# 3 Polynomials over the white noise

## 3a Convergence of moments

Measuring devices of Sect. 1 are linear; that is, $\sqrt{\varepsilon}\sum_k \varphi(k\varepsilon)\tau(k\varepsilon)$ is a linear function of our random signs $\tau(k\varepsilon)$. Accordingly, random variables $\int \varphi(x)\,dB(x)$ are linear functions of the white noise. (Recall, they are linear functions on $(\mathbb{R}^\infty, \gamma^\infty)$.) More general functions are nonlinear, of course. Say,[1]

$$\text{(3a1)} \qquad \operatorname{Lim} \Big( \sqrt{\varepsilon} \sum_k \varphi(k\varepsilon)\tau(k\varepsilon) \Big)^2 = \Big( \int \varphi(x)\,dB(x) \Big)^2 \,;$$

such a function is quadratic; the right-hand side is a continuous random variable; its distribution is not normal (in fact, it is a Gamma distribution).

The relation

$$\text{(3a2)} \qquad \operatorname{Lim} \mathbb{E} \Big( \sqrt{\varepsilon} \sum_k \varphi(k\varepsilon)\tau(k\varepsilon) \Big)^2 = \mathbb{E} \Big( \int \varphi(x)\,dB(x) \Big)^2$$

does not follow from (3a1), since convergence in distribution does not imply convergence of moments. Though, (3a2) holds anyway:

$$\mathbb{E} \Big( \sqrt{\varepsilon} \sum_k \varphi(k\varepsilon)\tau(k\varepsilon) \Big)^2 = \varepsilon \sum_{k,l} \varphi(k\varepsilon)\varphi(l\varepsilon)\mathbb{E}\tau(k\varepsilon)\tau(l\varepsilon) = \varepsilon \sum_k \varphi^2(k\varepsilon) \to \int \varphi^2(x)\,dx \,,$$
$$\mathbb{E} \Big( \int \varphi(x)\,dB(x) \Big)^2 = \int \varphi^2(x)\,dx \,.$$

More generally,

$$\text{(3a3)} \qquad \operatorname{Lim} \mathbb{E} \Big( \sqrt{\varepsilon} \sum_k \varphi(k\varepsilon)\tau(k\varepsilon) \Big)^n = \mathbb{E} \Big( \int \varphi(x)\,dB(x) \Big)^n$$

for all $n = 1, 2, 3, \dots$ Indeed, similarly to 1b, the approximation

$$\mathbb{E} \exp\big( z\sqrt{\varepsilon} \sum_k \varphi(k\varepsilon)\tau(k\varepsilon) \big) = \prod_k \frac{\exp(z\sqrt{\varepsilon}\varphi(k\varepsilon)) + \exp(-z\sqrt{\varepsilon}\varphi(k\varepsilon))}{2} \approx$$
$$\approx \prod_k \Big( 1 + \frac{1}{2} z^2 \varepsilon \varphi^2(k\varepsilon) \Big) \approx \exp\Big( \frac{1}{2} \varepsilon z^2 \sum_k \varphi^2(k\varepsilon) \Big) \approx \exp\Big( \frac{1}{2} z^2 \int \varphi^2(x)\,dx \Big)$$

suggests that

$$\operatorname{Lim} \mathbb{E} \exp\big( z\sqrt{\varepsilon} \sum_k \varphi(k\varepsilon)\tau(k\varepsilon) \big) = \exp\Big( \frac{1}{2} z^2 \int \varphi^2(x)\,dx \Big) = \mathbb{E} \exp\Big( z \int \varphi(x)\,dB(x) \Big)$$

[1] Convergence *in distribution* is meant.

for all *complex* numbers $z$ (not just $z = i\lambda$, $\lambda \in \mathbb{R}$). That is true, at least, for Riemann integrable (that is, continuous almost everywhere, compactly supported, and bounded) functions $\varphi$. The convergence is uniform on every disk ($|z| \le \text{const}$). It follows via contour integration that

$$\operatorname{Lim} \frac{d^n}{dz^n} \mathbb{E} \exp \Big( z\sqrt{\varepsilon} \sum_k \varphi(k\varepsilon)\tau(k\varepsilon) \Big) = \frac{d^n}{dz^n} \mathbb{E} \exp \Big( z \int \varphi(x)\, dB(x) \Big) ,$$

which means (3a3). A seemingly more general formula

$$\text{(3a4)} \quad \operatorname{Lim} \mathbb{E} \Big( \sqrt{\varepsilon} \sum_k \varphi_1(k\varepsilon)\tau(k\varepsilon) \Big) \dots \Big( \sqrt{\varepsilon} \sum_k \varphi_n(k\varepsilon)\tau(k\varepsilon) \Big) = \\ = \mathbb{E} \Big( \int \varphi_1(x)\, dB(x) \Big) \dots \Big( \int \varphi_n(x)\, dB(x) \Big)$$

follows from (3a3) by a simple algebra. You see, $a_1 a_2$ is a linear combination of $(a_1 + a_2)^2$ and $(a_1 - a_2)^2$; similarly, $a_1 \dots a_n$ is a linear combination[2] of $2^{n-1}$ terms $(a_1 \pm a_2 \pm \dots \pm a_n)^n$.

A still more general form of the same fact:

$$\text{(3a5)} \quad \operatorname{Lim} \mathbb{E} f \Big( \sqrt{\varepsilon} \sum_k \varphi_1(k\varepsilon)\tau(k\varepsilon), \dots, \sqrt{\varepsilon} \sum_k \varphi_d(k\varepsilon)\tau(k\varepsilon) \Big) = \\ = \mathbb{E} f \Big( \int \varphi_1(x)\, dB(x), \dots, \int \varphi_d(x)\, dB(x) \Big)$$

for every *polynomial* function $f : \mathbb{R}^d \to \mathbb{R}$ (of any degree $n$).[3]

The discrete model often helps to guess and prove useful formulas for the continuous model (while the continuous model often makes formulas simpler). The next result is just an equality for the multinomial distribution. However, it is instructive to get it by means of our scaling limit.

