Chapter 3
Radically Elementary Stochastic Integrals

3.1 Martingales and Itô Integrals

For any two processes ξ, η, the *stochastic integral* of η with respect to ξ is the process $\int \eta \mathrm{d}\xi$ defined by

$$\int_0^s \eta \mathrm{d}\xi = \int_0^s \eta(t)\mathrm{d}\xi(t) = \sum_{t<s} \eta(t)\mathrm{d}\xi(t)$$

for all $s \in \mathbf{T}$. Note that $\mathrm{d}(t) = t + \mathrm{d}t - t = \mathrm{d}t$ for all $t \in \mathbf{T}\setminus\{1\}$, whence for the process $\mathrm{id} = (t)_{t\in\mathbf{T}}$ we have $\int_0^s \eta \mathrm{d}\,\mathrm{id} = \int_0^s \eta(t)\mathrm{d}(t) = \int_0^s \eta(t)\mathrm{d}t$. (Since the radically elementary approach to stochastic processes does not use *conventional* Riemann integrals, there is no danger of confusion attached to the notation $\int_0^s \eta(t)\mathrm{d}t$.)

Note that since Ω and $\mathbf{T}$ are finite, the expectation operator E and the finite integral $\int \cdot \mathrm{d}t$ always commute.

Theorem 3.1. *Let $(\mathcal{G}_t)_{t\in\mathbf{T}}$ be a filtration. A process m is a $(\mathcal{G}, P)$-martingale if and only if $\int \eta \mathrm{d}m$ is a $(\mathcal{G}, P)$-martingale for all $\mathcal{G}$-adapted m.*

Proof. The constant deterministic process $(1)_{t\in\mathbf{T}}$ is clearly adapted and m can be written as $m = \int 1\mathrm{d}m$.

Conversely, suppose m is a martingale and let η be $\mathcal{G}$-adapted. Then for all $t \in \mathbf{T}\setminus\{1\}$,

$$E\left[\eta(t)\mathrm{d}m(t)|\,\mathcal{G}_t\right] = \eta(t)\underbrace{E\left[\mathrm{d}m(t)|\,\mathcal{G}_t\right]}_{=0} = 0,$$

so $\int \eta \mathrm{d}m$ is indeed a $(\mathcal{G}, P)$-martingale. □

Stochastic integrals with respect to W are also called *Itô integrals*. A martingale with respect to $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbf{T}}$ (the filtration generated by W) is just an Itô integral of some adapted process, and vice versa:

F. S. Herzberg, *Stochastic Calculus with Infinitesimals*, Lecture Notes in Mathematics 2067, DOI 10.1007/978-3-642-33149-7_3,

Theorem 3.2 (Martingale representation theorem and converse). *A stochastic process $(m_t)_{t\in\mathbf{T}}$ is an $(\mathcal{F}, P)$-martingale if and only if there exists a unique $\mathcal{F}$-adapted process $\phi = (\phi_t)_{t\in\mathbf{T}\setminus\{1\}}$ such that for all $s \in \mathbf{T}$,*

$$m(s) = m(0) + \int_0^s \phi(t)\mathrm{d}W(t).$$

$E\left[|m(s)|^2\right]$ *is limited for all $s \in \mathbf{T}$ if and only if* $E\left[\int_0^1 \left|\phi(t)^2\right| \mathrm{d}t\right]$ *is limited.*

Proof. First, let m be a martingale. Let $t \in \mathbf{T} \setminus \{1\}$. Since m is $\mathcal{F}$-adapted, $\mathrm{d}m(t)$ is $\mathcal{F}_{t+\mathrm{d}t}$-measurable and therefore, there is some $f : \mathbf{R}^{t/dt+1} \to \mathbf{R}$ such that

$$\mathrm{d}m(t)(\omega) = f\left(\mathrm{d}W(0)(\omega), \ldots, \mathrm{d}W(t)(\omega)\right)$$

for all $\omega \in \Omega$. Therefore, exploiting that W has independent increments, each with distribution $\frac{\delta_{\sqrt{\mathrm{d}t}}+\delta_{-\sqrt{\mathrm{d}t}}}{2}$, we obtain

$$\begin{aligned}
&E\left[\mathrm{d}m(t)|\,\mathcal{F}_t\right] \\
&\quad= E\left[f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W(t-\mathrm{d}t), \sqrt{\mathrm{d}t}\right)\chi_{\left\{\mathrm{d}W(t)=\sqrt{\mathrm{d}t}\right\}}\middle|\,\mathcal{F}_t\right] \\
&\qquad+E\left[f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W(t-\mathrm{d}t), -\sqrt{\mathrm{d}t}\right)\chi_{\left\{\mathrm{d}W(t)=-\sqrt{\mathrm{d}t}\right\}}\middle|\,\mathcal{F}_t\right] \\
&\quad= f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W(t-\mathrm{d}t), \sqrt{\mathrm{d}t}\right)P\left\{\mathrm{d}W(t) = \sqrt{\mathrm{d}t}\right\} \\
&\qquad+f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W(t-\mathrm{d}t), -\sqrt{\mathrm{d}t}\right)P\left\{\mathrm{d}W(t) = -\sqrt{\mathrm{d}t}\right\} \\
&\quad= \frac{1}{2}f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W(t-\mathrm{d}t), \sqrt{\mathrm{d}t}\right) \\
&\qquad+\frac{1}{2}f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W(t-\mathrm{d}t), -\sqrt{\mathrm{d}t}\right)
\end{aligned}$$

Since m is a martingale, $E\left[\mathrm{d}m(t)|\,\mathcal{F}_t\right] = 0$, hence

$$\begin{aligned}
&f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W(t-\mathrm{d}t), \sqrt{\mathrm{d}t}\right) \\
&\quad= -f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W(t-\mathrm{d}t), -\sqrt{\mathrm{d}t}\right).
\end{aligned}$$

Defining

$$\vartheta(t) = f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W(t-\mathrm{d}t), \sqrt{\mathrm{d}t}\right),$$

we get

$$\begin{aligned}
\mathrm{d}m(t) &= f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W\left(t - \mathrm{d}t\right), \sqrt{\mathrm{d}t}\right)\chi_{\{\mathrm{d}W(t)=\sqrt{\mathrm{d}t}\}} \\
&\quad + f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W\left(t - \mathrm{d}t\right), -\sqrt{\mathrm{d}t}\right)\chi_{\{\mathrm{d}W(t)=-\sqrt{\mathrm{d}t}\}} \\
&= f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W\left(t - \mathrm{d}t\right), \sqrt{\mathrm{d}t}\right)\chi_{\{\mathrm{d}W(t)=\sqrt{\mathrm{d}t}\}} \\
&\quad - f\left(\mathrm{d}W(0), \ldots, \mathrm{d}W\left(t - \mathrm{d}t\right), \sqrt{\mathrm{d}t}\right)\chi_{\{\mathrm{d}W(t)=-\sqrt{\mathrm{d}t}\}} \\
&= \vartheta(t)\chi_{\{\mathrm{d}W(t)=\sqrt{\mathrm{d}t}\}} - \vartheta(t)\chi_{\{\mathrm{d}W(t)=-\sqrt{\mathrm{d}t}\}} \\
&= \vartheta(t)\mathrm{d}W(t)/\sqrt{\mathrm{d}t}.
\end{aligned}$$

