13. DIFFUSIONS

In this section we introduce the concept of generalized diffusions and then extend the Itô integral to these processes. Then we use it to prove the corresponding Itô formula.

13.1 Generalized diffusions.

Under the sole assumption that $\mathcal{F}_0$ contains all P-null sets, the stochastic integral gives us a way to construct a large class of processes with continuous sample paths; namely,

$$\left\{\int_0^t Y_s \mathrm{d}B_s \colon t \geqslant 0\right\} \quad \text{for each } Y \in \mathcal{V}^{\mathrm{loc}} \tag{13.1}$$

We can of course include also processes that are "shifted" by adding integrals of the form $\int_0^t U_s \mathrm{d}s$ for U satisfying suitable conditions. These give rise to:

Definition 13.1 (Generalized diffusion) *A stochastic process* $\{X_t \colon t \geqslant 0\}$ *is a* generalized diffusion *if the underlying probability space carries a standard Brownian motion* $\{B_t \colon t \geqslant 0\}$ *associated with a Brownian filtration* $\{\mathcal{F}_t\}_{t\geqslant 0}$ *and also two stochastic processes* $\{U_t \colon t \geqslant 0\}$ *and* $\{Y_t \colon t \geqslant 0\}$ *such that the following holds:*

(1) *U, Y and X are jointly measurable and adapted to* $\{\mathcal{F}_t\}_{t\geqslant 0}$,

(2) *for all* $t \geqslant 0$,

$$\int_0^t |U_s| \mathrm{d}s < \infty \;\wedge\; \int_0^t Y_s^2 \mathrm{d}s < \infty \quad \text{a.s.} \tag{13.2}$$

(3) *for all* $t \geqslant 0$,

$$X_t = X_0 + \int_0^t U_s \mathrm{d}s + \int_0^t Y_s \,\mathrm{d}B_s \quad \text{a.s.} \tag{13.3}$$

(4) *X is continuous.*

The conditions (13.2) ensure that the integrals in (13.3) are well defined. The integral identity is often shortened to a differential/infinitesimal form

$$\mathrm{d}X_t = U_t \mathrm{d}t + Y_t \mathrm{d}B_t \tag{13.4}$$

where $U_t \mathrm{d}t$ is called the *drift term* while $Y_t \mathrm{d}B_t$ is referred to as the *diffusive term*. Writing (13.4) tacitly assumes that all of the conditions in Definition 13.1 hold as well.

The reason for adding the prefix "generalized" is that a somewhat more restricted concept of a "proper" diffusion arises in the physics literature:

Definition 13.2 *We say that a process* $\{X_t \colon t \geqslant 0\}$ *is a* diffusion *if there exist functions* $f, g \colon [0,\infty) \times \mathbb{R} \to \mathbb{R}$ *such that Definition 13.1 applies with*

$$U_t := f(t, X_t) \;\wedge\; Y_t := g(t, X_t) \tag{13.5}$$

This assumption ensures that the stochastic dynamics governing the process X has no memory and is local. This has been one of the governing principles of physics since the time of Newton's laws. Note, however, that since X itself enters the condition for its definition, finding the process X for a given f and g requires solving a *stochastic differential equation*, a topic of our future focused interest.

As it turns out, the processes U and Y are determined uniquely from (13.3) modulo changes on a null-set of times, but we will only prove this fact once we have introduced proper tools (see Lemma 14.5). The uniqueness nonetheless allows us to extend the notion of the stochastic integrals to those with respect to diffusions.

