Formalizing Brownian motion

February 4, 2025

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 $f, f_1, f_2, ... : E \to \mathbb{k}$. We say that $f_1, f_2, .. :$ converges to f boundedly pointwise if $f_n \xrightarrow{n\to\infty} f$ pointwise and $\sup_n ||f_n|| < \infty$. We write $f_n \xrightarrow{n\to\infty}_{bp} f$*

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

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. □

Definition 0.3. *Let $X, X_1, X_2, ...,$ all E-valued random variables.*

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

The two notions here are denoted `∀ᵐ (x : α) ∂P, Filter.Tendsto (fun n => X n x) Filter.atTop (nhds (X x))` and `MeasureTheory.TendstoInMeasure`, respectively.

l:aep **Lemma 0.4.** *Let $X, X_1, X_2, ...$ be E-valued random variables with $X_n \xrightarrow{n\to\infty}_{ae} X$. Then, $X_n \xrightarrow{n\to\infty}_p X$.*

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

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

Proof. We write, using monotone convergence and Lemma 0.4
$$\mathbf{P}(X \neq Y) = \lim_{\varepsilon \downarrow 0} \mathbf{P}(r(X,Y) > \varepsilon) \leq \lim_{\varepsilon \downarrow 0} \lim_{n\to\infty} \mathbf{P}(r(X, X_n) > \varepsilon/2) + \mathbf{P}(r(Y, X_n) > \varepsilon/2) = 0.$$
□

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 $\mathcal{M} \subseteq \mathcal{C}_b(E, \Bbbk)$.*

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 $\mathbf{P}, \mathbf{Q} \in \mathcal{P}(E)$,*
$$\mathbf{P} = \mathbf{Q} \quad \textit{iff} \quad \mathbf{P}[f] = \mathbf{Q}[f] \textit{ for all } f \in \mathcal{M},$$
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 $\Bbbk = \mathbb{C}$, 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.** *$\mathcal{M} := \mathcal{C}_b(E, \Bbbk)$ is separating.*

