Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis and Computation

Floyd B. Hanson
University of Illinois
Chicago, Illinois, USA

Chapter 13 Applied Guide to Abstract Stochastic Processes

Copyright © 2006 by the Society for Industrial and Applied Mathematics.

May 24, 2006

DRAFT

DRAFT

Chapter 13

Applied Guide to Abstract Theory of Stochastic Processes:

*Mathematicians are like Frenchmen: Whatever you say
to them they translate into their own language
and forthwith it is something entirely different.*
—Johann Wolfgang von Goethe.

*Since the mathematicians have invaded the theory of
relativity, I do not understand it myself anymore.*
—Albert Einstein (1879-1955),
`http://en.wikiquote.org/wiki/Mathematics`.

*Martingale (1589): Any of several systems of betting
in which a player increases his stake,
usually by doubling each time he loses a bet.*
—Digital Webster, definition 3, 1992.

*Martingales are treated because of their great importance,
but they are not used as a tool in this book.*
—William (Willy) Feller (1906-1970), p. 209 in [84].

*The concept of martingales is due to P. Lévy,
but it was J. L. Doob who realized its unexpected potential
and developed the theory.*
—William (Willy) Feller (1906-1970), p. 210 in [84].

*Our view of Brownian motion never focused too closely
on the underlying measure space, and, by and large,
we have profited from keeping a respectful distance.*
—J. Michael Steele, p. 218 in [251].

DRAFT

2006/5/24
page 484

This chapter briefly introduces more of the abstract analytical methods, such as measure theoretic methods, Martingale methods, Radon-Nikodým derivatives, Girsanov's theorem, Itô processes, Lévy processes, characteristic functions and exponents, Lévy-Klintchine formula, jump-diffusion process comparisons and other topics from the applied point of view as a bridge to more abstract methods.

The purpose of this chapter is to supply some insightful and useful background to make the more abstract literature on stochastic processes and control more accessible to general students in applied mathematics, statistics, computer science, applied science and engineering. The overall approach in this book is designed to start from the common calculus and analysis constructs of entry level graduate students in these applied areas by evolving these constructs to those of applied stochastic processes and control, much as genes have evolved by small but powerful changes. However, students still need to understand the important results that come from using more abstract methods.

The applied motivation is to solve problems with a combination of analytical and computational methods. These problems may have great complexity in terms on nonlinearities in the state and other dependencies. It is necessary to train both students and researchers from a broad range of areas in science and engineering to solve large scale problems. In the abstract approach the emphasis is not necessarily to solve applied problems, but to prove existence, uniqueness and convergence, often in very abstract language. However, sometimes the conditions of the proofs are too restrictive, so as to exclude many complex and large scale applications. Proofs as such are not given in this chapter, but some formal applied derivations are given and readers can refer to the list of references at the end of the chapter for more rigorous treatments.

13.1 Very Basic Probability Measure Background

In order to keep things simple and concise, it will be necessary to compromise on completeness, but keep sufficient detail for a coherent story. The notation will be somewhat different from the usual, if there is such a thing as usual notation, so that we can avoid conflict with the stochastic process notation where possible. The symbols are also selected so that they are related to what the quantity signifies, where possible.

13.1.1 Mathematical Measure Theory Basics

The starting point will be some notions of measure theory and its algebra, called σ-algebra. Measure theory provides an abstract generalization of integration theory including expectations, and distributions that are based on counts, intervals, areas, volumes and mass to that of general sets. The ultimate goal is **probability measure**, but the presentation begins with the foundations in the more general mathematical measure theory.

2006/5/24
page 485

Measure σ-Algebra Definition:

Let $\mathcal{U}$ be a nonempty set called the **universe**, but really is only the principal set of interest. Let $\mathbf{\Sigma}$ be a collection of subsets on $\mathcal{U}$.

Definition 13.1. *$\mathbf{\Sigma}$ is a **$\boldsymbol{\sigma}$-algebra** if*

- *$\emptyset \in \mathbf{\Sigma}$, i.e., the empty set $\emptyset$ is included.*
- *$\mathcal{U} \in \mathbf{\Sigma}$, i.e., the universe $\mathcal{U}$ is included.*
- *The set $\mathcal{S} \in \mathbf{\Sigma} \Longrightarrow \mathcal{S}^c \in \mathbf{\Sigma}$, i.e., its complement $\mathcal{S}^c$ with respect to $\mathcal{U}$ is included too, i.e., verifying that $\mathcal{S} \cup \mathcal{S}^c = \mathcal{U}$.*
- *If $\{\mathcal{S}_i \in \mathbf{\Sigma} : i = 1 : n\}$ is a sequence of subsets, then the union $\bigcup_{i=1}^{n} \mathcal{S}_i \in \mathbf{\Sigma}$, i.e., additive closure under unions.*
- *If so, then $\{\mathcal{U}, \mathbf{\Sigma}\}$ is called a **measurable space**.*

Often the symbol Ω is used for the general universe $\mathcal{U}$ and the symbol $\mathcal{F}$ is used for the σ-algebra $\mathbf{\Sigma}$. Recall that the **union** of two sets

$$\mathcal{S}_1 \cup \mathcal{S}_2 = \{\text{points } X : \ X \in \mathcal{S}_1 \text{ OR } X \in \mathcal{S}_2\},$$

the logical OR playing an important role when translated to probability measures.

A **Borel set** $\mathbf{\Sigma} = \mathcal{B} = \mathcal{B}(\mathbb{R}^{n_x})$ is the σ-algebra of open sets on $\mathcal{U} = \mathbb{R}^{n_x}$, so $\mathcal{B}(\mathbb{R}^{n_x})$ automatically contains all closed sets of $\mathbb{R}^{n_x}$ by complementarity.

Measure Definition:

Definition 13.2. ***The measure $\boldsymbol{\mathcal{M}}$ is a function on the measurable space $\{\mathcal{U}, \mathbf{\Sigma}\}$ that maps $\mathbf{\Sigma}$*** $\longrightarrow [0, \infty)$*, such that*

- *$\mathcal{M}(\emptyset) = 0$, i.e., the empty set $\emptyset$ has **measure zero**.*
- *If for any subset $\mathcal{S} \in \mathbf{\Sigma}$, then $\mathcal{M}(\mathcal{S}) \geq 0$, i.e., **nonnegativity**, as in mass.*
- *If $\{\mathcal{S}_i \in \mathbf{\Sigma} : i = 1, 2, \ldots,\}$ is any countable sequence of **disjoint subsets** (i.e., $\mathcal{S}_i \bigcap \mathcal{S}_j = \emptyset, \ i \neq j$, the intersection is empty), then the measure of the union is the sum of the measures,*

$$\mathcal{M}\left(\bigcup_{i=1}^{\infty} \mathcal{S}_i\right) = \sum_{i=1}^{\infty} \mathcal{M}(\mathcal{S}_i), \tag{13.1}$$

*i.e., **countable additivity**, as in preserving mass under partitioning.*

The triplet $\{\mathcal{U}, \mathbf{\Sigma}, \mathcal{M}\}$ is called a **measure space**. Often the symbol μ is used for the general measure symbol $\mathcal{M}$ used here, but the former conflicts with the use of μ as the mean or drift in this book. Recall that the **intersection** of two sets

$$\mathcal{S}_1 \cap \mathcal{S}_2 = \{\text{points } X : \ X \in \mathcal{S}_1 \text{ AND } X \in \mathcal{S}_2\},$$

the logical AND playing an important role when translated to probability measures.

The nonnegativity measure property $\mathcal{M}(\mathcal{S}) \geq 0$ means that **positive measure** has been defined. Positive meaures, among other things, facilitate convergence proofs, i.e., monotone convergence. However, if for any subset $\mathcal{S} \in \mathbf{\Sigma}$ and $\mathcal{M}(\mathcal{S}) \leq 0$, then $\mathcal{M}(\mathcal{S})$ would be a **negative measure** and negative measure may be needed for some applications in spite of the added awkwardness of the proofs.

Lebesgue Measure Introduction:

If the set $\mathcal{S}$ is measurable, the $\mathcal{M}(\mathcal{S})$ is called the **total mass of the set**, e.g., if $\mathcal{S}$ is an interval $[a, b]$ then it is the length $(b - a)$, if a rectangle $[a, b] \times [c, d]$ then it is its area $(b - a) \cdot (d - c)$, or if a cube $[a, b] \times [a, b] \times [a, b]$ then it is its volume $(b - a)^3$. The closed intervals $[a, b]$, open intervals (a, b) and semi-open intervals $[a, b)$ or $(a, b]$, have the same measure or mass or length of $(b - a)$, since they differ only by **points of zero measure**.

In general, a **Lebesgue measure** is a measure on an n_x dimensional space of real vectors, so the universe is $\mathcal{U} = \mathbb{R}^{n_x}$, a representative set is a hypercube

$$\mathcal{S} = (\mathbf{a}, \mathbf{b}) \equiv (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_{n_x}, b_{n_x}),$$

such that $-\infty < a_i < b_i < +\infty$ and the measure has the form

$$\mathcal{M}(\mathcal{S}) = \prod_{i=1}^{n_x} (b_i - a_i).$$

Alternatively,

$$\mathcal{M}(\mathcal{S}) = \int_{\mathcal{S}} d\mathbf{x}.$$

Lebesgue measure is a special case of Borel measure specialized to $\mathbb{R}^{n_x}$.

Often, the infinitesimal hypercube measure between vector positions from $\mathbf{x}$ to $\mathbf{x} + d\mathbf{x}$ is abbreviated as

$$\mathcal{M}(d\mathbf{x}) = \mathcal{M}((\mathbf{x}, \mathbf{x} + d\mathbf{x})),$$

for compact notation, letting $d\mathbf{x}$ represent the vector-interval set $(\mathbf{x}, \mathbf{x} + d\mathbf{x})$. This also recognizes the translation invariance of the measure of a generalized interval $(\mathbf{x}, \mathbf{x}+d\mathbf{x})$, since its generalized length $\prod_{i=1}^{n_x} dx_i$ is independent of the interval start at $\mathbf{x}$.

DRAFT

2006/5/24
page 487

Dirac Measures:

Another measure that complements the Lebesgue measure is the **Dirac measure** δ_x, for some point in $\mathcal{U}$, having the properties that for some set $\mathcal{S} \subseteq \mathcal{U}$,

$$\delta_x(\mathcal{S}) = \left\{ \begin{array}{ll} 1, & x \in \mathcal{S} \\ 0, & x \notin \mathcal{S} \end{array} \right\}. \tag{13.2}$$

This is the set version of the **Dirac delta function** and apparently the same basic definition as the indicator function $\mathbf{1}_{x\in\mathcal{S}}$, except without the measure infrastructure.

Counting Measures:

For Poisson processes and other discrete applications, there are also counting measures, i.e., when

$$\mathcal{M}(\mathcal{S}) = N(\mathcal{S}) \equiv \textit{ number of elements in set } \mathcal{S}. \tag{13.3}$$

This includes the points of zero measure that **do count**.

Some Properties of Measures:

- The measure space $\{\mathcal{U}, \mathbf{\Sigma}, \mathcal{M}\}$ is **finite** if $\mathcal{M}(\mathcal{U}) < \infty$ and real.
- The measure space $\{\mathcal{U}, \mathbf{\Sigma}, \mathcal{M}\}$ is **σ-finite** if there exists a countable sequence of measurable sets $\{\mathcal{S}_i \in \mathbf{\Sigma} : i = 1, 2, \ldots,\}$ such that $\mathcal{M}(\mathcal{S}_i) < \infty$ and real for all i, i.e., sets of finite measure, and

  $$\mathcal{U} = \bigcup_{i=1}^{\infty} \mathcal{S}_i,$$

  the union of a countable number of sets of finite measure. Note that σ-finite is not necessarily finite, since the set of real intervals $[i, i+1]$, have unit measure which is finite (a Lebesgue measure), but their union is the real line, $\mathcal{U} = \mathbb{R}^1$, which is infinite, so $\mathcal{U}$ is σ-finite while not finite.
- The measure $\mathcal{M}$ is a **montonic function** since if measurable sets $\mathcal{S}_1$ and $\mathcal{S}_2$ ordered $\mathcal{S}_1 \subseteq \mathcal{S}_2$ then $\mathcal{M}(\mathcal{S}_1) \le \mathcal{M}(\mathcal{S}_2)$.
- If $\{\mathcal{S}_i \in \mathbf{\Sigma} : i = 1, 2, \ldots,\}$ is any countable sequence of subsets that are **not necessarily disjoint**, then the measure of the union is only bounded by the sum of the measures,

  $$\mathcal{M}\left(\bigcup_{i=1}^{\infty} \mathcal{S}_i\right) \le \sum_{i=1}^{\infty} \mathcal{M}(\mathcal{S}_i),$$

  unlike the lack of redundancies of disjoint sets given in (13.1).

2006/5/24
page 488

- If $\{\mathcal{S}_i \in \mathbf{\Sigma} : i = 1, 2, \ldots,\}$ is any countable sequence of subsets that are **forward nested** so that $\mathcal{S}_i \subseteq \mathcal{S}_{i+1}$, then the limit of the **measure of the union** has the limiting measure,

$$\mathcal{M}\left(\bigcup_{i=1}^{\infty} \mathcal{S}_i\right) = \lim_{n\to\infty} \mathcal{M}(\mathcal{S}_n),$$

  noting that $\mathcal{M}(\mathcal{S}_i \cup \mathcal{S}_{i+1}) = \mathcal{M}(\mathcal{S}_{i+1})$.

- If $\{\mathcal{S}_i \in \mathbf{\Sigma} : i = 1, 2, \ldots,\}$ is any countable sequence of subsets that are **backward nested** so that $\mathcal{S}_{i+1} \subseteq \mathcal{S}_i$, then the limit of the **measure of the intersection** has the limiting measure,

$$\mathcal{M}\left(\bigcap_{i=1}^{\infty} \mathcal{S}_i\right) = \lim_{n\to\infty} \mathcal{M}(\mathcal{S}_n),$$

  noting that $\mathcal{M}(\mathcal{S}_i \cap \mathcal{S}_{i+1}) = \mathcal{M}(\mathcal{S}_{i+1})$.

- A **null set** $\mathcal{N} \in \mathbf{\Sigma}$ is a measurable set such that $\mathcal{M}(\mathcal{N}) = 0$, a **negligible set** is a subset of a null set and a measure $\mathcal{M}$ is **complete** if every negligible set is measurable. A σ-algebra $\mathbf{\Sigma}$ can always be completed by adding any missing null sets to it.

- A property P holds **almost everywhere (a. e.)** if the set of elements $\mathcal{S}$ in $\mathbf{\Sigma}$ for which the property does not hold is a null set, i.e., $\mathcal{S} = \mathcal{N}$ is a set with measure zero such that $\mathcal{M}(\mathcal{N}) = 0$.

- Given the measure space $\{\mathcal{U}, \mathbf{\Sigma}, \mathcal{M}_1\}$, another measure $\mathcal{M}_2$ on the measurable space $\{\mathcal{U}, \mathbf{\Sigma}\}$ is **absolutely continuous** with respect to $\mathcal{M}_1$ if for any measurable set $\mathcal{S} \in \mathbf{\Sigma}$

$$\mathcal{M}_1(\mathcal{S}) = 0 \Longrightarrow \mathcal{M}_2(\mathcal{S}) = 0,$$

  Absolute continuity is written symbolically as $\mathcal{M}_2(\mathcal{S}) \prec \mathcal{M}_1(\mathcal{S})$ (or as $\mathcal{M}_2(\mathcal{S}) \ll \mathcal{M}_1(\mathcal{S})$, but this conflicts with asymptotic notation). This property permits defining the ratio $\mathcal{M}_2(\mathcal{S})/\mathcal{M}_1(\mathcal{S})$ for comparison between two measures of a set.

  If $\mathcal{M}_2(\mathcal{S}) \prec \mathcal{M}_1(\mathcal{S})$ and $\mathcal{M}_1(\mathcal{S}) \prec \mathcal{M}_2(\mathcal{S})$, i.e., both are mutually absolutely continuous with respect to the other, then the measures $\mathcal{M}_1$ and $\mathcal{M}_2$ are said to be **equivalent** ($\mathcal{M}_1(\mathcal{S}) \overset{\text{a.c.}}{\equiv} \mathcal{M}_1(\mathcal{S})$). As Cont and Tankov [59] suggest, the term **equivalence** is perhaps misleading and should be called something like **comparable**.

Many of these properties are needed for proofs of existence and convergence, as well as for constructing stochastic processes.

Measurable Functions:

A prerequisite that a function f is integrable is that f is a **measurable function**.

