Métaphysique et de Morale*, in which he had already published (see note 100); another is *Revue de Synthèse*, founded in 1900. In June 1935 Lévy attended a colloquium on statistics organized by the editorial board of this latter journal. Among those editors, as of 1936, was Jean Ullmo, a former student of Lévy at the Polytechnique. Ullmo played an important role in establishing the curriculum and research direction of the Polytechnique after the war. If a collaborative project did in fact occur, it does not seem to have left any trace.

<sup>125</sup>Lévy's autobiographical book of 1970 returns to this point.

<sup>126</sup>This elementary remark was the origin of numerous developments during the next year.

Problème : Est-il possible, en désignant par  $\mathcal{L} = \mathcal{L}_1 \mathcal{L}_2$  la composition (par addition) de deux lois de probabilités, et ne considérant pas comme distinctes deux lois déduites l'une de l'autre par l'addition d'une constante, que l'on ait pour une même loi deux représentations différentes  $\mathcal{L}_1 \mathcal{L}_2 = \mathcal{L}'_1 \mathcal{L}'_2$ , les quatre facteurs ainsi mis en évidence étant premiers.

Réponse : oui. En effet

(1) (O ou 1) . (O ou 2 ou 4) équivaut à (O ou 1 ou 2). (O ou 3) chaque parenthèse désignant une variable aléatoire pour laquelle les différentes valeurs indiquées sont également probables.

Je vous cite surtout ce résultat car il vous fera bien comprendre mon point de vue. La formule (1) isolée n'aurait eu pour moi aucun intérêt, elle en prend à condition qu'on la mette à sa place dans l'arithmétique des lois de probabilité ; elle résout un problème qu'on ne pouvait pas ne pas se poser.

Plus généralement il peut arriver qu'une loi puisse être obtenue de  $n!$  manières différentes comme produit de  $n$  facteurs premiers.

Ces remarques se rattachent aussi à un problème qui, je crois, n'a pas été assez étudié : la numération généralisant la numération décimale. On se donne une suite de nombres premiers (restriction non essentielle)  $p_1, p_2, \dots, p_n$  et on représente chaque  $x$  entre 0 et 1 par

$$x = \sum_1^{\infty} \frac{a_n}{p_1 p_2 \cdots p_n} \quad (a_n = 0, 1 \dots \text{ou } p_n)$$

Je crois que je vous ai dit que je ferai le 12 Janvier<sup>127</sup> (sauf changement d'horaire, ils sont fréquents au séminaire H) une conférence sur l'arithmétique des lois de probabilité. Il y a de nombreux et nouveaux résultats, très curieux, de Khintchine et de son élève Raikoff.<sup>128</sup> Ainsi le théorème de Cramer sur la loi de Gauss<sup>129</sup> est maintenant étendu à la loi de Poisson.

Bien cordialement à vous

P.Lévy

Je vous parle toujours très franchement ; j'espère ne pas vous fâcher quand je n'attache pas à tel ou tel point la même importance que vous. Mais je serai toujours

<sup>127</sup>The lecture notes of this conference, as well as the proceedings of the Hadamard seminar are in Lévy's article published in the Hadamard Jubilee volume: (1938) *J. Math. Pures Appl.* 17:17–39. See the next letter.

<sup>128</sup>Dmitri Abramovitch Rajkov, born in 1905, became a student of Khinchin in 1935. The two men worked on the arithmetic of probability distributions and published a dozen or so articles between 1936 and 1938. In the 1940s, Rajkov turned to algebraic problems and, with Gelfand, was one of the founders of the theory of normed algebras.

<sup>129</sup>Cramer's theorem ((1936) *Math. Z.* 41(3):405–414) demonstrates that if a random Gaussian variable is the sum of two independent variables, each of them is Gaussian. Rajkov obtained the same result for the Poisson distribution ((1937) *Доклады Академии Наук* 14:9–11). In a note to the CRAS of January 11, 1937, Lévy found a more general theorem on the decomposition of characteristic functions that are entire.

d'accord sur l'intérêt d'une synthèse - comme celle que constitue votre œuvre sur le calcul fonctionnel abstrait- et sur le fait qu'un détail (même si pour moi il est banal) peut-être indispensable à cette synthèse.

### Lettre 34

Paris- le 8/1/37

Mon cher Collègue,

Je viens d'envoyer à Villat,<sup>130</sup> pour son Journal,<sup>131</sup> le texte de la conférence que je ferai Mardi sur l'arithmétique des lois de Probabilité. Je vous écris, non pour vous communiquer l'essentiel, que je m'efforcerai de faire ressortir (je vous réécrirai ultérieurement si vous ne pouvez pas venir Mardi), mais des détails dont je ne pense pas parler Mardi, qui se rattachent à ce genre de notions auxquelles vous attachez plus d'importance que moi-même.

- Soit  $\mathcal{L}$  une loi de probabilité,  $X$  la variable dépendant de cette loi. J'appelle toujours  $\omega(\alpha)$  la fonction de dispersion, c'est-à-dire la longueur minimum (minimum effectivement atteint) des intervalles  $(a, b)$  tels que  $\Pr\{a \leq X \leq b\} \geq \alpha$ . Suivant une idée que vous et moi avons souvent utilisée, j'introduis une fonction auxiliaire  $\lambda(\omega)$  définie pour  $\omega \geq 0$ , croissante et bornée ;<sup>132</sup> par exemple  $\frac{\omega}{\sqrt{1+\omega^2}}$ , ou  $\frac{\omega}{1+\omega}$ , ou  $\arctg \omega$ , ou  $\text{th } \omega \dots$  etc.

J'appelle dispersion moyenne et désigne par  $\delta = \delta(\mathcal{L})$  le nombre  $\delta$  défini par

$$\lambda(\delta) = \int_0^1 \lambda[\omega(\alpha)] d\alpha.$$

Les propriétés sont les suivantes :  $\delta = 0$  caractérise les lois unités ( $\mathcal{U}$ ), c'est-à-dire celles pour lesquelles  $X$  n'a qu'une valeur possible. Pour les lois dépendant d'un paramètre  $t$ ,  $\delta \rightarrow 0$  caractérise le cas où il existe une fonction non aléatoire  $f(t)$  telle que  $X_t - f(t)$  tende en probabilité vers zéro.  $\delta$  donne donc une définition commode de l'écart de la loi  $\mathcal{L}$  avec le type de loi ( $\mathcal{U}$ ).

La signification  $\delta \rightarrow \infty$  est évidente.<sup>133</sup>

Si  $\mathcal{L} = \mathcal{L}'\mathcal{L}''$ ,  $\delta(\mathcal{L}) \geq \delta(\mathcal{L}')$  et l'égalité n'est possible que si  $\mathcal{L}''$  est la loi unité.

Je dis qu'un produit infini  $\prod \mathcal{L}_n$  est quasi-convergent s'il existe une suite de constantes  $a_n$  telles que la série à termes aléatoires indépendants  $(X_n - a_n)$ , où  $X_n$  dépend de la loi  $\mathcal{L}_n$ , soit presque sûrement convergente.<sup>134</sup>

<sup>130</sup>In 1937, H. Villat was editor of the Journal de Mathématiques Pures et Appliquées.

<sup>131</sup>Lévy, Paul (1938) L'arithmétique des lois de probabilité. J. Math. Pures Appl. 17:17-39.

<sup>132</sup>'And continuous.' The published version includes this condition.

<sup>133</sup>The probability measure concentrates at infinity.

<sup>134</sup>According to the practice at the time, Lévy conflated the convergence of the laws and the convergence a.s. of the variables that have these laws. In the 1950s Skorokhod formalized these matters by changes of probability spaces.

La condition nécessaire et suffisante pour la quasi-convergence de  $\prod \mathcal{L}_n$  est que  $\delta(\mathcal{L}_1 \cdot \mathcal{L}_2 \cdots \mathcal{L}_n)$  (qui en tout cas croît avec  $n$ ) reste borné.<sup>135</sup>

Si chaque  $\mathcal{L}_n$  est définie par sa fonction caractéristique  $\varphi_n(z)$ , la quasi-convergence du produit  $\prod \mathcal{L}_n$  ne dépend que du module  $|\varphi_n(z)|$ . La condition nécessaire et suffisante pour cette quasi-convergence est qu'il existe un nombre  $Z > 0$  tel que pour  $|z| < Z$ ,  $\prod_1^n |\varphi_n(z)|$  tende uniformément vers une limite  $|\varphi(z)|$ . Mais, alors qu'un produit quasi-convergent définit une loi limite à un facteur unité près, il ne suffit pas de connaître  $|\varphi(z)|$  pour connaître la fonction caractéristique  $\varphi(z)$  de cette loi. En remplaçant notamment certains facteurs  $\varphi_n(z)$  par  $\varphi_n(-z)$  on change  $\varphi(z)$  et on ne change pas  $|\varphi(z)|$ .

- Je mentionne encore deux autres notions utiles.<sup>136</sup>

1° soit  $\mathcal{L} = \prod_1^n \mathcal{L}_v$  ( $n$  fini ou infini) une décomposition de  $\mathcal{L}$  en facteurs ;  $\bar{\delta} = \text{Max } \delta(\mathcal{L}_v)$ ;  $\eta(\mathcal{L}) =$  borne inférieure de  $\bar{\delta}$  quand la décomposition varie.

$\eta(\mathcal{L}) = 0$  caractérise les lois indéfiniment divisibles;

$\eta(L) = \delta(L)$  caractérise les lois indécomposables.

Dans tous les autres cas  $0 < \eta(\mathcal{L}) < \delta(\mathcal{L})$ .

2°  $\zeta(\mathcal{L}) =$  borne supérieure de  $\delta(\Lambda)$ ,  $\Lambda$  étant un diviseur indécomposable de  $\mathcal{L}$ .

On démontre facilement que  $\zeta(\mathcal{L}) \geq \eta(\mathcal{L})$ , donc que, si  $\eta > 0$ ,  $\mathcal{L}$  a au moins un diviseur indécomposable. On en déduit le théorème suivant, qui précise un peu le théorème de Khintchine.

Toute loi  $\mathcal{L}$  est de la forme  $\mathcal{L}'\mathcal{L}''$ ,  $\mathcal{L}'$  étant un produit de lois indécomposables,  $\mathcal{L}''$  étant une loi sans diviseur indécomposable.

Dans l'énoncé de Khintchine,  $\mathcal{L}''$  est une loi indéfiniment divisible ce qui est moins restrictif. La décomposition de Khintchine n'est pas toujours unique ; il se peut que la mienne le soit toujours ; c'est un problème non résolu ;<sup>137</sup> je n'ai remarqué que ces derniers jours que ma démonstration donnait un peu plus que le th. de Khintchine.

Je m'aperçois que je me laisse entraîner, en vous parlant de "l'essentiel" que je voulais réserver pour une autre lettre si je ne vous vois pas Mardi. Je m'arrête donc. Bien cordialement

P.Lévy

Je ne veux pas prolonger notre discussion. Il se peut qu'avant 1928 j'aie fait des confusions entre le point de vue de Bernoulli et celui de la loi forte des grands nombres. Mais depuis 1929 et en tout cas depuis 1930 je peux vous affirmer que,

<sup>135</sup>In Lévy, Paul (1938) L'arithmétique des lois de probabilité. J. Math. Pures Appl. 17:17-39, Lévy cited these results without proof. See Loève, *Probability Theory*, I, Chap. 18, Sect. 2 (4th ed., Springer (1978)).

<sup>136</sup>In Lévy, Paul (1938) L'arithmétique des lois de probabilité. J. Math. Pures Appl. 17:17-39, Lévy gave a proof of these results, which generalize Khinchin's.

<sup>137</sup>One year later, in his paper Lévy, Paul (1938) L'arithmétique des lois de probabilité. J. Math. Pures Appl. 17:17-39, Lévy remarked that the answer was negative.



sauf un lapsus toujours possible, je n'en ai pas fait, et qu'en tout cas il n'y en avait pas dans ce que je vous ai dit l'autre jour chez M. Borel.

- Mardi, je compte, au cours de mon exposé signaler 3 ou 4 problèmes que je considère comme pouvant être intéressants. Pour un ou deux, je n'exclus pas l'idée qu'il y ait, comme pour le théorème de Cramer, une solution simple à laquelle je ne pense pas ; mais il y en a au moins un qui est un problème que l'on peut faire progresser beaucoup avant d'arriver à une solution définitive "reconnaître si une loi est indécomposable".

## Lettre 35

[Un feuillet isolé daté en bas]

### 1 Simplification de la démonstration de Krasner<sup>138</sup>

$$f_p(x) = 1 + x + \dots + x^{p-1} = P(x)Q(x).$$

A partir du moment où l'on a observé que les coefficients sont entiers, il n'y a qu'à écrire  $f_p(1) = p = P(1)Q(1)$ , d'où on déduit que, pour  $p$  premier, un des facteurs se réduit à l'unité.

### 2. Remarques arithmétiques.

Soient  $P(x)$  et  $Q(x)$  deux diviseurs du polynôme  $f_p(x)$  réels à coefficients tous égaux à 0 ou à 1. L'équation

$$P(x) = Q(x)$$

n'a pas d'autre racines rationnelles<sup>139</sup> possibles que 1 ou  $-1$  (les coefficients extrêmes étant 1 ou  $-1$ ). Donc : *si  $n$  est un entier  $> 1$ , le nombre entier  $N = f_p(n)$  a au moins autant de diviseurs distincts que le polynôme  $f_p(x)$  a de diviseurs du type considéré.* Il ne peut être premier que si  $p$  est premier.

*Si  $p$  est premier, il a des chances singulières d'être premier.* Il est en effet premier pour

$$n = 2, p = 1, 2, 3, 5, 7, 13, 17 \text{ (pas pour 11)}$$

$$n = 3, p = 3 \text{ et } 7 \text{ (pas pour 2 et 5)}$$

$$n = 5p = 3 \text{ (pas pour 2 et 5)}$$

le 24 /11/37 P.Lévy

---

<sup>138</sup>During his conference on January 12, 1937, Lévy announced a question of polynomial decomposition related to the arithmetic of the uniform laws. A few days later, Krasner gave the answer to this simple question. It was later published as a note by Krasner and Ranulac (an anagram of *canular* (meaning practical joke)), (1937) CRAS 204:397–399. This jesting was no doubt inspired by Hadamard, who was always ready to ridicule the hieratic Emile Picard, perpetual secretary of the Academy of Science. After *canular* article was published, Hadamard remarked that Liénard, as well as Raikov had given an independent proof in the case of  $p$  prime. (Khinchin had alerted Lévy to Raikov's result in answering Lévy's letter on this topic.)

<sup>139</sup>Meaning a non-zero integer root.

P.S M.A.Liénard a, indépendamment de Krasner, démontré le théorème relatif au cas où  $p$  est premier, et fait indépendamment de moi la remarque ci-dessus.

### Lettre 36

Paris- le mars 1937<sup>140</sup>

Mon cher Collègue,

Je ne comprends pas bien votre objection au sujet du Wronskien.<sup>141</sup> Je n'aurai à reprendre cette question pour mon cours que dans 10 mois ; je précise que d'ici là j'aurai pu lire votre Note et y réfléchir.

Il ne me reste qu'un exemplaire du N° de l'enseignement scientifique qui contient mon article ;<sup>142</sup> ou plutôt je n'en ai eu qu'un (car ils ne donnent pas de tirage à part) ; je le déposerai un de ces jours à l'Institut Poincaré. Gardez-le le temps que vous voudrez, mais pas définitivement.

Bien cordialement à vous.

P.Lévy.

J'ai obtenu depuis 6 semaines un résultat, dont je suis presque sûr ; mais il subsiste un léger doute qui retarde sa publication :

Il peut exister des polynômes  $P(x)$  à coefficients non tous positifs, tels que la série  $F(x) = \exp P(x) = \sum A_n x^n$  ait tous ses coefficients  $\geq 0$ . Tel est le cas, si  $c$  est assez petit, pour  $P(x) = x^p - cx^{2p} + x^{2p+1} + x^{2p+2}$ . J'ai même à ce sujet la condition nécessaire et suffisante que doivent vérifier les degrés des termes négatifs pour que, si les modules des coefficients sont assez petits,  $P(x)$  ait la propriété indiquée. Alors  $F(e^{iZ})/F(1)$  est la fonction caractéristique d'une loi de probabilité intéressante à considérer. Avec les lois de ce genre on peut former des exemples résolvant tous les problèmes posés dans ma Note du 11 janvier.

### Lettre 37

Paris - le 8/1/39

Mon cher Collègue,

Je crains de vous avoir dit quelque chose d'inexact en ce qui concerne le plus grand des intervalles séparés sur l'intervalle fixe  $(0, 1)$  par  $n$  points choisis au hasard.

---

<sup>140</sup>Incomplete date.

<sup>141</sup>Fréchet had published a number of papers on the properties of a system of functions for which the Wronskian vanishes in an interval: see for example Fréchet, Maurice (1937) Bull. Math. Facultés Sci. Grandes Ecoles 3:105–109, or Fréchet, Maurice (1937) Bull. Calcutta Math. Soc. 29:105–109.

<sup>142</sup>L'Enseignement scientifique, a journal of instruction in science, was published by Eyrolles between 1927 and 1939, appearing ten times each year. It became an organ of debates and inquiry of all sorts about teaching (health, student overwork, programs, pedagogical ideas, analysis of contest reports, etc.). Hadamard, Lebesgue, Montel, Deltheil, Boll, Brunschvicg, Wallon all contributed to the journal. As part of a discussion of vector calculus, Lévy wrote an article (81:3–5 (October 1935)) to raise interest in vector notation, illustrated with Binet's formulas. The Wronskian, mentioned earlier in this letter, led Fréchet to write an article in Number 111 (1939) about problems posed by poor use of the idea. His article occasioned various comments by professors of Classes Préparatoires.

Pour que vous ne restiez pas sur cette erreur, je vous envoie ci-inclus un résumé de mes résultats sur cette question, en y remettant aussi deux énoncés, ceux où figurent des logarithmes itérés, dont je ne suis pas encore absolument sûr. Dans les questions de ce genre, il suffit en général de considérer des valeurs de  $n$  en progression géométrique pour prévoir le résultat ; la démonstration exige quelques précautions, et je n'ai encore terminé qu'une des vérifications nécessaires et, sans doute, la plus facile.

Je présenterai une Note<sup>143</sup> dès que je serai sûr, ou que j'aurai renoncé à préciser ces points.

Les problèmes analogues relatifs au plan et à l'espace sont bien plus difficiles ; pour le moment je ne vois rien\*. Vous pourriez peut-être le signaler à vos élèves, comme sujets de recherches possibles.

Croyez, mon cher Collègue, à mes sentiments dévoués.

P.Lévy.

\* en écrivant cette phrase, je m'aperçois qu'il y a au contraire un de mes énoncés dont l'extension est évidente ; mais c'est que  $Y_n$  n'y intervient que par son ordre de grandeur. Là où il faut une loi précise, je prévois plus de difficulté.

J'ajoute cet énoncé en P.S. de la feuille jointe.

Separated page attached to the letter:

Division d'un segment de droite par des points choisis au hasard.<sup>144</sup>

Soient  $x_1, x_2, \dots, x_n$  choisis au hasard entre 0 et 1.

$x'_1, x'_2, \dots, x'_n$  les mêmes rangés par ordre de grandeurs croissantes

Les lois dont dépendent les  $x'_v$  sont bien connues. Rappelons que, pour  $p$  fixé,  $n$  infini

$$\Pr\{nx'_p < x\} = \frac{1}{\Gamma(p)} \int_0^x e^{-\xi} \xi^{p-1} d\xi \quad (x > 0) \quad (\text{loi } \mathcal{L}_p)$$

La loi  $\mathcal{L}_1$  est la loi dont dépend  $|\xi|$  si  $\xi$  obéit à la 1ère loi de Laplace (Diverses propriétés des lois  $\mathcal{L}_1$  et  $\mathcal{L}_p$  sont indiquées dans ma lettre du 27/12/38<sup>145</sup>).

- Soient  $X_0 = x'_1, X_1 = x'_2 - x'_1, \dots, X_n = 1 - x'_n$  les  $n + 1$  intervalles séparés par les points  $x'_v$  ( $v = 1, 2, \dots, n$ ).

La loi à  $n + 1$  variables  $X_1, X_2, \dots, X_n$  (et  $X_0$  défini par  $\sum X_v = 1$ ) est identique à celle définie comme suit : Soit  $P_n$  le polyèdre régulier à  $n + 1$  sommets, de hauteur 1, dans l'espace à  $n$  dimensions ; les  $X_v$  sont les distances aux  $n + 1$  faces de  $P_n$  d'un point  $A$  choisi au hasard dans  $P_n$ . Donc, pour un quelconque

<sup>143</sup>(1939) CRAS 208:147-149.

<sup>144</sup>Fréchet had alerted Lévy to the results of Emil Julius Gumbel (then a refugee in Paris) on ordered samples and extreme values. Here we witness Lévy's first steps toward a stochastic integral based on Stieltjès sums with random points of subdivision.

<sup>145</sup>This letter seems to have been lost.

des  $X_v$ , comme pour  $x'_1$ ,

$$\Pr\{X_v > x\} = (1 - x)^n$$

$$\lim \Pr\{nX_v > x\} = e^{-x} \quad (c' \text{ est la loi } \mathcal{L}_1).$$

- Soit  $Y$  le plus petit des  $X_v$ .  $Y < h$  signifie que  $A$  est extérieur à un polyèdre homothétique à  $P_n$  et de hauteur  $1 - nh$ .<sup>146</sup> Donc  $\Pr\{Y > h\} = (1 - nh)^n$  ;  $nY$  dépend de la même loi que chacun des  $X_v$  ;  $n^2Y$  dépend d'une loi tendant vers la loi  $\mathcal{L}_1$  (pour  $n$  infini)
- Point de vue de la loi forte des grands nombres : la suite des  $x_n$  étant prolongée indéfiniment, et  $Y_n$  étant le plus petit des intervalles séparés par les  $n$  premiers de ces nombres. Soient  $a_1, a_2, \dots, a_n, \dots$  une suite de nombres positifs. Si  $\sum na_n$  converge, il est presque sûr que  $N$  existe tel que, pour  $n > N$ , on ait  $Y_n > a_n$  ; si  $na_n$  diverge, il est presque sûr qu'on a une infinité de fois  $Y_n < a_n$ .
- Sous réserve de vérification. Borne supérieure des  $Y_n$ . Il est presque sûr, si  $c > 1$ , qu'on a pour tout  $n$  assez grand  $Y_n < c \frac{\log \log n}{n^2}$  ; qu'au contraire on a une infinité de fois  $Y_n > \frac{\log \log n}{n^2}$   
T.S.V.P.
- Le plus grand des  $X_v$ . Soit  $Z = Z_n$  le plus grand des intervalles séparés par  $x_1, x_2, \dots, x_n$ . On a

$$\lim \Pr \left\{ Z_n < \frac{\log n - \log \theta}{n} \right\} = e^{-\theta}.$$

- Point de vue de la loi forte des grands nombres. Si  $c > 0, c' > 1$ , il est presque sûr que, pour  $n$  assez grand,

$$\frac{\log n - \log \log \log n - c}{n} < Z_n < \frac{\log n + c' \log \log n}{n},$$

et qu'au contraire ces inégalités sont une infinité de fois en défaut, la première si  $c \leq 0$ , la seconde si  $c' \leq 1$  (sous réserve de vérification en ce qui concerne l'inégalité de droite).

le 8/1/39.

P.Lévy.

P.S. Extension à l'espace à  $p$  dimensions. Soit  $Y_n$  la plus petite des distances des points  $A_1, A_2, \dots, A_n$  deux à deux, ces points étant choisis au hasard dans une région finie. Si  $\sum na_n^p$  converge on a presque sûrement  $Y_n > a_n$  pour tout  $n$  assez grand ; si  $\sum na_n^p$  diverge, on a presque sûrement une infinité de fois  $Y_n < a_n$ .

L'extension des autres énoncés me paraît difficile.

---

<sup>146</sup>Read  $1 - (n + 1)h$ . The published note includes this correction.

### Lettre 38

Girouard par le Mesnil St-Denis (S et O)

le 21-7-39

Mon cher collègue.

Je suis bien rentré dimanche à Girouard, et mon accès de fièvre n'aura pas eu d'autre inconvénient que de me faire manquer trois des séances les plus intéressantes.

J'ai fait quelques calculs relatifs à la question de la loi des erreurs.<sup>147</sup> Le résultat est que, pour une série de 500 expériences, on peut avec 5 % de chance d'erreur l'attribuer à l'une ou l'autre des lois de Laplace (si l'on considère comme sûr à priori qu'il s'agit d'une de ces lois).

Cette conclusion est loin de ce que je pensais, mais aussi, il me semble, un peu plus favorable à la thèse d'après laquelle il n'y a pas de difficulté excessive à distinguer ces deux lois.

J'ai rédigé une petite Note de 4 pages<sup>148</sup> que j'ai adressée à Fehr<sup>149</sup> pour l'enseignement mathématique, en ajoutant que si Wavre<sup>150</sup> et lui préféreraient l'ajouter au C.R du colloque je n'y ferai pas d'objection.

Bien cordialement à vous

P.Lévy

[*In leftside vertical margin, in Fréchet's handwriting:* Je lui écris le 23 Juillet que ma prévision est que les observations numériques suivent des lois différant des 2 lois de Laplace à peu près autant que celles-ci entre elles.]

### Lettre 39

Paris - le 15-10-39

Mon cher Collègue,

1° Je n'ai pas encore reçu l'article de Copeland<sup>151</sup> que vous m'annoncez ; mais j'ai trouvé son résumé dans le Zentralblatt. L'intégrale de C. se rattache à un cas particulier de la mienne.<sup>152</sup> Dans ce cas particulier, l'intégrale de C., comme la mienne, se réduit à celle de Riemann. Mais l'intérêt des notions que j'introduis

<sup>147</sup>The *Conférence Internationale sur les Applications des Probabilités* was held in Geneva from July 12 to 15, 1939. At this meeting, Fréchet raised once again the problem of distinguishing experimentally between a Laplace distribution with density  $\frac{1}{2}e^{-|x|}$  and a standard Gaussian distribution.

<sup>148</sup>(1940) *Enseign. Mat.* 38:227–231.

<sup>149</sup>H. Fehr was professor at the University of Geneva and co-director of the journal *L'Enseignement mathématique*.

<sup>150</sup>R. Wavre, a professor in Geneva, chaired the conference's local committee.

<sup>151</sup>See the references in the next letter.

<sup>152</sup>Lévy is referring to his note to (1939) CRAS 209:591–593. In his haste to point out the advantages of his construction, and to keep Fréchet from suggesting that Copeland had published first, he mistakenly conflated the two results, as he partially recognized in the next letter. The integral proposed by Copeland uses a measure with bounded variation (in fact, in the article, a distribution function), so that Copeland's integral is an extension of Stieltjès'. By contrast Lévy's integral allows him, for example, to integrate with respect to Brownian motion.

est qu'on peut définir des intégrales stochastiques de fonctions non aléatoires ayant un sens quand la notion d'intégrale (de Riemann, de Stieltjes, ... etc.) n'en a pas ; j'estime donc que C. a passé "à côté".<sup>153</sup>

Ainsi pour la courbe aléatoire du mouvement brownien plan, on peut parler de l'aire comprise entre la courbe et sa corde, au point de vue stochastique, mais non au point de vue de Riemann.

2° Pour les lois de Laplace, mon idée était la suivante.

Soit  $X$  une variable aléatoire ;

$$M = \mathcal{M}\{X\}, \quad M_\alpha = \mathcal{M}\{|X - M|^\alpha\}.$$

Il n'y a qu'à former  $\frac{M_4}{M_1^4}$  pour distinguer nettement les deux lois de Laplace.<sup>154</sup>

$\frac{M_2}{M_1^2}$  ne m'avait pas donné de résultats aussi satisfaisants que j'étais tenté a priori de le croire. Mais avec  $\frac{M_4}{M_1^4}$ , on peut fixer pour chacune des lois de Laplace un intervalle dont cette quantité a assez peu de chances de sortir (si on la calcule empiriquement d'après la moyenne de 500 expériences), et les deux intervalles sont extérieurs l'un à l'autre. Je dis "assez peu de chances" parce que je ne me rappelle pas le chiffre.

Le travail rédigé va paraître dans L'Enseignement Mathématique. Wavre m'a promis que j'aurais les épreuves en temps utile pour que je puisse vous les communiquer et que vous en teniez compte pour le C.R. du colloque. Il est vrai que c'était au début d'août. Mais le C.R. du colloque sera sans doute retardé par les événements plus que l'Ens<sup>t</sup> Math, si on l'imprime toujours en France. Je compte donc que la promesse de Wavre pourra être tenue.<sup>155</sup>

Bien cordialement P.Lévy

## Lettre 40

Paris - le 22-10-39

Mon cher Collègue,

J'ai lu le Mémoire de Copeland<sup>156</sup> que je vous rapporterai à notre prochaine réunion.

Je n'arrive pas à trouver très nouveau un théorème qui est un cas particulier de ce théorème connu : si une fonction de répartition a une limite pour laquelle la valeur probable soit finie, la valeur probable liée à la fonction variable tend vers celle liée

<sup>153</sup>Lévy softened this statement in the next letter.

<sup>154</sup>The advantage of such a quotient is that it is dimensionless and invariant under dilation.

<sup>155</sup>Lévy's doubts were realized. The corresponding volume was only printed, in Switzerland, at the end of 1940. It contained the article Lévy, Paul (1940) Sur la détermination expérimentale de la loi des erreurs. Enseig. Math. 38:227-231.

<sup>156</sup>Copeland, Arthur H. (1937) A New Definition of a Stieltjes Integral. Bull. Am. Math. Soc. 43:581-588.

à la fonction limite. Cela s'applique évidemment aussi bien si la fonction de répartition est une fonction de fréquence que si c'est une fonction des probabilités.<sup>157</sup>

Ce qui me paraît intéressant, c'est le rapprochement qu'on peut faire entre ce théorème et la loi forte des grands nombres pour la moyenne arithmétique, ou entre la suite des  $x_n$  de Copeland et une suite de nombres tirés au sort indépendamment les uns des autres et d'après une même loi. Des propriétés presque sûres d'une telle suite, Copeland en choisit une, qui caractérisera ce qu'on peut appeler un modèle de suite de nombres tirés au sort ; la notion est analogue à la notion de collectif. On n'imite pas le hasard ; mais un tel modèle imite quelques unes des ses propriétés, arbitrairement choisies.

Distinguons, d'une manière également arbitraire, des propriétés simples et des propriétés complexes d'une telle suite. Il est intéressant, lorsqu'on arrive à démontrer d'une propriété complexe qu'elle est presque sûre, de savoir si elle est indépendante des propriétés simples établies antérieurement, ou si elle en est la conséquence ; au point de vue axiomatique c'est un progrès. On est ainsi amené à énoncer les deux théorèmes relatifs à la suite de nombres tirés au sort d'après une même loi :

- a- la fonction des fréquences tend presque sûrement vers la fonction des probabilités.
- b- la moyenne vraie tend presque sûrement vers la moyenne théorique, si celle ci est finie.

et à remarquer que le théorème de Copeland comprend à la fois le théorème b, et le fait que la propriété presque sûre b résulte de la propriété presque sûre a.

J'étais arrivé de mon côté à l'idée que, sans partager les idées de de Mises, il ne faut pas proscrire absolument les modèles et les collectifs ; ils constituent un langage commode pour l'étude d'un point de vue axiomatique différent de celui où je m'étais placé jusqu'ici - et je m'en suis servi dans un mémoire récent (non encore imprimé).

- Quant à ce que j'ai écrit "il est passé à coté", je reconnais que le terme est inutilement péjoratif ; mais je me le suis appliqué à moi-même en jugeant rétrospectivement mes propres travaux - et il m'est arrivé de dire, par exemple, que Poincaré a passé à coté de la relativité restreinte (et de m'en étonner, d'ailleurs). Je suis évidemment plus sévère que vous dans mes jugements ; mais ce terme exprime assez bien l'impression que d'une idée qui pouvait conduire simplement à un résultat que je crois important on a tiré seulement une conséquence que, à tort ou à raison, je trouve moins intéressante. Evidemment, je pourrais dire "il est passé tout près" et admirer qu'il ait eu l'idée, ce qui est mieux que de ne pas l'avoir.

---

<sup>157</sup>Lévy must have been annoyed that Copeland did not chose a probabilistic framework and did not talk of random variables and laws. Rather, he worked in the spirit of von Mises' collectives, with a series of points ( $x_n$ ) in the integration interval whose properties "imitate" those of a series of independent random variables with the same distribution function  $f$ . The battle over collectives had just reached a culminating point in Paris with the defense of Ville's thesis at the start of 1939. Ville's work was a critical analysis of the idea of collectives (on this topic see Shafer 2009).

D'ailleurs l'idée de Copeland me paraît finalement assez éloignée de la mienne, et je me suis décidé lundi dernier, après avoir relu votre envoi, à présenter un Note à l'Académie en supprimant un renvoi où je citais Copeland et à le remplacer par une phrase précisant bien que mes intégrales  $\int u dv$  n'étaient pas des intégrales de Stieltjes,  $v$  n'étant pas à variation bornée.

Je me suis d'ailleurs aperçu depuis que sans introduire le hasard on peut définir  $\int u dv$  dans des cas où  $v$  n'est pas à variations bornées ; on a ainsi une nouvelle intégrale qui est une extension de celle de Stieltjes mais qui ne va pas aussi loin que mon intégrale stochastique. Je pense donc présenter prochainement une autre Note.<sup>158</sup>

Bien cordialement à vous P.Lévy

#### Lettre 41

*Carte postale adressée à Monsieur Maurice Fréchet 2 rue Emile Faguet Paris 14e, réexpédiée à l'Hôtel du Parc, Aulnay-en-Saintonge, Charente-Maritimes.<sup>159</sup> L'expéditeur est: M.Paul Lévy chez Madame Dullin, route de St Vérand par St Marcellin (Isère)*

St Marcellin - le 3-8-41

Mon cher Collègue,

Voici deux errata à mon mémoire sur le mouvement brownien. Le premier est important.

Page 515, lignes 17 à 23 - Cet alinéa est complètement faux. Le raisonnement ne s'applique qu'aux oscillations stochastiques (au sens de ma Note que vous avez transmise en juin), et non aux oscillations forcées.<sup>160</sup> Heureusement le résultat énoncé aux lignes 13 à 16 reste exact, parce que la définition donnée à la p.514 d'un modèle de mouvement brownien implique précisément l'existence de la limite dont il s'agit aux lignes 19 et 20 de la page 515.

---

<sup>158</sup>Lévy in fact made a presentation to the Société Mathématique de France on December 6, 1939, in which he discussed his paper Lévy, Paul (1941) Propriétés intrinsèques des fonctions et intégrales de Stieltjes. Bull. Soc. Math. Fr. 69:5-9.

<sup>159</sup>*Sic.*

<sup>160</sup>Lévy is referring to his long article Lévy, Paul (1940) Le Mouvement brownien plan. Am. J. Math. 62:487-550. For a sequence of random variables  $B(n)$  the forced oscillations are those of the sequence  $E(S(n))$ , the stochastic oscillations those of  $B(n) - E(B(n))$ . Here the subject is the oscillations of  $B(n) = \sum(\Delta X)_-$  where the increments  $\Delta X$  of the Brownian motion (for example) are calculated on a subdivision of the interval  $[0, T]$  by  $n$  successively chosen random points. The limit of  $B(n)$  when  $n$  tends to infinity is what Lévy calls the Brownian oscillation  $B(T)$ , which was introduced in a note to the CRAS of December 1938. It is the second order stochastic integral in Lévy's sense:  $\int (dX_s)^2$ . This Lévy integral equals  $T$  for almost every Brownian trajectory, and is a very sophisticated form of the process  $\langle X, X \rangle$  (see Sects. 7 and 8). In his 1940 article, Lévy also uses Brownian oscillation instead of quadratic variation for some random or deterministic curves for which it does not exist. Lévy corrected this error in articles and other works published after the war.



Page 534, lignes 10, 18 et 19 (chaque formule comptant pour une ligne), au lieu de  $\mu$  il faut lire  $m$  ; au lieu de  $\mu'$  il faut lire  $m'$ .<sup>161</sup>

Je viens de me décider à envoyer à Cramer, et à Tricomi<sup>162</sup> les deux derniers des 10 exemplaires que j'ai reçus. Je ne sais si M.Borel<sup>163</sup> a reçu celui que je lui ai envoyé. Votre exemplaire est peut-être le seul qui existe à Paris - Je n'ai pas été depuis deux mois dans une ville où il y ait une bibliothèque. Je compte être à Paris à la fin de septembre.

Bien cordialement P.Lévy

## Lettre 42

Hurigny- le 23-9-41

Mon cher Collègue,

Je reçois en même temps que votre lettre une lettre du Général Gouverneur de Polytechnique m'annonçant que, contrairement aux prévisions, l'Ecole reste à Lyon. J'avais depuis quelques jours une vague idée qu'il y avait un fait nouveau. Je vous avoue que j'en suis bien content. Je n'ai plus à me demander si ma décision était du courage ou du don Quichottisme- ou simplement la gêne de parler d'un danger que mes collègues ne semblaient pas soupçonner. Quoi qu'il en soit j'ai été très touché de l'intérêt que vous m'avez manifesté à cette occasion.

Vis à vis de Gau, cela change la situation. Je viens tout de même de lui téléphoner pour lui annoncer ma visite pour vendredi. J'irai de là en Hte Vienne, voir un de mes beaux-frères sur le point de repartir pour Paris (M.Dufour, Beauséjour, par St Mathieu, Hte Vienne). L'hiver, naturellement, mon adresse sera à l'Ecole Polytechnique, 14 avenue Berthelot, Lyon (d'où l'on fera suivre si après la fin des cours je quitte Lyon).

Comme vous serez sans doute à Paris avant mon beau-frère, je vous serais obligé de téléphoner au Secrétaire de l'Ecole des Mines - ou d'y entrer en passant - pour vous assurer que l'on a bien reçu l'avis que je restais cet hiver à Lyon. Je ne vous demande pas de téléphoner au Directeur, parce que j'ai déjà essayé de 3 autres moyens (carte directe, Vichy, et Polytechnique) et je ne voudrais pas l'agacer ; mais il n'est pas exclu que le 30 ou le 1er octobre vous soyez encore le premier à annoncer cette nouvelle.

Je vous adresse d'autre part par envoi séparé et recommandé un livre que je vous serais obligé de rendre de ma part à l'Institut Henri Poincaré (Nevanlinna - Eind.Funktionen<sup>164</sup>). Si vous croyiez qu'il y a risque à le passer, veuillez le renvoyer à ma fille. Mme Schwartz (Villa Costadau-Ceyrat, Puy-de-Dôme) ; elle ne sera pas fâchée de le garder jusqu'à nouvel ordre, mais je l'avais emprunté il y a 18 mois et ai des scrupules.

---

<sup>161</sup>Page 534 is at the heart of a proof of an important theorem on planar Brownian motion: the trajectories have vanishing area a.s.

<sup>162</sup>Francesco Tricomi was in Turin. His work was interrupted during the war.

<sup>163</sup>Borel was removed from office as mayor of Saint-Affrique on May 12, 1941. He was arrested in November 1941. On this topic see Mazliak and Shafer (2011).

<sup>164</sup>The book of Nevanlinna (1936).

En tout cas apportez mes souvenirs aux collègues parisiens - tout particulièrement MM. Borel, Montel, Cartan.

Bien cordialement,  
P. Lévy

### Lettre 43

Le 4 Novembre 1942<sup>165</sup>

Mon cher Collègue,

J'ai bien reçu votre lettre du 22 Octobre et l'envoi qu'elle m'annonçait. Je ne suis pas rentré à Grenoble ni à Montbonnot, et ne sais pas où je passerai l'hiver. Cela explique que votre envoi ait mis 12 jours à m'atteindre. Ecrivez moi à l'adresse *Paul Piron, Institut polytechnique, 46 Avenue Félix Viallet, Grenoble*- Robert Piron, qui a lui-même quitté Montbonnot, fera suivre.<sup>166</sup>

Bernstein- J'ai ici sa conférence du colloque de de Genève (Actualités scientifiques ; N° 738 ; 1938). C'est une des rares brochures échappées au désastre. Ne me l'envoyez donc pas. Par contre je vous ai demandé, et ai demandé depuis à Love, dont j'espère une prochaine réponse, des précisions sur la communication de B. au congrès de Zurich.

Copeland<sup>167</sup> - Je me suis maintenant rappelé qu'en effet c'est lorsque je vous ai parlé de mes idées sur l'intégrale stochastique, que vous m'avez dit qu'il y avait une idée analogue de C. sur l'intégrale de Stieltjes ; vous ne saviez plus exactement laquelle, et m'avez prêté la brochure.

La revoyant, je m'explique pourquoi je n'avais pas retenu qu'il y eut une antériorité de Copeland sur ces questions. C'est qu'en effet il n'y en a pas qui soit en rapport avec mon intégrale stochastique. L'idée de Copeland est absolument différente de la mienne. Je précise :

1° L'intégrale de Stieltjes étant désignée par  $\int_{\alpha}^{\beta} g(x) df(x)$ , *Copeland, contrairement à ce que vous croyez, ne fait pas intervenir le hasard*. Il n'y a de commun avec moi que l'idée d'une convergence non sûrement réalisée pour une suite quelconque. Il choisit une suite de points  $x_n$ , indépendante du hasard, mais dépendant de  $f(x)$  qu'il suppose à cet effet monotone. Dans mon int. stoch., les  $x_n$  dépendent du hasard, et non de  $f(x)$ .

Comme on ne change rien d'essentiel en prenant  $f(x)$  comme nouvelle variable, je suppose maintenant  $f(x) = x$ .

Les  $x_n$  de Copeland sont des nombres tels que la fréquence limite relative à chaque intervalle partiel  $(\alpha', \beta')$  soit  $\beta' - \alpha' / \beta - \alpha$ . C'est une condition presque sûrement vérifiée par le hasard, mais il ne parle pas du hasard.

<sup>165</sup>No address given, as was true of many letters during the Occupation.

<sup>166</sup>Lévy's apartment in Paris was searched and vandalized by the Gestapo and its French auxiliary. In August 1942, 'foreign' Jews in the South Zone were delivered to the Germans. Lévy and thousands of other French Jews took refuge in the Italian zone.

<sup>167</sup>This continues the discussion begun in 1939. Fréchet is being shown that Lévy's integral has nothing to do with Copeland's.

2° Il définit l'intégrale comme  $\lim \frac{g(x_1) + \dots + g(x_n)}{n}$ . C'est une moyenne de Césaro. Je la définis comme limite (presque sûre, ou en probabilité, ou en moyenne) d'une somme  $\sum \frac{g(\xi_{i-1}) + g(\xi_i)}{2} (\xi_i - \xi_{i-1})$ , les  $\xi_i$  étant  $x_1, x_2, \dots, x_n$  rangés par ordre de grandeur.

Au point de vue probabilités, le résultat auquel il arrive est classique. Les  $x_n$  étant supposés donnés par le hasard, la condition de Copeland est presque sûrement vérifiée : c'est la loi forte des grands nombres. De cette loi, on déduit que la moyenne vraie  $[g(x_1) + \dots + g(x_n)]/n$  tend (presque sûrement) vers la moyenne théorique  $\int_{\alpha}^{\beta} g(x) dx / \beta - \alpha$ . C'est l'équivalence de ces deux formes de la loi forte que Copeland retrouve.