By definition, $\vartheta(t)$ is $\mathcal{F}_t$-measurable. Hence, if we define $\phi(t) = \vartheta(t)/\sqrt{\mathrm{d}t}$, it is also $\mathcal{F}_t$-measurable and

$$\mathrm{d}m(t) = \phi(t)\mathrm{d}W(t).$$

Since t was arbitrary, this holds for any t, and yields, after writing $m(s) - m(0)$ as a telescoping sum,

$$m(s) = m(0) + \sum_{t<s} \mathrm{d}m(t) = m(0) + \sum_{t<s} \phi(t)\mathrm{d}W(t).$$

If there were another process $\tilde{\phi} = \left(\tilde{\phi}_t\right)_{t\in\mathbf{T}\setminus\{1\}}$ such that $\int \tilde{\phi}\mathrm{d}W = m = \int \phi\mathrm{d}W$, then for all $t \in \mathbf{T} \setminus \{1\}$,

$$\tilde{\phi}(t)\mathrm{d}W(t) = \mathrm{d}m(t) = \phi(t)\mathrm{d}W(t),$$

whence $\tilde{\phi}(t) = \phi(t)$ since $\mathrm{d}W(t) = \pm\sqrt{\mathrm{d}t} \neq 0$, therefore $\tilde{\phi} = \phi$, proving the uniqueness of ϕ.

Conversely, suppose $m(s) = m(0) + \int_0^s \phi(t)\mathrm{d}W(t)$ for all $s \in \mathbf{T}$ for some $\mathcal{F}$-adapted ϕ. The definition of the stochastic integral and the $\mathcal{F}$-adaptedness of W imply that m is $\mathcal{F}$-adapted. It remains to be shown that $E\left[\mathrm{d}m(t)|\,\mathcal{F}_t\right] = 0$ for all $t \in \mathbf{T} \setminus \{1\}$. This is straightforward:

$$E\left[\mathrm{d}m(t)|\,\mathcal{F}_t\right] = E\left[\phi(t)\mathrm{d}W(t)|\,\mathcal{F}_t\right] = \phi(t)E\left[\mathrm{d}W(t)|\,\mathcal{F}_t\right] = \phi(t)\underbrace{E\left[\mathrm{d}W(t)\right]}_{=0}.$$

By the Itô isometry (Lemma 3.4), one has

$$E\left[|m(s)|^2\right] = E\left[\int_0^s \left|\phi(t)^2\right|\mathrm{d}t\right],$$

and the right-hand side is monotonely increasing in s. Hence, $E\left[|m(s)|^2\right]$ is limited for all $s \in \mathbf{T}$ if and only if $E\left[\int_0^s \left|\phi(t)^2\right| \mathrm{d}t\right]$ is. □

Definition 3.3. A stochastic process $\xi = (\xi(t))_{t\in\mathbf{T}}$ is called a *normalized martingale* (or just *normalized*) if and only if

$$\forall t \in \mathbf{T}\setminus\{1\} \qquad E\left[\mathrm{d}\xi(t)|\,\mathcal{F}_t\right] = 0, \qquad E\left[(\mathrm{d}\xi(t))^2\middle|\,\mathcal{F}_t\right] = \mathrm{d}t.$$

The Wiener walk W, for example, is normalized.

Lemma 3.4 (Radically elementary Itô isometry). *Let m be a normalized martingale and η be an $\mathcal{F}$-adapted stochastic process. Then for all $s, v \in \mathbf{T}$ with $s \geq v$,*

$$E\left[\left|\int_v^s \eta(t)\mathrm{d}m(t)\right|^2\middle|\,\mathcal{F}_v\right] = E\left[\int_v^s \eta(t)^2\mathrm{d}t\middle|\,\mathcal{F}_v\right].$$

Proof.

$$\begin{aligned}
E\left[\left|\int_v^s \eta(t)\mathrm{d}m(t)\right|^2\middle|\,\mathcal{F}_v\right] &= E\left[\left(\sum_{t=v}^{s}\eta(t)\mathrm{d}m(t)\right)^2\middle|\,\mathcal{F}_v\right] \\
&= 2\sum_{v\leq u<t<s} E\left[\eta(t)\mathrm{d}m(t)\eta(u)\mathrm{d}m(u)|\,\mathcal{F}_v\right] + \sum_{v\leq t<s} E\left[\eta(t)^2\mathrm{d}m(t)^2\middle|\,\mathcal{F}_v\right] \\
&= 2\sum_{v\leq u<t<s} E\left[E\left[\eta(t)\mathrm{d}m(t)\eta(u)\mathrm{d}m(u)|\,\mathcal{F}_t\right]|\,\mathcal{F}_v\right] \\
&\quad + \sum_{v\leq t<s} E\left[E\left[\eta(t)^2\mathrm{d}m(t)^2\middle|\,\mathcal{F}_t\right]\middle|\,\mathcal{F}_v\right] \\
&= 2\sum_{v\leq u<t<s} E\left[\eta(t)E\left[\mathrm{d}m(t)|\,\mathcal{F}_t\right]\eta(u)\mathrm{d}m(u)|\,\mathcal{F}_v\right] \\
&\quad + \sum_{v\leq t<s} E\left[\eta(t)^2 E\left[\mathrm{d}m(t)^2\middle|\,\mathcal{F}_t\right]\middle|\,\mathcal{F}_v\right] \\
&= 0 + \sum_{v\leq t<s} E\left[\eta(t)^2\mathrm{d}t\middle|\,\mathcal{F}_v\right] = E\left[\sum_{v\leq t<s}\eta(t)^2\mathrm{d}t\middle|\,\mathcal{F}_v\right] \\
&= E\left[\int_v^s \eta(t)^2\mathrm{d}t\middle|\,\mathcal{F}_v\right].
\end{aligned}$$

□

Remark 3.5. Nelson [60, deliberations following Theorem 13.1, p. 55] has shown that if m is a normalized martingale such that $\mathrm{d}m(t)$ is infinitesimal for all $t \in \mathbf{T} \setminus \{1\}$, then m is P-a.s. continuous.

3.2 Radically Elementary Itô Processes

An Itô process is essentially an Itô integral plus an absolutely continuous process.

