A Festschrift for Herman Rubin
Institute of Mathematical Statistics
Lecture Notes – Monograph Series
Vol. 45 (2004) 75–91

A short history of stochastic integration and mathematical finance: The early years, 1880–1970

Robert Jarrow[1] **and Philip Protter**[*1]

Cornell University

Abstract: We present a history of the development of the theory of Stochastic Integration, starting from its roots with Brownian motion, up to the introduction of semimartingales and the independence of the theory from an underlying Markov process framework. We show how the development has influenced and in turn been influenced by the development of Mathematical Finance Theory. The calendar period is from 1880 to 1970.

The history of stochastic integration and the modelling of risky asset prices both begin with Brownian motion, so let us begin there too. The earliest attempts to model Brownian motion mathematically can be traced to three sources, each of which knew nothing about the others: the first was that of T. N. Thiele of Copenhagen, who effectively created a model of Brownian motion while studying time series in 1880 [81].[2]; the second was that of L. Bachelier of Paris, who created a model of Brownian motion while deriving the dynamic behavior of the Paris stock market, in 1900 (see, [1, 2, 11]); and the third was that of A. Einstein, who proposed a model of the motion of small particles suspended in a liquid, in an attempt to convince other physicists of the molecular nature of matter, in 1905 [21](See [64] for a discussion of Einstein's model and his motivations.) Of these three models, those of Thiele and Bachelier had little impact for a long time, while that of Einstein was immediately influential.

We go into a little detail about what happened to Bachelier, since he is now seen by many as the founder of modern Mathematical Finance. Ignorant of the work of Thiele (which was little appreciated in its day) and preceding the work of Einstein, Bachelier attempted to model the market noise of the Paris Bourse. Exploiting the ideas of the Central Limit Theorem, and realizing that market noise should be without memory, he reasoned that increments of stock prices should be independent and normally distributed. He combined his reasoning with the Markov property and semigroups, and connected Brownian motion with the heat equation, using that the Gaussian kernel is the fundamental solution to the heat equation. He was able to define other processes related to Brownian motion, such as the maximum change during a time interval (for one dimensional Brownian motion), by using random walks and letting the time steps go to zero, and by then taking

*Supported in part by NSF grant DMS-0202958 and NSA grant MDA-904-03-1-0092.

[1]School of Operations Research and Industrial Engineering, Cornell University, 219 Rhodes Hall, Ithaca, NY 14853, USA. e-mail: pep4@cornell.edu

Keywords and phrases: stochastic integration, semimartingales, martingales, Brownian motion, Markov processes, Black-Scholes, options, warrants, contingent claims, hedging strategies, Bachelier, homogeneous chaos, history of mathematics.

AMS 2000 subject classifications: 01A60, 60H05, 60H30, 60G44, 60G35, 60G46, 91B70, 91B28, 91B99, 60J45, 60J55, 60J65.

[2]This was called to our attention by Ragnar Norberg, whom we thank, and the contributions of Thiele are detailed in a paper of Hald [30].

limits. His thesis was appreciated by his mentor H. Poincaré, but partially due to the distaste of studying economics as an application of mathematics, he was unable to join the Paris elite, and he spent his career far off in the provincial capital of Besançon, near Switzerland in Eastern France. (More details of this sad story are provided in [11]).

Let us now turn to Einstein's model. In modern terms, Einstein assumed that Brownian motion was a stochastic process with continuous paths, independent increments, and stationary Gaussian increments. He did not assume other reasonable properties (from the standpoint of physics), such as rectifiable paths. If he had assumed this last property, we now know his model would not have existed as a process. However, Einstein was unable to show that the process he proposed actually did exist as a mathematical object. This is understandable, since it was 1905, and the ideas of Borel and Lebesgue constructing measure theory were developed only during the first decade of the twentieth century.

In 1913 Daniell's approach to measure theory (in which integrals are defined before measures) appeared, and it was these ideas, combined with Fourier series, that N. Wiener used in 1923 to construct Brownian motion, justifying after the fact Einstein's approach. Indeed, Wiener used the ideas of measure theory to construct a measure on the path space of continuous functions, giving the canonical path projection process the distribution of what we now know as Brownian motion. Wiener and others proved many properties of the paths of Brownian motion, an activity that continues to this day. Two key properties relating to stochastic integration are that (1) the paths of Brownian motion have a non zero finite quadratic variation, such that on an interval (s,t), the quadratic variation is $(t-s)$ and (2) the paths of Brownian motion have infinite variation on compact time intervals, almost surely. The second property follows easily from the first. Note that if Einstein were to have assumed rectifiable paths, Wiener's construction would have essentially proved the impossibility of such a model. In recognition of his work, his construction of Brownian motion is often referred to as *the Wiener process*. Wiener also constructed a multiple integral, but it was not what is known today as the "Multiple Wiener Integral": indeed, it was K. Itô, in 1951, when trying to understand Wiener's papers (not an easy task), who refined and greatly improved Wiener's ideas [36].

The next step in the groundwork for stochastic integration lay with A. N. Kolmogorov. The beginnings of the theory of stochastic integration, from the non-finance perspective, were motivated and intertwined with the theory of Markov processes, in which Kolmogorov, of course, played a fundamental role. Indeed, in 1931, two years before his famous book establishing a rigorous mathematical basis for Probability Theory using measure theory, Kolmogorov refers to and briefly explains Bachelier's construction of Brownian motion ([41], pp. 64, 102–103). It is this paper too in which he develops a large part of his theory of Markov processes. Most significantly, in this paper Kolmogorov showed that continuous Markov processes (diffusions) depend essentially on only two parameters: one for the speed of the drift and the other for the size of the purely random part (the diffusive component). He was then able to relate the probability distributions of the process to the solutions of partial differential equations, which he solved, and which are now known as "Kolmogorov's equations." Of course, Kolmogorov did not have the Itô integral available, and thus he relied on an analysis of the semigroup and its infinitesimal generator, and the resulting partial differential equations.[3]

[3] J. L. Doob [17] has complained that the PDE methods of Kolmogorov and Feller used to study Markov processes have often been called "analytic", whereas the method of stochastic differentials

