SMOOTHING IN HIERARCHICAL CONSTRUCTION

1. Haar construction of Brownian motion revisited

For simplicity we shall assume throughout that we work on a fixed time interval $[0, T]$ with $T = 1$.

With the Haar mother wavelet

$$\psi(t) := \begin{cases} 1, & 0 \le t < \frac{1}{2}, \\ -1, & \frac{1}{2} \le t < 1, \\ 0, & \text{else}, \end{cases} \tag{1}$$

we construct the Haar basis of $L^2([0,1])$ by setting

$$\psi_{-1}(t) := \mathbf{1}_{[0,1]}(t), \tag{2a}$$

$$\psi_{n,k}(t) := 2^{n/2}\psi\left(2^n t - k\right), \quad n \in \mathbb{N}_0,\ k = 0, \dots, 2^n - 1. \tag{2b}$$

We note that $\operatorname{supp}\psi_{n,k} = [2^{-n}k, 2^{-n}(k+1)]$. Moreover, we define a grid $\mathcal{D}^n := \left\{ t^n_\ell \mid \ell = 0, \dots, 2^{n+1} \right\}$ by $t^n_\ell := \frac{\ell}{2^{n+1}}$. Notice that the Haar functions up to level n are piece-wise constant with points of discontinuity given by $\mathcal{D}^n$.

Next we define the antiderivatives of the basis functions

$$\Psi_{-1}(t) := \int_0^t \psi_{-1}(s)ds, \tag{3a}$$

$$\Psi_{n,k}(t) := \int_0^t \psi_{n,k}(s)ds. \tag{3b}$$

For an i.i.d. set of standard normal random variables (*coefficients*) Z_{-1}, $Z_{n,k}$, $n \in \mathbb{N}_0$, $k = 0, \dots, 2^n - 1$, we can then define a standard Brownian motion

$$W_t := Z_{-1}\Psi_{-1}(t) + \sum_{n=0}^{\infty} \sum_{k=0}^{2^n-1} Z_{n,k}\Psi_{n,k}(t), \tag{4}$$

and the truncated version

$$W^N_t := Z_{-1}\Psi_{-1}(t) + \sum_{n=0}^{N} \sum_{k=0}^{2^n-1} Z_{n,k}\Psi_{n,k}(t). \tag{5}$$

Note that W^N already coincides with W along the grid $\mathcal{D}^N$. We define the corresponding increments for any function or process F by

$$\Delta^N_\ell F := F(t^N_{\ell+1}) - F(t^N_\ell). \tag{6}$$

2. Stochastic differential equations

For simplicity we consider a one-dimensional SDE X given by

$$dX_t = b(X_t)dW_t, \quad X_0 = x \in \mathbb{R}. \tag{7}$$

We assume that b is bounded and has bounded derivatives of all orders. Recall that we want to compute

$$E\left[g\left(X_T\right)\right]$$

for some function $g : \mathbb{R} \to \mathbb{R}$ which is not necessarily smooth. We also define the solution of the Euler scheme along the grid $\mathcal{D}^N$ by $X_0^N \coloneqq X_0 = x$ and

$$X_{\ell+1}^N \coloneqq X_\ell^N + b\left(X_\ell^N\right)\Delta_\ell^N W. \tag{8}$$

For convenience, we also define $X_T^N \coloneqq X_{2^N}^N$.

Clearly, the random variable X_ℓ^N is a deterministic function of the random variables Z_{-1} and $Z^N \coloneqq (Z_{n,k})_{n=0,\dots,N,\ k=0,\dots 2^n-1}$. Abusing notation, let us therefore write

$$X_\ell^N = X_\ell^N\left(Z_{-1}, Z^N\right)$$

for the appropriate (now deterministic) map $X_\ell^N : \mathbb{R}\times\mathbb{R}^{2^{N+1}-1} \to \mathbb{R}$. We shall write $y \coloneqq z_{-1}$ and z^N for the (deterministic) arguments of the function X_ℓ^N.

A note of caution is in order regarding convergence as $N \to \infty$: while the sequence of random processes $X_\cdot^N$ converges to the solution of (7) (under the usual assumptions on b), this is not true in any sense for the deterministic functions.

Define

$$H^N(z^N) \coloneqq E\left[g\left(X_T^N\left(Z_{-1}, z^N\right)\right)\right]. \tag{9}$$

We claim that H^N is analytic.