Definition 13.3 *Let $\{X_t : t \geqslant 0\}$ take the form $\mathrm{d}X_t = U_t\mathrm{d}t + Y_t\mathrm{d}B_t$. For any $t \geqslant 0$ and any $\{Z_s : s \geqslant 0\}$ which is adapted, jointly measurable and obeys*

$$\int_0^t |Z_s||U_s|\mathrm{d}s < \infty \ \wedge \ \int_0^t |Z_s|^2 Y_s^2 \mathrm{d}s < \infty \quad \textit{a.s.} \tag{13.6}$$

we then define

$$\int_0^t Z_s \mathrm{d}X_s := \int_0^t Z_s U_s \mathrm{d}s + \int_0^t Z_s Y_s \mathrm{d}B_s \tag{13.7}$$

The conditions (13.6) ensure that the integrals in (13.7) are well defined a.s. In particular, if Z has continuous paths, the function $s \mapsto Z_s$ is locally bounded and conditions (13.6) are thus ensured by those in (13.2). Note that, in the infinitesimal form, (13.7) amounts to a mere multiplication of both sides of (13.4) by Z_t.

13.2 Itô formula for diffusions.

With the notion of an integral with respect to a diffusion clarified, we will now generalize the proof of Itô formula to this case. Given a diffusion X with infinitesimal form

$$\mathrm{d}X_t = U_t\mathrm{d}t + Y_t\mathrm{d}B_t \tag{13.8}$$

we denote

$$\langle X \rangle_t := \int_0^t Y_s^2 \mathrm{d}s \tag{13.9}$$

The notation will later be given the meaning of a quadratic variation process associated with local martingales but, at this point, we only think of it as a notation. We now have:

Theorem 13.4 (Itô formula for diffusions) *Let $\{X_t : t \geqslant 0\}$ be a diffusion. Then for all $f \in C^2(\mathbb{R})$ also $\{f(X_t) : t \geqslant 0\}$ is a diffusion and, for all $t \geqslant 0$,*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)\mathrm{d}\langle X\rangle_s \quad \text{a.s.} \tag{13.10}$$

where the first integral is in the sense of Definition 13.3 while the second integral is an ordinary Stieltjes integral. Both integrals exist and are finite a.s.

In order to make the statement completely explicit, assume that X has the differential form (13.8). Then (13.10) means that, for each $t \geqslant 0$, a.s.,

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)U_s\mathrm{d}s + \int_0^t f'(X_s)Y_s\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(X_s)Y_s^2\mathrm{d}s \tag{13.11}$$

We will at times write (13.10) in differential form as

$$\mathrm{d}f(X_t) = f'(X_t)\mathrm{d}X_t + \frac{1}{2}f''(X_t)\mathrm{d}\langle X\rangle_t \tag{13.12}$$

Note that, for $X_t := B_t$, where B is the standard Brownian motion, we have $\langle B\rangle_t = t$ by Proposition 6.2 and so Theorem 13.4 subsumes Theorem 12.11.

Proof of Theorem 13.4 for "simple" processes. We will first prove the theorem for U and Y "simple" in the sense that, for each $t \geqslant 0$,

$$U_t = \sum_{j=1}^m U_{t_{j-1}} 1_{[t'_{j-1}, t'_j)}(t) \tag{13.13}$$

and

$$Y_t = \sum_{j=1}^m Y_{t_{j-1}} 1_{[t'_{j-1}, t'_j)}(t) \tag{13.14}$$

where $0 = t'_0 < t'_1 < \cdots < t'_m$ and $U_{t'_j}$ and $Y_{t'_j}$ are $\mathcal{F}_{t'_j}$-measurable bounded random variables for each $j = 0, \dots, m-1$. (We write "simple" in quotes because the above used a different convention about the endpoints of the intervals $[t'_{j-1}, t'_j)$ than what we postulated for simple processes.) Fix $t \geqslant 0$ and let $\Pi = \{t_i\}_{i=0}^n$ denote a generic partition of $[0, t]$ such that all t'_j's belong to Π.