Definition 3.6. Let $\bar{W}$ be a Wiener process, let $\xi(0) \in \mathbf{R}$, and let $\mu = (\mu(t))_{t \in \mathbf{T}\setminus\{1\}}$ and $\sigma = (\sigma(t))_{t \in \mathbf{T}\setminus\{1\}}$ be two $\mathcal{F}$-adapted processes. A stochastic process ξ is called an *Itô process on* (Ω, P) *with respect to* $\bar{W}$ *and with drift coefficient* μ, *diffusion coefficient* σ *and initial value* $\xi(0)$ if and only if

$$\xi(s) = \xi(0) + \int_0^s \mu(t)\mathrm{d}t + \int_0^s \sigma(t)\mathrm{d}\bar{W}(t)$$

for all $s \in \mathbf{T}$. The equation

$$\forall t \in \mathbf{T} \setminus \{1\} \qquad \mathrm{d}\xi(t) = \mu(t)\mathrm{d}t + \sigma(t)\mathrm{d}\bar{W}(t)$$

is called the *stochastic differential equation* solved by ξ.

The representation of an Itô process in the form $\xi = \xi(0) + \int \mu(t)\mathrm{d}t + \int \sigma(t)\mathrm{d}W(t)$ is called *Itô decomposition*. Under certain assumptions, the Itô decomposition is essentially unique. We give a proof under fairly restrictive assumptions (recall that ν denotes the normalized counting measure on $\mathbf{T} \setminus \{T\}$):

Theorem 3.7 (Uniqueness of the Itô decomposition). *Let $\mu_1, \mu_2, \sigma_1, \sigma_2$ be $\mathcal{F}$-adapted processes. Suppose for all $t \in \mathbf{T} \setminus \{1\}$, we have*

$$\mu_1(t)\mathrm{d}t + \sigma_1(t)\mathrm{d}W(t) = \mu_2(t)\mathrm{d}t + \sigma_2(t)\mathrm{d}W(t) + R\,(t + \mathrm{d}t)\,(\mathrm{d}t)^{3/2} \tag{3.1}$$

for some $R(t+\mathrm{d}t)$ such that $E\left[\int_0^1 R(t+\mathrm{d}t)^2\mathrm{d}t\right]$ is limited.[1] *Assume $E\left[\int_0^1 |\mu_1(t) - \mu_2(t)|^2\mathrm{d}t\right]$ is limited. Then for P-a.e. $\omega \in \Omega$ and ν-a.e. $t \in \mathbf{T} \setminus \{1\}$,*

$$\sigma_1(t)(\omega) \simeq \sigma_2(t)(\omega), \qquad \mu_1(t)(\omega) \simeq \mu_2(t)(\omega).$$

Proof. Put $\mu = \mu_1 - \mu_2$ and $\sigma = \sigma_1 - \sigma_2$. We need to verify that for P-a.e. $\omega \in \Omega$ and ν-a.e. $t \in \mathbf{T} \setminus \{1\}$, $\sigma(t)(\omega) \simeq 0 \simeq \mu(t)(\omega)$.

[1] We denote this random variable by $R\,(t + \mathrm{d}t)$ rather than $R\,(t)$ because it is $\mathcal{F}_{t+\mathrm{d}t}$-measurable, but in general not $\mathcal{F}_t$-measurable.

For this purpose, first note that by definition of μ and σ and by the assumption (3.1) in the Theorem, we have for all $t \in \mathbf{T} \setminus \{1\}$,

$$\mu(t)\mathrm{d}t = \sigma(t)\mathrm{d}W(t) + R\,(t + \mathrm{d}t)\,(\mathrm{d}t)^{3/2}. \tag{3.2}$$

Squaring both sides of this equality and afterwards rearranging terms yields $\mu(t)^2(\mathrm{d}t)^2 - R\,(t+\mathrm{d}t)^2\,(\mathrm{d}t)^3 - 2R\,(t+\mathrm{d}t)\,(\mathrm{d}t)^{3/2}\mathrm{d}W(t) = \sigma(t)^2\mathrm{d}t$, hence (dropping nonnegative terms and using the triangle inequality)

$$\begin{aligned}
E\left[\int_0^s \sigma(t)^2\mathrm{d}t\right] &= \int_0^s E\left[\sigma(t)^2\right]\mathrm{d}t \\
&= \sum_{t<s} E\left[\mu(t)^2\right](\mathrm{d}t)^2 - \sum_{t<s} E\left[R(t+\mathrm{d}t)^2\right](\mathrm{d}t)^3 \\
&\quad -2\sum_{t<s} E\left[R\,(t+\mathrm{d}t)\,\mathrm{d}W(t)\right](\mathrm{d}t)^{3/2} \\
&\leq \sum_{t<s} E\left[\mu(t)^2\right](\mathrm{d}t)^2 - 2\sum_{t<s} E\left[R\,(t+\mathrm{d}t)\,\mathrm{d}W(t)\right](\mathrm{d}t)^{3/2} \\
&\leq \sum_{t<s} E\left[\mu(t)^2\right](\mathrm{d}t)^2 \\
&\quad +2\sum_{t<s} \left|E\left[R\,(t+\mathrm{d}t)\,\mathrm{d}W(t)\right]\right|(\mathrm{d}t)^{3/2}.
\end{aligned}$$

However, the last expression can be estimated, due to Jensen's inequality, as follows:

$$\left|E\left[R\,(t+\mathrm{d}t)\,\mathrm{d}W(t)\right]\right| \leq E\left[R\,(t+\mathrm{d}t)^2\,\mathrm{d}t\right]^{1/2},$$

so we actually have shown that

$$\begin{aligned}
E\left[\int_0^s \sigma(t)^2\mathrm{d}t\right] \leq E\left[\int_0^s \mu(t)^2\mathrm{d}t\right]\mathrm{d}t& \\
+2\sum_{t<s}\left(E\left[R\,(t+\mathrm{d}t)^2\right]\mathrm{d}t\right)^{1/2}(\mathrm{d}t)^{3/2}.&
\end{aligned} \tag{3.3}$$

In order to further simplify the right-hand side, we apply Jensen's inequality again (this time for the average on $\mathbf{T} \cap [0,s)$ as expectation operator):

$$\begin{aligned}
&\sum_{t<s}\left(E\left[R\,(t+\mathrm{d}t)^2\right]\mathrm{d}t\right)^{1/2} \\
&= \mathrm{card}\,(\mathbf{T}\cap[0,s))\,\frac{1}{\mathrm{card}\,(\mathbf{T}\cap[0,s))}\sum_{t<s}\left(E\left[R\,(t+\mathrm{d}t)^2\right]\mathrm{d}t\right)^{1/2}
\end{aligned}$$

$$\leq \operatorname{card}(\mathbf{T} \cap [0,s)) \left(\frac{1}{\operatorname{card}(\mathbf{T} \cap [0,s))} \sum_{t<s} E\left[R(t+\mathrm{d}t)^2\right] \mathrm{d}t \right)^{1/2}$$

$$= \operatorname{card}(\mathbf{T} \cap [0,s))^{1/2} E\left[\int_0^s R(t+\mathrm{d}t)^2 \,\mathrm{d}t\right]^{1/2}.$$