Let us consider a mollified version g_δ of g and the corresponding function H_δ^N (defined by replacing g with g_δ in (9)). Tacitly assuming that we can interchange integration and differentiation, we have

$$\frac{\partial H_\delta^N(z^N)}{\partial z_{n,k}} = E\left[g_\delta'\left(X_T^N\left(Z_{-1}, z^N\right)\right)\frac{\partial X_T^N(Z_{-1}, z^N)}{\partial z_{n,k}}\right].$$

Multiplying and dividing by $\frac{\partial X_T^N(Z_{-1}, z^N)}{\partial y}$ and replacing the expectation by an integral w.r.t. the standard normal density, we obtain

$$\frac{\partial H_\delta^N(z^N)}{\partial z_{n,k}} = \int_{\mathbb{R}} \frac{\partial g_\delta\left(X_T^N(y, z^N)\right)}{\partial y}\left(\frac{\partial X_T^N}{\partial y}(y, z^N)\right)^{-1}\frac{\partial X_T^N}{\partial z_{n,k}}(y, z^N)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\,dy. \tag{10}$$

If we are able to do integration by parts, then we can get rid of the mollification and obtain smoothness of H^N since we get

$$\frac{\partial H^N(z^N)}{\partial z_{n,k}} = -\int_{\mathbb{R}} g\left(X_T^N(y, z^N)\right)\frac{\partial}{\partial y}\left[\left(\frac{\partial X_T^N}{\partial y}(y, z^N)\right)^{-1}\frac{\partial X_T^N}{\partial z_{n,k}}(y, z^N)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\right]dy.$$

We realize that there is a potential problem looming in the inverse of the derivative w.r.t. y.[1] Before we continue, let us introduce the following notation: for sequences of random variables F_N, G_N we say that $F_N = \mathcal{O}(G_N)$ if there is a random variable C with finite moments of all orders such that for all N we have $|F_N| \le C\,|G_N|$ a.s.

Assumption 2.1. There are positive random variables C_p with finite moments of all orders such that

$$\forall N \in \mathbb{N},\ \forall \ell_1, \dots, \ell_p \in \{0, \dots, 2^N - 1\} : \left|\frac{\partial^p X_T^N}{\partial X_{\ell_1}^N \cdots \partial X_{\ell_p}^N}\right| \le C_p \text{ a.s.}$$

In terms of the above notation, that means that $\frac{\partial^p X_T^N}{\partial X_{\ell_1}^N \cdots \partial X_{\ell_p}^N} = \mathcal{O}(1)$.

Remark 2.2. It is probably hard to argue that a deterministic constant C may exist.

[1] Let us assume that $X_T^N(y, z^N) = \cos(y) + z_{n,k}$. Then (10) is generally not integrable.

Assumption 2.1 is natural, but now we need to make a much more serious assumption, which is probably difficult to verify in practise.

Assumption 2.3. For any $p \in \mathbb{N}$ we have that

$$\left(\frac{\partial X_T^N}{\partial y}\left(Z_{-1}, Z^N\right)\right)^{-p} = O(1).$$

Lemma 2.4. *We have*

$$\frac{\partial X_T^N}{\partial z_{n,k}}(Z_{-1}, Z^N) = 2^{-n/2+1} O(1)$$

in the sense that the $O(1)$ term does not depend on n or k.

Proof. First let us note that Assumption 2.1 implies that $\frac{\partial X_T^N}{\partial \Delta_\ell^N W} = O(1)$. Indeed, we have

$$\frac{\partial X_T^N}{\partial \Delta_\ell^N W} = \frac{\partial X_T^N}{\partial X_{\ell+1}^N} \frac{\partial X_{\ell+1}^N}{\partial \Delta_\ell^N W} = O(1) b(X_\ell^N) = O(1).$$

Next we need to understand which increments Δ_ℓ^N do depend on $Z_{n,k}$. This is the case iff $\operatorname{supp} \psi_{n,k}$ has a non-empty intersection with $]t_\ell^N, t_{\ell+1}^N[$. Explicitly, this means that

$$\ell 2^{-(N-n+1)} - 1 < k < (\ell+1) 2^{-(N-n+1)}.$$

If we fix N, k, n, this means that the derivative of $\Delta_\ell^N W$ w.r.t. $Z_{n,k}$ does not vanish iff

$$2^{N-n+1} k \le \ell < 2^{N-n+1}(k+1).$$

Noting that

$$\left|\frac{\partial \Delta_\ell^N W}{\partial Z_{n,k}}\right| = \left|\Delta_\ell^N \Psi_{n,k}\right| \le 2^{-(N-n/2)}, \tag{11}$$

we thus have

$$\frac{\partial X_T^N}{\partial z_{n,k}}(Z_{-1}, Z^N) = \sum_{\ell=2^{N-n+1}k}^{2^{N-n+1}(k+1)-1} \frac{\partial X_T^N}{\partial \Delta_\ell^N W} \frac{\partial \Delta_\ell^N W}{\partial Z_{n,k}} = 2^{N-n+1} 2^{-(N-n/2)} O(1) = 2^{-n/2+1} O(1). \tag{12}$$

