CHAPTER 8

Stratonovich's Theory

From an abstract mathematical standpoint, Itô's theory of stochastic integration has as a serious flaw: it behaves dreadfully under changes of coordinates (cf. Remark 3.3.6). In fact, Itô's formula itself is the most dramatic manifestation of this problem.

The origin of the problems Itô's theory has with coordinate changes can be traced back to its connection with independent increment processes. Indeed, the very notion of an independent increment process is inextricably tied to the linear structure of Euclidean space, and anything but a linear change of coordinates will wreak havoc to that structure. Generalizations of Itô's theory like the one of Kunita and Watanabe do not really cure this problem, they only make it slightly less painful.

To make Itô's theory more amenable to coordinate changes, we will develop an idea which was introduced by R.L. Stratonovich. Stratonovich was motivated by applications to engineering, and his own treatment [34] had some mathematically awkward aspects. In fact, it is ironic, if not surprising, that Itô [12] was the one who figured out how to put Stratonovich's ideas on a firm mathematical foundation.

8.1 Semimartingales & Stratonovich Integrals

From a technical perspective, the coordinate change problem alluded to above is a consequence of the fact that non-linear functions destroy the martingale property. Thus, our first step will be to replace martingales with a class of processes which is invariant under composition with non-linear functions. Throughout, $(\Omega, \mathcal{F}, \mathbb{P})$ will be a complete probability space which comes equipped with a non-decreasing family $\{\mathcal{F}_t : t \geq 0\}$ of $\mathbb{P}$-complete σ-algebras.

8.1.1. Semimartingales. We will say that $Z : [0, \infty) \times \Omega \longrightarrow \mathbb{R}$ is a continuous *semimartingale*[1] and will write $Z \in \mathcal{S}(\mathbb{P}; \mathbb{R})$ if Z is can written in the form $Z(t, \omega) = M(t, \omega) + B(t, \omega)$, where $\big(M(t), \mathcal{F}_t, \mathbb{P}\big)$ is a continu-

[1] It is unfortunate that Doob originally adopted this term for the class of processes which are now called submartingales. However, the confusion caused by this terminological accident recedes along with the generation of probabilists for whom it was a problem.

ous local martingale and B is a progressively measurable function with the property that $B(\,\cdot\,,\omega)$ is a continuous function of locally bounded variation for $\mathbb{P}$-almost every ω.

Notice that when we insist that $M(0) \equiv 0$ the decomposition a semimartingale Z into to *martingale part* M and its *bounded variation part* B is almost surely unique. This is just an application of the first statement in Corollary 7.1.2. In addition, Itô's formula (cf. Theorem 7.2.9) shows that the class of continuous semimartingales is invariant under composition with twice continuously differentiable functions f. In fact, his formula says that if $Z(t) = \big(M_1(t)+B_1(t),\dots,Z_n(t)+B_n(t)\big)$ is an $\mathbb{R}^n$-valued, continuous semimartingale and $F \in C^2(\mathbb{R}^n;\mathbb{R})$, then the martingale and bounded variation parts of $F \circ Z$ are, respectively,

$$\sum_{i=1}^n \int_0^t \partial_i F\big(Z(\tau)\big)\, dM_i(\tau) \quad \text{and}$$
$$F\big(Z(0)\big) + \frac{1}{2}\sum_{i,j=1}^n \int_0^t \partial_i\partial_j F\big(Z(\tau)\big)\, \langle M_i, M_j\rangle(d\tau) + \sum_{i=1}^n \int_0^t \partial_i F\big(Z(\tau)\big)\, B_i(d\tau),$$

where, of course, the integrals in the first line are taken in the sense of Itô and the ones in the second are (or Riemann) Lebesgue integrals.

In various circumstances it is useful to have defined the integral

$$\int_0^t \theta(\tau)\, dZ(t) = \int_0^t \theta(\tau)\, dM(\tau) + \int_0^t \theta(\tau)\, B(d\tau)$$
$$\text{for } Z = M + B \text{ and } \theta \in \Theta^2_{\text{loc}}(\langle M\rangle, \mathbb{P}; \mathbb{R}),$$

where the $dM(\tau)$-integral is taken in the sense of Itô and the $B(d\tau)$ integral is a taken a la Lebesgue. Also, we take

$$\langle Z_1, Z_2\rangle(t) \equiv \langle M_1, M_2\rangle(t) \quad \text{if } Z_i = M_i + B_i \text{ for } i \in \{1,2\}. \tag{8.1.1}$$

Indeed, with this notation, Itô's formula 7.2.9 becomes

$$\begin{aligned} F\big(Z(t)\big) - F\big(Z(0)\big) = & \sum_{i=1}^n \int_0^t \partial_i F\big(Z(t)\big)\, dZ_i(t) \\ & + \frac{1}{2}\sum_{i,j=1}^n \int_0^t \partial_i\partial_j F\big(Z(\tau)\big)\, \langle Z_i, Z_j\rangle(d\tau). \end{aligned} \tag{8.1.2}$$

The notational advantage of (8.1.2) should be obvious. On the other hand, the disadvantage is that it mixes the martingale and bounded variation parts of the right hand side, and this mixing has to be disentangled before expectation values are taken.

Finally, it should be noticed that the extension of the bracket given in (8.1.1) is more than a notational device and has intrinsic meaning. Namely, by Exercise 7.2.13 and the easily checked fact that, as $N \to \infty$,

$$\sum_{m=0}^{2^N t}\Big|\Big(Z_1\big((m+1)2^{-N}\big) - Z_1\big(m2^{-N}\big)\Big)\Big(Z_2\big((m+1)2^{-N}\big) - Z_2\big(m2^{-N}\big)\Big)$$
$$- \Big(M_1\big((m+1)2^{-N}\big) - M_1\big(m2^{-N}\big)\Big)\Big(M_2\big((m+1)2^{-N}\big) - M_2\big(m2^{-N}\big)\Big)\Big|,$$

tends to 0 $\mathbb{P}$-almost surely uniformly for t in compacts, we know that

$$\begin{aligned}\sum_{m=0}^{\infty}&\Big(Z_1\big(t\wedge(m+1)2^{-N}\big) - Z_1\big(t\wedge m2^{-N}\big)\Big)\\ &\times\Big(Z_2\big(t\wedge(m+1)2^{-N}\big) - Z_2\big(t\wedge m2^{-N}\big)\Big) \longrightarrow \langle Z_1, Z_2\rangle(t)\end{aligned} \tag{8.1.3}$$

in $\mathbb{P}$-probability uniformly for t in compacts.

8.1.2. Stratonovich's Integral. Keeping in mind that Stratonovich's purpose was to make stochastic integrals better adapted to changes in variable, it may be best to introduce his integral by showing that it is integral at which one arrives if one adopts a somewhat naive approach, one which reveals its connection to the Riemann integral. Namely, given elements X and Y of $\mathcal{S}(\mathbb{P};\mathbb{R})$, we want to define the *Stratonovich integral*[2] $\int_0^t Y(\tau)\circ dX(\tau)$ of Y with respect to X as the limit of the Riemann integrals

$$\lim_{N\to\infty}\int_0^t Y^N(\tau)\,dX^N(\tau),$$

where, for $\alpha : [0,\infty) \longrightarrow \mathbb{R}$, we use α^N to denote the polygonal approximation of α obtained by linear interpolation on each interval $[m2^{-N}, (m+1)2^{-N}]$. That is,

$$\begin{gathered}\alpha^N(t) = \alpha(m2^{-N}) + 2^N(t - m2^{-N})\Delta_m^N\alpha\\ \text{where } \Delta_m^N\alpha \equiv \alpha\big((m+1)2^{-N}\big) - \alpha\big(m2^{-N}\big)\end{gathered}$$

for all m, $N \in \mathbb{N}$ and $t \in \big[m2^{-N}, (m+1)2^{-N}\big]$.

[2] This notation is only one of many which are used to indicate it is Stratonovich's, as opposed to Itô's , sense in which an integral is being taken. Others include the use of $\delta X(\tau)$ in place of $\circ dX(\tau)$. The lack of agreement about the choice of notation reflects the inadequateness of all the variants.

Of course, before we can adopt this definition, we are obliged to check that the this limit exists. For this purpose, write

$$\int_0^t Y^N(\tau)\,dX^N(\tau) = \int_0^t Y([\tau]_N)\,dX(\tau) + \int_0^t \big(Y^N(\tau) - Y([\tau]_N)\big)\,dX^N(\tau),$$

where the dX-integral on the right can be interpreted either as an Itô or a Riemann integral. By using Corollary 7.2.3, one can easily check that

$$\left\| \int_0^t Y([\tau]_N)\,dX(\tau) - \int_0^t Y(\tau)\,dX(\tau) \right\|_{[0,T]} \longrightarrow 0 \quad \text{in } \mathbb{P}\text{-probability.}$$

At the same time, by (8.1.3),

$$\begin{aligned} &\int_0^t \big(Y^N(\tau) - Y([\tau]_N)\big)\,dX^N(\tau) \\ &= 4^N \sum_{m \le 2^N t} \big(\Delta_m^N X\big)\big(\Delta_m^N Y\big) \int_{m2^{-N}}^{t \wedge (m+1)2^{-N}} (\tau - m2^{-N})\,d\tau \longrightarrow \frac{1}{2}\langle X, Y\rangle(t) \end{aligned}$$

in $\mathbb{P}$-probability uniformly for t in compacts. Hence, for computational purposes, it is best to present the Stratonovich integral as

$$\int_0^t Y(\tau) \circ dX(\tau) = \int_0^t Y(\tau)\,dX(\tau) + \frac{1}{2}\langle X, Y\rangle(t), \tag{8.1.4}$$

where integral on the right is taken in the sense of Itô .

So far as I know, the formula in (8.1.4) was first given by Itô in [12]. In particular, Itô seems to have been the one who realized that the problems posed by Stratonovich's theory could be overcome by insisting that the integrands be semimartingales, in which case, as (8.1.4) makes obvious, *the Stratonovich integral of Y with respect to X is again a continuous semimartingale*. In fact, if $X = M + B$ is the decomposition of X into its martingale and bounded variation parts, then

$$\int_0^t Y(\tau)\,dM(\tau) \quad \text{and} \quad \int_0^t Y(\tau)\,dB(\tau) + \frac{1}{2}\langle Y, M\rangle(t)$$

are the martingale and bounded variation parts of $I(t) \equiv \int_0^t Y(\tau) \circ dX(\tau)$, and so $\langle Z, I\rangle(dt) = Y(t)\langle Z, X\rangle(dt)$ for all $Z \in \mathcal{S}(\mathbb{P}; \mathbb{R})$.

To appreciate how clever Itô's observation is, notice that Stratonovich's integral is not really an integral at all. Indeed, in order to deserve being called an *integral*, an operation should result in a quantity which can be estimated in terms of zeroth order properties of the integrand. On the other hand, as (8.1.4) shows, no such estimate is possible. To see this, take $Y(t) = f\big(Z(t)\big)$, where $Z \in \mathcal{S}(\mathbb{P};\mathbb{R})$ and $f \in C^2(\mathbb{R};\mathbb{R})$. Then, because $\langle X, Y\rangle(dt) = f'\big(Z(t)\big)\langle X, Z\rangle(dt)$,

$$\int_0^t Y(\tau) \circ dX(\tau) = \int_0^t f\big(Z(\tau)\big)\, dX(\tau) + \frac{1}{2}\int_0^t f'\big(Z(\tau)\big)\,\langle X, Z\rangle(d\tau),$$

which demonstrates that there is, in general, no estimate of the Stratonovich integral in terms of the zeroth order properties of the integrand.

Another important application of (8.1.4) is to the behavior of Stratonovich integrals under iteration. That is, suppose that $X, Y \in \mathcal{S}(\mathbb{P};\mathbb{R})$, and set $I(t) = \int_0^t Y(\tau) \circ dX(\tau)$. Then $\langle Z, I\rangle(dt) = Y(t)\langle Z, X\rangle(dt)$ for all $Z \in \mathcal{S}(\mathbb{P};\mathbb{R})$, and so

$$\begin{aligned}
\int_0^t Z(\tau) \circ dI(\tau) &= \int_0^t Z(\tau)\, dI(\tau) + \frac{1}{2}\int_0^t Y(\tau)\,\langle Z, X\rangle(d\tau) \\
&= \int_0^t Z(\tau)Y(\tau)\, dX(\tau) + \frac{1}{2}\int_0^t \Big(Z(\tau)\langle Y, X\rangle(d\tau) + Y(\tau)\langle Z, X\rangle(d\tau)\Big) \\
&= \int_0^t Z(\tau)Y(\tau)\, dX(\tau) + \frac{1}{2}\langle ZY, X\rangle(t) = \int_0^t ZY(\tau) \circ dX(\tau),
\end{aligned}$$

since, by Itô's formula,

$$ZY(t) - ZY(0) = \int_0^t Z(\tau)\, dY(\tau) + \int_0^t Y(\tau)\, dZ(\tau) + \frac{1}{2}\langle Z, Y\rangle(t),$$

and therefore $\langle ZY, X\rangle(dt) = Z(t)\langle Y, X\rangle(dt) + Y(t)\langle Z, X\rangle(dt)$. In other words, we have now proved that

$$\int_0^t Z(\tau) \circ d\left(\int_0^\tau Y(\sigma) \circ X(\sigma)\right) = \int_0^t Z(\tau)Y(\tau) \circ dX(\tau). \tag{8.1.5}$$

8.1.3. Ito's Formula & Stratonovich Integration. Because the origin of Stratonovich's integral is in Riemann's theory, it should come as no surprise that Itô's formula looks deceptively like the fundamental theorem of calculus when Stratonovich's integral is used. Namely, let $Z = (Z_1, \dots, Z_n) \in \mathcal{S}(\mathbb{P};\mathbb{R})^n$ and $f \in C^3(\mathbb{R}^n;\mathbb{R})$ be given, and set $Y_i = \partial_i f \circ Z$ for $1 \le i \le n$. Then $Y_i \in \mathcal{S}(\mathbb{P};\mathbb{R})$ and, by Itô's formula applied to $\partial_i f$, the local martingale part of Y_i is given by the sum over $1 \le j \le n$ of the Itô stochastic integrals of $\partial_i\partial_j f \circ Z$ with respect the local martingale part of Z_j. Hence,

$$\langle Y_i, Z_i\rangle(dt) = \sum_{j=1}^n \partial_i\partial_j f \circ Z(\tau)\langle Z_i, Z_j\rangle(dt).$$

But, by Itô's formula applied to f, this means that

$$df\bigl(Z(t)\bigr) = \sum_{i=1}^{n}\Bigl(Y_i(\tau)\,dZ_i(\tau) + \frac{1}{2}\langle Y_i, Z_i\rangle(dt)\Bigr),$$

and so we have now shown that, in terms of Stratonovich integrals, Itô's formula does look like the "the fundamental theorem of calculus"

$$f\bigl(Z(t)\bigr) - f\bigl(Z(0)\bigr) = \sum_{i=1}^{n}\int_0^t \partial_i f\bigl(Z(\tau)\bigr)\circ dZ_i(\tau). \tag{8.1.6}$$

As I warned above, (8.1.6) is deceptive. For one thing, as its derivation makes clear, it, in spite of its attractive form, is really just Itô's formula (8.1.2) in disguise. In fact, if, as we will, one adopts Itô's approach to Stratonovich's integral, then it is not even that. Indeed, one cannot write (8.1.6) unless one knows that $\partial_i f \circ Z$ is a semi-martingale. Thus, in general, (8.1.6) requires us to assume that f is three times continuously differentiable, not just twice, as in the case with (8.1.2). Ironically, we have arrived at a first order fundamental theorem of calculus which applies only to functions with three derivatives. In view of these remarks, it is significant that, at least in the case of Brownian motion, Itô found a way (cf. [13] or Exercise 8.1.8 below) to make (8.1.6) closer to a *true* Fundamental Theorem of Calculus, at least in the sense that it applies to all $f \in C^1(\mathbb{R}^n;\mathbb{R})$.

