Chapter V

Brownian Motion, the Gaussian Lévy Process

What remains of the program initiated in Chapter IV is the construction of a Lévy process for the standard, normal distribution $\gamma_{\mathbf{0},\mathbf{I}}$, the infinitely divisible law whose Fourier transform is $e^{-\frac{|\boldsymbol{\xi}|^2}{2}}$. Indeed, if $\{\mathbf{Z}_\gamma(t) : t \geq 0\}$ is such a process, $\{\mathbf{Z}_\mu(t) : t \geq 0\}$ is an independent, Lévy process for the $\mu \in \mathcal{I}(\mathbb{R}^N)$ whose Fourier transform is given by (4.2.1), and $\{\mathbf{Z}_\gamma(t) : t \geq 0\}$ is independent of $\{\mathbf{Z}_\mu(t) : t \geq 0\}$, then it is an easy matter to check that

$$\mathbf{C}^{\frac{1}{2}}\mathbf{Z}_\gamma(t) + \mathbf{Z}_\mu(t)$$

will be an independent, homogeneous increment process for $\gamma \star \mu$, whose Fourier transform is

$$\exp\biggl(\sqrt{-1}\bigl(\boldsymbol{\xi}, \mathbf{m}\bigr)_{\mathbb{R}^N} - \tfrac{1}{2}\bigl(\boldsymbol{\xi}, \mathbf{C}\boldsymbol{\xi}\bigr)_{\mathbb{R}^N} \\ + \int_{\mathbb{R}^N} \Bigl[e^{\sqrt{-1}(\boldsymbol{\xi},\mathbf{y})_{\mathbb{R}^N}} - 1 - \sqrt{-1}\,\mathbf{1}_{[0,1]}(|\mathbf{y}|)\bigl(\boldsymbol{\xi}, \mathbf{y}\bigr)_{\mathbb{R}^N}\Bigr]\, M(d\mathbf{y})\biggr).$$

Because one its earliest applications was as a mathematical model for the motion "Brownian particles"* such the process $\{\mathbf{Z}_\gamma(t) : t \geq 0\}$ is called a **Brownian motion**. In recognition of its origins, we will adopt this terminology and will use the notation $\{\mathbf{B}(t) : t \geq 0\}$ instead of $\{\mathbf{Z}_\gamma(t) : t \geq 0\}$.

§ 5.1 A Construction of Brownian Motion

Before getting into the details, it may be helpful to think a little about what sort of properties we should expect the paths $t \rightsquigarrow \mathbf{B}(t)$ will possess. For this purpose, set $M_n = n\bigl(\delta_{n^{-\frac{1}{2}}} + \delta_{-n^{-\frac{1}{2}}}\bigr)^N$, and recall that we already saw that $\pi_{M_n} \Longrightarrow \gamma_{\mathbf{0},\mathbf{I}}$. Since a Poisson process associated with M_n has nothing but jumps of size $n^{-\frac{1}{2}}$, if one believes that the Lévy process for $\gamma_{\mathbf{0},\mathbf{I}}$ should be, in some sense, the limit

* R. Brown, an 18th Century English botanist, observed the motion of pollen articles in a dilute gas. His observations were interpreted by A. Einstein as evidence for the kinetic theory of gases. In his famous 1905 paper, Einstein took the first steps in a program, eventually completed by N. Wiener in 1923, to give a mathematical model of what Brown had seen.

of such Poisson processes, then it is reasonable to guess that its paths will have jumps of size 0. That is, they will be continuous.

Although the prediction that the paths of $\{\mathbf{B}(t) : t \geq 0\}$ will be continuous is correct, it turns out that, because it is based on the Central Limit Theorem, the heuristic reasoning just given does not lead to the easiest construction. The problem is that the Central Limit Theorem gives convergence of distributions, not random variables, and therefore one should not expect the paths, as opposed to their distributions, of the approximating Poisson processes to converge. For this reason, it is best to avoid the Central Limit Theorem and work with Gaussian random variables from the start.

§ 5.1.1. Deconstructing Brownian Motion. Our construction of Brownian motion is based on an idea of Lévy; and in order to explain Lévy's idea, we will begin with the following line of reasoning.

Assume that $\{\mathbf{B}(t) : t \geq 0\}$ is a Brownian motion in $\mathbb{R}^N$. That is, $\{\mathbf{B}(t) : t \geq 0\}$ starts at $\mathbf{0}$, has independent increments, any increment $\mathbf{B}(s+t) - \mathbf{B}(s)$ has distribution $\gamma_{\mathbf{0},t\mathbf{I}} \in \mathcal{N}(\mathbf{0}, t\mathbf{I})$, and the paths $t \rightsquigarrow \mathbf{B}(t)$ are continuous. Next, given $n \in \mathbb{N}$, let $t \rightsquigarrow \mathbf{B}_n(t)$ be the polygonal path obtained from $t \rightsquigarrow \mathbf{B}(t)$ by linear interpolation during each time interval $[m2^{-n}, (m+1)2^{-n}]$. Thus,

$$\mathbf{B}_n(t) = \mathbf{B}(m2^{-n}) + 2^n\big(t - m2^{-n}\big)\Big(\mathbf{B}\big((m+1)2^{-n}\big) - \mathbf{B}(m2^{-n})\Big)$$

for $m2^{-n} \leq t \leq (m+1)2^{-n}$. The distribution of $\{\mathbf{B}_0(t) : t \geq 0\}$ is very easy to understand. Namely, if $\mathbf{X}_{m,0} = \mathbf{B}(m) - \mathbf{B}(m-1)$ for $m \geq 1$, then the $\mathbf{X}_{m,0}$'s are independent, standard normal $\mathbb{R}^N$-valued random variable, $\mathbf{B}_0(m) = \sum_{1 \leq m \leq n} \mathbf{X}_{m,0}$, and $\mathbf{B}_0(t) = (m-t)\mathbf{B}_0(m-1) + (t-m+1)\mathbf{B}_0(m)$ for $m-1 \leq t \leq m$. To understand the relationship between successive $\mathbf{B}_n$'s, observe that $\mathbf{B}_{n+1}(m2^{-n}) = \mathbf{B}_n(m2^{-n})$ for all $m \in \mathbb{N}$ and that

$$\begin{aligned}
\mathbf{X}_{m,n+1} &\equiv 2^{\frac{n}{2}+1}\Big(\mathbf{B}_{n+1}\big((2m-1)2^{-n-1}\big) - \mathbf{B}_n\big((2m-1)2^{-n-1}\big)\Big) \\
&= 2^{\frac{n}{2}+1}\left(\mathbf{B}\big((2m-1)2^{-n-1}\big) - \frac{\mathbf{B}\big(m2^{-n}\big) + \mathbf{B}\big((m-1)2^{-n}\big)}{2}\right) \\
&= 2^{\frac{n}{2}}\Big[\Big(\mathbf{B}\big((2m-1)2^{-n-1}\big) - \mathbf{B}\big((m-1)2^{-n}\big)\Big) \\
&\qquad\qquad - \Big(\mathbf{B}\big(m2^{-n}\big) - \mathbf{B}\big((2m-1)2^{-n-1}\big)\Big)\Big],
\end{aligned}$$

and therefore $\{\mathbf{X}_{m,n} : m \geq 1\}$ is again a sequence of independent standard normal random variables. What is less obvious is that $\{\mathbf{X}_{m,n} : (m,n) \in \mathbb{Z}^+ \times \mathbb{N}\}$ is also a family of independent random variables. In fact, checking this requires us to make essential use of the fact that we are dealing with Gaussian random variables.

In preparation for proving the preceding independence assertion, say that $\mathfrak{G} \subseteq L^2(\mathbb{P};\mathbb{R})$ is a **Gaussian family** if $\mathfrak{G}$ is a linear subspace and each element of $\mathfrak{G}$ is a centered (i.e., mean-value 0) Gaussian random variable. Our interest in Gaussian families at this point is that the linear span $\mathfrak{G}(\mathbf{B})$ of $\big\{\big(\boldsymbol{\xi}, \mathbf{B}(t)\big)_{\mathbb{R}^N} : t \geq 0 \text{ and } \boldsymbol{\xi} \in \mathbb{R}^N\big\}$ is one. To see this, simply note that, for any $0 = t_0 < t_1 < \cdots t_n$ and $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n \in \mathbb{R}^N$,

$$\sum_{m=1}^{n} \big(\boldsymbol{\xi}_m, \mathbf{B}(t_m)\big)_{\mathbb{R}^N} = \sum_{\ell=1}^{n} \left(\sum_{m=\ell}^{n} \big(\boldsymbol{\xi}_m, \mathbf{B}(t_\ell) - \mathbf{B}(t_{\ell-1})\big)_{\mathbb{R}^N} \right)_{\mathbb{R}^N},$$

which, as a linear combination of independent centered Gaussians is itself a centered Gaussian.

The crucial fact about Gaussian families is the content of the next lemma.

LEMMA 5.1.1. *Suppose that $\mathfrak{G} \subseteq L^2(\mathbb{P};\mathbb{R})$ is a Gaussian family. Then the closure of $\mathfrak{G}$ in $L^2(\mathbb{P};\mathbb{R})$ is again a Gaussian family. Moreover, for any $S \subseteq \mathfrak{G}$, S is independent of $S^\perp \cap \mathfrak{G}$, where $S^\perp$ is the orthogonal complement of S in $L^2(\mathbb{P};\mathbb{R})$.*

PROOF: The first assertion is easy since, as we noted in the introduction to Chapter III, Gaussian random variables are closed under convergence in probability.

Turning to the second part, what we must show is that if $X_1, \dots, X_n \in S$ and $X'_1, \dots, X'_n \in S^\perp \cap \mathfrak{G}$,

$$\mathbb{E}^{\mathbb{P}}\left[\prod_{m=1}^{n} e^{\sqrt{-1}\,\xi_m X_m} \prod_{m=1}^{n} e^{\sqrt{-1}\,\xi'_m X'_m}\right] = \mathbb{E}^{\mathbb{P}}\left[\prod_{m=1}^{n} e^{\sqrt{-1}\,\xi_m X_m}\right] \mathbb{E}^{\mathbb{P}}\left[\prod_{m=1}^{n} e^{\sqrt{-1}\,\xi'_m X'_m}\right]$$

for any choice of $\{\xi_m : 1 \leq m \leq n\} \cup \{\xi'_m : 1 \leq m \leq n\} \subseteq \mathbb{R}$. But the expectation value on the left is equal to

$$\begin{aligned}
&\exp\left(-\frac{1}{2}\mathbb{E}^{\mathbb{P}}\left[\left(\sum_{m=1}^{n} \big(\xi_m X_m + \xi'_m X'_m\big)\right)^2\right]\right) \\
&\quad = \exp\left(-\frac{1}{2}\mathbb{E}^{\mathbb{P}}\left[\left(\sum_{m=1}^{n} \xi_m X_m\right)^2\right] - \frac{1}{2}\mathbb{E}^{\mathbb{P}}\left[\left(\sum_{m=1}^{n} \xi'_m X'_m\right)^2\right]\right) \\
&\quad = \mathbb{E}^{\mathbb{P}}\left[\prod_{m=1}^{n} e^{\sqrt{-1}\,\xi_m X_m}\right] \mathbb{E}^{\mathbb{P}}\left[\prod_{m=1}^{n} e^{\sqrt{-1}\,\xi'_m X'_m}\right],
\end{aligned}$$

since $\mathbb{E}^{\mathbb{P}}[X_m X'_{m'}] = 0$ for all $1 \leq m, m' \leq n$. $\square$

Armed with Lemma 5.1.1, we can now check that $\{\mathbf{X}_{m,n} : (m,n) \in \mathbb{Z}^+ \times \mathbb{N}\}$ is independent. Indeed, since, for all $(m,n) \in \mathbb{Z}^+ \times \mathbb{N}$ and $\boldsymbol{\xi} \in \mathbb{R}^N$, $\big(\boldsymbol{\xi}, \mathbf{X}_{m,n}\big)_{\mathbb{R}^N}$

a member of the Gaussian family $\mathfrak{G}(\mathbf{B})$, all that we have to do is check that, for each $(m,n) \in \mathbb{Z}^+ \times \mathbb{N}$, $\ell \in \mathbb{N}$, and $(\boldsymbol{\xi},\boldsymbol{\eta}) \in (\mathbb{R}^N)^2$,

$$\mathbb{E}^{\mathbb{P}}\big[\big(\boldsymbol{\xi},\mathbf{X}_{m,n+1}\big)_{\mathbb{R}^N}\big(\boldsymbol{\eta},\mathbf{B}(\ell 2^{-n})\big)_{\mathbb{R}^N}\big] = 0.$$

But, since, for $s \le t$, $\mathbf{B}(s)$ is independent of $\mathbf{B}(t) - \mathbf{B}(s)$,

$$\mathbb{E}^{\mathbb{P}}\big[\big(\boldsymbol{\xi},\mathbf{B}(s)\big)_{\mathbb{R}^N}\big(\boldsymbol{\eta},\mathbf{B}(t)\big)_{\mathbb{R}^N}\big] = \mathbb{E}^{\mathbb{P}}\big[\big(\boldsymbol{\xi},\mathbf{B}(s)\big)_{\mathbb{R}^N}\big(\boldsymbol{\eta},\mathbf{B}(s)\big)_{\mathbb{R}^N}\big] = s\big(\boldsymbol{\xi},\boldsymbol{\eta}\big)_{\mathbb{R}^N}$$

and therefore

$$\begin{aligned}
&2^{-\frac{n}{2}-1}\mathbb{E}^{\mathbb{P}}\big[\big(\boldsymbol{\xi},\mathbf{X}_{m,n}\big)_{\mathbb{R}^N}\big(\boldsymbol{\eta},\mathbf{B}(\ell 2^{-n})\big)_{\mathbb{R}^N}\big] \\
&\quad = \mathbb{E}^{\mathbb{P}}\Big[\Big(\boldsymbol{\xi},\mathbf{B}\big((2m-1)2^{-n-1}\big)\Big)_{\mathbb{R}^N}\Big(\boldsymbol{\eta},\mathbf{B}\big(\ell 2^{-n}\big)\Big)_{\mathbb{R}^N}\Big] \\
&\qquad\qquad - \frac{1}{2}\mathbb{E}^{\mathbb{P}}\Big[\Big(\boldsymbol{\xi},\mathbf{B}\big(m2^{-n}\big) - \mathbf{B}\big((m-1)2^{-n}\big)\Big)_{\mathbb{R}^N}\Big(\boldsymbol{\eta},\mathbf{B}\big(\ell 2^{-n}\big)\Big)_{\mathbb{R}^N}\Big] \\
&\quad = 2^{-n}\big(\boldsymbol{\xi},\boldsymbol{\eta}\big)_{\mathbb{R}^N}\left[\big(m-\tfrac{1}{2}\big)\wedge\ell - \frac{m\wedge\ell-(m-1)\wedge\ell}{2}\right] = 0.
\end{aligned}$$

§ **5.1.2. Lévy's Construction of Brownian Motion.** Lévy's idea was to invert the reasoning given in the preceding subsection. That is, start with a family $\{\mathbf{X}_{m,n} : (m,n) \in \mathbb{Z}^+ \times \mathbb{N}\}$ is independent $\mathcal{N}(\mathbf{0},\mathbf{I})$-random variables. Next, define $\{\mathbf{B}_n(t) : t \ge 0\}$ inductively so that $t \rightsquigarrow \mathbf{B}_n(t)$ is linear on each interval $[(m-1)2^{-n}, m2^{-n}]$, $\mathbf{B}_0(m) = \sum_{1\le\ell\le m}\mathbf{X}_{\ell,0}$, $m \in \mathbb{N}$, $\mathbf{B}_{n+1}(m2^{-n}) = \mathbf{B}_n(m2^{-n})$ for $m \in \mathbb{N}$, and

$$\mathbf{B}_{n+1}\big((2m-1)2^{-n}\big) = \mathbf{B}_n\big((2m-1)2^{-n-1}\big) + 2^{-\frac{n}{2}-1}\mathbf{X}_{m,n+1} \quad \text{for } m \in \mathbb{Z}^+.$$

If Brownian motion exists, then the distribution of $\{\mathbf{B}_n(t) : t \ge 0\}$ is the distribution of the process obtained by polygonalizing it on each of the intervals $[(m-1)2^{-n}, m2^{-n}]$, and so the limit $\lim_{n\to\infty}\mathbf{B}_n(t)$ should exist uniformly on compacts and should be Brownian motion.

To see that this procedure works, one must first verify that the preceding definition of $\{\mathbf{B}_n(t) : t \ge 0\}$ gives a process with the right distribution. That is, we need to show that $\big\{\mathbf{B}_n\big((m+1)2^{-n}\big) - \mathbf{B}_n\big(m2^{-n}\big) : m \in \mathbb{N}\big\}$ is a sequence of independent $\mathcal{N}(\mathbf{0}, 2^{-n}\mathbf{I})$-random variables. But, since this sequence is contained in the span of $\{\mathbf{X}_{m,n} : (m,n) \in \mathbb{Z}^+ \times \mathbb{N}\}$, Lemma 5.1.1 says that we will know this once we show that

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\Big[&\Big(\boldsymbol{\xi},\mathbf{B}_n\big((m+1)2^{-n}\big) - \mathbf{B}_n\big(m2^{-n}\big)\Big)_{\mathbb{R}^N} \\
&\times\Big(\boldsymbol{\xi}',\mathbf{B}_n\big((m'+1)2^{-n}\big) - \mathbf{B}_n\big(m'2^{-n}\big)\Big)_{\mathbb{R}^N}\Big] = 2^{-n}\delta_{m,m'}\big(\boldsymbol{\xi},\boldsymbol{\xi}'\big)_{\mathbb{R}^N}
\end{aligned}$$

for $\boldsymbol{\xi}$, $\boldsymbol{\xi}' \in \mathbb{R}^N$ and m, $m' \in \mathbb{N}$. When $n = 0$, this is obvious. Now assume that it is true for n, and observe that

$$\begin{aligned}\mathbf{B}_{n+1}(m2^n) - \mathbf{B}_{n+1}\big((2m-1)2^{-n-1}\big) \\ = \frac{\mathbf{B}_n(m2^{-n}) - \mathbf{B}_n\big((m-1)2^{-n}\big)}{2} - 2^{-\frac{n}{2}-1}\mathbf{X}_{m,n+1}\end{aligned}$$

and

$$\begin{aligned}\mathbf{B}_{n+1}\big((2m-1)2^{-n-1}\big) - \mathbf{B}_{n+1}\big((m-1)2^{-n}\big) \\ = \frac{\mathbf{B}_n(m2^{-n}) - \mathbf{B}_n\big((m-1)2^{-n}\big)}{2} + 2^{-\frac{n}{2}-1}\mathbf{X}_{m,n+1}.\end{aligned}$$

Using these expressions and the induction hypothesis, it is easy to check the required orthogonality.

Second, and more challenging, we must show that, $\mathbb{P}$-almost surely, these processes are converging uniformly on compact time intervals. For this purpose, consider the difference $t \rightsquigarrow \mathbf{B}_{n+1}(t) - \mathbf{B}_n(t)$. Since this path is linear on each interval $[m2^{-n-1}, (m+1)2^{-n-1}]$,

$$\begin{aligned}\max_{t\in[0,2^L]}\big|\mathbf{B}_{n+1}(t) - \mathbf{B}_n(t)\big| &= \max_{1\le m\le 2^{L+n+1}}\big|\mathbf{B}_{n+1}(m2^{-n-1}) - \mathbf{B}_n(m2^{-n-1})\big| \\ &= 2^{-\frac{n}{2}-1}\max_{1\le m\le 2^{L+n+1}}|\mathbf{X}_{m,n+1}| \le 2^{-\frac{n}{2}-1}\left(\sum_{m=1}^{2^{L+n+1}}|\mathbf{X}_{m,n+1}|^4\right)^{\frac{1}{4}}.\end{aligned}$$

Thus, by Jensen's inequality,

$$\mathbb{E}^{\mathbb{P}}\big[\|\mathbf{B}_{n+1} - \mathbf{B}_n\|_{[0,2^L]}\big] \le 2^{\frac{n}{2}-1}\left(\sum_{m=1}^{2^{L+n+1}}\mathbb{E}^{\mathbb{P}}\big[|\mathbf{X}_{m,n+1}|^4\big]\right)^{\frac{1}{4}} = 2^{-\frac{n-L+3}{4}}C_N$$

where $C_N \equiv \mathbb{E}^{\mathbb{P}}\big[|\mathbf{X}_{1,n+1}|^4\big]^{\frac{1}{4}} < \infty$.

Starting from the preceding, it is an easy matter to show that there is a measurable $\mathbf{B} : [0,\infty)\times\Omega \longrightarrow \mathbb{R}^N$ such that $\mathbf{B}(0) = \mathbf{0}$, $\mathbf{B}(\,\cdot\,,\omega) \in C\big([0,\infty);\mathbb{R}^N\big)$ for each $\omega \in \Omega$, and $\|\mathbf{B}_n - \mathbf{B}\|_{[0,t]} \longrightarrow 0$ in $\mathbb{P}$-almost surely and in $L^1(\mathbb{P};\mathbb{R})$ for every $t \in [0,\infty)$. Furthermore, since $\mathbf{B}(m2^{-n}) = \mathbf{B}_n(m2^{-n})$ $\mathbb{P}$-almost surely for all $(m,n) \in \mathbb{N}^2$, it is clear that $\big\{\mathbf{B}\big((m+1)2^{-n}\big) - \mathbf{B}(m2^{-n}) : m \ge 0\big\}$ is a sequence of independent $\mathcal{N}(\mathbf{0}, 2^{-n}\mathbf{I})$-random variables for all $n \in \mathbb{N}$. Hence, by continuity, it follows that $\{\mathbf{B}(t) : t \ge 0\}$ is a Brownian motion.

