Itô’s stochastic calculus: Its surprising power for applications

Hiroshi Kunita

Kyushu University

Available online 1 February 2010

Abstract

We trace Itô’s early work in the 1940s, concerning stochastic integrals, stochastic differential equations (SDEs) and Itô’s formula. Then we study its developments in the 1960s, combining it with martingale theory. Finally, we review a surprising application of Itô’s formula in mathematical finance in the 1970s. Throughout the paper, we treat Itô’s jump SDEs driven by Brownian motions and Poisson random measures, as well as the well-known continuous SDEs driven by Brownian motions.

MSC: 60-03; 60H05; 60H30; 91B28

Keywords: Itô’s formula; Stochastic differential equation; Jump–diffusion; Black–Scholes equation; Merton’s equation

0. Introduction

This paper is written for Kiyosi Itô of blessed memory. Itô’s famous work on stochastic integrals and stochastic differential equations started in 1942, when mathematicians in Japan were completely isolated from the world because of the war. During that year, he wrote a colloquium report in Japanese, where he presented basic ideas for the study of diffusion processes and jump–diffusion processes. Most of them were completed and published in English during 1944–1951.

Itô’s work was observed with keen interest in the 1960s both by probabilists and applied mathematicians. Itô’s stochastic integrals based on Brownian motion were extended to those...
based on martingales, through Doob–Meyer’s decomposition of positive submartingales. The famous Itô formula was extended to semimartingales.

Linking up with martingale theory, Itô’s stochastic calculus became a very useful tool for applied problems. In stochastic control and filtering problems, Itô’s stochastic differential equations are employed as models of dynamical systems disturbed by noise. Then conditional distributions of the filtering are obtained by solving stochastic partial differential equations.

Another important and thorough application of Itô’s stochastic calculus took place in mathematical finance. In 1973, Itô’s formula was applied in a striking way to the pricing of options by Black–Scholes and Merton. The martingale representation theorem was used for the hedging problem of derivatives and Girsanov’s theorem was applied for constructing a risk neutral measure. These are three pieces of tool kits or three jewels in mathematical finance.

In the first part of this paper, we will look back at Itô’s early work on stochastic calculus. It includes the Lévy–Itô decomposition of a Lévy process and stochastic differential equations based on Lévy processes. In Section 2, we will consider a combination of Itô’s work with martingale theory. Then we will explain in detail Itô’s stochastic integral based on a Poisson random measure. In Section 3, we will trace an application of Itô’s theory in mathematical finance. We will review the historical work by Black–Scholes and Merton on pricing options both for diffusion and jump–diffusion models.

The author was a student of K. Itô at Kyoto in 1958–1961. Itô always loved and encouraged students and young mathematicians. He recommended us to study an important but difficult paper. Following his suggestion, I read P. Lévy’s paper on Markov chains (1951). Though I could not understand the full importance of the paper, I got a small idea from the paper and was able to finish the graduate course. I wish to express my cordial thanks to Kiyosi Itô.

1. Itô’s early work

1.1. The Lévy–Itô decomposition of a Lévy process

Itô’s first work [7] was published in 1942. The title is “On stochastic processes”. In those days, a stochastic process was often understood as a family of multivariate distributions. A Markov process was considered as a system of transition probabilities. After the fundamental work of Kolmogorov [21], 1933, the view of stochastic processes was changed greatly. Doob and Itô defined them as a family of random variables \( \{X_t, t \in T\} \) on an infinite-dimensional probability space, whose finite-dimensional distributions coincide with given multivariate distributions.

In the first half of the paper, Itô showed how we can construct a stochastic process called a Lévy process associated with a given family of infinitely divisible distributions. A Markov process was considered as a system of transition probabilities. After the fundamental work of Kolmogorov [21], 1933, the view of stochastic processes was changed greatly. Doob and Itô defined them as a family of random variables \( \{X_t, t \in T\} \) on an infinite-dimensional probability space, whose finite-dimensional distributions coincide with given multivariate distributions.

In the first half of the paper, Itô showed how we can construct a stochastic process called a Lévy process associated with a given family of infinitely divisible distributions. In the second half, he discussed the decomposition of a Lévy process. Let \( X_t, t \in T = [0, T] (0 < T < \infty) \) be an \( R \)-valued stochastic process (where \( R \) is a Euclidean space) on the probability space \( (\Omega, \mathcal{F}, P) \). It is called a Lévy process if the following three properties are satisfied:

1. \( X_0 = 0 \) a.s. Sample paths are right continuous with left-hand limits a.s.
2. Let \( X_{t-} = \lim_{\epsilon \to 0^+} X_{t-\epsilon} \). Then \( X_t = X_{t-} \) holds a.s. for any \( t \).
3. It has independent increments, i.e., for any \( a \leq t_0 < t_1 < \cdots < t_n \leq b \), the random variables \( X_{t_i} - X_{t_{i-1}} , i = 1, \ldots, n \) are independent.

If the law of \( X_{t+h} - X_{s+h} \) does not depend on \( h \) for any \( s, t \), the Lévy process is called homogeneous. (Recently, a homogeneous Lévy process is simply called a Lévy process). If the sample paths \( X_t \) are continuous a.s., the Lévy process \( X_t \) is called a Brownian motion.
In particular a homogeneous Brownian motion \(W_t\) with mean 0 and covariance \(tI\) is called a standard Brownian motion.

Itô looked into the sample paths of a Lévy process profoundly and showed that any Lévy process consists of two components: a Brownian motion and a Poisson random measure. The idea goes back to P. Lévy. His intuitive work became clear through Itô’s work.

Before the description of Itô’s work, we will define a Poisson random measure on \(T \times R_0\), where \(T = T - \{0\}\) and \(R_0 = R - \{0\}\). Let \(\mathcal{B}\) be the Borel field of \(T \times R_0\) and let \(n(\text{d}r \text{d}a)\) be a \(\sigma\)-finite measure on \(T \times R_0\). A family of random variables \(N(B), B \in \mathcal{B}\) is called a Poisson random measure on \(T \times R_0\) with intensity measure \(n\) if

1. For every \(B\) such that \(0 < n(B) < \infty\), \(N(B)\) is Poisson distributed with mean \(n(B)\).
2. If \(B_1, \ldots, B_n\) are disjoint and \(0 < n(B_i) < \infty, i = 1, \ldots, n\), \(N(B_1), \ldots, N(B_n)\) are independent.
3. For every \(\omega\), \(N(\cdot, \omega)\) is a counting measure on \(T \times R_0\).

Now for a given Lévy process \(X_t\), we define a random counting measure \(N\) on \(T \times R_0\) by

\[
N((s, t] \times A) = \mathbb{P}[r \in (s, t]; \Delta X_r \in A],
\]

(1.1)

where \(\Delta X_r = X_r - X_r^-\). We set \(n(B) = E[N(B)]\) and define a compensated random measure by

\[
\tilde{N}(\text{d}s \text{d}z) = N(\text{d}s \text{d}z) - n(\text{d}s \text{d}z).
\]

(1.2)

**Theorem 1.1.** Let \(X_t\) be any Lévy process.

1. The counting measure \(N\) of (1.1) is a Poisson random measure on \(T \times R_0\) with intensity measure \(n\) satisfying

\[
\int_0^T \int_{R_0} \frac{|z|^2}{1+|z|^2} n(\text{d}s \text{d}z) < \infty.
\]

(2) For almost all \(\omega\),

\[
X^1_t = \int_0^t \int_{|z|>1} zN(\text{d}s \text{d}z) + \lim_{n \to \infty} \int_0^t \int_{\frac{1}{n} < |z| \leq 1} z\tilde{N}(\text{d}s \text{d}z),
\]

(1.3)

is defined for all \(t\) and the convergence is uniform with respect to \(t\).

3. \(W_t := X_t - X^1_t\) is a Brownian motion independent of the Poisson random measure \(N(\text{d}s \text{d}z)\).

The last term of (1.3) is denoted by \(\int_0^t \int_{0<|z| \leq 1} z\tilde{N}(\text{d}s \text{d}z)\).

Thus a Brownian motion and a Poisson random measure are the fundamental ingredients of any Lévy process, and furthermore the jumps of a Lévy process \(X_t\) are characterized by a Poisson counting measure \(N(\text{d}s \text{d}z)\). A surprising fact which impresses us is that the class of Lévy processes is large and rich, since the properties of Poisson random measures are quite different according to their intensity measures. Still now it is an attractive subject for research (Sato [37]).

A consequence of Theorem 1.1 is that the law of a Lévy process \(X_t\) is infinitely divisible and its characteristic function \(E[e^{i(\alpha, X_t)}]\) is given by the Lévy–Khintchine formula.

\[
\exp \left\{ i(\alpha, b(t)) - \frac{1}{2} (\alpha, A(t)\alpha) + \int_0^t \int_{|z|>1} (e^{i(\alpha, z)} - 1)n(\text{d}s \text{d}z) + \int_0^t \int_{0<|z| \leq 1} (e^{i(\alpha, z)} - 1 - i(\alpha, z))n(\text{d}s \text{d}z) \right\},
\]
where \( b(t) \) and \( A(t) \) are the mean and the covariance of the Brownian motion \( W_t \) and \( n(dsdz) \) is the intensity measure of the Poisson random measure \( N \). These satisfy the following conditions:

1. \( b(t) \) is continuous and \( b(0) = 0 \).
2. \( A(t) \) is continuous, increasing and \( A(0) = 0 \).
3. \( n(dsdz) \) is a measure on \( T_0 \times R_0 \) such that \( \int_0^T \int_{R_0} \frac{|z|^2}{1+|z|^2} n(dsdz) < \infty \) and \( n([t] \times R_0) = 0 \) for any \( t \).

The triplet \((b(t), A(t), n(drdz))\) is the characteristic triplet of the Lévy process. It determines the characteristic functions of \( X_t \) so that it determines the laws of \( X_t \).

A Lévy process \( X_t \) is homogeneous if and only if its characteristic triplet is given by \((bT, A, d\nu(dz))\) where \( b \), \( A \) are constants and \( \nu \) is a measure on \( R_0 \) satisfying \( \int \frac{|z|^2}{1+|z|^2} \nu(dz) < \infty \). The measure \( \nu \) is called Lévy measure.

The proof of Theorem 1.1 is by no means simple. A detailed exposition may be found in Itô’s lecture notes from Aarhus University [8]. Difficult parts of the proof would be: (a) \( N \) in (1.1) is a Poisson random measure, (b) Itô’s program for diffusions and jump-diffusions

1.2. Itô’s program for diffusions and jump–diffusions

After the work on Lévy processes, Itô’s interest moved towards Markov processes. At that time Kolmogorov’s work on diffusion processes [22] was known. He showed by an analytical method that the transition probabilities of a diffusion process \( X_t = (X_t^1, \ldots , X_t^d) \) on a Euclidean space satisfy a linear second order parabolic partial differential equation which is called Kolmogorov’s equation. The equation tells us that the coefficients \( a^{ij}(t, x) \) of the second order part and \( b^i(t, x) \) of the first order part of the equation are characterized by

\[
\begin{align*}
(a) \quad E[X_{t+h}^i - X_t^i | X_t = x] &= b^i(t, x)h + o(h), \\
(b) \quad E[(X_{t+h}^i - X_t^i)(X_{t+h}^j - X_t^j) | X_t = x] &= a^{ij}(t, x)h + o(h).
\end{align*}
\]

In words, the transition probabilities of a diffusion process should be determined by the above infinitesimal mean \( b(t, x) = (b^1(t, x), \ldots , b^d(t, x)) \) and the infinitesimal covariance \( A(t, x) = (a^{ij}(t, x)) \) of the diffusion process.

Again, Itô was interested in the sample paths of a diffusion process. His idea may be described as follows. The above formula indicates that for each point \((t, x)\), a Brownian motion with mean \( b(t, x) \) and covariance \( A(t, x) \) is tangent to the diffusion \( X_t \). The tangential Brownian motion at the point \((t, x)\) denoted by \( Z^{(t,x)}(\tau), \tau \geq t \) could be written as

\[
Z^{(t,x)}(\tau) = b(t, x)(\tau - t) + \sigma(t, x)(W_\tau - W_t),
\]

where \( \sigma(t, x) \) is the square root of \( A(t, x) \) and \( W_t \) is a standard Brownian motion. Then the paths of the diffusion process should be obtained by integrating the random tangential vector fields \( Z(x, dt) := Z^{(t,x)}(dt) \). Therefore the solution \( X_t \) of the stochastic differential equation

\[
dX_t = Z(X_t, dt) = b(t, X_t)dt + \sigma(t, X_t)dW_t
\]

should be a diffusion process satisfying (1.4).
To make the idea rigorous, he defined stochastic integrals $\int_0^t f_x dW_x$ based on the Brownian motion $W_t$ and then solved Eq. (1.5). Itô’s stochastic integral is not similar to the Lebesgue integral. He obtained a chain rule for stochastic integrals, which is now well known as Itô’s formula. Itô showed that the solution $X_t$ of Eq. (1.5) is a Markov process and its transition probabilities satisfy Kolmogorov’s forward equation.

Itô’s first work on this subject appeared during the second world war in 1942. It was a colloquium report at Osaka University, hand-written in Japanese. Later it was translated to English [9]. In the report, he defined the stochastic integral based on a Brownian motion and gave some formulas concerning the calculus of stochastic integrals, which differs from usual calculus. A formula given by him is:

$$\int_0^t F'(W_x) dW_x = F(W_t) - F(W_0) - \frac{1}{2} \int_0^t F''(W_x) dx,$$

where $W_t$ is a Brownian motion and $F$ is a $C^2$-function. It was the first version of Itô’s formula. Then he considered the stochastic differential (1.5). Assuming the Lipschitz continuity for the coefficients, he showed the existence and uniqueness of the solution and then the Markov property. Thus the solution is a diffusion process.

