5

Martingales and Stochastic Integrals

In this chapter we consider a class of continuous stochastic processes, called martingales, which play a central role in finance. We also define the gains realized from trading as a stochastic integral. Stochastic integration and martingales provide key tools for the analysis of the continuous time evolution of financial markets.

5.1 Martingales

One of the fundamental concepts in modern finance is the notion of a martingale. This is a stochastic process that, with its last observed value, provides the best forecast for its future values. Martingales exhibit the property of having no systematic trends in their dynamics. It is obvious that financial quantities, such as asset prices, are driven primarily by information. Forecasting a quantity, for example, the value of a derivative price when expressed in units of the market portfolio, is strongly dependent on the information that is available at the present time. This forces one to use a detailed notion for the information structure related to the evolution of the underlying stochastic processes.

Information Sets and Filtrations

On a given probability space $(\Omega, \mathcal{A}, P)$, as introduced in Sect. 1.1, let us consider a financial market model that is based on the observation of a continuous time stochastic vector process $\boldsymbol{X} = \{\boldsymbol{X}_t \in \Re^n,\, t \in [0,\infty)\},\ n \in \mathcal{N}$, typically expressing asset price processes. We denote by $\hat{\mathcal{A}}_t$ the time t *information set*, which is the sigma-algebra of events that are known to the market participants at time $t \in [0,\infty)$. Our interpretation of $\hat{\mathcal{A}}_t$ is that it represents the information obtained from the values of the vector process $\boldsymbol{X}$ up to time t. More precisely, it is the sigma-algebra

E. Platen, D. Heath, *A Benchmark Approach to Quantitative Finance*,
Springer Finance,

$$\hat{\mathcal{A}}_t = \sigma\{\boldsymbol{X}_s : s \in [0,t]\}$$

generated from all observations of $\boldsymbol{X}$ in the market up to time t. In a general financial market model the components of $\boldsymbol{X}$ could include diverse quantities, for instance, security prices, interest rates, indicators for certain political events, market activity, corporate data, employment figures, insurance claims, balance sheets of companies or trade balances.

Assuming that information is not lost, then the increasing family

$$\underline{\hat{\mathcal{A}}} = \{\hat{\mathcal{A}}_t,\, t \in [0,\infty)\}$$

of information sets $\hat{\mathcal{A}}_t$, which are sub-sigma-algebras of $\hat{\mathcal{A}}_\infty$ satisfy, for any sequence $0 \le t_1 < t_2 < \ldots < \infty$ of observation times, the relation $\hat{\mathcal{A}}_{t_1} \subseteq \hat{\mathcal{A}}_{t_2} \subseteq \ldots \subseteq \hat{\mathcal{A}}_\infty = \cup_{t\in[0,\infty)}\hat{\mathcal{A}}_t$.

Furthermore, to avoid technical subtleties, we introduce the information set $\mathcal{A}_t$ as the *augmented* sigma-algebra of $\hat{\mathcal{A}}_t$ for each $t \in [0,\infty)$. This means that it is augmented by every null set in $\hat{\mathcal{A}}_\infty$ such that it belongs to $\mathcal{A}_0$, and so to each $\mathcal{A}_t$. We define $\mathcal{A}_{t+} = \cap_{\varepsilon>0}\mathcal{A}_{t+\varepsilon}$ to be the sigma-algebra of events immediately after $t \in [0,\infty)$. We say that the family $\underline{\mathcal{A}} = \{\mathcal{A}_t,\, t \in [0,\infty)\}$ is *right continuous* if $\mathcal{A}_t = \mathcal{A}_{t+}$ holds for every $t \in [0,\infty)$. Such a right-continuous family $\underline{\mathcal{A}} = \{\mathcal{A}_t,\, t \in [0,\infty)\}$ of information sets we call a *filtration*. Thus, a filtration models the evolution of information as it becomes available over time. For simplicity, we define $\mathcal{A}$ as the smallest sigma-algebra that contains $\mathcal{A}_\infty = \cup_{t\in[0,\infty)}\mathcal{A}_t$.

The above technical assumptions allow convenient mathematical derivations and do not restrict our practical modeling potential. From now on, if not stated otherwise, we shall assume a *filtered probability space* $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$ to be given, where the filtration $\underline{\mathcal{A}}$ characterizes the evolution of the corresponding information. The capturing of the evolution of this information is essential for the modeling of financial markets since it is information that drives most of its dynamics.

Any given stochastic process $Y = \{Y_t,\, t \in [0,\infty)\}$ generates a filtration $\mathcal{A}^Y = \{\mathcal{A}_t^Y,\, t \in [0,\infty)\}$. Here $\mathcal{A}_t^Y = \sigma\{Y_s : s \in [0,t]\}$ is the information set, that is the sigma-algebra, generated by Y up to time t. This information set can be interpreted as a complete record of all movements of the process Y up until time t. $\mathcal{A}^Y$ is also called the *natural filtration* for the process Y. For a given model with a vector process $\boldsymbol{X}$ that describes the total evolution of the model and, thus, the corresponding increasing family of information sets, we write $\underline{\mathcal{A}} = \underline{\mathcal{A}}^{\boldsymbol{X}}$ and set $\mathcal{A}_t = \mathcal{A}_t^{\boldsymbol{X}}$, similarly as above.

If for a process $Z = \{Z_t,\, t \in [0,\infty)\}$ and each time $t \in [0,\infty)$ the random variable Z_t is $\mathcal{A}_t^{\boldsymbol{X}}$-measurable, then Z is called *adapted* to $\underline{\mathcal{A}}^{\boldsymbol{X}} = \{\mathcal{A}_t^{\boldsymbol{X}},\, t \in [0,\infty)\}$. In intuitive terms this means that the history of the process Z until time t is covered by the information set $\mathcal{A}_t^{\boldsymbol{X}}$. As a consequence, for an $\underline{\mathcal{A}}^{\boldsymbol{X}}$-adapted process Z the value Z_t is known, given the information set $\mathcal{A}_t^{\boldsymbol{X}}$ up to and including time t. We mention that the completeness of the information set $\mathcal{A}_t^{\boldsymbol{X}}$, which includes all null events, allows us to conclude that for two random variables Z_1 and Z_2, where $Z_1 = Z_2$ a.s. and Z_1 is $\mathcal{A}_t^{\boldsymbol{X}}$ measurable, Z_2 is also $\mathcal{A}_t^{\boldsymbol{X}}$-measurable.

If the process $\boldsymbol{X}$ is Markovian, then the relevant information needed to determine properties of its future values reduces to the knowledge of the value $\boldsymbol{X}_t$ at the present time t. This makes it possible to express and store the relevant information in a compact form. It also highlights the importance of Markovianity for the tractability of a wide range of financial models.

In financial modeling we shall typically use later a filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$, where the sources of continuous uncertainty are independent standard Wiener processes $W^1, W^2, \ldots, W^m$ and the sources of event driven uncertainty are independent Poisson processes $N^{m+1}, N^{m+2}, \ldots, N^d$, $d \in \{1, 2, \ldots\}$, $m \in \{1, 2, \ldots, d\}$. We shall always assume that these Wiener and Poisson processes are $\underline{\mathcal{A}}$-adapted and that their increments $(W_t^j - W_s^j)$ are independent of $\mathcal{A}_s$, see (1.1.16), for $t \in [0, \infty)$, $s \in [0, t]$ and $j \in \{1, 2, \ldots, m\}$. We call then $\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, W_t^2, \ldots, W_t^m)^\top,\ t \in [0, \infty)\}$ an m-dimensional standard Wiener process on $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$ or an $(\underline{\mathcal{A}}, P)$-Wiener process.

Continuous Time Martingales

In financial markets, investors have to determine best estimates for the actual value of future payoffs. If they were to use different information sets, then they might generate different value estimates. For simplicity, let us assume that they all use the same information sets. Furthermore, a value estimate needs to be based on a corresponding *benchmark* or *numeraire*, which provides the units in which the investor formulates his or her best estimates. We shall later discuss cases where one uses the savings account or the market portfolio as numeraire. Finally, an investor has also to employ a probability measure for forming some expectation when identifying the best estimate, as we shall see below. Let us use the numeraire for which it is appropriate to form an expectation under the real world probability measure when searching for the best estimate of a future payoff. We shall see later that an appropriate numeraire is the market portfolio. More generally, given an information set, a probability measure and a numeraire, we shall ask what is at present the best estimate for the value of a future cash flow or payoff.

To answer this question in a mathematically precise manner we define the quantity F_s for $s \in [0, \infty)$ as the least-squares estimate, see (1.3.72), of the future value X_t at the later time $t \in [s, \infty)$ under the information given by $\mathcal{A}_s$. This best estimate is $\mathcal{A}_s$-measurable and minimizes the expected least-squares error

$$\varepsilon_s = E\left((X_t - F_s)^2\right)$$

over all possible $\mathcal{A}_s$-measurable estimates. Note that we need here to assume that X_t is square integrable, see (1.3.7). The random variable F_s is simply the least-squares projection of X_t given the information at time $s \in [0, t]$. It is obtained by the conditional expectation

$$F_s = E(X_t \,\big|\, \mathcal{A}_s), \tag{5.1.1}$$

for all $s \in [0, t]$.

In a price system a candidate for a reasonable price at time $s \in [0, t]$ for the future value X_t at time t is the least-squares estimate F_s that can be formed on the basis of the information contained in $\mathcal{A}_s$. This means, one obtains realistic prices when setting $X_s = F_s$ by forming the price process $X = \{X_t,\, t \in [0, \infty)\}$ which satisfies the conditional expectation

$$X_s = E(X_t \,\big|\, \mathcal{A}_s) \tag{5.1.2}$$

for all $s \in [0, t]$ and $t \in [0, \infty)$.

Definition 5.1.1. *We call a continuous time stochastic process $X = \{X_t, t \in [0, \infty)\}$, which satisfies the property (5.1.2) and the integrability condition*

$$E(|X_t|) < \infty \tag{5.1.3}$$

for all $t \in [0, \infty)$, a martingale *or more precisely an $(\underline{\mathcal{A}}, P)$*-martingale*.*

If for a martingale X in addition the random variable X_t is square integrable for all $t \in [0, \infty)$, that is

$$E\left(|X_t|^2\right) < \infty \tag{5.1.4}$$

for all $t \in [0, \infty)$, then we call X a *square integrable martingale*. Note by (5.1.1) and (5.1.2) that for a square integrable martingale the least-squares estimate of its future values is always given by its last available observation.

A martingale is defined with respect to a given filtration $\underline{\mathcal{A}}$, which denotes the family of relevant information sets, and a probability measure P, which expresses the likelihood of events. The conditional expectation is then taken under P. Since both ingredients are essential we shall call a martingale an $(\underline{\mathcal{A}}, P)$-*martingale* if it is defined with respect to the filtration $\underline{\mathcal{A}}$ and the probability measure P. This is sometimes important because it is not always clear from the context which filtration and probability measure are chosen. If one changes the filtration $\underline{\mathcal{A}}$ or the probability measure P, then a given martingale will usually no longer remain a martingale.

The martingale relation (5.1.2) is fundamental in finance, in particular, in derivative pricing. Under the benchmark approach we shall ask later derivative prices, when expressed in units of the benchmark, to form martingales. Different pricing rules are obtained by selecting different reference units or numeraires, an issue that will be discussed later in detail.

Examples of Martingales

As an example of a continuous time martingale, let us consider a Wiener process $W = \{W_t,\, t \in [0, \infty)\}$ on a filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$.

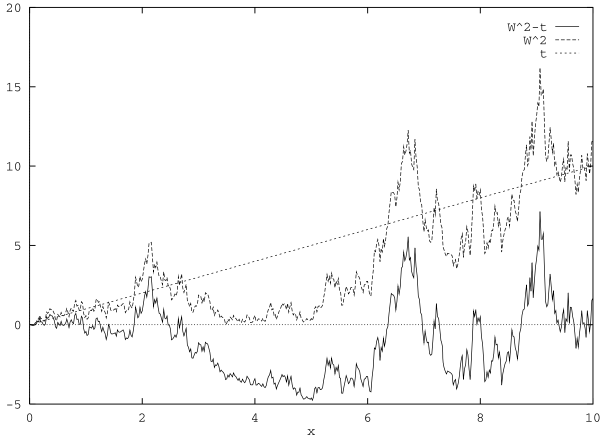

Fig. 5.1.1. Paths of $X_t = W_t^2 - t$, W_t^2 and t

Here, as previously mentioned, we assume that $W_{t+h} - W_t$ is independent of $\mathcal{A}_t$ for all $t \in [0, \infty)$ and $h \in [0, \infty)$. Furthermore, the natural filtration $\mathcal{A}^W$ of W is such that $\mathcal{A}_t^W \subseteq \mathcal{A}_t$ for each $t \in [0, \infty)$.

Note that W is $\underline{\mathcal{A}}$-adapted, which means that W_t is $\mathcal{A}_t$-measurable for $t \in [0, \infty)$. We can show by the linearity and independence properties of conditional expectations, see (1.3.69) and (1.3.67), that

$$\begin{aligned} E(W_t \,|\, \mathcal{A}_s) &= E\left(W_t - W_s \,\middle|\, \mathcal{A}_s\right) + E\left(W_s \,\middle|\, \mathcal{A}_s\right) \\ &= E\left(W_t - W_s\right) + E\left(W_s \,\middle|\, \mathcal{A}_s\right) \\ &= W_s \end{aligned} \tag{5.1.5}$$

for $s \in [0, \infty)$ and $t \in [s, \infty)$. From (5.1.5) it follows by Definition 5.1.1 that the above Wiener process W is a martingale, or more precisely an $(\underline{\mathcal{A}}, P)$-martingale.

