

A multiscale guide to Brownian motion

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2016 J. Phys. A: Math. Theor. 49 043001

(<http://iopscience.iop.org/1751-8121/49/4/043001>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.67.119.101

This content was downloaded on 28/12/2015 at 12:29

Please note that [terms and conditions apply](#).

Topical Review

A multiscale guide to Brownian motion

Denis S Grebenkov¹, Dmitry Belyaev² and Peter W Jones³¹Laboratoire de Physique de la Matière Condensée, CNRS—Ecole Polytechnique, F-91128 Palaiseau, France²Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, UK³Yale University, Mathematics Department, PO Box 208283, New Haven, CT 06520–8283, USAE-mail: denis.grebenkov@polytechnique.edu

Received 5 July 2015, revised 21 October 2015

Accepted for publication 13 November 2015

Published 18 December 2015



CrossMark

Abstract

We revise the Lévy construction of Brownian motion as a simple though rigorous approach to operate with various Gaussian processes. A Brownian path is explicitly constructed as a linear combination of wavelet-based ‘geometrical features’ at multiple length scales with random weights. Such a wavelet representation gives a closed formula mapping of the unit interval onto the functional space of Brownian paths. This formula elucidates many classical results about Brownian motion (e.g., non-differentiability of its path), providing an intuitive feeling for non-mathematicians. The illustrative character of the wavelet representation, along with the simple structure of the underlying probability space, is different from the usual presentation of most classical textbooks. Similar concepts are discussed for the Brownian bridge, fractional Brownian motion, the Ornstein-Uhlenbeck process, Gaussian free fields, and fractional Gaussian fields. Wavelet representations and dyadic decompositions form the basis of many highly efficient numerical methods to simulate Gaussian processes and fields, including Brownian motion and other diffusive processes in confining domains.

Keywords: multiscale, wavelet, Brownian motion, Gaussian free field, fractal

(Some figures may appear in colour only in the online journal)

1. Introduction

Diffusion is a fundamental transport mechanism in nature and industry, with applications ranging from physics to biology, chemistry, engineering, and economics. This process has attracted much attention during the last decades, particularly in statistical and condensed

matter physics: diffusion–reaction processes, transport in porous media and biological tissues, trapping in heterogeneous systems, kinetic and aggregation phenomena like DLA, to name a few fields. From an intuitive point of view, Brownian motion (also known as the Wiener process) is often considered as a continuous limit of lattice random walks. However, a more rigorous background is needed to answer subtle questions. In mathematical textbooks, Brownian motion is defined as an almost surely continuous process with independent normally distributed increments [1–6]. The deceptive simplicity of this definition relies on the notion of almost surely that, in turn, requires a sophisticated formalism of Wiener measures in the space of continuous functions, filtrations, sigma-algebra, etc. Although this branch of mathematics is well developed, it is rather difficult for non-mathematicians, that is, the majority of scientists dealing with diffusion in their every-day research.

In this review, we discuss a different, but still rigorous, approach to define and operate with Brownian motion as suggested by P. Lévy [7]. We construct from scratch a simple and intuitively appealing representation of this process that gives a closed formula mapping of the unit interval onto the functional space of Brownian paths. In this framework, sampling a Brownian path is nothing other than picking up uniformly a point from the unit interval. Figuratively speaking, Brownian motion is constructed here by adding random wavelet-based geometrical features at multiple length scales. The explicit formula elucidates many classical results about Brownian motion (e.g., the non-differentiability of its path). The illustrative character of the wavelet representation, along with the simple structure of the underlying probability space, is different from the usual presentation of most classical textbooks.

Among various amazing properties, Brownian motion is known to have a self-similar structure: when a fragment of its path is magnified, it ‘looks’ like the whole path. In other words, any fragment obeys the same probability law as the whole path. As a consequence, Brownian paths exhibit their features at an (infinitely) broad range of length scales. As a matter of fact, multiscale geometrical structures are ubiquitous in nature and materials sciences [8]. For instance, respiratory and cardiovascular systems start from large conduits (trachea and artery) that are then split into thinner and thinner channels, up to a size of a few hundred microns for the alveoli and several microns for the smallest capillaries [9]. Another example is a high-performance concrete which is made with grains of different sizes (from centimeters to microns), smaller grains filling empty spaces between larger ones. The adaptive description of such self-similar structures needs to capture their mechanical or transport properties at different length scales and thus relies on intrinsically multiscale functions. We illustrate this idea by constructing Brownian motion and a Brownian bridge using wavelets, a family of functions with compact support and well defined scaling [10–12]. Wavelets appear as the natural mathematical language to describe and analyze multiscale structures, from heterogeneous rocks to biological tissues [13–15]. The wavelet construction of Brownian motion naturally extends to fractional Brownian motion and other Gaussian processes and fields, allowing one to efficiently simulate, for instance, turbulent diffusion with high Reynolds numbers or financial markets. In particular, we discuss the Gaussian (or massless) free field and fractional Gaussian fields which appear as basic models in different areas of physics, from astrophysics (cosmic microwave background) to critical phenomena, quantum physics, and turbulence [16–19]. Written in a spectral form in one dimension, the Brownian bridge and the fractional Gaussian field look very similar, one of them being the fractional derivative of the other. Putting together these two processes reveals deep relations between them, and this correspondence carries over to higher dimensions.

Most importantly, the wavelet representation is a starting point for a number of highly efficient numerical methods to simulate various Gaussian processes and fields. Though wavelet representations and the related numerical methods are all known, they are not easily

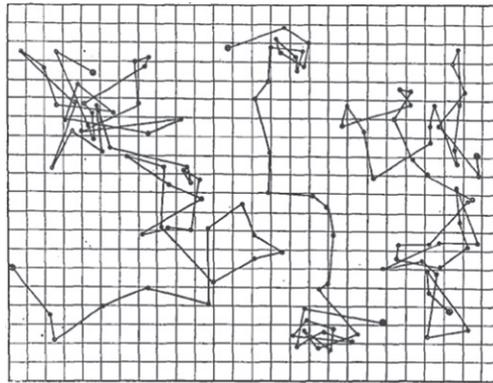


Figure 1. Three random trajectories of small mastic grains in water recorded by J. Perrin at 30-second intervals (reproduced from [22]).

available in a single source. Indeed the totality of these methods seems to be poorly understood, even amongst specialists. The purpose of this review is to present a unified and intuitive framework that is based on elementary mathematical structures like, for example, dyadic subdivision or a wavelet tree. We also discuss some simple number-theoretic shortcuts and consequent numerical algorithms. Finally, we describe fast simulations of restricted diffusion in confined domains, where one of the difficulties is the ability to quickly access the local geometry near the boundary. These techniques can be applied for studying Brownian motion and related processes or solving partial differential equations in complex multiscale media.

We hope that this didactic review will provoke interesting discussions amongst the experts and will help in a better understanding of both the theory and implementation of Brownian motion and other Gaussian processes and fields for a much broader community of their practical ‘users’, namely, physicists, biologists, chemists, engineers, and economists.

2. Brownian motion

In this section, we derive a wavelet representation of Brownian motion in a simple explicit way, allowing one to gain an intuitive feeling for this constructive approach.

2.1. Physical view: upscaling and downscaling

In order to illustrate the basic idea of a wavelet representation, we revisit the single particle tracking experiment by R. Brown who looked through a microscope at stochastic trajectories (now known as Brownian paths) of pollens of *Clarkia* (primrose family) [20]. The first examples of such trajectories for mastic grains in water were reported by J. Perrin [21, 22]. The essence of the wavelet representation can be recognized in his description of these trajectories (figure 1) that were recorded at 30-second intervals [22]: ‘*Ils ne donnent qu’une idée très affaiblie du prodigieux enchevêtrement de la trajectoire réelle. Si, en effet, on faisait des pointés de seconde en seconde, chacun de ces segments rectilignes se trouverait remplacé*

*par un contour polygonal de 30 côtés relativement aussi compliqué que le dessin ici reproduit, et ainsi de suite.*⁴

Following this idea, let us record the positions of a particle (e.g., pollen or grain) at successive time moments with a selected time resolution δ . Each particle submerged in water is permanently ‘bombarded’ by water molecules. Since the number of surrounding water molecules is very large and their tiny actions are mostly uncorrelated, net microscopic displacements of the particle cannot be considered deterministically as in classical mechanics, but random. Since the interaction of water molecules between them is very rapid as compared to the macroscopic resolution scale, there is no memory effect in their action on the heavy particle. As a consequence, the microscopic displacements of the particle are (almost) independent and have a finite variance σ^2 . The macroscopic displacement during the resolution time δ is the sum of a large number N of these displacements with zero mean (no coherent flow). Although the average displacement is also zero, the stochastic fluctuations around this value are of the order of $\sigma\sqrt{N}$. Moreover, the central limit theorem gives us a precise probabilistic description of the fluctuations, resulting in a normal (or Gaussian) distribution of macroscopic displacements of the particle [23]. It is worth stressing that the Gaussian character of the macroscopic displacements appears *without any specific knowledge about the microscopic interactions*. The only important information at the microscopic level is the stationary, uncorrelated character of the interaction, and the finite variance (if one of these conditions is missing, the resulting macroscopic process may exhibit anomalous behavior, see [24–26] and references therein). This is known as coarsening or upscaling: complex interactions and the specific features of the underlying microscopic dynamics are averaged out on macroscopic scales. This is the reason why Brownian motion (or diffusion in general) is so ubiquitous in nature and science. Once we know that the microscopic details are irrelevant (under the conditions mentioned above), we can extend the Gaussian behavior from macroscopic scale, where it has been established, to microscopic scale. This procedure can be called downscaling, when we explicitly and purposefully transpose the universal macroscopic behavior even onto smaller scales. The resulting *model* of microscopic dynamics exhibits Gaussian features at all scales. While the true dynamics and its Gaussian model can be completely different at microscopic scale, they become identical at macroscopic scale.

Knowing that a particle moves continuously, we connect its successive positions separated by time δ by a continuous line. Since the experimental setup is limited to the selected time scale δ , nothing can be said about the trajectory of the particle in between two data points. In other words, the only condition for the trajectory to pass through the recorded points leaves us a variety of choices for the shape of the connecting continuous line. The common choice is connecting the successive positions by linear segments. As one will see below, this choice fixes a particular wavelet representation, Haar wavelets. We shall show that other wavelets, corresponding to other choices of continuous connections, are also useful.

When the magnification and time resolution of the experimental setup are increased, smaller details of the particle’s trajectory appear, allowing one to refine the above piecewise linear approximation. Repeating this procedure, in theory up to infinity, one recovers all geometrical features and thus constructs the whole Brownian path. In what follows, we put this schematic description into a more rigorous mathematical frame.

⁴ ‘They provide only a very rough idea of the prodigious intricacy of the real trajectory. If one got points every second, each of the straight segments would be replaced by a polygonal contour of 30 sides having approximately the same complexity as the current plot, and so forth’ (translated by authors).

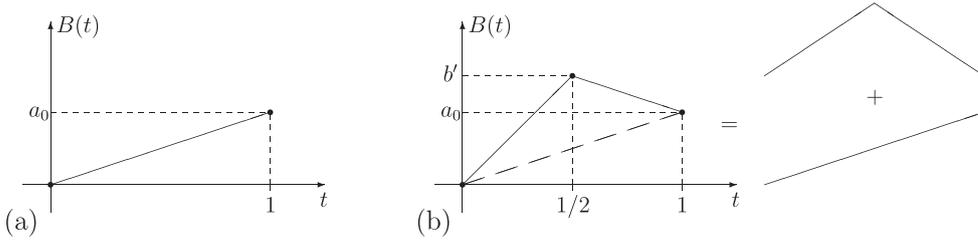


Figure 2. Iterative construction of Brownian motion: **(a)** linear approximation $B_0(t)$ by a segment at scale 2^0 ; **(b)** linear approximation $B_1(t)$ by two segments at scale 2^{-1} . The latter ‘skewed tent’ can be uniquely represented as the sum of a ‘symmetric tent’ and a linear shift.

2.2. Mathematical view: multiscale construction

We start with the position ‘records’ at every unit time: $t = 1, 2, 3, \dots$ (e.g., every second) along one coordinate (two- and higher dimensional Brownian motion is then obtained by taking d independent copies of the one-dimensional process). We focus on the time interval between 0 and 1, the construction being applicable to any interval $[\ell, \ell + 1]$. For convenience, Brownian motion is started at $t = 0$ from the origin: $B(0) = 0$. By definition (or as a consequence of the central limit theorem if one relies on the above physical reasoning), the position at time $t = 1$, $B(1)$, is a random variable a_0 distributed according to the standard normal (or Gaussian) law, $\mathcal{N}(0, 1)$, with mean zero and variance one

$$\mathbb{P}\{a_0 \in (x, x + dx)\} = dx \frac{e^{-x^2/2}}{\sqrt{2\pi}}. \tag{1}$$

In physics, the variance σ^2 is related to the diffusion coefficient, $D = \sigma^2/(2\tau)$, where τ is the one step duration; here, $\sigma = \tau = 1$ yielding $D = 1/2$ throughout the review. A linear approximation at this time scale ($\delta = 1$) is simply

$$B_0(t) = a_0 t,$$

that connects the positions $B(0) = 0$ and $B(1) = a_0$ by a linear segment.