Definition 13.3. *Given two measurable spaces,* $(\mathcal{U}_1, \mathbf{\Sigma}_1)$ *and* $(\mathcal{U}_2, \mathbf{\Sigma}_2)$*, a mapping of the* ***function*** f *from* $\mathcal{U}_1$ *to* $\mathcal{U}_2$ *is* ***measurable*** *with respect to* $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2)$ *if the* ***inverse (preimage)*** $f^{-1}(\mathcal{S}_2) \in \mathbf{\Sigma}_1$ *for all* $\mathcal{S}_2 \in \mathbf{\Sigma}_2$*, i.e., there is a* $\mathcal{S}_1 \in \mathbf{\Sigma}_1$ *such that* $f(\mathcal{S}_1) = \mathcal{S}_2$*.*

Just as in Riemann integration for general Riemann integrable functions, the integral is built up from the limit of finite Riemann sums, the integral with respect to a measurable function is built-up from sums of step functions called a **simple function**.

Definition 13.4.

- *A* ***simple function*** *is a finite linear combination of set indicator functions* $\{\mathbf{1}_{x\in\mathcal{S}_i}\}$ *of measurable sets* $\mathcal{S}_i$ *for* $i = 1\!:\!n$ *on a measurable space* $(\mathcal{U}, \mathbf{\Sigma})$*, with real coefficients (could also be complex)* c_i*, having the form*

$$f(x) = \sum_{i=1}^{n} c_i \mathbf{1}_{x\in\mathcal{S}_i},$$

where $x \in \mathcal{U}$*.*

- *The* ***integral with respect to the measure*** $\boldsymbol{\mathcal{M}}$ *for such a simple function is*

$$\mathcal{I}_{\mathcal{M}}[f] = \sum_{i=1}^{n} c_i \mathcal{M}(\mathcal{S}_i),$$

provided all the measures $\mathcal{M}(\mathcal{S}_i)$ *are finite, i.e., providing the analogy to the Riemann sums.*

- *For a general, positive measurable function* f*,* ***integrability*** *can be extended to* f *by* ***comparison to simple measurable functions*** *on* $\mathcal{U}$*, such as*

$$\mathcal{I}_{\mathcal{M}}[f] = \sup_{g} \left\{ \mathcal{I}_{\mathcal{M}}[g] : \; g(x) = \sum_{i=1}^{n} c_i \mathbf{1}_{x\in\mathcal{S}_i}, \; g(x) \le f(x), \; x \in \mathcal{U} \right\},$$

provided $\mathcal{I}_{\mathcal{M}}[f]$ *is finite. For functions that change sign, i.e., signed functions, the* ***positive-negative decomposition*** $f(x) = f_+(x) - f_-(x)$ *with the* $f_\pm(x) \equiv (|f|(x) \pm f(x))/2$ *for* $x \in \mathcal{U}$*, such that*

$$\mathcal{I}_{\mathcal{M}}[f] = \mathcal{I}_{\mathcal{M}}[f_+] - \mathcal{I}_{\mathcal{M}}[f_-],$$

provided the $\mathcal{I}_{\mathcal{M}}[f_\pm]$ *are finite. (The positive-negative decomposition is used in Chapt. 9 for numerical up-winding to ensure stability.)*

2006/5/24
page 490

- *If $\mathcal{M}$ is a Lebesgue measure, then the Lebesgue of the measure function f on $\mathcal{S} \in \mathcal{U}$ can be written,*

$$\mathcal{I}_{\mathcal{M}}[f] = \int_{\mathcal{S}} f(x)\mathcal{M}(dx) = \int_{\mathcal{U}} \mathbf{1}_{x\in\mathcal{S}}\ f(x)\mathcal{M}(dx),$$

where recall dx symbolizes the set $(x, x+dx)$.

- ***Monotone Convergence Theorem:***
Given the measure space $(\mathcal{U}, \mathbf{\Sigma}, \mathcal{M})$, if $\{f_n(x),\ f_n(x) \geq 0 \text{ for } n = 1, 2, \ldots\}$ is a countable sequence of 1-dimensional (non-negative) measurable functions on $\mathcal{U}$ that is a. e. monotone increasing and converging pointwise to $f(x)$ a. e., then

$$\lim_{n\to\infty} \int_{\mathcal{U}} f_n(x)\mathcal{M}(dx) = \int_{\mathcal{U}} f_n(x)\mathcal{M}(dx).$$

This basic convergence theorem leads to several others.

13.1.2 Change of Measure: Radon-Nikodým Theorem and Derivative:

The abstract analog of the change of variables, chain rule and Jacobian techniques for Riemann or Riemann-Stieltjes integral is the change of measures and the Radon-Nikodým derivative.

Theorem 13.5. ***Radon-Nikodým Change of Measures:***
Given the measure space $\{\mathcal{U}, \mathbf{\Sigma}, \mathcal{M}_1\}$ with σ-finite measure $\mathcal{M}_1$, if $\mathcal{M}_2$ is a finite measure that is absolutely continuous with respect to $\mathcal{M}_1$ ($\mathcal{M}_2 \prec \mathcal{M}_1$) then there exists a measurable real function $\mathbb{D}(x) > 0$ for $x \in \mathcal{U}$ such that for each measurable set $\mathcal{S} \in \mathbf{\Sigma}$

$$\mathcal{M}_2(\mathcal{S}) = \mathcal{I}_{\mathcal{M}_1}[\mathbb{D}\,\mathbf{1}_{*\in\mathcal{S}}] = \int_{\mathcal{U}} \mathbb{D}(x)\mathbf{1}_{x\in\mathcal{S}} d\mathcal{M}_1(x) = \int_{\mathcal{S}} \mathbb{D}(x)d\mathcal{M}_1(x), \quad (13.4)$$

*where $d\mathcal{M}_i(x) = \mathcal{M}_i(dx)$ is equivalent notation for $i = 1\!:\!2$. The function $\mathbb{D}$ is the **Radon-Nikodým derivative** of $\mathcal{M}_2$ with respect to $\mathcal{M}_1$, i.e.,*

$$\mathbb{D}(x) = \frac{d\mathcal{M}_2}{d\mathcal{M}_1}(x) \quad \text{or} \quad \mathbb{D}(\mathcal{S}) = \frac{d\mathcal{M}_2}{d\mathcal{M}_1}(\mathcal{S}). \quad (13.5)$$

Further, if η is integrable with respect to the measure $\mathcal{M}_2$, then

$$\begin{aligned}\mathcal{I}_{\mathcal{M}_2}[\eta] &= \int_{\mathcal{U}} \eta(x)d\mathcal{M}_2(x) = \int_{\mathcal{U}} \eta(x)\frac{d\mathcal{M}_2(x)}{d\mathcal{M}_1(x)}d\mathcal{M}_1(x)\\ &= \mathcal{I}_{\mathcal{M}_2}[\eta\,\mathbb{D}] = \int_{\mathcal{U}} \eta(x)\mathbb{D}(x)\mathcal{M}_1(x),\end{aligned}$$

i.e., using the Radon-Nikodým derivative in a measure-theoretic chain rule.

Thus, the Radon-Nikodým derivative is the analog of the Jacobian of the transformation (10.56) in an integral change of variables and leads to the absolutely continuous measure chain rule, symbolically substituting for $\mathbb{D}$,

$$d\mathcal{M}_2 = \frac{d\mathcal{M}_2}{d\mathcal{M}_1} d\mathcal{M}_1.$$

If $d\mathcal{M}_2$ and $d\mathcal{M}_1$ are mutually absolutely continuous, i.e., equivalent ($\mathcal{M}_1(\mathcal{S}) \overset{\text{a.c.}}{\equiv} \mathcal{M}_1(\mathcal{S})$), the Radon-Nikodým derivatives are mutual reciprocals,

$$\frac{d\mathcal{M}_1}{d\mathcal{M}_2} = 1 \Big/ \frac{d\mathcal{M}_2}{d\mathcal{M}_1}\,,$$

formally justified by common null sets.

See the probability measure Examples 13.13 illustrations of applied-oriented calculations for Radon-Nikodým derivatives in Subsect. 13.2.1.

13.1.3 Probability Measure Basics

Since the probability distribution function for the real random variable $\mathbf{X}$ on the real set $\mathcal{S} \subseteq \mathbb{R}^{n_x}$ has the property that

$$\Phi_{\mathbf{X}}(\mathcal{S}) = \text{Prob}[\mathbf{X} \in \mathcal{S}] \in [0, 1],$$

it is a natural candidate for a measure and the density $\phi_{\mathbf{X}}(\mathbf{x})$ could play the role of the Radon-Nikodým derivative. According to convention, we reset the universe as $\mathcal{U} = \Omega$, the σ-algebra as $\mathbf{\Sigma} = \mathcal{F}$ and the measure as $\mathcal{M} = \mathbb{P}$. For the jump part of jump-diffusions, counting or jump measures will also be needed.

Definition 13.6. ***Probability Measure:***
*A **probability space** $(\Omega, \mathcal{F}, \mathbb{P})$ is a measure space with elements $\omega \in \Omega$ called **sample points** of random events in the **sample space** Ω, elements $\mathcal{F}_i \in \mathcal{F}$ called random **events** and any **probability measure** $\mathbb{P}$ on the measurable space $(\Omega, \mathcal{F})$ has total mass of one, i.e., $\mathbb{P}(\Omega) = 1$.*

Summarizing the Kolmogorov axioms [33] of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$:

- $\mathbb{P}(\emptyset) = 0$ *and* $\mathbb{P}(\Omega) = 1$.
- $\mathbb{P}(\mathcal{S}) \geq 0$ *for all* $\mathcal{S} \in \Omega$.
- $\mathbb{P}(\cup_{i=1}^{\infty} \mathcal{S}_i) = \cup_{i=1}^{\infty} \mathbb{P}(\mathcal{S}_i)$, *assuming the $\{\mathcal{S}_i\}$ are disjoint and countable, i.e., there is **countable additivity**, so that if $\mathcal{S} \cup \mathcal{S}^c = \Omega$, then the complementarity property also holds, $\mathbb{P}(\mathcal{S}^c) = \mathbb{P}(\Omega) + \mathbb{P}(\mathcal{S})$.*
- *If $\mathcal{S}_2 \subseteq \mathcal{S}_2$ and $\mathbb{P}(\mathcal{S}_1) = 0$, then $\mathbb{P}(\mathcal{S}_2) = 0$, i.e., the probability space is **complete**.*

Some additional properties and nomenclature:

DRAFT

- *The $\omega \in \Omega$ are also called scenarios as well as* ***outcomes****, the* ***underlying*** *or* ***background random variables****, e.g., like the mark variable of the compound Poisson process or Poisson random measure.*

- *An event set $\mathcal{S}$ with probability $\mathbb{P}(\mathcal{S}) = 1$ is said to happen* ***almost surely*** *(****a.s.****) or* ***with probability one*** *(****w.p.o.****), equivalent to* ***almost everywhere*** *(****a.e.****) for mathematical measures. If an event $\mathcal{S}$ has probability $\mathbb{P}(\mathcal{S}) = 0$, the event is said to be* ***impossible****.*

- *Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then a (real)* ***random variable*** *$\mathbf{X}(\omega)$ is a measurable mapping from Ω to $\mathbb{R}^{n_x}$ such that the inverse (preimage) $\mathbf{X}^{-1}(\mathcal{S}) = \{\omega \in \Omega : \ \mathbf{X}(\omega) \in \mathcal{S}\}$ is $\mathcal{F}$-measurable for Borel (open) sets $\mathcal{S} \in \mathcal{B}(\mathbb{R}^{n_x})$, i.e., $\mathbf{X}(\omega)$ is the* ***realization*** *$\mathbf{X}$ upon event ω. If f is a (real) measurable function, then $f(\mathbf{X}(\omega))$ will also be a random variable.*

- *If the problem involves only a single probability measure $\mathbb{P}$ for the single random variable ω, then we can write in more usual notation,*

$$X \equiv \omega, \quad \mathrm{Prob}[X \in \mathcal{S}] = \mathrm{Pr}[X \in \mathcal{S}] \equiv \mathbb{P}(\mathcal{S}),$$

  i.e., the probability measure is the distribution $\Phi_\omega(\mathcal{S}) = \mathbb{P}(\mathcal{S})$ for $\mathcal{S} \subseteq \Omega$.

- *In general, if $\mathbf{X} = \mathbf{X}(\omega) \in \mathbb{R}^{n_x}$ for $\omega \in \Omega$, then let $\omega \in \mathcal{S}_\omega \subseteq \Omega$, $\mathbf{X}(\omega) \in \mathcal{S}_{\mathbf{X}} = \mathbf{X}(\mathcal{S}_\omega)$ and assuming the preimage $\mathcal{S}_\omega = \mathbf{X}^{-1}(\mathcal{S}_{\mathbf{X}})$ exists, then the* ***distribution*** *of $\mathbf{X}$ is the probability measure*

$$\Phi_{\mathbf{X}}(\mathcal{S}_{\mathbf{X}}) = \mathbb{P}(\mathbf{X}^{-1}(\mathcal{S}_{\mathbf{X}})),$$

  so $\Phi_{\mathbf{X}}(\mathbf{x}) = \mathbb{P}(\{\omega \ni \mathbf{X} \le \mathbf{x}\})$, the inequality $(\mathbf{X} \le \mathbf{x})$ meant element-wise.

- *The* ***expectation*** *for a measurable real function f of $X \in \mathbb{R}^{n_x}$ with $\omega \in \Omega$ is then*

$$\mathrm{E}[f(\mathbf{X})] = \int_\Omega f(\mathbf{X}(\omega))\mathbb{P}(d\omega) = \int_\Omega f(\mathbf{X}(\omega))d\mathbb{P}(\omega) = \int_{\mathbb{R}^{n_x}} f(\mathbf{x})\Phi_{\mathbf{X}}(d\mathbf{x}),$$

  provided f is absolutely integrable,

$$\int_\Omega |f(\mathbf{X}(\omega))|\mathbb{P}(d\omega) < \infty,$$

  noting that the $d\omega$ argument of $\mathbb{P}$ is an abbreviation for the interval $(\omega, \omega + d\omega)$ and that $d\mathbb{P}(\omega)$ and $\mathbb{P}(d\omega)$ will be assumed to be equivalent notation.

- ***Almost Sure Equivalence:*** *Let $\mathbf{X}_1(\omega)$ and $\mathbf{X}_2(\omega)$ be two random variables for $\omega \in \Omega$, then $\mathbf{X_1} \overset{\text{a.s.}}{=} \mathbf{X_2}$ if*

$$\mathbb{P}(\{\omega \in \Omega, \mathbf{X}_1(\omega) = \mathbf{X}_2(\omega)\}) = 1.$$

- ***Equivalence in Distribution:*** *Let* $\mathbf{X}_1(\omega)$ *and* $\mathbf{X}_2(\omega)$ *be two random variables for* $\omega \in \Omega$. *If the distribution satisfy*

$$\Phi_{\mathbf{X}_1} = \Phi_{\mathbf{X}_2},$$

then $\mathbf{X}_1(\omega)$ *and* $\mathbf{X}_2(\omega)$ *are called* ***equal in distribution*** *and we write*

$$\mathbf{X_1} \stackrel{\text{dist}}{=} \mathbf{X_2}.$$

(*Also called* ***equal in law*** *or* ***identically distributed***; *the notation* $\mathbf{X_1} \stackrel{\text{d}}{=} \mathbf{X_2}$ *is also used.*)

- *The set of* n *random variables* $\{X_i\}$ *are* ***independent*** *with respect to the measurable sets* $\mathcal{S}_i$ *for* $i = 1\!:\!n$ *if the probability of the union is the product of the probabilities,*

$$\mathbb{P}\left(\bigcup_{i=1}^{n}\{X_i \in \mathcal{S}_i\}\right) = \prod_{i=1}^{n}\mathbb{P}(\{X_i \in \mathcal{S}_i\}),$$

where the underlying random variable ω *has been suppressed. A more concrete and useful form as distribution in the vector* $\mathbf{X} = [X_i]_{n\times 1}$ *is*

$$\Phi_{\mathbf{X}}(\mathbf{x}) = \mathbb{P}\left(\bigcup_{i=1}^{n}\{X_i \le x_i\}\right) = \prod_{i=1}^{n}\mathbb{P}(\{X_i \le x_i\}) = \prod_{i=1}^{n}\Phi_{X_i}(x_i).$$

An immediate corollary is the multiplication rule for the expectation of a set of independent random variables,

$$\mathrm{E}\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n}\mathrm{E}[X_i],$$

assuming finite expectations, $\mathrm{E}[|X_i|] < \infty$ *for* $i = 1\!:\!n$.

For more background information, see Applebaum [12], Billingsley [32], Bingham and Kiesel [33], Cont and Tankov [59]. Cyganowski, Kloeden and Ombach [66], Øksendal [218] and Øksendal and Sulem [219].

Much of the further results, such as conditional expectations, follow the applied path in this book, except that matters like that of positivity and changes in sign have to be treated with care to account for particular abstract constructs and conditions that are designed to facilitate proofs rather than the wide variety of problem applications.