We have now completed the task described in the introduction to this section. However, before moving on, it is only proper to recognize that, clever as his method is, Lévy was not the first to construct a Brownian motion. Instead, it was N. Wiener who was the first. Indeed, in his famous* 1923 article "Differential Space" in *J. Math. Phys.* #**2** (pp. 131–174) contains three different approaches.

§ 5.1.3. Lévy's Construction in Context. There are elements of Lévy's construction which admit interesting generalizations, perhaps the most important of which is **Kolmogorov's Continuity Criterion**.

THEOREM 5.1.2. *Suppose that $\{X(t) : t \in [0,T]\}$ is a family of random variables taking values in a Banach space B, and assume that, for some $p \in [1,\infty)$, $C < \infty$, and $r \in (0,1]$,*

$$\mathbb{E}^{\mathbb{P}}\big[\|X(t) - X(s)\|_B^p\big]^{\frac{1}{p}} \le C|t-s|^{\frac{1}{p}+r} \quad \text{for all } s, t \in [0,T].$$

Then, there exists a family $\{\tilde{X}(t) : t \in [0,T]\}$ such that $X(t) = \tilde{X}(t)$ $\mathbb{P}$-almost surely for each $t \in [0,T]$ and $t \in [0,T] \longmapsto \tilde{X}(t,\omega) \in B$ is continuous for all $\omega \in \Omega$. In fact, for each $\alpha \in (0,r)$,

$$\mathbb{E}^{\mathbb{P}}\left[\sup_{0\le s<t\le T} \frac{\|\tilde{X}(t) - \tilde{X}(s)\|_B}{(t-s)^\alpha}\right] \le \frac{5CT^{\frac{1}{p}+r-\alpha}}{(1-2^{-r})(1-2^{\alpha-r})}.$$

PROOF: First note that, by rescaling time, it suffices to treat the case when $T = 1$.

Given $n \ge 0$, set $M_n = \max_{1\le m\le 2^n}\big\|X(m2^{-n}) - X\big((m-1)2^{-n}\big)\big\|_B$, and observe that

$$\mathbb{E}^{\mathbb{P}}\big[M_n\big] \le \mathbb{E}^{\mathbb{P}}\left[\left(\sum_{m=1}^{2^n}\big\|X(m2^{-n}) - X\big((m-1)2^{-n}\big)\big\|_B^p\right)^{\frac{1}{p}}\right] \le C2^{-rn}.$$

Next, let $t \rightsquigarrow X_n(t)$ be the polygonal path obtained by linearizing $t \rightsquigarrow X(t)$ on each interval $[(m-1)2^{-n}, m2^{-n}]$, and check that

$$\begin{aligned}&\max_{t\in[0,1]}\|X_{n+1}(t) - X_n(t)\|_B \\ &\quad = \max_{1\le m\le 2^n}\left\|X\big((2m-1)2^{-n-1}\big) - \frac{X\big((m-1)2^{-n}\big) - X(m2^{-n})}{2}\right\|_B \le M_{n+1}.\end{aligned}$$

Hence, $\mathbb{E}^{\mathbb{P}}\Big[\sup_{t\in[0,1]}\|X_{n+1}(t) - X_n(t)\|_B\Big] \le C2^{-rn}$, and so there exists a measurable $\tilde{X} : [0,1]\times\Omega \longrightarrow B$ such that $t \rightsquigarrow \tilde{X}(t,\omega)$ is continuous for all $\omega \in \Omega$ and

$$\mathbb{E}^{\mathbb{P}}\left[\sup_{t\in[0,1]}\|\tilde{X}(t) - X_n(t)\|_B\right] \le \frac{C2^{-rn}}{1-2^{-r}}.$$

* Wiener's article is remarkable, but I must admit that I have never been convinced that it is complete. Undoubtedly, my doubts are more a consequence of my own ineptitude than of his.

Moreover, because, for each $t \in [0,1]$, $\|X(\tau) - X(t)\|_B \longrightarrow 0$ in probability as $\tau \to t$, it is easy to check that, for each $t \in [0,1]$, $\tilde{X}(t) = X(t)$ $\mathbb{P}$-almost surely.

To prove the final estimate, note that for $2^{-n-1} \le t - s \le 2^{-n}$ one has that

$$\begin{aligned}\|\tilde{X}(t) - \tilde{X}(s)\|_B &\le \|\tilde{X}(t) - X_n(t)\|_B + \|X_n(t) - X_n(s)\|_B + \|X_n(s) - \tilde{X}(s)\|_B \\ &\le 2 \sup_{\tau \in [0,1]} \|\tilde{X}(\tau) - X_n(\tau)\|_B + 2^n (t-s) M_n,\end{aligned}$$

and therefore that

$$\frac{\|\tilde{X}(t) - \tilde{X}(s)\|_B}{(t-s)^\alpha} \le 2 2^{\alpha(n+1)} \sup_{\tau \in [0,1]} \|\tilde{X}(\tau) - X_n(\tau)\|_B + 2^n 2^{(\alpha-1)n} M_n.$$

But this means that

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[\sup_{0 \le s < t \le 1} \frac{\|\tilde{X}(t) - \tilde{X}(s)\|_B}{(t-s)^\alpha}\right] &\le C \sum_{n=0}^{\infty} \left(2 \frac{2^{\alpha(n+1)} 2^{-rn}}{1 - 2^{-r}} + 2^{\alpha n} 2^{-rn} \right) \\ &\le \frac{5C}{(1-2^{-r})(1-2^{\alpha - r})}. \quad \square\end{aligned}$$

Corollary 5.1.3. *If $\{\mathbf{B}(t) : t \ge 0\}$ is an $\mathbb{R}^N$-valued Brownian motion, then, for each $\alpha \in \left(0, \frac{1}{2}\right)$, $t \rightsquigarrow \mathbf{B}(t)$ is $\mathbb{P}$-almost surely Hölder continuous of order α. In fact, for each $T \in (0, \infty)$,*

$$\mathbb{E}^{\mathbb{P}}\left[\sup_{0 \le s < t \le T} \frac{|\mathbf{B}(t) - \mathbf{B}(s)|}{(t-s)^\alpha}\right] < \infty.$$

Proof: In view of Theorem 5.1.2, all that we have to do is note that for each $n \in \mathbb{Z}^+$, there is a $C_n < \infty$ such that

$$\mathbb{E}^{\mathbb{P}}\left[|\mathbf{B}(t) - \mathbf{B}(s)|^{2n}\right] \le C_n |t-s|^n. \quad \square$$

§ 5.1.4. Brownian Paths are Non-differentiable. Having shown that Brownian paths are Hölder continuous of every order strictly less than $\frac{1}{2}$, we will close this section by showing that they are nowhere Hölder continuous of any order strictly greater than $\frac{1}{2}$. In particular, this will prove Wiener's famous result that *Brownian paths are nowhere differentiable.* The proof which follows is due to A. Devoretzky.

Theorem 5.1.4. *Let $\{\mathbf{B}(t) : t \ge 0\}$ be an $\mathbb{R}^N$-valued Brownian motion. Then, for each $\alpha > \frac{1}{2}$,*

$$\mathbb{P}\left(\exists s \in [0,\infty)\ \varlimsup_{t \searrow s} \frac{|\mathbf{B}(t) - \mathbf{B}(s)|}{(t-s)^\alpha} < \infty\right) = 0.$$

PROOF: Because $\{\mathbf{B}(T+t)-\mathbf{B}(T) : t \geq 0\}$ is a Brownian motion for each $T \in [0,\infty)$, it suffices for us to show that

$$\mathbb{P}\left(\exists s \in [0,1)\ \varlimsup_{t\searrow s} \frac{|\mathbf{B}(t)-\mathbf{B}(s)|}{(t-s)^\alpha} < \infty\right) = 0.$$

To this end, note that, for every $L \in \mathbb{Z}^+$,

$$\left\{\exists s \in [0,1)\ \varlimsup_{t\searrow s} \frac{|\mathbf{B}(t)-\mathbf{B}(s)|}{(t-s)^\alpha} < \infty\right\}$$
$$\subseteq \bigcup_{M=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=0}^{n} \bigcap_{\ell=0}^{L-1} \left\{\left|\mathbf{B}\left(\tfrac{m+\ell+1}{n}\right) - \mathbf{B}\left(\tfrac{m+\ell}{n}\right)\right| \leq \tfrac{M}{n^\alpha}\right\}.$$

Thus, it enough to show that there is a choice of L such that

$$\lim_{n\to\infty} n\mathbb{P}\Big(\left|\mathbf{B}\left(\tfrac{\ell+1}{n}\right) - \mathbf{B}\left(\tfrac{\ell}{n}\right)\right| \leq \tfrac{M}{n^\alpha},\ 0 \leq \ell < L\Big) = 0.$$

But

$$\mathbb{P}\Big(\left|\mathbf{B}\left(\tfrac{\ell+1}{n}\right) - \mathbf{B}\left(\tfrac{\ell}{n}\right)\right| \leq \tfrac{M}{n^\alpha},\ 0 \leq \ell < L\Big)$$
$$= \gamma_{\mathbf{0},\frac{1}{n}\mathbf{I}}\Big(\overline{B\big(\mathbf{0}, \tfrac{M}{n^\alpha}\big)}\Big)^L = \left((2\pi)^{-\frac{N}{2}} \int_{B(\mathbf{0},Mn^{\frac{1}{2}-\alpha})} e^{-\frac{|\mathbf{y}|^2}{2}}\, d\mathbf{y}\right)^L \leq Cn^{(\frac{1}{2}-\alpha)NL}.$$

Hence, we need only take L so that $(\frac{1}{2}-\alpha)NL > 1$. $\square$

In spite of their being non-differentiable, "differentials" of Brownian paths display remarkable regularity properties. To wit, we give the following simple observation. In its statement, $\|\cdot\|_{\mathrm{H.S.}}$ denotes the Hilbert–Schmidt norm on $\mathrm{Hom}(\mathbb{R}^N;\mathbb{R}^N)$.

THEOREM 5.1.5. *If $\{\mathbf{B}(t) : t \geq 0\}$ is an $\mathbb{R}^N$-valued Brownian motion, then, for each $T \in (0,\infty)$*

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \left\| \sum_{m=1}^{[nt]} \big(\Delta_{m,n}\mathbf{B}\big) \otimes \big(\Delta_{m,n}\mathbf{B}\big) - t\mathbf{I} \right\|_{\mathrm{H.S.}} = 0 \quad \mathbb{P}\text{-almost surely,}$$

where $\Delta_{m,n}\mathbf{B} \equiv \mathbf{B}\left(\frac{m}{n}\right) - \mathbf{B}\left(\frac{m-1}{n}\right)$.

PROOF: Let $(\mathbf{e}_1,\dots,\mathbf{e}_N)$ be an orthonormal basis for $\mathbb{R}^N$, and set $X_i(k,n) = \big(\mathbf{e}_i, \Delta_{k,n}\mathbf{B}\big)_{\mathbb{R}^N}$. Then, what we have to show is that

$$(*)\qquad \lim_{n\to\infty} \sup_{1\leq m\leq nT} \left| \sum_{k=1}^{m} X_i(k,n)X_j(k,n) - \frac{m}{n}\delta_{i,j} \right| = 0 \quad \mathbb{P}\text{-almost surely.}$$

To this end, note that, for each $n \in \mathbb{Z}^+$ and $1 \le i, j \le N$, $\{X_i(k,n) : k \ge 1 \ \&\ 1 \le i \le N\}$ are independent $\mathcal{N}(0, n^{-1})$-random variables. Hence, for each $1 \le i \le N$, $\{X_i(k,n)^2 - n^{-1} : k \ge 1\}$ are independent random variables with mean value 0 and variance $2n^{-2}$, and therefore, by (1.4.21) and the second inequality in (1.3.2),

$$\mathbb{P}\left(\max_{1 \le m \le nT} \left| \sum_{k=1}^{m} \big(X_i(k,n)^2 - n^{-1} \big) \right| \ge \epsilon \right) \le \frac{12 M_4 T^2}{\epsilon^4 n^2},$$

where M_4 fourth moment of $X_1(1,1)^2 - 1$, and so the Borel–Cantelli Lemma can be used to check (*) when $i = j$. When $i \ne j$, the argument is essentially the same, only, because $X_i(k,n)X_j(k,n)$ has mean value 0, there is no need to substract off its mean. □

Just in case it is not obvious, it should be pointed out that the preceding theorem provides another proof that Brownian paths do not have locally bounded variation. Indeed, if $\psi \in C\big([0,T];\mathbb{R}\big)$ has bounded variation, then it is clear that

$$\sum_{m=1}^{[nT]} \Big(\psi\big(\tfrac{m}{n}\big) - \psi\big(\tfrac{m-1}{n}\big) \Big)^2 = 0.$$

Exercises for § 5.1

EXERCISE 5.1.6. This exercise deals with a few elementary facts about Brownian motion.

(i) Let $\{\mathbf{X}(t) : t \ge 0\}$ be an $\mathbb{R}^N$-valued stochastic process satisfying $\mathbf{X}(0,\omega) = \mathbf{0}$ and $\mathbf{X}(\,\cdot\,,\omega) \in C(\mathbb{R}^N)$ for all $\omega \in \Omega$, and show that $\{\mathbf{X}(t) : t \ge 0\}$ is an $\mathbb{R}^N$-valued Brownian motion if and only if the span of $\big\{ \big(\boldsymbol{\xi}, \mathbf{X}(t)\big)_{\mathbb{R}^N} : t \ge 0 \ \&\ \boldsymbol{\xi} \in \mathbb{R}^N \big\}$ is a Gaussian family with the property that, for all $t,\, t' \in [0,\infty)$ and $\boldsymbol{\xi},\, \boldsymbol{\xi}' \in \mathbb{R}^N$,

$$\mathbb{E}^{\mathbb{P}}\Big[\big(\boldsymbol{\xi}, \mathbf{X}(t)\big)_{\mathbb{R}^N} \big(\boldsymbol{\xi}', \mathbf{X}(t')\big)_{\mathbb{R}^N} \Big] = t \wedge t' (\boldsymbol{\xi}, \boldsymbol{\xi}')_{\mathbb{R}^N}.$$

(ii) Assuming that $\{\mathbf{B}(t) : t \ge 0\}$ is an $\mathbb{R}^N$-valued Brownian motion, show that $\{\mathcal{O}\mathbf{B}(t) : t \ge 0\}$ is also an $\mathbb{R}^N$-valued Brownian motion for any orthogonal transformation $\mathcal{O}$. That is, *Brownian motion is invariant under rotation.*

(iii) Assuming that $\{\mathbf{B}(t) : t \ge 0\}$ is an $\mathbb{R}^N$-valued Brownian motion, show that $\{\lambda^{-\frac{1}{2}}\mathbf{B}(\lambda t) : t \ge 0\}$ is also an $\mathbb{R}^N$-Brownian motion for each $\lambda \in (0,\infty)$. This is called the *Brownian scaling invariance* property.

EXERCISE 5.1.7. This exercise introduces the *time inversion invariance* property of Brownian motion.

(**i**) Suppose that $\{\mathbf{B}(t) : t \geq 0\}$ is an $\mathbb{R}^N$-valued Brownian motion, and set $\mathbf{X}(t) = t\mathbf{B}\big(\frac{1}{t}\big)$ for $t > 0$. As an application of Exercise 5.1.6, show that $\{\mathbf{X}(t) : t > 0\}$ has the same distribution as $\{\mathbf{B}(t) : t > 0\}$, and conclude from this that $\lim_{t\searrow 0} \mathbf{X}(t) = \mathbf{0}$ $\mathbb{P}$-almost surely. In particular, if $\tilde{\mathbf{B}}(0,\omega) = \mathbf{0}$ and, for $t \in (0,\infty)$,

$$\tilde{\mathbf{B}}(t,\omega) = \begin{cases} t\mathbf{B}\left(\frac{1}{t},\omega\right) & \text{when } \lim_{t\to 0} t\mathbf{B}\left(\frac{1}{t},\omega\right) = \mathbf{0} \\ \mathbf{0} & \text{otherwise,} \end{cases}$$

then $\{\tilde{\mathbf{B}}(t) : t \geq 0\}$ is an $\mathbb{R}^N$-valued Brownian motion.

(**ii**) As a consequence of part (**i**), prove the *Brownian strong law of large numbers*: $\lim_{t\to\infty} t^{-1}\mathbf{B}(t) = \mathbf{0}$.

Exercise 5.1.8. Let $\{\mathbf{B}(t) : t \geq 0\}$ be an $\mathbb{R}^N$-valued Brownian motion.

(**i**) As an application of Theorem 1.4.13, show that, for any $\mathbf{e} \in \mathbb{S}^{N-1}$ and $T \in (0,\infty)$,

$$\mathbb{P}\left(\sup_{t\in[0,T]} \big|\big(\mathbf{e},\mathbf{B}(t)\big)\big| \geq R\right) \leq 2\mathbb{P}\Big(\big|\big(\mathbf{e},\mathbf{B}(T)\big)_{\mathbb{R}^N}\big| \geq R\Big) \leq 2e^{-\frac{R^2}{2T}},$$

and conclude that

$$\mathbb{P}\big(\|\mathbf{B}\|_{[0,T]} \geq R\big) \leq 4Ne^{-\frac{R^2}{2NT}}. \tag{5.1.9}$$

(**ii**) Now assume that $N = 1$, and set $B^*(t) = \max_{\tau\in[0,t]} B(\tau)$. Just as in part (**i**), use Theorem 1.4.13 to show that $\mathbb{P}\big(B^*(1) \geq a\big) \leq 2\mathbb{P}\big(B(1) \geq a\big)$ for all $a > 0$. By examining the proof, one sees that the inequality comes from not knowing how far over a the partial sums jumps when they first exceed level a. Thus, because we are now dealing with "continuous partial sums," one should suspect that the inequality can be made an inequality. To verify this suspicion, let $\Gamma_n(\epsilon)$ denote the set of ω such that $|B(t,\omega) - B(s,\omega)| < \epsilon$ for all $0 \leq s < t \leq 1$ with $t - s \leq 2^{-n}$, and show that, for $0 < \epsilon < a$,

$$\begin{aligned}&\{B(1) \geq a\} \cap \Gamma_n(\epsilon)\\ &\qquad \subseteq \bigcup_{m=1}^{2^n-1} \left\{\max_{0\leq \ell<m} B(\ell 2^{-n}) < a - \epsilon \leq B(m2^{-n}) \;\&\; B(1) - B(m2^{-n}) > 0\right\},\end{aligned}$$

and conclude that

$$\mathbb{P}\big(\{B(1) \geq a\} \cap \Gamma_n(\epsilon)\big) \leq \tfrac{1}{2}\mathbb{P}\big(B^*(1) \geq a - \epsilon\big)$$

for all $n \in \mathbb{N}$. Now let $n \to \infty$ and then $\epsilon \searrow 0$ to arrive at $\mathbb{P}\big(B^*(1) \geq a\big) \leq 2\mathbb{P}\big(B(1) \geq a\big)$.

(**iii**) By combining the preceding with Brownian scaling invariance, arrive at

$$\mathbb{P}\big(B^*(t) \geq a\big) = 2\mathbb{P}\big(B(t) \geq a\big) = (2\pi)^{-\frac{1}{2}} \int_{at^{-\frac{1}{2}}}^{\infty} e^{-\frac{x^2}{2}}\, dx. \tag{5.1.10}$$

This beautiful result, which is sometimes called the **reflection principle for Brownian motion** seems to have appeared first in L. Bachelier's now famous 1900 thesis, where he used what we now call "Brownian motion" to model price fluctuations on the Paris Bourse.

EXERCISE 5.1.11. Let $\{B(t) : t \geq 0\}$ be an $\mathbb{R}$-valued Brownian. The goal of this exercise is to prove the *Brownian laws of the iterated logarithm*:

$$\overline{\lim_{t\to\infty}} \frac{B(t)}{\sqrt{2t\log_2 t}} = 1 = \overline{\lim_{t\searrow 0}} \frac{B(t)}{\sqrt{2t\log_2 t^{-1}}} \quad \mathbb{P}\text{-almost surely.}$$

Begin by checking that the second equality follows from the first applied to the time inverted process $\{\tilde{B}(t) : t \geq 0\}$ described in (**iii**) of Exercise 5.1.7. Next, observe that

$$\overline{\lim_{n\to\infty}} \frac{B(n)}{\sqrt{2n\log_2 n}} = 1 \quad \mathbb{P}\text{-almost surely}$$

is just the Law of the Iterated Logarithm for standard normal random variables. Thus, all that remains is to show that

$$\overline{\lim_{n\to\infty}} \sup_{t\in[n,n+1]} \left| \frac{B(t)}{\sqrt{2t\log_2 t}} - \frac{B(n)}{\sqrt{2n\log_2 n}} \right| = 0 \quad \mathbb{P}\text{-almost surely,}$$

which can be checked by a combination of the strong law for Brownian motion, the estimate in Exercise 5.1.8, and the easy half of the Borel–Cantelli Lemma.

EXERCISE 5.1.12. Given a stochastic process $\{X(t) :\geq 0\}$, the stochastic process $\{\tilde{X}(t) : t \geq 0\}$ is said to be a **modification** of $\{X(t) : t \geq 0\}$ if, for each $t \in [0,\infty)$, $\tilde{X}(t) = X(t)$ $\mathbb{P}$-almost surely. Further, given a stochastic process $\{X(t) : t \geq 0\}$ with values in a metric space (E,ρ), one says that $\{X(t) : t \geq 0\}$ is **stochastically continuous** if $X(t) \longrightarrow X(s)$ in probability for each $s \in [0,\infty)$.

