

# Chapter 1

## From Diffusions to Semimartingales

This chapter is a quick review of the theory of semimartingales, these processes being those for which statistical methods are considered in this book.

A *process* is a collection  $X = (X_t)$  of random variables with values in the Euclidean space  $\mathbb{R}^d$  for some integer  $d \geq 1$ , and indexed on the half line  $\mathbb{R}_+ = [0, \infty)$ , or a subinterval of  $\mathbb{R}_+$ , typically  $[0, T]$  for some real  $T > 0$ . The distinctive feature however is that all these variables are defined on *the same probability space*  $(\Omega, \mathcal{F}, \mathbb{P})$ . Therefore, for any outcome  $\omega \in \Omega$  one can consider the *path* (or “trajectory”), which is the function  $t \mapsto X_t(\omega)$ , and  $X$  can also be considered as a single random variable taking its values in a suitable space of  $\mathbb{R}^d$ -valued functions on  $\mathbb{R}_+$  or on  $[0, T]$ .

In many applications, including the modeling of financial data, the index  $t$  can be interpreted as *time*, and an important feature is the way the process evolves through time. Typically an observer knows what happens up to the current time  $t$ , that is, (s)he observes the path  $s \mapsto X_s(\omega)$  for all  $s \in [0, t]$ , and wants to infer what will happen later, after time  $t$ . In a mathematical framework, this amounts to associating the “history” of the process, usually called the *filtration*. This is the increasing family  $(\mathcal{F}_t)_{t \geq 0}$  of  $\sigma$ -fields associated with  $X$  in the following way: for each  $t$ ,  $\mathcal{F}_t$  is the  $\sigma$ -field generated by the variables  $X_s$  for  $s \in [0, t]$  (more precise specifications will be given later). Therefore, of particular interest is the law of the “future” after time  $t$ , that is of the family  $(X_s : s > t)$ , conditional on the  $\sigma$ -field  $\mathcal{F}_t$ .

In a sense, this amounts to specifying the dynamics of the process, which again is a central question in financial modeling. If the process were not random, that would consist in specifying a differential equation governing the time evolution of our quantity of interest, or perhaps a non-autonomous differential equation where  $dX_t = f(t, X_s : s \leq t) dt$  for a function  $f$  depending on time and on the whole “past” of  $X$  before  $t$ . In a random setting, this is replaced by a “stochastic differential equation.”

Historically, it took quite a long time to come up with a class of processes large enough to account for the needs in applications, and still amenable to some simple calculus rules. It started with the Brownian motion, or Wiener process, and then processes with independent and stationary increments, now called Lévy processes after P. Lévy, who introduced and thoroughly described them. Next, martingales were considered, mainly by J. L. Doob, whereas K. Itô introduced (after W. Feller and W. Doeblin) the stochastic differential equations driven by Brownian motions, and also by Poisson random measures. The class of semimartingales was finalized by P.-A. Meyer in the early 1970s only.

In many respects, this class of processes is the right one to consider by someone interested in the dynamics of the process in the sense sketched above. Indeed, this is the largest class with respect to which stochastic integration is possible if one wants to have something like the dominated (Lebesgue) convergence theorem. It allows for relatively simple rules for stochastic calculus. Moreover, in financial mathematics, it also turns out to be the right class to consider, because a process can model the price of an asset in a fair market where no arbitrage is possible only if it is a semimartingale. This certainly is a sufficient motivation for the fact that, in this book, we only consider this type of process for modeling prices, including exchange rates or indices, and interest rates.

Now, of course, quite a few books provide extensive coverage of semimartingales, stochastic integration and stochastic calculus. Our aim in this chapter is not to duplicate any part of those books, and in particular not the proofs therein: the interested reader should consult one of them to get a complete mathematical picture of the subject, for example, Karatzas and Shreve (1991) or Revuz and Yor (1994) for continuous processes, and Dellacherie and Meyer (1982) or Jacod and Shiryaev (2003) for general ones. Our aim is simply to introduce semimartingales and the properties of those which are going to be of constant use in this book, in as simple a way as possible, starting with the most commonly known processes, which are the Brownian motion or Wiener process and the diffusions. We then introduce Lévy processes and Poisson random

measures, and finally arrive at semimartingales, presented as a relatively natural extension of Lévy processes.

## 1.1 Diffusions

### 1.1.1 The Brownian Motion

The *Brownian motion* (or *Wiener process*), formalized by N. Wiener and P. Lévy, has in fact been used in finance even earlier, by T. N. Thiele and L. Bachelier, and for modeling the physical motion of a particle by A. Einstein. It is the simplest continuous-time analogue of a random walk.

Mathematically speaking, the one-dimensional Brownian motion can be specified as a process  $W = (W_t)_{t \geq 0}$ , which is Gaussian (meaning that any finite family  $(W_{t_1}, \dots, W_{t_k})$  is a Gaussian random vector), centered (i.e.  $\mathbb{E}(W_t) = 0$  for all  $t$ ), and with the covariance structure

$$\mathbb{E}(W_t W_s) = t \wedge s \tag{1.1}$$

where the notation  $t \wedge s$  means  $\min(t, s)$ . These properties completely characterize the law of the process  $W$ , by Kolmogorov's Theorem, which allows for the definition of a stochastic process through its finite-dimensional distributions, under conditions known as consistency of the finite-dimensional distributions. And, using for example the Kolmogorov continuity criterion (since  $\mathbb{E}(|W_{t+s} - W_s|^4) = 3s^2$  for all nonnegative  $t$  and  $s$ ), one can “realize” the Brownian motion on a suitable probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  as a process having *continuous paths*, i.e.  $t \mapsto W_t(\omega)$  is continuous and with  $W_0(\omega) = 0$  for all outcomes  $\omega$ . So we will take the view that a Brownian motion always starts at  $W_0 = 0$  and has continuous paths.

One of the many fundamental properties of Brownian motion is that it is a *Lévy process*, that is it starts from 0 and has independent and stationary increments: for all  $s, t \geq 0$  the variable  $W_{t+s} - W_t$  is independent of  $(W_r : r \leq t)$ , with a law which only depends on  $s$ : here, this law is the normal law  $\mathcal{N}(0, s)$  (centered with variance  $s$ ). This immediately follows from the above definition. However, a converse is also true: any Lévy process which is centered and continuous is of the form  $\sigma W$  for some constant  $\sigma \geq 0$ , where  $W$  is a Brownian motion.

Now, we need two extensions of the previous notion. The first one is straightforward: a  $d$ -dimensional Brownian motion is an  $\mathbb{R}^d$ -valued process  $W = (W_t)_{t \geq 0}$  with components  $W_t^i$  for  $i = 1, \dots, d$  (we keep the same notation  $W$  as in the one-dimensional case), such that each com-

ponent process  $W^i = (W_t^i)_{t \geq 0}$  is a one-dimensional Brownian motion, and all components are *independent* processes. Equivalently,  $W$  is a centered continuous Gaussian process with  $W_0 = 0$ , and with the following covariance structure:

$$\mathbb{E}(W_t^i W_s^j) = \begin{cases} t \wedge s & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases} \quad (1.2)$$

A  $d$ -dimensional Brownian motion retains the nice property of being a Lévy process.

The second extension is slightly more subtle, and involves the concept of a general *filtered probability space*, denoted by  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . Here  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space, equipped with a *filtration*  $(\mathcal{F}_t)_{t \geq 0}$ : this is an increasing family of sub- $\sigma$ -fields  $\mathcal{F}_t$  of  $\mathcal{F}$  (that is,  $\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}$  when  $t \leq s$ ), which is right-continuous (that is  $\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s$ ). The right-continuity condition appears for technical reasons, but is in fact an essential requirement. Again,  $\mathcal{F}_t$  can be viewed as the amount of information available to an observer up to (and including) time  $t$ .

We say that a process  $X$  is *adapted* to a filtration  $(\mathcal{F}_t)$ , or  $(\mathcal{F}_t)$ -adapted, if each variable  $X_t$  is  $\mathcal{F}_t$ -measurable. The *filtration generated by a process*  $X$  is the smallest filtration with respect to which  $X$  is adapted. It is denoted as  $(\mathcal{F}_t^X)$ , and can be expressed as follows:

$$\mathcal{F}_t^X = \bigcap_{s > t} \sigma(X_r : r \in [0, s])$$

(this is right-continuous by construction).

We suppose that the reader is familiar with the notion of a *martingale* (a real process  $M = (M_t)_{t \geq 0}$  is a martingale on the filtered space if it is adapted, if each variable  $M_t$  is integrable and if  $\mathbb{E}(M_{t+s} \mid \mathcal{F}_t) = M_t$  for  $s, t \geq 0$ ), and also with the notion of a *stopping time*: a  $[0, \infty]$ -valued random variable  $\tau$  is a stopping time if it is possible to tell, for any  $t \geq 0$ , whether the event that  $\tau$  has occurred before or at time  $t$  is true or not, on the basis of the information contained in  $\mathcal{F}_t$ ; formally, this amounts to saying that the event  $\{\tau \leq t\}$  belongs to  $\mathcal{F}_t$ , for all  $t \geq 0$ . Likewise,  $\mathcal{F}_\tau$  denotes the  $\sigma$ -field of all sets  $A \in \mathcal{F}$  such that  $A \cap \{\tau \leq t\} \in \mathcal{F}_t$  for all  $t \geq 0$ , and it represents the information available up to (and including) time  $\tau$ .

A process  $X$  is called *progressively measurable* if for any  $t$  the function  $(\omega, s) \mapsto X_s(\omega)$  is  $\mathcal{F}_t \otimes \mathcal{B}([0, t])$ -measurable on  $\Omega \times [0, t]$ ; here,  $\mathcal{B}([0, t])$  denotes the Borel  $\sigma$ -field of the interval  $[0, t]$ , that is the  $\sigma$ -field generated by the open sets of  $[0, t]$ . Then, given a stopping time  $\tau$  and a progressively measurable process  $X_t$ , the variable  $X_\tau 1_{\{\tau < \infty\}}$  is  $\mathcal{F}_\tau$ -measurable;

moreover one can define the new process  $X_{\tau \wedge t}$ , or  $X$  stopped at  $\tau$ , and this process is again adapted. We note the following useful property: if  $M$  is a martingale and  $\tau_1$  and  $\tau_2$  are two stopping times such that  $0 \leq \tau_1 \leq \tau_2 \leq T$  a.s. (almost surely), then  $\mathbb{E}(M_{\tau_2} | \mathcal{F}_{\tau_1}) = M_{\tau_1}$ .

Coming back to Brownian motion, we say that a  $d$ -dimensional process  $W = (W^i)_{1 \leq i \leq d}$  is a *Brownian motion* on the filtered space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , or an  $(\mathcal{F}_t)$ -Brownian motion, if it satisfies the following three conditions:

1. It has continuous paths, with  $W_0 = 0$ .
2. It is adapted to the filtration  $(\mathcal{F}_t)$ .
3. For all  $s, t \geq 0$  the variable  $W_{t+s} - W_t$  is independent of the  $\sigma$ -field  $\mathcal{F}_t$ , with centered Gaussian law  $\mathcal{N}(0, sI_d)$  ( $I_d$  is the  $d \times d$  identity matrix).

It turns out that a Brownian motion in the previously stated restricted sense is an  $(\mathcal{F}_t^W)$ -Brownian motion. This property is almost obvious: it would be obvious if  $\mathcal{F}_t^W$  were  $\sigma(W_r : r \in [0, t])$ , and its extension comes from a so-called 0–1 law which asserts that, if an event is in  $\mathcal{F}_s^W$  for all  $s > t$  and is also independent of  $\mathcal{F}_v^W$  for all  $v < t$ , then its probability can only equal 0 or 1.

Another characterization of the Brownian motion, particularly well suited to stochastic calculus, is the *Lévy Characterization Theorem*. Namely, a continuous  $(\mathcal{F}_t)$ -adapted process  $W$  with  $W_0 = 0$  is an  $(\mathcal{F}_t)$ -Brownian motion if and only if it satisfies

$$\text{the processes } W_t^i \text{ and } W_t^i W_t^j - \delta_{ij}t \text{ are } (\mathcal{F}_t) \text{ - martingales} \quad (1.3)$$

where  $\delta_{ij} = 1$  if  $i = j$  and 0 otherwise denotes the Kronecker symbol. The necessary part is elementary; the sufficient part is more difficult to prove.

Finally, we mention two well known and important properties of the paths of a one-dimensional Brownian motion:

- *Lévy modulus of continuity*: almost all paths satisfy, for any interval  $I$  of positive length,

$$\limsup_{r \rightarrow 0} \frac{1}{\sqrt{r \log(1/r)}} \sup_{s, t \in I, |s-t| \leq r} |W_t - W_s| = \sqrt{2}. \quad (1.4)$$

- *Law of iterated logarithm:* for each finite stopping time  $T$ , almost all paths satisfy

$$\limsup_{r \rightarrow 0} \frac{1}{\sqrt{r \log \log(1/r)}} (W_{T+r} - W_T) = \sqrt{2}. \quad (1.5)$$

By symmetry, the  $\liminf$  in the law of iterated logarithm is equal to  $-\sqrt{2}$ . Despite the appearances, these two results are not contradictory, because of the different position of the qualifier “almost all.” These facts imply that, for any  $\rho < 1/2$ , almost all paths are locally Hölder with index  $\rho$ , but nowhere Hölder with index  $1/2$ , and *a fortiori* nowhere differentiable.

### 1.1.2 Stochastic Integrals

A second fundamental concept is stochastic integration. The paths of a one-dimensional Brownian motion  $W$  being continuous but nowhere differentiable, *a priori* the “differential”  $dW_t$  makes no sense, and neither does the expression  $\int_0^t H_s dW_s$ . However, suppose that  $H$  is a “simple,” or piecewise constant, process of the form

$$H_t = \sum_{i \geq 1} H_{t_{i-1}} 1_{[t_{i-1}, t_i)}(t), \quad (1.6)$$

where  $0 = t_0 < t_1 < \dots$  and  $t_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Then it is natural to set

$$\int_0^t H_s dW_s = \sum_{i \geq 1} H_{t_{i-1}} (W_{t \wedge t_i} - W_{t \wedge t_{i-1}}). \quad (1.7)$$

This would be the usual integral if  $t \mapsto W_t$  were the distribution function of a (signed) measure on  $[0, t]$ , which it is of course not. Nevertheless it turns out that, due to the martingale properties (1.3) of  $W$ , this “integral” can be extended to all processes  $H$  having the following properties:

$$\begin{aligned} &H \text{ is progressively measurable, and not too big,} \\ &\text{in the sense that } \int_0^t H_s^2 ds < \infty \text{ for all } t. \end{aligned} \quad (1.8)$$

The extension is still denoted as  $\int_0^t H_s dW_s$  or, more compactly, as  $H \bullet W_t$ . It is called a *stochastic integral*, which emphasizes the fact that it is not a Stieltjes ( $\omega$ -wise) integral. In particular, it is only defined “up to a null set,” meaning that another variable  $Y$  is also a version of the stochastic integral if and only if we have  $Y = H \bullet W_t$  a.s. So we emphasize that *every* statement about a stochastic integral variable  $H \bullet X_t$  or process  $(H \bullet X_t)_{t \geq 0}$  is *necessarily* true “up to a null set”

only, although for simplicity we usually omit the qualifier “almost surely” (exactly as, for the conditional expectation, one simply writes an equality such as  $\mathbb{E}(Y \mid \mathcal{G}) = Z$  without explicitly mentioning “a.s.”).

Stochastic integrals enjoy the following properties:

- the process  $(H \bullet W_t)_{t \geq 0}$  is a continuous (local) martingale starting at 0
- the process  $(H \bullet W_t)^2 - \int_0^t H_s^2 ds$  is a (local) martingale
- the map  $H \mapsto H \bullet W$  is linear
- We have a “dominated convergence theorem”: if  $H^n \rightarrow H$  pointwise and  $|H^n| \leq H'$  with  $H'$  as in (1.8), then  $H^n \bullet W \xrightarrow{\text{u.c.P.}} H \bullet W$ .

In this statement, two notions need some explanation. First,  $\xrightarrow{\text{u.c.P.}}$  stands for “local uniform convergence in probability,” that is we write  $X^n \xrightarrow{\text{u.c.P.}} X$  if for all  $t$  we have  $\sup_{s \leq t} |X_s^n - X_s| \xrightarrow{\mathbb{P}} 0$  (convergence in probability). Second, a *local martingale* is a process  $M$  for which there exists an increasing sequence of stopping times  $T_n$  with infinite limit (called a “localizing sequence”), such that each “stopped” process  $t \mapsto M_{t \wedge T_n}$  is a martingale. In other words,  $M$  behaves like a martingale up to suitably chosen stopping times; martingales are local martingales but the converse is not true.

Let us also mention that the first statement (1.9) is shorthand for the more correct “one can find versions of the stochastic integrals  $H \bullet W_t$  such that almost all paths  $t \mapsto H \bullet W_t$  are continuous and start at 0, and further  $(H \bullet W_t)_{t \geq 0}$  is a local martingale.” And, as mentioned before, the third statement in (1.9) is true “up to null sets.”

More generally, let  $W$  be a  $d$ -dimensional Brownian motion. Then one can integrate “componentwise” a  $d$ -dimensional process  $H = (H^i)_{1 \leq i \leq d}$  whose components each satisfy (1.8), thus getting the following one-dimensional process:

$$H \bullet W_t = \int_0^t H_s dW_s = \sum_{i=1}^d \int_0^t H_s^i dW_s^i.$$

We still have the properties (1.9), with the norm of  $H^n$  instead of the absolute value in the fourth property. Moreover if  $H$  and  $K$  are two integrable processes, the process

$$(H \bullet W_t)(K \bullet W_t) - \int_0^t \sum_{i=1}^d H_s^i K_s^i ds \text{ is a local martingale.} \quad (1.10)$$

Before stating the “change of variable” formula for stochastic integrals, we give our first (restricted) definition of a semimartingale:

**Definition 1.1.** *A one-dimensional continuous Itô semimartingale (also called a “generalized diffusion,” or a “Brownian semimartingale” sometimes) is an adapted process  $X$  which can be written as*

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \quad (1.11)$$

where  $W$  is a  $q$ -dimensional Brownian motion, and  $\sigma$  is a  $q$ -dimensional process, integrable in the above sense (i.e., its components satisfy (1.8)), and  $b = (b_t)_{t \geq 0}$  is another progressively measurable process such that  $\int_0^t |b_s| ds < \infty$  for all  $t$ .

