A Formalization of the Ionescu-Tulcea theorem in Mathlib

Lean Together 2026

Etienne Marion

École Normale Supérieure de Lyon

Supervised by Sébastien Gouëzel (May-June 2024)

January 2026

Table of contents

Table of contents

Markov kernels

Let $(X, \mathcal{A}), (Y, \mathcal{B})$ be two measurable spaces.
A **Markov kernel** from X to Y is a measurable map $\kappa : X \to \mathcal{P}(Y)$.
$\approx$ a map from X to Y with random output.
If $\kappa : X \rightsquigarrow Y$ and $x \in X$, "$\kappa(x) \in Y$" is the random output.

- Push-forward: $\kappa : X \rightsquigarrow Y, f : Y \to Z$.

$$f_* \kappa : X \rightsquigarrow Z$$
$$\text{"}x \mapsto f(\kappa(x))\text{"}$$

- Product: $\kappa : X \rightsquigarrow Y$, $\eta : X \rightsquigarrow Z$.

$$\kappa \times \eta : X \rightsquigarrow Y \times Z$$
$$\text{"}x \mapsto (\kappa(x), \eta(x))\text{"}$$

Operations on Markov kernels

- Composition: $\kappa : X \rightsquigarrow Y$, $\eta : Y \rightsquigarrow Z$.

$$\eta \circ \kappa : X \rightsquigarrow Z$$
$$\text{"}x \mapsto\!\!\!\rightsquigarrow \eta(\kappa(x))\text{"}$$

- Composition-product: $\kappa : X \rightsquigarrow Y$, $\eta : X \times Y \rightsquigarrow Z$.

$$\kappa \otimes \eta : X \rightsquigarrow Y \times Z$$
$$\text{"}x \mapsto\!\!\!\rightsquigarrow (\kappa(x), \eta(x, \kappa(x)))\text{"}$$

Setup of the theorem

Let $(X_n, \mathcal{A}_n)_{n \in \mathbb{N}}$ be a family of measurable spaces.
Let $(\kappa_n : X_0 \times \ldots \times X_n \rightsquigarrow X_{n+1})_{n \in \mathbb{N}}$ be a family of Markov kernels.
Let $x_0 \in X_0$.

$$\begin{aligned} x_0 \mapsto\!\rightsquigarrow \kappa_0(x_0) \quad & = x_1 \in X_1 \\ \mapsto\!\rightsquigarrow \kappa_1(x_0, x_1) \quad & = x_2 \in X_2 \\ \mapsto\!\rightsquigarrow \kappa_2(x_0, x_1, x_2) \quad & = x_3 \in X_3 \\ & \vdots \\ x_0 \mapsto\!\rightsquigarrow \kappa_n(x_0, \ldots, x_n) \quad & = x_{n+1} \in X_{n+1} \end{aligned}$$

This is $(x_1, \ldots, x_{n+1}) = (\kappa_0 \otimes \ldots \otimes \kappa_n)(x_0)$, with

$$\kappa_0 \otimes \ldots \otimes \kappa_n : X_0 \rightsquigarrow X_1 \times \ldots \times X_{n+1}.$$

Does there exist a limit kernel $X_0 \rightsquigarrow \prod_{n \geqslant 1} X_n$?

Statement of the theorem

For $n \geqslant 1$ denote $\pi_n : \prod_{k \geqslant 1} X_k \to X_1 \times \ldots \times X_n$ the canonical projection.

Ionescu-Tulcea Theorem

There exists a unique Markov kernel $\xi : X_0 \rightsquigarrow \prod_{n \geqslant 1} X_n$ such that for all $n \in \mathbb{N}$,

$$\pi_{n+1*}\xi = \kappa_0 \otimes \ldots \otimes \kappa_n.$$

Consequences:

- If for all n, $X_n = X$ et $\kappa_n = \kappa : X \rightsquigarrow X$, then ξ is the Markov chain with transition kernel κ.
- If for all n, κ_n is constant, then ξ is the product measure $\bigotimes_{n \geqslant 1} \kappa_n$.

Already formalized in Isabelle/HOL [Hölzl, 2017].

Olav Kallenberg. Foundations of modern probability. 2021.

Donald L. Cohn. Measure theory. 2013.

Proof sketch

- Measurable cylinders (sets that only depend on first coordinates):

$$A \times \prod_{k \geqslant n+1} X_k \subseteq \prod_{k \geqslant 1} X_k.$$

- Additive content over cylinders:

$$P_0\left(x_0, A \times \prod_{k \geqslant n+1} X_k\right) := (\kappa_0 \otimes \ldots \otimes \kappa_{n-1})(x_0, A).$$

- Carathéodory's extension theorem

Key lemma

If $(A_n)_{n\in\mathbb{N}}$ is a nonincreasing sequence of measurable cylinders with empty intersection, then $P_0(x_0, A_n) \longrightarrow_{n\to\infty} 0$.

