NEW

Huai Jie (Dante) • 28 Apr 2026

Kolmogorov's Continuity Theorem and the Fine Structure of Brownian Motion

Expanded from lecture notes

Contents

Introduction

The goal of these notes is to answer two fundamental questions about stochastic processes:

1. Given a process defined by its finite-dimensional distributions, when can we guarantee that it has a *continuous* version (i.e., a modification with continuous sample paths)?
2. If a continuous version exists, how *smooth* are its sample paths? Can we measure their irregularity quantitatively?

The answers are provided by two landmark results:

- The **Kolmogorov continuity theorem** (also called Kolmogorov–Čentsov theorem) gives a simple moment condition that implies the existence of a Hölder-continuous modification.
- **Lévy's modulus of continuity** for Brownian motion tells us exactly how fast the increments can oscillate: the almost sure oscillation is of order $\sqrt{2h\log(1/h)}$, not $h^{1/2}$ as one might naively think.

We will also explore the **scaling and inversion symmetries** of Brownian motion, which are the source of its self-similarity and fractal nature. The problems at the end illustrate these concepts and lead to deeper insights.

Hölder Continuity – A Measure of Smoothness

Before stating the theorems, we need a way to grade the regularity of a continuous function.

Definition 1. *Let $f : [a,b] \to \mathbb{R}$ be continuous. For $\alpha \in (0,1]$, we say f is **Hölder continuous of order** α (or α-Hölder) if there exists a constant C such that*

$$|f(s) - f(t)| \le C\,|s-t|^{\alpha} \qquad \text{for all } s,t \in [a,b].$$

*The largest α for which this holds is the **Hölder exponent**.*

Example 2.

- A Lipschitz function is 1-Hölder. Example: $f(t) = t$.
- A function with a cusp, like $f(t) = |t|^{1/2}$, is 1/2-Hölder (with $C = 1$).
- The Weierstrass function is nowhere differentiable but Hölder continuous for some $\alpha < 1$.

Why is this important? For stochastic processes, sample paths are almost always irregular (e.g., Brownian motion is nowhere differentiable). Yet they often satisfy a Hölder condition with some exponent $\alpha < 1$. The Kolmogorov theorem tells us that if the moments of the increments grow like a power of the time difference, then the paths are automatically Hölder continuous – we don't need to construct them explicitly.

Kolmogorov's Continuity Theorem

Theorem 3 (Kolmogorov–Čentsov)**.** *Let $\{X_t : t \in [0,1]^d\}$ be a stochastic process on a probability space $(\Omega, \mathcal{F}, P)$. Suppose there exist constants $p > 0$, $\varepsilon > 0$ and $C < \infty$ such that for all $s,t \in [0,1]^d$,*

$$\mathbb{E}\big[|X_t - X_s|^p\big] \le C\,|t-s|^{d+\varepsilon}.$$

Then there exists a modification $\widetilde{X}$ of X (i.e., $\widetilde{X}_t = X_t$ a.s. for each t) whose sample paths are Hölder continuous of order α for every $\alpha \in (0, \varepsilon/p)$. Moreover, the modification is unique up to indistinguishability.

Intuition

The condition controls the p-th moment of the increment. By Markov's inequality,

$$P(|X_t - X_s| > \delta) \le \frac{\mathbb{E}[|X_t - X_s|^p]}{\delta^p} \le \frac{C}{\delta^p}\,|t-s|^{d+\varepsilon}.$$

If we take $\delta = |t-s|^{\alpha}$ with $\alpha < \varepsilon/p$, the right-hand side becomes $C|t-s|^{d+\varepsilon-p\alpha}$. For small $|t-s|$, this probability decays faster than $|t-s|^d$, which allows us to control the oscillations on a fine dyadic grid and then use a Borel–Cantelli argument. The result is a continuous, Hölder-regular modification.

