Chapter 5
Wong-Zakai

5.1. Itô versus Stratonovich

We recall that $B=(B_t)_{t\in[0,T]}$ is a Brownian motion in $\mathbb{R}^d$. Given the Itô rough path $\mathbb{B}=(\mathbb{B}^1,\mathbb{B}^2)$ over B constructed in Theorem 4.2, see (4.2), we can define a new rough path $\bar{\mathbb{B}}=(\bar{\mathbb{B}}^1,\bar{\mathbb{B}}^2)$ over B, called the *Stratonovich rough path*, given by

$$\bar{\mathbb{B}}^1_{st}:=\mathbb{B}^1_{st}\,,\qquad \bar{\mathbb{B}}^2_{st}:=\mathbb{B}^2_{st}+\frac{t-s}{2}\,\mathrm{Id}_{\mathbb{R}^d}\,,\qquad \forall 0\leqslant s\leqslant t\leqslant T, \tag{5.1}$$

that is

$$(\bar{\mathbb{B}}^2_{st})^{ij}:=(\mathbb{B}^2_{st})^{ij}+\frac{t-s}{2}\,\mathbb{1}_{\{i=j\}}=\begin{cases}\frac{B^i_t-B^i_s}{2} & \text{if } i=j,\\ \int_s^t (B^i_r-B^i_s)\,\mathrm{d}B^j_r & \text{if } i\neq j.\end{cases} \tag{5.2}$$

The fact that $\bar{\mathbb{B}}$ is an α-rough path over B, for any $\alpha\in\left]\frac{1}{3},\frac{1}{2}\right[$, is a consequence of Theorem 4.1 (note that $\bar{\mathbb{B}}^2_{st}=\mathbb{B}^2_{st}+\delta f_{st}$ with $f_t=\frac{t}{2}\mathrm{Id}_{\mathbb{R}^d}$, hence $\delta\bar{\mathbb{B}}^2=\delta\mathbb{B}^2$).

Remark 5.1. (Stratonovich integral) If $X,Y\colon[0,T]\to\mathbb{R}$ are continuous semimartingales, the Stratonovich integral of X with respecto to Y is defined by

$$\int_0^t X_s\circ\mathrm{d}Y_s:=\int_0^t X_s\,\mathrm{d}Y_s+\frac{1}{2}\langle X,Y\rangle_t,\qquad t\in[0,T], \tag{5.3}$$

where $\int_0^t X_s\,\mathrm{d}Y_s$ is the Itô integral and $\langle\cdot,\cdot\rangle$ is the quadratic covariation. For Brownian motion B on $\mathbb{R}^d$ we have $\langle B^i,B^j\rangle_t=t\,\mathbb{1}_{\{i=j\}}$, hence recalling (4.2) we see that

$$\bar{\mathbb{B}}^2_{st}:=\int_s^t \bar{\mathbb{B}}^1_{sr}\otimes\circ\mathrm{d}B_r,\qquad 0\leqslant s\leqslant t\leqslant T. \tag{5.4}$$

This explains why we call $\bar{\mathbb{B}}=(\bar{\mathbb{B}}^1,\bar{\mathbb{B}}^2)$ the Stratonovich rough path.

Let us consider now the Stratonovich version of the SDE (4.12):

$$\mathrm{d}Y_t=b(Y_t)\,\mathrm{d}t+\sigma(Y_t)\circ\mathrm{d}B_t,\qquad \text{i.e.}$$

$$Y_t=Y_0+\int_0^t b(Y_s)\,\mathrm{d}s+\int_0^t\sigma(Y_s)\circ\mathrm{d}B_s,\qquad t\geqslant 0, \tag{5.5}$$

where $b\colon\mathbb{R}^k\to\mathbb{R}^k$ and $\sigma\colon\mathbb{R}^k\to\mathbb{R}^k\otimes(\mathbb{R}^d)^*$ are given. This equation can be recast in the Itô form by the conversion rule (5.3): since the martingale part of $(\sigma(Y_t))_{t\geqslant 0}$ is $(\int_0^t\sigma_2(Y_s)\mathrm{d}B_s)_{t\geqslant 0}$ by the Itô formula, see (4.11), we obtain

$$Y_t=Y_0+\int_0^t\left(b(Y_s)+\frac{1}{2}\,\mathrm{Tr}_{\mathbb{R}^d}[\sigma_2(Y_s)]\right)\mathrm{d}s+\int_0^t\sigma(Y_s)\,\mathrm{d}B_s,\qquad t\geqslant 0.$$

This is precisely the SDE (4.12) with a different drift $\hat{b}(\cdot) := b(\cdot) + \frac{1}{2}\operatorname{Tr}_{\mathbb{R}^d}[\sigma_2(\cdot)]$.

As an immediate corollary of Theorem 4.4, we obtain the following result.

