1 Brownian Motion

Random Walks. Let $S_0 = 0$, $S_n = R_1 + R_2 + \cdots + R_n$, with R_k the Rademacher functions. We consider S_n to be a path with time parameter the discrete variable n. At each step the value of S goes up or down by 1 with equal probability, independent of the other steps. S_n is known as a *random walk*.

To find the rescaled, continuum limit of a random walk, define

$$f_n(k/n) = S_k/\sqrt{n}, \quad k \in \mathbb{Z}$$

and for $k/n \le t \le (k+1)/n$, define $f(t)$ to be linear. For $t = k/n$, the variance is

$$\mathrm{Var}\,(f_n(t)) = \mathbb{E}(S_k^2/n) = \mathbb{E}((R_1 + \cdots + R_k)^2/n) = k/n = t$$

The central limit theorem implies that $f_n(t)$ tends in probability law to a gaussian random variable. In other words,

$$\lim_{n\to\infty} P(a < f_n(t) < b) = \int_a^b \frac{e^{-x^2/2t}}{\sqrt{2\pi t}}\,dx$$

For each n, there is a unique $k = k(n,t)$ such that $k/n \le t \le (k+1)/n$, so that $S_k/\sqrt{n} \le f_n(t) \le S_{k+1}/\sqrt{n}$. Since both $S_k/\sqrt{n}$ and $S_{k+1}/\sqrt{n}$ tend in law to the gaussian with variance t as $n \to \infty$, it's not hard to show that $f_n(t)$ does so as well. More generally, we can describe the probability distribution of the entire path, that is, what happens at many different times.

Theorem 1 *Let $0 \le t_0 < t_1 < t_2 < \cdots < t_m$ and let $I_1 \times I_2 \times \cdots I_m \subset \mathbb{R}^m$ be a rectangle (product of intervals). Let $\sigma_j > 0$ be such that $\sigma_j^2 = t_j - t_{j-1}$. Then*

$$\lim_{n\to\infty} P[(f_n(t_1) - f_n(t_0), f_n(t_2) - f_n(t_1), \ldots, f_n(t_m) - f_n(t_{m-1})) \in I_1 \times \cdots I_m]$$
$$= \int_{I_1\times\cdots\times I_m} \prod_{j=1}^m g_{\sigma_j}(x_j)\,dx_1 \cdots dx_m$$

Proof. If $k_j = k_j(n)$ is the integer such that $k_j/n \le t_j < (k_j+1)/n$, then

$$f(t_j) - f(t_{j-1}) = X_j + O(1/\sqrt{n}) \quad \text{with } X_j = (S_{k_j} - S_{k_{j-1}})/\sqrt{n}$$

For each $j = 1, \ldots, m$, the central limit theorem implies X_j tends to a gaussian with mean zero, variance σ_j^2. Moreover, because for different j, X_j depends on the Rademacher functions R_ℓ, $k_{j-1} + 1 \le \ell \le k_j$, which do not overlap with the Rademacher functions used to define the other X_j, these m random variables are independent. Hence the limit of the joint distribution is the product of the individual limits, the appropriate product distribution of gaussians. Therefore, the m variables $f_n(t_j) - f_n(t_{j-1}) = X_j + O(1/\sqrt{n})$ converge jointly as $n \to \infty$ to the same limit. This proves Theorem 1.

The rest of this section is devoted to explaining how to describe the limiting paths of the random walk, a continuum stochastic process called *Brownian motion.* Brownian motion is a function

$$B : \Omega \times \mathbb{R}_+ \to \mathbb{R}, \quad (\omega, t) \in \Omega \times \mathbb{R}_+$$

First, a few words about notation. When we display the dependence on $\omega \in \Omega$, we will put it into a subscript, $B_\omega(t)$. The main focus is on B_ω, as a random function of t. The sample space Ω is rarely mentioned in probability theory, and the dependence of B on ω is omitted, so that one usually usually writes

$$B(t) = B_\omega(t).$$

The idea of trying to define $B(t)$ as a function of a continuous variable t, as opposed to just a discrete time variable n, is in the same spirit as the overall motivation for differential and integral calculus. Continuum formulas are more transparent and capture better the essence of the phenomenon. For example, we will show that for almost every ω, $B_\omega(t)$ is continuous in t. (Although we won't prove anything this precise, it turns out that the Brownian paths are almost surely of the Hölder class C^α for all $\alpha < 1/2$, almost surely nowhere differentiable, and almost never C^α for $\alpha \ge 1/2$. In higher level courses, one goes on to study so-called stochastic differential equations, in which $dB(t)$ the differential of $B(t)$ plays a role.)