Proof. We restrict ourselves to $\Bbbk = \mathbb{R}$, since the result for $\Bbbk = \mathbb{C}$ 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 $g_n(x) \xrightarrow{n \to \infty} 1_A(x)$. Then, we have by dominated convergence
$$\mathbf{P}(A) = \lim_{n \to \infty} \mathbf{P}[g_n] = \lim_{n \to \infty} \mathbf{Q}[g_n] = \mathbf{Q}(A),$$
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 = ⊤
```

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 [EK86, 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 $\mathcal{M} \subseteq \mathcal{C}_b(E,\Bbbk)$ 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 $g \in \mathcal{C}_b(E,\Bbbk)$. According to the Stone-Weierstrass Theorem, there is $(g_n)_{n=1,2,\dots}$ in $\mathcal{M}$ with

$$\sup_{x\in K} |g_n(x) - g(x)| \xrightarrow{n\to\infty} 0. \tag{1}$$ eq:wc9

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

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

We bound the first term by

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

and analogously for the third, fifth and last. The second and second to last vanish for $n\to\infty$ due to (1). Since $\mathcal{M}$ is an algebra, we can approximate, using dominated convergence,

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

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

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

Since g was arbitrary and $\mathcal{C}_b(E,\Bbbk)$ is separating by Lemma 1.2, we find $\mathbf{P}=\mathbf{Q}$. $\square$

We now come to characteristic functions and Laplace transforms.

Pr:char1 **Proposition 1.3** (Charakteristic 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

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

separates points in $\mathbb{R}^d$. Since $\mathcal{M} \subseteq \mathcal{C}_b(\mathbb{R}^d,\Bbbk)$ 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 1. $\square$

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 $C \in \mathbb{R}^{d\times d}$, some[1] $A \in \mathbb{R}^{d\times d}$ 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

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

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 1.3, 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.

In the concrete application of the finite dimensional distribution of Brownian Motion, consider $0 \le t_1 \le \dots \le t_n$ and $C = (t_i \wedge t_j)_{1\le i,j\le n}$. In order to show that C is positive semi-definite, there are two paths:

1. Find A with $A^\top A = C$: In fact, this A can be given explicitely by

$$A = \begin{pmatrix} \sqrt{t_1} & \sqrt{t_1} & \sqrt{t_1} & \cdots \\ 0 & \sqrt{t_2 - t_1} & \sqrt{t_2 - t_1} & \cdots \\ 0 & 0 & \sqrt{t_3 - t_2} & \\ \cdots & \cdots & \cdots & \ddots \end{pmatrix},$$

   such that

$$A^\top A = \begin{pmatrix} t_1 & t_1 & t_1 & \cdots \\ t_1 & t_2 & t_2 & \cdots \\ t_1 & t_2 & t_3 & \\ \cdots & \cdots & \cdots & \ddots \end{pmatrix}.$$

2. Use induction: Apparently (and this is implemented in `mathlib`), if X and Y are positively semidefinite, then $X + Y$ is positively semidefinite. We write

$$C_3 := \begin{pmatrix} t_1 & t_1 & t_1 \\ t_1 & t_2 & t_2 \\ t_1 & t_2 & t_3 \end{pmatrix} = \begin{pmatrix} t_1 & t_1 & t_1 \\ t_1 & t_2 & t_2 \\ t_1 & t_2 & t_2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & t_3 - t_2 \end{pmatrix} =: X_3 + Y_3,$$

   where Y_3 is obviously positively semidefinite. Now, for $x = (x_1, x_2, x_3)$, we find with $y :=$