**3a6 Exercise.**

$$\mathbb{E} \Big( \int \varphi_1(x)\, dB(x) \Big) \Big( \int \varphi_2(x)\, dB(x) \Big) \Big( \int \varphi_3(x)\, dB(x) \Big) \Big( \int \varphi_4(x)\, dB(x) \Big) = \\ = \langle \varphi_1, \varphi_2 \rangle \langle \varphi_3, \varphi_4 \rangle + \langle \varphi_1, \varphi_3 \rangle \langle \varphi_2, \varphi_4 \rangle + \langle \varphi_1, \varphi_4 \rangle \langle \varphi_2, \varphi_3 \rangle$$

for every $\varphi_1, \varphi_2, \varphi_3, \varphi_4 \in L_2(\mathbb{R})$; here $\langle \varphi_1, \varphi_2 \rangle = \int \varphi_1(x)\varphi_2(x)\, dx$.

Prove it. What about 6 factors, and more? What happens for $\varphi_1 = \varphi_2 = \dots = \varphi$?

Hint. Turn from the continuum to the discrete model (but what about Riemann integrability?), open the brackets, take the expectation, and note that $\mathbb{E}(\tau_k \tau_l \tau_m \tau_n) = (\mathbb{E}\tau_k\tau_l)(\mathbb{E}\tau_m\tau_n) + (\mathbb{E}\tau_k\tau_m)(\mathbb{E}\tau_l\tau_n) + (\mathbb{E}\tau_k\tau_n)(\mathbb{E}\tau_l\tau_m)$ unless $k = l = m = n$; the latter case disappears in the limit.

---

[2]Namely,

$$\frac{1}{2^n} \sum_{\tau_1, \dots, \tau_n = \pm 1} \tau_1 \dots \tau_n (\tau_1 a_1 + \dots + \tau_n a_n)^n = n! a_1 \dots a_n$$

(think, why).

[3]Of course, it holds for all bounded continuous functions; however, a polynomial is not bounded.

## 3b Orthogonal polynomials

For the discrete model we have

$$\Big(\sqrt{\varepsilon}\sum_k \varphi_1(k\varepsilon)\tau(k\varepsilon)\Big)\Big(\sqrt{\varepsilon}\sum_k \varphi_2(k\varepsilon)\tau(k\varepsilon)\Big) = \\ = \underbrace{\varepsilon\sum_{k\ne l}\varphi_1(k\varepsilon)\varphi_2(l\varepsilon)\tau(k\varepsilon)\tau(l\varepsilon)}_{\text{quadratic}} + \underbrace{\varepsilon\sum_k \varphi_1(k\varepsilon)\varphi_2(k\varepsilon)}_{\text{constant}};$$

interestingly, the second sum (containing only const/$\varepsilon$ terms) is not small in comparison to the first sum (containing const/$\varepsilon^2$ terms). You see, terms of the first sum are random and tend to cancel each other, while terms of the second sum are positive. Further,

$$\Big(\sqrt{\varepsilon}\sum_k \varphi_1(k\varepsilon)\tau(k\varepsilon)\Big)\Big(\sqrt{\varepsilon}\sum_k \varphi_2(k\varepsilon)\tau(k\varepsilon)\Big)\Big(\sqrt{\varepsilon}\sum_k \varphi_3(k\varepsilon)\tau(k\varepsilon)\Big) = \\ = \underbrace{\varepsilon^{3/2}\sum_{k\ne l,k\ne m,l\ne m}\varphi_1(k\varepsilon)\varphi_2(l\varepsilon)\varphi_3(m\varepsilon)\tau(k\varepsilon)\tau(l\varepsilon)\tau(m\varepsilon)}_{\text{cubic}} + \\ + \underbrace{\sqrt{\varepsilon}\sum_k \varphi_1(k\varepsilon)\tau(k\varepsilon)\cdot\varepsilon\sum_{m\ne k}\varphi_2(m\varepsilon)\varphi_3(m\varepsilon) + \text{ two similar terms }}_{\text{linear}} + \\ + \underbrace{\varepsilon^{3/2}\sum_k \varphi_1(k\varepsilon)\varphi_2(k\varepsilon)\varphi_3(k\varepsilon)\tau(k\varepsilon)}_{\text{small}}$$

etc. That is quite useful, since terms of different degree (constant, linear, quadratic, cubic, ...) are uncorrelated (orthogonal) random variables.

Following a physical tradition, we introduce so-called Wick product, denoted $:\dots:$, by

$$:\Big(\sqrt{\varepsilon}\sum_k \varphi_1(k\varepsilon)\tau(k\varepsilon)\Big)\dots\Big(\sqrt{\varepsilon}\sum_k \varphi_n(k\varepsilon)\tau(k\varepsilon)\Big): = \\ = \varepsilon^{n/2}\sum \varphi_1(k_1\varepsilon)\dots\varphi_n(k_n\varepsilon)\tau(k_1\varepsilon)\dots\tau(k_n\varepsilon),$$

where the sum (in the right-hand side) is taken over *pairwise distinct* $k_1,\dots,k_n$.