13.1.4 Stochastic Processes in Continuous Time on Filtered Probability Spaces

Since the emphasis of this book is on jump-diffusions, stochastic processes in continuous time are treated and the relatively simpler, but not simple, discrete time stochastic processes are omitted (see Pliska's [221] book or Bingham and Kiesel's [33, Chapt. 3] chapter devoted to discrete time processes). The main additional difficulty treating stochastic processes in continuous time is extending the notion of a single probability space to a family of probability spaces over the continuous time variable t which often has infinite range.

Definition 13.7. ***Filtered Probability Space:***

- *Based upon a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, *a* ***filtration*** *is a family of increasing* σ*-algebras*
$$\mathbb{F} = \{\mathcal{F}_t : \ t \geq 0; \ \mathcal{F}_s \subseteq \mathcal{F}_t, \ 0 \leq s \leq t < \infty\}$$
and the extended space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ *is called a* ***filtered probability space****. The sub-*σ*-algebra* $\mathcal{F}_t$ *represents the known information of the system on* $(0, t]$ *at time* t.

- *The usual filtration conditions (with jump-diffusions in mind) are*
  - *The initial sub-*σ*-algebra* $\mathcal{F}_0$ *the* $\mathbb{P}$*-null-sets of* $\mathcal{F}$.
  - *The filtration* $\mathbb{F}$ *is* ***right-continuous with left limits (RCLL or cádlág in French)****, i.e.,* $\mathcal{F}_t = \mathcal{F}_{t^+} = \lim_{\epsilon \to 0^+} \mathcal{F}_{t+\epsilon}$ *for the RC part and* $\mathcal{F}_{t^-} = \lim_{\epsilon \to 0^+} \mathcal{F}_{t-\epsilon}$ *for the LL part exists. The* ***jump*** *in the sub-*σ*-algebra at time* t *is* $[\mathcal{F}]_t = \mathcal{F}_{t^+} - \mathcal{F}_{t^-}$. *If only continuous processes such as diffusions are under consideration, then* ***right continuity (RC)*** *is sufficient.*

Definition 13.8. ***Stochastic Process:***

- *Given the filtered probability space* $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, *a* ***stochastic process in continuous time*** $\mathbf{X} = \{\mathbf{X}(t) : \ t \geq 0\}$ *and* X *is* $\mathcal{F}_t$***-adapted*** *to the filtration* $\mathbb{F}$ *if* $\mathbf{X}(t)$ *is* $\mathcal{F}_t$*-measurable* $(\mathbf{X}(t) \in \mathcal{F}_t)$ *for each* t.

- *The* ***natural filtration*** *for the stochastic process* $\mathbf{X}(t)$ *can be written as*
$$\mathcal{F}_{t,\mathbf{X}} = \widehat{\sigma}(\mathbf{X}(s), \ 0 \leq s \leq t),$$
with $\widehat{\sigma}$ *signifying the* σ*-field of* $\mathbf{X}(t)$, *or more loosely the information or history of the process* $\mathbf{X}(t)$ *up until time* t.

- *Including the dependence on the underlying random variable,* $\omega \in \Omega$, *the* $\mathbf{X}(t; \omega)$ *defines a random function of time, called the* ***sample path*** *and is a mapping from* $[0, t] \times \Omega$ *to* $\mathbb{R}^{n_x}$. *Usually,* $\mathbf{X}(t; \omega)$ *is denoted by* $\mathbf{X}_t(\omega)$ *or just* X_t, *however in this book real subscripts are reserved to denote partial derivatives, except for algebraic quantities like* $\mathcal{F}_t$ *that are not genuine functions.*

- *If* X *is adapted, i.e.,* $\mathcal{F}_t$***-adapted*** *to* $\mathbb{F}$, *for* $t \geq 0$, *then the conditional expectation satisfies*
$$\mathrm{E}[X(t) \mid \mathcal{F}_t] \stackrel{\text{a.s.}}{=} X(t),$$

DRAFT

since $X(t)$ is known from $\mathcal{F}_t$ (recall the symbol $\stackrel{\text{a.s.}}{=}$ denotes ***equals almost surely****). Saying that X or $X(t)$ is $\mathcal{F}_t$-****adapted*** *to $\mathbb{F}$ means the same as saying that $X(t)$ is* ***nonanticipating****.*

- *Two stochastic processes X_1 and X_2 are the* ***same with respect to a set of finite-dimensional distributions*** *if for some positive integer n and discrete time points $\{t_i : \ i = 1:n\}$, the random vectors $\mathbf{X}_j = [X_{i,j}]_{n\times 1}$ for $j = 1:2$ have the same n-dimensional distribution, corresponding to the stochastic processes X_j for $j = 1:2$, respectively.*

13.1.5 Martingales in Continuous Time

Martingales are processes with the property that the best predictor of the process future value is the present value given present knowledge, i.e., it represents a fair game of gambling, rather than a favorable or unfavorable one.

Definition 13.9. ***MartingaleProperties in Continuous Time:***

- *Given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ and $\mathcal{F}_t$-adapted process $X(t)$ on $[0, T]$, $T < \infty$, then $X(t)$ is a* ***martingale*** *if*

$$\mathrm{E}[X(t) \mid \mathcal{F}_s] \stackrel{\text{a.s.}}{=} X(s), \quad t > s \ge 0, \tag{13.6}$$

provided $X(t)$ is absolutely integrable, $\mathrm{E}[|X(t)|] < \infty$ on $[0, T]$, i.e., the best predictor of $X(t)$ with respect to the filter $\mathcal{F}_s$ is $X(s)$.

- *If instead of (13.6),*

$$\mathrm{E}[X(t) \mid \mathcal{F}_s] \stackrel{\text{a.s.}}{\le} X(s), \quad t > s \ge 0,$$

then $X(t)$ is a ***supermartingale****,*
but if

$$\mathrm{E}[X(t) \mid \mathcal{F}_s] \stackrel{\text{a.s.}}{\ge} X(s), \quad t > s \ge 0,$$

then $X(t)$ is a ***submartingale****. (The submartingale corresponds to the favorable game and the supermartingale corresponds to the unfavorable game, provided $X(t) - X(s)$ represents the gain.)*

- *Two martingales $\mathcal{M}_1(t)$ and $\mathcal{M}_2(t)$ which are also equivalent or mutually absolutely continuous measures, i.e., $\mathcal{M}_1(t) \stackrel{\text{a.c.}}{\equiv} \mathcal{M}_2(t)$, are called* ***equivalent martingale measures (EMM)*** *and they play an important role in mathematical finance.*

Examples 13.10. ***Diffusion, Jump and other Martingales:***
For this set of examples, the time interval $[0, T]$ as well as the coefficients will be finite, so there is no question that the stochastic processes will be absolutely integrable.

1. *Let $X(t)$ be a $\{\mu_0, \sigma_0\}$-constant coefficient,* ***diffusion process*** *with SDE,*

$$dX(t) = \mu_0 dt + \sigma_0 dW(t),$$

and it is of interest to know for what values of μ_0 is $X(t)$ a martingale, a supermartingale or a submartingale.

The solution by integrating over $[s, t]$ is

$$X(t) = X(s) + \mu_0(t - s) + \sigma_0(W(t) - W(s)),$$

noting that $(W(t) - W(s)) \stackrel{\text{dist}}{=} W(t-s)$ by stationary property and is independent of $W(s)$ and $\mathrm{E}[W(t-s)|W(s)] = 0$, so with $\mathcal{F}_t = \widehat{\sigma}(X(r),\ 0 \le r \le t)$, the natural filtration for $X(t)$,

$$\mathrm{E}[X(t) \mid \mathcal{F}_s] = X(s) + \mu_0(t - s),$$

$0 \le s < t$. Hence, $X(t)$ is a martingale if $\mu_0 = 0$ (the case of the zero infinitesimal mean diffusion, $d\widehat{X}(t) = \sigma_0 dW(t)$), a supermartingale if $\mu_0 < 0$ or a submartingale if $\mu_0 > 0$. Alternatively, the translated process

$$\widetilde{X}(t) = X(t) - \mu_0 t \stackrel{\text{dist}}{=} \widetilde{X}(s) + \sigma_0 W(t - s)$$

is a martingale.

2. *Let $X(t)$ be a $\{\mu_0, \sigma_0\}$-constant coefficient,* ***geometric diffusion process****,*

$$dX(t) = X(t)(\mu_0 dt + \sigma_0 dW(t)),$$

which has the Itô calculus solution,

$$X(t) \stackrel{\text{dist}}{=} X(s)\exp((\mu_0 - \sigma_0^2/2)(t - s) + \sigma_0 W(t - s)),$$

so

$$\mathrm{E}[X(t) \mid \mathcal{F}_s] = X(s)\exp(\mu_0(t - s)),$$

$0 \le s < t$. Again, $X(t)$ is a martingale if $\mu_0 = 0$, a supermartingale if $\mu_0 < 0$ or a submartingale if $\mu_0 > 0$. Alternatively, the scaled process

$$\widetilde{X}(t) = \exp(-\mu_0 t)X(t) \stackrel{\text{dist}}{=} \widetilde{X}(s)\exp(\sigma_0(W(t - s) - \sigma_0(t - s)/2)$$

is martingale, or more specifically an ***exponential martingale*** *[22] and the scaling corresponds to the Girsanov transformation of $W(t)$ that will be discussed in Subsect. 13.2.2.*

3. *Let $X(t)$ be a* ***simple Poisson process*** *$P(t)$ with additional drift and constant coefficients, $\{\mu_0, \nu_0, \lambda_0\}$,*

$$dX(t) = \mu_0 dt + \nu_0 dP(t),$$

where $\mathrm{E}[dP(t)] = \lambda_0 dt = \mathrm{Var}[dP(t)]$*. The solution is*

$$X(t) = X(s) + \mu_0(t-s) + \nu_0(P(t) - P(s)),$$

where $(P(t) - P(s)) \stackrel{\text{dist}}{=} P(t-s)$ *and the conditional expectation is*

$$\mathrm{E}[X(t) \mid \mathcal{F}_s] = X(s) + (\mu_0 + \lambda_0\nu_0)(t-s),$$

so $X(t)$ *is a martingale if* $\mu_0 = -\lambda_0\nu_0$ *(the zero infinitesimal mean Poisson,* $d\widehat{X}(t) = \nu_0 d\widehat{P}(t)$*, where* $d\widehat{P}(t) = dP(t) - \lambda_0 dt$*), a supermartingale if* $\mu_0 < -\lambda_0\nu_0$ *or a submartingale if* $\mu_0 > -\lambda_0\nu_0$*. Alternatively, the translated process*

$$\widetilde{X}(t) = X(t) - (\mu_0 + \lambda_0\nu_0)t \stackrel{\text{dist}}{=} \widetilde{X}(s) + \nu_0(P(t-s) - \lambda_0(t-s))$$

is a martingale.

4. *Let* $X(t)$ *be a* ***compound Poisson process*** *with additional drift and constant coefficients,* $\{\mu_0, \nu_0, \lambda_0, \mu_Q\}$*,*

$$dX(t) = \mu_0 dt + \nu_0 \sum_{i=1}^{dP(t)} Q_i,$$

where $\mathrm{E}[dP(t)] = \lambda_0 dt = \mathrm{Var}[dP(t)]$ *and the* Q_i *are IID random marks with mean* μ_Q *and variance* σ_Q^2 *which will not be needed. The solution is*

$$X(t) \stackrel{\text{dist}}{=} X(s) + \mu_0(t-s) + \nu_0 \sum_{i=1}^{P(t-s)} Q_i$$

and the conditional expectation, by iterated conditional expectations between the Poisson counting process and the marks, is

$$\mathrm{E}[X(t) \mid \mathcal{F}_s] = X(s) + (\mu_0 + \lambda_0\nu_0\mu_Q)(t-s),$$

so $X(t)$ *is a martingale if* $\mu_0 = -\lambda_0\nu_0\mu_Q$ *(the zero infinitesimal mean compound Poisson,* $d\widetilde{X}(t) = \nu_0\mu_Q d\widetilde{P}(t) + \nu_0 \sum_{i=1}^{dP(t)} \widetilde{Q}_i$ *where* $\widetilde{P}(t) \equiv (P(t) - \lambda_0 t)$ *and* $\widetilde{Q}_i \equiv (Q_i - \mu_Q)$*), a supermartingale if* $\mu_0 < -\lambda_0\nu_0\mu_Q$ *or a submartingale if* $\mu_0 > -\lambda_0\nu_0\mu_Q$*. The alternative process*

$$\widetilde{X}(t) = X(t) - (\mu_0 + \lambda_0\nu_0\mu_Q)t = \widetilde{X}(s) + \nu_0\mu_Q\widetilde{P}(t-s) + \nu_0 \sum_{i=1}^{P(t)} \widetilde{Q}_i$$

is a martingale, such that the difference $\widetilde{X}(t) - \widetilde{X}(s)$ *is a linear combination of zero-mean random processes or variables.*

2006/5/24
page 498

5. *As an exercise the reader can find the similar martingale properties as a function of the additional drift for the* ***geometric jump diffusion*** *problem with constant coefficients,*

$$dX(t) = X(t)\left(\mu_0 dt + \sigma_0 dW(t) + \nu_0 \sum_{i=1}^{dP(t)} (\exp(Q_i) - 1)\right),$$

where again the marks are IID with mean μ_Q and variance σ_Q^2, with the amplitude in the log-ready exponential form.

6. *The simplest, but trivial, example is the* ***constant process*** *$X(t) = c_0$ for $t \geq 0$, i.e., $dX(t) = 0$, so $X(t)$ is a martingale since $\mathrm{E}[X(t)|\mathcal{F}_s] = c_0 = X(s)$ for $s < t$.*

7. *Another example is the* ***closed martingale*** *that is constructed from an absolutely integrable random variable Y, independent of t, on the filtered probability space, such that a stochastic process is defined as*

$$X(t) \equiv \mathrm{E}[Y \mid \mathcal{F}_t], \quad t \geq 0.$$

Thus, by the ***tower law*** *([22, p. 34], [205, Rule 6, p. 72]),*

$$\mathrm{E}[X(t) \mid \mathcal{F}_s] = \mathrm{E}[\mathrm{E}[Y \mid \mathcal{F}_t] \mid \mathcal{F}_s] = \mathrm{E}[Y \mid \mathcal{F}_s] \equiv X(s),$$

for $s < t$, since the conditioning on $\mathcal{F}_t$ followed by the conditioning on $\mathcal{F}_s$ is the same as the original conditioning on $\mathcal{F}_s$, i.e., dependence is on the smaller of the conditioning filters.

13.1.6 Jump-Diffusion Martingale Representation:

For hedging in financial applications, martingale representations are heavily relied upon. There are many versions of martingale representation in the literature. Some have useful and elementary presentations. Many are restricted to diffusions except for a mention of jump processes. A selected sample is given by the references: Baxter and Rennie [22], Duffie [74], Glasserman [96], Øksendal [218] and Steele [251]. Here, a form of the martingale repesentation theorem is given for marked-jump-diffusion processes following Applebaum [12] and, particularly, Runggaldier [235]. Their formulation, after Jacod and Shiryaev [149], and Kunita and Watanabe [168], respectively, uses Poisson random measure $\mathcal{P}(\mathbf{dt}, \mathbf{dq})$ defined beginning in (5.1) on page 206 and whose integrals are related to compound Poisson processes (5.3) on mark-sample-space $\mathcal{Q}$ by

$$\int_{\mathcal{Q}} h(t, q)\mathcal{P}(\mathbf{dt}, \mathbf{dq}) = \sum_{i=1}^{dP(t)} h(T_i^-, Q_i), \tag{13.7}$$

2006/5/24
page 499

where the T_i^- are the pre-jump-times and the Q_i are the IID sampled marks, but often found in martingale form by using the centered or mean-zero Poisson random measure,

$$\widetilde{\mathcal{P}}(\mathbf{dt},\mathbf{dq}) \equiv \mathcal{P}(\mathbf{dt},\mathbf{dq}) - \mathrm{E}[\mathcal{P}(\mathbf{dt},\mathbf{dq})] = \mathcal{P}(\mathbf{dt},\mathbf{dq}) - \phi_Q(q;t)dq\lambda(t)dt,$$

where $\Phi_Q(dq;t) = \phi_Q(q;t)dq$ is jump-amplitude probability measure and $\lambda(t)$ is the Poisson jump-rate. The mean-zero relationship corresponding to the original relationship (13.7) is then

$$\int_{\mathcal{Q}} h(t,q)\widetilde{\mathcal{P}}(\mathbf{dt},\mathbf{dq}) = \sum_{i=1}^{dP(t)} h(T_i^-,Q_i) - \mathrm{E}_Q[h(t,Q)]\lambda(t)dt, \tag{13.8}$$

where $\mathrm{E}_Q[h(t,Q)] = \int_{\mathcal{Q}} h(t,q)\phi_Q(q;t)dq$.