(**i**) Show that the simple Poisson process $\{N(t) : t \geq 0\}$ is stochastically continuous. Thus, *stochastic continuity does not imply path continuity.*

(**ii**) Let $\mathbb{Q}$ denote the set of rational real numbers. Show that an $\mathbb{R}^N$-valued, stochastically continuous stochastic process $\{X(t) : t \geq 0\}$ admits a continuous modification if and only if, for each $T > 0$, $t \in [0,T] \cap \mathbb{Q} \longmapsto X(t)$ is uniformly continuous. Conclude that a stochastic process $\{X(t) : t \geq 0\}$ admits a continuous modification if and only if there exists a $\mu \in \mathbf{M}_1\big(C(\mathbb{R}^N)\big)$ such that the distribution of $\{X(t) : t \geq 0\}$ under $\mathbb{P}$ is the same as the distribution of $\{\boldsymbol{\psi}(t) : t \geq 0\}$ under μ. Equivalently, $\{X(t) : t \geq 0\}$ admits a continuous modification if and only if there exists a stochastic process $\{Y(t) : t \geq 0\}$, not necessarily on the same probability space, with the same distribution as $\{X(t) : t \geq 0\}$.

Exercise 5.1.13. It is important to realize that the insistence in Theorem 5.1.2 that pth moment of $|\mathbf{X}(t) - \mathbf{X}(s)|$ be dominated by $|t-s|$ to a power strictly greater than p is essential. To see this, recall the simple Poisson process $\{N(t) : t \geq 0\}$ in § 5.2.1. The paths of this process are non-constant, right-continuous, but piecewise constant process. Thus, they are certainly not continuous. On the other hand, show that

$$\mathbb{E}^{\mathbb{P}}\Big[\big(N(t) - N(s) - (t-s)\big)^2\Big] \leq t - s \quad \text{for } 0 \leq s < t.$$

Thus, knowing that a moment of $|\mathbf{X}(t) - \mathbf{X}(s)|$ is dominated by $|t-s|$ is not enough to conclude that there is a continuous modification of $t \rightsquigarrow \mathbf{X}(t)$.

Exercise 5.1.14. In this exercise we will examine a couple of the implications that Theorem 5.1.5 about any Riemann–Stieltjes type integration theory involving Brownian paths. For simplicity, we restrict our attention to the one dimensional case. Thus, let $\{B(t) : t \geq 0\}$ be an $\mathbb{R}$-valued Brownian motion. Because $t \rightsquigarrow B(t)$ is continuous, one knows that any function $\psi : [0,1] \longrightarrow \mathbb{R}$ of bounded variation is Riemann–Stieltjes integrable on $[0,1]$ with respect to $B \restriction [0,1]$. However, as the following shows, almost no Brownian path is Riemann–Stieltjes with respect to itself. Namely, using Theorem 5.1.5, show that $\mathbb{P}$-almost surely,

$$\lim_{n\to\infty} \sum_{m=1}^{n} B\big(\tfrac{m-1}{n}\big)\Big(B\big(\tfrac{m}{n}\big) - B\big(\tfrac{m-1}{n}\big)\Big) = \frac{B(1)^2 - 1}{2},$$
$$\lim_{n\to\infty} \sum_{m=1}^{n} B\big(\tfrac{m}{n}\big)\Big(B\big(\tfrac{m}{n}\big) - B\big(\tfrac{m-1}{n}\big)\Big) = \frac{B(1)^2 + 1}{2},$$

whereas

$$\lim_{n\to\infty} \sum_{m=1}^{n} B\big(\tfrac{2m-1}{2n}\big)\Big(B\big(\tfrac{m}{n}\big) - B\big(\tfrac{m-1}{n}\big)\Big) = B(1)^2.$$

§ 5.2 General Lévy Processes

Our original reason for constructing Brownian motion was to complete the program of constructing all the Lévy processes. In this section, we will do that.

Througout this section, $\mu \in \mathcal{I}(\mathbb{R}^N)$ has Fourier transform

$$\exp\bigg(\sqrt{-1}\,(\boldsymbol{\xi},\mathbf{m})_{\mathbb{R}^N} - \tfrac{1}{2}(\boldsymbol{\xi},\mathbf{C}\boldsymbol{\xi})_{\mathbb{R}^N} + \int\Big[e^{\sqrt{-1}(\boldsymbol{\xi},\mathbf{y})_{\mathbb{R}^N}} - 1 - \sqrt{-1}\,(\boldsymbol{\xi},\mathbf{y})_{\mathbb{R}^N}\Big]\,M(d\mathbf{y})\bigg), \tag{5.1.15}$$

where $\mathbf{m} \in \mathbb{R}^N$, $\mathbf{C} \in \mathrm{Hom}(\mathbb{R}^N;\mathbb{R}^N)$ is symmetric and non-negative definite, and $M \in \mathfrak{M}_2(\mathbb{R}^N)$. In addition, we will use μ_0 to denote $\gamma_{\mathbf{m},\mathbf{C}}$ and μ_1 to denote the element of $\mathcal{I}(\mathbb{R}^N)$ whose Fourier transform is

$$\exp\left(\int \left[e^{\sqrt{-1}(\boldsymbol{\xi},\mathbf{y})_{\mathbb{R}^N}} - 1 - \sqrt{-1}\,(\boldsymbol{\xi},\mathbf{y})_{\mathbb{R}^N}\right] M(d\mathbf{y})\right).$$

Thus, $\mu = \mu_0 \star \mu_1$.

Theorem 5.1.16. *There is a Lévy process $\{\mathbf{Z}(t) : t \geq 0\}$ for μ. Furthermore, there exist independent Lévy processes $\{\mathbf{Z}_0(t) : t \geq 0\}$ and $\{\mathbf{Z}_1(t) : t \geq 0\}$ for μ_0 and μ_1 such that $\mathbf{Z}(t) = \mathbf{Z}_0(t) + \mathbf{Z}_1(t)$, $t \geq 0$, $\mathbb{P}$-almost surely. In fact, if, for $r \in (0,1]$,*

$$\mathbf{Z}^{(r)}(t) = \int_{|\mathbf{y}|>r} \mathbf{y}\, j(t,d\mathbf{y},\mathbf{Z}) - t\int_{r<|\mathbf{y}|\leq 1} \mathbf{y}\, M(d\mathbf{y}),$$

then

$$\mathbb{P}\big(\|\mathbf{Z}^{(r)} - \mathbf{Z}_1\|_{[0,1]} \geq \epsilon\big) \leq \frac{Nt}{\epsilon^2}\int_{\overline{B(\mathbf{0},r)}} |\mathbf{y}|^2\, M(d\mathbf{y}).$$

Proof: Let $\{\mathbf{B}(t) : t \geq 0\}$ be a Brownian motion and $\{\mathbf{Z}_1(t) : t \geq 0\}$ an independent Lévy process for μ_1, and define $\mathbf{Z}_0(t) = t\mathbf{m} + \mathbf{C}^{\frac{1}{2}}\mathbf{B}(t)$ and $\mathbf{Z}(t) = \mathbf{Z}_0(t) + \mathbf{Z}_1(t)$. As we pointed out in the introduction to this chapter, $\{\mathbf{Z}_0(t) : t \geq 0\}$ is a Lévy process for μ_0 and $\{\mathbf{Z}(t) : t \geq 0\}$ is a Lévy process for μ. Furthermore, because $t \rightsquigarrow Z_0(t)$ is continuous, $j(t,\cdot,\mathbf{Z}) = j(t,\cdot,\mathbf{Z}_1)$. Hence, by the last part of Theorem 4.2.15, we know that the last part of the present theorem holds for this choice of $\{\mathbf{Z}(t)\, t \geq 0\}$. Finally, since every Lévy process for μ will have the same distribution as this one, there is nothing more to do. □

Corollary 5.1.17. *Let $\{\mathbf{Z}(t) : t \geq 0\}$ be a Lévy process for μ. Then $t \rightsquigarrow Z(t)$ is $\mathbb{P}$-almost surely continuous if and only if $M = 0$ and is $\mathbb{P}$-almost surely of locally bounded variation if and only if $\mathbf{C} = 0$ and $M \in \mathfrak{M}_1(\mathbb{R}^N)$. Finally, $t \rightsquigarrow \mathbf{Z}(t)$ is $\mathbb{P}$-almost surely an absolutely pure jump path if and only if $\mathbf{C} = 0$, $M \in \mathfrak{M}_1(\mathbb{R}^N)$, and $\mathbf{m} = \int_{\overline{B(0,1)}} \mathbf{y}\, M(d\mathbf{y})$.*

Proof: Let $\mathbf{Z}(t) = \mathbf{Z}_0(t) + \mathbf{Z}_1(t)$ be the decomposition described in Theorem 5.1.16, and let $\{j(t,\cdot\,) : t \geq 0\}$ be the jump process for $\{\mathbf{Z}(t) : t \geq 0\}$. If $M = 0$, then $\mathbf{Z}_1(t) = \mathbf{0}$, $t \geq 0$, $\mathbb{P}$-almost surely, and so $t \rightsquigarrow \mathbf{Z}(t) = \mathbf{Z}_0(t)$ is continuous $\mathbb{P}$-almost surely. Conversely, if $t \rightsquigarrow \mathbf{Z}(t)$ is continuous $\mathbb{P}$-almost surely, then $j(t,\cdot\,) = 0$, $t \geq 0$, $\mathbb{P}$-almost surely. Hence, since $\{j(t,\cdot\,) : t \geq 0\}$ is a Poisson jump process associated with M, we see that $M = 0$. Next, suppose that $\mathbf{C} = 0$ and that $M \in \mathfrak{M}_1(\mathbb{R}^N)$. Then $\mathbf{Z}(t) = \mathbf{Z}_1(t)$, $t \geq 0$ $\mathbb{P}$-almost surely and therefore, by Corollary 4.2.16, $t \rightsquigarrow \mathbf{Z}(t)$ has locally bounded variation $\mathbb{P}$-almost surely if and only if $M \in \mathfrak{M}_1(\mathbb{R}^N)$ and is $\mathbb{P}$-almost surely an absulutely pure jump path if and only if $M \in \mathfrak{M}_1(\mathbb{R}^N)$ and $\mathbf{m} = \int_{\overline{B(\mathbf{0},1)}} \mathbf{y}\, M(d\mathbf{y})$. Thus, all that remains is to show that $\mathbf{C} = 0$ if $t \rightsquigarrow \mathbf{Z}(t)$ $\mathbb{P}$-almost surely has locally bounded variation. But, if $t \rightsquigarrow Z(t)$ has locally bounded variation $\mathbb{P}$-almost surely, then, by (4.1.10), $\int |\mathbf{y}| j(t, d\mathbf{y}) < \infty$, $t \geq 0$, $\mathbb{P}$-almost surely and therefore, by Lemma 4.2.13, $M \in \mathfrak{M}_1(\mathbb{R}^N)$, which, by Corollary 4.2.16, implies that $t \rightsquigarrow \mathbf{Z}_1(t)$ has locally bounded variation $\mathbb{P}$-almost surely. But this means that $t \rightsquigarrow \mathbf{Z}_0(t)$ must also have locally bounded variation $\mathbb{P}$-almost surely, and, since $\{\mathbf{Z}_0(t) : t \geq 0\}$ has the same distribution as $\{t\mathbf{m} + \mathbf{C}^{\frac{1}{2}}\mathbf{B}(t) : t \geq 0\}$, both Theorems 5.1.4 and 5.1.5 show that this is possible only if $\mathbf{C} = 0$. $\square$

Remark 5.1.18**.** Recall the linear functional A_μ introduced in (3.2.10). As we showed in Lemma 3.2.14, the action of A_μ on φ decomposes into a local part and a non-local part, which, with 20-20 hindsight, we can write as, respectively,

$$\big(\mathbf{m}, \nabla\varphi\big)_{\mathbb{R}^N} + \frac{1}{2}\mathrm{Trace}\big(\mathbf{C}\nabla^2\varphi\big)$$

and

$$\int \Big[\varphi(\mathbf{y}) - \varphi(\mathbf{0}) - \mathbf{1}_{[0,1]}\big(|\mathbf{y}|\big)\Big]\, M(d\mathbf{y}).$$

In terms of this decomposition, Corollary 5.1.17 is saying that the local part of A_μ governs the continuous part of $\{\mathbf{Z}(t) : t \geq 0\}$ and that the non-local part governs the discontinuous part.

Exercises for § 5.2

Exercise 5.1.19. Say that a $D(\mathbb{R}^N)$-valued process $\{\mathbf{Z}(t) : t \geq 0\}$ is a Lévy process if $\mathbf{Z}(0) = \mathbf{0}$ and it has independent, homogeneous increments. Show that every Lévy process is a Lévy process for some $\mu \in \mathcal{I}(\mathbb{R}^N)$.

Exercise 5.1.20. Let $\{j(t, \cdot\,) : t \geq 0\}$ be a Poisson jump process associated with some $M \in \mathfrak{M}_\infty(\mathbb{R}^N)$. In Lemma 4.2.13, we showed that when $M \in \mathfrak{M}_2(\mathbb{R}^N)$, then $\int |\mathbf{y}|\, j(t, d\mathbf{y}) < \infty$, $t \geq 0$, with positive probability only if $M \in \mathfrak{M}_1(\mathbb{R}^N)$. In this exercise, we will show that the same is true for any $M \in \mathfrak{M}_\infty(\mathbb{R}^N)$. Thus, assume that $\int |\mathbf{y}|\, j(t, d\mathbf{y}) < \infty$, $t \geq 0$, with positive probability. We want to show that $M \in \mathfrak{M}_1(\mathbb{R}^N)$.

(**i**) As an application of Kolmogorov's 0–1 Law, show that $\int |\mathbf{y}|\, j(t, d\mathbf{y}) < \infty$ with positive probability implies it is finite with probability 1.

(**ii**) Let $\mathcal{N}$ be the set of $\omega \in \Omega$ for which there is a $t > 0$ such that $\int |\mathbf{y}|\, j(t, d\mathbf{y}, \omega) = \infty$. By (**i**), $\mathbb{P}(\mathcal{N}) = 0$. Define $\mathbf{Z}(t, \omega) = \int \mathbf{y}\, j(t, d\mathbf{y}, \omega)$ for $\omega \notin \mathcal{N}$ and $\mathbf{Z}(t, \omega) = \mathbf{0}$ for $\omega \in \mathcal{N}$, and show that $\{\mathbf{Z}(t) : t \geq 0\}$ is a Lévy process with absolutely pure jump paths.

(**iii**) Applying Exercise 5.1.19, first show that $\{\mathbf{Z}(t) : t \geq 0\}$ is a Lévy process for a μ with Lévy measure M, and then apply Corollary 5.1.17 to conclude that $M \in \mathfrak{M}_1(\mathbb{R}^N)$.

§ 5.3 Gaussian Aspect of Brownian Motion

We introduced Brownain motion to complete our program of constructing independent, homogeneous increment processes for infinitely divisibe laws. However, Lévy's construction relies at as heavily on properties of Gaussian random variables as it does on independent increments. In fact, as we will show in this section, by adopting a concertedly Gaussian perspective, one can see Lévy's construction as a very special case of a general procedure for handling Gaussian measures on infinite dimensional spaces. Although the ideas which we will use are implicit in Wiener's work, it was I. Segal and his school, especially L. Gross,* who gave them the form presented here.

In recognition of its provenance, we will call the distribution of an $\mathbb{R}^N$-valued Brownian motion **Wiener measure** and will use $\mathcal{W}^{(N)}$ to denote it. That is, $\mathcal{W}^{(N)}$ is the element of $\mathbf{M}_1\big(C(\mathbb{R}^N)\big)$ determined by

$$\mathcal{W}^{(N)}(\Gamma) = \mathbb{P}\big(\{\omega : \mathbf{B}(\,\cdot\,,\omega) \in \Gamma\}\big), \quad \Gamma \in \mathcal{B}_{C(\mathbb{R}^N)},$$

where $\{\mathbf{B}(t) : t \geq 0\}$ is an $\mathbb{R}^N$-valued Brownian motion on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

§ 5.3.1. The Feynman Representation. In order to get started, we begin with a somewhat fanciful presentation of Wiener's measure. Namely, given $n \in \mathbb{Z}^+$, $0 = t_0 < t_1 < \cdots < t_n$, and a set $A \in \big(\mathcal{B}_{\mathbb{R}^N}\big)^n$, we know that $\mathcal{W}^{(N)}$ assigns $\big\{\boldsymbol{\psi} : \big(\boldsymbol{\psi}(t_1), \dots, \boldsymbol{\psi}(t_n)\big) \in A\big\}$ probability

$$\frac{1}{Z(t_1,\dots,t_n)} \int_A \exp\left[-\sum_{m=1}^n \frac{|\mathbf{y}_m - \mathbf{y}_{m-1}|^2}{t_m - t_{m-1}}\right] d\mathbf{y}_1 \cdots d\mathbf{y}_n$$

where $\mathbf{y}_0 \equiv \mathbf{0}$ and $Z(t_1,\dots,t_n) = \prod_{m=1}^n \big(2\pi(t_m - t_{m_1})\big)^{\frac{N}{2}}$. Now rename the variable $\mathbf{y}_m$ to be "$\boldsymbol{\psi}(t_m)$," and rewrite the preceding as $Z(t_1,\dots,t_n)^{-1}$ times

$$\int_A \exp\left[-\sum_{m=1}^n \frac{t_m - t_{m-1}}{2} \left(\frac{\big|\boldsymbol{\psi}(t_m) - \boldsymbol{\psi}(t_{m-1})\big|}{t_m - t_{m-1}}\right)^2\right] d\boldsymbol{\psi}(t_1) \cdots d\boldsymbol{\psi}(t_n).$$

Obviously, nothing very significant has happened yet since nothing very exciting has been done yet. However, if we now close our eyes, suspend our disbelief, and *pass to the limit* as n tends to infinity and the t_k's become dense, we arrive at the *Feynman's representation*† of Wiener's measure:

$$\mathcal{W}^{(N)}\big(d\boldsymbol{\psi}\big) = \frac{1}{Z} \exp\left[-\frac{1}{2}\int_{[0,\infty)} \big|\dot{\boldsymbol{\psi}}(t)\big|^2\,dt\right] d\boldsymbol{\psi}, \tag{5.3.1}$$

---

* See I.E. Segal's "Distributions in Hilbert space and canonical systems of operators," *T.A.M.S.* **88** (1958) and L. Gross's "Abstract Wiener spaces," *Proc. 5th Berkeley Symp. on Prob. & Stat., 2* (1965). A good exposition of this topic can be found in H.-H. Kuo's *Gaussian Measures in Banach Spaces*, publ. by Springer–Verlag Math. Lec. Notes., no. **463**.

† In truth, Feynman himself never dabbled in considerations so mundane as the ones which follow.

where $\dot{\boldsymbol{\psi}}$ denotes the velocity (i.e., derivative) of $\boldsymbol{\psi}$. Of course, when we reopen our eyes and take a look at (5.3.1), we see that it is riddled with flaws. Not even one of the ingredients on the right-hand side (5.3.1) makes sense! In the first place, the constant Z must be 0 (or maybe ∞). Secondly, since the image of the "measure $d\boldsymbol{\psi}$" under

$$\boldsymbol{\psi} \in C(\mathbb{R}^N) \longmapsto \big(\boldsymbol{\psi}(t_1)\dots,\boldsymbol{\psi}(t_n)\big) \in \big(\mathbb{R}^N\big)^n$$

is Lebesgue's measure for every $n \in \mathbb{Z}^+$ and $0 < t_1 \cdots < t_n$, "$d\boldsymbol{\psi}$" must be the nonexistent *translation invariant measure* on the infinite dimensional space $C(\mathbb{R}^N)$. Finally, if it has any rigorous meaning at all, (5.3.1) certainly seems to be saying that $\mathcal{W}^{(N)}\big(\mathbf{H}(\mathbb{R}^N)\big) = 1$ where

$$\mathbf{H}(\mathbb{R}^N) \equiv \bigg\{\boldsymbol{\psi} \in \mathfrak{P}\big(\mathbb{R}^N\big) : \boldsymbol{\psi}(t) = \int_0^t \dot{\boldsymbol{\psi}}(s)\,ds,\ t \in [0,\infty), \\ \text{with } \dot{\boldsymbol{\psi}} \in L^2\big([0,\infty);\mathbb{R}^N\big)\bigg\}.$$

But even this is entirely wrong, because

$$\boldsymbol{\psi} \in \mathbf{H}(\mathbb{R}^N) \implies \operatorname{var}_{[0,T]}\big(\boldsymbol{\psi}(\,\cdot\,)\big) \le T^{\frac{1}{2}}\|\boldsymbol{\psi}\|_{\mathbf{H}(\mathbb{R}^N)}, \quad T \in [0,\infty),$$

where

$$\|\boldsymbol{\psi}\|_{\mathbf{H}(\mathbb{R}^N)} \equiv \big\|\dot{\boldsymbol{\psi}}\big\|_{L^2([0,\infty);\mathbb{R}^N)},$$

and yet (cf. Theorem 5.1.4) we know that $\mathcal{W}^{(N)}$-almost no $\boldsymbol{\psi}$ has bounded variation on any open interval. That is, although (5.3.1) would seem to be predicting that $\mathbf{H}(\mathbb{R}^N)$ should have full $\mathcal{W}^{(N)}$-measure, $\mathbf{H}(\mathbb{R}^N)$ is invisible to Wiener measure.