There are many other continuous time stochastic processes that form martingales. For example, using again the standard Wiener process W it can be demonstrated that the process

$$X = \left\{X_t = W_t^2 - t,\, t \in [0, \infty)\right\} \tag{5.1.6}$$

is an $(\underline{\mathcal{A}}, P)$-martingale. In Fig. 5.1.1 we show a typical path for this process together with W_t^2 and t.

The process

$$\bar{X} = \left\{\bar{X}_t = \exp\left\{\sigma W_t - \frac{1}{2}\,\sigma^2 t\right\},\, t \in [0, \infty)\right\},$$

which is an exponential of a transformed Wiener process, is also an $(\underline{\mathcal{A}}, P)$-martingale. Note that this is a specific geometric Brownian motion with

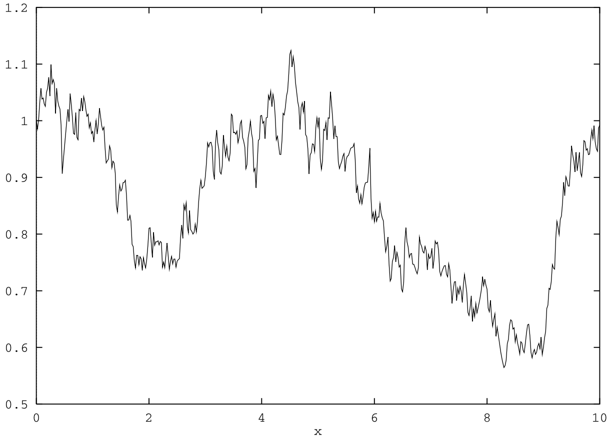

Fig. 5.1.2. Path of $\bar{X}_t = \exp\left\{\sigma W_t - \frac{1}{2}\,\sigma^2 t\right\}$

volatility σ, negative growth rate $\mu = -\frac{1}{2}\sigma^2$ and initial value $\bar{X}_0 = 1$. Figure 5.1.2 displays a sample path for this process with volatility $\sigma = 0.2$.

Super- and Submartingales

In practice asset prices are usually not completely trendless. For instance, the price of a zero coupon bond, which pays one dollar at a fixed maturity date, increases on average over time until it reaches at maturity the value one. These types of systematically trending stochastic processes are captured by the following definition of super- and submartingales.

Definition 5.1.2. *One calls an $\underline{\mathcal{A}}$-adapted process $X = \{X_t,\, t \in [0,\infty)\}$ an $(\underline{\mathcal{A}}, P)$-*supermartingale (submartingale) *if*

$$X_s \overset{(\leq)}{\geq} E\left(X_t \,\big|\, \mathcal{A}_s\right) \tag{5.1.7}$$

and

$$E(|X_t|) < \infty \tag{5.1.8}$$

for $s \in [0,\infty)$ and $t \in [s,\infty)$.

This means, on average, a supermartingale (submartingale) decreases (increases) its value over time. In comparison with a martingale the equality in (5.1.2) is replaced by the inequality (5.1.7). We call a supermartingale (submartingale) a *strict supermartingale* (submartingale) if the inequality in (5.1.7) is a strict inequality.

As an example for a submartingale we show in Fig. 5.1.3 for some geometric Brownian motion with

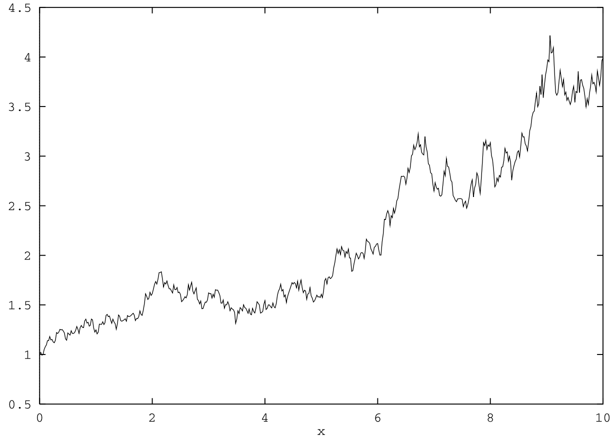

Fig. 5.1.3. Path of X_t for a submartingale

$$X_t = \exp\left\{\left(r - \frac{1}{2}\,\sigma^2\right) t + \sigma\, W_t\right\} \tag{5.1.9}$$

a sample path over a period of ten years with expected rate of return $r = 0.05$ and volatility $\sigma = 0.2$. This example illustrates some features that are typical for asset price scenarios. They seem to exhibit larger fluctuations for larger asset price values, as is the case for the S&P500 index shown in Fig. 3.1.1. If the submartingale X is discounted by the process $B = \{B_t = \exp\{r\,t\},\, t \in [0,\infty)\}$, which is simply a savings account with continuously compounding constant interest rate $r > 0$, then the discounted process $\bar{X} = \{\bar{X}_t = \frac{X_t}{B_t},\, t \in [0,\infty)\}$ is a martingale. Let us mention that Fig. 5.1.2 displays the sample path for $\bar{X}_t$, where X_t is shown in Fig. 5.1.3.

As we shall see later in Chaps. 9 to 14, in financial market models supermartingales play a natural role. They appear when securities are expressed in units of a particular benchmark, which is the, so-called, *growth optimal portfolio* (GOP). This is the portfolio that maximizes the expected logarithm of its value at future dates, see Kelly (1956), Long (1990). By interpreting a diversified market index as the GOP it has been suggested in Platen (2004c) that the savings account B, when expressed in units of the market index should be modeled to form a strict supermartingale and not a martingale, as the classical risk neutral theory assumes, and will be explained in Chap. 9.

Compensated Poisson Process

In Fig. 3.5.1 we plotted the path of a Poisson process $N = \{N_t,\, t \in [0,\infty)\}$ with intensity $\lambda > 0$, see Definition 3.5.1. We have assumed for any Poisson process N that N is $\underline{\mathcal{A}}$-adapted and such that for $t \in [0,\infty)$ and $h \in [0, T-t]$ the $\mathcal{A}_{t+h}$-measurable random variable $N_{t+h} - N_t$ is independent of $\mathcal{A}_t$. We can then show for $0 \le s < t < \infty$ by using (3.5.2) that

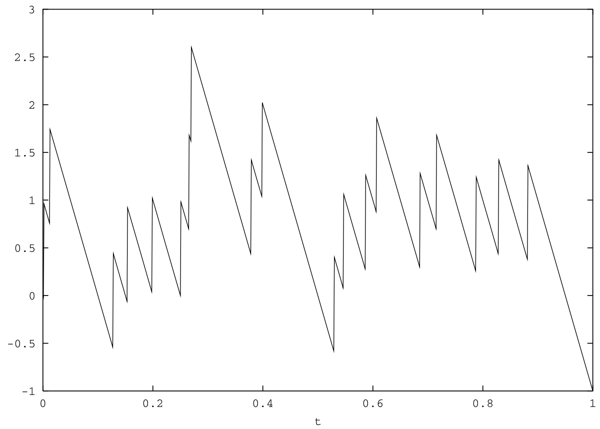

Fig. 5.1.4. Path of a compensated Poisson process

$$E\left(N_t \,\middle|\, \mathcal{A}_s\right) = E\left(N_t - N_s \,\middle|\, \mathcal{A}_s\right) + N_s = E\left(N_t - N_s\right) + N_s$$
$$= \lambda(t-s) + N_s \geq N_s, \tag{5.1.10}$$

which proves that the Poisson process is a submartingale.

On the other hand, the *compensated Poisson process* $q = \{q_t,\, t \in [0,\infty)\}$ with

$$q_t = N_t - \lambda\, t \tag{5.1.11}$$

is a martingale since we have by similar arguments as in (5.1.10)

$$E\left(q_t \,\middle|\, \mathcal{A}_s\right) = E\left(q_t - q_s \,\middle|\, \mathcal{A}_s\right) + q_s$$
$$= E\left(N_t - N_s\right) - \lambda\,(t-s) + q_s = q_s \tag{5.1.12}$$

for $0 \leq s \leq t < \infty$. In Fig. 5.1.4 we plot the path of a compensated Poisson process q with intensity $\lambda = 20$, where Fig. 3.5.1 shows the corresponding trajectory of the Poisson process N.

Stopping Times

Random times naturally appear in financial and insurance applications, for instance, as time of default of a company. We refer to Sect. 3.7 for an insurance example. Also the first hitting time of a critical barrier by an underlying asset price is a random time. Since the information structure is essential in stochastic modeling such random times have to be properly defined.

This leads us to the notion of stopping times. Let us consider a filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$ as introduced above.

Definition 5.1.3. *A random variable* $\tau : \Omega \to [0,\infty)$ *is called a* stopping time *with respect to the filtration* $\underline{\mathcal{A}}$ *if for all* $t \in [0,\infty)$

$$\{\tau \leq t\} \in \mathcal{A}_t. \tag{5.1.13}$$

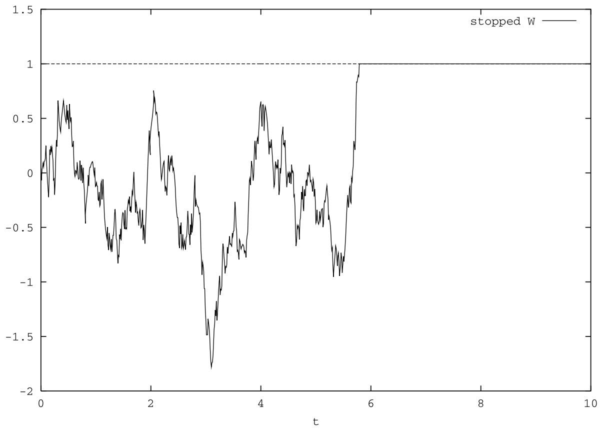

Fig. 5.1.5. First hitting time of a Wiener path

The relation (5.1.13) means that for all $\omega \in \Omega$ the event $\tau \leq t$ is in $\mathcal{A}_t$, which expresses the fact that it is $\mathcal{A}_t$-measurable and thus observable at time t. The information set, that is, the sigma-algebra associated with a stopping time τ is defined as

$$\mathcal{A}_\tau = \sigma\left\{A \in \mathcal{A} :\ A \cap \{\tau \leq t\} \in \mathcal{A}_t \quad \text{for} \quad t \in [0, \infty)\right\}. \tag{5.1.14}$$

It represents the information available before and at the stopping time τ. For instance, the kth jump time τ_k of a Poisson process N, as defined in Sect. 3.5, is a stopping time. This could be the time when the kth company collapses in a given year. One can show that a counting process is adapted if and only if the associated jump times are stopping times.

The first time

$$\tau(a) = \inf\{t \geq 0 :\ W_t = a\} \tag{5.1.15}$$

when a Wiener process W reaches a level $a \in \Re$ is a stopping time. In Fig. 5.1.5 we display the first time $\tau(1.0) \approx 5.8$ of a Wiener path hitting the level $a = 1.0$. Similarly, the default time of a company is a stopping time.

Predictable Processes

The allocation of assets in a portfolio can, in practice, only be performed in a predictable way. That means, the investor has to decide in advance what allocation will be pursued. To make this notion of predictability precise for stopping times, we call a sigma-algebra *predictable* when it is generated by left-continuous $\underline{\mathcal{A}}$-adapted processes with right hand limits. Roughly speaking, we exclude in a predictable sigma-algebra all information about the time instant when a sudden not predictable event, like a default, occurs. Note however, immediately after the event a predictable sigma-algebra already contains also this information. A stochastic process $X = \{X_t,\, t \in [0, \infty)\}$, where X_τ is for each stopping time τ measurable with respect to a predictable sigma-algebra, is called *predictable*. For instance, all continuous stochastic processes are predictable. From a right-continuous process with left hand limits $X = \{X_t,\, t \in [0,\infty)\}$ we obtain its predictable version $\tilde{X} = \{\tilde{X}_t,\, t \in [0,\infty)\}$ by taking at each time point the left hand limit, that is

$$\tilde{X}_t = X_{t-} \tag{5.1.16}$$

for all $t \in [0,\infty)$. Later when we form stochastic integrals we shall typically request that the integrands are predictable processes. In the case when a given potential integrand is not predictable, then its left-continuous version is chosen as integrand. This is similar to the natural request that an investor has to decide about his or her portfolio allocation of stocks at the beginning of any trading period and cannot revise it afterwards.

A stopping time is called *predictable*, if $\mathcal{A}_\tau$ is predictable. This means, $\mathcal{A}_\tau$ is generated by left-continuous stochastic processes with right hand limits. A stopping time that is not predictable is called *inaccessible*. The jump times of a Poisson process are inaccessible. Here $\mathcal{A}_\tau$ cannot be generated by left-continuous processes. However, the first hitting time $\tau(a)$ of the continuous Wiener process W, given in (5.1.15), is predictable.

Properties of Stopping Times (*)

For $a, b \in \Re$ we employ the notation $a \wedge b = \min(a,b)$ and $a \vee b = \max(a,b)$. One can derive the following useful properties of stopping times τ and τ', see Karatzas & Shreve (1991) and Elliott (1982).

(i) τ is $\mathcal{A}_\tau$-measurable.
(ii) For a continuous $\underline{\mathcal{A}}$-adapted process $X = \{X_t,\, t \in [0,\infty)\}$ the random variable X_τ is $\mathcal{A}_\tau$-measurable.
(iii) If $P(\tau \le \tau') = 1$, then $\mathcal{A}_\tau \subseteq \mathcal{A}_{\tau'}$.
(iv) The random variables $\tau \wedge \tau'$, $\tau \vee \tau'$ and $(\tau + \tau')$ are stopping times.
(v) If for a real valued random variable Y we have $E(|Y|) < \infty$ and $P(\tau \le \tau') = 1$, then

$$E(Y \,|\, \mathcal{A}_\tau) = E(Y \,|\, \mathcal{A}_{\tau \wedge \tau'}) \tag{5.1.17}$$

and

$$E\left(E(Y \,|\, \mathcal{A}_\tau) \,\middle|\, \mathcal{A}_{\tau'}\right) = E(Y \,|\, \mathcal{A}_\tau). \tag{5.1.18}$$

Optional Sampling Theorem (*)

If $X = \{X_t,\, t \in [0,\infty)\}$ is a right continuous $(\underline{\mathcal{A}}, P)$-supermartingale, then the supermartingale property (5.1.2) is also true if the times s and t in (5.1.2) are stopping times. More precisely, Doob's *Optional Sampling Theorem* states the following result, see Doob (1953), Elliott (1982) or Karatzas & Shreve (1991).

