

Formalizing Brownian motion

April 21, 2024

Our goal is to write down the steps necessary in order to formalize Brownian motions (or \mathbb{R}^d -valued Gaussian processes) in some generality using `mathlib`.

Remark 0.1 (Notation). We will write (E, r) for some extended pseudo-metric space, $\mathcal{P}(E)$ for the set of probability measures on the Borel σ -algebra on E , $\mathbb{k} \in \{\mathbb{R}, \mathbb{C}\}$, and $\mathcal{C}_b(E, \mathbb{k})$ the set of \mathbb{k} -valued bounded continuous functions on E . For some $\mathbf{P} \in \mathcal{P}(E)$ and $f \in \mathcal{C}_b(E, \mathbb{k})$, we let $\mathbf{P}[f] := \int f(x) \mathbf{P}(dx) \in \mathbb{k}$ be the expectation.

0 Some simple probability results

The following is a simple consequence of dominated convergence, and often needed in probability theory.

Definition 0.1. Let E be some set and $f, f_1, f_2, \dots : E \rightarrow \mathbb{k}$. We say that f_1, f_2, \dots converges to f boundedly pointwise if $f_n \xrightarrow{n \rightarrow \infty} f$ pointwise and $\sup_n \|f_n\| < \infty$. We write $f_n \xrightarrow{n \rightarrow \infty}_{bp} f$

lemma:bp

Lemma 0.2. Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability (or finite) measure space, and $X, X_1, X_2, \dots : \Omega \rightarrow \mathbb{k}$ such that $X_n \xrightarrow{n \rightarrow \infty}_{bp} X$. Then, $\mathbf{E}[X_n] \xrightarrow{n \rightarrow \infty} \mathbf{E}[X]$.

Proof. Note that the constant function $x \mapsto \sup_n \|f_n\|$ is integrable (since \mathbf{P} is finite), so the result follows from dominated convergence. \square

Definition 0.3. Let X, X_1, X_2, \dots , all E -valued random variables.

1. We say that $X_n \xrightarrow{n \rightarrow \infty} X$ almost everywhere if $\mathbf{P}(\lim_{n \rightarrow \infty} X_n = X) = 1$. We also write $X_n \xrightarrow{n \rightarrow \infty}_{ae} X$.
2. We say that $X_n \xrightarrow{n \rightarrow \infty} X$ in probability (or in measure) if, for all $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbf{P}(r(X_n, X) > \varepsilon) = 0.$$

The two notions here are denoted $\forall^m (x : \alpha) \partial \mathbf{P}, \text{Filter.Tendsto} (\text{fun } n \Rightarrow X \ n \ x) \text{Filter.atTop} (\text{nhds } (X \ x))$ and $\text{MeasureTheory.TendstoInMeasure}$, respectively.

l:aep

Lemma 0.4. Let X, X_1, X_2, \dots be E -valued random variables with $X_n \xrightarrow{n \rightarrow \infty}_{ae} X$. Then, $X_n \xrightarrow{n \rightarrow \infty}_p X$.

This result is called `MeasureTheory.tendstoInMeasure_of_tendsto_ae` in `mathlib`. We also need the (almost sure) uniqueness of the limit in measure, which is not formalized in `mathlib` yet:

l:puni

Lemma 0.5 (Uniqueness of a limit in probability). Let X, Y, X_1, X_2, \dots be E -valued random variables with $X_n \xrightarrow{n \rightarrow \infty}_p X$ and $X_n \xrightarrow{n \rightarrow \infty}_p Y$. Then, $X = Y$, almost surely.

Proof. We write, using monotone convergence and Lemma ??

$$\mathbf{P}(X \neq Y) = \lim_{\varepsilon \downarrow 0} \mathbf{P}(r(X, Y) > \varepsilon) \leq \lim_{\varepsilon \downarrow 0} \lim_{n \rightarrow \infty} \mathbf{P}(r(X, X_n) > \varepsilon/2) + \mathbf{P}(r(Y, X_n) > \varepsilon/2) = 0.$$

\square

l:supsum

Lemma 0.6. Let I be some (finite or infinite) set and $(X_t)_{t \in I}$ be a family of random variables with values in $[0, \infty)$. Then, $\sup_{t \in I} X_t \leq \sum_{t \in I} X_t$.

1 Separating algebras and characteristic functions

Definition 1.1 (Separating class of functions). Let $\mathcal{M} \subseteq \mathcal{C}_b(E, \mathbb{k})$.

1. If, for all $x, y \in E$ with $x \neq y$, there is $f \in \mathcal{M}$ with $f(x) \neq f(y)$, we say that \mathcal{M} separates points.
2. If, for all $\mathbf{P}, \mathbf{Q} \in \mathcal{P}(E)$,

$$\mathbf{P} = \mathbf{Q} \quad \text{iff} \quad \mathbf{P}[f] = \mathbf{Q}[f] \text{ for all } f \in \mathcal{M},$$

we say that \mathcal{M} is separating in $\mathcal{P}(E)$.