§ **5.3.2. Abstract Wiener Space.** At this point it would appear that (5.3.1) does not have very much to recommend it. On the other hand, it is such an intuitively appealing formula that one is reluctant to simply abandon it; and, for this reason, we are going to take another look at it from a slightly different perspective. Namely, notice that $\mathbf{H}(\mathbb{R}^N)$ becomes a separable, real Hilbert space under the norm $\|\cdot\|_{\mathbf{H}(\mathbb{R}^N)}$; the corresponding inner product being

$$\big(\boldsymbol{\varphi},\boldsymbol{\psi}\big)_{\mathbf{H}(\mathbb{R}^N)} = \big(\dot{\boldsymbol{\varphi}},\dot{\boldsymbol{\psi}}\big)_{L^2([0,\infty);\mathbb{R}^N)}.$$

Next, we rewrite (5.3.1) in the form

$$\mathcal{W}^{(N)}(d\mathbf{h}) = (2\pi)^{-\frac{\dim(\mathbf{H})}{2}} \exp\left[-\frac{\|\mathbf{h}\|^2_{\mathbf{H}(\mathbb{R}^N)}}{2}\right] d\mathbf{h} \tag{5.3.2}$$

and thereby come to the conclusion that (5.3.1) is the statement that $\mathcal{W}^{(N)}$ *is the standard Gaussian measure for* $\mathbf{H}(\mathbb{R}^N)$. We have said *for* $\mathbf{H}(\mathbb{R}^N)$ instead of *on* $\mathbf{H}(\mathbb{R}^N)$ since we already know that $\mathbf{H}(\mathbb{R}^N)$ is invisible to $\mathcal{W}^{(N)}$, and, more generally, that Exericise 3.1.10 rules out the possibility of putting a standard Gauss measure on an infinite dimensional Hilbert space. Persevering nonetheless, we attempt giving some substance to this interpretation by using (5.3.2) to guess what the *Fourier transorm* $\mathcal{W}^{(N)}$ looks like. That is, we want to see whether (5.3.2) can be used to guess the value of

$$\widehat{\mathcal{W}^{(N)}}(\mathbf{h}) \equiv \int\limits_{C(\mathbb{R}^N)} \exp\Big[\sqrt{-1}\,\big(\boldsymbol{\psi},\mathbf{h}\big)_{\mathbf{H}(\mathbb{R}^N)}\Big]\,\mathcal{W}^{(N)}(d\boldsymbol{\psi})$$

for $\mathbf{h}\in\mathbf{H}(\mathbb{R}^N)$, where, at least for the moment, we ignore the problem of giving a rigorous meaning to $\big(\boldsymbol{\psi},\mathbf{h}\big)_{\mathbf{H}(\mathbb{R}^N)}$ for $\boldsymbol{\psi}$'s which are not in $\mathbf{H}(\mathbb{R}^N)$. But, as soon as one poses the problem in this way, the answer is immediate: namely, by analogy with what we know about Gaussian measures in finite dimensions, we are compelled to guess that

$$\widehat{\mathcal{W}^{(N)}}(\mathbf{h}) = \exp\left[-\frac{\|\mathbf{h}\|^2_{\mathbf{H}(\mathbb{R}^N)}}{2}\right], \quad \mathbf{h}\in\mathbf{H}(\mathbb{R}^N). \tag{5.3.3}$$

With these heuristic preliminaries in place, we will now see what can be done to make mathematics out of them. From the point of view adopted by Segal's school, this means that we want to find a separable Banach space Θ which, on the one hand, is small enough that $\mathbf{H}(\mathbb{R}^N)$ is embedded as a dense subspace while, at the same time, it is large enough to support a measure for which (5.3.3) (properly interpreted) holds. To make all this more precise, we will need a few elementary facts about Banach spaces and measures on them.

LEMMA 5.3.4. *Let Θ with norm $\|\cdot\|_\Theta$ be a separable, real Banach space, and use*

$$(\theta,\lambda)\in\Theta\times\Theta^*\longmapsto\big\langle\theta,\lambda\big\rangle\in\mathbb{R}$$

to denote the duality relation between Θ and its dual space Θ^*. Then the Borel field $\mathcal{B}_\Theta$ coincides with the σ-algebra generated by the maps $\theta\in\Theta\longmapsto\big\langle\theta,\lambda\big\rangle$ as λ runs over Θ^*. In particular, if, for $\mu\in\mathbf{M}_1(\Theta)$, we define its **Fourier transform** $\hat\mu:\Theta^*\longrightarrow\mathbb{C}$ *by*

$$\hat\mu(\lambda) = \int_\Theta \exp\Big[\sqrt{-1}\,\big\langle\theta,\lambda\big\rangle\Big]\,\mu(d\theta), \quad \lambda\in\Theta^*,$$

then $\hat\mu$ is a continuous function of weak convergence on Θ^*, and $\hat\mu$ uniquely determines μ in the sense that if ν is a second element of $\mathbf{M}_1(\Theta)$ and $\hat\mu=\hat\nu$ then $\mu=\nu$.*

PROOF: Since it is clear that each of the maps $\theta \in \Theta \longmapsto \big\langle \theta, \lambda \big\rangle \in \mathbb{R}$ is continuous and therefore $\mathcal{B}_\Theta$-measurable, the first assertion will follow as soon as we show that the norm $\|\cdot\|_\Theta$ can be expressed as a measurable function of these maps. But, because Θ is separable, we know that Θ^* is separable with respect to the weak* topology and therefore that we can find a sequence $\{\lambda_n\}_1^\infty \subseteq \Theta^*$ so that

$$\|\theta\|_\Theta = \sup_{n \in \mathbb{Z}^+} \big\langle \theta, \lambda_n \big\rangle, \quad \theta \in \Theta.$$

Turning to the properties of $\hat{\mu}$, note that its continuity with respect to weak* convergence is an immediate consequence of Lebesgue's Dominated Convergence Theorem. Furthermore, in view of the preceding, we will know that $\hat{\mu}$ completely determines μ as soon as we show that, for each $n \in \mathbb{Z}^+$ and $\Lambda = \big(\lambda_1, \dots, \lambda_n\big) \in \big(\Theta^*\big)^n$, $\hat{\mu}$ determines the marginal distribution $\mu_\Lambda \in \mathbf{M}_1(\mathbb{R}^N)$ of

$$\theta \in \Theta \longmapsto \Big(\big\langle \theta, \lambda_1 \big\rangle, \dots, \big\langle \theta, \lambda_n \big\rangle\Big) \in \mathbb{R}^n$$

under μ. But this is clear (cf. Lemma 2.3.3), since

$$\widehat{\mu_\Lambda}(\mathbf{x}) = \hat{\mu}\left(\sum_1^n \boldsymbol{\xi}_m \lambda_m\right) \quad \text{for } \boldsymbol{\xi} \in \mathbb{R}^n. \quad \square$$

LEMMA 5.3.5. *Assume that H is a separable, real Hilbert space which is continuously embedded as a dense subspace of a separable, real Banach space Θ. For each $\lambda \in \Theta^*$, there is a unique $h_\lambda \in H$ with the property that*

$$\big(h, h_\lambda\big)_H = \big\langle h, \lambda \big\rangle, \quad h \in H;$$

the mapping $\lambda \in \Theta^* \longmapsto h_\lambda \in H$ is continuous from the weak topology on Θ^* into the weak topology on H; and $\{h_\lambda : \lambda \in \Theta^*\}$ is dense in H. Moreover, there is at most one $\mathcal{W}_H \in \mathbf{M}_1(\Theta)$ with the property that*

$$\widehat{\mathcal{W}_H}(\lambda) = \exp\left[-\frac{\|h_\lambda\|_H^2}{2}\right], \quad \lambda \in \Theta^*, \tag{5.3.6}$$

and in order for $\mathcal{W}_H$ to exist it is necessary that the inclusion mapping taking H into Θ be compact. Finally, if $\mathcal{W}_H$ exists, then there is a unique isometric mapping $h \in H \longmapsto \mathcal{I}(h) \in L^2(\mathcal{W}_H; \mathbb{R})$ with the property that

$$\big[\mathcal{I}(h_\lambda)\big](\theta) = \big\langle \theta, \lambda \big\rangle, \ \theta \in \Theta, \quad \text{for each } \lambda \in \Theta^*.$$

In fact, each $\mathcal{I}(h)$ is, under $\mathcal{W}_H$, an $\mathcal{N}\big(0, \|h\|_H^2\big)$ random variable, and $\{\mathcal{I}(h) : h \in H\}$ is a Gaussian family.

Proof: Because H is continuously embedded in Θ, there exists a $C \in (0,\infty)$ such that

$$\big|\langle h,\lambda\rangle\big| \le C\|h\|_H\|\lambda\|_{\Theta^*}, \quad h \in H.$$

Hence, the Riesz Representation Theorem for Hilbert spaces guarantees both the existence and the uniqueness of h_λ. In fact, $\|h_\lambda\|_H \le C\|\lambda\|_{\Theta^*}$. Moreover, if $\{\lambda_\alpha : \alpha \in I\}$ is a net in Θ^* which is weak* convergent to λ, then

$$\big(h, h_{\lambda_\alpha}\big)_H \longrightarrow \langle h,\lambda\rangle = \big(h,h_\lambda\big)_H$$

for every $h \in H$. Hence, $\{h_{\lambda_\alpha}\}$ tends weakly to h_λ in H, which means that we have the required continuity property of $\lambda \in \Theta^* \longmapsto h_\lambda \in H$. As for the density of $L \equiv \{h_\lambda : \lambda \in \Theta^*\}$, suppose that $h \perp L$. We would then know that $\big\langle h,\lambda\big\rangle = 0$ for all $\lambda \in \Theta^*$, and so h would have to be 0, first as an element of Θ and therefore also as an element of H.

Now assume that $\mathcal{W}_H$ exists. To prove that the inclusion map must be compact, note that, because $\widehat{\mathcal{W}_H}$ is continuous with respect to the weak* topology, (5.3.6) implies that $\lambda \in \Theta^* \longmapsto \|h_\lambda\|_H \in \mathbb{R}$ must also be continuous with respect to the weak* topology. But, after combining this with the continuity statement derived in the preceding paragraph, this implies that $\lambda \in \Theta^* \longmapsto h_\lambda \in H$ is continuous from the weak* topology on Θ^* into the strong topology on H; and, therefore, $A \equiv \big\{h_\lambda : \|\lambda\|_{\Theta^*} \le 1\big\}$ is compact with the respect to the strong topology on H. In particular, this means that if $\{g_\alpha : \alpha \in I\} \subseteq H$ converges weakly to 0 in H, then

$$\|g_\alpha\|_\Theta = \sup_{\|\lambda\|_{\Theta^*}\le 1} \big\langle g_\alpha,\lambda\big\rangle = \sup_{h\in A}\big(g_\alpha,h\big)_H \longrightarrow 0;$$

and so we have now proved that the embedding is a compact map.

Turning to the map $\mathcal{I}$, recall that, under $\mathcal{W}_H$, $\big\langle\,\cdot\,,\lambda\big\rangle$ is an $\mathfrak{N}\big(0,\|h_\lambda\|_H^2\big)$ random variable for each $\lambda \in \Theta^*$, and conclude that $\mathcal{I}$ is isometric on $\big\{h_\lambda : \lambda \in \Theta^*\big\}$. Hence, because $\big\{h_\lambda : \lambda \in \Theta^*\big\}$ is dense in H, the existence as well as the uniqueness of $\mathcal{I}$ are clear. In addition, if h is any element of H and we choose $\{\lambda_n\}_1^\infty$ so that $h_{\lambda_n} \longrightarrow h$ in H, then

$$\mathbb{E}^{\mathcal{W}_H}\Big[\exp\Big(\sqrt{-1}\,\xi\,\mathcal{I}(h)\Big)\Big] = \lim_{n\to\infty}\widehat{\mathcal{W}_H}\big(\xi\,h_{\lambda_n}\big) = \exp\left[-\frac{\xi^2\,\|h\|_H^2}{2}\right], \quad \xi \in \mathbb{R},$$

and therefore $\mathcal{I}(h)$ is an $\mathfrak{N}\big(0,\|h\|_H^2\big)$ random variable under $\mathcal{W}_H$.

Given the preceding, the final assertion is obvious. □

If H is a separable Hilbert space which is embedded as a dense subspace of the separable Banach space Θ and if $\mathcal{W}_H \in \mathbf{M}_1(\Theta)$ satisfies (5.3.6), the triple $(H,\Theta,\mathcal{W}_H)$ is called an **abstract Wiener space**. As we will see in the next subsection, although the Hilbert space H is *canonical*, the Banach space Θ is not. Nonetheless, L. Gross proved that every H admits a Θ on which there exists a $\mathcal{W}_H$ for which $(H, \Theta, \mathcal{W}_H)$ is an abstract Wiener space. Finally, the isometry $h \in H \longmapsto \mathcal{I}(h) \in L^2(\mathcal{W}_H)$ was introduced by Paley and Wiener and will be called the **Paley–Wiener map** and, for reasons explained in the next subsection, $\mathcal{I}(h)$ is called the **Paley–Wiener integral** of h.

Remark. The central issue being discussed in Lemma 5.3.5, as well as the paragraph which follows it, is that of understanding on what Θ's $\mathcal{W}_H$ can exist. In particular, when H is finite dimensional with dimension N, the whole issue disappears since we have have no choice but to take $\Theta = H$ and, after identifying H with $\mathbb{R}^N$, $\mathcal{W}_H = \gamma_{\mathbf{0},\mathbf{I}}$. However, when H is infinite dimensional, the situation is entirely different. In fact, because bounded subsets of H are relatively compact if and only if H is finite dimensional, we know from Lemma 5.3.5 that *the only time when we can take $\Theta = H$ is when H is finite dimensional.* The problem is, of course, that although $\mathcal{W}_H$ always exists (cf. Exercise 5.3.29) on H as a *finitely additive* measure defined on the *algebra* of subsets generated by the maps $h \in H \longmapsto (h, g)_H \in \mathbb{R}$ as g runs over H, *when H is infinite dimensional, $\mathcal{W}_H$ cannot be extended to $\mathcal{B}_H$ as a countably additive measure.*

At first this sort of issue may look a little unfamiliar and might be written off as the sort of pathology which one encounters only in infinite dimensional situations. However, this is not at all the case. For example, consider the problem of putting a translation invariant probability measure on the countable set $\mathbb{Q} \cap [0, 1)$ of rational $q \in [0, 1)$. That is, suppose one attempts to construct a $\mu \in \mathbf{M}_1\big(\mathbb{Q} \cap [0, 1)\big)$ with the property that μ is invariant with respect rational translations (i.e., addition) modulo 1. To this end, one might start by taking $\mathcal{A}$ to be the algebra over $\mathbb{Q} \cap [0, 1)$ which is generated by the collection

$$\big\{[p, q) \cap \mathbb{Q}_1 : p,\, q \in \mathbb{Q} \cap [0, 1) \text{ with } p \le q\big\}.$$

It is then an elementary matter to see that μ exists as the one and only finitely additive measure on $\mathcal{A}$ such that $\mu\big([p, q)\big) = q - p$ for all $p,\, q \in \mathbb{Q}_1$ with $p \le q$. On the other hand, it is equally elementary to see that μ cannot be extended to $\sigma(\mathcal{A})$ (i.e., the set of all subsets of $\mathbb{Q} \cap [0, 1)$) as a countably additive measure. In fact, because μ would have to assign measure 0 to each point, countable additivity would mean that $1 = \mu\big(\mathbb{Q} \cap [0, 1)\big) = 0$. Hence, in this *very* finite dimensional setting, $\mathbb{Q} \cap [0, 1)$ plays the role of H, the interval $[0, 1]$ plays the role of Θ, and Lebesgue's measure $\lambda_{[0,1)}$ that of $\mathcal{W}_H$.

§ 5.3.3. Brownian Motion as an Abstract Wiener Space. The notion of an abstract Wiener space provides us with a context in which to complete our interpretation of (5.3.3). However, we must first find an appropriate separable Banach space $\Theta(\mathbb{R}^N)$; and, because we already know that $\mathcal{W}^{(N)}$ lives on $C(\mathbb{R}^N)$ (which is not itself a Banach space), we should look for a suitable subspace of $C(\mathbb{R}^N)$. Actually, because of (**iii**) in Exercise 5.1.7, we need not look very far.

Namely, set

$$\Theta(\mathbb{R}^N) = \left\{\boldsymbol{\theta} \in C(\mathbb{R}^N) : \boldsymbol{\theta}(0) = 0 \quad \text{and} \quad \lim_{t\to\infty} \frac{|\boldsymbol{\theta}(t)|}{t} = 0\right\},$$

and define

$$\|\boldsymbol{\psi}\|_{\Theta(\mathbb{R}^N)} = \sup_{t\in[0,\infty)} \frac{|\boldsymbol{\psi}(t)|}{1+t} \in [0,\infty] \quad \text{for all } \boldsymbol{\psi} \in C(\mathbb{R}^N).$$

At the same time, let $\Lambda(\mathbb{R}^N)$ be the space of $\mathbb{R}^N$-valued Borel measures $\boldsymbol{\lambda}$ on $(0,\infty)$ satisfying

$$\|\boldsymbol{\lambda}\|_{\Lambda(\mathbb{R}^N)} \equiv \int_{(0,\infty)} (1+t)\,|\boldsymbol{\lambda}|(dt) < \infty, \tag{5.3.7}$$

where $|\boldsymbol{\lambda}|$ denotes the total variation measure determined by $\boldsymbol{\lambda}$.

Lemma 5.3.8. *The map*

$$\boldsymbol{\psi} \in C(\mathbb{R}^N) \longmapsto \|\boldsymbol{\psi}\|_{\Theta(\mathbb{R}^N)} \in [0,\infty]$$

is lower semicontinuous, and the pair $\big(\Theta(\mathbb{R}^N), \|\cdot\|_{\Theta(\mathbb{R}^N)}\big)$ *is a separable Banach space which is continuously embedded as a dense, measurable subset of* $C(\mathbb{R}^N)$. *In particular,* $\mathcal{B}_{\Theta(\mathbb{R}^N)}$ *coincides with* $\mathcal{B}_{C(\mathbb{R}^N)}[\Theta(\mathbb{R}^N)] = \big\{A \cap \Theta(\mathbb{R}^N) : A \in \mathcal{B}_{C(\mathbb{R}^N)}\big\}$*; and the dual space* $\Theta(\mathbb{R}^N)^*$ *of* $\Theta(\mathbb{R}^N)$ *can be identified with the* $\Lambda(\mathbb{R}^N)$ *via the duality relation given by*

$$\big\langle \boldsymbol{\theta}, \boldsymbol{\lambda}\big\rangle \equiv \int_{(0,\infty)} \boldsymbol{\theta}(t)\cdot\boldsymbol{\lambda}(dt), \quad \boldsymbol{\theta} \in \Theta(\mathbb{R}^N),$$

in which case (cf. (5.3.7)) $\|\boldsymbol{\lambda}\|_{\Lambda(\mathbb{R}^N)}$ *is the norm of* $\boldsymbol{\lambda}$ *as an element of* $\Theta(\mathbb{R}^N)^*$. *In addition,* $\mathbf{H}(\mathbb{R}^N)$ *is continuously embedded as a dense, measurable subset of* $\Theta(\mathbb{R}^N)$*; and, if* $\boldsymbol{\lambda} \in \Lambda(\mathbb{R}^N)$*, then, for all* $\boldsymbol{\lambda} \in \Lambda(\mathbb{R}^N)$ *:*

$$\mathbf{h}_{\boldsymbol{\lambda}}(t) \equiv \int_{(0,\infty)} t\wedge\tau\,\boldsymbol{\lambda}(d\tau) = \int_0^t \boldsymbol{\lambda}\big((s,\infty)\big)\,ds, \quad t\in[0,\infty), \tag{5.3.9}$$

is the unique element of $\mathbf{H}(\mathbb{R}^N)$ *satisfying*

$$\big\langle \mathbf{h}, \boldsymbol{\lambda}\big\rangle = \big(\mathbf{h}, \mathbf{h}_{\boldsymbol{\lambda}}\big)_{\mathbf{H}(\mathbb{R}^N)}, \quad \mathbf{h} \in \mathbf{H}(\mathbb{R}^N).$$

Proof: To see that $\|\cdot\|_{\Theta(\mathbb{R}^N)}$ is lower semicontinuous on $C(\mathbb{R}^N)$ and that $\Theta(\mathbb{R}^N) \in \mathcal{B}_{C(\mathbb{R}^N)}$, note that, for any $s \in [0,\infty)$ and $R \in (0,\infty)$,

$$C(s,R) \equiv \Big\{\boldsymbol{\psi} \in C(\mathbb{R}^N) : \big|\boldsymbol{\psi}(t)\big| \le R(1+t) \text{ for } t \ge s\Big\}$$

is closed in $C(\mathbb{R}^N)$. Hence, since $\|\boldsymbol{\psi}\|_{\Theta(\mathbb{R}^N)} \le R \iff \boldsymbol{\psi} \in C(0,R)$, $\|\cdot\|_{\Theta(\mathbb{R}^N)}$ is lower semicontinuous. In addition, since $\{\boldsymbol{\psi} \in C(\mathbb{R}^N) : \boldsymbol{\psi}(0) = \mathbf{0}\}$ is also closed,

$$\Theta(\mathbb{R}^N) = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \left\{\boldsymbol{\psi} \in C\left(m, \tfrac{1}{n}\right) : \boldsymbol{\psi}(0) = \mathbf{0}\right\} \in \mathcal{B}_{C(\mathbb{R}^N)};$$

and there is no doubt that the inclusion map from $\Theta(\mathbb{R}^N)$ into $C(\mathbb{R}^N)$ is continuous or that $\Theta(\mathbb{R}^N)$ is dense in $C(\mathbb{R}^N)$.

In order to analyze the space $\big(\Theta(\mathbb{R}^N), \|\cdot\|_{\Theta(\mathbb{R}^N)}\big)$, define

$$F : \Theta(\mathbb{R}^N) \longrightarrow C_0\big(\mathbb{R};\mathbb{R}^N\big) \equiv \left\{\boldsymbol{\psi} \in C\big(\mathbb{R};\mathbb{R}^N\big) : \lim_{|s|\to\infty} |\boldsymbol{\psi}(s)| = 0\right\}$$

by

$$\big[F(\boldsymbol{\theta})\big](s) = \frac{\boldsymbol{\theta}\,(e^s)}{1+e^s}, \quad s \in \mathbb{R}.$$

As is well-known, $C_0\big(\mathbb{R};\mathbb{R}^N\big)$ with the uniform norm is a separable Banach space; and it is obvious that F is an isometry from $\Theta(\mathbb{R}^N)$ onto $C_0\big(\mathbb{R};\mathbb{R}^N\big)$. Moreover, by the Riesz Representation Theorem for $C_0\big(\mathbb{R};\mathbb{R}^N\big)$, one knows that the dual of $C_0\big(\mathbb{R};\mathbb{R}^N\big)$ is isometric to the space $\mathbf{M}\big(\mathbb{R};\mathbb{R}^N\big)$ of totally finite, $\mathbb{R}^N$-valued measures on $\big(\mathbb{R};\mathcal{B}_{\mathbb{R}}\big)$ with the norm given by total variation. Hence, the identification of $\Theta(\mathbb{R}^N)^*$ with $\Lambda(\mathbb{R}^N)$ reduces to the obvious interpretation of the adjoint map F^* as a mapping from $\mathbf{M}\big(\mathbb{R};\mathbb{R}^N\big)$ onto $\Lambda(\mathbb{R}^N)$.