The sample space Ω is barely mentioned because we can identify $\omega \in \Omega$ with B_ω a continuous function. But there is one remark we need to make about the sigma field $\mathcal{F}$. $\mathcal{F}$ will be defined as the sigma field generated by sets of the form

$$\{\omega \in \Omega : (B_\omega(t_1), \ldots, B_\omega(t_m)) \in I_1 \times \cdots \times I_m\}$$

Let $0 \le t_0 < t_1 < \cdots t_m$, and let $R = I_1 \times I_2 \times \cdots \times I_m$. Our goal is to find $B_\omega(t)$ so that

$$\lim_{n\to\infty} P[(f_n(t_1), \ldots, f_n(t_m)) \in R] = P[(B(t_1), \ldots, B(t_m)) \in R] \tag{1}$$

This uniquely specifies the probability law of $B(t)$ on $\mathcal{F}$.

To get a picture of what (1) means, imagine we can simulate as many trials of $B(t)$ as we like on an $m \times m$ pixel video screen. (This turns out to be easy to carry out using the Fourier series formula for $B(t)$ due to Wiener and any fast Fourier transform package, such as the ones contained in Maple, Matlab, or Mathematica.) Suppose that $t_j = j/m$ (pixel width $1/m$) and I_j are chosen from intervals of the form $[Ck/m, C(k+1)/m)$, corresponding to pixel height. The event $(B(t_1), \ldots, B(t_m)) \in I_1 \times \cdots \times I_m$ specifies the graph to accuracy $1/m$, which we can think of as the level of resolution of the pixels. Thus (1) says that the collection of graphs obtained by simulating $B(t)$ look exactly the same as the collection of graphs simulating the functions $f_n(t)$ (up to some negligible probability for n sufficiently large).

Since $B(0) = 0$, the knowledge of the increments $B(t_j) - B(t_{j-1})$ is the same as the knowledge of the values of $B(t_j)$. Let $\sigma_j > 0$, $\sigma_j^2 = t_j - t_{j-1}$, We can always let $t_0 = 0$, so Theorem 1 implies (1) is equivalent to

$$P[(B(t_1) - B(t_0), \ldots, B(t_m) - B(t_{m-1})) \in R] = \int_R \prod_{j=1}^{m} g_{\sigma_j}(x_j)\, dx_1 \cdots dx_m \tag{2}$$

for all $R = I_1 \times \cdots \times I_m$

The main issue is to show that $B(t)$ exists. In around 1920, Norbert Wiener gave a formula for Brownian motion as a random Fourier series. Let a_k, $k = 0, 1, \ldots,$ be independent mean 0, variance 1 Gaussians. Let

$$W(t) = c_0 a_0 t + c_1 \sum_{k=1}^{\infty} a_k \frac{\sin kt}{k}, \quad 0 \le t \le \pi. \tag{3}$$

with $c_0 = \sqrt{1/\pi}$ and $c_1 = \sqrt{2/\pi}$. The rest of this section is devoted to proving Wiener's theorem.

Theorem 2 *Let W be given by* (3)*. Then $W(0) = 0$ and*

a) $B(t) = W(t)$ satisfies (2) *(or equivalently* (1)*) on $0 \le t \le \pi$.*

b) W is almost surely continuous in t.

To obtain Brownian motion on all $t \ge 0$, take a countable number of independent copies of $W_n(t)$ and let

$$B(t) = \begin{cases} W_1(t), & 0 \le t \le \pi \\ W_1(\pi) + W_2(t - \pi), & \pi \le t \le 2\pi \\ W_1(\pi) + W_2(\pi) + W_2(t - 2\pi), & 2\pi \le t \le 3\pi \\ \text{etc.} \end{cases}$$

We begin the proof of Theorem 2 with a lemma about gaussian random variables.

Proposition 1 *If X_k, are independent gaussians with mean 0 and variance σ_k^2 with*

$$\sum_k \sigma_k^2 < \infty$$

then $X_1 + X_2 + \cdots$ converges in $L^2(\Omega)$ to a gaussian random variable with mean zero and variance $\sigma^2 = \sum_k \sigma_k^2$.