After Kolmogorov we turn to the fascinating and tragic story of Vincent Doeblin (born Wolfgang Döblin) the son of the author Alfred Döblin, who wrote *Berlin Alexanderplatz* for example. The Döblin family fled the Nazis from Germany, first to Switzerland, and then to Paris. Wolfgang changed his name to Vincent Doeblin, and became a French citizen, finishing his schooling there and being quickly recognized as an extraordinary mathematical talent. In the late 1920's Probability Theory was becoming stylish among mathematicians, especially in the two centers, Moscow and Paris. Doeblin joined the probabilists, working on Markov chains and later Markov processes.[4] Doeblin wanted to construct a stochastic process with continuous paths that would be consistent with Kolmogorov's analytic theory of transition probabilities for Markov processes. He ultimately developed a framework to study them which was prescient in regards to future developments. However Doeblin was drafted, and he volunteered to go to the front. Before he went he sketched out his ideas and he put this work in the safe of the National Academy of Science of France, to be opened only by him or else after 100 years. As the Maginot line fell, to avoid sharing his ideas with the Nazis Doeblin first burned his notes, and then he took his own life. The academy safe was opened only in May 2000, at the request of his brother, Claude Doeblin. It was only then that the far reaching vision of his work became apparent. In those notes, he utilized the new concept of martingales proposed by J. Ville only in 1939 [84] and understood the importance of studying sample paths, instead of relying exclusively on distributional properties. One idea he had was to run Brownian motion by a random clock: what is known today as a time change. The change of time was then related to the diffusion coefficient, and in this way he was able to give a modern treatment of diffusions decades before it was developed otherwise. [5]

We turn now to Kiyosi Itô, the father of stochastic integration. We will not attempt to reproduce the beautiful summary of his work and contributions provided in 1987 by S. R. S. Varadhan and D. W. Stroock [83], but instead give a short synopsis of what we think were key moments.[6] No doubt an attempt to establish a true *stochastic differential* to be used in the study of Markov processes was one of Itô's primary motivations for studying stochastic integrals, just as it was Döblin's before him, although of course Döblin's work was secret, hidden away in the safe of the French Academy of Science. Wiener's integral did not permit stochastic processes as integrands, and such integrands would of course be needed

---

introduced by Itô has in contrast been called "probabilistic". Indeed, he writes, "It is considered by some mathematicians that if one deals with analytic properties and expectations then the subject is part of analysis, but that if one deals with sample sequences and sample functions then the subject is probability but not analysis". Doob then goes on to make his point convincingly that both methods are probability. (Doob's criticism is likely to have been partially inspired by comments of the second author.) Nevertheless, we contend that the methods of Itô changed the *probabilistic intuition* one develops when studying Markov processes.

[4] J. Doob references his fundamental work on Markov chains and Markov processes extensively in his book [17], for example. Paul Lévy wrote of him in an article devoted to an appreciation of his work after his death: "Je crois pouvoir dire, pour donner une idée du niveau où il convient de le situer, qu'on peut compter sur les doigts d'une seule main les mathématiciens qui, depuis Abel et Galois, sont morts si jeunes en laissant une oeuvre aussi important". Translated: 'I can say, to give an idea of Doeblin's stature, that one can count on the fingers of one hand the mathematicians who, since Abel and Galois, have died so young and left behind a body of work so important.' See [44]

[5] The second author is grateful to Marc Yor for having sent to him his beautiful article, written together with Bernard Bru [6]. This article, together with the companion (and much more detailed) article [7], are the sources for this discussion of Doeblin. In addition, the story of Doeblin has recently been turned into a book in biographical form [65].

[6] The interested reader can also consult [66].

if one were to represent (for example) a diffusion as a solution of a stochastic differential equation. Indeed, Itô has explained this motivation himself, and we let him express it: "In these papers[7] I saw a powerful analytic method to study the transition probabilities of the process, namely Kolmogorov's parabolic equation and its extension by Feller. But I wanted to study the paths of Markov processes in the same way as Lévy observed differential processes. Observing the intuitive background in which Kolmogorov derived his equation (explained in the introduction of the paper), I noticed that a Markovian particle would perform a time homogeneous differential process for infinitesimal future at every instant, and arrived at the notion of a stochastic differential equation governing the paths of a Markov process that could be formulated in terms of the differentials of a single differential process" [37].[8]

Itô's first paper on stochastic integration was published in 1944 ([34]), the same year that Kakutani published two brief notes connecting Brownian motion and harmonic functions. Meanwhile throughout the 1940's Doob, who came to probability from complex analysis, saw the connection between J. Ville's martingales and harmonic functions, and he worked to develop a martingale based probabilistic potential theory. In addition, H. Cartan greatly advanced potential theory in the mid 1940's, later followed by Deny's classic work in 1950. All these ideas swirling around were interrelated, and in the 1940s Doob, clearly explained, for the first time, what should be the strong Markov property. A few years later (in 1948) E. Hille and K. Yosida independently gave the structure of semigroups of strongly continuous operators, clarifying the role of infinitesimal generators in Markov process theory.

In his efforts to model Markov processes, Itô constructed a stochastic differential equation of the form:

$$dX_t = \sigma(X_t)dW_t + \mu(X_t)dt,$$

where of course W represents a standard Wiener process. He now had two problems: one was to make sense of the stochastic differential $\sigma(X_t)dW_t$ which he accomplished in the aforementioned article [34].[9] The second problem was to connect Kolmogorov's work on Markov processes with his interpretation. In particular, he wanted to relate the paths of X to the transition function of the diffusion. This amounted to showing that the distribution of X solves Kolmogorov's forward equation. This effort resulted in his spectacular paper [35] in 1951, where he stated and proved what is now known as **Itô's formula**:

$$f(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)d[X,X]_t.$$

Here the function f is of course assumed to be $\mathcal{C}^2$, and we are using modern notation.[10] Itô's formula is of course an extension of the change of variables formula for

[7] Here Itô is referring to the papers of Kolmogorov [41] and of Feller [26].

[8] Note that while Itô never mentions the work of Bachelier in his foreword, citing instead Kolmogorov, Lévy, and Doob as his main influences, it is reasonable to think he was aware of the work of Bachelier, since it is referenced and explained in the key paper of Kolmogorov ([41]) that he lists as his one of his main inspirations. While we have found no direct evidence that Itô ever read Bachelier's work, nevertheless Hans Föllmer and Robert Merton have told the authors in private communications that Itô had indeed been influenced by the work of Bachelier. Merton has also published this observation: see page 47 of [51].

[9] Here Itô cites the work of S. Bernstein [5] as well as that of Kolmogorov [41] and W. Feller [26] as antecedents for his work.