3. If (i) $1 \in \mathcal{M}$ and (ii) if \mathcal{M} is closed under sums and products, then we call \mathcal{M} a (sub-)algebra. If $\mathbb{k} = \mathbb{C}$, and (iii) if \mathcal{M} is closed under complex conjugation, we call \mathcal{M} a star-(sub-)algebra.

In mathlib, 1. and 3. of the above definition are already implemented:

```
structure Subalgebra (R : Type u) (A : Type v) [CommSemiring R] [Semiring A]
  [Algebra R A] extends Subsemiring :
  Type v

abbrev Subalgebra.SeparatesPoints {α : Type u_1} [TopologicalSpace α]
  {R : Type u_2} [CommSemiring R] {A : Type u_3} [TopologicalSpace A]
  [Semiring A] [Algebra R A] [TopologicalSemiring A] (s : Subalgebra R C(α, A))
  : Prop
```

The latter is an extension of `Set.SeparatesPoints`, which works on any set of functions. For the first result, we already need that (E, r) has a metric structure. There is a formalization of this result in https://github.com/pfaffelh/some_probability/tree/master.

l:unique

Lemma 1.2. $\mathcal{M} := \mathcal{C}_b(E, \mathbb{k})$ is separating.

Proof. We restrict ourselves to $\mathbb{k} = \mathbb{R}$, since the result for $\mathbb{k} = \mathbb{C}$ follows by only using functions with vanishing imaginary part. Let $\mathbf{P}, \mathbf{Q} \in \mathcal{P}(E)$. We will prove that $\mathbf{P}(A) = \mathbf{Q}(A)$ for all A closed. Since the set of closed sets is a π -system generating the Borel- σ -algebra, this suffices for $\mathbf{P} = \mathbf{Q}$. So, let A be closed and $g = 1_A$ be the indicator function. Let $g_n(x) := (1 - nr(A, x))^+$ (where $r(A, y) := \inf_{y \in A} r(y, x)$) and note that $g_n(x) \xrightarrow{n \rightarrow \infty} 1_A(x)$. Then, we have by dominated convergence

$$\mathbf{P}(A) = \lim_{n \rightarrow \infty} \mathbf{P}[g_n] = \lim_{n \rightarrow \infty} \mathbf{Q}[g_n] = \mathbf{Q}(A),$$

and we are done. □

We will use the Stone-Weierstrass Theorem below. Here is its version in mathlib. Note that this requires E to be compact.

```
theorem ContinuousMap.starSubalgebra_topologicalClosure_eq_top_of_separatesPoints
  {k : Type u_2} {X : Type u_1} [IsROrC k] [TopologicalSpace X] [CompactSpace X]
  (A : StarSubalgebra k C(X, k)) (hA : Subalgebra.SeparatesPoints A.toSubalgebra) :
  StarSubalgebra.topologicalClosure A = T
```

We also need (as proved in the last project):

```
theorem innerRegular_isCompact_isClosed_measurableSet_of_complete_countable
  [PseudoEMetricSpace α] [CompleteSpace α] [SecondCountableTopology α] [BorelSpace α]
  (P : Measure α) [IsFiniteMeasure P] :
  P.InnerRegular (fun s => IsCompact s ∧ IsClosed s) MeasurableSet
```

The proof of the following result follows [?, Theorem 3.4.5].

T:wc3

Theorem 1 (Algebras separating points and separating algebras).

Let (E, r) be a complete and separable extended pseudo-metric space, and $\mathcal{M} \subseteq \mathcal{C}_b(E, \mathbb{k})$ be a star-sub-algebra that separates points. Then, \mathcal{M} is separating.

Proof. Let $\mathbf{P}, \mathbf{Q} \in \mathcal{P}(E)$, $\varepsilon > 0$ and K compact, such that $\mathbf{P}(K) > 1 - \varepsilon$, $\mathbf{Q}(K) > 1 - \varepsilon$, and $g \in \mathcal{C}_b(E, \mathbb{k})$. According to the Stone-Weierstrass Theorem, there is $(g_n)_{n=1,2,\dots}$ in \mathcal{M} with

$$\sup_{x \in K} |g_n(x) - g(x)| \xrightarrow{n \rightarrow \infty} 0. \quad (1)$$

eq:wc9

So, (note that $C := \sup_{x \geq 0} x e^{-x^2} < \infty$)