[1] In 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, $A \top A = O^\top\sqrt{D}\sqrt{D}O = O^\top DO = C$.

$(x_1, x_2 + x_3)$, summing over all entries of the matrix

$$\begin{aligned} xX_3x^\top &= \sum \begin{matrix} x_1t_1x_1 & x_1t_1x_2 & x_1t_1x_3 \\ x_2t_1x_1 & x_2t_2x_2 & x_2t_2x_3 \\ x_3t_1x_1 & x_3t_2x_2 & x_3t_2x_3 \end{matrix} \\ &= \sum \begin{matrix} x_1t_1x_1 & x_1t_1x_2 & x_1t_1x_3 \\ (x_2+x_3)t_1x_1 & (x_2+x_3)t_2x_2 & (x_2+x_3)t_2x_3 \end{matrix} \\ &= \sum \begin{matrix} x_1t_1x_1 & x_1t_1(x_2+x_3) \\ (x_2+x_3)t_1x_1 & (x_2+x_3)t_2(x_2+x_3) \end{matrix} = yC_2y^\top. \end{aligned}$$

Therefore, using $yC_2y^\top \geq 0$ as the induction hypothesis, this shows that X_3 is positively semidefinite, and the same applies to C_3.

3 Projectivity

S:proj

For projectivity of finite-dimensional distributions of the BM, proceed as follows: (i) For $I = \{s_1, ..., s_n\} \subseteq \mathbb{R}^d$ (with $s_1 < ... < s_d$), define P_J 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 1.4) $\psi_I \circ g_J = e^{-g_J(.)C_I g_J(.)^\top} = e^{-.C_J-/2} = \psi_J$. In other words, this is the required projectivity of $(P_I)_{I \subseteq_f [0,\infty)}$.

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 $\mathbf{P}(.)$ for probabilities and $\mathbf{E}[.]$ for its expectations.

Definition 4.1 (Local Hölder). *Let $f : D \to E$ 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 $f : D \to E$ 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 $g : A \to E$ is Hölder, it can be extended to a Hölder-γ-function on D.*

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

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

□

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 $i : A \to D, x \mapsto x$.)

l:gauss

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

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

The following holds:

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

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 $\lfloor x \rfloor + 1 \le x$, in contradiction to the definition of $\lfloor x \rfloor$.
2. Without loss of generality, assume that $y \le x$ (which implies that $\lfloor y \rfloor \le \lfloor x \rfloor$). The proof is by contradition, so assume that $\lfloor x \rfloor - \lfloor y \rfloor > 1$. So, we find $n := \lfloor x \rfloor \in \mathbb{N}$ such that $y < n - 1 < n \le x$. This means that $x - y > n - (n-1) = 1$, in contradiction to $|x - y| \le 1$.
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 $1/2 \le x - \lfloor x \rfloor < 1$, then $1 \le 2x - 2\lfloor x \rfloor < 2$, so $\lfloor 2x \rfloor = 2\lfloor x \rfloor + 1$ and the result follows. □

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

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

Then, there is $n \ge m$ and $s_m, ..., s_n, t_m, ..., t_n$ such that

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

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

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

1. Since $|x - \lfloor x \rfloor| \le 1$ for all $x \in \mathbb{R}^d$ by Lemma 4.3.1, we have that

$$|s - s_k| = 2^{-k}|s2^k - \lfloor s2^k \rfloor| \le 2^{-k}, \qquad k = m, ..., n.$$

2. Using Lemma 4.3.3, write

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

3. Since $|t - s| \le 2^{-m}$, we have $|2^m t - s^m s| \le 1$, so by Lemma 4.3.2