**3b1 Exercise.**

$$\begin{aligned} :U_\varepsilon V_\varepsilon: &= U_\varepsilon V_\varepsilon - \mathbb{E}(U_\varepsilon V_\varepsilon), \\ :U_\varepsilon V_\varepsilon W_\varepsilon: &= U_\varepsilon :V_\varepsilon W_\varepsilon: - \mathbb{E}(U_\varepsilon V_\varepsilon)W_\varepsilon - \mathbb{E}(U_\varepsilon W_\varepsilon)V_\varepsilon + O(\varepsilon), \\ :U_\varepsilon V_\varepsilon W_\varepsilon X_\varepsilon: &= U_\varepsilon :V_\varepsilon W_\varepsilon X_\varepsilon: - \\ &\quad - \mathbb{E}(U_\varepsilon V_\varepsilon):W_\varepsilon X_\varepsilon: - \mathbb{E}(U_\varepsilon W_\varepsilon):V_\varepsilon X_\varepsilon: - \mathbb{E}(U_\varepsilon X_\varepsilon):V_\varepsilon W_\varepsilon: + O(\varepsilon) \end{aligned}$$

whenever $U_\varepsilon, V_\varepsilon, \dots$ are linear functions of the form $\sqrt{\varepsilon}\sum \varphi(k\varepsilon)\tau(k\varepsilon)$ each. Here $O(\varepsilon)$ stands for a random variable whose $L_2$-norm is $\le \text{const}\cdot\varepsilon$ for $\varepsilon \to 0$. (Functions $\varphi$ are bounded.)

Prove it. What about more than 4 factors?

Hint (for $U_\varepsilon V_\varepsilon W_\varepsilon X_\varepsilon$): the error contains $O(1/\varepsilon^2)$ terms of norm $O(\varepsilon^2)$ each; these terms are orthogonal (each one contains the product of 2 different $\tau(k\varepsilon)$).

Thus, we *define* Wick products over the white noise:

(3b2)
$$
\begin{aligned}
:U: &= U\,,\\
:UV: &= U:V: - \mathbb{E}(UV)\,,\\
:UVW: &= U:VW: - \mathbb{E}(UV):W: - \mathbb{E}(UW):V:\,,\\
:UVWX: &= U:VWX: - \mathbb{E}(UV):WX: - \mathbb{E}(UW):VX: - \mathbb{E}(UX):VW:\,,
\end{aligned}
$$

and so on. Here $U, V, \dots$ are linear functions of the white noise, that is, random variables of the form $\int \varphi(x)\,dB(x)$, $\varphi \in L_2(\mathbb{R})$.

In particular,

$$
\begin{aligned}
UV &= :UV: + \mathbb{E}(UV)\,;\\
UVW &= :UVW: + \mathbb{E}(UV)W + \mathbb{E}(UW)V + \mathbb{E}(VW)U\,;\\
UVWX &= :UVWX: + \mathbb{E}(UV):WX: + \mathbb{E}(UW):VX: + \mathbb{E}(UX):VW: +\\
&\quad + \mathbb{E}(VW):UX: + \mathbb{E}(VX):UW: + \mathbb{E}(WX):UV: +\\
&\quad + \mathbb{E}(UV)\mathbb{E}(WX) + \mathbb{E}(UW)\mathbb{E}(VX) + \mathbb{E}(UX)\mathbb{E}(VW)\,.
\end{aligned}
\tag{3b3}
$$

Note that $:UVW: \; = \; :UWV: \; = \; :VUW: \; = \dots$ in spite of the 'asymmetric' definition. In fact, the order of factors does not matter, for any number of factors. The case $U = V = \dots = X$, $\|X\| = 1$, gives

(3b4)
$$
\begin{aligned}
:X: &= X\,; & X &= :X:\,;\\
:X^2: &= X^2 - 1\,; & X^2 &= :X^2: + 1\,;\\
:X^3: &= X^3 - 3X\,; & X^3 &= :X^3: + 3:X:\,;\\
:X^4: &= X^4 - 6X^2 + 3\,; & X^4 &= :X^4: + 6:X^2: + 3\,.
\end{aligned}
$$

These are well-known Hermite polynomials,

(3b5)
$$
\begin{gathered}
:X^n: = H_n(X)\,; \qquad (\|X\| = 1)\\
H_n(x) = (-1)^n \exp\Big(\frac12 x^2\Big)\frac{d^n}{dx^n}\exp\Big(-\frac12 x^2\Big)\,.
\end{gathered}
$$

(Be warned: some authors use $H_n(x\sqrt{2})$ or $2^{n/2}H_n(x\sqrt{2})$, call these functions 'Hermite polynomials' and denote them $H_n(x)$.) Here are some classical formulas for Hermite polynomials:

(3b6)
$$
\begin{gathered}
\frac{d}{dx}H_n(x) = nH_{n-1}(x)\,;\\
H_{n+1}(x) = xH_n(x) - nH_{n-1}(x) = \Big(x - \frac{d}{dx}\Big)H_n(x)\,;\\
\sum_{n=0}^\infty \frac{y^n}{n!}H_n(x) = \exp\Big(\frac12 x^2 - \frac12 (x-y)^2\Big)\,.
\end{gathered}
$$

The condition $\|X\| = 1$ is essential for (3b4). In general, for $\|X\| = \sigma \in (0, \infty)$, we apply (3b4) to $\frac{1}{\sigma}X$ and get $:X^2: \, = X^2 - \sigma^2$, $:X^3: \, = X^3 - 3\sigma^2 X$, $:X^4: \, = X^4 - 6\sigma^2 X^2 + 3\sigma^4$ and so on. In any case, $:X^n:$ is a polynomial of $X$; coefficients of the polynomial depend on $\|X\|$. More generally, $:X_1 \dots X_n:$ is a polynomial of $X_1, \dots, X_n$; coefficients depend on numbers $\mathbb{E}(X_k X_l)$. (The continuous model is meant.)