$\square$

Lemma 2.5. *In the same sense as in Lemma 2.4 we have*

$$\frac{\partial^2 X_T^N}{\partial y \partial z_{n,k}}(Z_{-1}, Z^N) = 2^{-n/2+1} O(1).$$

Proof. $\Delta_\ell^N W$ is a linear function in Z_{-1} and Z^N, implying that all mixed derivatives $\frac{\partial^2 \Delta_\ell^N W}{\partial Z_{n,k} \partial Z_{-1}}$ vanish. From equation (12) we hence see that

$$\frac{\partial^2 X_T^N}{\partial z_{n,k} \partial y}(Z_{-1}, Z^N) = \sum_{\ell=2^{N-n+1}k}^{2^{N-n+1}(k+1)-1} \frac{\partial^2 X_T^N}{\partial \Delta_\ell^N W \partial Z_{-1}} \frac{\partial \Delta_\ell^N W}{\partial Z_{n,k}}.$$

Further,

$$\frac{\partial^2 X_T^N}{\partial \Delta_\ell^N W \partial Z_{-1}} = \sum_{j=0}^{2^{N+1}-1} \frac{\partial^2 X_T^N}{\partial \Delta_\ell^N W \partial \Delta_j^N W} \frac{\partial \Delta_j^N W}{\partial Z_{-1}}.$$

Note that

$$\frac{\partial^2 X_T^N}{\partial \Delta_\ell^N W \partial \Delta_j^N W} = \frac{\partial^2 X_T^N}{\partial X_{\ell+1}^N \partial X_{j+1}^N} b(X_\ell^N) b(X_j^N) + \mathbf{1}_{j<\ell} \frac{\partial X_T^N}{\partial X_\ell^N} b'(X_\ell^N) \frac{\partial X_\ell^N}{\partial X_{j+1}^N} b(X_j^N) = O(1) \tag{13}$$

by Assumption 2.1. We also have $\frac{\partial \Delta_j^N W}{\partial Z_{-1}} = O(2^{-N})$, implying the statement of the lemma. ◻

Remark 2.6. Lemma 2.4 and 2.5 also hold (mutatis mutandis) for $z_{n,k} = y$ (with $n = 0$).

Proposition 2.7. *We have $\frac{\partial H^N(z^N)}{\partial z_{n,k}} = O(2^{-n/2})$ in the sense that the constant in front of $2^{-n/2}$ does not depend on n or k.*

Proof. We have

$$\begin{aligned}\frac{\partial H^N(z^N)}{\partial z_{n,k}} &= -\int_{\mathbb{R}} g\left(X_T^N(y,z^N)\right)\frac{\partial}{\partial y}\left[\left(\frac{\partial X_T^N}{\partial y}(y,z^N)\right)^{-1}\frac{\partial X_T^N}{\partial z_{n,k}}(y,z^N)\right]dy\\ &= -\int_{\mathbb{R}} g\left(X_T^N(y,z^N)\right)\left[-\left(\frac{\partial X_T^N}{\partial y}(y,z^N)\right)^{-2}\frac{\partial^2 X_T^N}{\partial y^2}(y,z^N)\frac{\partial X_T^N}{\partial z_{n,k}}(y,z^N)+\right.\\ &\quad\left.+\left(\frac{\partial X_T^N}{\partial y}(y,z^N)\right)^{-1}\frac{\partial^2 X_T^N}{\partial z_{n,k}\partial y}(y,z^N) - y\left(\frac{\partial X_T^N}{\partial y}(y,z^N)\right)^{-1}\frac{\partial X_T^N}{\partial z_{n,k}}(y,z^N)\right]\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy.\end{aligned}$$

Notice that when $F^N(Z_{-1}, Z^N) = O(c)$ for some deterministic constant c, then this property is retained when integrating out one of the random variables, i.e., we still have

$$\int_{\mathbb{R}} F^N(y, Z^N)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy = O(c).$$

Hence, Lemma 2.4 and Lemma 2.5 together with Assumption 2.3 (for $p = 2$) imply that

$$\frac{\partial H^N(z^N)}{\partial z_{n,k}} = O(2^{-n/2})$$

with constants independent of n and k. ◻

For the general case we need

Lemma 2.8. *For any $p \in \mathbb{N}$ and indices $n_1, \dots, n_p$ and $k_1, \dots, k_p$ (satisfying $0 \le k_j < 2^{n_j}$) we have (with constants independent from n_j, k_j)*

$$\frac{\partial^p X_T^N}{\partial z_{n_1,k_1}\cdots\partial z_{n_p,k_p}}(Z_1, Z^N) = O\left(2^{-\sum_{j=1}^p n_j/2}\right).$$

The result also holds (mutatis mutandis) if one or several z_{n_j,k_j} are replaced by $y = z_{-1}$ (with n_j set to 0).