If the time resolution is doubled, a new, intermediate position $b' = B(1/2)$ can now be seen (figure 2). The random variable b' is conditioned by the fact that Brownian motion passes through the points $(0, 0)$ and $(1, a_0)$, the value of a_0 being already known. b' is distributed according to the normal law with mean value $\frac{1}{2}(B(0) + B(1)) = a_0/2$ and variance $1/4$ (see appendix A). In other words, one can write $b' = a_0/2 + a_{00}/2$, where the new random variable $a_{00} \in \mathcal{N}(0, 1)$ (i.e., distributed according to the standard normal law (1)) is independent of $B(0) = 0$ and $B(1) = a_0$.

The linear approximation at the time scale $\delta = 1/2$ connects three successive points $(0, 0)$, $(1/2, b')$ and $(1, a_0)$ by two linear segments:

$$B_1(t) = \begin{cases} 2tb', & 0 \leq t \leq \frac{1}{2}, \\ 2t(a_0 - b') + (2b' - a_0), & \frac{1}{2} \leq t \leq 1. \end{cases} \tag{2}$$

The shape of this approximation looks like a skewed tent (figure 2) that can be represented as the sum of a linear shift and a ‘symmetric tent’ function:

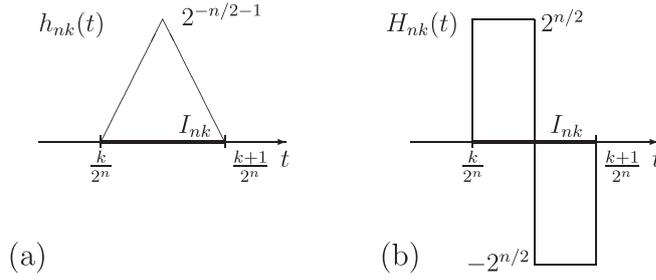


Figure 3. Tent function $h_{nk}(t)$ (a) and the related Haar function $H_{nk}(t)$ (b), both having the support on the interval $I_{nk} = [k2^{-n}, (k + 1)2^{-n})$.

$$B_1(t) = a_0 t + (2b' - a_0)h_{00}(t) = a_0 t + a_{00}h_{00}(t), \tag{3}$$

where the ‘symmetric tent’ function $h_{00}(t)$ is

$$h_{00}(t) = \begin{cases} t, & 0 \leq t \leq \frac{1}{2}, \\ 1 - t, & \frac{1}{2} \leq t \leq 1, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

The decomposition (3) into the linear function t and the tent function $h_{00}(t)$ is unique. The new approximation $B_1(t)$ is obtained from the previous one, $B_0(t)$, by adding the new term representing a smaller geometrical detail.

The same concept is applicable at every scale. Assume that an approximation $B_n(t)$ of Brownian motion is already constructed on the time scale $\delta = 2^{-n}$, i.e., the positions $b_k = B(t_k)$ are known at successive times $t_k = k2^{-n}$, k ranging from 0 to 2^n . The approximate Brownian path is a piecewise linear function passing successively through these points.

At the next time scale 2^{-n-1} , a new, intermediate position $b'_k = B(t'_k)$ of Brownian motion at time $t'_k = \frac{1}{2}(t_k + t_{k+1}) = (k + \frac{1}{2})2^{-n}$ should be determined for each k . As previously, the random variable b'_k is conditioned by the fact that Brownian motion is known to pass through the points (t_k, b_k) and (t_{k+1}, b_{k+1}) . It is again distributed according to the normal law, with mean value $\frac{1}{2}(b_k + b_{k+1})$ and variance $2^{-n}/4$. In other words, one can write b'_k as

$$b'_k = \frac{1}{2}(b_k + b_{k+1}) + 2^{-n/2-1}a_{nk}, \tag{5}$$

where the new normal random variable $a_{nk} \in \mathcal{N}(0, 1)$ is independent of the other positions. The linear segment between (t_k, b_k) and (t_{k+1}, b_{k+1}) is then replaced by two linear segments connecting the three successive points (t_k, b_k) , (t'_k, b'_k) , and (t_{k+1}, b_{k+1}) . This is a new ‘skewed tent’ function which can be uniquely decomposed as the sum of the previous linear segment and a symmetric tent function $h_{nk}(t)$ with the weight a_{nk} , where

$$h_{nk}(t) = 2^{-n/2}h_{00}(2^n t - k) \tag{6}$$

is a rescaled symmetric tent function on the interval $I_{nk} = [k2^{-n}, (k + 1)2^{-n})$ (see figure 3(a)). We stress again that the new approximation is obtained from the previous one by simply adding the tent function $h_{nk}(t)$, representing a smaller geometrical detail at the new scale 2^{-n-1} , weighted by a normally distributed coefficient a_{nk} which is independent of the previously determined positions.

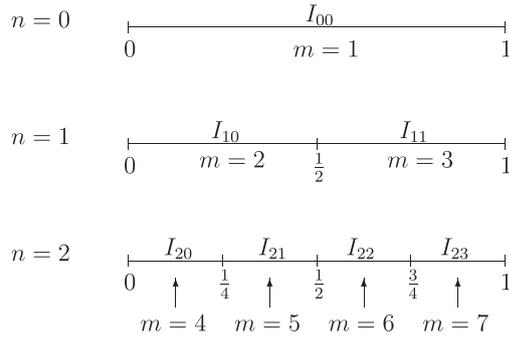


Figure 4. Enumeration of the dyadic subintervals using a single index m .

This construction is applicable to all linear segments (all k) at the given scale n , and it is valid for any scale. Repeating this procedure from the scale 2^0 up to infinity, one obtains the Lévy–Ciesielski representation of Brownian motion on the unit interval:

$$B(t) = a_0 t + \sum_{n=0}^{\infty} \sum_{k=0}^{2^n-1} a_{nk} h_{nk}(t), \tag{7}$$

where all weights a_0 and a_{nk} are independent $\mathcal{N}(0, 1)$ random variables. As discussed in appendix B, all these Gaussian weights can be related to a single random uniform number that provides a natural parameterization of Brownian paths.

The dyadic structure of the intervals implies that for any $n \in \mathbb{Z}$, there exists only one interval I_{nk} of length 2^{-n} containing a point $t : k2^{-n} \leq t < (k + 1)2^{-n}$. The index k is simply the integer part of $2^n t : k = \lfloor 2^n t \rfloor$ (i.e., the largest integer that does not exceed $2^n t$). As a consequence, the convergence in the above formula is very rapid. In fact, if one needs to obtain the value of the function $B(t)$ with a desired precision ε , it is sufficient to calculate the first $\log_2(1/\varepsilon)$ terms, $\log_2 x$ being the logarithm of x with the base 2.

Subtracting the linear term $a_0 t$ from equation (7) yields the Haar wavelet representation of a Brownian bridge \hat{B}_t on the unit interval, i.e., Brownian motion conditioned to return to 0 at time $t = 1 : \hat{B}_t = B_t - tB_1$ (it can also be defined as a Gaussian process with mean zero and covariance $\mathbb{E}\{\hat{B}_t \hat{B}_s\} = \min\{t, s\} - ts$).

2.3. Dyadic decomposition and interval subdivision

In the wavelet representation (7), the first sum is carried over all scales n , while the second sum covers all 2^n subintervals I_{nk} at the scale n . In some cases, it is convenient to enumerate all the dyadic subintervals I_{nk} using a single index,

$$m = 2^n + k, \tag{8}$$

as shown in figure 4. Since k ranges from 0 to $2^n - 1$, the new index m uniquely identifies the interval I_{nk} . In particular, one easily retrieves the pair (n, k) from m as

$$n = \lfloor \log_2 m \rfloor, k = m - 2^n.$$

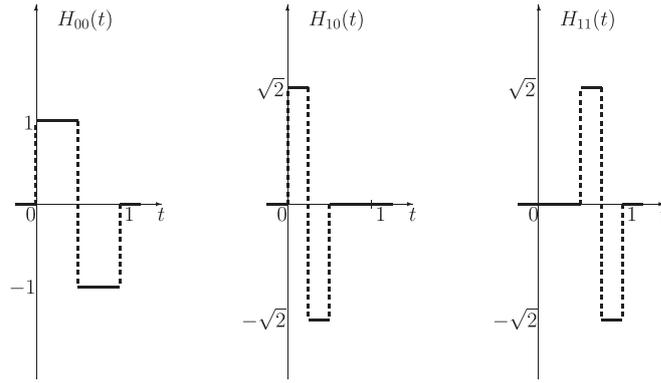


Figure 5. Haar wavelets $H_{nk}(t)$ are obtained by dilations and translations of the mother function $\phi^{1,1}(t) = H_{00}(t)$ (shown on the left).

Using the notation

$$\tilde{a}_m = \begin{cases} a_0, & m = 0, \\ a_{nk}, & m > 0, \end{cases} \quad \tilde{h}_m(t) = \begin{cases} t, & m = 0, \\ h_{nk}(t), & m > 0, \end{cases}$$

we can write equation (7) in a more compact form

$$B(t) = \sum_{m=0}^{\infty} \tilde{a}_m \tilde{h}_m(t). \tag{9}$$

As a result, the Brownian trajectory is decomposed into a sum of tent functions (plus a linear term) with random independent identically distributed Gaussian weights. As shown in appendix B, all these weights can be determined from a single uniformly distributed random variable.

2.4. Haar wavelets

In equations (7) or (9), Brownian motion is decomposed into a sum of a linear function and tent functions. Figure 3(a) illustrates that any tent function $h_{nk}(t)$ can actually be represented as the integral of a piecewise-constant function

$$h_{nk}(t) = \int_0^t dt' H_{nk}(t'), \tag{10}$$

where $H_{nk}(t)$ is called the Haar function and is defined to be 0 on the complement of I_{nk} and to take values $2^{n/2}$ and $-2^{n/2}$ on its left and right subintervals, respectively (figure 3(b)). In fact, all Haar functions are obtained by translations and dilations of a single ‘mother’ function $\phi^{1,1}(t)$:

$$\phi^{1,1}(t) = \begin{cases} 1, & 0 \leq t < \frac{1}{2}, \\ -1, & \frac{1}{2} < t \leq 1, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

As illustrated in figure 5, one has

$$H_{nk}(t) = 2^{n/2} \phi^{1,1}(2^n t - k). \tag{12}$$

As previously, it is convenient to use the single index m to denote different Haar functions (completed by a constant):

$$\tilde{H}_m(t) = \begin{cases} 1, & m = 0, \\ H_{nk}(t), & m > 0. \end{cases}$$

Equation (9) yields the following representation for Brownian motion

$$B(t) = \int_0^t dW(t'), \tag{13}$$

where $dW(t)$ denotes the Gaussian white noise which is defined here through Haar wavelet decomposition:

$$dW(t) = \sum_{m=0}^{\infty} \tilde{a}_m \tilde{H}_m(t). \tag{14}$$

It is easy to check that the Haar functions together with a constant form an orthonormal basis in the space $L^2([0, 1])$ of measurable and square integrable functions that is

$$\int_0^1 dt \tilde{H}_m(t) \tilde{H}_{m'}(t) = \delta_{m,m'}.$$

Moreover, this basis is known to be complete in $L^2([0, 1])$. This means that any function from $L^2([0, 1])$ can be decomposed into a linear combination of Haar functions (and a constant). The use of Haar wavelets for constructing stochastic processes was first proposed by Ciesielski [27].

2.5. General representation

The wavelet representation (14) can be extended to *any* complete orthonormal basis $\{\psi_j(t)\}_{j=1,2,\dots}$ of $L^2([0, 1])$. Although we started from the Haar basis, one can switch from one basis to another depending on the problem. In particular, this flexibility will help us to relate Brownian motion to a one-dimensional fractional Gaussian field with logarithmic corrections.

If the orthonormal basis $\{\psi_j(t)\}$ is complete, one can decompose any function $\tilde{H}_m(t)$ into a linear combination of $\psi_j(t)$:

$$\tilde{H}_m(t) = \sum_{j=1}^{\infty} c_{m,j} \psi_j(t),$$

where the coefficients $c_{m,j}$ satisfy the orthogonality relations

$$\sum_{m=0}^{\infty} c_{m,j}^2 = 1 \quad (j = 1, 2, 3, \dots). \tag{15}$$

Substitution of this decomposition into equation (14) gives

$$dW(t) = \sum_{m=0}^{\infty} \tilde{a}_m \sum_{j=1}^{\infty} c_{m,j} \psi_j(t) = \sum_{j=1}^{\infty} \hat{a}_j \psi_j(t), \tag{16}$$

with new random weights

$$\hat{a}_j = \sum_{m=0}^{\infty} \tilde{a}_m c_{m,j}.$$

The sum of independent Gaussian variables is a Gaussian variable, and its variance is the sum of the squared coefficients $c_{m,j}^2$, which is equal to 1 according to equation (15). In other words,

$\hat{a}_j \in \mathcal{N}(0, 1)$. Moreover, the new random variables \hat{a}_j are independent due to the orthogonality of the functions $\psi_j(t)$. We have thus shown that the Gaussian white noise can be decomposed into a linear combination with independent Gaussian weights in an arbitrary complete orthonormal basis of $L^2([0, 1])$.