Turning to the relationship between $\Theta(\mathbb{R}^N)$ and $\mathbf{H}(\mathbb{R}^N)$, first note that

$$\|\mathbf{h}\|_{\Theta(\mathbb{R}^N)} \le \sup_{t\in[0,\infty)} \frac{t^{\frac{1}{2}}}{1+t}\, \|\mathbf{h}\|_{\mathbf{H}(\mathbb{R}^N)} \le \|\mathbf{h}\|_{\mathbf{H}(\mathbb{R}^N)}, \quad \mathbf{h} \in \mathbf{H}(\mathbb{R}^N).$$

Hence $\mathbf{H}(\mathbb{R}^N)$ is certainly continuously embedded in $\Theta(\mathbb{R}^N)$. Moreover, to see that it is dense, simply use the obvious fact that $C_{\mathrm{c}}^{\infty}\big((0,\infty);\mathbb{R}^N\big)$ is already dense in $\Theta(\mathbb{R}^N)$; and to see that it is measurable as a subset of $\Theta(\mathbb{R}^N)$, note that the map $\boldsymbol{\theta} \in \Theta(\mathbb{R}^N) \longmapsto \Psi(\boldsymbol{\theta}) \in [0,\infty]$ given by

$$\Psi(\boldsymbol{\theta}) = \sup\bigg\{\int_{[0,\infty)} \big(\boldsymbol{\theta}(t), \dot{\boldsymbol{\varphi}}(t)\big)_{\mathbb{R}^N}\, dt \\ : \boldsymbol{\varphi} \in C_{\mathrm{c}}\big((0,\infty);\mathbb{R}^N\big) \text{ and } \|\boldsymbol{\varphi}\|_{L^2([0,\infty);\mathbb{R}^N)} \le 1\bigg\}$$

is a lower-semicontinuous function with the property that $\mathbf{H}(\mathbb{R}^N) = \big\{\boldsymbol{\theta} \in \Theta(\mathbb{R}^N) : \Psi(\boldsymbol{\theta}) < \infty\big\}$.

Finally, let $\boldsymbol{\lambda} \in \Lambda(\mathbb{R}^N) \longmapsto \mathbf{h}_{\boldsymbol{\lambda}} \in \mathbf{H}(\mathbb{R}^N)$ be defined as in (5.3.9). It is then an easy application of integration by parts to check that $\langle \mathbf{h}, \boldsymbol{\lambda} \rangle = \big(\mathbf{h}, \mathbf{h}_{\boldsymbol{\lambda}}\big)_H$ holds. $\square$

We now know that $\mathbf{H}(\mathbb{R}^N)$ is embedded in $\Theta(\mathbb{R}^N)$ as a dense, measurable subspace. At the same time, $\Theta(\mathbb{R}^N)$ is a measurable subset of $C(\mathbb{R}^N)$, and, by part (**iii**) of Exercise 5.1.7, it has $\mathcal{W}^{(N)}$-measure 1. Thus, in order to show that $\big(\mathbf{H}(\mathbb{R}^N), \Theta(\mathbb{R}^N), \mathcal{W}^{(N)}\big)$ is an abstract Wiener space, all that remains is to check that (5.3.6) holds for $\mathcal{W}^{(N)}$ when $H = \mathbf{H}(\mathbb{R}^N)$. To this end, let $\boldsymbol{\lambda} \in \Lambda(\mathbb{R}^N)$ be given, and use Fubini's Theorem to check that

$$\begin{aligned}
\|\mathbf{h}_{\boldsymbol{\lambda}}\|^2_{\mathbf{H}(\mathbb{R}^N)} &= \int_{(0,\infty)} \big|\boldsymbol{\lambda}\big((\tau,\infty)\big)\big|^2 \, d\tau \\
&= \iiint\limits_{(0,\infty)^3} \mathbf{1}_{(0,s\wedge t)}(\tau) \, \boldsymbol{\lambda}(ds) \cdot \boldsymbol{\lambda}(dt) \, d\tau = \iint_{(0,\infty)^2} s \wedge t \, \boldsymbol{\lambda}(ds) \cdot \boldsymbol{\lambda}(dt).
\end{aligned}$$

That is,

$$\|\mathbf{h}_{\boldsymbol{\lambda}}\|^2_{\mathbf{H}(\mathbb{R}^N)} = \iint\limits_{(0,\infty)^2} s \wedge t \, \boldsymbol{\lambda}(ds) \cdot \boldsymbol{\lambda}(dt), \quad \boldsymbol{\lambda} \in \Lambda(\mathbb{R}^N). \tag{5.3.10}$$

Given (5.3.10), checking (5.3.6) for $\mathcal{W}^{(N)}$ is easy. Indeed, because $\mathfrak{G}(\mathbf{B}) = \operatorname{span}\big\{\big(\boldsymbol{\xi}, \mathbf{B}(t)\big)_{\mathbb{R}^N} : t \geq 0 \ \& \ \boldsymbol{\xi} \in \mathbb{R}^N\big\}\big)$ is a Gaussian family, $\boldsymbol{\theta} \in \Theta(\mathbb{R}^N) \longmapsto \langle \boldsymbol{\theta}, \boldsymbol{\lambda} \rangle$ is a centered, Gaussian random variable under $\mathcal{W}^{(N)}$ and, for any orthonormal basis $\{\mathbf{e}_1, \dots, \mathbf{e}_N\}$ in $\mathbb{R}^N$,

$$\begin{aligned}
&\mathbb{E}^{\mathcal{W}^{(N)}}\big[\langle \boldsymbol{\theta}, \boldsymbol{\lambda} \rangle^2\big] \\
&= \iint\limits_{(0,\infty)^2} \sum_{i,j=1}^N \mathbb{E}^{\mathcal{W}^{(N)}}\big[\big(\mathbf{e}_i, \boldsymbol{\theta}(s)\big)_{\mathbb{R}^N} \big(\mathbf{e}_j, \boldsymbol{\theta}(s)\big)_{\mathbb{R}^N}\big] \big(\mathbf{e}_i, \boldsymbol{\lambda}(ds)\big)_{\mathbb{R}^N} \big(\mathbf{e}_j, \boldsymbol{\lambda}(ds)\big)_{\mathbb{R}^N} \\
&= \iint_{(0,\infty)^2} s \wedge t \, \boldsymbol{\lambda}(ds) \cdot \boldsymbol{\lambda}(dt) = \|\mathbf{h}_{\boldsymbol{\lambda}}\|^2_{\mathbf{H}(\mathbb{R}^N)},
\end{aligned}$$

this proves that

$$\widehat{\mathcal{W}^{(N)}}(\mathbf{h}_{\boldsymbol{\lambda}}) = e^{-\frac{\|\mathbf{h}_{\boldsymbol{\lambda}}\|^2_{\mathbf{H}(\mathbb{R}^N)}}{2}}, \quad \boldsymbol{\lambda} \in \Lambda(\mathbb{R}^N), \tag{5.3.11}$$

which is the final ingredient needed to prove the following theorem.

Theorem 5.3.12. *The triple $\big(\mathbf{H}(\mathbb{R}^N), \Theta(\mathbb{R}^N), \mathcal{W}^{(N)}\big)$ is an abstract Wiener space.*

Before moving on, there are a couple of points which should be made. First, $\Theta(\mathbb{R}^N)$ is by no means the only choice of Banach space which we could have made. For instance, with a little effort, for each $\alpha \in \left(0, \frac{1}{2}\right)$, we could have introduced a space $\Theta_\alpha(\mathbb{R}^N) \subseteq \Theta(\mathbb{R}^N)$ consisting of paths which are Hölder continuous with exponent α, and there are lots of other choices which readily come to mind. Thus, we have ample confirmation of our earlier assertion that, as distinguished from the Hilbert space, the Banach space appearing in an abstract Wiener triple is not canonical.

The second point is that the Paley–Wiener map $\mathcal{I}$ for $\left(\mathbf{H}(\mathbb{R}^N), \Theta(\mathbb{R}^N), \mathcal{W}^{(N)}\right)$ is a sort of integration theory. To understand this, suppose that $\boldsymbol{\lambda} \in \Lambda(\mathbb{R}^N)$ is supported in $[0, T)$. Then $t \rightsquigarrow \dot{\mathbf{h}}_{\boldsymbol{\lambda}}(t) = \boldsymbol{\lambda}\big((t, \infty)\big)$ a function of bounded variation which vanishes on $[T, \infty)$. In particular, not only is every $\boldsymbol{\theta} \in \Theta(\mathbb{R}^N)$ Riemann–Stieljes integrable with respect to $\dot{\mathbf{h}}_{\boldsymbol{\lambda}}$ but also

$$\langle \boldsymbol{\theta}, \boldsymbol{\lambda} \rangle = -\int_0^T \boldsymbol{\theta}(t) \cdot d\dot{\mathbf{h}}_{\boldsymbol{\lambda}}(t),$$

where the integral on the right is interpeted as a Riemann–Stieltjes integral. Hence, by the integration, by parts formula for Riemann–Stieltjes integration theory,* we know that $\dot{\mathbf{h}}_{\boldsymbol{\lambda}}$ is Riemann–Stieltjes integrable with respect to each $\boldsymbol{\theta}$ and that

$$\langle \boldsymbol{\theta}, \boldsymbol{\lambda} \rangle = \int_0^T \dot{\mathbf{h}}_{\boldsymbol{\lambda}}(t) \cdot d\boldsymbol{\theta}(t),$$

where again the integral on right hand side taken in the sense of Riemann–Stieltjes. Moreover, even if $\boldsymbol{\lambda}$ is not compactly supported, a simple argument shows that in general,

$$\langle \boldsymbol{\theta}, \boldsymbol{\lambda} \rangle = \lim_{T \to \infty} \int_0^T \dot{\mathbf{h}}_{\boldsymbol{\lambda}} \cdot d\boldsymbol{\theta}(t), \quad (\boldsymbol{\theta}, \boldsymbol{\lambda}) \in \Theta(\mathbb{R}^N) \times \Lambda(\mathbb{R}^N), \tag{5.3.13}$$

where, for each $T \in (0, \infty)$, the integral on the right is a Riemann-Stieltjes integral.

If nothing else, (5.3.13) gives further evidence that $\mathcal{I}(\mathbf{h})$ provides a reasonable interpretation of $\big(\boldsymbol{\theta}, \mathbf{h}\big)_{\mathbf{H}(\mathbb{R}^N)}$. In addition, it explains the origin of the terminology *integral* when referring to $\mathcal{I}(\mathbf{h})$. The explanation for the appearence of Paley and Wiener's names is historical: they were the first to consider these objects. See Exercise 5.3.30 for more information.

§ **5.3.4. Wiener Series.** Knowing that $\left(\Theta(\mathbb{R}^N), \mathbf{H}(\mathbb{R}^N), \mathcal{W}^{(N)}\right)$ is an abstract Wiener space, we have another way of thinking about Lévy's construction of Brownian. Namely, choose an orthonormal basis $\{\mathbf{h}_m : m \geq 0\}$ for $\mathbf{H}(\mathbb{R}^N)$. If

* See, for example, Theorem 1.2.7 in my *A Concise Introduction to the Theory of Integration* published by Birkhäuser (3rd edition, 1999).

the paths $\mathbf{B}(\,\cdot\,,\omega)$ of a Brownian motion $\{\mathbf{B}(t) : t \geq 0\}$ were elements of $\mathbf{H}(\mathbb{R}^N)$, then we would have

$$\mathbf{B}(t,\omega) = \sum_{m=0}^{\infty} \big(\mathbf{B}(\,\cdot\,,\omega), \mathbf{h}_m\big)_{\mathbf{H}(\mathbb{R}^N)} \mathbf{h}_m(t),$$

where the convergence would be in $\mathbf{H}(\mathbb{R}^N)$. However, since they $\mathbf{B}(\,\cdot\,,\omega)$ is almost never an element of $\mathbf{H}(\mathbb{R}^N)$, a literal interpretation of the series on the right is problematic.

Fortunately, as we have already (cf. the final paragraph in § 5.3.3) pointed out, an interpretation of $\big(\mathbf{B}, \mathbf{h}_m\big)_{\mathbf{H}(\mathbb{R}^N)}$ is provided by the Paley–Wiener map. If we adopt this interpretation, then (cf. Lemma HBlem) the coefficients of the $\mathbf{h}_m$'s become independent, standard normal variables, and so we are led to guess that if $\{X_m : m \geq 0\}$ is a sequence of indendent $\mathcal{N}(0,1)$ random variables and $\{\mathbf{h}_m : m \geq 0\}$ is an orthonormal basis in $\mathbf{H}(\mathbb{R}^N)$, then the series

$$\mathbf{S}(t) = \sum_{m=0}^{\infty} X_m \mathbf{h}_m(t) \tag{5.3.14}$$

should converge in a reasonable sense to a Brownian motion. For reasons which will become clear in the next paragraph, we will call a series of the form in (5.3.14) a **Wiener series**.

To convince oneself that this line of reasoning has a chance of working, one should observe that Lévy's construction corresponds to a particular choice of the orthonomal basis $\{\mathbf{h}_m : m \geq 0\}$.* To see this, define $f : \mathbb{R} \longrightarrow \{-1, 0, 1\}$ by $f = \mathbf{1}_{[0,\frac{1}{2})} - \mathbf{1}_{[\frac{1}{2},1)}$, and determine $\{\dot{h}_{k,n} : (k,n) \in \mathbb{N}^2\}$ by

$$\dot{h}_{k,0} = \mathbf{1}_{[k,k+1)} \text{ and } \dot{h}_{k,n} = 2^{\frac{n-1}{2}} \begin{cases} 1 & \text{on } \big[k2^{1-n}, (2k+1)2^{-n}\big) \\ -1 & \text{on } \big[(2k+1)2^{-n}, (k+1)2^{1-n}\big) \\ 0 & \text{elsewhere.} \end{cases}$$

Clearly, the $\dot{h}_{k\ell}$'s are orthonormal in $L^2\big([0,\infty);\mathbb{R}\big)$. In addition, for each $n \in \mathbb{N}$, the span of $\{\dot{h}_{k,n} : k \in \mathbb{N}\}$ equals that of of $\{\mathbf{1}_{[k2^{-n},(k+1)2^{-n})} : n \in \mathbb{N}\}$. Perhaps the easiest way to check this is to do so by diminsion counting. That is, for a given $\ell \in \mathbb{N}$, note that $\{\dot{h}_{k,m} : \ell 2^{m-1} \leq k < (\ell+1)2^{m-1} \ \&\ 0 \leq m \leq n\}$ has the same number of elements as $\{\mathbf{1}_{[k2^{-n},(k+1)2^{-n})} : \ell 2^n \leq k(\ell+1)2^n\}$ and that the

* The observation that Lévy's construction can be interpretted in terms of Wiener series is due to Z. Ciesielski. To be more precise, initially Ciesielski himself was thinking entirely in terms of orthogonal series and did not realize that he was giving a re-interpretation of Lévy's construction. Only later did the connection become clear.

first is contained in the span of the second. Hence, if $h_{k,n}(t) = \int \dot{h}_{k,n}(\tau)\,d\tau$ and $(\mathbf{e}_1, \dots, \mathbf{e}_N)$ is an orthonormal basis in $\mathbb{R}^N$, then

$$\left\{\mathbf{h}_{k,n,i} : (k,n,i) \in \mathbb{N}^2 \times \{1,\dots,N\}\right\}$$

is an orthonormal basis in $\mathbf{H}(\mathbb{R}^N)$, known as the **Haar basis**. Finally, if $\left\{X_{k,n,i} : (k,n,i) \in \mathbb{N}^2 \times \{1,\dots,N\}\right\}$ is a family of independent, $\mathcal{N}(0,1)$-random variables and $\mathbf{X}_{k,n} = \sum_{i=1}^N X_{k,n,i}\mathbf{e}_i$, then

$$\sum_{m=0}^{n}\sum_{k=0}^{\infty}\sum_{i=1}^{N} X_{k,m,i}\mathbf{h}_{k,m,i}(t) = \sum_{m=0}^{n}\sum_{k=0}^{\infty} h_{k,m}(t)\mathbf{X}_{k,m}$$

is precisely what we denoted by $\mathbf{B}_n(t)$ in Lévy's construction. Wiener's own construction of Brownian motion was essentially the same, only, he chose a different basis for $\mathbf{H}(\mathbb{R}^N)$. Namely, Wiener took $\dot{h}_{k,0}(t) = \mathbf{1}_{[k,k+1)}(t)$ for $k \in \mathbb{N}$ and $\dot{h}_{k,\ell}(t) = 2^{\frac{1}{2}}\mathbf{1}_{[\ell,k+1)}\cos\bigl(\pi\ell(t-k)\bigr)$ for $(k,\ell) \in \mathbb{N}\times\mathbb{Z}^+$, which means that he was looking at the series

$$\sum_{k=0}^{\infty}(t-k)\mathbf{1}_{[k,k+1)}(t)\mathbf{X}_{k,0} + \sum_{(k,\ell)\in\mathbb{N}\times\mathbb{Z}^+}\mathbf{1}_{[k,k+1)}(t)\frac{2^{\frac{1}{2}}\sin\bigl(\pi\ell(t-k)\bigr)}{\pi\ell}\mathbf{X}_{k,\ell},$$

where again $\{\mathbf{X}_{k,\ell} : (k,\ell) \in \mathbb{N}^2\}$ is a family of independent $\mathcal{N}(\mathbf{0},\mathbf{I})$-random variables. The reason why Lévy's choice is easier to handle than Wiener's is that, for each $n \in \mathbb{N}$ and $t \in [0,\infty)$, $h_{k,n}(t) \neq 0$ for precisely one $k \in \mathbb{N}$. Wiener's choice has no such property.

We now want to show that no matter how one chooses the basis, the corresponding Wiener series gives a realization of Brownian motion.

Theorem 5.3.15. *Let $\{X_m : m \geq 0\}$ be a sequence of independent standard normal random variables. Given an orthonormal basis $\{\mathbf{h}_m : m \geq 0\}$ in $\mathbf{H}(\mathbb{R}^N)$, set*

$$\mathbf{S}_n(t) = \sum_{m=0}^{n} X_m\mathbf{h}_m(t).$$

Then, there exists a Brownian motion $\{\mathbf{B}(t) : t \geq 0\}$ such that, $\mathbb{P}$-almost surely, $\mathbf{S}_n \longrightarrow \mathbf{B}$ in $\Theta(\mathbb{R}^N)$.

Proof: We begin by showing that, for each $t \in [0,\infty)$, $\{\mathbf{S}_n(t) : n \geq 0\}$ converges $\mathbb{P}$-almost surely. To this end, let $\mathbf{e} \in \mathbb{S}^{N-1}$ be given, and set $a_m(t,\mathbf{e}) = \bigl(\mathbf{e},\mathbf{h}_m(t)\bigr)_{\mathbb{R}^N}$. Then

$$\bigl(\mathbf{e},\mathbf{S}_n(t)\bigr)_{\mathbb{R}^N} = \sum_{m=0}^{n} a_m(t,\mathbf{e})X_m.$$

Thus, by Theorem 1.4.2, we will know that $\{\mathbf{S}_n(t) : n \geq 0\}$ converges $\mathbb{P}$-almost surely and in $L^2(\mathbb{P};\mathbb{R}^N)$ once we check that $\sum_{m=0}^{\infty} a_m(t,\mathbf{e})^2 < \infty$. For this purpose, set $\mathbf{h}_t(\tau) = t \wedge \tau\mathbf{e}$, and note that $a_m(t,\mathbf{e}) = \big(\mathbf{h}_t, \mathbf{h}_m\big)_{\mathbf{H}(\mathbb{R}^N)}$. Hence, $\sum_{m=0}^{\infty} a_m(t,\mathbf{e})^2 = \|\mathbf{h}_t\|^2_{\mathbf{H}(\mathbb{R}^N)} = t$.