Proof. For finite sums we have

$$\begin{aligned}
\mathbb{E}(e^{-i\xi(X_1+X_2+\cdots+X_n)}) &= \prod_{k=1}^n \mathbb{E}(e^{-i\xi X_k}) \\
&= \prod_{k=1}^n e^{-\sigma_k^2\xi^2/2} = e^{-\sum_{k=1}^n \sigma_k^2 \xi^2/2}
\end{aligned}$$

By the uniqueness of the Fourier transform of measures, $S = X_1 + \cdots + X_n$ is gaussian with variance $\sum_{k=1}^n \sigma_k^2$.

For the infinite sum, consider first the partial sums

$$S_n = X_1 + \cdots + X_n$$

For $n > m$,

$$\mathbb{E}(|S_n - S_m|^2) = \sum_{k=m+1}^n \mathbb{E}(X_k^2) = \sum_{k=m+1}^n \sigma_k^2$$

which tends to zero as $m \to \infty$. Therefore, S_n converges in $L^2(\Omega)$ to a random variable S.

Denote by $\rho_n^2 = \sum_1^n \sigma_k^2$. Fix $\epsilon > 0$.

$$P(a < S < b) \le P(a - \epsilon < S_n < b + \epsilon) + P(|S - S_n| \ge \epsilon)$$

and

$$P(|S - S_n| \ge \epsilon) = P(|S - S_n|^2 \ge \epsilon^2) \le \frac{1}{\epsilon^2} \|S - S_n\|_{L^2(\Omega)}^2 \to 0$$

as $n \to \infty$. Therefore, since $\rho_n \to \sigma$,

$$P(a < S < b) \le \lim_{n\to\infty} \int_{a-\epsilon}^{b+\epsilon} g_{\rho_n}(x)dx = \int_{a-\epsilon}^{b+\epsilon} g_\sigma(x)dx$$

Since $\epsilon > 0$ was arbitrary, we have

$$P(a < S < b) \le \int_a^b g_\sigma(x)dx$$

A similar argument gives the same lower bound, proving the proposition.

Covariance. The *covariance* of two random variables X and Y is defined as

$$\mathrm{Cov}\,(X, Y) = \mathbb{E}((X - \mathbb{E}X)(Y - \mathbb{E}Y))$$

Note that

$$\mathrm{Cov}\,(\sum_j X_j, \sum_k X_k) = \sum_{j,k} \mathrm{Cov}\,(X_j, X_k)$$

The polarization method says that to determine the covariance, of a family $X_1, X_2, \ldots, X_n$ of random variables is suffices to know

$$\mathrm{Var}\,(\sum_j a_j X_j) = \sum_{j,k} a_j a_k \mathrm{Cov}\,(X_j, X_k)$$

for all choices of a_j.

The mean and variance determine the distribution of a single gaussian random variable. The analogous statement for several variables involves the *covariance matrix* $\mathrm{Cov}\,(X_j, X_k)$. We will formulate it as follows.

Lemma 1 *Let $X = (X_1, \ldots, X_m)$ be independent gaussian random variables with mean zero. Let $A = (a_{jk})$ be an invertible (real-valued) matrix and define $Y = (Y_1, \ldots, Y_m)$ by*

$$Y_j = \sum_{k=1}^{m} a_{jk} X_k.$$

Then for every $a = (a_1, \ldots, a_m) \in \mathbb{R}^m$,

$$a \cdot Y = \sum_{j=1}^{m} a_j Y_j$$

is a gaussian random variable with mean 0. Conversely, if $Z = (Z_1, \ldots, Z_m)$ are random variables such that for every $a = (a_1, \ldots, a_m)$,

$$a \cdot Z = \sum_{j=1}^{m} a_j Z_j$$

is a gaussian random variable with mean 0 and covariances coincide with those of Y,

$$\mathbb{E}(Z_j Z_k) = \mathbb{E}(Y_j Y_k)$$

then the joint distribution of $Z = (Z_1, \ldots, Z_m)$ is the same as that of Y. In other words,

$$P(Z \in E) = P(Y \in E) \quad \textit{for all Borel sets } E \subset \mathbb{R}^m$$

Moreover, $A^{-1}Z$ has the same probability distribution as X.