REMARK 8.1.7. Putting (8.1.6) together with our earlier footnote about the notation for Stratonovich integrals, one might be inclined to think the *right* notation should be $\int_0^t Y(\tau)\dot{X}(\tau)\,d\tau$. For one thing, this notation recognizes that the Stratonovich integral is closely related to the notion of generalized derivatives *a la* Schwartz's distribution theory. Secondly, (8.1.6) can be summarized in differential form by the expression

$$df\bigl(Z(t)\bigr) = \sum_{i=1}^{n}\partial_i f\bigl(Z(t)\bigr)\circ dZ_i(t),$$

which would take the appealing form

$$\frac{d}{dt}f\bigl(Z(t)\bigr) = \sum_{i=1}^{n}\partial_i f\bigl(Z(t)\bigr)\dot{Z}(t).$$

Of course, the preceding discussion should also make one cautious about being too credulous about all this.

8.1.4. Exercises.

EXERCISE 8.1.8. In this exercise we will describe Itô's approach to extending the validity of (8.1.6).

(i) The first step is to give another description of Stratonovich integrals. Namely, given $X, Y \in \mathcal{S}(\mathbb{P};\mathbb{R})$, show that the Stratonovich integral of Y with respect to X over the interval $[0,T]$ is almost surely equal to

$$(*) \qquad \lim_{N\to\infty} \sum_{m=0}^{2^N-1} \frac{Y\left(\frac{(m+1)T}{2^N}\right) + Y\left(\frac{mT}{2^N}\right)}{2} \left(X\left(\frac{(m+1)T}{2^N}\right) - X\left(\frac{mT}{2^N}\right)\right).$$

Thus, even if Y is not a semimartingale, we will say that $Y : [0,T] \times \Omega \longrightarrow \mathbb{R}$ is *Stratonvich integrable* on $[0,T]$ with respect to X if the limit in (*) exists in $\mathbb{P}$-measure, in which case we will use $\int_0^T Y(t) \circ dX(t)$ to denote this limit. It must be emphasized that the definition here is "T by T" and not simultaneous for all T's in an interval.

(ii) Given an $X \in \mathcal{S}(\mathbb{P};\mathbb{R})$ and a $T > 0$, set $\check{X}^T(t) = X\big((T-t)^+\big)$, and suppose that $\big(\check{X}^T(t), \check{\mathcal{F}}_t^T, \mathbb{P}\big)$ is a semimartingale relative to some filtration $\{\check{\mathcal{F}}_t^T : t \geq 0\}$. Given a $Y : [0,T] \times \Omega \longrightarrow \mathbb{R}$ with the properties that $Y(\,\cdot\,,\omega) \in C\big([0,T];\mathbb{R}\big)$ for $\mathbb{P}$-almost every ω and that $\omega \rightsquigarrow Y(t,\omega)$ is $\mathcal{F}_t \cap \check{\mathcal{F}}_t^T$ for each $t \in [0,T]$, show that Y is Stratonovich integrable on $[0,T]$ with respect to X. In fact, show that

$$\int_0^T Y(t) \circ dX(t) = \frac{1}{2}\int_0^T Y(t)\, dX(t) - \frac{1}{2}\int_0^T Y(t)\, d\check{X}^T(t),$$

where each of the integrals on the right is the taken in the sense of Itô .

(iii) Let $\big(\beta(t), \mathcal{F}_t, \mathbb{P}\big)$ be an $\mathbb{R}^n$-valued Brownian motion. Given $T \in (0,\infty)$, set $\check{\beta}^T(t) = \beta\big((T-t)^+\big)$, $\check{\mathcal{F}}_t^T = \sigma\big(\{\check{\beta}^T(\tau) : \tau \in [0,t]\}\big)$, and show that, for each $\xi \in \mathbb{R}^n$, $\big((\xi, \check{\beta}^T(t))_{\mathbb{R}^n}, \check{\mathcal{F}}_t^T, \mathbb{P}\big)$ is a semimartingale with $\langle \check{\beta}^T, \check{\beta}^T \rangle(t) = t \wedge T$ and bounded variation part $t \rightsquigarrow -\int_0^t \frac{\check{\beta}_\tau^T}{T-\tau}\, d\tau$. In particular, show that, for each $\xi \in \mathbb{R}^n$ and $g \in C(\mathbb{R}^n;\mathbb{R})$, $t \rightsquigarrow g\big(\beta(t)\big)$ is Stratonovich integrable on $[0,T]$ with respect to $t \rightsquigarrow \big(\xi, \beta(t)\big)_{\mathbb{R}^n}$. Further, if $\{g_n\}_1^\infty \subseteq C(\mathbb{R}^n;\mathbb{R})$ and $g_n \longrightarrow g$ uniformly on compacts, show that

$$\int_0^T g_n\big(\beta(t)\big) \circ d\big(\xi, \beta(t)\big)_{\mathbb{R}^n} \longrightarrow \int_0^T g\big(\beta(t)\big) \circ d\big(\xi, \beta(t)\big)_{\mathbb{R}^n}$$

in $\mathbb{P}$-measure.

(iv) Continuing with the notation in **(iii)**, show that

$$f\big(\beta(T)\big) - f(0)) = \sum_{i=1}^n \int_0^T \partial_i f\big(\beta(t)\big) \circ d\beta_i(t)$$

for every $f \in C^1(\mathbb{R}^n;\mathbb{R})$. In keeping with the comment at the end of **(i)**, it is important to recognize that although this form of Itô 's formula holds for all continuously differentiable functions, it is, in many ways less useful than forms which we obtained previously. In particular, when f is no better than once differentiable, the right hand side is defined only up to a $\mathbb{P}$-null set for each T and not for all T's simultaneously.

(v) Let $\big(\beta(t), \mathcal{F}_t, \mathbb{P}\big)$ be an $\mathbb{R}$-valued Brownian motion, show that, for each $T \in (0,\infty)$, $t \rightsquigarrow \operatorname{sgn}\big(\beta(t)\big)$ is Stratonovich integrable on $[0,T]$ with respect to β, and arrive at

$$\big|\beta(T)\big| = \int_0^T \operatorname{sgn}\big(\beta(t)\big) \circ d\beta(t).$$

After comparing this with the result in (6.1.7), conclude that the local time $\ell(T,0)$ of β at 0 satisfies

$$\int_0^T \frac{|\beta(t)|}{t}\,dt - \ell(T,0) - \big|\beta(T)\big| = \int_0^T \operatorname{sgn}\big(\beta(T-t)\big)\,d\check{M}^T(t),$$

where $\check{M}^T$ is the martingale part of $\check{\beta}^T$. In particular, the expression on the left hand side is a centered Gaussian random variable with variance T.

EXERCISE 8.1.9. Itô's formula proves that $f \circ Z \in \mathcal{S}(\mathbb{P};\mathbb{R})$ whenever $Z = (Z_1,\dots,Z_n) \in \mathcal{S}(\mathbb{P};\mathbb{R})^n$ and $f \in C^2(\mathbb{R}^n;\mathbb{R})$. On the other hand, as Tanaka's treatment (cf. §6.1) of local time makes clear, it is not always necessary to know that f has two continuous derivatives. Indeed, both (6.1.3) and (6.1.7) provide examples in which the composition of a martingale with a continuous function leads to a continuous semimartingale even though the derivative of the function is discontinuous. More generally, as a corollary of the Doob-Meyer Decomposition Theorem (alluded to at the beginning of § 7.1) one can show that $f \circ Z$ will be a continuous semimartingale whenever $Z \in \mathcal{M}_{\text{loc}}(\mathbb{P};\mathbb{R}^n)$ and f is a continuous, convex function. Here is a more pedestrian approach to this result.

(i) Following Tanaka's procedure, prove that $f \circ Z \in \mathcal{S}(\mathbb{P};\mathbb{R})$ whenever $Z \in \mathcal{S}(\mathbb{P};\mathbb{R})^n$ and $f \in C^1(\mathbb{R}^n;\mathbb{R})$ is convex. That is, if M_i denotes the martingale part of Z_i, show that

$$f\big(Z(t)\big) - \sum_{i=1}^n \int_0^t \partial_i f\big(Z(\tau)\big)\,dM_i(\tau)$$

is the bounded variation part of $f \circ Z$.

Hint: Begin by showing that when $A(t) \equiv \big(\big(\langle Z_i, Z_j\rangle(t)\big)\big)_{1\le i,j\le n}$, $A(t)-A(s)$ is $\mathbb{P}$-almost surely non-negative definite for all $0 \le s \le t$, and conclude that if $f \in C^2(\mathbb{R}^n;\mathbb{R})$ is convex then

$$t \rightsquigarrow \sum_{i,j=1}^{n} \int_0^t \partial_i\partial_j f\big(Z(\tau)\big)\langle Z_i, Z_j\rangle(d\tau)$$

is $\mathbb{P}$-almost surely non-decreasing.

(ii) By taking advantage of more refined properties of convex functions, see if you can prove that $f \circ Z \in \mathcal{S}(\mathbb{P};\mathbb{R})$ when f is a continuous, convex function.

EXERCISE 8.1.10. If one goes back to the original way in which we described Stratonovich in terms of Riemann integration, it becomes clear that the only reason why we needed Y to be a semimartingale is that we needed to know that

$$\sum_{m\le 2^N t} \big(\Delta_m^N Y\big)\big(\Delta_m^N X\big)$$

converges in $\mathbb{P}$-probability to a continuous function of locally bounded variation uniformly for t in compacts.

(i) Let $Z = \big(Z_1,\dots,Z_n\big) \in \mathcal{S}(\mathbb{P};\mathbb{R})^n$, and set $Y = f \circ Z$, where $f \in C^1(\mathbb{R}^n;\mathbb{R})$. Show that, for any $X \in \mathcal{S}(\mathbb{P};\mathbb{R})$,

$$\sum_{m\le 2^N t} \big(\Delta_m^N Y\big)\big(\Delta_m^N X\big) \longrightarrow \sum_{i=1}^{n} \int_0^t \partial_i f\big(Z(\tau)\big)\langle Z_i, X\rangle(d\tau)$$

in $\mathbb{P}$-probability uniformly for t compacts.

(ii) Continuing with the notation in **(i)**, show that

$$\begin{aligned}\int_0^t Y(\tau)\circ dX(\tau) &\equiv \lim_{N\to\infty}\int_0^t Y^N(\tau)\,dX^N(\tau)\\ &= \int_0^t Y(\tau)\,dX(\tau) + \frac{1}{2}\sum_{i=1}^{n}\int_0^t \partial_i f\big(Z(\tau)\big)\langle Z_i, X\rangle(d\tau),\end{aligned}$$

where the convergence is in $\mathbb{P}$-probability uniformly for t in compacts.

(iii) Show that (8.1.6) continues to hold for $f \in C^2(\mathbb{R}^n;\mathbb{R})$ when the integral on the right hand side is interpreted using the extension of Stratonovich integration developed in **(ii)**.

EXERCISE 8.1.11. Let $X, Y \in \mathcal{S}(\mathbb{P};\mathbb{R})$. In connection with Remark 8.1.7, it is interesting to examine whether it is sufficient to mollify only X when defining the Stratonovich integral of Y with respect to X.

(i) Show $\int_0^1 Y([\tau]_N)\, dX^N(\tau)$ tends in $\mathbb{P}$-probability to $\int_0^1 Y(\tau)\, dX(\tau)$.

(ii) Define $\psi_N(t) = 1 - 2^N(t - [t]_N)$, set $Z_N(t) = \int_0^t \psi_N(\tau)\, dY(\tau)$, and show that

$$\int_0^1 Y(\tau)\, dX^N(\tau) - \int_0^1 Y([\tau]_N)\, dX^N(\tau) = \sum_{m=0}^{2^N-1} \big(\Delta_m^N X\big)\big(\Delta_m^N Z_N\big).$$

(iii) Show that

$$\sum_{m=0}^{2^N-1} \big(\Delta_m^N X\big)\big(\Delta_m^N Z_N\big) - \int_0^1 \psi_N(t)\langle X, Y\rangle(dt) \longrightarrow 0 \quad \text{in } \mathbb{P}\text{-probability.}$$

(iv) Show that for any Lebesgue integrable function $\alpha : [0,1] \longrightarrow \mathbb{R}$, $\int_0^1 \psi_N(\tau)\alpha(\tau)\, d\tau$ tends to $\frac{1}{2}\int_0^1 \alpha(\tau)\, d\tau$.

(v) Under the condition that $\langle X, Y\rangle(dt) = \beta(t)\, dt$, where $\beta : [0,\infty) \times \Omega \longrightarrow [0,\infty)$ is a progressively measurable function, use the preceding to see that $\int_0^1 Y(\tau)\, dX^N(\tau)$ tends in $\mathbb{P}$-probability to $\int_0^1 Y(\tau) \circ dX(\tau)$.

EXERCISE 8.1.12. One place where Stratonovich's theory really comes into its own is when it comes to computing the determinant of the solution to a linear stochastic differential equation. Namely, suppose that $A = ((A_{ij}))_{1\le i,j\le n} \in \mathcal{S}\big(\mathbb{P}; \mathrm{Hom}(\mathbb{R}^n;\mathbb{R}^n)\big)$ (i.e., $A_{ij} \in \mathcal{S}(\mathbb{P};\mathbb{R})$ for each $1 \le i, j \le n$), and assume that $X \in \mathcal{S}\big(\mathbb{P}; \mathrm{Hom}(\mathbb{R}^n;\mathbb{R}^n)\big)$ satisfies $dX(t) = X(t) \circ dA(t)$ in sense that

$$dX_{ij}(t) = \sum_{k=1}^n X_{ik}(t) \circ dA_{kj}(t) \quad \text{for all } 1 \le i, j \le n.$$

Show that

$$\det\big(X(t)\big) = \det\big(X(0)\big)e^{\mathrm{Trace}\left(A(t)-A(0)\right)}.$$

8.2 Stratonovich Stochastic Differential Equations

Seeing as every Stratonovich integral can be converted into an Itô integral, it might seem unnecessary to develop a separate theory of Stratonovich stochastic differential equations. On the other hand, the replacement of a Stratonovich equation by the equivalent Itô equation removes the advantages, especially the coordinate invariance, of Stratonovich's theory. With this in mind, we devote the present section to a *non-Itô* analysis of Stratonovich stochastic differential equations.

Let $\mathbb{P}^0$ denote the standard Wiener measure for r-dimensional paths $p = (p_1, \dots, p_r) \in C([0,\infty);\mathbb{R}^r)$. In order to emphasize the coordinate invariance of the theory, we will write our Stratonovich stochastic differential equations in the form

$$\begin{gathered} dX(t,x,p) = V_0\big(X(t,x,p)\big)\,dt + \sum_{k=1}^{r} V_k\big(X(t,x,p)\big) \circ dp_k(t) \\ \text{with } X(0,x,p) = x, \end{gathered} \tag{8.2.1}$$

where, for each $0 \le k \le r$, $V_k : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a smooth function. To see what the equivalent Itô equation is, notice that

$$\begin{aligned} \Big\langle (V_k)_i\big(X(\,\cdot\,,x,p)\big), p_k\Big\rangle(dt) &= \sum_{j=1}^{n} \partial_j (V_k)_i\big(X(t,x,p)\big)\langle X_j(\,\cdot\,,x,p), p_k\rangle(dt) \\ &= \sum_{j=1}^{n}\sum_{\ell=1}^{r} (V_\ell)_j \partial_j (V_k)_i\big(X(t,x,p)\big)\langle p_\ell, p_k\rangle(dt) = D_{V_k}(V_k)_i\big(X(t,x,p)\big)\,dt, \end{aligned}$$

where $D_{V_k} = \sum_{i=1}^{n}(V_k)_j\partial_j$ denotes the directional derivative operator determined by V_k. Hence, if we think of each V_k, and therefore $X(t,x,p)$, as a column vector, then the Itô equivalent to (8.2.1) is

$$\begin{gathered} dX(t,x,p) = \sigma\big(X(t,x,p)\big)dp(t) + b\big(X(t,x,p)\big)\,dt \\ \text{with } X(0,x,p) = x, \end{gathered} \tag{8.2.2}$$

where $\sigma(x) = \big(V_1(x), \dots, V_r(x)\big)$ is the $n \times r$-matrix whose kth column is $V_k(x)$ and $b = V_0 + \frac{1}{2}\sum_1^r D_{V_k}V_k$. In particular, if

$$L = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\partial_i\partial_j + \sum_{i=1}^{n} b_i\partial_i, \quad \text{where } a = \sigma\sigma^\top,$$

then L is the operator associated with (8.2.1) in the sense that for any $f \in C^{1,2}\big([0,T]\times\mathbb{R}^n\big)$,

$$\left(f\big(t\wedge T, X(t,x,p)\big) - \int_0^{t\wedge T} \big(\partial_\tau f + Lf\big)\big(X(\tau,x,p)\big)\,d\tau, \mathcal{B}_t, \mathbb{P}^0 \right)$$

is a local martingale. However, a better way to write this operator is directly in terms of the directional derivative operators D_{V_k}. Namely,

$$L = D_{V_0} + \frac{1}{2}\sum_{k=1}^{r}(D_{V_k})^2. \tag{8.2.3}$$

Hörmander's famous paper [10] was the first to demonstrate the advantage of writing a second order elliptic operator in the form given in (8.2.3), and, for this reason, (8.2.3) is often said to be the *Hörmander form* expression for L. The most obvious advantage is the same as the advantage that Stratonovich's theory has over Itô's: it behaves well under change of variables. To wit, if $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a diffeomorphism, then

$$(L\varphi) \circ F = D_{F_* V_0}\varphi + \frac{1}{2}\sum_{k=1}^{r}\big(D_{F_* V_k}\big)^2\varphi,$$

where

$$\big(F_* V_k\big)_i = D_{V_k} F_i, \quad 1 \le i \le n,$$

is the "pushforward" under F of V_k. That is, $D_{F_* V_k}\varphi = D_{V_k}(\varphi \circ F)$.