For the discrete model, $:U_\varepsilon V_\varepsilon:$ is still a polynomial of $U_\varepsilon, V_\varepsilon$; however, $:U_\varepsilon V_\varepsilon W_\varepsilon:$ is not a polynomial of $U_\varepsilon, V_\varepsilon, W_\varepsilon$, even if $U_\varepsilon = V_\varepsilon = W_\varepsilon$; that is, $:U_\varepsilon^3:$ is not a polynomial (in fact, not a function) of $U_\varepsilon$. For example, consider a linear combination of only 3 random signs, $U_\varepsilon = 2\tau_1 + \tau_2 + \tau_3$; then $:U_\varepsilon^3: \, = 12\tau_1\tau_2\tau_3$ (think, why), which is not a function of $2\tau_1 + \tau_2 + \tau_3$; indeed,

$$\begin{aligned} \tau_1 = +1,\ \tau_2 = \tau_3 = -1 \quad &\Longrightarrow \quad U_\varepsilon = 0,\ :U_\varepsilon^3: \, = +12\,;\\ \tau_1 = -1,\ \tau_2 = \tau_3 = +1 \quad &\Longrightarrow \quad U_\varepsilon = 0,\ :U_\varepsilon^3: \, = -12\,. \end{aligned}$$

This effect is caused by terms $O(\varepsilon)$ present in 3b1 but omitted in (3b2). When $\varepsilon \to 0$, the effect becomes small, and we get[4]

(3b7)
$$\begin{gathered} :UVW: \, = f(U,V,W)\,,\\ :U_\varepsilon V_\varepsilon W_\varepsilon: \, = f(U_\varepsilon, V_\varepsilon, W_\varepsilon) + o(1)\,,\\ \text{in other words,} \quad \| :U_\varepsilon V_\varepsilon W_\varepsilon: - f(U_\varepsilon, V_\varepsilon, W_\varepsilon)\| \xrightarrow[\varepsilon\to 0]{} 0\,; \end{gathered}$$

here $U_\varepsilon = \sqrt{\varepsilon}\sum_k \varphi_1(k\varepsilon)\tau(k\varepsilon)$, $U = \int \varphi(x)\, dB(x)$, and the same for $V_\varepsilon, V$ and $W_\varepsilon, W$; also, $f$ is a cubic polynomial.

**3b8 Exercise.**

$$\operatorname{Lim} \mathbb{E}\big( :U_\varepsilon V_\varepsilon W_\varepsilon: \big)\big( :X_\varepsilon Y_\varepsilon Z_\varepsilon: \big) = \mathbb{E}\big( :UVW: \big)\big( :XYZ: \big)\,;$$

prove it.

Hint: use (3b7) and 3a5.

Three factors are just an example; the statement holds for all cases; another example:

$$\operatorname{Lim} \mathbb{E}\big( :U_\varepsilon V_\varepsilon W_\varepsilon: \big)\big( :X_\varepsilon Y_\varepsilon: \big) = \mathbb{E}\big( :UVW: \big)\big( :XY: \big)\,.$$

However, $:U_\varepsilon V_\varepsilon W_\varepsilon:$ and $:X_\varepsilon Y_\varepsilon:$ are always orthogonal! Thus, $:UVW:$ and $:XY:$ are also orthogonal. Generally,

(3b9)
$$\begin{gathered} \mathbb{E}\big( :U_1 \dots U_m: \big)\big( :V_1 \dots V_n: \big) = 0 \quad \text{if } m \ne n\,;\\ \text{here } U_k = \int \varphi_k(x)\, dB(x)\,, \quad V_l = \int \psi_l(x)\, dB(x)\,, \text{ and } \varphi_k, \psi_l \in L_2(\mathbb{R})\,. \end{gathered}$$

The proof works for Riemann integrable functions, but these are dense in $L_2(\mathbb{R})$; the result extends by continuity. The case $U_k = X$, $V_l = X$, $\|X\| = 1$ is especially interesting:

(3b10)
$$\begin{gathered} \mathbb{E}\, H_m(X) H_n(X) = 0 \quad \text{if } m \ne n \text{ and } X \sim N(0,1)\,;\\ \text{that is, } \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H_m(x) H_n(x) e^{-x^2/2}\, dx = 0 \quad \text{for } m \ne n\,; \end{gathered}$$

the classical orthogonality property of Hermite polynomials.

[4] Why '$o(1)$' rather than '$O(\varepsilon)$'? Well, we may write $:U_\varepsilon V_\varepsilon W_\varepsilon: \, = f_\varepsilon(U_\varepsilon, V_\varepsilon, W_\varepsilon) + O(\varepsilon)$, and coefficients of $f_\varepsilon$ tend to coefficients of $f$, but maybe slower than $O(\varepsilon)$.

**3b11 Exercise.**

$$\begin{aligned}\mathbb{E}\big( :U_\varepsilon V_\varepsilon W_\varepsilon: \big)\big( :X_\varepsilon Y_\varepsilon Z_\varepsilon: \big) &= \mathbb{E}(U_\varepsilon X_\varepsilon)\mathbb{E}(V_\varepsilon Y_\varepsilon)\mathbb{E}(W_\varepsilon Z_\varepsilon) + \mathbb{E}(U_\varepsilon X_\varepsilon)\mathbb{E}(V_\varepsilon Z_\varepsilon)\mathbb{E}(W_\varepsilon Y_\varepsilon) + \\ &+ \mathbb{E}(U_\varepsilon Y_\varepsilon)\mathbb{E}(V_\varepsilon X_\varepsilon)\mathbb{E}(W_\varepsilon Z_\varepsilon) + \mathbb{E}(U_\varepsilon Y_\varepsilon)\mathbb{E}(V_\varepsilon Z_\varepsilon)\mathbb{E}(W_\varepsilon X_\varepsilon) + \\ &+ \mathbb{E}(U_\varepsilon Z_\varepsilon)\mathbb{E}(V_\varepsilon X_\varepsilon)\mathbb{E}(W_\varepsilon Y_\varepsilon) + \mathbb{E}(U_\varepsilon Z_\varepsilon)\mathbb{E}(V_\varepsilon Y_\varepsilon)\mathbb{E}(W_\varepsilon X_\varepsilon) + O(\varepsilon)\,.\end{aligned}$$

Prove it (under the same assumptions as 3b1).

Hint. Open the brackets, take the expectation; now the situation is somewhat similar to that of (the proof of) 2b2.