3° Toujours dans le cas  $\int g(x) dx$ - cette intégrale a un sens à mon point de vue si  $g(x)$  est sommable. Je retrouve dans ce cas l'intégrale de Lebesgue. Pour Copeland, elle n'en a que si  $g(x)$  a seulement des discontinuités de première espèce. C'est une extension banale de l'intégrale de Riemann.\*

[on the bottom and vertically in the left margin: \* Prenez le cas où  $g(x) = 0$  si  $x$  rationnel et 1 dans le cas contraire. Je retrouve l'intégrale de Lebesgue  $\int g(x) dx = 1$ . Pour Copeland tous les  $x_n$  peuvent être rationnels ; il peut trouver zéro. On se demande d'ailleurs à la seconde page de son exposé s'il n'entend pas se restreindre à une suite particulière de nombres  $x_n$ , qui sont précisément rationnels. Il s'exprime en tout cas mal ; il faut comprendre qu'il ne voit dans cette suite qu'un exemple.]

4° Retour au cas  $\int g(x) df(x)$ . Copeland suppose essentiellement  $f(x)$  à variation bornée ;  $g(x)$  même condition que ci-dessus. La définition de Copeland conduit seulement à une extension banale de l'intégrale de Stieltjes. L'intérêt de mon intégrale est qu'elle a un sens dans des cas bien plus étendus (surtout si on se contente d'une limite en probabilité) :  $f(x)$  n'a plus besoin d'être à variation bornée.

de Finetti- Je connaissais de lui les deux notes que je cite, et non le gros mémoire<sup>168</sup> dont vous me parlez. Je dois avouer que de ces notes je me rappelle seulement qu'il y a une méthode très différente de la mienne ; je n'arrive pas à la reconstituer exactement. . . Si ; en vous écrivant, je me dis qu'elle est sans doute dans Cramer- un des rares livres que j'ai pu sauver. Elle y est en effet.

Peut-être pourriez vous demander un jour à Loève de regarder s'il y a dans ce gros mémoire une idée importante non contenue dans les Notes.

Je passe maintenant à mon travail.

En ce qui concerne K.,<sup>169</sup> j'ai bon espoir que vous pourrez de nouveau communiquer librement avec lui. Je voudrais penser que Steinhaus, Marcinkiewicz, Doebelin, ont autant de chances que lui de nous être un jour rendus, mais je ne le crois pas.<sup>170</sup>

<sup>168</sup>Probably De Finetti, Bruno (1930) *Funzione caratteristica di un fenomeno aleatorio*. Mem. Accad. Lincei 4:251–300, published in 1930, which Lévy cites in his book of 1948 (Lévy 1948).

<sup>169</sup>Kolmogorov. The battle of Stalingrad had just begun and would not end until February 1943.

<sup>170</sup>Marcinkiewicz had just died. By a miracle, Steinhaus survived in hiding. Doebelin committed suicide in June 1940 to escape the Germans during their pincer operation against the Maginot

Il est certain que vous ne pouvez pas refuser d'imprimer K. ; tout au plus lui suggérer qu'il aura peut-être plus de facilités que moi pour proposer son travail à une autre collection.

Mais il me semble qu'à votre place je ne craindrais pas de présenter au public deux livres sur le même sujet. Une courte introduction de vous expliquerait que ce sont des travaux indépendants (en raison des circonstances, personne ne pourra en douter) et conçus dans des esprits bien différents (c'est au moins très probable).

En tout cas il ne faut pas s'illusionner et écrire que la prolongation de la guerre résoudra peut-être la question. Je crois et espère qu'elle sera finie avant un an ; mais en tout cas il est très peu probable que les événements qui permettront d'imprimer mon mémoire et ceux qui vous permettront de communiquer avec K. soient séparés par un long intervalle.

Quoiqu'il en soit, je prends note de vos réserves, et continue à travailler comme si elles n'existaient pas. S'il le fallait, je proposerais ma prose à Gauthier-Villars ou à Vuibert.

Au sujet de mon travail, j'envisage de plus en plus d'y faire des changements assez sérieux, qui vous donneront satisfaction. Il m'apparaît déjà nécessaire, je crois vous l'avoir déjà écrit, de modifier sensiblement l'introduction ; le N°10, 2° pour citer les résultats de Francis Perrin<sup>171</sup> (qui a dès 1928 dépassé, pour la sphère, les résultats que j'avais obtenus) ; le N°13 où il est inutile de séparer, comme je l'ai fait, la partie positive de la partie négative du flux.

Je verrai, après étude du mémoire de K, ce qu'il y a lieu de faire. Mais je pense que vous avez raison de me dire qu'il vaut mieux intégrer dans mon texte ce que j'en dirai.

Quant à ce j'appelle les intégrales stochastiques, c'est évidemment un peu en dehors du sujet. J'ai en tout cas du nouveau à dire -ayant seulement ou surtout considéré en 1939 le cas de "l'intégrale presque sûre". Maintenant il me semble important de parler surtout des cas de convergence en probabilité ou en moyenne quadratique. C'est la convergence en probabilité<sup>172</sup> qui donne le plus d'extension à l'intégrale de Stieltjes et Young. La relation avec les fonctions aléatoires est qu'elles donnent l'application la plus naturelle de cette notion (et pour moi l'importance et l'intérêt d'une notion abstraite peuvent dépendre de ses applications). C'est pour cela que j'envisagerais volontiers une note annexe. Mais je n'exclus pas non plus un mémoire séparé.

Bien cordialement à vous.

P.L.

---

Line. The opening of sealed letter 11668 in 2000 (Bru, Bernard and Marc Yor (2000) CRAS 331) gave revealing evidence of the profundity of Doebelin's work in stochastic calculus.

<sup>171</sup>In 1928, F. Perrin published Perrin, Francis (1928) *Etude mathématique du Mouvement Brownien de rotation*. Ann. Sci. Ec. Norm. Supér. 45:1-5. This work was the origin of the study of Brownian motion in non-Euclidean geometries. Lévy himself moved from the sphere to hyperbolic geometry in his book of 1948 (Lévy 1948).

<sup>172</sup>See Sect. 8.

## Lettre 44

Montbonnot - le 1er juin 1943

Mon cher Collège,

Je reviens sur le P.S. de ma lettre de dimanche.<sup>173</sup> Le résultat obtenu me paraît assez important et, comme je ne peux pas présenter en ce moment de Note à l'Académie,<sup>174</sup> 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  $\lambda t$  ( $a$  et  $b$  étant changés en  $\lambda a$  et  $\lambda b$ )
- 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.<sup>175</sup>).

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  $(t_0, \infty)$  à l'intervalle fini  $(0, t'_0)$ , l'origine étant d'ailleurs n'importe quel point intérieur à  $(t_0, \infty)$ .

Finalement, la nature stochastique de  $E$  étant invariante par les substitutions qui servent de base au groupe des homographies, on voit que

<sup>173</sup>This letter seems to have been lost.

<sup>174</sup>Lévy took refuge in Grenoble, where he made some efforts to conceal his identity. The theorem that he describes here was received in a sealed letter (number 11904) on June 16, 1943, and opened at his request on December 13, 1944, in the presence of Borel and de Broglie. This result has important consequences for the detailed study of Brownian motion. (See Sect. 7.)

<sup>175</sup>See Letter 41, note 160, Letter 47, and note 197.

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.<sup>176</sup>

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.

## Lettre 45

Montbonnot - le 6 août 1943

Mon cher Collègue,

Il faut d'abord que je vous dise que je crains d'avoir calomnié G.<sup>177</sup> Je l'ai rencontré hier dans le tramway de Montbonnot à Grenoble. Il a été si surpris quand je lui ai parlé de mon article refusé et du prétexte invoqué que je ne doute plus qu'il ait ignoré toute l'affaire. Il était avec madame G., dont un mot, qui ne peut pas ne pas être naturel, prouve une invraisemblable distraction, et fait par là même tomber un autre de mes petits griefs. Elle était venue un jour excuser son mari, qui malgré un rendez vous donné, ne pouvait me recevoir chez lui. Le souvenir qui lui en est resté, c'est que je suis venu prendre le thé chez elle avec mon beau-frère P.L.<sup>178</sup> (qu'elle avait effectivement invité la veille de ce jour). Moi je m'étais dit "malgré ses occupations, G reçoit facilement ; s'il ne m'invite pas, il deviendra assez probable que

---

<sup>176</sup>In his 1940 article of the American Journal of Mathematics, Lévy had already noted the closeness of Ville's results to a weaker form of his own theorem.

<sup>177</sup>René Gosse, nicknamed *Doyen Bâtitseur* (*Dean Builder*), was relieved of his functions as director of the Institut Polytechnique and Dean of the University of Grenoble on December 6, 1940, under the 'purification decree' of July 17, 1940. He was also removed from his municipal functions (he had been a town councilor since 1928, belonging to the French Section of the Socialist International (SFIO)). The German police arrested him and his son Jean in December 1943. They were assassinated shortly afterward, probably by men of the Milice.

<sup>178</sup>Pierre Lévy, who graduated from the Ecole Polytechnique in 1908, was the brother of Suzanne Lévy. She married Paul Lévy, of the same last name. (See Letter 51.)

cette impolitesse est voulue”. Le raisonnement tombe ; ce qui reste c’est que M.G. d’une part est très occupé, d’autre part assez distrait pour donner le même jour deux rendez-vous incompatibles, et que Mme l’est encore plus.

Des autres indices défavorables à G, si je les examine, il y a une chose que je n’arrive pas à m’expliquer : ce qu’il m’avait dit de ses deux successeurs,<sup>179</sup> le doyen et le Directeur de l’I.P et que je sais être faux. Mais il peut avoir été sincère, et c’est peut-être à tort que cela m’a disposé, dans une autre occasion, à accepter trop facilement l’interprétation défavorable.

Donc je vous prie de ne rien retenir de ce que j’ai pu écrire de défavorable à G ; j’en suis même assez confus. Tout ce que je vous ai dit en faveur du doyen et de E subsiste.

J’ai enfin reçu la thèse de Loève.<sup>180</sup> Ma première impression est très favorable. Il aborde le problème sous un point de vue qui me semble nouveau.

Il a, dans la 1<sup>ère</sup> page, une expression inexacte me concernant. Loève, comme Bernstein, introduit des probabilités a priori. Mais Bernstein, si je me rappelle bien, n’introduit que les moments  $\mathcal{M}\{X^n\}$  ( $n = 1, 2$ ) ; l’idée originale de L est d’introduire  $\sum \mathcal{M}(X_i X_j)$ . La mienne avait été d’introduire les probabilités conditionnelles évaluées juste avant l’expérience. Ce ne sont pas des conditions trop restrictives que j’introduisais, puisque j’ai donné des conditions nécessaires et suffisantes. L. a employé une expression incorrecte, pour expliquer, ce qui est exact, que son point de vue est nouveau.<sup>181</sup>

Je suis surpris que vous croyiez mon travail sans rapport avec celui de Petrowsky.<sup>182</sup> Dans mon livre, p.266, je cite en l’attribuant à Kolmogoroff un théorème qui précise définitivement la loi du logarithme itéré ;<sup>183</sup> d’ailleurs avec une faute d’impression ; il faut évidemment avec les notations dont je me sers, écrire  $\int \varphi(t)e^{-\varphi^2(t)/2} \frac{dt}{t}$  ; K a du m’écrire en utilisant d’autres notations, et écrire

<sup>179</sup>René Gosse was replaced by Maurice Gignoux as the head of the Faculté des Sciences. Gignoux, in turn, was relieved of his duties at the Liberation for ‘having profited from a law contrary to university traditions.’ Lucienne Gosse, in a book about her husband Gosse, Lucienne (1964) René Gosse, Chronique d’une vie française. Plon, Paris, depicted Gignoux as a blind follower of the Vichy regime. Gosse’s successor as head of the Institut Polytechnique was Félix Esclangon. The University of Grenoble, which housed many leading scientists, came under the close surveillance of the Vichy government.

<sup>180</sup>Fréchet directed Loève’s thesis, and was its main dispatcher to the mathematical community. The thesis was published in the Journal de Mathématiques pures et appliquées, but only after the Liberation.

<sup>181</sup>Lévy already recognized the value of Loève’s work. Loève’s results on covariance and random functions of second order, from the period 1944 to 1947, were included as an appendix to Lévy (1948).

<sup>182</sup>Lévy used various spellings of this name. In mathematics, he made reference to two articles by Petrovskij, I.G. (1933) Über das Irrfahrproblem. Math. Ann. 109:425–444 and Petrowsky, I.G. (1935) Zur ersten Randwertaufgabe der Wärmeleitungsgleichung. Compos. Math. 1:383–419. Using Petrovski’s methods, Kolmogorov found an improved form of the law of iterated logarithms, of which he informed Lévy but never published.

<sup>183</sup>Lévy (1937). The misprint was corrected in the post-war edition.

$\int \varphi(t) e^{-\varphi^2(t)} \frac{dt}{t}$ , ce que j'aurais eu le tort de recopier simplement. Or j'ai demandé des explications à K, qui m'a répondu "cela se déduit aisément des résultats de Petrowski sur l'équation de la chaleur" Or il est presque évident que demander s'il existe presque sûrement un  $\tau$  tel que pour  $t < \tau$  on ait  $X(t) < \sqrt{t}\varphi(t)$  [ $X(t)$  étant la fonction aléatoire du mouvement brownien] revient à se demander si la solution  $u(x, t)$  de  $\frac{\partial^2 u}{\partial x^2} = 2 \frac{\partial u}{\partial t}$ , nulle pour  $t > 0$ ,  $x(t) = \sqrt{t}\varphi(t)$ , et pour  $t = 0$ ,  $x = 0$ , singulière à l'origine, positive pour  $t > 0$ ,  $x < \sqrt{t}\varphi(t)$  (conditions qui la définissent à un facteur constant positif près) est telle que l'intégrale

$$U(t) = \int_0^{\sqrt{t}\varphi(t)} u(x, t) dx$$

qui décroît nécessairement avec  $t$  ait une valeur initiale finie ou infinie. Dans le second cas, il y a perte instantanée de chaleur à travers l'extrémité froide  $x = \sqrt{t}\varphi(t)$  ; si on part d'une chaleur initiale (ou d'une probabilité) égale à 1, on tombe instantanément à zéro. Cela veut dire qu'il est infiniment peu probable que la particule brownienne mobile ne franchisse pas cette limite. Dans l'autre cas au contraire, elle ne la franchit dans un intervalle  $(0, t)$  que dans des cas dont la probabilité  $\frac{U(t)-U(0)}{U(0)}$  tend vers 0 avec  $t$ .

De l'interprétation de  $u(x, y, t)$  ou  $u(x, y, z, t)$  (cas du plan ou de l'espace), en intégrant  $u(x, y, t) - u(x', y', t)$  par rapport à  $t$  de 0 à  $\infty$  (ou la même chose avec  $z$ ) on déduit très aisément aussi un très joli théorème : Une particule brownienne initialement en  $A$  à l'intérieur d'un contour  $C$  (ou d'une surface  $S$ ) finit presque sûrement par en sortir. La probabilité qu'elle en sorte à travers un élément de cette surface est proportionnelle à sa mesure harmonique, c'est à dire à  $\frac{dg}{dn} ds$ ,  $g$  étant la fonction de Green relative à l'équation de Laplace  $\Delta u = 0$ , au problème de Dirichlet, au point  $A$ , et au contour  $C$  (et à la surface  $S$ ).

Ce théorème m'a été indiqué par mon gendre Schwartz,<sup>184</sup> qui ne se rappelait pas où il l'avait vu. J'en ai très vite trouvé, pour le plan, une démonstration basée sur la représentation conforme.<sup>185</sup> Je viens de voir que, quelque soit le nombre des variables, cela se déduit immédiatement de l'équation de la chaleur. Connaissez-vous le résultat. De qui est-il ? Se trouve-t-il dans le mémoire de Petrowski ?

Bien cordialement à vous

P.L.

Pour la question Laplace contre Gauss, vous n'avez pas répondu à la question que je vous ai posée un jour : la loi de Gauss à  $n$  variables (si  $n > 1$ ) n'est-elle pas effectivement une découverte de Gauss. Si oui, cela rend assez difficile de trouver

<sup>184</sup>Schwartz mentioned several times that Lévy had the habit of attributing to him results that he had in fact found and explained himself to Schwartz. Kakutani proved the same theorem independently in Japan in 1944.

<sup>185</sup>In two dimensions, this problem can be conformally mapped into a circle, with  $A$  its center and the 'surface element' an arc of the circle. Then the theorem becomes obvious (see Sect. 6). Thus Lévy used the conformal invariance of planar Brownian motion, a feature of which he was probably already aware in 1941, but did not publish until 1945.



une attitude pratique satisfaisante ; on ne peut guère dire loi de Laplace si  $n = 1$  et loi [la suite en marge gauche verticale : de Gauss si  $n > 1$ . L'expression variable laplacienne n'est d'ailleurs pas très commode, en raison de l'usage très différent du mot laplacien. Aussi je serais d'avis dans ce cas de me conformer à l'usage, tout en ne perdant pas une occasion de rappeler les rôles de Moivre et Laplace. D'ailleurs la solution qui consiste à dire variable normale ne me satisfait pas ; le mot normal est trop mis à toutes les sauces, et je préfère pouvoir dire que des variables sont orthogonales et normales si elles vérifient  $\mathcal{M}(X_k^2) = 1$ ,  $\mathcal{M}(X_h X_k) = 0$ .<sup>186</sup> Mais surtout qu'on ne me dise pas (comme le fait encore Bass, que j'ai vu le 11 juillet) qu'elles sont indépendantes !]

[following on the first page, and vertically: D'accord pour Slutsky dont je retrouve le mémoire écrit à l'occasion du Colloque de Genève.<sup>187</sup>

Je suis content de voir que vous avez trouvé une solution pour ce mois d'août. J'espère que vous en êtes content. "La glacière" en août, c'est parfait<sup>188</sup> - et c'est une région où l'on est bien nourri-

Ici, on reste inquiet des répercussions locales possibles d'événements par eux-mêmes très heureux.]

[In left headings of the first page: Il se peut que je parte du 15 au 25.]

## Lettre 46

Montbonnot- le 27 août 1943

Mon cher Collègue,

Je me suis promis de vous récrire avant que vous quittiez le Puy de Dôme. Mais je suis confus de constater que la fin d'août arrive sans que j'ai beaucoup avancé, ni mon travail personnel, ni la lecture de la thèse de Loeve. Pourtant je n'ai pas eu d'autre repos que le repos agréable et forcé que me fait parfois prendre la visite d'amis.

Pour la thèse de Loeve, ce sont bien quelques pages du début qui ont provoqué une observation que la suite ne justifie pas. Cette suite me déçoit d'ailleurs un peu. Je n'ai fait que parcourir les 50 premières pages ; mais j'ai lu avec attention l'énoncé de son théorème fondamental\* (\*page 45). Il est sans doute, à la lettre, nouveau ; mais il ne m'apparaît, suivant l'expression que lui-même emploie quelque part, que comme une insignifiante généralisation d'un de mes théorèmes (th. 67,2 de mes Variables aléatoires, p.243).

Cela nécessite quelques explications.

1° Je suppose les moments du second ordre bornés ; mais j'ai commencé par indiquer<sup>189</sup> au début du §67 (p.238,1er alinéa) qu'il est facile de s'affranchir

---

<sup>186</sup>See Letter 50.

<sup>187</sup>Slutsky, Eugène (1938) Sur les fonctions aléatoires presque périodiques et sur la décomposition des fonctions aléatoires en composantes stationnaires. In: Colloque consacré à la théorie des Probabilités. Genève, 11-16 Octobre 1937. Hermann, Paris.

<sup>188</sup>St Eloy la Glacière, in the Livradois, 50 kilometers from Clermont-Ferrand, was Fréchet's regular summer vacation spot.

<sup>189</sup>Sic.

de cette condition par le procédé indiqué au § 34 c'est-à-dire au fond par l'utilisation de la notion de suites équivalentes (au sens de Khintchine) ; (il s'agit d'ailleurs ici de résultats au sens de Bernoulli). La condition (1) de Loève revient à dire qu'on est dans le cas où la réduction est possible (la "variable corrigée"  $X'_v$  ne diffère de  $X_v$  que dans des cas de probabilité  $\alpha_v$ , et  $\sum \alpha_v$  étant très petit).

2° Le passage des sommes

$$S_n = X_1 + X_2 + \dots + X_n$$

aux sommes

$$S_n = X_{n,1} + X_{n,2} + \dots + X_{n,n}$$

est banal.

3° Le seul point délicat est le suivant. J'introduis p.242 une fonction  $\mathcal{S}(t)$  et explique que le gain au bout du temps  $t$  et le gain après  $n$  coups sont choses très différentes ; et c'est  $\mathcal{S}(t)$  que j'introduis dans mon théorème 67,2.

Or l'hypothèse fondamentale (3.4) revient à dire que la valeur de

$$\sum_1^n \sigma_{n-1}^2 \{X'_v\}$$

diffère très peu d'une constante  $t'$  (fonction non aléatoire de  $n$ ). La correction à faire pour passer de  $S_n$  à  $\mathcal{S}(t)$  a donc un écart type très petit, et mon résultat relatif à  $\mathcal{S}(t)$  s'applique à  $S_n$ .

4° Enfin je fais l'hypothèse  $\mu'_v = 1$ ,  $\mathcal{M}_{v-1}\{X_v\} = 0$  que Loeve estime être assez restrictive. Mais je suis à peu près sûr d'avoir indiqué quelque part que s'il n'en est pas ainsi il n'y a qu'à écrire

$$X_v = \mu_v + (X_v - \mu_v)$$

et étudier séparément les deux termes (si je ne me trompe, c'est à la fin de mon mémoire publié en 1935 dans le Bulletin des Sciences Mathématiques, mémoire d'où mon th.67,2 est tiré). Or l'hypothèse (3) de Loève est que  $\sum X'_v$  est très petit.

Je considère donc que le théorème de Loeve peut être considéré comme banal par quelqu'un qui aurait bien compris l'esprit de mon mémoire de 1935 et certains passages de mon livre. Pour être plus large, on peut dire qu'il a tiré parti du "programme" que j'avais esquissé à la fin de ce mémoire.

Voilà tout ce que j'ai à vous dire, pour le moment, sur ce sujet.<sup>190</sup>

Quant à mon travail personnel, il s'allonge parce qu'il me paraît indispensable d'indiquer au moins sommairement quelques aspects de la question que j'avais d'abord laissés de côté ; notamment la dérivée aléatoire, les intégrales stochastiques,

---

<sup>190</sup>See the preceding letter.

et les équations différentielles stochastiques. En outre j'ai plus de mal que je ne le pensais à préciser mes idées sur les questions en rapport avec la note de Doebelin de Novembre 1939.<sup>191</sup>

A propos de la dérivée aléatoire (au sens de Slutsky) pouvez-vous me donner un exemple assez simple de fonction  $X(t)$  ayant une telle dérivée mais qui presque sûrement n'ait nulle part de dérivée proprement dite.

Ce que je trouve de plus simple part de la remarque que

$$f(x, \xi) = x^{1+\sin^2(\xi-\log x)} \quad (x > 0, \xi \text{ aléatoire})$$

a ce caractère à l'origine.<sup>192</sup> Je pose alors  $f(x, \xi) = 0$  si  $x < 0$ ,  $f(x, \xi)$  si  $x > 0$ , et j'étudie des séries de la forme

$$X(t) = \sum c_n f(t - a_n, \xi_n),$$

les  $a_n$  étant partout denses dans un intervalle. Si les  $c_n$  tendent assez rapidement vers zéro, les caractères cherchés sont réalisés *par les points*  $a_n$ . Il faut procéder à une discussion assez minutieuse pour savoir ce qui se passe en dehors de ces points.

- Je note avec intérêt votre réponse sur la loi de Gauss-Bravais.
- J'ai correspondu ces derniers temps avec Bass. Il a besoin d'être guidé, mais il ne manque pas d'idées.<sup>193</sup> Il y aurait peut-être intérêt à le mettre en rapport avec Ville ; leurs recherches sont pour le moment orientées tout à fait dans le même sens (mes nouvelles de Ville remontent d'ailleurs à 3 ou 4 mois). Je l'ai signalé à Bass.

Bien cordialement à vous

P.L.

Mon gendre Schwartz n'a pu disposer que de très peu d'exemplaires de sa thèse.<sup>194</sup> Je n'en ai que les épreuves.

Si elle vous intéresse, je pense qu'il n'est pas impossible qu'il ait bientôt d'autres exemplaires ; je ne me rappelle [*following in margin*: pas où en est cette question. En tout cas je pense bien qu'elle se trouve à la Sorbonne ; ou à défaut Valiron, qui fut président de jury, pourra vous la prêter.]

<sup>191</sup>Most probably Doebelin's note of October 24, 1938.

<sup>192</sup>Lévy never ceased to point out that only the quadratic mean derivative has the properties that allow it to play the role of the ordinary derivative. In his note to CRAS on December 13, 1949, he mentioned the error on this point that he attributed to Kampé de Fériet and Slutsky.

<sup>193</sup>Lévy shows here the severity of a mentor to a disciple. Ten years later, Bass nevertheless guided Lévy in a calculation of the characteristic function of the Lévy area.

<sup>194</sup>Schwartz defended his thesis in Lyon. It was published in 1943 by Hermann, which was directed by Freymann. In his autobiography (Schwartz 2001), Schwartz wrote: 'Freymann, director of the publisher Hermann, did not lack for courage in publishing and promoting the thesis of a Jew in 1943.'

**Lettre 47**

Le 31-8-43

Mon cher Collègue,

Je pense que vous avez bien reçu, avant de quitter St Eloy, une lettre que je vous ai écrite il y a quatre jours. J'ai déjà plusieurs raisons de vous écrire.

D'abord je voudrais vous reparler de la dérivée aléatoire. Voici un exemple curieux. Je prends pour  $X(t)$  la fonction du processus additif défini par :<sup>195</sup>

$$\psi(z, \tau) = \log \mathcal{M}\{e^{iz[X(t+\tau)-X(t)]}\} = \tau \int_0^1 (e^{izu} - 1) dN(u).$$

Il est presque sûr que cette fonction est constamment croissante, si  $\int_0^1 dN(u) = +\infty$ , quoique ne croissant que par sauts. La dérivée est nulle presque partout. Quant à la dérivée aléatoire, elle existe partout, et est nulle.

On ne peut donc pas espérer retrouver la primitive par intégration de la dérivée aléatoire.

J'ai toujours eu l'impression que Kampé de Fériet et *illegible name*<sup>196</sup> s'exagéraient l'importance de cette notion. Je vois maintenant que c'est une notion dangereuse ; le langage employé pourrait suggérer des conclusions fausses. Ne concluez pas comme moi que ce n'est pas, et ne doit pas être, une notion fondamentale ?

D'autre part j'ai enfin réussi à mettre au point mes idées sur les intégrales stochastiques, suivant le programme esquissé déjà à la fin de l'introduction de mon mémoire de l'AJM.<sup>197</sup> Mon premier théorème fondamental est le suivant : si  $X(t)$  est la fonction aléatoire du mouvement brownien, si  $f(t)$  vérifie la condition<sup>198</sup>

$$\begin{cases} |f(t + \tau') - f(t)| \leq \varphi(\tau) & (0 < \tau' < \tau) \\ \int_0^1 \varphi^2(\tau) \frac{d\tau}{\tau} < +\infty \end{cases} \quad (1)$$

il est presque sûr l'intégrale<sup>199</sup> stochastique

$$\int_{t'}^{t''} f(t) dX(t) \quad (2)$$

est bien définie, tant au sens de la convergence presque sûre qu'au sens de la convergence presque partout.<sup>200</sup>

<sup>195</sup>A special case of the Lévy-Khinchin formula, which had originally been proposed by Finetti.

<sup>196</sup>The illegible name may be Wehrlé, judging from handwriting here and in the following letters. Philippe Wehrlé used quadratic mean derivatives in meteorology. Lévy learned from Kampé de Fériet and Slutsky of the note that is cited in the footnote 192.

<sup>197</sup>Lévy, Paul (1940) Le Mouvement brownien plan. Am. J. Math. 62:487-550.

<sup>198</sup>Lévy hoped that if  $\varphi$  behaved as  $\sqrt{\tau}$ , he could integrate it with respect to Brownian motion (a reference to his stochastic extensions of Young's integral).

<sup>199</sup>*Sic*. Incorrect syntax.

<sup>200</sup>Lévy never succeeded in proving this result, which he had to bypass to progress further. It is not cited in his 1948 book (Lévy 1948).

Je vous rappelle que mon intégrale stochastique, bien différente de celle de Slutsky, introduit le hasard dans le choix des points de division ; il n'est pas nécessaire que le hasard intervienne dans l'expression intégrée. Cette intégrale, si cette expression n'est pas aléatoire, définit une grandeur qui n'est pas non plus aléatoire. L'intégrale (2) est une fonctionnelle non aléatoire de la fonction aléatoire  $X(t)$ .

La condition (1) est beaucoup moins restrictive que celle moyennant laquelle (2) a un sens au point de vue classique, c.à.d. celui de L.C.Young. Ainsi, si  $\varphi(t) = t^\alpha$ , elle s'applique pour  $\alpha > 0$ , tandis que pour Young il faut  $\alpha > \frac{1}{2}$ .

Je passe ensuite au cas où  $X(t)$  dépend d'un processus additif quelconque. Puis à celui des équations différentielles stochastiques telles que

$$dY = f(X, Y, t) dX + g(X, Y, t) dt$$

Mon point de vue diffère essentiellement de celui de S.Bernstein parce que  $Y$  (avec une valeur initiale donnée) est une fonctionnelle non aléatoire de  $X(t)$ . Bernstein définit la loi dont dépend  $Y(t)$  sans se préoccuper de la manière dont  $Y(t)$  dépend de  $X(t)$  ; plus exactement, il pose autrement le problème et n'introduit pas  $X(t)$ , ce qui à mon avis, masque l'essentiel.<sup>201</sup>

Je suis persuadé que tout cela est très important, et m'efforce de mener à bien la seconde partie de ce programme.

Bien cordialement à vous

P.L.

Je pense que vous avez ma conférence de Lyon (Janvier 1941, Soc. Math de France, section du Sud Est).<sup>202</sup> Dans le cas contraire vous pourriez la demander à Henri Cartan (dont j'ignore l'adresse à Paris ; j'ai profité du voisinage pour lui envoyer quelques brochures<sup>203</sup>).

Je généralise un peu ma définition de l'intégrale stochastique en associant une définition à chacun des modes de convergence du calcul des probabilités. En outre je montre qu'on peut modifier de bien des manières la règle suivant laquelle on choisit au hasard les points de division, sans changer le résultat.

## Lettre 48

Montbonnot - le 6 septembre 1943

Mon cher Collègue,

Je reçois votre nouveau fascicule des actualités et vous en remercie vivement.<sup>204</sup> J'ajoute que j'ai honte de mon ignorance, en voyant que vous avez écrit plus de 200

---

<sup>201</sup>Bernstein worked in distribution on stochastic differential equations (SDEs), using a Cauchy method to obtain the solutions. Lévy, on the other hand, demanded an a.s. convergence of the approximations, a feature that later was often overlooked in France. Skorokhod's 1961 book on random processes mentions the concept of SDEs in Bernstein's sense, and does point out that Lévy was the first to construct the trajectories of the solutions.

<sup>202</sup>Lévy, Paul (1941) Les intégrales stochastiques. Bull. Soc. Math. Fr. 67-74.

<sup>203</sup>Schwartz and Cartan were in contact.

<sup>204</sup>Fréchet, Maurice (1943) Les probabilités associées à un système d'événements compatibles et dépendants. Hermann, Paris.

pages sur un sujet que je devrais connaître, car il est très voisin de ceux que j'étudie, et dont je ne sais presque rien ; j'ai tout juste le souvenir des conférences de Pòlya, que j'ai écoutées mais n'ai pas étudiées la plume à la main.

Pour mon travail, ce qui prime actuellement, c'est de continuer la lecture de trois livres que Bass m'a prêtés, et de mettre au point mes idées sur les processus stochastiques continus.

Je vous ai écrit récemment deux lettres, l'une adressée à St Eloy la Glacière, l'autre à Paris, où je croyais que vous rentriez. Vous la fera-t-on suivre ? Dans le doute je répète les deux points essentiels.

Le premier concerne la dérivée aléatoire. J'ai toujours eu l'impression que Wehrlé et Kampé de Fériet s'exagéraient l'importance de cette notion<sup>205</sup> qui ne saurait à mon avis jouer en calcul des probabilités le même rôle que la dérivée ordinaire dans l'intégrale classique. Une remarque que j'ai faite récemment, en cherchant un exemple de fonction aléatoire admettant une dérivée aléatoire mais n'admettant de dérivée qu'en des points ne remplissant aucun intervalle, m'a confirmé dans cette opinion.

Je considère le processus additif défini par :<sup>206</sup>

$$\psi(z, t) = \log \mathcal{M}\{e^{iz[X(t+\tau)-X(t)]}\} = t \int_0^\infty (e^{izu} - 1) dN(u).$$

$N(u)$  étant non décroissant,  $N(0) = -\infty$ ,<sup>207</sup>  $N(\infty)$  fini,  $\int_0^1 u dN(u)$  fini.

La fonction aléatoire  $X(t)$  ainsi définie est, presque sûrement, constamment croissante (quoique ne croissant que par sauts) ; or elle admet une dérivée aléatoire constamment nulle. Il est donc absolument impossible d'espérer retrouver la fonction initiale par intégration de la dérivée aléatoire... sauf évidemment dans des cas particuliers ; ceux là sans doute où cette dérivée est une dérivée au sens classique.

Cela ne vous semble-t-il pas justifier ma conclusion ?<sup>208</sup>

Notez que cet exemple n'est pas du tout particulier ni artificiel. Si vous avez sous les yeux mes formules concernant la forme générale des lois indéfiniment divisibles, celles que l'on rencontre dans les processus additifs continus, vous vous en convaincrez aisément.

Je vous signalai d'autre part le théorème suivant que j'ai obtenu récemment.

<sup>205</sup>Before the war, the physical meteorologists Philippe Wehrlé and Georges Dedebeant developed their own stochastic calculus. In 1937 Fréchet alerted them to the work of Slutsky. See the book Wehrlé, Philippe (1956) *L'univers aléatoire*. Editions du Griffon and Bru, Bernard (1998) *A propos des modèles probabilistes de turbulence au début du XXème siècle*. Matapli 55:51–64, and also Locker (2001).

<sup>206</sup>Read  $\tau$  as in the preceding letter.

<sup>207</sup>*Sic.* Lévy was probably confusing limits and meant to write either  $N(0) = 0$  or  $N(-\infty) = 0$ .

<sup>208</sup>Fréchet was probably convinced by Lévy's arguments, which were published in a note to CRAS, December 13, 1944. With the use of Poisson processes, which he had employed in his 1948 book (Lévy 1948), Lévy constructed strictly increasing random functions that have a vanishing derivative at every point a.s.

Soit  $\varphi(\tau)$  une fonction paire nulle pour  $\tau = 0$  croissant avec  $|\tau|$  et telle que

$$\int \varphi^2(\tau) \frac{d\tau}{\tau} \text{ converge à l'origine}$$

(On peut donc prendre une fonction tendant lentement vers zéro, comme  $\log(\frac{1}{1+\tau})$ .)  
Soit  $f(t)$  une fonction vérifiant la condition de Lipschitz :

$$|f(t + \tau) - f(t)| \leq \varphi(\tau)$$

Soit enfin  $X(t)$  la fonction aléatoire du mouvement brownien linéaire, définie par  $\log \mathcal{M}\{e^{iz(X(t+\tau)-X(t))}\} = -\tau \frac{z^2}{2}$ . Alors il est presque sûr que l'intégrale

$$\int f(t) dX(t)$$

a un sens, dans tout intervalle fini, comme intégrale stochastique définie à la fois presque sûrement et en moyenne quadratique.

Pour qu'elle ait un sens au point de vue classique (c'est à dire au sens de L.C.Young), il faut prendre pour  $\varphi(t)$  une fonction tendant vers 0 bien plus rapidement, par exemple  $\sqrt{t}/[\log \frac{1}{|t|}]^\alpha$  avec  $\alpha$  assez grand ( $> 1$  ou même  $> 2$ , je ne sais pas au juste). Cela montre que ma notion d'intégrale stochastique a un champ d'application bien plus étendu que celui de l'intégrale classique.

Je suis de plus en plus convaincu que c'est sur cette notion qu'il faut fonder la théorie des équations différentielles stochastiques, et que les travaux de S.Bernstein sont à reprendre de ce point de vue.<sup>209</sup> Mais il faut d'abord que je travaille à étendre mon théorème au cas d'un processus additif continu quelconque, notamment au cas du processus brownien non homogène par rapport au temps  $\{-\tau \frac{z^2}{2}$  étant remplacé par  $[\psi(t + \tau) - \psi(t)] \frac{z^2}{2}$ . Il faut alors une condition de Lipschitz plus restrictive.

Bien cordialement à vous.

P.L.

L'état sanitaire est moins bon à Grenoble que je ne le pensais au moment de ma précédente lettre.<sup>210</sup> Sur l'avis concordant de plusieurs médecins, je pense aller un peu à une autre altitude. Comme je ne peux pas encore vous donner mon adresse, écrivez moi comme toujours à la même adresse ; ma fille fera suivre ; mais cela ralentira la correspondance.

[In vertical left margin: Par le même courrier que votre livre, je reçois une carte de Wavregnote<sup>211</sup> Elle a mis douze jours, mais elle est arrivée.]

<sup>209</sup>See the preceding letter.

<sup>210</sup>After the capitulation of Italy, the militia was unable to maintain 'order,' and three days later Grenoble was taken by the Germans. See Giolitto, Pierre (1999) Grenoble 40-44. Perrin, Paris.

<sup>211</sup>See note 150. Lévy kept in contact with Wavre, who was in Geneva, to be able to continue publishing his work. Lévy's results on the projective invariance of Brownian motion had just appeared, September 1, 1943, in *Commentarii helvetici*.

[*Note by Fréchet in left margin*: D'après lettre suivante du 9 sep erreur de calcul rend ce théorème incertain<sup>212</sup>]

### Lettre 49

Monbonnot- le 9 septembre 1943

Mon cher Collègue,

J'allais vous écrire, lorsque je reçois votre lettre ; je commence par y répondre.

- 1° La thèse de mon gendre a été soutenue le 9 Janvier 1943 à Clermont-Ferrand. Titre : étude des sommes exponentielles réelles.<sup>213</sup>

Je m'aperçois que, si je n'ai pas d'exemplaires de cette thèse, sous sa forme définitive, j'ai un double pris de placards. Je vous en envoie un.

Malheureusement il a négligé de transcrire les corrections, et il peut y avoir des fautes d'impression qui rendent la lecture difficile.<sup>214</sup>

- 2° Je m'étais déjà dit que mon travail actuel serait mieux dans une collection de fascicules que dans un périodique. Cartan en a parlé de ma part à Freymann,<sup>215</sup> qui l'a accepté en principe pour ses actualités scientifiques ; mais je ne suis pas engagé, et, si vous le désirez, je peux donner la priorité pour votre collection d' "Exposés d'analyse générale".

Je crois en effet, surtout aujourd'hui 9 septembre, que bien avant un an les difficultés qui s'opposent à l'impression actuelle n'existeront plus.

Ce qui m'effrayerait un peu, ce serait l'obligation de revoir les parties déjà rédigées et d'alourdir et allonger le texte pour donner trop d'explications élémentaires. Mais je pense que nous nous arrangerons toujours. Je ne m'opposerai jamais à l'addition de notes explicatives en bas de page, si vous les jugez nécessaires.

- 3° Loève. Je suis trop "polarisé" dans mon propre travail pour l'étudier à fond. J'ai tout "parcouru". Mon impression reste la même : peu d'idées vraiment nouvelles ; mais une bonne utilisation d'idées déjà connues, et des résultats dépassant légèrement les résultats antérieurs.

Je n'en reste pas moins très bien disposé en faveur de l'auteur qui, à propos des lois limites au sens de Khintchine (V p.192–194 de mes variables aléatoires) m'a montré qu'il pouvait avoir des idées- et ce que vous m'écrivez me confirme dans cette opinion. Je crois qu'il faut l'encourager à continuer.

- 4° Pour ce que vous me demandez pour Fortet,<sup>216</sup> je ne peux que vous renvoyer au N° 72 de mes variables aléatoires (p.277 à 289). Vous aviez lu ce passage sur les

<sup>212</sup>Lévy would later recognize that his proof was incorrect.

<sup>213</sup>*Sic*. The original title was Etude des sommes d'exponentielles réelles.

<sup>214</sup>See Letter 46, note 194. The Hermann edition was preceded by a typed version, of which Lévy certainly had a copy.

<sup>215</sup>See Letter 46, note 194, and Letter 47, note 203.

<sup>216</sup>In 1939 Fortet defended his thesis *Sur l'itération des substitutions algébriques linéaires à une infinité de variables et ses applications à la théorie des probabilités en chaîne*. At the time of this letter, Fortet was already a well-known probabilist, studying the connections between potential theory and Markov processes (see in particular Fortet (1943)).



épreuves, et m'avez même signalé l'inutilité (logique), à la p.278 de l'indication  $\lambda_1 = 1, 08 \dots$ , que j'ai toutefois maintenue, parce qu'il m'a paru utile de montrer comment ces facteurs  $\lambda_n$  se rapprochent vite de l'unité. Je pense que ces calculs ont quelque rapport avec ce qui vous intéresse. Les formules fondamentales du 1° sont en tout cas des formules rigoureuses, applicables quelque soit  $n$ .

Je ne sais rien d'autre sur ces questions.

- 5° J'avais hésité, en recevant votre 2° fascicule sur les fonctions aléatoires, à vous demandez<sup>217</sup> le premier. Mais puisque vous posez la question, je n'hésite plus à vous répondre que je ne l'ai jamais reçu, et serais très content de le recevoir.
- 6° J'ai envoyé le 20 août à Wavre le théorème sur le mouvement brownien, celui que les Annales de Grenoble ont refusé.<sup>218</sup> J'avais auparavant reçu et corrigé des épreuves pour les Commentaires.

Je voulais d'autre part vous écrire pour deux raisons.

- 1° Je vous ai donné dans ma dernière lettre l'énoncé d'un théorème sur l'intégrale stochastique  $\int f(t) dX(t)$ ,  $X(t)$  étant la fonction aléatoire du mouvement brownien. Or je viens de m'apercevoir que j'ai fait une grossière faute de calcul. Il n'est pas impossible que le théorème reste exact ; mais, pour le moment, je ne peux rien affirmer.
- 2° Je trouve dans S.Bernstein, fascicule 738 des Actualités p.6 "je ne reviendrai pas ici sur la question que j'ai discutée ailleurs..." et il renvoie à sa communication au congrès de Zurich, 1932.

Cela est très important pour moi. La question est voisine de celle de mes intégrales stochastiques. Peut-être celles ci sont -elles moins nouvelles que je le croyais.