The completeness of the basis $\{\psi_j(t)\}$ yields the usual covariance of the Gaussian white noise

$$\mathbb{E}\{dW(t_1)dW(t_2)\} = \sum_{j_1, j_2=1}^{\infty} \psi_{j_1}(t_1)\psi_{j_2}(t_2)\mathbb{E}\{\hat{a}_{j_1}\hat{a}_{j_2}\} = \sum_{j=1}^{\infty} \psi_j(t_1)\psi_j(t_2) = \delta(t_1 - t_2),$$

where δ is the Dirac distribution (or ‘ δ -function’).

Substituting equation (16) into equation (13), one gets a general representation of Brownian motion

$$B(t) = \sum_{j=1}^{\infty} \hat{a}_j \int_0^t dt' \psi_j(t'). \tag{17}$$

For instance, taking the Fourier basis on the unit interval,

$$\psi_j(t) = \sqrt{2} \cos(\pi(j - 1/2)t), \quad (j = 1, 2, 3, \dots),$$

one gets the Karhunen–Loève expansion of Brownian motion [28]:

$$B(t) = \sqrt{2} \sum_{j=1}^{\infty} \hat{a}_j \frac{\sin\left(\pi\left(j - \frac{1}{2}\right)t\right)}{\left(j - \frac{1}{2}\right)\pi}. \tag{18}$$

A different Fourier basis on the unit interval,

$$\psi_j(t) = \begin{cases} \sqrt{2} \cos(\pi jt), & (j > 0), \\ 1, & (j = 0), \end{cases}$$

yields another expansion

$$B(t) = \hat{a}_0 t + \sqrt{2} \sum_{j=1}^{\infty} \hat{a}_j \frac{\sin(\pi jt)}{\pi j}, \tag{19}$$

where the second term is the Karhunen–Loève expansion of the Brownian bridge $\hat{B}(t)$. In general, Karhunen–Loève expansions can be constructed for a broad class of centered square-integrable stochastic processes with continuous covariance functions [28].

2.6. Basic properties of Brownian motion

The Haar wavelet decomposition (7) and the general representation (17) have been explicitly constructed in order to reproduce the basic properties of Brownian motion. Alternatively, one could first postulate such a representation as a definition of Brownian motion and then check that the basic properties are fulfilled. To illustrate this point, we check several properties. Since this section is a little technical, it can be skipped at first reading.

(i) Brownian motion is a Gaussian process with independent increments. First, since $B(t)$ is a linear combination (17) of Gaussian variables, it is Gaussian. Let $t_1 < t_2$ and $t_3 < t_4$ define two increments, $B(t_2) - B(t_1)$ and $B(t_4) - B(t_3)$. If $t_2 \leq t_3$ (i.e., the increments do not ‘overlay’), they are independent (a similar statement holds if $t_4 \leq t_1$ by symmetry). To prove this statement, we decompose the unit interval as

$$[0, 1] = [0, t_1] \cup [t_1, t_2] \cup (t_2, t_3) \cup [t_3, t_4] \cup (t_4, 1], \quad (20)$$

(if one of the subintervals $[0, t_1]$, (t_2, t_3) or $(t_4, 1]$ is empty, it is ignored). The basis $\{\psi_j(t)\}$ in equation (17) can be chosen as a direct product of the Haar eigenbases on each subinterval. For instance, $\{\psi_m^{[t_1, t_2]}\}_{m=0,1,2,\dots}$ is the Haar basis of $L^2([t_1, t_2])$ which is extended to $[0, 1] \setminus [t_1, t_2]$ by zeros. In this particular representation, one has

$$\begin{aligned} B(t_2) - B(t_1) &= \int_{t_1}^{t_2} dt' \sum_{j=1}^{\infty} \hat{a}_j \psi_j(t') \\ &= \int_{t_1}^{t_2} dt' \sum_{m=0}^{\infty} \hat{a}_m^{[t_1, t_2]} \psi_m^{[t_1, t_2]}(t') = \hat{a}_0^{[t_1, t_2]} \sqrt{t_2 - t_1}. \end{aligned} \quad (21)$$

In the second equality, the Haar functions from other subintervals (except $[t_1, t_2]$) vanish by construction. In turn, all the Haar functions $\psi_m^{[t_1, t_2]}(t')$ on $[t_1, t_2]$ (with $m > 0$) vanish after integration due to their orthogonality to a constant. The only remaining contribution is the constant term which has the unit $L^2([t_1, t_2])$ norm: $\psi_0^{[t_1, t_2]}(t) = (t_2 - t_1)^{-1/2}$. Integrating this term, one gets the right-hand side of equation (21) which shows that the increment $B(t_2) - B(t_1)$ is a Gaussian variable with mean zero and variance $t_2 - t_1$, as expected. The same representation for $[t_3, t_4]$ yields

$$\begin{aligned} B(t_4) - B(t_3) &= \int_{t_3}^{t_4} dt' \sum_{j=1}^{\infty} \hat{a}_j \psi_j(t') \\ &= \int_{t_3}^{t_4} dt' \sum_{m=0}^{\infty} \hat{a}_m^{[t_3, t_4]} \psi_m^{[t_3, t_4]}(t') = \hat{a}_0^{[t_3, t_4]} \sqrt{t_4 - t_3}, \end{aligned} \quad (22)$$

and the random weights $\hat{a}_0^{[t_1, t_2]}$ and $\hat{a}_0^{[t_3, t_4]}$ are independent by construction. As a consequence, the increments $B(t_2) - B(t_1)$ and $B(t_4) - B(t_3)$ are independent.

(ii) The mean and covariance of Brownian motion are:

$$\mathbb{E}\{B(t)\} = 0, \quad \mathbb{E}\{B(t_1)B(t_2)\} = \min\{t_1, t_2\}. \quad (23)$$

The first statement is obvious from equation (17) given that all weights have mean zero. To prove the second statement, one assumes $t_2 > t_1$ and considers

$$\mathbb{E}\{B(t_1)B(t_2)\} = \mathbb{E}\{[B(t_1) - B(0)][B(t_2) - B(t_1)]\} + \mathbb{E}\{[B(t_1) - B(0)]^2\}.$$

The first term vanishes due to the independence of increments (and $B(0) = 0$), while the second term is equal to t_1 according to equation (21).

(iii) Brownian motion is continuous but nowhere differentiable almost surely. The proof relies on the simple fact that the Gaussian weights a_{nk} cannot be too large, e.g., the probability that $|a_{nk}| > n$ decays extremely fast (as $e^{-n^2/2}$ for large enough n). In turn, the norm of the tent functions decreases exponentially which ensures the continuity of Brownian motion and the convergence of a partial sum approximation in equation (7) or similar expressions to Brownian motion. Moreover, the remainder of this approximation decreases exponentially fast with the truncated scale N (for technical details, see appendix C).

2.7. Alpert–Rokhlin wavelets

As we mentioned in section 2.1, taking a particular orthonormal basis is equivalent to choosing a way to connect successive positions of Brownian motion at a finite scale δ . The Haar wavelets and the resulting tent functions present the simplest way of connection by linear segments. Such a piecewise linear approximation of Brownian motion introduces

singularities at the connection nodes (corners). For some problems, it is convenient to deal with a smooth approximation of Brownian motion at finite scales (although a true Brownian trajectory, the limiting curve, remains nowhere differentiable). For this purpose, one can use the Alpert–Rokhlin multiwavelet basis [29–32]. This basis is generated by a set of q functions $\phi^{q,1}(t), \dots, \phi^{q,q}(t)$ which are supported on the interval $[0, 1]$, are piecewise polynomials of degree $q - 1$ on $[0, 1/2]$ and on $[1/2, 1]$, and satisfy the moment cancellation conditions

$$\int_0^1 dt t^k \phi^{q,p}(t) = 0, \quad \begin{matrix} k = 0, 1, \dots, q - 1, \\ p = 1, 2, \dots, q. \end{matrix} \quad (24)$$

These mother functions generate the Alpert–Rokhlin multiwavelets of order q by translations and dilations:

$$\phi_{nk}^{q,p}(t) = 2^{-n/2} \phi^{q,p}(2^n t - k) \quad \begin{matrix} n = 0, 1, 2, \dots, \\ k = 0, 1, 2, \dots, 2^n - 1. \end{matrix}$$

The set of functions $\{\phi_{nk}^{q,p}(t)\}$, completed by the set of orthonormal polynomials of order $m < q$, forms a complete basis of $L^2([0, 1])$. This completion is necessary because all mother functions $\phi^{q,p}(t)$ (and thus all $\phi_{nk}^{q,p}(t)$) are orthogonal by construction to all polynomials of order $m < q$. Similarly, the Haar wavelets were completed by a constant function.

When $q = 1$, there is only one mother function $\phi^{1,1}(t)$ defined by equation (11) which generates the Haar wavelets by translations and dilations. For $q = 2$, there are two mother functions (figure 6), satisfying the moment cancellation conditions (24):

$$\phi^{2,1}(t) = \begin{cases} \sqrt{3} (1 - 4t), & 0 \leq t < \frac{1}{2}, \\ \sqrt{3} (4t - 3), & \frac{1}{2} < t \leq 1, \\ 0, & \text{otherwise} \end{cases}$$

$$\phi^{2,2}(t) = \begin{cases} 6t - 1, & 0 \leq t < \frac{1}{2}, \\ 6t - 5, & \frac{1}{2} < t \leq 1, \\ 0, & \text{otherwise} \end{cases}$$

Higher-order mother functions (with $q > 2$) can be constructed through an orthogonalization procedure (see [29, 33] for details and examples).

Note that the wavelet representation of Brownian motion involves the integral of wavelets

$$h_{00}^{q,p}(t) = \int_0^t dt' \phi^{q,p}(t'). \quad (25)$$

For instance, one gets for $q = 2$ (figure 6)

$$h_{00}^{2,1}(t) = \begin{cases} \sqrt{3} t (1 - 2t), & 0 \leq t < \frac{1}{2}, \\ \sqrt{3} (1 - t)(1 - 2t), & \frac{1}{2} < t \leq 1, \\ 0, & \text{otherwise} \end{cases}$$

$$h_{00}^{2,2}(t) = \begin{cases} t (3t - 1), & 0 \leq t < \frac{1}{2}, \\ (t - 1)(3t - 2), & \frac{1}{2} < t \leq 1, \\ 0, & \text{otherwise} \end{cases}$$

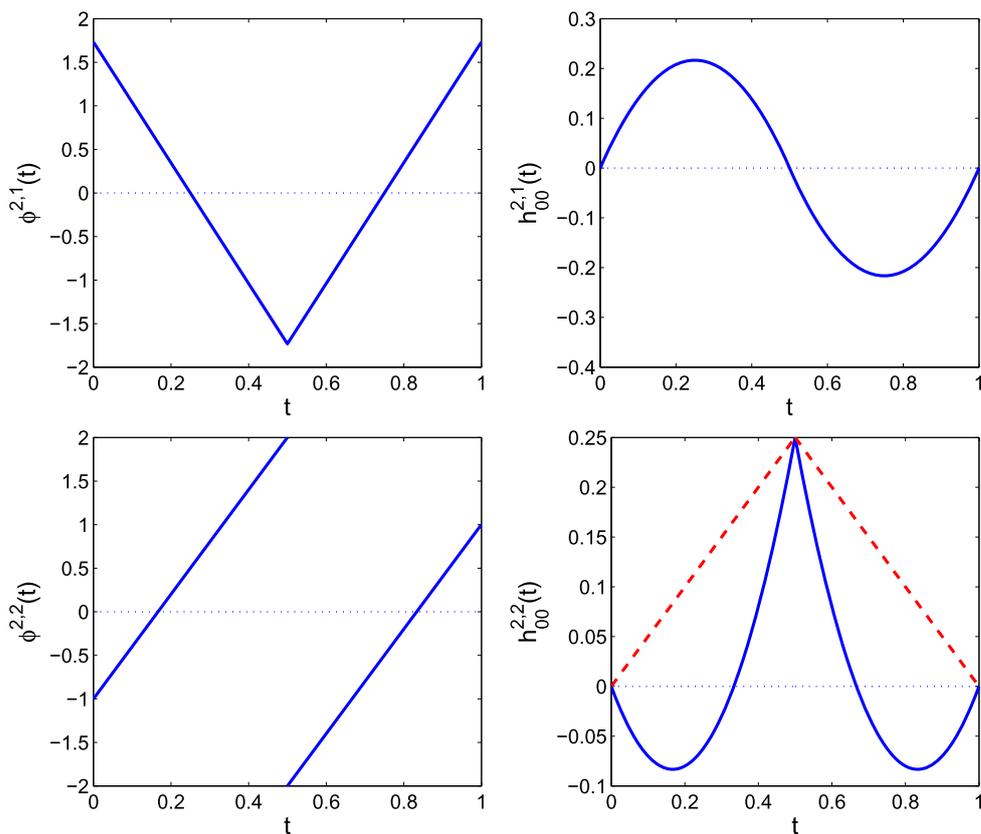


Figure 6. Two mother functions for Alpert–Rokhlin first-order multiwavelets (on the left) and their integrals $h_{00}^{q,p}(t)$ (on the right). The use of these wavelets corresponds to another type of connection (not by linear segments) between successive points of Brownian motion in the refinement procedure. For comparison, the tent function $h_{00}(t)$ is shown in the last plot by a dashed line. Note that horizontal and vertical scales are not matched here.

while the other functions $h_{nk}^{2,1}(t)$ and $h_{nk}^{2,2}(t)$ are obtained by dilations and translations. As a consequence, Brownian motion gets a closed formula in terms of the Alpert–Rokhlin multiwavelets of order 2:

$$B(t) = a_0 t + a_1 \sqrt{3} t(t - 1) + \sum_{n=0}^{\infty} \sum_{k=0}^{2^n-1} \sum_{p=1}^2 a_{nk}^{(p)} h_{nk}^{2,p}(t), \tag{26}$$

where all weights $a_0, a_1, a_{nk}^{(p)}$ are independent $\mathcal{N}(0, 1)$ variables, and the second term is the integral of the linear basis function $\sqrt{3}(2t - 1)$. An extension of this representation to the Alpert–Rokhlin multiwavelet basis of order q is straightforward.