Combining 3b11 and (3a5) we get

$$\begin{aligned}\mathbb{E}\big( :UVW: \big)\big( :XYZ: \big) &= \mathbb{E}(UX)\mathbb{E}(VY)\mathbb{E}(WZ) + \mathbb{E}(UX)\mathbb{E}(VZ)\mathbb{E}(WY) + \\ &+ \mathbb{E}(UY)\mathbb{E}(VX)\mathbb{E}(WZ) + \mathbb{E}(UY)\mathbb{E}(VZ)\mathbb{E}(WX) + \\ &+ \mathbb{E}(UZ)\mathbb{E}(VX)\mathbb{E}(WY) + \mathbb{E}(UZ)\mathbb{E}(VY)\mathbb{E}(WX)\end{aligned} \tag{3b12}$$

for all linear functions $U, V, W, X, Y, Z$ of the white noise. (Extended by continuity from the Riemann integrable case.) Generally, $\mathbb{E}\big( :U_1 \dots U_n: \big)\big( :V_1 \dots V_n: \big)$ is the sum of $n!$ terms, each term being the product of $n$ factors of the form $\mathbb{E}(U_k V_l)$ each.

The case $U_k = X$, $V_l = X$, $\|X\| = 1$ is especially interesting:

$$\begin{gathered}\mathbb{E}\, H_n^2(X) = n! \quad \text{if } X \sim N(0,1)\,; \\ \text{that is, } \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H_n^2(x) e^{-x^2/2}\, dx = n!\,.\end{gathered} \tag{3b13}$$

Another interesting case: $U_k = X$, $V_l = Y$, $\|X\| = 1$, $\|Y\| = 1$, $\mathbb{E}(XY) = \rho \in [-1, +1]$; then

$$\mathbb{E}\big( H_n(X) H_n(Y) \big) = n! \rho^n\,. \tag{3b14}$$

We see from (3b10) and (3b13) that functions $\frac{1}{\sqrt{n!}} H_n(\cdot)$ are orthonormal in $L_2(\gamma^1)$; in other words, functions $x \mapsto \frac{1}{\sqrt{n!}} H_n(\cdot)\sqrt{(2\pi)^{-1/2} e^{-x^2/2}}$ are orthonormal in $L_2(\mathbb{R})$. Whether they are a basis of the whole $L_2$, or only a subspace? That is, are polynomials dense in $L_2(\gamma^1)$? Of course, every continuous function can be approximated by polynomials on any bounded interval; but these polynomials are large outside the interval.

We have (recall (3b6))

$$\sum_{n=0}^{\infty} \frac{(i\lambda)^n}{n!} H_n(x) = e^{\lambda^2/2} e^{i\lambda x}\,; \tag{3b15}$$

the convergence is pointwise (in fact, uniform on bounded intervals). On the other hand, the series converges (to something!) in $L_2(\gamma^1)$, since

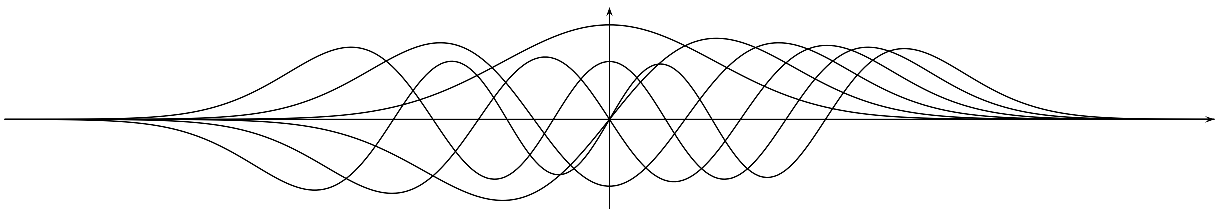
$$\sum_{n=0}^{\infty} \frac{(i\lambda)^n}{n!} H_n(x) = \sum_{n=0}^{\infty} \frac{(i\lambda)^n}{\sqrt{n!}} \cdot \frac{1}{\sqrt{n!}} H_n(x)\,, \quad \text{and } \sum \left| \frac{(i\lambda)^n}{\sqrt{n!}} \right|^2 = e^{\lambda^2} < \infty\,.$$

The two limits must conform (think, why);

$$\sum_{n=0}^{\infty} \frac{(i\lambda)^n}{n!} H_n(x) = e^{\lambda^2/2} e^{i\lambda x} \quad \text{in } L_2(\gamma^1) \,. \tag{3b16}$$

We see that the function $x \mapsto e^{i\lambda x}$ belongs to the subspace (spanned by Hermite polynomials), for every $\lambda \in \mathbb{R}$. It follows (think, why) that the subspace is the whole $L_2(\gamma^1)$. So,

$$\begin{gathered} \text{functions } \frac{1}{\sqrt{n!}} H_n(\cdot) \text{ are an orthonormal basis of } L_2(\gamma^1) \,; \\ \text{functions } x \mapsto \frac{1}{\sqrt{n!}} H_n(x) \sqrt{(2\pi)^{-1/2} e^{-x^2/2}} \text{ are an orthonormal basis of } L_2(\mathbb{R}) \,. \end{gathered} \tag{3b17}$$



**3b18 Exercise.** Let random variables $X, Y$ have a two-dimensional normal distribution such that $X \sim N(0,1)$, $Y \sim N(0,1)$, $\mathbb{E}(XY) = \rho \in [-1,+1]$. Then[5]

$$\operatorname{Cov}\big(f(X), g(Y)\big) \le \rho \sqrt{\operatorname{Var} f(X)} \sqrt{\operatorname{Var} g(Y)}$$

for all $f, g \in L_2(\gamma^1)$.

Prove it. Is the equality possible?

Hint. Use (3b14) and (3b17).