Proof sketch of the lemma

- section of sets: if $A \subseteq \prod_{k\geqslant 1} X_k$,

$$A(x_1, ..., x_n) := \left\{ y \in \prod_{k>n} X_k \mid (x_1, ..., x_n, y) \in A \right\}.$$

- Say $A_n = B_n \times \prod_{k>a_n} X_k$.

$$\begin{aligned}
&P_0(x_0, A_n)\\
=&(\kappa_0 \otimes ... \otimes \kappa_{a_n-1})(x_0, B_n)\\
=&\int_{X_1} (\kappa_1 \otimes ... \otimes \kappa_{a_n-1})(x_0, x_1, B_n(x_1))\kappa_0(x_0, \mathrm{d}x_1)\\
=&\int_{X_1}\int_{X_2} (\kappa_2 \otimes ... \otimes \kappa_{a_n-1})(x_0, x_1, x_2, B_n(x_1, x_2))\kappa_1(x_0, x_1, \mathrm{d}x_2)\kappa_0(x_0, \mathrm{d}x_1).
\end{aligned}$$

Assume $P_0(x_0, A_n) \geqslant \epsilon$. We can choose x_1, x_2 so that

$$\forall n, (\kappa_2 \otimes ... \otimes \kappa_{a_n-1})(x_0, x_1, x_2, B_n(x_1, x_2)) \geqslant \epsilon.$$

Table of contents

What was already there?

In Mathlib:

- Markov kernels and operations [Degenne, 2025]

In KolmogorovExtension4 [Degenne and Pfaffelhuber]:

- Carathéodory's extension theorem
- Cylinder sets
- Additive contents
- Projective families of measures

Composition of kernels

Let $\kappa : X \rightsquigarrow Y$, $\eta : X \times Y \rightsquigarrow Z$, $\xi : X \times (Y \times Z) \rightsquigarrow T$.

$$\kappa \otimes \eta : X \rightsquigarrow Y \times Z$$
$$(\kappa \otimes \eta) \otimes \xi : X \rightsquigarrow (Y \times Z) \times T$$

$\kappa \otimes (\eta \otimes \xi)$? No! $\xi' : (X \times Y) \times Z \rightsquigarrow T$.

$$\eta \otimes \xi' : X \times Y \rightsquigarrow Z \times T$$
$$\kappa \otimes (\eta \otimes \xi') : X \rightsquigarrow Y \times (Z \times T)$$

For $a < b$, we want to compose

$$(\kappa_0 \otimes \ldots \otimes \kappa_{a-1}) : X_0 \rightsquigarrow X^{[\![1,a]\!]} \quad \text{and} \quad (\kappa_a \otimes \ldots \otimes \kappa_{b-1}) : X^{\leqslant a} \rightsquigarrow X^{[\![a+1,b]\!]}$$

as a kernel $X_0 \rightsquigarrow X^{[\![1,b]\!]}$. We need

$$X^{\leqslant a} \simeq X_0 \times X^{[\![1,a]\!]} \quad \text{and} \quad X^{[\![1,a]\!]} \times X^{[\![a+1,b]\!]} \simeq X^{[\![1,b]\!]}.$$

Solution

For $a, b \in \mathbb{N}$ consider

$$\eta_{a,b} : X^{\leqslant a} \rightsquigarrow X^{\leqslant b}$$
$$(x_0, \dots, x_a) \mapsto \begin{cases} (x_0, \dots, x_b) & \text{if } b \leqslant a \\ (x_0, \dots, x_a, (\kappa_a \otimes \dots \otimes \kappa_{b-1})(x_0, \dots, x_a)) & \text{else} \end{cases}.$$

Using isomorphisms to be perfectly formal, denote

$$\varphi_b : X_{b+1} \to \prod_{b<k\leqslant b+1} X_k \quad \text{and} \quad \psi_{m,n} : X^{\leqslant m} \times \prod_{m<k\leqslant n} X_k \to X^{\leqslant n}$$

the two obvious isomorphisms. Then we define

$$\eta_{a,b+1} := \psi_{b,b+1*}((\delta \times \varphi_{b*}\kappa_b) \circ \eta_{a,b}).$$

If $a \leqslant b \leqslant c$ then $\eta_{b,c} \circ \eta_{a,b} = \eta_{a,c}$.

```
theorem partialTraj_comp_partialTraj (hab : a ≤ b) (hbc : b ≤ c) :
    partialTraj κ b c ∘ₖ partialTraj κ a b = partialTraj κ a c
```

If $b \leqslant c$, $\pi_{c\to b*}\eta_{a,c} = \eta_{a,b}$.