Proof. Denote by V the covariance matrix of $X = (X_1, \ldots, X_m)$, that is $v_{jk} = \mathbb{E}(X_j X_k)$. Then V is diagonal with entries $\sigma_1^2, \ldots, \sigma_m^2$ (the variance of $X_1, \ldots, X_m$, respectively) along the diagonal. Denote the covariance matrix of Y by $C = (c_{jk})$, that is, $c_{jk} = \mathrm{Cov}\,(Y_j, Y_k) = \mathbb{E}(Y_j Y_k)$. Then

$$c_{jk} = \mathbb{E}(Y_j Y_k) = \sum_{\ell} a_{j\ell} a_{k\ell} \sigma_\ell^2 \implies \quad C = (c_{jk}) = AVA^T$$

The random variable $a \cdot Y$ is a linear combination of the independent gaussians X_j. Therefore by Proposition 1, $a \cdot Y$ is a gaussian random variable with mean zero. The variance of $a \cdot Y$ is

$$v(a) = \mathbb{E}((a \cdot Y)^2) = \sum_{j,k=1}^{m} a_j a_k c_{jk}$$

The mean and variance specify the distribution of $a \cdot Y$ completely, and it follows from the formula for the Fourier transform of the gaussian that

$$\mathbb{E}\left(e^{-ita \cdot Y}\right) = e^{-v(a)t^2/2}$$

Specializing to $t = 1$ and $a = \xi$ we find the Fourier transform of μ_Y, the joint probability distribution of Y on $\mathbb{R}^m$, is

$$\int_{\mathbb{R}^m} e^{-i\xi \cdot x}\, d\mu_Y(x) = \mathbb{E}\left(e^{-i\xi \cdot Y}\right) = e^{-v(\xi)/2}$$

If Z has the property that $a \cdot Z$ is gaussian with mean zero and the variances $\mathbb{E}(Z_j Z_k) = \mathbb{E}(Y_j Y_k) = c_{jk}$, then the same reasoning leads to the conclusion that the Fourier transform on $\mathbb{R}^m$ of μ_Z the joint probability distribution of Z is also equal to $e^{-v(\xi)/2}$. Therefore, by uniqueness of the Fourier transform for measures, $\mu_Y = \mu_Z$. This concludes the proof.

We make use of Lemma 1 in a special case, in order to characterize $B(t)$.

Proposition 2 *Suppose that $B(t)$ is such that $B(0) = 0$. Then B satisfies property (2) if and only if*

a) $\sum_{j=1}^{m} \xi_j B(t_j)$ *is a Gaussian random variable with mean* 0.

b) $\mathbb{E}(B(s)B(t)) = s \wedge t$, *($s \wedge t = \min(s,t)$).*

Proof. Assume that B satisfies (2). To prove (a), note that for any ξ_j one can find b_j such that

$$\sum_j \xi_j B(t_j) = \xi_1 (B(t_1) - B(0)) + \sum_\ell b_\ell (B(t_{\ell+1}) - B(t_\ell))$$

The latter sum is a sum of independent gaussians, so the fact that the sum is gaussian follows from Proposition 1. To prove (b), note first that (2) implies $B(s)$ is gaussian with mean 0 and variance s, i. e., $\mathbb{E}(B(s)^2) = s$. More generally, for $s \le t$,

$$\mathbb{E}(B(s)B(t)) = \mathbb{E}(B(s)(B(t) - B(s))) + \mathbb{E}(B(s)^2) = 0 + \mathbb{E}(B(s)^2) = s,$$

because independence gives $\mathbb{E}(B(s)(B(t) - B(s))) = \mathbb{E}(B(s))\mathbb{E}((B(t) - B(s))) = 0$.

Conversely, suppose $C(t)$ satisfies a) and b) and $C(0) = 0$. Define B by $B(0) = 0$ and (2), then we have just shown that $X = (B(t_1) - B(t_0), \dots, B(t_m) - B(t_{m-1}))$ is a sequence of independent gaussians of mean zero, and $Y = (B(t_1), \dots, B(t_m))$ has correlation matrix $\mathbb{E}(B(t_j)B(t_k)) = t_j \wedge t_k$. Therefore $Z = (C(t_1), \dots, C(t_m))$ satisfies the hypotheses of Lemma 1 with the same correlation matrix as Y, and by Lemma 1, Z satisfies (2).

We are now ready to prove part (a) of Theorem 2. Let $0 \le t_0 < t_1 < \cdots < t_m \le \pi$. The fact that

$$\sum_{j=1}^{m} \xi_j W(t_j)$$

is gaussian of mean zero follows from Proposition 1. The fact that $W(0) = 0$ is obvious. According to Proposition 2 it remains to show that for $0 \le s \le t \le \pi$,

$$\mathbb{E}(W(s)W(t)) = s \wedge t$$

We could do this all at once, but we carry out a slightly simpler calculation $\mathbb{E}(W(t)^2) = t$ first.