Theorem 13.11. ***Marked-Jump-Diffusion Martingale Representation Theorem:***
Given the Wiener process $W(t)$ and compound Poisson triplet

$$\{dP(t),\lambda(t),\phi_Q(q;t)\}$$

or else a Poisson random measure $\mathcal{P}(\mathbf{dt},\mathbf{dq})$ on the sigma-field

$$\mathbb{F} = \mathcal{F}_t^{(W,P,Q)} = \sigma\{W(s),P(t),\mathcal{S}_Q,\ \mathcal{S}_N : 0\le s\le t,\ \mathcal{S}_Q\in\mathcal{Q},\ \mathcal{S}_N\in\mathcal{N}_1\},$$

$\mathcal{N}_1$ is the collection of null-sets of $\mathbb{P}$. Then, any $(\mathbb{P},\mathbb{F})$-martingale $\mathcal{M}(t)$ has the representions

$$\begin{aligned}\mathcal{M}(t) = {} & \mathcal{M}(0) + \textstyle\int_0^t \Gamma^{(\mathrm{D})}(s)dW(s) + \int_0^t\int_{\mathcal{Q}} \Gamma^{(\mathrm{MJ})}(s,q)\widetilde{\mathcal{P}}(\mathbf{ds},\mathbf{dq}) \\ = {} & \mathcal{M}(0) + \textstyle\int_0^t \Gamma^{(\mathrm{D})}(t)dW(s) + \sum_{i=1}^{dP(t)} \Gamma^{(\mathrm{MJ})}\left(T_i^-,Q_i\right) \\ & -\mathrm{E}_Q\left[\Gamma^{(\mathrm{MJ})}(t,Q)\right]\lambda(t)dt,\end{aligned} \tag{13.9}$$

where $\Gamma^{(\mathrm{D})}(t)$ is a predictable (measurable with respect to $\mathbb{P}$), square-integrable process, while $\Gamma^{(\mathrm{MJ})}(t,q)$ is a $\mathcal{F}_t^{(W,P,Q)}$-predictable, $\mathcal{Q}$-marked process, such that

$$\mathrm{E}_Q\left[\Gamma^{(\mathrm{MJ})}(t,Q)\right] = \int_{\mathcal{Q}} \Gamma^{(\mathrm{MJ})}(t,q)\phi_Q(q;t)dq < \infty.$$

The martingale representation theorem is used in the following Subsect. 13.2.2 for two versions of Girsanov's stochastic process transformation theorem, one for the diffusion process alone, i.e., without the Poisson terms in (13.9), and another for marked-jump-diffusion processes using the full form in (13.9).

The martingale approach may be a favored approach to solving SDE problems, but Heath and Schweizer [132] show the equivalence of the martingale and PDE approaches for a number of financial applications. The Feynmann-Kac formula (see (8.71) in the item on p. 329 here or the appendix of Duffie [74, Appendix E.] for more background) is used to solve the corrresponding PDE problem that is derived from the SDE.

DRAFT

13.2 Change in Probability Measure: Radon-Nikodým Derivatives and Girsanov's Theorem

13.2.1 Radon-Nikodým Theorem and Derivative for Change of Probability Measure:

Here, a version of the Radon-Nikodým Theorem 13.5 and derivative is formulated especially for probability measures and expectations. The abstract analog of the change of distribution corresponding to a change in random variables presented in Eq. (0.5) for the distribution and (0.6) for the density on p. 4 in preliminaries Chapt. 0.

Theorem 13.12. ***Radon-Nikodým Change of Probability Measures:*** *Given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ with σ-finite measure $\mathbb{P}_1$, if $\mathbb{P}_2$ is a finite measure that is mutually absolutely continuous with $\mathbb{P}_1$ (equivalent, $\mathbb{P}_2 \overset{\text{a.c.}}{\equiv} \mathbb{P}_1$) then there exists a positive measurable real function*

$$\mathbb{D}(x) = \frac{d\mathbb{P}_2}{d\mathbb{P}_1}(x) \quad \text{or} \quad \mathbb{D}(\mathcal{S}) = \frac{d\mathbb{P}_2}{d\mathbb{P}_1}(\mathcal{S}). \tag{13.10}$$

*called the **Radon-Nikodým derivative** of $\mathbb{P}_2$ with respect to $\mathbb{P}_1$, for $x \in \Omega$ such that for each measurable set $\mathcal{S} \in \mathcal{F}$*

$$\mathbb{P}_2(\mathcal{S}) = \mathrm{E}_{\mathbb{P}_1}[\mathbb{D}(X)\mathbf{1}_{X\in\mathcal{S}}] = \int_\Omega \mathbb{D}(x)\mathbf{1}_{x\in\mathcal{S}} d\mathbb{P}_1(x) = \int_{\mathcal{S}} \mathbb{D}(x) d\mathbb{P}_1(x), \tag{13.11}$$

where $d\mathbb{P}_i(x) = \mathbb{P}_i(dx)$ is equivalent notation for $i = 1{:}2$.

Further, if η is absolutely integrable with respect to the measure $\mathbb{P}_2$, then

$$\begin{aligned}\mathrm{E}_{\mathbb{P}_2}[\eta(X)] = \int_\Omega \eta(x) d\mathbb{P}_2(x) &= \int_\Omega \eta(x)\frac{d\mathbb{P}_2(x)}{d\mathbb{P}_1(x)} d\mathbb{P}_1(x) \\ &= \mathrm{E}_{\mathbb{P}_2}[\eta(X)\mathbb{D}(X)] = \int_\Omega \eta(x)\mathbb{D}(x)\mathbb{P}_1(x),\end{aligned}$$

i.e., using the Radon-Nikodým derivative in a measure-theoretic chain rule.

Thus, the Radon-Nikodým derivative is the analog of the Jacobian of the transformation (10.56) in an integral change of variables and leads to the absolutely continuous measure chain rule, symbolically substituting for g,

$$d\mathbb{P}_2 = \frac{d\mathbb{P}_2}{d\mathbb{P}_1} d\mathbb{P}_1.$$

If $d\mathbb{P}_2$ and $d\mathbb{P}_1$ are mutually absolutely continuous, i.e., equivalent ($\mathbb{P}_1(\mathcal{S}) \overset{\text{a.c.}}{\equiv} \mathbb{P}_1(\mathcal{S})$), the Radon-Nikodým derivatives are mutual reciprocals,

$$\frac{d\mathbb{P}_1}{d\mathbb{P}_2} = 1 \Big/ \frac{d\mathbb{P}_2}{d\mathbb{P}_1} \,,$$

DRAFT

formally justified by common null sets.

Examples 13.13. ***Radon-Nikodým Derivative Calculations:***

- ***Normal distributions***:
  Suppose a transformation from a standard normal distribution with density

$$\phi_1(x) = \exp\left(-x^2/2\right)\Big/\sqrt{2\pi}$$

to a mean-μ, variance-σ^2 normal distribution with density,

$$\phi_2(x) = \exp\left(-(x-\mu)^2/\left(2\sigma^2\right)\right)\Big/\sqrt{2\pi\sigma^2}\,.$$

The change in measure coincides with a change of drift and a change of scale. Thus, $\mathbb{P}_1(x) = \Phi_1(x)$ is the first probability measure and the second is

$$\mathbb{P}_2(x) = \Phi_2(x) = \int_{-\infty}^{\infty} \mathbb{D}(y)\phi_1(y)dy = \int_{-\infty}^{\infty} \mathbb{D}(y)\phi_1(y)dy,$$

or $\phi_2(x) = \mathbb{D}(x)\phi_1(x)$ upon differentiating according to the fundamental theorem of integral calculus and the Radon-Nikodým derivative is

$$\mathbb{D}(x) = \frac{d\mathbb{P}_2(x)}{d\mathbb{P}_1(x)} = \frac{d\Phi_2(x)}{d\Phi_1(x)} = \frac{\phi_2(x)}{\phi_1(x)} = \frac{\exp\left(-(x-\mu)^2/\left(2\sigma^2\right)\right)\Big/\sqrt{2\pi\sigma^2}}{\exp\left(-x^2/2\right)\big/\sqrt{2\pi}}$$
$$= \frac{1}{\sigma}\exp\left(-\frac{\left(1-\sigma^2\right)x^2 - 2\mu x + \mu^2}{2\sigma^2}\right). \tag{13.12}$$

Hence, under measure $\mathbb{P}_1$ the random variable X has mean 0 and variance 1, but under measure $\mathbb{P}_2$ the random variable X has mean μ and variance σ^2. If $\sigma = 1$, then there is only a change of drift and the Radon-Nikodým derivative is simpler:

$$\mathbb{D}(x) = \frac{d\mathbb{P}_2(x)}{d\mathbb{P}_1(x)} = \exp\left(\frac{2\mu x - \mu^2}{2}\right).$$

The more general form (13.12), formally justified here, can be transformed to the form in a proposition of Cont and Tankov [59, p. 306, Prop. 9.7] for two diffusion or Brownian motion processes, both denoted by $X = X(T)$, with parameters $\mu \to \mu_j T$ for the drifts and $\sigma^2 \to \sigma_1^2 T = \sigma^2 T = \sigma_2^2 T$ for a common variance on $(\Omega, \mathcal{F}_T, \mathbb{P}_j, \mathbb{F})$ for $j = 1:2$. Hence, using the fact the

Radon-Nikodým derivative is the ratio of the two densities,

$$\begin{aligned}\mathbb{D}(X(T),T) &= \frac{d\mathbb{P}_2(X(T),T)}{d\mathbb{P}_1(X(T),T)} \\ &= \frac{\exp\left(-(X(T)-\mu_2 T)^2/\left(2\sigma_2^2 T\right)\right)\Big/\sqrt{2\pi\sigma_2^2 T}}{\exp\left(-(X(T)-\mu_1 T)^2/\left(2\sigma_1^2 T\right)\right)\Big/\sqrt{2\pi\sigma_1^2 T}} \\ &= \exp\left(\frac{2(\mu_2-\mu_1)X(T)-(\mu_2^2-\mu_1^2)T}{2\sigma^2}\right). \end{aligned} \tag{13.13}$$

This corrects an error in [59, p. 306, Prop. 9.7]. They also convert this to the Cameron-Martin theorem form, by letting $X(T) = \mu_1 T + \sigma W_1(T)$, in the notation here, so

$$\mathbb{D}(T) = \frac{d\mathbb{P}_2(T)}{d\mathbb{P}_1(T)} = \exp\left(\frac{2(\mu_2-\mu_1)\sigma W_1(T)-(\mu_2-\mu_1)^2 T}{2\sigma^2}\right), \tag{13.14}$$

which is correct in [59, p. 306, following Prop. 9.7].

- ***Sets of Independent Random Variables:***
Let $\mathbf{X} = [X_i]_{n\times 1}$ be a set on n independent random variables with vector mean $\boldsymbol{\mu}^{(1)} = \left[\mu_i^{(1)}\right]_{n\times 1}$ and variance vector $\mathbf{V}^{(1)} = \left[\sigma_i^{(1)}\right]_{n\times 1}$, with product density

$$\phi^{(1)}(\mathbf{x}) = \prod_{i=1}^{n} \phi_i^{(1)}(x_i),$$

due to the independence property. The relationship between the measure, the distribution

$$\Phi^{(1)}(\mathbf{x}) = \mathrm{Prob}_{\mathbb{P}_1}[\mathbf{X} \le \mathbf{x}]$$

and the density can be written formally as

$$\frac{d\mathbb{P}_1(\mathbf{x})}{d\mathbf{x}} = \left(\prod_{i=1}^{n}\frac{\partial}{\partial x_i}\right)\Phi^{(1)}(\mathbf{x}) = \phi^{(1)}(\mathbf{x}),$$

where $\mathbf{X} \le \mathbf{x}$ means $X_i \le x_i$ for $i = 1\!:\!n$ and $d\mathbf{x} = \prod_{i=1}^{n} dx_i$ is the infinitesimal n-dimensional Euclidean measure, not a vector differential.

Let there be a function $\mathbb{D}(\mathbf{x})$ that generates a second distribution or measure,

$$\begin{aligned}\Phi^{(2)}(\mathbf{x}) = \mathrm{Prob}_{\mathbb{P}_2}[\mathbf{X}\le\mathbf{x}] &= \left(\prod_{i=1}^{n}\int_{-\infty}^{x_i} dy_i \phi_i^{(2)}(y_i)\right) \\ &= \left(\prod_{i=1}^{n}\int_{-\infty}^{x_i} dy_i \phi_i^{(1)}(y_i)\right)\mathbb{D}(\mathbf{y}),\end{aligned}$$

DRAFT

so

$$\frac{d\mathbb{P}_2(\mathbf{x})}{d\mathbf{x}} = \left(\prod_{i=1}^{n} \frac{\partial}{\partial x_i}\right) \Phi^{(2)}(\mathbf{x}) = \prod_{i=1}^{n} \int_{-\infty}^{x_i} dy_i \phi_i^{(2)}(x_i) = \phi_i^{(2)}(\mathbf{x})$$

$$= \mathbb{D}(\mathbf{x}) \prod_{i=1}^{n} \int_{-\infty}^{x_i} dy_i \phi_i^{(1)}(x_i) = \mathbb{D}(\mathbf{x})\phi^{(1)}(\mathbf{x}).$$

Solving produces

$$\mathbb{D}(\mathbf{x}) = \frac{d\mathbb{P}_2(\mathbf{x})}{d\mathbb{P}_1(\mathbf{x})} = \frac{\phi_i^{(2)}(\mathbf{x})}{\phi^{(1)}(\mathbf{x})} = \prod_{i=1}^{n} \frac{\phi_i^{(2)}(x_i)}{\phi_i^{(1)}(x_i)}. \tag{13.15}$$

This result is important for stochastic processes $X(t)$ for $t \in [0, T]$, since a Radon-Nikodým derivative cannot be computed for a random variable over an infinite-dimensional interval, but it is possible to sample $X(t)$ at sample times $t_i = (i-1)T/n$ using $X_i = X(t_i)$ for $i = 1{:}n$, assuming the process of interest has independent increments.

As a more concrete example, suppose that the X_i have a standard normal distribution, i.e., IID with $\mu_i^{(1)} = 0$ and $\left(\sigma_i^{(1)}\right)^2 = 1$, and a nonstandard distribution is sought with mean $\mu_i^{(2)} = \mu_i$ and $\left(\sigma_i^{(2)}\right)^2 = \sigma_i^2$, then using (13.12),

$$\mathbb{D}(\mathbf{x}) = \frac{d\mathbb{P}_2(\mathbf{x})}{d\mathbb{P}_1(\mathbf{x})} = \frac{1}{\prod_{j=1}^{n} \sigma_j} \exp\left(-\sum_{i=1}^{n}\left(\frac{(1-\sigma_i^2)x_i^2 - 2\mu_i x_i + \mu_i^2}{2\sigma_i^2}\right)\right). \tag{13.16}$$

This example is similar to one in Glasserman [96], except there the $\sigma_i \equiv 1$.

- ***Poisson Distribution, a Discrete Analogy:***
  Next consider a Poisson cumulative distribution with parameter Λ_1 for the discrete variable N_1,

$$\Phi_n^{(1)} = \mathrm{Prob}[N_1 < n] = e^{-\Lambda_1} \sum_{k=0}^{n} \frac{\Lambda_1^n}{n!}$$

which has increment (discrete derivative analog)

$$\Delta\Phi_{n-1}^{(1)} \equiv \Phi_n^{(1)} - \Phi_{n-1}^{(1)} = e^{-\Lambda_1} \frac{\Lambda_1^n}{n!},$$

the numerical forward difference notation, corresponding to a discrete density and consistent with Itô rules. The change of measure from variable N_1 with parameter Λ_1 to variable N_2 with parameter Λ_2 is given by

$$\Phi_n^{(2)} = \mathrm{Prob}[N_2 < n] = e^{-\Lambda_2} \sum_{k=0}^{n} \frac{\Lambda_2^n}{n!} = e^{-\Lambda_1} \sum_{k=0}^{n} \mathbb{D}_n \frac{\Lambda_1^n}{n!},$$

2006/5/2
page 504

with the Radon-Nikodým discrete derivative satisfying

$$\Delta\Phi^{(2)}_{n-1} = e^{-\Lambda_2}\frac{\Lambda_2^n}{n!} = \mathbb{D}_n e^{-\Lambda_1}\frac{\Lambda_1^n}{n!},$$

and solving yields

$$\mathbb{D}_n = \frac{\Delta\mathbb{P}_2(n-1)}{\Delta\mathbb{P}_1(n-1)} = \frac{\Delta\Phi^{(2)}_{n-1}}{\Delta\Phi^{(2)}_{n-1}} = e^{\Lambda_1 - \Lambda_2}\left(\frac{\Lambda 2}{\Lambda 1}\right)^n$$

$$= e^{\Lambda_1 - \Lambda_2 + n\ln(\Lambda_2/\Lambda_1)}. \tag{13.17}$$

Thus, with the change in measure from $\mathbb{P}_1$ to $\mathbb{P}_2$, the mean or average jump count changes from Λ_1 to Λ_2.