Proof. We start noting that each $\Delta_\ell^N W$ is a linear function of (Z_{-1}, Z^N) implying that all higher derivatives of $\Delta_\ell^N W$ w.r.t. (Z_{-1}, Z^N) vanish. Hence,

$$\frac{\partial^p X_T^N}{\partial Z_{n_1,k_1}\cdots\partial Z_{n_p,k_p}} = \sum_{\ell_1=2^{N-n_1+1}k_1}^{2^{N-n_1+1}(k_1+1)-1}\cdots\sum_{\ell_p=2^{N-n_p+1}k_p}^{2^{N-n_p+1}(k_p+1)-1}\frac{\partial^p X_T^N}{\partial\Delta_{\ell_1}^N\cdots\partial\Delta_{\ell_p}^N W}\frac{\partial\Delta_{\ell_1}^N W}{\partial Z_{n_1,k_1}}\cdots\frac{\partial\Delta_{\ell_p}^N W}{\partial Z_{n_p,k_p}}.$$

By a similar argument as in (13) we see that

$$\frac{\partial^p X_T^N}{\partial\Delta_{\ell_1}^N\cdots\partial\Delta_{\ell_p}^N W} = O(1).$$

By (11) we see that each summand in the above sum is of order $\prod_{j=1}^p 2^{-(N-n_j/2)}$. The number of summands in total is $\prod_{j=1}^p 2^{N-n_j+1}$. Therefore, we obtain the desired result. ◻

Theorem 2.9. *For any $p \in \mathbb{N}$ and indices $n_1, \dots, n_p$ and $k_1, \dots, k_p$ (satisfying $0 \le k_j < 2^{n_j}$) we have (with constants independent from n_j, k_j)*

$$\frac{\partial^p H^N}{\partial z_{n_1,k_1} \cdots \partial z_{n_p,k_p}}(Z^N) = O\left(2^{-\sum_{j=1}^p n_j/2}\right).$$

The result also holds (mutatis mutandis) if one or several z_{n_j,k_j} are replaced by $y = z_{-1}$ (with n_j set to 0). In particular, H^N is a smooth function.

Remark 2.10. We actually expect that H^N is analytic, but a formal proof seems difficult. In particular, note that our proof below relies on successively applying the above tricke for enabling integration by parts: divide by $\frac{\partial X_T^N}{\partial y}$ and then integrate by parts. This means that the number of terms (denoted by $\blacksquare$ below) increases fast as p increases by the product rule of differentiation. Hence, the constant in fron of the $O\left(2^{-\sum_{j=1}^p n_j/2}\right)$ term will depend on p and increase in p. In that sense, Theorem 2.9 needs to be understood as an assertion about the anisotropy in the variables $z_{n,k}$ rather than a statement about the behaviour of higher and higher derivatives of H^N. In fact, one can see that in our proof the number of summands increases as $p!$ in p. Therefore, the statement of the theorem does not already imply analyticity. Of course, this problem is an artefact of our construction, and there is no reason to assume such a behaviour in general.

Sketch of a proof of Theorem 2.9. We apply integration by parts p times as in the proof of Proposition 2.7, which shows that we can again replace the mollified payoff function g_δ by the true, non-smooth one g. Moreover, from the procedure we obtain a formula of the form

$$\frac{\partial^p H^N}{\partial z_{n_1,k_1} \cdots \partial z_{n_p,k_p}}(z^N) = \int_{\mathbb{R}} g\left(X_T^N(y, z^N)\right) \blacksquare \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy,$$

where $\blacksquare$ represents a long sum of products of various terms. However, it is quite easy to notice the following structure: ignoring derivatives w.r.t. y, each summand contains all derivatives w.r.t. $z_{n_1,k_1}, \dots, z_{n_p,k_p}$ exactly once. (Generally speaking, each summand will be a product of derivatives of X_T^N w.r.t. some z_{n_j,k_j}s, possibly with other terms such as polynomials in y and derivatives w.r.t. y included.) As all other terms are assumed to be of order $O(1)$ by Assumptions 2.1 and 2.3, this implies the claimed result by Lemma 2.8. □