8.2.1. Commuting Vector Fields. Until further notice, we will be dealing with vector fields $V_0, \dots, V_r$ which have two uniformly bounded, continuous derivatives. In particular, these assumptions are more than enough to assure that the equation (8.2.2), and therefore (8.2.1), admits a $\mathbb{P}^0$-almost surely unique solution. In addition, by Corollary 4.2.6 or Corollary 7.3.6, the $\mathbb{P}^0$-distribution of the solution is the unique solution $\mathbb{P}^L_x$ to the martingale for L starting from x.

In this section we will take the first in a sequence of steps leading to an alternative (especially, non-Itô) way of thinking about solutions to (8.2.1).

Given $\xi = (\xi_0, \dots, \xi_r) \in \mathbb{R}^{r+1}$, set $V_\xi = \sum_{k=0}^{r} \xi_k V_k$, and determine $E(\xi, x)$ for $x \in \mathbb{R}^n$ so that $E(0, x) = x$ and

$$\frac{d}{dt}E(t\xi, x) = V_\xi\big(E(t\xi, x)\big).$$

From elementary facts (cf. §4 of Chapter 2 in [4]) about ordinary differential equations, we know that $(\xi, x) \in \mathbb{R}^{r+1} \times \mathbb{R}^n \longmapsto E(\xi, x) \in \mathbb{R}^n$ is a twice continuously differentiable function which satisfies estimates of the form

$$\begin{aligned}&\big|E(\xi,x) - x\big| \le C|\xi|, \quad \big|\partial_{\xi_k}E(\xi,x)\big| \vee \big|\partial_{x_i}E(\xi,x)\big| \le Ce^{\nu|\xi|}\\ &\big|\partial_{\xi_k}\partial_{\xi_\ell}E(\xi,x)\big| \vee \big|\partial_{\xi_k}\partial_{x_i}E(\xi,x)\big| \vee \big|\partial_{x_i}\partial_{x_j}E(\xi,x)\big| \le Ce^{\nu|\xi|},\end{aligned} \tag{8.2.4}$$

for some $C < \infty$ and $\nu \in [0, \infty)$. Finally, define

$$V_k(\xi, x) = \partial_{\xi_k}E(\xi, x) \quad \text{for } 0 \le k \le r \text{ and } (\xi, x) \in \mathbb{R}^{r+1} \times \mathbb{R}^n. \tag{8.2.5}$$

For continuously differentiable, $\mathbb{R}^n$-valued functions V and W on $\mathbb{R}^n$, we will use the notation $[V, W]$ to denote the $\mathbb{R}^n$-valued function which is determined so that $D_{[V,W]}$ is equal to the commutator, $\big[D_V, D_W\big] = D_V \circ D_W - D_W \circ D_V$, of D_V and D_W. That is, $[V, W] = D_V W - D_W V$.

8.2.6 Lemma. *$V_k(\xi, x) = V_k\big(E(\xi, x)\big)$ for all $0 \le k \le r$ and $(\xi, x) \in \mathbb{R}^{r+1} \times \mathbb{R}^n$ if and only if $[V_k, V_\ell] \equiv 0$ for all $0 \le k, \ell \le r$.*

Proof: First assume that $V_k(\xi, x) = V_k\big(E(\xi, x)\big)$ for all k and (ξ, x). Then, if $\mathbf{e}_\ell$ is the element of $\mathbb{R}^{r+1}$ whose ℓth coordinate is 1 and whose other coordinates are 0, we see that

$$\frac{d}{dt} E(\xi + t\mathbf{e}_\ell, x) = V_\ell(\xi + t\mathbf{e}_\ell, x) = V_\ell\big(E(\xi + t\mathbf{e}_\ell, x)\big),$$

and so, by uniqueness, $E(\xi + t\mathbf{e}_\ell, x) = E\big(t\mathbf{e}_\ell, E(\xi, x)\big)$. In particular, this leads first to

$$E\big(t\mathbf{e}_\ell, E(s\mathbf{e}_k, x)\big) = E(t\mathbf{e}_\ell + s\mathbf{e}_k, x) = E\big(s\mathbf{e}_k, E(t\mathbf{e}_\ell, x)\big),$$

and thence to

$$\begin{aligned} D_{V_k} V_\ell(x) &= \frac{\partial^2}{\partial s \partial t} E\big(t\mathbf{e}_\ell, E(s\mathbf{e}_k, x)\big)\Big|_{s=t=0} \\ &= \frac{\partial^2}{\partial t \partial s} E\big(s\mathbf{e}_k, E(t\mathbf{e}_\ell, x)\big)\Big|_{s=t=0} = D_{V_\ell} V_k(x). \end{aligned}$$

To go the other direction, first observe that $V_k(\xi, x) = V_k\big(E(\xi, x)\big)$ is implied by $E(\xi + t\mathbf{e}_k, x) = E\big(t\mathbf{e}_k, E(\xi, x)\big)$. Thus, it suffices to show that $E(\xi + \eta, x) = E\big(\eta, E(\xi, x)\big)$ follows from $[V_\xi, V_\eta] \equiv 0$. Second, observe that $E(\xi + \eta, x) = E\big(\eta, E(\xi, x)\big)$ is implied by $E\big(\xi, E(\eta, x)\big) = E\big(\eta, E(\xi, x)\big)$. Indeed, if the second of these holds and $F(t) \equiv E\big(t\xi, E(t\eta, x)\big)$, then $\dot F(t) = (V_\xi + V_\eta)\big(F(t)\big)$, and so, by uniqueness, $F(t) = E\big(t(\xi + \eta), x\big)$. In other words, all that remains is to show that $E\big(\xi, E(\eta, x)\big) = E\big(\eta, E(\xi, x)\big)$ follows from $[V_\xi, V_\eta] \equiv 0$. To this end, set $F(t) = E\big(\xi, E(t\eta, x)\big)$, and note that $\dot F(t) = E(\xi, \cdot\,)_* V_\eta\big(E(t\eta, x)\big)$. Hence, by uniqueness, we will know that $E\big(\xi, E(t\eta, x)\big) = F(t) = E\big(t\eta, E(\xi, x)\big)$ once we show that $E(\xi, \cdot\,)_* V_\eta = V_\eta\big(E(\xi, \cdot\,)\big)$. But

$$\frac{d}{ds} E(s\xi, \cdot\,)_*^{-1} V_\eta\big(E(s\xi, \cdot\,)\big) = [V_\xi, V_\eta]\big(E(s\xi, \cdot\,)\big),$$

and so we are done. □

8.2.7 Theorem. *Assume that the vector fields V_k commute. Then the one and only solution to (8.2.1) is $(t, p) \rightsquigarrow X(t, x, p) \equiv E\Big(\big(t, p(t)\big), x\Big)$.*

Proof: Let $X(\,\cdot\,, x, p)$ be given as in the statement. By Itô's formula,

$$dX(t, x, p) = V_0\Big(\big(t, p(t)\big), X(t, x, p)\Big) dt + \sum_{k=1}^r V_k\Big(\big(t, p(t)\big), X(t, x, p)\Big) \circ dp_k(t),$$

and so, by Lemma 8.2.6, $X(\,\cdot\,, x, p)$ is a solution. As for uniqueness, simply re-write (8.2.1) in its Itô equivalent form, and conclude, from Theorem 5.2.2, that there can be at most one solution to (8.2.1). □

REMARK 8.2.8. A significant consequence of Theorem 8.2.7 is that the solution to (8.2.1) is a *smooth* function of p when the V_k's commute. In fact, for such vector fields, $p \rightsquigarrow X(\,\cdot\,, x, p)$ is the unique continuous extention to $C\big([0,\infty);\mathbb{R}^r\big)$ of the solution to the ordinary differential equation

$$\dot{X}(t,x,p) = V_0\big(X(t,x,p)\big) + \sum_{k=1}^{r} V_k\big(X(t,x,p)\big)\,\dot{p}_k(t)$$

for $p \in C^1\big([0,\infty);\mathbb{R}^r\big)$. This fact should be compared to the examples given in § 3.3.

8.2.2. General Vector Fields. Obviously, the preceding is simply not going to work when the vector fields do not commute. On the other, it indicates how to proceed. Namely, the commuting case plays in Stratonovich's theory the role that the constant coefficient case plays in Itô's . In other words, we should suspect that the commuting case is correct locally and that the general case should be handled by perturbation. We continue with the assumption that the V_k's are smooth and have two bounded continuous derivatives.

With the preceding in mind, for each $N \ge 0$, set $X^N(0,x,p) = x$ and

$$\begin{gathered} X^N(t,x,p) = E\big(\Delta^N(t,p), X^N([t]_N, x, p)\big) \\ \text{where } \Delta^N(t,p) \equiv \big(t - [t]_N, p(t) - p([t]_N)\big). \end{gathered} \tag{8.2.9}$$

Equivalently (cf. (8.2.5)):

$$\begin{aligned} dX^N(t,x,p) = V_0\big(\Delta^N(t,p), X^N([t]_N,x,p)\big)\,dt& \\ + \sum_{k=1}^{r} V_k\big(\Delta^N(t,p), X^N([t]_N,x,p)\big) \circ dp_k(t);& \end{aligned}$$

Note that, by Lemma 8.2.6, $X^N(t,x,p) = E\big((t,p(t)),x\big)$ for each $N \ge 0$ when the V_k's commute. In order to prove that $\{X^N(\,\cdot\,,x,p) : N \ge 0\}$ is $\mathbb{P}^0$-almost surely convergent even when the V_k's do not commute, we proceed as follows. Set $D^N(t,x,p) \equiv X(t,x,p) - X^N(t,x,p)$ and $W_k(\xi,x) \equiv V_k\big(E(\xi,x)\big) - V_k(\xi,x)$, and note that $D^N(0,x,p) = 0$ and

$$\begin{aligned} dD^N(t,x,p) =& \Big(V_0\big(X(t,x,p)\big) - V_0\big(X^N(t,x,p)\big)\Big)\,dt \\ &+ \sum_{k=1}^{r}\Big(V_k\big(X(t,x,p)\big) - V_k\big(X^N(t,x,p)\big)\Big) \circ dp_k(t) \\ &+ W_0\big(\Delta^N(t,p), X^N([t]_N,x,p)\big)\,dt \\ &+ \sum_{k=1}^{r} W_k\big(\Delta^N(t,p), X^N([t]_N,x,p)\big) \circ dp_k(t). \end{aligned} \tag{8.2.10}$$

8.2.11 LEMMA. *There exists a $C < \infty$ and $\nu > 0$ such that*

$$\left|V_k(\eta, E(\xi, x)) - V_k(\xi + \eta, x) - \tfrac{1}{2}\left[V_\xi, V_k\right](x)\right| \le C\left(|\xi| + |\eta|\right)^2 e^{\nu(|\xi|+|\nu|)}$$

for all $\xi, \eta \in \mathbb{R}^{r+1}$, $x \in \mathbb{R}^n$, and $0 \le k \le r$. In particular, if

$$\tilde{W}_k(\xi, x) \equiv W_k(\xi, x) - \sum_{\ell \ne k} \xi_\ell \left[V_\ell, V_k\right](x),$$

then $\tilde{W}_k(0, x) = 0$ and $\left|\partial_\xi \tilde{W}_k(\xi, x)\right| \le C|\xi| e^{\nu|\xi|}$ for all $(\xi, x) \in \mathbb{R}^{r+1} \times \mathbb{R}^n$ and some $C < \infty$ and $\nu > 0$.

PROOF: In view of the estimates in (8.2.4), it suffices for us to check the first statement. To this end, observe that

$$E\big(\eta, E(\xi, x)\big) = E(\xi, x) + V_\eta\big(E(\xi, x)\big) + \tfrac{1}{2} D_{V_\eta} V_\eta\big(E(\xi, x)\big) + R_1(\xi, \eta, x),$$

where $R_1(\xi, \eta, x)$ is the remainder term in the second order Taylor's expansion of $\eta \rightsquigarrow E\big(\,\cdot\,, E(\xi, x)\big)$ around 0 and is therefore (cf. (8.2.4)) dominated by constant $C_1 < \infty$ times $|\eta|^3 e^{\nu|\eta|}$. Similarly,

$$\begin{aligned}
&E(\xi, x) = x + V_\xi(x) + \tfrac{1}{2} D_{V_\xi} V_\xi(x) + R_2(\xi, x) \\
&V_\eta\big(E(\xi, x)\big) = V_\eta(x) + D_\xi V_\eta(x) + R_3(\xi, \eta, x) \\
&D_{V_\eta} V_\eta\big(E(\xi, x)\big) = D_{V_\eta} V_\eta(x) + R_4(\xi, \eta, x) \\
&E(\xi + \eta, x) = x + D_{V_{\xi+\eta}} V_{\xi+\eta}(x) + \tfrac{1}{2} D^2_{V_{\xi+\eta}} V_{\xi+\eta}(x) + R_2(\xi + \eta, x),
\end{aligned}$$

where

$$|R_2(\xi, x)| \le C_2 |\xi|^3 e^{\nu|\xi|}, \quad |R_3(\xi, \eta, x)| \le C_3 |\xi|^2 |\eta| e^{\nu|\xi|},$$
$$|R_4(\xi, \eta, x)| \le C_4 |\xi| |\eta|^2 e^{\nu|\xi|}.$$

Hence

$$E\big(\eta, E(\xi, x)\big) - E(\xi + \eta, x) = \tfrac{1}{2}\left[V_\xi, V_\eta\right](x) + R_5(\xi, \eta, x),$$
$$\text{where } |R_5(\xi, \eta, x)| \le C_5\big(|\xi| + |\eta|\big)^3 e^{\nu(|\xi|+|\eta|)},$$

and so the required estimate follows. □

Returning to (8.2.10), we now write

$$\begin{aligned}
dD^N(t, x, p) = {} & W_0\big(\Delta^N(t, p), X^N([t]_N, x, p)\big)\, dt \\
& + \frac{1}{2} \sum_{1 \le k \ne \ell \le r} \left[V_\ell, V_k\right]\big(X^N([t]_N, x, p)\big) \Delta^N_\ell(t, p)\, dp_k(t) \\
& + \sum_{k=1}^r \tilde{W}_k\big(\Delta^N(t, p), X^N([t]_N, x, p)\big) \circ dp_k(t) \\
& + \Big(V_0\big(X(t, x, p)\big) - V_0\big(X^N(t, x, p)\big)\Big)\, dt \\
& + \sum_{k=1}^r \Big(V_k\big(X(t, x, p)\big) - V_k\big(X^N(t, x, p)\big)\Big) \circ dp_k(t),
\end{aligned}$$

where, for $1 \le \ell \le r$, we have used $\Delta_\ell^N(t,p) = p_\ell(t) - p_\ell([t]_N)$ to denote the kth coordinate of $\Delta^N(t,p)$. Also, it is important to observe that, because $\langle p_k, p_\ell\rangle \equiv 0$ for $\ell \ne k$, we were able to replace the Stratonovich stochastic integral $\Delta_\ell^N(t,p) \circ dp_k(t)$ by the Itô stochastic integral $\Delta_\ell^N(t,p)\, dp_k(t)$. In fact, it is this replacement which makes what follows possible.