A  $d$ -dimensional continuous Itô semimartingale is a process whose each one of the  $d$  components is a continuous Itô semimartingale.

If  $X$  is as above, one can integrate a process  $K$  with respect to it, by setting

$$K \bullet X_t = \int_0^t K_s dX_s = \int_0^t K_s b_s ds + \int_0^t K_s \sigma_s dW_s$$

(with an obvious interpretation when  $X$  and  $K$  are  $q$ -dimensional). We need  $K\sigma$  to be integrable with respect to  $W$ , as in (1.8), and also  $Kb$  to be integrable with respect to the Lebesgue measure on each finite interval  $[0, t]$  (this integral is an “ordinary,” or  $\omega$ -wise, integral). A precise description of all processes  $K$  which can thus be integrated is somewhat complicated, but in any case we can integrate all processes  $K$  which are progressively measurable and *locally bounded* (meaning we have  $|K_s| \leq n$  for any  $0 < s \leq T_n$ , where  $T_n$  is a sequence of stopping times increasing to  $\infty$ ); note that no condition on  $K_0$  is implied. In this case, the integral process  $K \bullet X$  is also a continuous Itô semimartingale.

The last process on the right hand side of (1.11) is called the *continuous martingale part* of  $X$ , although it usually is a local martingale only, and it is denoted as  $X^c$  (with components  $X^{i,c}$  in the multidimensional case). We also associate the family of processes, for  $j, l = 1, \dots, q$  and  $q$  the dimension of  $X$ :

$$C_t^{jl} = \sum_{i=1}^d \int_0^t \sigma_s^{ji} \sigma_s^{li} ds. \quad (1.12)$$

Another notation for  $C_t^{jl}$  is  $\langle X^{j,c}, X^{l,c} \rangle_t$  and it is called the *quadratic variation-covariation process*. From this formula, the  $q^2$ -dimensional pro-

cess  $C = (C^{jl})_{1 \leq j, l \leq q}$  takes its values in the cone of all positive semi-definite symmetric  $q \times q$  matrices, and is continuous in  $t$  and increasing (in  $t$  again) for the strong order in this cone (that is,  $C_t - C_s$  is also positive semi-definite if  $t \geq s$ ). Note also the following obvious but important fact:

If  $X = W$  is a Brownian motion, then  $X^c = W$  and  $C_t^{jl} = \delta_{jl}t$ .

This definition of the quadratic variation is based upon the definition (1.11). However, there are two other characterizations. First, one can rewrite (1.10) as

$$X^{j,c} X^{l,c} - C^{jl} \text{ is a local martingale.} \quad (1.13)$$

Conversely,  $C$  is the *unique* (up to null sets) adapted continuous process, starting at  $C_0 = 0$ , with path  $t \mapsto C_t^{jl}$  having finite variation over compact intervals, and such that (1.13) holds.

Second, the name “quadratic variation” comes from the following property: let  $((t(n, i) : i \geq 0) : n \geq 1)$  be a sequence of subdivisions on  $\mathbb{R}_+$  with meshes going to 0. This means that for each  $n$  we have  $t(n, 0) = 0 < t(n, 1) < \dots$  and  $t(n, i) \rightarrow \infty$  as  $i \rightarrow \infty$ , and also  $\sup_{i \geq 1} (t(n, i) - t(n, i-1)) \rightarrow 0$  as  $n \rightarrow \infty$ . Then we have

$$\sum_{i \geq 1: t(n, i) \leq t} \left( X_{t(n, i)}^j - X_{t(n, i-1)}^j \right) \left( X_{t(n, i)}^l - X_{t(n, i-1)}^l \right) \xrightarrow{\mathbb{P}} C_t^{jl}. \quad (1.14)$$

Historically this is the way the quadratic variation has been introduced, indeed as a tool for defining stochastic integrals. We will give below a (simple) proof of this property, deduced from Itô’s formula, and for general semimartingales. The reason for giving the proof is that from an applied viewpoint (and especially for financial applications) (1.14) is an important property in high-frequency statistics: the left hand side, say when  $j = l = 1$ , is the so-called *realized quadratic variation*, or realized volatility, of the component  $X^1$ , along the observation times  $t(n, i)$  for  $i \geq 0$ , whereas the process  $C^{11}$  is what is called “integrated (squared) volatility.”

We are now ready to state the change of variable formula, more commonly called *Itô’s formula*: for any  $\mathcal{C}^2$  real-valued function  $f$  on  $\mathbb{R}^d$  (twice continuously differentiable) and any  $d$ -dimensional continuous Itô semimartingale  $X = (X^j)_{1 \leq j \leq d}$ , the process  $Y = f(X)$  is also a continuous Itô semimartingale; moreover, if  $f'_i$  and  $f''_{ij}$  denote the first and

second partial derivatives of  $f$ , we have

$$\begin{aligned}
 f(X_t) &= f(X_0) + \sum_{i=1}^d \int_0^t f'_i(X_s) dX_s^i \\
 &\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t f''_{ij}(X_s) d\langle X^{i,c}, X^{j,c} \rangle_s .
 \end{aligned}
 \tag{1.15}$$

Note that the processes  $f'_i(X_s)$  and  $f''_{ij}(X_s)$  are continuous, hence locally bounded, and adapted; so the first integrals in the right hand side above are stochastic integrals with respect to the Itô semimartingales  $X^i$ , and the second integrals are ordinary (Stieltjes) integrals with respect to the functions  $t \mapsto C_t^{ij} = \langle X^{i,c}, X^{j,c} \rangle_t$ , which are absolutely continuous by (1.12).

When  $X^c = 0$ , that is, when the functions  $t \mapsto X_t^i$  are absolutely continuous, the last sum in (1.15) vanishes, and the formula reduces to the usual change of variable formula (then of course  $f$  being  $C^1$  would be sufficient). The fact that in general this additional last term is present is one of the key observations made by K. Itô.

### 1.1.3 A Central Example: Diffusion Processes

As the other name “generalized diffusion processes” for continuous Itô semimartingales suggests, the main examples of such processes are diffusions processes. Historically speaking, they were the first relatively general semimartingales to be considered, and they play a central role in modeling, in the physical sciences and in finance, although in many cases they can be far from sufficient to account for all encountered empirical features of the processes being measured.

Going from general to particular, one can characterize *diffusions* as those continuous Itô semimartingales which are Markov processes. These may be homogeneous or not, and for simplicity we only consider the homogeneous case below, since they are by far the most common ones. More specifically, following for example Çinlar and Jacod (1981), if a continuous  $d$ -dimensional Itô semimartingale of the form (1.11) is a homogeneous Markov process, then the two random “infinitesimal coefficients”  $b_t$  and  $\sigma_t$  take the form

$$b_t = b(X_t), \quad \sigma_t = \sigma(X_t),$$

where  $b = (b^i)_{1 \leq i \leq d}$  and  $\sigma = (\sigma^{ij})_{1 \leq i \leq d, 1 \leq j \leq d}$  are functions on  $\mathbb{R}^d$ . That

is,

$$X_t = X_0 + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dW_s^j \quad (1.16)$$

or, componentwise,

$$X_t^i = X_0^i + \int_0^t b^i(X_s) ds + \sum_{j=1}^q \int_0^t \sigma^{ij}(X_s) dW_s^j, \quad \forall i = 1, \dots, d.$$

Now, the law of a Markov processes is also characterized by the law of its initial value  $X_0$  and its transition semi-group  $(\mathcal{P}_t)_{t \geq 0}$  (defined as the operator which returns  $\mathcal{P}_t f(x) = \mathbb{E}(f(X_t) | X_0 = x)$  when applied to any Borel bounded test function  $f$ , and  $x \in \mathbb{R}^d$ ), and in turn the semi-group is characterized by its infinitesimal generator (in general an unbounded operator defined on a suitable domain). Whereas there is no hope in general to have an explicit expression for the semi-group, one can easily compute the infinitesimal generator, at least when the test function is  $C^2$ . Namely, with the notation  $c(x) = \sigma(x)\sigma(x)^*$  (where  $\sigma^*$  is the transpose of  $\sigma$ ), we observe that, by virtue of (1.12),  $\langle X^{i,c}, X^{j,c} \rangle_t = C_t^{ij} = \int_0^t c_s^{ij} ds$ . Then Itô's formula (1.15) implies that

$$\begin{aligned} M_t^f &= f(X_t) - f(X_0) - \sum_{i=1}^d \int_0^t b(X_s)^i f'_i(X_s) ds \\ &\quad - \frac{1}{2} \sum_{i,j=1}^d \int_0^t c(X_s)^{ij} f''_{ij}(X_s) ds \end{aligned}$$

is a local martingale. With the notation

$$\mathcal{A}f(x) = \sum_{i=1}^d b(x)^i f'_i(x) + \frac{1}{2} \sum_{i,j=1}^d c(x)^{ij} f''_{ij}(x), \quad (1.17)$$

we then have  $M_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{A}f(X_s) ds$ . If further  $f$  has compact support, say, and if the coefficients  $b$  and  $\sigma$  are locally bounded, then  $\mathcal{A}f$  is bounded. Hence  $M^f$  is a martingale and not only a local martingale, and by taking the expectation and using Fubini's Theorem we obtain

$$\mathcal{P}_t f(x) = f(x) + \int_0^t \mathcal{P}_s \mathcal{A}f(x) ds.$$

In other words,  $f$  belongs to the domain of the infinitesimal generator, which is  $\mathcal{A}f$  (as defined above) for such a function.

Of course, this does not fully specify the infinitesimal generator: for this we should exhibit its full domain and say how it acts on functions

which are in the domain but are not  $C^2$  with compact support (and the domain always contains plenty of such functions). But in the “good” cases, the complete infinitesimal generator is simply the closure of the operator  $\mathcal{A}$  acting on  $C^2$  functions with compact support as in (1.17).

Now, diffusions can also be viewed, and historically have been introduced, as solutions of *stochastic differential equation*, or SDE in short. Coming back to (1.16), a convenient way of writing it is in “differential form,” as follows, and with the convention  $Y = X_0$ :

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = Y \quad (1.18)$$

despite the fact that the differential  $dW_t$  is *a priori* meaningless. Now, we can consider (1.18) as an equation. The “solution” will be a  $d$ -dimensional process  $X$ , whereas  $W$  is a  $q$ -dimensional Brownian motion, and the following ingredients are given: the initial condition  $Y$  (an  $\mathbb{R}^d$ -valued random vector, most often a constant  $Y = x \in \mathbb{R}^d$ ) and the coefficients  $b$  and  $\sigma$  which are Borel functions on  $\mathbb{R}^d$ , with the suitable dimensions (they are respectively called the *drift* and *diffusion* coefficients).

The word “solution” for an SDE like (1.18) may have several different meanings. Here we consider the simplest notion, sometimes called solution-process or strong solution. Namely, we are given a  $q$ -dimensional  $(\mathcal{F}_t)$ -Brownian motion  $W$  and a  $d$ -dimensional  $\mathcal{F}_0$ -measurable variable  $Y$  on some filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , and a *solution* is a continuous adapted process  $X$  which satisfies (1.18), which is a shorthand way of writing (1.16) with  $X_0 = Y$ .

In particular, the integrals on the right side of (1.16) should make sense: so the functions  $b$  and  $\sigma$  should be measurable and not “too big”; more important, the process  $\sigma(X_t)$  should be progressively measurable, so (unless of course  $\sigma$  is constant) we need  $X$  itself to be progressively measurable, and since it is continuous in time this is the same as saying that it is adapted. Finally, it follows in particular that  $X_0$  is  $\mathcal{F}_0$ -measurable, which is the reason why we impose the  $\mathcal{F}_0$ -measurability to the initial condition  $Y$ .

Of course, as ordinary differential equations, not all SDEs have a solution. Let us simply mention that, if the two functions  $b$  and  $\sigma$  are *locally Lipschitz and with at most linear growth* on  $\mathbb{R}^d$ , then (1.18) has a solution, and furthermore this solution is unique: the uniqueness here means that any two solutions  $X$  and  $X'$  satisfy  $X_t = X'_t$  a.s. for all  $t$ , which is the best one can hope since in any case stochastic integrals are defined only up to null sets. An interesting feature is that, under the

previous assumptions on  $b$  and  $\sigma$ , existence and uniqueness hold for all initial conditions  $Y$ .

These Lipschitz and growth conditions are by far not the only available conditions implying existence and/or uniqueness. In fact, existence and/or uniqueness is somehow related to the fact that the operator  $\mathcal{A}$  defined by (1.17) for  $\mathcal{C}^2$  functions  $f$  with compact support can be extended as a *bona fide* infinitesimal generator.

**Example 1.2.** *The following one-dimensional example ( $d = q = 1$ ):*

$$dX_t = \kappa(\nu - X_t)dt + \eta\sqrt{X_t}dW_t, \quad X_0 = Y \quad (1.19)$$

where  $\kappa, \eta \in \mathbb{R}$  and  $\nu \in \mathbb{R}_+$  are given constants, is known as Feller's equation or in finance the Cox-Ingersoll-Ross (CIR) model, and it shows some of the problems that can occur. It does not fit the previous setting, for two reasons: first, because of the square root, we need  $X_t \geq 0$ ; so the natural state space of our process is not  $\mathbb{R}$ , but  $\mathbb{R}_+$ . Second, the diffusion coefficient  $\sigma(x) = \eta\sqrt{x}$  is not locally Lipschitz on  $[0, \infty)$ .

Let us thus provide some comments:

1. In the general setting of (1.18), the fact that the state space is not  $\mathbb{R}^d$  but a domain  $D \subset \mathbb{R}^d$  is not a problem for the formulation of the equation: the coefficients  $b$  and  $\sigma$  are simply functions on  $D$  instead of being functions on  $\mathbb{R}^d$ , and the solution  $X$  should be a  $D$ -valued process, provided of course that the initial condition  $Y$  is also  $D$ -valued. Problems arise when one tries to solve the equation. Even with Lipschitz coefficients, there is no guarantee that  $X$  will not reach the boundary of  $D$ , and here anything can happen, like the drift or the Brownian motion forcing  $X$  to leave  $D$ . So one should either make sure that  $X$  cannot reach the boundary, or specify what happens if it does (such as, how the process behaves along the boundary, or how it is reflected back inside the domain  $D$ ).
2. Coming back to the CIR model (1.19), and assuming  $Y > 0$ , one can show that with the state space  $D = (0, \infty)$  (on which the coefficients are locally Lipschitz), then the solution  $X$  will never reach 0 if and only if  $2\kappa\nu > \eta^2$ . Otherwise, it reaches 0 and uniqueness fails, unless we specify that the process reflects instantaneously when it reaches 0, and of course we need  $\kappa > 0$ .

All these problems are often difficult to resolve, and sometimes require *ad hoc* or model-specific arguments. In this book, when we have a diffusion we suppose in fact that it is well defined, and that those difficult

problems have been solved beforehand. Let us simply mention that the literature on this topic is huge, see for example Revuz and Yor (1994) and Karatzas and Shreve (1991).

## 1.2 Lévy Processes

As already said, a *Lévy process* is an  $\mathbb{R}^d$ -valued process starting at 0 and having stationary and independent increments. More generally:

**Definition 1.3.** *A filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  being given, a Lévy process relative to the filtration  $(\mathcal{F}_t)$ , or an  $(\mathcal{F}_t)$ -Lévy process, is an  $\mathbb{R}^d$ -valued process  $X$  satisfying the following three conditions:*

1. *Its paths are right-continuous with left limit everywhere (we say, a càdlàg process, from the French acronym “continu à droite, limité à gauche”), with  $X_0 = 0$ .*
2. *It is adapted to the filtration  $(\mathcal{F}_t)$ .*
3. *For all  $s, t \geq 0$  the variable  $X_{t+s} - X_t$  is independent of the  $\sigma$ -field  $\mathcal{F}_t$ , and its law only depends on  $s$ .*

These are the same as the conditions defining an  $(\mathcal{F}_t)$ -Brownian motion, except that the paths are càdlàg instead of continuous, and the laws of the increments are not necessarily Gaussian. In particular, we associate with  $X$  its jump process, defined (for any càdlàg process, for that matter) as

$$\Delta X_t = X_t - X_{t-} \tag{1.20}$$

where  $X_{t-}$  is the left limit at time  $t$ , and by convention  $\Delta X_0 = 0$ .

Regarding the nature of the sample paths of a Lévy process, note that the càdlàg property is important to give meaning to the concept of a “jump” defined in (1.20) where we need as a prerequisite left and right limits at each  $t$ . When a jump occurs at time  $t$ , being right-continuous means that the value after the jump,  $X_{t+}$ , is  $X_t$ . In finance, it means that by the time a jump has occurred, the price has already moved to  $X_t = X_{t+}$  and it is no longer possible to trade at the pre-jump value  $X_{t-}$ .

Although the process can have infinitely many jumps on a finite interval  $[0, T]$ , the càdlàg property limits the total number of jumps to be at most countable. It also limits the number of jumps larger than any fixed

value  $\varepsilon > 0$  on any interval  $[0, t]$  to be at most finite. Further, the paths of a Lévy process are stochastically continuous, meaning that  $X_s \xrightarrow{\mathbb{P}} X_t$  (convergence in probability) for all  $t > 0$ , and as  $s \rightarrow t$  (here,  $t$  is non-random). This condition does not make the paths of  $X$  continuous, but it excludes jumps at fixed times. All it says is that at any given *non-random* time  $t$ , the probability of seeing a jump at this time is 0.

### 1.2.1 The Law of a Lévy Process

One can look at Lévy processes *per se* without consideration for the connection with other processes defined on the probability space or with the underlying filtration. We then take the viewpoint of the law, or equivalently the family of all finite-dimensional distributions, that is the laws of any  $n$ -tuple  $(X_{t_1}, \dots, X_{t_n})$ , where  $0 \leq t_1 < \dots < t_n$ .

Because of the three properties defining  $X$  above, the law of  $(X_{t_1}, \dots, X_{t_n})$  is the same as the law of  $(Y_1, Y_1 + Y_2, \dots, Y_1 + \dots + Y_n)$ , where the  $Y_j$ 's are independent variables having the same laws as  $X_{t_j - t_{j-1}}$ . Therefore the law of the whole process  $X$  is completely determined, once the one-dimensional laws  $G_t = \mathcal{L}(X_t)$  are known.