He was also interested in jump–diffusions. If $X_t$ is a jump–diffusion, the tangential process at $(t, x)$ should be a homogeneous Lévy process $Z^{(t, x)}(\tau), \tau \geq t$ with $(b(t, x), A(t, x), d\nu^{(t, x)}(dz))$ as its characteristics. It is known that for the family of Lévy measures $\nu^{(t, x)}(dz), (t, x) \in T \times \mathbb{R}^d$, there exists a common Lévy measure $\nu$ and a family of maps $g^{(t, x)} : R_0 \rightarrow R_0$, measurable with respect to $(t, x, z)$, such that $\nu^{(t, x)} = (g^{(t, x)})_* \nu$, i.e., $\nu^{(t, x)}(E) = \nu(\{z : g^{(t, x)}(z) \in E\})$ holds for any Borel set $E$. The maps $g^{(t, x)}$ satisfy $\int \frac{|g^{(t, x)}(z)|^2}{1 + |g^{(t, x)}(z)|^2} \nu(dz) < \infty$. Common Lévy measures are not unique, obviously. Itô took $\nu(dz) = \frac{1}{|z|^2} dz$ in the one-dimensional case and Stroock [38] took $\nu(dz) = \frac{1}{|z|^{d+1}} dz$ in the $d$-dimensional case.

In any case we can define a family of tangential homogeneous Lévy processes by a pair of a standard Brownian motion and a homogeneous Poisson random measure with Lévy measure $\nu(dz)$. Indeed, let $N(d\tau dz)$ be a homogeneous Poisson random measure with Lévy measure $\nu(dz)$, which is independent of a standard Brownian motion $W_t$. Set $N^{(t, x)} := (g^{(t, x)})_* N$. Then it is a homogeneous Poisson random measure with the Lévy measure $\nu^{(t, x)}(dz)$. It holds

$$\int_0^\tau \int_{|z'| \leq 1} z' \tilde{N}^{(t, x)}(dsdz') = \int_0^\tau \int_{|g^{(t, x)}| \leq 1} g^{(t, x)}(z) \tilde{N}(dsdz).$$

Therefore the family of homogeneous Lévy processes $Z^{(t, x)}(\tau), \tau \geq t$ with characteristics $(b(t, x), A(t, x), d\nu^{(t, x)}(dz))$ can be given by making use of a common Brownian motion $W_t$ and a common Poisson random measure $N(d\tau dz)$ as

$$Z^{(t, x)}(\tau) = b(t, x)(\tau - t) + \sigma(t, x)(W_\tau - W_t) + \int_t^\tau \int_{|g^{(t, x)}| \leq 1} g^{(t, x)}(z) \tilde{N}(dsdz) + \int_t^\tau \int_{|g^{(t, x)}| > 1} g^{(t, x)}(z) N(d\tau dz).$$

Then the jump–diffusion should be obtained by integrating the above random vector fields $Z(x, dt) := Z^{(t, x)}(dt)$, i.e., the solution of the stochastic integral equation

$$dX_t = b(t, X_t-)dt + \sigma(t, X_t-)dW_t$$
\[ + \int_{R_0} g_1(t, X_t, z) \tilde{N}(dr dz) + \int_{R_0} g_2(t, X_t, z) N(dr dz). \]  

Here \( g_1(t, x, z) = g(t, x, z)1_{|g(t,x,z)| \leq 1} \) and \( g_2(t, x, z) = g(t, x, z)1_{|g(t,x,z)| > 1} \), where \( g(t, x, z) = g^{t,x}(z) \).

The colloquium report was not a final one. Some discussions were insufficient. However, most of the work was completed and published later in English in a series of papers [10–14] and as a monograph [15].

### 1.3. Stochastic integrals based on Brownian motion and Itô’s formula

Stochastic integrals based on a Brownian motion are discussed in [10,13]. In the definition, Itô neither used terms such as filtration nor sub-\(\sigma\)-field nor martingale. Instead, he introduced the property \( \alpha \). Let \( \xi = \{\xi_x(t), \lambda \in \Lambda\} \) and \( \eta = \{\eta_x(t); \mu \in M\} \) be two families of stochastic processes. \( \xi \) is said to have the property \( \alpha \) with respect to \( \eta \), if for any \( t \) the following two families of random variables are independent:

\[ \{\xi_x(\tau), \lambda \in \Lambda, \eta_x(\tau), \mu \in M, 0 \leq \tau \leq t\}, \quad \{\eta_x(\sigma) - \eta_x(t); \mu \in M, t \leq \sigma\}. \]

Let \( W_t \) be a one-dimensional standard Brownian motion and \( \xi(t) \) be a measurable process, which has the property \( \alpha \) with respect to \( W_t \) and satisfies \( \int_0^t \xi(t)^2 \, dt < \infty \) a.s. for any \( t \). Itô defined the stochastic integral \( \int_0^t \xi(t) \, dW_t \). If \( \xi(t) \) is in \( L^2 \), i.e., \( E[\int_0^T \xi(t)^2 \, dt] < \infty \), his definition of the stochastic integral \( \int_0^t \xi(t) \, dW_t \) is exactly the same as that used nowadays. Indeed, if we introduce the filtration \( \{\mathcal{F}_t\} \) by \( \mathcal{F}_t = \sigma(\xi, W, 0 \leq t \leq t) \), then \( W_t \) is an \( \{\mathcal{F}_t\} \)-Brownian motion and \( \xi(t) \) is a measurable \( \{\mathcal{F}_t\} \)-adapted process. Itô’s stochastic integral is then a square-integrable \( \{\mathcal{F}_t\} \)-martingale.

Further, if \( \xi(t) \) does not satisfy the above square integrability condition, Itô introduces a sequence of processes \( \xi_n(t) \) by

\[ \xi_n(t) = n \left( \int_0^n \frac{\xi(t)^2 \, dt}{n} \right) \xi(t) \]

where \( \phi_n(\lambda) = 1 \) if \( |\lambda| \leq n \) and \( \phi_n(\lambda) = 0 \) if \( |\lambda| > n \). Then the stochastic integral for \( \xi \) was defined by

\[ \int_0^t \xi(t) \, dW_t = \lim_{n \to \infty} \int_0^t \xi_n(t) \, dW_t, \]

when the limit exists in probability uniformly in \( t \in [0, T] \). Later in 1965, Itô and Watanabe [18] introduced the notion of a local martingale. The above Itô stochastic integral is indeed a local martingale.

Now let \( W = (W_1, \ldots, W_r) \) be an \( r \)-dimensional standard Brownian motion. Suppose that the family of measurable processes

\[ \{\beta^i(t), \alpha^j(t), i = 1, \ldots, d, j = 1, \ldots, r\} \]

has the property \( \alpha \) with respect to the Brownian motion \( W_t \). Suppose furthermore: \( \int_0^T |\beta^i(t)| \, dt < \infty \), \( \int_0^T |\alpha^j(t)|^2 \, dt < \infty \). We may define a \( d \)-dimensional stochastic process \( X_t = (X_1(t), \ldots, X_r(t)) \).
by
\[ X_t^i - X_s^i = \int_s^t \beta^i(\tau) \, d\tau + \sum_j \int_s^t \alpha^i_j(\tau) \, dW^j_\tau, \quad 0 \leq s \leq t \leq T, \quad 1 \leq i \leq d. \]

It is simply written as
\[ dX_t = \beta(t) \, dt + \alpha(t) \, dW_t. \]

The stochastic process \( X_t \) is now called an \textit{Itô process}. In [13] he showed the celebrated \textit{Itô formula}. In [13] he showed the celebrated \textit{Itô formula}.

**Theorem 1.2.** Let \( F(t, x^1, \ldots, x^d) \) be a function of class \( C^{1,2} \). Then \( \eta(t) = F(t, X_t) \) satisfies
\[ d\eta_t = \left( F_t(t, X_t) + \sum_i F_{x_i}(t, X_t) \beta^i(t) \right) dt + \sum_{i,j} F_{x_i}(t, X_t) \alpha^i_j(t) dW^j_t, \]
where \( F_t, F_{x_i} \) are the partial derivatives of \( F \) with respect to \( t \) and \( x_i \), respectively and \( F_{x_i x_j} \) are the second order partial derivatives of \( F \) with respect to \( x_i, x_j \).

Itô’s formula is a change of variable formula or a chain rule for the calculus of stochastic integrals. It has been applied to many types of stochastic calculus. Another important value of Itô’s formula is that we may find an explicit form of the generator of a diffusion process through Itô’s formula. It will be discussed in the next subsection.

### 1.4. Stochastic differential equations

Stochastic differential equations (SDE) (1.5) are studied in [11,12,15,13,14]. A continuous stochastic process \( X_t \) with values in \( \mathbb{R}^d \) is called a solution of Eq. (1.5) if \( X_t \) has the property \( \alpha \) with respect to \( W_t \) and satisfies Eq. (1.5). Assuming that the coefficients \( b^i(t, x), \sigma^i_j(t, x) \) are of linear growth and Lipschitz continuous with respect to \( x \), Itô proved the existence and the uniqueness of the solution \( X_t \) by the method of successive approximations.

Let \( X_t^{(s,x)}, t \geq s \) be the solution starting from \( x \) at time \( s \). In view of the linear growth property of the coefficients, it has the following moment property: For any \( p \geq 2 \), there exists a positive constant \( C_p \) such that [23]
\[ E[|X_t^{(s,x)}|^p] \leq C_p(1 + |x|)^p, \quad 0 \leq s, t \leq T, \quad \forall x \in \mathbb{R}^d. \]

The solution is a Markov process. Set
\[ P_{s,t}(x, E) := P(X_t^{(s,x)} \in E), \quad P_{s,t}f(x) := \int P_{s,t}(x, dy) f(y). \]

Then \( P_{s,t}(x, E) \) are transition probabilities and satisfy the \textit{Chapman–Kolmogorov equation}
\[ P_{s,u}P_{u,t}f(x) = P_{s,t}f(x), \quad \forall s < t < u \]
for any measurable function of polynomial growth.
Let $f$ be a $C^2$ function such that $f, f_{x_i}, f_{x_ix_j}$ are of polynomial growth. Set $F = f$ and apply Itô's formula. Then we have

$$
f(X_t) = f(x) + \int_s^t \mathcal{L}(u) f(X_u)du + \sum_{i,k} \int_s^t f_{x_i}(u, X_u)\sigma^i_k(u, X_u)dW^k_u,
$$

where

$$
\mathcal{L}(t)f(x) = \frac{1}{2} \sum_{i,j} a^{ij}(t, x) f_{x_ix_j}(x) + \sum_i b^i(t, x) f_{x_i}(x),
$$

and $a^{ij}(t, x) = \sum_k \sigma^i_k(t, x)\sigma^j_k(t, x)$. Each of the above terms in (1.7) is integrable. The expectation of the last term of (1.7) is 0. Then we find that $P_{s,t}$ satisfies the equality

$$
P_{s,t}f = f + \int_s^t P_{s,u}\mathcal{L}(u)fdu.
$$

Differentiating each term of the above equality (1.9) with respect to $t$, we get Kolmogorov's forward equation

$$
\frac{\partial}{\partial t} P_{s,t}f(x) = P_{s,t}\mathcal{L}(t)f(x).
$$

The differential operator (1.8) is called the generator of the diffusion process.

Now, we will fix a point $x' \in \mathbb{R}^d$ and set $X_{t+h} = X_{t+h}(t, x')$. Eq. (1.9) implies

$$
\lim_{h \downarrow 0} \frac{1}{h} (E[f(X_{t+h}(t, x'))] - f(x')) = \mathcal{L}(t)f(x'),
$$

for any $C^2$ function $f$ of polynomial growth. Take $f(x) = x_i - x'_i$. Then $\mathcal{L}(t)f = b^i(t)$. Eq. (1.11) is written as

$$
\lim_{h \downarrow 0} \frac{1}{h} E[X_{t+h}^i - x'_i] = b^i(t, x').
$$

Take $f(x) = (x_i - x'_i)(x_j - x'_j)$. Then $\mathcal{L}(t)f = a^{ij}(t) + (x_j - x'_j)b^i(t) + (x_i - x'_i)b^j(t)$. Then (1.11) implies

$$
\lim_{h \downarrow 0} \frac{1}{h} E[(X_{t+h}^i - x'_i)(X_{t+h}^j - x'_j)] = a^{ij}(t, x').
$$

The Eqs. (1.12) and (1.13) are equivalent to Eq. (1.4). Hence the solution of the SDE is exactly what Itô wanted to construct.

In [12] and [14], he studied stochastic differential equations on a differentiable manifold. The equation is written in each local coordinate neighborhood using local coordinates. He explained clearly how the coefficients $b^i$ and $\sigma^i_j$ of the equation should be changed according to the change of the local coordinates. Then, assuming that the coefficients are uniformly bounded in any canonical coordinates, he showed the existence and the uniqueness of the solution.

In [15], stochastic differential equations with jumps are discussed. The equations can be written as in (1.6). Itô studied the one-dimensional equation in the case where the Lévy measure is given by $\frac{1}{\sqrt{t}}dz$. He defined stochastic integrals based on the Poisson random measure $N(dt, dz)$.
and the compensated Poisson random measure \( \tilde{N}(dtdz) \), and then proved the existence, the uniqueness and the Markov property of the solution, under some regularity condition for the coefficients \( b, \sigma, g_1, g_2 \). We will return to Eq. (1.6) in Section 2.

Itô’s aim of introducing stochastic differential equations was the construction of Markov diffusion processes. After the 1960s, stochastic differential equations and diffusion processes were studied in great detail. Topics include weak and strong solutions of SDE, SDE with boundary conditions, stochastic flow of diffeomorphisms, Malliavin calculus for SDE, etc. A survey of work on these topics is beyond the scope of this paper. We refer to Watanabe’s comprehensive paper [40] on these subjects. See also [19,36,39]. Further, a different approach using Dirichlet forms was taken by Fukushima et al. [4] for the study of diffusions and jump–diffusions with the symmetry property.

Around 1960, stochastic differential equations were applied to engineering problems. In stochastic control and filtering problems, SDEs were taken as a model of a dynamical system, which is disturbed by noise. A dynamical system is supposed to move according to a differential equation \( dX_t = b(t, X_t)dt \), where \( X = \{X_s; s \in T\} \). But it is actually disturbed by a random noise \( \sigma(t, X_t)dW_t \). Its motion is then described by a SDE

\[
dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t.\]

The coefficients \( b(t), \sigma(t) \) may depend on the whole past of the solution \( (X_s; s \leq t) \). Thus the solution is no longer a Markov process.