STEP 1 (Taylor's theorem): Given any $f \in C^2(\mathbb{R})$, Taylor's theorem gives

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_{i=1}^n \big[f(X_{t_i}) - f(X_{t_{i-1}})\big] \\ &= \sum_{i=1}^n f'(X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}) \\ &\quad + \sum_{i=1}^n \int_{t_{i-1}}^{t_i} f''\big(sX_{t_i} + (1-s)X_{t_{i-1}}\big)(1-s)(X_{t_i} - X_{t_{i-1}})^2 \end{aligned} \tag{13.15}$$

The second term on the right can further be separated to the form

$$\begin{aligned} &\frac{1}{2}\sum_{i=1}^n f''(X_{t_{i-1}}) Y_{t_{i-1}}^2 (X_{t_i} - X_{t_{i-1}})^2 \\ &\quad + \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \Big(f''\big(sX_{t_i} + (1-s)X_{t_{i-1}}\big) - f''(X_{t_{i-1}})\Big)(1-s)(X_{t_i} - X_{t_{i-1}})^2 \end{aligned} \tag{13.16}$$

Since Y and U are constant on each interval induced by the partition Π, we have

$$X_{t_i} - X_{t_{i-1}} = U_{t_{i-1}}(t_i - t_{i-1}) + Y_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}) \tag{13.17}$$

Plugging this into the above formula and using the usual expression for the square of the sum yields

$$f(X_t) - f(X_0) = J_t^{(1)}(\Pi) + \cdots + J_t^{(6)}(\Pi) \tag{13.18}$$

where

$$\begin{aligned}J_t^{(1)}(\Pi) &:= \sum_{i=1}^n f'(X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}) \\ J_t^{(2)}(\Pi) &:= \frac{1}{2}\sum_{i=1}^n f''(X_{t_{i-1}})Y_{t_{i-1}}^2(t_i - t_{i-1})\end{aligned} \tag{13.19}$$

and

$$\begin{aligned}J_t^{(3)}(\Pi) &:= \frac{1}{2}\sum_{i=1}^n f''(X_{t_{i-1}})Y_{t_{i-1}}^2\Big[(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\Big] \\ J_t^{(4)}(\Pi) &:= \sum_{i=1}^n f''(X_{t_{i-1}})U_{t_{i-1}}Y_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})(t_i - t_{i-1}) \\ J_t^{(5)}(\Pi) &:= \sum_{i=1}^n f''(X_{t_{i-1}})U_{t_{i-1}}(t_i - t_{i-1})^2 \\ J_t^{(6)}(\Pi) &:= \sum_{i=1}^n \int_{t_{i-1}}^{t_i}\Big(f''\big(sX_{t_i} + (1-s)X_{t_{i-1}}\big) - f''(X_{t_{i-1}})\Big)(1-s)(X_{t_i} - X_{t_{i-1}})^2\end{aligned} \tag{13.20}$$

We will now control each of these five terms in the limit as $\|\Pi\| \to 0$ and show that $J_t^{(1)}(\Pi)$ and $J_t^{(2)}(\Pi)$ give rise to the terms on the right of (13.10) while the remaining terms vanish in the limit as $\|\Pi\| \to 0$.

STEP 2 (Limit of $J_t^{(1)}(\Pi)$): first term will give rise to the first term on the right of (13.10). To show this we first use additivity of the integral to get

$$\begin{aligned}J_t^{(1)}(\Pi) - \int_0^t f'(X_s)\mathrm{d}X_s &= \sum_{i=1}^n \int_{t_{i-1}}^{t_i}\big(f'(X_{t_{i-1}}) - f'(X_s)\big)\mathrm{d}X_s \\ &= \sum_{i=1}^n \int_{t_{i-1}}^{t_i}\big(f'(X_{t_{i-1}}) - f'(X_s)\big)U_s\mathrm{d}s + \sum_{i=1}^n \int_{t_{i-1}}^{t_i}\big(f'(X_{t_{i-1}}) - f'(X_s)\big)Y_s\mathrm{d}B_s\end{aligned} \tag{13.21}$$