Inserting this into Eq. (3.3) and exploiting that $\operatorname{card}(\mathbf{T} \cap [0,s)) = s/\mathrm{d}t \leq 1/\mathrm{d}t$, hence $\operatorname{card}(\mathbf{T} \cap [0,s))^{1/2} = (\mathrm{d}t)^{-1/2}$, we conclude that

$$E\left[\int_0^s \sigma(t)^2 \mathrm{d}t\right] \leq E\left[\int_0^1 \mu(t)^2 \mathrm{d}t\right] \mathrm{d}t + 2E\left[\int_0^1 R(t+\mathrm{d}t)^2 \,\mathrm{d}t\right]^{1/2} \mathrm{d}t.$$

However, by assumption, both $E\left[\int_0^1 R(t+\mathrm{d}t)^2 \,\mathrm{d}t\right]$ and $E\left[\int_0^1 \mu(t)^2 \mathrm{d}t\right]$ are limited, whence

$$E\left[\int_0^s \sigma(t)^2 \mathrm{d}t\right] = \mathcal{O}(\mathrm{d}t) \simeq 0.$$

This entails that for P-a.e. $\omega \in \Omega$ and ν-a.e. $t \in \mathbf{T} \setminus \{1\}$, $\sigma(t)(\omega) \simeq 0$ (by Theorem 2.7).

In order to complete the proof, we also need to verify that $\mu(t)(\omega) \simeq 0$ for P-a.e. $\omega \in \Omega$ and ν-a.e. t. To achieve this, we first compute (the conditional expectation of) $\mu(t)\mathrm{d}t$. Now, according to Eq. (3.2), the latter term is the same as $\sigma(t)\mathrm{d}W(t) + R(t+\mathrm{d}t)(\mathrm{d}t)^{3/2}$, hence, using the $\mathcal{F}_t$-linearity of the operator $E[\cdot|\,\mathcal{F}_t]$, we get

$$\begin{aligned}\mu(t)\mathrm{d}t = E[\mu(t)\mathrm{d}t|\,\mathcal{F}_t] &= \sigma(t) \underbrace{E[\mathrm{d}W(t)|\,\mathcal{F}_t]}_{=0} \\ &\quad + E[R(t+\mathrm{d}t)|\,\mathcal{F}_t](\mathrm{d}t)^{3/2}.\end{aligned}$$

Therefore, $\mu(t) = E[R(t+\mathrm{d}t)|\,\mathcal{F}_t](\mathrm{d}t)^{1/2}$, hence (applying the conditional Jensen inequality)

$$\begin{aligned}\mu(t)^2 &= E[R(t+\mathrm{d}t)|\,\mathcal{F}_t]^2 \,\mathrm{d}t \\ &\leq E\left[R(t+\mathrm{d}t)^2 \middle|\, \mathcal{F}_t\right] \mathrm{d}t.\end{aligned}$$

It follows that

$$\begin{aligned}E\left[\int_0^1 \mu(t)^2 \mathrm{d}t\right] &\leq E\left[\int_0^1 E\left[R(t+\mathrm{d}t)^2 \middle|\, \mathcal{F}_t\right] \mathrm{d}t\right] \mathrm{d}t \\ &= \int_0^1 E\left[E\left[R(t+\mathrm{d}t)^2 \middle|\, \mathcal{F}_t\right] \mathrm{d}t\right] \mathrm{d}t\end{aligned}$$

$$= \int_0^1 E\left[R\,(t + \mathrm{d}t)^2 \,\mathrm{d}t\right] \mathrm{d}t$$

$$= E\left[\int_0^1 R\,(t + \mathrm{d}t)^2 \,\mathrm{d}t\right] \mathrm{d}t.$$

Since $E\left[\int_0^1 R\,(t+\mathrm{d}t)^2 \,\mathrm{d}t\right]$ was assumed to be limited, we deduce $E\left[\int_0^1 \mu(t)^2 \mathrm{d}t\right] = \mathcal{O}(\mathrm{d}t) \simeq 0$. This, however, means—again by Theorem 2.7—that $\mu(t)(\omega) \simeq 0$ for P-a.e. $\omega \in \Omega$ and ν-a.e. $t \in \mathbf{T} \setminus \{1\}$. □

For special Itô processes one can prove their a.s. limitedness:

Lemma 3.8. *If ξ is an Itô process with respect to W, with limited initial value $\xi(0)$, with drift coefficient μ and diffusion coefficient σ. Suppose that $E\left[\int_0^1 \sigma(t)^2 \mathrm{d}t\right]$ is limited and that μ is a.s. limited. Then ξ is a.s. limited.*

Proof. Since μ is a.s. limited, it follows that a.s. $\int_0^s \mu(t)\mathrm{d}t$ is limited (because a.s. $\max_{t\in\mathbf{T}\setminus\{1\}} |\mu(t)|$ is limited and $\max_{s\in\mathbf{T}} \left|\int_0^s \mu(t)\mathrm{d}t\right| \leq \max_{t\in\mathbf{T}\setminus\{1\}} |\mu(t)|$), and hence so is $\xi(0)+\int_0^s \mu(t)\mathrm{d}t$. What remains to be shown is that $\int \sigma \mathrm{d}W$ is a.s. limited. However,

$$E\left[\left|\int_0^1 \sigma \mathrm{d}W\right|^2\right] = E\left[\int_0^1 \sigma(t)^2 \mathrm{d}t\right]$$

by the Itô isometry (Lemma 3.4), hence by the Cauchy–Schwarz inequality,

$$E\left[\left|\int_0^1 \sigma \mathrm{d}W\right|\right] \leq E\left[\int_0^1 \sigma(t)^2 \mathrm{d}t\right]^{1/2},$$

and the right-hand side is limited by assumption. Since $\int \sigma \mathrm{d}W$ is a martingale (Theorem 3.2), we may apply the corollary to Nelson's martingale inequality (Corollary 2.13) and obtain that $\int \sigma \mathrm{d}W$ is a.s. limited. Since we have already seen that $\int \mu(t)\mathrm{d}t$ is a.s. limited, we conclude that ξ is a.s. limited. □

3.3 A Basic Radically Elementary Itô Formula

A function $f : \mathbf{R} \to \mathbf{R}$ is said to be *uniformly limited* if and only if there is some limited real C such that $|f(x)| \leq C$ for all $x \in \mathbf{R}$. f is said to be *limited* if and only if $f(x)$ is limited for all limited $x \in \mathbf{R}$.