Later in the 1970s, the dynamic behavior of an asset price \( S_t \) was described by a SDE

\[
\frac{dS_t}{S_t} = b(t, S_t)dt + \sigma(t, S_t)dW_t.\]

Then Itô’s stochastic calculus encountered with new big objects, which Itô did not design and which he even might not have wanted to. It seems that he was not very much interested in mathematical finance and he was not involved in it, either.

In Section 3, we will see how Itô’s formula is applied for pricing options.

2. Reinforcement of Itô’s work with martingales

2.1. Stochastic integrals with respect to semimartingales and Itô’s formula

In the 1960s, Itô’s theory on stochastic calculus was reinforced for martingales. In combination with martingale theory, Itô’s stochastic calculus turned out to be a powerful tool not only for mathematical problems related to stochastic analysis but also for applications to engineering and financial problems.

In 1962–1963, Meyer [30,31] showed that any positive submartingale can be decomposed into the sum of a martingale and of a natural (predictable) increasing process, which is now called the Doob–Meyer decomposition. The decomposition theorem was then applied to the stochastic integrals based on martingales, and Itô’s formula was extended to a chain rule for semimartingales, by Kunita–Watanabe [26] and Meyer [32] in 1965–1967.

Let \((\Omega, \mathcal{F}, P)\) be a complete probability space. Let \(\{\mathcal{F}_t, t \in [0, T]\}\) be a family of sub-\(\sigma\)-fields of \(\mathcal{F}\). It is called a filtration if \(\mathcal{F}_s \subset \mathcal{F}_t\) holds for any \(s < t\). A filtration is said to satisfy the usual conditions if \(\mathcal{F}_0\) contains all null sets of \(\mathcal{F}\) and satisfies \(\mathcal{F}_t = \wedge_{\epsilon > 0} \mathcal{F}_{t+\epsilon}\) (right continuous). In this subsection, we will consider problems on a probability space equipped with a filtration \(\{\mathcal{F}_t\}\) with the usual conditions.
Filtrations play an important role in martingale theory and the theory of Markov processes. Further, in applications to engineering problems such as stochastic control and filtering or to mathematical finance, each $\mathcal{F}_t$ is understood as available data up to time $t$ of a random time evolution that we have been observing. Therefore the analysis of the filtration is important in applications, too.

A stochastic process $X_t$, $t \in [0, T]$ is said to be adapted (to $\{\mathcal{F}_t\}$) if $X_t$ is $\mathcal{F}_t$ measurable for any $t$. It is called progressively measurable if for every $t$, the map $X : [0, t] \times \Omega \ni (u, \omega) \rightarrow X(u, \omega)$ is $\mathcal{B}[0, t] \times \mathcal{F}_t$-measurable, where $\mathcal{B}[0, t]$ is the Borel field of $[0, t]$. Let $\mathcal{P}$ be the smallest $\sigma$-field on $[0, \infty) \times \Omega$ with respect to which all left continuous adapted processes are measurable. A stochastic process $X_t$ is said to be predictable if it is measurable with respect to $\mathcal{P}$.

A real valued adapted process $X_t$ is called a submartingale if it is integrable for any $t$ and satisfies $E[X_t | \mathcal{F}_s] \geq X_s$ for any $s < t$, where $E[X | \mathcal{G}]$ denotes the conditional expectation of the integrable random variable $X$ with respect to the sub-$\sigma$-field $\mathcal{G}$ of $\mathcal{F}$. If the equality holds for any $t, s$ with $t > s$, $X_t$ is called a martingale. Let $\tau$ be a nonnegative random variable. It is called a stopping time if $\{\tau \leq t\} \in \mathcal{F}_t$ holds for any $t > 0$.

Doob studied martingales in detail. The beautiful classical theory of martingales is found in the books [3,5,35]. Doob showed that any martingale has a modification such that its sample paths are right continuous with left-hand limits. In our discussion we will always consider such a modification.

In [3], Doob pointed out that Itô’s stochastic integral can be extended to certain martingales. Let $M_t$ be a square-integrable martingale. Suppose that there exists a right continuous (deterministic) increasing function $F$ satisfying

$$E[(M_t - M_s)^2 | \mathcal{F}_s] = F(t) - F(s), \quad \forall s < t. \quad (2.1)$$

Let $\phi(t)$ be a progressively measurable process with $E[\int_0^T |\phi(t)|^2 dF(t)] < \infty$. Doob showed that the stochastic integral $\mathcal{N}_t = \int_0^t \phi(s) dM_s$ is well defined for any $t$ and it is a square-integrable martingale.

If $M_t$ is a square-integrable Lévy process with mean 0, it is a martingale satisfying (2.1), where $F(t) = at$ is the variance of $M_t$. We will give later an example of a square-integrable martingale which is not a Lévy process but satisfies (2.1) with a suitable $F$. However, if $M_t$ is not a process with independent increments, we cannot expect in general that there exists a function $F$ satisfying (2.1).

Meyer [30,31] showed that if $X_t$ is a positive submartingale there exists a unique predictable increasing process $A_t$ such that $X_t - A_t$ is a martingale. Now let $M_t$ be a square-integrable martingale. Then $M^2_t$ is a submartingale. We denote the associated $\{\mathcal{F}_t\}$-predictable increasing process $A_t$ by $\langle M \rangle_t$. Then the martingale $M_t$ satisfies the equality

$$E[(M_t - M_s)^2 | \mathcal{F}_s] = E[\langle M \rangle_t - \langle M \rangle_s | \mathcal{F}_s], \quad \forall s < t. \quad (2.2)$$

A similar equality had been known for additive functionals of a Markov process $\xi_t$. Here a process $X_t$ is called an additive functional if it satisfies the equality $X_{t+s} = X_t + X_t \circ \theta_s$ for any $s, t > 0$ a.s., where $\theta_s$ is the shift of $\xi_t$ such that $\xi_t \circ \theta_s = \xi_{s+t}$ holds for any $t$. It tells us that if $M_t$ is a square-integrable martingale additive functional of a Markov process $\xi_t$, there exists a unique continuous increasing additive functional $\langle M \rangle_t$ satisfying (2.2). Using this equality, Motoo–Watanabe [33] defined the stochastic integral $\int_0^t \phi(\xi_s) dM_s$ as a continuous martingale additive functional. The integral was then extended to arbitrary square-integrable martingales in [26].
Following [26,32], we shall define the stochastic integral based on $M_t$ for a predictable process $\phi_t$ whose norm
\[
\|\phi\|_{(M)} := E\left[\int_0^T |\phi(t)|^2 \, d\langle M \rangle_t\right]^{1/2}
\]
is finite. Let $\phi(t) = \sum_{i=1}^m \phi_i 1_{(t_{i-1}, t_i]}(t)$ be a simple process such that each $\phi_i$ is bounded $\mathcal{F}_{t_{i-1}}$-measurable, where $0 = t_0 < t_1 < \cdots < t_m = T$. Then the stochastic integral of $\phi$ with respect to $dM_t$ is defined by
\[
\int_0^t \phi \, dM := \sum_{i=1}^m \phi_i (M_{t \wedge t_i} - M_{t \wedge t_{i-1}}).
\]
It is a square-integrable martingale and satisfies $E[\int_0^T \phi \, dM]^2 = \|\phi\|^2_{(M)}$. Next let $\phi_t$ be a predictable process such that $\|\phi\|_{(M)} < \infty$. Then there exists a sequence of simple processes $\phi^{(n)}_t$ such that $\|\phi - \phi^{(n)}\|_{(M)} \to 0$. Then the sequence of stochastic integrals $\{\int_0^t \phi^{(n)} \, dM\}$ is an $L^2(P)$-Cauchy sequence. The limit is a square-integrable martingale. We denote it by $\int_0^t \phi(s) \, dM_s$ or $\int_0^t \phi \, dM$.

Local martingales were defined by Itô–Watanabe [18]. Let $X_t$ be a right continuous adapted process. It is called a local (or locally square-integrable) martingale if there exists an increasing sequence of stopping times $\tau_n$ such that $\tau_n \to \infty$ and the stopped processes $M_{t \wedge \tau_n}$ are (square-integrable) martingales. For a locally square-integrable martingale $M_t$, there exists a unique predictable increasing process $\langle M \rangle_t$ such that $M_t^2 - \langle M \rangle_t$ is a local martingale.

Stochastic integrals are extended to a locally square-integrable martingale $M_t$ and a predictable process $\phi(t)$ such that $\int_0^T |\phi(s)|^2 \, d\langle M \rangle_s < \infty$. In fact, there exists a locally square-integrable martingale $I_t(\phi)$ such that the equality $I_t(\phi) = \int_0^t \phi \, dM^T$ holds for any stopping time $\tau$ such that $M_\tau^T := M_{t \wedge \tau}$ is a square-integrable martingale and satisfies $E[\int_0^\tau (1 + \phi(s)^2) \, d\langle M \rangle_s] < \infty$. We denote such $I_t(\phi)$ by $\int_0^t \phi \, dM$.

In the following, a filtration $\{\mathcal{F}_t\}$ is fixed. The $\langle \mathcal{F}_t \rangle$-martingales are simply called martingales.

A right continuous adapted process $A_t$ is called a process of finite variation if it is written as a difference of two adapted increasing processes. A right continuous adapted process $X_t$ is called a semimartingale if it is written as a sum of a locally square-integrable martingale $M_t$ and a process of finite variation $A_t$. It is called a continuous semimartingale if both $A_t$ and $M_t$ are continuous processes. Let $\phi(t)$ be a predictable process such that $\int_0^T \phi(s) \, dA_s$ (Lebesgue integral) and $\int_0^T \phi(s) \, dM_s$ are well defined a.s. The sum of these two integrals is denoted by $\int_0^T \phi(s) \, dX_s$ or simply as $\int_0^T \phi \, dX$. It is again a semimartingale.

The bracket process for two square-integrable martingales was introduced in [33] for the study of additive functionals of a Markov process. The authors defined the orthogonality of square-integrable martingales and discussed the orthogonal decomposition of a martingale. We will recall their arguments. Let $M_t$ and $N_t$ be locally square-integrable martingales. Their bracket process is defined by $\langle M, N \rangle_t = \frac{1}{4} \{\langle M + N \rangle_t - \langle M - N \rangle_t\}$. Then we have the relation
\[
\left(\int_0^T \phi \, dX, \int_0^T \phi \, dY\right)_t = \int_0^T \phi \, d\langle X, Y \rangle_s.
\]

Two locally square-integrable martingales $M, N$ are called orthogonal if $\langle M, N \rangle_t = 0$ holds for any $t$ a.s. We denote by $\mathcal{M}_{loc}^2$ the set of all continuous locally square-integrable martingales.
and \( \mathcal{M}^d_{loc} \), the set of all locally square-integrable martingales which are orthogonal to any element of \( \mathcal{M}^c_{loc} \). Elements of \( \mathcal{M}^d_{loc} \) are called purely discontinuous. For any locally square-integrable martingale \( M_t \), there exist \( M^c_t \in \mathcal{M}^c_{loc} \) and \( M^d_t \in \mathcal{M}^d_{loc} \), and \( M_t \) is decomposed as \( M_t = M^c_t + M^d_t \). Such a decomposition is unique.

Let \( X_t \) be a semimartingale. Then its sample paths are right continuous with left-hand limits. We denote the left limits of \( X_t \) by \( X_{t-} \). Then \( X_{t-} \) is a left continuous predictable process. For a semimartingale \( X_t \), we may define the stochastic integral \( \int_0^t X_s dX_s \). Another bracket process \([X]_t \) was introduced in [32]. It is defined by \([X]_t = X_t^2 - 2 \int_0^t X_s dX_s \). We define \([X, Y] = \frac{1}{2}([X + Y] - [X - Y]) \).

Let \( X, Y \) be semimartingales. They are decomposed as \( X_t = M^c_t + M^d_t + A_t \) and \( Y_t = N^c_t + N^d_t + B_t \), where \( M^c, N^c, M^d, N^d \in \mathcal{M}^c_{loc}, M^d, N^d \in \mathcal{M}^d_{loc} \) and \( A_t, B_t \) are processes of finite variation. We have the formula

\[
[X, Y]_t = (M^c, N^c)_t + \sum_{s \leq t} (\Delta X_s)(\Delta Y_s).
\]

(2.4)

Hence \([X, Y]_t \) is a process of finite variation. Its continuous part \([X, Y]_t^c \) coincides with \((M^c, N^c)_t \).

Itô’s formula can be extended to semimartingales. A \( d \)-dimensional stochastic process \( X_t = (X^1_t, \ldots, X^d_t) \) is called a semimartingale if all its components \( X^i_t, i = 1, \ldots, d \) are semimartingales. When a semimartingale has jumps, several versions of the Itô formula are used. The following is due to Meyer [32,34].

**Theorem 2.1.** Let \( X_t = (X^1_t, \ldots, X^d_t) \) be a semimartingale. Let \( F(t, x) \) be a function of the \( C^{1,2} \)-class. Then \( F(t, X_t) \) is again a semimartingale and the following formula holds:

\[
F(t, X_t) = F(0, X_0) + \int_0^t F_t(s, X_s^-) ds + \sum_i \int_0^t F_{x_i}(s, X_s^-) dX^i_s \\
+ \frac{1}{2} \sum_{i,j} \int_0^t F_{x_ix_j}(s, X_s^-) d[X^i, X^j]_s \\
+ \sum_{0 \leq s \leq t} \left\{ F(s, X_s) - F(s, X_s^-) - \sum_i F_{x_i}(s, X_s^-) \Delta X^i_s \right\}.
\]

(2.5)

**2.2. Stochastic integrals based on compensated Poisson random measure**

Let \((\Omega, \mathcal{F}, P)\) be a probability space, where a standard \( m \)-dimensional Brownian motion \( W_t = (W^1_t, \ldots, W^m_t) \) and a homogeneous Poisson random measure \( N(drdz) \) on \( T_0 \times R_0 \) are defined. \( T_0 = [0, T] - \{0\} \) and \( R_0 = R^d - \{0\} \). These two processes are assumed to be independent. Let \( B \) be the Borel field of \( R_0 \). For \( B \in \mathcal{B} \), set \( N_t(B) = N((0, t] \times B) \). If it is integrable for any \( t \), it is a Poisson process. Its mean is written as \( tv(B) \) for a measure \( \nu \), called the Lévy measure. Let \( \{\mathcal{F}_t\} \) be the smallest filtration such that for any \( t \), the family of random variables \( W_s, N_s(B), B \in \mathcal{B}, s \leq t \) are measurable with respect to \( \mathcal{F}_t \). In Sections 2.2–2.4, we will consider problems under this filtration.