THEOREM 5.2. (STRATONOVICH SDE & RDE) *f* $\sigma(\cdot)$ *is of class* C^2 *and* $b(\cdot)$ *is continuous, then almost surely any solution* $Y = (Y_t)_{t\in[0,T]}$ *of the Stratonovich SDE (5.5) is also a solution of the following RDE, for* $0 \leqslant s \leqslant t \leqslant T$:

$$\begin{aligned} \delta Y_{st} &= b(Y_s)\,(t-s) + \sigma(Y_s)\,\bar{\mathbb{B}}^1_{st} + \sigma_2(Y_s)\,\bar{\mathbb{B}}^2_{st} + o(t-s) \\ &= \left(b(Y_s) + \frac{1}{2}\operatorname{Tr}_{\mathbb{R}^d}[\sigma_2(Y_s)]\right)(t-s) + \sigma(Y_s)\,\mathbb{B}^1_{st} + \sigma_2(Y_s)\,\mathbb{B}^2_{st} + o(t-s). \end{aligned} \tag{5.6}$$

If $\sigma(\cdot)$, $\sigma_2(\cdot)$, $b(\cdot)$ *are of class* C^3 *and, furthermore,* $\sigma(\cdot)$, $\sigma_2(\cdot)$, $b(\cdot)$ *are globally Lipschitz, i.e.* $\|\nabla\sigma\|_\infty + \|\nabla\sigma_2\|_\infty + \|\nabla b\|_\infty < \infty$, *then almost surely the SDE (5.5) and the RDE (5.6) have the same unique solution* $Y = (Y_t)_{t\in[0,T]}$.

In conclusion, if the coefficients $b(\cdot)$ and $\sigma(\cdot)$ are sufficiently regular, the Itô equation (4.12) can be reintepreted as the RDE

$$\delta Y_{st} = b(Y_s)\,(t-s) + \sigma(Y_s)\,\mathbb{B}^1_{st} + \sigma_2(Y_s)\,\mathbb{B}^2_{st} + o(t-s), \qquad 0 \leqslant s \leqslant t \leqslant T,$$

while the Stratonovich equation (5.5) can be reintepreted as the RDE

$$\delta Y_{st} = b(Y_s)\,(t-s) + \sigma(Y_s)\,\bar{\mathbb{B}}^1_{st} + \sigma_2(Y_s)\,\bar{\mathbb{B}}^2_{st} + o(t-s), \qquad 0 \leqslant s \leqslant t \leqslant T.$$

In other words, rough paths allow to describe the Itô and the Stratonovich SDEs as *the same equation* where only the second level of the rough path has been changed. This shows that, in a sense, the *relevant noise* for a SDE is not only the Brownian path $(B_t)_{t\geqslant 0}$, but rather the rough path $\mathbb{B}$ or $\bar{\mathbb{B}}$.

5.2. WONG-ZAKAI

In this section we show the following application of the previous results. We consider a family $(\rho_\varepsilon)_{\varepsilon>0}$ of compactly supported mollifiers on $\mathbb{R}$, namely $\rho\colon \mathbb{R} \to [0,\infty)$ is smooth, compactly supported in $[-1,1]$, satisfies $\int_{\mathbb{R}} \rho(x)\,\mathrm{d}x = 1$ and we set

$$\rho_\varepsilon(x) := \frac{1}{\varepsilon}\,\rho\!\left(\frac{x}{\varepsilon}\right), \qquad \varepsilon > 0, x \in \mathbb{R}. \tag{5.7}$$

(We do not assume that ρ is even.) We consider a d-dimensional two-sided Brownian motion $(B_t)_{t\in\mathbb{R}}$, namely a Gaussian centered process with values in $\mathbb{R}^d$ such that

$$B_0 = 0, \qquad \mathbb{E}[B^i_s\,B^j_t] = \mathbb{1}_{(i=j)}\,\mathbb{1}_{(st\geqslant 0)}\,(|s| \wedge |t|),$$

which is equivalent to say that $(B_t)_{t\geqslant 0}$ and $(B_{-t})_{t\geqslant 0}$ are two independent d-dimensional Brownian motions.

We consider the following problem: we define a regularization of $(B_t)_{t\geqslant 0}$ by

$$B^\varepsilon_t := (\rho_\varepsilon * B)_t = \int_{\mathbb{R}} \rho_\varepsilon(u)\,B_{t-u}\,\mathrm{d}u, \qquad t \geqslant 0, \tag{5.8}$$

and we consider the integral equation (3.3) controlled by B^ε, namely

$$Z_t^\varepsilon = Z_0 + \int_0^t \sigma(Z_s^\varepsilon)\, \dot{B}_s^\varepsilon \, \mathrm{d}s, \qquad 0 \leqslant t \leqslant T. \tag{5.9}$$

It is easy to check that $(B_t^\varepsilon)_{t\geqslant 0}$ converges to $(B_t)_{t\geqslant 0}$ as $\varepsilon \downarrow 0$ uniformly for $t \in [0,T]$ (and even in C^α for any $\alpha < \frac{1}{2}$; see below). Then we want to understand whether $(Z_t^\varepsilon)_{t\geqslant 0}$ also converges, and especially to which limit.