If $\omega \in \Omega$ and ξ is a stochastic process, then $\xi(\omega)$ will also be called the *ω-trajectory of* ξ; a trajectory $\lambda : \mathbf{T}' \to \mathbf{R}$ is said to be *limited* if and only if $\lambda(t)$ is limited for all $t \in \mathbf{T}'$.

Let now $p \in \mathbf{R}$. A trajectory $\lambda : \mathbf{T}' \to \mathbf{R}$ is said to be $\mathfrak{o}\left((\mathrm{d}t)^p\right)$ (limited, respectively) if and only if $\max_{t\in\mathbf{T}'} |\lambda(t)|$ is $\mathfrak{o}\left((\mathrm{d}t)^p\right)$ (limited, respectively).

The following result, a basic radically elementary version of the Itô–Doeblin formula, is essentially due to Benoît [10, Proposition 4.6.1]. It allows to calculate the increment process of a function of a Wiener walk plus linear drift.

Lemma 3.9 (Itô–Doeblin formula for Wiener walks with additive linear drift). *Let $L(t) = \mu t + \sigma W(t)$ for all $t \in \mathbf{T}$ for limited $\mu, \sigma \in \mathbf{R}$, and let f be a thrice continuously differentiable function. Then for every $s \in \mathbf{T}$ and every ω such that the ω-trajectories of $f''(L)$ and $f'''(L)$ are $\mathfrak{o}\left((\mathrm{d}t)^{-1/2}\right)$,*

$$f(L(s)(\omega)) - f(L(0)(\omega)) \simeq \int_0^s f'(L(t)(\omega))\mathrm{d}L(t)(\omega) + \frac{\sigma^2}{2}\int_0^s f''(L(t)(\omega))\mathrm{d}t. \tag{3.4}$$

In particular, if f'' and f''' are uniformly limited, then the above formula (3.4) holds for all $\omega \in \Omega$.

Proof. Let us suppress the argument ω. Fix $t \in \mathbf{T} \setminus \{1\}$. Then, by the third-order Taylor formula,

$$\mathrm{d}f(L(t)) = f'(L(t))\mathrm{d}L(t) + \frac{1}{2}f''(L(t))\left(\mathrm{d}L(t)\right)^2 + \frac{1}{6}f'''(\zeta(t))\left(\mathrm{d}L(t)\right)^3,$$

for some $\zeta(t) \in [L(t), L(t+\mathrm{d}t)]$. By assumption on L,

$$(\mathrm{d}L(t))^2 = \mu^2(\mathrm{d}t)^2 + 2\mu\sigma\mathrm{d}t\,\mathrm{d}W(t) + \sigma^2\mathrm{d}t,$$

hence

$$(\mathrm{d}L(t))^3 = \left(\mu^2\mathrm{d}t + 2\mu\sigma\mathrm{d}W(t) + \sigma^2\right)^{3/2}(\mathrm{d}t)^{3/2}.$$

By assumption,

$$\max_{t\in\mathbf{T}\cap[0,s]} \left|f''(L(t))\right| \vee \left|f'''(L(t))\right| = \mathfrak{o}\left((\mathrm{d}t)^{-1/2}\right),$$

so

$$\begin{aligned}
&\left|f(L(s)) - f(L(0)) - \int_0^s f'(L(t))\mathrm{d}L(t) - \frac{\sigma^2}{2}\int_0^s f''(L(t))\mathrm{d}t\right| \\
&= \left|\sum_{t<s} \mathrm{d}f(L(t))\, f'(L(t))\mathrm{d}L(t) - \frac{\sigma^2}{2}f''(L(t))\mathrm{d}t\right| \\
&= \left|\sum_{t<s}\left(\frac{1}{2}f''(L(t))\left(\mu^2(\mathrm{d}t)^2 + 2\mu\sigma\mathrm{d}t\,\mathrm{d}W(t)\right)\right.\right.
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{1}{6} f'''(\zeta(t)) \left(\mu^2 \mathrm{d}t + 2\mu\sigma \mathrm{d}W(t) + \sigma^2\right)^{3/2} (\mathrm{d}t)^{3/2} \Big) \Big| \\
&\leq C \sum_{t<s} \frac{1}{2} \left|\mu^2 (\mathrm{d}t)^2 + 2\mu\sigma \mathrm{d}t \, \mathrm{d}W(t)\right| \\
&\quad + \frac{1}{6} \left| \left(\mu^2 \mathrm{d}t + 2\mu\sigma \mathrm{d}W(t) + \sigma^2\right)^{3/2} (\mathrm{d}t)^{3/2} \right| \\
&\leq C \frac{s}{\mathrm{d}t} \left(\frac{1}{2} \mu^2 (\mathrm{d}t)^2 + |\mu\sigma| \mathrm{d}t \sqrt{\mathrm{d}t} + \frac{1}{6} \left(\mu^2 \mathrm{d}t + 2|\mu\sigma| \sqrt{\mathrm{d}t} + \sigma^2 \right)^{3/2} (\mathrm{d}t)^{3/2} \right) \\
&\leq C s \left(\frac{1}{2} \mu^2 \mathrm{d}t + |\mu\sigma| \sqrt{\mathrm{d}t} + \frac{1}{6} \left(\mu^2 \mathrm{d}t + 2|\mu\sigma| \sqrt{\mathrm{d}t} + \sigma^2 \right)^{3/2} \sqrt{\mathrm{d}t} \right) \\
&\simeq 0.
\end{aligned}
$$

□

In applications, one will rather often not be able to literally apply this version of the Itô–Doeblin formula in Lemma 3.9, as it is usually not obvious how to establish sufficient upper bounds on $f''(L)$ of $f'''(L)$. Nevertheless, the proof idea—i.e. a third-order Taylor expansion—will usually be applicable even in those settings. An important example will be studied in Sect. 3.4 of Chap. 3, which is concerned with a particularly simple class of Itô processes.

3.4 Analytic Excursion: A Radically Elementary Treatment of Geometric Itô Processes with Monotone Drift

Geometric Itô processes are processes which satisfy a stochastic differential equation of the form

$$
\forall t \in \mathbf{T} \setminus \{1\} \qquad \mathrm{d}\xi(t) = \xi(t)\mu(t)\mathrm{d}t + \xi(t)\sigma(t)\mathrm{d}W(t) \tag{3.5}
$$

for some limited $\xi(0)$. For limited μ, σ, one can show that $\xi(t) > 0$ for all $t \in \mathbf{T}$. (See the proof of Lemma 3.10.) Hence, whenever $\mu(t) \geq 0$ for all $t \in \mathbf{T}$ or $\mu(t) \leq 0$ for all $t \in \mathbf{T}$, the drift coefficient of the Itô process ξ will be either monotonely increasing or decreasing in t (for every fixed $\omega \in \Omega$).