Let \( \tilde{N}(dz) = N(dz) - d\nu(dz) \) be the compensated Poisson random measure. Suppose \( \nu(B) < \infty \). Then \( \tilde{N}_t(B) = \tilde{N}((0, t] \times B) = N_t(B) - tv(B) \) is a compensated Poisson
process. Its mean is 0 and its covariance $tv(B)$. Hence it is a square-integrable martingale with 
$(\tilde{N}(B))_t = tv(B)$.

We shall define a stochastic integral based on the compensated Poisson random measure. A
random function $\psi(t, z, \omega), (t, z, \omega) \in T_0 \times R \times \Omega$ is called a predictable process if it is $\mathcal{P} \times B$-
measurable. We shall define a stochastic integral of the form $\int_0^t \int_{R_0} \psi(s, z) \tilde{N}(dsdz)$ or simply $\int_0^t \int \psi d\tilde{N}$ for such $\psi$. We first introduce some functional spaces:

$$
\psi_{loc}^2 = \left\{ \psi(t, z) \text{ predictable } | \int_0^T \int_{R_0} |\psi(t, z)|^2 d\nu(dz) < \infty \right\},
$$

$$
\psi^2 = \left\{ \psi(t, z) \in \psi_{loc}^2 \text{ | } E\left[ \int_0^T \int_{R_0} |\psi(t, z)|^2 d\nu(dz) \right] < \infty \right\},
$$

$$
\psi_{loc} = \left\{ \psi(t, z) \text{ predictable } | \int_0^T \int_{R_0} \frac{|\psi(t, z)|^2}{1 + |\psi(t, z)|} d\nu(dz) < \infty \right\},
$$

$$
\psi = \left\{ \psi(t, z) \in \psi_{loc} \text{ | } E\left[ \int_0^T \int_{R_0} \frac{|\psi(t, z)|^2}{1 + |\psi(t, z)|} d\nu(dz) \right] < \infty \right\}.
$$

We denote by $\psi_{loc}^1$ the set of all $\psi$’s such that $|\psi|^{1/2} \in \psi_{loc}^2$. $\psi^1$ is defined similarly. Then all
the above spaces are vector spaces.

It holds $\psi^1 \cup \psi^2 \subset \psi$ and $\psi_{loc}^1 \cup \psi_{loc}^2 \subset \psi_{loc}$. Further, a predictable process $\psi$ belongs to $\psi$ (or $\psi_{loc}$) if and only if $\psi_1 := \psi_{|\psi|>1} \in \psi^1$ (or $\psi_{loc}^1$) and $\psi_2 := \psi_{|\psi|\leq 1} \in \psi^2$ (or $\psi_{loc}^2$).

If $\nu$ is a finite measure, we have $\psi^2 \subset \psi^1 = \psi$ and $\psi_{loc}^1 \subset \psi_{loc}^2 = \psi_{loc}$.

We set

$$
\int_{R_0} \psi(z) \tilde{N}((s, t], dz) := \int_{R_0} \psi(z) \tilde{N}_t(dz) - \int_{R_0} \psi(z) \tilde{N}_s(dz).
$$

**Lemma 2.2.** Let $\psi(z)$ be a $\mathcal{F}_s \times B$-measurable random variable such that $E[\int_{R_0} \psi(z)^2 \nu(dz)] < \infty$. Then it holds for any $s < t$,

$$
E\left[ \int_{R_0} \psi(z) \tilde{N}((s, t], dz) \bigg| \mathcal{F}_s \right] = 0, \text{ a.s.} \quad (2.6)
$$

$$
E\left[ \left( \int_{R_0} \psi(z) \tilde{N}((s, t], dz) \right)^2 \bigg| \mathcal{F}_s \right] = (t-s) \int_{R_0} \psi(z)^2 \nu(dz), \text{ a.s.} \quad (2.7)
$$

**Proof.** Since $\tilde{N}((s, t], dz)$ is independent of $\mathcal{F}_s$, we have

$$
E\left[ \int_{R_0} \psi(z) \tilde{N}((s, t], dz) \bigg| \mathcal{F}_s \right] = \int_{R_0} \psi(z) E[\tilde{N}((s, t], dz)] = 0, \text{ a.s.}
$$

Suppose that $\psi(z)$ is a simple function written as $\sum_{i=1}^n c_i \chi_{A_i}$, where $A_i, i = 1, \ldots, n$ are disjoint
Borel sets in $R_0$. Then

$$
E\left[ \left( \int_{R_0} \psi(z) \tilde{N}((s, t], dz) \right)^2 \bigg| \mathcal{F}_s \right] = \sum_{i,j} c_ic_j E[\tilde{N}((s, t] \times A_i) \tilde{N}((s, t] \times A_j)|\mathcal{F}_s].
$$

Note that $\tilde{N}((s, t] \times A_i), i = 1, \ldots, n$ are Poisson random variables with respective intensities
$(t-s)\nu(A_i)$, and they are mutually independent and are also independent of $\mathcal{F}_s$. Then the above
is equal to $\sum_i c_i^2(t-s)v(A_i) = (t-s) \int \psi^2 \nu(dz)$, showing (2.7). The equality (2.7) is extended to arbitrary square integrable $\psi$. \hfill \Box

Now let $\psi(t, z) = \sum \psi_i(z)1_{(t_{i-1}, t_i)}(t)$ be a simple predictable process, where $\psi_i$ are $\mathcal{F}_{t_{i-1}} \times \mathcal{B}$-measurable functions with $E[\int |\psi_i(z)|^2 \nu(dz)] < \infty$. The stochastic integral of $\psi$ with respect to $d\tilde{N}$ is defined by

$$
\int_0^t \psi \circ d\tilde{N} := \sum_i \int \psi_i(z) \tilde{N}((t_{i-1}, t_i], dz).
$$

It is a square-integrable martingale and the equality

$$
\left\{ \int \int \psi \circ d\tilde{N} \right\}_t = \int_0^t \int_{\mathbb{R}_0} \psi(s, z)^2 ds \nu(dz)
$$

holds in view of the above lemma.

For $\psi(t, z) \in \psi^2$, there exists a sequence $\psi_n(t, x)$ of simple predictable processes such that $E[\int_0^T \int_{\mathbb{R}_0} |\psi - \psi_n|^2 d\nu(dz)] \to 0$. Then the sequence of stochastic integrals $\int_0^t \psi d\tilde{N}$ converges to a square-integrable martingale $M_t$. We denote it by $\int_0^t \psi d\tilde{N}$.

The stochastic integral can be extended for $\psi \in \psi^2_{\text{loc}}$ as a locally square-integrable martingale, which we denote by $\int_0^t \psi d\tilde{N}$. Next let $\psi \in \psi_{\text{loc}}$. Set $\psi_1 = \psi_{1, \psi > 1}$ and $\psi_2 = \psi_{1, \psi \leq 1}$. Then $\psi_2 \in \psi^2_{\text{loc}}$ and the integral $\int_0^t \psi_2 d\tilde{N}$ is well defined as a locally square-integrable martingale. For $\psi_1$, both integrals $\int_0^t \psi_1 d\tilde{N}$ and $\int_0^t \nu_1 d\nu$ are well defined. Denote the difference of the above two integrals by $\int_0^t \psi_1 d\tilde{N}$. It is a local martingale. We define the stochastic integral $\int_0^t \psi d\tilde{N}$ as the sum of the two integrals $\int_0^t \psi_1 d\tilde{N}, i = 1, 2$. It is a local martingale.

If $\psi \in \Psi$, then $\psi_2 \in \psi^2$. The stochastic integral of $\psi_2$ is a square-integrable martingale. Further since $\psi_1 \in \psi^1$, $\int_0^t \psi_1 d\tilde{N}$ is a martingale. Hence the stochastic integral $\int_0^t \psi d\tilde{N}$ is a martingale.

It is convenient to introduce the Lévy process

$$
J_t = \int_0^t \int_{0 < |z| \leq 1} z \tilde{N}(dsd|z|) + \int_0^t \int_{|z| > 1} z \tilde{N}(dsd|z|).
$$

Then we may rewrite the Poisson random measure as the counting measure of the jumps of $J_t$. We have

$$
\int_0^t \int \psi d\tilde{N} = \sum_{s \leq t} \psi(s, \Delta J_s).
$$

We summarize the above theory on stochastic integrals based on compensated Poisson random measure.

**Proposition 2.3.** Suppose $\psi \in \psi_{\text{loc}}$ (or $\Psi$). Then the stochastic integral $\int_0^t \psi d\tilde{N}$ is defined as a local martingale (or martingale). It holds

$$
\int_0^t \psi d\tilde{N} = \lim_{\epsilon \downarrow 0} \left\{ \sum_{s \leq t, |\Delta J_s| > \epsilon} \psi(s, \Delta J_s) - \int_0^t \int_{|z| > \epsilon} \psi(s, z) ds \nu(dz) \right\}.
$$

If $\psi \in \psi^2_{\text{loc}}$ (or $\psi^2$), $\int_0^t \psi d\tilde{N}$ is a locally square-integrable (or square-integrable) martingale.
Now, for vector processes $\phi(t) = (\phi^1(t), \ldots, \phi^m(t))$, we introduce
\[
\phi^2_{loc} = \left\{ \phi(t) \text{ predictable} \mid \int_0^T |\phi(t)|^2 dt < \infty \right\},
\]
\[
\phi^2 = \left\{ \phi(t) \in \Phi^2_{loc} \mid E\left[ \int_0^T |\phi(t)|^2 dt \right] < \infty \right\}.
\]
The stochastic integrals $\int_0^t \phi^j dW_t^j$ are well defined. These are locally square martingales if $\phi \in \Phi^2_{loc}$ and are square-integrable martingales if $\phi \in \Phi^2$.

**Proposition 2.4.** For $\phi, \phi' \in \Phi^2_{loc}$ and $\psi, \psi' \in \Psi^2_{loc}$, we set
\[
M_t = \sum_j \int_0^t \phi^j dW_t^j + \int_0^t \int \psi d\tilde{N}, \quad M'_t = \sum_j \int_0^t \phi'^j dW_t^j + \int_0^t \int \psi' d\tilde{N}.
\]
These are locally square-integrable martingales and
\[
\langle M, M' \rangle_t = \sum_j \int_0^t \phi^j \phi'^j ds + \int_0^t \int \psi \psi' ds dv,
\]
\[
[M, M']_t = \sum_j \int_0^t \phi^j \phi'^j ds + \int_0^t \int \psi \psi' dN.
\]

2.3. **Itô’s formula, martingale representation theorem and Girsanov’s theorem**

We give three theorems, which are useful in applications.

Let $X_t = (X_1^i, \ldots, X_i^d)$ be a semimartingale such that
\[
X_s^i - X_s^j = \int_s^t \beta^i(\tau) d\tau + \sum_j \int_s^t \alpha^j_i(\tau) dW_t^j + \int_s^t \int_{R_0^d} \gamma^i(\tau, z) d\tilde{N}(d\tau dz),
\]
where $\beta(t) \in \Phi^2_{loc}$ and $\gamma(t, z) \in \Psi_{loc}$. We call it an Itô jump process. Itô’s formula (Theorem 2.1) may be rewritten for Itô jump processes. The following is another version of Itô’s formula due to [26,5]. The formula is useful for the study of jump–diffusions.

**Theorem 2.5.** Let $X_t = (X_1^i, \ldots, X_i^d)$ be an Itô jump process represented above. Let $F(t, x_1, \ldots, x_d)$ be a $C^{1,2}$-function such that $\frac{F}{1 + |x|}$ is bounded. Then $\eta_t = F(t, X_t)$ is an Itô jump process and is decomposed as
\[
d\eta_t = \left[ F_t(t, X_t) + \sum_i F_{x_i}(t, X_t) \beta^i(t) + \frac{1}{2} \sum_{i,j} F_{x_i x_j}(t, X_t) \left( \sum_k \alpha^i_k(t) \alpha^j_k(t) \right) \right] dt
\]
\[
+ \int_{R_0^d} \left\{ F(t, X_t + \gamma(t, z)) - F(t, X_t) - \sum_i \gamma^i(t, z) F_{x_i}(X_t) \right\} d\tilde{N}(dz) dt
\]
\[
+ \sum_{i,j} F_{x_i}(t, X_t) \alpha^j_i(t) dW_t^j + \int_{R_0^d} \{ F(t, X_t + \gamma(t, z)) - F(t, X_t) \} d\tilde{N}(dz).
\]
Proof. In Theorem 2.1, we may rewrite

\[ F_{x_i}(t, X_{t-})dX^i_t = F_{x_i}(t, X_{t-})\beta^i(t)dt + \sum_j F_{x_i}(t, X_{t-})\alpha^j(t)dW^j_t + \int F_{x_i}(t, X_{t-})\gamma^i(t, z)\tilde{N}(drdz), \]

\[ F_{x_ix_j}(t, X_{t-})d[X^i, X^j]_t = F_{x_ix_j}(t, X_{t-})(\sum_k \alpha_k^i(t)\alpha_k^j(t))dt, \]

because

\[ [X^i, X^j]_t = \sum_{k,l} \left( \int \alpha_k^i dW^k, \int \alpha_l^j dW^l \right)_t = \sum_{k,l} \int_0^t \alpha_k^i \alpha_l^j d\langle W^k, W^l \rangle_s = \sum_k \int_0^t \alpha_k^i \alpha_k^j ds. \]