$$\begin{aligned} |\mathbf{P}[g e^{-\varepsilon g^2}] - \mathbf{Q}[g e^{-\varepsilon g^2}]| &\leq |\mathbf{P}[g e^{-\varepsilon g^2}] - \mathbf{P}[g e^{-\varepsilon g^2}; K]| \\ &\quad + |\mathbf{P}[g e^{-\varepsilon g^2}; K] - \mathbf{P}[g_n e^{-\varepsilon g_n^2}; K]| \\ &\quad + |\mathbf{P}[g_n e^{-\varepsilon g_n^2}; K] - \mathbf{P}[g_n e^{-\varepsilon g_n^2}]| \\ &\quad + |\mathbf{P}[g_n e^{-\varepsilon g_n^2}] - \mathbf{Q}[g_n e^{-\varepsilon g_n^2}]| \\ &\quad + |\mathbf{Q}[g_n e^{-\varepsilon g_n^2}] - \mathbf{Q}[g_n e^{-\varepsilon g_n^2}; K]| \\ &\quad + |\mathbf{Q}[g_n e^{-\varepsilon g_n^2}; K] - \mathbf{Q}[g e^{-\varepsilon g^2}; K]| \\ &\quad + |\mathbf{Q}[g e^{-\varepsilon g^2}; K] - \mathbf{Q}[g e^{-\varepsilon g^2}]| \end{aligned}$$

We bound the first term by

$$|\mathbf{P}[g e^{-\varepsilon g^2}] - \mathbf{P}[g e^{-\varepsilon g^2}; K]| \leq \frac{C}{\sqrt{\varepsilon}} \mathbf{P}(K^c) \leq C\sqrt{\varepsilon},$$

and analogously for the third, fifth and last. The second and second to last vanish for $n \rightarrow \infty$ due to (??). Since \mathcal{M} is an algebra, we can approximate, using dominated convergence,

$$\mathbf{P}[g_n e^{-\varepsilon g_n^2}] = \lim_{m \rightarrow \infty} \underbrace{\mathbf{P}[g_n \left(1 - \frac{\varepsilon g_n^2}{m}\right)^m]}_{\in \mathcal{M}} = \lim_{m \rightarrow \infty} \underbrace{\mathbf{Q}[g_n \left(1 - \frac{\varepsilon g_n^2}{m}\right)^m]}_{\in \mathcal{M}} = \mathbf{Q}[g_n e^{-\varepsilon g_n^2}],$$

so the fourth term vanishes for $n \rightarrow \infty$ as well. Concluding,

$$|\mathbf{P}[g] - \mathbf{Q}[g]| = \lim_{\varepsilon \rightarrow 0} |\mathbf{P}[g e^{-\varepsilon g^2}] - \mathbf{Q}[g e^{-\varepsilon g^2}]| \leq 4C \lim_{\varepsilon \rightarrow 0} \sqrt{\varepsilon} = 0.$$

Since g was arbitrary and $\mathcal{C}_b(E, \mathbb{k})$ is separating by Lemma ??, we find $\mathbf{P} = \mathbf{Q}$. \square

We now come to characteristic functions and Laplace transforms.

Pr:char1

Proposition 1.3 (Characteristic functions determine distributions uniquely).

A probability measure $\mathbf{P} \in \mathcal{P}(\mathbb{R}^d)$ is uniquely given by its characteristic function.

In other words, if $\mathbf{P}, \mathbf{Q} \in \mathcal{P}(\mathbb{R}^d)$ are such that $\int e^{itx} \mathbf{P}(dx) = \int e^{itx} \mathbf{Q}(dx)$ for all $t \in \mathbb{R}^d$. Then, $\mathbf{P} = \mathbf{Q}$.

Proof. The set

$$\mathcal{M} := \left\{ x \mapsto \sum_{k=1}^n a_k e^{it_k x}; n \in \mathbb{N}, a_1, \dots, a_n \in \mathbb{C}, t_1, \dots, t_n \in \mathbb{R}^d \right\}$$

separates points in \mathbb{R}^d . Since $\mathcal{M} \subseteq \mathcal{C}_b(\mathbb{R}^d, \mathbb{k})$ contains 1, is closed under sums and products, and closed under complex conjugation, it is a star-subalgebra of $\mathcal{C}_b(E, \mathbb{C})$. So, the assertion directly follows from Theorem ??. \square

rem:proj

Remark 1.4. We also need to show the following: For $J \subseteq I$, where I is finite, let ψ be the characteristic function for some distribution on \mathbb{R}^I . Then, for the projection π_J , the characteristic function of the image measure under π_J is given by $\psi \circ g_J$, where $(g_J(t))_j = t_j$ for $j \in J$ and $(g(t))_j = 0$ otherwise. In other words, when computing the characteristic function of a projection, just set the coordinates in $t \mapsto \psi(t)$ which need to be projected out to 0.