Moreover,  $G_{t+s}$  is equal to the convolution product  $G_t * G_s$  for all  $s, t \geq 0$ , so the laws  $G_t$  have the very special property of being *infinitely divisible*: the name comes from the fact that, for any integer  $n \geq 1$ , we can write  $X_t = \sum_{j=1}^n Y_j$  as a sum of  $n$  i.i.d. random variables  $Y_j$  whose distribution is that of  $X_{t/n}$ ; equivalently, for any  $n \geq 1$ , the law  $G_t$  is the  $n$ -fold convolution power of some probability measure, namely  $G_{t/n}$  here. This property places a restriction on the possible laws  $G_t$ , and is called infinite divisibility. Examples include the Gaussian, gamma, stable and Poisson distributions. For instance, in the Gaussian case, any  $X_t \sim \mathcal{N}(m, v)$  can be written as the sum of  $n$  i.i.d. random variables  $Y_j \sim \mathcal{N}(m/n, v/n)$ .

Infinite divisibility implies that the characteristic function  $\widehat{G}_t(u) = \mathbb{E}(e^{iu^* X_t}) = \int e^{iu^* x} G_t(dx)$  (where  $u \in \mathbb{R}^d$  and  $u^* x$  is the scalar product) does not vanish and there exists a function  $\Psi : \mathbb{R}^d \rightarrow \mathbb{R}$ , called the characteristic exponent of  $X$ , such that the characteristic function takes the form

$$\widehat{G}_t(u) = \exp(t\Psi(u)), \quad u \in \mathbb{R}^d, \quad t > 0. \quad (1.21)$$

Indeed we have  $\widehat{G}_{t+s}(u) = \widehat{G}_t(u)\widehat{G}_s(u)$  and  $\widehat{G}_t(u)$  is càdlàg in  $t$ , which imply that the logarithm of  $\widehat{G}_t(u)$  is linear, as a function of  $t$ ; of course these are complex numbers, so some care must be taken to justify the statement.

In fact it implies much more, namely that  $\Psi(u)$  has a specific form given by the *Lévy-Khintchine* formula, which we write here for the variable  $X_1$ :

$$\widehat{G}_1(u) = \exp\left(u^*b - \frac{1}{2}u^*cu + \int(e^{iu^*x} - 1 - iu^*x1_{\{\|x\|\leq 1\}})F(dx)\right). \quad (1.22)$$

In this formula, the three ingredients  $(b, c, F)$  are as follows:

- $b = (b^i)_{i \leq d} \in \mathbb{R}^d$ ,
  - $c = (c^{ij})_{i,j \leq d}$  is a symmetric nonnegative matrix,
  - $F$  is a positive measure on  $\mathbb{R}^d$
- (1.23)
- with  $F(\{0\}) = 0$  and  $\int(\|x\|^2 \wedge 1) F(dx) < \infty$ .

and  $u^*cu = \sum_{i,j=1}^d u_i c_{ij} u_j$ . The integrability requirement on  $F$  in (1.23) is really two different constraints written in one: it requires  $\int_{\|x\|\leq 1} \|x\|^2 F(dx) < \infty$ , which limits the rate at which  $F$  can diverge near 0, and  $\int_{\|x\|\geq 1} F(dx) < \infty$ . These two requirements are exactly what is needed for the integral in (1.22) to be absolutely convergent, because  $e^{iu^*x} - 1 - iu^*x1_{\{\|x\|\leq 1\}} \sim (u^*x)^2$  as  $x \rightarrow 0$ , and  $e^{iu^*x} - 1$  is bounded.

We have written (1.2) at time  $t = 1$ , but because of (1.21), we also have it at any time  $t$ . *A priori*, one might think that the triple  $(b_t, c_t, F_t)$  associated with  $G_t$  would depend on  $t$  in a rather arbitrary way, but this is not so. We have, for any  $t \geq 0$ , and with the same  $(b, c, F)$  as in (1.22),

- $\widehat{G}_t(u) = e^{t\Psi(u)}$ ,
  - $\Psi(u) = u^*b - \frac{1}{2}u^*cu + \int(e^{iu^*x} - 1 - iu^*x1_{\{\|x\|\leq 1\}}) F(dx)$ ,
- (1.24)

and  $\Psi(u)$  is called the *characteristic exponent* of the Lévy process.

In other words, the law of  $X$  is completely characterized by the triple  $(b, c, F)$ , subject to (1.23), thus earning it the name *characteristic triple* of the Lévy process  $X$ . And we do have a converse: if  $(b, c, F)$  satisfies the conditions (1.23), then it is the characteristic triple of a Lévy process. The measure  $F$  is called the *Lévy measure* of the process,  $c$  is called the diffusion coefficient (for reasons which will be apparent later), and  $b$  is the drift (a slightly misleading terminology, as we will see later again). In Section 1.2.5 below, we will see that the different elements on the right-hand side of (1.22) correspond to specific elements of the canonical decomposition of a Lévy process in terms of drift, volatility, small jumps and big jumps.

We end this subsection by pointing out some moment properties of

Lévy processes. First, for any reals  $p > 0$  and  $t > 0$  we have

$$\mathbb{E}(\|X_t\|^p) < \infty \iff \int_{\{\|x\|>1\}} \|x\|^p F(dx) < \infty. \quad (1.25)$$

In particular the  $p$ th moments are either finite for all  $t$  or infinite for all  $t > 0$ , and they are all finite when  $F$  has compact support, a property which is equivalent to saying that the jumps of  $X$  are bounded, as we will see later.

Second, cumulants and hence moments of the distribution of  $X_t$  of integer order  $p$  can be computed explicitly using (1.24) by differentiation of the characteristic exponent and characteristic function. For example, in the one-dimensional case and when  $\mathbb{E}(|X_t|^n) < \infty$ , the  $n$ th cumulant and  $n$ th moment of  $X_t$  are

$$\begin{aligned} \kappa_{n,t} &= \frac{1}{i^n} \frac{\partial^n}{\partial u^n} (t\Psi(u))|_{u=0} = t\kappa_{n,1}, \\ \varphi_{n,t} &= \frac{1}{i^n} \frac{\partial^n}{\partial u^n} (e^{t\Psi(u)})|_{u=0}. \end{aligned}$$

The first four cumulants, in terms of the moments and of the centered moments  $\mu_{n,t} = \mathbb{E}((X_t - \phi_{1,t})^n)$ , are

$$\begin{aligned} \kappa_{1,t} &= \varphi_{1,t} = \mathbb{E}(X_t), \\ \kappa_{2,t} &= \mu_{2,t} = \varphi_{2,t} - \varphi_{1,t}^2 = \text{Var}(X_t), \\ \kappa_{3,t} &= \mu_{3,t} = \varphi_{3,t} - 3\varphi_{2,t}\varphi_{1,t} + 2\varphi_{1,t}^3, \\ \kappa_{4,t} &= \mu_{4,t} - 3\mu_{2,t}^2. \end{aligned}$$

In terms of the characteristics  $(b, c, F)$  of  $X$ , we have

$$\begin{aligned} \kappa_{1,t} &= t\left(b + \int_{|x|\geq 1} x F(dx)\right), \\ \kappa_{2,t} &= t\left(c + \int x^2 F(dx)\right), \\ \kappa_{p,t} &= t \int x^p F(dx) \text{ for all integers } p \geq 3. \end{aligned}$$

All infinitely divisible distributions with a non-vanishing Lévy measure are leptokurtic, that is  $\kappa_4 > 0$ . The skewness and excess kurtosis of  $X_t$ , when the third and fourth moments are finite, are

$$\begin{aligned} \text{skew}(X_t) &= \frac{\kappa_{3,t}}{\kappa_{2,t}^{3/2}} = \frac{\text{skew}(X_1)}{t^{1/2}} \\ \text{kurt}(X_t) &= \frac{\kappa_{4,t}}{\kappa_{2,t}^2} = \frac{\text{kurt}(X_1)}{t} \end{aligned}$$

which both increase as  $t$  decreases.

## 1.2.2 Examples

The Brownian motion  $W$  is a Lévy process, with characteristic triple  $(0, I_d, 0)$ . Another (trivial) example is the pure drift  $X_t = bt$  for a vector  $b \in \mathbb{R}^d$ , with the characteristic triple  $(b, 0, 0)$ . Then  $X_t = bt + \sigma W_t$  is also a Lévy process, with characteristic triple  $(b, c, 0)$ , where  $c = \sigma\sigma^*$ . Those are the only *continuous* Lévy processes – all others have jumps – and below we give some examples, starting with the simplest one. Note that the sum of two independent  $d$ -dimensional Lévy processes with triples  $(b, c, F)$  and  $(b', c', F')$  is also a Lévy process with triple  $(b', c+c', F+F')$  for a suitable number  $b''$ , so those examples can be combined to derive further ones.

**Example 1.4 (Poisson process).** *A counting process is an  $\mathbb{N}$ -valued process whose paths have the form*

$$N_t = \sum_{n \geq 1} 1_{\{T_n \leq t\}}, \quad (1.26)$$

where  $T_n$  is a strictly increasing sequence of positive times with limit  $+\infty$ . The usual interpretation is that the  $T_n$ 's are the successive arrival times of some kind of “events,” and  $N_t$  is the number of such events occurring within the interval  $[0, t]$ . The paths of  $N$  are piecewise constant, and increase (or jump) by 1. They are càdlàg by construction.

A Poisson process is a counting process such that the inter-arrival times  $S_n = T_n - T_{n-1}$  (with the convention  $T_0 = 0$ ) are an i.i.d. sequence of variables having an exponential distribution with intensity parameter  $\lambda$ . Using the memoryless property of the exponential distribution, it is easy to check that a Poisson process is a Lévy process, and  $N_t$  has a Poisson distribution with parameter  $\lambda t$ , that is,  $\mathbb{P}(N_t = n) = \exp(-\lambda t) \frac{(\lambda t)^n}{n!}$  for  $n \in \mathbb{N}$ . The converse is also easy: any Lévy process which is also a counting process is a Poisson process.

In particular,  $\mathbb{E}(N_t) = \lambda t$  so  $\lambda$  represents the expected events arrival rate per unit of time, and also  $\text{Var}(N_t) = \lambda t$ . The characteristic function of the Poisson random variable  $N_t$  is

$$\widehat{G}_t(u) = \exp(t\lambda(e^{iu} - 1)), \quad (1.27)$$

and the characteristic triple of  $N$  is  $(\lambda, 0, \lambda\varepsilon_1)$ , where  $\varepsilon_a$  stands for the Dirac mass sitting at  $a$ ; note that (1.27) matches the general formula (1.24). When  $\lambda = 1$  it is called the standard Poisson process.

Another property is important. Assume that  $N$  is an  $(\mathcal{F}_t)$ -Poisson process on the filtered space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ ; by this, we mean a Poisson

process which is also an  $(\mathcal{F}_t)$ -Lévy process. Because of the independence of  $N_{t+s} - N_t$  from  $\mathcal{F}_t$ , we have  $\mathbb{E}(N_{t+s} - N_t \mid \mathcal{F}_t) = \lambda s$ . In other words,

$$N_t - \lambda t \text{ is an } (\mathcal{F}_t)\text{-martingale.} \quad (1.28)$$

This is the analogue of the property (1.3) of the Brownian motion. And, exactly as for the Lévy characterization of the Brownian motion, we have the following: if  $N$  is a counting process satisfying (1.28), then it is an  $(\mathcal{F}_t)$ -Poisson process. This is called the Watanabe characterization of the Poisson process.

$N_t - \lambda t$  is called a compensated Poisson process. Note that the compensated Poisson process is no longer  $\mathbb{N}$ -valued. The first two moments of  $(N_t - \lambda t) / \lambda^{1/2}$  are the same as those of Brownian motion:  $\mathbb{E}((N_t - \lambda t) / \lambda^{1/2}) = 0$  and  $\text{Var}((N_t - \lambda t) / \lambda^{1/2}) = t$ . In fact, we have convergence in distribution of  $(N_t - \lambda t) / \lambda^{1/2}$  to Brownian motion as  $\lambda \rightarrow \infty$ .

Finally, if  $N_t$  and  $N'_t$  are two independent Poisson processes of intensities  $\lambda$  and  $\lambda'$ , then  $N_t + N'_t$  is a Poisson process of intensity  $\lambda + \lambda'$ .

**Example 1.5 (Compound Poisson process).** A compound Poisson process is a process of the form

$$X_t = \sum_{n \geq 1} Y_n 1_{\{T_n \leq t\}}, \quad (1.29)$$

where the  $T_n$ 's are like in the Poisson case (i.e., the process  $N$  associated by (1.26) is Poisson with some parameter  $\lambda > 0$ ) and the two sequences  $(T_n)$  and  $(Y_n)$  are independent, and the variables  $(Y_n)$  are i.i.d. with values in  $\mathbb{R}^d \setminus \{0\}$ , with a law denoted by  $G$ . The  $T_n$ 's represent the jump times and the  $Y_n$ 's the jump sizes.

On the set  $\{N_t = n\}$ , the variable  $X_t$  is the sum of  $n$  i.i.d. jumps  $Y_n$ 's with distribution  $F$ , so

$$\mathbb{E}\left(e^{iu^* X_t} \mid N_t = n\right) = \mathbb{E}\left[e^{iu^*(Y_1 + \dots + Y_n)}\right] = \mathbb{E}\left[e^{iu^* Y_1}\right]^n = \widehat{G}(u)^n$$

where  $\widehat{G}(u) = \int_{\mathbb{R}^d} e^{iu^* x} G(dx)$  is the characteristic function of  $Y_1$ . Therefore the characteristic function of  $X_t$  is

$$\begin{aligned} \mathbb{E}\left(e^{iu^* X_t}\right) &= \sum_{n=0}^{\infty} \mathbb{E}\left(e^{iu^* X_t} \mid N_t = n\right) \mathbb{P}(N_t = n) \\ &= e^{-t\lambda} \sum_{n=0}^{\infty} \frac{(\lambda \widehat{G}(u) t)^n}{n!} = e^{-\lambda t(1 - \widehat{G}(u))} \\ &= \exp\left(t\lambda \int (e^{iu^* x} - 1) G(dx)\right). \end{aligned}$$

Again, proving that a compound Poisson process is Lévy is a (relatively) easy task, and the above formula shows that its characteristic triple is  $(b, 0, \lambda G)$ , where  $b = \lambda \int_{\{\|x\| \leq 1\}} xG(dx)$ . The converse is also true, although more difficult to prove: any Lévy process whose paths are piecewise constant, that is, have the form (1.29), is a compound Poisson process.

**Example 1.6 (Symmetrical stable process).** A symmetrical stable process is by definition a one-dimensional Lévy process such that  $X$  and  $-X$  have the same law, and which has the following scaling (or self-similarity) property, for some index  $\beta > 0$ :

$$\text{for all } t > 0, \text{ the variables } X_t \text{ and } t^{1/\beta} X_1 \text{ have the same law.} \quad (1.30)$$

By virtue of the properties of Lévy processes, this implies that for any  $a > 0$ , the two processes  $(X_{at})_{t \geq 0}$  and  $(a^{1/\beta} X_t)_{t \geq 0}$  have the same (global) law.

By (1.24), the log-characteristic function of  $X_t$  satisfies  $\Psi(u) = \Psi(-u)$  because of the symmetry, so (1.30) immediately yields  $\Psi(u) = -\phi|u|^\beta$  for some constant  $\phi$ . The fact that  $\psi$  is the logarithm of a characteristic function has two consequences, namely that  $\phi \geq 0$  and that  $\beta \in (0, 2]$ , and  $\phi = 0$  will be excluded because it corresponds to having  $X_t = 0$  identically. Now, we have two possibilities:

1.  $\beta = 2$ , in which case  $X = \sqrt{2\phi}W$ , with  $W$  a Brownian motion.
2.  $\beta \in (0, 2)$ , and most usually the name “stable” is associated with this situation. The number  $\beta$  is then called the index, or stability index, of the stable process. In this case the characteristic triple of  $X$  is  $(0, 0, F)$ , where

$$F(dx) = \frac{a\beta}{|x|^{1+\beta}} dx \quad (1.31)$$

for some  $a > 0$ , which is connected with the constant  $\phi$  by

$$\phi = \begin{cases} a\pi & \text{if } \beta = 1 \\ \frac{2a\beta \sin(\frac{(1-\beta)\pi}{2})}{(1-\beta)\Gamma(2-\beta)} & \text{if } \beta \neq 1 \end{cases}$$

( $\Gamma$  is the Euler gamma function). The requirement in (1.23) that  $F$  integrates the function  $1 \wedge x^2$  is exactly the property  $0 < \beta < 2$ .

The case  $\beta = 1$  corresponds to the Cauchy process, for which the density of the variable  $X_t$  is explicitly known and of the form  $x \mapsto 1/(t\pi^2(1 + (x/t\pi)^2))$ .

Note that the value of  $\beta$  controls the rate at which  $F$  diverges near 0: the higher the value of  $\beta$ , the faster  $F$  diverges, and, as we will see later, the higher the concentration of small jumps of the process. But the same parameter also controls the tails of  $F$  near  $\infty$ . In the case of a stable process, these two behaviors of  $F$  are linked. Of course, this is the price to pay for the scaling property (1.30).

Finally, we note that these processes are stable under addition: if  $X^{(1)}, \dots, X^{(n)}$  are  $n$  independent copies of  $X$ , then there exist numbers  $a_n > 0$  such that  $X_t^{(1)} + \dots + X_t^{(n)} \stackrel{d}{=} a_n X_t$ .

**Example 1.7 (General stable process).** As said before, we exclude the case of the Brownian motion here. The terminology in the non-symmetrical case is not completely well established. For us, a stable process will be a Lévy process  $X$  having the characteristic triple  $(b, 0, F)$ , where  $b \in \mathbb{R}$  and  $F$  is a measure of the same type as (1.31), but not necessarily symmetrical about 0:

$$F(dx) = \left( \frac{a^{(+)}\beta}{|x|^{1+\beta}} 1_{\{x>0\}} + \frac{a^{(-)}\beta}{|x|^{1+\beta}} 1_{\{x<0\}} \right) dx, \quad (1.32)$$

where  $a^{(+)}, a^{(-)} \geq 0$  and  $a^{(+)} + a^{(-)} > 0$ , and  $\beta \in (0, 2)$  is again called the index of the process. This includes the symmetrical stable processes (take  $b = 0$  and  $a^{(+)} = a^{(-)} = a$ ).

