Random Fourier series and Brownian motion

P. B. Kronheimer, for Math 114

December 5, 2010

1. Introduction

These notes introduce some language from probability with the aim of understanding the statement and proof of the following result, which is due to Wiener.

Theorem 1.1. *Let b_m, $m \in \mathbb{Z}$, be complex-valued, independent random variables, all identically distributed with a standard normal distribution. Then, with probability 1, the terms*

$$\frac{b_m}{m}, \qquad m \in \mathbb{Z},\ m \neq 0,$$

are the Fourier coefficients of a continuous function σ. Furthermore, with probability 1, the Fourier series for σ converges pointwise:

$$\sigma(t) = \sum_{m \neq 0} \left(\frac{b_m}{m}\right) \exp(2\pi i m t), \qquad \forall t.$$

This result is part of the story of Wiener measure and Brownian motion. For a fuller account of these ideas you should turn elsewhere.

2. Definitions

Let $(\Omega, \mathcal{M}, \mathbb{P})$ be a probability space, i.e. a measure space with $\mathbb{P}(\Omega) = 1$.

Event. An event is a measurable subset of Ω.

Probability. The probability of an event E is its measure, $\mathbb{P}(E)$.

Almost sure. An event E is almost sure (or occurs almost surely) if $\mathbb{P}(E) = 1$. We say the event has probability 1.

Random variable. A random variable is a measurable function $a : \Omega \to \mathbb{R}$. Sometimes we will also use complex-valued random variables (which we regard as a pair of real-valued random variables, the real and imaginary parts).

Expectation. The expectation (or expected value) of a random variable $a \in L^1(\Omega)$ is its integral: $\mathbb{E}(a) = \int_\Omega a$. Random variables that are not in L^1 do not have an expectation, though if the negative part of a is integrable and the positive part is not, then one may say that the expected value is $+\infty$. The expectation, or expected value, is also called the mean.

Variance and standard deviation. If a has expectation μ, and $a \in L^2(\Omega)$, then the variance of a is

$$\mathrm{Var}(a) = \mathbb{E}((a - \mu)^2).$$

In particular if a has expectation 0, then $\mathrm{Var}(a)$ is its squared L^2 norm, $\int_\Omega a^2$. The square root of the variance is called the standard deviation.

Covariance. If $\mathbb{E}(a_i) = \mu_i$ for $i = 1, 2$, and a_1, a_2 are both in L^2, then the *covariance* of a_1 and a_2 is

$$\mathrm{Cov}(a_1, a_2) = \mathbb{E}((a_1 - \mu_1)(a_2 - \mu_2)).$$

As a special case, if $\mu_1 = \mu_2 = 0$, then the covariance is the inner product $\int_\Omega a_1 a_2$.

Example. If $a_1, \ldots, a_n$ each have zero mean and unit variance and $\mathrm{Cov}(a_i, a_j) = 0$ for $i \neq j$, then the a_i are an orthonormal set in $L^2(\Omega)$.

Distribution function. The distribution function of a random variable a is the function $F : \mathbb{R} \to [0, 1]$ defined by

$$F(x) = \mathbb{P}(a \leq x).$$

It is a monotonic increasing function. It need not be continuous.

Joint distribution function. Given random variable $a_1, \ldots, a_N$ on Ω, their joint distribution function is the function $F : \mathbb{R}^N \to [0, 1]$ defined by

$$F(x_1, \ldots, x_N) = \mathbb{P}\left(\cap_{i=1}^N (a_i \leq x_i)\right).$$

Independent random variables. Random variables $a_1, \ldots a_N$ are independent if, for all $\lambda_1, \ldots, \lambda_N$ we have

$$\mathbb{P}\big(\cap_{i=1}^N (a_i \leq \lambda_i)\big) = \prod_{i=1}^N \mathbb{P}(a_i \leq \lambda_i).$$

In other words, the joint distribution function, as a function of N variables, is the product of the distribution functions of the individual random variables. This implies in particular that

$$\mathbb{E}(a_1^{p_1} a_2^{p_2} \ldots a_n^{p_n}) = \prod_{i=1}^n \mathbb{E}(a_i^{p_i}), \tag{1}$$

as long as all these expectations are defined (for example if $a_i \in L^p$ for all i and all $p \geq 1$). Thus, if a_1 and a_2 belong to L^2 and are independent, then $\mathrm{Cov}(a_1, a_2) = 0$.

Borel-Cantelli lemma. Let E_n, $n \in \mathbb{N}$, be a sequence of events. For any N, the event

$$\cup_{n \geq N} E_n$$

is, of course, the set of $\omega \in \Omega$ belonging to at least one E_n with $n \geq N$. The event

$$\cap_{N=1}^\infty \cup_{n \geq N} E_n$$

is the set of ω belonging to at least one n in each interval $[N, \infty)$. In other words, it is the set of ω belonging to infinitely many E_n. It is the event "infinitely many E_n occur." We record the following lemma.