- ***Poisson Distribution with Fixed Size Jumps:***
  Now, consider a Poisson distribution for discrete variable N_1 with parameter Λ_1 and constant jump size $\nu_1 \neq 0$, so

  $$X = \nu_1 N_1.$$

  Given the primary measure

  $$\mathbb{P}_1(x) = \mathrm{Prob}[X \le x] = \mathrm{Prob}[N_1 \le x/\nu_1] = e^{-\Lambda_1}\sum_{k=0}^{\infty}\frac{\Lambda_1^k}{k!}\mathbf{1}_{\{k\le x/\nu_1\}},$$

  a change in measure with parameters $\{\Lambda_2, \nu_2\}$ is sought such that

  $$\mathbb{P}_2(x) = \mathrm{Prob}[X \le x] = \mathrm{Prob}[N_2 \le x/\nu_2] = e^{-\Lambda_2}\sum_{k=0}^{\infty}\frac{\Lambda_2^{k_2}}{k_2!}\mathbf{1}_{\{k_2\le x/\nu_2\}}$$

  $$= e^{-\Lambda_1}\sum_{k=0}^{\infty}\frac{\Lambda_1^{k_1}}{k_1!}\mathbf{1}_{\{k_1\le x/\nu_1\}}\mathbb{D}_{k_1}.$$

  In lieu of a proper derivative for the indicator functions $\mathbf{1}_{\{k_j\le x/\nu_j\}}$ for $j = 1{:}2$, consider the increment at $x = (n-1)\nu_2$,

  $$\Delta\mathbb{P}_2((n-1)\nu_2) = \mathbb{P}_2((n-1)\nu_2 + \Delta x) - \mathbb{P}_2((n-1)\nu_2)$$

  $$= e^{-\Lambda_2}\sum_{k=0}^{\infty}\frac{\Lambda_2^{k_2}}{k_2!}\mathbf{1}_{\{n-1<k_2\le n-1+\Delta x/\nu_2\}}$$

  $$= e^{-\Lambda_1}\sum_{k=0}^{\infty}\frac{\Lambda_1^{k_1}}{k_1!}\mathbf{1}_{\{(n-1)\nu_2/\nu_1<k_1\le(n-1)\nu_2/\nu_1+\Delta x/\nu_1\}}\mathbb{D}_{k_1}.$$

  Aside from the coupling of the potential Radon-Nikodým discrete derivatives $\mathbb{D}_{k_1}$, as Cont and Tankov [59, Prop. 9.5] state that the two measures will not

DRAFT

be equivalent since their null sets will in general not coincide unless the jump sizes are the same, $\nu_2 = \nu_1$.

Thus, with $\nu_2 = \nu_1$ and $\Delta x = \nu_1$ for a semi-open unit step $(n-1, n]$, the new measure increment becomes

$$\begin{aligned}\Delta\mathbb{P}_2((n-1)\nu_1) = \mathbb{P}_2(n\nu_1) - \mathbb{P}_2((n-1)\nu_1) \\ &= e^{-\Lambda_2}\sum_{k=0}^{\infty}\frac{\Lambda_2^{k_2}}{k_2!}\mathbf{1}_{\{n-1<k_2\le n\}} = e^{-\Lambda_2}\frac{\Lambda_2^n}{n!} \\ &= e^{-\Lambda_1}\sum_{k=0}^{\infty}\frac{\Lambda_1^{k_1}}{k_1!}\mathbf{1}_{\{n-1<k_1\le n\}}\mathbb{D}_{k_1} = e^{-\Lambda_1}\frac{\Lambda_1^n}{n!}\mathbb{D}_n,\end{aligned}$$

so obtaining the same Radon-Nikodým discrete derivative as in the previous unit step example (13.17)

$$\mathbb{D}_n = \frac{\Delta\mathbb{P}_2(n-1)}{\Delta\mathbb{P}_1(n-1)} = e^{\Lambda_1 - \Lambda_2 + n\ln(\Lambda_2/\Lambda_1)}. \tag{13.18}$$

*Note that although the original measures $\mathbb{P}_j(n\nu_j)$ are RCLL as they should be, inherited from the indicators $\mathbf{1}_{\{k\le n\}}$, the increment $\Delta\mathbb{P}_1((n-1)\nu_1)$ is **LCRL** (**left continuous, right limits**) due to the indicator increments $\mathbf{1}_{\{n-1<k_2\le n\}}$, but they precisely allow the selection of just the nth jump term in the Poisson distribution sum since the indicator increments are closed at n and open at $n-1$.*

This Poisson distribution example is an applied justification of the proposition in Cont and Tankov [59, Prop. 9.5] for two Poisson processes $n = N_j = P(T) = N(T)$ with parameters $\Lambda_j = \lambda_j T$ on $(\Omega, \mathcal{F}_T, \mathbb{P}_j, \mathbb{F})$ for $j = 1{:}2$, i.e.,

$$\frac{\Delta\mathbb{P}_2(N(T)-1)}{\Delta\mathbb{P}_1(N(T)-1)} = e^{(\lambda_1-\lambda_2)T + N(T)\ln(\lambda_2/\lambda_1)}, \tag{13.19}$$

but only for the same size, $\nu_2 = \nu_1$, which has an explicit form as given here.

13.2.2 Change in Measure for Stochastic Processes: Girsanov's Theorem

There are many versions of Girsanov's theorem for changing a probability measure to change the drift of a stochastic diffusion process and some of these variants are not very distinguishable from the Radon-Nikodým theorem. Here, a modification of Runggaldier's [235] (see also Brémaud [43] for even more details) version will be followed since it has been found to be the most useful, the Radon-Nikodým derivative being relatively easy to calculate and comes with an extension to jump-diffusions. The application of this theorem is determining the measure change for a relative change $\gamma(t)$ of the drift from $\mu_1(t)$ to a drift $\mu_2(t)$ appropriate for the problem of interest, e.g., the change of the drift coefficient $\mu_1(t) = \mu$ in the Black-Scholes [34] method to the current market rate $\mu_2(t) = r$.

Diffusion Girsanov Transformations

Let the reference $\mathbb{P}_1$-SDE for a state diffusion process $X(t)$ be

$$dX(t) = \mu_1(t)dt + \sigma(t)dW_1(t) \tag{13.20}$$

with time-dependent coefficients $\{\mu_1(t), \sigma(t)\}$, whose integrabilities are implied by the following Girsanov diffusion theorem, on a finite time interval $[0, T]$ on the filtered probability space $(\Omega, \mathcal{F}_t, \mathbb{P}_1, \mathbb{F})$ with $W_1(t)$ being a $\mathbb{P}_1$-Wiener process. In addition, let the target $\mathbb{P}_2$-SDE objective for this state diffusion process $X(t)$ be

$$dX(t) = \mu_2(t)dt + \sigma(t)dW_2(t) \tag{13.21}$$

with the same volatility $\sigma(t)$ but changed to drift $\mu_2(t)$, integrability also implied, on the finite time interval $[0, T]$ on the filtered probability space $(\Omega, \mathcal{F}_t, \mathbb{P}_2, \mathbb{F})$ with $W_1(t)$ being a corresponding $\mathbb{P}_2$-Wiener process.

Theorem 13.14. ***Girsanov's Theorem for Changing the Probability Measure of a Diffusion Process to Change the Drift:***
Let $(\Omega, \mathcal{F}_t, \mathbb{P}_1, \mathbb{F})$ be a filtered probability space with $\mathbb{F} = \cup_t \mathcal{F}_t$, symbolically over t. Let $\gamma(t)$ be a square integrable predictable (measurable with respect to $\mathbb{P}_1$, i.e., knowable given $\mathcal{F}_t$) drift process

$$\int_0^t \gamma^2(s)ds < \infty$$

for all $t \in [0, T]$. Then, the Radon-Nikodým derivative $\mathbb{D}(t)$ at time t for the process $X(t)$ is given by the martingale representatiion (13.9),

$$d\mathbb{D}(t) = \mathbb{D}(t)\gamma(t)dW_1(t), \quad \mathbb{D}(0) \overset{\text{w.p.o.}}{=} 1, \tag{13.22}$$

supposing that $\mathrm{E}_{\mathbb{P}_1}[\mathbb{D}(t)] = 1$ and there exists a second probability measure $\mathbb{P}_2$ on $\mathbb{F}$ that is equivalent to $\mathbb{P}_1$ (mutually absolutely continuous, $\mathbb{P}_2 \overset{\text{a.c.}}{\equiv} \mathbb{P}_1$), such that

$$d\mathbb{P}_2 = \mathbb{D}(t)d\mathbb{P}_1$$

and

$$dW_2(t) = dW_1(t) - \gamma(t)dt, \tag{13.23}$$

where $W_1(t)$ is a $\mathbb{P}_1$-Wiener process as in (13.20) while $W_2(t)$ is a $\mathbb{P}_2$-Wiener process as in (13.21).

The Radon-Nikodým derivative is explicitly given by

$$\mathbb{D}(t) = \frac{d\mathbb{P}_2(t)}{d\mathbb{P}_1(t)} = \exp\left(\int_0^t \gamma(s)\left(-\frac{1}{2}\gamma(s)ds + dW_1(s)\right)\right) \tag{13.24}$$

and the relative drift change is

$$\gamma(t) = \frac{\mu_2(t) - \mu_1(t)}{\sigma(t)}. \tag{13.25}$$

If the filtration

$$\mathbb{F} = \mathcal{F}_t^{(W_1)} = \sigma\{W_1(s),\ \mathcal{S}_N : 0 \le s \le t,\ \mathcal{S}_N \in \mathcal{N}_1\},$$

$\mathcal{N}_1$ *is the collection of null-sets of* $\mathbb{P}_1$*, then conversely every probability measure* $\mathbb{P}_2 \overset{\text{a.c.}}{\equiv} \mathbb{P}_1$ *has the same Radon-Nikodým derivative structure.*

Substituting (13.23) the Wiener process shift into the original SDE,

$$dX(t) = \mu_1(t)dt + \sigma(t)dW_1(t) = (\mu_1(t) + \sigma(t)\gamma(t))dt + \sigma(t)dW_2(t),$$

so comparing to the second SDE $\mu_2(t) = \mu_1(t) + \sigma(t)\gamma(t)$ and (13.25) for $\gamma(t)$ is immediate, given common volatilities $\sigma_1(t) = \sigma(t) = \sigma_2(t)$.

Upon applying the Itô stochastic chain rule to solve the $\mathbb{D}$-SDE (13.22),

$$d\ln(\mathbb{D}(t)) \overset{\text{dt}}{=} \frac{\gamma\mathbb{D}dW_1}{\mathbb{D}} - \frac{(\gamma\mathbb{D})^2 dt}{2\mathbb{D}^2} = \gamma\left(-\frac{1}{2}\gamma dt + dW_1\right),$$

integrating with $\mathbb{D}(0) = 1$ and inverting the logarithm, the answer for $\mathbb{D}(t)$ in (13.24) follows. Note that the assumption of common volatility is essential for obtaining the simple linear SDE in $\mathbb{D}(t)$ given in (13.22), since from just one of the independent example terms i in (13.16) it is seen that there is a quadratic term in x of the ith exponent unless $\sigma_i^{(2)} = \sigma_i = \sigma_i^{(1)} = 1$, the common σ in this example. Hence, this Girsanov theorem is quite simple and special. The crudely derived constant coefficient case in (13.14), as an example for Radon-Nikodým derivatives, can be properly recovered from the Girsanov form (13.24) by setting $t = T$ and replacing the time-dependent coefficients by constants, i.e., $\mu_j(s) \to \mu_j$ for $j = 1:2$ and $\sigma(s) \to \sigma$.

Note that the relative drift shift (13.25), being state independent, is also the same for the linear diffusion case,

$$dX(t) = X(t)\left(\mu_1(t)dt + \sigma(t)dW_1(t)\right), \tag{13.26}$$

which is important for applications in finance. This is a linear SDE for **geometric Brownian motion (GBM) or multiplicative diffusion noise** of the Black-Scholes-Merton[34, 197] option pricing model, while the reference SDE 13.26 for the Theorem 13.14 is for **arithmetic Brownian motion (ABM) or additive diffusion noise** of the historic 1900 Bachelier [16] model. The multiplicative model is better for compounded effects, while the additive model is better for strictly cumulative effects. It is well-known that the multiplicative model can be transformed into an additive one by the logarithmic transformation using Itô rules,

$$d\ln(X(t)) = \left(\mu_1 - \sigma^2(t)/2\right)dt + \sigma(t)dW_1(t). \tag{13.27}$$

Since the diffusion coefficient shift, $\sigma^2(t)/2$, of the drift would be the same for the GBM target model (2) as for the GBM reference model (1), it is clear that the diffusion Girsanov transformation of the drift will be the same as for the ABM model, i.e.,

$$\gamma^{(\mathrm{GBM})}(t) = \frac{\mu_2(t) - \mu_1(t)}{\sigma(t)}. \tag{13.28}$$

DRAFT

Marked-Jump-Diffusion Girsanov Transformations

Now consider the case of marked-jump-diffusions or compound-jump-diffusion. Let the reference $\mathbb{P}_1$-SDE for a state marked-jump-diffusion process $X(t)$ be

$$dX(t) = \mu_1(t)dt + \sigma(t)dW_1(t) + \sum_{i=1}^{dP_1(t)} h_1(T_i, Q_i) \tag{13.29}$$

with $\mathbb{P}_1$-Wiener process $W_1(t)$, $\mathbb{P}_1$-Poisson process $P_1(t)$, $\mathrm{E}_{\mathbb{P}_1}[dP_1(t)] = \lambda_1(t)dt$ defines the jump-rate, integrable time-dependent coefficients $\{\mu_1(t), \sigma(t), \lambda_1(t)\}$, (time, mark)-dependent jump-amplitude $h_1(t, q)$, whose integrability is implied by the following theorem, $\mathbb{P}_1$-Poisson jump-times T_i and IID sample marks Q_i distributed with density $\phi_Q^{(1)}(q; t)$ on the filtered probability space $(\Omega, \mathcal{F}_t, \mathbb{P}_1, \mathbb{F})$ with $W_1(t)$ over on a finite time-interval $[0, T]$. In addition, let the target $\mathbb{P}_2$-SDE objective for this state marked-jump-diffusion process $X(t)$ be

$$dX(t) = \mu_2(t)dt + \sigma(t)dW_2(t) + \sum_{i=1}^{dP_2(t)} h_2(T_i, Q_i) \tag{13.30}$$

with $\mathbb{P}_2$-Wiener process $W_2(t)$, $\mathbb{P}_2$-Poisson process $P_2(t)$, $\mathrm{E}_{\mathbb{P}_2}[dP_2(t)] = \lambda_2(t)dt$ defines the jump-rate, the same volatility $\sigma(t)$ but changed to drift $\mu_2(t)$ and changed jump-rate $\lambda_2(t)$, (time, mark)-dependent jump-amplitude $h_2(t, q)$, integrability also implied, $\mathbb{P}_2$-Poisson jump-times T_i and IID sample marks Q_i distributed with density $\phi_Q^{(2)}(q; t)$ on the finite time interval $[0, T]$ on the filtered probability space $(\Omega, \mathcal{F}_t, \mathbb{P}_2, \mathbb{F})$.

The following theorem follows Runggaldier [235], but is also presented more in the notation of this book.

Theorem 13.15. ***Girsanov's Theorem for Changing the Probability Measure of a Jump-Diffusion Process to Change the Drift and the Jump-Rate:***
Let $(\Omega, \mathcal{F}_t, \mathbb{P}_1, \mathbb{F})$ a filtered probability space on the finite time-interval $[0, T]$ with the mark space $\mathcal{Q} = \mathbb{R}$. and the (jump-rate, mark-density)-characteristics

$$(\lambda_1(t), \Phi_Q(\mathbf{dq}; t) = \phi_Q(q; t)dq).$$

Let $\gamma^{(\mathrm{D})}(t)$ be the square integrable diffusion drift change given in (13.25)

$$\gamma^{(\mathrm{D})}(t) = \frac{\mu_2(t) - \mu_1(t)}{\sigma(t)} \tag{13.31}$$

of Theorem 13.14. Let $\gamma^{(\mathrm{J}))}(t)$ be a nonnegative, $\mathcal{F}_t$-predictable jump-rate scaling process such that

$$\int_0^t \lambda_1(s)\gamma^{(\mathrm{J}))}(s)ds < \infty$$

2006/5/24
page 509

for all $t \in [0,T]$ and let $\gamma^{(\mathrm{M})}(q;t)$ be a nonnegative, $\mathcal{F}_t$-predictable, $\mathcal{Q}$-space dependent mark-distribution scaling process such that

$$\int_{\mathcal{Q}} \gamma^{(\mathrm{M})}(q;s)\phi_Q^{(1)}(q;t)ds = 1,$$

i.e., mark-space probability is conserved.