2 Gaussian random variables

Define an arbitrary family of Gaussian rvs with values in \mathbb{R}^d by (i) defining a standard normal distribution on \mathbb{R} with the correct density, (ii) show that its characteristic function is given by $\psi(t) = e^{-t^2/2}$, (iii) define an independent finite family of standard normal Gaussians using finite product measures and (iv) define a general independent family by taking some symmetric, positive definite $C \in \mathbb{R}^{d \times d}$, some¹ $A \in \mathbb{R}^{d \times d}$ with $C = A^\top A$, and define the Gaussian measure as the image measure of the independent family Y under the map $X = AY + \mu$. Show that

$$\begin{aligned} \mathbb{E}[e^{itX}] &= \mathbb{E}[e^{it(\mu + AY)}] = e^{it\mu} \mathbb{E}[e^{itAY}] = e^{it\mu} \mathbb{E}\left[\exp\left(i \sum_{kl} t_k A_{kl} Y_l\right)\right] \\ &= e^{it\mu} \prod_l \mathbb{E}\left[\exp\left(i \left(\sum_k t_k A_{kl}\right) Y_l\right)\right] = e^{it\mu} \prod_l \mathbb{E}[e^{i(tA_{\cdot l})Y_l}] \\ &= e^{it\mu} \prod_l e^{-(tA_{\cdot l})^2/2} = e^{it\mu} e^{-\sum_l (tA_{\cdot l})(A_{\cdot l}^\top t^\top)/2} = e^{it\mu - tCt^\top/2}. \end{aligned}$$

In particular, this shows that the distribution does not depend on the choice of A as long as $A^\top A = C$. Together with Proposition ??, this shows that there is a unique probability measure on \mathbb{R}^d with characteristic function $t \mapsto e^{it\mu - tCt^\top/2}$ for any vector μ and symmetric and positive definite matrix C .

3 Projectivity

S:proj

For projectivity of finite-dimensional distributions of the BM, proceed as follows: (i) For $I = \{s_1, \dots, s_n\} \subseteq \mathbb{R}^d$ (with $s_1 < \dots < s_d$), define P_I as the unique probability measure with characteristic function $\psi_I(t) = e^{-tC_I t^\top/2}$ with $C_{ij} = s_i \wedge s_j$. For $J \subseteq I$, we then have that the characteristic function of the projection to coordinates in J is (see Remark ??) $\psi_I \circ g_J = e^{-g_J(\cdot)C_I g_J(\cdot)^\top} = e^{-C_J/2} = \psi_J$. In other words, this is the required projectivity of $(P_I)_{I \subseteq [0, \infty)}$.

4 The Kolmogorov-Chentsov criterion

In this section, let $(D, r_D), (E, r_E)$ be extended pseudo-metric spaces. In addition, we will only have a single probability measure in this section, so we write $\mathbf{P}(\cdot)$ for probabilities and $\mathbf{E}[\cdot]$ for its expectations.

Definition 4.1 (Local Hölder). *Let $f : D \rightarrow E$ and $s \in D$. If there is $\tau > 0$ and some $C < \infty$ with $r_E(f(s), f(t)) \leq Cr_D(s, t)^\gamma$ for all t with $r_D(s, t) < \tau$, we call f locally Hölder of order γ at s .*

Hölder is implemented as `HolderOnWith` (on a set) and `HolderWith`. Moreover, locally Hölder at a point is used for $\gamma = 1$ (i.e. Lipschitz continuity) e.g. in `continuousAt_of_locally_lipschitz` (Every function, which is locally Lipschitz at a point, is continuous.)

l:holderext

Lemma 4.2. *Let D, E be extended pseudo-metric spaces and $f : D \rightarrow E$ and $s \in D$.*

1. *If f is locally Hölder at x , it is continuous at x .*
2. *If E is complete, $A \subseteq D$ is dense, and $g : A \rightarrow E$ is Hölder, it can be extended to a Hölder- γ -function on D .*

¹In order to see that such an A exists, consider some orthogonal O and a diagonal matrix D with $C = O^\top DO$ and set $A := \sqrt{D}O$, where \sqrt{D} is the diagonal matrix with entries $\sqrt{\lambda_i}$ for all eigenvalues λ_i of C . Then, $A^\top A = O^\top \sqrt{D} \sqrt{D} O = O^\top DO = C$.

Proof. 1. Since f is locally Hölder at s , choose $\tau > 0$ and $C < \infty$ such that $r_E(f(s), f(t)) \leq Cr(s, t)^\gamma$ for all t with $r_D(s, t) < \tau$. For $\varepsilon > 0$, there is $\delta' > 0$ such $r_D(s, t)^\gamma < \varepsilon/C$ for all $t \in B_{\delta'}(s)$. Choose $\delta := \tau \wedge \delta'$ in order to see, for $t \in B_\delta(s)$

$$r_E(f(s), f(t)) \leq Cr(s, t)^\gamma < \varepsilon.$$

2. For $s \in D$, choose $s_1, s_2, \dots \in A$ with $s_n \xrightarrow{n \rightarrow \infty} s$. Then, note that $r_E(f(s_n), f(s_m)) \leq Cr_D(s_n, s_m) \xrightarrow{m, n \rightarrow \infty} 0$, so $(f(s_n))_{n=1,2,\dots}$ is a Cauchy-sequence in E . We define $f(s)$ to be its limit. Then, for $s, t \in D$ and the sequences $s_1, s_2, \dots \in D, t_1, t_2, \dots \in D$ with $s_n \xrightarrow{n \rightarrow \infty} s, t_n \xrightarrow{n \rightarrow \infty} t$,

$$r_E(f(s), f(t)) = \lim_{n \rightarrow \infty} r_E(f(s_n), f(t_n)) \leq \lim_{n \rightarrow \infty} Cr_D(s_n, t_n) = Cr_D(s, t).$$

□

For 1., `continuousAt_of_locally_lipschitz` must be adapted for Hölder instead of Lipschitz, i.e. for $\gamma < 1$.