Theorem 5.1.4. (Doob) *If $X = \{X_t,\, t \in [0,\infty)\}$ is a right continuous $(\underline{\mathcal{A}}, P)$-supermartingale on $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$, then it holds for two bounded stopping times τ and τ' with $\tau \le \tau'$ almost surely that*

$$E(X_{\tau'} \,|\, \mathcal{A}_\tau) \le X_\tau \tag{5.1.19}$$

almost surely. Furthermore, if X is also an $(\underline{\mathcal{A}}, P)$-martingale, then equality holds in (5.1.19).

This theorem is important if one wants to apply a pricing rule at a stopping time or the payoff that one aims to price matures at a stopping time. Such case arises for American options that allow exercising the payoff at any time prior to maturity.

Martingale Inequalities (*)

For a given underlying financial quantity, or more generally, a given stochastic process X, it is important to have some upper bounds for its maximum. If $X = \{X_t,\, t \in [0,\infty)\}$ is a right continuous supermartingale, then it can be shown, see Doob (1953) or Elliott (1982), that for any $\lambda > 0$ it holds

$$\lambda\, P\left(\sup_{t\in[0,\infty)} X_t \ge \lambda \,\Big|\, \mathcal{A}_0\right) \le E\left(X_0 \,\big|\, \mathcal{A}_0\right) + E\left(\max(0, -X_0) \,\big|\, \mathcal{A}_0\right). \tag{5.1.20}$$

By exploiting the martingale property (5.1.19) one can prove the following powerful *martingale inequalities*, see Doob (1953) or Elliott (1982). A continuous martingale $X = \{X_t,\, t \in [0,\infty)\}$ with finite pth moment satisfies the *maximal martingale inequality*

$$P\left(\sup_{s\in[0,t]} |X_s| > a\right) \le \frac{1}{a^p}\, E(|X_t|^p) \tag{5.1.21}$$

and the *Doob inequality*

$$E\left(\sup_{s\in[0,t]} |X_s|^p\right) \le \left(\frac{p}{p-1}\right)^p E(|X_t|^p) \tag{5.1.22}$$

for $a > 0$, $p > 1$ and $t \in [0,\infty)$. If X is a continuous martingale, then the maximal martingale inequality provides an estimate for the probability that a level a will be exceeded by the maximum of X. In particular the Doob inequality provides for $p = 2$ for the squared maximum the estimate

$$E\left(\sup_{s\in[0,t]} |X_s|^2\right) \le 4\, E\left(|X_t|^2\right)$$

for $t \in [0,\infty)$. These inequalities are important for deriving a number of fundamental results in stochastic calculus and quantitative finance.

5.2 Quadratic Variation and Covariation

Quadratic Variation

The notion of the, so-called, *quadratic variation* of a given stochastic process X plays a fundamental role in stochastic calculus and, therefore, in finance as well. It is a characteristic of the fluctuating part of a stochastic process and can be easily observed. In this capacity it will be useful for measuring locally in time the risk of an asset price.

To introduce this notion in a simple manner let us consider an *equidistant time discretization*

$$\{t_k = k\,h :\ k \in \{0, 1, \ldots\}\}, \tag{5.2.1}$$

with small time steps of lengths $h > 0$, such that $0 = t_0 < t_1 < t_2 < \ldots$. Thus, we have the discretization times $t_k = k\,h$ for $k \in \{0, 1, \ldots\}$. The specific structure of the time discretization is in fact not essential for the definition of the quadratic variation that we shall use, as long as the maximum time step size vanishes a.s. when approaching the limit. We employ the equidistant time discretization here to simplify our presentation. Other time discretizations with vanishing step size yield the same limit.

For a given stochastic process X the *quadratic variation process* $[X] = \{[X]_t,\ t \in [0, \infty)\}$ is defined as the limit in probability, see (2.7.1), as $h \to 0$ of the sums of squared increments of the process X, provided this limit exists and is unique. For details we refer to Jacod & Shiryaev (2003) and Protter (2004). For instance, for semimartingales, which form a very general class of stochastic processes that we shall introduce in Sect. 5.5, the quadratic variation is uniquely defined. We have at time t the *quadratic variation*

$$[X]_t \stackrel{P}{=} \lim_{h \to 0} [X]_{h,t}, \tag{5.2.2}$$

where the *approximate quadratic variation* $[X]_{h,t}$ is given by the sum

$$[X]_{h,t} = \sum_{k=1}^{i_t} (X_{t_k} - X_{t_{k-1}})^2. \tag{5.2.3}$$

Here i_t denotes the integer

$$i_t = \max\{k \in \mathcal{N} :\ t_k \le t\} \tag{5.2.4}$$

of the last discretization point before or including $t \in [0, \infty)$.

Examples of Quadratic Variations

As an example, Fig. 5.2.1 shows for a standard Wiener process $W = \{W_t, t \in [0, \infty)\}$ a sample path and its approximate quadratic variation $[W]_{h,t}$ with time step size $h = 0.02$ on the interval $[0, 10]$. Note that the approximate

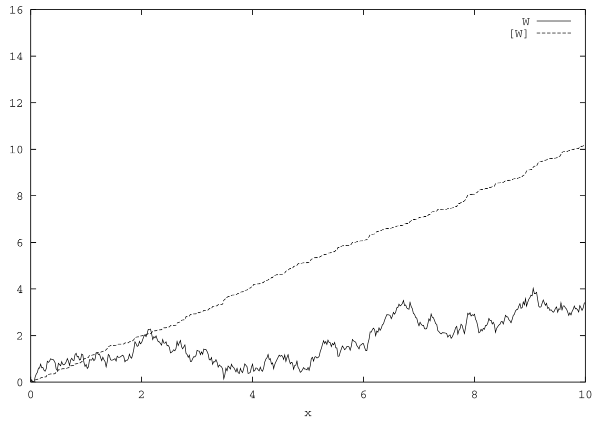

Fig. 5.2.1. A Wiener path W_t and its approximate quadratic variation $[W]_{h,t}$

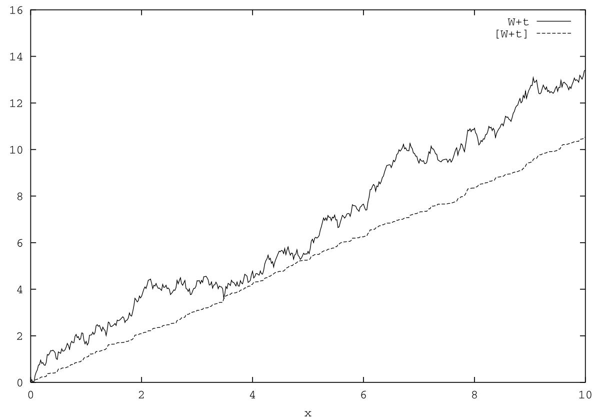

Fig. 5.2.2. Transformed Wiener process Y_t and its approximate quadratic variation $[Y]_{h,t}$

quadratic variation in Fig. 5.2.1 forms almost a straight line with slope one. Indeed, it can be shown, see Karatzas & Shreve (1991) or Elliott (1982), that the value of the quadratic variation process $[W] = \{[W]_t,\, t \in [0,\infty)\}$ at time t for a standard Wiener process W is given by the relation

$$[W]_t = t \tag{5.2.5}$$

for $t \in [0,\infty)$. Thus, for finer time discretizations, the approximate quadratic variation becomes almost a perfect straight line.

In Fig. 5.2.2, a sample path of a transformed Wiener process $Y = \{Y_t,\, t \in [0,\infty)\}$ with values

$$Y_t = W_t + t$$

together with its approximate quadratic variation are displayed. Observe that the drift, which was added to the Wiener process, had practically no impact on the approximate quadratic variation, when compared to Fig. 5.2.1. This effect can be explained by noting that for a stochastic process $F = \{F_t,\, t \in [0,\infty)\}$ its, so-called, *total variation* is

$$[F]_t^{\frac{1}{2}} \stackrel{P}{=} \lim_{h\to 0} \sum_{k=1}^{i_t} \left| F_{t_k} - F_{t_{k-1}} \right| \tag{5.2.6}$$

for $t \in [0,\infty)$. Note that in the case where $F_t = t$ for $t \in [0,\infty)$ the total variation $[F]_t^{\frac{1}{2}} = t$ is bounded. However, $F_t = t$ has zero quadratic variation since

$$[F]_t \stackrel{P}{=} [t]_t \stackrel{P}{=} \lim_{h\to 0} \sum_{k=1}^{i_t} (t_k - t_{k-1})^2 = 0. \tag{5.2.7}$$

One notes that a differentiable function has finite total variation but zero quadratic variation. In contrast to that one can show that the strongly fluctuating Wiener process has no finite total variation but some finite quadratic variation.

It is then possible to show that the above transformed Wiener process Y has the finite quadratic variation

$$\begin{aligned}
[Y]_t &\stackrel{P}{=} \lim_{h\to 0} \sum_{k=1}^{i_t} (Y_{t_k} - Y_{t_{k-1}})^2 \\
&\stackrel{P}{=} \lim_{h\to 0} \sum_{k=1}^{i_t} \Big((W_{t_k} - W_{t_{k-1}})^2 + 2\,(W_{t_k} - W_{t_{k-1}})\,(t_k - t_{k-1}) + \,(t_k - t_{k-1})^2 \Big) \\
&\stackrel{P}{=} [W]_t,
\end{aligned} \tag{5.2.8}$$

for $t \in [0,\infty)$, which is the same as that for the Wiener process. Here only the sum of the squared Wiener process increments does not vanish asymptotically. We note that only the martingale term in the transformed Wiener process, which is in the above example the Wiener process itself, contributes to the quadratic variation.

Another Martingale

Starting with a continuous, square integrable $(\underline{\mathcal{A}}, P)$-martingale X, another $(\underline{\mathcal{A}}, P)$-martingale can be constructed by using its quadratic variation $[X]$ if $E([X]_T) < \infty$ for each $T \in [0,\infty)$. More precisely, a new continuous $(\underline{\mathcal{A}}, P)$-martingale $Y = \{Y_t,\, t \in [0,\infty)\}$ is obtained by setting

$$Y_t = (X_t)^2 - [X]_t \tag{5.2.9}$$

for $t \in [0,\infty)$, see Protter (2004).

In the case of a standard Wiener process W, we obtain the martingale $Y = \{Y_t = (W_t)^2 - t,\, t \in [0,\infty)\}$, see (5.1.6). The type of martingale property of Y given in (5.2.9) is fundamental to stochastic calculus.

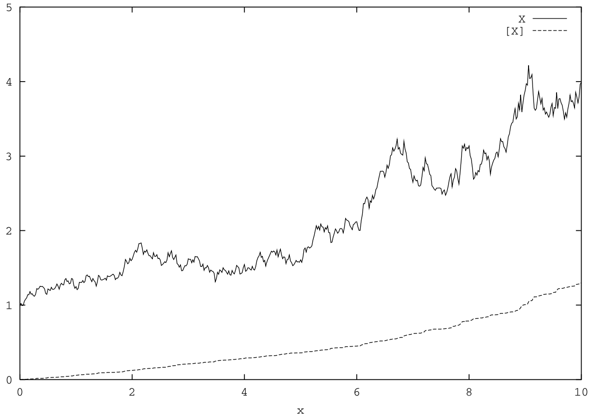

Fig. 5.2.3. Path of a geometric Brownian motion and its quadratic variation

Quadratic Variation and Geometric Brownian Motion

The quadratic variation turns out to be one of the most important characteristics of a martingale. The standard market model for an asset price is the Black-Scholes (BS) model, given by a geometric Brownian motion. To highlight the usefulness of the quadratic variation in such a financial context we consider as a model for an asset price X_t at time t the BS model, see (4.1.2), which we write in the form

$$X_t = X_0 \exp\{L_t\}, \tag{5.2.10}$$

where

$$L_t = g\, t + \sigma\, W_t \tag{5.2.11}$$

for $t \in [0, \infty)$. Here $W = \{W_t,\, t \in [0, \infty)\}$ denotes again a standard Wiener process. With the choice of the growth rate $g = r - \frac{1}{2}\,\sigma^2$ this provides the same dynamics as was given in (5.1.9). When we use the initial value $X_0 = 1$, the expected rate of return $r = 0.05$ and the volatility $\sigma = 0.2$, then the quadratic variation $[X]$ for X is shown in Fig. 5.2.3. Also displayed in Fig. 5.2.3 is the sample path for X, see also Fig. 5.1.3. Note that the quadratic variation is not linear. However, if we visualize the quadratic variation of the logarithm $\ln(X_t)$ of X_t, then we obtain, as can be seen in Fig. 5.2.4, an almost perfect straight line. The reason for this effect can be directly seen when using the following identities

$$[\ln(X)]_t = [L]_t = \sigma^2\, [W]_t = \sigma^2\, t \tag{5.2.12}$$

for $t \in [0, \infty)$. These relations hold because $L_t = \ln(X_t)$ forms a linearly transformed Wiener process and we can use the fact that $[W]_t = t$, see (5.2.5).