Example: Brownian motion

For standard Brownian motion in $d = 1$, we have $X_t - X_s \sim N(0, |t-s|)$. Hence

$$\mathbb{E}[|X_t - X_s|^p] = c_p|t-s|^{p/2},$$

where $c_p = \mathbb{E}[|Z|^p]$ for $Z \sim N(0,1)$. Take p large and set $\varepsilon = p/2 - d = p/2 - 1$. For $p > 2$ we have $\varepsilon > 0$. Then the theorem guarantees Hölder continuity for any $\alpha < (p/2-1)/p = \frac{1}{2} - \frac{1}{p}$. Letting $p \to \infty$, we obtain Hölder continuity for all $\alpha < 1/2$. So Brownian paths are α-Hölder for every $\alpha < 1/2$, but *not* for $\alpha = 1/2$ (this is the content of Lévy's modulus).

Lévy's Modulus of Continuity for Brownian Motion

For Brownian motion, the optimal Hölder exponent is $1/2$ in a very precise, almost-sure sense.

Theorem 4 (Lévy 1937). *Let $\{W_t : t \geq 0\}$ be a standard Brownian motion. Then*

$$\limsup_{h \to 0} \frac{\sup_{0 \leq t \leq 1-h} |W_{t+h} - W_t|}{\sqrt{2h \log(1/h)}} = 1 \qquad \textit{almost surely.}$$

Interpretation

For each fixed h, the maximum increment over a sliding window of length h is approximately $\sqrt{2h\log(1/h)}$. This is slightly larger than $h^{1/2}$ because $\sqrt{\log(1/h)}$ grows slowly as $h \to 0$. The theorem says:

- If $c > 1$, then eventually (for all sufficiently small h),
- $$\sup_t |W_{t+h} - W_t| \leq c\sqrt{2h\log(1/h)} \quad \text{a.s.}$$

  If $c < 1$, then infinitely often (for a sequence $h_n \to 0$),

  $$\sup_t |W_{t+h_n} - W_{t_n}| > c\sqrt{2h_n\log(1/h_n)} \quad \text{a.s.}$$

Thus the function $\psi(h) = \sqrt{2h\log(1/h)}$ is the **exact modulus of continuity** for Brownian motion.

Consequences

1. Brownian motion is nowhere locally α-Hölder for any $\alpha > 1/2$ (because the oscillation is too large). For $\alpha = 1/2$ it fails as well (the additional $\sqrt{\log(1/h)}$ factor diverges). Hence the critical exponent is exactly $1/2$.
2. The result is a cornerstone of stochastic calculus and the theory of Gaussian processes. It shows that Brownian paths are *just barely* continuous – but with an infinite $1/2$-variation (quadratic variation is non-zero).

Scaling and Inversion Symmetries of Brownian Motion

Brownian motion enjoys two important invariance properties that make it a “random fractal”. They are often used to deduce properties of first passage times, local times, and fractal dimensions.

Scaling invariance

Lemma 5. *If* $\{W_t\}$ *is a standard BM and* $c > 0$*, then the process*

$$X_t = \frac{1}{c} W_{c^2 t}$$

is also a standard BM.

Proof. For any $0 \le s < t$,

$$X_t - X_s = \frac{1}{c}(W_{c^2 t} - W_{c^2 s}) \sim N(0, t-s),$$

and the increments are independent. Path continuity is preserved. The scaling factor c^2 on time is necessary to keep the variance linear in t. □

Meaning: If you zoom out horizontally by c^2 and vertically by c, the path looks like a Brownian motion again. This self-similarity is the key to many fractal properties.

First exit times

Let $T(a,b) = \inf\{t > 0 : W_t = a \text{ or } W_t = b\}$, where $a < 0 < b$. Using scaling, we can relate expectations and probabilities.

Example 6. Take $a = -b$ (symmetry). Then

$$\mathbb{E}[T(-b,b)] = b^2\, \mathbb{E}[T(-1,1)].$$

So the expected exit time scales like b^2. This is intuitive: for a simple random walk in discrete time, the expected exit time from $\pm b$ is b^2. For BM, one can compute $\mathbb{E}[T(-1,1)] = 1$.