Lemma 2.1. *Suppose* $\sum_{n=1}^\infty \mathbb{P}(E_n) < \infty$. *Then*

$$\mathbb{P}\big(\cap_{N=1}^\infty \cup_{n \geq N} E_n\big) = 0.$$

That is, with probability 1*, only finitely many* E_n *occur.*

Proof. We have

$$\mathbb{P}\big(\cup_{n \geq N} E_n\big) \leq \sum_{n \geq N} \mathbb{P}(E_n)$$

which converges to 0 as $N \to \infty$ because of the hypothesis of the lemma. So the nested sequence of events $F_N = \cup_{n \geq N} E_n$ have probability converging to zero. The intersection of this nested sequence therefore has probability zero, by continuity of measure. □

3. Standard normal random variables and Gaussian measures

Normal and standard normal random variables

A random variable a has a *standard normal* distribution if

$$\mathbb{P}(a \le \lambda) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\lambda} e^{-x^2/2}\, dx.$$

If a is a standard normal random variable then it has expectation 0 and variance 1. A random variable a has a normal distribution with mean μ and variance σ^2 if

$$\mathbb{P}(a \le \lambda) = \frac{1}{(2\pi)^{1/2}\sigma} \int_{-\infty}^{\lambda} e^{-(x-\mu)^2/(2\sigma^2)}\, dx.$$

If a is a standard normal random variable, then σa has a normal distribution with mean 0 and variance σ^2.

A *complex-valued* random variable is standard normal if its real and imaginary parts are independent (real) standard normal random variables.

Gaussian measure: independent standard normals r.v.'s.

The standard Gaussian measure on $\mathbb{R}^N$ is the measure

$$\mathbb{P}_N(E) = \frac{1}{(2\pi)^{N/2}} \int_E e^{-|x|^2/2}\, dx_1 \cdots dx_N,$$

defined on the σ-algebra of Borel subsets of $\mathbb{R}^N$. There are two things to notice about the density function that appears here,

$$G_N(x) = \frac{1}{(2\pi)^{N/2}} e^{-|x|^2/2}.$$

First, it is invariant under orthogonal linear transformations of $\mathbb{R}^N$ (those that preserve $|x|^2$). Second, it has a product form:

$$G_N(x) = G_1(x_1)G_1(x_2)\cdots G_1(x_N).$$

This second property means that if we equip $\mathbb{R}^N$ with its Gaussian measure, then the coordinate functions

$$a_i : \mathbb{R}^N \to \mathbb{R}, \quad (i = 1, \ldots, N)$$
$$a_i(x) = x_i$$

are independent standard normal random variables (ISNRV's). The first property means that the same is true for any collection of linear functions a_i obtained from the coordinate functions by an *orthogonal transformation*.

We spell out this last conclusion. Suppose we have an $N \times M$ matrix

$$\lambda = (\lambda_{nm})$$

(not necessarily square) whose columns are orthonormal, so that

$$\sum_{n=1}^{N} \lambda_{nm}\lambda_{nm'} = \begin{cases} 1, & m = m' \\ 0, & m \neq m', \end{cases}$$

or in more compact notation,

$$\lambda^T \lambda = I_M.$$

(the $M \times M$ identity matrix). Then we have:

Proposition 3.1. *Let $a_1, \ldots, a_N$ be a collection of* ISNRV*'s, and let $\lambda = (\lambda_{nm})$ be a real $N \times M$ matrix with orthonormal columns as above. Then the random variables $b_1, \ldots, b_M$ defined by*

$$b_m = \sum_{n=1}^{N} \lambda_{nm} a_n, \quad m = 1, \ldots, M,$$

are also ISNRV*'s. The same holds for complex-valued* ISNRV*'s when λ is a complex matrix with orthonormal columns, i.e. satisfying $\lambda^T \bar{\lambda} = I_M$.*

Infinitely many ISNRVs.

Now consider the space

$$\Omega = \mathbb{R}^{\mathbb{N}},$$

the set of sequence $a = (a_n)_{n \in \mathbb{N}}$. We can define a version of Gaussian measure on Ω as follows. For each N, let $\Pi_N : \Omega \to \mathbb{R}^N$ be the projection onto the first N coordinates. Let $\mathscr{A} \subset \mathscr{P}(\Omega)$ be the collection of subsets of the form

$$E = \Pi_N(S)$$

for $S \subset \mathbb{R}^N$ a Borel subset. For E of this form, define

$$\mathbb{P}(E) = \mathbb{P}_N(S)$$

where $\mathbb{P}_N$ is the Gaussian measure on $\mathbb{R}^N$. This defines a premeasure on $\mathscr{A}$ and hence it extends to a measure on Ω. The coordinate functions

$$a_i : \Omega \to \mathbb{R}, \quad i \in \mathbb{N}$$

are independent standard normal random variables (ISNRVs). We can also construct a probability space with a countable collection of *complex-valued* ISNRV's this way.

