The Work of Kyosi Itô

Philip Protter

The *Notices* solicited the following article describing the work of Kiyosi Itô, recipient of the 2006 Gauss Prize. The International Mathematical Union also issued a news release, which appeared in the November 2006 issue of the *Notices*.

Photo used with permission of Junko Itô.

Kyosi Itô, 1954, when he was a Fellow at the Institute for Advanced Study, Princeton.

On August 22, 2006, the International Mathematical Union awarded the Carl Friedrich Gauss Prize at the opening ceremonies of the International Congress of Mathematicians in Madrid, Spain. The prizewinner is Kyosi Itô. The Gauss prize was created to honor mathematicians whose research has had a profound impact not just on mathematics itself but also on other disciplines.

To understand the achievements of Itô, it is helpful to understand the context in which they were developed. Bachelier in 1900, and Einstein in 1905, proposed mathematical models for the phenomenon known as Brownian motion. These models represent the random motion of a very small particle in a liquid suspension. Norbert Wiener and collaborators showed in the 1920s that Einstein's model exists as a stochastic process, using the then-new ideas of Lebesgue measure theory. Many properties of the process were established in the 1930s, the most germane for this article being that its sample paths are of infinite variation on any compact time interval, no matter how small. This made the Riemann-Stieltjes integration theory inapplicable. Wiener wanted to use such integrals to study filtering theory and signal detection, important during the second world war. Despite these problems he developed a theory of integrals, known today as Wiener integrals, where the integrands are non-random functions. This served his purpose but was unsatisfying because it ruled out the study of stochastic differential equations, among other things.

The problem in essence is the following: how can one define a stochastic integral of the form $\int_0^t H_s dW_s$, where H has continuous sample paths and W is a Wiener process (another name for Brownian motion), as the limit of Riemann-style sums? That is, to define an integral as the limit of sums such as $\sum_{1\le i\le n} H_{\xi_i}(W_{t_{i+1}} - W_{t_i})$, with convergence for all such H. Unfortunately as a consequence of the Banach-Steinhaus theorem, W must then have sample paths of finite variation on compact time intervals. What Itô saw, and Wiener missed, was that if one restricts the class of potential integrands H to those that are adapted to the underlying filtration of sigma algebras generated by the Wiener process, and if one restricts the choice of $\xi_i \in [t_i, t_{i+1})$ to t_i, then one can use the independence of the increments of the Wiener process in a clever way to obtain the convergence of the sums to a limit. This became the stochastic integral of Itô. One should note that Itô did this in the mathematical isolation of Japan during the second world war and was one of the pioneers (along with G. Maruyama) of modern probability in Japan, which has since spawned some of the world's leading probabilists. Moreover since Jean Ville had named martingales as such only in 1939, and J. L. Doob had started developing his theory of martingales only in the 1940s, Itô was unaware of the spectacular developments in this area that were happening in the U.S., France, and the Soviet Union. Thus modern tools such as Doob's martingale inequalities were unavailable to Itô, and his

Philip Protter is professor of operations research at Cornell University. His email address is `pep4@cornell.edu`.

creativity in the proofs, looked at today, is impressive. But the key result related to the stochastic integral was Itô's change of variables formula.

Indeed, one can argue that most of applied mathematics traditionally comes down to changes of variable and Taylor-type expansions. The classical Riemann-Stieltjes change of variables, for a stochastic process A with continuous paths of finite variation on compacts, and $f \in C^1$ is of course

$$f(A_t) = f(A_0) + \int_0^t f'(A_s)dA_s.$$

With the Itô integral it is different and contains a "correction term". Indeed, for $f \in C^2$ Itô proved

$$f(W_t) = f(W_0) + \int_0^t f'(W_s)dW_s + \frac{1}{2}\int_0^t f''(W_s)ds.$$

This theorem has become ubiquitous in modern probability theory and is astonishingly useful. Moreover Itô used this formula to show the existence and uniqueness of solutions of stochastic ordinary differential equations:

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt; \quad X_0 = x_0,$$

when σ and b are Lipschitz continuous. This approach provided methods with an alternative intuition to the semigroup/partial differential equations approaches of Kolmogorov and Feller, for the study of continuous strong Markov processes, known as diffusions. These equations found applications without much delay: for example as approximations of complicated Markov chains arising in population and ecology models in biology (W. Feller), in electrical engineering where dW models white noise (N. Wiener, I. Gelfand, T. Kailath), in chemical reactions (e.g., L. Arnold), in quantum physics (P. A. Meyer, L. Accardi, etc.), in differential geometry (K. Elworthy, M. Emery), in mathematics (harmonic analysis (Doob), potential theory (G. Hunt, R. Getoor, P. A. Meyer), PDEs, complex analysis, etc.), and, more recently and famously, in mathematical finance (P. Samuelson, F. Black, R. Merton, and M. Scholes).

When Wiener was developing his Wiener integral, his idea was to study random noise, through sums of iterated integrals, creating what is now known as "Wiener chaos". However his papers on this were a mess, and the true architect of Wiener chaos was (of course) K. Itô, who also gave it the name "Wiener chaos". This has led to a key example of Fock spaces in physics, as well as in filtering theory, and more recently to a fruitful interpretation of the Malliavin derivative and its adjoint, the Skorohod integral.

Itô also turned his talents to understanding what are now known as Lévy processes, after the renowned French probabilist Paul Lévy. He was able to establish a decomposition of a Lévy process into a drift, a Wiener process, and an integral mixture of compensated compound Poisson processes, thus revealing the structure of such processes in a more profound way than does the Lévy-Khintchine formula.

In the late 1950s Itô collaborated with Feller's student H. P. McKean Jr. Together Itô and McKean published a complete description of one-dimensional diffusion processes in their classic tome, *Diffusion Processes and Their Sample Paths* (Springer-Verlag, 1965). This book was full of original research and permanently changed our understanding of Markov processes. It developed in detail such notions as local times and described essentially all of the different kinds of behavior the sample paths of diffusions could manifest. The importance of Markov processes for applications, and especially that of continuous Markov processes (diffusions), is hard to overestimate. Indeed, if one is studying random phenomena evolving through time, relating it to a Markov process is key to understanding it, proving properties of it, and making predictions about its future behavior.

Later in life, when conventional wisdom holds that mathematicians are no longer so spectacular, Itô embraced the semimartingale-based theory of stochastic integration, developed by H. Kunita, S. Watanabe, and principally P. A. Meyer and his school in France. This permitted him to integrate certain processes that were no longer adapted to the underlying filtration. Of course, this is a delicate business, due to the sword of Damocles Banach-Steinhaus theorem. In doing this, Itô began the theory of expansion of filtrations with a seminal paper and then left it to the work of Meyer's French school of the 1980s (Jeulin, Yor, etc.). The area became known as *grossissements de filtrations*, or in English as "the expansions of filtrations". This theory has recently undergone a revival, due to applications in finance to insider trading models, for example.

A much maligned version of the Itô integral is due to Stratonovich. While others were ridiculing this integral, Itô saw its potential for explaining parallel transport and for constructing Brownian motion on a sphere (which he did with D. Stroock), and his work helped to inspire the successful use of the integral in differential geometry, where it behaves nicely when one changes coordinate maps. These ideas have also found their way into other domains, for example in physics, in the analysis of diamagnetic inequalities involving Schrödinger operators (D. Hundertmark, B. Simon).

It is hard to imagine a mathematician whose work has touched so many different areas of applications, other than Isaac Newton and Gottfried Leibniz. The legacy of Kyosi Itô will live on a long, long time.

Construction of Brownian Motion

Mark H.A. Davis

22 November 2004

1 Definition

Brownian motion is a stochastic process $(W_t,\, t \geq 0)$ such that

- $W_0 = 0$.
- $(W_{t_2} - W_{t_1})$ and $(W_{t_4} - W_{t_3})$ are independent, for any $t_1 < t_2 \leq t_3 < t_4$.
- $(W_{t_2} - W_{t_1}) \sim N(0, t_2 - t_1)$.
- For almost all ω, the sample function $t \mapsto W_t(\omega)$ is continuous.

Since W_t has stationary independent increments, we already know that $EW_t = 0$ and that the covariance function is $r(t,s) = t \wedge s$. Thus the covariance matrix of the random vector $X = (W_{t_1}, \ldots, W_{t_n})$ is

$$Q = \begin{bmatrix} t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & \cdots & t_2 \\ \vdots & \vdots & \vdots & \vdots \\ t_1 & t_2 & \cdots & t_n \end{bmatrix}.$$

X therefore has characteristic function

$$\psi_X(u) = E\left[e^{iu^T X}\right] = e^{-\frac{1}{2}u^T Q u}, \tag{1}$$

corresponding to the normal distribution with mean zero and covariance matrix Q. One can check that (1) satisfies the Kolmogorov consistency conditions, so a process with finite-dimensional distributions given by (1) exists on some probability space. The point of what follows is to show that we can construct it in such a way that the sample functions are continuous, which does *not* follow from the consistency theorem.

2 Orthonormal bases in $L_2[0,1]$

Recall that any function $f \in L_2[0,1]$ can be expanded in a Fourier series

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin 2\pi nt + \sum_{i=1}^{\infty} b_i \cos 2\pi nt.$$

This is a special case of the more general result below. In fact, the trigonometric functions $\{1, \sqrt{2} sin 2\pi nt, \sqrt{2}\cos 2\pi nt, n = 1, \ldots\}$ provide an *orthonormal basis* of $H = L_2[0,1]$.

For $f, g \in H$ the *norm* and *inner product* are defined by

$$\begin{aligned} ||f|| &= \sqrt{\int_0^1 f^2(t)dt} \\ < f, g > &= \int_0^1 f(t)g(t)dt \end{aligned}$$

A countable set of functions $\Phi = \{\phi_1, \phi_2, \ldots\}$ is *orthonormal* if

$$\begin{aligned} ||\phi_i|| &= 1 \quad \text{for all } i \\ <\phi_i, \phi_j> &= 0, \quad i \neq j \end{aligned}$$

Let

$$H_n = \left\{ \sum_{i=1}^n \alpha_i \phi_i : \alpha \in R^n \right\}$$

be the linear subspace spanned by $(\phi_1, \ldots, \phi_n)$. The projection onto H_n of an arbitrary $f \in H$ is

$$\hat{f}_n = \sum_1^n < f, \phi_i > \phi_i.$$

Indeed, $<\hat{f}, \phi_j> = \sum_{i=1}^n < f, \phi_i >< \phi_i, \phi_j > = < f, \phi_j >$, so that $< f - \hat{f}_n, \phi_j > = 0$ for all j, so that[1] $(f - \hat{f}_n) \perp H_n$.