[10] The book by H. P. McKean, Jr., published in 1969 [47], had a great influence in popularizing the Itô integral, as it was the first explanation of Itô's and others' related work in book form. But McKean referred to Itô's formula as *Itô's lemma*, a nomenclature that has persisted in some

Riemann-Stieltjes integration, and it reveals the difference between the Itô stochastic calculus and that of the classical path by path calculus available for continuous stochastic processes with paths of bounded variation on compact time sets. That formula is, of course, where A denotes such a process and f is $\mathcal{C}^1$:

$$df(A_t) = f'(A_t)dA_t.$$

It can be shown that if one wants to define a path by path integral of the form $\int_0^t H_s dA_s$ as the limit of sums, where H is any process with continuous sample paths, then as a consequence of the Banach Steinhaus theorem A *a fortiori* has sample paths of bounded variation on compacts. (See, for example, [67].) Since Brownian motion has paths of unbounded variation almost surely on *any* finite time interval, Itô knew that it was not possible to integrate all continuous stochastic processes. One of his key insights was to limit his space of integrands to those that were, as he called it, *non anticipating*. That is, he only allows integrands that are adapted to the underlying filtration of σ-algebras generated by the Brownian motion. This allowed him to make use of the independence of the increments of Brownian motion to establish the L^2 isometry

$$E\left(\left(\int_0^t H_s dW_s\right)^2\right) = E\left(\int_0^t H_s^2 ds\right).$$

Once the isometry is established for continuous non-anticipating processes H, it then extends to jointly measurable non-anticipating processes.[11]

J. L. Doob realized that Itô's construction of his stochastic integral for Brownian motion did not use the full strength of the independence of the increments of Brownian motion. In his highly influential 1953 book [16] he extended Itô's stochastic integral for Brownian motion first to processes with orthogonal increments (in the L^2 sense), and then to processes with conditionally orthogonal increments, that is, martingales. What he needed, however, was a martingale M such that $M_t^2 - F(t)$ is again a martingale, where the increasing process F is *non-random*. He established the now famous Doob decomposition theorem for submartingales: *If X_n is a (discrete time) submartingale, then there exists a unique decomposition $X_n = M_n + A_n$ where M is a martingale, and A is a process with non-decreasing paths, $A_0 = 0$, and with the special measurability property that A_n is $\mathcal{F}_{n-1}$ measurable.* Since M^2 is a submartingale when M is a martingale, he needed an analogous decomposition theorem in continuous time in order to extend further his stochastic integral. As it was, however, he extended Itô's isometry relation as follows:

$$E\left(\left(\int_0^t H_s dM_s\right)^2\right) = E\left(\int_0^t H_s^2 dF(s)\right),$$

where F is non-decreasing and non-random, $M^2 - F$ is again a martingale, and also the stochastic integral is also a martingale. (See Chapter IX of [16].)

circles to this day. Obviously this key theorem of Itô is much more important than the status the lowly nomenclature "lemma" affords it, and we prefer Itô's own description: "formula".

[11] Indeed, this is how the theory is presented in the little 1969 book of McKean [47]. Unfortunately it is not as simple as McKean thought at this early stage of the theory, to determine exactly which processes are included in this procedure; the natural σ-algebra generated by the simple integrands is today known as the *predictable* σ-algebra, and the predictably measurable processes are a strict subset of jointly measurable, non-anticipating processes. This point is clarified in (for example) the book of K. L. Chung and R. Williams [9], p. 63.

Thus it became an interesting question, if only for the purpose of extending the stochastic integral to martingales in general, to see if one could extend Doob's decomposition theorem to submartingales indexed by continuous time. However there were other reasons as well, such as the development of probabilistic potential theory, which began to parallel the development of axiomatic potential theory, especially with the publication of G. A. Hunt's seminal papers in 1957 and 1958 [31, 32, 33]. It took perhaps a decade for these papers to be fully appreciated, but in the late 1960's and early 1970's they led to even greater interest in Itô's treatment of Markov processes as solutions of stochastic differential equations, involving both Brownian motion and what is today known as Poisson random measure.

The issue was resolved in two papers by the (then) young French mathematician P. A. Meyer in 1962. Indeed, as if to underline the importance of probabilistic potential theory in the development of the stochastic integral, Meyer's first paper, establishing the existence of the Doob decomposition for continuous time submartingales [52], is written in the language of potential theory. Meyer showed that the theorem is false in general, but true if and only if one assumes that the submartingale has a uniform integrability property when indexed by stopping times, which he called "Class (D)", clearly in honor of Doob. Ornstein had shown that there were submartingales not satisfying the Class (D) property[12], and G. Johnson and L. L. Helms [40] quickly provided an example in print in 1963, using three dimensional Brownian motion. Also in 1963, P. A. Meyer established the uniqueness of the Doob decomposition [53], which today is known as the Doob-Meyer decomposition theorem. In addition, in this second paper Meyer provides an analysis of the structure of L^2 martingales, which later will prove essential to the full development of the theory of stochastic integration. Two years later, in 1965, Itô and S. Watanabe, while studying multiplicative functionals of Markov processes, define *local martingales* [39]. This turns out to be the key object needed for Doob's original conjecture to hold. That is, any submartingale X, whether it is of Class (D) or not, has a unique decomposition

$$X_t = M_t + A_t,$$

where M is a local martingale, and A is a non-decreasing, predictable process with $A_0 = 0$.

Returning however to P. A. Meyer's original paper [52], at the end of the paper, as an application of his decomposition theorem, he proposes an extension of Doob's stochastic integral, and thus *a fortiori* an extension of Itô's integral. His space of integrands is that of "well adapted" processes, meaning jointly measurable and adapted to the underlying filtration of σ-algebras. He makes the prescient remark at the end of his paper that "it seems hard to show (though it is certainly true) that the full class of well adapted processes whose "norm" is finite has been attained by this procedure." This anticipates the oversight of McKean six years later (see footnote 11), and it is this somewhat esoteric measurability issue that delays the full development of stochastic integration for martingales which have jumps, as we shall see.

Before we continue our discussion of the evolution of the theory of stochastic integration, however, let us digress to discuss the developments in economics. It is curious that Peter Bernstein, in his 1992 book [4], states "Despite its importance, Bachelier's thesis was lost until it was rediscovered quite by accident in the 1950's by Jimmie Savage, a mathematical statistician at Chicago." He goes on a little later to say "Some time around 1954, while rummaging through a university library, Savage

[12] See, for example, [59], p. 823

chanced upon a small book by Bachelier, published in 1914, on speculation and investment." We know however that Kolmogorov and also Doob explicitly reference Bachelier, and Itô certainly knew of his work too; but perhaps what was "lost" was Bachelier's contributions to economics.[13] Bernstein relates that Savage alerted the economist Paul Samuelson to Bachelier's work, who found Bachelier's thesis in the MIT library, and later remarked "Bachelier seems to have had something of a one-track mind. But what a track!" [73]. See also [74].

After a decade of lectures around the country on warrant pricing and how stock prices must be random,[14] Samuelson then went on to publish, in 1965, two papers of ground breaking work. In his paper [72] he gives his economics arguments that prices must fluctuate randomly, 65 years after Bachelier had assumed it! This paper, along with Fama's [24] work on the same topic, form the basis of what has come to be known as "the efficient market hypothesis." The efficient market hypothesis caused a revolution in empirical finance; the debate and empirical investigation of this hypothesis is still continuing today (see [25]). Two other profound insights can be found in this early paper that subsequently, but only in a modified form, became the mainstay of option pricing theory. The first idea is the belief (postulate) that discounted futures prices follow a martingale[15]. From this postulate, Samuelson proved that changes in futures prices were uncorrelated across time, a generalization of the random walk model (see [46], and also [13]). The second insight is that this proposition can be extended to arbitrary functions of the spot price, and although he did not state it explicitly herein, this forebodes an immediate application to options.