To proceed, first use an obvious Gaussian computation to see that there exists for each $q \in [1,\infty)$ a $A_q < \infty$ such that

$$\mathbb{E}^{\mathbb{P}^0}\Big[\big|\Delta^N(t,p)\big|^{2q} e^{2q\nu|\Delta^N(t,p)|}\Big]^{\frac{1}{q}} \le A_q\big(t - [t]_N\big).$$

Hence, there exist constants $C_q < \infty$ such that

$$\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot} W_0\big(\Delta^N(t,p), X^N([t]_N,x,p)\big)\,dt\right\|_{[0,T]}^{2q}\right]^{\frac{1}{q}} \le C_q T^2 2^{-N}$$

$$\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot} \big[V_k, V_\ell\big]\big(X^N([t]_N,x,p)\big)\Delta_\ell^N(t,p)\,dp_k(t)\right\|_{[0,T]}^{2q}\right]^{\frac{1}{q}} \le C_q T 2^{-N}$$

$$\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot} \tilde{W}_k\big(\Delta^N(t,p), X^N(t,x,p)\big)\circ dp_k(t)\right\|_{[0,T]}^{2q}\right]^{\frac{1}{q}} \le C_q\big(T + T^2\big)2^{-N}.$$

In the derivation of the last two of these, we have used Burkholder's Inequality (cf. Exercises 5.1.27 or 7.2.15). Of course, in the case of the last, we had to first convert the Stratonovich integral to its Itô equivalent and then had to also apply the final part of Lemma 8.2.11. Similarly, we obtain

$$\begin{aligned}\mathbb{E}^{\mathbb{P}^0}&\left[\left\|\int_0^{\cdot} \Big(V_k\big(X(t,x,p)\big) - V_k\big(X^N(t,x,p)\big)\Big)\circ dp_k(t)\right\|_{[0,T]}^{2q}\right]^{\frac{1}{q}} \\ &\le C_q(1+T)\int_0^T \mathbb{E}^{\mathbb{P}^0}\Big[\big\|X(\,\cdot\,,x,p) - X^N(\,\cdot\,,x,p)\big\|_{[0,T]}^{2q}\Big]^{\frac{1}{q}}\,dt.\end{aligned}$$

After combining the preceding with Gronwall's inequality, we arrive at the goal toward which we have been working.

8.2.12 THEOREM. *For each $T > 0$ and $q \in [1,\infty)$ there exists a $C(T,q) < \infty$ such that*

$$\mathbb{E}^{\mathbb{P}^0}\Big[\big\|X(\,\cdot\,,x,p) - X^N(\,\cdot\,,x,p)\big\|_{[0,t]}^{2q}\Big]^{\frac{1}{q}} \le C(T,q)t2^{-N} \quad \textit{if } t \in [0,T].$$

8.2.3. Another Interpretation. In order to exploit the result in Theorem 8.2.12, it is best to first derive a simple corollary of it, a corollary which contains an important result which was proved, in a somewhat different form and by entirely different methods, originally by Wong and Zakai [42]. Namely, when p is a locally absolutely continuous element of $C\big([0,\infty);\mathbb{R}^n\big)$, then it is easy to check that, for each $T>0$,

$$\lim_{N\to\infty}\big\|X^N(\,\cdot\,,x,p)-X(\,\cdot\,,x,p)\big\|_{[0,T]}=0,$$

where here $X(\,\cdot\,,x,p)$ is the unique locally absolutely continuous function such that

$$X(t,x,p)=x+\int_0^t\left(V_0\big(X(\tau,x,p)\big)+\sum_{k=1}^r V_k\big(X(\tau,x,p)\big)\dot p_k(\tau)\right)d\tau. \tag{8.2.13}$$

8.2.14 Theorem. *For each $N\ge 0$ and $p\in C\big([0,\infty);\mathbb{R}^r\big)$, let $p^N\in C\big([0,\infty);\mathbb{R}^r\big)$ be determined so that*

$$p^N(m2^{-N})=p(m2^{-N})\text{ and }p\restriction I_{m,N}\equiv[m2^{-N},(m+1)2^{-N}]\text{ is linear}$$

for each $m\ge 0$. Then, for all $T>0$ and $\epsilon>0$,

$$\lim_{N\to\infty}\mathbb{P}^0\Big(\big\|X(\,\cdot\,,x,p)-X(\,\cdot\,,x,p^N)\big\|_{[0,T]}\ge\epsilon\Big)=0$$

where $X(\,\cdot\,,x,p^N)$ is the unique, absolutely continuous solution to (8.2.13).

Proof: The key to this result is the observation is that $X^N(m2^{-N},x,p)=X(m2^{-N},x,p^N)$ for all $m\ge 0$. Indeed, for each $m\ge 0$, both the maps $t\in(m2^{-N},(m+1)2^{-N})\longmapsto X(t,x,p^N)$ and

$$\begin{aligned}t\in&\big(m2^{-N},(m+1)2^{-N}\big)\\ &\longrightarrow E\Big(t\big(2^{-N},p\big((m+1)2^{-N}\big)-p(m2^{-N})\big),X(m2^{-N},x,p)\Big)\in\mathbb{R}^n\end{aligned}$$

are solutions to the same ordinary differential equation. Thus, by induction on $m\ge 0$, the asserted equality follows from the standard uniqueness theory for the solutions to such equations.

In view of the preceding and the result in Theorem 8.2.12, it remains only to prove that

$$\begin{aligned}\lim_{N\to\infty}\mathbb{P}^0\bigg(\sup_{t\in[0,T]}&\big|X(t,x,p^N)-X([t]_N,x,p^N)\big|\\ &\vee\big|X^N(t,x,p)-X^N([t]_N,x,p)\big|\ge\epsilon\bigg)=0\end{aligned}$$

for all $T\in(0,\infty)$ and $\epsilon>0$. But

$$\begin{aligned}&\big|X(t,x,p^N)-X([t]_N,x,p^N)\big|\le C2^{-N}\vee\big|p([t]_N+2^{-N})-p([t]_N)\big|\quad\text{and}\\ &\big|X^N(t,x,p)-X^N([t]_N,x,p)\big|\le Ce^{\nu\|p\|_{[0,T]}}2^{-N}\vee\big|p(t)-p([t]_N)\big|,\end{aligned}$$

and so there is nothing more to do. $\square$

An important dividend of preceding is best expressed in terms of the *support* in $C\big([0,\infty);\mathbb{R}^n\big)$ of the solution $\mathbb{P}^L_x$ to the martingale problem for L (cf. (8.2.3)) starting from x. Namely, take (cf. Exercise 8.2.16 below for more information)

$$S(x;V_0,\dots,V_r)=\big\{X(\,\cdot\,,x,p):\,p\in C^1\big([0,\infty);\mathbb{R}^r\big)\big\}.$$

8.2.15 COROLLARY. *Let L be the operator described by (8.2.3), and (cf. Corollary 4.2.6) let $\mathbb{P}^L_x$ be the unique solution to the martingale problem for L starting at x. Then the support of $\mathbb{P}^L_x$ in $C\big([0,\infty);\mathbb{R}^n\big)$ is contained in the closure there of the set $S(x;V_0,\dots,V_r)$.*

PROOF: First observe that $\mathbb{P}^L_x$ is the $\mathbb{P}^0$-distribution of $p \rightsquigarrow X(\,\cdot\,,x,p)$. Thus, by Theorem 8.2.14, we know that

$$\mathbb{P}^L_x\big(\overline{S(x;V_0,\dots,V_r)}\big)\geq\lim_{N\to\infty}\mathbb{P}^0\big(\{p:\,X(\,\cdot\,,x,p^N)\in\overline{S(x;V_0,\dots,V_r)}\}\big).$$

But, for each $n\in\mathbb{N}$ and $p\in C\big([0,\infty);\mathbb{R}^r\big)$, it is easy to construct $\{p^N_\epsilon:\,\epsilon>0\}\subseteq C^\infty\big([0,\infty);\mathbb{R}^r\big)$ so that $p^N_\epsilon(0)=p(0)$ and

$$\lim_{\epsilon\searrow 0}\int_0^T\big|\dot{p}^N_\epsilon(t)-\dot{p}^N(t)\big|^2\,dt=0$$

and to show that $\lim_{\epsilon\searrow 0}\|X(\,\cdot\,,x,p^N_\epsilon)-X(\,\cdot\,,x,p^N)\|_{[0,T]}=0$ for all $T>0$. Hence, $X(\,\cdot\,,x,p^N)\in\overline{S(x;V_0,\dots,V_r)}$ for all $n\in\mathbb{N}$ and $p\in C\big([0,\infty);\mathbb{R}^n\big)$. □

8.2.4. Exercises.

EXERCISE 8.2.16. Except when $\operatorname{span}\big(\{V_1(x),\dots,V_r(x)\}\big)=\mathbb{R}^n$ for all $x\in\mathbb{R}^n$, in which case $\overline{S(x;V_0,\dots,V_r)}$ is the set of $p\in C\big([0,\infty);\mathbb{R}^n\big)$ with $p(0)=x$, it is a little difficult to get a feeling for what paths are and what paths are not contained in $\overline{S(x;V_0,\dots,V_r)}$. In this exercise we hope to give at least some insight into this question. For this purpose, it will be helpful to introduce the space $\mathcal{V}(V_0,\dots,V_r)$ of bounded $V\in C^1\big([0,\infty);\mathbb{R}^n\big)$ with the property that for all $\alpha\in\mathbb{R}$ and $x\in\mathbb{R}^n$ the integral curve of $V_0+\alpha V$ starting at x is an element of $\overline{S(x;V_0,\dots,V_r)}$. Obviously, $V_k\in\mathcal{V}(V_0,\dots,V_r)$ for each $1\leq k\leq r$.

(i) Given $V,W\in\mathcal{V}(V_0,\dots,V_r)$ and $(T,x)\in(0,\infty)\times\mathbb{R}^n$, determine $X\in C\big([0,\infty);\mathbb{R}^n\big)$ so that

$$X(t)=\begin{cases} x+\int_0^t V\big(X(\tau)\big)\,d\tau & \text{if } t\in[0,T]\\ X(T)+\int_T^t W\big(X(\tau)\big)\,d\tau & \text{if } t>T,\end{cases}$$

and show that $X\in\overline{S(x;V_0,\dots,V_r)}$.

(ii) Show that if $V, W \in \mathcal{V}(V_0, \dots, V_r)$ and $\varphi, \psi \in C^1_{\mathrm{b}}(\mathbb{R}^n; \mathbb{R})$, then $\varphi V + \psi W \in \mathcal{V}(V_0, \dots, V_r)$.
Hint: Define $X : \mathbb{R}^3 \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ so that $X\big((0,0,0), x\big) = x$ and

$$\frac{d}{dt} X\big(t(\alpha,\beta,\gamma), x\big) = (\alpha V_0 + \beta V + \gamma W)\big(X\big(t(\alpha,\beta,\gamma), x\big).$$

Given $N \in \mathbb{N}$, define $Y_N : [0,\infty) \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ so that $Y_N(t,x)$ equals

$$\begin{cases} X\big(t(1, 2\varphi(x), 0), x\big) & \text{for } 0 \le t \le 2^{-N-1} \\ X\big((t - 2^{-N-1})(1, 0, 2\psi(x)), Y_N(2^{-N-1}\big) & \text{for } 2^{-N-1} \le t \le 2^{-N} \\ Y_N\big(t - [t]_N, Y_M([t]_N, x)\big) & \text{for } t \ge 2^{-N}. \end{cases}$$

Show that $Y_N(\,\cdot\,, x) \longrightarrow Y(\,\cdot\,, x)$ in $C\big([0,\infty); \mathbb{R}^n\big)$, where $Y(\,\cdot\,, x)$ is the integral curve of $V_0 + \varphi V + \psi W$ starting at x.

(iii) If $V, W \in \mathcal{V}(V_0, \dots, V_r)$ have two bounded continuous derivatives, show that $[V, W] \in \mathcal{V}(V_0, \dots, V_r)$.
Hint: Use the notation in the preceding. For $N \in \mathbb{N}$, define $Y_N : [0,\infty) \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ so that $Y_N(t,x)$ is equal to

$$\begin{aligned} &X\big((t, 2^{\frac{N}{2}+2} t, 0), x\big) \\ &X\big((t - 2^{-N-2}, 0, 2^{\frac{N}{2}+2}(t - 2^{N-2}), Y_N(2^{-N-2}, x)\big) \\ &X\big((t - 2^{-N-1}, -2^{\frac{N}{2}+2}(t - 2^{-N-1}), 0), Y_N(2^{-N-1}, x)\big) \\ &X\big((t - 32^{-N-2}, 0, -2^{\frac{N}{2}+2}(t - 32^{N-2})), Y_N(32^{-N-2}, x)\big) \\ &Y_N\big(t - [t]_N, Y_N([t]_N, x)\big) \end{aligned}$$

according to whether $0 \le t \le 2^{-N-2}$, $2^{-N-2} \le t \le 2^{N-1}$, $2^{-N-1} \le t \le 32^{-N-2}$, $32^{-N-2} \le t \le 2^{-N}$, or $t \ge 2^{-N}$. Show that $Y_N(\,\cdot\,, x) \longrightarrow Y(\,\cdot\,, x)$ in $C\big([0,\infty); \mathbb{R}^n\big)$, where $Y(\,\cdot\,, x)$ is the integral curve of $V_0 + [V, W]$ starting at x.

(iv) Suppose that M is a closed submanifold of $\mathbb{R}^n$ and that, for each $x \in M$, $\{V_0(x), \dots, V_r(x)\}$ is a subset of the tangent space $T_x M$ to M at x. Show that, for each $x \in M$, $\overline{S(x; V_0, \dots, V_r)} \subseteq C\big([0,\infty); M\big)$ and therefore that $\mathbb{P}^L_x\big(\forall t \in [0,\infty)\ p(t) \in M\big) = 1$. Moreover, if $\{V_1, \dots, V_r\} \subseteq C^\infty(\mathbb{R}^n; \mathbb{R}^n)$ and $\mathrm{Lie}(V_1, \dots, V_r)$ is the smallest Lie algebra of vectors fields on $\mathbb{R}^n$ containing $\{V_1, \dots, V_r\}$, show that

$$\overline{S(x; V_0, \dots, V_r)} = \big\{p \in C\big([0,\infty); M\big) : p(0) = x\big\}$$

if $M \ni x$ is a submanifold of $\mathbb{R}^n$ with the property that

$$V_0(y) \in T_y M = \big\{V(y) : V \in \mathrm{Lie}(V_1, \dots, V_r)\big\} \quad \text{for all } y \in M.$$

8.3 The Support Theorem

Corollary 8.2.15 is the easier half the following result which characterizes the support of the measure $\mathbb{P}_x^L$. In its statement, and elsewhere, we use the notation (cf. Theorem 8.2.14) $\|q\|_{M,I} \equiv \|\dot{q}^M\|_{L^2(I;\mathbb{R}^r)}$ for $M \in \mathbb{N}$, $q \in C^1\big([0,\infty);\mathbb{R}^r\big)$), and closed intervals $I \subseteq [0,\infty)$.