The first term on the right is bounded by

$$\sum_{i=1}^n \int_{t_{i-1}}^{t_i}\big|f'(X_{t_{i-1}}) - f'(X_s)\big||U_s|\mathrm{d}s \leqslant \mathrm{osc}_{f'\circ X}\big([0,t],\|\Pi\|\big)\int_0^t |U_s|\mathrm{d}s \tag{13.22}$$

The oscillation tends to zero as $\|\Pi\| \to 0$ by continuity of $f' \circ X$. For the second term on the right we note

$$\sum_{i=1}^n \int_{t_{i-1}}^{t_i}\big|f'(X_{t_{i-1}}) - f'(X_s)\big|^2 Y_s^2\mathrm{d}s \leqslant \mathrm{osc}_{f'\circ X}\big([0,t],\|\Pi\|\big)^2\int_0^t Y_s^2\mathrm{d}s \tag{13.23}$$

Again, this tends to zero as $\|\Pi\| \to 0$ which by Lemma 12.9 implies that the second term on the right of (13.21) tends in probability to zero as $\|\Pi\| \to 0$. Hereby we conclude

$$J_t^{(1)}(\Pi) \underset{\|\Pi\|\to 0}{\overset{P}{\longrightarrow}} \int_0^t f'(X_s)\mathrm{d}X_s \tag{13.24}$$

as desired.

STEP 3 (Limit of $J_t^{(2)}(\Pi)$): Since $s \mapsto f''(X_s)Y_s^2$ is piecewise continuous, the standard facts about Riemann integral give

$$J_t^{(2)}(\Pi) \underset{\|\Pi\|\to 0}{\longrightarrow} \frac{1}{2}\int_0^t f''(X_s)Y_s^2 \mathrm{d}s \tag{13.25}$$

Noting that $\mathrm{d}\langle X\rangle_t = Y_t^2\mathrm{d}t$, this reproduces the second term on the right of (13.10).

STEP 4 (Limit of $J_t^{(3)}(\Pi)$): Abbreviate $Z_i := Y_{t_{i-1}}^2[(f''(X_{t_{i-1}}) \wedge M) \vee (-M)]$. Then

$$\begin{aligned} P\big(|J_t^{(3)}(\Pi)| > \epsilon\big) \leqslant P\Big(\sup_{s\leqslant t}\big|f''(X_s)\big| > M\Big) \\ + P\left(\left|\sum_{i=1}^n Z_{i-1}\Big[(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\Big]\right| > \epsilon\right) \end{aligned} \tag{13.26}$$

The key point is that the quantity inside the absolute value in the second probability on the right is a martingale. Using Doob's L^2-inequality (10.3) and the fact that Z_{i-1} is, conditional on $\mathcal{F}_{t_{i-1}}$, independent of $B_{t_i} - B_{i-1}$, the probability is thus at most

$$\begin{aligned} &\frac{1}{\epsilon^2}\sum_{i=1}^n E\left(Z_{i-1}^2\Big[(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\Big]^2\right) \\ &\leqslant \frac{1}{\epsilon^2}\max_{i=1,\dots,n}\|Z_i\|_\infty^2 \sum_{i=1}^n \mathrm{Var}\big((B_{t_i} - B_{t_{i-1}})^2\big) = \frac{3}{\epsilon^2}\max_{i=1,\dots,n}\|Z_i\|_\infty^2 \sum_{i=1}^n |t_i - t_{i-1}|^2 \end{aligned} \tag{13.27}$$

the sum on the right is bounded by $t\|\Pi\|$ and so this expression vanishes as $\|\Pi\| \to 0$. Hence so does the second probability on the right of (13.26), regardless of M. Taking $M \to \infty$ then gives

$$J_t^{(3)}(\Pi) \underset{\|\Pi\|\to 0}{\overset{P}{\longrightarrow}} 0 \tag{13.28}$$

as desired.