In his companion paper [71], he combined forces with H.P. McKean Jr.[16] (who the same year published his tome together with K. Itô [38]) who wrote a mathematical appendix to the paper, to show essentially that a good model for stock price movements is what is today known as geometric Brownian motion. Samuelson explains that Bachelier's model failed to ensure that stock prices always be positive, and that his model leads to absurd inconsistencies with economic principles, whereas geometric Brownian motion avoids these pitfalls. This paper also derived valuation formulas for both European and American options.[17] The derivation was almost identical to that used nearly a decade later to derive the Black-Scholes formula, except that instead of invoking a no arbitrage principle to derive the valuation formula, he again postulated the condition that the discounted options payoffs follow a martingale (see [71] p. 19), from which the valuation formulae easily followed.

[13] It is possible that L. J. Savage read Bachelier's work because Doob's book had appeared only one year earlier and had referenced it, and then he might have been surprised by the economics content of Bachelier's work. But this is pure speculation. Also, Samuelson wrote in [73] (p. 6) that "this was largely lost in the literature, even though Bachelier does receive occasional citation in standard works in probability."

[14] These lectures lead to other papers being published by researchers following up on Samuelson's ideas, for example the renowned paper of Osborne [62].

[15] See the Theorem of Mean Percentage Price Drift on page 46 and the subsequent discussion.

[16] Samuelson combined forces with McKean, and later R. C. Merton, because he did not feel comfortable with the newly developed stochastic calculus (see [4] p. 215). This insight was also confirmed by private communications with R. C. Merton.

[17] This is the paper that first coined the terms "European" and "American" options. According to a private communication with R.C. Merton, prior to writing the paper, P. Samuelson went to Wall Street to discuss options with industry professionals. His Wall Street contact explained that there were two types of options available, one more complex - that could be exercised any time prior to maturiy, and one more simple - that could be exercised only at the maturity date, and that only the more sophisticated European mind (as opposed to the American mind) could understand the former. In response, when Samuelson wrote the paper, he used these as prefixes and reversed the ordering.

The much later insights of Black, Scholes, and Merton, relating prices of options to perfect hedging strategies, is of course not discussed in this article. Furthermore, it is also noteworthy that within this paper, Samuelson and McKean determine the price of an American option by discovering the relation of an American option to a free boundary problem for the heat equation. This is the first time that this connection is made. Interestingly, Samuelson and McKean do not avail themselves of the tools of stochastic calculus, at least not explicitly. The techniques McKean uses in his appendix are partial differential equations in the spirit of Kolmogorov, coupled with stopping times and the potential theoretic techniques pioneered by G. Hunt and developed by Dynkin.

The final precursor to the Black, Scholes and Merton option pricing formulaes can be found in the paper of Samuelson and Merton [75]. Following similar mathematics to [71], instead of invoking the postulate that discounted option payoffs follow a martingale, they derived this postulate as an implication of a utility maximizing investor's optimization decision. Herein, they showed that the option's price could be viewed as its discounted expected value, where instead of using the actual probabilities to compute the expectation, one uses utility or risk adjusted probabilities[18]. These risk adjusted probabilities later became known as "risk-neutral" or "equivalent martingale" probabilities. It is interesting to note that, contrary to common belief, this use of "equivalent martingale probabilities" under another guise predated the paper by Cox and Ross [12] by nearly 10 years. In fact, Merton (footnote 5 page 218, [50]) points out that Samuelson knew this fact as early as 1953! Again, by not invoking the no arbitrage principle, this paper just missed obtaining the famous Black Scholes formula. The first use of the no arbitrage principle to prove a pricing relation between various financial securities can be found in Modigliani and Miller [60] some eleven years earlier, where they showed the equivalence between two different firms' debt and equity prices, generating the famous M&M Theorem. Both Samuelson and Merton were aware of this principle, Modigliani being a colleague at M.I.T., but neither thought to apply it to this pricing problem until many years later.

Unrelated to finance, and almost as an aside in the general tide of the development of the theory of stochastic integration, were the insights of Herman Rubin. At the Third Berkeley Symposium in 1955, Rubin gave a talk on stochastic differential equations. The following year, he presented an invited paper at the Seattle joint meetings of the Institute of Mathematical Statistics, the American Mathematical Society , the Biometric Society, the Mathematical Association of America, and the Econometrics Society. In this paper he outlined what was later to become D. L. Fisk's Ph.D. thesis, which invented both quasimartingales and what is now known as the Stratonovich integral. To quote his own recollections, "I was unhappy with the Itô integral because of the lack of invariance with nonlinear change of coordinate systems, no matter how smooth, and, observing that using the average of the right and left endpoints gave exactly the right results for the integral of XdX for any X (even discontinuous), it seemed that this was, for continuous X with sufficiently good properties, the appropriate candidate for the integral...Quasimartingales seemed the natural candidate for the class of processes, but I did not see a clear proof. I gave the problem to Fisk to work on for a Ph.D. thesis, and he did come up with what was needed" [69].

Indeed, in D. L. Fisk's thesis [27], written under Rubin when he was at Michigan State University, Fisk developed what is now known as the Stratonovich integral,

[18] See especially expression (20) on page 26.

and he also coined the phrase and developed the modern theory of quasimartingales, later used by K. M. Rao [68] to give an elegant proof that a quasimartingale is the difference of two submartingales, and also used by S. Orey [63] in a paper extending the idea and which foreshadowed modern day semimartingales. Fisk submitted his thesis for publication, but the editor did not believe there was much interest in stochastic integration, again according to the recollections of Herman Rubin [69]. So Fisk dropped that part of the thesis and did not pursue it, publishing instead only the part on quasimartingales, which appeared as [28].

Returning now to the historical development of stochastic integration, we mention that P. A. Meyer's development of the stochastic integral in [52] is skeletal at best, and a more systematic development is next put forward by Philippe Courrège in 1963 [10]. The motivation clearly arises from potential theory, and the paper of Courrège is published not in a journal, but in the (at the time) widely circulated *Séminaire Brélot-Choquet-Dény (Théorie du Potentiel)*. Many reasonable Markov processes, and in particular those treated by Hunt ([31, 32, 33]), have the property that they are *quasi-left continuous*. That is, they have paths which are right continuous with left limits a.s., and if there is a jump at a stopping time T, then that time T must be totally inaccessible. Intuitively, T must come as a complete surprise. One can formulate the condition of quasi-left continuity in terms of the underlying filtration of σ-algebras of the Markov process as well. This seems to be a reasonable property for the filtration of a time homogeneous Markov process to have, and is satisfied for a vast collection of examples.