## 3c Wiener chaos

We take some orthonormal basis $(\varphi_1, \varphi_2, \dots)$ of $L_2(\mathbb{R})$. (It may be the Haar basis, as in 1b, or (3b17), or whatever.) Linear random variables $\int \varphi_k(x)\, dB(x)$ are an orthonormal system in $L_2(\Omega, \mathcal{F}, P)$; here $(\Omega, \mathcal{F}, P)$ is the probability space supporting the white noise. On the other hand, (nonlinear) random variables $\frac{1}{\sqrt{n!}} :\big(\int \varphi_1(x)\, dB(x)\big)^n: \; = \frac{1}{\sqrt{n!}} H_n\big(\int \varphi_1(x)\, dB(x)\big)$ are another orthonormal system in $L_2(\Omega, \mathcal{F}, P)$. The two systems have a common element $\int \varphi_1(x)\, dB(x)$; except for that, they are orthogonal (recall (3b9)). Clearly, no one is a basis of the whole $L_2(\Omega, \mathcal{F}, P)$.

Denote $U_k = \int \varphi_k(x)\, dB(x)$. We have $:U_k^2: \; = U_k^2 - 1$ and $:U_k U_l: \; = U_k U_l$ for $k < l$ (recall (3b3)). The formula

$$\mathbb{E}(:UV:)(:XY:) = \mathbb{E}(UX)\mathbb{E}(VY) + \mathbb{E}(UY)\mathbb{E}(VX)$$

[5] Do not forget, $\operatorname{Var} U = \mathbb{E}U^2 - (\mathbb{E}U)^2 = \mathbb{E}\big(U - \mathbb{E}U\big)^2$, and $\operatorname{Cov}(U,V) = \mathbb{E}(UV) - \mathbb{E}(U)\mathbb{E}(V) = \mathbb{E}\big((U - \mathbb{E}U)(V - \mathbb{E}V)\big)$.

(similar to (3b12)) gives

$$\mathbb{E}(\,:U_kU_l:\,)(\,:U_mU_n:\,) = \begin{cases} 2 & \text{if } k=l=m=n, \\ 1 & \text{if } k=m<l=n, \\ 0 & \text{otherwise} \end{cases}$$

(think, why). It means that the system $\big(\,:U_kU_l:\,\big)_{k\le l}$ is orthogonal, and the system

$$\Big(\tfrac{1}{\sqrt{2}}:U_k^2:\Big)_k \cup \Big(\,:U_kU_l:\,\Big)_{k<l} = \Big(\tfrac{1}{\sqrt{2}}(U_k^2-1)\Big)_k \cup \Big(U_kU_l\Big)_{k<l}$$

is orthonormal. The system depends on the choice of the basis $(\varphi_k)$, but the spanned subspace does not depend, according to the next result.

**3c1 Exercise.** The following three sets of random variables span the same linear subspace in $L_2(\Omega,\mathcal{F},P)$:
$:U_kU_l:$ for $k\le l$;
$:\big(\int\varphi(x)\,dB(x)\big)^2:$ for $\varphi\in L_2(\mathbb{R})$;
$:\big(\int\varphi(x)\,dB(x)\big)\big(\int\psi(x)\,dB(x)\big):$ for $\varphi,\psi\in L_2(\mathbb{R})$.
Prove it.

The subspace described in 3c1 is called the second Wiener chaos.

As you may guess, the first Wiener chaos is spanned by $U_k$, and consists of all linear random variables $\int\varphi(x)\,dB(x)$, $\varphi\in L_2(\mathbb{R})$. These two chaoses (first and second) are orthogonal (recall (3b9)).

A better understanding of the second chaos is gained from the discrete model. Namely,

$$U_\varepsilon = \sqrt{\varepsilon}\sum_k \varphi(k\varepsilon)\tau(k\varepsilon)\,, \qquad V_\varepsilon = \sqrt{\varepsilon}\sum_k \psi(k\varepsilon)\tau(k\varepsilon)\,,$$
$$:U_\varepsilon V_\varepsilon: \,= \varepsilon\sum_{k\ne l}\varphi(k\varepsilon)\psi(l\varepsilon)\tau(k\varepsilon)\tau(l\varepsilon) =$$
$$= \sum_{k<l}\big(\varphi(k\varepsilon)\psi(k\varepsilon)+\psi(k\varepsilon)\varphi(l\varepsilon)\big)\sqrt{\varepsilon}\tau(k\varepsilon)\sqrt{\varepsilon}\tau(l\varepsilon)\,,$$

which suggests for $U=\int\varphi(x)\,dB(x)$, $V=\int\psi(x)\,dB(x)$ the equality

$$:UV: \,= \iint_{x<y}\big(\varphi(x)\psi(y)+\psi(x)\varphi(y)\big)\,dB(x)dB(y)\,;$$

in order to give it a meaning, we should define

$$\iint_{x<y}\xi(x,y)\,dB(x)dB(y)$$

for an arbitrary function $\xi\in L_2(\Delta_2)$, where $\Delta_2=\{(x,y)\in\mathbb{R}^2 : x<y\}$.

An orthonormal basis $(\varphi_k)$ of $L_2(\mathbb{R})$ gives us an orthonormal basis $(\varphi_k\otimes\varphi_l)$ of $L_2(\mathbb{R}^2)$; here

$$(\varphi\otimes\psi)(x,y)=\varphi(x)\psi(y)\,.$$

Thus, every $\xi \in L_2(\mathbb{R}^2)$ may be written as

$$\xi = \sum_{k,l} c_{k,l}\varphi_k \otimes \varphi_l \,, \qquad \sum_{k,l} |c_{k,l}|^2 = \|\xi\|^2 < \infty \,,$$

and we *define*

$$\iint_{x\ne y} \xi(x,y)\, dB(x)dB(y) = \sum_{k,l} c_{k,l} :U_k U_l: \tag{3c2}$$

(the right-hand side converges in $L_2(\Omega,\mathcal{F},P)$).