The scaling property (1.30) is lost here, unless either  $\beta = 1$  and  $b \in \mathbb{R}$  and  $a^{(+)} = a^{(-)}$ , or  $\beta \neq 1$  and  $b = \frac{\beta(a^{(+)} - a^{(-)})}{1 - \beta}$ . The variables  $X_t$  for  $t > 0$  have a density, unfortunately almost never explicitly known, but one knows exactly the behavior of this density at infinity, and also at 0, as well as the explicit (but complicated) form of the characteristic function; see for example the comprehensive monograph of Zolotarev (1986). Note also that, by a simple application of (1.25), we have for all  $t > 0$

$$p < \beta \quad \Rightarrow \quad \mathbb{E}(|X_t|^p) < \infty, \quad p \geq \beta \quad \Rightarrow \quad \mathbb{E}(|X_t|^p) = \infty.$$

Finally, let us mention that the density of the variable  $X_t$  is positive on  $\mathbb{R}$  when  $\beta \geq 1$ , and also when  $\beta < 1$  and  $a^{(+)}, a^{(-)} > 0$ . When  $\beta < 1$  and  $a^{(-)} = 0 < a^{(+)}$ , (resp.  $a^{(+)} = 0 < a^{(-)}$ ), the density of  $X_t$  is positive on  $(b't, \infty)$ , (resp.  $(-\infty, b't)$ ), and vanishes elsewhere, where  $b' = b - \int_{\{|x| \leq 1\}} x F(dx)$  is the “true drift.” If  $\beta < 1$  and  $b' \geq 0$  and  $a^{(-)} = 0 < a^{(+)}$ , almost all paths of  $X$  are strictly increasing, and we say that we have a subordinator.

**Example 1.8 (Tempered stable process).** A tempered stable process of index  $\beta \in (0, 2)$  is a Lévy process whose characteristic triple is  $(b, 0, F)$ , where  $b \in \mathbb{R}$  and  $F$  is

$$F(dx) = \left( \frac{a^{(+)} \beta e^{-B_+|x|}}{|x|^{1+\beta}} 1_{\{x>0\}} + \frac{a^{(-)} \beta e^{-B_-|x|}}{|x|^{1+\beta}} 1_{\{x<0\}} \right) dx,$$

for some  $a^{(+)}, a^{(-)} \geq 0$  with  $a^{(+)} + a^{(-)} > 0$ , and  $B_-, B_+ > 0$ . The reason for introducing tempered stable processes is that, although they somehow behave like stable processes, as far as “small jumps” are concerned, they also have moments of all orders (a simple application of (1.25) again). Those processes were introduced by Novikov (1994) and extended by Rosiński (2007) to a much more general situation than what is stated here.

**Example 1.9 (Gamma process).** The gamma process is in a sense a “tempered stable process with index 0.” It is an increasing Lévy process  $X$  having the characteristics triple  $(0, 0, F)$ , where

$$F(dx) = \frac{a e^{-Bx}}{x} 1_{\{x>0\}} dx,$$

with  $a, B > 0$ . The name comes from the fact that the law of  $X_t$  is the gamma distribution with density  $x \mapsto \frac{1}{\Gamma(ta)} e^{-Bx} B^{ta} x^{ta-1} 1_{\{x>0\}}$ .

### 1.2.3 Poisson Random Measures

In this subsection we switch to a seemingly different topic, whose (fundamental) connection with Lévy processes will be explained later. The idea is to count the number of jumps of a given size that occur between times 0 and  $t$ . We start with a sketchy description of general Poisson random measures, also called Poisson point processes, or “independently scattered point processes.” We consider a measurable space  $(L, \mathcal{L})$ . A random measure on  $L$ , defined on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , is a transition measure  $\underline{p} = \underline{p}(\omega, dz)$  from  $(\Omega, \mathcal{F})$  into  $(L, \mathcal{L})$ . If for each  $\omega$  the measure  $\underline{p}(\omega, \cdot)$  is an at most countable sum of Dirac masses sitting at pairwise distinct points of  $L$ , depending on  $\omega$ , we say that  $\underline{p}$  is associated with a “simple point process” in  $L$ , and  $\underline{p}(A)$  is simply the number of points falling into  $A \subset L$ .

**Definition 1.10.** A random measure  $\underline{p}$  associated with a simple point process is called a Poisson random measure if it satisfies the following two properties:

1. For any pairwise disjoint measurable subsets  $A_1, \dots, A_n$  of  $L$  the variables  $\underline{p}(A_1), \dots, \underline{p}(A_n)$  are independent.
2. The intensity measure  $\underline{q}(A) = \mathbb{E}(\underline{p}(A))$  is a  $\sigma$ -finite measure on  $(L, \mathcal{L})$ , without atoms (an atom is a measurable set which has positive measure and contains no subset of smaller but positive measure).

In this case, if  $\underline{q}(A) = \infty$  we have  $\underline{p}(A) = \infty$  a.s., and if  $\underline{q}(A) < \infty$  the variable  $\underline{p}(A)$  is Poisson with parameter  $\underline{q}(A)$ , that is,

$$\mathbb{P}(\underline{p}(A) = n) = \exp(-\underline{q}(A)) \frac{(\underline{q}(A))^n}{n!}$$

for  $n \in \mathbb{N}$ . Hence the “law” of  $\underline{p}$  is completely characterized by the intensity measure  $\underline{q}$ . It is also characterized by the “Laplace functional,” which is

$$\Phi(f) := \mathbb{E}\left(e^{-\int f(x)\underline{p}(dx)}\right) = \exp\left(-\int (1 - e^{-f(x)})\underline{q}(dx)\right).$$

for any Borel nonnegative (non-random) function  $f$  on  $L$ , with the convention  $e^{-\infty} = 0$ .

A last useful (and simple to prove) property is the following one: if  $A_1, A_2, \dots$  are pairwise disjoint measurable subsets of  $L$ , we have:

the restrictions of  $\underline{p}$  to the  $A_n$ 's are *independent* Poisson random measures, whose intensity measures  
(1.33)are the restrictions of  $\underline{q}$  to the  $A_n$ 's.

The situation above is quite general, but in this book we specialize as follows: we let  $L = \mathbb{R}_+ \times E$ , where  $(E, \mathcal{E})$  is a “nice” topological space with its Borel  $\sigma$ -field, typically  $E = \mathbb{R}^q$  (it could be a general Polish space as well): for the measure associated with the jumps of a process, see below,  $\mathbb{R}_+$  is the set of times and  $\mathbb{R}^q$  the set of jump sizes.

Moreover, we only consider Poisson random measures  $\underline{p}$  having an intensity measure of the form

$$\underline{q}(dt, dx) = dt \otimes Q(dx), \tag{1.34}$$

where  $Q$  is a  $\sigma$ -finite measure on  $(E, \mathcal{E})$ . In this case it turns out that, outside a null set, the process  $a_t = \underline{p}(\{t\} \times E)$  (which *a priori* takes its values in  $\overline{\mathbb{N}}$ ) actually takes only the values 0 and 1, and the (random) set  $D = \{t : a_t = 1\}$  is countable: this is the set of times where points occur.

Upon deleting the above null set, it is thus not a restriction to assume that  $\underline{p}$  has the representation

$$\underline{p} = \sum_{t \in D} \varepsilon_{(t, Z_t)}, \tag{1.35}$$

where as usual  $\varepsilon_a$  denotes the Dirac measure sitting at  $a$ , and  $Z = (Z_t)_{t \geq 0}$  is a measurable process.

When  $Q(A) < \infty$ , the process  $\underline{p}([0, t] \times A)$  is a Poisson process with parameter  $Q(A)$ . Exactly as in the previous subsection, in which Lévy processes (and in particular Poisson processes) relative to a filtration  $(\mathcal{F}_t)$  were defined, we introduce a similar notion for Poisson random measures whose intensity has the form (1.34). Suppose that our random measure  $\underline{p}$  is defined on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ ; we then say that  $\underline{p}$  is a *Poisson measure relative to  $(\mathcal{F}_t)$* , or an  *$(\mathcal{F}_t)$ -Poisson random measure*, if it is a Poisson measure satisfying also:

1. The variable  $\underline{p}([0, t] \times A)$  is  $\mathcal{F}_t$ -measurable for all  $A \in \mathcal{E}$  and  $t \geq 0$ .
2. The restriction of the measure  $\underline{p}$  to  $(t, \infty) \times E$  is independent of  $\mathcal{F}_t$ .

This implies that for all  $A$  with  $Q(A) < \infty$ , the process  $\underline{p}([0, t] \times A)$  is an  $(\mathcal{F}_t)$ -Poisson process, and we have a (not completely trivial) converse: if  $\underline{p}$  is a random measure of the form (1.35) such that for any  $A$  with  $Q(A) < \infty$  the process  $\underline{p}([0, t] \times A)$  is an  $(\mathcal{F}_t)$ -Poisson process with parameter  $Q(A)$ , then  $\underline{p}$  is an  $(\mathcal{F}_t)$ -Poisson random measure with the intensity measure given by (1.34).

If  $E = \{1\}$  the measure  $\underline{p}$  is entirely characterized by the process  $N_t = \underline{p}([0, t] \times \{1\})$ . In this case  $\underline{p}$  is an  $(\mathcal{F}_t)$ -Poisson random measure if and only if  $N$  is an  $(\mathcal{F}_t)$ -Poisson process; this is the simplest example of a Poisson random measure.

**Example 1.11.** *If  $X$  is an  $\mathbb{R}^d$ -valued càdlàg process, we associate its jump measure, defined as follows (recall the notation (1.20) for the jumps):*

$$\mu^X = \sum_{s > 0: \Delta X_s \neq 0} \varepsilon_{(s, \Delta X_s)}, \tag{1.36}$$

*which is (1.35) when  $E = \mathbb{R}^d$  and  $Z_t = \Delta X_t$  and  $D = \{t : \Delta X_t \neq 0\}$ . Note that  $\mu^X([0, t] \times A)$  is the number of jumps of size falling in the measurable set  $A \subset E$ , between times 0 and  $t$ :*

$$\mu^X([0, t] \times A) = \sum_{0 < s \leq t} 1_{\{\Delta X_s \in A\}}.$$

Then it turns out that the measure  $\mu^X$  is a Poisson random measure when  $X$  is a Lévy process and an  $(\mathcal{F}_t)$ -Poisson random measure when  $X$  is an  $(\mathcal{F}_t)$ -Lévy process, and in those cases the measure  $Q$  in (1.34) is equal to the Lévy measure of the process; we thus have independence of the number of jumps both serially (over two disjoint time intervals  $[t_1, t_2]$  and  $[t_3, t_4]$ ) and cross-sectionally (jump sizes in two disjoint sets  $A_1$  and  $A_2$ ).

The cross-sectional independence is not simple to prove, but the time independence is quite intuitive: the value  $\mu^X((t, t+s] \times A) = \mu^X((0, t+s] \times A) - \mu^X((0, t] \times A)$  is the number of jumps of size in  $A$ , in the time interval  $(t, t+s]$ , so it only depends on the increments  $(X_{t+v} - X_t)_{v \geq 0}$ . Then, the  $(\mathcal{F}_t)$ -Lévy property, say, implies that this variable  $\mu^X((t, t+s] \times A)$  is independent of  $\mathcal{F}_t$ , and also that its law only depends on  $s$  (by stationarity of the increments of  $X$ ). Therefore the process  $\mu^X((0, t] \times A)$  is an  $(\mathcal{F}_t)$ -Lévy process, and also a counting process, hence an  $(\mathcal{F}_t)$ -Poisson process.

### 1.2.4 Integrals with Respect to Poisson Random Measures

For any random measure  $\mu$  on  $\mathbb{R}_+ \times E$ , where  $(E, \mathcal{E})$  is a Polish space, and for any measurable function  $U$  on  $\Omega \times \mathbb{R}_+ \times E$ , we set

$$U * \mu_t(\omega) = \int_0^t \int_E U(\omega, s, x) \mu(\omega; ds, dx), \quad (1.37)$$

whenever this makes sense, for example when  $U \geq 0$  and  $\mu$  is positive.

Suppose that  $\underline{p}$  is an  $(\mathcal{F}_t)$ -Poisson random measure, with intensity measure  $\underline{q}$  given by (1.34). If  $U(\omega, t, x) = 1_A(x)$  with  $Q(A) < \infty$ , then both  $U * \underline{p}_t$  and  $U * \underline{q}_t$  are well defined. Keep in mind that  $\underline{p}$  is random, but that  $\underline{q}$ , its compensator, is not. And by (1.28) the compensated difference  $U * \underline{p} - U * \underline{q}$  is a martingale on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . This property extends to any finite linear combination of such  $U$ 's, and in fact extends much more, as we shall see below.

To this end, we first recall that we can endow the product space  $\Omega \times \mathbb{R}_+$  with the *predictable*  $\sigma$ -field  $\mathcal{P}$ , that is, the  $\sigma$ -field generated by the sets  $B \times \{0\}$  for  $B \in \mathcal{F}_0$  and  $B \times (s, t]$  for  $s < t$  and  $B \in \mathcal{F}_s$ , or equivalently (although this is not trivial) the  $\sigma$ -field generated by all processes that are adapted and left-continuous, or the  $\sigma$ -field generated by all processes that are adapted and continuous. By extension, the product  $\sigma$ -field  $\tilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{E}$  is also called the predictable  $\sigma$ -field on  $\Omega \times \mathbb{R}_+ \times E$ , and a  $\tilde{\mathcal{P}}$ -measurable function on this space is called a *predictable function*.

If  $U$  is a predictable function on  $\Omega \times \mathbb{R}_+ \times E$  such that  $\mathbb{E}(|U| * \underline{q}_t) < \infty$  for all  $t$ , the difference  $U * \underline{p} - U * \underline{q}$  is again a martingale (this is an easy result, because the linear space spanned by all functions of the form  $U(\omega, t, x) = 1_B(\omega)1_{(u,v]}(t)1_A(x)$  is dense in the sets of all predictable functions in  $\mathbb{L}^1(\mathbb{P} \otimes \underline{q})$ ). Slightly more generally, if  $|U| * \underline{q}_t < \infty$  for all  $t$ , then we have  $|U| * \underline{p}_t < \infty$  as well, and the difference  $U * \underline{p} - U * \underline{q}$  is a local martingale. Moreover, the càdlàg process  $U * \underline{p} - U * \underline{q}$  has jumps obviously satisfying for  $t > 0$ :

$$\Delta(U * \underline{p} - U * \underline{q})_t = \int_E U(t, x) \underline{p}(\{t\} \times dx) = U(t, Z_t)$$

(where we use the representation (1.35) for the last equality, and  $\Delta Y$  is the jump process of any càdlàg process  $Y$ ).

At this stage, and somewhat similar to stochastic integrals with respect to a Brownian motion, one can define *stochastic integrals* with respect to the Poisson random measure  $\underline{p}$ , or rather with respect to the compensated measure  $\underline{p} - \underline{q}$ , as follows: if  $U$  is a *predictable* function on  $\Omega \times \mathbb{R}_+ \times E$  such that

$$(|U| \wedge U^2) * \underline{q}_t < \infty \quad \forall t \geq 0, \tag{1.38}$$

there exists a local martingale  $M$  having the following properties:

- $M$  is orthogonal to all continuous martingales, meaning that the product  $MM'$  is a local martingale for any continuous martingale  $M'$ ;
- outside a null set,  $M_0 = 0$  and  $t > 0$   
 $\Rightarrow \Delta M_t = \int_E U(t, x) \underline{p}(\{t\} \times dx) = U(t, Z_t)$ .

This local martingale is unique (up to a null set again), and we use either one of the following notations:

$$M_t = U * (\underline{p} - \underline{q})_t = \int_0^t \int_E U(s, x) (\underline{p} - \underline{q})(ds, dx). \tag{1.39}$$

We have the following four properties, quite similar to (1.9):

- the map  $U \mapsto U * (\underline{p} - \underline{q})$  is linear;
- we have a “dominated convergence theorem”:  
 if  $U^n \rightarrow U$  pointwise and  $|U^n| \leq V$  and  $V$  satisfies (1.38), then  $U^n * (\underline{p} - \underline{q}) \xrightarrow{\text{u.c.P.}} U * (\underline{p} - \underline{q})$ ;
- if  $|U| * \underline{q}_t < \infty$  for all  $t$ , then  $U * (\underline{p} - \underline{q}) = U * \underline{p} - U * \underline{q}$  (otherwise, the processes  $U * \underline{p}$  and  $U * \underline{q}$  may be ill-defined);
- if  $U^2 * \underline{q}_t < \infty$  for all  $t$ , then  $U * (\underline{p} - \underline{q})$  is a locally square-integrable local martingale.

$$\tag{1.40}$$

Sometimes, the local martingale  $M = U * (\underline{p} - \underline{q})$  is called a “purely discontinuous” local martingale, or a “compensated sum of jumps”; the reason is that in the third property (1.40),  $U * \underline{p}_t = \sum_{s \in D} U(s, Z_s) 1_{\{s \leq t\}}$  is a sum of jumps, and  $U * \underline{q}$  is the unique predictable process starting at 0 and which “compensates”  $U * \underline{p}$ , in the sense that the difference becomes a (local) martingale.

The notion of stochastic integral can be extended to random measures of the type (1.35) which are not necessarily Poisson, and below we consider the only case of interest for us, which is the jump measure  $\mu = \mu^X$  of an  $\mathbb{R}^d$ -valued càdlàg adapted process  $X$ ; see (1.36).

For any Borel subset  $A$  of  $\mathbb{R}^d$  at a positive distance of 0, the process  $\mu([0, t] \times A)$  is an adapted counting process, taking only finite values because for any càdlàg process the number of jumps with size bigger than any  $\varepsilon > 0$  and within the time interval  $[0, t]$  is finite. Therefore one can “compensate” this increasing process by a *predictable* increasing càdlàg process  $Y(A)$  starting at 0, in such a way that the difference  $\mu([0, t] \times A) - Y(A)_t$  is a martingale (this is like  $U * \underline{p} - U * \underline{q}$  in the previous paragraph), and  $Y(A)$  is almost surely unique (this is a version of the celebrated Doob-Meyer decomposition of a submartingale). The map  $A \mapsto \mu([0, t] \times A)$  is additive, and thus so is the map  $A \mapsto Y(A)$ , up to null sets. Hence it is not a surprise (although it needs a somewhat involved proof, because of the  $\mathbb{P}$ -negligible sets) that there exists an almost surely unique random measure  $\nu$  on  $\mathbb{R}_+ \times \mathbb{R}^d$  such that, for all  $A \in \mathcal{R}^d$  at a positive distance of 0,

$$\nu([0, t] \times A) \text{ is predictable, and is a version of } Y(A). \quad (1.41)$$

The measure  $\nu = \nu^X$  is called the *compensating measure* of  $\mu$ . Of course, when  $\mu$  is further a Poisson random measure, its compensating measure  $\nu$  is also its intensity, and is thus not random.