Φ is *complete* if, for any $f \in H$, $\hat{f}_n \to f$ as $n \to \infty$. This is the same thing as saying that if $f \perp \phi_i$ for all i then $f = 0$. In this case Φ is said to be a *complete orthonormal basis* (CONB). For a CONB we have the *Parseval equality*

$$||f||^2 = \sum_1^\infty < f, \phi_i^2 >, \tag{2}$$

and, as a corollary, the following identity

$$< f, g > = \sum_1^\infty < f, \phi_i >< g, \phi_i > \tag{3}$$

Indeed, (3) follows from (2) and the fact that

$$< f, g > = \frac{1}{4}\left(||f + g||^2 - ||f - g||^2 \right).$$

3 Construction of Brownian motion: the L_2 theory

Let $\{\phi_i\}$ be an arbitrary CONB of H and let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables defined on a probability space $(\Omega, \mathcal{F}, P)$, with $X_i \sim N(0, 1)$. For $n = 1, 2, \ldots$, define

$$W_t^n = \sum_{i=1}^n X_i \int_0^t \phi_i(s) ds. \tag{4}$$

Theorem 1 *For each t, W_t^n is a Cauchy sequence in $L_2(\Omega, \mathcal{F}, P)$ whose limit W_t is a normal random variable with mean zero and variance t. For any two times t, s, $E[W_t W_s] = t \wedge s$, where $t \wedge s = \min(t, s)$.*

Proof: Define

$$I_t(s) = \begin{cases} 1, & s < t \\ 0, & s \geq t. \end{cases}$$

Then

$$\int_0^t \phi_i(s) ds = < I_t, \phi_i > .$$

[1] $f \perp g$ means $< f, g > = 0$, and $f \perp H_n \iff f \perp g$ for all $g \in H_n$.

Since ϕ_i is a CONB,

$$I_t = \sum_i^{\infty} < I_t, \phi_i > \phi_i \quad \text{and} \quad t = ||I_t||^2 = \sum_1^{\infty} < I_t, \phi_i >^2 . \tag{5}$$

Thus for $n > m$

$$\begin{aligned} E\left(W_t^n - W_t^m\right)^2 &= E\left(\sum_{i=m+1}^{n} X_i \int_0^t \phi_i(s)ds\right)^2 \\ &= \sum_{i=m+1}^{n} < I_t, \phi_i >^2 \\ &\to 0 \quad \text{as } m, n \to \infty. \end{aligned}$$

Thus W_t^n is a Cauchy sequence in $L_2(\Omega, \mathcal{F}, P)$. Denoting the limit W_t we see from (5) that $\mathrm{var}(W_t) = \lim \mathrm{var}(W_t^n) = t$. It now follows from (3) that

$$\begin{aligned} E[W_t W_s] &= \sum_1^{\infty} < I_t, \phi_i >< I_s, \phi_i > \\ &= < I_t, I_s > \\ &= t \wedge s. \end{aligned}$$

It remains to show that W_t is normal. Note that W_t^n is a finite sum of normal random variables and is therefore normal, with variance $\sigma_n^2 = \sum_1^n < I_t, \phi_i >^2$. Hence the characteristic function of W_t^n is $\chi_n(u) = E[\exp(iuW_t^n)] = \exp(-\sigma_n^2 u^2/2)$, which converges as $n \to \infty$ to $\chi(u) \equiv \exp(-tu^2/2)$. Now $W_t^n \to W_t$ in L_2 implies that there is a sub-sequence $W_t^{n_k}$ such that $W_t^{n_k} \to W_t$ a.s. as $k \to \infty$. It follows from the bounded convergence theorem that $E[\exp(iuW_t^{n_k})] \to E[\exp(iuW_t)]$ and hence that $E[\exp(iuW_t)] = \chi(u)$. Thus $W_t \sim N(0, t)$. $\diamondsuit$

Theorem 1 shows that Brownian motion 'exists' in the sense that we have a gaussian process W_t with the right covariance function, but we have not shown a key property of Brownian motion, namely that it has continuous sample paths. To do this we need to introduce a special ONB, the *Haar functions.*

4 The Haar Functions

The Haar functions $\{f_0, f_{j,n}, j = 1, \ldots, 2^{n-1}, n = 1, 2, \ldots\}$ are defined by $f_0(t) \equiv 1$ and, with $k = 2j - 1$,

$$f_{j,n} = \begin{cases} 2^{(n-1)/2}, & \frac{k-1}{2^n} \le t \le \frac{k}{2^n} \\ -2^{(n-1)/2}, & \frac{k}{2^n} < t \le \frac{k+1}{2^n} \\ 0 \quad \text{elsewhere.} \end{cases}$$

Theorem 2 *The Haar functions are a complete orthonormal basis in $L_2[0, 1]$.*

It is easy to check that the Haar functions are orthonormal. We have to show that they are complete, i.e. that if for any $f \in H$ we have $< f, f_0 >= 0$ and $< f, f_{j,n} >= 0$ for all (j, n) then $f = 0$. Suppose f satisfies these conditions. Fix an integer n and define

$$J_k = \int_{k/2^n}^{(k+1)/2^n} f(t)dt, \quad k = 0, 1, \ldots, 2^n - 1.$$

Now $f \perp f_{0,n}$ implies $J_0 = J_1$, $f \perp f_{1,n}$ implies $J_2 = J_3$, and similarly $J_4 = J_5$ etc. Moving from n to $n-1$, we see that $f \perp f_{0,n-1}$ is equivalent to

$$(J_0 + J_1) - (J_2 + J_3) = 0,$$

implying that $J_0 = J_1 = J_2 = J_3$. Similarly, $J_4 = J_5 = J_6 = J_7$ etc. Now move to $n-2$ to show that

$$\begin{aligned} J_i &= J_0, \ i = 1, 2, \dots 7 \\ J_i &= J_8, \ i = 9, 10, \dots 15, \end{aligned}$$

and so forth. Continuing in this way we conclude that all the J_i are equal to J_0. But now

$$0 =< f, f_0 >= \sum_0^{2^n-1} J_i = 2^n J_0,$$

so that $J_0 = 0$. We have proved the following:

$$\int_a^b f(t)dt = 0 \quad \text{for all dyadic rational numbers } a, b. \tag{6}$$

(A dyadic rational number is one of the form $j/2^n$ for some j, n.) Since for any real number a there is a sequence a_n of dyadic rational numbers a_n converging to a, (6) implies that

$$\int_a^b f(t)dt = 0 \quad \text{for all real numbers } a, b. \tag{7}$$

A 'monotone class' argument as in Problems II now shows that

$$\int_A f(t)dt = 0 \qquad \text{for all Borel sets } B \in \mathcal{B}. \tag{8}$$

However, any integrable function f satisfying (8) is equal to zero almost surely. This completes the proof. $\diamondsuit$

5 The Lévy-Ciesielski Construction

This argument is described in McKean [1]. It is based on the following lemma from real analysis.

Lemma 1 *Suppose that, for $n = 1, 2, \dots$, $f_n : [0,1] \to R$ is a continuous function, and that f_n converges uniformly to a function f, i.e. given $\epsilon > 0$ there is a number N such that $n \geq N$ implies $|f_n(t) - f(t)| < \epsilon$ for any $t \in [0,1]$. Then f is a continuous function.*

Proof: For any $t_0, s \in [0,1]$ we can write

$$|f(t_0) - f(s)| \leq |f(t_0) - f_n(t_0)| + |f_n(t_0) - f_n(s)| + |f_n(s) - f(s)|.$$

Given $\epsilon > 0$ we can find n such that the first and third terms on the right are each less than $\epsilon/3$ (whatever t_0, s). Now f_n is continuous, so for fixed t_0 we can choose δ so that then second term is less than $\epsilon/3$ for all s such that $|t_0 - s| < \delta$. Consequently, f is continuous at t_0. $\diamondsuit$

The idea now is to show that $W_t^n \to W_t$ uniformly almost surely when we take the ONB ϕ_i to be the Haar functions.

The *Schauder functions* $\{F_0, F_{n,j}\}$ are the indefinite integrals of the Haar functions, i.e.

$$F_0(t) = t$$

and

$$F_{n,j}(t) = \begin{cases} 2^{(n-1)/2}(t - (k-1)/2^n), & t \in [(k-1)/2^n, k/2^n[\\ 2^{(n-1)/2}((k+1)/2^n - t), & t \in [k/2^n, (k+1)/2^n[\\ 0 \quad \text{elsewhere.} \end{cases}$$

As McKean puts it, the Schauder functions are "little tents" of height $2^{-(n+1)/2}$, as shown in figure 1.

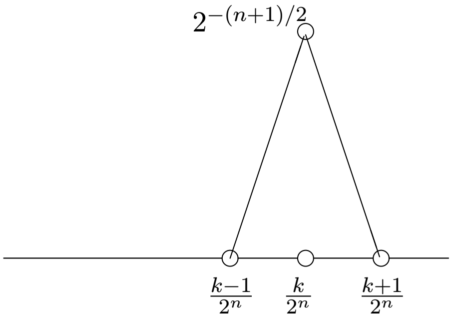

Figure 1: Schauder function

Now let $\{X_0, X_{n,j}, n = 1, 2, \ldots, j = 1, 2, \ldots, 2^{n-1}\}$ be independent $N(0,1)$ random variables defined on some probability space $(\Omega, \mathcal{F}, P)$, and for $t \in [0,1], N = 1, 2, \ldots$ define

$$W_t^N(\omega) = X_0 F_0(t) + \sum_{n=1}^{N} Y_n(t, \omega), \tag{9}$$

where

$$Y_n(t, \omega) = \sum_{j=1}^{2^{n-1}} X_{n,j}(\omega) F_{n,j}(t).$$

For each N, ω the sample function $t \mapsto W^N(t, \omega)$ is a continuous function. We want to show that $W_t^N \to W_t$ a.s. as $N \to \infty$ for some continuous-path process W_t.