Or les Verhandlungen du congrès de Zurich n'existent pas à Grenoble. Pourriez vous, quand vous serez à Paris, jeter un coup d'œil sur cette Note de S.Bernstein ; si possible m'en faire un court résumé ; à défaut me dire si cela vous paraît avoir un rapport avec mes intégrales stochastiques.<sup>219</sup>

Je vous rappelle que pour ces intégrales, la fonction n'a pas besoin d'être aléatoire. J'introduis le hasard dans le choix des points de subdivision ; et il peut y avoir convergence en moyenne quadratique, ou presque sûre, vers une limite indépendante de ce choix. J'ai introduit cette notion (sans le mot) dans mon mémoire de l'A.J.M., et le mot dans une Note à l'Académie et dans une conférence à Lyon, que j'ai dû vous envoyer, mais que je vous renvoie avec le papier de Schwartz.

Bien cordialement à vous

P.L.

Je crains qu'il ne me faille deux bons mois, ou plus, pour terminer mon travail. Je voudrais bien mettre au point quelques idées qui sont encore bien confuses dans mon esprit.

---

<sup>217</sup> *Sic.*

<sup>218</sup> Politics at the University of Grenoble under the Vichy regime is evoked in Letter 45.

<sup>219</sup> In fact, Bernstein's memoire had nothing to do with Lévy's integral. As usual, Lévy was worried about a possible dispute over who had published first.

Pour la dérivée aléatoire, je vous ai indiqué ma conclusion. Je viens [*la suite en marge verticale gauche* : d'ailleurs, grâce à Bass, de relire différents mémoires ou notes de Slutsky, Wehrlé et Kampé de Fériet. Le mémoire de Kampé de Fériet (Annales de la Soc. Sc. de Bruxelles, 1939) est très bien. Mais il me semble qu'au N° 13 il se trompe sur la définition de l'intégrale aléatoire d'après Slutsky ;<sup>220</sup> et au N°24 il donne d'un théorème exact une démonstration complètement fausse. De plus, à partir du moment où il a indiqué le théorème de Khintchine  $R(h) = \int_{-\infty}^{+\infty} \cos(\omega h) dF(\omega)$ ,<sup>221</sup> il est bien évident qu'il peut appliquer tous les résultats classiques ou du moins connus sur les fonctions caractéristiques. Il a l'air de les redécouvrir.

Pour Laplace, Gauss, Bravais, je ne doute pas que vous ayez raison. Mais comment faire, pratiquement ? Si je dis "variable gaussienne" tout le monde comprend. Et comment pratiquement dire loi de Laplace s'il y a une variable et faire intervenir Bravais s'il y en a trois. S'il y avait un nom comme Bourbaki<sup>222</sup> qui signifie Laplace + Gauss + Bravais, cela vaudrait mieux. Je ne suis satisfait d'aucune solution.]

## Lettre 50

le 27-9-43<sup>223</sup>

Mon cher Collègue,

J'ai bien reçu votre lettre du 13 avant de quitter Monbonnot, où je ne retournerai sans doute pas avant un mois. J'ai beaucoup de choses à vous dire et commence par répondre à votre lettre.

Loi de Laplace. Je me décide à adopter cette dénomination, en disant pourquoi. Mais alors je dis aussi loi de Laplace à  $n$  variables - de même que par exemple on parle de coefficients de Fourier dans des cas que Fourier n'avait pas prévus.

J'ai écrit à Loève pour lui demander quelques renseignements bibliographiques.

---

<sup>220</sup>Lévy did not know exactly what Slutsky had done. For a certain time after the war the integral that Slutsky discovered in 1928 would be called Intégrale stochastique définie au sens de Riemann m.q. (Stochastic Integral Defined in Riemann's sense in quadratic mean). At least since 1937, Slutsky knew how to give a trajectorial sense to some of these integrals.

<sup>221</sup>Real part of the characteristic function of a random variable with distribution function  $F$  for which the regularity at 0 governs the properties of  $F$ . This function was used in the spectral analysis of turbulence. If  $R(h) = R(t, t + h)$  is the autocorrelation function of the stationary process  $X(t)$ , there exists a measure  $dF$  such that the integral written is exactly  $R(h)$ . Today this measure is called the 'spectral measure of the process.' It has some analytical connections (via the intermediary of an  $L^2$  space with a stochastic spectral measure  $G$  which allows a representation of  $X(t)$  in the form  $\int e^{iat} dG(t)$ ). These questions were taken up again by Lévy in his 1948 book (Lévy 1948).

<sup>222</sup>Amusing reference to the group Bourbaki, which was established in 1935 and continued during the war. Needless to mention, Lévy's approach to mathematics was very different than Bourbaki's.

<sup>223</sup>No indication of place.

Je n'ai pas souvenir que vous m'ayez parlé déjà de Copeland à propos de mes intégrales stochastiques.<sup>224</sup> S'il y a une priorité de lui, il faudra évidemment la mentionner dans mon travail.

Pour la dérivée aléatoire, comme c'est une notion moins restrictive que la dérivée ordinaire, je n'ai pas songé à être choqué du fait qu'elle puisse exister sans qu'il y ait de dérivée ordinaire. Mais je me demandais quels renseignements elle donne, puisqu'elle peut être nulle dans le cas d'une fonction presque sûrement constamment croissante.

Mais je me suis aperçu depuis que cette circonstance paradoxale cesse d'être possible s'il s'agit d'une dérivée aléatoire en moyenne quadratique, c'est-à-dire si l'on a

$$\lim_{h \rightarrow 0} \mathcal{M} \left\{ \left[ \frac{X(t+h) - X(t)}{h} - \dot{X}(t) \right]^2 \right\} = 0.$$

On peut définir de même une intégrale aléatoire m.q. et les relations entre cette intégrale et cette dérivée sont celles de l'analyse classique. Dans ces conditions, ces notions deviennent, à mon avis, très importantes.<sup>225</sup>

Au sujet du P.S. de votre lettre, j'ai envoyé il y a plus d'un an à Wavre un mémoire sur les suites de Rademacher ;<sup>226</sup> je viens d'en corriger les épreuves. Je lui ai envoyé en août un mémoire sur le mouvement brownien, et j'attends avec impatience son accusé de réception.<sup>227</sup>

Quant à mon nouveau travail, destiné à votre collection, il se trouve tout à coup terminé. Voici comment. D'une part, mes idées sur la dérivée aléatoire étant maintenant bien au point, cela a été vite fait de rédiger 9 pages sur cette dérivée et la condition nécessaire et suffisante de son existence, particulièrement simple dans le cas des processus stationnaires au sens de Khintchine.

Par contre mes idées sur les intégrales et équations différentielles stochastiques ne sont pas au point. La question est plus difficile que je ne le pensais d'abord, et c'est peut-être un an aussi bien que 3 mois qu'il me faudra pour arriver à y voir clair. Je me décide donc à faire un tout de ce qui est actuellement rédigé. Je signale dans une introduction ce qui manque pour donner une idée un peu complète de la théorie des processus continus "que j'exposerai dans un mémoire ou dans un autre fascicule". Si par extraordinaire c'était prêt en temps utile, j'en ferai une note annexe.

<sup>224</sup>Lévy sounds very weak. He had forgotten his previous discussions with Fréchet about Copeland. Risking a dangerous journey, he was going to spend several days with his sister (see following letters).

<sup>225</sup>See Letter 46, note 192. Having clarified the difficulties of the a.s. derivative, Lévy shows the advantages of the derivative in quadratic mean.

<sup>226</sup>These series, which use dyadics, were introduced in 1922 by Rademacher. In 1925 Khinchin and Kolmogorov revisited them and found—simultaneously with Steinhaus—an example of a sequence of independent random variables. Lévy knew that in 1934 Wiener used this sequence as a way of integrating deterministic functions with respect to Brownian motion.

<sup>227</sup>See Letter 48, note 211.

A propos des questions que je laisse de côté, je reste convaincu de l'exactitude du théorème dont je vous avais parlé sur l'intégrale  $\int f(t) dX(t)$ , mais je n'arrive pas à le démontrer.

- Je viendrai donc vous remettre bientôt mon manuscrit, si vous êtes encore dans le Rhône.<sup>228</sup> Ce n'est pas que je me fasse des illusions sur la possibilité d'une impression rapide. Mais je pense que, comme directeur d'une collection, vous pourriez le dater à l'arrivée comme le font les périodiques, et j'y attacherais un certain prix. Il doit y avoir pas mal d'idées nouvelles (je ne sais pas au juste ce qui est nouveau) ; mais, sauf le théorème qui a fait l'objet de ma Note de juin, je ne vois rien d'assez sûrement important et nouveau pour faire l'objet d'une nouvelle Note.<sup>229</sup>

En raison des changements à prévoir dans l'organisation de la censure, je ne vois pas d'intérêt à ce que vous transmettiez mon manuscrit à la dite censure, dans l'état actuel. Votre idée de le faire examiner par un de vos élèves est au contraire excellente. Sans parler des initiatives qu'il pourrait prendre, je serais très heureux de lui voir rédiger un résumé en 12 à 15 pages du mémoire fondamental de Kolmogoroff (Math. Annalen, 1931, sauf erreur). Ce serait une note annexe signée par lui. Je ne pense guère pouvoir avant deux mois travailler dans des bibliothèques et faire moi-même ce compte rendu ; mais si vous pensiez que c'est préférable je pense pouvoir le faire avant Noël.

Je demanderai aussi à votre élève de compléter quelques blancs dans ma bibliographie où je mets au crayon les indications suffisantes pour qu'il sache à quels mémoires je pense et en recherche les titres exacts.

Bien entendu si vous pensez pour cela à Ville,<sup>230</sup> je n'ai aucune objection. Si j'ai pu vous paraître admirer peu ses travaux (et en fait ils ne me semblent jamais vraiment originaux, c'est surtout un bon élève) je me rends bien compte qu'il est très sérieux, a un grand sens de la rigueur et connaît à fond les questions dont il parle. J'aurai toute confiance en lui pour revoir mon manuscrit.

Au point de vue des questions de priorité, j'ai une correspondance récente avec Bachelier, qui me dit avoir donné l'équation dite de Chapman dès 1906 dans un mémoire du Journal de mathématiques. Pouvez-vous vérifier si c'est exact,<sup>231</sup> ou le faire vérifier par un élève ?

*[following in left vertical margin: Il m'a donné d'autre part sur le mouvement brownien sur la surface d'une sphère, qui aurait été étudié par Perrin, une indication que j'ai demandée à Loève de vérifier. Correspondance suivie avec Bass, qui*

---

<sup>228</sup>This manuscript was apparently neither completed nor submitted.

<sup>229</sup>The sealed letter delivered by Fréchet to the Academy and opened at the Liberation.

<sup>230</sup>Unlike Fréchet, Lévy did not have a very high opinion of Ville. Lévy only somewhat tardily recognized Ville's role in the opposition to von Mises' collectives. See the various papers on Ville in the issue 5.1 of the Electronic Journal for History of Probability and Statistics (2009).

<sup>231</sup>See next letter.



l'Université. En tout cas, je ne sacrifierai pas la sécurité à la rapidité, puisqu'il n'y a rien d'urgent.

Je ne sais pas trop comment je m'y prendrais pour fondre dans le texte ce que moi ou d'autres tireront du mémoire de K. Même si c'était possible, je vous avoue que ce travail m'effrayerait un peu. Je rédige toujours difficilement, et ai noirci peut-être 500 pages de papier pour arriver à 105 pages dont je suis à peu près satisfait. Une re-fonte serait un gros travail. J'ai seulement modifié quelques lignes de l'introduction pour annoncer une note annexe.

Je vous ai peut-être déjà dit -mais je suis un cas "psychologique" tellement exceptionnel qu'il est difficile que d'autres le comprennent bien- que je ne peux que suivre le cours naturel de mes pensées - qui parfois me jette avec passion sur un problème et d'autres fois me laisse pendant un an incapable de penser à autre chose qu'aux événements. Tout ce que j'ai essayé pour me discipliner, m'obliger à lire un peu systématiquement ce qui devrait m'intéresser, n'a abouti qu'à diminuer mon rendement. Il faut me prendre tel que je suis.

Si je vous dis cela, c'est pour que vous compreniez que je me résigne à publier un exposé qui contient des lacunes. Je crois malgré ces lacunes qu'il est intéressant. Je les signale au lecteur, pour qu'il puisse au moins servir de guide pour une étude plus complète. Mais j'aurais bien du mal à faire un exposé complet, tel que vous pourriez le concevoir.

Ceci dit, j'accepte avec plaisir l'idée de faire moi-même l'analyse du mémoire de K., si vous pouvez me le procurer. Comme j'en ai fait autrefois une première lecture (K m'avait envoyé les épreuves même avant sa publication) je suis à peu près sûr de pouvoir en un ou deux mois faire une analyse convenable. Quant à la question de savoir si j'arriverai à le posséder vraiment, comme c'est le cas pour les questions vers lesquelles mon esprit est naturellement porté, je ne puis le savoir à l'avance.

Ce que vous me dites du mémoire de Copeland, que vous m'avez prêté ne m'étonne pas, bien que j'ai pu l'oublier complètement. En cherchant dans le tréfonds de ma mémoire, j'y retrouve une vague impression visuelle, qui doit se rapporter à cette note, et c'est tout. Je suis très capable de l'avoir lue, un jour où mon esprit était ailleurs, et sans en saisir l'intérêt, et de ne pas m'en être souvenu le jour où cette idée s'est présentée sur mon chemin. Elle ne pouvait d'ailleurs pas ne pas se présenter à propos du mouvement brownien. Ma première idée, et je crois la plus naturelle, a été de voir qu'en particulierisant les points de division, on obtenait une aire, variable aléatoire dépendant d'une loi déterminée. Mais quand j'ai vu que ce n'était pas une aire véritable, il a bien fallu me demander ce qu'elle signifiait. L'intégrale stochastique s'est imposée à moi certainement sans qu'aucune réminiscence y soit pour rien.<sup>236</sup> Par contre je crois que si après avoir acquis cette idée, j'avais eu connaissance de la priorité de Copeland, je ne l'aurais pas oublié.

Je reverrai avec intérêt son article. Cela ne presse pas, puisque l'intégrale stochastique se trouve reléguée dans mon travail au rang des notions mentionnées et

---

<sup>236</sup>Lévy explains his reasoning. It was while giving a sense to the area determined by an arc of a planar Brownian trajectory and its chord ('Lévy Area') that he had the idea for his stochastic integral (see Sect. 8).

non développées. Mais si vous m'envoyez le mémoire de K, cela ne sera pas difficile d'y joindre celui de C.<sup>237</sup>

Je crois que les envois d'imprimés par la poste d'une zone à l'autre ne sont toujours pas possibles. J'espère que vous trouverez des moyens. Veuillez l'envoyer à Montbonnot, à l'adresse habituelle. Je pense y être de retour dans une quinzaine.

Bien cordialement à vous

P.L.

Je réponds à d'autres points de votre lettre.

Dérivée aléatoire m.q. Je n'ai jamais cru l'avoir inventée. Si je me rappelle bien une note de Slutsky, il l'a considérée au moins implicitement, mais s'est surtout intéressé à la notion la plus générale. Ce que je pense dire de nouveau, c'est que la dérivée m.q. est plus intéressante, parce qu'elle comporte un théorème sur la relation entre la dérivée et l'intégrale.

- Pour Bachelier, cela m'étonnerait qu'il m'ait écrit avoir obtenu en 1906 l'équation de Chapman Kolmogoroff, si ce n'était pas vrai.<sup>238</sup> Il faudra le vérifier. Je joins d'ailleurs à mon manuscrit une liste de vérifications à faire.
- Khintchine, dans son livre de 1933, parle du théorème de Laplace-Liapounoff, que Laplace n'aurait indiqué que dans un cas particulier, celui de Bernoulli, si je comprends bien. Cela m'étonne. Je croyais me rappeler que son résultat dépassait beaucoup celui de de Moivre Liapounoff ?

[*in left vertical margin*: Je n'ai aucune objection pour le choix de Loève. Je m'en rapporte entièrement à vous.]

### **Lettre 53**

le 13 octobre 1943

Mon cher Collègue,

Mon neveu a remis mon travail vendredi dernier à M. Boutet,<sup>239</sup> ingénieur en chef des Ponts et Chaussées à Mâcon, qui espère pouvoir vous le faire remettre dans un délai de quinze jours au plus.

Depuis la lecture de Khintchine d'une part, d'autre part une lettre de Loève<sup>240</sup> répondant à une de mes questions, me fait déjà envisager quelques modifications. Je les rédigerai ultérieurement, à loisir. Mais je vous en envoie ci-joint un résumé que je vous serais obligé de joindre au manuscrit pour le cas où ces corrections devraient être faites par quelqu'un d'autre.

Bien cordialement à vous

P.L.

<sup>237</sup>In Letter 50 Lévy appeared to have completely forgotten about Copeland. Fréchet had no doubt refreshed his memory.

<sup>238</sup>Twelve years after Khinchin and Kolmogorov, Lévy at last recognized Bachelier's contributions, and finally reconciled with him. See Sect. 7.

<sup>239</sup>See the preceding letter.

<sup>240</sup>Thus Loève, as well as Bass, remained in contact with Lévy up to the last possible moment. Later Lévy would remember the epistolary presence of those who stood by him through very difficult times.

## Lettre 54

Le 29 novembre 1943<sup>241</sup>

Mon cher collègue,

J'ai bien reçu votre lettre du 9 avec le mémoire de S. Bernstein, dont je vous remercie. Il contient beaucoup d'idées intéressantes, mais pas celle que je craignais au point de vue priorité - Pour ne pas oublier, je commence par vous dire que j'ai cru pouvoir prêter la traduction de K. par W.D.<sup>242</sup> à Bass (à Villerest, Loire). Il vous la renverra directement. Il me reste donc à vous les brochures de Bernstein et de Copeland.

Pour Copeland, je vous concède bien volontiers que sa suite de nombres  $x_n$  ressemble fort à une suite choisie au hasard, et qu'il ne pouvait pas ne pas y penser. Pourquoi ne l'a-t-il pas dit ? Pour une raison bien simple, et qui marque la différence entre lui et moi qui à la réflexion me semble la plus essentielle. Il voulait montrer que dans le cas de Stieltjes, il y a convergence sûre (et non seulement presque sûre) de la moyenne  $\frac{1}{n}(g(x_1) + \dots + g(x_n))$  vers l'intégrale. Il n'a donc pas besoin d'introduire le calcul des probabilités, et mes résultats relatifs à la convergence presque sûre ne comprennent pas le sien. Mais du fait qu'il ne considère que la convergence sûre, et non les convergences presque sûres, ou en probabilité, ou en moyenne quadratique, résulte que lui (comme son titre l'indique bien) ne fait que donner une nouvelle définition de l'intégrale de Stieltjes ; moi j'en réalise l'extension. Et c'est bien la preuve que la différence est essentielle. Et on ne peut pas dire, comme vous avez pu le dire de l'intégrale de Radon, que son principal intérêt provient de ce qu'elle permet l'extension au calcul des intégrales abstraites. Ici, pour réaliser l'extension, il faut modifier la définition.

Ceci dit j'aurais mauvaise grâce, puisque vous le désirez, à refuser de mettre une note disant que vous me signalez la Note de Copeland comme présentant un lien avec mon intégrale stochastique, et précisant l'analogie que vous voyez et les différences que je constate.

Mes idées en 1939, quand je vous en ai parlé, étaient déjà assez nettes. Je n'ai pas pu vous en parler avant octobre 39, date à laquelle mon mémoire sur le mouvement brownien était en Amérique ; c'est aussi, à quelques semaines près la date de ma note à l'Académie. Si mes idées ont un peu évolué récemment, c'est parce que je m'étais surtout attaché à la convergence presque sûre, et que maintenant je m'intéresse davantage à la convergence en mesure qui réalise l'extension maxima. Je n'ai d'ailleurs toujours pas résolu de manière qui me satisfasse le problème des conditions suffisantes pour que cette intégrale stochastique ait un sens.

J'ai vécu depuis trois mois dans des conditions peu favorables au travail. Mais ce qui m'inquiète davantage, c'est que je suis dans une de ces périodes, qui reviennent de temps en temps chez moi, où je n'arrive pas à penser à mon travail ; il faut sans doute que les cellules de mon cerveau se reforment, c'est du moins l'impression que

---

<sup>241</sup>The last of Lévy's wartime letters. Cognizant of the dangers of the most turbulent time of the war, six months before the Normandy landings and the fall of the Vichy regime, he was apparently putting his affairs in order before going into hiding.

<sup>242</sup>Certainly Kolmogorov and Wolfgang Doeblin.



cela me donne, et je ne sais jamais le temps qu'il faudra. Il y a trois problèmes que je voudrais encore résoudre, pour en parler dans la rédaction définitive de mon exposé.

Après trois mois de vie errante, je suis à nouveau installé dans une maison où j'espère passer l'hiver. Grenoble, comme vous devez le savoir, est très agité ; il y a constamment des incidents : attentats, sanctions, repréailles. Ma famille a insisté il y a trois mois pour que je ne reste pas au voisinage immédiat d'une ville si exposée, d'autant plus qu'étant connu comme je l'étais, je pouvais être particulièrement exposé. Au début de 1943, j'ai vu la dispersion de beaucoup de mes amis lyonnais, dont je n'ai souvent plus de nouvelles ; depuis trois mois, l'on fuit aussi Grenoble. Et nous avons dû abandonner des provisions intransportables, pommes de terre, haricots, bois de chauffage, péniblement cultivés et amassés pendant six mois ; je ne sais pas si, cet hiver, le froid ne m'empêchera pas de travailler.

D'autre part, je viens d'apprendre que je ne suis plus professeur à l'Ecole Polytechnique. Au début de novembre, un chèque postal émis pour disposer de mon traitement d'octobre m'est renvoyé comme étant sans provision. J'écris au trésor de l'Ecole pour demander des explications. Je reçois le 24 novembre, copie d'une décision du 30 juin par laquelle l'Ecole me remet le 1er octobre, à la disposition du Corps des Mines ; il ressort d'ailleurs du texte de cette décision que le 29 avril le Gouverneur de l'Ecole savait que je n'étais pas "réinvesti dans mes fonctions" et demandait au Ministre jusqu'à quand il devait me payer !!! Et tout cela est légal, si l'on considère comme légal un statut de 1941 soumettant les professeurs à la réélection décennale, sans tenir compte des droits acquis. Et maintenant que va faire de moi le Corps des Mines, dont j'étais détaché depuis 1913, et qui ne peut me confier des travaux pour lesquels je n'ai plus guère d'aptitude- et à la disposition duquel je ne me mets qu'avec deux mois de retard.<sup>243</sup> Va-t-il me payer ? J'ajoute que ce qui eût été pour moi une catastrophe, ne l'est plus, d'autant plus que j'espère bien que 1944 verra la fin de ces misères.

Bien cordialement à vous.

P.L.

Ecrivez toujours : Paul Piron<sup>244</sup> - Institut Polytechnique - 46 Avenue Félix Viallet - Grenoble

## Lettre 55

Pour Monsieur Fréchet Extrait de Paul Lévy - Variables aléatoires - p.135

Théorème 43,2 - 1° Pour toute suite de variables aléatoires  $S_n$  tendant au sens de Bernoulli vers une limite  $S$ , si  $\mu$  est une valeur d'accumulation pour les valeurs

---

<sup>243</sup>The Ministère des Communications, which governed the Ecole Polytechnique, and the Ministère de la Production Industrielle, which governed the Corps des Mines sent Lévy's dossier back and forth for months. The dossier Lévy (771390, art. 597) held at the Centre des Archives de Fontainebleau gives evidence of the details of this sordid affair. See the thesis of B. Locker (Locker 2001).

<sup>244</sup>As of 1942 Lévy could be reached at Robert Piron's address, under the name Paul Piron (see the letter of November 4, 1942). Letters addressed this way were passed on to Lévy. Later, in 1943, Lévy used the pseudonym Paul Lengé, which was also on his identity card.

médianes  $\mu_n$  des variables  $S_n$ , c'est une valeur médiane de  $S$  (donc si la valeur médiane de  $S$  est unique,  $\mu_n$  tend vers cette valeur).

2° Même dans le cas particulier des séries presque sûrement convergentes à termes indépendants,  $\mu_n$  peut ne pas avoir de limite.

Le résultat du 1° est attribué à Slutsky, sans référence précise. Mes souvenirs ne me disent rien de plus. Peut-être est-ce dans le mémoire de Metron (1925).

En tout cas mon livre est antérieur à la note de Ville sur le même sujet.<sup>245</sup>  
le 2.3.45

P.L.

[*On the bottom, in Fréchet's handwriting: répondu le 5/3 que c'est un théor ≠ de celui de Ville*]

## Lettre 56

Paris - 147 Boulevard Raspail

le 12 mars 1945

Mon cher Collègue,

J'ai été stupéfait en recevant votre lettre. J'ai toujours fait une confusion sur le résultat de Ville dont vous m'avez parlé en 1936, au moment de sa publication.<sup>246</sup>

Je m'en excuse, mais cela ne change pas beaucoup mon opinion sur le peu d'originalité de cette note. La loi forte des grands nombres, c'est-à-dire le théorème suivant : "La loi de la distribution des valeurs trouvées au cours d'expériences indépendantes tend presque sûrement vers la loi de probabilité théorique", était connu depuis longtemps (1917 ou même 1909).<sup>247</sup>

D'autre part, dans mon théorème 43,2, peu importe que les lois considérées soient des lois de répartition théoriques ou des lois de distribution empiriques. L'hypothèse de mon théorème étant vérifiée, sauf dans des cas de probabilité nulle, pour ces lois de fréquence, il en est de même du mien. Compte tenu de la loi forte des grands nombres, le résultat de Ville apparaît donc comme une application de mon théorème 43,2.

Bien entendu, je ne peux pas reprocher à Ville de n'avoir pas connu mon livre à l'époque où j'en corrigeais les épreuves. Mais mon théorème 43,2 n'a jamais été dans mon idée qu'une remarque banale que je n'ai énoncée explicitement que parce que j'en avais besoin. Je ne peux de même considérer le théorème de Ville que comme un corollaire évident de la loi forte des grands nombres. T.S.V.P.

- D'autre part j'ai toujours oublié de vous reparer de *mes tirages à part qui vous ont été rendus pour moi par le fils de M Lebesgue*.<sup>248</sup>

<sup>245</sup>Continuation of the discussion in Letter 31. The two letters following this one, and Letter 61 show that Lévy remained sensitive to questions of who had published first and was not shy about expressing his low opinion of Ville. See Mazliak (2009).

<sup>246</sup>Fréchet had probably alerted Lévy that Ville had announced his theorem on empirical means.

<sup>247</sup>The strong law of large numbers appeared in Borel, Emile (1909) Les probabilités dénombrables et leur application arithmétique. Rend. Circ. Mat. Palermo 26:247–271 and in Cantelli, Paolo (1917) Sulla probabilita come limita della frequenza. Rend. Accad. Lincei 26:39–45. See the following letter.

<sup>248</sup>Lebesgue died in July 1941. His son, who sorted through his father's papers, probably tried to contact Lévy via Fréchet.

Pourriez vous me les apporter à l'AFNOR<sup>249</sup> lundi prochain, ou les remettre à Fèvre<sup>250</sup> à qui je les réclamerai ?

Bien cordialement à vous P.Lévy

[*In left margin in Fréchet's handwriting: répondu le 23 mars 45*]

## Lettre 57

Montbonnot - le 23 avril 45

Mon cher Collègue,

La date de 1909, que j'ai mentionnée, est celle du mémoire de Borel.<sup>251</sup> Je trouve qu'on est injuste pour lui en n'appelant pas la loi forte des grands nombres loi de Borel-Cantelli.

Pour ma part, j'ai appris les premiers éléments du calcul des probabilités au printemps de 1919, grâce à Carvallo<sup>252</sup> qui m'avait demandé de faire trois conférences sur ce sujet aux élèves de l'Ecole Polytechnique. Je suis d'ailleurs arrivé en trois semaines à des résultats nouveaux. Mais jamais je ne revendiquerai pour mes travaux de calcul des probabilités une date antérieure à 1919. Je peux même ajouter, et je l'ai dit un jour à M.Borel, que je n'ai guère vu qu'en 1929 l'importance des problèmes nouveaux que posait la théorie des probabilités dénombrable.<sup>253</sup> Mais j'étais préparé par le calcul fonctionnel à l'étude des fonctions d'une infinité de variables et beaucoup de mes idées sur l'analyse fonctionnelle sont devenues sans effort des idées applicables au calcul des probabilités.

C'est vrai que notre discussion est stérile et je ne l'aurais pas rouverte si je n'avais pas cru trouver un fait nouveau, en voyant un théorème que je ne me rappelais même pas avoir mis dans mon livre. Je ne le regrette pas, puisque cela vous a permis de rectifier une erreur que je faisais sur l'énoncé même du théorème de Ville.

Bien cordialement à vous P.Lévy

Je serai à Paris dimanche prochain.

---

<sup>249</sup>In 1926 the Association Française de Normalisation was formed. From 1938 until his death in 1948 the Association was directed by Ernest Lhoste, former captain "inspecteur des élèves" at the Ecole Polytechnique. It was Lhoste who, in 1919, set Lévy on course for the discovery of the stable distributions (see Sect. 5). In March 1944 Lhoste created a *Commission de terminologie statistique* (*Commission on Statistical Terminology*) at AFNOR. This commission met more than fifty times between 1944 and 1948. It was charged with standardizing the vocabulary and notation of statistics, both theoretical and applied. Fréchet played the dominant role at these meetings. At the Liberation, Lévy, Darmon, and a few others were named permanent members of the commission. It is no surprise that Lévy participated in the commission, given his interest in terminology, often manifested in his correspondence with Fréchet. See N. Hadjaji Seddik-Ameur, *Les tests de normalité de Lhoste (Lhoste's Criteria for the Standardization of Language)*. *Math. Sci. Hum.* 162:19–43 (2003).

<sup>250</sup>At the Institut Henri Poincaré, the Fèvre family 'dynasty', father and son, was employed as janitor, guard, and then assistant at the library.

<sup>251</sup>The article by Borel cited in note 247.

<sup>252</sup>Emmanuel Carvallo, then *Directeur des Etudes* (*Director of Studies*) at the Ecole Polytechnique. His memoir on the theory of probability was published in 1912 by Gauthier-Villars.

<sup>253</sup>*Sic.*

**Lettre 58**

Paris - le 28 octobre 45

38 Av. Théophile Gautier Auteuil-03-09

Mon cher Collègue,

Je veux d'abord vous dire combien j'ai admiré le travail considérable que vous avez fourni (avec Denjoy, Valiron, et Bouligand)<sup>254</sup> pour rendre compte de l'effort des mathématiciens français ou vivant en France, pendant la guerre, et pour donner à chacun la place qui lui revient. En ce qui concerne mes travaux, j'ai particulièrement admiré que vous ayez pu parler avec précision de ceux relatifs au mouvement brownien à plusieurs paramètres, alors que, il me semble, vous n'avez pas assisté à la conférence où j'en ai parlé pour la première fois.<sup>255</sup>

Au sujet de cette question, j'ai commis jeudi une erreur assez grave. Si  $X(A)$  est connu sur une surface fermée de dimensions bornées,  $\mu = E\{X(A)\}$  devient harmonique à l'intérieur quand le nombre  $n$  des dimensions de l'espace augmente indéfiniment, tandis que la part du hasard tend vers zéro. J'ai vécu 48 heures dans l'illusion que, même pour  $n$  fini,  $\mu$  était une fonction harmonique de  $A$ . Je crains de l'avoir dit. Or cela est faux.<sup>256</sup>

D'autre part, à propos de mes intégrales stochastiques, vous avez indiqué une idée de Slutsky que je ne connaissais pas, et qui est très voisine de la mienne (beaucoup plus, à mon avis, que celle de Copeland).<sup>257</sup> Pouvez me donner la référence ?

Je connaissais, de Slutsky, l'intégrale aléatoire d'une fonction aléatoire, mais je ne savais pas qu'il avait introduit le hasard dans les  $\xi_i$ , bien qu'en fait cela aille moins loin qu'en l'introduisant dans les  $x_i$ , c'est une priorité à mentionner dans mon exposé.

Enfin un mot pour Loève. Vous l'avez cité plusieurs fois. Mais je me demande si vous connaissez ses Notes du début de 1945 (janvier et mars, je crois) sur les fonctions aléatoires et spécialement les fonctions stationnaires. Ne devraient-elles pas être citées à côté des recherches de Blanc-Lapierre ?<sup>258</sup>

<sup>254</sup>The 'Congrès de la Victoire' was held in Paris, October 20–26, 1945, under the auspices of the Association Française pour l'Avancement des Sciences. Its purpose was to present recent contributions of French mathematicians, especially those who had to remain in hiding during the Occupation. Fréchet gave a talk published as Fréchet, Maurice (1947) Les contributions françaises récentes au calcul des probabilités et à la statistique mathématique. Inter Rech. Math. Suppl. 9:107–210.

<sup>255</sup>On March 26, 1945 Lévy published a note on Brownian motion with  $n$  parameters ((1945) CRAS 220:420–422). He developed the theory further in Chap. VIII of his 1948 book (Lévy 1948). In these two publications he referred to "Laplacian" variables, whereas in his 1955 article on the same subject (Proceedings 7th Symposium on Applied Mathematics. Am. Math. Soc., pp. 1–20 (1955)), he used the classic designation of "Gaussian" random variable.

<sup>256</sup>This is an odd remark, because in his note to CRAS of March 26, 1945, Lévy himself explained that only in finite dimension does one have a harmonic function.

<sup>257</sup>On Copeland, see Letter 43 and Sect. 8.

<sup>258</sup>Trained as a physicist, André Blanc-Lapierre (1915–2001) defended a second thesis on electrical fluctuations due to electronic structure, directed by Fortet. From 1943 to 1945 he published three notes in CRAS, inspired by the study of shocks in an amplifier. This example is further developed in Blanc-Lapierre and Fortet (1953). It led Blanc-Lapierre to an interest in random processes of second order. An interesting interview of Blanc-Lapierre by J. Ségala may be found at

Bien cordialement à vous

P.Lévy

### **Lettre 59**

Paris- 5-11-45

Mon cher Collègue - Je viens d'apprendre, par un ami qui revient des Indes, que le professeur Dr.F.W.Lévi<sup>259</sup> (allemand réfugié depuis 1933, qui a à peu près mon âge) (5 Bright Street, P.O.Ballygange, Calcutta), cherche à rétablir le contact France-Indes.

Je lui écris et demande à G.V. et à Hermann de lui envoyer, si possible, des catalogues.

Mais des tirages à part et éventuellement d'autres lettres que la mienne seraient bienvenues. Veuillez le dire à l'occasion aux collègues que vous verrez.

Il voudrait avoir un aperçu de ce qu'on a fait en France depuis 1939. J'avoue ne pas avoir essayé de donner dans une lettre idée d'un effort qui vient d'être résumé en 5 conférences. Si vous lui donnez quelques indications, elles seront bienvenues.

F.W.Lévi est un spécialiste de l'algèbre. Mais il s'agit pour lui d'un programme d'achat pour les bibliothèques des Indes, pour l'ensemble des mathématiques.

En raison d'un très prochain congrès à Delhi,<sup>260</sup> au début de décembre, quelques indications par avion seraient utiles.

Bien cordialement à vous

P.Lévy

### **Lettre 60**

[*In left margin, in Fréchet's handwriting: rép. le 4 avril*]

Paris - 3-4-47

Mon cher collègue,

Votre carte arrive trop tard. Vous ayant initialement appelé au téléphone, sachant Montel absent, j'ai rapporté les épreuves à Gauthier Villars.

---

[www.mpiwg-berlin.mpg.de/staff/segal/thesis/thesehtm/entret/blancclap.htm](http://www.mpiwg-berlin.mpg.de/staff/segal/thesis/thesehtm/entret/blancclap.htm). Meanwhile on March 5 and 19, 1945, Loève had published two notes to CRAS on covariance functions of a random function, which are related to Blanc-Lapierre's work. Lévy apparently wanted to support Loève and asked him to write a chapter on this subject to complement Lévy's 1948 book (Lévy 1948).

<sup>259</sup>F.W. Lévi fled to India to escape the Nazis. During the war, he taught at the Tata Institute in Mumbai and engaged in mathematical research at a high level. He remained in India after 1945, and was one of the founders of the Tata Institute. His work on Lie groups with associative Lie brackets and on certain tensorial algebras is still inspiring mathematicians. The origin of the term 'combinatorics' is attributed to him.

<sup>260</sup>Hadamard went to India in 1947 and was received by Nehru. The great traveler Fréchet, who kept up relations with Indian statisticians, went to India in 1952. A record of Fréchet's voyage is held by the archives of the Academy in the 'Cartons Fréchet.'

Il m'avait semblé aussi que Pollaczek<sup>261</sup> avait fait trop de corrections. J'en ai ajouté 2 ou 3, mais ne pouvant pas tout relire, je n'ai fait dans ce sens aucun changement qui risque d'avoir des répercussions sur d'autres passages.

Je vous renvoie le papier de Pollaczek avec mes observations. La remarque que  $e^{z \cos t}$  n'est pas une fonction caractéristique étant incontestablement exacte, je n'ai pas eu d'autre solution que de supprimer tout le n° 17 (soit 14 lignes [Quelque soit  $z \neq \theta$ , on n'a pas pour tout  $t$  réel,  $e^{z \cos t} \leq 1$ ]).<sup>262</sup>

A part cela, il me semble aussi qu'il faut supprimer au pt.7 une formule sans rapport avec le reste. Le reste n'est que détails. J'ai supprimé un barbarisme que P. n'avait pas remarqué (banaux).

P. a oublié de mettre les renvois dont les n°s manquent. Je le regrette, parce qu'il avait sûrement fait une étude du mémoire assez sérieuse pour le faire sans peine.

Après le 18 mai, si vous en avez besoin, je pourrai sans doute faire pour revoir la mise en page un travail plus sérieux que celui que je viens de faire.

Bien cordialement à vous

P.Lévy

Je préviens Lefort<sup>263</sup> d'attendre vos instructions. Je pense qu'il n'aura encore rien fait.

## Lettre 61

Nancy - le 20 août 47

Mon cher collègue,

Votre lettre m'arrive chez mon gendre,<sup>264</sup> que je quitte demain, et qui est lui-même en plein déménagement, ce qui fait que je n'ai pas mon livre sous la main (bien qu'il le possède) et ne peux guère le consulter en temps utile pour vous répondre. D'ailleurs cela n'a pas grande importance. Permettez-moi de vous dire franchement qu'il y a des détails auxquels je n'arrive pas à attacher autant d'importance que vous le faites.

Ce qui est important, c'est la découverte de la loi forte des grands nombres (Borel, 1909 ; Cantelli 1917). Je n'ai rien à revendiquer pour moi.

Je l'ai énoncée dans ma précédente lettre sous la forme qui me paraît le plus commode, et que vous me rappelez : "Dans le cas des épreuves répétées, la loi de ré-

---

<sup>261</sup>An Austrian born in Vienna, Felix Pollaczek (1892–1981) was an expert in queuing theory. After finishing his studies he was mobilized in World War I. In 1920 he obtained a doctorate in mathematics (in number theory) at the University of Berlin. He remained in Berlin until 1933, primarily doing research for the post office. With the rise of the Nazis, he emigrated to Paris, then went to Czechoslovakia. He returned to Paris in 1938, managing to escape persecution during the Occupation. At the Liberation he obtained the position, however precarious, of Maître de Recherches in the CNRS. He became a French citizen in 1947.

<sup>262</sup>The closing parenthesis is missing.

<sup>263</sup>Guy Lefort (1910(?)–1979) was a student of Georges Darmois and acerbic assistant to Zaman-ski at the Faculté des Sciences de Paris. Assistants were expected to proofread the books written by their professors. Lefort was later named professor of statistics at the Institut National de l'Agronomie in Paris.

<sup>264</sup>In 1947 Laurent Schwartz was named professor in Nancy.

partition expérimentale tend presque sûrement vers la loi de probabilité théorique”. Si ce n’est pas contenu dans l’énoncé initial de Borel, cela s’en déduit aisément ; malgré cela je conçois que vous attachiez de l’importance à savoir qui l’a énoncé le premier explicitement. Mais puisque vous la citez dans un livre contemporain du mien, en l’appelant th. de Glivenko Cantelli,<sup>265</sup> c’est que vous avez une référence antérieure, à laquelle mon livre de 1937 ne peut rien changer.

Ce que je peux vous dire, c’est que je le savais depuis longtemps. Si vous avez sous la main le *Bulletin de la Société math de France 1929*,<sup>266</sup> peut-être trouverez-vous une précision à ce sujet. Dans mon *mémoire sur les fractions continues*, j’énonce et j’applique la loi forte des grands nombres sous une forme qui contient peut-être celle-là\*. J’en reparle dans le dernier chapitre de mon livre de 1937 en disant qu’à ce moment j’avais cru pouvoir considérer comme connu un énoncé qui ne l’était peut-être pas, et que c’est en 1935 (dans des travaux sur la loi forte des grands nombres) que je suis revenu sur ce sujet. Voilà ce que je me rappelle.

Je suis à peu près sûr qu’il n’y a rien sur ces sujets dans mes travaux antérieurs à 1929.

Ceci dit, je ne suis on ne peut plus d’accord avec vous sur le fait que pour le théorème de Ville il suffit d’invoquer le th. de Glivenko-Cantelli. C’est ce que, en ce qui me concerne sans savoir le nom de ce théorème, Darmois et moi savions depuis longtemps quand vous avez parlé du th. de Ville, et je crois qu’en effet j’ai pu vous faire remarquer que le dit théorème s’étendait évidemment au cas, non d’un rang  $r$  déterminé, mais d’un rang  $r$  tel que  $r/n$  ait une limite déterminée.

J’ai quelquefois le tort de ne pas préciser des choses qui me paraissent évidentes et qui ne le sont pas pour d’autres. J’ai aussi laissé échapper certaines priorités qu’il est difficile de réclamer ensuite. Dans le cas dont il s’agit, la seule chose que je vous aie dite est que je connaissais depuis longtemps le résultat de Ville. Mais par le fait même qu’il est un corollaire évident de celui de Glivenko-Cantelli, je ne prétendais en tirer aucune gloire, ni l’appeler “mon théorème”.

Ceci dit, j’espère que vous pourrez partir l’esprit libre de ces préoccupations, et je vous souhaite un heureux séjour en Amérique.

Bien cordialement à vous.

P.Lévy

\* Il devait s’agir de variables enchaînées. Mais ce que je me demande c’est si je ne donne pas là un énoncé qui appliqué aux variables indépendantes donne celui qui vous intéresse.

## Lettre 62

Paris - 23 octobre 49

---

<sup>265</sup>Lévy, clearly annoyed, is being somewhat underhanded. As he himself said in Letter 32, the essential point of the Glivenko-Cantelli theorem concerns the uniform convergence of empirical distribution functions.

<sup>266</sup>Lévy, Paul (1929) Sur les lois de probabilités dont dépendent les quotients complets et incomplets d’une fraction continue. Bull. Soc. Math. Fr. 57:178–194. In this article, Lévy proves a strong law of large numbers for independent variables as well as variables in chain.

Mon cher Collègue,

J'ai lu (ou relu ?) avec intérêt votre mémoire de 1915 sur les fonctionnelles bilinéaires.<sup>267</sup> Il y a en effet quelque chose à corriger dans mes Leçons d'analyse fonctionnelle, p. 62, et je suis rétrospectivement un peu confus de n'avoir en 1922 ni connu ni retrouvé votre résultat de 1915.