STEP 5 (Limit of $J_t^{(4)}(\Pi)$): Denoting $C := \max_{i=1,\dots,n}\max\{\|U_{t_{i-1}}\|_\infty, \|Y_{t_{i-1}}\|_\infty\}$, we have

$$\big|J_t^{(4)}(\Pi)\big| \leqslant C\Big(\sup_{s\leqslant t}\big|f''(X_s)\big|\Big)\mathrm{osc}_B\big([0,t], \|\Pi\|\big)t \tag{13.29}$$

Thanks to continuity of B, the oscillation tends to zero as $\|\Pi\| \to 0$ showing that

$$J_t^{(4)}(\Pi) \underset{\|\Pi\|\to 0}{\overset{P}{\longrightarrow}} 0 \tag{13.30}$$

as desired.

STEP 6 (Limit of $J_t^{(5)}(\Pi)$): Continuing to use the above notation, here we get

$$\big|J_t^{(5)}(\Pi)\big| \leqslant C^2\Big(\sup_{s\leqslant t}\big|f''(X_s)\big|\Big)t\|\Pi\| \tag{13.31}$$

and so

$$J_t^{(5)}(\Pi) \underset{\|\Pi\|\to 0}{\overset{P}{\longrightarrow}} 0 \tag{13.32}$$

as desired.

STEP 7 (Limit of $J_t^{(6)}(\Pi)$): Here we invoke the bound

$$\big|J_t^{(6)}(\Pi)\big| \leqslant \frac{1}{2}\mathrm{osc}_{f''}\big(X([0,t]),\mathrm{osc}_X([0,t],\|\Pi\|)\big)V_t^{(2)}(X,\Pi) \tag{13.33}$$

The continuity of f'' and X ensures that the inner as well as outer oscillations tend to zero as $\|\Pi\| \to 0$. For the second variation we invoke the inequality $(a+b)^2 \leqslant 2a^2+2b^2$ along with (13.17) to get

$$V_t^{(2)}(X,\Pi) \leqslant 2C^2 V_t^{(2)}(B,\Pi) + 2C^2 t\|\Pi\| \tag{13.34}$$

This stays bounded uniformly (in probability) as $\|\Pi\| \to 0$ thus showing that

$$J_t^{(6)}(\Pi) \underset{\|\Pi\|\to 0}{\overset{P}{\longrightarrow}} 0 \tag{13.35}$$

as desired. This completes the proof of (13.10) for U and V of the form (13.13–13.14). $\square$

13.3 Proof of Itô formula for diffusion, general case.

In order to prove the general case of (13.10), suppose U and Y are such that, for some processes $\{U_t^{(n)}\colon t \geqslant 0\}$ and $\{Y_t^{(n)}\colon t \geqslant 0\}$ of the form (13.13–13.14) we have

$$\int_0^t |U_s^{(n)} - U_s|\,\mathrm{d}s \underset{n\to\infty}{\overset{P}{\longrightarrow}} 0 \quad \wedge \quad \int_0^t |Y_s^{(n)} - Y_s|^2\,\mathrm{d}s \underset{n\to\infty}{\overset{P}{\longrightarrow}} 0 \tag{13.36}$$

Use these to define

$$X_u^{(n)} := X_0 + \int_0^u U_s^{(n)}\mathrm{d}s + \int_0^u Y_s^{(n)}\mathrm{d}B_s \tag{13.37}$$

for all $u \in [0,t]$, where we assume continuous versions of the integrals. We now observe:

Lemma 13.5 *Under* (13.36)*,*

$$\sup_{0\leqslant u\leqslant t}\big|X_u^{(n)} - X_u\big| \underset{n\to\infty}{\overset{P}{\longrightarrow}} 0 \tag{13.38}$$

Proof. We have

$$\sup_{0\leqslant u\leqslant t}\big|X_u^{(n)} - X_u\big| \leqslant \int_0^t |U_s^{(n)} - U_s|\,\mathrm{d}s + \sup_{0\leqslant u\leqslant t}\Big|\int_0^t (Y_s^{(n)} - Y_s)\mathrm{d}B_s\Big| \tag{13.39}$$