8.3.1 THEOREM. *The support of $\mathbb{P}_x^L$ in $C([0,\infty);\mathbb{R}^n)$ is the closure there of (cf. Corollary 8.2.15) $S(x;V_0,\dots,V_r)$. In fact, if for each smooth $g \in C^1\big([0,\infty);\mathbb{R}^n\big)$ $X(\,\cdot\,,x,g)$ is the solution to (8.2.13) with $p = g$, then*

$$\lim_{M\to\infty}\lim_{\delta\searrow 0}\mathbb{P}^0\Big(\big\|X(\,\cdot\,,x,p) - X(\,\cdot\,,x,g)\big\|_{[0,T]} < \epsilon \;\Big|\; \|p-g\|_{M,[0,T]} \le \delta\Big) = 1 \tag{8.3.2}$$

for all $T \in (0,\infty)$ and $\epsilon > 0$.[3]

Notice that, because we already know $\operatorname{supp}(\mathbb{P}_x^L) \subseteq \overline{S(x;V_0,\dots,V_r)}$, the first assertion will follow as soon as we prove (8.3.2). Indeed, since

$$p(0) = 0 \implies \big\|p-g\big\|^2_{M,[0,T]} \le 2^{M+1}\sum_{m=1}^{[2^MT]+1}\big|p(m2^{-M}) - g(m2^{-M})\big|^2$$

and, for any $\ell \in \mathbb{Z}^+$, the $\mathbb{P}^0$-distribution of

$$p \in C\big([0,\infty);\mathbb{R}^n\big) \longmapsto \big(p(2^{-M}),\dots,p(\ell 2^{-M})\big) \in (\mathbb{R}^n)^\ell$$

has a smooth, strictly positive density, we know that

$$\mathbb{P}^0\big(\big\|p-g\big\|_{M,[0,T]} \le \delta\big) > 0 \text{ for all } \delta > 0. \tag{8.3.3}$$

Hence, (8.3.2) is much more than is needed to conclude that

$$\mathbb{P}^0\big(\big\|X(\,\cdot\,,x,p) - X(\,\cdot\,,x,g)\big\|_{[0,T]} < \epsilon\big) > 0$$

for all $\epsilon > 0$.

---

[3] This sort of statement was proved for the first time in [39]. However, the conditioning there was somewhat different. Namely, the condition was that $\max_{1\le i\le n}\|p_i - g_i\|_{[0,T]} < \delta$. Ikeda and Watanabe [16] follow the same basic strategy in their treatment, although they introduce an observation which not only greatly simplifies the most unpleasant part of the argument in [39] but also allows them to use the more natural condition $\|p-g\|_{[0,T]} < \delta$. The strategy of the proof followed below was worked out in [38]. Finally, if all that one cares about is the support characterization, [40] shows that one need not assume that the operator L can be expressed in Hörmander's form.

To begin the proof of (8.3.2), first observe that, by Theorem 8.2.12 and (8.3.3),

$$\begin{aligned}
&\mathbb{P}^0\Big(\big\|X(\,\cdot\,,x,p)-X(\,\cdot\,,x,g)\big\|_{[0,T]}>\epsilon\Big|\|p-g\|_{M,[0,T]}\le\delta\Big)\\
&\quad\le\varlimsup_{N\to\infty}\mathbb{P}^0\Big(\big\|X^N(\,\cdot\,,x,p)-X(\,\cdot\,,x,g)\big\|_{[0,T]}>\epsilon\Big|\|p-g\|_{M,[0,T]}\le\delta\Big)\\
&\quad\le\sup_{N>M}\mathbb{P}^0\Big(\big\|X^N(\,\cdot\,,x,p)-X^M(\,\cdot\,,x,g)\big\|_{[0,T]}>\tfrac{\epsilon}{2}\Big|\|p-g\|_{M,[0,T]}\le\delta\Big)\\
&\qquad\quad+\mathbb{P}^0\Big(\big\|X^M(\,\cdot\,,x,p)-X(\,\cdot\,,x,g)\big\|_{[0,T]}>\tfrac{\epsilon}{2}\Big|\|p-g\|_{M,[0,T]}\le\delta\Big).
\end{aligned}$$

Hence, since

$$\begin{aligned}
&\big\|X^M(\,\cdot\,,x,p)-X(\,\cdot\,,x,g)\big\|_{[0,T]}\\
&\quad\le\big\|X^M(\,\cdot\,,x,p)-X(\,\cdot\,,x,p^M)\big\|_{[0,T]}+\big\|X(\,\cdot\,,x,p^M)-X(\,\cdot\,,x,g^M)\big\|_{[0,T]}\\
&\qquad+\big\|X(\,\cdot\,,x,g^M)-X(\,\cdot\,,x,g)\big\|_{[0,T]}\\
&\quad\le C\Bigg(2^{-M}+e^{\nu\|p\|_{[0,T]}}\sup_{\substack{0\le s<t\le T\\ t-s\le 2^{-M}}}\big|p(t)-p(s)\big|\\
&\qquad\qquad+\|p-g\|_{M,[0,T]}+2^{-\frac{M}{2}}\|g\|_{M,[0,T]}\Bigg),
\end{aligned}$$

it follows that we need only show that, for each $\epsilon>0$,

$$\lim_{M\to\infty}\sup_{\delta\in(0,1]}\mathbb{P}^0\left(\sup_{\substack{0\le s<t\le T\\ t-s\le 2^{-M}}}\big|p(t)-p(s)\big|\ge\epsilon\,\middle|\,\|p-g\|_{M,[0,T]}\le\delta\right)=0\tag{8.3.4}$$

and that

$$\begin{aligned}
\lim_{M\to\infty}\sup_{\substack{N>M\\ \delta\in(0,1]}}\mathbb{P}^0\Big(&\big\|X^N(\,\cdot\,,x,p)-X^M(\,\cdot\,,x,g)\big\|_{[0,T]}\ge\epsilon\\
&\Big|\|p-g\|_{M,[0,T]}\le\delta\Big)=0.
\end{aligned}\tag{8.3.5}$$

8.3.1. The Support Theorem, Part I. The key to our proof of both (8.3.4) and (8.3.5) is the following lemma about Wiener measure. As the astute reader will recognize, this lemma is a manifestation of the same property of Gaussian measures on which (6.2.5) is based.

8.3.6 Lemma. *For $M\in\mathbb{N}$ and $p\in C\big([0,\infty);\mathbb{R}^n\big)$, set $\tilde{p}^M=p-p^M$. Then $\sigma\big(\{\tilde{p}^M(t):\ t\ge 0\}\big)$ is $\mathbb{P}^0$-independent of $\mathcal{B}^M\equiv\sigma\big(\{p(m2^{-M}):\ m\in\mathbb{N}\}\big)$. Hence, if, for some Borel measurable $F:C\big([0,\infty);\mathbb{R}^n\big)\longrightarrow[0,\infty)$,*

$$\tilde{F}^M(q)\equiv\int F\big(\tilde{p}^M+q^M\big)\,\mathbb{P}^0(dp),\quad q\in C\big([0,\infty);\mathbb{R}^r\big),$$

then $\tilde F^M = \mathbb{E}^{\mathbb{P}^0}[F|\mathcal{B}^M]$ $\mathbb{P}^0$*-almost surely. In particular,*

$$\mathbb{E}^{\mathbb{P}^0}\big[F(p)\big|\,\|p-h\|_{M,[0,T]}\le\delta\big] = \frac{\mathbb{E}^{\mathbb{P}^0}\big[\tilde F^M(p),\ \|p-h\|_{M,[0,T]}\le\delta)\big]}{P^0(\|p-h\|_{M,[0,T]}\le\delta)}.$$

PROOF: Clearly, the second part of this lemma is an immediate consequence of the first part. Thus (cf. part **(i)** in Exercise 4.2.39 in [36]) because all elements of $\mathrm{span}\big(\{p(t) : t\ge 0\}\big)$ are centered Gaussian random variables, it suffices to prove that $\mathbb{E}^{\mathbb{P}^0}\big[\tilde p^M(s)p(m2^{-M})\big]=0$ for all $s\in[0,\infty)$ and $m\in\mathbb{N}$. Equivalently, what we need to show is that

$$\mathbb{E}^{\mathbb{P}^0}\big[p(s)p^M(m2^{-M})\big] = \mathbb{E}^{\mathbb{P}^0}\big[p^M(s)p^M(m2^{-M})\big] \text{ for all } s\in[0,\infty) \ \&\ m\in\mathbb{N}.$$

But, after choosing $\ell\in\mathbb{N}$ so that $\ell 2^{-M}\le s\le(\ell+1)2^{-M}$ and using the fact that $\mathbb{E}^{\mathbb{P}^0}\big[p(u)p(v)\big]=u\wedge v$, this becomes an elementary computation. □

Knowing Lemma 8.3.6, one gets (8.3.4) is easily. Namely, given $M\in\mathbb{N}$ and $(p,h)\in C\big([0,\infty);\mathbb{R}^r\big)$, set

$$\tilde p_h^M(t)\equiv\tilde p^M(t)+h^M(t).$$

Then, we will have proved (8.3.4) once we show that

$$\sup_{\|h-g\|_{M,[0,T]}\le 1}\mathbb{E}^{\mathbb{P}^0}\left[\sup_{0\le s<t\le T}\frac{|\tilde p_h^M(t)-\tilde p_h^M(s)|}{(t-s)^{\frac14}}\right]<\infty.$$

But, $|h^M(t)-h^M(s)|\le|t-s|^{\frac12}\|h\|_{M,[0,T]}$,

$$\sup_{0\le s<t\le T}\frac{|\tilde p^M(t)-\tilde p^M(s)|}{(t-s)^{\frac14}}\le 4\sup_{0\le s<t\le T+1}\frac{|p(t)-p(s)|}{(t-s)^{\frac14}},$$

and so the required estimate follows immediately from part **(ii)** of Exercise 2.4.16.

8.3.2. The Support Theorem, Part II. Unfortunately, the proof of (8.3.5) requires much more effort. To get started, fix $x\in\mathbb{R}^n$ and set $D^{M,N}(t,x,p)=X^N(t,x,p)-X^M(t,x,p)$ for $N>M$. Then, in view of Lemma 8.3.6, (8.3.5) will follow once we show that

$$\lim_{M\to\infty}\sup_{\substack{N>M\\ \|h-g\|_{M,[0,T]}\le 1}}\mathbb{E}^{\mathbb{P}^0}\Big[\big\|D^{M,N}(\,\cdot\,,x,\tilde p_h^M)\big\|_{[0,T]}^2\Big]=0 \tag{8.3.7}$$

Because

$$\begin{aligned}X^N(t,x,p)&=X^N\big(t-[t]_M,X^N([t]_M,x,p),\delta_{[t]_M}p\big)\\ \text{and } X^M(t,x,p)&=X^M\big(t-[t]_M,X^M([t]_M,x,p),\delta_{[t]_M}p\big),\end{aligned}$$

where, for each $s \geq 0$, $\delta_s : C\big([0,\infty);\mathbb{R}^r\big) \longrightarrow C\big([0,\infty);\mathbb{R}^r\big)$ is defined so that $\delta_s p(t) = p(s+t) - p(s)$, we have that $D^{M,N}(t,x,p)$ can be decomposed into

$$\begin{aligned}
&D^{M,N}\big(t-[t]_M, X^N([t]_M,x,p), \delta_{[t]_M}p\big)\\
&\quad + E\big(\Delta^M(t,p), X^N([t]_M,x,p)\big) - E\big(\Delta^M(t,p), X^M([t]_M,x,p)\big)\\
&= D^{M,N}([t]_M,x,p) + D^{M,N}\big(t-[t]_M, X^N([t]_M,x,p), \delta_{[t]_M}p\big)\\
&\quad + \bar{E}\big(\Delta^M(t,p), X^N([t]_M,x,p)\big) - \bar{E}\big(\Delta^M(t,p), X^M([t]_M,x,p)\big),
\end{aligned}$$

where $\bar{E}(\xi,x) \equiv E(\xi,x) - x$. Hence, we can write

$$\begin{aligned}
D^{M,N}(t,x,p) = \int_0^t W_0^{M,N}(\tau,x,p)\,d\tau + \sum_{k=1}^r \int_0^t W_k^{M,N}(\tau,x,p) \circ dp_k(\tau)\\
+ \tilde{D}^{M,N}(t,x,p),
\end{aligned} \tag{8.3.8}$$

where $W_k^{M,N}(t,x,p)$ is used to denote

$$V_k\big(\Delta^M(t,p), X^N([t]_M,x,p)\big) - V_k\big(\Delta^M(t,p), X^M([t]_M,x,p)\big)$$

and $\tilde{D}^{M,N}(t,x,p)$ is given by

$$\sum_{m=0}^{[2^M t]} D^{M,N}\big((t-m2^{-M}) \wedge 2^{-M}, X^N(m2^{-M},x,p), \delta_{m2^{-M}}p\big).$$

The terms involving the $W_k^{M,N}$'s are relatively easy to handle. Namely, because (cf. (5.1.25) and use Brownian scaling) for any $\lambda \geq 0$ and $T \in (0,\infty)$

$$\mathbb{E}^{\mathbb{P}^0}\Big[e^{\lambda T^{-\frac{1}{2}}\|p\|_{[0,T]}}\Big] = \mathbb{E}^{\mathbb{P}^0}\Big[e^{\lambda\|p\|_{[0,1]}}\Big] \equiv K(\lambda) < \infty, \tag{8.3.9}$$

and, for any $(m,M) \in \mathbb{N}^2$, (cf. the notation in Theorem 8.2.14)

$$\big\|\Delta^M(\,\cdot\,,\tilde{p}_h^M)\|_{I_{m,M}} \leq 2\big\|\Delta^M(\,\cdot\,,p)\big\|_{I_{m,M}} + 2^{-\frac{M}{2}}\|h\|_{M,I_{m,M}},$$

we have that

$$\mathbb{E}^{\mathbb{P}^0}\Big[e^{\lambda 2^{\frac{M}{2}}\|\Delta^M(\,\cdot\,,\tilde{p}_h^M)\|_{I_{m,M}}}\,\Big|\,\mathcal{B}_{m2^{-M}}\Big] \leq K(2\lambda)e^{\lambda\|h\|_{M,I_{m,M}}}. \tag{8.3.10}$$

Hence, since (cf. (8.2.4))

$$\begin{aligned}
\Big\|\int_0^{\cdot} W_0^{M,N}(\tau,x,p)\,d\tau\Big\|_{[0,T]}^2 &\leq T\int_0^T \big|W_0^{M,N}(\tau,x,p)\big|^2\,d\tau\\
&\leq CT\int_0^T e^{2\nu|\Delta^M(\tau,p)|}\big|D^{M,N}([\tau]_M,x,p)\big|^2\,d\tau,
\end{aligned}$$

it is clear that there are finite constants K and K' such that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot} W_0^{M,N}(\tau,x,\tilde{p}_h^M)\,d\tau\right\|_{[0,T]}^2\right] \\
&\qquad \le KT\int_0^T \mathbb{E}^{\mathbb{P}^0}\left[e^{2\nu|\Delta^M(\tau,\tilde{p}_h^M)|}\big|D^{M,N}([\tau]_M,x,\tilde{p}_h^M)\big|^2\right]d\tau \\
&\qquad \le K'Te^{2\nu\|h\|_{M,[0,T]}}\int_0^T \mathbb{E}^{\mathbb{P}^0}\left[\big|D^{M,N}([\tau]_M,x,\tilde{p}_h^M)\big|^2\right]d\tau.
\end{aligned}$$

Next, suppose that $1 \le k \le r$. Then, after converting Stratonovich integrals to their Itô equivalents, we have

$$\begin{aligned}
&\int_0^t W_k^{M,N}(\tau,x,\tilde{p}_h^M)\circ d(\tilde{p}_h^M)_k(\tau) \\
&\quad = \int_0^t W_k^{M,N}(\tau,x,\tilde{p}_h^M)\,dp_k(\tau) + \frac{1}{2}\int_0^t W_{k,k}^{M,N}(\tau,x,\tilde{p}_h^M)\,d\tau \\
&\qquad - \int_0^t W_k^{M,N}(\tau,x,\tilde{p}_h^M)\dot{p}_k^M(\tau)\,d\tau + \int_0^t W_k^{M,N}(\tau,x,\tilde{p}_h^M)\dot{h}_k^M(\tau)\,d\tau,
\end{aligned}$$

where $W_{k,k}^{M,N}(t,x,p)$ denotes

$$V_{k,k}\big(\Delta^M(t,p),X^N([t]_M,x,p)\big) - V_{k,k}\big(\Delta^M(t,p),X^M([t]_M,x,p)\big)$$

when $V_{k,k}(\xi,x) \equiv \big[\partial_k V_k(\,\cdot\,,x)\big](\xi)$. The second and fourth terms on the right are handled in precisely the same way as the one involving $W_0^{M,N}$ and satisfy estimates of the form