It was natural for someone working in potential theory to make the assumption that the filtration is quasi-left continuous, and such an assumption has the fortuitous consequence to imply that if X is a submartingale and $X = M + A$ is its Doob-Meyer decomposition, then A has continuous sample paths. What this means is that in the L^2 isometry

$$E\left(\left(\int_0^t H_s\, dM_s\right)^2\right) = E\left(\int_0^t H_s^2\, dA_s\right),$$

where A is the increasing process corresponding to the submartingale $X = M^2$, one extends the Itô-Doob technique to general L^2 martingales, and the resultant increasing random process A has continuous paths. This, it turns out, greatly simplifies the theory. And it is precisely this assumption that Courrège makes. Courrège also works with integrands which have left continuous paths, and he considers the space of processes that are measurable with respect to the σ-algebra they generate, on $\mathbb{R} \times \Omega$, calling it processes which are "fortement bien adapté". Thus Courrège had, in effect, pioneered the *predictable σ-algebra*, although he did not use it as P. A. Meyer did, as we shall see. As it turns out, if dA_t is path by path absolutely continuous with respect to dt (this is usually written $dA_t << dt$), almost surely, then there ends up being essentially no difference which σ-algebra one uses: the predictable σ-algebra, or the progressive σ-algebra,[19] or even jointly measurable adapted processes. However if A is merely continuous and does not necessarily have absolutely continuous paths a.s., then one needs at least the progressive σ-algebra. We now know that what happens is that the difference between one such process and its predictable projection is a process that has a stochastic integral which is

[19]The *progressive σ-algebra* is defined later in the theory, and it has the property that if a process H_s is progressively measurable, and if τ is a finite valued stopping time, then H_τ is $\mathcal{F}_\tau$ measurable.

the zero process a.s, and this is why it does not matter. (For a detailed explanation see Liptser and Shiryaev [45], or alternatively Chung and Williams [9]).

One important thing that Courrège did not do, however, was to prove a change of variables formula, analogous to Itô's formula for stochastic integration with respect to Brownian motion. This was done in 1967 in an influential paper of H. Kunita and S. Watanabe [42]. Whereas the approach of Courrège was solidly in the tradition of Doob and Itô, that of establishing an L^2 isometry, the approach pioneered by M. Motoo and S. Watanabe two years later in 1965 was new: they treated the stochastic integral as an operator on martingales having specific properties, utilizing the Hilbert space structure of L^2 by using the Doob-Meyer increasing process to inspire an inner product through the quadratic variation of martingales. (See [61]). In the same paper Motoo and Watanabe established a martingale representation theorem which proved to be prescient of what was to come: they showed that all L^2 martingales defined on a probability space obtained via the construction of a type of Markov process named a Hunt process (in honor of the fundamental papers of G. Hunt mentioned earlier)were generated by a collection of additive functionals which were also L^2 martingales, and which were obtained in a way now associated with Dynkin's formula and "martingale problems."

The important paper of Motoo and Watanabe, however, was quickly overshadowed by the subsequent and beautifully written paper of H. Kunita and S. Watanabe, published in 1967 [42]. Here Kunita and Watanabe developed the ideas on orthogonality of martingales pioneered by P. A. Meyer, and Motoo and Watanabe, and they developed a theory of stable spaces of martingales which has proved fundamental to the theory of martingale representation, known in Finance as "market completeness." They also clarified the idea of quadratic variation as a pseudo inner product, and used it to prove a general change of variables formula, profoundly extending Itô's formula for Brownian martingales. The formula was clean and simple for martingales with continuous paths, but when it came to the general case (i.e., martingales that can have jump discontinuities in their sample paths)the authors retreated to the rich structure available to them in the Hunt process setting, and they expressed the jumps in terms of the Lévy system of the underlying Markov process. (Lévy systems for Markov processes, a structure which describes the jump behavior of a Hunt process, had only been developed a few years earlier in 1964 by S. Watanabe [85], and extended much later by A. Benveniste and J. Jacod [3]). This "retreat" must have seemed natural at the time, since stochastic integrals were, as noted previously, seen as intimately intertwined with Markov processes. And also, as an application of their change of variables formula, Kunita and Watanabe gave simple and elegant proofs of Lévy's theorem characterizing Brownian motion among continuous martingales via its quadratic variation process, as well as an extension from one to N dimensions of the spectacular 1965 theorem of L. Dubins and G. Schwarz [18] and K. E. Dambis [14] that a large class of continuous martingales can be represented as time changes of Brownian motion.

This remarkable paper of Kunita and Watanabe was quickly appreciated by P.A. Meyer, now in Strasbourg. He helped to start, with the aid of Springer-Verlag, the *Séminaire de Probabilités*, which is one of the longest running seminars to be published in Springer's famed *Lecture Notes in Mathematics* series. In the first issue, which is only Volume 39 in the *Lecture Notes* series, he published four key papers inspired by the article of Kunita and Watanabe [54, 55, 56, 57].[20] In

[20] A large number of the historically important works on stochastic integration were published in the Séminaire de Probabilités series, and these papers have been recently reprinted in a new

these papers he made two important innovations: he went beyond the "inner product" of Kunita and Watanabe (which is and was denoted $< X, Y >$, and which is tied to the Doob-Meyer decomposition), and expanding on an idea of Austin for discrete parameter martingales he created the "square bracket" (*le crochet droit*) pseudo inner product, denoted $[X, Y]$. Unlike the bracket process $< X, Y >$, which exists for all locally square integrable martingales (and therefore all continuous ones), the square bracket process exists for all martingales, and even all local martingales. This turned out to be important in later developments, such as the invention of semimartingales, and of course is key to the extension of the stochastic integral to all local martingales, and not only locally square integrable ones.

The second major insight of Meyer in these papers is his realization of the importance of the predictable σ-algebra. Going far beyond Courrège he realized that when a martingale also had paths of finite variation (of necessity a martingale with jumps), the stochastic integral should agree with a path by path construction using Lebesgue-Stieltjes integration. He showed that this holds if and only if the integrand is a predictable process. Moreover, he was able to analyze the jumps of the stochastic integral, observing that the stochastic integral has the same jump behavior as does the Lebesgue-Stieltjes integral if the integrand is predictably measurable. This laid the groundwork for the semimartingale theory that was to come a few years later.

We should further note at this point that Meyer was able to discard the Markov process framework used by Kunita and Watanabe in the first two of the four papers, and he established the general change of variables formula used today without using Lévy systems. Meyer then applied his more general results to Markov processes in the latter two of his four papers. Again, this was natural, since one of Meyer's primary interests was to resolve the many open questions raised by Hunt's seminal papers. It was research in Markov processes that was driving the interest in stochastic integration, from Itô on, up to this point. Nevertheless, Doob had begun to isolate the martingale character of processes independent of Markov processes, and Meyer's approach in his classic papers of 1962 and 1963 (already discussed [52] and [53]) was to use the techniques developed in Markov process potential theory to prove purely martingale theory results.