This question has a very natural answer in the context of rough paths. We define the *canonical rough path* over B^ε (see section 8.7 below for more on this notion):

$$\mathbb{B}_{st}^{\varepsilon,1} := B_t^\varepsilon - B_s^\varepsilon, \qquad \mathbb{B}_{st}^{\varepsilon,2} := \int_s^t \mathbb{B}_{su}^{\varepsilon,1} \otimes \dot{B}_u^\varepsilon \, \mathrm{d}u, \qquad 0 \leqslant s \leqslant t. \tag{5.10}$$

Then we can prove the following result.

Theorem 5.3. (Wong-Zakai) *As $\varepsilon \downarrow 0$, $\mathbb{B}^\varepsilon$ converges in probability to the Stratonovich rough path $\bar{\mathbb{B}}$, see (5.1), namely for any $\alpha < \frac{1}{2}$*

$$\|\mathbb{B}^{\varepsilon,1} - \bar{\mathbb{B}}^1\|_\alpha + \|\mathbb{B}^{\varepsilon,2} - \bar{\mathbb{B}}^2\|_{2\alpha} \xrightarrow[\varepsilon \downarrow 0]{} 0 \qquad \textit{in probability}. \tag{5.11}$$

The convergence holds almost surely along sequences $\varepsilon = \varepsilon_n \downarrow 0$ exponentially fast.

Moreover let $(Z_t^\varepsilon)_{t\in[0,T]}$ be the solution to the controlled equation

$$Z_t^\varepsilon = Z_0 + \int_0^t \sigma(Z_s^\varepsilon)\, \dot{B}_s^\varepsilon \, \mathrm{d}s, \qquad t \geqslant 0.$$

Assume that $\sigma: \mathbb{R}^k \to \mathbb{R}^k \otimes (\mathbb{R}^d)^*$ is of class C^3, with $\|\nabla\sigma\|_\infty + \|\nabla^2\sigma\|_\infty + \|\nabla^3\sigma\|_\infty + \|\nabla\sigma_2\|_\infty + \|\nabla^2\sigma_2\|_\infty < +\infty$. Then, for any $\alpha \in \left]0, \frac{1}{2}\right[$, we have $Z^\varepsilon \to Z$ in probability in $C^\alpha([0,T];\mathbb{R}^k)$ as $\varepsilon \downarrow 0$, where Z is the unique solution to the Stratonovich SDE

$$Z_t = Z_0 + \int_0^t \sigma(Z_s) \circ \mathrm{d}B_s = Z_0 + \int_0^t \sigma(Z_s)\, \mathrm{d}B_s + \frac{1}{2}\int_0^t \mathrm{Tr}_{\mathbb{R}^d}[\sigma_2(Z_s)]\, \mathrm{d}s.$$

Proof. Fix $\alpha \in \left(\frac{1}{3}, \frac{1}{2}\right)$. Let $\mathbb{B}^\varepsilon$ be the canonical smooth rough path associated with B^ε as in (3.9). Suppose we have proved that $\mathbb{B}^\varepsilon$ converges to $\bar{\mathbb{B}}$ as in (5.11). By Proposition 3.5, the solution Z^ε to the controlled equation (5.9) is equal to the (unique by Theorem 3.10) solution to the rough finite difference equation (3.19) associated with the α-rough path $\mathbb{B}^\varepsilon$. In the notation (3.51), we have $Z^\varepsilon = \Phi(Z_0, \mathbb{B}^\varepsilon)$, and by Theorem 5.2 we have $Z = \Phi(Z_0, \bar{\mathbb{B}})$. By the continuity result Theorem 3.11 we obtain that $Z^\varepsilon = \Phi(Z_0, \mathbb{B}^\varepsilon) \to \Phi(Z_0, \bar{\mathbb{B}}) = Z$ a.s. as $\varepsilon \downarrow 0$.

It remains now to prove (5.11). We first observe that by (5.8)

$$\mathbb{B}_{st}^{\varepsilon,1} = \int_{\mathbb{R}} \rho_\varepsilon(u)\, \delta B_{s-u,t-u}\, \mathrm{d}u. \tag{5.12}$$

Let us fix $\alpha < \frac{1}{2}$ and set $\|\delta B\|_\alpha := \|\delta B\|_{\alpha,[-1,T+1]}$, so that $|\delta B_{ab}| \leqslant \|\delta B\|_\alpha (b-a)^\alpha$ for all $-1 \leqslant a < b \leqslant T+1$. Then, uniformly for $\varepsilon \in (0,1)$ and $0 \leqslant s < t \leqslant T$, we can bound

$$|\mathbb{B}_{st}^{\varepsilon,1}| \leqslant \|\delta B\|_\alpha (t-s)^\alpha. \tag{5.13}$$