Proposition 1 implies $W(t)$ is gaussian with mean zero and variance

$$\mathbb{E}(W(t)^2) = c_0^2 t^2 + c_1^2 \sum_{k=1}^{\infty} \frac{\sin^2(kt)}{k^2}$$

The case $t = \pi$ identifies c_0:

$$\pi = c_0^2 \pi^2 \implies c_0 = 1/\sqrt{\pi}$$

Denote

$$u(t) = \sum_{k=1}^{\infty} \frac{\sin^2(kt)}{k^2}$$

Then

$$u'(t) = \sum_{k=1}^{\infty} \frac{2k \sin(kt)\cos(kt)}{k^2} = \sum_{k=1}^{\infty} \frac{\sin(2kt)}{k}$$

and

$$u''(t) \sim \sum_{k=1}^{\infty} \frac{2k\cos(2kt)}{k} = 2\sum_{k=1}^{\infty} \cos(2kt) = -1 + \sum_{k \in \mathbb{Z}} e^{2ikt}$$

The last series is periodic of period π. The standard delta function of period π has Fourier coefficients

$$\frac{1}{\pi}\int_{-\pi/2}^{\pi/2} \delta(t)e^{-2int}\,dt = 1/\pi$$

Thus

$$u''(t) = -1 + \pi \sum_{n\in\mathbb{Z}} \delta(t - n\pi)$$

The function $u'(t)$ is odd and we can find its formula by integrating $u''(t)$. We get

$$u'(t) = \begin{cases} \pi/2 - t & 0 < t < \pi \\ -\pi/2 - t & -\pi < t < 0 \end{cases}$$

One way to check your arithmetic is to evaluate $u'(t)$ places where we know what to expect. For instance,

$$u'(\pm\pi/2) = \sum_{k=1}^{\infty} \frac{\sin(\pm 2k\pi/2)}{k} = 0$$

and the formula above gives $u'(\pi/2) = \pi/2 - \pi/2 = 0$ and $u'(-\pi/2) = -\pi/2 - (-\pi/2) = 0$. (You can also confirm the periodicity of period π.)

Next integrate $u'(t)$ to get $u(t)$, which is even and satisfies $u(0) = 0$. Thus,

$$u(t) = \begin{cases} (\pi/2)t - t^2/2 & 0 \le t \le \pi \\ -(\pi/2)t - t^2/2 & -\pi \le t \le 0 \end{cases}$$

You can check your arithmetic in this case by confirming that $u(t)$ is continuous and periodic of period 2π so that the values at $t = \pm\pi$ must agree. (The series is absolutely convergent, so u must be continuous everywhere.)

Now inserting the values of $u(t)$ into the formula for the variance we have for $0 \le t \le \pi$,

$$\mathbb{E}(W(t)^2) = (1/\pi)t^2 + c_1^2[(\pi/2)t - t^2/2] = t$$

provided that $c_1 = \sqrt{2/\pi}$.

Now let's do the full calculation, which is very similar.

$$\mathbb{E}(W(s)W(t)) = c_0^2 st + c_1^2 \sum_{k=1}^{\infty} \frac{\sin(ks)\sin(kt)}{k^2}$$

Since $\sin A \sin B = \frac{1}{2}[\cos(A - B) - \cos(A + B)]$,

$$\sum_{k=1}^{\infty} \frac{\sin(ks)\sin(kt)}{k^2} = \frac{1}{2}\sum_{k=1}^{\infty} \frac{\cos(k(s-t)) - \cos(k(s+t))}{k^2} = \frac{1}{4}[v(s-t) - v(s+t)]$$

with

$$v(t) = 2\sum_{k=1}^{\infty} \frac{\cos(kt)}{k^2}$$

We evaluate $v(t)$ by a similar procedure to the one above.

$$v'(t) = -2\sum_{k=1}^{\infty} \frac{\sin(kt)}{k}$$

$$v''(t) \sim -2\sum_{k=1}^{\infty} \cos kt = 1 - \sum_{k\in\mathbb{Z}} e^{ikt}$$

This time v'' has period 2π and

$$1 - \sum_{k\in\mathbb{Z}} e^{ikt} = 1 - 2\pi\sum_{n\in\mathbb{Z}} \delta(t - 2\pi n)$$

Integrating, the odd function $v'(t)$ is given by

$$v'(t) = \begin{cases} -\pi + t & 0 < t < \pi \\ \pi + t & -\pi < t < 0 \end{cases}$$

Integrating a second time,

$$v(t) - v(0) = -\pi|t| + t^2/2, \quad |t| \le \pi$$

We could calculate $v(0)$, but we only need difference $v(s-t) - v(s+t)$, so the value is not relevant. If we extend $v(t) - v(0)$ as a periodic function of period 2π, then we get

$$v(t) - v(0) = -\pi t + t^2/2, \quad 0 \le t \le 2\pi$$

(We need this range in order to evaluate $v(s+t)$.)