Time inversion

Lemma 7. *Define* $X_0 = 0$ *and for* $t > 0$*,* $X_t = t\,W_{1/t}$*. Then* $\{X_t\}$ *is also a standard BM.*

Proof. Finite-dimensional distributions are Gaussian with mean zero and

$$\mathrm{Cov}(X_t, X_s) = ts\,\mathrm{Cov}(W_{1/t}, W_{1/s}) = ts \cdot \frac{1}{\max(t,s)} = \min(t,s).$$

Continuity at 0 follows from the law of large numbers: $\lim_{t\to 0} X_t = \lim_{u\to\infty} \frac{W_u}{u} = 0$ a.s. □

Significance: The behaviour of BM near 0 (infinitesimal times) and near ∞ (large times) are the same after inversion. This is why properties like “almost every path hits every real number” can be derived from the fact that $\limsup_{t\to\infty} W_t = \infty$ a.s.

Supplementary Examples and Applications

Example: Using scaling to compute exit probabilities

Let $p(a,b) = P(W_t \text{ hits } a \text{ before } b)$ with $a < 0 < b$. By scaling,

$$p(a,b) = P\left(\frac{1}{b}W_{b^2t} \text{ hits } a/b \text{ before } 1\right) = p(a/b, 1).$$

Thus the probability depends only on the ratio $|a|/b$. In fact, for BM, $p(-1,1) = \frac{1}{2}$ by symmetry, and one can show $p(-c,1) = \frac{1}{1+c}$ for $c > 0$. This matches the gambler’s ruin formula.

Example: Hölder regularity of BM via Kolmogorov

Take $p = 4$. Then $\mathbb{E}[|W_t - W_s|^4] = 3|t-s|^2$. Here $d = 1$, so we have exponent $2 = d+1$ (so $\varepsilon = 1$). Hence $\alpha < \varepsilon/p = 1/4$. This already gives Hölder of order $< 1/4$. By taking p larger we approach $1/2$. So the theorem gives existence of a version that is $1/2$-Hölder? No, it gives for any $\alpha < 1/2$ but not $\alpha = 1/2$ (the theorem would require $\varepsilon/p > 1/2$ which is impossible). Lévy’s modulus shows that $\alpha = 1/2$ fails.

Example: The law of the iterated logarithm (LIL)

A related result is Khinchin’s LIL:

$$\limsup_{t\to 0} \frac{|W_t|}{\sqrt{2t\log\log(1/t)}} = 1 \quad \text{a.s.}$$

This is a pointwise version (at a fixed t) whereas Lévy’s modulus is uniform over the whole interval.

Problems with Detailed Solutions/ Hints

Below are the problems from the original notes, expanded with full solutions or extensive hints.

Problem 1. Show that almost all Brownian sample paths are *not* Hölder continuous of order 1/2 using Lévy's modulus of continuity.

Solution 1. Lévy's theorem states that

$$\limsup_{h\to 0} \frac{\sup_{t\in[0,1-h]} |W_{t+h} - W_t|}{\sqrt{2h\log(1/h)}} = 1.$$

If the paths were 1/2-Hölder, there would exist a constant C such that for all small h,

$$\sup_t |W_{t+h} - W_t| \le Ch^{1/2}.$$

But $\sqrt{2h\log(1/h)}/h^{1/2} = \sqrt{2\log(1/h)} \to \infty$ as $h \to 0$. Hence for any fixed C, for sufficiently small h we would have

$$\sqrt{2h\log(1/h)} > Ch^{1/2} \quad \Longrightarrow \quad \frac{\sup_t |W_{t+h} - W_t|}{h^{1/2}} > C.$$

The limsup being 1 in the Lévy modulus implies that the ratio $\frac{\sup_t |W_{t+h}-W_t|}{h^{1/2}}$ is unbounded as $h \to 0$. Therefore no finite C works, i.e., the path is not 1/2-Hölder. (It is α-Hölder for any $\alpha < 1/2$ by the Kolmogorov theorem, so 1/2 is the critical exponent.)

Problem 2. Let $Z \sim N(0,1)$. Prove the Gaussian tail bounds:

(a) $\dfrac{a}{a^2+1}\dfrac{e^{-a^2/2}}{\sqrt{2\pi}} \le P(Z \ge a) \le \dfrac{1}{a}\dfrac{e^{-a^2/2}}{\sqrt{2\pi}}$ for all $a > 0$.