Dans la nouvelle édition, je remplacerai toute la première partie de mon livre par un résumé de peut-être 50 pages. Il est actuellement chez G.V., et j'avoue ne pas me rappeler ce que j'y ai mis à propos des fonctionnelles bilinéaires. En tout cas maintenant que j'ai bien compris la question je ferai à la correction des épreuves les corrections nécessaires.

Cette question est liée à une autre pour laquelle vous pourriez peut-être m'aider. Soit  $f(s, t)$  à variation bornée en  $t$ . Il s'agit de savoir à quelles conditions l'opération

$$\varphi(s) = \int_0^1 \psi(t) d_t f(s, t)$$

fait à toute fonction continue  $\psi(t)$  correspondre une fonction continue  $\varphi(s)$ .\*

J'ai mis des bêtises à ce sujet dans mon livre, p.65. Je ne sais plus ce qui m'a conduit à écrire ce que j'ai écrit ; c'est complètement faux.<sup>268</sup>

Cela m'a été signalé vers 1930 par A.Fouillade,<sup>269</sup> un élève de Bouligand.<sup>270</sup> J'ai alors réfléchi à la question, et ai eu la chance de trouver une solution très simple ; chacune des deux parties de la démonstration tenait en quelques lignes, sans rien de vague et sans recours à des lemmes antérieurs.

Je n'ai pas dû vous en parler. Mais j'ai communiqué cette solution à Fouillade, et aussi à Fr. Riesz, qui m'a répondu que je trouverais la solution de ce problème dans son mémoire de 1910,<sup>271</sup> mais qu'il me félicitait de ma solution, beaucoup plus simple et élégante.

Je n'ai rien publié à cette époque, pensant attendre la 2<sup>ième</sup> édition. Or je n'arrive pas à retrouver ma solution (ni l'énoncé ni la démonstration). J'ai écrit à Riesz, qui se rappelait bien que je lui avais écrit sur ce sujet, mais n'a pas retrouvé ma lettre

<sup>267</sup>Fréchet, Maurice (1915) Sur les fonctionnelles bilinéaires. Trans. Am. Math. Soc. 16:215–234 extended Riesz's representation theorem for linear functionals to the bilinear case. In the first version of his 1922 book (Lévy 1922, p. 62), Lévy seemed to assert, incorrectly, that Fréchet's result is immediate.

<sup>268</sup>Once again, Lévy had too hastily come up with a geometrical condition which appeared to him sufficient.

<sup>269</sup>André Fouillade defended a thesis in 1937 at Poitiers on the iteration of linear functional substitutions.

<sup>270</sup>Georges Bouligand (1889–1979) did significant research in potential theory and analytic geometry. See Sect. 6.

<sup>271</sup>It is not at all clear that Riesz's paper (Riesz, Frygies (1911) Sur certains systèmes singuliers d'équations intégrales. Ann. Sci. Ec. Norm. Supér. 28:33–62) even addresses this problem. More likely, Riesz told Lévy that he thought the ideas in this article could be used to solve the problem.



(il y a de cela peut-être 3 ans). La trace de Fouillade était perdue. Bouligand vient de me dire qu'il l'avait retrouvée, et je lui écris. Mais il ne fait plus de science et je doute qu'il retrouve ma lettre, après 20 ans, ou se souvienne de son contenu.

La situation actuelle est donc la suivante : je sais qu'il y a quelque chose de beaucoup plus simple que ce qui est dans Riesz, et je serai reconnaissant à qui m'aidera à le retrouver.

Si j'ai une réponse satisfaisante de Fouillade, je vous le ferai savoir.

Bien cordialement à vous

P.Lévy.

P.S.- Des circonstances du même ordre me font être sûr qu'il y a une démonstration élémentaire du théorème des nombres premiers, beaucoup plus simple que celle qu'Erdős nous a exposée l'année dernière,<sup>272</sup> 6 pages de la grande écriture de Landau, d'une lettre qu'il m'a écrite en 1907. J'ai lu la lettre, sans difficulté, mais sans voir l'idée directrice. L'idée que Landau mourrait sans avoir publié sa démonstration ne m'avait pas traversé l'esprit. C'est cependant ce qui est arrivé. Et sa lettre a disparu- sans doute dans le pillage de mon appartement en 42 (mais je l'avais déjà en vain cherchée avant la guerre, en apprenant la mort de Landau). Et je n'ai pas de souvenirs suffisants pour reconstituer la dite démonstration.

Pólya refuse de croire à cette histoire : mais je n'invente rien. Maintenant il est possible qu'il y ait eu une faute qui m'ait échappé, et que Landau s'en soit aperçu en revoyant sa démonstration. Mais il est alors étrange qu'il ait commencé par ces mots, dont je suis sûr "Le théorème des nombres premiers n'est pas aussi difficile à démontrer que vous le pensez." Cela ne semble pas indiquer-surtout chez Landau- une démonstration récemment découverte et non revue avec soin.

[*Note in margin and disposed vertically*: \*Naturellement cela se ramène à la condition de convergence, pour toute fonction continue  $\psi(t)$ , de la suite  $\int_0^1 \psi(t) df_n(t)$ , pour laquelle vous donnez, d'après Riesz, une condition nécessaire, mais non suffisante (p.217 de votre mémoire de 1915).]

### Lettre 63

Paris -38 Av. Théophile Gautier

le 23-1-50<sup>273</sup>

<sup>272</sup>Erdős, Pal (1949) On a new method in elementary number theory which leads to an elementary proof of the prime number theorem. Proc. Natl. Acad. Sci. USA 35:374-384.

<sup>273</sup>Lévy and Fréchet returned to some of their prewar work. Lévy would soon discover his area law and rewrite his 1922 book *Leçons d'analyse fonctionnelle*. He consulted Fréchet about certain delicate points in analysis. His book was almost completely rewritten, and appeared under a new title, *Problèmes concrets d'analyse fonctionnelle*. Only the earlier Chap. II (on Lévy's work from before World War I) remained intact. Generalities about measure theory were removed from the new edition. Since the theory of analytic functionals had undergone considerable development, Lévy had Franco Pellegrino of the Italian school give an overview of the work begun in 1922 and developed since then by Luigi Fantappie. Fréchet began to work on a new edition of his fascicule in Borel's series *Traité du calcul des probabilités et ses applications*, called *Recherches théoriques modernes sur le calcul des probabilités*, and many remarks concern it in this year's correspondence. The edition of 1952, like that of 1938, included a six page note by Lévy on distance between random variables, and between probability distributions.

Mon cher collègue,

Quand j'ai écrit, dans le passage que vous trouvez peu clair "rien ne distingue ces variables les unes des autres",<sup>274</sup> c'est un peu comme si je disais "je mets des boules identiques entre elles dans une urne ; rien ne les distingue ; elles ont tout de même chacune son identité".

Maintenant cet alinéa de 7 lignes n'apporte rien d'essentiel. Si vous croyez que c'est plus clair en le supprimant, je veux bien. Cela s'enchaînera très bien sans lui.

Naturellement, si la pagination du livre est changée, il faudra corriger les renvois que j'y fais (p.291, l.4 ; même page l.5).

A la p.292, l.1, au lieu de "seul intéressant" il vaudrait peut-être mieux mettre "surtout intéressant". (Je dis cela pour ne pas paraître trop tranchant ; au fond mes idées n'ont pas changé<sup>275</sup>). Enfin même p.292 l.45, j'avais mis "p.84 et suiv. ; p.87". Cela voulait dire que dans le mémoire cité qui commence p.84 il faut chercher la p.87. Mais est-ce clair ? On peut peut-être supprimer la 1ère indication et laisser seulement "p.87".

Je vous remercie de me signaler le mémoire de Mark<sup>276</sup> en relation avec mon mémoire des Comm.math.helvetici.<sup>277</sup> Je n'avais pensé à aucune généralisation et même votre lettre m'a d'abord surpris. Mais en revoyant mon mémoire je vois qu'une des deux démonstrations du théorème fondamental s'applique presque sans changement aux autres lois stables.

Vous vous rappelez peut-être mon mémoire sur le mouvement brownien plan, où j'ai étudié l'aire  $S$  comprise entre un arc  $AB$  de la courbe du mouvement brownien plan et sa corde.<sup>278</sup> Je n'avais pas pu déterminer sa loi de probabilité. Je viens d'y

<sup>274</sup>The expression in quotes is indeed in Lévy's note on page 291 of the 1938 edition cited above, at the end of a long note which was included by Lévy on Fréchet's requirement for readers of the first edition to clarify certain questions about topological notions and the theoretical formulation of probability. The citation of Fréchet's booklet in quotes in the letter refers to the distances introduced by Lévy to define convergence in distribution, which induce a metric structure on the distribution functions but do not distinguish two variables with the same law. The comments by Lévy in the early edition of the book were perhaps necessary, given the shifting understanding (even by Fréchet) of the foundations of the theory of probability. By 1950 they became less important.

<sup>275</sup>On the cited page, the only interesting and practical case is that of random variables that are functions of a denumerable infinity of independent random variables, resulting in identifying the space of random variables to the space of measurable functions, and, in practice, in considering as probability space the interval  $[0, 1]$  equipped with the Lebesgue measure. This is the only probability space  $\Omega$  that Lévy ever made use of.

<sup>276</sup>Kac, Mark (1947) Random walks and the theory of Brownian motion. Am. Math. Mon. 54:369–391. Kac (1914–1984) was born in Poland and as a student of Hugo Steinhaus, defended his thesis in Lvov in 1937. He became an American citizen in 1943. The reference to Kac's memoir is the prelude to a discussion about random (principally Markov) processes and work by Doebelin that is developed in the following letters.

<sup>277</sup>Lévy, Paul (1944) Un théorème d'invariance projective du mouvement brownien. Comment. Math. Helv. 16:242–248. This article was mentioned in the letter of September 9th, 1943.

<sup>278</sup>See footnote 160. In this article, Lévy gave the partial differential equations satisfied by the distribution function of the pair consisting of the area and the polar angle (see Sects. 8 and 7). Lévy

arriver. C'est (à un facteur constant près) la loi de fonction caractéristique  $\frac{1}{\text{ch } z}$ .<sup>279</sup> La loi conditionnelle relative au cas où les points  $A$  et  $B$  coïncident est celle de f.c.  $= \frac{z}{\text{sh } z}$ . Ces lois, et plus généralement la loi de f.c.  $= \frac{1}{\alpha^2 \text{ch } z + \beta^2 \frac{\text{sh } z}{z}}$  ( $\alpha^2 + \beta^2 = 1$ ) me semblent, à cause au moins de cette application, dignes de retenir l'attention. Pensez-vous qu'elles aient déjà été considérées ?<sup>280</sup>

Bien cordialement à vous

P.Lévy.

## Lettre 64

Paris - le octobre 50<sup>281</sup>

Mon cher Collègue,

Vos observations m'ont paru très judicieuses, et j'en tiens le plus grand compte. Je ne vois pas très bien comment je pourrai introduire dans la 1ère partie la notion d'espace séparable ; mais j'espère y arriver dans la 3ième partie. Pour le reste, je ne vois pas de difficulté du même genre. J'ajoute un n° 11 bis<sup>282</sup> pour y parler de votre mémoire de 1937.<sup>283</sup> Mais il y a dans ce mémoire quelque chose qui ne me paraît pas exact. . .

A moins qu'il n'y ait un malentendu. Quand vous parlez de votre définition de la différentielle, où intervient la distance, je comprends, sauf mention expresse du contraire, qu'on peut se placer dans n'importe quel espace avec une définition appropriée de la distance. Or il y a, p.244–245 (phrase à cheval sur deux pages) une phrase qui ne me paraît exacte que s'il s'agit de l'espace des fonctions continues et

---

had just discovered the characteristic function  $F$  of his area. To find it, he abandoned the partial differential equation approach in favor of an application of the Paley-Wiener series, described in his note to CRAS, January 30, 1950. That note does not demonstrate the equivalence between the definition of Lévy's area employing his stochastic integral and the definition he used with integrals of Wiener's type. Lévy later proved this equivalence and presented it at the Berkeley Symposium on Probability and Statistics in 1950. In the second (1965) edition of *Processus stochastiques et Mouvement Brownien (Stochastic Processes and Brownian Motion)* Lévy reviewed the various methods for determining  $F$ . The Paley-Wiener series was replaced by an application of the Cameron-Martin integral. Yor has given an elegant and modern proof of Lévy's formula in 1980.

<sup>279</sup>If the time interval is  $[0, 2]$ ,  $\Phi(z) = \frac{1}{\text{ch } z}$ .

<sup>280</sup>This question sounds strange, since Lévy had done so much work on characteristic functions as Fourier transforms. The result quoted appeared for the first time in a 1937 thesis defended by Geza Kunetz, a Hungarian student of Darmois. Fréchet apparently did not remember the result, even though he was a member of the thesis committee. Lévy was informed of it by Darmois. This evident from a footnote in *Propriétés des lois dont les fonctions caractéristiques sont  $\frac{1}{\text{ch } z}$ ,  $\frac{z}{\text{sh } z}$ ,  $\frac{1}{\text{ch}^2 z}$* . CRAS. February 27, 1950, co-authored by Lévy and Bass (the latter having performed most of the analytical calculations).

<sup>281</sup>The day of the month was not specified. This letter certainly precedes that of November 28, 1950, as Lévy mentioned reflecting on a memoir by Fréchet that he announced to take back to the IHP, in the November letter in which he also finished the current discussion.

<sup>282</sup>See Sect. 4 and the following letter.

<sup>283</sup>(1937) *J. Math. Pures Appl.* 16:233–250. See Sect. 4.

d'une distance liée à la convergence uniforme (et je ne vois cette restriction nulle part).<sup>284</sup>

Soit en effet la fonctionnelle  $x'(\tau)$  [ $x'$  dérivée de  $x$  ;  $\tau$  valeur particulière de  $t$ ]. Si  $x(t) = x(t, \lambda)$ , il est bien évident que, même si  $x'(t)$  existe et est continue en  $\lambda$  et  $t$ ,  $\frac{dx}{d\lambda}$  peut exister sans que  $\frac{dx'(\tau)}{d\lambda}$  existe. Donc  $x'(\tau)$  n'a pas une variation au sens d'Hadamard (généralisé par vous<sup>285</sup>).

D'autre part, dans votre sens, n'importe quelle fonctionnelle linéaire, définie et continue dans un espace fonctionnel convenable avec une définition convenable de la distance, est sa propre variation.

Donc  $x'(\tau)$  est différentiable à votre sens et non à celui d'Hadamard. C'est en contradiction avec votre affirmation.

Ma conclusion est :

Votre objection à ma définition est fondée. Mais on y échappe en ajoutant à ma définition une condition de Lipschitz

$$|U[y(t)] - U[x(t)]| \leq K \cdot \text{distance de } x(t) \text{ et de } y(t).<sup>286</sup>$$

Mais alors on introduit une distance ; et il vaut mieux revenir à votre définition, plus simple et très naturelle.

Je m'étais demandé un instant si ma définition ainsi modifiée n'était pas équivalente à la vôtre. Il n'en est rien. Ainsi, à l'origine, la fonction

$$u(x) = x^2 \sin \frac{1}{x^2}$$

est différentiable, mais ne vérifie pas une condition de Lipschitz au voisinage de ce point puisque sa dérivée n'y est pas bornée. Pourtant en tout point, dans votre sens, elle a une différentielle en tout point<sup>287</sup> (mais sans uniformité).

Quant à la différentielle d'Hadamard généralisée, je la mentionne aussi sans la retenir, parce qu'avec cette définition une fonctionnelle linéaire pourrait n'avoir pas de variation.

Si vous n'êtes pas d'accord avec les remarques qui précèdent, veuillez me prévenir d'urgence. Je pensais rapporter demain à G.V. les épreuves de la troisième partie ; mais j'attendrai votre réponse 2 à 3 jours.

Bien cordialement

<sup>284</sup>At the spot Lévy indicated, Fréchet had written:

“Tout d'abord la notion de continuité imposée suppose qu'on a affaire à un champ fonctionnel déterminé  $C$ , où la limite d'une suite de fonctions a une signification déterminée (suivant les cas, on supposera qu'il y a convergence uniforme, en moyenne quadratique etc. . .) ( . . . ) D'autre part il faudra supposer que  $f'_\lambda(t)$  appartienne aussi à  $C$ . Il suffira pour cela de supposer que la convergence de  $\frac{f(t, \lambda) - f(t, 0)}{\lambda}$  soit de nature admise dans  $C$ .”

At a first reading, Lévy might well have had difficulty accepting such a high level of generality.

<sup>285</sup>Lévy had not read Fréchet with sufficient care. See Sect. 4 for remarks on differentiability in the “generalized Hadamard” sense.

<sup>286</sup>This sort of Lipschitz-type condition is intended to reconcile the Lévy, Fréchet and Hadamard definitions. However, it does not help much, since it limits the generality of Lévy's differential.

<sup>287</sup>*Sic.*

P.Lévy

Madame Doblin<sup>288</sup> ne sait pas quoi faire de la bibliothèque de son fils.<sup>289</sup> Si vous avez quelque suggestion à lui soumettre, peut-être est-elle encore à Paris (rue St Jacques, hôtel en face de la Faculté de Droit).

J'ai pris le parti d'éviter le "Monsieur" quand la personne citée intervient comme auteur. Ainsi je dis le th. de J. Hadamard. Mais à l'occasion je dirai : c'est grâce à une observation de M.Hadamard que...

## Lettre 65

Paris - 28-11-50

Mon cher Collègue,

Je vous confirme que nous sommes bien d'accord.

J'avais eu tort de ne pas faire attention aux explications qui suivent dans votre mémoire de 1937<sup>290</sup> la définition en italiques de la p.244. Au fond, il y a 3 définitions différentes suivant qu'on ne regarde que la définition en italiques, ou qu'on lise ensuite les deux phrases suivantes (sans compter celle entre (...)), ou qu'on aille jusqu'au bout de l'alinéa, en se plaçant dans le cas où la condition suffisante indiquée est réalisée.

J'ai refait mon n° 11 bis. J'ai donné brièvement 7 définitions possibles, mais pour conclure que tout cela est sans doute plus compliqué qu'il n'est nécessaire, et qu'il vaut mieux s'en tenir à votre première définition, au fond très naturelle.<sup>291</sup>

Si elle n'est pas vérifiée - du moins si sa seconde condition n'est pas vérifiée, c'est qu'on peut trouver  $K$  fixe et, pour  $\rho$  arbitrairement petit, un  $\delta x(t)$  tel que  $r[x, \delta x(t)] < \rho$  et

$$|U[x + \delta x(t)] - U[x(t)] - \delta U| > Kr[x, \delta x(t)]$$

Dans les cas de ce genre, j'aurai toujours l'impression que  $\delta U$  ne mérite pas d'être appelé différentielle.

Peut-être suis plus royaliste que le roi ? Je n'ai d'ailleurs pas mis cette dernière remarque pour ne pas trop allonger mon exposé.

Bien cordialement

P.Lévy

Cette lettre n'ayant plus rien d'urgent, je me contenterai de la déposer à l'Institut H.P. - en rendant le Journal de Math. - 1937 que j'avais emprunté.

<sup>288</sup> *Sic.* Doebelin changed his name to the more French-sounding Vincent Doblin.

<sup>289</sup> As mentioned (see letter of November 4, 1942), Doebelin committed suicide on June 21, 1940, to avoid capture by the Germans. Doebelin's work appears in the letters of 1951.

<sup>290</sup> See preceding letter.

<sup>291</sup> In the 1951 edition of number "11 bis," Lévy wrote: "Il y a une hiérarchie des différentes définitions possibles [;] on peut préférer la définition qui s'applique le plus souvent, c'est la moins restrictive, mais c'est aussi celle dont on peut déduire le moins de conséquences[.] La définition de Gateaux semble trop peu restrictive. Il est naturel d'exiger que la variation soit une fonctionnelle linéaire." In 1951 Lévy stated clearly that on passing to "second order variations" it is Fréchet's definition that 'applies naturally to maxima and minima.' See Sect. 4.

## Lettre 66

Paris - 14-1-51

Mon cher collègue,

J'ai sous les yeux le mémoire de Doebelin dont vous me parlez<sup>292</sup>

Ce que je vois nettement est ceci : si le lemme I s'applique uniformément quand  $\tau$  varie dans un petit intervalle  $(\tau_1, \tau_2)$ , il n'y a plus de difficulté jusqu'à la fin du n° 3.

En effet, d'après la fin du dit n°,  $b$  étant rationnel et  $\rightarrow 0$ ,  $A^{(b\tau)}$  et  $A^{(b\tau_1)}$  ont la même limite  $A$ . Si on pose  $b'\tau_1 = b\tau$ ,  $b'$  est un nombre tendant vers zéro d'une manière quelconque (puisque  $\tau$  est quelconque entre  $\tau_1$  et  $\tau_2$ ), et  $A^{(b'\tau_1)} \rightarrow A$ .

Il n'y a alors aucune difficulté si on suppose les  $a_{ik}(t)$  continus sauf peut-être à l'origine.<sup>293</sup> Mais sans la continuité, je ne vois pas encore clairement ce qu'on peut dire.

Au début du n° 4, je ne vois pas d'où vient la formule initiale. Il n'a dém. . . Excusez ; je confondais  $a_{i,k}(t)$  et  $a_{i,k}(t; \tau)$ . Il faudra que je revoie cela plus posément.

- Comme vous n'avez pas assisté à mes récentes conférences sur les systèmes à une infinité dénombrable d'états possibles, je vous signale qu'un des résultats qui m'a d'abord surpris est le suivant : il peut arriver que le problème de Cauchy, pour les équations de Kolmogoroff (p.26, l.2 dans Doebelin), défini par  $a_{ik}(0) = \delta_{ik}$  et  $a'_{ik}(0) = D_{ik}$  donnés, ait une infinité de solutions.

Dans une note du 27 novembre 1950 je me suis contenté de donner deux exemples curieux.<sup>294</sup> La correction des épreuves de mon analyse fonctionnelle<sup>295</sup> m'a jusqu'ici empêché de rédiger mon mémoire développé sur ces questions.<sup>296</sup>

<sup>292</sup>Doebelin, Wolfgang (1938) Sur l'équation matricielle  $a(t+s) = a(t)a(s)$  et ses applications aux probabilités en chaîne. Bull. Soc. Math. Fr. 62:21–32. Doebelin's prewar work, related to the Kolmogorov equation and Markov processes, are considered in the present letter. Before the war Fréchet and Lévy worked on Markov processes, but subsequently went on to other topics. Nevertheless, they were aware that the Americans, especially Doob, were focussing on these subjects. At the 1950 Berkeley Symposium, Lévy met Kai Lai Chung who inspired him to work on stationary Markov chains. In Doob's *Stochastic Processes*, 1953, five of Doebelin's memoirs on Markov processes are cited but only one of Fréchet's. Feller's *An introduction to Probability theory and its applications*, 1964, includes an investigation Doebelin's results and their consequences. The influence of Doebelin's work on Chung is explained in Chung's memoir Chung, Kai Lai (1964) The general theory of Markov processes according to Doebelin. Zeit. Wahrscheinlichkeitstheor. 2:230–254. See also Mazliak (2007) and, above all, B.Bru and M.Yor (2000) CRAS 331, the special number dedicated to the sealed letter sent by Doebelin just before his 1940 suicide to escape capture by the Germans.

<sup>293</sup>The  $a_{ik}(t)$  are the coefficients of the Markov process transition matrix  $a(t)$ . The equation  $a(t+s) = a(t)a(s)$  is a consequence of the definition of the Markov process for which the future and past are independent, conditioned on the present. This property gives the definition of Markovian semigroups. It plays a fundamental role to define an infinitesimal generator for operators associated to a diffusion.

<sup>294</sup>Lévy, Paul (1951) Deux nouveaux exemples de processus stochastiques. CRAS 231:1208–1210.

<sup>295</sup>See the three preceding letters.

<sup>296</sup>In 1951 Lévy wrote the two notes Lévy, Paul (1951) Fonctions aléatoires dépendant de processus à la fois markoviens et stationnaires. CRAS 232:1400–1402, and Lévy, Paul (1951) Processus

A ce sujet, je vous signale que votre définition d'espace vectoriel n'est pas adoptée par les jeunes (d'après Schwartz que je crois bien informé).<sup>297</sup> C'est pour cela qu'après des corrections suggérées par votre remarque, et ennuyé à l'idée de les annuler, j'ai mis une note disant que d'autres auteurs n'adoptaient pas cette définition, et que je ne l'adoptais moi-même que pour un chapitre (il aurait autrement fallu faire diverses corrections dans la suite du livre ; mais à ce chapitre cela m'était commode).

Avez-vous jamais remarqué l'ambiguïté de l'expression "fonction de carré sommable" ; on devrait dire "fonction mesurable et de carré sommable". En effet une fonction toujours égale à  $-1$  ou à  $+1$  est de carré sommable et peut n'être pas mesurable.<sup>298</sup>

Je m'en suis aperçu trop tard pour faire dans mon livre les corrections voulues. Ou plutôt je l'avais oublié. Car j'ai trouvé cette remarque faite explicitement dans mon fascicule du mémorial, rédigé il y peut-être 15 mois, dont je viens de relire les épreuves.<sup>299</sup>

Bien cordialement à vous

P.Lévy

P.S. - Avez-vous jamais songé au problème suivant ?

Je pose  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$ . et, pour l'étude d'une fonction aléatoire  $X(t)$

$\Pr\{\dots\} = \text{probabilité } a \text{ priori}$

$\Pr_{t_0}\{\dots\} = \text{pr. estimée à l'instant } t_0, \text{ en connaissant } X(u) \text{ pour tous les } u \leq t_0.$

On peut définir le mouvement brownien, soit par

$$\Pr_{t_0}\{X(t) - X(t_0) < \xi \sqrt{t - t_0}\} = \Phi(\xi)$$

(pour  $t > t_0$ ) soit par  $\Pr\{\dots\} = \Phi(\xi)$  et cette probabilité est indépendante des valeurs antérieures de  $X(u)$  ( $u \leq t_0$ ).

Ceci est une réponse tardive à une remarque que vous aviez faite sur mes épreuves en 1948. Vous m'aviez suggéré d'effacer les mots soulignés, ce qui à mon sens n'est possible que si on écrit  $\Pr_{t_0}$  et non  $\Pr$ .

de Markov : cas dénombrable. CRAS 232:1803–1805, followed by a 54 page memoir on stationary Markov systems Lévy, Paul (1951) Systèmes markoviens et stationnaires. Cas dénombrable. Ann. Ec. Norm. Supér. 68:327–381.

<sup>297</sup>Lévy's remark is not just teasing. At the time, the vocabulary of linear algebra in France was far from standardized, despite the efforts of Bourbaki. Resistance to modern usage and to the axiomatization of vector spaces, as found in the work of Fréchet and Banach, was very strong and supported by the traditional geometric interpretation of the notion. See Godement's remarks in his Godement, Roger (1963) Cours d'Algèbre. Hermann, Paris. p. 650.

<sup>298</sup>If  $E$  is a non Lebesgue-measurable set, the function that is 1 on  $E$  and  $-1$  on its complement is an example. Lévy, who often had to respond to criticism of his supposed lack of rigor, was very aware of these questions. At times he was even somewhat excessive in preventing possible problems.

<sup>299</sup>Lévy, Paul (1952) Le mouvement brownien. Méml. Sci. Math. 126.

Mais alors je me suis demandé : existe-t-il effectivement d'autres processus pour lesquels

$$\Pr\{X(t) - X(t_0) < \xi \sqrt{|t - t_0|}\} = \Phi(\xi) ?$$

Je n'ai pas encore résolu cette question, ni même trouvé le temps d'y penser sérieusement.<sup>300</sup>

P.L.

### Lettre 67

Paris - 11 avril 1951

Mon cher collègue,

1° Merci pour votre lettre du 10. Je suis bien d'accord.

J'avais lu trop vite, et cru que vous utilisez la remarque que si  $p_{h,k}^{(n)}$  est une solution du système (1), p.105, et si,  $|s| = 1$ ,  $p_{h,k}^{(n+1)} = sp_{h,k}^{(n)}$  est une solution de même module. Alors mon objection aurait porté. Mais il y a, à cheval sur les p. 107-108, un autre principe de raisonnement, très simple, et que j'avais lu un peu vite, ayant cru à tort que les deux parties de ce paragraphes étaient indépendantes.

Ceci dit, il me semble qu'il y a tout de même une petite confusion, à rectifier dans la seconde édition, entre le systèmes associés

$$(1) \quad sw_k = \sum_j p_{jk} w_j \quad (1') \quad sw_k = \sum_j p_{kj} w_j$$

Si je ne me trompe, vous n'utilisez que le second ; c'est-à-dire que dans votre formule (1) il y a simplement à remplacer  $p_{jk}$  par  $p_{kj}$ .

2° J'ai de mon côté été conduit aux remarques suivantes :

Théorème de Frobenius. - La solution générale de

$$sw_k = \sum_k p_{jk} w_j \tag{1}$$

est

$$w_k = \sum_k c_k p_{h,k}^{(1)} = \sum_k c_k p_{h,k}$$

et la solution des équations itérées est  $\sum_k c_k p_{h,k}^{(n)}$ . Elle reste finie pour  $n$  infini. Donc on n'a que des racines  $|s| \leq 1$ , et si  $x = e^{i\theta}$ , il faut exclure les solutions qui introduiraient des puissances de  $n$ .

Cela n'est peut-être pas plus simple que le raisonnement que vous donnez p.105-106. Mais cela me paraît bien intuitif.

---

<sup>300</sup>The question is whether there exists a stochastic process with increments  $X(t) - X(s)$  that are Gaussian of mean zero and variance  $t - s$  that is not a Brownian motion. In Stoyanov, Jordan. M. (1997) Counterexamples in Probability, 2nd edn. Wiley, New York. Stoyanov left this question open. However, in 2000, in Föllmer, Hans and Marc Yor. On weak Brownian motion of arbitrary order. Ann. Inst. Henri Poincaré 36:447-487, the authors constructed such a process.



Votre théorème de la p.102. - Je le démontre comme suit, en utilisant les principes que j'ai établis en 1939 (Bull.Soc.Math.)<sup>301</sup> sur l'addition des variables aléatoires définies sur une circonférence, et notamment : soit  $S_n = X_1 + \dots + X_n$ , les  $X_\nu$  étant indépendants, dépendant de la même loi et zéro étant une valeur possible. On étudie la répartition de  $S_n$  (pour  $n$  très grand)  $S_n$  sur une circonférence de longueur 1. Alors, ou bien tous les  $X_n$  sont multiples de  $\frac{1}{p}$ , mais ne le sont pas tous de  $\frac{k}{p}$  (quel que soit  $k > 1$  et diviseur de  $p$ ) ; alors il y a tendance vers la répartition uniforme entre les valeurs  $0, \frac{1}{p}, \dots, 1 - \frac{1}{p}$ . Ou bien il n'en est ainsi pour aucun entier  $p$ , alors la répartition de  $S_n$  tend vers la répartition uniforme sur la circonférence.

Soit alors  $N_1, N_2, \dots$ , la suite des  $n$  pour lesquels  $H_n = k$ . Tous les  $N_{\nu+1} - N_\nu$  dépendant de la même loi, les théorèmes précédents s'appliquent. D'ailleurs  $p_{h,k}^{(n)} = \sum_{\nu=1}^n \{N_\nu \text{ existe et } = n, \text{ dans l'hypothèse } H_0 = h\}$ . Pour  $n$  très grands, les  $\nu$  très grands interviennent seuls, à une erreur près  $< \varepsilon$ . On en déduit la tendance à l'équipartition, par rapport à n'importe quel module, sauf dans le seul cas cyclique où il y a une période sûre. Cela donne à la fois votre théorème de la p.108 et celui que vous indiquez plus loin sur l'impossibilité de termes trigonométriques dans le cas d'une variable continue  $t$ .

3° Merci de vos indications bibliographiques. J'ai eu des déceptions depuis 8 mois que j'étudie ces questions en retrouvant dans Kolmogoroff les plus importants de mes théorèmes (retrouvés en août dernier et depuis), et celui dont je viens de parler dans votre livre. J'en aurai peut-être encore, car je vois dans Fortet la référence d'un travail de Doob que je ne connais pas encore. Mais je pense qu'elles seront limitées, car j'ai commencé ces recherches pour répondre à une question que Doob et Chung avaient en vain cherché à résoudre, et plusieurs conversations avec Chung m'ont convaincu que j'étais bien sur une voie nouvelle.<sup>302</sup>

En tout cas, si vous ne consacrez pas une troisième partie de votre 2ième édition à ces questions, un travail didactique, où la partie nouvelle sera plus ou moins importante, me semble n'être pas inutile.

Bien cordialement à vous

P.Lévy

Je viens de présenter une 3ième note sur ces questions (9 avril).<sup>303</sup>

Je viens de remettre à la bibliothèque de l'Institut H.P. le texte d'une conférence que j'ai faite à l'X sur ces questions. Je n'en ai que 5 ; aussi je ne vous en envoie pas ; j'ai prié Fèvre<sup>304</sup> de vous le montrer.

<sup>301</sup>Lévy, Paul (1939) L'addition des variables aléatoires définies sur une circonférence. Bull. Soc. Math. Fr. 67:1-41.

<sup>302</sup>See Kai Lai Chung's Preface to this volume.

<sup>303</sup>In 1950 and 1951 Lévy published five notes to CRAS on Markov processes; 15 pages in the proceedings of the Congrès International des Mathématiciens of 1950; and 54 pages in the annals of the ENS (number 68, 1951).

<sup>304</sup>See note 250.

Je me suis aperçu que j'ai eu tort de dire cycle (final). J'adopte maintenant groupe avec Doebelin et Fortet. Je comprends pourquoi vous évitez ce mot ; mais j'ai l'impression que l'usage s'établit de dire groupe. Par contre, je propose délibérément un néologisme en disant au pluriel cycles finaux (finaux sonne trop mal)

## Lettre 68

Paris - 19 mai 1951

Mon cher collègue,

Je vous renvoie votre feuille complétée. Pour les mémoires de Doob, entièrement consacrés au sujet mentionné, je ne sais pas si vous désirez une page particulière. Je vous ai d'ailleurs rendu le "Markoff chains" : je l'ai mis le 9 mai dans votre case à l'Institut Poincaré, avec une lettre attirant votre attention sur le fait que, dans "Topics. . ." (p.52), Doob démontre très simplement que si les fonctions  $P_{l,k}(t)$  sont mesurables, elles sont continues. C'est pour une autre question qu'il renvoie à Doebelin.

J'ai regretté de ne pas vous voir hier. La conférence de Kampé de Fériet a été très intéressante. Il est toujours très intéressant. Il est très estimé en Amérique, mais je me demande si en France on se rend bien compte de sa valeur.<sup>305</sup>

Ceci dit, je voudrais vous parler de l'Académie.<sup>306</sup> Je ne l'aurais pas fait si tôt si je n'avais pas été mercredi présenter une Note, et plusieurs personnes m'ont déjà demandé mes intentions, notamment M.Hadamard. Je lui ai répondu qu'il ne s'agissait pas de me "jeter en travers de votre chemin", mais qu'une abstention de ma part équivaldrait à une renonciation définitive.<sup>307</sup> Mes amis ne comprendraient pas. Vous n'êtes pas sans savoir qu'un troisième candidat<sup>308</sup> a dès maintenant une position très forte. Je ne crois pas qu'il ait des chances sérieuses contre vous, mais

---

<sup>305</sup>Joseph Kampé de Fériet was one of the leading European experts on turbulence and random functions. The preceding letters show Lévy's great regard for the Berkeley school. He knew that they were well ahead of the French school on many aspects of random functions. His colleagues Loève and Bass shared this opinion. Lévy's felt isolated in France but Lévy's candidacy to the Academy, anticipated here, received strong international (primarily American) support.

<sup>306</sup>Elie Cartan died on May 6, 1951, at the age of 82, so an election to the Geometry Section, the first since 1942, was in the offing. Fréchet's candidacy (he was 74) was recorded on January 28, 1952; Lévy's (he was 66) on February 4, 1952; and René Garnier's (65) on February 18, 1952. Fréchet was in the first tier (the candidates were ranked by the 'secret committee'; see the letter of February 26, 1952) with Lévy and Garnier tied in second place along with Georges Valiron (68). Mandelbrojt (53) and the young Henri Cartan (48) made up the third tier.

<sup>307</sup>Since 1938 Lévy had made clear his wish to join the Academy as a Free Academician, and was four times a favorite. From his visits to the Academy Lévy realized that many of the Academicians thought he belonged in the Geometry Section (see his letter of July 1951 to the members of the Academy, which is attached to the next letter). He was thus concerned, and with good reason, that not registering for the election to the Geometry Section about to take place would lose him support and block any future advancement.

<sup>308</sup>The third candidate, clearly Garnier, had very strong support. His reputation in algebraic geometry and his older work on the Plateau problem put him squarely in the academic tradition of the time. The letters to come show how a negotiated alliance arose between Lévy and Fréchet, in opposition to Garnier.

le crois assuré de 15 à 20 voix (un peu plus si je ne me présente pas) ; en tout cas à la section. Si je ne me présente pas, ce serait lui abandonner sans combat de très grandes chances pour la fois suivante.

Je me présenterai donc. Sans l'insistance de géomètres, de Chazy<sup>309</sup> et d'autres, j'aurais renoncé dès 1944 à l'autre solution. Je compte expliquer les raisons de mes attitudes successives assez clairement pour qu'on ne puisse pas me reprocher mes changements. Mais je compte dire aussi à mes amis que je comprendrais très bien qu'on ne vous fasse pas attendre plus longtemps et qu'il n'y a en tout cas pas à hésiter entre vous et les autres. Je vous ai déjà exprimé mon sentiment à ce sujet, et je tiens à vous assurer que je ne parlerai pas autrement lors de mes visites de candidature.

Bien cordialement à vous,

P.Lévy

Je ne compte d'ailleurs pas faire la tournée complète des visites. Il y a des membres de l'Académie que j'ai déjà vu 2 fois ; cela me suffit et peut-être à eux aussi.

### **Lettre 69**

Paris - 38 Av.Th.Gautier

2 juillet 1951

Mon cher Collègue,

Je crois que, tout en défendant chacun nos titres, nous devrions nous aider, en donnant l'impression qu'il ne "devrait" y avoir de compétition sérieuse qu'entre nous deux.

J'envoie à tout le monde, dès maintenant, une lettre circulaire dont je vous envoie une copie ci-joint. J'ai une vingtaine environ d'amis sûrs auquel je dis explicitement que, si après 2 tours je n'avais aucune chance, c'est à vous qu'ils devraient se rallier ; et je compte qu'ils le feront.

Julia<sup>310</sup> fait une campagne active pour Garnier, et contre nous - ce qui n'est pas dans les usages de l'Académie.<sup>311</sup> Il raconte que vous êtes un philosophe, et pas un mathématicien ;<sup>312</sup> il me l'a dit à moi même, et j'ai protesté. Si vous savez ce

<sup>309</sup>Jean Chazy (1882–1955) encountered Fréchet and Lévy at Hadamard's seminar. In addition to his mathematical work, Chazy was, in the 1950s, the last representative of the French school of celestial mechanics. He was known for having extended Poincaré's work on the three body problem. He also worked on developing relativity theory; a relativistic formula for the advance of the perihelion of a planet bears his name.

<sup>310</sup>Gaston Julia (1893–1978), a hero of the Great War who carried the scars of his facial injuries, remained very influential at the Academy. It was Garnier who wrote Julia's obituary (CRAS 286:22–25 (1978)).

<sup>311</sup>Lebesgue died in 1941, resulting in an election in 1942. Lévy, having fled Paris, could obviously not be a candidate. Denjoy was then elected with 21 votes, followed by Garnier with 16. Fréchet only got 2 votes, Valiron none. Perhaps Julia thought that Garnier's time had come. Henri Cartan was tainted by his association with Bourbaki, representing a new generation little appreciated by a very traditional Academy. Mandelbrojt's third tier position was purely honorary.

<sup>312</sup>An allusion to Fréchet's interest in the philosophy of mathematics and probably to his contributions to the Congrès de Philosophie of 1949 (see the letter of November 19, 1951).

qu'il dit contre moi, je compte que vous me le répéterez, pour me permettre de me défendre. La vraie raison de sa rancune contre moi, j'en suis sûr, est que je suis professeur à Polytechnique depuis 1920. Il avait 27 ans, et s'était présenté.<sup>313</sup>

J'aurais dû reconnaître le génie naissant et m'effacer. Je ne l'ai pas fait et il me l'a pas pardonné. La rancune a éclaté quand, nommé à son tour à l'X, il était sûr de n'avoir plus besoin de moi.<sup>314</sup>

Jolibois<sup>315</sup> me dit que l'active campagne de Julia est déjà commencée, et qu'il faut me hâter, même si je ne peux pas faire maintenant de visites, de faire savoir à tout le monde que je suis candidat. Vous devriez faire de même, avant l'été.

Bien cordialement à vous

P.Lévy.

Je pense que vous savez que vous serez présenté par la Section. Il n'y a aucun doute.

C'est ma pleurésie d'il y a trois ans dont les suites m'obligent de nouveau à quelques précautions.<sup>316</sup> Je n'ai pas de raison d'être très inquiet, mais je ne peux pas commencer mes visites dès maintenant, et le peu de temps qui séparaera l'été de l'élection suffira difficilement, car il me faut réduire mon activité.

*Joint to the letter of July 2nd, 1951:*

Monsieur,<sup>317</sup>

L'objet de la présente lettre est de vous informer de ma candidature à la section de géométrie de l'Académie des Sciences, et en même temps, non de vous exposer mes titres (ce qui a été fait ou sera fait autrement), mais de vous expliquer pourquoi je renonce en le faisant à une autre voie dans laquelle je m'étais engagé.

J'avais toujours espéré être un jour membre de la section de géométrie. C'est il y a vingt ans environ que le géomètre M.d'Ocagne, Académicien libre, me conseilla de suivre la même voie que lui, insinuant que j'arriverais plus vite de cette manière. Je crois que sans lui je n'y aurais jamais songé. Mais cette idée trouva dans mon esprit un terrain favorable à la suite de la déception que je venais d'éprouver. Sur la foi d'un des membres de la Section de géométrie, j'avais compté être proposé, en troisième ligne ex æquo avec d'autres candidats, sur la liste des candidats présentés pour la succession du Recteur P.Appell à la dite section.<sup>318</sup> Je n'ai figuré sur cette liste que beaucoup plus tard.

C'est après la mort de M.d'Ocagne, en 1938, que je me décidai à suivre son conseil. Il y avait quatre fauteuils d'Académiciens libres simultanément vacants, et je pensais que l'Académie envisagerait de remplacer un mathématicien par un autre. Mes amis me dirent que j'aurais

<sup>313</sup>Lévy was 34 at the time, a teaching assistant at the Polytechnique as of October 13, 1913, and professor of analysis, geometry and topography at the Ecole Supérieure des Mines as of August 21, 1914.

<sup>314</sup>See next letter.

<sup>315</sup>Pierre Jolibois (1884–1954), chemist and professor of general and applied chemistry at the Ecole des Mines, was elected to the Academy of Sciences on June 26, 1944. Remind that all the Sections of the Academy participate to the election of a new member (but the non-specialists generally follow the advice of the section to which the candidate would belong.

<sup>316</sup>Once again, Lévy had serious pulmonary problems (see following letters).

<sup>317</sup>Let us recall that the first woman to be elected in Paris Academy of Sciences was Yvonne Choquet Bruhat in 1979.