2.8. Numerical implementation

The wavelet representation of Brownian motion can be easily implemented in practice. To carry out computations with a (fixed) desired precision ε , it is sufficient to truncate the first sum in equation (7) or equivalent relations up to $N = \lfloor \log_2(1/\varepsilon) \rfloor$, because higher-order

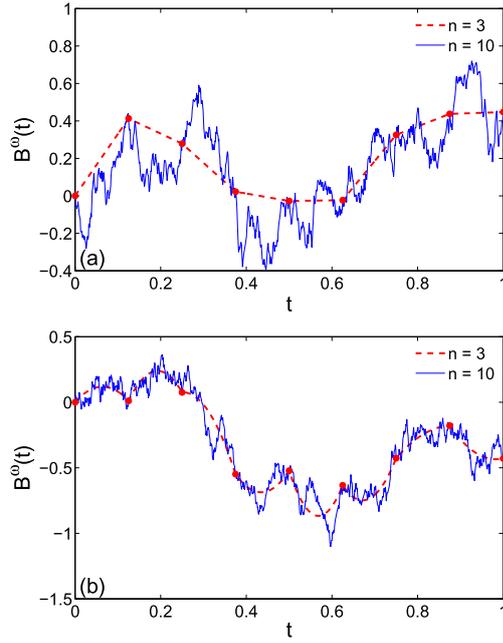


Figure 7. A random Brownian path at scales $n = 3$ (dashed line) and $n = 10$ (solid line) with Haar wavelets (a) and Alpert–Rohkin wavelets with $q = 2$ (b).

terms describe geometrical details at smaller scales. The remainder of this series can be estimated using equation (C.4) (see appendix C). For Haar wavelets, such a truncation qualitatively corresponds to an approximation of Brownian motion by a broken line composed of linear segments of length close to ε , while the Alpert–Rohlin wavelets yield smoother approximations (figure 7). This is a fascinating feature of wavelets that allows one to capture geometrical details at different scales.

A realization of a Brownian path is completely determined by a set of random coefficients a_{nk} (or \tilde{a}_m). In appendix B, we discussed an explicit scheme to generate all these coefficients from a randomly chosen number from the unit interval. In practice, one can use standard routines to generate pseudo-random normally distributed weights a_{nk} (or \tilde{a}_m). The computation of tent functions $h_{nk}(t)$ can be easily implemented. Consequently, the computation of $B(t)$ at any time t requires only $\log_2(1/\varepsilon)$ operations, each of them consisting of finding $h_{nk}(t)$, multiplying it by a_{nk} , and summing their contributions.

It is instructive to compare the wavelet approach to conventional techniques. We consider the computation of all positions $b_k = B(t_k)$ at equidistant times $t_k = k2^{-N}$ ($k = 0, \dots, 2^N$) at some scale $\varepsilon = 2^{-N}$. In a classical scheme, Brownian motion is modeled by a sequence of small random jumps

$$b_0 = 0, \quad b_{k+1} = b_k + 2^{-N/2}a'_k \quad (k = 0, 1, \dots, 2^N - 1), \quad (27)$$

with 2^N independent normally distributed random variables $a'_k \in \mathcal{N}(0, 1)$. Similar computation relying on wavelet representations requires one random variable for a linear shift and 2^k random variables at each scale k , k ranging from 0 to $N - 1$. The total number is then $1 + (1 + 2 + 4 + \dots + 2^{N-1}) = 2^N$. It is not surprising that both schemes require the same degree of randomness to represent a Brownian path at a chosen scale. The wavelet

representation does not reduce the complexity or randomness, but re-organizes the data into a hierarchical structure to facilitate their use. For instance, formula (7) accesses approximate positions of Brownian motion at any time point t , not necessarily t_k . In a classical scheme, one could use a linear interpolation between two neighboring points to get the same result. Again, the wavelets do not bring new features which are not available using conventional techniques, but provide another, structured and efficient, representation.

Although we focused on Brownian motion on the unit interval, the above construction can be extended for any positive t either by rescaling the interval, or by piecing together independent copies B_i of Brownian motion on successive intervals [2]:

$$B(t) = \begin{cases} B_0(t), & (0 \leq t < 1), \\ B_0(1) + B_1(t - 1), & (1 \leq t < 2), \\ B_0(1) + B_1(1) + B_2(t - 2), & (2 \leq t < 3), \dots \end{cases}$$

Finally, an extension to isotropic Brownian motion in \mathbb{R}^d is obtained by taking d independent samples of one-dimensional Brownian motion.

3. Beyond Brownian motion

3.1. Fractional Brownian motion

A similar technique can be applied to construct and study fractional Brownian motion which is also known as a random fractal velocity field [34, 35]. For instance, a random fractal velocity field with the Hurst exponent $H = 1/3$ (defined below) corresponds to the Kolmogorov spectrum in high Reynolds number turbulence [8, 36–38]. Fractional Brownian motion, as a Gaussian stochastic process with long-range correlation, has found numerous applications in different fields, ranging from transport phenomena in porous media [33, 39–42] to financial markets [43].

P. Lévy proposed the first extension of equation (13) by using the Riemann–Liouville fractional integration which can be thought of as a moving average of Gaussian white noise [44]

$$\tilde{B}_H(t) = \frac{1}{\Gamma\left(H + \frac{1}{2}\right)} \int_0^t (t - s)^{H - \frac{1}{2}} dW(s), \tag{28}$$

where $0 < H < 1$ is the Hurst exponent, and $\Gamma\left(H + \frac{1}{2}\right)$ is the normalization factor ($\Gamma(z)$ being the Gamma function). Mandelbrot and van Ness discussed the limitations of this definition (e.g., its strong emphasis on the origin) and proposed to use the Weyl fractional integral that yields [35]

$$B_H(t) = \frac{1}{\Gamma\left(H + \frac{1}{2}\right)} \times \left\{ \int_{-\infty}^0 \left[(t - s)^{H - \frac{1}{2}} - (-s)^{H - \frac{1}{2}} \right] dW(s) + \int_0^t (t - s)^{H - \frac{1}{2}} dW(s) \right\} \tag{29}$$

for $t > 0$ (and similar for $t < 0$). The last representation can also be written as [45]

$$B_H(t) = \int_0^t K_H(t, s) dW(s), \tag{30}$$

where

$$K_H(t, s) = \frac{(t-s)^{H-\frac{1}{2}}}{\Gamma\left(H+\frac{1}{2}\right)} {}_2F_1\left(H-\frac{1}{2}, \frac{1}{2}-H; H+\frac{1}{2}; 1-\frac{t}{s}\right), \quad (31)$$

and ${}_2F_1(a, b; c; z)$ is the hypergeometric function. The ordinary Brownian motion is retrieved at $H = 1/2$ for all cases.

Using the general representation (16) for Gaussian white noise, one obtains

$$B_H(t) = \sum_{j=1}^{\infty} \hat{a}_j \int_0^t ds K_H(t, s) \psi_j(s). \quad (32)$$

The integrals can be evaluated using an appropriate basis $\{\psi_j(t)\}$. Moreover, the moment cancellation property (24) for the Alpert–Rokhlin multiwavelets with a large enough order q guarantees that the integrals in equation (32) are highly localized, yielding a rapid convergence of the above sum. This convergence is a key point for efficient numerical algorithms for simulation of fractional Brownian motion (see [33, 39–41]). Among other numerical methods, we mention alternative wavelet representations [46–49] (e.g., the method by Arby and Sellan is implemented in the Matlab function ‘wfm’), circulant embedding of the covariance matrix [50–52], and the random midpoint displacement method [53] which is often used in computer graphics to generate random two-dimensional landscapes.

In general, the kernel $K_H(t, s)$ can be replaced by any convenient kernel to extend this approach to various Gaussian processes. For instance, setting $K(t, s) = e^{-\theta(t-s)}$ yields the Ornstein–Uhlenbeck process [54, 55]. Similarly, one can deal with various stochastic dynamics generated by Langevin equations [56] while appropriate wavelet decompositions can be constructed for a general class of elliptic Gaussian processes [57].

3.2. Gaussian free field and its extensions

Brownian motion is the integral of Gaussian white noise which, in turn, is obtained as a linear combination of orthonormal functions $\{\psi_j\}$ forming a complete basis of the space $L^2([0, 1])$, with standard Gaussian weights. This construction can be extended to any separable Hilbert space H . However, whatever the functional space H is taken, a linear combination of its orthonormal basis functions with standard Gaussian weights does not belong to this space (the argument is the same as for the L^2 space, the norm of such a linear combination being infinite). In particular, Gaussian white noise is not a function but a distribution. In order to construct reasonable extensions of Gaussian white noise with desired properties, one needs to carefully choose the Hilbert space H . In this section, we briefly discuss two such extensions: the Gaussian free field (GFF) [58] and the fractional Gaussian field (FGF) with logarithmic correlations [59, 60]. The GFF appears as the basic description of massless non-interacting particles in field theories. Both the GFF and FGF constitute important models in different areas of physics, from astrophysics (describing stochastic anisotropy in the cosmic microwave background) to critical phenomena, quantum physics, and turbulence [16–19]. While the ‘sequence’ of random variables of Brownian motion $B(t)$ was naturally parameterized by ‘time’ t (a real number from the unit interval or, in general, from \mathbb{R}), random variables of a field can in general be parameterized by points from a Euclidean domain, a manifold, or a graph. For instance, one can speak about random surfaces which can model landscapes (e.g., mountains) or the ocean’s water surface, in which the height is parameterized by two coordinates. The geometrical structure of ‘smooth’ random surfaces has been thoroughly investigated [61–63], especially for Gaussian fields which are fully characterized by their mean and

covariance. Important examples of smooth Gaussian fields are random plane waves and random spherical harmonics (see [64, 65]). In turn, the GFF and FGF are examples of highly irregular random fields. As Brownian motion can be obtained as the limit of discrete random walks, the Gaussian free field in two dimensions appears as the limit of discrete random surfaces [66, 67]. Fast algorithms for generating Gaussian random processes and fields often involve Fourier and wavelet transforms [68–70].