We can write similarly

$$\mathbb{B}^{\varepsilon,1}_{st} - \bar{\mathbb{B}}^1_{st} = \int_{\mathbb{R}} \rho_\varepsilon(u)\,(\delta B_{s-u,t-u} - \delta B_{st})\,\mathrm{d}u\,, \tag{5.14}$$

hence for any $\alpha' \in \left]\alpha, \frac{1}{2}\right[$ we can estimate, by the triangle inequality,

$$|\mathbb{B}^{\varepsilon,1}_{st} - \bar{\mathbb{B}}^1_{st}| \leqslant 2\,\|\delta B\|_{\alpha'}\,(t-s)^{\alpha'}.$$

At the same time, since $\delta B_{st} - \delta B_{s-u,t-u} = \delta B_{t-u,t} - \delta B_{s-u,s}$, we can also bound

$$|\mathbb{B}^{\varepsilon,1}_{st} - \bar{\mathbb{B}}^1_{st}| \leqslant 2\,\|\delta B\|_{\alpha'} \int_{\mathbb{R}} \rho_\varepsilon(u)\,u^{\alpha'}\,\mathrm{d}u \leqslant 2\,\|\delta B\|_{\alpha'}\,\varepsilon^{\alpha'}$$

because ρ_ε is supported in $[-\varepsilon, \varepsilon]$. Overall, we have shown that

$$\begin{aligned}\|\mathbb{B}^{\varepsilon,1} - \bar{\mathbb{B}}^1\|_{\alpha,[0,T]} &\leqslant 2\,\|\delta B\|_{\alpha',[-1,T+1]} \sup_{0\leqslant s<t\leqslant T} \frac{(t-s)^{\alpha'} \vee \varepsilon^{\alpha'}}{(t-s)^{\alpha}} \\ &= 2\,\|\delta B\|_{\alpha',[-1,T+1]}\,\varepsilon^{\alpha'-\alpha} \xrightarrow[\varepsilon\downarrow 0]{} 0, \qquad \forall \alpha < \alpha'\end{aligned} \tag{5.15}$$

(for the equality, consider separately $t - s > \varepsilon$ and $t - s \leqslant \varepsilon$). We stress that the previous arguments are *pathwise*. Since $\|\delta B\|_{\alpha',[-1,T+1]} < \infty$ *almost surely* for any $\alpha' < \frac{1}{2}$, it follows that $\|\mathbb{B}^{\varepsilon,1} - \bar{\mathbb{B}}^1\|_\alpha \to 0$ *almost surely* for any $\alpha < \frac{1}{2}$.

To complete the proof of (5.11), it remains to show that $\|\mathbb{B}^{\varepsilon,2} - \bar{\mathbb{B}}^2\|_{2\alpha} \to 0$ in probability as $\varepsilon\downarrow 0$. We distinguish $(\mathbb{B}^{\varepsilon,2} - \bar{\mathbb{B}}^2)^{ij}$ for $i = j$ (diagonal terms) and for $i \neq j$ (off-diagonal terms, in case $d > 1$). To lighten notation, we fix $i \neq j$ and abbreviate $X = B^i$ and $Y = B^j$, which are *independent Brownian motions*.

Diagonal terms are easy: by (5.10) and integration by parts (since X^ε is smooth)

$$(\mathbb{B}^{\varepsilon,2})^{ii}_{st} = \int_s^t (X^\varepsilon_u - X^\varepsilon_s)\,\dot{X}^\varepsilon_u\,\mathrm{d}u = \frac{(X^\varepsilon_t - X^\varepsilon_s)^2}{2}.$$

Similarly $(\bar{\mathbb{B}}^2)^{ii}_{st} = \frac{(X_t - X_s)^2}{2}$ by definition (5.2) of Stratonovich Brownian motion. Since $(\delta X^\varepsilon_{st})^2 - (\delta X_{st})^2 = 2\,\delta X_{st}\,\delta(X^\varepsilon - X)_{st} + (\delta(X^\varepsilon - X)_{st})^2$, by what we already proved on $\|(\mathbb{B}^{\varepsilon,1} - \bar{\mathbb{B}}^1)^i\|_\alpha = \|\delta(X^\varepsilon - X)\|_\alpha$, see (5.15), we have *almost surely*

$$\|(\mathbb{B}^{\varepsilon,2} - \bar{\mathbb{B}}^2)^{ii}\|_{2\alpha} \leqslant \|\delta X\|_\alpha\,\|\delta(X^\varepsilon - X)\|_\alpha + \frac{\|\delta(X^\varepsilon - X)\|^2_\alpha}{2} \xrightarrow[\varepsilon\downarrow 0]{} 0.$$

We next turn to off-diagonal terms $(\mathbb{B}^{\varepsilon,2} - \bar{\mathbb{B}}^2)^{ij} = L^\varepsilon - L$, where we set

$$L_{st} := \int_s^t \delta X_{sw}\,\mathrm{d}Y_w, \qquad L^\varepsilon_{st} := \int_s^t \delta X^\varepsilon_{su}\,\mathrm{d}Y^\varepsilon_u = \int_s^t \delta X^\varepsilon_{su}\,\dot{Y}^\varepsilon_u\,\mathrm{d}u. \tag{5.16}$$

The core of the proof is the following second moment bound, that we prove below.