Theorem 3 *The sequence W_t^N defined by (9) converges* uniformly in t, almost surely. *Thus the process $W_t = \lim_{N\to\infty} W_t^N$ is a stochastic process with continuous sample paths.*

PROOF: The proof is an application of the Borel-Cantelli Lemma. Define $H_n(\omega) = \max_{t\in[0,1]} |Y_n(t, \omega)|$. Since for fixed n the Schauder functions $F_{n,1}, F_{n,2}, \ldots$ are non-zero on disjoint intervals we see that

$$H_n = 2^{-(n+1)/2} \max_{1\le j\le 2^{n-1}} |X_{n,j}|.$$

Thus for any constant c_n,

$$\begin{aligned} P\left[H_n > 2^{-(n+1)/2} c_n\right] &= P\left[\max_j |X_{n,j}| > c_n\right] \\ &= P \bigcup_j [|X_{n,j}| > c_n] \\ &\le \sum_j P[|X_{n,j}| > c_n] \\ &\le 2^{n-1} \times \frac{2}{c_n\sqrt{2\pi}} e^{-\frac{1}{2}c_n^2}. \end{aligned} \tag{10}$$

(The last inequality follows from Lemma 2 below.) We now ingeniously choose $c_n = \theta\sqrt{2n\log 2}$ for some $\theta > 1$. Then the right hand side of (10) is

$$\text{const} \times 2^{(1-\theta^2)n} \frac{1}{\sqrt{n}},$$

which is the general term in a convergent series, and

$$b_n := 2^{-(n+1)/2} c_n = \theta\sqrt{n2^{-n}\log 2},$$

also a convergent series. From (10) and the Borel-Cantelli Lemma,

$$P[H_n > b_n \text{ infinitely often}] = 0,$$

i.e. for almost all ω there exists $N(\omega)$ such that

$$H_n(\omega) \le b_n \quad \text{for } n \ge N(\omega).$$

This shows that H_n is a convergent series and completes the proof.

Lemma 2 *For a standard normal r.v.* $X \sim N(0,1)$ *and any* $c > 0$

$$P[|X| > c] < \frac{2}{c\sqrt{2\pi}} e^{-\frac{1}{2}c^2}$$

Proof: Since $x/c > 1$ for $x > c$,

$$P[X > c] = \frac{1}{\sqrt{2\pi}} \int_c^\infty e^{-x^2/2} dx < \frac{1}{\sqrt{2\pi}} \int_c^\infty \frac{x}{c} e^{-x^2/2} dx = \frac{1}{c\sqrt{2\pi}} e^{-\frac{1}{2}c^2},$$

and by symmetry $P[|X| > c] = 2P[X > c]$ for $c > 0$.

References

[1] H.P. McKean, *Stochastic Integrals*, Academic Press 1969

7

Appendix: Itô Calculus Without Probabilities

Séminaire de Probabilités XV 1979/80

ITO CALCULUS WITHOUT PROBABILITIES

by H. Föllmer

The aim of this note is to show that the Itô calculus can be developed "path by path" in the strict meaning of this term. We will derive Itô's formula as an exercise in analysis for a class of real functions of quadratic variation, including the construction of the stochastic integral $\int F'(X_{s-})dX_s$, by means of Riemann sums. Only afterwards we shall speak of probabilities in order to verify that for certain stochastic processes (semimartingales, processes of finite energy,...) almost all paths belong to this class.

Let x be a real function on $[0, \infty[$ which is right continuous and has left limits (also called *càdlàg*). We will use the following notation: $x_t = x(t)$, $\triangle x_t = x_t - x_{t-}$, $\triangle x_t^2 = (\triangle x_t)^2$.

We define a *subdivision* to be any finite sequence $\tau = (t_o, \cdots, t_k)$ such that $0 \leq t_o < \cdots < t_k < \infty$, and we put $t_{k+1} = \infty$ and $x_\infty = 0$. Let $(\tau_n)_{n=1,2,\cdots}$ be a sequence of subdivisions whose meshes converge to 0 on each compact interval. We say that x is of *quadratic variation along* (τ_n) if the discrete measures

$$\xi_n = \sum_{t_i \in \tau_n} (x_{t_{i+1}} - x_{t_i})^2 \varepsilon_{t_i} \tag{1}$$

converge weakly to a Radon measure ξ on $[0, \infty[$ whose atomic part is given by the quadratic jumps of x:

$$[x,x]_t = [x,x]_t^c + \sum_{s \le t} \triangle x_s^2, \tag{2}$$

where $[x,x]$ denotes the distribution function of ξ and $[x,x]^c$ its continuous part.

Theorem. *Let x be of quadratic variation along (τ_n) and F a function of class C^2 on $\mathbb{R}$. Then the Itô formula*

$$F(x_t) = F(x_o) + \int_0^t F(x_{s-}) dx_s + \frac{1}{2} \int_{]0,t]} F''(x_{s-}) d[x,x]_s \tag{3}$$

$$+ \sum_{s \le t} [F(x_s) - F(x_{s-}) - F'(x_{s-}) \triangle x_s - \frac{1}{2} F''(x_{s-}) \triangle x_s^2],$$

holds with

$$\int_0^t F'(x_{s-}) dx_s = \lim_n \sum_{\tau_n \ni t_i \le t} F'(x_{t_i})(x_{t_{i+1}} - x_{t_i}), \tag{4}$$

and the series in (4) is absolutely convergent.

Remark. Due to (2) the last two terms of (3) can be written as

$$\frac{1}{2} \int_0^t F''(x_{s-}) d[x,x]_s^c + \sum_{s \le t} [F(x_s) - F(x_{s-}) - F'(x_{s-}) \triangle x_s], \tag{5}$$

and we have

$$\int_0^t F''(x_{s-}) d[x,x]_s^c = \int_0^t F''(x_s) d[x,x]_s^c, \tag{6}$$

since x is a *càdlàg* function.

Proof. Let $t > 0$. Since x is right continuous we have

$$F(x_t) - F(x_o) = \lim_n \sum_{\tau_n \ni t_i \le t} [F(x_{t_{i+1}}) - F(x_{t_i})].$$

1) For the sake of clarity we first treat the particularly simple case where x is a continuous function. Taylor's formula can be written as

$$\sum_{\tau_n \ni t_i \le t} [F(x_{t_{i+1}}) - F(x_{t_i})] = \sum F'(x_{t_i})(x_{t_{i+1}} - x_{t_i})$$
$$+ \frac{1}{2} \sum F''(x_{t_i})(x_{t_{i+1}} - x_{t_i})^2 + \sum r(x_{t_i}, x_{t_{i+1}}),$$

where

$$r(a,b) \le \varphi(|b-a|)(b-a)^2, \tag{7}$$

and where $\varphi(\cdot)$ is an increasing function on $[0, \infty[$ such that $\varphi(c) \to 0$ for $c \to 0\cdot$. For $n \uparrow \infty$ the second sum of the right hand side converges to

$$\frac{1}{2} \int_{[0,t]} F''(x_s) d[x,x]_s = \frac{1}{2} \int_{]0,t]} F''(x_{s-}) d[x,x]_s$$

due to the weak convergence of the discrete measures (ξ_n); note that by (2) the continuity of x implies the continuity of $[x,x]$. The third sum, which is dominated by

$$\varphi(\max_{\tau_n \ni t_i \le t} |x_{t_{i+1}} - x_{t_i}|) \sum_{\tau_n \ni t_i \le t} (x_{t_{i+1}} - x_{t_i})^2,$$

converges to 0 since x is continuous. Thus one obtains the existence of the limit (4) and Itô's formula (3).

2) Consider now the general case. Let $\varepsilon > 0$. We divide the jumps of x on $[0,t]$ into two classes: a finite class $C_1 = C_1(\varepsilon,t)$, and a class $C_2 = C_2(\varepsilon,t)$ such that $\sum\limits_{s\in C_2} \triangle x_s^2 \le \varepsilon^2$. Let us write

$$\sum_{\tau_n \ni t_i \le t} [F(x_{t_{i+1}}) - F(x_{t_i})] = \sum\nolimits_1 [F(x_{t_{i+1}} - F(x_{t_i})] + \sum\nolimits_2 [F(x_{t_{i+1}} - F(x_{t_i})]$$

where $\sum_1$ indicates the summation over those $t_i \in \tau_n$ with $t_i \le t$ for which the interval $]t_i, t_{i+1}]$ contains a jump of class C_1. We have

$$\lim_n \sum\nolimits_1 [F(x_{t_{i+1}}) - F(x_{t_i})] = \sum_{s\in C_1} [F(x_s) - F(x_{s-})].$$

On the other hand, Taylor's formula allows us to write

$$\sum\nolimits_2 [F(x_{t_{i+1}}) - F(x_{t_i})] =$$

$$\sum_{\tau_n \ni t_i \le t} F'(x_{t_i})(x_{t_{i+1}} - x_{t_i}) + \frac{1}{2} \sum_{\tau_n \ni t_i \le t} F''(x_{t_i})(x_{t_{i+1}} - x_{t_i})^2$$

$$- \sum\nolimits_1 [F'(x_{t_i})(x_{t_{i+1}} - x_{t_i}) + \frac{1}{2} F''(x_{t_i})(x_{t_{i+1}} - x_{t_i})^2] + \sum\nolimits_2 r(x_{t_i}, x_{t_{i+1}})$$

We will show below that the second sum on the right hand side converges to

$$\frac{1}{2} \int_{]0,t]} F''(x_{s-})[x,x]_s,$$

as $n \uparrow \infty$; see (9). The third sum converges to

$$\sum_{s\in C_1} [F'(x_{s-})\triangle x_s + \frac{1}{2} F''(x_{s-})\triangle x_s^2].$$

Due to the uniform continuity of F'' on the bounded set of values $x_s \ (0 \le s \le t)$ we can assume (7), and this implies

$$\limsup_n \sum\nolimits_2 r(x_{t_i}, x_{t_{i+1}}) \le \varphi(\varepsilon+)[x,x]_{t+}. \tag{8}$$

Let ε converge to 0. Then (8) converges to 0, and

$$\sum_{s\in C_1(\varepsilon,t)} [F(x_s) - F(x_{s-}) - F'(x_{s-})\triangle x_s] - \frac{1}{2} \sum_{s\in C_1(\varepsilon,t)} F''(x_{s-})\triangle x_s^2$$

converges to the series in (3). Furthermore the series converges absolutely since

$$\sum_{s\leq t} |F(x_s) - F(x_{s-}) - F'(x_{s-})\triangle x_s| \leq \text{ const } \sum_{s\leq t} \triangle x_s^2$$

by Taylor's formula. Thus we obtain the existence of the limit in (4) and Itô's formula (3).