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

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 $t_n = t$. □

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,

$$t = t_n = s_n + \Big(\sum_{k=m+1}^n t_k - t_{k-1} - (s_k - s_{k-1}) \Big) + t_m - s_m,$$

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

The proof of the continuity theorem follows the version in [KS91].

T:kolchen **Theorem 2** (Continuous version; Kolmogorov, Chentsov)**.** *For some $d \in \mathbb{N}$ and $\sigma_1, \tau_1, ..., \sigma_d, \tau_d > 0$, 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*

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

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

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

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

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

for $n = 0, 1, ...,$ $D = \bigcup_{n=0}^\infty D_n$. 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 0.6,

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

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

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

is finite, almost surely. From this and Remark 4.5, we conclude with $C' = \frac{1}{1-2^{-\gamma}}$,

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

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

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

i.e. X is locally Hölder-γ on D.

With this and Lemma 4.2.2, 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 $t_1, t_2, ... \in D$ with $t_n \to t$ as $n \to \infty$. Then, for all $\varepsilon > 0$,

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

i.e. $X_{t_n} \xrightarrow{n\to\infty}_p X_t$. Moreover, due to continuity of Y, we have $Y_{t_n} \xrightarrow{n\to\infty}_{fs} Y_t$. In particular, since $X_{t_n} = Y_{t_n}$ for all n, we have $\mathbf{P}(X_t = Y_t) = 1$ by Lemma 0.5 , which concludes the proof. □

l:chain **Lemma 4.6.** *Let (I, q) and (E, r) be metric spaces, and $f : I \to E$. Moreover, let $J \subseteq I$ be finite, $a, b, c \in \mathbb{R}_+$ with $a \geq 1$ and $n \in \{1, 2, ...\}$ such that $|J| \leq ba^n$. Then, there is $K \subseteq J^2$ such that*

$$|K| \leq a|J|, \tag{2}$$ eq:chain1

$$(s, t) \in K \Rightarrow q(s, t) \leq cn, \tag{3}$$ eq:chain2

$$\sup_{s,t\in J, q(s,t)\leq c} |f(t) - f(s)| \leq 2 \sup_{(s,t)\in K} |f(s) - f(t)|. \tag{4}$$ eq:chain3

Proof. Start with $V_1 = J$ and some $t_1 \in V_1$. We iteratively construct tuples $(V_\ell, t_\ell \in I, r_\ell \in \{1, ..., d\}, B_\ell \subseteq V_\ell, C_\ell \subseteq V_\ell, K_\ell \subseteq V_\ell^2)$ such that $V_{\ell+1} = V_\ell \setminus B_\ell$ (hence, $\ell \mapsto V_\ell$ is decreasing), $t_\ell \in V_\ell$ is some arbitrary element, and $(r_\ell, B_\ell, C_\ell, K_\ell)$ are given by first finding $r_\ell \in \{1, 2, ...\}$ minimal with

$$|C_\ell| \leq ba^{r_\ell} \text{ for } C_\ell := \{s \in V_\ell : q(s, t_\ell) \leq r_\ell c\}$$

(since $|V_\ell| \leq ba^d$, this r_ℓ exists uniquely) and

$$B_\ell := \{s \in V_\ell : r(s, t_\ell) \leq (r_\ell - 1)c\} \subseteq C_\ell, \qquad K_\ell := \{t_\ell\} \times C_\ell.$$

Note that this implies

$$|B_\ell| \geq ba^{r_\ell - 1}, \qquad |K_\ell| = |C_\ell| \leq ba^{r_\ell}$$

by definition of r_ℓ, and since $t_\ell \in B_\ell$ in all cases. We continue this construction until $V_m = \emptyset$. We claim that