The GFF is constructed by choosing the Dirichlet Hilbert space $H_{\nabla}(\Omega)$, in which the scalar product of two functions f and g is defined as

$$\langle f, g \rangle_{H_{\nabla}(\Omega)} = \int_{\Omega} dx (\nabla f \cdot \nabla g) \quad (33)$$

for a given Euclidean domain $\Omega \subset \mathbb{R}^d$. When Ω is bounded, an orthonormal basis of this space can be obtained by setting $\psi_j(x) = \lambda_j^{-1/2} u_j(x)$, where $\{u_j(x)\}_{j=1,2,3,\dots}$ are the L^2 -normalized Dirichlet eigenfunctions of the Laplace operator forming a complete basis of $L^2(\Omega)$:

$$\begin{cases} \Delta u_j(x) + \lambda_j u_j(x) = 0 & (x \in \Omega), \\ u_j(x) = 0 & (x \in \partial\Omega), \end{cases} \quad (34)$$

λ_j are the corresponding eigenvalues, and $\partial\Omega$ is the boundary of Ω . The GFF on Ω is defined as

$$F(x) = \sum_{j=1}^{\infty} a_j \psi_j(x), \quad (35)$$

where $a_j \in \mathcal{N}(0, 1)$ are independent Gaussian weights. Given that the eigenvalues λ_j asymptotically grow as $j^{2/d}$ according to Weyl's law [74, 75], the sum in equation (35) is convergent for $d = 1$ (in which case the GFF is simply a Brownian bridge) but diverges in higher dimensions. In the plane, this sum barely misses convergence, being logarithmically diverging. In quantum field theory, it is related to the infra-red divergence for massless particles. As a consequence, $F(x)$ is not a function but a distribution for $d > 1$. Being a linear combination of normal variables, the field $F(x)$ is Gaussian and thus is fully characterized by its mean $\mathbb{E}\{F(x)\} = 0$ and the covariance

$$\mathbb{E}\{F(x_1)F(x_2)\} = \sum_{j=1}^{\infty} \lambda_j^{-1} u_j(x_1)u_j(x_2) = G(x_1, x_2), \quad (36)$$

where the right-hand side can be recognized as the Green function $G(x_1, x_2)$ of the Laplace operator in the domain Ω . Alternatively, one could define the GFF by setting the covariance equal to the Green function (in which case the definition holds even for unbounded domains). Strictly speaking, since the sum in equation (35) diverges for $d > 1$, the GFF should be treated as a distribution by its action on every fixed test function $\phi(x)$. In particular, the covariance should be given by the covariance of the actions of F on two test functions ϕ_1 and ϕ_2 :

$$\mathbb{E}\{F\phi(x_1), F\phi(x_2)\} = \int_{\Omega \times \Omega} dx_1 dx_2 G(x_1, x_2) \phi_1(x_1) \phi_2(x_2). \quad (37)$$

An extension to the GFF with the inhomogeneous Dirichlet boundary condition, i.e., $F(x) = f(x)$ on $\partial\Omega$ for a given $f(x)$, is easily obtained by adding to equation (35) the harmonic function $u_0(x)$ satisfying $u_0(x) = f(x)$ on the boundary. In particular, in one dimension, harmonic functions are linear so that Brownian motion can also be understood as

an extension of a Brownian bridge to connect $B(0) = 0$ and $B(1) = \hat{a}_0$ (i.e., with a random inhomogeneous boundary condition $B(1) = \hat{a}_0$).

FGFs can be obtained by replacing the gradient operators ∇ in the scalar product (33) by fractional Laplacians²[71].

$$\langle f, g \rangle_{H_\nabla^\nu(\Omega)} = \int_\Omega dx \left((-\Delta)^\nu f \cdot (-\Delta)^\nu g \right), \quad (38)$$

for a positive ν . The Dirichlet Hilbert space is retrieved for $\nu = 1/2$. In a similar way, the functions $\psi_j(x) = \lambda_j^{-\nu/2} u_j(x)$ form a basis of the space $H_\nabla^\nu(\Omega)$, from which the FGF is constructed as in equation (35). This sum is convergent for $\nu > d/4$ and divergent otherwise. In the particular case $\nu = d/4$, the FGF is logarithmically divergent in all dimensions. Note that the FGF coincides with the GFF in the plane ($d = 2$). In particular, the FGF with logarithmic correlations on the unit interval reads explicitly

$$F(t) = \sqrt{2} \sum_{j=1}^{\infty} a_j \frac{\sin(\pi j t)}{\sqrt{\pi j}}. \quad (39)$$

Comparing this representation to equation (19), one concludes that the FGF with logarithmic corrections in one dimension appears as the half-derivative of a Brownian bridge:

$$F(t) = \left(\frac{d}{dt} \right)^{1/2} \hat{B}(t), \quad (40)$$

where the half-derivative, $\left(\frac{d}{dt} \right)^{1/2} = (-\Delta)^{1/4}$, is defined through the Laplacian eigenbasis. This formula reveals a very close relation between these two processes which are often considered as distinct objects. Note that the series in equation (40) diverges logarithmically, while the derivative of order $1/2 - \epsilon$ would lead to a converging series. In other words, the Brownian bridge (as well as Brownian motion) belongs to the Hölder space $H_{1/2-\epsilon}$ (for any $\epsilon > 0$) so that its derivatives of any order less than $1/2$ exist, but the derivatives of order $1/2$ and higher do not. As for Brownian motion, the explicit closed formula (35) allows one to sample random realizations of the FGF by picking up a uniform number from the unit interval, i.e., the probability space for this process is nothing other than the unit interval with uniform measure.

4. Restricted diffusion

In this section, we briefly discuss how multiscaling and dyadic decompositions help simulate restricted diffusion. This is a ubiquitous problem in physics (e.g., transport in porous media), chemistry (e.g., heterogeneous catalysis), biology and physiology (e.g., diffusion in cells, tissues, and organs). When Brownian motion is restricted, physico-chemical or biological interactions between the diffusing particle and the interface of a confining medium should be taken into account. For instance, paramagnetic impurities dispersed at a liquid/solid interface

² For a function f on the whole space $\Omega = \mathbb{R}^d$, the fractional Laplacian, also known as the Riesz fractional derivative, can be defined through the Fourier transform, in particular,

$$\left[(-\Delta)^\nu f \right](x) = \mathcal{F}^{-1} \left\{ |k|^{2\nu} \mathcal{F}\{f\} \right\} = \int_{\mathbb{R}^d} \frac{dk}{(2\pi)^d} e^{-ikx} |k|^{2\nu} \int_{\mathbb{R}^d} dx' e^{ikx'} f(x').$$

Even in one dimension, the fractional Laplacian of order $1/2$ differs from the ordinary derivative due to the absolute value $|k|$ (the ordinary derivative corresponding to ik). Note also that the Riesz fractional derivative is different from other fractional derivatives such as Riemann–Liouville or Caputo derivatives [72, 73]. For bounded domains, the Laplacian eigenbasis yields a straightforward definition of $(-\Delta)^\nu$.

cause surface relaxation in nuclear magnetic resonance (NMR) experiments [76, 77]; cellular membranes allow for semi-permeable transport through the boundary [78–80]; a chemical reaction may transform the particle or alter its diffusive properties [81, 82]. While an accurate description of these processes at the microscopic level is challenging, the contact with the interface is very rapid at the macroscopic scale, allowing one to resort to an effective description of the surface transport by an absorption/reflection mechanism [83, 84]. This mathematical process is known as (partially) reflected Brownian motion [85–87].

The presence of a boundary drastically changes the properties of Brownian motion (e.g., the reflecting boundary forces the process to remain inside the domain). As a consequence, earlier wavelet representations cannot directly incorporate the effect of the boundary. Since restricted diffusion is relevant for most physical, chemical and biological applications, various Monte Carlo methods have been developed for simulating this stochastic process, computing the related statistics (e.g., the first passage times [88, 89]), and solving the underlying boundary value problems [90–93]. The slow convergence of Monte Carlo techniques (typically of the order of $M^{-1/2}$ in the number of trials) requires the fast generation of many Brownian paths. The simplest generation of a Brownian path at successive times $\delta, 2\delta, 3\delta, \dots$ by adding normally distributed displacements and checking the boundary effects at each step becomes inefficient in multiscale media. In fact, tiny geometrical details of the medium require the use of comparably small displacements, resulting in a very large number of steps needed to model large-scale excursions.

To overcome this limitation, the concept of fast random walks was proposed [94]. The basic idea consists in adapting displacements to the local geometrical environment, performing as large as possible displacements without violating the properties of Brownian motion. When the walker is at point x , the largest displacement is possible at the distance $|x - \partial\Omega|$ between x and the boundary $\partial\Omega$ of a Euclidean domain $\Omega \subset \mathbb{R}^d$. In fact, the ball $B(x, |x - \partial\Omega|)$ of radius $|x - \partial\Omega|$ does not contain any ‘obstacle’ (e.g., piece of boundary) to the walker. Since Brownian motion is continuous, it must leave the ball before approaching the boundary of the confining domain. The rotation symmetry implies that the exit points are distributed uniformly over the boundary of the ball. Instead of modeling the fully-resolved trajectory of Brownian motion inside the ball, one can just pick at random a point x' on the sphere of radius $|x - \partial\Omega|$ and move the random walker to this new position. The random duration of this displacement can be easily generated [88, 93]. From here, one draws a new ball $B(x', |x' - \partial\Omega|)$, and so on, until the walker approaches the boundary $\partial\Omega$ closer than a chosen threshold. From this point, an appropriate boundary effect (e.g., absorption, relaxation, chemical transformation, permeation, reflection, etc.) is implemented. Due to its efficiency, fast random walk algorithms have been used to simulate diffusion-limited aggregates (DLAs) [95, 96], to generate a harmonic measure on fractals [86, 97, 98], to model diffusion–reaction phenomena in spherical packs [99, 100], to compute the signal attenuation in pulsed-gradient spin-echo experiments [101, 102], etc. In this section, we focus on multiscale tools to estimate the distance, while other aspects of fast random walk algorithms can be found elsewhere [93].

4.1. Distance to a boundary

The efficiency of fast random walk algorithms fully relies on the ability to rapidly estimate the distance between any point (e.g., the current position of the walker) and the boundary. Multiscale dyadic decompositions provide an efficient way to these estimates. To illustrate the idea, we first consider the one-dimensional case and then discuss its straightforward extension to the multidimensional case.

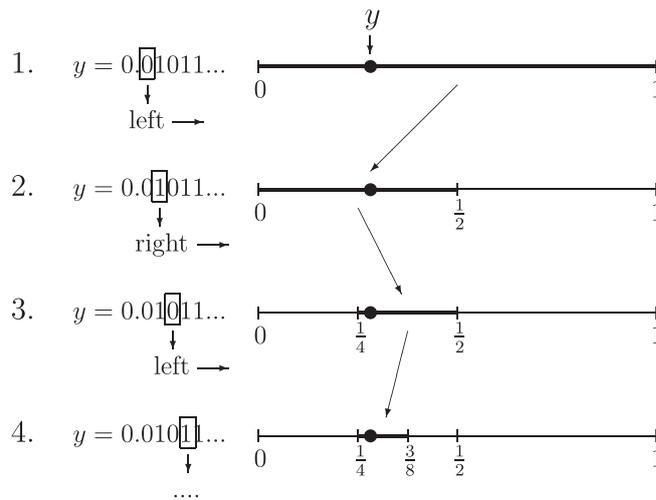


Figure 8. Construction of a dyadic tree of subintervals for a given boundary point y by its binary expansion.

In one dimension, the problem can be formulated as follows: given a set $\{y_m\}$ of N ‘boundary’ points on the unit interval, how can the distance to this set from another point x be estimated in a rapid way? Successive computation of the distances $|x - y_m|$ and finding their minimum is of course the simplest but the slowest way (of the order of N). Instead of computing the distances to all boundary points, one can split the unit interval into two half-intervals, and check the distance to the points belonging to the half-interval that contains x . If the distribution of points y_m on $[0, 1]$ is more or less uniform, this division approximately halves the number of computations. In the same spirit, splitting on subintervals of length $1/4$, $1/8$, etc would reduce the number of computations roughly by factors 4, 8, etc. Using such dyadic decompositions, one needs approximately $\log_2(N)$ splitting to achieve the level where one (or a few) point y_m belongs to the same subinterval as x . The number of computations is then of the order of $\log_2(N)$ (assuming the distribution of boundary points is more or less uniform). Moreover, if computation is carried with a desired precision ε (to consider y_m as ‘pointlike’, one needs $\varepsilon \ll 1/N$), one can continue splitting up to the level $\log_2(1/\varepsilon)$ so that the length of subintervals becomes smaller than ε . Since the points $\{y_m\}$ are stored with precision ε , one cannot distinguish two points at any scale smaller than ε . Consequently, any subinterval of length ε can be either vacant, or occupied by only one point y_m (two points from the same subinterval would be indistinguishable). In this case, the number of computations, $\log_2(1/\varepsilon)$, is actually independent of whether the distribution of points y_m is uniform or not. In other words, this algorithm can be applied for any finite set of points y_m that are all distinguishable at scale ε , i.e., $|y_m - y_{m'}| \geq \varepsilon$ for any m and m' .

4.2. Dyadic decomposition

For practical implementation, the boundary points y_m are used to generate a dyadic tree of subintervals at scales ranging from 1 to $\log_2(1/\varepsilon)$. In fact, one can associate to a boundary point $y \in [0, 1]$ a sequence of dyadic intervals $I_{n, [2^n y]}$ such that $y \in I_{n, [2^n y]}$ at any scale n (figure 8). At each scale n , one chooses the left or the right subinterval depending on whether the n th bit is 0 or 1. Applying this procedure to all boundary points y_m , one generates a dyadic

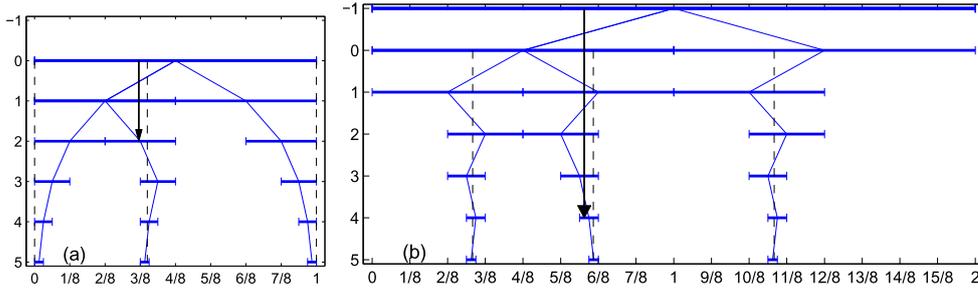


Figure 9. (a) Example of dyadic decomposition of the unit interval with three boundary points $\{0, 0.4, 1\}$ (shown by vertical dashed lines) and the related tree of subintervals at five levels ($\epsilon = 2^{-5}$). For a test point $x = 0.37$ (shown by arrow), one uses its binary expansion $x = 0.01011\dots$ to navigate over the tree. The descent is stopped at level $n = 2$ since $I_{21} = [1/4, 1/2]$ is the largest interval containing both 0.37 and 0.4 (note that $x \notin I_{33} = [3/8, 4/8]$). However, the naive estimate, $2^{-n-1} = 0.125$, of the distance between x and the boundary, $|x - \partial\Omega| = |0.37 - 0.4| = 0.03$, obviously fails. To get the correct lower estimate, one applies the one-third trick by constructing the dyadic decomposition for the boundary points shifted by $1/3$ (b). In this new tree, the descent for the point $x + 1/3$ is stopped at level $n' = 4$. The combined lower estimate $2^{-\max\{n, n'\}}/6 = 2^{-4}/6 \approx 0.0104$ is valid.

tree representing the boundary. Figure 9(a) shows an example with three boundary points $\{0, 0.4, 1\}$ at five scales, from 2^0 to the smallest one $\epsilon = 2^{-5}$. The dyadic decomposition is stored as a tree, where a vertex is associated with a subinterval. Each vertex can be connected to one, two, or three other vertices (see figure 9(a)). The ‘height’ of the vertex from the ‘root’ determines the scale of the corresponding subinterval.