Let

$$\mathbb{D}(t) = \mathbb{D}^{(\mathrm{D})}(t)\mathbb{D}^{(\mathrm{MJ})}(t),$$

where the diffusion martingale representation factor $\mathbb{D}^{(\mathrm{D})}(t)$ is given in (13.24) with stochastic differential in (13.22) and the marked-jump factor $\mathbb{D}^{(\mathrm{MJ})}(t)$. Letting

$$K^{(\mathrm{MJ})}(q;t) \equiv \gamma^{(\mathrm{J}))}(t)\gamma^{(\mathrm{M})}(q;t) - 1 \tag{13.32}$$

$$\overline{K}^{(\mathrm{MJ})}(t) = \int_{\mathcal{Q}} K^{(\mathrm{MJ})}(q;t)\phi_Q^{(1)}(q;t)ds, \tag{13.33}$$

then $\mathbb{D}^{(\mathrm{MJ})}(t)$ factor is given by the marked-jump martingale representation (13.9),

$$d\mathbb{D}^{(\mathrm{MJ})}(t) = \sum_{i=1}^{dP_1(t)} K^{(\mathrm{MJ})}(Q_i;T_i)\mathbb{D}^{(\mathrm{MJ})}(T_i^-) - \overline{K}^{(\mathrm{MJ})}(t)\mathbb{D}^{(\mathrm{MJ})}(t)\lambda_1(t)dt, \tag{13.34}$$

where $\mathrm{E}_{\mathbb{P}_1}[\mathbb{D}^{(\mathrm{MJ})}] = 1$ is assumed and the changed quantities are

$$dW_2(t) = dW_1(t) - \gamma^{(\mathrm{D})}(t)dt, \tag{13.35}$$

$$\lambda_2(t) = \gamma^{(\mathrm{J}))}(t)\lambda_1(t) \tag{13.36}$$

$$\phi_Q^{(2)}(q;t) = \gamma^{(\mathrm{M})}(t)\phi_Q^{(1)}(q;t). \tag{13.37}$$

Thus, the explicit form of the marked-jump-diffusion Radon-Nikodým derivative is

$$\begin{aligned}\mathbb{D}(t) &= \frac{d\mathbb{P}_2(t)}{d\mathbb{P}_1(t)} \\ &= \exp\left(\int_0^t \gamma^{(\mathrm{D})}(s)\left(dW_1(s) - \gamma^{(\mathrm{D})}(s)ds/2\right) - \overline{K}^{(\mathrm{MJ})}(s)\lambda_1(s)ds\right) \\ &\quad \cdot \prod_{i=1}^{P_1(t)} \left(K^{(\mathrm{MJ})}(Q_i;T_i) + 1\right).\end{aligned} \tag{13.38}$$

If the filtration

$$\mathbb{F} = \mathcal{F}_t^{(W_1,P_1,Q)} = \sigma\{W_1(s), P_1(t), \mathcal{S}_Q,\ \mathcal{S}_N : 0 \le s \le t,\ \mathcal{S}_Q \in \mathcal{Q},\ \mathcal{S}_N \in \mathcal{N}_1\},$$

$\mathcal{N}_1$ is the collection of null-sets of $\mathbb{P}_1$, then conversely every probability measure $\mathbb{P}_2 \overset{\mathrm{a.c.}}{\equiv} \mathbb{P}_1$ has the same Radon-Nikodým derivative structure.

Note that the Wiener process W_1 is independent of the marked Poisson process double (P_1, Q), but the mark random variables Q are only conditionally independent of P_1 and that condition is that there exists a jump of the state X in time, so the factoring $\mathbb{D}(t) = \mathbb{D}^{(\mathrm{D})}(t)\mathbb{D}^{(\mathrm{MJ})}(t)$ into only two parts makes sense. Also, using Itô's stochastic chain rule,

$$\begin{aligned} d\mathbb{D}(t) &= d\mathbb{D}^{(\mathrm{D})}(t)\mathbb{D}^{(\mathrm{MJ})}(t) + \mathbb{D}^{(\mathrm{D})}(t)d\mathbb{D}^{(\mathrm{MJ})}(t) + d\mathbb{D}^{(\mathrm{D})}(t)d\mathbb{D}^{(\mathrm{MJ})}(t) \\ &\stackrel{\mathrm{dt}}{=} d\mathbb{D}^{(\mathrm{D})}(t)\mathbb{D}^{(\mathrm{MJ})}(t) + \mathbb{D}^{(\mathrm{D})}(t)d\mathbb{D}^{(\mathrm{MJ})}(t) \end{aligned}$$

Since (13.34) for $\mathbb{D}^{(\mathrm{MJ})}$ is linear,

$$d\ln(\mathbb{D}^{(\mathrm{MJ})}(t) = -\overline{K}^{(\mathrm{MJ})}(t)\lambda_1(t)dt + \sum_{i=1}^{dP_1(t)} \ln\left(K^{(\mathrm{MJ})}(Q_i;T_i)+1\right),$$

since if dP_1 jumps the jump $[\mathbb{D}^{(\mathrm{MJ})}] = K^{(\mathrm{MJ})}\mathbb{D}^{(\mathrm{MJ})}$, then the jump

$$[\ln(\mathbb{D}^{(\mathrm{MJ})})] = \ln(\mathbb{D}^{(\mathrm{MJ})} + K^{(\mathrm{MJ})}\mathbb{D}^{(\mathrm{MJ})}) - \ln(\mathbb{D}^{(\mathrm{MJ})}) = \ln(K^{(\mathrm{MJ})}+1),$$

so

$$\mathbb{D}^{(\mathrm{MJ})}(t) = \exp\left(-\int_0^t \overline{K}^{(\mathrm{MJ})}(s)\lambda_1(s)ds + \sum_{i=1}^{P_1(t)} \ln\left(K^{(\mathrm{MJ})}(Q_i;T_i)+1\right)\right). \quad (13.39)$$

Finally, combining equations (13.22), (13.24), (13.34) and (13.39), along with converting the exponential of a sum to a product yields the result (13.38) for $\mathbb{D}(t)$ for the marked-jump-diffusion change from measure $\mathbb{P}_1$ to $\mathbb{P}_2$ according to the recipe (13.35) to (13.37).

For the geometric or linear marked-jump-diffusion,

$$dX(t) = X(t)\left(\mu_1(t)dt + \sigma(t)dW_1(t) + \sum_{i=1}^{dP_1(t)} h_1(T_i, Q_i)\right), \quad (13.40)$$

the logarithmic change of variable can transform the geometric model to an arithmetic one like (13.29),

$$d\ln(X(t)) = \left(\mu_1(t) - \sigma^2(t)/2\right)dt + \sigma(t)dW_1(t) + \sum_{i=1}^{dP_1(t)} \ln\left(h_1(T_i, Q_i)+1\right). \quad (13.41)$$

Again assuming a common volatility $\sigma(t)$, the Itô rule diffusion coefficient shift of the drift coefficient will be common in both target (2) and reference (1) models, while the jump-rate $\lambda_1(t)$ and jump-amplitude distribution is unchanged, then the Girsanov tranformation triplet,

$$\begin{aligned} dW_2(t) &= dW_1(t) - \gamma^{(\mathrm{D})}(t)dt, \\ \lambda_2(t) &= \gamma^{(\mathrm{J}))}(t)\lambda_1(t) \\ \phi_Q^{(2)}(q;t) &= \gamma^{(\mathrm{M})}(t)\phi_Q^{(1)}(q;t), \end{aligned} \quad (13.42)$$

will be preserved for the geometric case.

Refer to Runggaldier's jump-diffusion handbook article [235] for more information on the multidimensional case, Poisson random measure formulation and financial applications.

13.3 Itô, Lévy and Jump-Diffusion Comparisons

13.3.1 Itô Processes and Jump-Diffusion Processes

Many authors, Bingham and Kiesel [33] Duffie [74], Glasserman [96], Hull [144], Merton [199], Mikosch [205], Øksendal [218] and others, mostly refer to Brownian motion or Wiener-driven processes with Wiener scaling by a factor $\sigma(t)$ and translated by drift $\mu(t)$,

$$dX(t) = \mu(t)dt + \sigma(t)dW(t), \tag{13.43}$$

at least as basic definition of an **Itô process**. Some such as Glasserman [96], Hull [144], Merton [199] and Mikosch [205] would explicitly allow the composite interpretation of the coefficient functions in the basic definition (13.43) to include dependence on the state $X(t)$, such that $\mu(t) = f(X(t), t)$, $\sigma(t) = g(X(t), t)$ and

$$dX(t) = f(X(t), t)dt + g(X(t), t)dW(t). \tag{13.44}$$

Others extend the basic Itô process (13.43) to include (13.44) by application of the Itô chain rule using a transformation like $\widehat{X}(t) = F(X(t), t)$ to obtain

$$d\widehat{X}(t) = f(X(t), t)dt + g(X(t), t)dW(t),$$

where

$$f(X(t), t) = \left(F_t + \mu(t)F_x + \frac{1}{2}\sigma^2(t)F_{xx}^2)\right)(X(t), t)$$

and

$$g(X(t), t) = \sigma(t)F_x(X(t), t).$$

Thus, state dependent formula (13.44) will be taken as an acceptable **definition of the Itô process**.

However, in his stochastic differential equation classic 1951 memoir [146], Itô also correctly includes jumps in his discussion of simple Markov processes. Itô referred to simple Markov processes, specified by a stochastic differential equation, which for general Poisson noise with distributed jump-amplitudes might be called stochastic integral differential equations,

$$dX(t) = f(X(t), t)dt + g(X(t), t)dW(t) + \int_{\mathcal{Q}} h(X(t), t, q)\mathcal{P}(dt, dq)\ , \tag{13.45}$$

in our notation, or preferably by a stochastic integral equation,

$$\begin{aligned} X(t) = X(t_0) + \int_{t_0}^{t} \Big[& f(X(s), s)ds + g(X(s), s)dW(s) \\ & + \int_{\mathcal{Q}} h(X(s), s, q)\mathcal{P}(ds, dq) \Big]\ , \end{aligned} \tag{13.46}$$

DRAFT

2006/5/2
page 512

again in our notation. Hence, there is a historical basis for calling the jump-diffusion processes that are the focus of this book as **Itô processes**.

Still others, for instance, Tavella and Randall [259] refer to a jump-diffusion processes as a superposition of an Itô process and a Poisson jump process, while Øksendal and Sulem [219] refer to a similar combination as an Itô-Lévy process, but see the next subsection on Lévy processes for the differences between jump-diffusion and Lévy processes. Applebaum [12] and others more precisely call diffusion processes like (13.44) **Itô diffusion processes**.

Although diffusion processes are easier to treat since they have continuous sample paths, jump processes and jump-diffusion processes have discontinuous sample paths so are relatively more difficult to prove theorems for. Some of the most significant changes occur with jumps, such as extreme financial crashes and natural disasters.

Hence, according to the more of less standard Itô process usage (13.44),

$$\textit{\textbf{Itô processes}} \subset \textit{\textbf{Jump-diffusion processes}}. \tag{13.47}$$

13.3.2 Lévy Processes and Jump-Diffusion Processes

Lévy processes are essentially jump-diffusion processes, but extended to processes with infinite jump rates. There have been much recent efforts in the literature studying and applying Lévy processes, such as Carr, Geman, Madan and Yor (CGMY model) [46], Carr and Madan (VG model) [47] and Rydberg (NIG model) [239]. Sometimes the term non-Gaussian processes is used as in Barndorff-Nielsen and Shepherd (GIG model) [20], but may not necessarily mean strict Lévy processes. There also several recent books on Lévy processes such as that of Applebaum [12], as well as others on Lévy processes but with jump processes or jump-diffusions in the titles such as those of Cont and Tankov [59] and Øksendal and Sulem [219]. As with other abstract concepts, there are many different definitions of a Lévy process, and some attempt has been made to merge them within the spirit of this book.

Definition 13.16. ***Basic Lévy Process Conditions:***
*A **Lévy process** satisfies the following conditions:*

- ***RCLL stochastic process**: $\{\mathbf{X}(t),\ t \geq 0\}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^{n_x}$ (the term **cádlág** means RCLL in French but is used in English probability texts too).*

- ***Initial condition**: $\mathbf{X}(0) \overset{\text{a.s.}}{=} \mathbf{0}$.*

- ***Independent increments**: for every partition $0 = t_0 < t_1 < t_2 < \cdots < t_{n_t} < \infty$, the increments*

$$\Delta\mathbf{X}(t_j) \equiv \mathbf{X}(t_{j+1}) - \mathbf{X}(t_j),\ \text{for } j = 0{:}n-1 \tag{13.48}$$

are independent.

- ***Stationary increments:*** *Together with independence,*

$$\Delta\mathbf{X}(t_j) \stackrel{\text{dist}}{=} \mathbf{X}(\Delta t_j), \tag{13.49}$$

where $\Delta t_j \equiv t_{j+1} - t_j$.

- ***Stochastic Continuity:*** *The increments of $\mathbf{X}(t)$ satisfy,*

$$\lim_{\Delta t\to 0} \mathrm{Prob}[\mathbf{X}(t+\Delta t) - \mathbf{X}(t)| \ge \epsilon] = 0, \quad \forall\, \epsilon > 0 \quad \text{and} \quad t \ge 0. \tag{13.50}$$

All but the last condition (13.50) are standard for the processes dealt with here when the **coefficients are constant**, so it is usually sufficient to show stochastic continuity (*note that continuous in probability is not the same as continuous*). However, when the process coefficients are not constant, then the process will in general not be stationary as Lévy condition (13.49) requires. For many real problems the process coefficients, as in financial markets, time-dependence is important (for instance, see Hanson and Westman [126]), so (13.49) will not be valid in these problems. It is clear that the IID Wiener vector process $\mathbf{W}(t)$ or the Wiener driven vector Gaussian process with constant coefficients

$$\mathbf{G}(t) = \boldsymbol{\mu}_0 t + \sigma_0 \mathbf{W}(t)$$

and the Poisson vector process $\mathbf{P}(t)$ with constant jump rates $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}_0$ will all be Lévy processes, as well as any linear combination that is the simple constant coefficient jump-diffusion n_x-vector process,

$$\mathbf{X}(t) = \boldsymbol{\mu}_0 t + \sigma_0 \mathbf{W}(t) + \nu_0 \mathbf{P}(t),$$

where $\sigma_0 \in \mathbb{R}^{n_x \times n_w}$ and $\nu_0 \in \mathbb{R}^{n_x \times n_p}$ consistent with IID $\mathbf{W}(t) \in \mathbb{R}^{n_w}$ and IID $\mathbf{P}(t) \in \mathbb{R}^{n_p}$. Adding the compound Poisson process to the combination will be discussed in the sequel.

There are some preliminary definitions that are important for further properties of Lévy processes.

Definition 13.17. ***Infinitely Divisible Distribution:***
A probability distribution $\Phi_{\mathbf{X}}$ on $\mathbb{R}^{n_x}$ is infinitely divisible if for each positive integer n there exists of a set of IID random variable $\mathbf{Y}_j$ for $j = 1\!:\!n$ such that the sum

$$\mathbf{S}_n = \sum_{j=1}^{n} \mathbf{Y}_j \stackrel{\text{dist}}{=} \mathbf{X},$$

where $\mathbf{X}$ has distribution $\Phi_{\mathbf{X}}$.

Infinitely Divisibility can be related to the central limit theorem and is closely connected to Lévy processes via compound Poisson processes as follows [59].

Proposition 13.18. ***Lévy processes and Infinitely Divisibility:***
Let $\mathbf{X}(t)$ be a Lévy process for $t \geq 0$ on $\mathbb{R}^{n_x}$, then for every t, $X(t)$ has an infinitely divisible distribution. Conversely, if Φ is an infinitely divisible distribution, then there exists a Lévy process $\mathbf{X}(t)$ with the distribution Φ.

The compound Poisson process is included firmly as a Lévy process by the following result proved in Cont and Tankov [59],

Proposition 13.19. ***Compound Poisson Processes as Lévy Processes:***
*The process $CP(t)$ for $t \geq 0$ is a **compound Poisson process**, i.e.,*

$$CP(t) = \sum_{j=1}^{P(t)} Q_i, \tag{13.51}$$

*where $P(t)$ is a **simple Poisson process** with constant rate λ_0 and the Q_j are IID random jump-amplitudes with common distribution $\Phi_Y(y)$ such that λ_0 are independent of the Q_i,*

if and only if

*$CP(t)$ is a Lévy process and its sample paths are **piecewise constant functions**.*

Characteristic Functions and Lévy Characteristic Exponents

Definition 13.20. ***Characteristic Function:***
The characteristic function of a random vector $\mathbf{X}$ on $\mathbb{R}^{n_x}$ is the complex-valued function,

$$\mathcal{C}_{\mathbf{X}}(\mathbf{z}) \equiv \mathrm{E}_{\mathbf{X}}\left[\exp\left(i\mathbf{z}^\top \mathbf{X}\right)\right] \tag{13.52}$$

for all $\mathbf{z} \in \mathbb{R}^{n_x}$ and i is the imaginary unit.