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot} W_{k,k}^{M,N}(\tau,x,\tilde{p}_h^M)\,d\tau\right\|_{[0,T]}^2\right] \\
&\qquad \le KTe^{2\nu\|h\|_{M,[0,T]}}\int_0^T \mathbb{E}^{\mathbb{P}^0}\left[\big|D^{M,N}([\tau]_M,x,\tilde{p}_h^M)\big|^2\right]d\tau
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot} W_k^{M,N}(\tau,x,\tilde{p}_h^M)\dot{h}_k^M(\tau)\,d\tau\right\|_{[0,T]}^2\right] \\
&\qquad \le K\|h\|_{M,[0,T]}^2 e^{2\nu\|h\|_{M,[0,T]}}\int_0^T \mathbb{E}^{\mathbb{P}^0}\left[\big|D^{M,N}([\tau]_M,x,\tilde{p}_h^M)\big|^2\right]d\tau.
\end{aligned}$$

In addition, because the first term is an Itô integral, we can apply Doob's inequality followed by the above reasoning to obtain the estimate

$$\begin{aligned}\mathbb{E}^{\mathbb{P}^0}&\left[\left\|\int_0^{\cdot} W_k^{M,N}(\tau,x,\tilde{p}_h^M)\,dp_k(\tau)\right\|_{[0,T]}^2\right]\\&\le Ke^{2\nu\|h\|_{M,[0,T]}}\int_0^T \mathbb{E}^{\mathbb{P}^0}\left[\left|D^{M,N}([\tau]_M,x,\tilde{p}_h^M)\right|^2\right]d\tau.\end{aligned}$$

As for the third term on the right, it is best to first decompose it into

$$\begin{aligned}&\int_0^t W_k^{M,N}([\tau]_M,x,\tilde{p}_h^M)\,dp_k(\tau)\\&\qquad+\int_0^t\big(W_k^{M,N}(\tau,x,p)-W_k^{M,N}([\tau]_M,x,p)\big)\dot{p}_k^M(\tau)\,d\tau.\end{aligned}$$

Since the first of these is an Itô integral, it presents no new problem. To deal with the second, observe that (cf. (8.2.4))

$$\begin{aligned}&\left\|\int_0^{\cdot}\big(W_k^{M,N}(\tau,x,p)-W_k^{M,N}([\tau]_M,x,p)\big)\dot{p}_k^M(\tau)\,d\tau\right\|_{[0,T]}^2\\&\le C^2 4^M T\sum_{m=0}^{[2^MT]}\big(p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\big)^2\\&\quad\times\int_{I_{m,M}} e^{2\nu|\Delta^M(\tau,\tilde{p}_h^M)|}|\Delta^M(\tau,\tilde{p}_h^M)|^2\,d\tau\big|D^{M,N}(m2^{-M},x,\tilde{p}_h^M)\big|^2,\end{aligned}$$

and pass from this to the same sort of estimate at which we arrived for the other terms. Hence, by combining these, we can now replace (8.3.8) by

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}^0}\left[\big\|D^{M,N}(\,\cdot\,,x,\tilde{p}_h^M)\big\|_{[0,T]}^2\right]\\&\le C\big(T,\|h\|_{M,[0,T]}\big)\int_0^T\mathbb{E}^{\mathbb{P}^0}\left[\big\|D^{M,N}(\,\cdot\,,x,\tilde{p}_h^M)\big\|_{[0,t]}^2\right]dt\\&\quad+2\mathbb{E}^{\mathbb{P}^0}\big[\|\tilde{D}^{M,N}(\,\cdot\,,x,\tilde{p}_h^M)\|_{[0,T]}^2\big],\end{aligned}\tag{8.3.11}$$

where $(s,t)\in[0,\infty)\longmapsto C(s,t)\in(0,\infty)$ is non-decreasing in each variable separately.

8.3.3. The Support Theorem, Part III. After applying Gronwall's inequality to (8.3.11), we know that, there is a $(s,t)\in[0,\infty)^2\longmapsto K(s,t)\in(0,\infty)$, which is non-decreasing in each variable separately, such that

$$\begin{aligned}&\sup_{\|h-g\|_{M,[0,T]}\le 1}\mathbb{E}^{\mathbb{P}^0}\left[\big\|D^{M,N}(\,\cdot\,,x,\tilde{p}_h^M)\big\|_{[0,T]}^2\right]\\&\qquad\le K(T,g)\sup_{\|h-g\|_{M,[0,T]}\le 1}\mathbb{E}^{\mathbb{P}^0}\left[\big\|\tilde{D}^{M,N}(\,\cdot\,,x,\tilde{p}_h^M)\big\|_{[0,T]}^2\right].\end{aligned}$$

Hence, what remains is to show that

$$\lim_{M\to\infty} \sup_{\substack{N>M \\ \|h-g\|_{M,[0,T]}\le 1}} \mathbb{E}^{\mathbb{P}^0}\left[\big\|\tilde{D}^{M,N}(\,\cdot\,,x,\tilde{p}_h^M)\big\|_{[0,T]}^2\right] = 0, \tag{8.3.12}$$

and this turns out to be quite delicate.

To get started on the proof of (8.3.12), we again break the computation into parts. Namely, because

$$X^N\big(t-[\tau]_M, X^N([\tau]_M,x,p),\delta_{[\tau]_M}p\big) = X^N\big(t-[\tau]_N, X^N([\tau]_N,x,p),\delta_{[\tau]_N}p\big),$$

$\tilde{D}^{M,N}\big(t,x,\tilde{p}_h^M\big)$ can be written as

$$\begin{aligned}&\int_0^t \tilde{W}_0^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\,d\tau \\ &\quad + \sum_{k=1}^r \int_0^t \tilde{W}_k^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\,dp_k(\tau) + \frac{1}{2}\sum_{k=1}^r \int_0^t \tilde{W}_{k,k}^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\,d\tau \\ &\quad - \sum_{k=1}^r \int_0^t \tilde{W}_k^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\dot{p}_k^M(\tau)\,d\tau + \sum_{k=1}^r \int_0^t \tilde{W}_k^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\dot{h}_k^M(\tau),\end{aligned}$$

where $\tilde{W}_k^{M,N}(t,x,p)$ denotes

$$V_k\big(\Delta^N(t,p),X^N([t]_N,x,p)\big) - V_k\big(\Delta^M(t,p),X^N([t]_M,x,p)\big)$$

and $\tilde{W}_{k,k}^{M,N}(t,x,p)$ is equal to

$$V_{k,k}\big(\Delta^N(t,p),X^N([t]_N,x,p)\big) - V_{k,k}\big(\Delta^M(t,p),X^N([t]_M,x,p)\big).$$

With the exception of those involving $\dot{p}_k^M(\tau)$, none of these terms is very difficult to estimate. Indeed,

$$\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot} \tilde{W}_0^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\,d\tau\right\|_{[0,T]}^2\right] \le T\int_0^T \mathbb{E}^{\mathbb{P}^0}\left[\big|\tilde{W}_0^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\big|^2\right]d\tau,$$

$$\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot} \tilde{W}_k^{M,N}\big(\tau,x,\tilde{p}_h^M\big)dp_k(\tau)\right\|_{[0,T]}^2\right] \le 4\int_0^T \mathbb{E}^{\mathbb{P}^0}\left[\big|\tilde{W}_k^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\big|^2\right]d\tau,$$

$$\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot} \tilde{W}_{k,k}^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\,d\tau\right\|_{[0,T]}^2\right] \le T\int_0^T \mathbb{E}^{\mathbb{P}^0}\left[\big|\tilde{W}_{k,k}^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\big|^2\right]d\tau,$$

and

$$\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot} \tilde{W}_k^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\dot{h}_k^M(\tau)\right\|_{[0,T]}^2\right] \le \|h\|_{M,[0,T]}^2 \int_0^T \mathbb{E}^{\mathbb{P}^0}\left[\big|\tilde{W}_k^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\big|^2\right] d\tau.$$

At the same time, by (8.2.4), $\tilde{W}_k^{M,N}(t,x,\tilde{p}_h^M)$ is dominated by a constant times

$$\Big(\big|\tilde{p}_h^M([\tau]_N)-\tilde{p}_h^M([\tau]_M)\big| + \big|X^N\big([\tau]_N,x,\tilde{p}_h^M\big)-X^N\big([\tau]_M,x,\tilde{p}_h^M\big)\big|\Big)e^{\nu\|\Delta^M(\,\cdot\,,\tilde{p}_h^M)\|_{I_{m,M}}},$$

and so, in view of (8.3.9) and (8.3.10), all the above terms will have been handled once we check that, for each $q\in[1,\infty)$, there exists a $C_q<\infty$ such that

$$\mathbb{E}^{\mathbb{P}^0}\Big[\big\|X^N(\,\cdot\,,x,\tilde{p}_h^M)-X^N(m2^{-M},x,\tilde{p}_h^M)\big\|_{I_{m,M}}^q\Big]^{\frac{1}{q}} \le C_q\big(1+\|h\|_{M,I_{m,M}}\big)2^{-\frac{M}{2}} \tag{8.3.13}$$

for all $x\in\mathbb{R}^n$, $(m,M)\in\mathbb{N}^2$, and $N\ge M$. To this end, first note that, because $X^N(\,\cdot\,+m2^{-M},x,p)=X^N\big(\,\cdot\,,X^N(m2^{-M},x,p),\delta_{m2^{-M}}p\big)$, it suffices to treat the case when $m=0$. Second, write $X^N(t,x,\tilde{p}_h^M)-x$ as

$$\begin{aligned}&\int_0^t V_0\big(\Delta^N(\tau,\tilde{p}_h^M),x,\tilde{p}_h^M\big)\,d\tau\\ &+\sum_{k=1}^r\left(\int_0^t V_k\big(\Delta^N(\tau,\tilde{p}_h^M),x,\tilde{p}_h^M\big)\,dp_k(\tau)+\frac{1}{2}\int_0^t V_{k,k}\big(\Delta^N(\tau,\tilde{p}_h^M),x,\tilde{p}_h^M\big)\,d\tau\right)\\ &-\sum_{k=1}^r\int_0^t\Big(V_k\big(\Delta^N(\tau,\tilde{p}_h^M),x,\tilde{p}_h^M\big)\dot{p}_k^M(\tau)-V_k\big(\Delta^N(\tau,\tilde{p}_h^M),x,\tilde{p}_h^M\big)\dot{h}_k^M(\tau)\Big)\,d\tau,\end{aligned}$$

and apply (8.2.4) and standard estimates to conclude from this that

$$\mathbb{E}^{\mathbb{P}^0}\big[\big\|X^N(\,\cdot\,,x,\tilde{p}_h^M)-x\big\|_{I_{0,M}}^q\big]^{\frac{1}{q}}$$

is dominated by an expression of the form on the right hand side of (8.3.13).

8.3.4. The Support Theorem, Part IV. The considerations in § 8.3.3 reduce the proof of (8.3.12) to showing that, for each $T\in[0,\infty)$,

$$\lim_{M\to\infty}\sup_{\substack{N>M\\ \|h-g\|_{M,[0,T]}\le 1}}\mathbb{E}^{\mathbb{P}^0}\left[\left\|\int_0^{\cdot}\tilde{W}_k^{M,N}\big(\tau,x,\tilde{p}_h^M\big)\dot{p}_k^M(\tau)\,d\tau\right\|_{0,T]}^2\right]=0, \tag{8.3.14}$$

and for this purpose it is best to begin with yet another small reduction. Namely, dominate $\left\|\int_0^{\cdot} \tilde{W}_k^{M,N}\big(\tau, x, \tilde{p}_h^M\big)\dot{p}_k^M(\tau)\,d\tau\right\|_{[0,T]}$ by

$$\max_{0\le m\le 2^M T}\int_{I_{m,M}}\big|\tilde{W}_k^{M,N}\big(\tau, x, \tilde{p}_h^M\big)\big|\big|\dot{p}_k^M(\tau)\big|\,d\tau \\ + \max_{0\le m\le 2^M T}\left|\int_0^{m2^{-M}} \tilde{W}_k^{M,N}\big(\tau, x, \tilde{p}_h^M\big)\dot{p}_k^M(\tau)\,d\tau\right|.$$

Again, the first of these is easy, since, by lines of reasoning which should be familiar by now:

$$\begin{aligned} &\mathbb{E}^{\mathbb{P}^0}\left[\max_{0\le m\le [2^M T]}\left(\int_{I_{m,M}}\big|\tilde{W}_k^{M,N}\big(\tau, x, \tilde{p}_h^M\big)\big|\big|\dot{p}_k^M(\tau)\big|\,d\tau\right)^2\right] \\ &\le C\mathbb{E}^{\mathbb{P}^0}\left[\sum_{m=0}^{[2^M T]-1} e^{4\nu\|\Delta^M \tilde{p}_h^M\|_{I_{m,M}}}\left|p_k\big((m+1)2^{-M}\big) - p_k(m2^{-M})\right|^4\right]^{\frac{1}{2}} \\ &\le C' T^{\frac{1}{2}} e^{2^{1-\frac{M}{2}}\nu\|h\|^2_{M,[0,T]}} 2^{-\frac{M}{2}} \end{aligned}$$

for appropriate, finite constants C and C'. Hence, the proof of (8.3.14), and therefore (8.3.12), is reduced to showing that

$$\lim_{M\to\infty}\sup_{N>M}\mathbb{E}^{\mathbb{P}^0}\left[\max_{m\le 2^M T}\left|\int_0^{m2^{-M}} \tilde{W}_k^{M,N}\big(\tau, x, \tilde{p}_h^M\big)\dot{p}_k^M(\tau)\,d\tau\right|^2\right] = 0 \tag{8.3.15}$$

uniformly with respect to h with $\|h-g\|_{M,[0,T]} \le 1$.