3) Let us show that

$$\lim_n \sum_{\tau_n \ni t_i \leq t} f(x_{t_i})(x_{t_{i+1}} - x_{t_i})^2 = \int_{]0,t]} f(x_{s-}) d[x,x]_s \tag{9}$$

for any continuous function f on $\mathbb{R}$. Let $\varepsilon > 0$, and denote by z the distribution function of the jumps in class $C_1 = C_1(\varepsilon, t)$, i. e.,

$$z_u = \sum_{C_1 \ni s \leq u} \triangle x_s \quad (u \geq 0).$$

We have

$$\lim_n \sum_{\tau_n \ni t_i \leq u} f(x_{t_i})(z_{t_{i+1}} - z_{t_i})^2 = \sum_{C_1 \ni s \leq u} f(x_{s-})\triangle x_s^2 \tag{10}$$

for each $u \geq 0$. Denote by ζ_n and η_n the discrete measures associated with z and $y = x - z$ in the sense of (1). By (10) the measures ζ_n converge weakly to the discrete measure

$$\zeta = \sum_{s\in C_1} \triangle x_s^2 \varepsilon_s.$$

Since the last sum of

$$\sum_{\tau_n \ni t_i \le u} (x_{t_{i+1}} - x_{t_i})^2 = \sum (y_{t_{i+1}} - y_{t_i})^2 + \sum (z_{t_{i+1}} - z_{t_i})^2$$
$$+2 \sum (y_{t_{i+1}} - y_{t_i})(z_{t_{i+1}} - z_{t_i})$$

converges to 0, the measures η_n converge weakly to the measure $\eta = \xi - \zeta$ whose atomic part has total mass $\le \varepsilon^2$. Hence the function $f \circ x$ is almost surely continuous with respect to the continuous part of η, and this implies

$$\limsup_n \Big| \sum_{\tau_n \ni t_i \le t} f(x_{t_i})(y_{t_{i+1}} - y_{t_i})^2 - \int_{]0,t]} f(x_{s-}) d\eta \Big| \le 2 \parallel f \parallel_t \varepsilon^2 \quad (11)$$

where $\parallel f \parallel_t = \sup\{f(x_s); 0 \le s \le t\}$. Combining (10) and (11) we obtain (9), and this completes the proof. Let us emphasize that we have followed closely the "classical" argument; see Meyer [4]. The only new contribution is the use of *weak convergence*, which allows us to give a completely analytic version.

Remarks.

1) Let $x = (x^1, \cdots, x^n)$ be a *càdlàg* function on $[0, \infty[$ with values in $\mathbb{R}^n$. We say that x a *is of quadratic variation along* (τ_n) if this holds for all real functions $x^i, x^i + x^j$ $(1 \le i, j \le n)$. In this case we put

$$[x^i, x^j]_t = \frac{1}{2}([x^i + x^j, x^i + x^j]_t - [x^i, x^i]_t - [x^j, x^j]_t)$$
$$= [x^i, x^j]^c_t + \sum_{s \le t} \triangle x^i_s \triangle x^j_s.$$

Then we have the Itô formula

$$F(x_t) = F(x_o) + \int_0^t DF(x_{s-}) dx_s + \frac{1}{2} \sum_{i,j} \int_0^t D_i D_j F(x_{s-}) d[x^i, x^j]^c_s$$
$$+ \sum_{s \le t} [F(x_s) - F(x_{s-}) - \sum_i D_i F(x_{s-}) \triangle x^i_s] \quad (12)$$

for any function F of class C^2 on $\mathbb{R}^n$, where

$$\int_0^t DF(x_{s-})dx_s = \lim_n \sum_{\tau_n \ni t_i \leq t} < DF(x_{t_i}), x_{t_{i+1}} - x_{t_i} > \tag{13}$$

($< \cdot, \cdot >=$ scalar product on $\mathbb{R}^n$). The proof is the same as above, but with more cumbersome notation.

2) The class of functions of quadratic variation is *stable with respect to* C^1 *- operations*. More precisely, if $x = (x^1, \cdots, x^n)$ is of quadratic variation along (τ_n) and F a continuously differentiable function on $\mathbb{R}^n$ then $y = F \circ x$ is of quadratic variation along (τ_n), with

$$[y, y]_t = \sum_{i,j} \int_0^t D_i F(x_s) D_j F(x_s) d[x^i, x^j]_s^c + \sum_{s \leq t} \triangle y_s^2. \tag{14}$$

This is the analytic version of a result of Meyer for semimartingales, see [4] p. 359. The proof is analogous to the previous one.
Let us now turn to stochastic processes. Let $(X_t)_{t \geq 0}$ be a semimartingale. Then, for any $t \geq 0$, the sums

$$S_{\tau,t} = \sum_{\tau \ni t_i \leq t} (X_{t_{i+1}} - X_{t_i})^2 \tag{15}$$

converge in probability to

$$[X, X]_t =< X^c, X^c >_t + \sum_{s \leq t} \triangle X_s^2$$

when the mesh of the subdivision τ converges to 0 on $[0, t]$; see Meyer [4] p. 358. For each sequence there exists thus a subsequence (τ_n) such that, almost surely,

$$\lim_n S_{\tau_n,t} = [X, X]_t \tag{16}$$

for each rational t . This implies that *almost all paths are of quadratic variation along* (τ_n). Furthermore the relation (16) is valid for all $t \geq 0$ due to (9). The Itô formula (3), applied strictly pathwise, does not depend on the sequence (τ_n). In particular, we obtain the convergence in probability of the Riemann sums in (4) to the stochastic integral

$$\int_0^t F'(X_{s-})dX_s,$$

when the mesh of τ goes to 0 on $[0, t]$.

Remarks.

1) For *Brownian motion* and an arbitrary sequence of subdivisions (τ_n) with mesh tending to 0 on each compact interval, almost all paths are of quadratic variation along (τ_n). Indeed, by Lévy's theorem we have (16) without passing to subsequences.

2) For the above argument it suffices to know that the sums (15) converge in probability to an increasing process $[X, X]$ which has paths of the form (2). The class of *processes of quadratic variation* is clearly larger than the class of semimartingales: Just consider a deterministic process of quadratic variation which is of unbounded variation. Let us mention also the *processes of finite energy* $X = M + A$ where M is a local martingale and A is a process with paths of quadratic variation 0 along the dyadic subdivisions. These processes occur in the probabilistic study of Dirichlet spaces: see Fukushima [3].

3) For a semimartingale it is known how to construct the stochastic integral $\int H_{s-}dX_s$ (H càdlàg and adapted) pathwise as a limit of Riemann sums, in the sense that the sums converge almost surely outside an exceptional set which depends on H; see Bichteler [1]. We have just shown that for the particular needs of Itô calculus, where $H = f \circ X$ (f of class C^1), the exceptional set can be chosen in advance, independently of H. It is possible to go beyond the class C^1 by treating local times "path by path". But not too far beyond: Stricker [5] has just shown that an extension to continuous functions is only possible for processes with paths of finite variation.

Traducción tomada del texto de Dieter Sondermann

579 LECTURE NOTES IN ECONOMICS AND MATHEMATICAL SYSTEMS

Dieter Sondermann

Introduction to Stochastic Calculus for Finance

A New Didactic Approach

Springer

W. Schachermayer
Encyclopedia of Quantitative Finance, Vol. 2 (2010), pp. 792-801.

The Fundamental Theorem of Asset Pricing

The subsequent theorem is one of the pillars supporting the modern theory of Mathematical Finance.

Fundamental Theorem of Asset Pricing:

The following two statements are *essentially* equivalent for a model S of a financial market:
(i) S does not allow for arbitrage (**NA**)
(ii) There exists a probability measure Q on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which is equivalent to $\mathbb{P}$ and under which the process is a martingale (**EMM**).

We have formulated this theorem in vague terms which will be made precise in the sequel: we shall formulate versions of this theorem below which use precise definitions and avoid the use of the word *essentially* above. In fact, the challenge is precisely to turn this vague "meta-theorem" into sharp mathematical results.

The story of this theorem started - like most of modern Mathematical Finance - with the work of F. Black, M. Scholes [3] and R. Merton [25]. These authors consider a model $S = (S_t)_{0 \leq t \leq T}$ of geometric Brownian motion proposed by P. Samuelson [30], which today is widely known under the name of Black–Scholes model. Presumably every reader of this article is familiar with the by now wellknown technique to price options in this framework (compare **eqf04/003: Risk Neutral Pricing**): one changes the underlying measure $\mathbb{P}$ to an equivalent measure Q under which the discounted stock price process is a martingale. Subsequently one prices options (and other derivatives) by simply taking expectations with respect to this "risk neutral" or "martingale" measure Q.

In fact, this technique was *not* the novel feature of [3] and [25]. It was used by actuaries for some centuries and it was also used by L. Bachelier [2] in 1900 who considered Brownian motion (which, of course, is a martingale) as a model $S = (S_t)_{0 \leq t \leq T}$ of a stock price process. In fact, the prices obtained by Bachelier by this method were - at least for the empirical data considered by Bachelier himself - very close to those derived from the celebrated Black–Merton–Scholes formula (compare [34]).