Next, checking that

\[ F(t, X_{t-} + \gamma(t, z)) - F(t, X_{t-}) \in \Psi^2_{loc}, \quad \sum_i \gamma^i(t, z)F_{x_i}(X_{t-}) \in \Psi^1_{loc}, \]

\[ G(t, z) := F(t, X_{t-} + \gamma(t, z)) - F(t, X_{t-}) - \sum_i \gamma^i(t, z)F_{x_i}(X_{t-}) \in \Psi^1_{loc}, \]

and noting that \( \Delta X_t = \gamma(t, \Delta J_t) \), we have by Proposition 2.3,

\[ \sum_{0 \leq s \leq t} \left\{ F(s, X_s) - F(s, X_{s-}) - \sum_j F_{x_j}(s, X_{s-})\Delta X^j_s \right\} \]

\[ = \int_0^t \int G(s, z)\tilde{N}(dsdz) + \int_0^t \left( \int G(s, z)v(dz) \right)ds \]

\[ = \int_0^t \int \{ F(s, X_{s-} + \gamma(s, z)) - F(s, X_{s-}) \} \tilde{N}(dsdz) \]

\[ - \sum_i \int_0^t \int F_{x_i}(s, X_{s-})\gamma^i(s, z)\tilde{N}(dsdz) + \int_0^t \left( \int G(s, z)v(dz) \right)ds. \]

Therefore we get the formula. \( \square \)

Associated with the filtration \( \{ \mathcal{F}_t \} \) generated by the Brownian motion \( W_t \) and Poisson random measure \( N(dr dz) \), we have the martingale representation theorem

**Theorem 2.6.** ([26,24]) Let \( M_t \) be a locally square-integrable martingale. Then there exists \( \phi(t) \in \Phi^2_{loc} \) and \( \psi(t, z) \in \Psi^2_{loc} \) and \( M_t \) is represented by

\[ M_t - M_0 = \sum_i \int_0^t \phi_i dW^i + \int_0^t \psi \tilde{N}. \quad (2.12) \]

Martingale representations are also known for processes which are not Lévy processes. Here we shall consider one such martingale. Let \( W_t \) be a one-dimensional standard Brownian motion and let \( \gamma_t = \sup \{ s \leq t; W_s = 0 \} \). We define \( \mu_t := \text{sgn}(W_t)\sqrt{t - \gamma_t} \). Then \( \mu_t \) is a purely discontinuous martingale, called the Azéma martingale. It holds \( \langle \mu \rangle_t = t/2 \). Hence the martingale \( \mu_t \) has the property (2.1) with \( F(t) = t/2 \). For the proof of the theorem, Itô’s work [16,17] is used. Let \( \{ \mathcal{S}_t \} \) be the filtration generated by \( \mu_t \). Then for any square-integrable
If \( (2.14) \) and let \( (2.13) \)

Conversely suppose we are given a pair of theorem for Itô jump processes.

Then \( Q \)

\[ Q(\Omega) = \int_{0}^{T} \phi(t)^{2} dt < \infty \]

and \( M_{t} \) is represented as

\[ M_{t} = M_{0} + \int_{0}^{t} \phi d\mu. \]

For details we refer to Mansuy–Yor [28].

Now, let \( \alpha_{t} \) be a local martingale such that \( \alpha_{t} > 0 \) a.s. for any \( t \). It is called a positive local martingale. We assume \( \alpha_{0} = 1 \).

(1) \((24)\) We may represent it as the solution of a linear SDE

\[ d\alpha_{t} = \alpha_{t-} dZ_{t}, \quad (2.13) \]

where \( Z_{t} \) is a local martingale represented as

\[ Z_{t} = \sum_{i} \int_{0}^{t} \phi_{i} dW_{i} + \int_{0}^{t} \psi d\tilde{N}, \quad (2.14) \]

with \( \phi \in \Phi_{loc}^{2} \) and \( \psi \in \Psi_{loc} \) such that \( 1 + \psi(t, z) > 0 \) for any \( t, z \).

(2) Conversely suppose we are given a pair of \( \phi \in \Phi_{loc}^{2} \) and \( \psi \in \Psi_{loc} \) such that \( \psi(t, z) + 1 > 0 \) for any \( t, z \). Define \( Z_{t} \) by \((2.14)\) and let \( \alpha_{t} = \alpha_{t}(\phi, \psi) \) be the solution of \((2.13)\). Then \( \alpha_{t} \) is a positive local martingale.

(3) If \( \phi(t) \) and \( \int |\psi(t, z)|^{2} v(dz) \) are bounded processes, then \( \alpha_{t} \) is a martingale. This fact is shown as follows. Set \( \tau_{n} = \inf\{0 < t < T; \alpha_{t} \geq n\} (= \infty \) if \( \{\cdots\} = \emptyset \). Then \( \tau_{n}, n = 1, 2, \ldots \) is an increasing sequence of stopping times such that \( \tau_{n} \to \infty \) a.s. The stopped processes \( \alpha_{t}^{(n)} = \alpha_{t \wedge \tau_{n}} \) satisfy \( \alpha_{t}^{(n)} - 1 = \int_{0}^{t \wedge \tau_{n}} \alpha_{s}^{(n)} dZ_{s} \). The stochastic integral in the right-hand side is a square-integrable martingale, because its bracket process satisfies

\[ E\left[ \int \alpha_{s}^{(n)} dZ_{s} \right]_{t \wedge \tau_{n}} = E\left[ \int_{0}^{t \wedge \tau_{n}} \alpha_{s}^{(n)} d\langle Z \rangle_{s} \right] \]

\[ = E\left[ \int_{0}^{t \wedge \tau_{n}} \alpha_{s}^{(n)} \phi(s)^{2} + \int \psi(s, z)^{2} v(dz) \right] ds \]

\[ \leq nKt < \infty, \]

where \( K = \sup_{s, \alpha_{t}} (|\phi(s)|^{2} + \int \psi(s, z)^{2} v(dz)) \). Therefore \( \alpha_{t}^{(n)} \) is a square-integrable martingale. Further, the above computation yields another inequality

\[ E[(\alpha_{t}^{(n)} - 1)^{2}] = E\left[ \int \alpha_{s}^{(n)} dZ_{s} \right]_{t \wedge \tau_{n}} \leq K \int_{0}^{t} E[(\alpha_{s}^{(n)})^{2}] ds. \]

Therefore \( \sup_{t < T} E[(\alpha_{t}^{(n)})^{2}] < \infty. \) Then for any \( 0 \leq t \leq T \), the sequence of random variables \( \{\alpha_{t}^{(n)}\} \) is uniformly integrable and hence it converges to \( \alpha_{t} \) in \( L^{1} \). Therefore \( \alpha_{t} \) is a martingale.

Now assume \( \alpha_{t} = \alpha_{t}(\phi, \psi) \) is a martingale. Define

\[ Q(B) = E[\alpha_{T} 1_{B}], \quad \forall B \in \mathcal{F}_{T}. \quad (2.15) \]

Then \( Q \) is a probability measure on \((\Omega, \mathcal{F}_{T})\). We denote it by \( \alpha T \cdot P \).

Since \( Q \) is equivalent (mutually absolutely continuous) to \( P \), a stochastic process on \((\Omega, \mathcal{F}_{T}, P)\) can be regarded as a stochastic process on \((\Omega, \mathcal{F}_{T}, Q)\). The following is a Girsanov theorem for Itô jump processes.
**Theorem 2.7.** ([24]) With respect to $Q$, we have

1. $W_t^\psi := W_t - \int_0^t \phi(s)ds$ is a standard Brownian motion.
2. The compensator of $N$ is $n^\psi (dxdz) = (1 + \psi(s, z))dsv(dz)$, i.e.,

$$
\bar{N}_t^\psi (g) := \int_0^t \int_{R_x} g(s, z)\{N(dxdz) - n^\psi (dxdz)\}
$$

is a local martingale, if $g(s, z)$ is square integrable with respect to $n^\psi (dxdz)$ a.s.

### 2.4. SDE with jumps and jump–diffusions

In Section 1 we mentioned that Itô constructed a jump–diffusion by solving an SDE. His idea was to integrate tangential Lévy processes $Z(x, dt)$, which consist of a Brownian motion $W_t$ and of a Poisson random measure $\bar{N}(dtdz)$. In this subsection, we shall study his approach in detail.

Let us consider Eq. (1.6). Assume that $b, \sigma$ are of linear growth and Lipschitz continuous. Assume further that $g$ is of linear growth and Lipschitz continuous with respect to $\nu$,

$$
\begin{align*}
|g(t, x, z)| & \leq K(z)(1 + |x|), \\
|g(t, x, z) - g(t, y, z)| & \leq L(z)|x - y|,
\end{align*}
$$

where $f(K(z)^2 + L(z)^2)\nu(dz) < \infty$. Then we may rewrite Eq. (1.6) simply as

$$
dX_t = b(t, X_{t-})dt + \sigma(t, X_{t-})dW_t + \int_{R_x} g(t, X_{t-}, z)\bar{N}(dtdz),
$$

where the drift vector $b(t, x)$ of the above equation is changed from the one in Eq. (1.6) by adding the term $\int_{|g(t, x, z)| > 1} g(t, x, z)\nu(dz)$.

More precisely, an $\{\mathcal{F}_t\}$-semimartingale $X_t$ is called a solution of the equation starting from $X_{t_0}$ at time $t_0$ if it satisfies the equality

$$
\begin{align*}
X_t = X_{t_0} + \int_{t_0}^t b(s, X_{s-})ds + \int_{t_0}^t \sigma(s, X_{s-})dW_s \\
+ \int_{t_0}^t \int_{R_x} g(s, X_{s-}, z)\bar{N}(dtdz).
\end{align*}
$$

The existence and the (pathwise) uniqueness of the solution may be shown under the above conditions for coefficients.

Let $X_t^{(s,x)}$ be the solution starting from $x$ at time $s$. We have the relation $X_u^{(s,x)} = X_u^{(t, X_t^{(s,x)})}$ a.s. for $s < t < u$. Since $X_u^{(t,x)}$ is independent of $\mathcal{F}_t$ and since $X_t^{(s,x)}$ is $\mathcal{F}_t$-measurable, we get

$$
E[f(X_u^{(s,x)})|\mathcal{F}_t] = E[f(X_u^{(t,y)})]_{y = X_t^{(s,x)}}, \quad a.s.
$$

Therefore for any $s, x$, $X_t^{(s,x)}$ is a Markov process. Set $P_{s,t}(x, E) = P(X_t^{(s,x)} \in E)$ and denote $P_{s,t}f(x) = \int f(y)P_{s,t}(x, dy)$. Then the above equality is written as

$$
E[f(X_u^{(s,x)})|\mathcal{F}_t] = P_{t,u}f(X_t^{(s,x)}), \quad a.s.
$$

Taking expectations, we get the equality $P_{s,u}f(x) = P_{s,t}P_{t,u}f(x)$, which is called the Chapman–Kolmogorov equation. The solution $X_t^{(s,x)}$ is called a jump–diffusion.
We may obtain Kolmogorov’s forward and backward equations for its transition probabilities by applying Itô’s formula to jump–diffusion processes. We shall apply Itô’s formula (Theorem 2.5) to a \( C^2 \)-function \( f \) such that \( f_{x_i}, f_{x_ix_j} \) are bounded. Set
\[
\mathcal{L}(t) f(x) := \frac{1}{2} \sum_{i,j} a^{ij}(t,x) f_{x_i,x_j}(x) + \sum_i b^i(t,x) f_{x_i}(x)
+ \int_{R_0} \left\{ f(x + g(t,x,z)) - f(x) - \sum_i g^i(t,x,z) f_{x_i}(x) \right\} v(dz).
\tag{2.18}
\]
The last integral is well defined since
\[
|f(x + g(t,x,z)) - f(x) - \sum_i g^i(t,x,z) f_{x_i}(x)| \leq \frac{1}{2} \sum_{i,j} \|f_{x_i,x_j}\||g(t,x,z)|^2
\]
and \( g \) satisfies (2.16). Further, each term on the right-hand side of (2.18) is dominated by \( C(1 + |x|)^2 \).

Now, define a family of measures by \( \nu^{(t,x)} = (g^{(t,x)})_x \nu \), where \( g^{(t,x)}(z) = g(t,x,z) \). Then \( (b(t,x), A(t,x), d\tau\nu^{(t,x)}(dz)) \) is the characteristic triplet of the tangential Lévy process at the point \((t,x)\) for Eq. (2.17). The above operator \( \mathcal{L}(t) \) can be rewritten by
\[
\mathcal{L}(t) f(x) = \frac{1}{2} \sum_{i,j} a^{ij}(t,x) f_{x_i,x_j}(x) + \sum_i b^i(t,x) f_{x_i}(x)
+ \int_{R_0} \left\{ f(x + y) - f(x) - \sum_i y^i f_{x_i}(x) \right\} \nu^{(t,x)}(dy).
\tag{2.19}
\]
For \( X_t = X_t^{(s,x)} \), Itô’s formula is written as
\[
f(X_t) = f(x) + \int_s^t \mathcal{L}(u) f(X_{u-}) du + \sum_{i,j} \int_s^t f_{x_i}(X_{u-}) \sigma^{ij}(u,X_{u-}) dW_u^j
+ \int_s^t \int_{R_0} \left\{ f(X_{u-} + g(u,X_{u-},z)) - f(X_{u-}) \right\} \tilde{N}(dudz).
\tag{2.20}
\]
We shall consider the expectation of each term of the above equation. We have \( E[f(X_t)] = P_{s,t} f(x) \) and
\[
E\left[ \int_s^t \mathcal{L}(u) f(X_{u-}) du \right] = \int_s^t E[\mathcal{L}(u) f(X_{u-})] du = \int_s^t P_{s,u} \mathcal{L}(u) f(x) du.
\]
Here we used that \( P(X_u = X_{u-}) = 1 \) holds for any \( u \). Processes \( \phi^j(u) = \sum_i f_{x_i}(X_{u-}) \sigma^{ij}(u,X_{u-}), j = 1, \ldots, d \) belong to \( \Phi^2 \). Then their stochastic integrals with respect to \( dW_u^j \) are martingales with mean 0. Further, the process \( \psi(u,z) = f(X_{u-} + g(u,X_{u-},z)) - f(X_{u-}) \) satisfies
\[
|\psi(u,z)|^2 \leq \left( \sum \|f_{x_i}\|^2 \right)^2 (1 + |X_{u-}|^2) K(z)^2.
\]
Hence $\psi \in \Psi^2$ and $\int_0^t \int \psi d\tilde{N}$ is a square-integrable martingale with mean 0. Then Eq. (2.20) implies that the transition probabilities $P_{s,t}$ satisfy

$$P_{s,t} f(x) = f(x) + \int_s^t P_{s,u} \mathcal{L}(f(x)) du.$$  

Consequently, we have

**Theorem 2.8.** Let $X_t$ be the jump–diffusion determined by SDE (2.17). Then its transition probabilities $P_{s,t}$ satisfy Kolmogorov’s forward integro-differential equation

$$\frac{\partial}{\partial t} P_{s,t} f(x) = P_{s,t} \mathcal{L}(f)(x), \quad 0 < s < t < T.$$  

where $\mathcal{L}(f)$ is given by (2.18) or (2.19).