Clearly, the characteristic function of a **continuous random variable X** is the **Fourier transform** of the density of **X**, i.e.,

$$\mathcal{C}_{\mathbf{X}}(\mathbf{z}) = \int_{\mathbb{R}^{n_x}} e^{i\mathbf{z}^\top \mathbf{x}} \phi_{\mathbf{X}}(\mathbf{x}) d\mathbf{x},$$

while if X is a **discrete scalar random variable** with distribution given by the countable sequence of probabilities $\pi_k = \mathrm{Prob}[X = k]$, then the characteristic function is the discrete Fourier transform,

$$\mathcal{C}_X(z) = \sum_{k=0}^{\infty} e^{izk} \cdot \pi_k.$$

This is the basic random vector definition, but here the interest will be the same definition when the random vector is a function of time t, i.e., a stochastic process $\mathbf{X}(t)$,

$$\mathcal{C}_{\mathbf{X}(t)}(\mathbf{z}) \equiv \mathrm{E}_{\mathbf{X}(t)}\left[\exp\left(i\mathbf{z}^\top \mathbf{X}(t)\right)\right].$$

One of the most important features of a Lévy process is that the characteristic function has relatively simple form [59, 12]

Proposition 13.21. ***Lévy Characteristic Functions and Exponents:*** *If $\mathbf{X}(t)$ is a Lévy process for $t \geq 0$ on $\mathbb{R}^{n_x}$, then there exist a real-valued continuous function $\eta_{\mathbf{X}(t)}(\mathbf{z})$ of the characteristic vector $\mathbf{z} \in \mathbb{R}^{n_x}$, called the* ***characteristic exponent****, such that*

$$\mathcal{C}_{\mathbf{X}(t)}(\mathbf{z}) = \mathrm{E}_{\mathbf{X}(t)}\left[\exp\left(i\mathbf{z}^\top \mathbf{X}(t)\right)\right] = \exp\left(it\eta_{\mathbf{X}(t)}(\mathbf{z})\right). \tag{13.53}$$

However, for nonstationary problems without the Lévy stationarity condition (13.50), then it would be expected that in general the exponent will not be linear in t,

$$\mathcal{C}_{\mathbf{X}(t)}(\mathbf{z}) = \exp\left(i\overline{\eta}_{\mathbf{X}(t)}(\mathbf{z};t)\right).$$

It is well-known that Fourier transforms, and the characteristic function, is mainly useful for constant coefficients, with few exceptions.

Examples 13.22. ***Characteristic Functions and Exponents of Lévy Processes:***

- ***Standard Wiener Process*** $W(t)$ ***on*** $\mathbb{R}$***:***

$$\mathcal{C}_{W(t)}(z) = \mathrm{E}_{W(t)}\left[e^{izW(t)}\right] = \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty} e^{izw}e^{-w^2/(2t)}dw$$
$$= e^{-(tz)^2/(2t)}\frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty} e^{-(w-itz)^2/(2t)}dw = e^{-tz^2/2},$$

using the completing the square technique, so the Lévy characteristic exponent is

$$\eta_{W(t)}(z) = -\frac{1}{2}z^2. \tag{13.54}$$

- ***IID Wiener Vector Process*** $\mathbf{W}(t)$ ***on*** $\mathbb{R}^{n_w}$ ***with*** $\mathrm{Cov}[\mathbf{W}(t), \mathbf{W}^\top(t)] = tI_{n_w}$***:***

$$\mathcal{C}_{\mathbf{W}(t)}(\mathbf{z}) = \mathrm{E}_{\mathbf{W}(t)}\left[e^{i\mathbf{z}^\top\mathbf{W}(t)}\right] = \prod_{j=1}^{n_w}\frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty} e^{iz_jw_j}e^{-w_j^2/(2t)}dw_j$$
$$= \prod_{j=1}^{n_w}\mathcal{C}_{W_j(t)}(z_j) = \exp\left(-t\sum_{j=1}^{n_x}z_j^2/2\right) = \exp\left(-t|\mathbf{z}|^2/2\right),$$

so the Lévy characteristic exponent is

$$\eta_{\mathbf{W}(t)}(\mathbf{z}) = -\frac{1}{2}|\mathbf{z}|^2. \tag{13.55}$$

- ***IID Gaussian Vector Process*** $\mathbf{G}(t) = \boldsymbol{\mu}_0 t + \sigma_0 \mathbf{W}(t)$ ***on*** $\mathbb{R}^{n_x}$ ***with*** $\mathrm{Cov}[\mathbf{W}(t), \mathbf{W}^\top(t)] = tI_{n_w}$***, Constant*** $\boldsymbol{\mu}_0 \in \mathbb{R}^{n_x}$ ***and Constant*** $\sigma_0 \in \mathbb{R}^{n_x \times n_w}$***:***

$$\begin{aligned}
\mathcal{C}_{\mathbf{G}(t)}(\mathbf{z}) &= \mathrm{E}_{\mathbf{W}(t)}\left[e^{i\mathbf{z}^\top(\boldsymbol{\mu}_0 t + \sigma_0 \mathbf{W}(t))}\right] \\
&= e^{i\mathbf{z}^\top \boldsymbol{\mu}_0 t} \prod_{k=1}^{n_w} \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \exp\left(i \sum_{j=1}^{n_w} z_j \sigma_{0,j,k} w_k\right) e^{-w_k^2/(2t)} dw_k \\
&= \exp\left(it\mathbf{z}^\top \boldsymbol{\mu}_0 - \sum_{j=1}^{n_x} z_j \sum_{\ell=1}^{n_x} z_\ell \sum_{k=1}^{n_w} \sigma_{0,j,k}\sigma_{0,\ell,k}/2\right) \\
&= \exp\left(it\mathbf{z}^\top \boldsymbol{\mu}_0 - t\mathbf{z}^\top(\sigma_0\sigma_0^\top)\mathbf{z}\right),
\end{aligned}$$

so the Lévy characteristic exponent is

$$\eta_{\mathbf{G}(t)}(\mathbf{z}) = i\mathbf{z}^\top \boldsymbol{\mu}_0 - \frac{1}{2}\mathbf{z}^\top(\sigma_0\sigma_0^\top)\mathbf{z}/2. \tag{13.56}$$

- ***Simple Poisson Process*** $P(t)$ ***on*** $\mathbb{R}$ ***with Constant Jump-Rate*** λ_0***:***

$$\begin{aligned}
\mathcal{C}_{P(t)}(z) = \mathrm{E}_{P(t)}\left[e^{izP(t)}\right] &= e^{-\lambda_0 t} \sum_{k=0}^{\infty} \frac{(\lambda_0 t)^k}{k!} e^{izk} \\
&= e^{-\lambda_0 t} \sum_{k=0}^{\infty} \frac{\left(\lambda_0 t e^{iz}\right)^k}{k!} = e^{-\lambda_0 t + \lambda_0 t e^{iz}} = e^{\lambda_0 t\,(e^{iz} - 1)},
\end{aligned}$$

so the Lévy characteristic exponent is

$$\eta_{P(t)}(z) = \lambda_0(e^{iz} - 1). \tag{13.57}$$

- ***Centered or Martingale Form of Poisson Process*** $\widetilde{P}(t) \equiv P(t) - \lambda_0 t$ ***on*** $\mathbb{R}$ ***with Constant Jump-Rate*** λ_0***:***

$$\begin{aligned}
\mathcal{C}_{\widetilde{P}(t)}(z) = \mathrm{E}_{P(t)}\left[e^{iz(P(t) - \lambda_0 t)}\right] &= e^{-\lambda_0 t} \sum_{k=0}^{\infty} \frac{(\lambda_0 t)^k}{k!} e^{iz(k - \lambda_0 t)} \\
&= e^{-\lambda_0 t(1 + iz)} \mathcal{C}_{P(t)}(z) = e^{\lambda_0 t(e^{iz} - 1 - iz)},
\end{aligned}$$

so the Lévy characteristic exponent is

$$\eta_{\widetilde{P}(t)}(z) = \lambda_0(e^{iz} - 1 - iz). \tag{13.58}$$

- ***Simple Poisson Vector Process*** $\mathbf{P}(t)$ ***on*** $\mathbb{R}^{n_p}$ ***with Independent Components and Constant Jump-Rate Vector*** $\boldsymbol{\lambda}_0 = [\lambda_{0,j}]_{n_p\times 1}$***:***

$$\begin{aligned}
\mathcal{C}_{\mathbf{P}(t)}(\mathbf{z}) = \mathrm{E}_{\mathbf{P}(t)}\left[\exp\left(i\mathbf{z}^\top\mathbf{P}(t)\right)\right] &= \prod_{j=1}^{n_p} e^{-\lambda_{0,j}t}\sum_{k_j=0}^{\infty}\frac{(\lambda_{0,j}t)^{k_j}}{k_j!}e^{iz_jk_j} \\
&= \prod_{j=1}^{n_p}\exp(\lambda_{0,j}t\,(\exp(iz_j)-1)) = \exp\left(t\sum_{j=1}^{n_p}\lambda_{0,j}\left(\exp(iz_j)-1\right)\right) \\
&= \exp\Big(tn_p\left(\overline{\lambda_0\exp(iz)}-\overline{\lambda_0}\right)\Big),
\end{aligned}$$

where $\overline{\lambda_0} \equiv \sum_{j=1}^{n_p}\lambda_{0,j}/n_p$ *and* $\overline{\lambda_0\exp(iz)} \equiv \sum_{j=1}^{n_p}\lambda_{0,j}\exp(iz_j)/n_p$*, so the Lévy characteristic exponent is*

$$\eta_{\mathbf{P}(t)}(\mathbf{z}) = n_p\left(\overline{\lambda_0\exp(iz)}-\overline{\lambda_0}\right). \tag{13.59}$$

- ***Simple Compound Poisson Process*** $CP(t) = \sum_{\ell=1}^{P(t)} Q_\ell$ ***on*** $\mathbb{R}$ ***with Constant Jump-Rate*** λ_0 ***and IID Jump-Amplitudes*** Q_ℓ ***with Distribution*** $\Phi_Q(q)$***:***

$$\begin{aligned}
\mathcal{C}_{CP(t)}(z) = \mathrm{E}_{P(t),Q}\left[e^{izX(t)}\right] &= e^{-\lambda_0 t}\sum_{k=0}^{\infty}\frac{(\lambda_0 t)^k}{k!}\mathrm{E}_Q\left[\exp\left(iz\sum_{\ell=1}^{k}Q_\ell\right)\right] \\
&= e^{-\lambda_0 t}\sum_{k=0}^{\infty}\frac{(\lambda_0 t)^k}{k!}\prod_{\ell=1}^{k}\mathrm{E}_Q\left[\exp(izQ_\ell)\right] \\
&= e^{-\lambda_0 t}\sum_{k=0}^{\infty}\frac{(\lambda_0 t)^k}{k!}\mathrm{E}_Q^k\left[\exp(izQ)\right] = \exp\left(\lambda_0 t\left(\mathrm{E}_Q\left[\exp(izQ)\right]-1\right)\right),
\end{aligned}$$

using the iterated conditional expectation technique and IID, so the Lévy characteristic exponent, substituting $\mathrm{E}_Q\left[\exp(izQ)\right] = \mathcal{C}_Q(z)$*, is*

$$\eta_{CP(t)}(z) = \lambda_0(\mathcal{C}_Q(z)-1) \tag{13.60}$$

and the simple Poisson process exponent is recovered if $Q_\ell \overset{\text{w.p.o.}}{=} 1 \;\; \forall \ell \ge 1$.

- ***Vector Compound Poisson Process*** $\mathbf{CP}(t) = \sum_{\ell=1}^{P(t)}\mathbf{Q}_\ell$ ***on*** $\mathbb{R}^{n_x}$ ***with Constant Jump-Rate*** λ_0 ***and IID Vector Jump-Amplitudes*** $\mathbf{Q}_\ell$ ***with Distribution*** $\Phi_{\mathbf{Q}}(\mathbf{q})$***:*** *Note that the* $\mathbf{Q}_\ell$ *are IID as vectors not necessarily as* *components, thus,*

$$\begin{aligned}
\mathcal{C}_{\mathbf{CP}(t)}(\mathbf{z}) = \mathrm{E}_{P(t),\mathbf{Q}}\left[e^{i\mathbf{z}^\top \mathbf{CP}(t)}\right] &= e^{-\lambda_0 t}\sum_{k=0}^{\infty}\frac{(\lambda_0 t)^k}{k!}\mathrm{E}_{\mathbf{Q}}\left[\exp\left(i\mathbf{z}^\top\sum_{\ell=1}^{k}\mathbf{Q}_\ell\right)\right] \\
&= e^{-\lambda_0 t}\sum_{k=0}^{\infty}\frac{(\lambda_0 t)^k}{k!}\prod_{\ell=1}^{k}\mathrm{E}_{\mathbf{Q}}\left[\exp\left(i\mathbf{z}^\top\mathbf{Q}_\ell\right)\right] \\
&= e^{-\lambda_0 t}\sum_{k=0}^{\infty}\frac{(\lambda_0 t)^k}{k!}\mathrm{E}^k_{\mathbf{Q}}\left[\exp\left(i\mathbf{z}^\top\mathbf{Q}_\ell\right)\right] \\
&= \exp\bigl(\lambda_0 t\left(\mathrm{E}_{\mathbf{Q}}\left[\exp\left(i\mathbf{z}^\top\mathbf{Q}_\ell\right)\right]-1\right)\bigr),
\end{aligned}$$

using the iterated conditional expectation technique and IID again, so the Lévy characteristic exponent, substituting $\mathrm{E}_{\mathbf{Q}}\left[\exp(i\mathbf{z}^\top\mathbf{Q})\right] = \mathcal{C}_{\mathbf{Q}}(\mathbf{z})$*, is*

$$\eta_{\mathbf{CP}(t)}(z) = \lambda_0(\mathcal{C}_{\mathbf{Q}}(\mathbf{z})-1). \tag{13.61}$$

Lévy-Klintchine Jump-Diffusion Formula

In these examples, the ingredients for the fundamental theorem of the **Lévy-Klintchine representation formula** specialized to jump-diffusion processes has been derived, based on the vector Gaussian process exponent result in (13.56) and the vector compound Poisson process exponent result in (13.61).

Theorem 13.23. ***Lévy-Klintchine Formula for Jump-Diffusion Processes:*** *Let* $\mathbf{X}(t)$ *be the jump-diffusion process on* $\mathbb{R}^{n_x}$ *for* $t \geq 0$,

$$\mathbf{X}(t) = \mathbf{X}(0) + \boldsymbol{\mu}_0 t + \sigma_0\mathbf{W}(t) + \sum_{\ell=1}^{P(t)}\mathbf{Q}_\ell, \tag{13.62}$$

with ***Lévy characteristic triplet*** $(\sigma_0\sigma_0^\top, \lambda_0\Phi_{\mathbf{Q}}(d\mathbf{q})dt, \boldsymbol{\mu}_0)$*, where* $\boldsymbol{\mu}_0 \in \mathbb{R}^{n_x}$ *is a constant,* $\sigma_0 \in \mathbb{R}^{n_x\times n_w}$ *is a constant,* $\mathbf{W}(t) \in \mathbb{R}^{n_w}$ *is a vector Wiener process,* $\mathbf{P}(t) \in \mathbb{R}$ *is a simple Poisson process with constant and finite jump-rate* $\lambda_0 \in \mathbb{R}$ *and compounded with IID vector jump-amplitudes* $\mathbf{Q}_\ell \in \mathbb{R}^{n_x}$ *with distribution* $\Phi_{\mathbf{Q}}(\mathbf{q})$*. The random triplet* $(\mathbf{W}(t), P(t), Q)$ *are independent variables, except that the jump-amplitude* Q *requires the existence of a jump of the Poisson process.*

Then, the characteristic function with $\mathbf{z} \in \mathbb{R}^{n_x}$ *for the initial condition translated process*

$$\mathbf{Y}(t) \equiv \mathbf{X}(t) - \mathbf{X}(0) \tag{13.63}$$

is

$$\mathcal{C}_{\mathbf{Y}(t)}(\mathbf{z}) = \mathrm{E}_{\mathbf{Y}(t)}\left[\exp(i\mathbf{z}^\top\mathbf{Y}(t)\right] = \exp\left(t\eta_{\mathbf{Y}(t)}(\mathbf{z})\right),$$

where the Lévy characteristic exponent is

$$\eta_{\mathbf{Y}(t)}(\mathbf{z}) = i\boldsymbol{\mu}_0 t - \frac{1}{2}\mathbf{z}^\top \sigma_0 \sigma_0^\top \mathbf{z} + \lambda_0 \int_{\mathbb{R}^{n_x}} \left(\exp(i\mathbf{z}^\top \mathbf{q}) - 1\right) \phi_{\mathbf{Q}}(\mathbf{q}) d\mathbf{q}. \quad (13.64)$$

Except for the technical details, the Lévy characteristic exponent result (13.64) follows from (13.56) for $\mathbf{G}(t)$ and from (13.61) for $\mathbf{CP}(t)$ by the independence properties between $\mathbf{G}(t)$ and $(P(t), Q)$ and by iterative conditional expectation between $P(t)$ and Q that is conditioned on the existence of a jump as for (13.61). Thus,

$$\begin{aligned}
\mathcal{C}_{\mathbf{Y}(t)}(\mathbf{z}) &= \mathrm{E}_{\mathbf{Y}(t)}\left[\exp\left(i\mathbf{z}^\top \mathbf{Y}(t)\right)\right] \\
&= \mathrm{E}_{\mathbf{W}(t)}\left[\exp\left(i\mathbf{z}^\top \mathbf{G}(t)\right)\right] \cdot \mathrm{E}_{P(t),\mathbf{Q}}\left[\exp\left(i\mathbf{z}^\top \mathbf{CP}(t)\right)\right] \\
&= \mathcal{C}_{\mathbf{G}(t)}(\mathbf{z}) \cdot \mathcal{C}_{\mathbf{P}(t)}(\mathbf{z}) \\
&= \exp\left(t\eta_{\mathbf{G}(t)}(\mathbf{z})\right) \cdot \exp\left(t\eta_{\mathbf{CP}(t)}(\mathbf{z})\right) \\
&= \exp\left(t\left(\eta_{\mathbf{G}(t)}(\mathbf{z}) + \eta_{\mathbf{CP}(t)}(\mathbf{z})\right)\right)
\end{aligned}$$

so substituting (13.56) and (13.61) and expanding the expectations leads directly to the main result (13.64). It should be noted that embedded in this derivation is the semi-group property [12, 59] of the characteristic function in the case of constant coefficients.