<sup>318</sup>Paul Appell died in 1930.

quelques chances, au moins pour le quatrième poste. Enfin l'approche d'événements qui n'étaient que trop prévisibles acheva de me décider ; je ne pouvais guère, comme les autres, compter sur l'avenir, et attendre patiemment.

La guerre survint, puis la libération. La raison qui m'avait décidé n'existait heureusement plus. Je rentrai à Paris, presque décidé à renoncer à la voie où je m'étais engagé en 1938. Mais je me trouvai prisonnier de mon attitude - Mes amis m'engageaient à persévérer, et l'évènement faillit leur donner raison. Quatre fois je fus présenté en première ligne, et une fois il s'en fallut d'une voix que je fusse élu.

J'avais constaté au cours de mes visites qu'un grand nombre d'Académiciens estimaient que ma place serait plutôt à la Section de géométrie, et il m'était d'autant plus difficile de leur répondre qu'au fond de moi je n'ai jamais cessé de leur donner raison. Même parmi ceux qui me promettaient leur voix, beaucoup le faisait à regret. Ainsi, en étant aujourd'hui candidat à la Section de géométrie, je suis le conseil qui m'a été donné par beaucoup d'Académiciens, et qui concorde avec mon sentiment intime.

J'espère que vous ne me reprocherez pas un changement d'attitude qui a peut-être été une maladresse, mais qui n'a pas été un caprice, et que vous accepterez de juger ma candidature avant tout sur mes titres scientifiques. Je vous rappelle d'ailleurs que je ne pouvais pas être candidat en 1942 et qu'il n'y a eu finalement à la Section de géométrie aucune élection dont je fusse volontairement absent.

Je tiens à ajouter que je ne serai ni surpris ni humilié si l'Académie ne croit pas devoir faire attendre plus longtemps un de mes collègues et amis, plus âgé que moi, et dont les éminents titres scientifiques sont incontestables. Mais je serai heureux et honoré si, malgré la présence d'un tel concurrent, vous voulez bien m'accorder votre suffrage.

Enfin je vous prie de m'excuser si ma santé et mes occupations que j'espère pouvoir ne pas interrompre m'empêchent de faire toutes les visites d'usage. Je compte bien entendu me présenter à tous ceux des membres parisiens de l'Académie qui ne me connaissent pas ou ne m'ont vu qu'une fois, mais je ne suis pas sûr de pouvoir faire plus.

## Lettre 70

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<sup>319</sup> lui avait dit "présentez-vous ; vous aurez la 2<sup>ième</sup> ligne

---

<sup>319</sup>Georges Humbert (1859–1921), *polytechnicien*, became president of the SMF in 1893. He was professor of analysis at the Ecole Polytechnique in 1895, successor to Charles Hermite in 1901 at the Academy of Sciences, and in 1912 successor to Jordan at the Collège de France. Humbert and Lucien Lévy, Paul's father, were colleagues at the Polytechnique. In 1920 Lévy replaced Humbert as professor at the Ecole Polytechnique, Humbert having taken early retirement for health reasons. Humbert intervened in the 1918 contest involving Julia and Pierre Fatou, who were vying for the Grand Prix of the Academy of Sciences (on this topic see the book by Michèle Audin: Fatou, Julia, Montel, the great prize of Mathematical Sciences of 1918, and beyond. Springer (2011)). The topic, chosen in 1915, was the iteration of functions. Fatou announced his results in CRAS, December 1917. Julia claimed he had published first, requesting, on December 31, 1917, that Humbert open the sealed letter that he had deposited at the Academy. Julia won the Grand Prix with his *Mémoire sur l'itération des fonctions rationnelles* (*Memoir on the Iteration of Rational Functions*), published in the J. Math. Pures Appl. 8:47–245 (1918). It remains one of his most recognized achievements.

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 combatif ; 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.<sup>320</sup>

Il ne reste donc qu’un tout petit point de ma lettre précédente\* [*on the bottom of the page*: \* 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

## Lettre 71

Paris - 38 Av.Th.Gautier

1-10-51

Mon cher Collègue,

F.P.Cantelli (105 Via Merulana, Roma) me demande en vue d’une conférence qu’il doit faire à Messine le 25 octobre,

une demi-page sur les contributions les plus originales suggérées par le calcul des probabilités, dues à M.Fréchet.

Je lui adresse une réponse provisoire - pour le cas où vous seriez en voyage et où ma lettre vous rejoindrait trop tard. Mais j’ai peur d’oublier des choses que vous jugeriez importantes. Le mieux serait que vous lui envoyiez lui-même<sup>321</sup> une meilleure rédaction de la Note demandée. Je lui écris que je vous le demande.

Bien cordialement à vous      P.Lévy.

L’élection à l’Académie n’est, paraît-il, qu’en janvier.

Malgré l’optimisme des amis de G., et l’inquiétude de M.Borel, je reste convaincu que cette place et la suivante seront pour vous et moi, dans l’un ou l’autre des 2! ordres possibles.

J’oublie de vous dire que pour Cantelli, il s’agit de l’aspect analytique “Contributions à l’analyse suggérées par le calcul des probabilités”.

---

<sup>320</sup>Lévy in fact, as we shall see, received 18 votes in the first round.

<sup>321</sup>*Sic.*

## Lettre 72

*Fréchet's handwriting with a pencil: answered on November 23rd*

Paris -19-11-51

Mon cher Collègue -

Excusez-moi d'avoir tant tardé à vous répondre. Votre lettre m'avait pourtant d'autant plus touché que votre sympathie contraste avec la superbe indifférence de l'Ecole Polytechnique qui m'a simplement prévenu que ma demande de congé était acceptée. Quoique j'aie vu pire en 43, j'ai trouvé que c'était peu.<sup>322</sup>

Je ne suis pas inquiet pour l'avenir. Grâce aux progrès très récents de la thérapeutique, ces affections pulmonaires se guérissent bien. Mais il faut une longue patience, et on exige un repos absolu du corps et de l'esprit, qui ne permet guère un travail sérieux. Tout au plus quelques lettres, et celle-ci m'occupera peut-être 2 jours, car j'ai pas mal de choses à vous dire.

Je pense que vous avez remarqué, dans *Econometrica*, la notice sur Wald,<sup>323</sup> et dans ma Notice complémentaire, le résultat d'Ottaviani, qui doit vous intéresser, puisque ma communication de Lyon vous avait intéressé.

Votre Notice m'a rappelé notre discussion de 1928,<sup>324</sup> et votre mémoire où j'ai précisément retrouvé l'exemple de Hausdorff dont vous me parliez récemment. Je pense que vous êtes convaincu maintenant que vous aviez interprété cet exemple d'une manière incorrecte. Mais votre Notice, qui vient d'être largement distribuée, donne de la deuxième partie de notre discussion une interprétation qui ne me semble pas correcte, et je me réserve de demander à Caquot<sup>325</sup> d'y répondre brièvement.

Je vous surprends peut-être, parce que je n'avais pas dit cela aussi nettement dans ma réponse de 1929 ni dans ma Notice de 1935. Je savais bien ne pas m'être trompé en pensée. Mais en vous lisant, au lieu de commencer par me relire, j'ai dû me persuader que si vous n'aviez pas compris, c'était que je m'étais mal exprimé. Or en relisant mon livre de 1925, il me semble que vous auriez dû me comprendre.

Je précise : vous m'avez reproché d'avoir dit d'une manière trop absolue que les erreurs accidentelles obéissent à la loi de Gauss.

1° Vous dites : on peut l'expliquer, mais pas le démontrer.

Or vous trouverez dans mon livre à ce sujet :

p.70, vers le bas : "S'il est exact que les erreurs accidentelles obéissent à la loi de Gauss, il y a certainement à cela une raison que nous pouvons essayer de comprendre"

<sup>322</sup>See the letter of November 29, 1943.

<sup>323</sup>Abraham Wald (1902–1950), who took refuge in the USA during the war, had died, tragically, in India at age 47 in an airplane accident. He worked in both pure mathematics and econometrics. His work in statistics was on statistical decision theory.

<sup>324</sup>The debate on the law of errors (see Sect. 4). Lévy was worried that this dialogue with Fréchet would become ammunition for his adversaries. At Lévy's request, a letter was addressed by Fréchet and Lévy to the members of the Academy to forestall this possibility. A first draft of this letter, written by Lévy and corrected by Fréchet, was found among the letters of this period and is reproduced below.

<sup>325</sup>Albert Caquot (1881–1976), the 'best living French engineer' was 70 at the time. His mathematical work on geotechnics was considered authoritative.

p.72, 1. 1–3 : “Si nous arrivons à démontrer que. . . , c’est un peu parce que nous avons mis dans la définition des erreurs accidentelles, ce qu’il fallait pour arriver à ce résultat”. Suit une nouvelle critique de cette définition.

p.74 1.3–5 ; je note que la vérification expérimentale est tout de même le plus souvent nécessaire (on est rarement sûr a priori que toutes les conditions sont vérifiées).

2° La phrase en italique p.73–74 montre bien que l’énoncé n’est vrai qu’asymptotiquement. Pour un nombre fini de termes, on ne doit pas s’attendre à trouver la loi de Gauss. Je l’avais dit avant vous (v. aussi le haut de la p.73).

3° Vous avez insisté sur l’importance de l’hypothèse de l’additivité des erreurs partielles. Or cette hypothèse n’est nullement restrictive. On peut toujours représenter l’erreur totale par  $y_1 + y_2 + \dots + y_n$ , la somme partielle  $y_1 + \dots + y_\nu$  représentant l’effet des  $\nu$  premières causes d’erreur (v.p.73, Note (1)). On peut donc dire que les erreurs  $y_\nu$  s’ajoutent.

On peut aussi (v.p.279), exprimer les  $y_\nu$  par des fonctions  $\varphi = \varphi(x_1, \dots, x_\nu)$  de variables auxiliaires  $x_\nu$  indépendantes les unes des autres.

Donc, on peut, soit appeler les  $x_\nu$  erreurs partielles et dire que ces erreurs sont indépendantes, soit appeler les  $y_\nu$  erreurs partielles et dire qu’elles s’ajoutent. Aucune de ces hypothèses n’est restrictive. Ce qui est restrictif, c’est de supposer les deux et cela je l’ai fait remarquer au milieu de la p.73.

Mais j’explique que je considère la mesure comme assez précise si l’erreur possible (donc en particulier  $y_1 + \dots + y_{\nu-1}$ ) est assez petite pour que, dans les limites d’incertitude qui en résultent, les conditions de la mesure varient peu. J’aurais dû préciser que cela veut dire en particulier que la loi de  $y_\nu$  varie peu. Donc les variables  $y_\nu$  qui s’ajoutent sont sensiblement indépendantes et la conclusion est absolument correcte.

Conclusion de cette discussion. Il y aurait peut-être intérêt à préciser et développer certaines remarques indiquées d’une manière trop concise, mais si on ne s’attache pas à une phrase isolée du contexte il n’y a guère de risque d’erreur, et toutes les remarques de la partie critique de votre mémoire de 1928 se trouvent déjà dans mon livre.

Je reviens à l’exemple de Hausdorff. Il y avait longtemps que j’avais considéré des séries presque sûrement convergentes de cette nature (par ex.  $\sum \pm \frac{1}{n}$ <sup>326</sup>) et savais que la fonction caractéristique de la somme  $\sum a_n X_n$  est  $\varphi(z) = \prod \varphi_n(a_n z)$ . La formule de Hausdorff n’est qu’une application triviale de cette remarque, et j’avais été stupéfait que vous la citiez à l’appui de votre affirmation qu’une loi n’appartient au domaine d’attraction de la loi de Gauss que si elle lui ressemble suffisamment.

Je ne m’étais pas attaché au fait que c’était la première fois qu’on signalait cette loi dont la f.c. est  $\frac{1}{\text{ch } z}$  et je vous remercie de me l’avoir rappelé.<sup>327</sup>

<sup>326</sup>In Lévy (1970, p. 86), Lévy reports his 1922 conversation with Wiener on this subject. When the + and – signs are chosen successively and independently at random with probability 1/2, this series converges a.s., because it is square summable. For series  $\sum \pm a_n \frac{1}{n}$  with sufficient conditions on the coefficients  $a_n$  there is an analogous result.

<sup>327</sup>This is also the characteristic function of the Lévy Area. See the letter of January 23, 1950.



J'ai lu votre exposé au Congrès de Philosophie de 1949.<sup>328</sup> Naturellement nous ne sommes pas tout à fait d'accord.

Pour moi, la meilleure schématisation d'un problème de probabilité s'obtient lorsque :

- a) il y a une raison précise de considérer les différents cas comme indiscernables (symétrie ; identité des cartes au point de vue du toucher, etc.).
- b) il n'y a aucune raison de s'attendre à l'un plutôt qu' à l'autre (et pour cela, j'ai souvent insisté sur la différence entre le point de vue subjectif et le point de vue objectif. Il peut arriver qu'un jeu de cartes me semble bien battu après 3 ou 4 répétitions d'un même geste ; pourtant il est bien connu que cela ne suffit pas pour que les 52 ! permutations soient également probables ; sans doute ne sont-elles pas toutes possibles).

Vous remarquerez que la condition b) n'est pratiquement pas bien réalisée au jeu de pile ou face. Je suis d'accord. C'est pour cela que je préfère parler du jeu de carte, ou de la roulette, où elle est au moins asymptotiquement réalisée.

Si, dans une expérience, je veux savoir si cette condition est bien réalisée, il faut évidemment regarder les choses de plus près, et voir de quelle catégorie d'épreuves il s'agit. Mais il faut expliquer pourquoi on peut dans ces conditions considérer qu'il y a une fréquence limite presque sûre, et c'est ce que vous ne faites pas.

Je reviens à la première question. Je ne voudrais naturellement pas engager une polémique. Mais j'espère que vous comprendrez que je désire que Caquot dise quelque chose pour compenser l'impression que donne votre Notice qu'un des résultats importants de mon livre de 1925 est faux- et que j'aie voulu vous prévenir pour qu'on ne puisse pas me reprocher une manœuvre de la dernière heure.

Caquot me croira sur parole et je ne lui donnerai pas beaucoup de détails. Je compte mettre aussi Montel au courant ; il est compétent et impartial ; et si quelqu'un me reprochait une manœuvre de la dernière heure, il interviendrait pour assurer que ce que Caquot aura dit est bien exact.

Je pense que vous irez à la réunion du Comité des mathématiciens. Voulez-vous m'y excuser.

Je pense que vous partirez peu après pour les Indes.<sup>329</sup> J'espère que vous y ferez un bon voyage.

Bien cordialement

P.Lévy

### **Lettre 73**

Paris - 20-11-51

[Fréchet's handwritten with a pencil: Répondu les 23 et 25 novembre]

---

<sup>328</sup>Mathematics was an important topic at this meeting. Apéry presented a case for intuitionism, and criticized the axiomatic approach of the Bourbaki's kind (he had himself refused to join the Bourbakis).

<sup>329</sup>The dossier on his voyage to India is conserved in the archives of the Academy in the *Fonds Fréchet*.

Mon cher Collègue,

Je me reproche déjà de m'être bien trop appesanti sur notre discussion de 1928. Je demanderai simplement à Caquot de dire sans insister qu'à mon avis les remarques de votre mémoire de 1928 rappelées dans votre Notice n'ont fait que développer et non contredire des remarques contenues dans mon livre. Cela ne risquera pas de créer un incident et dans ces conditions je juge inutile de prévenir Montel.

A propos de Montel, il m'avait dit un jour il y a 20 ans au moins que la notion d'ensemble compact était due à Arzéla.<sup>330</sup> Votre notice me prouve qu'il s'est trompé. Vous pourriez à l'occasion vous assurer qu'il sait maintenant à quoi s'en tenir.

Peut-être y avait-il eu un malentendu, et avait-il voulu parler de l'application aux ensembles de fonctions. Mais je ne le crois pas.

Inutile de vous dire que je considère qu'il s'agit d'une notion très importante. Personne ne peut le contester.

Savez-vous quelque chose au sujet du problème suivant, résolu dans mon livre (de 1937<sup>331</sup>) dans le cas où  $a_n = 1$ .

On se donne une suite de constantes  $a_n$ , une loi  $\mathcal{L}$ , des variables indépendantes  $X_n$  dépendant de cette loi. A quelles conditions peut-on déterminer  $c_n$  et  $N = \varphi(n)$  de manière que

$$\frac{1}{N} \sum_1^N (X_n - c_n)$$

dépende d'une loi tendant vers celle de Laplace-Gauss ?

Vu hier un "grand pontife". Il me demande de me soumettre chaque jour pendant un mois à une piqûre intraveineuse d'une durée de 3 heures - parce que je n'ai pas pu supporter par voie buccale le remède qu'il s'agit de prendre ! Les remèdes sont parfois pires que les maladies ! Heureusement que nous ne sommes plus au temps de Molière et qu'on peut penser qu'on ne nous demande ces choses là qu'à bon escient.

Bien cordialement

P.Lévy

## Lettre 74

Paris - 6-12-51

Mon cher collègue,

---

<sup>330</sup>Paul Montel (1876–1975) was successively professor at the Ecole Polytechnique; professor at the Faculté des Sciences de Paris; professor at the Ecole des Beaux Arts; president of the Ecole Pratique des Hautes Etudes, first section; and president of the Palais de la Découverte. He was elected to the Academy of Sciences in 1937. In 1907 he had defended his thesis *Sur les suites infinies de fonctions* (*On Infinite Series of Functions*), which treats 'normal families of functions' along the lines of the work of the Italians Arzela and Ascoli. Montel's thesis includes a close study of compactness, which was introduced by Fréchet in metric spaces. A Montel space is now defined as a Fréchet space, in which the closure of each bounded subset is compact.

<sup>331</sup>Théorie de l'addition des variables aléatoires.

Je n'ai rien à objecter aux précisions que vous apportez à ma rédaction.

Je m'en rapporte donc à vous pour la suite à donner. Veuillez simplement expliquer à M.Borel que je ne veux pas "créer un incident", mais que je ne voudrais pas qu'on puisse croire que j'ai commis une erreur qui serait particulièrement ridicule, puisque j'ai commencé à m'occuper de calcul des probabilités précisément pour élucider cette question.

Vous voudrez bien, S.V.P., me tenir au courant de la réponse de M.Borel. J'attends sans impatience ; vous jugerez peut-être qu'il ne faut pas l'ennuyer pour cela avant son voyage aux Indes.

Bien cordialement à vous

P.Lévy.

*[Letter jointly signed by Paul Lévy and Maurice Fréchet and joint to the letter from December 6th, 1951]*

Les soussignés demandent à Monsieur... de lire à l'Académie des Sciences la Note ci-dessous, ou du moins de la résumer, au moment de la discussion sur les titres des candidats à la Section de géométrie.

« M.Fréchet a, dans son mémoire de 1928, attiré l'attention sur ce que l'énoncé connu : "les erreurs accidentelles obéissent à la loi de Gauss" ne doit être considéré que comme un énoncé approché, valable seulement si certaines conditions restrictives sont vérifiées. Il a notamment attiré l'attention sur l'importance de l'hypothèse de l'additivité des erreurs partielles, et à cet effet il a développé une théorie nouvelle relative au cas où l'on adopte une formule de composition autre que l'addition.

Monsieur Paul Lévy fait remarquer que, dans sa Notice de 1933, M.Fréchet donne l'impression qu'il s'agissait, dans ce Mémoire de 1928, d'une mise au point rendue nécessaire parce que, dans ses travaux antérieurs, Paul Lévy aurait méconnu la nécessité des restrictions mentionnées ci-dessus. Or il n'en est rien. Toutes ces restrictions sont développées aux p.72 à 74 du Calcul des Probabilités de Paul Lévy, publié en 1925 ; on y trouve aussi un exemple de composition par une formule autre que l'addition, peut-être moins original que celui de M.Fréchet mais dont les études ultérieures de Gibrat<sup>332</sup> sur les inégalités économiques ont montré l'importance.

M.Fréchet se déclare d'accord avec les remarques de Paul Lévy.

Il regrette d'avoir pu *[added in Fréchet's handwriting: contrairement à son intention]* donner l'impression *[added in Fréchet's handwriting: fût-ce seulement à quelques uns]* que Paul Lévy aurait commis une erreur dans un des résultats fondamentaux de son livre de 1925. »

P.Lévy      M.Fréchet

## **Lettre 75**

Paris - 26-2-52

---

<sup>332</sup>Robert Gibrat (1904–1980), a *polytechnicien* like Lévy, defended a thesis in economics at Lyon. In 1931 he published *Les inégalités économiques*. He believed that the log-normal distribution was capable of representing most of the probability distributions that apply to phenomena of economic inequality (income, inheritances, etc.). Others, including Fréchet and subsequently Mandelbrot or Allais, delegate this role mainly to distributions in the domain of attraction of different stable distributions, namely those of Pareto-Lévy (see Sect. 5). In modern western economies, the data sometimes favors Gibrat (income), sometimes Fréchet (urban population concentration, for example). In April 1942, Gibrat was made Communications Minister in the second Laval government in Vichy. Despite his resignation from that post in November 1942, he faced serious problems at the Liberation.

Mon cher Collègue,

Jolibois m'a téléphoné hier, après le Comité secret.<sup>333</sup> Il pense que l'Académie sera impressionnée par les exposés concordants de MM. Hadamard, Borel et Denjoy, et que votre élection est assurée ; que je serai après en très bonne posture, malgré Garnier. Je pense, bien que Jolibois ne soit pas prophète, que cela vous fera plaisir de le savoir. Je pense d'ailleurs qu'il est bien placé ; il est peut-être le seul de mes amis à avoir fait campagne pour moi, et sait quelles sont les voix que je peux espérer.

Je voulais d'ailleurs vous écrire pour une autre raison. M.Borel m'a suggéré de dire à mes amis que je ne demandais pas de voix ; c'est ce que Denjoy avait fait en 34 et 38 et ce que vous avez fait en 42 ;<sup>334</sup> il eût été assez naturel que je fasse de même. Je vous ai déjà écrit en juillet pourquoi je ne le faisais pas. Si Garnier donne des inquiétudes à M. Borel, c'est à cause des 16 voix qu'il a eues en 42 ;<sup>335</sup> si je ne me présentais pas, il en aurait 25, peut-être 30,<sup>336</sup> et on le considérerait comme désigné pour la fois d'après. J'espère au contraire, me présentant, arriver à 20 voix dont 5 lui seront enlevées, et, même s'il a 1 ou 2 voix de plus que moi, cela sera largement compensé la prochaine fois par l'appui de la Section.

Enfin je n'arrive pas à croire que M.Borel ait raison en pensant que ma candidature augmente les chances de Garnier. Je ne peux pas le démontrer, mais je ne doute pas que mes amis n'aient envisagé l'hypothèse que les premiers tours aboutissent à mon élimination, et que, s'ils ont à choisir entre Garnier et vous, le résultat soit à peu près le même, au 2<sup>o</sup> tour ou au 3<sup>o</sup>, qu'il eût été sans moi au 1er. Il y a bien une grande influence du 1er tour ; mais elle consiste, dans une élection à 3, à éliminer un des trois candidats.

J'oubliais de vous dire aussi qu'une autre raison de ne pas changer d'attitude est que j'en ai déjà changé 2 fois, en 38, et en juillet dernier. Un troisième changement déconcerterait mes amis.

Pour parler d'autre chose, je pense que vous avez remarqué dans le Journal de Cantelli le mémoire d'Ottaviani dont je vous avais parlé.

Bien cordialement à vous

P.Lévy.

Ne vous donnez pas la peine de me répondre. Vous m'avez déjà répondu en juillet.

## Lettre 76

38 avenue Théophile Gautier

<sup>333</sup>The meeting of the secret committee preceded the investigation leading to Fréchet's position in the first tier.

<sup>334</sup>As mentioned, Fréchet had in fact received 2 votes in 1942.

<sup>335</sup>See the commentaries to the letter of July 2, 1951.

<sup>336</sup>Lévy thought that if he stopped campaigning, the 20 votes that he would have received would be evenly divided between Fréchet and Garnier. This possibly would allow a majority for Fréchet but eliminate Lévy himself in favor of Garnier at the next election. Thus Lévy maintained his candidacy, to preserve his chances in the next election.

Paris - 38 Av. Théophile Gautier

24 octobre 1952

Mon cher Collègue,

J'apprends seulement, avec presque quatre mois de retard, que vous avez enfin été rattaché à l'Académie des Sciences comme membre correspondant. Je m'excuse de ne pas vous avoir plus tôt adressé mes félicitations.

J'avais d'ailleurs appris, peu de jours avant mon opération, en juin, que vous aviez, modestement, accepté cette solution.<sup>337</sup> Mais l'Académie n'est en général pas si pressée et je ne croyais pas que ce fût chose faite.

M. Borel me dit que vous êtes satisfait ; je ne voudrais donc pas vous donner de regrets. Mais je ne peux pas m'empêcher de constater qu'il est sans exemple depuis que l'Académie existe qu'un mathématicien français, ayant les titres que vous avez et ayant eu l'influence que vous avez eue, n'arrive pas à être membre d'une des deux premières sections. Malgré l'avantage qui peut un jour en résulter pour moi, ce n'est pas sans un peu de regret que j'apprends la conséquence actuelle - maintenant normale - des erreurs antérieures.

Mais je veux surtout vous féliciter, et exprimer l'espoir que vous arriverez bientôt à l'échelon supérieur dans cette nouvelle voie.<sup>338</sup>

Bien cordialement à vous

P. Lévy.

M. Borel m'a demandé, un peu prématurément, de vous appuyer à l'occasion auprès de mes amis. Il est bien entendu que, lorsque le moment sera venu, vous pourrez compter sur moi.

Je pense avoir de l'influence sur Jolibois, Becquerel, Leprince Ringuet, qui savent déjà ce que je pense de vous.

Je continue à aller aussi bien que possible. J'espère faire mon cours à l'Institut Poincaré au printemps prochain. Je travaille maintenant à une Note sur la loi des grands nombres ;<sup>339</sup> j'ai quelques résultats nouveaux. Je me rappelle en écrivant

<sup>337</sup>Fréchet and Lévy's difficult struggle ended with Garnier's election to the Academy. In the first round, with 66 votes cast, Garnier came in first with 29 votes, followed by Lévy with 18, Fréchet 17 and Valiron 2. Cartan and Mandelbrojt did not receive any votes. The second round (67 votes cast) gave Garnier an absolute majority with 39, Fréchet with 16 and Lévy with 12. Subsequently, Lévy was seriously ill for several months.

<sup>338</sup>A letter of May 5, 1952 (archives of the Academy of Sciences), from Borel to Fréchet shows that because of these electoral complications, Borel suggested to Fréchet that he present himself as a non-resident member. Borel had just learned of the death of Castelnuovo and explained to Fréchet that he could replace Castelnuovo as a corresponding member of the Geometry Section, an obligatory first step toward non-resident membership. In the same letter, Borel advised Fréchet to get an address outside Paris, because a non-resident member could not be a Parisian. It seems clear from the letter that Borel had explained this strategy to Lévy. The *échelon supérieur* (next step) in question is thus that of a non-resident member. In the end, Fréchet was elected on March 14, 1956, to the seat left vacant by the death of Borel.

<sup>339</sup>It would become Lévy, Paul (1952) *Loi faible et forte des grands nombres*. CRAS 235:1186–1188.

que vous le saviez déjà par Dugué;<sup>340</sup> mais les difficultés de rédaction m’ont arrêté plus longtemps que je ne le pensais, et je n’ai pas tout à fait terminé.

### Lettre 77

Paris - 38 Av. Th.Gautier  
9 juin 1956<sup>341</sup>

Mon cher collègue,

Merci beaucoup pour votre livre “les mathématiques et le concret”,<sup>342</sup> que je reçois en même temps que le n° de *Compositio math.* qui contient votre article en esperanto.<sup>343</sup> Il me semble reconnaître un ou deux articles déjà lus, mais j’en ignorais le plus grand nombre, et les lirai avec intérêt.

Je ne sais pas dans quels délais une 2e édition est possible pour ce genre de livre. Je vous signale à tout hasard une faute d’impression, p.154, Note (1). Il faut lire 1933 et non 1923.<sup>344</sup>

J’ai bien reçu votre lettre sur l’intégrale de Hellinger,<sup>345</sup> et vous en remercie.

Je profite de l’occasion pour vous parler d’autre chose. Je viens de recevoir la visite de M. Vilentchuk de l’Université de Haïfa (Israël). Il cherche un professeur, d’un niveau scientifique élevé, pour prendre la direction de la partie mathématique du “Technion”, Israel Institute of Technology, Haïfa. Il dit que le niveau des étudiants, qui ne viennent là qu’après leur service militaire, et sont de ce fait très mûrs (le service est de 2 ans 1/2), est très élevé.<sup>346</sup> Ils demandent un engagement de 2 ans ; s’ils ne trouvent pas, ils accepteraient peut-être un engagement de 3 semestres en 3 ans. Cela serait bien payé; 600 ou 700 dollars américains par mois, +200 £ israéliennes qui doivent suffire pour la vie courante, plus le logement et le voyage (famille

<sup>340</sup>Daniel Dugué was close to both Fréchet and Lévy. He organized and edited Lévy’s *Œuvres Complètes*.

<sup>341</sup>On May 14, 1956, Fréchet was elected to the Geometry Section of the Academy of Sciences. He was the sole first-tier candidate, Lévy the only one in the second tier, with Cartan and Mandelbrojt equally ranked in the third. A fourth tier was comprised of Jean Favard and Laurent Schwartz, and Paul Dubreil was added by the Academy after these four groupings. One round of votes was sufficient, and Fréchet won with 54 votes out of 63 (Lévy received 5, Cartan 1, Dubreil 1, and there were 2 blank ballots).

<sup>342</sup>Fréchet’s book (Fréchet 1955) remains today his best known work by the general public. It takes up several of his favorite themes: the foundations of mathematics, the philosophy of science, the theory of probability, the teaching and applications of statistics, and the usefulness of these disciplines in the social sciences. Fréchet offered an exemplary of the book to the Academy of Sciences, and it was reviewed in CRAS on January 30, 1956.

<sup>343</sup>Fréchet published an impressive number of articles in Esperanto. The present one refers to Fréchet, Maurice (1954) La kanonaj formoj de la 2–3–4 dimensiaj paraanalitikaj funkcioj. *Compos. Math.* 12:81–96 (with a resume in French).

<sup>344</sup>The typo refers to the year when Kolmogorov’s *Grundbegriffe* appeared.

<sup>345</sup>The integral named for Hellinger (1907) was transformed into a Stieltjès integral by Hahn in 1912. The content of the letter from Fréchet to Lévy is not known.

<sup>346</sup>Israel was then in a state of alert. In July 1956 Nasser announced the nationalization of the Suez canal. The Israeli army intervened on October 29 and Franco-English troops arrived on November 5.

comprise, éventuellement). Il estime que quelqu'un qui viendrait avec femme et 2 enfants pourrait mettre de côté environ 500 \$ U.S.A. par mois. Enseignement en anglais ; pas de condition religieuse ou raciale. Age optimum 30 ou 40 ans.

Si vous aviez une idée, que l'intéressé prenne contact avant samedi avec M. Vientchuk, Hôtel Balmoral, 6 rue du Général Lanrezac, Etoile 30.50, ou après samedi écrive à Dr M.Lévy,<sup>347</sup> Secretary for Academic Staff, Technion Israël Inst.of technology. Haïfa.

Bien cordialement

P.Lévy

## Lettre 78

L'Auberdrière - par Baugé - Maine et Loire

2 avril 1958

Mon cher collègue,

Les vacances de Pâques me donnent l'occasion de regarder tranquillement le volume "d'hommage" que Madame Soury<sup>348</sup> m'a remis récemment, et je tiens à vous remercier tout spécialement de votre belle contribution à ce volume.

Je me sens d'ailleurs spécialement confus de n'avoir pas fait pour vous ce que vous avez fait pour moi. J'aurais voulu le faire, mais je ne suis pas assez souple. Quand Villat m'a demandé de participer à un volume qui vous serait dédié,<sup>349</sup> mon esprit était polarisé vers une tentative de théorie des fonctions à corrélation linéaire - qui finalement n'a pas donné tout ce que j'espérais - et je n'ai pas pu mieux faire que d'envoyer un travail sur ce sujet.

J'ai ainsi pu apprécier ce que vous aviez fait, en prenant comme point de départ de votre mémoire ce que vous appelez ma première définition de la distance de deux lois de probabilité (en fait ce n'était pas la première).<sup>350</sup> Vous l'avez d'ailleurs généralisée, en ce sens que je ne l'avais associée qu'à une de vos définitions de deux variables aléatoires (au moins je crois me le rappeler ; je n'ai pas ici le texte sous les yeux). Et j'ai beaucoup admiré comment avec votre quatrième définition, vous arrivez à faire quelque chose de maniable d'une idée qui pour moi était surtout théorique, vu la difficulté de déterminer le minimum de la distance de deux variables aléatoires ayant les répartitions marginales données.

Votre note du 14 mai 1956 m'avait frappé comme un exemple d'une de ces choses très simples à côté desquelles tant de chercheurs passent sans les voir ;<sup>351</sup> c'est

<sup>347</sup>Lévy automatically put an accent on the "e" in this name and then crossed it out.

<sup>348</sup>Madame Soury, an IHP employee, was known for her agreeable and courteous service. The teas that she prepared after each seminar have acquired legendary status at the IHP.

<sup>349</sup>Two volumes of the *Journal de Mathématiques Pures et Appliquées* (1958.37; 1959.38) were dedicated to Fréchet. Lévy's contribution was Lévy, Paul (1959) Sur quelques classes de fonctions aléatoires. *J. Math. Pures Appl.* 38:1–23.

<sup>350</sup>See footnotes to Letters 65 and Sect. 4.

<sup>351</sup>Fréchet, Maurice (1956) Sur les tableaux de corrélation dont les marges sont données. *CRAS* 242:2426–2428. Given the marginals  $F(x)$  and  $G(y)$ , Fréchet showed that (except in the trivial case) there is an infinity of possible correlations. He also used this article once again to condemn the unjustified use of linear correlation coefficients in statistics.

encourageant pour les jeunes de constater qu'il y en a encore. Aussi j'ai été spécialement intéressé de voir dans cette idée simple le point de départ d'une théorie nouvelle.

Bien cordialement à vous - et encore merci

P.Lévy

P.S.- Je crois que vous n'étiez pas il y a 2 ans aux conférences de Kolmogorov.<sup>352</sup>

Il a donné une définition beaucoup plus stricte de la convergence d'une suite de lois vers une limite, basée sur la définition suivante de la distance : c'est le maximum de la différence des probabilités  $P(E)$  et  $P^*(E)$  pour un même ensemble  $E$ .<sup>353</sup>

Cela me semblait trop restrictif pour être utile. Or, avec cette définition, il a énoncé le théorème suivant, de Prohorov : "à toute loi  $\mathcal{L}$  correspond une loi indéfiniment divisible  $\mathcal{L}'$  telle que la distance  $r(\mathcal{L}^n, \mathcal{L}'^n)$  tende vers zéro pour  $n$  infini." Si ce n'est pas tout à fait cela, cela y ressemble beaucoup. C'est en tout cas un énoncé très surprenant. . . il en a donné d'autres<sup>354</sup> non moins surprenants, dans son cours de cette [*En marge et à la verticale* : année. J'ai regretté de ne pas vous voir à sa 1ère et à sa 3ième leçons. Je n'ai d'ailleurs pas pu continuer à le suivre.]

## Lettre 79

L'Auberdrière - par Baugé

9 avril 1958

Mon cher collègue,

J'avoue ne pas très bien comprendre l'énoncé qui vous embarrasse. Etant donné un ensemble (fini ou non) de variables aléatoires indépendantes  $X_n$ , comment définit-on l'espace des événements élémentaires qui lui correspond ?

Quand vous parliez de catégories d'épreuves, j'ai toujours compris qu'il s'agissait de ces considérations de bon sens qui s'intercalent entre le calcul des probabilités et ses applications. Mais si cette notion intervient dans un énoncé mathématique, si la catégorie d'épreuves devient un ensemble défini avec précision, je demande à connaître cette définition.<sup>355</sup>

Peut-être l'énoncé qui vous étonne et à première vue me paraît étrange vient-il simplement d'un malentendu sur la définition.

Bien cordialement à vous

P.Lévy

---

<sup>352</sup>Kolmogorov came to France in 1956, his first visit since 1931. He gave several talks at the Institut Henri Poincaré.

<sup>353</sup>Convergence in Prohorov's sense. In the case of distribution functions, this implies uniform convergence.

<sup>354</sup>Prohorov's paper Prokhorov, Yuri V. (1956) Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.* 1:177–238 was published in 1956 by the Academy of Sciences of the USSR. In this work Yuri Vasilievich Prohorov (1929–2013) posed the basic conditions for the theorem proving the existence of a convergent subsequence in a sequence of bounded measure.

<sup>355</sup>Fréchet's *catégorie d'épreuves* is very close to today's space  $\Omega$ , the space that Lévy doggedly avoided (see Letter 63 and Sect. 4).



Je ferai le 24 une conférence sur les processus strictement markoviens. J'exposerai quelques résultats qui me semblent bien simples ; or je ne suis pas d'accord avec Chung, Doob, et beaucoup d'autres qui, sans contester les résultats, appellent heuristiques des raisonnements qui me paraissent rigoureux.<sup>356</sup> Votre avis m'intéresserait.

D'ailleurs vous recevrez le texte- si vous préférez lire plutôt qu'écouter, ou si vous n'êtes pas libre le 24, je serai heureux d'avoir votre avis ensuite.

P.L.

### **Lettre 80**

Paris - 38 Av. Th.Gautier

18 avril 1958

Mon cher collègue,

J'ai corrigé hier les épreuves de ma conférence du 24 avril, et demandé à Madame Soury de vous l'envoyer dès qu'elle serait tirée.

Vous n'aurez qu'à lire l'introduction, et le n° 7 (intitulé "instant initial de Chung") pour vous rendre compte de la nature du débat.<sup>357</sup>

J'ai ajouté le n° 8 pour répondre à une objection que vous m'aviez faite à propos de mes variables aléatoires. Il ne s'agit que d'un théorème classique, et je ne crois pas que pour Chung la difficulté vienne de ce théorème ; les probabilistes ne se sont peut-être pas encore donné la peine de le démontrer ; mais les analystes l'avaient fait (d'après Schwartz).

Ceci répond à la première partie de votre lettre, qu'heureusement j'avais lue hier avant de prendre mon journal. Comme je vous l'ai dit hier, je l'avais mise de côté pour lire plus attentivement la partie mathématique, et elle a disparu mystérieusement. Je ne peux accuser qu'un geste inconscient qui m'a fait la placer en quelque lieu bien sûr où je la retrouverai un jour. Mais j'ai beau me casser la tête, je ne trouve pas, et me trouve dans l'impossibilité de vous répondre.

Si vous trouvez que cela vaut la peine de récrire cette seconde partie de votre lettre, je tâcherai de ne pas égarer cette seconde édition.

Toutes mes excuses - et bien cordialement

P.Lévy

### **Lettre 81**

Paris - 38 Av. Théophile Gautier

2 janvier 1959

Mon cher ami,

Permettez-moi d'abord de vous souhaiter une bonne année. En ce qui me concerne, je ne m'inquiète pas trop de mon état pulmonaire. Avec la thérapeutique moderne, on doit me guérir assez vite. Mais un grand repos est d'autant plus indispensable que cela fait suite à une grippe qui s'est portée successivement sur tous les points faibles de mon organisme.

---

<sup>356</sup>See the next letter.

<sup>357</sup>Chung's optional starting times play a central role in the distinction between Markov processes and strictly Markov processes. Lévy explained these properties in Lévy, Paul (1959) Processus strictement ou presque strictement markoviens. Compos. Math. 14(2):172–193.

Je vous envoie une Note que je vous serais obligé de présenter à l'Académie.<sup>358</sup> Je serais content que vous y jetiez un coup d'œil. Il s'agit d'un paradoxe qui m'intrigue depuis assez longtemps ; je crois avoir trouvé enfin la bonne manière de le présenter. Ma note commence par un exemple très simple, afin que le lecteur soit convaincu que ces éléments aléatoires paradoxaux existent réellement.

J'ai transmis votre lettre à J.P. Kahane, 11 rue du Val de Grâce (chez son père) et su qu'elle était bien arrivée. Je pense qu'il vous a récrit, pour quelques modifications que je lui avais conseillé d'apporter à sa rédaction ; au moins pour corriger une erreur matérielle dans une des formules.<sup>359</sup>

Bien cordialement à vous

P.Lévy

## Lettre 82

21 mars 1937<sup>360</sup>

Mon cher collègue,

J'ai bien reçu votre lettre jeudi, et vendredi le livre de Borel.<sup>361</sup> Je vous remercie et tacherai de ne pas le garder trop longtemps. Je n'ai pas beaucoup travaillé depuis 5 jours ; une grippe à début brutal, et deux séries d'épreuves à corriger. J'ai tout de même regardé le chapitre.<sup>362</sup>

---

<sup>358</sup>Lévy, Paul (1959) Un paradoxe de la théorie des ensembles aléatoires. CRAS 248:181–184 was presented on January 5, 1959, but published with the session of January 12. Lévy constructed non-trivial probability distributions with no atoms on points of its range but for which all events have probability 0 or 1. This has the consequence that one cannot consider the probability even of some very simple events involving such variables.

<sup>359</sup>Kahane, Jean-Pierre (1959) Sur le recouvrement d'un cercle par des arcs disposés au hasard. CRAS 248:184–186 gives necessary and sufficient conditions for an infinite series of arcs of independently given lengths placed at random on the unit circle to cover it a.s. This note was presented by Fréchet after Lévy's one.

<sup>360</sup>This date, written in Lévy's hand, is the greatest mystery of the entire correspondence. This letter can only have been from the end of 1960 or the beginning of 1961, because it fits perfectly with the letters from that period. One possibility, perhaps a little far-fetched, is that Lévy used an old piece of letter paper that had been dated then put aside. The ink used to write the date and that used for the letter itself may indeed be different.

<sup>361</sup>Borel, Emile (1949) *Eléments de la théorie des ensembles*. Albin-Michel, Paris. The parts questioned by Lévy are on sets of zero measure, Chap. VI, pp. 152–198, and possibly (see the following letters on the same subject-rarefaction of sets of zero measure) in the part on the additivity of sets of measure zero, pp. 287–313.

<sup>362</sup>This question, which Borel called (at least after 1935) the rarefaction of sets with zero measure, had in fact appeared in Borel's publications since 1913, first under the name "asymptotic order of a set with measure zero" (see Borel, Emile (1913) *Les ensembles de mesure nulle*. Bull. Soc. Math. Fr. 41:1–19, subsequently in 1919 as 'asymptotic measure'—see Borel, Emile (1919) *Sur la classification des ensembles de mesure nulle*. Bull. Soc. Math. Fr. 47:97–125. The verb *se raréfier* for a set of zero measure appeared for the first time, without formal definition, in 1935, in Borel, Emile (1935) *Sur les ensembles de mesure nulle*. Fundam. Math. 25:7–12. The noun *raréfaction* seems to have been used for the first time in 1948 CRAS 227:453–455. Borel apparently navigated between several definitions of this notion, leading to the ambiguities for which Lévy and Fréchet took him to task. Fréchet gave an excellent exposition of this question in the extensive review

D'abord à propos de Borel, je fais amende honorable ; il a bien dit (p.164) qu'il fallait choisir la série qui converge le plus rapidement possible. Ce qu'il ajoute sur la possibilité de ce choix me semble inexact ; vous même avez exprimé des doutes par une remarque en marge.