Such processes are of paramount importance in applications of Girsanov's theorem, in particular to mathematical finance, and therefore merit to be studied in some detail. (For instance, the radically elementary analogue of the stock price process of the classical Black–Scholes [18] model satisfies Eq. (3.5) for constant limited μ, σ.) However, the main parts of the book—in particular our version of Girsanov's theorem—do not depend on the results of this Sect. 3.4.

Lemma 3.10. *Let μ, σ be $\mathcal{F}$-adapted limited processes, let ξ be the process given by*

$$\forall t \in \mathbf{T} \setminus \{1\} \qquad \mathrm{d}\xi(t) = \xi(t)\mu(t)\mathrm{d}t + \xi(t)\sigma(t)\mathrm{d}W(t)$$

for some limited $\xi(0) \in \mathbf{R}_{>0}$. Suppose that either $\mu(t) \geq 0$ for all $t \in \mathbf{T}$ or $\mu(t) \leq 0$ for all $t \in \mathbf{T}$. Then, for all $s \in \mathbf{T}$, $\xi(s)$ is $L^1(P)$ with limited second moment. Moreover, with probability 1, one has $\xi(s) > 0$ for all $s \in \mathbf{T}$.

The proof uses a radically elementary analogue of the Harnack inequality.

Lemma 3.11 (Harnack inequality). *Let $\alpha, \gamma \in \mathbf{R}_{>0}$ and $v : \mathbf{T} \to \mathbf{R}$. If*

$$\forall s \in \mathbf{T} \qquad v(s) \leq \alpha + \gamma \int_0^s v(t)\mathrm{d}t,$$

then

$$\forall s \in \mathbf{T} \qquad v(s) \leq \alpha \mathrm{e}^{\gamma s}.$$

Proof of the Harnack inequality. The proof proceeds by induction on $s \in \mathbf{T}$. Let $C = \mathrm{e}^{\gamma}$ and suppose $v(s) \leq \alpha C^t$ for all $t < s$. Then, using that $\mathrm{e}^{\gamma\,\mathrm{d}t} = \sum_{n=0}^{\infty} \frac{\gamma^n (\mathrm{d}t)^n}{n!} \geq 1 + \gamma\,\mathrm{d}t$, one obtains

$$\begin{aligned}
v(s) &= \alpha + \gamma \int_0^s \alpha C^t \mathrm{d}t = \alpha + \gamma\alpha \sum_{\ell=0}^{s/\mathrm{d}t - 1} C^{\ell\,\mathrm{d}t}\mathrm{d}t \\
&= \alpha + \gamma\alpha \frac{C^s - 1}{C^{\mathrm{d}t} - 1}\mathrm{d}t = \alpha\left(1 + \gamma\,\mathrm{d}t\,\frac{\mathrm{e}^{\gamma s} - 1}{\mathrm{e}^{\gamma\,\mathrm{d}t} - 1}\right) \\
&\leq \alpha\left(1 + \gamma\,\mathrm{d}t\,\frac{\mathrm{e}^{\gamma s} - 1}{\gamma\,\mathrm{d}t}\right) \\
&= \alpha \mathrm{e}^{\gamma s} = \alpha C^s.
\end{aligned}$$

□

Proof of Lemma 3.10. Since $\mu, \sigma, \xi(0)$ are limited, there must be some limited $C \in \mathbf{R}_{>0}$ such that $|\mu(t)| \vee |\sigma(t)| \vee \xi(0) \leq C$ for all $t \in \mathbf{T}$. Combining this estimate with the fact that Itô integrals are martingales (Theorem 3.2), the Itô isometry (Lemma 3.4) and the Cauchy–Schwarz inequality, we may calculate for all $s \in \mathbf{T}$,

$$\begin{aligned}
&E\left[\xi(s)^2\right] \\
&= \xi(0)^2 + 2\xi(0)\left(E\left[\int_0^s \xi(t)\mu(t)\mathrm{d}t\right] + \underbrace{E\left[\int_0^s \xi(t)\sigma(t)\mathrm{d}W(t)\right]}_{=0}\right)
\end{aligned}$$

$$
\begin{aligned}
&+E\left[\left|\int_0^s \xi(t)\mu(t)\mathrm{d}t\right|^2\right] + E\left[\left|\int_0^s \xi(t)\sigma(t)\mathrm{d}W(t)\right|^2\right] \\
&+2E\left[\left(\int_0^s \xi(t)\mu(t)\mathrm{d}t\right)\left(\int_0^s \xi(t)\sigma(t)\mathrm{d}W(t)\right)\right] \\
\le\ & \xi(0)^2 + 2\xi(0)\, E\left[\left|\int_0^s \xi(t)\mu(t)\mathrm{d}t\right|^2\right]^{1/2} \\
&+E\left[\left|\int_0^s \xi(t)\mu(t)\mathrm{d}t\right|^2\right] + E\left[\int_0^s \xi(t)^2\sigma(t)^2\mathrm{d}t\right] \\
&+2E\left[\left|\int_0^s \xi(t)\mu(t)\mathrm{d}t\right|^2\right]^{1/2} E\left[\left|\int_0^s \xi(t)\sigma(t)\mathrm{d}W(t)\right|^2\right]^{1/2} \\
\le\ & \xi(0)^2 + 2\xi(0)\, E\left[\left|\int_0^s \xi(t)\mu(t)\mathrm{d}t\right|^2\right]^{1/2} \\
&+E\left[\left|\int_0^s \xi(t)\mu(t)\mathrm{d}t\right|^2\right] + E\left[\int_0^s \xi(t)^2\sigma(t)^2\mathrm{d}t\right] \\
&+2E\left[\left|\int_0^s \xi(t)\mu(t)\mathrm{d}t\right|^2\right]^{1/2} E\left[\int_0^s \xi(t)^2\sigma(t)^2\mathrm{d}t\right]^{1/2}.
\end{aligned}
$$

Note that $\frac{1}{s}\int_0^s \cdot\mathrm{d}t$ defines an expectation operator on $\mathbf{T} \cap [0, s)$. Applying Jensen's inequality, we find for arbitrary η and $s \in \mathbf{T}$,

$$
\begin{aligned}
\left|\frac{1}{s}\int_0^s \eta(t)\mathrm{d}t\right|^2 &= s^2\left|\frac{1}{s}\int_0^s \eta(t)\mathrm{d}t\right|^2 \\
&\le s^2\frac{1}{s}\int_0^s \eta(t)^2\mathrm{d}t = s\int_0^s \eta(t)^2\mathrm{d}t \\
&\le \int_0^s \eta(t)^2\mathrm{d}t.
\end{aligned}
$$