Sums of independent normal random variables: convergence in square norm

Let a_n, $n \in \mathbb{N}$, be a collection independent standard normal random variables (on a probability space Ω). Let $(\lambda_n)_{n\in\mathbb{N}}$ be a fixed sequence of real numbers and consider the formal series

$$\sum_1^\infty \lambda_n a_n.$$

In what sense does this series converge? The random variables a_n are functions on a probability space Ω, and as such they are an orthonormal sequence in $L^2(\Omega)$. Our knowledge of Hilbert spaces tells us the following:

Proposition 3.2. *The above series converges in norm in $L^2(\Omega)$ provided that*

$$\sum_1^\infty \lambda_n^2 < \infty,$$

and so defines a random variable

$$s = \sum_1^\infty \lambda_n a_n.$$

Remark. The proposition does not assert that the series converges almost surely (i.e. almost everywhere in Ω), though this is true.

What can we say about the random variable s that arises in the limit? It has expectation 0 and its variance is

$$\begin{aligned}\mathbb{E}(s^2) &= \int_\Omega s^2 \\ &= \sum_1^\infty \lambda_n^2 \\ &= \Lambda^2,\end{aligned}$$

say. In fact, s has a normal distribution; so in particular, if $\sum \lambda_n^2 = 1$, then s is a standard normal random variable. More generally, orthonormal linear combinations of ISNRV's are also ISNRV's, just as in the finite case (cf. Proposition 3.1 above). That is, suppose we have a matrix of scalars $\lambda = (\lambda_{nm})$ (either real or complex) indexed by $(n, m) \in \mathbb{N} \times \mathbb{N}$. Suppose that the columns are orthonormal:

$$\sum_{n=1}^{\infty} \lambda_{nm} \bar{\lambda}_{nm'} = \begin{cases} 1, & m = m' \\ 0, & m \neq m'. \end{cases}$$

Then we have the following result.

Proposition 3.3. *Let $\{a_n\}_{n \in \mathbb{N}}$ be a collection of* ISNRV*'s, either real or complex-valued. Let a matrix of scalars λ_{nm} be given (real or complex respectively) be given, with orthonormal columns as above. Then the random variables*

$$b_m = \sum_{n=1}^{\infty} \lambda_{nm} a_n, \qquad m \in \mathbb{N},$$

are also ISNRV*'s, either real or complex respectively.*

Proof. It is only necessary to check that the first M of the random variables b_m are orthonormal, for all M. We want to deduce the proposition from the case of a finite sum (Proposition 3.1) by taking a limit. Let for each $N \in \mathbb{N}$, let $\lambda^N = (\lambda_{nm}^N)$ be the $N \times M$ matrix obtained taking the first $N \times M$ entries of λ and applying Gram-Schmidt orthogonalization to the columns. Consider the following approximation to b_m:

$$b_m^N = \sum_{n=1}^{N} \lambda_{nm}^N a_n, \qquad m \in \{1, \ldots, M\}.$$

It is straightforward to verify that, for each m, we have

$$\lim_{N \to \infty} b_m^N = b_m$$

in the sense of convergence in $L^2(\Omega)$. Furthermore, for each N, the random variables $(b_1^N, \ldots, b_M^N)$ are ISNRV's by Proposition 3.1. The result of the proposition now follows from the following lemma, which says tells us (in particular) that the property of being ISNRV's is preserved under L^2 limits. □

Lemma 3.4. *Let s_N be a sequence of random variables, all having the same distribution function F. Suppose that $s_N \to s$ as $N \to \infty$, either pointwise almost surely, or in L^p norm for some p. Then s also has distribution function F.*

The same holds for a collection of M random variables $(s_1^N, \ldots, s_M^N)$ which converge as $N \to \infty$,

$$(s_1^N, \ldots, s_M^N) \to (s_1, \ldots, s_M)$$

and which all share the same joint distribution function $F(x_1, \ldots, x_M)$.

Proof. This is an exercise in standard material about distribution functions. We will suppose that the convergence is pointwise, because in the L^p case we can always pass to a subsequence that converges almost surely. We will also do only the case of single variable, because the ideas are the same.