Knowing the preceding convergence result, we will know that the convergence takes place $\mathbb{P}$-almost surely in $\Theta(\mathbb{R}^N)$ once we show that, $\mathbb{P}$-almost surely, $\{\mathbf{S}_n \restriction [0,T] : n \geq 0\}$ is equicontinuous for each $T > 0$ and that $\sup_{n\geq 0} t^{-1}|\mathbf{S}_n(t)| \longrightarrow 0$ as $t \to \infty$. To this end, we first observe that both these will follow once we show that, for all $\mathbf{e} \in \mathbb{S}^{N-1}$,

$$(*) \qquad \sup_{n\geq 0} \mathbb{E}^{\mathbb{P}}\big[|\mathbf{S}_n(t,\mathbf{e}) - \mathbf{S}_n(s,\mathbf{e})|^4\big] \leq 3(t-s)^2,$$

where $\mathbf{S}_n(t,\mathbf{e}) \equiv \big(\mathbf{S}_n(t),\mathbf{e}\big)_{\mathbb{R}^N}$. To see this, note that, because they are partial sums of centered, independent random variables, (1.4.21) with $p = 2$ implies that

$$\mathbb{E}^{\mathbb{P}}\left[\sup_{n\geq 0} |\mathbf{S}_n(t,\mathbf{e}) - \mathbf{S}_n(s,\mathbf{e})|^4\right] \leq 4 \sup_{n\geq 0} \mathbb{E}^{\mathbb{P}}\big[|\mathbf{S}_n(t,\mathbf{e}) - \mathbf{S}_n(s,\mathbf{e})|^4\big].$$

Thus, (*) implies

$$\mathbb{E}^{\mathbb{P}}\left[\sup_{n\geq 0} |\mathbf{S}_n(t,\mathbf{e}) - \mathbf{S}_n(s,\mathbf{e})|^4\right] \leq 12(t-s)^2,$$

which, by Theorem 5.1.2 applied with $B = \ell^\infty(\mathbb{N};\mathbb{R})$, implies that, for each $\alpha \in \big[0,\frac{1}{4}\big)$, there is a $C_\alpha < \infty$ such that

$$\mathbb{E}^{\mathbb{P}}\left[\sup_{0\leq s<t\leq T}\sup_{n\geq 0} \frac{|\mathbf{S}_n(t,\mathbf{e}) - \mathbf{S}_n(s,\mathbf{e})|}{|t-s|^\alpha}\right] \leq C_\alpha T^{\frac{1}{2}-\alpha}$$

for all $T \in (0,\infty)$. By taking $\alpha = \frac{1}{8}$, we get the required equicontinuity. By taking $\alpha = 0$, we get

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[\sup_{t\geq 2^L}\sup_{n\geq 0} \frac{|\mathbf{S}_n(t,\mathbf{e})|}{t}\right] &\leq \sum_{\ell=L}^{\infty} 2^{-\ell}\mathbb{E}^{\mathbb{P}}\left[\sup_{2^\ell\leq t\leq 2^{\ell+1}}\sup_{n\geq 0}|\mathbf{S}_n(t,\mathbf{e})|\right] \\ &\leq 2^{\frac{1}{2}} C_0 \sum_{\ell=L}^{\infty} 2^{-\frac{\ell}{2}} \longrightarrow 0 \quad \text{as } L \to \infty.\end{aligned}$$

To prove (*), note that because $\mathbf{S}_n(t,\mathbf{e}) - \mathbf{S}_n(s,\mathbf{e})$ is a centered Gaussian, we will know that such an estimate holds as soon as we show that

$$\sup_{n\in\mathbb{N}} \mathbb{E}^{\mathbb{P}}\Big[\big(\mathbf{S}_n(t,\mathbf{e}) - \mathbf{S}_n(s,\mathbf{e})\big)^2\Big] \leq (t-s), \quad 0 \leq s < t.$$

For this purpose, set $\mathbf{h}_{s,t}(\tau) = (t \wedge \tau - s \wedge \tau)\mathbf{e}$. Then

$$\mathbf{S}_n(t, \mathbf{e}) - \mathbf{S}_n(s, \mathbf{e}) = \sum_{m=0}^{n} \big(\mathbf{h}_{s,t}, \mathbf{h}_m\big)_{\mathbf{H}(\mathbb{R}^N)} X_m,$$

and therefore

$$\mathbb{E}^{\mathbb{P}}\Big[\big(\mathbf{S}_n(t, \mathbf{e}) - \mathbf{S}_n(s, \mathbf{e})\big)^2\Big] = \sum_{m=0}^{n} \big(\mathbf{h}_{s,t}, \mathbf{h}_m\big)^2_{\mathbf{H}(\mathbb{R}^N)} \le \|\mathbf{h}_{s,t}\|^2_{\mathbf{H}(\mathbb{R}^N)} = t - s.$$

To complete the proof, let $\mathcal{N}$ be the set of $\omega \in \Omega$ for which $\{\mathbf{S}_n(\,\cdot\,, \omega) : n \ge 0\}$ fails to converge in $\Theta(\mathbb{R}^N)$, and define $\mathbf{B}(t, \omega) = \lim_{n\to\infty} \mathbf{S}_n(t, \omega)$ for $\omega \notin \mathcal{N}$ and $\mathbf{B}(t) = \mathbf{0}$ for $\omega \in \mathcal{N}$. Clearly, for each $t \ge 0$ and $\boldsymbol{\xi} \in \mathbb{R}^N$, $\big(\boldsymbol{\xi}, \mathbf{B}(t)\big)_{\mathbb{R}^N}$ is in the Gaussian family $\mathfrak{G}$ which is obtained by taking the closure of $\mathrm{span}\big(\{X_m : m \ge 0\}\big)$ in $L^2(\mathbb{P}; \mathbb{R})$. Hence, by part (**i**) of Exercise 5.1.6, we will know that $\{\mathbf{B}(t) : t \ge 0\}$ is a Brownian motion once we show that

$$\mathbb{E}^{\mathbb{P}}\big[\big(\boldsymbol{\xi}, \mathbf{B}(s)\big)_{\mathbb{R}^N} \big(\boldsymbol{\xi}', \mathbf{B}(t)\big)_{\mathbb{R}^N}\big] = s \wedge t (\boldsymbol{\xi}, \boldsymbol{\xi}')_{\mathbb{R}^N}.$$

To do this, assume that $s \le t$ and set $\mathbf{h}(\tau) = s \wedge \tau\, \boldsymbol{\xi}$ and $\mathbf{h}'(\tau) = t \wedge \tau\, \boldsymbol{\xi}'$. Then

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}\big[\big(\boldsymbol{\xi}, \mathbf{B}(s)\big)_{\mathbb{R}^N} \big(\boldsymbol{\xi}', \mathbf{B}(t)\big)_{\mathbb{R}^N}\big] &= \lim_{n\to\infty} \sum_{m=1}^{n} \big(\mathbf{h}, \mathbf{h}_m\big)_{\mathbf{H}(\mathbb{R}^N)} \big(\mathbf{h}', \mathbf{h}_m\big)_{\mathbf{H}(\mathbb{R}^N)} \\ &= \big(\mathbf{h}, \mathbf{h}'\big)_{\mathbf{H}(\mathbb{R}^N)} = s(\boldsymbol{\xi}, \boldsymbol{\xi}')_{\mathbb{R}^N}. \quad \square \end{aligned}$$

The following corollary is simply a restatement of the preceding result.

Corollary 5.3.16. *Let $\Gamma = {\gamma_{0,1}}^{\mathbb{N}}$ on $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}_{\mathbb{R}}^{\mathbb{N}})$. Given an orthonormal basis $\{\mathbf{h}_m : m \ge 0\}$ in $\mathbf{H}(\mathbb{R}^N)$, let $\mathcal{N}$ be the set of $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$ such that $\sum_{m=0}^{\infty} x_m \mathbf{h}_m$ fails to converge in $\Theta(\mathbb{R}^N)$. Then $\Gamma(\mathcal{N}) = 0$. Moreover, if $W(t, \mathbf{x}) = \sum_{m=0}^{\infty} x_m \mathbf{h}_m(t)$ for $(t, \mathbf{x}) \in [0, \infty) \times \mathcal{N}\complement$ and $W(t, \mathbf{x}) = \mathbf{0}$ for $(t, \mathbf{x}) \in [0, \infty) \times \mathcal{N}$, then $\mathcal{W}^{(N)} = W_*\Gamma$.*

Corollary 5.3.17. *Let $\{\mathbf{h}_m : m \ge 0\}$ be an orthonormal basis in $\mathbf{H}(\mathbb{R}^N)$, and, for each $m \ge 0$, let $X_m = \mathcal{I}(\mathbf{h}_m)$ be a Paley–Wiener integral of $\mathbf{h}_m$ relative to $\mathcal{W}^{(N)}$. Then, for $\mathcal{W}^{(N)}$-almost every $\boldsymbol{\theta} \in \Theta(\mathbb{R}^N)$, $\sum_{m=0}^{\infty} X_m(\boldsymbol{\theta})\mathbf{h}_m$ converges in $\Theta(\mathbb{R}^N)$ to $\boldsymbol{\theta}$. Moreover, for each $\mathbf{h} \in \mathbf{H}(\mathbb{R}^N)$, $\sum_{m=0}^{\infty} \big(\mathbf{h}, \mathbf{h}_m\big)_{\mathbf{H}(\mathbb{R}^N)} X_m$ converges to $\mathcal{I}(\mathbf{h})$ both $\mathcal{W}^{(N)}$-almost surely and in $L^2(\mathcal{W}^{(N)}; \mathbb{R})$.*

Proof: We begin by verifying the final assertion. To this end, remember that $\{X_m : m \ge 0\}$ is a sequence of independent, standard normal random variables and therefore that, for any $\mathbf{h} \in \mathbf{H}(\mathbb{R}^N)$, Theorem 1.4.2 guarantees

$\sum_{m=0}\big(\mathbf{h},\mathbf{h}\big)_{\mathbf{H}(\mathbb{R}^N)}$ converges both $\mathbb{P}$-almost surely and in $L^2(\mathcal{W}^{(N)};\mathbb{R})$ to a random variable $\mathcal{J}(\mathbf{h})$. Moreover, the map $\mathbf{h}\in\mathbf{H}(\mathbb{R}^N)\longmapsto\mathcal{J}(\mathbf{h})\in L^2(\mathcal{W}^{(N)};\mathbb{R})$ is a linear isometry, and clearly $\mathcal{J}(\mathbf{h}_m)=X_m=\mathcal{I}(\mathbf{h}_m)$ $\mathcal{W}^{(N)}$-almost surely for all $m\ge 0$. Hence, since $\mathbf{h}\rightsquigarrow\mathcal{I}(\mathbf{h})$ is also a linear isometry, it follows that they are equal.

To complete the proof, let $\mathcal{N}$ be the set of $\boldsymbol{\theta}$ for which $\sum_{m=0}^\infty X_m(\boldsymbol{\theta})\mathbf{h}_m$ fails to converge in $\Theta(\mathbb{R}^N)$. By Theorem 5.3.15, $\mathcal{W}^{(N)}(\mathcal{N})=0$. Now set $\mathbf{B}(\,\cdot\,,\boldsymbol{\theta})=\sum_{m=0}^\infty X_m(\boldsymbol{\theta})\mathbf{h}_m$ for $\boldsymbol{\theta}\notin\mathcal{N}$ and $\mathbf{B}(\,\cdot\,,\boldsymbol{\theta})=\mathbf{0}$ for $\theta\in\mathcal{N}$. Since both $t\rightsquigarrow\mathbf{B}(t,\boldsymbol{\theta})$ and $t\rightsquigarrow\boldsymbol{\theta}(t)$ are continuous for each $\boldsymbol{\theta}\in\Theta(\mathbb{R}^N)$, we will know that $\boldsymbol{\theta}=\mathbf{B}(\,\cdot\,,\boldsymbol{\theta})$ for $\mathcal{W}^{(N)}$-almost every $\boldsymbol{\theta}$ once we show that, for each $t\in(0,\infty)$ and $\mathbf{e}\in\mathbb{S}^{N-1}$, $\big(\mathbf{e},\boldsymbol{\theta}(t)\big)_{\mathbb{R}^N}=\big(\mathbf{e},\mathbf{B}(t,\boldsymbol{\theta})\big)_{\mathbb{R}^N}$ for $\mathcal{W}^{(N)}$-almost every $\boldsymbol{\theta}$. To check this, let $t\in(0,\infty)$ and $\mathbf{e}\in\mathbb{S}^{N-1}$ be given, and set $\boldsymbol{\lambda}_{t,\mathbf{e}}=\delta_t\mathbf{e}$. Then

$$\big(\mathbf{e},\boldsymbol{\theta}(t)\big)_{\mathbb{R}^N}=\langle\boldsymbol{\theta},\boldsymbol{\lambda}_{t,\mathbf{e}}\rangle=\big[\mathcal{I}(\mathbf{h}_{\boldsymbol{\lambda}_{t,\mathbf{e}}})\big](\boldsymbol{\theta})=\big[\mathcal{J}(\mathbf{h}_{\boldsymbol{\lambda}_{t,\mathbf{e}}})\big](\boldsymbol{\theta})=\big(\mathbf{e},\mathbf{B}(t,\boldsymbol{\theta})\big)_{\mathbb{R}^N}$$

for $\mathcal{W}^{(N)}$-almost every $\boldsymbol{\theta}$. $\square$

§ 5.3.5. Pinned Brownian Motion. Set

$$\mathfrak{G}\big(\mathbf{H}(\mathbb{R}^N)\big)=\{\mathcal{I}(\mathbf{h}):\,\mathbf{h}\in\mathbf{H}(\mathbb{R}^N)\}. \tag{5.3.18}$$

Then $\mathfrak{G}\big(\mathbf{H}(\mathbb{R}^N)\big)$ is a Gaussian family in $L^2(\mathcal{W}^{(N)};\mathbb{R})$. In fact, because $\mathbf{h}\in\mathbf{H}(\mathbb{R}^N)\rightsquigarrow\mathcal{I}(\mathbf{h})\in l^2(\mathcal{W}^{(N)\prime}\mathbb{R})$ is an isometry, it is a closed Gaussian family.

Given a finite dimensional subspace L of $\mathbf{H}(\mathbb{R}^N)$, let Π_L denote orthogonal projection onto L. In order to define the "adjoint" action Π_L^* on $\Theta(\mathbb{R}^N)$, let $\ell=\dim(L)$, choose an orthonormal basis $\{\mathbf{h}_k:\,1\le k\le\ell\}$ for L, and define $\Pi_L^*:\Theta(\mathbb{R}^N)\longrightarrow\mathbf{H}(\mathbb{R}^N)$ by

$$\Pi_L^*\boldsymbol{\theta}=\sum_{k=1}^{\ell}[\mathcal{I}(\mathbf{h}_k)](\boldsymbol{\theta})\mathbf{h}_k.$$

The following simple lemma explains the sense in which this is a reasonable definition.

Lemma 5.3.19. *Referring to the preceding, $\{\boldsymbol{\theta}(t)-\Pi_L^*\boldsymbol{\theta}(t):\,t\ge 0\}$ is independent of $\{\mathcal{I}(\mathbf{h}):\,\mathbf{h}\in L\}$. Hence, if $\mathbf{H}_L:\,\mathbb{R}^\ell\longrightarrow\mathbf{H}(\mathbb{R}^N)$ is given by $\mathbf{H}_L(\mathbf{y})=\sum_{k=1}^\ell y_k\mathbf{h}_k$, then for any measurable $F:\Theta(\mathbb{R}^N)\times\mathbf{H}(\mathbb{R}^N)\longrightarrow[0,\infty)$,*

$$\begin{aligned}&\int_{\Theta(\mathbb{R}^N)}F\big(\boldsymbol{\theta},\Pi_L^*\boldsymbol{\theta}\big)\,\mathcal{W}N(d\boldsymbol{\theta})\\&\quad=\iint_{\Theta(\mathbb{R}^N)^2}F\big((\boldsymbol{\theta}-\Pi_L^*\boldsymbol{\theta})+\Pi_L^*\boldsymbol{\theta}',\Pi_L^*\boldsymbol{\theta}'\big)\,\mathcal{W}^{(N)}(d\boldsymbol{\theta})\mathcal{W}^{(N)}(d\boldsymbol{\theta}')\\&\quad=\int_{\mathbb{R}^\ell}\mathbb{E}^{\mathbb{P}}\Big[F\big((\boldsymbol{\theta}-\Pi_L^*\boldsymbol{\theta})+\mathbf{H}_L(\mathbf{y}),\mathbf{H}_L(\mathbf{y})\big)\Big]\,\gamma_{\mathbf{0},\mathbf{I}}(d\mathbf{y}).\end{aligned}$$

PROOF: Let $\Pi_L^\perp = \mathbf{I} - \Pi_L$ be orthogonal projection onto the perpendicular complement $L^\perp$ of L. Then, for each $t \in (0,\infty)$ and $\boldsymbol{\xi} \in \mathbb{R}^N$,

$$\big(\boldsymbol{\xi}, \boldsymbol{\theta}(t) - (\Pi_L^*\boldsymbol{\theta})(t)\big)_{\mathbb{R}^N} = \mathcal{I}(\Pi_L^\perp \mathbf{h}_{t,\boldsymbol{\xi}}) \quad \mathcal{W}^{(N)}\text{-almost surely,}$$

where $\mathbf{h}_{t,\boldsymbol{\xi}}(\tau) = \tau \wedge t\,\boldsymbol{\xi}$, and therefore $\big(\boldsymbol{\xi}, \boldsymbol{\theta}(t) - (\Pi_L^*\boldsymbol{\theta})(t)\big)_{\mathbb{R}^N}$ is perpendicular in $L^2(\mathcal{W}^{(N)};\mathbb{R})$ to $\mathcal{I}(\mathbf{h})$ for all $\mathbf{h} \in L$. Hence, the first assertion follows from Lemma 5.1.1. As for the second assertion, the first equality is a simply a restatement of independence, and the second equality follows from the first combined with the observation that the distribution of $\big(\mathcal{I}(\mathbf{h}_1), \dots, \mathcal{I}(\mathbf{h}_\ell)\big)$ under $\mathcal{W}^{(N)}$ is that of an $\mathbb{R}^\ell$-valued, standard normal random variable. $\square$

When we discuss conditional expectations in Chapter VII, we will see that the conclusion drawn in Lemma 5.3.19 is a statment about the conditional distribution under $\mathcal{W}^{(N)}$ of $\boldsymbol{\theta}$ given the sigma algebra $\mathcal{F}_L$ generated by $\{\mathcal{I}(\mathbf{h}) : \mathbf{h} \in L\}$. More precisely, if $\mu_{\mathbf{y}} \in \mathbf{M}_1\big(\Theta(\mathbb{R}^N)\big)$ is the distribution of $\boldsymbol{\theta} \rightsquigarrow \boldsymbol{\theta} - \Pi_L^*\boldsymbol{\theta} + \mathbf{H}_L(\mathbf{y})$ under $\mathcal{W}^{(N)}$, then $\mu_{(\mathcal{I}(\mathbf{h}_1),\dots,\mathcal{I}(\mathbf{h}_\ell))}$ is a regular conditional distribution of $\mathcal{W}^{(N)}$ given $\mathcal{F}_L$. For the present, it suffices to note that if $\mathbf{Y} = \big(\mathcal{I}(\mathbf{h}_1), \dots, \mathcal{I}(\mathbf{h}_\ell)\big)$, then for any $F \in C_{\mathrm{b}}\big(\Theta(\mathbb{R}^N);\mathbb{R}\big)$ and $\mathbf{y} \in \mathbb{R}^N$, then Lemma 5.3.19 implies that

$$\lim_{r \searrow 0} \frac{\mathbb{E}^{\mathcal{W}^{(N)}}\big[F,\, |\mathbf{Y} - \mathbf{y}| < r\big]}{\mathcal{W}^{(N)}(|\mathbf{Y} - \mathbf{y}| < r)} = \mathbb{E}^{\mu_{\mathbf{y}}}[F].$$

A particularly interesting case of the preceding is the one when L is the nN-dimensional space spanned by $\{h_{t_{m-1},t_m}\boldsymbol{\xi} : 1 \le m \le n \ \&\ \boldsymbol{\xi} \in \mathbb{R}^N\}$, where $0 = t_0 < \cdots t_n$ and $h_{s,t}(\tau) \equiv (\tau - s)^+ \wedge (t-s)$ for $s < t$. In this case,

$$\Pi_L^*\boldsymbol{\theta} = \sum_{m=1}^n \frac{h_{t_{m-1},t_m}}{t_m - t_{m-1}}\big(\boldsymbol{\theta}(t_m) - \boldsymbol{\theta}(t_{m-1})\big),$$

and so

$$\begin{aligned} &\boldsymbol{\theta}_{(t_1,\dots,t_n)}(\tau) \equiv [\boldsymbol{\theta} - \Pi_L^*\boldsymbol{\theta}](\tau) \\ &\quad = \begin{cases} \boldsymbol{\theta}(t) - \boldsymbol{\theta}(t_{m-1}) - \frac{t - t_{m-1}}{t_m - t_{m-1}}\big(\boldsymbol{\theta}(t_m) - \boldsymbol{\theta}(t_{m-1})\big) & \text{if } t \in [t_{m-1}, t_m] \\ \boldsymbol{\theta}(t) - \boldsymbol{\theta}(t_n) & \text{if } t \in [t_n, \infty). \end{cases} \end{aligned} \tag{5.3.20}$$

Thus, if $\mathbf{H}_{(t_1,\dots,t_n)} : (\mathbb{R}^N)^n \longrightarrow \mathbf{H}(\mathbb{R}^N)$ is given by

$$\mathbf{H}_{(t_1,\dots,t_n)}(\vec{\mathbf{y}}) = \sum_{m=1}^n \frac{h_{t_{m-1},t_m}}{t_m - t_{m-1}}(\mathbf{y}_m - \mathbf{y}_{m-1}) \text{ where } \vec{\mathbf{y}} = (\mathbf{y}_1, \dots, \mathbf{y}_n) \text{ and } \mathbf{y}_0 \equiv \mathbf{0},$$

then, for any Borel measurable $F : \Theta(\mathbb{R}^N) \times (\mathbb{R}^N)^n \longrightarrow [0, \infty)$,

$$
\begin{aligned}
&\int_{\Theta(\mathbb{R}^N)} F\Big(\boldsymbol{\theta}, \big(\boldsymbol{\theta}(t_1), \dots, \boldsymbol{\theta}(t_n)\big)\Big)\, \mathcal{W}^{(N)}(d\boldsymbol{\theta}) \\
&= \int_{(\mathbb{R}^N)^n} \bigg(\int_{\Theta(\mathbb{R}^N)} F\big(\boldsymbol{\theta} + \mathbf{H}_{(t_1,\dots,t_n)}\vec{\mathbf{y}}, \vec{\mathbf{y}}\big)\, \mathcal{W}^{(N)}(d\boldsymbol{\theta})\bigg) \mathcal{W}^{(N)}_{(t_1,\dots,t_n)}(d\vec{\mathbf{y}}),
\end{aligned}
\tag{5.3.21}
$$

where $\mathcal{W}^{(N)}_{(t_1,\dots,t_n)}$ is the $\mathcal{W}^{(N)}$-distribution of $\boldsymbol{\theta} \rightsquigarrow (\boldsymbol{\theta}(t_1), \dots, \boldsymbol{\theta}(t_n))$. Equivalently, if

$$
\bar{\mathbf{H}}_{(t_1,\dots,t_n)}(\mathbf{y}_1, \dots, \mathbf{y}_m) = \sum_{m=1}^{n} \frac{h_{t_{m-1},t_m}}{t_m - t_{m-1}} \mathbf{y}_m,
$$

then

$$
\begin{aligned}
&\int_{\Theta(\mathbb{R}^N)} F\Big(\boldsymbol{\theta}, \big(\boldsymbol{\theta}(t_1) - \boldsymbol{\theta}(t_0), \dots, \boldsymbol{\theta}(t_n) - \boldsymbol{\theta}(t_{n-1})\big)\Big)\, \mathcal{W}^{(N)}(d\boldsymbol{\theta}) \\
&= \int_{(\mathbb{R}^N)^n} \bigg(\int_{\Theta(\mathbb{R}^N)} F\Big(\boldsymbol{\theta} + \bar{\mathbf{H}}_{(t_1,\dots,t_n)}(\vec{\mathbf{y}}), \vec{\mathbf{y}}\Big)\, \mathcal{W}^{(N)}(d\boldsymbol{\theta})\bigg) \gamma_{\mathbf{0},\mathbf{C}(t_1,\dots,t_n)}(d\vec{\mathbf{y}}),
\end{aligned}
\tag{5.3.22}
$$

where $\mathbf{C}(t_1, \dots, t_n)$ is the diagonal matrix whose mth diagonal entry is $t_m - t_{m-1}$.