Peeling off integrals

$$P_0(x_0, A_n) = (\kappa_0 \otimes \ldots \otimes \kappa_{a_n-1})(x_0, B_n)$$
$$= \int_{X_1} (\kappa_1 \otimes \ldots \otimes \kappa_{a_n-1})(x_0, x_1, B_n(x_1))\kappa_0(x_0, \mathrm{d}x_1).$$

$\Rightarrow$ `lmarginal` [Van Doorn and Macbeth, 2024].
Let $(X_i, \mu_i)_{i\in\iota}$, $f : \prod_{i\in\iota} X_i \to [0, +\infty]$ and $s \subseteq \iota$.

$$\int_{\prod_{i\in\iota} X_i} f(x) \bigotimes_{i\in\iota} \mu_i(\mathrm{d}x) = \int_{\prod_{i\in s^c} X_i} \int_{\prod_{i\in s} X_i} f(x_s, x_{s^c}) \bigotimes_{i\in s} \mu_i(\mathrm{d}x_s) \bigotimes_{i\in s^c} \mu_i(\mathrm{d}x_{s^c})$$

Does not work in lean. `lmarginal μ s f` :=

$$\prod_{i\in\iota} X_i \to [0, +\infty]$$
$$x \mapsto \int_{\prod_{i\in s} X_i} f(y_s \hookrightarrow x) \bigotimes_{i\in s} \mu_i(\mathrm{d}y_s)$$

Peeling off integrals

Let $a, b \in \mathbb{N}$, $f : \prod_{n\in\mathbb{N}} X_n \to [0, +\infty]$ and $x \in \prod_{n\in\mathbb{N}} X_n$. Then define `lmarginalPartialTraj κ a b f x` as

$$\int_{X^{\leqslant b}} f(y \hookrightarrow_b x)\eta_{a,b}(x_0, ..., x_a, \mathrm{d}y).$$

We can write

```
lmarginalPartialTraj κ a b (lmarginalPartialTraj κ b c f).
```

```
theorem lmarginalPartialTraj_self (hab : a ≤ b) (hbc : b ≤ c)
    {f : (Π n, X n) → ℝ≥0∞} (hf : Measurable f) :
    lmarginalPartialTraj κ a b (lmarginalPartialTraj κ b c f) =
    lmarginalPartialTraj κ a c f
```

$$\begin{aligned} P_0(x_0, A_n \times X^{>n}) &= \int_{X^{\leqslant n}} \mathbb{1}_{A_n} \mathrm{d}\eta_{0,n}(x_0) \\ &= \int_{X_1} \int_{X^{\leqslant n}} \mathbb{1}_{A_n}(y)\eta_{1,n}(x_0, x_1, \mathrm{d}y)\kappa_0(x_0, \mathrm{d}x_1). \end{aligned}$$

Output object

We defined:

$$\eta_{a,b} : X^{\leqslant a} \rightsquigarrow X^{\leqslant b} \approx \delta \times (\kappa_a \otimes \ldots \otimes \kappa_{b-1})$$

"Letting b go to infinity" we get

$$\xi_a : X^{\leqslant a} \rightsquigarrow \prod_{n \in \mathbb{N}} X_n \approx \delta \times (\kappa_a \otimes \ldots).$$

```
def traj (a : ℕ) : Kernel (Π i : Iic a, X i) (Π n, X n)
```

$$\pi_{b*}\xi_a = \eta_{a,b}.$$

```
lemma traj_map_frestrictLe (a b : ℕ) :
      (traj κ a).map (frestrictLe b) = partialTraj κ a b
```

$$\xi_b \circ \eta_{a,b} = \xi_a.$$

```
theorem traj_comp_partialTraj {a b : ℕ} (hab : a ≤ b) :
    (traj κ b) ∘ₖ (partialTraj κ a b) = traj κ a
```

Table of contents

Conclusion

- What's next?
  - Product measure: already used in Mathlib, for instance #31908 by Yaël (Erdös-Renyi graphs)
  - Stochastic algorithms:
    - LeanBandits (Rémy Degenne, Paulo Ruber)
    - A Unifying Framework for Global Optimization: From Theory to Formalization, Serré et al. 2025
    - Towards Formalizing Reinforcement Learning Theory, Zhang 2025

- Acknowledgements:
  - Sébastien Gouëzel
  - Rémy Degenne
  - The Mathlib community