**3c3 Exercise.** For all $\varphi,\psi \in L_2(\mathbb{R})$,

$$\iint_{x\ne y} \varphi(x)\psi(y)\, dB(x)dB(y) = \,:\Big( \int \varphi(x)\, dB(x) \Big) \Big( \int \psi(y)\, dB(y) \Big): \,.$$

Prove it.

**3c4 Exercise.** If $\xi$ is symmetric (that is, $\xi(x,y)=\xi(y,x)$), then

$$\mathbb{E} \Big( \iint_{x\ne y} \xi(x,y)\, dB(x)dB(y) \Big)^2 = 2 \iint |\xi(x,y)|^2 \, dxdy \,.$$

Prove it. Does it hold for non-symmetric $\xi$?

Hint:

$$\begin{aligned} \mathbb{E} \Big( \iint_{x\ne y} \varphi_k(x)\varphi_l(y)\, dB(x)dB(y) \Big) \Big( \iint_{x\ne y} \varphi_m(x)\varphi_n(y)\, dB(x)dB(y) \Big) = \\ = \mathbb{E} \big( :U_k U_l: \big) \big( :U_m U_n: \big) = \langle \varphi_k, \varphi_m \rangle \langle \varphi_l, \varphi_n \rangle + \langle \varphi_k, \varphi_n \rangle \langle \varphi_l, \varphi_m \rangle = \\ = 2 \iint \frac{\varphi_k(x)\varphi_l(y) + \varphi_l(x)\varphi_k(y)}{2} \cdot \frac{\varphi_m(x)\varphi_n(y) + \varphi_n(x)\varphi_m(y)}{2} \, dxdy \,. \end{aligned}$$

Given $\xi \in L_2(\Delta_2)$, we define

$$\iint_{x<y} \xi(x,y)\, dB(x)dB(y) = \frac12 \iint_{x\ne y} \tilde\xi(x,y)\, dB(x)dB(y) \,,$$

$$\text{where } \tilde\xi(x,y) = \begin{cases} \xi(x,y) & \text{if } x<y, \\ \xi(y,x) & \text{if } x>y \end{cases}$$

(that is, $\tilde\xi$ is the symmetric extension of $\xi$). We have now

$$\mathbb{E} \Big( \iint_{x<y} \xi(x,y)\, dB(x)dB(y) \Big)^2 = \iint_{x<y} |\xi(x,y)|^2 \, dxdy \,.$$

**3c5 Exercise.** The map

$$L_2(\Delta_2) \ni \xi \mapsto \iint_{x<y} \xi(x,y)\, dB(x)dB(y) \in L_2(\Omega, \mathcal{F}, P)$$

is uniquely determined by two properties:

(a) linearity: for all $\xi_1, \xi_2 \in L_2(\Delta_2)$ and $c_1, c_2 \in \mathbb{R}$,

$$\iint_{x<y} \big( c_1\xi_1(x,y) + c_2\xi_2(x,y) \big)\, dB(x)dB(y) =$$
$$= c_1 \iint_{x<y} \xi_1(x,y)\, dB(x)dB(y) + c_2 \iint_{x<y} \xi_2(x,y)\, dB(x)dB(y)\,;$$

(b) for all $\varphi, \psi \in L_2(\mathbb{R})$,

$$\iint_{x<y} \big( \varphi(x)\psi(y) + \psi(x)\varphi(y) \big)\, dB(x)dB(y) = \,:\Big( \int \varphi(x)\, dB(x) \Big) \Big( \int \psi(x)\, dB(x) \Big):\,.$$

Prove it.

We see that $\iint \xi(x,y)\, dB(x)dB(y)$ does not depend on the choice of a basis $(\varphi_k)$, as far as the map $L_2(\mathbb{R}) \ni \varphi \mapsto \int \varphi(x)\, dB(x) \in L_2(\Omega, \mathcal{F}, P)$ is given (recall 1b10).

All said about the second chaos can be generalized to the third chaos, and higher. The $n$-th Wiener chaos is the subspace of $L_2(\Omega, \mathcal{F}, P)$ spanned by Wick products of $n$ linear random variables (recall 3c1). The map (here $\Delta_n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 < \dots < x_n\}$)

$$L_2(\Delta_n) \ni \xi \mapsto \int_{x_1<\dots<x_n} \!\!\!\cdots\! \int \xi(x_1, \dots, x_n)\, dB(x_1) \dots dB(x_n) \in L_2(\Omega, \mathcal{F}, P)$$

may be defined similarly to (3c2), or as the only linear map $L_2(\Delta_n) \to L_2(\Omega, \mathcal{F}, P)$ such that

$$\int_{x_1<\dots<x_n} \!\!\!\cdots\! \int \sum \varphi_{k_1}(x_1) \dots \varphi_{k_n}(x_n)\, dB(x_1) \dots dB(x_n) =$$
$$= \,:\Big( \int \varphi_1(x)\, dB(x) \Big) \dots \Big( \int \varphi_n(x)\, dB(x) \Big):$$

where the sum is taken over all permutations $(k_1, \dots, k_n)$ of numbers $1, \dots, n$. The map is isometric,

$$\mathbb{E} \Big( \int_{x_1<\dots<x_n} \!\!\!\cdots\! \int \xi(x_1, \dots, x_n)\, dB(x_1) \dots dB(x_n) \Big)^2 = \int_{x_1<\dots<x_n} \!\!\!\cdots\! \int |\xi(x_1, \dots, x_n)|^2\, dx_1 \dots dx_n\,,$$

and maps $L_2(\Delta_n)$ onto the whole $n$-th chaos.