From the (5.3.21), we see that, for any $F \in C_{\mathrm{b}}\big(\Theta(\mathbb{R}^N)\big)$ and $(\mathbf{y}_1, \dots, \mathbf{y}_n) \in (\mathbb{R}^N)^n$,

$$
\begin{aligned}
&\lim_{r \searrow 0} \frac{\mathbb{E}^{\mathcal{W}^{(N)}}\big[F(\boldsymbol{\theta}),\, |\boldsymbol{\theta}(t_m) - \mathbf{y}_m| \le r,\, 1 \le m \le n\big]}{\mathcal{W}^{(N)}\big(|\boldsymbol{\theta}(t_m) - \mathbf{y}_m| \le r,\, 1 \le m \le n\big)} \\
&\qquad = \mathbb{E}^{\mathcal{W}^{(N)}}\Big[F\big(\boldsymbol{\theta} + \mathbf{H}_{(t_1,\dots,t_n)}(\mathbf{y}_1, \dots, \mathbf{y}_n)\big)\Big],
\end{aligned}
$$

which, in the language of conditional probability, is a strong statement that the distribution of $\boldsymbol{\theta} \rightsquigarrow \boldsymbol{\theta}_{(t_1,\dots,t_n)} + \mathbf{H}(\mathbf{y}_1, \dots, \mathbf{y}_n)$ under $\mathcal{W}^{(N)}$ is *the conditional distribution of $\boldsymbol{\theta}$ under $\mathcal{W}^{(N)}$ given that $\boldsymbol{\theta}(t_m) = \mathbf{y}_m$ for $1 \le m \le n$.* For this reason, $\boldsymbol{\theta}_{(t_1,\dots,t_n)} + \mathbf{H}(\mathbf{y}_1, \dots, \mathbf{y}_m)$ is called a **pinned Brownian motion** : it is *pinned* to the point $\mathbf{y}_m$ at time t_m for each $1 \le m \le n$.

§ 5.3.6. The Cameron–Martin Formula. If one believes that there is truth hidden in (5.3.2), then one should believe that it makes correct predictions when those predictions make *sense*. In this subsection, we will give an example of such a prediction and show that it is correct. Namely, given $\mathbf{h} \in \mathbf{H}(\mathbb{R}^N)$, consider the translation map $\mathbf{h}' \in \mathbf{H}(\mathbb{R}^N) \longmapsto T_{\mathbf{h}}(\mathbf{h}') \equiv \mathbf{h} + \mathbf{h}' \in \mathbf{H}(\mathbb{R}^N)$. Then (5.3.2) predicts that

$$
\big((T_{\mathbf{h}})_* \mathcal{W}^{(N)}\big)(d\mathbf{h}') = R_{\mathbf{h}}(\mathbf{h}')\, \mathcal{W}^{(N)}(d\mathbf{h}') \text{ where } R_{\mathbf{h}}(\mathbf{h}') = e^{(\mathbf{h},\mathbf{h}')_{\mathbf{H}(\mathbb{R}^N)} - \frac{1}{2}\|\mathbf{h}\|^2_{\mathbf{H}(\mathbb{R}^N)}}.
$$

Of course, since $\mathcal{W}^{(N)}$ lives on $\Theta(\mathbb{R}^N)$, not $\mathbf{H}(\mathbb{R}^N)$, one wants to extend $T_{\mathbf{h}}$ to $\Theta(\mathbb{R}^N)$ so that $T_{\mathbf{h}}\boldsymbol{\theta} = \mathbf{h} + \boldsymbol{\theta}$. In addition, one has to worry about the meaning of the $\mathbf{H}(\mathbb{R}^N)$ inner product appearing in $R_{\mathbf{h}}$. However, we have already seen that the Paley–Wiener map $\mathcal{I}(\mathbf{h})$ provides a rational way of interpreting $\boldsymbol{\theta} \rightsquigarrow (\boldsymbol{\theta}, \mathbf{h})_{\mathbf{H}(\mathbb{R}^N)}$, and so we are led to guess that

$$\text{(5.3.23)}\quad \big((T_{\mathbf{h}})_*\mathcal{W}^{(N)}\big)(d\boldsymbol{\theta}) = R_{\mathbf{h}}(\boldsymbol{\theta})\,\mathcal{W}^{(N)}(d\boldsymbol{\theta}) \text{ where } R_{\mathbf{h}}(\boldsymbol{\theta}) = e^{\mathcal{I}(\boldsymbol{\theta}) - \frac{1}{2}\|\mathbf{h}\|^2_{\mathbf{H}(\mathbb{R}^N)}}.$$

That (5.3.23) is correct was discovered by R. Cameron and T. Martin, and is therefore known as the **Cameron–Martin formula**. Moreover, given Lemma 5.3.19, its proof is easy. Indeed, when looked at from the point of view of that lemma, it is really just the formala for the Radon–Nikodym derivate of $\gamma_{a,1}$ with respect to $\gamma_{0,1}$ with $a = \|\mathbf{h}\|_{\mathbf{H}(\mathbb{R}^N)}$. To see this, let $\mathbf{h} \in \mathbf{H}(\mathbb{R}^N)\setminus\{\mathbf{0}\}$ be given, and set $\bar{\mathbf{h}} = \frac{\mathbf{h}}{\|boldh\|}$. If $L = \operatorname{span}(\mathbf{h})$, then $\Pi_L^*\boldsymbol{\theta} = [\mathcal{I}(\bar{\mathbf{h}})](\boldsymbol{\theta})\bar{\mathbf{h}}$, and so, by (5.3.20),

$$\begin{aligned}
\mathbb{E}^{\mathcal{W}^{(N)}}\big[F\circ T_{\mathbf{h}}\big] &= \int_{\mathbb{R}} \mathbb{E}^{\mathcal{W}^{(N)}}\Big[F\big((\boldsymbol{\theta} - \Pi_L^*\boldsymbol{\theta}) + (y + \|\mathbf{h}\|_{\mathbf{H}(\mathbb{R}^N)})\bar{\mathbf{h}}\big)\Big]\,\gamma_{0,1}(dy)\\
&= \mathbb{E}^{\mathcal{W}^{(N)}}\Big[e^{\|\mathbf{h}\|_{\mathbf{H}(\mathbb{R}^N)}y - \frac{1}{2}\|\mathbf{h}\|^2_{\mathbf{H}(\mathbb{R}^N)}} F\big((\boldsymbol{\theta} - \Pi_L^*\boldsymbol{\theta}) + y\bar{\mathbf{h}}\big)\Big]\,\gamma_{0,1}(dy)\\
&= \mathbb{E}^{\mathcal{W}^{(N)}}\Big[e^{\|\mathbf{h}\|_{\mathbf{H}(\mathbb{R}^N)}\mathcal{I}(\bar{\mathbf{h}}) - \frac{\|\mathbf{h}\|^2_{\mathbf{H}(\mathbb{R}^N)}}{2}} F\Big] = \mathbb{E}^{\mathcal{W}^{(N)}}\big[R_{\mathbf{h}}F\big]
\end{aligned}$$

for any Borel measurable $F : \Theta(\mathbb{R}^N) \longrightarrow [0,\infty)$. In addition, noting that $\mathbb{E}\mathcal{W}^{(N)}\big[R_{\mathbf{h}}^p\big]^{\frac{1}{p}} = e^{\frac{p-1}{2}\|\mathbf{h}\|^2_{\mathbf{H}(\mathbb{R}^N)}}$ for any $p \in (0,\infty)$, we can use Hölder's inequality to show that

$$\text{(5.3.24)}\qquad \begin{aligned}
&e^{-\frac{p}{2}\|\mathbf{h}\|^2_{\mathbf{H}(\mathbb{R}^N)}}\mathbb{E}^{\mathcal{W}^{(N)}}\Big[F^{\frac{p-1}{p}}\Big]^{\frac{p}{p-1}}\\
&\qquad \le \mathbb{E}^{\mathcal{W}^{(N)}}\big[F\circ T_{\mathbf{h}}\big] \le e^{\frac{p-1}{2}\|\mathbf{h}\|_{\mathbf{H}(\mathbb{R}^N)^2}}\mathbb{E}^{\mathcal{W}^{(N)}}\Big[F^{\frac{p}{p-1}}\Big]^{\frac{p-1}{p}}
\end{aligned}$$

for any such F.

It should be remarked that the argument just given works equally well for any abstract Wiener space. In addition, and more interesting, we will show in Exercise (?) that translates of $\mathcal{W}^{(N)}$ in directions other than those from $\mathbf{H}(\mathbb{R}^N)$ are singular to $\mathcal{W}^{(N)}$. Because of their contribution in revealing the fundamental role that it plays, the space $\mathbf{H}(\mathbb{R}^N)$ is sometimes called the **Cameron–Martin subspace** for Wiener measure.

§ 5.3.7. The Support of Wiener Measure. We close this section by showing that, in a rather precise sense, Wiener paths fill up $\Theta(\mathbb{R}^N)$.

THEOREM 5.3.25. *The support of $\mathcal{W}^{(N)}$ is the whole of $\Theta(\mathbb{R}^N)$. That is, for each $\boldsymbol{\theta} \in \Theta(\mathbb{R}^N)$ and $r > 0$, $\mathcal{W}^{(N)}$ assigns the ball $B_{\Theta(\mathbb{R}^N)}(\boldsymbol{\theta}, r)$ in $\Theta(\mathbb{R}^N)$ around $\boldsymbol{\theta}$ of radius r positive probability.*

PROOF: We begin by making several reductions. In the first place, because $\mathbf{H}(\mathbb{R}^N)$ is dense in $\Theta(\mathbb{R}^N)$ we need only show that $\mathcal{W}^{(N)}\big(B_{\Theta(\mathbb{R}^N)}(\mathbf{h},r)\big)>0$ for $\mathbf{h}\in\mathbf{H}(\mathbb{R}^N)$. Furthermore, by (5.3.24) with $p=2$, we know that

$$\mathcal{W}^{(N)}\big(B_{\Theta(\mathbb{R}^N)}(\mathbf{h},r)\big)\geq e^{-\|\mathbf{h}\|^2_{\mathbf{H}(\mathbb{R}^N)}}\mathcal{W}^{(N)}\big(B_{\Theta(\mathbb{R}^N)}(\mathbf{0},r)\big)^2,$$

and therefore that it suffices to handle the case when $\mathbf{h}=\mathbf{0}$.

Next, observe that, for any $T\in[1,\infty)$,

$$\begin{aligned}\mathcal{W}^{(N)}\big(B_{\Theta(\mathbb{R}^N)}(\mathbf{0},r)\big)&\geq\mathcal{W}^{(N)}\left(\|\boldsymbol{\theta}\|_{[0,T]}<\frac{r}{3}\ \&\ \sup_{t\geq T}\frac{|\boldsymbol{\theta}(t)-\boldsymbol{\theta}(T)|}{t}<\frac{r}{3}\right)\\&=\mathcal{W}^{(N)}\left(\|\boldsymbol{\theta}\|_{[0,T]}<\frac{r}{3}\right)\mathcal{W}^{(N)}\left(\sup_{t\geq T}\frac{|\boldsymbol{\theta}(t)-\boldsymbol{\theta}(T)|}{t}<\frac{r}{3}\right),\end{aligned}$$

and, by Brownian scaling $\mathcal{W}^{(N)}\big(\|\boldsymbol{\theta}\|_{[0,T]}<\frac{r}{3}\big)=\mathcal{W}^{(N)}\big(\|\boldsymbol{\theta}\|_{[0,1]}<\frac{r}{3}T^{-\frac12}\big)$ and

$$\begin{aligned}\mathcal{W}^{(N)}\left(\sup_{t\geq T}\frac{|\boldsymbol{\theta}(t)-\boldsymbol{\theta}(T)|}{t}<\frac{r}{3}\right)&=\mathcal{W}^{(N)}\left(\sup_{t\geq 1}\frac{|\boldsymbol{\theta}(t)-\boldsymbol{\theta}(1)|}{t}<\frac{rT^{\frac12}}{3}\right)\\&=\mathcal{W}^{(N)}\left(\|\boldsymbol{\theta}\|_{\Theta(\mathbb{R}^N)}<\frac{rT^{\frac12}}{3}\right)\longrightarrow 1\text{ as }T\to\infty.\end{aligned}$$

Hence, it is enough to show that $\mathcal{W}^{(N)}\big(\|\boldsymbol{\theta}\|_{[0,1]}<r\big)>0$ for all $r>0$.

To finish the proof, refer to (5.3.22), and, for each $n\in\mathbb{Z}^+$, take $\boldsymbol{\theta}^{(n)}=\boldsymbol{\theta}_{(\frac1n,\frac2n,\dots,1)}$, $[\bar{\mathbf{H}}^{(n)}(\vec{\mathbf{y}})](t)=n\sum_{m=1}^n\big(t-\frac{m-1}{n}\big)^+\wedge\frac1n\mathbf{y}_m$. Then, by (5.3.22),

$$\begin{aligned}\mathcal{W}^{(N)}\big(\|\boldsymbol{\theta}\|_{[0,1]}<r\big)&=\int_{(\mathbb{R}^N)^n}\mathcal{W}^{(N)}\big(\|\boldsymbol{\theta}^{(n)}+\bar{\mathbf{H}}^{(n)}(\vec{\mathbf{y}})\|_{[0,1]}<r\big)\,\gamma_{\mathbf{0},\frac1n\mathbf{I}}(d\vec{\mathbf{y}})\\&\geq\mathcal{W}^{(N)}\big(\|\boldsymbol{\theta}^{(n)}\|_{[0,1]}<\tfrac{r}{2}\big)\gamma_{\mathbf{0},\frac1n\mathbf{I}}\big(\|\bar{\mathbf{H}}^{(n)}\|_{[0,1]}<\tfrac{r}{2}\big).\end{aligned}$$

Since $\|\bar{\mathbf{H}}^{(n)}(\vec{\mathbf{y}})\|_{[0,1]}\leq n^{\frac12}|\vec{\mathbf{y}}|$,

$$\gamma_{\mathbf{0},\frac1n\mathbf{I}}\big(\|\bar{\mathbf{H}}^{(n)}\|_{[0,1]}<r\big)\geq\gamma_{\mathbf{0},\mathbf{I}}\big(B_{(\mathbb{R}^N)^n}(\mathbf{0},r)\big)>0$$

for all $n\in\mathbb{Z}^+$ and $r>0$, we are left with proving that, for each $r>0$ there is an $n\in\mathbb{Z}^+$ such that $\mathcal{W}^{(N)}\big(\|\boldsymbol{\theta}^{(n)}\|_{[0,1]}<r\big)>0$. To this end, use (5.3.20) to check that the processes $\big\{\{\boldsymbol{\theta}^{(n)}(t):\frac{m-1}{n}\leq t\leq\frac{m}{n}\}:1\leq m\leq n\big\}$ are independent and identically distributed, and therefore

$$\mathcal{W}^{(N)}\big(\|\boldsymbol{\theta}^{(n)}\|_{[0,1]}<r\big)=\mathcal{W}^{(N)}\big(\|\boldsymbol{\theta}^{(n)}\|_{[0,\frac1n]}<r\big)^n.$$

Finally, since $\|\boldsymbol{\theta}^{(n)}\|_{[0,\frac{1}{n}]} \le 2\|\boldsymbol{\theta}\|_{[0,\frac{1}{n}]}$, we see that $\mathcal{W}^{(N)}\big(\|\boldsymbol{\theta}^{(n)}\|_{[0,\frac{1}{n}]} < r\big) \longrightarrow 1$ as $n \to \infty$. $\square$

§ 5.3.8. The Ornstein–Uhlenbeck Process. Given $\mathbf{x} \in \mathbb{R}^N$ and $\boldsymbol{\theta} \in \Theta(\mathbb{R}^N)$, consider the integral equation

$$\mathbf{U}(t,x,\boldsymbol{\theta}) = \mathbf{x} + \boldsymbol{\theta}(t) - \frac{1}{2}\int_0^t \mathbf{U}(\tau,\mathbf{x},\boldsymbol{\theta})\,d\tau, \quad t \ge 0. \tag{5.3.26}$$

A completely elementary argument shows that, for each $\mathbf{x}$ and $\boldsymbol{\theta}$, there is at most one solution. Futhermore, integration by parts allows one to check that if

$$\mathbf{U}_0(t,\boldsymbol{\theta}) = \int_0^t e^{\frac{\tau}{2}}\,d\boldsymbol{\theta}(\tau),$$

where the integral is taken in the sense of Riemann-Stieltjes, then

$$\mathbf{U}(t,\mathbf{x},\boldsymbol{\theta}) = e^{-\frac{t}{2}}\big(\mathbf{x} + \mathbf{U}_0(t,\boldsymbol{\theta})\big)$$

is one, and therefore the only, solution.

The stocastic process $\{\mathbf{U}(t,\mathbf{x}) :\ge 0\}$ under $\mathcal{W}^{(N)}$ was introduced by L. Ornstein and G. Uhlenbeck* is known as the **Ornstein–Uhlenbeck process**. From our immediate point of view, its importance is that it provides another interesting process which lends itself to detailed computations.

From a physical standpoint, $\mathbf{U}(t,\mathbf{0},\boldsymbol{\theta})$ as a Brownian motion which is subjected to a linear restoring force. Thus, locally it should behave very much like a Brownian motion. However, over long time intervals, it should feel the effect of the restoring force, which is always pushing it back toward the origin. To see how these physical ideas are reflected in the distribution of $\{\mathbf{U}(t,\mathbf{0},\boldsymbol{\theta}) : t \ge 0\}$, we begin by noting that, for each $\mathbf{e} \in \mathbb{S}^{N-1}$,

$$\begin{gathered}\big(\mathbf{e},\mathbf{U}_0(t)\big)_{\mathbb{R}^N} = \langle \boldsymbol{\theta}, \boldsymbol{\lambda}^t_{\mathbf{e}}\rangle = [\mathcal{I}(\mathbf{h}^t_{\mathbf{e}})](\boldsymbol{\theta}) \text{ where}\\ \boldsymbol{\lambda}^t_{\mathbf{e}}(d\tau) = \Big(e^{\frac{t}{2}}\delta_t(d\tau) - \tfrac{1}{2}\mathbf{1}_{[0,t)}(\tau)e^{\frac{\tau}{2}}\,d\tau\Big)\mathbf{e} \text{ and } \mathbf{h}^t_{\mathbf{e}}(\tau) = 2\Big(e^{\frac{\tau\wedge t}{2}} - 1\Big)\mathbf{e}.\end{gathered}$$

Hence, the span of $\big\{\big(\boldsymbol{\xi},\mathbf{U}(t,\mathbf{0})\big)_{\mathbb{R}^N} : t \ge 0 \;\&\; \boldsymbol{\xi} \in \mathbb{R}^N\big\}$ is a Gaussian family in $L^2(\mathcal{W}^{(N)};\mathbb{R})$, and

$$\mathbb{E}^{\mathcal{W}^{(N)}}\big[\mathbf{U}(s,\mathbf{0}) \otimes \mathbf{U}(t,\mathbf{0})\big] = \big(e^{-\frac{|t'-t|}{2}} - e^{-\frac{t'+t}{2}}\big)\mathbf{I}.$$

The key to understanding this process is the observation that it has the same distribution as the process $\big\{e^{-\frac{t}{2}}\mathbf{B}\big(e^t - 1\big) : t \ge 0\big\}$, where $\{\mathbf{B}(t) : t \ge 0\}$ is

---

* In their article "On the theory of Brownian motion," *Phys. Reviews* **36** (3), L. Ornstein & G. Uhlenbeck introduced this process in an attempt to reconcile some of the more disturbing properties of Wiener paths with physical reality.

a Brownian motion. In particular, by combining this with law of the iterated logarithm proved in Exercise 5.1.11, we see that, for each $\mathbf{e} \in \mathbb{S}^{N-1}$,

$$\varlimsup_{t\to\infty} \frac{\big(\mathbf{e}, \mathbf{U}(t,\mathbf{x})\big)_{\mathbb{R}^N}}{\sqrt{2\log t}} = 1 = -\varliminf_{t\to\infty} \frac{\big(\mathbf{U}(t,\mathbf{x})\big)_{\mathbb{R}^N}}{\sqrt{2\log t}} \tag{5.3.27}$$

$\mathcal{W}^{(N)}$-almost surely, which confirms the suspicion that the restoring force dampens the Brownian excursions out toward infinity.