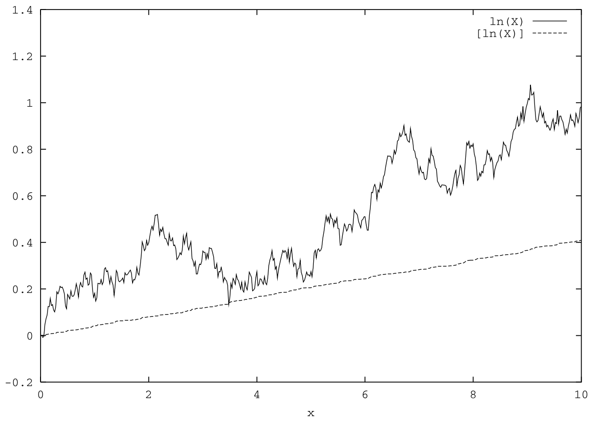

Fig. 5.2.4. Path of $\ln(X)$ and $[\ln(X)]$

Volatility

The key quantity for the parametrization of the BS model, which was the standard market model for many decades, is the *volatility*. We observe in (5.2.12) that under the BS model the squared volatility is the time derivative of the quadratic variation of the logarithm of the asset price. We can express this important observation in the form

$$\sigma^2 = \frac{d}{dt}\,[\ln(X)]_t. \tag{5.2.13}$$

This relation can still be used theoretically as a definition for the volatility of a continuous asset price process, even if its dynamics is not that of a geometric Brownian motion.

To be more precise, we define the historical volatility $\mathrm{Vol}_X(t)$ at a given time $t \in [0,\infty)$ of a given continuous asset price process X, as the square root of the left hand derivative of the quadratic variation of the logarithm of X. That is, we define the historical volatility in the form

$$\mathrm{Vol}_X(t) = \sqrt{\frac{d}{dt}\,[\ln(X)]_t} \tag{5.2.14}$$

for $t \in [0,\infty)$. A common market practice for estimating squared volatility, see for instance Hull (2000), is that one estimates the sample variance of log-returns, see (2.1.19). Note that $\mathrm{Vol}_X(t)$ is by (5.2.3) and (5.2.2) asymptotically equivalent to the way that volatility is calculated in practice. However, it is well-known, see for instance, Corsi, Zumbach, Müller & Dacorogna (2001) and Barndorff-Nielsen & Shephard (2003), that the estimation of volatility is in practice a very delicate task.

The definition of volatility in (5.2.14) is quite general and can be used for all continuous asset price processes. It has the advantage that it is independent

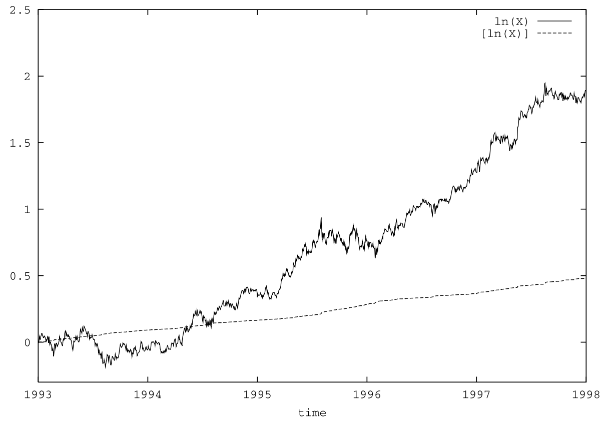

Fig. 5.2.5. IBM log-share price and its quadratic variation

of the specific choice of the underlying asset price model and also the time discretization employed. In the particular case of geometric Brownian motion it leads us directly to the constant volatility of the BS model, as can be seen from (5.2.12). We shall see, that the above definition of historical volatility is useful for the study of the actual volatility dynamics in asset price models. Furthermore, the approximate quadratic variation (5.2.3) can be directly used to construct a volatility estimator.

It is well-known that in reality, volatility is stochastic, as can be seen from the changing slope of the quadratic variation of the logarithm of asset prices. This indicates that the standard market model with its constant volatility can only be considered to be used as a first, rough approximation of the existing market dynamics. As another example for an application of the above definition of historical volatility, Fig. 5.2.5 shows the logarithm $\ln(\frac{X_t}{X_{t_0}})$ of the IBM share price X_t from 1993 up until 1998 together with its approximate quadratic variation based on daily observations. According to the definition of historical volatility in (5.2.14) we can interpret the square root of the slope in Fig. 5.2.5 as an empirical volatility estimate of the IBM share price during the corresponding time period. By estimating the observed slope of the quadratic variation in Fig. 5.2.5, an annualized average volatility of approximately $\sqrt{\frac{0.5}{5}} = \sqrt{0.1} \approx 0.32$ is inferred.

To illustrate further the type of information that the quadratic variation provides we show in Fig. 5.2.6 the logarithm $\ln(\frac{X_t}{X_{t_0}})$ of the S&P500 index for the period from 1993 up until 1998 together with its quadratic variation. Note that the average slope of the quadratic variation in Fig. 5.2.6, that is its squared volatility, is much smaller than that for the IBM share price in Fig. 5.2.5. This is mainly due to the effect of diversification for the index. Again, an approximate estimate for the average volatility of the S&P500 index can be obtained from the square root of the slope of the quadratic variation shown in Fig. 5.2.6. Thus, we estimate an annualized average volatility of

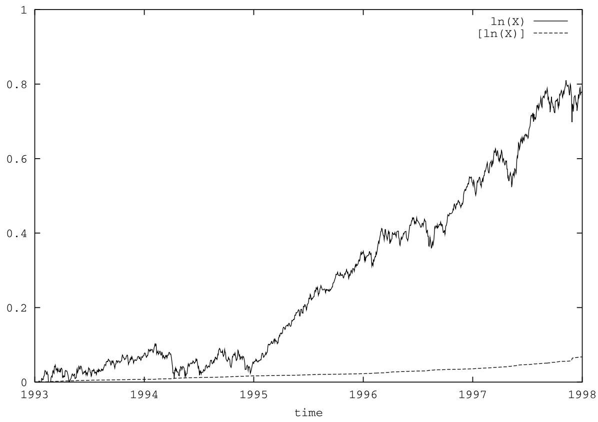

Fig. 5.2.6. Logarithm of S&P500 and its quadratic variation

about $\sqrt{\frac{0.05}{5}} = \sqrt{0.01} = 0.1$, which is about a third of the estimated volatility of the IBM share price.

Covariation

In a similar manner as the quadratic variation the covariation of two continuous stochastic processes can be defined. This is another important tool which turns out to be useful for the characterization of dependencies between two stochastic processes, for instance, between asset prices. It allows the, locally in time, measurement of associations between the random fluctuations of two different continuous processes.

For the definition of covariation the same equidistant time discretization, as given in (5.2.1), is now used. That is, we set $t_k = kh$ for $k \in \{0, 1, \ldots\}$, $h > 0$. For continuous stochastic processes Z_1 and Z_2 the *covariation process* $[Z_1, Z_2] = \{[Z_1, Z_2]_t,\, t \in [0, \infty)\}$ is defined as the limit in probability, see (2.7.2), as $h \to 0$ of the values of the *approximate covariation process* $[Z_1, Z_2]_{h,\cdot}$ with

$$[Z_1, Z_2]_{h,t} = \sum_{k=1}^{i_t} (Z_1(t_k) - Z_1(t_{k-1}))(Z_2(t_k) - Z_2(t_{k-1})) \tag{5.2.15}$$

for $t \in [0, \infty)$ and $h > 0$, given by the sums of the products of the increments of the processes Z_1 and Z_2. Here the integer i_t is as introduced in (5.2.4). More precisely, we define at time $t \in [0, \infty)$ the *covariation*

$$[Z_1, Z_2]_t \overset{P}{=} \lim_{h \to 0} [Z_1, Z_2]_{h,t}, \tag{5.2.16}$$

where $[Z_1, Z_2]_{h,t}$ is the approximate covariation.

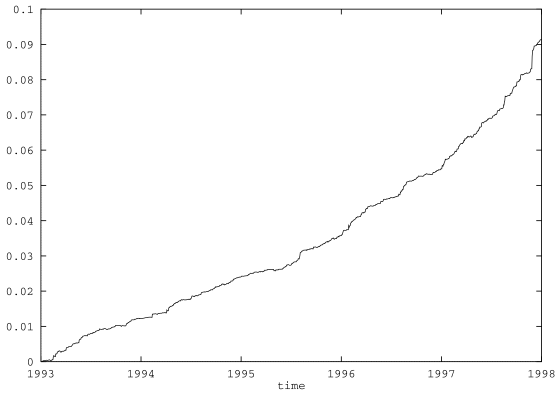

Fig. 5.2.7. Covariation between logarithms of S&P500 and IBM share price

As an example, we display in Fig. 5.2.7 the approximate covariation between the logarithms of the S&P500 index, see Fig. 5.2.6, and the IBM share price, see Fig. 5.2.5, for the period from 1993 up until 1998 using daily observations. Note that the average slope of the covariation seems to be here almost always positive, which indicates some association between the movements of the IBM share price and those of the S&P500 index. Summarizing these observations, it appears that the covariation provides a useful tool for measuring the degree of association of the fluctuations of two stochastic processes locally in time. Obviously, if the processes Z_1 and Z_2 are identical, then their covariation coincides with their quadratic variation.

Covariation for Processes with Jumps (*)

For any right-continuous stochastic process $\xi = \{\xi(t),\, t \in [0, \infty)\}$ we denote by

$$\xi(t-) \stackrel{\text{a.s.}}{=} \lim_{h \to 0+} \xi(t-h) \tag{5.2.17}$$

the almost sure *left hand limit* of $\xi(t)$ at time $t \in (0, \infty)$. The *jump size* $\Delta\xi(t)$ at time t is then defined as

$$\Delta\xi(t) = \xi(t) - \xi(t-) \tag{5.2.18}$$

for $t \in (0, \infty)$.

In the case of a pure jump process $p = \{p_t,\, t \in [0, \infty)\}$ the corresponding quadratic variation is obtained as

$$[p]_t = \sum_{0 \le s \le t} (\Delta p_s)^2 \tag{5.2.19}$$

for $t \in [0, \infty)$, where $\Delta p_s = p_s - p_{s-}$. In the case when p is a Poisson process, its quadratic variation equals the process itself, that is, $[N]_t = N_t$ for all

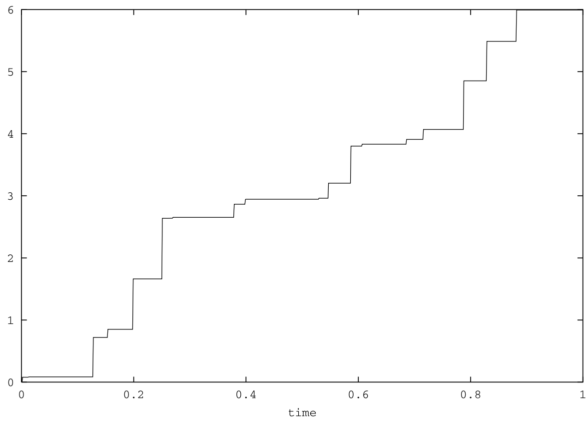

Fig. 5.2.8. Quadratic variation of a compound Poisson process

$t \in [0, \infty)$. We show in Fig. 5.2.8 the quadratic variation $[Y]_t$ of the trajectory of the compound Poisson process Y shown in Fig. 3.5.2.

It is preferable to separate the jump part of a process when computing its quadratic variation. For a general stochastic process the quadratic variation consists of the sum of the quadratic variations of its continuous and its pure jump part. This will be made more precise below.

Let us denote by Z_1 and Z_2 two stochastic processes with continuous part

$$Z_i^c(t) = Z_i(t) - Z_i(0) - \sum_{0<s\leq t} \Delta Z_i(s) \tag{5.2.20}$$

for $t \in [0, \infty)$ and $i \in \{1, 2\}$. Here the *jump size* at time s is given as

$$\Delta Z_i(s) = Z_i(s) - Z_i(s-) \tag{5.2.21}$$

for $s \in [0, \infty)$ and we assume that the sum in (5.2.20) is almost surely finite. The covariation $[Z_1, Z_2]_t$ of Z_1 and Z_2 at time t is then defined as

$$[Z_1, Z_2]_t = [Z_1^c, Z_2^c]_t + \sum_{0<s\leq t} (\Delta Z_1(s)) (\Delta Z_2(s)) \tag{5.2.22}$$

for $t \in [0, \infty)$, as long as the quantities involved are almost surely finite. This also means that the quadratic variation of a process Z_1 equals the quadratic variation $[Z_1^c]_t$ of its continuous part Z_1^c plus the sum of the squares of its jumps, that is

$$[Z_1]_t = [Z_1^c]_t + \sum_{0<s\leq t} (\Delta Z_1(s))^2 \tag{5.2.23}$$

for $t \in [0, \infty)$. Again, we assume that the expressions involved are almost surely finite. The above notion of covariation for processes with jumps is convenient and useful. Obviously, if the processes Z_1 and Z_2 are identical, then their quadratic variation coincides with their covariation. The quadratic variation $[q]_t$ of the compensated Poisson process $q = \{q_t, \, t \in [0, \infty)\}$, shown in Fig. 5.1.4, equals that of the Poisson process N shown in Fig. 3.5.1, that is $[N]_t = [q]_t$ for $t \in [0, \infty)$.

We emphasize that the covariation of a process with continuous and jump part is an important characteristic in financial modeling, see Cont & Tankov (2004) and Ait-Sahalia (1996).