The decisive *novel feature* of the Black–Merton–Scholes approach was the argument which links this pricing technique with the notion of arbitrage: the pay-off function of an option can be precisely *replicated* by trading dynamically in the underlying stock. This idea, which is credited in footnote 3 of

[3] to R. Merton, opened a completely new perspective on how to deal with options, as it linked the pricing issue with the idea of hedging, i.e., dynamically trading in the underlying asset.

The technique of replicating an option is completely absent in Bachelier's early work; apparently the idea of "spanning" a market by forming linear combinations of primitive assets first appears in the Economics literature in the classic paper by K. Arrow [1]. The mathematically delightful situation, that the market is complete in the sense that all derivatives can be replicated, occurs in the Black–Scholes model as well as in Bachelier's original model of Brownian motion (compare **eqf04/008: Second Fundamental Asset Pricing Theorem**). Another example of a model in continuous time sharing this property is the compensated Poisson process, as observed by J. Cox and S. Ross [4]. Roughly speaking, these are the only models in continuous time sharing this seducingly beautiful "martingale representation property" (see [16] and [39] for a precise statement on the uniqueness of these families of models).

Appealing as it might be, the consideration of "complete markets" as above is somewhat dangerous from an economic point of view: the precise replicability of options, which is a sound mathematical theorem in the framework of the above models, may lead to the illusion that this is also true in economic reality. But, of course, these models are far from matching reality in a one-to-one manner. Rather they only highlight important aspects of reality; therefore they should not be considered as ubiquitously appropriate.

For many purposes it is of crucial importance to put oneself into a more general modeling framework.

When the merits as well as the limitations of the Black–Merton–Scholes approach unfolded in the late 70's, the investigations on the Fundamental Theorem of Asset Pricing started. As J. Harrison and S. Pliska formulate it in their classic paper [15]: "it was a desire to better understand their formula which originally motivated our study,..."

The challenge was to obtain a deeper insight into the relation of the following two aspects: on the one side the methodology of pricing by taking expectations with respect to a properly chosen "risk neutral" or "martingale" measure Q; on the other hand the methodology of pricing by "no arbitrage" considerations. Why, after all, do these two seemingly unrelated approaches yield identical results in the Black–Merton–Scholes approach? Maybe even more importantly: how far can this phenomenon be extended to more involved models?

To the best of my knowledge the first one to take up these questions in a systematic way was S. Ross ([29]; see also [4], [28], and [27]).

He chose the following setting to formalize the situation: fix a topological, ordered vector space (X, τ), modeling the possible cash flows (e.g. the pay-off function of an option) at a fixed time horizon T. A good choice is, e.g. $X = L^p(\Omega, \mathcal{F}, \mathbb{P})$, where $1 \leq p \leq \infty$ and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ is the underlying filtered probability space. The set of marketed assets M is a subspace of X.

In the context of a stock price process $S = (S_t)_{0 \leq t \leq T}$ as above, one might think of M as all the outcomes of an initial investment $x \in \mathbb{R}$ plus the result of subsequent trading according to a predictable trading strategy $H = (H_t)_{0 \leq t \leq T}$. This yields (in discounted terms) an element

$$m = x + \int_0^T H_t dS_t \tag{1}$$

in the set M of marketed claims. It is natural to price the above claim m by setting $\pi(m) = x$, as this is the net investment necessary to finance the above claim m.

For notational convenience we shall assume in the sequel that S is a one-dimensional process. It is straightforward to generalize to the case of d risky assets by assuming that S is $\mathbb{R}^d$-valued and replacing the above integral by

$$m = x + \int_0^T \sum_{i=1}^{d} H_t^i dS_t^i.$$

Some words of warning about the stochastic integral (1) seem necessary. The precise admissibility conditions which should be imposed on the stochastic integral (1), in order to make sense both mathematically as well as economically, are a subtle issue. Much of the early literature on the Fundamental Theorem of Asset Pricing struggled exactly with this question. An excellent reference is [14]. In [29] S. Ross circumvented this problem by deliberately leaving this issue aside and simply starting with the modeling assumption that the subset $M \subseteq X$ as well as a pricing operator $\pi : M \to \mathbb{R}$ are *given*.

Let us now formalize the notion of arbitrage. In the above setting, we say that the **no arbitrage** assumption is satisfied if, for $m \in M$, satisfying $m \geq 0$, $\mathbb{P}$-a.s. and $\mathbb{P}[m > 0] > 0$, we have $\pi(m) > 0$. In prose this means that it is not possible to find a claim $m \in M$, which bears no risk (as $m \geq 0, \mathbb{P}$-a.s.), yields some gain with strictly positive probability (as $\mathbb{P}[m > 0] > 0$), and such that its price $\pi(m)$ is less than or equal to zero.

The question now arises whether it is possible to extend $\pi : M \to \mathbb{R}$ to a non-negative, continuous linear functional $\pi^* : X \to \mathbb{R}$.

What does this have to do with the issue of martingale measures? This theme was developed in detail by M. Harrison and D. Kreps [14]. Suppose that $X = L^p(\Omega, \mathcal{F}, \mathbb{P})$ for some $1 \leq p < \infty$, that the price process $S = (S_t)_{0 \leq t \leq T}$ satisfies $S_t \in X$, for each $0 \leq t \leq T$, and that M contains (at least) the "simple integrals" on the process $S = (S_t)_{0 \leq t \leq T}$ of the form

$$m = x + \sum_{i=1}^{n} H_i (S_{t_i} - S_{t_{i-1}}). \tag{2}$$

Here $x \in \mathbb{R}$, $\quad 0 = t_0 < t_1 < \cdots < t_n = T$ and $(H_i)_{i=1}^n$ is a (say) bounded process which is *predictable*, i.e. H_i is $\mathcal{F}_{t_{i-1}}$-measurable. The sums in (2) are the Riemann sums corresponding to the stochastic integrals (1). The Riemann sums (2) have a clear-cut economic interpretation (see [14]). In (2) we do not have to bother about subtle convergence issues as only finite sums are involved in the definition. It therefore is a traditional (minimal) requirement that the Riemann sums of the form (2) are in the space M of marketed claims; naturally, the price of a claim m of the form (2) should be defined as $\pi(m) = x$.

Now suppose that the functional π, which is defined for the claims of the form (2) can be extended to a continuous, non-negative functional π^* defined on $X = L^p(\Omega, \mathcal{F}, \mathbb{P})$. If such an extension π^* exists, it is induced by some function $g \in L^q(\Omega, \mathcal{F}, \mathbb{P})$, where $\frac{1}{p} + \frac{1}{q} = 1$. The non-negativity of π^* is tantamount to $g \geq 0, \mathbb{P}$-a.s., and the fact that $\pi^*(1) = 1$ shows that g is the density of a probability measure Q with Radon–Nikodym derivative $\frac{dQ}{d\mathbb{P}} = g$.

If we can find such an extension π^* of π, we thus find a probability measure Q on $(\Omega, \mathcal{F}, \mathbb{P})$ for which

$$\pi^* \Big(\sum_{i=1}^{n} H_i (S_{t_i} - S_{t_{i-1}}) \Big) = \mathbb{E}_Q \Big[\sum_{i=1}^{n} H_i (S_{t_i} - S_{t_{i-1}}) \Big]$$

for every bounded predictable process $H = (H_i)_{i=1}^n$ as above, which is tantamount to $(S_t)_{0 \leq t \leq T}$ being a martingale (see [14, Th. 2], or [11, Lemma 2.2.6]).

Summing up: in the case $1 \leq p < \infty$, finding a continuous, non-negative extension $\pi^* : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ of π amounts to finding a $\mathbb{P}$-absolutely continuous measure Q with $\frac{dQ}{d\mathbb{P}} \in L^q$ and such that $(S_t)_{0 \leq t \leq T}$ is a martingale under Q.

At this stage it becomes clear that in order to find such an extension π^* of π, the Hahn–Banach theorem should come into play in some form, e.g., in one of the versions of the separating hyperplane theorem.

In order to be able to do so, S. Ross assumes ([29, p. 472]) that "...we will endow X with a strong enough topology to insure that the positive orthant $\{x \in X | x > 0\}$ is an open set,...". In practice, the only infinite-dimensional ordered topological vector space X, such that the positive orthant has non-empty interior, is $X = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, endowed with the topology induced by $\|.\|_\infty$.

Hence the two important cases, applying to S. Ross' hypothesis, are when either the probability space Ω is finite, so that $X = L^p(\Omega, \mathcal{F}, \mathbb{P})$ simply is finite dimensional and its topology does not depend on $1 \leq p \leq \infty$, or if $(\Omega, \mathcal{F}, \mathbb{P})$ is infinite and $X = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ equipped with the norm $\|.\|_\infty$.

After these preparations we can identify the two convex sets to be separated: let $A = \{m \in M : \ \pi(m) \leq 0\}$ and B be the interior of the positive cone of X. Now make the easy, but crucial, observation: these sets are disjoint if and only if the no arbitrage condition is satisfied. As one always can separate an *open* convex set from a disjoint convex set, we find a functional $\tilde{\pi}$ which is strictly positive on B, while $\tilde{\pi}$ takes non-positive values on A. By normalizing $\tilde{\pi}$, i.e., letting $\pi^* = \tilde{\pi}(1)^{-1}\tilde{\pi}$ we have thus found the desired extension.

In summary, the first precise version of the Fundamental Theorem of Asset Pricing is established in [29], the proof relying on the Hahn–Banach theorem. There are, however, serious limitations: in the case of infinite $(\Omega, \mathcal{F}, \mathbb{P})$ the present result only applies to $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ endowed with the norm topology. In this case the continuous linear functional π^* only is in $L^\infty(\Omega, \mathcal{F}, \mathbb{P})^*$ and not necessarily in $L^1(\Omega, \mathcal{F}, \mathbb{P})$; in other words we cannot be sure that π^* is induced by a probability measure Q, as it may happen that $\pi^* \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})^*$ also has a singular part.