For the proof, we need to understand first that random variables r and s have the same distribution function if and only if

$$\mathbb{E}(\psi \circ r) = \mathbb{E}(\psi \circ s) \tag{2}$$

for all continuous, bounded functions $\psi : \mathbb{R} \to \mathbb{R}$. To see this in one direction, fix $c \geq 0$, and let ψ_n be a uniformly bounded sequence of functions converging pointwise to the discontinuous function $\chi_{(-\infty,c]}$. If the equality (2) holds for all ψ_n, then we can apply the dominated convergence theorem to both sides to see that

$$\mathbb{E}(\chi_{(-\infty,c]} \circ r) = \mathbb{E}(\chi_{(-\infty,c]} \circ s),$$

or in other words,

$$\mathbb{P}(r \leq c) = \mathbb{P}(s \leq c).$$

This holds for all c, which is to say that r and s have the same distribution function. The other direction is left as another exercise in the dominated convergence theorem.

We can now complete the proof of the lemma. From what we have just learned, we know that, for any continuous bounded function ψ, the terms

$$\mathbb{E}(\psi \circ s_N)$$

are independent of N. The integrands $\psi \circ s_N$ are uniformly bounded (by $\sup|\psi|$) and converge pointwise to $\psi \circ s$, because $s_N(\omega) \to s(\omega)$ and ψ is continuous. Applying the dominated convergence theorem again, we have

$$\mathbb{E}(\psi \circ s_N) = \mathbb{E}(\psi \circ s).$$

So by our earlier observation, s and s_N have the same distribution function. $\square$

Estimating normal random variables

Let a be a standard normal random variables. We have

$$\mathbb{P}(a > M) = \frac{1}{(2\pi)^{1/2}} \int_M^\infty e^{-x^2/2}\, dx.$$

How small is this quantity when M is large? One estimate for this integral is:

$$\mathbb{P}(a > M) \leq e^{-M^2/2}. \tag{3}$$

(This estimate is not hard to prove.) When we consider a sequence of standard normal random variables, this estimate is useful when combined with the Borel-Cantelli lemma. It yields the following.

Proposition 3.5. *Let a_n, $n \in \mathbb{N}$, be a sequence of standard normal random variables. Let β be a fixed number larger than 1. Then with probability 1,*

$$|a_n| \leq (2\beta \log(n))^{1/2}$$

for all but finitely many n.

Proof. The inequality (3) gives us

$$\begin{aligned}\mathbb{P}\left(|a_n| \geq (2\beta\log(n))^{1/2}\right) &= 2\mathbb{P}\left(a_n \geq (2\beta\log(n))^{1/2}\right) \\ &\leq 2n^{-\beta}.\end{aligned}$$

Since $\beta > 1$, the sum $\sum n^{-\beta}$ is convergent, so the Borel-Cantelli lemma (Lemma 2.1) applies. □

4. Random Fourier series

We now turn to Theorem 1.1. We will prove the theorem in two steps. In the first step, we discard the usual basis of exponential functions $\exp(2\pi imt)$ for $L^2[0,1]$ and use an orthonormal system that is better suited to the problem. In the second step, we consider effect of the "change of basis", to return to the exponential functions.

Proof of Theorem 1.1: step 1 of 2

Let $\mathbb{T}$ denote the circle $\mathbb{R}/\mathbb{Z}$ and let $L^2(\mathbb{T})$ denote the Hilbert space of periodic locally square-integrable functions with the usual norm,

$$\|f\|_2^2 = \int_0^1 |f|^2.$$

The functions $\exp(2\pi imt)$ for $m \neq 0$ that appear in the statement of Theorem 1.1 are a complete orthonormal system not in the Hilbert space $L^2(\mathbb{T})$ but in the codimension-1 subspace $L^2(\mathbb{T})'$ which we define as

$$L^2(\mathbb{T})' = \{ f \in L^2(\mathbb{T}) \mid \int f = 0 \}.$$

Let us similarly write $C(\mathbb{T})'$ for the continuous periodic functions with integral 0.

Rather than the exponentials, we consider the following complete orthonormal system, $\{d_n\}_{n\in\mathbb{N}}$, in $L^2(\mathbb{T})'$. For $n \geq 1$ we write $n = 2^l + k$ for $l \geq 0$ and $0 \leq k < 2^l$, and we define $d_n(t)$ by

$$d_n(t) = \begin{cases} 2^{l/2}, & 2^l t \in [k, k+1/2) \\ -2^{l/2}, & 2^l t \in [k+1/2, k+1) \\ 0, & \text{otherwise.} \end{cases}$$

We then define $\delta_n(t)$ as the the unique antiderivative of d_n belonging to $C(\mathbb{T})'$. That is, we first define

$$\tilde{\delta}_n(t) = \int_0^t d_n(\tau)\, d\tau$$

(which is the unique antiderivative with $\tilde{\delta}_n(0) = 0$), and we then define

$$\delta_n(t) = \tilde{\delta}_n(t) - \int_0^1 \tilde{\delta}_n.$$

For $n = 2^l + k$ we have

$$\tilde{\delta}_n(t) = \begin{cases} 2^{-l/2}(2^l t - k), & 2^l t \in [k, k+1/2) \\ -2^{-l/2}(k+1-2^l t), & 2^l t \in [k+1/2, k+1) \\ 0, & \text{otherwise.} \end{cases}$$