As $n \to \infty$, the first term tends to zero in probability by assumption (13.36). For the second term, we introduce

$$T_n := \inf\Big\{u \geqslant 0\colon \int_0^u |Y_s^{(n)} - Y_s|^2\,\mathrm{d}s \geqslant 1\Big\} \tag{13.40}$$

and note that

$$\int_0^t (Y_s^{(n)} - Y_s)\mathrm{d}B_s = \int_0^t (Y_s^{(n)} - Y_s)1_{\{T_n>s\}}\mathrm{d}B_s \quad \text{a.s. on } \{T_n > t\} \tag{13.41}$$

Doob's L^2-inequality (10.3) along with Itô isometry then give

$$P\biggl(\sup_{0\leqslant u\leqslant t}\Bigl|\int_0^t (Y_s^{(n)}-Y_s)\,\mathrm{d}B_s\Bigr|>\epsilon\biggr) \leqslant P(T_n\leqslant t)+\frac{1}{\epsilon^2}E\biggl(\int_0^t (Y_s^{(n)}-Y_s)^2 1_{\{T_n>s\}}\mathrm{d}s\biggr) \tag{13.42}$$

As $n\to\infty$, the first probability tends to zero in light of $T_n\to\infty$ a.s. The expectation tends to zero by the Bounded Convergence Theorem and the second part of (13.36). □

We now use these to prove:

Lemma 13.6 *Assuming* (13.36)*, for all* $h\in C(\mathbb{R})$*,*

$$\int_0^t h(X_s^{(n)})\mathrm{d}X_s^{(n)} \ \underset{n\to\infty}{\overset{P}{\longrightarrow}}\ \int_0^t h(X_s)\mathrm{d}X_s \tag{13.43}$$

Proof. By (13.7), the integrals are sums of two integrals and so it suffices to prove the limit for each of the integrals separately. Here we observe

$$\begin{aligned}\Bigl|\int_0^t h(X_s^{(n)})U_s^{(n)}\mathrm{d}s-\int_0^t h(X_s)U_s\mathrm{d}s\Bigr| &\leqslant \int_0^t \bigl|h(X_s^{(n)})U_s^{(n)}-h(X_s)U_s\bigr|\mathrm{d}s\\ &\leqslant \Bigl(\sup_{s\leqslant t}\bigl|h(X_s^{(n)})-h(X_s)\bigr|\Bigr)\int_0^t |U_s^{(n)}|\mathrm{d}s+\Bigl(\sup_{s\leqslant t}\bigl|h(X_s)\bigr|\Bigr)\int_0^t |U_s^{(n)}-U_s|\,\mathrm{d}s\end{aligned} \tag{13.44}$$

Here the first supremum tends to zero by Lemma 13.5 and the fact that $\int_0^t|U_s^{(n)}|\mathrm{d}s$ is bounded uniformly in probability. Hence this term tends to zero by (13.36).

For the second integral contributing to those in the claim we invoke the inequality $(a+b)^2\leqslant 2a^2+2b^2$ to get

$$\begin{aligned}&\int_0^t \bigl|h(X_s^{(n)})Y_s^{(n)}-h(X_s)Y_s\bigr|^2\mathrm{d}s\\ &\quad\leqslant 2\Bigl(\sup_{s\leqslant t}\bigl|h(X_s^{(n)})-h(X_s)\bigr|\Bigr)^2\int_0^t (Y_s^{(n)})^2\mathrm{d}s+2\Bigl(\sup_{s\leqslant t}\bigl|h(X_s)\bigr|\Bigr)^2\int_0^t |Y_s^{(n)}-Y_s|^2\,\mathrm{d}s\end{aligned} \tag{13.45}$$