The operator $\mathcal{L}(f)$ is called the generator of the jump–diffusion $X_t$. To obtain Kolmogorov’s backward equation, we need some smoothness conditions on the coefficients.

**Theorem 2.9.** Assume that the coefficients $b, \sigma, g$ are twice continuously differentiable with respect to $x$ and their derivatives (up to the second order) are of linear growth and Lipschitz continuous. Then if $f$ is a $C^2$-function with bounded derivatives, $P_{s,t} f(x)$ is a $C^{1,2}$ function of $(s, x)$. It satisfies Kolmogorov’s backward integro-differential equation

$$\frac{\partial}{\partial s} P_{s,t} f(x) + \mathcal{L}(s) P_{s,t} f(x) = 0, \quad 0 < s < t < T.$$  

**Proof.** It is known that the solution $X_t^{(s,x)}$ is twice continuously differentiable with respect to $x$ (see [23]). Then $P_{s,t} f(x)$ is also twice continuously differentiable with respect to $x$, if $f$ is a smooth function. It is actually a $C^{1,2}$ function of $(s, x)$.

We will prove (2.22). For a fixed $t$, set $M_s = P_{s,t} f(X_s), t_0 \leq s \leq t$, where $X_s, t_0 \leq s \leq t$ is a solution of the SDE. It is a martingale, because for $t_0 \leq s \leq u \leq t$,

$$E[M_u | \mathcal{F}_s] = E[P_{u,t} f(X_u) | \mathcal{F}_s] = P_{s,u} P_{u,t} f(X_s) = P_{s,t} f(X_s) = M_s.$$  

Applying Itô’s formula to the function $F(s, x) := P_{s,t} f(x)$, we have

$$M_u = F(t_0, X_{t_0}) + \int_{t_0}^u \left\{ \frac{\partial}{\partial s} F(s, X_{s-}) + \mathcal{L}(s) F(s, X_{s-}) \right\} ds + \text{a local martingale}.$$  

Therefore the integral $\int_{t_0}^u \left\{ \frac{\partial}{\partial s} + \mathcal{L}(s) \right\} F(s, X_{s-}) ds$ is a local martingale, so that it is equal to 0 a.s. Therefore $\left\{ \frac{\partial}{\partial s} + \mathcal{L}(s) \right\} F(s, X_{s-}) = 0$ a.s. for any $t_0 \leq s \leq t$. Since this is valid for any solution $X_s$, we have $\left\{ \frac{\partial}{\partial s} + \mathcal{L}(s) \right\} F(s, x) = 0 \text{ for any } t_0 \leq s \leq t \text{ and } x \in \mathbb{R}^d.$

Let us transform the probability measure $P$ to an equivalent measure $Q = \alpha_T (\phi, \psi) \cdot P$. Assume that the pair $(\phi(t), \psi(t, z))$ is given by $\phi(t) = \phi(t, X_t)$ and $\psi(t, z) = \psi(t, X_{t-}, z)$, where $\phi(t, x), \psi(t, x, z)$ are continuous functions such that both $\phi(t, x)$ and $\int \psi^2(t, x, z) v(dz)$ are bounded with respect to $t, x$. Then $\alpha_t (\phi, \psi)$ is a martingale. Hence $Q = \alpha_T (\phi, \psi) \cdot P$ is well defined.

---

1 An alternative and rigorous proof is given in [25], where the backward chain rule (backward Itô formula) for the stochastic flow $X_t^{(s,x)}$ is used.
We are interested in the change of generator for $X$ through Girsanov's transformation. We set

$$b^{(\phi, \psi)}(t, x) := b(t, x) + \phi(t, x)\sigma(t, x) + \int_{R_0} \psi(t, x, z)g(t, x, z)d\nu(z) \quad (2.23)$$

and define

$$L^Q(t) f(x) := \frac{1}{2} \sum_{i, j} a^{ij}(t, x) f_{x_i x_j}(x) + \sum_i b^{(\phi, \psi), i}(t, x) f_{x_i}(x)$$

$$+ \int_{R_0} \left( f(x + g(t, x, z)) - f(x) - \sum_i g^i(t, x, z) f_{x_i}(x) \right)(1 + \psi(t, x, z))\nu(dz). \quad (2.24)$$

Each term of the above is well defined and is dominated by $C(1 + |x|)^2$.

**Theorem 2.10.** $X_t$ is a jump–diffusion with respect to $Q$. Its generator is given by (2.24).

**Proof.** Let $X_t = X_{(t, x)}^{(s, x)}$ be the solution starting from $x$ at time $s$. We can rewrite Itô's formula (Theorem 2.5) as

$$f(X_t) = f(X_s) + \int_s^t L^Q(u) f(X_{u-})du + \sum_{i, j} \int_s^t f_{x_i}(X_{u-})\sigma^{ij}(u, X_{u-})dW_{u}^{\phi, j}$$

$$+ \int_s^t \int \{ f(X_{u-} + g(u, X_{u-}, z)) - f(X_{u-}) \} \tilde{N}^{\psi}(dudz).$$

Take the expectation of each term with respect to $Q$. The expectations of the last two stochastic integrals with respect to $dW^{\phi}$ and $d\tilde{N}^{\psi}$ are both 0. Therefore we get the equality

$$P_{s,t}^Q f(x) = f(x) + \int_s^t P_{s,u}^Q L^Q(u) f(x)du,$$

where $P_{s,t}^Q f(x) = E^Q[ f(X_{t}^{(s,x)}))] = \int f(X_{t}^{(s,x)})dQ$. Then we see that $L^Q(t)$ is the generator of $X_t$ under $Q$.  \quad \square

### 3. Application to mathematical finance

#### 3.1. Asset, option and non-arbitrage market

We will introduce two processes. One is the price of the riskless asset. It is defined by $B^0_t = e^{rt}$, where $r$ is a positive constant. It is called a bank account process. The other one is the price process of a risky asset which may have jumps.

We need some notation. Let $(\Omega, \mathcal{F}, P)$ be a probability space, where the following two processes are defined. $W_t$ is a one-dimensional standard Brownian motion and $N(dr dz)$ is a homogeneous Poisson random measure on $R_0$ with intensity measure $dr d\nu(dz)$, which are mutually independent. Let $\{\mathcal{F}_t\}$ be the smallest filtration such that for any $t$, $W_s, N(s, B), s \leq t, B \in \mathcal{B}$ are measurable with respect to $\mathcal{F}_t$. The price process $S_t$ of a risky asset is defined by the following one-dimensional SDE:

$$dS_t = S_{t-}dY(S_{t-}, dr), \quad (3.1)$$
where \( Y(x, \, dt) \) is a random vector field given by

\[
Y(x, \, dt) = b(t, x) \, dt + \sigma(t, x) \, dW_t + \int_{R_0} g(t, x, z) \, \tilde{N}(drdz).
\]  
(3.2)

For the coefficients \( b, \sigma, g \), we assume the following:

(1) \( b, \sigma \) are bounded and Lipschitz continuous. Further, \( \sigma \) is uniformly positive.

(2) \( g(t, x, z) \) is bounded and Lipschitz continuous with respect to \( v \). Further, it holds \( g(t, x, z) + 1 > 0 \) for any \( t, x, z \).

Then Eq. (3.1) has a unique solution. Let \( S_{x,t}(x) \) be the solution starting from \( x > 0 \) at time \( s \). Then it holds \( S_{x,t}(x) > 0 \) for any \( t > s \) a.s.

The random vector field \( Y(x, \, dt) \) is called the return process. Its mean \( b(t, x) \, dt \) shows the expected rate of return of the asset and the variance \( \sigma(t, x)^2 + \int g(t, x, z)^2 \nu(dz) \, dt \) shows the rate of risk of the asset. \( \sigma(t, x) \) is called the volatility. \( \int g(t, x, z)^2 \nu(dz) \) is the rate of jump risk.

If the coefficients \( b, \sigma \) and \( g \) do not depend on the state \( x \) and time \( t \), the random vector field \( Y(x, \, dt) \) does not depend on \( x \). It is a Lévy process. Then \( \log S_{x,t}(x) \) is also a Lévy process. The solution \( S_{x,t}(x) \) is called a geometric Lévy process. In particular if the Poisson part is \( 0 \) (\( g = 0 \)), the solution is called a geometric Brownian motion or Black–Scholes model. If the Lévy measure \( \nu \) has a finite total mass, it is called a Merton model.

A nonnegative \( \mathcal{F}_T \)-measurable random variable \( h \) is called a contingent claim. If it is written as \( h = h(S_T) \) with a fixed time \( T \), it is called a European option and the function \( h \) is called a pay-off function. Let \( K \) be a positive constant. In the case where \( h(x) = (x - K)^+ \), it is called a European call option with exercise price \( K \) and in the case where \( h(x) = (K - x)^+ \), it is called a European put option with exercise price \( K \). In this paper, we assume that \( h \) is a continuous function and \( \frac{h(x)}{1+|x|} \) is bounded.

We shall consider a European option. Suppose that we have a guarantee to be paid the amount \( h(S_T) \) at time \( t = T \). An important question in finance is to find the price \( P_t \) of the option at each time \( 0 \leq t < T \), which both the buyer and the seller of the option can agree with. The problem was solved in 1973 by Black–Scholes and Merton in the case where the price process is a diffusion (Black–Scholes model). In 1975 Merton proposed a price in the case where the price process is a jump–diffusion (Merton model). We will review their arguments from a slightly different viewpoint.

We will assume that the price of the option is given in the form \( P_t = F(t, s) \) with a function \( F(t, s), t \in [0, T], s > 0 \) of the \( C^{1,2} \)-class such that \( F(T, x) = h(x) \). The function \( F \) is called a pricing function. In their arguments, they applied Itô’s formula in a striking way. They found a partial differential equation or an integro-differential equation for the pricing function \( F(t, x) \) in such a way that the market \((B^0_t, S_t, P_t)\) becomes non-arbitrage, through Itô’s formula.

Now we shall consider a market consisting of the triple \((B^0_t, S_t, P_t)\). By a strategy or a portfolio, we mean a triple of progressively measurable processes \( \theta(t) = (\theta_0(t), \theta_1(t), \theta_2(t)) \). The value process of the portfolio \( \theta(t) \) is defined by

\[
V_t(\theta) = \theta_0(t) B^0_t + \theta_1(t) S_t + \theta_2(t) P_t.
\]  
(3.3)

Suppose that \( \int_0^T (|\theta_0| + |\theta_1|^2 + |\theta_2|^2) \, dt < \infty \) and that the pair \((\theta_1(t), \theta_2(t))\) is predictable. If the equality

\[
V_t(\theta) = V_0(\theta) + \int_0^t \theta_0(u) \, dB^0_u + \int_0^t \theta_1(u) \, dS_u + \int_0^t \theta_2(u) \, dP_u,
\]  
(3.4)
holds for any $t$ a.s., the portfolio $θ(t)$ is called self-financing.

A self-financing portfolio $θ(t)$ is called an arbitrage opportunity, provided that the three properties $V_0(θ) = 0$, $V_T(θ) ≥ 0$ a.s. and $P(V_T(θ) > 0) > 0$ hold. If there is no arbitrage opportunity, the market is called non-arbitrage. A self-financing portfolio is called riskless if the value process $V_t(θ)$ does not include the martingale part.

We define normalized (or discounted) processes $\tilde{S}_t$, $\tilde{P}_t$ and $\tilde{V}_t(θ)$ by $\tilde{S}_t = e^{-rt}S_t$, $\tilde{P}_t = e^{-rt}P_t$ and $\tilde{V}_t(θ) = e^{-rt}V_t(θ)$. It is not difficult to see that if the market is non-arbitrage, any riskless normalized value process $\tilde{V}_t(θ)$ is equal to $V_0(θ)$ for any $0 ≤ t ≤ T$.

The next lemma is easily verified.

**Lemma 3.1.** Let $(θ_1(t), θ_2(t))$ be a pair of predictable processes such that $\int_0^T (|θ_1|^2 + |θ_2|^2)dt < ∞$ and $V_0$ be an $F_0$-measurable random variable. Set

$$θ_0(t) = V_0 + \int_0^t θ_1 d\tilde{S}_u + \int_0^t θ_2 d\tilde{P}_u - θ_1(t)\tilde{S}_t - θ_2(t)\tilde{P}_t.$$  

(3.5)

Then the triple $θ(t) = (θ_0(t), θ_1(t), θ_2(t))$ is a self-financing portfolio. Further, its normalized value process $\tilde{V}_t(θ)$ satisfies

$$\tilde{V}_t(θ) = V_0 + \int_0^t θ_1 d\tilde{S}_u + \int_0^t θ_2 d\tilde{P}_u.$$  

(3.6)

We will consider the problem of finding the pricing function $F(t, x)$, which both the seller and the buyer of the option could agree with. Discussions for the case where $S_t$ is a diffusion and the case where $S_t$ is a jump–diffusion are quite different. In Sections 3.2 and 3.3, we shall consider the case where $S_t$ is a diffusion. The case where $S_t$ has jumps will be discussed in Sections 3.4 and 3.5.

### 3.2. The Black–Scholes partial differential equation

In this and the next subsections, we assume that the price process is a diffusion. Eq. (3.1) is written simply as

$$dS_t = S_t b(t, S_t) dt + S_t σ(t, S_t) dW_t.$$  

(3.7)

We assume further that the market $(B^0_t, S_t, P_t)$ is non-arbitrage in this subsection.