In the case to the geometric or linear jump-diffusion process (5.35) with constant rate coefficients for $X(t) \in \mathbb{R}$ with SDE,

$$dX(t) = X(t)\left(\mu_0 dt + \sigma_0 dW(t) + \sum_{k=1}^{dP(t)} \left(e^{Q_k} - 1\right)\right),$$

the solution is exponential via a logarithmic change of variable technique,

$$X(t) = X(0)\exp\left((\mu_0 - \sigma_0^2/2)t + \sigma_0 W(t) + \sum_{k=1}^{P(t)} Q_k\right), \quad (13.65)$$

with $X(0) > 0$, is obviously not a Lévy process due to the exponential time-dependence, without further transformation:

Corollary 13.24. ***Lévy-Klintchine Transformed Geometric Jump-Diffusion Formula:***

Assuming the hypotheses of Th. 13.23, except that, $n_x = 1$, $n_w = 1$ *and that the Lévy characteristic triplet is* $(\sigma_0^2, \lambda_0\Phi_Q(dq)dt, \mu_0 - \sigma_0^2)$, *then the characteristic function with* $z \in \mathbb{R}$ *of the the logarithmic-translated process* $Y(t)$*.,*

$$Y(t) \equiv \ln(X(t)/X(0)), \quad (13.66)$$

2006/5/2
page 520

corresponding to the geometric process (13.65), is

$$\mathcal{C}_{Y(t)}(z) = \mathrm{E}_{Y(t)}[\exp(izY(t)] = \exp\left(t\eta_{Y(t)}(z)\right),$$

where the Lévy characteristic exponent is

$$\eta_{Y(t)}(z) = i(\mu_0 - \sigma_0^2/2)t - \frac{1}{2}\sigma_0^2 z^2 + \lambda_0 \int_{\mathbb{R}} (\exp(izq) - 1)\, \phi_Q(q)dq. \quad (13.67)$$

Lévy-Klintchine Lévy Process Formula including Infinite Rate Processes

So far the jump-rate λ_0 has been assumed to be constant and either explicitly or implicitly finite in this Subsection on Lévy processes. However, the infinite jump-rates is a distinguishing feature of Lévy processes, so that, in general, it is not valid to write the jump-rate symbol λ_0 in Lévy process formulas. Instead, it is necessary to refer to the number of jumps rather than to the jump-rate.

Recall the definition (0.176) on page 64 of the **jump function** of a process:

$$[\mathbf{X}](t) \equiv \mathbf{X}(t) - \mathbf{X}(t-),$$

written here for RCLL vector processes (caution: in some of the literature $\Delta\mathbf{X}(t)$ is used, but can be confused with the analytic or numerical time increment, $\Delta\mathbf{X}(t) \equiv \mathbf{X}(t+\Delta t) - \mathbf{X}(t)$). At points where $\mathbf{X}(t)$ is continuous, then $[\mathbf{X}](t) = 0$.

Definition 13.25. ***Number of Jumps of a Process, Poisson Random Measure and Lévy Measure:*** *The number of jumps in the open set $\mathcal{S}$, assuming a bounded number of jumps and excluding zero jumps ($0 \notin \mathcal{S}$) on the interval $(0,t]$, is*

$$\mathcal{P}((0,t],\mathcal{S}) = \sum_{s\in(0,t]} \mathbf{1}_{[\mathbf{X}](s)\in\mathcal{S}}. \quad (13.68)$$

Here, $\mathcal{P}((0,t],\mathcal{S})$ is the Poisson random or jump measure [219]. The differential form is denoted by $\mathcal{P}(dt,d\mathbf{q}) = \mathcal{P}((t,t+dt,(\mathbf{q},\mathbf{q}+d\mathbf{q}])$, as previous used in Chapt. 5. An alternate form [228] uses a sequence of stopping or jump times,

$$T_{k+1}(\mathcal{S}) = \inf\{t \mid t > T_k(\mathcal{S}), [X](t) \in \mathcal{S}\};\ T_0(\mathcal{S}) \equiv 0,$$

such that

$$\mathcal{P}((0,t],\mathcal{S}) = \sum_{k=1}^{\infty} \mathbf{1}_{T_k(\mathcal{S})\le t}.$$

The zero-mean (centered or Martingale) form is denoted by

$$\widetilde{\mathcal{P}}(dt,\mathbf{q}) = \mathcal{P}(dt,d\mathbf{q}) - \nu^{(L)}(\mathbf{dq})dt, \quad (13.69)$$

*where now $\nu^{(L)}(\mathbf{dq})dt = \mathrm{E}[\mathcal{P}(dt,d\mathbf{q})]$ and $\nu^{(L)}$ is called the **Lévy measure** in general .*

DRAFT

2006/5/24
page 521

A fundamental tool for separating out the large jumps in the presence of infinite jump-rates is the following decomposition after the concise form of Øksendal and Sulem [219]:

Theorem 13.26. ***Lévy-Itô Decompsition:*** *Let $0 \le R < \infty$ be a jump-amplitude cutoff, then a Lévy process $\mathbf{X}^{(L)}(t)$ on $\mathbb{R}^{n_x}$ has the decomposition,*

$$\mathbf{X}^{(L)}(t) = \widetilde{\boldsymbol{\mu}}_{0,R}t + \sigma_0\mathbf{W}(t) + \int_{|\mathbf{q}|<R} \mathbf{q}\widetilde{\mathcal{P}}(t, d\mathbf{q}) + \int_{|\mathbf{q}|\ge R} \mathbf{q}\mathcal{P}(t, d\mathbf{q}), \quad (13.70)$$

where $\mathbf{W}(t) \in \mathbb{R}^{n_w}$ is an independent vector Wiener process, $\widetilde{\boldsymbol{\mu}}_{0,R} \in \mathbb{R}^{n_x}$ is a constant adjusted with R from the original drift $\boldsymbol{\mu}_0 \in \mathbb{R}^{n_x}$, $\sigma_0 \in \mathbb{R}^{n_x\times n_w}$ is a constant .

In particular, the Lévy-Itô decomposition states that the Lévy process is, as is the jump-diffusion, decomposable into a continuous process and a discontinuous process,

$$\begin{aligned}
\mathbf{X}^{(L)}(t) &= \mathbf{X}^{(cont)}(t) + \mathbf{X}^{(discont)}(t);\\
\mathbf{X}^{(cont)}(t)) &= \widetilde{\boldsymbol{\mu}}_{0,R}t + \sigma_0\mathbf{W}(t);\\
\mathbf{X}^{(discont)}(t)) &= \mathbf{X}^{(L)}(t) - \mathbf{X}^{(cont)}(t).
\end{aligned}$$

One consequence of this Lévy-Itô decomposition is another fundamental result [219, 228]:

Theorem 13.27. ***Lévy-Klintchine Representation Formula for Lévy Processes:*** *Let $\mathbf{X}^{(L)}(t)$ be a Lévy process for $t \ge 0$ with Lévy measure $\nu^{(L)}$ on $\mathbb{R}^{n_x}$, given constants $\widetilde{\boldsymbol{\mu}}_{0,R} \in \mathbb{R}^{n_x}$ and $\sigma_0 \in \mathbb{R}^{n_x\times n_w}$, then the jump-count satisfies*

$$\int_{\mathbb{R}^{n_x}} \min(|\mathbf{q}|^2, R)\nu^{(L)}(d\mathbf{q}) < \infty$$

and the characteristic function on $\mathbf{z} \in \mathbb{R}^{n_x}$ for $\mathbf{X}(t) = \mathbf{X}^{(L)}(t)$ is

$$\mathcal{C}_{\mathbf{X}(t)}(\mathbf{z}) = \mathrm{E}_{\mathbf{X}(t)}\left[\exp\left(i\mathbf{z}^\top\mathbf{X}(t)\right)\right] = \exp\left(t\eta_{\mathbf{X}(t)}(\mathbf{z})\right),$$

where the Lévy characteristic exponent is

$$\begin{aligned}
\eta_{\mathbf{X}(t)}(\mathbf{z}) = {} & i\widetilde{\boldsymbol{\mu}}_{0,R}\, t - \tfrac{1}{2}\mathbf{z}^\top\sigma_0\sigma_0^\top\mathbf{z}\\
&+ \textstyle\int_{|\mathbf{q}|<R} \left(\exp\left(i\mathbf{z}^\top\mathbf{q}\right) - 1 - i\mathbf{z}^\top\mathbf{q}\right)\nu^{(L)}(d\mathbf{q})\\
&+ \textstyle\int_{|\mathbf{q}|\ge R} \left(\exp\left(i\mathbf{z}^\top\mathbf{q}\right) - 1\right)\nu^{(L)}(d\mathbf{q}).
\end{aligned} \quad (13.71)$$

Conversely, given constants $\widetilde{\boldsymbol{\mu}}_{0,R} \in \mathbb{R}^{n_x}$ and $\sigma_0 \in \mathbb{R}^{n_x\times n_w}$, along with the Lévy measure $\nu^{(L)}$ on $\mathbb{R}^{n_x}$ such that the jump-count satisfies

$$\int_{\mathbb{R}^{n_x}} \min(|\mathbf{q}|^2, R)\nu^{(L)}(d\mathbf{q}) < \infty,$$

then there exists a Lévy process $\mathbf{X}^{(L)}$ *that is unique in distribution such that the Lévy characteristic is (13.71) for* $\mathbf{z} \in \mathbb{R}^{n_x}$.

Note that the extra linear term $i\mathbf{z}^\top\mathbf{q}$ in the first or inner integral of (13.71) is related to the zero-mean Poisson $\widetilde{P}(t)$ form iz found in (13.58) but not in (13.57) for $P(t)$.

Although jump processes time-dependent coefficients, like drift and volatility coefficients, do not strictly satisfy the stationary increment condition (13.49) for a Lévy process, Øksendal and Sulem [219] define **Lévy-driven processes** which satisfy the Lévy-Itô decomposition formula (13.70), but not the constant coefficient condition. For example, analogous to the **Wiener-driven Itô process** (13.44), there is the **Lévy-driven Itô-Lévy process** [219, Th. 1.14, p. 6] on $\mathbb{R}$ with time-random coefficients,

$$\begin{aligned} dX(t) = {} & \widetilde{\mu}_{0,R}(t;\omega)dt + \sigma_0(t;\omega)dW(t) \\ & + \int_{|q|<R} h(t,q;\omega)\widetilde{\mathcal{P}}(dt,dq) + \int_{|q|\geq R} h(t,q;\omega)\mathcal{P}(dt,dq), \end{aligned} \tag{13.72}$$

for some $R \in [0,\infty)$, $(\widetilde{\mu}_{0,R}(t;\omega), \sigma_0(t;\omega), h(t,q\omega)$ are integrable function and ω is some background random variable.

The **Lévy-driven geometric Lévy process** [219, Example 1.15, p. 7] is similarly defined,

$$\begin{aligned} dX(t) = {} & X(t)\Big(\widetilde{\mu}_{0,R}(t;\omega)dt + \sigma_0(t;\omega)dW(t) \\ & + \int_{|q|<R} h(t,q;\omega)\widetilde{\mathcal{P}}(dt,dq) + \int_{|q|\geq R} h(t,q;\omega)\mathcal{P}(dt,dq)\Big), \end{aligned} \tag{13.73}$$

where, in addition, the jump-amplitude $h(t,q;\omega) \geq -1$ to preserve positivity assuming $X(0) > 0$, with more potential uses in financial applications.

In general, these processes are special cases of what Øksendal and Sulem [219, Th. 1.19,p. 10] call **Lévy diffusions** governed by **Lévy stochastic differential equations**,

$$\begin{aligned} d\mathbf{X}(t) = {} & \widetilde{\mu}(t,\mathbf{X}(t))dt + \sigma(t,\mathbf{X}(t))d\mathbf{W}(t) \\ & + \int_{|q|<R} h(t,q;\omega)\widetilde{\mathcal{P}}(dt,dq) + \int_{|q|\geq R} h(t,\mathbf{X}(t),q)\mathcal{P}(dt,dq), \end{aligned} \tag{13.74}$$

where $0 \leq t \leq T$, $\mathbf{X} \in \mathbb{R}^{n_x}$, $\widetilde{\mu} \in \mathbb{R}^{n_x}$, $\mathbf{W} \in \mathbb{R}^{n_w}$, $\sigma \in \mathbb{R}^{n_x \times n_w}$, $\mathcal{P} \in \mathbb{R}^{n_p}$, $\mathbf{Q} \in \mathbb{R}^{n_x}$ and $h \in \mathbb{R}^{n_x \times n_p}$, subject to the usual linear growth and Lipschitz continuity conditions.

For many other Lévy process models, including models which push the limits of the assumptions here, see Applebaum [12, Subsect. 5.4.7, p. 286ff]

Concluding this subsection like the last, the size of the Lévy processes is compared to that of jump-diffusions. According to the strict Lévy process definition leading to a restriction to constant coefficients,

$$\left\{ \begin{matrix} \textbf{\textit{constant coefficient}} \\ \textbf{\textit{jump-diffusion processes}} \end{matrix} \right\} \subset \textbf{\textit{Lévy processes}}, \tag{13.75}$$

since ordinarily jump-diffusions based upon Poisson processes do not allow for infinite jump-rates on $[0,t]$. However, if the infinite jump activity is controlled for, then

$$\left\{ \begin{array}{c} \textbf{\textit{finite jump-rate}} \\ \textbf{\textit{Lévy processes}} \end{array} \right\} \subset \textbf{\textit{jump-diffusion processes}}, \tag{13.76}$$

since jump-diffusions in general include variable coefficients and nonlinear terms.

If the comparison is made to the Lévy-driven processes discussed by Øksendal and Sulem [219] and summarized here, then

$$\{\textbf{\textit{jump-diffusion processes}}\} \subset \textbf{\textit{Lévy-driven processes}}, \tag{13.77}$$

due to the inclusion of infinite jump-rates with nonlinear and time-dependent coefficients in Lévy-driven processes.

Suggested References for Further Reading

- Applebaum, 2004 [12].
- Bain, 2006 [17].
- Baxter and Rennie, 1996 [22].
- Billingsley, 1986 Billingsley86.
- Bingham and Kiesel, 2004 [33]
- Bossaerts, 2002 [41].
- Brémaud, 1981 [43].
- Cont and Tankov, 2004 [59].
- Cyganowski, Kloeden and Ombach, 2002 [66].
- Doob, 1953 [69].
- Duffie, 1992 [74].
- Gihman and Skorohod, 1972 [94].
- Glasserman, 2003 [96].
- Heath and Schweizer, 2000 [132].
- Hull, 2000 [144].
- Itô, 1951 [146].
- Karlin and Taylor, 1981 [158].
- Klebaner, 1998 [160].

DRAFT

2006/5/24
page 524

- Mikosch, 1998 [205].
- Neftci, 2000 [213].
- Øksendal, 1998 [218].
- Øksendal and Sulem, 2005 [219].
- Pliska, 1997 [221].
- Protter, 2004 [228].
- Runggaldier, 2003 [235].
- Rogers and Williams, 2000 [232].
- Steele, 2001 [251].
- Yong and Zhou, 1999 [283].

DRAFT