We can rephrase the previous statement as follows, with the notation (1.37): If  $U(\omega, t, x) = 1_A(x)$ , with  $A$  Borel and at a positive distance of 0, then  $U * \nu$  is predictable and the difference  $U * \mu - U * \nu$  is a local martingale. Exactly as in the previous paragraph, this extends to any  $U$  which is *predictable* and such that

$$(|U| \wedge U^2) * \nu_t < \infty \quad \forall t \geq 0. \quad (1.42)$$

Namely, there exists an almost surely unique local martingale  $M$  satis-

fying

- $M$  is orthogonal to all continuous martingales
- outside a null set, we have  $M_0 = 0$  and, for all  $t > 0$ ,
 
$$\Delta M_t = \int_E U(t, x)\mu(\{t\} \times dx) - \int_E U(t, x)\nu(\{t\} \times dx). \tag{1.43}$$

As in (1.39), we use either one of the following notations:

$$M_t = U * (\mu - \nu)_t = \int_0^t \int_E U(s, x)(\mu - \nu)(ds, dx),$$

and all four properties in (1.40) are valid here, with  $\underline{p}$  and  $\underline{q}$  substituted with  $\mu$  and  $\nu$ .

There is a difference, though, with the Poisson case: the process  $\gamma_t = \nu(\{t\} \times \mathbb{R}^d)$  takes its values in  $[0, 1]$ , but it is not necessarily vanishing everywhere. When it is (for example when  $\mu$  is a Poisson measure), the second property in (1.43) can be rewritten as

$$\Delta M_t = \int_E U(t, x)\mu(\{t\} \times dx) = U(t, \Delta X_t) 1_{\{\Delta X_t \neq 0\}} \tag{1.44}$$

and the condition (1.42) describes the biggest possible class of predictable integrands  $U$ . When  $\gamma_t$  is not identically 0, (1.44) is wrong, and it is possible to define  $U * (\mu - \nu)$  for a slightly larger class of integrands (see for example Jacod (1979) for more details).

### 1.2.5 Path Properties and Lévy-Itô Decomposition

Now we come back to our  $d$ -dimensional Lévy processes  $X$ , defined on the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ .

A fundamental property, already mentioned in Example 1.11, is that the jump measure  $\mu = \mu^X$  of  $X$  is a *Poisson random measure* on  $L = \mathbb{R}_+ \times \mathbb{R}^d$ , with the intensity measure

$$\nu(dt, dx) = dt \otimes F(dx),$$

where  $F$  is the Lévy measure of  $X$ . Below, we draw some consequences of this fact.

First, since  $\mu(A)$  is Poisson with parameter  $\nu(A)$  if  $\nu(A) < \infty$  and

$\mu(A) = \infty$  a.s. otherwise, we see that

$$\bullet F = 0 \Rightarrow X \text{ is continuous (we knew this already)} \quad (1.45)$$

$$\bullet 0 < F(\mathbb{R}^d) < \infty \Rightarrow X \text{ has a.s. finitely many jumps on any interval } [0, t] \text{ and a.s. infinitely many on } \mathbb{R}_+ \quad (1.46)$$

$$\bullet F(\mathbb{R}^d) = \infty \Rightarrow X \text{ has a.s. infinitely many jumps on any interval } [t, t + s] \text{ such that } s > 0. \quad (1.47)$$

**Definition 1.12.** *In the case of (1.46) we say that we have finite activity for the jumps, whereas if (1.47) holds we say that we have infinite activity.*

Next, let  $g$  be a nonnegative Borel function on  $\mathbb{R}^d$  with  $g(0) = 0$ . By using the Laplace functional for the functions  $f(r, x) = \lambda(g(x) \wedge 1)1_{(t, t+s]}(r)$  and the fact that a nonnegative variable  $Y$  is a.s. finite if  $\mathbb{E}(e^{-\lambda Y}) \rightarrow 1$  as  $\lambda \downarrow 0$  and a.s. infinite if and only if  $\mathbb{E}(e^{-\lambda Y}) = 0$  for all  $\lambda > 0$ , we deduce

$$\begin{aligned} \int (g(x) \wedge 1)F(dx) < \infty &\Leftrightarrow \sum_{s \leq t} g(\Delta X_s) < \infty, \\ &\text{a.s. } \forall t > 0 \\ \int (g(x) \wedge 1)F(dx) = \infty &\Leftrightarrow \sum_{t < r \leq t+s} g(\Delta X_r) = \infty, \\ &\text{a.s. } \forall t \geq 0, s > 0 \end{aligned} \quad (1.48)$$

which is particularly useful for the absolute power functions  $g(x) = \|x\|^p$  where  $p > 0$ .

We now set

$$I = \left\{ p \geq 0 : \int_{\{\|x\| \leq 1\}} \|x\|^p F(dx) < \infty \right\}, \quad \beta = \inf(I). \quad (1.49)$$

The number  $\beta$  defined above is called the *Blumenthal-Gettoor index* of the process  $X$ , as introduced by Blumenthal and Gettoor (1961), precisely for studying the path properties of Lévy processes. Note that, since the function  $p \mapsto \|x\|^p$  is decreasing when  $\|x\| \leq 1$ , the set  $I$  is necessarily of the form  $[\beta, \infty)$  or  $(\beta, \infty)$ , whereas  $2 \in I$  always by (1.23), hence  $\beta \in [0, 2]$ .

There is no conflicting notation here: for a stable or tempered stable process, the stability index and the Blumenthal-Gettoor index agree. Those are examples where  $I = (\beta, \infty)$ . A gamma process has Blumenthal-Gettoor index  $\beta = 0$  and again  $I = (\beta, \infty)$ . For a compound Poisson process we have  $\beta = 0$  and  $I = [\beta, \infty)$ . More generally, the jumps have finite activity if and only if  $0 \in I$ . Later on, we will see that  $\beta$  can be quite naturally generalized, and interpreted as a *jump activity index*: processes

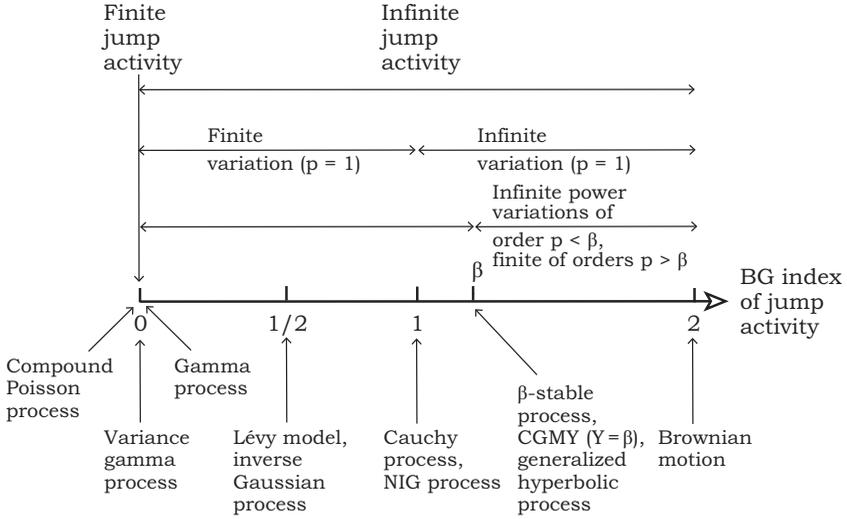


Figure 1.1: Examples of processes and their corresponding BG index of jump activity  $\beta$ .

with higher  $\beta$  tend to jump more frequently. Figure 1.1 provides some examples of processes and the corresponding values of  $\beta$ .

$\beta$  is a lower bound for the values of  $p$  for which the  $p$ th power of the jumps are summable. Observing that  $\sum_{s \leq t} \|\Delta X_s\|^p 1_{\{\|\Delta X_s\| > 1\}} < \infty$  for all  $t$  because  $X$  has finitely many jumps bigger than 1 on any interval  $[0, t]$  (a property of all càdlàg processes), we deduce

$$\begin{aligned} p \in I &\Rightarrow \sum_{s \leq t} \|\Delta X_s\|^p < \infty \quad \text{a.s. } \forall t \geq 0 \\ p \notin I &\Rightarrow \sum_{s \leq t} \|\Delta X_s\|^p = \infty \quad \text{a.s. } \forall t > 0. \end{aligned} \tag{1.50}$$

With  $\mu = \mu^X$  and  $\nu$  and above, we observe that the predictable function  $U(\omega, t, x) = x 1_{\{\|x\| \leq 1\}}$  satisfies (1.38), or equivalently (1.38) with  $\underline{q} = \nu$ . Hence the stochastic integral  $U * (\underline{p} - \underline{q})$  is well defined, and will be written below as  $(x 1_{\{\|x\| \leq 1\}}) * (\mu - \nu)$ . With the notation (1.37) we also clearly have

$$(x 1_{\{\|x\| > 1\}}) * \mu_t = \sum_{s \leq t} \Delta X_s 1_{\{\|\Delta X_s\| > 1\}}$$

(a finite sum for each  $t$ ). On the other hand the symmetric nonnegative matrix  $c$  occurring in the Lévy-Khintchine formula can be written as  $c = \sigma \sigma^*$  for a  $d \times q$  matrix  $\sigma$ , where  $q$  is the rank of  $c$ .

With all this notation, and with  $b$  as in (1.24), one can show that on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  there is a  $q$ -dimensional Brownian motion  $W$ , *independent* of the Poisson measure  $\mu$ , and such that the Lévy process  $X$  is

$$X_t = bt + \sigma W_t + (x1_{\{\|x\| \leq 1\}}) * (\mu - \nu)_t + (x1_{\{\|x\| > 1\}}) * \mu_t. \quad (1.51)$$

This is called the *Lévy-Itô decomposition* of  $X$ . This decomposition is quite useful for applications, and also provides a lot of insight on the structure of a Lévy process.

A few comments are in order here:

1. When  $c = 0$  there is no  $\sigma$ , and of course the Brownian motion  $W$  does not show in this formula.
2. The independence of  $W$  and  $\mu$  has been added for clarity, but one may show that if  $W$  and  $\mu$  are an  $(\mathcal{F}_t)$ -Brownian motion and an  $(\mathcal{F}_t)$ -Poisson measure on some filtered space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , then they necessarily are independent.
3. The four terms on the right in the formula (1.51) correspond to a canonical decomposition of  $X_t$  into a sum of a pure drift term, a continuous martingale, a purely discontinuous martingale consisting of “small” jumps (small meaning smaller than 1) that are compensated, and the sum of the “big” jumps (big meaning bigger than 1). As we will see, this is also the structure of a general semimartingale.
4. The four terms in (1.51) are *independent* of each other; for the last two terms, this comes from (1.33).
5. These four terms correspond to the decomposition of the characteristic function (1.24) into four factors, that is,

$$\mathbb{E} \left( e^{iu^* X_t} \right) = \prod_{j=1}^4 \phi_j(u),$$

where

$$\begin{aligned} \phi_1(u) &= e^{iu^* bt}, & \phi_2(u) &= e^{-\frac{1}{2} tu^* cu}, \\ \phi_3(u) &= e^{t \int_{\{\|x\| \leq 1\}} (e^{iu^* x} - 1 - iu^* x) F(dx)}, \\ \phi_4(u) &= e^{t \int_{\{\|x\| > 1\}} (e^{iu^* x} - 1) F(dx)}. \end{aligned}$$

The terms  $\phi_1$  and  $\phi_2$  are the characteristic functions of  $bt$  and  $\sigma W_t$ , and the last one is the characteristic function of the compound Poisson variable which is the last term in (1.51). For the

third factor, one can observe that on the one hand the variable  $(x1_{(1/n) < \|x\| \leq 1}) * (\mu - \nu)_t$  is a compound Poisson variable minus  $t \int_{\{(1/n) < \|x\| \leq 1\}} xF(dx)$ , whose characteristic function is

$$\exp t \int_{\{(1/n) < \|x\| \leq 1\}} \left( e^{iu^*x} - 1 - iu^*x \right) F(dx),$$

whereas on the other hand it converges to the third term in (1.51) as  $n \rightarrow \infty$  by the dominated convergence theorem in (1.40).

6. Instead of truncating jumps at 1 as in (1.51), we can decide to truncate them at an arbitrary fixed  $\varepsilon > 0$ , in which case the decomposition formula becomes

$$X_t = b_\varepsilon t + \sigma W_t + (x1_{\{\|x\| \leq \varepsilon\}}) * (\mu - \nu)_t + (x1_{\{\|x\| > \varepsilon\}}) * \mu_t$$

with the drift term changed to

$$b_\varepsilon = b + \int x (1_{\{\|x\| \leq \varepsilon\}} - 1_{\{\|x\| \leq 1\}}) F(dx).$$

We can more generally employ a truncation function  $h(x)$  in lieu of  $1_{\{\|x\| \leq \varepsilon\}}$ , as long as  $h(x) = 1 + o(\|x\|)$  near 0 and  $h(x) = O(1/|x|)$  near  $\infty$ , so that

$$\begin{aligned} e^{iu^*x} - 1 - iu \cdot xh(x) &= O(\|x\|^2) & \text{as } x \rightarrow 0 \\ e^{iu^*x} - 1 - iu \cdot xh(x) &= O(1) & \text{as } \|x\| \rightarrow \infty \end{aligned}$$

and  $\int (\|x\|^2 \wedge 1)F(dx) < \infty$  ensures that

$$\int |e^{iu^*x} - 1 - iu^*xh(x)|F(dx) < \infty.$$

The drift needs again to be adjusted to

$$b_h = b + \int x (h(x) - 1_{\{\|x\| \leq 1\}}) F(dx)$$

Different choices of  $\varepsilon$  or  $h$  do not change  $(c, F)$  but they are reflected in the drift  $b_h$ , which is therefore a somewhat arbitrary quantity. Since the choice of truncation is essentially arbitrary, so is the distinction between small vs. big jumps. The only distinguishing characteristic of big jumps is that there are only a finite number of them, at the most.

7. Recall that the reason we cannot simply take  $x * \mu_t$  is that, when the process has infinite jump activity, the series  $x * \mu_t = \sum_{s \leq t} \Delta X_s$  may be divergent even though the number of jumps is at most countable. The variable  $(x 1_{\{\|x\| > \varepsilon\}}) * \mu_t = \sum_{s \leq t} \Delta X_s 1_{\{\|\Delta X_s\| > \varepsilon\}}$ , being a finite sum for all  $\varepsilon > 0$ , is well defined, but its limit as  $\varepsilon \rightarrow 0$  may not exist because of the infinitely many jumps. Compensating solves that problem and  $(x 1_{\{\varepsilon < \|x\| \leq 1\}}) * (\mu - \nu)_t$  converges as  $\varepsilon \rightarrow 0$ . On the other hand, we cannot simply compensate without truncating and take  $x * (\mu - \nu)_t$ , because  $(x 1_{\{\|x\| > 1\}}) * \nu_t$  may be divergent; hence the solution which consists in breaking the sum (or integrals) into two parts, compensated small jumps, and un-compensated large jumps, both of which are convergent.
8. If the process has finite variation, that is, if  $\|x\| * \mu_t < \infty$  almost surely for all  $t$ , or equivalently if  $\int (\|x\| \wedge 1) F(dx) < \infty$  (note the absence of a square on  $\|x\|$ ), then the integral  $x * \mu_t$  is convergent, and the compensation is not needed. The difference between the representation

$$X_t = b't + \sigma W_t + x * \mu_t$$

and (1.51) for a finite variation process consists in a change of drift from  $b$  to  $b'$ .

## 1.3 Semimartingales

### 1.3.1 Definition and Stochastic Integrals

We now reach the main topic of this chapter. Among all processes, the class of semimartingales plays a very special role. For example they are the most general processes with respect to which a (stochastic) integration theory, having the usual “nice” properties like a Lebesgue convergence theorem, can be constructed. This fact may even be used as the definition of semimartingales, according to the Bichteler-Dellacherie-Mokobodski theorem. In mathematical finance they also play a special role, since one of the most basic results (the so-called fundamental asset pricing theorem) says that if no arbitrage is allowed, then the price process should be a semimartingale.

**Definition 1.13.** *A real-valued process  $X$  on the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is called a semimartingale if it can be written as*

$$X = A + M, \tag{1.52}$$

where  $M$  is a local martingale and  $A$  is an adapted càdlàg process “with finite variation,” which means that the total variation of each path  $t \mapsto A_t(\omega)$  is bounded over each finite interval  $[0, t]$ .

Intuitively, the finite variation process plays the role of a “drift” term (such as  $\int_0^t b_s ds$ ) whereas the local martingale part is “pure randomness” (such as  $\int_0^t \sigma_s dW_s$ ). A deterministic process is a semimartingale if and only if it has finite variation. This decomposition  $X = M + A$  is essentially *not* unique, since we can always add to  $A$  and subtract from  $M$  the same martingale with finite variation. Both  $A$  and  $M$  may be discontinuous.

A  $d$ -dimensional semimartingale  $X = (X^i)_{1 \leq i \leq d}$  is a process whose components are real-valued semimartingales. A semimartingale is always adapted and càdlàg. Every adapted process of finite variation (e.g. Poisson or pure drift) is a semimartingale. Every martingale (e.g. Brownian motion) is a semimartingale. By virtue of the Lévy-Itô decomposition (1.51), any Lévy process is a semimartingale, since it is the sum of a square integrable martingale and a finite variation process: one may for example take  $A_t = bt + (x1_{\{\|x\|>1\}}) * \mu_t$  to get a decomposition (1.52).

Now we introduce stochastic integrals with respect to a semimartingale  $X$ , starting with the one-dimensional case. Exactly as for the Brownian motion, and substituting  $W$  with  $X$ , one may define the integral  $\int_0^t H_s dX_s$  by (1.7) for any simple process  $H$  of the form (1.6).