Once the dyadic decomposition for a boundary is constructed, it can be used to estimate the distance to the boundary from any point x . In fact, one can easily (and very rapidly) find the smallest subinterval $I_{n, [2^n x]}$ containing x and at least one boundary point. For this purpose, one starts from the ‘root’ vertex and descends through the tree using the bits of x to choose the left or right edges at each scale. The descent is stopped when there is no edge to follow (figure 9(a)). Once the smallest common interval $I_{n, [2^n x]}$ is found, there are two options: either $n = \lfloor \log_2(1/\epsilon) \rfloor$ so that the point x is indistinguishable from some boundary point at scale ϵ , and the distance estimate is set to 0; or $n < \lfloor \log_2(1/\epsilon) \rfloor$, and the distance to the boundary can be roughly estimated as 2^{-n-1} since the next subdivision must separate the point x from the boundary points. However, this simplistic argument fails when the point x and the closest boundary point lie on opposite sides of the midpoint of the smallest common interval $I_{n, [2^n x]}$, but very close to each other. Although the next subdivision separates these two points, the distance between them can be arbitrarily small so that the naive estimate 2^{-n-1} is wrong, as illustrated in figure 9(a). To get the correct lower estimate, one can apply the so-called ‘one-third trick’.

4.3. The one-third trick

The one-third trick can be easily illustrated for two points x and y . Let $I_{n, [2^n x]}$ be the smallest common interval containing both points x and y . Suppose that the distance between these points is smaller than $2^{-n}/6$. We consider the points $x' = x + 1/3$ and $y' = y + 1/3$ (shifted by $1/3$), and determine their smallest common interval $I_{n', [2^{n'} x']}$. As shown in

appendix D, the distance between the points x' and y' (and thus between the points x and y) is larger than $2^{-n'}/6$. It is thus sufficient to find the smallest common intervals for the pair x, y and its shifted counterpart x', y' , and the distance between the points is bounded below as

$$|x - y| = |x' - y'| \geq 2^{-\max\{n, n'\}}/6. \quad (41)$$

This simple fact allows one to rapidly estimate the distance to boundary points. Let us consider a new boundary $\partial\Omega'$ which is obtained by shifting the old one by $1/3$: $\partial\Omega' = \{x \in \mathbb{R} : x - 1/3 \in \partial\Omega\}$. For the new boundary, another dyadic tree of sub-intervals is constructed (figure 9(b)) to estimate the distance between $\partial\Omega'$ and the shifted point $x' = x + 1/3$ at a given scale ε . While this construction may appear redundant at first thought because $|x - \partial\Omega| = |x' - \partial\Omega'|$, the crucial point is that we search for a lower estimate at a finite scale at which two dyadic trees are different. Performing the descent over both dyadic trees (using the binary expansions of x and x' , respectively), one identifies the level n (resp. n') of the smallest common interval of x (resp. x') and the closest boundary point (resp. shifted closest boundary point). The lower estimate of the distance is then (see appendix D):

$$|x - \partial\Omega| = |x' - \partial\Omega'| \geq 2^{-\max\{n, n'\}}/6. \quad (42)$$

4.4. Higher dimensions

Similar constructions are applicable in higher dimensions which are more relevant for applications. In the simplest approach, the Euclidean distance between two points $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$ in \mathbb{R}^d can be bounded from below as

$$|x - y| = \left(\sum_{k=1}^d (x_k - y_k)^2 \right)^{1/2} \geq \max_{1 \leq k \leq d} \{|x_k - y_k|\}, \quad (43)$$

i.e., by the largest distance among all projections. This inequality reduces the computation back to the one-dimensional setting. However, since the maximum can be achieved along any direction, one needs to repeat the one-third trick in all directions.

The practical implementation consists of several steps. First, a prescribed boundary $\partial\Omega$ is discretized into a set of boundary points at a desired scale N (i.e., $\partial\Omega$ is approximated by a set of hypercubes of size 2^{-N}). Second, $2^d - 1$ replicas of the boundary points are obtained by shifting the original set by all vectors whose coordinates are either 0 or $1/3$. There are $2^d - 1$ such vectors (excluding the zero vector) that produce $2^d - 1$ shifted boundaries. Third, the binary expansions of the coordinates of the boundary points are used to construct a 2^d -adic tree which replaces the dyadic tree used in one dimension. Similar construction is repeated for all shifted boundaries, resulting in total in 2^d trees.

Once this pre-construction is completed, one can use the trees to estimate the distance from an arbitrary point x to the boundary $\partial\Omega$. As previously, the binary expansions of the coordinates of the point x navigate the descent over the first tree to determine the scale n of the smallest hypercube that contains x and at least one boundary point. Similarly, the binary expansions of the coordinates of the shifted point x' are used to descend over the corresponding shifted tree to get the scale n' . Performing 2^d descents over all trees, one can estimate the distance between x and $\partial\Omega$ as previously (see appendix D):

$$|x - \partial\Omega| \geq 2^{-\max\{n, n', n'', \dots\}}/6, \quad (44)$$

where n, n', n'', \dots are the scales for each tree.

Although the complexity of this basic algorithm rapidly grows with the space dimension d , it remains very fast in two and three dimensions. The quality of estimation and the efficiency can be further improved, in particular, by implementing random rotations and translations of the boundary. Although these transformations preserve the distance, they can improve the lower bound at a finite scale. Note that one can use other multiscale constructions and related searchable data structures such as Whitney decompositions of the computational domain (in which the size of each square (or cube) paving the domain is comparable to the distance to the boundary), quadtrees (or Q-trees), k-d trees, multigrid methods, etc [103, 104].

4.5. Overall efficiency

The advantages of multiscale dyadic trees are numerous: the simplicity of construction, the generality of boundary shapes, the rapidity of distance estimation, the flexibility for shape modification, and low memory usage. In fact, for a given precision ε , the storage of a dyadic tree requires at worst $N \log_2(1/\varepsilon)$ intervals (i.e., $\log_2(1/\varepsilon)$ levels for each boundary point, for N points). In practice, this number is much smaller since many boundary points share the same interval at larger scales (e.g., the interval of size l is shared by all boundary points).

The crucial point is that the geometrical structure of the boundary does not matter at all: the method works for a random Cantor dust as well as for a circle. Moreover, the tree-like representation is highly adaptive, allowing one to modify the boundary from one set of simulations to another (or even from one run to another). This feature can be very useful to study diffusion-controlled growth processes like DLA or transport phenomena in domains with moving boundaries.

5. Conclusions

In this review, we revised the multiscale construction of Gaussian processes and fields. First, the Haar wavelet representation of Brownian motion was explicitly constructed as a natural way to refine the geometrical features of a Brownian path under magnification. Since the Haar functions form a complete basis in the space $L^2([0, 1])$ and their weights are Gaussian, such a representation can be extended to any complete basis $\{\psi_j(t)\}$ of $L^2([0, 1])$, wavelet-like or not. In other words, the construction of Brownian motion has two separate ‘ingredients’: deterministic functions $\psi_j(t)$ capturing geometrical details, and their random weights \hat{a}_j . The voluntary choice of the functions $\psi_j(t)$ offers a certain freedom and flexibility in dealing with different problems. Qualitatively, this choice determines the way of connecting successive positions of Brownian motion at a given scale. On the other hand, the random weights determine the intrinsic stochastic properties of Brownian motion, independent of our choice of basis $\{\psi_j(t)\}$.

The multiscale construction gives not only a simple closed formula for Brownian motion, but reveals its fundamental properties. For instance, the continuity and non-differentiability of Brownian paths naturally follow from this construction. In addition, we discussed a closed mapping from the unit interval onto the space of Brownian paths. Sampling Brownian paths can therefore be formally reduced to picking up a real number with a uniform measure. This construction does not require elaborate notions from modern probability theory such as Wiener measures, sigma-algebras, filtrations, etc. Although these notions are useful, the explicit multiscale construction is much easier for non-mathematicians.

These concepts are not limited to Brownian motion. We illustrated how fractional Brownian motion, Gaussian free fields, and fractional Gaussian fields can be constructed in a very similar way. Extensions to other Gaussian processes were also mentioned. Finally, we

briefly discussed how multiscale concepts can be used for simulating restricted diffusion (i.e., Brownian motion in confining domains). The dyadic subdivision and the related hierarchical (multiscale) tree of subintervals allow one to rapidly estimate the distance to the boundary and thus to generate random displacements adapted to the local geometrical environment. Such fast random walk algorithms have found numerous applications. As for usual Brownian motion, dyadic decompositions appear as natural tools to store and rapidly access geometrical information, resulting in fast algorithms.

Acknowledgments

DG acknowledges the financial support from the French National Research Agency (ANR Project ANR-13-JSV5-0006-01). DB was partially funded by an EPSRC Fellowship ref. EP/M002896/1.

Appendix A. Conditional law

We explain why the distribution of the position of Brownian motion $y = B(1/2)$ at time $t = 1/2$ under conditions $B(0) = 0$ and $B(1) = x$ is given by the normal law $\mathcal{N}(x/2, 1/4)$. Since the increments of Brownian motion on the unit intervals $(0, 1/2)$ and $(1/2, 1)$ are independent, the joint probability density $\mathbb{P}\{B(1/2) = y, B(1) = x\}$ is simply equal to the product of the probability density $\mathbb{P}\{B(1/2) - B(0) = y\}$ to have the first increment equal to y and the probability density $\mathbb{P}\{B(1) - B(1/2) = x - y\}$ to have the second increment equal to $x - y$. These densities are given by normal laws with variance $1/2$, yielding

$$\begin{aligned} \mathbb{P}\{B(1/2) = y, B(1) = x\} &= \left(\frac{e^{-\frac{y^2/2}{1/2}}}{\sqrt{2\pi} \sqrt{1/2}} \right) \left(\frac{e^{-\frac{(x-y)^2/2}{1/2}}}{\sqrt{2\pi} \sqrt{1/2}} \right) \\ &= \left(\frac{e^{-\frac{(y-x/2)^2/2}{1/4}}}{\sqrt{2\pi} \sqrt{1/4}} \right) \left(\frac{e^{-x^2/2}}{\sqrt{2\pi}} \right). \end{aligned}$$

The second factor is simply the probability density for $x = B(1)$, while the first factor is the conditional probability density we are looking for:

$$\mathbb{P}\{B(1/2) = y \mid B(1) = x\} = \frac{e^{-\frac{(y-x/2)^2/2}{1/4}}}{\sqrt{2\pi} \sqrt{1/4}}.$$

This is the normal distribution with the mean $x/2$ and the variance $1/4$.

Appendix B. Uniform sampling of Brownian paths

In this appendix, we illustrate how all random Gaussian weights \tilde{a}_m can be explicitly related to a single uniformly distributed random number ω . Sampling a Brownian path or another Gaussian process or field is therefore reduced to picking a uniform random number. In other words, we show that the complicated abstract probability space of Brownian paths can have a simple parameterization. However, this construction remains formal and is thus not relevant from a practical point of view.

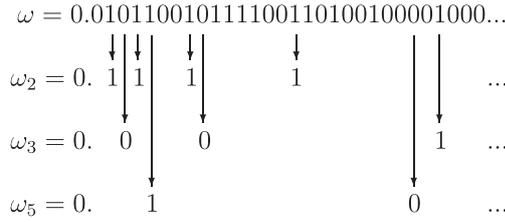


Figure 10. Generation of an infinite sequence of independent uniformly distributed variables $\omega_2, \omega_3, \omega_5, \dots$ using binary expansion of a single number ω from the unit interval. For instance, ω_2 is constituted of the 2nd, 4th, 8th, ... bits of ω .