For 2., there is `LipschitzOnWith.extend_real`, which does not require the set A to be dense, but $\gamma = 1$ and $E = \mathbb{R}$. Also, there is `DenseInducing.continuous_extend` which gives a condition when a function can continuously be extended. (It needs a `DenseInducing` function, which in our case is $i : A \rightarrow D, x \mapsto x$.)

l:gauss

Lemma 4.3. For $x \in \mathbb{R}$, let

$$\lfloor x \rfloor := \max\{n \in \mathbb{N} : n \leq x\}.$$

The following holds:

1. $0 \leq x - \lfloor x \rfloor < 1$;
2. If $|x - y| \leq 1$, then $|\lfloor x \rfloor - \lfloor y \rfloor| \leq 1$.
3. $|\lfloor 2x \rfloor - 2\lfloor x \rfloor| \leq 1$.

Proof. 1. The first inequality is clear that $\lfloor x \rfloor$ is defined as a maximum over a set of numbers bounded above by x . The second inequality holds since otherwise we would have $\lfloor x \rfloor + 1 \leq x$, in contradiction to the definition of $\lfloor x \rfloor$.

2. Without loss of generality, assume that $y \leq x$ (which implies that $\lfloor y \rfloor \leq \lfloor x \rfloor$). The proof is by contradiction, so assume that $\lfloor x \rfloor - \lfloor y \rfloor > 1$. So, we find $n := \lfloor x \rfloor \in \mathbb{N}$ such that $y < n - 1 < n \leq x$. This means that $x - y > n - (n - 1) = 1$, in contradiction to $|x - y| \leq 1$.

3. If $x - \lfloor x \rfloor < 1/2$, then $2x - 2\lfloor x \rfloor < 1$, which implies that $\lfloor 2x \rfloor = 2\lfloor x \rfloor$. Last, if $1/2 \leq x - \lfloor x \rfloor < 1$, then $1 \leq 2x - 2\lfloor x \rfloor < 2$, so $\lfloor 2x \rfloor = 2\lfloor x \rfloor + 1$ and the result follows. □

Lemma 4.4. Let $I = [0, 1]^d$ and $|s - t| := \max_{i=1,\dots,d} |s_i - t_i|$ for $s, t \in I$. Let

- $D_n := \{0, 1, \dots, 2^n\}^n \cdot 2^{-n} \subseteq I$ for $n = 0, 1, \dots$, and $D = \bigcup_{n=0}^{\infty} D_n$;
- $m \in \mathbb{N}$ and $s, t \in D$ with $|t - s| \leq 2^{-m}$.

Then, there is $n \geq m$ and $s_m, \dots, s_n, t_m, \dots, t_n$ such that

1. $s_k, t_k \in D_k$ with $|s - s_k|, |t - t_k| \leq 2^{-k}$ for all $k = m, \dots, n$
2. $|s_k - s_{k-1}|, |t_k - t_{k-1}| \leq 2^{-k}$,
3. $|t_m - s_m| \leq 2^{-m}$,
4. $s_n = s, t_n = t$.

Proof. Since $s, t \in D = \bigcup_n D_n$, and $D_n \subseteq D_m$ for $n \geq m$, there is some $n \geq m$ with $s, t \in D_n$. For $k \in m, \dots, n$, we set

$$s_k := \lfloor s2^k \rfloor 2^{-k}, \quad t_k := \lfloor t2^k \rfloor 2^{-k} \in D_k.$$

1. Since $|x - \lfloor x \rfloor| \leq 1$ for all $x \in \mathbb{R}^d$ by Lemma ??1, we have that

$$|s - s_k| = 2^{-k} |s2^k - \lfloor s2^k \rfloor| \leq 2^{-k}, \quad k = m, \dots, n.$$

2. Using Lemma ??3, write

$$|s_k - s_{k-1}| = 2^{-k} |\lfloor 2s2^{k-1} \rfloor - 2\lfloor s2^{k-1} \rfloor| \leq 2^{-k}.$$

3. Since $|t - s| \leq 2^{-m}$, we have $|2^m t - s^m s| \leq 1$, so by Lemma ??2

$$|t_m - s_m| = 2^{-m} |\lfloor t2^m \rfloor - \lfloor s2^m \rfloor| \leq 2^{-m}.$$