(b) $\displaystyle\lim_{a\to\infty} \frac{1-\Phi(a)}{a^{-1}\phi(a)} = 1.$

(c) $\Phi(t) - \Phi(s) < t - s$ for all $s < t$.

Solution 2. Let $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ and $\Phi(t) = \int_{-\infty}^{t} \phi(u)du$. For $a > 0$,

$$P(Z \geq a) = \int_a^\infty \phi(t)dt.$$

Notice that $\phi'(t) = -t\phi(t)$. For the upper bound, use the inequality $\int_a^\infty \phi(t)dt \leq \int_a^\infty \frac{t}{a}\phi(t)dt = \frac{1}{a}\int_a^\infty t\phi(t)dt = \frac{1}{a}\phi(a)$ because $t\phi(t) = \phi'(t)$ up to sign. Indeed, $\int_a^\infty t\phi(t)dt = [-\phi(t)]_a^\infty = \phi(a)$. This gives the right-hand side.
For the lower bound, consider

$$P(Z \geq a) = \int_a^\infty \frac{1}{t} \cdot t\phi(t)dt \geq \frac{1}{a + 1/a} \int_a^\infty t\phi(t)dt \quad \text{(since for } t \geq a,\ 1/t \geq 1/(a + 1/a)\text{? Not}$$

Better: write $\int_a^\infty \phi(t)dt = \int_a^\infty \frac{t}{t}\phi(t)dt$. For $t \geq a$, $\frac{1}{t} \geq \frac{1}{a+\frac{1}{a}}$? That's false for large t. Instead, integrate the inequality $\frac{d}{dt}\left[\left(1 + \frac{1}{t^2}\right)\phi(t)\right] = -\phi(t) + \text{remainder}$. The classic proof: note that

$$\left(1 - \frac{1}{t^2}\right)\phi(t) \leq \phi(t) \leq \left(1 + \frac{1}{t^2}\right)\phi(t)$$

and integrate from a to ∞. More directly, one can use that for $x > 0$,

$$\frac{x}{1 + x^2} e^{-x^2/2} \leq \int_x^\infty e^{-t^2/2}dt \leq \frac{1}{x} e^{-x^2/2}.$$

This is a known Mills' ratio bound. The proof uses the substitution $u = t^2$ and integration by parts. We'll give a succinct verification: Let $I(a) = \int_a^\infty e^{-t^2/2}dt$. Then

$$I(a) = \int_a^\infty \frac{1}{t} te^{-t^2/2}dt = \frac{e^{-a^2/2}}{a} - \int_a^\infty \frac{1}{t^2} e^{-t^2/2}dt.$$

The integral term is positive, so $I(a) < e^{-a^2/2}/a$. Also,

$$I(a) = \frac{e^{-a^2/2}}{a} - \int_a^\infty \frac{1}{t^2} e^{-t^2/2}dt \geq \frac{e^{-a^2/2}}{a} - \frac{1}{a^2}\int_a^\infty e^{-t^2/2}dt = \frac{e^{-a^2/2}}{a} - \frac{1}{a^2} I(a).$$

Solving for $I(a)$ gives $I(a)(1 + 1/a^2) \geq e^{-a^2/2}/a$, i.e., $I(a) \geq \frac{e^{-a^2/2}}{a(1+1/a^2)} = \frac{a}{a^2+1} e^{-a^2/2}$. Multiplying by $1/\sqrt{2\pi}$ yields the desired lower bound.
(b) From the bounds, dividing $1 - \Phi(a)$ by $a^{-1}\phi(a)$ gives something that tends to 1 because both upper and lower bounds converge to 1. More precisely,

$$\frac{a}{a^2 + 1} \leq \frac{1 - \Phi(a)}{a^{-1}\phi(a)} \leq 1,$$

and the left side tends to 1 as $a \to \infty$.
(c) Since $\Phi'(t) = \phi(t) \leq 1$ (because $\phi(t) \leq 1/\sqrt{2\pi} < 1$), the mean value theorem gives $\Phi(t) - \Phi(s) = \phi(\xi)(t - s)$ for some $\xi \in (s, t)$, and $\phi(\xi) \leq 1$. Moreover, the inequality is strict unless $t = s$ because $\phi(\xi) < 1$ for all finite ξ. Hence $\Phi(t) - \Phi(s) < t - s$.
Problem 3. For $T(a) = \inf\{t > 0 : W_t = a\}$ with $a \in \mathbb{R}$, prove:

$$T(-a) \overset{d}{=} T(a) \quad \text{and} \quad T(a) \overset{d}{=} a^2 T(1).$$