Note that for $2^l \leq n < 2^{l+1}$, the function $\tilde{\delta}_n$ is supported in an interval of length 2^{-l} and

$$\sup_t |\tilde{\delta}_n(t)| = \tfrac{1}{2} 2^{-l/2}.$$

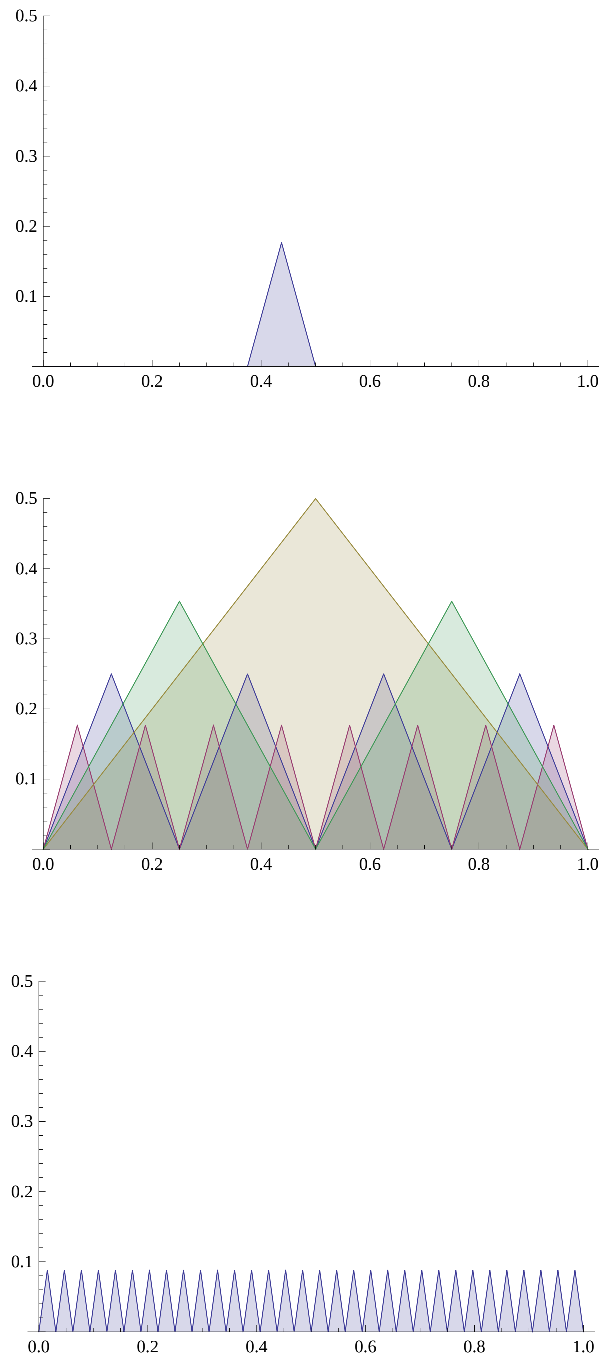

Figure 1: The graph of a typical $\tilde{\delta}_n$ (top, with $n = 11$), the superposition of the graphs of $\tilde{\delta}_n$ for $n = 1, \ldots, 15$ (middle), and the sum $\sum_{32}^{63} \tilde{\delta}_n$ (bottom).

Figure 1 shows the graph of a typical $\tilde{\delta}_n$ (top, with $n = 11$), the superposition of the graphs of $\tilde{\delta}_n$ for $n = 1, \ldots, 15$ (middle), and the sum $\sum_{32}^{63} \tilde{\delta}_n$. The following proposition is the counterpart of Theorem 1.1 for the orthonormal sequence d_n. The convergence is now a little better.

Proposition 4.1. *Let a_n, $n \in \mathbb{N}$, be a sequence of* ISNRV*'s. Then with probability* 1*, the series*

$$\sigma(t) = \sum_0^\infty a_n \delta_n(t)$$

converges uniformly for $t \in [0, 1]$. In particular, σ is almost surely a continuous function of t.