PROPOSITION 5.4. (SECOND MOMENT BOUND) *For all $\varepsilon > 0$, $s < t$ we have*

$$\mathbb{E}[(L^\varepsilon_{st} - L_{st})^2] \leqslant 10\,(t-s)^2 \min\left\{1, \frac{\varepsilon}{t-s}\right\}. \tag{5.17}$$

We derive from (5.17) a bound for moments of order $p \geqslant 2$ exploiting a key property known as *hypercontractivity*, that we state in the special case which is relevant for us. The proof is given below.

Proposition 5.5. (Hypercontractivity) *Consider the stochastic integral*

$$W := \int_{-\infty}^{\infty} \left(\int_{-\infty}^{t} g(s,t)\, \mathrm{d}X_s \right) \mathrm{d}Y_t \tag{5.18}$$

for a deterministic function $g \in L^2(\mathbb{R}^2 \to \mathbb{R})$. Then the following bound holds:

$$\forall p \in [2,\infty): \qquad \mathbb{E}[|W|^p] \leqslant c_p^2\, \mathbb{E}[W^2]^{\frac{p}{2}}, \tag{5.19}$$

with $c_p := \mathbb{E}[|\mathcal{N}(0,1)|^p] < \infty$.

We can now apply (5.19) to $L_{st}^{\varepsilon} - L_{st}$, which is of the form (5.18) (see (5.24) below): plugging (5.17) into (5.19) we obtain

$$\mathbb{E}[|L_{st}^{\varepsilon} - L_{st}|^p] \leqslant 10\, c_p^2\, (t-s)^p \min\left\{1, \left(\frac{\varepsilon}{t-s}\right)^{\frac{p}{2}}\right\}. \tag{5.20}$$

Since $\min\{1,x\} \leqslant x^{\kappa}$ for all $x \geqslant 0$ and $\kappa \in [0,1]$, it follows that

$$\forall \kappa \in (0,1]: \qquad \mathbb{E}[|L_{st}^{\varepsilon} - L_{st}|^p] \leqslant 10\, c_p^2\, (t-s)^{p\left(1-\frac{\kappa}{2}\right)}\, \varepsilon^{p\frac{\kappa}{2}}. \tag{5.21}$$

We now fix $\alpha < \frac{1}{2}$ and exploit Theorem 4.6 for $A_{st} := L_{st}^{\varepsilon} - L_{st}$ with $\rho = \alpha$ and $\gamma = 2\alpha$. We need to control the *random constants $Q_{2\alpha}$ and $K_{\alpha,2\alpha}$* from (4.20)-(4.21).

- For $Q_{2\alpha}$ we apply Proposition 4.8 with $\gamma_0 = 1 - \frac{\kappa}{2}$: if we take $\kappa > 0$ small and $p \geqslant 2$ large, so that (4.26) is satisfied, by (5.21) and (4.27) we get

$$\mathbb{E}[Q_{2\alpha}^p] \leqslant \mathfrak{C}\, \varepsilon^{p\frac{\kappa}{2}} \qquad \text{with} \quad \mathfrak{C} = \mathfrak{C}_{p,\alpha,\kappa} := \frac{10\, c_p^2}{1 - 2^{1-p\left(1-2\alpha-\frac{\kappa}{2}\right)}}. \tag{5.22}$$

  This implies that $Q_{2\alpha} \to 0$ *in probability* as $\varepsilon \downarrow 0$, and even *almost surely* along sequences $\varepsilon = \varepsilon_n \downarrow 0$ which vanish exponentially fast.

- For $K_{\alpha,2\alpha}$ we note that, by the Chen relation,

$$\delta A_{sut} = \delta X_{su}^{\varepsilon}\, \delta Y_{ut}^{\varepsilon} - \delta X_{su}\, \delta Y_{ut} = \delta X_{su}^{\varepsilon}\, (\delta Y_{ut}^{\varepsilon} - \delta Y_{ut}) + \delta Y_{ut}\, (\delta X_{su}^{\varepsilon} - \delta X_{su}),$$

  therefore by (5.13) and (5.15), if we fix any $\alpha' \in \left(\alpha, \frac{1}{2}\right)$, we can bound

$$\begin{aligned} K_{\alpha,2\alpha} &\leqslant \|\delta X^{\varepsilon}\|_{\alpha}\, \|\delta(Y^{\varepsilon} - Y)\|_{\alpha} + \|\delta Y\|_{\alpha}\, \|\delta(X^{\varepsilon} - X)\|_{\alpha} \\ &\leqslant 2\, (\|\delta X\|_{\alpha}\, \|\delta Y\|_{\alpha'} + \|\delta Y\|_{\alpha}\, \|\delta X\|_{\alpha'})\, \varepsilon^{\alpha' - \alpha}. \end{aligned} \tag{5.23}$$

  This shows that $Q_{2\alpha} \to 0$ *almost surely* as $\varepsilon \downarrow 0$.