Solution 3. Symmetry: $-W_t$ is also a BM, so the hitting time of $-a$ for W equals the hitting time of a for $-W$, which has the same distribution as $T(a)$.
Scaling: Define $\widetilde{W}_t = (1/a)W_{a^2t}$. By scaling invariance, $\widetilde{W}$ is a BM. Then

$$T(a) = \inf\{t > 0 : W_t = a\} = \inf\{t > 0 : a\widetilde{W}_{t/a^2} = a\} = \inf\{t > 0 : \widetilde{W}_{t/a^2} = 1\}.$$

Let $u = t/a^2$; then $t = a^2u$ and the condition becomes $\inf\{a^2u > 0 : \widetilde{W}_u = 1\} = a^2T(1)$ (where $T(1)$ is the hitting time for $\widetilde{W}$ to reach 1). Since $\widetilde{W}$ is a BM, $\widetilde{T}(1)$ has the same distribution as $T(1)$. Thus $T(a) \overset{d}{=} a^2T(1)$.

Problem 4. Fix $a > 0$. Show that $B_t := W_{t+a} - W_a$ is again a standard BM.

Solution 4. By the independent increments property of BM, the process $\{W_{t+a} - W_a\}_{t\geq 0}$ is independent of $\mathcal{F}_a$ and has the same finite-dimensional distributions as $\{W_t\}$. Its paths are continuous. Hence it is a BM. (This is the Markov property restarted at time a.)

Problem 5. Show that $P(\sup_{t\geq 0} W_t = \infty) = 1$ and consequently $P(\inf_{t\geq 0} W_t = -\infty) = 1$.

Solution 5. Let $Z = \sup_{t\geq 0} W_t$. By scaling, for any $c > 0$,

$$Z \overset{d}{=} \sup_{t\geq 0} \frac{1}{c}W_{c^2t} = \frac{1}{c}\sup_{u\geq 0} W_u = \frac{1}{c}Z.$$

Thus Z has the same distribution as Z/c. This can only happen if Z is either 0 or ∞ almost surely (because if Z were a finite positive constant with positive probability, scaling would change it). The reflection principle gives $P(Z = 0) = 0$ (since BM will hit any positive level with probability 1). More directly, by the law of the iterated logarithm, $\limsup_{t\to\infty} W_t/\sqrt{2t\log\log t} = 1$ a.s., so $\limsup_{t\to\infty} W_t = \infty$ a.s. Hence $P(Z = \infty) = 1$. By symmetry, $\inf_{t\geq 0} W_t = -\infty$ a.s.

Problem 6. A process $\{X_t\}$ with $\mathbb{E}[X_t^2] < \infty$ is weakly stationary if $\mathbb{E}[X_t]$ is constant and $\mathrm{Cov}(X_s, X_t) = g(t - s)$ for some even function g. Determine which of the following are stationary:

(a) $U_t = W_t^2 - t$

(b) $X_t = e^{-at}W_{e^{2at}}$

(c) $Y_t = W_{t+h} - W_t$ (for fixed $h > 0$)

(d) $Z_t = W_{e^t}$ for $t > 0$

Solution 6.