*p*th Variation (*)

We call for $p > 0$ and a stochastic process $X = \{X_t, \, t \in [0, \infty)\}$ the process $[X]_h^{\frac{p}{2}} = \left\{[X]_{h,t}^{\frac{p}{2}}, \, t \in [0, \infty)\right\}$ with

$$[X]_{h,t}^{\frac{p}{2}} = \sum_{k=1}^{i_t} |X_{t_k} - X_{t_{k-1}}|^p \tag{5.2.24}$$

for $t \in [0, \infty)$ the *approximate pth variation process* of X. Then the pth *variation process* $[X]^{\frac{p}{2}} = \left\{[X]_t^{\frac{p}{2}}, \, t \in [0, \infty)\right\}$ is for each $t \in [0, \infty)$ defined as the limit in probability

$$[X]_t^{\frac{p}{2}} \stackrel{P}{=} \lim_{h \to 0} [X]_{h,t}^{\frac{p}{2}}, \tag{5.2.25}$$

see (2.7.1). The first order variation process $[X]^{\frac{1}{2}}$ is called *total variation*, see (5.2.6). For instance, the time t with $X_t = t$ is a process with total variation $[X]_t^{\frac{1}{2}} = t < \infty$ a.s. for $t \in [0, \infty)$. Furthermore, any differentiable process can be shown to have finite total variation. Note that the Wiener process $W = \{W_t, \, t \in [0, \infty)\}$ does *not* have finite total variation, however, it has finite quadratic variation, as shown in (5.2.5). On the other hand, a Poisson process with finite intensity does have finite total variation.

Local Martingales (*)

As we shall see later, in quantitative finance stochastic processes naturally appear that are not martingales but become martingales if they are properly stopped. These *local martingales* are locally in time similar to martingales.

Definition 5.2.1. *A stochastic process $X = \{X_t, \, t \in [0, \infty)\}$ is an $(\underline{\mathcal{A}}, P)$-* local martingale *if there exists an increasing sequence $(\tau_n)_{n \in \mathcal{N}}$ of stopping times, that may depend on X, such that $\lim_{n \to \infty} \tau_n \stackrel{a.s.}{=} \infty$ and each stopped process*

$$X^{\tau_n} = \{X_t^{\tau_n} = X_{t \wedge \tau_n}, \, t \in [0, \infty)\} \tag{5.2.26}$$

is an $(\underline{\mathcal{A}}, P)$-martingale, where $t \wedge \tau_n = \min(t, \tau_n)$.

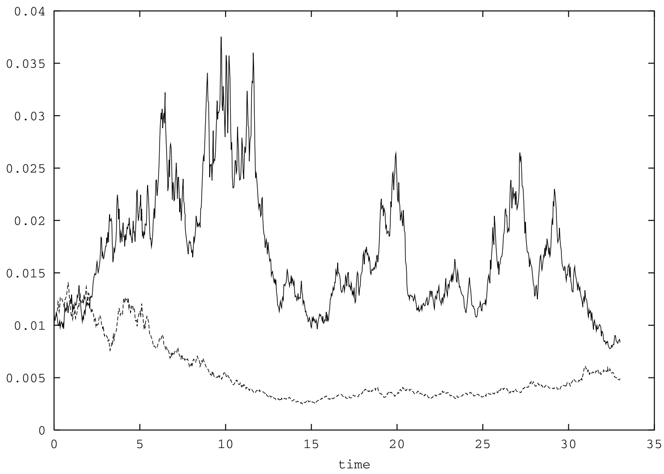

Fig. 5.2.9. Two trajectories of a strict local martingale

If X is a local martingale, then the value X_s does, in general, not equal the conditional expectation $E(X_t|\mathcal{A}_s)$ for $s \in [0,\infty)$ and $t \in [s,\infty)$. Note that an $(\underline{\mathcal{A}}, P)$-martingale is also an $(\underline{\mathcal{A}}, P)$-local martingale. However, an $(\underline{\mathcal{A}}, P)$-local martingale is not always an $(\underline{\mathcal{A}}, P)$-martingale. A local martingale that is not a martingale is called a *strict local martingale*.

To provide an example for such a strict local martingale $X = \{X_t, t \in [0,\infty)\}$ we form the sum of the squares of four independent Wiener processes W^1, W^2, W^3, W^4 which start each at the value $W_0^i = 5$, $i \in \{1,2,3,4\}$. By taking the inverse of this sum, that is

$$X_t = \left(\sum_{i=1}^{4} \left(W_t^i + 5 \right)^2 \right)^{-1} \tag{5.2.27}$$

for $t \in [0,\infty)$, we shall show later that this inverse of a squared Bessel process of dimension four forms a strict local martingale, see Revuz & Yor (1999). Two paths of such a process $X = \{X_t, t \in [0,\infty)\}$ are shown in Fig. 5.2.9. They both look rather different but are both constructed according to (5.2.27). It appears that they can mimic very different behaviors, in particular, over the initial time period. As we discuss later in the context of squared Bessel processes, this process has peculiar properties that differentiate it from a martingale, see Revuz & Yor (1999).

One can formulate the following statements, see Protter (2004), that will become relevant when dealing with local martingales in financial modeling under the benchmark approach.

Lemma 5.2.2.

(i) *An almost surely nonnegative (negative) $(\underline{\mathcal{A}}, P)$-local martingale is an $(\underline{\mathcal{A}}, P)$-supermartingale (submartingale).*

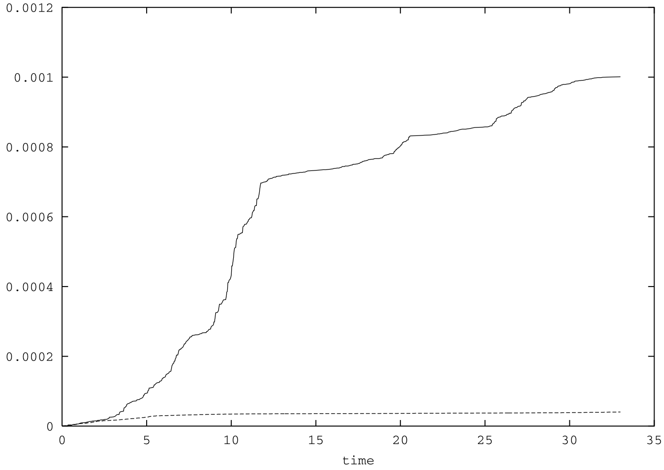

Fig. 5.2.10. Quadratic variations of two trajectories of a strict local martingale

(ii) *An a.s. uniformly bounded* $(\underline{\mathcal{A}}, P)$*-local martingale is an* $(\underline{\mathcal{A}}, P)$*-martingale.*

(iii) *A square integrable* $(\underline{\mathcal{A}}, P)$*-local martingale* X *is a square integrable* $(\underline{\mathcal{A}}, P)$*-martingale if and only if*

$$E([X]_T) < \infty \tag{5.2.28}$$

for all $T \in [0, \infty)$.

We prove the assertion (i) at the end of this section. The Definition 5.2.1 of a local martingale is rather technical and somehow difficult to verify. However, the statement (iii) of the above lemma is quite useful in practice because local martingales that one typically faces in finance seem to be square integrable. The statement (iii) means that if the fluctuations of a square integrable local martingale are so strong that the mean of its quadratic variation does not exist, then it cannot be a martingale and is therefore a strict local martingale. In Fig. 5.2.10 we show the quadratic variation of the two paths of the strict local martingale shown in Fig. 5.2.9. Note that its quadratic variation appears to be highly dependent on the particular path. As we shall see later, one can show for the given example that $E([X]_t) = \infty$ for $t \in (0, \infty)$. This means that the quadratic variation of different paths varies so strongly that no finite expectation can be calculated.

We face here a subtle but important property, which will be highly relevant for the understanding of the typical dynamics of financial markets as it becomes visible under the benchmark approach.

Nonnegative Local Martingales are Supermartingales (*)

As we shall see later the statement (i) in Lemma 5.2.2 is crucial for the benchmark approach when it establishes no-arbitrage. For completeness we provide here a proof.

Lemma 5.2.3. *A nonnegative* $(\underline{\mathcal{A}}, P)$*-local martingale* $X = \{X_t,\, t \in [0,\infty)\}$ *with* $E(X_t|\mathcal{A}_s) < \infty$ *for all* $0 \le s \le t < \infty$ *is an* $(\underline{\mathcal{A}}, P)$*-supermartingale.*

Proof: Consider a nonnegative $(\underline{\mathcal{A}}, P)$-local martingale $X = \{X_t,\, t \in [0,\infty)\}$. Then there exists an increasing sequence $(\tau_n)_{n\in\mathcal{N}}$ of stopping times, with respect to the filtration $\underline{\mathcal{A}}$, such that each stopped process $X^{\tau_n} = \{X_t^{\tau_n} = X_{t\wedge\tau_n},\, t \in [0,\infty)\}$ is an $(\underline{\mathcal{A}}, P)$-martingale and we have $\tau_n \to \infty$ almost surely. Consequently, for each $n \in \mathcal{N}$ and $0 \le s \le t < \infty$ we have

$$
\begin{aligned}
E(X_t \,|\, \mathcal{A}_s) &= E\left(\mathbf{1}_{\{\tau_n \ge t\}}\, X_t \,\middle|\, \mathcal{A}_s\right) + E\left(\mathbf{1}_{\{\tau_n < t\}}\, X_t \,\middle|\, \mathcal{A}_s\right) \\
&= E\left(\mathbf{1}_{\{\tau_n \ge t\}}\, X_t^{\tau_n} \,\middle|\, \mathcal{A}_s\right) + E\left(\mathbf{1}_{\{\tau_n < t\}}\, X_t \,\middle|\, \mathcal{A}_s\right) \\
&\le E\left(X_t^{\tau_n} \,\middle|\, \mathcal{A}_s\right) + E\left(\mathbf{1}_{\{\tau_n < t\}}\, X_t \,\middle|\, \mathcal{A}_s\right) \\
&= X_s^{\tau_n} + E\left(\mathbf{1}_{\{\tau_n < t\}}\, X_t \,\middle|\, \mathcal{A}_s\right).
\end{aligned}
\tag{5.2.29}
$$

Since we have for each $t \in [0,\infty)$ by definition that $\mathbf{1}_{\{\tau_n \ge t\}} X_t$ approaches X_t almost surely from below as $n \to \infty$ it follows by (5.2.29) and the Monotone Convergence Theorem, see (2.7.9), that the difference

$$
E(X_t \,|\, \mathcal{A}_s) - E\left(\mathbf{1}_{\{\tau_n < t\}}\, X_t \,\middle|\, \mathcal{A}_s\right) = E\left(\mathbf{1}_{\{\tau_n \ge t\}}\, X_t \,\middle|\, \mathcal{A}_s\right)
$$

approaches almost surely the conditional expectation $E(X_t \,|\, \mathcal{A}_s)$ from below as $n \to \infty$. As a consequence of that, the conditional expectation $E(\mathbf{1}_{\{\tau_n < t\}}\, X_t \,|\, \mathcal{A}_s)$ is for $n \to \infty$ decreasing and converges almost surely to zero. By using the fact that $\lim_{n\to\infty} X_s^{\tau_n} \overset{\text{a.s.}}{=} X_s$ yields in (5.2.29) the inequality $E(X_t \,|\, \mathcal{A}_s) \le X_s$ when letting n tend to infinity. This proves the lemma. $\square$

In Rogers & Williams (2000) one can find an alternative proof of this result based on Fatou's Lemma, see (2.7.11). We emphasize that it is essential in Lemma 5.2.3 that one defines the local martingale over the infinite time interval $[0,\infty)$ and not on $[0,\infty]$ or $[0,T]$ with $T \in (0,\infty)$ since the above result does not hold in these cases.

Fig. 5.3.1. Gains from trade of one share of IBM stock during 1993 - 1998

5.3 Gains from Trade as Stochastic Integral

One of the most fundamental notions in finance is that of gains from trade. In stochastic calculus this corresponds exactly to the notion of a stochastic integral, the Itô integral, which is therefore highly relevant in finance.

Gains from Trade

Let us consider an investor who holds during the time period $[0, T]$ a constant number $\xi(0)$ of units of an asset with price process $X = \{X_t,\, t \in [0, T]\}$. The investor's allocation strategy $\xi = \{\xi(t) = \xi(0),\, t \in [0, T]\}$, characterized by the number of units of the asset held, is assumed to be constant in this case. Then the investor's *gains from trade* over the period $[0, t]$ equals

$$I_{\xi,X}(t) = \xi(0)\, \{X_t - X_0\}, \tag{5.3.1}$$

for $t \in [0, T]$. This provides the first step towards an appropriate definition of a stochastic integral, which we shall call later Itô integral. Formally, we interpret the above gains from trade $I_{\xi,X}(t)$ as an *Itô integral* of the *integrand* ξ with respect to the *integrator* X over the time interval $[0, t]$, and use the following notation

$$I_{\xi,X}(t) = \int_0^t \xi(s)\, dX_s. \tag{5.3.2}$$

To illustrate the above construction we show in Fig. 5.3.1 the gains from trade obtained from IBM share holdings over the period from 1993 to 1998. This refers to a constant allocation strategy which is holding $\xi(t) = 1$ share.

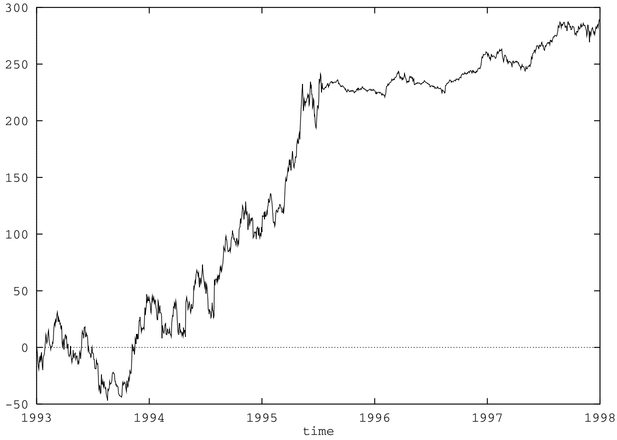

Fig. 5.3.2. Gains from trade of ten and later one share of IBM stock

Piecewise Constant Allocation Strategies

Now, let us allow the investor to change his or her strategy so that it becomes a piecewise constant allocation process $\xi = \{\xi(t),\, t \in [0, T]\}$ with $\xi(t) = \xi(t_k)$ units of shares held at time $t \in [t_k, t_{k+1})$, $k \in \{0, 1, \ldots\}$ and $t_k = kh$ for $h > 0$. Here the reallocation times t_k form an equidistant time discretization, as given in (5.2.1). Obviously, the gains from trade over the period $[0, t]$ can be expressed in the form

$$\int_0^t \xi(s)\, dX_s = \sum_{k=1}^{i_t} \xi(t_{k-1})\, \{X_{t_k} - X_{t_{k-1}}\} + \xi(t_{i_t})\, \{X_t - X_{t_{i_t}}\}, \qquad (5.3.3)$$

where

$$i_t = \max\{k \in \mathcal{N} :\, t_k \leq t\} \qquad (5.3.4)$$

is the integer index of the latest discretization time before and including t, see (5.2.4). Here we formally interpret the gains from trade as an Itô integral in the same form as in (5.3.2) with integrand ξ and integrator X covering the interval $[0, t]$.