We can finally apply (4.22) to conclude that, by (5.22) and (5.23),

$$\|(\mathbb{B}^{\varepsilon,2} - \bar{\mathbb{B}}^2)^{ij}\|_{2\alpha} = \|L^{\varepsilon} - L\|_{2\alpha} \leqslant C_{\alpha,2\alpha}\, (Q_{2\alpha} + K_{\alpha,2\alpha}) \xrightarrow[\varepsilon \downarrow 0]{} 0 \qquad \text{in probability.}$$

This completes the proof of (5.11). $\square$

Proof. (OF THEOREM 5.4) Recalling that $X^\varepsilon = \rho_\varepsilon * X$ and $Y^\varepsilon = \rho_\varepsilon * Y$, an integration by parts for the stochastic (Wiener) integral yields for $s < t$

$$\begin{aligned}\delta X^\varepsilon_{st} &:= \int_{\mathbb{R}} (\rho_\varepsilon(t-v) - \rho_\varepsilon(s-v))\, X_v\, \mathrm{d}v = \int_{\mathbb{R}} \left(\int_{s-v}^{t-v} \rho_\varepsilon(r)\, \mathrm{d}r \right) \mathrm{d}X_v,\\ \dot{Y}_\varepsilon(t) &:= \int_{\mathbb{R}} (\rho_\varepsilon)'(t-w)\; Y_w\, \mathrm{d}w = \int_{\mathbb{R}} \rho_\varepsilon(t-w)\, \mathrm{d}Y_w.\end{aligned}$$

Recalling the definition (5.16) of L_{st} and L^ε_{st}, we can write

$$L^\varepsilon_{st} - L_{st} = \iint (g_\varepsilon^{(s,t)}(v,w) - \mathbb{1}_{(s\le v\le w\le t)})\, \mathrm{d}X_v\, \mathrm{d}Y_w\,, \tag{5.24}$$

where we set

$$\begin{aligned} g_\varepsilon^{(s,t)}(v,w) &:= \int_s^t \rho_\varepsilon(u-w) \left(\int_{s-v}^{u-v} \rho_\varepsilon(r)\, \mathrm{d}r \right) \mathrm{d}u\\ &= \int \mathbb{1}_{(s\leqslant r\leqslant u\leqslant t)}\, \rho_\varepsilon(r-v)\, \rho_\varepsilon(u-w)\, \mathrm{d}r\, \mathrm{d}u.\end{aligned}$$

Since $0 \leqslant g_\varepsilon^{(s,t)}(v,w) \leqslant 1$ (recall that $\rho_\varepsilon(\cdot)$ is a probability density), it follows that

$$\begin{aligned}\mathbb{E}[(L^\varepsilon_{st} - L_{st})^2] &= \iint (g_\varepsilon^{(s,t)}(v,w) - \mathbb{1}_{(s\le v\le w\le t)})^2\, \mathrm{d}v\, \mathrm{d}w\\ &\leqslant \iint |g_\varepsilon^{(s,t)}(v,w) - \mathbb{1}_{(s\le v\le w\le t)}|\, \mathrm{d}v\, \mathrm{d}w. \end{aligned}\tag{5.25}$$

To estimate this integral, we give a probabilistic representation of $g_\varepsilon(v,w)$: denoting by Q_1 and Q_2 two independent random variables with density $\rho(\cdot)$, since $\rho_\varepsilon(\cdot - v)$ and $\rho_\varepsilon(\cdot - w)$ are the densities of $\varepsilon\, Q_1 + v$ and $\varepsilon\, Q_2 + w$, we can write

$$g_\varepsilon^{(s,t)}(v,w) = \mathbb{P}(s \le \varepsilon Q_1 + v \le \varepsilon Q_2 + w \leqslant t)\,.$$

Writing $v = s + a\,(t-s)$ and $w = s + b\,(t-s)$, for new variables a, b, we note that

$$g_\varepsilon^{(s,t)}(s + a\,(t-s), s + b\,(t-s)) = g_\delta^{(0,1)}(a,b) \qquad \text{with} \quad \delta := \frac{\varepsilon}{t-s}\,.$$

A change of variables in the integral (5.25) then yields

$$\mathbb{E}[(L^\varepsilon_{st} - L_{st})^2] \leqslant (t-s)^2 \iint |g_\delta^{(0,1)}(a,b) - \mathbb{1}_{(0\le a\le b\le 1)}|\, \mathrm{d}a\, \mathrm{d}b\,.$$