Again as for the Brownian motion, this elementary integral can be extended to a much larger class of integrands. An important difference with the Brownian case is the fact that we need the integrand to be *predictable*, a fundamental requirement without which the whole theory breaks down. More precisely, we consider processes  $H$  with the following properties:

$$\begin{aligned} &H \text{ is predictable, and locally bounded in the sense that} \\ &\text{we have } |H_t(\omega)| \leq n \text{ for all } 0 < t \leq T_n(\omega), \\ &\text{where } (T_n) \text{ is a sequence of stopping times increasing to } \infty. \end{aligned} \tag{1.53}$$

This is *not* the largest possible class of integrands, but it will be sufficient for our purposes. The extension is of course still denoted by  $\int_0^t H_s dX_s$  or  $H \bullet X_t$ , and is uniquely defined up to a null set again. It has properties

analogous to (1.9):

- the process  $H \bullet X_t$  is a semimartingale starting at 0
  - if  $X$  is a local martingale, then so is  $H \bullet X$
  - the maps  $H \mapsto H \bullet X$  and  $X \mapsto H \bullet X$  are linear
  - we have a “dominated convergence theorem”:
- if  $H^n \rightarrow H$  pointwise and  $|H^n| \leq H'$   
 where  $H'$  satisfies (1.53), then  $H^n \bullet X \xrightarrow{\text{u.c.p.}} H \bullet X$ .
- (1.54)

When  $X$  is  $d$ -dimensional, one can integrate componentwise a  $d$ -dimensional predictable locally bounded process  $H = (H^i)_{1 \leq i \leq d}$ , thus getting the following one-dimensional process:

$$H \bullet X_t = \int_0^t H_s dX_s = \sum_{i=1}^d \int_0^t H_s^i dX_s^i.$$

Why the predictability requirement? There are two reasons. A first “mathematical” reason is that, for a general semimartingale, it is impossible to extend integrals of simple integrands to all càdlàg adapted processes while preserving the dominated convergence theorem. Second, in financial applications, trading strategies must be predictable to avoid arbitrage opportunities. Imagine an investor trading an asset with price process  $X$  at times  $T_i$ ,  $0 = T_0 < T_1 < \dots < T_{n+1} = T$ , and holding  $H_{T_i}$  shares of the asset between  $T_i$  and  $T_{i+1}$ . The capital gain from that strategy is

$$\sum_{i=0}^n H_{T_i} (X_{T_{i+1}} - X_{T_i}) = \int_0^T H_t dX_t.$$

The transaction times  $T_i$  may be fixed, or more generally they may be non-anticipating random times (that is, stopping times). For example the investor may trade the first time  $X$  crosses a barrier: this would be a limit order. Also,  $H_{T_i}$  is chosen based on information known at  $T_i$ : it is  $\mathcal{F}_{T_i}$ -measurable. The investor’s holdings at each time  $t$  are given by

$$H_t = H_0 1_0(t) + \sum_{i=0}^n H_{T_i} 1_{(T_i, T_{i+1}]}(t).$$

The investor decides to trade at  $T_i$ ; immediately afterwards, the portfolio changes from  $H_{T_{i-1}}$  to  $H_{T_i}$ . Therefore  $H_t$  is left-continuous with right-limits (càglàd). This makes it a *predictable* process.

Trading strategies must be predictable when jumps are present, otherwise there may be arbitrage opportunities. Consider for example  $X_t = \lambda t - N_t$ , where  $N$  is a Poisson process with intensity  $\lambda$ . Let  $T_1$  denote

the time of the first jump. At that time,  $X$  jumps by  $-1$ . Consider the strategy that buys 1 unit of the asset at time 0 for price  $X_0 = 0$  and sells it right before the crash at  $T_1$ .  $H$  is not a predictable process, since it is impossible to implement this strategy without knowing ahead the time when  $T_1$  will happen. The holdings  $H_t$  are given by 1 between  $[0, T_1)$ . The fact that the interval is open on the right means that this strategy is not predictable because of the properties of the jump times of a Poisson process. The strategy generates sure profits since

$$\int_0^t H_s dX_s = \begin{cases} \lambda t & \text{for } t < T_1 \\ \lambda T_1 & \text{for } t \geq T_1 \end{cases}$$

and does that with zero initial investment, so it is an arbitrage. This explains why, in the presence of jumps, only predictable strategies are admissible.

### 1.3.2 Quadratic Variation

For defining the quadratic variation, one needs to recall some properties. The first is that a local martingale  $M$  can always be written as  $M_t = M_0 + M_t^c + M_t^d$ , where  $M_0^c = M_0^d = 0$  and  $M^c$  is a continuous local martingale, and  $M^d$  is a local martingale orthogonal to each continuous (local) martingale. The second is that a local martingale starting from 0, which has bounded variation in the sense explained after (1.52), and which is continuous, is almost surely vanishing everywhere. Therefore, if we consider two decompositions  $X = M + A = M' + A'$  as (1.52), then necessarily  $M^c = M'^c$  a.s. In other words, we can write the semimartingale  $X$  as

$$X_t = X_0 + X_t^c + M_t + A_t,$$

where  $A_0 = M_0 = 0$  and where  $A$  is of finite variation and  $M$  is a local martingale orthogonal to all continuous martingales, and  $X^c$  is a continuous local martingale starting at 0. In this decomposition the two processes  $M$  and  $A$  are still not unique, but the process  $X^c$  is unique (up to null sets), and it is called the *continuous martingale part* of  $X$  (although it usually is a local martingale only). When  $X$  is  $d$ -dimensional, so are  $X^c$ ,  $M$  and  $A$ , and the components of  $X^c$  are denoted  $X^{i,c}$ .

Now, we saw in Subsection 1.1.2 that, when two continuous local martingales  $M$  and  $M'$  are stochastic integrals with respect to a (possibly multidimensional) Brownian motion, one can define the quadratic covariation (or variation, if  $M' = M$ ). The same is true of all continuous local martingales: based on the Doob-Meyer decomposition of submartingales

again, we have a unique (up to null sets) continuous adapted process  $\langle M, M' \rangle$  starting at 0, with finite variation, and such that

$$M M' - \langle M, M' \rangle \text{ is a local martingale,}$$

and further  $\langle M, M \rangle$  is increasing.

At this point, we can introduce the *quadratic variation* of a one-dimensional semimartingale  $X$  as being

$$[X, X]_t = \langle X^c, X^c \rangle_t + \sum_{s \leq t} (\Delta X_s)^2. \quad (1.55)$$

The sum above makes sense, since it is a sum of positive numbers on the countable set  $\{s : \Delta X_s \neq 0\} \cap [0, t]$ . What is not immediately obvious is that it is a.s. finite, but this fact is one of the main properties of semimartingales. Hence the process  $[X, X]$  is increasing and càdlàg, and also adapted (another intuitive but not mathematically obvious property). Another name for  $[X, X]$  is the “square bracket.” Note that  $[X, X] = \langle X, X \rangle$  when  $X$  is a continuous local martingale, and in general  $[X^c, X^c] = \langle X^c, X^c \rangle$  is the “continuous part” of the increasing process  $[X, X]$  (not to be confused with its “continuous martingale part,” which is identically 0).

For example, if  $X_t = \sigma W_t$ , where  $W$  is Brownian motion, then  $[X, X]_t = \sigma^2 t$ . So  $[X, X]_t$  is not random, and coincides with the variance of  $X_t$ . This is not the case in general:  $[X, X]_t$ , unlike the variance, is a random variable. It is not defined by taking expectations. For example, for a Poisson process, since  $N$  jumps by 1 whenever it does,  $[N, N]_t = N_t$  is the number of jumps of the process between 0 and  $t$ , and we also have  $[X, X]_t = N_t$  for the martingale  $X_t = N_t - \lambda t$  if  $\lambda$  is the parameter of the Poisson process  $N$ . Moreover,  $[X, X]_t$  is well defined for all semimartingales, including those with infinite variance.

If now  $X$  and  $X'$  are two real-valued semimartingales we set

$$[X, X']_t = \langle X^c, X'^c \rangle_t + \sum_{s \leq t} \Delta X_s \Delta X'_s. \quad (1.56)$$

Here again the sum above is a.s. absolutely convergent, by the finiteness in (1.55) for  $X$  and  $X'$  and the Cauchy-Schwarz inequality. The process  $[X, X']$  is adapted and of finite variation, but not necessarily increasing any more, and is called the *quadratic covariation process* of  $X$  and  $X'$ . For example, if  $X_t = \int_0^t \sigma_s dW_s$  and  $X'_t = \int_0^t \sigma'_s dW'_s$ , where  $W$  and  $W'$  are two Brownian motions with correlation coefficient  $\rho$ , then  $[X, X']_t = \int_0^t \rho \sigma_s \sigma'_s ds$ .

For any real  $a$  and any other semimartingale  $X''$  we have

$$[X + aX', X''] = [X, X''] + a[X', X''], \quad [X, X'] = [X', X].$$

Another useful property, which immediately follows from this, is the *polarization identity*:

$$[X, X'] = \frac{1}{4} ([X + X', X + X'] - [X - X', X - X']) \quad (1.57)$$

which expresses the quadratic covariation in terms of quadratic variations only. Finally, the following is obvious:

$$[X, X'] = [X - X_0, X' - X'_0]. \quad (1.58)$$

When  $X$  is  $d$ -dimensional, we thus have a  $d \times d$  matrix-valued process  $[X, X] = ([X^i, X^j])_{1 \leq i, j \leq d}$ . For all  $s, t \geq 0$  the matrix  $[X, X]_{t+s} - [X, X]_t$  is symmetric nonnegative.

We end this subsection with a set of inequalities, known under the name of *Burkholder-Gundy inequalities* when  $p > 1$  and *Davis-Burkholder-Gundy* when  $p = 1$  (when  $p = 2$  it is also a version of Doob's inequality). These inequalities assert that, if  $X$  is a *local martingale* and  $p \geq 1$  and  $S \leq T$  are two arbitrary stopping times, then

$$\begin{aligned} & \mathbb{E} \left( \sup_{s \in [S, T]} |X_{S+s} - X_s|^p \mid \mathcal{F}_S \right) \\ & \leq K_p \mathbb{E} \left( ([X, X]_T - [X, X]_S)^{p/2} \mid \mathcal{F}_S \right), \end{aligned} \quad (1.59)$$

where  $K_p$  is a universal constant depending on  $p$  only, and  $[X, X]_T$  stands on the set  $\{T = \infty\}$  for the increasing (possibly infinite) limit of  $[X, X]_t$  as  $t$  increases to infinity, and  $[X, X]_T - [X, X]_S = 0$  on the set where  $S = T = \infty$ . As a matter of fact, we also have the inequality (1.59) in the other direction, and with another constant  $K_p$ , but this will not be useful for us.

### 1.3.3 Itô's Formula

We are now ready to state the general form of *Itô's formula*, which extends (1.15). From their very definition, semimartingales form a vector space; linear combinations of a finite number of semimartingales are semimartingales. But the class of semimartingales is closed under much more general transformations, and much of its usefulness comes from this fact. If  $f$  is a  $\mathcal{C}^2$  function on  $\mathbb{R}^d$  and  $X$  is a  $d$ -dimensional semimartingale, the

process  $Y = f(X)$  is also a semimartingale and is given by

$$\begin{aligned} f(X_t) &= f(X_0) + \sum_{i=1}^d \int_0^t f'_i(X_{s-}) dX_s^i \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t f''_{ij}(X_{s-}) d\langle X^{i,c}, X^{j,c} \rangle_s \\ &\quad + \sum_{s \leq t} \left( f(X_s) - f(X_{s-}) - \sum_{i=1}^d f'_i(X_{s-}) \Delta X_s^i \right). \end{aligned} \tag{1.60}$$

The reader will notice that all processes  $f'_i(X_{t-})$  and  $f''_{ij}(X_{s-})$  are left-continuous with right limits, so they are locally bounded and predictable and the first (stochastic) and second (ordinary) integrals make sense. Moreover, the  $s$ th summand in the last sum is smaller than  $K_n \|\Delta X_s\|^2$  on the set  $\{\sup_{s \leq t} \|X_s\| \leq n\}$ , for a constant  $K_n$ . Since  $\sum_{s \leq t} \|\Delta X_s\|^2 < \infty$  (because the quadratic variation is finite), this last sum is in fact absolutely convergent. In other words, all terms on the right of (1.60) are well defined.

We end this subsection with the promised proof of (1.14), in the general setting of semimartingales. The result is stated in the form of a theorem, which is unusual in this chapter but motivated by its importance in the econometrics literature and more generally for high-frequency statistics.

**Theorem 1.14.** *Let  $X$  and  $X'$  be two semimartingales. For each  $n$ , let  $T(n, 0) = 0 < T(n, 1) < T(n, 2) < \dots$  be a strictly increasing sequence of stopping time with infinite limit, and suppose that the mesh  $\pi_n(t) = \sup_{i \geq 1} (T(n, i) \wedge t - T(n, i-1) \wedge t)$  goes to 0 for all  $t$ , as  $n \rightarrow \infty$ . Then we have the following convergence in probability:*

$$\sum_{i \geq 1} (X_{T(n,i) \wedge t} - X_{T(n,i-1) \wedge t}) (X'_{T(n,i) \wedge t} - X'_{T(n,i-1) \wedge t}) \xrightarrow{u.c.p.} [X, X']_t. \tag{1.61}$$

Moreover, for any given  $t$  we also have

$$\sum_{i \geq 1, T(n,i) \leq t} (X_{T(n,i)} - X_{T(n,i-1)}) (X'_{T(n,i)} - X'_{T(n,i-1)}) \xrightarrow{\mathbb{P}} [X, X']_t. \tag{1.62}$$

in restriction to the set  $\{\Delta X_t = 0\} \cup \{\Delta X'_t = 0\}$  on which either  $X$  or  $X'$  have no jump at time  $t$ , and on the whole set  $\Omega$  when further, for any  $n$ , there is an index  $i$  such that  $T(n, i) = t$ .

The convergence (1.62) may actually fail on the set  $\{\Delta X_t \neq 0, \Delta X'_t \neq 0\}$ . For example, if for all  $n$  there is no  $i$  such that  $T(n, i) = t$ , the left side of (1.62) converges in probability (on  $\Omega$ ) to the left limit  $[X, X']_{t-}$ , which equals  $[X, X']_t - \Delta X_t \Delta X'_t$ .

The theorem (and its proof below) assumes that  $\pi_n(t) \rightarrow 0$  pointwise, but the condition  $\pi_n(t) \xrightarrow{\mathbb{P}} 0$  is indeed enough for the results to hold.

*Proof.* In view of (1.58) we can replace  $X$  and  $X'$  by  $X - X_0$  and  $X' - X'_0$ , or equivalently assume that  $X_0 = X'_0 = 0$ . The proof is based on the elementary equality  $(x - y)(x' - y') = xx' + yy' - y(x' - y') - y'(x - y)$  applied with  $x = X_{T(n,i) \wedge t}$  and  $y = X_{T(n,i-1) \wedge t}$  and  $x' = X'_{T(n,i) \wedge t}$  and  $y' = X'_{T(n,i-1) \wedge t}$ . Summing over all  $i \geq 1$ , we deduce that the left side of (1.61) is equal to

$$X_t X'_t - \int_0^t H_s^n dX'_s + \int_0^t H_s'^n dX_s$$

(recall  $X_0 = X'_0 = 0$ ), where we have set

$$H_s^n = \sum_{i \geq 1} X_{T(n,i-1)} \mathbf{1}_{(T(n,i-1), T(n,i)]}(s)$$

and a similar formula for  $H_s'^n$ , with  $X'$  instead of  $X$ . The processes  $H^n$  are adapted and left-continuous, hence predictable, and  $|H^n| \leq Z$  where  $Z$  is the predictable locally bounded process defined by  $Z_s = \sup(|X_r| : r \in [0, s])$ . Moreover, since the mesh of the subdivision goes to 0, we have  $H^n \rightarrow X_-$  pointwise. Then the dominated convergence theorem for stochastic integrals, see (1.54), yields  $H^n \bullet X' \xrightarrow{\text{u.c.p.}} X_- \bullet X'$ , and  $H^n \bullet X \xrightarrow{\text{u.c.p.}} X'_- \bullet X$  holds by the same argument. In other words, the left side of (1.61) converges in the u.c.p. sense to

$$X_t X'_t - \int_0^t X_{s-} dX_s + \int_0^t X'_{s-} dX_s.$$

It remains to apply Itô's formula to the two-dimensional semimartingale with components  $X$  and  $X'$  and the function  $f(x, x') = xx'$ : we deduce from (1.56) that the above expression is equal to  $[X, X']_t$ , and thus (1.61) is proved. When for each  $n$  there is  $i$  such that  $T(n, i) = t$ , the left sides of (1.61) and (1.62) are the same, so (1.62) is proved. Finally, in general, the difference between the left sides of (1.61) and (1.62) is smaller than  $\rho(\varepsilon) = \sup(|X_t - X_{t-s}| |X'_t - X'_{t-s}| : s \in [0, \varepsilon])$  as soon as  $\pi_n(t) \leq \varepsilon$ . On the set  $\{\Delta X_t = 0\} \cup \{\Delta X'_t = 0\}$  we have  $\rho(\varepsilon) \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . Then the convergence (1.62) in restriction to this set readily follows from the fact that  $\pi_n(t) \rightarrow 0$ .  $\square$

### 1.3.4 Characteristics of a Semimartingale and the Lévy-Itô Decomposition

Here,  $X$  is a  $d$ -dimensional semimartingale on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , and we are now almost ready to define the characteristics of  $X$ .

The process  $\sum_{s \leq t} \Delta X_s 1_{\{\|\Delta X_s\| > 1\}}$ , or equivalently  $(x 1_{\{\|x\| > 1\}}) * \mu$  where  $\mu = \mu^X$  is the jump measure of  $X$  defined by (1.36), is of finite variation. Then we can rewrite (1.52) as

$$X_t = X_0 + A'_t + M_t + \sum_{s \leq t} \Delta X_s 1_{\{\|\Delta X_s\| > 1\}}$$

where  $M_0 = A'_0 = 0$  and  $A'$  is of finite variation and  $M$  is a local martingale. Now, one can show that the semimartingale  $A' + M$ , which has jumps smaller than 1 by construction, can be written in a *unique* (up to null sets) way as  $A' + M = B + N$ , where again  $N_0 = B_0 = 0$  and  $N$  is a local martingale and  $B$  is a *predictable* process of finite variation.