Substituting the formula for $v(t) - v(0)$, we obtain for $0 \le s \le t \le \pi$,

$$\begin{aligned}
\mathbb{E}(W(s)W(t)) &= c_0^2 st + \frac{1}{4}c_1^2[v(s-t) - v(s+t)] \\
&= \frac{st}{\pi} + \frac{1}{2\pi}[-\pi(t-s) + (t-s)^2/2 - (-\pi(s+t) + (s+t)^2/2)] \\
&= s
\end{aligned}$$

This finishes the proof of part (a) of Theorem 2.

To get started with part (b) we need some lemmas.

Lemma 2 *If a_k are mean zero variance 1 gaussians, then*

$$\mathbb{E}(|a_{i_1}a_{i_2}a_{i_3}a_{i_4}|) \le \mathbb{E}(|a_1|^4) = 3 < \infty$$

Proof. All we really care about is that this is finite, which is easy because all the distributions involved are rapidly decreasing. But we can also give an explicit bound as follows. For equidistributed random variables, applying the Schwarz inequality twice,

$$\mathbb{E}(X_1X_2X_3X_4) \le [\mathbb{E}(|X_1X_2|^2)]^{1/2}[\mathbb{E}(|X_3X_4|^2)]^{1/2} \le \prod_{j=1}^{4}[\mathbb{E}(|X_j|^4)^{1/4} = \mathbb{E}(|X_1|^4)$$

One can also get this by applying a version of Hölder's inequality with several factors.

$$\mathbb{E}(|a_j a_{j'} a_k a_{k'}|) \le \mathbb{E}(|a_1|^4) < \infty$$

$\mathbb{E}(a_1^4) = 3$ is calculated as follows. (We only need finiteness, but this calculation is a nice trick to know.) Change of variables of the formula saying the the standard gaussian has integral 1 to get

$$\int_{-\infty}^{\infty} \frac{e^{-\lambda x^2/2}}{\sqrt{2\pi}}\,dx = \lambda^{-1/2}.$$

Differentiate with respect to λ to obtain

$$\int_{-\infty}^{\infty} (-x^2/2)\frac{e^{-\lambda x^2/2}}{\sqrt{2\pi}}\,dx = -(1/2)\lambda^{-3/2}$$

Differentiate a second time to get

$$\int_{-\infty}^{\infty} (-x^2/2)^2\frac{e^{-\lambda x^2/2}}{\sqrt{2\pi}}\,dx = (1/2)(3/2)\lambda^{-5/2}$$

Thus, $\mathbb{E}(a_1^2) = 1$, $\mathbb{E}(a_1^4) = 3$ and, more generally, $\mathbb{E}((a_1)^{2n}) = (2n-1)(2n-3)\cdots 3\cdot 1$.

Lemma 3 *For $m \ge 1$ and $0 \le \beta \le 2$,*

$$R_\beta(m) = \int_0^1 r^m(1-r)^\beta\,dr \le 100m^{-1-\beta}$$

Proof. This integral can be evaluated in terms of what is known as Euler's beta integral. The answer is a product of gamma functions and the asymptotics are easy to read off from Stirling's formula. We won't rely on any of that, but rather prove the upper bound directly.

For $1-(k+1)/m \le r \le 1-k/m$,

$$r^m(1-r)^\beta \le (1-k/m)^m[(k+1)/m]^\beta \le e^{-k}[(k+1)/m]^\beta$$

Therefore,

$$\begin{aligned}\int_0^1 r^m(1-r)^\beta\,dr &\le \frac{1}{m}\sum_{k=0}^{m-1} e^{-k}[(k+1)/m]^\beta \\ &= m^{-1-\beta}\sum_{k=0}^{m-1}(k+1)^\beta e^{-k} \\ &\le m^{-1-\beta}\int_0^\infty (x+2)^2 e^{-x}\,dx \le 10m^{-1-\beta}\end{aligned}$$

Let

$$F(z) = \sum_{k=1}^\infty a_k z^k/k$$

with a_k independent gaussians with mean zero and variance 1. Note that

$$B(t) = c_0 a_0 t + c_1 \sum_{k=1}^\infty a_k \frac{\sin(kt)}{k} = c_0 a_0 t + c_1 \mathrm{Im}\, F(e^{it})$$