We consider the binary expansion of a given real number ω from the unit interval

$$\omega = 0. b_1 b_2 b_3 b_4 \dots \tag{B.1}$$

where b_i are equal to 0 or 1 (note that $\omega = 1$ is expanded as 0.111... instead of its equivalent form 1.000...). A uniform picking of ω in $[0,1]$ is equivalent to *independent* random choice of its binary digits (or bits) b_i . Then we choose a prime number p and construct another number ω_p using the bits of ω at positions p, p^2, p^3, \dots

$$\omega_p = 0. b_p b_{p^2} b_{p^3} b_{p^4} \dots \tag{B.2}$$

For example, $\omega_2 = 0. b_2 b_4 b_8 b_{16} \dots$, $\omega_3 = 0. b_3 b_9 b_{27} b_{81} \dots$, etc (figure 10). If p and q are two different prime numbers then ω_p and ω_q are independent as being constructed from separate sets of independent bits b_i . Moreover, if ω is chosen uniformly from the unit interval, then each ω_p is also uniformly distributed on the unit interval. Consequently, a single random number ω gives rise to an infinite sequence $\{\omega_p\}$ of independent uniformly distributed random variables. In a more formal way, the real numbers ω_p can be written as

$$\omega_p = \sum_{n=1}^{\infty} 2^{-n-1} (1 + R(2^n \omega)), \tag{B.3}$$

where $R(x) = (-1)^{|x|}$ is the Rademacher function.

Finally, we need to pass from uniformly distributed to normally distributed variables. For this purpose, we define the inverse $\Phi(x)$ of the error function: for $x \in [0, 1]$, the value y of the function $\Phi(x)$ satisfies

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^y dz e^{-z^2/2} = x. \tag{B.4}$$

Although there is no simple analytic form for the function $\Phi(x)$, many properties can be easily derived, and the whole function can be tabulated with any required precision.

If p_m denotes the $(m + 1)^{\text{th}}$ prime (e.g., $p_0 = 2, p_1 = 3, p_2 = 5$), then we set

$$\tilde{a}_m = \Phi(\omega_{p_m}), m = 0, 1, 2, \dots \tag{B.5}$$

By construction, \tilde{a}_m are independent normally distributed random variables. In other words, equations (B.3), (B.4) map a uniformly distributed ω onto a sequence of Gaussian weights \tilde{a}_m . As a result, the Brownian path $B(t)$ is constructed as a mapping from the unit interval, $\omega \in [0, 1]$, onto the space of real-valued functions (more precisely, the Hölder space $H_{1/2-\epsilon}$ with any $\epsilon > 0$). In other words, any Brownian path is explicitly encoded by the real number ω . Picking the real number ω from the unit interval (with uniform measure) is thus equivalent to choosing a Brownian path (with a Wiener measure). In this representation, the probability space for Brownian motion is nothing other than a unit interval with uniform measure. It is

intuitively much simpler than the classical construction of the probability space by means of a Wiener measure, filtrations, etc. At the same time, this mapping is evidently neither continuous, nor injective (e.g., two numbers ω and ω' that differ only at the 6th bit correspond to the same sequence of Gaussian random numbers). The mapping remains a formal construction whose only purpose was to show the mathematical equivalence between two spaces.

Appendix C. Continuity and non-differentiability of Brownian motion

In this section, we illustrate how the wavelet representation of Brownian motion can be used to prove its basic properties such as continuity and non-differentiability.

C.1. Continuity and convergence

First, we show that a partial sum approximation converges to Brownian motion in both L^∞ and L^2 norms. Let us denote

$$F_n(t) = \sum_{k=0}^{2^n-1} a_{nk} h_{nk}(t) \quad (\text{C.1})$$

so that $B(t) = a_0 t + F_0(t) + F_1(t) + \dots$ according to equation (7). Each function F_n is a sum of hat functions with disjoint supports. To estimate the size of F_n we need some upper bound on a_{nk} . We are going to show that $|a_{nk}| < n$ for all sufficiently large n .

Since $a_{nk} \in \mathcal{N}(0, 1)$ are standard normal variables, we observe that

$$\mathbb{P}(|a_{nk}| \geq n) = 2 \int_n^\infty \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz \leq e^{-n^2/2}$$

for sufficiently large n . This inequality yields

$$\sum_{n=0}^\infty \sum_{k=0}^{2^n-1} \mathbb{P}(|a_{nk}| \geq n) \leq \sum_{n=0}^\infty 2^n e^{-n^2/2} < \infty.$$

By the Borel–Cantelli lemma this implies that $|a_{nk}| < n$ for all but finitely many coefficients a_{nk} . In other words, a.s., there is a finite, but random, N^* such that $|a_{nk}| < n$ if $n > N^*$. For such n we have

$$\|F_n\|_{L^\infty([0,1])} = 2^{-n/2-1} \max_k \{|a_{nk}|\} < n 2^{-n/2-1}.$$

Since the sum of these norms converges, the series $a_0 t + \sum_n F_n(t)$ converges to $B(t)$ in L^∞ norm (uniformly), i.e. for any $\epsilon > 0$

$$\lim_{N \rightarrow \infty} \mathbb{P}\{\|B(t) - B_N(t)\|_{L^\infty([0,1])} > \epsilon\} = 0, \quad (\text{C.2})$$

where

$$B_N(t) = a_0 t + \sum_{n=0}^N \sum_{k=0}^{2^n-1} a_{nk} h_{nk}(t) \quad (\text{C.3})$$

is a partial sum approximation of Brownian motion at the scale 2^{-N} . This proves that Brownian motion is a.s. continuous. Moreover, the remainder of a partial sum approximation at the scale 2^{-N} is exponentially small:

$$\|B(t) - B_N(t)\|_{L^\infty((0,1))} < \sum_{n=N+1}^{\infty} \|F_n(t)\|_{L^\infty((0,1))} < \frac{3(N+3)}{2^{1+N/2}}, \tag{C.4}$$

when N is large enough.

The L^2 convergence can be shown in the same way. We have

$$\|F_n\|_{L^2((0,1))}^2 = \sum_{k=0}^{2^n-1} |a_{nk}|^2 \|h_{nk}\|_{L^2((0,1))}^2 < 2^n n^2 2^{-2n},$$

where we used $|a_{nk}| < n$ for large enough n . This inequality implies the convergence of $\sum_n \|F_n\|_{L^2((0,1))}$ with probability one and hence the series in equation (7) converges a.s. in L^2 norm:

$$\lim_{N \rightarrow \infty} \mathbb{E} \{ \|B(t) - B_N(t)\|_{L^2((0,1))} \} = 0. \tag{C.5}$$

Both statements (C.2, C.5) can be extended to arbitrary spectral representation of Brownian motion. Note also that these statements are applicable pointwise, e.g.,

$$\lim_{N \rightarrow \infty} \mathbb{E} \{ |B(t) - B_N(t)| \} = 0 \tag{C.6}$$

for any $t \in [0, 1]$.

C.2. Nowhere differentiability

Since Brownian motion is a sum of hat functions it is easy to believe that $B(t)$ is not differentiable almost everywhere. One can prove a much stronger statement that $B(t)$ is nowhere differentiable with probability one. The proof goes along the same lines as for convergence (the argument follows the proof from [105]).

We are going to show that a.s. for all t at least one of the two limits,

$$\bar{B}'(t) = \limsup \frac{B(t+h) - B(t)}{h}, \quad \underline{B}'(t) = \liminf \frac{B(t+h) - B(t)}{h}.$$

is infinite. This obviously implies that $B(t)$ is a.s. nowhere differentiable. Note that at local extrema, one of these limits can be finite, so it is not true that *both* of them are always infinite.

Let us assume that there is a t_0 such that both limits are finite at t_0 . This implies that there is a random finite constant M such that

$$\left| \frac{B(t_0+h) - B(t_0)}{h} \right| < M \tag{C.7}$$

for all h . For a given scale n , let k be such that t_0 is between dyadic points $t_{n,k-1}$ and $t_{n,k}$, where $t_{n,k} \equiv 2^{-n}k$. The triangle inequality implies that for any j

$$|B(t_{n,k+j}) - B(t_{n,k+j-1})| < |B(t_{n,k+j}) - B(t_0)| + |B(t_{n,k+j-1}) - B(t_0)| < M(2j+1)2^{-n}.$$

Let E_{nk} be the event that this inequality holds for $j = 1, 2, 3$. Since increments are independent normal variables, one gets $\mathbb{P}\{E_{nk}\} \leq c(2^{-n/2})^3$, where c is a constant. The probability that these inequalities hold for *some* k from 0 to $2^n - 1$ is then bounded by $2^n(c2^{-3n/2}) = c2^{-n/2}$. The sum of these probabilities over n is finite, hence by the Borel–Cantelli lemma, with probability one only finitely many of them will occur. On the other hand the assumed inequality (C.7) implies that infinitely many of E_{nk} will occur. This yields the contradiction and proves that (C.7) cannot be true.

Appendix D. The one-third trick

Although the one-third trick is classical in analysis [106, 107], we provide some explanations which may be instructive for non-experts.

We aim at proving the lower estimate (42) for the distance from a given point x to the set of boundary points $\{y_m\}$. Let I_{nk} be the largest interval that contains x and does not contain any boundary point y_m . This means that there exists a single dyadic point $k2^{-n}$ that separates x and the nearest boundary point $y \in \{y_m\}$. Suppose that the lower estimate (42) does not hold, in particular,

$$|x - y| < 2^{-n}/3. \tag{D.1}$$

We note that the smallest interval containing both x and at least one boundary condition (considered in section 4.3) is twice as large as the largest interval I_{nk} considered here. As a consequence, the right-hand side of the inequality (42) includes an extra factor $1/2$ to compensate this notational difference.

The inequality (D.1) implies that there is no dyadic point $k'2^{-n}$ between the shifted points $x' = x + 1/3$ and $y' = y + 1/3$ or, equivalently, there is no shifted dyadic point $\tilde{k}2^{-n} - 1/3 = \hat{k}2^{-n} - (-1)^n 2^{-n}/3$ between x and y (here k' , \tilde{k} , and \hat{k} are integers). The last statement is elementary because the shift of the original dyadic point $k2^{-n}$ moves it beyond the interval between x and y , while the neighboring dyadic points $(k \pm 1)2^{-n}$ are too far from this interval to be able to move in. Since there is no dyadic point between x' and y' , they belong to the same dyadic interval on the scale 2^{-n} , and the smaller scale $2^{-n'}$ with $n' > n$ is needed to separate them. The distance $|x' - y'|$ should be at least $2^{-n'}/3$, as otherwise the same argument would imply that $n' < n$, in contradiction to the above inequality $n' > n$. We conclude that the inequality (D.1) implies $|x' - y'| \geq 2^{-n'}/3$.

The same argument is valid if we assume that $|x' - y'| < 2^{-n'}/3$. In this case we would get that $n > n'$ and $|x - y| \geq 2^{-n}/3$. This proves that if $|x - y| = |x' - y'|$ is smaller than one of $2^{-n'}/3$ and $2^{-n}/3$, then it is larger than the other, i.e.,

$$|x - y| \geq \min \left\{ 2^{-n'}/3, 2^{-n}/3 \right\} = 2^{-\max\{n, n'\}}/3$$

which completes the proof.

We emphasize that the lower estimate $2^{-\max\{n, n'\}}/3$ is sharp. For instance, setting $x = 2/3 - \varepsilon$ and $y = 1 + \varepsilon$, one gets $n = n' = 0$ from which the lower estimate is $1/3$, while the distance $|x - y| = 1/3 + 2\varepsilon$ can be made arbitrarily close to $1/3$.

The same argument carries over in higher dimensions. For a given set of boundary points (and their $2^d - 1$ shifted copies), one constructs 2^d 2^d -adic trees. For a point $x \in \mathbb{R}^d$, one determines scales n_1, n_2, \dots, n_{2^d} by descending over these trees. Let n be the maximum over these scales. Since all trees provide equivalent representations, we suppose that $n = n_1$. Assume that the lower estimate does not hold, i.e., there exists a boundary point y such that $|x - y| < 2^{-n}/3$. Since the scale n corresponds to the largest hypercube $U_{n,k} = I_{n,k_1} \times \dots \times I_{n,k_d}$ that contains x but does not contain any boundary point, there is at least one direction $i \in \{1, \dots, d\}$ such that $y_i \notin I_{n,k_i}$ (while $x_i \in I_{n,k_i}$). One gets

$$|x_i - y_i| \leq \max_{1 \leq k \leq d} \left\{ |x_k - y_k| \right\} \leq |x - y| < 2^{-n}/3.$$

We consider the shifted point x' and the shifted boundary point y' along the direction i . Repeating the above one-dimensional construction for the direction i , one shows that both x'_i and y'_i belong to the same interval $I_{nj'}$ that contradicts the assumption that n was the maximum.