**Definition 1.15.** *The characteristics of the semimartingale  $X$  are the following triple  $(B, C, \nu)$ :*

- (i)  $B = (B^i)_{1 \leq i \leq d}$  is the predictable process of finite variation defined above;
- (ii)  $C = (C^{ij})_{1 \leq i, j \leq d}$  is the quadratic variation of the continuous local martingale part  $X^c$  of  $X$ , that is,  $C^{ij} = \langle X^{i,c}, X^{j,c} \rangle$ ;
- (iii)  $\nu$  is the (predictable) compensating measure of the jump measure  $\mu = \mu^X$  of  $X$ , as defined in (1.41).

Sometimes one says “predictable characteristics” or “local characteristics” of the semimartingale  $X$ . The name comes from the fact that, when  $X$  is a Lévy process with characteristics triple  $(b, c, F)$ , then its characteristics in the semimartingale sense are

$$B_t(\omega) = bt, \quad C_t(\omega) = ct, \quad \nu(\omega, dt, dx) = dt \otimes F(dx). \quad (1.63)$$

So  $(B, C, \nu)$  and  $(b, c, F)$  convey the same information. Note that in this case the characteristics  $(B, C, \nu)$  are *not random*. This turns out to be “necessary and sufficient.” More precisely, a semimartingale  $X$  has non-random characteristics if and only if it has  $(\mathcal{F}_t)$ -independent increments; and it has characteristics of the form (1.63) if and only if it is an  $(\mathcal{F}_t)$ -Lévy process.

The reader should not be misled: unlike for Lévy processes (or more generally for processes with independent increments), the characteristics do *not* characterize the law of the process in general. Whether they do

characterize the law in specific cases is an important problem, closely related to the uniqueness of (weak) solutions of some associated SDEs, and we will come back to this point later.

Now, although the characteristics do not always characterize the process  $X$ , they provide useful information, especially on the jumps. For example, we have a (partial) analogue of (1.48): for any nonnegative Borel function  $g$  on  $\mathbb{R}^d$  with  $g(0) = 0$  and any  $t > 0$ ,

$$\begin{aligned} &\text{the two sets } \{(g \wedge 1) * \nu_t < \infty\} \text{ and } \{g * \mu_t < \infty\} \\ &\text{are a.s. equal} \end{aligned} \tag{1.64}$$

and in particular, similar to (1.50), for any  $p \geq 0$  we have

$$\begin{aligned} &\text{the two sets } \{(\|x\|^p \wedge 1) * \nu_t < \infty\} \\ &\text{and } \left\{ \sum_{s \leq t} \|\Delta X_s\|^p < \infty \right\} \text{ are a.s. equal.} \end{aligned} \tag{1.65}$$

The triple  $(B, C, \nu)$  satisfies a number of structural properties, coming from its definition. The process  $C$  is such that  $C_{t+s} - C_t$  is a symmetric nonnegative  $d \times d$  matrix, as already mentioned. Next,  $B$  and  $\nu$  are predictable, with finite variation for  $B$ , but there are other necessary requirements, namely there is a version of  $(B, \nu)$  satisfying identically

$$\begin{aligned} &\|x\|^2 \wedge 1 * \nu_t(\omega) < \infty, \quad \nu(\omega, \{t\} \times \mathbb{R}^d) \leq 1, \\ &\|\Delta B_t(\omega)\| \leq 1, \quad \nu(\omega, \{t\} \times \mathbb{R}^d) = 0 \Rightarrow \Delta B_t(\omega) = 0. \end{aligned} \tag{1.66}$$

The analogy with Lévy processes goes further; for example we have a formula similar to the Lévy-Itô decomposition, that is,

$$X_t = X_0 + B_t + X_t^c + (x 1_{\{\|x\| \leq 1\}}) * (\mu - \nu)_t + (x 1_{\{\|x\| > 1\}}) * \mu_t, \tag{1.67}$$

where the stochastic integral above makes sense because of the first property in (1.66). However, this may look like (1.51), but  $\mu$  is not a Poisson measure and  $X^c$  is not a Brownian motion here.

## 1.4 Itô Semimartingales

### 1.4.1 The Definition

As seen above, quite a few properties of Lévy processes extend to general semimartingales, but there are also big differences, like the fact that  $\nu(\omega, \{t\} \times \mathbb{R}^d)$  may be positive. There is, however, a class of semimartingales which is a more direct extension of Lévy processes:

**Definition 1.16.** *A  $d$ -dimensional semimartingale  $X$  is an Itô semimartingale if its characteristics  $(B, C, \nu)$  are absolutely continuous with respect to the Lebesgue measure, in the sense that*

$$B_t = \int_0^t b_s ds, \quad C_t = \int_0^t c_s ds, \quad \nu(dt, dx) = dt F_t(dx), \quad (1.68)$$

where  $b = (b_t)$  is an  $\mathbb{R}^d$ -valued process,  $c = (c_t)$  is a process with values in the set of all  $d \times d$  symmetric nonnegative matrices, and  $F_t = F_t(\omega, dx)$  is for each  $(\omega, t)$  a measure on  $\mathbb{R}^d$ . The terms  $(b_t, c_t, F_t)$  are called the spot characteristics of  $X$ .

These  $b_t$ ,  $c_t$  and  $F_t$  necessarily have some additional measurability properties, so that (1.68) makes sense: we may choose  $b_t$  and  $c_t$  predictable (or simply progressively measurable, this makes no difference in the sequel and does not change the class of Itô semimartingales), and  $F_t$  is such that  $F_t(A)$  is a predictable process for all  $A \in \mathcal{R}^d$  (or progressively measurable, again this makes no difference). The last three requirements of (1.66) are automatically fulfilled here, and we can and will choose a version of  $F_t$  which satisfies identically

$$\begin{aligned} \int (\|x\|^2 \wedge 1) F_t(\omega, dx) &< \infty \\ \text{and } \int_0^t ds \int (\|x\|^2 \wedge 1) F_t(\omega, dx) &< \infty. \end{aligned} \quad (1.69)$$

In view of (1.63), any Lévy process is an Itô semimartingale.

There is an apparent contradiction between Definitions 1.1 and 1.16: a continuous Itô semimartingale in the former sense (or Brownian semimartingale) is also Itô in the latter sense, but the converse is not obvious when  $X$  is continuous. However, this contradiction is only apparent.

To be more specific, assume for example that  $d = 1$  and that  $X$  is continuous, and an Itô semimartingale in the sense of (1.16), and also that  $B_t = 0$  identically, so  $X = X^c$ . The question becomes: can we find a Brownian motion  $W$  and a progressively measurable process  $H$  such that  $X_t = \int_0^t H_s dW_s$ ? If this is the case, we necessarily have  $H_t^2 = c_t$ , so since we know that  $c_t \geq 0$ , natural candidates are

$$H_t = \sqrt{c_t}, \quad W_t = \int_0^t \frac{1}{H_s} dX_s.$$

Of course the second integral does not make sense if  $c_t$  vanishes somewhere, and it should thus be replaced by

$$W_t = \int_0^t \frac{1}{H_s} 1_{\{H_s \neq 0\}} dX_s.$$

This defines a continuous martingale  $W$  starting at 0, and with quadratic variation  $\int_0^t 1_{\{c_s > 0\}} ds$ , which may be different from  $t$ : hence  $W$  is *not* necessarily a Brownian motion; when for example  $c_t = 0$  for all  $t \leq 1$  it is quite possible that the  $\sigma$ -field  $\mathcal{F}_1$  is trivial and thus there simply does *not* exist any  $(\mathcal{F}_t)$ -Brownian motion on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ .

The solution to this problem needs an extension of the original probability space. The need for an extension also arises in other contexts in this book, so we devote the next subsection to the general question of “extending” the probability space.

### 1.4.2 Extension of the Probability Space

The space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is fixed. Let  $(\Omega', \mathcal{F}')$  be another measurable space, and  $\mathbb{Q}(\omega, d\omega')$  be a transition probability from  $(\Omega, \mathcal{F})$  into  $(\Omega', \mathcal{F}')$ . We can define the products

$$\tilde{\Omega} = \Omega \times \Omega', \quad \tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}', \quad \tilde{\mathbb{P}}(d\omega, d\omega') = \mathbb{P}(d\omega)\mathbb{Q}(\omega, d\omega'). \quad (1.70)$$

The probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  is called an *extension* of  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . Any variable or process which is defined on either  $\Omega$  or  $\Omega'$  is extended in the usual way to  $\tilde{\Omega}$ , with the same symbol; for example  $X_t(\omega, \omega') = X_t(\omega)$  if  $X_t$  is defined on  $\Omega$ . In the same way, a set  $A \subset \Omega$  is identified with the set  $A \times \Omega' \subset \tilde{\Omega}$ , and we can thus identify  $\mathcal{F}_t$  with  $\mathcal{F}_t \otimes \{\emptyset, \Omega'\}$ , so  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\mathcal{F}_t)_{t \geq 0}, \tilde{\mathbb{P}})$  is a filtered space.

The filtration  $(\mathcal{F}_t)$  on the extended space does not incorporate any information about the second factor  $\Omega'$ . To bridge this gap we consider a bigger filtration  $(\tilde{\mathcal{F}}_t)_{t \geq 0}$  on  $(\tilde{\Omega}, \tilde{\mathcal{F}})$ , that is with the inclusion property

$$\mathcal{F}_t \subset \tilde{\mathcal{F}}_t, \quad \forall t \geq 0.$$

The filtered space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}})$  is then called a *filtered extension* of  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ .

In many, but not all cases the filtration  $(\tilde{\mathcal{F}}_t)$  has the product form

$$\tilde{\mathcal{F}}_t = \bigcap_{s > t} \mathcal{F}_s \otimes \mathcal{F}'_s \quad (1.71)$$

where  $(\mathcal{F}'_t)$  is a filtration on  $(\Omega', \mathcal{F}')$ . Quite often also, the transition probability  $\mathbb{Q}$  has the simple form  $\mathbb{Q}(\omega, d\omega') = \mathbb{P}'(d\omega')$  for some probability on  $(\Omega', \mathcal{F}')$ . In the latter case we say that the extension is a *product extension*, and if further (1.71) holds we say that we have a *filtered product extension*, which is simply the product of two filtered spaces.

A filtered extension is called *very good* if it satisfies

$$\begin{aligned} \omega \mapsto \int 1_A(\omega, \omega') \mathbb{Q}(\omega, d\omega') \text{ is } \mathcal{F}_t\text{-measurable} \\ \text{for all } A \in \tilde{\mathcal{F}}_t, \text{ all } t \geq 0. \end{aligned} \tag{1.72}$$

Under (1.71), this is equivalent to saying that  $\omega \mapsto \mathbb{Q}(\omega, A')$  is  $\mathcal{F}_t$ -measurable for all  $A' \in \mathcal{F}'_t$  and  $t \geq 0$ . A very good filtered extension is very good because it has the following nice properties:

- any martingale, local martingale, submartingale, supermartingale on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is also a martingale, local martingale, submartingale, supermartingale on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}})$
- a semimartingale on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is a semimartingale on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}})$ , with the same characteristics.

Statement (1.72) is equivalent to saying that any bounded martingale on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is a martingale on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}})$ . For example a Brownian motion on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is also a Brownian motion on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}})$  if the extension is very good, and the same for Poisson random measures.

Many extensions are *not* very good: for example take  $\mathbb{Q}(\omega, \cdot)$  to be the Dirac mass  $\varepsilon_{U(\omega)}$ , on the space  $(\Omega', \mathcal{F}') = (\mathbb{R}, \mathcal{R})$  endowed with the filtration  $\mathcal{F}'_t = \mathcal{F}'$  for all  $t$ , and where  $U$  is an  $\mathbb{R}$ -valued variable on  $(\Omega, \mathcal{F})$  which is not measurable with respect to the  $\mathbb{P}$ -completion of  $\mathcal{F}_1$ , say. Then  $\mathbb{Q}(\omega, A') = 1_{A'}(U(\omega))$  is not  $\mathcal{F}_1$ -measurable in general, even when  $A' \in \mathcal{F}'_1$ , and the extension is not very good.

### 1.4.3 The Grigelionis Form of an Itô Semimartingale

We are now ready to give our representation theorem. The difficult part comes from the jumps of our semimartingale, and it is fundamentally a representation theorem for integer-valued random measure in terms of a Poisson random measure, a result essentially due to Grigelionis (1971). The form given below is Theorem (14.68) of Jacod (1979), and we will call the representation given here the *Grigelionis form* of the semimartingale  $X$ .

We have the  $d$ -dimensional Itô semimartingale  $X$  with characteristics  $(B, c, \nu)$  given by (1.68). Moreover,  $d'$  is an arbitrary integer with  $d' \geq d$ , and  $E$  is an arbitrary Polish space with a  $\sigma$ -finite and infinite measure

$\lambda$  having no atom, and  $\underline{q}(dt, dx) = dt \otimes \lambda(dx)$ . Then one can construct a very good filtered extension  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}})$ , on which are defined a  $d'$ -dimensional Brownian motion  $W$  and a Poisson random measure  $\underline{p}$  on  $\mathbb{R}_+ \times E$  with intensity measure  $\lambda$ , such that

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + (\delta 1_{\{\|\delta\| \leq 1\}}) \star (\underline{p} - \underline{q})_t + (\delta 1_{\{\|\delta\| > 1\}}) \star \underline{p}_t, \quad (1.74)$$

where  $\sigma_t$  is an  $\mathbb{R}^d \otimes \mathbb{R}^{d'}$ -valued process on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  which is predictable (or only progressively measurable), and  $\delta$  is a predictable  $\mathbb{R}^d$ -valued function on  $\Omega \times \mathbb{R}_+ \times E$ , both being such that the integrals in (1.74) make sense.

The process  $b_t$  is the same here and in (1.68), and we have close connections between  $(\sigma_t, \delta(t, z))$  and  $(c_t, F_t)$ . Namely, a version of the spot characteristics  $c_t$  and  $F_t$  is given by the following:

$$\begin{aligned} \bullet c_t(\omega) &= \sigma_t(\omega) \sigma_t^*(\omega) \\ \bullet F_t(\omega, \cdot) &= \text{the image of the measure } \lambda \\ &\text{restricted to the set } \{x : \delta(\omega, t, x) \neq 0\} \\ &\text{by the map } x \mapsto \delta(\omega, t, x). \end{aligned} \quad (1.75)$$

Conversely, any process of the form (1.74) (with possibly  $b$ ,  $\sigma$  and  $\delta$  defined on the extension instead of  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ ) is an Itô semimartingale on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}})$ , and on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  as well if it is further adapted to  $(\mathcal{F}_t)$ . Therefore, the formula (1.74) may serve as the definition of Itô semimartingales, if we do not mind extending the space, and for practical applications we do not mind! Therefore, in the sequel we freely use the Grigelionis form above, pretending that it is defined on our original filtered space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ .

There is a lot of freedom in the choice of the extension, of the space  $E$  and the function  $\delta$ , and even of the dimension  $d'$  and the process  $\sigma$ : for the latter, for example, the requirement being  $\sigma_t \sigma_t^* = c_t$ , we can always take an arbitrary  $d' \geq d$ , or more generally a  $d'$  not smaller than the maximal rank of the matrices  $c_t(\omega)$ . A natural choice for  $E$  is  $E = \mathbb{R}^d$ , but this is not compulsory and we may take in all cases  $E = \mathbb{R}$  with  $\lambda$  being the Lebesgue measure. For example, if we have several Itô semimartingales, and even countably many of them, we can use the same measure  $\underline{p}$  for representing all of them at once.

For a Lévy process, the Grigelionis form coincides with its Lévy-Itô representation, upon taking  $\underline{p} = \mu$  and  $\delta(\omega, t, x) = x$ . More generally, in equation (1.74), the term  $(\delta 1_{\{\|\delta\| \leq 1\}}) \star (\underline{p} - \underline{q})_t$  corresponds to the small

jumps of the process, while the term  $(\delta 1_{\{\|\delta\|>1\}}) \star \underline{p}_t$  corresponds to the big jumps of the process.

### 1.4.4 A Fundamental Example: Stochastic Differential Equations Driven by a Lévy Process

We have already mentioned stochastic differential equations (SDE) of the form (1.18), driven by a Brownian motion. Natural extensions of them are SDEs driven by a Lévy process  $Z$ , written as

$$dX_t = a(X_{t-}) dZ_t, \quad X_0 = Y. \tag{1.76}$$

Here  $Z$  is a  $q$ -dimensional Lévy process, the “solution”  $X$  will be  $d$ -dimensional, so  $Y$  is an  $\mathcal{F}_0$ -measurable  $\mathbb{R}^d$ -valued variable, and  $a$  is a function from  $\mathbb{R}^d$  into  $\mathbb{R}^d \times \mathbb{R}^q$ . As for (1.18), a (strong) solution is a càdlàg adapted process  $X$  which satisfies the following, written componentwise:

$$X_t^i = Y^i + \sum_{j=1}^q \int_0^t a(X_{s-})^{ij} dZ_s^j. \tag{1.77}$$

The fact that we take the left limit  $a(X_{s-})$  in the integral above, and not  $a(X_s)$ , is absolutely crucial, because the integrand needs to be predictable, otherwise the stochastic integral *a priori* makes no sense. Of course when  $Z$  is continuous, hence the solution  $X$  as well, we have  $a(X_{s-}) = a(X_s)$ .

Note that (1.18) is a special case of (1.76), although it may not be apparent at first glance: in (1.18), if  $W$  is  $q'$ -dimensional, we take  $q = q' + 1$  and  $Z^j = W^j$  for  $j \leq q'$  and  $Z_t^q = t$ , and the coefficient  $a$  defined by  $a^{ij} = \sigma^{ij}$  when  $j \leq q'$  and  $a^{iq} = b^i$ .

Here again, a wide variety of conditions on  $a$  imply existence and/or uniqueness of the solution of (1.76). The simplest one is that  $a$  is locally Lipschitz with at most linear growth, but many other conditions exist, sometimes related with the specific properties of the driving Lévy process  $Z$ .