The development of stochastic integration as recounted so far seems to be primarily centered in Japan and France. But important parallel developments were occurring in the Soviet Union. The books of Dynkin on Markov processes appeared early, in 1960 [19] and in English as Springer Verlag books in 1965 [20]. The famed Moscow seminar (reconstituted at least once on October 18 and 19, 1996 in East Lansing, Michigan, with Dynkin, Skorohod, Wentzell, Freidlin, Krylov, etc.), and Girsanov's work on transformations of Brownian motion date to 1960 and earlier [29].[21] Stratonovich developed a version of the Itô integral which obeys the usual Riemann-Steiltjes change of variables formula, but sacrifices the martingale property as well as much of the generality of the Itô integral.[22] [80] While popular

volume of the Séminaire, with a small amount of commentary as well [23].

[21] Girsanov's work extends the much earlier work first of Cameron and Martin [8], who in 1949 transformed Brownian paths for both deterministic translations and also some random translations, keeping the old and new distributions of the processes equivalent (in the sense of having the same sets of probability zero); these ideas were extended to Markov processes first by Maruyama [48] in 1954, and then by Girsanov in 1960. It was not until 1974 that Van Schuppen and Wong [82] extended these ideas to martingales, followed in 1976 by P. A. Meyer [58] and in 1977 Lenglart [43] for the current modern versions. See also (for example) pages 132–136 of [67] for an exposition of the modern results.

[22] Indeed, the Stratonovich integral was not met with much excitement. In a book review of the

in some engineering circles, the Stratonovich integral seemed to be primarily a curiosity, until much later when it was shown that if one approximates the paths of Brownian motion with differentiable curves, the resultant integrals converge to the Stratonovich integral; this led to it being an intrinsic object in stochastic differential geometry (see, e.g., [22]).

The primary works of interest in the Soviet Union were the series of articles of Skorohod. Again mainly inspired by the developing theory of Markov processes, Skorohod generalized the Itô integral in ways startlingly parallel to those of Courrège and Kunita and Watanabe. In 1963 Skorohod, squarely in the framework of Markov processes and clearly inspired by the work of Dynkin, developed a stochastic integral for martingales which is analogous to what Courrège had done in France, although he used changes of time [76]. In 1966, while studying additive functionals of continuous Markov processes, he developed the idea of quadratic variation of martingales, as well as what is now known as the Kunita-Watanabe inequality, and the same change of variables formula that Kunita and Watanabe established [77]. He extended his results and his change of variables formula to martingales with jumps (always only those defined on Markov processes) in 1967 [79]. The jump terms in the change of variables formula are expressed with the aid of a kernel reminiscent of the Lévy systems of S. Watanabe.[23]

We close this short history with a return to France. After the paper of Kunita and Watanabe, and after P. A. Meyer's four papers extending their results, there was a hiatus of three years before the paper of C. Doléans-Dade and P. A. Meyer appeared [15]. Prior to this paper the development of stochastic integration had been tied rather intimately to Markov processes, and was perhaps seen as a tool with which one could more effectively address certain topics in Markov process theory. A key assumption made by the prior work of H. Kunita and S. Watanabe, and also of P. A. Meyer, was that the underlying filtration of σ algebras was *quasi left continuous*, alternatively stated as saying that the filtration had no fixed times of discontinuity. Doléans-Dade and Meyer were able to remove this hypothesis, thus making the theory a purely martingale theory, and casting aside its relation to Markov processes. This can now be seen as a key step that led to the explosive growth of the theory in the 1970's and also in finance to the fundamental papers of Harrison-Kreps and Harrison-Pliska, towards the end of the next decade. Last, in this same paper Doléans-Dade and Meyer coined the modern term *semimartingale*, to signify the most general process for which one knew (at that time) there existed a stochastic integral.[24]

---

time Skorohod wrote "The proposed integral, when it exists, may be expressed rather simply using the Itô integral. However the class of functions for which this integral exists is extremely narrow and artificial. Although some of the formulas are made more simple by using the symmetrized integral (while most of them are made more complicated which will be made clear from what follows), its use is extremely restricted by its domain of definition. Thus this innovation is completely unjustified." [78] The Stratonovich integral was developed simultaneously by D. Fisk in the United States, as part of his PhD thesis. However it was rejected for publication as being too trivial. In the second half of his thesis he invents quasimartingales, and that half was indeed published [28].

[23] P. A. Meyer's work ([54, 55, 56, 57]), which proved to be highly influential in the West, references Courrège, Motoo and Watanabe, Watanabe, and Kunita and Watanabe, but not Skorohod, of whose work Meyer was doubtless unaware. Unfortunately this effectively left Skorohod's work relatively unknown in the West for quite some time.

[24] As we will see in a sequel to this paper, the description of semimartingales of Doléans-Dade and Meyer of 1970 turned out to be prescient. In the late 1970's C. Dellacherie and K. Bichteler simultaneously proved a characterization of semimartingales: they showed that given a right continuous process X with left limits, if one defined a stochastic integral in the obvious way on simple predictable processes, and if one insisted on having an extremely weak version of a bounded convergence theorem, then X was *a fortiori* a semimartingale.

Acknowledgements

The authors are grateful for help with this short history given by H. Föllmer, K. Itô,[25] J. Jacod, J. Pitman, H. Rubin, A. N. Shiryaev, S. Watanabe, M. Yor, and M. Zakai.

References

[1] Bachelier, L. (1900). *Théorie de la Spéculation*, Annales Scientifiques de l'École Normale Supérieure, 21–86.

[2] Bachelier, L. (1900). *Théorie de la Spéculation*, Gauthier-Villars, Paris. {Note: This book has been reprinted by the Paris publisher Éditions Jacques Gabay (1995).} MR1397712

[3] Benveniste, A. and Jacod, J. (1973). *Systèmes de Lévy des processus de Markov*, Invent. Math., **21**, 183–198. MR343375

[4] Bernstein, P. L. (1992). *Capital Ideas: The Improbable Origins of Modern Wall Street*, The Free Press, New York.

[5] Bernstein, S. (1938). *Equations différentielles stochastiques*, Actualités Sci. Ind., **738**, 5–31.

[6] Bru, B. and Yor, M. (2001). *La vie de W. Doeblin et le Pli cacheté 11 668*, La Lettre de L'Académie des Sciences, **2**, 16–17.