In Fig. 5.3.2, the gains from trade are displayed when during the first half of the time period, that is until mid 1995, ten shares of IBM were held and in the second half only one share. One observes during the first period strong fluctuations of the gains from trade when compared to the second half of that time period.

Itô Integral as a Limit

It is sufficient in many applications to use a Wiener process as integrator. Therefore, we use in a standard setting often the Itô integral with respect to the Wiener process $W = \{W_t,\, t \in [0,\infty)\}$ as integrator over the interval $[0,t]$ for a wide range of integrands $\xi = \{\xi(t),\, t \in [0,\infty)\}$.

Definition 5.3.1. *For a left continuous stochastic process $\xi = \{\xi(t),\, t \in [0,\infty)\}$ as integrand with*

$$\int_0^T \xi(s)^2\, ds < \infty \tag{5.3.5}$$

for all $T \in [0,\infty)$ almost surely, the Itô integral with respect to the Wiener process W is defined as the left continuous limit in probability

$$\int_0^t \xi(s)\, dW_s \stackrel{P}{=} \lim_{h\to 0} \sum_{k=1}^{i_t} \xi(t_{k-1})\, \{W_{t_k} - W_{t_{k-1}}\} \tag{5.3.6}$$

of the sequence of corresponding approximating sums for $t \in [0,\infty)$.

For details on the definition of Itô integrals we refer to Karatzas & Shreve (1991), Kloeden & Platen (1999) or Protter (2004). We see that the right hand sides of both (5.3.3) and (5.3.6) are very similar and coincide in the case of piecewise constant integrands. Consequently, the Itô integral can be seen as a limit in probability of gains from trade, taken over progressively finer time discretizations.

An important characteristic of the Itô integral is that the evaluation point t_{k-1} for the integrand ξ is always taken at the left hand side of the discretization interval $[t_{k-1}, t_k)$. This feature is natural for finance applications because an investor needs to decide at the beginning of an investment period how many units of a security he or she wants to hold. It distinguishes the Itô integral from other stochastic integrals, see Protter (2004). The choice of the evaluation point at the left hand side corresponds in finance to the economically given fact that once an allocation is made it remains constant for some period of time and cannot be changed retrospectively in a legal manner. As we shall see later, this fact is essential for establishing the martingale property for Itô integrals with respect to Wiener processes.

The above definition of Itô integrals can be extended to include more general classes of integrators rather than just the Wiener process, see Protter (2004), which will be discussed later.

Explicit Value for an Itô Integral

To give a simple example of how the Itô integral differs from the classical, say, Riemann-Stieltjes integral, let us consider a trading strategy, where the number of shares held in an asset equals its price. For simplicity, we assume the asset price to be modeled by the Wiener process W. Then according to (5.3.6) we obtain

$$\int_0^t W_s\, dW_s \stackrel{P}{=} \lim_{h\to 0} \sum_{k=1}^{i_t} W_{t_{k-1}} (W_{t_k} - W_{t_{k-1}}) \tag{5.3.7}$$

for $t \in [0, \infty)$. If W were differentiable, then we would obtain from the deterministic integration rule the quantity $\frac{1}{2} W_t^2$ as the value of this integral at time t. However, in the stochastic case the correct value will be much less, as we shall see. This means that the gains from trade under this strategy do not accumulate in the same way as they would for differentiable asset prices or under the classical integration rule.

The following calculation demonstrates this important effect in more detail. By subtracting and adding $W_{t_k}^2$ and completing the square on the right hand side of (5.3.7) for each time step we see that

$$\int_0^t W_s \, dW_s \overset{P}{=} \lim_{h \to 0} \frac{1}{2} \sum_{k=1}^{i_t} \left\{ \left(W_{t_k}^2 - W_{t_{k-1}}^2 \right) - (W_{t_k} - W_{t_{k-1}})^2 \right\}$$

$$\overset{P}{=} \frac{1}{2} W_t^2 - \frac{1}{2} W_0^2 - \lim_{h \to 0} \frac{1}{2} \sum_{k=1}^{i_t} (W_{t_k} - W_{t_{k-1}})^2,$$

where all except the first and last terms in the first sum cancel each other. From the definition of the approximate quadratic variation of standard Wiener processes in (5.2.3) we have $[W]_t = t$, see (5.2.5), and $W_0 = 0$, see (3.2.6). Consequently, the value of the Itô integral (5.3.7) is

$$\int_0^t W_s \, dW_s = \frac{1}{2} W_t^2 - \frac{1}{2} [W, W]_t = \frac{1}{2} W_t^2 - \frac{1}{2} [W]_t = \frac{1}{2} W_t^2 - \frac{1}{2} t. \quad (5.3.8)$$

The quantity on the right hand side of this equation is clearly less than $\frac{1}{2} W_t^2$, which would be expected for a differentiable function under classical integration. Note that the difference is equal to half the covariation of integrand and integrator. We shall see below that this property holds more generally.

The above example exhibits striking differences between the Itô integral and the classical integral. Since, in practice, asset price processes with properties similar to those of Wiener processes are typically encountered, these differences turn out to be crucial for the rigorous modeling in finance. For instance, the computation of derivative prices, values of portfolios and other financial quantities may become incorrect, if these differences were ignored. Stochastic calculus which we introduce in this and the following two chapters will allow us to obtain correct quantities.

To illustrate these differences we show in Fig. 5.3.3 the path of a Wiener process together with half of its squared value and the Itô integral

$$I_{W,W}(t) = \int_0^t W_s \, dW_s = \int_0^t \int_0^s dW_z \, dW_s, \quad (5.3.9)$$

for $t \in [0, 1]$. Note in this figure the significant difference between the Itô integral $I_{W,W}(t)$ and the value $\frac{1}{2} W_t^2$ that would be obtained under the classical integration rule.

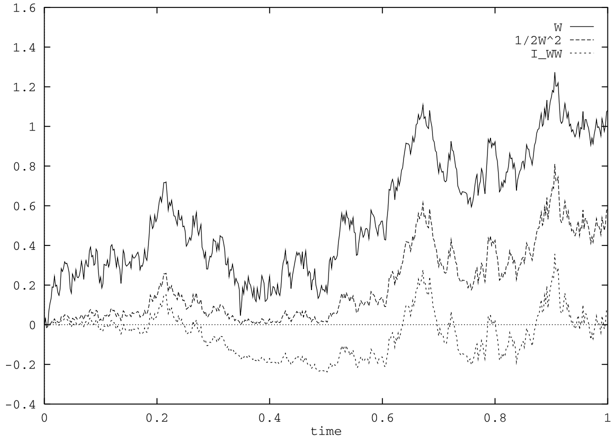

Fig. 5.3.3. Paths of W, $\frac{1}{2}W^2$ and $I_{W,W}$

In the following analysis it will be shown that the differences between Itô and classical integration relate to the covariation of the processes involved as integrand and integrator. These differences are crucial and impact significantly the area of quantitative finance due to the nature of asset prices.

General Itô Integrals and Differentials

The definition of an Itô integral as gains from trade, given in (5.3.6), can naturally be extended to include more general integrators. Let us again use, for simplicity, the equidistant time discretization (5.2.1) and denote, as previously, by Y_{t-} the left hand limit of the value of a process $Y = \{Y_t,\, t \in [0,\infty)\}$ at time $t \in [0,\infty)$. We define for a stochastic process $X = \{X_t,\, t \in [0,\infty)\}$ as integrator and a predictable process $\xi = \{\xi(t),\, t \in [0,\infty)\}$ as integrand with

$$\int_0^T \xi(s)^2 \, d[X]_s < \infty \tag{5.3.10}$$

for all $T \in [0,\infty)$ a.s., the Itô integral as the limit in probability

$$\int_0^t \xi(s)\, dX_s \stackrel{P}{=} \lim_{h\to 0} \sum_{k=1}^{i_t} \xi(t_{k-1})\,(X_{t_k} - X_{t_{k-1}}) \tag{5.3.11}$$

for $t \in [0,\infty)$, provided this limit exists. For details we refer the reader to Protter (2004). Here i_t is the integer index given by (5.3.4) for $t \in [0,\infty)$. We emphasize that in financial applications the Itô integral can be naturally interpreted as gains from trade. Furthermore, one can use almost any adapted process ξ, which satisfies (5.3.10), to form an integrand by using its predictable version with left hand limits.

Let $e = \{e_t,\, t \in [0,\infty)\}$ and $f = \{f_t,\, t \in [0,\infty)\}$ be predictable stochastic processes. Consider a stochastic process $Y = \{Y_t,\, t \in [0,\infty)\}$, where

$$Y_t = y_0 + \int_0^t e_s\, ds + \int_0^t f_s\, dW_s \tag{5.3.12}$$

for $t \in [0,\infty)$ and initial value $Y_0 = y_0$. Here $W = \{W_t,\, t \in [0,\infty)\}$ is a standard Wiener process and we assume that appropriate measurability and integrability conditions apply so that the above integrals exist. In particular, the first integral is a random ordinary Riemann-Stieltjes integral for $t \in [0,\infty)$. It exists if

$$\int_0^t |e_s|\, ds < \infty \tag{5.3.13}$$

for all $t \in [0,\infty)$ a.s. The second integral is an Itô integral with respect to the Wiener process W, see (5.3.6), where we assume that

$$\int_0^t |f_s|^2\, ds < \infty \tag{5.3.14}$$

for all $t \in [0,\infty)$ a.s. It is common to use the following more compact way of expressing the integral equation (5.3.12): The *Itô differential* dY_t of Y at time t is given by the expression

$$dY_t = e_t\, dt + f_t\, dW_t \tag{5.3.15}$$

for $t \in [0,\infty)$ with $Y_0 = y_0$. This is simply another symbolic way of writing (5.3.12), where one should not forget to add the specification of the initial value Y_0. The processes e and f are called *drift* and *diffusion coefficients* of the Itô differential (5.3.15), respectively. The concept of an Itô differential is very powerful. It leads to a compact characterization that can be used to succinctly express the dynamics of rather complicated stochastic processes. Note that no Markovianity is required to characterize a process Y via its stochastic differential. This allows the modeling of very general dynamics and corresponding gains from trade.

For the above process Y, given in (5.3.12), consider the Itô integral defined in (5.3.11) with Y replacing X. Under rather general conditions it can be shown that

$$\int_0^t \xi(s)\, dY_s = \int_0^t \xi(s)\, e_s\, ds + \int_0^t \xi(s)\, f_s\, dW_s \tag{5.3.16}$$

a.s. for all $t \in [0,\infty)$, see Protter (2004). Therefore, an Itô integral of the above type can be expressed as the sum of a random ordinary Riemann-Stieltjes integral with respect to time and a standard Itô integral with respect to the Wiener process W.

These definitions and formulations extend to the case of multi-dimensional integrands ξ and integration with respect to several independent standard Wiener processes. Furthermore, they can be generalized also to hold for more

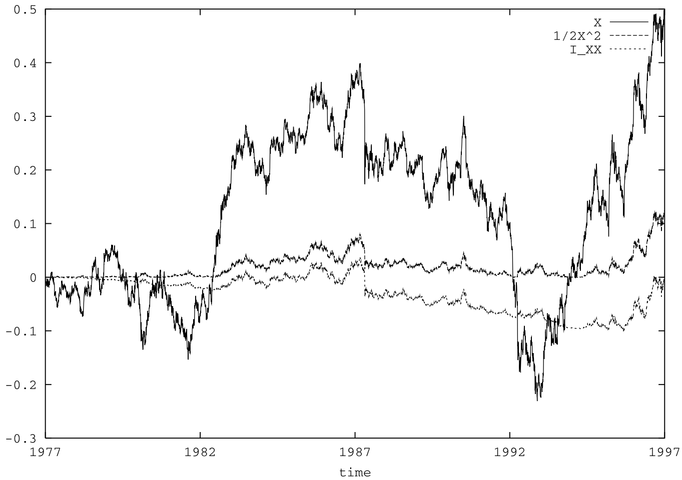

Fig. 5.3.4. Log IBM share price X_t, $\frac{1}{2}(X_t)^2$ and $I_{X,X}(t)$

general processes as integrators including those with jumps, as we shall see later.

In Fig. 5.3.4 we consider the logarithm X_t of the IBM share price between 1977 and 1997 when normalized to the value one at the beginning. Using X_t we compute also half of its squared value, that is $\frac{1}{2}(X_t)^2$, and plot these values in Fig. 5.3.4 together with the Itô integral $I_{X,X}(t)$ of X with respect to itself. One notes in Fig. 5.3.4 that there is a clear difference between the Itô integral $I_{X,X}(t)$ and what one would expect from a classical integral of a function, which would result in the value $\frac{1}{2}(X_t)^2$ at time t. The Itô integral provides here the smaller values, similar as in Fig. 5.3.3.