An example:

(3c6)
$$\int\!\!\cdots\!\!\int_{0<x_1<\dots<x_n<1} dB(x_1)\dots dB(x_n) = \frac{1}{n!} \mathbf{:} \Big( \int_0^1 dB(x) \Big) \dots \Big( \int_0^1 dB(x) \Big) \mathbf{:} = \frac{1}{n!} H_n\big(B(1)\big) ;$$
$$\mathbb{E} \Big( \frac{1}{n!} H_n\big(B(1)\big) \Big)^2 = \frac{1}{n!} = \int\!\!\cdots\!\!\int_{0<x_1<\dots<x_n<1} dx_1 \dots dx_n .$$

Interestingly, the integral of $dB(x_1)\dots dB(x_n)$ over pairwise distinct $x_1,\dots,x_n \in (0,1)$ is equal to $H_n\big(B(1)\big)$ rather than $B^n(1)$; the difference is the singular contribution of degenerate cases.

Here is a general form of a square integrable function over the white noise:

(3c7)
$$L_2(\Delta) \ni \xi \mapsto \sum_{n=0}^\infty \int\!\!\cdots\!\!\int_{x_1<\dots<x_n} \xi(x_1,\dots,x_n)\, dB(x_1)\dots dB(x_n) \in L_2(\Omega,\mathcal{F},P) ;$$

here $\Delta = \uplus \Delta_n$ is the disjoint union of all $\Delta_n$ (a single point $\Delta_0$, a line $\Delta_1$, a half-plane $\Delta_2$ and so on); each $\Delta_n$ is equipped with the $n$-dimensional Lebesgue measure (for $\Delta_0$ it is the unit mass at the point), thus, $\Delta$ is equipped with a measure, and $L_2(\Delta) = L_2(\Delta_0) \oplus L_2(\Delta_1) \oplus \dots$ is well-defined. The right-hand side converges in $L_2$; and the map is isometric:

$$\mathbb{E} \Big( \sum_{n=0}^\infty \int\!\!\cdots\!\!\int_{x_1<\dots<x_n} \xi(x_1,\dots,x_n)\, dB(x_1)\dots dB(x_n) \Big)^2 =$$
$$= \sum_{n=0}^\infty \int\!\!\cdots\!\!\int_{x_1<\dots<x_n} |\xi(x_1,\dots,x_n)|^2 \, dx_1 \dots dx_n .$$

Does (3c7) cover the whole $L_2(\Omega,\mathcal{F},P)$? It depends on the choice of $(\Omega,\mathcal{F},P)$. The answer is positive for $(\Omega,\mathcal{F},P) = (\mathbb{R}^\infty,\gamma^\infty)$ (recall 1b), which follows from completeness of Hermite polynomials (recall (3b17)). In general, (3c7) covers $L_2(\Omega,\mathcal{F}_B,P|_{\mathcal{F}_B})$, where $\mathcal{F}_B \subset \mathcal{F}$ is the sub-$\sigma$-field generated by the white noise (or equivalently, the Brownian motion).

An example (recall (3c6) and (3b15)):

$$e^{i\lambda B(1)} = e^{-\lambda^2/2} \sum_{n=0}^\infty \frac{(i\lambda)^n}{n!} H_n\big(B(1)\big) = e^{-\lambda^2/2} \sum_{n=0}^\infty (i\lambda)^n \int\!\!\cdots\!\!\int_{0<x_1<\dots<x_n<1} dB(x_1)\dots dB(x_n) ;$$

(3c8)
$$\mathbb{E} \big| e^{i\lambda B(1)} \big|^2 = 1 = (e^{-\lambda^2/2})^2 \sum_{n=0}^\infty |i\lambda|^{2n} \underbrace{\int\!\!\cdots\!\!\int_{0<x_1<\dots<x_n<1} dx_1 \dots dx_n}_{=1/n!} .$$

All said in 3c till now concerns the continuous model (the white noise). Returning to discrete models, we may consider nonlinear 'spin-measuring devices' of the form

$$X_{\varepsilon,M,N,\xi} = \sum_{n=0}^N \varepsilon^{n/2} \sum_{-M/\varepsilon \le k_1 < \dots < k_n \le M/\varepsilon} \xi(k_1\varepsilon,\dots,k_n\varepsilon)\tau(k_1\varepsilon)\dots\tau(k_n\varepsilon) ;$$

a 'test function' $\xi$ is defined on $\Delta = \Delta_0 \uplus \Delta_1 \uplus \Delta_2 \uplus \dots$; its restriction to $\Delta_1 = \mathbb{R}$ is nothing but $\varphi$, the one-dimensional test function introduced in Sect. 1 for the linear case. Similarly to Sect. 1, the scaling limit stipulates not only $\varepsilon \to 0$, but also $M \to \infty$ and (a new element) $N \to \infty$, which leads to a number of possible setups. I choose

$$\operatorname{Lim}(\dots) = \lim_{M,N\to\infty} \lim_{\varepsilon\to 0}(\dots)\,.$$

Similarly to Sect. 1 we assume that the test function $\xi$ not only belongs to $L_2(\Delta)$, but also is locally Riemann integrable; it means Riemann integrability of its restriction to every bounded domain of every $\Delta_n$; or equivalently, for every $n$, $\xi|_{\Delta_n}$ must be continuous almost everywhere and locally bounded. Similarly to 1b6 and 1b12, one can prove that[6]

$$\operatorname{Lim} \mathbb{E}\, f(X_{\varepsilon,M,N,\xi_1}, \dots, X_{\varepsilon,M,N,\xi_d}) = \mathbb{E}\, f(X_{\xi_1}, \dots, X_{\xi_d})$$

for every $d \in \{1, 2, \dots\}$, every bounded continuous function $f : \mathbb{R}^d \to \mathbb{R}$, and every locally Riemann integrable $\xi_1, \dots, \xi_d \in L_2(\Delta)$; here

$$X_\xi = \sum_{n=0}^{\infty} \int \dots \int_{x_1<\dots<x_n} \xi(x_1,\dots,x_n)\, dB(x_1) \dots dB(x_n)\,.$$

[6]However, I do not prove it now.