(a) $\mathbb{E}[U_t] = \mathbb{E}[W_t^2] - t = t - t = 0$. Covariance: $\mathrm{Cov}(U_s, U_t) = \mathbb{E}[W_s^2 W_t^2] - st - \mathbb{E}[W_s^2]\mathbb{E}[W_t^2] + st$? Actually $\mathbb{E}[U_s U_t] = \mathbb{E}[(W_s^2 - s)(W_t^2 - t)]$. For $s \leq t$, $W_t = W_s + (W_t - W_s)$ independent of W_s. Then $\mathbb{E}[W_s^2 W_t^2] = \mathbb{E}[W_s^2(W_s^2 + 2W_s(W_t - W_s) + (W_t - W_s)^2)] = \mathbb{E}[W_s^4] + 0 + \mathbb{E}[W_s^2]\mathbb{E}[(W_t - W_s)^2]$. $\mathbb{E}[W_s^4] = 3s^2$, $\mathbb{E}[W_s^2] = s$, $\mathbb{E}[(W_t - W_s)^2] = t - s$. So $\mathbb{E}[W_s^2 W_t^2] = 3s^2 + s(t-s) = 2s^2 + st$. Then $\mathbb{E}[U_s U_t] = 2s^2 + st - st - st + st = 2s^2 - st$. This depends on both s and t individually, not just on $t - s$. Hence not stationary (except trivial). Actually the correct known result: $W_t^2 - t$ is a martingale but its increments are not stationary; the covariance is $\mathrm{Cov}(U_s, U_t) = 2\min(s,t)^2$? Let's compute carefully: For $s \leq t$,

$$\mathbb{E}[U_s U_t] = \mathbb{E}[(W_s^2 - s)((W_s + W_t - W_s)^2 - t)] = \mathbb{E}[W_s^4 - 2W_s^2(W_t - W_s) + \cdots].$$

Better: $\mathbb{E}[W_t^2 W_s^2] = s^2 + 2s(t-s)$ for $s < t$ (using W_s and independent increment). Then $\mathbb{E}[(W_s^2 - s)(W_t^2 - t)] = s^2 + 2s(t-s) - st - st + st = s^2 + 2s(t-s) - st = s^2 + 2st - 2s^2 - st = st - s^2$. So covariance $= s(t-s) = s(t-s)$, which is not a function of $t - s$ alone. Therefore not stationary.

(b) $X_t = e^{-at} W_{e^{2at}}$. Mean zero. Covariance:

$$\mathbb{E}[X_s X_t] = e^{-a(s+t)}\mathbb{E}[W_{e^{2as}} W_{e^{2at}}] = e^{-a(s+t)} \min(e^{2as}, e^{2at}) = e^{-a(s+t)} e^{2a\min(s,t)} = e^{-a|t-s|}.$$

This depends only on $|t - s|$ and is even. So $\{X_t\}$ is stationary (it is the Ornstein–Uhlenbeck process).

(c) $Y_t = W_{t+h} - W_t$ for fixed h. Mean zero. Covariance:

$$\mathbb{E}[Y_t Y_{t+\tau}] = \mathbb{E}[(W_{t+h} - W_t)(W_{t+\tau+h} - W_{t+\tau})].$$

For $\tau \geq h$ the intervals $[t, t+h]$ and $[t+\tau, t+\tau+h]$ are disjoint, so covariance $= 0$. For $0 \leq \tau < h$ we have overlapping increments; a direct computation gives $\mathbb{E}[Y_t Y_{t+\tau}] = h - \tau$ (when $\tau \geq 0$). This depends only on $\tau = |t-s|$ (and is symmetric). So $\{Y_t\}$ is stationary (it is a moving average of white noise, in fact the process of non-overlapping increments is i.i.d., but overlapping increments have a triangular covariance).

(d) $Z_t = W_{e^t}$ for $t > 0$. Mean zero. Covariance:

$$\mathbb{E}[Z_s Z_t] = \min(e^s, e^t) = e^{\min(s,t)}.$$

This is not a function of $|t - s|$ (e.g., take $s = 0, t = 1$ gives $e^0 = 1$; $s = 1, t = 2$ gives $e^1 = e$). Hence not stationary.

So the stationary processes are (b) and (c).

Further Reading

- Kolmogorov's original paper: "Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung" (1931).
- For Brownian motion: Karatzas & Shreve, *Brownian Motion and Stochastic Calculus*.

- For Gaussian processes and Hölder regularity: Adler & Taylor, *Random Fields and Geometry*.