5.4 Itô Integral for Wiener Processes

The Itô integral exhibits a number of important properties and features that are essential in stochastic calculus and thus also for many applications in quantitative finance. The following properties will be repeatedly exploited later, for instance, in the context of pricing and hedging of derivatives.

Properties of Itô Integrals with Respect to Wiener Processes

Let us consider two $\underline{\mathcal{A}}$-adapted independent Wiener processes W^1 and W^2. Recall that $(W_t^i - W_s^i)$ is independent of $\mathcal{A}_s$ for $t \in [0,\infty)$, $s \in [0,t]$ and $i \in \{1,2\}$.

It is useful to specify for $T \in [0,\infty)$ the class $\mathcal{L}_T^2$ of predictable, square integrable stochastic processes $f = \{f_t,\, t \in [0,T]\}$ in the form that

$$\int_0^T E\left((f_t)^2\right)\, dt < \infty. \tag{5.4.1}$$

Note that it is convenient to work in a world of square integrable stochastic processes as long as this is possible for the problem at hand. Let us now summarize some fundamental properties of Itô integrals, which are essential and often used in derivations in quantitative finance.

1. *Linearity property*: For $T \in (0, \infty)$, $t \in [0, T]$, $s \in [0, t]$, $Z_1, Z_2 \in \mathcal{L}_T^2$ and $\mathcal{A}_s$-measurable, square integrable random variables A and B it is
$$\int_s^t (A\, Z_1(u) + B\, Z_2(u))\, dW_u^1 = A \int_s^t Z_1(u)\, dW_u^1 + B \int_s^t Z_2(u)\, dW_u^1. \tag{5.4.2}$$
2. *Martingale property*: For $T \in (0, \infty)$, $t \in [0, T]$, $s \in [0, t]$ and $\xi \in \mathcal{L}_T^2$ one has
$$E\left(\int_0^t \xi(u)\, dW_u^1 \,\Big|\, \mathcal{A}_s \right) = \int_0^s \xi(u)\, dW_u^1. \tag{5.4.3}$$
3. *Correlation property*: For $T \in (0, \infty)$, $t \in [0, T]$, independent Wiener processes W^1 and W^2 and $Z_1, Z_2 \in \mathcal{L}_T^2$ the conditional correlation of two Itô integrals is given by
$$E\left(\int_0^t Z_1(u)\, dW_u^i \int_0^t Z_2(u)\, dW_u^j \,\Big|\, \mathcal{A}_s \right) = \begin{cases} \int_0^t E\left(Z_1(u)\, Z_2(u) \,\Big|\, \mathcal{A}_s \right) du & \text{for } i = j \\ 0 & \text{otherwise} \end{cases} \tag{5.4.4}$$
with $i, j \in \{1, 2\}$.
4. *Covariation property*: For $t \in [0, \infty)$, independent Wiener processes W^1 and W^2 and predictable integrands Z_1 and Z_2 with $\int_0^t |Z_1(u)\, Z_2(u)|\, du < \infty$ a.s. the covariation of two Itô integrals is
$$\left[\int_0 Z_1(u)\, dW_u^i, \int_0 Z_2(u)\, dW_u^j \right]_t = \begin{cases} \int_0^t Z_1(u)\, Z_2(u)\, du & \text{for } i = j \\ 0 & \text{otherwise} \end{cases} \tag{5.4.5}$$
with $i, j \in \{1, 2\}$.
5. *Finite variation property*: For $t \in [0, \infty)$ the covariation between an Itô and a random ordinary Riemann-Stieltjes integral with respect to time vanishes. That is, for predictable Z_1 and Z_2 one has
$$\left[\int_0 Z_1(u)\, dW_u^1, \int_0 Z_2(u)\, du \right]_t = 0. \tag{5.4.6}$$

In (5.4.5) and (5.4.6) we take the upper end of the integration interval as the time parameter when forming the covariation. Using the martingale property (5.4.3) it can be shown that an Itô integral process is an $(\underline{\mathcal{A}}, P)$-martingale if the integrand is in $\mathcal{L}_T^2$. The above imposed measurability and integrability conditions can be weakened for some of the above stated properties, see Protter (2004).

The following important property of an Itô integral with respect to a Wiener process is very useful in finance. It involves again the notion of a predictable process, as was introduced in Sect. 5.1.

Lemma 5.4.1. *If ξ is predictable and it holds for this integrand that*

$$\int_0^T \xi(u)^2\,du < \infty \tag{5.4.7}$$

a.s. for all $T \in [0,\infty)$, then the corresponding Itô integral process $I_{\xi,W} = \{I_{\xi,W}(t) = \int_0^t \xi(s)dW_s,\ t \in [0,\infty)\}$ is an $(\underline{\mathcal{A}}, P)$-local martingale.

The proofs for the above properties and lemma take advantage of the properties of increments of Wiener processes and their relationship to the filtration $\underline{\mathcal{A}}$. Details can be found in Karatzas & Shreve (1991), Kloeden & Platen (1999) or Protter (2004). By application of the Statement (iii) of Lemma 5.2.2 one can derive directly the following result.

Corollary 5.4.2. *Assume that $I_{\xi,W}$ is square integrable, then $I_{\xi,W}$ is a square integrable $(\underline{\mathcal{A}}, P)$-martingale if and only if*

$$E\left(\int_0^T \xi(u)^2\,du\right) < \infty \tag{5.4.8}$$

for all $T \in [0,\infty)$.

Covariation Property

To illustrate the covariation property (5.4.5), Fig. 5.4.1 shows a sample path of a Wiener process together with an Itô integral with respect to this Wiener process using an integrand with value 10 for the first half of the period and value 1 for the rest of the period. The covariation of the Wiener process with this Itô integral is then shown in Fig. 5.4.2. Note that the slope of the covariation is proportional to the integrand of the Itô integral, as is suggested by formula (5.4.5). In Fig. 5.4.3, the quadratic variation of the Itô integral, shown in Fig. 5.4.1, is displayed. Again, as indicated by (5.4.5), the time derivative of the quadratic variation is proportional to the square of the integrand. Consequently, the slope of the quadratic variation in the second period is rather small, about 1% of that of the first period.

Itô and Deterministic Calculus

The rules that apply to Itô integrals form most of what is called the *Itô* or *stochastic calculus*. This calculus specifies rules for handling stochastic quantities, which involves integration over time. The key relationships in quantitative finance are strongly influenced by these rules.

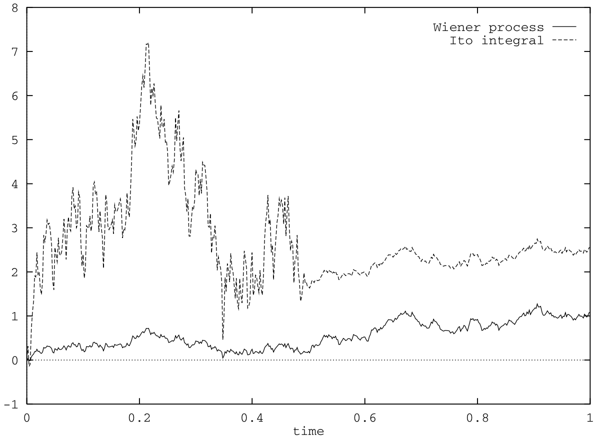

Fig. 5.4.1. Wiener process and Itô integral

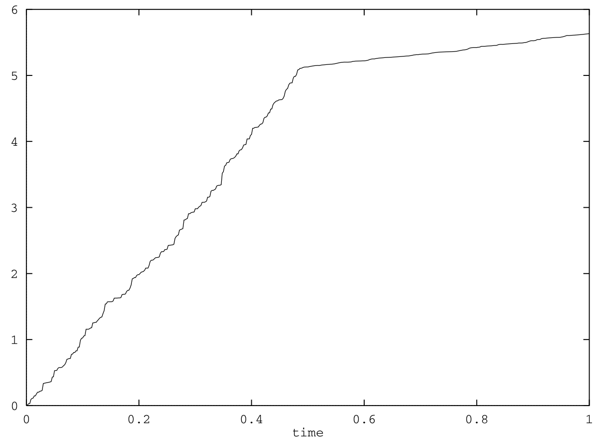

Fig. 5.4.2. Covariation of Wiener process and Itô integral

As previously mentioned the rules of Itô calculus are different from those of classical calculus, which is, in general, built on Riemann-Stieltjes integration requiring finite total variation of the integrator. The differences are primarily due to the fact that the Wiener process is of infinite total variation and has trajectories of non-zero, finite quadratic variation, see (5.2.2) and (5.2.25). Thus, the Itô integral has, in general, non-vanishing covariation between its integrand and integrator. The Wiener process and the Itô integral with respect to the Wiener process are continuous processes but not differentiable. Therefore, to ask for the slope or time derivative of a Wiener process, an Itô integral or an asset price when modeled by such process, is a meaningless question.

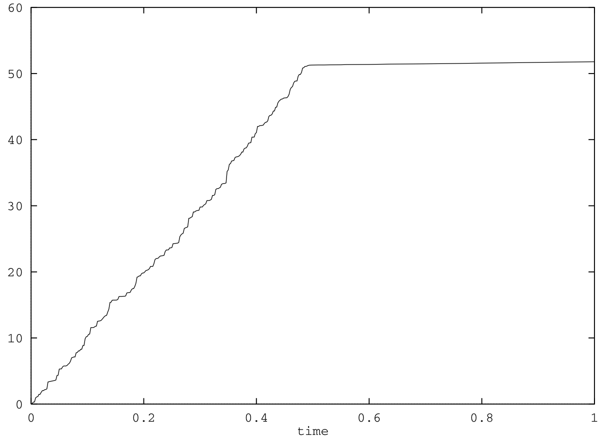

Fig. 5.4.3. Quadratic variation of an Itô integral

However as described earlier, the Itô integral is well defined without having to require differentiability of its integrator. As we shall see later, the rules of stochastic calculus provide answers to important problems in quantitative finance, such as how the pricing and hedging of a derivative can be performed or what is the typical dynamics of an asset price or optimal portfolio.

5.5 Stochastic Integrals for Semimartingales (*)

In this section we introduce general Itô integrals. These are useful for the formulation of general statements. The most general class of stochastic processes that we mention is that of *semimartingales*. For details on the following results we refer to Protter (2004).

Semimartingales (*)

From the practical point of view the following class of *semimartingales* is a very rich class of processes. It turns out to be sufficient for the modeling of most finite dimensional problems that appear in finance, insurance, portfolio optimization and other areas of risk management. As we shall see later, staying within this class, is rewarded by rather general and elegant results.

As usual, we assume a filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$ as introduced in Sect. 5.1. In the following definition of a semimartingale we refer to several notions that we have introduced earlier in this chapter.

Definition 5.5.1. *A* semimartingale *is an $\underline{\mathcal{A}}$-adapted, right-continuous stochastic process $X = \{X_t, t \in [0, \infty)\}$ with left hand limits, where X_t can be expressed as a sum of the form*

$$X_t = X_0 + A_t + M_t \tag{5.5.1}$$

for all $t \in [0, \infty)$*. Here* $A = \{A_t, t \in [0, \infty)\}$ *is a process of finite total variation and* $M = \{M_t, t \in [0, \infty)\}$ *is an* $(\underline{\mathcal{A}}, P)$*-local martingale.*

When A is predictable, then X is called a *special semimartingale* and the decomposition (5.5.1) is unique. If the discontinuous part

$$A_t^d = \sum_{0 \leq s \leq t} \Delta A_s$$

of A and the discontinuous part

$$M_t^d = \sum_{0 \leq s \leq t} \Delta M_s$$

of M are almost surely finite, then each of the processes A and M can be split into a continuous and discontinuous part, that is,

$$A_t = A_t^c + A_t^d \tag{5.5.2}$$

and

$$M_t = M_t^c + M_t^d \tag{5.5.3}$$

for $t \in [0, \infty)$, respectively.

The above defined class of semimartingales includes all stochastic processes that we have introduced so far, in particular, it covers discrete and continuous time Markov chains, diffusion processes, compound Poisson processes and Lévy processes. Note that semimartingales do not need to be Markovian.

For instance, the Wiener process $W = \{W_t, t \in [0, \infty)\}$, given in Definition 3.2.2, is a semimartingale. Here the decomposition (5.5.1) is simply so that $X_0 = 0$, $A_t = 0$ and $M_t = M_t^c = W_t$. The Wiener process is a martingale and, thus, by Definition 5.1.2 a local martingale.

A Poisson process N with intensity λ, as given by Definition 3.5.1, and the compensated Poisson process q, defined in (5.1.11), are semimartingales. For the latter, we have $q_0 = X_0 = 0$, where $M_t = M_t^d = N_t - \lambda t = q_t$ is the local martingale, which is here a martingale. Furthermore, $A_t = A_t^c = \lambda t$ characterizes the predictable process A of finite total variation.

In the case of a Lévy process X with the notation given in (3.6.2) and almost surely finite discontinuous martingale part

$$M_t^d = \int_0^t \int_{\mathcal{E}} v(p_\varphi(dv, ds) - \varphi(dv)\, ds), \tag{5.5.4}$$

the initial value is $X_0 = 0$. The continuous local martingale part of X is then of the form

$$M_t^c = \beta\, W_t \tag{5.5.5}$$

and the predictable finite total variation term is continuous and equals

$$A_t = A_t^c = \alpha\, t + \int_0^t \int_{|v|\geq 1} v\, \varphi(dv)\, ds \tag{5.5.6}$$

for $t \in [0, \infty)$, assuming A_t to be finite.

It turns out that the class of semimartingales is stable with respect to important operations and transformations. It is closed with respect to stochastic integration, which when applied to semimartingales as integrands and integrators, generates again semimartingales. Further examples of transformations that map into the class of semimartingales include the application of smooth functions, equivalent changes of measure and time changes. These properties show that the class of semimartingales is a very special class and also highly suitable for financial modeling. The class of semimartingales includes all financial models that we shall cover. However, there are non-semimartingale models being actively studied, such as those based on fractional Brownian motion, see Heyde (1999), Heyde & Liu (2001) and Elliott & van der Hoek (2003).