Il y a d'ailleurs dans ce chapitre plusieurs choses douteuses ; p.152, à un endroit que j'ai marqué par un 8 au crayon, il donne comme sûr un énoncé qui équivaut au théorème du continu. Il y a une confusion plus grave p.163–165. Il considère un ensemble parfait discontinu  $E$  qu'il appelle ensemble ternaire de Cantor (je l'avais cru de Lebesgue), recouvert par la réunion  $E_n$  de  $2^n$  intervalles  $I_{n,v}$  de longueurs  $1/3^n$  et il confond la convergence de la série  $\sum m(I_{n,v})$  avec celle de  $\sum m(E_n)$ ,  $m(\cdot)$  désignant la mesure. Il définit en effet la raréfaction par la rapidité de la convergence de  $\sum m(I_{n,v})$ , tandis qu'il faut considérer  $\sum m(E_n)$  pour dire comme il le fait, que cette série converge comme une progression géométrique de raison  $\frac{2}{3}$ .

Ce qu'il fait sans le dire me paraît d'ailleurs intéressant, et peut se généraliser comme suit. Pour chaque  $\varepsilon > 0$ , je cherche le minimum  $\mu(\varepsilon)$  de la mesure d'une réunion de segments de longueurs  $\geq \varepsilon$ , cette réunion recouvrant  $E$ . Cette fonction  $\mu(E)$  est une caractéristique intéressante de  $E$ , dans le cas où la fermeture  $\bar{E}$  est un ensemble parfait dis-continu<sup>363</sup> (de mesure nulle).

On peut établir à son sujet des règles simples :

$$\mu(\varepsilon, E \cup E') \leq \mu(\varepsilon, E) + \mu(\varepsilon, E').$$

En se donnant une suite  $\{\varepsilon_n\}$  de valeurs décroissantes de  $\varepsilon$ , et si  $m(E) = 0$ , on peut trouver une série majorante de  $E$  (avec votre définition),  $\sum E_n$  ( $E_n$  formé avec  $\varepsilon_n$ ). C'est une suite de ce genre que considère Borel dans l'exemple qu'il étudie. Or, plus la suite des  $\varepsilon_n$  tend rapidement vers zéro, plus  $\sum m(E_n)$  converge rapidement, de sorte qu'au point de vue de Borel il n'y pas de série convergeant plus rapidement que toutes les autres. Les choses sont un peu différentes avec le vôtre, qui consiste à ranger tous les  $I_{n,v}$  en une série unique. Le nombre  $\varphi(n)$  des  $I_{n,v}$  qui composent  $E_n$ , et  $\mu(\varepsilon_n)$ , sont des fonctions bien déterminées de  $\varepsilon_n$ . Or si  $m(I_{n,v}) = m_n^*$  (indépendant de  $v$ ), le reste de la série  $\sum_{k,v} m(I_{k,v})$  arrêtée juste avant  $m(I(n, v))$  est  $> m(E_n)$ , et chaque terme de plus ne peut le diminuer que de  $m_n^*$ . La rapidité de convergence a donc une borne supérieure.

J'ai cru d'abord pouvoir baser là dessus la démonstration du fait qu'on pouvait limiter supérieurement la rapidité de la convergence des séries majorantes de  $E$ , dans le cas des ensembles de nombres n'ayant que les décimales 0 et 9 (ou 0 et  $p-1$  dans le système en base  $p$ ). En prenant ensuite les sous ensembles pour lesquels la décimale 9 (ou  $p-1$ ) n'est possible que pour les  $n$  d'une suite  $\{n_k\}$  rapidement croissante, on aurait une suite transfinie d'ensembles de plus en plus en plus raréfiés (avec vos définitions).

---

that he consecrated to the life and work of Borel in his introduction to Œuvres complètes d'Emile Borel. CNRS, Paris (4 volumes, 1967–1972). See pp. 60–65, the paragraph entitled 'Raréfaction d'un ensemble de mesure nulle.'

<sup>363</sup> *Sic*.

Mais cela ne marche pas. D’abord il faut que  $\mu(\varepsilon, E) \rightarrow 0$  avec  $\varepsilon$ , ce qui implique que  $\bar{E}$  soit un ensemble parfait discontinu, de mesure nulle. C’est donc une théorie inapplicable pour l’ensemble des nombres dont la représentation décimale ne comprend qu’un nombre fini de 7. Même quand elle est applicable, il n’est pas évident que les  $\sum(I_{n,v})$ , déduits des séries  $\sum E_n$  dans lesquelles chaque  $E_n$  est une réunion finie d’intervalles recouvrant  $E$  soient celles qui donnent la convergence la plus rapide\*.

C’est pour cela qu’après avoir cru à une idée intéressante, je deviens sceptique sur son intérêt.<sup>364</sup> Mais comme cela touche aux questions qui vous intéressent, je vous en fait part.

Je veux y réfléchir un peu encore. Puis je vous renverrai votre note et le livre de Borel, ou vous le rapporterai. Je n’ai pas revu M. Hadamard depuis mon accident,<sup>365</sup> et en suis un peu confus ; j’espère le voir d’ici un mois et cela ne sera pas difficile de passer chez vous.

Bien cordialement

P.Lévy.

\* Ainsi, je pourrais imaginer une suite telle que

$$\begin{array}{ll}
 I_{2,1}; I_{2,2}(\text{et non } I_{2,3} \text{ et } I_{2,4}); & \text{il manque donc encore} \\
 I_{3,5}, I_{3,6} (\text{dont la réunion est } I_{2,3}); & I_{2,4} \\
 I_{4,13}, I_{4,14} (\dots\dots\dots I_{3,7} & I_{3,8} \\
 I_{5,27}, I_{5,28} (\dots\dots\dots I_{4,15}) & I_{4,16}
 \end{array}$$

et ainsi de suite, de manière à reculer indéfiniment les intervalles qui terminent le premier recouvrement de  $E$ , en diminuant leur longueur totale. C’est évidemment moins rapidement convergent que

$$I_{0,1} + 0 + 0 + 0 \dots$$

ou

$$I_{1,1} + I_{1,2} + 0 + 0 + \dots$$

Aussi je ne crois pas que des procédés de ce genre puissent arriver à améliorer la convergence. Mais ce n’est pas évident du tout.

<sup>364</sup>Later, Lévy became less skeptical. He dedicated a substantial part of his 1961 to 1964 correspondence to this subject (see Letters 83, 89, 90, 91, 93, 98, 99 and 102), and made considerable effort to render the mathematical status of a related notion more precise and more useful under the name *probabilités plus ou moins nulles*. His last two publications were on this subject: Lévy, Paul (1969) Esquisse d’un calcul des probabilités plus ou moins nulles. Rev. Roum. Math. Pures Appl. 14(6):813–818 and Lévy, Paul (1969) Une hiérarchie des probabilités plus ou moins nulles. Application à certains nuages de points. Enseign. Math. 15:217–225.

<sup>365</sup>Lévy was knocked down by a car on November 25, 1960 (see Letter 83).

**Lettre 83**

Paris - 38 Avenue Théophile Gautier.

3-1-61

Mon cher Collègue,

Je m'excuse de ne pas avoir répondu plus tôt à vos dernières lettres. La mort d'une sœur de ma femme, indépendamment de mon état de santé, en sont la cause. Peut-être n'avez vous pas su que j'ai été renversé par une voiture le 25 Novembre. Je ne crains aucune infirmité permanente, mais je suis très contusionné et marche péniblement. Je croyais vous l'avoir écrit, mais comme vos lettres n'y font pas allusion, je dois me tromper.

Je joins à ma lettre un texte en français de ma communication au 4ième symposium de Berkeley (1960).<sup>366</sup> Ayant maintenant la traduction anglaise, qui doit être imprimée, je peux m'en dessaisir.<sup>367</sup> Ce n'est pas exactement la question qui vous intéresse ; il s'agit de l'extension de la notion de mesure à certains ensembles non mesurables ; extension évidemment arbitraire, comme l'est la notion de raréfaction de Borel. Si cela ne vous intéresse pas ou quand vous l'aurez lu, pourrez vous déposer ce papier avec un mot d'explication chez Henri Cartan. Cela doit l'intéresser, car je cite un théorème de Choquet et Cartan, résolvant la question que je m'étais posée, et qui ne sera peut-être pas publié autrement.

Ceci dit, je reviens à la raréfaction. Comme il y a dans vos lettres plusieurs points que je ne comprends, je me dis que la raison en est peut-être que j'ai mal lu Borel, et comme je ne peux pas en ce moment retourner à l'Institut H.P. (où j'espère d'ailleurs pouvoir faire le 20 janvier une conférence primitivement annoncée pour le 16 décembre), je pense que le mieux est de vous expliquer exactement mes idées à la suite de cette lecture, bien ou mal faite.

Soit  $E$  un ensemble, que je suppose pour fixer les idées intérieur au segment  $[0, 1]$  de l'axe réel. Indépendamment de tout système de numération, il peut arriver qu'il soit intérieur à un ensemble  $\overline{E}$  (sa fermeture) obtenu en enlevant de  $[0, 1]$  une suite d'intervalles  $i_1, i_2, \dots, i_n, \dots$ , de longueurs  $l_1, l_2, \dots, l_n, \dots$ , la mesure de l'ensemble  $E_n$  obtenu en enlevant  $i_1, i_2, \dots, i_n$  est  $\mu_n = 1 - (l_1 + l_2 + \dots + l_n)$ . On a évidemment  $E_n \supset E_{n+1} \supset \dots \supset \overline{E} \supset E$  et cela revient au même de parler de la rapidité de la convergence de  $\sum l_n$  ou de la rapidité de décroissance de la suite des  $\mu_n$ .

Il faut d'ailleurs supposer  $\lim \mu_n = m(\overline{E}) = 0$ . Autrement  $\overline{E}$  qui est mesurable, serait de mesure positive, et rien ne distinguerait  $E$  d'un ensemble de mesure positive. Or la raréfaction prétend établir une distinction parmi les ensembles de mesure nulle.\*

Ceci dit, on obtient une définition qui n'a rien d'arbitraire en supposant que les extrémités des  $i_n$  appartiennent à  $\overline{E}$  (ce qui exclut l'intervention d'un système de

<sup>366</sup>Lévy, Paul (1960) An extension of the Lebesgue measure of linear sets. In: Fourth Berkeley Symposium. pp. 273–287.

<sup>367</sup>This text, part handwritten and part typed, is in the *Fonds Fréchet* of the archives of the Academy of Sciences.

numération donné d'avance) et en rangeant les  $i_n$  dans l'ordre. On peut d'ailleurs définir la raréfaction en se donnant  $\mu_n$  en fonction de  $l$ , ou de  $l_n$ , ou bien  $l_n$  en fonction de  $n$ .

Passons maintenant au cas où l'on introduit avec Borel la numération décimale. On introduit les intervalles  $i'_{p,h} = (\frac{h-1}{10^p}, \frac{h}{10^p})$  qui ne sont pas des intervalles  $i_n$  ci dessus. Chaque  $i_n$  à extrémité non décimale est la réunion d'une infinité de  $i'_{p,h}$ ; de même si une seule des deux extrémités est décimale. Ainsi le choix du système de numération a pour un même ensemble  $e$  une importance essentielle. Et le système de notation qui consistait tout à l'heure à définir  $\mu_n$  en fonction de  $n$ , ou bien de  $l_n$ , prend une forme différente.  $l_n$  est remplacé par  $\frac{1}{10^p}$ , mais  $n$  l'est par l'expression moins simple  $10^{p-1} + h$ .

Je précise ensuite ce que je disais dans ma dernière lettre de la comparaison entre les systèmes de base 10 et 100. Se donner les  $p$  premières décimales de  $x$  dans le système à base 100 revient à se donner  $2p$  décimales dans le système à base 10. Prenons alors pour  $E$  l'ensemble des nombres qui dans le système à base 10 s'écrivent sans faire usage du chiffre 7. Quand on connaît  $q = 2p$  premières décimales,  $E \subset E'_p$ , réunion des intervalles distingués qui contiennent les points de  $E$ , est  $(\frac{9}{10})^{2p} = (\frac{9}{10})^q$ . Cette mesure est bien déterminée, quel que soit le système de numération. Mais on n'a pas la même fonction suivant qu'on l'exprime en fonction de  $p$  ou de  $q$ . Avec le système à base 10, il est naturel de l'exprimer par  $\varphi(q) = (\frac{9}{10})^q$ . Mais avec le système à base 100, c'est  $\psi(p) = (\frac{9}{10})^{2p} = \varphi^2(p)$ . C'est ce que j'avais voulu dire en disant que, dans Borel, la raréfaction dépendait du système de numération. Si vous voulez, la mesure n'a pas changé mais la fonction qui l'exprime dépend de la variable.

La question est différente si on prend comme variable la longueur  $\frac{1}{10^p}$  des derniers intervalles enlevés. En effet, si  $\varphi(2p) = \omega(\frac{1}{10^{2p}})$ , on a  $\frac{\log \omega(\lambda)}{\log \lambda} = -9 \frac{\log \frac{9}{10}}{\log 10}$ , et il importe peu qu'on remplace  $\lambda$  par  $q$  ou par  $2p$ . La seule différence est que, pour le système à base 10, on a à interpoler les  $\omega(\frac{1}{10^{2p}})$  par les  $\omega(\frac{1}{10^{2p+1}})$ .

Cette remarque faite, ne considérons que le système à base 10, et soit  $\mathcal{E}_k$  l'ensemble des nombres de  $[0, 1]$  comprenant  $k$  décimales quelconques suivies de décimales autres que 7. Si  $p$  est le nombre de décimales, pris comme variable, la fonction de raréfaction est  $(\frac{9}{10})^{q-k}$  ( $q \geq k$ ). Mais pour l'ensemble de tous les  $\mathcal{E}_k$  ( $k = 0, 1, \dots$ ), je ne vois pas comment on peut définir la raréfaction. Il est partout dense, et il est impossible (contrairement à ce que vous semblez croire) de l'enfermer dans un ensemble  $\bar{\mathcal{E}}$  obtenu en enlevant de  $[0, 1]$  des intervalles  $i_n$ .

Voilà ce que je voulais vous dire. Mais peut-être ai je mal lu Borel ou vous ai je mal compris ?

Mes meilleurs vœux pour 1961 et bien cordialement à vous

P.Lévy

L'éloignement et mon accident m'ont empêché depuis 2 mois d'aller voir M. Hadamard. Je pense que vous vous doutez que ce serait une bonne action de monter

quelquefois chez lui, ne fût-ce que pour une 1/2 heure.<sup>368</sup> Cela vous est plus facile qu'à moi.

\* Il en résulte bien que  $E$  ne peut être dense dans aucun intervalle et est de mesure de Jordan nulle. Pour  $n$  assez grand,  $E \subset E_n$  qui est de mesure arbitrairement petite et réunion d'un nombre fini d'intervalles ;  $\overline{E}$  ne contient aucun intervalle.

#### Lettre 84

Paris - 38 Av. Théophile Gautier

9 février 1961

Mon cher collègue,

Il me sera difficile de venir demain à votre conférence, et je vous prie de m'excuser.

J'ai bien reçu votre lettre du 7.2., et ne vous ai pas répondu tout de suite parce que je crois vous avoir dit l'essentiel de ce que je voulais vous dire, et je pense que l'avez maintenant bien compris.

Je ne me suis jamais étonné que vous ne confondiez pas une courbe et une fonction. Ce que j'ai dit c'est qu'en disant que vous ajoutez des courbes, vous risquez de faire oublier que ce sont des fonctions que vous ajoutez, ou du moins que l'intermédiaire des fonctions est nécessaire, et que cet oubli fait apparaître comme paradoxaux des faits qui sont bien simples si on souvient des définitions.

Ainsi, en conservant des notations de mes premières lettres, soit  $x_p + y_p = z_q$ , je ne m'étonne pas que  $z_p - y_p$  ne corresponde pas à la même courbe que  $x_p = z_q - y_p$ . Je m'en étonne si on écrit que  $x + y = z$  n'entraîne pas  $x = y - z$ .

Mais ce sont de questions de définition, et, je vous le répète, vos conclusions, que je n'avais pas bien comprises à votre conférence mais que j'ai comprises le soir en regardant votre texte dactylographié, sont tout à fait correctes.

Je ne crois pas que pour les courbes continues plus générales, on puisse, comme pour les courbes rectifiables, trouver un paramètre intrinsèque qui s'impose.

Je vous signale dans cet ordre d'idées le fait suivant, assez curieux. Soit la courbe du mouvement brownien (linéaire, ou plan, mais donnée en faisant abstraction du paramètre) ; je veux dire "soit une réalisation de cette courbe". Soit d'autre part une horloge. J'ai donné dans mes "processus stochastiques" (n° 41, pour le cas linéaire ; p.259 pour le plan) un théorème qui permet de reconnaître presque sûrement si l'horloge marche bien,<sup>369</sup> et, dans le cas contraire, de la corriger pour avoir le temps vrai. D.G. Kendall est revenu récemment sur cette question. Mais il ne s'agit que d'un résultat presque sûr. Etant donnée la courbe, un malin génie pourrait nous proposer une horloge, c'est-à-dire une paramétrisation initiale, qui mette ce résultat en défaut. J'ai cherché sans succès une pararamétrisation intrinsèque dont je sois sûr qu'elle évite les pièges du malin génie. Je n'ai pas trouvé et Kendall n'a pas non plus résolu cette question.

---

<sup>368</sup>Hadarnard was by then 95, and Lévy probably wanted to resume his longstanding relations with his mentor.

<sup>369</sup>Lévy is alluding to his Theorem 41.2, which mentions that what he calls *oscillation brownienne* (Brownian oscillation), i.e. oscillation in quadratic mean, is a.s. equal to  $t$ .

Bien cordialement à vous  
P.Lévy.

### Lettre 85

12 mars 1961

Mon cher collègue,

Je n'ai reçu qu'hier votre note et votre lettre du 8 mars. Je vous écris tout de suite un mot pour le cas, peu probable, où il serait encore temps de corriger une faute d'impression qui semble vous avoir échappé. Au bas de la p.2, formule qui suit la formule (2), on a écrit Rar au lieu de Rap.

Vous avez incontestablement réalisé un grand progrès par rapport à Borel, auquel je reprochais surtout de n'avoir considéré qu'un recouvrement particulier lié à la numération décimale, sans voir qu'il s'agissait de chercher un recouvrement convergent aussi rapidement que possible (ou du moins une borne supérieure, atteinte ou non, pour la rapidité des recouvrements). Tant sur ce point, que sur la suppression des restrictions non essentielles (qui est toujours une idée essentielle dans votre œuvre scientifique), vous avez réalisé de grands progrès, et je vous félicite d'autant plus sincèrement qu'il est rare qu'un mathématicien fasse du bon travail passé 80 ans.

Je n'ai encore rien à vous dire en réponse à vos questions. La première me semble très importante, puisque l'importance de votre théorie, qui semble d'aspect très séduisant, dépend de la réponse qu'on doit y faire. Mais mes premières réflexions ne m'ont conduit qu'à la conclusion que c'est un problème difficile. Je veux tâcher d'y réfléchir encore et vous réécrirai s'il y a lieu. En tout cas je vous renverrai votre note avant la fin de la semaine.

Je ne vous promets rien pour les autres questions, qui exigent un travail à l'institut H.P., à moins que je ne puisse trouver à l'X le livre de Borel. Je croyais l'avoir, mais décidément je ne le retrouve plus.

Encore une remarque, dont je ne veux pas exagérer l'importance. J'ai pris l'habitude de dire que  $\sum u_n$  converge rapidement si  $c > 0$  et  $q < 1$  existent tels que  $|u_n| < cq^n$  ; en d'autres termes si  $\sum u_n z^n$  a un rayon de convergence  $> 1$  ; que  $\sum u_n$  converge beaucoup plus rapidement que  $\sum v_n$  si  $\sum u_n z^n$  a un rayon de convergence  $>$  celui de  $\sum v_n z^n$ . Cela ne s'accorde pas avec vos définitions. Je considérais d'ailleurs que convergence plus rapide signifiait  $u_n/v_n \rightarrow 0$  et je considérais les deux rap. de convergence comme égales (ou presque égales) si  $cu_n \leq v_n \leq c'u_n$ .

Pour exprimer vos idées avec ce langage, j'aurais parlé de convergence uniformément plus rapide (ou beaucoup plus rapide) qu'une autre. Ainsi  $\sum u_n$  convergerait beaucoup plus rapidement, et cela uniformément, que  $\sum v_n$ , si  $\sum \frac{u_n}{v_n} x^n \rightarrow 0$ <sup>370</sup> pour  $|x| < 1 + \varepsilon$  ( $\varepsilon > 0$  arbitrairement petit).

Evidemment, je ne peux ni vous demander de changer pour cela une note déjà composée, ni m'étonner que vous ne connaissiez pas un langage que j'ai seulement employé quelquefois à l'X. J'attache plus d'importance à ma critique sur votre définition de la somme de deux courbes. Je ne pense pas y revenir, mais n'exclus tout

<sup>370</sup>The summation sign  $\sum$  is superfluous.



de même pas absolument de le faire, quand votre mémoire des Ann.Ec.Norm aura paru.

Bien cordialement à vous P.Lévy.

J'ai écouté vendredi avec intérêt une conférence d'Hennequin<sup>371</sup> au séminaire Fortet. Il a enfin les éléments d'une bonne thèse. Il avait suivi mon cours, une des 2 ou 3 années où je l'avais fait à l'institut H.P. et m'avait paru avec Mandelbrot un des meilleurs auditeurs. J'ai été surpris du temps qu'il lui a fallu.

J'ai maintenant parlé à Fortet, Dugué, Loève, Hennequin et Tortrat de la difficulté qui m'arrête, et j'espère que l'un d'eux me dépannera. C'est vexant d'avoir deux raisonnements qui se contredisent et de ne pas trouver la faute.

## Lettre 86

Paris - 38 Av. Théophile Gautier

31 mars 1961

Mon cher collègue,

Les précisions que vous me donnez prouvent que l'énoncé de Borel est faux.<sup>372</sup>

La convergence de

$$\int_a^\infty e^{-a}[x'(a) - x(a)] da = [e^{-a}x(a)]_0^\infty$$

équivalait à  $e^{-a}x(a) \rightarrow$  limite (pour  $a \rightarrow +\infty$ ).

Prenons par exemple

$$\begin{aligned} x(a) &= e^a + \sin e^a \\ x'(a) &= e^a + e^a \cos e^a \end{aligned}$$

Alors

$$f(a) = e^{-a}[x'(a) - x(a)] = \cos e^a - e^{-a} \sin e^a$$

n'a pas de limite.

Pour  $x(a) = e^a + \sin e^{a^2}$ , les oscillations seraient encore plus fortes.

---

<sup>371</sup>Shortly thereafter, Paul-Louis Hennequin became professor at the University of Clermont-Ferrand. The French community of probabilists has his enthusiasm and generosity to thank for the flourishing Ecole d'Été de Calcul des Probabilités de Saint-Flour (Saint-Flour Summer School in Probability Theory).

<sup>372</sup>Lévy is probably alluding to a somewhat offhand remark on p. 97 of Borel's 1901 book *Leçons sur les séries divergentes* (*Lessons in Divergent Series*), which seems to affirm that the convergence of the integral  $\int_0^\infty \varphi(a)e^{-a} da$  and the fact that  $\varphi$  is positive imply that  $\varphi(a)e^{-a} \rightarrow 0$  as  $a$  goes to infinity. Lévy must have had the first edition rather than the second, which had been completely revised by Bouligand. In the second edition, the statement above was changed and the sufficient condition supposedly proved by Borel was proved to be false.

Pour rendre l'énoncé de Borel exact, il faudrait, je pense, se limiter aux fonctions entières d'ordre  $< 1$ , avec en outre des termes tels que  $c'e^{cx}$  ( $c \leq 1$ ). Mais ce n'était évidemment pas ce qu'il avait dans l'idée.

Je suis d'ailleurs un peu surpris. J'avais autrefois étudié et admiré la sommation exponentielle de Borel, et la faute qu'il a commise m'avait échappé. Elle ne réduit d'ailleurs pas l'intérêt de la définition initiale, et, à première vue, le théorème fondamental sur le polygone de sommabilité subsiste. J'avoue que je n'en ai plus la démonstration présente à l'esprit. Il est étonnant que la faute commise dans la transformation de la définition initiale ait échappé jusqu'ici aux nombreux lecteurs des leçons sur les séries divergentes. Peut-être serait-il utile de recueillir l'avis de Bouligand, qui a collaboré à la 2ième édition.

Bien cordialement à vous

P.Lévy

## Lettre 87

Paris - 38 Av. Th.Gautier - 4 avril 1961

Mon cher collègue,

Je vous renvoie séparément le livre de Borel<sup>373</sup> et votre Note,<sup>374</sup> et vous remercie de me les avoir communiqués.

Je n'ai pas tout lu, mais ce que j'ai lu du livre de Borel m'a suffi pour trouver beaucoup de choses qui m'ont surpris.

p. 44 il donne du point d'accumulation une définition qui n'est pas celle que j'ai toujours vue ;<sup>375</sup> p.152 il donne un énoncé équivalent au théorème du continu,<sup>376</sup> qui, je crois, et quoiqu'en dise Eyraud,<sup>377</sup> n'est pas démontré.<sup>378</sup>

Ce qui m'a le plus gêné, c'est ce qu'il dit des ensembles analytiques, p.220 et suiv. J'avais su ce que c'était, mais j'oublie vite les choses quand je n'ai pas l'occasion d'y repenser, et espérais m'instruire à nouveau. Or il écrit : "Un ensemble analytique est l'ensemble des valeurs que prend une fonction continue [ $x = f(t)$ ]

<sup>373</sup>Borel, Emile (1949) *Théorie des ensembles*. Albin Michel, Paris. which appeared in 1949.

<sup>374</sup>On February 27, 1961, Fréchet presented a note Fréchet, Maurice (1961) *Une généralisation de la raréfaction*.

<sup>375</sup>Was Lévy annoyed by the demand for uncountability of the intersection with a neighborhood? As Borel himself remarked, a series does not have an accumulation point with such a definition.

<sup>376</sup>Borel affirmed that an uncountable set of measure zero has the cardinality of the continuum.

<sup>377</sup>Henri Eyraud, an influential statistician in Lyon before World War II. Eyraud, in particular, played an important role in supporting Gumbel, before Gumbel left for the USA (see the thesis by S. Hertz, University of Lyon I, 1997). During the 1950s Eyraud became interested in the foundations of mathematics. F. Le Lionnais had him write an article about the transfinite for the book *Les Grands Courants de la Pensée Mathématique (Main Themes of Mathematical Thought)*. In this article Eyraud announced, somewhat vaguely, that a proof of the continuum hypothesis was forthcoming. A little later in his correspondence (see note 411) Lévy mentioned that he was to review the book. Perhaps he had already seen Eyraud's claim and found it rash.

<sup>378</sup>The conjecture that there is no cardinality between that of  $\mathbb{N}$  and that of  $\mathbb{R}$ , the first of Hilbert's 23 problems. Its undecidability in Zermelo-Fraenkel set theory was finally proved in 1963 by Paul J. Cohen.

d'une variable  $t$ , lorsque la variable prend elle même les variables<sup>379</sup> comprises entre 0 et 1". C'est donc un intervalle de l'axe des  $x$  (ouvert, semi-ouvert ou fermé). 2 lignes plus loin : "l'ensemble analytique le plus général est donc défini par une relation telle que  $x = f(t) = \sum P_n(t)$ ." Alors maintenant l'ensemble est une fonction !<sup>380</sup> Mais une 1/2 page plus loin c'est de nouveau un ensemble de valeurs de  $x$  ! Ensuite il emploie le procédé diagonal pour définir un ensemble qui n'est pas analytique.<sup>381</sup> Or il suffit de prendre n'importe quel ensemble qui ne soit pas un continu.

Je ne peux pas douter qu'il y ait une erreur dans la définition initiale. Je me demande s'il ne faudrait pas s'entendre avec l'éditeur pour corriger cette erreur dans les éditions futures, s'il doit y en avoir.

J'ai d'ailleurs été surpris par beaucoup d'autres choses, mais moins graves, et qu'on peut hésiter corriger.<sup>382</sup>

L'axiome de Zermelo. Il n'est pas étonnant que Borel se soit durci dans une position qui a toujours été la sienne, mais qui, je crois, n'a aucun adepte parmi les jeunes et qui pour nous est incompréhensible. Mais j'ai été surpris par le mal qu'il se donne p.229–232, pour définir un ensemble dénombrable qu'on ne sache pas dénombrer effectivement.<sup>383</sup> Il fait intervenir la suite des décimales de  $\pi$ . Il n'y a qu'à prendre l'ensemble des entiers qui peuvent être définis par les décimales consécutives de  $\pi$ . Il a toutes les propriétés de celui qu'il considère : on en connaît des éléments, mais en dehors de ces éléments connus, on ne sait pas si un nombre donné en fait partie, de sorte qu'on ne peut pas dire quel rang aura un nombre donné (sauf pour les premiers nombres). Plus simplement encore il n'y a qu'à prendre l'ensemble des exposants entiers pour lesquels le théorème de Fermat est vrai.

Probabilités dénombrables p.80 et suiv. J'ai là un peu une déception personnelle. En lisant autrefois le mémoire de Borel sur ce sujet, j'avais été frappé de la manière dont il démontrait deux lemmes évidents par des méthodes compliquées. Dans mes variables aléatoires (p.126–128), je les démontre bien plus simplement.<sup>384</sup> Il l'a lu, et c'est même lui qui m'a fait remarquer un lapsus (p.128,l.12). J'espérais qu'il abandonnerait ses anciennes méthodes. Or c'est toujours d'elles qu'il se sert en 1949 !

Fractions continues. Impression un peu analogue. Je pensais qu'il avait regardé le chapitre de mes Variables aléatoires où je traite de ce sujet, car, là aussi, il m'avait

<sup>379</sup>*Sic*. In Borel's text this word is *valeurs*.

<sup>380</sup>Lévy is being somewhat insincere. Borel's text is clear, even if it is incorrect (see below)!

<sup>381</sup>Borel defines an analytic set as the continuous image of the interval  $[0, 1]$ . Thus, for example, it is necessarily connected (or, in Lévy's parlance, 'continu'). In fact the analytic sets of  $\mathbb{R}$  are projections on the first coordinate of the Borel sets of  $\mathbb{R} \times [0, 1]$ .

<sup>382</sup>*Sic*.

<sup>383</sup>*Le mal qu'il se donne (the effort he makes)* is a little exaggerated. Borel considered a series of integers formed by the decimals of  $\pi$  between two consecutive 5s.

<sup>384</sup>In his 1937 book, Lévy proved the Borel-Cantelli Lemma in the modern manner, whereas Borel retained his rather clumsy method. Lévy was annoyed that Borel had not taken the trouble to adopt his proof.

fait une observation (celle-là venait de ce qu'il avait mal lu). Or, p.260, il indique comme difficile, mais peut-être pas insoluble, un problème qui est complètement résolu dans mon livre.<sup>385</sup>

Ai-je raison de vous signaler tout cela ? Croyez que j'ai été un grand admirateur de Borel, dès la lecture de ses premiers livres (séries à termes positifs ; séries divergentes ; fonctions entières), et je ne peux pas m'étonner de ce que le Borel de 1949 n'ait plus été celui du début du siècle.<sup>386</sup> Si je vous signale ces erreurs, c'est à cause des problèmes qui se poseraient en cas de réédition.

Bien cordialement à vous

P.Lévy

## Lettre 88

Paris - 38 Av. Théophile Gautier

6 avril 1961

Mon cher collègue,

Je reçois votre lettre au moment où j'allais vous renvoyer le livre de Borel, votre note,<sup>387</sup> et quelques observations sur ce livre, dont plusieurs parties me paraissent regrettables.<sup>388</sup> Alors je ne le fais pas. Je pense aller voir à la fin du mois M.Hadamard et déposerai le tout chez vous.

Je crois que vous n'avez pas bien compris mon énoncé sur l'indépendance possible de  $X$  et  $Y = f(X)$ .<sup>389</sup> Sauf bien entendu si  $\Pr\{X = a\} = 1$ , il est absolument essentiel que la fonction  $f(x)$  ne soit pas mesurable. On peut supposer  $X$  uniformément réparti dans  $(0, 1)$ . Alors l'ensemble des  $x$  tels que  $f(x) < b$  n'est pas mesurable. Mais on peut dans les cas que j'indique prolonger la mesure de Lebesgue de manière que la mesure de cet ensemble soit bien définie et que  $X$  et  $Y$  apparaissent comme indépendants.

J'ai été heureux que mon mémoire vous ait intéressé. Dans ce genre de questions, je me demande toujours si mes résultats méritent d'être publiés. Je n'ai pas les mêmes hésitations lorsqu'il s'agit, par exemple, de mon dernier mémoire de Palerme.<sup>390</sup>

Puisque vous recevrez cette lettre à Messine, voulez-vous leur répéter que j'ai beaucoup regretté de n'avoir pas pu l'année dernière accepter une invitation trop tardive. Nos billets pour le tour de la Sicile dans le sens trigonométrique étaient pris

<sup>385</sup>Chap. IX of *Théorie de l'Addition des Variables aléatoires* is entirely devoted to the application of the theory of probability to continued fractions. Lévy, whom Borel did not even cite, was understandably irked.

<sup>386</sup>*In cauda venenum*. Admirably, Lévy held his barbed comment until the end.

<sup>387</sup>See footnote 374.

<sup>388</sup>See Letter 87.

<sup>389</sup>In Part 7 of his Lévy, Paul (1950) An extension of the Lebesgue measure of linear sets. In: 4th Berkeley Symposium. pp. 273–287, Lévy showed how to construct an example of a non-measurable function  $f$  and a random variable  $X$  such that  $X$  and  $f(X)$  are independent.

<sup>390</sup>Lévy, Paul (1959) Le mouvement brownien fonction d'un point de la sphère de Riemann. *Rend. Circ. Mat. Palermo* 8:1–14.

et payés, et je n'ai pas voulu demander à ma femme de renoncer à un voyage dont elle se réjouissait.

J'hésite à dire que ce sera pour une autre fois. Je ne suis pas encore tout à fait remis de mon accident, et je ne sais pas quels voyages je pourrai envisager dans l'avenir.

Bien cordialement à vous

P.Lévy.

## Lettre 89

Beauséjour - par St Mathieu - Hte Vienne.

18 septembre 1961

Mon cher collègue,

Votre lettre m'arrive dans la Haute Vienne, où je suis jusqu'à la fin du mois. Je réponds d'abord à votre seconde question.

A part une conférence de 1921 que je considère un peu comme un péché de jeunesse,<sup>391</sup> j'ai toujours soutenu l'insuffisance des théories empiriques de la probabilité et combattu les idées de von Mises. Je n'ai aucun souvenir de lettre sur ce sujet qui puisse être dans les papiers de Borel. Si à l'occasion vous me montrez ce dont il s'agit, je verrai si je reconnais ma prose, ou du moins mes idées.

Pour l'autre question, elle m'embarrasse. Je me rappelle avoir fait incidemment une remarque sur les probabilités nulles, mais ne crois pas qu'elle ait jamais été l'objet principal d'un travail que j'aie publié.

En gros, voici mes idées sur cette question : 0 signifiant une probabilité nulle, et  $\infty$  un infini non dénombrable,  $\infty \times 0$  est une forme indéterminée, qui n'a aucune espèce de sens si les deux symboles se rapportent à des objets indépendants. Mais s'il y a une relation entre eux, il peut être question de lever l'indétermination, comme on le fait en algèbre pour la forme  $\infty \times 0$ . De même qu'on distingue des fonctions tendant vers zéro plus ou moins rapidement, on pourrait distinguer des probabilités plus ou moins nulles.<sup>392</sup> Du moins une telle distinction pourrait être utile ; mais je ne suis pas sûr qu'une théorie cohérente soit possible.

Quelques remarques simples peuvent préciser cette idée. Soit  $X(t)$  la fonction du mouvement brownien, défini dans  $(0, \infty)$ , avec la valeur initiale 0. On a, pour n'importe quel  $t_0$  donné  $\Pr\{X(t_0) = 0\} = 0$  ; mais la probabilité que  $X(t)$  soit constamment nul apparaît comme infiniment plus nulle.

---

<sup>391</sup>A session on the foundations of the theory of probability, organized by Lévy at a meeting of the Société de Philosophie, in late 1921. At the last moment Lévy replaced Borel, who was ill, at Borel's request. In the introduction to Volume III of his *Œuvres* (Lévy 1973), Lévy explained why he was opposed to the reprinting of this 'sin of his youth.'

<sup>392</sup>First mention in the correspondence of 'probabilities more or less zero.' See note 361. This letter is also interesting because it demonstrates why Lévy cared about rarefaction: the study of multiple points of the trajectory of Brownian motion. This theme is frequently revisited in the correspondence.

Soit  $\mathcal{E}$  l'ensemble des  $t$  correspondant aux maxima relatifs de  $X(t)$ , dans  $(0, \infty)$  et  $\mathcal{E}'$  celui des nombres rationnels  $> 0$ . La probabilité que  $\mathcal{E} \cap \mathcal{E}'$  ne soit pas vide est nulle. Celle que  $\mathcal{E}$  et  $\mathcal{E}'$  est<sup>393</sup> identique est infiniment plus nulle.

Je vous cite des exemples simples\*. Mais il y a des problèmes plus difficiles. Soit  $C_n$  la courbe du mouvement brownien à  $n$  dimensions, et  $M_t$  la position du mobile à l'instant  $t$ . La probabilité que  $M(t_0)$  soit un point double est nulle (si  $n > 1$ ) ; la probabilité qu'il y ait des points doubles est  $\infty \times 0$ . Elle ne peut que décroître avec  $n$  ; doit donc passer brusquement de 1 à 0. En fait, Dvoretzky, Erdős et Kakutani ont démontré qu'elle est 1 pour  $n = 3$  et 0 pour  $n = 4$ . Ce n'était pas du tout évident.<sup>394</sup>

Les mêmes auteurs ont démontré un résultat vraiment stupéfiant : la courbe  $C_2$  a des points multiples de tous les ordres, et même d'ordre infini non dénombrable (ce qui ne veut pas dire qu'il y en ait d'ordre infini dénombrable). C'est bien difficile à imaginer, mais je ne crois pas qu'ils se soient trompés. [*Following in left vertical margin*: A propos, je vous signale que j'ai démontré, il y a peut-être 10 ou 15 ans, qu'il peut arriver que le problème de Plateau ait une infinité non dénombrable de solutions.<sup>395</sup> Je ne sais pas s'il peut arriver qu'il y ait une infinité dénombrable, sans qu'il en existe une infinité non dénombrable.]

Au fond, je ne crois plus qu'une théorie générale des probabilités plus ou moins nulles puisse faire disparaître la difficulté de ces problèmes.

Bien cordialement

P.Lévy

\*La numérotation décimale nous en donne d'autres. La probabilité que la représentation que  $x \in (0, 1)$  ne contienne pas le chiffre 7 est nulle. Celle qu'il n'y ait ni 7 ni 8 apparaît comme plus nulle. L'une étant 0, l'autre serait  $0^2$ .

## Lettre 90

Paris -38 Av.Théophile Gautier - 9-1-62

Mon cher Collègue,

Recevant à la fois, ce matin, deux lettres de vous, je constate que l'une d'elles est datée du 12-10-61 avec un cachet de la poste du 3-1-62 !

Avant de vous répondre, je dirai un mot d'une autre question, celle de la discussion "Loi de Laplace ou loi de Gauss ?" Je vous ai entendu dire un jour que "malgré la précocité de Gauss, il est impossible qu'il ait devancé les travaux de Laplace." Or Gauss (d'après Larousse) est né en 1777. Le calcul des probabilités de Laplace, 1ère édition, est de 1812 = 1777 + 35. D'après les premières lignes de son introduction, cette introduction est le développement d'une conférence faite en 1795. Mais elle ne contient aucune formule, et je ne vois rien dans ce livre, que j'ai chez moi, qui permette d'affirmer que Laplace ait connu la loi en question avant le moment où Gauss

---

<sup>393</sup>*Sic.*

<sup>394</sup>Lévy is referring to the article Dvoretzky, A., Pál Erdős and Shinzo Kakutani (1950) Double points of paths of Brownian motion in  $n$ -space. *Acta Scient. Math.* Szeged 12(B):75–81.

<sup>395</sup>In 1947, Lévy gave examples of contours for which the Plateau problem has  $2p + 1$  solutions (CRAS 224:325–327). In 1948, in the Rumanian review *Matematica Timisoara* (number 23, pp. 1–45), Lévy gave an example with a countable infinity of solutions.

avait 33 ans (en comptant largement les délais d'impression). Dans ces conditions, maintenez-vous votre point de vue ?

Je me souviens que van Dantzig,<sup>396</sup> qui n'avait aucune raison d'être partial, m'avait dit en 1956, qu'il ne croyait pas votre point de vue justifié. Cela ne m'avait pas impressionné, parce que je n'avais pas mis en doute ce que vous aviez dit, ou du moins que je crois avoir entendu, sur les dates. Si la date du travail de Gauss est antérieure à 1812, ce que j'ignore, nous ne pouvons pas maintenir la priorité de Laplace.

---

Cela dit, je reviens à vos lettres.

Au sujet de la fonction  $x(a) = e^a + \sin e^a$ , je vous ferai simplement remarquer que  $\sin e^a$  est une fonction entière d'ordre infini dont les coefficients d'ordre élevé de son développement en série l'emportent infiniment sur ceux de  $e^a$  (sauf peut-être quelques coefficients par hasard plus petits ; mais je ne pense pas que ce soit le cas). Cela est vrai en particulier pour les coefficients négatifs (autrement on aurait  $x(a) \rightarrow \infty$  pour  $a \uparrow \infty$ ).<sup>397</sup>

Pour votre lettre sur les probabilités nulles, je ne peux pour le moment vous faire qu'une réponse très partielle.

D'abord, depuis que mon appartement a été vidé par les agents (français ou allemands) d'Hitler, je n'ai plus tous mes travaux antérieurs à 1936, et n'ai pas en particulier celui dont vous parlez. Je le rechercherai un jour à Poincaré, où d'ailleurs je vais rarement, depuis 1 an.

Pour la notion de probabilité conditionnelle, je me rallie au point de vue suivant qui me semble très net : si  $X$  est une variable aléatoire de f. de répartition  $F(x)$ , on n'a pas le droit de dire "j'appelle  $g(x)$  la probabilité de  $A$  dans l'hypothèse  $X = x$ ", et d'ajouter que le théorème des probabilités composées donne

$$\Pr(A) = \int_{-\infty}^{+\infty} g(x) dF(x),$$

puisque cette intégrale peut n'avoir aucun sens. Si donc je détermine d'abord  $X$ , et fais ensuite l'expérience qui peut réaliser  $A$  avec la probabilité  $g(x)$  si  $X = x$ , l'événement  $(X = x) \cup A$  peut n'être pas un événement aléatoire ayant une probabilité définie.

On évite toute difficulté en disant, dans le cas de deux variables aléatoires : je me donne la loi à deux variables  $X, Y$  définie par la fonction  $H(x - 0, y - 0) = \Pr(X < x, Y < y)$  ; soit  $F(x - 0) = \Pr(X < x)$ .