Proof. It is in fact the case that

$$\sum_0^\infty |a_n||\delta_n(t)|$$

converges uniformly for $t \in [0, 1]$ almost surely. This is easy to understand by recalling Proposition 3.5 and then looking at the bottom graph in Figure 1. To spell it out, Proposition 3.5 tells us that, almost surely, $|a_n| \leq 2(\log(n))^{1/2}$ for all but finitely many n, or more straightforwardly but a bit less sharply,

$$|a_n| \leq \log_2 n$$

for all but finitely many n. It is therefore enough to check that the series

$$\sum_1^\infty \log_2(n)|\delta_n(t)|$$

converges uniformly. Group the terms as follows:

$$\begin{aligned} &\sum_{l=0}^\infty \sum_{n=2^l}^{2^{l+1}-1} \log_2(n)|\delta_n(t)| \\ &\qquad \leq \sum_{l=0}^\infty (l+1) \sum_{n=2^l}^{2^{l+1}-1} |\delta_n(t)|. \end{aligned} \tag{4}$$

In the range $2^l \leq n < 2^{l+1}$, the functions $\tilde{\delta}_n$ are supported on disjoint intervals in $[0, 1]$, and each has supremum $2^{-l/2}/2$. The difference between δ_n and $\tilde{\delta}_n$ is even smaller, being of size $2^{-2-3l/2}$. So it is straightforward to verify that

$$\sup_t \left(\sum_{n=2^l}^{2^{l+1}-1} |\delta_n(t)| \right) < (3/4)2^{-(l/2)}.$$

So the series (4) converges uniformly because the sum

$$\sum_{l=0}^{\infty}(l+1)2^{-l/2}$$

is finite. □

The limit $\sigma(t)$ is a bit better than just continuous. With probability 1, it satisfies a Hölder continuity condition with exponent α, for any $\alpha < 1/2$. Recall that, for $\alpha \in (0,1)$, the Hölder-α norm of a function $f : [0,1] \to \mathbb{R}$ is the quantity

$$\|f\|_\alpha = \sup_t |f(t)| + \sup_{t_1 \neq t_2} \frac{|f(t_2) - f(t_1)|}{|t_2 - t_1|^\alpha}.$$

The space of functions for which this norm is finite is the Hölder space $C^{0,\alpha}[0,1]$. The functions $\delta_n(t)$ have uniformly-bounded norm in $C^{0,1/2}$, essentially because they gain height $\epsilon^{1/2}$ in range of size ϵ.

With this in mind, look again at the sum defining $\sigma(t)$, and group the terms once more:

$$\begin{aligned} \sigma(t) &= \sum_{n=0}^{\infty} a_n \delta_n(t) \\ &= a_0 t + \sum_{l=0}^{\infty} \left(\sum_{n=2^l}^{2^{l+1}-1} a_n \delta_n(t) \right) \\ &= a_0 t + \sum_{l=0}^{\infty} A_l(t). \end{aligned} \tag{5}$$

Figure 2 shows a typical A_l for randomly sampled coefficients a_n. We can bound the Hölder-α norm of A_l by

$$\begin{aligned} \|A_l\| &\le M_l \frac{2^{-l/2}}{2^{-\alpha l}} \\ &= M_l 2^{(\alpha - 1/2)l}, \end{aligned}$$

where $M_l = \max |a_n|$ for n in the range $2^l \le n < 2^{l+1} - 1$. As above, with probability 1, we have $M_l \le (l+1)$ for all but finitely many l. So the sum on the last line of (5) converges absolutely in $C^{0,\alpha}$. Thus:

Proposition 4.2. *In the situation of Proposition 4.1, the series*

$$\sigma(t) = \sum_0^{\infty} a_n \delta_n(t)$$

converges in $C^{0,\alpha}$ norm with probability 1, for any $\alpha < 1/2$. The function $\sigma(t)$ is therefore almost surely Hölder continuous with exponent α. □

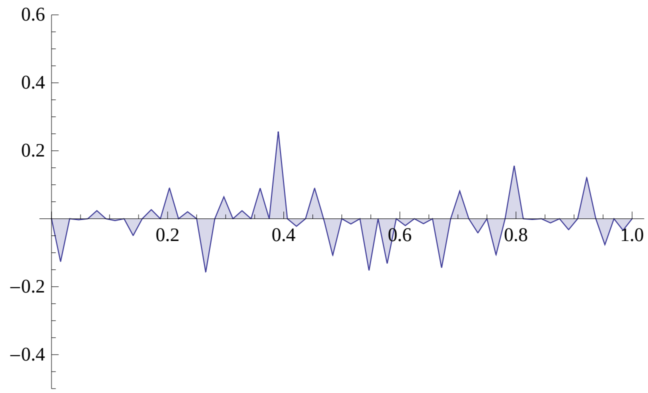

Figure 2: The graph of a typical A_l, with $l = 5$.