[7] Bru, B. and Yor, M. (2002). *Comments on the life and mathematical legacy of Wolfgang Doeblin*, Finance and Stochastics **6**, 3–47. MR1885582

[8] Cameron, R. H. and Martin, W. T. (1949). *Transformation of Wiener integrals by non-linear transformations*, Transactions of the American Math. Society **66**, 253–283. MR31196

[9] Chung, K. L. and Williams, R. (1990). *Introduction to Stochastic Integration, Second Edition*, Birkhäuser, Boston. MR1102676

[10] Courrège, Ph. (1963). *Intégrales stochastiques et martingales de carré intégrable*, Séminaire Brélot-Choquet-Dény (Théorie du Potentiel), **7e année, 1962/63**, 7-01–7-20.

[11] Courtault, J.-M., Kabanov, Y., Bru, B., Crépel, P., Lebon, I., and Le Marchand, A. (2000). *Louis Bachelier: On the Centenary of Théorie de la Spéculation*, Mathematical Finance **10**, 341–353. MR1800320

[12] Cox, J. and Ross, S. A. (1976). *The Valuation of Options for Alternative Stochastic Processes*, Journal of Financial Economics, **3** (1/2), 145–166.

[13] Csörgo. (1998). *Random walking around financial mathematics*, Random walks (Budapest), edited by Pál Révész, Bálint Tóth. Bolyai Soc. Math. Stud., **9**, 59–111. MR1752891

[14] Dambis, K. E. (1965). *On the decomposition of continuous martingales*, Theor. Proba. Applications, **10**, 401–410. MR202179

[25] The second author wishes to express his thanks to Junko Itô who facilitated communication with K. Itô in the preparation of [66], and this information was also used in this article.

[15] Doléans-Dade, C. and Meyer, P. A. (1970). *Intégrales stochastiques par rapport aux martingales locales*, Séminaire de Probabilités IV, Lecture Notes in Mathematics, **124** 77–107. MR270425

[16] Doob, J. L. (1953). *Stochastic Processes*, John Wiley and Sons, New York. MR58896

[17] Doob, J. L. (1996). *The Development of Rigor in Mathematical Probability (1900–1950)*, in J.-P. Pier, ed., *Development of Mathematics 1900–1950*, Birkhauser Verlag AG, Basel. MR1404084

[18] Dubins, L. and Schwarz, G. (1965). *On continuous martingales*, Proc. National Acad. Sciences USA, **53**, 913–916. MR178499

[19] Dynkin, E. (1960). *Theory of Markov Processes*, Pergamon Press, Oxford. MR193669

[20] Dynkin, E. (1965). *Markov Processes* (two volumes) Springer-Verlag, Berlin, 1965. MR193671

[21] Einstein, A. (1905). *On the movement of small particles suspended in stationary liquid demanded by the molecular-kinetic theory of heat*, Ann. d. Physik **17** {In *Investigations of the theory of Brownian movement*, ed. R. Fürth, Dover, New York, 1956}.

[22] Emery, M. (1989). *Stochastic calculus in manifolds, with an appendix by P. A. Meyer*, Springer-Verlag, Berlin. MR1030543

[23] Emery, M. and Yor, M., eds. (2002). *Séminaire de Probabilités 1967–1980: A Selection in Martingale Theory*, Lecture Notes in Mathematics, **1771**. MR1925827

[24] Fama, E. (1965). *The Behavior of Stock Prices*, Journal of Business, **38** 34–105.

[25] Fama, E. (1998). *Market Efficiency, Long Term Returns, and Behavioral Finance,* Journal of Financial Economics, **49**, 283–306.

[26] Feller, W. (1936). *Zur Theorie der Stochastichen Prozesse (existenz-und Eindeutigkeitssatze)*, Math. Ann. **113**.

[27] Fisk, D. (1963). *Quasi-martingales and stochastic integrals*, Ph.D. thesis, Michigan State University, Department of Statistics.

[28] Fisk, D. (1965). *Quasimartingales*, Transactions of the American Math. Soc., **120**, 369–389. MR192542

[29] Girsanov, I. V. (1960). *On transforming a certain class of stochastic processes by absolutely continuous changes of measures*, Theory Proba. Appl., **5**, 285–301. MR133152

[30] Hald, A. (1981). *T. N. Thiele's contributions to Statistics*, International Statistic Review **49**, 1–20. MR623007

[31] Hunt, G. A. (1957) *Markoff processes and potentials I*, Illinois J. Math. **1**, 44–93. MR91349

[32] Hunt, G. A. (1957). *Markoff processes and potentials II*, *Illinois J. Math.* **1**, 316–369. MR91349

[33] Hunt, G. A. (1958). *Markoff processes and potentials III*, *Illinois J. Math.* **2**, 151–213. MR107097

[34] Itô, K. (1944). *Stochastic Integral*, *Proc. Imp. Acad. Tokyo* **20**, 519–524. MR14633

[35] Itô, K. (1951). *On a formula concerning stochastic differentials*, *Nagoya Math. J.* **3**, 55–65. MR44063

[36] Itô, K. (1951). *Multiple Wiener integral*, *J. Math. Society of Japan* **3**, 157–169. MR44064

[37] Itô, K. (1987). *Foreword*, *K. Itô Collected Papers*, Springer-Verlag, Heidelberg, xiii–xvii. MR931775

[38] Itô, K. and McKean, H. P., Jr. (1965). *Diffusion Processes and Their Sample Paths*, Springer-Verlag, New York; new edition by Springer-Verlag, 1996.

[39] Itô, K. and Watanabe, S. (1965). *Transformation of Markov processes by multiplicative functionals*, *J. Math. Kyoto Univ.* **4**, 1–75. MR184282

[40] Johnson, G. and Helms, L. L. (1963). *Class (D) Supermartingales*, *Bull. American Math. Society* **69**, 59–62. MR142148

[41] Kolmogorov, A. N. (1931). *On Analytic Methods in Probability Theory*, in A. N. Shiryaev, ed., *Selected Works of A. N. Kolmogorov; Volume II: Probability Theory and Mathematical Statistics*, Kluwer, Dordrecht, 1992, 62–108. [Original: Uber die analytischen Methoden in der Wahrscheinlichkeitsrechnung, *Math. Ann.* **104**, 1931, 415–458.] MR1153022

[42] Kunita, H. and Watanabe, S. (1967). *On Square Integrable Martingales*, *Nagoya Math. J.* **30**, 209–245. MR217856

[43] Lenglart, E. (1977). *Transformation des martingales locales par changement absolument continu de probabilités*, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **39**, 65–70. MR448541

[44] Lévy, P. (1955). *W. Döblin (V. Doeblin) (1915-1940)* *Revue d'Histoire des Sciences et de leurs Applications* **8**, 107–115. MR72808

[45] Liptser, R. Sh., and Shiryaev, A. S.; A. B. Aries, translator, 1977, 1978, (2nd, revised and expanded edition, 2001) *Statistics of Random Processes*, Two volumes, Springer-Verlag, Heidelberg.