---

<sup>396</sup>Dutch mathematician David Van Danzig (1900–1959) was mainly active in algebraic topology. After World War II he became interested in statistics and probability.

<sup>397</sup>This somewhat arcane passage appears to be related to the discussion of Borel's text (see Letter 86).

J'appelle probabilité conditionnelle de  $Y < y$  n'importe quelle fonction  $G(y - 0, x)$  telle que

$$\int_{-\infty}^{+\infty} G(y - 0, u) dF(u) = \Pr(X < x, Y < y) = H(x - 0, y - 0).$$

Alors  $G$  est bien définie, sauf peut-être sur un ensemble de valeurs de  $x$  sur lequel la variation de  $F(x)$  est nulle.

Cela est très net, assez voisin des idées de Doob, et réfute l'objection de Finetti. Etes vous d'avis que cela mérite une petite note que j'enverrais à M.Montel pour son Bulletin ?<sup>398</sup>

D'autre part, pour vous mettre en garde contre le risque de trop exiger d'une classification des ensembles de mesure nulle, j'attire votre attention sur le fait suivant : pour chaque  $x \in (0, 1)$ , je considère une variable aléatoire  $U(x)$  uniformément répartie dans  $(0, 1)$ , et les choix des  $U(x)$  sont absolument indépendants. Alors la probabilité que  $U(x)$  prenne au moins une fois une valeur donnée  $x \in (0, 1)$  n'a aucune espèce de signification.

Plus généralement, si on considère un événement  $A$  de probabilité nulle, par exemple  $X \in E$  dans le cas où  $X$  est uniformément répartie dans  $(0, 1)$ , et où  $E$  est  $\subset (0, 1)$  et de mesure nulle, et une infinité non dénombrable d'épreuves susceptibles de réaliser  $A$ , et indépendantes, on ne peut pas considérer le fait que  $A$  soit réalisé au moins une fois comme un événement ayant une probabilité bien définie.

Dans un ordre d'idées voisin\*, je me demande si vous connaissez des théorèmes très curieux de A.Dvoretzky, P.Erdős et S.Kakutani.<sup>399</sup> Si  $n > 3$ , la courbe du mouvement brownien dans l'espace à  $n$  dimensions n'a aucun point double ; elle en a presque sûrement, formant un ensemble partout dense, si  $n \leq 3$ . Si  $n > 2$ , elle n'a aucun point triple ; si  $n = 2$ , elle en a une infinité, et elle a même une infinité de points multiples de n'importe quel ordre entier. Ce qui est bien plus curieux, et que je n'arrive pas à bien m'imaginer, c'est qu'elle a même des points doubles d'ordre infini non dénombrable. Cela signifie que, si  $X(t)$  et  $Y(t)$  sont deux fonctions de Wiener indépendantes l'une de l'autre, on peut presque sûrement (et naturellement d'une infinité de manières), trouver  $x$  et  $y$  tels que  $X(t) = x$  et  $Y(t) = y$ , soient réalisés simultanément sur un ensemble non dénombrable de valeurs de  $t$ . (Ces points  $x, y$  forment un ensemble partout dense sur la courbe, donc dans le plan.) Si je me rappelle bien, ce dernier point (sauf les remarques entre () que je fais en vous écrivant), a été obtenu par une collaboration des mêmes auteurs avec S. J. Taylor.<sup>400</sup> Je n'ai pas approfondi la démonstration ; mais elle m'a paru exacte. Une erreur qui aurait échappé à ces 3 auteurs est d'ailleurs une hypothèse bien invraisemblable.

Bien cordialement

<sup>398</sup>At the time Montel was editor of the Bulletin des Sciences Mathématiques. However, the note was apparently never submitted. Perhaps Fréchet dissuaded Lévy from publishing it.

<sup>399</sup>These theorems were mentioned in Letter 89.

<sup>400</sup>Dvoretzky, A., Paul Erdős, Shinzo Kakutani and S.J. Taylor (1957) Triple points of Brownian paths in 3-space. Proc. Camb. Philos. Soc. 53:856–862.



P.Lévy

\* voisin pour la raison suivante : si  $t$  et  $t'$  sont donnés  $X(t) = X(t')$  est infiniment peu probable ; la répétition des expériences rend cet événement presque sûr. La conclusion change si vous considérez la réalisation de 4 événements simultanés  $X_i(t) = X_i(t')$  ( $t' - t \neq 0$  ;  $i = 1, 2, 3, 4$ ). Il faut donc y regarder de très près. Naturellement, cela n'exclut pas absolument la possibilité de progresser dans la voie signalée par Borel, mais je suis plutôt sceptique.

## Lettre 91

Paris - 38 Av. Théophile Gautier- 24-1-62

Mon cher collègue,

Vous avez bien fait de me relancer, car je crois être arrivé, non à la solution de votre problème lui-même, mais à celle du problème obtenu en modifiant une de vos définitions. Le but, qui consiste à établir une hiérarchie dans les ensembles linéaires de mesures nulles, n'en est pas moins atteint.

Je considère, comme vous, une suite d'intervalles  $i_n$ , de longueur  $l_n$ , majorant un ensemble  $E$ , c'est-à-dire le recouvrant une infinité de fois. Mais, 1° au lieu de considérer le reste de la série  $\sum l_n$ , supposée convergente, je considère la convergence vers 0 des  $l_n$  eux-mêmes. Je les suppose non croissants.

2° si  $l'_n/l_n$  est toujours compris entre deux nombres positifs finis, je dis que les deux suites convergent vers zéro avec la même vitesse. Si une de ces bornes existe, et non l'autre, une des suites converge plus vite que l'autre. Cela n'implique pas que  $\frac{l'_n}{l_n}$  tende vers 0 ou  $\infty$ .

Cela dit, je définis dans  $[0, 1]$  un ensemble  $E_\alpha$  ( $0 < \alpha < 1$ ) qui, pour  $\alpha = \frac{1}{3}$ , se réduit à l'ensemble ternaire de Cantor. Dans le cas général, au lieu de diviser  $[0, 1]$  en 3 parties égales, je le divise en 3 parties  $[0, \alpha]$ ,  $(\alpha, 1 - \alpha)$ ,  $[1 - \alpha, 1]$ , puis, l'intervalle du milieu étant enlevé, je divise de même les deux intervalles conservés. Après  $p$  opérations, j'ai  $2^p$  intervalles, de longueur  $\alpha^p$ , dont la réunion est un ensemble  $E_\alpha^p$ . L'intersection de tous les  $E_\alpha^p$  ( $p = 1, 2, 3, \dots$ ) est  $E_\alpha$ . Si  $\alpha = \frac{1}{r}$  ( $r$  entier  $> 2$ ),  $E_\alpha$  est l'ensemble des nombres qui peuvent s'écrire dans le système de numération de base  $r$  en n'employant que les chiffres 0 et  $r - 1$ .

Cet ensemble  $E_\alpha$  est d'autant plus raréfié que  $\alpha$  est plus petit. Il est toujours moins raréfié que n'importe quel ensemble dénombrable.

Je n'ai pas voulu tarder à vous prévenir. Il me reste à revoir dans tous ses détails le raisonnement que je viens d'esquisser dans mon esprit. Je crois que le tout tiendra dans les limites d'une note que je pense envoyer avant 15 jours, si tout va bien.<sup>401</sup>

Il faut aussi que je prépare des conférences que je dois faire à Rome avant Pâques. Quoique mon activité soit assez réduite, je pense pouvoir faire les deux choses en temps utile.

Bien cordialement

P.Lévy.

---

<sup>401</sup>Lévy eventually renounced to write this note (see next letter).

**Lettre 92**

Paris - 25-1-62

Mon cher collègue,

Je vous ai écrit bien trop vite. Je ne crois plus qu'on puisse se passer de la considération des restes ; je ne crois pas non plus qu'ils suffisent.

Je pense à une condition telle que celle-ci : si  $u_n, u'_n > 0$ , si  $\sum u_n$  et  $\sum u'_n$  convergent et ont des restes  $\rho_n$  et  $\rho'_n$ , on dira que  $\sum u_n$  converge plus vite que  $\sum u'_n$  si :

- 1° il existe un  $k > 0$  tel que  $u'_n > ku_n$ , d'où  $\rho'_n > k\rho_n$
- 2° Aucun des deux rapports  $\frac{u'_n}{u_n}$  et  $\frac{\rho'_n}{\rho_n}$  n'est borné supérieurement (ce sont deux conditions distinctes ; peut-être suffit-il de considérer la seconde).

En tout cas, je crois qu'il faut faire intervenir la condition  $u'_n > ku_n$ .

J'ai l'impression très nette qu'avec une définition convenable de la raréfaction, qu'il s'agit de trouver, mes ensembles  $E_\alpha$  ( $0 < \alpha < \frac{1}{2}$ ) sont d'autant plus raréfiés que  $\alpha$  est plus petit.

Malheureusement, autant je pensais trouver le temps de rédiger une note, autant je crains de ne pas pouvoir étudier à fond avant mon voyage à Rome une question plus complexe que je ne l'avais cru. J'espère que dans 3 mois je n'aurai pas oublié les quelques idées que j'ai eues.

Bien cordialement,

P.Lévy

**Lettre 93**

Paris - 38 Av. Théophile Gautier

février 62<sup>402</sup>

Mon cher collègue,

La préparation des conférences que je dois faire à Rome avant Pâques m'a empêché de penser ces derniers jours à la raréfaction des ensembles de mesure nulle. Mais en y repensant, je ne vois plus ce qui m'avait arrêté. Il me semble que l'idée que j'avais eue est valable... mais elle ne résout pas votre problème.

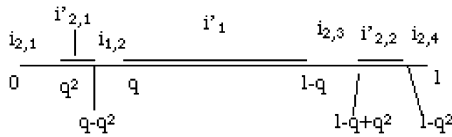
Il doit être entendu que, si une suite d'intervalles  $i_n$ , de longueurs  $l_n$ , majore un ensemble  $E$ , il faut les supposer rangés de manière que  $l_{n+1} < l_n$ . Sans cela on n'aurait qu'à changer l'ordre des termes pour empêcher de dire que la suite considérée converge plus rapidement qu'une autre suite donnée. Alors on peut appliquer le théorème de Borel :  $nl_n$  tend vers 0.

J'appelle  $E_q$  ( $0 < q < \frac{1}{2}$ ), l'ensemble ternaire généralisé, défini comme suit. Je pars de  $(0, 1)$ . J'enlève l'intervalle  $i'_1$ , soit  $(q, 1 - q)$ , et conserve  $(0, q)$  et  $(1 - q, 1)$

---

<sup>402</sup>Despite the incomplete date, this letter was clearly begun on February 4 (see end of letter).

que j'appelle  $i_{1,1}$  et  $i_{1,2}$ .



J'enlève de nouveau des intervalles  $i'_{2,1}$  et  $i'_{2,2}$  (fig.) et conserve après cette deuxième opération 4 intervalles de longueurs  $q^2$ , soit  $i_{2,1}, \dots, i_{2,4}$ . Après  $n$  opérations, j'aurai  $2^n$  intervalles  $i_{n,h}$ , ( $h = 1, 2, \dots, 2^n$ ) de longueur  $q^n$ . A la limite j'ai  $E_q$ , de mesure nulle.

Or  $E_q$  est la réunion de  $2^n$  ensembles partiels  $e_{n,k} = i_{n,k} \cap E_q$ , disjoints et superposables. On peut alors définir sur  $E$  une "mesure relative"  $\mu(e)$  bien définie par la condition  $\mu(e_{n,k}) = 2^{-n}$ , et, pour qu'une suite d'intervalles  $I_p$  majore  $E_q$ , il faut que  $\sum \mu(I_p) = \infty$ , avec  $\sum l_p < \infty$ .

Or, au moins si  $q \leq \frac{1}{3}$ , il est évident que les  $I$  de longueur  $q^n$  qui rendent  $\mu(I)$  maximum sont les  $i_{n,k}$ . D'ailleurs, si un  $I_l$  a une longueur comprise entre  $q^n$  et  $q^{n+1}$ , on peut l'agrandir et lui donner la longueur  $q^n$  ; il couvrira au moins la même partie de  $E_q$ , et, avec vos définitions, ce sera sans effet sur la rapidité de la convergence de  $\sum l_p$ . Il en résulte qu'on a les meilleures couvertures possibles en ne considérant que des suites d'intervalles  $i_{n_p, k_p}$ . Les  $n_p$  doivent vérifier les conditions  $\sum 2^{-n_p} = \infty$ ,  $\sum q^{n_p} < \infty$ , et il me semble facile de voir que ces conditions sont suffisantes pour qu'on puisse déterminer les  $k_p$  de manière que la suite des  $i_{n_p, k_p}$  majore  $E_q$ .

Une grosse étape est ainsi franchie, la condition  $\sum q^{n_p} < \infty$  étant d'autant moins restrictive que  $q$  est plus petit.

Soit alors  $q < q' \leq \frac{1}{3}$ . Nous pouvons déterminer les  $n_p$  vérifiant les conditions  $\sum q^{n_p} < \infty$  et  $q^{m_p} \geq \frac{k}{\rho}$  ( $k > 0$ ). Alors, d'après le lemme de Borel, la série  $\sum q^{m_p}$  est divergente.

5 février- Je me suis arrêté hier soir, en retrouvant la même difficulté qu'il y a quelques semaines. En rédigeant, je me suis aperçu que mon idée intuitive était fautive.<sup>403</sup> Le lemme de Borel donne une égalité qui n'a pas le sens voulu pour donner  $\text{Rar}.E_q > \text{Rar}.E_{q'}$ , avec vos définitions.

Je vous envoie tout de même cette lettre, parce que je crois que ma première étape donne un résultat d'un certain intérêt.

J'ai de nouveau l'impression qu'il faudrait trouver une autre définition de la raréfaction, telle que mon résultat entraîne  $\text{Rar}.E_q > \text{Rar}.E_{q'}$ . Le résultat n'a de sens que pour les ensembles  $E$  ayant la propriété suivante : il existe une suite de nombres croissants  $n_l$  tels que  $E$  soit la réunion de  $n_l$  ensembles partiels, disjoints et superposables. Il y en a beaucoup d'autres que les  $E_q$  ; mais cela ne donne pas un résultat applicable à tous les ensembles de mesure nulle.

<sup>403</sup>See Letters 90 and 91.

Conclusion pratique. Il faut que je me remette à mes conférences de Rome, qu'on me demande de rédiger,<sup>404</sup> et je dois laisser la raréfaction de côté, jusqu'à Pâques\*. Si vous voyez le moyen d'utiliser mon idée, dont l'essentiel est l'introduction d'une mesure relative sur certains ensembles de mesure nulle, je serais heureux que vous le fassiez.

Bien cordialement à vous

P.Lévy

\* Je réfléchis que j'ai déjà introduit en 1938-39 une idée analogue pour certains ensembles aléatoires dénombrables partout denses, ceux des points de discontinuité des processus stochastiques homogènes liés aux lois stables. Voyez Processus stochastiques, n° 47 à 50 ; j'y définis un nombre  $s$  qui a presque sûrement le caractère d'une mesure de cet ensemble. Mais je ne pense pas que la lecture, assez ardue, de ces 16 pages, vous aiderait pour votre problème.

#### Lettre 94

Paris - 38 Av. Théophile Gautier

8 octobre 1962

Mon cher collègue,

Pouvez-vous me dire s'il y a des travaux de faits sur les fonctions analytiques dans l'espace de Hilbert. J'en ai besoin pour mes recherches sur le mouvement brownien,<sup>405</sup> et je me demande s'il faut reprendre la question à ses débuts, ou considérer les premiers éléments comme connus.

Il s'agirait de démontrer le théorème suivant, que je crois exact : si la fonction brownienne d'un point  $A$  de l'espace de Hilbert est connue sur un ouvert d'une surface analytique, elle est déterminée sur toute cette surface. Ces "surfaces" sont des lieux de points dépendant d'une infinité de paramètres ; il ne peut rien y avoir d'analogie si le nombre des paramètres est fini.

Bien cordialement à vous

P.Lévy.

#### Lettre 95

Paris - 38 Av. Théophile Gautier

10 Octobre 1962

Mon cher collègue,

---

<sup>404</sup>The proceedings of these conferences held at the Istituto Guido Castelnuovo of the University 'La Sapienza' were published in Lévy, Paul (1962) Le mouvement brownien fonction d'un ou plusieurs paramètres. Rend. Mat. 22:24-101.

<sup>405</sup>Lévy presented a note to the Academy on June 4, 1962, on Brownian motion with a parameter in a Hilbert space. This theme was developed in his Lévy, Paul (1962) Le déterminisme de la fonction brownienne dans l'espace de Hilbert. Ann. Sci. Ec. Norm. Supér. 79:377-398. Lévy extended these results with a second paper: (1962) Ann. Sci. Ec. Norm. Supér. 80:193-212, as well as a note (1963) CRAS 256:1444-1446, in which he proved the theorem mentioned in a particular case.

Je pense que vous irez à l'Institut lundi (15 octobre). Si oui, voulez vous y apporter le livre de Michal<sup>406</sup> que vous voulez bien me prêter ; je serai heureux de vous voir au moins quelques instants.

Sinon, voulez-vous me prévenir en téléphonant à BAG(atelle) 77-35, ou par lettre (j'ai demandé qu'on change mon appareil qui est défectueux ; en attendant, il vaut mieux appeler chez ma fille, qui occupe une partie de mon appartement).

J'avais d'ailleurs essayé tout à l'heure de vous appeler ; mais vous étiez sorti.

Bien cordialement à vous P.Lévy.

P.S. Est-ce que je me trompe en attribuant à Banach et Kuratowski<sup>407</sup> le théorème d'après lequel la mesure de Lebesgue (sur une droite, par exemple) ne peut pas être prolongée par une mesure complètement additive ?

Je pense que vous le savez, dans le cas contraire ne vous donnez pas la peine de le rechercher. Je vois quelquefois Loève qui pourra sans doute me renseigner.

## Lettre 96

[In left margin, in Fréchet's handwriting: rép. le 2.11.62]

Paris - 38 Av. Théophile Gautier

23-10-62

Mon cher collègue,

Je vous remercie pour le livre de Michal.<sup>408</sup> Il est bien fait, et intéressant.

Je vous signale seulement

- 1° une erreur dans la définition des fonctions analytiques. Il faut préciser qu'il s'agit de polynômes de degrés croissants. Sans cette condition, pour  $x$  réel, la série de polynômes homogènes  $\sum P_n(x)$ , supposée uniformément convergente, peut représenter n'importe quelle fonction continue dans un intervalle fini.
- 2° la première partie du théorème III,3, qui vous est attribuée avec la date de 1937, se trouve déjà dans mes Leçons d'analyse fonctionnelle de 1922, p. 52, n° 38.<sup>409</sup>

Je répète ces remarques sur une feuille séparée, ci-jointe. Peut-être jugerez-vous utile de la transmettre à M<sup>elle</sup> Mourier, puisque c'est évidemment elle qui aura éventuellement à s'occuper d'une nouvelle édition, ou même sans attendre cette éventualité, à faire établir une liste d'errata.

---

<sup>406</sup>See Letter 96.

<sup>407</sup>In any case, it was the Polish school of topology that took on the task of axiomatizing the Lebesgue measure. The standard reference is Sierpinski, Waclav (1918) Sur une définition axiomatique des ensembles mesurables (L). Bull. Intl. Acad. Sci. Crac. 173-178. See also Shafer and Volk (2006).

<sup>408</sup>Michal, Aristotle D. (1958) Le calcul différentiel dans les Espaces de Banach. Gauthier-Villars, Paris, by the American mathematician Aristotle D. Michal was published in 1958 in the series of monographs begun by Borel. It was translated into French by Edith Mourier, who was at Poitiers at the time, with a preface by Fréchet.

<sup>409</sup>This theorem concerns the differentiability of the supremum norm on the space of real continuous functions on  $[a, b]$ .

J'ajoute que l'opinion exprimée par Michal, p.18, sur votre mémoire de 1906, est à ma connaissance partagée par beaucoup de savants.<sup>410</sup> Pour ma part, en faisant l'analyse du livre de Le Lionnais<sup>411</sup> où figurent deux articles de vous, j'ai écrit, à propos du second : je dois ajouter ce qu'il a négligé de dire :<sup>412</sup> l'analyse générale, qui est peut être aujourd'hui la branche la plus importante des mathématiques (qui l'est sûrement si on évalue son importance par le nombre des travaux qui s'y rapportent) est née de son mémoire de 1906.

Je suis un peu plus prudent que Michal, parce que je suis incapable de savoir si ce n'était pas un peu une idée "dans l'air", et si Moore,<sup>413</sup> par exemple, n'aurait pas écrit sa general analysis si vous n'aviez pas écrit votre mémoire. Mais je n'ai jamais entendu contester qu'en fait vous avez été le premier.

Bien cordialement à vous

P.Lévy.

J'ai un nouveau téléphone, qui rend les conversations plus faciles.

[*Feuille jointe*] :

Errata ou observations pour le livre de Michal : *Le Calcul différentiel pour les espaces de Banach*

p.38, l.8. Après polynômes homogènes mettre de degrés croissants.

Sans cette restriction, on pourrait aisément obtenir des fonctions non analytiques (dans le cas  $x$  réel).

p.56 Théorème III.3

La seconde partie m'étonne ; je ne comprends pas très bien.<sup>414</sup>

Mais pour la première, c'est un résultat que j'ai déjà indiqué en 1922 (Leçons d'analyse fonctionnelle, p.52, n° 38)

Je signale enfin que, p.19 et p.146, le nom de Lichnerowicz a été mal écrit.<sup>415</sup>

le 22 octobre 1963<sup>416</sup> P.Lévy

## Lettre 97

Paris-38 Av. Théophile Gautier

22 février 1963

---

<sup>410</sup>Michal wrote: 'Ce travail [...] a influencé l'orientation de la pensée mathématique plus que tout autre travail mathématique du XXème siècle (This work had more influence on mathematical thought than any other mathematical work of the 20th century.)' One feels that Lévy found this opinion slightly exaggerated.

<sup>411</sup>François Le Lionnais' extraordinary collection of articles, *Les grands courants de la pensée mathématique*, published by Hermann in 1948 (with a new edition in 1998). Fréchet contributed two articles, on numbers and space.

<sup>412</sup>*Sic*. The syntax of this sentence is weak.

<sup>413</sup>The American mathematician Eliakim Hastings Moore (1862–1932) was one of the great pioneers of spaces of functions. Birkhoff was his student.

<sup>414</sup>The first part considers the differentiability, in Gateaux's sense, of the functional  $x \mapsto \max_{s \in [0, 1]} x(s)$ , and the second stipulates that this functional is not Fréchet-differentiable.

<sup>415</sup>The spelling in the book is *Lichnerowicz*.

<sup>416</sup>*Sic*, instead of 1962.

Mon cher collègue,

Je vous adresse ci-inclus une note de mon fils, Jean-Claude Lévy, ingénieur en chef de l'artillerie navale.

Il n'indique que des solutions partielles et approchées des problèmes posés. Mais je crois que ces problèmes sont réellement difficiles, et que ces résultats partiels méritent d'être signalés.<sup>417</sup>

J'en ai parlé à Claude Berge, spécialiste de l'analyse combinatoire, et à Choquet. Aucun d'eux n'a connaissance de travaux déjà faits sur ce sujet.

Au cas où vous voudriez avant de présenter cette note lui faire quelques observations, mon fils habite Versailles, 103 Bould de la Reine, Téléph. 950-37-49. Il déjeune en général à Paris chez sa belle-mère : TRO-68-68. Enfin son bureau est à Issy : MIC-28-45.

Bien cordialement

P.Lévy.

Je vous ai envoyé hier un exposé sur les ensembles laplaciens de variables aléatoires ; je pense qu'il résout la question de savoir ce qu'il convient d'appeler une fonction laplacienne. Vous avez peut-être vu ma dernière note sur la fonction brownienne définie dans l'espace de Hilbert.<sup>418</sup> Je l'ai présentée un peu trop vite. Je suis d'ailleurs convaincu de l'exactitude du résultat énoncé (théorème 3) ; mais la démonstration que j'ai esquissée ne s'applique que si le plan  $\Omega_2$  n'a qu'un nombre fini de dimensions ; c'est-à-dire s'il ne manque qu'un nombre fini de dimensions à la surface  $S$  pour remplir un volume de l'espace de Hilbert.

Je pars le 21 pour les Etats Unis où je passerai deux mois ; New York et Washington.

P.L.

## Lettre 98

[*Up left margin: 15.1* Je retrouve cette lettre et vous l'envoie bien que l'autre soit partie, à cause de la lettre de Fortet.]

Paris - 38 Av. Théophile Gautier

le 12 janvier 1964

Mon cher collègue,

Je vous remercie pour vos deux mémoires, et vous renvoie une lettre de Fortet qui se trouvait dans la même enveloppe.

J'admire votre activité. L'âge a déjà singulièrement diminué la mienne, et je me demande ce que je serai capable de faire quand j'aurai franchi l'étape des 80 ans, à supposer que j'y arrive.

---

<sup>417</sup>In fact, the publication of this work took considerable time. Jean-Claude Lévy published *Théorie statistique des combinaisons discernables* (1966) CRAS 262:479–481, on a study of neural networks. He returned to this subject in a supplementary note also published in 1966, and in a later article Lévy, Jean-Claude (1969) *Sur la possibilité de simuler par le calcul électronique les réseaux de cellules nerveuses*. *Elektro. Infverarb. Kybern.* 5:13–26.

<sup>418</sup>Lévy, Paul (1963) *Ensembles laplaciens de variables aléatoires et fonction brownienne définie dans l'espace de Hilbert*. CRAS 256:1444–1446. A few weeks later Lévy published corrections in CRAS 256:5495.

Au sujet de la raréfaction, son intérêt pourrait à mon sens encore être accru par des applications, et je songe à la suivante.<sup>419</sup> Soit  $A(t)$  le point qui décrit la courbe du mouvement brownien. Soit  $E_n$  l'ensemble des  $t$  pour lesquels  $A(t)$  est un point multiple d'ordre  $n$  de cette courbe. Si on sait que  $A(t_1) = \dots = A(t_n) = A(t)$ ,  $t$  appartient à  $E_n$ , mais n'a aucune chance d'appartenir à  $E_{n+1}$ . Les probabilités  $\Pr\{t \in E_n\}$  sont donc "de plus en plus nulles". Pourtant, les  $E_n$  ne sont jamais vides, leur intersection  $E_\omega$  ne l'est pas, et dans  $E_\omega$ , on peut encore (d'après A.Dvoretzky, P.Erdős et S. Kakutani) trouver un ensemble partout dense de points multiples dont l'ordre a la puissance du continu. Tout cela peut-il être précisé par la théorie de la raréfaction ? Peut-on par exemple dire que  $E_{n+1}$  est presque sûrement plus raréfié que  $E_n$  ?\*

C'est naturellement un problème très difficile ; je pense que les trois auteurs cités pourraient s'y intéresser. Si vous êtes de cet avis, je pourrai écrire à l'un d'eux. Pour moi, je n'ose pas me lancer dans une question si difficile. Voilà 3 mois que je réfléchis à un "théorème évident" sur une classe de fonctions entières aléatoires et que je n'arrive pas à le démontrer !

Bien cordialement

P.Lévy.

Mes visites me donnent en général satisfaction. Même au Collège de France, où je pensais que M<sup>420</sup> aurait des appuis, je suis très bien accueilli. Au moins 3 fois sur 4, on me dit "Je sais quels sont vos titres et je voterai pour vous" ou bien "je suivrai la section", ce qui revient au même.

Il me reste environ 30 visites à faire. Je n'ai commencé ni le museum, ni l'Institut Pasteur. Je dois voir Trefouel<sup>421</sup> demain et Heim<sup>422</sup> mercredi.

Je posterai cette lettre demain si je ne vous vois pas à l'Académie.

\* Une question analogue se pose pour l'ensemble des points  $A(t_n)$  d'ordre  $n$  si on définit la raréfaction dans le plan. Ces ensembles sont aussi partout denses dans le plan.

## Lettre 99

Paris - 38 Av. Théophile Gautier

14 janvier 1964

Mon cher Collègue,

Je vous ai écrit dimanche<sup>423</sup> pour vous remercier de vos tirés à part, et pour d'autres raisons aussi. Je pensais mettre cette lettre à la poste hier soir si je ne vous rencontrais pas à l'Institut, et je ne l'ai plus retrouvée.

<sup>419</sup>An application mentioned in the letters since 1961. See especially Letters 89 and 90.

<sup>420</sup>Mandelbrojt.

<sup>421</sup>Jacques Tréfouël (1897–1977), in the 1940s director of the Institut Pasteur, was in 1964 Vice President of the Academy of Sciences.

<sup>422</sup>The renowned mycologist Roger Heim (1900–1979) was then director of the Museum d'Histoire Naturelle de Paris.

<sup>423</sup>The preceding Letter 98. Apparently, after just two days, it had not reached Fréchet, so Lévy repeated his explanation of the application of rarefaction to multiple points on Brownian trajectories. His insistence shows the importance he attached to this application.



Aurait-elle été mise à la poste par erreur, non timbrée ? Dans ce cas je n'ai qu'à m'excuser de cette distraction.

L'hypothèse qu'elle soit perdue est plus probable. Alors une lettre de Fortet, glissée par erreur dans vos brochures, et que je vous renvoyais aussi, serait perdue aussi. J'espère que cela n'aura pas de conséquences.

Je vous disais d'autre part que, pour moi, l'intérêt de la notion de raréfaction pourrait être accru par quelques applications, et je me demande si elle pourrait être appliquée à l'étude des points multiples de la courbe du mouvement brownien. J'appelle  $E_n$  l'ensemble des  $t$  pour lesquels le point  $A(t)$  qui décrit la courbe est multiple d'ordre  $n$  ;  $E_\omega$  l'intersection de tous les  $E_n$  ;  $E_c$  le sous-ensemble de  $E^\omega$  pour lequel l'ordre de multiplicité est infini non dénombrable. Je ne sais pas si j'ai déjà attiré votre attention sur un étonnant théorème de A.Dvoretzky, P.Erdős et S. Kakutani<sup>424</sup> d'après lequel  $E_c$  n'est pas vide \*[Dans la marge et à la verticale : \* On déduit aisément qu'il est presque sûrement partout dense sur l'axe des  $t$ , et que l'ensemble des points multiples qui lui correspondent est partout dense dans le plan.] (énoncé presque sûr). On a évidemment  $E_2 \supset E_3 \supset \dots \supset E_n \supset \dots \supset E_\omega \supset E_c$ , et si on sait, pour  $n$  valeurs distinctes  $t_\nu$  de  $t$  que  $A(t_1) = A(t_2) = \dots = A(t_\nu)$ , on sait que ces  $t_\nu$  appartiennent à  $E_n$ , et que la probabilité qu'ils appartiennent à  $E_{n+1}$  est nulle. Cela suggère l'idée que les  $E_n$  doivent être de plus en plus raréfiés, et que peut-être  $E_c$  est plus raréfié que  $E_\omega$ .

Je ne m'attaquerai pas à ce problème qui me semble trop ardu. Mais je pense à attirer sur lui l'attention de Dvoretzky.

Mes visites progressent favorablement. Même au Collège de France, je trouve un terrain bien préparé par Leray, Lichnerowicz, et par vous (au moins en ce qui concerne Laval). Je ne pense pas que Mandelbrojt puisse avoir beaucoup de voix ; ses visites lui serviront peut-être pour une élection ultérieure (il a jusqu'ici été classé ex-æquo avec Henri Cartan).

Bien cordialement à vous

P.Lévy

## Lettre 100

Paris - 38 Av. Théophile Gautier

22 avril 1964<sup>425</sup>

Mon cher collègue,

J'ai bien reçu votre mot et vous remercie. Il est curieux que j'aie eu 35 voix sur 70, en 1949, sans être élu ; et cette fois je suis élu avec 35 voix sur 62, et si je me rappelle bien, Garnier a été élu avec 35 voix sur 64.

<sup>424</sup>The paper by Dvoretzky, A and Shinzo Kakutani (1958) Points of multiplicity  $c$  of plane Brownian motion. Bull. Res. Council. Isr. 7F:175-180.

<sup>425</sup>Lévy was elected to the Geometry Section of the Academy of Sciences on April 20, 1964, replacing Hadamard. Prior to that, the most recently elected member of this Section had been Fréchet, who acceded to Borel's position in 1956. See the lively discussions in the letters from 1950 and 1952.

Je vous avoue que l'accueil que j'avais reçu au cours de mes visites m'avait fait penser que j'aurais 45 ou 50 voix au premier tour. J'en suis loin ; mais comme vous me le dites, tout cela n'a plus grande importance.

J'ai reçu une lettre des Secrétaires Perpétuels me disant que j'avais à rédiger une notice sur M.Hadamard. Je ne m'en doutais pas, croyant que celle que vous avez lue suffirait. Pour cette notice, j'aurais à vous demander quelques conseils et surtout de me communiquer la documentation que vous avez du réunir. Mais cela ne presse pas ; je veux d'abord terminer le travail nécessaire pour la réédition de mes Processus stochastiques.

Bien cordialement

P.Lévy.

### Lettre 101

Paris - 38 Av. Théophile Gautier

28 avril 1964

Mon cher collègue,

Je n'ai jamais bien compris la première définition des collectifs donnée par Ville ;<sup>426</sup> Loève et Khintchine m'avaient dit et écrit n'avoir pas compris. C'est en 1950, à Berkeley, que j'ai appris par Loève que les processus qu'on appelait martingales sont ceux que j'avais considérés dès 1935 ; d'après votre lettre, sa deuxième définition, p.99, coïncide bien avec la mienne\*<sup>427</sup>

Naturellement, je n'ai pas employé ce mot que je n'avais pas connu en 1937 dans la réédition en 1954 de mon livre de 1937 ; pour permettre la reproduction photographique, je n'ai guère fait que corriger quelques erreurs et ajouter deux notes.

Mais la condition (C), introduite p.238, revient à dire que la suite des  $X_\nu$  est une martingale. Cette condition intervient dans la suite : théorèmes 67,1; 67,2; 67,3; 68; n° 69, 1° et 2°. <sup>428</sup> J'ai ainsi esquissé une théorie, développée depuis par Doob, et qui généralise les suites de variables aléatoires indépendantes à valeurs probables nulles.

Quant à la théorie des collectifs, malgré les mérites que je reconnais à Von Mises, je l'ai toujours trouvée absurde, ce que je n'ai pas caché à Wald quand il l'a exposée

<sup>426</sup>Another mention of Ville's thesis and Lévy's low opinion of it (see Letter 40, note 157 and the rest of this letter). Ville introduced the term "martingale" in 1939. Doob developed the corresponding theory as of 1940. Dellacherie and Meyer *Probabilités et Potentiel (Probabilities and Potential)*, Chaps. V-VIII, Hermann, 1980 (p. 443) credit Ville with making the idea of stopping time quasi-explicit, however Lévy had already made use of the fundamental properties of this notion (see Lévy 1937). See also the various papers in the issue 5.1 of the Electronic Journal for History of Probability and Statistics (2009).

<sup>427</sup>Carried away by his low regard for Ville's work, Lévy goes a bit too far here. His condition (C), on p. 238 of Lévy (1937), amounts to saying that the sequence of partial sums of a random series is a martingale. Ville in his 1939 thesis *Critique de la Notion de Collectif*, p. 99, proposed the modern definition of the idea in full generality. See details in Mazliak (2009).

<sup>428</sup>In the results already cited, Lévy made use of the martingale properties (without calling them that) of partial sums of a series of random variables to obtain classical inequalities and convergence results.

à Genève. Je sais gré à Ville de m'avoir aidé à la combattre. Mais cela ne suffit pas pour le classer au même niveau que... disons Fortet et Dugué, pour ne parler que des probabilistes de la Sorbonne.

De même le fait qu'André Metz ait fort justement critiqué les erreurs de Bergson sur la relativité<sup>429</sup> ne m'empêche pas d'être un peu surpris de le voir candidat à l'Académie. Nous sommes beaucoup à souhaiter la candidature d'un mathématicien aux Académiciens libres ; mais aucun de ceux qui le méritent ne veut poser sa candidature ; et alors il faut s'attendre à ne plus voir d'ici 30 ans un moins de 70 ans à la Section de géométrie !

Bien cordialement

P.Lévy.

Au moment de signer cette lettre, j'ai été interrompu par un long coup de téléphone de Bass, et l'idée m'est venue qu'il serait mieux à sa place que Metz aux Académiciens libres ; mais je n'ai pas prononcé de paroles irréfléchies.<sup>430</sup>

\* ou du moins s'y ramène par l'addition de constantes

## Lettre 102

Marlotte - 16 juin 1965

Mon cher ami,

Je me félicite de m'être décidé à emporter à la campagne votre brochure sur Borel ; je l'ai lue hier sans pouvoir m'en détacher, ce qui ne veut pas dire que je n'aurai pas à reprendre certains passages pour y réfléchir plus longuement. Mais je voudrais vous faire part de quelques impressions.

Inutile de vous dire que, sauf quelques points que je veux vous signaler, je suis d'accord avec vous. Vous avez admirablement marqué les rôles respectifs de Borel et de Lebesgue dans la mise au point des théories de la mesure et de l'intégration. Vous avez su mentionner l'étrangeté de ses conceptions sur l'axiome du choix, avec une modération dont je n'aurais pas été capable. Et vous avez sans doute bien fait de passer sous silence quelques erreurs - au moins trois que je ne m'explique pas chez un savant de cette classe. Je ne parle que des erreurs dans lesquelles il s'est entêté ; car je ne m'étonne jamais de celles qui ne sont que des distractions ; je me suis moi-même souvent trompé.

Cela dit, je veux d'abord vous signaler trois erreurs matérielles, qu'il faudrait peut-être corriger sur les exemplaires qui vous restent

p.38 - dernière ligne

au lieu de description lire construction

p.74 - un peu après le milieu

---

<sup>429</sup>In 1923 André Metz wrote a book on relativity, followed by another, Metz, André (1928) Temps, Espace, Relativité. Beauchesne, Paris. In the latter he attempted a popularization of Einstein's theory, to counter its poor reception by philosophers, especially Bergson in *Durée et simultanéité*. Later, Metz wrote various scientific works related to relativity. See for instance Metz, André (1967) Einstein et la philosophie des sciences. Arch. Int. Hist. Sci. 20:225-234.

<sup>430</sup>Finally it was René Lucas, director of the Ecole de Physique et Chimie de la Ville de Paris, who was elected, June 8, 1964, Free Academician.

au lieu de a son sens lire à son sens.

La troisième erreur, sur le théorème de M.Hadamard, à la p.64 est plus grave. L'énoncé correct est : les seuls points singuliers possibles de  $f(z)$  sont les produits  $\alpha\beta$ . La fausseté de votre énoncé résulte de l'exemple suivant :  $a_n$ , variable aléatoire uniformément répartie sur la circonférence  $|a_n| = 1$  ; les différents  $a_n$  indépendants ;  $b_n$  imaginaire conjugué de  $a_n$ . Donc  $a_n b_n = 1$  et  $f(z) = \frac{1}{1-z}$  n'a qu'un point singulier, tandis que  $\varphi(z)$  et  $\psi(z)$  ont presque sûrement le cercle  $|z| = 1$  comme coupure.

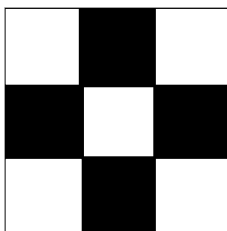
Je crois qu'il y a aussi une erreur au début de la p.65. Borel n'attribuait pas au mot "en général" le sens que vous dites, et le théorème auquel vous pensez n'a été signalé que par Steinhaus, vers 1930 je crois\*. [*in left margin*: \* pour les séries  $\sum c_n e^{-\theta_n} e^{ni\theta_n}$ , les  $\theta_n$  unif. répartis dans  $(0, 2\pi)$ ] Mais la démonstration simple est celle de Schwartz : la probabilité de l'existence d'un point singulier sur l'arc  $(\theta, \theta + \alpha)$  ne dépend que de  $\alpha$  et ne peut, d'après le théorème de Kolmogorov, être que 0 ou 1. La première hypothèse étant absurde, tout arc contient un point singulier.

Si je vous écris, c'est surtout pour vous reparler de la raréfaction. Certainement, c'est une notion importante, et j'ai eu tort d'en douter. Mais je crois qu'une meilleure classification des ensembles de mesure nulle est celle de Hausdorff.

Vous savez sans doute qu'étant donné un ensemble  $E$  dans  $\mathbb{R}^N$ , et  $\mathcal{F}$  une famille de sphères  $S_n$  de rayon  $r_n$  dont la réunion recouvre  $E$  (il s'agit des volumes des sphères et non des surfaces), la mesure de Hausdorff d'ordre  $\alpha$  de  $E$  est la borne inférieure de  $\sum r_n^\alpha$  quand on fait varier les sphères. Si  $N = 1$ , on remplace les sphères par des intervalles. On peut généraliser en introduisant des  $\varphi$ -mesures, bornes inférieures de  $\sum \varphi(r_n)$ ,  $\varphi(r)$  tendant vers zéro avec  $r$  ; il y a intérêt à supposer sa décroissance régulière. A condition de préciser cette notion de régularité, il y a une fonction régulière  $\varphi$  et une seule, à un facteur constant près, telle que la  $\varphi$ -mesure d'un ensemble  $E$  soit  $> 0$  et  $< \infty$ . Elle caractérise la raréfaction de  $E$ . Si on l'appelle  $\mathcal{M}_\varphi(E)$ , on a

$$\mathcal{M}_\varphi(E \cup E') \leq \mathcal{M}_\varphi(E) + \mathcal{M}_\varphi(E').$$

Si par exemple on considère la généralisation suivante de l'ensemble ternaire de Lebesgue : un carré, divisé en 9 carrés égaux ; on en raye 4 suivant le schéma ci-contre



et on recommence dans chaque carré conservé, et ainsi de suite indéfiniment. On obtient un ensemble auquel est associé la fonction  $r^\alpha$ ,  $\alpha$  étant défini par  $5.3^{-\alpha} = 1$ ,  $\alpha = \frac{\log 5}{\log 3}$ .

Si on multiplie les abscisses par  $k$  et les ordonnées par  $1/k$ , la mesure change, mais est multipliée par un facteur  $< c < \infty$  et  $> c' > 0$ , si  $\frac{\log \varphi(r)}{\log r}$  a une limite  $\alpha > 0$  et  $< \infty$ .

Comme application, j'ai résolu un problème de Dvoretzky en démontrant que la mesure d'ordre 2 d'un arc fini de la courbe d'un mouvement brownien dans  $\mathbb{R}^N$  est presque sûrement nulle. J'ai émis une hypothèse dont l'exactitude a été démontrée par S.J. Taylor concernant la fonction  $\varphi$  associée à cette courbe. C'est quelque chose comme  $\frac{r^2}{\log |\log r|}$  ou  $\frac{r^2}{\sqrt{\log |\log r|}}$ . Et cela pour  $N > 2$ . Je ne me rappelle plus exactement. Pour  $N = 2$ , le problème est plus difficile, et n'est pas résolu. Mon hypothèse était dans ce cas qu'il faut remplacer  $\log |\log |r|$  par  $|\log r|$  (Voyez 1952 ou 53 - Palerme ou le journal de Cantelli).<sup>431</sup>

L'étude au même point de vue de l'ensemble des points multiples d'ordre donné  $p$  de la courbe du mouvement brownien plan dans  $\mathbb{R}^2$  ( $p$  pouvant être infini ; ou bien  $p = 2$  ou  $3$  dans  $\mathbb{R}^3$ ) serait à mon avis très intéressante.