Another drawback, which already appears in the case of finite-dimensional Ω (in which case π^* certainly is induced by some Q with $\frac{dQ}{d\mathbb{P}} = g \in L^1(\Omega, \mathcal{F}, \mathbb{P})$) is the following: we cannot be sure that the function g is strictly positive $\mathbb{P}$-a.s. or, in other words, that Q is equivalent to $\mathbb{P}$.

After this early work by S. Ross a major advance in the theory was achieved between '79 and '81 by three seminal papers [14], [15], [24] by M. Harrison, D. Kreps and S. Pliska. In particular, [14] is a landmark in the field. It uses a similar setting as [29], namely an ordered topological vector space (X, τ) and a linear functional $\pi : M \to \mathbb{R}$, where M is a linear subspace of X. Again the question is whether there exists an extension of π to a linear, continuous, strictly positive $\pi^* : X \to \mathbb{R}$. This question is related in [14] to the issue whether (M, π) is viable *as a model of economic equilibrium*. Under proper assumptions on the convexity and continuity of the preferences of agents this is shown to be equivalent to the extension discussed above.

The paper [14] also analyses the case when Ω is finite. Of course, only processes $S = (S_t)_{t=0}^T$ indexed by finite, discrete time $\{0, 1, ..., T\}$ make sense in this case. For this easier setting the following precise theorem was stated and proved in the subsequent paper [15] by J. Harrison and S. Pliska:

Theorem 1. *([15, Th. 2. 7.]): Suppose the stochastic process $S = (S_t)_{t=0}^T$ is based on a finite, filtered, probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, \mathbb{P})$. The market model contains no arbitrage possibilities if and only if there is an equivalent martingale measure for S.*

The proof again relies on a (finite-dimensional version) of the Hahn–Banach theorem plus an extra argument making sure to find a measure Q which is *equivalent to* $\mathbb{P}$. M. Harrison and S. Pliska thus have achieved a precise version of the above meta-theorem in terms of equivalent martingale measures which does not use the word "essentially". Actually, the theme of the Harrison–Pliska theorem goes back much further, to the work of A. Shimony [35] and J. Kemeny [22] on symbolic logic in the tradition of R. Carnap, B. de Finetti, and F. Ramsey. These authors showed that, in a setting with only finitely many states of the world, a family of possible bets does not allow (by taking linear combinations) for making a riskless profit (i.e. one certainly does not lose but wins with strictly positive probability), if and only if there is a probability measure Q on these finitely many states, which prices the possible bets by taking conditional Q-expectations.

The restriction to finite Ω is very severe in applications: the flavor of the theory, building on Black–Scholes–Merton, is precisely the concept of *continuous time*. Of course, this involves infinite probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$.

Many interesting questions were formulated in the papers [14] and [15] hinting on the difficulties to prove a version of the Fundamental Theorem of Asset Pricing beyond the setting of finite probability spaces.

A major break-through in this direction was achieved by D. Kreps [24]: as above, let $M \subseteq X$ and a linear functional $\pi : M \to \mathbb{R}$ be given. The typical choice for X will now be $X = L^p(\Omega, \mathcal{F}, \mathbb{P})$, for $1 \leq p \leq \infty$, equipped with the topology τ of convergence in norm, or, if $X = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, equipped with the Mackey topology τ induced by $L^1(\Omega, \mathcal{F}, \mathbb{P})$. This setting will make sure that a continuous linear functional on (X, τ) will be induced by a measure Q which is absolutely continuous with respect to $\mathbb{P}$.

The no arbitrage assumption means that $M_0 := \{m \in M : \pi(m) = 0\}$ intersects the positive orthant X_+ of X only in $\{0\}$. In order to obtain an extension of π to a continuous, linear functional $\pi^* : X \to \mathbb{R}$ we have to find an element in $(X, \tau)^*$ which separates the convex set M_0 from the disjoint convex set $X_+ \backslash \{0\}$, i.e., the positive orthant of X with 0 deleted.

Easy examples show that, in general, this is not possible. In fact, this is not much of a surprise (if X is infinite-dimensional) as we know that some topological condition is needed for the Hahn–Banach theorem to work.

It is always possible to separate a *closed* convex set from a disjoint *compact* convex set by a continuous linear functional. In fact, one may even get strict separation in this case. It is this version of the Hahn–Banach theorem which D. Kreps eventually applies.

But how? After all, neither M_0 nor $X_+\backslash\{0\}$ are closed in (X,τ), let alone compact.

Here is the ingenious construction of D. Kreps: define

$$A = \overline{M_0 - X_+}\ , \tag{3}$$

where the bar denotes the closure with respect to the topology τ. We shall require that A still satisfies

$$A \cap X_+ = \{0\}. \tag{4}$$

This property is baptized as *"no free lunch"* by D. Kreps:

Definition 2. *[24]: The financial market defined by (X,τ), M, and π admits a free lunch if there are nets $(m_\alpha)_{\alpha\in I} \in M_0$ and $(h_\alpha)_{\alpha\in I} \in X_+$ such that*

$$\lim_{\alpha\in I}\,(m_\alpha - h_\alpha) = x \tag{5}$$

for some $x \in X_+\backslash\{0\}$.

It is easy to verify that the negation of the above definition is tantamount to the validity of (4).

The economic interpretation of the "no free lunch" condition is a sharpening of the "no arbitrage condition". If the latter is violated, we can simply find an element $x \in X_+\backslash\{0\}$ which also lies in M_0. If the former fails, we cannot quite guarantee this, but we can find $x \in X_+\backslash\{0\}$ which *can be approximated* in the τ-topology by elements of the form $m_\alpha - h_\alpha$. The passage from m_α to $m_\alpha - h_\alpha$ means that agents are allowed to "throw away money", i.e. to abandon a positive element $h_\alpha \in X_+$. This combination of the "free disposal" assumption with the possibility of passing to limits is crucial in Kreps' approach (3) as well as in most of the subsequent literature. It was shown in [32, Ex. 3.3]; (compare also [33]) that the (seemingly ridiculous) "free disposal" assumption cannot be dropped.

Definition (3) is tailor-made for the application of Hahn–Banach. If the no free lunch condition (4) is satisfied, we may, for any $h \in X_+$, separate the τ-closed, convex set A from the one-point set $\{h\}$ by an element $\pi_h \in (X,\tau)^*$. As $0 \in A$ we may assume that $\pi_h|_A \leq 0$ while $\pi_h(h) > 0$. We thus have obtained a non-negative (as $-X_+ \subseteq A$), continuous linear functional π_h which is strictly positive on a given $h \in X_+$. Supposing that X_+ is τ-separable (which is the case in the above setting of L^p-spaces if $(\Omega, \mathcal{F}, \mathbb{P})$ is countably generated), fix a dense sequence $(h_n)_{n=1}^\infty$ and find strictly positive scalars $\mu_n > 0$ such that $\pi^* = \sum_{n=1}^\infty \mu_n \pi_{h_n}$ converges to a probability measure in $(X, \tau)^* = L^q(\Omega, \mathcal{F}, \mathbb{P})$, where $\frac{1}{p} + \frac{1}{q} = 1$. This yields the desired extension π^* of π which is *strictly positive* on $X_+ \setminus \{0\}$.

We still have to specify the choice of (M_0, π). The most basic choice is to take for given $S = (S_t)_{0 \le t \le T}$ the space generated by the "simple integrands" (2) as proposed by J. Harrison and D. Kreps [14]. We thus may deduce from Kreps' arguments in [24] the following version of the Fundamental Theorem of Asset pricing.

Theorem 3. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be countably generated and $X = L^p(\Omega, \mathcal{F}, \mathbb{P})$ endowed with the norm topology τ, if $1 \le p < \infty$, or the Mackey topology induced by $L^1(\Omega, \mathcal{F}, \mathbb{P})$, if $p = \infty$.*

Let $S = (S_t)_{0 \le t \le T}$ be a stochastic process taking values in X. Define $M_0 \subseteq X$ to consist of the simple stochastic integrals $\sum_{i=1}^n H_i(S_{t_i} - S_{t_{i-1}})$ as in (2).

Then the "no free lunch" condition (3) is satisfied if and only if there is a probability measure Q with $\frac{dQ}{d\mathbb{P}} \in L^q(\Omega, \mathcal{F}, \mathbb{P})$, where $\frac{1}{p} + \frac{1}{q} = 1$, such that $(S_t)_{0 \le t \le T}$ is a Q-martingale.

This remarkable theorem of D. Kreps sets new standards. For the first time, we have a mathematically precise statement of our meta-theorem applying to a general class of models *in continuous time*. There are still some limitations, however.

When applying the theorem to the case $1 \le p < \infty$ we find the requirement $\frac{dQ}{d\mathbb{P}} \in L^q(\Omega, \mathcal{F}, \mathbb{P})$ for some $q > 1$, which is not very pleasant. After all, we want to know: what exactly corresponds (in terms of some no arbitrage condition) to the existence of an equivalent martingale measure Q? The q-moment condition is unnatural in most applications. In particular, it is not invariant under equivalent changes of measures as is done often in the applications.

The most interesting case of the above theorem is $p = \infty$. But in this case the requirement $S_t \in X = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ is unduly strong for most applications. In addition, for $p = \infty$ we run into the subtleties of the Mackey topology τ (or the weak-star topology, which does not make much of a difference) on $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. We shall discuss this issue below.

The "heroic period" of the development of the Fundamental Theorem of Asset Pricing marked by S. Ross [29], Harrison–Kreps [14], Harrison–Pliska [15] and D. Kreps [24], put the issue on safe mathematical grounds and brought some spectacular results. However, it still left many questions open; quite a number of them were explicitly stated as open problems in these papers.

Subsequently a rather extensive literature developed, answering these problems and opening new perspectives. We cannot give a full account on all of this literature and refer, e.g., to the monograph [11] for more extensive information. We can give an outline.