$$K := \bigcup_{\ell=1}^m K_\ell = \{(t_\ell, s) : s \in V_\ell, q(t_\ell, s) \leq cr_\ell \text{ for some } \ell = 1, 2, ...\}$$

satisfies (2), (3) and (4). In order to show (2), we have, since $B_1, B_2, ...$ are disjoint,

$$\sum_\ell ba^{r_\ell - 1} \leq \sum_\ell |B_\ell| \leq |J|.$$

Hence, (2) follows from

$$|K| \leq \sum_\ell |K_\ell| = \sum_\ell |C_\ell| \leq \sum_\ell ba^{r_\ell} \leq a|J|.$$

For (3), we have for $(t_\ell, s) \in K_\ell \subseteq K$,

$$q(t_\ell, s) \leq cr_\ell \leq cd.$$

Last, for (4), consider $(s, t) \in J$ with $q(s, t) \leq c$. (Recall that $\ell \mapsto V_\ell$ is decreasing with $V_1 = J$ and $V_m = \emptyset$.) Find ℓ maximal with $s, t \in V_\ell$. Assume wlog that $s \notin V_{\ell+1}$, which implies $s \in B_\ell$ (since $V_{\ell+1} = V_\ell \setminus B_\ell$), which further implies $q(s, t_\ell) \leq (r_\ell - 1)c$. (This implies $(t_\ell, s) \in K_\ell$.) Since $q(s, t) \leq c$, this gives $q(t, t_\ell) \leq q(t, s) + q(s, t_\ell) \leq r_\ell c$, so $s, t \in C_\ell$. From here, $(t_\ell, s), (t_\ell, t) \in K_\ell \subseteq K$, hence

$$q(f(s), f(t)) \leq q(f(s), f(t_\ell)) + q(f(t_\ell), f(t)) \leq 2 \sup_{s',t'\in K} q(f(s'), f(t'))$$

and we are done by taking $\sup_{s,t\in J.q(s,t)\leq c}$ on the left hand side. □

Definition 4.7 (ε-cover)**.** *Let $\varepsilon > 0$ and (D, r) be some pseudometric space.*
A set $D' \subseteq D$ is said to be an ε-cover of D if $D = \bigcup_{x \in D'} B_\varepsilon(x)$. It is called minimal, if $D' \setminus \{x\}$ is no ε-cover for all $x \in D'$.

Lemma 4.8. *Let $\varepsilon > 0$, (D, r) be some pseudometric space and $D_\varepsilon \subseteq D$ a minimal ε-cover of D. In addition, for $n = 1, 2, ...,$ 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 $\varepsilon > 0$ and $x' \neq x$ with $x' \in D_\varepsilon$ and $r(x, x') < 3\varepsilon$.*
3. *Let $m \le k$ and $x \in D_k$. Then, there is a sequence $x_k := x \in D_k, x_{k-1} \in D_{k-1}, ..., x_{m+1} \in D_{m+1}$ with $r(x_\ell, x_{\ell+1}) < 2^{-\ell-1}$ for $\ell = k, ..., m$.*
4. *Let $m \le k, \ell$, $x \in D_k, y \in D_\ell$ with $r(x, y) < 2^{-m}$. Then, there are sequences $x_k := x, x_{k-1} \in D_{k-1}, ..., x_m \in D_m$ and $y_\ell := y, y_{\ell-1} \in D_{\ell-1}, ..., y_m \in D_m$ as in 2. with $r(x_m, y_m) < 3 \cdot 2^{-m}$.*

Proof. 1. This follows from the definition of a ε-cover.
2. Since x is not isolated, there is $y \in D$ with $r(x, y) \in (\varepsilon, 2\varepsilon)$ for some $\varepsilon > 0$. Let $x' \in D_\varepsilon$ with $r(x', y) < \varepsilon$, which exists by 1. Then, $r(x, x') \le r(x, y) + r(y, x') < 3\varepsilon$.
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,

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

□

Remark 4.9. Ok, this is simple, but, for $d > 0$

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

T:kolchen_general **Theorem 3** (Continuous version; Kolmogorov, Chentsov)**.** *Let (I, q) be a compact metric space and for $\varepsilon > 0$, let I_ε be a finite ε-cover of I. Assume that, for some $d \in \{1, 2, ...\}$, we have $c_1 > 0$ with*

$$|I_\varepsilon| \le c_1 \varepsilon^{-d}$$

for ε small enough. Assume that $X = (X_t)_{t \in I}$ is an E-valued stochastic process and there are $\alpha, \beta, c_2 > 0$ with

$$\mathbf{E}[r(X_s, X_t)^\alpha] \le c_2 q(s, t)^{d+\beta}, \qquad s, t \in I.$$

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