4. We have $s2^n, t2^n \in \mathbb{Z}^d$ since $s, t \in D_n$, so $s_n = 2^{-n} \lfloor s2^n \rfloor = 2^{-n} s2^n = s$ and $t_n = t$. \square

rem1

Remark 4.5. Assume that $r(x_s, x_t) \leq 2^{-\gamma k}$ for all s, t with $|t - s| = 2^{-k}$ for $k \geq m$. Then, for some $s, t \in D$ with $|t - s| \leq 2^{-m}$, with s_k, t_k as in the above result and the triangle inequality,

$$\begin{aligned} t &= t_n = s_n + \left(\sum_{k=m+1}^n t_k - t_{k-1} - (s_k - s_{k-1}) \right) + t_m - s_m, \\ r(x_t, x_s) &\leq \left(\sum_{k=m+1}^n r(x_{t_k}, x_{t_{k-1}}) + r(x_{s_k}, x_{s_{k-1}}) \right) + r(x_{t_m}, x_{s_m}) \\ &\leq 2 \sum_{k=m}^n 2^{-\gamma k} \leq \frac{1}{1-2^{-\gamma}} 2^{-\gamma m}. \end{aligned}$$

The proof of the continuity theorem follows the version in [?].

T:kolchen

Theorem 2 (Continuous version; Kolmogorov, Chentsov). *For some $d \in \mathbb{N}$ and $\sigma_1, \tau_1, \dots, \sigma_d, \tau_d > 0$, let $I = \prod_{i=1}^d [\sigma_i, \tau_i]$, and $X = (X_t)_{t \in I}$ an E -valued stochastic process. Assume that there are $\alpha, \beta, C > 0$ with*

$$\mathbf{E}[r(X_s, X_t)^\alpha] \leq C|t - s|^{d+\beta}, \quad 0 \leq s, t \leq \tau.$$

There there exists a version $Y = (Y_t)_{t \in I}$ of X such that, for some random variables $H > 0$ and $K < \infty$,

$$\mathbf{P}\left(\sup_{s \neq t, |t-s| \leq H} r(Y_s, Y_t)/|t-s|^\gamma \leq K\right) = 1,$$

for every $\gamma \in (0, \beta/\alpha)$. In particular, Y almost surely is locally Hölder of all orders $\gamma \in (0, \beta/\alpha)$, and has continuous paths.

Proof. It suffices to show the assertion for $I = [0, 1]^d$. The general case then follows by some scaling of I . We consider the set of times

$$D_n := \{0, 1, \dots, 2^n\}^n \cdot 2^{-n}$$

for $n = 0, 1, \dots$, $D = \bigcup_{n=0}^\infty D_n$. Using the Markov inequality, we write for any $n \in \mathbb{N}$ (note that $|\{s, t \in D_n, |t - s| = 2^{-n}\}| \leq d2^{nd}$), using Lemma ??,

$$\begin{aligned} \mathbf{P}\left(\sup_{s, t \in D_n, |t-s|=2^{-n}} r(X_s, X_t) \geq 2^{-\gamma n}\right) &\leq 2^{\gamma \alpha n} \mathbf{E}\left[\sup_{s, t \in D_n, |t-s|=2^{-n}} r(X_s, X_t)^\alpha\right] \\ &\leq \sum_{s, t \in D_n, |t-s|=2^{-n}} 2^{\gamma \alpha n} \mathbf{E}[r(X_t, X_s)^\alpha] \leq Cd2^{nd} 2^{\gamma \alpha n} 2^{-(d+\beta)n} = Cd2^{(\gamma\alpha-\beta)n}, \end{aligned}$$

and we see that the right hand side is summable. By the Borel-Cantelli Lemma,

$$N := \max \left\{ n : \sup_{s, t \in D_n, |t-s|=2^{-n}} r(X_s, X_t) \geq 2^{-\gamma n} \right\} + 1$$

is finite, almost surely. From this and Remark ??, we conclude with $C' = \frac{1}{1-2^{-\gamma}}$,

$$\sup_{s, t \in D, s \neq t, |t-s| \leq 2^{-N}} r(X_s, X_t) \leq \sup_{m \geq N} \left(\sup_{s, t \in D, |t-s| \leq 2^{-m}} r(X_s, X_t) \right) \leq C' \sup_{m \geq N} 2^{-\gamma m} = C' 2^{-\gamma N}.$$

In other words, we see with $H = 2^{-N}$ and $K = C' 2^{-\gamma N}$,

$$\mathbf{P} \left[\sup_{s, t \in D, s \neq t, |t-s| \leq H} r(X_s, X_t) / |t-s|^\gamma \leq K \right] = 1,$$

i.e. X is locally Hölder- γ on D .