In order to show that $B(t)$ is almost surely continuous, it suffices to show the same for $F(e^{it})$. This is accomplished by estimating $F(z)$ in $|z|<1$. Note that

$$F_x(z) = \frac{\partial}{\partial x}F(z) = \sum_{j=0}^\infty a_{j+1}z^j$$

since for $z = x+iy$, $(\partial/\partial x)z = 1$. Moreover, since $(\partial/\partial y)z = i$, we also have $F_y = -iF_x$ and

$$\frac{1}{2}|\nabla F|^2 = |F_x|^2 = \left|\sum_{j=0}^\infty a_{j+1}z^j\right|^2 = \sum_{j,j'=0}^\infty a_{j+1}a_{j'+1}z^j\bar{z}^{j'}$$

Lemma 4 *For any $\beta > 1$,*

$$\mathbb{E}\int_0^{2\pi}\int_0^1 |\nabla F(re^{it})|^4(1-r)^\beta r\,dr\,dt < \infty,$$

and, consequently

$$\int_0^{2\pi}\int_0^1 |\nabla F(re^{it})|^4(1-r)^\beta r\,dr\,dt < \infty,$$

almost surely.

Proof. Think of the expectation as a triple integral (over $\omega \in \Omega$, the probability sample) and r and t. The monotone convergence theorem implies

$$\mathbb{E}\int_0^{2\pi}\int_0^1 |\nabla F(re^{it})|^4(1-r)^\beta r\,dr\,dt = \lim_{r_0\to 1^-}\mathbb{E}\int_0^{2\pi}\int_0^{r_0} |\nabla F(re^{it})|^4(1-r)^\beta r\,dr\,dt$$

Therefore it suffices to bound the integral restricted to $0 \le r \le r_0 < 1$, uniformly as $r_0 \to 1$.

Next, apply Fubini's theorem (justified subsequently)

$$\begin{aligned}
\mathbb{E}\int_0^{2\pi}\int_0^{r_0} &|\nabla F(re^{it})|^4(1-r)^\beta r\,dr\,dt \\
&= \mathbb{E}\int_0^{2\pi}\int_0^{r_0}\sum_{j,j',k,k'=0}^{\infty} a_j a_{j'} a_k a_{k'} r^{j+j'+k+k'} e^{i(j-j'+k-k')t}(1-r)^\beta r\,dr\,dt \\
&= \int_0^{2\pi}\int_0^{r_0}\sum_{j,j',k,k'=0}^{\infty} \mathbb{E}(a_j a_{j'} a_k a_{k'}) r^{j+j'+k+k'} e^{i(j-j'+k-k')t}(1-r)^\beta r\,dr\,dt \\
&= 4\pi\int_0^{2\pi}\int_0^{r_0}\sum_{j,k=0}^{\infty}\mathbb{E}(a_j^2 a_k^2) r^{2j+2k+1}(1-r)^\beta\,dr \\
&\le 12\pi\int_0^1\sum_{j,k=0}^{\infty} r^{2j+2k+1}(1-r)^\beta\,dr \\
&= 12\pi\sum_{j,k=0}^{\infty} R_\beta(2j+2k+1) \\
&\le C\sum_{j,k=0}^{\infty}\frac{1}{(2j+2k+1)^{1+\beta}} \approx \int_{x\in\mathbb{R}^2}\frac{dx}{(1+|x|)^{1+\beta}} < \infty
\end{aligned}$$

provided $\beta > 1$, so that the exponent $1+\beta > 2$. We need to justify Fubini's theorem so as to bring the expectation inside the integrals and the sum. It is in order to justify this step that the integral was restricted to $0 \le r \le r_0$. In fact,

$$\int_0^{2\pi}\int_0^{r_0}\sum_{j,j',k,k'=0}^{\infty}\mathbb{E}|a_j a_{j'} a_k a_{k'} r^{j+j'+k+k'} e^{i(j-j'+k-k')t}|(1-r)^\beta r\,dr\,dt < \infty$$

By Lemma 2, $\mathbb{E}(|a_j a_{j'} a_k a_{k'}|) \le 3$. Moreover,

$$\begin{aligned}
\int_0^{2\pi}\int_0^{r_0}\sum_{j,j',k,k'=0}^{\infty} r^{j+j'+k+k'+1} 3(1-r)^\beta\,dr\,dt &\le 6\pi\sum_{j,j',k,k'=0}^{\infty} r_0^{j+j'+k+k'} \\
&= \frac{6\pi}{(1-r_0)^4} < \infty
\end{aligned}$$