Now, assuming (1.77), the process  $X$  is of course a semimartingale, and even an Itô semimartingale. If  $(\beta, \gamma, F)$  is the characteristic triple of  $Z$ , and using the Lévy-Itô representation of  $Z$  with  $\underline{p} = \nu^Z$  and  $\underline{q} =$

$\nu^X = dt \otimes F(dz)$  and  $\sigma$  such that  $\sigma\sigma^* = \gamma$ , we can rewrite (1.77) as

$$\begin{aligned} X_t^i &= X_0^i + \sum_{j=1}^q \int_0^t a(X_{s-})^{ij} \beta^j ds + \sum_{j,k=1}^q \int_0^t a(X_{s-})^{ij} \sigma^{jk} dW_s^k \\ &\quad + \sum_{j=1}^q \int_0^t \int_{\mathbb{R}^q} a(X_{s-})^{ij} z^j 1_{\{\|z\| \leq 1\}} (\underline{p} - \underline{q})(ds, dz) \\ &\quad + \sum_{j=1}^q \int_0^t \int_{\mathbb{R}^q} a(X_{s-})^{ij} z^j 1_{\{\|z\| > 1\}} \underline{p}(ds, dz). \end{aligned}$$

Therefore, the characteristics  $(B, C, \nu)$  of  $X$  take the form (1.68) with the following (where  $a(x)z$  stands for the  $d$ -dimensional vector with components  $\sum_{j=1}^q a(x)^{ij} z^j$ ):

$$\begin{aligned} b_t^i &= \sum_{j=1}^q a(X_{t-})^{ij} (\beta^j + \int_{\mathbb{R}^q} z^j (1_{\{\|z\| \leq 1\}} - 1_{\{\|a(X_{t-})z\| \leq 1\}}) F(dz)) \\ c_t^{ij} &= \sum_{k,l=1}^q a(X_{t-})^{ik} \gamma^{kl} a(X_{t-})^{jl} \\ F_t(\omega, dx) &= \text{the image of the measure } F \\ &\quad \text{by the map } z \mapsto a(X_{t-}(\omega))z. \end{aligned} \tag{1.78}$$

These characteristics have a complicated form, although they indeed come naturally as functions of the coefficient  $a$ : the problem of finding a process which is an Itô semimartingale with characteristics given *a priori* in the form (1.78) reduces in fact to solving Equation (1.76).

Now, we also have, in a somewhat more immediate way, the Grigelionis form of  $X$ , provided we take for  $W$  and  $\underline{p}$  the terms coming in the Lévy-Itô decomposition of  $Z$ . Namely, in (1.74) the process  $b_t$  is the same (complicated) process as above, but  $\sigma_t$  and  $\delta$  take the simple form

$$\sigma_t^{ik} = \sum_{j=1}^q a(X_{t-})^{ij} \sigma^{jk}, \quad \delta(t, z)^i = \sum_{j=1}^q a(X_{t-})^{ij} z^j.$$

In fact, equations like (1.68) are not really the most natural ones to consider in a discontinuous setting. It is also useful to consider equations which are driven directly by a Brownian motion  $W$  and a (general) Poisson random measure  $\underline{p}$  on  $\mathbb{R}_+ \times E$  for some “abstract” space  $E$  and with intensity measure  $\underline{q} = dt \otimes \lambda(dz)$ . This amounts to considering an

equation of the form

$$\begin{aligned}
 X_t &= Y + \int_0^t b(X_{s-}) ds + \int_0^s \sigma(X_{s-}) dW_s \\
 &+ \int_0^t \int_E v(X_{s-}, z) 1_{\{\|v(X_{s-}, z)\| \leq 1\}} (\underline{p} - \underline{q})(ds, dz) \quad (1.79) \\
 &+ \int_0^t \int_E v(X_{s-}, z) 1_{\{\|v(X_{s-}, z)\| > 1\}} \underline{p}(ds, dz),
 \end{aligned}$$

where  $b$  is an  $\mathbb{R}^d$ -valued function on  $\mathbb{R}^d$  and  $\sigma$  is an  $\mathbb{R}^d \otimes \mathbb{R}^q$ -valued function on  $\mathbb{R}^d$  (where  $q$  is the dimension of  $W$ ) and  $v$  is an  $\mathbb{R}^d$ -valued function on  $\mathbb{R}^d \times E$ . This type of equation includes (1.77) and immediately gives the Grigelionis form of the solution. When existence and uniqueness hold for all initial conditions  $Y$ , the solution is a homogeneous Markov process, and the restriction of its infinitesimal generator to the  $C^2$  functions takes the form

$$\begin{aligned}
 \mathcal{A}f(x) &= \sum_{i=1}^d b(x)^i f'_i(x) + \frac{1}{2} \sum_{i,j=1}^d c(x)^{ij} f''_{ij}(x) \\
 &+ \int_E \left( f(x + v(x, z)) - f(x) \right. \\
 &\quad \left. - \sum_{i=1}^d f'_i(x) v(x, z)^i 1_{\{\|v(x, z)\| \leq 1\}} \right) \lambda(dz), \quad (1.80)
 \end{aligned}$$

where  $c = \sigma \sigma^*$ . This extends (1.17), and one may show that any homogeneous Markov process which is an Itô semimartingale is of this form.

More generally even, one can interpret the Grigelionis form (1.74) as a generalized SDE similar to (1.79), but with “coefficients”  $b_t(\omega)$ ,  $\sigma_t(\omega)$  and  $\delta(\omega, t, z)$  which may depend on the whole past of  $X$  before time  $t$ , and also on  $\omega$  in an arbitrary (predictable) way. In this setting the infinitesimal generator is replaced by the so-called extended generator

$$\begin{aligned}
 \mathcal{A}_t f(x) &= \sum_{i=1}^d b_t^i f'_i(x) + \frac{1}{2} \sum_{i,j=1}^d c_t^{ij} f''_{ij}(x) \\
 &+ \int_E \left( f(x + \delta(t, z)) - f(x) \right. \\
 &\quad \left. - \sum_{i=1}^d f'_i(x) \delta(t, z)^i 1_{\{\|\delta(t, z)\| \leq 1\}} \right) \lambda(dz),
 \end{aligned}$$

so  $\mathcal{A}_t$  is a second order integro-differential operator mapping the  $C^2$  functions into the set of random variables. This extended generator is no

longer the generator of a semi-group, but the characteristic martingale property of the generator is preserved, and it reads as follows: for any  $C^2$  function  $f$ , the process

$$M_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{A}_s f ds$$

is a local martingale.

As a consequence of all these considerations, one may state (in a somewhat heuristic way) that the characteristics  $(B, C, \nu)$ , or equivalently (and perhaps more appropriately)  $(b_t, c_t, F_t)$ , determine the dynamics of the process. They are thus of fundamental importance for modeling purposes. More precisely, the problem of describing the process  $X$  is often considered as “solved” when one knows the characteristics  $(B, C, \nu)$ , in connection with  $X$  itself and perhaps with other random inputs.

## 1.5 Processes with Conditionally Independent Increments

In most problems considered in this book we start with an underlying process  $X$  which is an Itô semimartingale on some given filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . We then consider various functionals of  $X$ , such as the approximate quadratic variation defined as the left side of (1.61) or (1.62). We are interested first in the convergence in probability of these functionals, as in Theorem 1.14, and ultimately in the rate of convergence and, whenever possible, in the “second order” asymptotic behavior, or Central Limit Theorem, associated with the convergence in probability. As it turns out, the limit when such a CLT holds will almost always be defined on an extension of  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , as introduced in (1.70), and what will be available is the law of the limit under the measures  $\mathbb{Q}(\omega, d\omega')$ , or equivalently the  $\mathcal{F}$ -conditional law of the limit; the reader can look immediately at Chapter 3 for a (relatively) simple case for which this situation occurs.

Actually, the putative limits will (almost) always be a stochastic process  $U$  belonging to a special and relatively restricted class, namely the processes having  $\mathcal{F}$ -conditionally independent increments. The aim of this section is to describe those processes, which have an interest by themselves and can in fact be defined independently of any limiting procedure. Toward this aim, we first need to consider processes which extend Lévy processes.

### 1.5.1 Processes with Independent Increments

A  $q$ -dimensional process  $U$  defined on a space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is called a *process with independent increments*, relative to  $(\mathcal{F}_t)$ , if it is càdlàg, with  $U_0 = 0$ , adapted, and the increment  $U_{t+s} - U_t$  is independent of  $\mathcal{F}_t$  for all  $s, t \geq 0$ . In other words, we have all properties of Definition 1.1 except that the law of  $U_{t+s} - U_t$  may depend on both  $s$  and  $t$ . When  $(\mathcal{F}_t) = (\mathcal{F}_t^U)$  is the filtration generated by  $U$  we simply say a *process with independent increments*.

When the increments are stationary,  $U$  is simply a Lévy process. Otherwise, it still exhibits similar features. In particular, when it is a semimartingale, its characteristics relative to  $(\mathcal{F}_t)$  are deterministic (this property characterizes semimartingales with independent increments). We can thus decompose its third (deterministic) characteristic  $\nu$  as a sum  $\nu = \nu^c + \nu^d$ , where, using the notation  $D = \{t > 0 : \nu(\{t\} \times \mathbb{R}^q) > 0\}$  and its complement  $D^c$  in  $\mathbb{R}_+$ ,

$$\nu^c(dt, dx) = \nu(dt, dx)1_{D^c}(t), \quad \nu^d(dt, dx) = \sum_{s \in D} \varepsilon_s(dt) \otimes \nu(\{s\}, dx)$$

(there is no randomness here). We thus have two types of jumps: those occurring at a time outside the countable set  $D$ , which are like the jumps of a Lévy process except that the associated Poisson measure is non-homogeneous (it admits  $\nu^c$  as its intensity measure), and those occurring at a time in  $D$ , which are called *fixed times of discontinuity* because we have

$$\mathbb{P}(\Delta U_t \neq 0) > 0 \iff t \in D.$$

Although it is possible to describe all processes with independent increments, we restrict our attention to those which are encountered as limits in this book. This class of processes, denoted for short as  $\mathcal{L}_0^q$  below ( $q$  stands for the dimension), is in fact rather special: it is the class of all processes of the form

$$U_t = \sum_{n: t_n \leq t} v(t_n) Y_n + \int_0^t v'(s) dW'_s \tag{1.81}$$

where

- $v$  is a measurable function on  $\mathbb{R}_+$ , with dimension  $q \times Q$
- $v'$  is a measurable locally square-integrable function on  $\mathbb{R}_+$ , with dimension  $q \times Q'$
- $W'$  is a  $Q'$ -dimensional Brownian motion
- $t_n \in (0, \infty]$ ,  $t_n \neq t_m$  if  $n \neq m$  and  $t_n < \infty$
- the  $\mathbb{R}^Q$ -valued variables  $Y_n$  are i.i.d. and independent of  $W'$ ,

and we suppose that at least one of the following two sets of conditions holds:

- (i)  $0 < \mathbb{E}(\|Y_1\|) < \infty$  and  $\sum_{n: t_n \leq t} \|v(t_n)\| < \infty$
- (ii)  $\mathbb{E}(Y_1) = 0$ ,  $0 < \mathbb{E}(\|Y_1\|^2) < \infty$  and  $\sum_{n: t_n \leq t} \|v(t_n)\|^2 < \infty$ .

Under (i) the first sum in (1.81) is absolutely convergent. Under (ii) this is no longer necessarily true, but it converges in  $\mathbb{L}^2$  and the resulting process is a square-integrable martingale. Since  $\mathbb{P}(Y_n \neq 0) > 0$ , the set of fixed times of discontinuity of  $U$  is the set of all finite  $t_n$ 's such that  $v(t_n) \neq 0$ .

The following (obvious) fact is important:

Any  $U$  in  $\mathcal{L}_0^q$  has independent increments, and its law is determined by the functions  $v$  and  $v'$ , the set (1.82)  
 $D = \{t_n : n \geq 1\} \cap (0, \infty)$ , and the law  $\eta$  of the variables  $Y_n$ ,

and we can relabel the sequence  $t_n$  at will. Finally, under the previous conditions we have, relative to the filtration  $(\mathcal{F}_t^U)$  generated by  $U$ :

- $U$  is a centered Gaussian process  
 $\Leftrightarrow$  either  $\eta$  is a centered Gaussian law or  $v = 0$  on  $D$
- $U$  is a continuous centered Gaussian martingale (1.83)  
 $\Leftrightarrow v = 0$  on  $D$ .

### 1.5.2 A Class of Processes with $\mathcal{F}$ -Conditionally Independent Increments

Generally speaking, a  $q$ -dimensional process  $U$  on an extension  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  of the original probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , see (1.68), is called a *process with  $\mathcal{F}$ -conditionally independent increments* if, for  $\mathbb{P}$ -almost all  $\omega$ , it is a process with independent increments under the measure  $\mathbb{Q}(\omega, \cdot)$ .

As previously, we restrict our attention to those  $U$  which, under  $\mathbb{Q}(\omega, \cdot)$ , belong to the class  $\mathcal{L}_0^q$  described above, and we even suppose that the law  $\eta$  (as in (1.82)) is the same for all  $\omega$ . Therefore, in order to construct such a process, we start with the following ingredients:

- a progressively measurable process  $V$  with dimension  $q \times Q$
- a progressively measurable, locally square-integrable process  $V'$  with dimension  $q \times Q'$
- a probability measure  $\eta$  on  $\mathbb{R}^Q$
- a sequence  $T_n$  of stopping times, with  $T_n \neq T_m$  on the set  $\{T_n < \infty\}$  if  $n \neq m$ .

We also suppose that  $V$  and  $\eta$  satisfies at least one of the following two conditions:

$$\begin{aligned} \text{(i)} \quad & 0 < \int \|x\| \eta(dx) < \infty \text{ and } \sum_{n: T_n \leq t} \|V_{T_n}\| < \infty \\ \text{(ii)} \quad & 0 < \int \|x\|^2 \eta(dx) < \infty, \quad \int x \eta(dx) = 0, \\ & \text{and } \sum_{n: T_n \leq t} \|V_{T_n}\|^2 < \infty. \end{aligned} \tag{1.84}$$

The construction of a version of the process  $U$  is now quite simple. We choose an extra space  $(\Omega', \mathcal{F}', \mathbb{P}')$  endowed with the following objects:

- a  $q'$ -dimensional Brownian motion  $W'$
  - a sequence  $(Y_i)_{i \geq 1}$  of i.i.d. variables, independent of  $W'$ , with law  $\eta$ .
- (1.85)

Then we define the product extension

$$\tilde{\Omega} = \Omega \times \Omega', \quad \tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}', \quad \tilde{\mathbb{P}} = \mathbb{P} \otimes \mathbb{P}'. \tag{1.86}$$

We also consider the smallest filtration  $(\tilde{\mathcal{F}}_t)_{t \geq 0}$  on  $\tilde{\Omega}$  which contains  $(\mathcal{F}_t)_{t \geq 0}$  (recall that any variable on  $\Omega$  or  $\Omega'$  can be considered as a variable on the product  $\tilde{\Omega}$ , so  $\mathcal{F}_t$  is also a  $\sigma$ -field on  $\tilde{\Omega}$ ), and to which  $W'$  is adapted, and such that each  $Y_n$  is  $\tilde{\mathcal{F}}_{T_n}$ -measurable. Due to the independence built in (1.85) and (1.86), the process  $W'$  is an  $(\tilde{\mathcal{F}}_t)$ -Brownian motion and the processes  $V, V'$  are  $(\tilde{\mathcal{F}}_t)$ -progressively measurable. Then it is an easy job to prove that the next formula defines a process  $U$  on the extension, with all the required properties:

$$U_t = \sum_{n: T_n \leq t} V_{T_n} Y_n + \int_0^t V'_s dW'_s. \tag{1.87}$$

It is useful to have explicit formulas for the conditional first and second moments of the process  $U$  in this setting. Letting

$$\begin{aligned} M_j &= \int x^j \eta(dx), \quad M_{jk} = \int x^j x^k \eta(dx), \\ C_t'^{jk} &= \sum_{l=1}^{q'} \int_0^t V_s'^{jl} V_s'^{kl} ds, \end{aligned} \tag{1.88}$$

we get

$$\begin{aligned} \text{(1.84)-(i)} \quad & \Rightarrow \quad \mathbb{E}(U_t^j \mid \mathcal{F}) = \sum_{s \leq t} \sum_{k=1}^q V_s^{jk} M_k \\ \text{(1.84)-(ii)} \quad & \Rightarrow \quad \begin{cases} \mathbb{E}(U_t^j \mid \mathcal{F}) = 0 \\ \mathbb{E}(U_t^j U_t^k \mid \mathcal{F}) = C_t'^{jk} \\ \quad + \sum_{s \leq t} \sum_{l,m=1}^q V_s^{jl} V_s^{km} M_{lm}. \end{cases} \end{aligned}$$

Finally, if  $D(\omega) = \{T_n(\omega) : n \geq 1\} \cap (0, \infty)$ , the next two equivalences follow from (1.83):

- $U$  is  $\mathcal{F}$ -conditionally a centered Gaussian process  
 $\Leftrightarrow \eta$  is centered Gaussian or  $V = 0$  on  $D$
  - $U$  is  $\mathcal{F}$ -conditionally a continuous  
 centered Gaussian martingale  $\Leftrightarrow V = 0$  on  $D$ .
- (1.89)

We conclude with an important particular case. In many instances, the process  $U$  given by (1.87) is continuous, that is,  $U_t = \int_0^t V'_s dW'_s$ . The law of a continuous centered process  $U$  with independent increments is completely specified by its covariance matrix  $\mathbb{E}(U_t^i U_t^j)$ , as a function of time. This is of course still true when  $U$  is continuous, with  $\mathcal{F}$ -conditional mean 0 and with  $\mathcal{F}$ -conditionally independent increments. That is, the  $\mathcal{F}$ -conditional law of  $U$  in this case is completely specified by the process  $C'$  given in (1.88).

In practice, one usually goes the other way around: one starts with a  $q \times q$ -dimensional process  $C'$  of the form

$$C'_t = \int_0^t c'_s ds,$$

with  $c'$  adapted, with values in the set of nonnegative symmetric  $q \times q$  matrices. Then, what precedes gives us a process  $U$  which is continuous and  $\mathcal{F}$ -conditionally centered and with independent increments (equivalently:  $\mathcal{F}$ -conditionally a centered Gaussian martingale). For this, it suffices to choose a  $q \times Q$ -dimensional process  $V$  which is adapted and is a square root of  $c'$ , that is,  $c'_t = V_t V_t^*$ ; this is always possible by a measurable selection theorem, up to null sets, and with any choice of  $Q$  with  $Q \geq q$ , or even with  $Q \geq \sup_{\omega, t} \text{rank}(c'_t(\omega))$ .