Let $θ(t)$ be a self-financing portfolio given in Lemma 3.1. We shall compute the normalized value process $\tilde{V}_t(θ)$ written by (3.6). Since $S_t$ satisfies the above equation, $\tilde{S}_t = e^{-rt}S_t$ satisfies

$$d\tilde{S}_t = e^{-rt}S_t (b(t) - r)dt + e^{-rt}S_t σ(t) dW_t,$$  

(3.8)

where $b(t) = b(t, S_t)$, $σ(t) = σ(t, S_t)$, etc. Set $\tilde{F}(t, x) = e^{-rt}F(t, x)$ and apply Itô’s formula. Then $\tilde{P}_t := \tilde{F}(t, S_t)$ satisfies

$$d\tilde{P}_t = e^{-rt} \left( F_t + S_t b(t) F_x + \frac{1}{2} S_t^2 σ(t)^2 F_{xx} - r F \right) dt + e^{-rt} S_t σ(t) F_x dW_t.$$  

(3.9)

Then, $d\tilde{V}_t(θ)$ can be written as

$$d\tilde{V}_t(θ) = θ_1(t)d\tilde{S}_t + θ_2(t)d\tilde{P}_t = β(t) dt + α(t) dW_t,$$
where

\[ \alpha(t) = e^{-rt} (\theta_1(t) S_t \sigma(t) + \theta_2(t) S_t \sigma(t) F_x), \]
\[ \beta(t) = e^{-rt} \left( \theta_1(t) (S_t b(t) - \sigma(t) r) + \theta_2(t) (F_t + S_t b(t) F_x + \frac{1}{2} S_t^2 \sigma(t)^2 F_{xx} - r F) \right). \]

Now, in order to eliminate the risk from the normalized value process \( \tilde{V}_t(\theta) \), choose portfolios \( \theta_1, \theta_2 \) such that \( \alpha(t) = 0 \) holds for any \( t \). For example, choose \( \theta_1'(t) = F_x(t, S_t), \theta_2'(t) = -1 \) and define \( \theta_0(t) \) by (3.5) with setting \( V_0 = 0 \). Set \( \theta'(t) = (\theta_0'(t), \theta_1'(t), \theta_2'(t)) \). Then the random component due to \( dW_t \) is eliminated in \( d\tilde{V}_t(\theta') \). Further the drift coefficient \( \beta(t) \) is written as

\[ \beta(t) = -e^{-rt} \left( F_t(t) + \frac{1}{2} S_t^2 \sigma(t)^2 F_{xx}(t) + S_t r F_x(t) - r F(t) \right). \]

Then the riskless normalized value process \( \tilde{V}_t(\theta') \) should be equal to the initial value \( \tilde{V}_0(\theta') \), since the market is non-arbitrage. This implies that \( \beta(t) = 0 \) for all \( t \). Since the support of the law of \( S_t \) is the whole space \( (0, \infty) \), we have the following assertion.

**Theorem 3.2 (Black–Scholes [11]).** Suppose that the market \( (B_t^0, S_t, F(t, S_t)) \) is non-arbitrage. Then the pricing function \( F(t, x), 0 \leq t \leq T, x > 0 \) satisfies the following backward partial differential equation.

\[ F_t(t, x) + \frac{1}{2} x^2 \sigma(t, x)^2 F_{xx}(t, x) + x r F_x(t, x) - r F(t, x) = 0, \quad (3.10) \]

\[ F(T, x) = h(x). \quad \text{(terminal condition)} \]

We shall find a hedging portfolio for the option \( h(S_T) \). The triple \( \theta'(t) = (\theta_0'(t), \theta_1'(t), \theta_2'(t)) \) chosen above is a self-financing and riskless portfolio such that \( V_0(\theta') = 0 \). Then because of non-arbitrage, we have

\[ \theta_0'(t) B^0_t + F_x(t, S_t) S_t - F(t, S_t) = V_t(\theta') = 0. \]

This yields

\[ \theta_0'(t) = e^{-rt} (F(t, S_t) - F_x(t, S_t) S_t). \]

Since \( \theta'(t) \) is self-financing and \( V_t(\theta') = 0 \), we get from formula (3.4)

\[ \int_t^T e^{-ru} (F(u, S_u) - F_x(u, S_u) S_u) dB^0_u + \int_t^T F_x(u, S_u) dS_u - \int_t^T dP_u = 0. \]

Since \( \int_t^T dP_u = h(S_T) - F(t, S_t) \), \( h(S_T) \) is represented by

\[ h(S_T) = F(t, S_t) + \int_t^T e^{-ru} (F(u, S_u) - F_x(u, S_u) S_u) dB^0_u + \int_t^T F_x(u, S_u) dS_u. \]

The above formula indicates that if we (buyer and seller of the option) hold the self-financing portfolio \( (\theta_0'(t), \theta_1'(t)) \) for \( (B_t^0, S_t) \), its value at the terminal time \( T \) coincides with the option \( h(S_T) \). Consequently the option \( h(S_T) \) is completely hedged by the above portfolio. In other words, the portfolio \( (\theta_0'(t), \theta_1'(t)) \) is a replica of the option. Therefore we can agree about \( F(t, S_t) \) as the price of the option at \( t \) without any risk. For further informations on hedging problems, we refer to [20,27].
3.3. Equivalent martingale measure for diffusions

Let \( Q \) be another probability defined on the same measurable space (\( \Omega, \mathcal{F} \)). If \( P \) and \( Q \) are equivalent (mutually absolutely continuous) on \( \mathcal{F}_T \), the price process \( S_t \) can be regarded as a stochastic process with respect to \( Q \). An equivalent measure \( Q \) is called a martingale measure if \( \tilde{S}_t := e^{-rt}S_t \) is a local martingale with respect to \( Q \).

We shall construct a martingale measure. Set \( \phi(t) = \frac{b(t, S_t) - r}{\sigma(t, S_t)} \) and define

\[
\alpha_t = \alpha_t(\phi) = \exp\left\{ \int_0^t \phi dW_s - \frac{1}{2} \int_0^t |\phi|^2 ds \right\}.
\]

Since \( \phi(t) \) is a bounded process, \( \alpha_t \) is a positive martingale. Let \( Q = \alpha_T \cdot P \). By Girsanov’s theorem, \( W_t^\phi := W_t - \int_0^t \phi ds \) is a \( Q \)-Brownian motion. Note that \( \tilde{S}_t \) satisfies (3.8). It is rewritten as

\[
d\tilde{S}_t = \tilde{S}_t \alpha(t) dW_t^\phi.
\]

Therefore \( \tilde{S}_t \) is a martingale with respect to \( Q \). Consequently \( Q \) is a martingale measure.

Apply Theorem 2.10 to the case of a diffusion, i.e., \( \nu \equiv 0 \). Then we find that with respect to \( Q, \tilde{S}_t \) is a diffusion process with generator

\[
\mathcal{L}Q(t)F(x) = \frac{1}{2} \sigma^2(t, x)^2 F''(x) + x r F'(x).
\]

The Black–Scholes backward partial differential equation (3.10) is written as

\[
\left( \frac{\partial}{\partial t} + \mathcal{L}Q(t) - r \right) F = 0, \tag{3.11}
\]

\[
F(T, x) = h(x).
\]

It is known that for any continuous function \( h \) with polynomial growth, the solution of the above equation exists uniquely. It is given by

\[
F(t, x) = E^Q[e^{-r(T-t)}h(S_t^{(s, x)})], \tag{3.12}
\]

where \( S_t^{(s, x)} \) is the solution of Eq. (3.7) starting from \( x \) at time \( s \). In other words, the Black–Scholes equation (3.11) is equal to Kolmogorov’s backward equation stated in Theorem 2.9 for the discounted transition function

\[
\bar{P}_{S, t}^Q f(x) = e^{-r(t-s)} E^Q[f(S_t^{(s, x)})].
\]

We will check this fact again in Section 3.5.

The next theorem shows that the Black–Scholes equation for the pricing function \( F(t, x) \) is a necessary and sufficient condition for the market \((B^0_t, S_t, F(t, S_t))\) to be non-arbitrage. Thus the Black–Scholes equation is characterized as the unique non-arbitrage price of the European option with pay-off function \( h(x) \).

**Theorem 3.3.** Let \( F(t, x), 0 \leq t \leq T, x \in \mathbb{R}^+ \) be a \( C^{1,2} \)-function with \( h(T, x) = h(x) \) and \( P_t = F(t, S_t) \). We set \( \tilde{S}_t = e^{-rt}S_t \) and \( \bar{P}_t = e^{-rt}P_t \) as before. The following three statements are equivalent.

1. \( \bar{P}_t \) is a \( Q \)-martingale.
(2) $F$ satisfies the Black–Scholes equation.

(3) The triple $(B_0^0, S_t, P_t)$ is non-arbitrage.

**Proof.** We first show the equivalence of (1) and (2). We can rewrite Eq. (3.9) as

$$d\tilde{P}_t = e^{-rt} \left( \frac{\partial}{\partial t} + \mathcal{L}(t) - r \right) F(t, S_t) dt + e^{-rt} S_t \sigma(t, S_t) dW_t^\phi.$$  \hfill (3.13)

Hence if $\tilde{P}_t$ is a martingale with respect to $Q$, then $(\frac{\partial}{\partial t} + \mathcal{L}(t) - r) F(t, S_t) = 0$, proving the Black–Scholes equation. Conversely if (2) is satisfied, then $(\frac{\partial}{\partial t} + \mathcal{L}(t) - r) F = 0$, so that $\tilde{P}_t$ is a martingale with respect to $Q$ by formula (3.13).

Next we will show that (1) implies (3). If (1) is satisfied, $\tilde{V}_t(\theta)$ is a continuous martingale. If it is riskless, it should be a continuous process of finite variation. Then it should be equal to $\tilde{V}_0(\theta)$, proving (3).

Finally suppose (3). We saw in **Theorem 3.2** that $F$ satisfies the Black–Scholes equation. Then $\tilde{P}_t$ is a $Q$-martingale, in view of (3.13). □

### 3.4. Merton’s integro-differential equation

In this subsection we assume that the price process $S_t$ is a jump–diffusion determined by (3.1) and (3.2). We will again consider the triple $(\tilde{B}_0^0, S_t, P_t)$ where $P_t = F(t, S_t)$ and $F$ is a pricing function of the $C^{1,2}_0$-class. We assume that the function $F(t, x)$ is of linear growth with respect to $x$.

Let $\tilde{S}_t = e^{-rt} S_t$ and $\tilde{P}_t = e^{-rt} P_t$ as before. We have

$$d\tilde{S}_t = e^{-rt} S_t \left( b(t, S_t) - r \right) dt$$

$$+ e^{-rt} S_t \left[ \sigma(t, S_t) dW_t + \int_{R_0} g(t, S_t, z) \tilde{N}(drdz) \right].$$ \hfill (3.14)

Apply Itô’s formula (**Theorem 2.5**) to the function $\tilde{F}(t, x) = e^{-rt} F(t, x)$. Then we have

$$d\tilde{P}_t = e^{-rt} \left( \frac{\partial}{\partial t} + \mathcal{L}(t) - r \right) F(t, S_t) dt$$

$$+ e^{-rt} \left[ S_t \sigma(t, S_t) F_x(t, S_t) dW_t + \int_{R_0} G d\tilde{N} \right],$$ \hfill (3.15)

where $\mathcal{L}(t)$ is the generator of the jump–diffusion $S_t$, given by

$$\mathcal{L}(t) f(x) = \frac{1}{2} x^2 \sigma(t, x)^2 f_{xx}(x) + x b(t, x) f_x(x)$$

$$+ \int_{R_0} \{ f(x(1 + g(t, x, z))) - f(x) - x g(t, x, z) f_x(x) \} \nu(dz),$$ \hfill (3.16)

and $G = F(t, x(1 + g(t, x, z))) - F(t, x)$. Let $\theta(t) = (\theta_0(t), \theta_1(t), \theta_2(t))$ be a self-financing portfolio for $(\tilde{B}_t^0, S_t, P_t)$. Then the normalized value process $\tilde{V}_t(\theta)$ of (3.6) satisfies

$$d\tilde{V}_t(\theta) = \beta'(t) dt + \alpha'(t) dW_t + \int_{R_0} \gamma'(t, z) \tilde{N}(drdz).$$
where

\[ \alpha'(t) = e^{-rt} \left[ \theta_1(t) S_t \sigma(t) + \theta_2(t) S_t \sigma(t) F_x \right]. \]

\[ \beta'(t) = e^{-rt} \left[ \theta_1(t) (S_t - b(t)) - S_t \gamma r + \theta_2(t) \left( \frac{\partial}{\partial t} + L(t) - r \right) F \right]. \]

Here \( b(t) = b(t, \sigma(t)) = \sigma(t), \sigma(t) = \sigma(t, S_t). \)

Now, in order to eliminate the continuous risk from the normalized value process \( \tilde{V}_t(\theta) \), take the portfolio \( \theta''_1(t) = F_x(t, S_t) \) and \( \theta''_2(t) = -1 \). Then we have \( \alpha'(t) = 0 \). Further, we can rewrite \( \beta'(t) \) explicitly. In fact, the normalized value process is written as

\[ \tilde{V}_t(\theta'') = V_0 + \int_0^t e^{-rs} A(s) F(s, S_s) ds + \int_0^t \int_{R_0} \gamma'(r, z) \tilde{N}(drdz), \]

where

\[ A(t) F(t, x) = F_t(t, x) + \frac{1}{2} x^2 \sigma(t, x)^2 F_{xx}(t, x) + x r F_x(t, x) - r F(t, x) \]

\[ + \int_{R_0} \left( F(t, x + xg(t, x, z)) - F(t, x) - xg(t, x, z) F_x(t, x) \right) d\nu(z). \tag{3.17} \]

We have thus eliminated the continuous risk concerned with \( dW_t \). However, we cannot eliminate \( \gamma'(r, z) \), i.e., the jump risk by any portfolio \( (\theta_1(t), \theta_2(t)) \). Then, the jump risk should be neutral to the seller and the buyer of the option. Hence the normalized value process \( \tilde{V}_t(\theta'') \) should be risk neutral, i.e., it should be a local martingale. Then we have \( A(t) F(t, S_t) = 0 \) for any \( t \). Merton [29] adopted such \( F(t, x) \) as the pricing function of the option.