For the next step we need the mean value property for harmonic functions. If u is harmonic in $|z| < 1$, and continuous on $|z| \le 1$, then for $0 \le r < 1$,

$$u(re^{it}) = \sum_{n \in \mathbb{Z}} a_n r^{|n|} e^{int}$$

Integrating in t,

$$\frac{1}{2\pi} \int_0^{2\pi} u(re^{it})\, dt = a_0 = u(0)$$

Now integrating with respect to r on $0 \le r \le \rho \le 1$,

$$\frac{1}{2\pi} \int_0^{2\pi} \int_0^{\rho} u(re^{it})\, rdr\, dt = \int_0^{\rho} u(0)\, rdr = u(0)\rho^2/2$$

Thus

$$u(0) = \frac{1}{\pi\rho^2} \int_0^{\rho} \int_0^{2\pi} u(re^{it})\, rdr\, dt = \frac{1}{\pi\rho^2} \int_{|z| \le \rho} u(z)\, dxdy$$

where $z = x + iy$. A similar argument (or a change of variable) shows that if u is harmonic in $|z - z_0| \le \rho$, then

$$u(z_0) = \frac{1}{\pi\rho^2} \int_{|z-z_0| \le \rho} u(z)\, dxdy \qquad \text{(Mean value property)} \tag{4}$$

We now apply (4) to ∇F. Almost surely,

$$C_* = \int_0^{2\pi} \int_0^1 |\nabla F(re^{it})|^4 (1-r)^{\beta} r\, dr\, dt < \infty$$

Let $\alpha = 1 - (2 + \beta)/4$. For F satisfying the bound above, we will show that there is a constant C depending on $C_* = C_*(F)$ such that for all z in the unit disk,

$$|\nabla F(z)| \le C(1 - |z|)^{-1+\alpha} \tag{5}$$

We will then deduce Hölder continuity with exponent α. Since β is any real number greater than 1, we see that the Hölder exponent of $W(t)$ at least α for any $\alpha < 1/4$. (Estimates using higher powers than $|\nabla F|^4$ and higher moments of the gaussian coefficients $a_k(\omega)$ can be used to show that W is Hölder continuous for any exponent $\alpha < 1/2$.)

Let $|z_0| = r_0$, and let $1 - r_0 = 2\rho$. Then

$$|\nabla F(z_0)| \le \frac{1}{\pi\rho^2} \int_{|z-z_0| \le \rho} |\nabla F(z)|\, dxdy \le \left(\frac{1}{\pi\rho^2} \int_{|z-z_0| \le \rho} |\nabla F(z)|^4\, dxdy \right)^{1/4}$$

On the disk $|z-z_0| \le \rho$, $|z| = r \le 1-\rho$. Therefore,

$$\int_{|z-z_0|\le\rho} |\nabla F(z)|^4 \, dxdy \le 10\rho^{-\beta} C_*$$

It follows that

$$|\nabla F(z_0)| \le (10 C_* \rho^{-2-\beta})^{1/4}$$

which is the same as (5).

Lemma 5 *If F satisfies (5) on $|z| \le 1$, for some α, $0 < \alpha \le 1$, then $F(e^{it})$ is Hölder continuous with exponent α, that is,*

$$|F(e^{t_1}) - F(e^{it_2})| \le C|t_1 - t_2|^\alpha$$

Proof. Given any two points t_1 and t_2 such that $t_2 - t_2 = \rho$. Consider $1 - r_0 = \rho$ and the line segment L_1 from e^{it_1} to $r_0 e^{it_1}$, L_2 is the circular arc of length less than ρ along $|z| = r_0$ from $r_0 e^{it_1}$ to $r_0 e^{it_2}$, and L_3 is the segment from $r_0 e^{it_2}$ to e^{it_2}. The integral of $|\nabla F|$ on L_1 is at most

$$C \int_0^\rho s^{-1+\alpha} \, ds = (C/\alpha)\rho^\alpha$$

and similarly on L_3. The integral on the circular arc L_2 is at most its length of the arc times the bound on $|\nabla F|$ along that arc, namely $\rho\, O(\rho^{-1+\alpha}) = O(\rho^\alpha)$. Thus

$$|F(e^{it_2}) - F(e^{it_1})| \le C\rho^\alpha$$