A second indication of that $\mathbf{U}(\,\cdot\,,\mathbf{x})$ tends to spend more time than Brownian motion near the origin is that its distribution at time t will be $\gamma_{e^{-\frac{t}{2}}\mathbf{x},(1-e^{-t})\mathbf{I}}$, and so, as distinguished from Brownian motion itself, its distribution at time t tends to a limit, namely $\gamma_{\mathbf{0},\mathbf{I}}$, as $t \to \infty$. This observation suggests that it might be interesting to look at an *ancient* Ornstein–Uhlenbeck process, one that has been already going for an infinite amount of time. To be more precise, since the distribution of an ancient Ornstein–Uhlenbeck at time 0 would be $\gamma_{\mathbf{0},\mathbf{I}}$, what we should look at is the process which we get by making the $\mathbf{x}$ in $\mathbf{U}(\,\cdot\,,\mathbf{x},\boldsymbol{\theta})$ a standard normal random variable. Thus, we will call the process $\{\mathbf{U}(\,\cdot\,) : t \ge 0\}$ under $\gamma_{\mathbf{0},\mathbf{I}} \times \mathcal{W}^{(N)}$ an **ancient Ornstein–Uhlenbeck process**. Clearly, the distribution $\mathcal{U}^{(N)}$ of an ancient Ornstein–Uhlenbeck process is that same as that of $\big\{e^{-\frac{t}{2}}\mathbf{X}_0 + e^{-\frac{t}{2}}\mathbf{B}\big(e^t - 1\big) : t \ge 0\big\}$, where $\mathbf{X}_0$ is a standard normal random variable which is independent of the Brownian motion $\{\mathbf{B}(t) : t \ge 0\}$. Again this process spans a Gaussian family. In addition,

$$\mathbb{E}^{\gamma_{\mathbf{0},\mathbf{I}}\times\mathcal{W}^{(N)}}\big[\mathbf{U}(s)\otimes\mathbf{U}(t)\big] = e^{-\frac{|t-s|}{2}}\mathbf{I}.$$

Thus, as we suspected, the ancient Orstein–Uhlenbeck process is a **stationary process** in the sense that, for each $T > 0$, the distribution of $\{\mathbf{U}(t+T) : t \ge 0\}$ is the same as that of $\{\mathbf{U}(t) : t \ge 0\}$.

In fact, even more is true: it is time reversible in the sense that $\{\mathbf{U}(t) : t \in [0,T]\}$ has the same distribution as $\{\mathbf{U}(T-t) : t \in [0,T]\}$. This observation suggests that we can give the ancient Ornstein-Uhlenbeck its past by running it backwards. That is, consider $\mathbf{U}_R : [0,\infty) \times \mathbb{R}^N \times \Theta(\mathbb{R}^N)^2 \longrightarrow \mathbb{R}^N$ given by

$$\mathbf{U}_R(t,\mathbf{x},\boldsymbol{\theta}_+,\boldsymbol{\theta}_-) = \begin{cases} \mathbf{U}(t,\mathbf{x},\boldsymbol{\theta}_+) & \text{if } t \ge 0 \\ \mathbf{U}(-t,\mathbf{x},\boldsymbol{\theta}_-) & \text{if } t < 0, \end{cases}$$

and consider the process $\{\mathbf{U}_R(t) : t \in \mathbb{R}\}$ under $\gamma_{\mathbf{0},\mathbf{I}} \times \mathcal{W}^{(N)} \times \mathcal{W}^{(N)}$. This process also spans a Gaussian family, and it is still true that

$$\mathbb{E}^{\gamma_{\mathbf{0},\mathbf{I}}\times\mathcal{W}^{(N)}\times\mathcal{W}^{(N)}}\big[\mathbf{U}_S(s)\otimes\mathbf{U}_R(t)\big] = u(s,t)\mathbf{I}, \text{ where } u(s,t) \equiv e^{-\frac{|t-s|}{2}}, \tag{5.3.28}$$

only now for all $s,\,t \in \mathbb{R}$. The advantage of having added the past is that the statement of reversibility takes more appealing form. Namely, $\{\mathbf{U}_R(t) : t \in \mathbb{R}\}$ is **reversible** in the sense that its distribution is the same whether one runs it forward or backward in time. That is, $\{\mathbf{U}_R(-t) : t \in \mathbb{R}\}$ has the same distribution as $\{\mathbf{U}_R(t) : t \in \mathbb{R}\}$. For this reason, we will call it the **reversible Ornstein–Uhlenbeck process**, although it is usually called the stationary Ornstein–Uhlenbeck process.

Exercises for § 5.3

EXERCISE 5.3.29. Let H be a separable Hilbert space, and, for each $n \in \mathbb{Z}^+$ and subset $\{g_1, \dots, g_n\} \subseteq H$, let $\mathcal{A}(g_1, \dots, g_n)$ denote the σ-algebra over H generated by the mapping

$$h \in H \longmapsto \big((h, g_1)_H, \dots, (h, g_n)_H\big) \in \mathbb{R}^n,$$

and check that

$$\mathcal{A} = \bigcup \big\{\mathcal{A}(g_1, \dots, g_n) : n \in \mathbb{Z}^+ \text{ and } g_1, \dots, g_n \in H\big\}$$

is an algebra which generates $\mathcal{B}_H$. Show that there *always* exists a *finitely additive* $\mathcal{W}_H$ on $\mathcal{A}$ which is uniquely determined by the properties that it is σ-additive on $\mathcal{A}(g_1, \dots, g_n)$ for every $n \in \mathbb{Z}^+$ and $\{g_1, \dots, g_n\} \subseteq H$ and

$$\int_H \exp\Big[\sqrt{-1}\,(h, g)_H\Big]\,\mathcal{W}_H(dh) = \exp\left[-\frac{\|g\|_H^2}{2}\right], \quad g \in H.$$

On the other hand, as already know this finitely additive measure admits a countably additive extension to $\mathcal{B}_H$ if and only if H is finite dimensional.

EXERCISE 5.3.30. Given $\boldsymbol{\lambda} \in \Lambda(\mathbb{R}^N)$, we pointed out (cf. (5.3.13)) that the Paley–Wiener integral $[\mathcal{I}(\mathbf{h}_{\boldsymbol{\lambda}})](\boldsymbol{\theta})$ can be interpreted as the Riemann–Stieltjes integral of $\boldsymbol{\lambda}\big((s, \infty)\big)$ with respect to $\boldsymbol{\theta}(s)$. In this exercise, we will use this observation as the starting point for what is called **stochastic integration**.

(**i**) Given $\boldsymbol{\lambda} \in \Lambda(\mathbb{R}^N)$ and $t > 0$, set $\boldsymbol{\lambda}^t(d\tau) = \mathbf{1}_{[0,t)}(\tau)\boldsymbol{\lambda}(d\tau) + \delta_t \boldsymbol{\lambda}\big([t, \infty)\big)$, and show that

$$[\mathcal{I}(\boldsymbol{\lambda}^t)](\boldsymbol{\theta}) = \int_0^t \boldsymbol{\lambda}\big((\tau, \infty)\big) \cdot d\boldsymbol{\theta}(\tau),$$

where the integral on the right is taken in the sense of Riemann–Stieltjes. In particular, conclude that $t \rightsquigarrow [\mathcal{I}(\boldsymbol{\lambda}^t)](\boldsymbol{\theta})$ is continuous.

(**ii**) Given $\mathbf{f} \in C^1_c\big([0, \infty); \mathbb{R}^N\big)$, set $\boldsymbol{\lambda}^t_{\mathbf{f}}(d\tau) = \mathbf{1}_{[t,\infty)}(\tau)\dot{\mathbf{f}}(\tau)\,d\tau$, and show that

$$[\mathcal{I}(\boldsymbol{\lambda}^t_{\mathbf{f}})](\boldsymbol{\theta}) = \int_0^t \mathbf{f}(\tau) \cdot d\boldsymbol{\theta}(\tau),$$

where again the integral on the right is Riemann–Stieltjes. Use this to see that the process

$$\left\{\int_0^t \mathbf{f}(\tau) \cdot d\boldsymbol{\theta}(\tau) : t \geq 0\right\}$$

has the same distribution under $\mathcal{W}^{(N)}$ as

$$(*) \qquad \left\{B\left(\int_0^t |\mathbf{f}(\tau)|^2\,d\tau\right) : t \geq 0\right\},$$

where $\{B(t) : t \geq 0\}$ is an $\mathbb{R}$-valued Brownian motion.

(iii) Given $\mathbf{f} \in L^2_{\text{loc}}\big([0,\infty);\mathbb{R}^N\big)$ and $t > 0$, set $\mathbf{h}^t_{\mathbf{f}}(\tau) = \int_0^{t\wedge\tau} \mathbf{f}(s)\,ds$. Show that the $\mathcal{W}^{(N)}$-distribution of the process

$$\big\{\mathcal{I}(\mathbf{h}^t) : t \geq 0\big\}$$

is the same as that in (*). In particular, conclude (cf. part (**ii**) of Exercise 5.1.12) that there is a continuous modification of the process $\{\mathcal{I}(\mathbf{h}^t_{\mathbf{f}}) : t \geq 0\}$. For reasons made clear in (**ii**), such a continuous modification is denoted by

$$\left\{\int_0^t \mathbf{f}(\tau)\cdot d\boldsymbol{\theta}(\tau) : t \geq 0\right\}.$$

Of course, unless $\mathbf{f}$ has bounded variation, the integrals in the preceding is no longer interpretable a Riemann-Stieltjes integral. In fact, it is not even defined a path by path but only as a stochastic process. For this reason, it is called a **stochastic integral**.

EXERCISE 5.3.31. Let $N = 1$. Referring to § 5.3.4 and using Corollary 5.3.17, take Wiener's choice of orthonormal basis and check that there are independent, standard normal random variables $\{X_m : m \geq 1\}$ under $\mathcal{W}^{(1)}$, such that, for $\mathcal{W}^{(1)}$-almost almost every θ,

$$\theta(t) = tX_0(\theta) + 2^{\frac{1}{2}} \sum_{m=1}^{\infty} X_m(\theta)\frac{\sin(\pi m t)}{m\pi}, \quad t \in [0,1],$$

where the convergence is uniform. From this, conclude that, $\mathcal{W}^{(1)}$-almost surely,

$$\int_0^1 \theta(t)^2\,dt = \frac{X_0(\theta)^2}{3} + \frac{1}{\pi^2}\sum_{m=1}^{\infty} \frac{X_m(\theta)^2 + \sqrt{8}X_0(\theta)X_m(\theta)}{n^2},$$

where the convergence of the series is absolute. Using the preceding, conclude that, for any $\alpha \in (0,\infty)$,

$$\begin{aligned} &\mathbb{E}^{\mathcal{W}^{(1)}}\left[-\alpha\int_0^1 \theta(t)^2\,dt\right] \\ &\qquad = \left[\prod_{m=1}^{\infty}\left(1 + \frac{2\alpha}{m^2\pi^2}\right)\right]^{-\frac{1}{2}} \left[1 + 4\alpha\sum_{m=1}^{\infty}\frac{1}{m^2\pi^2 + 2\alpha}\right]^{-\frac{1}{2}}. \end{aligned}$$

Finally, recall Euler's product formula

$$\sinh z = \prod_{m=1}^{\infty}\left(1 + \frac{z^2}{m^2\pi^2}\right), \quad z \in \mathbb{C},$$

and arrive first at

$$\mathbb{E}^{\mathcal{W}^{(1)}}\left[\exp\left(-\alpha\int_0^1 \theta(t)^2\,dt\right)\right] = \left[\cosh\sqrt{2\alpha}\right]^{-\frac{1}{2}}$$

and then, after rescaling, at

$$\mathbb{E}^{\mathcal{W}^{(1)}}\left[\exp\left(-\alpha\int_0^T \theta(t)^2\,dt\right)\right] = \left[\cosh\sqrt{2\alpha}\,T\right]^{-\frac{1}{2}}.$$

This is a famous calculation which can be made using many different methods. We will return to it in Exercise (?).

Hint: Use Euler's product formula to see that

$$\frac{d}{dt}\log\frac{\sinh t}{t} = 2t\sum_{n=1}^{\infty}\frac{1}{n^2\pi^2+t^2} \quad \text{for} \quad t\in\mathbb{R}.$$

Exercise 5.3.32. Related to the preceding exercise, but easier, is finding the Laplace transform of the variance

$$V_T(\theta) \equiv \frac{1}{T}\int_0^T \theta(t)^2\,dt - \left(\frac{1}{T}\int_0^T \theta(t)\,dt\right)^2$$

of a Brownian path over the interval $[0,T]$. To do this caluculation, first use Brownian scaling to show that

$$\mathbb{E}^{\mathcal{W}^{(1)}}\left[e^{-\alpha V_T}\right] = \mathbb{E}^{\mathcal{W}^{(1)}}\left[e^{-\alpha T V_1}\right].$$

Next, use elementary Fourier series to see that (cf. part (**iii**) of Exercise 5.3.30)

$$V_1(\theta) = 2\sum_{k=1}^{\infty}\left(\int_0^1 \theta(t)\cos(k\pi t)\,dt\right)^2 = \sum_{k=1}^{\infty}\frac{\left(\int_0^1 f_k(t)\,d\theta(t)\right)^2}{k^2\pi^2},$$

where $f_k(t) = 2^{\frac{1}{2}}\sin(k\pi t)$ for $k\ge 1$. Since the f_k are orthonormal in $L^2\big([0,\infty);\mathbb{R}\big)$, this leads to

$$\mathbb{E}^{\mathcal{W}^{(1)}}\left[e^{-\alpha V_1}\right] = \prod_{k=1}^{\infty}\left(1+\frac{2\alpha}{k^2\pi^2}\right)^{-\frac{1}{2}}.$$

Thus,

$$\mathbb{E}^{\mathcal{W}}\left[e^{-\alpha V_T}\right] = \sqrt{\frac{\sqrt{2\alpha T}}{\sinh(\sqrt{2\alpha T})}}.$$

Finally, using Wiener's choice of basis, show that $\theta \rightsquigarrow V_1(\theta)$ has the same distribution as $\theta \rightsquigarrow \int_0^1\big(\theta(t)-t\theta(1)\big)^2\,dt$ under $\mathcal{W}^{(1)}$.

EXERCISE 5.3.33. In this exercise, we discuss some properties of pinned Brownian motion. Given $T > 0$, set $\boldsymbol{\theta}_T(t) = \boldsymbol{\theta}(t) - \frac{t\wedge T}{T}\boldsymbol{\theta}(T)$. As we pointed out in § 5.3.5, the $\mathcal{W}^{(N)}$-distribution of $\boldsymbol{\theta}_T$ is that of a Brownian motion conditioned to be back at $\mathbf{0}$ at time T. Next set $\Theta_T(\mathbb{R}^N)$ to be the space of continuous paths $\boldsymbol{\theta} : [0,T] \longrightarrow \mathbb{R}^N$ satisfying $\boldsymbol{\theta}(0) = \mathbf{0} = \boldsymbol{\theta}(T)$, and let $\mathcal{W}_T^{(N)}$ denote the $\mathcal{W}^{(N)}$-distribution of $\boldsymbol{\theta} \in \Theta(\mathbb{R}^N) \longmapsto \boldsymbol{\theta}_T \in \Theta_T(\mathbb{R}^N)$.

(**i**) Show that the $\mathcal{W}^{(N)}$-distribution of $\{\boldsymbol{\theta}_T(t) : t \geq 0\}$ is the same as that of $\{T^{\frac{1}{2}}\boldsymbol{\theta}_1(T^{-1}t) : t \geq 0\}$.

(**ii**) Let $\mathbf{H}_T(\mathbb{R}^N) = \{\mathbf{h}\backslash[0,T] : \mathbf{h} \in \mathbf{H}(\mathbb{R}^N) \;\&\; \mathbf{h}(T) = 0\}$, and define $\|\mathbf{h}\|_{\mathbf{H}_T(\mathbb{R}^N)} = \|\dot{\mathbf{h}}\|_{L^2([0,T];\mathbb{R}^N)}$. Show that the triple $\big(\mathbf{H}_T(\mathbb{R}^N), \Pi_T(\mathbb{R}^N), \mathcal{W}_T^{(N)}\big)$ is an abstract Wiener space. In addition, show that $\mathcal{W}_T^{(N)}$ is invariant under **time reversal**. That is, $\{\boldsymbol{\theta}(t) : t \in [0,T]\}$ and $\{\boldsymbol{\theta}(T-t) : t \in [0,T]\}$ have the same distribution under $\mathcal{W}_T^{(N)}$.

(**iii**) Given an orthonormal basis $\{\mathbf{h}_m : m \geq 1\}$ in $\mathbf{H}_T(\mathbb{R}^N)$, extend each $\mathbf{h}_m$ as an element of $\mathbf{H}(\mathbb{R}^N)$ by taking it equal $\mathbf{0}$ on (T,∞), and show that, for $\mathcal{W}^{(N)}$-almost every $\boldsymbol{\theta}$,

$$\boldsymbol{\theta}_T \restriction [0,T] = \sum_{m=1}^{\infty} [\mathcal{I}(\mathbf{h}_m)](\boldsymbol{\theta})\mathbf{h}_m \restriction [0,T],$$

where the convergence is uniform on $[0,T]$. Conclude from this that for any sequence $\{X_m : m \geq 1\}$ of independent, standard normal random variables, $\sum_{m=1}^{\infty} X_m\mathbf{h}_m$ is $\mathbb{P}$-almost surely convergent uniformly on $[0,T]$ to an element of $\Theta_T(\mathbb{R}^N)$ which has distribution $\mathcal{W}_T^{(N)}$.

EXERCISE 5.3.34. Let $\mathbf{H}^U(\mathbb{R}^N) = \mathbf{H}(\mathbb{R}^N) \cap L^2\big([0,\infty);\mathbb{R}^N\big)$, and make $\mathbf{H}^U(\mathbb{R}^N)$ into a Hilbert space with norm

$$\|\mathbf{h}^U\|_{\mathbf{H}^U(\mathbb{R}^N)} \equiv \sqrt{\|\mathbf{h}^U\|^2_{\mathbf{H}(\mathbb{R}^N)} + \frac{1}{4}\|\mathbf{h}^U\|^2_{L^2([0,\infty);\mathbb{R}^N)}}.$$

(**i**) Given $\boldsymbol{\lambda} \in \Lambda(\mathbb{R}^N)$, define

$$\mathbf{h}^U_{\boldsymbol{\lambda}}(t) = \int_{(0,\infty)} u(t,\tau)\,\boldsymbol{\lambda}(d\tau),$$

where $u(s,t)$ as in (5.3.28). Show that $\mathbf{h}^U_{\boldsymbol{\lambda}} \in \mathbf{H}^U(\mathbb{R}^N)$ and that

$$\|\mathbf{h}^U_{\boldsymbol{\lambda}}\|^2_{\mathbf{H}^U(\mathbb{R}^N)} = \iint_{(0,\infty)^2} u(s,t)\,\boldsymbol{\lambda}(ds)\cdot\boldsymbol{\lambda}(dt).$$

(**ii**) Let $\mathcal{U}^{(N)} \in \mathbf{M}_1\big(\Theta(\mathbb{R}^N)\big)$ denote the $\mathcal{W}^{(N)}$-distribution of the Ornstein–Unlenbeck process $\{\mathbf{U}(t,\mathbf{0}) : t \geq 0\}$, and show that $\big(\mathbf{H}^U(\mathbb{R}^N), \Theta(\mathbb{R}^N), \mathcal{U}^{(N)}\big)$ is an abstract Wiener space.

(**iii**) Given $\mathbf{h} \in \mathbf{H}(\mathbb{R}^N)$, set $\mathbf{h}^U(t) = e^{-\frac{t}{2}}\mathbf{h}\big(e^t - 1\big)$, and show that $\mathbf{h} \rightsquigarrow \mathbf{h}^U$ is an isometry from $\mathbf{H}(\mathbb{R}^N)$ onto $\mathbf{H}^U(\mathbb{R}^N)$. In light of the discussion in § 5.3.8, this should come as no surprise.

Exercise 5.3.35. Let $\mathbf{H}^U_R(\mathbb{R}^N)$ be the space of absolutely continuous $\mathbf{h}^U_R$: $\mathbb{R} \longrightarrow \mathbb{R}^N$ with the property that

$$\|\mathbf{h}^U_R\|_{\mathbf{H}^U_R(\mathbb{R}^N)} \equiv \sqrt{\|\dot{\mathbf{h}}^U_R\|^2_{L^2(\mathbb{R};\mathbb{R}^N)} + \tfrac{1}{4}\|\mathbf{h}^U_R\|^2_{L^2(\mathbb{R};\mathbb{R}^N)}} < \infty.$$

If $\mathcal{U}^{(N)}_R$ denotes the distribution of a reversible Ornstein–Uhlenbeck process, show that $\big(\mathbf{H}^U_R(\mathbb{R}^N), \Theta_R(\mathbb{R}^N), \mathcal{U}^{(N)}_R\big)$ is an abstract Wiener space when $\Theta_R(\mathbb{R}^N)$ is the Banach space of $\boldsymbol{\theta}_R \in C(\mathbb{R};\mathbb{R}^N)$ with the property that $\|\boldsymbol{\theta}_R\|_{\Theta_R(\mathbb{R}^N)} \equiv \sup_{t\in\mathbb{R}}(1+|t|)^{-1}|\boldsymbol{\theta}_R(t)| < \infty$.

Hint: Identify $\Theta_R(\mathbb{R}^N)^*$ with the space $\Lambda_R(\mathbb{R}^N)$ of $\mathbb{R}^N$-valued, Borel measures $\boldsymbol{\lambda}_R$ on $\mathbb{R}$ satisfying $\int_{\mathbb{R}}(1+|t|)|\boldsymbol{\lambda}_R|(dt) < \infty$, and proceed as in Exercise 5.3.33.