Voici quelques autres remarques :<sup>432</sup>

p. 55, dernier alinéa avant les exemples.

Ce qui caractérise les ensembles boréliens, c'est qu'ils sont toujours mesurables, pourvu que les intervalles le soient. En les complétant au sens de Lebesgue, on obtient des ensembles probabilisables ou non suivant la loi de probabilité considérée.

p.31, au milieu. Je ne comprends pas bien votre phrase sur les probabilités dénombrables, qui sont au contraire très bien définies dans votre Chapitre III. Je n'ai jamais compris pourquoi les probabilités absolument continues étaient appelées géométriques : mais il y a dans les probabilités dénombrables une idée différente de celle du pont entre les deux sortes de probabilités, ce pont étant réalisé par la fonction de répartition, que j'ai le premier, je crois, utilisé systématiquement.

Enfin une dernière remarque purement subjective. Il est historiquement certain que Borel est le premier à avoir énoncé en 1909 ses premiers mémoires sur les probabilités dénombrables.<sup>433</sup> Mais je peux affirmer que, sans énoncer explicitement ses théorèmes, je les appliquais sans hésiter quand j'en avais besoin. C'est sur la loi forte des grands nombres de Borel Cantelli que reposait ma communication de décembre 1907 à la Soc.math.,<sup>434</sup> sur une courbe sans tangente dont je ne savais pas encore qu'elle était de von Koch. J'ai déjà dit d'ailleurs (Bull.Inst. Statistique,

<sup>431</sup>Lévy, Paul (1953) La mesure de Hausdorff de la courbe du mouvement brownien. G. Ist. Ital. Attuari 16:1-37. The article of the Circolo di Palermo from 1952 considers random continued fractions.

<sup>432</sup>Clearly this refers to Fréchet, Maurice (1965) La vie et l'œuvre d'Emile Borel. Kundig, Geneva.

<sup>433</sup>See Borel, Emile (1909) Les probabilités dénombrables et leurs applications arithmétiques. Rend. Circ. Mat. Palermo 27:247-270.

<sup>434</sup>This conference was presented on March 11, 1908 to the Société mathématique de France.

1957)<sup>435</sup> que je regrettais de ne pas avoir publié mes idées sur cette courbe, tant en 1902 qu'en 1908.\*<sup>436</sup>

Malgré quelques objections, j'admire beaucoup l'effort que vous avez fait pour donner un exposé complet de l'œuvre de Borel, et votre souci d'impartialité. Je croyais connaître cette œuvre, et je m'aperçois que j'avais beaucoup de lacunes.

Excusez cette longue lettre. Jamais je n'aurais pu à l'Académie vous dire tout cela. Je ne demande de réponse que sur un point. J'ai peut-être fait quelques remarques nouvelles sur la mesure de Hausdorff. Je me demande si elles valent la peine d'être publiées. Avez-vous une opinion sur ce point ?

Bien cordialement à vous

P.Lévy

\* Et j'ajoute qu'en lisant le mémoire de Borel, mon impression a été, pour les 2 premiers chapitres : "Comment peut-on compliquer à ce point des choses si simples." Et, plus tard, je suis arrivé à la même conclusion pour le Chap.III<sup>437</sup> Pourtant, en général, Borel était simple et clair.

### Lettre 103

Marlotte -18 juin 1965

Mon cher ami,

Je vous récris tout de suite un mot parce que j'ai commis dans ma dernière lettre<sup>438</sup> quelques erreurs sur la  $\varphi$ -mesure de Hausdorff.

D'abord, la définition intéressante, que ce soit ou non celle de Hausdorff, est

$$\mu_\varphi(E) = \lim_{r \downarrow 0} \mu_\varphi(E | r)$$

avec  $\mu_\varphi(E | r) = \inf \sum_n \varphi(r_n)$ , les  $r_n$  étant les rayons des sphères dont la réunion recouvre  $E$ .<sup>439</sup>

Pour  $\varphi(r) = \varphi_n(r) = \omega_n r^n$ , volume de la sphère de rayon  $r$  dans l'espace  $\mathbb{R}^n$ , la  $\varphi_n$ -mesure est la mesure extérieure de Lebesgue dans  $\mathbb{R}^n$  ; ce qui prouve qu'il n'y a pas l'intérêt qu'on pourrait croire à remplacer les sphères par des cubes.

Toujours dans  $\mathbb{R}^n$ , si on a toujours  $\mu(E \cup E') \leq \mu(E) + \mu(E')$ , ce qui exclut les fonctions telles  $\varphi(r) = r^\alpha$  ( $\alpha < n$ ), la restriction que les sphères aient leur rayon

<sup>435</sup>Lévy is alluding to Lévy, Paul (1957) Réponse à Monsieur Dugué. Publ. Inst. Stat. Univ. Paris VI 3:163–176.

<sup>436</sup>In his autobiography (Lévy 1970, pp. 17–23), Lévy relates his independent discovery at age 16 of Von Koch's continuous curve without tangents. He found it the very year that Von Koch published his result. Clearly, as a student at the Lycée Saint-Louis, Lévy had not seen Von Koch's paper. Lévy's further work, up to 1908, on choosing at random a point on the curve is very close to the notion of a Brownian trajectory.

<sup>437</sup>Chap. III of the memoir cited in footnote 433. It treats the probability distribution of the successive integers in the continued fraction representation of a random number from the interval [0, 1]. Letter 82 makes the same criticism of Borel's book on set theory.

<sup>438</sup>This letter appears to have been lost.

<sup>439</sup>With  $r_n \leq r$  implied.

très petit ne change rien. Mais, pour  $\varphi(r) = r^\alpha$  ( $\alpha < n$ ), sans cette restriction, un ensemble situé dans un volume fini aurait toujours une  $\varphi$ -mesure finie, tandis qu'avec cette restriction sa mesure peut être infinie.

D'autre part, je vous ai dit que, pour tout ensemble  $E$ , il y avait toujours une fonction  $\varphi$  telle que  $0 < \mu_\varphi(E) < \infty$ . Or, d'une part, je ne pensais qu'aux ensembles situés dans un volume fini. D'autre part cette affirmation est liée à un problème que j'ai posé en 1928, qui n'est pas résolu (je crois) et que je crois très difficile.

J'appelle échelle de croissance toute famille  $\mathcal{F}$  de fonctions continues  $\varphi(x)$  définies chacune pour  $x$  assez grand, telles que, si  $\varphi$  et  $\psi \in \mathcal{F}$ ,  $\varphi - \psi$  soit  $\neq 0$  et par suite de signe constant pour  $x$  assez grand.

Je dis qu'une échelle est complète si on ne peut pas y mettre une fonction de plus.<sup>440</sup> Il est évident, pour ceux qui admettent l'axiome de Zermelo, qu'il existe des échelles complètes - et probable qu'il y en a une seule qui caractérise les croissances parfaitement régulières.<sup>441</sup> Mais le problème d'en former une est évidemment difficile.

Or, s'il y a une fonction  $\varphi$  associée à chaque ensemble borné  $E$ , cela serait une étape formidable dans la formation d'une échelle complète de croissance. Je ne crois pas que ce soit un résultat actuellement acquis.

Bien cordialement à vous P.Lévy

Je pense rentrer dimanche et être lundi à l'Académie.

Naturellement, la publication éventuelle d'une Note<sup>442</sup> dépendra maintenant de l'évolution de mes réflexions sur ce dont je vous parle aujourd'hui.

## Lettre 104

Paris - 38 Av. Th.Gautier

23 juin 1965

Mon cher Collègue,

Je réponds à votre question sur les fonctions à croissance régulière. C'est une notion un peu vague, que j'ai essayé de préciser en 1928, dans les *Annali di Matematica*.<sup>443</sup> Je n'ai pas réussi ; mais il y a tout de même quelques idées à retenir.

Je parlerai de fonctions  $f(x)$  régulières à l'infini, chacune étant définie dans  $(a, \infty)$ ,  $a$  pouvant varier d'une fonction à l'autre.

<sup>440</sup>Lévy introduced this definition in CRAS. February 22, 1926. In the ensuing years, he used the definition in several papers, especially in Lévy, Paul (1928) Fonctions à croissance régulière et itération d'ordre fractionnaire. *Ann. Mat.* 6:269–298. There he attempted, unsuccessfully, the explicit construction of a totally ordered family of subsets without using Zermelo's Lemma which belongs, in Lévy's words, to the idealistic point of view. Lévy's acerbic footnote in his note, CRAS 261:298 (1965), was an amusing condemnation of the *axiomatique moderne*. See Ageron, Pierre (2002) L'autre axiome du choix. *Rev. Hist. Math.* 8(1):113–140, and the next letter.

<sup>441</sup>See the next letter.

<sup>442</sup>It was published in CRAS on July 12th, 1965, and followed by an erratum ((1965) CRAS 261:2577).

<sup>443</sup>Lévy, Paul (1928) Fonctions à croissance régulière et itération d'ordre fractionnaire. *Ann. Mat.* 6:269–298.

Deux fonctions  $f(x)$  et  $g(x)$  ne peuvent pas être considérées à la fois comme régulières si  $f(x) - g(x)$  change de signe une infinité de fois. Ainsi  $e^x$  et  $f(x) = e^x + \sin x$ . C'est la première qui est régulière. L'irrégularité de la seconde se manifeste directement par le fait que  $f(x)$  et  $f'(x)$  [ou  $f(x)$  et  $f(x+a)e^{-a}$ , pour  $\sin \frac{a}{2} \neq 0$ ] ne peuvent pas être à la fois régulières.

De même l'irrégularité de  $e^x + e^{-x} \sin x$  se manifeste par deux dérivations, celle de  $x + \frac{1}{x} \sin \log x$  par deux intégrations.

Cela me conduit à l'idée d'un ensemble  $E$  de fonctions "régulières" vérifiant les conditions suivantes : toutes les fonctions algébriques réelles appartiennent à  $E$ . Si  $f(x)$  et  $g(x) \in E$ , il en est de même de  $f'(x)$  (dont l'existence est exigée), de  $\int^x f(u) du$ , de  $f[g(x)]$  si  $g(x) \rightarrow \infty$ ,  $f(x) + g(x)$ ,  $f(x)g(x)$ ,  $\frac{f(x)}{g(x)}$ . Toutes les fonctions obtenues par ces opérations en partant des fonctions algébriques vérifient la condition indiquée au début : la différence de deux de ces fonctions a un signe constant pour  $x$  assez grand. Je les considère comme parfaitement régulières.

Mais on n'a pas encore une échelle complète ; il faut encore extrapoler et interpoler. Là est la difficulté, que je n'ai pas résolue. J'ai seulement fait quelques remarques qui peuvent être utiles, notamment la suivante : si  $F(x) = x + \sum_2^{+\infty} a_n x^n$  est croissante et analytique dans  $(0, \infty)$ , il y a une fonction analytique et une seule  $f(x) = x + \frac{a_0}{2} x^2 + \dots$  qui vérifie l'équation  $f(f(x)) = F(x)$ . On peut ainsi définir l'itérée régulière d'ordre  $\frac{1}{2}$  de  $e^x - 1$  ; pour celle de  $e^x$ , c'est plus difficile.<sup>444</sup>

Après 1928, j'ai renoncé à ces problèmes, à cause de leur difficulté. Je me demande maintenant si les remarques de ma précédente lettre ne donnent pas un moyen de préciser la notion d'échelle complète, mais toujours, bien entendu, en utilisant l'axiome de Zermelo.

Bien cordialement

P.Lévy

J'ai cherché hier à l'Institut le mémoire d'Appert ; il est dans le Bull. des Sciences Math., 1936.<sup>445</sup> Il considère en effet différentes mesures possibles, dont celle de Hausdorff.

## Lettre 105

Mont Pèlerin s. Vevey - 18 août 1965

Mon cher confrère,

J'ai bien reçu votre lettre du 11 août. Vous ne me demandez de réponse que pour la mesure de Hausdorff généralisée, que j'appelle la  $\varphi$ -mesure, et désigne par  $\mu_\varphi(E)$ . La fonction  $\varphi$  n'intervenant que par son allure près du point  $r = 0$ ,  $\psi(r)/\varphi(r) \rightarrow c (r \downarrow 0)$  entraîne

$$\mu_\psi(E) = c \cdot \mu_\varphi(E),$$

<sup>444</sup>Iterating  $e^x$  led to a note to CRAS, February 28, 1927.

<sup>445</sup>Appert, Antoine (1936) Mesures normales dans les espaces distancés. Bull. Sci. Math. 60:329-352 & 368-380.



et cela en particulier si  $c = 0$  ou  $\infty$ . Donc, si la mesure de Lebesgue, qui est  $\mu_\varphi$  pour  $\varphi(r) = r$ ,<sup>446</sup> est finie (nulle ou positive), et si  $\psi(r) = r^\alpha$  ( $\alpha > 1$ ), la  $\psi$ -mesure est nulle, comme vous le dites. C'est l'ABC de la théorie.

Pour la courbe du mouvement brownien, un arc fini a une longueur infinie. Le résultat précédent ne s'applique pas. J'ai démontré en 1940 que, pour le mouvement brownien plan, la mesure de Lebesgue est nulle.<sup>447</sup> Le problème que Dvoretzky m'a posé en 1950 restait difficile pour les espaces à  $n$  dimension ( $n > 2$ ). Je l'ai résolu en 1951 ou 52,<sup>448</sup> et en même temps ai énoncé une hypothèse concernant la fonction  $\varphi$  telle que  $\mu_\varphi(E) > 0$  et  $< \infty$  pour un arc fini de cette courbe ; mon hypothèse, pour des raisons que j'indiquais, n'était pas la même pour  $n = 2$  et pour  $n > 2$ . Son exactitude, pour  $n > 2$ , a été établie par S.J.Taylor.<sup>449</sup> Le cas  $n = 2$  n'est pas résolu à ma connaissance.<sup>450</sup>

Naturellement, je suis d'accord avec ce que vous me dites sur la différence entre l'idée intuitive d'un théorème et sa démonstration. On peut discuter indéfiniment sur la valeur de l'impression de certitude que donne parfois une idée intuitive (le cas de Poincaré est connu ; pour moi cette impression a été très grande dans le cas du th.de Cramer ;<sup>451</sup> moins forte dans d'autres cas ; dans un cas, l'intuition subite a porté sur la démonstration du th. de Fisher et Riesz dont je connaissais l'énoncé ; je crois en avoir vu tous les détails en 2 ou 3 secondes).

Je ne réagis pas comme vous pour le th. de Picard. La précision apportée par Carathéodory au th. de Landau prouve que l'introduction de la fonction modulaire est dans la nature des choses, et la démonstration utilisant cette fonction est, sans doute moins élémentaire, mais plus naturelle. Celle de Borel me semble artificielle. Ce que j'admire, c'est l'important complément qu'il a par cette méthode apporté au th. de Picard.<sup>452</sup>

Enfin je reviens sur un point de vos lettres antérieures auquel je crois n'avoir pas répondu. Dans ma dernière phrase, p. 43, vous m'avez suggéré de mettre "abstraites"

<sup>446</sup>Lévy is thinking of one dimension.

<sup>447</sup>Theorem 12 of Lévy (1940).

<sup>448</sup>In his note, CRAS (September 10, 1951).

<sup>449</sup>Taylor, S.J. (1964) The exact Hausdorff measure of the sample path for planar Brownian motion. Proc. Camb. Philos. Soc. 60:253–258.

<sup>450</sup>Le Gall finally resolved it in 1984; see Le Gall, Jean-François (1985) Sur la mesure de Hausdorff de la courbe brownienne. In: Séminaire de Probabilités, vol. XIX, 1983-84, Lecture Notes in Mathematics, vol. 1123. Springer, Heidelberg.

<sup>451</sup>The theorem on the decomposition of Gaussian variables. See Letter 33, as well as the foreword by Dugué to the edition of Lévy's Oeuvres (Lévy 1973).

<sup>452</sup>This paragraph refers to the article written by Lévy in the commemorative volume for the tricentennial of the Academy of Sciences (see next letter—footnote 458). On pp. 178–180, Lévy discusses the theorem of Picard on the image of an entire function, which was clarified by Borel, Landau, Schottky and Carathéodory.

après “variables aléatoires”. Mais il y déjà “à valeurs dans un espace abstrait”. Je pense donc la répétition inutile.<sup>453</sup>

Bien cordialement à vous

P.Lévy.

Adresse j. 3 septembre

Hôtel Villa Castagnola

Lugano Suisse

retour à Paris vers le 15 septembre. Je serais heureux de recevoir à Paris entre le 15 et le 20 les précisions que je vous ai demandées sur 2 de vos théorèmes (sauf s'ils sont dans votre Notice que j'ai ; sans cela j'aurais du mal à les retrouver).

### Lettre 106

Paris - 38 Av. Théophile Gautier

19 septembre 1965

Mon cher collègue,

Je pense vous voir demain à l'Institut. Mais il vaut peut-être mieux que je mette sur le papier ce que j'ai à vous dire. Je vous écris donc cette lettre que je vous remettrai ou mettrai à la poste suivant que je vous voie ou non.

Ma note du 12 juillet En la relisant, je trouve plusieurs fautes. Je remettrai demain des erratum<sup>454</sup> dont voici la copie

	9	au lieu de grandes	lire petites
p.295, L	17	$\varphi(0)$	$\varphi'(0)$
	19	$< 0$	$> 0$
	13	$E_1$	$E_\mu$
p.297	15	$N_\mu$	$E_\mu$
L3 du	n°4	$\in$	$\subset$

[*En marge gauche* : p.296, après la formule (2), tous les  $r_\varphi$  doivent être  $\leq \rho$ .]

#### Réponse à votre lettre

Fonction d'allure régulière :<sup>455</sup> dont la courbe représentative n'est pas trop sinuose. Notion intuitive, un peu comme celle de raréfaction. J'ai essayé de la préciser dans un mémoire de 1928. Je pense qu'il peut y avoir encore quelque chose à dire sur ce sujet.

Le rôle d'Appert.<sup>456</sup> J'ai assisté au séminaire Hadamard. Ce qu'il a dit était incontestablement une généralisation de la mesure de Hausdorff.

Définition de  $S_\rho(E)$ . Il y a en effet un oubli, que je répare, grâce à vous.<sup>457</sup>

<sup>453</sup>Finally Lévy gave way (p. 212 of his article in the commemorative volume—see next letter), inserting the word “abstract” into the paragraph on Fréchet and Mourier’s work on variables with values in general spaces.

<sup>454</sup>CRAS 261:2577. This note of July 12, annotated by Lévy himself, can also be found in the Mathematical Research library at University Paris VI.

<sup>455</sup>See Letter 104.

<sup>456</sup>See Letter 104.

<sup>457</sup>The fact that the radii are  $\leq \rho$ . This correction is included with the errata.

p. 297 après la formule (3) c'est bien  $\mu_\varphi(E)$  qu'il faut lire. Un peu plus loin on a

$$\mu_\varphi/\mu_\psi = \lim_{r \rightarrow 0} \varphi/\psi$$

cette limite existant toujours pour les fonctions  $\varphi, \psi$  de la classe  $\mathcal{C}_N$ . Donc, si  $\mu_\psi = 1$ ,  $\varphi/\psi$  tend bien vers  $\mu_\varphi$ , comme je l'ai écrit.

#### Mon travail pour le tricentenaire<sup>458</sup>

J'ai reçu de nombreuses lettres de Denjoy, Montel, Mandelbrojt, Dieudonné, Schwartz, qui toutes réclament des additions. Je ne peux pas satisfaire tout le monde.

Il me semble qu'il faut tenir un grand compte de l'avis des trois derniers, qui sont plus au courant des progrès récents de l'analyse qu'aucun des 6 membres actuels de la Section de géométrie. C'est parce que j'avais conscience de l'impossibilité où j'étais d'exposer certains chapitres de l'analyse moderne, et que la collaboration obtenue était insuffisante que j'ai pris le parti de faire tirer une rédaction provisoire, pour provoquer des réactions. Elles sont venues. Il faut en tenir compte.

J'ajoute - j'espère ne pas vous fâcher, que les additions demandées par Mandelbrojt, Dieudonné et Schwartz portent sur de grandes choses - et il me semble que plusieurs de vos observations portent sur de petites choses. Je ne peux donc faire qu'une partie des additions ou modifications que vous m'avez suggérées.

Naturellement, je rectifierai tout ce qui paraît inexact. Mais j'ai été surpris que vous puissiez être intervenu dans la décomposition d'une fonction à variations bornées en trois termes. En êtes vous sûr ?<sup>459</sup>

D'accord avec Denjoy, je pense que cette décomposition est triviale, à partir du moment où l'on a

- 1° la notion de fonction absolument continue
- 2° celle de fonction continue ne variant que sur un ensemble de mesure nulle
- 3° celle de fonction totalement discontinue.

Je serais surpris si ces notions - du moins les deux premières- ne se trouvaient pas dans Lebesgue. La 3ième est plus ancienne.

Je pense que je suis, peut-être avec von Mises, le 1er à avoir appliqué cette décomposition au calcul des probabilités.<sup>460</sup> Mais je n'en parle pas. Il n'y a pas à être fier d'avoir appliqué des résultats dûs à d'autres.

Je reviens à vos travaux. J'ai trouvé dans votre Notice la décomposition d'une fonction analytique en deux termes, dont vous m'avez parlé. Je vais essayer de

<sup>458</sup>In 1966 the Academy of Sciences of Paris celebrated its 300th anniversary. Two commemorative volumes retracing the history of the different disciplines were published by Gauthier-Villars. The article on the Geometry Section (Vol. I, pp. 143–212) treated mathematics and was written by Lévy, with the collaboration of Cartan, Dieudonné, Ehresmann, Schwartz and Vincensini.

<sup>459</sup>Lévy gave in and in the Geometry Section article, p. 166, cited Fréchet along with Lebesgue and de La Vallée Poussin.

<sup>460</sup>During the establishment of the Lévy-Khinchin formula for Lévy processes.

l'ajouter. Je dis essayer, parce j'ai un peu peur que mon exposé soit jugé trop long et qu'on exige des suppressions.

J'ajoute aussi, comme vous le désirez, l'énoncé du théorème d' Hadamard sur les nombres premiers, et une indication sur le rôle de Poincaré dans le calcul des probabilités.

Par contre, il ne me paraît pas possible d'entrer dans les détails sur la définition de la différentielle. Pour les conditions de Cauchy, elles sont suffisantes pour l'holomorphie dans un domaine à *contour simple*, restriction que j'avais oubliée et que je rajoute.

Pour l'arithmétique des lois de probabilités, je crois correct de dire qu'elle recherche les décompositions  $X = U + V$ ,  $U$  et  $V$  étant indépendantes. Il est bien entendu qu'il y a toujours la solution triviale  $X = (X - a) + a$ , et que la difficulté est de chercher les autres. Mais je n'ai pas voulu entrer dans des détails.

A ce sujet, un élève de D.G.Kendall vient de démontrer que si  $X$  est uniformément réparti dans un intervalle, il est impossible que les lois de  $U$  et  $V$  soient toutes les deux absolument continues.

Voilà l'essentiel de ce que je voulais vous dire. J'espère que vous ne serez pas fâché que je n'ai suivi qu'une partie de vos suggestions. Vous m'en avez d'ailleurs déjà assuré.

Je ne suis plus préoccupé du fait que j'ai réveillé la vieille querelle Julia Montel pour les familles normales.<sup>461</sup> Je ne peux pas, dans un exposé fait un peu au nom de l'Académie, me poser en arbitre ; ni supprimer toute allusion aux familles normales. Je crains de les mécontenter tous les deux.

Bien cordialement à vous

P.Lévy

[*In left vertical margin*: On a dû vous dire que je vous avais demandé à Sceaux vendredi. J'y étais allé faire la connaissance de mon arrière petite fille (qui a 1 mois aujourd'hui). C'était près du Boulevard Colbert.]

## Lettre 107

Paris - 38 Av.Th.Gautier

19-10-65

Mon cher collègue et confrère,

La Commission du tricentenaire<sup>462</sup> s'est réunie hier après le Comité secret (jusqu'à 18h15!). Il a été décidé que chaque Section devrait prendre la responsabilité du travail fait en son nom.

Montel, Garnier et moi avons tenu une réunion préliminaire. Denjoy est dans le Midi. Je préviens Julia qui voudra peut-être nous réunir le mercredi 27- après la

---

<sup>461</sup>On page 180 of his article for the tricentenary of the Academy of Sciences (see note 458), Lévy, in a note at the bottom of the page, affirmed Montel's prior publication on this subject, written in 1914 and published in 1916. Nevertheless in 1918 Julia won a prize from the Academy on a related subject. In 1965 both Julia and Montel were at the Academy. See M.Audin: *Fatou, Julia, Montel, The Great Prize of Mathematical Sciences of 1918, and Beyond*. Springer (2011).

<sup>462</sup>See Letter 106.

séance (remise à mercredi à cause de la séance de l'Institut du lundi). Je crois bien faire de vous prévenir de cette éventualité.

J'ai dû exposer longuement à Julia les critiques auxquelles mon travail a donné lieu. Il est trop long, et je n'aurais pas dû citer les vivants !<sup>463</sup> Excusez-moi de ne pas vous donner plus de détails. Il est trop tard pour tout recommencer.

Bien cordialement

P.Lévy

---

<sup>463</sup>It is somewhat comic to observe that in his article for the tricentennial of the Academy (see Letter 106, note 458), Lévy did not omit to frequently quote his own works. This might well have irritated Julia.

# References

- Aldrich J (2007) *J Électron Hist Probab Stat* 3(2)
- Armatte M (2002) Maurice Fréchet statisticien, enquêteur et agitateur public. *Rev Histoire Math* 7:7–66
- Bachelier L (1900) Théorie de la spéculation. *Ann Sci Éc Norm Super* 17:21–86
- Barbut M, Mazliak L (2008) Commentary on the notes for Paul Lévy's 1919 lectures on the probability calculus at the Ecole Polytechnique. *J Électron Hist Probab Stat* 4(1)
- Бернштейн СН (1917) Опыт аксиоматического обоснования теории вероятностей. Сообщение Харьковского Математического Общества 15
- Бернштейн СН (1927) Современное состояние теории вероятностей и её приложений. In: Труды всероссийского съезда математиков. pp 50–63
- Betti R (2010) *La matematica come abitudine di pensiero*. PRISTEM, Milano
- Blanc-Lapierre A, Fortet R (1953) *Théorie des fonctions aléatoires*. Masson, Paris
- Borel E (1972) *Œuvres, Introduction et Bibliographie* par M. Fréchet en 4 Tomes. CNRS, Paris
- Bouleau N (1986) La jonction entre la théorie du potentiel et des probabilités. *Sémin Hist Math* 8:43–62
- Bouligand G (1926) *Fonctions harmoniques. Principes de Picard et de Dirichlet*. Mémorial des Sciences Mathématiques, vol XI. Gauthier-Villars, Paris
- Brelot M (1964) *Eléments de théorie classique du potentiel*. Centre de documentation universitaire, Paris
- Bru B (1999) Borel, Lévy, Neyman, Pearson et les autres. *Mataplí* 60:51–60
- Bru B (2001) Bachelier et son époque: une conversation avec Bernard Bru par Mourad Taqqu. *J Soc Fr Stat* 142:3–40
- Bru B (2003) Souvenirs de Bologne. *J Soc Fr Stat* 144(1-2):135–226
- Bru B, Bru M-F, Chung KL (2009) Borel and the St Petersburg martingale. *J Électron Hist Probab Stat* 5(1)
- Catellier R, Mazliak L (2012) The emergence of French probabilistic statistics. Borel and the Institut Henri Poincaré around the 1920s. *Rev Histoire Math* 18:271–335
- Chapman S (1928) On the Brownian displacements and thermal diffusion of grains suspended in a non-uniform fluid. *Proc R Soc Lond Ser A, Math Phys Sci* 119:34–54
- Chaumont L, Mazliak L, Yor M (2007) Some aspects of the probabilistic works. In: Charpentier E, Lesne A, Nikolski NK (eds) *Kolmogorov's heritage in mathematics*. Springer, London
- Choquet G (1990) La vie et l'œuvre de Marcel Brelot. *Sémin Hist Math* 11:1–31
- Chung KL (1995) *Green, Brown and probability*. World Scientific, Singapore
- Collective (1948) *Математика в СССР за 30 лет*. Государственное издание технической теоретической литературы, Москва
- Courtault J-M, Kabanov Y (2002) *Louis Bachelier: Aux origines de la finance mathématique*. Presses Univ. Franc-Comtoises, Besancon

- De la Pradelle A (1988) Potentiel. Encyclopedia Universalis, Paris
- Dellacherie C (1980) Un survol de la théorie de l'intégrale stochastique. *Stoch Process Appl* 10:115–144
- Dellacherie C, Meyer P-A (1975) Probabilités et Potentiel (4 tomes, 1975–1990). Hermann, Paris
- Демидов СС, Левшин БВ (1999) Дело академика Н.Н. Лузина. РХГИ, Москва
- Демидов СС, Паршин АН, Половинкин СМ (1989) О переписке Н.Н. Лузина с П.А. Флоренским. *Историческо-Математические Исследования* 31:116–124
- Dieudonné J (1978) *Abrégé d'histoire des mathématiques*. Hermann, Paris
- Dixmier J (1985) *Topologie générale*. Presses Universitaires de France, Paris
- Doob JL (1953) *Stochastic processes*. Wiley, New York
- Doob JL (1980) *Classical potential theory and its probabilistic counterpart*. Springer, New York
- Dugac P (2003) *Les mathématiques dans la vie de leur temps, autour de la notion de limite et de voisinage*. Vuibert, Paris
- Durand A, Mazliak L (2011) Volterra's Proclusionone as a source of Borel's interest for probability. *Centaurus* 53:306–333
- Einstein A (1905) Über die von der molekularkinetischen Theorie des Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann Phys* 19:549–560
- Fortet R (1943) Les fonctions aléatoires du type de Markoff associées à certaines équations aux dérivées partielles de type parabolique. *J Math Pures Appl* 22:177–243
- Fortet R (1948) Opinions modernes sur les fondements du Calcul des Probabilités. In: Le Lionnais F (ed) *Les grands courants de la pensée mathématique*. Hermann, Paris
- Fréchet M (1928) *Les espaces abstraits*. Gauthier-Villars, Paris
- Fréchet M (1955) *Les mathématiques et le concret*. Presses Universitaires de France, Paris
- Fréchet M, Halbwachs M (1924) *Le calcul des probabilités à la portée de tous*. Dunod, Paris
- Graham LR (1993) *Science in Russia and the Soviet Union*. Cambridge University Press, Cambridge
- Guiraldencq P (1999) Emile Borel. L'espace et le temps d'une vie sur deux siècles. Librairie Blanchard, Paris
- Havlova V, Mazliak L, Šišma P (2005) Le début des relations mathématiques franco-tchécoslovaques vu à travers la correspondance Fréchet-Hostinský. *J Électron Hist Probab Stat* 1(1)
- Hawkins T (1970) *Lebesgue's theory of integration*. Chelsea, New York
- Itô K, Mc Kean HP Jr (1965) *Diffusion processes and their sample paths*. Springer, New York
- Kahane J-P (1998) *Le Mouvement Brownien, Matériaux pour l'histoire des mathématiques au XXème siècle*. Société Mathématique de France, Paris
- Kakutani S (1944a) On Brownian motion in  $n$ -space. *Proc Imp Acad (Tokyo)* 20:648–652
- Kakutani S (1944b) Two-dimensional Brownian motion. *Proc Imp Acad (Tokyo)* 20:706–714
- Хинчин АЯ (1926) Идеи интуиционизма и борьба за предмет в современной математике. *Вестник Коммунистической Академии* 16:184–192
- Колмогоров АН (1925) О принципе tertium non datur. *Математический Сборник*
- Колмогоров АН (1933) *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin
- Колмогоров АН (1947) Роль русской науки в развитии теории вероятностей. *Ученые Записки МГУ* 91
- Lévy P (1911) *Les équations intégro-différentielles définissant des fonctions de ligne*. Thèse de la Faculté des sciences, Paris
- Lévy P (1922) *Leçons d'Analyse Fonctionnelle*. Gauthier-Villars, Paris
- Lévy P (1925) *Calcul des probabilités*. Gauthier-Villars, Paris
- Lévy P (1937) *Théorie de l'Addition des Variables aléatoires*. Gauthier-Villars, Paris
- Lévy P (1940) *Le mouvement brownien plan*. *Am J Math* 62:487–550
- Lévy P (1948) *Processus stochastiques et Mouvement brownien*. Gauthier-Villars, Paris
- Lévy P (1970) *Quelques aspects de la pensée d'un mathématicien*. Blanchard, Paris
- Lévy P (1973) *Œuvres, éditées sous sa direction par D. Dugué, avec la collaboration de P. Dehevels et M. Ibero (1973–1980)*. Gauthier-Villars, Paris
- Locker B (2001) *Paul Lévy: la période de guerre*. Thèse, Université Paris V

- Loève M (1973) Paul Lévy (1886–1971). *Ann Probab* 1:1–18
- Maasshal M (1999) Bourbaki, une secte de mathématiciens. Pour la Science, Paris
- Mandelbrot B (1975) Les objets fractals. Flammarion, Paris
- Maz'ja V, Shaposhnikova T (1998) Jacques Hadamard: a universal mathematician. Am Math Soc, Providence
- Mazliak L (2007) On the exchanges between Wolfgang Doeblin and Bohuslav Hostinský. *Rev Histoire Math* 13:155–180
- Mazliak L (2009) How Paul Lévy saw Jean Ville and martingales. *J Électron Hist Probab Stat* 5(1)
- Mazliak L (2011) The ghosts of the école normale, life, death and destiny of René Gateaux (to appear). <http://arxiv.org/abs/math/0701490>
- Mazliak L (2014) Poincaré's odds. In: Duplantier B, Rivasseau V (eds) 'Poincaré, 1912–2012', Poincaré seminar XVI, 24 November 2012. Progress in mathematical physics. Birkhäuser, Basel
- Mazliak L, Shafer G (2011) What does the arrest and release of Emile Borel and his colleagues in 1941 tell us about the German Occupation of France? *Sci Context* 24:587–623
- Mazliak L, Tazzioli R (2009) Mathematicians at war. Springer, Dordrecht
- Nevanlinna R (1936) Eindeutige analytische Funktionen. Springer, Berlin
- Perrin J (1912) Les atomes. Félix Alcan, Paris
- Petrovski IG (1933) Über das Irrfahrtproblem. *Math Ann* 109:425–444
- Petrovski IG (1935) Zür ersten Randvertaufgabe der wärmeleitungsgleichung. *Compos Math* 1:383–419
- Picard E (1941) La vie et l'œuvre de Paul Villard et de Georges Gouy. *Mém Acad Sci Paris* 63:21–24
- Pier J-P (1994) Developments of mathematics 1900–1950. Birkhäuser, Basel
- Protter P (1990) Stochastic integration and differential equations. Springer, New York
- Roussel Y (1989) L'histoire d'une politique des inventions 1887–1918. Cahiers pour l'histoire du CNRS 1939–1989. Cahier 1989-3
- Seneta E, Markov AA (2001). In: Heyde CC, Seneta E (eds) Statisticians of the centuries. Springer, New York
- Schwartz L (2001) A mathematician grappling with his century. Birkhäuser, Basel
- Shafer G (2009) The education of Jean-André Ville. *J Electron Hist Probab Stat* 5(1)
- Shafer G, Volk V (2006) The sources of Kolmogorov's Grundbegriffe. *Stat Sci* 21(1):70–98
- Shiryayev AN (1989) Kolmogorov: life and creative activities. *Ann Probab* 17(3):866–944
- Слущкий ЕЕ (1912) Теория корреляции и элементы учения о кривых распределения. Известия Киевского Коммерческого Института
- Taton R (1961) Histoire générale des Sciences. La science contemporaine (tomes III-1 et III-2; 1961–1964). Presses Universitaires de France, Paris
- Вавилов С (1946) Советская наука на новом этапе. Издательство Академия Наук, СССР
- Von Plato J (1994) Creating modern probability. Cambridge University Press, Cambridge
- Vucinich A (1999) Mathematics and dialectics in the Soviet Union: the pre-Stalin period. *Hist Math* 26:107–124
- Wiener N (1923) Differential space. *J Math Phys* 58:131–174
- Юшкевич АА (1968) История математики в России до 1917 года. Гл. 19: Теория вероятностей. Наука, Москва
- Žust M (2002) À la recherche de la vérité vivante. L'expérience religieuse de Pavel A Florenskij (1882–1937). Липа, Роме



# Index of Persons Quoted in the Letters

## A

Abel, 79  
Appell, 156  
Appert, 200, 202  
Arzela, 162

## B

Bachelier, 132, 135  
Baer, 71  
Banach, 189, 190  
Bass, 123, 126, 130, 132, 133, 136, 195  
Bauer, 79  
Becquerel, 165  
Berge, 191  
Bergson, 195  
Bernoulli, 78, 97, 99, 100, 104, 122, 135  
Bernstein, F., 71  
Bernstein, S.N., 114, 119, 125, 127, 129, 136  
Bertrand, 78  
Bienaymé, 71, 78, 97  
Blanc-Lapierre, 140  
Borel, 63, 78, 97, 99, 100, 105, 113, 114, 139, 142, 143, 158, 163–165, 170–174, 176–181, 185, 187, 195–198, 201  
Bouligand, 140, 144, 145, 178  
Bourbaki, 130  
Boutet, 135  
Bravais, 130

## C

Cantelli, 97–101, 139, 142, 143, 158, 164, 197  
Cantor, 171, 185  
Caquot, 159, 161, 162  
Caratheodory, 201  
Cartan, 114, 125, 128, 173, 193  
Carvallo, 139  
Cauchy, 78, 79, 150, 204

Cerf, 77

Césaro, 115  
Chapman, 132, 135  
Chazy, 155  
Chittenden, 68, 69  
Choquet, 173, 191  
Chung, 153, 169  
Copeland, 109–112, 114, 115, 131, 134–136, 140  
Cramer, 83, 100, 102, 105, 113, 115, 201

## D

Daniell, 70, 72  
Darmois, 97, 99, 101, 143  
De Finetti, 93, 115, 184  
De Moivre, 121, 135  
Denjoy, 140, 164, 203, 204  
Dickstein, 95  
Dieudonné, 203  
Dirichlet, 62, 63, 120  
D'Ocagne, 156  
Doebelin, 96, 115, 123, 136, 149, 150, 154  
Doob, 153, 154, 169, 194  
Dufour, 113  
Dugué, 166, 177, 195  
Dvoretzky, 182, 184, 192, 193, 197, 201

## E

Erdős, 145, 182, 184, 192, 193  
Esclangon, 119  
Evans, 63  
Eyraud, 178

## F

Fehr, 109  
Feldheim, 96  
Fermat, 179

Fèvre, 139, 153

Fischer, 91

Fisher, 201

Fortet, 128, 153, 154, 177, 191, 193, 195

Fouillade, 144, 145

Fourier, 76, 79, 130

Freyman, 128

Frobenius, 152

## G

Garnier, 155, 158, 164, 193, 204

Gateaux, 57, 59–62, 66, 67, 69, 190

Gau, 113

Gauss, 71, 75, 80, 81, 83, 87, 90, 99, 102, 120, 121, 130, 159, 160, 163, 182, 183

Gibrat, 163

Glivenko, 98–101, 143

Gosse, 118, 119

Green, 56, 59, 60, 63, 65, 120

## H

Hadamard, 55–57, 63, 68, 83, 85, 148, 149, 154, 164, 172, 174, 180, 194, 196, 202, 204

Hamilton, 55, 63

Hausdorff, 76, 80, 90, 159, 160, 196, 198, 200, 202

Heim, 192

Hellinger, 166

Hennequin, 177

Hilbert, 188, 191

Hitler, 183

Hostinsky, 83, 85

Humbert, 157

## J

Jacobi, 55, 63

Jessen, 101

Jolibois, 156, 164, 165

Jordan, 99, 175

Julia, 155–157, 204, 205

## K

Kac, 146

Kahane, 170

Kakutani, 182, 184, 192, 193

Kampé de Fériet, 124, 126, 130, 154

Kendall, 175, 204

Khinchin, 92, 102, 104, 122, 128, 130, 131, 133, 135, 194

Kolmogorov, 92, 95–97, 115, 116, 119, 120, 134–136, 150, 153, 168, 196

Krasner, 105, 106

Kuratowski, 189

## L

Landau, 145, 201

Langevin, 78

Laplace, 59, 60, 78, 79, 81, 90, 98, 99, 107, 109, 110, 120, 121, 130, 135, 182, 183

Laval, 193

Le Lionnais, 190

Lebesgue, 64, 72, 78, 79, 99, 115, 138, 171, 180, 189, 195, 196, 198, 201, 203

Lefort, 142

Leprince-Ringuet, 165

Leray, 193

Levi, 141

Lévy, J.-C., 191

Levy, M., 167

Lichnerowicz, 190, 193

Liénard, 106

Lindeberg, 94

Liouville, 78

Lipschitz, 148

Loève, 114, 115, 119, 121, 122, 128, 130, 132, 135, 177, 189, 194

Lyapunov, 76, 135

## M

Mandelbrojt, 192, 193, 203

Mandelbrot, 177

Marcinkiewicz, 115

Markov, 154

Metron, 138

Metz, 195

Michal, 189, 190

Mineur, 77

Molière, 162

Montel, 114, 141, 161, 162, 184, 203, 204

Moore, E.H., 68, 190

Mourier, 189

## N

Nevanlinna, 113

## O

Ottaviani, 159, 164

## P

Peccot, 55

Perrin, F., 116, 132

Petrovski, 119, 120

Picard, 78, 89, 90, 201

Piron, 114, 137

Plateau, 182

Poincaré, 78, 83, 85, 111, 201, 204

Poisson, 78, 99, 102

Pollaczek, 142

Polya, 78, 100, 101, 126, 145  
Prohorov, 168

**R**

Rademacher, 131  
Radon, 136  
Rajkov, 102  
Riemann, 109, 110, 115  
Riesz, 91, 144, 145, 201

**S**

Schwartz, 113, 120, 123, 128, 129, 142, 151,  
169, 203  
Schwarzshild, 82  
Slutski, 121, 123, 125, 130, 135, 138, 140  
Soury, 167, 169  
Steinhaus, 92, 93, 95, 100, 101, 115  
Stieltjes, 110, 112, 114–116, 136  
Subbotin, 71–73

**T**

Taylor, S.J., 184, 197, 201  
Tchebichev, 78  
Tornier, 99  
Tortrat, 177  
Trefouël, 192  
Tricomi, 113

**V**

Valiron, 123, 140  
Van Dantzig, 183  
Vilentchuk, 166, 167  
Villat, 70, 103, 167  
Ville, 98, 99, 101, 117, 118, 123, 132, 138,  
139, 143, 194, 195  
Vitali, 94  
Volterra, 61, 62  
Von Koch, 197  
Von Mises, 111, 181, 194, 203

**W**

Wahl, 133  
Wald, 159, 194  
Wavre, 109, 110, 127, 129, 131  
Wehrlé, 124, 126, 130  
Whittaker, 76  
Wiener, 70, 72, 92, 184

**Y**

Young, 116, 125, 127

**Z**

Zermelo, 179, 199, 200