To carry out the proof of (8.3.15), we decompose $\tilde{W}_k^{M,N}(\tau, x, p)$ into the sum

$$\sum_{L=M+1}^{N}\Big(\hat{V}_k^{M,L}\big(\tau, X^N([\tau]_M, x, p), p\big) + \hat{W}_k^{M,L}\big(\tau, X^N([\tau]_M, x, p), p\big)\Big),$$

where

$$\begin{aligned} \hat{V}_k^{M,L}(\tau, x, p) \equiv{}& V_k\big(\Delta^L(\tau, p), X^L([\tau]_L - [\tau]_M, x, \delta_{[\tau]_M} p)\big) \\ &- V_k\big(\Delta^L(\tau, p), X^{L-1}([\tau]_L - [\tau]_M, x, \delta_{[\tau]_M} p)\big). \end{aligned}$$

and

$$\begin{aligned} \hat{W}_k^{M,L}(\tau, x, p) \equiv{}& V_k\big(\Delta^L(\tau, p), X^{L-1}([\tau]_L - [\tau]_M, x, \delta_{[\tau]_M} p)\big) \\ &- V_k\big(\Delta^{L-1}(\tau, p), X^{L-1}([\tau]_{L-1} - [\tau]_M, x, \delta_{[\tau]_M} p)\big) \end{aligned}$$

After making this decomposition, it is clear that (8.3.15) will be proved once we show that

$$\lim_{M\to\infty}\sup_{\substack{N>M\\ \|h-g\|_{M,[0,T]}\le 1}}\mathbb{E}^{\mathbb{P}^0}\left[\left(\sum_{m=0}^{[2^MT]-1}\left|p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\right|\right.\right.$$
$$\left.\left.\times F_k^{M,N}\big(m,x,\tilde{p}_h^M\big)\right)^2\right]=0, \tag{8.3.16}$$

where $F_k^{M,N}(m,x,p)$ is given by

$$\sum_{L=M+1}^{N}2^M\int_{I_{m,M}}\left|\hat{V}_k^{M,L}\big(\tau,X^N(m2^{-M},x,p),p\big)\right|d\tau,$$

and that

$$\lim_{M\to\infty}\sup_{\substack{N>M\\ \|h-g\|_{M,[0,T]}\le 1}}\mathbb{E}^{\mathbb{P}^0}\left[\max_{m<[2^MT]}\left|J_k^{M,N}\big(m,x,(p,h)\big)\right|^2\right]=0, \tag{8.3.17}$$

where $J_k^{M,N}\big(m,x,(p,h)\big)$ is equal to

$$\sum_{m'=0}^{m}\sum_{L=M+1}^{N}\int_{I_{m',M}}\hat{W}_k^{M,L}\big(\tau,X^N(m'2^{-M},x,\tilde{p}_h^M),\tilde{p}_h^M\big)\dot{p}_k(\tau)\,d\tau.$$

By Schwarz's inequality, the expectation value in (8.3.16) is dominated by

$$2^MT\sum_{m=0}^{[2^MT]-1}\mathbb{E}^{\mathbb{P}^0}\left[\Big(p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\Big)^2F_k^{M,N}(m,x,\tilde{p}_h^M)^2\right]$$
$$\le 2T\sum_{m=0}^{[2^MT]-1}\mathbb{E}^{\mathbb{P}^0}\big[F_k^{M,N}(m,x,\tilde{p}_h^M)^2\big].$$

In addition, by Minkowski's and Jensen's inequalities,

$$\mathbb{E}^{\mathbb{P}^0}\big[F_k^{M,N}(m,x,\tilde{p}_h^M)^2\big]^{\frac12}$$
$$\le\sum_{L=M+1}^{N}\sup_{x\in\mathbb{R}^n}\mathbb{E}^{\mathbb{P}^0}\left[\left(2^M\int_{I_{m,M}}\left|\hat{V}^{M,L}(\tau,x,\tilde{p}_h^M)\right|d\tau\right)^2\right]^{\frac12}$$
$$\le 2^{\frac{M}{2}}\sum_{L=M+1}^{N}\left(\sup_{x\in\mathbb{R}^n}\int_{I_{m,M}}\mathbb{E}^{\mathbb{P}^0}\left[\left|\hat{V}^{M,L}(\tau,x,\tilde{p}_h^M)\right|^2\right]d\tau\right)^{\frac12}.$$

Thus, (8.3.16) will be proved once we show that

$$\mathbb{E}^{\mathbb{P}^0}\left[\int_{I_{m,M}} \left|\hat{V}_k^{M,L}\big(\tau,x,\tilde{p}_h^M\big)\right|^2 d\tau\right] \le C\big(\|h\|_{M,[0,T]}\big)2^{-L-\frac{3M}{2}} \tag{8.3.18}$$

for some non-decreasing $r \in [0,\infty) \longmapsto C(r) \in [0,\infty)$ and all $x \in \mathbb{R}^n$ and $0 \le m < [2^M T]$. To this end, first observe that the left hand side of (8.3.18) is dominated by

$$C\big(\|h\|_{M,[0,T]}\big) \sup_{x\in\mathbb{R}^n} \int_{I_{0,M}} \mathbb{E}^{\mathbb{P}^0}\left[\left|D^{L-1,L}([\tau]_L, x, \delta_{m2^{-M}}\tilde{p}_h^M)\right|^2\right] d\tau.$$

With the help of the following lemma, we will be able to use Theorem 8.2.12 to control this last expression.

8.3.19 LEMMA. *Given $0 \le \tau < 2^{-M}$ and a $\mathcal{B}_\tau$-measurable, continuous $\Psi : C\big([0,\infty);\mathbb{R}^r\big) \longrightarrow [0,\infty)$,*

$$\mathbb{E}^{\mathbb{P}^0}\left[\Psi\big(\delta_{m2^{-M}}\tilde{p}_h^M\big)\right] \le \left(\frac{2^M}{1-2^M\tau}\right)^{\frac{n}{2q}} e^{\frac{\|h\|^2_{M,I_{m,M}}}{2q}} \|\Psi\|_{L^q(P^0)}$$

for every $q \in [1,\infty]$.

PROOF: First observe that (cf. Lemma 8.3.6)

$$\mathbb{E}^{\mathbb{P}^0}\left[\Psi\big(\delta_{m2^{-M}}\tilde{p}_h^M\big)\right] = \mathbb{E}^{\mathbb{P}^0}\left[\Psi\big(\tilde{p}^M_{\delta_{m2^{-M}}h}\big)\right] = \mathbb{E}^{\mathbb{P}^0}\left[\Psi\big(\tilde{p}^M + h_m^M\big)\right] = \tilde{\Psi}^M(h_m^M),$$

where $h_m^M(t) = (1 \wedge 2^M t)\big(h((m+1)2^M) - h(m2^M)\big)$. Indeed, the first of these is just the time shift-invariance of Brownian increments, the second comes from the fact that $\delta_{m2^{-M}}h^M$ coincides with h_m^M on $[0, 2^{-M}]$, and the last is the definition of $\tilde{\Psi}^M$ in Lemma 8.3.6.

The next step is to show that another expression for $\tilde{\Psi}^M(h_m^M)$ is

$$\left(\frac{2^M}{1-2^M\tau}\right)^{\frac{n}{2}} e^{2^{M-1}|h_m^M(2^{-M})|^2}\mathbb{E}^{\mathbb{P}^0}\left[\Psi(p)\exp\left(-\frac{2^M|p(\tau)-h_m^M(2^{-M})|^2}{2(1-2^M\tau)}\right)\right],$$

and, while doing this, we may and will assume that Ψ is not only continuous and non-negative but also bounded. But, by Lemma 8.3.6, we know that, for any bounded continuous $f : \mathbb{R}^r \longrightarrow \mathbb{R}$,

$$\mathbb{E}^{\mathbb{P}^0}\left[\tilde{\Psi}^M(p)f\big(p(2^{-M})\big)\right] = \mathbb{E}^{\mathbb{P}^0}\left[\Psi(p)f\big(p(2^{-M})\big)\right].$$

Further, if $y \in \mathbb{R}^r \longmapsto \ell_y^M \in C\big([0,\infty);\mathbb{R}^r\big)$ is given by $\ell_y^M(t) = (t \wedge 2^{-M})y$, then

$$\mathbb{E}^{\mathbb{P}^0}\left[\tilde{\Psi}^M(p)f\big(p(2^{-M})\big)\right] = \int_{\mathbb{R}^r} \tilde{\Psi}^M(\ell_y^M)f(y)\gamma_{2^{-M}}(y)\,dy$$

while

$$\mathbb{E}^{\mathbb{P}^0}\big[\Psi(p)f\big(p(2^{-M})\big)\big] = \int_{\mathbb{R}^r} \mathbb{E}^{\mathbb{P}^0}\Big[\Psi(p)\gamma_{2^{-M}-\tau}\big(y - p(\tau)\big)\Big] f(y)\,dy,$$

where $\gamma_t(y) = (2\pi t)^{-\frac{r}{2}} \exp\left(-\frac{|y|^2}{2t}\right)$. Hence, because Ψ, and therefore also $\tilde{\Psi}^M$, is bounded and continuous, the asserted equality follows by choosing a sequence of f's which form an approximate identity at $h((m+1)2^{-M}) - h(m2^{-M})$.

Given the preceding, there are two ways in which the proof can be completed. One is to note that, trivially,

$$\mathbb{E}^{\mathbb{P}^0}\big[\Psi\big(\tilde{p}^M + h_m^M\big)\big] \le \|\Psi\|_{L^\infty(\mathbb{P}^0)},$$

whereas, by the preceding,

$$\mathbb{E}^{\mathbb{P}^0}\big[\Psi\big(\tilde{p}^M + h_m^M\big)\big] \le \|\Psi\|_{L^1(\mathbb{P}^0)} \left(\frac{2^M}{1 - 2^M\tau}\right)^{\frac{n}{2}} e^{2^{M-1}|h_m^M(2^{-M})|^2}.$$

Hence interpolation provides the desired estimate. Alternatively, one can apply Hölder's inequality to get

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^0}\left[\Psi(p)\exp\left(-\frac{2^M|p(\tau) - h_m^M(2^{-M})|^2}{2(1-2^M\tau)}\right)\right] \\
&\qquad \le \|\Psi\|_{L^q(\mathbb{P}^0)} \left(\int_{\mathbb{R}^r} \exp\left(-\frac{q'2^M|y - h_m^M(2^{-M})|^2}{2(1-2^M\tau)}\right)\gamma_\tau(y)\,dy\right)^{\frac{1}{q'}},
\end{aligned}$$

perform an elementary Gaussian integration, and arrive at the desired result after making some simple estimates. $\square$

If we now apply Lemma 8.3.19 with

$$\Psi(p) = \big|X^L([\tau]_L, x, p) - X^{L-1}([\tau]_L, x, p)\big|^2$$

and $q = n$, we see from Theorem 8.2.12 that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^0}\Big[\big|X^L([\tau]_L, x, \tilde{p}_h^M) - X^{L-1}([\tau]_L, x, \tilde{p}_h^M)\big|^2\Big] \\
&\qquad \le C\big(\|h\|_{M,I_{m,M}}\big)(2^{-L}\tau)\left(\frac{2^M}{1-2^M\tau}\right)^{\frac{1}{2}},
\end{aligned}$$

for an appropriate choice of non-decreasing $r \rightsquigarrow C(r)$. In view of the discussion prior to Lemma 8.3.19, (8.3.18), and therefore (8.3.16), is now an easy step.

8.3.5. The Support Theorem, Part V. What remains is to check (8.3.17). For this purpose, we have to begin by rewriting $\hat{W}_k^{M,L}(\tau,x,p)$ as

$$V_k\Big(\Delta^L(\tau,p), E\big(\xi^L(\tau,p), X^{L-1}([\tau]_{L-1}-[\tau]_M, x, \delta_{[\tau]_M}p)\big)\Big) \\ - V_k\Big(\Delta^L(\tau,p)+\xi^L(\tau,p), X^{L-1}([\tau]_{L-1}-[\tau]_M, x, \delta_{[\tau]_M}p)\Big),$$

where $\xi^L(\tau,p) \equiv \big([\tau]_L-[\tau]_{L-1}, p([\tau]_L)-p([\tau]_{L-1})\big)$. Having done so, one sees that Lemma 8.2.11 applies and shows $\hat{W}_k^{M,L}(\tau,x,p)$ can be written as

$$\frac{1}{2}\sum_{\ell\neq k}[V_\ell,V_k]\big(X^{L-1}([\tau]_{L-1}-[\tau]_M,x,\delta_{[\tau]_M}p)\big)\xi_\ell^L(\tau,p)+R_{k,\ell}^{M,L}(\tau,x,p),$$

where

$$\big|R_{k,\ell}^{M,L}(\tau,x,p)\big| \le Ce^{2\nu\|\Delta^{L-1}(\,\cdot\,,p)\|_{I_{m,M}}}\|\Delta^{L-1}(\,\cdot\,,p)\|_{I_{m,M}}^2 \quad \text{for } \tau\in I_{m,M}.$$

Thus, since

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^0}\Bigg[\Bigg(\sum_{m=0}^{[2^MT]-1}\big|p_k((m+1)2^{-M})-p_k(m2^{-M})\big| \\
&\qquad\qquad \times\sum_{L=M+1}^{N} e^{2\nu\|\Delta^{L-1}(\,\cdot\,,\tilde{p}_h^M)\|_{I_{m,M}}}\|\Delta^{L-1}(\,\cdot\,,\tilde{p}_h^M)\|_{I_{m,M}}^2\Bigg)^2\Bigg] \\
&\le 2^MT\sum_{m=0}^{[2^MT]-1}\Bigg(\sum_{L=M+1}^{N}\mathbb{E}^{\mathbb{P}^0}\Big[\big|p_k((m+1)2^{-M})-p_k(m2^{-M})\big|^2 \\
&\qquad\qquad \times e^{4\nu\|\Delta^{L-1}(\,\cdot\,,\tilde{p}_h^M)\|_{I_{m,M}}}\|\Delta^{L-1}(\,\cdot\,,\tilde{p}_h^M)\|_{I_{m,M}}^4\Big]^{\frac{1}{2}}\Bigg)^2 \\
&\le 2^MT^2C\Bigg(\sum_{L=M+1}^{N}\mathbb{E}^{\mathbb{P}^0}\Big[e^{8\nu\|\Delta^{L-1}(\,\cdot\,,\tilde{p}_h^M)\|_{I_{0,M}}}\|\Delta^{L-1}(\,\cdot\,,\tilde{p}_h^M)\|_{I_{0,M}}^8\Big]^{\frac{1}{4}}\Bigg)^2 \\
&\le C(\|h\|_{M,[0,T]})T^2 2^{-M},
\end{aligned}$$

∎

we are left with showing that for each $1\le k\le r$ and $\ell\neq k$,

$$\text{(8.3.20)}\quad \lim_{M\to\infty}\sup_{\substack{N>M\\ \|h-g\|_{M,[0,T]}\le 1}}\mathbb{E}^{\mathbb{P}^0}\left[\max_{m<[2^MT]}\left|\sum_{m'=0}^{m}H_{k,\ell}^{M,N}\big(m',x,(p,h)\big)\right|^2\right]=0,$$

where $H_{k,\ell}^{M,N}\big(m,x,(p,h)\big)$ equals

$$2^M\Big(p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\Big) \\ \times\int_{I_{m,M}}\sum_{L=M+1}^{N}\alpha_{k,\ell}^{M,L,N}([\tau]_{L-1},x,\tilde p_h^M)\xi_\ell^L(\tau,\tilde p_h^M)\,d\tau$$

with

$$\alpha_{k,\ell}^{M,L,N}(\tau,x,p)\equiv[V_\ell,V_k]\Big(X^{L-1}\big(\tau-[\tau]_M,X^N([\tau]_M,x,p),\delta_{[\tau]_M}p\big)\Big).$$

Since

$$\big|H_{k,0}^{M,N}\big(m,x,(p,h)\big)\big|\le C2^{-M}\big|p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\big|,$$

the case when $\ell=0$ is easy. Thus, we will restrict our attention to $1\le k\ne\ell\le r$, in which case we decompose $H_{k,\ell}^{M,N}\big(m,x,(p,h)\big)$ into the sum

$$\begin{aligned}&\Big(p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\Big)A_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\\ &+\Big(p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\Big)\\ &\qquad\times\Big((h-p)_\ell\big((m+1)2^{-M}\big)-(h-p)_\ell(m2^{-M})\Big)B_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\\ &+\Big(p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\Big)\\ &\qquad\times\Big((h-p)_\ell\big((m+1)2^{-M}\big)-(h-p)_\ell(m2^{-M})\Big)C_{k,\ell}^{M,N}\big(m,x,(p,h)\big),\end{aligned}$$

where $A_{k,\ell}^{M,N}\big(m,x,(p,h)\big)$ equals

$$2^M\int_{I_{m,M}}\sum_{L=M+1}^{N}\alpha_{k,\ell}^{M,L,N}\big([\tau]_{L-1},x,\tilde p_h^M\big)\big(p_\ell([\tau]_L)-p_\ell([\tau]_{L-1})\big)\,d\tau,$$

$B_{k,\ell}^{M,N}\big(m,x,(p,h)\big)$ equals

$$4^M\int_{I_{m,M}}\sum_{L=M+1}^{N}\alpha_{k,\ell}^{M,L,N}\big(m2^{-M},x,\tilde p_h^M\big)\big([\tau]_L-[\tau]_{L-1}\big)\,d\tau,$$

and $C_{k,\ell}^{M,N}\big(m,x,(p,h)\big)$ equals

$$4^M\int_{I_{m,M}}\sum_{L=M+1}^{N}\Big(\alpha_{k,\ell}^{M,L,N}\big([\tau]_{L-1},x,\tilde p_h^M\big) \\ -\alpha_{k,\ell}^{M,L,N}\big(m2^{-M},x,\tilde p_h^M\big)\Big)\big([\tau_L]-[\tau]_{L-1}\big)\,d\tau$$