Proof of Theorem 1.1: step 2 of 2

We have proved that there is a null set $\mathcal{N} \subset \Omega$ such that for all $\omega \in \Omega \setminus \mathcal{N}$, the series

$$\sum_{n=1}^{\infty} a_n(\omega)\delta_n$$

converges uniformly on $[0, 1]$ to define a Hölder-continuous periodic function σ. To return to the exponential functions, we now examine the Fourier coefficients of σ. Since $\int \sigma = 0$, the Fourier coefficient $\hat{\sigma}(0)$ is zero. For the others, because the series converges uniformly,

$$\hat{\sigma}(m) = \sum_{n=1}^{\infty} a_n(\omega)\hat{\delta}_n(m).$$

Since $\hat{\delta}_n$ is an antiderivative of d_n, we have

$$\begin{aligned}\hat{\delta}_n(m) &= \frac{1}{2\pi im}\hat{d}_n(m)\\ &= \frac{1}{2\pi im}\langle d_n, e_m\rangle_{L^2(\mathbb{T})}.\end{aligned}$$

Let us write

$$\lambda_{nm} = \langle d_n, e_m\rangle.$$

We can think of this as a "change of basis" matrix. We therefore have

$$\hat{\sigma}(m) = \frac{1}{2\pi im}\sum_{n=1}^{\infty} \lambda_{nm} a_n(\omega),$$

and the sum converges for all $\omega \in \Omega \setminus \mathcal{N}$. Let us write

$$b_m : \Omega \setminus \mathcal{N} \to \mathbb{C}$$

for the sum

$$b_m(\omega) = \sum_{n=1}^{\infty} \lambda_{nm} a_n(\omega).$$

The change-of-basis matrix (λ_{nm}) is has orthonormal columns, because the collections $\{d_n\}$ and $\{e_m\}$ are two complete orthonormal systems in $L^2(\mathbb{T})'$. So the random variables b_m are independent standard normal random variables, just like the a_n, by Proposition 3.3.

The Fourier coefficient $\hat{\sigma}(m)$ is thus $b_m/(2\pi i m)$. Because $\hat{\sigma}$ is Hölder continuous, it satisfies Dini's criterion and its Fourier series converges pointwise. We therefore have, for all t,

$$\sigma(t) = \frac{1}{2\pi i} \sum_{m \neq 0} \left(\frac{b_m(\omega)}{2\pi i m} \right) e_m(t).$$

So for all $\omega \in \Omega \setminus N$, the quantities b_m/m are the Fourier coefficients of a continuous function whose Fourier series converges pointwise. This concludes the proof of Theorem 1.1. □

Brownian paths

So far we have been working with $L^2(\mathbb{T})'$ and $C(\mathbb{T})'$, the functions with average value 0 on $\mathbb{T}$. Let us go back the orthonormal system d_n defined above and adjoin the additional element

$$d_0 = 1$$

so as to have a complete orthonormal system in $L^2(\mathbb{T})$. Let $\tilde{\delta}_n(t)$ as before be the indefinite integral of $\tilde{\delta}_n(0) = 0$, so that have

$$\tilde{\delta}_0(t) = t.$$

Note that, unlike the others, $\tilde{\delta}_0$ does not extend to a continuous periodic function. We regard all the functions $\tilde{\delta}_n$ now as continuous functions on $[0, 1]$.

Let a_n be again a collection of ISNRV's, indexed now by $n \in \mathbb{N} \cup \{0\}$, and consider the function

$$B(t) = \sum_{n=0}^{\infty} a_n \tilde{\delta}_n. \tag{6}$$

On the set $\Omega \setminus \mathcal{N}$ where the series for σ converged uniformly, this new function B is related very simply to our previous function $\sigma(t)$:

$$B(t) = a_0 t + \sigma(t) - \sigma(0).$$

Because the Fourier series of σ converges pointwise, we also have

$$\sigma(0) = \sum_{m \neq 0} \frac{b_m}{2piim},$$

which leads to an alternative formula for B in terms of the exponentials:

$$B(t) = b_0 t + \sum_{m \neq 0} b_m \left(\frac{e_m(t) - 1}{2\pi imt} \right).$$

Here $b_0 = a_0$ and b_m $(m \neq 0)$ are as before.

For $\omega \in \Omega \setminus \mathcal{N}$, the function $B(t)$ is a continuous function on $[0, 1]$ and is Hölder continuous with exponent α for all $\alpha < 1/2$.

We can think of $B(t)$ as a "random continuous function of t" on $[0, 1]$. Writing Ω^* for the probability space $\Omega \setminus \mathcal{N}$, we can regard $B(t)$ for each fixed $t \in [0, 1]$ as a random variable (over the probability space Ω^*). To emphasize its dependence on the variable $\omega \in \Omega^*$ (which is usually hidden), we may write it as $B(\omega, t)$. Here are some properties of the random variable $B(t)$.