[46] Malkiel, B. G. (2003). *A Random Walk Down Wall Street*, 7th edition, WW Norton, New York.

[47] McKean, H. P., Jr. (1969). *Stochastic Integrals*, Academic Press, New York. MR247684

[48] Maruyama, G. (1954). *On the transition probability functions of Markov processes*, *Nat. Sci. Rep. Ochanomizu Univ.* **5**, 10–20. MR67400

[49] Maruyama, G. (1955). *Continuous time processes and stochastic equations*, *Rend. Circ. Math. Palermo* **4**, 1–43. MR71666

[50] Merton, R. C. (1990). *Continuous Time Finance*, Basil Blackwell, Cambridge, Massachusetts.

[51] Merton, R. C. (2002). *Future possibilities in Finance Theory and Finance practice, Mathematical Finance – Bachelier Congress 2000*, eds. Geman, H., Madan, D., Pliska, S. R., and T. Vorst; Springer-Verlag, Heidelberg, 47–73.

[52] Meyer, P. A. (1962). *A decomposition theorem for supermartingales*, *Ill. J. Math.* **6**, 193–205. MR159359

[53] Meyer, P. A. (1963). *Decomposition of supermartingales: the uniqueness theorem*, *Ill. J. Math.* **7**, 1–17. MR144382

[54] Meyer, P. A. (1967). *Intégrales Stochastiques I*, Séminaire de Probabilités I, *Lecture Notes in Mathematics*, **39**, 72–94. MR231445

[55] Meyer, P. A. (1967). *Intégrales Stochastiques II*, Séminaire de Probabilités I, *Lecture Notes in Mathematics*, **39**, 95–117. MR231445

[56] Meyer, P. A. (1967). *Intégrales Stochastiques III*, Séminaire de Probabilités I, *Lecture Notes in Mathematics*, **39**, 118–141. MR231445

[57] Meyer, P. A. (1967). *Intégrales Stochastiques IV*, Séminaire de Probabilités I, *Lecture Notes in Mathematics*, **39**, 142–162. MR231445

[58] Meyer, P. A. (1976). *Un cours sur les intégrales stochastiques*, Séminaire de Probabilités X, *Lecture Notes in Mathematics*, **511**, 246–400. MR501332

[59] Meyer, P. A. (2000). *Les Processus Stochastiques de 1950 à Nos Jours*, in *Development of Mathematics 1950-2000*, edited by Jean-Paul Pier; Birkhäuser, Boston, MA. 813–848. MR1796860

[60] Modigliani, F. and Miller, M. H. *The Cost of Capital, Corporation Finance, and the Theory of Investment*, *American Economic Review*, **48**, 261–297.

[61] Motoo, M. and Watanabe, S. (1965). *On a class of aditive functionals of Markov process*, *J. Math. Kyoto Univ.* **4**, 429–469. MR196808

[62] Osborne, M. F. M. (1959). *Brownian motion in the stock market*, Operations Research, **7**, 145–173. MR104513

[63] Orey, S. (1965). *F-processes*, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, **2** 301–313, University of California Press, Berkeley. MR214124

[64] Pais, A. (1982). *'Subtle is the Lord. . . ' The Science and Life of Albert Einstein*, Oxford University Press, Oxford. MR690419

[65] Petit, M. (2003). *L'équation de Kolmogoroff*, Editions Ramsay, Paris.

[66] Protter, P. (2003). *A new prize in honor of Kiyosi Itô*, *Stochastic Processes and their Applications*, **108**, 151–153. MR2016970

[67] Protter, P. (2004). *Stochastic Integration and Differential Equations; Second Edition*, Springer Verlag, Heidelberg. MR2020294

[68] Rao, K. M. (1969). *Quasimartingales*, *Math. Scand.*, **24**, 79–92. MR275511

[69] Rubin, H. (2003). Personal communication by electronic mail.

[70] Rubin, H. (1956). *Quasi-martingales and stochastic integrals*, title of an invited talk at the Seattle Meeting of the IMS, August 21–24, 1956; see page 1206 of the *Annals Math. Statist.*, **27**, 1198–1211.

[71] Samuelson, P. (1965). *Rational Theory of Warrant Pricing*, *Industrial Management Review*, **6**, 13–39.

[72] Samuelson, P. (1965). *Proof That Properly Anticipated Prices Fluctuate Randomly*, *Industrial Management Review*, **6**, 41–49.

[73] Samuelson, P. (1973). *Mathematics of Speculative Price*, *SIAM Review*, **15**, 1–42. MR323315

[74] Samuelson, P. (2002). *Modern finance theory within one lifetime*, *Mathematical Finance – Bachelier Congress 2000*, eds. Geman, H., Madan, D., Pliska, S. R., and T. Vorst; Springer-Verlag, Heidelberg, 41–46. MR1960557

[75] Samuelson, P. and Merton, R. C. (1969). *A Complete Model of Warrant Pricing that Maximizes Utility, Industrial Management Review*, **10**(2), 17–46.

[76] Skorokhod, A. V. (1963). *On homogeneous continuous Markov proceses that are martingales*, *Theory of Probability and its Applications*, **8**, 355–365. MR158432

[77] Skorokhod, A. V. (1966). *On the local structure of continuous Markov processes*, *Theory of Probability and its Applications*, **11**, 336–372. MR203815

[78] Skorokhod, A. V. (1967). Review of R. L. Stratonovich, *Conditional Markov Processes and Their Application to the Theory of Optimal Control*, *Theory of Probability and its Applications*, **12**, 154–156.

[79] Skorokhod, A. V. (1967). *Homogeneous Markov processes without discontinuities of the second kind*, *Theory of Probability and its Applications*, **12**, 222–240. MR230372

[80] Stratonovich, R. L. (1966). *Conditional Markov Processes and Their Application to the Theory of Optimal Control*, Izd. Moscow University Press, Moscow. MR197209

[81] Thiele, T. N. (1880). *Sur la compensation de quelques erreurs quasi-systématiques par la méthode des moindres carrés*, Reitzel, Copenhagen. {Note: This article was published simultaneously in Danish and French; for the Danish reference see [30].}

[82] Van Schuppen, J. H. and Wong, E. (1974). *Transformations of local martingales under a change of law*, *Annals of Probability* **2**, 879–888. MR358970

[83] Varadhan, S. R. S. and Stroock, D. W. (1987). *Introduction*, *K. Itô Collected Papers*, Springer-Verlag, Heidelberg, vii–xii. MR868862

[84] Ville, J. (1939). *Étude critique de la notion de collectif*, Gauthier-Villars, Paris.

[85] Watanabe, S. (1964). *On discontinuous additive functionals and Lévy measures of a Markov process*, *Japanese J. Math.*, **36**, 53–70. MR185675