Looking at our goal (5.17), it only remains to show that

$$\iint |g_\delta^{(0,1)}(a,b) - \mathbb{1}_{(0\le a\le b\le 1)}|\, \mathrm{d}a\, \mathrm{d}b \leqslant 10 \min\{1, \delta\}\,. \tag{5.26}$$

We define the subset

$$D := \{(a,b) \in \mathbb{R}^2 \colon \quad 0 \le a \le b \le 1\}$$

so that we can write

$$g_\delta^{(0,1)}(a,b) = \mathbb{P}(0 \le \delta\, Q_1 + a \le \delta\, Q_2 + b \leqslant 1) = \mathbb{E}[\mathbb{1}_{D-\delta Q}(a,b)] \qquad \text{with } Q := (Q_1, Q_2)\,.$$

We can express the integral in (5.26) as

$$\begin{aligned}\iint |g_\delta^{(0,1)}(a,b) - \mathbb{1}_{(0\le a\le b\le 1)}|\,\mathrm{d}a\mathrm{d}b &= \mathbb{E}\bigg[\int_{\mathbb{R}^2} |\mathbb{1}_{D-\delta Q}(z) - \mathbb{1}_D(z)|\,\mathrm{d}z\bigg]\\ &= \mathbb{E}[|(D-\delta\,Q)\triangle D|]\end{aligned}$$

where $|\cdot|$ denotes Lebesgue measure in $\mathbb{R}^2$ and $A\triangle B := (A\cap B^c)\cup(A^c\cap B)$ is the symmetric difference between sets. Note that $z\in(D-y)\triangle D$ means that either $z\in D$ but $z+y\in D^c$, or $z\in D^c$ but $z+y\in D$, and in both cases $\mathrm{dist}(z,\partial D)\leqslant|y|$, where ∂D is the boundary of D. In other terms, for any $y\in\mathbb{R}^2$ we have the inclusion

$$(D-y)\triangle D\subseteq\{z\in\mathbb{R}^2:\quad \mathrm{dist}(z,\partial D)\leqslant|y|\}\,.$$

Since ∂D is a triangle with perimeter $2+\sqrt{2}$, the area of $\{z\in\mathbb{R}^2:\quad \mathrm{dist}(z,\partial T)\leqslant|y|\}$ is bounded above by $2\,(2+\sqrt{2})\,|y|$, hence

$$\mathbb{E}[|(D-\delta\,Q)\triangle D|]\leqslant 2\,(2+\sqrt{2})\,\mathbb{E}[|\delta\,Q|]\leqslant 2\,(2+\sqrt{2})\,\sqrt{2}\,\delta,$$

because $|Q|=\sqrt{Q_1^2+Q_2^2}\leqslant\sqrt{2}$ (we recall that $\rho(\cdot)$ is supported in $[-1,1]$, hence $|Q_1|,|Q_2|\leqslant 1$). Since $2\,(2+\sqrt{2})\,\sqrt{2}\leqslant 10$, the proof of (5.26) is completed. □

Proof. (of Proposition 5.5) By (5.18) we can write $W=\int_{-\infty}^{\infty} h(t)\,\mathrm{d}Y_t$ where $h(t)=h(X,t):=\int_{-\infty}^{t} g(s,t)\,\mathrm{d}X_s$ depends only on X. Since X and Y are independent, it follows that W *is a Gaussian random variable conditionally on* X, as a Wiener integral. Recalling that $c_p:=\mathbb{E}[|\mathcal{N}(0,1)|^p]$, we can thus write

$$\mathbb{E}[|W|^p|X]=c_p\,\mathbb{E}[W^2|X]^{\frac{p}{2}},$$

We now denote by $E=C(\mathbb{R},\mathbb{R})$ the standard path space for X and Y, so that can write $W=f(X,Y)$ for a suitable measurable function $f\colon E\times E\to\mathbb{R}$. Denoting by μ the law of X, i.e. the two-sided Wiener measure, Fubini's theorem yields

$$\mathbb{E}[W^2|X]=\mathbb{E}[f(x,Y)^2]|_{x=X}=(\|f(x,y)\|_{L^2(\mu(\mathrm{d}y))}^2)|_{x=X},$$

hence

$$\mathbb{E}[|W|^p]=c_p\,\mathbb{E}\Big[\mathbb{E}[|W|^2|X]^{\frac{p}{2}}\Big]=c_p\,(\|\|f(x,y)\|_{L^2(\mu(\mathrm{d}y))}\|_{L^p(\mu(\mathrm{d}x))})^p\,.$$