Itô Integral for Semimartingales (*)

For an $\underline{\mathcal{A}}$-adapted, right-continuous stochastic process $\xi = \{\xi(t),\, t \in [0, \infty)\}$ let

$$\xi(t-) \stackrel{\text{a.s.}}{=} \lim_{h \to 0+} \xi(t-h)$$

denote again the almost sure left hand limit of $\xi(t)$ at time t. Similarly as in (5.3.11), we define for semimartingales $X = \{X_t,\, t \in [0, \infty)\}$ and $\xi = \{\xi(s),\, s \in [0, \infty)\}$ the corresponding *Itô integral* as limit in probability

$$I_{\xi,X}(t) = \int_0^t \xi(s-)\, dX_s \stackrel{P}{=} \lim_{h \to 0} \sum_{k=1}^{i_t} \xi(t_{k-1})\, (X_{t_k} - X_{t_{k-1}}), \tag{5.5.7}$$

using an equidistant time discretization with step size h. What is important here is that the integrand is effectively a predictable stochastic process, see Sect. 5.1, since we take always the left hand value in a discretization interval. We could have asked ξ to be a predictable process and could then write in (5.5.7) instead of $\xi(s-)$ simply $\xi(s)$.

The Itô integral enjoys important properties. Most importantly, it is again a semimartingale. If the integrator X is an $(\underline{\mathcal{A}}, P)$-local martingale, then the Itô integral is also an $(\underline{\mathcal{A}}, P)$-local martingale if the integrand is, for example, continuous or locally bounded, see Protter (2004). In the case when X is of finite total variation, then the Itô integral coincides with the random ordinary Riemann-Stieltjes integral.

Itô Integral for Jump Processes (*)

Let us consider the case when a semimartingale X has jumps, that is, the difference

$$\Delta X_t = X_t - X_{t-} \tag{5.5.8}$$

does not vanish for all $t \in [0, \infty)$. In this case, the following important property of the jumps of the Itô integral $I_{\xi,X}(t)$ applies:

$$\Delta I_{\xi,X}(t) = I_{\xi,X}(t) - I_{\xi,X}(t-) = \xi(t-)\, \Delta X_t \tag{5.5.9}$$

for $t \in [0, \infty)$. This means that at a jump time the value of the Itô integral increases by the value of the integrand before the jump multiplied by the jump size of the integrator. For example, if N is a Poisson process, as given in Definition 3.5.1, then at its kth jump time τ_k we have

$$\Delta N_{\tau_k} = N_{\tau_k} - N_{\tau_{k-1}} = 1$$

for $k \in \{1, 2, \ldots\}$. Consequently, it follows in this case from (5.5.9) that the Itô integral for an integrand $\xi = \{\xi(t),\, t \in [0, \infty)\}$ takes simply the form

$$I_{\xi,N}(t) = \int_0^t \xi(s-)\, dN_s = \sum_{k=1}^{N_t} \xi(\tau_k-)\, \Delta N_{\tau_k} = \sum_{k=1}^{N_t} \xi(\tau_k-) \tag{5.5.10}$$

for $t \in [0, \infty)$. Consider the special case of a finite *pure jump process* $X = \{X_t = \sum_{0 \le s \le t} \Delta X_s,\, t \in [0, \infty)\}$, which jumps at the jump times $\tau_1, \tau_2, \ldots$ of a counting process $p = \{p_t,\, t \in [0, \infty)\}$ with jump size $\Delta X_{\tau_k} = c(k, \tau_k-)$, we obtain for an integrand $\xi = \{\xi(t),\, t \in [0, \infty)\}$ the Itô integral

$$I_{\xi,X}(t) = \int_0^t \xi(s-)\, dX_s = \sum_{k=1}^{p_t} \xi(\tau_k-)\, \Delta X_{\tau_k} = \sum_{k=1}^{p_t} \xi(\tau_k-)\, c(k, \tau_k-) \tag{5.5.11}$$

for $t \in [0, \infty)$. Here it is important to assume that the terms involved are almost surely finite. This means that the sums $\sum_{k=1}^{p_t} \xi(\tau_k-)\, \Delta X_{\tau_k}$ and $\sum_{k=1}^{p_t} \Delta X_{\tau_k}$ almost surely converge to a finite value for all $t \in [0, \infty)$. Note that the integral (5.5.11) covers also the cases of inaccessible, predictable, as well as, deterministic jump times. Thus, discrete time Markov chains and continuous time Markov chains are covered as integrators by the above formula.

Itô Integral for Poisson Measures (*)

In the case when jump sizes are continuously distributed, as is the case for general Lévy processes, we need to consider the stochastic integration with respect to a Poisson measure. Assume that $p_\varphi(dv, dt)$ is the Poisson measure on $\mathcal{E} \times [0, \infty)$ with intensity measure $q_\varphi(dv, dt) = \varphi(dv)\, dt$, as introduced at the end of Sect. 3.5, satisfying condition (3.5.13). Here $\mathcal{E} = \Re \backslash \{0\}$ is the mark set. We again assume that a Poisson measure is such that for all $h \in [0,\infty)$ and any set $B \in \mathcal{B}(\mathcal{E})$ the $\mathcal{A}_{t+h}$-measurable random variable $p_\varphi(B,[0,t+h]) - p_\varphi(B,[0,t])$ is independent of $\mathcal{A}_t$ for all $t \in [0,\infty)$.

In generalization of relation (5.5.9), we define for a family $(\xi(v))_{v\in\mathcal{E}}$ of a.s. finite adapted processes $\xi(v) = \{\xi(v,t),\, t \in [0,\infty)\}$ with $v \in \mathcal{E}$ the Itô integral

$$I_{\xi,p_\varphi}(t) = \int_0^t \int_{\mathcal{E}} \xi(v,s-)\, p_\varphi(dv,ds) \tag{5.5.12}$$

with respect to p_φ, such that

$$\begin{aligned} \Delta I_{\xi,p_\varphi}(t) &= \int_0^t \int_{\mathcal{E}} \xi(v,s-)\, p_\varphi(dv,ds) - \int_0^{t-} \int_{\mathcal{E}} \xi(v,s-)\, p_\varphi(dv,ds) \\ &= \int_{\mathcal{E}} \xi(v,t-)\, p_\varphi(dv,\{t\}) \end{aligned} \tag{5.5.13}$$

for all $t \in [0,\infty)$. This means that if at a jump time τ the Poisson measure p_φ generates an event with mark v, then the change of the value of the corresponding Itô integral is given by the value $\xi(v,\tau-)$ of the integrand ξ for the mark v just before the jump time. Note that we do not have to write always $\xi(v,s-)$ for the integrands in (5.5.9) and (5.5.13) if $\xi(v,\cdot)$ is predictable. However, to emphasize the fact that the integrand has in the case of a jump its value always taken before the jump time, we prefer often the above notation. We refer to Protter (2004) for more details on Itô integrals for Poisson measures.

To illustrate the above definition for the case when $\mathcal{E} = (0,\lambda)$ with $\lambda \in (0,\infty)$, where

$$\varphi(v) = \begin{cases} 1 \text{ for } \ v \in \mathcal{E} \\ 0 \text{ otherwise,} \end{cases}$$

we obtain for the special case $\xi(v,t) = 1$ the Itô integral

$$I_{1,p_\varphi}(t) = \int_0^t \int_0^\lambda p_\varphi(dv,ds) = N_t \tag{5.5.14}$$

for $t \in [0,\infty)$. Here

$$N = \{N_t = p_\varphi((0,\lambda) \times [0,t]),\, t \in [0,\infty)\} \tag{5.5.15}$$

is a Poisson process with intensity λ.

As another example let us form the Itô integral for the simple integrand $\xi(v,t) = t$ and use the previous integrator p_φ. This leads to the Itô integral

$$I_{t,p_\varphi}(t) = \int_0^t \int_0^\lambda s\, p_\varphi(dv,ds) = \sum_{k=1}^{N_t} \tau_k, \tag{5.5.16}$$

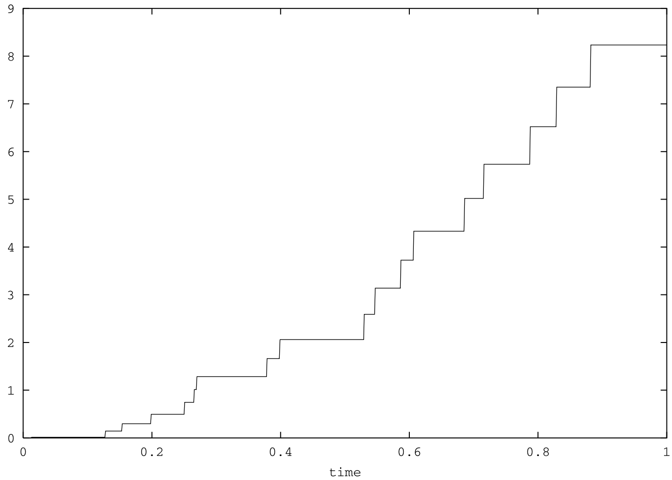

Fig. 5.5.1. Itô integral of $\xi(v,t) = t$ with respect to p_φ

which equals the sum of the jump times of the above Poisson process N given in (5.5.15). Figure 5.5.1 displays for the path of the Poisson process N, shown in Fig. 3.5.1, the resulting value of the Itô integral over time. Note that the jump sizes in Fig. 5.5.1 increase as the jump times increase.

Finally, let us discuss an example where the integrand depends on the mark v. We choose as integrand the simple function $\xi(v,t) = \frac{v}{\lambda}$. This leads to the Itô integral $I_{\frac{v}{\lambda},p_\varphi}(t)$, which is equivalent to a compound Poisson process, as defined in (3.5.9). Here we have uniformly $U(0,1)$ distributed jump sizes. An example for a path of such an Itô integral can be found in Fig. 3.5.2.

Itô Integral for a Lévy Process (*)

Similarly as above, we can introduce for a Lévy process $X = \{X_t,\, t \in [0,\infty)\}$ as integrator with decomposition (3.6.2) and for some stochastic process $\xi = \{\xi(t),\, t \in [0,\infty)\}$ the Itô integral

$$\begin{aligned}\int_0^t \xi(s-)\,dX_s = \int_0^t \xi(s)\,\alpha\,ds &+ \int_0^t \xi(s)\,\beta\,dW_s \\ &+ \int_0^t \int_{|v|<1} \xi(s-)\,v\,(p_\varphi(dv,ds) - \varphi(dv)\,ds) \\ &+ \int_0^t \int_{|v|\geq 1} \xi(s-)\,v\,p_\varphi(dv,ds) \qquad (5.5.17)\end{aligned}$$

for $t \in [0,\infty)$, see Protter (2004). Recall that the Poisson measure $p_\varphi(\cdot,\cdot)$ is specified in (3.6.2) under the condition (3.5.13). Here we have split the jump terms according to the representation (3.6.2). In the same manner as for Lévy processes one obtains the Itô integral for general semimartingales by using the decomposition (5.5.1)–(5.5.3) and calculating the different contributing terms.

5.6 Exercises for Chapter 5

5.1. If we assume that $W^i = \{W_t^i,\, t \in [0,\infty)\}$, $i \in \{1,2\}$, are standard Wiener processes, is the process $Y = \{Y_t = \alpha_1 W_t^1 + \alpha_2 W_t^2,\, t \in [0,\infty)\}$ for $\alpha_1, \alpha_2 \in \Re$ a martingale?

5.2. For a standard Wiener process W, is the process $Y = \{Y_t = (W_t)^2,\, t \in [0,\infty)\}$ a martingale, submartingale or supermartingale?

5.3. Show that $M = \{M_t = (W_t)^2 - t,\, t \in [0,\infty)\}$ is a martingale, if W is a standard Wiener process.

5.4. Let $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$ be a filtered probability space with standard Wiener process W, geometric Brownian motion $X = \{X_t = \exp\{(r - \frac{1}{2}\sigma^2)\,t + \sigma W_t\},\, t \in [0,\infty)\}$ and a money account $B = \{B_t = \exp\{r\,t\},\, t \in [0,\infty)\}$ with interest rate r. Is the discounted process $\bar{X} = \{\bar{X}_t = \frac{X_t}{B_t},\, t \in [0,\infty)\}$ a martingale, submartingale or supermartingale? Use the fact that an $N(0,1)$ distributed Gaussian random variable Y has Laplace transform

$$E(\exp\{\sigma Y\}) = \exp\left\{\frac{1}{2}\,\sigma^2\right\},$$

see (1.3.76).

5.5. Compute the quadratic variation $[Y]$ for a transformed Wiener process $Y = \{Y_t = a\,t + b\,W_t,\, t \in [0,\infty)\}$, where W is a standard Wiener process.

5.6. Determine the covariation $[Y, W]$ between the transformed Wiener process Y from Exercise 5.5 and the standard Wiener process W.

5.7. If $X = \{X_t,\, t \in [0,\infty)\}$ is a martingale and $g(\cdot)$ a convex function, is the process

$$g(X) = \{g(X_t),\, t \in [0,\infty)\}$$

a martingale, supermartingale, submartingale or none of these?

5.8. (*) Prove the martingale property for Itô integrals with piecewise constant deterministic integrands and the Wiener process as integrator.

5.9. (*) Show that the correlation property holds for Itô integrals with piecewise constant deterministic integrands and the Wiener process as integrator.

5.10. (*) Derive the linearity property for Itô integrals with piecewise constant deterministic integrands and the Wiener process as integrator.

5.11. (*) For a Lévy process $X = \{X_t,\, t \in [0,\infty)\}$ with $E(X_t \,|\, \mathcal{A}_0) = 0$ for all $t \in [0,\infty)$ prove that X is a martingale.