Using the same reasoning as above, this again tends to zero as $n\to\infty$ in probability. Lemma 12.9 then shows

$$\int_0^t h(X_s^{(n)})Y_s^{(n)}\mathrm{d}B_s \ \underset{n\to\infty}{\overset{P}{\longrightarrow}}\ \int_0^t h(X_s)Y_s\mathrm{d}B_s \tag{13.46}$$

thus proving the claim. □

Lemma 13.7 *Assuming* (13.36)*, for all* $h\in C(\mathbb{R})$*,*

$$\int_0^t h(X_s^{(n)})\mathrm{d}\langle X^{(n)}\rangle_s \ \underset{n\to\infty}{\overset{P}{\longrightarrow}}\ \int_0^t h(X_s)\mathrm{d}\langle X\rangle_s \tag{13.47}$$

Proof. Using the explicit form of the bracket processes,

$$\begin{aligned}\left|\int_0^t h(X_s^{(n)})\mathrm{d}\langle X^{(n)}\rangle_s - \int_0^t h(X_s)\mathrm{d}\langle X\rangle_s\right| &\leqslant \int_0^t \left|h(X_s^{(n)})(Y_s^{(n)})^2 - h(X_s)Y_s^s\right|\mathrm{d}s \\ &\leqslant \Big(\sup_{s\leqslant t}\left|h(X_s^{(n)}) - h(X_s)\right|\Big)\int_0^t (Y_s^{(n)})^2\mathrm{d}s + \Big(\sup_{s\leqslant t}\left|h(X_s)\right|\Big)\int_0^t |(Y_s^{(n)})^2 - Y_s^2|\,\mathrm{d}s\end{aligned} \tag{13.48}$$

In light of

$$\int_0^t |(Y_s^{(n)})^2 - Y_s^2|\,\mathrm{d}s \leqslant \left(\int_0^t (Y_s^{(n)})^2 + Y_s^2)\,\mathrm{d}s\right)^{1/2}\left(\int_0^t |Y_s^{(n)} - Y_s|^2\,\mathrm{d}s\right)^{1/2} \tag{13.49}$$

the assumption (13.36) again shows that the quantity on the right of (13.48) tends to zero in probability as $n \to \infty$. □

We are now finally ready to give:

Proof of Theorem 13.4. Let U and Y be processes satisfying the conditions in Definition 13.1. Using the reasoning underlying the proof of Theorem 10.5 albeit without expectation, we can find simple processes $\{U_t^{(n)}\colon t \geqslant 0\}$ and $\{Y_t^{(n)}\colon t \geqslant 0\}$ such that (13.36) hold. Modifying the processes at the partition points brings them to the form assumed in (13.13–13.14). By the previous part of the proof we then have

$$f(X_t^{(n)}) = f(X_0) + \int_0^t f'(X_s^{(n)})\mathrm{d}X_s^{(n)} + \frac{1}{2}\int_0^t f''(X_s^{(n)})\mathrm{d}\langle X^{(n)}\rangle_s \quad \text{a.s.} \tag{13.50}$$

whenever $f \in C^2(\mathbb{R})$. Using Lemmas 13.5–13.7, both sides converge in probability to those of the desired expression (13.10). □

13.4 Further extensions.

The result of Itô formula can be encoded by the following mnemonic rules for operations with differentials:

$$(\mathrm{d}t)^2 = 0, \quad (\mathrm{d}t)(\mathrm{d}B_t) = 0 \quad \text{but} \quad (\mathrm{d}B_t)^2 = \mathrm{d}t \tag{13.51}$$

Similarly as differentiation in one variable extends to partial derivatives with respect to multiple variables, the Itô formula extends even to functions of multiple diffusions that arise from a multiple of Brownian motions. The statement, which we leave without proof, is as follows:

Theorem 13.8 *Let $B^{(1)},\dots,B^{(d)}$ be independent standard Brownian motions with a common Brownian filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$. Let $X = (X^{(1)},\dots,X^{(m)})$ be and $\mathbb{R}^m$-valued stochastic process whose coordinates are given in differential form by*

$$\mathrm{d}X_t^{(i)} = U_t^{(i)}\mathrm{d}t + \sum_{k=1}^d Y_t^{(i,k)}\mathrm{d}B_t^{(k)} \tag{13.52}$$

where $\{U^{(i)}\colon i = 1,\dots,m\}$ and $\{Y^{(i,k)}\colon i = 1,\dots,m \wedge k = 1,\dots,d\}$ are as specified in Definition 13.1 for the above to make sense. Then for each $f\colon [0,\infty)\times\mathbb{R}^m \to \mathbb{R}$ that is C^1 in the *time variable and C^2 in the spatial variables and each $t \geqslant 0$,*

$$f(t,X_t) = f(0,X_0) + \int_0^t \frac{\partial f}{\partial s}(s,X_s)\mathrm{d}s + \sum_{i=1}^m \int_0^t \frac{\partial f}{\partial x_i}(s,X_s)U_s^{(i)}\mathrm{d}s + \sum_{i=1}^k \sum_{k=1}^d \int_0^t \frac{\partial f}{\partial x_i}(s,X_s)Y_s^{(i,k)}\mathrm{d}B_s^{(k)} + \frac{1}{2}\sum_{i,j=1}^m \sum_{k=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(s,X_s)Y_s^{(i,k)}Y_s^{(j,k)}\mathrm{d}s \tag{13.53}$$

As a further generalization, one can even consider the case when the Brownian motions $B^{(1)},\dots,B^{(d)}$ are not independent meaning that the covariance of this vector is not a multiple of the identity matrix. This changes (13.53) only in the last term, which in turn becomes

$$\frac{1}{2}\sum_{i,j=1}^m \sum_{k,k'=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(s,X_s)Y_s^{(i,k)}Y_s^{(j,k)}C_{k,k'}\mathrm{d}s \tag{13.54}$$

where $C_{k,k'} = \mathrm{Cov}(B_1^{(k)},B_1^{(k')})$. Here the mnemonic rules become

$$(\mathrm{d}t)^2 = 0, \quad (\mathrm{d}t)(\mathrm{d}B_t^{(k)}) = 0 \quad \text{but} \quad (\mathrm{d}B_t^{(k)})(\mathrm{d}B_t^{(k')}) = C_{k,k'}\mathrm{d}t \tag{13.55}$$

We leave the details to the reader.

To demonstrate an example of where the multidimensional setting is particularly useful, recall that the d-dimensional standard Brownian motion is the $\mathbb{R}^d$-valued process $B = (B^{(1)},\dots,B^{(d)})$ where $B^{(1)},\dots,B^{(d)}$ are independent (1-dimensional) standard Brownian motions. The radial variable is then denoted as

$$R_t := \Bigl(\sum_{i=1}^d [B_t^{(i)}]^2\Bigr)^{1/2} \tag{13.56}$$

We can interpret this as $R_t = f\circ B$ for $f(x_1,\dots,x_d) := (x_1^2+\dots+x_d^2)^{1/2}$. This function is not C^2 at the origin but if we stop the process before hitting the origin, Theorem 13.8 can be applied. Since, for each $k=1,\dots,d$,

$$\frac{\partial f}{\partial x_k} = \frac{x_k}{f(x)} \quad \wedge \quad \frac{\partial^2 f}{\partial x_k^2} = \frac{1}{f(x)} - \frac{x_k^2}{f(x)^3} \tag{13.57}$$

a calculation shows that R admits the differential form

$$\mathrm{d}R_t = \frac{d-1}{2R_t}\mathrm{d}t + \frac{1}{R_t}\sum_{i=1}^d B_t^{(i)}\mathrm{d}B_t^{(i)} \tag{13.58}$$

We will see that, with proper rewrite of the second term on the right hand side, this becomes stochastic differential equation for the so called *d-dimensional Bessel process*.

Further reading: Section 3.3A of Karatzas-Shreve