We now apply the *Minkowski integral inequality* (see Remark 5.6 below), which states that for $p\geqslant 2$ switching the two norms yields an upper bound:

$$\begin{aligned}\mathbb{E}[|W|^p] &\leqslant c_p\,(\|\|f(x,y)\|_{L^p(\mu(\mathrm{d}x))}\|_{L^2(\mu(\mathrm{d}y))})^p\\ &= c_p\Big(\Big\|\mathbb{E}[|f(X,y)|^p]^{\frac{1}{p}}\Big\|_{L^2(\mu(\mathrm{d}y))}\Big)^p.\end{aligned}\tag{5.27}$$

We finally observe that $f(X,y)$ is a Gaussian random variable, i.e. $W=f(X,Y)$ *is Gaussian conditionally on* Y (because $W=\int\tilde{h}(s)\,\mathrm{d}X_s$ with $\tilde{h}(t):=\int_s^{\infty} g(s,t)\,\mathrm{d}Y_t$ is a Wiener integral conditionally on Y, by independence of X and Y). It follows that

$$\mathbb{E}[|f(X,y)|^p]=c_p\,\mathbb{E}[f(X,y)^2]^{\frac{p}{2}}=c_p\,(\|f(x,y)\|_{L^2(\mu(\mathrm{d}x))})^p.\tag{5.28}$$

Plugging (5.28) into (5.27) we obtain (5.19), since

$$\| \| f(x,y)\|_{L^2(\mu(\mathrm{d}x))}\|_{L^2(\mu(\mathrm{d}y))} = \mathbb{E}[W^2]^{1/2}$$

by Fubini's theorem. □

Remark 5.6. (MINKOWSKI'S INTEGRAL INEQUALITY) Given σ-finite measure spaces (E,μ) and (F,ν) and a measurable function $f: E\times F\to\mathbb{R}$, Minkowski's integral inequality states that for any $0<q\leqslant p\leqslant\infty$

$$\| \| f(x,y)\|_{L^q(E,\mu(\mathrm{d}x))}\|_{L^p(F,\nu(\mathrm{d}y))} \leqslant \| \| f(x,y)\|_{L^p(F,\nu(\mathrm{d}y))}\|_{L^q(E,\mu(\mathrm{d}x))}\,. \tag{5.29}$$

For $q=p$ this holds an equality, as a consequence of Fubini's theorem. If $q<p$, the proof goes as follows: if the left-hand side of (5.29) is equal to zero, there is nothing to prove; if it is not, then raising it to power p gives, by Fubini's theorem,

$$\begin{aligned}
\int_F\biggl(\int_E |f|^q\,\mathrm{d}\mu\biggr)^{\frac{p}{q}}\mathrm{d}\nu &= \int_F\biggl[\int_E |f|^q\,\biggl(\int_E |f|^q\,\mathrm{d}\mu\biggr)^{\frac{p}{q}-1}\mathrm{d}\mu\biggr]\mathrm{d}\nu\\
&= \int_E\biggl[\int_F |f|^q\,\biggl(\int_E |f|^q\,\mathrm{d}\mu\biggr)^{\frac{p}{q}-1}\mathrm{d}\nu\biggr]\mathrm{d}\mu\\
&\leqslant \int_E\Biggl[\biggl(\int_F |f|^p\,\mathrm{d}\nu\biggr)^{\frac{q}{p}}\Biggl\{\int_F\biggl(\int_E |f|^q\,\mathrm{d}\mu\biggr)^{\frac{p-q}{q}\frac{p}{p-q}}\mathrm{d}\nu\Biggr\}^{\frac{p-q}{p}}\Biggr]\mathrm{d}\mu\\
&= \Biggl\{\int_F\biggl(\int_E |f|^q\,\mathrm{d}\mu\biggr)^{\frac{p}{q}}\mathrm{d}\nu\Biggr\}^{\frac{p-q}{p}}\int_E\biggl(\int_F |f|^p\,\mathrm{d}\nu\biggr)^{\frac{q}{p}}\mathrm{d}\mu
\end{aligned}$$

where we have used the Hölder inequality on (F,ν) with conjugated exponents $\frac{p}{q}$ and $\frac{p}{p-q}$. The first term in the last line is the left-hand side raised to power $\frac{p-q}{p}$: dividing by such term (which is not zero by assumption) we obtain (5.29).

Note that for $q=1$ we have additionally, since $|\int_E f\,\mathrm{d}\mu| \leqslant \int_E |f|\,\mathrm{d}\mu$,

$$\biggl[\int_F\biggl|\int_E f\,\mathrm{d}\mu\biggr|^p\mathrm{d}\nu\biggr]^{\frac{1}{p}} \leqslant \int_E\biggl|\int_F |f|^p\,\mathrm{d}\nu\biggr|^{\frac{1}{p}}\mathrm{d}\mu.$$

In the special case $E=\{1,2\}$ with $\mu=\delta_1+\delta_2$, if we set $f_i(\cdot):=f(i,\cdot)$, then for $p\geqslant 1$ we recover the usual Minkowski inequality $\|f_1+f_2\|_{L^p}\leqslant\|f_1\|_{L^p}+\|f_2\|_{L^p}$.