**Theorem 3.4** (Merton [29]). Let \( (B^0_t, S_t, F(t, S_t)) \) be the triple where \( S_t \) is a jump–diffusion determined by SDE (3.1). Then there exists a self-financing portfolio which eliminates the continuous risk and makes the jump risk neutral, if and only if the pricing function \( F(t, x) \) satisfies the backward integro-differential equation

\[ A(t) F(t, x) = 0, \quad \forall(t, x) \in [0, T] \times (0, \infty), \tag{3.18} \]

\[ F(T, x) = h(x), \quad \text{(terminal condition)}. \]

We will see in Proposition 3.9 that the above pricing function \( F \) makes \( (B^0_t, S_t, F(t, S_t)) \) non-arbitrage.

A self-financing portfolio \( (\theta''_0(t), \theta''_1(t), \theta''_2(t)) \) which hedges the continuous risk is given by

\[ \theta''_0(t) = e^{-rt} (F(t, S_t) - F_x(t, S_t) S_t), \quad \theta''_1(t) = F_x(t, S_t), \quad \theta''_2(t) = -1. \]

It takes the same form as the portfolio in the diffusion case.

**3.5. Equivalent martingale measures for jump–diffusions**

In the previous subsection, we reviewed the approach by Merton for pricing options. We obtained Merton’s integro-differential equation (3.18) for the pricing function \( F(t, x) \). In this subsection, we will show the existence and uniqueness of the solution of Merton’s equation by using an equivalent martingale measure.

We will show that there are infinitely many equivalent martingale measures for a jump–diffusion, in contrast to the unique martingale measure for a diffusion. Let \( Q \) be an
equivalent probability measure and let \( \alpha_t \) be the Radon–Nikodym density \( \frac{dQ}{dP} |_{\mathcal{F}_t} \). We saw in Section 2.3 that it is represented as the unique solution of a SDE \( d\alpha_t = \alpha_t \cdot dZ_t \) with the initial condition \( \alpha_0 = 1 \), where \( Z_t \) is a local martingale written as

\[
Z_t = \int_0^t \phi(s) dW_s + \int_0^t \int_0^s \psi(s, z) N(dsdz),
\]
with \( \phi \in \Phi_{loc}^2, \psi \in \Psi_{loc} \). We denote it as \( \alpha_t(\phi, \psi) \).

**Lemma 3.5.** ([24]) \( \tilde{S}_t \) is a local martingale with respect to \( Q \) if and only if \( \psi(t, z) g(t, S_{t-}, z) \in \Psi_{loc}^1 \) and satisfies a.e. \( dt \otimes P \),

\[
b(t, S_{t-}) + \phi(t) \sigma(t, S_{t-}) + \int_{R_0} \psi(t, z) g(t, S_{t-}, z) \nu(dz) = r. \tag{3.19}
\]

**Proof.** The measure \( Q \) is a martingale measure if and only if \( \alpha_t \tilde{S}_t \) is a local martingale with respect to \( P \). It holds

\[
d(\alpha_t \tilde{S}_t) = \tilde{S}_t d\alpha_t + \alpha_t \cdot d\tilde{S}_t + d[\alpha, \tilde{S}]_t = \alpha_t \cdot \tilde{S}_{t-} \cdot (dZ_t + dY_t + d[Z, Y]_t).
\]

Since \( [Z, Y]_t = \int_0^t \phi \sigma ds + \int_0^t \int \psi g dN \), we have

\[
dZ_t + dY_t + d[\alpha, \tilde{S}]_t = dZ_t + \sigma(t- \cdot dW_t + (b(t- \cdot) + \phi(t) \sigma(t- \cdot) dt + \int_{R_0} \psi g dN.
\]

The above is a local martingale if and only if \( \psi g \in \Psi_{loc}^1 \) and the equality

\[
(b(t- \cdot) + \phi(t) \sigma(t- \cdot) + \int_{R_0} \psi(t, z) g(t, S_{t-}, z) \nu(dz) = 0, \quad dt \text{ a.e.}
\]

holds. \( \square \)

If \( \nu = 0 \), Eq. (3.19) is equivalent to \( b(t- \cdot) + \phi(t) \sigma(t- \cdot) = r, \ \text{dt a.e.} \). Then the solution \( \phi(t) \) exists uniquely as an element of \( \Phi_{loc}^2 \). It is a bounded function \( dt \text{ a.e.} \) Hence an equivalent martingale measure exists uniquely. On the other hand, if \( \nu \neq 0 \), there exist infinitely many pairs \( (\phi(t), \psi(t)) \) such that \( \psi, \phi \) and \( \int \psi(t, z)^2 \nu(dz) \) are bounded and satisfy Eq. (3.19). Therefore we have

**Proposition 3.6.** (1) If \( \nu = 0 \), a martingale measure exists uniquely.

(2) If \( \nu \neq 0 \), there are infinitely many martingale measures.

Now let \( \psi(t, x, z) \) be a function such that \( \int \psi(t, x, z)^2 \nu(dz) \) is bounded with respect to \( t, x \). There exists a bounded \( \phi(t, x) \) satisfying

\[
b(t, x) + \phi(t, x) \sigma(t, x) + \int_{R_0} \psi(t, x, z) g(t, x, z) \nu(dz) = r. \tag{3.20}
\]

Then the local martingale \( \alpha_t = \alpha_t(\phi, \psi) \) is a martingale and \( Q = \alpha_T \cdot P \) is a martingale measure. We denote by \( Q \) the set of all martingale measures \( Q = \alpha_T(\phi, \psi) \cdot P \) such that the pair \( (\phi, \psi) \) satisfies the above property.

The next proposition follows from Theorem 2.10 since \( b^\phi \cdot \psi(t, x) = r \) holds by (3.20).
Proposition 3.7. With respect to $Q \in \mathcal{Q}$, $S_t$ is a jump–diffusion process. Its generator is given by
\[
\mathcal{L}_Q^0(t) f(x) = \frac{1}{2} x^2 \sigma(t, x)^2 f''(x) + x r f'(x)
\]
\[
+ \int_{R_0} \left\{ f(x + xg(t, x, z)) - f(x) - xg(t, x, z) f'(x) \right\} (1 + \psi(t, x, z)) \nu(dz).
\] (3.21)

A simple martingale measure is obtained in the case $\psi \equiv 0$. Then $\phi_0(t) = (b(t) - r)/\sigma(t)$ is the unique solution of Eq. (3.20). We denote the corresponding martingale measure by $Q_0$. With respect to $Q_0$, $S_t$ is a jump–diffusion. Its generator $\mathcal{L}_Q^0(t)$ is given by (3.21) where $\psi(t, x, z) = 0$. Therefore $(\frac{\partial}{\partial t} + \mathcal{L}_Q^0(t) - r)F$ coincides with $A(t)F$ given by (3.17). Consequently Merton’s integro-differential equation (3.18) is rewritten as
\[
\left( \frac{\partial}{\partial t} + \mathcal{L}_Q^0(t) - r \right) F = 0,
\] (3.22)
\[
F(T, x) = h(x).
\]

Theorem 3.8. Let $S_t^{(x,x)}$ be the solution of Eq. (3.1) starting from $x$ at time $s$. If it is a geometric Lévy process, then
\[
F(t, x) := E^{Q_0}[e^{-r(T-t)} h(S_T^{(x,x)})]
\] (3.23)
is a $C^{1,2}$-function. It is the unique $C^{1,2}$-solution of the integro-differential equation (3.22).

Proof. The solution is represented as $S_t^{(x,x)} = x \exp(Z_{t} - Z_t)$, where $Z_t$ is a homogeneous Lévy process. Since the law of the Lévy process $Z_T - Z_t - r(T - t)$ has a smooth density $p_{t,T}(z)$, the law of $e^{r(T-t) S_T^{(x,x)}}$ has also a density given by $p_{t,T}(\log z - \log x)z^{-1}$. It is a $C^{1,2}$ function of $(t, x)$. The function $F$ of (3.23) is then written as $F(t, x) = \int h(z)p_{t,T}(\log z - \log x)z^{-1}dz$. Consequently it is a $C^{1,2}$-function.

We will show that the function $F$ satisfies Eq. (3.22). Let $S_t, t \geq 0$ be any solution starting at time 0. Observe that $\tilde{P}_t = e^{-rt} F(t, S_t)$ is a $Q_0$-martingale because we have for any $t > s$,
\[
E^{Q_0}[e^{-rt} F(t, S_t)|\mathcal{F}_s] = E^{Q_0}[e^{-rT} h(S_T)|\mathcal{F}_s],
\]
in view of the Markov property of $S_t$. Apply Itô’s formula to $e^{-rt} F(t, S_t)$ under the measure $Q_0$. Then we have, similarly as (3.15)
\[
d\tilde{P}_t = e^{-rt} \left( \frac{\partial}{\partial t} + \mathcal{L}_Q^0(t) - r \right) F(t, S_t)dt + dM_t^0,
\]
where $M_t^0$ is a local $Q_0$-martingale. Since $\tilde{P}_t$ is also a $Q_0$-martingale, the drift part of the above is equal to 0 a.s. Then we get $(\frac{\partial}{\partial t} + \mathcal{L}_Q^0(t) - r)F = 0$.

We show the uniqueness of the solution. Let $F'(t, x)$ be any $C^{1,2}$-solution of Eq. (3.22) with linear growth. Set $\tilde{P}_t' = e^{-rt} F'(t, S_t)$. Applying Itô’s formula and noting that $F'$ is a solution of Eq. (3.22), we see that $\tilde{P}_t'$ is a local martingale. It holds $\tilde{P}_T' = h(S_T)$. Then we have $\tilde{P}_T' = \tilde{P}_T$. Since $\tilde{P}_T$ and $\tilde{P}_T'$ are local martingales, we have $\tilde{P}_t = \tilde{P}_T'$ for any $t$. Therefore we get $F(t, S_t) = F'(t, S_t)$ a.s. for any $t$. Since the support of the law of $S_t$ is the whole space $R^+ = (0, \infty)$, we get $F(t, x) = \tilde{F}(t, x)$ for all $x \in R^+$. The proof is complete. $\Box$
Remark. In the case where the coefficients $b, \sigma, g$ depend on $x$, we cannot apply the above theorem directly. However, if these coefficients are smooth functions, say $C^\infty$, the solution $S_u^{(t,x)}$ is a $C^\infty$-function of $x$, in view of the theory of stochastic flows [23]. Thus if the value function $h$ in (3.23) is a $C^2$-function, $F(t,x)$ is also a $C^{1,2}$-function and the assertion of the lemma is valid.

If the value function $h$ is not smooth as for call or put options, we may need Malliavin calculus on the Wiener–Poisson space [6]. Discussions are close to Watanabe on the Wiener space [41]. Here we give a very rough argument. The solution $X = S_u^{(t,x)}$ is a smooth random variable $(\in D_\infty)$ for any $t < u, x$. Let $\Phi$ be a tempered distribution. We may define the composition $\Phi \circ S_u^{(t,x)}$ as a generalized random variable in $D'_\infty$ (dual space of $D_\infty$) with parameter $s, t, x$. Further it can be shown that $\langle \Phi \circ S_u^{(t,x)}, G \rangle$ is infinitely differentiable with respect to $x$, if $G \in D_\infty$. In particular $\langle \Phi \circ S_u^{(t,x)} , 1 \rangle$ is smooth with respect to $x$. It coincides with the expectation $E[\Phi \circ S_u^{(t,x)}]$ if $\Phi$ is a function of polynomial growth. Details will be discussed elsewhere [25].

Finally we shall consider other martingale measures $Q \in \mathcal{Q}$. Let $F(t, x), 0 \leq t \leq T, x \in R^+$ be a $C^{1,2}$-function such that $F(T, x) = h(x)$ and let $P_t = F(t, S_t)$. Apply Itô’s formula again to $e^{-rt} F(t, S_t)$. With respect to the measure $Q$, $\tilde{P}_t = e^{-rt} P_t$ satisfies similarly as in (3.15),

$$d\tilde{P}_t = e^{-rt} \left( \frac{\partial}{\partial t} + \mathcal{L}_Q (t) - r \right) F(t, S_t) dt + dM_t,$$

where $\mathcal{L}_Q$ is the operator defined by (3.21) and $M_t$ is a local $Q$-martingale. Then $\tilde{P}_t$ is a local $Q$-martingale if and only if $\left( \frac{\partial}{\partial t} + \mathcal{L}_Q (t) - r \right) F = 0$.

The following is verified similarly as Theorem 3.3.

**Theorem 3.9.** Let $Q$ be any martingale measure in $\mathcal{Q}$. Then the stochastic process $\tilde{P}_t = e^{-rt} F(t, S_t)$ is a $Q$-martingale if and only if $F$ satisfies

$$\left( \frac{\partial}{\partial t} + \mathcal{L}_Q (t) - r \right) F = 0. \tag{3.24}$$

Further, if $F$ satisfies (3.24), then the triple $(B^0_t, S_t, F(t, S_t))$ is non-arbitrage.

The above theorem tells us that in case of a jump–diffusion, there are infinitely many pricing functions $F$ for which the triple $(B^0_t, S_t, F(t, S_t))$ is non-arbitrage. For any martingale measure $Q \in \mathcal{Q}$, a solution of the equation $\left( \frac{\partial}{\partial t} + \mathcal{L}_Q (t) - r \right) F = 0$ with the terminal condition $F(T, x) = h(x)$ gives us a non-arbitrage price of the European option with pay-off function $h$. On the other hand, in case of a diffusion, such $F$ is unique and it is the solution of the Black–Scholes equation, in view of Theorem 3.3.

Concerning the choice of one out of many martingale measures $Q$, there are some alternative proposals: super-hedging, utility maximization, minimal variance hedging, use of a minimal entropy martingale measure and so on. For these subjects, we refer to R. Cont-P. Tankov [2] and the references therein.

**Acknowledgement**

Marc Yor read carefully preliminary versions of this paper and gave me many useful comments and advices. I should like to express my gratitude to him.

**References**