As regards the situation for $1 \leq p \leq \infty$ in Kreps' theorem, this issue was further developed by D. Duffie and C.F. Huang [12] and, in particular, by C. Stricker [36]. This author related the no free lunch condition of D. Kreps to a theorem by J.A. Yan [37] obtained in the context of the Bichteler–Dellacherie theorem on the characterisation of semi-martingales. Using Yan's theorem, Stricker gave a different proof of Kreps' theorem which does not need the assumption that $(\Omega, \mathcal{F}, \mathbb{P})$ is countably generated.

A beautiful extension of the Harrison–Pliska theorem was obtained in 1990 by R. Dalang, A. Morton and W. Willinger [5]. They showed that, for an $\mathbb{R}^d$-valued process $(S_t)_{t=0}^T$ in finite discrete time, the no arbitrage condition is indeed equivalent to the existence of an equivalent martingale measure. The proof is surprisingly tricky, at least for the case $d \geq 2$. It is based on the measurable selection theorem (the suggestion to use this theorem is acknowledged to F. Delbaen). Different proofs of the Dalang–Morton–Willinger theorem have been given in [31], [20], [26], [17], and [21].

A important question left unanswered by D. Kreps was whether one can, in general, replace the use of nets $(m_\alpha - h_\alpha)_{\alpha \in I}$, indexed by α ranging in a general ordered set I, simply by sequences $(m_n - h_n)_{n=1}^\infty$. In the context of continuous processes $S = (S_t)_{0 \leq t \leq T}$ a positive answer was given by F. Delbaen in [6], if one is willing to make the harmless modification to replace the deterministic times $0 = t_0 \leq t_1 \leq \cdots \leq t_n = T$ in (2) by *stopping times* $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_n = T$. A second case, where the answer to this question is positive are processes $S = (S_t)_{t=0}^\infty$ in infinite, discrete time as shown in [32].

The Banach–Steinhaus theorem implies that, for a *sequence* $(m_n - h_n)_{n=1}^\infty$ converging in $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the weak-star (or Mackey) topology, the norms $(||m_n - h_n||_\infty)_{n=1}^\infty$ remain bounded ("uniform boundedness principle"). Therefore it follows that in the above two cases of continuous processes $S = (S_t)_{0 \leq t \leq T}$ or processes $(S_t)_{t=0}^\infty$ in infinite, discrete time, the "no free lunch" condition of D. Kreps can be equivalently replaced by the "no free lunch with bounded risk" condition introduced in [32]: in (5) above we additionally impose that $(||m_\alpha - h_\alpha||_\infty)_{\alpha \in I}$ remains bounded. In this case we have that there is a constant $M > 0$ such that $m_\alpha \geq -M$, $\mathbb{P}$-a.s. for each

$\alpha \in I$, which explains the wording "bounded risk".

However, in the context of general semi-martingale models $S = (S_t)_{0 \leq t \leq T}$, a counter-example was given by F. Delbaen and the present author in ([7, Ex. 7.8]) showing that the "no free lunch with bounded risk" condition does not imply the existence of an equivalent martingale measure. Hence, in a general setting and by only using simple integrals, there is no hope to get any more precise information on the free lunch condition than the one provided by Kreps' theorem.

At this stage it became clear that, in order to obtain sharper results, one has to go beyond the framework of simple integrals (2) and rather use general stochastic integrals (1). After all, the simple integrals only are a technical gimmick, analogous to step functions in measure theory. In virtually all the applications, e.g., the replication strategy of an option in the Black–Scholes model, one uses general integrals of the form (1).

General integrands pose a number of questions to be settled. First of all, the integral (1) has to be mathematically well-defined. The theory of stochastic calculus starting with K. Itô, and developed in particular by the Strasbourg school of probability around P.-A. Meyer, provides very precise information on this issue: there is a good integration theory for a given stochastic process $S = (S_t)_{0 \leq t \leq T}$ *if and only if* S is a semi-martingale (theorem of Bichteler–Dellacherie).

Hence mathematical arguments lead to the model assumption that S has to be a semi-martingale. But what about an economic justification of this assumption? Fortunately the economic reasoning hints in the same direction. It was shown by F. Delbaen and the present author that, for a locally bounded stochastic process $S = (S_t)_{0 \leq t \leq T}$, a very weak form of Kreps' no free lunch condition involving simple integrands (2), implies already that S is a semi-martingale (see [7, Theorem 7.2], for a precise statement).

Hence it is natural to assume that the model $S = (S_t)_{0 \leq t \leq T}$ of stock prices is a semi-martingale so that the stochastic integral (2) makes sense mathematically, for all S-integrable, predictable processes $H = (H_t)_{0 \leq t \leq T}$. As pointed out, [14] and [15] impose in addition an admissibility condition to rule out doubling strategies and similar schemes.

Definition 4. *([7, Def. 2.7]): An S-integrable predictable process $H = (H_t)_{0 \leq t \leq T}$ is called admissible if there is a constant $M > 0$ such that*

$$\int_0^t H_u dS_u \geq -M, \quad a.s., \textit{for } 0 \leq t \leq T. \tag{6}$$

The economic interpretation is that the economic agent, trading according to the strategy, has to respect a finite credit line M.

Let us now sketch the approach of [7]. Define

$$K = \left\{ \int_0^T H_t dS_t : H \text{ admissible} \right\} \tag{7}$$

which is a set of (equivalence classes of) random variables. Note that by (6) the elements $f \in K$ are uniformly bounded from below, i.e., $f \geq -M$ for some $M \geq 0$. On the other hand, there is no reason why the positive part f_+ should obey any boundedness or integrability assumption.

As a next step we "allow agents to throw away money" similarly as in Kreps' work [24]. Define

$$\begin{aligned} C &= \{g \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) : g \leq f \text{ for some } f \in K\} \\ &= [K - L^0_+(\Omega, \mathcal{F}, \mathbb{P})] \cap L^\infty(\Omega, \mathcal{F}, \mathbb{P}), \end{aligned} \tag{8}$$

where $L^0_+(\Omega, \mathcal{F}, \mathbb{P})$ denotes the set of non-negative measurable functions.

By construction, C consists of bounded random variables, so that we can use the functional analytic duality theory between L^∞ and L^1. The difference of the subsequent definition to Kreps' approach is that it pertains to the norm topology $\|.\|_\infty$ rather than to the Mackey topology on $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 5. *([11, 2.8]): A locally bounded semi-martingale $S = (S_t)_{0 \leq t \leq T}$ satisfies the no free lunch with vanishing risk condition if*

$$\bar{C} \cap L^\infty_+(\Omega, \mathcal{F}, \mathbb{P}) = \{0\}, \tag{9}$$

where $\bar{C}$ denotes the $\|.\|_\infty$-closure of C.

Here is the translation of (9) into prose: the process S fails the above condition iff there is a function $g \in L^\infty_+(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{P}[g > 0] > 0$ and a sequence $(f^n)_{n=1}^\infty$ of the form

$$f^n = \int_0^T H^n_t dS_t,$$

where H^n are admissible integrands, such that

$$f_n \geq g - \tfrac{1}{n}, \qquad a.s. \tag{10}$$

Hence the condition of no free lunch with vanishing risk is intermediate between the (stronger) no free lunch condition of D. Kreps and the (weaker) no arbitrage condition. The latter would require that there is a non-negative function g with $\mathbb{P}[g > 0] > 0$ which is of the form

$$g = \int_0^T H_t dS_t,$$

for an admissible integrand H. Condition (10) does not quite guarantee this, but something - at least from an economic point of view - very close: we can *uniformly* approximate from below such a g by the outcomes f_n of admissible trading strategies.
The main result of F. Delbaen and the author [7] reads as follows.

Theorem 6. *([7, Corr. 1.2]): Let $S = (S_t)_{0 \le t \le T}$ be a locally bounded real-valued semi-martingale.*

There is a probability measure Q on $(\Omega, \mathcal{F})$ which is equivalent to $\mathbb{P}$ and under which S is a local martingale if and only if S satisfies the condition of no free lunch with vanishing risk.

This is a mathematically precise theorem which, in my opinion, is quite close to the vague "meta-theorem" at the beginning of this article. The difference to the intuitive "no arbitrage" idea is that the agent has to be willing to sacrifice (at most) the quantity $\frac{1}{n}$ in (10), where we may interpret $\frac{1}{n}$ as, say, 1 Cent.

The proof of the above theorem is rather longish and technical and a more detailed discussion goes beyond the scope of the present article. To the best of my knowledge, no essential simplification of this proof has been achieved so far (compare [19]).

Mathematically speaking, the statement of the theorem looks very suspicious at first glance: after all, the *no free lunch with vanishing risk* condition pertains to the *norm topology* of $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. Hence it seems that, when applying the Hahn–Banach theorem, one can only obtain a linear functional in $L^\infty(\Omega, \mathcal{F}, \mathbb{P})^*$, which is not necessarily of the form $\frac{dQ}{d\mathbb{P}} \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, as we have seen in Ross' work [29].

The reason why the above theorem nevertheless is true is a little miracle: it turns out ([7, Th. 4.2]) that, under the assumption of no free lunch with vanishing risk, the set C defined in (8) is *automatically* weak-star closed in $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. This pleasant fact is not only a crucial step in the proof of the above theorem; maybe even more importantly, it also found other applications. For example, to find general existence results in the theory of utility optimisation (**eqf04/009: Expected utility maximization**) it is of crucial importance to have a closedness property of the set over which one optimizes: for these applications the above result is very useful (see, e.g.,

[23]).

Without going into the details of the proof let me point out the importance of certain elements in the set K. The admissibility rules out the use of doubling strategies. The opposite of such a strategy can be called a suicide strategy. It is the mathematical equivalent of making a bet at the roulette, leaving it as well as all gains on the table as long as one keeps winning, and wait until one loses for the first time. Such strategies, although admissible, do not reflect economic efficiency. More precisely we define

Definition 7. *: An admissible outcome $\int_0^T H_t dS_t$ is called maximal if there is no other admissible strategy H' such that $\int_0^T H'_t dS_t \geq \int_0^T H_t dS_t$ with $\mathbb{P}[\int_0^T H'_t dS_t > \int_0^T H_t dS_t] > 0$*

In the proof of Theorem 6, these elements play a crucial role and the heart of the proof consists in showing that every element in K is dominated by a maximal element. But besides their mathematical relevance they also have a clear economic interpretation. There is no use in implementing a strategy that is not maximal as one can do better. Non-maximal elements can also be seen as bubbles, see [18].