With this and Lemma ??2, we can extend X Hölder-continuously on I , and call this extension $Y = (Y_t)_{t \in I}$. In order to show that Y is a modification of X , fix $t \in I$ and consider a sequence $t_1, t_2, \dots \in D$ with $t_n \rightarrow t$ as $n \rightarrow \infty$. Then, for all $\varepsilon > 0$,

$$\mathbf{P}(r(X_{t_n}, X_t) > \varepsilon) \leq \mathbf{E}[r(X_{t_n}, X_t)^\alpha] / \varepsilon^\alpha \xrightarrow{n \rightarrow \infty} 0,$$

i.e. $X_{t_n} \xrightarrow{n \rightarrow \infty}_p X_t$. Moreover, due to continuity of Y , we have $Y_{t_n} \xrightarrow{n \rightarrow \infty}_{fs} Y_t$. In particular, since $X_{t_n} = Y_{t_n}$ for all n , we have $\mathbf{P}(X_t = Y_t) = 1$ by Lemma ??, which concludes the proof. \square

Definition 4.6 (ε -separated set). *Let $\varepsilon > 0$ and (D, r) be some pseudometric space.*

A set $D' \subseteq D$ is said to be ε -separating if $r(x, y) > \varepsilon$ for all $x \neq y$. It is called maximal, if $D' \cup \{x\}$ is not ε -separating for all $x \notin D'$.

Lemma 4.7. *Let $\varepsilon > 0$, (D, r) be some pseudometric space and $D_\varepsilon \subseteq D$ a maximal ε -separated set. In addition, for $n = 1, 2, \dots$, set $D_n := D_{2^{-n}}$.*

1. *For any $x \in D$ there is $x' \in D_\varepsilon$ such that $r(x, x') < \varepsilon$.*
2. *If $x \in D_\varepsilon$ is not isolated in D , there is $x' \neq x$ with $x' \in D_\varepsilon$ and $r(x, x') < 3\varepsilon$.*
3. *Let $m \leq k$ and $x \in D_k$. Then, there is a sequence $x_k := x \in D_k, x_{k-1} \in D_{k-1}, \dots, x_{m+1} \in D_{m+1}$ with $r(x_\ell, x_{\ell+1}) < 2^{-\ell-1}$ for $\ell = k, \dots, m$.*
4. *Let $m \leq k, \ell$, $x \in D_k, y \in D_\ell$ with $r(x, y) < 2^{-m}$. Then, there are sequences $x_k := x, x_{k-1} \in D_{k-1}, \dots, x_m \in D_m$ and $y_\ell := y, y_{\ell-1} \in D_{\ell-1}, \dots, y_m \in D_m$ as in 2. with $r(x_m, y_m) < 3 \cdot 2^{-m}$.*

Proof. 1. We proceed by contradiction. So, assume there is $x \in D$ such that for all $x' \in D'$ we have $r(x, x') \geq \varepsilon$. This implies that $D' \cup \{x\}$ is ε -separated, in contradiction to the maximality of D' .

2. Since x is not isolated, there is $y \in D$ with $r(x, y) \in (\varepsilon, 2\varepsilon)$. Let $x' \in D_\varepsilon$ with $r(x', y) < \varepsilon$, which exists by 1. Then, $r(x, x') \leq r(x, y) + r(y, x') < 3\varepsilon$.

3. As 1. shows, there is $x_{k+1} \in D_{k+1}$ with $r(x, x_{k+1}) < 2^{-k-1}$. The assertion follows inductively.

4. Using the triangle inequality,

$$\begin{aligned} r(x_m, y_m) &\leq r(x, y) + \sum_{i=m}^{k-1} r(x_i, x_{i+1}) + \sum_{i=m}^{\ell-1} r(y_i, y_{i+1}) \\ &< 2^{-m} + \sum_{i=m}^{\infty} 2^{-i-1} + \sum_{i=m}^{\infty} 2^{-i-1} = 3 \cdot 2^{-m} \end{aligned}$$

\square

Remark 4.8. Ok, this is simple, but, for $d > 0$

$$\sum_{j=0}^n 2^{dj} = \frac{2^{d(n+1)} - 1}{2^d - 1} \leq \frac{2^d}{2^d - 1} 2^{dn}.$$

Theorem 3 (Continuous version; Kolmogorov, Chentsov). *Let (I, r') be a compact metric space and for $\varepsilon > 0$, let J_ε be a maximal ε -separated set in I and $J'_\varepsilon := \{s, t \in J_\varepsilon : r'(s, t) < 3\varepsilon\}$. Assume that, for some $d > 0$, we have $c_1, c_2 > 0$ with*

$$|J_\varepsilon| \leq c_1 \varepsilon^{-d}, \quad |J'_\varepsilon| \leq c_2 \varepsilon^{-d}$$

as $\varepsilon \downarrow 0$. Assume that $X = (X_t)_{t \in I}$ is an E -valued stochastic process and there are $\alpha, \beta, c_3 > 0$ with

$$\mathbf{E}[r(X_s, X_t)^\alpha] \leq c_3 d(s, t)^{d+\beta}, \quad 0 \leq s, t \in I.$$

Then, there exists a version $Y = (Y_t)_{t \in I}$ of X such that, for some random variables $H > 0$ and $K < \infty$,

$$\mathbf{P}\left(\sup_{s \neq t, |t-s| \leq H} r(Y_s, Y_t)/d(s, t)^\gamma \leq K\right) = 1,$$

for every $\gamma \in (0, \beta/\alpha)$. In particular, Y almost surely is locally Hölder of all orders $\gamma \in (0, \beta/\alpha)$, and has continuous paths.