$$\mathbf{P}\Big(\sup_{s \neq t, q(s,t) \le H} r(Y_s, Y_t)/q(s, t)^\gamma \le K \Big) = 1,$$

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 $I_n := I_{2^{-n}}$. (So, $|I_n| \le c_1 2^{dn}$.) In addition, $J_n := \bigcup_{j=0}^n I_j$, hence $|J_n| \le c_1 \sum_{i=0}^n 2^{di} \le c_3 2^{dn}$ with $c_3 = \frac{2^d}{2^d - 1} c_1$. We use Lemma 4.6 with $J = J_n$, $a = 2^d$, $b = c_3$, $c = 2^{-n}$ and $n = n$. This gives some $K_n \subseteq J_n^2$ with $|K_n| \le c_3 2^d 2^{dn}$ such that $(s, t) \in K_n \Rightarrow q(s, t) \le n \cdot 2^{-n}$ and

$$\sup_{s, t \in J_n, q(s,t) \le 2^{-n}} r(X_s, X_t) \le 2 \sup_{(s,t) \in K_n} r(X_s, X_t).$$

Using the Markov inequality, we write for any $n \in \mathbb{N}$, using Lemma 0.6,

$$\begin{aligned}
\mathbf{P}\Big(\sup_{s,t\in J_n, q(s,t)<2^{-n}} r(X_s,X_t) \geq 2^{-\gamma n}\Big) &\leq \mathbf{P}\Big(2 \sup_{(s,t)\in K_n} r(X_s,X_t) \geq 2^{-\gamma n}\Big) \\
&= \mathbf{P}\Big(\sup_{(s,t)\in K_n} r(X_s,X_t)^\alpha \geq 2^{-\alpha}2^{-\gamma\alpha n}\Big) \leq 2^\alpha 2^{\gamma\alpha n}\mathbf{E}\Big[\sup_{(s,t)\in K_n} r(X_s,X_t)^\alpha\Big] \\
&\leq \sum_{(s,t)\in K_n} 2^\alpha 2^{\gamma\alpha n}\mathbf{E}[r(X_s,X_t)^\alpha] \leq c_3 2^d 2^{nd} 2^\alpha 2^{\gamma\alpha n} c_2(3n\cdot 2^{-n})^{d+\beta} = cn^{d+\beta}2^{(\gamma\alpha-\beta)n}
\end{aligned}$$

with $c = c_3 2^d c_2$. So, we see that the right hand side is summable. By the Borel-Cantelli Lemma,

$$N := \max\Big\{n : \sup_{s,t\in J_n, q(s,t)<2^{-n}} r(X_s,X_t) \geq 2^{-\gamma n}\Big\} + 1$$

is finite, almost surely. We set $J := \bigcup_n J_n$ and $H_m := 2^{-(N+m)}$ (with $H_m > 0$, almost surely). For $s \in J_k$ and $k > N+m$, let $s_k := s, s_{k-1}, ..., s_{N+m}$ be as in Lemma xxx, and analogously for $t \in K_\ell$ with $\ell > N+m$. From this and Remark 4.5, we conclude with $c' = 1 + \frac{2}{1-2^{-\gamma}}$ and

$$\begin{aligned}
&\sup_{s,t\in J, q(s,t)\leq H_m} r(X_s,X_t) \\
&\leq \sup_{k,\ell\geq N+m} \sup_{s\in K_k, t\in K_\ell, q(s,t)\leq H_m} r(x_{s_{N+m}}, y_{t_{N+m}}) + \sum_{i=N+m}^{k-1} r(X_{s_i}, X_{s_{i+1}}) + \sum_{i=N+m}^{\ell-1} r(X_{t_i}, X_{t_{i+1}}) \\
&\leq \sup_{s,t\in J_{N+m}, q(s,t)\leq 3\cdot H_m} r(X_s,X_t) + 2^{-\gamma N} + \sum_{i=N}^{\infty} 2^{-\gamma i} + \sum_{i=N}^{\infty} 2^{-\gamma i} = c' 2^{-\gamma N}.
\end{aligned}$$

In other words, we see with $H = 3\cdot 2^{-N}$ and $K = c'2^{-\gamma N}$,

$$\mathbf{P}\Big[\sup_{s,t\in J, s\neq t, q(s,t)\leq H} r(X_s,X_t)/q(s,t)^\gamma \leq K\Big] = 1,$$

i.e. X is locally Hölder-γ on K.

Then,

$$\begin{aligned}
&\sup\Big\{\frac{r(X_s,X_t)}{q(s,t)^\gamma} : s,t\in J, q(s,t)\leq 3\cdot 2^{-N}\Big\} \\
&= \sup_{m=0,1,2,\dots} \sup\Big\{\frac{r(X_s,X_t)}{q(s,t)^\gamma} : s,t\in J, H_{m+1} < q(s,t) \leq H_m\Big\} \\
&\leq \sup_{m=0,1,\dots} 2^{(N+m+1)\gamma} \sup\big\{r(X_s,X_t) : s,t\in J, q(s,t)\leq H_m\big\} \\
&\leq \sup_{m=0,1,\dots} 2^{(N+m+1)\gamma} c' 2^{-\gamma(N+m)} \\
&= 2^\gamma c'.
\end{aligned}$$

With this and Lemma 4.2.2, 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 $t_1, t_2, ... \in D$ with $t_n \to t$ as $n \to \infty$. Then, for all $\varepsilon > 0$,

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

i.e. $X_{t_n} \xrightarrow{n\to\infty}_p X_t$. Moreover, due to continuity of Y, we have $Y_{t_n} \xrightarrow{n\to\infty}_{fs} Y_t$. In particular, since $X_{t_n} = Y_{t_n}$ for all n, we have $\mathbf{P}(X_t = Y_t) = 1$ by Lemma 0.5 , 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.