Proposition 4.3. *As a collection of random variables indexed by $t \in [0, 1]$, the $B(t)$ have the following features.*

- *(a)* $B(0) = 0$.
- *(b) $B(t)$ is almost surely a continuous function of t: that is, for all $\omega \in \Omega^*$, the function $t \mapsto B(\omega, t)$ is continuous.*
- *(c) $B(t)$ has independent increments. This means that if $0 = t_0 < t_1 < \dots t_N = 1$, then the random variables $B(t_k) - B(t_{k-1})$, for $k = 1, \dots, N$, are independent.*
- *(d) For $t_1 < t_2$, the random variable $B(t_2) - B(t_1)$ has a normal distribution with mean 0 and variance $t_2 - t_1$.*

Proof. We only sketch the proof. The first item is clear and the second is what we have devoted our efforts to, in the proofs above. For the second and third properties, we can start with the illustrative simple case that (t_0, t_1, t_2) are $(0, 1/2, 1)$. Since all the $\tilde{\delta}_n$ vanish at these points t_i once $n \geq 2$, we have (from the formulae for $\tilde{\delta}_0$ and $\tilde{\delta}_1$),

$$\begin{aligned} B(t_1) - B(t_0) &= \frac{1}{2}(a_1 + a_0) \\ B(t_2) - B(t_1) &= \frac{1}{2}(-a_1 + a_0). \end{aligned}$$

The random variables on the right, when multiplied by $\sqrt{2}$, are independent standard normal random variables, because they are obtained from a_0 and a_1 by an orthogonal change of basis. So $B(t_1) - B(t_0)$ and $B(t_2) - B(t_1)$ are independent normal random variables, both with mean 0 and standard deviation $1/\sqrt{2}$ (i.e. variance $1/2$, as the proposition asserts).

For the case that $N = 2^l$ and $t_k = k/(2^l)$, the situation is similar, and the proposition can be verified essentially by repeating the above calculation. From this case, one derives the result for any case in which all the t_k are dyadic rationals. Finally, the general case follows from this case by an approximation argument, using the almost sure continuity of $B(t)$ and Lemma 3.4. □

The "independent increments" property described in the proposition is characteristic of a random walk: the change in $B(t)$ in any one interval is independent of its change in any other collection of intervals, as long as the intervals do not overlap. If the exchange rate between the British pound and the US dollar were a Brownian path, then looking at the past exchange rates would tell you nothing about whether the exchange rate might go up or down over the coming month.

In the language of probability theory, a collection of random variables indexed by a parameter t ("time", discrete or continuous), is a *stochastic process*. The particular stochastic process $B(t)$ that we have arrived at is known as the Wiener process, or as Brownian motion. The sample paths $t \mapsto B(t)$ describe the motion of a particle along a single axis when the particle is being buffeted at random from the left and the right. Figure 3 shows a sample Brownian path $B(t)$ (actually a partial sum of the first 2^{14} terms of the series (6), with the coefficients a_m pseudo-randomly generated according to a normal distribution).

Wiener measure

With a slight change of language, we can regard this construction of $B(t)$ or $\sigma(t)$ as defining a probability measure, $\mathbb{P}_W$, on a space of continuous functions: either the

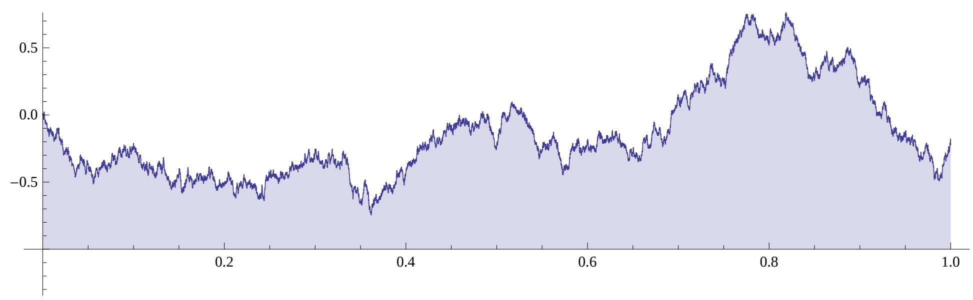

Figure 3: The graph at the top shows the graph of a sample Brownian path $B(t)$, or more accurately a partial sum of the series for one. The graph below shows the currency exchange rate GBP/USD in recent months.

continuous functions $B : [0, 1] \to \mathbb{R}$ with $B(0) = 0$ or the continuous, periodic functions with integral zero, the Banach space $C(\mathbb{T})'$. (The subscript W is for Wiener.) Thus, given a Borel subset E of the Banach space $C(\mathbb{T})'$, we define

$$\mathbb{P}_W(E) = \mathbb{P}\{\, \omega \mid t \mapsto \sigma(\omega, t) \text{ belongs to } E \,\}.$$

With respect to the Wiener measure, we can phrase questions such as, "What is the expected value of $\sup_t \sigma(t)$?" Or, "What is the probability that $\sigma(t) \leq 1$ for all t?"