Proof. With a slight abuse of notation, we set $J_n := J_{2^{-n}}$ and $J'_n := J'_{2^{-n}}$. (So, $|J_n| \leq c_1 2^{dn}$ and $|J'_n| \leq c_2 2^{dn}$.) In addition, $K_n := \bigcup_{j=0}^n J_j$ and $K'_n := \bigcup_{j=0}^n J'_j$ (hence $|K_n| \leq c'_1 2^{dn}$ and $|K'_n| \leq c'_2 2^n$ for $c'_i = \frac{c_i 2^d}{2^d - 1}$, $i = 1, 2$). Using the Markov inequality, we write for any $n \in \mathbb{N}$, using Lemma ??,

$$\begin{aligned} \mathbf{P}\left(\sup_{s, t \in K'_n} r(X_s, X_t) \geq 2^{-\gamma n}\right) &= \mathbf{P}\left(\sup_{s, t \in K'_n} r(X_s, X_t)^\alpha \geq 2^{-\gamma \alpha n}\right) \leq 2^{\gamma \alpha n} \mathbf{E}\left[\sup_{s, t \in K'_n} r(X_s, X_t)^\alpha\right] \\ &\leq \sum_{s, t \in K'_n} 2^{\gamma \alpha n} \mathbf{E}[r(X_s, X_t)^\alpha] \leq c'_2 2^{nd} 2^{\gamma \alpha n} c_3 (3 \cdot 2^{-n})^{d+\beta} = c'_2 c_3 3^{d+\beta} 2^{(\gamma \alpha - \beta)n}, \end{aligned}$$

and we see that the right hand side is summable. By the Borel-Cantelli Lemma,

$$N := \max\left\{n : \sup_{s, t \in K'_n} r(X_s, X_t) \geq 2^{-\gamma n}\right\} + 1$$

is finite, almost surely. From this and Remark ??, we conclude with $C = 1 + \frac{2}{1-2^{-\gamma}}$, with $K = \bigcup_{n=1}^\infty K'_n$,

$$\begin{aligned} \sup_{s, t \in K, r'(s, t) \leq 2^{-N}} r(X_s, X_t) &= \sup_{k, \ell \geq N} \sup_{s \in K_k, t \in K_\ell, r'(s, t) \leq 2^{-N}} r(X_s, X_t) \\ &\leq \sup_{k, \ell \geq N} \sup_{s \in K_k, t \in K_\ell, r'(s, t) \leq 2^{-N}} r(x_{s_N}, y_{t_N}) + \sum_{i=N}^{k-1} r(X_{s_i}, X_{s_{i+1}}) + \sum_{i=N}^{\ell-1} r(X_{t_i}, X_{t_{i+1}}) \\ &< 2^{-\gamma N} + \sum_{i=N}^\infty 2^{-\gamma i} + \sum_{i=N}^\infty 2^{-\gamma i} = C 2^{-\gamma N}. \end{aligned}$$

In other words, we see with $H = 2^{-N}$ and $K = C 2^{-\gamma N}$,

$$\mathbf{P}\left[\sup_{s, t \in K, s \neq t, r'(s, t) \leq H} r(X_s, X_t)/r'(s, t)^\gamma \leq K\right] = 1,$$

i.e. X is locally Hölder- γ on K .

With this and Lemma ??2, we can extend X Hölder-continuously on I , and call this extension $Y = (Y_t)_{t \in I}$. In order to show that Y is a modification of X , fix $t \in I$ and consider a sequence $t_1, t_2, \dots \in D$ with $t_n \rightarrow t$ as $n \rightarrow \infty$. Then, for all $\varepsilon > 0$,

$$\mathbf{P}(r(X_{t_n}, X_t) > \varepsilon) \leq \mathbf{E}[r(X_{t_n}, X_t)^\alpha] / \varepsilon^\alpha \xrightarrow{n \rightarrow \infty} 0,$$

i.e. $X_{t_n} \xrightarrow{n \rightarrow \infty}_p X_t$. Moreover, due to continuity of Y , we have $Y_{t_n} \xrightarrow{n \rightarrow \infty}_{fs} Y_t$. In particular, since $X_{t_n} = Y_{t_n}$ for all n , we have $\mathbf{P}(X_t = Y_t) = 1$ by Lemma ??, which concludes the proof. \square

References

- [EK86] S.N. Ethier and T.G. Kurtz. *Markov Processes. Characterization and Convergence*. John Wiley, New York, 1986.
- [KS91] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 1991.