Because $\left|C_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\right|$ is dominated by a constant times

$$\left\|X^{L-1}\big(\,\cdot\,,X^N(m2^{-M},x,\tilde p_h^M),\delta_{m2^{-M}}\tilde p_h^M\big)-X^N(m2^{-M},x,\tilde p_h^M)\right\|_{I_{0,M}},$$

a simple application of (8.3.13) leads to

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^0}\Bigg[\max_{m<[2^MT]}\Bigg|\sum_{m'=0}^{m}\Big(p_k\big((m'+1)2^{-M}\big)-p_k(m'2^{-M})\Big)\\
&\qquad\times\Big((h-p)_\ell\big((m'+1)2^{-M}\big)-(h-p)_\ell(m'2^{-M})\Big)C_{k,\ell}^{M,N}\big(m',x,(p,h)\big)\Bigg|^2\Bigg]\\
&\le 2^MTC\sup_x\mathbb{E}^{\mathbb{P}^0}\Bigg[\sum_{m=0}^{[2^MT]-1}\Bigg|\Big(p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\Big)\\
&\qquad\qquad\times\Big((h-p)_\ell\big((m+1)2^{-M}\big)-(h-p)_\ell(m2^{-M})\Big)\\
&\qquad\qquad\qquad\times\Big\|X^{L-1}\big(\,\cdot\,,X^N(m2^{-M}x,\tilde p_h^M),\delta_{m2^{-M}}\tilde p_h^M\big)\\
&\qquad\qquad\qquad\qquad-X^N(m2^{-M},x,\tilde p_h^M)\Big\|_{I_{0,M}}\Bigg|^2\Bigg]\\
&\le\big(1+\|h\|_{M,[0,T]}^2\big)CT^2\sup_{\substack{x\in\mathbb{R}^n\\ N\ge M}}\mathbb{E}^{\mathbb{P}^0}\Big[\|X^N(\,\cdot\,,x,\tilde p_{h_m^M}^M)-x\|_{I_{0,M}}^4\Big]^{\frac12}\\
&\le C\big(\|h\|_{M,[0,T]}\big)T^22^{-M}.
\end{aligned}$$

∎

The key to handling the other terms is the observation that, because $\ell\ne k$,

$$\mathbb{E}^{\mathbb{P}^0}\Big[\Big(p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\Big)A_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\Big|\,\mathcal{B}_{m2^{-M}}\Big]=0$$

and

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^0}\Big[\Big(p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\Big)\\
&\qquad\times\Big((h-p)_\ell\big((m+1)2^{-M}\big)-(h-p)_\ell(m2^{-M})\Big)\,\Big|\,\mathcal{B}_{m2^{-M}}\Big]=0.
\end{aligned}$$

Thus, since $B_{k,\ell}^{M,N}\big(m,x,(p,h)\big)$ is $\mathcal{B}_{m2^{-M}}$, both

$$\sum_{m'=0}^{m-1}\Big(p_k\big((m'+1)2^{-M}\big)-p_k(m'2^{-M})\Big)A_{k,\ell}^{M,N}\big(m',x,(p,h)\big)$$

and

$$\begin{aligned}
&\sum_{m'=0}^{m-1}\Big(p_k\big((m+1)2^{-M}\big)-p_k(m2^{-M})\Big)\\
&\qquad\times\Big((h-p)_\ell\big((m+1)2^{-M}\big)-(h-p)_\ell(m2^{-M})\Big)B_{k,\ell}^{M,N}\big(m,x,(p,h)\big)
\end{aligned}$$

are $\mathbb{P}^0$-martingales relative to $\{\mathcal{B}_{m2^{-M}} : m \geq 0\}$. In particular, by Doob's Inequality, the $\mathbb{P}^0$-expected square of the maximums over $m < [2^M T]$ of these are dominated by the sum from $m = 0$ to $[2^M T] - 1$ of, respectively,

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}^0}\left[\Big(p_k\big((m+1)2^{-M}\big) - p_k(m2^{-M})\Big)^2 \big|A_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\big|^2\right]\\&\qquad \leq C2^{-M}\mathbb{E}^{\mathbb{P}^0}\left[\big|A_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\big|^4\right]^{\frac{1}{2}}\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}^0}\Big[\Big(p_k\big((m+1)2^{-M}\big) - p_k(m2^{-M})\Big)^2\\&\qquad \times \Big((h-p)_\ell\big((m+1)2^{-M}\big) - (h-p)_\ell(m2^{-M})\Big)^2 \big|B_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\big|^2\Big]\\&\quad \leq C4^{-M}\big(1 + \|h\|_{M,[0,T]}^2\big)\mathbb{E}^{\mathbb{P}^0}\left[\big|B_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\big|^2\right]\end{aligned}$$

Because $\big|B_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\big|$ is uniformly bounded, we will be done if we can show that

$$\mathbb{E}^{\mathbb{P}^0}\left[\big|A_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\big|^4\right]^{\frac{1}{2}} \leq C2^{-M}.$$

To this end, note that

$$\begin{aligned}&\left|\int_{I_{m,M}} \sum_{L=M+1}^{N} \alpha_{k,\ell}^{M,L,N}\big([\tau]_{L-1}, x, \tilde{p}_h^M\big)\big(p_\ell([\tau]_L) - p_\ell([\tau]_{L-1})\big)\,d\tau\right|^4\\&\quad \leq 8^{-M}\int_{I_{m,M}}\left|\sum_{L=M+1}^{N} \alpha_{k,\ell}^{M,L,N}\big([\tau]_{L-1}, x, \tilde{p}_h^M\big)\big(p_\ell([\tau]_L) - p_\ell([\tau]_{L-1})\big)\right|^4 d\tau.\end{aligned}$$

Thus

$$\mathbb{E}^{\mathbb{P}^0}\left[\big|A_{k,\ell}^{M,N}\big(m,x,(p,h)\big)\big|^4\right] \leq 2^M \int_{I_{m,M}} \mathbb{E}^{\mathbb{P}^0}\left[\big|G_{k,\ell}^{M,N}\big(\tau,x,(p,h)\big)\big|^4\right] d\tau,$$

where

$$\begin{aligned}G_{k,\ell}^{M,N}\big(\tau,x,(p,h)\big) &\equiv \sum_{L=M+1}^{N} \alpha_{k,\ell}^{M,L,N}\big([\tau]_{L-1}, x, \tilde{p}_h^M\big)\big(p_\ell([\tau]_L) - p_\ell([\tau]_{L-1})\big)\\&= \int_{[\tau]_M}^{[\tau]_N} \hat{\alpha}_{k,\ell}^{M,N}(\sigma,\tau,x,\tilde{p}_h^M)\,dp_\ell(\sigma)\end{aligned}$$

when $\hat{\alpha}_{k,\ell}^{M,N}(\sigma,\tau,x,p) = \alpha_{k,\ell}^{M,N}([\tau]_{L-1},x,p)$ for $[\tau]_{L-1} \leq \sigma < [\tau]_L$. Finally, by Burkholder's Inequality,

$$\mathbb{E}^{\mathbb{P}^0}\left[\big|G_{k,\ell}^{M,N}\big(\tau,x,(p,h)\big)\big|^4\right] \leq C\mathbb{E}^{\mathbb{P}^0}\left[\left(\int_{[\tau]_M}^{[\tau]_N} \big|\hat{\alpha}_{k,\ell}^{M,N}(\sigma,\tau,x,\tilde{p}_h^M)\big|^2\,d\sigma\right)^2\right],$$

which is dominated by a constant times 4^{-M}. Hence, we have at last completed the proof of Theorem 8.3.1.

8.3.6. Exercises.

EXERCISE 8.3.21. It should be pointed out that the characterization of $\text{supp}(\mathbb{P}_x^L)$ is *much* easier when L is *strictly elliptic* in the sense that $\{V_1(x), \dots, V_r(x)\}$ spans $\mathbb{R}^n$ at each $x \in \mathbb{R}^n$. Of course, in this case, the statement is that $\text{supp}(\mathbb{P}_x^L)$ is the space of all continuous, $\mathbb{R}^n$-valued paths which start at x. What follows is an outline giving steps for a simple proof of this statement.

For the considerations here, Itô's theory is preferable to Stratonovich's. Thus, we will consider the solution $p \rightsquigarrow X(\,\cdot\,, x, p)$ to the Itô stochastic integral equation (8.2.2). Our ellipticity assumption becomes the assumption that the matrix $a(x) = \sigma(x)\sigma^\top(x)$ is strictly positive definite for each $x \in \mathbb{R}^n$.

(i) If $\big(M(t), \mathcal{F}_t, \mathbb{P}\big)$ is an $\mathbb{R}^n$-valued continuous martingale with $M(0) = 0$ satisfying

$$\sum_{j=1}^n \langle M_j \rangle(dt) \le K dt$$

for some $K < \infty$, show that

$$\mathbb{P}\big(\|M\|_{[0,T]} < R\big) \ge e^{-\frac{KT}{R^2}}$$

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

Hint: Set

$$u_R(t, x) = e^{\frac{Kt}{R^2}} \left(1 - \frac{|x|^2}{R^2}\right),$$

and show that $\big(u(t, M(t)), \mathcal{F}_t, \mathbb{P}\big)$ is a submartingale. In particular, by Doob's Stopping Time Theorem, conclude that

$$e^{\frac{KT}{R^2}} \mathbb{P}(\zeta_R > T) \ge \mathbb{E}^{\mathbb{P}}\Big[u\big(T \wedge \zeta_R, M(t \wedge \zeta_R)\big)\Big] \ge 1$$

where $\zeta_R = \inf\{t \ge 0 : |M(t)| \ge R\}$.

(ii) Given $h \in C^1\big([0, \infty); \mathbb{R}^n\big)$ with $h(0) = x$, set $\rho = \|h - x\|_{[0,T]}$, and choose a bounded, measurable $W : [0, \infty) \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ so that

$$\sigma(y)W(t, y) = b(y) - \dot{h}(t) \quad \text{for } (t, y) \in [0, \infty) \times B_{\mathbb{R}^n}(x, \rho + 1).$$

Next, take $\theta(t, p) = -W\big(t, X(t, x, p)\big)$ and $\zeta(p) = \inf\{t \ge 0 : |X(t, x, p) - x| \ge \rho\}$, set

$$E_\theta(t, p) = \exp\left(\int_0^t \big(\theta(\tau, p), dp(\tau)\big)_{\mathbb{R}^n} - \frac{1}{2}\int_0^t |\theta(\tau, p)|^2 \, d\tau\right),$$

determine (cf. Theorem 6.2.1) the probability measure $\mathbb{Q}$ so that $d\mathbb{Q} \restriction \mathcal{B}_t = E_\theta(t) d\mathbb{P}^0 \restriction \mathcal{B}_t$ for $t \ge 0$, and apply Corollary 6.2.2 together with Doob's Stopping Time Theorem to see that $\big(X(t \wedge \zeta_\rho, x, p) - h(t \wedge \zeta_\rho), \mathcal{B}_t, \mathbb{Q}\big)$ is an $\mathbb{R}^n$-valued martingale which satisfies the conditions in **(i)** with $K = \sup_{|y-x| \le \rho} \text{Trace}\big(\sigma\sigma^\top(y)\big)$.

(iii) By combining **(i)** with **(ii)**, show that $\mathbb{P}^0\big(\|X(\,\cdot\,,x,p)-h\|_{[0,T]}<\epsilon\big)>0$ for all $\epsilon>0$.

EXERCISE 8.3.22. Without much more effort, it is possible to refine (8.3.2) a little. Namely, given $\alpha\in(0,\frac{1}{2})$ and $T\in(0,\infty)$, set

$$\|p\|^{(\alpha)}_{[0,T]}=\sup_{0\le s<t\le T}\frac{|p(t)-p(s)|}{(t-s)^\alpha}\quad\text{for } p\in C\big([0,\infty);\mathbb{R}^n\big).$$

The purpose of this exercise is to show that (8.3.2) can be replaced by the statement that

$$\lim_{M\to\infty}\lim_{\delta\searrow 0}\mathbb{P}^0\Big(\big\|X(\,\cdot\,,x,p)-X(\,\cdot\,,x,g)\big\|^{(\alpha)}_{[0,T]}<\epsilon\,\Big|\,\|p-g\|_{M,[0,T]}\le\delta\Big)=1$$

for all $g\in C^1\big([0,\infty);\mathbb{R}^r\big)$ with $g(0)=0$ and $T>0$.

(i) Begin by showing that for each $q\in[1,\infty)$ and $T>0$ there is a non-decreasing $r\rightsquigarrow K_q(T,r)$ such that

$$\mathbb{E}^{\mathbb{P}^0}\Big[\big|X^N(t,x,\tilde p^M_h)-X^N(s,x,\tilde p^M_h)\big|^{2q}\Big]^{\frac{1}{2q}}\le K_q\big(T,\|h\|_{M,[0,T]}\big)(t-s)^{\frac12}$$

for all $0\le s<t\le T$, $N\ge M$, $x\in\mathbb{R}^n$, and $h\in C\big([0,\infty);\mathbb{R}^n\big)$.

(ii) Using the preceding, Lemma 8.3.6, and Exercise 2.4.17, show that for each $\alpha\in(0,\frac12)$, $q\in[1,\infty)$, and $T\in[0,\infty)$,

$$\sup_{\substack{x\in\mathbb{R}^n\\ \delta\in(0,1]}}\mathbb{E}^{\mathbb{P}^0}\Big[\Big(\big\|X(\,\cdot\,,x,p)\big\|^{(\alpha)}_{[0,T]}\Big)^q\,\Big|\,\|p-g\|\le\delta\Big]<\infty.$$

(iii) Given $0<\alpha<\beta$, show that

$$\|p\|^{(\alpha)}_{[0,T]}\le\big(2\|p\|_{[0,T]}\big)^{1-\frac{\alpha}{\beta}}\big(\|p\|^{(\beta)}_{[0,T]}\big)^{\frac{\alpha}{\beta}};$$

and use this, part **(ii)** above, and (8.3.2) to complete the program.

EXERCISE 8.3.23. Perhaps the single most important application of Theorem 8.3.1 is to the *strong minimum principle* for degenerate elliptic operators. Namely, let $\mathcal{G}$ is an open subset of $\mathbb{R}\times\mathbb{R}^n$, and, given $(s,x)\in\mathcal{G}$, define $\mathcal{G}_L(s,x)$ to be the set of $\big(t,X(t-s)\big)$ where $t\ge s$ and $X\in S(x;V_0,\dots,V_r)$ with $\big(\tau,X(\tau-s)\big)\in\mathcal{G}$ for $\tau\in[s,t]$. The minimum principle for L is the statement that if $u\in C^{1,2}(\mathcal{G};\mathbb{R})$ is a (cf. (8.2.3)) (∂_t+L)-supersolution in $\mathcal{G}_L(s,x)$ (i.e., $(\partial_t+L)u\le 0$ on $\mathcal{G}_L(s,x)$), then

$$u\restriction\mathcal{G}_L(s,x)\ge u(s,x)\implies u\restriction\mathcal{G}_L(s,x)=u(s,x).$$

Here are some steps which lead to this conclusion.

(i) Suppose that $(t, y) \in \mathcal{G}_L(s, x)$ and that $u(t, y) > u(s, x)$. Choose $X \in S(x; V_0, \ldots, V_r)$ so that $X \restriction [0, t-s] \subseteq \mathcal{G}$ and $y = X(t-s)$. Next, choose $\rho > 0$ so that

$$\big\{(\tau, z) : \tau \in [s, t] \text{ and } |z - X(\tau - s)| \le \rho\big\} \subseteq \mathcal{G}$$

and there exists an $\epsilon > 0$ such that $u(t, z) \ge u(s, x) + \epsilon$ when $|z - y| \le \rho$. Set

$$\zeta(p) = \inf\big\{\tau \ge 0 : |p(\tau) - X(\tau)| \ge \rho\big\} \wedge (t - s),$$

and show that $\big(u\big(s + \tau \wedge \zeta(p), p\big(\tau \wedge \zeta\big), \mathcal{B}_\tau, \mathbb{P}_x^L\big)$ is a supermartingale. In particular, conclude that

$$u(s, x) \ge \mathbb{E}^{\mathbb{P}_x^L}\Big[u\big(s + \zeta(p), p\big(\zeta\big)\big)\Big].$$

(ii) Show that $\alpha \equiv \mathbb{P}_x^L(\zeta = t - s\big) > 0$, and combine this with the conclusion reached in **(i)** to arrive at the contradiction $u(s, x) \ge u(s, x) + \epsilon\alpha > u(s, x)$.