References

- [1] Revuz D R J and Yor M 1999 *Continuous Martingales and Brownian Motion* (Berlin: Springer)
- [2] Itô K and McKean H P 1965 *Diffusion Processes and Their Sample Paths* (Berlin: Springer-Verlag)
- [3] McKean H P 1969 *Stochastic Integrals* (New York: Academic Press)
- [4] Port S C and Stone C J 1978 *Brownian Motion and Classical Potential Theory* (New York: Academic Press)
- [5] Bass R F 1998 *Diffusions and Elliptic Operators* (New York: Springer)
- [6] Borodin A N and Salminen P 1996 *Handbook of Brownian Motion: Facts and Formulae* (Birkhauser Verlag: Basel-Boston-Berlin)
- [7] Lévy P 1965 *Processus Stochastiques et Mouvement Brownien* (Paris: Gauthier-Villard)
- [8] Mandelbrot B B 1982 *The Fractal Geometry of Nature* (New York: Freeman)
- [9] Weibel E R 1984 *The Pathway for oxygen. Structure and function in the mammalian respiratory system* (Cambridge, Massachusetts, and London: Harvard University Press)
- [10] Jaffard S, Meyer Y and Ryan R D 2001 *Wavelets: Tools for Science and Technology* (Philadelphia: SIAM)
- [11] Daubechies I 1992 *Ten Lectures on Wavelets, CBMS-NSF Regional Conf. Series in Applied Mathematics* vol. 61 (Philadelphia: SIAM)
- [12] Mallat S 2008 *A Wavelet Tour of Signal Processing: The Sparse Way* 3rd edn (New York: Academic Press)
- [13] Mehrabi A R and Sahimi M 1997 *Phys. Rev. Lett.* **79** 4385–8
- [14] Ebrahimi F and Sahimi M 2004 *Trans. Porous Media* **57** 75–102
- [15] Friedrich R, Peinke J, Sahimi M and Tabar M R R 2011 *Phys. Rep.* **506** 87–162
- [16] Bardeen J M, Bond J R, Kaiser N and Szalay A S 1986 *Astrophys. J.* **304** 15–61
- [17] Kobayashi N, Yamazaki Y, Kuninaka H, Katori M, Matsushita M, Matsushita S and Chiang L-Y 2010 *J. Phys. Soc. Japan* **80** 074003
- [18] Fernandez R, Fröhlich J and Sokal A D 1992 *Random Walks, Critical Phenomena, and Triviality in Quantum Field Theory* (Berlin Heidelberg, New York: Texts and Monographs in Physics. Springer)
- [19] Dodelson S 2003 *Modern Cosmology* (Amsterdam: Academic Press)
- [20] Brown R 1828 *Edinburgh New Phil. J.* **5** 358–71
- [21] Perrin J 1908 *Compt. Rendus Herbo. Seances Acad. Sci. Paris* **146** 967
- [22] Perrin J 1909 *Ann. Chim. Phys.* **18** 1–14
- [23] Feller W 1971 *An Introduction to Probability Theory and Its Applications* 2nd edn (New York: John Wiley & Sons) Volumes I and II
- [24] Bouchaud J-P and Georges A 1990 *Phys. Rep.* **195** 127–293
- [25] Metzler R and Klafter J 2000 *Phys. Rep.* **339** 1–77
- [26] Shlesinger M F, Klafter J and Zumofen G 1999 *Am. J. Phys.* **67** 1253–9
- [27] Ciesielski Z 1961 *Trans. Am. Math. Soc.* **99** 403–13 (www.ams.org/journals/tran/1961-099-03/s0002-9947-1961-0132591-2/)
- [28] Loève M 1978 *Probability Theory* 4th edn (New York: Springer-Verlag) Graduate Texts in Mathematics, Vol. 46
- [29] Alpert B 1990 Sparse representation of smooth linear operators *Ph.D. Thesis* (Department of Computer Science Yale University)
- [30] Alpert B Construction of simple multi-scale bases for fastMatrix operations (*Wavelets and Their Applications*) ed Ruskai, Beylkin, Coifman, Daubechies, Mallat, Mayer and Raphael (Boston: Jones & Bartlett) p 121
- [31] Beylkin G, Coifman R and Rokhlin V 1991 *Comm. Pure Appl. Math.* **44** 141–83
- [32] Beylkin G, Coifman R and Rokhlin V 1992 Wavelets in numerical analysis (*Wavelets and Their Applications*) ed Ruskai, Beylkin, Coifman, Daubechies, Mallat, Mayer and Raphael (Boston: Jones & Bartlett) p 181
- [33] Elliott F W and Majda A J 1994 *J. Comput. Phys.* **113** 82–109
- [34] Kolmogorov A N 1940 *C. R. (Doklady) Acad. Sci. URSS (N. S.)* **26** 115–8
- [35] Mandelb B B and Van Ness J W 1968 *SIAM Rev.* **10** 422–37
- [36] Lesieur M 1990 *Turbulence in Fluids* (Boston: Kluwer Academic)
- [37] McComb W 1990 *The Physics of Fluid Turbulence* (Oxford: Clarendon Press)
- [38] Majda A J and Kramer P R 1999 *Phys. Rep.* **314** 237–574

- [39] Elliott F W and Majda A J 1995 *J. Comput. Phys.* **117** 146–62
- [40] Elliott F W, Majda A J, Hornthrop D J and McLaughlin R M 1995 *J. Stat. Phys.* **81** 717–35
- [41] Elliott F W, Hornthrop D J and Majda A J 1997 *J. Comput. Phys.* **132** 384–408
- [42] Bunde A, Kropp J and Schellnhuber H J (ed) 2002 (*The Science of Disasters: Climate Disruptions, Heart Attacks, and Market Crashes*) (Berlin, Heidelberg: Springer-Verlag)
- [43] Mandelbrot B B 1997 *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk* (New York: Springer)
- [44] Lévy P 1953 *Univ. California Publ. in Statist.* **1** 331–90
- [45] Decreusefond L and Üstünel A 1999 *Poten. Anal.* **10** 177–214
- [46] Wornell G W 1990 *IEEE Trans. Inform. Theory* **36** 859–61
- [47] Flandrin P 1992 *IEEE Trans. Inform. Theory* **38** 910–7
- [48] Sellan F 1995 *Compte Rendus Acad. Sci. Paris Série I* **321** 351–8
- [49] Abry P and Sellan F 1996 *Appl. Comput. Harm. Anal.* **3** 377–83
- [50] Davies R B and Harte D S 1987 *Biometrika* **74** 95–102
- [51] Wood A T A and Chan G 1994 *J. Comput. Graph. Stat.* **3** 409–32
- [52] Dietrich C R and Newsam G N 1997 *SIAM Journal Sci. Comput.* **18** 1088–107
- [53] Fournier A, Fussel D and Carpenter L 1982 *Commun. ACM* **25** 371–84
- [54] Uhlenbeck G E and Ornstein L S 1930 *Phys. Rev.* **36** 823
- [55] Risken H 1996 *The Fokker-Planck Equation: Methods of Solution and Applications* 3rd edn (Berlin: Springer)
- [56] Coffey W T, Kalmykov Y P and Waldron J T 2004 *The Langevin Equation: with Applications to Stochastic Problems in Physics, Chemistry and Electrical Engineering* 2nd edn (Singapore: World Scientific Publishing)
- [57] Benassi A, Jaffard S and Roux D 1997 *Rev. Mat. Ibero.* **13** 19–90
- [58] Sheffield S 2007 *Probab. Theory Relat. Fields* **139** 521–41
- [59] Lodhia A, Sheffield S, Sun X and Watson S S 2014 *Fractional Gaussian Fields: a Survey* ArXiv [1407.5598](https://arxiv.org/abs/1407.5598) [math.PR].
- [60] Duplantier B, Rhodes R, Sheffield S and Vargas V 2014 *Log-Correlated Gaussian Fields: An Overview* ArXiv [1407.5605v1](https://arxiv.org/abs/1407.5605v1) [math.PR].
- [61] Adler R J 1981 *The Geometry of Random Fields* (Chichester, New York: Wiley & Sons)
- [62] Adler R J and Taylor J E 2007 *Random Fields and Geometry* (New York: Springer Monographs in Mathematics)
- [63] Vanmarcke E 2010 *Random Fields: Analysis and Synthesis* (Singapore: World Scientific Publishing)
- [64] Berry M V 1977 *J. Phys. A: Math. Gen.* **10** 2083
- [65] Nazarov F and Sodin M 2010 *Random complex zeroes and random nodal lines*, *Proc. Int. Congress Math* Volume III (New Delhi: Hindustan Book Agency) 1450–84
- [66] Kondev J and Henley C L 1995 *Phys. Rev. Lett.* **74** 4580
- [67] Schramm O and Sheffield S 2009 *Acta Math.* **202** 21–137
- [68] Kramer P R, Kurbanmuradov O and Sabelfeld K 2007 *J. Comput. Phys.* **226** 897–924
- [69] Ruan F and McLaughlin D 1998 *Adv. Water Resources* **21** 385–99
- [70] Lang A and Potthoff J 2011 *Monte Carlo Methods Appl.* **17** 195–214
- [71] Chen Z-Q, Kim P and Song R 2010 *J. Eur. Math. Soc.* **12** 1307–29
- [72] Podlubny I 1999 *Fractional Differential Equations* (San Diego: Academic Press)
- [73] Samko S G, Kilbas A A and Marichev O I 2006 *Fractional Integrals and Derivatives: Theory and Applications* (New York: Gordon and Beach)
- [74] Courant R and Hilbert D 1937-1989 *Methods of Mathematical Physics* Vol. I (New York: John Wiley & Sons)
- [75] Grebenkov D S and Nguyen B-T 2013 *SIAM Rev.* **55** 601–67
- [76] Callaghan P T 1991 *Principles of Nuclear Magnetic Resonance Microscopy* (Oxford: Clarendon)
- [77] Grebenkov D S 2007 *Rev. Mod. Phys.* **79** 1077–137
- [78] Alberts B, Bray D, Lewis J, Raff M, Roberts K and Watson J D 1994 *Molecular Biology of the Cell* 3rd edn (New York: Garland)
- [79] Bressloff P C and Newby J M 2013 *Rev. Mod. Phys.* **85** 135–96
- [80] Bénichou O and Voituriez R 2014 *Phys. Rep.* **539** 225–84
- [81] Wilemski G and Fixman M 1973 *J. Chem. Phys.* **58** 4009–19
- [82] Coppens M-O 1999 *Catalysis Today* **53** 225–43
- [83] Sapoval B 1994 *Phys. Rev. Lett.* **73** 3314–6

- [84] Sapoval B 1996 Transport across irregular interfaces: fractal electrodes, membranes and catalysts (*Fractals and Disordered Systems*) ed A Bunde and S Havlin (Berlin: Springer)
- [85] Grebenkov D S 2006 (*Focus on Probability Theory*) ed L R Velle (Hauppauge, NY: NOVA Science Publishers) 135–69
- [86] Grebenkov D S 2006 *Fractals* **14** 231–43
- [87] Singer A, Schuss Z, Osipov A and Holcman D 2008 *SIAM J. Appl. Math.* **68** 844–68
- [88] Redner S 2001 *A Guide to First-Passage Processes* (Cambridge: Cambridge University Press)
- [89] Majumdar S 2005 *Curr. Sci.* **89** 2076–92
- [90] Sabelfeld K K 1991 *Monte Carlo Methods in Boundary Value Problems* (New York, Heidelberg, Berlin: Springer-Verlag)
- [91] Sabelfeld K K and Simonov N A 1994 *Random Walks on Boundary for Solving PDEs* (Utrecht: VSP)
- [92] Milstein G N 1995 *Numerical Integration of Stochastic Differential Equations* (Dordrecht: Kluwer)
- [93] Grebenkov D S 2014 (*First-Passage Phenomena and Their Applications*) ed R Metzler, G Oshanin and S Redner (Singapore: World Scientific Press)
- [94] Muller M E 1956 *Annals Math. Statist.* **27** 569–89
- [95] Meakin P 1985 *J. Phys. A* **18** L661
- [96] Ossadnik P 1991 *Physica A* **176** 454
- [97] Grebenkov D S 2005 *Phys. Rev. Lett.* **95** 200602
- [98] Grebenkov D S, Lebedev A A, Filoche M and Sapoval B 2005 *Phys. Rev. E* **71** 056121
- [99] Torquato S and Kim I C 1989 *Appl. Phys. Lett.* **55** 1847
- [100] Zheng L H and Chiew Y C 1989 *J. Chem. Phys.* **90** 322–7
- [101] Leibig M 1993 *J. Phys. A: Math. Gen.* **26** 3349
- [102] Grebenkov D S 2011 *J. Magn. Reson.* **208** 243–55
- [103] de Berg M, van Kreveld M, Overmars M and Schwarzkopf O 2000 *Computational Geometry: algorithms and Applications* 2nd edn (Berlin: Springer-Verlag)
- [104] Trottenberg U, Oosterlee C W and Schuller A 2001 *Multigrid* (London: Academic Press)
- [105] Mörters P and Peres Y 2010 *Brownian Motion* (Cambridge Series in Statistical and Probabilistic Mathematics) (New York: Cambridge University Press)
- [106] Jones P W 1990 *Invent. Math.* **102** 1–5
- [107] Okikiolu K 1992 *J. London Math. Soc.* **46** 336–48