Applying this to $\eta = \xi\mu$ in the above estimates, we obtain

$$
\begin{aligned}
E\left[\xi(s)^2\right] \le \xi(0)^2 + 2\xi(0)\, & E\left[\int_0^s \xi(t)^2\mu(t)^2\mathrm{d}t\right]^{1/2} \\
&+E\left[\int_0^s \xi(t)^2\mu(t)^2\mathrm{d}t\right] + E\left[\int_0^s \xi(t)^2\sigma(t)^2\mathrm{d}t\right]
\end{aligned}
$$

$$
\begin{aligned}
&+2E\left[\int_0^s \xi(t)^2\mu(t)^2 \mathrm{d}t\right]^{1/2} E\left[\int_0^s \xi(t)^2\sigma(t)^2 \mathrm{d}t\right]^{1/2} \\
&\leq C^2 + 2C^2\, E\left[\int_0^s \xi(t)^2 \mathrm{d}t\right]^{1/2} \\
&+C^2\, E\left[\int_0^s \xi(t)^2 \mathrm{d}t\right] + C^2\, E\left[\int_0^s \xi(t)^2 \mathrm{d}t\right] \\
&+2C^2\, E\left[\int_0^s \xi(t)^2 \mathrm{d}t\right]^{1/2} E\left[\int_0^s \xi(t)^2 \mathrm{d}t\right]^{1/2} \\
&\leq C^2 + 2C^2\, E\left[\int_0^s \xi(t)^2 \mathrm{d}t\right]^{1/2} + 4C^2\, E\left[\int_0^s \xi(t)^2 \mathrm{d}t\right].
\end{aligned}
$$

Now, clearly $x^{1/2} \leq 1 + x$ for all $x \geq 0$, whence

$$
E\left[\xi(s)^2\right] \leq 3C^2 + 6C^2\, E\left[\int_0^s \xi(t)^2 \mathrm{d}t\right].
$$

Applying the Harnack inequality (Lemma 3.11) with $v : t \mapsto E\left[\xi(t)^2\right]$ and suitable α and γ, we find that $E\left[\xi(s)^2\right]$ is limited (as C is limited). Therefore, $\xi(s)$ is $L^1(P)$ by Remark 2.9, and $E\left[|\xi(s)|\right]$ is limited (by the Cauchy–Schwarz inequality).

Now one can prove that $\xi(t) > 0$ for all $t \in \mathbf{T}$. Indeed, let $\omega \in \Omega$ be such that $\{t \in \mathbf{T} : \xi(t)(\omega) \leq 0\}$ is nonempty, and let $t_\omega + \mathrm{d}t$ be its least element (which must be $\geq \mathrm{d}t$, as $\xi(0) > 0$). Then, $\xi(t_\omega)(\omega) > 0$ while $0 \geq \xi(t_\omega + \mathrm{d}t)(\omega) = \xi(t_\omega)(\omega)\,(1 + \mu(t)(\omega)\mathrm{d}t + \sigma(t_\omega)(\omega)\mathrm{d}W(t_\omega)(\omega))$, so $1 + \mu(t)(\omega)\mathrm{d}t + \sigma(t_\omega)(\omega)\mathrm{d}W(t_\omega)(\omega) \leq 0$, hence either $\sigma(t_\omega)(\omega) \leq -(1 + \mu(t)(\omega)\mathrm{d}t)\,/\sqrt{\mathrm{d}t}$ (if $\mathrm{d}W(t_\omega)(\omega) = \sqrt{\mathrm{d}t}$) or $\sigma(t_\omega)(\omega) \geq (1 + \mu(t)(\omega)\mathrm{d}t)\,/\sqrt{\mathrm{d}t}$ (if $\mathrm{d}W(t_\omega)(\omega) = -\sqrt{\mathrm{d}t}$). In either case, $\sigma(t_\omega)(\omega)$ is unlimited (as μ is limited and thus $1 + \mu(t)(\omega)\mathrm{d}t \simeq 1$). Hence the set of ω such that $\xi(t)(\omega) > 0$ for all $t \in \mathbf{T}$ is for every limited $C' > 0$ a superset of the set of all $\omega \in \Omega$ such that $|\sigma(t)(\omega)| \leq C'$, and for sufficiently large limited C', this set has probability 1, as σ is a limited process.

Therefore, since $\mu(t)$ is either nonpositive for all $t \in \mathbf{T}$ or nonnegative for all $t \in \mathbf{T}$, $\left(\int_0^s \xi(t)\mu(t)\mathrm{d}t\right)_{s\in\mathbf{T}}$ is either a decreasing or an increasing process. On the other hand, $\int \xi\sigma \mathrm{d}W$ is a martingale (by the converse of the martingale representation theorem, Theorem 3.2) as the recursive definition of ξ ensures its adaptedness, so $\xi = \xi(0) + \int \xi(t)\mu(t)\mathrm{d}t + \int \xi(t)\sigma(t)\mathrm{d}W(t)$ is a submartingale or a supermartingale. Therefore, we may apply the corollary to Nelson's super-/submartingale inequality (Corollary 2.13), which, combined with the limitedness of $\xi(0)$ and $E\left[|\xi(s)|\right]$ (see above), yields that ξ is a.s. limited. $\square$

Lemma 3.12. *Let μ, σ be limited $\mathcal{F}$-adapted stochastic processes, and let ξ be the process defined by*

$$
\mathrm{d}\xi(t) = \xi(t)\mu(t)\,\mathrm{d}t + \xi(t)\sigma(t)\,\mathrm{d}W(t)
$$

for all $t \in \mathbf{T} \setminus \{1\}$*, wherein* $\xi(0)$ *is a limited real number* > 0*. Suppose that either* $\mu(t) \geq 0$ *for all* $t \in \mathbf{T}$ *or* $\mu(t) \leq 0$ *for all* $t \in \mathbf{T}$*. Then, a.s. for all* $s \in \mathbf{T}$*,*

$$\xi(s) \simeq \xi(0) \exp\left(\int_0^s \mu(t)\mathrm{d}t + \int_0^s \sigma(t)\, \mathrm{d}W(t) - \frac{1}{2}\int_0^s \sigma(t)^2 \mathrm{d}t\right). \tag{3.6}$$

Hence, if $\xi(0) \gg 0$*, then a.s. for all* $s \in \mathbf{T}$*,* $\xi(s) \gg 0$*.*