In Theorem 6 we only assert that S is a *local* martingale under Q. In fact, this technical concept cannot be avoided in this setting. Indeed, fix an S-integrable, predictable, admissible process $H = (H_t)_{0\leq t\leq T}$ as well as a bounded, predictable, strictly positive process $(k_t)_{0\leq t\leq T}$. The subsequent identity trivially holds true.

$$\int_0^t H_u dS_u = \int_0^t \frac{H_u}{k_u} d\tilde{S}_u, \quad 0 \leq t \leq T, \tag{11}$$

where

$$\tilde{S}_u = \int_0^u k_v dS_v, \qquad 0 \leq u \leq T. \tag{12}$$

The message of (11) and (12) is that the class of processes obtained by taking admissible stochastic integrals on S or $\tilde{S}$ simply coincide. An easy interpretation of this rather trivial fact is that the possible investment opportunities do not depend on whether stock prices are denoted in Euros or in Cents (this corresponds to taking $k_t \equiv 100$ above).

But it may very well happen that $\tilde{S}$ is a martingale while S only is a local martingale. In fact, the concept of local martingales may even be *characterized* in these terms ([10, Proposition 2.5]): a semi-martingale S is a local martingale if and only if there is a strictly positive, decreasing, predictable process k such that $\tilde{S}$ defined in (12) is a martingale.

Again we want to emphasize the role of the maximal elements. It turns out, see [8] and [11], that if $\int_0^T H_t dS_t$ is maximal, *if and only if* there is an equivalent local martingale measure Q such that the process $\int_0^t H_u dS_u$ is a martingale and not just a local martingale under Q. One can show, see [9] and [11], that for a given sequence of maximal elements $\int_0^T H_t^n dS_t$, one can find one and the same equivalent local martingale measure Q such that *all* the processes $\int_0^t H_u^n dS_u$ are Q-martingales. Another useful and related characterisation, see [8] and [11], is that if a process $V_t = x + \int_0^t H_u dS_u$ defines a maximal element $\int_0^T H_u dS_u$ and remains strictly positive, the whole financial market can be rewritten in terms of V as a new numéraire without losing the no-arbitrage properties. The change of numéraire and the use of the maximal elements allows to introduce a numéraire invariant concept of admissibility, see [9] for details. An important result in this paper is that the sum of maximal elements is again a maximal element.

Theorem 6 above still contains one severe limitation of generality, namely the local boundedness assumption on S. As long as we only deal with continuous processes S, this requirement is, of course, satisfied. But if one also considers processes with jumps, in most applications it is natural to drop the local boundedness assumption.

The case of general semi-martingales S (without any boundedness assumption) was analyzed in [10]. Things become a little trickier as the concept of local martingales has to be weakened even further: we refer to **eqf04/007: Equivalent Martingale Measure and Ramifications** for a discussion of the concept of sigma-martingales. This concept allows to formulate a result pertaining to a perfectly general setting.

Theorem 8. *([7, Corr. 1.2]): Let $S = (S_t)_{0 \leq t \leq T}$ be an $\mathbb{R}^d$-valued semi-martingale.*

There is a probability measure Q on $(\Omega, \mathcal{F})$ which is equivalent to $\mathbb{P}$ and under which S is a sigma-martingale if and only if S satisfies the condition of no free lunch with vanishing risk with respect to admissible strategies.

One still may ask whether it is possible to formulate a version of the fundamental theorem which does not rely on the concepts of local or sigma-, but rather on "true" martingales.

This was achieved by J. Yan [38] by applying a clever change of numéraire technique, (**eqf04/010: Change of Numéraire** compare also [13, Section 5]): let us suppose that $(S_t)_{0 \leq t \leq T}$ is a *positive* semi-martingale, which is natural if we model, e.g., prices of shares (while the previous setting of not necessarily positive price processes also allows for the modeling of forwards, futures etc.).

Let us weaken the admissibility condition (6) above, by calling a predictable, S-integrable process *allowable* if

$$\int_0^t H_u dS_u \geq -M(1+S_t) \quad a.s., \text{for } 0 \leq t \leq T. \tag{13}$$

The economic idea underlying this notion is wellknown and allows for the following interpretation: an agent holding M units of stock and bond may, in addition, trade in S according to the trading strategy H satisfying (13); she will then remain liquid during $[0, T]$.

By taking $S+1$ as new numéraire and replacing admissible by allowable trading strategies Yan obtains the following Theorem.

Theorem 9. *([38, Theorem 3.2]) Suppose that S is a positive semi-martingale. There is a probability measure Q on $(\Omega, \mathcal{F})$ which is equivalent to $\mathbb{P}$ and under which S is a martingale if and only if S satisfies the condition of no free lunch with vanishing risk with respect to allowable trading strategies.*

Walter Schachermayer

References

[1] Arrow, K. (1964) The role of securities in the optimal allocation of risk-bearing. *Rev. Econ. Stud.* **31**, 91–96 (1964).

[2] Bachelier L. (1964) Théorie de la Spéculation. *Ann. Sci. Ecole Norm. Sup.* **17**, 21–86 (1900) English translation in: *The Random Character of stock market prices* (P. Cootner, editor), MIT Press.

[3] Black, F, Scholes, M. (1973) The pricing of options and corporate liabilities. *Journal of Political Economy* **81**, 637–659.

[4] Cox, J, Ross, S. (1976) The valuation of options for alternative stochastic processes. *J. Financial Econom.* **3**, 145–166.

[5] Dalang, R.C, Morton, A, Willinger, W. (1990) Equivalent Martingale measures and no-arbitrage in stochastic securities market model. *Stochastics and Stochastic Reports* **29**, 185–201.

[6] Delbaen, F. (1992) Representing Martingale Measures when Asset Prices are Continuous and Bounded. *Mathematical Finance* **2**, 107–130.

[7] Delbaen, F, Schachermayer, W. (1994) A General Version of the Fundamental Theorem of Asset Pricing. *Mathematische Annalen* **300**, 463–520.

[8] Delbaen, F, Schachermayer, W. (1995) The no-arbitrage condition under a change of numéraire. *Stochastics Stochastics Rep.* **53**, 213–226.

[9] Delbaen, F, Schachermayer, W. (1997) The Banach space of workable contingent claims in arbitrage theory. *Ann IHP Prob Stat* **33**, 113–144.

[10] Delbaen, F, Schachermayer, W. (1998) The Fundamental Theorem of Asset Pricing for Unbounded Stochastic Processes. *Mathematische Annalen* **312**, 215–250.

[11] Delbaen, F, Schachermayer, W. (2006) The Mathematics of Arbitrage. *Springer Finance*, 371.

[12] Duffie, D, Huang, C.F. (1986) Multiperiod security markets with differential information; martingales and resolution times. *Journal of Mathematical Economics* **15**, 283–303.

[13] Guasoni, P, Rásonyi, M, Schachermayer, W. (2009) The Fundamental Theorem of Asset Pricing for Continuous Processes under Small Transaction Costs. *Annals of Finance*.

[14] Harrison, J.M, Kreps, D.M. (1979) Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory* **20**, 381–408.

[15] Harrison, J.M, Pliska, S.R. (1981) Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and their Applications* **11**, 215–260.

[16] Harrison, J.M, Pliska, S.R. (1983) A stochastic calculus model of continuous trading: Complete markets. *Stochastic Processes and their Applications* **11**, 313–316.

[17] Jacod, J, Shiryaev, A.N. (1998) Local martingales and the fundamental asset pricing theorems in the discrete-time case. *Finance and Stochastics* (3) **2**, 259–273.

[18] Jarrow, R, Protter, P, Shimbo, K. (2007) Asset price bubbles in complete markets. *Advances in Math. Finance* Appl. Numer. Harmon. Anal. Birkhäuser, Boston, Boston MA, 97–121.

[19] Kabanov, Y.M. (1997) On the FTAP of Kreps-Delbaen-Schachermayer (English). *Y.M. Kabanov (ed.) et al., Statistics and control of stochastic processes.* The Liptser Festschrift. Papers from the Steklov seminar held in Moscow, Russia, 1995-1996. Singapore: World Scientific, 191–203.

[20] Kabanov, Y.M, Kramkov, D. (1994) No-arbitrage and equivalent martingale measures: An elementary proof of the Harrison–Pliska theorem. *Theory Prob. Appl.* (3) **39**, 523–527.

[21] Kabanov, Y.M, Stricker, Ch. (2001) A teachers' note on no-arbitrage criteria. *Séminaire de Probabilités XXXV, Springer Lecture Notes in Mathematics* 1755, 149–152.

[22] Kemeny, J.G. (1955) Fair bets and inductive probabilities. *Journal of Symbolic Logic* (3) **20**, 263–273.

[23] Kramkov, D, Schachermayer, W. (1999) The Asymptotic Elasticity of Utility Functions and Optimal Investment in Incomplete Markets. *Annals of Applied Probability* (3) **9**, 904–950.

[24] Kreps, D.M. (1981) Arbitrage and Equilibrium in Economics with infinitely many Commodities. *Journal of Mathematical Economics* **8**, 15–35.

[25] Merton, R.C. (1973) The theory of rational option pricing. *Bell J. Econ. Manag. Sci.* **4**, 141–183.

[26] Rogers, L.C.G. (1994) Equivalent martingale measures and no-arbitrage. *Stochastics and Stochastic Reports* (1–2) **51**, 41–49.

[27] Ross, S. (1976) The arbitrage theory of capital asset pricing. *J. Econ. Theor.* **13**, 341–360.

[28] Ross, S. (1977) Return, Risk and Arbitrage. *Risk and Return in Finance* **1**, 189–218.

[29] Ross, S. (1978) A simple approach to the valuation of risky streams. *J. Business* **51**, 453–475.

[30] Samuelson, P.A. (1965) Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* **6**, 41–50.

[31] Schachermayer, W. (1992) A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time. *Insurance: Mathematics and Economics* (4) **11**, 249–257.