Proof. Since $\frac{1}{\xi(t)}\mathrm{d}\xi(t) = \int_0^s \mu(t)\mathrm{d}t + \int_0^s \sigma(t)\, \mathrm{d}W(t)$ (the subtrahend in the argument of the exponential function in Eq. (3.7)) it is enough to prove that

$$\ln \xi(s) - \ln \xi(0) \simeq \int_0^s \frac{1}{\xi(t)}\mathrm{d}\xi(t) - \frac{1}{2}\int_0^s \sigma(t)^2 \mathrm{d}t,$$

and since

$$\frac{1}{\xi(t)^2}(\mathrm{d}\xi(t))^2 = \mu(t)^2(\mathrm{d}t)^2 + 2\mu(t)\sigma(t)\mathrm{d}t\mathrm{d}W(t) + \sigma^2\mathrm{d}t = \sigma(t)^2\mathrm{d}t + \mathcal{O}\left((\mathrm{d}t)^{3/2}\right),$$

it is actually enough to show that

$$\ln \xi(s) - \ln \xi(0) \simeq \int_0^s \frac{1}{\xi(t)}\mathrm{d}\xi(t) - \frac{1}{2}\int_0^s \frac{1}{\xi(t)^2}(\mathrm{d}\xi(t))^2. \tag{3.7}$$

Now, since $\ln' : x \mapsto 1/x$, $\ln'' : x \mapsto -1/x^2$, $\ln''' : x \mapsto 2/x^3$, the third-order Taylor formula yields for every $t \in \mathbf{T}$

$$\mathrm{d}(\ln \xi(t)) = \frac{1}{\xi(t)}\mathrm{d}\xi(t) - \frac{1}{2}\frac{1}{\xi(t)^2}(\mathrm{d}\xi(t))^2 + \frac{1}{3}\frac{1}{\bar{\xi}(t)^3}(\mathrm{d}\xi(t))^3$$

for some $\bar{\xi}(t) \in [\xi(t), \xi(t + \mathrm{d}t)] \cup [\xi(t + \mathrm{d}t), \xi(t)]$, hence

$$\begin{aligned}\ln \xi(s) - \ln \xi(0) = \int_0^s \mathrm{d}(\ln \xi(t)) &= \int_0^s \frac{1}{\xi(t)}\mathrm{d}\xi(t) - \frac{1}{2}\int_0^s \frac{1}{\xi(t)^2}(\mathrm{d}\xi(t))^2 \\ &\quad + \frac{1}{3}\int_0^s \frac{1}{\bar{\xi}(t)^3}(\mathrm{d}\xi(t))^3\end{aligned}$$

for all $s \in \mathbf{T}$. All we need to prove therefore is that a.s. for all $s \in \mathbf{T}$,

$$\int_0^s \frac{1}{\bar{\xi}(t)^3}(\mathrm{d}\xi(t))^3 = \int_0^s \frac{\xi(t)^3}{\bar{\xi}(t)^3}(\mu(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t))^3 \simeq 0.$$

However, combining $\xi(t) > 0$ with the fact that $\bar{\xi}(t) \in [\xi(t), \xi(t + \mathrm{d}t)] \cup [\xi(t + \mathrm{d}t), \xi(t)]$, one gets the following uniform bound:

$$\left|\frac{\xi(t)}{\bar{\xi}(t)}\right| \leq \frac{\xi(t)}{\xi(t) \wedge \underbrace{\xi(t+\mathrm{d}t)}_{=\xi(t)+\mathrm{d}\xi(t)}} \leq 1 \vee \frac{1}{1+\underbrace{\mu(t)\mathrm{d}t+\sigma(t)\mathrm{d}W(t)}_{\simeq 0}} \simeq 1 \ll 2.$$

Moreover, $\mu(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t) = \mathcal{O}\left((\mathrm{d}t)^{1/2}\right)$, therefore we obtain indeed a.s.

$$\int_0^s \frac{\xi(t)^3}{\bar{\xi}(t)^3} \underbrace{(\mu(t)\mathrm{d}t+\sigma(t)\mathrm{d}W(t))^3}_{=\mathcal{O}\left((\mathrm{d}t)^{3/2}\right)} = \mathcal{O}\left((\mathrm{d}t)^{1/2}\right) \simeq 0.$$

□

3.5 The Radically Elementary Version of Lévy's Characterization of Wiener Processes

One of the most remarkable results in Nelson's *Radically elementary probability theory* is a single, unified theorem, called "de Moivre–Laplace–Lindeberg–Feller–Wiener–Lévy–Doob–Erdős–Kac–Donsker–Prokhorov theorem" by Nelson [60, Chap. 18], which entails:

- The necessity and sufficiency of the Lindeberg–Feller condition for the central limit theorem of de Moivre and Laplace.
- Wiener's result about the a.s. continuity of the trajectories of Wiener processes.
- Donsker's invariance principle.
- Lévy's martingale characterization of Wiener processes.

The last item (Lévy's martingale characterization of Wiener processes) is of great importance in stochastic analysis and its applications. It means that whenever a martingale (with respect to the filtration generated by a given Wiener process) has the same quadratic variation as the Wiener process, it already is the Wiener process; a related result is the theorem that the only path-continuous and square-integrable martingale which has stationary and independent increments (i.e. is a *Lévy process*[2]) is a (constant multiple of a) Wiener process.

Keeping in mind that the filtration generated by the Wiener process is a particularly simple and natural one, Lévy's martingale characterization informally asserts that any martingale which has a few desirable properties will already be, up to multiplicative constants, a Wiener process or the exponential of a Wiener process plus a linear drift term (a geometric Wiener process). As a consequence, Lévy's martingale characterization can be fruitfully applied both within pure mathematics

[2]For more on Lévy processes—from the perspective of radically elementary probability theory—see Chap. 9.

(for instance, in the proof of Girsanov's theorem, which establishes a relation between changing the probability measure and adding a linear drift term to the Wiener process) and in mathematical finance (as a mathematical rationale for the adequacy of the Samuelson–Black–Scholes model).

Nelson's unified result, which entails a radically elementary version of Lévy's martingale characterization, can be stated as follows:

Remark 3.13. (Cf. Nelson [60, Theorem 18.1, p. 75].) For a normalized martingale $(\xi(t))_{t\in\mathbf{T}}$ with $\xi(0) = 0$, the following three conditions are equivalent:

- ξ is a Wiener process,
- $\xi(1)$ is $L^2(P)$ and ξ is P-a.s. continuous,
- ξ satisfies the *(near) Lindeberg condition*, i.e.

$$E\left[\sum_{t\in\mathbf{T}\setminus\{1\}} (\mathrm{d}\xi(t))^2\right] \simeq E\left[\sum_{t\in\mathbf{T}\setminus\{1\}} \left(\mathrm{d}\xi(t)\,\chi_{\{|\mathrm{d}\xi(t)|\leq\varepsilon\}}\right)^2\right]$$

for all $\varepsilon \gg 0$.

http://www.springer.com/978-3-642-33148-0

Stochastic Calculus with Infinitesimals
Herzberg, F.S.
2013, XVIII, 112 p., Softcover
ISBN: 978-3-642-33148-0